

EDITORIAL BOARD

Honorary Editor

Lawrence Michael Brown

Editor in Chief

Mathematics

Ayşe Çiğdem Özcan (Hacettepe University, Mathematics - hjms@hacettepe.edu.tr)

Statistics

Çağdaş Hakan Aladağ (Hacettepe University, Statistics - statisticshjms@gmail.com)

Associate Editors

Bülent Saraç (Hacettepe University, Mathematics - bsarac@hacettepe.edu.tr)

Aslı Yıldız (Hacettepe University, Mathematics - aslipekcan@hacettepe.edu.tr)

Ceren Eda Can (Hacettepe University, Statistics - cerencan@hacettepe.edu.tr)

Production Editors

Gülbanu Tekbulut (Hacettepe University - dergilerhacettepe@gmail.com)

O. Oğulcan Tuncer (Hacettepe University, Mathematics - otuncer@hacettepe.edu.tr)

Talha Arıkan (Hacettepe University, Mathematics - tarikan@hacettepe.edu.tr)

Nurbanu Bursa (Hacettepe University, Statistics - nurbanubursa@hacettepe.edu.tr)

Area Editors

Mathematics

Evrım Akalan (Associative Rings and Algebras - eakalan@hacettepe.edu.tr)

Okay Çelebi (Partial Differential Equations - acelebi@yeditepe.edu.tr)

Angel Del Rio (Algebra, Group Theory - adelrio@um.es)

Gülin Ercan (Algebra, Group Theory - ercan@metu.edu.tr)

Sergio Estrada (Homological Algebra - sestrada@um.es)

Varga Kalantarov (Partial Differential Equations - vkalantarov@ku.edu.tr)

Emre Mengi (Numerical Analysis - emengi@ku.edu.tr)

Cihan Orhan (Analysis, Summability - orhan@science.ankara.edu.tr)

Murad Özyayın (Differential Geometry, Global Analysis - mozaydin@math.ou.edu)

Abdullah Özbekler (Ordinary Differential Equations - aozbekler@gmail.com)

Sibel Özkan (Combinatorics - s.ozkan@gtu.edu.tr)

Ekin Özman (Number Theory, Algebraic Geometry - ekin.ozman@boun.edu.tr)

Serap Öztıp Kaptanođlu (Abstract Harmonic Analysis - oztops@istanbul.edu.tr)

Mehmetçik Pamuk (Topology, Manifolds and Cell Complexes - mpamuk@metu.edu.tr)

Ivan Reilly (Topology - i.reilly@auckland.ac.nz)

Alexander P. Šostak (General Topology - sostaks@latnet.lv)

Statistics

Ali Allahverdi (Operational Research & Statistics - ali.allahverdi@ku.edu.kw)

Olcay Arslan (Robust Statistics - oarslan@ankara.edu.tr)

Narayanaswany Balakrishnan (Applied Statistics, Theory of Statistics - bala@mcmaster.ca)

Adil Baykasoglu (Operational Research - adil.baykasoglu@deu.edu.tr)

Sat Gupta (Sampling, Time Series - sngupta@uncg.edu)

Birdal Şenođlu (Experimental Design, Statistical Distributions - senoglu@science.ankara.edu.tr)

Tahir Hanaliođlu (Stochastic Processes Theory, Probability Theory - tahirkhaniyev@etu.edu.tr)

Zeynep Işıl Kalaylıođlu (Bayesian Inference, Model Selection - kzeynep@metu.edu.tr)



Chebyshev Wavelet Method for Numerical Solutions of Coupled Burgers' Equation

Ö. Oruç^{*1}, F. Bulut², A. Esen³

¹*Eğil Vocational and Technical Anatolian High School, Diyarbakır, Turkey*

²*İnönü University, Department of Physics, Malatya, Turkey*

³*İnönü University, Department of Mathematics, Malatya, Turkey*

Abstract

This paper deals with the numerical solutions of one dimensional time dependent coupled Burgers' equation with suitable initial and boundary conditions by using Chebyshev wavelets in collaboration with a collocation method. The proposed method converts coupled Burgers' equations into system of algebraic equations by aid of the Chebyshev wavelets and their integrals which can be solved easily with a solver. Benchmarking of the proposed method with exact solution and other known methods already exist in the literature is made by three test problems. The feasibility of the proposed method is demonstrated by test problems and indicates that the proposed method gives accurate results in short cpu times. Computer simulations show that the proposed method is computationally cheap, fast and quite good even in the case of less number of collocation points.

Mathematics Subject Classification (2010). 35Q35, 65M70

Keywords. Chebyshev wavelet method, Chebyshev collocation, coupled Burgers' equation, nonlinear phenomena, numerical solution

1. Introduction

We consider the Coupled Burgers' problem

$$u_t - u_{xx} + \eta uu_x + \alpha(u.v)_x = 0, \quad x \in [0, 1], \quad t \in [0, T] \quad (1.1)$$

$$v_t - v_{xx} + \xi vv_x + \beta(u.v)_x = 0, \quad x \in [0, 1], \quad t \in [0, T] \quad (1.2)$$

with the initial conditions

$$u(x, 0) = \psi_1(x), \quad v(x, 0) = \psi_2(x), \quad x \in [a, b]$$

and the boundary conditions

$$u(0, t) = f_1(t), \quad u(1, t) = f_2(t), \quad t \in [0, T]$$

$$v(0, t) = g_1(t), \quad v(1, t) = g_2(t), \quad t \in [0, T]$$

^{*}Corresponding Author.

Email addresses: omeroruc0@gmail.com (Ö. Oruç), fatih.bulut@inonu.edu.tr (F. Bulut), alaattin.esen@inonu.edu.tr (A. Esen)

Received: 16.02.2017; Accepted: 12.07.2017

where η , ξ are real constants and α , β are arbitrary constants depend on system parameters such as Peclet number, Stokes velocity of particles due to gravity and the Brownian diffusivity [30]. $u(x, t)$ and $v(x, t)$ are the velocity components to be determined; ψ_i , f_i and g_i ($i = 1, 2$) are the known functions; uu_x is the nonlinear convection term, u_t is unsteady term and u_{xx} is diffusion term.

Coupled Burgers' equation was first derived by Esipov [12] which is a simple model of sedimentation or evolution of scaled volume concentrations of two kinds of particles in fluid suspensions or colloids, under the effect of gravity [30]. This equation has been solved by various approaches such as; Khater et al. [21] used the Chebyshev spectral collocation method to solve the equation and M. Dehghan et al. [11] applied Adomian–Pade technique for solving the coupled Burgers equations and more recently Kutluay and Ucar [23] solved coupled Burgers' equation by using the Galerkin quadratic B-spline method. In order to solve Eqs. (1.1), (1.2) Mittal and Arora [28] used a cubic B-spline collocation scheme. Rashid and Ismail [34] have used Fourier Pseudospectral method to solve the equation numerically. Srivastava et al. [42] obtained numerical solutions of the Eqs.(1.1), (1.2) by implicit finite-difference method. Zhang et al. [37] applied local discontinuous Galerkin method to solve coupled Burgers' equations. Siraj-ul-Islam et al. [17] solved coupled Burgers' equation by mesh free interpolation method. Kelleci and Yildırım [20] have solved the equation by combining homotopy perturbation and Pade techniques and Inan et al. [16] have applied Bäcklund transformation to the Eqs.(1.1), (1.2). In the studies [22, 29], coupled Burgers' equations are solved by Haar wavelet method. Rashid et al. have solved the coupled viscous Burgers' equation by Chebyshev–Legendre Pseudo-Spectral method in [33].

Kaya [19] obtained the exact solution of the equation by Adomian Decomposition method and Soliman [41] used a modified extended tanh-function method to obtain its exact solution. Abdou and Soliman [2] used Variational iteration method to solve the coupled viscous Burgers' equation.

The wavelet methods were first applied for solving differential equations at the beginning of 1990s. Until now a vast number of papers devoted to this topic. In most cases the wavelet coefficients were calculated by the Galerkin or collocation method. But there is a drawback in these methods since we have to evaluate integrals of some combinations of the wavelet functions (connection coefficients). This is a very sophisticated problem, since for most wavelet families we do not have an explicit form for these integrals [25]. Due to these facts, researchers have focused on more simple wavelets such as Haar wavelets, Legendre wavelets and Chebyshev wavelets for obtaining numerical solutions of differential and integral equations. There are a lot of studies on application of Haar wavelets in solving differential and integral equations numerically [6, 7, 9, 18, 22, 24, 26, 27, 29, 31, 32, 40]. Nowadays, Legendre and Chebyshev wavelets are studied by many researchers [3, 4, 8, 13, 14, 35, 36, 38, 39, 44–47]. In this paper we propose a Chebyhev wavelet method for solving coupled Burgers' equations numerically.

The outline of this paper is as follows. In Section 2, preliminaries about Chebyshev wavelets are given. In section 3, we show how to use Chebyshev wavelet method for solving coupled Burgers' equation. In Section 4, proposed method tested by three examples, obtained numerical results tabulated and numerical solutions depicted graphically. Finally we conclude the paper in Section 5.

2. Preliminaries and notations

In this section, we give some necessary definitions and mathematical preliminaries of Chebyshev wavelets.

2.1. Chebyshev wavelets

Wavelets constitute a family of functions which are generated from dilation and translation of a single function which is called as mother wavelet $\psi(x)$. If the dilation parameter a and the translation parameter b vary continuously we have the following family of continuous wavelets [10]:

$$\psi_{a,b}(x) = |a|^{-1/2} \psi\left(\frac{x-b}{a}\right),$$

where $a, b \in \mathbb{R}$ and $a \neq 0$. Chebyshev wavelets $\psi_{nm} = \psi(k, n, m, x)$ defined as follows:

$$\psi_{nm}(x) = \begin{cases} \gamma_m \frac{2^{(k-1)/2}}{\sqrt{\pi}} T_m(2^k x - 2n + 1), & \frac{n-1}{2^{k-1}} \leq x < \frac{n}{2^{k-1}} \\ 0, & \text{else} \end{cases} \quad (2.1)$$

where

$$\gamma_m = \begin{cases} \sqrt{2}, & m = 0 \\ 2, & m = 1, 2, \dots \end{cases}$$

and $m = 0, 1, \dots, M - 1$. Here $n = 1, 2, \dots, 2^{k-1}$, k can take any positive integer, m is the degree of Chebyshev polynomials of first kind and x is the normalized time. $T_m(x)$ are Chebyshev polynomials of the first kind of degree m and satisfy the following recursive formula:

$$T_0(x) = 1, T_1(x) = x, \quad T_{m+1}(x) = 2xT_m(x) - T_{m-1}(x).$$

which are orthogonal with respect to the weight function $\omega(x) = 1/\sqrt{1-x^2}$. We should remind that Chebyshev wavelets are orthogonal with respect to the weight function $\omega_n(x) = \omega(2^k x - 2n + 1)$.

2.2. Function approximation

Any function $u(x) \in L^2_\omega[0, 1)$ can be expanded into Chebyshev wavelets as follows:

$$u(x) = \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} c_{nm} \psi_{nm}(x). \quad (2.2)$$

Here wavelet coefficients are $c_{nm} = \langle u(x), \psi_{nm}(x) \rangle$, where $\langle \cdot, \cdot \rangle$ represents the inner product with respect to $\omega_n(x)$.

In practice, one needs the truncated version of the Eq. (2.2), namely:

$$u(x) = \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} \psi_{nm}(x) = C^T \Psi(x), \quad (2.3)$$

where C and $\Psi(x)$ are $2^{k-1}M \times 1$ matrices given as

$$\begin{aligned} C &= [c_{10}, c_{11}, \dots, c_{1(M-1)}, c_{20}, c_{21}, \dots, c_{2(M-1)}, \dots, \\ &\quad c_{2^{k-1}0}, c_{2^{k-1}1}, \dots, c_{2^{k-1}(M-1)}]^T, \\ \Psi(x) &= [\psi_{10}(x), \psi_{11}(x), \dots, \psi(x)_{1(M-1)}, \psi_{20}(x), \psi_{21}(x), \dots, \psi(x)_{2(M-1)}, \dots, \\ &\quad \psi_{2^{k-1}0}(x), \psi_{2^{k-1}1}(x), \dots, \psi(x)_{2^{k-1}(M-1)}]^T. \end{aligned}$$

Convergence analysis of Chebyshev wavelets is given in [3, 45].

2.3. Integrals of Chebyshev wavelets [8]

We denote the first integral of the Eq. (2.1) as $p_{nm}(x) = \int_0^x \psi_{nm}(s)ds$ and the second integral of the Eq. (2.1) as $q_{nm}(x) = \int_0^x p_{nm}(s)ds$. The first integral $p_{nm}(x)$ is given for $m = 0$, $m = 1$ and $m > 1$ as follows:

$$p_{n0}(x) = \begin{cases} 0 & 0 \leq x < \frac{n-1}{2^{k-1}} \\ \gamma_0 \frac{2^{-(k-1)/2-1}}{\sqrt{\pi}} [T_1(t) + T_0(t)], & \frac{n-1}{2^{k-1}} \leq x < \frac{n}{2^{k-1}} \\ \gamma_0 \frac{2^{-(k-1)/2}}{\sqrt{\pi}} T_0(t), & \frac{n}{2^{k-1}} \leq x < 1 \end{cases}$$

$$p_{n1}(x) = \begin{cases} 0 & 0 \leq x < \frac{n-1}{2^{k-1}} \\ \gamma_1 \frac{2^{-(k-1)/2-3}}{\sqrt{\pi}} [T_2(t) - T_0(t)], & \frac{n-1}{2^{k-1}} \leq x < \frac{n}{2^{k-1}} \\ 0, & \frac{n}{2^{k-1}} \leq x < 1 \end{cases}$$

$$p_{nm}(x) = \begin{cases} 0 & 0 \leq x < \frac{n-1}{2^{k-1}} \\ \gamma_m \frac{2^{-(k-1)/2-2}}{\sqrt{\pi}} \left[\frac{T_{m+1}(t) - (-1)^{m+1}}{m+1} - \frac{T_{m-1}(t) - (-1)^{m-1}}{m-1} \right], & \frac{n-1}{2^{k-1}} \leq x < \frac{n}{2^{k-1}} \\ \gamma_m \frac{2^{-(k-1)/2-2}}{\sqrt{\pi}}, & \frac{n}{2^{k-1}} \leq x < 1 \end{cases}$$

where $t = 2^k x - 2n + 1$. The second integral $q_{nm}(x)$ is given for $m = 0$, $m = 1$, $m = 2$ and $m > 2$ as follows:

$$q_{n0}(x) = \begin{cases} 0 & 0 \leq x < \frac{n-1}{2^{k-1}} \\ \gamma_0 \frac{2^{-3(k-1)/2-4}}{\sqrt{\pi}} [T_2(t) + 4T_1(t) + 3T_0(t)], & \frac{n-1}{2^{k-1}} \leq x < \frac{n}{2^{k-1}} \\ \gamma_0 \frac{2^{-(k-1)/2}}{\sqrt{\pi}} \left(\frac{1}{2^k} + x - \frac{n}{2^{k-1}} \right), & \frac{n}{2^{k-1}} \leq x < 1 \end{cases}$$

$$q_{n1}(x) = \begin{cases} 0 & 0 \leq x < \frac{n-1}{2^{k-1}} \\ \gamma_1 \frac{2^{-3(k-1)/2-4}}{\sqrt{\pi}} \left[\frac{T_3(t)}{6} - \frac{3T_1(t)}{2} - \frac{4T_0(t)}{3} \right], & \frac{n-1}{2^{k-1}} \leq x < \frac{n}{2^{k-1}} \\ \gamma_1 \frac{2^{-3(k-1)/2-1}}{-3\sqrt{\pi}}, & \frac{n}{2^{k-1}} \leq x < 1 \end{cases}$$

$$q_{n2}(x) = \begin{cases} 0 & 0 \leq x < \frac{n-1}{2^{k-1}} \\ \gamma_2 \frac{2^{-3(k-1)/2-3}}{\sqrt{\pi}} \left[\frac{T_4(t) - 1}{24} - \frac{T_2(t) - 1}{3} - \frac{2T_1(t)}{3} - \frac{2T_0(t)}{3} \right], & \frac{n-1}{2^{k-1}} \leq x < \frac{n}{2^{k-1}} \\ \gamma_2 \frac{2^{-(k-1)/2}}{-3\sqrt{\pi}} \left(\frac{1}{2^k} + x - \frac{n}{2^{k-1}} \right), & \frac{n}{2^{k-1}} \leq x < 1 \end{cases}$$

$$q_{nm}(x) = \begin{cases} 0 & 0 \leq x < \frac{n-1}{2^{k-1}} \\ \gamma^m \frac{2^{-3(k-1)/2-3}}{\sqrt{\pi}} \left[\begin{aligned} & \frac{T_{m+2}(t) - (-1)^{m+2}}{2(m+1)(m+2)} - \frac{T_m(t) - (-1)^m}{2(m+1)m} \\ & - \frac{T_m(t) - (-1)^m}{2m(m-1)} \\ & + \frac{T_{m-2}(t) - (-1)^{m-2}}{2(m-1)(m-2)} + \\ & (1 + T_1(t)) \left[\frac{(-1)^{m-1}}{m-1} - \frac{(-1)^{m+1}}{m+1} \right] \end{aligned} \right], & \frac{n-1}{2^{k-1}} \leq x < \frac{n}{2^{k-1}} \\ \gamma^m \frac{2^{-3(k-1)/2-3}}{\sqrt{\pi}} \left[\begin{aligned} & \frac{1 - (-1)^{m+2}}{2(m+1)(m+2)} - \frac{1 - (-1)^m}{2(m+1)m} - \frac{1 - (-1)^m}{2m(m-1)} \\ & + \frac{1 - (-1)^{m-2}}{2(m-1)(m-2)} + 2 \left[\frac{(-1)^{m-1}}{m-1} - \frac{(-1)^{m+1}}{m+1} \right] \\ & + 2^k \left(x - \frac{n}{2^{k-1}} \right) \left[\frac{1 - (-1)^{m+1}}{m+1} - \frac{1 - (-1)^{m-1}}{m-1} \right] \end{aligned} \right], & \frac{n}{2^{k-1}} \leq x < 1 \end{cases}$$

where $t = 2^k x - 2n + 1$. We will use these integrals in solution procedure later.

3. Method of solution for coupled Burgers' equation

Consider the equations (1.1), (1.2) with the initial conditions

$$u(x, 0) = \psi_1(x), \quad v(x, 0) = \psi_2(x), \quad x \in [0, 1]$$

and the boundary conditions

$$\begin{aligned} u(0, t) &= f_1(t), & u(1, t) &= f_2(t), & t &\in [0, T] \\ v(0, t) &= g_1(t), & v(1, t) &= g_2(t), & t &\in [0, T] \end{aligned}$$

Let us divide the interval $[0, T]$ into N equal parts of length $\Delta t = T/N$ and denote $t_s = (s - 1)\Delta t$, $s = 1, 2, \dots, N$. In order to use the Chebyshev integrals given in the previous section we expand the highest derivatives that appeared in the Eqs. (1.1) and (1.2) into Chebyshev wavelets. Therefore assume that $\dot{u}''(x, t)$ and $\dot{v}''(x, t)$ can be expanded in terms of Chebyshev wavelets as

$$\dot{u}''(x, t) = \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} \psi_{nm}(x) \tag{3.1}$$

$$\dot{v}''(x, t) = \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} \psi_{nm}(x) \tag{3.2}$$

where $\dot{}$ and $'$ means differentiation with respect to t and x , respectively, the row vectors c_{nm} and d_{nm} are constants in the sub-interval $t \in [t_s, t_{s+1}]$. We discretize $u(x, t)$ below, same procedure can be applied to $v(x, t)$.

Integrating equation (3.1) with respect to t from t_s to t and twice with respect to x from 0 to x , we have following equations:

$$u''(x, t) = (t - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} \psi_{nm}(x) + u''(x, t_s), \quad (3.3)$$

$$u'(x, t) = (t - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} p_{nm}(x) + u'(x, t_s) - u'(0, t_s) + u'(0, t), \quad (3.4)$$

$$u(x, t) = (t - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(x) + u(x, t_s) - u(0, t_s) + x [u'(0, t) - u'(0, t_s)] + u(0, t), \quad (3.5)$$

$$\dot{u}(x, t) = \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(x) + \dot{u}(0, t) + x \dot{u}'(0, t). \quad (3.6)$$

By using boundary conditions, we obtain

$$\begin{aligned} u(0, t) &= f_1(t), & u(0, t_s) &= f_1(t_s), & \dot{u}(0, t_s) &= f_1'(t_s) \\ u(1, t) &= f_2(t), & u(1, t_s) &= f_2(t_s), & \dot{u}(1, t_s) &= f_2'(t_s) \end{aligned}$$

At $x = 1$ in the formulae (3.5) and (3.6) and by using conditions, we have

$$\begin{aligned} u'(0, t) - u'(0, t_s) &= - (t - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(1) + f_2(t) \\ &\quad - f_2(t_s) + f_1(t_s) - f_1(t) \end{aligned} \quad (3.7)$$

$$\dot{u}'(0, t) = - \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(1) - f_1'(t) + f_2'(t). \quad (3.8)$$

Substituting (3.7) and (3.8) into (3.4)-(3.6) and discretizing the results by assuming $x \rightarrow x_l$ and $t \rightarrow t_{s+1}$ we obtain

$$u''(x_l, t_{s+1}) = (t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} \psi_{nm}(x_l) + u''(x_l, t_s), \quad (3.9)$$

$$\begin{aligned} u'(x_l, t_{s+1}) &= (t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} p_{nm}(x_l) + u'(x_l, t_s) \\ &\quad - (t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(1) + f_2(t_{s+1}) \\ &\quad - f_2(t_s) + f_1(t_s) - f_1(t_{s+1}), \end{aligned} \quad (3.10)$$

$$\begin{aligned}
 u(x_l, t_{s+1}) = & (t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(x_l) + u(x_l, t_s) \\
 & + f_1(t_{s+1}) - f_1(t_s) \\
 & + x_l \left[-(t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(1) \right] \\
 & + x_l [f_2(t_{s+1}) - f_2(t_s) + f_1(t_s) - f_1(t_{s+1})], \tag{3.11}
 \end{aligned}$$

$$\begin{aligned}
 \dot{u}(x_l, t_{s+1}) = & \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(x_l) + f_1'(t_{s+1}) \\
 & + x_l \left[- \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} q_{nm}(1) - f_1'(t_{s+1}) + f_2'(t_{s+1}) \right]. \tag{3.12}
 \end{aligned}$$

Similarly we obtain

$$v''(x_l, t_{s+1}) = (t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} \psi_{nm}(x_l) + v''(x_l, t_s), \tag{3.13}$$

$$\begin{aligned}
 v'(x_l, t_{s+1}) = & (t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} p_{nm}(x_l) + v'(x_l, t_s) \\
 & - (t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} q_{nm}(1) \\
 & + g_2(t_{s+1}) - g_2(t_s) + g_1(t_s) - g_1(t_{s+1}), \tag{3.14}
 \end{aligned}$$

$$\begin{aligned}
 v(x_l, t_{s+1}) = & (t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} q_{nm}(x_l) + v(x_l, t_s) + g_1(t_{s+1}) - g_1(t_s) \\
 & + x_l \left[-(t_{s+1} - t_s) \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} q_{nm}(1) \right] \\
 & + x_l [g_2(t_{s+1}) - g_2(t_s) + g_1(t_s) - g_1(t_{s+1})], \tag{3.15}
 \end{aligned}$$

$$\begin{aligned}
 \dot{v}(x_l, t_{s+1}) = & \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} q_{nm}(x_l) + g_1'(t_{s+1}) \\
 & + x_l \left[- \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} q_{nm}(1) - g_1'(t_{s+1}) + g_2'(t_{s+1}) \right]. \tag{3.16}
 \end{aligned}$$

for $v(x, t)$. Based on the Eqs. (3.9)-(3.12) and (3.13)-(3.16) we will use following equations.

$$\begin{aligned}
 \dot{u}(x_l, t_{s+1}) = & u''(x_l, t_{s+1}) - \eta u(x_l, t_s) u'(x_l, t_s) - \alpha [u(x_l, t_s) v(x_l, t_s)]_x \\
 \dot{v}(x_l, t_{s+1}) = & v''(x_l, t_{s+1}) - \xi v(x_l, t_s) v'(x_l, t_s) - \beta [u(x_l, t_s) v(x_l, t_s)]_x \tag{3.17}
 \end{aligned}$$

Now by substituting (3.9)-(3.12) and (3.13)-(3.16) into (3.17) we obtain

$$\begin{aligned}
\sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} [q_{nm}(x_l) - x_l q_{nm}(1) - \Delta t \psi_{nm}(x_l)] &= u''(x_l, t_s) - \eta u(x_l, t_s) u'(x_l, t_s) \\
&\quad - \alpha [u(x_l, t_s) v(x_l, t_s)]_x - f'_1(t_{s+1}) \\
&\quad + x_l [f'_1(t_{s+1}) - f'_2(t_{s+1})] \\
\sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} d_{nm} [q_{nm}(x_l) - x_l q_{nm}(1) - \Delta t \psi_{nm}(x_l)] &= v''(x_l, t_s) - \eta v(x_l, t_s) v'(x_l, t_s) \\
&\quad - \alpha [u(x_l, t_s) v(x_l, t_s)]_x - g'_1(t_{s+1}) \\
&\quad + x_l [g'_1(t_{s+1}) - g'_2(t_{s+1})] \quad (3.18)
\end{aligned}$$

where $\Delta t = t_{s+1} - t_s$. From the system (3.18) the wavelet coefficients c_{nm} and d_{nm} can be successively calculated. This computation starts with the following initial values.

$$\begin{aligned}
u(x_l, 0) &= \psi_1(x_l), & v(x_l, 0) &= \psi_2(x_l), \\
u'(x_l, 0) &= \psi'_1(x_l), & v'(x_l, 0) &= \psi'_2(x_l), \\
u''(x_l, 0) &= \psi''_1(x_l), & v''(x_l, 0) &= \psi''_2(x_l).
\end{aligned}$$

We collocate x at $x_l = (l - 0.5)/m'$, $l = 1, 2, \dots, m'$ collocation points where $m' = 2^{k-1}M$.

4. Numerical results and discussion

To show the performance of suggested method as compared with the exact solution we are going to use the norms L_2 and L_∞ defined by

$$\begin{aligned}
L_2 &= \sqrt{\frac{\sum_{i=1}^{m'} |u_i^{\text{exact}} - u_i^{\text{num}}|^2}{\sum_{i=1}^{m'} |u_i^{\text{exact}}|^2}}, \\
L_\infty &= \max_i |u_i^{\text{exact}} - u_i^{\text{num}}|.
\end{aligned}$$

We have executed our computations on Intel Core i5-2410M 2.3Ghz and 4GB (667Mhz) of RAM with the codes implemented in free software package GNU Octave and Python programming language. Graphical outputs were generated by Matplotlib package [15].

4.1. Problem 1.

We firstly consider the coupled Burgers' equation (1.1), (1.2) for $\alpha = \beta = \frac{5}{2}$ and $\eta = \xi = -2$, [1]. So we have

$$\begin{aligned}
u_t - u_{xx} - 2uu_x + \frac{5}{2}(u.v)_x &= 0, \\
v_t - v_{xx} - 2vv_x + \frac{5}{2}(u.v)_x &= 0.
\end{aligned}$$

The exact solution of the coupled system for $x \in [0, 1]$, is

$$u(x, t) = v(x, t) = \lambda \left[1 - \tanh \left(\frac{3}{2} \lambda (40(x - 0.5) - 3\lambda t) \right) \right],$$

boundary conditions and initial conditions are determined from the exact solution and λ is an arbitrary constant. This problem has large gradients moving rightward with constant velocity. In Table 1, we show the obtained results for various values of λ . It can be seen from the Table 1 that by increasing the number of collocation points one can achieve more accurate results. On the other hand as time grows errors get larger. Measured cpu times

are also given in Table 1 which are quite small. In Fig. 1, we plot the numerical solution for $\lambda = 0.05, 0.1, 0.4, 0.8$ and $\Delta t = 0.005$ at $t = 3$ and $t = 5$. We see that for greater values of λ , large gradient regions occur in the solution. The present method is capable of analyzing the large gradient regions that occur in the solution which is an indicator of the efficiency of a numerical method according to the Vasilyev and Paolucci [43] and Basdevant *et al.* [5].

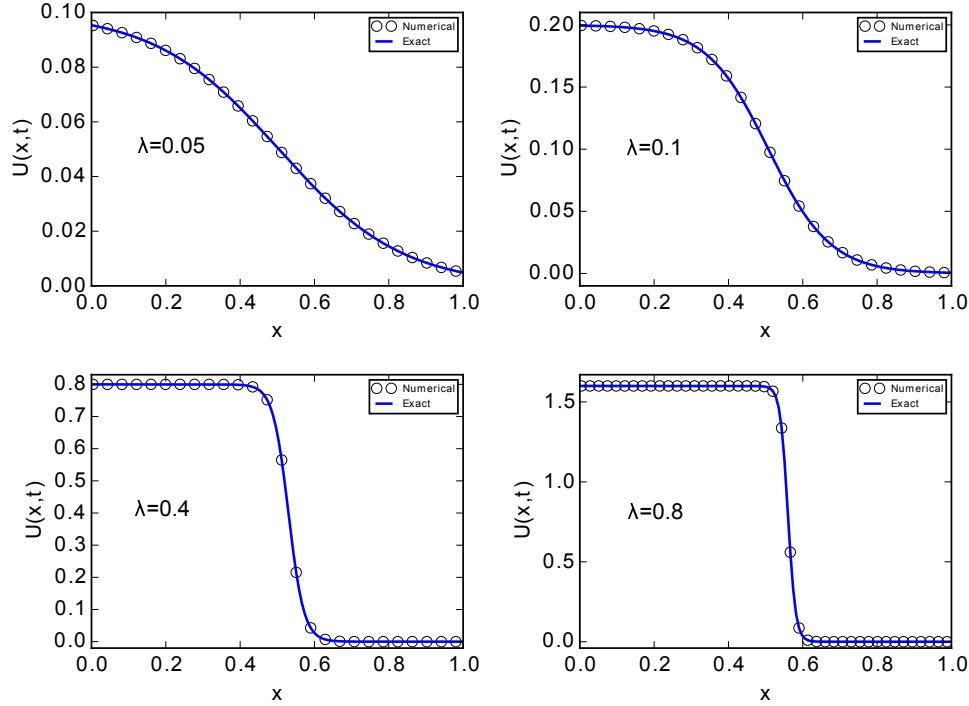


Figure 1. Numerical solution of problem 1 for $\lambda = 0.05, 0.1, 0.4, 0.8$ and $\Delta t = 0.01$ at $t = 1$ with $m' = 128$.

Table 1. Errors for various values of parameters of problem 1.

		$k = 5, M = 4, m' = 64$		$k = 6, M = 4, m' = 128$		$k = 7, M = 4, m' = 256$	
		$t = 3$	$t = 5$	$t = 3$	$t = 5$	$t = 3$	$t = 5$
$\lambda = 0.1$	L_2	8.5052e-6	1.2956e-5	8.4321e-6	1.2802e-5	8.4272e-6	1.2792e-5
	L_∞	2.7482e-6	4.2070e-6	2.7140e-6	4.1204e-6	2.7133e-6	4.1136e-6
$\lambda = 0.8$	L_2	1.0348e-2	1.3720e-2	1.1782e-3	1.3820e-3	7.6982e-4	7.0642e-4
	L_∞	7.7311e-2	1.1248e-1	9.5027e-3	1.2746e-2	7.0736e-3	7.0315e-3
Cpu time		0.458	0.796	0.963	1.505	2.633	4.081

4.2. Problem 2.

We consider the coupled Burgers' equation (1.1), (1.2) for $\eta = \xi = -2$ so that equations (1.1), (1.2) take the following form:

$$\begin{aligned}
 u_t - u_{xx} - 2uu_x + \alpha(u.v)_x &= 0, \\
 v_t - v_{xx} - 2vv_x + \beta(u.v)_x &= 0.
 \end{aligned}$$

The exact solution of this equation is given in [41]

$$u(x, t) = a_0 (1 - \tanh(A(20(x - 0.5) - 2At)))$$

$$v(x, t) = a_0 \left(\left(\frac{2\beta - 1}{2\alpha - 1} \right) - \tanh(A(20(x - 0.5) - 2At)) \right)$$

where

$$a_0 = 0.05, A = \frac{1}{2}a_0 \left(\frac{4\alpha\beta - 1}{2\alpha - 1} \right)$$

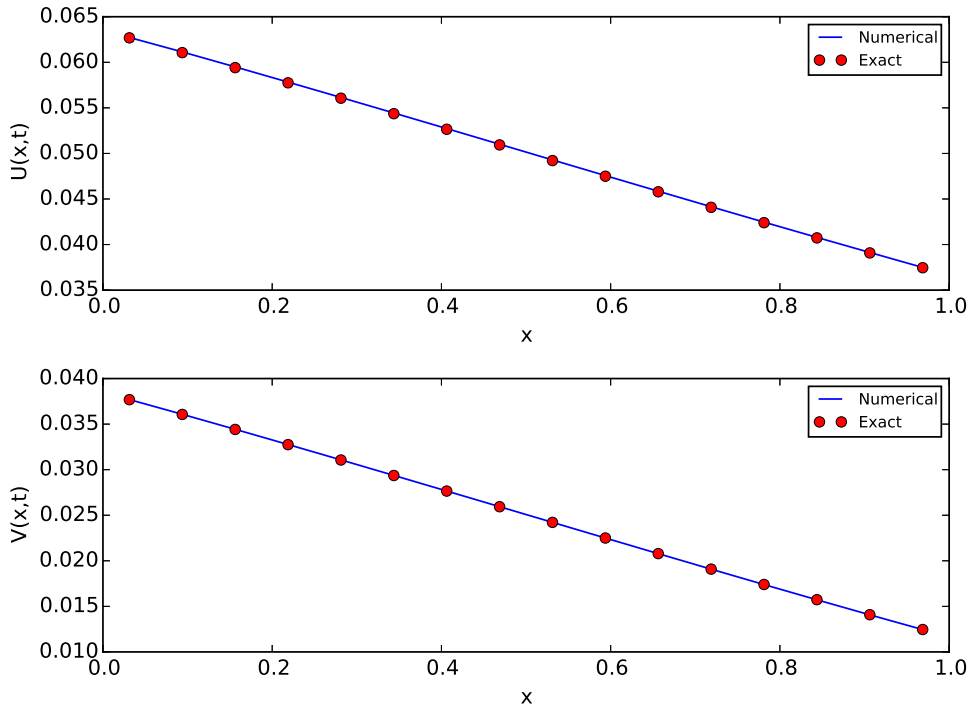
and $x \in [0, 1]$. The initial and boundary conditions are taken from the exact solution. In Table 2, we tabulated and compared the obtained results from the present method with Chebyshev spectral collocation method [21], cubic B-spline collocation method [28], Galerkin quadratic B-spline finite element method [23] and Fourier pseudospectral method [34]. We take $k = 4$, $M = 2$ and $\Delta t = 0.01$. It can be seen from the Table 2, the L_∞ error norm obtained by the present method is smaller than those obtained by the existing methods and even for the small number of collocation points one can achieve the accuracy of the existing methods in the literature. In Table 3, we compare the L_∞ error norms for $\alpha = 0.1$, $\beta = 0.3$, $k = 4$, $M = 2$ and $\Delta t = 0.001$ with the ones obtained by Haar wavelet method [22]. The superiority of the present method over Haar wavelet method in the sense of accuracy is clearly seen from Table 3. The numerical solutions at $t = 1$, are shown in Fig. 2 for $k = 4$, $M = 2$ and $\Delta t = 0.01$.

Table 2. Comparisons of error norms of $u(x, t)$ and $v(x, t)$, for various parameters for Problem 2.

$u(x, t)$		Present			Khater [21]			Mittal [28]			Kutluay [23]			Rashid [34]		
		t	α	β	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞
					$N = 10$ partitions			$N = 100$ partitions			$N = 100$ partitions			$N = 16$ partitions		
					L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞
0.5	0.1	0.3	6.7900e-4	4.1584e-5	1.44e-3	4.38e-5	6.736e-4	4.167e-5	6.783e-4	4.208e-5	6.783e-4	4.208e-5	6.783e-4	3.245e-5	9.619e-4	
	0.3	0.03	7.4791e-4	4.5846e-5	6.68e-4	4.58e-5	7.326e-4	4.590e-5	7.609e-4	4.703e-5	7.609e-4	4.703e-5	7.609e-4	2.733e-5	4.310e-4	
1.0	0.1	0.3	1.3322e-3	8.2147e-5	1.27e-3	8.66e-5	1.325e-3	8.258e-5	1.334e-3	8.320e-5	1.334e-3	8.320e-5	1.334e-3	2.405e-5	1.153e-3	
	0.3	0.03	1.4720e-3	9.1737e-5	1.30e-3	9.16e-5	1.452e-3	9.182e-5	1.500e-3	9.409e-5	1.500e-3	9.409e-5	1.500e-3	2.832e-5	1.268e-3	
$v(x, t)$		Present			Khater [21]			Mittal [28]			Kutluay [23]			Rashid [34]		
		t	α	β	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞
					$N = 10$ partitions			$N = 100$ partitions			$N = 100$ partitions			$N = 16$ partitions		
					L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞	L_2	L_∞
0.5	0.1	0.3	2.0314e-4	9.3908e-6	5.42e-4	4.99e-5	9.057e-4	1.480e-4	5.101e-4	0.221e-4	5.101e-4	0.221e-4	5.101e-4	2.746e-5	3.332e-4	
	0.3	0.03	1.3284e-3	1.8067e-4	1.20e-3	1.81e-4	1.591e-3	5.729e-4	1.327e-3	1.818e-4	1.327e-3	1.818e-4	1.327e-3	2.454e-4	1.148e-3	
1.0	0.1	0.3	9.8534e-4	4.1270e-5	1.29e-3	9.92e-5	1.251e-3	4.770e-5	0.995e-3	4.255e-5	0.995e-3	4.255e-5	0.995e-3	3.745e-5	1.162e-3	
	0.3	0.03	2.6158e-3	3.6136e-4	2.35e-3	3.62e-4	2.250e-3	3.617e-4	2.617e-3	3.636e-4	2.617e-3	3.636e-4	2.617e-3	4.525e-4	1.638e-3	

Table 3. Comparison of error norms for $\alpha = 0.1$, $\beta = 0.3$ and $\Delta t = 0.001$ at various time levels for Problem 2.

t	$L_\infty(u)$		$L_\infty(v)$		Cpu times for the present method
	Present	Haar	Present	Haar	
	$m' = 16$	$N = 64$	$m' = 16$	$N = 64$	
0.5	4.1638e-5	5.675e-5	2.1915e-5	3.679e-5	0.2089
2	1.6239e-4	2.085e-4	7.9455e-5	1.359e-4	0.8045
3	2.3962e-4	3.006e-4	1.1427e-4	2.049e-4	1.2080

**Figure 2.** Numerical solutions of problem 2 for $\alpha = 0.1$, $\beta = 0.3$, $\Delta t = 0.01$ at $t = 1$.

4.3. Problem 3.

Lastly we consider the coupled Burgers' equation (1.1), (1.2) for $\alpha = \beta = 1$ and $\eta = \xi = -2$ so that equations (1.1), (1.2) take the following form:

$$\begin{aligned} u_t - u_{xx} - 2uu_x + (u.v)_x &= 0, \\ v_t - v_{xx} - 2vv_x + (u.v)_x &= 0. \end{aligned}$$

The exact solution is $u(x, t) = v(x, t) = e^{-t} \sin(2\pi(x - 0.5))$, $x \in [0, 1]$. The initial and boundary conditions are taken from the exact solution. Comparison of the error norms at each time for $k = 6$, $M = 2$ and $\Delta t = 0.001$ is given in Table 4. The obtained results by the present method are in good agreement with Haar wavelet method [29] and are better than Finite element method [23]. The physical behavior of numerical solutions for $\alpha = \beta = 1$ and $\eta = \xi = -2$ between $t = 0$ and $t = 2$ and for $\alpha = 3$, $\beta = 2$ and $\eta = 1$, $\xi = -2$ between $t = 0$ and $t = 1.5$ are depicted with contour forms in Fig. 3 and Fig. (4) respectively.

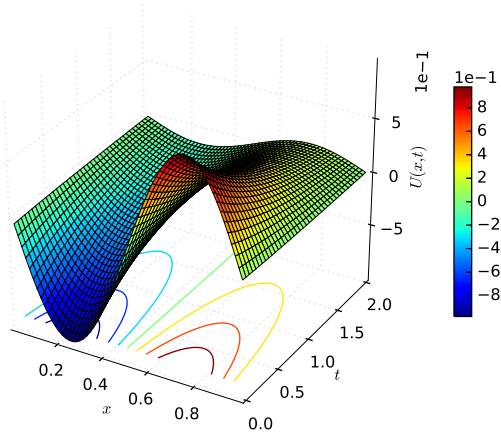


Figure 3. Numerical solution $u(x,t)$ of problem 3 for $\Delta t = 0.025$, $m' = 64$, $\alpha = \beta = 1$ and $\eta = \xi = -2$.

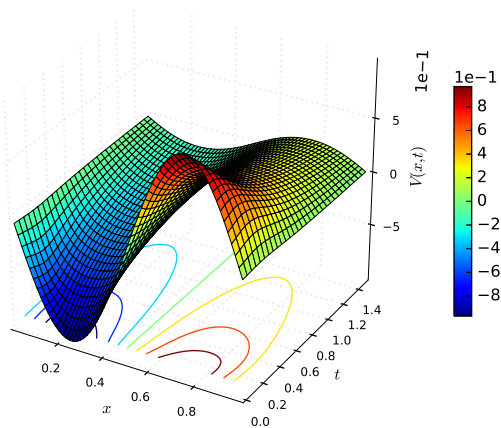


Figure 4. Numerical solution $v(x,t)$ of problem 3 for $\Delta t = 0.025$, $m' = 64$, $\alpha = 3$, $\beta = 2$ and $\eta = 1$, $\xi = -2$.

Table 4. Comparisons of error norms for $\Delta t = 0.001$ at different times for problem 3.

	Kutluay [23]		Mittal [29]		Present method		Cpu Time of Present Method
	(100 partitions)		(64 partitions)		(64 partitions)		
	L_2	L_∞	L_2	L_∞	L_2	L_∞	
$t = 0.01$			1.876e-4	1.986e-5	4.9971e-6	5.0040e-6	0.048
$t = 0.1$	1.396e-4	3.984e-4	1.943e-4	1.382e-5	5.0122e-5	4.9327e-5	0.119
$t = 0.5$	2.473e-4	2.869e-4	2.232e-4	2.119e-5	2.1336e-4	1.4195e-4	0.385
$t = 1$	3.530e-4	1.786e-4	2.676e-4	1.552e-5	2.5548e-4	9.8263e-5	0.712

5. Conclusion

In this paper Chebyshev wavelet method is used to get numerical solutions of one dimensional coupled Burgers' equation. In the solution procedure, the highest derivatives that appeared in the equations are expanded into Chebyshev wavelets and with aid of the integrals of Chebyshev wavelets the considered partial differential equations are converted to algebraic system of equations. The proposed method is tested by three examples and obtained results are compared with the exact solution and with those existed in the literature such as Finite element method, Haar wavelet method and Spectral methods. The comparisons show that the present method is quite satisfactory and competitive with other methods. We can give the highlights of the present method as follows:

- The present method can handle boundary conditions easily.
- Computer simulations show that the proposed method is computationally cheap, fast and gives accurate results even in the case of a small number of collocation points.
- The computer implementation of the proposed method is simple and straightforward.
- The present method can also be used for similar partial differential equations from different branches of science and engineering with suitable modifications.

References

- [1] R. Abazari and A. Borhanifar, *Numerical study of the solution of the Burgers and coupled Burgers equations by a differential transformation method*, Comput. Math. Appl. **59**, 2711-2722, 2010.
- [2] M.A. Abdou and A.A. Soliman, *Variational iteration method for solving Burger's and coupled Burger's equations*, J. Comput. Appl. Math. **181**, 245-251, 2005.
- [3] H. Adibi and P. Assari *Chebyshev wavelet method for numerical solution of Fredholm integral equations of the first kind*. Math. Prob. Eng. **2010**, 1-17, 2010, doi:10.1155/2010/138408.
- [4] E. Babolian and F. Fattahzadeh *Numerical solution of differential equations by using Chebyshev wavelet operational matrix of integration*, Appl. Math. Comput. **188**, 417-426, 2007.
- [5] C. Basdevant, M. Deville, P. Haldenwang and J. M. Lacroix *Spectral and finite difference solutions of the Burgers' equation*, Comput. Fluids **14**, 23-41, 1986.
- [6] F. Bulut, Ö. Oruç and A. Esen, *Numerical Solutions of Fractional System of Partial Differential Equations By Haar Wavelets*, Comput. Model. Eng. Sci. **108** (4), 263-284, 2015.
- [7] I. Celik, *Haar wavelet approximation for magnetohydrodynamic flow equations*, Appl. Math. Model. **37**, 3894-3902, 2013.
- [8] I. Celik, *Chebyshev Wavelet collocation method for solving generalized Burgers-Huxley equation*, Math. Methods. Appl. Sci. **39** (3), 366-377, 2015.
- [9] C. Chen and C.H. Hsiao, *Haar wavelet method for solving lumped and distributed parameter systems*, IEE Proc., Control Theory Appl. **144**, 87-94, 1997.
- [10] I. Daubechies, *Ten Lectures on Wavelet*, SIAM, Philadelphia, 1992.
- [11] M. Dehghan, A. Hamidi and M. Shakourifar, *The solution of coupled Burgers' equations using Adomian-Pade technique*, Appl. Math. Comput. **189**, 1034-1047, 2007.
- [12] S.E. Esipov, *Coupled Burgers' equations: a model of polydispersive sedimentation*, Phys. Rev. Lett. **52**, 3711-3718, 1995.
- [13] A.K. Gupta and S. Saha Ray, *Travelling wave solution of fractional KdV-Burger-Kuramoto equation describing nonlinear physical phenomena*, AIP Adv. **4**, 2014, doi: 10.1063/1.4895910.097120-1-11.

- [14] M.H. Heydari, M.R. Hooshmandasl and F.M. Maalek Ghaini, *A new approach of the Chebyshev wavelets method for partial differential equations with boundary conditions of the telegraph type*, Appl. Math. Model. **38**, 1597-1606, 2014.
- [15] J. D. Hunter, *Matplotlib: A 2D graphics environment*, Comput. Sci. Eng. **9** (3), 90-95, 2007.
- [16] I.E. Inan, D. Kaya and Y. Ugurlu, *Auto-Bäcklund transformation and similarity reductions for coupled Burger's equation*, Appl. Math. Comput. **216**, 2507-2511, 2010.
- [17] S. Islam, S. Haq and M. Uddin, *A mesh free interpolation method for the numerical solution of the coupled nonlinear partial differential equations*, Eng. Anal. Bound. Elem. **33**, 399-409, 2009.
- [18] R. Jiwari, *A Haar wavelet quasilinearization approach for numerical simulation of Burgers' equation*, Comput. Phys. Commun. **183**, 2413-2423, 2012.
- [19] D. Kaya, *An explicit solution of coupled viscous Burgers' equation by the decomposition method*. Int. J. Math. Math. Sci. **27** (11), 675-680, 2001.
- [20] A. Kelleci and A. Yıldırım, *An efficient numerical method for solving coupled Burgers' equation by combining homotopy perturbation and pade techniques*, Numer. Methods Partial Differ. Equ. **27** (4), 982-995, 2011.
- [21] A.H. Khater, R.S. Temsah and M.M. Hassan, *A Chebyshev spectral collocation method for solving Burgers'-type equations*, J. Comput. Appl. Math. **222** (2), 333-350, 2008.
- [22] M. Kumar and S. Pandit, *A composite numerical scheme for the numerical simulation of coupled Burgers' equation*, Comput. Phys. Commun. **185** (3), 809-817, 2014.
- [23] S. Kutluay and Y. Ucar, *Numerical solutions of the coupled Burgers' equation by the Galerkin quadratic B-spline finite element method*, Math. Methods Appl. Sci. **36** (17) 2403-2415, 2013.
- [24] U. Lepik, *Numerical solution of differential equations using Haar wavelets*, Math. Comput. Simul. **68**, 127-143, 2005.
- [25] U. Lepik, *Application of the Haar wavelet transform to solving integral and differential Equations*, Proc. Estonian Acad. Sci. Phys. Math. **56** (1), 28-46, 2007.
- [26] U. Lepik, *Numerical solution of evolution equations by the Haar wavelet method*, Appl. Math. Comput. **185**, 695-704, 2007.
- [27] U. Lepik, *Solving PDEs with the aid of two-dimensional Haar wavelets*, Comput. Math. Appl. **61**, 1873-1879, 2011.
- [28] R.C. Mittal and G. Arora, *Numerical solution of the coupled viscous Burgers' equation*, Commun. Nonlinear Sci. Numer. Simul. **16**, 1304-1313, 2011.
- [29] R.C. Mittal, H. Kaur, and V. Mishra, *Haar wavelet-based numerical investigation of coupled viscous Burgers' equation*, Int. J. Comput. Math. **92** (8), 1643-1659, 2015, doi: 10.1080/00207160.2014.957688
- [30] J. Nee and J. Duan, *Limit set of trajectories of the coupled viscous Burgers' equations*, Appl. Math. Lett. **11** (1), 57-61, 1998.
- [31] Ö. Oruç, F. Bulut and A. Esen, *A haar wavelet-finite difference hybrid method for the numerical solution of the modified Burgers' equation*, J. Math. Chem. **53** (7), 1592-1607, 2015.
- [32] Ö. Oruç, F. Bulut and A. Esen, *Numerical Solutions of Regularized Long Wave Equation By Haar Wavelet Method*, Mediterr. J. Math. **13** (5), 3235-3253, 2016.
- [33] A. Rashid, M. Abbas, A.I.Md. Ismail and A. Abd Majid, *Numerical solution of the coupled viscous Burgers equations by Chebyshev-Legendre Pseudo-Spectral method*, Appl. Math. Comput. **245**, 372-381, 2014.
- [34] A. Rashid and A.I.B.M. Ismail, *A Fourier pseudospectral method for solving coupled viscous Burgers equations*, Comput. Methods Appl. Math. **9** (4), 412-420, 2009.
- [35] M. Razzaghi and S. Yousefi, *Legendre wavelets direct method for variational problems*, Math. Comput. Simul. **53**, 185-192, 2000.

- [36] M. Razzaghi and S. Yousefi, *Legendre wavelets operational matrix of integration*, Int. J. Syst. Sci. **32** (4), 49-502, 2001.
- [37] Z. Rong-Pei, Y. Xi-Jun and Z. Guo-Zhong, *Local discontinuous Galerki nmethod for solving Burgers and coupled Burgers equations*, Chin. Phys. B. **20** (11), 110205-1-6, 2011.
- [38] P.K. Sahu and S. Saha Ray, *Legendre wavelets operational method for the numerical solutions of nonlinear Volterra integro-differential equations system*, Appl. Math. Comput. **256**, 715-723, 2015.
- [39] P.K. Sahu and S. Saha Ray, *Two dimensional Legendre wavelet method for the numerical solutions of fuzzy integro-differential equations*, J. Intell. Fuzzy Syst. **28**, 1271-1279, 2015.
- [40] Z. Shi, Y. Cao, and Q.J. Chen, *Solving 2D and 3D Poisson equations and biharmonic equations by the Haar wavelet method*, Appl. Math. Model. **36**, 5143-5161, 2012.
- [41] A.A. Soliman, *The modified extended tanh-function method for solving Burgers-type equations*, Physica A **361**, 394-404, 2006.
- [42] V.K. Srivastava, M.K. Awasthi and M. Tamsir, *A fully implicit Finite-difference solution to one dimensional Coupled Nonlinear Burgers' equations*, Int. J. Math. Sci. **7** (4) 2013.
- [43] O.V. Vasilyev and S. Paolucci, *A Dynamically Adaptive Multilevel Wavelet Collocation Method for Solving Partial Differential Equations in a Finite Domain*, J. Comput. Phys. **125**, 498-512, 1996.
- [44] Y. Wang and Q. Fan, *The second kind Chebyshev wavelet method for solving fractional differential equations*, Appl. Math. Comput. **218**, 8592-8601, 2012.
- [45] C. Yang and J. Hou, *Chebyshev wavelets method for solving Bratu' s problem*, Bound. Value. Probl. **142**, 1-9, 2013.
- [46] F. Zhou and X. Xu, *Numerical solution of the convection diffusion equations by the second kind Chebyshev wavelets*, Appl. Math. Comput. **247**, 353-367, 2014.
- [47] L. Zhu and Q. Fan, *Solving fractional nonlinear Fredholm integro-differential equations by the second kind Chebyshev wavelet*, Commun. Nonlinear Sci Numer. Simul. **17**, 2333-2341, 2012.



Comparison of some set open and uniform topologies and some properties of the restriction maps

Lamia Harkat¹, Abderrahmane Bouchair*², Smail Kelaiaia¹

¹*LMA Laboratory, University of Badji Mokhtar Annaba, 23000 Annaba, Algeria*

²*LMAM Laboratory, University of Mohamed Seddik Ben Yahia-Jijel, 18000 Jijel, Algeria*

Abstract

Let X be a Tychonoff space, Y an equiconnected space and $C(X, Y)$ be the set of all continuous functions from X to Y . In this paper, we provide a criterion for the coincidence of set open and uniform topologies on $C(X, Y)$ when these topologies are defined by a family α consisting of Y -compact subsets of X . For a subspace Z of a topological space X , we also study the continuity and the openness of the restriction map $\pi_Z : C(X, Y) \rightarrow C(Z, Y)$ when both $C(X, Y)$ and $C(Z, Y)$ are endowed with the set open topology.

Mathematics Subject Classification (2010). 54C35

Keywords. function space, set open topology, topology of uniform convergence on a family of sets, restriction map, Y -compact

1. Introduction

Let X, Y be topological spaces and $C(X, Y)$ be the set of all continuous functions from X to Y . The set $C(X, Y)$ has a number of classical topologies; among them the topology of uniform convergence and the set open topology. Since their introduction by Arens and Dugundji [1], set open topologies have been studied and the comparison between them and the topology of uniform convergence have been considered by many authors (see, for example, [4, 7, 9, 10]).

In [4], Bouchair and Kelaiaia have established a criterion for the coincidence of the set open topology and the topology of uniform convergence on $C(X, Y)$ defined on a family α of compact subsets of X . They also have studied the comparison between some set open topologies on $C(X, Y)$ for various families α . In this paper we continue the study of the comparison between these topologies in the case when α is a family consisting of Y -compact sets and give a criterion for their coincidence.

One of the most useful tools normally used for studying function spaces is the concept of restriction map. If Z is a subspace of a topological space X , then the restriction map $\pi_Z : C(X, Y) \rightarrow C(Z, Y)$ is defined by $\pi_Z(f) = f|_Z$ for any $f \in C(X, Y)$. The properties of the restriction map $\pi_Z : C(X, \mathbb{R}) \rightarrow C(Z, \mathbb{R})$, when both $C(X, \mathbb{R})$ and $C(Z, \mathbb{R})$ are endowed with the topology of the pointwise convergence, have been studied by Arhangel'skii in

*Corresponding Author.

Email addresses: harkat-lamia@yahoo.fr (L. Harkat), abdebouchair@yahoo.com (A. Bouchair), kelaiaiasmail@yahoo.fr (S. Kelaiaia)

Received: 27.09.2016; Accepted: 31.07.2017

[2, 3]. In the present paper, we give a criteria for the continuity and for the openness of the restriction map in the case when Y is an equiconnected topological space and $C(X, Y)$ and $C(Z, Y)$ are equipped with the set open topology.

Our paper is organized as follows. In Section 3, we prove that the set open and uniform topologies on $C(X, Y)$ coincide if and only if α is a functional refinement family. Section 4 is devoted to compare the spaces $C_\alpha(X, Y)$ and $C_\beta(X, Y)$ for two given families α and β of Y -compact subsets of X . In Section 5, we consider, for a subspace Z of a topological space X , the restriction map $\pi_Z : C_\alpha(X, Y) \rightarrow C_\beta(Z, Y)$ and we give necessary and sufficient condition for the continuity and for the openness of the restriction map in the framework of set open topology. We prove that, if α is a functional refinement family consisting of closed Y -compact subsets of X and β is a family of closed Y -compact subsets of Z , then π_Z is continuous if and only if the quadruplet (β, α, X, Y) satisfies the property (P) . We also show that, if α and β are two admissible families of compact subsets of X and Z respectively, then $\pi_Z : C_\alpha(X, Y) \rightarrow C_\beta(Z, Y)$ is open onto its image if and only if β approximates $\alpha|_Z$.

2. Definitions and preliminaries

Throughout this paper, X is a Tychonoff space, Y is an equiconnected topological space, $C(X, Y)$ is the set of all continuous functions from X to Y , and α is always a nonempty family of subsets of X . The set open topology on $C(X, Y)$ has a subbase consisting of all sets of the form $[A, V] = \{f \in C(X, Y) : f(A) \subseteq V\}$, where $A \in \alpha$ and V is an open subset of Y , and the function space $C(X, Y)$ endowed with this topology is denoted by $C_\alpha(X, Y)$. If V is not arbitrary but is restricted to some collection \mathcal{B} of open subsets of Y , then we denote by $C_\alpha^{\mathcal{B}}(X, Y)$ the corresponding function space.

For a metric space (Y, ρ) , the topology of uniform convergence on members of α has as base at each point $f \in C(X, Y)$ the family of all sets of the form

$$\langle f, A, \epsilon \rangle = \{g \in C(X, Y) : \sup_{x \in A} \rho(f(x), g(x)) < \epsilon\},$$

where $A \in \alpha$ and $\epsilon > 0$. The space $C(X, Y)$ having the topology of uniform convergence on α is denoted by $C_{\alpha, u}(X, Y)$.

The symbols \emptyset and \mathbb{N} will stand for the empty set and the positive integers, respectively. We denote by \mathbb{R} the real numbers with the usual topology. The complement and the closure of a subset A in X is denoted by A^c and \bar{A} , respectively. If $A \subseteq X$, the restriction of a function $f \in C(X, Y)$ to the set A is denoted by $f|_A$. Let Z be a subspace of X , then $\alpha|_Z$ denotes the family $\{A \cap Z : A \in \alpha\}$.

Let β be a nonempty family of subsets of X . We say that α refines β if every member of α is contained in some member of β . We say that β approximates α provided that for every $A \in \alpha$ and every open neighborhood U of A in X , there exist $B_1, \dots, B_n \in \beta$ such that $A \subseteq B_1 \cup B_2 \cup \dots \cup B_n \subseteq U$. A family α is said to be admissible if for every $A \in \alpha$ and every finite sequence U_1, \dots, U_n of open subsets of X such that $A \subseteq \bigcup_{i=1}^n U_i$, there exists a finite sequence A_1, \dots, A_m of members of α which refines U_1, \dots, U_n and whose union contains A . For example, the family of all compact sets as well as the family of all finite sets in a topological space is an admissible family.

A topological space Y is said to be equiconnected [6] if there exists a continuous map $\Psi : Y \times Y \times [0, 1] \rightarrow Y$ such that $\Psi(x, y, 0) = x$, $\Psi(x, y, 1) = y$, and $\Psi(x, x, t) = x$ for all $x, y \in Y$ and $t \in [0, 1]$. The map Ψ is called an equiconnecting function. A subset V of an equiconnected space Y is called a Ψ -convex subset of Y if $\Psi(V, V, [0, 1]) \subseteq V$. It is a known fact that any topological vector space or any convex subset of any topological vector space is an equiconnected space, and any equiconnected space is a pathwise connected space.

For topological space X and Y , we write $X = Y$ ($X \leq Y$) to mean that X and Y have the same underlying set and the topology on Y is the same to (finer than or equal to) the topology on X .

Definition 2.1. ([10]). Let $A \subseteq X$ and let Y be an arbitrary topological space. For a fixed natural number n , we will say that A is Y^n -compact if, for any continuous function $f \in C(X, Y^n)$, the set $f(A)$ is compact in Y^n .

We would like to mention that there are Y -compact sets which are not closed. Indeed, it is proved in [10, Example 1] that if X is the set of all ordinals that are less than or equal to ω_1 and $Y = \mathbb{R}$, then the subset of all countable ordinals from X is \mathbb{R} -compact but it is not closed in X . It was proved also that there are closed sets that are not Y -compact, see [10, Example 4]. So in our comparison of topologies on $C(X, Y)$ we consider the family α in the class of closed and Y -compact sets. Notice that, in the case when $A = X$ and $Y = \mathbb{R}$, the \mathbb{R} -compactness of the set A coincides with the pseudocompactness of the space X .

Definition 2.2. ([11]). A space Y is called *cub*-space (or *quadra*-space) if for any $x \in Y \times Y$ there are a continuous map f from $Y \times Y$ to Y and a point $y \in Y$ such that $f^{-1}(y) = x$.

For example, any Tychonoff space with G_δ -diagonal containing a nontrivial path or a zero-dimensional space with G_δ -diagonal containing a nontrivial convergent sequence is a cub-space. Also a pathwise connected metric space is a *cub*-space.

Proposition 2.3. Let X be a topological space, and (Y, ρ) be a pathwise connected metric space. If A is an Y -compact subset of X and n is a natural number, then A is also an Y^n -compact subset of X .

Proof. For the proof see Proposition 2.2 in [11]. □

The following lemma will be useful in the sequel which is a particular case of the Proposition 2.4 in [11].

Lemma 2.4. Let X be a topological space and (Y, ρ) be a pathwise connected metric space. Then the intersection of an Y -compact subset of X and the inverse image, by any continuous function from X to Y , of any closed subset of Y is Y -compact.

Proof. Let $A \subseteq X$ be an Y -compact set, F a closed subset of Y , and $g \in C(X, Y)$. We will show that $A \cap g^{-1}(F)$ is Y -compact. Take an arbitrary element f of $C(X, Y)$ and prove that $f(A \cap g^{-1}(F))$ is compact in Y^2 . Define the function $h : X \rightarrow Y^2$ by $h(x) = (f(x), g(x))$, for every $x \in X$. It is clear that the function h is continuous. By Proposition 2.3, the set $h(A)$ is compact in Y^2 . Consider the set $T = Y \times F$. We claim that $h(A \cap g^{-1}(F)) = h(A) \cap T$. Indeed, if $z \in h(A \cap g^{-1}(F))$, then $z \in h(A)$. On the other hand, there exists a point $x \in A \cap g^{-1}(F)$ such that $z = h(x) = (f(x), g(x))$. Therefore $z \in T$. Conversely, let $y \in h(A) \cap T$. Then, $y = h(x) = (f(x), g(x))$ for some $x \in A$. Since $y \in T$, we have $x \in g^{-1}(F)$. Therefore, $x \in A \cap g^{-1}(F)$ and $y \in h(A \cap g^{-1}(F))$. Thus, $h(A \cap g^{-1}(F)) = h(A) \cap T$ which is compact as the intersection of the compact set $h(A)$ with the closed set T . To finish the proof of the lemma it suffices to see that $f(A \cap g^{-1}(F))$ is the projection of the set $h(A \cap g^{-1}(F))$ on T . □

We give the following definition.

Definition 2.5. A family α of subsets of X is called a functional refinement if for every $A \in \alpha$, every finite sequence U_1, \dots, U_n of open subsets of Y , and every $f \in C(X, Y)$ such that $A \subseteq \bigcup_{i=1}^n f^{-1}(U_i)$, there exists a finite sequence A_1, \dots, A_m of members of α which refines $f^{-1}(U_1), \dots, f^{-1}(U_n)$ and whose union contains A .

It is clear that every admissible family is a functional refinement family.

Proposition 2.6. *Let X be a topological space, and (Y, ρ) be a metric space. Then, the family of all Y -compact subsets of X is a functional refinement family.*

Proof. Let $A \subseteq X$ be an Y -compact set, $\{U_1, \dots, U_n\}$ a finite sequence of open subsets of Y , and $f \in C(X, Y)$ such that $A \subseteq \bigcup_{i=1}^n f^{-1}(U_i)$. For every $y \in f(A)$, there exist $i_y \in \{1, \dots, n\}$ and $\varepsilon_y > 0$ such that $\overline{B(y, \varepsilon_y)} \subseteq U_{i_y}$, where $B(y, \varepsilon_y)$ is the open ball with center y and radius ε_y . Then the family $\{B(y, \varepsilon_y) : y \in f(A)\}$ is an open covering of $f(A)$. From this covering, let us choose a finite subcover $\{B(y_j, \varepsilon_{y_j})\}_{j=1}^m$ and for each $j = 1, \dots, m$, let us choose some i_{y_j} such that $\overline{B(y_j, \varepsilon_{y_j})} \subseteq U_{i_{y_j}}$. By Lemma 2.4, the set $A_j = f^{-1}(\overline{B(y_j, \varepsilon_{y_j})}) \cap A$ is Y -compact, for each j . Therefore, the family A_1, \dots, A_m of Y -compact subsets of X covers A and refines $f^{-1}(U_1), \dots, f^{-1}(U_n)$. Hence, the family of all Y -compact subsets of X is a functional refinement. \square

3. Coincidence of set open and uniform topologies

In this section, we study necessary and sufficient condition for the coincidence of the set open topology and the uniform topology on $C(X, Y)$ in the case when the family α consists of closed Y -compact sets. We first give subbase for the space $C_\alpha(X, Y)$ that help us to study the comparison of these topologies.

Theorem 3.1. *Let α be a functional refinement family consisting of Y -compact subsets of X and \mathcal{B} be an arbitrary base for Y . Then, the family*

$$\{[A, V] : A \in \alpha, V \in \mathcal{B}\}$$

is a subbase for the space $C_\alpha(X, Y)$.

Proof. Let $f \in C(X, Y)$ and take a subbasic open neighborhood $[K, U]$ of f in $C_\alpha(X, Y)$, where $K \in \alpha$ and U open in Y . The open set U will be written as the union of some subfamily $\{V_i : i \in I\}$ of \mathcal{B} which covers $f(K)$. Since $f(K)$ is compact, there exists $n \in \mathbb{N}^*$ such that $f(K) \subseteq \bigcup_{i=1}^n V_i$. Since α is a functional refinement, there exists a sequence K_1, \dots, K_m of members of α which refines $\{f^{-1}(V_i) : i = 1, \dots, n\}$ and whose union contains K . For each $j \in \{1, \dots, m\}$, let us choose $i_j \in \{1, \dots, n\}$ such that $K_j \subseteq f^{-1}(V_{i_j})$. It is easy to see that $f \in \bigcap_{j=1}^m [K_j, V_{i_j}] \subseteq [K, U]$. \square

The following result was obtained in [11, Theorem 3.3].

Theorem 3.2. *For every Hausdorff space X and any uniform cub-space (Y, \mathcal{U}) the topology on $C(X, Y)$ induced by the uniformity $\hat{\mathcal{U}}|_\alpha$ of uniform convergence on the saturation family α coincides with the set open topology on $C(X, Y)$, where Y has the topology induced by \mathcal{U} .*

Because every equiconnected metric space is a cub-space, the following result follows immediately from the above theorem.

Theorem 3.3. *Let X be a topological space, and (Y, ρ) be an equiconnected metric space. If α is a functional refinement family consisting of Y -compact subsets of X , then $C_\alpha(X, Y) = C_{\alpha, u}(X, Y)$.*

Corollary 3.4. *Let X be a topological space, and (Y, ρ) be an equiconnected metric space. If α is a family consisting of Y -compact subsets of X which contains all Y -compact subsets of its elements, then $C_\alpha(X, Y) = C_{\alpha, u}(X, Y)$.*

We will now give a necessary and sufficient condition for which $C_\alpha^B(X, Y) = C_{\alpha, u}(X, Y)$. To this end, we will introduce, for a family α of subsets of X , the following family

$$\alpha_1 = \{A'/A' \text{ is } Y\text{-compact subset of } X \text{ and } \exists A \in \alpha : A' \subseteq A\}.$$

Proposition 2.6 leads us to the following corollary.

Corollary 3.5. *If α is a family of Y -compact subsets of X , then the family α_1 is always a functional refinement family.*

We give a definition.

Definition 3.6. Let X and Y be two topological spaces. Let α and β be two families of subsets of X . We will say that the quadruplet (α, β, X, Y) satisfies the property (P) provided that for every $A \in \alpha$, every open subset U in Y , and every $f \in C(X, Y)$ such that $A \subseteq f^{-1}(U)$, there exist $B_1, \dots, B_n \in \beta$ with $A \subseteq \bigcup_{i=1}^n B_i \subseteq f^{-1}(U)$.

From the above definition, we observe that if a family β approximates α then (α, β, X, Y) satisfies the property (P) .

Lemma 3.7. *Let α be a family of Y -compact subsets of X . Then α is a functional refinement if and only if the quadruplet (α_1, α, X, Y) satisfies the property (P) .*

Proof. Suppose that α is a functional refinement family. We will show that (α_1, α, X, Y) satisfies the property (P) . Let $A' \in \alpha_1$, U be an open subset of Y and $f \in C(X, Y)$ such that $A' \subseteq f^{-1}(U)$. Let $A \in \alpha$, with $A' \subseteq A$. If $A \subseteq f^{-1}(U)$ the proof is finished. If $A \not\subseteq f^{-1}(U)$; then the family $\{f^{-1}(f(A')^c), f^{-1}(U)\}$ is an open cover of A . Since α is a functional refinement family, there exists a finite sequence A_1, \dots, A_n of elements of α which refines $\{f^{-1}(f(A')^c), f^{-1}(U)\}$ and whose union contains A . Put $J = \{i : A_i \subseteq f^{-1}(U)\}$. It is clear that $A' \subseteq \bigcup_{i \in J} A_i \subseteq f^{-1}(U)$.

Conversely, suppose that (α_1, α, X, Y) satisfies the property (P) . Let $A \in \alpha$, $\{U_1, \dots, U_n\}$ a finite family of open subsets of Y and let $f \in C(X, Y)$ such that $A \subseteq \bigcup_{i=1}^n f^{-1}(U_i)$. From Proposition 2.6, there exists a finite sequence A_1, \dots, A_m of Y -compact subsets of X which refines $\{f^{-1}(U_1), \dots, f^{-1}(U_n)\}$ and whose union contains A . We set $A'_j = A_j \cap A$ for each $1 \leq j \leq m$. Then the subfamily $\{A'_1, \dots, A'_m\}$ of α_1 covers A and refines $f^{-1}(U_1), \dots, f^{-1}(U_n)$. For each $j = 1, \dots, m$, let us choose some i_j such that $A'_j \subseteq f^{-1}(U_{i_j})$. By our hypothesis there is, for every $j = 1, \dots, m$, a finite family $\mathcal{A}_j = \{A_1^j, \dots, A_{m_j}^j\}$ of members of α , such that $A'_j \subseteq \bigcup_{k=1}^{m_j} A_k^j \subseteq f^{-1}(U_{i_j})$. Put $\mathcal{A} = \bigcup_{i=1}^m \mathcal{A}_i$. This is a finite family of elements of α which covers A and refines $f^{-1}(U_1), \dots, f^{-1}(U_n)$. Therefore α is a functional refinement family. \square

Corollary 3.8. *For any family α of Y -compact subsets of X , we have $C_\alpha(X, Y) \leq C_{\alpha_1}(X, Y) = C_{\alpha_1, u}(X, Y) = C_{\alpha, u}(X, Y)$.*

Let $\bar{\alpha} = \{\bar{A} : A \in \alpha\}$. We have the following result.

Proposition 3.9. *Let X be a topological space, and (Y, ρ) be a metric space. For any family α of Y -compact subsets of X , we have $C_{\alpha, u}(X, Y) = C_{\bar{\alpha}, u}(X, Y)$.*

Proof. Let us prove that $C_{\bar{\alpha}, u}(X, Y) = C_{\alpha, u}(X, Y)$. Let $A \in \alpha$, $\varepsilon > 0$ and $f \in C(X, Y)$. As $A \subseteq \bar{A}$, we have $\langle f, \bar{A}, \varepsilon \rangle \subseteq \langle f, A, \varepsilon \rangle$, then $C_{\alpha, u}(X, Y) \leq C_{\bar{\alpha}, u}(X, Y)$. Let $\langle f, \bar{A}, \varepsilon \rangle$ be an open subbasic set of $C_{\bar{\alpha}, u}(X, Y)$. Let us show that $\langle f, A, \frac{\varepsilon}{3} \rangle \subseteq \langle f, \bar{A}, \varepsilon \rangle$. Let $g \in \langle f, A, \frac{\varepsilon}{3} \rangle$ and x be an arbitrary point of the set \bar{A} . Since g and f are continuous, then for every $\varepsilon > 0$, there exists a point $y_\varepsilon \in A$ such that $\rho(f(x), f(y_\varepsilon)) < \frac{\varepsilon}{3}$ and $\delta(g(x), g(y_\varepsilon)) < \frac{\varepsilon}{3}$

(this is possible, since $f^{-1}(B(f(x), \frac{\varepsilon}{3})) \cap g^{-1}(B(g(x), \frac{\varepsilon}{3}))$ is a neighborhood of the point x and $x \in \bar{A}$). Hence $\rho(f(x), g(x)) \leq \rho(f(x), f(y)) + \rho(f(y), g(y)) + \rho(g(x), g(y)) < \varepsilon$; therefore $C_{\bar{\alpha}, u}(X, Y) \leq C_{\alpha, u}(X, Y)$. \square

Theorem 3.10. *Let X be a topological space, and (Y, ρ) be an equiconnected metric space with a bounded metric ρ and having a base \mathcal{B} consisting of Ψ -convex sets. Let α be a family of closed Y -compact subsets of X . Then $C_{\alpha}^{\mathcal{B}}(X, Y) = C_{\alpha, u}(X, Y)$ if and only if α is a functional refinement family.*

Proof. If α is a functional refinement family, then by Theorems 3.1 and 3.3 we have $C_{\alpha}^{\mathcal{B}}(X, Y) = C_{\alpha}(X, Y) = C_{\alpha, u}(X, Y)$. Conversely, suppose that $C_{\alpha}^{\mathcal{B}}(X, Y) = C_{\alpha, u}(X, Y)$ and let us show that α is a functional refinement family. From Lemma 3.7 it suffices to prove that (α_1, α, X, Y) satisfies the property (P). Let $A' \in \alpha_1$, $f \in C(X, Y)$ and let U be an open subset in Y such that $A' \subseteq f^{-1}(U)$. Since $f(A')$ is compact, there exists a continuous function $g : Y \rightarrow [0, 1]$ such that $g(f(A')) = \{1\}$ and $g(U^c) = \{0\}$. Take a nontrivial path p in Y with $p(0) \neq p(1)$, and put $h = p \circ g \circ f$. Since the topologies of the spaces $C_{\alpha_1, u}(X, Y)$ and $C_{\alpha, u}(X, Y)$ coincide, the set $\langle h, A', \epsilon \rangle$, where ϵ is strictly inferior to the distance between $p(0)$ and $p(1)$ in Y , is an open neighborhood of h in $C_{\alpha, u}(X, Y)$. Moreover, since $C_{\alpha}^{\mathcal{B}}(X, Y) = C_{\alpha, u}(X, Y)$ there exist $A_1, \dots, A_n \in \alpha$ and $V_1, \dots, V_n \in \mathcal{B}$ such that

$$h \in \bigcap_{i=1}^n [A_i, V_i] \subseteq \langle h, A', \epsilon \rangle.$$

Equiconnectedness of Y leads us, by [12, Corollary 1], to the fact that $A' \subseteq \bigcup_{i=1}^n A_i$. We set $\mathcal{J} = \{i : A_i \subseteq f^{-1}(U)\}$. By the same argument as in [4, Theorem 3], it follows that $A' \subseteq \bigcup_{i \in \mathcal{J}} A_i$. This means that (α_1, α, X, Y) satisfies the property (P), and so the family α is a functional refinement. \square

Corollary 3.11. *Let X be a topological space, and (Y, ρ) be an equiconnected metric space with a bounded metric ρ and having a base \mathcal{B} consisting of Ψ -convex sets. Let α be a family of closed Y -compact subsets of X . Then $C_{\alpha}^{\mathcal{B}}(X, Y) = C_{\alpha_1}(X, Y)$ if and only if α is a functional refinement family.*

4. Comparison of $C_{\alpha}(X, Y)$ and $C_{\beta}(X, Y)$

In this section, we are going to compare the topologies of $C_{\alpha}(X, Y)$ and $C_{\beta}(X, Y)$ when α and β are two families of Y -compact subsets of X .

Theorem 4.1. *Let α and β be two families of subsets of X . If (α, β, X, Y) satisfies the property (P), then $C_{\alpha}(X, Y) \leq C_{\beta}(X, Y)$.*

Proof. The proof is the same of [4, Theorem 5]. \square

Theorem 4.2. *Let α and β be two families of closed Y -compact subsets of X , and Y be an equiconnected topological space having a base \mathcal{B} consisting of Ψ -convex sets. If $C_{\alpha}(X, Y) \leq C_{\beta}^{\mathcal{B}}(X, Y)$, then the quadruplet (α, β, X, Y) satisfies the property (P).*

Proof. Let $A \in \alpha$, V an open subset in Y and $f \in C(X, Y)$ such that $A \subseteq f^{-1}(V)$. Since $f(A)$ is compact in Y , there exists a continuous function $g : Y \rightarrow [0, 1]$ such that $g(f(A)) = \{0\}$ and $g(V^c) = \{1\}$; Let $p : [0, 1] \rightarrow Y$ be a path in Y with $p(0) \neq p(1)$, and let $h = p \circ g \circ f$. Let $W \in \mathcal{B}$ which contains the point $p(0)$ and does not contain $p(1)$. Then $[A, W]$ is an open neighborhood of h in $C_{\alpha}^{\mathcal{B}}(X, Y)$. Since the topology of $C_{\beta}^{\mathcal{B}}(X, Y)$ is finer than the topology of $C_{\alpha}(X, Y)$, there exist $B_1, \dots, B_n \in \beta$ and $V_1, \dots, V_n \in \mathcal{B}$ such that

$$h \in \bigcap_{i=1}^n [B_i, V_i] \subseteq [A, W].$$

We have then $A \subseteq \bigcup_{i=1}^n B_i$. Put $I = \{i : B_i \subseteq f^{-1}(V)\}$. As in the proof of Theorem 3.10, we obtain that $A \subseteq \bigcup_{i \in I} B_i$ and hence (α, β, X, Y) satisfies the property (P). \square

Corollary 4.3. *Let α and β be two families consisting of closed Y -compact subsets of X , and Y be an equiconnected topological space having a base \mathcal{B} consisting of Ψ -convex sets. If β is a functional refinement family, then $C_\alpha(X, Y) \leq C_\beta(X, Y)$ if and only if (α, β, X, Y) satisfies the property (P).*

5. Restriction map

In this section, we use the results obtained above to study and generalize some results due to Arhangel'skii about the properties of the so-called restriction map on function spaces. Let Z be a subspace of a topological space X . The restriction map $\pi_Z : C(X, Y) \rightarrow C(Z, Y)$ is defined by $\pi_Z(f) = f|_Z$ for any $f \in C(X, Y)$. We begin by examining the continuity of π_Z . The following result is stated in [3].

Proposition 5.1. [3, Proposition 1] *Let X, Y be topological spaces, and Z be a subspace of X . Let α be a network in X and β be a network in Z . If $\beta \subset \alpha$, then the restriction map $\pi_Z : C_\alpha(X, Y) \rightarrow C_\beta(Z, Y)$ is continuous.*

Proposition 5.1 can be strengthened as follows.

Proposition 5.2. *Let X, Y be topological spaces, and Z be a subspace of X . Let α be a family of subsets of X and β be a family of subsets of Z . If (β, α, X, Y) satisfies the property (P), then $\pi_Z : C_\alpha(X, Y) \rightarrow C_\beta(Z, Y)$ is continuous.*

Proof. Let $f \in C(X, Y)$ and $[B, V]$ be an open neighborhood of $f|_Z$ in $C_\beta(Z, Y)$, where $B \in \beta$ and V open in Y . Then $B \subseteq f^{-1}(V)$. Since (β, α, X, Y) satisfies the property (P), there exist $A_1, \dots, A_n \in \alpha$ such that $B \subseteq \bigcup_{i=1}^n A_i \subseteq f^{-1}(V)$. Thus $\bigcap_{i=1}^n [A_i, V]$ is an open neighborhood of f in $C_\alpha(X, Y)$. It is easy to see that $\pi_Z(\bigcap_{i=1}^n [A_i, V]) \subseteq [B, V]$. Therefore π_Z is continuous. \square

Proposition 5.3. *Let X, Y be topological spaces, and Z be a subspace of X . Let α be a family of subsets of X and β be a family of subsets of Z . If Z is dense in X and (β, α, X, Y) satisfies the property (P), then $\pi_Z : C_\alpha(X, Y) \rightarrow \pi_Z(C_\alpha(X, Y))$ is a bijective continuous map, i.e. a condensation.*

Proof. Let f and g be distinct elements in $C(X, Y)$. The continuity of the functions f and g and the fact that $\overline{Z} = X$ imply that $f|_Z \neq g|_Z$. Hence $\pi_Z(f) \neq \pi_Z(g)$. This means that π_Z is one-to-one. By Proposition 5.2, π_Z is continuous. \square

Corollary 5.4. *Let X, Y be topological spaces, and Z be a subspace of X . Let α be a family of subsets of X and β be a family of subsets of Z . If α approximates β , then $\pi_Z : C_\alpha(X, Y) \rightarrow C_\beta(Z, Y)$ is continuous.*

Let $B \subseteq Z \subseteq X$ and $V \subseteq Y$. Recall that $[B, V] = \{f \in C(X, Y) : f(B) \subseteq V\}$; let us denote by $[B, V]_Z = \{f \in C(Z, Y) : f(B) \subseteq V\}$. For the converse of Theorem 5.2, we have the following.

Proposition 5.5. *Let X be topological space, Y an equiconnected topological space having a base \mathcal{B} consisting of Ψ -convex sets, and Z be a subspace of X . Let α be a functional refinement family consisting of closed Y -compact subsets of X and β be a family of closed Y -compact subsets of Z . If $\pi_Z : C_\alpha(X, Y) \rightarrow C_\beta(Z, Y)$ is continuous, then (β, α, X, Y) satisfies the property (P).*

Proof. By Corollary 4.3, it suffices to show that $C_\beta(X, Y) \leq C_\alpha(X, Y)$. Let $f \in C_\beta(X, Y)$ and $[B, V]$ be an open neighborhood of it in $C_\beta(X, Y)$, where $B \in \beta$ and V open in Y . Then $\pi_Z(f) = f|_Z \in [B, V]_Z$. The continuity of π_Z leads to the existence of $A_1, \dots, A_n \in \alpha$ and open subsets V_1, \dots, V_n of Y such that

$$f \in \bigcap_{i=1}^n [A_i, V_i] \subseteq [B, V].$$

Hence $C_\beta(X, Y) \leq C_\alpha(X, Y)$, and so (β, α, X, Y) satisfies the property (P). \square

Now, to find out when π_Z is open we first recall the following result obtained by Arhangel'skii [2, Proposition 3] for the topology of pointwise convergence in the case $Y = \mathbb{R}$.

Theorem 5.6. *If Z is a closed subset of X , then π_Z maps the space $C_p(X)$ openly onto the subspace $\pi_Z(C_p(X))$ of $C_p(Z)$.*

In order to study the openness of the restriction map when $C(X, Y)$ and $C(Z, Y)$ are equipped with set open topologies, we will need the following lemmas.

Lemma 5.7. *Let X be a topological space, and Y be an equiconnected space. Let K be a compact subset of X , F be a closed subset of X , and let $f : X \rightarrow Y$ be a continuous function such that $f(K \cap F) \subseteq V$, where V is an open Ψ -convex subset of Y . Then there exists a continuous function $f_1 : X \rightarrow Y$ such that $f_1(K) \subseteq V$ and $f_1|_F = f|_F$.*

Proof. We observe that V is pathwise connected. Let $p : [0, 1] \rightarrow V$ be a path in V . Put $K_1 = K \cap f^{-1}(V^c)$ which is compact. Let $g : X \rightarrow [0, 1]$ be a continuous function such that $g(F) = \{1\}$ and $g(K_1) = \{0\}$. Define the function $f_1 : X \rightarrow Y$ by $f_1(z) = \Psi(p \circ g(z), f(z), g(z))$ for each $z \in X$. It is clear that f_1 is continuous, and one can easily verify that $f_1|_F = f|_F$ and $f_1(K) \subseteq V$. \square

Lemma 5.8. *Let X be a topological space, Y an equiconnected space with equiconnecting function Ψ , Z a subspace of X , α a family of compact subsets of X with $A \cap Z = A \cap \bar{Z}$ for each $A \in \alpha$ and let $g \in C(Z, Y)$ be a function continuously extendable over X . Let $A_1, \dots, A_n \in \alpha$, and V_1, \dots, V_n are Ψ -convex open subsets of Y such that $g(A_i \cap \bar{Z}) \subseteq V_i$ for each $i = 1, \dots, n$. Then there exists a continuous extension $g' : X \rightarrow Y$ of g such that $g'(A_i) \subseteq V_i$ for each $i = 1, \dots, n$.*

Proof. We proceed by recurrence. Let $g \in C(Z, Y)$, $A \in \alpha$ and V be an open Ψ -convex subset of Y with $g(A \cap \bar{Z}) \subseteq V$. By applying Lemma 5.7 with $F = \bar{Z}$ and $K = A$, we obtain a continuous extension g' of g over X with $g'(A) \subseteq V$.

Suppose that the property is true up to n . We show that it remains true for $n + 1$; let $A_1, \dots, A_{n+1} \in \alpha$, and V_1, \dots, V_{n+1} are Ψ -convex open subsets of Y such that $g(A_i \cap \bar{Z}) \subseteq V_i$ for each $i = 1, \dots, n + 1$, and let us show the existence of a continuous extension $g' \in C(X, Y)$ of g such that $g'(A_i) \subseteq V_i$ for each $i = 1, \dots, n$. By our assumption, we have $g((A_i \cap A_{n+1}) \cap \bar{Z}) \subseteq V_i \cap V_{n+1}$, for each $i = 1, \dots, n$. Then the family $\{A_1 \cap A_{n+1}, \dots, A_n \cap A_{n+1}\}$ verifies $(A_i \cap A_{n+1}) \cap Z = (A_i \cap A_{n+1}) \cap \bar{Z}$ for each $i = 1, \dots, n$. Therefore, by the recurrence hypothesis, we find a function $g'_1 \in C(X, Y)$ extending g and such that $g'_1(A_i \cap A_{n+1}) \subseteq V_i \cap V_{n+1}$, for each $i = 1, \dots, n$. We set $X_1 = X \cup (\bigcup_{i=1}^n (A_i \cap A_{n+1}))$. Clearly we have $A_i \cap X_1 = A_i \cap \bar{X}_1$ and $g'_1(A_i \cap \bar{X}_1) \subseteq V_i$ for each $i = 1, \dots, n$. Applying Lemma 5.7 once again for $F = \bar{X}_1$ and $K = A_{n+1}$, we get an extension $g'_2 \in C(X, Y)$ of $g'_1|_{\bar{X}_1}$ such that $g'_2(A_{n+1}) \subseteq V_{n+1}$. We observe that $g'_2(A_i \cap A_{n+1}) \subseteq V_i$, for each i . Again we put $X_2 = X_1 \cup A_{n+1}$. Then, we have, for each $i = 1, \dots, n$, $A_i \cap X_2 = A_i \cap \bar{X}_2$ and $g'_2(A_i \cap \bar{X}_2) \subseteq V_i$. By the recurrence hypothesis, there exists $g'_3 \in C(X, Y)$ which extending $g'_2|_{\bar{X}_2}$ such that $g'_3(A_i) \subseteq V_i$, for each $i = 1, \dots, n$, and we have $g'_3(A_{n+1}) = g'_2(A_{n+1}) \subseteq V_{n+1}$. Whence $g' = g'_3$ is our required function. \square

Lemma 5.9. *Let X be a topological space, Z a subspace of X , α and β are two families of compact subsets of X and Z , respectively. If β approximates $\alpha|_{\bar{Z}}$, then $A \cap Z = A \cap \bar{Z}$ for each $A \in \alpha$.*

Proof. Suppose that there exists $A \in \alpha$ such that $A \cap (\bar{Z} \setminus Z) \neq \emptyset$. Let $x \in A \cap (\bar{Z} \setminus Z)$. Then there is no member of β for which x belongs. Therefore $A \cap \bar{Z}$ does not contained in any finite union of members of β ; this contradicts the fact that β approximates $\alpha|_{\bar{Z}}$. \square

Theorem 5.10. *Let X be a topological space, Z a subspace of X , and Y is an equiconnected space with a base \mathcal{B} consisting of Ψ -convex sets. Let α be an admissible family of compact subsets of X and β be a family of compact subsets of Z . If β approximates $\alpha|_{\bar{Z}}$, then π_Z is an open map from $C_\alpha(X, Y)$ onto the subspace $\pi_Z(C_\alpha(X, Y))$ of $C_\beta(Z, Y)$.*

Proof. Let $\cap_{i=1}^n [A_i, V_i]$, where $A_1, \dots, A_n \in \alpha$ and $V_1, \dots, V_n \in \mathcal{B}$, be a basic open subset of $C_\alpha(X, Y)$ and $f \in \pi_Z(\cap_{i=1}^n [A_i, V_i])$. Let $f' \in C(X, Y)$ be an extension of f over X such that $f' \in \cap_{i=1}^n [A_i, V_i]$. Since β approximates $\alpha|_{\bar{Z}}$, there exists, for each $i = 1, \dots, n$, a finite subfamily β_i of β such that

$$A_i \cap \bar{Z} \subseteq \bigcup \{B : B \in \beta_i\} \subseteq f'^{-1}(V_i).$$

Then

$$f \in \left(\bigcap_{i=1}^n \bigcap_{B \in \beta_i} [B, V_i] \right) \cap \pi_Z(C(X, Y)) = W.$$

We have $W \subseteq \pi_Z(\cap_{i=1}^n [A_i, V_i])$. Indeed, let $g \in W$. Because $g(\bigcup \{B : B \in \beta_i\}) \subseteq V_i$, then $g(A_i \cap \bar{Z}) \subseteq V_i$ for every $i = 1, \dots, n$. Also, from Lemma 5.9, we have $A \cap Z = A \cap \bar{Z}$ for each $A \in \alpha$. Then, by Lemma 5.8, there exists a function $g' \in C(X, Y)$ which agrees with g on Z and belongs to $\cap_{i=1}^n [A_i, V_i]$. We have then $g \in \pi_Z(\cap_{i=1}^n [A_i, V_i])$. Therefore $W \subseteq \pi_Z(\cap_{i=1}^n [A_i, V_i])$, which means that $\pi_Z : C_\alpha(X, Y) \rightarrow \pi_Z(C_\alpha(X, Y))$ is open. \square

Lemma 5.11. *Let X be a topological space, Z a subspace of X , and α, β are two families of compact subsets of X and Z , respectively. Let Y be an equiconnected T_1 -space, with equiconnecting function Ψ . If π_Z is an open map from $C_\alpha(X, Y)$ onto the subspace $\pi_Z(C_\alpha(X, Y))$ of $C_\beta(Z, Y)$, then $A \cap Z = A \cap \bar{Z}$ for each $A \in \alpha$.*

Proof. Suppose that there exists $A \in \alpha$ such that $A \cap (\bar{Z} \setminus Z) \neq \emptyset$. Let $x \in A \cap (\bar{Z} \setminus Z)$. Let $p : [0, 1] \rightarrow Y$ be a path in Y with $p(0) \neq p(1)$ and put $V = Y \setminus \{p(0)\}$. Let $f \in \pi_Z([A, V])$, then we have $f = f'|_Z$ for some $f' \in [A, V]$. Since $\pi_Z : C_\alpha(X, Y) \rightarrow \pi_Z(C_\alpha(X, Y))$ is open, there exist $B_1, \dots, B_n \in \beta$ and V_1, \dots, V_n open subsets in Y such that

$$f \in \left(\bigcap_{i=1}^n [B_i, V_i] \right) \cap \pi_Z(C(X, Y)) \subseteq \pi_Z([A, V]).$$

Since $x \in A \cap (\bar{Z} \setminus Z)$, we have $x \notin \cup_{i=1}^n B_i$. Complete regularity of X gives us a continuous function $h : X \rightarrow [0, 1]$ such that $h(x) = 0$ and $h(\cup_{i=1}^n B_i) = \{1\}$. Consider the function $h_1 : X \rightarrow Y$ defined by $h_1(z) = \Psi(p \circ h(z), f'(z), h(z))$ for each $z \in X$. It is clear that h_1 is continuous, $h_1(\cup_{i=1}^n B_i) \subseteq V$ and $h_1(x) = p(0) \notin V$. So $h_1|_Z \in (\bigcap_{i=1}^n [B_i, V_i]) \cap \pi_Z(C(X, Y))$ and h_1 does not belong to $[A, V]$. Assume that $h_1|_Z$ admits another extension $h_2 \in C(X, Y)$. By continuity we have $h_1|_{\bar{Z}} = h_2|_{\bar{Z}}$. Thus $h_1(x) = h_2(x) \notin V$, and so $h_2 \notin [A, V]$. This leads that no continuous extension of $h_1|_Z$ over X belongs to $[A, V]$, which contradicts the fact that $(\bigcap_{i=1}^n [B_i, V_i]) \cap \pi_Z(C(X, Y)) \subseteq \pi_Z([A, V])$. Hence, we have $A \cap Z = A \cap \bar{Z}$ for every $A \in \alpha$. \square

Theorem 5.12. *Let X be a topological space, Z a subspace of X , and Y is an equiconnected T_1 -space with a base \mathcal{B} consisting of Ψ -convex sets. Let α be a family of compact subsets of X and β an admissible family of compact subsets of Z . If π_Z is an open map from $C_\alpha(X, Y)$ onto the subspace $\pi_Z(C_\alpha(X, Y))$ of $C_\beta(Z, Y)$, then β approximates $\alpha|_{\bar{Z}}$.*

Proof. Notice first that, from Lemma 5.11, we have $A \cap \bar{Z} = A \cap Z$ for every $A \in \alpha$. Furthermore, $C_\beta(Z, Y) = C_\beta^{\mathcal{B}}(Z, Y)$ because β is a functional refinement family. Let $A \in \alpha$, G an open subset of Z with $A \cap \bar{Z} = A \cap Z \subseteq G$. Let G_1 be an open subset in X with $G_1 \cap Z = G$. Now the subset $G_2 = G_1 \cup (X \setminus \bar{Z})$, which is open in X , contains A and verifies $G_2 \cap Z = G_1 \cap Z = G$. Let $f : X \rightarrow [0, 1]$ be a continuous function such that $f(A) = \{1\}$ and $f(X \setminus G_2) = \{0\}$. Let $V = Y \setminus \{p(0)\}$, where p is a path in Y , and put $g = p \circ f$. Then $g^{-1}(V) \subseteq G_2$. Thus $g|_Z^{-1}(V) \subseteq G$. Consider in $C_\alpha(X, Y)$ the subbasic open subset $[A, V]$. We have then $g \in [A, V]$. Thus $\pi_Z(g) = g|_Z$ belongs to $\pi_Z([A, V])$ which is open in $\pi_Z(C_\alpha(X, Y))$, by our assumption. Therefore, there exist $B_1, \dots, B_n \in \beta$ and open sets $V_1, \dots, V_n \in \mathcal{B}$ such that

$$g|_Z \in (\cap_{i=1}^n [B_i, V_i]) \cap \pi_Z(C_\alpha(X, Y)) \subseteq \pi_Z([A, V]).$$

By the same reasoning as in the proof of Lemma 5.11, we obtain that $A \cap Z = A \cap \bar{Z} \subseteq \cup_{i=1}^n B_i$. To continue our proof we will introduce the following notation. By B^1 we denote a subset $B \subseteq X$, and by B^0 its complementary in X , i.e., B^c . Let $\mathcal{J} = \{1, \dots, n\}$ and define the following set

$$\Delta = \left\{ (\delta_1, \delta_2, \dots, \delta_n) \in \{0, 1\}^n \setminus \{(0, \dots, 0)\} : \left(\bigcap_{i \in \mathcal{J}} B_i^{\delta_i} \right) \cap (A \cap \bar{Z}) \neq \emptyset \right\}.$$

We have then

$$A \cap \bar{Z} = A \cap Z \subseteq \bigcup \left\{ \bigcap_{i=1}^n B_i^{\delta_i} / (\delta_1, \dots, \delta_n) \in \Delta \right\}.$$

Fixing an element $(\delta_1, \dots, \delta_n)$ in Δ and let us show that $\bigcap_{\delta_i=1} V_i \subseteq V$. Assume the contrary. Let $y_0 \in \bigcap_{\delta_i=1} V_i \setminus V$. Let $x_0 \in \left(\bigcap_{i \in \mathcal{J}} B_i^{\delta_i} \right) \cap (A \cap \bar{Z})$, then $x_0 \notin (\cup_{\delta_i=0} B_i)$ and $g(x_0) \in \bigcap_{\delta_i=1} V_i$. By continuity of g and the fact that $x_0 \notin (\cup_{\delta_i=0} B_i)$, we can take an open neighborhood U of x_0 such that $g(U) \subseteq \bigcap_{\delta_i=1} V_i$ and $U \cap (\cup_{\delta_i=0} B_i) = \emptyset$. Consider a continuous function $\varphi : X \rightarrow [0, 1]$ such that $\varphi(x_0) = 0$ and $\varphi(U^c) = \{1\}$, and $0 \leq \varphi(x) \leq 1$ for all $x \in X$. Then the function $h : X \rightarrow Y$ defined by $h(x) = \Psi(y_0, g(x), \varphi(x))$, for all $x \in X$, is continuous and $h|_Z$ does not belong to $\pi_Z([A, V])$ because $h(x_0) = y_0$. But $h|_Z \in (\cap_{i=1}^n [B_i, V_i]) \cap \pi_Z(C_\alpha(X, Y))$. In fact, if $x \in B_i \cap U$ then $h(x) = \Psi(y_0, g(x), \varphi(x)) \in V_i$, because V_i is Ψ -convex subset, for each $1 \leq i \leq n$. If $x \in B_i \setminus U$, then $h(x) = g(x) \in V_i$ for each i with $1 \leq i \leq n$. This gives us a contradiction, so we have $\bigcap_{\delta_i=1} V_i \subseteq V$. Moreover, since $\bigcap_{i \in \mathcal{J}} B_i^{\delta_i} \subseteq \bigcap_{\delta_i=1} B_i$ and $A \subseteq \bigcup \left\{ \bigcap_{i \in \mathcal{J}} B_i^{\delta_i} / (\delta_i) \in \Delta \right\}$ we obtain that

$$A \cap Z \subseteq \bigcup \left\{ \bigcap_{\delta_i=1} B_i / (\delta_i) \in \Delta \right\} \subseteq f|_Z^{-1}(V).$$

Moreover, the admissibility of β gives us, by [4, Lemma 1], that for each $(\delta_i) \in \Delta$, there exist $\beta_{(\delta_i)}$ finite subfamily of β such that

$$\bigcap_{\delta_i=1} B_i \subseteq \bigcup \left\{ B : B \in \beta_{(\delta_i)} \right\} \subseteq f|_Z^{-1}(V).$$

Hence

$$A \cap \bar{Z} = A \cap Z \subseteq \bigcup_{(\delta_i) \in \Delta} \left(\bigcup_{B \in \beta_{(\delta_i)}} B \right) \subseteq f|_Z^{-1}(V) \subseteq G.$$

Thus, we have β approximates $\alpha|_{\overline{Z}}$. \square

Corollary 5.13. *Let X be a topological space, Z a subspace of X , and Y is an equiconnected T_1 -space with a base \mathcal{B} consisting of Ψ -convex sets. Let α be an admissible family of compact subsets of X and β an admissible family of compact subsets of Z . Then π_Z is an open map from $C_\alpha(X, Y)$ onto the subspace $\pi_Z(C_\alpha(X, Y))$ of $C_\beta(Z, Y)$ if and only if β approximates $\alpha|_{\overline{Z}}$.*

Acknowledgment. The last section of this work was done when the second author was visiting the Department of Mathematics of Gazi University (Turkey). He want to thank Professor Çetin Vural for his warm hospitality and useful discussions.

References

- [1] R. Arens and J. Dugundji, *Topologies for function spaces*. Pacific J. Math. **1**, 5-31, 1951.
- [2] A.V. Arhangel'skii, *Topological Function Spaces*, Kluwer Academic Publishers, 1992.
- [3] A.V. Arhangel'skii, *General topology III*, Encyclopaedia of Mathematical Sciences **51**, Springer-Verlag Berlin, 1995.
- [4] A. Bouchair and S. Kelaiaia, *Comparison of some set open topologies on $C(X, Y)$* , Top. Appl. **178**, 352-359, 2014.
- [5] R. Engelking, *General Topology*, Polish Scientific Publishing, 1977.
- [6] R. Fox, *On fibre spaces II*, Bull. Amer. Math. Soc. **49**, 733-735, 1943.
- [7] S. Kelaiaia, *Propriétés de certaines topologies set-open sur $C(X)$* , Thèse de Doctorat de l'Université de Rouen, 1995.
- [8] R.A. McCoy and I. Ntantu, *Countability properties of function spaces with set-open topology*. Top. proc. **10**, 329-345, 1985.
- [9] R.A. McCoy and I. Ntantu, *Topological properties of spaces of continuous functions*. Lecture Note in Math. **1315**, Springer Verlag Germany, 1988.
- [10] S.E. Nokhrin and A.V. Osipov, *On the coincidence of set-open and uniform topologies*. Proc. steklov Inst. Math. Suppl. **3**, 184-191, 2009. Top. Appl. **150**, 255-265, 2005.
- [11] A.V. Osipov, *Uniformity of uniform convergence on the family of sets*, Top. Proc. **50**, 79-86, 2017.
- [12] S. Önal and C. Vural, *Some cardinal invariants on the space $C_\alpha(X, Y)$* , Top. Appl. **150**, 255-265, 2005.



Conformal slant submersions

Mehmet Akif Akyol^{*1}, Bayram Şahin²

¹*Bingöl University, Faculty of Arts and Science, Department of Mathematics, 12000, Bingöl, Turkey*

²*Ege University, Department of Mathematics, 35100, Izmir, Turkey*

Abstract

As a generalization of conformal holomorphic submersions and conformal anti-invariant submersions, we introduce a new conformal submersion from almost Hermitian manifolds onto Riemannian manifolds, namely conformal slant submersions. We give examples and find necessary and sufficient conditions for such maps to be harmonic morphism. We also investigate the geometry of foliations which are arisen from the definition of a conformal submersion and obtain a decomposition theorem on the total space of a conformal slant submersion. Moreover, we find necessary and sufficient conditions of a conformal slant submersion to be totally geodesic.

Mathematics Subject Classification (2010). Primary 53C15; Secondary 53C43

Keywords. Riemannian submersion, slant-submersion, conformal submersion, conformal slant-submersion

1. Introduction

A submanifold of a complex manifold is a complex (invariant) submanifold if the tangent space of the submanifold at each point is invariant with respect to the almost complex structure of the Kähler manifold. Besides complex submanifolds of a complex manifold, there is another important class of submanifolds called totally real submanifolds. A totally real submanifold of a complex manifold is a submanifold of such that the almost complex structure of ambient manifold carries the tangent space of the submanifold at each point into its normal space. Many authors have studied totally real submanifolds in various ambient manifolds and many interesting results were obtained, see ([45], page: 199) for a survey on all these results. As a generalization of holomorphic and totally real submanifolds, slant submanifolds were introduced by Chen in [13]. We recall that the submanifold M is called slant [14] if for any $p \in M$ and any $X \in T_pM$, the angle between JX and T_pM is a constant $\theta(X) \in [0, \frac{\pi}{2}]$, i.e, it does not depend on the choice of $p \in M$ and $X \in T_pM$. It follows that invariant and totally real immersions are slant immersions with slant angle $\theta = 0$ and $\theta = \frac{\pi}{2}$ respectively.

On the other hand, Riemannian submersions between Riemannian manifolds were studied by O’Neill [37] and Gray [25]. Since then Riemannian submersions have been an effective tool to obtain new manifolds and compare certain manifolds within differential

*Corresponding Author.

Email addresses: mehmetakifakyol@bingol.edu.tr (M.A. Akyol), bayram.sahin@ege.edu.tr (B. Şahin)

Received: 20.06.2017; Accepted: 03.08.2017

geometry, see [8], [12] and [21]. It is also known that Riemannian submersions have many applications in different areas such as Kaluza-Klein theory [22], [10], statistical machine learning processes [46], medical imaging [36], statistical analysis on manifolds [9] and the theory of robotics [3]. As analogue of holomorphic submanifolds, holomorphic submersions were introduced by Watson [44] in seventies by using the notion of almost complex map. This notion has been extended to other manifolds, see [21] for holomorphic submersions and their extensions to other manifolds. Although holomorphic submersions have been studied widely, however this research area is still an active research area, see a recent paper [43]. The main property of such maps is that the horizontal distribution and the vertical distribution of holomorphic submersions are invariant with respect to the almost complex map of the total manifold. Thus holomorphic submersions include only those submersions whose vertical distribution is invariant under the almost complex structure of the total manifold. Therefore, the second author of the present paper considered a new submersion defined on an almost Hermitian manifold such that the vertical distribution is anti-invariant with respect to almost complex structure [41]. He showed that such submersions have rich geometric properties and they are useful for investigating the geometry of the total space. This new class of submersions called anti-invariant submersions can be seen as an analogue of totally real submanifolds in the submersion theory. As a generalization of anti-invariant submersions, slant submersions were defined in [42] and it is shown that such maps are useful for obtaining new conditions for harmonicity, see also [4, 5, 7, 18–20, 24, 28–32, 34, 35, 38] and [40] for new submersions in other total spaces.

As a generalization of Riemannian submersions, horizontally conformal submersions are defined as follows [6]: Suppose that (M, g_M) and (B, g_B) are Riemannian manifolds and $F : M \rightarrow B$ is a smooth submersion, then F is called a horizontally conformal submersion, if there is a positive function λ such that

$$\lambda^2 g_M(X, Y) = g_B(F_*X, F_*Y)$$

for every $X, Y \in \Gamma((\ker F_*)^\perp)$. It is obvious that every Riemannian submersion is a particular horizontally conformal submersion with $\lambda = 1$. One can see that Riemannian submersions are very special maps comparing with conformal submersions. We note that horizontally conformal submersions are special horizontally conformal maps which were introduced independently by Fuglede [23] and Ishihara [33]. We also note that a horizontally conformal submersion $F : M \rightarrow B$ is said to be horizontally homothetic if the gradient of its dilation λ is vertical, i.e.,

$$\mathcal{H}(\text{grad}\lambda) = 0 \tag{1.1}$$

at $p \in M$, where \mathcal{H} is the projection on the horizontal space $(\ker F_*)^\perp$. Although conformal maps do not preserve distance between points contrary to isometries, they preserve angles between vector fields. This property enables one to transfer certain properties of a manifold to another manifold by deforming such properties.

As a generalization of holomorphic submersions, conformal holomorphic submersions were studied by Gudmundsson and Wood [27]. They obtained necessary and sufficient conditions for conformal holomorphic submersions to be a harmonic morphism, see also [15–17] for the harmonicity of conformal holomorphic submersions.

Recently, we introduced conformal anti-invariant submersions [2] from almost Hermitian manifolds onto Riemannian manifolds, as a generalization of anti-invariant submersions, and investigated the geometry of such submersions. (see also: [1]) We showed that the

geometry of such submersions are different from their counterparts anti-invariant submersions and semi-invariant submersions. In this paper, we study conformal slant submersions as a generalization of both conformal holomorphic submersions and conformal anti-invariant submersions and investigate the geometry of the total space and the base space for the existence of such submersions.

The paper is organized as follows. In the second section, we present the basic information needed for this paper. In section 3, we give definition of conformal slant submersions from almost Hermitian manifolds onto Riemannian manifolds, provide examples and give a sufficient condition for conformal slant submersions to be harmonic. We also investigate the geometry of leaves of $(\ker F_*)^\perp$ and $(\ker F_*)$. Moreover we obtain a decomposition theorem on the total space of a conformal slant submersion. Finally, we give necessary and sufficient conditions for a conformal slant submersion to be totally geodesic.

2. Preliminaries

In this section, we define almost Hermitian manifolds, recall the notion of (horizontally) conformal submersions between Riemannian manifolds and give a brief review of basic facts of (horizontally) conformal submersions.

Let (M, g_M) be an almost Hermitian manifold. This means [45] that M admits a tensor field J of type (1,1) on M such that, $\forall X, Y \in \Gamma(TM)$, we have

$$J^2 = -I, \quad g_M(X, Y) = g_M(JX, JY). \quad (2.1)$$

An almost Hermitian manifold M is called Kähler manifold if

$$(\nabla_X J)Y = 0, \quad \forall X, Y \in \Gamma(TM), \quad (2.2)$$

where ∇ is the Levi-Civita connection on M . Conformal submersions belong to a wide class of conformal maps that we are going to recall their definition, but we will not study such maps in this paper.

Definition 2.1. ([6]). Let $\varphi : (M^m, g) \rightarrow (N^n, h)$ be a smooth map between Riemannian manifolds, and let $x \in M$. Then φ is called horizontally weakly conformal or semi conformal at x if either

- (i) $d\varphi_x = 0$, or
- (ii) $d\varphi_x$ maps the horizontal space $\mathcal{H}_x = (\ker(d\varphi_x))^\perp$ conformally onto $T_{\varphi(x)}N$, i.e., $d\varphi_x$ is surjective and there exists a number $\Lambda(x) \neq 0$ such that

$$h(d\varphi_x X, d\varphi_x Y) = \Lambda(x)g(X, Y) \quad (X, Y \in \mathcal{H}_x). \quad (2.3)$$

We shall call a point x of type (i) in Definition 2.1 *critical point*. Also we shall call a point of type (ii) a *regular point*. At a critical point, $d\varphi_x$ has rank 0; at a regular point, $d\varphi_x$ has rank n and φ is submersion. The number $\Lambda(x)$ is called the *square dilation* (of φ at x); it is necessarily non-negative; its square root $\lambda(x) = \sqrt{\Lambda(x)}$ is called the *dilation* (of φ at x). The map φ is called *horizontally weakly conformal* or *semi conformal* (on M) if it is horizontally weakly conformal at every point of M . It is clear that if φ has no critical points, then we call it a (*horizontally*) conformal submersion.

Next, we recall the following definition from [26]. Let $\pi : M \rightarrow N$ be a submersion. A vector field E on M is said to be projectable if there exists a vector field \check{E} on N , such that $d\pi(E_x) = \check{E}_{\pi(x)}$ for all $x \in M$. In this case E and \check{E} are called π -related. A horizontal vector field Y on (M, g) is called basic, if it is projectable. It is well known fact, that if \check{Z} is a vector field on N , then there exists a unique basic vector field Z on M , such that Z and \check{Z} are π -related. The vector field Z is called the horizontal lift of \check{Z} .

The fundamental tensors of a submersion were introduced in [37]. They play a similar role to that of the second fundamental form of an immersion. More precisely, O'Neill's tensors T and A defined for vector fields E, F on M by

$$A_E F = \mathcal{V}\nabla_{\mathcal{H}E}\mathcal{H}F + \mathcal{H}\nabla_{\mathcal{H}E}\mathcal{V}F, \quad T_E F = \mathcal{H}\nabla_{\mathcal{V}E}\mathcal{V}F + \mathcal{V}\nabla_{\mathcal{V}E}\mathcal{H}F \quad (2.4)$$

where \mathcal{V} and \mathcal{H} are the vertical and horizontal projections (see [21]). On the other hand, from (2.4), we have

$$\nabla_V W = T_V W + \hat{\nabla}_V W \quad (2.5)$$

$$\nabla_V X = \mathcal{H}\nabla_V X + T_V X \quad (2.6)$$

$$\nabla_X Y = \mathcal{H}\nabla_X Y + A_X Y \quad (2.7)$$

for $X, Y \in \Gamma((\ker F_*)^\perp)$ and $V, W \in \Gamma(\ker F_*)$, where $\hat{\nabla}_V W = \mathcal{V}\nabla_V W$. If X is basic, then $\mathcal{H}\nabla_V X = A_X V$. It is easily seen that for $x \in M$, $X \in \mathcal{H}_x$ and \mathcal{V}_x the linear operators $T_V, A_X : T_x M \rightarrow T_x M$ are skew-symmetric. We see that the restriction of T to the vertical distribution $T|_{\mathcal{V} \times \mathcal{V}}$ is exactly the second fundamental form of the fibres of π . Since T_V is skew-symmetric we get: π has totally geodesic fibres if and only if $T \equiv 0$. For the special case when π is horizontally conformal we have the following:

Proposition 2.2. ([26]). *Let $\pi : (M^m, g) \rightarrow (N^n, h)$ be a horizontally conformal submersion with dilation λ and X, Y be horizontal vectors, then*

$$A_X Y = \frac{1}{2} \{ \mathcal{V}[X, Y] - \lambda^2 g(X, Y) \text{grad}_{\mathcal{V}}(\frac{1}{\lambda^2}) \}. \quad (2.8)$$

We now recall the notion of harmonic maps between Riemannian manifolds. Let (M, g_M) and (N, g_N) be Riemannian manifolds and suppose that $\varphi : M \rightarrow N$ is a smooth map between them. Then the differential φ_* of φ can be viewed a section of the bundle $\text{Hom}(TM, \varphi^{-1}TN) \rightarrow M$, where $\varphi^{-1}TN$ is the pullback bundle which has fibres $(\varphi^{-1}TN)_p = T_{\varphi(p)}N$, $p \in M$. $\text{Hom}(TM, \varphi^{-1}TN)$ has a connection ∇ induced from the Levi-Civita connection ∇^M and the pullback connection. Then the second fundamental form of φ is given by

$$(\nabla\varphi_*)(X, Y) = \nabla_X^\varphi \varphi_*(Y) - \varphi_*(\nabla_X^M Y) \quad (2.9)$$

for $X, Y \in \Gamma(TM)$, where ∇^φ is the pullback connection. It is known that the second fundamental form is symmetric. A smooth map $\varphi : (M, g_M) \rightarrow (N, g_N)$ is said to be harmonic if $\text{trace}(\nabla\varphi_*) = 0$. On the other hand, the tension field of φ is the section $\tau(\varphi) \in \Gamma(\varphi^{-1}TN)$ defined by

$$\tau(\varphi) = \text{div}\varphi_* = \sum_{i=1}^m (\nabla\varphi_*)(e_i, e_i), \quad (2.10)$$

where $\{e_1, \dots, e_m\}$ is an orthonormal frame on M . Then it follows that φ is harmonic if and only if $\tau(\varphi) = 0$, for details, see [6]. Now, we recall the following lemma from [6].

Lemma 2.3. *Suppose that $\varphi : M \rightarrow N$ is a horizontally conformal submersion. Then, for any horizontal vector fields X, Y and vertical fields V, W we have*

- (i) $\nabla d\varphi(X, Y) = X(\ln\lambda)d\varphi(Y) + Y(\ln\lambda)d\varphi(X) - g(X, Y)d\varphi(\text{grad}\ln\lambda)$;
- (ii) $\nabla d\varphi(V, W) = -d\varphi(A_V^\mathcal{V}W)$;
- (iii) $\nabla d\varphi(X, V) = -d\varphi(\nabla_X^M V) = d\varphi((A_X^{\mathcal{H}})^*V)$.

Here $(A_X^{\mathcal{H}})^*$ is the adjoint of $A_X^{\mathcal{H}}$ characterized by

$$\langle (A_X^{\mathcal{H}})^*E, F \rangle = \langle E, A_X^{\mathcal{H}}F \rangle \quad (E, F \in \Gamma(TM)).$$

Let g_B be a Riemannian metric tensor on the manifold $B = B_1 \times B_2$ and assume that the canonical foliations D_{B_1} and D_{B_2} intersect perpendicularly everywhere. Then g_B is the metric tensor of a usual product of Riemannian manifolds if and only if D_{B_1} and D_{B_2} are totally geodesic foliations [39].

3. Conformal Slant submersions

In this section, we define conformal slant submersions from an almost Hermitian manifold onto a Riemannian manifold, investigate the effect of the existence of conformal slant submersions on the source manifold and the target manifold. But we first present the following notion.

Definition 3.1. Let F be a horizontally conformal submersion from an almost Hermitian manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . If for any non-zero vector $X \in \Gamma(\ker F_{*p}); p \in M_1$, the angle $\theta(X)$ between JX and the space $(\ker F_{*p})$ is a constant, i.e. it is independent of the choice of the point $p \in M_1$ and choice of the tangent vector X in $(\ker F_{*p})$, then we say that F is a conformal slant submersion. In this case, the angle θ is called the slant angle of the conformal slant submersion.

We note that it is known that the distribution $\ker F_*$ is integrable. In fact, its leaves are $F^{-1}(q)$, $q \in M_2$, i.e., fibers. Thus it follows from above definition that the fibres of a conformal slant submersion are slant submanifolds of M_1 , for slant submanifolds, see [13]. We now give some examples of conformal slant submersions.

Example 3.2. Every Hermitian submersion from an almost Hermitian manifold onto an almost Hermitian manifold is a conformal slant submersion with $\lambda = 1$ and $\theta = 0$.

Example 3.3. Every conformal anti-invariant submersion from an almost Hermitian manifold to a Riemannian manifold is a conformal slant submersion with $\lambda = 1$ and $\theta = \frac{\pi}{2}$.

Example 3.4. Every slant submersion from an almost Hermitian manifold onto Riemannian manifold is a conformal slant submersion with $\lambda = 1$.

A conformal slant submersion is said to be proper if it is neither Hermitian nor conformal anti-invariant submersion. We now present two examples of a proper conformal slant submersion. We denote by J_α the compatible almost complex structure on R^4 defined by

$$J_\alpha(a, b, c, d) = (\cos \alpha)(-c, -d, a, b) + (\sin \alpha)(-b, a, d, -c), \quad 0 < \alpha \leq \frac{\pi}{2}$$

Example 3.5. Consider the following submersion given by

$$F : \begin{array}{ccc} R^4 & \longrightarrow & R^2 \\ (x_1, x_2, x_3, x_4) & & (e^{x_1} \sin x_2, e^{x_1} \cos x_2), \end{array}$$

where $x_2 \in \mathbb{R} - \{k\frac{\pi}{2}, k\pi\}$, $k \in \mathbb{Z}$. Then it follows that

$$\ker F_* = \text{span}\{V_1 = \partial x_3, V_2 = \partial x_4\}$$

and

$$\begin{aligned} (\ker F_*)^\perp &= \text{span}\{X_1 = e^{x_1} \sin x_2 \partial x_1 + e^{x_1} \cos x_2 \partial x_2, \\ &X_2 = e^{x_1} \cos x_2 \partial x_1 - e^{x_1} \sin x_2 \partial x_2\}. \end{aligned}$$

Then by direct computations for any $0 < \theta \leq \frac{\pi}{2}$, F is a slant submersion with slant angle θ . On the other hand,

$$F_* X_1 = (e^{x_1})^2 \partial y_1, \quad F_* X_2 = (e^{x_1})^2 \partial y_2.$$

Hence, we have

$$g_2(F_* X_1, F_* X_1) = (e^{x_1})^2 g_1(X_1, X_1), \quad g_2(F_* X_2, F_* X_2) = (e^{x_1})^2 g_1(X_2, X_2),$$

where g_1 and g_2 denote the standard metrics (inner products) of R^4 and R^2 . Thus F is a conformal slant submersion with $\lambda = e^{x_1}$.

Example 3.6. Let F be a submersion defined by

$$F : \begin{array}{ccc} R^4 & \longrightarrow & R^2 \\ (x_1, x_2, x_3, x_4) & & (\cosh x_1 \sin x_3, \sinh x_1 \cos x_3), \end{array}$$

where $x_3 \in \mathbb{R} - \{k\frac{\pi}{2}, k\pi\}$, $k \in \mathbb{Z}$. Then it follows that

$$\ker F_* = \text{span}\{V_1 = \partial x_2, V_2 = \partial x_4\}$$

and

$$(\ker F_*)^\perp = \text{span}\{X_1 = \sinh x_1 \sin x_3 \partial x_1 + \cosh x_1 \cos x_3 \partial x_2, \\ X_2 = \cosh x_1 \cos x_3 \partial x_1 - \sinh x_1 \sin x_3 \partial x_2\}.$$

Then by direct computations for any $0 < \theta \leq \frac{\pi}{2}$, F is a slant submersion with slant angle θ . On the other hand,

$$F_* X_1 = (\sinh^2 x_1 \sin^2 x_3 + \cosh^2 x_1 \cos^2 x_3) \partial y_1$$

and

$$F_* X_2 = (\sinh^2 x_1 \sin^2 x_3 + \cosh^2 x_1 \cos^2 x_3) \partial y_2.$$

Hence, we have

$$g_2(F_* X_1, F_* X_1) = (\sinh^2 x_1 \sin^2 x_3 + \cosh^2 x_1 \cos^2 x_3) g_1(X_1, X_1)$$

and

$$g_2(F_* X_2, F_* X_2) = (\sinh^2 x_1 \sin^2 x_3 + \cosh^2 x_1 \cos^2 x_3) g_1(X_2, X_2),$$

where g_1 and g_2 denote the standard metrics (inner products) of R^4 and R^2 . Thus F is a conformal slant submersion with $\lambda^2 = \sinh^2 x_1 \sin^2 x_3 + \cosh^2 x_1 \cos^2 x_3$.

Let F be a conformal slant submersion from an almost Hermitian manifold (M_1, g_1, J) onto a Riemannian manifold (M_2, g_2) . Then for $U \in \Gamma(\ker F_*)$, we write

$$JU = \phi U + \omega U \tag{3.1}$$

where ϕU and ωU are vertical and horizontal parts of JU . Also for $X \in \Gamma((\ker F_*)^\perp)$, we have

$$JX = \mathcal{B}X + \mathcal{C}X, \tag{3.2}$$

where $\mathcal{B}X$ and $\mathcal{C}X$ are vertical and horizontal components. Using (2.5), (2.6), (3.1) and (3.2) we obtain

$$(\nabla_U \omega)V = \mathcal{C}T_U V - T_U \phi V \tag{3.3}$$

$$(\nabla_U \phi)V = \mathcal{B}T_U V - T_U \omega V, \tag{3.4}$$

where ∇ is the Levi-Civita connection on M_1 and

$$(\nabla_U \omega)V = \mathcal{H} \nabla_U \omega V - \omega \hat{\nabla}_U V$$

$$(\nabla_U \phi)V = \hat{\nabla}_U \phi V - \phi \hat{\nabla}_U V$$

for $U, V \in \Gamma(\ker F_*)$. Let F be a proper conformal slant submersion from an almost Hermitian manifold (M_1, g_1, J) onto a Riemannian manifold (M_2, g_2) , then we say that ω is parallel with respect to the Levi-Civita connection ∇ on $(\ker F_*)$ if its covariant derivative with respect to ∇ vanishes, i.e., we have

$$(\nabla_U \omega)V = \nabla_U \omega V - \phi \hat{\nabla}_U V$$

for $U, V \in \Gamma(\ker F_*)$. The proof of the following result is exactly same with slant immersions (see [11] and [13]), therefore we omit its proof.

Theorem 3.7. *Let F be a conformal slant submersion from an almost Hermitian manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then F is a proper conformal slant submersion if and only if there exists a constant $\lambda_1 \in [-1, 0]$ such that*

$$\phi^2 U = \lambda_1 U$$

for $U \in \Gamma(\ker F_*)$. If F is a proper conformal slant submersion, then $\lambda_1 = -\cos^2 \theta$.

By using above theorem, it is easy to see the following lemma.

Lemma 3.8. *Let F be a proper conformal slant submersion from an almost Hermitian manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) with slant angle θ . Then, for any $U, V \in \Gamma(\ker F_*)$, we have*

$$g_1(\phi U, \phi V) = \cos^2 \theta g_1(U, V), \quad (3.5)$$

and

$$g_1(\omega U, \omega V) = \sin^2 \theta g_1(U, V). \quad (3.6)$$

We now denote complementary distribution of $\omega(\ker F_*)$ in $(\ker F_*)^\perp$ by μ . The proof of the following result is exactly same with slant submersion (see [42]), therefore we omit its proof.

Proposition 3.9. *Let F be a proper conformal slant submersion from an almost Hermitian manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then μ is invariant with respect to J_1 .*

Corollary 3.10. *Let F be a proper conformal slant submersion from an almost Hermitian manifold (M_1^m, g_1, J_1) onto a Riemannian manifold (M_2^n, g_2) . Let*

$$\{e_1, \dots, e_{m-n}\}$$

be a local orthonormal basis of $(\ker F_*)$, then $\{\csc \theta \omega e_1, \dots, \csc \theta \omega e_{m-n}\}$ is a local orthonormal basis of $\omega(\ker F_*)$.

Proof. It will be enough to show that $g_1(\csc \theta \omega e_i, \csc \theta \omega e_j) = \delta_{ij}$, for any $i, j \in \{1, \dots, \frac{m-n}{2}\}$. By using (3.6), we have

$$g_1(\csc \theta \omega e_i, \csc \theta \omega e_j) = \csc^2 \theta \sin^2 \theta g_1(e_i, e_j) = \delta_{ij},$$

which proves the assertion. \square

We note that above Proposition 3.9 tells that the distributions μ and $(\ker F_*) \oplus \omega(\ker F_*)$ are even dimensional. In fact it implies that the distribution $(\ker F_*)$ is even dimensional. More precisely, we have the following result whose proof is similar to the above corollary.

Lemma 3.11. *Let F be a proper conformal slant submersion from an almost Hermitian manifold (M_1^m, g_1, J_1) onto a Riemannian manifold (M_2^n, g_2) . If $e_1, e_2, \dots, e_{\frac{m-n}{2}}$ are orthogonal unit vector fields in $(\ker F_*)$, then*

$$\{e_1, \sec \theta \phi e_1, e_2, \sec \theta \phi e_2, \dots, e_{\frac{m-n}{2}}, \sec \theta \phi e_{\frac{m-n}{2}}\}$$

is a local orthonormal basis of $(\ker F_*)$.

Let F be a proper conformal slant submersion from an almost Hermitian manifold (M_1^{2n}, g_1, J_1) onto a Riemannian manifold (M_2^n, g_2) . As in the case of slant immersions, we call such an orthonormal frame

$$\{e_1, \sec \theta \phi e_1, e_2, \sec \theta \phi e_2, \dots, e_n, \sec \theta \phi e_n, \csc \theta \omega e_1, \csc \theta \omega e_2, \dots, \csc \theta \omega e_n\}$$

an adapted slant frame for conformal slant submersions. In the sequel, we show that the conformal slant submersion puts some restrictions on the dimensions of the distributions and the base manifold.

Proposition 3.12. *Let F be a proper conformal slant submersion from an almost Hermitian manifold (M_1^m, g_1, J_1) onto a Riemannian manifold (M_2^n, g_2) . Then $\dim(\mu) = 2n - m$. If $\mu = 0$, then $n = \frac{m}{2}$.*

Proof. First note that $\dim(\ker F_*) = m - n$. Thus using Corollary 3.10, we have $\dim((\ker F_*) \oplus \omega(\ker F_*)) = 2(m - n)$. Since M_1 is m - dimensional, we get $\dim(\mu) = 2n - m$. Second assertion is clear. \square

We now check the harmonicity of conformal slant submersions. But we first give a preparatory lemma.

Lemma 3.13. *Let F be a proper conformal slant submersion from a Kähler manifold onto a Riemannian manifold. If ω is parallel with respect to ∇ on $(\ker F_*)$, then we have*

$$T_{\phi U} \phi U = -\cos^2 \theta T_U U \tag{3.7}$$

for $U \in \Gamma(\ker F_*)$.

Proof. If ω is parallel, then from (3.3) we have $\mathcal{C}T_U V = T_U \phi V$ for $U, V \in \Gamma(\ker F_*)$. Interchanging the role of U and V , we get $\mathcal{C}T_V U = T_V \phi U$. Thus we have

$$\mathcal{C}T_U V - \mathcal{C}T_V U = T_U \phi V - T_V \phi U.$$

Since T is symmetric, we derive

$$T_U \phi V = T_V \phi U. \tag{3.8}$$

Then substituting V by ϕU we get $T_U \phi^2 U = T_{\phi U} \phi U$. Finally using Theorem 3.7 we obtain (3.7). \square

Theorem 3.14. *Let $F : (M_1^{2(m+r)}, g_1, J_1) \longrightarrow (M_2^{m+2r}, g_2)$ be a conformal slant submersion, where (M_1, g_1, J_1) is a Kähler manifold and (M_2, g_2) is a Riemannian manifold. Then the tension field τ of F is*

$$\tau(F) = -\frac{1}{m} F_* \left(T_{e_i} e_i + \sec^2 \theta T_{\phi e_i} \phi e_i \right) + \left(\frac{2}{\lambda^2} - (m + 2r) \right) F_*(\text{grad} \ln \lambda). \tag{3.9}$$

Proof. Let $\{e_1, \dots, e_m, \sec \theta \phi e_1, \dots, \sec \theta \phi e_m, \csc \theta \omega e_1, \dots, \csc \theta \omega e_m, \mu_1, \dots, \mu_r, J_1 \mu_1, \dots, J_1 \mu_r\}$ be orthonormal basis of $\Gamma(TM_1)$ such that $\{e_1, \dots, e_m, \sec \theta \phi e_1, \dots, \sec \theta \phi e_m\}$ be orthonormal basis of $\Gamma(\ker F_*)$, $\{\csc \theta \omega e_1, \dots, \csc \theta \omega e_m\}$ be orthonormal basis of $\Gamma(\omega(\ker F_*))$ and $\{\mu_1, \dots, \mu_r, J_1 \mu_1, \dots, J_1 \mu_r\}$ be orthonormal basis of $\Gamma(\mu)$. Then the trace of second fundamental form (restriction to $\ker F_* \times \ker F_*$) is given by

$$\begin{aligned} \text{trace}^{\ker F_*} \nabla F_* &= \sum_{i=1}^m (\nabla F_*)(e_i, e_i) + (\nabla F_*)(\sec \theta \phi e_i, \sec \theta \phi e_i) \\ &= \sum_{i=1}^m (\nabla F_*)(e_i, e_i) + \sec^2 \theta (\nabla F_*)(\phi e_i, \phi e_i). \end{aligned}$$

Then using (2.9) we obtain

$$\begin{aligned} \text{trace}^{\ker F_*} \nabla F_* &= -\frac{1}{m} F_*(T_{e_i} e_i) - \frac{1}{m} F_*(\sec^2 \theta T_{\phi e_i} \phi e_i) \\ &= -\frac{1}{m} F_*(T_{e_i} e_i + \sec^2 \theta T_{\phi e_i} \phi e_i). \end{aligned} \tag{3.10}$$

In a similar way, we have

$$\begin{aligned} \text{trace}^{(\ker F_*)^\perp} \nabla F_* &= \sum_{i=1}^m (\nabla F_*)(\csc \theta \omega e_i, \csc \theta \omega e_i) + \sum_{i=1}^{2r} (\nabla F_*)(\mu_i, \mu_i) \\ &= \csc^2 \theta \sum_{i=1}^m (\nabla F_*)(\omega e_i, \omega e_i) + \sum_{i=1}^{2r} (\nabla F_*)(\mu_i, \mu_i). \end{aligned}$$

Using Lemma 2.3 we arrive at

$$\begin{aligned}
\text{trace}^{(\ker F_*)^\perp} \nabla F_* &= \csc^2 \theta \sum_{i=1}^m \{ \omega e_i (\ln \lambda) F_* \omega e_i + \omega e_i (\ln \lambda) F_* \omega e_i \\
&\quad - g_1(\omega e_i, \omega e_i) F_*(\text{grad} \ln \lambda) \} \\
&\quad + \sum_{i=1}^{2r} \{ \mu_i (\ln \lambda) F_* \mu_i + \mu_i (\ln \lambda) F_* \mu_i - g_1(\mu_i, \mu_i) F_*(\text{grad} \ln \lambda) \} \\
&= \csc^2 \theta \sum_{i=1}^m 2g_1(\mathcal{H} \text{grad} \ln \lambda, \omega e_i) F_* \omega e_i \\
&\quad - \csc^2 \theta g_1(\omega e_i, \omega e_i) F_*(\text{grad} \ln \lambda) \\
&\quad + \sum_{i=1}^{2r} 2g_1(\mathcal{H} \text{grad} \ln \lambda, \mu_i) F_* \mu_i - 2r F_*(\text{grad} \ln \lambda).
\end{aligned}$$

Since F is a conformal slant submersion, we derive

$$\begin{aligned}
\text{trace}^{(\ker F_*)^\perp} \nabla F_* &= \csc^2 \theta \sum_{i=1}^m \frac{2}{\lambda^2} g_2(F_*(\text{grad} \ln \lambda), F_* \omega e_i) F_* \omega e_i \\
&\quad + \sum_{i=1}^{2r} \frac{2}{\lambda^2} g_2(F_*(\text{grad} \ln \lambda), F_* \mu_i) F_* \mu_i - (m+2r) F_*(\text{grad} \ln \lambda) \\
&= \frac{2}{\lambda^2} F_*(\text{grad} \ln \lambda) - (m+2r) F_*(\text{grad} \ln \lambda) \\
&= \left(\frac{2}{\lambda^2} - (m+2r) \right) F_*(\text{grad} \ln \lambda). \tag{3.11}
\end{aligned}$$

Then proof follows from (3.10) and (3.11). \square

We note that for any C^2 real valued function f defined on an open subset of a Riemannian manifold M , the equation $\Delta f = 0$ is called Laplace's equation and solutions are called harmonic functions on U . Let $F : M \rightarrow N$ be a smooth map between Riemannian manifolds. Then F is called harmonic morphism if, for every harmonic function $f : V \rightarrow \mathbf{R}$ defined on an open subset V of N with $F^{-1}(V)$ non-empty, the composition $f \circ F$ is harmonic on $F^{-1}(V)$. It is known that a smooth map $F : M \rightarrow N$ between Riemannian manifolds is a harmonic morphism if and only if F is both harmonic and horizontally weakly conformal [23] and [33]. Thus from Theorem 3.14 we deduce the following result.

Theorem 3.15. *Let $F : (M_1^{2(m+r)}, g_1, J_1) \rightarrow (M_2^{m+2r}, g_2)$ be a conformal slant submersion such that $\frac{2}{(m+2r)} \neq \lambda^2$ where (M_1, g_1, J_1) is a Kähler manifold and (M_2, g_2) is a Riemannian manifold. Then any two conditions below imply the third:*

- (i) F is a harmonic morphism
- (ii) ω is parallel with respect to ∇ on $(\ker F_*)$
- (iii) F is a horizontally homotetic map.

We also have the following result.

Corollary 3.16. *Let F be a conformal slant submersion from a Kähler manifold $(M_1^{2(m+r)}, g_1, J_1)$ to a Riemannian manifold (M_2^{m+2r}, g_2) . If $\frac{2}{(m+2r)} = \lambda^2$ then F is harmonic morphism if and only if ω is parallel with respect to ∇ on $(\ker F_*)$.*

Remark 3.17. By arguing as in [6, Proposition 3.5.1, Theorem 4.5.4], one can see that Theorem 3.15 and Corollary 3.16 are also valid for a horizontally weakly conformal map.

We note that if $\frac{2}{(m+2r)} = \lambda^2$ is satisfied, then F is certainly horizontally homothetic. We now study the integrability of the distribution $(\ker F_*)^\perp$ and then we investigate the geometry of leaves of $(\ker F_*)^\perp$ and $(\ker F_*)$. We note that it is known that the distribution $\ker F_*$ is integrable.

Theorem 3.18. *Let F be a proper conformal slant submersion from a Kähler manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then the following assertions are equivalent to each other;*

- (i) $(\ker F_*)^\perp$ is integrable,
- (ii) $\frac{1}{\lambda^2} \{g_2(\nabla_Y^F F_* \mathcal{C}X - \nabla_X^F F_* \mathcal{C}Y, F_* \omega V) - g_2(\nabla_Y^F F_* X - \nabla_X^F F_* Y, F_* \omega \phi V)\}$
 $= g_1(A_X \mathcal{B}Y - A_Y \mathcal{B}X$
 $- \mathcal{C}Y(\ln \lambda)X + \mathcal{C}X(\ln \lambda)Y$
 $- 2g_1(\mathcal{C}X, Y) \ln \lambda, \omega V)$

for $X, Y \in \Gamma((\ker F_*)^\perp)$, $V \in \Gamma(\ker F_*)$.

Proof. For $X, Y \in \Gamma((\ker F_*)^\perp)$ and $V \in \Gamma(\ker F_*)$, using (2.1), (2.2) and (3.1) we have

$$g_1([X, Y], V) = -g_1(\nabla_X Y, J\phi V) + g_1(\nabla_X JY, \omega V) + g_1(\nabla_Y X, J\phi V) - g_1(\nabla_Y JX, \omega V).$$

Then by using (3.2), we get

$$\begin{aligned} g_1([X, Y], V) &= -g_1(\nabla_X Y, \phi^2 V) - g_1(\nabla_X Y, \omega \phi V) + g_1(\nabla_X \mathcal{B}Y, \omega V) \\ &\quad + g_1(\nabla_X \mathcal{C}Y, \omega V) \\ &\quad + g_1(\nabla_Y X, \phi^2 V) + g_1(\nabla_Y X, \omega \phi V) - g_1(\nabla_Y \mathcal{B}X, \omega V) \\ &\quad - g_1(\nabla_Y \mathcal{C}X, \omega V). \end{aligned}$$

Since F is a conformal submersion, using (2.7), Theorem 3.7 and Lemma 2.3 we arrive at

$$\begin{aligned} g_1([X, Y], V) &= \cos^2 \theta g_1([X, Y], V) + g_1(A_X \mathcal{B}Y - A_Y \mathcal{B}X, \omega V) \\ &\quad + \frac{1}{\lambda^2} g_2((\nabla F_*)(X, Y), F_* \omega \phi V) \\ &\quad - \frac{1}{\lambda^2} g_2(\nabla_X^F F_* Y, F_* \omega \phi V) - g_1(\text{grad} \ln \lambda, X) g_1(\mathcal{C}Y, \omega V) - g_1(\text{grad} \ln \lambda, \mathcal{C}Y) g_1(X, \omega V) \\ &\quad + g_1(X, \mathcal{C}Y) g_1(\text{grad} \ln \lambda, \omega V) + \frac{1}{\lambda^2} g_2(\nabla_X^F F_* \mathcal{C}Y, F_* \omega V) - \frac{1}{\lambda^2} g_2((\nabla F_*)(Y, X), F_* \omega \phi V) \\ &\quad + \frac{1}{\lambda^2} g_2(\nabla_Y^F F_* X, F_* \omega \phi V) + g_1(\text{grad} \ln \lambda, Y) g_1(\mathcal{C}X, \omega V) + g_1(\text{grad} \ln \lambda, \mathcal{C}X) g_1(Y, \omega V) \\ &\quad - g_1(Y, \mathcal{C}X) g_1(\text{grad} \ln \lambda, \omega V) - \frac{1}{\lambda^2} g_2(\nabla_Y^F F_* \mathcal{C}X, F_* \omega V). \end{aligned}$$

Since ∇F_* is symmetric, we have

$$\begin{aligned} \sin^2 \theta g_1([X, Y], V) &= g_1(A_X \mathcal{B}Y - A_Y \mathcal{B}X - \mathcal{C}Y(\ln \lambda)X + \mathcal{C}X(\ln \lambda)Y \\ &\quad - 2g_1(\mathcal{C}X, Y) \ln \lambda, \omega V) \\ &\quad + \frac{1}{\lambda^2} \{g_2(\nabla_Y^F F_* X - \nabla_X^F F_* Y, F_* \omega \phi V) \\ &\quad - g_2(\nabla_Y^F F_* \mathcal{C}X - \nabla_X^F F_* \mathcal{C}Y, F_* \omega V)\} \end{aligned}$$

which proves assertion. \square

From Theorem 3.18, we deduce the following which shows that a conformal slant submersion with integrable $(\ker F_*)^\perp$ turns out to be a horizontally homothetic submersion.

Theorem 3.19. *Let F be a proper conformal slant submersion from a Kähler manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then any two conditions below imply the three;*

- (i) $(\ker F_*)^\perp$ is integrable
- (ii) F is horizontally homotetic.
- (iii) $\frac{1}{\lambda^2} \{g_2(\nabla_Y^F F_* \mathcal{C}X - \nabla_X^F F_* \mathcal{C}Y, F_* \omega V) - g_2(\nabla_Y^F F_* X - \nabla_X^F F_* Y, F_* \omega \phi V)\}$
 $= g_1(A_X \mathcal{B}Y - A_Y \mathcal{B}X, \omega V)$

for $X, Y \in \Gamma((\ker F_*)^\perp)$, $V \in \Gamma(\ker F_*)$.

Proof. For $X, Y \in \Gamma((\ker F_*)^\perp)$, $V \in \Gamma(\ker F_*)$, from Theorem 3.18, we have

$$\begin{aligned} \sin^2 \theta g_1([X, Y], V) &= g_1(A_X \mathcal{B}Y - A_Y \mathcal{B}X - \mathcal{C}Y(\ln \lambda)X + \mathcal{C}X(\ln \lambda)Y \\ &\quad - 2g_1(\mathcal{C}X, Y) \ln \lambda, \omega V) \\ &\quad + \frac{1}{\lambda^2} \{g_2(\nabla_Y^F F_* X - \nabla_X^F F_* Y, F_* \omega \phi V) \\ &\quad - g_2(\nabla_Y^F F_* \mathcal{C}X - \nabla_X^F F_* \mathcal{C}Y, F_* \omega V)\}. \end{aligned}$$

Now, if we have (i) and (iii), then we arrive at

$$\begin{aligned} &-g_1(\mathcal{H}grad \ln \lambda, \mathcal{C}Y)g_1(X, \omega V) + g_1(\mathcal{H}grad \ln \lambda, \mathcal{C}X)g_1(Y, \omega V) \\ &-2g_1(\mathcal{C}X, Y)g_1(\mathcal{H}grad \ln \lambda, \omega V) = 0. \end{aligned} \quad (3.12)$$

Now, taking $Y = JV$ in (3.12) for $V \in \Gamma(\ker F_*)$, we get

$$g_1(\mathcal{H}grad \ln \lambda, \mathcal{C}X)g_1(\omega V, \omega V) = 0.$$

Hence λ is a constant on $\Gamma(\mu)$. On the other hand, taking $Y = \mathcal{C}X$ in (3.12) for $X \in \Gamma(\mu)$, we derive

$$\begin{aligned} &-g_1(\mathcal{H}grad \ln \lambda, \mathcal{C}^2 X)g_1(X, \omega V) + g_1(\mathcal{H}grad \ln \lambda, \mathcal{C}X)g_1(\mathcal{C}X, \omega V) \\ &-2g_1(\mathcal{C}X, \mathcal{C}X)g_1(\mathcal{H}grad \ln \lambda, \omega V) = 0, \end{aligned}$$

hence, we arrive at

$$g_1(\mathcal{C}X, \mathcal{C}X)g_1(\mathcal{H}grad \ln \lambda, \omega V) = 0.$$

From above equation, λ is a constant on $\Gamma(\omega(\ker F_*))$. Similarly, one can obtain the other assertions. \square

Theorem 3.20. *Let F be a proper conformal slant submersion from a Kähler manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then the distribution $(\ker F_*)^\perp$ defines a totally geodesic foliation on M_1 if and only if*

$$\begin{aligned} \frac{1}{\lambda^2} \{g_2(\nabla_X^F F_* Y, F_* \omega \phi V) - g_2(\nabla_X^F F_* \mathcal{C}Y, F_* \omega V)\} &= g_1(A_X \mathcal{B}Y, \omega V) \\ &\quad + g_1(grad \ln \lambda, X)g_1(Y, \omega \phi V) \\ &\quad + g_1(grad \ln \lambda, Y)g_1(X, \omega \phi V) \\ &\quad - g_1(X, Y)g_1(grad \ln \lambda, \omega \phi V) \\ &\quad - g_1(grad \ln \lambda, X)g_1(\mathcal{C}Y, \omega V) \\ &\quad - g_1(grad \ln \lambda, \mathcal{C}Y)g_1(X, \omega V) \\ &\quad + g_1(X, \mathcal{C}Y)g_1(grad \ln \lambda, \omega V) \end{aligned}$$

for $X, Y \in \Gamma((\ker F_*)^\perp)$, $V \in \Gamma(\ker F_*)$.

Proof. For $X, Y \in \Gamma((\ker F_*)^\perp)$ and $V \in \Gamma(\ker F_*)$, using (2.1), (2.2), (3.1) and (3.2) we have

$$g_1(\nabla_X Y, V) = -g_1(\nabla_X Y, \phi^2 V) - g_1(\nabla_X Y, \omega \phi V) + g_1(\nabla_X \mathcal{B}Y, \omega V) + g_1(\nabla_X \mathcal{C}Y, \omega V).$$

Since F is a conformal submersion, using (2.7), Theorem 3.7 and Lemma 2.3 we arrive at

$$\begin{aligned} g_1(\nabla_X Y, V) &= \cos^2 \theta g_1(\nabla_X Y, V) + g_1(A_X \mathcal{B}Y, \omega V) + g_1(\text{grad} \ln \lambda, X) g_1(Y, \omega \phi V) \\ &\quad + g_1(\text{grad} \ln \lambda, Y) g_1(X, \omega \phi V) - g_1(X, Y) g_1(\text{grad} \ln \lambda, \omega \phi V) \\ &\quad - \frac{1}{\lambda^2} g_2(\nabla_X^F F_* Y, F_* \omega \phi V) \\ &\quad - g_1(\text{grad} \ln \lambda, X) g_1(\mathcal{C}Y, \omega V) - g_1(\text{grad} \ln \lambda, \mathcal{C}Y) g_1(X, \omega V) \\ &\quad + g_1(X, \mathcal{C}Y) g_1(\text{grad} \ln \lambda, \omega V) + \frac{1}{\lambda^2} g_2(\nabla_X^F F_* \mathcal{C}Y, F_* \omega V). \end{aligned}$$

Hence we have

$$\begin{aligned} \sin^2 \theta g_1(\nabla_X Y, V) &= g_1(A_X \mathcal{B}Y, \omega V) + g_1(\text{grad} \ln \lambda, X) g_1(Y, \omega \phi V) \\ &\quad + g_1(\text{grad} \ln \lambda, Y) g_1(X, \omega \phi V) \\ &\quad - g_1(X, Y) g_1(\text{grad} \ln \lambda, \omega \phi V) - g_1(\text{grad} \ln \lambda, X) g_1(\mathcal{C}Y, \omega V) \\ &\quad - g_1(\text{grad} \ln \lambda, \mathcal{C}Y) g_1(X, \omega V) + g_1(X, \mathcal{C}Y) g_1(\text{grad} \ln \lambda, \omega V) \\ &\quad - \frac{1}{\lambda^2} \{g_2(\nabla_X^F F_* Y, F_* \omega \phi V) - g_2(\nabla_X^F F_* \mathcal{C}Y, F_* \omega V)\} \end{aligned}$$

which proves assertion. \square

In a similar way we have the following.

Theorem 3.21. *Let F be a proper conformal slant submersion from a Kähler manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then the distribution $(\ker F_*)$ defines a totally geodesic foliation on M_1 if and only if*

$$\begin{aligned} \frac{1}{\lambda^2} \{g_2((\nabla F_*)(U, \omega \phi V), F_* Z) - g_2(\nabla_{\omega V}^F F_* \omega U, F_* J C Z)\} &= g_1(A_{\omega V} \phi U \\ &\quad + g_1(\omega U, \omega V) \text{grad} \ln \lambda, J C Z) \\ &\quad + g_1(T_U \mathcal{B}Z, \omega V) \end{aligned}$$

for $U, V \in \Gamma(\ker F_*)$ and $Z \in \Gamma((\ker F_*)^\perp)$.

From Theorem 3.21, we deduce that:

Theorem 3.22. *Let F be a proper conformal slant submersion from a Kähler manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then any two conditions below imply the three:*

- (i) $\ker F_*$ defines a totally geodesic foliation on M_1 .
- (ii) λ is a constant on $\Gamma(\mu)$.
- (iii) $\frac{1}{\lambda^2} \{g_2((\nabla F_*)(U, \omega \phi V), F_* Z) - g_2(\nabla_{\omega V}^F F_* \omega U, F_* J C Z)\} = g_1(A_{\omega V} \phi U, J C Z) + g_1(T_U \mathcal{B}Z, \omega V)$

for $U, V \in \Gamma(\ker F_*)$ and $Z \in \Gamma((\ker F_*)^\perp)$.

Proof. For $U, V \in \Gamma(\ker F_*)$ and $Z \in \Gamma((\ker F_*)^\perp)$, from Theorem 3.21, we have

$$\begin{aligned} \sin^2 \theta g_1(\nabla_U V, Z) &= g_1(T_U \omega V, \mathcal{B}Z) - g_1(A_{\omega V} \phi U, J C Z) \\ &\quad - g_1(\omega V, \omega U) g_1(\mathcal{H} \text{grad} \ln \lambda, J C Z) \\ &\quad + \frac{1}{\lambda^2} \{g_2((\nabla F_*)(U, \omega \phi V), F_* Z) - g_2(\nabla_{\omega V}^F F_* \omega U, F_* J C Z)\}. \end{aligned}$$

Now, if we have (i) and (iii), then we get

$$g_1(\omega V, \omega U) g_1(\mathcal{H} \text{grad} \ln \lambda, J C Z) = 0.$$

From above equation, λ is a constant on $\Gamma(\mu)$. Similarly, one can obtain the other assertions. \square

From Theorem 3.20 and Theorem 3.21 we have the following result.

Corollary 3.23. *Let F be a proper conformal slant submersion from a Kähler manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then M_1 is a locally product Riemannian manifold if and only if*

$$\begin{aligned} \frac{1}{\lambda^2} \{g_2(\nabla_X^F F_* Y, F_* \omega \phi V) - g_2(\nabla_X^F F_* \mathcal{C}Y, F_* \omega V)\} &= g_1(A_X \mathcal{B}Y, \omega V) \\ &+ g_1(\text{grad} \ln \lambda, X) g_1(Y, \omega \phi V) \\ &+ g_1(\text{grad} \ln \lambda, Y) g_1(X, \omega \phi V) \\ &- g_1(X, Y) g_1(\text{grad} \ln \lambda, \omega \phi V) \\ &- g_1(\text{grad} \ln \lambda, X) g_1(\mathcal{C}Y, \omega V) \\ &- g_1(\text{grad} \ln \lambda, \mathcal{C}Y) g_1(X, \omega V) \\ &+ g_1(X, \mathcal{C}Y) g_1(\text{grad} \ln \lambda, \omega V) \end{aligned}$$

and

$$\begin{aligned} \frac{1}{\lambda^2} \{g_2((\nabla F_*)(U, \omega \phi V), F_* Z) - g_2(\nabla_{\omega V}^F F_* \omega U, F_* J \mathcal{C}Z)\} &= g_1(A_{\omega V} \phi U \\ &+ g_1(\omega U, \omega V) \text{grad} \ln \lambda, J \mathcal{C}Z) \\ &+ g_1(T_U \mathcal{B}Z, \omega V) \end{aligned}$$

for $X, Y, Z \in \Gamma((\ker F_*)^\perp)$ and $U, V \in \Gamma(\ker F_*)$.

Finally we obtain necessary and sufficient condition for a conformal slant submersion to be totally geodesic. We recall that a differentiable map F between two Riemannian manifolds is called totally geodesic if

$$(\nabla F_*)(X, Y) = 0 \quad \forall X, Y \in \Gamma(TM).$$

A geometric interpretation of a totally geodesic map is that it maps every geodesic in the total manifold into a geodesic in the base manifold in proportion to arc lengths.

Theorem 3.24. *Let F be a proper conformal slant submersion from a Kähler manifold (M_1, g_1, J_1) onto a Riemannian manifold (M_2, g_2) . Then F is a totally geodesic map if and only if*

- (i) $\frac{1}{\lambda^2} \{g_2((\nabla F_*)(U, \omega \phi V), F_* Z) - g_2(\nabla_{\omega V}^F F_* \omega U, F_* J \mathcal{C}Z)\} = g_1(A_{\omega V} \phi U, J \mathcal{C}Z) + g_1(T_U \omega V, \mathcal{B}Z)$,
- (ii) $\frac{1}{\lambda^2} \{g_2((\nabla F_*)(U, \omega \mathcal{B}X), F_* Y) + g_2((\nabla F_*)(U, \mathcal{C}X), F_* \mathcal{C}Y)\} = g_1(T_U \phi \mathcal{B}X, Y) - g_1(T_U \mathcal{C}X, \mathcal{B}Y)$,
- (iii) F is a horizontally homothetic map

for $U, V \in \Gamma(\ker F_*)$ and $X, Y, Z \in \Gamma((\ker F_*)^\perp)$.

Proof. (i) For $U, V \in \Gamma(\ker F_*)$ and $Z \in \Gamma((\ker F_*)^\perp)$, using (2.1), (2.2), (3.1), (3.2 and Lemma 2.3 we have

$$\begin{aligned} \frac{1}{\lambda^2} g_2((\nabla F_*)(U, V), F_* Z) &= g_1(\nabla_U \phi^2 V, Z) + g_1(\nabla_U \omega \phi V, Z) \\ &- g_1(\nabla_U \omega V, \mathcal{B}Z) - g_1(\nabla_U \omega V, \mathcal{C}Z) \\ &= g_1(\nabla_U \phi^2 V, Z) + g_1(\nabla_U \omega \phi V, Z) \\ &- g_1(\nabla_U \omega V, \mathcal{B}Z) + g_1(\nabla_{\omega V} J U, J \mathcal{C}Z). \end{aligned}$$

Using (2.5), Theorem 3.7 and Lemma 2.3 we arrive at

$$\begin{aligned} \frac{1}{\lambda^2}g_2((\nabla F_*)(U, V), F_*Z) &= -\cos^2\theta g_1(\nabla_U V, Z) - g_1(T_U\omega V, \mathcal{B}Z) + g_1(A_{\omega V}\phi U, J\mathcal{C}Z) \\ &\quad - \frac{1}{\lambda^2}g_2((\nabla F_*)(U, \omega\phi V), F_*Z) - g_1(\text{grad}\ln\lambda, \omega V)g_1(\omega U, J\mathcal{C}Z) \\ &\quad - g_1(\text{grad}\ln\lambda, \omega U)g_1(\omega V, J\mathcal{C}Z) + g_1(\omega V, \omega U)g_1(\text{grad}\ln\lambda, J\mathcal{C}Z) \\ &\quad + \frac{1}{\lambda^2}g_2(\nabla_{\omega V}^F F_*\omega U, F_*J\mathcal{C}Z). \end{aligned}$$

Hence we have

$$\begin{aligned} \sin^2\theta \frac{1}{\lambda^2}g_2((\nabla F_*)(U, V), F_*Z) &= g_1(A_{\omega V}\phi U, J\mathcal{C}Z) - g_1(\omega V, \omega U)g_1(\text{grad}\ln\lambda, J\mathcal{C}Z) \\ &\quad - g_1(T_U\omega V, \mathcal{B}Z) \\ &\quad + \frac{1}{\lambda^2}\{g_2((\nabla F_*)(U, \omega\phi V), F_*Z) + g_2(\nabla_{\omega V}^F F_*\omega U, F_*J\mathcal{C}Z)\}. \end{aligned}$$

(ii) For $X, Y \in \Gamma((\ker F_*)^\perp)$ and $U \in \Gamma(\ker F_*)$, in a similar way

$$\begin{aligned} \frac{1}{\lambda^2}g_2((\nabla F_*)(U, X), F_*Y) &= g_1(\nabla_U\phi\mathcal{B}X, Y) + g_1(\nabla_U\omega\mathcal{B}X, Y) \\ &\quad - g_1(\nabla_U\mathcal{C}X, \mathcal{B}Y) - g_1(\nabla_U\mathcal{C}X, \mathcal{C}Y). \end{aligned}$$

Using also (2.5), Theorem 3.7 and Lemma 2.3 we arrive at

$$\begin{aligned} \frac{1}{\lambda^2}g_2((\nabla F_*)(U, X), F_*Y) &= g_1(T_U\phi\mathcal{B}X, Y) + \frac{1}{\lambda^2}g_2(F_*(\nabla_U\omega\mathcal{B}X), F_*Y) \\ &\quad - g_1(T_U\mathcal{C}X, \mathcal{B}Y) - \frac{1}{\lambda^2}g_2(F_*(\nabla_U\mathcal{C}X), F_*\mathcal{C}Y) \\ &= g_1(T_U\phi\mathcal{B}X, Y) + \frac{1}{\lambda^2}g_2(-(\nabla F_*)(U, \omega\mathcal{B}X) + \nabla_U^F\omega\mathcal{B}X, F_*Y) \\ &\quad - g_1(T_U\mathcal{C}X, \mathcal{B}Y) - \frac{1}{\lambda^2}g_2(-(\nabla F_*)(U, \mathcal{C}X) + \nabla_U^F\mathcal{C}X, F_*\mathcal{C}Y) \\ &= g_1(T_U\phi\mathcal{B}X, Y) - g_1(T_U\mathcal{C}X, \mathcal{B}Y) \\ &\quad + \frac{1}{\lambda^2}\{g_2((\nabla F_*)(U, \mathcal{C}X), F_*\mathcal{C}Y) - g_2((\nabla F_*)(U, \omega\mathcal{B}X), F_*Y)\}. \end{aligned}$$

(iii) For $X, Y \in \Gamma(\mu)$, from Lemma 2.3, we have

$$(\nabla F_*)(X, Y) = X(\ln\lambda)F_*Y + Y(\ln\lambda)F_*X - g_1(X, Y)F_*(\text{grad}\ln\lambda).$$

From above equation, taking $Y = JX$ for $X \in \Gamma(\mu)$ we obtain

$$\begin{aligned} (\nabla F_*)(X, JX) &= X(\ln\lambda)F_*JX + JX(\ln\lambda)F_*X - g_1(X, JX)F_*(\text{grad}\ln\lambda) \\ &= X(\ln\lambda)F_*JX + JX(\ln\lambda)F_*X. \end{aligned}$$

If $(\nabla F_*)(X, JX) = 0$, we obtain

$$X(\ln\lambda)F_*JX + JX(\ln\lambda)F_*X = 0. \quad (3.13)$$

Taking inner product in (3.13) with F_*JX we have

$$g_1(\text{grad}\ln\lambda, X)g_2(F_*JX, F_*JX) + g_1(\text{grad}\ln\lambda, X)g_2(F_*X, F_*JX) = 0.$$

From above equation, it follows that λ is a constant on $\Gamma(\mu)$. In a similar way, for $U, V \in \Gamma(\ker F_*)$, using Lemma 2.3 we have

$$(\nabla F_*)(\omega U, \omega V) = \omega U(\ln\lambda)F_*\omega V + \omega V(\ln\lambda)F_*\omega U - g_1(\omega U, \omega V)F_*(\text{grad}\ln\lambda).$$

From above equation, taking $V = U$ we obtain

$$\begin{aligned} (\nabla F_*)(\omega U, \omega U) &= \omega U(\ln \lambda)F_*\omega U + \omega U(\ln \lambda)F_*\omega U - g_1(\omega U, \omega U)F_*(\text{grad} \ln \lambda) \\ &= 2\omega U(\ln \lambda)F_*\omega U - g_1(\omega U, \omega U)F_*(\text{grad} \ln \lambda). \end{aligned} \quad (3.14)$$

Taking inner product in (3.14) with $F_*\omega U$ and since F is a conformal submersion, we derive

$$2g_1(\text{grad} \ln \lambda, \omega U)g_2(F_*\omega U, F_*\omega U) - g_1(\omega U, \omega U)g_2(F_*(\text{grad} \ln \lambda), F_*\omega U) = 0.$$

From above equation, it follows that λ is a constant on $\Gamma(\omega(\ker F_*))$. Thus λ is a constant on $\Gamma((\ker F_*)^\perp)$. On the other hand, if F is a horizontally homothetic map, it is obvious that $(\nabla F_*)(X, Y) = 0$. Thus proof is complete. \square

Acknowledgment. This paper is supported by The Scientific and Technological Council of Turkey (TUBITAK) with number (114F339).

References

- [1] M.A. Akyol, *Conformal anti-invariant submersions from cosymplectic manifolds*, Hacet. J. Math. Stat. **46** (2), 177-192, 2017.
- [2] M.A. Akyol and B. Şahin, *Conformal anti-invariant submersions from almost Hermitian manifolds*, Turk J. Math. **40** (1), 43-70, 2016.
- [3] C. Altafini, *Redundant robotic chains on Riemannian submersions*, 20, IEEE Transactions on Robotics and Automation, 335-340, 2004.
- [4] S. Ali and T. Fatima, *Anti-invariant Riemannian submersions from nearly Kaehler manifolds*, Filomat, **27**, 1219-1235, 2013.
- [5] S. Ali and T. Fatima, *Generic Riemannian submersions*, Tamkang J. Math. **44**, 395-409, 2013.
- [6] P. Baird and J.C. Wood, *Harmonic Morphisms Between Riemannian Manifolds*, London Mathematical Society Monographs, 29, Oxford University Press, The Clarendon Press, Oxford, 2003.
- [7] A. Beri, I.K. Erken and C. Murathan, *Anti-Invariant Riemannian submersions from Kenmotsu manifolds onto Riemannian manifolds*, Turk. J. Math. **40**, 540-552, 2016.
- [8] A.L. Besse, *Einstein Manifolds*, Springer, Berlin, 1987.
- [9] R. Bhattacharyya and V. Patrangenarub, *Nonparametric estimation of location and dispersion on Riemannian manifolds*, J. Statist. Plann. Inference **108**, 23-35, 2002.
- [10] J.P. Bourguignon and H.B. Lawson, *A mathematician's visit to Kaluza-Klein theory*, Rend. Sem. Mat. Univ. Politec. Torino, Special Issue, 143-163, 1989.
- [11] J.L. Cabrerizo, A. Carriazo, L.M. Fernandez and M. Fernandez, *Slant submanifolds in Sasakian manifolds*. Glasg. Math. J. **42** (1), 125-138, 2000.
- [12] B.Y. Chen, *Pseudo-Riemannian geometry, δ -invariants and applications*, World Scientific, Singapore, 2011.
- [13] B.Y. Chen, *Slant immersions*, Bull. Austral. Math. Soc. **41** (1), 135-147, 1990.
- [14] B.Y. Chen, *Geometry of Slant Submanifolds*, Katholieke Universiteit Leuven, Leuven, 1990.
- [15] D. Chinea, *Harmonicity on maps between almost contact metric manifolds*, Acta Math. Hungar. **126**, 352-365, 2010.
- [16] D. Chinea, *Harmonicity of holomorphic maps between almost Hermitian manifolds*, Canad. Math. Bull. **52**, 18-27, 2009.
- [17] D. Chinea, *On horizontally conformal $(\varphi, \hat{\varphi})$ -holomorphic submersions*, Houston J. Math. **34**, 721-737, 2008.
- [18] I.K. Erken and C. Murathan, *On slant Riemannian submersions for cosymplectic manifolds*, Bull. Korean Math. Soc. **51** (6), 1749-1771, 2014.

- [19] I.K. Erken and C. Murathan, *Anti-Invariant Riemannian submersions from Cosymplectic manifolds onto Riemannian manifolds*, Filomat **29**, 1429-1444, 2015.
- [20] I.K. Erken and C. Murathan, *Slant Riemannian submersions from Sasakian manifolds*, Arab J. Math. Sci. **22**, 250-264, 2016.
- [21] M. Falcitelli, S. Ianus and A.M. Pastore, *Riemannian submersions and Related Topics*, World Scientific, River Edge, NJ, 2004.
- [22] M. Falcitelli, S. Ianus, A.M. Pastore and M. Visinescu, *Some applications of Riemannian submersions in physics*, Revue Roumaine de Physique **48**, 627-639, 2003.
- [23] B. Fuglede, *Harmonic Morphisms Between Riemannian Manifolds*, Ann. Inst. Fourier (Grenoble) **28**, 107-144, 1978.
- [24] P. Gilkey, M. Itoh and J.H. Park, *Anti-invariant Riemannian submersions: A Lie-theoretical approach*, Taiwanese J. Math. **20** (4), 787-800, 2016.
- [25] A. Gray, *Pseudo-Riemannian almost product manifolds and submersions*, J. Math. Mech. **16**, 715-737, 1967.
- [26] S. Gundmundsson, *The Geometry of Harmonic Morphisms*, Ph.D. Thesis, University of Leeds, 1992.
- [27] S. Gundmundsson and J.C. Wood, *Harmonic Morphisms between almost Hermitian manifolds*, Boll. Un. Mat. Ital. B. **11** (2), 185-197, 1997.
- [28] Y. Gündüzalp, *Slant submersions from almost product Riemannian manifolds*, Turk. J. Math. **37** (5), 863-873, 2013.
- [29] Y. Gündüzalp, *Anti-invariant semi-Riemannian submersions from almost para-Hermitian manifolds*, J. Funct. Spaces Appl **2013**, Article ID 720623, 7 pages, 2013
- [30] Y. Gündüzalp, *Anti-invariant Riemannian submersions from almost product manifolds*, Mathematical Sciences and Applications E-Notes (MSAEN) **1**, 58-66, 2013.
- [31] Y. Gündüzalp, *Slant submersions from Lorentzian almost paracontact manifolds*, Gulf J. Math. **3**, 18-28, 2015.
- [32] Y. Gündüzalp, *Slant submersions from almost paracontact Riemannian manifolds*, Kuwait J. Sci. **42**, 17-29, 2015.
- [33] T. Ishihara, *A mapping of Riemannian manifolds which preserves harmonic functions*, J. Math. Kyoto Univ. **19**, 215-229, 1979.
- [34] J.W. Lee, *Anti-invariant ξ^\perp Riemannian submersions from almost contact manifolds*, Hacet. J. Math. Stat. **42** (3), 231-241, 2013.
- [35] J.C. Lee, J.H. Park, B. Şahin and D.Y. Song, *Einstein conditions for the base space of anti-invariant Riemannian submersions and Clairaut Submersions*, Taiwanese J. Math. **19**, 1145-1160, 2015.
- [36] F. Memoli, G. Sapiro and P. Thompson, *Implicit brain imaging*, NeuroImage **23**, 179-183, 2004.
- [37] B. O'Neill, *The fundamental equations of a submersion*, Mich. Math. J. **13**, 458-469, 1966.
- [38] K.S. Park and R. Prasad, *Semi-slant submersions*, Bull. Korean Math. Soc. **50** (3), 951-962, 2013.
- [39] R. Ponge and H. Reckziegel, *Twisted products in pseudo-Riemannian geometry*, Geom. Dedicata. **48** (1), 15-25, 1993.
- [40] S.A. Sepet and M. Ergüt, *Pointwise slant submersions from cosymplectic manifolds*, Turk. J. Math. **40**, 582-593, 2016.
- [41] B. Şahin, *Anti-invariant Riemannian submersions from almost Hermitian manifolds*, Central European J. Math **3**, 437-447, 2010.
- [42] B. Şahin, *Slant submersions from almost Hermitian manifolds*, Bull. Math. Soc. Sci. Math. Roumanie **1**, 93-105, 2011.
- [43] A.D. Vilcu and G.E. Vilcu, *Statistical manifolds with almost quaternionic Structures and quaternionic Kähler-like statistical submersions*, Entropy **17**, 6213-6228, 2015.

- [44] B. Watson, B. *Almost Hermitian submersions*. J. Differential Geom. **11** (1), 147-165, 1976.
- [45] K. Yano and M. Kon, *Structures on Manifolds*, World Scientific, Singapore, 1984.
- [46] H. Zhao, A.R. Kelly, J. Zhou, L. Lu and Y.Y. Yang, *Graph attribute embedding via Riemannian submersion learning*, Comput. Vis. Image Underst. **115**, 962-975, 2011.



Composition in EL -hyperstructures

Michal Novák¹, Irina Cristea^{*2}

¹*Department of Mathematics, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 10, 616 00 Brno, Czech Republic*

²*Centre for Information Technologies and Applied Mathematics, University of Nova Gorica, Vipaska 13, 5000 Nova Gorica, Slovenia*

Abstract

The link between ordered sets and hyperstructures is one of the classical areas of research in the hyperstructure theory. In this paper we focus on EL -hyperstructures, i.e. a class of hyperstructures constructed from quasi-ordered semigroups. In our paper we link this concept to the concept of a *composition hyperring*, a recent hyperstructure generalization of the classical notion of a composition ring.

Mathematics Subject Classification (2010). 06E20, 20N20, 06F15, 13B99

Keywords. Composition ring, composition hyperring, hyperstructure theory, partially ordered semigroup, quasi-ordered semigroup

1. Introduction

Since the times of elementary algebra, the scope of this mathematical discipline has widened considerably. Already in 1930s, a step from the study of single-valued structures to the study of multi-valued structures was made. This new creation, the *hyperstructure theory*, has since then grown to a fully established branch of algebra with numerous far-reaching applications in geometry, graph-theory, coding theory, medicine, number theory, physics, chemistry, etc. For basic introduction to the theory and applications see [9, 11].

Two important multi-valued analogues of classical topics of algebra intersect in this paper: *the study of ordered sets* and their connection to hyperstructures and *the study of ring-like hyperstructures*.

The ordered sets have been in the focus of attention of the hyperstructure theory since works of Nieminen, Corsini, Rosenberg, Krasner, Mittas, Davvaz, Leoreanu or Chvalina of 1960s to 1990s. Notice that one of the first chapters of [9], a canonical book of the hyperstructure theory, is dedicated to ordered sets. Selected reading on some aspects of the topic includes also works such as [3, 4, 8, 16]. Furthermore, Heidari and Davvaz [16] have recently introduced the notion of *partially ordered semihypergroups*, i.e. have transferred the concept of partially ordered semigroups to hyperstructures.

Krasner [20] introduced the notion of the *hyperfield* and then *hyperring* in order to approximate a local field of positive characteristic by a system of local fields of characteristic

*Corresponding Author.

Email addresses: novakm@feec.vutbr.cz (M. Novák), irinacri@yahoo.co.uk (I. Cristea)

Received: 02.06.2017; Accepted: 21.08.2017

zero. The additive part of this hyperring was a special hypergroup while the multiplicative part was a semigroup. Constructions of these structures can be found in [19, 22, 28]. While studying polynomials over Krasner's hyperrings, Mittas [27] introduced *superrings*, in which both parts, additive and multiplicative, were hyperstructures. G. Massouros, approaching the theory of languages and automata from the point of view of hypercompositional algebra, was led to the introduction of the concepts of *hyperringoid* and *join hyperring* [23, 24]. Also, Vougiouklis [33] generalizing Mittas' superring introduced *hyperrings in the general sense*. Some recent papers on the topic include [2, 7, 13, 26] and a book [11].

Motivated by the study of properties of the hyperring of polynomials [18], Cristea and Jančić-Rašović in [10] introduced the concept of *composition hyperring* as a multi-valued generalisation of an older concept of the *composition ring* introduced in [1]. Notice that as regards single-valued rings, composition leads to interesting applications in rings of polynomials, power series or in the field of rational functions. In [12], the concept of composition is used to construct *composition (m, n, k) -hyperrings*.

In this paper we study composition, suggested by Cristea and Jančić-Rašović, in *EL*-hyperstructures, i.e. in a class of hyperstructures constructed from quasi-ordered semigroups. The authors of [10] define the *composition hyperoperation* in hyperrings in the general case of [32], i.e. in multivalued systems $(R, +, \cdot)$, where $(R, +)$ is a hypergroup, (R, \cdot) is a semihypergroup and the multiplication is distributive with respect to the addition. In our paper we partly broaden this environment by suggesting implications also for cases of $(R, +)$ being a semihypergroup (making use of results achieved in [30]).

2. *EL*-hyperstructures: construction and use

There exist numerous constructions of hyperstructures from given single-valued algebraic structures. The concept of *EL-hyperstructures* was coined by Chvalina in [4] and explored in e.g. [15, 29, 31]. The construction is based on validity of a rather simple and straightforward Lemma 2.1. However, when looking for examples of *EL*-hyperstructures, the simplicity and straightforwardness disappear. Naturally, there are obvious intuitive face-value examples such as $(\mathbb{N}, +, \leq)$ or $(\mathcal{P}(S), \cap, \subseteq)$. *EL*-hyperstructures have also been used in papers such as [5, 6, 14] or Sections 8.3 and 8.4 of book [11] in the context of quasi-ordered semigroups such that the nature of their elements and the operation and ordering follow from the application task. In this respect also notice [21], where *EL*-hyperstructures have been used to construct a class of H_v -matrices. Finally, there is another layer of possible uses: Suppose that we have a set of elements, properties of which can be described by means of numerical values (such as length, cardinality, number of elements of a sequence, etc.). Since number domains with a suitably chosen operation and the natural ordering with respect to size often form quasi-ordered semigroups, Lemma 2.1 presents a natural way of constructing (associative and commutative) hyperstructures out of them. In this paper we intentionally demonstrate our results using the simplest possible examples. For a deeper insight and less obvious and straightforward uses of the construction see the above mentioned references.

Further on we work with *principal ends* (hence *EL* which stands for "Ends lemma"), i.e. for an arbitrary $a \in (S, \leq)$ we set $[a]_{\leq} = \{x \in S; a \leq x\}$.

Lemma 2.1. ([4], Theorem 1.3 & Theorem 1.4, pp. 146–147). *Let (S, \cdot, \leq) be a partially ordered semigroup. The binary hyperoperation $* : S \times S \rightarrow \mathcal{P}^*(S)$ defined by*

$$a * b = [a \cdot b]_{\leq} \tag{2.1}$$

*is associative. The semihypergroup $(S, *)$ is commutative if and only if the semigroup (S, \cdot) is commutative. Furthermore, the following conditions are equivalent:*

1⁰: For any pair $(a, b) \in S^2$ there exists a pair $(c, c') \in S^2$ such that $b \cdot c \leq a$ and $c' \cdot b \leq a$.

2⁰: The associated semihypergroup $(S, *)$ is a hypergroup.

Remark 2.2. If (S, \cdot, \leq) is a partially ordered group, then if we take $c = b^{-1} \cdot a$ and $c' = a \cdot b^{-1}$, then condition 1⁰ is valid. Therefore, if (S, \cdot, \leq) is a partially ordered group, then its associated hyperstructure is a hypergroup. In fact, it is a transposition hypergroup, i.e. our reasoning results in *transposition hyperrings*, which can suggest another line of further research. For the use of transposition axiom in hypercompositional structures see [25]. Cases of (S, \cdot) not being a group yet resulting in a hypergroup $(S, *)$ are discussed in [31]. It can also be easily verified that we can assume *quasi-ordered* structures instead of *partially ordered* ones in Lemma 2.1 (however, beware that in this case commutativity of the hyperoperation does not imply commutativity of the single-valued operation). For details see e.g. [29].

3. Basic notions and concepts, notation

Throughout the paper we work with the following definitions and concepts. By a *hyperring in the general sense* and by a *semihyperring in the general sense* we mean systems $(R, +, \cdot)$ discussed e.g. in [33].

Definition 3.1. ([33], p. 21, included as plain text) $(R, +, \cdot)$ is a hyperring in the general sense if $(R, +)$ is a hypergroup, (\cdot) is associative hyperoperation and the distributive law $x(y + z) \subseteq xy + xz$, $(x + y)z \subseteq xz + yz$ is satisfied for every x, y, z of R . [...] $(R, +, \cdot)$ will be called semihyperring if $(+), (\cdot)$ are associative hyperoperations, where (\cdot) is distributive with respect to $(+)$. The rest of definitions are analogous. If the equality in the distributive law is valid, then the hyperring is called strong or good.

By a *hyperring* and by a *semihyperring* we mean a *good* hyperring, or a *good* semihyperring in the sense of Definition 3.1, respectively. Notice that this means that our concept of *hyperring* is the same as the concept used in [10, 18, 32] yet it permits a generalisation in the sense of inclusions.

Composition hyperrings were introduced in [10] as a special class of hyperrings with one additional property.

Definition 3.2. ([10], Def. 3.1) A composition hyperring is an algebraic structure $(R, +, \cdot, \circ)$, where $(R, +, \cdot)$ is a commutative hyperring and the hyperoperation \circ satisfies the following properties, for any $x, y, z \in R$:

- (1) $(x + y) \circ z = x \circ z + y \circ z$
- (2) $(x \cdot y) \circ z = (x \circ z) \cdot (y \circ z)$
- (3) $x \circ (y \circ z) = (x \circ y) \circ z$.

The binary hyperoperation \circ having the previous properties is called the composition hyperoperation of the hyperring $(R, +, \cdot)$.

To be consistent with the background and reasoning of [1, 10] we further on deal with commutative hyperoperations and composition property only. Notice that in the construction using Lemma 2.1 commutativity of the single-valued operation implies commutativity of the hyperoperation and antisymmetry of \leq turns this implication into equivalence. If $x \circ y$ is a one-element set for all $x, y \in R$, we will speak about an operation rather than a hyperoperation even though it will have to be at certain point applied in an element-wise manner on sets (see below in e.g. (5.7) Theorem 5.10). Throughout the paper we will be interested in the composition (hyper)operation in various types of hyperstructures $(R, +, \cdot)$ – not only in hyperrings but also in hyperrings in the general sense, semihyperrings or semihyperrings in the general sense.

Since we construct hyperoperations from single-valued operations on the same set, we have to alter the standard notation of hyperoperations in ring-like hyperstructures. Thus in our context the symbols $+$ and \cdot will be reserved for single-valued operations and the hyperoperations will be denoted by \oplus and \bullet . The hyperoperations will be constructed from *single-valued quasi-ordered semigroups* using Lemma 2.1, i.e. for all $x, y \in R$, where $(R, +, \leq)$ and (R, \cdot, \leq) are quasi-ordered semigroups, we define

$$a \oplus b = [a + b]_{\leq} = \{x \in R; a + b \leq x\} \quad (3.1)$$

and

$$a \bullet b = [a \cdot b]_{\leq} = \{y \in R; a \cdot b \leq y\} \quad (3.2)$$

and get hyperstructures (R, \oplus, \bullet) which we then study. Since (R, \oplus) and (R, \bullet) are *EL*-hyperstructures, it is possible to apply results achieved in [29–31] and immediately state further properties of both (R, \oplus) , (R, \bullet) and (R, \oplus, \bullet) .

4. *EL*-hyperstructures with two hyperoperations

First we show the variety of *EL*-hyperstructures with two hyperoperations which can be obtained using hyperoperations (3.1) and (3.2). Thus the following lemma, included in [30] as Theorems 5.2, 5.4 and 5.5., bounds the area of our future considerations.

Lemma 4.1. *Let $(R, +, \leq)$ and (R, \cdot, \leq) be quasi-ordered semigroups and \oplus, \bullet hyperoperations defined by (3.1) and (3.2) respectively. Furthermore, let \cdot distribute over $+$ from both left and right.*

- (1) *(R, \oplus, \bullet) is a semihyperring in the general sense.*
- (2) *If $(R, +)$ is a group or if (R, \oplus) is a hypergroup, then (R, \oplus, \bullet) is a hyperring in the general sense.*
- (3) *If (R, \cdot) is a group, then (R, \oplus, \bullet) is a semihyperring.*
- (4) *If $(R, +)$ is a group with neutral element 0 and $(R \setminus \{0\}, \cdot)$ is a group, then (R, \oplus, \bullet) is a hyperring.*

Proof. The proof is included in [30] and is based on use of [30], Lemma 4.1, Lemma 4.4, which discuss distributivity, and Remark 4.8, which discusses the role of the absorbing element of the single-valued ring-like structures. Since Lemma 4.1 is important in the context of this paper and not including at least a sketch of its proof would not be correct, we include the main idea of the proof here.

First we show that, for all $a, b, c \in R$, where $(R, +, \leq)$ and (R, \cdot, \leq) are quasi-ordered semigroups, there is

$$\begin{aligned} a \cdot (b + c) = a \cdot b + a \cdot c &\Rightarrow a \bullet (b \oplus c) \subseteq a \bullet b \oplus a \bullet c \\ (a + b) \cdot c = a \cdot c + b \cdot c &\Rightarrow (a \oplus b) \bullet c \subseteq a \bullet c \oplus b \bullet c \end{aligned} \quad (4.1)$$

This is done in the usual way of rewriting both sides of the inclusions using (3.1) and (3.2) and then proving that an arbitrary element from one side of the inclusion is included in the other one.

If we now suppose that (R, \cdot, \leq) is a quasi-ordered group, then with the help of inverse elements we are able to prove the opposite inclusions, i.e.

$$\begin{aligned} a \cdot (b + c) = a \cdot b + a \cdot c &\Rightarrow a \bullet (b \oplus c) \supseteq a \bullet b \oplus a \bullet c \\ (a + b) \cdot c = a \cdot c + b \cdot c &\Rightarrow (a \oplus b) \bullet c \supseteq a \bullet c \oplus b \bullet c \end{aligned} \quad (4.2)$$

for all $a, b, c \in R$.

To complete the proof we need to discuss the role of the potentially existing absorbing elements. Suppose $a = 0$ (or $c = 0$ in the second inclusion) in (4.1). We get $[0]_{\leq} \subseteq \bigcup_{x, y \in [0]_{\leq}} [x + y]_{\leq}$ for $a = 0$ or $\{0\} \subseteq [0]_{\leq}$ for $c = 0$. Since the relation \leq is reflexive, this

obviously holds and does not cause any problems. If we suppose $a = 0$ (or $c = 0$) in (4.2), we get that

$$\bigcup_{x,y \in [0]_{\leq}} [x + y]_{\leq} \subseteq \bigcup_{h \in [b+c]_{\leq}} [0 \cdot h]_{\leq} = [0]_{\leq}.$$

However $x, y \in [0]_{\leq}$ means that $0 \leq x, 0 \leq y$, i.e. $0 = 0 + 0 \leq x + y$, i.e.

$$\bigcup_{x,y \in [0]_{\leq}} [x + y]_{\leq} = [0]_{\leq},$$

i.e. we get equality $[0]_{\leq} = [0]_{\leq}$. If in the second inclusion $c = 0$, then we get the same equality $[0]_{\leq} = [0]_{\leq}$.

Thus we have shown the respective parts on distributivity. The rest follows from Lemma 2.1 and definitions of the respective ring-like hyperstructures. \square

Remark 4.2. Notice that [31] discusses conditions under which Lemma 2.1 applied on a quasi-ordered semigroup which is not a group constructs a hypergroup. In this respect Lemma 4.1, item 2, could be made stronger – see Example 4.3. The same holds for analogous situations, e.g. below in Theorem 6.2.

Example 4.3. Regard an arbitrary set S and its power set $\mathcal{P}(S)$. The operations \cap, \cup of set intersection and set union are associative, thus $(\mathcal{P}(S), \cap)$ and $(\mathcal{P}(S), \cup)$ are semigroups. The relation \subseteq on $\mathcal{P}(S)$ is obviously reflexive and transitive and for arbitrary $A, B, C \in \mathcal{P}(S)$ such that $A \subseteq B$ there is $A \cap C \subseteq B \cap C$ and $A \cup C \subseteq B \cup C$. Thus if we define hyperoperations \oplus, \bullet for arbitrary $A, B \in \mathcal{P}(S)$ by

$$A \oplus B = [A \cup B]_{\leq} = \{X \in \mathcal{P}(S); A \cup B \subseteq X\} \quad (4.3)$$

and

$$A \bullet B = [A \cap B]_{\leq} = \{Y \in \mathcal{P}(S); A \cap B \subseteq Y\}, \quad (4.4)$$

we get semihypergroups $(\mathcal{P}(S), \oplus)$ and $(\mathcal{P}(S), \bullet)$. Moreover, as set intersection is distributive with respect to set union, $(\mathcal{P}(S), \oplus, \bullet)$ is a semihyperring in the general sense.

5. The composition hyperoperation in various EL -ring-like hyperstructures

In this section we study the potential and limitations of hyperstructures suggested in Section 4 with respect to the composition hyperoperation (or operation). Since the hyperstructures are constructed from single-valued structures, we concentrate on properties of the hyperstructures which follow from properties of the single-valued structures.

In the text below notice the precise meaning of symbols \oplus and \bullet . When applied on single elements, they are used in the meanings (3.1) and (3.2) respectively. However, for all sets $A, B \subseteq R$ there is

$$A \oplus B = \bigcup_{\substack{a \in A \\ b \in B}} [a + b]_{\leq} = \bigcup_{\substack{a \in A \\ b \in B}} \{x \in R; a + b \leq x\} \quad (5.1)$$

and

$$A \bullet B = \bigcup_{\substack{a \in A \\ b \in B}} [a \cdot b]_{\leq} = \bigcup_{\substack{a \in A \\ b \in B}} \{y \in R; a \cdot b \leq y\}. \quad (5.2)$$

First of all we discuss a rather trivial case of *constant composition*.

Definition 5.1. If there is $x \circ y = r \circ s$ for an arbitrary quadruple of elements $x, y, r, s \in R$, we call the composition operation (hyperoperation) \circ , defined in Definition 3.2, constant composition operation (hyperoperation).

The following theorem holds for all types of hyperstructures discussed in Lemma 4.1.

Theorem 5.2. *Let (R, \oplus, \bullet) be a semihyperring in the general sense constructed in Lemma 4.1 from idempotent quasi-ordered semigroups $(R, +, \leq)$ and (R, \cdot, \leq) . Consider $r \in R$ arbitrary. Then \circ defined by*

$$a \circ b = [r]_{\leq} \quad (5.3)$$

for all $a, b \in R$, is a constant composition hyperoperation on (R, \oplus, \bullet) . It is a constant operation if \leq is antisymmetric and r is the greatest element of (R, \leq) .

Proof. In the \oplus, \bullet notation, the left-hand side of Definition 3.2, property 1, reads $(x \oplus y) \circ z$. This is

$$[x + y]_{\leq} \circ z = \bigcup_{\substack{\text{number of elements} \\ \text{of } [x+y]_{\leq} \text{-times}}} [r]_{\leq} = [r]_{\leq}.$$

The right-hand side reads $(x \circ z) \oplus (y \circ z)$, which is

$$[r]_{\leq} + [r]_{\leq} = \bigcup_{a, b \in [r]_{\leq}} [a + b]_{\leq} = \bigcup_{\substack{r \leq a \\ r \leq b}} [a + b]_{\leq}.$$

Since $r \leq a, r \leq b$ implies $r + r \leq a + b$ and the relation \leq is reflexive, there is $[r]_{\leq} + [r]_{\leq} = [r + r]_{\leq}$. For idempotent $+$ there is $r + r = r$, i.e. $[r]_{\leq} + [r]_{\leq} = [r]_{\leq}$.

The same reasoning can be applied on property 2 of Definition 3.2. Property 3 holds obviously. Finally, if r is the greatest element of (R, \leq) , then $[r]_{\leq} = \{r\}$, thus we can speak about an operation instead of a hyperoperation. \square

Example 5.3. If we continue with Example 4.3, where the semihyperring in the general sense of the power set $\mathcal{P}(S)$ is discussed, and define

$$A \circ B = [R]_{\subseteq} = \{T \in \mathcal{P}(S); R \subseteq T\}$$

for an arbitrary pair of $A, B \in \mathcal{P}(S)$, we get a constant composition hyperoperation on $\mathcal{P}(S)$. If $R = S$, then \circ becomes a constant composition operation.

Theorem 5.2 obviously does not hold when operations $+$ or \cdot are non-idempotent. Not even one of the inclusions holds because neither $r \in [r + r]_{\leq}$ nor $r + r \in [r]_{\leq}$ in a general case. Yet for all types of hyperstructures discussed in Lemma 4.1 we might prove the following.

Theorem 5.4. *Let (R, \oplus, \bullet) be a semihyperring in the general sense constructed in Lemma 4.1 from partially ordered semigroups $(R, +, \leq)$ and (R, \cdot, \leq) . If they exist, denote e_s the neutral element of $(R, +)$ and e_p the neutral element of (R, \cdot) .*

- (1) *If simultaneously $e_p \leq e_p + e_p$ and $e_s \leq e_s \cdot e_s$, then $\circ_{\min e}$ defined by*

$$a \circ_{\min e} b = [\min\{e_s, e_p\}]_{\leq} \quad (5.4)$$

for all $a, b \in R$, is a constant composition hyperoperation on (R, \oplus, \bullet) .

- (2) *If simultaneously $e_p + e_p \leq e_p$ and $e_s \cdot e_s \leq e_s$, then $\circ_{\max e}$ defined by*

$$a \circ_{\max e} b = [\max\{e_s, e_p\}]_{\leq} \quad (5.5)$$

for all $a, b \in R$, is a constant composition hyperoperation on (R, \oplus, \bullet) .

Before proving the theorem, agree that, if the elements e_s, e_p are incomparable, then since the minimum does not exist, we set $a \circ_{\min e} b = \emptyset$. Moreover, if only e_s exists, then we set $\min\{e_s, e_p\} = e_s$ (and the same for e_p). And make the similar agreement for the maxima.

Proof. We will prove the theorem for $\circ_{\min e}$ only. The proof for $\circ_{\max e}$ is analogous.

In the \oplus, \bullet notation the left-hand-side of Definition 3.2, property 1, reads $(x \oplus y) \circ z$. This is

$$\begin{aligned} [x + y]_{\leq} \circ_{\min e} z &= \bigcup_{\substack{\text{number of elements} \\ \text{of } [x+y]_{\leq}\text{-times}}} [\min\{e_s, e_p\}]_{\leq} \\ &= [\min\{e_s, e_p\}]_{\leq} \end{aligned}$$

while the right-hand side, which reads $(x \circ z) \oplus (y \circ z)$, is

$$[\min\{e_s, e_p\}]_{\leq} + [\min\{e_s, e_p\}]_{\leq} = \bigcup_{\substack{\min\{e_s, e_p\} \leq a \\ \min\{e_s, e_p\} \leq b}} [a + b]_{\leq}.$$

Now the following cases are possible:

$e_s \leq e_p$: : This means that $\min\{e_s, e_p\} = e_s$; the left-hand side is $[e_s]_{\leq}$ while the right-hand side is $\bigcup_{\substack{e_s \leq a \\ e_s \leq b}} [a + b]_{\leq} = [e_s + e_s]_{\leq} = [e_s]_{\leq}$, i.e. the same.

$e_p < e_s$: : This means that $\min\{e_s, e_p\} = e_p$; the left-hand side is $[e_p]_{\leq}$ while the right hand side is $\bigcup_{\substack{e_p \leq a \\ e_p \leq b}} [a + b]_{\leq} = [e_p + e_p]_{\leq}$. Suppose now an arbitrary $x \in [e_p]_{\leq}$, i.e. such

$x \in R$ that $e_p \leq x$. Since we assume that $e_p < e_s$, there is also $e_p + e_p < x + e_s = x$, i.e. $x \in [e_p + e_p]_{\leq}$. If on the other hand we suppose an arbitrary $x \in [e_p + e_p]_{\leq}$, i.e. $e_p + e_p \leq x$, then on condition assumed in the theorem, i.e. $e_p \leq e_p + e_p$, there is from transitivity that $e_p \leq x$, which means that $x \in [e_p]_{\leq}$. Altogether $[e_p]_{\leq} = [e_p + e_p]_{\leq}$.

If neither e_s nor e_p exists or if e_s and e_p are incomparable, we end up with $\emptyset = \emptyset$. If only e_s exists, we get the same as when $e_s \leq e_p$. If only e_p exists, we get the same as when $e_p < e_s$.

The proof of Definition 3.2 property 2, is completely analogous. The proof of property 3 is obvious. \square

Example 5.5. Since $(\mathbb{Z}, +, \leq)$, where \leq is the natural ordering of integers, is a partially ordered group, $(\mathbb{Z}, \cdot, \leq)$ a partially ordered semigroup and $e_s = 0$, $e_p = 1$, the hyperoperation \circ defined for all $a, b \in \mathbb{Z}$ by $a \circ b = [0]_{\leq}$ is an example of a constant composition hyperoperation on the hyperring in the general sense $(\mathbb{Z}, \oplus, \bullet)$, where \oplus and \bullet are defined by (3.1) and (3.2) respectively, in a context when the single-valued operations $+$, \cdot are non-idempotent. The conditions of Theorem 5.4 obviously hold since $1 \leq 1 + 1$ and $0 \leq 0 \cdot 0$.

The constant compositions are rather trivial and degenerated cases yet even there the limits of applying the composition property in the context of the ‘‘Ends lemma’’, i.e. on hyperoperations based on the sets of the $[a]_{\leq}$ type, can be seen. It is rather difficult to achieve equality in properties 1 and 2 since the addition (or multiplication) on the left-hand side is applied on elements while on the right-hand side it is (in a general case) applied on sets – and this is done in a context where neither $a \in [a + a]_{\leq}$ nor $a + a \in [a]_{\leq}$ holds generally.

Let us therefore adjust the composition hyperoperation defined in Definition 3.2 to suit EL -hyperstructures better. In order to keep notation uniform with Definition 3.2 we use symbols $+$, \cdot for the hyperoperations even though below we are going to use Definition 5.6 only in context of hyperoperations \oplus, \bullet .

In the following definition we speak of ‘‘semihyperrings in the general sense’’. This is because they are the weakest of hyperstructures discussed in Lemma 4.1. Thus we make sure that the future considerations are valid for all types of relevant hyperstructures.

Definition 5.6. A binary operation (hyperoperation) on a semihyperring in the general sense $(R, +, \cdot)$, where $+$ and \cdot are hyperoperations on R , is called a left weak composition operation (hyperoperation) and denoted \circ_{lw} if, for all $x, y, z \in R$,

- (1) $(x + y) \circ_{lw} z \subseteq (x \circ_{lw} z) + (y \circ_{lw} z)$
- (2) $(x \cdot y) \circ_{lw} z \subseteq (x \circ_{lw} z) \cdot (y \circ_{lw} z)$
- (3) $x \circ_{lw} (y \circ_{lw} z) = (x \circ_{lw} y) \circ_{lw} z$.

or the right weak composition operation (hyperoperation) and denoted \circ_{rw} if, for all $x, y, z \in R$:

- (1) $(x \circ_{rw} z) + (y \circ_{rw} z) \subseteq (x + y) \circ_{rw} z$
- (2) $(x \circ_{rw} z) \cdot (y \circ_{rw} z) \subseteq (x \cdot y) \circ_{rw} z$
- (3) $x \circ_{rw} (y \circ_{rw} z) = (x \circ_{rw} y) \circ_{rw} z$.

The hyperstructure $(R, +, \cdot, \circ_W)$ (regardless of type) is called a weak composition hyperstructure (i.e. weak composition semihyperring / weak composition hyperring / etc.) regardless of whether $\circ_W = \circ_{lw}$ or $\circ_W = \circ_{rw}$ or whether \circ_W is single- or multi-valued.

Chvalina has in [3,4] and subsequent papers introduced and studied the concept of *quasi-order hypergroups*, which has been studied by a number of authors since. In the following theorem we not only give necessary conditions for the existence of a left (right) weak composition hyperoperation but also establish a link between quasi-order hypergroups and *EL*-hyperstructures by defining the composition hyperoperation by $a \circ b = [a]_{\leq} \cup [b]_{\leq}$ for all $a, b \in R$, i.e. by a condition used when testing whether a hypergroupoid (H, \circ) is a quasi-order hypergroup. (For details see e.g. [9], chapter 3, §1). Notice that thanks to reflexivity of relation \leq the set $[a]_{\leq} \cup [b]_{\leq}$ has for $a \neq b$ always at least two elements.

Theorem 5.7. *Let (R, \oplus, \bullet) be a semihyperring in the general sense constructed in Lemma 4.1 from quasi-ordered semigroups $(R, +, \leq)$ and (R, \cdot, \leq) . If, for all $r \in R$, there is $r + r \leq r$ and $r \cdot r \leq r$, then there exists a left weak composition hyperoperation \circ_{lw} on (R, \oplus, \bullet) .*

Proof. Define $a \circ_{lw} b = [a]_{\leq} \cup [b]_{\leq}$ for all $a, b \in R$. In this context the left-hand side of property 1 of Definition 5.6 is

$$[x + y]_{\leq} \circ_{lw} z = \bigcup_{x+y \leq a} [a]_{\leq} \cup [z]_{\leq} = [x + y]_{\leq} \cup [z]_{\leq}$$

while the right-hand side is

$$\begin{aligned} (x \circ_{lw} z) \oplus (y \circ_{lw} z) &= ([x]_{\leq} \cup [z]_{\leq}) \oplus ([y]_{\leq} \cup [z]_{\leq}) \\ &= \bigcup_{\substack{a \in [x]_{\leq} \cup [z]_{\leq} \\ b \in [y]_{\leq} \cup [z]_{\leq}}} [a + b]_{\leq}, \end{aligned}$$

i.e. $(x \circ_{lw} z) \oplus (y \circ_{lw} z) = \{d \in R; a + b \leq d; (x \leq a \text{ or } z \leq a) \text{ and } (y \leq b \text{ or } z \leq b)\}$. Suppose an arbitrary $c \in [x + y]_{\leq} \circ_{lw} z$. There are two options: $c \in [x + y]_{\leq}$ or $c \in [z]_{\leq}$. If $c \in [x + y]_{\leq}$, then obviously $c \in (x \circ_{lw} z) \oplus (y \circ_{lw} z)$ because $a \in [x]_{\leq}$, $b \in [y]_{\leq}$, i.e. $x \leq a$, $y \leq b$ implies $x + y \leq a + b$ which thanks to transitivity of \leq means that $x + y \leq c$ which is what we suppose. If $c \in [z]_{\leq}$, i.e. $z \leq c$, then if we suppose that $z + z \leq z$, we get from transitivity of \leq that $z + z \leq c$. Yet this is on the right-hand side the case of $a \in [z]_{\leq}$, $b \in [z]_{\leq}$, i.e. $z + z \leq a + b$.

The proof of property 2 is analogous, the proof of property 3 is obvious. \square

Corollary 5.8. *If $(R, +, \leq)$ and (R, \cdot, \leq) are idempotent quasi-ordered semigroups, then there always exists a left weak composition hyperoperation \circ_{lw} on (R, \oplus, \bullet) . The same holds if $r + r \leq r$ for all $r \in R$ and (R, \cdot, \leq) is an idempotent quasi-ordered semigroup or if $r \cdot r \leq r$ for all $r \in R$ and $(R, +, \leq)$ is an idempotent quasi-ordered semigroup.*

Proof. Conditions $r + r \leq r$, $r \cdot r \leq r$ included in Theorem 5.7 in this case turn into $r \leq r$. However, since the relation \leq is reflexive, they hold trivially. \square

Remark 5.9. If both $(R, +, \leq)$ and (R, \cdot, \leq) are quasi-ordered groups, then simultaneous validity of $r + r \leq r$ and $r \cdot r \leq r$ for all $r \in R$ is equivalent to the fact that $r \leq e_s$ and $r \leq e_p$, where e_s and e_p are neutral elements of $(R, +)$ and (R, \cdot) respectively. Thus e_s and e_p are the greatest elements of (R, \leq) , which means that for groups $(R, +, \leq)$ and (R, \cdot, \leq) validity of the conditions in Theorem 3 implies that $e_s = e_p$.

Theorem 5.10. *There exists a right weak composition operation \circ_{rw} on all types of hyperstructures (R, \oplus, \bullet) discussed in Lemma 4.1 which are constructed from a quasi-ordered semigroup $(R, +, \leq)$ and a commutative idempotent quasi-ordered semigroup (R, \cdot, \leq) .*

Proof. For arbitrary $A, B \subseteq R$ denote

$$A \circ_{rw} B = \{a \cdot b; a \in A, b \in B\}, \quad (5.6)$$

where \cdot is the single-valued product of (R, \cdot, \leq) . One-element sets A, B will be represented by the elements themselves, i.e. $\{a\} \circ_{rw} \{b\} = a \cdot b$, which will allow us to write

$$a \circ_{rw} b = a \cdot b \quad (5.7)$$

for all $a, b \in R$.

Now in property 1 of Definition 5.6 we get on the left-hand side, which reads $(x \circ_{rw} z) \oplus (y \circ_{rw} z)$, the set $[x \cdot z + y \cdot z]_{\leq}$ which thanks to distributivity of the single-valued structure $(R, +, \cdot)$ is $[(x + y) \cdot z]_{\leq}$. On the right-hand side, which reads $(x \oplus y) \circ_{rw} z$, we get $[x + y]_{\leq} \circ_{rw} z$, which equals $\bigcup_{x+y \leq s} \{s \cdot z\}$. Yet since the relation \leq is reflexive, there is $x + y \leq x + y$ and $[(x + y) \cdot z]_{\leq} \subseteq \bigcup_{x+y \leq s} \{s \cdot z\}$.

In property 2 of Definition 5.6 we get that (thanks to commutativity and idempotency)

$$\begin{aligned} (x \circ_{rw} z) \bullet (y \circ_{rw} z) &= (x \cdot z) \bullet (y \cdot z) = [x \cdot z \cdot y \cdot z]_{\leq} \\ &= [x \cdot y \cdot z \cdot z]_{\leq} = [x \cdot y \cdot z]_{\leq}. \end{aligned}$$

On the left-hand side we get that $[x \cdot y]_{\leq} \circ_{rw} z = \bigcup_{x \cdot y \leq r} \{r \cdot z\}$. Thus thanks to reflexivity of the relation \leq property 2 holds.

In property 3 of Definition 5.6 there is $x \circ_{rw} (y \circ_{rw} z) = x \circ_{rw} (y \cdot z) = x \cdot y \cdot z$ and $(x \circ_{rw} y) \circ_{rw} z = (x \cdot y) \circ_{rw} z = x \cdot y \cdot z$. \square

Example 5.11. If we continue with Example 4.3 and define

$$A \circ_{lw} B = [A]_{\subseteq} \cup [B]_{\subseteq} = \{R \in \mathcal{P}(S); A \subseteq R \text{ or } B \subseteq R\}$$

for all $A, B \in \mathcal{P}(S)$, then since both set intersection and set union are idempotent, the above defines a left weak composition hyperoperation on $(\mathcal{P}(S), \oplus, \bullet)$, i.e. $(\mathcal{P}(S), \oplus, \bullet, \circ_{lw})$ is a weak composition semihyperring in the general sense.

Example 5.12. If we continue with Example 4.3 and define $A \circ_{rw} B = A \cap B$ for all $A, B \in \mathcal{P}(S)$, then since the set intersection is both commutative and idempotent (and distributive with respect to set union), this defines a weak composition operation on $(\mathcal{P}(S), \oplus, \bullet)$, i.e. that $(\mathcal{P}(S), \oplus, \bullet, \cap)$ is a weak composition semihyperring in the general sense.

Examples 5.13 and 5.14 are partly motivated by the classical *interval binary hyperoperation* on a linearly ordered group discussed e.g. in [17] and defined by

$$\begin{aligned} a * b &= [\min\{a, b\}]_{\leq} \cap (\max\{a, b\}]_{\leq} \\ &= \{x \in G; \min\{a, b\} \leq x \leq \max\{a, b\}\} \end{aligned}$$

for all $a, b \in G$.

Example 5.13. Regard the ordered semiring of natural numbers, i.e. a distributive structure $(\mathbb{N}, +, \cdot)$, where $(\mathbb{N}, +)$ and (\mathbb{N}, \cdot) are semigroups and \leq is the usual ordering of natural numbers with the smallest element 1. Obviously $(\mathbb{N}, +, \leq)$ and $(\mathbb{N}, \cdot, \leq)$ are quasi-ordered semigroups, which enables us to construct semihypergroups (\mathbb{N}, \oplus) and (\mathbb{N}, \bullet) using (3.1) and (3.2) respectively. Thus we get a semihyperring in the general sense $(\mathbb{N}, \oplus, \bullet)$.

For arbitrary $a, b \in \mathbb{N}$ define

$$a \circ_{rw} b = [\max\{a, b\}]_{\leq}. \quad (5.8)$$

Obviously, the maximum always exists and $a \circ_{rw} b$ is never empty or a one-element set. In the proof of Theorem 6.2 we will show that (5.8) is a weak composition hyperoperation on $(\mathbb{N}, \oplus, \bullet)$, or rather on every set where there hold implications used in Theorem 6.2, such that it is different from the hyperoperation considered in the proof of Theorem 5.10.

Example 5.14. One can easily show that when changing in (5.8) $\max\{a, b\}$ to $\min\{a, b\}$, we get another weak composition hyperoperation on $(\mathbb{N}, \oplus, \bullet)$.

6. Existence theorems

Using Lemma 4.1, results of section 5 might be summed up as follows. Notice that definitions of composition hyperstructures are analogies of Definition 3.2, only the carrier hyperstructure is different.

Theorem 6.1. *Let $(R, +, \leq)$ and (R, \cdot, \leq) be quasi-ordered semigroups and (R, \oplus) , (R, \bullet) their associated EL-hyperstructures constructed using (3.1) and (3.2) respectively. Furthermore, let \cdot distribute over $+$ from both left and right.*

- (1) *If operations $+$ and \cdot are idempotent, then there exists a composition hyperoperation \circ on (R, \oplus, \bullet) such that $(R, \oplus, \bullet, \circ)$ is a composition semihyperring in the general sense.*
- (2) *The same holds if $(R, +)$ or (R, \cdot) are monoids with neutral elements e_s, e_p respectively, and either $e_p \leq e_p + e_p$, $e_s \leq e_s \cdot e_s$ or $e_p + e_p \leq e_p$, $e_s \cdot e_s \leq e_s$.*
- (3) *If $(R, +)$ is a group or (R, \oplus) is a hypergroup, then in 1 and 2 $(R, \oplus, \bullet, \circ)$ is a composition hyperring in the general sense.*
- (4) *If (R, \cdot) is a group, then in 1 and 2 $(R, \oplus, \bullet, \circ)$ is a composition semihyperring.*
- (5) *If $(R, +)$ is a group with neutral element 0 and $(R \setminus \{0\}, \cdot)$ is a group, then in 1 and 2 $(R, \oplus, \bullet, \circ)$ is a composition hyperring.*

Proof. Follows immediately from Lemma 4.1, Theorem 5.2 and Theorem 5.4. \square

Analogous theorems can be formulated for weak composition hyperstructures using Theorem 5.7, Corollary 5.8 or Theorem 5.10. Or – which is more important – immediately after finding suitable (weak) composition operations (hyperoperations) in some special contexts. An example of this is the following case of linearly ordered commutative semigroups used in Example 5.13 or Example 5.14.

Theorem 6.2. *Let (R, \oplus) and (R, \bullet) be two semihypergroups constructed from linearly ordered commutative semigroups $(R, +, \leq)$ and (R, \cdot, \leq) by (3.1) and (3.2) respectively. Furthermore, let \cdot distribute over $+$ from both left and right. If implications $a + a \leq b \Rightarrow a \leq b$ and $a \cdot a \leq b \Rightarrow a \leq b$ hold for all $a, b \in R$, then there exists a weak composition hyperoperation \circ_{rw} on R such that $(R, \oplus, \bullet, \circ_{rw})$ becomes a weak composition semihyperring in the general sense.*

Proof. The fact that (R, \oplus, \bullet) is a semihyperring follows from Lemma 4.1. The weak composition hyperoperation in question will be (5.8).

Suppose arbitrary $x, y, z \in R$. First we discuss the meaning of property 1 of Definition 5.6 based on definitions of \oplus and \circ_{rw} . In our notation the left-hand side reads

$(x \circ_{rw} z) \oplus (y \circ_{rw} z)$. This is

$$\begin{aligned} [\max\{x, z\}]_{\leq} \oplus [\max\{y, z\}]_{\leq} &= \bigcup_{\substack{a \in [\max\{x, z\}]_{\leq} \\ b \in [\max\{y, z\}]_{\leq}}} [a + b]_{\leq} \\ &= \bigcup_{\substack{\max\{x, z\} \leq a \\ \max\{y, z\} \leq b}} [a + b]_{\leq}, \end{aligned}$$

which results in the following four cases based on the relations between x , y and z . Notice that reasoning in cases C) and D) is analogous to reasoning in case B).

A) $x \leq z, y \leq z$: In this case $\max\{x, z\} = z$, $\max\{y, z\} = z$ and moreover $x + y \leq z + z$. Thus

$$\begin{aligned} \bigcup_{\substack{\max\{x, z\} \leq a \\ \max\{y, z\} \leq b}} [a + b]_{\leq} &= \bigcup_{\substack{z \leq a \\ z \leq b}} [a + b]_{\leq} = \\ &= \{c \in R; a + b \leq c; z \leq a, z \leq b\}. \end{aligned}$$

At the same time conditions $z \leq a$, $z \leq b$ result in $z + z \leq a + b$ and from transitivity of \leq we get that $z + z \leq c$. Finally

$$\begin{aligned} (x \circ_{rw} z) \oplus (y \circ_{rw} z) &= \{c \in R; x + y \leq c\} = \\ &= \{c \in R; z + z \leq c\}. \end{aligned} \tag{6.1}$$

B) $x \leq z, z \leq y$: In this case $\max\{x, z\} = z$, $\max\{y, z\} = y$ and moreover from transitivity of \leq there is $x \leq y$. Thus

$$\begin{aligned} \bigcup_{\substack{\max\{x, z\} \leq a \\ \max\{y, z\} \leq b}} [a + b]_{\leq} &= \bigcup_{\substack{z \leq a \\ y \leq b}} [a + b]_{\leq} = \\ &= \{c \in R; a + b \leq c; z \leq a, y \leq b\}. \end{aligned}$$

At the same time conditions $z \leq a$, $y \leq b$ result in $z + y \leq a + b$ and from transitivity of \leq we get that $z + y \leq c$. Finally

$$(x \circ_{rw} z) \oplus (y \circ_{rw} z) = \{c \in R; z + y \leq c\}. \tag{6.2}$$

C) $z \leq x, y \leq z$: This results in $(x \circ_{rw} z) \oplus (y \circ_{rw} z) = \{c \in R; x + z \leq c\}$.

D) $z \leq x, z \leq y$: This results in

$$(x \circ_{rw} z) \oplus (y \circ_{rw} z) = \{c \in R; x + y \leq c\} = \{c \in R; z + z \leq c\}$$

The right-hand side of property 1 of Definition 5.6 reads $(x \oplus y) \circ_{rw} z$. Based on definitions of \oplus and \circ_{rw} this is

$$[x + y]_{\leq} \circ_{rw} z = \bigcup_{r \in [x+y]_{\leq}} [\max\{r, z\}]_{\leq} = \bigcup_{x+y \leq r} [\max\{r, z\}]_{\leq}.$$

However, in our case this is the same as $[\max\{x + y, z\}]_{\leq}$, which is

$$\{d \in R; \max\{x + y, z\} \leq d\}. \tag{6.3}$$

Now we verify the inclusion in property 1 of Definition 5.6. Suppose an arbitrary $c \in (x \circ_{rw} z) \oplus (y \circ_{rw} z)$ and let us find out whether $c \in (x \oplus y) \circ_{rw} z$. We have to test each of the cases **A** – **D**.

ad A: The element c is such that $z + z \leq c$, $x + y \leq c$ and at the same time $x \leq z$, $y \leq z$. Thus

(1) if $\max\{x + y, z\} = x + y$, then (6.3) turns into $\{d \in R; x + y \leq d\}$. Thus $c \in (x \oplus y) \circ_{rw} z$ obviously holds.

(2) if $\max\{x + y, z\} = z$, then (6.3) turns into $\{d \in R; z \leq d\}$ and we have to show that $z \leq c$. Yet since $z + z \leq c$, there is – thanks to the assumption of the theorem – also $z \leq c$ and $c \in (x \oplus y) \circ_{rw} z$.

- ad B:** The element c is such that $z + y \leq c$ and at the same time $x \leq z$, $z \leq y$. Thus
- (1) if $\max\{x + y, z\} = x + y$, then (6.3) turns into $\{d \in R; x + y \leq d\}$. Since $x \leq z$, there is $x + y \leq z + y$ and from transitivity we get that $x + y \leq c$. Thus $c \in (x \oplus y) \circ_{rw} z$.
 - (2) if $\max\{x + y, z\} = z$, then (6.3) turns into $\{d \in R; z \leq d\}$ and we have to show that $z \leq c$. Since $z \leq y$, there is $z + z \leq z + y$ and from transitivity of \leq , there is $z + z \leq c$. Yet this means – thanks to the assumption of the theorem – that $z \leq c$ and $c \in (x \oplus y) \circ_{rw} z$.
- ad C:** The element c is such that $x + z \leq c$ and at the same time $z \leq x$, $y \leq z$. Thus
- (1) if $\max\{x + y, z\} = x + y$, then (6.3) turns into $\{d \in R; x + y \leq d\}$ and we have to show that $x + y \leq c$. Suppose on contrary that $c < x + y$. Since $y \leq z$, there is $c < x + z$. Yet since simultaneously $x + z \leq c$, we get from transitivity that $c < c$ which is impossible. Thus $x + y \leq c$ and $c \in (x \oplus y) \circ_{rw} z$.
 - (2) if $\max\{x + y, z\} = z$, then (6.3) turns into $\{d \in R; z \leq d\}$ and we have to show that $z \leq c$. Since $z \leq x$, there is $z + z \leq x + z$ and from transitivity of \leq , there is $z + z \leq c$. Yet this – thanks to the assumption of the theorem – means that $z \leq c$ and $c \in (x \oplus y) \circ_{rw} z$.
- ad D:** The element c is such that $x + y \leq c$, $z + z \leq c$ and at the same time $z \leq x$, $z \leq y$. Thus
- (1) if $\max\{x + y, z\} = x + y$, then (6.3) turns into $\{d \in R; x + y \leq d\}$ and we have to show that $x + y \leq c$. Yet this is one of our assumptions. Thus $c \in (x \oplus y) \circ_{rw} z$ holds trivially.
 - (2) if $\max\{x + y, z\} = z$, then (6.3) turns into $\{d \in R; z \leq d\}$ and we have to show that $z \leq c$. Yet since $z + z \leq c$, there is also – thanks to the assumption of the theorem – that $z \leq c$ and $c \in (x \oplus y) \circ_{rw} z$.

Thus we have verified validity of property 1 of Definition 5.6. The proof of property 2 is completely analogous.

Verifying property 3 is rather straightforward. The left-hand side $x \circ_{rw} (y \circ_{rw} z)$ is

$$\begin{aligned} x \circ_{rw} [\max\{y, z\}]_{\leq} &= \bigcup_{r \in [\max\{y, z\}]_{\leq}} [\max\{x, r\}]_{\leq} \\ &= \bigcup_{\max\{y, z\} \leq r} [\max\{x, r\}]_{\leq} \end{aligned}$$

while the right-hand side $(x \circ_{rw} y) \circ_{rw} z$ is

$$\begin{aligned} [\max\{x, y\}]_{\leq} \circ_{rw} z &= \bigcup_{s \in [\max\{x, y\}]_{\leq}} [\max\{s, z\}]_{\leq} \\ &= \bigcup_{\max\{x, y\} \leq s} [\max\{s, z\}]_{\leq}. \end{aligned}$$

Yet since the relation \leq is reflexive, i.e. $\max\{y, z\} \leq \max\{y, z\}$, $\max\{x, y\} \leq \max\{x, y\}$, both sides equal $[\max\{x, y, z\}]_{\leq}$.

Thus finally (5.8) is a weak composition hyperoperation on (R, \oplus, \bullet) with the assumed properties. \square

Remark 6.3. Notice that as regards number domains, the implications used in Theorem 6.2 which obviously hold in \mathbb{N} or \mathbb{Z} , do not hold for other number domains. The transition to \mathbb{Q} or \mathbb{R} is not possible as e.g. $0.1 \cdot 0.1 \leq 0.02$ yet $0.1 \not\leq 0.02$. Notice that if we expanded Example 5.13 to $R = \mathbb{R}^+$ or considered this in the theorem, then e.g. in case **C2** of the proof the conditions would not hold for multiplication and $x = 0.1$, $y = 0.02$, $z = 0.1$. Naturally, we could expand Theorem 6.2 by including analogies of parts 4 and 5 of Theorem 6.1.

Acknowledgment. The first author was supported by the project FEKT-S-14-2200 of Brno University of Technology. The second author acknowledges the financial support from the Slovenian Research Agency (research core funding No. P1 - 0285).

References

- [1] I. Adler, *Composition rings*, Duke Math. J. **29**, 607-623, 1962.
- [2] S. M. AnvariyeH and B. Davvaz, *Strongly transitive geometric spaces associated to hypermodules*, J. Algebra **332**, 1340-1359, 2009.
- [3] J. Chvalina, *Commutative hypergroups in the sense of Marty and ordered sets*, General Algebra and Ordered Sets, Proc. Int. Conf. Olomouc, 19-30, 1994.
- [4] J. Chvalina, *Functional Graphs, Quasi-ordered Sets and Commutative Hypergroups* (in Czech), Masaryk University, Brno, 1995.
- [5] J. Chvalina and L. Chvalinová, *Transposition hypergroups formed by transformation operators on rings of differentiable functions*, Ital. J. Pure Appl. Math. **15**, 93-106, 2004.
- [6] J. Chvalina, Š. Hořková–Mayerová and A.D. Nezhad, *General actions of hyperstructures and some applications*, An. Șt. Univ. Ovidius Constanța, **21** (1), 59-82, 2013.
- [7] A. Connes and C. Consani, *The hyperring of adèle classes*, J. Number Theory **131** (2), 159-194, 2011.
- [8] P. Corsini, *Hyperstructures associated with ordered sets*, Bull. Greek Math. Soc. **48**, 7-18, 2003.
- [9] P. Corsini and V. Leoreanu, *Applications of Hyperstructure Theory*, Kluwer Academic Publishers, Dodrecht, 2003.
- [10] I. Cristea and S. Jančić-Rašović, *Composition hyperrings*, An. Șt. Univ. Ovidius Constanța **21** (2), 81-94, 2013.
- [11] B. Davvaz and V. Leoreanu Fotea, *Applications of Hyperring Theory*, International Academic Press, Palm Harbor, 2007.
- [12] B. Davvaz, N. Rakhsh-Khorshid and K.P. Shum, *Construction of composition (m, n, k) -hyperrings*, An. Șt. Univ. Ovidius Constanța **24** (1), 177-188, 2016.
- [13] B. Davvaz and A. Salasi, *A realization of hyperrings*, Comm. Algebra **34**, 4389-4400, 2006.
- [14] A. Deghan Nezhad and B. Davvaz, *An Introduction to the Theory of H_v -Semilattices*, Bull. Malays. Math. Sci. Soc. **32** (3), 375-390, 2009.
- [15] S.H. Ghazavi and S.M. AnvariyeH, *EL -hyperstructures associated to n -ary relations*, Soft Comput. **21** (19), 5841-5850, 2017.
- [16] D. Heidari and B. Davvaz, *On ordered hyperstructures*, U.P.B. Sci. Bull. Series A, **73** (2), 2011.
- [17] K. Iwasawa, *On linearly ordered groups*, J. Math. Soc. **1** (1), 1-9, 1948.
- [18] S. Jančić-Rašović, *About the hyperring of polynomials*, Ital. J. Pure Appl. Math. **28**, 223-234, 2007.
- [19] M. Krasner, *A class of hyperrings and hyperfields*, Int. J. Math. Sci. **6** (2), 307-312, 1983.
- [20] M. Krasner, *Approximation des corps values complets de caracteristique $p \neq 0$ par ceux de caracteristique 0*, Colloque d'Algebre Superieure, CBRM, Bruxelles, 1957.
- [21] Š. Křehlík and M. Novák, *From lattices to H_v -matrices*, An. Șt. Univ. Ovidius Constanța **24** (3), 209-222, 2016.
- [22] Ch.G. Massouros, *On the theory of hyperrings and hyperfields*, Algebra i Logika **24** (6), 728-742, 1985.
- [23] G.G. Massouros, *The hyperringoid*, Multiple Valued Logic **3**, 217-234, 1998.

- [24] Ch.G. Massouros and G.G. Massouros, *On join hyperrings*, Proceedings of the 10th International Congress on Algebraic Hyperstructures and Applications, Brno, Czech Republic, 203-215, 2009.
- [25] Ch.G. Massouros and G. G. Massouros, *The transposition axiom in hypercompositional structures*, Ratio Mathematica **21**, 75-90, 2011.
- [26] S. Mirvakili and B. Davvaz, *Applications of the a^* -relation to Krasner hyperrings*, J. Algebra **362**, 145-156, 2012.
- [27] J.D. Mittas, *Sur certaines classes de structures hypercompositionnelles*, Proceedings of the Academy of Athens, **48**, 298-318, 1973.
- [28] A. Nakasis, *Recent results in hyperring and hyperfield theory*, Internat. J. of Math. & Math. Sci. **11** (2), 209-220, 1988.
- [29] M. Novák, *Some basic properties of EL-hyperstructures*, European J. Combin. **34**, 446-459, 2013.
- [30] M. Novák, *Potential of the "Ends lemma" to create ring-like hyperstructures from quasi-ordered (semi)groups*, South Bohemia Mathem. Letters **17** (1), 39-50, 2009.
- [31] M. Novák, *On EL-semihypergroups*, European J. Combin. **44** Part B, 274-286, 2015.
- [32] S. Spartalis, *A class of hyperrings*, Riv. Mat. Pura Appl. **4**, 55-64, 1989.
- [33] Th. Vougiouklis, *On some representations of hypergroups*, Annales scientifiques de l'Université de Clermont-Ferrand 2, Série Mathématiques **26**, 21-29, 1990.



Lyapunov-type inequalities for third order linear differential equations with two points boundary conditions

Mustafa Fahri Aktaş¹, Devrim Çakmak*²

¹*Gazi University, Faculty of Sciences, Department of Mathematics, 06500 Teknikokullar, Ankara, Turkey*

²*Gazi University, Faculty of Education, Department of Mathematics Education, 06500 Teknikokullar, Ankara, Turkey*

Abstract

In this paper, by using Green's functions for second order differential equations, we establish new Lyapunov-type inequalities for third order linear differential equations with two points boundary conditions. By using such inequalities, we obtain sharp lower bounds for the eigenvalues of corresponding equations.

Mathematics Subject Classification (2010). 34C10, 34B05, 34L15

Keywords. Lyapunov-type inequalities, Green's functions, two points boundary conditions

1. Introduction

In [15], Lyapunov obtained the following remarkable result: If $q \in C([0, \infty), \mathbb{R}^+)$ and $y(t)$ is a nontrivial solution of

$$y'' + q(t)y = 0 \tag{1.1}$$

with Dirichlet boundary condition

$$y(a) = y(b) = 0 \tag{1.2}$$

where $a, b \in \mathbb{R}$ with $a < b$, and $y(t) \neq 0$ for $t \in (a, b)$, then the following inequality

$$\frac{4}{b-a} \leq \int_a^b q(s) ds \tag{1.3}$$

holds. The inequality (1.3) is the best possible in the sense that if the constant 4 in the left hand side of (1.3) is replaced by any larger constant, then there exists an example of (1.1) for which (1.3) no longer holds (see [12, p. 345], [14, p. 267]). The inequality (1.3) provides a lower bound for the distance between two consecutive zeros of y . Furthermore, this result has found many applications in areas like eigenvalue problems, stability, oscillation theory, disconjugacy, etc. Since then, there have been several results to generalize the above linear equation in many directions [1–19]. Before stating many efforts, it is worth to the mention following work.

*Corresponding Author.

Email addresses: mfahri@gazi.edu.tr (M.F. Aktaş), dcakmak@gazi.edu.tr (D. Çakmak)

Received: 23.03.2017; Accepted: 23.08.2017

By using Green's function, Hartman [12] obtained the generalized inequality as follows: If $q \in C([0, \infty), \mathbb{R})$ and $y(t)$ is a nontrivial solution on (a, b) for problem (1.1)-(1.2), then

$$1 \leq \int_a^b \frac{(s-a)(b-s)}{b-a} q^+(s) ds \quad (1.4)$$

holds, where $q^+(t) = \max\{q(t), 0\}$. It is easy to see that the function $M(t) = (t-a)(b-t)$ takes the maximum value at $\frac{a+b}{2}$, i.e.

$$M(t) \leq \max_{a \leq t \leq b} M(t) = M\left(\frac{a+b}{2}\right) = \left(\frac{b-a}{2}\right)^2. \quad (1.5)$$

Thus, from (1.5), the inequality (1.4) is a natural generalization of the inequality (1.3).

In this paper, we prove new Lyapunov-type inequalities for third order linear differential equation of the form

$$y''' + q(t)y = 0, \quad (1.6)$$

where $q \in C(\mathbb{R}, \mathbb{R})$ and $y(t)$ is a real solution of (1.6) satisfying the following linearly independent two-point boundary conditions

$$\begin{cases} Y_1(y) := \gamma_{11}y(a) + \gamma_{12}y'(a) + \gamma_{13}y(b) + \gamma_{14}y'(b) = 0 \\ Y_2(y) := \gamma_{21}y(a) + \gamma_{22}y'(a) + \gamma_{23}y(b) + \gamma_{24}y'(b) = 0 \\ Y_3(y) := y''(a) + y''(b) = 0 \end{cases} \quad (1.7)$$

or

$$\begin{cases} Y_4(y) := \gamma_{11}y'(a) + \gamma_{12}y''(a) + \gamma_{13}y'(b) + \gamma_{14}y''(b) = 0 \\ Y_5(y) := \gamma_{21}y'(a) + \gamma_{22}y''(a) + \gamma_{23}y'(b) + \gamma_{24}y''(b) = 0 \\ Y_6(y) := y(a) + y(b) = 0 \end{cases} \quad (1.8)$$

where $a, b \in \mathbb{R}$ with $a < b$, and $y(t) \not\equiv 0$ for $t \in (a, b)$.

Now, we present Green's functions to be used in the proofs of our main results. Assume that $y(t)$ is a nontrivial solution of (1.1) satisfying the linearly independent two-point boundary conditions $Y_1(y) = Y_2(y) = 0$. Thus, this condition implies that, of six determinants contained in the matrix

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} \end{bmatrix}, \quad (1.9)$$

not all are zero. Therefore, either

$$\begin{vmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{vmatrix} \neq 0 \text{ or } \begin{vmatrix} \gamma_{13} & \gamma_{14} \\ \gamma_{23} & \gamma_{24} \end{vmatrix} \neq 0 \quad (1.10)$$

or else [13, p. 216]. We know that the solution of (1.1) satisfying $Y_1(y) = Y_2(y) = 0$ is given by

$$y(t) = \int_a^b G_1(t, s) y''(s) ds \quad (1.11)$$

with Green's function

$$G_1(t, s) = \begin{cases} \frac{A_1(s)B_2(t) - A_2(s)B_1(t)}{C} + t - s & ; \quad a \leq s \leq t \\ \frac{A_1(s)B_2(t) - A_2(s)B_1(t)}{C} & ; \quad t \leq s \leq b \end{cases} \quad (1.12)$$

where

$$\gamma_{11} + \gamma_{13} \neq 0, \quad (1.13)$$

$$A_i(t) = (b-t)\gamma_{i3} + \gamma_{i4}, \quad (1.14)$$

$$B_i(t) = (t - a)(\gamma_{i1} + \gamma_{i3}) - (\gamma_{i2} + (b - a)\gamma_{i3} + \gamma_{i4}) \quad (1.15)$$

for $i = 1, 2$, and

$$C = \begin{vmatrix} \gamma_{11} + \gamma_{13} & \gamma_{12} + \gamma_{13}(b - a) + \gamma_{14} \\ \gamma_{21} + \gamma_{23} & \gamma_{22} + \gamma_{23}(b - a) + \gamma_{24} \end{vmatrix} \quad (1.16)$$

(See the proof of the following Lemma 2.1 for the construction of the Green's function (1.12)). We also know that non-homogeneous linear boundary value problem $y''(t) = g(t)$ satisfying $Y_1(y) = Y_2(y) = 0$ has only the trivial solution under the condition

$$D(Y) = \begin{vmatrix} Y_1(y_1) & Y_1(y_2) \\ Y_2(y_1) & Y_2(y_2) \end{vmatrix} \neq 0, \quad (1.17)$$

where $y_1(t) = 1$ and $y_2(t) = t$ are the solutions of the corresponding homogeneous linear equation. Thus, we have the following condition

$$D(Y) = \begin{vmatrix} \gamma_{11} + \gamma_{13} & \gamma_{11}a + \gamma_{12} + \gamma_{13}b + \gamma_{14} \\ \gamma_{21} + \gamma_{23} & \gamma_{21}a + \gamma_{22} + \gamma_{23}b + \gamma_{24} \end{vmatrix} \neq 0 \quad (1.18)$$

instead of (1.17). It is clear that $D(Y) = C$. Here we note that the condition (1.18) is also valid for the problem (1.6) with the two-point boundary conditions (1.7) or (1.8). We also know that if the problem (1.1) satisfying $Y_1(y) = Y_2(y) = 0$ is well posed (if, in other words, the problem (1.1) satisfying $Y_1(y) = Y_2(y) = 0$ has only the trivial solution $y(t) \equiv 0$), then it has a unique Green's function.

It is easy to see that under the condition

$$\begin{vmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{vmatrix} = \begin{vmatrix} \gamma_{13} & \gamma_{14} \\ \gamma_{23} & \gamma_{24} \end{vmatrix}, \quad (1.19)$$

the Green's function $G_1(t, s)$ is symmetric, that is, $G_1(t, s) = G_1(s, t)$ for $t, s \in [a, b]$. Moreover, we know that this symmetry is a result of self-adjoint of the equation (1.1) satisfying $Y_1(y) = Y_2(y) = 0$ [13, p. 215]. Thus, if the condition (1.19) holds, then we have

$$y(t) = \int_a^b G(t, s) y''(s) ds, \quad (1.20)$$

where

$$G(t, s) = \begin{cases} \frac{A_1(t) B_2(s) - A_2(t) B_1(s)}{C}; & a \leq s \leq t \\ \frac{A_1(s) B_2(t) - A_2(s) B_1(t)}{C}; & t \leq s \leq b \end{cases} \quad (1.21)$$

is a symmetrized Green's function instead of (1.12). Therefore, in this paper, by using the symmetrized Green's function (1.21) for the equation (1.1) satisfying $Y_1(y) = Y_2(y) = 0$ under the condition (1.19), we prove new Lyapunov-type inequalities for third order linear differential equation (1.6) with the two-point boundary conditions (1.7) or (1.8). By using such inequalities, we obtain sharp lower bounds for the eigenvalues of corresponding equations.

2. Main results

We state some important lemmas which we will be used in the proofs of our main results. In the following first lemma, we construct Green's function for the second order nonhomogeneous differential equation

$$y'' = g(t) \quad (2.1)$$

with two-point boundary conditions $Y_1(y) = Y_2(y) = 0$.

Lemma 2.1. *If $y(t)$ is a solution of (2.1) satisfying $Y_1(y) = Y_2(y) = 0$, then the integral equation (1.11) holds.*

Proof. Integrating Eq. (2.1) from a to t to find y , we get

$$y'(t) = d_1 + \int_a^t g(s) ds \quad (2.2)$$

and

$$y(t) = d_0 + d_1(t-a) + \int_a^t (t-s)g(s) ds, \quad (2.3)$$

where d_0 and d_1 are arbitrary constants. Thus, the general solution of (2.1) is (2.3). Now, by using the boundary conditions $Y_1(y) = Y_2(y) = 0$, we can find the constants d_0 and d_1 . Thus, we have

$$d_1 = \int_a^b \frac{(\gamma_{11} + \gamma_{13})A_2(s) - (\gamma_{21} + \gamma_{23})A_1(s)}{C} g(s) ds \quad (2.4)$$

and

$$d_0 = - \int_a^b \left[\frac{(\gamma_{12} + (b-a)\gamma_{13} + \gamma_{14})[(\gamma_{11} + \gamma_{13})A_2(s) - (\gamma_{21} + \gamma_{23})A_1(s)] + CA_1(s)}{C(\gamma_{11} + \gamma_{13})} \right] g(s) ds, \quad (2.5)$$

where $A_i(t)$, $i = 1, 2$, and C are given in (1.14) and (1.16), respectively. Substituting the constants d_0 and d_1 in the general solution (2.3), we get

$$y(t) = \int_a^t \left[\frac{A_1(s)B_2(t) - A_2(s)B_1(t)}{C} + t - s \right] g(s) ds + \int_t^b \frac{A_1(s)B_2(t) - A_2(s)B_1(t)}{C} g(s) ds. \quad (2.6)$$

This completes the proof. \square

Lemma 2.2. *Let (1.19) hold. If $y(t)$ is a solution of (1.6) satisfying the two-point boundary conditions (1.8), then the following inequality*

$$|y(t)| \leq \int_a^b G(s) |y'''(s)| ds \quad (2.7)$$

holds, where

$$G(t) = \frac{1}{2} \int_a^b |G(u, t)| du \quad (2.8)$$

and $G(t, s)$ is given in (1.21).

Proof. Assume that $y(t)$ is a solution of (1.6) satisfying $Y_4(y) = Y_5(y) = Y_6(y) = 0$. It is easy to see that, by using $Y_4(y) = Y_5(y) = 0$ and proceeding as in the proof of Lemma 2.1, we have

$$y'(t) = \int_a^b G(t, s) y'''(s) ds, \quad (2.9)$$

where $G(t, s)$ is the Green's function (1.21). Integrating (2.9) from a to t , we get

$$y(t) = y(a) + \int_a^t \left(\int_a^b G(u, s) y'''(s) ds \right) du. \quad (2.10)$$

Similarly, integrating (2.9) from t to b , we get

$$y(t) = y(b) + \int_t^b \left(- \int_a^b G(u, s) y'''(s) ds \right) du. \quad (2.11)$$

Adding (2.10) and (2.11), and by using $Y_6(y) = 0$, we have

$$y(t) = \frac{1}{2} \left\{ \int_a^t \left(\int_a^b G(u, s) y'''(s) ds \right) du + \int_t^b \left(- \int_a^b G(u, s) y'''(s) ds \right) du \right\}. \quad (2.12)$$

By taking the absolute value of (2.12), we obtain

$$|y(t)| \leq \frac{1}{2} \int_a^b \left(\int_a^b |G(u, s)| |y'''(s)| ds \right) du \quad (2.13)$$

and hence

$$|y(t)| \leq \frac{1}{2} \int_a^b |y'''(s)| \left(\int_a^b |G(u, s)| du \right) ds, \quad (2.14)$$

where

$$G(u, s) = \begin{cases} \frac{A_1(u) B_2(s) - A_2(u) B_1(s)}{C}; & s \leq u \leq b \\ \frac{A_1(s) B_2(u) - A_2(s) B_1(u)}{C}; & a \leq u \leq s. \end{cases} \quad (2.15)$$

Therefore, we have the inequality (2.7). This completes the proof. \square

By using the inequality (2.7), we have the following result which is an useful tool to determine a lower bound of distance between a and b points of solution of the equation (1.6) under the boundary conditions (1.8).

Theorem 2.3. *Let (1.19) hold. If $y(t)$ is a nontrivial solution of (1.6) satisfying the two-point boundary conditions (1.8), then the following Lyapunov-type inequality*

$$1 \leq \int_a^b G(s) |q(s)| ds \quad (2.16)$$

holds, where $G(t)$ is given in (2.8).

Proof. Assume that $y(t)$ is a solution of (1.6) satisfying $Y_4(y) = Y_5(y) = Y_6(y) = 0$ and y is not identically zero on (a, b) . From (1.6) and (2.7), we get

$$|y'''(t)| = |q(t)| |y(t)| \leq |q(t)| \int_a^b G(s) |y'''(s)| ds. \quad (2.17)$$

Multiplying both sides of (2.17) by $G(t)$ and integrating from a to b , we get

$$\int_a^b G(s) |y'''(s)| ds \leq \int_a^b G(s) |y'''(s)| ds \int_a^b G(s) |q(s)| ds. \quad (2.18)$$

Next, we prove that

$$0 < \int_a^b G(s) |y'''(s)| ds. \quad (2.19)$$

If (2.19) is not true, then we have

$$\int_a^b G(s) |y'''(s)| ds = 0. \quad (2.20)$$

From (2.7), we get

$$|y(t)| \leq \int_a^b G(s) |y'''(s)| ds = 0. \quad (2.21)$$

It follows from (2.21) that $y(t) \equiv 0$ for $t \in (a, b)$, which contradicts with (1.8) since $y(t) \neq 0$ for all $t \in (a, b)$. Thus, by using (2.19) in (2.18), we get the inequality (2.16). \square

Now, we give another main result for the equation (1.6) under the boundary conditions (1.7).

Theorem 2.4. *Let (1.19) hold. If $y(t)$ is a nontrivial solution of (1.6) satisfying the two-point boundary conditions (1.7), then the following Lyapunov-type inequality*

$$1 \leq G_0 \int_a^b |q(s)| ds \quad (2.22)$$

holds, where $G_0 = \frac{1}{2} \int_a^b |G(t_0, s)| ds$ and $|y(t_0)| = \max\{|y(t)| : a \leq t \leq b\}$.

Proof. Assume that $y(t)$ is a solution of (1.6) satisfying $Y_1(y) = Y_2(y) = Y_3(y) = 0$ and y is not identically zero on (a, b) . By integrating $y'''(t)$ from a to t , we get

$$y''(t) = y''(a) + \int_a^t y'''(s) ds. \quad (2.23)$$

Similarly, by integrating $y'''(t)$ from t to b , we have

$$y''(t) = y''(b) - \int_t^b y'''(s) ds. \quad (2.24)$$

Adding the inequalities (2.23) and (2.24), and by using $Y_3(y) = 0$, we have

$$2y''(t) = \int_a^t y'''(s) ds - \int_t^b y'''(s) ds. \quad (2.25)$$

By taking the absolute value of (2.25), we obtain

$$|y''(t)| \leq \frac{1}{2} \int_a^b |y'''(s)| ds. \quad (2.26)$$

Next, pick $t_0 \in (a, b)$ so that $|y(t_0)| = \max\{|y(t)| : a \leq t \leq b\}$. From (1.20), (2.26), and (1.6), we get

$$\begin{aligned} |y(t_0)| &\leq \int_a^b |G(t_0, s)| |y''(s)| ds \\ &\leq \frac{1}{2} \int_a^b |G(t_0, s)| ds \int_a^b |y'''(s)| ds \end{aligned} \quad (2.27)$$

$$\begin{aligned} &= G_0 \int_a^b |q(s)| |y(s)| ds \\ &\leq G_0 |y(t_0)| \int_a^b |q(s)| ds. \end{aligned} \quad (2.28)$$

Dividing both sides by $|y(t_0)|$, we get the inequality (2.22). \square

Remark 2.5. To the best of our knowledge, the inequality (2.16) (or (2.22)) is new Lyapunov-type inequality for third order linear differential equation (1.6) under the two-point boundary conditions (1.8) (or (1.7)).

It is easy to see that since

$$G(t) = \frac{1}{2} \int_a^b |G(u, t)| du \leq \frac{\tilde{C}(t)}{2|C|}, \quad (2.29)$$

where

$$\begin{aligned} \tilde{C}(t) &= \tilde{A}_1(t) \int_a^t \tilde{B}_2(u) du + \tilde{A}_2(t) \int_a^t \tilde{B}_1(u) du \\ &\quad + \tilde{B}_2(t) \int_t^b \tilde{A}_1(u) du + \tilde{B}_1(t) \int_t^b \tilde{A}_2(u) du, \end{aligned} \quad (2.30)$$

$$\tilde{A}_i(t) = (b-t)|\gamma_{i3}| + |\gamma_{i4}|, \quad (2.31)$$

and

$$\tilde{B}_i(t) = (|\gamma_{i1}| + |\gamma_{i3}|)(t-a) + |\gamma_{i2}| + |\gamma_{i3}|(b-a) + |\gamma_{i4}| \quad (2.32)$$

for $i = 1, 2$, we have the following result from Theorem 2.3 and hence the proof is omitted.

Corollary 2.6. *Let (1.19) hold. If $y(t)$ is a nontrivial solution of (1.6) satisfying the two-point boundary conditions (1.8), then the following Lyapunov-type inequality*

$$1 \leq \int_a^b \frac{\tilde{C}(s)}{2|C|} |q(s)| ds \quad (2.33)$$

holds, where C and $\tilde{C}(t)$ are given in (1.16) and (2.30), respectively.

Remark 2.7. Note that if we take $\gamma_{11} = \gamma_{13} = \gamma_{23} = 1$, $\gamma_{21} = -1$, and $\gamma_{i2} = \gamma_{i4} = 0$ for $i = 1, 2$ in (1.8), we have

$$8 \leq \int_a^b (b-s)(b-4a+3s)|q(s)| ds \quad (2.34)$$

from (2.33), and hence

$$\frac{6}{(b-a)^2} \leq \int_a^b |q(s)| ds. \quad (2.35)$$

Now, we give another result for the equation (1.6) by using the following inequality

$$|G(t, s)| \leq \frac{\tilde{A}_1(s)\tilde{B}_2(s) + \tilde{A}_2(s)\tilde{B}_1(s)}{|C|} \quad (2.36)$$

obtained by taking the absolute value of (1.21). Thus, we have the following result from Theorem 2.4 and hence the proof is omitted.

Corollary 2.8. *Let (1.19) hold. If $y(t)$ is a nontrivial solution of (1.6) satisfying the two-point boundary conditions (1.7), then the following Lyapunov-type inequality*

$$1 \leq \int_a^b \frac{\tilde{A}_1(s)\tilde{B}_2(s) + \tilde{A}_2(s)\tilde{B}_1(s)}{2|C|} ds \int_a^b |q(s)| ds \quad (2.37)$$

holds, where C , $\tilde{A}_i(t)$, $\tilde{B}_i(t)$, $i = 1, 2$, are given in (1.16), (2.31), (2.32), respectively.

Now, we give an application of the obtained Lyapunov-type inequalities for the following eigenvalue problem

$$y''' + \lambda k(t)y = 0 \quad (2.38)$$

under the boundary conditions (1.7). Thus, if there exists a nontrivial solution $y(t)$ of linear homogeneous problem (2.38), then we have

$$\frac{2|C|}{\int_a^b \left(\tilde{A}_1(s)\tilde{B}_2(s) + \tilde{A}_2(s)\tilde{B}_1(s) \right) ds \int_a^b |k(s)| ds} \leq |\lambda|, \quad (2.39)$$

where C , $\tilde{A}_i(t)$, $\tilde{B}_i(t)$, $i = 1, 2$, are given in (1.16), (2.31), (2.32), respectively.

References

- [1] M.F. Aktaş, *Lyapunov-type inequalities for n -dimensional quasilinear systems*, Electron. J. Differential Equations **67**, 1-8, 2013.
- [2] M.F. Aktaş, D. Çakmak and A. Tiryaki, *A note on Tang and He's paper*, Appl. Math. Comput. **218**, 4867-4871, 2012.
- [3] M.F. Aktaş, D. Çakmak and A. Tiryaki, *Lyapunov-type inequality for quasilinear systems with anti-periodic boundary conditions*, J. Math. Inequal. **8**, 313-320, 2014.
- [4] M.F. Aktaş, D. Çakmak and A. Tiryaki, *On the Lyapunov-type inequalities of a three-point boundary value problem for third order linear differential equations*, Appl. Math. Lett. **45**, 1-6, 2015.
- [5] G. Borg, *On a Liapounoff criterion of stability*, Amer. J. Math. **71**, 67-70, 1949.
- [6] A. Cabada, J.A. Cid and B. Maquez-Villamarin, *Computation of Green's functions for boundary value problems with Mathematica*, Appl. Math. Comput. **219**, 1919-1936, 2012.
- [7] D. Çakmak, *Lyapunov-type integral inequalities for certain higher order differential equations*, Appl. Math. Comput. **216**, 368-373, 2010.
- [8] D. Çakmak, *On Lyapunov-type inequality for a class of nonlinear systems*, Math. Inequal. Appl. **16**, 101-108, 2013.
- [9] D. Çakmak, M.F. Aktaş and A. Tiryaki, *Lyapunov-type inequalities for nonlinear systems involving the (p_1, p_2, \dots, p_n) -Laplacian*, Electron. J. Differential Equations **128**, 1-10, 2013.
- [10] D. Çakmak and A. Tiryaki, *On Lyapunov-type inequality for quasilinear systems*, Appl. Math. Comput. **216**, 3584-3591, 2010.
- [11] D. Çakmak and A. Tiryaki, *Lyapunov-type inequality for a class of Dirichlet quasilinear systems involving the (p_1, p_2, \dots, p_n) -Laplacian*, J. Math. Anal. Appl. **369**, 76-81, 2010.
- [12] P. Hartman, *Ordinary Differential Equations*, Wiley, New York, 1964 and Birkhäuser, Boston 1982.
- [13] E.L. Ince, *Ordinary Differential Equations*, Dover Publications, New York, 1926.
- [14] W.G. Kelley and A.C. Peterson, *The Theory of Differential Equations, Classical and Qualitative*, Springer, New York, 2010.
- [15] A. Liapunov, *Probleme general de la stabilite du mouvement*, Annales de la Faculte des Sciences de Toulouse pour les Sciences Mathematiques et les Sciences Physiques **2**, 203-474, 1907.
- [16] A. Tiryaki, D. Çakmak and M.F. Aktaş, *Lyapunov-type inequalities for a certain class of nonlinear systems*, Comput. Math. Appl. **64**, 1804-1811, 2012.
- [17] A. Tiryaki, D. Çakmak and M.F. Aktaş, *Lyapunov-type inequalities for two classes of Dirichlet quasilinear systems*, Math. Inequal. Appl. **17**, 843-863, 2014.
- [18] A. Tiryaki, M. Ünal and D. Çakmak, *Lyapunov-type inequalities for nonlinear systems*, J. Math. Anal. Appl. **332**, 497-511, 2007.
- [19] X. Yang, *On inequalities of Lyapunov type*, Appl. Math. Comput. **134**, 293-300, 2003.



On QF rings and artinian principal ideal rings

Alejandro Alvarado-García¹, César Cejudo-Castilla², Hugo Alberto Rincón-Mejía^{*1}, Ivan Fernando Vilchis-Montalvo², Manuel Gerardo Zorrilla-Noriega¹

¹*Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito Exterior, C.U., C.P. 04510, Ciudad de México, México.*

²*Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y 18 Sur, Col. San Manuel, Ciudad Universitaria, C.P. 72570, Puebla, México*

Abstract

In this work we give sufficient conditions for a ring R to be quasi-Frobenius, such as R being left artinian and the class of injective cogenerators of R -Mod being closed under projective covers. We prove that R is a division ring if and only if R is a domain and the class of left free R -modules is closed under injective hulls. We obtain some characterizations of artinian principal ideal rings. We characterize the rings for which left cyclic modules coincide with left cocyclic R -modules. Finally, we obtain characterizations of left artinian and left coartinian rings.

Mathematics Subject Classification (2010). 16L60, 16D80

Keywords. left artinian ring, artinian principal ideal ring, conoetherian ring, coartinian ring, QF ring, perfect ring, semiartinian ring

1. Introduction

In [1] and [3] the authors obtained characterizations of artinian principal ideal rings using big lattices of classes of modules closed under certain closure properties. Also, in [4] the authors deal with rings over which all injective hulls of left simple modules are noetherian. These rings are called left coartinian rings. In this work we further investigate these notions, among others.

In the sequel, R denotes an associative ring with identity and R -Mod denotes the category of left unitary R -modules, to which all “modules” and “ R -modules” will belong, unless otherwise specified. A *left uniserial* ring will be a ring whose left ideals are linearly ordered. By “QF” we mean “quasi-Frobenius”. Also, “ $N \leq_e M$ ” will stand for “ N is essential in M ”.

*Corresponding Author.

Email addresses: alejandroalvaradogarcia@gmail.com (A. Alvarado-García), cesarcc@ciencias.unam.mx (C. Cejudo-Castilla), hurincon@gmail.com (H.A. Rincón-Mejía), vilchis.f@gmail.com (I.F. Vilchis-Montalvo), mgzn@hotmail.com (M.G. Zorrilla-Noriega)

Received: 08.04.2017; Accepted: 10.09.2017

2. Artinian principal ideal rings

Recall that an artinian principal ideal ring is a left and right artinian, left and right principal ideal ring.

Definition 2.1. We will say that ${}_R M$ is a *paraprojective* module if ${}_R M$ embeds in ${}_R N$ whenever ${}_R M$ is an epimorphic image of ${}_R N$. Dually, we say that ${}_R M$ is a *parainjective* module if ${}_R M$ is quotient of each module ${}_R N$ in which ${}_R M$ embeds.

Theorem 2.2. *The following statements are equivalent for a ring R .*

- 1) R is an artinian principal ideal ring.
- 2) Every class of R -modules closed under submodules and direct sums is also closed under quotients.
- 3) Every class of R -modules closed under quotients and direct products is also closed under submodules.
- 4) There exists an epimorphism $P \twoheadrightarrow M$ precisely when there exists a monomorphism $M \hookrightarrow P$ for each R -module M and for each projective R -module P .
- 5) R is left noetherian, and every cyclic R -module C is parainjective.

Proof. 1) \Rightarrow 2), 3), 4) and 5) They follow from [3, Theorem 38].

4) \Rightarrow 1) For each module M there exists an epimorphism $R^{(X)} \twoheadrightarrow M$ for some set X . Hence, by hypothesis there exists a monomorphism $M \hookrightarrow R^{(X)}$. Therefore, by [6, Corollary 24.15], R is a QF ring. Now, let us take a left ideal I of R . By hypothesis, there exists an epimorphism $R \twoheadrightarrow I$. Thus, $I = Rx$ for some $x \in I$. Then, R is a left principal ideal ring. Therefore, by [5, Sec. 4, Theorem 1], R is an artinian principal ideal ring.

2) \Rightarrow 1) Consider the class of modules

$$\mathcal{C} = \{M \mid \text{there exists a monomorphism } M \hookrightarrow R^{(X)} \text{ for some set } X\}.$$

It is clear that \mathcal{C} is closed under submodules and direct sums. Then, by hypothesis, \mathcal{C} is closed under quotients. Also, $R^{(X)} \in \mathcal{C}$ for each set X , so $\mathcal{C} = R\text{-Mod}$. Then for each module M , there exists a monomorphism $M \hookrightarrow R^{(X)}$ for some set X , so, by [6, Corollary 24.15], R is a QF ring. Let I be any two sided ideal of R . It is straightforward to verify that the ring R/I also satisfies 2). It follows that R/I is QF. By [6, P. 217], R is an artinian principal ideal ring.

3) \Rightarrow 1) Let E be a minimal injective cogenerator. Consider the class of modules $\mathcal{C} = \{M \mid \text{there exists an epimorphism } E^X \twoheadrightarrow M \text{ for some set } X\}$. It is clear that \mathcal{C} is closed under quotients and direct products. By hypothesis, \mathcal{C} is then closed under submodules. Of course, $E^X \in \mathcal{C}$ for each set X . But, for each module M , there exists a monomorphism $M \hookrightarrow E^X$ for some set X . Then, $\mathcal{C} = R\text{-Mod}$. So, for each projective module P , there exists an epimorphism $E^X \twoheadrightarrow P$, so that P is a direct summand of E^X . Therefore, P is an injective module. Thus, R is a QF ring. Moreover, as the ring R/I also satisfies 3), R/I is a QF ring for each two sided ideal I of R . Then, by [6, P. 217], R is an artinian principal ideal ring.

5) \Rightarrow 1) As there exists a monomorphism $R \hookrightarrow E(R)$, by hypothesis there exists an epimorphism $E(R) \twoheadrightarrow R$, so R is left self-injective and left noetherian. Therefore, R is a QF ring. As the ring R/I also satisfies 5), R/I is a QF ring for each two sided ideal I of R . Then, R is an artinian principal ideal ring. \square

Theorem 2.3. *Let R be an artinian principal ideal ring. Then the following conditions are equivalent for an R -module M .*

- 1) M is finitely generated.
- 2) M is finitely cogenerated.
- 3) M is artinian.
- 4) M is noetherian.

Proof. 4) \Rightarrow 1) and 3) \Rightarrow 2) They are clear.

1) \Rightarrow 4) This is true over every left noetherian ring.

2) \Rightarrow 3) Consider the class of modules $\mathcal{C} = \{M \mid M \text{ is finitely cogenerated}\}$, which is closed under submodules. By hypothesis and [3, Theorem 38] the class \mathcal{C} is closed under quotients, so M is artinian for each $M \in \mathcal{C}$.

4) \Rightarrow 2) By hypothesis R is left semiartinian, so that $\text{soc}(M) \leq_e M$ for each non zero module M . If M is noetherian, then $\text{soc}(M)$ is finitely generated, so M is finitely cogenerated.

3) \Rightarrow 1) For this part, we will use freely [3, Theorem 38]. If M is artinian, then M is finitely cogenerated. Thus $\text{soc}(M) \leq_e M$ and $\text{soc}(M)$ is finitely generated. Therefore $E(\text{soc}(M)) = E(M)$ and $\text{soc}(M) = \bigoplus_{i=1}^n S_i$ with S_i a simple module for each $i \in \{1, \dots, n\}$.

Thus, $E(M) = E(\text{soc}(M)) = E(\bigoplus_{i=1}^n S_i) = \bigoplus_{i=1}^n E(S_i)$. We claim that each $E(S_i)$ is cyclic. Indeed, take any simple S . By hypothesis, there exists a monomorphism $S \hookrightarrow R$. Moreover, since every artinian principal ideal ring is QF and thus left self-injective, there exists a monomorphism $E(S) \hookrightarrow R$. Then there exists an epimorphism $R \twoheadrightarrow E(S)$. Therefore, $E(S)$ is cyclic, as we claim. Thus $E(M)$ is finitely generated, and as the class of finitely generated modules is closed under submodules by hypothesis, M is finitely generated. \square

Recall that an R -module M is cocyclic if M contains an essential simple submodule.

Theorem 2.4. *The classes of non-zero cyclic R -modules and of cocyclic R -modules coincide if and only if R is a left uniserial artinian principal ideal ring.*

Proof. \Rightarrow] Let us first prove that R must be left artinian. This is equivalent to every quotient of ${}_R R$ being finitely cogenerated. By the hypothesis, all we need to prove is that every cocyclic module is finitely cogenerated. Take then any cocyclic M . There is some simple $S \leq_e M$. It follows that $\text{soc}(M) = S$. Thus, M has a finitely generated essential socle, a condition well-known to be equivalent to M being finitely cogenerated.

We now show that R is left self-injective. Suppose otherwise, that is, that $R \not\leq E(R)$. The hypothesis gives that ${}_R R$ is cocyclic, so there is some simple $S \leq_e R$. Hence, $E(R) = E(S)$, which is obviously cocyclic. Using the hypothesis, we get that $E(R)$ is cyclic, so that there is an epimorphism $R \xrightarrow{f} E(R)$. Consider the following commutative diagram, where i and j are inclusion maps.

$$\begin{array}{ccc} f^{-1}(R) & \xhookrightarrow{i} & R \\ f|_{f^{-1}(R)} \downarrow & = & \downarrow f \\ R & \xrightarrow{j} & E(R) \\ & \neq & \end{array}$$

Note that if i were surjective, j would also be so. Thus, $f^{-1}(R) \leq R$. Now we may construct another level of the diagram. Let us write $f|_i$ for appropriate restrictions of f .

$$\begin{array}{ccc} f^{-1}(f^{-1}(R)) & \xhookrightarrow{\quad} & f^{-1}(R) \\ f|_i \downarrow & = & \downarrow f|_i \\ f^{-1}(R) & \xhookrightarrow{i} & R \\ f|_i \downarrow & = & \downarrow f \\ R & \xrightarrow{j} & E(R) \\ & \neq & \end{array}$$

As above, the newest inclusion must be proper. Continuing in this manner, we obtain an infinite descending chain $R \supsetneq f^{-1}(R) \supsetneq f^{-1}(f^{-1}(R)) \supsetneq \dots$, contradicting that ${}_R R$ is artinian.

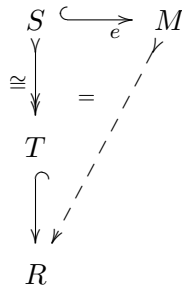
Now, R being left artinian and left self-injective is equivalent to R being QF. Take any two-sided ideal I . It is straightforward to verify (using that cocyclic modules are precisely those modules having simple essential submodules) that the ring R/I satisfies the hypothesis, i.e. that an arbitrary R/I -module is cyclic if and only if it is cocyclic. It follows that for each two-sided ideal I of R , R/I is QF. But, according to [6, P. 217], this is equivalent to R being an artinian principal ideal ring.

(One can prove directly that R is a left principal ideal ring. Indeed, take some non-zero left ideal I . By the hypothesis, it suffices to show that I is cocyclic. Note that the hypothesis gives some simple $S \leq_e R$. Also, as ${}_R R$ is artinian, there is some simple $T \leq I$. But then $T \leq R$, so necessarily $S = T$. And of course, $S \leq I \leq R$ implies that $S \leq_e I$. This establishes that R is a left principal ideal ring. As we have already shown R to be QF, [5, Sec. 4, Theorem 1] grants that R is an artinian principal ideal ring.)

As every nonzero quotient of R is cocyclic, then every nonzero quotient of R is uniform. Therefore by [7, Proposition 2.7] R is left uniserial.

⇐] Take any non-zero cyclic module, say R/I for some left ideal $I \leq R$. The submodule lattice of R/I is isomorphic to $[I, R]$, which is a chain. As ${}_R R$ is artinian, then there exists I' minimal such that $I \leq I' \leq R$. Linearity ensures that I' is essential in $[I, R]$. Therefore, I'/I is an essential simple submodule of R/I , proving its cocyclicity.

Conversely, let M be a cocyclic module. There is some simple $S \leq_e M$. Also, since ${}_R R$ is artinian, there is some simple $T \leq R$. From the hypothesis on linearity it follows that R must be local and thus left local, so that $S \cong T$. Now, any artinian principal ideal ring is in particular QF and then in particular left self-injective, so we may extend $S \cong T \hookrightarrow R$ to a mapping $M \rightarrow R$, which is monic due to the fact that $S \leq_e M$. The situation is depicted below.



Thus, M is isomorphic to some left ideal of R , which by hypothesis is principal, i.e., cyclic. □

Lemma 2.5. *If every semisimple R -module M is parainjective and paraprojective, then $R = R_1 \times R_2$, where R_1 is a semisimple ring and R_2 is a finite direct product of left local left artinian rings with all simple modules singular.*

Proof. By [2, Theorem 4.7] R is a finite direct product of left local, left and right perfect rings. Thus, R is a left semiartinian ring. Let Rx be a cyclic module. Then $\text{soc}(Rx) \leq_e Rx$ and there exists an epimorphism $Rx \twoheadrightarrow \text{soc}(Rx)$ by hypothesis, so $\text{soc}(Rx)$ is finitely generated. Therefore, Rx is a finitely cogenerated module. Then R is a left artinian ring.

Now, if M is a projective semisimple R -module, there exists an epimorphism $E(M) \twoheadrightarrow M$ by hypothesis. Then M is injective. Analogously, if M is a semisimple injective R -module, M is projective. Thus, M is projective if and only if M is injective, for each semisimple R -module M .

Write $R = R_1 \times \dots \times R_n$, where each R_i is a left local, left and right perfect ring. Let $1 \leq i \leq n$. Note that R_i is left artinian (either because R is left artinian or because

R_i satisfies the hypothesis on R), that $\text{soc}(R_i) \leq_e R_i$ (R_i being left semiartinian), that $\text{soc}(R_i)$ is a direct sum of copies of S_i (where $R_i\text{-simp} = \{S_i\}$) and is precisely the S_i -socle of R , and that for each $M \in R_i\text{-Mod}$, M is projective (respectively, injective) if and only if it is a projective (respectively, injective) R -module. We claim that S_i is projective if and only if R_i is a semisimple ring. Sufficiency is clear, because over any semisimple ring every module is projective. Conversely, suppose that S_i is projective. Then so is $\text{soc}(R_i)$, which is thus injective. but this makes $\text{soc}(R_i)$ an essential direct summand of R_i , whence $\text{soc}(R_i) = R_i$. This proves the claim.

Set $I_{SS} = \{i \in \{1, \dots, n\} \mid R_i \text{ is semisimple}\}$, and put $R_I = \prod_{i \in I_{SS}} R_i$ (which is the projective socle of R), and $R_{II} = \prod_{i \in \{1, \dots, n\} \setminus I_{SS}} R_i$. Then $R = R_I \times R_{II}$, R_I is a semisimple ring and R_{II} is a finite direct product of left local artinian rings over each of which all simple modules are singular. \square

Observe that the hypothesis of Lema 2.5 holds also for R/I for each two sided ideal I of R .

Recall that a ring R is called left quasi-duo if each maximal left ideal is two sided.

Remark 2.6. For a ring R the following conditions are equivalent.

- (1) R is a left quasi-duo ring.
- (2) For each simple R -module S , and for all $x \in S$, $(0 : x)$ is a two-sided ideal of R .

Theorem 2.7. For a left quasi-duo ring R , if every semisimple R -module is parainjective and paraprojective, then R is a finite direct product of left local left artinian rings and for each left ideal I of the factor ring R_i , $I = \text{rad}(R_i)^m$ for some $m \in \mathbb{N}$.

Proof. Consider the decomposition supplied by [2, Theorem 4.7]. Let R be a factor ring. Note that R inherits the current hypotheses. Then, there exists an epimorphism $R \rightarrow \text{soc}(R)$, so $Rx = \text{soc}(R) = S_1 \oplus S_2 \oplus \dots \oplus S_n$, where S_i is a simple module $\forall i \in \{1, \dots, n\}$. Write $x = x_1 + x_2 + \dots + x_n$, where each $x_i \in S_i \setminus \{0\}$. It is clear that $(0 : x) \subseteq (0 : x_i), \forall i \in \{1, \dots, n\}$. Let $j \in \{2, \dots, n\}$. Since R is left local, there is an isomorphism $f_j : S_1 \rightarrow S_j$. As $S_j = Rf_j(x_1)$, there is $r_j \in R$ such that $x_j = r_j f_j(x_1)$. Then, $x = x_1 + r_2 f_2(x_1) + \dots + r_n f_n(x_1)$. Note that, for $2 \leq j \leq n$, $(0 : x_1)Rf_j(x_1) = 0$ because $(0 : x_1)$ is a two sided ideal by hypothesis and Remark 2.6. Thus, $(0 : x) = (0 : x_1)$. Then, $Rx \cong Rx_1$. Therefore, $Rx = \text{soc}(R)$ is a simple module.

As established in the proof of Lemma 2.5, R is a left artinian ring, so that $\text{rad}(R)$ is nilpotent. Let us prove by induction on the nilpotency index that for each left ideal I of R , $I = \text{rad}(R)^m$ for some $m \in \mathbb{N}$. If $n = 1$, then $\text{rad}(R) = 0$. Since R is left artinian, it is semilocal, so in this case it is semisimple and thus a division ring, so that the only two left ideals are $0 = \text{rad}(R)$ and $R = \text{rad}(R)^0$. Let us suppose that $n > 1$ is the nilpotency index. As $\text{rad}(R)^n = 0$, $\text{rad}(R)^{n-1} \neq 0$ is annihilated by $\text{rad}(R)$, so $\text{rad}(R)^{n-1}$ is a semisimple module (again by semilocality). Then, $\text{rad}(R)^{n-1} \subseteq \text{soc}(R)$, but $\text{soc}(R)$ is a simple module, so that $\text{rad}(R)^{n-1} = \text{soc}(R)$. Let I be a left ideal of R . Note that, R being left artinian and having a simple socle, $\text{soc}(R) \leq I$. Then $R/\text{soc}(R) = R/(\text{rad}(R)^{n-1})$ is a ring with the same hypothesis of R whose radical has nilpotency index $n - 1$. Therefore, $I/\text{soc}(R) = \text{rad}(R)^m/\text{soc}(R)$ and by the Correspondence Theorem $I = \text{rad}(R)^m$, for some $m \in \mathbb{N}$. \square

Theorem 2.8. For a commutative ring R the following statements are equivalent.

- (1) Every semisimple R -module is parainjective and paraprojective.
- (2) R is a finite direct product of uniserial artinian principal ideal rings.

Proof. 1) \Rightarrow 2) Let R be any of the factor rings in the decomposition supplied by Theorem 2.7. We know that, for each ideal I of R , there exists $m \in \mathbb{N}$ such that $I = \text{rad}(R)^m$. In

the proof of Theorem 2.7, we showed that if $\text{rad}(R) = 0$, then R must be a division ring, and thus in particular an uniserial artinian principal ideal ring. Suppose then that there is $x \in \text{rad}(R)$ but $x \notin \text{rad}(R)^2$. Then $\text{rad}(R)^2 < Rx \leq \text{rad}(R)$. Thus, $\text{rad}(R) = Rx$. Then by an induction argument $\text{rad}(R)^m = Rx^m, \forall m \in \mathbb{N}$. Now we prove that R is a self-injective ring. Consider the following diagram:

$$\begin{array}{ccc} Rx^n & \xhookrightarrow{i} & R \\ \downarrow f & & \\ R & & \end{array}$$

where $f : Rx^n \rightarrow R$ is any homomorphism and $i : Rx^n \hookrightarrow R$ is the inclusion. Then $f(Rx^n) = Rx^m$ with $m \geq n$. Put $f(x^n) = rx^m$. Consider the homomorphism $h : R \rightarrow R$ such that $h(s) = s(rx^{m-n}), \forall s \in R$. Thus, for $t \in R$, $(hi)(tx^n) = h(i(tx^n)) = h(tx^n) = tx^n(rx^{m-n}) = trx^m = f(tx^n)$. Therefore, R is self-injective. As was established in the proof of Lemma 2.5, R is artinian. Then R is a QF-ring. Thus, by [5, Sec. 4, Theorem 1], R is an artinian principal ideal ring.

2) \Rightarrow 1) Follows by [3, Theorem 38]. \square

3. Coartinian, conoetherian and quasi-Frobenius rings

The ring R is said to be left *coartinian* if for every $S \in R\text{-simp}$, $E(S)$ is noetherian.

Proposition 3.1. *Let R be a ring.*

- 1) R is left artinian if and only if every finitely generated R -module is finitely cogenerated.
- 2) R is left coartinian if and only if every finitely cogenerated R -module is finitely generated.

Proof. 1) Suppose that R is left artinian and take some finitely generated $M \in R\text{-Mod}$. As R is left noetherian, $\text{soc}(M)$, being a submodule of M , is also finitely generated. Also, R being left semiartinian implies that $\text{soc}(M) \leq_e M$. Therefore, M is finitely cogenerated.

Conversely, suppose that every finitely generated module is finitely cogenerated. Any quotient of ${}_R R$, being cyclic, is by hypothesis finitely cogenerated. Thus, ${}_R R$ is artinian.

2) Suppose that R is left coartinian and take some finitely cogenerated $M \in R\text{-Mod}$.

Then, there are some simple S_1, \dots, S_n such that $\bigoplus_{i=1}^n S_i = \text{soc}(M) \leq_e M$. But then

$E(M) = E(\text{soc}(M)) = \bigoplus_{i=1}^n E(S_i)$ is, by hypothesis, noetherian, so that its submodule M is finitely generated.

Conversely, suppose that every finitely cogenerated module is finitely generated. For every $S \in R\text{-simp}$, any submodule of $E(S)$, being finitely cogenerated, is by hypothesis finitely generated. Thus, $E(S)$ is noetherian. \square

A ring R is called left *conoetherian* if for every $S \in R\text{-simp}$, $E(S)$ is artinian. Accordingly, let us call R left *strongly conoetherian* if every indecomposable[†] injective R -module is artinian.

Theorem 3.2. *Let R be a ring. The following statements are equivalent.*

- 1) R is left artinian and left coartinian.
- 2) The classes of finitely generated and of finitely cogenerated R -modules coincide.

[†]By "indecomposable" we mean "directly indecomposable".

3) R is left noetherian and left strongly conoetherian.

Proof. 1) \Leftrightarrow 2) Direct from Proposition 3.1.

1) \Rightarrow 3) Suppose that 1) holds. Of course, every left artinian ring is left noetherian. Let $E \in R\text{-Mod}$ be injective and indecomposable. Since R is artinian, there is some simple S such that $E = E(S)$. Then, as R is left coartinian, E is noetherian, and in particular finitely generated. But over a left artinian ring, every finitely generated module is artinian. Therefore, 3) holds.

3) \Rightarrow 1) Suppose now that 3) holds. Since R is left noetherian, in order to prove that it is left artinian it suffices to show that it is left semiartinian. Take then some non-zero $M \in R\text{-Mod}$. As is well-known, left noetherian rings are characterized by the fact that over them, every injective module is a direct sum of indecomposable modules. We can apply this to $E(M)$ and then use the fact that R is left strongly conoetherian to obtain some simple $S \leq E(M)$. Then, by simplicity, $S \leq M$.

Let now $S \in R\text{-simp}$. Let us write $J = \text{rad}(R)$. We have already established that R is left artinian, so J is nilpotent. Then there is a least $n \in \mathbb{N}$ such that $J^n E(S) = 0$. (Of course, $n > 0$.) Observe that both of $J^{n-1} E(S)$ and $J^{n-2} E(S)/J^{n-1} E(S)$ are artinian and semisimple. Indeed, they are subquotients of $E(S)$, an artinian module, and they are annihilated by J , i.e. they are R/J -modules (since R is left artinian, it is semilocal). Thus, $J^{n-1} E(S)$ and $J^{n-2} E(S)/J^{n-1} E(S)$ are noetherian, so that the short exact sequence

$$0 \rightarrow J^{n-1} E(S) \rightarrow J^{n-2} E(S) \rightarrow J^{n-2} E(S)/J^{n-1} E(S) \rightarrow 0$$

shows that $J^{n-2} E(S)$ is noetherian. Next, we use

$$0 \rightarrow J^{n-2} E(S) \rightarrow J^{n-3} E(S) \rightarrow J^{n-3} E(S)/J^{n-2} E(S) \rightarrow 0$$

to show that $J^{n-3} E(S)$ is noetherian, and so on. At the n -th step, we obtain that $E(S)$ is noetherian. \square

Theorem 3.3. *Let R be a ring. The following conditions are equivalent.*

- 1) R is a domain[‡] and the class of free R -modules is closed under taking injective hulls.
- 2) R is a division ring.

Proof. 1) \Rightarrow 2) We claim that every free module is injective. Consider $R^{(X)}$ for some set X . Suppose first that $|X| > |R|$. By hypothesis, $E(R^{(X)}) = R^{(Y)}$ for some set Y , so that $|R^{(X)}| \leq |R^{(Y)}|$. Let us verify that $|X| \leq |Y|$.

In case both of $R^{(X)}$ and $R^{(Y)}$ are infinite, we have that

$$\max\{|R|, |Y|\} = |R^{(Y)}| \geq |R^{(X)}| = \max\{|R|, |X|\} = |X| > |R|,$$

so that it must happen that $\max\{|R|, |Y|\} = |Y|$. Thus, $|X| \leq |Y|$.

In case both of $R^{(X)}$ and $R^{(Y)}$ are finite, we have that

$$|R|^{|X|} = |R^{(X)}| \leq |R^{(Y)}| = |R|^{|Y|},$$

so that necessarily $|X| \leq |Y|$.

Lastly, in case $R^{(X)}$ is finite and $R^{(Y)}$ is infinite, we must have that $|R|$ and $|X|$ are finite, and thus that $|Y|$ is infinite (seeing as, for any set A , $R^{(A)}$ is finite if and only if R and A are finite).

Thus, we always have that $|R| < |X| \leq |Y|$.

Let us write $\{\delta_y\}_{y \in Y}$ for the canonical basis of $R^{(Y)}$ §. Since $R^{(X)} \hookrightarrow_e R^{(Y)}$, for each $y \in Y$ there is an $r_y \in R$ such that $0 \neq r_y \delta_y \in R^{(X)}$. By the hypothesis on R , the set

[‡]That is, every product of non-zero elements of R is non-zero.

[§]That is, for $y \in Y$, $\delta_y : Y \rightarrow R$ is such that $\delta_y : z \mapsto \begin{cases} 1 & \text{if } z = y \\ 0 & \text{if } z \neq y \end{cases}$, although any basis will do.

$\{r_y\delta_y\}_{y\in Y}$ is linearly independent. Thus, the submodule of $R^{(X)}$ spanned by $\{r_y\delta_y\}_{y\in Y}$ is a free module. Therefore,

$$R^{(Y)} \cong R^{\langle\{r_y\delta_y\}_{y\in Y}\rangle} \cong \langle\{r_y\delta_y\}_{y\in Y}\rangle \leq R^{(X)},$$

so that $|R^{(Y)}| \leq |R^{(X)}|$. As we already had the reverse inequality, we obtain that $|R^{(X)}| = |R^{(Y)}|$. This cardinality may or may not be finite, but it is now easy to show that, in any case, $|X| = |Y|$. This implies that $R^{(X)} \cong R^{(Y)}$, whence $R^{(X)}$ is injective.

Now, in case $|X| \leq |R|$, simply take some set Z such that $|Z| > |R|$ and note that $R^{(X)}$ embeds as a direct summand in $R^{(Z)}$, which, by the above argument, is injective. Therefore, the claim is proved.

Since every projective module is a direct summand of a free module, every projective module is injective. This condition is well-known to be equivalent to R being QF. Therefore R is a left artinian domain; thus, there exists a minimal left ideal Rx , which is isomorphic to R . Therefore R is a division ring.

2) \Rightarrow 1) It is clear. □

Theorem 3.4. *Suppose that R is left artinian and that the class of injective cogenerators of $R\text{-Mod}$ is closed under taking projective covers. Then R is a QF ring.*

Proof. Take some injective cogenerator of $R\text{-Mod}$, and let P stand for its projective cover, which exists because R is left perfect. By hypothesis, P is a projective and injective cogenerator. Let $S \in R\text{-simp}$. As S is cogenerated by P , by simplicity S embeds in P , so that $E(S)$ embeds in P as a direct summand, which makes it projective. Also observe that, over any artinian ring, any injective indecomposable module E is the injective hull of some simple $S \leq E$.

Let $M \in R\text{-Mod}$ be injective. As R is left noetherian, M is a direct sum of injective indecomposable modules. By the above remark, M is a direct sum of injective hulls of simple modules, which we know are projective. Therefore, every injective module is projective. This condition is well-known to be equivalent to R being QF. □

Acknowledgment. The authors would like to express their gratitude to the referees for their helpful suggestions, which improved the contents of this paper.

References

- [1] A. Alvarado-García, C. Cejudo-Castilla, H.A. Rincón Mejía, I.F. Vilchis-Montalvo, *On classes of modules closed under injective hulls and artinian principal ideal rings*, Int. Electron. J. Algebra **15**, 1-12, 2014.
- [2] A. Alvarado-García, C. Cejudo-Castilla, H.A. Rincón-Mejía, I.F. Vilchis-Montalvo, *Pseudocomplements and strong pseudocomplements in lattices of module classes*, J. Algebra Appl. **17** (1), 1850016 (14 pages), 2018.
- [3] A. Alvarado-García, H.A. Rincón-Mejía, J. Ríos-Montes, *On some lattices of module classes*, J. Algebra Appl. **5** (1), 105-117, 2006.
- [4] A. El Mejdani, H. Essannouni, A. Kaidi, *Rings with nice artinian modules*, Int. J. Algebra. **2** (18), 895-904, 2008.
- [5] C. Faith, *On Köthe rings*, Math. Annalen **164**, 207-212, 1966.
- [6] C. Faith, *Algebra II, Ring Theory*, Grundlehren der Mathematischen Wissenschaften **191**, Springer-Verlag, 1976.
- [7] L. Shen and J.L. Chen, *On strong Goldie dimension*, Comm. Algebra **35**, 3018-3025, 2007.



Caristi type fixed point theorems in fuzzy metric spaces

Hakan Karayılan, Mustafa Telci*

Department of Mathematics, Science Faculty, Trakya University, 22030 Edirne, Turkey

Abstract

In this paper, we extend the generalized Caristi's fixed point theorem proved by Bollenbacher and Hicks to p -orbitally complete fuzzy metric spaces by considering the fuzzy metric spaces in the sense of George and Veeramani. We also give some illustrative examples that support our results.

Mathematics Subject Classification (2010). 54A40, 47H10

Keywords. Caristi's fixed point theorem, fixed point, fuzzy metric space, p -complete fuzzy metric space

1. Introduction

In 1976, Caristi [8] proved the following fixed point theorem on a complete metric space, which is one of the most important generalization of famous Banach contraction principle and is equivalent the Ekeland's variational principle [13].

Let T is a self-mapping of a complete metric space (X, d) such that there is a lower semi continuous function φ from X into $[0, \infty)$ satisfying

$$d(x, Tx) \leq \varphi(x) - \varphi(Tx)$$

for all $x \in X$, then T has a fixed point.

In this theorem, saying that φ is lower semi continuous at x if for any sequence $\{x_n\} \subset X$, we have $\lim x_n = x$ implies $\varphi(x) \leq \liminf \varphi(x_n)$.

Several authors have obtained various extensions and generalizations of Caristi's theorem by considering Caristi type mappings on many different spaces. For example, [1–7, 9, 23–25, 27, 28, 30, 31, 33, 38, 40], and others.

In this paper, we extend the results in [7] to fuzzy metric spaces.

Several notions of fuzzy metric spaces have been introduced and discussed in different directions by many mathematicians, see [10, 14, 29, 34, 39]. In particular, Kramosil and Michalek [34] introduced and studied the notion of fuzzy metric space which is closely related to a class of probabilistic metric spaces. In [15, 17] George and Veeramani modified the concept of fuzzy metric space of Kramosil and Michalek, and obtained a Hausdorff and first countable topology on the modified fuzzy metric space. In [16, 20], it was proved that the topology induced by a fuzzy metric space in George and Veeramani's sense is metrizable. Grabiec [18] obtained a fuzzy version of the Banach contraction principle

*Corresponding Author.

Email addresses: hakankarayilan@trakya.edu.tr (H. Karayılan), mtelci@trakya.edu.tr (M. Telci)

Received: 21.04.2017; Accepted: 11.09.2017

in fuzzy metric spaces in Kramosil and Michalek's sense, and since then many author [11, 12, 22, 26, 32, 35, 41, 42] have proved fixed point theorems in fuzzy metric spaces in the sense of Kramosil and Michalek and George and Veeramani, using one of the two different types of completeness in Grabiec's sense [18] or George and Veeramani's sense [15].

In [36] Miheţ defined a concept weaker than convergence called p -convergence and proved a fixed point theorem for fuzzy contractive mappings. Then, in [19] Gregori et al. introduced the concept of p -Cauchy sequence and showed that p -Cauchy sequence and Cauchy sequence are two different concepts even in principal fuzzy metric spaces and they also defined the concept p -completeness.

In this paper, we consider $(X, M, *)$ fuzzy metric space in George and Veeramani's sense and prove some fixed point theorems for Caristi type mappings orbitally p -complete fuzzy metric spaces.

2. Preliminaries

In this section, we give some known basic notion of fuzzy metric space in the sense of George and Veeramani. Throughout this paper, we denote by \mathbb{N} the set of positive integers.

Definition 2.1 ([39]). A binary operation $*$: $[0, 1] \times [0, 1] \rightarrow [0, 1]$ is a continuous t -norm if satisfies the following conditions:

- (i) $*$ is associative and commutative,
- (ii) $*$ is continuous,
- (iii) $a * 1 = a$ for every $a \in [0, 1]$,
- (iv) $a * b \leq c * d$ if $a \leq c$ and $b \leq d$ for all $a, b, c, d \in [0, 1]$.

Definition 2.2 ([15]). The 3-tuple $(X, M, *)$ is said to be a fuzzy metric space if X is an arbitrary set, $*$ is a continuous t -norm and M is a fuzzy set on $X \times X \times (0, +\infty)$ satisfying the following conditions, for all $x, y, z \in X$ and $t, s > 0$:

- (i) $M(x, y, t) > 0$,
- (ii) $M(x, y, t) = 1$ iff $x = y$,
- (iii) $M(x, y, t) = M(y, x, t)$,
- (iv) $M(x, y, t) * M(y, z, s) \leq M(x, z, t + s)$,
- (v) $M(x, y, \cdot) : (0, +\infty) \rightarrow (0, 1]$ is continuous.

If $(X, M, *)$ is a fuzzy metric space, we will say that $(M, *)$ is a fuzzy metric on X . If we replace (iv) by

- (vi) $M(x, y, t) * M(y, z, s) \leq M(x, z, \max\{t, s\})$,

then 3-tuple $(X, M, *)$ is called a non-Archimedean fuzzy metric space. Since (vi) implies (iv) then each non-Archimedean fuzzy metric space is a fuzzy metric space.

Example 2.3. Let (X, d) be a metric space. Denote by $a.b$ the usual multiplication for all $a, b \in [0, 1]$, and let M_d be the function defined on $X \times X \times (0, +\infty)$ by

$$M_d(x, y, t) = \frac{t}{t + d(x, y)}.$$

Then (X, M_d, \cdot) is a fuzzy metric space called standard fuzzy metric space and (M_d, \cdot) is called the standard fuzzy metric of d (see [15]).

George and Veeramani proved in [15] that ever fuzzy metric $(M, *)$ on X generates a topology τ_M on X which has as a base the family of sets of the form

$$\{B(x, r, t) : x \in X, 0 < r < 1, t > 0\},$$

where

$$B(x, r, t) = \{y \in X : M(x, y, t) > 1 - r\}$$

for all $r \in (0, 1)$ and $t > 0$. They proved also that (X, τ_M) is a Hausdorff first countable topological space.

Definition 2.4 ([21]). A fuzzy metric M on X is said to be stationary if M does not depend on t , i.e. if for each $x, y \in X$, the function $M_{x,y}(t) = M(x, y, t)$ is constant. In this case we write $M(x, y)$ instead of $M(x, y, t)$.

Theorem 2.5 ([15]). A sequence $\{x_n\}$ in a fuzzy metric space $(X, M, *)$ converges to x if and only if $M(x_n, x, t) \rightarrow 1$ as $n \rightarrow +\infty$.

The following definition was given by Miheţ.

Definition 2.6 ([36]). A sequence $\{x_n\}$ in a fuzzy metric space $(X, M, *)$ is called p -convergent to $x_0 \in X$ (we write $x_n \rightarrow_p x_0$) if $\lim_n M(x_n, x_0, t_0) = 1$ for some $t_0 > 0$.

If $\{x_n\}$ is p -convergent to x_0 , then

(1) $\{x_n\}$ in X has at most one limit.

(2) Every subsequence of $\{x_n\}$ is also convergent and has the same limit as the whole sequence, see [36].

Note that $\{x_n\}$ is convergent to x_0 if and only if $\{x_n\}$ is p -convergent to x_0 for all $t > 0$, see [19].

In [36] the author gave an example that there exist p -convergent but not convergent sequences.

Definition 2.7 ([18]). A sequence $\{x_n\}$ in a fuzzy metric space $(X, M, *)$ is G -Cauchy sequence iff $\lim_{n \rightarrow \infty} M(x_{n+p}, x_n, t) = 1$ for all $t > 0$ and $p \in \mathbb{N}$. A fuzzy metric space $(X, M, *)$ is G -complete if every G -Cauchy sequence is convergent in X .

Definition 2.8 ([15]). A sequence $\{x_n\}$ in a fuzzy metric space $(X, M, *)$ is Cauchy sequence iff for each $\varepsilon \in (0, 1)$ and each $t > 0$, there exists $n_0 \in \mathbb{N}$ such that $M(x_n, x_m, t) > 1 - \varepsilon$ for all $n, m \geq n_0$. A fuzzy metric space $(X, M, *)$ is complete if every Cauchy sequence is convergent in X .

In [19] Gregori et al. gave the following definition of Cauchyness and completeness in a natural way from the p -convergence concept.

Definition 2.9 ([19]). A sequence $\{x_n\}$ in a fuzzy metric space $(X, M, *)$ is called p -Cauchy if there exists $t_0 > 0$ such that for each $\varepsilon \in (0, 1)$ there exists $n_0 \in \mathbb{N}$ such that $M(x_n, x_m, t_0) > 1 - \varepsilon$ for all $n, m \geq n_0$, or equivalently $\lim_{n, m \rightarrow \infty} M(x_n, x_m, t_0) = 1$ for some $t_0 > 0$. A fuzzy metric space $(X, M, *)$ is p -complete if every p -Cauchy sequence in X is p -convergent to some point of X .

Note that $\{x_n\}$ is a Cauchy sequence if and only if $\{x_n\}$ is p -Cauchy for all $t > 0$ and, obviously, p -convergent sequences are p -Cauchy.

p -completeness and completeness are equivalent concepts in stationary fuzzy metrics, see [19].

Remark 2.10 ([19]). Let $(X, M_d, *)$ be a standard fuzzy metric space as in Example 2.3. Then $(X, M_d, *)$ is p -complete if and only if the metric space (X, d) is complete.

Definition 2.11 ([12]). Let $(X, M, *)$ be a fuzzy metric space. The fuzzy metric M is triangular if it satisfies the condition

$$\frac{1}{M(x, y, t)} - 1 \leq \frac{1}{M(x, z, t)} - 1 + \frac{1}{M(z, y, t)} - 1$$

for every $x, y, z \in X$ and every $t > 0$.

Note that every standard fuzzy metric (M_d, \cdot) is triangular.

Theorem 2.12 ([37]). *Let $(X, M, *)$ be a fuzzy metric space. Then M is a continuous function on $X \times X \times (0, +\infty)$.*

Definition 2.13. Let $(X, M, *)$ be a fuzzy metric space and $T : X \rightarrow X$ a mapping. The set $O_T(x, \infty) = \{x, Tx, T^2x, \dots\}$ is called the orbit of x . If for an $x \in X$, every p -Cauchy sequence in $O_T(x, \infty)$ is p -converges to a point in X , then the fuzzy metric space $(X, M, *)$ is said to be (x, T) -orbitally p -complete.

Definition 2.14. Let $(X, M, *)$ be a fuzzy metric space and $T : X \rightarrow X$ a mapping. A real-valued function $G : X \times (0, +\infty) \rightarrow [0, \infty)$ is said to be (x, T) -orbitally p -weak lower semi-continuous (p -w.l.s.c.) at u iff $\{x_n\}$ is a sequence in $O_T(x, \infty)$ and

$$x_n \rightarrow_p u \quad \text{implies} \quad G(u, t_0) \leq \limsup_{n \rightarrow \infty} G(x_n, t_0)$$

for some $t_0 > 0$. That is, $G(\cdot, t_0)$ is p -w.l.s.c on X in Ćirić's sense, see [9].

3. Main results

In this section, we state and prove our main results in orbitally p -complete fuzzy metric spaces. Now, we give the first main result as follows.

Theorem 3.1. *Let $(X, M, *)$ be a fuzzy metric space with M is triangular, $T : X \rightarrow X$ and $\Phi : X \times (0, +\infty) \rightarrow [0, \infty)$. Suppose there exist $x \in X$ and $t_0 > 0$ such that $(X, M, *)$ is (x, T) -orbitally p -complete, and*

$$\frac{1}{M(y, Ty, t_0)} - 1 \leq \Phi(y, t_0) - \Phi(Ty, t_0) \quad (3.1)$$

for all $y \in O_T(x, \infty)$. Then:

- (i) $T^n x \rightarrow_p x'$ exists,
- (ii) $\frac{1}{M(T^n x, x', t_0)} - 1 \leq \Phi(T^n x, t_0)$,
- (iii) $Tx' = x'$ if and only if $G(z, t_0) = \frac{1}{M(z, Tz, t_0)} - 1$ is (x, T) -orbitally p -w.l.s.c. at x' ,
- (iv) $\frac{1}{M(T^n x, x, t_0)} - 1 \leq \Phi(x, t_0)$ and $\frac{1}{M(x', x, t_0)} - 1 \leq \Phi(x, t_0)$.

Proof. (i) Using inequality (3.1) we have

$$\begin{aligned} S_n &= \sum_{i=0}^n \left(\frac{1}{M(T^i x, T^{i+1} x, t_0)} - 1 \right) \leq \sum_{i=0}^n [\Phi(T^i x, t_0) - \Phi(T^{i+1} x, t_0)] \\ &= \Phi(x, t_0) - \Phi(T^{n+1} x, t_0) \leq \Phi(x, t_0) \end{aligned}$$

for $n = 0, 1, 2, \dots$. Therefore, $\{S_n\}$ is bounded from above and also non-decreasing and so convergent.

Let $m > n$. Since M is triangular, we have

$$\frac{1}{M(T^n x, T^m x, t_0)} - 1 \leq \sum_{k=n}^{m-1} \left(\frac{1}{M(T^k x, T^{k+1} x, t_0)} - 1 \right) \quad (3.2)$$

Since $\{S_n\}$ is convergent, for every $1 > \varepsilon > 0$, we can choose a sufficiently large $N \in \mathbb{N}$ such that

$$\sum_{k=n}^{\infty} \left(\frac{1}{M(T^k x, T^{k+1} x, t_0)} - 1 \right) < \varepsilon$$

for all $n \geq N$. Thus, we get from inequality (3.2) that

$$\frac{1}{M(T^n x, T^m x, t_0)} - 1 \leq \sum_{k=n}^{m-1} \left(\frac{1}{M(T^k x, T^{k+1} x, t_0)} - 1 \right) < \varepsilon$$

and so

$$\frac{1}{M(T^n x, T^m x, t_0)} < 1 + \varepsilon.$$

Since $1 - \varepsilon^2 < 1$, it follows that

$$M(T^n x, T^m x, t_0) > \frac{1}{1 + \varepsilon} = \frac{1 - \varepsilon}{1 - \varepsilon^2} > 1 - \varepsilon$$

for all $n, m \geq N$. Hence, $\{T^n x\}$ is a p -Cauchy sequence in $O_T(x, \infty)$. Since $(X, M, *)$ is (x, T) -orbitally p -complete, $T^n x \rightarrow_p x' \in X$ exists.

(ii) Let $m > n$. Using inequalities (3.1) and (3.2) we have

$$\begin{aligned} \frac{1}{M(T^n x, T^m x, t_0)} - 1 &\leq \sum_{k=n}^{m-1} \left(\frac{1}{M(T^k x, T^{k+1} x, t_0)} - 1 \right) \\ &\leq \sum_{k=n}^{m-1} [\Phi(T^k x, t_0) - \Phi(T^{k+1} x, t_0)] \\ &= \Phi(T^n x, t_0) - \Phi(T^m x, t_0) \leq \Phi(T^n x, t_0). \end{aligned}$$

Letting m tend to infinity, we have from (i) and Theorem 2.12

$$\frac{1}{M(T^n x, x', t_0)} - 1 \leq \Phi(T^n x, t_0).$$

(iii) Assume that $Tx' = x'$ and $\{x_n\}$ is a sequence in $O_T(x, \infty)$ with $x_n \rightarrow_p x'$. Then $G(x', t_0) = \frac{1}{M(x', Tx', t_0)} - 1 = 0 \leq \limsup \left(\frac{1}{M(x'_n, Tx'_n, t_0)} - 1 \right) = \limsup G(x_n, t_0)$, and so G is (x, T) -orbitally p -w.l.s.c. at x' .

Now let $x_n = T^n x$ and G is (x, T) -orbitally p -w.l.s.c. at x' . Then from (i) we have

$$\begin{aligned} 0 \leq \frac{1}{M(x', Tx', t_0)} - 1 = G(x', t_0) &\leq \limsup G(T^n x, t_0) \\ &= \limsup \left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right) = 0 \end{aligned}$$

which implies $\frac{1}{M(x', Tx', t_0)} - 1 = 0$. Thus $M(x', Tx', t_0) = 1$ and so $Tx' = x'$.

(iv) We first of all prove by induction that

$$\frac{1}{M(T^n x, x, t_0)} - 1 \leq \sum_{k=0}^{n-1} \left(\frac{1}{M(T^k x, T^{k+1} x, t_0)} - 1 \right) \tag{3.3}$$

for all $n = 1, 2, 3, \dots$

Inequality (3.3) is trivial when $n = 1$ and so we will assume that inequality (3.3) holds for $n - 1$. Since M is triangular, it follows from inequality (3.1) we have

$$\begin{aligned} \frac{1}{M(T^n x, x, t_0)} - 1 &\leq \frac{1}{M(T^n x, T^{n-1} x, t_0)} - 1 + \frac{1}{M(T^{n-1} x, x, t_0)} - 1 \\ &\leq \sum_{k=0}^{n-2} \left(\frac{1}{M(T^k x, T^{k+1} x, t_0)} - 1 \right) + \frac{1}{M(T^n x, T^{n-1} x, t_0)} - 1 \\ &= \sum_{k=0}^{n-1} \left(\frac{1}{M(T^k x, T^{k+1} x, t_0)} - 1 \right). \end{aligned}$$

It therefore follows by induction that inequality (3.3) holds.

Using inequalities (3.1) and (3.3) we have

$$\begin{aligned} \frac{1}{M(T^n x, x, t_0)} - 1 &\leq \sum_{k=0}^{n-1} \left(\frac{1}{M(T^k x, T^{k+1} x, t_0)} - 1 \right) \\ &\leq \sum_{k=0}^{n-1} [\Phi(T^k x, t_0) - \Phi(T^{k+1} x, t_0)] \\ &= \Phi(x, t_0) - \Phi(T^n x, t_0) \leq \Phi(x, t_0). \end{aligned}$$

Letting n tend to infinity we have

$$\frac{1}{M(x', x, t_0)} - 1 \leq \Phi(x, t_0).$$

□

Corollary 3.2. *Let $(X, M, *)$ be a fuzzy metric space with M is triangular and T be a self-mapping of X . Suppose there exist $x \in X$ and $t_0 > 0$ such that $(X, M, *)$ is (x, T) -orbitally p -complete, and*

$$\frac{1}{M(Ty, T^2y, t_0)} - 1 \leq k \left(\frac{1}{M(y, Ty, t_0)} - 1 \right) \quad (3.4)$$

for all $y \in O_T(x, \infty)$. Then:

- (i) $T^n x \rightarrow_p x'$ exists,
- (ii) $\frac{1}{M(T^n x, x', t_0)} - 1 \leq k^n (1 - k)^{-1} \left(\frac{1}{M(x, Tx, t_0)} - 1 \right)$,
- (iii) $Tx' = x'$ if and only if $G(z, t_0) = \frac{1}{M(z, Tz, t_0)} - 1$ is (x, T) -orbitally p -w.l.s.c. at x' ,
- (iv) $\frac{1}{M(T^n x, x, t_0)} - 1 \leq \frac{1}{1-k} \left(\frac{1}{M(x, Tx, t_0)} - 1 \right)$, $\frac{1}{M(x', x, t_0)} - 1 \leq \frac{1}{1-k} \left(\frac{1}{M(x, Tx, t_0)} - 1 \right)$.

Proof. Put $\Phi(y, t) = (1 - k)^{-1} \left(\frac{1}{M(y, Ty, t)} - 1 \right)$ for $y \in O_T(x, \infty)$. Let $y = T^n x$ in (3.4). Then we have,

$$\frac{1}{M(T^{n+1} x, T^{n+2} x, t_0)} - 1 \leq k \left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right)$$

and

$$\begin{aligned} \left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right) - k \left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right) &\leq \\ \left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right) - \left(\frac{1}{M(T^{n+1} x, T^{n+2} x, t_0)} - 1 \right) & \end{aligned}$$

and so

$$\begin{aligned} \frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 &\leq \\ (1 - k)^{-1} \left[\left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right) - \left(\frac{1}{M(T^{n+1} x, T^{n+2} x, t_0)} - 1 \right) \right]. & \end{aligned}$$

Thus, we get

$$\frac{1}{M(y, Ty, t_0)} - 1 \leq \Phi(y, t_0) - \Phi(Ty, t_0)$$

so (i), (iii) and (iv) are immediate from Theorem 3.1.

Using inequality (3.4) we have

$$\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \leq k^n \left(\frac{1}{M(x, Tx, t_0)} - 1 \right)$$

and then from Theorem 3.1 (ii) we get

$$\begin{aligned} \frac{1}{M(T^n x, x', t_0)} - 1 &\leq \Phi(T^n x, t_0) \\ &= (1 - k)^{-1} \left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right) \\ &\leq k^n (1 - k)^{-1} \left(\frac{1}{M(x, Tx, t_0)} - 1 \right) \end{aligned}$$

and this gives (ii). \square

In the following theorem, we will show that if $(M, *)$ is non-Archimedean fuzzy metric, where the continuous t -norm is defined as $a * b = \min\{a, b\}$ for all $a, b \in [0, 1]$, then (i) and (iii) of Theorem 3.1 can be obtained without the triangular property of M .

Theorem 3.3. *Let $(X, M, *)$ be a non-Archimedean fuzzy metric space, where the continuous t -norm is defined as $a * b = \min\{a, b\}$ for all $a, b \in [0, 1]$. Let $T : X \rightarrow X$ and $\Phi : X \times (0, +\infty) \rightarrow [0, \infty)$. Suppose there exist $x \in X$ and $t_0 > 0$ such that $(X, M, *)$ is (x, T) -orbitally p -complete, and satisfying the inequality (3.1) for all $y \in O_T(x, \infty)$. Then:*

- (i) $T^n x \rightarrow_p x'$ exists,
- (ii) $Tx' = x'$ if and only if $G(z, t) = \frac{1}{M(z, Tz, t_0)} - 1$ is (x, T) -orbitally p -w.l.s.c. at x' .

Proof. (i) Using the same procedure as in the proof of Theorem 3.1, we obtain that

$$S_n = \sum_{i=0}^n \left(\frac{1}{M(T^i x, T^{i+1} x, t_0)} - 1 \right)$$

is convergent. Therefore we have

$$\sum_{n=0}^{\infty} \left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right) < \infty \quad \text{and so} \quad \lim_{n \rightarrow \infty} \left(\frac{1}{M(T^n x, T^{n+1} x, t_0)} - 1 \right) = 0$$

Thus, $\lim_{n \rightarrow \infty} M(T^n x, T^{n+1} x, t_0) = 1$. Hence for $0 < \varepsilon < 1$, there exists $n_0 \in \mathbb{N}$ such that $M(T^n x, T^{n+1} x, t_0) > 1 - \varepsilon$ for all $n > n_0$. Let $n_0 < n < m$. Using (vi) of Definition 2.2, we have

$$\begin{aligned} M(T^n x, T^m x, t_0) &\geq \overbrace{M(T^n x, T^{n+1} x, t_0) * \cdots * M(T^{m-1} x, T^m x, t_0)}^{m-n} \\ &= \min \{ M(T^n x, T^{n+1} x, t_0), \dots, M(T^{m-1} x, T^m x, t_0) \} > 1 - \varepsilon \end{aligned}$$

and so the sequence $\{T^n x\}$ is a p -Cauchy sequence in $O_T(x, \infty)$. Since $(X, M, *)$ is (x, T) -orbitally p -complete, $T^n x \rightarrow_p x'$ exists.

Using the same procedure as in the proof of Theorem 3.1 (iii), we obtain (ii). \square

Similarly, using the same procedure as in the proof of Corollary 3.2 (i) and (iii), we obtain the following result.

Corollary 3.4. *Let $(X, M, *)$ be a non-Archimedean fuzzy metric space, where the continuous t -norm is defined as $a * b = \min\{a, b\}$ for all $a, b \in [0, 1]$. Let T be a self-mapping of X . Suppose there exists an $x \in X$ such that $(X, M, *)$ is (x, T) -orbitally complete, and satisfying the inequality (3.4), for all $y \in O_T(x, \infty)$. Then:*

- (i) $T^n x \rightarrow_p x'$ exists,
- (ii) $Tx' = x'$ if and only if $G(z, t) = \frac{1}{M(z, Tz, t_0)} - 1$ is (x, T) -orbitally p -w.l.s.c. at x' .

If we replace non-Archimedean fuzzy metric by stationary fuzzy metric in the Theorem 3.3, then using the same procedure as in the proof of Theorem 3.3 and Corollary 3.2 we obtain the following results.

Theorem 3.5. *Let $(X, M, *)$ be a stationary fuzzy metric space, where the continuous t -norm is defined as $a * b = \min\{a, b\}$ for all $a, b \in [0, 1]$. Let $T : X \rightarrow X$ and $\Phi : X \rightarrow [0, \infty)$. Suppose there exists an $x \in X$ such that $(X, M, *)$ is (x, T) -orbitally p -complete, and*

$$\frac{1}{M(y, Ty)} - 1 \leq \Phi(y) - \Phi(Ty)$$

for all $y \in O_T(x, \infty)$. Then:

- (i) $T^n x \rightarrow_p x'$ exists,
- (ii) $Tx' = x'$ if and only if $G(z) = \frac{1}{M(z, Tz)} - 1$ is (x, T) -orbitally p -w.l.s.c. at x' .

Corollary 3.6. *Let $(X, M, *)$ be a stationary fuzzy metric space, where the continuous t -norm is defined as $a * b = \min\{a, b\}$ for all $a, b \in [0, 1]$. Let T be a self-mapping of X . Suppose there exists an $x \in X$ such that $(X, M, *)$ is (x, T) -orbitally p -complete, and*

$$\frac{1}{M(Ty, T^2y)} - 1 \leq k \left(\frac{1}{M(y, Ty)} - 1 \right)$$

for all $y \in O_T(x, \infty)$. Then:

- (i) $T^n x \rightarrow_p x'$ exists,
- (ii) $Tx' = x'$ if and only if $G(z) = \frac{1}{M(z, Tz)} - 1$ is (x, T) -orbitally p -w.l.s.c. at x' .

Note that Theorem 3.5 and Corollary 3.6 are true for complete fuzzy metric spaces since p -completeness and completeness are equivalent concepts in stationary fuzzy metrics.

The following theorem is slight generalization of Theorem 3 in [7].

Theorem 3.7 ([7]). *Let (X, d) be a metric space, $T : X \rightarrow X$ and $\varphi : X \rightarrow [0, \infty)$. Suppose there exists an $x \in X$ such that*

$$d(y, Ty) \leq \varphi(y) - \varphi(Ty) \tag{3.5}$$

for all $y \in O_T(x, \infty)$, and (X, d) is (x, T) -orbitally complete. Then:

- (i) $\lim_{n \rightarrow \infty} T^n x = x'$ exists,
- (ii) $d(T^n x, x') \leq \varphi(T^n x)$,
- (iii) $Tx' = x'$ if and only if $F(z) = d(z, Tz)$ is (x, T) -orbitally w.l.s.c. at x' ,
- (iv) $d(T^n x, x) \leq \varphi(x)$ and $d(x', x) \leq \varphi(x)$.

Proof. We consider the (M_d, \cdot) standard fuzzy metric induced by d on X as in Example 2.3. By Remark 2.10 (X, M_d, \cdot) is (x, T) -orbitally p -complete since (X, d) orbitally complete. Also (M_d, \cdot) is triangular.

Since $M_d(x, y, t) = \frac{t}{t+d(x, y)}$, we have $d(x, y) = \frac{t}{M_d(x, y, t)} - t$ for all $x, y \in X$ and $t > 0$.

Define $\Phi(x, t_0) = \frac{1}{t_0} \varphi(x)$ for all $x \in X$. Then from inequality (3.5) we have

$$\frac{t_0}{M_d(y, Ty, t_0)} - t_0 \leq t_0(\Phi(y, t_0) - \Phi(Ty, t_0))$$

and so

$$\frac{1}{M_d(y, Ty, t_0)} - 1 \leq \Phi(y, t_0) - \Phi(Ty, t_0).$$

Thus T satisfies inequality (3.1) of Theorem 3.1.

(i) From Theorem 3.1 (i) we have $T^n x \rightarrow_p x'$ exists and so $\lim_{n \rightarrow \infty} T^n x = x'$ (in the metric space).

(ii) From Theorem 3.1 (ii) we have

$$\frac{1}{M(T^n x, x', t_0)} - 1 \leq \Phi(T^n x, t_0),$$

and so

$$\frac{1}{\frac{t_0}{t_0 + d(T^n x, x')}} - 1 = \frac{t_0 + d(T^n x, x') - t_0}{t_0} = \frac{d(T^n x, x')}{t_0} \leq \frac{1}{t_0} \varphi(T^n x).$$

Thus $d(T^n x, x') \leq \varphi(T^n x)$.

(iii) From Theorem 3.1 (iii) we have

$$\frac{1}{M_d(x, Tx, t_0)} - 1 = \frac{d(x, Tx)}{t_0}.$$

If $G(x, t_0) = \frac{1}{M_d(x, Tx, t_0)} - 1$ is (x, T) orbitally p -w.l.s.c. at x' , then $t_0 G(x, t_0) = d(x, Tx)$ is (x, T) -orbitally w.l.s.c. at x' too. Thus (iii) follows from Theorem 3.1 (iii).

(iv) From Theorem 3.1 (iv) we have

$$\frac{1}{M_d(T^n x, x, t_0)} - 1 \leq \Phi(x, t_0) \quad \text{and so} \quad \frac{d(T^n x, x')}{t_0} \leq \frac{1}{t_0} \varphi(x).$$

Thus $d(T^n x, x) \leq \varphi(x)$. Similarly $\frac{1}{M_d(x', x, t_0)} - 1 \leq \Phi(x, t_0)$ and so $d(x', x) \leq \varphi(x)$. \square

By considering the (M_d, \cdot) standard fuzzy metric induced by d on X in Corollary 3.2 we obtain the following corollary.

Corollary 3.8 ([7]). *Let (X, d) be a metric space and T be a self mapping of X . Suppose there exists an $x \in X$ such that*

$$d(Ty, T^2 y) \leq d(y, Ty)$$

for all $y \in O_T(x, \infty)$, and (X, d) is (x, T) -orbitally complete. Then:

- (i) $\lim_{n \rightarrow \infty} T^n x = x'$ exists,
- (ii) $d(T^n x, x') \leq \varphi(T^n x)$,
- (iii) $Tx' = x'$ if and only if $F(z) = d(z, Tz)$ is (x, T) -orbitally w.l.s.c. at x' ,
- (iv) $d(T^n x, x) \leq \varphi(x)$ and $d(x', x) \leq \varphi(x)$.

4. Some examples

We finally give some examples which illustrate our results.

Example 4.1. Let $X = [0, \infty)$, $a * b = \min\{a, b\}$ for all $a, b \in [0, 1]$ and let

$$M(x, y, t) = \frac{t}{t + |x - y|},$$

for all $x, y \in X$, $t > 0$, then $(X, *)$ is triangular. Define $T : X \rightarrow X$ by

$$T(x) = \begin{cases} x/2 & \text{if } 0 \leq x < 1, \\ x + 1 & \text{if } 1 \leq x < \infty \end{cases}$$

for all $x \in X$.

If we take $0 \leq x_0 < 1$, then $O_T(x_0, \infty) = \{x_0, x_0/2, x_0/2^2, \dots, x_0/2^{n-1}, \dots\}$ for $n = 1, 2, \dots$ and so $(X, M, *)$ is (x_0, T) -orbitally p -complete. Also define $\Phi(x, t) = x$ and put $t_0 = 1$. Then for all $y \in O_T(x_0, \infty)$ we have

$$\begin{aligned} \frac{1}{M(y, Ty, t_0)} - 1 &= \frac{1}{\frac{1}{1+|y-Ty|}} - 1 = |y - Ty| \\ &= \left| \frac{x_0}{2^{n-1}} - \frac{x_0}{2^n} \right| = \frac{x_0}{2^n} = \Phi(y, t_0) - \Phi(Ty, t_0). \end{aligned}$$

Moreover,

$$G(z, t_0) = \frac{1}{M(z, Tz, t_0)} - 1 = |z - Tz| = \begin{cases} z/2 & \text{if } 0 \leq z < 1, \\ 1 & \text{if } 1 \leq z < \infty \end{cases}$$

is (x_0, T) -orbitally p -w.l.s.c. at $x = 0$.

All the conditions of Theorem 3.1 are therefore satisfied and $x = 0$ is a fixed point of T .

Example 4.2. Let $X = (0, \infty)$, $a * b = \min\{a, b\}$ for all $a, b \in [0, 1]$ and let

$$M(x, y, t) = \frac{t}{t + |x - y|},$$

for all $x, y \in X$, $t > 0$. Define $T : X \rightarrow X$ by

$$T(x) = \begin{cases} x/2 & \text{if } 0 < x < 1, \\ 1 & \text{if } 1 \leq x < \infty \end{cases}$$

for all $x \in X$.

If we take $0 < x_0 < 1$, then $O_T(x_0, \infty) = \{x_0, x_0/2, x_0/2^2, \dots, x_0/2^{n-1}, \dots\}$ for $n = 1, 2, \dots$. But $(X, M, *)$ is not (x_0, T) -orbitally p -complete.

Now we take $1 \leq x_0 < \infty$. Then $O_T(x_0, \infty) = \{x_0, 1, 1, 1, \dots\}$. Thus $(X, M, *)$ is (x_0, T) -orbitally p -complete.

Define $\Phi(x, t) = x$ and put $t_0 = 1$. Then for $y = x_0 \neq 1$ we have

$$\frac{1}{M(y, Ty, t_0)} - 1 = |y - Ty| = |x_0 - 1| = x_0 - 1 = \Phi(y, t_0) - \Phi(Ty, t_0).$$

Also inequality (3.1) is satisfied for $y = 1$. Moreover,

$$G(z, t_0) = \frac{1}{M(z, Tz, t_0)} - 1 = |z - Tz| = \begin{cases} z/2 & \text{if } 0 < z < 1, \\ z - 1 & \text{if } 1 \leq z < \infty \end{cases}$$

is (x_0, T) -orbitally p -w.l.s.c. at $x = 1$.

All the conditions of Theorem 3.1 are therefore satisfied and $x = 1$ is a fixed point of T .

Example 4.3. Let $X = [0, \infty)$, $a * b = \min\{a, b\}$ for all $a, b \in [0, 1]$ and let

$$M(x, y, t) = \begin{cases} \frac{1}{1+\max\{x, y\}}, & \text{if } x \neq y, \\ 1 & \text{if } x = y, \end{cases}$$

for all $x, y \in X$, $t > 0$. $(X, M, *)$ is non-Archimedean fuzzy metric space. Define $T : X \rightarrow X$ by $T(x) = x/2$ for all $x \in X$.

If we take $x_0 = 1$, then $O_T(1, \infty) = \{1, 1/2, 1/2^2, \dots, 1/2^{n-1}, \dots\}$ for $n = 1, 2, \dots$ and so $(X, M, *)$ is $(1, T)$ -orbitally p -complete. Also define $\Phi(x, t) = 2x$. Then for all $y \in O_T(1, \infty)$ we have

$$\frac{1}{M(y, Ty, t_0)} - 1 = y = \frac{1}{2^{n-1}} = \frac{2}{2^{n-1}} - \frac{2}{2^n} = \Phi(y, t_0) - \Phi(Ty, t_0).$$

Moreover, $G(z, t_0) = \begin{cases} z, & \text{if } z \neq Tz, \\ 0 & \text{if } z = Tz \end{cases}$ is $(1, T)$ -orbitally p -w.l.s.c. at $x = 0$.

All the conditions of Theorem 3.3 are therefore satisfied and $x = 0$ is a fixed point of T .

References

- [1] T. Abdeljawad and E. Karapınar, *Quasiconic metric spaces and generalizations of Caristi Kirk's theorem*, Fixed Point Theory Appl. **Vol. 2009**, ID 574387, 9 p, 2009.
- [2] Ö. Acar and İ. Altun, *Some generalizations of Caristi type fixed point metric spaces*, Filomat **26**(4), 833-837, 2012.
- [3] Ö. Acar, İ. Altun and S. Romaguera, *Caristi's type mappings on complete partial metric spaces*, Fixed Point Theory **14**(1), 3-10, 2013.
- [4] R.P. Agarwal and M.A. Khamsi, *Extension of Caristi's fixed point theorem to vector valued metric spaces*, Nonlinear Anal. **74**, 141-145, 2011.
- [5] İ. Altun and D. Mihet, *Ordered non-archimedean fuzzy metric spaces and some fixed point results*, Fixed Point Theory Appl. **Vol. 2010**, ID 782680, 11 p., 2010.
- [6] J.S. Bae, *Fixed point theorems for weakly contractive multivalued maps*, J. Math. Anal. Appl. **284**, 690-697, 2003.
- [7] A. Bollenbacher and T.L. Hicks, *A fixed point theorem revisited*, Proc. Amer. Math. Soc. **102**, 898-900, 1988.
- [8] J. Caristi, *Fixed point theorems for mappings satisfying inwardness conditions*, Trans. Amer. Math. Soc. **215**, 241-251, 1976.
- [9] L. Ćirić, *Periodic and fixed point theorems in a quasi-metric space*, J. Austral. Math. Soc. Ser. A **54**, 80-85, 1993.
- [10] Z. Deng, *Fuzzy pseudometric spaces*, J. Math. Anal. Appl. **86**, 74-95, 1982.
- [11] C. Di Bari and C. Vetro, *A fixed point theorem for a family of mappings in a fuzzy metric space*, Rend. Circ. Mat. Palermo **52**, 315-321, 2003.
- [12] C. Di Bari and C. Vetro, *Fixed points, attractors and weak fuzzy contractive mappings in a fuzzy metric space*, J. Fuzzy Math. **13**, 973-982, 2005.
- [13] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. **47**, 324-353, 1974.
- [14] M.A. Erceg, *Metric spaces in fuzzy set theory*, J. Math. Anal. Appl. **69**, 205-230, 1979.
- [15] A. George and P. Veeramani, *On some results in fuzzy metric spaces*, Fuzzy Sets and Systems **64**, 395-399, 1994.
- [16] A. George and P. Veeramani, *Some theorems in fuzzy metric spaces*, J. Fuzzy Math. **3**, 933-940, 1995.
- [17] A. George and P. Veeramani, *On some results of analysis for fuzzy metric spaces*, Fuzzy Sets and Systems **90**, 365-368, 1997.
- [18] M. Grabiec, *Fixed points in fuzzy metric spaces*, Fuzzy Sets and Systems **27**, 385-389, 1988.
- [19] V. Gregori, A. López-Crevillén, S. Morillas, and A. Sapena, *On convergence in fuzzy metric spaces*, Topology Appl. **156**, 3002-3006, 2009.
- [20] V. Gregori and S. Romaguera, *Some properties of fuzzy metric spaces*, Fuzzy Sets and Systems **115**, 485-489, 2000.
- [21] V. Gregori and S. Romaguera, *Characterizing completable fuzzy metric spaces*, Fuzzy Sets and Systems **144**, 411-420, 2004.
- [22] V. Gregori and A. Sapena, *On fixed-point theorems in fuzzy metric spaces*, Fuzzy Sets and Systems **125**, 245-252, 2002.
- [23] A.M. Harder and L.M. Saliga, *Periodic and fixed point theorems in d -complete topological spaces*, Indian J. Pure Appl. Math. **26**, 787-796, 1995.
- [24] T.L. Hicks, *Fixed point theorems for quasi-metric spaces*, Math. Japon. **33**, 231-236, 1988.

- [25] T.L. Hicks, *Fixed point theorems for d -complete topological spaces I*, Internat. J. Math. Math. Sci. **15**, 435-440, 1992.
- [26] C. Ionescu, Sh. Rezapour and M.E. Samei, *Fixed points of a class of contractive-type multifunctions on fuzzy metric spaces*, U.P.B. Sci. Bull., Series A **76**, 3-12, 2014.
- [27] J.R. Jachymski, *Caristi's fixed point theorem and selections of set-valued contractions*, J. Math. Anal. Appl. **227**, 55-67, 1998.
- [28] O. Kada, T. Suzuki and W. Takahashi, *Nonconvex minimization theorems and fixed point theorems in complete metric spaces*, Math. Japon. **44**, 381-391, 1996.
- [29] O. Kaleva and S. Seikkala, *On fuzzy metric spaces*, Fuzzy Sets and Systems **12**, 215-229, 1984.
- [30] E. Karapinar, *Generalizations of Caristi Kirk's theorem on partial metric spaces*, Fixed Point Theory Appl. **2011**:4, 2011.
- [31] E. Karapinar, İ.M. Erhan and A. Öztürk, *Fixed point theorems on quasi-partial metric spaces*, Math. Com. Mod. **57**, 2442-2448, 2013.
- [32] H. Karayilan and M. Telci, *Common fixed point theorems for contractive type mappings in a fuzzy metric spaces*, Rend. Circ. Mat. Palermo **60**, 145-152, 2011.
- [33] M.A. Khamsi, *Remarks on Caristi's fixed point theorem*, Nonlinear Anal. **71**, 227-231, 2009.
- [34] I. Kramosil and J. Michalek, *Fuzzy metric and statistical metric spaces*, Kybernetika **11**, 336-344, 1975.
- [35] D. Mihet, *A Banach contraction theorem in fuzzy metric spaces*, Fuzzy Sets and Systems **144**, 431-439, 2004.
- [36] D. Mihet, *On fuzzy contractive mappings in fuzzy metric spaces*, Fuzzy Sets and Systems **158**, 915-921, 2007.
- [37] J. Rodríguez-López and S. Romaguera, *The Hausdorff fuzzy metric on compact sets*, Fuzzy Sets and Systems **147**, 273-283, 2004.
- [38] S. Romaguera, *A Kirk type characterization of completeness for partial metric spaces*, Fixed Point Theory Appl. **Vol. 2010**, ID 493298, 6 p, 2010.
- [39] I. Schweizer and A. Sklar, *Statistical metric spaces*, Pac. J. Math. **10**, 314-334, 1960.
- [40] T. Suzuki, *Generalized Caristi's fixed point theorems by Bae and others*, J. Math. Anal. Appl. **302**, 502-508, 2005.
- [41] R. Vasuki, *A common fixed point theorem in a fuzzy metric space*, Fuzzy Sets and Systems **97**, 395-397, 1998.
- [42] C. Vetro and P. Vetro, *Common fixed points for discontinuous mappings in fuzzy metric spaces*, Rend. Circ. Mat. Palermo **57**, 295-303, 2008.



The uniform and pointwise estimates for polynomials on the weighted Lebesgue spaces in the general regions of complex plane

P. Özkartepe¹, F.G. Abdullayev^{*2,3}

¹*Gaziantep University Education Faculty in Nizip Department of Mathematics and Science Education, Nizip/Gaziantep-TURKEY.*

²*Kyrgyz-Turkish Manas University, Faculty of Science, Department of Applied Mathematics and Informatics, 720044 Bishkek-KYRGYZ REPUBLIC.*

³*Mersin University, Faculty of Arts and Science, Department of Mathematics, 33343 Mersin-TURKEY.*

Abstract

The estimation of the modulus of algebraic polynomials on the boundary contour with weight function, having some singularities, with respect to the their quasinorm, on the weighted Lebesgue space was studied in this current work.

Mathematics Subject Classification (2010). Primary 30A10, 30C10; Secondary 41A17

Keywords. algebraic polynomials, conformal mapping, quasicircle

1. Introduction

Let \mathbb{C} be a complex plane, and $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$; $G \subset \mathbb{C}$ be the bounded Jordan region, with $0 \in G$ and the boundary $L := \partial G$ be a closed Jordan curve, $\Omega := \overline{\mathbb{C}} \setminus \overline{G} = extL$. Let \wp_n denotes the class of arbitrary algebraic polynomials $P_n(z)$ of degree at most $n \in \mathbb{N} := \{1, 2, \dots\}$.

Let $0 < p \leq \infty$. For a rectifiable Jordan curve L , we denote

$$\|P_n\|_{\mathcal{L}_p} : = \|P_n\|_{\mathcal{L}_p(h,L)} := \left(\int_L h(z) |P_n(z)|^p |dz| \right)^{1/p}, \quad 0 < p < \infty;$$

$$\|P_n\|_{\mathcal{L}_\infty} : = \|P_n\|_{\mathcal{L}_\infty(1,L)} := \max_{z \in L} |P_n(z)|, \quad p = \infty.$$

Clearly, $\|\cdot\|_{\mathcal{L}_p}$ is the quasinorm (i.e. a norm for $1 \leq p \leq \infty$ and a p -norm for $0 < p < 1$).

Denoted by $w = \Phi(z)$, the univalent conformal mapping of Ω onto $\Delta := \{w : |w| > 1\}$ with normalization $\Phi(\infty) = \infty$, $\lim_{z \rightarrow \infty} \frac{\Phi(z)}{z} > 0$ and $\Psi := \Phi^{-1}$. For $t \geq 1$ we set

$$L_t := \{z : |\Phi(z)| = t\}, \quad L_1 \equiv L, \quad G_t := intL_t, \quad \Omega_t := extL_t.$$

*Corresponding Author.

Email addresses: pelinozkartepe@gmail.com (P. Özkartepe), fabdul@mersin.edu.tr; fahreddinabdullayev@gmail.com (F.G. Abdullayev)

Received: 17.11.2016; Accepted: 19.09.2017

Let $\{z_j\}_{j=1}^m$ be a fixed system of distinct points on curve L which is located in the positive direction. For some fixed R_0 , $1 < R_0 < \infty$, and $z \in G_{R_0}$, consider a so-called generalized Jacobi weight function $h(z)$ being defined as follows

$$h(z) := h_0(z) \prod_{j=1}^m |z - z_j|^{\gamma_j}, \quad (1.1)$$

where $\gamma_j > -1$, for all $j = 1, 2, \dots, m$, and h_0 is uniformly separated from zero in G_{R_0} , i.e. there exists a constant $c_0 := c_0(G_{R_0}) > 0$ such that for all $z \in G_{R_0}$

$$h_0(z) \geq c_0 > 0.$$

In many areas, we often need to study the following inequality

$$\|P_n\|_{\mathcal{L}_q(h,L)} \leq c\mu_n(L, h, p, q) \|P_n\|_{\mathcal{L}_p(h,L)}, \quad 0 < p < q \leq \infty, \quad (1.2)$$

where $c = c(G, p, q) > 0$ is the constant, independent from n , P_n , and $\mu_n(L, h, p, q) \rightarrow \infty$, $n \rightarrow \infty$, depending on the geometrical properties of curve L and weight function h in the neighborhood of the points $\{z_j\}_{j=1}^m$. In particular, it was studied the behavior of the $|P_n(z)|$ on L ($q = \infty$), where the boundary curve L and weight function h having some singularity on the L . First result of the (1.2)-type, in case of $h(z) \equiv 1$ and $L = \{z : |z| = 1\}$ for $0 < p < \infty$ was found in [13]. The other results, similar to (1.2), for the sufficiently smooth curve, were obtained in [25] ($h(z) \equiv 1$), and in [27] ($h(z) \neq 1$). The estimations of the (1.2)-type for $0 < p < \infty$ and for $h(z) \equiv 1$ when L is a rectifiable Jordan curve were investigated in [16], [17], [19, pp.122-133], [23], [26] and for $h(z) \neq 1$ in [10, Theorem 6], [1, 2, 4–8] and others, respectively. More references regarding the inequality of the (1.2)-type can also be found cited above and in Milovanovic et al. [18, Sect.5.3].

Let a rectifiable Jordan curve L has a natural parametrization $z = z(s)$, $0 \leq s \leq l := \text{mes}L$. It is said that $L \in C(1, \lambda)$, $0 < \lambda < 1$, if $z(s)$ is continuously differentiable and $z'(s) \in \text{Lip}\lambda$. Let L belongs to $C(1, \lambda)$ everywhere except for a single point $z_1 \in L$, i.e. the derivative $z'(s)$ satisfies the Lipschitz condition on the $[0, l]$ and $z(0) = z(l) = z_1$, but $z'(0) \neq z'(l)$. It is assumed that L has a corner at z_1 with exterior angle $\nu_1\pi$, $0 < \nu_1 \leq 2$, and denoted the set of such curves by $C(1, \lambda, \nu_1)$.

In [28], this problem was investigated in case $p = 2$ for orthonormal on the curve of $L \in C(1, \lambda, \nu_1)$ polynomials $Q_n(z)$ with the weight function h defined in (1.1) and can be shown that the condition of "pay off" singularity curve and weight function at the points z_1 can be given as $(1 + \gamma_1)\nu_1 = 1$. In [28], the case, where $(1 + \gamma_1)\nu_1 \neq 1$ was investigated by the author. It is shown, that if the singularity of a curve and weight function at the points z_1 satisfy the following condition

$$(1 + \gamma_1)\nu_1 < 1, \quad (1.3)$$

then for $|Q_n(z)|$, the following estimation is true

$$|Q_n(z)| \leq c(L) (n^{s_1} + |z - z_1|^{\sigma_1} \sqrt{n}), \quad z \in L, \quad (1.4)$$

where

$$s_1 = \frac{1}{2} (1 + \gamma_1)\nu_1, \quad \sigma_1 = \frac{1}{2} \left(\frac{1}{\nu_1} - 1 - \gamma_1 \right),$$

and $c(L) > 0$ is the constant independent on n .

In this work, the estimations of the (1.2)-type, in particular (1.4)-type for more general curves of the complex plane were studied and the analog of the estimate (1.4) under the condition (1.3) was obtained.

2. Definitions and main results

Throughout this paper, c, c_0, c_1, c_2, \dots are positive and $\varepsilon_0, \varepsilon_1, \varepsilon_2, \dots$ are sufficiently small positive constants (generally, different in different relations), depending on G in general and on parameters inessential for the argument. Otherwise, such dependence will be explicitly stated.

Without loss of generality, the number R_0 in the definition of the weight function, $R_0 = 2$ can be taken. Otherwise, the number n can be chosen as $n \geq \left\lceil \frac{\varepsilon_0}{R_0 - 1} \right\rceil$, where $\varepsilon_0, 0 < \varepsilon_0 < 1$, some fixed small constant.

For any $k \geq 0$ and $m > k$, notation $i = \overline{k, m}$ means $i = k, k + 1, \dots, m$.

Before go any further, some definitions and the notations are needed to be given. Let $z = \psi(w)$ be the univalent conformal mapping of $B := \{w : |w| < 1\}$ onto the G normalized by $\psi(0) = 0, \psi'(0) > 0$.

By [21, pp.286-294], it is said that, a bounded Jordan region G is called a κ -quasidisk, $0 \leq \kappa < 1$, if any conformal mapping ψ can be extended to a K -quasiconformal, $K = \frac{1+\kappa}{1-\kappa}$, homeomorphism of the plane \overline{B} on the plane \overline{G} . In that case, the curve $L := \partial G$ is called a κ -quasicircle. The region G (curve L) is called a quasidisk (quasicircle), if it is κ -quasidisk (κ -quasicircle) for some $0 \leq \kappa < 1$.

We denote the class of κ -quasicircle by $Q(\kappa), 0 \leq \kappa < 1$, and say that $L \in Q$, if $L \in Q(\kappa)$, for some $0 \leq \kappa < 1$.

It is well-known that, the quasicircle may not even be locally rectifiable [14, p.104].

Definition 2.1. We say that $L \in \tilde{Q}, 0 \leq \kappa < 1$, if L is a quasicircle and rectifiable.

Definition 2.2. We say that $L \in Q_\alpha, 0 < \alpha \leq 1$, if L is a quasicircle and $\Phi \in Lip_\alpha, z \in \overline{\Omega}$.

We note that the class Q_α is sufficiently wide. A detailed information on this and the related topics are contained in [15], [22], [30] (see also the references cited therein). We consider only some cases:

Remark 2.3. a) If $L = \partial G$ is a Dini-smooth curve [22, p.48], then $L \in Q_1$.

b) If $L = \partial G$ is a piecewise Dini-smooth curve and largest exterior angle at L has opening $\alpha\pi, 0 < \alpha \leq 1$, [22, p.52], then $L \in Q_\alpha$.

c) If $L = \partial G$ is a smooth curve having continuous tangent line, then $L \in Q_\alpha$ for all $0 < \alpha < 1$.

d) If L is quasismooth (in the sense of Lavrentiev), that is, for every pair $z_1, z_2 \in L, s(z_1, z_2)$ represents the smallest of the lengths of the arcs joining z_1 to z_2 on L , there exists a constant $c > 1$ such that $s(z_1, z_2) \leq c|z_1 - z_2|$, then $\Phi \in Lip_\alpha$ for $\alpha = \frac{1}{2} \left(1 - \frac{1}{c} \arcsin \frac{1}{c}\right)^{-1}$ [30].

e) If L is "c-quasiconformal" (see, for example, [15]), then $\Phi \in Lip_\alpha$ for $\alpha = \frac{\pi}{2(\pi - \arcsin \frac{1}{c})}$. Also, if L is an asymptotic conformal curve, then $\Phi \in Lip_\alpha$ for all $0 < \alpha < 1$ [15].

Definition 2.4. We say that $L \in \tilde{Q}_\alpha, 0 < \alpha \leq 1$, if $L \in Q_\alpha$ and L is rectifiable.

In this case the following can be had:

Theorem A. ([20]) Let $p > 0$. Suppose that $L \in \tilde{Q}_\alpha$, for some $0 < \alpha \leq 1$ and $h(z)$ defined as in (1.1) with $\gamma_j = 0$, for all $j = \overline{1, m}$. Then, for any $P_n \in \wp_n, n \in \mathbb{N}$, there exists $c_1 = c_1(L, p) > 0$ such that

$$\|P_n\|_{\mathcal{L}_\infty} \leq c_1 \|P_n\|_{\mathcal{L}_p(h_0, L)} \begin{cases} n^{\frac{1}{\alpha p}}, & \frac{1}{2} \leq \alpha \leq 1, \\ n^{\frac{\delta}{p}}, & 0 < \alpha < \frac{1}{2}, \end{cases} \quad (2.1)$$

where $\delta = \delta(L), \delta \in [1, 2]$, is a certain number.

Therefore, according to 2.3, α can be calculated in the right parts of estimation (2.1) for each case, respectively.

Now, we assume that the weight function h have "singularity" at the points $\{z_i\}_{i=1}^m$, i.e., $\gamma_i \neq 0$ for all $i = \overline{1, m}$. In this case, the following "local" (at the points $\{z_i\}_{i=1}^m \in L$) estimation is holds:

Theorem 2.5. *Let $p > 0$. Suppose that $L \in \tilde{Q}_\alpha$, for some $\frac{1}{2} \leq \alpha \leq 1$ and $h(z)$ defined as in (1.1). Then, for any $\gamma_i > -1$, $i = \overline{1, m}$, and $P_n \in \wp_n$, $n \in \mathbb{N}$, there exists $c_2 = c_2(L, p, \gamma_i, \alpha) > 0$ such that*

$$|P_n(z_i)| \leq c_2 n^{\frac{\gamma_i+1}{\alpha p}} \|P_n\|_{\mathcal{L}_p(h, L)}. \quad (2.2)$$

The inequality (2.2) is sharp. For the polynomial $P_n^*(z) = 1 + z + \dots + z^n$, $h^*(z) = h_0(z)$, $h^{**}(z) = |z - 1|^\gamma$, $\gamma > 0$, $L := \{z : |z| = 1\}$ and any $n \in \mathbb{N}$, there exists $c_3 = c_3(h^*, p) > 0$, $c'_3 = c'_3(h^{**}, p) > 0$ such that

$$\begin{aligned} a) \quad \|P_n^*\|_{\mathcal{L}_\infty} &\geq c_3 n^{\frac{1}{p}} \|P_n^*\|_{\mathcal{L}_p(h^*, L)}, \quad p > 1; \\ b) \quad \|P_n^*\|_{\mathcal{L}_\infty} &\geq c'_3 n^{\frac{\gamma+1}{p}} \|P_n^*\|_{\mathcal{L}_p(h^{**}, L)}, \quad p > \gamma + 1. \end{aligned}$$

Let's introduce the special "singular" points on the curve L and then lets give the following definition. For $\delta > 0$ and $z \in \mathbb{C}$, let us set $B(z, \delta) := \{\zeta : |\zeta - z| < \delta\}$, $\Omega(z, \delta) := \Omega \cap B(z, \delta)$.

Definition 2.6. We say that $L \in Q_{\alpha, \beta_1, \dots, \beta_m}$, $0 < \beta_i \leq \alpha \leq 1$, $i = \overline{1, m}$, if

i) for every sequence non-crossing in pairs circles $\{D(\zeta_i, \delta_i)\}_{i=1}^m$ restriction of the function Φ on $\Omega(\zeta_i, \delta_i)$ belongs to $Lip\beta_i$ ($\Phi|_{\Omega(\zeta_i, \delta_i)} \in Lip\beta_i$), and restriction

$$\Phi \Big|_{\Omega \setminus \bigcup_{i=1}^m \Omega(\zeta_i, \delta_i)} \in Lip\alpha,$$

ii) there exists a sequence non-crossing in pairs circles $\{D(\zeta_i, \delta_i^*)\}_{i=1}^m$, such that $\forall i = \overline{1, m}$, $\delta_i^* > \delta_i$ and $\forall \xi, z \in \Omega(\zeta_i, \delta_i^*)$, $z \neq \zeta_i \neq \xi$ is fulfilled estimation

$$|\Phi(z) - \Phi(\xi)| \leq k_i(z, \xi) |z - \xi|^\alpha, \quad (2.3)$$

where

$$k_i(z, \xi) = c_i \max \left(|\xi - \zeta_i|^{\beta_i - \alpha}; |z - \zeta_i|^{\beta_i - \alpha} \right),$$

and c_i are independed on z and ξ .

Definition 2.7. We say that $L \in \tilde{Q}_{\alpha, \beta_1, \dots, \beta_m}$, $0 < \beta_i \leq \alpha \leq 1$, $i = \overline{1, m}$, if $L \in Q_{\alpha, \beta_1, \dots, \beta_m}$, $0 < \beta_i \leq \alpha \leq 1$, $i = \overline{1, m}$ and $L = \partial G$ is rectifiable.

It is clear that, from Definition 2.6 (2.7), each region $L \in \tilde{Q}_{\alpha, \beta_1, \dots, \beta_m}$, $0 < \beta_i \leq \alpha \leq 1$, $i = \overline{1, m}$, may have the "singularity" at the points $\{\zeta_i\}_{i=1}^m \in L$. If the region L does not have such "singularity", i.e. if $\beta_i = \alpha$, for all $i = \overline{1, m}$, then, it can be written as $G \in \tilde{Q}_\alpha$, $0 < \alpha \leq 1$.

Throughout this work, we will assume that the points $\{z_i\}_{i=1}^m \in L$ defined in (1.1), and $\{\zeta_i\}_{i=1}^m \in L$ defined in Definitions 2.6, and 2.7 are coincide. Without loss of generality, the points $\{z_i\}_{i=1}^m$ ordered in the positive direction on the curve L .

We assume that the curve L has "singularity" on the boundary points $\{z_i\}_{i=1}^m$, i.e., $\beta_i \neq \alpha$, for all $i = \overline{1, m}$, and the weight function h has "singularity" at the same points, i.e., $\gamma_i \neq 0$ for some $i = \overline{1, m}$. In [20], the following result was shown:

Theorem B. Let $p > 0$. Suppose that $L \in \tilde{Q}_{\alpha, \beta_1, \dots, \beta_m}$, for some $\frac{1}{2} \leq \beta_i \leq \alpha \leq 1$, $i = \overline{1, m}$ and $h(z)$ defined as in (1.1) and

$$\gamma_i + 1 = \frac{\beta_i}{\alpha}, \quad (2.4)$$

for each points $\{z_i\}_{i=1}^m$. Then, for any $P_n \in \wp_n$, $n \in \mathbb{N}$, there exists $c_4 = c_4(L, p, \gamma_i, \alpha) > 0$ such that

$$\|P_n\|_{\mathcal{L}_\infty} \leq c_4 n^{\frac{1}{\alpha p}} \|P_n\|_{\mathcal{L}_p(h, L)}. \quad (2.5)$$

In comparisons of Theorem B and Theorem 2.5, it is seen that, if interference conditions (2.4) are satisfied for any "singular" points $\{z_i\}_{i=1}^m$ of the weight function and the curve L , then the growth rate of the polynomials $P_n(z)$ on L does not depend on whether or not the weight function $h(z)$ and the curve L .

Now, assume that the equality (2.4) does not hold for each singular points $\{z_i\}_{i=1}^m$. Let us suppose

$$\gamma_i + 1 \neq \frac{\beta_i}{\alpha}, \quad (2.6)$$

for each singular point $\{z_i\}_{i=1}^m$. In the present paper the case is investigated

$$\gamma_i + 1 < \frac{\beta_i}{\alpha}.$$

For simplicity of our next calculations, will be taken as $i = 1$.

Theorem 2.8. Let $p > 0$. Suppose that $G \in \tilde{Q}_{\alpha, \beta_1}$, for some $\frac{1}{2} \leq \beta_1 \leq \alpha \leq 1$; $h(z)$ defined by (1.1) and

$$\gamma_1 + 1 < \frac{\beta_1}{\alpha}. \quad (2.7)$$

Then, for every $z \in L$ and $P_n \in \wp_n$, $n \in \mathbb{N}$, there exists $c_5 = c_5(L, p, \gamma_1, \beta_1, \alpha) > 0$, such that

$$|P_n(z)| \leq c_5 (n^{s_1} + |z - z_1|^{\sigma_1} n^{1/p\alpha}) \|P_n\|_{\mathcal{L}_p(h, L)}, \quad (2.8)$$

and

$$|P_n(z_1)| \leq c_6 n^{\frac{\gamma_1+1}{p\beta_1}} \|P_n\|_{\mathcal{L}_p(h, L)}, \quad (2.9)$$

where

$$s_1 = \frac{1 + \gamma_1}{p\beta_1}, \quad \sigma_1 = \frac{\beta_1}{2\alpha} - \frac{1 + \gamma_1}{2}. \quad (2.10)$$

Since $\alpha \geq \beta_1$, (2.7) will be satisfied when $-1 < \gamma_1 < 0$. So, from (2.8) it can be seen that, the order of the height of P_n in point z_1 and points $z \in L$, $z \neq z_1$, where $h(z) \rightarrow \infty$ and curve L not having singularity, acts itself identically. Thus, the conditions (2.7) can be called "algebraic pole" conditions of the order $\lambda_1 = 1 - \frac{\alpha}{\beta_1}(1 + \gamma_1)$.

In case, if L and $h(z)$ have of two singular points, it can be written as

$$|P_n(z)| \leq c_7 (|z - z_1|^{\sigma_1} n^{s_2} + |z - z_2|^{\sigma_2} n^{s_1} + |z - z_1|^{\sigma_1} |z - z_2|^{\sigma_2} n^{1/p\alpha}) \|P_n\|_{\mathcal{L}_p(h, L)}, \quad z \in L, \quad (2.11)$$

where s_i, σ_i , $i = 1, 2$, defined as in (2.10), for $i = 1$ and analogously for $i = 2$, respectively.

Theorem 2.8 is also correct if the curve L has at point z_1 algebraic pole and at points $\{z_k\}$, $k \geq 2$, singularity, in which satisfying the interference conditions (2.4).

Corollary 2.9. If $L \in C(1, \lambda, \nu_1)$, then $L \in \tilde{Q}_{\alpha, \beta_1}$ for $\alpha = 1$ (2.3) and $\beta_1 = \frac{1}{\nu_1}$ [15]. Consequently, if the condition

$$(\gamma_1 + 1) \nu_1 < 1$$

satisfies on the point z_1 , then, for $p = 2$, from (2.8) and (2.10), we have

$$|P_n(z)| \leq (c_6 n^{s_1} + c_7 |z - z_1|^{\sigma_1} \sqrt{n}) \|P_n\|_{\mathcal{L}_2(h, L)}, \quad (2.12)$$

where

$$s_1 = \frac{1}{2}(1 + \gamma_1)\nu_1, \quad \sigma_1 = \frac{1}{2}\left(\frac{1}{\nu_1} - 1 - \gamma_1\right). \quad (2.13)$$

For $P_n \equiv Q_n$, estimation (2.12) coincides from the result by P.K. Suetin in [28, Theorem 2]. Therefore, the Theorem 2.8 generalizes the result in [28, Th2] for $1 \leq \nu_1 \leq 2$ and extends this result to the more general curves of the complex plane.

3. Some auxiliary results

For $a > 0$ and $b > 0$, we will use the notations “ $a \preceq b$ ” (order inequality), if $a \leq cb$ and “ $a \succcurlyeq b$ ” are equivalent to $c_1a \leq b \leq c_2a$ for some constants c, c_1, c_2 (independent of a and b), respectively.

The following definitions of the K -quasiconformal curves are well-known (see, for example, [9], [14, p.97] and [24]).

Definition 3.1. The Jordan arc (or curve) L is called K -quasiconformal ($K \geq 1$), if there is a K -quasiconformal mapping f of the region $D \supset L$ in such that $f(L)$ is a line segment (or circle).

Let $F(L)$ denotes the set of all sense preserving plane homeomorphisms f of the region $D \supset L$ such that $f(L)$ is a line segment (or circle) and lets define

$$K_L := \inf \{K(f) : f \in F(L)\},$$

where $K(f)$ is the maximal dilatation of a such mapping f . L is a quasiconformal curve, if $K_L < \infty$, and L is a K -quasiconformal curve, if $K_L \leq K$.

Remark 3.2. It is well-known that if we are not interested in the coefficients of quasiconformality of the curve, then the definitions of "quasircle" and "quasiconformal curve" are identical whereas, if we are also interested in the coefficients of quasiconformality of the given curve, then a consideration is taken if the curve L is K -quasiconformal; therefore, it is κ -quasircle with $\kappa = \frac{K^2-1}{K^2+1}$.

Following of the Remark 3.2, the both terms will be used for simplicity, depending on the situation.

For $z \in \mathbb{C}$ and $M \subset \mathbb{C}$, we set $\rho(z, M) = \text{dist}(z, M) := \inf \{|z - \zeta| : \zeta \in M\}$.

Lemma 3.3. [3] *Let L be a K -quasiconformal curve, $z_1 \in L$, $z_2, z_3 \in \Omega \cap \{z : |z - z_1| \preceq \rho(z_1, L_{r_0})\}$; $w_j = \Phi(z_j)$, $j = 1, 2, 3$. Then,*

- a) *The statements $|z_1 - z_2| \preceq |z_1 - z_3|$ and $|w_1 - w_2| \preceq |w_1 - w_3|$ are equivalent. So statements $|z_1 - z_2| \succcurlyeq |z_1 - z_3|$ and $|w_1 - w_2| \succcurlyeq |w_1 - w_3|$ also are equivalent.*
- b) *If $|z_1 - z_2| \preceq |z_1 - z_3|$, then*

$$\left| \frac{w_1 - w_3}{w_1 - w_2} \right|^\varepsilon \preceq \left| \frac{z_1 - z_3}{z_1 - z_2} \right| \preceq \left| \frac{w_1 - w_3}{w_1 - w_2} \right|^c,$$

where $\varepsilon < 1$, $c > 1$, $0 < r_0 < 1$ are the constants depending on G and $L_{r_0} := \{z = \psi(w) : |w| = r_0\}$.

Lemma 3.4. *Let $G \in Q(\kappa)$ for some $0 \leq \kappa < 1$. Then,*

$$|\Psi(w_1) - \Psi(w_2)| \succeq |w_1 - w_2|^{1+\kappa},$$

for all $w_1, w_2 \in \bar{\Delta}$.

This fact follows from [21, p.287, Lemma 9.9] and estimation for the Ψ' (see, for example, [11, Th.2.8]):

$$|\Psi'(\tau)| \asymp \frac{\rho(\Psi(\tau), L)}{|\tau| - 1}. \quad (3.1)$$

Let $\{z_j\}_{j=1}^m$ be the fixed system of the points on L and the weight function $h(z)$ defined as (1.1).

Lemma 3.5. [8] Let L be a rectifiable Jordan curve and $h(z)$ is defined as in (1.1). Then, for arbitrary $P_n(z) \in \wp_n$, any $R > 1$ and $n \in \mathbb{N}$, we have

$$\|P_n\|_{\mathcal{L}_p(h, L_R)} \leq R^{n + \frac{1+\gamma^*}{p}} \|P_n\|_{\mathcal{L}_p(h, L)}, \quad p > 0. \quad (3.2)$$

Remark 3.6. In case of $h(z) \equiv 1$, the estimation (3.2) has been proved in [12].

4. Proofs of theorems

4.1. Proof of Theorem 2.5

Proof. Suppose that $L \in \tilde{Q}_\alpha$, for some $\frac{1}{2} \leq \alpha \leq 1$, $i = \overline{1, m}$ be given and $h(z)$ is defined as in (1.1). Let $w = \varphi_R(z)$ be the univalent conformal mapping of G_R , $R > 1$, onto the B , normalized by $\varphi_R(0) = 0$, $\varphi'_R(0) > 0$, and let $\{\xi_j\}$, $1 \leq j \leq m \leq n$, be zeros of $P_n(z)$ lying on G_R . Let

$$B_{m,R}(z) := \prod_{j=1}^m B_{j,R}(z) = \prod_{j=1}^m \frac{\varphi_R(z) - \varphi_R(\xi_j)}{1 - \overline{\varphi_R(\xi_j)} \varphi_R(z)}, \quad (4.1)$$

denotes the Blaschke function with respect to zeros $\{\xi_j\}$, $1 \leq j \leq m \leq n$, of $P_n(z)$. Clearly,

$$|B_{m,R}(z)| \equiv 1, \quad z \in L_R, \quad (4.2)$$

and

$$|B_{m,R}(z)| < 1, \quad z \in G_R. \quad (4.3)$$

For any $p > 0$ and $z \in G_R$, let us set:

$$T_n(z) := \left[\frac{P_n(z)}{B_{m,R}(z)} \right]^{p/2}. \quad (4.4)$$

The function $T_n(z)$ is analytic in G_R and continuous on \overline{G}_R , not having zeros in G_R . We take an arbitrary continuous branch of the $T_n(z)$ and we maintain the same designation for this branch. Then, the Cauchy integral representation for the $T_n(z)$ in G_R gives

$$T_n(z) = \frac{1}{2\pi i} \int_{L_R} T_n(\zeta) \frac{d\zeta}{\zeta - z}, \quad z \in G_R, \quad (4.5)$$

and then

$$T_n(z_j) = \frac{1}{2\pi i} \int_{L_R} T_n(\zeta) \frac{d\zeta}{\zeta - z_j}.$$

Now, let $z \in L$. Multiplying the numerator and the denominator of the integrand by $h^{1/2}(\zeta)$, according to the Hölder inequality, from (4.2) and (4.3), we obtain:

$$\begin{aligned} |P_n(z_j)| &\leq \left(\frac{1}{2\pi} \right)^{\frac{2}{p}} \left(\int_{L_R} h(\zeta) |P_n(\zeta)|^p |d\zeta| \right)^{1/p} \\ &\times \left(\int_{L_R} \frac{|d\zeta|}{\prod_{j=1}^m |\zeta - z_j|^{2+\gamma_j}} \right)^{1/p} =: \left(\frac{1}{2\pi} \right)^{\frac{2}{p}} J_{n,1} \times J_{n,2}, \end{aligned} \quad (4.6)$$

where

$$J_{n,1} := \|P_n\|_{\mathcal{L}_p(h, L_R)}, \quad J_{n,2} := \left(\int_{L_R} \frac{|d\zeta|}{\prod_{j=1}^m |\zeta - z_j|^{2+\gamma_j}} \right)^{1/p}.$$

Then, from Lemma 3.5, for the each points $\{z_j\}_{j=1}^m \in L$, we have:

$$|P_n(z_j)| \preceq \|P_n\|_{\mathcal{L}_p} \cdot (J_{n,2})^{1/p}. \quad (4.7)$$

To estimate the integral $J_{n,2}$, we introduce

$$w_j := \Phi(z_j), \quad \varphi_j := \arg w_j, \quad L^j := L \cap \overline{\Omega}^j, \quad L_R^j := L_R \cap \overline{\Omega}^j, \quad j = \overline{1, m},$$

where $\Omega^j := \Psi(\Delta'_j)$

$$\begin{aligned} \Delta'_1 &:= \left\{ t = Re^{i\theta} : R > 1, \frac{\varphi_m + \varphi_1}{2} \leq \theta < \frac{\varphi_1 + \varphi_2}{2} \right\}, \\ \Delta'_m &:= \left\{ t = Re^{i\theta} : R > 1, \frac{\varphi_{m-1} + \varphi_m}{2} \leq \theta < \frac{\varphi_m + \varphi_1}{2} \right\}, \end{aligned}$$

and, for $j = \overline{2, m-1}$

$$\Delta'_j := \left\{ t = Re^{i\theta} : R > 1, \frac{\varphi_{j-1} + \varphi_j}{2} \leq \theta < \frac{\varphi_j + \varphi_{j+1}}{2} \right\}.$$

Since the points $\{z_j\}_{j=1}^m \in L$ are distinct, we get:

$$(J_{n,2})^p = \sum_{i=1}^m \int_{L_R^i} \frac{|d\zeta|}{\prod_{j=1}^m |\zeta - z_j|^{2+\gamma_j}} \asymp \sum_{i=1}^m \int_{L_R^i} \frac{|d\zeta|}{|\zeta - z_i|^{2+\gamma_i}} =: \sum_{i=1}^m J_{n,2}^i, \quad (4.8)$$

where

$$J_{n,2}^i := \int_{L_R^i} \frac{|d\zeta|}{|\zeta - z_i|^{2+\gamma_i}}, \quad i = \overline{1, m}. \quad (4.9)$$

It remains to estimate the integrals $J_{n,2}^i$ for each $i = \overline{1, m}$. For the simplicity of our next calculations, we assume that

$$m = 1; \quad R = 1 + \frac{\varepsilon_0}{n}. \quad (4.10)$$

Let the numbers δ_1, δ_1^* , $0 < \delta_1 < \delta_1^* < \delta_0 < \text{diam } \overline{G}$, are chosen from the Definition 2.6. By denoting

$$l_{R,1}^1 := L_R^1 \cap \Omega(z_1, \delta_1), \quad l_{R,2}^1 := L_R^1 \setminus l_{R,1}^1, \quad F_{R,i}^1 := \Phi(l_{R,i}^1), \quad i = 1, 2,$$

from (4.9), we get

$$J_{n,2}^1 := \int_{L_R^1} \frac{|d\zeta|}{|\zeta - z_1|^{2+\gamma_1}} = \int_{l_{R,1}^1} \frac{|d\zeta|}{|\zeta - z_1|^{2+\gamma_1}} + \int_{l_{R,2}^1} \frac{|d\zeta|}{|\zeta - z_1|^{2+\gamma_1}}. \quad (4.11)$$

Then, by applying Lemma 3.3, we have:

$$\begin{aligned} \int_{l_{R,1}^1} \frac{|d\zeta|}{|\zeta - z_1|^{2+\gamma_1}} &= \int_{\Phi(l_{R,1}^1)} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w_1)|^{2+\gamma_1} (|\tau| - 1)} \\ &\asymp \int_{\Phi(l_{R,1}^1)} \frac{|d\tau|}{|\Psi(\tau) - \Psi(w_1)|^{1+\gamma_1} (|\tau| - 1)} \\ &\asymp n \int_{\Phi(l_{R,1}^1)} \frac{|d\tau|}{|\tau - w_1|^{\frac{\gamma_1+1}{\alpha}}} \preceq n^{\frac{\gamma_1+1}{\alpha}}, \end{aligned} \quad (4.12)$$

$$\int_{l_{R,2}^1} \frac{|d\zeta|}{|\zeta - z_1|^{2+\gamma_1}} \preceq (\delta_1)^{2+\gamma_1} \text{mes} l_{R,1}^1 \preceq 1. \quad (4.13)$$

Then from (4.11)-(4.13), we get:

$$J_{n,2}^1 \preceq n^{\frac{\gamma_1+1}{\alpha}}. \quad (4.14)$$

By combining (4.7)-(4.14), we obtain:

$$|P_n(z_1)| \preceq n^{\frac{\gamma_1+1}{p\alpha}} \|P_n\|_{\mathcal{L}_p},$$

and, according to (4.10), the proof is completed. \square

4.2. Proof of Theorem 2.8

Proof. Suppose that $L \in \tilde{Q}_{\alpha,\beta_1}$, for some $\frac{1}{2} \leq \beta_1 \leq \alpha \leq 1$, be given and $h(z)$ defined as in (1.1). Let the functions $B_{m,R}(z)$ and $T_R(z)$ be constructed as in the beginning to the proof of Theorem 2.5 by (4.1)- (4.4). By taking the arbitrary fixed branch of $T_R(z)$, for any $z \in G_R$ from (4.5), we get:

$$\begin{aligned} T_n(z) - T_n(z_1) &= \frac{1}{2\pi i} \int_{L_R} T_n(\zeta) \left[\frac{1}{\zeta - z} - \frac{1}{\zeta - z_1} \right] d\zeta, \\ &= \frac{1}{2\pi i} \int_{L_R} \left[\frac{P_n(\zeta)}{B_{m,R}(\zeta)} \right]^{p/2} \left[\frac{z - z_1}{(\zeta - z)(\zeta - z_1)} \right] d\zeta. \end{aligned}$$

For arbitrary $z \in L$, $z \neq z_1$, multiplying both sides of the equality by $(z - z_1)^{-\sigma_1}$, we obtain:

$$\left| \frac{T_n(z) - T_n(z_1)}{(z - z_1)^{\sigma_1}} \right| \leq \frac{1}{2\pi} \int_{L_R} |P_n(\zeta)|^{p/2} \left| \frac{(z - z_1)^{1-\sigma_1}}{(\zeta - z)(\zeta - z_1)} \right| |d\zeta|,$$

since $|B_{m,R}(\zeta)| = 1$, for $\zeta \in L_R$. By multiplying the numerator and the denominator of the integrand via $h^{1/2}(\zeta)$ and by applying the Hölder inequality, we obtain:

$$\begin{aligned} \left| \frac{T_n(z) - T_n(z_1)}{(z - z_1)^{\sigma_1}} \right| &\leq \frac{1}{2\pi} \left(\int_{L_R} h(\zeta) |P_n(\zeta)|^p |d\zeta| \right)^{1/2} \\ &\times \left(\int_{L_R} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \right)^{1/2} =: \frac{1}{2\pi} (J_{n,1} \times J_{n,2})^{1/2}, \end{aligned}$$

where

$$\begin{aligned} J_{n,1} &: = \int_{L_R} h(\zeta) |P_n(\zeta)|^p |d\zeta| = \|P_n\|_{\mathcal{L}_p(h,L_R)}^p, \\ J_{n,2} &: = \int_{L_R} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta|. \end{aligned} \quad (4.15)$$

Then, for $z \in L$, from Lemma 3.5, we have:

$$\left| \frac{T_n(z) - T_n(z_1)}{(z - z_1)^{\sigma_1}} \right| \preceq \|P_n\|_{\mathcal{L}_p}^{\frac{p}{2}} \cdot (J_{n,2})^{1/2}, \quad z \in L \setminus \{z_1\}.$$

From (4.4), we obtain:

$$\left| \frac{P_n(z)}{B_{m,R}(z)} \right|^{p/2} \leq c_9 \left| \frac{P_n(z_1)}{B_{m,R}(z_1)} \right|^{p/2} + c_{10} \|P_n\|_{\mathcal{L}_p}^{\frac{p}{2}} \cdot (J_{n,2})^{1/2}. \quad (4.16)$$

According to well-known inequalities [29, p.121]

$$\begin{aligned} |A+B|^p &\leq 2^{p-1}(|A|^p + |B|^p), \quad p > 1, \\ |A+B|^p &\leq |A|^p + |B|^p, \quad 0 < p \leq 1, \quad A > 0, \quad B > 0, \end{aligned} \quad (4.17)$$

from (4.16), we obtain:

$$|P_n(z)| \leq c_{11} \left| \frac{B_{m,R}(z)}{B_{m,R}(z_1)} \right| |P_n(z_1)| + c_{12} \|P_n\|_{\mathcal{L}_p} \cdot (J_{n,2})^{1/p}. \quad (4.18)$$

Since $|B_{m,R}(z)| < 1$, for $z \in L$ and $|B_{m,R}(\zeta)| = 1$, for $\zeta \in L_R$, then there exists ε_1 , where $0 < \varepsilon_1 < 1$, in such that the following is fulfilled:

$$|B_{m,R}(z_1)| > 1 - \varepsilon_1. \quad (4.19)$$

Then, from (4.18) and (4.19), for each $z \in L \setminus \{z_1\}$, we have

$$|P_n(z)| \leq c_{13} |P_n(z_1)| + c_{12} \|P_n\|_{\mathcal{L}_p} \cdot (J_{n,2})^{1/p}, \quad p > 0. \quad (4.20)$$

By the Theorem 2.5, the estimation for $|P_n(z_1)|$ is known. Therefore, finding of an estimation of $J_{n,2}$

$$J_{n,2} = \int_{L_R} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta|.$$

is needed for completion. We set:

$$\begin{aligned} L_{R,1}^1 &:= L_R^1 \cap \Omega(z_1, \delta_1), \quad L_{R,2}^1 := L_R^1 \cap (\Omega(z_1, \delta_1^*) \setminus \Omega(z_1, \delta_1)), \\ L_{R,3}^1 &:= L_R \setminus (L_{R,1}^1 \cup L_{R,2}^1); \quad F_{R,j}^1 := \Phi(L_{R,j}^1), \\ L_1^1 &:= L^1 \cap D(z_1, \delta_1), \quad L_2^1 := L^1 \cap (D(z_1, \delta_1^*) \setminus D(z_1, \delta_1)), \\ L_3^1 &:= L \setminus (L_1^1 \cup L_2^1); \quad F_i^1 := \Phi(L_i^1), \quad i = 1, 2, 3, \end{aligned}$$

then, from (4.15), we get:

$$J_{n,2} = \sum_{i=1}^3 J_{n,2}(L_{R,i}^1), \quad (4.21)$$

where

$$J_{n,2}(l) := \int_l \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta|, \quad (4.22)$$

for $l \subset L$. There are two possible cases: the point z may lie on L^1 or L^2 . Suppose first, $z \in L^1$. If $z \in L_i^1$, then $w \in F_i^1$, for $i = 1, 2, 3$. Consider the individual cases:

1) Let $z \in L_1^1$.

1.1) By applying (4.17), we have:

$$\begin{aligned} J_{n,2}(L_{R,1}^1) &= \int_{L_{R,1}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\ &\preceq \int_{L_{R,1}^1} \frac{[|\zeta - z| + |\zeta - z_1|]^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\ &= \int_{L_{R,1}^1} \frac{|d\zeta|}{|\zeta - z|^{2\sigma_1} |\zeta - z_1|^{2+\gamma_1}} + \int_{L_{R,1}^1} \frac{|d\zeta|}{|\zeta - z|^2 |\zeta - z_1|^{2\sigma_1 + \gamma_1}}. \end{aligned} \quad (4.23)$$

Lets

$$\begin{aligned} L_{R,j}^{1,1} &:= \left\{ \zeta \in L_{R,j}^1 : |\zeta - z_1| \geq |\zeta - z| \right\}, \\ L_{R,j}^{1,2} &:= L_{R,j}^1 \setminus L_{R,j}^{1,1}, \quad F_{R,j}^{1,i} := \Phi(L_{R,j}^{1,i}), \quad i, j = 1, 2. \end{aligned}$$

Then, from (4.23), we get:

$$\begin{aligned}
 J_{n,2}(L_{R,1}^1) &\preceq \int_{L_{R,2}^{1,1}} \frac{|d\zeta|}{|\zeta - z|^{2\sigma_1+2+\gamma_1}} + \int_{L_{R,2}^{1,2}} \frac{|d\zeta|}{|\zeta - z_1|^{2\sigma_1+2+\gamma_1}} \\
 &\preceq \int_{F_{R,2}^{1,1}} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^{2\sigma_1+2+\gamma_1} (|\tau| - 1)} \\
 &\quad + \int_{F_{R,2}^{1,2}} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^{2\sigma_1+2+\gamma_1} (|\tau| - 1)} \\
 &\preceq n \int_{F_{R,1}^{1,1}} \frac{|d\tau|}{|\tau - w|^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}}} + n \int_{F_{R,1}^{1,2}} \frac{|d\tau|}{|\tau - w|^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}}} \\
 &\preceq n \cdot n^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}-1} + n \cdot n^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}-1} \preceq n^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}}.
 \end{aligned} \tag{4.24}$$

1.2) For any $\zeta \in L_{R,2}^1$ and $z \in L_1^1$, $|\zeta - z_1| \geq \delta_1$ and by Definition 2.6, we obtain:

$$\begin{aligned}
 |\zeta - z|^\alpha &\succeq \max \left\{ |\zeta - z_1|^{\alpha-\beta_1}; |z - z_1|^{\alpha-\beta_1} \right\} |w - \tau| \\
 &= |\zeta - z_1|^{\alpha-\beta_1} |w - \tau| \geq \delta_1^{\alpha-\beta_1} |w - \tau| \succeq |w - \tau|.
 \end{aligned}$$

Then, for this case we get:

$$\begin{aligned}
 J_{n,2}(L_{R,2}^1) &= \int_{L_{R,2}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\
 &\preceq \int_{L_{R,2}^1} \frac{[|\zeta - z| + |\zeta - z_1|]^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\
 &= \int_{L_{R,2}^1} \frac{|d\zeta|}{|\zeta - z|^{2\sigma_1} |\zeta - z_1|^{2+\gamma_1}} + \int_{L_{R,2}^1} \frac{|d\zeta|}{|\zeta - z|^2 |\zeta - z_1|^{2\sigma_1+\gamma_1}} \\
 &\preceq \frac{1}{\delta_1^{2+\gamma_1}} \int_{L_{R,2}^1} \frac{|d\zeta|}{|\zeta - z|^{2\sigma_1}} + \frac{1}{\delta_1^{2\sigma_1+\gamma_1}} \int_{L_{R,2}^1} \frac{|d\zeta|}{|\zeta - z|^2} \\
 &\preceq \int_{F_{R,2}^1} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^{2\sigma_1} (|\tau| - 1)} \\
 &\quad + \int_{F_{R,2}^1} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^2 (|\tau| - 1)} \\
 &\preceq n \int_{F_{R,1}^{1,1}} \frac{|d\tau|}{|\tau - w|^{\frac{2\sigma_1-1}{\alpha}}} + n \int_{F_{R,1}^{1,2}} \frac{|d\tau|}{|\tau - w|^{\frac{1}{\alpha}}} \\
 &\preceq n \cdot n^{\frac{2\sigma_1-1}{\alpha}-1} + n \cdot n^{\frac{1}{\alpha}-1} \preceq n^{\frac{1}{\alpha}}.
 \end{aligned} \tag{4.25}$$

1.3) For any $\zeta \in L_{R,3}^1$ and $z \in L_1^1$, $|\zeta - z_1| \geq \delta_1^*$ and $|\zeta - z| \geq \delta_1^* - \delta_1$. Then, we have

$$\begin{aligned} J_{n,2}(L_{R,3}^1) &= \int_{L_{R,3}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\ &\leq \frac{(\text{diam} \overline{G_R})^{2-2\sigma_1} \text{mes} L}{(\delta_1^* - \delta_1)^{2-2\sigma_1} (\delta_1^*)^2} \leq 1. \end{aligned} \quad (4.26)$$

2) Let $z \in L_2^1$.

2.1) According to $|\zeta - z_1| < |z - z_1|$, by Definition 2.6, we obtain:

$$\begin{aligned} |\zeta - z|^\alpha &\succeq \max \left\{ |\zeta - z_1|^{\alpha-\beta_1}; |z - z_1|^{\alpha-\beta_1} \right\} |w - \tau| \\ &= |z - z_1|^{\alpha-\beta_1} |w - \tau| \geq \delta_1^{\alpha-\beta_1} |w - \tau| \succeq |w - \tau|. \end{aligned}$$

Then, we get:

$$\begin{aligned} J_{n,2}(L_{R,1}^1) &= \int_{L_{R,1}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\ &\preceq \int_{L_{R,1}^1} \frac{[|\zeta - z| + |\zeta - z_1|]^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\ &= \int_{L_{R,1}^1} \frac{|d\zeta|}{|\zeta - z|^{2\sigma_1} |\zeta - z_1|^{2+\gamma_1}} + \int_{L_{R,1}^1} \frac{|d\zeta|}{|\zeta - z|^2 |\zeta - z_1|^{2\sigma_1+\gamma_1}} \\ &\preceq \int_{L_{R,1}^{1,1}} \frac{|d\zeta|}{|\zeta - z|^{2\sigma_1+2+\gamma_1}} + \int_{L_{R,1}^{1,2}} \frac{|d\zeta|}{|\zeta - z_1|^{2\sigma_1+2+\gamma_1}} \\ &\preceq \int_{F_{R,2}^{1,1}} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^{2\sigma_1+2+\gamma_1} (|\tau| - 1)} + \int_{F_{R,2}^{1,2}} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w_1)|^{2\sigma_1+2+\gamma_1} (|\tau| - 1)} \\ &\preceq n \int_{F_{R,1}^{1,1}} \frac{|d\tau|}{|\tau - w|^{\frac{2\sigma_1+1+\gamma_1}{\alpha}}} + n \int_{F_{R,1}^{1,2}} \frac{|d\tau|}{|\tau - w_1|^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}}} \\ &\preceq n \cdot n^{\frac{2\sigma_1+1+\gamma_1}{\alpha}-1} + n \cdot n^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}-1} \preceq n^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}}. \end{aligned}$$

Therefore, in this case we have

$$J_{n,2}(L_{R,1}^1) \preceq n^{\frac{2\sigma_1+1+\gamma_1}{\beta_1}}. \quad (4.27)$$

2.2) For any $\zeta \in L_{R,2}^1$ and $z \in L_2^1$, $|\zeta - z_1| \geq \delta_1$ and analogously to the case 1.2), in this case from Definition 2.6, we obtain

$$\begin{aligned} |\zeta - z|^\alpha &\succeq \max \left\{ |\zeta - z_1|^{\alpha-\beta_1}; |z - z_1|^{\alpha-\beta_1} \right\} |w - \tau| \\ &\geq \delta_1^{\alpha-\beta_1} |w - \tau| \succeq |w - \tau|, \end{aligned}$$

and then,

$$J_{n,2}(L_{R,2}^1) = \int_{L_{R,1}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta|$$

$$\begin{aligned}
 & \asymp \frac{(\delta_1^*)^{2-2\sigma_1}}{(\delta_1)^{2+\gamma_1}} \int_{L_{R,1}^1} \frac{|d\zeta|}{|\zeta - z|^2} \tag{4.28} \\
 & \asymp \int_{F_{R,2}^1} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^2 (|\tau| - 1)} \preceq n \int_{F_{R,2}^1} \frac{|d\tau|}{|\Psi(\tau) - \Psi(w)|} \\
 & \asymp n \int_{F_{R,2}^1} \frac{|d\tau|}{|\tau - w|^{\frac{1}{\alpha}}} \preceq n^{\frac{1}{\alpha}}.
 \end{aligned}$$

2.3) For any $\zeta \in L_{R,3}^1$ and $z \in L_2^1$, $|\zeta - z_1| \geq \delta_1^*$ and $|\zeta - z| \geq \delta_1^* - \delta_1$. Then, we have:

$$\begin{aligned}
 J_{n,2}(L_{R,3}^1) &= \int_{L_{R,3}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\
 & \asymp \frac{(\delta_1^*)^{2-2\sigma_1}}{(\delta_1)^{2+\gamma_1}} \int_{L_{R,3}^1} \frac{|d\zeta|}{|\zeta - z|^2} \asymp \int_{L_{R,3}^1} \frac{|d\zeta|}{|\zeta - z|^2} \tag{4.29} \\
 & \asymp \int_{F_{R,3}^1} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^2 (|\tau| - 1)} \preceq n \int_{F_{R,3}^1} \frac{|d\tau|}{|\Psi(\tau) - \Psi(w)|} \\
 & \asymp n \int_{F_{R,3}^1} \frac{|d\tau|}{|\tau - w|^{\frac{1}{\alpha}}} \preceq n^{\frac{1}{\alpha}}.
 \end{aligned}$$

3) Let $z \in L_3^1$.

3.1) Analogously to previous cases, we get:

$$\begin{aligned}
 J_{n,2}(L_{R,1}^1) &= \int_{L_{R,1}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\
 & \leq \frac{(2\text{diam}\overline{G})^{2-2\sigma_1}}{(\delta_1^* - \delta_1)^2} \int_{L_{R,1}^1} \frac{|d\zeta|}{|\zeta - z_1|^{2+\gamma_1}} \tag{4.30} \\
 & \asymp \int_{F_{R,1}^1} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^{2+\gamma_1} (|\tau| - 1)} \\
 & \asymp n \int_{F_{R,1}^1} \frac{|d\tau|}{|\tau - w|^{\frac{2+\gamma_1}{\beta_1}}} \preceq n^{\frac{2+\gamma_1}{\beta_1}}.
 \end{aligned}$$

3.2) Since, from Definition 2.6,

$$\begin{aligned}
 |\zeta - z|^\alpha & \succeq \max \left\{ |\zeta - z_1|^{\alpha-\beta_1}; |z - z_1|^{\alpha-\beta_1} \right\} |w - \tau| \\
 & \geq \delta_1^{\alpha-\beta_1} |w' - \tau| \succeq |w - \tau|,
 \end{aligned}$$

then, we have:

$$J_{n,2}(L_{R,2}^1) = \int_{L_{R,2}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta|$$

$$\begin{aligned}
&\leq \frac{(\delta_1^*)^{2-2\sigma_1}}{(\delta_1)^{2+\gamma_1}} \int_{L_{R,2}^1} \frac{|d\zeta|}{|\zeta - z|^2} \tag{4.31} \\
&\preceq \int_{F_{R,2}^1} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^2 (|\tau| - 1)} \preceq n \int_{F_{R,2}^1} \frac{|d\tau|}{|\Psi(\tau) - \Psi(w)|} \\
&\preceq n \int_{F_{R,2}^1} \frac{|d\tau|}{|\tau - w|^{\frac{1}{\alpha}}} \preceq n^{\frac{1}{\alpha}}.
\end{aligned}$$

3.3) Also, we get:

$$\begin{aligned}
J_{n,2}(L_{R,3}^1) &= \int_{L_{R,3}^1} \frac{|z - z_1|^{2-2\sigma_1}}{|\zeta - z|^2 |\zeta - z_1|^{2+\gamma_1}} |d\zeta| \\
&\leq (\delta_1^*)^{-2\sigma_1 - \gamma_1} \int_{L_{R,3}^1} \frac{|d\zeta|}{|\zeta - z|^2} \tag{4.32} \\
&\preceq \int_{F_{R,3}^1} \frac{\rho(\Psi(\tau), L) |d\tau|}{|\Psi(\tau) - \Psi(w)|^2 (|\tau| - 1)} \preceq n \int_{F_{R,3}^1} \frac{|d\tau|}{|\Psi(\tau) - \Psi(w)|} \\
&\preceq n \int_{F_{R,3}^1} \frac{|d\tau|}{|\tau - w|^{\frac{1}{\alpha}}} \preceq n^{\frac{1}{\alpha}}.
\end{aligned}$$

By combining the estimations of (4.18), (4.20)-(4.32), finally we obtain:

$$|P_n(z)| \leq c_{13} |P_n(z_1)| + c_{12} \|P_n\|_{\mathcal{L}_p} \cdot n^{\frac{1}{\alpha p}}, \quad p > 0.$$

and, then the proof of (2.8) is completed.

The estimation of (2.9) follows from the Theorem 2.5 for $\alpha = \beta_1$. \square

Acknowledgment. This work is supported by TUBITAK under Project No: 115F652.

References

- [1] F.G. Abdullayev, *On the some properties on orthogonal polynomials over the regions of complex plane 1*. Ukr. Math. J. **52** (12), 1807-1817, 2000.
- [2] F.G. Abdullayev, *On the interference of the weight boundary contour orthogonal polynomials over the region*, J. of Comp. Anal. and Appl. **6** (1), 31-42, 2004.
- [3] F.G. Abdullayev and V.V. Andrievskii, *On the orthogonal polynomials in the domains with K -quasiconformal boundary*, Izv. Akad. Nauk Azerb. SSR., Ser. FTM, **1**, 3-7, 1983.
- [4] F.G. Abdullayev and C.D. Gün, *On the behavior of the algebraic polynomials in regions with piecewise smooth boundary without cusps*, Ann. Polon. Math. **111**, 39-58, 2014.
- [5] F.G. Abdullayev and N.P. Özkartepe, *On the Behavior of the Algebraic Polynomial in Unbounded Regions with Piecewise Dini -Smooth Boundary*, Ukr. Math. J., **66** (5), 579-597, 2014.
- [6] F.G. Abdullayev and N.P. Özkartepe, *On the growth of algebraic polynomials in the whole complex plane*, J. Korean Math. Soc. **52** (4) 699-725, 2015.
- [7] F.G. Abdullayev and N.P. Özkartepe, *Uniform and pointwise polynomial inequalities in regions with cusps in the weighted Lebesgue space*, Jaen J. Approx. **7** (2), 231-261, 2015.

- [8] F.G. Abdullayev, N.P. Özkartepe, C.D. Gün, *Uniform and pointwise polynomial inequalities in regions without cusps in the weighted Lebesgue space*, Bulletin of Tbilisi ICMC **18** (1), 146-167, 2014.
- [9] L. Ahlfors, *Lectures on Quasiconformal Mappings*, Princeton, NJ: Van Nostrand, 1966.
- [10] V.V. Andrievskii, *Weighted Polynomial Inequalities in the Complex Plane*, J. Approx. Theory **164**(9), 1165-1183, 2012.
- [11] V.V. Andrievskii, V.I. Belyi, and V.K. Dzyadyk, *Conformal invariants in constructive theory of functions of complex plane*, Atlanta: World Federation Publ.Com., 1995.
- [12] E. Hille, G. Szegö, and J.D. Tamarkin, *On some generalization of a theorem of A. Markoff*, Duke Math. **3**, 729-739, 1937.
- [13] D. Jackson, *Certain problems on closest approximations*, Bull. Amer. Math. Soc. **39**, 889-906, 1933.
- [14] O. Lehto and K.I. Virtanen, *Quasiconformal Mapping in the Plane*, Springer Verlag, Berlin, 1973.
- [15] F.D. Lesley, *Hölder continuity of conformal mappings at the boundary via the strip method*, Indiana Univ. Math. J. **31**, 341-354, 1982.
- [16] D.I. Mamedhanov, *Inequalities of S.M.Nikol'skii type for polynomials in the complex variable on curves*, Soviet Math. Dokl. **15**, 34-37, 1974.
- [17] D.I. Mamedhanov, *On Nikol'skii-type inequalities with new characteristic*, Dokl. Math. **82**, 882-883, 2010.
- [18] G.V. Milovanovic, D.S. Mitrinovic, and Th.M. Rassias, *Topics in Polynomials: Extremal Problems, Inequalities, Zeros*, World Scientific, Singapore, 1994.
- [19] S.M. Nikol'skii, *Approximation of Function of Several Variable and Imbedding Theorems*, Springer-Verlag, New-York, 1975.
- [20] N.P. Özkartepe and F.G. Abdullayev, *Interference of the weight and boundary contour for algebraic polynomials in the weighted Lebesgue spaces I*, Ukr. Math. J. **68**(10), 1574-1590, 2017.
- [21] Ch. Pommerenke, *Univalent Functions*, Göttingen, Vandenhoeck & Ruprecht, 1975.
- [22] Ch. Pommerenke, *Boundary Behavior of Conformal Maps*, Springer-Verlag, Berlin, 1992.
- [23] I. Pritsker, *Comparing Norms of Polynomials in One and Several Variables*, J. Math. Anal. Appl. **216**, 685-695, 1997.
- [24] S. Rickman, *Characterisation of quasiconformal arcs*, Ann. Acad. Sci. Fenn., Ser. A, Mathematica **395**, 30 p, 1966.
- [25] G. Szegö and A. Zigmund, *On certain mean values of polynomials*, J. Anal. Math. **3**, 225-244, 1954.
- [26] P.K. Suetin, *The ordinally comparison of various norms of polynomials in the complex domain*, Matematicheskie zapiski Uralskogo Gos. Universiteta **5** (4), 1966 (in Russian).
- [27] P.K. Suetin, *Main properties of the orthogonal polynomials along a circle*, Uspekhi Math. Nauk **21** (2 (128)), 41-88, 1966.
- [28] P.K. Suetin, *On some estimates of the orthogonal polynomials with singularities weight and contour*, Sib. Math. J **VIII** (3), 1070-1078, 1967 (in Russian).
- [29] J.L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, AMS, 1960.
- [30] S.E. Warschawski, *On differentiability at the boundary in conformal mapping*, Proc. Amer. Math. Soc., **12**, 614-620, 1961.



A note on the embedding properties of p -subgroups in finite groups

Boru Zhang, Xiuyun Guo*

Department of Mathematics, Shanghai University, Shanghai 200444, P. R. China

Abstract

In this note, we prove that a finite group G is p -supersolvable if and only if there exists a power d of p with $p^2 \leq d < |P|$ such that $H \cap O^p(G_p^*)$ is normal in $O^p(G)$ for all non-cyclic normal subgroups H of P with $|H| = d$, where P is a Sylow p -subgroup of G . Moreover, we also prove that either $l_p(G) \leq 1$ and $r_p(G) \leq 2$ or else $|P \cap O^p(G)| > d$ if there exists a power d of p with $1 \leq d < |P|$ such that $H \cap O^p(G_{p^2}^*)$ is normal in $O^p(G)$ for all non-meta-cyclic normal subgroups H of P with $|H| = d$.

Mathematics Subject Classification (2010). 20D10, 20D20, 20D40

Keywords. Finite p -group of maximal class, p -supersolvable group, meta-cyclic p -group

1. Introduction

All groups considered in this note are finite. We use conventional notions and notation, as in [9].

It is quite interesting to investigate the structure of a group by using some kind of the embedding properties of subgroups and many interesting results have been given (for example, see [1, 6, 8, 13]). Recently, Guo and Isaacs [6] proved the following theorem by using the embedding condition “ $H \cap O^p(G) \trianglelefteq O^p(G)$ ”.

Theorem 1.1. ([6, Theorem B]). *Let $P \in \text{Syl}_p(G)$, and let d be a power of p such that $1 \leq d < |P|$. Assume that $H \cap O^p(G) \trianglelefteq O^p(G)$ for all subgroups $H \trianglelefteq P$ with $|H| = d$. Then either G is p -supersolvable or else $|P \cap O^p(G)| > d$.*

An interesting idea of [6] is that in the hypothesis of the theorem only the normal subgroups of order d are considered, not necessarily the family of all subgroups of order d . Recall that a subgroup H of a group G is said to be S -semipermutable in G (see [12]) if H permutes with all Sylow q -subgroups of G for the primes q not dividing $|H|$. Ballester-Bolinches etc in their paper [1] proved an analogous result, but they only assume that $H \cap O^p(G)$ are S -semipermutable in $O^p(G)$ instead of assuming that these subgroups are normal in $O^p(G)$.

Theorem 1.2. ([1, Theorem 2]). *Let $P \in \text{Syl}_p(G)$, and let d be a power of p such that $1 \leq d < |P|$. Assume that $H \cap O_p(G)$ is S -semipermutable in G for all subgroups $H \trianglelefteq P$ with $|H| = d$. Then either G is p -supersolvable or else $|P \cap O_p(G)| > d$.*

*Corresponding Author.

Email addresses: brzhang@live.com (B. Zhang), xyguo@staff.shu.edu.cn (X. Guo)

Received: 03.07.2017; Accepted: 03.10.2017

More recently, Yu in his paper [13] use the subgroup G_p^* of a group G , and consider the embedding condition $O^p(G_p^*) \cap H \trianglelefteq O^p(G)$ to prove the following result, where G_p^* is the unique smallest normal subgroup of a group G for which the corresponding factor group is abelian of exponent dividing $p - 1$.

Theorem 1.3. ([13, Theorem 2]). *Let $P \in \text{Syl}_p(G)$, and let d be a power of p such that $1 \leq d < |P|$. Then G is p -supersolvable if and only if $|P \cap O^p(G_p^*)| \leq d$ and $H \cap O^p(G_p^*) \trianglelefteq O^p(G)$ for all subgroups $H \trianglelefteq P$ with $|H| = d$.*

Remark 2.3 and Example 2.4 in [13] show that it is better to use the embedding condition $O^p(G_p^*) \cap H \trianglelefteq O^p(G)$ to investigate the p -supersolvability of groups. On the other hand, in all of the above results all normal subgroups of order d in P are considered. So we wonder whether we may reduce the number of normal subgroups of order d in P ?

In fact, we have the following results.

Theorem 1.4. *Let $P \in \text{Syl}_p(G)$, and let d be a power of p such that $p^2 \leq d < |P|$. Then G is p -supersolvable if and only if $|P \cap O^p(G_p^*)| \leq d$ and $H \cap O^p(G_p^*) \trianglelefteq O^p(G)$ for all non-cyclic subgroups $H \trianglelefteq P$ with $|H| = d$.*

Theorem 1.5. *Let p be a prime dividing the order of a group G of odd order, let d be a power of p such that $1 \leq d < |P|$ and $P \in \text{Syl}_p(G)$ with $|P| > d$. And suppose that $H \cap O^p(G_{p^2}^*) \trianglelefteq O^p(G)$ for all non-meta-cyclic normal subgroups H in P with $|H| = d$. Then either p -length $l_p(G) \leq 1$ and p -rank $r_p(G) \leq 2$ or else $|P \cap O^p(G_{p^2}^*)| > d$, where $G_{p^2}^*$ is the unique smallest normal subgroup of the group G for which the corresponding factor group is abelian of exponent dividing $p^2 - 1$.*

We should notice that we assume $d \geq p^2$ in Theorem 1.4. In fact, if $p = 2$ and $d = 2$, then the result is still true. Since $|P \cap O^p(G_p^*)| \leq 2$, it follows from Burnside Theorem [9, IV, 2.8] that $O^p(G_p^*)$ is 2-nilpotent, and thus G_p^* is 2-nilpotent. Hence G is 2-supersolvable. However, the result is not true if p is odd prime and $d = p$ in Theorem 1.4. In fact, let D be a non-abelian simple group such that Sylow p -subgroups of D are cyclic of order p , and let $G = D \times C$ with a cyclic group C of order p . Clearly, $G_p^* = G$ and $H \cap O^p(G_p^*)$ is normal in $O^p(G)$ for every non-cyclic normal subgroup H of P of order d , where P is a Sylow p -subgroup. But $|P \cap O^p(G_p^*)| = p$ and G is not p -supersolvable.

We also notice that the hypothesis “ G is an odd order group” in Theorem 1.5 can not be removed. In fact, let D be a non-abelian simple group such that Sylow p -subgroups of D are cyclic of order p^m ($d \geq p^m \geq 1$), and let $G = D \times C$ with a cyclic group C of order p^n ($n \geq 1$). Clearly, $H \cap O^p(G_{p^2}^*)$ is normal in $O^p(G)$ for every non-meta-cyclic normal subgroup H of P of order d , where P is a Sylow p -subgroup of G . But $|P \cap O^p(G_{p^2}^*)| = p^m \leq d$ and G is not p -solvable.

Furthermore, the following example tells us that G may not be p -supersolvable if G satisfies the hypotheses in Theorem 1.5. Let p be an odd prime with $p \neq 2^k - 1$ for all positive integer k . Write $P = \langle a \rangle \times \langle b \rangle \simeq C_p^{2n}$. So $\text{Aut}(\Omega_1(P)) \simeq GL(2, p)$ and there exists a cyclic subgroup T of order $p^2 - 1$ in $\text{Aut}(\Omega_1(P))$. Note that $p+1$ is not a power of 2, then we can choose an automorphism $\bar{\alpha} \in T$ with order q such that $q|p+1$ and $q \neq 2$. Now, considering the surjective homomorphism $\phi : \text{Aut}(P) \rightarrow \text{Aut}(\Omega_1(P))$; we can choose an automorphism α of P such that $\phi(\alpha) = \bar{\alpha}$. Write the semidirect product $G = P \rtimes \langle \alpha \rangle$, it is clear that $H \cap O^p(G_{p^2}^*)$ is normal in $O^p(G)$ for every non-meta-cyclic normal subgroup H of P of order $d = p^m$ ($m < 2n$). Now we prove that G is not p -supersolvable. If the action of $\bar{\alpha}$ on $\Omega_1(P) = \langle a_1 \rangle \times \langle b_1 \rangle \simeq C_p^2$ is reducible, then it follows from $p \neq q$ and Maschke’s Theorem that $\langle a_1 \rangle, \langle b_1 \rangle$ are $\bar{\alpha}$ -invariant. It follows from $\text{g.c.d.}(p+1, p-1) = 2$ that $q \nmid p-1$, and therefore $\bar{\alpha}$ acts trivially on both $\langle a \rangle$ and $\langle b \rangle$, that is, $\bar{\alpha}$ acts trivially on $\Omega_1(P)$, a contradiction. Hence $\bar{\alpha}$ acts irreducibly on $\Omega_1(P)$, implying that α acts irreducibly on

$\Omega_1(P)$. Then we have $\Omega_1(P) \simeq C_p^2$ is a minimal normal subgroup of G and so $r_p(G) = 2$. It follows that G is not p -supersolvable.

2. Preliminary results

In this section we list some basic and known results which will be used in our proofs.

Definition 2.1. ([7, Definition 1.9]). Let p be prime. A group G is said to be a $CS(p, n)$ -group if G is a p -group with a characteristic series

$$1 = G_0 < G_1 < \cdots < G_m = G$$

such that $|G_i/G_{i-1}| \leq p^n$ for all $i \geq 1$.

It is clear that meta-cyclic p -groups and p -groups of maximal class are both $CS(p, 2)$ -groups.

Lemma 2.2. ([3, Lemma 1.4]). *Let p be a prime, let P be a p -group and let d be a power of p such that $p^2 \leq d < |P|$. Let $N \trianglelefteq P$, where $|N| \geq d$, and suppose that every normal subgroup of P that has order d and is contained in N is cyclic. Then N is cyclic, dihedral, semidihedral or generalized quaternion.*

Lemma 2.3. ([2, Lemma 2.4]). *Let $P \trianglelefteq G$, where P is a p -group. Also, let $N \leq G$ be a p -subgroup with $|N| \leq |P|$ and $N \not\leq P$. Then $N < PN$, and every subgroup H with $N < H \leq NP$ is non-cyclic.*

Lemma 2.4. ([8, Lemma 2.5]). *If a group P of order $2^n > 2^3$ has a subgroup M of order 2^{n-1} of maximal class, then either P is of maximal class or $P/P' \simeq C_2^3$, and P' is cyclic.*

Lemma 2.5. ([3, Exercise 1(b), p.114]). *Let P be dihedral, semidihedral or generalized quaternion, then P has the only one normal subgroup N of order 2^a for every $1 < 2^a < |P|/2$ and N is cyclic.*

Lemma 2.6. ([4, Corollary 69.5]). *Let p be an odd prime and d be a power of p such that $d \geq p^3$, and let N be a normal subgroup of a p -group P with $|N| \geq d$. If every normal subgroup of P that has order d and is contained in N is meta-cyclic, then N is a meta-cyclic group or a 3-group of maximal group.*

Lemma 2.7. *Let p be a odd prime, and let P be a meta-cyclic p -group or a 3-group of maximal class. If N is normal in P , then $\Omega_1(N) \lesssim C_p \times C_p$ or N is a 3-group of maximal class.*

Proof. If P is meta-cyclic, then $\Omega_1(N) \lesssim C_p \times C_p$. Now assume that P is a 3-group of maximal class. It follows from [3, Exercise 9.1] that N is a 3-group of maximal class or absolutely regular, where a p -group N is absolutely regular if $|G/\mathcal{U}_1(G)| < p^p$ (see [3, List of definitions and notations]). If N is absolutely regular, then $|\Omega_1(N)| = |N/\mathcal{U}_1(N)| \leq 3^2$, and thus $\Omega_1(N) \lesssim C_p \times C_p$ by [3, Lemma 1.4]. \square

Lemma 2.8. *Let p be a prime and d be a power of p such that $p^3 \leq d$, and let P be a p -group. Also, let N and P_1 be normal subgroups of P with $N \lesssim C_p \times C_p$ and $N < P_1$. If P_1 contains a non-meta-cyclic normal subgroup of order d of P , then there exists a non-meta-cyclic normal subgroup H of order d of P such that $N < H \leq P_1$.*

Proof. Let H_1 be a non-meta-cyclic normal subgroup of order d of P with $H_1 \leq P_1$. If $N \not\leq H_1$, then $|N \cap H_1| = 1$ or p by $N \lesssim C_p \times C_p$, that is, $|N : N \cap H_1| = p^2$ or p . First, we assume that $|N : N \cap H_1| = p$. Since $N \cap H_1$ is normal in P , there exists a maximal subgroup M of H_1 such that $M \trianglelefteq P$ and $N \cap H_1 \leq M$, and so $H = NM$ is normal in P and $|H| = d$. Noting that H_1 is non-meta-cyclic, we have that M is non-cyclic. It follows from $N \not\leq M$ and $N \lesssim C_p^2$ that $\Omega_1(H) > \Omega_1(M) \geq p^2$. Thus H is non-meta-cyclic by Lemma 2.7 and $H \leq P_1$. Now assume that $|N : N \cap H_1| = p^2$ and take a subgroup M_1 of H_1

with $|M_1| = d/p^2$ and $M_1 \trianglelefteq P$. Then $H = NM_1$ is a normal subgroup of P with $|H| = d$. Noticing that $N \simeq C_p \times C_p$ and $N \cap M = 1$, we see $|\Omega_1(H)| \geq |\Omega_1(N)||\Omega_1(M)| \geq p^3$. Hence H is non-meta-cyclic by Lemma 2.7, as we wanted. \square

Lemma 2.9. ([7, Lemma 2.2]). *Let P be a p -group. If P has a meta-cyclic maximal subgroup and P is not isomorphic to C_p^3 , then P is a $CS(p, 2)$ -group.*

Lemma 2.10. ([7, Lemma 3.2]). *Let an odd order group A act on a $CS(p, 2)$ -group P . Then P is centralized by $OP(A_{p^2}^*)$.*

Lemma 2.11. *Let G be a group and let p be a prime of $|G|$. If $G_{p^2}^*$ is p -nilpotent, then G is p -solvable with $l_p(G) \leq 1$ and $r_p(G) \leq 2$.*

Proof. Since $G_{p^2}^*$ is p -nilpotent, G is p -solvable of $l_p(G) \leq 1$. We see G has a chief series

$$1 = K_0 < \cdots < H_0 = OP(G_{p^2}^*) < H_1 < \cdots < H_n = G_{p^2}^* < \cdots < G$$

Noticing that $OP(G_{p^2}^*) \leq C_G(H_{i+1}/H_i)$ ($0 \leq i \leq n-1$), we have $A_G(H_{i+1}/H_i) \simeq G/C_G(H_{i+1}/H_i) \in \mathfrak{D}_p \mathfrak{U}_{p^2-1}$, where \mathfrak{D}_p is the formation consisting of all p -groups and \mathfrak{U}_{p^2-1} is the formation consisting of all abelian groups with exponent dividing p^2-1 . Since $O_p(A_G(H_{i+1}/H_i)) = 1$ by [5, A, Lemma 13.6], it follows that $A_G(H_{i+1}/H_i) \in \mathfrak{U}_{p^2-1}$, and so $A_G(H_{i+1}/H_i)$ is abelian with exponent dividing p^2-1 .

Write $|H_{i+1}/H_i| = p^m$. By the faithful and irreducible action of the abelian group $A_G(H_{i+1}/H_i)$ on H_{i+1}/H_i , we see that $A_G(H_{i+1}/H_i)$ is cyclic and m is the smallest positive integer such that $|A_G(H_{i+1}/H_i)|$ divides p^m-1 by [9, II, Lemma 3.10], and thus $m \leq 2$ since the exponent of $A_G(H_{i+1}/H_i)$ divides p^2-1 . Then $r_p(G) \leq 2$. \square

3. Proof of Theorem 1.4

Lemma 3.1. *Let p be a prime, and let $P \in Syl_p(G)$, where G is a group. If P is a cyclic group, then either G is p -supersolvable or else $P \cap OP(G_p^*) = P$.*

Proof. Without loss of generality, we assume $P \cap OP(G_p^*) < P$. If $P \cap OP(G_p^*) = 1$, then G_p^* is a p -nilpotent, and thus G is p -supersolvable. So $1 < P \cap OP(G_p^*) < P$, then it follows from [11, Theorem 2.1] that G is p -supersolvable. \square

Proof of Theorem 1.4. Note that G is p -supersolvable if and only if G_p^* is p -nilpotent, and so we only need to prove the sufficient. Now assume that G is a counterexample of minimal order. Then G is not p -supersolvable. In particular, G_p^* is not p -nilpotent, and therefore $N = P \cap OP(G_p^*) > 1$. For convenience, we write

$$\mathfrak{H} = \{H \trianglelefteq P \mid H \text{ is a non-cyclic subgroup with } |H| = d\}$$

and

$$\mathfrak{Y} = \{Y < \cdot P \mid N \not\leq Y\}.$$

It is easy to see that $H \cap OP(G_p^*) \trianglelefteq G$ for all $H \in \mathfrak{H}$. We proceed in a number of steps to derive a contradiction.

Step 1. P is not cyclic, dihedral, semidihedral or generalized quaternion.

If P is cyclic, then, by Lemma 3.1, G is p -supersolvable, a contradiction. Now assume that P is dihedral, semidihedral or generalized quaternion. If N is a cyclic subgroup, then it follows from Burnside's theorem [9, IV, 2.8] and $p = 2$ that $OP(G_p^*)$ is 2-nilpotent, and thus G_p^* is 2-nilpotent, a contradiction. Thus, by Lemma 2.5, we may assume that N is a non-cyclic maximal subgroup of P and $|N| = d$. In this case $P = D_{2^n}$ ($n \geq 3$), Q_{2^n} ($n \geq 4$) or SD_{2^n} , and thus there exists a non-cyclic maximal subgroup N_1 of P such that $N \not\leq N_1$. For convenience, we write $M_1 = N \cap N_1$ and have

$$M_1 = N \cap OP(G_p^*) \cap N_1 \cap OP(G_p^*) \trianglelefteq G.$$

Since $|P : M_1| = 2^2$, M_1 is cyclic by Lemma 2.5. It follows that $O^p(G_p^*)$ is 2-supersolvable and therefore $O^p(G_p^*)$ is 2-nilpotent. Hence G_p^* is 2-nilpotent, a contradiction.

Step 2. $\mathfrak{H} \neq \phi$.

Suppose not, that is, all normal subgroups of P with order d are cyclic. Now by Lemma 2.2, P is cyclic, dihedral, semidihedral or generalized quaternion, in contradiction to Step 1.

Step 3. $O_{p'}(G_p^*) = 1$.

Write $D = O_{p'}(G_p^*)$ and $\bar{G} = G/D$, and note that $O^p(\bar{G}_p^*) = \overline{O^p(G_p^*)}$ by [13, Lemma 2.9]. It follows from Dedekind's lemma that $O^p(G_p^*) \cap DH = D(O^p(G_p^*) \cap H)$ for $H \in \mathfrak{H}$. In addition, both D and $O^p(G_p^*) \cap H$ are normalized by $O^p(G)$, we see that $O^p(G)$ normalizes $O^p(G_p^*) \cap DH$, or equivalently, $\overline{O^p(G)}$ normalizes $\overline{O^p(G_p^*) \cap H}$. Since $PD \cap O^p(G_p^*) = D(P \cap U) = DN$ and $|N| \leq d$, we see that $|\bar{P} \cap \overline{O^p(G_p^*)}| \leq d$. Then \bar{G} satisfies the hypotheses, and therefore \bar{G} is p -supersolvable. It is clear that the subgroups of \bar{G} corresponding to the members of \mathfrak{H} are exactly the subgroups \bar{H} for $H \in \mathfrak{H}$. Hence \bar{G} is p -supersolvable. Furthermore, we see that G is p -supersolvable, which is a contradiction. So we conclude that $D = 1$.

Step 4. N is normal in G . In fact, G is p -solvable and $P \trianglelefteq G$.

Since $H \cap O^p(G_p^*) \trianglelefteq O^p(G)$ for $H \in \mathfrak{H}$ and $O^p(G_p^*) \leq O^p(G)$, we see that $H \cap O^p(G_p^*) \trianglelefteq O^p(G_p^*)$ for $H \in \mathfrak{H}$. Then G_p^* satisfies the hypotheses of [8, Theorem 3.2], and thus G_p^* is p -supersolvable. Hence G is p -solvable. Noticing that $O_{p'}(G_p^*) = 1$ and G_p^* is p -supersolvable, we have $P \trianglelefteq G_p^*$ by [9, VI, 6.6]. Then it follows from $P \in \text{Syl}_p(G_p^*)$ and $G_p^* \trianglelefteq G$ that P is normal in G . So $N = P \cap O^p(G_p^*)$ is normal in G by $O^p(G_p^*) \trianglelefteq G$.

Step 5. There exists a maximal subgroup $Y \in \mathfrak{Y}$ with $L = N \cap Y$ is not normal in G and L is cyclic.

If $N \leq \Phi(P)$, then it follows from Tate's theorem [9, IV, 4.7] that $O^p(G_p^*)$ is p -nilpotent, and therefore G_p^* is p -nilpotent, a contradiction. Thus there exists a maximal subgroup Y of P with $N \not\leq Y$.

Next we prove that there exists $Y \in \mathfrak{Y}$ such that $L = N \cap Y$ is not normal in G . If not, then $L = N \cap Y$ is normal in G and $|N : L| = p$ for all $Y \in \mathfrak{Y}$. So $G_p^* \leq C_G(N/L)$. Noticing that N/L is a normal Sylow p -subgroup of $O^p(G_p^*)/L$, we see $N/L \leq Z(O^p(G_p^*)/L)$, and therefore $O^p(G_p^*)/L$ is p -nilpotent by Burnside's theorem [9, IV, 2.6]. Hence $O^p(O^p(G_p^*)) < O^p(G_p^*)$, a contradiction.

Finally, we prove that L is cyclic. If L is non-cyclic, then there exists $H \in \mathfrak{H}$ such that $L < H \leq Y$. So

$$L = H \cap L = H \cap Y \cap N = H \cap N = H \cap P \cap O^p(G_p^*) = H \cap O^p(G_p^*) \trianglelefteq G,$$

which is a contradiction.

Step 6. Y is a cyclic, dihedral, semidihedral or generalized quaternion group.

Let Y and L be as in Step 5. If there exists a subgroup S in Y such that $S \in \mathfrak{H}$, then, since $|L| < |N| \leq d = |S|$, there exists $H \in \mathfrak{H}$ such that $L < H \leq LS \leq Y$ by Lemma 2.3. In this case, we have

$$L = H \cap L = H \cap Y \cap N = H \cap N = H \cap P \cap O^p(G_p^*) = H \cap O^p(G_p^*) \trianglelefteq G,$$

in contradiction to Step 5. So every normal subgroup of P that has order d and is contained in Y is cyclic. By Lemma 2.2, the Step 6 is true.

Step 7. The final contradictions.

If Y is a cyclic maximal subgroup of P , then it follows from [2, Lemma 2.1(b)] and Step 1 that $O^p(G_p^*)$ acts trivially on P , and therefore G_p^* is p -nilpotent, a contradiction. Now assume that Y is a dihedral, semidihedral or generalized quaternion group. If $|Y| = d$, then $Y \in \mathfrak{H}$, and therefore $L = N \cap Y = P \cap O^p(G_p^*) \cap Y$ is normal in G , in contradiction to Step 5. The remaining case is $|Y| > d$. In this case, since Y is of maximal class, we see

that $|Y : Y'| = 2^2$. Furthermore, by $|Y| > d \geq |N|$ and $|N : L| = 2$, we see that $L \leq Y'$ by Lemma 2.5. It follows from Lemma 2.4 that P' is cyclic, and therefore $L \leq Y' \leq P'$ is normal in G by $P \trianglelefteq G$, in contradiction to Step 5, which is the final contradiction. So the proof is complete. \square

Now we present some application of Theorem 1.4.

Lemma 3.2. *Let $P \in \text{Syl}_p(G)$ with $|P| > p^3$. If P has exactly one non-cyclic maximal subgroup M and $M \trianglelefteq G_p^*$, then G is p -supersolvable.*

Proof. It is easy to see that the hypotheses are inherited by $G/O_{p'}(G_p^*)$ and $P^{G_p^*}$, so we can assume that $O_{p'}(G_p^*) = 1$. If $P^{G_p^*} < G$, then $P^{G_p^*}$ is p -supersolvable by induction. It follows from $O_{p'}(G_p^*) = 1$ and [9, VI, 6.6] that P is normal in $P^{G_p^*}$, and thus $P = P^{G_p^*}$. And since $G_p^* \trianglelefteq G$ and $P \in \text{Syl}_p(G^*)$, we see that $P \trianglelefteq G$. Noticing that there exists a cyclic maximal subgroup in P , we see, by [2, Lemma 2.1], that $O^p(G_p^*)$ acts trivially on P . Thus G_p^* is p -nilpotent, and therefore G is p -supersolvable. Now we can assume that $P^{G_p^*} = G$, and in particular, $G_p^* = G$. Then it follows from [8, Lemma 4.1] that G is p -supersolvable. \square

Lemma 3.3. *Let a Sylow p -subgroup P of G be a non-cyclic subgroup with $|P| > p^3$. If every non-cyclic maximal subgroup of P is normal in G_p^* , then G is p -supersolvable.*

Proof. By Lemma 3.2, we can assume that P has two distinct non-cyclic maximal subgroups. Then P is normal in G_p^* . In addition, G_p^* is normal in G and $P \in \text{Syl}_p(G_p^*)$. Thus P is normal in G . Since $|P| > p^3$, we see, by [2, Theorem A], that $O^p(G_p^*)$ acts trivially on P . Then G_p^* is p -nilpotent, and therefore G is p -supersolvable. \square

Corollary 3.4. *Let P be a non-cyclic Sylow p -subgroup of G with $|P| > p^3$, and suppose for every non-cyclic maximal subgroup H of P that $H \cap O^p(G_p^*) \trianglelefteq O^p(G)$. Then G is p -supersolvable.*

Proof. Assume that G is not p -supersolvable. Applying Theorem 1.4 with $d = |P|/p$, we deduce that $O^p(G_p^*) = G_p^*$, and thus every non-cyclic maximal subgroup of P is normal in G_p^* . It follows from Lemma 3.3 that G is p -supersolvable, a contradiction. \square

Corollary 3.5. *Let p be an odd prime and $P \in \text{Syl}_p(G)$, where P is non-cyclic. Let d be a power of p such that $p^2 \leq d < |P|$, and let \mathfrak{H} be the set of all non-cyclic normal subgroups H of P with $|H| = d$. Assume that $H \cap O^p(G_p^*) \trianglelefteq O^p(G)$ for all $H \in \mathfrak{H}$. If $N_G(H)$ is p -supersolvable for all $H \in \mathfrak{H}$, then G is p -supersolvable.*

Proof. If $|P \cap O^p(G_p^*)| \leq d$, then G is p -supersolvable by Theorem 1.4. Now we can assume that $|P \cap O^p(G_p^*)| > d$. In this case, if there exists $H \in \mathfrak{H}$ such that $H \leq O^p(G_p^*)$, then $H \trianglelefteq O^p(G)$, and thus $H \trianglelefteq PO^p(G) = G$. Hence $G = N_G(H)$ is p -supersolvable. Now we may assume that $N = P \cap O^p(G_p^*)$ is cyclic by Lemma 2.2. Let L be a subgroup of N with order d/p . Since P is non-cyclic, there exists $H \in \mathfrak{H}$ such that $L \leq H$ by Lemma 2.2 and [8, Lemma 2.4], and thus $L = N \cap H$. Noticing that $L = N \cap H = H \cap O^p(G_p^*)$ is normal in $O^p(G)$, we have that L is normal in G . It follows from [11, Theorem 2.1] that $O^p(G_p^*)$ is p -supersolvable, and therefore $N \trianglelefteq O^p(G_p^*)$. In addition, $O^p(G_p^*)$ is normal in G and $N \in \text{Syl}_p(O^p(G_p^*))$. Then N is normal in G . Hence, by [2, lemma 2.1], $O^p(G_p^*)$ acts trivially on N . Furthermore, we see that G_p^* is p -nilpotent and G is p -supersolvable. The proof of the corollary is complete. \square

4. Proof of Theorem 1.5

Lemma 4.1. *Let G be a group of odd order and P be a Sylow p -subgroup of G . If P is a meta-cyclic group or 3-group of maximal class, then $l_p(G) \leq 1$ and $r_p(G) \leq 2$.*

Proof. we proceed by induction on $|G|$. It is easy to see that the hypotheses are inherited by $G/O_{p'}(G)$. So we assume that $O_{p'}(G) = 1$. Then $O_p(G) \neq 1$ since G is p -solvable. By Lemma 2.7, we see that $O_p(G)$ is a 3-group of maximal class or $\Omega_1(O_p(G)) \lesssim C_p \times C_p$. If $O_p(G)$ is of maximal class, then $O_p(G)$ is a $CS(p, 2)$ -group. Hence $O^p(G_{p^2}^*)$ acts trivially on $O_p(G)$ by Lemma 2.10. It follows from Hall-Higman lemma [10, Theorem 3.21] that $O^p(G_{p^2}^*) \leq C_G(O_p(G)) \leq O_p(G)$, and thus $G_{p^2}^*$ is p -group. Furthermore, by Lemma 2.11, $l_p(G) \leq 1$ and $r_p(G) \leq 2$. Now we assume that $\Omega_1(O_p(G)) \lesssim C_p \times C_p$. It follows from Lemma 2.10 that $O^p(G_{p^2}^*)$ act trivially on $\Omega_1(O_p(G))$, and thus $O^p(G_{p^2}^*)$ act trivial on $O_p(G)$ by [9, IV, 5.12]. So $l_p(G) \leq 1$ and $r_p(G) \leq 2$ by using the arguments above. \square

Proof of Theorem 1.5. Suppose that G is a counterexample of minimal order. Then $|P \cap O^p(G_{p^2}^*)| \leq d$ and $l_p(G) \not\leq 1$ or $r_p(G) \not\leq 2$. By Lemma 2.11, we see that $G_{p^2}^*$ is not p -nilpotent, and $N = P \cap O^p(G_{p^2}^*) > 1$. For convenience, we write

$$\mathfrak{H}_1 = \{H \trianglelefteq P \mid H \text{ is a non-meta-cyclic subgroup with } |H| = d\}$$

and

$$\mathfrak{Y} = \{Y < \cdot P \mid N \not\leq Y\}.$$

It is easy to see $H \cap O^p(G_{p^2}^*) \trianglelefteq G$ for all $H \in \mathfrak{H}_1$. We proceed in a number of steps to derive a contradiction.

Step 1. $O_{p'}(G) = 1$.

Write $D = O_{p'}(G)$ and $\overline{G} = G/D$. We argue that \overline{G} satisfies the hypotheses of the theorem. The subgroups of \overline{G} corresponding to the members of \mathfrak{H}_1 are exactly the subgroups \overline{H} for $H \in \mathfrak{H}_1$, and since $O^p(\overline{G}) = \overline{O^p(G)}$ and $O^p(\overline{G_{p^2}^*}) = \overline{O^p(G_{p^2}^*)}$, we must show that $\overline{O^p(G)}$ normalizes $\overline{O^p(G_{p^2}^*)} \cap \overline{H}$. On the other hand, $\overline{O^p(G_{p^2}^*)} \cap \overline{H} = (O^p(G_{p^2}^* \cap H)D)/D$ by [13, Lemma 2.8]. Then $O^p(G)$ normalizes $O^p(G_{p^2}^*) \cap H$. Since D and $O^p(G_{p^2}^*) \cap H$ are normalized by $O^p(G)$, this shows that \overline{G} satisfies the hypotheses, as claimed.

If $D > 1$, then $l_p(\overline{G}) \not\leq 1$ or $r_p(\overline{G}) \not\leq 2$, and thus $|\overline{P} \cap \overline{O^p(G_{p^2}^*)}| > d$ by the minimality of G . Hence $|PD \cap O^p(G_{p^2}^*)| > d|D|$. Since $PD \cap U = D(P \cap O^p(G_{p^2}^*)) = DN$, we see that $|N| > d$, which is a contradiction with $|N| \leq d$. So we conclude that $D = 1$.

Step 2. $d \geq p^3$.

If $d \leq p^2$, then $|N| \leq p^2$. Since G is an odd order group, we see that G is p -solvable. Then it follows from [9, VI, 6.6] that $l_p(O^p(G_{p^2}^*)) \leq 1$. In addition, $O_{p'}(G_{p^2}^*) = 1$ since $O_{p'}(G) = 1$. Thus $N \trianglelefteq O^p(G_{p^2}^*)$, and therefore $N \trianglelefteq G$. It follows that $O^p(G_{p^2}^*)$ acts trivially on N by Lemma 2.10, and so $O^p(G_{p^2}^*)$ is p -nilpotent by Burnside's theorem [9, IV, 2.6]. Hence $G_{p^2}^*$ is p -nilpotent, a contradiction.

Step 3. $\mathfrak{H}_1 \neq \emptyset$.

Suppose not, that is, all subgroups of P with order d are meta-cyclic. Now by Lemma 2.6, P is a meta-cyclic group or a 3-group of maximal class. Then it follows from Step 2 and Lemma 4.1 that $l_p(G) \leq 1$ and $r_p(G) \leq 2$, a contradiction.

Step 4. N is non-meta-cyclic and is normal in G .

Suppose that N is meta-cyclic, that is, N is a cyclic group or a meta-cyclic group with $d(N) = 2$, where $d(N)$ is a minimal number of generators of N . If N is cyclic, and let A be a subgroup of N with order p , then A is normal in P by $N \trianglelefteq P$, and therefore there exists $H \in \mathfrak{H}_1$ such that $A \leq H$ by Lemma 2.8 and $H \cap N \neq 1$. Hence, by $H \cap N = H \cap P \cap O^p(G_{p^2}^*) \trianglelefteq G$ and [11, Theorem 2.1], $O^p(G_{p^2}^*)$ is p -supersolvable. Furthermore, it follows from $O_{p'}(G) = 1$ and [9, VI, 6.6] that N is normal in $O^p(G_{p^2}^*)$. By Lemma 2.9 and 2.10, we see that $O^p(G_{p^2}^*)$ centralizes N , and thus $O^p(G_{p^2}^*)$ is p -nilpotent and $G_{p^2}^*$ is p -nilpotent, a contradiction. Now we assume that N is a metacyclic subgroup of P with $d(N) = 2$. Then $\Omega_1(N) \simeq C_p \times C_p$, and thus, by Lemma 2.8, there exists $H \in \mathfrak{H}_1$ such that $\Omega_1(N) \subseteq H$ and $H \cap N \neq 1$. Hence $T = H \cap N = H \cap P \cap O^p(G_{p^2}^*) \trianglelefteq G$. Noticing

that $\Omega_1(N) = \Omega_1(T)$ char T , we have that $\Omega_1(N)$ is normal in G , and therefore $O^p(G_{p^2}^*)$ centralizes $\Omega_1(N)$ by Lemma 2.10. Since p is odd, we see that $O^p(G_{p^2}^*)$ centralizes N by [9, IV, 5.12]. Then $O^p(G_{p^2}^*)$ is p -nilpotent by Burnside's Theorem [9, IV, 2.6], and thus $G_{p^2}^*$ is p -nilpotent, a contradiction.

Hence N is non-meta-cyclic, and thus there exists $H \in \mathfrak{H}_1$ such that $N \subseteq H$. We see

$$N = N \cap H = O^p(G_{p^2}^*) \cap P \cap H \trianglelefteq G.$$

Step 5. There exists a maximal subgroup $Y \in \mathfrak{Y}$ such that $N \not\leq Y$.

If $N \leq \Phi(P)$, then it follows from Tate's theorem [9, IV, 4.7] that $O^p(G_{p^2}^*)$ is p -nilpotent, and therefore $G_{p^2}^*$ is p -nilpotent, a contradiction. Thus there exists a maximal subgroup Y of P with $N \not\leq Y$.

Step 6. For any $Y \in \mathfrak{Y}$, $L = N \cap Y$ is not normal in G and L is meta-cyclic.

First, we prove that $L = N \cap Y$ is not normal in G for any $Y \in \mathfrak{Y}$. If not, then there exists $Y \in \mathfrak{Y}$ such that $L = N \cap Y \trianglelefteq G$. Since $|N : L| = p$ for all $Y \in \mathfrak{Y}$, $G_{p^2}^* \leq C_G(N/L)$. In addition, N/L is a normal Sylow p -subgroup of $O^p(G_{p^2}^*)/L$, then $N/L \leq Z(O^p(G_{p^2}^*)/L)$, and therefore $O^p(G_{p^2}^*)/L$ is p -nilpotent by Burnside's theorem [9, IV, 2.6]. Hence $O^p(O^p(G_{p^2}^*)) < O^p(G_{p^2}^*)$, a contradiction.

Next, we prove that L is meta-cyclic. If L is non-meta-cyclic, then there exists $H \in \mathfrak{H}_1$ such that $L < H \leq Y$. So

$$L = H \cap L = H \cap Y \cap N = H \cap N = H \cap P \cap O^p(G_{p^2}^*) = H \cap O^p(G_{p^2}^*) \trianglelefteq G,$$

which is a contradiction.

Step 7. $N \simeq C_p \times C_p \times C_p$.

If not, then, since L is a meta-cyclic maximal subgroup of N , we see that N is a $CS(p, 2)$ -group by Lemma 2.9, and thus N is centralized by $O^p(G_{p^2}^*)$ by Lemma 2.10. Hence $G_{p^2}^*$ is p -nilpotent, a contradiction.

Step 8. The final contradiction.

It is easy to see that $G_{p^2}^*/N$ is p -nilpotent. If $N \leq \Phi(G)$, then $G_{p^2}^*$ is p -nilpotent, a contradiction. Hence there exists a maximal subgroup M of G such that $N \not\leq M$. It is easy to see that N is a minimal normal subgroup of G . If not, there is nothing to be proved. Then $G = NM$ and $N \cap M = 1$. It follows that $P = N(P \cap M)$ by Dedekind's lemma. For convenience, write $S = P \cap M$. Noticing that there exists a maximal subgroup P_1 of P such that $S \leq P_1$ and $N \not\leq P_1$. Write $K = N \cap P_1$ is normal in P and $K \simeq C_p \times C_p$ by Step 7. If there exists $H_1 \in \mathfrak{H}_1$ such that $H_1 \leq P_1$, then, by Lemma 2.8, there exists $H \in \mathfrak{H}_1$ such that $K \leq H \leq P_1$, and thus $K = N \cap P_1 \cap H = H \cap O^p(G_{p^2}^*) \trianglelefteq G$, which contradicts Step 6. Then it follows from Lemma 2.6 and Lemma 4.1 that P_1 is a meta-cyclic group of $d(P_1) = 2$ or a 3-group of maximal class. If P_1 is meta-cyclic of $d(P_1) = 2$, then $\Omega_1(P_1) \simeq C_p \times C_p$, and therefore $\Omega_1(S) \leq \Omega_1(P_1) = K \leq N$. In addition, we know that $\Omega_1(S) \leq S \leq M$ and $N \cap M = 1$. Then $\Omega_1(S) = 1$, and thus $S = 1$. Hence $N = P$, which is a contradiction with $|N| \leq d < |P|$. Now we assume that P_1 is a 3-group of maximal class. Since $p^3 = |N| \leq d < |P|$, we see that $|P_1| \geq p^3$. If $|P_1| \geq 3^4$, then $K \leq \Phi(P_1)$ by [3, Exercise 9.1.]. It follows from Dedekind's lemma that $P_1 = (P_1 \cap N)S$ and $P_1 = S$, which is a contradiction with $P = NS > P$. Now we assume that $|P_1| = 3^3$ and $|P| = 3^4$. Then it follows from $p^3 = |N| \leq d < |P| = p^4$ that $d = p^3$. Hence $P_1 \in \mathfrak{H}_1$. Furthermore, we see that $K = N \cap P_1 = P_1 \cap O^p(G_{p^2}^*) \trianglelefteq G$, which is a contradiction with Step 6. This final contradiction completes the proof. \square

Now we may present some applications of Theorem 1.5.

Lemma 4.2. *Let G be a group of odd order and $P \in \text{Syl}_p(G)$ with $|P| > p^4$. If P has exactly one non-meta-cyclic maximal subgroup M and $M \trianglelefteq G$, then $l_p(G) \leq 1$ and $r_p(G) \leq 2$.*

Proof. We proceed by induction on $|G|$. It is clear that the hypotheses are inherited by $G/O_{p'}(G)$ and P^G , so we can assume that $O_{p'}(G) = 1$. If $P^G < G$, then $l_p(P^G) = 1$ by induction. In addition, $O_{p'}(G) = 1$, then P is normal in P^G . Since $P^G \trianglelefteq G$ and $P \in \text{Syl}_p(P^G)$, we see $P = P^G \trianglelefteq G$. Notice that P has a meta-cyclic maximal subgroup and $|P| > p^4$. Then P is a $CS(p, 2)$ -group by Lemma 2.11, and thus P is centralized by $O^p(G_{p^2}^*)$ by Lemma 2.10. Then it follows from Burnside's theorem [9, IV, 2.6] that $O^p(G_{p^2}^*)$ is p -nilpotent. Hence $G_{p^2}^*$ is p -nilpotent, and therefore $l_p(G) \leq 1$ and $r_p(G) \leq 2$ by Lemma 2.11.

Now we can assume that $P^G = G$, and in particular, $G_{p^2}^* = G$. Applying Theorem 1.5, we may assume that $d = |P|/p$ and $|P \cap O^p(G_{p^2}^*)| > d$, and therefore $O^p(G_{p^2}^*) = G_{p^2}^*$. Since M is the unique non-meta-cyclic maximal subgroup of P , we see that M has a meta-cyclic maximal subgroup by [7, Lemma 2.3]. On the other hand, M is a $CS(p, 2)$ -group since $|M| > p^3$. Then, by Lemma 2.10, $O^p(G_{p^2}^*) = G$ acts trivially on M . Thus P is abelian and $P \simeq C_{p^m} \times C_p \times C_p$ ($m \geq 3$). Let $N = N_G(P)$. We see that N/P acts on the P and centralizes M . It follows from Fitting's lemma [10, Lemma 4.28] and $P \simeq C_{p^m} \times C_p \times C_p$ that the action of N/P on P is trivial, and therefore $P \leq Z(N)$. So G is p -nilpotent by Burnside's theorem [9, IV, 2.6], and thus $l_p(G) \leq 1$ and $r_p(G) \leq 2$ by Lemma 2.11. \square

Lemma 4.3. *Let G be a group of odd order, and let P be a Sylow p -subgroup of G with $|P| > p^4$. If every non-meta-cyclic maximal subgroup of P is normal in G , then $l_p(G) \leq 1$ and $r_p(G) \leq 2$.*

Proof. We proceed by induction on $|G|$. It is easy to see that the hypotheses are inherited by $G/O_{p'}(G)$. so we can assume that $O_{p'}(G) = 1$. It follows from Lemma 2.6 and 4.1 that P has a non-meta-cyclic maximal subgroup. By Lemma 4.2, we can assume that P has two distinct non-meta-cyclic maximal subgroups, and therefore P is normal in G . Since $|P| > p^4$, $O^p(G_{p^2}^*)$ acts trivially on P by [7, Theorem A]. Hence $G_{p^2}^*$ is p -nilpotent, and thus $l_p(G) \leq 1$ and $r_p(G) \leq 2$. \square

Corollary 4.4. *Let G be an odd order group and P be a Sylow p -subgroup of G with $|P| > p^4$, and suppose for every non-cyclic maximal subgroup H of P that $H \cap U \leq U$, where $U = O^p(G)$. Then $l_p(G) \leq 1$ and $r_p(G) \leq 2$.*

Proof. Applying Theorem 1.5 with $d = |P|/p$, we deduce that $O^p(G_{p^2}^*) = G_{p^2}^*$, and thus every non-cyclic maximal subgroup of P is normal in G . It follows from Lemma 4.3 that $l_p(G) \leq 1$ and $r_p(G) \leq 2$, a contradiction. \square

Acknowledgment. The authors would like to thank the referee for his/her valuable suggestions and useful comments which contributed to the final version of this article. The research of the work was partially supported by the National Natural Science Foundation of China(11771271).

References

- [1] A. Ballester-Bolinches, R. Esteban-Romero and S. Qiao, *A note on a result of Guo and Isaacs about p -supersolvability of finite groups*, Arch. Math. (Basel) **106** (6), 501-506, 2016.
- [2] Y. Berkovich and I.M. Isaacs, *p -Supersolvability and actions on p -groups stabilizing certain subgroups*, J. Algebra **414** 82-94, 2014.
- [3] Y. Berkovich and Z. Janko, *Groups of prime power order Vol. 1*, Walter de Gruyter, Berlin, New York, 2008.
- [4] Y. Berkovich and Z. Janko, *Groups of prime power order Vol 2*, Walter de Gruyter, Berlin, New York, 2008.
- [5] K. Doerk and T. Hawkes, *Finite soluble groups*, Walter de Gruyter, Berlin, 1992.

- [6] Y. Guo and I.M. Isaacs, *Conditions on p -subgroups implying p -nilpotence or p -supersolvability*, Arch. Math. (Basel) **105** (3), 215-222, 2015.
- [7] X. Guo and H. Meng, *Actions on p -groups with the kernel containing the p -residuals*, Comm. Algebra, **45** (7), 3022-3033, 2017.
- [8] X. Guo and B. Zhang, *Conditions on p -subgroups implying p -supersolvability*, J. Algebra Appl. **16** (10) 1750196, 9 pages, 2017.
- [9] B. Huppert, *Endliche Gruppen I*, Springer-Verlag, Berlin, 1967.
- [10] I.M. Isaacs, *Finite group theory*, Graduate Studies in Math. 92 AMS Providence, 2008.
- [11] R.M. Peacock, *Groups with a cyclic sylow subgroup*, J. Algebra **56**, 506-509, 1979.
- [12] L. Wang and Y. Wang, *On s -semipermutable maximal and minimal subgroups of Sylow p -subgroups of finite groups*, Comm. Algebra **34**, 143-149, 2006.
- [13] H. Yu, *Some sufficient and necessary conditions for p -supersolvability and p -nilpotence of a finite group*. J. Algebra Appl. **16** (2), 1750052, 9 pages, 2017.



Multi-objective Sustainable Fuzzy Economic Production Quantity (SFEPQ) Model with Demand as Type-2 Fuzzy Number: A Fuzzy Differential Equation Approach

B. K. Debnath*, P. Majumder, U. K. Bera

Department of Mathematics, National Institution of Technology, Agartala, Tripura(west), India

Abstract

A sustainable fuzzy economic production quantity (SFEPQ) inventory model is formulated by introducing the concept of fuzzy differential equation (FDE) due to dynamic behavior of the production-demand system. Generalized Hukuhara (gH) differentiability procedure is applied to solve FDE. Since the demand parameter is taken as trapezoidal type-2 fuzzy number, to get corresponding defuzzified values, first critical value (CV)-based reduction method is applied on demand function to transfer into type-1 fuzzy variable which turns to hexagonal fuzzy number in form. After that α -cut of a hexagonal fuzzy number is used to find the upper and lower value of demand. To apply the α -cut operation on FDE, we divided the interval $[0,1]$ into two sub-intervals $[0,0.5]$ and $[0.5,1]$ and gH-differentiation is applied on this sub-intervals. The objective of this paper is to maximize the profit and simultaneously minimize the carbon emission cost occurring due to the process of inventory management. Finally, the non-linear objective functions are solved by using of multi-objective genetic algorithm and sensitivity analyses on various parameters are also performed in numerically and graphically.

Mathematics Subject Classification (2010). 90B05, 90B30

Keywords. Sustainable economic production quantity model, type-2 fuzzy demand, α -cut of hexagonal fuzzy number, fuzzy differential equation, generalized Hukuhara differentiation

1. Introduction

Nowadays, many countries have implemented various carbon emission taxes as a part of damage to the environment caused by industry on the inventory process. Therefore, it is a challenge for every manufacturing company or organizations to reduce the carbon emission cost on waste management, excess energy use and obsolescence management by producing sustainable products as well as maintain the profit which motivates the researchers to apply carbon emission factors in their models.

*Corresponding Author.

Email addresses: deb.bijoy91@yahoo.com (B. K. Debnath), pinki.mjmdr@rediffmail.com (P. Majumder), bera_uttam@yahoo.co.in (U. K. Bera)

Received: 31.07.2017; Accepted: 10.10.2017

Due to the complex environment during the business management, some critical parameters in the inventory problem are always treated as uncertain variables to meet the practical situations. For instance, if one needs to make a decision on inventory management for the next month, the demand and other relevant costs related to inventory are often required to be estimated by professional judgments or probability statistics because of no precise prior information. But, representing demand parameter for an inventory control problem by fuzzy set is considered difficult since it can be determined from many expert's opinion in different ways, and sometimes it is tough to determine the exact membership function. In these cases, each expert's opinion is a membership function of type-1 fuzzy set and thus, this membership function again becomes fuzzy. The final opinion of all experts is expressed by a type-2 fuzzy set (T2 FS).

Also, when the behavior of a dynamical system is not certain, i.e. when the production and demand are fuzzy, the governing differential equation is called fuzzy differential equation (FDE) of instantaneous state of inventory level and the parameters are characterized by a fuzzy number. Hence, we take the demand parameter as trapezoidal type-2 fuzzy number. In case of a T2 FS, complete defuzzification process consists of two parts-type reduction and defuzzification. Type reduction is a procedure by which a T2 FS is transformed to the corresponding T1 FS, known as type reduced set (TRS). The TRS is then easily defuzzified to a crisp value. Using CV-base reduction method we defuzzified the type-2 fuzzy amount.

The major contribution of this research can be stated as follows:

- (i) A profit maximization and carbon emission cost minimization multi-objective partially backlogging fuzzy economic production quantity model is developed where the demand function is taken as trapezoidal type-2 fuzzy variable.
- (ii) Fuzzy differential equation proposed by Kandel and Byatt [13] is considered because of the dynamic nature of the system.
- (iii) Generalized Hukuhara (gH) derivative approach proposed by Stefanini and Bede [33] is used to solve the fuzzy differential equation.
- (iv) Critical value (CV) -based reduction method is used for trapezoidal type-2 fuzzy variable which become hexagonal fuzzy number in form and α -cut of hexagonal fuzzy number is used to get the corresponding crisp value of demand.
- (v) Multi-objective genetic algorithm is used to get the corresponding lower and upper bound of profit and carbon emission cost of the non-linear objective function.

2. Literature survey

In the literature, it is found that Stock et al. [34] showed that transport and warehouse operations generate large amounts of carbon emission. Hovelaque and Bironneau [11] formulated a carbon constrained integrated economic order quantity (EOQ) model which maximizes a retailer's profit and minimizes carbon emission cost. They investigated the link between inventory policy, total carbon emission and both price and environmental dependent demands. Kazemi et al. [14] formulated an economic order quantity models for items with imperfect quality considering the effect of emission. Battini et al. [3] constructed a new model on sustainable economic order quantity (SEOQ) considering ordering and holding cost of inventory and obsolescence costs and also considered emissions of obsolescence cost for transportation problem. Jonas et al. [12] discussed about the uncertainty present in the greenhouse gas and formulated a fuzzy model in greenhouse gas inventory. Recently, Aljazzar et al. [1] formulated a strategy to reduce carbon emissions from supply chains.

One of the first economic production quantity (EPQ) models with fuzzy parameters was developed by Lee and Yao [17]. In a similar paper, Chang [6] applied the methodology in Lee and Yao [17] and analyzed a condition that the production quantity is a triangular

fuzzy number (TFN). He deduced that fuzzy and crisp approaches lead to the same result in the investigated model. Another identical research was treated by Lin and Yao [18] who assumed that production quantity is a trapezoidal fuzzy number (TPFN). In this direction, Shekarian et al. [27], [28],[30],[31] formulated different fuzzy EOQ/EPQ models considering different holding costs for imperfect quality items with backorders and rework for a single stage system. Soni et al. [32] formulated a fuzzy inventory model with demand uncertainty and learning in a continuous process. Sadeghi et al. [25] proposed a two-tuned metaheuristics approach for a fuzzy random EPQ problem with shortage and redundancy allocation. The readers could read the extensive survey paper by Shekarian et al. [29] on fuzzy inventory models. All the above investigations assumed the fuzzy parameters/ variables to be of type-1 fuzzy set (T1 FS). T2 FSs are extensions of T1 FSs was first introduced by Zadeh [38], [39]. The membership grade of a T2 FS is a fuzzy number with a support bounded by the interval $[0, 1]$. The logical operations of T2 FS were explored by Mizumoto and Tanaka [22] and Dubois and Prade [8]. Many authors e.g., [19], [26], [35] contributed a large number of theoretical research works on the property of T2 FS and the applications of T2 FS on operations research e.g., [15], [16], [21]. There are several method for type reduction. Qin et al. [23] introduced three kinds of reduction methods called optimistic CV, pessimistic CV and CV reduction methods for T2 FVs based on critical values (CVs) of regular FVs. α -cut and the extension principle forms a methodology for extending mathematical concepts from crisp sets to fuzzy sets. These have been applied to many operations and have also been extended to interval valued fuzzy sets. Dubois and Prade [8] has defined fuzzy number as a fuzzy subset of the real line. So far, fuzzy numbers like TFN, TPFN, Hexagonal fuzzy number [24] have been introduced with its membership functions. These numbers have got many applications in practical field and many operations were performed using fuzzy numbers.

The presence of fuzzy demand as well as fuzzy production rate leads to FDE of instantaneous state of inventory level. Till now, FDE is less used to solve fuzzy inventory models though the topics on FDE have been rapidly growing in the recent years. The first impetus on solving FDE was made by Kandel and Byatt [13]. Furthermore, different approaches have been made by several authors to solve FDE [2], [9]. In the FDE, all derivatives are deliberated as either Hukuhara or generalized derivatives. The Hukuhara differentiability [5] has a deficiency that the solution turns fuzzifier as time goes on. Bede [4] exhibited that a large class of Boundary Value Problems (BVPs) has no solution if the Hukuhara derivative is applied. To remove this difficulty, the concept of a generalized derivative was developed and fuzzy differential equations were smeared using this concept. Stefanini and Bede [33] introduced the concept of generalization of the Hukuhara difference for compact convex set, introduced generalized Hukuhara differentiability for fuzzy valued function and they displayed that, this concept of differentiability have relationships with weakly generalized differentiability and strongly generalized differentiability. Villamizar-Roa et al. [36] studied the existence and uniqueness of solution for fuzzy differential equation problems in the setting of a generalized Hukuhara derivative. Guchhait et al. [10] formulated a fuzzy production inventory model using fuzzy differential equation and the corresponding inventory costs and components are calculated using fuzzy Riemann integration. Trade credit financing is one of the central features in supply chain management. In real life situations retailer offers trade credit to his/her customers to boost the demand. This real phenomenon is depicted in our present model. Also, Majumder et al. [20] formulated a fuzzy production inventory model with partial trade credit and solve in fuzzy environment via Generalized Hukuhara derivative approach.

Some papers of the above literature survey and our proposed model are summarized and presented in Table 1.

The rest of the paper is organized as follows: In Section 3, we define all the preliminary concepts relating to fuzzy sets. Section 4, discusses various notations and assumptions.

Authors	Crisp	Fuzzy	EOQ	EPQ	Carbon emission	FDE	T2FS	gH-differentiability
Hovelaque and Bironneau [11]	✓		✓		✓			
Battini et al. [3]	✓		✓		✓			
Jonas et al. [12]		✓			✓			
Villamizar-Roa et al. [36]		✓				✓		✓
Guchhait et al. [10]		✓		✓		✓		
Majumder et al. [20]		✓		✓		✓		✓
Shekarian et al. [30]		✓	✓	✓				
Kazemi et al. [14]	✓		✓		✓			
Proposed Model		✓		✓	✓	✓	✓	✓

Table 1. Contribution of Different Authors

section 5 is about mathematical formulation of the model. In section 6, we discuss about the solution procedure for solving multi-objective non-linear problems. In Section 7 real life numerical data and solutions are represented. Discussion on the solution are presented in section 8. Finally brief conclusions and future research work are drawn in section 9.

3. Preliminaries

3.1. Type-1 Fuzzy set (T1FS) [38]:

A fuzzy set \tilde{A} is defined by $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) : x \in A, \mu_{\tilde{A}}(x) \in [0, 1]\}$. In the pair $(x, \mu_{\tilde{A}}(x))$ the first element x belong to the classical set A , the second element $\mu_{\tilde{A}}(x)$, belong to the interval $[0, 1]$, called Membership function.

3.2. Type-2 Fuzzy Set (T2FS) [38]:

Type-2 fuzzy set \tilde{A} defined on a universe of discourse X , which is denoted as $\tilde{A} \subseteq X$, is a set of pairs $\{(x, \mu_{\tilde{A}}(x))\}$, where x an element of a fuzzy set is, and its grade of membership $\{\mu_{\tilde{A}}(x)\}$ in the fuzzy set \tilde{A} is a type-1 fuzzy set defined in the interval $J_x \subset [0, 1]$, i.e. A T2 FS \tilde{A} is defined as

$$\tilde{A} = \{(x, u), \mu_{\tilde{A}}(x, u) : \forall x \in X, J_x \subset [0, 1]\}, \quad (3.1)$$

where $0 \leq \mu_{\tilde{A}}(x, u) \leq 1$ is the type-2 membership function.

3.3. Regular fuzzy variable (RFV) [8]:

For a possibility space (φ, p, Pos) , a regular fuzzy variable $\tilde{\xi}$ is defined as a measurable map from φ to $[0, 1]$ in the sense that for every $t \in [0, 1]$, one has $\{\gamma \in \varphi \mid \tilde{\xi}(\gamma) \leq t\} \in p$. A discrete RFV is represented as $\tilde{\xi} \sim \begin{pmatrix} r_1, r_2 \cdots r_n \\ \mu_1, \mu_2 \cdots \mu_n \end{pmatrix}$ where $r_i \in [0, 1]$ and $\mu_i > 0, \forall i$ and $\max_i(\mu_i) = 1$.

If $\tilde{\xi} = (r_1, r_2, r_3, r_4)$ with $0 \leq r_1 < r_2 < r_3 < r_4 \leq 1$, then $\tilde{\xi}$ is called a trapezoidal RFV.

Example 3.1. Let us take $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) : x \in X\}$ where $X = 3, 6, 9$ and the primary memberships of the points 3, 6, 9 are given by $J_3 = 0.4, 0.8, 0.9$, $J_6 = 0.3, 0.7, 0.8, 0.9$ and $J_9 = 0.2, 0.7, 1.0$ respectively. Then the secondary grade of the point 3 is

$$\mu_{\tilde{A}}(3) = \mu_{\tilde{A}}(3, u) = (0.5/0.4) + (0.7/0.8) + (0.3/0.9) \sim \begin{pmatrix} 0.4 & 0.8 & 0.9 \\ 0.5 & 0.7 & 0.3 \end{pmatrix}$$

That means, $\mu_{\tilde{A}}(3, 0.4) = 0.5$, $\mu_{\tilde{A}}(3, 0.8) = 0.7$, $\mu_{\tilde{A}}(3, 0.9) = 0.3$.

More specifically $\mu_{\tilde{A}}(3, 0.4) = 0.5$ means that the membership grade which is named as secondary membership grade that the point 3 has the primary membership 0.4 is 0.5.

So \tilde{A} considers on the value 3 with membership grade $\begin{pmatrix} 0.4 & 0.8 & 0.9 \\ 0.5 & 0.7 & 0.3 \end{pmatrix}$, which is a RFV.

3.4. Critical values (CVs) for RFVs [23]

Qin et al. [23] introduced three kinds of critical values (CVs). Let $\tilde{\xi}$ be a RFV. Then,

I. The optimistic CV of $\tilde{\xi}$, denoted by $CV^*[\tilde{\xi}]$ is given by,

$$CV^*[\tilde{\xi}] = \sup_{\alpha \in [0,1]} [\alpha \wedge Pos(\tilde{\xi} \geq \alpha)] \quad (3.2)$$

II. The pessimistic CV of $\tilde{\xi}$, denoted by $CV_*[\tilde{\xi}]$ is given by,

$$CV_*[\tilde{\xi}] = \sup_{\alpha \in [0,1]} [\alpha \wedge Nec(\tilde{\xi} \geq \alpha)] \quad (3.3)$$

III. The CV of $\tilde{\xi}$, denoted by $CV[\tilde{\xi}]$ is given by,

$$CV[\tilde{\xi}] = \sup_{\alpha \in [0,1]} [\alpha \wedge Cr(\tilde{\xi} \geq \alpha)] \quad (3.4)$$

Example 3.2. Let $\tilde{\xi}$ be a discrete RFV define as $\xi \sim \begin{pmatrix} 0.3 & 0.5 & 0.8 & 0.9 \\ 0.1 & 0.9 & 0.6 & 0.3 \end{pmatrix}$

Then we can find out that,

$$Pos(\tilde{\xi} \geq \alpha) = \begin{cases} 0, & \text{if } \alpha \leq 0.2 \\ 0.9, & \text{if } 0.2 < \alpha \leq 0.5 \\ 0.6, & \text{if } 0.5 < \alpha \leq 0.8 \\ 0.3, & \text{if } 0.8 < \alpha \leq 1.0 \end{cases}$$

$$Nec(\tilde{\xi} \geq \alpha) = \begin{cases} 0.9, & \text{if } \alpha \leq 0.3 \\ 0.1, & \text{if } 0.3 < \alpha \leq 0.5 \\ 0.4, & \text{if } 0.5 < \alpha \leq 0.8 \\ 0.7, & \text{if } 0.8 < \alpha \leq 1.0 \end{cases}$$

and

$$Cr(\tilde{\xi} \geq \alpha) = \begin{cases} 0.9, & \text{if } \alpha \leq 0.3 \\ 0.5, & \text{if } 0.3 < \alpha \leq 0.5 \\ 0.5, & \text{if } 0.5 < \alpha \leq 0.8 \\ 0.5, & \text{if } 0.8 < \alpha \leq 1.0 \end{cases}$$

Then by the definitions of CVs, from (3.2)-(3.4), we have

$$\begin{aligned} CV^*[\tilde{\xi}] &= \sup_{\alpha \in [0,1]} [\alpha \wedge Pos(\tilde{\xi} \geq \alpha)] \\ &= \sup_{\alpha \in [0,0.2]} [\alpha \wedge 0] \vee \sup_{\alpha \in (0.2,0.5]} [\alpha \wedge 0.9] \vee \sup_{\alpha \in (0.5,0.8]} [\alpha \wedge 0.6] \vee \sup_{\alpha \in [0.8,1]} [\alpha \wedge 0.3] \\ &= 0 \vee 0.5 \vee 0.6 \vee 0.3 \\ &= 0.6 \end{aligned}$$

$$\begin{aligned} CV_*[\tilde{\xi}] &= \sup_{\alpha \in [0,1]} [\alpha \wedge Nec(\tilde{\xi} \geq \alpha)] \\ &= \sup_{\alpha \in [0,0.3]} [\alpha \wedge 0.9] \vee \sup_{\alpha \in (0.3,0.5]} [\alpha \wedge 0.1] \vee \sup_{\alpha \in (0.5,0.8]} [\alpha \wedge 0.4] \vee \sup_{\alpha \in (0.8,1]} [\alpha \wedge 0.7] \\ &= 0 \vee 0.1 \vee 0.4 \vee 0.7 \\ &= 0.7 \end{aligned}$$

and

$$\begin{aligned} CV[\tilde{\xi}] &= \sup_{\alpha \in [0,1]} [\alpha \wedge Cr(\tilde{\xi} \geq \alpha)] \\ &= \sup_{\alpha \in [0,0.3]} [\alpha \wedge 0.9] \vee \sup_{\alpha \in (0.3,0.5]} [\alpha \wedge 0.5] \vee \sup_{\alpha \in (0.5,0.8]} [\alpha \wedge 0.5] \vee \sup_{\alpha \in (0.8,1]} [\alpha \wedge 0.5] \\ &= 0 \vee 0.5 \vee 0.5 \vee 0.5 \\ &= 0.5 \end{aligned}$$

3.5. Critical values (CVs) of trapezoidal RFVs [23]

The following theorems introduced the critical values (CVs) of trapezoidal RFVs.

Theorem 3.3. (Qin et al. [23]) *Let $\tilde{\xi} = (r_1, r_2, r_3, r_4)$ be a trapezoidal RFV. Then we have,*

- (1) *The optimistic CV of $\tilde{\xi}$ is $CV^*[\tilde{\xi}] = \frac{r_4}{(1+r_4-r_3)}$.*
- (2) *The pessimistic CV of $\tilde{\xi}$ is $CV_*[\tilde{\xi}] = \frac{r_2}{(1+r_2-r_1)}$.*
- (3) *The CV of $\tilde{\xi}$ is $CV[\tilde{\xi}] = \begin{cases} \frac{2r_2-r_1}{1+2(r_2-r_1)}, & \text{if } r_2 \geq \frac{1}{2} \\ \frac{1}{2}, & \text{if } r_2 \leq \frac{1}{2} < r_3 \\ \frac{r_4}{1+2(r_4-r_3)}, & \text{if } r_3 \leq \frac{1}{2} \end{cases}$*

Example 3.4. Let $\tilde{\xi} = (0.3, 0.4, 0.8, 0.9)$ be a trapezoidal RFV. Then according to the theorem 3.3 we have,

$$CV^*[\tilde{\xi}] = \frac{9}{11}, \quad CV_*[\tilde{\xi}] = \frac{4}{11}, \quad CV[\tilde{\xi}] = \frac{1}{2}.$$

3.6. Proposed CV based defuzzification for trapezoidal type-2 fuzzy variable

According to Chen et al. [7] for a trapezoidal type-2 fuzzy variable $\tilde{\xi} = (r_1, r_2, r_3, r_4; \theta_l, \theta_r)$, where $r_i \in R, \forall i$ and $\theta_l, \theta_r \in [0, 1]$ are the two parameters that characterize the degree of uncertainty that $\tilde{\xi}$ takes a value say x and the corresponding secondary possibility distribution function $\tilde{\mu}_{\tilde{\xi}}(x)$ is given by,

For any $x \in [r_1, r_2]$,

$$\tilde{\mu}_{\tilde{\xi}}(x) = \left(\frac{x-r_1}{r_2-r_1} - \theta_l \min \left\{ \frac{x-r_1}{r_2-r_1}, \frac{r_2-x}{r_2-r_1} \right\}, \frac{x-r_1}{r_2-r_1}, \frac{x-r_1}{r_2-r_1} + \theta_r \min \left\{ \frac{x-r_1}{r_2-r_1}, \frac{r_2-x}{r_2-r_1} \right\} \right)$$

for $x \in (r_2, r_3)$, $\tilde{\mu}_{\tilde{\xi}}(x) = \tilde{1}$ and

$$\tilde{\mu}_{\tilde{\xi}}(x) = \left(\frac{r_4-x}{r_4-r_3} - \theta_l \min \left\{ \frac{r_4-x}{r_4-r_3}, \frac{x-r_3}{r_4-r_3} \right\}, \frac{r_4-x}{r_4-r_3}, \frac{r_4-x}{r_4-r_3} + \theta_r \min \left\{ \frac{r_4-x}{r_4-r_3}, \frac{x-r_3}{r_4-r_3} \right\} \right)$$

For any $x \in [r_3, r_4]$.

Theorem 3.5. *Let $\tilde{\xi} = (r_1, r_2, r_3, r_4; \theta_l, \theta_r)$ be a type -2 trapezoidal fuzzy variable. Then we have,*

- I. *Using the optimistic CV reduction method, the reduction ξ_1 of $\tilde{\xi}$ has the following possibility distribution,*

$$\mu_{\xi_1}(x) = \begin{cases} \frac{(1+\theta_r)(x-r_1)}{r_2-r_1+\theta_r(x-r_1)}, & \text{if } x \in \left[r_1, \frac{r_1+r_2}{2} \right] \\ \frac{(1-\theta_r)x-r_1+\theta_r r_2}{r_2-r_1+\theta_r(r_2-x)}, & \text{if } x \in \left(\frac{r_1+r_2}{2}, r_2 \right] \\ \tilde{1}, & \text{if } x \in (r_2, r_3) \\ \frac{(1+\theta_r)(r_4-x)}{r_4-r_3+\theta_r(r_4-x)}, & \text{if } x \in \left(r_3, \frac{r_3+r_4}{2} \right] \\ \frac{(-1+\theta_r)x-\theta_r r_3+r_4}{r_4-r_3+\theta_r(x-r_4)}, & \text{if } x \in \left(\frac{r_3+r_4}{2}, r_4 \right]. \end{cases} \quad (3.5)$$

II. Using the pessimistic CV reduction method, the reduction ξ_2 of $\tilde{\xi}$ has the following possibility distribution,

$$\mu_{\xi_2}(x) = \begin{cases} \frac{(x-r_1)}{r_2-r_1+\theta_l(x-r_1)}, & \text{if } x \in \left[r_1, \frac{r_1+r_2}{2} \right] \\ \frac{(x-r_1)}{r_2-r_1+\theta_l(r_1-x)}, & \text{if } x \in \left(\frac{r_1+r_2}{2}, r_2 \right) \\ \tilde{1}, & \text{if } x \in (r_2, r_3) \\ \frac{(r_4-x)}{r_4-r_3+\theta_l(x-r_3)}, & \text{if } x \in \left(r_3, \frac{r_3+r_4}{2} \right) \\ \frac{(r_4-x)}{r_4-r_3+\theta_l(r_4-x)}, & \text{if } x \in \left(\frac{r_3+r_4}{2}, r_4 \right). \end{cases} \quad (3.6)$$

III. Using the CV reduction method, the reduction ξ_3 of $\tilde{\xi}$ has the following possibility distribution,

$$\mu_{\xi_3}(x) = \begin{cases} \frac{(1+\theta_r)(x-r_1)}{r_2-r_1+2\theta_r(x-r_1)}, & \text{if } x \in \left[r_1, \frac{r_1+r_2}{2} \right] \\ \frac{(1-\theta_l)x+\theta_l r_2-r_1}{r_2-r_1+2\theta_l(r_2-x)}, & \text{if } x \in \left(\frac{r_1+r_2}{2}, r_2 \right) \\ \tilde{1}, & \text{if } x \in (r_2, r_3) \\ \frac{(-1+\theta_l)x-\theta_l r_3+r_4}{r_4-r_3+2\theta_l(x-r_3)}, & \text{if } x \in \left(r_3, \frac{r_3+r_4}{2} \right) \\ \frac{(1+\theta_r)(r_4-x)}{r_4-r_3+2\theta_r(r_4-x)}, & \text{if } x \in \left(\frac{r_3+r_4}{2}, r_4 \right). \end{cases} \quad (3.7)$$

From the above theorem we can conclude that when the reduction of trapezoidal type-2 fuzzy variable is made by optimistic CV-reduction method, the possibility distribution function is constructed by use of θ_r and similarly for construction of pessimistic CV-reduction of trapezoidal type-2 fuzzy variable θ_l is use. But in case of CV-reduction of trapezoidal type-2 fuzzy variable both θ_l and θ_r are used. Thus, CV-reduction gives more accurate normal value rather than optimistic or pessimistic CV-reduction. Therefore, we take CV-reduction method for our future calculation.

3.7. Fuzzy number [37]

A fuzzy number is an extension of a regular number in the sense that it does not refer to one single value but rather than a connected set of possible values. Thus, a fuzzy number is a fuzzy set like $u : R \rightarrow I = [0, 1]$ which satisfies

- (1) u is upper semi-continuous.
- (2) $u(x) = 0$ outside the interval $[c, d]$.
- (3) There are real numbers a, b such $c \leq a \leq b \leq d$ and
 - (i) $u(x)$ is monotonic increasing on $[c, a]$,
 - (ii) $u(x)$ is monotonic decreasing on $[b, d]$,
 - (iii) $u(x) = 1$, $a \leq x \leq b$.

3.8. Hexagonal fuzzy number [24]

A fuzzy number \tilde{A}_h is called a hexagonal fuzzy number, denoted by $\tilde{A}_h = (a_1, a_2, a_3, a_4, a_5, a_6)$ where $a_1, a_2, a_3, a_4, a_5, a_6$ are real numbers and its membership function $\mu_{\tilde{A}_h}(x)$ is given below

$$\mu_{\tilde{A}_h}(x) = \begin{cases} 0, & \text{if } x < a_1 \\ \frac{1}{2} \left[\frac{x-a_1}{a_2-a_1} \right], & \text{if } a_1 \leq x \leq a_2 \\ \frac{1}{2} + \frac{1}{2} \left[\frac{x-a_2}{a_3-a_2} \right], & \text{if } a_2 \leq x \leq a_3 \\ 1, & \text{if } a_3 \leq x \leq a_4 \\ 1 - \frac{1}{2} \left[\frac{x-a_4}{a_5-a_4} \right], & \text{if } a_4 \leq x \leq a_5 \\ \frac{1}{2} \left[\frac{a_5-x}{a_5-a_4} \right], & \text{if } a_5 \leq x \leq a_6 \\ 0, & \text{if } x > a_6 \end{cases}$$

Remark: In other words, a hexagonal fuzzy number \tilde{A}_h is an ordered quadruple $(P_1(u), Q_1(v), Q_2(v), P_2(u))$, for $u \in [0, 0.5]$ and $v \in [0.5, 1]$ where,

- (1) $P_1(u) = \frac{1}{2} \left[\frac{u-a_1}{a_2-a_1} \right]$ is a bounded continuous non-decreasing function over $[0, 0.5]$.
- (2) $Q_1(v) = \frac{1}{2} + \frac{1}{2} \left[\frac{v-a_2}{a_3-a_2} \right]$ is a bounded continuous non-decreasing function over $[0.5, 1]$.
- (3) $Q_2(v) = 1 - \frac{1}{2} \left[\frac{v-a_4}{a_5-a_4} \right]$ is a bounded continuous non-decreasing function over $[0.5, 1]$.
- (4) $P_2(u) = \frac{1}{2} \left[\frac{a_6-u}{a_6-a_5} \right]$ is a bounded continuous non-decreasing function over $[0, 0.5]$.

3.9. α -cut of fuzzy set [37]

The α -level set (or interval of confidence at level α or α -cut) of the fuzzy set \tilde{A} of X is a crisp set A_α that contains all the elements of X that have membership values greater than or equal to α , i.e. $\tilde{A} = \{x : \mu_{\tilde{A}}(x) \geq \alpha, x \in X, \alpha \in [0, 1]\}$.

In case of hexagonal fuzzy number $\tilde{A}_h = (a_1, a_2, a_3, a_4, a_5, a_6)$, the α -cut of \tilde{A}_h is defined as

$$\begin{aligned} A_\alpha &= \{x \in X : \mu_{\tilde{A}_h}(x) \geq \alpha\} \\ &= \begin{cases} [P_1(\alpha), P_2(\alpha)] & \text{for } \alpha \in [0, 0.5] \\ [Q_1(\alpha), Q_2(\alpha)] & \text{for } \alpha \in [0.5, 1] \end{cases} \end{aligned}$$

3.10. α -cut operations [37]:

If we get crisp interval by α -cut operations interval A_α shall be obtained as follows, for all $\alpha \in [0, 1]$

Consider, $Q_1(x) = \alpha$

$$\frac{1}{2} + \frac{1}{2} \left[\frac{x-a_2}{a_3-a_2} \right] = \alpha$$

Hence, $Q_1(\alpha) = 2\alpha(a_3 - a_2) + 2a_2 - a_3$

Similarly, $Q_2(x) = \alpha$, $Q_2(\alpha) = 2a_5 - a_4 - 2\alpha(a_5 - a_4)$, $P_1(\alpha) = 2\alpha(a_2 - a_1) + a_1$

$P_2(\alpha) = a_6 - 2\alpha(a_6 - a_5)$

Hence,

$$A_\alpha = \begin{cases} [2\alpha(a_2 - a_1) + a_1, a_6 - 2\alpha(a_6 - a_5)] & \text{for } \alpha \in [0, 0.5] \\ [2\alpha(a_3 - a_2) + 2a_2 - a_3, 2a_5 - a_4 - 2\alpha(a_5 - a_4)] & \text{for } \alpha \in [0.5, 1] \end{cases}$$

3.11. α -cut operation on reduction of a trapezoidal type-2 fuzzy variable

A trapezoidal type-2 fuzzy variable is defined as $\tilde{\xi} = (r_1, r_2, r_3, r_4; \theta_l, \theta_r)$, then we have already discussed about reduction method of trapezoidal type-2 fuzzy variable by optimistic CV , pessimistic CV and CV reduction.

Now according to the definition of α -cut [37], we have the following α -cuts of the reductions of $\tilde{\xi}$

I. Using the optimistic *CV* reduction method,

$$\xi_{1L}(\alpha) = \begin{cases} \frac{(1+\theta_r)r_1+(r_2-r_1-\theta_r r_1)\alpha}{(1+\theta_r)-\theta_r\alpha}, & \text{if } 0 \leq \alpha \leq 0.5 \\ \frac{(r_1-\theta_r r_1)+(r_2-r_1-\theta_r r_2)\alpha}{(1-\theta_r)+\theta_r\alpha}, & \text{if } 0.5 < \alpha \leq 1 \end{cases} \quad (3.8)$$

$$\xi_{1R}(\alpha) = \begin{cases} \frac{(r_4-\theta_r r_3)-(r_4-r_3-\theta_r r_3)\alpha}{(1-\theta_r)-\theta_r\alpha}, & \text{if } 0.5 \leq \alpha \leq 1 \\ \frac{(1+\theta_r)r_4-(r_4-r_3+\theta_r r_4)\alpha}{(1+\theta_r)-\theta_r\alpha}, & \text{if } 0 \leq \alpha < 0.5 \end{cases} \quad (3.9)$$

II. Using the pessimistic *CV* reduction method,

$$\xi_{2L}(\alpha) = \begin{cases} \frac{r_1+(r_2-r_1-\theta_r r_1)\alpha}{1-\theta_l\alpha}, & \text{if } 0 \leq \alpha \leq 0.5 \\ \frac{r_1+(r_2-r_1+\theta_r r_2)\alpha}{1+\theta_l\alpha}, & \text{if } 0.5 < \alpha \leq 1 \end{cases} \quad (3.10)$$

$$\xi_{2R}(\alpha) = \begin{cases} \frac{r_4-(r_4-r_3-\theta_l r_2)\alpha}{1+\theta_l\alpha}, & \text{if } 0.5 \leq \alpha \leq 1 \\ \frac{r_4-(r_4-r_3+\theta_l r_4)\alpha}{1-\theta_l\alpha}, & \text{if } 0 \leq \alpha < 0.5 \end{cases} \quad (3.11)$$

III. Using the *CV* reduction method,

$$\xi_{3L}(\alpha) = \begin{cases} \frac{(1+\theta_r)r_1+(r_2-r_1-2\theta_r r_1)\alpha}{(1+\theta_r)-2\theta_r\alpha}, & \text{if } 0 \leq \alpha \leq 0.5 \\ \frac{(r_1-\theta_l r_2)+(r_2-r_1-\theta_l r_2)\alpha}{(1-\theta_l)+2\theta_l\alpha}, & \text{if } 0.5 < \alpha \leq 1 \end{cases} \quad (3.12)$$

$$\xi_{3R}(\alpha) = \begin{cases} \frac{(r_4-\theta_l r_3)+(r_4-r_3-2\theta_l r_3)\alpha}{(1-\theta_l)+2\theta_l\alpha}, & \text{if } 0.5 \leq \alpha \leq 1 \\ \frac{(1+\theta_r)r_4-(r_4-r_3+2\theta_r r_4)\alpha}{(1+\theta_r)-2\theta_r\alpha}, & \text{if } 0 \leq \alpha < 0.5 \end{cases} \quad (3.13)$$

By *CV* reduction method, membership function of type two fuzzy variable $\tilde{\xi} = (r_1, r_2, r_3, r_4; \theta_l, \theta_r)$ reduces to membership function of type one variable which is just like a hexagonal fuzzy number. Therefore α -cut of $\tilde{\xi}$ is

$$[\tilde{\xi}]_\alpha = \begin{cases} [P_1(\alpha), P_2(\alpha)] & \text{for } \alpha \in [0, 0.5] \\ [Q_1(\alpha), Q_2(\alpha)] & \text{for } \alpha \in [0.5, 1] \end{cases}$$

where,

$$P_1(\alpha) = \frac{(1+\theta_r)r_1+(r_2-r_1-2\theta_r r_1)\alpha}{(1+\theta_r)-2\theta_r\alpha}, \quad P_2(\alpha) = \frac{(1+\theta_r)r_4-(r_4-r_3+2\theta_r r_4)\alpha}{(1+\theta_r)-2\theta_r\alpha}$$

$$Q_1(\alpha) = \frac{(r_1-\theta_l r_2)+(r_2-r_1-\theta_l r_2)\alpha}{(1-\theta_l)+2\theta_l\alpha}, \quad Q_2(\alpha) = \frac{(r_4-\theta_l r_3)+(r_4-r_3-2\theta_l r_3)\alpha}{(1-\theta_l)+2\theta_l\alpha}$$

4. Notations and assumptions

To formulate the mathematical model for the proposed inventory system, the following notations and assumptions are made.

4.1. Notation

Decision Variables:

M : Permissible delay period (time) for the retailer offered by the wholesaler, $M > 0$.
 N : Permissible delay period (time) for the customer offered by the retailer, $0 < N < M$.
 T : Business period i.e., time period for the cycle of the system, $T > 0$.

Parameters:

C_3 = Fixed set-up cost (\$ /set up).
 C_s = Unit selling price (\$ /unit).
 C'_s = Scrap price per unit (\$ /unit).
 p = Unit production cost (\$ /unit).
 α' = Obsolescence rate of inventory (percent).
 b' = Required space for each unit of product (m^3 /unit).
 a' = The weight of the obsolescence product stored in the warehouse (ton/ m^3).
 C_b = Backordering cost per unit quantity per unit of time (\$ /unit/time).

C_{mc} = The emission cost of carbon for manufacturing each unit (\$ / m^3).

C_{oc} = Average disposal, waste collection and emission cost for inventory obsolescence (\$ / m^3).

C_h = Holding cost per unit item per unit time (\$ /unit).

C_{hc} = Average emission cost of carbon for holding inventory (\$ / m^3).

i_p = Rate of interest per year per unit to be paid for the unsold inventory after the credit period M , $i_p > 0$ (\$ /year/unit).

i_e = Rate of interest per year per unit to be earned from the revenue sold till the time horizon T (\$ /year/unit) ($i_e < i_p$).

4.2. Assumptions

The model is developed with the following assumptions.

- (1) Production system involves only one non-deteriorating item.
- (2) Shortages are allowed with partial backordering.
- (3) D , rate of demand depends on the production price and stock i.e.,

$$D(p, q) = \begin{cases} p^{-\epsilon}(\tilde{a} + bq) & \text{if } q > 0 \\ \tilde{a}p^{-\epsilon} & \text{if } q \leq 0 \end{cases}$$

Where $\tilde{a} = (r_1, r_2, r_3, r_4, \theta_l, \theta_r)$ is a trapezoidal type two fuzzy number and $0 < \epsilon < 1$, $r_1, r_2, r_3, r_4 > 0$, $0 < \theta_l, \theta_r < 1$ and b is any positive real number.

- (4) K , rate of production is linearly demand dependent i.e., of the form $K = \mu D$ and $\mu > 1$
- (5) Rate of earning interest by the retailer is lesser than the rate of interest paid to the wholesaler by the retailer, i.e. $i_e < i_p$.
- (6) Credit period offered by the retailer is smaller than that offered by the wholesaler, i.e. $N < M$.
- (7) Customer maintains the trade credit policy offered by the retailer.

5. Mathematical formulation of the model

Let the retailer fails to fulfill the demand initially and hence shortages arise from time $t = 0$ to the time $t = t_1$ and maximum shortage level Q_s occur at $t = t_1$. After that production process starts to backlog the shortage quantities with partial backordering process and at time $t = t_2$ the shortage level reaches to zero. In the mean time inventory accumulates upto time $t = t_3$ of amount Q_m . At that time production process being stop and the accumulated inventory declines to meet up the customers demand and reaches to zero at time T .

The governing differential equations of the stock level at any instant t for this model is given by

$$\frac{dq}{dt} = \begin{cases} -D & 0 \leq t \leq t_1 \\ K - D & t_1 \leq t \leq t_2 \\ K - D & t_2 \leq t \leq t_3 \\ -D & t_3 \leq t \leq T \end{cases}$$

With the boundary conditions, $q(0) = q(t_2) = q(T) = 0$

Bede and Gal [4] applied fuzzy number valued function in fuzzy differential equation and hence, the above equation can be rewritten in fuzzy form

$$\frac{d\tilde{q}}{dt} = \begin{cases} -\tilde{a} \odot p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) \odot \tilde{a} \odot p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) \odot p^{-\epsilon} \odot (\tilde{a} + b \odot \tilde{q}) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon} \odot (\tilde{a} + b \odot \tilde{q}) & t_3 \leq t \leq T \end{cases}$$

For using CV based reduction method for trapezoidal type-2 fuzzy number, we divided the interval for $\alpha \in [0, 1]$ into two sub-intervals like $\alpha \in [0, 0.5]$ and $\alpha \in [0.5, 1]$ as discussed

in preliminaries. Also Stefanini and Bede [33] shows that any general interval differential equation can be formulated via generalized Hukuhara (gH) derivative in terms of system of ordinary differential equations.

Now, the Generalized Hukuhara derivative [33] of a fuzzy valued function $f : [a, b] \mapsto R$ at t_0 is defined as

$$f'(t_0) = \lim_{h \rightarrow 0} \frac{f(t_0 + h) \ominus_{gH} f(t_0)}{h}$$

In parametric form we say that $f(t)$ is gH-(i) differentiable at t_0 if

$$[f'(t_0)]_\alpha = [f'_L(t_0, \alpha), f'_R(t_0, \alpha)] \tag{5.1}$$

Also, $f(t)$ is gH-(ii) differentiable at t_0 if

$$[f'(t_0)]_\alpha = [f'_R(t_0, \alpha), f'_L(t_0, \alpha)] \tag{5.2}$$

Depending upon the value of α two possibility arises

(i) $\alpha \in [0, 0.5]$

(ii) $\alpha \in [0.5, 1]$

If $\alpha \in [0, 0.5]$, then the above equation takes the form

$$\left[\frac{dq_L}{dt}, \frac{dq_R}{dt} \right] = \begin{cases} -[a_L, a_R] \odot p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) \odot [a_L, a_R] \odot p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) \odot p^{-\epsilon}([a_L, a_R] + b \odot [q_L, q_R]) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon}([a_L, a_R] + b \odot [q_L, q_R]) & t_3 \leq t \leq T \end{cases} \tag{5.3}$$

Therefore two cases arise

Case 1: gH-(i) differentiable

Case 2: gH-(ii) differentiable

Case-1: Therefore, solving the above system via gH-(i) differentiable is equivalent to solve the corresponding simultaneous system (see [33])

$$\frac{dq_L}{dt} = \begin{cases} -a_R p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) a_L p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) p^{-\epsilon} (a_L + b q_L) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon} (a_R + b q_R) & t_3 \leq t \leq T \end{cases}$$

$$\frac{dq_R}{dt} = \begin{cases} -a_L p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) a_R p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) p^{-\epsilon} (a_R + b q_R) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon} (a_L + b q_L) & t_3 \leq t \leq T \end{cases}$$

After solving using the boundary conditions, $q(0) = q(t_2) = q(T) = 0$, we get

$$q_L(t) = \begin{cases} -a_R p^{-\epsilon} t & 0 \leq t \leq t_1 \\ (\mu - 1) a_L p^{-\epsilon} (t - t_2) & t_1 \leq t \leq t_2 \\ \frac{a_L}{b} [e^{x(t-t_2)} - 1] & t_2 \leq t \leq t_3 \\ -\frac{a_R}{b} [e^{y(T-t)} - 1] & t_3 \leq t \leq T \end{cases} \tag{5.4}$$

$$q_R(t) = \begin{cases} -a_L p^{-\epsilon} t & 0 \leq t \leq t_1 \\ (\mu - 1) a_R p^{-\epsilon} (t - t_2) & t_1 \leq t \leq t_2 \\ \frac{a_R}{b} [e^{x(t-t_2)} - 1] & t_2 \leq t \leq t_3 \\ -\frac{a_L}{b} [e^{y(T-t)} - 1] & t_3 \leq t \leq T \end{cases} \tag{5.5}$$

where, $x = b(\mu - 1)p^{-\epsilon}$ and $y = bp^{-\epsilon}$

$$a_L(\alpha) = \frac{(1 + \theta_r)r_1 + (r_2 - r_1 - 2\theta_r r_1)\alpha}{(1 + \theta_r) - 2\theta_r \alpha}$$

$$a_R(\alpha) = \frac{(1 + \theta_r)r_4 - (r_4 - r_3 + 2\theta_r r_4)\alpha}{(1 + \theta_r) - 2\theta_r \alpha}$$

Similarly, to find the lower and upper limit of inventory accumulated upto time $t = t_3$ is given by the condition

$$(Q_m)_L = \frac{a_L}{b} [e^{x(t_3-t_2)} - 1] = -\frac{a_R}{b} [e^{y(T-t_3)} - 1]$$

$$(Q_m)_U = \frac{a_R}{b} [e^{x(t_3-t_2)} - 1] = -\frac{a_L}{b} [e^{y(T-t_3)} - 1]$$

Now, the inventory related costs are as follows

Total obsolescence cost of inventory

$$TOC_L = \alpha' (C_s - C'_s) \left[\int_{t_2}^{t_3} q_L(t) dt + \int_{t_3}^T q_L(t) dt \right] = \alpha' (C_s - C'_s) \left[\frac{a_L}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_R}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

$$TOC_R = \alpha' (C_s - C'_s) \left[\int_{t_2}^{t_3} q_R(t) dt + \int_{t_3}^T q_R(t) dt \right] = \alpha' (C_s - C'_s) \left[\frac{a_R}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_L}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

Total cost of emission of inventory obsolescence

$$TEO_L = \alpha' a' C_{oc} \left[\int_{t_2}^{t_3} q_L(t) dt + \int_{t_3}^T q_L(t) dt \right] = \alpha' a' C_{oc} \left[\frac{a_L}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_R}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

$$TEO_R = \alpha' a' C_{oc} \left[\int_{t_2}^{t_3} q_R(t) dt + \int_{t_3}^T q_R(t) dt \right] = \alpha' a' C_{oc} \left[\frac{a_R}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_L}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

Total backordering cost

$$TBC_L = C_b \left[\int_0^{t_1} q_L(t) dt + \int_{t_1}^{t_2} q_L(t) dt \right] = -\frac{C_b}{2} \left[a_R p^{-\epsilon} t_1^2 + (\mu - 1) a_L p^{-\epsilon} (t_1 - t_2)^2 \right]$$

$$TBC_R = C_b \left[\int_0^{t_1} q_R(t) dt + \int_{t_1}^{t_2} q_R(t) dt \right] = -\frac{C_b}{2} \left[a_L p^{-\epsilon} t_1^2 + (\mu - 1) a_R p^{-\epsilon} (t_1 - t_2)^2 \right]$$

Total goodwill loss for back-order is

$$TGC_L = C_g \left[\int_0^{t_1} q_L(t) dt + \int_{t_1}^{t_2} q_L(t) dt \right] = -\frac{C_g}{2} \left[a_R p^{-\epsilon} t_1^2 + (\mu - 1) a_L p^{-\epsilon} (t_1 - t_2)^2 \right]$$

$$TGC_R = C_g \left[\int_0^{t_1} q_R(t) dt + \int_{t_1}^{t_2} q_R(t) dt \right] = -\frac{C_g}{2} \left[a_L p^{-\epsilon} t_1^2 + (\mu - 1) a_R p^{-\epsilon} (t_1 - t_2)^2 \right]$$

Total holding cost

$$THC_L = \int_{t_2}^{t_3} C_h q_L(t) dt + \int_{t_3}^T C_h q_L(t) dt = C_h \left[\frac{a_L}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_R}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

$$THC_R = \int_{t_2}^{t_3} C_h q_R(t) dt + \int_{t_3}^T C_h q_R(t) dt = C_h \left[\frac{a_R}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_L}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

Total emission cost of carbon for holding inventory

$$TEH_L = C_{hc} b' \left[\int_{t_2}^{t_3} q_L(t) dt + \int_{t_3}^T q_L(t) dt \right] = C_{hc} b' \left[\frac{a_L}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_R}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

$$TEH_R = C_{hc} b' \left[\int_{t_2}^{t_3} q_R(t) dt + \int_{t_3}^T q_R(t) dt \right] = C_{hc} b' \left[\frac{a_R}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_L}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

Total set-up cost = \bar{C}_3

Total revenue earned

$$TRE_L = c_S \left[\mu a_L p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_L + b q_L) dt \right] = c_S \left[\mu a_L p^{-\epsilon} (t_2 - t_1) + \mu p^{-\epsilon} a_L e^{-xt_2} \int_{t_2}^{t_3} e^{xt} dt \right]$$

$$= c_S \mu a_L p^{-\epsilon} \left[t_2 - t_1 + \frac{1}{x} (e^{x(t_3-t_2)} - 1) \right]$$

$$TRE_R = c_S \left[\mu a_{RP} p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_R + bq_R) dt \right] = c_S \left[\mu a_{RP} p^{-\epsilon} (t_2 - t_1) + \mu p^{-\epsilon} a_R e^{-xt_2} \int_{t_2}^{t_3} e^{xt} dt \right]$$

$$= c_S \mu a_{RP} p^{-\epsilon} \left[t_2 - t_1 + \frac{1}{x} (e^{x(t_3-t_2)} - 1) \right]$$

Total Production cost

$$TPC_L = p \left[\mu a_{LP} p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_L + bq_L) dt \right] = p \left[\mu a_{LP} p^{-\epsilon} (t_2 - t_1) + \mu p^{-\epsilon} a_L e^{-xt_2} \int_{t_2}^{t_3} e^{xt} dt \right]$$

$$= p \mu a_{LP} p^{-\epsilon} \left[t_2 - t_1 + \frac{1}{x} (e^{x(t_3-t_2)} - 1) \right]$$

$$TPC_R = p \left[\mu a_{RP} p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_R + bq_R) dt \right] = p \left[\mu a_{RP} p^{-\epsilon} (t_2 - t_1) + \mu p^{-\epsilon} a_R e^{-xt_2} \int_{t_2}^{t_3} e^{xt} dt \right]$$

$$= p \mu a_{RP} p^{-\epsilon} \left[t_2 - t_1 + \frac{1}{x} (e^{x(t_3-t_2)} - 1) \right]$$

The emission cost of carbon for manufacturing total units

$$TEP_L = c_{mc} \left[\mu a_{LP} p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_L + bq_L) dt \right] = c_{mc} \left[\mu a_{LP} p^{-\epsilon} (t_2 - t_1) + \mu p^{-\epsilon} a_L e^{-xt_2} \int_{t_2}^{t_3} e^{xt} dt \right]$$

$$= c_{mc} \mu a_{LP} p^{-\epsilon} \left[t_2 - t_1 + \frac{1}{x} (e^{x(t_3-t_2)} - 1) \right]$$

$$TEP_R = c_{mc} \left[\mu a_{RP} p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_R + bq_R) dt \right] = c_{mc} \left[\mu a_{RP} p^{-\epsilon} (t_2 - t_1) + \mu p^{-\epsilon} a_R e^{-xt_2} \int_{t_2}^{t_3} e^{xt} dt \right]$$

$$= c_{mc} \mu a_{RP} p^{-\epsilon} \left[t_2 - t_1 + \frac{1}{x} (e^{x(t_3-t_2)} - 1) \right]$$

Therefore three sub-cases arise depending upon the values of changing time periods.

Sub-case 1.1: $0 \leq t_1 \leq N \leq M \leq t_2 \leq t_3 \leq T$

In this case if the amount is paid within M by the retailer, then there is no interest payable. Otherwise, the retailer will pay for the rest of the inventory. Hence, the total amount of interest paid and interest earned by the retailer is calculated.

Hence, total interest paid

$$TIP_L = pi_p \left[\int_M^{t_2} (1 + T - t_2) q_L(t) dt + \int_{t_2}^{t_3} (1 + T - t_3) q_L(t) dt + \int_{t_3}^T q_L(t) dt \right]$$

$$= pi_p \left[\frac{1}{2} (1 + T - t_2) (\mu - 1) a_{LP} p^{-\epsilon} (t_2 - M)^2 + (1 + T - t_3) \frac{a_L}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_R}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

$$TIP_R = pi_p \left[\int_M^{t_2} (1 + T - t_2) q_R(t) dt + \int_{t_2}^{t_3} (1 + T - t_3) q_R(t) dt + \int_{t_3}^T q_R(t) dt \right]$$

$$TIP_R = C_p i_p \left[\frac{1}{2} (1 + T - t_2) (\mu - 1) a_{RP} p^{-\epsilon} (t_2 - M)^2 + (1 + T - t_3) \frac{a_R}{b} \left\{ \frac{1}{x} (e^{x(t_3-t_2)} - 1) - (t_3 - t_2) \right\} + \frac{a_L}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right]$$

And, total interest earned

$$TIE_L = C_s i_e \left[(T - N) \int_{t_1}^N D_L(p, q) dt + (1 + T - M) \int_N^M D_L(p, q) (M - t) dt + (1 + T - t_2) \int_M^{t_2} D_L(p, q) (t_2 - t) dt + (1 + T - t_3) \int_{t_2}^{t_3} D_L(p, q) (t_3 - t) dt + \int_{t_3}^T D_L(p, q) (T - t) dt \right]$$

$$= C_s i_e a_{LP} p^{-\epsilon} \left[(T - N)(N - t_1) + \frac{1}{2} (1 + T - M)(M - N)^2 + \frac{1}{2} (1 + T - t_2)(t_2 - M)^2 + p^{-\epsilon} \{ a_L - a_R (e^{y(T-t_3)} - 1) \} \right]$$

$$TIE_R = C_s i_e \left[(T - N) \int_{t_1}^N D_R(p, q) dt + (1 + T - M) \int_N^M D_R(p, q) (M - t) dt + (1 + T - t_2) \int_M^{t_2} D_R(p, q) (t_2 - t) dt + (1 + T - t_3) \int_{t_2}^{t_3} D_R(p, q) (t_3 - t) dt + \int_{t_3}^T D_R(p, q) (T - t) dt \right]$$

$$= C_s i_e a_{LP} p^{-\epsilon} \left[(T - N)(N - t_1) + \frac{1}{2} (1 + T - M)(M - N)^2 + \frac{1}{2} (1 + T - t_2)(t_2 - M)^2 + p^{-\epsilon} \{ a_L - a_R (e^{y(T-t_3)} - 1) \} \right]$$

Sub-case 1.2: $0 \leq t_1 \leq t_2 \leq N \leq M \leq t_3 \leq T$

In this case the total amount of interest paid and interest earned by the retailer is calculated.

Hence, total interest paid

$$\begin{aligned} TIP_L &= pi_p \left[\int_M^{t_3} (1 + T - t_3) q_L(t) dt + \int_{t_3}^T q_L(t) dt \right] \\ &= pi_p \left[(1 + T - t_3) \frac{a_L}{b} \left\{ \frac{1}{x} (e^{x(t_3-M)} - 1) - (t_3 - M) \right\} + \frac{a_R}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right] \\ TIP_L &= pi_p \left[\int_M^{t_3} (1 + T - t_3) q_R(t) dt + \int_{t_3}^T q_R(t) dt \right] \\ &= pi_p \left[(1 + T - t_3) \frac{a_L}{b} \left\{ \frac{1}{x} (e^{x(t_3-M)} - 1) - (t_3 - M) \right\} + \frac{a_L}{b} \left\{ \frac{1}{y} (1 - e^{y(T-t_3)}) + (T - t_3) \right\} \right] \end{aligned}$$

Total interest earned

$$\begin{aligned} TIE_L &= C_s i_e \left[(T - t_2) \int_{t_1}^{t_2} D_L(p, q) dt + (T - N) \int_{t_2}^N D_L(p, q) dt + (1 + T - t_3) \int_N^{t_3} D_L(p, q) (t_3 - t) dt + (1 + T - M) \int_{t_3}^M D_L(p, q) (M - t) dt + \int_M^T D_L(p, q) (T - t) dt \right] \\ &= C_s i_e a_L p^{-\epsilon} \left[(T - t_2)(t_2 - t_1) + (T - N) \frac{a_L}{x} (e^{x(N-t_2)} - 1) + \frac{1}{2} (1 + T - t_3)(t_3 - N)^2 + \frac{1}{2} (1 + T - M)(M - t_3)^2 + p^{-\epsilon} \{ a_L - a_R (e^{y(T-t_3)} - 1) \} \right] \\ TIE_R &= C_s i_e \left[(T - t_2) \int_{t_1}^{t_2} D_R(p, q) dt + (T - N) \int_{t_2}^N D_R(p, q) dt + (1 + T - t_3) \int_N^{t_3} D_R(p, q) (t_3 - t) dt + (1 + T - M) \int_{t_3}^M D_R(p, q) (M - t) dt + \int_M^T D_R(p, q) (T - t) dt \right] \\ &= C_s i_e a_R p^{-\epsilon} \left[(T - t_2)(t_2 - t_1) + (T - N) \frac{a_L}{x} (e^{x(N-t_2)} - 1) + \frac{1}{2} (1 + T - t_3)(t_3 - N)^2 + \frac{1}{2} (1 + T - M)(M - t_3)^2 + p^{-\epsilon} \{ a_R - a_L (e^{y(T-t_3)} - 1) \} \right] \end{aligned}$$

Sub-case 1.3: $0 \leq t_1 \leq t_2 \leq N \leq t_3 \leq T \leq M$

In this case interest payable by the retailer is zero, i.e., $TIP_L = TIP_R = 0$

But, interest earned by the retailer is given by

$$\begin{aligned} TIE_L &= C_s i_e \left[(M - t_2) \int_{t_1}^{t_2} D_L(p, q) dt + (M - N) \int_{t_2}^N D_L(p, q) dt + (1 + M - t_3) \int_N^{t_3} D_L(p, q) (t_3 - t) dt + (1 + M - T) \int_{t_3}^T D_L(p, q) (T - t) dt \right] \\ &= C_s i_e a_L p^{-\epsilon} \left[(M - t_2)(t_2 - t_1) + (M - N)(N - t_1) + \frac{1}{2} (1 + M - t_3)(t_3 - N)^2 + \frac{1}{2} (1 + M - T)(T - t_3)^2 + \{ a_L - a_R (e^{y(T-t_3)} - 1) \} \right] \\ TIE_R &= C_s i_e \left[(M - t_2) \int_{t_1}^{t_2} D_R(p, q) dt + (M - N) \int_{t_2}^N D_R(p, q) dt + (1 + M - t_3) \int_N^{t_3} D_R(p, q) (t_3 - t) dt + (1 + M - T) \int_{t_3}^T D_R(p, q) (T - t) dt \right] \\ &= C_s i_e a_R p^{-\epsilon} \left[(M - t_2)(t_2 - t_1) + (M - N)(N - t_1) + \frac{1}{2} (1 + M - t_3)(t_3 - N)^2 + \frac{1}{2} (1 + M - T)(T - t_3)^2 + \{ a_R - a_L (e^{y(T-t_3)} - 1) \} \right] \end{aligned}$$

Here, a carbon emissions integrated fuzzy EPQ model with type-2 fuzzy variable is considered, where the objectives are maximizing the profit and minimizing various carbon emission costs associated with inventory management. By using gH-differentiability, we get a range for profit and a range of emission rather than an exact value of profit and emission, which is more realistic in practical value. Hence, the objective functions are defined as follows:

$$\text{Max Profit } (TP) = \frac{1}{T} [TRE + TIE - TPC - THC - TOC - TIP - TBC - TGC - C_3] \quad (19)$$

$$\text{Min Carbon Emission Cost } (TE) = TEO + TEH + TEP \quad (20)$$

On taking α -cut over the total profit and emission per unit time is a interval crisp set and is defined by

$$TP = [TP_L, TP_R] \text{ where,} \quad (21)$$

$$TP_L = \frac{1}{T} [TRE_L + TIE_L - TPC_L - THC_L - TOC_L - TIP_L - TBC_L - TGC_L - C_3]$$

$$TP_R = \frac{1}{T} [TRE_R + TIE_R - TPC_R - THC_R - TOC_R - TIP_R - TBC_R - TGC_R - C_3] \quad (22)$$

And, $TE = [TE_L, TE_R]$ where,

$$TE_L = TEO_L + TEH_L + TEP_L$$

$$TE_R = TEO_R + TEH_R + TEP_R, \text{ when } 0 \leq \alpha \leq 0.5$$

Case-2: If $\alpha \in [0, 0.5]$, then on taking the α -cut of fuzzy differential equation reduces to interval fuzzy differential equation via gH-(ii) differentiability

$$\left[\frac{dq_L}{dt}, \frac{dq_R}{dt} \right] = \begin{cases} -[a_L, a_R] \odot p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) \odot [a_L, a_R] \odot p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) \odot p^{-\epsilon} ([a_L, a_R] + b \odot [q_L, q_R]) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon} ([a_L, a_R] + b \odot [q_L, q_R]) & t_3 \leq t \leq T \end{cases} \quad (23)$$

Thus the above system is equivalent to

$$\frac{dq_L}{dt} = \begin{cases} -a_L p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) a_R p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) p^{-\epsilon} (a_R + b q_R) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon} (a_L + b q_L) & t_3 \leq t \leq T \end{cases}$$

$$\frac{dq_R}{dt} = \begin{cases} -a_R p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) a_L p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) p^{-\epsilon} (a_L + b q_L) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon} (a_R + b q_R) & t_3 \leq t \leq T \end{cases}$$

After solving using the boundary conditions, $q(0) = q(t_2) = q(T) = 0$ and $q(t_2) = Q_m$, we get

$$q_L(t) = \begin{cases} -a_L p^{-\epsilon} t & 0 \leq t \leq t_1 \\ (\mu - 1) a_R p^{-\epsilon} (t - t_2) & t_1 \leq t \leq t_2 \\ c_1 e^{bkt} + c_2 e^{-bkt} - \frac{a_L}{b} & t_2 \leq t \leq t_3 \\ k_1 e^{r_2 t} + k_2 e^{-r_2 t} + \frac{a_L}{b} & t_3 \leq t \leq T \end{cases} \quad (24)$$

$$q_R(t) = \begin{cases} -a_R p^{-\epsilon} t & 0 \leq t \leq t_1 \\ (\mu - 1) a_L p^{-\epsilon} (t - t_2) & t_1 \leq t \leq t_2 \\ c_1 e^{bkt} - c_2 e^{-bkt} - \frac{a_R}{b} & t_2 \leq t \leq t_3 \\ -k_1 e^{r_2 t} + k_2 e^{-r_2 t} + \frac{a_R}{b} & t_3 \leq t \leq T \end{cases} \quad (25)$$

Where $k = (\mu - 1) p^{-\epsilon}$, $r_2 = -b p^{-\epsilon}$

$$c_1 = \frac{1}{e^{2bkt_3} - e^{2bkt_2}} \left[\left(Q_m + \frac{a_L}{b} \right) e^{bkt_3} - \frac{a_L}{b} e^{bkt_2} \right]$$

$$c_2 = \frac{a_L}{b} e^{bkt_2} - \frac{1}{e^{2bk(t_3-t_2)}} \left[\left(Q_m + \frac{a_L}{b} \right) e^{bkt_3} - \frac{a_L}{b} e^{bkt_2} \right]$$

$$k_1 = \frac{a_R - a_L}{2b} e^{-c_2 T}$$

$$k_2 = -\frac{a_L + a_R}{2b} e^{c_2 T}$$

$$a_L(\alpha) = \frac{(1 + \theta_r) r_1 + (r_2 - r_1 - 2\theta_r r_1) \alpha}{(1 + \theta_r) - 2\theta_r \alpha}$$

$$a_R(\alpha) = \frac{(1 + \theta_r) r_4 - (r_4 - r_3 + 2\theta_r r_4) \alpha}{(1 + \theta_r) - 2\theta_r \alpha}$$

Various costs related to the above system are

Total obsolescence cost of inventory

$$\begin{aligned}
TOC_L &= \alpha' (C_s - C'_s) \left[\int_{t_2}^{t_3} q_L(t) dt + \int_{t_3}^T q_L(t) dt \right] \\
&= \alpha' (C_s - C'_s) \left[\int_{t_2}^{t_3} [c_1 e^{bkt} + c_2 e^{-bkt} - \frac{a_L}{b}] dt + \int_{t_3}^T [k_1 e^{r_2 t} + k_2 e^{-r_2 t} + \frac{a_L}{b}] dt \right] \\
&= \alpha' (C_s - C'_s) \left[\frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) - \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_L}{b} (t_3 - t_2) + \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \right. \\
&\quad \left. \frac{k_2}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) + \frac{a_L}{b} (T - t_3) \right]
\end{aligned}$$

$$\begin{aligned}
TOC_R &= \alpha' (C_s - C'_s) \left[\int_{t_2}^{t_3} q_R(t) dt + \int_{t_3}^T q_R(t) dt \right] \\
&= \alpha' (C_s - C'_s) \left[\int_{t_2}^{t_3} [c_1 e^{bkt} - c_2 e^{-bkt} - \frac{a_R}{b}] dt + \int_{t_3}^T [-k_1 e^{r_2 t} + k_2 e^{-r_2 t} + \frac{a_R}{b}] dt \right] \\
&= \alpha' (C_s - C'_s) \left[\frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) + \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_R}{b} (t_3 - t_2) - \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \right. \\
&\quad \left. \frac{k_2}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) + \frac{a_R}{b} (T - t_3) \right]
\end{aligned}$$

Total cost of emission of inventory obsolescence

$$\begin{aligned}
TEO_L &= \alpha' a' C_{oc} \left[\int_{t_2}^{t_3} q_L(t) dt + \int_{t_3}^T q_L(t) dt \right] \\
&= \alpha' a' C_{oc} \left[\int_{t_2}^{t_3} [c_1 e^{bkt} + c_2 e^{-bkt} - \frac{a_L}{b}] dt + \int_{t_3}^T [k_1 e^{r_2 t} + k_2 e^{-r_2 t} + \frac{a_L}{b}] dt \right] \\
&= \alpha' a' C_{oc} \left[\frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) - \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_L}{b} (t_3 - t_2) + \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{k_2}{r_2} (e^{-r_2 T} - \right. \\
&\quad \left. e^{-r_2 t_3}) + \frac{a_L}{b} (T - t_3) \right]
\end{aligned}$$

$$\begin{aligned}
TEO_R &= \alpha' a' C_{oc} \left[\int_{t_2}^{t_3} q_R(t) dt + \int_{t_3}^T q_R(t) dt \right] \\
&= \alpha' a' C_{oc} \left[\int_{t_2}^{t_3} [c_1 e^{bkt} - c_2 e^{-bkt} - \frac{a_R}{b}] dt + \int_{t_3}^T [-k_1 e^{r_2 t} + k_2 e^{-r_2 t} + \frac{a_R}{b}] dt \right] \\
&= \alpha' a' C_{oc} \left[\frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) + \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_R}{b} (t_3 - t_2) - \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{k_2}{r_2} (e^{-r_2 T} - \right. \\
&\quad \left. e^{-r_2 t_3}) + \frac{a_R}{b} (T - t_3) \right]
\end{aligned}$$

Total backordering cost

$$TBC_L = C_b \left[\int_0^{t_1} q_L(t) dt + \int_{t_1}^{t_2} q_L(t) dt \right] = c_b \left[-\frac{1}{2} a_L p^{-\epsilon} t_1^2 - \frac{1}{2} (\mu - 1) a_R p^{-\epsilon} (t_2 - t_1)^2 \right]$$

$$TBC_R = C_b \left[\int_0^{t_1} q_R(t) dt + \int_{t_1}^{t_2} q_R(t) dt \right] = c_b \left[-\frac{1}{2} a_R p^{-\epsilon} t_1^2 - \frac{1}{2} (\mu - 1) a_L p^{-\epsilon} (t_2 - t_1)^2 \right]$$

Total goodwill loss for back-order is

$$TGC_L = C_g \left[\int_0^{t_1} q_L(t) dt + \int_{t_1}^{t_2} q_L(t) dt \right] = c_g \left[-\frac{1}{2} a_L p^{-\epsilon} t_1^2 - \frac{1}{2} (\mu - 1) a_R p^{-\epsilon} (t_2 - t_1)^2 \right]$$

$$TGC_R = C_g \left[\int_0^{t_1} q_R(t) dt + \int_{t_1}^{t_2} q_R(t) dt \right] = c_g \left[-\frac{1}{2} a_R p^{-\epsilon} t_1^2 - \frac{1}{2} (\mu - 1) a_L p^{-\epsilon} (t_2 - t_1)^2 \right]$$

Total holding cost

$$\begin{aligned}
THC_L &= \int_{t_2}^{t_3} C_h q_L(t) dt + \int_{t_3}^T C_h q_L(t) dt \\
&= C_h \left[\frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) - \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_L}{b} (t_3 - t_2) + \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{k_2}{r_2} (e^{-r_2 T} - \right. \\
&\quad \left. e^{-r_2 t_3}) + \frac{a_L}{b} (T - t_3) \right]
\end{aligned}$$

$$\begin{aligned}
THC_R &= \int_{t_2}^{t_3} C_h q_R(t) dt + \int_{t_3}^T C_h q_R(t) dt \\
&= C_h \left[\frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) + \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_R}{b} (t_3 - t_2) - \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{k_2}{r_2} (e^{-r_2 T} - \right. \\
&\quad \left. e^{-r_2 t_3}) + \frac{a_R}{b} (T - t_3) \right]
\end{aligned}$$

Total emission cost of carbon for holding inventory

$$\begin{aligned}
TEH_L &= C_{hc} b' \left[\int_{t_2}^{t_3} q_L(t) dt + \int_{t_3}^T q_L(t) dt \right] \\
&= C_{hc} b' \left[\frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) - \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_L}{b} (t_3 - t_2) + \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{k_2}{r_2} (e^{-r_2 T} - \right. \\
&\quad \left. e^{-r_2 t_3}) + \frac{a_L}{b} (T - t_3) \right]
\end{aligned}$$

$$TEHR = Chcb' \left[\int_{t_2}^{t_3} q_R(t) dt + \int_{t_3}^T q_R(t) dt \right]$$

$$= Chcb' \left[\frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) + \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_R}{b} (t_3 - t_2) - \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{k_2}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) + \frac{a_R}{b} (T - t_3) \right]$$

Total set-up cost = C_3

Total revenue earned

$$TRE_L = c_S [\mu a_L p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_L + bq_L) dt]$$

$$= c_S [\mu a_L p^{-\epsilon} (t_2 - t_1) + \frac{\mu p^{-\epsilon}}{k} \{c_1 (e^{bkt_3} - e^{bkt_2}) - c_2 (e^{-bkt_3} - e^{-bkt_2})\}]$$

$$TRE_R = c_S [\mu a_R p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_R + bq_R) dt]$$

$$= c_S [\mu a_R p^{-\epsilon} (t_2 - t_1) + \frac{\mu p^{-\epsilon}}{k} \{c_1 (e^{bkt_3} - e^{bkt_2}) + c_2 (e^{-bkt_3} - e^{-bkt_2})\}]$$

Total Production cost

$$TPC_L = p [\mu a_L p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_L + bq_L) dt]$$

$$= p [\mu a_L p^{-\epsilon} (t_2 - t_1) + \frac{\mu p^{-\epsilon}}{k} \{c_1 (e^{bkt_3} - e^{bkt_2}) - c_2 (e^{-bkt_3} - e^{-bkt_2})\}]$$

$$TPC_R = p [\mu a_R p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_R + bq_R) dt]$$

$$= p [\mu a_R p^{-\epsilon} (t_2 - t_1) + \frac{\mu p^{-\epsilon}}{k} \{c_1 (e^{bkt_3} - e^{bkt_2}) + c_2 (e^{-bkt_3} - e^{-bkt_2})\}]$$

The emission cost of carbon for manufacturing total units

$$TEP_L = c_{mc} [\mu a_L p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_L + bq_L) dt]$$

$$= c_{mc} [\mu a_L p^{-\epsilon} (t_2 - t_1) + \frac{\mu p^{-\epsilon}}{k} \{c_1 (e^{bkt_3} - e^{bkt_2}) - c_2 (e^{-bkt_3} - e^{-bkt_2})\}]$$

$$TEP_R = c_{mc} [\mu a_R p^{-\epsilon} (t_2 - t_1) + \mu \int_{t_2}^{t_3} p^{-\epsilon} (a_R + bq_R) dt]$$

$$= c_{mc} [\mu a_R p^{-\epsilon} (t_2 - t_1) + \frac{\mu p^{-\epsilon}}{k} \{c_1 (e^{bkt_3} - e^{bkt_2}) + c_2 (e^{-bkt_3} - e^{-bkt_2})\}]$$

Therefore three sub-cases may arise depending upon the values of changing time periods.

Sub-case 2.1: $0 \leq t_1 \leq N \leq M \leq t_2 \leq t_3 \leq T$

In this case if the amount is paid within M by the retailer, then there is no interest payable. Otherwise, the retailer will pay for the rest of the inventory. In this case the total amount of interest paid and interest earned by the retailer is calculated.

Hence, total interest paid

$$TIP_L = pi_p [\int_M^{t_2} (1 + T - t_2) q_L(t) dt + \int_{t_2}^{t_3} (1 + T - t_3) q_L(t) dt + \int_{t_3}^T q_L(t) dt]$$

$$= pi_p \left[\frac{1}{2} (1 + T - t_2) (\mu - 1) a_L p^{-\epsilon} (t_2 - M)^2 + (1 + T - t_3) \left\{ \frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) - \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_L}{b} (t_3 - t_2) \right\} + \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{k_2}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) + \frac{a_L}{b} (T - t_3) \right]$$

$$TIP_R = pi_p \left[\int_M^{t_2} (1 + T - t_2) q_R(t) dt + \int_{t_2}^{t_3} (1 + T - t_3) q_R(t) dt + \int_{t_3}^T q_R(t) dt \right]$$

$$= pi_p \left[\frac{1}{2} (1 + T - t_2) (\mu - 1) a_R p^{-\epsilon} (t_2 - M)^2 + (1 + T - t_3) \left\{ \frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) + \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) - \frac{a_R}{b} (t_3 - t_2) \right\} - \frac{k_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{k_2}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) + \frac{a_R}{b} (T - t_3) \right]$$

And, total interest earned

$$TIE_L = C_s i_e [(T - N) \int_{t_1}^N D_L(p, q) dt + (1 + T - M) \int_N^M D_L(p, q) (M - t) dt + (1 + T - t_2) \int_M^{t_2} D_L(p, q) (t_2 - t) dt + (1 + T - t_3) \int_{t_2}^{t_3} D_L(p, q) (t_3 - t) dt + \int_{t_3}^T D_L(p, q) (T - t) dt]$$

$$= C_s i_e \left[(T - N) a_L p^{-\epsilon} (N - t_1) + \frac{1}{2} (1 + T - M) a_L p^{-\epsilon} (M - N)^2 + \frac{1}{2} (1 + T - t_2) a_L p^{-\epsilon} (t_2 - M)^2 + p^{-\epsilon} t_3 b (1 + T - t_3) \left\{ \frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) - \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) \right\} - p^{-\epsilon} b (1 + T - t_3) \left\{ \frac{1}{bk} (t_3 e^{bkt_3} - t_2 e^{bkt_2}) - \frac{1}{b^2 k^2} (e^{bkt_3} - e^{bkt_2}) \right\} + T \left\{ \frac{bk_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{bk_2}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) + 2a_L (T - t_3) \right\} - \right]$$

$$bk_1\left\{\frac{1}{r_2}(Te^{r_2T} - t_3e^{r_2t_3}) - \frac{1}{r_2}(e^{r_2T} - e^{r_2t_3})\right\} - bk_2\left\{-\frac{1}{r_2}(Te^{-r_2T} - t_3e^{-r_2t_3}) - \frac{1}{r_2}(e^{-r_2T} - e^{-r_2t_3})\right\} - 2a_L(T - t_3)]$$

$$TIE_R = C_s i_e [(T - N) \int_{t_1}^N D_R(p, q) dt + (1 + T - M) \int_N^M D_R(p, q) (M - t) dt + (1 + T - t_2) \int_M^{t_2} D_R(p, q) (t_2 - t) dt + (1 + T - t_3) \int_{t_2}^{t_3} D_R(p, q) (t_3 - t) dt + \int_{t_3}^T D_R(p, q) (T - t) dt]$$

$$= C_s i_e \left[(T - N) a_R p^{-\epsilon} (N - t_1) + \frac{1}{2} (1 + T - M) a_R p^{-\epsilon} (M - N)^2 + \frac{1}{2} (1 + T - t_2) a_R p^{-\epsilon} (t_2 - M)^2 + p^{-\epsilon} t_3 b (1 + T - t_3) \left\{ \frac{c_1}{bk} (e^{bkt_3} - e^{bkt_2}) + \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkt_2}) \right\} - p^{-\epsilon} b (1 + T - t_3) \left\{ \frac{1}{bk} (t_3 e^{bkt_3} - t_2 e^{bkt_2}) - \frac{1}{b^2 k^2} (e^{bkt_3} - e^{bkt_2}) \right\} + T \left\{ \frac{bk_1}{r_2} (e^{r_2T} - e^{r_2t_3}) - \frac{bk_2}{r_2} (e^{-r_2T} - e^{-r_2t_3}) + 2a_R (T - t_3) \right\} \right] + bk_1 \left\{ \frac{1}{r_2} (Te^{r_2T} - t_3 e^{r_2t_3}) - \frac{1}{r_2} (e^{r_2T} - e^{r_2t_3}) \right\} - bk_2 \left\{ -\frac{1}{r_2} (Te^{-r_2T} - t_3 e^{-r_2t_3}) - \frac{1}{r_2} (e^{-r_2T} - e^{-r_2t_3}) \right\} - 2a_L (T - t_3)]$$

Sub-case 2.2: $0 \leq t_1 \leq t_2 \leq N \leq M \leq t_3 \leq T$

In this case the total amount of interest paid and interest earned by the retailer is calculated.

Hence, total interest paid

$$TIP_L = pi_p \left[\int_M^{t_3} (1 + T - t_3) q_L(t) dt + \int_{t_3}^T q_L(t) dt \right] \\ = pi_p \left[(1 + T - t_3) \left\{ \frac{c_1}{bk} (e^{bkt_3} - e^{bkM}) - \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkM}) - \frac{a_L}{b} (t_3 - M) \right\} + \frac{k_1}{r_2} (e^{r_2T} - e^{r_2t_3}) - \frac{k_2}{r_2} (e^{-r_2T} - e^{-r_2t_3}) + \frac{a_L}{b} (T - t_3) \right]$$

$$TIP_R = pi_p \left[\int_M^{t_3} (1 + T - t_3) q_R(t) dt + \int_{t_3}^T q_R(t) dt \right] \\ = pi_p \left[(1 + T - t_3) \left\{ \frac{c_1}{bk} (e^{bkt_3} - e^{bkM}) + \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkM}) - \frac{a_R}{b} (t_3 - M) \right\} - \frac{k_1}{r_2} (e^{r_2T} - e^{r_2t_3}) - \frac{k_2}{r_2} (e^{-r_2T} - e^{-r_2t_3}) + \frac{a_R}{b} (T - t_3) \right]$$

Total interest earned

$$TIE_L = C_s i_e [(T - t_2) \int_{t_1}^{t_2} D_L(p, q) dt + (1 + T - t_3) \int_{t_2}^{t_3} D_L(p, q) (t_3 - t) dt + \int_{t_3}^T D_L(p, q) (T - t) dt]$$

$$= C_s i_e \left[a_L p^{-\epsilon} (T - t_2) (t_2 - t_1) + (1 + T - t_3) p^{-\epsilon} b t_3 \left\{ \frac{c_1}{k} (e^{bkt_3} - e^{bkt_2}) - \frac{c_2}{k} (e^{-bkt_3} - e^{-bkt_2}) \right\} - p^{-\epsilon} (1 + T - t_3) \left[c_1 b \left\{ \frac{1}{bk} (t_3 e^{bkt_3} - t_2 e^{bkt_2}) - \frac{1}{b^2 k^2} (e^{bkt_3} - e^{bkt_2}) \right\} - c_2 b \left\{ -\frac{1}{bk} (t_3 e^{-bkt_3} - t_2 e^{-bkt_2}) - \frac{1}{b^2 k^2} (e^{-bkt_3} - e^{-bkt_2}) \right\} \right] + T \left\{ -\frac{k_1}{r_2} (e^{r_2T} - e^{r_2t_3}) - \frac{k_2}{r_2} (e^{-r_2T} - e^{-r_2t_3}) + \frac{2a_R}{b} (T - t_3) \right\} - k_1 \left\{ -\frac{1}{r_2} (Te^{r_2T} - t_3 e^{r_2t_3}) + \frac{1}{r_2} (e^{r_2T} - e^{r_2t_3}) \right\} - k_2 \left\{ -\frac{1}{r_2} (Te^{-r_2T} - t_3 e^{-r_2t_3}) - \frac{1}{r_2} (e^{-r_2T} - e^{-r_2t_3}) - \frac{a_R}{b} (T^2 - t_3^2) \right\} \right]$$

$$TIE_R = C_s i_e [(T - t_2) \int_{t_1}^{t_2} D_R(p, q) dt + (1 + T - t_3) \int_{t_2}^{t_3} D_R(p, q) (t_3 - t) dt + \int_{t_3}^T D_R(p, q) (T - t) dt]$$

$$= C_s i_e \left[a_R p^{-\epsilon} (T - t_2) (t_2 - t_1) + (1 + T - t_3) p^{-\epsilon} b t_3 \left\{ \frac{c_1}{k} (e^{bkt_3} - e^{bkt_2}) + \frac{c_2}{k} (e^{-bkt_3} - e^{-bkt_2}) \right\} - p^{-\epsilon} (1 + T - t_3) \left[c_1 b \left\{ \frac{1}{bk} (t_3 e^{bkt_3} - t_2 e^{bkt_2}) - \frac{1}{b^2 k^2} (e^{bkt_3} - e^{bkt_2}) \right\} + c_2 b \left\{ -\frac{1}{bk} (t_3 e^{-bkt_3} - t_2 e^{-bkt_2}) - \frac{1}{b^2 k^2} (e^{-bkt_3} - e^{-bkt_2}) \right\} \right] + T \left\{ \frac{k_1}{r_2} (e^{r_2T} - e^{r_2t_3}) - \frac{k_2}{r_2} (e^{-r_2T} - e^{-r_2t_3}) + \frac{2a_R}{b} (T - t_3) \right\} + k_1 \left\{ -\frac{1}{r_2} (Te^{r_2T} - t_3 e^{r_2t_3}) + \frac{1}{r_2} (e^{r_2T} - e^{r_2t_3}) \right\} - k_2 \left\{ -\frac{1}{r_2} (Te^{-r_2T} - t_3 e^{-r_2t_3}) - \frac{1}{r_2} (e^{-r_2T} - e^{-r_2t_3}) - \frac{a_R}{b} (T^2 - t_3^2) \right\} \right]$$

Sub-case 2.3: $0 \leq t_1 \leq t_2 \leq N \leq t_3 \leq T \leq M$

In this case interest payable by the retailer is zero, i.e., $TIP_L = TIP_R = 0$

But, interest earned by the retailer is given by

$$\begin{aligned}
TIE_L &= C_s i_e \left[(M - t_2) \int_{t_1}^{t_2} D_L(p, q) dt + (M - N) \int_{t_2}^N D_L(p, q) dt \right. \\
&+ (1 + M - t_3) \int_N^{t_3} D_L(p, q) (t_3 - t) dt + (1 + M - T) \int_{t_3}^T D_L(p, q) (T - t) dt \left. \right] \\
&= C_s i_e p^{-\epsilon} \left[a_L (M - t_2) (t_2 - t_1) + b (M - N) \left\{ \frac{c_1}{bk} (e^{bkN} - e^{bkt_2}) - \frac{c_2}{bk} (e^{-bkN} - e^{-bkt_2}) \right\} + (1 + M - t_3) t_3 \left\{ \frac{c_1}{bk} (e^{bkt_3} - e^{bkN}) - \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkN}) \right\} + (1 + M - t_3) \left[c_1 \left\{ \frac{1}{bk} (t_3 e^{bkt_3} - N e^{bkN}) - \frac{1}{b^2 k^2} (e^{bkt_3} - e^{bkN}) \right\} + c_2 \left\{ -\frac{1}{bk} (t_3 e^{-bkt_3} - N e^{-bkN}) - \frac{1}{b^2 k^2} (e^{-bkt_3} - e^{-bkN}) \right\} \right] + (1 + M - T) T \left\{ \frac{bk_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{bk_2}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) + 2a_L (T - t_3) \right\} - (1 + M - T) [bk_1 \left\{ \frac{1}{r_2} (T e^{r_2 T} - t_3 e^{r_2 t_3}) - \frac{1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) \right\} + bk_2 \left\{ -\frac{1}{r_2} (T e^{-r_2 T} - t_3 e^{-r_2 t_3}) - \frac{1}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) \right\} + a_L (T^2 - t_3^2)] \right] \\
TIE_R &= C_s i_e \left[(M - t_2) \int_{t_1}^{t_2} D_R(p, q) dt + (M - N) \int_{t_2}^N D_R(p, q) dt \right. \\
&+ (1 + M - t_3) \int_N^{t_3} D_R(p, q) (t_3 - t) dt + (1 + M - T) \int_{t_3}^T D_R(p, q) (T - t) dt \left. \right] \\
&= C_s i_e p^{-\epsilon} \left[a_R (M - t_2) (t_2 - t_1) + b (M - N) \left\{ \frac{c_1}{bk} (e^{bkN} - e^{bkt_2}) + \frac{c_2}{bk} (e^{-bkN} - e^{-bkt_2}) \right\} + (1 + M - t_3) t_3 \left\{ \frac{c_1}{bk} (e^{bkt_3} - e^{bkN}) + \frac{c_2}{bk} (e^{-bkt_3} - e^{-bkN}) \right\} + (1 + M - t_3) \left[c_1 \left\{ \frac{1}{bk} (t_3 e^{bkt_3} - N e^{bkN}) - \frac{1}{b^2 k^2} (e^{bkt_3} - e^{bkN}) \right\} - c_2 \left\{ -\frac{1}{bk} (t_3 e^{-bkt_3} - N e^{-bkN}) - \frac{1}{b^2 k^2} (e^{-bkt_3} - e^{-bkN}) \right\} \right] + (1 + M - T) T \left\{ -\frac{bk_1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) - \frac{bk_2}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) + 2a_R (T - t_3) \right\} - (1 + M - T) [-bk_1 \left\{ \frac{1}{r_2} (T e^{r_2 T} - t_3 e^{r_2 t_3}) - \frac{1}{r_2} (e^{r_2 T} - e^{r_2 t_3}) \right\} + bk_2 \left\{ -\frac{1}{r_2} (T e^{-r_2 T} - t_3 e^{-r_2 t_3}) - \frac{1}{r_2} (e^{-r_2 T} - e^{-r_2 t_3}) \right\} + a_R (T^2 - t_3^2)] \right]
\end{aligned}$$

In this case also the objective functions are defined as follows:

Max Profit

$$(TP) = \frac{1}{T} [TRE + TIE - TPC - THC - TOC - TIP - TBC - TGC - C_3] \quad (26)$$

$$\text{Min Emission } (TE) = TEO + TEH + TEP \quad (27)$$

On taking α -cut over the total profit and emission per unit time is a crisp interval and is defined by

$$TP = [TP_L, TP_R] \text{ where,} \quad (28)$$

$$TP_L = \frac{1}{T} [TRE_L + TIE_L - TPC_L - THC_L - TOC_L - TIP_L - TBC_L - TGC_L - C_3]$$

$$TP_R = \frac{1}{T} [TRE_R + TIE_R - TPC_R - THC_R - TOC_R - TIP_R - TBC_R - TGC_R - C_3]$$

$$\text{And, } TE = [TE_L, TE_R] \text{ where,} \quad (29)$$

$$TE_L = TEO_L + TEH_L + TEP_L$$

$$TE_R = TEO_R + TEH_R + TEP_R, \text{ when } 0 \leq \alpha \leq 0.5$$

Again, if $\alpha \in [0.5, 1]$, then the system of interval fuzzy differential equation is given by

$$\left[\frac{dq_L}{dt}, \frac{dq_R}{dt} \right] = \begin{cases} -[a_L, a_R] \odot p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) \odot [a_L, a_R] \odot p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) \odot p^{-\epsilon} ([a_L, a_R] + b \odot [q_L, q_R]) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon} ([a_L, a_R] + b \odot [q_L, q_R]) & t_3 \leq t \leq T \end{cases} \quad (30)$$

Therefore two cases arise

Case 1: gH-(i) differentiability

Case 2: gH-(ii) differentiability

In case 1: gH-(i) differentiability, the above system reduces to

$$\frac{dq_L}{dt} = \begin{cases} -a_R p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1) a_L p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1) p^{-\epsilon} (a_L + b q_L) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon} (a_R + b q_R) & t_3 \leq t \leq T \end{cases}$$

$$\frac{dq_R}{dt} = \begin{cases} -a_L p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1)a_R p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1)p^{-\epsilon}(a_R + bq_R) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon}(a_L + bq_L) & t_3 \leq t \leq T \end{cases}$$

After solving using the boundary conditions, $q(0) = q(t_2) = q(T) = 0$, we get

$$q_L(t) = \begin{cases} -a_R p^{-\epsilon} t & 0 \leq t \leq t_1 \\ (\mu - 1)a_L p^{-\epsilon}(t - t_2) & t_1 \leq t \leq t_2 \\ \frac{a_L}{b} [e^{x(t-t_2)} - 1] & t_2 \leq t \leq t_3 \\ -\frac{a_R}{b} [e^{y(T-t)} - 1] & t_3 \leq t \leq T \end{cases} \quad (31)$$

$$q_R(t) = \begin{cases} -a_L p^{-\epsilon} t & 0 \leq t \leq t_1 \\ (\mu - 1)a_R p^{-\epsilon}(t - t_2) & t_1 \leq t \leq t_2 \\ \frac{a_R}{b} [e^{x(t-t_2)} - 1] & t_2 \leq t \leq t_3 \\ -\frac{a_L}{b} [e^{y(T-t)} - 1] & t_3 \leq t \leq T \end{cases} \quad (32)$$

where, $x = b(\mu - 1)p^{-\epsilon}$ and $y = bp^{-\epsilon}$

$$a_L(\alpha) = \frac{(r_1 - \theta_l r_2) + (r_2 - r_1 - \theta_l r_2)\alpha}{(1 - \theta_l) + 2\theta_l \alpha},$$

$$a_R(\alpha) = \frac{(r_4 - \theta_l r_3) + (r_4 - r_3 - 2\theta_l r_3)\alpha}{(1 - \theta_l) + 2\theta_l \alpha}$$

In this case all the relevant costs and total interest earned and payable are same as the previous case 1 when $\alpha \in [0, 0.5]$ but the values of $a_L(\alpha)$ and $a_R(\alpha)$ are different from the previous case 1. In this case

$$a_L(\alpha) = \frac{(r_1 - \theta_l r_2) + (r_2 - r_1 - \theta_l r_2)\alpha}{(1 - \theta_l) + 2\theta_l \alpha},$$

$$a_R(\alpha) = \frac{(r_4 - \theta_l r_3) + (r_4 - r_3 - 2\theta_l r_3)\alpha}{(1 - \theta_l) + 2\theta_l \alpha}$$

Here the objective function for $\alpha \in [0.5, 1]$ have the same expression as previous Case 1 where $\alpha \in [0, 0.5]$ with different values of $a_L(\alpha)$ and $a_R(\alpha)$

In case 2: gH-(ii) differentiability, the above system reduces to

$$\frac{dq_L}{dt} = \begin{cases} -a_L p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1)a_R p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1)p^{-\epsilon}(a_R + bq_R) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon}(a_L + bq_L) & t_3 \leq t \leq T \end{cases}$$

$$\frac{dq_R}{dt} = \begin{cases} -a_R p^{-\epsilon} & 0 \leq t \leq t_1 \\ (\mu - 1)a_L p^{-\epsilon} & t_1 \leq t \leq t_2 \\ (\mu - 1)p^{-\epsilon}(a_L + bq_L) & t_2 \leq t \leq t_3 \\ -p^{-\epsilon}(a_R + bq_R) & t_3 \leq t \leq T \end{cases}$$

After solving using the boundary conditions, $q(0) = q(t_2) = q(T) = 0$ and $q(t_2) = Q_m$, we get

$$q_L(t) = \begin{cases} -a_L p^{-\epsilon} t & 0 \leq t \leq t_1 \\ (\mu - 1)a_R p^{-\epsilon}(t - t_2) & t_1 \leq t \leq t_2 \\ c_1 e^{bkt} + c_2 e^{-bkt} - \frac{a_L}{b} & t_2 \leq t \leq t_3 \\ k_1 e^{r_2 t} + k_2 e^{-r_2 t} + \frac{a_L}{b} & t_3 \leq t \leq T \end{cases} \quad (33)$$

$$q_R(t) = \begin{cases} -a_R p^{-\epsilon} t & 0 \leq t \leq t_1 \\ (\mu - 1)a_L p^{-\epsilon}(t - t_2) & t_1 \leq t \leq t_2 \\ c_1 e^{bkt} - c_2 e^{-bkt} - \frac{a_R}{b} & t_2 \leq t \leq t_3 \\ -k_1 e^{r_2 t} + k_2 e^{-r_2 t} + \frac{a_R}{b} & t_3 \leq t \leq T \end{cases} \quad (34)$$

Where $k = (\mu - 1)p^{-\epsilon}$ $r_2 = -bp^{-\epsilon}$

$$c_1 = \frac{1}{e^{2bkt_3} - e^{2bkt_2}} \left[\left(Q_m + \frac{a_L}{b} \right) e^{bkt_3} - \frac{a_L}{b} e^{bkt_2} \right]$$

$$c_2 = \frac{a_L}{b} e^{bkt_2} - \frac{1}{e^{2bk(t_3-t_2)}} \left[\left(Q_m + \frac{a_L}{b} \right) e^{bkt_3} - \frac{a_L}{b} e^{bkt_2} \right]$$

$$k_1 = \frac{a_R - a_L}{2b} e^{-c_2 T}$$

$$k_2 = -\frac{a_L + a_R}{2b} e^{c_2 T}$$

$$a_L(\alpha) = \frac{(r_1 - \theta_l r_2) + (r_2 - r_1 - \theta_l r_2)\alpha}{(1 - \theta_l) + 2\theta_l \alpha},$$

$$a_R(\alpha) = \frac{(r_4 - \theta_l r_3) + (r_4 - r_3 - 2\theta_l r_3)\alpha}{(1 - \theta_l) + 2\theta_l \alpha}$$

In this case also all the relevant costs and total interest earned and payable are same as the previous case 2 when $\alpha \in [0, 0.5]$ but the values of $a_L(\alpha)$ and $a_R(\alpha)$ are different from the previous case 2. In this case

$$Z a_L(\alpha) = \frac{(r_1 - \theta_l r_2) + (r_2 - r_1 - \theta_l r_2)\alpha}{(1 - \theta_l) + 2\theta_l \alpha},$$

$$a_R(\alpha) = \frac{(r_4 - \theta_l r_3) + (r_4 - r_3 - 2\theta_l r_3)\alpha}{(1 - \theta_l) + 2\theta_l \alpha}$$

Here the objective function for $\alpha \in [0.5, 1]$ have the same expression as previous Case 2 where $\alpha \in [0, 0.5]$ with different values of $a_L(\alpha)$ and $a_R(\alpha)$

6. Solution procedure for solving multi-objective non-linear problem

The foregoing discussion provides a methodology for converting interval valued fuzzy differential equation into system of ordinary differential equation via generalized Hukuhara derivative approach. The α -cut on the profit function and the emission function leads to a system of objective functions which have been solved by multi-objective genetic algorithm. As the developed problem arise so many parameters and handle this problem with classical methods will be very critical. Hence we applied the meta-heuristic multi-objective genetic algorithm method.

Multi-objective genetic algorithm

Genetic algorithm (GA) is a heuristic search algorithm used in computing to find true or approximate solutions in optimization which mimics the process of natural genetics i.e., survival of the fittest. It has five phases i.e., initial population, fitness function, selection, crossover, mutation. Parents are selected according to their fitness values. The better chromosomes have more chances to be selected. In this method, a few good chromosomes are used for creating new offspring in every iteration. Then some bad chromosomes are removed and the new offspring is placed in their places. The rest of population migrates to the next generation without going through the selection process. A multi-objective optimization problem involves a number of objective functions which are to be either minimized or maximized. As in a single-objective optimization problem, the multi-objective optimization problem may contain a number of constraints which have feasible solution (including all optimal solutions) to be satisfy. Since objectives can be either minimized or maximized, the multi-objective optimization problem in its general form can be written

as

Minimize/Maximize $f(x) = \{f_1(x), f_2(x), f_3(x), \dots, f_k(x)\}$

$x = \{x_1, x_2, x_3, \dots, x_z\}$

subject to:

$g(x) \geq 0$

$h(x) = 0$

$x^l \leq x \leq x^u$

Where f is a vector comprising of k objective functions and x is a vector comprising of z solutions. g and h are vectors corresponding to inequality and equality constraints respectively. The lower bound and upper bound of the vector x is x^l and x^u . The solutions of a multi-objective optimization problem are known as pareto optimal solutions.

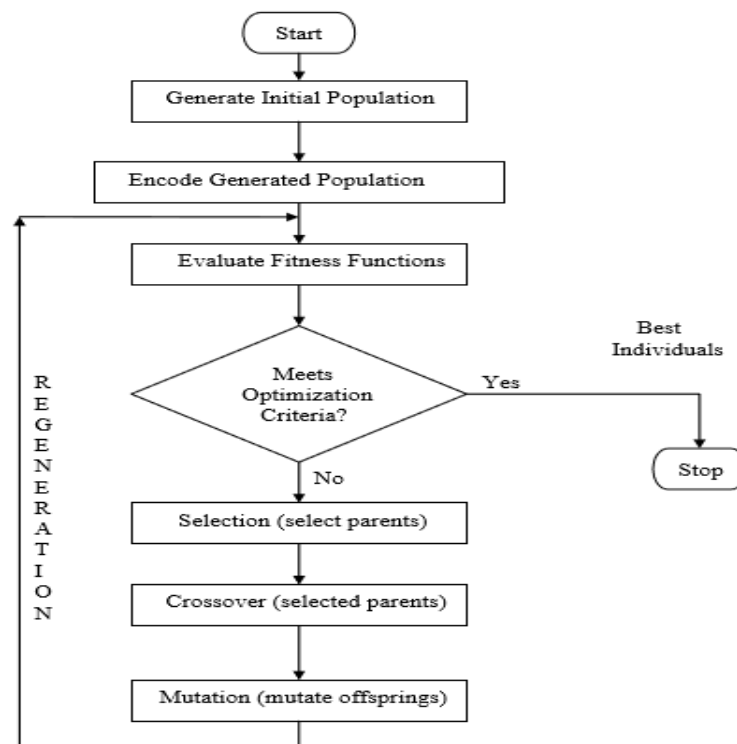


Figure 1. Graphical representation of procedure of GA

7. Real life numerical data and estimation to type-2 fuzzy data

"TATA Motors Limited" a well famous Indian multinational automotive manufacturing company manufactures passengers cars, trucks, vans, buses, sports car, construction equipment etc. The demands of these items from the suppliers are not fixed in every month. A group of managements decisions over the demand of these items are fuzzy in nature and the final decisions by chief production manager over the expert's decision is taken as type-2 fuzzy variable, more precisely trapezoidal type-2 fuzzy variable. Also, the company have to pay carbon emission cost due to the emission creates for obsolescence products, production units and to hold the manufacturing products. We have collected the data for January, 2017 and the corresponding input values in reduced and approximate form are given values.

Let $C_s = 45$ \$/unit, $\mu = 1.8$, $p = \$30$ $C_{hc} = 7$ \$/m³, $b' = 0.5m^3/unit$, $\alpha' = 0.05$, $C'_s = 20$ \$/unit, $C_3 = 25$ \$/unit, $C_{oc} = 5$ \$/ton, $i_p = 0.61$, $i_e = 1.75$, $\alpha = 0.3$ (when $\alpha \in [0, 0.5]$), $\alpha = 0.7$ (when $\alpha \in [0.5, 1]$), $b = 0.8$ $\epsilon = 0.7$ $C_{mc} = 3$ \$/unit, $C_h = 6$ \$/m³, $C_b = 1.02$ \$/m³ $a' = 0.8$ ton/unit, $(r_1, r_2, r_3, r_4, \theta_l, \theta_r) = (10, 12, 14, 16, 0.5, 0.3)$

Table-2: Optimization results for different sub cases for type-2 fuzzy demand by using Multi-objective Genetic algorithm

Different cases		sub cases	T	M	N	$[TP_L, TP_R]$	$[TE_L, TE_R]$
$\alpha \in [0, 0.5]$	Case 1	1.1	1.2567	0.8257	0.6758	[213.91, 368.08]	[106.82, 171.12]
		1.2	1.3459	0.9256	0.7584	[433.67, 597.31]	[213.56, 297.65]
		1.3	1.1346	1.5682	0.8496	[609.68, 779.84]	[305.57, 387.69]
	Case 2	2.1	1.1438	0.8391	0.7143	[232.58, 378.37]	[118.64, 185.97]
		2.2	1.2337	0.9123	0.7852	[445.72, 605.38]	[221.37, 307.82]
		2.3	1.3258	1.7235	1.1324	[618.67, 796.57]	[312.67, 396.78]
$\alpha \in (0.5, 1]$	Case 1	1.1	1.3457	0.8472	0.7523	[198.78, 275.86]	[118.23, 185.79]
		1.2	1.5676	0.8472	0.7523	[412.37, 576.28]	[227.52, 307.64]
		1.3	1.1256	1.4328	0.9726	[582.57, 680.19]	[331.24, 415.28]
	Case 2	2.1	1.058	0.7581	0.7058	[223.81, 353.57]	[127.64, 196.67]
		2.2	1.123	0.7638	0.7253	[432.35, 575.83]	[231.58, 316.97]
		2.3	1.321	1.873	1.0357	[608.67, 721.64]	[327.67, 409.61]

Table-3: Effects of unit selling price C_s on profit function via gH-(i) differentiability of different sub cases of Case 1 for type-2 fuzzy demand when $\alpha \in (0, 0.5]$

C_s (\$)	Sub-case 1.1	Sub-case 1.2	Sub-case 1.3
	$[TP_L, TP_R]$	$[TP_L, TP_R]$	$[TP_L, TP_R]$
45	[213.91,368.08]	[433.67,597.31]	[609.68,779.84]
47	[237.81,381.15]	[457.28,605.35]	[602.13,797.23]
49	[251.26,398.57]	[478.63,621.46]	[641.25,809.15]
51	[273.43,418.93]	[493.54,639.56]	[657.81,822.21]
53	[287.19,429.35]	[506.23,652.21]	[662.23,832.32]
55	[293.25,444.61]	[517.41,667.82]	[671.82,843.67]
57	[302.61,459.82]	[529.35,679.81]	[682.56,857.67]

Table-4: Effects of unit purchasing cost p on profit function via gH-(ii) differentiability of different sub cases of Case 2 for type-2 fuzzy demand when $\alpha \in (0, 0.5]$

p (\$)	Sub-case 2.1	Sub-case 2.2	Sub-case 2.3
	$[TP_L, TP_R]$	$[TP_L, TP_R]$	$[TP_L, TP_R]$
30	[232.58,378.37]	[445.72,605.38]	[618.67,796.57]
32	[227.64,369.87]	[438.79,591.25]	[611.52,787.23]
34	[221.23,362.51]	[432.81,583.71]	[603.15,779.14]
36	[218.21,356.10]	[426.75,575.23]	[596.45,769.23]
38	[211.37,349.58]	[420.12,562.14]	[589.64,761.42]
40	[202.51,341.78]	[413.25,557.69]	[581.23,756.21]
42	[196.25,335.62]	[402.72,551.13]	[571.51,749.17]

Table-5: Effects of unit obsolescence rate α' on emission function via gH-(i) differentiability of different sub cases of Case 1 for type-2 fuzzy demand when $\alpha \in (0.5, 1]$

α'	Sub-case 1.1	Sub-case 1.2	Sub-case 1.3
	$[TE_L, TE_R]$	$[TE_L, TE_R]$	$[TE_L, TE_R]$
0.05	[118.23,185.97]	[227.52,307.64]	[331.24,415.28]
0.10	[123.19,192.25]	[233.67,315.19]	[338.75,423.18]
0.15	[129.20,198.75]	[237.15,322.50]	[343.16,429.11]
0.20	[136.49,207.26]	[242.62,329.23]	[349.21,436.07]
0.25	[142.15,211.27]	[249.13,337.19]	[356.16,442.18]
0.30	[149.35,220.05]	[256.27,342.56]	[361.23,451.95]
0.35	[156.07,227.18]	[261.81,347.25]	[369.09,459.67]

Table-6: Effects of weight of obsolescence product a' on emission function via gH-(ii) differentiability of different sub cases of Case 2 for type-2 fuzzy demand when $\alpha \in (0.5, 1]$

a'	Sub-case 2.1	Sub-case 2.2	Sub-case 2.3
	$[TE_L, TE_R]$	$[TE_L, TE_R]$	$[TE_L, TE_R]$
0.8	[127.64,196.67]	[231.58,316.97]	[327.67,409.61]
1.0	[131.25,201.25]	[237.05,321.84]	[334.25,416.87]
1.2	[136.71,212.08]	[239.17,329.15]	[339.51,421.75]
1.4	[141.81,216.25]	[244.23,336.91]	[343.61,429.82]
1.6	[146.34,218.31]	[249.71,341.28]	[349.71,437.71]
1.8	[153.81,225.83]	[255.82,346.17]	[356.82,442.82]
2.0	[157.35,231.19]	[261.09,351.25]	[360.08,446.72]

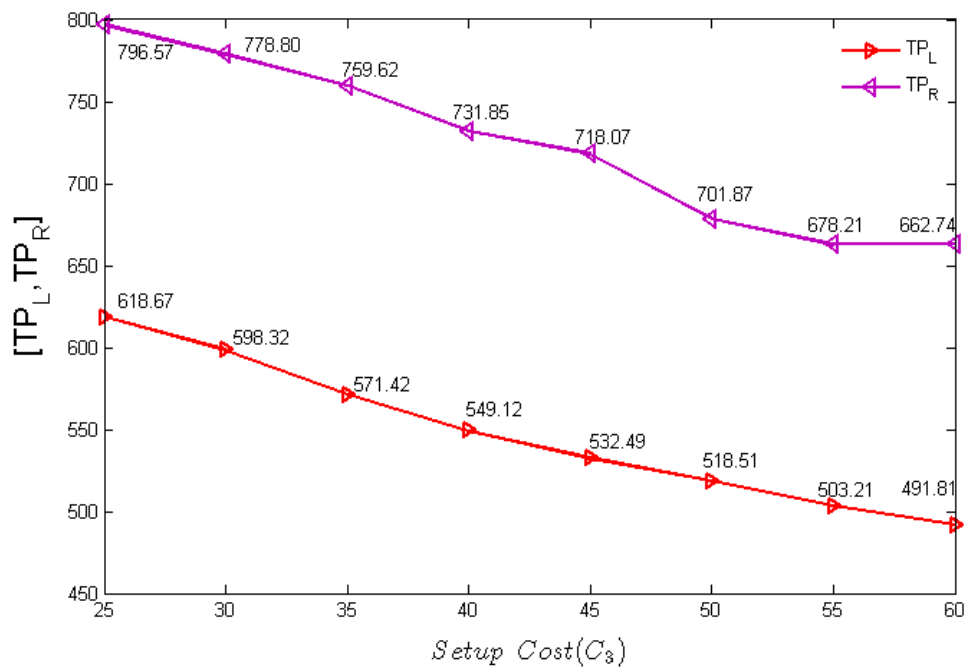


Figure 2. Effect of set-up cost on profit function for sub-case 2.3 when $\alpha \in [0, 0.5]$

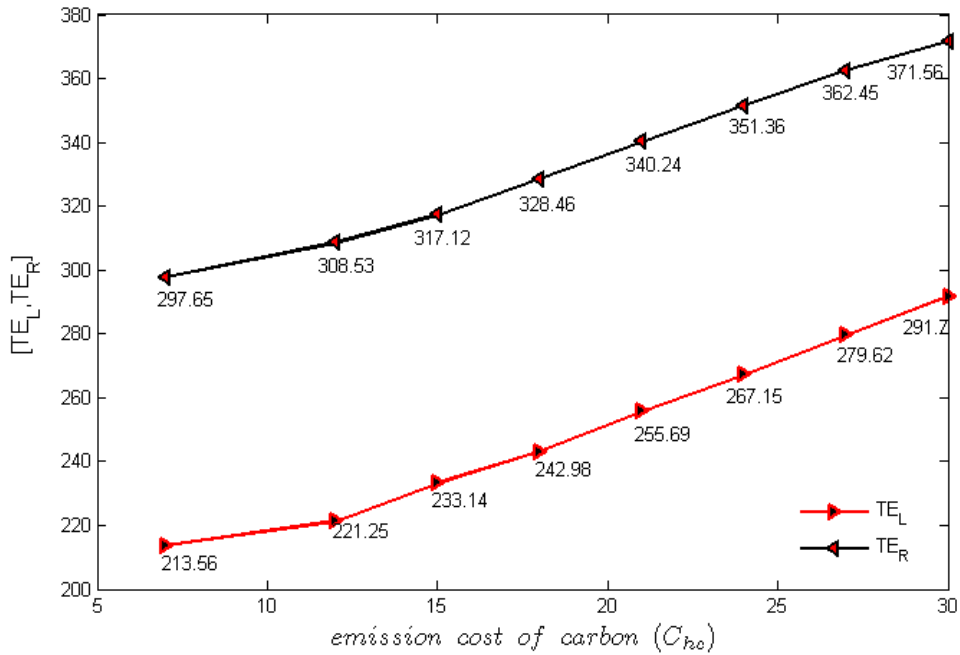


Figure 3. Effect of emission cost of carbon on emission function for sub-case 1.1 when $\alpha \in [0, 0.5]$

8. Discussion

Table 2 describes the optimal results for the profit function and the emission function in interval form and concluded that when the credit period of retailer is greater than total cycle time T via gH-(i) differentiability, the profit is maximum, i.e. $[609.68, 779.84]$ and minimum emission is calculated as $[106.82, 171.12]$ for sub-case 1.1 when $\alpha \in [0, 0.5]$. When the profit and emission is calculated via gH-(ii) differentiability, we also observe that as the credit period of retailer is greater than total cycle time T , the profit is maximum, i.e. $[618.67, 796.57]$ and minimum emission is calculated as $[118.64, 185.97]$ for sub-case 2.1 when $\alpha \in [0, 0.5]$. As one can easily observed from Table 2 that the same scenario is depicted for $\alpha \in [0.5, 1]$. In this case sub-case 2.3 gives the maximum profit, i.e. $[608.67, 721.64]$. We can also conclude that in case of sub-case 1.1, the emission cost is minimum, i.e. $[106.82, 171.12]$ when $\alpha \in [0, 0.5]$. We observe the effect of unit selling price (C_s) on profit function and can conclude that with the increase of unit selling price, the profit function is also increasing as depicted in Table 3. Table 4 analyses the effect of unit purchasing cost (p) on profit function via gH-(ii) differentiability when $\alpha \in [0, 0.5]$ and observe that with the increase of unit purchasing cost the total profit of each sub-case is decrease. Table 5 shows that if the unit obsolescence rate α' is increase for type 2 fuzzy demand over time for $\alpha \in [0.5, 1]$ corresponding cost of emission is also increase. From Table 6 it observed that with the increasing values of weight of obsolescence product a' , total emission cost for each sub-cases are also increased. With the increase of set-up cost, the profit function is decreasing as depicted in Figure 2 when $\alpha \in [0, 0.5]$. We can observe from Figure 3 that the total emission cost is increasing as the emission cost of carbon is increasing.

9. Conclusions and future research work:

The present analyses of the model specifically introduce the concept of type-2 fuzzy variable can be taken as a key factor for a decision maker (DM) engaged with the demand

unit. Demand of an item in market is always fluctuating and in this present model this fluctuation is measured by trapezoidal type-2 fuzzy variable. The present model illustrated a new direction in the field of inventory modeling applying the adventure of Mathematics. The DM is able to take more appropriate precise decisions with the help of present analyses.

In this paper, some useful ideas are presented to deal with inventory control problem with type-2 fuzzy parameters. Along with the main contributions discussed in introduction some more aspects are as follows.

1. CV based reduction method proposed by Qin et al. [23] is discussed and successfully applied to the proposed model to find the total profit function and emission function.
2. According to literature survey for the first time in a single mathematical formulation, we introduced an economic production quantity model with demand depends on the production price and stock in fuzzy environment where demand is taken as trapezoidal type-2 fuzzy number. With the use of CV based reduction method and α -cut of hexagonal fuzzy number the proposed model is solved to find maximum profit and minimum cost of emission of carbon.
3. Some new real life based important facts are provided and discussed in this paper, which will help in developing the business management.

As a future work the presented models can be extended to different types of inventory problems including price discounts, quantity discounts, taking selling price, ordering cost as triangular fuzzy number, intuitionistic number, triangular type-2 fuzzy number, gamma type-2 fuzzy number, Gaussian type-2 fuzzy number etc .

As it is assumed that the unit selling price is greater than the unit purchasing price, the retailer must have sufficient amounts before the end of business period and to pay the dues to the wholesaler some time before the end of the total cycle and in this situation, he will have to pay less interest to the wholesaler. Moreover, the retailer can earn more interest after that time up to the end of the business period. This new approach to calculate the interest earned by the retailer may also apply in this model and the result can be compared with the conventional approach also. The concept of immediate part payment and the delay-payment for the rest can also allowed by the wholesaler for an item over a finite planning horizon or random planning horizon In addition, against an immediate part payment (variable) to the wholesaler, there is a provision for (i) borrowing money from a money lending source and (ii) earning some relaxation on credit period from the wholesaler. The models can also be developed with respect to the retailer for maximum profit. We can also extend the current model for partial trade credit i.e. supplier offers partial trade credit to retailer and retailer offers full trade credit to customers.

References

- [1] S.M. Aljazzar, A. Gurtu and M.Y. Jaber, *Delay in payments- A strategy to reduce carbon emissions from supply chains*, J. Clean. Prod. **170**, 636-644, 2018.
- [2] T. Allahviranloo and M. Afsher Kermani, *Numerical methods for fuzzy linear partial differential equations under new definition for derivative*, Iran. J. Fuzzy Syst. **7** (3), 33-50, 2010.
- [3] D. Battini, A. Persona and F. Sgarbossa, *A Sustainable EOQ model: Theoretical formulation and applications*, Int. J. Prod. Econ. **149**, 145-153, 2014.
- [4] B. Bede, *A note on two-point boundary value problems associated with non-linear fuzzy differential equations*, Fuzzy Sets and Systems **157**, 986-989, 2006.
- [5] B. Bede and S.G. Gal, *Generalizations of the differentiability of fuzzy-number-valued functions with applications to fuzzy differential equations*, Fuzzy Sets and Systems **151**, 581-599, 2005.
- [6] S.C. Chang, *Fuzzy production inventory for fuzzy product quantity with triangular fuzzy number*, Fuzzy Sets and Systems **107**, 37-57, 1999.

- [7] Y. Chen and L. Zhang, *Some new results about arithmetic of type-2 fuzzy variables*, Journal of Uncertain System **5** (3), 227-240, 2011.
- [8] D. Dubois and H. Prade, *Operations on fuzzy numbers*, Int. J. Syst. Sci. **9** (6), 613-626, 1978.
- [9] R. Ezzati, K. Maleknejad and S. Khezertoo, *Convergence, Consistency and stability in fuzzy differential equations*, Iran. J. Fuzzy Syst. **12** (3), 95-112, 2015.
- [10] P. Guchhait, M.K. Maiti and M. Maiti, *A production inventory model with fuzzy production and demand using fuzzy differential equation: An interval compared genetic algorithm approach*, Eng. Appl. Artif. Intell. **26**, 766-778, 2013.
- [11] V. Hovelaque and L. Bironneau, *The carbon-constrained EOQ model with carbon emission dependent demand*, Int. J. Prod. Econ. **164**, 285-291, 2015.
- [12] M. Jonas, M. Marland, W. Winiwarter, T. White, Z. Nahorski, R. Bun and S. Nilsson, *Benefits of dealing with uncertainty in greenhouse gas inventories: introduction*, Climatic Change **103** (1-2), 175-213, 2010b.
- [13] A. Kandel and W.J. Byatt, *Fuzzy differential equations*, In Proceedings of the International Conference on Cybernetics and Society, Tokyo, 1213-1216, 1978.
- [14] N. Kazemi, S.H. Abdul-Rashid, R.A.R. Ghazila, E. Shekarian and S. Zaroni, *Economic order quantity models for items with imperfect quality and emission considerations*, Int. J. Syst. Sci.: Oper. & Logist. **5** (2), 99-115, 2018.
- [15] P. Kundu, S. Kar and M. Maiti, *Fixed charge transportation problem with type-2 fuzzy variables*, Inf. Sci. **255**, 170-184, 2014.
- [16] P. Kundu, S. Kar and M. Maiti, *Multi-item solid transportation problem with type-2 fuzzy parameters*, Appl. Soft Comput. **31**, 61-80, 2015.
- [17] H.M. Lee and J.S. Yao, *Economic production quantity for fuzzy demand quantity and fuzzy production quantity*, European J. Oper. Res. **109**, 203-211, 1998.
- [18] D.C. Lin and J.S. Yao, *Fuzzy economic production for production inventory*, Fuzzy Sets and Systems **111**, 465-495, 2000.
- [19] Z.Q. Liu and Y.K. Liu, *Type-2 fuzzy variables and their arithmetic*, Soft Computing **14**, 729-747, 2010.
- [20] P. Majumder, S.P. Mondal, U.K. Bera and M. Maiti, *Application of Generalized Hukuhara derivative approach in an economic production quantity model with partial trade credit policy under fuzzy environment*, Operations Research Perspective **3**, 77-91, 2016.
- [21] S. Miller, M. Gongora and R. John, *Interval type-2 fuzzy modeling and simulated annealing for real world inventory management*, International Conference on Hybrid Artificial Intelligence Systems, 231-238, Springer-Verlag Bertin, Heidelberg, 2011.
- [22] M. Mizumoto and K. Tanaka, *Fuzzy sets of type-2 under algebraic product and algebraic sum*, Fuzzy Sets and Systems **5** (3), 277-280, 1981.
- [23] R. Qin, Y.K. Liu and Z.Q. Liu, *Methods of critical value reduction for type-2 fuzzy variables and their applications*, J. Comput. Appl. Math. **235**, 1454-1481, 2011.
- [24] P. Rajarajeswari, A.S. Sudha and R. Karthika, *A New Operation on Hexagonal Fuzzy Number*, Int. J. Fuzzy Log. Syst. **3** (3), 15-26, 2013.
- [25] J. Sadeghi, S.T.A. Niaki, M. Malekian and Y. Wang, *A Lagrangian relaxation for a fuzzy random EPQ problem with shortages and redundancy allocation: two tuned meta-heuristics*, Int. J. Fuzzy Syst. **20** (2), 515-533, 2018.
- [26] S. Sharan, S.P. Tiwary and V.K. Yadav, *Interval type-2 fuzzy rough sets and interval type-2 fuzzy closure spaces*, Iran. J. Fuzzy Syst. **12** (3), 113-125, 2015.
- [27] E. Shekarian, C.H. Glock, S.M.P. Amiri and K. Schwindl, *Optimal manufacturing lot size for a single-stage production system with rework in a fuzzy environment*, J. Intell. Fuzzy Syst. **27** (6), 3067-3080, 2014.

- [28] E. Shekarian, M.Y. Jaber, N. Kazemi and E. Ehsani, *A fuzzified version of the economic production quantity (EPQ) model with backorders and rework for a single stage system*, Eur. J. Ind. Eng. **8** (3), 291-324, 2014.
- [29] E. Shekarian, N. Kazemi, S.H. Abdul-Rashid and E.U. Olugu, *Fuzzy inventory models: A comprehensive review*, Appl. Soft Comput. **55**, 588-621, 2017.
- [30] E. Shekarian, E.U. Olugu, S.H. Abdul-Rashid and E. Bottani, *A fuzzy reverse logistics inventory system integrating economic order/production quantity models*, Int. J. Fuzzy Syst. **18** (6), 1141-1161, 2016.
- [31] E. Shekarian, E.U. Olugu, S.H. Abdul-Rashid and N. Kazemi, *An economic order quantity model considering different holding costs for imperfect quality items subject to fuzziness and learning*, J. Intell. Fuzzy Syst. **30** (5), 1985-2997, 2016.
- [32] H.N. Soni, B. Sarkar and M. Joshi, *Demand uncertainty and learning in fuzziness in a continuous review inventory model*, J. Intell. Fuzzy Syst. **33** (4), 2595-2608, 2017.
- [33] L. Stefanini and B. Bede, *Generalized Hukuhara differentiability of interval-valued functions and interval differential equations*, Nonlinear Anal. **71**, 1311-1328, 2009.
- [34] J.R. Stock, S.L. Boyer and T. Harmon, *Research opportunities in supply chain management*, J. Acad. Mark. Sci. **38** (1), 32-41, 2010.
- [35] Z. Takac, *Inclusion and subethood measure for interval-valued fuzzy sets and for continuous type-2 fuzzy sets*, Fuzzy Sets and Systems **224**, 106-120, 2013.
- [36] E.J. Villamizar-Roa, V. Angulo-Castilo and Y. Chaleo-Cano, *Existence of solutions to fuzzy differential equation with generalized Hukuhara derivative via contractive-like mapping principles*, Fuzzy Sets and Systems **265**, 24-38, 2015.
- [37] H.C. Wu, *The central limit theorems for fuzzy random variables*, Inf. Sci. **120**, 239-256, 1999.
- [38] L.A. Zadeh, *The concept of a linguistic variable and its application to approximate reasoning - I*, Inf. Sci. **8**, 199-249, 1975.
- [39] L.A. Zadeh, *The concept of a linguistic variable and its application to approximate reasoning - II*, Inf. Sci. **8**, 301-357, 1975.



Pullback crossed modules in the category of racks

Kadir Emir*, Hatice Gülsün Akay

Department of Mathematics and Computer Science, Eskişehir Osmangazi University, Turkey.

Abstract

In this paper, we define the pullback crossed modules in the category of racks that are mainly based on a pullback diagram of rack morphisms with extra crossed module data on some of its arrows. Furthermore, we prove that the conjugation functor, which is defined between the category of crossed modules of groups and of racks, preserves the pullback crossed modules.

Mathematics Subject Classification (2010). 20N02, 18D05, 18A30, 18A40

Keywords. rack, crossed module, limit

1. Introduction

A rack R is a set equipped with a non-associative binary operation satisfying:

$$(x \triangleleft y) \triangleleft z = (x \triangleleft z) \triangleleft (y \triangleleft z)$$

for all $x, y, z \in R$, and one additional property of this binary operation. Moreover, a rack is called “quandle” if it further satisfies $x \triangleleft x = x$, for all $x \in R$. These total quandle axioms are related to the Reidemeister moves of knot diagrams, and this yields a connection between knot theory and the theory of quandles (hence racks) [9]. Racks have been variously studied under plenty of names and a variety of terminology in literature. They are called automorphic sets [1], crystals [8], left distributive left quasigroups [10] and racks (as a modification of wrack) [4]. The most important example of racks comes from the conjugation in a group G where $g \triangleleft h = h^{-1}gh$, for all $g, h \in G$. This property yields a functor $\text{Conj}: \text{Grp} \rightarrow \text{Rack}$ from the category of groups to the category of racks. Moreover, there exists an adjunction [7] between these two categories with:

$$\text{Hom}_{\text{Grp}}(\text{As}(X), G) \cong \text{Hom}_{\text{Rack}}(X, \text{Conj}(G)),$$

where the functor $\text{As}: \text{Rack} \rightarrow \text{Grp}$ is left adjoint to the functor Conj .

A crossed module of groups [11] $\mathcal{G} = (\partial: E \rightarrow G, \cdot)$ is defined by a group homomorphism $\partial: E \rightarrow G$, together with a (right) group action of G on E satisfying the Peiffer relations, i.e. $\partial(e \cdot g) = g^{-1}\partial(e)g$ and $f \cdot \partial(e) = e^{-1}fe$, for all $e, f \in E$ and $g \in G$. Crossed modules of racks [5] generalize the notion of crossed modules of groups satisfying two parallel Peiffer conditions. An interesting result of this notion is the functors As and Conj preserving the crossed module structures, see [5]. Therefore, we can also consider them as the (induced) functors between the category of crossed modules of groups XGrp and

*Corresponding Author.

Email addresses: kadiremir86@gmail.com (K. Emir), hgulsun@ogu.edu.tr (H. Gülsün Akay)

Received: 20.12.2016; Accepted: 11.10.2017

the category of crossed modules of racks XRack , denoted by As^* and Conj^* respectively. Then, the previous adjunction leads to the following extended adjunction:

$$\text{Hom}_{\text{XGrp}}(\text{As}^*(\mathcal{X}), \mathcal{G}) \cong \text{Hom}_{\text{XRack}}(\mathcal{X}, \text{Conj}^*(\mathcal{G})).$$

Consequently, one can say that the functor Conj^* preserves limits and As^* preserves colimits.

Crossed modules of groups or racks, which have the same fixed codomain A will be called crossed_A modules, and lead to full subcategories of the corresponding categories. We denote these categories by XGrp_A and XRack_A , respectively.

Pullback crossed modules in the category of groups are introduced in [3] which is derived originally from [2]. Explicitly, let $\phi: S \rightarrow R$ be a fixed group homomorphism and $\partial: P \rightarrow R$ be a crossed module. Let A be the pullback in the category of groups with the diagram:

$$\begin{array}{ccc} A & \xrightarrow{\beta} & P \\ \partial^* \downarrow & & \downarrow \partial \\ S & \xrightarrow{\phi} & R \end{array}$$

Then, S acts on $A \subset P \times S$ by the rule $a^s = ((\beta a)^{\phi s}, s^{-1}(\partial a)s)$ for all $s \in S$ and $a \in A$ that makes $\partial^*: A \rightarrow S$ a crossed module and (β, ϕ) a crossed module morphism. This morphism is universal for morphisms from crossed_R modules to crossed_S modules that induce $\phi: S \rightarrow R$. Writing $A = \phi^*P$ we obtain a functor $\phi^*: \text{XGrp}_R \rightarrow \text{XGrp}_S$ which is called restriction that is left adjoint to the induced functor introduced in [3].

In this paper, we construct the pullback crossed modules in the category of racks that will generalize the pullback crossed modules of groups. Furthermore, we see that the functor Conj^* preserves the pullback crossed module structure in the sense of the following commutative diagram:

$$\begin{array}{ccc} \text{XGrp}_R & \xrightarrow{\text{Conj}^*} & \text{XRack}_R \\ \phi^* \downarrow & & \downarrow \phi^* \\ \text{XGrp}_S & \xrightarrow{\text{Conj}^*} & \text{XRack}_S \end{array}$$

for any arbitrary but fixed group homomorphism $\phi: S \rightarrow R$.

2. Preliminaries

We recall some notions from [5, 7] that will be used in the sequel.

2.1. Category of racks

Definition 2.1. A (right) rack R is a set equipped with a (right) binary operation satisfying the following conditions:

- for each $a, b \in R$, there is a unique $c \in R$ such that:

$$c \triangleleft a = b,$$

- for all $a, b, c \in R$, we have:

$$(a \triangleleft b) \triangleleft c = (a \triangleleft c) \triangleleft (b \triangleleft c).$$

A pointed rack is a rack R with an element $1 \in R$ such that (for all $a \in R$):

$$1 \triangleleft a = 1 \quad \text{and} \quad a \triangleleft 1 = a.$$

From now on, all racks will be pointed.

Let R, S be two racks. A rack homomorphism is a map $f: R \rightarrow S$ such that:

$$f(a \triangleleft b) = f(a) \triangleleft f(b) \quad \text{and} \quad f(1) = 1,$$

for all $a, b \in R$. Thus we have the category of racks denoted by Rack . Alternatively, for a point of view on racks where the two right and left rack operations are treated on an equal basis, see [6].

Examples:

1) Given a group G , there exists a rack structure on G where the binary operation is:

$$g \triangleleft h = h^{-1}gh,$$

for all $g, h \in G$. This rack is called the conjugation rack of G , from which we get the functor:

$$\text{Conj}: \text{Grp} \rightarrow \text{Rack}.$$

2) The core rack on a group G is defined by:

$$g \triangleleft h = hg^{-1}h,$$

for all $g, h \in G$; however this construction is not functorial.

3) Let P, R be two racks, we have a rack structure on $P \times R$ defined by:

$$(p, r) \triangleleft (p', r') = (p \triangleleft p', r \triangleleft r'),$$

which is also the product object in the category of racks.

2.2. Rack action

Definition 2.2. Let R be a rack and X be a set. We say that X is an R -set when there are bijections $(\cdot r): X \rightarrow X$ for all $r \in R$ such that:

$$(x \cdot r) \cdot r' = (x \cdot r') \cdot (r \triangleleft r'), \quad (2.1)$$

for all $x \in X$ and $r, r' \in R$.

Definition 2.3. Let R be a rack and X be an R -set. The hemi-semi-direct product $X \rtimes R \subset X \times R$ is the rack defined by:

$$(x, r) \triangleleft (x', r') = (x \cdot r', r \triangleleft r'),$$

for all $x, x' \in X$ and $r, r' \in R$.

Remark that x' disappears in the hemi-semi direct operation which is the main technical difference from the semi-direct product of groups and causes various problems when we deal with it.

Definition 2.4. Let R, S be two racks. We say that S acts on R by automorphisms when there is a (right) action of S on R and:

$$(r \triangleleft r') \cdot s = (r \cdot s) \triangleleft (r' \cdot s) \quad (2.2)$$

for all $s \in S$ and $r, r' \in R$.

2.3. Crossed modules of racks

Definition 2.5. A crossed module of racks (R, S, ∂) is a rack homomorphism $\partial: R \rightarrow S$ together with a (right) rack action of S on R such that following two Peiffer relations hold (for all $r, r' \in R$ and $s \in S$):

- X1) $\partial(r \cdot s) = \partial(r) \triangleleft s$,
- X2) $r \cdot \partial(r') = r \triangleleft r'$.

If (R, S, ∂) and (R', S', ∂') are two crossed module of racks, a crossed module morphism:

$$(f_1, f_0) : (R, S, \partial) \rightarrow (R', S', \partial')$$

is a tuple which consists of rack homomorphisms $f_1 : R \rightarrow R'$, $f_0 : S \rightarrow S'$ such that:

- $\partial' f_1 = f_0 \partial$,
- $f_1(r \cdot s) = f_1(r) \cdot f_0(s)$,

for all $r \in R$, $s \in S$. Thus we get the category of crossed modules of racks, denoted by XRack.

Examples:

1) Let $N \subset R$ be a normal subrack of R (i.e. $n \triangleleft r \in N$ for all $n \in N, r \in R$). The inclusion map $N \rightarrow R$ is a crossed module (inclusion crossed module) where the action is defined by the main rack operation.

2) Let $\mu : M \rightarrow N$ be a crossed module of groups. We obtain a crossed module of racks by passing to the associated conjugation racks of M and N .

3. Fiber product of racks

Definition 3.1. Let $\alpha: P \rightarrow R$ and $\beta: S \rightarrow R$ be two rack homomorphisms. The fiber product $P \times_R S$ is the subrack of the rack $P \times S$ defined by:

$$P \times_R S = \{(p, s) \mid \alpha(p) = \beta(s)\}.$$

From the categorical point of view, the fiber product is the equalizer of the parallel rack homomorphisms:

$$P \times S \begin{array}{c} \xrightarrow{\alpha \circ \pi_1} \\ \xrightarrow{\beta \circ \pi_2} \end{array} R.$$

Proposition 3.2. Let (P, R, α) and (S, R, β) be two crossed modules of racks. The map $\partial: P \times_R S \rightarrow R$ given by:

$$\partial(p, s) = \alpha(p) = \beta(s)$$

yields a crossed module $(P \times_R S, R, \partial)$ with the (right) rack action:

$$\begin{array}{ccc} (P \times_R S) \times R & \rightarrow & P \times_R S \\ ((p, s), r) & \mapsto & (p, s) \cdot r = (p \cdot r, s \cdot r) \end{array}$$

Proof. The action of R is well-defined, i.e. it preserves $P \times_R S$. This follows directly from $\alpha(p) \triangleleft r = \beta(s) \triangleleft r$. Moreover, it satisfies the conditions (2.1) and (2.2) since:

$$\begin{aligned} ((p, s) \cdot r) \cdot r' &= (p \cdot r, s \cdot r) \cdot r' \\ &= ((p \cdot r) \cdot r', (s \cdot r) \cdot r') \\ &= ((p \cdot r') \cdot (r \triangleleft r'), (s \cdot r') \cdot (r \triangleleft r')) \\ &= ((p \cdot r'), (s \cdot r')) \cdot (r \triangleleft r') \\ &= ((p, s) \cdot r') \cdot (r \triangleleft r'), \end{aligned}$$

and

$$\begin{aligned}
((p, s) \triangleleft (p', s')) \cdot r &= (p \triangleleft p', s \triangleleft s') \cdot r \\
&= ((p \triangleleft p') \cdot r, (s \triangleleft s') \cdot r) \\
&= ((p \cdot r) \triangleleft (p' \cdot r), (s \cdot r) \triangleleft (s' \cdot r)) \\
&= ((p \cdot r), (s \cdot r)) \triangleleft ((p' \cdot r), (s' \cdot r)) \\
&= ((p, s) \cdot r) \triangleleft ((p', s') \cdot r),
\end{aligned}$$

for all $(p, s), (p', s') \in P \times_R S$ and $r, r' \in R$.

Also the map $\partial : P \times_R S \rightarrow R$ is a rack homomorphism since:

$$\begin{aligned}
\partial((p, s) \triangleleft (p', s')) &= \partial(p \triangleleft p', s \triangleleft s') \\
&= \alpha(p \triangleleft p') \\
&= \alpha(p) \triangleleft \alpha(p') \\
&= \partial(p, s) \triangleleft \partial(p', s').
\end{aligned}$$

Finally $(P \times_R S, R, \partial)$ is a crossed module of racks since:

X1)

$$\begin{aligned}
\partial((p, s) \cdot r) &= \partial(p \cdot r, s \cdot r) \\
&= \alpha(p \cdot r) \\
&= \alpha(p) \triangleleft r \quad (\because X1 \text{ condition of } \alpha) \\
&= \partial(p, s) \triangleleft r,
\end{aligned}$$

X2)

$$\begin{aligned}
(p, s) \cdot \partial(p', s') &= (p, s) \cdot \alpha(p') \\
&= (p \cdot \alpha(p'), s \cdot \alpha(p')) \\
&= (p \cdot \alpha(p'), s \cdot \beta(s')) \quad (\because \alpha(p') = \beta(s')) \\
&= (p \triangleleft p', s \triangleleft s') \quad (\because X2 \text{ condition of } \alpha, \beta) \\
&= (p, s) \triangleleft (p', s'),
\end{aligned}$$

for all $(p, s), (p', s') \in P \times_R S$ and $r \in R$. □

4. Pullback crossed modules in the category of racks

4.1. Idea

Suppose that we have a crossed module of racks (P, R, ∂) and a rack homomorphism $\phi : S \rightarrow R$. The pullback crossed module of racks:

$$\phi^*(P, R, \partial) = (\phi^*(P), S, \partial^*)$$

is a crossed module of racks satisfying the following universal property:

For a given crossed module morphism of racks:

$$(f, \phi) : (X, S, \mu) \rightarrow (P, R, \partial)$$

there exists a unique crossed module morphism:

$$(f^*, \text{id}_S) : (X, S, \mu) \rightarrow (\phi^*(P), S, \partial^*)$$

which makes the following diagram commutative:

$$\begin{array}{ccc}
 & & (X, S, \mu) \\
 & \swarrow (f^*, \text{id}_S) & \downarrow (f, \phi) \\
 (\phi^*(P), S, \partial^*) & \xrightarrow{(\phi', \phi)} & (P, R, \partial)
 \end{array}$$

Remark 4.1. In other words, the previous definition can be seen as a pullback of rack homomorphisms:

$$\begin{array}{ccc}
 X & \xrightarrow{f} & P \\
 \downarrow \mu & \swarrow f^* & \nearrow \phi' \\
 & \phi^*(P) & \\
 & \swarrow \partial^* & \downarrow \partial \\
 S & \xrightarrow{\phi} & R
 \end{array} \tag{4.1}$$

where the arrows ϕ, ϕ' have crossed module structures. It is clear that pullback crossed modules are not the pullback objects in the category \mathbf{XRack} .

4.2. Construction

Let (P, R, ∂) be a crossed module and let $\phi : S \rightarrow R$ be a rack homomorphism. Define $\phi^*(P) = P \times_R S$ and $\partial^* : \phi^*(P) \rightarrow S$ by $\partial^*(p, s) = s$. Then ∂^* turns into a crossed module where the action of S on $\phi^*(P)$ is defined by:

$$\begin{aligned}
 \phi^*(P) \times S &\rightarrow \phi^*(P) \\
 ((p, s), s') &\mapsto (p, s) \cdot s' = (p \cdot \phi(s'), s \triangleleft s')
 \end{aligned}$$

First of all, the action given above is well-defined, i.e. it preserves the set $\phi^*(P)$, which follows directly from $\partial(p) \triangleleft \phi(s') = \phi(s) \triangleleft \phi(s')$. Moreover, ∂^* is a rack homomorphism since:

$$\begin{aligned}
 \partial^*((p, s) \triangleleft (p', s')) &= \partial^*(p \triangleleft p', s \triangleleft s') \\
 &= s \triangleleft s' \\
 &= \partial^*(p, s) \triangleleft \partial^*(p', s'),
 \end{aligned}$$

for all $(p, s), (p', s') \in \phi^*(P)$. Furthermore the action conditions are satisfied since:

$$\begin{aligned}
 ((p, s) \cdot s') \cdot s'' &= (p \cdot \phi(s'), s \triangleleft s') \cdot s'' \\
 &= ((p \cdot \phi(s')) \cdot \phi(s''), (s \triangleleft s') \triangleleft s'') \\
 &= ((p \cdot \phi(s'')) \cdot (\phi(s') \triangleleft \phi(s'')), (s \triangleleft s'') \triangleleft (s' \triangleleft s'')) \\
 &= ((p \cdot \phi(s'')) \cdot \phi(s' \triangleleft s''), (s \triangleleft s'') \triangleleft (s' \triangleleft s'')) \\
 &= (p \cdot \phi(s''), s \triangleleft s'') \cdot (s' \triangleleft s'') \\
 &= ((p, s) \cdot s'') \cdot (s' \triangleleft s'')
 \end{aligned}$$

and

$$\begin{aligned}
((p, s) \triangleleft (p', s')) \cdot s'' &= (p \triangleleft p', s \triangleleft s') \cdot s'' \\
&= ((p \triangleleft p') \cdot \phi(s''), (s \triangleleft s') \triangleleft s'') \\
&= ((p \cdot \phi(s'') \triangleleft p' \cdot \phi(s'')), (s \triangleleft s'') \triangleleft (s' \triangleleft s'')) \\
&= (p \cdot \phi(s''), (s \triangleleft s'')) \triangleleft (p' \cdot \phi(s''), (s' \triangleleft s'')) \\
&= ((p, s) \cdot s'') \triangleleft ((p', s') \cdot s'')
\end{aligned}$$

for all $(p, s), (p', s') \in \phi^*(P)$ and $s'' \in S$.

Finally ∂^* is a crossed module:

X1)

$$\begin{aligned}
\partial^*((p, s) \cdot s') &= \partial^*(p \cdot \phi(s'), s \triangleleft s') \\
&= s \triangleleft s' \\
&= \partial^*(p, s) \triangleleft s'
\end{aligned}$$

X2)

$$\begin{aligned}
(p, s) \cdot \partial^*(p', s') &= (p, s) \cdot s' \\
&= (p \cdot \phi(s'), s \triangleleft s') \\
&= (p \cdot \partial(p'), s \triangleleft s') \quad (\because \partial(p') = \phi(s')) \\
&= (p \triangleleft p', s \triangleleft s') \quad (\because X2 \text{ condition of } \partial) \\
&= (p, s) \triangleleft (p', s')
\end{aligned}$$

for all $(p, s), (p', s') \in \phi^*(P)$.

Furthermore, this construction satisfies the universal property. To state it, we need the crossed module morphism:

$$(\phi', \phi) : (\phi^*(P), S, \partial^*) \rightarrow (P, R, \partial)$$

where $\phi' : \phi^*(P) \rightarrow P$ is given by $\phi'(p, s) = p$.

Suppose that (X, S, μ) is an arbitrary crossed module with a crossed module morphism:

$$(f, \phi) : (X, S, \mu) \rightarrow (P, R, \partial)$$

We need to prove that there exists a unique crossed module morphism:

$$(f^*, \text{id}_S) : (X, S, \mu) \rightarrow (\phi^*(P), S, \partial^*)$$

such that:

$$(\phi', \phi)(f^*, \text{id}_S) = (f, \phi).$$

Define $f^* : X \rightarrow \phi^*(P)$ by $f^*(x) = (f(x), \mu(x))$, for all $x \in X$. Then the tuple (f^*, id_S) becomes a crossed module morphism, since (for all $s \in S$ and $x \in X$):

$$\begin{aligned}
f^*(x \cdot s) &= (f(x \cdot s), \mu(x \cdot s)) \\
&= (f(x) \cdot \phi(s), \mu(x \cdot s)) \quad (\because (f, \phi) \text{ crossed module morphism}) \\
&= (f(x) \cdot \phi(s), \mu(x) \triangleleft s) \quad (\because X1 \text{ condition of } \mu) \\
&= (f(x), \mu(x)) \cdot s \\
&= f^*(x) \cdot \text{id}_S(s)
\end{aligned}$$

and

$$\begin{aligned}\partial^* f^*(x) &= \partial^*(f(x), \mu(x)) \\ &= \mu(x) \\ &= \text{id}_S \mu(x).\end{aligned}$$

Finally the diagram (4.1) commutes, since (for all $x \in X$):

$$\begin{aligned}\partial^* f^*(x) &= \partial^*(f(x), \mu(x)) \\ &= \mu(x) \\ \phi' f^*(x) &= \phi'(f(x), \mu(x)) \\ &= f(x)\end{aligned}$$

and also $\phi \partial^* = \partial \phi'$ by the definition of $\phi^*(P)$.

Let $(f', \text{id}_S): (X, S, \mu) \rightarrow (\phi^*(P), S, \partial^*)$ be a crossed module morphism of racks with the same property as (f^*, id_S) . Define p and s by $f'(x) = (p, s)$. Then we get:

$$\begin{aligned}\phi' f'(x) &= f(x) \Leftrightarrow \phi'(p, s) = f(x) \Leftrightarrow p = f(x) \\ \partial^* f'(x) &= \mu(x) \Leftrightarrow \partial^*(p, s) = \mu(x) \Leftrightarrow s = \mu(x)\end{aligned}$$

leading to:

$$f'(x) = (p, s) = (f(x), \mu(x)) = f^*(x)$$

which implies that (f^*, id_S) is unique and completes the construction.

Definition 4.2. Let us fix a rack R as a codomain for all crossed modules and construct the related category which is the full subcategory of XRack . These kinds of crossed modules will be called as crossed_R modules and denote the corresponding category by XRack_R .

Corollary 4.3. As a consequence of the pullback crossed module structure in the category of racks, we have the functor:

$$\phi^*: \text{XRack}_R \rightarrow \text{XRack}_S.$$

Example 4.4. Let $\partial: N \rightarrow R$ be an inclusion crossed module and $\phi: S \rightarrow R$ be a rack homomorphism. Then the pullback crossed module is defined by:

$$\begin{aligned}\phi^*(N) &= \{(n, s) \mid \partial(n) = \phi(s), n \in N, s \in S\} \\ &\cong \{s \in S \mid \phi(s) = n, n \in N\} \\ &= \phi^{-1}(N)\end{aligned}$$

with the following commutative diagram:

$$\begin{array}{ccc}\phi^{-1}(N) & \xrightarrow{\phi'} & N \\ \partial^* \downarrow & & \downarrow \partial \\ S & \xrightarrow{\phi} & R\end{array}$$

where the preimage $\phi^{-1}(N)$ is a normal subrack of S .

It follows that:

Example 4.5. If $N = \{1\}$ and R is a rack, then:

$$\phi^*(\{1\}) \cong \{s \in S \mid \phi(s) = 1\} = \ker \phi.$$

Thus $(\ker \phi, S, \partial^*)$ is a pullback crossed module which implies $\ker \phi$ is a normal subrack.

Corollary 4.6. *The kernel of a rack homomorphism is a particular case of a pullback crossed module.*

Example 4.7. If $N = R$ and ϕ is surjective, then:

$$\phi^*(R) = R \times S.$$

5. Functorial approach

Let R be a rack. The associated group $As(R)$ is the quotient of the free group $F(R)$ by the normal subgroup generated by the elements $y^{-1}x^{-1}y(x \triangleleft y)$ for all $x, y \in R$, see [7]. This property leads to the functor:

$$As: Rack \rightarrow Grp,$$

which is left adjoint to the functor $Conj$.

The major property of these functors is; they both preserve the crossed module structure that is proven in [5]. Consequently:

Corollary 5.1. *We have the functors:*

$$As^*: XRack \rightarrow XGrp \quad Conj^*: XGrp \rightarrow XRack,$$

which are induced by As and $Conj$, respectively.

Theorem 5.2. *There exists an adjunction between the categories of crossed modules of racks and of crossed modules of groups:*

$$Hom_{XGrp}(As^*(\mathcal{X}), \mathcal{G}) \cong Hom_{XRack}(\mathcal{X}, Conj^*(\mathcal{G})), \tag{5.1}$$

which is induced by

$$Hom_{Grp}(As(X), G) \cong Hom_{Rack}(X, Conj(G)). \tag{5.2}$$

Proof. Let X be a rack and G be a group. We know from [7] that; for a given rack homomorphism $f: X \rightarrow Conj(G)$, there exists a unique group homomorphism $f_{\#}: As(X) \rightarrow G$ such that the following diagram commutes:

$$\begin{array}{ccc} X & \xrightarrow{\mu} & As(X) \\ \downarrow f & & \downarrow f_{\#} \\ Conj(G) & \xrightarrow{id} & G \end{array}$$

where μ is the natural map. This diagram leads to (5.1).

One level further, let \mathcal{X} be a crossed module of racks and \mathcal{G} be a crossed module of groups. Given a crossed module morphism of racks $(f, g): \mathcal{X} \rightarrow Conj^*(\mathcal{G})$, there exists a unique crossed module morphism of groups $(f_{\#}, g_{\#}): As^*(\mathcal{X}) \rightarrow \mathcal{G}$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathcal{X} & \xrightarrow{(\mu, \mu)} & As^*(\mathcal{X}) \\ \downarrow (f, g) & & \downarrow (f_{\#}, g_{\#}) \\ Conj^*(\mathcal{G}) & \xrightarrow{(id, id)} & \mathcal{G} \end{array}$$

which induces two forms of (4.1) based on rack homomorphisms f, g and proves the adjunction (5.2). □

As another main outcome of the paper, we have the following:

Theorem 5.3. *We have the following commutative diagram:*

$$\begin{array}{ccc}
 \mathbf{XGrp}_R & \xrightarrow{\text{Conj}^*} & \mathbf{XRack}_R \\
 \downarrow \phi^* & & \downarrow \phi^* \\
 \mathbf{XGrp}_S & \xrightarrow{\text{Conj}^*} & \mathbf{XRack}_S
 \end{array}$$

Proof. It follows at once from the known fact that, Conj preserves limits and As preserves colimits since the adjunction (5.2), see also Remark 4.1. \square

Acknowledgment. The authors are thankful to Friedrich Wagemann and Ronald Brown for their kind assistance through invaluable discussions, and also the anonymous referees for their useful comments that helped improve the paper.

References

- [1] E. Brieskorn, *Automorphic sets and braids and singularities*, Braids, AMS-IMS-SIAM Jt. Summer Res. Conf., Santa Cruz/Calif. 1986, Contemp. Math. **78**, 45–115, 1988.
- [2] R. Brown and P. J. Higgins, *On the connection between the second relative homotopy groups of some related spaces*, Proc. Lond. Math. Soc. **36** (3), 192–212, 1978.
- [3] R. Brown and C. D. Wensley, *On finite induced crossed modules and the homotopy 2-type of mapping cones*, Theory Appl. Categ. **1**, 57–71, 1995.
- [4] J. Conway and G. Wraith, *Unpublished correspondence*, 1959.
- [5] A. S. Crans and F. Wagemann, *Crossed modules of racks*, Homology Homotopy Appl. **16** (2), 85–106, 2014.
- [6] J. Faria Martins and R. Picken, *Link invariants from finite categorical groups*, Homology Homotopy Appl. **17** (2), 205–233, 2015.
- [7] R. Fenn and C. Rourke, *Racks and links in codimension two*, J. Knot Theory Ramifications, **1** (4), 343–406, 1992.
- [8] L. Kauffman, *Knots and Physics*, K & E series on knots and everything, World Scientific, 2001.
- [9] S. Nelson, *What is...a quandle?*, Notices of AMS, **63**, 378–380, 2016.
- [10] D. Stanovský, *Left distributive left quasigroup*, Ph.D. thesis, Charles University, Prague, 2004.
- [11] J. H. C. Whitehead, *On adding relations to homotopy groups*, Ann. Math. **42** (2), 409–428, 1941.



A Related Fixed Point Theorem for F -Contractions on Two Metric Spaces

Murat Olgun^{*1}, Özge Biçer², Tuğçe Alyıldız¹, Ishak Altun³

¹Department of Mathematics, Faculty of Science, Ankara University, 06100, Tandogan, Ankara, Turkey

²Istanbul Medipol University, Vocational School, Department of Electronic Communication Technology, Istanbul, Turkey

³Department of Mathematics, Kirikkale University, 71450 Yahsihan, Kirikkale, Turkey

Abstract

Recently, Wardowski in [Fixed points of a new type of contractive mappings in complete metric spaces, Fixed Point Theory Appl. 2012] introduced the concept of F -contraction on complete metric space which is a proper generalization of Banach contraction principle. In the present paper, we proved a related fixed point theorem with F -contraction mappings on two complete metric spaces.

Mathematics Subject Classification (2010). 54H25, 47H10

Keywords. fixed point, F -Contractions, complete metric space

1. Introduction and preliminaries

The Banach contraction mapping principle is one of the pivotal results of analysis. It is widely considered as the source of metric fixed point theory and its significance lies in its vast applicability in a number of branches of mathematics. There are a lot of generalization of Banach contraction mapping principle in the literature. One of a different way of this generalization is to consider two metric space. In 1981, Fisher defined related fixed points of mappings on two metric spaces and obtained some related fixed point theorems. Let (X, d) and (Y, ρ) be two metric space, $T : X \rightarrow Y$ and $S : Y \rightarrow X$ be two mappings. If there exist $x \in X$ and $y \in Y$ such that $Tx = y$ and $Sy = x$, then the pair of (T, S) is said to be has related fixed points. Thereafter many authors obtained some related fixed point theorems (see [1, 3–5, 10]).

In 1994, Namdeo et al. [9] proved the following:

Theorem 1.1. *Let (X, d) and (Y, ρ) be two complete metric spaces, $T : X \rightarrow Y$ and $S : Y \rightarrow X$ mappings satisfying the following equations:*

$$\begin{aligned}d(Sy, STx) &\leq c\phi(x, y) \\ \rho(Tx, TSy) &\leq c\psi(x, y)\end{aligned}$$

^{*}Corresponding Author.

Email addresses: olgun@ankara.edu.tr (M. Olgun), obicer@medipol.edu.tr (Ö. Biçer), tugcekavuzlu@hotmail.com (T. Alyıldız), ishakaltun@yahoo.com (I. Altun)

Received: 19.04.2017; Accepted: 11.10.2017

for all $x \in X$ and $y \in Y$ for which

$$g(x, y) \neq 0 \neq h(x, y)$$

where $0 \leq c < 1$

$$\phi(x, y) = \frac{f(x, y)}{g(x, y)}, \quad \psi(x, y) = \frac{f(x, y)}{h(x, y)} \tag{1.1}$$

and

$$\begin{aligned} f(x, y) &= \max\{d(x, Sy)\rho(y, Tx), d(x, STx)\rho(y, TSy), d(Sy, STx)\rho(y, Tx)\} \\ g(x, y) &= \max\{d(x, STx), \rho(y, TSy), d(x, Sy)\} \\ h(x, y) &= \max\{d(x, STx), \rho(y, TSy), \rho(y, Tx)\}. \end{aligned}$$

Then, ST has a unique fixed point $z \in X$ and TS has a unique fixed point $w \in Y$. Further, $Tz = w$ and $Sw = z$.

In this paper, by taking into account the recent proof technique, which is first used by Wardowski [16], we will present a related fixed point result for two single valued mappings on two complete metric spaces. For the sake of completeness, we consider the following notion due to [16].

Let \mathcal{F} be the set of all functions $F : (0, \infty) \rightarrow \mathbb{R}$ satisfying the following:

(F1) F is strictly increasing, that is for all $\alpha, \beta \in (0, \infty)$ such that $\alpha < \beta$, $F(\alpha) < F(\beta)$;

(F2) For each sequence $\{\alpha_n\}_{n \in \mathbb{N}}$ of positive numbers $\lim_{n \rightarrow \infty} \alpha_n = 0$ if and only if

$$\lim_{n \rightarrow \infty} F(\alpha_n) = -\infty;$$

(F3) There exists $k \in (0, 1)$ such that $\lim_{\alpha \rightarrow 0^+} \alpha^k F(\alpha) = 0$.

Some examples of the functions belonging to \mathcal{F} are $F_1(\alpha) = \ln \alpha$, $F_2(\alpha) = \alpha + \ln \alpha$, $F_3(\alpha) = -\frac{1}{\sqrt{\alpha}}$ and $F_4(\alpha) = \ln(\alpha^2 + \alpha)$.

Definition 1.2 ([16]). Let (X, d) be a metric space and $T : X \rightarrow X$ be a mapping. Then, we say that T is an F -contraction if $F \in \mathcal{F}$ and there exists $\tau > 0$ such that

$$\forall x, y \in X [d(Tx, Ty) > 0 \Rightarrow \tau + F(d(Tx, Ty)) \leq F(d(x, y))]. \tag{1.2}$$

If we take $F(\alpha) = \ln \alpha$ in Definition 1.2, the inequality (1.2) turns to

$$d(Tx, Ty) \leq e^{-\tau} d(x, y), \text{ for all } x, y \in X, Tx \neq Ty. \tag{1.3}$$

It is clear that for $x, y \in X$ such that $Tx = Ty$, the inequality $d(Tx, Ty) \leq e^{-\tau} d(x, y)$ also holds. Thus T is a Banach contraction with contractive constant $L = e^{-\tau}$. Therefore, every Banach contraction is also F -contraction, but the converse may not be true as shown in the Example 2.5 of [16]. If we choose some different functions from \mathcal{F} in (1.2), we can obtain some new as well as existing contractive conditions. In addition, Wardowski showed that every F -contraction T is a contractive mapping, i.e.,

$$d(Tx, Ty) < d(x, y), \text{ for all } x, y \in X, Tx \neq Ty.$$

Thus, every F -contraction is a continuous map. We can find some important properties about F -contractions in [2, 6–8, 11–15, 17]. In the light of these informations, we can see that the following theorem is a proper generalization of Banach Contraction Principle.

Theorem 1.3 ([16]). Let (X, d) be a complete metric space and $T : X \rightarrow X$ be an F -contraction. Then, T has a unique fixed point.

2. Main result

In this section, we present a new kind of related fixed point theorems using the concept of F -contraction.

Theorem 2.1. *Let (X, d) and (Y, ρ) be two complete metric spaces, $T : X \rightarrow Y$ and $S : Y \rightarrow X$ be two mappings. Suppose that there exist $F \in \mathcal{F}$ and $\tau > 0$ such that*

$$d(Sy, STx) > 0 \Rightarrow \tau + F(d(Sy, STx)) \leq F(\phi(x, y)) \quad (2.1)$$

$$\rho(Tx, TSy) > 0 \Rightarrow \tau + F(\rho(Tx, TSy)) \leq F(\psi(x, y)) \quad (2.2)$$

hold for all $x \in X$ and $y \in Y$ for which

$$g(x, y) \neq 0 \neq h(x, y),$$

where ϕ and ψ are as in Theorem 1.1. Then, ST has a unique fixed point $z \in X$ and TS has a unique fixed point $w \in Y$. Further, $Tz = w$ and $Sw = z$.

Proof. Let $x \in X$ be an arbitrary point. Define sequences $\{x_n\} \subset X$ and $\{y_n\} \subset Y$ by

$$(ST)^n x = x_n, \quad T(ST)^{n-1} x = y_n$$

and define $\alpha_n = d(x_n, x_{n+1})$ and $\beta_n = \rho(y_n, y_{n+1})$, $n = 1, 2, 3, \dots$

If there exist $n_0 \in \mathbb{N}$ for which $x_{n_0+1} = x_{n_0}$ or $y_{n_0+1} = y_{n_0}$ then the proof is finished. Indeed, if $x_{n_0+1} = x_{n_0}$, then $(ST)^{n_0+1} x = (ST)^{n_0} x$ and so $(ST)(ST)^{n_0} x = (ST)^{n_0} x$. Therefore, $(ST)^{n_0} x := z$ is a fixed point of ST . Also, if $x_{n_0+1} = x_{n_0}$, then $Tx_{n_0+1} = Tx_{n_0}$ and so $T(ST)^{n_0+1} x = T(ST)^{n_0} x$ or equivalently we have

$$TST(ST)^{n_0} x = T(ST)^{n_0} x.$$

Therefore, $T(ST)^{n_0} x := w$ is a fixed point TS . In this case we have $Tz = w$ and $Sw = z$. Similar result can be obtained when $y_{n_0+1} = y_{n_0}$ for some n_0 .

Now suppose that $x_n \neq x_{n+1}$ and $y_n \neq y_{n+1}$ for every $n \in \mathbb{N}$. Applying inequality (2.1) we get

$$d(x_n, x_{n+1}) = d(Sy_n, STx_n) > 0$$

so we can write

$$F(d(Sy_n, STx_n)) \leq F(\phi(x_n, y_n)) - \tau$$

from which it follows that

$$F(\alpha_n) \leq F(\beta_n) - \tau. \quad (2.3)$$

Applying inequality (2.2) we get

$$\rho(y_n, y_{n+1}) = \rho(Tx_{n-1}, TSy_n) > 0$$

so we can write

$$F(\rho(Tx_{n-1}, TSy_n)) \leq F(\psi(x_{n-1}, y_n)) - \tau$$

from which it follows that

$$F(\beta_n) \leq F(\alpha_{n-1}) - \tau. \quad (2.4)$$

From (2.3) and (2.4) we get

$$\begin{aligned} F(\alpha_n) &\leq F(\beta_n) - \tau \\ &\leq F(\alpha_{n-1}) - 2\tau \\ &\leq \\ &\quad \vdots \\ &\leq F(\alpha_0) - 2n\tau \\ &\leq F(\beta_0) - (2n + 1)\tau \end{aligned} \quad (2.5)$$

for all $n \in \mathbb{N}$. From (2.5) we obtain $\lim_{n \rightarrow \infty} F(\alpha_n) = -\infty$ and with (F2) we get

$$\lim_{n \rightarrow \infty} \alpha_n = 0. \tag{2.6}$$

Similarly, we get $\lim_{n \rightarrow \infty} F(\beta_n) = -\infty$ from (2.4) and with (F2) we find

$$\lim_{n \rightarrow \infty} \beta_n = 0. \tag{2.7}$$

From (F3) there exist $k \in (0, 1)$ such that

$$\lim_{n \rightarrow \infty} \alpha_n^k F(\alpha_n) = 0 \text{ and } \lim_{n \rightarrow \infty} \beta_n^k F(\beta_n) = 0. \tag{2.8}$$

By (2.5) the following holds for all $n \in \mathbb{N}$

$$\begin{aligned} \alpha_n^k F(\alpha_n) &\leq \alpha_n^k [F(\alpha_{n-1}) - 2\tau] \\ &\leq \\ &\vdots \\ &\leq \alpha_n^k [F(\alpha_0) - 2n\tau] \end{aligned}$$

and so

$$\alpha_n^k F(\alpha_n) - \alpha_n^k F(\alpha_0) \leq -2\alpha_n^k n\tau \leq 0. \tag{2.9}$$

Letting $n \rightarrow \infty$ in (2.9), using (2.6) and (2.8) we obtain

$$\lim_{n \rightarrow \infty} \alpha_n^k n = 0 \tag{2.10}$$

Similarly by (2.5) we get

$$\beta_n^k F(\beta_n) - \beta_n^k F(\beta) \leq -\beta_n^k (2n + 1)\tau \leq 0. \tag{2.11}$$

Letting $n \rightarrow \infty$ in (2.11), using (2.7) and (2.8) we obtain

$$\lim_{n \rightarrow \infty} (2n + 1)\beta_n^k = \lim_{n \rightarrow \infty} n\beta_n^k = 0. \tag{2.12}$$

Now let us observe that from (2.10) there exist $n_1 \in \mathbb{N}$ such that $n\alpha_n^k \leq 1$ for all $n \geq n_1$ and from (2.12) there exist $n_2 \in \mathbb{N}$ such that $n\beta_n^k \leq 1$ for all $n \geq n_2$. Let $n_0 = \max\{n_1, n_2\}$, then we have for all $n > n_0$

$$\alpha_n^k \leq \frac{1}{n} \text{ and } \beta_n^k \leq \frac{1}{n}. \tag{2.13}$$

In order to show that $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences consider $m, n \in \mathbb{N}$ such that $m > n > n_0$. From (2.13) and triangular inequality, we write

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{m-1}, x_m) \\ &< \sum_{i=n}^{\infty} \alpha_i \\ &\leq \sum_{i=n}^{\infty} \frac{1}{i^{\frac{1}{k}}} \end{aligned}$$

and

$$\begin{aligned} \rho(y_n, y_m) &\leq \rho(y_n, y_{n+1}) + \rho(y_{n+1}, y_{n+2}) + \dots + \rho(y_{m-1}, y_m) \\ &< \sum_{i=n}^{\infty} \beta_i \\ &\leq \sum_{i=n}^{\infty} \frac{1}{i^{\frac{1}{k}}}. \end{aligned}$$

From the convergence of the serie $\sum_{i=1}^{\infty} \frac{1}{i^{\frac{1}{k}}}$ we receive that $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences with limits $z \in X$ and $w \in Y$ respectively.

Now, suppose $z \neq STz$ and $w \neq TS w$. The following two cases arise:

Case 1. Let $z = Sw$ and $w = Tz$. Then, $w = TS w$ and $z = STz$, which is a contradiction.

Case 2. Let $z \neq Sw$ or $w \neq Tz$. If $z \neq Sw$, then there exists a subsequence $\{x_{n(k)}\}$ of $\{x_n\}$ such that $d(Sw, x_{n(k)}) > 0$ for all $k \in \mathbb{N}$. Therefore, applying inequality (2.1) we have

$$F(d(Sw, STx_{n(k)})) \leq F(\phi(x_{n(k)}, w)) - \tau \quad (2.14)$$

where $\phi(x_{n(k)}, w) = \frac{f(x_{n(k)}, w)}{g(x_{n(k)}, w)}$. Since

$$\lim_{n \rightarrow \infty} g(x_{n(k)}, w) = \lim_{n \rightarrow \infty} \max\{d(x_{n(k)}, x_{n(k)+1}), \rho(w, TS w), d(x_{n(k)}, Sw)\} > 0$$

and

$$\begin{aligned} \lim_{n \rightarrow \infty} f(x_{n(k)}, w) &= \lim_{n \rightarrow \infty} \max \left\{ \begin{array}{l} d(x_{n(k)}, Sw)\rho(w, Tx_{n(k)}), \\ d(x_{n(k)}, STx_{n(k)})\rho(w, TS w), \\ d(Sw, STx_{n(k)})\rho(w, Tx_{n(k)}) \end{array} \right\} \\ &= \lim_{n \rightarrow \infty} \max \left\{ \begin{array}{l} d(x_{n(k)}, Sw)\rho(w, y_{n(k)+1}), \\ d(x_{n(k)}, x_{n(k)+1})\rho(w, TS w), \\ d(Sw, x_{n(k)+1})\rho(w, y_{n(k)}) \end{array} \right\} \\ &= 0 \end{aligned}$$

we get $\lim_{n \rightarrow \infty} \phi(x_{n(k)}, w) = 0$. Therefore, from (2.14) and (F2) we have

$$\lim_{n \rightarrow \infty} d(Sw, STx_{n(k)}) = 0$$

and so $Sw = z$, which is a contradiction. If $w \neq Tz$, then similar contradiction can be occur.

Therefore, either $z = STz$ or $w = TS w$. If $z = STz$, then z is a fixed point of ST and Tz is a fixed point of TS . Similarly, if $w = TS w$, then w is a fixed point of TS and Sw is a fixed point of ST .

To prove uniqueness, suppose that z and z' are two fixed points of ST . Then, since

$$\phi(z', Tz) = \frac{f(z', Tz)}{g(z', Tz)} = \rho(Tz, Tz')$$

and

$$\psi(z', Tz) = \frac{f(z', Tz)}{h(z', Tz)} = d(z, z'),$$

it follows from inequality (2.1) and (2.2) that

$$\begin{aligned} F(d(STz, STz')) &\leq F(\phi(z', Tz)) - \tau \\ &= F(\rho(Tz, Tz')) - \tau \\ &\leq F(\psi(z, Tz')) - 2\tau \\ &= F(d(z, z')) - 2\tau, \end{aligned}$$

which is a contradiction. Therefore, ST (similarly TS) has a unique fixed point in X . \square

We can obtain the following corollaries.

Corollary 2.2. *Theorem 1.1 is immediate from Theorem 2.1.*

Proof. The proof is clear, by taking $F(\alpha) = \ln \alpha$ in Theorem 2.1. \square

Corollary 2.3. *Let (X, d) be a complete metric space and $T : X \rightarrow X$ be a mapping. Suppose that there exist $F \in \mathcal{F}$ and $\tau > 0$ such that*

$$d(Ty, T^2x) > 0 \Rightarrow \tau + F(d(Ty, T^2x)) \leq F(\phi(x, y))$$

holds for all $x, y \in X$ for which $\max\{d(x, T^2x), d(y, T^2y), d(x, Ty)\} > 0$, where

$$\phi(x, y) = \frac{\max\{d(x, Ty)d(y, Tx), d(x, T^2x)d(y, T^2y), d(Ty, T^2x)d(y, Tx)\}}{\max\{d(x, T^2x), d(y, T^2y), d(x, Ty)\}}.$$

Then, T has a unique fixed point.

Proof. Take $X = Y$, $d = \rho$ and $T = S$ in Theorem 2.1. □

Corollary 2.4. Let (X, d) be a complete metric space and $T : X \rightarrow X$ be a mapping. Suppose that there exist $F \in \mathcal{F}$ and $\tau > 0$ such that

$$\begin{aligned} d(y, Tx) > 0 &\Rightarrow \tau + F(d(y, Tx)) \leq F(\phi(x, y)) \\ d(Tx, Ty) > 0 &\Rightarrow \tau + F(d(Tx, Ty)) \leq F(\psi(x, y)) \end{aligned}$$

hold for all $x, y \in X$ for which

$$g(x, y) \neq 0 \neq h(x, y),$$

where

$$\phi(x, y) = \frac{f(x, y)}{g(x, y)}, \quad \psi(x, y) = \frac{f(x, y)}{h(x, y)}$$

and

$$\begin{aligned} f(x, y) &= \max\{d(x, y)d(y, Tx), d(x, Tx)d(y, Ty), d^2(y, Tx)\} \\ g(x, y) &= \max\{d(x, Tx), d(y, Ty), d(x, y)\} \\ h(x, y) &= \max\{d(x, Tx), d(y, Ty), d(y, Tx)\}. \end{aligned}$$

Then, T has a unique fixed point.

Proof. Take $X = Y$, $d = \rho$ and $S = I$ (the identity mapping) in Theorem 2.1. □

Example 2.5. Let $X = Q$ and $Y = I$, where Q is the set of rational numbers and I is the set of irrational numbers. Consider the discrete metric d on X , and a metric defined by

$$\rho(x, y) = \begin{cases} 0 & , \quad x = y \\ 1 + |x - y| & , \quad x \neq y \end{cases}$$

on Y . Then, it is clear that (X, d) and (Y, ρ) are complete metric spaces. Define two mappings $T : X \rightarrow Y$ by $Tx = \sqrt{2}$ and $S : Y \rightarrow X$ by $Sy = 0$. Then, for all $x \in X$ and $y \in Y$, we have

$$d(Sy, STx) = 0 = \rho(Tx, TSy).$$

This shows that the conditions (2.1) and (2.2) are satisfied for all $F \in \mathcal{F}$ and $\tau > 0$. Therefore, by Theorem 2.1, ST has a unique fixed point $z \in X$ and TS has a unique fixed point $w \in Y$. Further, $Tz = w$ and $Sw = z$.

References

- [1] A. Aliouche and B. Fisher, *A related fixed point theorem for two pairs of mappings on two complete metric spaces*, Hacet. J. Math. Stat. **34**, 39-45, 2005.
- [2] I. Altun, G. Durmaz, G. Minak and S. Romaguera, *Multivalued almost F -contractions on complete metric spaces*, Filomat **30** (2), 441-448, 2016.
- [3] B. Fisher, *Fixed points on two metric spaces*, Glas. Mat. Ser. III **16** (36), 333-337, 1981.
- [4] B. Fisher, *Related fixed points on two metric spaces*, Math. Sem. Notes Kobe Univ. **10**, 17-26, 1982.
- [5] B. Fisher, R.K. Jain and H.K. Sahu, *Related fixed point theorems for three metric spaces*, Novi Sad J. Math. **26** (1), 11-17, 1996.

- [6] M. Imdad, R. Gubran, M. Arif and D. Gopal, *An observation on α -type F -contractions and some ordered-theoretic fixed point results*, Math. Sci. **11** (3), 247-255, 2017.
- [7] G. Minak, A. Helvacı and I. Altun, *Ciric type generalized F -contractions on complete metric spaces and fixed point results*, Filomat **28** (6), 1143-1151, 2014.
- [8] G. Minak, M. Olgun and I. Altun, *A new approach to fixed point theorems for multi-valued contractive maps*, Carpathian J. Math. **31** (2), 241-248, 2015.
- [9] R.K. Namdeo, D. Gupta and B. Fisher, *A related fixed point theorem on two metric spaces*, Punjab Univ. J. Math. **27**, 109-112, 1994.
- [10] R.K. Namdeo, S. Jain and B. Fisher, *A related fixed point theorem for two pairs of mappings on two complete metric spaces*, Hacet. J. Math. Stat. **32**, 7-11, 2013.
- [11] M. Olgun, G. Minak and I. Altun, *A new approach to Mizoguchi-Takahashi type fixed point theorems*, J. Nonlinear Convex Anal. **17** (3), 579-587, 2016.
- [12] H. Piri and P. Kumam, *Some fixed point theorems concerning F -contraction in complete metric spaces*, Fixed Point Theory Appl. **2014**, 210, 2014.
- [13] M. Sgrio and C. Vetro, *Multi-valued F -contractions and the solution of certain functional and integral equations*, Filomat **27** (7), 1259-1268, 2013.
- [14] D. Singh, V. Joshi, M. Imdad and P. Kumam, *Fixed point theorems via generalized F -contractions with applications to functional equations occurring in dynamic programming*, J. Fixed Point Theory Appl. **19** (2), 1453-1479, 2017.
- [15] F. Vetro, *F -contractions of Hardy-Rogers type and application to multistage decision processes*, Nonlinear Anal. Model. Control **21** (4), 531-546, 2016.
- [16] D. Wardowski, *Fixed points of a new type of contractive mappings in complete metric spaces*, Fixed Point Theory Appl. **2012**, 94, 2012.
- [17] D. Wardowski and N.V. Dung, *Fixed points of F -weak contractions on complete metric spaces*, Demonstratio Math. **47** (1), 146-155, 2014.



Ulam-Hyers-Stability for nonlinear fractional neutral differential equations

Azmat Ullah Khan Niazi^{*1}, Jiang Wei¹, Mujeeb Ur Rehman², Du Jun¹

¹*School of Mathematical Sciences, Anhui University, Hefei, Anhui 230039, China*

²*Department of Mathematics, School of Natural Sciences (SNS), National University of Sciences and Technology (NUST), H-12 Sector Islamabad, Pakistan*

Abstract

We discuss Ulam-Hyers stability, Ulam-Hyers-Rassias stability and Generalized Ulam-Hyers-Rassias stability for a class of nonlinear fractional functional differential equations with delay involving Caputo fractional derivative by using Picard operator. An example is also given to show the applicability of our results.

Mathematics Subject Classification (2010). 34K37, 34B15

Keywords. fractional functional differential equation, Ulam-Hyers stability, Ulam-Hyers-Rassias stability

1. Introduction

Fractional differential equations is the area of concentration of recent research and there has been significant progress in this area. However, the concept of fractional derivative is as old as differential equations. L' Hospital in 1695 wrote a letter to Leibniz related to his generalization of differentiation and raised the question about fractional derivative. Nowadays the fractional order differential equations has proved to be the most valuable tools in the modeling of many phenomena in various fields of science and engineering. Indeed, we can find many applications in electromagnetic, control, electrochemistry etc. (see[5–8]). For more details on this area, one can see the monograph of Kilbas et al. [14], I. Podulbny [23], Miller and Ross [17], Li et al. [15, 16], Rehman et al. [25], Chen et al. [1–4], Saeed [27] and the references therein.

Over last three decades, the stability theory for functional equations developed and it got popularity so quickly. It started in 1940, when the stability of functional equations were originally raised by Ulam at Wisconsin University. The problem posed by Ulam was the following: “Under what conditions does there exist an additive mapping near an approximately additive mapping”? (for more details see [29]). The first answer to the question of Ulam [9] was given by Hyers in 1941 in the case of Banach spaces. Thereafter, this type of stability is called the Ulam-Hyers stability. In 1978, Rassias [24] provided a remarkable generalization of the Ulam-Hyers stability of mappings by considering variables.

*Corresponding Author.

Email addresses: azmatmath@yahoo.com (A. U. K. Niazi), jiangwei@ahu.edu.cn (J. Wei), mujeeburrehman345@yahoo.com (M. U. Rehman), djwlm@163.com (D. Jun)

Received: 28.02.2017; Accepted: 11.10.2017

Subsequently, a large number of mathematicians took these two types of stabilities, the Ulam-Hyers stability and the Ulam-Hyers-Rassias stability to carry on their researches and the study of this area has grown to be one of the central and most essential subjects in the mathematical analysis area. For more details on the recent advances on the Ulam-Hyers stability and the Ulam-Hyers-Rassias stability of differential equations, one can see the monographs [10, 11] and the research papers [12, 13, 18–21, 26, 28, 30, 32–35]. However, to the best of our knowledge, most of the authors discuss the stability for implicit functions but in our paper we take neutral functions and as far as we know Ulam's type stability results for a class of nonlinear neutral functional differential equations involving Caputo fractional derivatives have not been investigated yet. Wherefore, motivated by the above articles, we have discussed Ulam-Hyers stability for initial value problems of fractional differential equations with delay

$$\begin{cases} {}^c D_0^\gamma x(t) = f(t, x_t, {}^c D_0^\delta x_t), & t \in I, \\ x(t) = \psi(t), & t \in [-\tau, 0] \\ x(0) = x_0, \quad x'(0) = x_1. \end{cases}$$

where ${}^c D_0^\gamma$ and ${}^c D_0^\delta$ are Caputo derivatives with $I = [0, 1]$, $1 < \gamma < 2$, $0 < \delta < 1$, and x_0, x_1 are real constants, $f : [0, 1] \times C_\tau \times C_\tau \rightarrow \mathbb{R}$ and $\psi : [-\tau, 0] \rightarrow \mathbb{R}$ are continuous, we denote C_τ the Banach space of all continuous functions $\phi : [-\tau, 1] \rightarrow \mathbb{R}$, endowed with the maximum norm $\|\phi\| = \max\{|\phi(s)|; -\tau \leq s \leq 1\}$. If $x : [-\tau, 1] \rightarrow \mathbb{R}$, then for any $t \in I$, and $x_t \in C_\tau$ we denote x_t by $x_t(\theta) = x(t + \theta)$, for $\theta \in [-\tau, 0]$, $\tau > 0$.

The paper is arranged as follows: In Section 2 we review some basic definitions and lemmas used throughout this paper. In the third section we establish Ulam-Hyers stability, Ulam-Hyers-Rassias stability, Generalized Ulam-Hyers-Rassias stability for the above initial value problem and in the last section an example is given to show the applicability of our results.

2. Preliminaries

This part includes some basic definitions and results used throughout this paper.

Definition 2.1. [14] The Gamma function is defined as,

$$\Gamma(\gamma) = \int_0^\infty e^{-t} t^{\gamma-1} dt, \quad \gamma > 0.$$

One of the basic property of Gamma function is that it satisfies the following functional equation: $\Gamma(\gamma + 1) = \gamma\Gamma(\gamma)$.

Definition 2.2. [14] The fractional integral for a function f with lower limit t_0 and order γ can be defined as

$$I_t^\gamma f(t) = \frac{1}{\Gamma(\gamma)} \int_{t_0}^t \frac{f(s)}{(t-s)^{1-\gamma}} ds, \quad \gamma > 0, \quad t > t_0.$$

where Γ is the Gamma function, and right hand side is point-wise defined on \mathbb{R}^+ .

Definition 2.3. [14] The left Caputo fractional derivative of order γ is given by

$${}^c D_t^\gamma f(t) = \frac{1}{\Gamma(n-\gamma)} \int_a^t \frac{f^{(n)}(s)}{(t-s)^{\gamma+1-n}} ds.$$

where $n = [\gamma] + 1$ ($[\gamma]$ stands for the bracket function of γ). Here we define one of the important property of Caputo derivative as the composition of the fractional integration operator I_t^γ with the fractional differentiation operator ${}^c D_t^\gamma$. Let $\gamma > 0$, $n = [\gamma] + 1$ and let $f_{n-\gamma}(t) = {}^c D_t^{n-\gamma} f(t)$ then if $f \in C^n[a, b]$ then

$$I_t^\gamma {}^c D_t^\gamma f(t) = f(t) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (t-a)^k.$$

Lemma 2.4. (Gronwall lemma)[22] Let $\mu, \nu \in C([0, 1], \mathbb{R}^+)$. Suppose that μ is increasing. If $x \in C([0, 1], \mathbb{R}^+)$ is a solution to the inequality

$$x(t) \leq \mu(t) + \int_0^t \nu(s)x(s)ds, \quad t \in [0, 1],$$

then

$$x(t) \leq \mu(t) \exp \left(\int_0^t \nu(s)ds \right), \quad t \in [0, 1].$$

Definition 2.5. [22] Let (X, d) be a metric space. An operator $A : X \rightarrow X$ is a Picard operator if there exists $u^* \in X$ such that

- (i) $F_A = \{u^*\}$ where $F_A = \{\mu \in X : A(\mu) = \mu\}$ is the fixed point set of A .
- (ii) The sequence $(A^n(\mu_0))_{n \in \mathbb{N}}$ converges to u^* for all $\mu_0 \in X$.

Definition 2.6. [22] Let (X, d, \leq) be an ordered metric space. An operator $A : X \rightarrow X$ is an increasing Picard operator $F_A = \mu^*$, then for $\mu \in X$, $\mu \leq A(\mu) \Rightarrow \mu \leq \mu^*$ while $\mu \geq A(\mu) \Rightarrow \mu \geq \mu^*$.

3. Stability

In this section, we will discuss Ulam-Hyers stability, Ulam-Hyers-Rassias stability and Generalized Ulam-Hyers-Rassias stability for a class of fractional neutral differential equations. Let ϵ be a positive real number, $T : X \rightarrow X$ is a continuous operator and $f : [0, 1] \times C_\tau \times C_\tau \rightarrow \mathbb{R}$ is a continuous function, we consider the following differential equation

$$\begin{cases} {}^c D_0^\gamma x(t) = f(t, x_t, {}^c D_0^\delta x_t), & t \in [0, 1], \\ x(t) = \psi(t), & t \in [-\tau, 0], \\ x(0) = x_0, \quad x'(0) = x_1. \end{cases} \tag{3.1}$$

For equation (3.1), for some $\epsilon > 0$, $\phi \in C([-\tau, 1], \mathbb{R}^+)$, we focus on the following inequalities:

$$|{}^c D_0^\gamma y(t) - f(t, y_t, {}^c D_0^\delta y_t)| \leq \epsilon, \quad t \in [0, 1]. \tag{3.2}$$

$$|{}^c D_0^\gamma y(t) - f(t, y_t, {}^c D_0^\delta y_t)| \leq \phi(t), \quad t \in [0, 1]. \tag{3.3}$$

$$|{}^c D_0^\gamma y(t) - f(t, y_t, {}^c D_0^\delta y_t)| \leq \epsilon\phi(t), \quad t \in [0, 1]. \tag{3.4}$$

Definition 3.1. [31] Equation(3.1) is Ulam-Hyers stable if there exists a positive real number c_1 such that for each positive ϵ and for every solution $y \in C^1([-\tau, 1], \mathbb{R})$ of (3.2) there exists a solution $x \in C^1([-\tau, 1], \mathbb{R})$ of (3.1) with $|y(t) - x(t)| \leq c_1\epsilon$, $t \in [-\tau, 1]$.

Definition 3.2. [31] Equation(3.1) is Generalized Ulam-Hyers-Rassias stable with respect to ϕ if there exists $c_{1\phi} > 0$ such that for each solution $y \in C^1([-\tau, 1], \mathbb{R})$ to (3.3) there exists a solution $x \in C^1([-\tau, 1], \mathbb{R})$ to (3.1) with $|y(t) - x(t)| \leq c_{1\phi}\phi(t)$, $t \in [-\tau, 1]$.

Definition 3.3. [31] Equation(3.1) is Ulam-Hyers-Rassias stable with respect to ϕ if there exists $c_{1\phi} > 0$ such that for each solution $y \in C^1([-\tau, 1], \mathbb{R})$ to (3.4) there exists a solution $x \in C^1([-\tau, 1], \mathbb{R})$ to (3.1) with $|y(t) - x(t)| \leq c_{1\phi}\epsilon\phi(t)$, $t \in [-\tau, 1]$.

Remark 3.4. A solution of differential equation is stable (asymptotically stable) if it attracts all other solutions with sufficiently close initial values.

On the other hand, in Hyers-Ulam stability, we compare solution of given differential equation with the solution of differential inequality. We say solution of differential equation is stable if it stays close to solution of differential inequality.

Hyers-Ulam stability may not imply the asymptotic stability.

Remark 3.5. [31] A function $y \in C^1([0, 1], \mathbb{R})$ is a solution of the inequality (3.2) if and only if there exist $h \in C^1([0, 1], \mathbb{R})$ such that

- (i) $|h(t)| \leq \epsilon, \quad t \in [0, 1],$
- (ii) ${}^c D_0^\gamma y(t) = f(t, y_t, {}^c D_0^\delta y_t) + h(t), \quad t \in [0, 1].$

Also a function $y \in C^1([0, 1], \mathbb{R})$ is a solution of the inequality (3.3) if and only if there exist $\tilde{h} \in C^1([0, 1], \mathbb{R})$ such that

- (i) $|\tilde{h}(t)| \leq \phi(t), \quad t \in [0, 1],$
- (ii) ${}^c D_0^\gamma y(t) = f(t, y_t, {}^c D_0^\delta y_t) + \tilde{h}(t), \quad t \in [0, 1].$

Similarly for (3.4) there exist a function $g \in C^1([0, 1], \mathbb{R})$ such that

- (i) $|g(t)| \leq \epsilon\phi(t), \quad t \in [0, 1],$
- (ii) ${}^c D_0^\gamma y(t) = f(t, y_t, {}^c D_0^\delta y_t) + g(t), \quad t \in [0, 1].$

Remark 3.6. Let $1 < \gamma < 2$ and $0 < \delta < 1$ if $y \in C^1([0, 1], \mathbb{R})$ is a solution of inequality (3.2) then y is a solution of the following inequality

$$\left| y(t) - y(0) - y'(0)t - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right| \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)}, \quad t \in [0, 1].$$

From Remark(3.5) we have

$${}^c D_0^\gamma y(t) = f(t, y_t, {}^c D_0^\delta y_t) + h(t).$$

Then

$$\begin{aligned} y(t) - y(0) - y'(0)t &= \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \\ &+ \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} h(s) ds. \end{aligned}$$

Therefore

$$\begin{aligned} \left| y(t) - y(0) - y'(0)t - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right| &\leq \frac{1}{\Gamma(\gamma)} \frac{t^\gamma \epsilon}{\gamma} \\ &\leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)}. \end{aligned}$$

If $y \in C^1([0, 1], \mathbb{R})$ is a solution of inequality (3.4) then y is a solution of the following inequality

$$\begin{aligned} \left| y(t) - y(0) - y'(0)t - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right| \\ \leq \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \phi(s) ds, \quad t \in [0, 1]. \end{aligned}$$

And for inequality (3.4)

$$\begin{aligned} \left| y(t) - y(0) - y'(0)t - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right| \\ \leq \frac{\epsilon}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \phi(s) ds, \quad t \in [0, 1]. \end{aligned}$$

In the following theorems we will prove the Ulam-Hyers stability, Generalized Ulam-Hyers-Rassias stability and Ulam-Hyers-Rassias stability for equation(3.1) on the interval $I = [0, 1]$.

Theorem 3.7. Suppose that

- (a) $f \in C(I \times \mathbb{R}^2, \mathbb{R}), |{}^c D_0^\delta x(t)| \leq \frac{1}{\Gamma(2-\delta)} |x(t)|;$

(b) there exists $Q > 0$ such that for every $t \in [0, 1]$, $\mu_i, \nu_i \in \mathbb{R}$, $i = 1, 2$

$$|f(t, \mu_1, \mu_2) - f(t, \nu_1, \nu_2)| \leq Q \sum_{i=1}^2 |\mu_i - \nu_i|,$$

and $\frac{Q}{\Gamma(\gamma)}[\frac{1}{\gamma} + \frac{1}{\Gamma(2-\delta)}] = k < 1$. Then

- (i) Problem (3.1) has a unique solution in $C^1([-\tau, 1], \mathbb{R}) \cap C^1([0, 1], \mathbb{R})$.
- (ii) Equation (3.1) is Ulam-Hyers stable.

Proof. (i) Under the condition (a), (3.1) is equivalent to the integral equation

$$x(t) = \begin{cases} x_0 + x_1 t + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds, & t \in [0, 1] \\ \psi(t), & t \in [-\tau, 0]. \end{cases}$$

Let $X = \{x \in C[-\tau, 1]; {}^c D_0^\delta x \in C^1[-\tau, 1]\}$ with $\|x\| = \max_{t \in I} |x(t)| + \max_{t \in I} |{}^c D_0^\delta x(t)|$, here $C[-\tau, 1]$ and $C^1[-\tau, 1]$ are denoted as continuous and continuously differentiable sets and $T : X \rightarrow X$ be given by

$$Tx(t) = \begin{cases} x_0 + x_1 t + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds, & t \in [0, 1] \\ \psi(t), & t \in [-\tau, 0]. \end{cases}$$

Here we will show that T is a contraction on X .

$|Tx(t) - Ty(t)| = 0$, $x, y \in C([-\tau, 1], \mathbb{R})$, $t \in [-\tau, 0]$. And for $t \in [0, 1]$, by using

$$\begin{aligned} \max_{0 \leq s \leq t} |x_s - y_s| &= \max_{0 \leq s \leq t} |x(s + \theta) - y(s + \theta)| \\ &= \max_{\theta \leq s + \theta \leq t + \theta} |x(s + \theta) - y(s + \theta)| \\ &\leq \max_{-\tau \leq \bar{s} \leq t} |x(\bar{s}) - y(\bar{s})|, \text{ where } s + \theta = \bar{s}, \text{ and } -\tau \leq \theta < 0 \\ &\leq \max_{-\tau \leq \bar{s} \leq 1} |x(\bar{s}) - y(\bar{s})| \\ &= \|x - y\|. \end{aligned}$$

Thus

$$\begin{aligned} &|Tx(t) - Ty(t)| \\ &\leq \frac{1}{\Gamma(\gamma)} \left| \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds - \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right| \\ &\leq \frac{Q}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \left(|x_s - y_s| + |{}^c D_0^\delta x_s - {}^c D_0^\delta y_s| \right) ds \\ &\leq \frac{Q}{\Gamma(\gamma)} \left(\max_{0 \leq s \leq t} |x_s - y_s| + \max_{0 \leq s \leq t} |{}^c D_0^\delta x_s - {}^c D_0^\delta y_s| \right) \int_0^t (t-s)^{\gamma-1} ds \\ &\leq \frac{Q}{\Gamma(\gamma+1)} \left(\max_{-\tau \leq \bar{s} \leq 1} |x(\bar{s}) - y(\bar{s})| + {}^c D_0^\delta \max_{-\tau \leq \bar{s} \leq 1} |x(\bar{s}) - y(\bar{s})| \right) \\ &\leq \frac{Q}{\Gamma(\gamma+1)} \left(\|x - y\| + {}^c D_0^\delta \|x - y\| \right) \\ &\leq \frac{Q}{\Gamma(\gamma+1)} \|x - y\|. \end{aligned}$$

Also

$$\begin{aligned}
& |D^\delta Tx(t) - D^\delta Ty(t)| \\
& \leq \frac{1}{\Gamma(1-\delta)} \left| \int_0^t (t-s)^{-\delta} \left(\frac{\gamma-1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-2} f(s, x_s, {}^c D_0^\delta x_s) ds \right. \right. \\
& \quad \left. \left. - \frac{\gamma-1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-2} f(s, y_s, {}^c D_0^\delta y_s) ds \right) ds \right| \\
& \leq \frac{Q}{\Gamma(1-\delta)\Gamma(\gamma)} \int_0^t (t-s)^{-\delta} ds \|x-y\| \\
& \leq \frac{Q}{\Gamma(2-\delta)\Gamma(\gamma)} \|x-y\|.
\end{aligned}$$

So,

$$\begin{aligned}
\|Tx - Ty\| & \leq \frac{Q}{\Gamma(\gamma+1)} \|x-y\| + \frac{Q}{\Gamma(2-\delta)\Gamma(\gamma)} \|x-y\| \\
& \leq \frac{Q}{\Gamma(\gamma)} \left[\frac{1}{\gamma} + \frac{1}{\Gamma(2-\delta)} \right] \|x-y\|.
\end{aligned}$$

Therefore $\|Tx(t) - Ty(t)\| \leq k\|x-y\|$. Hence by Banach contraction principle T is a contraction.

(ii) Let $y \in C^1([0, 1], \mathbb{R})$ be the solution of (3.2), let us denote by $x \in C^1([0, 1], \mathbb{R})$ the unique solution of equation(3.1) i.e

$$\begin{cases}
{}^c D_0^\gamma x(t) = f(t, x_t, {}^c D_0^\delta x_t), & t \in [0, 1], \\
x(t) = y(t), & t \in [-\tau, 0], \\
x(0) = x_0, & x'(0) = x_1.
\end{cases}$$

then we have

$$\begin{aligned}
x(t) & = x(0) + x'(0)t + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds \\
& = y(0) + y'(0)t + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds.
\end{aligned}$$

We can see that $|y(t) - x(t)| = 0$ for $t \in [-\tau, 0]$. For $t \in [0, 1]$ we have

$$\begin{aligned}
 & |y(t) - x(t)| \\
 & \leq \left| y(t) - y(0) - y'(0)t - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds \right| \\
 & \leq \left| y(t) - y(0) - y'(0)t - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right. \\
 & \quad \left. + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right. \\
 & \quad \left. - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds \right| \\
 & \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} |f(s, y_s, {}^c D_0^\delta y_s) \\
 & \quad - f(s, x_s, {}^c D_0^\delta x_s)| ds.
 \end{aligned}$$

Using (b)

$$\begin{aligned}
 & \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \frac{Q}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} (|y_s - x_s| + |{}^c D_0^\delta y_s - {}^c D_0^\delta x_s|) ds \\
 & \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \frac{Q}{\Gamma(\gamma)} \left(\int_0^t (t-s)^{\gamma-1} |y_s - x_s| ds \right. \\
 & \quad \left. + \int_0^t (t-s)^{\gamma-1} |{}^c D_0^\delta y_s - {}^c D_0^\delta x_s| ds \right).
 \end{aligned} \tag{3.5}$$

According to the last inequality for $\mu \in C([-\tau, 1], \mathbb{R}^+)$. We consider the operator $A : C([-\tau, 1], \mathbb{R}^+) \rightarrow C([-\tau, 1], \mathbb{R}^+)$ defined by

$$A\mu(t) = \begin{cases} 0, & t \in [-\tau, 0], \\ \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \frac{Q}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \mu_s ds \\ \quad + \frac{Q}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} {}^c D_0^\delta \mu_s ds, & t \in [0, 1]. \end{cases}$$

For proving A is a Picard operator, we prove that A is a contraction.

$$\begin{aligned}
 & |A\mu(t) - A\nu(t)| \\
 & \leq \frac{Q}{\Gamma(\gamma)} \left(\int_0^t (t-s)^{\gamma-1} |\mu_s - \nu_s| ds + \int_0^t (t-s)^{\gamma-1} |{}^c D_0^\delta \mu_s - {}^c D_0^\delta \nu_s| ds \right) \\
 & \leq \frac{Qt^\gamma}{\gamma\Gamma(\gamma)} \left(\max_{0 \leq s \leq t} |\mu_s - \nu_s| + \max_{0 \leq s \leq t} |{}^c D_0^\delta \mu_s - {}^c D_0^\delta \nu_s| \right) \\
 & \leq \frac{Q}{\Gamma(\gamma+1)} \|\mu - \nu\|.
 \end{aligned}$$

And

$$\begin{aligned}
& |D^\delta A\mu(t) - D^\delta A\nu(t)| \\
& \leq \frac{1}{\Gamma(1-\delta)} \left[\int_0^t (t-s)^{-\delta} \left(\frac{Q(\gamma-1)}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-2} |\mu_s - \nu_s| ds \right) ds \right. \\
& \quad \left. + \int_0^t (t-s)^{-\delta} \left(\frac{Q(\gamma-1)}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-2} |{}^c D_0^\delta \mu_s - {}^c D_0^\delta \nu_s| ds \right) ds \right] \\
& \leq \frac{Q(\gamma-1)}{\Gamma(1-\delta)(\gamma-1)\Gamma(\gamma)} \|\mu - \nu\| \int_0^t (t-s)^{-\delta} ds \\
& \leq \frac{Q}{\Gamma(2-\delta)\Gamma(\gamma)} \|\mu - \nu\|.
\end{aligned}$$

So

$$\begin{aligned}
\|A(\mu) - A(\nu)\| & \leq \left[\frac{Q}{\Gamma(\gamma+1)} + \frac{Q}{\Gamma(2-\delta)\Gamma(\gamma)} \right] \|\mu - \nu\| \\
& \leq \frac{Q}{\Gamma(\gamma)} \left[\frac{1}{\gamma} + \frac{1}{\Gamma(2-\delta)} \right] \|\mu - \nu\|.
\end{aligned}$$

for all $\mu, \nu \in C([-\tau, 1], \mathbb{R}^+)$. Therefore $\|A(\mu) - A(\nu)\| \leq k\|\mu - \nu\|$ for all $\mu, \nu \in C([-\tau, 1], \mathbb{R}^+)$. Hence A is a contraction with respect to the norm on X . By applying the Banach contraction principle, we can say that A is a Picard operator and $F_A = \{\mu^*\}$, then

$$u^*(t) \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \frac{Q}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} u_s^* ds + \frac{Q}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1c} D_0^\delta u_s^* ds.$$

The solution $u^*(t)$ is increasing and ${}^c D_0^\delta u^* > 0$, also

$$\begin{aligned}
u^*(t) & \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \frac{Q}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} u^*(s) ds \\
& \quad + \frac{Q}{\Gamma(\gamma)\Gamma(2-\delta)} \int_0^t (t-s)^{\gamma-1} u^*(s) ds \\
u^*(t) & \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} + \frac{Q}{\Gamma(\gamma)} \left(1 + \frac{1}{\Gamma(2-\delta)} \right) \int_0^t (t-s)^{\gamma-1} u^*(s) ds.
\end{aligned}$$

therefore by using Gronwall lemma, we can say that

$$\begin{aligned}
u^*(t) & \leq \frac{t^\gamma \epsilon}{\Gamma(\gamma+1)} \exp \frac{Q}{\Gamma(\gamma)} \left(1 + \frac{1}{\Gamma(2-\delta)} \right) \int_0^t (t-s)^{\gamma-1} ds, \quad t \in [0, 1] \\
& \leq \frac{\epsilon}{\Gamma(\gamma+1)} \exp \frac{Q}{\Gamma(\gamma+1)} \left(1 + \frac{1}{\Gamma(2-\delta)} \right) \\
u^*(t) & \leq c_1 \epsilon, \quad \text{where } c_1 = \frac{1}{\Gamma(\gamma+1)} \exp \frac{Q}{\Gamma(\gamma+1)} \left(1 + \frac{1}{\Gamma(2-\delta)} \right).
\end{aligned}$$

So particularly, from (3.5) if $u = |y - x|$, then $u(t) \leq Au(t)$ and by applying the abstract Gronwall lemma we get $u(t) \leq u^*(t)$. Thus it follows

$$|y(t) - x(t)| \leq c_1 \epsilon, \quad t \in [-\tau, 1].$$

Hence equation (3.1) is Ulam-Hyers stable. \square

Theorem 3.8. *If*

(a) $f \in C(I \times \mathbb{R}^2, \mathbb{R})$, $\tilde{h} \in C^1([0, 1], \mathbb{R})$. $|\tilde{h}(t)| \leq \phi(t)$, $\tilde{h} > 0$;

(b) there exists $Q_f \in L^1([0, 1], \mathbb{R}^+)$ such that for all $t \in [0, 1]$, $\mu_i, \nu_i \in \mathbb{R}$, $i = 1, 2$

$$|f(t, \mu_1, \mu_2) - f(t, \nu_1, \nu_2)| \leq Q_f(t) \sum_{i=1}^2 |\mu_i - \nu_i|.$$

(c) the function $\phi \in C([0, 1], \mathbb{R}^+)$ is an increasing function and there exists $\lambda_\phi > 0$ such that

$$\frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \phi(s) ds \leq \lambda_\phi \phi(t); \text{ for all } t \in [0, 1].$$

Then equation (3.1) has a unique solution in $C^1([-\tau, 1], \mathbb{R}) \cap C^1([0, 1], \mathbb{R})$ which is generalized Ulam-Hyers-Rassias stable with respect to ϕ .

Proof. The proof follows the same steps as in Theorem (3.7). Let $y \in C^1([-\tau, 1], \mathbb{R}) \cap C^1([0, 1], \mathbb{R})$ be a solution to (3.3) then by previous theorem $x \in C^1([-\tau, 1], \mathbb{R}) \cap C^1([0, 1], \mathbb{R})$ is a unique solution to the Cauchy problem

$$\begin{cases} {}^c D_0^\gamma x(t) = f(t, x_t, {}^c D_0^\delta x_t), t \in [0, 1], \\ x(t) = y(t), t \in [-\tau, 0], \\ x(0) = x_0, x'(0) = x_1. \end{cases}$$

So

$$x(t) = \begin{cases} x(0) + x'(0)t + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds, & t \in [0, 1] \\ y(t). & t \in [-\tau, 0]. \end{cases}$$

Remarks (3.5) and (3.6) imply

$$\begin{aligned} \left| y(t) - y(0) - y'(0)t - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right| \\ \leq \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \phi(s) ds \\ \leq \lambda_\phi \phi(t), \quad t \in [0, 1]. \end{aligned}$$

From Theorem (3.7) we can see that $|y(t) - x(t)| = 0$, for $t \in [-\tau, 0]$. For $t \in [0, 1]$ we have

$$\begin{aligned} & |y(t) - x(t)| \\ & \leq \left| y(t) - y(0) - y'(0)t - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds \right. \\ & \quad \left. + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, y_s, {}^c D_0^\delta y_s) ds - \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s, x_s, {}^c D_0^\delta x_s) ds \right| \\ & \leq \lambda_\phi \phi(t) + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} |f(s, y_s, {}^c D_0^\delta y_s) - f(s, x_s, {}^c D_0^\delta x_s)| ds \\ & \leq \lambda_\phi \phi(t) + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} Q_f(s) |y_s - x_s| ds \\ & \quad + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} Q_f(s) |{}^c D_0^\delta y_s - {}^c D_0^\delta x_s| ds. \end{aligned}$$

From the proof of Theorem (3.7) it follows that

$$\begin{aligned} |y(t) - x(t)| &\leq \lambda_\phi \phi(t) \exp \frac{Q_f(s)}{\Gamma(\gamma+1)} \left(1 + \frac{1}{\Gamma(2-\delta)}\right) (t)^\gamma, \quad t \in [0, 1] \\ &\leq c_{1\phi} \phi(t), \quad \text{where } c_{1\phi} = \lambda_\phi \exp \frac{Q_f(s)}{\Gamma(\gamma+1)} \left(1 + \frac{1}{\Gamma(2-\delta)}\right). \end{aligned}$$

Hence equation (3.1) is generalized Ulam-Hyers-Rassias stable. \square

Theorem 3.9. *Suppose that*

(a) $f \in C(I \times \mathbb{R}^2, \mathbb{R})$, $g \in C^1([0, 1], \mathbb{R})$, $|g(t)| \leq \epsilon \phi(t)$, $g > 0$;

(b) *there exists* $l_f > 0$ *such that for all* $t \in [0, 1]$, $\mu_i, \nu_i \in \mathbb{R}$, $i = 1, 2$

$$|f(t, \mu_1, \mu_2) - f(t, \nu_1, \nu_2)| \leq l_f \sum_{i=1}^2 |\mu_i - \nu_i|$$

with $\frac{l_f}{\Gamma(\gamma)} \left[\frac{1}{\gamma} + \frac{1}{\Gamma(2-\delta)}\right] < 1$.

(c) *the function* $\phi \in C([0, 1], \mathbb{R}^+)$ *is an increasing function there exists* $\lambda_\phi > 0$ *such that*

$$\frac{\epsilon}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} \phi(s) ds \leq \lambda_\phi \phi(t); \quad \text{for all } t \in [0, 1].$$

Then equation (3.1) has a unique solution in $C^1([-\tau, 1], \mathbb{R}) \cap C^1([0, 1], \mathbb{R})$ *and is Ulam-Hyers-Rassias stable with respect to* ϕ .

Proof. Following the same steps as in Theorems (3.7) and (3.8), we can find both the results, i.e here we will get

$$\begin{aligned} &|y(t) - x(t)| \\ &\leq \epsilon \lambda_\phi \phi(t) + \frac{l_f}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} |y_s - x_s| ds \\ &\quad + \frac{l_f}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} |{}^c D_0^\delta y_s - {}^c D_0^\delta x_s| ds \\ &\leq \epsilon \lambda_\phi \phi(t) \exp \frac{l_f}{\Gamma(\gamma+1)} \left(1 + \frac{1}{\Gamma(2-\delta)}\right) (t)^\gamma, \quad t \in [0, 1] \\ &\leq c_{1\phi} \epsilon \phi(t), \quad \text{where } c_{1\phi} = \lambda_\phi \exp \frac{l_f}{\Gamma(\gamma+1)} \left(1 + \frac{1}{\Gamma(2-\delta)}\right). \end{aligned}$$

So equation (3.1) is Ulam-Hyers-Rassias stable. \square

4. Examples

In this section, we present an example to explain the applicability of main results.

Example 4.1. Consider the initial value problem

$$\begin{cases} D_0^{\frac{3}{2}} x(t) = Ax(t) + Bx(t-0.1) + CD_0^{\frac{1}{2}} x(t-0.1), t \in [0, 1] \\ x(t) = 0.2, t \in [-0.1, 0], \\ x(0) = x'(0) = 0. \end{cases} \quad (4.1)$$

and the inequalities

$$|D_0^{\frac{3}{2}} y(t) - Ay(t) - By(t-0.1) - CD_0^{\frac{1}{2}} y(t-0.1)| \leq \epsilon_1, \quad t \in [0, 1], \quad (4.2)$$

$$|D_0^{\frac{3}{2}} y(t) - Ay(t) - By(t-0.1) - CD_0^{\frac{1}{2}} y(t-0.1)| \leq \phi(t), \quad t \in [0, 1], \quad (4.3)$$

$$|D_0^{\frac{3}{2}}y(t) - Ay(t) - By(t - 0.1) - CD_0^{\frac{1}{2}}y(t - 0.1)| \leq \epsilon_1\phi(t), \quad t \in [0, 1], \quad (4.4)$$

where $A = \frac{1}{9}$, $B = \frac{1}{9}$, $C = \frac{1}{9}$. For proving that equation (4.1) is Ulam-Hyers stable, we take the conditions as in Theorem (3.7) i.e a function $y \in C^1([0, 1], \mathbb{R})$ is a solution of the inequality (3.6) if and only if there exists $h \in C^1([0, 1], \mathbb{R})$ such that

$$\begin{cases} |h(t)| \leq \epsilon_1, & t \in [0, 1], \\ D_0^{\frac{3}{2}}y(t) = Ay(t) + By(t - 0.1) + CD_0^{\frac{1}{2}}y(t - 0.1) + h(t), & t \in [0, 1]. \end{cases} \quad (4.5)$$

Here $\gamma = \frac{3}{2}$, $\delta = \frac{1}{2}$, and $Q = \frac{1}{3}$ also $\frac{Q}{\Gamma(\gamma)}[\frac{1}{\gamma} + \frac{1}{\Gamma(2-\delta)}] \approx 0.6752 < 1$. Furthermore all the assumptions of Theorem (3.7) are satisfied, thus problem (4.1) has a unique solution and is Ulam-Hyers stable with

$$|y(t) - x(t)| \leq c_1\epsilon, \quad t \in [-0.1, 1].$$

where $c_1 \approx 1.2823 > 0$.

Remark 4.2. If we replace equation(4.5) by the inequality

$$\begin{aligned} |\tilde{h}(t)| &\leq \phi(t), \quad t \in [0, 1], \\ D_0^{\frac{3}{2}}y(t) &= Ay(t) + By(t - 0.1) + CD_0^{\frac{1}{2}}y(t - 0.1) + \tilde{h}(t), \quad t \in [0, 1]. \end{aligned}$$

By repeating the same process as in above example one can easily verify the main results of Theorem (3.8). Similarly replace ϵ by $\epsilon\phi(t)$ and $h(t)$ by $g(t)$ we can get the results for Theorem (3.9).

5. Conclusions

We present some new results about stability of a class of fractional neutral differential equations with Caputo fractional derivative by using Picard operator. We discuss the Ulam-Hyers stability, Ulam-Hyers-Rassias stability and Generalized Ulam-Hyers-Rassias stability, which maybe provide a new way for the researchers to discuss such interesting problems in the mathematical analysis area.

The current concepts have significant applications since it means that if we are studying Hyers-Ulam-Rassias stable (or Hyers-Ulam stable) system then one does not have to reach the exact solution. We just need to get a function which satisfies a suitable approximation inequality. In other words, Hyers-Ulam-Rassias stability (or Hyers-Ulam stability) guarantees that there exists a close exact solution. This is altogether useful in many applications where finding the exact solution is quite difficult such as optimization, numerical analysis, biology and economics. It also helps, if the stochastic effects are small, to use deterministic model to approximate a stochastic one.

Acknowledgment. We would like to thank the anonymous referees for their comments and suggestions which have considerably improved the quality and presentation of this paper. This research had been supported by National Natural Science Foundation of China (no.11371027, no.11471015 and no.11601003).

References

- [1] L. Chen, Y. He, Y. Chai and R. Wu, *New results on stability and stabilization of a class of nonlinear fractional-order systems*, *Nonlinear Dyn.* **75** (4), 633–641, 2014.
- [2] L. Chen, C. Liu, R. Wu, Y. He and Y. Chai, *Finite-time stability criteria for a class of fractional-order neural networks with delay*, *Neural Comput. Appl.* **27** (3), 549–556, 2016.
- [3] L. Chen, R. Wu, J. Cao and J.B. Liu, *Stability and synchronization of memristor-based fractional-order delayed neural networks*, *Neural Netw.*, **71**, 37–44, 2015.

- [4] L. Chen, R. Wu, Z. Chu and Y. He, *Stabilization of fractional-order coupled systems with time delay on networks*, *Nonlinear Dyn.* **88** (1), 521–528, 2017.
- [5] K. Diethelm and A.D. Freed, *On the solution of nonlinear fractional-order differential equations used in the modeling of viscoplasticity*, in: *Scientific Computing in Chemical Engineering II*, 217–224, Springer, Berlin, Heidelberg, 1999.
- [6] A.M. El-Sayed, *Fractional-order diffusion-wave equation*, *Internat. J. Theoret. Phys.* **35** (2), 311–322, 1996.
- [7] W.G. Glockle and T.F. Nonnenmacher, *A fractional calculus approach to self-similar protein dynamics*, *Biophys. J.* **68** (1), 46–53, (1995).
- [8] R. Hilfer, ed., *Applications of fractional calculus in physics*, World Scientific, 2000.
- [9] D.H. Hyers, *On the stability of the linear functional equation*, *Proc. Natl. Acad. Sci. USA* **27** (4), 222–224, 1941.
- [10] D.H. Hyers, G. Isac and T.M. Rassias, *Stability of functional equations in several variables*, Birkhuser Boston. Inc., Boston, MA, 1998.
- [11] S.M. Jung, *Hyers-Ulam-Rassias stability of functional equations in mathematical analysis*, Hadronic Press, 2001.
- [12] S.M. Jung, *Hyers-Ulam stability of linear differential equations of first order*, *Appl. Math. Lett.* **17** (10), 1135–1140, 2004.
- [13] S. Jung, *A fixed point approach to the stability of differential equations $y' = F(x, y)$* , *Bull. Malays. Math. Sci. Soc.* **33**, 47–56, 2010.
- [14] A.A. Kilbas, H.M. Srivastava and J.J. Trujillo, *Theory and Applications of Fractional Differential Equations*, North-Holland Mathematics Studies, **204**, 2006.
- [15] X. Li, S. Liu and W. Jiang, *Positive solutions for boundary value problem of nonlinear fractional functional differential equations*, *Appl. Math. Comput.* **217** (22), 9278–9285, 2011.
- [16] S. Liu, X. Li, X.F. Zhou and W. Jiang, *Synchronization analysis of singular dynamical networks with unbounded time-delays*, *Adv. Difference Equ.* **193**, 1–9, 2015.
- [17] K.S. Miller and B. Ross, *An introduction to the fractional calculus and fractional differential equations*, *Mathematics in Science and Engineering* **198**, 1993.
- [18] T. Miura, S. Miyajima and S.E. Takahasi, *A characterization of Hyers-Ulam stability of first order linear differential operators*, *J. Math. Anal. Appl.* **286** (1), 136–146, 2003.
- [19] T. Miura, S. Miyajima and S.E. Takahasi, *Hyers-Ulam stability of linear differential operator with constant coefficients*, *Math. Nachr.* **258** (1), 90–96, 2003.
- [20] M. Obłozza, *Hyers stability of the linear differential equation*, *Rocznik Nauk.-Dydakt. Prace Mat.* **13**, 259–270, 1993.
- [21] M. Obłozza, *Connections between Hyers and Lyapunov stability of the ordinary differential equations*, *Rocznik Nauk.-Dydakt. Prace Mat.* **14**, 141–146, 1997.
- [22] D. Otrocol and V. Ilea, *Ulam stability for a delay differential equation*, *Open Math.* **11** (7), 1296–1303, 2013.
- [23] I. Podlubny, *Fractional Differential Equations*, Academic Press, London, 1999.
- [24] T.M. Rassias, *On the stability of the linear mapping in Banach spaces*, *Proc. Amer. Math. Soc.* **72** (2), 297–300, 1978.
- [25] M.U. Rehman, R.A. Khan and N.A. Asif, *Three point boundary value problems for nonlinear fractional differential equations*, *Acta Math. Sci.* **31** (4), 1337–1346, 2011.
- [26] I.A. Rus, *Ulam stability of ordinary differential equations*, *Stud. Univ. Babe(s)-Bolyai Math.* **54**, 125–133, 2009.
- [27] U. Saeed, *CAS Picard method for fractional nonlinear differential equation*, *Appl. Math. Comput.* **307**, 102–112, 2017.
- [28] S.E. Takahasi, T. Miura and S. Miyajima, *On the Hyers-Ulam stability of the Banach spacevalued differential equation $y' = \lambda y$* , *Bull. Korean Math. Soc.* **39** (2), 309–315, 2002.

- [29] S.M. Ulam, *A Collection of Mathematical Problems*, Interscience Publishers. Inc., New York, 1968.
- [30] J. Wang, M. Fec and Y. Zhou, *Ulam's type stability of impulsive ordinary differential equations*, J. Math. Anal. Appl. **395** (1), 258–264, 2012.
- [31] J. Wang, L. Lv and Y. Zhou, *Ulam stability and data dependence for fractional differential equations with Caputo derivative*, Electron. J. Qual. Theory Differ. Equ. **63**, 1–10, 2011.
- [32] J. Wang, L. Lv and Y. Zhou, *New concepts and results in stability of fractional differential equations*, Commun. Nonlinear Sci. Numer. Simul. **17** (6), 2530–2538, 2012.
- [33] J. Wang and Y. Zhou, *Mittag-Leffler–Ulam stabilities of fractional evolution equations*, Appl. Math. Lett. **25** (4), 723–728, 2012.
- [34] J. Wang, Y. Zhou and M. Fec, *Nonlinear impulsive problems for fractional differential equations and Ulam stability*, Comput. Math. Appl. **64** (10), 3389–3405, 2012.
- [35] W. Wei, X. Li and X. Li, *New stability results for fractional integral equation*, Comput. Math. Appl. **64** (10), 3468–3476, 2012.



Improved oscillation results for second-order half-linear delay differential equations

George E. Chatzarakis¹, Irena Jadlovská*²

¹*Department of Electrical and Electronic Engineering Educators, School of Pedagogical and Technological Education (ASPETE), 14121, N. Heraklio, Athens, Greece*

²*Department of Mathematics and Theoretical Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 04200 Košice, Slovakia*

Abstract

In this paper, we study the second-order half-linear delay differential equation of the form

$$(r(t) (y'(t))^\alpha)' + q(t)y^\alpha(\tau(t)) = 0. \quad (E)$$

We establish new oscillation criteria for (E), which improve a number of related ones in the literature. Our approach essentially involves establishing sharper estimates for the positive solutions of (E) than those presented in known works and a comparison principle with first-order delay differential inequalities. We illustrate the improvement over the known results by applying and comparing our method with the other known methods on the particular example of Euler-type equations.

Mathematics Subject Classification (2010). 34C10, 34K11

Keywords. half-linear differential equation, delay, second-order, oscillation

1. Introduction

Consider the second-order half-linear delay differential equation of the form

$$(r(t) (y'(t))^\alpha)' + q(t)y^\alpha(\tau(t)) = 0, \quad t \geq t_0 > 0. \quad (E)$$

Throughout the paper, it is assumed that the following conditions hold:

- (i) $\alpha > 0$ is a quotient of odd positive integers;
- (ii) $\tau \in \mathcal{C}^1([t_0, \infty))$, $\tau'(t) > 0$, $\tau(t) \leq t$ and $\lim_{t \rightarrow \infty} \tau(t) = \infty$;
- (iii) $q \in \mathcal{C}([t_0, \infty))$ is nonnegative and does not vanish identically on any half line of the form $[t_*, \infty)$;
- (iv) $r \in \mathcal{C}^1([t_0, \infty))$ is positive and satisfies

$$R(t, t_0) := \int_{t_0}^t r^{-1/\alpha}(s) ds \rightarrow \infty \quad \text{as } t \rightarrow \infty.$$

*Corresponding Author.

Email addresses: geaxatz@otenet.gr (G. E. Chatzarakis), irena.jadlovska@tuke.sk (I. Jadlovská)

Received: 24.08.2017; Accepted: 17.10.2017

Under the solution of equation (E) we mean a function $y \in \mathcal{C}([t_a, \infty), \mathbb{R})$ with $t_a = \tau(t_b)$, for some $t_b \geq t_0$, which has the property $r(y')^\alpha \in \mathcal{C}^1([t_a, \infty), \mathbb{R})$ and satisfies (E) on $[t_b, \infty)$. We consider only those solutions of (E) which exist on some half-line $[t_b, \infty)$ and satisfy the condition $\sup\{|x(t)| : t_c \leq t < \infty\} > 0$ for any $t_c \geq t_b$.

As is customary, a solution $y(t)$ of (E) is said to be oscillatory if it is neither eventually positive nor eventually negative. Otherwise, it is said to be nonoscillatory. The equation itself is termed oscillatory if all its solutions oscillate.

The problem of establishing oscillation criteria for differential equations with deviating arguments has been a very active research area over the past decades and several references and reviews of known results can be found in the monographs by Agarwal et al. [1–4], Došlý and Řehák [6] and Győri and Ladas [11].

The oscillation problem for (E) and its particular cases (or its generalizations on dynamic, neutral, nonlinear equations, etc.) has been studied extensively, see, e.g., [5, 7, 8, 12, 15–18, 20–26] and the references therein.

One of the basic techniques in oscillation theory is to acquire criteria by comparing the given differential equation with first-order delay differential equations or inequalities, whose oscillatory behavior is known in advance.

The first results in this direction for second-order delay equations were given by Koplatadze [15] in 1986 and Wei [22] in 1988, who proved that the equation

$$y''(t) + q(t)y(\tau(t)) = 0 \tag{1.1}$$

is oscillatory if

$$\liminf_{t \rightarrow \infty} \int_{\tau(t)}^t q(s)\tau(s)ds > \frac{1}{e}. \tag{1.2}$$

In 2000, Koplatadze, Kvinikadze and Stavroulakis [14, Theorem 1] presented an improved oscillation criterion for (1.1), namely,

$$\liminf_{t \rightarrow \infty} \int_{\tau(t)}^t \left(\tau(s) + \int_{t_0}^{\tau(s)} \xi \tau(\xi)q(\xi)d\xi \right) ds > \frac{1}{e}. \tag{1.3}$$

In 1995, Kusano and Wang [16, Theorem 2] used a variant of the Mahfoud’s comparison principle [19] with the ordinary second-order differential equation

$$((x'(t))^\alpha)' + \frac{q(\tau^{-1}(t))}{\tau'(\tau^{-1}(t))}x^\alpha(t) = 0$$

and proved that (E) is oscillatory if

$$\liminf_{t \rightarrow \infty} R^\alpha(\tau(t), t_0) \int_t^\infty q(s)ds > \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}}. \tag{1.4}$$

Condition (1.4) extends the well-known Hille’s criterion

$$\liminf_{t \rightarrow \infty} t \int_t^\infty q(s)ds > \frac{1}{4} \tag{1.5}$$

for a linear ordinary differential equation

$$y''(t) + q(t)y(t) = 0.$$

The most oscillation results for (E) existing in the literature use the Riccati transformation to reduce the second-order equation to a first-order Riccati inequality.

In 2006, Sun and Meng [20, Theorem 2.1] improved the oscillation result of Džurina and Stavroulakis [8] by employing the Riccati transformation

$$\omega(t) = R^\alpha(\tau(t), t_0) \frac{r(t)(y'(t))^\alpha}{y^\alpha(\tau(t))}, \tag{1.6}$$

which led to the following criterion for (E) to be oscillatory:

$$\int_{t_0}^{\infty} \left(R^{\alpha}(\tau(t), t_0)q(t) - \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1} \frac{\tau'(t)}{R(\tau(t), t_0)r^{1/\alpha}(\tau(t))} \right) dt = \infty. \quad (1.7)$$

Similar conditions to (1.7) have been obtained in a number of papers, see, for instance, [9, 10, 21, 23, 24] and the references cited therein. It is useful to note (see the proof of Theorem 2.5 below), that we can get Hille-type condition from (1.7).

Despite the fact that the above-mentioned oscillation results were proven by different techniques, they all have in common that their strength depends on sharpness of the estimates for nonoscillatory, say positive solutions of (E).

The purpose of this article is to further study the oscillatory behavior of solutions of (E) and to obtain new criteria which improve the known ones mentioned above. Our approach is essentially based on establishing sharper estimates for positive solutions of (E) than those used in the known works [8–10, 14, 16, 20–24], using an iterative technique, and a comparison principle with first-order delay differential inequalities. If, in some iteration step, the comparison result fails to apply, we are able to improve (in delay case only) conditions of (1.7)-type.

The effectiveness of our results is illustrated by means of various examples.

2. Main results

As is customary, we state here that all the functional inequalities considered through the rest of the paper are assumed to hold eventually, that is, they are satisfied for all t large enough.

For a clear and compact presentation of our results, we will adopt the following notation to be used in the whole paper. Let the number ρ be defined by

$$\rho := \liminf_{t \rightarrow \infty} \int_{\tau(t)}^t q(s)R^{\alpha}(\tau(s), t_0)ds,$$

and $\lambda(\eta)$ be the smaller positive root of the transcendental equation

$$\lambda = e^{\eta\lambda}, \quad 0 < \eta \leq 1/e.$$

Also, let us define the sequence of constants ρ_k as follows: set

$$\rho_1 := \rho$$

and, for $\rho_i \in (0, 1/e]$, $i \in \mathbb{N}$, let

$$\rho_{i+1} := \liminf_{t \rightarrow \infty} \int_{\tau(t)}^t q(s)R_i^{\alpha}(\tau(s), t_0)ds,$$

where

$$R_i(t, t_0) = R(t, t_0) + \frac{\lambda(\rho_i)}{\alpha} \int_{t_0}^t R(s, t_0)R^{\alpha}(\tau(s), t_0)q(s)ds.$$

We start by stating a simple, but useful result for the first-order delay differential inequality

$$x'(t) + q(t)x(\tau(t)) \leq 0, \quad t \geq t_0, \quad (2.1)$$

where τ and q are assumed to satisfy (ii) and (iii), respectively.

Lemma 2.1. *Let the number k be defined by*

$$k := \liminf_{t \rightarrow \infty} \int_{\tau(t)}^t q(s)ds.$$

Suppose that $k > 0$ and (2.1) has an eventually positive solution. Then $k \leq 1/e$ and

$$\liminf_{t \rightarrow \infty} \frac{x(\tau(t))}{x(t)} \geq \lambda(k).$$

Proof. The proof is almost the same as that given for the corresponding delay differential equation (see [13, Lemma 1]), hence we omit it. \square

Lemma 2.2. *Suppose that $\rho > 0$ and (E) has an eventually positive solution. Then $\rho \leq 1/e$ and*

$$\liminf_{t \rightarrow \infty} \frac{r(\tau(t)) (y'(\tau(t)))^\alpha}{r(t) (y'(t))^\alpha} \geq \lambda(\rho). \tag{2.2}$$

Proof. Pick $t_1 \in [t_0, \infty)$ so that $y(\tau(t)) > 0$ on $[t_1, \infty)$. Since $y(t)$ is a positive solution of (E), we have

$$(r(t) (y'(t))^\alpha)' = -q(t)y^\alpha(\tau(t)) \leq 0$$

on $[t_1, \infty)$, which means that $r(t) (y'(t))^\alpha$ is eventually nonincreasing and does not change its sign.

We claim that $r(t) (y'(t))^\alpha > 0$ on $[t_1, \infty)$. Indeed, for the sake of contradiction, assume that $r(t) (y'(t))^\alpha < 0$ on $[t_1, \infty)$. Then there exists a $t'_1 \geq t_1$ such that

$$r(t) (y'(t))^\alpha \leq r(t'_1) (y'(t'_1))^\alpha := c < 0 \quad \text{on } [t'_1, \infty).$$

Integrating the above inequality from t'_1 to t and taking (iv) into account, we have

$$y(t) \leq y(t'_1) + c^{1/\alpha} \int_{t'_1}^t r^{-1/\alpha}(s) ds \rightarrow -\infty \quad \text{as } t \rightarrow \infty,$$

which contradicts the fact that $y(t)$ is a positive solution of (E). Thus, we have

$$y(t) > 0, \quad r(t) (y'(t))^\alpha > 0, \quad (r(t) (y'(t))^\alpha)' \leq 0 \quad \text{on } [t_1, \infty).$$

Since $r^{1/\alpha}(t)y'(t)$ is nonincreasing, there exists a finite limit

$$\lim_{t \rightarrow \infty} r^{1/\alpha}(t)y'(t) = \ell \geq 0.$$

If we assume $\ell > 0$, then $r^{1/\alpha}(t)y'(t) \geq \ell > 0$ and $y(t) \geq \ell R(t, t_1) > 0$ on $[t_1, \infty)$. Noting that $\rho > 0$, we have that

$$\int_{t_0}^\infty q(s)R^\alpha(\tau(s), t_0) ds = \infty.$$

Integrating (E) from t_1 to t yields

$$r(t_1) (y'(t_1))^\alpha \geq \ell^\alpha \int_{t_1}^t q(s)R^\alpha(\tau(s), t_1) ds \rightarrow \infty \quad \text{as } t \rightarrow \infty.$$

This contradiction implies that

$$\lim_{t \rightarrow \infty} r^{1/\alpha}(t)y'(t) = 0. \tag{2.3}$$

On the other hand, it is obvious that

$$r^{1/\alpha}(s)y'(s) \geq r^{1/\alpha}(t)y'(t) \quad \text{for every } s \in [t_1, t].$$

Therefore,

$$\begin{aligned} y(t) &= y(t_1) + \int_{t_1}^t r^{-1/\alpha}(s)r^{1/\alpha}(s)y'(s) ds \\ &\geq y(t_1) + r^{1/\alpha}(t)y'(t)R(t, t_1) \\ &= y(t_1) - r^{1/\alpha}(t)y'(t)R(t_1, t_0) + r^{1/\alpha}(t)y'(t)R(t, t_0). \end{aligned} \tag{2.4}$$

Combining (2.3) and (2.4), we have

$$y(t) > r^{1/\alpha}(t)y'(t)R(t, t_0) \quad \text{on } [t_2, \infty), \tag{2.5}$$

for some $t_2 \in [t_1, \infty)$ large enough. Using (2.5) in (E), it is easy to see that $x(t) := r(t) (y'(t))^\alpha$ is a positive solution of the first-order delay differential inequality

$$x'(t) + q(t)R^\alpha(\tau(t), t_0)x(\tau(t)) < 0. \tag{2.6}$$

To complete the proof, it suffices to apply Lemma 2.1 to (2.6). □

Application of Lemma 2.2 allows us to obtain various important oscillation results. Theorem 2.3 below is a simple generalization of (1.2) for a half-linear differential equation, while Theorems 2.4 and 2.5 essentially improve the known criteria (1.7) and (1.4), respectively.

Theorem 2.3. *If $\rho > 1/e$, then (E) is oscillatory.*

Theorem 2.4. *Assume that $0 < \rho \leq 1/e$. If*

$$\limsup_{t \rightarrow \infty} \int_{t_0}^t \left(R^\alpha(\tau(s), t_0)q(s) - \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1} \frac{1}{(\lambda(\rho) - \epsilon)} \frac{\tau'(s)}{R(\tau(s), t_0)r^{1/\alpha}(\tau(s))} \right) ds = \infty \tag{2.7}$$

for some $\epsilon > 0$, then (E) is oscillatory.

Proof. Suppose to the contrary that (E) has a nonoscillatory solution $y(t)$ on $[t_0, \infty)$. Without loss of generality, we can assume that there exists a $t_1 \geq t_0$ such that $y(t) > 0$ and $y(\tau(t)) > 0$ on $[t_1, \infty)$. Define $\omega(t)$ as in (1.6), i.e.,

$$\omega(t) = R^\alpha(\tau(t), t_0) \frac{r(t) (y'(t))^\alpha}{y^\alpha(\tau(t))}. \tag{2.8}$$

We see that $\omega > 0$ for $t \geq t_1$. Differentiating (2.8) and using (E), we get

$$\begin{aligned} \omega'(t) &= \frac{\alpha\tau'(t)R^{\alpha-1}(\tau(t), t_0)}{r^{1/\alpha}(\tau(t))} \frac{r(t) (y'(t))^\alpha}{y^\alpha(\tau(t))} - R^\alpha(\tau(t), t_0) \frac{(r(t) (y'(t))^\alpha)'}{y^\alpha(\tau(t))} \\ &\quad - R^\alpha(\tau(t), t_0) \frac{\alpha r(t) (y'(t))^\alpha y'(\tau(t))\tau'(t)}{y^{\alpha+1}(\tau(t))} \\ &= \frac{\alpha\tau'(t)}{R(\tau(t), t_0)r^{1/\alpha}(\tau(t))} \omega(t) - R^\alpha(\tau(t), t_0)q(t) \\ &\quad - R^\alpha(\tau(t), t_0) \frac{\alpha r(t) (y'(t))^\alpha y'(\tau(t))\tau'(t)}{y^{\alpha+1}(\tau(t))}. \end{aligned} \tag{2.9}$$

Lemma 2.2 implies that, for each $\epsilon > 0$, there is $t_2 \in [t_1, \infty)$ large enough such that

$$y'(\tau(t)) \geq \left((\lambda(\rho) - \epsilon) \frac{r(t)}{r(\tau(t))} \right)^{1/\alpha} y'(t), \quad \text{on } [t_2, \infty). \tag{2.10}$$

Combining (2.8)–(2.10), we obtain

$$\begin{aligned} \omega'(t) &\leq \frac{\alpha\tau'(t)}{R(\tau(t), t_0)r^{1/\alpha}(\tau(t))} \omega(t) - R^\alpha(\tau(t), t_0)q(t) \\ &\quad - \alpha\tau'(t) \left(\frac{\lambda(\rho) - \epsilon}{R^\alpha(\tau(t), t_0)r(\tau(t))} \right)^{1/\alpha} \omega^{(\alpha+1)/\alpha}(t). \end{aligned} \tag{2.11}$$

Using the inequality

$$Au - Bu^{(\alpha+1)/\alpha} \leq \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}} \frac{A^{\alpha+1}}{B^\alpha}, \quad A \geq 0, B > 0, u \geq 0$$

with

$$A := \frac{\alpha\tau'(t)}{R(\tau(t), t_0)r^{1/\alpha}(\tau(t))} \quad \text{and} \quad B := \alpha\tau'(t) \left(\frac{\lambda(\rho) - \epsilon}{R^\alpha(\tau(t), t_0)r(\tau(t))} \right)^{1/\alpha}$$

in (2.11), we obtain

$$w'(t) \leq -R^\alpha(\tau(t), t_0)q(t) + \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1} \frac{1}{\lambda(\rho) - \epsilon} \frac{\tau'(t)}{R(\tau(t), t_0)r^{1/\alpha}(\tau(t))}.$$

Integrating the above inequality from t_2 to t , we arrive at

$$\int_{t_2}^t \left(R^\alpha(\tau(s), t_0)q(s) - \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1} \frac{1}{\lambda(\rho) - \epsilon} \frac{\tau'(s)}{R(\tau(s), t_0)r^{1/\alpha}(\tau(s))} \right) ds \leq \omega(t_2) - \omega(t) \leq \omega(t_2),$$

which contradicts (2.7) as $t \rightarrow \infty$. The proof is complete. □

Theorem 2.5. Assume that $0 < \rho \leq 1/e$. If

$$\liminf_{t \rightarrow \infty} R^\alpha(\tau(t), t_0) \int_t^\infty q(s)ds > \frac{1}{\lambda(\rho)} \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}}, \tag{2.12}$$

then (E) is oscillatory.

Proof. It suffices to prove that (2.12) implies (2.7). If we admit that (2.7) fails, then for all $\tilde{\epsilon} > 0$ there exists a $t_1 \in [t_0, \infty)$ such that for any $t \geq t_1$,

$$\int_t^\infty \left(R^\alpha(\tau(s), t_0)q(s) - \frac{1}{\lambda(\rho) - \epsilon} \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1} \frac{\tau'(s)}{R(\tau(s), t_0)r^{1/\alpha}(\tau(s))} \right) ds < \tilde{\epsilon}.$$

Since $R(\tau(t), t_0)$ is increasing, it is easy to see that

$$R^\alpha(\tau(t), t_0) \times \int_t^\infty \left(q(s) - \frac{1}{\lambda(\rho) - \epsilon} \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1} \frac{\tau'(s)}{R^{\alpha+1}(\tau(s), t_0)r^{1/\alpha}(\tau(s))} \right) ds < \tilde{\epsilon},$$

or

$$R^\alpha(\tau(t), t_0) \int_t^\infty \left(q(s) + \frac{1}{\lambda(\rho) - \epsilon} \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}} \left(\frac{1}{R^\alpha(\tau(s), t_0)} \right)' \right) ds < \tilde{\epsilon}.$$

Hence,

$$R^\alpha(\tau(t), t_0) \int_t^\infty q(s)ds < \tilde{\epsilon} + \frac{1}{\lambda(\rho) - \epsilon} \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}}$$

for all $\tilde{\epsilon} > 0$, which contradicts (2.12). The proof is complete. □

Lemma 2.6. Suppose that $\rho > 0$ and (E) has an eventually positive solution. Then, for any $k \in \mathbb{N}$, $0 < \rho_k \leq 1/e$ and

$$\liminf_{t \rightarrow \infty} \frac{r(\tau(t)) (y'(\tau(t)))^\alpha}{r(t) (y'(t))^\alpha} \geq \lambda(\rho_k). \tag{2.13}$$

Proof. By Lemma 2.2, it is clear that the statement holds for $k = 1$, i.e., $\rho_1 \leq 1/e$ and, for each $\epsilon > 0$, there is $t_2 \in [t_1, \infty)$ such that

$$\frac{r(\tau(t)) (y'(\tau(t)))^\alpha}{r(t) (y'(t))^\alpha} \geq \lambda(\rho_1) - \epsilon, \quad \text{on } [t_2, \infty). \tag{2.14}$$

Now, employing the chain rule

$$(r(t) (y'(t))^\alpha)' = \alpha (r^{1/\alpha}y'(t))^{\alpha-1} (r^{1/\alpha}y'(t))'$$

in the equality

$$(y(t) - r^{1/\alpha}(t)y'(t)R(t, t_0))' = -R(t, t_0) (r^{1/\alpha}y'(t))',$$

we get

$$\left(y(t) - r^{1/\alpha}(t)y'(t)R(t, t_0)\right)' = -\frac{1}{\alpha}R(t, t_0) \left(r^{1/\alpha}y'(t)\right)^{1-\alpha} (r(t) (y'(t))^\alpha)',$$

which, by virtue of (E), becomes

$$\left(y(t) - r^{1/\alpha}(t)y'(t)R(t, t_0)\right)' = \frac{1}{\alpha}R(t, t_0) \left(r^{1/\alpha}y'(t)\right)^{1-\alpha} q(t)y^\alpha(\tau(t)). \quad (2.15)$$

Integrating (2.15) from t_2 to t yields to

$$\phi(t) = \phi(t_2) + \frac{1}{\alpha} \int_{t_2}^t R(s, t_0) \left(r^{1/\alpha}y'(s)\right)^{1-\alpha} q(s)y^\alpha(\tau(s))ds, \quad (2.16)$$

where we set $\phi(t) = y(t) - r^{1/\alpha}(t)y'(t)R(t, t_0)$. It is clear from (2.5) that $\phi(t)$ is positive on $[t_2, \infty)$. Now, using (2.5) and (2.14) in (2.16), we arrive at

$$\begin{aligned} \phi(t) &\geq \phi(t_2) + \frac{1}{\alpha} \int_{t_2}^t R(s, t_0) \left(r^{1/\alpha}(s)y'(s)\right)^{1-\alpha} \times \\ &\quad \times q(s)r(\tau(s)) (y'(\tau(s)))^\alpha R^\alpha(\tau(s), t_0)ds \\ &\geq \phi(t_2) + \frac{\lambda(\rho_1) - \epsilon}{\alpha} \int_{t_2}^t R(s, t_0) \left(r^{1/\alpha}(s)y'(s)\right)^{1-\alpha} \times \\ &\quad \times q(s)r(s) (y'(s))^\alpha R^\alpha(\tau(s), t_0)ds \\ &= \phi(t_2) + \frac{\lambda(\rho_1) - \epsilon}{\alpha} \int_{t_2}^t R(s, t_0)R^\alpha(\tau(s), t_0)r^{1/\alpha}(s)y'(s)q(s)ds. \end{aligned}$$

Using the nondecreasing character of $r(t) (y'(t))^\alpha$ in the latter inequality, we obtain

$$\begin{aligned} \phi(t) &\geq \phi(t_2) + \frac{\lambda(\rho_1) - \epsilon}{\alpha} r^{1/\alpha}(t)y'(t) \int_{t_2}^t R(s, t_0)R^\alpha(\tau(s), t_0)q(s)ds \\ &= \phi(t_2) - \frac{\lambda(\rho_1) - \epsilon}{\alpha} r^{1/\alpha}(t)y'(t) \int_{t_0}^{t_2} R(s, t_0)R^\alpha(\tau(s), t_0)q(s)ds \\ &\quad + \frac{\lambda(\rho_1) - \epsilon}{\alpha} r^{1/\alpha}(t)y'(t) \int_{t_0}^t R(s, t_0)R^\alpha(\tau(s), t_0)q(s)ds. \end{aligned} \quad (2.17)$$

By virtue of (2.3) and the positivity of ϕ , we have

$$\phi(t) > \frac{\lambda(\rho_1) - \epsilon}{\alpha} r^{1/\alpha}(t)y'(t) \int_{t_0}^t R(s, t_0)R^\alpha(\tau(s), t_0)q(s)ds \quad \text{on } [t_3, \infty), \quad (2.18)$$

for some $t_3 \in [t_2, \infty)$ large enough. Hence,

$$y(t) > r^{1/\alpha}(t)y'(t) \left(R(t, t_0) + \frac{\lambda(\rho_1) - \epsilon}{\alpha} \int_{t_0}^t R(s, t_0)R^\alpha(\tau(s), t_0)q(s)ds \right) \quad (2.19)$$

or

$$y(t) > r^{1/\alpha}(t)y'(t)R_1(t, t_0, \epsilon) \quad \text{on } [t_3, \infty), \quad (2.20)$$

where

$$R_1(t, t_0, \epsilon) = R(t, t_0) + \frac{\lambda(\rho_1) - \epsilon}{\alpha} \int_{t_0}^t R(s, t_0)R^\alpha(\tau(s), t_0)q(s)ds.$$

Using (2.20) in (E), we see that $x(t) := r(t) (y'(t))^\alpha$ is a positive solution of the first-order delay differential inequality

$$x'(t) + R_1^\alpha(\tau(t), t_0, \epsilon)q(t)x(\tau(t)) < 0. \quad (2.21)$$

Applying Lemma 2.1 to (2.21), it is clear that the conclusion holds for $k = 2$, that is, $\rho_2 \leq 1/e$ and, for each $\epsilon > 0$, there is $t_4 \in [t_1, \infty)$ such that

$$\frac{r(\tau(t)) (y'(\tau(t)))^\alpha}{r(t) (y'(t))^\alpha} \geq \lambda(\rho_2) - \epsilon, \quad \text{on } [t_4, \infty). \quad (2.22)$$

Repeating the above process with (2.5) being replaced by (2.20), one can show that

$$y(t) > r^{1/\alpha}(t)y'(t)R_2(t, t_0, \epsilon) \quad \text{on } [t_5, \infty), \tag{2.23}$$

for some $t_5 \in [t_4, \infty)$, where

$$R_2(t, t_0, \epsilon) = R(t, t_0) + \frac{\lambda(\rho_2) - \epsilon}{\alpha} \int_{t_0}^t R(s, t_0)R^\alpha(\tau(s), t_0)q(s)ds.$$

Using (2.20) in (E) and applying Lemma 2.1 to a resulting inequality, we see that the Lemma conclusion holds for $k = 3$. By induction, it is not hard to show that the same conclusion holds for any $k \in \mathbb{N}$. The proof is complete. \square

Using Lemma 2.6 instead of Lemma 2.2, we are ready to improve Theorems 2.3–2.5. Since the proofs are the same, we omit them.

Theorem 2.7. *If $\rho_k > 1/e$ for some $k \in \mathbb{N}$, then (E) is oscillatory.*

Theorem 2.8. *Assume that $0 < \rho_i \leq 1/e$, $i = 1, 2, \dots, k$, for some $k \in \mathbb{N}$. If*

$$\limsup_{t \rightarrow \infty} \int_{t_0}^t \left(R^\alpha(\tau(s), t_0)q(s) - \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1} \frac{1}{(\lambda(\rho_k) - \epsilon)} \frac{\tau'(s)}{R(\tau(s), t_0)r^{1/\alpha}(\tau(s))} \right) ds = \infty \tag{2.24}$$

for some $\epsilon > 0$, then (E) is oscillatory.

Theorem 2.9. *Assume that $0 < \rho_i \leq 1/e$, $i = 1, 2, \dots, k$, for some $k \in \mathbb{N}$. If*

$$\liminf_{t \rightarrow \infty} R^\alpha(\tau(t), t_0) \int_t^\infty q(s)ds > \frac{1}{\lambda(\rho_k)} \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}}, \tag{2.25}$$

then (E) is oscillatory.

Finally, we give an example to illustrate the efficiency of our results.

Example 2.10. Consider the second-order delay differential equation of the Euler type:

$$((y'(t))^\alpha)' + \frac{q_0}{t^{\alpha+1}}y^\alpha(mt) = 0, \quad t \geq 1, \tag{2.26}$$

where α is a quotient of odd positives integers, $q_0 > 0$, $m \in (0, 1)$.

Note that the known condition (1.7) (or (1.4)) requires

$$q_0m^\alpha > \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}} \tag{2.27}$$

for (2.26) to be oscillatory.

By Theorem 2.3, we have that Eq. (2.26) is oscillatory if

$$\rho := q_0m^\alpha \ln \frac{1}{m} > \frac{1}{e}. \tag{2.28}$$

Now consider the case that (2.28) fails, that is, if $\rho \leq 1/e$. By Theorem 2.4 (or Theorem 2.5), we deduce that (2.26) is oscillatory if

$$q_0m^\alpha > \frac{1}{\lambda(\rho)} \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}}. \tag{2.29}$$

Since $\lambda(\rho) \in [1, e)$, our result improves (2.27).

Next, let us illustrate how Theorems 2.7 and 2.8 apply when both condition (2.28) and (2.29) fail. By Theorem 2.7, equation (2.26) is oscillatory if, for some $k \in \mathbb{N}$,

$$\rho_k > \frac{1}{e}, \tag{2.30}$$

where

$$\rho_1 := \rho$$

and, for $\rho_i \in (0, 1/e]$, $i \in \mathbb{N}$,

$$\rho_{i+1} := q_0 m^\alpha \left(1 + \frac{\lambda(\rho_i)}{\alpha} m^\alpha q_0 \right)^\alpha \ln \frac{1}{m}, \quad i \in \mathbb{N}.$$

If, in k th iteration step, condition (2.30) fails, then, by Theorem 2.8 (or Theorem 2.9), we have that (2.26) is oscillatory if

$$q_0 m^\alpha > \frac{1}{\lambda(\rho_k)} \frac{\alpha^\alpha}{(\alpha + 1)^{\alpha+1}}. \quad (2.31)$$

Now, let us consider a particular case of (2.26), namely,

$$\left((y'(t))^3 \right)' + \frac{11}{t^4} y^3(0.2t) = 0. \quad (2.32)$$

Note that condition (2.27) is not applicable, since $0.088 \not\geq 0.105469$. Using the definition of ρ_k , we get $\rho_1 = 0.141631 \not\geq 1/e$, $\rho_2 = 0.156883 \not\geq 1/e$. Then condition (2.31) with $k = 2$ gives $0.106376 > 0.105469$. Hence, by Theorem 2.8, (2.32) is oscillatory.

Finally, we consider another particular case of (2.26), namely,

$$\left((y'(t))^{1/3} \right)' + \frac{0.4}{t^{4/3}} y^{1/3}(0.4t) = 0. \quad (2.33)$$

Note that condition (2.27) is not applicable, since $0.294723 \not\geq 0.47247$. Using the definition of ρ_k , we get $\rho_1 = 0.270052 \not\geq 1/e$, $\rho_2 = 0.357776 \not\geq 1/e$, $\rho_3 = 0.386561 > 1/e$. Thus, condition (2.30) is satisfied for $k = 3$, and by Theorem 2.7, we conclude that (2.33) is oscillatory.

We remark that none of the oscillation criteria presented in [8–10, 14, 16, 20–24] can be applied to equation (2.32) or (2.33).

Remark 2.11. The results presented in this paper strongly depend on the properties of first-order delay differential equations. An interesting problem for further research is to establish different iterative techniques for testing oscillations in (E) independently on the constant $1/e$.

Acknowledgment. The authors would like to thank the editors and three referees for the constructive remarks which improved the presentation of the paper. The work of the second author was supported by the grant project KEGA 035TUKE-4/2017.

References

- [1] R. P. Agarwal, M. Bohner and W.-T. Li, *Nonoscillation and oscillation: theory for functional differential equations*, Monographs and Textbooks in Pure and Applied Mathematics **267**, Marcel Dekker, Inc., New York, 2004.
- [2] R. P. Agarwal, S. R. Grace and D. O'Regan, *Oscillation theory for second order linear, half-linear, superlinear and sublinear dynamic equations*, Kluwer Academic Publishers, Dordrecht, 2002.
- [3] R. P. Agarwal, S. R. Grace and D. O'Regan, *Oscillation theory for second order dynamic equations*, Series in Mathematical Analysis and Applications **5**, Taylor & Francis, Ltd., London, 2003.
- [4] R. P. Agarwal, S. R. Grace and D. O'Regan, *Oscillation theory for difference and functional differential equations*, Springer Science & Business Media, 2013.
- [5] R. P. Agarwal, C. Zhang and T. Li, *Some remarks on oscillation of second order neutral differential equations*, Appl. Math. Comput. **274**, 178–181, 2016.
- [6] O. Došlý and P. Rehák, *Half-linear differential equations*, North-Holland Mathematics Studies **202**, Elsevier Science B.V., Amsterdam, 2005.
- [7] J. Džurina and I. Jadlovská, *A note on oscillation of second-order delay differential equations*, Appl. Math. Lett. **69**, 126–132, 2017.

- [8] J. Džurina and I. P. Stavroulakis, *Oscillation criteria for second-order delay differential equations*, Appl. Math. Comput. **140** (2-3), 445–453, 2003.
- [9] L. Erbe, A. Peterson and S. H. Saker, *Kamenev-type oscillation criteria for second-order linear delay dynamic equations*, Dynam. Systems Appl. **15** (1), 65–78, 2006.
- [10] S. Fišnarová and R. Mařík, *Oscillation of half-linear differential equations with delay*, Abstr. Appl. Anal. **2013**, Article ID: 583147, 1–6, 2013.
- [11] I. Gyóri and G. Ladas, *Oscillation theory of delay differential equations*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1991.
- [12] Z. Han, T. Li, S. Sun and Y. Sun, *Remarks on the paper [Appl. Math. Comput. 207 (2009) 388–396] [mr2489110]*, Appl. Math. Comput. **215** (11), 3998–4007, 2010.
- [13] J. Jaroš and I. P. Stavroulakis, *Oscillation tests for delay equations*, Rocky Mountain J. Math. **29** (1), 197–207, 1999.
- [14] R. Koplatadze, G. Kvinikadze and I. P. Stavroulakis, *Oscillation of second order linear delay differential equations*, Funct. Differ. Equ. **7** (1-2), 121–145, 2000.
- [15] R. G. Koplatadze, *Criteria for the oscillation of solutions of differential inequalities and second-order equations with retarded argument*, Tbiliss. Gos. Univ. Inst. Prikl. Mat. Trudy, **17**, 104–121, 1986.
- [16] T. Kusano and J. Wang, *Oscillation properties of half-linear functional-differential equations of the second order*, Hiroshima Math. J. **25** (2), 371–385, 1995.
- [17] T. Li, Y. V. Rogovchenko and C. Zhang, *Oscillation of second-order neutral differential equations*, Funkcial. Ekvac. **56** (1), 111–120, 2013.
- [18] T. Li, Y. V. Rogovchenko and C. Zhang, *Oscillation results for second-order nonlinear neutral differential equations*, Adv. Difference Equ. **2013** (336), 1–13, 2013.
- [19] W. E. Mahfoud, *Comparison theorems for delay differential equations*, Pacific J. Math. **83** (1), 187–197, 1979.
- [20] Y. G. Sun and F. W. Meng, *Note on the paper of Džurina and Stavroulakis: “Oscillation criteria for second-order delay differential equations” [Appl. Math. Comput. 140 (2003), no. 2-3, 445–453; mr1953915]*, Appl. Math. Comput. **174** (2), 1634–1641, 2006.
- [21] A. Tiryaki, *Oscillation criteria for a certain second-order nonlinear differential equations with deviating arguments*, Electron. J. Qual. Theory Differ. Equ. **2009** (61), 1–11, 2009.
- [22] J. J. Wei, *Oscillation of second order delay differential equation*, Ann. Differential Equations, **4** (4), 473–478, 1988.
- [23] H. Wu, L. Erbe and A. Peterson, *Oscillation of solution to second-order half-linear delay dynamic equations on time scales*, Electron. J. Differential Equations, **2016** (71), 1–15, 2016.
- [24] R. Xu and F. Meng, *Some new oscillation criteria for second order quasi-linear neutral delay differential equations*, Appl. Math. Comput. **182** (1), 797–803, 2006.
- [25] L. Ye and Z. Xu, *Oscillation criteria for second order quasilinear neutral delay differential equations*, Appl. Math. Comput. **207** (2), 388–396, 2009.
- [26] C. Zhang, R. P. Agarwal, M. Bohner and T. Li, *Oscillation of second-order nonlinear neutral dynamic equations with noncanonical operators*, Bull. Malays. Math. Sci. Soc. **38** (2), 761–778, 2015.



A note on distributivity of the lattice of L -ideals of a ring

Dilek Bayrak^{*1}, Sultan Yamak²

¹*Department of Mathematics, Tekirdağ Namık Kemal University, 59030, Tekirdağ, Turkey*

²*Department of Mathematics, Karadeniz Technical University, 61080, Trabzon, Turkey*

Abstract

Many studies have investigated the lattice of fuzzy substructures of algebraic structures such as groups and rings. In this study, we prove that the lattice of L -ideals of a ring is distributive if and only if the lattice of its ideals is distributive, for an infinitely \vee -distributive lattice L .

Mathematics Subject Classification (2010). 08A72, 13A15

Keywords. lattice, distributive lattice, fuzzy ideal

1. Introduction

The lattice-theoretic aspects of algebraic substructures and L -algebraic substructures have been a topic of discussion in the literature for quite some time. It follows as a consequence of the subdirect product theorem formulated by Professor Tom Head in [9] that the properties of the lattice of algebraic substructures and that of corresponding fuzzy algebraic substructures are almost identical. However before the emergence of the subdirect product theorem, the modularity of the lattice of fuzzy normal subgroups of a group and the modularity of the lattice of fuzzy ideals of a ring have been established in [1–5, 10, 17].

The distributivity constitutes a very powerful property of a lattice. On the other hand, Tarnaucedanu [15] worked on finite groups and proved that a group is cyclic iff its lattice of fuzzy subgroups is distributive. Majumdar and Sultana [13] proved that the lattice of fuzzy ideals of a ring is distributive. However, Kumar [12] has obtained just the opposite of this result. Also Zhang and Meng [18] gave a counter example for the result of Majumdar and Sultana. Recently in [11] the modularity of L -ideals of a ring is established, where the subdirect product theorem of Tom Head does not apply. Finally, the lattice of L -fuzzy extended ideals is studied in [7]. We ask: is the lattice of all L -ideals of a ring distributive whose lattice of all ideals is distributive? This paper will answer the question for an infinitely \vee -distributive lattice. In this paper, we propose an analogous connection between the lattice of L -ideals and the lattice of ideals of a ring. We first describe some properties of the lattice of L -ideals that are tools to obtain some results. Using these results, we prove that the lattice of L -ideals is distributive when the lattice of ideals is distributive for an infinitely \vee -distributive lattice L .

^{*}Corresponding Author.

Email addresses: dbayrak@nku.edu.tr (D. Bayrak), syamak@ktu.edu.tr (S. Yamak)

Received: 14.03.2013; Accepted: 18.10.2017

2. Preliminaries

In this section, we briefly recall some basic concepts of lattices, L -subsets and rings. Throughout this paper, L is a completely lattice with the least element 0 and the greatest element 1. For every family $\{b_i \mid i \in \Delta\}$, we can popularize some operations such as

$$\bigvee_{i \in \Delta} b_i = \sup\{b_i \mid i \in \Delta\}, \quad \bigwedge_{i \in \Delta} b_i = \inf\{b_i \mid i \in \Delta\}.$$

A complete lattice L is called infinitely \vee -distributive lattice if for all $\alpha \in L$ and $\Delta \subseteq L$,

$$\alpha \wedge \left(\bigvee_{\beta \in \Delta} \beta \right) = \bigvee_{\beta \in \Delta} (\alpha \wedge \beta).$$

For a nonempty set X , an L -subset is any function from X into L , which is introduced by Goguen [8] as a generalization of the notion of Zadeh's fuzzy subset [16]. The class of L -subsets of X will be denoted by $F(X, L)$. In particular, if $L = [0, 1]$, it is appropriate to replace fuzzy subset with L -subset. In this case the set of all fuzzy subsets of X is denoted by $F(X)$. Let μ and ν be L -subsets of X . We say that μ is contained in ν if $\mu(x) \leq \nu(x)$ for every $x \in X$, denoted $\mu \leq \nu$. Then \leq is a partial ordering on $F(X, L)$.

For each $\alpha \in L$, we define the level subset

$$\mu_\alpha = \{x \in X \mid \alpha \leq \mu(x)\}$$

Let μ_i ($i \in \Delta$) be an L -subset of X . Define the intersection as follows:

$$\left(\bigcap_{i \in \Delta} \mu_i \right)(x) = \bigwedge_{i \in \Delta} \mu_i(x)$$

for all $x \in X$. The characteristic function of a set $A \subseteq X$ is denoted by 1_A .

Throughout this paper, R stands for a commutative ring with identity. $I(R)$ stands for all ideals of R , is a complete lattice with respect to set inclusion, called the ideals lattice of R . Note that $I(R)$ has initial element $\{0\}$ and final element R , and its binary operations \wedge, \vee are defined by $I \wedge J = I \cap J$ and $I \vee J = I + J$, for all $I, J \in I(R)$. $I(R)$ may not be a distributive lattice. For example, let $R = \mathbb{Z} \times \mathbb{Z}$, \mathbb{Z} is the ring of integers, we define the operations as follows:

$$(a, b) + (c, d) = (a + c, b + d) \text{ and } (a, b) \cdot (c, d) = (0, 0)$$

for any $(a, b), (c, d) \in \mathbb{Z} \times \mathbb{Z}$. Then $(R, +, \cdot)$ form a ring with zero $(0, 0)$.

$$\{(x, x) \mid x \in \mathbb{Z}\} \cap ((\mathbb{Z} \times \{0\}) + (\{0\} \times \mathbb{Z})) = \{(x, x) \mid x \in \mathbb{Z}\},$$

whereas

$$(\{(x, x) \mid x \in \mathbb{Z}\} \cap (\mathbb{Z} \times \{0\})) + (\{(x, x) \mid x \in \mathbb{Z}\} \cap (\{0\} \times \mathbb{Z})) = \{(0, 0)\}.$$

The further knowledge about lattices and rings required in this paper can be found in [6, 14].

3. L -ideals

In this section, we investigate the lattice structure of L -ideals of a ring R .

Definition 3.1. [14] Let μ be an L -subset in a ring R . Then μ is called an L -ideal of R if

$$\mu(x - y) \geq \mu(x) \wedge \mu(y) \text{ and } \mu(xy) \geq \mu(x) \vee \mu(y)$$

for all $x, y \in R$. The family of all L -ideals is denoted by $FI(R, L)$. In particular, when $L = [0, 1]$, an L -ideal of R is referred to as a fuzzy ideal of R . The family of all fuzzy ideals is denoted by $FI(R)$.

The following lemma easily obtained from Proposition 2.2.[17].

Lemma 3.2. Let $\mu \in F(R, L)$. Then μ is an L -ideal of R iff $\mu_\alpha = \emptyset$ or μ_α is a classical ideal of R , for any $\alpha \in L$.

Theorem 3.3. [11] Let μ_i ($i \in \Delta$) be an L -ideal of a ring R . Then $\bigcap_{i \in \Delta} \mu_i$ is an L -ideal of R .

By the Theorem 3.3., we immediately get the next corollary.

Corollary 3.4. $FI(R, L)$ is a complete lattice under the ordering of L -set inclusion such that $\bigwedge_{i \in \Delta} \mu_i = \bigcap_{i \in \Delta} \mu_i$ for all $\mu_i \in FI(R, L)$ ($i \in \Delta$).

Lemma 3.5. [2] Let A, B be subsets of R . Then

- (1) A is an ideal of R if and only if 1_A is an L -ideal of R ,
- (2) If A, B are ideals of R , then $1_A \vee 1_B = 1_{A+B}$ and $1_A \wedge 1_B = 1_{A \cap B}$.
- (3) $\{1_A \mid A \text{ is an ideal of } R\}$ is a sublattice of $FI(R, L)$

4. The distributivity of $FI(R, L)$

In this section we will investigate some conditions related to distributivity of the lattice of L -ideals of a ring R .

Definition 4.1. Let μ and ν be L -subsets of a ring R . Define $\mu \oplus \nu$ as follows:

$$(\mu \oplus \nu)(x) = \mu(x) \vee \nu(x) \vee \bigvee_{x=y+z} \mu(y) \wedge \nu(z)$$

for all $x \in R$.

Lemma 4.2. Let L be an infinitely \vee -distributive lattice and $\mu, \nu \in FI(L, R)$. Then $\mu \vee \nu = \mu \oplus \nu$.

Proof. Let $x, y \in R$. Then

$$\begin{aligned} & \mu \oplus \nu(x) \wedge \mu \oplus \nu(y) \\ &= [\mu(x) \vee \nu(x) \vee (\bigvee_{x=a+b} \mu(a) \wedge \nu(b))] \wedge [\mu(y) \vee \nu(y) \vee (\bigvee_{y=c+d} \mu(c) \wedge \nu(d))] \\ &= [(\mu(x) \vee \nu(x)) \wedge (\mu(y) \vee \nu(y))] \vee [(\mu(x) \vee \nu(x)) \wedge (\bigvee_{y=c+d} \mu(c) \wedge \nu(d))] \\ & \quad \vee [(\mu(y) \vee \nu(y)) \wedge (\bigvee_{x=a+b} \mu(a) \wedge \nu(b))] \vee [(\bigvee_{y=c+d} \mu(c) \wedge \nu(d)) \wedge (\bigvee_{x=a+b} \mu(a) \wedge \nu(b))] \\ &= (\mu(x) \wedge \mu(y)) \vee (\mu(x) \wedge \nu(y)) \vee (\nu(x) \wedge \mu(y)) \vee (\nu(x) \wedge \nu(y)) \\ & \quad \vee (\bigvee_{y=c+d} \mu(x) \wedge \mu(c) \wedge \nu(d)) \vee (\bigvee_{y=c+d} \nu(x) \wedge \mu(c) \wedge \nu(d)) \vee (\bigvee_{x=a+b} \mu(a) \wedge \nu(b) \wedge \mu(y)) \\ & \quad \vee (\bigvee_{x=a+b} \mu(a) \wedge \nu(b) \wedge \nu(y)) \vee (\bigvee_{\substack{x=a+b \\ y=c+d}} \mu(a) \wedge \nu(b) \wedge \mu(c) \wedge \nu(d)) \\ & \leq \mu(x+y) \vee \nu(x+y) \vee (\mu(x) \wedge \nu(y)) \vee (\mu(y) \wedge \nu(x)) \vee (\bigvee_{y=c+d} \mu(x+c) \wedge \nu(d)) \\ & \quad \vee (\bigvee_{y=c+d} \mu(c) \wedge \nu(d+x)) \vee (\bigvee_{x=a+b} \mu(a+y) \wedge \nu(b)) \\ & \quad \vee (\bigvee_{x=a+b} \mu(a) \wedge \nu(b+y)) \vee (\bigvee_{\substack{x=a+b \\ y=c+d}} \mu(a+c) \wedge \nu(b+d)) \\ & \leq \mu(x+y) \vee \nu(x+y) \vee (\bigvee_{x+y=u+v} \mu(u) \wedge \nu(v)) \\ & = \mu \oplus \nu(x+y) \end{aligned}$$

Hence $\mu \oplus \nu(x) \wedge \mu \oplus \nu(y) \leq \mu \oplus \nu(x + y)$.

$$\begin{aligned} \mu \oplus \nu(-x) &= \mu(-x) \vee \nu(-x) \vee \left(\bigvee_{-x=a+b} \mu(a) \wedge \nu(b) \right) \\ &= \mu(-x) \vee \nu(-x) \vee \left(\bigvee_{x=(-a)+(-b)} \mu(-(-a)) \wedge \nu(-(-b)) \right) \\ &\leq \mu(x) \vee \nu(x) \vee \left(\bigvee_{x=u+v} \mu(u) \wedge \nu(v) \right) \\ &= \mu \oplus \nu(x) \end{aligned}$$

Hence $\mu \oplus \nu(-x) \leq \mu \oplus \nu(x)$.

$$\begin{aligned} \mu \oplus \nu(x) &= \mu(x) \vee \nu(x) \vee \left(\bigvee_{x=a+b} \mu(a) \wedge \nu(b) \right) \\ &\leq \mu(xy) \vee \nu(xy) \vee \left(\bigvee_{x=a+b} \mu(ay) \wedge \nu(by) \right) \\ &\leq \mu(xy) \vee \nu(xy) \vee \left(\bigvee_{xy=u+v} \mu(u) \wedge \nu(v) \right) \\ &= \mu \oplus \nu(xy) \end{aligned}$$

Similarly, we have $\mu \oplus \nu(y) \leq \mu \oplus \nu(xy)$. Thus $\mu \oplus \nu \in FI(R, L)$. It is clear that $\mu \leq \mu \oplus \nu$ and $\nu \leq \mu \oplus \nu$.

Let $\theta \in FI(R, L)$ such that $\mu \leq \theta$ and $\nu \leq \theta$. Then

$$\mu(a) \wedge \nu(b) \leq \theta(a) \wedge \theta(b) \leq \theta(a + b) = \theta(x)$$

for all $x = a + b$. By the definition of $\mu \oplus \nu$, it follows that $\mu \oplus \nu \leq \theta$. Hence $\mu \vee \nu = \mu \oplus \nu$ \square

The following theorem gives the main results of this section.

Theorem 4.3. *If L is an infinitely \vee -distributive lattice, then the following conditions are equivalent:*

- (1) $I(R)$ is a distributive lattice,
- (2) $FI(R, L)$ is a distributive lattice.

Proof. (2) \Rightarrow (1) By Lemma 3.5, it is clear.

(1) \Rightarrow (2) Let $\mu, \nu, \theta \in FI(R)$. Since the distributive inequality is valid for every lattice, we have

$$(\mu \wedge \nu) \vee (\mu \wedge \theta) \leq \mu \wedge (\nu \vee \theta).$$

And by Lemma 4.2 and Corollary 3.4,

$$\begin{aligned} (\mu \wedge (\nu \vee \theta))(x) &= (\mu \wedge (\nu \oplus \theta))(x) \\ &= \mu(x) \wedge [\nu(x) \vee \theta(x) \vee \left(\bigvee_{x=a+b} \nu(a) \wedge \theta(b) \right)] \\ &= (\mu(x) \wedge \nu(x)) \vee (\mu(x) \wedge \theta(x)) \vee \left(\bigvee_{x=a+b} \nu(a) \wedge \theta(b) \wedge \mu(x) \right) \end{aligned}$$

Let $\lambda = \nu(a) \wedge \theta(b) \wedge \mu(x)$ for some $a, b \in R$ such that $x = a + b$

Thus we have $x \in \mu_\lambda$, $a \in \nu_\lambda$, $b \in \theta_\lambda$. Then $x \in \mu_\lambda \cap (\nu_\lambda + \theta_\lambda)$. Due to distributivity of $I(R)$,

$$x \in (\mu_\lambda \cap \nu_\lambda) + (\mu_\lambda \cap \theta_\lambda).$$

It follows that there exist $u, v \in R$ such that $x = u + v$,

$$u \in \mu_\lambda \cap \nu_\lambda \quad \text{and} \quad v \in \mu_\lambda \cap \theta_\lambda.$$

Thus we have $\lambda \leq \mu(u)$, $\lambda \leq \nu(u)$, $\lambda \leq \mu(v)$, $\lambda \leq \theta(v)$. Hence,

$$\lambda \leq (\mu \wedge \nu)(u) \wedge (\mu \wedge \theta)(v).$$

Now it follows that

$$\lambda \leq \bigvee_{x=u+v} (\mu \wedge \nu)(u) \wedge (\mu \wedge \theta)(v).$$

Hence we obtain

$$\bigvee_{x=a+b} \nu(a) \wedge \theta(b) \wedge \mu(x) \leq \bigvee_{x=u+v} (\mu \wedge \nu)(u) \wedge (\mu \wedge \theta)(v).$$

Therefore,

$$\begin{aligned} (\mu \wedge (\nu \vee \theta))(x) &= (\mu(x) \wedge \nu(x)) \vee (\mu(x) \wedge \theta(x)) \vee (\bigvee_{x=a+b} \nu(a) \wedge \theta(b) \wedge \mu(x)) \\ &\leq (\mu \wedge \nu)(x) \vee (\mu \wedge \theta)(x) \vee (\bigvee_{x=u+v} (\mu \wedge \nu)(u) \wedge (\mu \wedge \theta)(v)) \\ &= ((\mu \wedge \nu) \oplus (\mu \wedge \theta))(x) \\ &= (\mu \wedge \nu) \vee (\mu \wedge \theta)(x). \end{aligned}$$

and the proof is completed. \square

By the Theorem 4.3, we immediately get the next corollary.

Corollary 4.4. *$I(R)$ is a distributive lattice if and only if $FI(R)$ is a distributive lattice.*

5. Conclusion

Many researches studied the lattice structure (distributive or modular) of fuzzy algebraic substructures. In future work, the same results could also be studied under a t-norm operation on L . Also, we will try to expose some classes of algebra whose lattices of L -subalgebras constitute distributive lattice.

References

- [1] N. Ajmal and K.V. Thomas, *The lattices of fuzzy subgroups and fuzzy normal subgroups*, Inform. Sci. **76**, 1–11, 1994.
- [2] N. Ajmal and K.V. Thomas, *The lattice of fuzzy ideals of a ring*, Fuzzy Sets and Systems **74**, 371–379, 1995.
- [3] N. Ajmal and K.V. Thomas, *The lattices of fuzzy normal subgroups is modular*, Inform. Sci. **83**, 199–218, 1995.
- [4] D. Bayrak and S. Yamak, *The lattice of generalized normal L -subgroups*, J. Intell. Fuzzy Syst. **27**, 1143–1152, 2014.
- [5] D. Bayrak and S. Yamak *Distributivity and pseudocomplementation of lattices of generalized L -subgroups*, Int. J. Algebra Stat. **5**, 107–114, 2016.
- [6] G. Birkhoff, *Lattice Theory*, Amer. Math. Soc. Colloq. Publ., Rhode Island, 1967.
- [7] N. Gao, Q. Li and Z. Li, *When do L -fuzzy ideals of a ring generate a distributive lattice?*, Open Math. **4** (1), 531–542, 2016.
- [8] J.A. Goguen, *L -fuzzy sets*, J. Math. Anal. Appl. **18**, 145–174, 1967.
- [9] T. Head, *A meta theorem for deriving fuzzy theorems from crisp versions*, Fuzzy Sets and Systems **73**, 349–358, 1995.
- [10] I. Jahan, *Modularity of Ajmal for the lattices of fuzzy ideals of a ring*, Iran. J. Fuzzy Syst. **5**, 71–78, 2008.
- [11] I. Jahan, *The lattice of L -ideals of a ring is modular*, Fuzzy Sets and Systems **199**, 121–129, 2012.
- [12] R. Kumar, *Non-distributivity of the lattice of fuzzy ideals of a ring*, Fuzzy Sets and Systems **97**, 393–394, 1998.
- [13] S. Majumdar and Q.S. Sultana, *The lattice of fuzzy ideals of a ring*, Fuzzy Sets and Systems **81**, 271–273, 1996.
- [14] J.N. Mordeson and D.S. Malik, *Fuzzy Commutative Algebra*, World Scientific, 1998.

- [15] M. Tarnauceanu, *Distributivity in lattices of fuzzy subgroups*, Inform. Sci. **179**, 1163–1168, 2009.
- [16] L.A. Zadeh, *Fuzzy sets*, Inform. Control **8**, 338–353, 1965.
- [17] Q. Zhang, *The lattice of fuzzy (left, right) ideals of a ring is modular*, Fuzzy Sets and Systems **125**, 209–214, 2002.
- [18] Q. Zhang and G. Meng, *On the lattice of fuzzy ideals of a ring*, Fuzzy Sets and Systems **112**, 349–353, 2000.



Hom-coalgebra cleft extensions and braided tensor Hom-categories of Hom-entwining structures

Chen Quanguo*, Wang Dingguo

School of Mathematical Sciences, Qufu Normal University, Shandong, Qufu 273165, China

Abstract

We investigate how the category of Hom-entwined modules can be made into a monoidal category. The sufficient and necessary conditions making the category of Hom-entwined modules have a braiding are given. Also, we formulate the concept of Hom-cleft extension for a Hom-entwining structure, and prove that if (A, α) is a (C, γ) -cleft extension, then there is an isomorphism of Hom-algebras between (A, α) and a crossed product Hom-algebra of A^{coC} and C .

Mathematics Subject Classification (2010). 16W30

Keywords. Hom-Hopf algebra, Hom-entwining structure, cleft extension

1. Introduction

Entwined modules were introduced by Brzeziński and Majid [2, 3], which contained the Long modules, Yetter-Drinfeld modules and Doi-Koppinen modules, etc. So it is very important to study entwined module. As a generalization of entwined modules, Hom-entwined modules were defined by Karacuha [14] as special examples of Hom-corings.

As we know, braided monoidal categories are special categories, whose importance is that the “braiding” structures provide a class of solutions to quantum Yang-Baxter equations. Thus constructing a class of braided monoidal categories is an interesting job. Caenepeel et al. studied how the category of Doi-Hopf modules can be made into a braided monoidal category [5], which have been generalized to entwined modules and Doi-Hom-Hopf modules [13, 17].

The definition of the normal basis for extension associated to a Hopf algebra was introduced by Kreimer and Takeuchi [15]. Using this notion, Doi and Takeuchi [11] characterized H -Galois extensions with normal basis in terms of H -cleft extensions. This result can be extended for Hopf algebras living in symmetric closed categories [12]. A more general formulation in the context of (weak)entwining structures can be found in [1, 3].

The main goal of this paper shall discuss how to make the category of Hom-entwined modules into a monoidal category, and introduce a definition of cleft extension for Hom-entwining structures and with it to obtain a general cleft extension theory. In Section 3, we construct a monoidal category of Hom-entwined modules and give the sufficient and

*Corresponding Author.

Email addresses: cqg211@163.com (C. Quanguo), dingguo95@126.com (W. Dingguo)

Received: 31.12.2016; Accepted: 18.10.2017

necessary conditions making the monoidal category into a braided category. In Section 4, we introduce the notion of (C, γ) -Hom-cleft extension $(A^{coC}, \alpha|_{A^{coC}}) \hookrightarrow (A, \alpha)$, being (A, α) a Hom-algebra, (C, γ) a Hom-coalgebra and A^{coC} a sub-Hom-algebra of A . We prove that if (A, α) is a (C, γ) -Hom-cleft extension, then there is an isomorphism of Hom-algebras between (A, α) and a crossed product Hom-algebra of A^{coC} and C .

2. Preliminaries

Throughout this paper, k will be a field. More knowledge about monoidal Hom-(co)algebra, monoidal Hopf Hom-algebra, Hom-entwined modules, etc. can be found in [4, 6–10, 13, 14, 16, 18–24]. Let $\mathcal{M} = (\mathcal{M}, \otimes, k, a, l, r)$ be the monoidal category of vector spaces over k . We can construct a new monoidal category $\mathcal{H}(\mathcal{M})$ whose objects are ordered pairs (M, μ) with $M \in \mathcal{M}$ and $\mu \in \text{Aut}(M)$ and morphisms $f : (M, \mu) \rightarrow (N, \nu)$ are morphisms $f : M \rightarrow N$ in \mathcal{M} satisfying $\nu \circ f = f \circ \mu$. The monoidal structure is given by $(M, \mu) \otimes (N, \nu) = (M \otimes N, \mu \otimes \nu)$ and (k, id_k) . All monoidal Hom-structures are objects in the tensor category $\tilde{\mathcal{H}}(\mathcal{M}) = (\mathcal{H}(\mathcal{M}), \otimes, (k, id_k), \tilde{a}, \tilde{l}, \tilde{r})$ introduced in [4] with the associativity and unit constraints given by

$$\tilde{a}_{M,N,C}((m \otimes n) \otimes p) = \mu(m) \otimes (n \otimes \gamma^{-1}(c)),$$

$$\tilde{l}(x \otimes m) = \tilde{r}(m \otimes x) = x\mu(m),$$

for $(M, \mu), (N, \nu)$ and (C, γ) . The category $\tilde{\mathcal{H}}(\mathcal{M})$ is termed Hom-category associated to \mathcal{M} .

2.1. Monoidal Hom-algebra

Recall from [4] that a monoidal Hom-algebra is an object $(A, \alpha) \in \tilde{\mathcal{H}}(\mathcal{M})$ together with a linear map $m_A : A \otimes A \rightarrow A$, $m_A(a \otimes b) = ab$ and an element $1 \in A$ such that

$$\alpha(ab) = \alpha(a)\alpha(b), \alpha(a)(bc) = (ab)\alpha(c), \tag{2.1}$$

$$\alpha(1) = 1, a1 = \alpha(a) = 1a, \tag{2.2}$$

for all $a, b, c \in A$.

A right (A, α) -Hom-module consists of an object $(M, \mu) \in \tilde{\mathcal{H}}(\mathcal{M})$ together with a linear map $\psi : M \otimes A \rightarrow M$, $\psi(m \otimes a) = ma$ satisfying the following conditions:

$$\mu(m)(ab) = (ma)\alpha(b), m1 = \mu(m), \tag{2.3}$$

for all $m \in M$ and $a, b \in A$. For ψ to be a morphism in $\tilde{\mathcal{H}}(\mathcal{M})$ means

$$\mu(ma) = \mu(m)\alpha(a). \tag{2.4}$$

We call that ψ is a right Hom-action of (A, α) on (M, μ) .

Let (M, μ) and (M', μ') be two right (A, α) -Hom-modules. We call a morphism $f : M \rightarrow M'$ right (A, α) -linear, if $f \circ \mu = \mu' \circ f$ and $f(ma) = f(m)a$. \mathcal{M}_A denotes the category of all right (A, α) -Hom-modules.

2.2. Monoidal Hom-coalgebras

Recall from [4] that a monoidal Hom-coalgebra is an object $(C, \gamma) \in \tilde{\mathcal{H}}(\mathcal{M})$ together with two linear maps $\Delta_C : C \rightarrow C \otimes C$, $\Delta_C(c) = c_1 \otimes c_2$ (summation implicitly understood) and $\varepsilon_C : C \rightarrow k$ such that

$$\gamma^{-1}(c_1) \otimes \Delta_C(c_2) = c_{11} \otimes (c_{12} \otimes \gamma^{-1}(c_2)), \Delta_C(\gamma(c)) = \gamma(c_1) \otimes \gamma(c_2), \tag{2.5}$$

$$\varepsilon_C(\gamma(c)) = \varepsilon_C(c), c_1\varepsilon_C(c_2) = \gamma^{-1}(c) = \varepsilon_C(c_1)c_2, \tag{2.6}$$

for all $c \in C$.

A right (C, γ) -Hom-comodule consists of an object $(M, \mu) \in \tilde{\mathcal{H}}(\mathcal{M})$ together with a linear map $\rho_M : M \rightarrow M \otimes C, \rho_M(m) = m_{[0]} \otimes m_{[1]}$ (summation implicitly understood) satisfying the following conditions:

$$\mu^{-1}(m_{[0]}) \otimes \Delta(m_{[1]}) = m_{[0][0]} \otimes (m_{[0][1]} \otimes \gamma^{-1}(m_{[1]})), \tag{2.7}$$

$$m_{[0]}\varepsilon_C(m_{[1]}) = \gamma^{-1}(m), \tag{2.8}$$

$$\mu(m)_{[0]} \otimes \mu(m)_{[1]} = \mu(m_{[0]}) \otimes \gamma(m_{[1]}), \tag{2.9}$$

for all $m \in M$. We call that ρ_M is a right Hom-coaction of (A, α) on (M, μ) .

Let (M, μ) and (M', μ') be two right (C, γ) -Hom-comodules. We call a morphism $f : M \rightarrow M'$ right (A, α) -colinear, if $f \circ \mu = \mu \circ f$ and $f(m)_{[0]} \otimes f(m)_{[1]} = f(m_{[0]}) \otimes m_{[1]}$. \mathcal{M}^C denotes the category of all right (C, γ) -Hom-comodules.

2.3. Monoidal Hom-Hopf algebra

A monoidal Hom-bialgebra $H = (H, \beta, m_H, 1, \Delta_H, \varepsilon_H)$ is a bialgebra in the category $\tilde{\mathcal{H}}(\mathcal{M})$. This means that $(H, \beta, m_H, 1)$ is a monoidal Hom-algebra and $(H, \beta, \Delta_H, \varepsilon_H)$ is a monoidal Hom-coalgebra such that Δ_H and ε_H are Hom-algebra maps, that is, for any $h, g \in H$,

$$\Delta_H(hg) = \Delta_H(h)\Delta_H(g), \Delta_H(1) = 1 \otimes 1, \tag{2.10}$$

$$\varepsilon_H(hg) = \varepsilon_H(h)\varepsilon_H(g), \varepsilon_H(1) = 1. \tag{2.11}$$

A monoidal Hom-bialgebra (H, β) is called a monoidal Hom-Hopf algebra, if there exists a morphism (called the antipode) $S : H \rightarrow H$ in $\tilde{\mathcal{H}}(\mathcal{M})$ such that

$$S(h_1)h_2 = \varepsilon_H(h)1 = h_1S(h_2), \tag{2.12}$$

for all $h \in H$.

2.4. Hom-Doi-Koppinen datum

Let (H, β) be a monoidal Hom-bialgebra. Recall from [14] that a right (H, β) -Hom-comodule algebra (A, α) is a monoidal Hom-algebra and a right (H, β) -Hom-comodule with a Hom-coaction ρ_A such that ρ_A is a Hom-algebra morphism, i.e., for any $a, a' \in A$,

$$(aa')_{[0]} \otimes (aa')_{[1]} = a_{[0]}a'_{[0]} \otimes a_{[1]}a'_{[1]}, \tag{2.13}$$

$$\rho_A(1) = 1 \otimes 1, \rho_A \circ \alpha = (\alpha \otimes \beta) \circ \rho_A. \tag{2.14}$$

A right (H, β) -Hom-module coalgebra (C, γ) is a monoidal Hom-coalgebra and a right (H, β) -Hom-module such that, for any $c \in C$ and $h \in H$,

$$(ch)_1 \otimes (ch)_2 = c_1h_1 \otimes c_2h_2, \tag{2.15}$$

$$\varepsilon_C(ch) = \varepsilon_C(c)\varepsilon_H(h), \gamma(ch) = \gamma(c)\beta(h). \tag{2.16}$$

A Hom-Doi-Koppinen datum is a triple $[(H, \beta), (A, \alpha), (C, \gamma)]$, where (H, β) is a monoidal Hom-Hopf algebra, (A, α) a right (H, β) -Hom-comodule algebra and (C, γ) a left (H, β) -Hom-module coalgebra. A Doi-Koppinen Hom-Hopf module (M, μ) is a left (A, α) -Hom-module which is also a right (C, γ) -Hom-comodule with the coaction structure ρ_M such that

$$\rho_M(ma) = m_{[0]}a_{[0]} \otimes m_{[1]}a_{[1]},$$

for all $m \in M$ and $a \in A$.

2.5. Hom-entwining structure

A (right-right) Hom-entwining structure is a $[(A, \alpha), (C, \gamma)]_\psi$ consisting of a monoidal Hom-algebra (A, α) , a monoidal Hom-coalgebra (C, γ) and a linear map $\psi : C \otimes A \rightarrow A \otimes C$ in $\tilde{\mathcal{H}}(\mathcal{M})$ satisfying the following conditions, for all $a, a' \in A, c \in C$,

$$(aa')_\psi \otimes \gamma(c)^\psi = a_\psi a'_\Psi \otimes \gamma(c^{\psi\Psi}), \quad (2.17)$$

$$\alpha^{-1}(a_\psi) \otimes c^{\psi_1} \otimes c^{\psi_2} = \alpha^{-1}(a)_{\psi\Psi} \otimes c_1^\Psi \otimes c_2^\psi, \quad (2.18)$$

$$1_{A\psi} \otimes c^\psi = 1_A \otimes c, \quad (2.19)$$

$$a_\psi \varepsilon_C(c^\psi) = a \varepsilon_C(c). \quad (2.20)$$

Here we use the following notation $\psi(c \otimes a) = a_\psi \otimes c^\psi$ for the so-called entwining map ψ . $\psi \in \tilde{\mathcal{H}}(\mathcal{M})$ means that the relation

$$\alpha(a)_\psi \otimes \gamma(c)^\psi = \alpha(a_\psi) \otimes \gamma(c^\psi). \quad (2.21)$$

If the map ψ occurs more than once in the same expression, then we use different sub- and superscripts: $\psi, \Psi, \psi_1, \psi_2, \dots$.

Given a Hom-entwining structure $[(A, \alpha), (C, \gamma)]_\psi$. A right-right $[(A, \alpha), (C, \gamma)]_\psi$ -entwined Hom-module is an object (M, μ) in $\tilde{\mathcal{H}}(\mathcal{M})$ is a right (A, α) -Hom-module, and a right (C, γ) -Hom-comodule with coaction $\rho_M : M \rightarrow M \otimes C, m \mapsto m_{[0]} \otimes m_{[1]}$ satisfying the condition, for any $m \in M, a \in A$,

$$\rho_M(ma) = m_{[0]} \alpha^{-1}(a)_\psi \otimes \gamma(m_{[1]}^\psi).$$

We use $\tilde{\mathcal{M}}_A^C(\psi)$ to denote the category of $[(A, \alpha), (C, \gamma)]_\psi$ -entwined Hom-modules together with the morphisms in which are both right (A, α) -linear and right (C, γ) -colinear.

3. Braiding on the Hom-category of Hom-entwined modules

Definition 3.1. We call $[(A, \alpha), (C, \gamma)]_\psi$ a monoidal Hom-entwining datum, if $[(A, \alpha), (C, \gamma)]_\psi$ is a Hom-entwining structure and A and C are monoidal Hom-bialgebras with the additional compatibility relations, for all $a \in A$ and $c, c' \in C$,

$$a_{1\psi} \otimes a_{2\Psi} \otimes c^\psi c'^\Psi = \Delta_A(a_\psi) \otimes (cc')^\psi, \quad (3.1)$$

$$\varepsilon_A(a)1_C = \varepsilon_A(a_\psi)1_C^\psi. \quad (3.2)$$

Proposition 3.2. Let $[(A, \alpha), (C, \gamma)]_\psi$ be a monoidal Hom-entwining structure. Then the tensor product of two Hom-entwined modules (M, μ) and (N, ν) is again a Hom-entwined module $(M \otimes N, \mu \otimes \nu)$ with the structure maps given by

$$\rho_{M \otimes N}(m \otimes n) = m_{[0]} \otimes n_{[0]} \otimes m_{[1]} n_{[1]}, \quad (3.3)$$

$$(m \otimes n)a = ma_1 \otimes na_2, \quad (3.4)$$

for all $m \in M, n \in N$ and $a \in A$. Thus the category $\tilde{\mathcal{M}}_A^C(\psi)$ is a Hom-category.

Proof. We show that $(M \otimes N, \mu \otimes \nu)$ is a Hom-entwined module. For all $m \in M, n \in N$ and $a \in A$, we have

$$\begin{aligned} \rho_{M \otimes N}((m \otimes n)a) &= (ma_1)_{[0]} \otimes (na_2)_{[0]} \otimes (ma_1)_{[1]} (na_2)_{[1]} \\ &= m_{[0]} \alpha^{-1}(a_1)_\psi \otimes n_{[0]} \alpha^{-1}(a_2)_\Psi \otimes \gamma(m_{[1]}^\psi) \gamma(n_{[1]}^\Psi) \\ &= m_{[0]} \alpha^{-1}(a)_{1\psi} \otimes n_{[0]} \alpha^{-1}(a)_{2\Psi} \otimes \gamma(m_{[1]}^\psi n_{[1]}^\Psi) \\ &= m_{[0]} \alpha^{-1}(a)_{\psi_1} \otimes n_{[0]} \alpha^{-1}(a)_{\psi_2} \otimes \gamma((m_{[1]} n_{[1]})^\psi) \text{ (by (3.1))} \\ &= (m_{[0]} \otimes n_{[0]}) \alpha^{-1}(a)_\psi \otimes \gamma((m_{[1]} n_{[1]})^\psi). \end{aligned}$$

Thus $(M \otimes N, \mu \otimes \nu)$ is an object of $\widetilde{\mathcal{M}}_A^C(\psi)$. Let (M, μ) , (N, ν) and (W, ς) be Hom-entwined modules. The isomorphisms

$$\begin{aligned} \widetilde{a}_{M,N,W} &: (M \otimes N) \otimes W \rightarrow M \otimes (N \otimes W) \\ &(m \otimes n) \otimes w \mapsto \mu(m) \otimes (\nu(n) \otimes \varsigma^{-1}(w)), \\ \widetilde{r}_M &: M \otimes k \rightarrow M, m \otimes x \mapsto x\mu(m), \\ \widetilde{l}_M &: k \otimes M \rightarrow M, x \otimes m \mapsto x\mu(m), \end{aligned}$$

obviously satisfy the pentagon axiom and the triangle axiom. We observe that (k, id) is an object of $\widetilde{\mathcal{M}}_A^C(\psi)$ via the trivial (A, α) -Hom-action and (C, γ) -Hom-coaction given by $xa = \varepsilon_A(a)x$ and $\rho_k = x \otimes 1_C$. It is clear that (k, id) is a unit object of $\widetilde{\mathcal{M}}_A^C(\psi)$. Hence $\widetilde{\mathcal{M}}_A^C(\psi)$ is a Hom-category. \square

Let $[(A, \alpha), (C, \gamma)]_\psi$ be a monoidal Hom-entwining datum. We know that a braiding on $\widetilde{\mathcal{M}}_A^C(\psi)$ is a natural family of isomorphisms

$$t_{M,N} : M \otimes N \rightarrow N \otimes M$$

in $\widetilde{\mathcal{M}}_A^C(\psi)$ such that, for all (M, μ) , (N, ν) and (W, ς) ,

$$(id_N \otimes t_{M,W}) \circ \widetilde{a}_{N,M,W} \circ (t_{M,N} \otimes id_W) \circ \widetilde{a}_{M,N,W}^{-1} = \widetilde{a}_{N,W,M} \circ t_{M,N \otimes W}, \quad (3.5)$$

$$\widetilde{a}_{P,M,N}^{-1} \circ t_{M,P} \otimes id_N \circ \widetilde{a}_{M,P,N}^{-1} \circ id_M \otimes t_{N,P} \circ \widetilde{a}_{M,N,P} = t_{M \otimes N, P}. \quad (3.6)$$

Consider a map $Q : C \otimes C \rightarrow A \otimes A$ in $\widetilde{\mathcal{H}}(\mathcal{M})$ with twisted convolution inverse R . We use the following notations $Q(c \otimes d) = Q^1(c \otimes d) \otimes Q^2(c \otimes d)$ and $R(c \otimes d) = R^1(c \otimes d) \otimes R^2(c \otimes d)$, for all $c, d \in C$. Thus we have

$$Q^1(c_2 \otimes d_2)R^1(c_1 \otimes d_1) \otimes Q^2(c_2 \otimes d_2)R^2(c_1 \otimes d_1) = \varepsilon_C(c)1_A \otimes \varepsilon_C(d)1_A, \quad (3.7)$$

$$R^1(c_2 \otimes d_2)Q^1(c_1 \otimes d_1) \otimes R^2(c_2 \otimes d_2)Q^2(c_1 \otimes d_1) = \varepsilon_C(c)1_A \otimes \varepsilon_C(d)1_A. \quad (3.8)$$

Consider two Hom-entwined modules (M, μ) and (N, ν) , we define

$$t_{M,N} : M \otimes N \rightarrow N \otimes M, m \otimes n \mapsto (n_{[0]} \otimes m_{[0]})Q(n_{[1]} \otimes m_{[1]}),$$

for all $m \in M, n \in N$. It follows from (3.7) and (3.8) that $t_{M,N}$ is bijective.

Example 3.3. Let $[(A, \alpha), (C, \gamma)]_\psi$ a Hom-entwining structure. The $(A \otimes C, \alpha \otimes \gamma)$ can become a Hom-entwined module with the right (A, α) -Hom-action and right (C, γ) -Hom-coaction given by

$$(a \otimes c)b = a\alpha^{-1}(b) \otimes \gamma(c), \quad (3.9)$$

$$\rho_{A \otimes C}(a \otimes c) = (\alpha^{-1}(a)_\psi \otimes c_1) \otimes \gamma(c_2^\psi), \quad (3.10)$$

for all $a \in A$ and $c \in C$.

Proof. It is straightforward to check that $(A \otimes C, \alpha \otimes \gamma)$ is a right (A, α) -Hom-module. Here we shall check that $(A \otimes C, \alpha \otimes \gamma)$ is also a right (C, γ) -Hom-comodule. In fact, for $a \in A$ and $c \in C$,

$$\begin{aligned} &(\alpha^{-1} \otimes \gamma^{-1})((a \otimes c)_{[0]}) \otimes \Delta_C((a \otimes c)_{[1]}) \\ &= \alpha^{-1}(\alpha^{-1}(a)_\psi) \otimes \gamma^{-1}(c_1) \otimes (\gamma(c_2^\psi) \otimes \gamma(c_2^\psi)) \\ &= \alpha^{-2}(a)_{\psi\Psi} \otimes \gamma^{-1}(c_1) \otimes (\gamma(c_2^\Psi) \otimes \gamma(c_2^\psi)) \\ &= \alpha^{-1}(\alpha^{-1}(a)_{\psi\Psi}) \otimes \gamma^{-1}(c_1) \otimes (\gamma(c_2^\Psi) \otimes \gamma(c_2^\psi)) \\ &= \alpha^{-1}(\alpha^{-1}(a)_{\psi\Psi}) \otimes c_{1_1} \otimes (\gamma(c_{1_2}^\Psi) \otimes c_2^\psi) \\ &= \alpha^{-1}(\alpha^{-1}(a)_\psi)_\Psi \otimes c_{1_1} \otimes (\gamma(c_{1_2}^\Psi) \otimes c_2^\psi) \\ &= (a \otimes c)_{[0][0]} \otimes ((a \otimes c)_{[0][1]} \otimes \gamma^{-1}((a \otimes c)_{[1]})), \end{aligned}$$

which proves that (2.7) holds. The other conditions can be checked straightforwardly. The compatibility can be proved as follows: for $a, b \in A, c \in C$,

$$\begin{aligned}
\rho_{A \otimes C}((b \otimes c)a) &= \rho_{A \otimes C}(b\alpha^{-1}(a) \otimes \gamma(c)) \\
&= (\alpha^{-1}(b\alpha^{-1}(a))_\psi \otimes \gamma(c_1)) \otimes \gamma(\gamma(c_2)^\psi) \\
&= ((\alpha^{-1}(b)\alpha^{-2}(a))_\psi \otimes \gamma(c_1)) \otimes \gamma(\gamma(c_2)^\psi) \\
&= (\alpha^{-1}(b)_\psi \alpha^{-2}(a)_\Psi \otimes \gamma(c_1)) \otimes \gamma(\gamma(c_2^{\psi\Psi})) \\
&= (\alpha^{-1}(b)_\psi \alpha^{-1}(\alpha^{-1}(a)_\Psi) \otimes \gamma(c_1)) \otimes \gamma(\gamma(c_2^\psi)^\Psi) \\
&= (\alpha^{-1}(b)_\psi \otimes c_1) \alpha^{-1}(a)_\Psi \otimes \gamma(\gamma(c_2^\psi)^\Psi)
\end{aligned}$$

as desired. \square

Lemma 3.4. *With notations as above, the map $t_{M,N}$ is right (A, α) -linear for all Hom-entwined modules (M, μ) and (N, ν) if and only if*

$$(b_{2\psi} \otimes b_{1\Psi})Q(c'^\psi \otimes c^\Psi) = Q(c' \otimes c)\Delta_A(b), \quad (3.11)$$

for all $b \in A$ and $c, c' \in C$.

Proof. Suppose that $t_{A \otimes C, A \otimes C}$ is (A, α) -linear. Then, for $a, a', b \in A$ and $c, c' \in C$, we have

$$t_{A \otimes C, A \otimes C}((a \otimes c) \otimes (a' \otimes c'))b = t_{A \otimes C, A \otimes C}((a \otimes c) \otimes (a' \otimes c'))b. \quad (3.12)$$

Since

$$\begin{aligned}
\text{LHS} &= t_{A \otimes C, A \otimes C}((a \otimes c)b_1 \otimes (a' \otimes c')b_2) \\
&= t_{A \otimes C, A \otimes C}((a\alpha^{-1}(b_1) \otimes \gamma(c)) \otimes (a'\alpha^{-1}(b_2) \otimes \gamma(c'))) \\
&= (\alpha^{-1}(a'\alpha^{-1}(b_2))_\psi \otimes \gamma(c'_1)Q^1(\gamma(\gamma(c')_2^\psi) \otimes \gamma(\gamma(c)_2^\Psi))) \\
&\quad \otimes (\alpha^{-1}(a\alpha^{-1}(b_1))_\Psi \otimes \gamma(c)_1)Q^2(\gamma(\gamma(c')_2^\psi) \otimes \gamma(\gamma(c)_2^\Psi)) \\
&= (\alpha^{-1}(a'\alpha^{-1}(b_2))_\psi \alpha^{-1}(Q^1(\gamma(\gamma(c')_2^\psi) \otimes \gamma(\gamma(c)_2^\Psi))) \otimes \gamma(\gamma(c')_1)) \\
&\quad \otimes (\alpha^{-1}(a\alpha^{-1}(b_1))_\Psi \alpha^{-1}(Q^2(\gamma(\gamma(c')_2^\psi) \otimes \gamma(\gamma(c)_2^\Psi))) \otimes \gamma(\gamma(c)_1))
\end{aligned}$$

and

$$\begin{aligned}
\text{RHS} &= (((\alpha^{-1}(a')_\psi \otimes c'_1) \otimes (\alpha^{-1}(a)_\Psi \otimes c_1))Q(\gamma(c'_2^\psi) \otimes \gamma(c_2^\Psi)))b \\
&= ((\alpha^{-1}(a')_\psi \otimes c'_1)Q^1(\gamma(c'_2^\psi) \otimes \gamma(c_2^\Psi)))b_1 \\
&\quad \otimes ((\alpha^{-1}(a)_\Psi \otimes c_1)Q^2(\gamma(c'_2^\psi) \otimes \gamma(c_2^\Psi)))b_2 \\
&= ((\alpha^{-1}(a')_\psi \alpha^{-1}(Q^1(\gamma(c'_2^\psi) \otimes \gamma(c_2^\Psi))))\alpha^{-1}(b_1) \otimes \gamma^2(c'_1)) \\
&\quad \otimes ((\alpha^{-1}(a)_\Psi \alpha^{-1}(Q^2(\gamma(c'_2^\psi) \otimes \gamma(c_2^\Psi))))\alpha^{-1}(b_2) \otimes \gamma^2(c_1)),
\end{aligned}$$

we have

$$\begin{aligned}
&(\alpha^{-1}(a'\alpha^{-1}(b_2))_\psi \alpha^{-1}(Q^1(\gamma(\gamma(c')_2^\psi) \otimes \gamma(\gamma(c)_2^\Psi))) \otimes \gamma(\gamma(c')_1)) \\
&\quad \otimes (\alpha^{-1}(a\alpha^{-1}(b_1))_\Psi \alpha^{-1}(Q^2(\gamma(\gamma(c')_2^\psi) \otimes \gamma(\gamma(c)_2^\Psi))) \otimes \gamma(\gamma(c)_1)) \\
&= ((\alpha^{-1}(a')_\psi \alpha^{-1}(Q^1(\gamma(c'_2^\psi) \otimes \gamma(c_2^\Psi))))\alpha^{-1}(b_1) \otimes \gamma^2(c'_1)) \\
&\quad \otimes ((\alpha^{-1}(a)_\Psi \alpha^{-1}(Q^2(\gamma(c'_2^\psi) \otimes \gamma(c_2^\Psi))))\alpha^{-1}(b_2) \otimes \gamma^2(c_1)).
\end{aligned}$$

By taking $a = a' = 1_A$ in the above equality and then applying $id_A \otimes \varepsilon_C \otimes id_A \otimes \varepsilon_C$ to both sides, we can get

$$(b_{2\psi} \otimes b_{1\Psi})Q(c'^\psi \otimes c^\Psi) = Q(c' \otimes c)\Delta_A(b). \quad (3.13)$$

Conversely, suppose that (3.11) holds, and consider two Hom-entwined modules (M, μ) and (N, ν) . For all $m \in M, n \in N$ and $a \in A$, we have

$$\begin{aligned}
t_{M,N}((m \otimes n)a) &= t_{M,N}(ma_1 \otimes na_2) \\
&= ((na_2)_{[0]} \otimes (ma_1)_{[0]})Q((na_2)_{[1]} \otimes (ma_1)_{[1]}) \\
&= (n_{[0]}\alpha^{-1}(a_2)_\psi \otimes m_{[0]}\alpha^{-1}(a_1)_\Psi)Q(\gamma(n_{[1]}^\psi) \otimes \gamma(m_{[1]}^\Psi)) \\
&= (n_{[0]}\alpha^{-1}(a)_{2\psi} \otimes m_{[0]}\alpha^{-1}(a)_{1\Psi})Q(\gamma(n_{[1]}^\psi) \otimes \gamma(m_{[1]}^\Psi)) \\
&= \nu(n_{[0]})\alpha^{-1}(a_{2\psi}Q^1(\gamma(n_{[1]}^\psi) \otimes \gamma(m_{[1]}^\Psi))) \\
&\quad \otimes \mu(m_{[0]})\alpha^{-1}(a_{1\Psi}Q^2(\gamma(n_{[1]}^\psi) \otimes \gamma(m_{[1]}^\Psi))) \\
&= \nu(n_{[0]})\alpha^{-1}(Q^1(\gamma(n_{[1]}) \otimes \gamma(m_{[1]}))a_1) \\
&\quad \otimes \mu(m_{[0]})\alpha^{-1}(Q^2(\gamma(n_{[1]}) \otimes \gamma(m_{[1]}))a_2) \\
&= (n_{[0]}\alpha^{-1}(Q^1(\gamma(n_{[1]}) \otimes \gamma(m_{[1]}))))a_1 \\
&\quad \otimes (m_{[0]}\alpha^{-1}(Q^2(\gamma(n_{[1]}) \otimes \gamma(m_{[1]}))))a_2 \\
&= (n_{[0]}Q^1(n_{[1]} \otimes m_{[1]}))a_1 \otimes (m_{[0]}Q^2(n_{[1]} \otimes m_{[1]}))a_2 \\
&= t_{M,N}(m \otimes n)a,
\end{aligned}$$

which follows that $t_{M,N}$ is (A, α) -linear. \square

Lemma 3.5. *With notations as above, the map $t_{M,N}$ is right (C, γ) -colinear for all Hom-entwined modules (M, μ) and (N, ν) if and only if*

$$Q^1(c'_2 \otimes c_2)_\psi \otimes Q^2(c'_2 \otimes c_2)_\Psi \otimes c_1^\psi c_1^\Psi = Q^1(c'_1 \otimes c_1) \otimes Q^2(c'_1 \otimes c_1) \otimes c_2 c'_2, \quad (3.14)$$

for all $c, c' \in C$.

Proof. Suppose that $t_{A \otimes C, A \otimes C}$ is (C, γ) -colinear. Then, for $c, c' \in C$, we have

$$\begin{aligned}
&(\alpha^{-1}(Q^1(\gamma(c'_2) \otimes \gamma(c_2))_\psi) \otimes \gamma(c'_{11})) \\
&\quad \otimes (\alpha^{-1}(Q^2(\gamma(c'_2) \otimes \gamma(c_2))_\Psi) \otimes \gamma(c_{11})) \otimes \gamma^2(c'_{12})^\psi \gamma^2(c_{12})^\Psi \\
&= (Q^1(\gamma(c'_{12}) \otimes \gamma(c_{12})) \otimes \gamma(c'_{11})) \otimes (Q^2(\gamma(c'_{12}) \otimes \gamma(c_{12})) \otimes \gamma(c_{11})) \otimes \gamma(c_2)\gamma(c'_2).
\end{aligned}$$

Applying $id_A \otimes \varepsilon_C \otimes id_A \otimes \varepsilon_C \otimes id_C$ to both sides, we can have (3.14).

Conversely, assume that (3.14) holds. Take two Hom-entwined modules (M, μ) and (N, ν) . Then, for $m \in M$, and $n \in N$, we have

$$\begin{aligned}
\rho_{M \otimes N}(t_{M,N}(m \otimes n)) &= \rho_{M \otimes N}((n_{[0]} \otimes m_{[0]})Q(n_{[1]} \otimes m_{[1]})) \\
&= (n_{[0][0]}\alpha^{-1}(Q^1(n_{[1]} \otimes m_{[1]}))_\psi \\
&\quad \otimes m_{[0][0]}\alpha^{-1}(Q^2(n_{[1]} \otimes m_{[1]}))_\Psi) \otimes \gamma(n_{[0][1]}^\psi)\gamma(m_{[0][1]}^\Psi) \\
&= (\nu^{-1}(n_{[0]})\alpha^{-1}(Q^1(\gamma(n_{[1]2}) \otimes \gamma(m_{[1]2})))_\psi \\
&\quad \otimes \mu^{-1}(m_{[0]})\alpha^{-1}(Q^2(\gamma(n_{[1]2}) \otimes \gamma(m_{[1]2})))_\Psi) \otimes \gamma(n_{[1]1}^\psi)\gamma(m_{[1]1}^\Psi) \\
&= ((\nu^{-1}(n_{[0]})\alpha^{-1}(Q^1(\gamma(n_{[1]2}) \otimes \gamma(m_{[1]2})))_\psi \\
&\quad \otimes (\mu^{-1}(m_{[0]})\alpha^{-1}(Q^2(\gamma(n_{[1]2}) \otimes \gamma(m_{[1]2})))_\Psi)) \otimes \gamma(n_{[1]1}^\psi)\gamma(m_{[1]1}^\Psi) \\
&= ((\nu^{-1}(n_{[0]})\alpha^{-1}(Q^1(\gamma(n_{[1]1}) \otimes \gamma(m_{[1]1})))) \\
&\quad \otimes (\mu^{-1}(m_{[0]})\alpha^{-1}(Q^2(\gamma(n_{[1]1}) \otimes \gamma(m_{[1]1})))) \otimes \gamma(n_{[1]1}^\psi)\gamma(m_{[1]1}^\Psi) \\
&= (n_{[0][0]}\alpha^{-1}(Q^1(\gamma(n_{[0][1]}) \otimes \gamma(m_{[0][1]}))) \\
&\quad \otimes (m_{[0][0]}\alpha^{-1}(Q^2(\gamma(n_{[0][1]}) \otimes \gamma(m_{[0][1]})))) \otimes m_{[1]}n_{[1]}
\end{aligned}$$

$$\begin{aligned}
&= ((n_{[0][0]}Q^1(n_{[0][1]} \otimes m_{[0][1]})) \otimes (m_{[0][0]}Q^2(n_{[0][1]} \otimes m_{[0][1]}))) \otimes m_{[1]}n_{[1]} \\
&= (t_{M,N} \otimes id_C)(\rho_{M \otimes N}(m \otimes n)),
\end{aligned}$$

which follows that $(t_{M,N})$ is (C, γ) -colinear. \square

Lemma 3.6. *With notations as above, (3.5) holds for all Hom entwined modules (M, μ) , (N, ν) and (W, ς) if and only if*

$$(\Delta_A \otimes id_A)Q(c'c'' \otimes c) = Q^1(c' \otimes c_2) \otimes Q^1(c'' \otimes c_1^\psi) \otimes Q^2(c' \otimes c_2)_\psi Q^2(c'' \otimes c_1^\psi), \quad (3.15)$$

for all $c, c', c'' \in C$.

Proof. Suppose that (3.5) holds. We take $M = N = W = (A \otimes C, \alpha \otimes \gamma)$. For $c, c', c'' \in C$, on the one hand, we have

$$\begin{aligned}
&(id_N \otimes t_{M,W}) \circ \tilde{a}_{N,M,W} \circ (t_{M,N} \otimes id_W) \circ \tilde{a}_{M,N,W}^{-1}((1 \otimes c) \otimes ((1 \otimes c') \otimes (1 \otimes c''))) \\
&= (\alpha(Q^1(\gamma(c'_2) \otimes c_2)) \otimes \gamma^2(c'_1)) \otimes ((Q^1(\gamma(c''_2) \otimes \gamma(c_{1_2}^\psi)) \otimes \gamma(c''_1)) \\
&\quad \otimes \alpha^{-1}(Q^2(\gamma(c'_2) \otimes c_2)_\psi Q^2(\gamma(c''_2) \otimes \gamma(c_{1_2}^\psi)))) \otimes \gamma(c_{1_1})) \\
&= (\alpha(Q^1(\gamma(c'_2) \otimes \gamma(c_{2_2}))) \otimes \gamma^2(c'_1)) \otimes ((Q^1(\gamma(c''_2) \otimes \gamma(c_{2_1}^\psi)) \otimes \gamma(c''_1)) \\
&\quad \otimes \alpha^{-1}(Q^2(\gamma(c'_2) \otimes \gamma(c_{2_2}))_\psi Q^2(\gamma(c''_2) \otimes \gamma(c_{2_1}^\psi)))) \otimes c_1).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\tilde{a}_{N,W,M} \circ t_{M,N \otimes W}((1 \otimes c) \otimes ((1 \otimes c') \otimes (1 \otimes c''))) \\
&= \alpha(Q^1(\gamma(c'_2 c''_2) \otimes \gamma(c_2))_1) \otimes \gamma^2(c'_1) \\
&\quad \otimes ((Q^1(\gamma(c'_2 c''_2) \otimes \gamma(c_2))_2 \otimes \gamma(c''_1)) \otimes (\alpha^{-1}(Q^2(\gamma(c'_2 c''_2) \otimes \gamma(c_2))) \otimes c_1)) \\
&= (\alpha(Q^1(\gamma(c'_2) \otimes \gamma(c_{2_2}))) \otimes \gamma^2(c'_1)) \otimes ((Q^1(\gamma(c''_2) \otimes \gamma(c_{2_1}^\psi)) \otimes \gamma(c''_1)) \\
&\quad \otimes \alpha^{-1}(Q^2(\gamma(c'_2) \otimes \gamma(c_{2_2}))_\psi Q^2(\gamma(c''_2) \otimes \gamma(c_{2_1}^\psi)))) \otimes c_1).
\end{aligned}$$

Applying $id_A \otimes \varepsilon_C \otimes id_A \otimes \varepsilon_C \otimes id_A \otimes \varepsilon_C$ to both sides, we get (3.15).

Conversely, if (3.15) holds. Let (M, μ) , (N, ν) and (W, ς) be Hom-entwined modules. We easily compute that

$$\begin{aligned}
&(id_N \otimes t_{M,W}) \circ \tilde{a}_{N,M,W} \circ (t_{M,N} \otimes id_W) \circ \tilde{a}_{M,N,W}^{-1}(m \otimes (n \otimes w)) \\
&= \nu(n_{[0]})\alpha(Q^1(n_{[1]} \otimes \gamma^{-1}(m_{[1]}))) \otimes (w_{[0]}Q^1(w_{[1]} \otimes \gamma((\mu^{-1}(m_{[0]})_{[1]})^\psi)) \\
&\quad \otimes (\mu^{-1}(m_{[0]})_{[0]}\alpha^{-1}(Q^2(n_{[1]} \otimes \gamma^{-1}(m_{[1]})))_\psi Q^2(w_{[1]} \otimes \gamma((\mu^{-1}(m_{[0]})_{[1]})^\psi)))) \\
&= \nu(n_{[0]})\alpha(Q^1(n_{[1]} \otimes \gamma^{-1}(m_{[1]}))) \otimes (w_{[0]}Q^1(w_{[1]} \otimes \gamma((\gamma^{-1}(m_{[0]})_{[1]})^\psi))) \\
&\quad \otimes (\mu^{-1}(m_{[0]})_{[0]}\alpha^{-1}(Q^2(n_{[1]} \otimes \gamma^{-1}(m_{[1]})))_\psi Q^2(w_{[1]} \otimes \gamma((\gamma^{-1}(m_{[0]})_{[1]})^\psi)))) \\
&= \nu(n_{[0]})\alpha(Q^1(n_{[1]} \otimes \gamma^{-1}(m_{[1]}))) \otimes (w_{[0]}Q^1(w_{[1]} \otimes (m_{[0]})_{[1]}^\psi) \\
&\quad \otimes (\mu^{-1}(m_{[0]})_{[0]}\alpha^{-1}(Q^2(n_{[1]} \otimes \gamma^{-1}(m_{[1]})))_\psi) Q^2(w_{[1]} \otimes (m_{[0]})_{[1]}^\psi)) \\
&= \nu(n_{[0]})\alpha(Q^1(n_{[1]} \otimes m_{[1]_2})) \otimes (w_{[0]}Q^1(w_{[1]} \otimes (m_{[1]_1})^\psi) \\
&\quad \otimes (\mu^{-2}(m_{[0]})\alpha^{-1}(Q^2(n_{[1]} \otimes m_{[1]_2})_\psi) Q^2(w_{[1]} \otimes (m_{[1]_1})^\psi)) \\
&= \nu(n_{[0]})\alpha(Q^1(n_{[1]} \otimes m_{[1]_2})) \otimes (w_{[0]}Q^1(w_{[1]} \otimes (m_{[1]_1})^\psi) \\
&\quad \otimes \mu^{-1}(m_{[0]})\alpha^{-1}(Q^2(n_{[1]} \otimes m_{[1]_2})_\psi)\alpha^{-1}(Q^2(w_{[1]} \otimes (m_{[1]_1})^\psi))) \\
&= \tilde{a}_{N,W,M} \circ t_{M,N \otimes W}(m \otimes (n \otimes w)),
\end{aligned}$$

which proves that (3.5) holds. \square

The proof of The next lemma is similar to the proof of Lemma 3.6, so we omit it.

Lemma 3.7. *With notations as above, (3.6) holds for all Hom entwined modules (M, μ) , (N, ν) and (W, ς) if and only if*

$$(id_A \otimes \Delta_A)Q(c \otimes c'c'') = Q^1(c_2 \otimes c'')_{\psi} Q^1(c_1^{\psi} \otimes c') \otimes Q^2(c_1^{\psi} \otimes c') \otimes Q^2(c_2 \otimes c''), \quad (3.16)$$

for all $c, c', c'' \in C$.

We summarize our results as follows:

Theorem 3.8. *Let $[(A, \alpha), (C, \gamma)]_{\psi}$ a monoidal Hom-entwining datum, and $Q : C \otimes C \rightarrow A \otimes A$ a twisted convolution invertible map in $\tilde{\mathcal{H}}(\mathcal{M})$. Then the family of maps*

$$t_{M,N} : M \otimes N \rightarrow N \otimes M, m \otimes n \mapsto (n_{[0]} \otimes m_{[0]})Q(n_{[1]} \otimes m_{[1]})$$

defines a braiding on the category of Hom-entwined modules $\tilde{\mathcal{M}}_A^C(\psi)$ if and only if Q satisfies Equations (3.11) and (3.14)-(3.16).

Now, we shall apply Theorem 3.8 to Doi-Koppinen Hom-Hopf modules. Given a Hom-Doi-Koppinen datum $[(H, \beta), (A, \alpha), (C, \gamma)]$, we have a Hom-entwining datum $[(A, \alpha), (C, \gamma)]_{\psi}$ with ψ given by

$$\psi : C \otimes A \rightarrow A \otimes C, c \otimes a \mapsto \alpha(a_{[0]}) \otimes \gamma^{-1}(c)a_{[1]} = a_{\psi} \otimes c^{\psi}. \quad (3.17)$$

The Hom-category $\tilde{\mathcal{M}}_A^C(\psi)$ of Hom-entwined modules associated to the induced Hom-entwining datum $[(A, \alpha), (C, \gamma)]_{\psi}$ is denoted by $\tilde{\mathcal{M}}(H)_A^C$.

A Hom-Doi-Koppinen datum $[(H, \beta), (A, \alpha), (C, \gamma)]$ is called a monoidal Hom-Doi-Koppinen datum, if it satisfies the following condition,

$$a_{1[0]} \otimes a_{2[0]} \otimes (ca_{1[1]})(c'a_{2[1]}) = a_{[0]1} \otimes a_{[0]2} \otimes (cc')a_{[1]}, \quad (3.18)$$

for all $a \in A$ and $c \in C$.

From Theorem 3.8, we have the following result.

Corollary 3.9. *Let $[(H, \beta), (A, \alpha), (C, \gamma)]$ be a monoidal Hom-Doi-Koppinen datum, and $Q : C \otimes C \rightarrow A \otimes A$ a twisted convolution invertible map in $\tilde{\mathcal{H}}(\mathcal{M})$. Then the family of maps*

$$t_{M,N} : M \otimes N \rightarrow N \otimes M, m \otimes n \mapsto (n_{[0]} \otimes m_{[0]})Q(n_{[1]} \otimes m_{[1]})$$

defines a braiding on the category of Doi-Koppinen Hom-Hopf modules $\tilde{\mathcal{M}}(H)_A^C$ if and only if Q satisfies the following equations, for any $b \in A$ and $c, c', c'' \in C$,

- (1) $(\alpha(b_{2[0]}) \otimes \alpha(b_{1[0]}))Q(c'b_{2[1]} \otimes cb_{1[1]}) = Q(\gamma(c') \otimes \gamma(c))\Delta_A(b)$,
- (2)

$$\begin{aligned} & \alpha(Q^1(c'_2 \otimes c_2)_{[0]}) \otimes \alpha(Q^2(c'_2 \otimes c_2)_{[0]}) \\ & \quad \otimes (\gamma^{-1}(c'_1)Q^1(c'_2 \otimes c_2)_{[1]})(\gamma^{-1}(c_1)Q^2(c'_2 \otimes c_2)_{[1]}) \\ & = Q^1(c'_1 \otimes c_1) \otimes Q^2(c'_1 \otimes c_1) \otimes c_2c'_2, \end{aligned}$$

(3)

$$\begin{aligned} (\Delta_A \otimes id_A)Q(c'c'' \otimes c) & = Q^1(c' \otimes c_2) \otimes Q^1(c'' \otimes \gamma^{-1}(c_1)Q^2(c' \otimes c_2)_{[1]}) \\ & \quad \otimes \alpha(Q^2(c' \otimes c_2)_{[0]})Q^2(c'' \otimes \gamma^{-1}(c_1)Q^2(c' \otimes c_2)_{[1]}), \end{aligned}$$

(4)

$$\begin{aligned} (id_A \otimes \Delta_A)Q(c \otimes c'c'') & = \alpha(Q^1(c_2 \otimes c'')_{[0]})Q^1(\gamma^{-1}(c_1)Q^1(c_2 \otimes c'')_{[1]} \otimes c') \\ & \quad \otimes Q^2(\gamma^{-1}(c_1)Q^1(c_2 \otimes c'')_{[1]} \otimes c') \otimes Q^2(c_2 \otimes c''). \end{aligned}$$

4. Hom-coalgebra cleft extensions for Hom-entwining structures

Let (A, α) be a object of $\widetilde{\mathcal{M}}_A^C(\psi)$ with the Hom-coaction ρ_A . For $(M, \mu) \in \widetilde{\mathcal{M}}_A^C(\psi)$, The Hom-invariants of C on M are the set

$$M^{coC} = \{m \in M \mid \rho_M(m) = \mu^{-2}(m)1_{[0]} \otimes \gamma(1_{[1]})\}.$$

Specially, we have $A^{coC} = \{a \in A \mid \rho_A(a) = \alpha^{-2}(a)1_{[0]} \otimes \gamma(1_{[1]})\}$. For $m \in M^{coC}$, it follows that $\mu(a) \in M^{coC}$. We use $\mu|_{M^{coC}}$ for denoting the restriction map of μ on M^{coC} .

Lemma 4.1. For $(A, \alpha), (M, \mu)$ in $\widetilde{\mathcal{M}}_A^C(\psi)$, we have

- (1) $(A^{coC}, \alpha|_{A^{coC}}, 1)$ is a Hom-algebra.
- (2) $(M^{coC}, \mu|_{M^{coC}})$ is a right $(A^{coC}, \alpha|_{A^{coC}})$ -Hom-module.

Proof. Straightforward. □

Let us put $\text{Hom}^C(C, A)$ consisting of right (C, γ) -colinear morphisms $f : C \rightarrow A$, that is, $f(c)_{[0]} \otimes f(c)_{[1]} = f(c_{[0]}) \otimes c_{[1]}$, for $c \in C$ and $f \circ \gamma = \alpha \circ f$.

Lemma 4.2. $\text{Hom}^C(C, A)$ is an associative algebra with the unit $\varepsilon_C 1_A$ and multiplication

$$(f * g)(c) = f(c_1)g(c_2),$$

for $f, g \in \text{Hom}^C(C, A)$ and $c \in C$.

Proof. Straightforward. □

By $\text{Reg}(C, A)$ we denote the set of morphisms $\omega \in \text{Hom}^C(C, A)$ which are invertible under the convolution $*$ in Lemma 4.2.

Definition 4.3. We say that $(A^{coC}, \alpha|_{A^{coC}}) \hookrightarrow (A, \alpha)$ is a (C, γ) -Hom-cleft extension, if there exists a morphism $\omega \in \text{Reg}(C, A)$.

Proposition 4.4. If $(A^{coC}, \alpha|_{A^{coC}}) \hookrightarrow (A, \alpha)$ is a (C, γ) -Hom-cleft extension, we have

$$\omega^{-1}(c_2)_\psi \otimes c_1^\psi = \alpha^{-2}(\omega^{-1}(c))1_{[0]} \otimes \gamma(1_{[1]}), \quad (4.1)$$

for all $c \in C$.

Proof. Since $(A, \alpha) \in \widetilde{\mathcal{M}}_A^C(\psi)$, the Hom-coaction can be written as $\rho_A(a) = 1_{[0]}\alpha^{-2}(a)_\psi \otimes \gamma(1_{[1]}^\psi)$. Then we have, for any $c \in C$,

$$\begin{aligned} \varepsilon_C(c)\alpha(1_{[0]}) \otimes 1_{[1]} &= 1_{[0]}\psi(1_{[1]}) \otimes \omega(c_1)\omega^{-1}(c_2) \\ &= 1_{[0]}(\omega(c_1)\omega^{-1}(c_2))_\psi \otimes 1_{[1]}^\psi \\ &= \alpha(1_{[0]})(\omega(c_1)\omega^{-1}(c_2))_\psi \otimes \gamma(1_{[1]}^\psi) \\ &= \alpha(1_{[0]})(\omega(c_1)_\psi\omega^{-1}(c_2)_\Psi) \otimes \gamma(1_{[1]}^{\psi\Psi}) \\ &= (1_{[0]}\omega(c_1)_\psi)\alpha(\omega^{-1}(c_2)_\Psi) \otimes \gamma(1_{[1]}^{\psi\Psi}) \\ &= \alpha^2(\omega(c_1)_{[0]})\alpha(\omega^{-1}(c_2)_\Psi) \otimes \gamma(\gamma(\omega(c_1)_{[1]})^\Psi) \\ &= \alpha^2(\omega(c_1))\alpha(\omega^{-1}(c_2)_\Psi) \otimes \gamma(\gamma(c_1)_2^\Psi) \\ &= \alpha(\omega(c_1))\alpha(\omega^{-1}(\gamma(c_2)_2)_\Psi) \otimes \gamma(\gamma(c_2)_1^\Psi), \end{aligned}$$

which implies that Eq (4.1) holds. □

Lemma 4.5. Assume that $(A^{coC}, \alpha|_{A^{coC}}) \hookrightarrow (A, \alpha)$ is a (C, γ) -Hom-cleft extension via and $(M, \mu) \in \widetilde{\mathcal{M}}_A^C(\psi)$. Then, for any $m \in M$, $m_{[0]}\omega^{-1}(m_{[1]}) \in M^{coC}$. As a consequence, if $M = A$, we have $a_{[0]}\omega^{-1}(a_{[1]}) \in A^{coC}$

Proof. We compute

$$\begin{aligned}
\rho_M(m_{[0]}\omega^{-1}(m_{[1]})) &= m_{[0][0]}\alpha^{-1}(\omega^{-1}(m_{[1]}))_\psi \otimes \gamma(m_{[0][1]}^\psi) \\
&= \mu^{-1}(m_{[0]})\alpha^{-1}(\omega^{-1}(\gamma(m_{[1]2})))_\psi \otimes \gamma(m_{[1]1}^\psi) \\
&= \mu^{-1}(m_{[0]})\alpha^{-1}(\omega^{-1}(\gamma(m_{[1]2})_\psi) \otimes (\gamma(m_{[1]1})_1)^\psi \\
&= \mu^{-1}(m_{[0]})(\alpha^{-2}(\omega^{-1}(m_{[1]}))\alpha^{-1}(1_{[0]})) \otimes \gamma(1_{[1]}) \\
&= (\mu^{-2}(m_{[0]})\alpha^{-2}(\omega^{-1}(m_{[1]})))1_{[0]} \otimes \gamma(1_{[1]}) \\
&= \mu^{-2}(m_{[0]}\omega^{-1}(m_{[1]}))1_{[0]} \otimes \gamma(1_{[1]}).
\end{aligned}$$

Hence $m_{[0]}\omega^{-1}(m_{[1]}) \in M^{coC}$. \square

Theorem 4.6. *Suppose that $(A^{coC}, \alpha|_{A^{coC}}) \hookrightarrow (A, \alpha)$ is a (C, γ) -Hom-cleft extension via ω . For $(M, \mu) \in \mathcal{M}_A^C(\psi)$, then $(M, \mu) \cong (M^{coC} \otimes C, \mu|_{M^{coC} \otimes C} \otimes \gamma)$ as right (C, γ) -Hom-comodules, where the (C, γ) -Hom-coaction on $(M^{coC} \otimes C, \mu|_{M^{coC} \otimes C} \otimes \gamma)$ is*

$$\rho_{M^{coC} \otimes C}(m \otimes c) = (\mu^{-1}(m) \otimes c_1) \otimes \gamma(c_2).$$

In particular, if we consider $M = A$, we have $(A, \alpha) \cong (A^{coC} \otimes C, \alpha|_{A^{coC} \otimes C} \otimes \gamma)$ as both right (C, γ) -Hom-comodules and left $(A^{coC}, \alpha|_{A^{coC}})$ -Hom-modules, where the $(A^{coC}, \alpha|_{A^{coC}})$ -Hom-action on $A^{coC} \otimes C$ defined by $b \cdot (a \otimes c) = \alpha^{-1}(b)a \otimes \gamma(c)$, for all $a, b \in A^{coC}$ and $c \in C$.

Proof. We define $\Theta_M : (M^{coC} \otimes C, \mu|_{M^{coC} \otimes C} \otimes \gamma) \rightarrow (M, \mu)$ by $\Theta_M(m \otimes c) = m\omega(c)$ and $\Omega_M : (M, \mu) \rightarrow (M^{coC} \otimes C, \mu|_{M^{coC} \otimes C} \otimes \gamma)$ by $\Omega_M(m) = m_{[0][0]}\omega^{-1}(m_{[0][1]}) \otimes m_{[1]}$. For $m \in M$, we have

$$\begin{aligned}
\Theta_M \circ \Omega_M(m) &= (m_{[0][0]}\omega^{-1}(m_{[0][1]}))\omega(m_{[1]}) \\
&= (\mu^{-1}(m_{[0]})\omega^{-1}(m_{[1]1}))\omega(\gamma(m_{[1]2})) \\
&= (\mu^{-1}(m_{[0]})\omega^{-1}(m_{[1]1}))\alpha(\omega(m_{[1]2})) \\
&= m_{[0]}(\omega^{-1}(m_{[1]1})\omega(m_{[1]2})) \\
&= m_{[0]}\varepsilon_C(m_{[1]})1_A = m,
\end{aligned}$$

which follows that $\Theta_M \circ \Omega_M = id$. Next, we check that $\Omega_M \circ \Theta_M = id$ holds. In fact, for any $m \in M^{coC}$ and $c \in C$, we compute

$$\begin{aligned}
\Omega_M \circ \Theta_M(m \otimes c) &= (m\omega(c))_{[0][0]}\omega^{-1}((m\omega(c))_{[0][1]}) \otimes (m\omega(c))_{[1]} \\
&= (m_{[0][0]}\alpha^{-1}(\alpha^{-1}(\omega(c))_\psi)_\Psi)\omega^{-1}(\gamma(m_{[0][1]}^\Psi)) \otimes \gamma(m_{[1]}^\psi) \\
&= (m_{[0][0]}\alpha^{-1}(\alpha^{-1}(\omega(c))_\psi)_\Psi)\omega^{-1}(\gamma(m_{[0][1]}^\Psi)) \otimes \gamma(m_{[1]}^\psi) \\
&= (\mu^{-1}(m_{[0]})\alpha^{-1}(\alpha^{-1}(\omega(c))_\psi)_\Psi)\omega^{-1}(\gamma(m_{[1]1}^\Psi)) \otimes \gamma(\gamma(m_{[1]2}^\psi)) \\
&= (\mu^{-1}(\mu^{-2}(m)1_{[0]})\alpha^{-1}(\alpha^{-1}(\omega(c))_\psi)_\Psi)\omega^{-1}(\gamma^2(1_{[1]1}^\Psi)) \otimes \gamma(\gamma^2(1_{[1]2}^\psi)) \\
&= (\mu^{-2}(m)(\alpha^{-1}(1_{[0]})\alpha^{-2}(\alpha^{-1}(\omega(c))_\psi)_\Psi))\omega^{-1}(\gamma^2(1_{[1]1}^\Psi)) \otimes \gamma(\gamma^2(1_{[1]2}^\psi)) \\
&= (\mu^{-2}(m)(\alpha^{-1}(1_{[0]})\alpha^{-3}(\omega(c)_\psi)))\omega^{-1}(\gamma^2(1_{[1]1}^\Psi)) \otimes \gamma(\gamma^2(1_{[1]2}^\psi)) \\
&= (\mu^{-2}(m)\alpha^{-3}(1_{[0]}\omega(c)_\psi))\omega^{-1}(1_{[1]1}^\Psi) \otimes \gamma(1_{[1]2}^\psi) \\
&= (\mu^{-2}(m)\alpha^{-1}(\omega(c)_{[0]}))\omega^{-1}(\gamma(\omega(c)_{[1]1})) \otimes \gamma(\gamma(\omega(c)_{[1]2})) \\
&= (\mu^{-2}(m)\alpha^{-1}(\omega(c_1)))\omega^{-1}(\gamma(c_2)_1) \otimes \gamma(\gamma(c_2)_2) \\
&= (\mu^{-2}(m)\alpha^{-1}(\omega(c_1)))\omega^{-1}(\gamma(c_{21})) \otimes \gamma(\gamma(c_{22}))
\end{aligned}$$

$$\begin{aligned}
&= (\mu^{-2}(m)\omega(c_{1_1}))\omega^{-1}(\gamma(c_{1_2})) \otimes \gamma(c_2) \\
&= \mu^{-1}(m)(\omega(c_{1_1})\omega^{-1}(c_{1_2})) \otimes \gamma(c_2) \\
&= \mu^{-1}(m)1_A \otimes c \\
&= m \otimes c,
\end{aligned}$$

as desired. \square

Let $(A^{coC}, \alpha|_{A^{coC}}) \hookrightarrow (A, \alpha)$ be a (C, γ) -Hom-cleft extension via ω . From Theorem 4.6, we have that Ω_A is an isomorphism of monoidal Hom-algebras, where the monoidal Hom-algebra structure on $A^{coC} \otimes C$ can be induced by Ω_A :

$$1_{A^{coC} \times C} = \Omega_A(1_A), \tilde{m}_{A^{coC} \times C} = \Omega_A \circ m_A \circ (\Omega_A^{-1} \otimes \Omega_A^{-1}).$$

The induced monoidal Hom-algebras on $A^{coC} \otimes C$ is called a crossed product Hom-algebra of A^{coC} and C , and denoted by $A^{coC} \times C$.

Next, we can obtain $\tilde{m}_{A^{coC} \times C}$ in other way. First, we need some preliminary results.

Lemma 4.7. *Suppose that $(A^{coC}, \alpha|_{A^{coC}}) \hookrightarrow (A, \alpha)$ is a (C, γ) -Hom-cleft extension via ω . We define a morphism $\varpi : C \otimes A \rightarrow A$ by*

$$\varpi(c, a) = (\omega(c_1)\alpha^{-1}(a)_\psi)\omega^{-1}(\gamma(c_2^\psi)).$$

Then $\rho_A(\varpi(c, a)) \in A^{coC}$, for all $c \in C$ and $a \in A$.

Proof. For if $c \in C, a \in A$, then

$$\begin{aligned}
\rho_A(\varpi(c, a)) &= \alpha(\omega(c_1))_{[0]}\alpha^{-1}(\alpha^{-1}(a)_\psi\omega^{-1}(c_2^\psi))_\Psi \otimes \gamma(\alpha(\omega(c_1))_{[1]})^\Psi \\
&= \alpha(\omega(c_1)_{[0]})\alpha^{-1}(\alpha^{-1}(a)_\psi\omega^{-1}(c_2^\psi))_\Psi \otimes \gamma(\gamma(\omega(c_1)_{[1]})^\Psi) \\
&= \alpha(\omega(c_1)_{[0]})\alpha^{-1}(\alpha^{-1}(a)_\psi\omega^{-1}(c_2^\psi))_\Psi \otimes \gamma(\gamma(\omega(c_1)_{[1]})^\Psi) \\
&= \alpha(\omega(c_1)_{[0]})\alpha^{-1}((\alpha^{-1}(a)_\psi\omega^{-1}(c_2^\psi))_\Psi \otimes \gamma(\gamma(\omega(c_1)_{[1]})^\Psi))^\Psi \\
&= \alpha(\omega(c_1)_{[0]})\alpha^{-1}(\alpha^{-1}(a)_\psi\omega^{-1}(c_2^\psi)_{\Psi'}) \otimes \gamma(\gamma(\omega(c_1)_{[1]})^{\Psi\Psi'}) \\
&= \alpha(\omega(c_1))\alpha^{-1}(\alpha^{-1}(a)_\psi\omega^{-1}(c_2^\psi)_{\Psi'}) \otimes \gamma(\gamma(c_1)_{[1]}^{\Psi\Psi'}) \\
&= \omega(c_1)\alpha^{-1}(\alpha^{-1}(a)_\psi\omega^{-1}(\gamma(c_2)_{[2]}^\psi)_{\Psi'}) \otimes \gamma(\gamma(c_2)_{[1]}^{\Psi\Psi'}) \\
&= \omega(c_1)\alpha^{-1}(\alpha^{-1}(a)_\psi)\omega^{-1}(\gamma(c_2)_{[2]}^\psi)_{\Psi'} \otimes \gamma(\gamma(c_2)_{[1]}^{\Psi\Psi'}) \\
&= \omega(c_1)\alpha^{-1}(\alpha^{-1}(a)_\psi)(\alpha^{-2}(\omega^{-1}(\gamma(c_2)^\psi)1_{[0]})) \otimes \gamma(\gamma(1_{[1]})) \\
&= (\alpha^{-1}(\omega(c_1))\alpha^{-3}(a_\psi\omega^{-1}(\gamma(c_2)^\psi)))1_{[0]} \otimes \gamma(1_{[1]}) \\
&= ((\alpha^{-2}(\omega(c_1))\alpha^{-3}(a_\psi))\alpha^{-2}(\omega^{-1}(\gamma(c_2)^\psi)))1_{[0]} \otimes \gamma(1_{[1]}) \\
&= \alpha^{-2}((\omega(c_1)\alpha^{-1}(a_\psi))\omega^{-1}(\gamma(c_2)^\psi))1_{[0]} \otimes \gamma(1_{[1]}) \\
&= \alpha^{-2}((\omega(c_1)\alpha^{-1}(a)_\psi)\omega^{-1}(\gamma(c_2^\psi)))1_{[0]} \otimes \gamma(1_{[1]}).
\end{aligned}$$

Thus $\varpi(c, a) \in A^{coC}$. \square

Now, we construct a morphism Λ as follows:

$$\Lambda : C \otimes A \rightarrow A \otimes C, \Lambda(c \otimes d) = \varpi(c_1 \otimes \alpha^{-1}(\omega(d))_\psi) \otimes \gamma(c_2^\psi).$$

By Lemma 4.7, we have $\Lambda(c \otimes d) \in A^{coC} \otimes C$. Using Λ , we define a multiplication $m_{A^{coC} \otimes C}$ on $A^{coC} \otimes C$ by

$$\begin{aligned}
m_{A^{coC} \otimes C} &= (m_A \otimes id_C) \circ (m_A \otimes \Lambda \circ (id_C \otimes \omega)) \circ \tilde{a}_{A, A, C \otimes C} \\
&\quad \circ (id_C \otimes \tilde{a}_{A, C, C}) \circ (id_C \otimes \Lambda \otimes id_C) \circ (id_A \otimes \tilde{a}_{C, A, C}^{-1}) \circ \tilde{a}_{A, C, A \otimes C}.
\end{aligned}$$

Concretely,

$$(a \otimes c)(b \otimes d) = (\alpha^{-1}(a)((\alpha^{-2}(\omega(c_1))\alpha^{-2}(\alpha^{-1}(b)_\psi)\omega^{-1}(c_2^\psi_1))) \\ \times ((\omega(\gamma(c_2^\psi_{2_1}))\alpha^{-2}(\omega(d)_\Psi)\omega^{-1}(\gamma^3(c_2^\psi_{2_2})^\Psi_1)) \otimes \gamma^5(c_2^\psi_{2_2})^\Psi_2).$$

Proposition 4.8. *Suppose that $(A^{coC}, \alpha|_{A^{coC}}) \hookrightarrow (A, \alpha)$ is a (C, γ) -Hom-cleft extension via ω . Then $m_{A^{coC} \otimes C} = \tilde{m}_{A^{coC} \otimes C}$.*

Proof. It suffice to prove that $m_{A^{coC} \times C} = \Omega_A \circ m_A \circ (\Omega_A^{-1} \otimes \Omega_A^{-1})$ holds. Indeed, for any $a, b \in A$ and $c, d \in C$, we have

$$\begin{aligned} & \Omega_A^{-1} \circ m_{A^{coC} \otimes C}((a \otimes c) \otimes (b \otimes d)) \\ &= ((\alpha^{-1}(a)((\alpha^{-2}(\omega(c_1))\alpha^{-1}(b)_\psi)\omega^{-1}(c_2^\psi_1))) \\ & \quad \times ((\omega(\gamma(c_2^\psi_{2_1}))\alpha^{-2}(\omega(d)_\Psi)\omega^{-1}(\gamma^3(c_2^\psi_{2_2})^\Psi_1))\omega(\gamma^5(c_2^\psi_{2_2})^\Psi_2)) \\ &= (a((\alpha^{-1}(\omega(c_1)\alpha^{-1}(b)_\psi)\alpha(\omega^{-1}(c_2^\psi_1)))) \\ & \quad \times (((\omega(\gamma(c_2^\psi_{2_1}))\alpha^{-2}(\omega(d)_\Psi)\omega^{-1}(\gamma^3(c_2^\psi_{2_2})^\Psi_1))\omega(\gamma^4(c_2^\psi_{2_2})^\Psi_2)) \\ &= (a((\alpha^{-1}(\omega(c_1)\alpha^{-1}(b)_\psi)\alpha(\omega^{-1}(c_2^\psi_1)))) \\ & \quad \times ((\alpha(\omega(\gamma(c_2^\psi_{2_1})))\alpha^{-1}(\omega(d)_\Psi)(\omega^{-1}(\gamma^3(c_2^\psi_{2_2})^\Psi_1)\omega(\gamma^3(c_2^\psi_{2_2})^\Psi_2))) \\ &= (a((\alpha^{-1}(\omega(c_1)\alpha^{-1}(b)_\psi)\alpha(\omega^{-1}(c_2^\psi_1))))(\alpha^2(\omega(c_2^\psi_2)\omega(d))) \\ &= \alpha(a)((\alpha^{-1}(\omega(c_1)\alpha^{-1}(b)_\psi)\alpha(\omega^{-1}(c_2^\psi_1)))(\alpha(\omega(c_2^\psi_2))\alpha^{-1}(\omega(d)))) \\ &= \alpha(a)((\omega(c_1)\alpha^{-1}(b)_\psi)(\alpha(\omega^{-1}(c_2^\psi_1))(\omega(c_2^\psi_2)\alpha^{-2}(\omega(d)))) \\ &= \alpha(a)((\omega(c_1)\alpha^{-1}(b)_\psi)((\omega^{-1}(c_2^\psi_1)\omega(c_2^\psi_2))\alpha^{-1}(\omega(d)))) \\ &= \alpha(a)((\omega(\gamma^{-1}(c))\alpha^{-1}(b))\omega(d)) \\ &= \alpha(a)(\omega(c)(\alpha^{-1}(b)\alpha^{-1}(\omega(d)))) \\ &= (a\omega(c))(b\omega(d)) \\ &= m_A \circ (\Omega_A^{-1} \otimes \Omega_A^{-1})((a \otimes c) \otimes (b \otimes d)). \end{aligned}$$

Thus we gain the desired result. \square

Acknowledgment. The authors sincerely thank the referee for his/her numerous very valuable comments and suggestions on this article. This work was supported by the National Natural Science Foundation of China (Nos.11261063, 11471186 and 11501317) and the Natural Foundation of Shandong Province (No. ZR2018MA012).

References

- [1] J.N. Alonso Álvarez, J.M. Fernández Vilaboa, R. González Rodríguez, et al. *Weak C-cleft extensions, weak entwining structures and weak Hopf algebras*, J. Algebra **284**, 679–704, 2005.
- [2] T. Brzeziński and S. Majid, *Coalgebra bundles*, Comm. Math. Phys. **191**, 467–492, 1998.
- [3] T. Brzeziński, *On modules associated to coalgebra Galois extensions*, J. Algebra **215**, 290–317, 1999.
- [4] S. Caenepeel and I. Goyvaerts *Monoidal Hom-Hopf algebras*, Comm. Algebra **39** (6), 2216–2240, 2011.
- [5] S. Caenepeel, F. van Oystaeyen and B. Zhou, *Making the category of Doi-Hopf modules into a braided monoidal category*, Algebr. Represent. Theory **1** (1), 75–96, 1998.
- [6] Q.G. Chen and D.G. Wang, *Constructing quasitriangular Hopf algebras*, Comm. Algebra **43** (4), 1698–1722, 2015

- [7] Q.G. Chen and D.G. Wang, *A class of coquasitriangular Hopf group algebras*, Comm. Algebra **44** (1), 310–335, 2016.
- [8] Q.G. Chen and D.G. Wang, *A duality Theorem for L-R crossed product*, Filomat **30** (5), 1305–1313, 2016.
- [9] Q.G. Chen, D.G. Wang and X.D. Kang, *Twisted partial coactions of Hopf algebras*, Front. Math. China **12**, 63–86, 2017.
- [10] Y.Y. Chen, Z.W. Wang and L.Y. Zhang, *The fundamental theorem and Maschke's theorem in the category of relative Hom-Hopf modules*, Colloq. Math. **144** (1), 55–71, 2016.
- [11] Y. Doi and M. Takeuchi, *Cleft comodule algebras for a bialgebra*, Comm. Algebra **14**, 801–817, 1986.
- [12] J.M. Fernández Vilaboa and E. Villanueva Novoa, *A Characterization of the cleft comodule triples*, Comm. Algebra **16**, 613–622, 1988.
- [13] S.J. Guo, X.H. Zhang and S.X. Wang, *Braided monoidal categories and Doi-Hopf modules for monoidal Hom-Hopf algebras*, Colloq. Math. **143** (1), 79–103, 2016.
- [14] S. Karacuha, *Hom-entwining structures and Hom-Hopf-type modules*, arXiv:1412.2002v2.
- [15] H.F. Kreimer and M. Takeuchi, *Hopf algebras and Galois extensions of an algebra*, Indiana Univ. Math. J. **30**, 675–691, 1981.
- [16] L. Liu and B.L. Shen, *Radford's biproducts and Yetter-Drinfel'd modules for monoidal Hom-Hopf algebras*, J. Math. Phys. **55**, 031701, 2014.
- [17] L. Liu and S.H. Wang, *Making the category of entwined modules into a braided monoidal category*, J. Southeast Univ. (English Ed.) **24** (2), 250–252, 2008.
- [18] A. Makhlouf and F. Panaite, *Yetter-Drinfeld modules for Hom-bialgebras*, J. Math. Phys. **55**, 013501, 2014.
- [19] A. Makhlouf and S.D. Silvestrov, *Hom-algebra structures*, J. Gen. Lie Theory Appl. **2**(2), 51–64, 2008.
- [20] A. Makhlouf and S.D. Silvestrov, *Hom-algebras and Hom-coalgebras*, J. Algebra Appl. **9** (4), 553–589, 2010.
- [21] D. Yau, *Hom-Yang-Baxter-equation, Hom-Lie algebras and quasitriangular bialgebras*, J. Phys. A **42**, 165202, 2009.
- [22] D. Yau, *Hom-algebras and homology*, J. Lie Theory **19**, 409–421, 2009.
- [23] D. Yau, *Hom-quantum groups I: Quasitriangular Hom-bialgebras*, J. Phys A **45**, 065203, 2012.
- [24] X.F. Zhao and X.H. Zhang, *Lazy 2-cocycles over monoidal Hom-Hopf algebras*, Colloq. Math. **142**, 61–81, 2016.

Two classes of risk model with diffusion and multiple thresholds: the discounted dividends

Wuyuan Jiang^{*†}

Abstract

In this paper, we consider the present value of total dividends until ruin in a perturbed risk model with two independent classes of risks under multiple thresholds, in which both of the two inter-claim times have phase-type distributions. We obtain the integro-differential equations for the moment-generating function and the r th moment of discounted dividend payments. Explicit expressions for the expectation of discounted dividend payments are derived if the two classes claim amount distributions both belong to the rational family.

Keywords: Two classes of claims, Diffusion process, Dividend payments, Multiple thresholds, Phase-type distribution.

2010 AMS Classification: 62P05; 91B30

Received : 22.09.2017 *Accepted :* 28.06.2018 *Doi :* 10.15672/HJMS.2018.602

1. Introduction

The discounted sum of dividend payments until ruin is an important quantity in assessing the quality of a dividend barrier strategy in insurance risk theory, which has been studied in some papers and books, see e.g. [1], [3], [5], [8], [10], [13].

Recently, some researchers consider the ruin measures for a risk model involving two independent classes of risks in the actuarial literature. Among them, [11] considered the expected discounted penalty functions by assuming that the two claim number processes are independent Poisson and generalized Erlang(2) processes. [15] supposed that the claim number processes are independent Poisson and generalized Erlang(n) processes, respectively, in which the Laplace transforms of the expected discounted penalty functions are obtained. As an extension to these papers, [7] investigated the same ruin measures in the risk model with two classes of renewal risk processes by assuming that both of the two claim number processes have phase-type inter-claim times.

^{*}Department of Mathematics, Hunan Institute of Science and Technology, Yueyang 414006, China, Email: csujw@163.com

[†]Corresponding Author.

There is a particular attention in considering risk models with multi-threshold dividend strategies. For instance, [12] discussed the Gerber-Shiu expected discounted penalty function in the compound Poisson risk model with multiple thresholds. [14] extended the corresponding results to a Sparre Andersen model with generalized Erlang(n)-distributed inter-claim times. In insurance risk models with multiple thresholds, the premium rate is a step function of the insurer's surplus. The premium policy is effective when the insurer intend to keep a fixed retention ratio on its revenues and pays bonuses as an incentive to its policyholders.

[9] investigated the discounted penalty function for two classes of risk processes with diffusion and multiple thresholds, where both of the two claim number processes have phase-type inter-claim times. It is natural to ask for the results on the discounted sum of dividend payments until ruin for a corresponding risk model. The rest of the paper is structured as follows. Section 2 describes the risk model. In Section 3, we derive systems of integro-differential equations for the moment generating function. In Section 4, integro-differential equations for the moments of discounted dividend payments are obtained. Section 5 presents the main results and derives explicit expressions for the expectation of discounted dividend payments when two classes claim amount distributions both belong to the rational family. Section 6 gives a numerical example.

2. Notation and model description

The surplus process $R(t)$ perturbed by diffusion satisfies

$$(2.1) \quad R(t) = u + ct - S(t) + \sigma B(t), \quad t \geq 0,$$

where $u \geq 0$ is the initial surplus, c denotes the insurer's premium income per unit time, $\{B(t); t \geq 0\}$ is a standard Brownian motion and $\sigma > 0$ is the dispersion parameter, and the aggregate-claim process $\{S(t); t \geq 0\}$ is defined by

$$S(t) = \sum_{i=1}^{N_1(t)} X_i + \sum_{i=1}^{N_2(t)} Y_i, \quad t \geq 0,$$

where $\{X_1, X_2, \dots\}$ and $\{Y_1, Y_2, \dots\}$ are independent and identically distributed (i.i.d.) positive random variables representing the successive individual claim amounts from the first and the second class, respectively. The random variables $\{X_1, X_2, \dots\}$ are assumed to have common cumulative distribution function $F(x) = 1 - \bar{F}(x), x \geq 0$, with probability density function $f(x) = F'(x)$, of which the Laplace transform is $\tilde{f}(s) = \int_0^\infty e^{-sx} f(x) dx, s \in \mathbb{C}, \mathbb{C}$ denotes the complex space. Similarly, common cumulative distribution function, density function and the Laplace transform of the density function of $\{Y_1, Y_2, \dots\}$ are given by $G(x) = 1 - \bar{G}(x), x \geq 0, g(x) = G'(x)$ and $\tilde{g}(s) = \int_0^\infty e^{-sx} g(x) dx$. The renewal processes $\{N_1(t); t \geq 0\}$ and $\{N_2(t); t \geq 0\}$ denote the number of claims up to time t caused by the first and the second class of claim respectively, and are defined as follows.

$$N_1(t) = \sup\{n : T_1 + T_2 + \dots + T_n \leq t\},$$

$$N_2(t) = \sup\{n : V_1 + V_2 + \dots + V_n \leq t\},$$

where the i.i.d. interclaim times $\{T_1, T_2, \dots\}$ have common cumulative distribution function $K_1(t), t \geq 0$ and density function $k_1(x) = K_1'(x)$, and $\{V_1, V_2, \dots\}$ have common cumulative distribution function $K_2(t), t \geq 0$ and density function $k_2(x) = K_2'(x)$.

In addition, we suppose that $\{X_1, X_2, \dots\}, \{Y_1, Y_2, \dots\}, \{N_1(t); t \geq 0\}, \{N_2(t); t \geq 0\}$ and $\{B(t); t \geq 0\}$ are mutually independent, and $c > E(X_1)/E(T_1) + E(Y_1)/E(V_1)$, providing a positive safety loading factor.

Under the multi-threshold risk model, there are L thresholds $0 = d_0 < d_1 < \cdots < d_{L-1} < d_L = \infty$ such that when the surplus is between the thresholds d_{l-1} and d_l , dividends are paid continuously at a constant rate $\eta_l \geq 0$. Furthermore, we assume $\eta_1 = 0$, namely, when the surplus is below the level d_1 , no dividends are paid, and $\eta_l > 0$ for $l = 2, 3, \dots, L$. Correspondingly, let c_l denote the premium rate when $d_{l-1} \leq u < d_l$, thus, the net premium rate after dividend payments is $c_{l+1} = c_l - \eta_{l+1} \geq 0$. Thus the surplus process $\{R(t); t \geq 0\}$ can be expressed as

$$(2.2) \quad dR(t) = c_l dt + \sigma dB(t) - dS(t), \quad d_{l-1} \leq R(t) < d_l.$$

The time of (ultimate) ruin is defined as $T = \inf\{t | R(t) \leq 0\}$, where $T = \infty$ if $R(t) > 0$ for all $t \geq 0$. The probability of ruin is $\psi(u) = Pr(T < \infty)$.

Denote by $D(t)$ the cumulative amount of dividends paid out up to time t and $\delta > 0$ the force of interest, then $\mathbb{D} = \int_0^T e^{-\delta t} dD(t)$ is the present value of all dividends until ruin time T . In the following text, we turn to the moment generating function under multiple thresholds,

$$M(u, y) = E[e^{y\mathbb{D}} | R(0) = u]$$

(for those values of y where it exists) and the r th moment

$$W(u, r) = E[\mathbb{D}^r | R(0) = u], \quad r \in \mathbb{N}.$$

Note that $W(u, 0) \equiv 1$. We will always assume that $M(u, y)$ and $W(u, r)$ are sufficiently smooth functions in u and y , respectively.

Throughout the text of the paper, all bold-faced letters represent either vectors or matrices and all vectors are column vectors. We assume that the distribution $K_1(t)$ of the inter-claim time random variable T_1 is phase-type with representation $(\boldsymbol{\alpha}^\top, \mathbf{A}, \mathbf{a})$, where $\boldsymbol{\alpha}^\top = (\alpha_1, \alpha_2, \dots, \alpha_n)$, with $\alpha_i \geq 0$, $\sum_{i=1}^n \alpha_i = 1$, $\mathbf{A} = (a_{ij})_{i,j=1}^n$ is an $n \times n$ matrix with $a_{ii} < 0$, $a_{ij} \geq 0$, for $i \neq j$, $\sum_{j=1}^n a_{ij} \leq 0$, for any $i = 1, 2, \dots, n$, and $\mathbf{a} = (a_1, a_2, \dots, a_n)^\top$ with $\mathbf{a} = -\mathbf{A}\mathbf{e}_n$, where \mathbf{x}^\top denotes the transpose of \mathbf{x} and \mathbf{e}_n denotes a column vector of length n with all elements being one. Following [2], we have $K_1(t) = 1 - \boldsymbol{\alpha}^\top e^{\mathbf{A}t} \mathbf{e}_n$, $k_1(t) = \boldsymbol{\alpha}^\top e^{\mathbf{A}t} \mathbf{a}$, $t \geq 0$, and

$$(2.3) \quad \tilde{k}_1(s) = \int_0^\infty e^{-st} k_1(t) dt = \boldsymbol{\alpha}^\top (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{a}.$$

By the definition of phase-type distributions, each of the inter-claim times $T_i, i = 1, 2, \dots$, corresponds to the time to absorption in a terminating continuous-time Markov Chain, say, $I_t^{(i)}$ with n transient states $\{E_1, E_2, \dots, E_n\}$ and one absorbing state E_0 .

The distribution $K_2(t)$ of the inter-claim time random variable V_1 is phase-type with representation $(\boldsymbol{\beta}^\top, \mathbf{B}, \mathbf{b})$, where $\boldsymbol{\beta}^\top = (\beta_1, \beta_2, \dots, \beta_m)$, $\mathbf{B} = (b_{ij})_{i,j=1}^m$ is an $m \times m$ matrix, $\mathbf{b} = (b_1, b_2, \dots, b_m)^\top$ with $\mathbf{b} = -\mathbf{B}\mathbf{e}_m$. Then we have $K_2(t) = 1 - \boldsymbol{\beta}^\top e^{\mathbf{B}t} \mathbf{e}_m$, $k_2(t) = \boldsymbol{\beta}^\top e^{\mathbf{B}t} \mathbf{b}$, $t \geq 0$, and $\tilde{k}_2(s) = \int_0^\infty e^{-st} k_2(t) dt = \boldsymbol{\beta}^\top (s\mathbf{I} - \mathbf{B})^{-1} \mathbf{b}$. $J_t^{(i)}$ denotes the terminating continuous-time Markov Chain of $V_i, i = 1, 2, \dots$, with m transient states $\{F_1, F_2, \dots, F_m\}$ and one absorbing state F_0 .

Now, we construct a two-dimensional Markov process $\{(I(t), J(t)); t \geq 0\}$ by piecing the $\{I_t^{(i)}; i = 1, 2, \dots\}$ and $\{J_t^{(i)}; i = 1, 2, \dots\}$ together,

$$I(t) = \{I_t^{(1)}\}, 0 \leq t < T_1, \quad I(t) = \{I_{t-T_1}^{(2)}\}, T_1 \leq t < T_1 + T_2, \dots,$$

$$J(t) = \{J_t^{(1)}\}, 0 \leq t < V_1, \quad J(t) = \{J_{t-V_1}^{(2)}\}, V_1 \leq t < V_1 + V_2, \dots.$$

So $\{(I(t), J(t)); t \geq 0\}$ is the underlying state process with states $\{(E_1, F_1), (E_2, F_1), \dots, (E_n, F_1), (E_1, F_2), (E_2, F_2), \dots, (E_n, F_2), \dots, (E_1, F_m), (E_2, F_m), \dots, (E_n, F_m)\}$, initial distribution $\boldsymbol{\gamma} = \boldsymbol{\beta} \otimes \boldsymbol{\alpha}$, where \otimes denotes the Kronecker product of two matrices.

For $k = 1, 2; i = 1, 2, \dots, n; j = 1, 2, \dots, m$, let $M^{(k)}(u, y)$ denote the moment generating function of \mathbb{D} if the ruin is caused by a claim from class k and $R(0) = u$. $M_{ij}^{(k)}(u, y)$ denotes the moment generating function of \mathbb{D} when the ruin is caused by a claim from class k and initial state $(I_0^{(1)}, J_0^{(1)}) = (E_i, F_j)$, then the moment generating function can be written as

$$(2.4) \quad M^{(k)}(u, y) = \boldsymbol{\gamma}^\top \mathbf{M}^{(k)}(u, y),$$

where $\mathbf{M}^{(k)}(u, y) \equiv \left(M_{11}^{(k)}(u, y), M_{21}^{(k)}(u, y), \dots, M_{n1}^{(k)}(u, y), M_{12}^{(k)}(u, y), M_{22}^{(k)}(u, y), \dots, M_{n2}^{(k)}(u, y), \dots, M_{1m}^{(k)}(u, y), M_{2m}^{(k)}(u, y), \dots, M_{nm}^{(k)}(u, y) \right)^\top$. Thus

$$(2.5) \quad M(u, y) = \boldsymbol{\gamma}^\top \mathbf{M}(u, y) = \boldsymbol{\gamma}^\top [\mathbf{M}^{(1)}(u, y) + \mathbf{M}^{(2)}(u, y)].$$

$W_{ij}(u, r)$ denotes the r th moment of \mathbb{D} if $(I_0^{(1)}, J_0^{(1)}) = (E_i, F_j)$, then the moment can be computed by

$$(2.6) \quad W(u, r) = \boldsymbol{\gamma}^\top \mathbf{W}(u, r),$$

where $\mathbf{W}(u, r) \equiv (W_{11}(u, r), W_{21}(u, r), \dots, W_{n1}(u, r), W_{12}(u, r), W_{22}(u, r), \dots, W_{n2}(u, r), \dots, W_{1m}(u, r), W_{2m}(u, r), \dots, W_{nm}(u, r))^\top$.

3. The moment generating function

Let $\frac{\partial}{\partial u}$ and $\frac{\partial}{\partial y}$ denote the differentiation operators with respect to (w.r.t.) u and y , respectively.

3.1. Theorem. *The vectors $\mathbf{M}^{(k)}(u, y), d_{l-1} \leq u < d_l, l = 1, 2, \dots, L, k = 1, 2$ satisfy the following partial integro-differential system, respectively,*

$$(3.1) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{\partial^2}{\partial u^2} + c_i \frac{\partial}{\partial u} - y\delta \frac{\partial}{\partial y} + y\eta_i \right) \mathbf{M}^{(1)}(u, y) + I_{m \times m} \otimes \mathbf{A} \mathbf{M}^{(1)}(u, y) + \\ & \mathbf{B} \otimes I_{n \times n} \mathbf{M}^{(1)}(u, y) + I_{m \times m} \otimes (\mathbf{a} \boldsymbol{\alpha}^\top) \int_0^u \mathbf{M}^{(1)}(u-x, y) f(x) dx + \\ & (\mathbf{b} \boldsymbol{\beta}^\top) \otimes I_{n \times n} \int_0^u \mathbf{M}^{(1)}(u-x, y) g(x) dx + (\mathbf{e}_m \otimes \mathbf{a}) \bar{F}(u) = \mathbf{0}, \end{aligned}$$

and

$$(3.2) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{\partial^2}{\partial u^2} + c_i \frac{\partial}{\partial u} - y\delta \frac{\partial}{\partial y} + y\eta_i \right) \mathbf{M}^{(2)}(u, y) + I_{m \times m} \otimes \mathbf{A} \mathbf{M}^{(2)}(u, y) + \\ & \mathbf{B} \otimes I_{n \times n} \mathbf{M}^{(2)}(u, y) + I_{m \times m} \otimes (\mathbf{a} \boldsymbol{\alpha}^\top) \int_0^u \mathbf{M}^{(2)}(u-x, y) f(x) dx + \\ & (\mathbf{b} \boldsymbol{\beta}^\top) \otimes I_{n \times n} \int_0^u \mathbf{M}^{(2)}(u-x, y) g(x) dx + (\mathbf{b} \otimes \mathbf{e}_n) \bar{G}(u) = \mathbf{0}, \end{aligned}$$

where $I_{n \times n}$ denotes the $n \times n$ identity matrix, $\mathbf{0}$ denotes a column vector of length mn with all elements being 0. $\bar{F}(u) = \int_u^\infty f(x) dx$ and $\bar{G}(u) = \int_u^\infty g(x) dx$.

Proof. Taking into account an infinitesimal time interval $(0, dt)$ for $d_{l-1} \leq u < d_l, l = 1, 2, \dots, L$, there are four possible events regarding to the occurrence of the claim and change of the environment: (1) no claim arrival and no change of state; (2) a claim arrival but no change of state; (3) a change of state but no claim arrival; (4) two or more events

occur. Using the total expectation formula, yields

$$\begin{aligned}
& M_{ij}^{(1)}(u, y) \\
&= e^{y\eta_1 dt} \left\{ (1 + a_{ii}dt)(1 + b_{jj}dt) E[M_{ij}^{(1)}(u + c_l dt + \sigma B(dt), ye^{-\delta dt})] \right. \\
&+ (1 + b_{jj}dt) \sum_{k=1, k \neq i}^n (a_{ik}dt) E[M_{kj}^{(1)}(u + c_l dt + \sigma B(dt), ye^{-\delta dt})] \\
&+ (1 + a_{ii}dt) \sum_{h=1, h \neq j}^m (b_{jh}dt) E[M_{ih}^{(1)}(u + c_l dt + \sigma B(dt), ye^{-\delta dt})] \\
(3.3) \quad &+ (1 + b_{jj}dt)(a_{ii}dt) \\
&\times E \left[\sum_{s=1}^n \alpha_s \int_0^{u+c_l dt + \sigma B(dt)} M_{sj}^{(1)}(u + c_l dt + \sigma B(dt) - x, ye^{-\delta dt}) f(x) dx \right. \\
&+ \left. \int_{u+c_l dt + \sigma B(dt)}^{\infty} f(x) dx \right] + (1 + a_{ii}dt)(b_{jj}dt) \\
&\times E \left[\sum_{r=1}^m \beta_r \int_0^{u+c_l dt + \sigma B(dt)} M_{ir}^{(1)}(u + c_l dt + \sigma B(dt) - x, ye^{-\delta dt}) g(x) dx \right] \left. \right\} \\
&+ o(dt).
\end{aligned}$$

By the aid of Taylor expansion, we have

$$\begin{aligned}
& E[M_{ij}^{(1)}(u + c_l dt + \sigma B(dt), ye^{-\delta dt})] \\
(3.4) \quad &= M_{ij}^{(1)}(u, y) + c_l dt \frac{\partial M_{ij}^{(1)}(u, y)}{\partial u} + y(e^{-\delta dt} - 1) \frac{\partial M_{ij}^{(1)}(u, y)}{\partial y} \\
&+ \frac{\sigma^2}{2} dt \frac{\partial^2 M_{ij}^{(1)}(u, y)}{\partial u^2} + o(dt).
\end{aligned}$$

Substituting (3.4) into (3.3), after some careful calculations, it follows that

$$\begin{aligned}
& \left(\frac{\sigma^2}{2} \frac{\partial^2}{\partial u^2} + c_l \frac{\partial}{\partial u} - y\delta \frac{\partial}{\partial y} + y\eta_1 \right) M_{ij}^{(1)}(u, y) + \sum_{k=1}^n a_{ik} M_{kj}^{(1)}(u, y) + \\
(3.5) \quad & \sum_{h=1}^m b_{jh} M_{ih}^{(1)}(u, y) + a_i \left(\sum_{s=1}^n \alpha_s \int_0^u M_{sj}^{(1)}(u - x, y) f(x) dx + \int_u^{\infty} f(x) dx \right) \\
&+ b_j \sum_{r=1}^m \beta_r \int_0^u M_{ir}^{(1)}(u - x, y) g(x) dx = 0.
\end{aligned}$$

Rewriting (3.5) in matrix form, we conclude (3.1). By similar arguments, we can obtain (3.2). \square

4. The moments of discounted dividend payments

4.1. Theorem. *The vector $\mathbf{W}(u, r)$, $d_{l-1} \leq u < d_l$, $l = 1, 2, \dots, L$, satisfies the following integro-differential system,*

$$\begin{aligned}
(4.1) \quad & \left(\frac{\sigma^2}{2} \frac{d^2}{du^2} + c_l \frac{d}{du} - r\delta \right) \mathbf{W}(u, r) + r\eta_l \mathbf{W}(u, r-1) + I_{m \times m} \otimes \mathbf{A} \mathbf{W}(u, r) + \\
& \mathbf{B} \otimes I_{n \times n} \mathbf{W}(u, r) + I_{m \times m} \otimes (\mathbf{a}\boldsymbol{\alpha}^\top) \int_0^u \mathbf{W}(u-x, r) f(x) dx + \\
& (\mathbf{b}\boldsymbol{\beta}^\top) \otimes I_{n \times n} \int_0^u \mathbf{W}(u-x, r) g(x) dx = \mathbf{0},
\end{aligned}$$

with boundary conditions

$$(4.2) \quad \mathbf{W}(u, r)|_{u=0} = \mathbf{0}, \quad \mathbf{W}(u, r)|_{u=d_l^-} = \mathbf{W}(u, r)|_{u=d_l^+},$$

and

$$(4.3) \quad \left. \frac{d\mathbf{W}(u, r)}{du} \right|_{u=d_l^-} = \left. \frac{d\mathbf{W}(u, r)}{du} \right|_{u=d_l^+}.$$

Proof. Adding (3.1) to (3.2) and noting that $\mathbf{M}(u, y) = \mathbf{M}^{(1)}(u, y) + \mathbf{M}^{(2)}(u, y)$, which results in

$$(4.4) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{\partial^2}{\partial u^2} + c_l \frac{\partial}{\partial u} - y\delta \frac{\partial}{\partial y} + y\eta_l \right) \mathbf{M}(u, y) + I_{m \times m} \otimes \mathbf{A}\mathbf{M}(u, y) + \\ & \mathbf{B} \otimes I_{n \times n} \mathbf{M}(u, y) + I_{m \times m} \otimes (\mathbf{a}\boldsymbol{\alpha}^\top) \int_0^u \mathbf{M}(u-x, y) f(x) dx + \\ & (\mathbf{b}\boldsymbol{\beta}^\top) \otimes I_{n \times n} \int_0^u \mathbf{M}(u-x, y) g(x) dx + (\mathbf{e}_m \otimes \mathbf{a}) \bar{F}(u) \\ & + (\mathbf{b} \otimes \mathbf{e}_n) \bar{G}(u) = \mathbf{0}. \end{aligned}$$

Since $W(u, r) = E[\mathbb{D}^r | R(0) = u]$, we have the following representation

$$(4.5) \quad \mathbf{M}(u, y) = \mathbf{e}_{mn} + \sum_{r=1}^{\infty} \frac{y^r}{r!} \mathbf{W}(u, r).$$

Substituting (4.5) into (4.4) and equating the coefficients of y^r ($r \in \mathbb{N}$). Then by virtue of $\mathbf{a} = -\mathbf{A}\mathbf{e}_n$ and $\mathbf{b} = -\mathbf{B}\mathbf{e}_m$. Further, $I_{m \times m} \otimes \mathbf{A}\mathbf{e}_{mn} = -I_{m \times m} \otimes (\mathbf{a}\boldsymbol{\alpha}^\top)\mathbf{e}_{mn} = -\mathbf{e}_m \otimes \mathbf{a}$ and $\mathbf{B} \otimes I_{n \times n}\mathbf{e}_{mn} = -(\mathbf{b}\boldsymbol{\beta}^\top) \otimes I_{n \times n}\mathbf{e}_{mn} = -\mathbf{b} \otimes \mathbf{e}_n$. Hence, we achieve (4.1).

When $u = 0$, ruin is immediate and no dividends are paid. That is to say $\mathbf{W}(u, r)|_{u=0} = \mathbf{0}$. Utilizing the continuity of $\mathbf{W}(u, r)$ and thanks to [16], we have the boundary conditions (4.2) and (4.3). \square

4.2. Remark. We assume that $m = 1$ and $G(0) = 1$, from Eq.(4.1), which yields

$$(4.6) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{d^2}{du^2} + c_l \frac{d}{du} - r\delta \right) \mathbf{W}(u, r) + r\eta_l \mathbf{W}(u, r-1) \\ & + \mathbf{A}\mathbf{W}(u, r) + (\mathbf{a}\boldsymbol{\alpha}^\top) \int_0^u \mathbf{W}(u-x, r) f(x) dx = \mathbf{0}, \end{aligned}$$

where $d_{l-1} \leq u < d_l, l = 1, 2, \dots, L$. Furthermore, when $L = 2$ and the distribution $K_1(t)$ is a generalized Erlang(n) distribution, we recover Theorem 4.1 in [5] from (4.6), which consider the perturbed renewal risk model with a threshold dividend strategy.

5. The expectation of discounted dividend payments

In what follows, we consider the case $r = 1$ for $\mathbf{W}(u, r)$, the expectation of discounted dividend payments. For notational convenience, let $\mathbf{W}(u) \equiv \mathbf{W}(u, 1)$. From Theorem 4.1, we have for $d_{l-1} \leq u < d_l, l = 1, 2, \dots, L$, that

$$(5.1) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{d^2}{du^2} + c_l \frac{d}{du} - \delta \right) \mathbf{W}(u) + \eta_l \mathbf{e}_{mn} + I_{m \times m} \otimes \mathbf{A}\mathbf{W}(u) + \\ & \mathbf{B} \otimes I_{n \times n} \mathbf{W}(u) + I_{m \times m} \otimes (\mathbf{a}\boldsymbol{\alpha}^\top) \int_0^u \mathbf{W}(u-x) f(x) dx + \\ & (\mathbf{b}\boldsymbol{\beta}^\top) \otimes I_{n \times n} \int_0^u \mathbf{W}(u-x) g(x) dx = \mathbf{0}, \end{aligned}$$

and with boundary conditions $\mathbf{W}(u)|_{u=0} = \mathbf{0}$, $\mathbf{W}(u)|_{u=d_l^-} = \mathbf{W}(u)|_{u=d_l^+}$, $\frac{d\mathbf{W}(u)}{du} \Big|_{u=d_l^-} = \frac{d\mathbf{W}(u)}{du} \Big|_{u=d_l^+}$.

5.1. Laplace transforms. Motivating by [12], we relax the constraint $d_{l-1} \leq u < d_l$ in (5.1) and consider the case of $u \geq d_{l-1}$. Let $\mathbf{W}_l(u), u \geq d_{l-1}, l = 1, \dots, L$ be the solutions of the following non-homogeneous integro-differential equations:

$$(5.2) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{d^2}{du^2} + c_l \frac{d}{du} - \delta \right) \mathbf{W}_l(u) + I_{m \times m} \otimes \mathbf{A}\mathbf{W}_l(u) + \mathbf{B} \otimes I_{n \times n} \mathbf{W}_l(u) + \\ & I_{m \times m} \otimes (\mathbf{a}\boldsymbol{\alpha}^\top) \left[\int_0^{u-d_{l-1}} \mathbf{W}_l(u-x) f(x) dx + \int_{u-d_{l-1}}^u \mathbf{W}(u-x) f(x) dx \right] \\ & + (\mathbf{b}\boldsymbol{\beta}^\top) \otimes I_{n \times n} \left[\int_0^{u-d_{l-1}} \mathbf{W}_l(u-x) g(x) dx + \int_{u-d_{l-1}}^u \mathbf{W}(u-x) g(x) dx \right] \\ & + \eta_l \mathbf{e}_{mn} = \mathbf{0}, \quad u \geq d_{l-1}. \end{aligned}$$

From the theory of differential equations, it follows that

$$(5.3) \quad \mathbf{W}(u) = \mathbf{W}_l(u) + \sum_{j=1}^{mn} k_{lj} \Theta_{lj}(u), \quad d_{l-1} \leq u < d_l,$$

where k_{lj} is constant coefficient for each l and j , and $\Theta_{lj}(u), j = 1, 2, \dots, mn$, are mn linearly independent solutions to the associated homogeneous integro-differential equations

$$(5.4) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{d^2}{du^2} + c_l \frac{d}{du} - \delta \right) \Theta_l(u) + I_{m \times m} \otimes \mathbf{A} \Theta_l(u) + \\ & \mathbf{B} \otimes I_{n \times n} \Theta_l(u) + I_{m \times m} \otimes (\mathbf{a} \boldsymbol{\alpha}^\top) \int_0^{u-d_{l-1}} \Theta_l(u-x) f(x) dx + \\ & (\mathbf{b} \boldsymbol{\beta}^\top) \otimes I_{n \times n} \int_0^{u-d_{l-1}} \Theta_l(u-x) g(x) dx = \mathbf{0}, \quad u \geq d_{l-1}. \end{aligned}$$

5.1. Remark. When $u \rightarrow \infty$, ruin does not happen all the time and dividends are always paid at a constant rate η_L . So we have $\lim_{u \rightarrow \infty} \mathbf{W}(u) = \frac{\eta_L}{\delta} \mathbf{e}_{mn}$. We can find that $\frac{\eta_L}{\delta} \mathbf{e}_{mn}$ are really particular solutions of (5.2). It follows from the general theory of differential equations that

$$(5.5) \quad \mathbf{W}(u) = \frac{\eta_L}{\delta} \mathbf{e}_{mn} + \sum_{j=1}^{mn} k_{Lj} \Theta_{Lj}(u), \quad u \geq d_{L-1},$$

Taking a change of variables $z = u - d_{l-1}$ and $\Phi_l(z) \equiv \mathbf{W}_l(u) = \mathbf{W}_l(z + d_{l-1})$, then we obtain from (5.2),

$$(5.6) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{d^2}{dz^2} + c_l \frac{d}{dz} - \delta \right) \Phi_l(z) + I_{m \times m} \otimes \mathbf{A} \Phi_l(z) + \\ & \mathbf{B} \otimes I_{n \times n} \Phi_l(z) + I_{m \times m} \otimes (\mathbf{a} \boldsymbol{\alpha}^\top) \int_0^z \Phi_l(z-x) f(x) dx + \\ & (\mathbf{b} \boldsymbol{\beta}^\top) \otimes I_{n \times n} \int_0^z \Phi_l(z-x) g(x) dx + \Gamma_l(z) = \mathbf{0}, \quad z \geq 0, \end{aligned}$$

where

$$(5.7) \quad \begin{aligned} \Gamma_l(z) = & I_{m \times m} \otimes (\mathbf{a} \boldsymbol{\alpha}^\top) \int_0^{d_{l-1}} \mathbf{W}(x) f(z + d_{l-1} - x) dx \\ & + (\mathbf{b} \boldsymbol{\beta}^\top) \otimes I_{n \times n} \int_0^{d_{l-1}} \mathbf{W}(x) g(z + d_{l-1} - x) dx + \eta_l \mathbf{e}_{mn}. \end{aligned}$$

Next define the following Laplace transforms: $\tilde{\Phi}_l(s) = \int_0^\infty e^{-sx} \Phi_l(x) dx$, $\tilde{\Gamma}_l(s) = \int_0^\infty e^{-sx} \Gamma_l(x) dx$. Taking Laplace transforms on both sides of (5.6) and rearranging, we have

$$(5.8) \quad \begin{aligned} & \left[\left(\frac{\sigma^2}{2} s^2 + c_l s - \delta \right) I_{mn \times mn} + I_{m \times m} \otimes \mathbf{A} + \mathbf{B} \otimes I_{n \times n} + \right. \\ & \left. I_{m \times m} \otimes (\mathbf{a} \boldsymbol{\alpha}^\top) \tilde{f}(s) + (\mathbf{b} \boldsymbol{\beta}^\top) \otimes I_{n \times n} \tilde{g}(s) \right] \tilde{\Phi}_l(s) = \\ & \frac{\sigma^2}{2} \Phi_l'(0) + p_l(s) \Phi_l(0) - \tilde{\Gamma}_l(s), \end{aligned}$$

where $p_l(s) = \frac{\sigma^2}{2} s + c_l$, $\Phi_l(0) = \mathbf{W}_l(d_{l-1})$, $\Phi_l'(0) = \mathbf{W}_l'(d_{l-1})$.

Let $\mathbf{L}_l(s) = \left(\frac{\sigma^2}{2} s^2 + c_l s - \delta \right) I_{mn \times mn} + I_{m \times m} \otimes \mathbf{A} + \mathbf{B} \otimes I_{n \times n} + I_{m \times m} \otimes (\mathbf{a} \boldsymbol{\alpha}^\top) \tilde{f}(s) + (\mathbf{b} \boldsymbol{\beta}^\top) \otimes I_{n \times n} \tilde{g}(s)$, and $\mathbf{L}_l^*(s)$ is the adjoint of matrix $\mathbf{L}_l(s)$ for $l = 1, 2, \dots, L$. Thus, when $\det[\mathbf{L}_l(s)] \neq 0$, we get from (5.8)

$$(5.9) \quad \tilde{\Phi}_l(s) = \frac{\mathbf{L}_l^*(s)}{\det[\mathbf{L}_l(s)]} \left(\frac{\sigma^2}{2} \Phi_l'(0) + p_l(s) \Phi_l(0) - \tilde{\Gamma}_l(s) \right).$$

For a given l the generalized Lundberg's equations $\det[\mathbf{L}_l(s)] = 0$ has exactly mn roots in the right half of the complex plane when $\delta > 0$, see e.g. [7] for details. We denote them by $\rho_{l1}, \rho_{l2}, \dots, \rho_{l,mn}$ respectively, and for simplicity, we assume that they are different from each other.

Divided difference plays an important role in the present paper. Now we recall divided differences of a matrix $\mathbf{L}(s)$ w.r.t. distinct numbers r_1, r_2, \dots , which are defined recursively as follows:

$$\mathbf{L}[r_1, s] = \frac{\mathbf{L}(s) - \mathbf{L}(r_1)}{s - r_1}, \quad \mathbf{L}[r_1, r_2, s] = \frac{\mathbf{L}[r_1, s] - \mathbf{L}[r_1, r_2]}{s - r_2},$$

and so on.

Since each element of $\tilde{\Phi}_l(s)$ is finite for all $\Re(s) > 0$, $\rho_{l1}, \rho_{l2}, \dots, \rho_{l,mn}$ are also roots of numerator in (5.9). Utilizing a similar technique to Theorem 4.2 in [7], we obtain from (5.9) the following theorem.

5.2. Theorem. *The Laplace transforms of $\Phi_l(y)$ for $l = 1, 2, \dots, L$ are represented by*

$$(5.10) \quad \begin{aligned} \tilde{\Phi}_l(s) = & \frac{\prod_{j=1}^{mn} (s - \rho_{lj})}{\det[\mathbf{L}_l(s)]} \left\{ \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}, s] \left(\frac{\sigma^2}{2} \Phi'_l(0) + p_l(s) \Phi_l(0) - \tilde{\Gamma}_l(s) \right) + \right. \\ & \left. \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}] \frac{\sigma^2}{2} \Phi_l(0) - \sum_{i=1}^{mn} \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{li}] \tilde{\Gamma}_l[\rho_{li}, \dots, \rho_{l,mn}, s] \right\}. \end{aligned}$$

5.2. The homogeneous integro-differential equations. The solutions to the associated homogeneous integro-differential equations (5.4) are uniquely determined by the initial conditions $\Theta_l(d_{l-1})$ and $\Theta'_l(d_{l-1})$. In the following, we apply Laplace transforms to find the solutions of (5.4).

let $z = u - d_{l-1}$ and $\Xi_l(z) \equiv \Theta_l(u) = \Theta_l(z + d_{l-1})$, $l = 1, 2, \dots, L$, then (5.4) can be rewritten as

$$(5.11) \quad \begin{aligned} & \left(\frac{\sigma^2}{2} \frac{d^2}{dz^2} + c_l \frac{d}{dz} - \delta \right) \Xi_l(z) + I_{m \times m} \otimes \mathbf{A} \Xi_l(z) + \\ & \mathbf{B} \otimes I_{n \times n} \Xi_l(z) + I_{m \times m} \otimes (\mathbf{a} \mathbf{a}^\top) \int_0^z \Xi_l(z-x) f(x) dx + \\ & (\mathbf{b} \mathbf{b}^\top) \otimes I_{n \times n} \int_0^z \Xi_l(z-x) g(x) dx = \mathbf{0}, \quad z \geq 0. \end{aligned}$$

Taking Laplace transforms on both sides of (5.11) yields

$$(5.12) \quad \begin{aligned} & \left[\left(\frac{\sigma^2}{2} s^2 + c_l s - \delta \right) I_{mn \times mn} + I_{m \times m} \otimes \mathbf{A} + \mathbf{B} \otimes I_{n \times n} + \right. \\ & \left. I_{m \times m} \otimes (\mathbf{a} \mathbf{a}^\top) \tilde{f}(s) + (\mathbf{b} \mathbf{b}^\top) \otimes I_{n \times n} \tilde{g}(s) \right] \tilde{\Xi}_l(s) = \frac{\sigma^2}{2} \Xi'_l(0) + p_l(s) \Xi_l(0), \end{aligned}$$

where $\tilde{\Xi}_l(s) = \int_0^\infty e^{-sx} \Xi_l(x) dx$. Then, we have

$$(5.13) \quad \tilde{\Xi}_l(s) = \frac{\mathbf{L}_l^*(s)}{\det[\mathbf{L}_l(s)]} \left(\frac{\sigma^2}{2} \Xi'_l(0) + p_l(s) \Xi_l(0) \right).$$

Since $\Theta_l(d_{l-1}) = \Xi_l(0)$, $\Theta'_l(d_{l-1}) = \Xi'_l(0)$, invert (5.13) leads to

$$(5.14) \quad \Theta_l(u) = \mathcal{L}^{-1} \left\{ \frac{\mathbf{L}_l^*(s)}{\det[\mathbf{L}_l(s)]} \left(\frac{\sigma^2}{2} \Theta'_l(d_{l-1}) + p_l(s) \Theta_l(d_{l-1}) \right) \right\}, \quad u \geq d_{l-1}.$$

5.3. Claim sizes with rational Laplace transform. Let us now restrict the further analysis to the case of the claim amount distributions $F(x)$ and $G(x)$ both with rational Laplace transforms, that is,

$$\tilde{f}(s) = \frac{\mathfrak{q}_{m_1-1}(s)}{\mathfrak{q}_{m_1}(s)}, \quad \tilde{g}(s) = \frac{\mathfrak{r}_{m_2-1}(s)}{\mathfrak{r}_{m_2}(s)}, \quad m_1, m_2 \in \mathbb{N}^+,$$

where $\mathfrak{q}_{m_1-1}(s)$, $\mathfrak{r}_{m_2-1}(s)$ are polynomials of degree m_1-1 and m_2-1 or less, respectively, while $\mathfrak{q}_{m_1}(s)$ and $\mathfrak{r}_{m_2}(s)$ are polynomials of degree m_1 and m_2 with only negative roots, and satisfy $\mathfrak{q}_{m_1-1}(0) = \mathfrak{q}_{m_1}(0)$, $\mathfrak{r}_{m_2-1}(0) = \mathfrak{r}_{m_2}(0)$. Without loss of generality, we assume that $\mathfrak{q}_{m_1}(s)$ and $\mathfrak{r}_{m_2}(s)$ have leading coefficient 1. This wide class of distributions includes the Erlang, Coxian and phase-type distributions, and also the mixtures of these.

Multiplying both numerator and denominator of (5.13) by $\mathfrak{h}(s)$, where $\mathfrak{h}(s) = [\mathfrak{q}_{m_1}(s)\mathfrak{r}_{m_2}(s)]^{mn}$. We get for $l = 1, 2, \dots, L$ that

$$(5.15) \quad \tilde{\Xi}_l(s) = \frac{\mathbf{L}_l^*(s)}{\mathfrak{h}(s)\det[\mathbf{L}_l(s)]} \left(\mathfrak{h}(s) \frac{\sigma^2}{2} \Xi'_l(0) + \mathfrak{h}(s)p_l(s)\Xi_l(0) \right).$$

It is obvious that the factor $\mathfrak{h}(s)\det[\mathbf{L}_l(s)]$ of the denominator is a polynomial of degree $mn(m_1 + m_2 + 2)$ with leading coefficient $(\sigma^2/2)^{mn}$. Therefore, the equation $\mathfrak{h}(s)\det[\mathbf{L}_l(s)] = 0$ has $mn(m_1 + m_2 + 2)$ roots on the complex plane. We can factorize $\mathfrak{h}(s)\det[\mathbf{L}_l(s)]$ as follows

$$(5.16) \quad \mathfrak{h}(s)\det[\mathbf{L}_l(s)] = \left(\frac{\sigma^2}{2} \right)^{mn} \prod_{j=1}^{mn} (s - \rho_{lj}) \prod_{j=1}^{mn(m_1+m_2+1)} (s + R_{lj}),$$

where R_{lj} for each l and j has positive real part and we assume that all of them are distinct from each other.

Since $p_l(s)$ with degree 1, the numerator $\mathbf{L}_l^*(s) \left(\mathfrak{h}(s) \frac{\sigma^2}{2} \Xi'_l(0) + \mathfrak{h}(s)p_l(s)\Xi_l(0) \right)$ in (5.15) is a polynomial with degree less than $mn(m_1 + m_2 + 2)$ for each l . By the partial fraction decomposition, we get

$$(5.17) \quad \tilde{\Xi}_l(s) = \sum_{j=1}^{mn} \frac{\boldsymbol{\vartheta}_{lj}}{s - \rho_{lj}} + \sum_{j=1}^{mn(m_1+m_2+1)} \frac{\boldsymbol{\chi}_{lj}}{s + R_{lj}}, \quad s \in \mathbb{C},$$

where $\boldsymbol{\vartheta}_{lj}$, for $j = 1, 2, \dots, mn$, and $\boldsymbol{\chi}_{lj}$, for $j = 1, 2, \dots, mn(m_1 + m_2 + 1)$, are the coefficient matrices defined respectively by

$$(5.18) \quad \boldsymbol{\vartheta}_{lj} = - \frac{\mathbf{L}_l^*(\rho_{lj}) \left(\mathfrak{h}(\rho_{lj}) \frac{\sigma^2}{2} \Xi'_l(0) + \mathfrak{h}(\rho_{lj})p_l(\rho_{lj})\Xi_l(0) \right)}{\left(\frac{\sigma^2}{2} \right)^{mn} \left[\prod_{k=1}^{mn(m_1+m_2+1)} (R_{lk} + \rho_{lj}) \right] \left[\prod_{i=1, i \neq j}^{mn} (\rho_{li} - \rho_{lj}) \right]},$$

and

$$(5.19) \quad \boldsymbol{\chi}_{lj} = \frac{\mathbf{L}_l^*(-R_{lj}) \left(\mathfrak{h}(-R_{lj}) \frac{\sigma^2}{2} \Xi'_l(0) + \mathfrak{h}(-R_{lj})p_l(-R_{lj})\Xi_l(0) \right)}{\left(\frac{\sigma^2}{2} \right)^{mn} \left[\prod_{k=1}^{mn} (\rho_{lk} + R_{lj}) \right] \left[\prod_{i=1, i \neq j}^{mn(m_1+m_2+1)} (R_{li} - R_{lj}) \right]}.$$

Inverting (5.17) yields

$$(5.20) \quad \Xi_l(z) = \sum_{j=1}^{mn} \boldsymbol{\vartheta}_{lj} e^{\rho_{lj}z} + \sum_{j=1}^{mn(m_1+m_2+1)} \boldsymbol{\chi}_{lj} e^{-R_{lj}z}, \quad z \geq 0.$$

To conclude, we have

5.3. Theorem. *If the claim-size distributions $F(x)$ and $G(x)$ both belong to the rational family, then the solutions of the associated homogeneous integro-differential equations (5.4) are given by*

$$(5.21) \quad \Theta_l(u) = \sum_{j=1}^{mn} \boldsymbol{\vartheta}_{lj} e^{\rho_{lj}(u-d_{l-1})} + \sum_{j=1}^{mn(m_1+m_2+1)} \boldsymbol{\chi}_{lj} e^{-R_{lj}(u-d_{l-1})},$$

$u \geq d_{l-1}, l = 1, 2, \dots, L,$

where $\boldsymbol{\vartheta}_{lj}$ and $\boldsymbol{\chi}_{lj}$ are given by (5.18) and (5.19), respectively.

Next, we turn to derive the expressions of $\mathbf{W}_l(u)$, for $l = 1, 2, \dots, L$. For this purpose, multiplying both numerator and denominator of (5.10) by $\mathfrak{h}(s)$, by virtue of (5.16) and then canceling the same factor $\prod_{j=1}^{mn} (s - \rho_{lj})$, we derive from (5.10) that

$$(5.22) \quad \begin{aligned} \tilde{\Phi}_l(s) &= \frac{1}{\left(\frac{\sigma^2}{2}\right)^{mn} \prod_{j=1}^{mn(m_1+m_2+1)} (s+R_{lj})} \times \\ &\left\{ \mathfrak{h}(s) \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}, s] \left(\frac{\sigma^2}{2} \Phi'_l(0) - \tilde{\Gamma}_l(s) \right) + \right. \\ &\mathfrak{h}(s) \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}, s] p_l(s) \Phi_l(0) + \mathfrak{h}(s) \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}] \frac{\sigma^2}{2} \Phi_l(0) \\ &\left. - h(s) \sum_{i=1}^{mn} \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{li}] \tilde{\Gamma}_l[\rho_{li}, \dots, \rho_{l,mn}, s] \right\}. \end{aligned}$$

Thanks to [9], which can be rewritten as

$$(5.23) \quad \begin{aligned} \tilde{\Phi}_l(s) &= \frac{1}{\left(\frac{\sigma^2}{2}\right)^{mn}} \sum_{j=1}^{mn(m_1+m_2+1)} \frac{1}{s+R_{lj}} \left\{ \mathbf{Q}_{lj} \left(\frac{\sigma^2}{2} \Phi'_l(0) - \tilde{\Gamma}_l(s) \right) + \mathbf{H}_{lj} \Phi_l(0) \right. \\ &\left. + D_{lj} \left(\mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}] \frac{\sigma^2}{2} \Phi_l(0) - \sum_{i=1}^{mn} \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{li}] \tilde{\Gamma}_l[\rho_{li}, \dots, \rho_{l,mn}, s] \right) \right\}. \end{aligned}$$

where \mathbf{Q}_{ij} , D_{ij} and \mathbf{H}_{ij} are given respectively by

$$(5.24) \quad \mathbf{Q}_{lj} = \frac{\mathfrak{h}(-R_{lj}) \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}, -R_{lj}]}{\prod_{i=1, i \neq j}^{mn(m_1+m_2+1)} (R_{li} - R_{lj})}, \quad D_{lj} = \frac{\mathfrak{h}(-R_{lj})}{\prod_{i=1, i \neq j}^{mn(m_1+m_2+1)} (R_{li} - R_{lj})},$$

and

$$(5.25) \quad \mathbf{H}_{lj} = \frac{\mathfrak{h}(-R_{lj}) \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}, -R_{lj}] p_l(-R_{lj})}{\prod_{i=1, i \neq j}^{mn(m_1+m_2+1)} (R_{li} - R_{lj})}.$$

In order to obtain the Laplace inverses of (5.23), we recall the operator T_r for a real-valued integrable function $f(x)$ defined by $T_r f(x) = \int_x^\infty e^{-r(u-x)} f(u) du$, $r \in \mathbb{C}$, $x \geq 0$. For properties of the operator T_r , see [4]. Now, we extend the definition of operator T_r for a real-valued integrable function to a matrix function w.r.t. a complex number r . If each element is a real-valued integrable function of x in matrix $\Psi(x)$, we define $T_r \Psi(x) = \int_x^\infty e^{-r(u-x)} \Psi(u) du$, $r \in \mathbb{C}$, $x \geq 0$, and it is easy to see that

$$T_{r_1} T_{r_2} \Psi(x) = T_{r_2} T_{r_1} \Psi(x) = \frac{T_{r_1} \Psi(x) - T_{r_2} \Psi(x)}{r_1 - r_2}, r_1 \neq r_2 \in \mathbb{C}, x \geq 0.$$

Furthermore, from [6], we can get the Laplace inverse of $\tilde{\Psi}[r_1, r_2, \dots, r_n, s]$ as follows

$$(5.26) \quad \mathcal{L}^{-1} \left(\tilde{\Psi}[r_1, r_2, \dots, r_n, s] \right) = (-1)^n \left(\prod_{i=1}^n T_{r_i} \right) \Psi(x).$$

Using (5.26) and inverting (5.23), which results in

$$(5.27) \quad \begin{aligned} \Phi_l(z) &= \frac{1}{\left(\frac{\sigma^2}{2}\right)^{mn}} \sum_{j=1}^{mn(m_1+m_2+1)} \left\{ \left(\frac{\sigma^2}{2} \mathbf{Q}_{lj} \Phi'_l(0) + \mathbf{H}_{lj} \Phi_l(0) \right) e^{-R_{lj}z} \right. \\ &+ \frac{\sigma^2}{2} D_{lj} \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{l,mn}] \Phi_l(0) e^{-R_{lj}z} - e^{-R_{lj}z} \star \left[\mathbf{Q}_{lj} \Gamma_l(z) \right. \\ &\left. \left. - D_{lj} \sum_{i=1}^{mn} \mathbf{L}_l^*[\rho_{l1}, \dots, \rho_{li}] (-1)^{mn-i} \left(\prod_{j=i}^{mn} T_{\rho_{lj}} \right) \Gamma_l(z) \right] \right\}, \quad z \geq 0, \end{aligned}$$

where \star represents the convolution operator.

Since $\Phi_l(z) = \mathbf{W}_l(u) = \mathbf{W}_l(z + d_{l-1})$, we can obtain the following theorem from (5.27).

5.4. Theorem. *If the claim-size distributions $F(x)$ and $G(x)$ both belong to the rational family, for $l = 1, 2, \dots, L$, when $u \geq d_{l-1}$, the solutions of the equations (5.2) are given by*

$$(5.28) \quad \begin{aligned} \mathbf{W}_l(u) = & \frac{1}{\left(\frac{\sigma^2}{2}\right)^{mn}} \sum_{j=1}^{mn(m_1+m_2+1)} \left\{ \left(\frac{\sigma^2}{2} \mathbf{Q}_{lj} \mathbf{W}'_l(d_{l-1}) + \mathbf{H}_{lj} \mathbf{W}_l(d_{l-1}) \right) e^{-R_{lj}(u-d_{l-1})} \right. \\ & + \frac{\sigma^2}{2} D_{lj} \mathbf{L}_l^* [\rho_{l1}, \dots, \rho_{l,mn}] \mathbf{W}_l(d_{l-1}) e^{-R_{lj}(u-d_{l-1})} \\ & \left. - e^{-R_{lj}(u-d_{l-1})} \star [\mathbf{Q}_{lj} \Gamma_l(u-d_{l-1}) \right. \\ & \left. - D_{lj} \sum_{i=1}^{mn} \mathbf{L}_l^* [\rho_{l1}, \dots, \rho_{li}] (-1)^{mn-i} \left(\prod_{j=i}^{mn} T_{\rho_{lj}} \right) \Gamma_l(u-d_{l-1}) \right\}, \end{aligned}$$

where \mathbf{Q}_{lj} , D_{lj} and \mathbf{H}_{lj} are given by (5.24), (5.25), respectively.

5.5. Remark. Let $l = 1$ in (5.28), we have

$$(5.29) \quad \mathbf{W}_1(u) = \frac{1}{\left(\frac{\sigma^2}{2}\right)^{mn-1}} \sum_{j=1}^{mn(m_1+m_2+1)} \mathbf{Q}_{1j} \mathbf{W}'_1(0) e^{-R_{1j}u}, \quad u \geq 0.$$

Obviously, $\mathbf{W}(u) = \mathbf{W}_1(u)$ for $0 \leq u < d_1$. By virtue of $\mathbf{W}'_r(u) = r \mathbf{W}_{r-1}(u)$. Thus, when $r = 1$, $\mathbf{W}'(u) = \mathbf{e}_{mn}$, that is, $\mathbf{W}'(d_{l-1}) = \mathbf{e}_{mn}$, $l = 2, 3, \dots, L$. So, differentiating (5.29) w.r.t. u and letting $u = d_1$, we can determine $\mathbf{W}'_1(0)$. Thus, $\mathbf{W}(u)$, $0 \leq u < d_1$ can be obtain.

6. Numerical illustrations

We now illustrate an application of the main conclusions in this paper with a numerical example. We suppose that the claim amounts from class 1 and class 2 have density functions, respectively,

$$f(x) = \mu_1 e^{-\mu_1 x}, \quad \mu_1 > 0, x > 0, \quad g(y) = \mu_2 e^{-\mu_2 y}, \quad \mu_2 > 0, y > 0.$$

Hence, the Laplace transforms $\tilde{f}(s) = \frac{\mu_1}{s+\mu_1}$, $\tilde{g}(s) = \frac{\mu_2}{s+\mu_2}$. The inter-claim times from class 1 occur following a Poisson process with parameter λ , i.e. $\boldsymbol{\alpha} = (1)$, $\mathbf{A} = (-\lambda)$, $\mathbf{a} = (\lambda)$, and inter-claim times from class 2 occur following a phase-type distribution with the following parameters: $\boldsymbol{\beta} = (1/2, 1/2)^\top$, $\mathbf{B} = \text{diag}(-\lambda_1, -\lambda_2)$, $\mathbf{b} = (\lambda_1, \lambda_2)^\top$. In addition, we assume that the multi-threshold layers $L = 2$ with $0 = d_0 < d_1 < d_2 = \infty$. So, we have $\mathfrak{h}(s) = [(s + \mu_1)(s + \mu_2)]^2$ and $\mathbf{L}_l(s)$, $l = 1, 2$ are given by

$$\left(\begin{array}{cc} \kappa(s) - \lambda_1 + \frac{\lambda \mu_1}{s + \mu_1} + \frac{\lambda_1 \mu_2}{2(s + \mu_2)} & \frac{\lambda_1 \mu_2}{2(s + \mu_2)} \\ \frac{\lambda_2 \mu_2}{2(s + \mu_2)} & \kappa(s) - \lambda_2 + \frac{\lambda \mu_1}{s + \mu_1} + \frac{\lambda_2 \mu_2}{2(s + \mu_2)} \end{array} \right),$$

where $\kappa(s) = \frac{\sigma^2}{2} s^2 + c_1 s - \delta - \lambda$. Let $\mu_1 = 1, \mu_2 = 2, \lambda = 2, \lambda_1 = 1, \lambda_2 = 3, \sigma = 1, \delta = 0.01, d_1 = 2, c_1 = 2$ and $c_2 = 1.5$. So, $\eta_2 = 0.5$, and the positive security loading conditions are satisfied. Under this hypothesis, the solutions of $\mathfrak{h}(s) \det[\mathbf{L}_1(s)] = 0$ are $-R_{11} = -6.0539, -R_{12} = -5.3665, -R_{13} = -2.0000, -R_{14} = -1.5891, -R_{15} = -0.5883, -R_{16} = -0.0127, \rho_{11} = 0.3244, \rho_{12} = 1.2861$, and the solutions of $\mathfrak{h}(s) \det[\mathbf{L}_2(s)] = 0$ are $-R_{21} = -5.3915, -R_{22} = -4.6356, -R_{23} = -2.0000, -R_{24} = -1.5608, -R_{25} = -0.5600, -R_{26} = -0.0079, \rho_{21} = 0.6041, \rho_{22} = 1.5517$.

Differentiating (5.29) w.r.t. u , then letting $u = d_1$ and using $\mathbf{W}'(d_1) = \mathbf{e}_{mn}$, we have $\mathbf{W}'_1(0) = (90.7431, 97.5835)^\top$. Substituting the value of $\mathbf{W}'_1(0)$ into (5.29) and noting that the root $s = -2.0000$ is Singular, we have the expression of $\mathbf{W}(u)$ for $0 \leq u < 2$,

$$(6.1) \quad \mathbf{W}(u) = \begin{pmatrix} -56.9599 & 48.0343 & -0.0078 & -6.5895 & 15.5229 \\ -85.1765 & 78.8988 & -1.2463 & -6.0170 & 13.5409 \end{pmatrix} \begin{pmatrix} e^{-6.0539u} \\ e^{-5.3665u} \\ e^{-1.5891u} \\ e^{-0.5883u} \\ e^{-0.0127u} \end{pmatrix}.$$

Letting the initial conditions $\mathbf{\Xi}'_2(0) = (1, 0)^\top$, $\mathbf{\Xi}_2(0) = (0, 1)^\top$ and $\mathbf{\Xi}'_2(0) = (0, 1)^\top$, $\mathbf{\Xi}_2(0) = (1, 0)^\top$, respectively, by virtue of the asymptotic behaviour of $\mathbf{W}(u)$, $u \geq 2$, we get the following two linearly independent solutions from (5.21) when $u \geq 2$,

$$\Theta_{21}(u) = \begin{pmatrix} 6.5218 & -2.9357 & -0.0069 & -0.0053 & -1.4173 \\ 59.4285 & 1.1493 & -0.0164 & 0.0118 & -1.4228 \end{pmatrix} \begin{pmatrix} e^{-5.3915(u-2)} \\ e^{-4.6356(u-2)} \\ e^{-1.5608(u-2)} \\ e^{-0.5600(u-2)} \\ e^{-0.0079(u-2)} \end{pmatrix},$$

$$\Theta_{22}(u) = \begin{pmatrix} 2.1362 & 21.4218 & -0.0057 & -0.0046 & -3.4694 \\ 19.4654 & -8.3868 & -0.0135 & 0.0103 & -3.4830 \end{pmatrix} \begin{pmatrix} e^{-5.3915(u-2)} \\ e^{-4.6356(u-2)} \\ e^{-1.5608(u-2)} \\ e^{-0.5600(u-2)} \\ e^{-0.0079(u-2)} \end{pmatrix}.$$

Combining (5.5) with (6.1) and utilizing the boundary condition $\mathbf{W}(u)|_{u=d_1^-} = \mathbf{W}(u)|_{u=d_1^+}$, then solving the linear equations, we have $k_{21} = -0.4243$, $k_{22} = -1.7921$. Thus, we obtain $\mathbf{W}(u)$, $u \geq 2$,

$$(6.2) \quad \mathbf{W}(u) = \begin{pmatrix} 50 \\ 50 \end{pmatrix} + \begin{pmatrix} -6.5956 & -37.1448 & 0.0131 & 0.0105 & 6.8190 \\ -60.1005 & 14.5425 & 0.0312 & -0.0235 & 6.8457 \end{pmatrix} \begin{pmatrix} e^{-5.3915(u-2)} \\ e^{-4.6356(u-2)} \\ e^{-1.5608(u-2)} \\ e^{-0.5600(u-2)} \\ e^{-0.0079(u-2)} \end{pmatrix}.$$

Last, since $\boldsymbol{\gamma} = \boldsymbol{\beta} \otimes \boldsymbol{\alpha} = (1/2, 1/2)^\top$, we can obtain $W(u)$ by $W(u) = \boldsymbol{\gamma}^\top \mathbf{W}(u)$, viz,

$$(6.3) \quad W(u) = \begin{cases} -71.0682e^{-6.0539u} + 63.4665e^{-5.3665u} - 0.6271e^{-1.5891u} \\ \quad - 6.3033e^{-0.5883u} + 14.5319e^{-0.0127u}, & 0 \leq u < 2, \\ 50 - 33.3481e^{-5.3915(u-2)} - 11.3011e^{-4.6356(u-2)} \\ \quad + 0.0222e^{-1.5608(u-2)} - 0.0065e^{-0.5600(u-2)} \\ \quad + 6.8323e^{-0.0079(u-2)}, & u \geq 2. \end{cases}$$

Acknowledgements

This work was supported by the National social Science Foundation of China (17BJY206).

References

- [1] Albrecher, H. Claramunt, M.M. and Mármol, M. *On the distribution of dividend payments in a Sparre Andersen model with generalized Erlang(n) interclaim times*, Insurance: Mathematics and Economics **37**, 324-334, 2005.
- [2] Asmussen, S. and Albrecher, H. *Ruin probabilities*, second ed. World Scientific, New Jersey, 2010.
- [3] Chueng, E.C.K. and Landriault, D. *Perturbed MAP risk models with dividend barrier strategies*, Journal of Applied Probability **46**, 521-541, 2009.
- [4] Dickson, D.C.M. and Hipp, C. *On the of ruin for Erlang(2) risk processes*, Insurance: Mathematics and Economics **29**, 333-344, 2001.
- [5] Gao, H. and Yin, C. *The perturbed Sparre Andersen model with a threshold dividend strategy*, Journal of Computational and Applied Mathematics **220**, 394-408, 2008.
- [6] Gerber, H.U. and Shiu, E.S.W. 2005. *The time value of ruin in a Sparre Andersen model*, North American Actuarial Journal **9**, 49-69, 2005.
- [7] Ji, L. and Zhang, C. *The Gerber-Shiu penalty functions for two classes of renewal risk processes*, Journal of Computational and Applied Mathematics **233**, 2575-2589, 2010.
- [8] Jiang, W.Y. and Yang, Z.J. *The phase-type risk model perturbed by diffusion under a threshold dividend strategy*, Acta Mathematicae Applicatae Sinica English Series **29**, 215-224, 2013.
- [9] Jiang, W.Y. and Yang, Z.J. *The expected discounted penalty function for two classes of risk processes perturbed by diffusion with multiple thresholds*, Indian Journal of Pure and Applied Mathematics **45**, 479-295, 2014.
- [10] Li, S. *The distribution of the dividend payments in the compound Poisson risk model perturbed by diffusion*, Scandinavian Actuarial Journal **26**, 73-85, 2006.
- [11] Li, S. and Lu, Y. *On the expected discounted penalty functions for two classes of risk processes*, Insurance: Mathematics and Economics **36**, 179-193, 2005.
- [12] Lin, X.S. and Sendova, K.P. *The compound Poisson risk model with multiple thresholds*, Insurance: Mathematics and Economics **42**, 617-627, 2008.
- [13] Lu, Y. and Li, S. *The Markovian regime-switching risk model with a threshold dividend strategy*, Insurance: Mathematics and Economics **44**, 296-303, 2009.
- [14] Yang, H. and Zhang, Z.M. *Gerber-Shiu discounted penalty function in a Sparre Andersen model with multi-layer dividend strategy*, Insurance: Mathematics and Economics **42**, 984-991, 2008.
- [15] Zhang, Z.M. Li, S. and Yang, H. *The Gerber-Shiu discounted penalty functions for a risk model with two classes of claims*, Journal of Computational and Applied Mathematics **230**, 643-655, 2009.
- [16] Wan, N. *Dividend payments with a threshold dividend strategy in the compound Poisson risk model perturbed by diffusion*, Insurance: Mathematics and Economics **40**, 509-532, 2007.

Variable selection for high dimensional partially linear varying coefficient errors-in-variables models

Zhaoliang Wang^{*†} and Liugen Xue[‡]

Abstract

In this paper, we consider variable selection procedure for the high dimensional partially linear varying coefficient models where the parametric part covariates are measured with additive errors. The penalized bias-corrected profile least squares estimators are conducted, and their asymptotic properties are also studied under some regularity conditions. The rate of convergence and the asymptotic normality of the resulting estimates are established. We further demonstrate that, with proper choices of the penalty functions and the regularization parameter, the resulting estimates perform asymptotically as well as an oracle property. Choice of smoothing parameters is also discussed. Finite sample performance of the proposed variable selection procedures is assessed by Monte Carlo simulation studies.

Keywords: High dimensionality, Measurement error, Semiparametric models, Local linear regression, Variable selection.

2000 AMS Classification: 62G05, 62G20

Received : 09.06.2017 *Accepted :* 28.06.2018 *Doi :* 10.15672/HJMS.2018.603

1. Introduction

With the development of applied sciences, semiparametric regression models have been well researched and popularly used for their flexibility and interpretability. [16] present diverse semiparametric regression models along with their inference procedures and applications. Of particular interests to us in this paper is the partially linear varying coefficient (PLVC) model. Let $\{(Y_i, X_i, Z_i, T_i), i = 1, \dots, n\}$ be an iid copies of (Y, X, Z, T) , where

^{*}School of Mathematics and Information Science, Henan Polytechnic University and College of Applied Sciences, Beijing University of Technology, Email: wangzhaolinag@hpu.edu.cn

[†]Corresponding Author.

[‡]College of Applied Sciences, Beijing University of Technology, Email: lgxue@bjut.edu.cn

Y is a scalar response variable and $(X, Z, T) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}$ is its associated regressors. The PLVC models take the form

$$(1.1) \quad Y_i = X_i^\top \beta + Z_i^\top \alpha(T_i) + \varepsilon_i,$$

where $\beta = (\beta_1, \dots, \beta_p)^\top$ is a p -dimensional vector of unknown parameters, $\alpha(\cdot) = (\alpha_1(\cdot), \dots, \alpha_q(\cdot))^\top$ is a q -dimensional vector of unknown coefficient functions, and ε_i 's are iid model error with $E(\varepsilon_i | X_i, Z_i, T_i) = 0$. In this model, the dependence of $\alpha(\cdot)$ on T implies a special kind of interaction between the covariate Z and T . Due to the curse of dimensionality, we assume, for simplicity, that T is univariate. This model presents a novel and general structure, which indeed covers many well-studied, important semiparametric regression models, e.g. linear model, partially linear model and varying coefficient model.

Model (1.1) has been studied by many authors recently. Examples include but are not limited to [1, 26, 13, 12, 10, 3, 23]. An essential assumption in their papers is that all data can be observed directly. However, measurement error data are often encountered in many fields, including engineering, economics, biomedical sciences and epidemiology. Simply ignoring measurement errors, known as the naive method, will result in biased estimators. There is a long standing literature on statistical modeling subject to measurement errors. Comprehensive reviews can be found in [2, 7]. PLVC models have been used to study measurements with errors, see, for instance, [21, 8, 20, 19, 6].

Concerns about model bias often prompt us to build models that contain many variables, especially when the sample size becomes large. A reasonable way to capture such a tendency is to consider the situation where the dimension of the parameter increases along with the sample size. On the other hand, to enhance predictability and to select significant variables is practically interesting, but is always a tricky task for data analysis. When the number of covariates is large, traditional variable selection methods such as stepwise regression and best subset selection is computationally infeasible and statistical properties of the estimators are difficult to analyze, as argued in [14], this is part of the reason why penalization based method (e.g., Lasso [17], Elastic net [28], Adaptive Lasso [27], SCAD [4], MCP [22], among others) has gained popularity in recent years. There has been much work on variable selection for semiparametric regression models. In particular, examples for fixed dimensional PLVC models include [25, 24, 11, 18] and references therein.

In these studies, however, high dimensional vector X , variable selection in X and measurement error problem were not considered at the same time. The goal of this paper intends to develop an unified estimation and variable selection method for high dimensional PLVC errors-in-variables models. To be precise, we allow $p \rightarrow \infty$ as the sample size $n \rightarrow \infty$ and denote it by p_n whenever necessary, but q is a fixed and finite integer in (1.1). In addition, the covariate X is measured with additive errors, while Z and T are errors free. More specific, we cannot observe X_i but we can observe W_i with

$$(1.2) \quad W_i = X_i + U_i,$$

and U_i 's are iid measurement error, which is independent of $(X_i, Z_i, T_i, \varepsilon_i)$, and has mean zero and the known covariance $\text{Cov}(U_i) = \Sigma_U$ (for simplicity). If Σ_U is unknown, its estimation usually requires multiple observations of W or instrumental variables, see [15] for details. We term (1.1) and (1.2) with PLVCE models. To our best knowledge, variable selection for PLVCE models with high dimension has not been systematically studied yet.

We propose penalized bias-corrected profile least squares estimator and systematically study the asymptotic properties of the estimators. It is worth pointing out that theoretic results in this paper provide explicit results on the asymptotic properties under the setting in which both the dimension of the true non-zero components of β and the total length of β tend to infinity as n goes to infinity. This resonates with the perspective that a more complex statistical model can be fit when more data are collected. The issue of a diverging number of parameters has also been considered in [5] in the context of penalized likelihood. This advances the results in current literature, where estimation and inference are studied only for fixed finite dimensional parameters for measurement error models. We demonstrate how the convergence rate of the resulting estimator depends on the regularization parameter. Furthermore, with a proper choice of the regularization parameters and the penalty function, we show that this variable selection procedure is consistent, and the regularized estimators of the regression coefficients have oracle property. This indicates that the penalized estimators work as well as if the subset of true zero coefficients were already known. In addition, we address issues of practical implementation of the proposed methodology. Monte-Carlo simulation studies are conducted to assess finite sample performance.

The rest of this paper is organized as follows. A variable selection procedure for PLVCE models is proposed in Section 2, assumptions and the asymptotic properties of the proposed estimators are given in this section. We give the computational algorithms and discuss the selections of tuning parameters in Section 3. In Section 4, some simulations are conducted to illustrate the performance of our methodology. Given in Section 5 are conclusions. All technical proofs are relegated to Section 6.

Notation: The gradient and hessian matrix of a function $f(x)$ are denoted by $\nabla f(x)$ and $\nabla^2 f(x)$ respectively. We write $\|f\|_2$ and $\|f\|_\infty$ for the L_2 and sup norm of a function f , respectively. The L_q norm of a p -vector v is defined as $\|v\|_q = (\sum_{j=1}^p |v_j|^q)^{1/q}$ for $q \geq 1$ with $\|v\|_\infty = \max_{1 \leq j \leq p} |v_j|$, and $\|v\|_0 = |\text{supp}(v)|$ where $\text{supp}(v) = \{j : v_j \neq 0\}$ and $|S|$ is the cardinality of a set S . Let $M_{i\cdot}$, $M_{\cdot j}$ and M_{ij} be the i th row, j th column and (i, j) entry of the matrix M , respectively. Let $\|M\|_q = \sup_{\|v\|_q=1} \|Mv\|_q$ be the matrix L_q operator norm. We use $\|\cdot\|$ as a shorthand for $\|\cdot\|_2$. We use c and C to denote generic positive constants that may vary from place to place. Moreover, the operator \xrightarrow{P} denotes convergence in probability, and \xrightarrow{D} denotes convergence in distribution.

2. Methods and results

2.1. Penalized bias-corrected profile least squares estimator. As in [3], if X_i is observable we can apply the profile least squares estimation to estimate the parametric component and apply the local polynomial estimation to estimate the nonparametric component. Profile least squares is a useful approach and will be showed to be semi-parametrically efficient for model (1.1). When $\varepsilon_i \sim N(0, \sigma^2)$, the approach becomes profile likelihood estimation. For the paper to be self-contained, we summarize the main ingredients as follows. If β is known, (1.1) can be written as

$$(2.1) \quad Y_i - X_i^\top \beta = Z_i^\top \alpha(T_i) + \varepsilon_i,$$

which can be treated as a varying coefficient model. Thus, we may apply a local linear regression technique to estimate the varying coefficient functions $\{\alpha_j(\cdot), j = 1, \dots, q\}$. For T in a small neighbourhood of t , approximate each $\alpha_j(t)$ by $\alpha_j(T) \approx \alpha_j(t) + \alpha_j'(t)(T - t)$, $j = 1, \dots, q$. This leads to the following weighted local least-squares problem: find

$\alpha_j(t), \alpha'_j(t)$ to minimize

$$(2.2) \quad \sum_{i=1}^n \left[Y_i - X_i^\top \beta - \sum_{j=1}^q Z_{ij} \{ \alpha_j(t) + \alpha'_j(t)(T_i - t) \} \right]^2 K_h(T_i - t),$$

where $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ is a kernel function and h is a bandwidth.

For the sake of descriptive convenience, we denote $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, write $\mathbf{X}, \mathbf{Z}, \boldsymbol{\varepsilon}$ in a similar fashion. Let $\omega_t = \text{diag}\{K_h(T_1 - t), \dots, K_h(T_n - t)\}$ and

$$D_t = \begin{pmatrix} Z_1 & \cdots & Z_n \\ \frac{T_1-t}{h} Z_1 & \cdots & \frac{T_n-t}{h} Z_n \end{pmatrix}^\top.$$

It is easy to show that the minimizers of (2.2) are given by

$$(\tilde{\alpha}(t)^\top, h\tilde{\alpha}'(t)^\top)^\top = \{D_t^\top \omega_t D_t\}^{-1} D_t^\top \omega_t (\mathbf{Y} - \mathbf{X}\beta).$$

This solutions depend on β implicitly. Then we can estimate $\alpha(t)$, when β is given, by

$$(2.3) \quad \tilde{\alpha}(t; \beta) = (I_{q \times q}, 0_{q \times q}) \{D_t^\top \omega_t D_t\}^{-1} D_t^\top \omega_t (\mathbf{Y} - \mathbf{X}\beta),$$

where $I_{q \times q}$ denote the q by q identity matrix, and $0_{q \times q}$ denote a q by q matrix of zeros. Substituting $\tilde{\alpha}(t; \beta)$ into model (2.1), we can obtain the profile least square estimator of β by the following regression problem

$$(2.4) \quad \tilde{\beta} = \arg \min_{\beta} \frac{1}{2} \sum_{i=1}^n (Y_i - X_i^\top \beta - Z_i^\top \tilde{\alpha}(T_i; \beta))^2$$

Moreover, plug $\tilde{\beta}$ into (2.3), the estimators of $\alpha(t)$ can be obtained, see [3] for details.

However, in our case, X_i cannot be exactly observed. If one ignores the measurement error and replaces X_i by W_i in (2.4), one can show that the resulting estimator is inconsistent. By the correction for attenuation technique as in [21], the bias-corrected profile least squares estimator of β can be defined by minimizing

$$(2.5) \quad \hat{L}_n(\beta) = \frac{1}{2} \sum_{i=1}^n (Y_i - W_i^\top \beta - Z_i^\top \hat{\alpha}(T_i, \beta))^2 - \frac{n}{2} \beta^\top \Sigma_U \beta,$$

where $\hat{\alpha}(T_i, \beta)$ is obtained by replace \mathbf{X} with \mathbf{W} in the right hand side of (2.3). The second term is included to correct the bias in the squared loss function due to measurement error.

In high dimensional data analysis, to perform variable selection and estimation simultaneously, based on (2.5) we propose the penalized bias-corrected profile least squares function defined as

$$(2.6) \quad \hat{Q}_n(\beta) = \hat{L}_n(\beta) + n \sum_{j=1}^{p_n} p_\lambda(|\beta_j|),$$

where $p_\lambda(\cdot)$ is a prespecified penalty function with a tuning parameter λ , which may be chosen by a data-driven method. It is worth noting that the penalty functions and the tuning parameters are not necessarily the same for all coefficients. For instance, we want to keep important variables in the final model, and therefore we should not penalize their coefficients. For ease of presentation, we assume that the penalty functions and the regularization parameters are the same for all coefficients in this paper.

The choice of the penalty functions has been studied in [4] in depth. A property of (2.6) is that with a proper choice of penalty functions, such as the SCAD and Lasso penalty, the resulting estimate contains some exact zero coefficients. This is equivalent to excluding the corresponding variables from the final selected model, thus variable selection is achieved at the same time as parameter estimation. Solving for $\hat{\beta}$ from (2.6)

gives the estimate of β . Moreover, the fact that $E(Y_i - X_i^\top \beta | T_i) = E(Y_i - W_i^\top \beta | T_i)$ suggests us to estimate $\alpha(\cdot)$ by

$$(2.7) \quad \hat{\alpha}(t) = (I_{q \times q}, 0_{q \times q}) \{D_t^\top \omega_t D_t\}^{-1} D_t^\top \omega_t (\mathbf{Y} - \mathbf{W} \hat{\beta}).$$

2.2. Asymptotic properties. In this subsection we consider the large sampling properties of the proposed estimator. For convenience of notation, we assume the true value $\beta^* = (\beta_I^{*\top}, \beta_{II}^{*\top})^\top$, where β_I^* consists of all nonzero components of β^* and $\beta_{II}^* = 0$. Let s_n denote the dimension of β_I^* . Furthermore, denote

$$B = (p'_{\lambda_n}(|\beta_1^*|) \text{sign}(\beta_1^*), \dots, p'_{\lambda_n}(|\beta_{s_n}^*|) \text{sign}(\beta_{s_n}^*))^\top \quad \text{and} \\ \Sigma_{\lambda_n} = \text{diag}\{p''_{\lambda_n}(|\beta_1^*|), \dots, p''_{\lambda_n}(|\beta_{s_n}^*|)\},$$

where we write λ as λ_n to emphasize its dependence on the sample size n . To give the asymptotic results, here are regularity conditions required.

(C1) The random variable T has a bounded support \mathcal{T} . Its density function $f_T(t)$ is Lipschitz continuous and bounded away from 0 on \mathcal{T} .

(C2) The $q \times q$ matrix $E(ZZ^\top | T)$ is nonsingular for each $T \in \mathcal{T}$. $E(XX^\top | T)$, $E(ZZ^\top | T)$ and $E(XZ^\top | T)$ are all Lipschitz continuous.

(C3) There is an $\kappa > 2$ such that $E\|X\|^{2\kappa} < \infty$, $E\|Z\|^{2\kappa} < \infty$, $E\|\varepsilon\|^{2\kappa} < \infty$ and $E\|U\|^{2\kappa} < \infty$, and for some $\delta < 2 - \kappa^{-1}$ there is $n^{2\delta-1}h \rightarrow \infty$ as $n \rightarrow \infty$.

(C4) All of the coefficient functions $\{\alpha_j(\cdot), j = 1, \dots, q\}$ are Lipschitz continuous and have continuous second order derivatives on \mathcal{T} .

(C5) The function $K(\cdot)$ is a symmetric density function with compact support and the bandwidth h satisfies $nh^8 \rightarrow 0$ and $nh^2/(\log n)^2 \rightarrow \infty$ as $n \rightarrow \infty$.

(C6) $\min\{|\beta_j^*|, j = 1, \dots, s_n\}/\lambda_n \rightarrow \infty$ as $n \rightarrow \infty$.

(C7) There exist constant c and C such that $0 < c < \Lambda_{\min}(\Sigma_1) < \Lambda_{\max}(\Sigma_1) < C < \infty$ for all n , where $\Lambda_{\min}(M)$ and $\Lambda_{\max}(M)$ denote respectively the smallest and largest eigenvalues of symmetric matrix M .

(P1) Let $a_n = \max_{1 \leq j \leq p_n} \{p'_{\lambda_n}(|\beta_j^*|), \beta_j^* \neq 0\}$ and $b_n = \max_{1 \leq j \leq p_n} \{p''_{\lambda_n}(|\beta_j^*|), \beta_j^* \neq 0\}$. Assume that $a_n = O(n^{-1/2})$ and $b_n \rightarrow 0$ as $n \rightarrow \infty$. In addition, there exist constants c and C such that, when $\theta_1, \theta_2 \geq c\lambda_n$, $|p''_{\lambda_n}(\theta_1) - p''_{\lambda_n}(\theta_2)| \leq C|\theta_1 - \theta_2|$.

(P2) $\liminf_{n \rightarrow \infty} \liminf_{\theta \rightarrow 0^+} p'_{\lambda_n}(\theta)/\lambda_n > 0$.

These conditions, while a little bit lengthy at first look, are actually quite mild and may be further relaxed. Conditions (C1)–(C5) are also used by [3]. Conditions (C6)–(C7) and (P1)–(P2) are adopted from [5], see [5] for details. Condition (C6) gives the rate at which the penalized estimator can distinguish nonvanishing parameters from 0, which is necessary for obtaining the oracle property. In the finite-parameter situation this condition is implicitly assumed, and is in fact stronger than that imposed here. Condition (C7) assumes that the Σ_1 is positive definite and its eigenvalues are uniformly bounded. Conditions (P1)–(P2) are regularity conditions on penalty function.

The following theorem demonstrates that the convergence rate for the penalized bias-corrected estimator depends on the penalty function and the regularization parameter λ_n through a_n .

2.1. Theorem. (*Existence*) Suppose the penalty function satisfies condition (P1). Under regularity conditions (C1)–(C5), if $\lambda_n \rightarrow 0$ and $p_n^4/n \rightarrow 0$ as $n \rightarrow \infty$, then with probability tending to 1, there is a local minimizer $\hat{\beta}$ of (2.6) such that $\|\hat{\beta} - \beta^*\| = O_P\{\sqrt{p_n}(n^{-1/2} + a_n)\}$.

The proof of this theorem is given in Section 6. As it can be seen from the statement of Theorem 2.1, it requires that λ_n and the penalty function must be chosen such that $a_n = O(n^{-1/2})$ to achieve $\sqrt{n/p_n}$ convergence rate (or \sqrt{n} convergence rate for finite and

fixed p). For the L_1 penalty, $a_n = \lambda_n$. Thus, the $\sqrt{n/p_n}$ convergence rate requires that $\lambda_n = O(n^{-1/2})$. This requirement will make it difficult to choose λ_n in practice. However, if condition (C6) is satisfied, it is clear that $a_n = 0$ as when n is large enough for the SCAD penalty. Thus, the resulting estimator is $\sqrt{n/p_n}$ consistent, and no requirements are imposed on the convergence rate of λ_n . Note that the optimal bandwidth $h = O(n^{-1/5})$ is included in Theorem 2.1. Hence $\sqrt{n/p_n}$ -consistency is achieved without the need of undersmoothing of the nonparametric component.

2.2. Theorem. (*Oracle property*). *Suppose the penalty function satisfies conditions (P1)-(P2). Under regularity conditions (C1)-(C7), if $\lambda_n \rightarrow 0$, $p_n^5/n \rightarrow 0$ and $\sqrt{n/p_n}\lambda_n \rightarrow \infty$ as $n \rightarrow \infty$, then with probability tending to 1, the $\sqrt{n/p_n}$ -consistent local minimizer $\hat{\beta} = (\hat{\beta}_I^\top, \hat{\beta}_{II}^\top)^\top$ in Theorem 2.1 must satisfy: (i) (Sparsity) $\hat{\beta}_{II} = 0$; (ii) (Asymptotic normality) Let A_n be a deterministic $l \times s_n$ matrix with l fixed and $A_n A_n^\top \rightarrow G$, a positive definite matrix. Then*

$$\sqrt{n} A_n \Sigma_{2I}^{-1/2} \{ \Sigma_{1I} + \Sigma_{\lambda_n} \} [\hat{\beta}_I - \beta_I^* + \{ \Sigma_{1I} + \Sigma_{\lambda_n} \}^{-1} B] \xrightarrow{D} N(0_l, G),$$

where Σ_{1I} and Σ_{2I} are the top left-hand $s_n \times s_n$ submatrix of Σ_1 and Σ_2 , respectively.

Theorem 2.2 is proved in Section 6. It is easy to see that sparsity and asymptotic normality are still valid when the number of parameter diverges in PLVCE models. For some penalty functions, including the SCAD penalty, B and Σ_{λ_n} are zero when n is large enough. Hence the results in Theorem 2.2 imply that the proposed procedure has the celebrated oracle property, i.e., $\hat{\beta}_{II} = 0$ and $\sqrt{n} A_n \Sigma_{2I}^{-1/2} \Sigma_{1I} (\hat{\beta}_I - \beta_I^*) \xrightarrow{D} N(0_l, G)$. On the other hand, the L_q penalty, $q \geq 1$, cannot simultaneously satisfy the conditions $\lambda_n = O_P(n^{-1/2})$ and $\sqrt{n/p_n}\lambda_n \rightarrow \infty$ as $n \rightarrow \infty$. These penalty functions cannot produce estimators with the oracle property. The L_q penalty, $q < 1$, may satisfy these two conditions at same time, but the bias term in Theorem 2.2(ii) cannot be ignored.

To make statistical inference on β_I^* , we need to estimate the standard error of the estimator of $\hat{\beta}_I$. The standard errors for estimated parameters can be obtained directly because we are estimating parameters and selecting variables at the same time. From Theorem 2.2, we can further approximate the estimation variance of the resulting estimator by the sandwich formula. Namely

$$(2.8) \quad \frac{1}{n} \{ \hat{\Sigma}_{1I} + \Sigma_{\lambda_n}(\hat{\beta}_I) \}^{-1} \hat{\Sigma}_{2I} \{ \hat{\Sigma}_{1I} + \Sigma_{\lambda_n}(\hat{\beta}_I) \}^{-1},$$

where $\hat{\Sigma}_{1I}$, a consistent estimate of Σ_{1I} , is defined as

$$\hat{\Sigma}_{1I} = \frac{1}{n} \nabla^2 \hat{L}_{nI}(\hat{\beta}_I) = \frac{1}{n} \sum_{i=1}^n \left(W_{Ii} + \frac{\partial \hat{\alpha}(T_i; \hat{\beta}_I)}{\partial \beta_I} Z_i \right)^{\otimes 2} - \Sigma_{UI},$$

and $\hat{\Sigma}_{2I} = \text{Cov}(\nabla \hat{L}_{nI}(\hat{\beta}_I))$ is given by

$$\frac{1}{n} \sum_{i=1}^n \left\{ (Y_i - W_{Ii}^\top \hat{\beta}_I - Z_i^\top \hat{\alpha}(T_i; \hat{\beta}_I)) (W_{Ii} + \frac{\partial \hat{\alpha}(T_i; \hat{\beta}_I)}{\partial \beta_I} Z_i) + \Sigma_{UI} \hat{\beta}_I \right\}^{\otimes 2},$$

furthermore, $\Sigma_{\lambda_n}(\hat{\beta}_I)$ is obtained by replacing β_I^* by $\hat{\beta}_I$ in Σ_{λ_n} .

The consistency of the proposed sandwich formula can be shown by using similar techniques as in [5]. The accuracy of this sandwich formula will be tested in our simulation studies.

3. Issues in practical implementation

In this section, we present a computational algorithm for obtaining the estimator and selection methods for the tuning parameters.

3.1. Computational algorithm. Since some penalty functions such as the SCAD penalty and $L_q, 0 \leq q \leq 1$ penalty are singular at the origin, it is challenging to minimize the penalized bias-corrected least squares function of (2.6). Following the idea of [4], we apply iterative algorithm based on the local quadratic approximation (LQA) of the penalty function. More specifically, suppose that at the k th step of the iteration, we obtain the value $\hat{\beta}^{(k)}$ that is close to the true value β^* . If $\hat{\beta}_j^{(k)}$ is very close to 0, then set $\hat{\beta}_j^{(k+1)} = 0$, and exclude the corresponding covariate from the model. Otherwise, an approximation of the penalty function at value $\hat{\beta}_j^{(k)}$ can be given by

$$p_\lambda(|\beta_j|) \approx p_\lambda(|\hat{\beta}_j^{(k)}|) + \frac{1}{2} \frac{p'_\lambda(|\hat{\beta}_j^{(k)}|)}{|\hat{\beta}_j^{(k)}|} (\beta_j^2 - \hat{\beta}_j^{(k)2}),$$

Consequently, with a slight abuse of notation, removing irrelevant terms we update the estimate of β repeatedly until convergence with

$$(3.1) \quad \beta^{(k+1)} = \arg \min_{\beta} \left\{ \widehat{L}_n(\beta) + \frac{n}{2} \beta^\top \Sigma_{\lambda_n}^{\text{LQA}}(\hat{\beta}^{(k)}) \beta \right\},$$

where $\Sigma_{\lambda_n}^{\text{LQA}}(\hat{\beta}^{(k)}) = \text{diag}\{p'_{\lambda_n}(|\hat{\beta}_1^{(k)}|)/|\hat{\beta}_1^{(k)}|, \dots, p'_{\lambda_n}(|\hat{\beta}_{p_n}^{(k)}|)/|\hat{\beta}_{p_n}^{(k)}|\}$. Hence, the foregoing discussion leads to the following iterating algorithm:

- Step 1. Given an initial estimate $\hat{\beta}^{(0)}$.
- Step 2. Update $\hat{\beta}^{(1)}$ by (3.1).
- Step 3. Set $\hat{\beta}^{(0)} = \hat{\beta}^{(1)}$. Iterate Step 1 and 2 until convergence, and denote the final estimator $\hat{\beta}$.

In the initialization step, the initial estimators do not affect the degree of sparsity of the solution and the accuracy of the final estimator, but they will affect the speed of convergence of our iterative algorithm. In the following simulations, we obtain an initial estimator using a bias-corrected ordinary least-squares method based on (2.5). The simulation results show that such a choice is workable. During the iterations, to avoid numerical instability we need to keep track of zero coefficients and modify the penalty terms accordingly once $|\hat{\beta}_j^{(0)}|$ drops below a certain threshold ϵ ($\epsilon = 10^{-4}$ in our implementation). Specifically, in Step 2, if $|\hat{\beta}_j^{(0)}| < \epsilon$, then set $\hat{\beta}_j^{(1)} = 0$, delete the j th component of the covariates from the iteration.

3.2. Tuning parameters selection. To implement the proposed method, the bandwidth h and the tuning parameters λ_n in the penalty functions should be chosen. It is desirable to have automatic, data-driven methods to select h and λ_n .

Bandwidth selection. Condition (C5) reveal the rate of h . Any bandwidths with this rate lead to the same limiting distribution for $\hat{\beta}$. Therefore, the bandwidth selection can be done in a standard routine. For simple calculation, the bandwidth h is taken to be $h = 0.5n^{-1/5}$ in this paper, which we find to work satisfactorily in a variety of setting. We also conduct a sensitivity analysis by shifting bandwidths around the selected values, and found that the results are stable. Thus, the simulation results are not sensitive to the choice of h within certain range.

Regularization parameters selection. Here, given h , we use the "leave one sample out" method to select the tuning parameter λ_n . This method has been widely applied in

practice. The cross-validation score for λ_n is defined as

$$(3.2) \quad \text{CV}(\lambda_n) = \sum_{i=1}^n (Y_i - W_i^\top \hat{\beta}^{-i} - Z_i^\top \hat{\alpha}^{-i}(T_i))^2 - \sum_{i=1}^n (\hat{\beta}^{-i})^\top \Sigma_U \hat{\beta}^{-i}$$

where $\hat{\beta}^{-i}$ is the solution based on (2.6) after deleting the i th observation, and $\hat{\alpha}^{-i}(T_i)$ is the estimator defined in (2.7) with $\hat{\beta}$ replaced by $\hat{\beta}^{-i}$. The CV tuning parameter λ_n^{CV} is selected to minimize (3.2), that is, $\lambda_n^{\text{CV}} = \arg \min_{\lambda_n} \text{CV}(\lambda_n)$.

We also can use any other appropriate selection method to select the tuning parameters such as GCV, AIC and BIC. However, the definition of the degrees of freedom for the effective parameters in our variable selection procedure poses great challenges. Then, it is inconvenient to use such selection criteria for our variable selection procedure. In addition, from our simulation experience, we found that the CV method used in this paper works well. Further study of the asymptotic property of the proposed tuning parameter selection is needed, but it is outside the scope of this paper.

4. Simulation studies

In this section we corroborate our theoretical results with numerical experiments on synthetic data examples. That is, we conduct simulations to evaluate the finite sample performance of the proposed methods. We focus on only the SCAD penalty and referred to the proposed procedure as C_{SCAD} . The C_{SCAD} is compared with four alternative procedures as follows. The first is the naive penalized procedures with a direct replacement of X by W ignoring measurement error (N_{SCAD}). The second is the estimators with considering measurement errors, but not penalized for complexity (Full). As a benchmark, two oracle methods in which the nonzero subset of slope β were known are implemented. In particular, the first (Oracle₁) serves as the gold standard, in which X can be observed. The second (Oracle₂) is another type, in which using W based on bias-corrected due to measurement errors.

We simulate data from model (1.1) and (1.2) with $q = 2$ and $p_n = \lfloor 1.8n^{1/3} \rfloor$ where $\lfloor k \rfloor$ denote the largest integer not greater than k , in which $\alpha_1(t) = 2 \sin(2\pi t)$ and $\alpha_2(t) = 16t(1-t) - 2$, and $\beta = (2, -1.5, 4, 0, \dots, 0)^\top$. Thus the first $s_n = 3$ regression variables were significant, but the remaining were not. The rate $p_n = \lfloor 1.8n^{1/3} \rfloor$ is not the same as presented in the theorems in Section 2, but we use this to show the capability of handling a higher rate of parameters growth for proposed method. The index variable T is sampled uniformly on $[0, 1]$. The covariates (X, Z) are taken from multivariate normal distribution $N_{p_n+q}(0, \Sigma)$. We consider Toeplitz covariance matrices $\Sigma_{ij} = \varrho^{|i-j|}$, in which both independent ($\varrho = 0$) and correlated cases ($\varrho = 0.5$) are taken into account. Y is generated according to the model, where noise term $\varepsilon \sim N(0, \sigma^2)$, and two different value $\sigma^2 = 0.5$ and 1, which represent strong and weak signal-to-noise ratios, were considered. Moreover, we assume that measurement error $U \sim N(0, \sigma_U^2 I_{p_n})$, where we take $\sigma_U = 0.2$ and 0.4 to represent different level of measurement errors. We perform 1000 simulations for all configurations with sample size $n = 100$ and $n = 400$ respectively. In all simulations, as a commonly adopted strategy we use the Epanechnikov kernel function $K(t) = 0.75(1 - t^2)_+$.

To assess the performance of different methods, we adopt the following criteria. For model error, the performance of estimator $\hat{\beta}$ will be assessed by using the generalized mean square error (GMSE), defined as

$$\text{GMSE} = (\hat{\beta} - \beta^*)^\top (EWW^\top - \Sigma_U)(\hat{\beta} - \beta^*).$$

Table 1. Simulation results with different methods for $\sigma^2 = 1$ over 1000 repetitions

ρ	Method	$\sigma_U = 0.2$					$\sigma_U = 0.5$				
		True	C	IC	GMSE	RASE	True	C	IC	GMSE	RASE
$(n, p_n) = (100, 8)$											
0	C _{SCAD}	99.9	5.000	0.001	0.072	0.556	91.0	4.970	0.063	0.546	1.030
	N _{SCAD}	99.6	5.000	0.004	0.100	0.550	80.0	5.000	0.205	1.377	0.897
	Full	0.00	0.020	0.000	0.188	0.571	0.00	0.016	0.000	1.060	1.107
	Oracle ₁	100	5	0	0.033	0.551	100	5	0	0.033	0.950
	Oracle ₂	100	5	0	0.071	0.556	100	5	0	0.399	1.018
0.5	C _{SCAD}	99.6	5.000	0.004	0.076	0.621	69.3	4.447	0.078	2.051	1.448
	N _{SCAD}	88.7	5.000	0.114	0.244	0.622	1.70	4.999	1.112	2.071	0.985
	Full	0.00	0.008	0.000	0.202	0.647	0.00	0.006	0.000	2.801	1.638
	Oracle ₁	100	5	0	0.033	0.614	100	5	0	0.035	1.075
	Oracle ₂	100	5	0	0.071	0.620	100	5	0	0.607	1.220
$(n, p_n) = (400, 13)$											
0	C _{SCAD}	100	10.00	0.000	0.015	0.274	99.9	10.00	0.001	0.072	0.469
	N _{SCAD}	100	10.00	0.000	0.046	0.271	97.7	10.00	0.023	0.982	0.430
	Full	0.0	0.12	0.000	0.067	0.277	0.0	0.055	0.000	0.300	0.481
	Oracle ₁	100	10	0	0.008	0.273	100	10	0	0.007	0.463
	Oracle ₂	100	10	0	0.015	0.274	100	10	0	0.071	0.469
0.5	C _{SCAD}	99.9	9.999	0.000	0.016	0.314	99.2	10.00	0.009	0.099	0.548
	N _{SCAD}	98.7	10.00	0.013	0.085	0.310	0.0	10.00	1.008	1.873	0.475
	Full	0.0	0.090	0.000	0.073	0.321	0.0	0.032	0.000	0.384	0.589
	Oracle ₁	100	10	0	0.007	0.313	100	10	0	0.007	0.534
	Oracle ₂	100	10	0	0.016	0.314	100	10	0	0.087	0.548

The performance of estimator $\hat{\alpha}(\cdot)$ will be assessed by using the square root of average errors (RASE)

$$\text{RASE} = \left\{ N_{\text{grid}}^{-1} \sum_{k=1}^{N_{\text{grid}}} \|\hat{\alpha}(t_k) - \alpha(t_k)\|^2 \right\}^{1/2},$$

over $N_{\text{grid}} = 200$ grid points $\{t_k\}$. Table 1 presents the mean of GMSE and RASE over the 1000 simulations. For the selected model, the model complexity is summarized in terms of the number of zero coefficients for the parametric components, as also reported in Table 1. In Table 1, the column labeled "C" is the average numbers of zero coefficients correctly estimated to be zero, and the column labeled "IC" depicts the average numbers of nonzero coefficients erroneously set to zero. Furthermore, the column labeled "True" is the proportion of times the true model is exactly identified.

From Table 1, we can make the following observations: (i) The performances of both C_{SCAD} and N_{SCAD} procedures become better in terms of model error and model complexity as the level of measurement error decreases. (ii) Both variable selection procedures perform very similarly when the level of measurement error is small. However, when the level of measurement error is large, the performance of C_{SCAD} is significantly better than that of N_{SCAD}. The latter cannot eliminate some unimportant variables and gives larger model errors. This implies that the estimators based on the N_{SCAD} procedure are biased. (iii) In addition, as expected, the performance of the Oracle₁ procedure is best in all cases in terms of model error. Furthermore, the performance of C_{SCAD} becomes increasingly closer to that based on the Oracle₂ procedure as the level of measurement

error decreases or ϱ decreases. (iv) As the sample size increases, the performance of all methods becomes better. To save space the simulation results, for others settings with $\sigma^2 = 0.5$, are not showed here. The above conclusions can also be drawn similarly except now all approaches perform better than they done when $\sigma^2 = 1$ as presented in Table 1. These findings imply that the model selection result based on the C_{SCAD} approach is satisfactory and the selected model is very close to the true model in terms of nonzero coefficients.

Table 2. Bias and standard deviations of estimators for $\sigma^2 = 1$, $\sigma_U = 0.5$ and $\varrho = 0.5$

Method	$\hat{\beta}_1$			$\hat{\beta}_2$			$\hat{\beta}_3$		
	Bias	SD	SDE(sd(SDE))	Bias	SD	SDE(sd(SDE))	Bias	SD	SDE(sd(SDE))
$(n, p_n) = (100, 8)$									
C_{SCAD}	0.769	3.689	0.687(2.701)	1.378	9.414	0.973(4.749)	1.532	13.419	0.823(3.894)
N_{SCAD}	0.875	0.492	0.251(0.063)	1.471	0.178	0.263(0.049)	1.215	0.252	0.228(0.033)
Oracle ₁	0.101	0.127	0.119(0.049)	1.111	0.139	0.139(0.146)	0.100	0.126	0.121(0.084)
Oracle ₂	0.430	0.532	0.785(6.753)	0.597	0.678	1.364(4.364)	0.521	0.585	0.837(8.114)
Full	0.847	3.695	1.091(7.934)	1.534	9.477	1.784(8.565)	1.806	13.476	1.211(9.255)
$(n, p_n) = (400, 13)$									
C_{SCAD}	0.162	0.200	0.193(0.023)	0.212	0.282	0.241(0.039)	0.180	0.217	0.206(0.026)
N_{SCAD}	0.758	0.145	0.117(0.021)	1.500	0.000	0.125(0.010)	1.253	0.101	0.111(0.008)
Oracle ₁	0.046	0.058	0.058(0.004)	0.055	0.068	0.064(0.004)	0.050	0.063	0.058(0.004)
Oracle ₂	0.159	0.193	0.194(0.023)	0.204	0.249	0.243(0.045)	0.177	0.211	0.207(0.026)
Full	0.181	0.208	0.200(0.025)	0.249	0.277	0.249(0.044)	0.263	0.280	0.215(0.030)

We now verify the consistency of the estimators and test the accuracy the standard error formula. Table 2 displays the bias (columns labeled Bias) and sample standard deviation (columns labeled SD) of the estimates for three nonzero coefficients, over 1000 simulations. These can be regard as the true standard errors and compared with 1000 estimated standard errors. The 1000 estimated standard errors by using the sandwich formula are summarized by their mean (columns labeled SDE) and the sample standard deviations (sd(SDE)). The accuracy gets better when n increases. We omit here the results for other configurations, only for case $\sigma^2 = 1$, $\sigma_U = 0.5$ and $\varrho = 0.5$. Overall, the estimators are consistent and the sandwich formula works well.

5. Discussion

In this paper, we have proposed a variable selection procedure for the high dimensional PLVCE models. Our method extends the variable selection procedure to the setting, in which high dimension, measurement error, semiparametric models are considered at the same time. We have shown that the proposed method is consistent in variable selections, and the estimators of the regression coefficients have oracle property. Simulation studies indicate that the proposed method seems rather encouraging. To conclude this article, we would like to discuss some interesting topics for future study. Firstly, in this paper, we assume that the covariance matrix of measurement errors is known. However, it is usually unknown in many applications. If the covariance matrix is unknown, the variable selection procedure proposed by this paper will not work any more unless repeated measurements of the data are available. As a future research topic, it is interest to consider the variable selection for the high dimensional PLVCE models when the covariance matrix of measurement errors is unknown. Secondly, it is interesting to perform variable

selection for $p_n \gg n$. Variable selection for large p_n , small n setting is a very active research topic. However, it is challenging to extend the existing procedures for large p_n , small n problems to measurement error data. The details will also be further investigated in the future.

6. Proofs

In order to prove the main results, we first introduce several lemmas. Let $\mu_k = \int t^k K(t) dt$, $\nu_k = \int t^k K^2(t) dt$, $c_n = h^2 + [\log(1/h)/nh]^{1/2}$. Set $\Psi(T_1) = E(X_1 Z_1^\top | T_1)$, $\Upsilon(T_1) = E(Z_1 Z_1^\top | T_1)$ and $\Xi(T_1; \beta) = E[Z_1(Y_1 - X_1^\top \beta) | T_1]$. Furthermore, denote by $\alpha(t; \beta)$ the 'least favorable curve' of the nonparametric function $\alpha(t)$, which is defined as

$$(6.1) \quad \alpha(t; \beta) = \arg \min_{\eta} E[(Y_i - W_i^\top \beta - Z_i^\top \eta)^2 | T_i = t] = \Upsilon^{-1}(t) \Xi(t; \beta),$$

and let $Q_n(\beta) = L_n(\beta) + n \sum_{j=1}^{p_n} p_\lambda(|\beta_j|)$, where $2L_n(\beta) = \sum_{i=1}^n (Y_i - W_i^\top \beta - Z_i^\top \alpha(T_i; \beta))^2 - n\beta^\top \Sigma_U \beta$. Apparently, $\alpha(t; \beta^*) = \alpha^*(t)$ and $\frac{\partial \alpha(t; \beta)}{\partial \beta} = \Psi(t) \Upsilon^{-1}(t)$ is a p_n by q matrix. The following Lemma 6.1 can be found in [3].

6.1. Lemma. *Let (X_i, Y_i) , $i = 1, \dots, n$ be i.i.d. random vectors, where the Y_i are scale random variables. Further assume that $E|y|^\kappa < \infty$ and $\sup_x \int |y|^\kappa f(x, y) dy < \infty$, where f denotes the joint density of (X, Y) . Let K be a bounded positive function with a bounded support, satisfying a Lipschitz condition. Given that $n^{2\delta-1}h \rightarrow \infty$ for some $\delta < 1 - \kappa^{-1}$, then*

$$\sup_x \left| \frac{1}{n} \sum_{i=1}^n \left\{ K_h(X_i - x) Y_i - E[K_h(X_i - x) Y_i] \right\} \right| = O_P \left(\left\{ \frac{\log(1/h)}{nh} \right\}^{1/2} \right).$$

6.2. Lemma. *Under regularity conditions (C1)-(C5), the following holds uniformly in $t \in \mathcal{T}$,*

$$\begin{aligned} \hat{\alpha}(t; \beta) - \alpha(t; \beta) &= O_P(c_n), \\ \frac{\partial \hat{\alpha}(t; \beta)}{\partial \beta_k} - \frac{\partial \alpha(t; \beta)}{\partial \beta_k} &= O_P(c_n), \text{ for } k = 1, \dots, p_n. \end{aligned}$$

Proof. From Lemma 6.1, we have that

$$\begin{aligned} \frac{1}{n} D_t^\top \omega_t D_t &= \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} Z_i Z_i^\top & Z_i Z_i^\top \frac{T_i - t}{h} \\ Z_i Z_i^\top \frac{T_i - t}{h} & Z_i Z_i^\top \left(\frac{T_i - t}{h} \right)^2 \end{pmatrix} K_h(T_i - t) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 \end{pmatrix} \otimes \Upsilon(t) f_T(t) \{1 + O_P(c_n)\} \quad \text{and} \\ \frac{1}{n} D_t^\top \omega_t (\mathbf{Y} - \mathbf{W}\beta) &= \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} Z_i (Y_i - W_i^\top \beta) \\ Z_i (Y_i - W_i^\top \beta) \frac{T_i - t}{h} \end{pmatrix} K_h(T_i - t) \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \Xi(t; \beta) f_T(t) \{1 + O_P(c_n)\} \end{aligned}$$

hold uniformly in $t \in \mathcal{T}$. Here the symbol \otimes represent the Kronecker product between matrices. Hence, invoking equation (6.1) and $\hat{\alpha}(t; \beta)$ in Section 2, the first conclusion follows. The second assertion can get similarly. \square

6.3. Lemma. *Under regularity conditions (C1)-(C5), if $p_n^\kappa/n \rightarrow 0$ for $\kappa > 5/4$, $h = O(n^{-\varsigma})$ with $(4\kappa)^{-1} < \varsigma < 1 - \kappa^{-1}$, then for any β ,*

$$n^{-1/2} \|\nabla \hat{L}_n(\beta) - \nabla L_n(\beta)\| = o_P(1).$$

Proof. Invoking Lemma 6.2, the column vector $n^{-1/2}(\nabla\widehat{L}_n(\beta) - \nabla L_n(\beta))$ has the k th component equals

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ (Y_i - W_i^\top \beta - Z_i^\top \widehat{\alpha}(T_i; \beta)) \left(-W_{ik} - \frac{\partial \widehat{\alpha}(T_i; \beta)}{\partial \beta_k} Z_i \right) \right. \\ & \quad \left. - (Y_i - W_i^\top \beta - Z_i^\top \alpha(T_i; \beta)) \left(-W_{ik} - \frac{\partial \alpha(T_i; \beta)}{\partial \beta_k} Z_i \right) \right\} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ (Y_i - W_i^\top \beta - Z_i^\top \alpha(T_i; \beta)) \frac{\partial \alpha(T_i; \beta)}{\partial \beta_k} Z_i + Z_i^\top \alpha(T_i; \beta) W_{ik} \right\} O_P(c_n) \\ &= O_P(c_n). \end{aligned}$$

Hence we have shown

$$n^{-1/2} \|\nabla\widehat{L}_n(\beta) - \nabla L_n(\beta)\| = O_P(\sqrt{p_n} c_n) = o_P(1),$$

and the proof is complete. \square

6.4. Lemma. *Under the conditions of Theorem 1, we have*

$$\frac{[n^{-1/2} \nabla^\top L_n(\beta^*)](\Sigma_2)^{-1} [n^{-1/2} \nabla L_n(\beta^*)] - p_n}{\sqrt{2p_n}} \xrightarrow{D} N(0, 1),$$

where $\Sigma_2 = \mathbb{E}[(\varepsilon_i - U_i^\top \beta^*)(\Psi(T_i) \Upsilon^{-1}(T_i) Z_i - X_i) - \Sigma_U \beta^*]^{\otimes 2}$. In addition, $\nabla L_n(\beta^*) = O_P(\sqrt{np_n})$. Likewise, the results above hold also by $L_n(\beta^*)$ replaced with $\widehat{L}_n(\beta^*)$.

Proof. From (6.1), we get the following formulas $\mathbb{E}[Z_i(Y_i - X_i^\top \beta - Z_i^\top \alpha(T_i; \beta)) | T_i = t] = 0$ and $\mathbb{E}[X_i Z_i^\top + \frac{\partial \alpha(T_i; \beta)}{\partial \beta} Z_i Z_i^\top | T_i = t] = 0$. Then $\mathbb{E}[\nabla L_n(\beta)] = 0$ follows. Direct calculation yields

$$\nabla L_n(\beta) = \sum_{i=1}^n (Y_i - W_i^\top \beta - Z_i^\top \alpha(T_i; \beta)) (\Psi(T_i) \Upsilon^{-1}(T_i) Z_i - W_i) - n \Sigma_U \beta.$$

Thus,

$$\frac{1}{\sqrt{n}} \nabla L_n(\beta^*) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ (\varepsilon_i - U_i^\top \beta^*) (\Psi(T_i) \Upsilon^{-1}(T_i) Z_i - W_i) - \Sigma_U \beta^* \right\}.$$

By applying the martingale central limit theorem as given in [9], we can easily obtain the first part. The second part follows from Lemma 6.3. \square

6.5. Lemma. *Under regularity conditions C1-C5, and $p^4/n = o(1)$,*

$$\begin{aligned} & \left\| \frac{1}{n} \nabla^2 L_n(\beta) - \Sigma_1 \right\| = o_P(p_n^{-1}), \\ & \left\| \frac{1}{n} \nabla^2 \widehat{L}_n(\beta) - \Sigma_1 \right\| = o_P(p_n^{-1}) + O_P(p_n c_n), \end{aligned}$$

where $\Sigma_1 = \mathbb{E}(X_1 X_1^\top) - \mathbb{E}\{\Psi(T_1) \Upsilon^{-1}(T_1) \Psi^\top(T_1)\}$.

Proof. Direct calculation yields $n^{-1} \nabla^2 L_n(\beta) = n^{-1} \sum_{i=1}^n (W_i - \Psi(T_i) \Upsilon^{-1}(T_i) Z_i) (W_i - \Psi(T_i) \Upsilon^{-1}(T_i) Z_i)^\top - \Sigma_U$. Then $\mathbb{E}[n^{-1} \nabla^2 L_n(\beta)] = \mathbb{E}\{\mathbb{E}[(W_i - \Psi(T_i) \Upsilon^{-1}(T_i) Z_i) (W_i - \Psi(T_i) \Upsilon^{-1}(T_i) Z_i)^\top | T_i]\} - \Sigma_U = \Sigma_1$. The first conclusion follows from

$$\begin{aligned} \mathbb{E} p_n^2 \left\| \frac{1}{n} \nabla^2 L_n(\beta) - \Sigma_1 \right\|^2 &= p_n^2 \mathbb{E} \sum_{j,k=1}^{p_n} \left\{ \frac{1}{n} \nabla^2 L_n(\beta) - \Sigma_1 \right\}_{jk}^2 \\ &= O\left(\frac{p_n^4}{n}\right) = o(1). \end{aligned}$$

From this, triangle inequality immediately gives the second conclusion if we can show that

$$(6.2) \quad \left\| \frac{1}{n} \nabla^2 \widehat{L}_n(\beta) - \frac{1}{n} \nabla^2 L_n(\beta) \right\| = O_P(p_n c_n).$$

To this end, for $k = 1, \dots, p_n$,

$$\begin{aligned} & n^{-1} \frac{\partial}{\partial \beta_k} (\nabla \widehat{L}_n(\beta) - \nabla L_n(\beta)) \\ &= n^{-1} \frac{\partial}{\partial \beta_k} \sum_{i=1}^n \left\{ (Y_i - W_i^\top \beta - Z_i^\top \hat{\alpha}(T_i; \beta)) \left(-W_i - \frac{\partial \hat{\alpha}(T_i; \beta)}{\partial \beta} Z_i \right) \right. \\ & \quad \left. - (Y_i - W_i^\top \beta - Z_i^\top \alpha(T_i; \beta)) \left(-W_i - \frac{\partial \alpha(T_i; \beta)}{\partial \beta} Z_i \right) \right\} \\ &= n^{-1} \sum_{i=1}^n \left\{ \left(W_{ik} + \frac{\partial \hat{\alpha}(T_i; \beta)}{\partial \beta_k} Z_i \right) \left(W_i + \frac{\partial \hat{\alpha}(T_i; \beta)}{\partial \beta} Z_i \right) \right. \\ & \quad \left. - \left(W_{ik} + \frac{\partial \alpha(T_i; \beta)}{\partial \beta_k} Z_i \right) \left(W_i + \frac{\partial \alpha(T_i; \beta)}{\partial \beta} Z_i \right) \right\} \\ &= O_P(\sqrt{p_n} c_n) \end{aligned}$$

where the last line follows from Lemma 6.2. Hence (6.2) follows and the proof completes. \square

Proof of Theorem 2.1. Let $\vartheta_n = \sqrt{p_n}(n^{-1/2} + a_n)$ and set $\|v\| = C$, where C is a large enough constant. Our aim is to show that for any given $\epsilon > 0$ there is a large constant C such that, for large n we have

$$(6.3) \quad \Pr \left\{ \inf_{\|v\|=C} \widehat{Q}_n(\beta^* + \vartheta_n v) > \widehat{Q}_n(\beta^*) \right\} \geq 1 - \epsilon.$$

This implies that with probability tending to 1 there is a local minimizer $\hat{\beta}$ in the ball $\{\beta^* + \vartheta_n v : \|v\| \leq C\}$ such that $\|\hat{\beta} - \beta^*\| = O_P(\vartheta_n)$.

Let $\Delta_n(v) = \widehat{Q}_n(\beta^* + \vartheta_n v) - \widehat{Q}_n(\beta^*)$. Recall that the first s_n components of β^* are nonzero, and $p_\lambda(\cdot)$ is nonnegative and $p_\lambda(0) = 0$. By the Taylor expansion and the fact that $\widehat{L}_n(\beta)$ is quadratic, we have

$$\begin{aligned} \Delta_n(v) &\geq \widehat{L}_n(\beta^* + \vartheta_n v) - \widehat{L}_n(\beta^*) + n \sum_{j=1}^{s_n} \{p_\lambda(|\beta_j^* + \vartheta_n v_j|) - p_\lambda(|\beta_j^*|)\} \\ &\geq \vartheta_n v^\top \nabla \widehat{L}_n(\beta^*) + \frac{1}{2} \vartheta_n^2 v^\top \nabla^2 \widehat{L}_n(\beta^*) v \\ & \quad + \sum_{j=1}^{s_n} n \vartheta_n p'_\lambda(|\beta_j^*|) \text{sign}(\beta_j^*) v_j + \frac{1}{2} \sum_{j=1}^{s_n} n \vartheta_n^2 p''_\lambda(|\beta_j^*|) v_j^2 \{1 + o(1)\} \\ &\triangleq D_1 + D_2 + D_3 + D_4. \end{aligned}$$

By Lemma 6.4 and $\sqrt{p_n} \leq \sqrt{n} \vartheta_n$, we get

$$\begin{aligned} |D_1| &= |\vartheta_n v^\top \nabla \widehat{L}_n(\beta^*)| \leq \vartheta_n \|\nabla \widehat{L}_n(\beta^*)\| \|v\| \\ &\leq O_P(\vartheta_n \sqrt{n p_n}) \|v\| \leq O_P(n \vartheta_n^2) \|v\| \end{aligned}$$

Next we consider D_2 , An application of Lemma 6.5 yields that

$$\begin{aligned} D_2 &= \frac{1}{2}\vartheta_n^2 v^\top \nabla^2 \widehat{L}_n(\beta^*) v = \frac{1}{2}n\vartheta_n^2 v^\top \left[\frac{1}{n} \nabla^2 \widehat{L}_n(\beta^*) - \Sigma_1 \right] v + \frac{1}{2}n\vartheta_n^2 v^\top \Sigma_1 v \\ &= \frac{1}{2}n\vartheta_n^2 v^\top \Sigma_1 v + o_P(1)n\vartheta_n^2 \|v\|^2. \end{aligned}$$

With regard to D_3 and D_4 , for $\sqrt{s_n}a_n \leq \sqrt{s_n}(n^{-1/2} + a_n) \leq \vartheta_n$, we have

$$\begin{aligned} |D_3| &\leq \sum_{j=1}^{s_n} |n\vartheta_n p'_\lambda(|\beta_j^*|) \text{sign}(\beta_j^*) v_j| \\ &\leq n\vartheta_n a_n \sum_{j=1}^{s_n} |v_j| \leq n\vartheta_n a_n \sqrt{s_n} \|v\| \leq n\vartheta_n^2 \|v\|, \quad \text{and} \\ |D_4| &= \frac{1}{2} \sum_{j=1}^{s_n} n\vartheta_n^2 p''_\lambda(|\beta_j^*|) v_j^2 \{1 + o(1)\} \leq b_n n\vartheta_n^2 \|v\|^2. \end{aligned}$$

Therefore, under the condition (P1), by allowing C to be large enough, all terms D_1 , D_3 , D_4 are dominated by D_2 , which is positive. This proves (6.3) and completes the proof. \square

Proof of Theorem 2.2. Let $\zeta_n = C\sqrt{p_n/n}$. It is sufficient to show that with probability tending to 1 as $n \rightarrow \infty$, for any β satisfying $\|\beta - \beta^*\| = O_P(\sqrt{p_n/n})$ we have, for $j = s_n + 1, \dots, p_n$,

$$(6.4) \quad \frac{\partial \widehat{Q}_n(\beta)}{\partial \beta_j} < 0 \text{ for } \beta_j \in (-\zeta_n, 0) \quad \text{and} \quad \frac{\partial \widehat{Q}_n(\beta)}{\partial \beta_j} > 0 \text{ for } \beta_j \in (0, \zeta_n).$$

By Taylor expansion and the fact that $\widehat{L}_n(\beta)$ is quadratic in β , we get

$$\begin{aligned} \frac{\partial \widehat{Q}_n(\beta)}{\partial \beta_j} &= \frac{\partial \widehat{L}_n(\beta)}{\partial \beta_j} + n p'_\lambda(|\beta_j|) \text{sign}(\beta_j) \\ &= \frac{\partial \widehat{L}_n(\beta^*)}{\partial \beta_j} + \sum_{k=1}^{p_n} \frac{\partial^2 \widehat{L}_n(\beta^*)}{\partial \beta_j \partial \beta_k} (\beta_k - \beta_k^*) + n p'_\lambda(|\beta_j|) \text{sign}(\beta_j) \\ &\triangleq J_1 + J_2 + J_3. \end{aligned}$$

Next, we consider J_1, J_2 . Invoking Lemma 6.4, we have

$$J_1 = O_P(\sqrt{n}) = O_P(\sqrt{np_n}).$$

The term J_2 can be written as $J_2 = \sum_{k=1}^{p_n} \left\{ \frac{\partial^2 \widehat{L}_n(\beta^*)}{\partial \beta_j \partial \beta_k} - n \Sigma_{1,jk} \right\} (\beta_k - \beta_k^*) + n \sum_{k=1}^{p_n} \Sigma_{1,jk} (\beta_k - \beta_k^*) \triangleq J_{21} + J_{22}$. Using the Cauchy-Schwarz inequality and $\|\beta - \beta^*\| = O_P(\sqrt{p_n/n})$, we have

$$|J_{22}| \leq n \sum_{k=1}^{p_n} |\Sigma_{1,jk} (\beta_k - \beta_k^*)| \leq n O_P(\sqrt{p_n/n}) \left[\sum_{k=1}^{p_n} (\Sigma_{1,jk})^2 \right]^{1/2}.$$

As the eigenvalues of Σ_1 are bounded according to condition (C7), we have $\sum_{k=1}^{p_n} (\Sigma_1)_{jk}^2 = O(1)$. This entails that $J_{22} = O_P(\sqrt{np_n})$. For J_{21} , applying the Cauchy-Schwarz inequality,

$$|J_{21}| \leq \|\beta - \beta^*\| \left[\sum_{k=1}^{p_n} \left\{ \frac{\partial^2 \widehat{L}_n(\beta^*)}{\partial \beta_j \partial \beta_k} - n \Sigma_{1,jk} \right\}^2 \right]^{1/2}.$$

By a standard argument from condition (C7), we have

$$\left[\sum_{k=1}^{p_n} \left\{ \frac{\partial^2 L_n(\beta^*)}{\partial \beta_j \partial \beta_k} - n \Sigma_{1,jk} \right\}^2 \right]^{1/2} = O_P(n).$$

Then $J_{21} = O_P(\sqrt{np_n})$ follows from $\|\hat{\beta} - \beta^*\| = O_P(\sqrt{p_n/n})$. Now we have

$$J_2 = O_P(\sqrt{np_n}).$$

Hence we have

$$\frac{\partial \widehat{Q}_n(\beta)}{\partial \beta_j} = n\lambda \left\{ \frac{p'_\lambda(|\beta_j|)}{\lambda} \text{sign}(\beta_j) + O_P\left(\frac{\sqrt{p_n/n}}{\lambda}\right) \right\}.$$

Because of $\sqrt{p_n/n}/\lambda \rightarrow 0$ and (P2), the sign of β_j completely determines the sign of $\partial \widehat{Q}_n(\beta)/\partial \beta_j$. Then (6.4) follows from the continuity of $\partial \widehat{Q}_n(\beta)/\partial \beta_j$. Combining with the result of Theorem 2.1, there is a $\sqrt{n/p_n}$ -consistent local minimizer $\hat{\beta}$ of $\widehat{Q}_n(\beta)$ and $\hat{\beta}$ has the form $(\hat{\beta}_I^\top, 0^\top)^\top$, i.e. part (i) holds.

Now we prove part (ii). As shown in Theorem 2.1, we let λ_n be sufficiently small so that $a_n = o(n^{-1/2})$, then $\hat{\beta}$ is $\sqrt{n/p_n}$ consistent. By part (i), each component of $\hat{\beta}_I$ stays away from zero for a sufficiently large sample size n because β_I^* is away from zero. At the same time, $\hat{\beta}_{I^c} = 0$ with probability tending to 1. As a consequence, the estimate $\hat{\beta}_I$ based on the penalized estimation are necessarily the solution of the following estimation equation

$$(6.5) \quad \nabla \widehat{L}_{nI}(\hat{\beta}_I) + nP'_\lambda(|\hat{\beta}_I|) = 0$$

where $P'_\lambda(|\hat{\beta}_I|)$ is a s_n -vector whose j th element is $p'_\lambda(|\hat{\beta}_j|)\text{sign}(\hat{\beta}_j)$. Applying a Taylor expansion to (6.5) and re-arranging the resulting terms, we have

$$(\Sigma_{1I} + \Sigma_{\lambda_n})(\hat{\beta}_I - \beta_I^*) + P'_\lambda(|\hat{\beta}_I|) = -\frac{1}{n} \nabla L_{nI}(\beta_I^*) + R_1 + R_2$$

where $R_1 = -\left[\frac{1}{n} \nabla^2 \widehat{L}_{nI}(\beta_I^*) + P''_\lambda(|\hat{\beta}_I|) - \Sigma_{1I} - \Sigma_{\lambda_n}\right](\hat{\beta}_I - \beta_I^*)$ and $R_2 = \frac{1}{n} \nabla L_{nI}(\beta_I^*) - \frac{1}{n} \nabla \widehat{L}_{nI}(\beta_I^*)$. By Lemma 6.5 and Cauchy-Schwarz inequality, $\|R_1\| = o_P((np_n)^{-1/2}) + O_P(\sqrt{p_n^3/n}c_n) = o_P(n^{-1/2})$. By Lemma 6.3, we have $R_2 = o_P(n^{-1/2})$. Hence, we have

$$\begin{aligned} & \sqrt{n} A_n \Sigma_{2I}^{-1/2} \{ \Sigma_{1I} + \Sigma_\lambda \} \{ (\hat{\beta}_I - \beta_I^*) + \{ \Sigma_{1I} + \Sigma_\lambda \}^{-1} B \} \\ &= -\frac{1}{\sqrt{n}} A_n \Sigma_{2I}^{-1/2} \nabla L_{nI}(\beta_I^*) + o_P(1), \end{aligned}$$

Since $\|A_n \Sigma_{2I}^{-1/2}\| = O(1)$ by conditions of this theorem.

Next, we verify the Lindeberg-Feller Central Limit Theorem for the last term above. Let

$$\psi_{ni} = \frac{1}{\sqrt{n}} A_n \Sigma_{2I}^{-1/2} \nabla L_{nIi}(\beta_I^*), \quad i = 1, \dots, n,$$

where $\nabla L_{nIi}(\beta_I^*) = \left\{ (Y_i - W_{iI}^\top \beta_I^* - Z_i^\top \alpha(T_i; \beta_I^*)) (W_{iI} + \frac{\partial \alpha(T_i; \hat{\beta}_I)}{\partial \beta_I}) + \Sigma_{UI} \beta_I^* \right\}$. For any $\epsilon > 0$,

$$\begin{aligned} \sum_{i=1}^n \mathbb{E} \|\psi_{ni}\|^2 I\{\|\psi_{ni}\| > \epsilon\} &= n \mathbb{E} \|\psi_{n1}\|^2 I\{\|\psi_{n1}\| > \epsilon\} \\ &\leq n \{ \mathbb{E} \|\psi_{n1}\|^4 \}^{1/2} \{ \Pr(\|\psi_{n1}\| > \epsilon) \}^{1/2}. \end{aligned}$$

Using Chebyshev's inequality, we have $\Pr(\|\psi_{n1}\| > \epsilon) \leq \frac{\mathbb{E}\|\psi_{n1}\|^2}{\epsilon^2}$
 $= \frac{\mathbb{E}\|A_n \Sigma_{2I}^{-1/2} \nabla L_{nIi}(\beta_I^*)\|^2}{n\epsilon^2} = O(n^{-1})$ and $\mathbb{E}\|\psi_{n1}\|^4 = \mathbb{E}(\psi_{ni}^\top \psi_{ni})^2 \leq \frac{1}{n^2} \Lambda_{\max}^2(A_n A_n^\top)$
 $\Lambda_{\max}^2(\Sigma_{2I}^{-1}) \mathbb{E}\|\nabla L_{nIi}^\top(\beta_I^*) \nabla L_{nIi}(\beta_I^*)\|^2 = O(\frac{s_n}{n^2})$, by condition (C7). Hence, we get

$$\sum_{i=1}^n \mathbb{E}\|\psi_{ni}\|^2 I\{\|\psi_{ni}\| > \epsilon\} = O(n \frac{s_n}{n} \frac{1}{\sqrt{n}}) = o(1).$$

Also, note that $\mathbb{E}\psi_{ni} = 0$ and

$$\sum_{i=1}^n \text{Cov}(\psi_{ni}) = n \text{Cov}(\psi_{n1}) = \text{Cov}(A_n \Sigma_{2I}^{-1/2} \nabla L_{nIi}(\beta_I^*)) = A_n A_n^\top \rightarrow G.$$

From the foregoing argument, ψ_{ni} satisfies the conditions of the Lindeberg-Feller central limit theorem, then we complete the proof of part (ii). \square

Acknowledgments

This work was supported by the National Natural Science Foundation of China (11571025, Key grant: 11331011), and the Beijing Natural Science Foundation (1182002).

References

- [1] Ahmad, I., Leelahanon, S. and Li, Q. *Efficient estimation of a semiparametric partially linear varying coefficient model*, The Annals of Statistics **33** (1), 258–283, 2005.
- [2] Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. *Measurement Error in Nonlinear Models (2nd ed)*, New York: Chapman and Hall, 2006.
- [3] Fan, J.Q. and Huang, T. *Profile likelihood inferences on semiparametric varying coefficient partially linear models*, Bernoulli **11** (6), 1031–1057, 2005.
- [4] Fan, J.Q. and Li, R.Z. *Variable selection via nonconcave penalized likelihood and its oracle properties*, Journal of the American Statistical Association **96** (456), 1348–1360, 2001.
- [5] Fan, J.Q. and Peng, H. *Nonconcave penalized likelihood with a diverging number of parameters*, The Annals of Statistics **32** (3), 928–961, 2004.
- [6] Feng, S.Y. and Xue, L.G. *Bias-corrected statistical inference for partially linear varying coefficient errors-in-variables models with restricted condition*, Annals of the Institute of Statistical Mathematics **66** (1), 121–140, 2014.
- [7] Fuller, W. A. *Measurement Error Models*. New York: John Wiley, 1987.
- [8] Hu, X.M., Wang, Z.Z. and Zhao, Z.Z. *Empirical likelihood for semiparametric varying-coefficient partially linear errors-in-variables models*, Statistics and Probability Letters **79** (8), 1044–1052, 2009.
- [9] Hall, P. and Heyde, C. (1980). *Martingale limit theory and its application*. Academic Press.
- [10] Huang, Z. and Zhang, R.Q. *Empirical likelihood for nonparametric parts in semiparametric varying-coefficient partially linear models*, Statistics and Probability Letters **79** (16), 1798–1808, 2009.
- [11] Kai, B., Li, R. Z. and Zou, H. *New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models*, The Annals of Statistics **39** (1), 305–332, 2011.
- [12] Li, G.R., Feng, S.Y. and Peng, H. *Profile-type smoothed score function for a varying coefficient partially linear model*, Journal of Multivariate Analysis **102** (2), 372–385, 2011.
- [13] Li, G.R., Lin, L. and Zhu, L.X. *Empirical likelihood for varying coefficient partially linear model with diverging number of parameters*, Journal of Multivariate Analysis **105** (1), 85–111, 2012.
- [14] Li, R.Z. and Liang, H. *Variable selection in semiparametric regression modeling*, The Annals of Statistics **36** (1), 261–286, 2008.
- [15] Liang, H., Härdle, W. and Carroll, R.J. *Estimation in a semiparametric partially linear errors-in-variables model*, The Annals of Statistics **27** (5), 1519–1535, 1999.

- [16] Ruppert, D., Wand, M. and Carroll, R. *Semiparametric Regression*. Cambridge University Press, 2003.
- [17] Tibshirani, R. J. *Regression shrinkage and selection via the Lasso*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **58** (1), 267–288, 1996.
- [18] Wang, K. and Lin, L. *Simultaneous structure estimation and variable selection in partial linear varying coefficient models for longitudinal data*, Journal of Statistical Computation and Simulation **85** (7), 1459–1473, 2015.
- [19] Wang, X. L., Li, G. R. and Lin, L. *Empirical likelihood inference for semiparametric varying-coefficient partially linear EV models*, Metrika **73** (2), 171–185, 2011.
- [20] Wei, C. H. *Statistical inference for restricted partially linear varying coefficient errors-in-variables models*, Journal of Statistical Planning and Inference **142** (8), 2464–2472, 2012.
- [21] You, J. H. and Chen, G. M. *Estimation of a semiparametric varying-coefficient partially linear errors-in-variables model*, Journal of Multivariate Analysis **97** (2), 324–341, 2006.
- [22] Zhang, C.H. *Nearly unbiased variable selection under minimax concave penalty*, The Annals of Statistics **38** (2), 894–942, 2010.
- [23] Zhang, W., Lee, S., Y. and Song, X. *Local polynomial fitting in semivarying coefficient models*, Journal of Multivariate Analysis **82** (1), 166–188, 2002.
- [24] Zhao, P.X. and Xue, L.G. *Variable selection for semiparametric varying coefficient partially linear models*, Statistics and Probability Letters **79** (20), 2148–2157, 2009.
- [25] Zhao, P.X. and Xue, L.G. *Variable selection for semiparametric varying coefficient partially linear errors-in-variables models*, Journal of Multivariate Analysis **101** (8), 1872–1883, 2010.
- [26] Zhou, Y. and Liang, H. *Statistical inference for semiparametric varying-coefficient partially linear models with generated regressors*, The Annals of Statistics **37** (1), 427–458, 2009.
- [27] Zou, H. *The adaptive lasso and its oracle properties*, Journal of the American Statistical Association **101** (476), 1418–1429, 2006.
- [28] Zou, H. and Hastie, T. *Regularization and variable selection via the elastic net*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **67** (2), 301–320, 2005.

Calibration of the empirical likelihood for semiparametric varying-coefficient partially linear models with diverging number of parameters

Junwei He* and Xia Chen^{†‡}

Abstract

This article is concerned with the calibration of the empirical likelihood for semiparametric varying-coefficient partially linear models with diverging number of parameters. However, there is always substantial lack-of-fit, when the empirical likelihood ratio is calibrated by a bias-corrected empirical likelihood, producing tests with type I errors much larger than nominal levels. So we consider an effective calibration method and study the asymptotic behavior of this bias-corrected empirical likelihood ratio function. Some simulation studies are conducted to illustrate our approach.

Keywords: Varying-coefficient partially linear models, Empirical likelihood, Bias correction, Asymptotic normality, Coverage accuracy.

2000 AMS Classification: 62G08, 62G15

Received : 15.06.2017 *Accepted :* 28.06.2018 *Doi :* 10.15672/HJMS.2018.604

1. Introduction

Consider the following semiparametric varying-coefficient partially linear models

$$(1.1) \quad Y = X^T \alpha(U) + Z^T \beta + \varepsilon$$

where $\alpha(\cdot) = (\alpha_1(\cdot), \dots, \alpha_q(\cdot))^T$ is a q -dimensional vector of unknown regression functions, $\beta = (\beta_1, \dots, \beta_p)^T$ is a p -dimensional of unknown regression coefficients, and ε is an independent random error with $E(\varepsilon|X, Z, U) = 0$ almost surely. Without loss of generality, we assume that the variable U is defined on the unit interval $[0, 1]$.

As the extension of the usual linear regression model and partially linear regression model, semiparametric varying-coefficient partially linear model (1.1) has attracted great

*School of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710119, China, Email: junwei-he@snnu.edu.cn

[†]School of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710119, China, Email: xchen80@snnu.edu.cn

[‡]Corresponding Author.

research interest. For example, Fan and Huang [4] proposed a profile-kernel inference and established the asymptotic normality of the profile least-square estimator for this model. You and Zhou [16] studied the model (1.1) using the empirical likelihood method when p is fixed. When dimensionality p of the parameters tends to infinity as the sample size $n \rightarrow \infty$, this generalized varying-coefficient partially linear model was considered by Lam and Fan [7]. More relevant works on the varying-coefficient partially linear model can be found in Huang and Zhang [6], Li et al. [8] and references therein.

Empirical likelihood method has taken much attention in literatures since it was introduced and developed by Owen [10,11]. One of the motivation is that the empirical likelihood-based confidence regions not only have natural shape and respect the range of the parameter, but also have the advantages of studentising automatically. In many cases, empirical likelihood-based confidence regions are shown to be Bartlett correctable(DiCiccio et al. [3], Chen and Cui [1]). Owen [12] and Xue and Zhu [15] are fairly comprehensive references.

However, in practical application, there is always lack-of-fit for the asymptotic normality distribution of empirical likelihood ratio with expectation p and variance $2p$ when we refer to the coverage probability, especially when p/n is not small. We find that this is mainly due to the underestimation of the expectation and variance of the empirical likelihood ratio, producing tests with type I errors much larger than the nominal level. And this inspires us to look for an effective estimation of the expectation and variance. Liu et al. [9] proposed a new method which is fitted for the calibration of empirical likelihood for high-dimensional data. Through the calibration of the expectation and variance of the empirical likelihood for the population mean, they got a considerable improvements for the coverage probabilities. Guo et al. [5] considered this calibration method for high-dimensional data in linear models and discussed the asymptotic behavior of the empirical likelihood ratio function in random and fixed design cases, respectively. Recently, Li et al. [8] showed that under some conditions, the bias-correction empirical likelihood for the semiparametric varying-coefficient partially linear models is asymptotic normal.

Taking these issues into account, in this paper, we consider a new calibration of empirical likelihood for semiparametric varying-coefficient partially linear models with diverging number of parameters and investigate the asymptotic behavior of this bias-corrected empirical likelihood ratio function. Numerical studies show that this new calibration method will have a great improvement.

The rest of this paper is organized as follows. In Section 2, we introduce the bias-corrected empirical likelihood(BCEL) for semiparametric varying-coefficient partially linear models. A new calibration of bias-corrected empirical likelihood is given in Section 3. In Section 4, some simulations are carried out to assess the performance of the proposed method. Technical proofs are stated in Section 5.

2. Bias-corrected Empirical Likelihood

Let $(Y_i; X_i^T, Z_i^T, U_i, 1 \leq i \leq n)$ be an independent identically distributed(i.i.d) random sample which come from the model (1.1) with the β and Z_i having the dimension $p \rightarrow \infty$ as $n \rightarrow \infty$. Then for any given β , we get

$$(2.1) \quad Y_i - Z_i^T \beta = X_i^T \alpha(U_i) + \varepsilon_i$$

Following Fan and Huang [4], we apply a local linear regression technique to estimate the varying-coefficient functions $\alpha_j(\cdot), j = 1, \dots, q$. For v in a small neighborhood of u , one can approximate $\alpha_j(v)$ by

$$(2.2) \quad \alpha_j(v) \approx \alpha_j(u) + \alpha_j'(u)(v - u) \equiv a_j + b_j(v - u) \quad j = 1, \dots, q$$

This leads to the following weighted local least squares problem: find $\{(a_j, b_j), j = 1, \dots, q\}$ to minimize

$$(2.3) \quad \sum_{i=1}^n \{Y_i - X_i^T(a + b(U_i - u)) - Z_i^T \beta\}^2 K_h(U_i - u)$$

where $K(\cdot)$ is a kernel function, h is a bandwidth and $K_h(\cdot) = K_h(\cdot/h)/h$.

The solution of problem (2.3) is

$$(2.4) \quad \hat{\alpha}(u, \beta) = (I_q, O_q)(D_u^T W_u D_u)^{-1} D_u^T W_u (Y - Z^* \beta)$$

where I_q denotes a q -dimensional identity matrix, O_q is the $q \times q$ matrix with all the entries being 0 and Let

$$D_u = \begin{pmatrix} X_1^T & \frac{U_1 - u}{h} X_1^T \\ \vdots & \vdots \\ X_n^T & \frac{U_n - u}{h} X_n^T \end{pmatrix}, \quad Z^* = (Z_1, \dots, Z_n) = \begin{pmatrix} Z_{11} & \dots & Z_{1p} \\ \vdots & \ddots & \vdots \\ Z_{n1} & \dots & Z_{np} \end{pmatrix}$$

$$Y = (Y_1, \dots, Y_n), \quad W_u = \text{diag}(K_h(U_1 - u), \dots, K_h(U_n - u))$$

and

$$\mu(u) = (E(XX^T|U = u))^{-1} E(XZ|U = u)$$

So we can write the auxiliary random vectors as follows

$$(2.5) \quad \hat{\eta}_i(\beta) = (Z_i - \hat{\mu}^T(U_i)X_i)(Y_i - X_i^T \hat{\mu}(U_i, \beta) - Z_i^T \beta)$$

where $\hat{\mu}(u) = (\hat{E}(X_i X_i^T | U_i = u))^{-1} \hat{E}(X_i Z_i^T | U_i = u)$ is the estimator of $\mu(u)$. $E(X_i X_i^T | U_i = u)$ and $E(X_i Z_i^T | U_i = u)$ can be estimated easily by using the kernel smoothing method. For convenience, we can also define the estimator of $X_i^T \mu(U_i)$ directly as follows

$$(2.6) \quad X_i^T \hat{\mu}(U_i) = \sum_{k=1}^n S_{ik} Z_k$$

where S_{ik} is the (i, k) -th element of the smoothing matrix S , which depends only on the observations $\{(U_i, X_i), i = 1, \dots, n\}$, with

$$S = \begin{pmatrix} (X_1^T, O)(D_{u_1}^T W_{u_1} D_{u_1})^{-1} D_{u_1}^T W_{u_1} \\ \vdots \\ (X_n^T, O)(D_{u_n}^T W_{u_n} D_{u_n})^{-1} D_{u_n}^T W_{u_n} \end{pmatrix}$$

Thus, the bias-corrected auxiliary random vectors can be expressed as

$$(2.7) \quad \hat{\eta}_i(\beta) = (Z_i - \hat{\mu}^T(U_i)X_i)(Y_i - X_i^T \hat{\mu}(U_i, \beta) - Z_i^T \beta) \triangleq \hat{Z}_i(\hat{Y}_i - \beta^T \hat{Z}_i)$$

where $\hat{Z}_i = Z_i - \sum_{k=1}^n S_{ik} Z_k$, $\hat{Y}_i = Y_i - \sum_{k=1}^n S_{ik} Y_k$.

Therefore, a bias-corrected empirical log-likelihood ratio is defined as

$$(2.8) \quad l_n(\beta) = -2 \max \left\{ \sum_{i=1}^n \log(n\omega_i) \mid \omega_i \geq 0, \sum_{i=1}^n \omega_i = 1, \sum_{i=1}^n \omega_i \hat{\eta}_i(\beta) = 0 \right\}$$

By the Lagrange multiplier method, we can obtain

$$(2.9) \quad l_n(\beta) = 2 \sum_{i=1}^n \log(1 + \lambda^T \hat{\eta}_i(\beta))$$

where $\lambda = \lambda(\beta)$ is determined by

$$(2.10) \quad \frac{1}{n} \sum_{i=1}^n \frac{\hat{\eta}_i(\beta)}{1 + \lambda^T \hat{\eta}_i(\beta)} = 0$$

According to Qin and Lawless [13], if $\alpha(u)$ is given and p is fixed, under some conditions, $l_n(\beta)$ is asymptotically χ^2 with p degree of freedom, which is a non-parametric version of Wilks' theorem. And when the number of p grows with the sample size n , Li et al. [8] showed that under some conditions, the conclusion below is valid.

$$(2.11) \quad \frac{l_n(\beta_0) - p}{\sqrt{2p}} \xrightarrow{d} N(0, 1), \quad \text{as } n \rightarrow \infty$$

where β_0 is the true value of the parameter vector β .

3. A new Calibration method for BCEL

When testing hypotheses with the BCEL method, we would calculate the critical values based on normal approximation (2.11). However, these critical values often deviate from the true ones when p/n is not small. We find that this awkward fact is mainly due to the large difference between the true expectation and variance pair (E_n, V_n) of $l_n(\beta_0)$ and $(p, 2p)$. And our simulation also indicates that this method is not good. We know that the foundation of using (2.11) to calibrate the BCEL are that $l_n(\beta_0)$ is close to $K_n = n\tilde{\eta}_n^T \Sigma^{-1} \tilde{\eta}_n$, and that $E(K_n) = p$, $Var(K_n) \approx 2p$. But in practice, we always use the moment estimation of K_n , which is, $T_n = n\tilde{\eta}_n^T S_n^{-1} \tilde{\eta}_n$, whose expectation and variance are denoted as $(\hat{E}_{n1}, \hat{V}_{n1})$, for statistical inference and it can always get a better approximation to $l_n(\beta_0)$. But when (E_n, V_n) deviates from $(p, 2p)$ or $(\hat{E}_{n1}, \hat{V}_{n1})$, these calibration methods do not work any more.

We expect that replacing $(p, 2p)$ with $(\hat{E}_{n2}, \hat{V}_{n2})$, the expectation and variance of T_{nc} (see (3.2)), in (2.11), will improve the performance of the usual normal calibration. Let

$$f(\lambda) = 2 \sum_{i=1}^n \log(1 + \lambda^T \hat{\eta}_i(\beta))$$

Obviously, $l_n(\beta_0) = \sup_{\lambda} f(\lambda) = f(\lambda_*)$, and λ_* is the maximum point of $f(\lambda)$. By second-order Taylor expansion, we have

$$(3.1) \quad f(\lambda) \approx g_1(\lambda) = 2 \sum_{i=1}^n \left\{ \lambda^T \hat{\eta}_i - \frac{1}{2} (\lambda^T \hat{\eta}_i)^2 \right\}$$

provided $\lambda^T \hat{\eta}_i$'s are small. So an approximation of $l_n(\beta_0)$ is

$$l_n(\beta_0) \approx \sup_{\lambda} f(\lambda) = \sup_{\lambda} g_1(S_n^{-1} \tilde{\eta}_n) = T_n$$

However, in the case of moderate n and large p , this approximation may not work any more. The remainder of each Taylor expansion in (3.1) is under control only for $\lambda^T \hat{\eta}_i \in (-1, 1)$. We find in our simulation that when p/n is not small, some of $\lambda^T \hat{\eta}_i$'s are greater than 1 with a large probability. Note that when

$$x \in (-1, 1), \log(1 + x) \approx x - \frac{x^2}{2}$$

while if

$$x > 1, \log(1 + x) > \log(2) > x - \frac{x^2}{2}$$

Therefore, roughly we have $f(\lambda) \geq g_1(\lambda)$ in the neighborhood of 0. This finding also restrict us to approximate $l_n(\beta_0)$ by two terms Taylor expansion, because Taylor expansion of (3.1) would deviate from $l_n(\beta_0)$ if more terms are extracted and some of $\lambda_* \hat{\eta}_i$ are not small.

To reduce the approximation error of $g_1(\lambda)$, we add a high-order term $(\lambda^T \hat{\eta}_i)^2$ to $g_1(\lambda)$. Intuitively $g_2(\lambda) = g_1(\lambda) + (\lambda^T \hat{\eta}_i)^2$ is the better approximate to $f(\lambda)$. So is $\sup_\lambda g_2(\lambda)$ to $l_n(\beta_0) = \sup_\lambda f(\lambda)$. It can be verified

$$(3.2) \quad \sup_\lambda g_2(\lambda) = n \bar{\eta}_n^T S_{nc}^{-1} \bar{\eta}_n = T_{nc}$$

with

$$S_{nc} = \frac{1}{n} \sum_{i=1}^n (\hat{\eta}_i - \bar{\eta}_n)(\hat{\eta}_i - \bar{\eta}_n)^T$$

The following theorem establishes the asymptotic behavior of $l_n(\beta_0) - T_{nc}$.

3.1. Theorem. *Under Conditions (C1) – (C9) in Section 5, if $p^{3+4/(k-2)}/n \rightarrow 0$, for $k \geq 4$, then we have*

$$(l_n(\beta_0) - T_{nc})/p^{\frac{1}{2}} = o_p(1)$$

This theorem implies that using T_{nc} to approximate $l_n(\beta_0)$ is equivalent to using K_n or T_n from the asymptotic viewpoints. However, these approximations exhibit quite different finite-sample behaviors, especially when p/n is not small. Based on some simulations, we find that T_{nc} is amazingly close to $l_n(\beta_0)$ regardless of the choices of (n, p) in the sense that $(l_n(\beta_0) - T_{nc})/p^{\frac{1}{2}} = o_p(1)$ is always pretty small. To appreciate this, Fig.1 shows the scatter plots of 200 simulated values of $(l_n(\beta_0), T_n)$ and $(l_n(\beta_0), T_{nc})$ for the model(4.1) with the $\varepsilon_i \sim N(0, 1)$. We choose $p=10, 16$ for $n=200$. From Fig.1, we can see that the value of $(l_n(\beta_0), T_{nc})$ are always around the line $y = x$, but T_n tends to under-approximate $l_n(\beta_0)$. See Sect.4 for more analysis and comparison.

Given the foregoing discussion and evidence, we expect that the expectation and variance of T_{nc} are good approximations of E_n and V_n , respectively. Let $(\hat{E}_{n2}, \hat{V}_{n2})$ be the moment estimation of (E_n, V_n) . We may calculate critical values according to

$$(3.3) \quad l_n(\beta_0) - A_n/\sqrt{B_n} \xrightarrow{d} N(0, 1)$$

where (A_n, B_n) could be chosen as $(p, 2p)$ or $(\hat{E}_{ni}, \hat{V}_{ni})(i = 1, 2)$. We will show that the method based on $(\hat{E}_{n2}, \hat{V}_{n2})$ is the best. Hence, it is our final recommendation.

4. Numerical Analysis

Here we report a simulation study designed to evaluate the performance of the proposed calibration method of BCEL. Throughout this section, we use the Epanechnikow kernel $K(u) = 0.75(1 - u^2)_+$, and use the "leave-one-out" cross-validation method to select the optimal bandwidth h_{opt} .

Consider the following semiparametric varying-coefficient partially linear model

$$(4.1) \quad Y_i = X_i^T \alpha(U_i) + Z_i^T \beta + \varepsilon_i, \quad i = 1, \dots, n$$

In our simulations, $\beta = [0.5, 0.3, -0.5, 1, 0.1, -0.25, 0, \dots, 0]^T$, the covariate U_i is uniformly distributed on $[0, 1]$, the nonparametric component $\alpha(u) = (\alpha_1(u), \alpha_2(u))^T$ with $\alpha_1(u) = 4 + \sin(2\pi u)$, $\alpha_2(u) = 2u(1 - u)$, $X_i = (X_{i1}, X_{i2})^T$ with $X_{i1} = 1$ and $X_{i2} \sim N(0, 1)$, the covariates Z_i is a p -dimensional normal random vector with mean zero and covariance matrix (σ_{ij}) with $\sigma_{ij} = 0.5^{|i-j|}$.

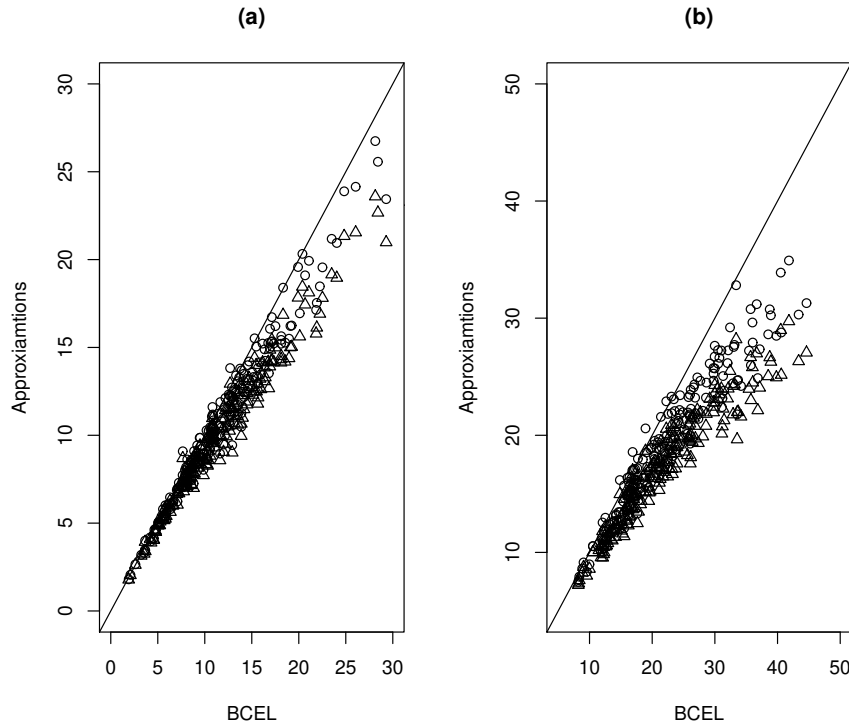


Fig.1. Scatter plots of n simulated values $(l_n(\beta_0), T_n)$ (triangles), $(l_n(\beta_0), T_{nc})$ (circles) for the model (4.1) with (a) $n=200$ and $p=10$, (b) $n=200$ and $p=16$. The solid line is $y = x$.

4.1. Simulation I. For this simulation, we evaluate the asymptotic normality of BCEL ratio using the following methods. The proposed method is based on the calibrated $l_n(\beta_0)$ with the sample mean and variance of T_{nc} obtained from 500 Bootstrap samples for each simulation data set (denoted as MEL). The normal calibration is based on the calibrated $l_n(\beta_0)$ with the sample mean and variance of T_n obtained from 500 Bootstrap samples for each simulation data set (denoted as SEL). And the standard normal calibration is based on the calibrated $l_n(\beta_0)$ with $(A_n, B_n) = (p, 2p)$ (denoted as STEL). Through QQ-plots, we will demonstrate the advantages of MEL in different growth rates of p for each sample size. Here we only consider the case of noise $\varepsilon_i \sim N(0, 1)$.

We draw 1000 random samples of size 200, 400 or 600 from model (4.1). For comparison, we here take the dimensionality of the parametric component as $p = \lceil cn^{1/3} \rceil$. By assigning $c = 1.8, 2.8$ and 3.8 , the corresponding dimensions $p = 10, 16$ and 22 for $n = 200$; $p = 13, 20$ and 27 for $n = 400$; $p = 15, 23$ and 32 for $n = 600$. The results are reported in Fig.2.

From Fig 2., we can observe from the QQ-plots that the MEL outperforms better than SEL and STEL as n increases or p decreases. Therefore, the MEL can be regarded as a reasonable alternative for the calibration of the BCEL in practice.

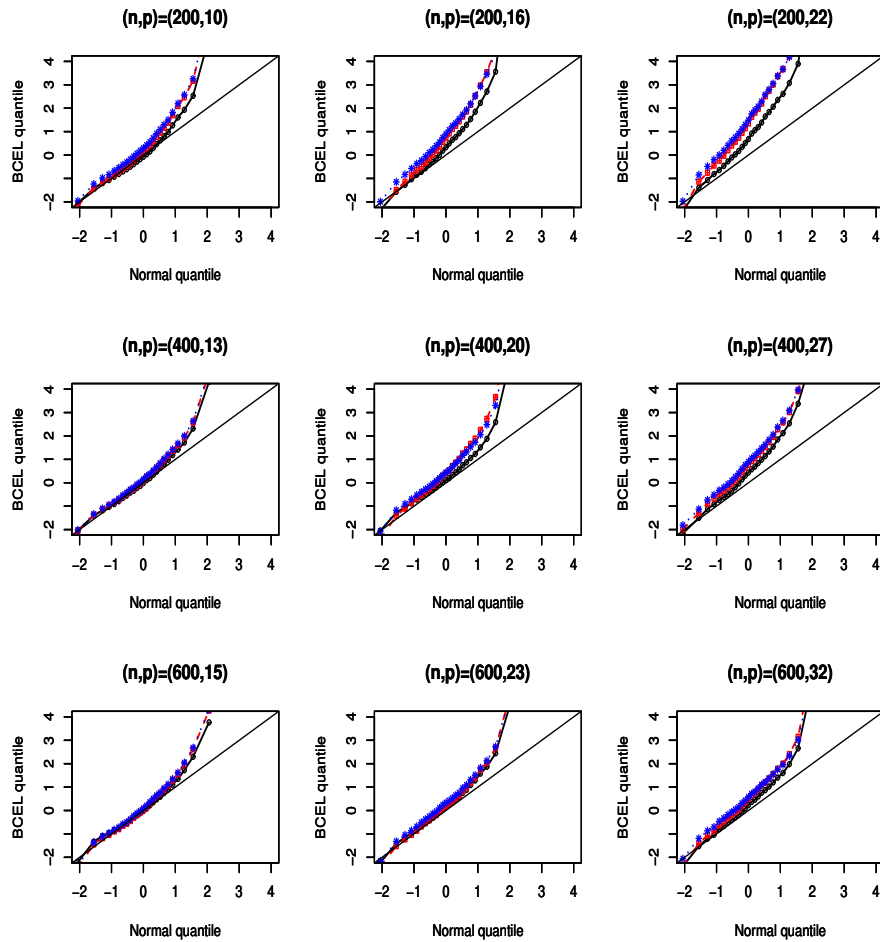


Fig.2. Normal QQ-plots of the BCEL ratio with $\epsilon_i \sim N(0,1)$:
MEL(black and $-\circ-$), SEL(red and $\cdot\boxplus\cdot$), STEL(blue and $-\cdot*\cdot-$)

4.2. Simulation II. In this simulation, We draw 1000 random samples of size 200, 400 and 600, respectively. The choice of (n, p) is the same as Simulation I. As for noise, two error distributions were chosen: (i) the standard normal distribution; (ii) the chi-square distribution with freedom 3.

In this simulation, we will compare four calibration methods for the BCEL. Besides the MEL, SEL and STEL methods mentioned in Section 4.1, there also consider the ordinary χ_p^2 calibration(denoted as OEL). Tables 1 and 2 report the coverage probability comparison for constructing confidence region on parameter β with nominal level 0.95.

It can be concluded from Tables 1 and 2 that the empirical coverage probabilities based on MEL are higher than that based on OEL, STEL and SEL. Especially for the case of $n = 600$, $p = 15$ and $\epsilon_i \sim N(0, 1)$, the coverage probabilities of MEL is closed to the nominal level. Thus the calibration method of MEL is a good alternative. We can

Table 1. Coverage percentages for model (4.1) with the $\epsilon_i \sim N(0, 1)$

n	p	MEL	OEL	SEL	STEL	\hat{E}_{n1}	\hat{V}_{n1}	\hat{E}_{n2}	\hat{V}_{n2}
200	10	0.920	0.838	0.846	0.854	10.67	19.15	11.69	25.80
	16	0.838	0.726	0.756	0.750	17.23	26.81	19.04	37.04
	22	0.764	0.552	0.593	0.615	23.46	38.75	26.79	56.09
400	13	0.937	0.899	0.910	0.914	13.24	26.39	13.81	29.77
	20	0.925	0.846	0.864	0.842	20.42	31.33	21.94	53.69
	27	0.841	0.741	0.777	0.789	27.98	51.72	30.24	58.93
600	15	0.936	0.898	0.904	0.911	15.36	29.69	15.62	37.41
	23	0.921	0.873	0.893	0.899	23.98	42.90	24.10	50.65
	32	0.896	0.836	0.872	0.849	32.88	54.57	34.72	64.91

Table 2. Coverage percentages for model (4.1) with the $\epsilon_i \sim \chi_3^2$

n	p	MEL	OEL	SEL	STEL	\hat{E}_{n1}	\hat{V}_{n1}	\hat{E}_{n2}	\hat{V}_{n2}
200	10	0.863	0.796	0.810	0.821	11.01	17.60	11.38	22.29
	16	0.803	0.694	0.721	0.698	16.73	25.31	18.27	34.76
	22	0.755	0.576	0.610	0.599	23.32	33.63	26.08	57.37
400	13	0.908	0.878	0.889	0.866	13.34	20.50	14.05	27.17
	20	0.844	0.772	0.798	0.763	20.45	31.61	22.07	41.41
	27	0.828	0.692	0.728	0.720	27.68	47.28	29.80	61.13
600	15	0.916	0.868	0.888	0.878	15.46	26.70	16.08	33.56
	23	0.890	0.852	0.869	0.871	23.83	43.75	24.62	47.74
	32	0.839	0.745	0.785	0.776	33.07	53.54	34.60	64.03

also observed from Table 1 and Table 2 that the MEL has improving coverage accuracy along with the increasing sample size. However, when the dimension p increases, the coverage probabilities of both MEL, OEL, STEL and SEL decrease. When $n = 200$ and $p = 22$, the performances of OEL, SEL and STEL are unacceptable. In comparison, our proposed method, MEL, can always attain the desired coverage percent and outperform the other three methods. The advantages get more remarkable when n decreases or p increases.

5. Proof of main results

Throughout the paper, we denote $\gamma_1(A) \leq \dots \leq \gamma_p(A)$ as the eigenvalues and $tr(A)$ as the trace operator of a matrix A . To derive our main results, the following conditions required to be made.

(C1) The random variable U has a compact support Ω . The density function $f_U(u)$ of the U has a continuous second derivative and is uniformly bounded away from zero.

(C2) The $q \times q$ matrix $E(XX^T|U = u)$ is non-singular for each $U \in \Omega$. Furthermore, $E(XX^T|U = u)^{-1}$ and $E(XZ|U = u)$ are all Lipschitz continuous and each element of $E(XX^T|U = u)^{-1}$ and $E(XZ|U = u)$ is bounded.

(C3) $\{\alpha_i(\cdot), i = 1, \dots, q\}$ has continuous second derivatives in $u \in \Omega$.

(C4) The kernel $K(\cdot)$ is bounded symmetric density function with bounded support.

(C5) The bandwidth h satisfies that $nh^6 \rightarrow 0$ and $nh^3/(\log(n))^3 \rightarrow \infty$.

(C6) $\Sigma = E[\varepsilon^2(Z - \mu^T(U)X)(Z - \mu^T(U)X)^T]$ is a positive definite matrix with all the eigenvalues being uniformly bounded away from zero and infinity.

(C7) For some integer $k \geq 4$, $E(\|X\varepsilon\|^k) < \infty$, $E(\|X\|^k) < \infty$, $E(\|\varepsilon\|^k) < \infty$.

(C8) Let $\eta = \varepsilon(Z - \mu^T(U)X)$, and η_j be the j -th component of η , $j = 1 \dots p$. For k of condition (C7), there is a positive constant c such that

$$E(\|\eta\sqrt{p}\|^k) < c, E(\|ZX^T/\sqrt{p}\|^k) < c, E(\|\mu(U)XX^T/\sqrt{p}\|^k) < c$$

and

$$\frac{1}{p} \sum_{l_1=1}^p E(|\eta_{l_1}| (\|ZX^T/\sqrt{p}\|^4 + \|\mu(U)XX^T/\sqrt{p}\|^4)) < c$$

(C9) $\max_{1 \leq l_1, l_2, l_3 \leq p} E(\eta_{l_1} \eta_{l_2} \eta_{l_3})^2$ is bounded, where η_{l_i} are the components of η .

In order to prove the main results, we introduce the following notations. Simple calculation yields that

$$(5.1) \quad \hat{\eta}_i(\beta) = \eta_i(\beta) + \sum_{k=1}^3 M_{i,k} =: \eta_i(\beta) + R_i$$

where

$$\eta_i(\beta) = (Z_i - \mu^T(U_i)X_i)(Y_i - X_i^T \alpha(U_i) - Z_i^T \beta) = (Z_i - \mu^T(U_i)X_i)\epsilon_i$$

$$M_{i,1} = (Z_i - \mu^T(U_i)X_i)X_i^T(\alpha(U_i) - \hat{\alpha}(U_i, \beta))$$

$$M_{i,2} = (\mu(U_i) - \hat{\mu}(U_i))^T X_i \epsilon_i$$

$$M_{i,3} = [(\mu(U_i) - \hat{\mu}(U_i))^T X_i][X_i^T(\alpha(U_i) - \hat{\alpha}(U_i, \beta))]$$

5.1. Lemma. *Suppose that Conditions (C1)-(C5) hold. If $h \rightarrow 0$ and $nh \rightarrow \infty$ as $n \rightarrow \infty$, then letting $c_n = \{\frac{\log n}{nh}\}^{1/2} + h^2$ and $d_n = \{\frac{\log n}{nh}\}^{1/2}$,*

$$\sup_{u \in \Omega} \frac{1}{n} \sum_{i=1}^n K_h(U_i - u) \left(\frac{U_i - u}{h} \right)^l X_{ij} \epsilon_i = O_p(d_n)$$

$$\sup_{u \in \Omega} \left| \frac{1}{n} \sum_{i=1}^n K_h(U_i - u) \left(\frac{U_i - u}{h} \right)^l X_{ij_1} X_{ij_2} - f(u) \mu_1 \Gamma_{j_1 j_2}(u) \right| = O_p(c_n)$$

$$\sup_{u \in \Omega} \left| \frac{1}{n} \sum_{i=1}^n K_h(U_i - u) \left(\frac{U_i - u}{h} \right)^l X_{ij} Z_{ik} - f(u) \phi_{jk}(u) \right| = O_p(c_n)$$

where $j_1, j_2, j = 1, \dots, q, k = 1, \dots, p, l = 0, 1, 2, 4$, $\Gamma_{j_1 j_2}(u)$ is the (j_1, j_2) -th the element of $\Gamma(u)$ and $\phi_{jk}(u)$ is the (j, k) -th element of $\phi(u)$.

We refer to Xia and Li [14] for details.

5.2. Lemma. *Under the Conditions of Lemma 5.1, we have,*

$$(5.2) \quad \|\hat{\alpha}(u, \beta) - \alpha(u)\| = O_p(c_n)$$

and

$$(5.3) \quad \max_{1 \leq j \leq q} \sup_{u \in \Omega} |\hat{\alpha}_j(u, \beta) - \alpha_j(u)| = O_p(c_n)$$

holds uniformly in $u \in \Omega$, the support of U .

Proof. We first give the proof of (5.2). Let

$$S_{n,l} = \sum_{i=1}^n K_h(U_i - u) X_i X_i^T \left(\frac{U_i - u}{h} \right)^l, \quad l = 0, 1, 2$$

Then, we can rewrite

$$D_u^T W_u D_u = \begin{pmatrix} S_{n,0} & S_{n,1} \\ S_{n,1} & S_{n,2} \end{pmatrix}$$

The elements of the above matrix are in the form of a kernel regression. From Lemma 5.1 and some simple calculation, we have

$$(5.4) \quad S_{n,l} = nf(u)\mu_l\Gamma(u)(1 + O_p(c_n))$$

holds uniformly in $u \in \Omega$. So

$$(5.5) \quad \hat{\alpha}(u, \beta) = [nf(u)\Gamma(u)]^{-1} \sum_{i=1}^n K_h(U_i - u) X_i \{X_i^T \alpha(U_i) + \varepsilon_i\} + O_p(c_n)$$

Applying Lemma 5.1 and (5.4), we can easily get

$$(5.6) \quad \frac{1}{n} \sum_{i=1}^n K_h(U_i - u) X_i X_i^T \alpha(U_i) = f(u)\Gamma(u)\alpha(u)\{1 + O_p(c_n)\}$$

and

$$(5.7) \quad \frac{1}{n} \sum_{i=1}^n K_h(U_i - u) X_i \varepsilon_i = o_p(1)$$

holds uniformly in $u \in \Omega$. From (5.5)-(5.7), $\hat{\alpha}(u, \beta) = \alpha(u) + O_p(c_n)$ holds uniformly in $u \in \Omega$. This completes the proof of (5.2).

By the similar method of Xia and Li [14], we can conclude the result (5.3), so we omit the details here. \square

5.3. Lemma. *Under the Conditions of Lemma 5.1, we have*

$$\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n R_i \right\| = O_p(n^{1/2} p^{1/2} c_n^2)$$

where $R_i = \sum_{k=1}^3 M_{i,k}$ can be found in (5.1).

The proof of Lemma 5.3 is similar as that of Lemma B.3 in Li et al. [5].

5.4. Lemma. *Under conditions (C1)-(C8), we have*

$$(5.8) \quad \text{tr}[(S_{nc} - \Sigma)^2] = O_p(p^2(c_n^4 + 1/n))$$

Proof. From the definition of η_i and S_{nc} , we can get

$$S_{nc} - \Sigma = \frac{1}{n} \sum_{i=1}^n \eta_i \eta_i^T + \frac{1}{n} \sum_{i=1}^n (R_i \eta_i^T + \eta_i R_i^T + R_i R_i^T) - \bar{\eta}_n \bar{\eta}_n^T = J_1 + J_2 + J_3$$

It is easy to see that

$$\begin{aligned} \text{tr}[(S_{nc} - \Sigma)^2] &= \text{tr}[(J_1 + J_2 + J_3)^2] \leq 4\text{tr}[(J_1)^2] + 4\text{tr}[(J_2)^2] + 2\text{tr}[(J_3)^2] \\ &= I_1 + I_2 + I_3 \end{aligned}$$

Thus, we know that

$$I_1 = O_p(p/n^2), \quad I_2 = O_p(p^2 c_n^4)$$

For I_3 , first we can get $\bar{\eta}_n = O_p(\sqrt{p/n})$, then

$$I_3 = \text{tr}[(\bar{\eta}_n \bar{\eta}_n^T)^2] = O_p(p^3/n^2) = \frac{p}{n} O_p(p^2/n) = O_p(p^2/n)$$

Therefore, we have

$$\text{tr}[(S_{nc} - \Sigma)^2] = I_1 + I_2 + I_3 = O_p(p^2/n) + O_p(p^2 c_n^4) + O_p(p^2/n) = O_p(p^2(c_n^4 + 1/n))$$

The proof is complete. \square

5.5. Lemma. *Under conditions (C1)-(C8), if $p^{3+4/(k-2)}/n \rightarrow 0$, we have*

$$(5.9) \quad n \left\{ \left(\frac{1}{n} \sum_{i=1}^n \hat{\eta}_i \right)^\top (S_{nc}^{-1} - \Sigma^{-1}) \left(\frac{1}{n} \sum_{i=1}^n \hat{\eta}_i \right) \right\} = o_p(p^{1/2})$$

Proof. Let $\hat{D}_n = \Sigma^{-1/2} S_{nc} \Sigma^{-1/2} - I_p$, similar arguments used in the proof of Lemma 6 in Chen et al. [2] yield

$$\begin{aligned} S_{nc}^{-1} - \Sigma^{-1} &= \Sigma^{-1/2} (\Sigma^{1/2} S_n^{-1} \Sigma^{1/2} - I_p) \Sigma^{-1/2} \\ &= \Sigma^{-1/2} [-\hat{D}_n + \hat{D}_n^2 + \hat{D}_n^2 \{\Sigma^{1/2} S_n^{-1} \Sigma^{1/2} - I_p\}] \Sigma^{-1/2} \end{aligned}$$

Note that

$$\begin{aligned} \text{tr}((S_{nc} - \Sigma)^2) &= \text{tr}((\Sigma^{1/2} (\Sigma^{-1/2} S_n \Sigma^{-1/2} - I_p) \Sigma^{1/2})^2) \\ &= \text{tr}(\hat{D}_n \Sigma \hat{D}_n \Sigma) \\ &\geq \gamma_1^2(\Sigma) \text{tr}(\hat{D}_n^2) \end{aligned}$$

By Lemma 5.4, we have

$$\text{tr}(\hat{D}_n^2) \leq \frac{1}{\gamma_1^2(\Sigma)} \text{tr}((S_{nc} - \Sigma)^2) = O_p(p^2(c_n^4 + 1/n))$$

Thus, we have

$$(5.10) \quad \begin{aligned} \text{tr}(S_{nc}^{-1} - \Sigma^{-1})^2 &\leq 2\text{tr}\{\Sigma^{-2}(-\hat{D}_n + \hat{D}_n^2)^2\} + 2\text{tr}\{\hat{D}_n^4(S_{nc}^{-1} - \Sigma^{-1})^2\} \\ &\leq 2\text{tr}\{\Sigma^{-2}(-\hat{D}_n + \hat{D}_n^2)^2\} + 2[\text{tr}\hat{D}_n^2]^2 \text{tr}\{(S_{nc}^{-1} - \Sigma^{-1})^2\} \\ &= 2\text{tr}\{\Sigma^{-2}(-\hat{D}_n + \hat{D}_n^2)^2\} + o_p(\text{tr}\{(S_{nc}^{-1} - \Sigma^{-1})^2\}) \\ &= o_p(p^2(c_n^4 + 1/n)) \end{aligned}$$

Then

$$(5.11) \quad \left\| \frac{1}{n} \sum_{i=1}^n \hat{\eta}_i^\top \right\| = O_p(\sqrt{p/n})$$

This together with $p^{3+4/(k-2)}/n \rightarrow 0$, $c_n^2 = o(1/\sqrt{n})$ and condition (C5), we can obtain

$$\begin{aligned} n \left\{ \left(\frac{1}{n} \sum_{i=1}^n \hat{\eta}_i \right)^\top (S_{nc}^{-1} - \Sigma^{-1}) \left(\frac{1}{n} \sum_{i=1}^n \hat{\eta}_i \right) \right\} &\leq n \left\| \frac{1}{n} \sum_{i=1}^n \hat{\eta}_i^\top \right\|^2 \sqrt{\text{tr}(S_{nc}^{-1} - \Sigma^{-1})^2} \\ &= o_p(p^2(c_n^2 + 1/\sqrt{n})) \\ &= o_p(p^{1/2}) \end{aligned}$$

The proof is finished. \square

Proof of Theorem 3.1 Applying the Taylor expansion to (2.9) and invoking Lemmas 5.3-5.5, we obtain that

$$\begin{aligned} l_n(\beta_0) &= 2 \sum_{i=1}^n \log(1 + \lambda^\top \hat{\eta}_i(\beta)) = n \left\{ \bar{\eta}_n^\top \Sigma^{-1} \bar{\eta}_n \right\} + o_p(p^{1/2}) \\ &= n \left\{ \bar{\eta}_n^\top (\Sigma^{-1} - S_{nc}^{-1}) \bar{\eta}_n \right\} + n \left\{ \bar{\eta}_n^\top S_{nc}^{-1} \bar{\eta}_n \right\} + o_p(p^{1/2}) \end{aligned}$$

From Lemma 5.4, we have

$$n \left\{ \bar{\eta}_n^\top (\Sigma^{-1} - S_{nc}^{-1}) \bar{\eta}_n \right\} = o_p(p^{1/2})$$

So

$$\begin{aligned} \frac{l_n(\beta_0) - T_{nc}}{p^{1/2}} &= \frac{n \left\{ \bar{\eta}_n^T (\Sigma^{-1} - S_{nc}^{-1}) \bar{\eta}_n \right\} + n \left\{ \bar{\eta}_n^T S_{nc}^{-1} \bar{\eta}_n \right\} + o_p(p^{1/2}) - n \left\{ \bar{\eta}_n^T S_{nc}^{-1} \bar{\eta}_n \right\}}{p^{1/2}} \\ &= \frac{n \left\{ \bar{\eta}_n^T (\Sigma^{-1} - S_{nc}^{-1}) \bar{\eta}_n \right\}}{p^{1/2}} \\ &= o_p(1) \end{aligned}$$

The proof is complete. \square

Acknowledgments

This work is supported by the National Natural Science Foundation of China (11201276, 61772017), the Natural Science Basic Research Plan in Shaanxi Province of China (2014JQ1042), and the Fundamental Research Funds for the Central Universities (GK201503015).

References

- [1] Chen, S. X. and Cui, H. J. *On Bartlett correction of empirical likelihood in the presence of nuisance parameters*, *Biometrika* **93**, 215-220, 2006.
- [2] Chen, S. X., Peng, L. and Qin, Y. L. *Effects of data dimension on empirical likelihood*, *Biometrika* **9**, 711-722, 2009.
- [3] DiCiccio, T. J., Hall, P. and Romano, J. P. *Empirical likelihood is bartlett-correctable*, *Ann. Statist.* **19**, 1053-1061, 1991.
- [4] Fan, J. and Huang, T. *Profile likelihood inference on semiparametric varying-coefficient partially linear models*, *Bernoulli* **11**, 1031-1057, 2005.
- [5] Guo, H., Zou, C. and Chen, B. *Empirical likelihood for high-dimensional linear regression models*, *Metrika* **77**, 921-945, 2014.
- [6] Huang, Z. S. and Zhang, R. Q. *Empirical likelihood for nonparametric parts in semiparametric varying-coefficient partially linear models*, *Statist. Probab. Lett.* **79**, 1798-1808, 2009.
- [7] Lam, C. and Fan, J. *Profile-kernel likelihood inference with diverging number of parameters*, *Ann. Statist.* **36**, 2232-2260, 2008.
- [8] Li, G., Lin, L. and Zhu, L. *Empirical likelihood for a varying coefficient partially linear model with diverging number of parameters*, *J. Multivar. Anal.* **105**, 85-111, 2012.
- [9] Liu, Y., Zou, C. and Wang, Z. *Calibration of the empirical likelihood for high-dimensional data*, *Ann. Inst. Stat. Math.* **65**, 529-550, 2013.
- [10] Owen, A. B. *Empirical likelihood ratio confidence intervals for single functionals*, *Biometrika* **75**, 237-249, 1988.
- [11] Owen, A. B. *Empirical likelihood ratio for linear models*, *Ann. Statist.* **19**, 1725-1747, 1991.
- [12] Owen, A. B. *Empirical Likelihood*, Chapman & Hall, New York, 2001.
- [13] Qin, J. and Lawless, J. *Empirical likelihood and general estimating equations*, *Ann. Statist.* **22**, 300-325, 1994.
- [14] Xia, Y. and Li, W.K. *On the estimation and testing of functional-coefficient linear models*, *Statistica Sinica* **9**, 735-757, 1999.
- [15] Xue, L. G. and Zhu, L. X. *Empirical Likelihood in Nonparametric and Semiparametric Models*, Science Press, Beijing, 2010.
- [16] You, J. and Zhou, Y. *Empirical likelihood for semiparametric varying-coefficient partially linear regression models*, *Statist. Probab. Lett.* **76**, 412-422, 2006.

A kernel density approach for replacing rounded zeros in compositional data sets

Jiajia Chen^{*†}, Xiaoqin Zhang[‡] and Shengjia Li[§]

Abstract

The logratio methodology widely used in compositional data analysis is not applicable when some components have rounded zeros. There are many univariate and multivariate methods that have been used to deal with rounded zeros. However, both of them have restrictions: the univariate methods replaced the rounded zeros only using the information of the corresponding component; the multivariate methods need to assume the distribution of transformed data. When the form of the distribution function is unknown, a multivariate nonparametric replacement approach is proposed in this paper. The proposed method uses conditional expected value based on isometric logratio coordinates to replace rounded zeros, in which the conditional density is estimated through multivariate Gauss kernel function. The permutation invariance and invariance under change of orthonormal basis are also presented. Simulation studies show that the proposed method has better performance than previous methods as the percentage of rounded zeros increases. The proposed method is also applied on the moss data from the Kola project.

Keywords: Compositional data, Isometric logratio coordinates, Rounded zeros, Gauss kernel function, Detection limit.

2000 AMS Classification: 62-07, 62G07, 62H11.

Received : 10.07.2017 *Accepted :* 28.06.2018 *Doi :* 10.15672/HJMS.2018.605

^{*}School of Statistics, Shanxi University of Finance and Economics, Taiyuan 030006, Shanxi, P.R. China Email: chenjiajia0401@163.com

[†]Corresponding Author.

[‡]School of Mathematical Sciences, Shanxi University, Taiyuan 030006, Shanxi, P.R. China Email: zhangxiaoqin@sxu.edu.cn

[§]School of Mathematical Sciences, Shanxi University, Taiyuan 030006, Shanxi, P.R. China Email: shjiali@sxu.edu.cn

1. Introduction

Compositional data, or compositions, are vectors in which all components are positive real numbers and carry only relative information [1]. These vectors can be represented as proportions using closure operation, that is, they multiplied by the appropriate scaling factors. Two vectors are compositional equivalent if they can be expressed in the same proportion, thus compositions can be viewed as equivalence classes, in which all vectors convey the same compositional information [18]. This type of data often occurs in geosciences, biosciences, economics and many other disciplines [1, 16, 18].

Compositional data provide information only about the relative magnitudes of the components, the logratio methodology plays a key role in compositional data analysis. Three logratio transformations including additive logratio (alr) transformation [1], centered logratio (clr) transformation [1] and isometric logratio (ilr) transformation [6] were proposed. The relationship between alr transformation and clr transformation is well known [1], and ilr transformation can be represented by means of alr transformation or clr transformation [6]. Because the alr transformation is non-isometric, and the clr transformation results in singular covariance matrix, the ilr transformation which can avoid the above drawbacks is suggested. The logratio transformations transform compositional data to coordinates in real space. However, zeros may exist in some components, thus the logratio transformations fail.

There are three kinds of zeros in compositional data set: rounded zeros, count zeros and essential zeros [9]. In this paper, we are interested in the rounded zero which is not true zero and results from the existence of value below a threshold. When the threshold is rounding-off error, the component is present in a very small quantity and rounded to zero; when the threshold is detection limit, the value below the detection limit cannot be observed and is commonly reported as zero. There are many classic methods in rounded zeros problem. Aitchison proposed the additive replacement strategy [1], but the ratios of components having no rounded zeros are not preserved, later the multiplicative replacement strategy [8] was proposed. Instead of replacing rounded zeros in a component by a fixed value, the multiplicative lognormal replacement method [13] allowing for random imputation was suggested. The multivariate method is the modified EM algorithm [15, 12], which assumed that the alr coordinates follow multivariate normal distribution. Later the robust modified EM algorithm working on ilr coordinates [10] was introduced. In addition, there are other algorithms, for example, the multiplicative Kaplan-Meier method [14] was proposed, which is a univariate method. The implementations of all these methods discussed above are available in the R package `zCompositions` [14].

The previous univariate methods replace rounded zeros based on the data of the corresponding component and perform poorly when the proportion of rounded zeros is high. The multivariate methods for rounded zeros usually rely on the underlying assumption of multivariate normality in the space of coordinates. Furthermore, the modified EM algorithm based on alr coordinates requires that at least one component has no rounded zeros. To avoid these disadvantages, a new multivariate nonparametric replacement method based on multivariate Gauss kernel density estimation is proposed in this paper. To illustrate the performance of proposed method compared with the existing methods, this method is applied to both simulation and example analysis.

The rest of this paper is organized as follows. Some basic concepts about compositional data are reviewed in Section 2. In Section 3, the proposed approach is presented. Simulation study and real example are given in Section 4 to verify the effectiveness and usefulness of proposed method. Section 5 concludes this paper.

2. Preliminaries

Let $\mathbf{x} = [x_1, x_2, \dots, x_D]$ be a row vector denoting a D -part composition represented with constant sum k , its sample space is the simplex \mathcal{S}^D [1] defined as

$$\mathcal{S}^D = \left\{ \mathbf{x} = [x_1, x_2, \dots, x_D] \left| x_i > 0, i = 1, 2, \dots, D; \sum_{i=1}^D x_i = k \right. \right\},$$

where the constant k is an arbitrary positive real number and is usually 1 or 100 depending on the units of measurement. The simplex is a Euclidean vector space structure [1, 17, 3] when defining inner product with its related norm and Aitchison distance [2]. The distance between two compositions \mathbf{x} and $\mathbf{y} \in \mathcal{S}^D$ is

$$d_a(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^D \left(\ln \frac{x_i}{g_m(\mathbf{x})} - \ln \frac{y_i}{g_m(\mathbf{y})} \right)^2 \right)^{1/2},$$

where $d_a(\cdot, \cdot)$ stands for the Aitchison distance in \mathcal{S}^D , and $g_m(\mathbf{x})$ denotes the geometric mean of the parts of \mathbf{x} .

The ilr transformation [6] assigns coordinates with respect to the given orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{D-1}\}$ of the simplex \mathcal{S}^D . An orthonormal basis can be obtained through sequential binary partition of parts of a composition [5]. Following the reference [5], we can construct a $(D-1) \times D$ matrix Ψ in which rows are

$$(2.1) \quad \psi_i = \sqrt{\frac{D-i}{D-i+1}} \left[0, \dots, 0, 1, \underbrace{-\frac{1}{D-i}, \dots, -\frac{1}{D-i}}_{D-i} \right], \quad i = 1, 2, \dots, D-1,$$

respectively. An orthonormal basis can be obtained through $\mathbf{e}_i = \mathcal{C}(\exp \psi_i)$ ($i = 1, 2, \dots, D-1$), where \mathcal{C} is the closure operation. Thus the composition $\mathbf{x} \in \mathcal{S}^D$ is transformed to ilr coordinates $\mathbf{z} = \text{ilr}(\mathbf{x}) = [z_1, z_2, \dots, z_{D-1}] \in \mathbb{R}^{D-1}$, where

$$(2.2) \quad z_i = \sqrt{\frac{D-i}{D-i+1}} \ln \frac{x_i}{\sqrt[D-i]{\prod_{j=i+1}^D x_j}}, \quad i = 1, 2, \dots, D-1.$$

The ilr coordinates guarantee the invariance of distance, that is, $d_a(\mathbf{x}, \mathbf{y}) = d(\text{ilr}(\mathbf{x}), \text{ilr}(\mathbf{y}))$, where $d(\cdot, \cdot)$ is the Euclidean distance in real space. The inverse mapping of any real-valued vector $\mathbf{z} \in \mathbb{R}^{D-1}$ to the original composition \mathbf{x} is then given by

$$(2.3) \quad \begin{cases} x_1 = \exp \left\{ \sqrt{\frac{D-1}{D}} z_1 \right\}, \\ x_i = \exp \left\{ -\sum_{j=1}^{i-1} \frac{1}{\sqrt{(D-j+1)(D-j)}} z_j + \sqrt{\frac{D-i}{D-i+1}} z_i \right\}, \quad i = 2, \dots, D-1, \\ x_D = \exp \left\{ -\sum_{j=1}^{D-1} \frac{1}{\sqrt{(D-j+1)(D-j)}} z_j \right\}. \end{cases}$$

The compositions can be viewed as equivalence classes, therefore the obtained composition \mathbf{x} can be represented as constant sum vectors.

For any random composition $\mathbf{x} = [x_1, x_2, \dots, x_D]$, the measure of dispersion is the variation matrix [1] defined as

$$(2.4) \quad \mathbf{T} = [t_{ij}]_{D \times D}, \quad t_{ij} = \text{Var} \left(\ln \frac{x_i}{x_j} \right),$$

where the element in variation matrix is the logratio variance for any two parts i and j of a D -part composition \mathbf{x} .

3. Kernel density replacement approach

Consider a random composition $\mathbf{x} = [x_1, x_2, \dots, x_D]$, the sample data set is \mathbf{X} with n compositions and D -part, that is

$$\mathbf{X} = [x_{ij}]_{n \times D} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1D} \\ x_{21} & x_{22} & \cdots & x_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nD} \end{pmatrix}.$$

Suppose that the compositional data set \mathbf{X} has rounded zeros, the corresponding threshold matrix is denoted as $\mathbf{E} = [e_{ij}]_{n \times D}$, where e_{ij} is the threshold of x_{ij} . Let $R_j \subset \{1, 2, \dots, n\}$ be the row indices referring to the rounded zeros of the j th component ($j \in \{1, 2, \dots, D\}$), then $O_j = \{1, 2, \dots, n\} \setminus R_j$ refers to the remaining row indices of the j th component, that is, $R_j = \{i : i \in \{1, 2, \dots, n\}, x_{ij} \leq e_{ij}\}$, $O_j = \{i : i \in \{1, 2, \dots, n\}, x_{ij} > e_{ij}\}$.

Firstly, we initialize the rounded zeros by multiplicative replacement strategy in which the rounded zero is equal to 65% of the threshold [8], thus \mathbf{X} denotes the replaced data set. Denote the ilr coordinates in Equation (2.2) of random composition \mathbf{x} as $\mathbf{z} = \text{ilr}(\mathbf{x}) = [z_1, z_2, \dots, z_{D-1}] = [z_1, \mathbf{z}_{-1}]$, where \mathbf{z}_{-1} refers to the remaining components of \mathbf{z} except for the first component. Then initialized data set \mathbf{X} is transformed to real data set $\mathbf{Z} = [z_{ij}]_{n \times (D-1)}$, where each row in \mathbf{Z} is the ilr coordinates of the corresponding composition in \mathbf{X} . For the element e_{i1} in threshold set \mathbf{E} , the ilr transformation of rounded zero $x_{i1} < e_{i1}$ can result in the unknown value z_{i1} less than ψ_{i1} , where

$$\psi_{i1} = \sqrt{\frac{D-1}{D}} \ln \frac{e_{i1}}{\sqrt[D-1]{\prod_{j=2}^D x_{ij}}}.$$

In the proposed approach, the unknown data z_{i1} ($i \in R_1$) is imputed by conditional expected value

$$(3.1) \quad E(z_1 | \mathbf{z}_{-1} = \mathbf{z}_{i,-1}, z_1 < \psi_{i1}) = \frac{\int_{-\infty}^{\psi_{i1}} z_1 f(z_1 | \mathbf{z}_{-1} = \mathbf{z}_{i,-1}) dz_1}{\int_{-\infty}^{\psi_{i1}} f(z_1 | \mathbf{z}_{-1} = \mathbf{z}_{i,-1}) dz_1},$$

where $\mathbf{z}_{i,-1}$ is the i th row of \mathbf{Z} except for the first column, the conditional density function $f(z_1 | \mathbf{z}_{-1} = \mathbf{z}_{i,-1})$ can be calculated as follows

$$(3.2) \quad f(z_1 | \mathbf{z}_{-1} = \mathbf{z}_{i,-1}) = \frac{f(z_1, \mathbf{z}_{-1} = \mathbf{z}_{i,-1})}{f(\mathbf{z}_{-1} = \mathbf{z}_{i,-1})} = \frac{f(z_1, \mathbf{z}_{-1} = \mathbf{z}_{i,-1})}{\int_{-\infty}^{+\infty} f(z_1, \mathbf{z}_{-1} = \mathbf{z}_{i,-1}) dz_1}.$$

Regardless the distribution of multivariate random variable \mathbf{z} , the density function $f(z_1, \mathbf{z}_{-1} = \mathbf{z}_{i,-1})$ can be estimated by multivariate Gauss kernel density [4]. In this

paper, the same bandwidth h is applied to different coordinate direction, thus

$$\begin{aligned}
& \widehat{f}(z_1, \mathbf{z}_{-1} = \mathbf{z}_{i,-1}) \\
&= \frac{1}{n(\sqrt{2\pi}h)^{D-1}} \sum_{k=1}^n \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right. \\
&\quad \left. - \frac{1}{2} \left(\frac{\mathbf{z}_{i,-1} - \mathbf{z}_{k,-1}}{h} \right) \left(\frac{\mathbf{z}_{i,-1} - \mathbf{z}_{k,-1}}{h} \right)^T \right\} \\
(3.3) \quad &= \frac{1}{n(\sqrt{2\pi}h)^{D-1}} \sum_{k=1}^n \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 - \frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\}.
\end{aligned}$$

The bandwidth h is given by $h = \sigma \left(\frac{4}{n(D+1)} \right)^{\frac{1}{D+3}}$ [20], where $\sigma^2 = \frac{1}{D-1} \sum_{j=1}^{D-1} \text{Var}(z_j) = \frac{1}{D-1} \text{tr}(\text{Var}(\mathbf{z}))$, tr represents the trace of matrix $\text{Var}(\mathbf{z})$.

It follows from Equation (3.2) and Equation (3.3) that

$$\begin{aligned}
& \widehat{f}(z_1 | \mathbf{z}_{-1} = \mathbf{z}_{i,-1}) \\
&= \frac{\sum_{k=1}^n \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\}}{\sum_{k=1}^n \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} dz_1} \\
(3.4) \quad &= \frac{\sum_{k=1}^n \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\}}{\sum_{k=1}^n \sqrt{2\pi}h \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\}},
\end{aligned}$$

By conditional density function in Equation (3.4), Equation (3.1) can be expressed as

$$(3.5) \quad \frac{\sum_{k=1}^n \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\} \int_{-\infty}^{\psi_{i1}} z_1 \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} dz_1}{\sum_{k=1}^n \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\} \int_{-\infty}^{\psi_{i1}} \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} dz_1}.$$

Since

$$\int_{-\infty}^{\psi_{i1}} \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} dz_1 = \sqrt{2\pi}h\Phi \left(\frac{\psi_{i1} - z_{k1}}{h} \right)$$

and

$$\begin{aligned}
& \int_{-\infty}^{\psi_{i1}} z_1 \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} dz_1 \\
&= \int_{-\infty}^{\psi_{i1}} (z_1 - z_{k1}) \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} dz_1 + z_{k1} \int_{-\infty}^{\psi_{i1}} \exp \left\{ -\frac{1}{2} \left(\frac{z_1 - z_{k1}}{h} \right)^2 \right\} dz_1 \\
&= \sqrt{2\pi}h \left(-h\phi \left(\frac{\psi_{i1} - z_{k1}}{h} \right) + z_{k1}\Phi \left(\frac{\psi_{i1} - z_{k1}}{h} \right) \right),
\end{aligned}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and distribution function of the standard normal distribution, respectively. Thus Equation (3.5) can be simplified as

$$(3.6) \quad \frac{E(z_1 | \mathbf{z}_{-1} = \mathbf{z}_{i,-1}, z_1 < \psi_{i1}) = \sum_{k=1}^n \left(-h\phi \left(\frac{\psi_{i1} - z_{k1}}{h} \right) + z_{k1}\Phi \left(\frac{\psi_{i1} - z_{k1}}{h} \right) \right) \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\}}{\sum_{k=1}^n \Phi \left(\frac{\psi_{i1} - z_{k1}}{h} \right) \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) \right\}}.$$

Hence, the unknown data z_{i1} is imputed by Equation (3.6). For the ilr coordinates in Equation (2.2), since $d(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1}) = d_a(\mathbf{x}_{i,-1}, \mathbf{x}_{k,-1})$, the imputed value z_{i1} is related with the Aitchison distance between subcompositions $\mathbf{x}_{i,-1}$ and $\mathbf{x}_{k,-1}$, where $\mathbf{x}_{i,-1}$ and $\mathbf{x}_{k,-1}$ denote the remaining components of compositions \mathbf{x}_i and \mathbf{x}_k except for the first component, respectively.

3.1. Property. *The imputed value $E(z_1|\mathbf{z}_{-1} = \mathbf{z}_{i,-1}, z_1 < \psi_{i1})$ in Equation (3.6) has the following properties:*

- (1) *It is below the threshold, that is, $E(z_1|\mathbf{z}_{-1} = \mathbf{z}_{i,-1}, z_1 < \psi_{i1}) < \psi_{i1}$.*
- (2) *It is unchanged when the remaining components of \mathbf{x} except for the first component are arbitrarily permuted.*
- (3) *It is invariant under change of orthonormal basis $\{\mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_{D-1}\}$.*

Property 3.1 is quite obvious. It follows from $\mathbf{z} = \mathbf{x}\Psi^T$ that $\text{tr}(\text{Var}(\mathbf{z})) = \text{tr}(\text{Var}(\mathbf{x}\Psi^T)) = \text{tr}(\Psi\text{Var}(\mathbf{x})\Psi^T) = \text{tr}(\text{Var}(\mathbf{x})\Psi^T\Psi) = \text{tr}(\text{Var}(\mathbf{x})\mathbf{G}_D)$ [18], where $\mathbf{G}_D = \mathbf{I}_D - \frac{1}{D}\mathbf{J}_D$, \mathbf{I}_D is a identity matrix, \mathbf{J}_D is a matrix of units. Therefore all the underlying elements ($d(\mathbf{z}_{i,-1}, \mathbf{z}_{k,-1})$, h , z_{k1} and ψ_{i1}) are invariant by permutation and change of basis, thus the imputed value in Equation (3.6) is unchanged.

Property 3.1 (2) and (3) point out that $E(z_1|\mathbf{z}_{-1} = \mathbf{z}_{i,-1}, z_1 < \psi_{i1})$ satisfies permutation invariance and invariance under change of orthonormal basis, but $E(z_l|\mathbf{z}_{-l} = \mathbf{z}_{i,-l}, z_l < \psi_{il})$ ($l = 2, \dots, D-1$) may not satisfy these two properties, for example, z_{kl} may changed when the remaining components of \mathbf{x} except for the l th component are arbitrarily permuted. To replace the rounded zeros in the l th component of \mathbf{x} , we define the permuted composition $\mathbf{x}^{(l)} = [x_1^{(l)}, x_2^{(l)}, \dots, x_l^{(l)}, x_{l+1}^{(l)}, \dots, x_D^{(l)}] = [x_l, x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_D]$. The ilr coordinates are denoted as $\mathbf{z}^{(l)} = \text{ilr}(\mathbf{x}^{(l)}) = [z_1^{(l)}, z_2^{(l)}, \dots, z_{D-1}^{(l)}] = [z_1^{(l)}, \mathbf{z}_{-1}^{(l)}]$, the corresponding ilr data set is $\mathbf{Z}^{(l)} = [z_{ij}^{(l)}]_{n \times (D-1)}$. According to Equation (3.6), the unknown data $z_{i1}^{(l)}$ ($i \in R_l$) resulting from the rounded zero in the i th row and the l th component of \mathbf{X} can be imputed by

$$(3.7) \quad \frac{E(z_1^{(l)}|\mathbf{z}_{-1}^{(l)} = \mathbf{z}_{i,-1}^{(l)}, z_1^{(l)} < \psi_{i1}^{(l)}) = \sum_{k=1}^n \left(-h\phi \left(\frac{\psi_{i1}^{(l)} - z_{k1}^{(l)}}{h} \right) + z_{k1}^{(l)} \Phi \left(\frac{\psi_{i1}^{(l)} - z_{k1}^{(l)}}{h} \right) \right) \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}^{(l)}, \mathbf{z}_{k,-1}^{(l)}) \right\}}{\sum_{k=1}^n \Phi \left(\frac{\psi_{i1}^{(l)} - z_{k1}^{(l)}}{h} \right) \exp \left\{ -\frac{1}{2h^2} d^2(\mathbf{z}_{i,-1}^{(l)}, \mathbf{z}_{k,-1}^{(l)}) \right\}},$$

where $\psi_{i1}^{(l)} = \sqrt{\frac{D-1}{D}} \ln \frac{e_{il}}{\sqrt[D-1]{\prod_{j=2}^D x_{ij}^{(l)}}}$.

The specific steps of the proposed method, similar to the modified EM algorithm based on ilr coordinates [10], are as follows:

Step 1: Sort the parts of compositional data set according to the number of rounded zeros of each part. The ilr coordinates in Equation (2.2) is used in the proposed method, the first component is only included in the first ilr coordinate. In order to reduce the error, the component with more zeros should be put in the first column. Without loss of generality, assume that the parts are already sorted, i.e. $|R_1| \geq |R_2| \geq \dots \geq |R_D|$, where $|R_j|$ denotes the number of elements of R_j ($j = 1, 2, \dots, D$).

Step 2: Initialize the rounded zeros by multiplicative replacement strategy.

Step 3: Set $l = 1$.

Step 4: Replace the unknown data $z_{i1}^{(l)}$ ($i \in R_l$) using Equation (3.7).

Step 5: Inverse the every row of updated data set using Equation (2.3).

Step 6: Carry out Steps 4-5 for each $l = 2, 3, \dots, |C|$, where $C = \{j : j \in \{1, 2, \dots, D\}, |R_j| \neq 0\}$ is the index set of parts containing at least one rounded zero.

Step 7: Repeat Steps 3-6 until the Euclidean distance between the variation matrix of compositional data set from the present and the previous iteration is smaller than a certain boundary.

Step 8: Sort the parts of replaced compositional data set in the original order.

If the data set $\mathbf{X} = [x_{ij}]_{n \times D}$ is closed to a constant, then the replaced data set is $\hat{\mathbf{X}} = [\hat{x}_{ij}]_{n \times D}$ obtained from the above algorithm, otherwise, we should rescale the replaced value \hat{x}_{ij} using the expression [14]

$$(3.8) \quad \hat{x}_{ij}^* = \hat{x}_{ij} \frac{x_{ik}}{\hat{x}_{ik}}, \quad j \in C, i \in R_j,$$

where \hat{x}_{ij}^* is the rescaled value, x_{ik} is the originally observed element in the i th row and k th column of compositional data set \mathbf{X} , \hat{x}_{ik} is the corresponding replaced value in $\hat{\mathbf{X}}$.

4. Simulation and Example

In this section we present simulation study and real example in order to illustrate the good performance of proposed method (multK), which is compared with the multiplicative replacement strategy (multR), the multiplicative Kaplan-Meier method (multKM), the multiplicative lognormal replacement method (multLN), the modified EM algorithm working on alr coordinates (alrEM) and the robust modified EM algorithm working on ilr coordinates (ilrEM). Given the original compositional data set \mathbf{X} which has no rounded zeros, we set the value below the threshold as zero, the replaced compositional data set is denoted as \mathbf{X}^* . We consider two measures of distortion, standardized residual sum of squares (STRESS) [8] and relative difference in variation matrix (RDVM) [13]. Denote the variation matrix in Equation (2.4) of original data set \mathbf{X} and imputed data set \mathbf{X}^* as $\mathbf{T} = [t_{ij}]_{D \times D}$ and $\mathbf{T}^* = [t_{ij}^*]_{D \times D}$, the two measures STRESS and RDVM are defined as

$$\text{STRESS} = \frac{\sum_{i < j} (d_a(\mathbf{x}_i, \mathbf{x}_j) - d_a(\mathbf{x}_i^*, \mathbf{x}_j^*))^2}{\sum_{i < j} d_a^2(\mathbf{x}_i, \mathbf{x}_j)},$$

and

$$\text{RDVM} = \frac{1}{2|C|D - |C|^2} \sum_{i, j \in C} \frac{|t_{ij}^* - t_{ij}|}{t_{ij}},$$

respectively, where \mathbf{x}_i is the i th row of data set \mathbf{X} . The two measures STRESS and RDVM represent the distance difference and variation difference, respectively.

4.1. Simulation Study. In this subsection, several simulation studies were conducted. We first simulated real data set with sample size 300 from multivariate normal distribution $N_4(\mu, \Sigma)$, then the compositional data set \mathbf{X} can be obtained through ilr-inverse transformation in Equation (2.3). Suppose that the rounded zero is resulting from value below the detection limit, and the detection limits of same part-different compositions are the same, so the detection limit set is denoted as a vector, that is, $\mathbf{E} = [e_1, e_2, \dots, e_5]$, where e_j ($j = 1, 2, \dots, 5$) is the α_j quantile of the j th component in \mathbf{X} .

We set mean $\mu = [-2, -1.5, -1, -0.3]$ and covariance matrix $\Sigma = [\rho^{|i-j|}]_{4 \times 4}$. To describe different levels of correlations among the components, take $\rho = 0.3, 0.5, 0.7$ and 0.9 . Ten situations of detection limit set are conducted, where $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ range from 0.05 to 0.5 by 0.05, 0.04 to 0.4 by 0.04, 0.03 to 0.3 by 0.03, 0.02 to 0.2 by 0.02, respectively,

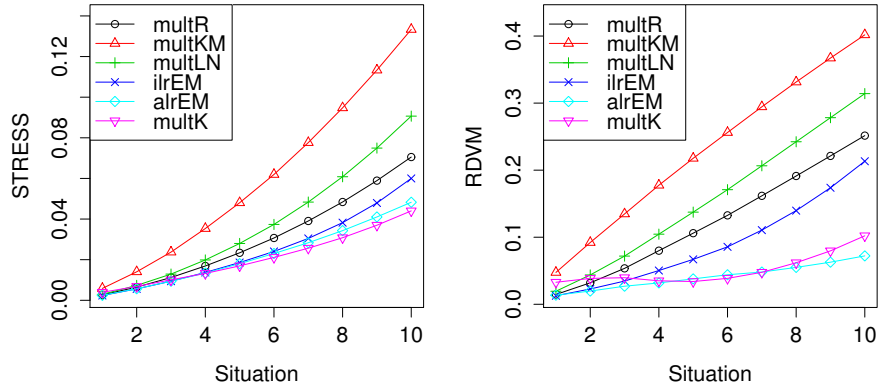
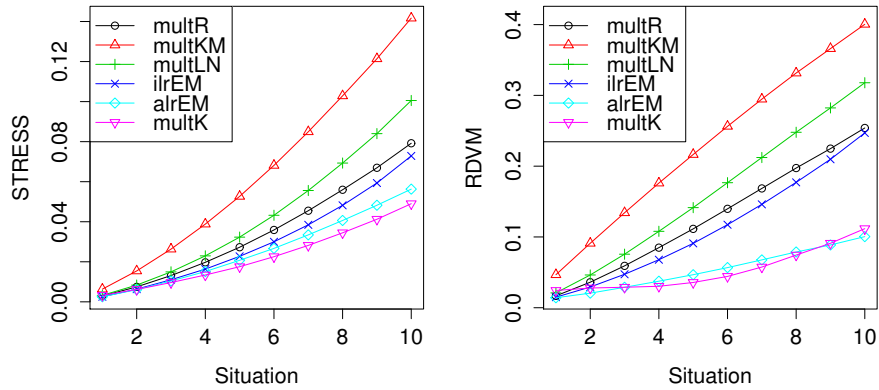
(a) $\rho = 0.3$.(b) $\rho = 0.5$.

Figure 1. Two measures of distortion STRESS and RDVM for six methods (multR, multKM, multLN, ilrEM, alrEM, multK) under ten situations of detection limit set when $\rho = 0.3$ (a) and $\rho = 0.5$ (b).

and $\alpha_5 = 0$. Set each data in the j th component smaller than e_j ($j = 1, 2, 3, 4$) to a zero value, then the percentage of rounded zeros in the first four components approximately range from 5% to 50% by 5%, 4% to 40% by 4%, 3% to 30% by 3%, 2% to 20% by 2%, respectively, and the last component has no rounded zeros, therefore the corresponding percentage of rounded zeros in compositional data set approximately ranges from 2.8% to 28% by 2.8%.

We run 100 Monte Carlo simulations for each setting described above. The performance comparisons among previous methods and proposed method with varying percentage of rounded zeros corresponding to situations are showed in Figure 1 and Figure

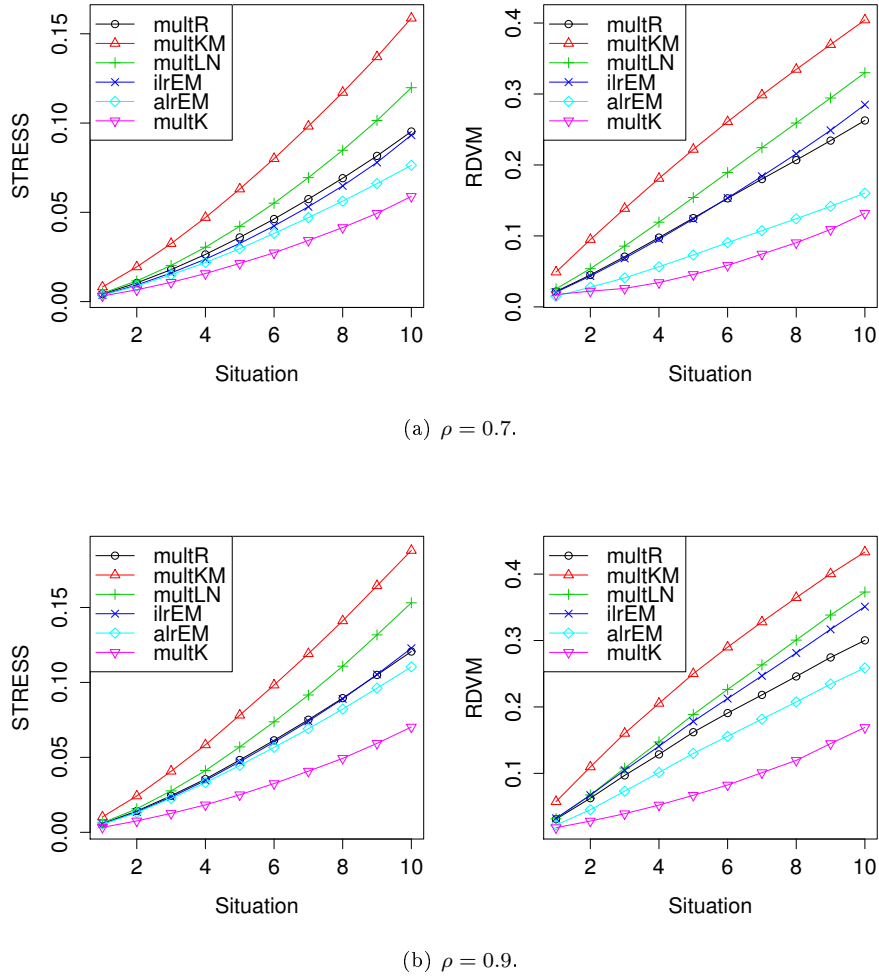


Figure 2. Two measures of distortion STRESS and RDVM for six methods (multR, multKM, multLN, ilrEM, alrEM, multK) under ten situations of detection limit set when $\rho = 0.7$ (a) and $\rho = 0.9$ (b).

2. The values in Figure 1 and Figure 2 are the average STRESS or RDVM of 100 simulations. Figure 1(a) and Figure 1(b) depict the trends in two performance measures under ten situations of detection limit set when $\rho = 0.3$ and 0.5 . It can be seen from Figure 1(a) and Figure 1(b) that the ilrEM and alrEM have smaller STRESS and RDVM than those of multR, however, the STRESS and RDVM of multKM and multLN are greater than those of multR. Moreover, when the percentage of rounded zeros increases, the STRESS value of multK is lower than those of previous methods. The multK method performs worse than previous methods in the measure RDVM when $\rho = 0.3$, whereas it performs better under some situations when $\rho = 0.5$. Figure 2 shows the trends in two measures

among different methods when $\rho = 0.7$ and 0.9 . From Figure 2(a) and Figure 2(b), we see that the multK method outperforms the other methods in two measures STRESS and RDVM. The STRESS value of ilrEM is very close to that of multR, while ilrEM performs worse than multR in measure RDVM. To sum up, when the percentage of rounded zeros increases, the proposed method has better performance than other methods in the two measures STRESS and RDVM.

4.2. Real example. The proposed method discussed in the previous section will be applied to the moss data from the Kola project available in the R package StatDA [7] and compared with the previous methods (multR, multKM, multLN, ilrEM and alrEM). The moss data set consists of more than 50 chemical elements and 594 observations. We focus on the 7-part subcomposition [Al, Ca, Fe, K, Mg, Na, Si] denoted as compositional data set $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_7]$ with constant sum 100%, which has no rounded zeros. Similar to the simulation analysis, we give the detection limit set, the value below detection limit is set as zero. The aim of this study is to replace rounded zeros using different methods.

Suppose that the components $\mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_6$ and \mathbf{u}_7 have rounded zeros. Eight situations of detection limit set are given in Table 1 in which e_j ($j = 1, 3, 6, 7$) is the detection limit of the j th component. Table 1 also gives the percentages of rounded zeros of components $\mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_6, \mathbf{u}_7$ and the total percentage of rounded zeros of compositional data set \mathbf{U} . Table 2 gives the computed results of STRESS and RDVM for six methods (multR, multKM, multLN, ilrEM, alrEM, multK) under eight situations. According to Table 2, we can find that the proposed method has smaller STRESS value than those of other methods except the first two situations, and the RDVM value of proposed method for each situation is always smaller than other methods. In addition, multR performs better than ilrEM and alrEM as the percentage of rounded zeros increases, of which alrEM has larger STRESS and RDVM than ilrEM. This is because that the ilrEM and alrEM all assume the distribution of compositional data set. In fact, compositional data set \mathbf{U} departures from normal distribution on the simplex [11], which is tested using the energy test [19] or the test based on SVD including the marginal univariate tests, the bivariate tests and radius tests [21]. Because the ilrEM is a robust method, which performs better than alrEM. These results suggest that the proposed method is superior to the others in the case of moss data set.

Table 1. Eight situations of detection limit set for compositional data set \mathbf{U} . The value in parentheses is the percentage of rounded zeros of the corresponding component. The last column ZR represents the total percentage of rounded zeros of the corresponding situation (Unit: %).

situation	e_1	e_3	e_6	e_7	ZR
1	1.39(14.14)	1.41(13.97)	0.41(14.65)	1.41(14.31)	8.15
2	1.51(18.86)	1.56(18.69)	0.46(19.53)	1.55(19.19)	10.89
3	1.63(23.57)	1.72(23.23)	0.51(24.41)	1.68(23.91)	13.59
4	1.76(28.28)	1.85(27.95)	0.56(29.29)	1.76(28.62)	16.31
5	1.84(33.00)	1.96(32.49)	0.60(34.18)	1.86(33.50)	19.02
6	1.93(37.71)	2.04(37.21)	0.66(39.06)	1.98(38.22)	21.74
7	2.01(42.42)	2.22(41.75)	0.72(43.94)	2.05(42.93)	24.43
8	2.12(47.14)	2.38(46.46)	0.78(48.82)	2.13(47.81)	27.18

Table 2. Two evaluation indexes STRESS and RDVM of methods (multR, multKM, multLN, ilrEM, alrEM, multK) for compositional data set \mathbf{U} under eight situations of detection limit set.

	situation	multR	multKM	multLN	ilrEM	alrEM	multK
STRESS	1	0.0179	0.0317	0.0166	0.0148	0.0158	0.0159
	2	0.0216	0.0426	0.0213	0.0182	0.0212	0.0189
	3	0.0244	0.0567	0.0275	0.0222	0.0260	0.0218
	4	0.0283	0.0706	0.0348	0.0265	0.0336	0.0257
	5	0.0328	0.0833	0.0422	0.0312	0.0444	0.0302
	6	0.0372	0.0994	0.0518	0.0372	0.0592	0.0358
	7	0.0425	0.1185	0.0638	0.0468	0.0737	0.0421
	8	0.0494	0.1382	0.0774	0.0613	0.1042	0.0493
RDVM	1	0.0623	0.1626	0.0645	0.0478	0.0465	0.0389
	2	0.0671	0.2033	0.0864	0.0544	0.0679	0.0396
	3	0.0551	0.2475	0.1165	0.0724	0.0839	0.0401
	4	0.0538	0.2849	0.1423	0.0863	0.1025	0.0376
	5	0.0576	0.3161	0.1653	0.0979	0.1442	0.0425
	6	0.0688	0.3513	0.1959	0.1292	0.1988	0.0630
	7	0.0821	0.3891	0.2311	0.1771	0.2532	0.0793
	8	0.0954	0.4252	0.2681	0.2376	0.3568	0.0950

5. Conclusions

The logratio transformations do not apply when compositional data have zeros. In this paper, a nonparametric method based on the multivariate Gauss kernel density estimation is suggested to deal with the rounded zeros. Because the clr coordinates add to zero, the ilr coordinates are applied in the proposed method. Under the ilr coordinates in Equation (2.2), the multivariate Gauss kernel function is related with the Aitchison distance between subcompositions. In the simulation study and real example, the proposed method is compared with the multiplicative replacement strategy, the multiplicative Kaplan-Meier method, the multiplicative lognormal replacement method, the modified EM algorithm based on alr coordinates and the robust modified EM algorithm based on ilr coordinates. The results in simulation study show that the proposed method presents a good performance in comparison with other methods in the two measures

STRESS and RDVM as the percentage of rounded zeros increases. Furthermore, in the real example, the performance of proposed method is obvious. The feature of our framework is that the proposed method works when the distribution function form is unknown. Future work will be dedicated to the study of bandwidth matrix in multivariate kernel function.

Acknowledgements

This work was supported by the Natural Science Foundation of Shanxi Province of China (No. 2015011044), Shanxi International Science and Technology Cooperation Project (No. 2015081020), and Graduate Education Innovation Project of Shanxi Province (No. 2017BY001).

References

- [1] Aitchison, J. *The statistical analysis of compositional data*, Monographs on Statistics and Applied Probability, Chapman & Hall, London, 1986.
- [2] Aitchison, J., Barceló-Vidal, C., Martín-Fernández, J.A. and Pawlowsky-Glahn, V. *Logratio analysis and compositional distance*, *Mathematical Geosciences* **32** (3), 271-275, 2000.
- [3] Billheimer, D., Guttorp, P. and Fagan, W.F. *Statistical interpretation of species composition*, *Journal of the American Statistical Association* **96** (456), 1205-1214, 2001.
- [4] Chacón, J.E., Mateu-Figueras, G. and Martín-Fernández, J.A. *Gaussian kernels for density estimation with compositional data*, *Computers & Geosciences* **37** (5), 702-711, 2011.
- [5] Egozcue, J.J. and Pawlowsky-Glahn, V. *Groups of parts and their balances in compositional data analysis*, *Mathematical Geosciences* **37** (7), 795-828, 2005.
- [6] Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G. and Barceló-Vidal, C. *Isometric logratio transformations for compositional data analysis*, *Mathematical Geosciences* **35** (3), 279-300, 2003.
- [7] Filzmoser, P. *StatDA: statistical analysis for environmental data, R package version 1.6.3*, 2011.
- [8] Martín-Fernández, J.A., Barceló-Vidal, C. and Pawlowsky-Glahn, V. *Dealing with zeros and missing values in compositional data sets using nonparametric imputation*, *Mathematical Geosciences* **35** (3), 253-278, 2003.
- [9] Martín-Fernández, J.A., Palarea-Albaladejo, J. and Olea, R.A. *Dealing with zeros*, in *Compositional Data Analysis: Theory and Applications*, V. Pawlowsky-Glahn and A. Buccianti, eds., John Wiley & Sons Ltd., Chichester, 47-62, 2011.
- [10] Martín-Fernández, J.A., Hron, K., Templ, M., Filzmoser, P. and Palarea-Albaladejo, J. *Model-based replacement of rounded zeros in compositional data: classical and robust approaches*, *Computational Statistics & Data Analysis* **56** (9), 2688-2704, 2012.
- [11] Mateu-Figueras, G., Pawlowsky-Glahn, V. and Egozcue, J.J. *The normal distribution in some constrained sample spaces*, *Sort-statistics and Operations Research Transactions* **37** (1), 29-56, 2008.
- [12] Palarea-Albaladejo J. and Martín-Fernández, J.A. *A modified EM algorithm for replacing rounded zeros in compositional data sets*, *Computers & Geosciences* **34** (8), 902-917, 2008.
- [13] Palarea-Albaladejo, J. and Martín-Fernández, J.A. *Values below detection limit in compositional chemical data*, *Analytica Chimica Acta* **764**, 32-43, 2013.
- [14] Palarea-Albaladejo, J. and Martín-Fernández, J.A. *zCompositions-R package for multivariate imputation of left-censored data under a compositional approach*, *Chemometrics and Intelligent Laboratory Systems* **143**, 85-96, 2015.
- [15] Palarea-Albaladejo, J., Martín-Fernández, J.A. and Gómez-García, J. *A parametric approach for dealing with compositional rounded zeros*, *Mathematical Geosciences* **39** (7), 625-645, 2007.
- [16] Pawlowsky-Glahn, V. and Buccianti, A. *Compositional Data Analysis: Theory and Applications*, John Wiley & Sons Ltd., Chichester, 2011.

- [17] Pawlowsky-Glahn, V. and Egozcue, J.J. *Geometric approach to statistical analysis on the simplex*, Stochastic Environmental Research and Risk Assessment **15** (5), 384-398, 2001.
- [18] Pawlowsky-Glahn, V., Egozcue, J.J. and Tolosana-Delgado, R. *Modeling and analysis of compositional data*, Statistics in Practice, John Wiley & Sons, Ltd., Chichester, 2015.
- [19] Rizzo, M.L. and Székely, G.J. *Energy: E-statistics (energy statistics)*, R package version 1.1-0, 2008.
- [20] Silverman, B.W. *Density estimation for statistics and data analysis*, Chapman & Hall, London, 1986.
- [21] van den Boogaart, K.G. and Tolosana-Delgado, R. *Analyzing compositional data with R*, Springer, Heidelberg, 2013.

Equivariant estimation of quantile vector of two normal populations with a common mean

Manas Ranjan Tripathy^{*†}, Adarsha Kumar Jena[‡] and Somesh Kumar[§]

Abstract

The problem of estimating quantile vector $\underline{\theta} = (\theta_1, \theta_2)$ of two normal populations, under the assumption that the means (μ_i s) are equal has been considered. Here $\theta_i = \mu + \eta\sigma_i$, $i = 1, 2$, denotes the p^{th} quantile of the i^{th} population, where $\eta = \Phi^{-1}(p)$, $0 < p < 1$, and Φ denotes the c.d.f. of a standard normal random variable. The loss function is taken as sum of the quadratic losses. First, a general result has been proved which helps in constructing some improved estimators for the quantile vector $\underline{\theta}$. Further, classes of equivariant estimators have been proposed and sufficient conditions for improving estimators in these classes are derived. In the process, two complete class results have been proved. A numerical comparison of these estimators are done and recommendations have been made for the use of these estimators. Finally, we conclude our results with some practical examples.

Keywords: Equivariant estimator, Estimation of quantiles, Complete class results, Common mean, Inadmissibility, Relative risk comparison.

2000 AMS Classification: 62C15, 62F10, 62C20

Received : 22.08.2017 *Accepted :* 28.06.2018 *Doi :* 10.15672/HJMS.2018.606

^{*}Department of Mathematics, National Institute of Technology Rourkela, Rourkela-769 008, India, Email: manas@nitrkl.ac.in

[†]Corresponding Author.

[‡]Department of Mathematics, National Institute of Technology Rourkela, Rourkela-769 008, India, Email: adarsha@gmail.com

[§]Department of Mathematics, Indian Institute of Technology Kharagpur, Kharagpur-721 302, India, Email: smsh@iitkgp.ac.in

1. Introduction

Let $\underline{X} = (X_1, X_2, \dots, X_m)$ and $\underline{Y} = (Y_1, Y_2, \dots, Y_n)$ be independent random samples drawn from two normal populations $N(\mu, \sigma_1^2)$ and $N(\mu, \sigma_2^2)$ respectively. Here the common mean μ , and the variances σ_1^2, σ_2^2 are unknown. The p^{th} quantile of the first and second populations are $\theta_1 = \mu + \eta\sigma_1$ and $\theta_2 = \mu + \eta\sigma_2$ respectively where $\eta = \Phi^{-1}(p)$; $0 < p < 1$. Here $\Phi(\cdot)$ denotes the cumulative distribution function of a standard normal random variable. The problem is to estimate the quantile vector $\underline{\theta} = (\theta_1, \theta_2)$ with respect to the sum of the quadratic losses given by,

$$(1.1) \quad L(\underline{d}, \underline{\theta}) = \sum_{i=1}^2 \left(\frac{d_i - \theta_i}{\sigma_i} \right)^2,$$

where $\underline{d} = (d_1, d_2)$ is an estimator of $\underline{\theta} = (\theta_1, \theta_2)$.

The problem of estimation of quantiles has attracted several researchers in the recent past due to its real life applications. For example, quantiles of exponential populations are widely used in the study of reliability, life testing, survival analysis and some related areas. We refer to Keating and Tripathi [7] and Saleh [16] for some applications of exponential quantiles.

We note that, in the literature most of the results on quantile estimation are for a single parameter, $\theta = \mu + \eta\sigma$, whereas the current work is for simultaneous estimation of a vector $\underline{\theta} = (\theta_1, \theta_2)$ of two quantiles. Probably, Zidek [21] was the first to consider the estimation of quantile of normal population with respect to a quadratic loss function. Zidek [21, 22] proved that the best affine equivariant estimator of the quantile $\theta = \mu + \eta\sigma$ is inadmissible if $|\eta|$ is chosen very large. Rukhin [14] derived a class of minimax estimators for quantile θ , each of which improves upon the best equivariant estimator. For some decision theoretic results on estimation of quantiles of an exponential population one may refer to Rukhin [15] and the references therein.

Some study also has been done in estimating the quantile θ_1 , when two or more populations are available from normal populations. Kumar and Tripathy [9] considered the estimation of $\theta_1 = \mu + \eta\sigma_1$ under a quadratic loss function using a decision theoretic approach. Exploiting the information available for the common mean, they could obtain improved estimators for quantiles θ_1 . They also derived some inadmissibility conditions for estimators belonging to equivariant classes. A similar type of results have been obtained by Sharma and Kumar [17] in the case of exponential populations while estimating the quantile θ_1 of the first population.

The problem under consideration has its importance in the sense that it uses the information available for estimating a common mean. The problem of estimating the common mean of normal populations is an age old problem and has its origin in the study of recovery of inter-block information in balance incomplete block designs. In the literature, this problem is also referred as Meta-Analysis, where samples (data) from multiple sources are combined with a common objective. One may refer to Vazquez et al. [20] for application of Meta-Analysis in clinical trials. For a detailed review on inference on common mean of two or more normal populations one may refer to Moore and Krishnamoorthy [11], Lin and Lee [10], Chang and Pal [5], Tripathy and Kumar [18] and the references therein.

It should be noted that, the underlying model has been considered previously by Kumar and Tripathy [9], and estimated the first component θ_1 . We in this paper, consider the simultaneous estimation of quantiles, that is, the vector $\underline{\theta} = (\theta_1, \theta_2)$, which is important from theoretical as well as application point of view. For some results on simultaneous estimation of location and scale parameters with application we refer to Bai and Durairajan [2], Alexander and Chandrasekar [1] and Tsukuma [19]. The rest

of our work is organized as follows. In Section 2, we derive a basic result which helps in constructing improved estimators for quantile vector $\underline{\theta}$. In Section 3, we derive affine and location equivariant estimators. Sufficient conditions for improving estimators in the class have been obtained. In the process, two complete class results proved. An extensive simulation study has been done in order to numerically compare the relative risk performances of various proposed estimators in Section 4. We conclude with some practical examples in Section 5.

2. A General Result and Some Improved Estimators

In this section we discuss the model and prove a general result which will be handy in constructing some good estimators for the quantile vector $\underline{\theta} = (\theta_1, \theta_2)$.

Suppose $\underline{X} = (X_1, X_2, \dots, X_m)$ and $\underline{Y} = (Y_1, Y_2, \dots, Y_n)$ be independent random samples taken from two normal populations $N(\mu, \sigma_1^2)$ and $N(\mu, \sigma_2^2)$ respectively. Here the parameters μ, σ_1^2 and σ_2^2 are unknown. Our aim is to estimate the vector $\underline{\theta} = (\theta_1, \theta_2)$, where $\theta_i = \mu + \eta\sigma_i$, ($\eta \neq 0$ and $i = 1, 2$) with respect to the loss function (1.1). Obviously, θ_i is the p^{th} quantile of the i^{th} population that is, $\eta = \Phi^{-1}(p)$, ($0 < p < 1$) where $\Phi(\cdot)$ is the cumulative distributive function of a standard normal random variable. A minimal sufficient statistic for this problem is $(\bar{X}, \bar{Y}, S_1^2, S_2^2)$ where

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \quad \bar{Y} = \frac{1}{n} \sum_{j=1}^n Y_j, \quad S_1^2 = \sum_{i=1}^m (X_i - \bar{X})^2 \quad \text{and} \quad S_2^2 = \sum_{j=1}^n (Y_j - \bar{Y})^2.$$

It is well known that the maximum likelihood estimator (MLE) for μ , is not obtainable in a closed form (see Pal et al. [12]). Also the minimal sufficient statistics for this problem are not complete, hence the usual approaches to find uniformly minimum variance unbiased estimator (UMVUE) for individual quantile do not work as ancillary statistics may carry relevant information for the parameter of interest. Therefore, it is not known if a UMVUE exists or not, and it is difficult to find even if one exists. Further, it is known that when we have only one population (say \underline{X}) the best affine equivariant estimator for estimating quantile $\theta_1 = \mu + \eta\sigma_1$ is minimax (see Kiefer [8]). When we have both the populations \underline{X} and \underline{Y} the problem of estimating the first component θ_1 has been considered by Kumar and Tripathy [9]. Following their arguments, a natural way to construct improved estimators for $\underline{\theta}$ is to combine the improved estimators for the common mean and the improved estimators for the respective standard deviations. Hence we first propose a basic estimator for $\underline{\theta}$ as,

$$\underline{d} = (d_1, d_2), \quad \text{where } d_i = \bar{X} + cS_i, \quad i = 1, 2.$$

Let us define

$$(2.1) \quad c_{m+n} = \frac{\eta\sqrt{2}}{m+n-2} \left[\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m-1}{2})} + \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \right].$$

2.1. Theorem. *If we estimate the quantiles $\underline{\theta}$ by $\underline{d} = (\bar{X} + cS_1, \bar{X} + cS_2)$ with respect to the loss function (1.1), then the value of c for which the risk is minimum is found to be c_{m+n} .*

Let us denote $\underline{d}^X = (\bar{X} + c_{m+n}S_1, \bar{X} + c_{m+n}S_2)$. Next, we give a general result which in parallel to Theorem 2.1 of Kumar and Tripathy [9] that valid for estimating only θ_1 .

2.2. Theorem. *Suppose $\underline{d}_M = (d_M, d_M)$ be an estimator for $\underline{\mu} = (\mu, \mu)$, and $\underline{d}_S = (d_{S_1}, d_{S_2})$ be an estimator for $\underline{\sigma} = (\sigma_1, \sigma_2)$. Consider $\underline{d}_Q = (d_{Q1}, d_{Q2}) = \underline{d}_M + \eta\underline{d}_S$ as an*

estimator for $\underline{\theta}$. Further, assume that given d_{S_1} , and d_{S_2} , d_M is conditionally unbiased for μ , that is

$$(2.2) \quad E(d_M | d_{S_1}) = E(d_M | d_{S_2}) = \mu,$$

then,

$$(2.3) \quad E(d_{Q_1} - \theta_1)^2 + E(d_{Q_2} - \theta_2)^2 = 2E(d_M - \mu)^2 + \eta^2 \{E(d_{S_1} - \sigma_1)^2 + E(d_{S_2} - \sigma_2)^2\}.$$

Proof. The proof is similar to the arguments used in proving Theorem 2.1 of Kumar and Tripathy [9], hence omitted. \square

2.3. Remark. It is easy to observe that, condition (2.3) will satisfy if we choose d_M to be an unbiased estimator for μ and both d_{S_1} and d_{S_2} are independent of d_M . For example we may take $d_M = \bar{X}$ and $d_{S_1} = S_1$, $d_{S_2} = S_2$.

2.4. Remark. As a consequence of Theorem 2.2, to construct a good estimator for $\underline{\theta}$, it is sufficient to have a good estimator for μ and/or a good estimator for σ_1 or/and a good estimator for σ_2 .

2.5. Remark. Let $d_M = d_\phi$, where $d_\phi = \phi(S_1, S_2)\bar{X} + (1 - \phi(S_1, S_2))\bar{Y}$ be any unbiased estimator for μ , and $d_{S_1} = cS_1/\eta$, $d_{S_2} = cS_2/\eta$ ($\eta \neq 0$), it is easy to see that, the condition of Theorem 2.2 satisfies and we prove the following result.

2.6. Theorem. Let $d_\phi = \phi(S_1, S_2)\bar{X} + (1 - \phi(S_1, S_2))\bar{Y}$ be an estimator for the common mean μ . Consider the estimator $\underline{d}_\phi(c) = (d_\phi + cS_1, d_\phi + cS_2)$ for estimating quantile vector $\underline{\theta}$. Then $\underline{d}_\phi(c)$ has smaller risk than \underline{d} with respect to the sum of quadratic loss (1.1) if and only if d_ϕ has smaller risk than \bar{X} . Further, $\underline{d}_\phi(c)$ has minimum risk with respect to the loss (1.1) when $c = c_{m+n}$.

We note that, the minimizing choice of c is c_{m+n} which is symmetric in both m and n . One may construct an estimator for the quantile $\underline{\theta}$ using \bar{Y} for the common mean. Let us denote $\underline{d}^* = (\bar{Y} + cS_1, \bar{Y} + cS_2)$. The results of Theorem 2.6 will remain true if we replace \underline{d} by \underline{d}^* . Hence we have the following remark.

2.7. Remark. Let $d_\phi = \phi(S_1, S_2)\bar{X} + (1 - \phi(S_1, S_2))\bar{Y}$ be an estimator for the common mean μ . Consider the estimator $\underline{d}_\phi(c) = (d_\phi + cS_1, d_\phi + cS_2)$ for estimating quantile vector $\underline{\theta}$. Then $\underline{d}_\phi(c)$ has smaller risk than \underline{d}^* with respect to the sum of quadratic loss (1.1) if and only if d_ϕ has smaller risk than \bar{Y} . Further, $\underline{d}_\phi(c)$ has minimum risk with respect to the loss (1.1) when $c = c_{m+n}$. Let us denote $\underline{d}^Y = (\bar{Y} + c_{m+n}S_1, \bar{Y} + c_{m+n}S_2)$.

2.8. Remark. Following Theorem 2.6, one can easily construct good estimators for $\underline{\theta}$ by replacing \bar{X} in \underline{d}^X or \bar{Y} in \underline{d}^Y by any improved estimator of the form d_ϕ for the common mean μ .

Following the above remarks and Theorem 2.2, we propose the following estimators for $\underline{\theta}$ which have smaller risk than \underline{d}^X or/and \underline{d}^Y under certain conditions on the sample

sizes.

$$\begin{aligned}
\tilde{d}^{GM} &= (\hat{\mu}_{GM} + c_{m+n}S_1, \hat{\mu}_{GM} + c_{m+n}S_2), \\
\tilde{d}^{GD} &= (\hat{\mu}_{GD} + c_{m+n}S_1, \hat{\mu}_{GD} + c_{m+n}S_2), \\
\tilde{d}^{KS} &= (\hat{\mu}_{KS} + c_{m+n}S_1, \hat{\mu}_{KS} + c_{m+n}S_2), \\
\tilde{d}^{CS} &= (\hat{\mu}_{CS} + c_{m+n}S_1, \hat{\mu}_{CS} + c_{m+n}S_2), \\
\tilde{d}^{MK} &= (\hat{\mu}_{MK} + c_{m+n}S_1, \hat{\mu}_{MK} + c_{m+n}S_2), \\
\tilde{d}^{TK} &= (\hat{\mu}_{TK} + c_{m+n}S_1, \hat{\mu}_{TK} + c_{m+n}S_2), \\
\tilde{d}^{BC1} &= (\hat{\mu}_{BC1} + c_{m+n}S_1, \hat{\mu}_{BC1} + c_{m+n}S_2), \\
\tilde{d}^{BC2} &= (\hat{\mu}_{BC2} + c_{m+n}S_1, \hat{\mu}_{BC2} + c_{m+n}S_2).
\end{aligned}$$

Here we denote $\hat{\mu}_{GM} = \frac{m\bar{X} + n\bar{Y}}{m+n}$, $\hat{\mu}_{TK} = \frac{\sqrt{m} b_{n-1}S_2\bar{X} + \sqrt{n} b_{m-1}S_1\bar{Y}}{\sqrt{m} b_{n-1}S_2 + \sqrt{n} b_{m-1}S_1}$, and $\hat{\mu}_{GD}, \hat{\mu}_{KS}, \hat{\mu}_{BC1}, \hat{\mu}_{BC2}, \hat{\mu}_{CS}, \hat{\mu}_{MK}$, are estimators for the common mean μ , as defined in Tripathy and Kumar [18]. Although the closed form of the MLE of μ is not available, one can obtain it numerically by solving a system of three equations in three unknowns. Let us denote $\hat{\mu}_{ML}$ as the MLE of the common mean. Using this estimator for the common mean we propose an estimator for the quantile vector $\underline{\theta}$ as,

$$\tilde{d}^{ML} = (\hat{\mu}_{ML} + c_{m+n}S_1, \hat{\mu}_{ML} + c_{m+n}S_2).$$

All these estimators belong to the class $\underline{d}^{\phi}(c_{m+n})$ and will be compared numerically in Section 4.

2.9. Theorem. *Let the estimators $\underline{d}^X, \underline{d}^Y, \underline{d}^{GD}, \underline{d}^{KS}, \underline{d}^{BC1}, \underline{d}^{BC2}$, and \underline{d}^{CS} as defined above for estimating $\underline{\theta}$. The loss function be taken as the sum of the quadratic losses (1.1).*

- (i) *The estimator \underline{d}^{GD} performs better than both \underline{d}^X and \underline{d}^Y if and only if $m, n \geq 11$.*
- (ii) *The estimator \underline{d}^{KS} performs better than both \underline{d}^X and \underline{d}^Y if and only if $(m - 7)(n - 7) \geq 16$.*
- (iii) *The estimator \underline{d}^{BC1} performs better than \underline{d}^X if and only if $m \geq 2, n \geq 3$ and for $0 < b_1 < b_{\max}(m, n)$.*
- (iv) *The estimator \underline{d}^{BC2} performs better than \underline{d}^X if and only if $m \geq 2, n \geq 6$ and for $0 < b_2 < b_{\max}(m, n - 3)$.*
- (v) *The estimator \underline{d}^{CS} performs better than \underline{d}^X if $m = n \geq 7$.*

Here b_1, b_2 and $b_{\max}(m, n)$ are as defined in Kumar and Tripathy [9].

Proof. The proof of (i)-(v) can be done by using Theorem 2.6 and the arguments given in the proof of Theorem 2.4 in Kumar and Tripathy [9]. \square

2.10. Remark. The estimator \underline{d}^{MK} uses the estimator proposed by Moore and Krishnamoorthy [11] that uses the estimates of standard deviation instead of variance. Their estimator does not improve upon \bar{X} uniformly. The estimator \underline{d}^{TK} proposed by Tripathy and Kumar [18], also does not improve upon \bar{X} uniformly. As our numerical results shows (in Section 4), these two estimators perform quite well for moderate values of $\sigma_2/\sigma_1 > 0$ and also they are good competitor of each other.

3. Inadmissibility Results for Equivariant Estimators

In this section, we introduce the concept of invariance to the problem of simultaneous estimation of quantiles of two normal populations and derive classes of affine and location

equivariant estimators. Further sufficient conditions for improving estimators in these classes have been derived. Consequently some complete class results are also proved.

Consider the group $G_A = \{g_{a,b} : g_{a,b}(x) = ax + b, a > 0, b \in R\}$ of affine transformations. Under the transformation, $\bar{X} \rightarrow a\bar{X} + b$, $\bar{Y} \rightarrow a\bar{Y} + b$, $S_i^2 \rightarrow a^2 S_i^2$, $\mu \rightarrow a\mu + b$, $\sigma_i^2 \rightarrow a^2 \sigma_i^2$, and $\underline{\theta} \rightarrow a\underline{\theta} + b\underline{e}$, where $\underline{e} = (1, 1)$ and $\underline{\theta} = (\theta_1, \theta_2)$, $\theta_i = \mu + \eta\sigma_i$, $i = 1, 2$. The problem considered is invariant if we choose the loss function as the sum of affine invariant loss functions (1.1). Based on the sufficient statistics $(\bar{X}, \bar{Y}, S_1^2, S_2^2)$ the form of an affine equivariant estimator for estimating the vector $\underline{\theta}$ is obtained as,

$$\begin{aligned} (d_1(\bar{X}, \bar{Y}, S_1^2, S_2^2), d_2(\bar{X}, \bar{Y}, S_1^2, S_2^2)) &= (\bar{X} + S_1\Psi_1(T_1, T_2), \bar{X} + S_1\Psi_2(T_1, T_2)) \\ &= (d_{\Psi_1}, d_{\Psi_2}) \\ (3.1) \qquad \qquad \qquad &= \underline{d}_{\Psi} \text{ say,} \end{aligned}$$

where $T_1 = \frac{\bar{Y} - \bar{X}}{S_1}$ and $T_2 = \frac{S_2^2}{S_1^2}$.

Denote $M_1 = \min(t_1, 0)$, and $M_2 = \max(t_1, 0)$. Let us define the following functions for any affine equivariant estimator \underline{d}_{Ψ} .

$$(3.2) \qquad \Psi^0 = (\min(\max(\Psi_1, M_1), M_2), \min(\max(\Psi_2, M_1), M_2))$$

$$(3.3) \qquad \Psi^1 = (\max\{M_1 + \eta b_{m+n}, \Psi_1\}, \max\{M_1 + \eta b_{m+n}\sqrt{t_2}, \Psi_2\}),$$

$$(3.4) \qquad \Psi^2 = (\min\{M_2 + \eta b_{m+n}, \Psi_1\}, \min\{M_2 + \eta b_{m+n}\sqrt{t_2}, \Psi_2\}).$$

Next we prove the following inadmissibility result for affine equivariant estimators.

3.1. Theorem. *Let \underline{d}_{Ψ} be an affine equivariant estimator of the form (3.1) of a quantile vector $\underline{\theta}$, and the loss function be the sum of quadratic loss (1.1) or the sum of squared errors. Let the functions Ψ^0 , Ψ^1 and Ψ^2 be defined as in (3.2), (3.3) and (3.4) respectively. Let $\underline{\alpha} = (\mu, \sigma_1^2, \sigma_2^2)$.*

- (i) *When $\eta = 0$, the estimator \underline{d}_{Ψ} is improved by \underline{d}_{Ψ^0} if $P_{\underline{\alpha}}(\Psi^0 \neq \Psi) > 0$ for some choices of $\underline{\alpha}$.*
- (ii) *When $\eta > 0$, the estimator \underline{d}_{Ψ} is improved by \underline{d}_{Ψ^1} if $P_{\underline{\alpha}}(\Psi^1 \neq \Psi) > 0$ for some choices of $\underline{\alpha}$.*
- (iii) *When $\eta < 0$, the estimator \underline{d}_{Ψ} is improved by \underline{d}_{Ψ^2} if $(\Psi^2 \neq \Psi) > 0$ for some choices of $\underline{\alpha}$.*

Proof. To prove this theorem we use a result due to Brewster and Zidek [3]. Consider the conditional risk function of \underline{d}_{Ψ} given $\underline{T} = (T_1, T_2)$:

$$\begin{aligned} R((\underline{d}_{\Psi}, \underline{\theta})|\underline{T}) &= E\{L(\underline{d}_{\Psi}, \underline{\theta})|\underline{T}\} \\ &= \frac{1}{\sigma_1^2} E\{(\bar{X} + S_1\Psi_1(\underline{T}) - \mu - \eta\sigma_1)^2|\underline{T} = \underline{t}\} \\ (3.5) \qquad \qquad \qquad &+ \frac{1}{\sigma_2^2} E\{(\bar{X} + S_1\Psi_2(\underline{T}) - \mu - \eta\sigma_2)^2|\underline{T} = \underline{t}\}. \end{aligned}$$

The above risk function (3.5) is a sum of two convex functions in Ψ_1 and Ψ_2 , which is a convex function. The minimizing choices of $\Psi_1(\underline{t})$ and $\Psi_2(\underline{t})$, are obtained respectively as,

$$\Psi_1(\underline{t}) = -\frac{E\{(\bar{X} - \mu)S_1|\underline{T}\}}{E(S_1^2|\underline{T})} + \eta\sigma_1 \frac{E(S_1|\underline{T})}{E(S_1^2|\underline{T})}$$

and

$$\Psi_2(\underline{t}) = -\frac{E\{(\bar{X} - \mu)S_1|\underline{T}\}}{E(S_1^2|\underline{T})} + \eta\sigma_2 \frac{E(S_1|\underline{T})}{E(S_1^2|\underline{T})}.$$

Using the conditional expectations derived in Kumar and Tripathy [9], the minimizing choices for $\Psi_1(\underline{t})$ and $\Psi_2(\underline{t})$ are simplified and are given by

$$(3.6) \quad \Psi_1(\underline{t}, \rho) = \frac{t_1}{1 + \rho} + \eta b_{m+n} \sqrt{\lambda}$$

and

$$(3.7) \quad \Psi_2(\underline{t}, \rho) = \frac{t_1}{1 + \rho} + \eta b_{m+n} \sqrt{\frac{n\rho}{m}} \sqrt{\lambda}.$$

Here $\lambda = \frac{mt_1^2}{1+\rho} + \frac{mt_2}{n\rho} + 1$, $b_{m+n} = \frac{\Gamma(\frac{m+n}{2})}{\sqrt{2}\Gamma(\frac{m+n+1}{2})}$ and $\rho = \frac{m\sigma_2^2}{n\sigma_1^2}$.

In order to prove the theorem, we need to find the infimum and supremum values of $\Psi_1(\underline{t}, \rho)$ and $\Psi_2(\underline{t}, \rho)$ with respect to $\rho > 0$, for all values of η and \underline{t} . After analyzing the terms $\Psi_1(\underline{t}, \rho)$ and $\Psi_2(\underline{t}, \rho)$, for separate values of η , we have the following cases:

(i) When $\eta = 0$, and $t_1 \in \mathbb{R}$,

$$(3.8) \quad \begin{aligned} \inf_{\rho} \Psi_1(\underline{t}, \rho) &= M_1 \quad \text{and} \quad \sup_{\rho} \Psi_1(\underline{t}, \rho) = M_2 \\ \inf_{\rho} \Psi_2(\underline{t}, \rho) &= M_1 \quad \text{and} \quad \sup_{\rho} \Psi_2(\underline{t}, \rho) = M_2. \end{aligned}$$

(ii) When $\eta > 0$, and $t_1 \in \mathbb{R}$, we have

$$(3.9) \quad \begin{aligned} \inf_{\rho} \Psi_1(\underline{t}, \rho) &\geq M_1 + \eta b_{m+n} \quad (\text{equality holds if } t_1 > 0) \\ &\quad \text{and} \quad \sup_{\rho} \Psi_1(\underline{t}, \rho) = +\infty \\ \inf_{\rho} \Psi_2(\underline{t}, \rho) &\geq M_1 + \eta b_{m+n} \sqrt{t_2} \quad (\text{equality holds if } t_1 < 0) \\ &\quad \text{and} \quad \sup_{\rho} \Psi_2(\underline{t}, \rho) = +\infty. \end{aligned}$$

(iii) When $\eta < 0$, $t_1 \in \mathbb{R}$, we have

$$(3.10) \quad \begin{aligned} \sup_{\rho} \Psi_1(\underline{t}, \rho) &\leq M_2 + \eta b_{m+n} \quad (\text{equality holds if } t_1 < 0) \\ &\quad \text{and} \quad \inf_{\rho} \Psi_1(\underline{t}, \rho) = -\infty \\ \sup_{\rho} \Psi_2(\underline{t}, \rho) &\leq M_2 + \eta b_{m+n} \sqrt{t_2} \quad (\text{equality holds if } t_1 > 0) \\ &\quad \text{and} \quad \inf_{\rho} \Psi_2(\underline{t}, \rho) = -\infty. \end{aligned}$$

Utilizing the expressions (3.8)-(3.10), for $\eta = 0$, $\eta > 0$ and $\eta < 0$, respectively, for an affine equivariant estimator $\underline{d}_{\Psi} = (d_{\Psi_1}, d_{\Psi_2})$, we can easily define the functions $\underline{\Psi}^0$, $\underline{\Psi}^1$, $\underline{\Psi}^2$ as in (3.2)-(3.4) respectively. An application of orbit-by-orbit improvement technique for improving equivariant estimators of Brewster and Zidek [3], proves the theorem. \square

3.2. Remark. The above theorem is basically a complete class result. It tells that for an equivariant estimator of the form (3.1),

(i) if $P_{\underline{\alpha}}(\{\Psi_1 \in [\min(T_1, 0), \max(T_1, 0)]^c\} \cup \{\Psi_2 \in [\min(T_1, 0), \max(T_1, 0)]^c\}) > 0$, then the estimator \underline{d}_{Ψ} is improved by \underline{d}_{Ψ^0} , when $\eta = 0$.

(ii) if $P(\{\Psi_1 < \min(T_1, 0) + \eta b_{m+n}\} \cup \{\Psi_2 < \min(T_1, 0) + \eta b_{m+n} \sqrt{T_2}\}) > 0$, then the estimator \underline{d}_{Ψ^1} will improve upon \underline{d}_{Ψ} , when $\eta > 0$,

- (iii) if $P(\{\Psi_1 > \max(T_1, 0) + \eta b_{m+n}\} \cup \{\Psi_2 > \max(T_1, 0) + \eta b_{m+n} \sqrt{T_2}\}) > 0$, then the estimator \underline{d}_{Ψ_2} will improve upon \underline{d}_{Ψ} when $\eta < 0$.

Here $[a, b]^c$ stands for complement of the interval $[a, b]$ in \mathbb{R} .

3.3. Remark. All the estimators discussed in Section 2 (except \underline{d}^{ML} whose closed form does not exist), belong to the class (3.1). But it has been seen that for none of these estimators, the choices of Ψ_1 and Ψ_2 satisfy the above conditions in Remark 3.2. So the estimators considered can not be improved by using Theorem 3.1, but they form a complete class. The result we write as a theorem below.

3.4. Theorem. *Let the loss function be (1.1).*

- (i) *The class of estimators $\{\underline{d}_{\Psi} : \Psi_1 \in [\min(T_1, 0), \max(T_1, 0)]$ and $\Psi_2 \in [\min(T_1, 0), \max(T_1, 0)]\}$ is complete for $\eta = 0$.*
- (ii) *The class of estimators $\{\underline{d}_{\Psi} : \Psi_1 > \min(T_1, 0) + \eta b_{m+n}$ and $\Psi_2 > \min(T_1, 0) + \eta b_{m+n} \sqrt{T_2}\}$ is complete for $\eta > 0$.*
- (ii) *The class of estimators $\{\underline{d}_{\Psi} : \Psi_1 < \max(T_1, 0) + \eta b_{m+n}$ and $\Psi_2 < \max(T_1, 0) + \eta b_{m+n} \sqrt{T_2}\}$ is complete for $\eta < 0$.*

Next, we consider a smaller group of transformations and hence a larger class of estimators for estimating the vector $\underline{\theta}$. Consider the group $G_L = \{g_c : g_c(x) = c+x, c \in \mathbb{R}\}$ of location transformations. Under the transformation, $\bar{X} \rightarrow \bar{X} + c$, $\bar{Y} \rightarrow \bar{Y} + c$, $S_i^2 \rightarrow S_i^2$, $\mu \rightarrow \mu + c$, $\sigma_i \rightarrow \sigma_i$, $\theta_i = \mu + \eta\sigma_i \rightarrow \theta_i + c$ where $i = 1, 2$.

The estimation problem is invariant if we take the loss function as the sum of squared error losses (1.1), and the form of a location equivariant estimator for estimating the vector $\underline{\theta}$ based on the sufficient statistics $(\bar{X}, \bar{Y}, S_1^2, S_2^2)$, is obtained as

$$(3.11) \quad \underline{d}_{\psi} = (\bar{X} + \psi_1(U), \bar{X} + \psi_2(U)),$$

where $U = (T, S_1^2, S_2^2)$ and $T = \bar{Y} - \bar{X}$.

Let us denote $N_1 = \min(t, 0)$ and $N_2 = \max(t, 0)$. For a location equivariant estimator \underline{d}_{ψ} , define the functions $\underline{\psi}^0$, $\underline{\psi}^1$ and $\underline{\psi}^2$ as,

$$(3.12) \quad \underline{\psi}^0(u) = (\min(\max(\psi_1, N_1), N_2), \min(\max(\psi_2, N_1), N_2))$$

$$(3.13) \quad \underline{\psi}^1(u) = (\max\{N_1, \psi_1\}, \max\{N_1, \psi_2\}),$$

$$(3.14) \quad \underline{\psi}^2(u) = (\min\{N_2, \psi_1\}, \min\{N_2, \psi_2\}).$$

Next, we prove a theorem regarding inadmissibility of location equivariant estimators.

3.5. Theorem. *Let \underline{d}_{ψ} be a location equivariant estimator of the quantile $\underline{\theta}$ and the loss function be the sum of quadratic losses (1.1) or the sum of squared error. Let the functions $\underline{\psi}^0$, $\underline{\psi}^1$ and $\underline{\psi}^2$ be defined as in (3.12), (3.13) and (3.14) respectively.*

- (i) *When $\eta = 0$, the estimator \underline{d}_{ψ} is improved by \underline{d}_{ψ^0} if $P_{\alpha}(\underline{\psi}^0 \neq \underline{\psi}) > 0$ for some choices of α .*
- (ii) *When $\eta > 0$, the estimator \underline{d}_{ψ} is improved by \underline{d}_{ψ^1} if $P_{\alpha}(\underline{\psi}^1 \neq \underline{\psi}) > 0$ for some choices of α .*
- (iii) *When $\eta < 0$, the estimator \underline{d}_{ψ} is improved by \underline{d}_{ψ^2} if $P_{\alpha}(\underline{\psi}^2 \neq \underline{\psi}) > 0$ for some choices of α .*

Proof. The proof is similar to the arguments used in proving Theorem 3.1. The details of the proof is omitted. \square

3.6. Remark. Similar to Theorem 3.1 above Theorem 3.5 is also a complete class result. It tells that for an estimator of the form (3.11),

- (i) if $P(\{\psi_1 \in [\min(T, 0), \max(T, 0)]^c\} \cup \{\psi_2 \in [\min(T, 0), \max(T, 0)]^c\}) > 0$ then the estimator $\underline{d}_{\underline{\psi}}$ is improved by $\underline{d}_{\underline{\psi}^0}$, when $\eta = 0$,
- (ii) if $P(\{\psi_1 < \min(T, 0)\} \cup \{\psi_2 < \min(T, 0)\}) > 0$, then the estimator $\underline{d}_{\underline{\psi}^1}$ will improve upon $\underline{d}_{\underline{\psi}}$, for $\eta > 0$, and
- (iii) if $P(\{\psi_1 > \max(T, 0)\} \cup \{\psi_2 > \max(T, 0)\}) > 0$, then the estimator $\underline{d}_{\underline{\psi}^2}$ will improve upon $\underline{d}_{\underline{\psi}}$ when $\eta < 0$.

3.7. Remark. All the estimators discussed in Section 2 (except \underline{d}^{ML} whose closed form does not exist), belong to the class (3.11). But it has also been seen that for none of these estimators the choices of ψ_1 and ψ_2 satisfy the above conditions in Remark 3.6. So the estimators considered can not be improved by using Theorem 3.5, but they form a complete class. This we write as a theorem.

3.8. Theorem. *Let the loss function be (1.1).*

- (i) *The class of estimators $\{\underline{d}_{\underline{\psi}} : \psi_1 \in [\min(T, 0), \max(T, 0)]$ and $\psi_2 \in [\min(T, 0), \max(T, 0)]\}$ is complete for $\eta = 0$.*
- (ii) *The class of estimators $\{\underline{d}_{\underline{\psi}} : \psi_1 > \min(T, 0)$ and $\psi_2 > \min(T, 0)\}$ is complete for $\eta > 0$.*
- (ii) *The class of estimators $\{\underline{d}_{\underline{\psi}} : \psi_1 < \max(T, 0)$ and $\psi_2 < \max(T, 0)\}$ is complete for $\eta < 0$.*

4. Numerical Comparisons

In the previous sections we have derived several estimators for the quantile vector $\underline{\theta}$ such as \underline{d}^X , \underline{d}^Y , \underline{d}^{GD} , \underline{d}^{GM} , \underline{d}^{KS} , \underline{d}^{BC1} , \underline{d}^{BC2} , \underline{d}^{CS} , \underline{d}^{MK} , \underline{d}^{TK} , and \underline{d}^{ML} . We have also shown that these well structured estimators, except \underline{d}^{ML} , belong to the class (3.1) and (3.11). It seems quite difficult to compare the risk values of all these estimators analytically. But for practical purposes, one needs the estimator to be used. Taking the advantages of computational resources, we in this section compare numerically the simulated risk values of all these estimators which may be handy for practical purposes. For evaluating the risk function, we use the loss function (1.1). For numerical comparison purpose, we have generated 20,000 random samples \underline{X} of sizes m and 20,000 random samples \underline{Y} of sizes n from normal populations with equal mean and different variances. It can be easily checked that all the risks values are functions of $\tau = \frac{\sigma_2^2}{\sigma_1^2} > 0$, for fixed values of m , n and $|\eta|$. The approximate value of π is taken to be 3.1416. We have computed the risk values of all the estimators taking various choices of τ and the sample sizes. However, for illustration purpose we present the risk values for some selected choices of τ and m, n . We also observe that when the values of τ increase from 0 to ∞ the risk values converge for all the estimators except \underline{d}^{GM} and \underline{d}^Y . As the sample sizes increases the risk values of all the estimators decrease for fixed $|\eta|$. Further, the risk values increase as η increases for fixed values of τ and sample sizes. If we choose the value of b_1 and b_2 near 0 the estimators \underline{d}^{BC1} and \underline{d}^{BC2} tends to \underline{d}^X . Also if we choose the value of b_2 near 1 the estimator \underline{d}^{BC2} tends to \underline{d}^{GD} . So for numerical comparison a convenient choice would be an intermediate value which we take as $\frac{1}{2}b_{max}$. The value of $b_{max}(m, n)$ have been taken from the tabulated values given in Brown and Cohen [4]. We also note that, when the sample sizes are equal the estimator \underline{d}^{GD} becomes same as \underline{d}^{KS} and \underline{d}^{MK} becomes same as \underline{d}^{TK} . When the sample sizes are unequal the estimator \underline{d}^{CS} is not defined, so for

unequal sample sizes we do not include it for numerical comparison purpose. A massive simulation study has been conducted separately for the cases $m = n$, $m > n$ and $m < n$. The simulated risk values have been plotted against τ for all the estimators in Figure 1 and Figure 2. In Figure 1 the sample sizes have been taken as equal, whereas in Figure 2, the simulated risk values have been plotted for unequal sample sizes. In Figures 1, and 2 we label X , Y , GM , GD , KS , $BC1$, $BC2$, CS , MK , TK and ML for the estimators \underline{d}^X , \underline{d}^Y , \underline{d}^{GM} , \underline{d}^{GD} , \underline{d}^{KS} , \underline{d}^{BC1} , \underline{d}^{BC2} , \underline{d}^{CS} , \underline{d}^{MK} , \underline{d}^{TK} \underline{d}^{ML} respectively. In Tables 1-3, we have presented the simulated values of the percentage of relative risk improvement of all the estimators with respect to \underline{d}^X , which are defined as

$$\begin{aligned} PR1 &= \left(1 - \frac{Risk(\underline{d}^Y)}{Risk(\underline{d}^X)}\right) \times 100, \quad PR2 = \left(1 - \frac{Risk(\underline{d}^{GM})}{Risk(\underline{d}^X)}\right) \times 100, \\ PR3 &= \left(1 - \frac{Risk(\underline{d}^{GD})}{Risk(\underline{d}^X)}\right) \times 100, \quad PR4 = \left(1 - \frac{Risk(\underline{d}^{KS})}{Risk(\underline{d}^X)}\right) \times 100, \\ PR5 &= \left(1 - \frac{Risk(\underline{d}^{BC1})}{Risk(\underline{d}^X)}\right) \times 100, \quad PR6 = \left(1 - \frac{Risk(\underline{d}^{BC2})}{Risk(\underline{d}^X)}\right) \times 100, \\ PR7 &= \left(1 - \frac{Risk(\underline{d}^{CS})}{Risk(\underline{d}^X)}\right) \times 100, \quad PR8 = \left(1 - \frac{Risk(\underline{d}^{MK})}{Risk(\underline{d}^X)}\right) \times 100, \\ PR9 &= \left(1 - \frac{Risk(\underline{d}^{TK})}{Risk(\underline{d}^X)}\right) \times 100, \quad PR10 = \left(1 - \frac{Risk(\underline{d}^{ML})}{Risk(\underline{d}^X)}\right) \times 100. \end{aligned}$$

The following observations can be made from the Tables 1-3 and the Figures 1-2 as well as from our simulation study. For illustration purpose, we have presented the risk functions only for the case $\eta = 1.960$.

Case 1: $m = n$.

- (i) Figure 1 represents the risk values of all the estimators for the equal sample sizes and $\eta = 1.960$. In Figure 1, (a)-(c) it represents the risk values for sample sizes small to moderate that is (6,6), (8,8) and (12,12) whereas (d)-(f) the sample sizes are taken as moderate to large (20,20), (30,30) and (40,40). It has been noticed that the risk values of the estimators \underline{d}^X , \underline{d}^{BC1} , \underline{d}^{BC2} and \underline{d}^{CS} decreasing as τ increases from 0 to ∞ . The estimator \underline{d}^{GD} first increases and attains maximum value then decreases. The estimators \underline{d}^{GM} , and \underline{d}^{MK} first decrease attains minimum (in the neighborhood of $\tau = 1$) then increases. The estimator \underline{d}^Y increases as τ varies from 0 to ∞ . It has also been noticed that all the estimators (except \underline{d}^{GM} and \underline{d}^Y) converge to the estimator \underline{d}^X which is true as these estimators are consistent.
- (ii) The percentage of relative risk performances of all the estimators with respect to \underline{d}^X decrease as τ varies from 0 to ∞ . Let us first consider the case of small sample sizes ($m, n \leq 10$). For small values of τ ($\tau < 0.25$) the estimators \underline{d}^Y and \underline{d}^{ML} has the maximum percentage of relative risk improvement and it is seen near to 98.88%. For moderate values of τ ($0.75 < \tau < 2.5$) the estimators \underline{d}^{GM} and \underline{d}^{MK} compete each other however when $\tau = 1$, the estimator \underline{d}^{GM} has the maximum percentage of relative risk improvement and it is seen near to 15.68%. For large values of τ , the estimator \underline{d}^{BC1} has the maximum percentage of relative risk improvement.

Consider the case of moderate sample sizes ($12 \leq m, n \leq 20$). For small values of τ , the estimator \underline{d}^{ML} has the best performance and the percentage of relative risk improvement is seen near to 89.78%. For moderate values of τ

($0.75 < \tau < 2.5$) the estimators \underline{d}^{MK} and \underline{d}^{GD} perform equally well, however for $\tau = 1$, the estimator \underline{d}^{GM} has the maximum percentage of relative risk performances. For large values of τ , ($\tau > 3.5$) the estimators \underline{d}^{BC1} and \underline{d}^{ML} compete with each other.

Consider the case of large sample sizes ($m, n \geq 30$). For small values of τ the estimators \underline{d}^{ML} and \underline{d}^{GD} compete with each other and the percentage of relative risk performance has been noticed near to 90.40%. For moderate values of τ ($0.75 < \tau < 2.5$), the estimators \underline{d}^{GD} , \underline{d}^{ML} and \underline{d}^{MK} compete with each other, however for $\tau = 1$, the estimator \underline{d}^{GM} has the best performance. For large values of τ , the estimators \underline{d}^{BC1} and \underline{d}^{BC2} compete with \underline{d}^{ML} .

Case 2: $m < n$.

- (i) Figure 2, ((a), (c) and (e)) represents the risk values of all the estimators for $\eta = 1.960$ and the sample sizes (4,10), (12,20) and (30,40). The risk values of the estimators \underline{d}^X , is decreasing as τ increases. The risk values of \underline{d}^{GD} , \underline{d}^{KS} increase and attains maximum then decrease as τ increases. The risk values of all the estimators converge to the risk of \underline{d}^X except \underline{d}^Y and \underline{d}^{GM} .
- (ii) Consider the small sample sizes ($m, n \leq 10$). For small values of $\tau < 0.25$, the estimator \underline{d}^Y and \underline{d}^{ML} compete with each other and the percentage of relative risk improvement is seen near to 98.88%. For moderate values of τ ($0.75 < \tau < 3$), the estimators \underline{d}^{TK} and \underline{d}^{GM} compete each other, however for $\tau = 1$, the estimator \underline{d}^{GM} has the best performance. For large values of τ ($\tau > 3.0$), the estimator \underline{d}^{BC1} performs the best and the percentage of relative risk performance.

Consider the case of moderate sample sizes ($12 \leq m, n \leq 20$). For small values of τ the estimator \underline{d}^{ML} has the maximum percentage of relative risk performance and it is seen near to 98.88%. For moderate values of τ ($0.75 < \tau < 3$) the estimators \underline{d}^{TK} , \underline{d}^{MK} and \underline{d}^{KS} compete each other, however for $\tau = 1$, \underline{d}^{GM} has the best performance. For large values of τ ($\tau > 3$) the estimator \underline{d}^{BC1} has the maximum percentage of relative risk improvement.

Consider the case of large sample sizes ($m, n \geq 30$). For small values of τ ($\tau \leq 0.25$), the estimators \underline{d}^{KS} , \underline{d}^{GD} and \underline{d}^{ML} compete each other. For moderate values of τ ($0.25 < \tau < 3$), the estimators \underline{d}^{GD} , \underline{d}^{KS} , \underline{d}^{TK} , \underline{d}^{MK} and \underline{d}^{ML} compete each other. For large values of τ the estimators \underline{d}^{ML} and \underline{d}^{BC1} compete each other.

Case-3: $m > n$.

- (i) Figure 2, ((b), (d) and (f)) represent the risk values of all the estimators for $\eta = 1.960$ and for the sample sizes (10,4), (20,12) and (40,30). The risk values of \underline{d}^X is decreasing as τ increases. The risk values of \underline{d}^{GD} , \underline{d}^{KS} , \underline{d}^{BC1} and \underline{d}^{BC2} decrease as τ increases. The risk values of estimators \underline{d}^{GM} , and \underline{d}^Y first decrease attains minimum then increase with respect to τ .
- (ii) Consider the case of small sample sizes ($m, n \leq 10$). For small values of τ ($\tau \leq 0.25$) the estimator \underline{d}^{ML} has maximum percentage of relative risk performance and it is noticed near to 97.7%, for moderate values of τ ($0.75 < \tau < 2.0$) the estimators \underline{d}^{TK} and \underline{d}^{GM} compete each other, however for $\tau = 1$, the estimator \underline{d}^{GM} has the best performance. For large values of τ , ($\tau > 3$) the estimator \underline{d}^{BC1} has the best performance.

Consider the case of moderate sample sizes ($12 \leq m, n \leq 20$). For small values of τ ($\tau < 0.25$) the estimator \underline{d}^{ML} has the best performance, for moderate values

of τ ($0.75 \leq \tau < 2.0$), the estimator \underline{d}^{KS} and \underline{d}^{GD} compete each other. For $\tau = 1$ the estimator \underline{d}^{GM} performs the best. For large values of τ the estimator \underline{d}^{BC1} and \underline{d}^{ML} compete each other.

Consider the case of large sample sizes ($m, n \geq 30$). For small values of τ the estimator \underline{d}^{ML} has the maximum percentage of risk improvement, for moderate values of τ the estimators \underline{d}^{ML} , \underline{d}^{GD} , \underline{d}^{KS} , \underline{d}^{TK} , and \underline{d}^{MK} compete each other. However for $\tau = 1$ the estimator \underline{d}^{GM} has the best performance. For large values of τ the estimators \underline{d}^{ML} , \underline{d}^{GD} , \underline{d}^{BC1} , \underline{d}^{BC2} and \underline{d}^{KS} perform equally well.

On the basis of the above discussion and observations the following recommendations may be done for the use of the estimators.

- (i) We conclude from the above discussion that, none of the estimators completely dominate others in terms of the risk function for the full range of the parameters.
- (ii) When the sample sizes are small that is $m, n \leq 10$, the estimators \underline{d}^{ML} and \underline{d}^Y can be used if τ is near to 0. For values of τ in the neighborhood of 1, the estimators \underline{d}^{MK} and \underline{d}^{TK} may be used, however for $\tau = 1$, that is, when the variances of the two populations are same, the estimator \underline{d}^{GM} should be used. For large values of τ we recommend to use \underline{d}^{BC1} .
- (iii) When the sample sizes are from moderate to large the estimators \underline{d}^{ML} , \underline{d}^{GD} , or \underline{d}^{KS} may be used if τ is near to 0, however for moderate values of τ we recommend to use either of the estimators \underline{d}^{GD} , \underline{d}^{KS} , \underline{d}^{MK} , \underline{d}^{TK} , or \underline{d}^{ML} . For values of $\tau = 1$, the estimator \underline{d}^{GM} is strongly recommended to use. For large values of τ , the estimators \underline{d}^{ML} , \underline{d}^{BC1} , or \underline{d}^{BC2} may be used.
- (iv) A similar type of observations have been made for other combinations of sample sizes and η .

Table 1: Percentage of Relative Risk Improvements of Various Estimators of Normal Quantiles with $\eta = 1.960$, $(m, n) = (8, 8), (12, 12), (20, 20), (40, 40)$

$\tau \downarrow$	<i>PR1</i>	<i>PR2</i>	<i>PR3</i>	<i>PR5</i>	<i>PR6</i>	<i>PR7</i>	<i>PR8</i>	<i>PR10</i>
0.05	98.72	74.20	98.72	55.40	53.86	30.17	98.46	98.73
	98.76	74.20	98.76	70.66	75.75	63.20	98.52	98.76
	98.76	74.19	98.76	83.70	88.49	84.65	98.54	98.76
	98.80	74.22	98.80	92.81	94.90	95.05	98.58	98.80
0.15	89.50	68.21	89.45	50.46	48.60	27.28	88.06	89.50
	89.73	68.29	89.74	64.02	68.62	57.31	88.45	89.76
	89.75	68.39	89.79	76.20	80.40	76.96	88.59	89.79
	90.16	68.66	90.20	84.69	86.62	86.77	89.05	90.20
0.25	75.29	59.03	75.28	42.83	40.68	22.89	73.34	75.43
	75.77	59.25	75.89	54.65	57.81	48.33	73.97	75.93
	76.07	59.45	76.28	64.74	68.12	65.22	74.43	76.29
	76.32	59.75	76.57	71.96	73.48	73.62	74.77	76.58
0.50	40.53	36.89	41.99	24.76	22.66	12.60	41.28	41.92
	41.25	37.75	43.22	31.97	33.03	27.52	42.35	43.21
	41.60	38.23	44.02	37.94	39.44	37.74	43.01	44.03
	41.79	38.54	44.60	42.16	42.88	42.95	43.48	44.60
0.75	17.86	24.31	23.92	15.07	13.37	07.07	24.71	23.45
	18.06	24.87	25.04	19.29	19.57	15.87	25.52	24.77
	17.79	24.86	25.47	22.42	22.99	21.74	25.63	25.40
	17.45	25.02	25.81	24.66	24.94	24.90	25.81	25.79
1.00	-0.86	15.68	13.55	09.61	08.38	04.16	15.01	12.70
	-0.18	16.58	15.10	12.37	12.39	09.65	16.14	14.64
	01.15	17.20	16.38	14.65	14.93	13.72	16.98	16.24
	00.56	16.99	16.51	15.83	15.97	15.76	16.85	16.48
1.25	-15.31	10.10	08.78	06.77	05.86	02.75	10.05	07.96
	-17.90	09.50	09.08	08.06	08.01	06.04	09.81	08.69
	-16.36	10.49	10.32	09.62	09.72	08.80	10.81	10.18
	-16.82	10.67	11.02	10.71	10.78	10.60	11.17	10.97
1.50	-31.46	04.87	05.88	05.07	04.33	01.97	06.51	05.19
	-32.18	04.77	06.44	06.00	05.87	04.38	06.62	06.29
	-34.66	04.40	06.96	06.76	06.80	06.21	06.68	06.88
	-34.23	05.07	08.09	07.90	07.92	07.80	07.54	08.08
2.00	-67.82	-5.69	02.45	02.94	02.54	01.13	01.45	02.30
	-70.73	-6.03	03.48	03.54	03.47	02.61	01.95	03.46
	-69.55	-4.90	04.53	04.35	04.37	03.99	03.00	04.51
	-72.31	-5.85	04.57	04.52	04.53	04.48	02.63	04.57
2.50	-116.05	-18.14	01.11	01.89	01.68	00.77	-1.87	01.24
	-115.39	-17.45	02.27	02.46	02.40	01.82	-0.84	02.38
	-120.07	-18.71	02.59	02.68	02.66	02.47	-1.00	02.68
	-119.29	-18.49	02.74	02.78	02.77	02.75	-0.81	02.76
3.00	-169.15	-31.75	00.49	01.45	01.21	00.57	-4.27	01.09
	-170.01	-31.47	01.27	01.56	01.57	01.22	-3.49	01.42
	-172.46	-31.42	02.11	02.12	02.09	01.93	-2.55	02.17
	-176.39	-32.49	02.20	02.17	02.17	02.14	-2.59	02.20
4.00	-293.23	-61.96	00.38	00.88	00.78	00.36	-6.55	00.80
	-304.09	-65.67	00.48	00.81	00.82	00.66	-7.06	00.68
	-311.09	-66.74	01.19	01.20	01.19	01.10	-6.02	01.23
	-319.53	-69.42	01.03	01.06	01.06	01.05	-6.51	01.05

Table 2: Percentage of Relative Risk Improvements of Various Estimators of Normal Quantiles with $\eta = 1.960$, $(m, n) = (4, 10), (12, 20), (30, 40)$

τ	<i>PR1</i>	<i>PR2</i>	<i>PR3</i>	<i>PR4</i>	<i>PR5</i>	<i>PR6</i>	<i>PR8</i>	<i>PR9</i>	<i>PR10</i>
0.05	98.74	90.72	98.72	98.74	48.01	53.65	98.51	98.53	98.74
	98.69	84.86	98.69	98.69	79.81	83.67	98.53	98.54	98.69
	98.67	80.62	98.67	98.67	91.94	93.71	98.50	98.50	98.67
0.15	89.99	82.96	89.77	89.92	43.60	48.41	88.72	88.83	89.98
	89.34	77.33	89.35	89.35	72.34	75.63	88.45	88.46	89.36
	89.21	73.63	89.24	89.24	83.16	84.76	88.31	88.31	89.24
0.25	76.50	71.03	75.95	76.33	36.96	40.55	74.39	74.57	76.48
	75.22	66.08	75.27	75.28	61.08	63.67	73.81	73.84	75.30
	74.94	62.96	75.07	75.07	69.93	71.23	73.52	73.53	75.08
0.50	46.57	44.83	45.59	46.46	22.71	23.98	44.43	44.65	46.53
	43.61	41.04	44.23	44.28	36.01	37.19	42.89	42.93	44.31
	42.16	39.03	43.48	43.49	40.56	41.16	42.17	42.17	43.50
0.75	29.00	29.96	28.53	29.48	14.74	15.03	28.38	28.58	28.88
	24.71	27.02	27.10	27.17	22.40	22.85	26.75	26.77	27.11
	20.91	25.07	25.53	25.53	24.10	24.37	25.31	25.32	25.52
1.00	18.38	21.72	19.56	20.22	10.59	10.54	20.24	20.38	19.29
	11.35	18.16	17.45	17.48	14.86	15.01	17.78	17.79	17.36
	07.2	17.09	16.77	16.77	15.93	16.05	16.96	16.96	16.74
1.25	09.62	15.57	13.89	14.02	07.97	07.77	15.03	15.12	13.32
	02.77	13.30	12.91	12.93	11.06	11.09	13.50	13.51	12.74
	-4.45	11.25	11.45	11.45	11.02	11.02	11.77	11.77	11.41
1.50	03.39	11.56	10.57	10.33	06.29	06.10	12.03	12.07	09.84
	-8.37	07.64	09.18	09.15	08.16	08.13	09.66	09.65	08.99
	-15.42	06.25	08.30	08.29	07.98	07.99	08.34	08.33	08.27
2.00	-13.23	01.65	05.59	03.91	04.05	03.82	06.86	06.78	04.87
	-29.23	-1.70	05.12	05.03	04.86	04.83	05.06	05.03	05.00
	-44.68	-4.36	04.66	04.64	04.65	04.64	03.58	03.57	04.65
2.50	-29.19	-6.74	03.43	00.82	02.96	02.77	04.27	04.11	02.86
	-54.92	-12.23	03.36	03.25	03.32	03.29	02.20	02.15	03.32
	-73.59	-14.60	02.94	02.93	02.98	02.97	00.74	00.73	02.95
3.00	-49.24	-17.00	02.39	-0.89	02.31	02.15	02.43	02.18	01.92
	-87.01	-25.22	02.26	02.13	02.37	02.34	-0.28	-0.36	02.31
	-110.99	-26.79	02.34	02.33	02.34	02.31	-0.94	-0.96	02.36
3.50	-76.25	-31.59	00.89	-3.42	01.63	01.51	-0.37	-0.76	00.76
	-120.65	-38.15	01.77	01.66	01.85	01.83	-1.67	-1.76	01.82
	-159.48	-43.07	01.61	01.59	01.63	01.64	-3.08	-3.10	01.62
3.75	-83.63	-34.56	01.41	-2.41	01.64	01.53	00.17	-0.21	01.22
	-138.71	-45.64	01.33	01.21	01.49	01.47	-2.67	-2.77	01.41
	-181.46	-49.60	01.60	01.59	01.59	01.58	-3.22	-3.25	01.61
4.00	-99.21	-42.8	00.90	-3.31	01.42	01.31	-0.97	-1.42	00.81
	-157.38	-53.73	00.97	00.86	01.18	01.16	-3.7	-3.80	01.08
	-199.50	-55.59	01.39	01.39	01.38	01.37	-3.54	-3.57	01.40

Table 3: Percentage of Relative Risk Improvements of Various Estimators of Normal Quantiles with $\eta = 1.960$, $(m, n) = (10, 4), (20, 12), (40, 30)$

τ	<i>PR1</i>	<i>PR2</i>	<i>PR3</i>	<i>PR4</i>	<i>PR5</i>	<i>PR6</i>	<i>PR8</i>	<i>PR9</i>	<i>PR10</i>
0.05	96.66	47.61	96.65	96.61	31.74	06.15	96.10	96.04	96.66
	97.65	59.62	97.65	97.65	73.60	76.84	97.28	97.28	97.65
	98.16	66.31	98.16	98.16	90.21	93.39	97.90	97.89	98.16
0.15	74.93	38.57	74.89	73.82	25.03	04.75	73.28	73.07	75.01
	81.95	51.36	82.00	81.98	62.19	64.41	80.44	80.41	82.01
	85.51	58.82	85.58	85.57	78.61	81.31	84.22	84.21	85.58
0.25	51.02	28.76	51.30	49.01	17.79	03.27	50.24	50.00	51.33
	60.73	40.31	61.25	61.21	46.87	48.12	59.58	59.54	61.27
	66.65	47.91	67.08	67.08	61.81	63.78	65.37	65.36	67.09
0.50	11.92	13.43	17.02	15.51	07.06	01.24	18.18	18.11	16.37
	23.63	21.38	27.08	26.99	21.28	21.32	26.88	26.85	27.00
	30.09	27.26	33.27	33.26	30.78	31.58	32.56	32.55	33.27
0.75	-7.17	07.28	05.65	05.53	03.32	00.57	07.09	07.16	04.80
	01.51	12.33	12.52	12.51	10.49	10.37	13.11	13.11	12.31
	07.33	16.57	17.34	17.34	16.42	16.67	17.52	17.52	17.31
1.00	-22.64	04.15	01.39	02.30	01.87	00.34	02.23	02.41	01.16
	-13.11	07.75	06.84	06.88	06.14	05.98	07.26	07.27	06.68
	-9.14	10.03	09.55	09.55	09.27	09.34	09.76	09.77	09.52
1.25	-37.14	01.91	-0.45	00.84	01.11	00.21	-0.44	-0.18	-0.40
	-27.24	04.17	04.00	04.06	03.82	03.70	03.75	03.78	03.96
	-21.59	06.40	06.57	06.58	06.36	06.38	06.46	06.46	06.55
1.50	-50.31	00.63	-0.75	00.55	00.87	00.17	-1.31	-1.01	-0.17
	-42.87	00.85	02.20	02.30	02.41	02.32	01.18	01.23	02.27
	-38.20	02.26	04.45	04.46	04.37	04.39	03.55	03.56	04.45
2.00	-89.75	-3.66	-2.23	-0.43	00.41	00.09	-4.51	-4.06	-0.24
	-77.46	-4.60	01.17	01.26	01.41	01.33	-1.18	-1.11	01.33
	-72.94	-5.01	02.44	02.45	02.46	02.45	00.30	00.31	02.46
2.50	-130.42	-7.14	-2.33	-0.55	00.26	00.06	-5.74	-5.21	-0.18
	-125.44	-11.86	00.60	00.68	00.85	00.81	-3.25	-3.16	00.78
	-115.91	-13.22	1.64	01.65	01.64	01.63	-1.63	-1.61	01.66
3.00	-190.10	-12.31	-2.55	-0.66	00.13	00.03	-7.50	-6.85	-0.24
	-175.49	-19.09	00.36	00.42	00.56	00.53	-4.33	-4.23	00.50
	-162.38	-21.35	01.32	01.32	01.29	01.28	-2.49	-2.47	01.32
3.50	-250.22	-17.03	-2.10	-0.49	00.10	00.03	-8.03	-7.31	-0.17
	-234.56	-27.38	00.29	00.34	00.43	00.42	-5.11	-5.00	00.40
	-225.80	-34.25	00.75	00.76	00.78	00.77	-4.54	-4.51	00.78
3.75	-285.75	-19.42	-1.88	-0.42	00.16	00.04	-7.97	-7.23	00.06
	-264.67	-31.69	00.20	00.25	00.36	00.33	-5.41	-5.30	00.32
	-264.47	-41.68	00.55	00.56	00.59	00.60	-5.42	-5.39	00.57
4.00	-328.00	-23.98	-2.36	-0.77	00.08	00.02	-9.20	-8.41	-0.09
	-296.86	-35.41	00.37	00.40	00.44	00.40	-5.16	-5.04	00.45
	-296.24	-47.50	00.50	00.51	00.53	00.53	-5.68	-5.65	00.52

5. Concluding Remarks and Illustrative Examples

We note here that, in the literature most of the results on estimation of quantiles are for a single parameter $\theta = \mu + \eta\sigma$ either using one or more populations. In this article, we consider the simultaneous estimation of the quantile vector $\underline{\theta} = (\theta_1, \theta_2)$ which is important from an application point of view. The loss function is taken as the sum of the quadratic loss functions. It should be noted that, Kumar and Tripathy [9] considered this model and estimated the first component θ_1 with respect

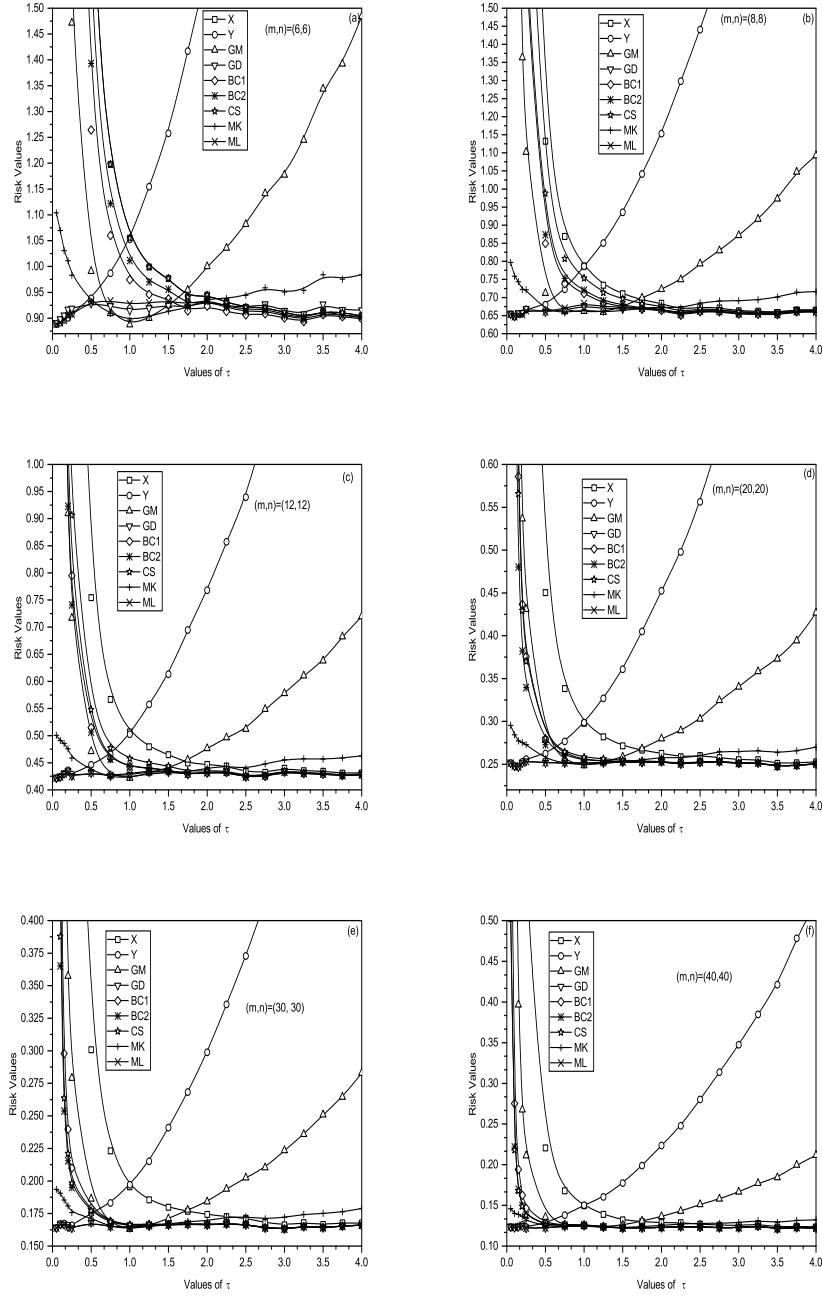


Figure 1. Comparison of risk values of various estimators for quantile vector (θ_1, θ_2) when $\eta = 1.960$ and the sample sizes are equal

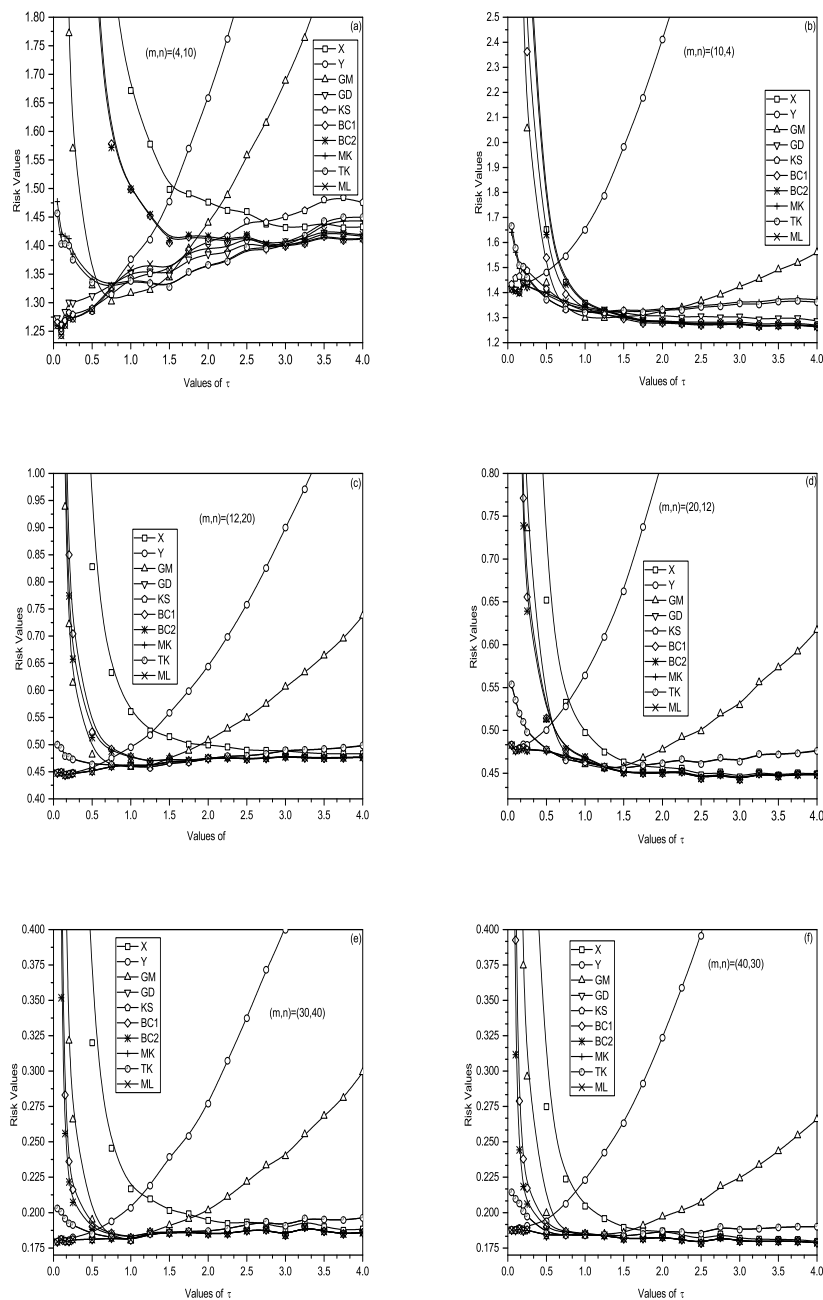


Figure 2. Comparison of risk values of various estimators for quantile vector (θ_1, θ_2) when $\eta = 1.960$ and the sample sizes are unequal

to a quadratic loss function. We have implemented the Brewster and Zidek [3] technique to the case of estimating a vector parameter, which is interesting. Further we have proposed some new estimators such as the \underline{d}^Y , \underline{d}^{GM} , and \underline{d}^{ML} which was not considered by them. First, we derived sufficient conditions for improving equivariant estimators and in the process some complete class results obtained. We have constructed some improved estimators using one of our result obtained in Section 2. However, the analytical comparison of these estimators is not possible. We have conducted a detailed simulation study to numerically compare these estimators which can be used in practice. Our conclusions regarding the use of the estimators are completely based on the simulation study as no analytical comparison is possible among all the estimators. It will be interesting to generalize the results to the case of $k \geq 3$ normal populations, where proving inadmissibility of these estimators will be challenging. Below we present some examples where our model fits well and also compute the estimates for practical purposes. In the examples below we have taken the value of $\eta = 1.960$ for convenient.

5.1. Example. We consider the example discussed in Hines et al. [6], (p. 290). Suppose a manufacturer of video display units produces two micro circuit designs design A and design B. He wants to test whether the two design produce same current flow. The summarized data for design A are given by $m = 15$, $\bar{x} = 24.2$, $s_1^2 = 10$ where as the data for design B are given by $n = 10$, $\bar{y} = 23.9$, $s_2^2 = 20$. It is also given that both the data follow normal distributions with a common mean. The experimental conditions ensures that the variances are unequal. This is a situation where our model will be very much useful. The several estimators for quantiles are calculated as $\underline{d}^X = (25.97, 26.71)$, $\underline{d}^Y = (25.67, 26.41)$, $\underline{d}^{GM} = (25.85, 26.59)$, $\underline{d}^{GD} = (25.92, 26.65)$, $\underline{d}^{KS} = (25.92, 26.66)$, $\underline{d}^{BC1} = (25.97, 26.71)$, $\underline{d}^{BC2} = (25.94, 26.68)$, $\underline{d}^{MK} = (25.88, 26.61)$, $\underline{d}^{TK} = (25.88, 26.61)$ and $\underline{d}^{ML} = (25.92, 26.65)$. If the variances of both the data set differ significantly we may use either the estimator \underline{d}^{GD} , \underline{d}^{ML} , or \underline{d}^{BC1} . If the variances differ marginally we may use either \underline{d}^{KS} , or \underline{d}^{MK} .

5.2. Example. Rohatgi and Saleh [13], (p.515) discussed one example regarding the mean life time (in hours) of light bulbs. Suppose a random sample of 9 bulbs has sample mean 1309 hours with standard deviation of 420 hours. A second sample of 16 bulbs chosen from a different batch has sample mean 1205 hours and standard deviation 390 hours. A two sample t-test fails to reject the hypothesis that the means are equal. This is a situation where our model will be useful. Suppose we want to know the life time of both the bulbs at any instant of time then we can use our estimators. The various estimators are calculated as $\underline{d}^X = (1543.45, 1526.70)$, $\underline{d}^Y = (1439.45, 1422.70)$, $\underline{d}^{GM} = (1476.89, 1460.14)$, $\underline{d}^{GD} = (1460.82, 1444.08)$, $\underline{d}^{KS} = (1458.47, 1441.73)$, $\underline{d}^{BC1} = (1501.44, 1484.69)$, $\underline{d}^{BC2} = (1498.38, 1481.64)$, $\underline{d}^{MK} = (1474.51, 1457.77)$, $\underline{d}^{TK} = (1474.18, 1457.43)$ and $\underline{d}^{ML} = (1457.08, 1440.33)$. Also a F-test fails to reject the hypothesis that the population variances are equal. In this situation we recommend to use either \underline{d}^{TK} , or \underline{d}^{MK} .

References

- [1] Alexander, T. L. and Chandrasekar, B. *Simultaneous equivariant estimation of the parameters of matrix scale and matrix location-scale models*, Statistical Papers **46**, 483-507, 2005.
- [2] Bai, S. K. and Durairajan, T. M. *Simultaneous equivariant estimation of the parameters of linear models*, Statistical Papers **39**, 125-134, 1998.
- [3] Brewster, J. F. and Zidek, J. V. *Improving on equivariant estimators*, Annals of Statistics **2**, 21-38, 1974.
- [4] Brown, L. D. and Cohen, A. *Point and confidence estimation of a common mean and recovery of interblock information*, Annals of Statistics **2** (5), 963-976, 1974.
- [5] Chang, C. H. and Pal, N. *Testing on the common mean of several normal distributions*, Computational Statistics and Data Analysis **53**, 321-333, 2008.
- [6] Hines, W. W., Montgomery, D. C., Goldsman, D. M. and Borror, D. M. *Probability and Statistics in Engineering*, John Wiley, New Work, 2008.
- [7] Keating, J. P. and Tripathi, R. C. *Percentiles, estimation of 'Encyclopedia of Statistical Sciences*, VI, 668-674, 1985.
- [8] Kiefer, J. *Invariance minimax sequential estimation and continuous time processes*, Annals of Mathematical Statistics **28**, 573-601, 1957.
- [9] Kumar, S. and Tripathy, M. R. *Estimating quantiles of normal populations with a common mean*, Communications in Statistics-Theory and Methods **26**, 115-118, 2011.
- [10] Lin, S. H. and Lee, J. C. *Generalized inferences on the common mean of several normal populations*, Journal of Statistical Planning and Inferences **134**, 568-582, 2005.
- [11] Moore, B. and Krishnamoorthy, K. *Combining independent normal sample means by weighting with their standard errors*, Journal of Statistical Computation and Simulations **58**, 145-153, 1997.
- [12] Pal, N., Lin, J. J., Chang, C. H. and Kumar, S. *A revisit to the common mean problem: Comparing the maximum likelihood estimator with the Graybill-Deal estimator*, Computational Statistics and Data Analysis **51**, 5673-5681, 2007.
- [13] Rohatgi, V. K. and Saleh, A. Md. E. *An Introduction to Probability and Statistics*, John Wiley, 2nd Edition, New Work, 2003.
- [14] Rukhin, A. L. *A class of minimax estimators of a normal quantile*, Statistics and Probability Letters **1**, 217-221, 1983.
- [15] Rukhin, A. L. *Admissibility and minimaxity results in the estimation problem of exponential quantiles*, Annals of Statistics **14**, 220-237, 1986.
- [16] Saleh, A. K. Md. E. *Estimating quantiles of exponential distributions*, In: Statistics and Related Topics, Csorgo, M., Dawson, D., Rao, J. N. K., Saleh, A. K. Md. E. (Eds.), North Holland, Amsterdam, 279-283, 1981.
- [17] Sharma, D. and Kumar, S. *Estimating quantiles of exponential populations*, Statistics and Decisions, **12**, 343-352, 1994.
- [18] Tripathy, M. R. and Kumar, S. *Estimating a common mean of two normal populations*, Journal of Statistical Theory and Applications **9** (2), 197-215, 2010.
- [19] Tsukuma, T. *Simultaneous estimation of restricted location parameters based on permutation and sign-change*, Statistical Papers, **53**, 915-934, 2012.
- [20] Vazquez, G., Duval, S., Jacobs Jr, D. R. and Silventoinen, K. *Comparison of body mass index, waist circumference and waist/hip ratio in predicting incident diabetes: A meta analysis*, Epidemiologic Reviews **29** (1), 115-128, 2007.
- [21] Zidek, J. V. *Inadmissibility of the best invariant estimators of extreme quantiles of the normal law under squared error loss*, Annals of Mathematical Statistics **40** (5), 1801-1808, 1969.
- [22] Zidek, J. V. *Inadmissibility of a class of estimators of a normal quantile*, The Annals of Mathematical Statistics **42** (4), 1444-1447, 1971.

Estimation of the waiting time of patients in a hospital with simple Markovian model using order statistics

Soma Dhar*, Lipi B. Mahanta^{†‡} and Kishore K. Das[§]

Abstract

In this paper, consider a single server queue in a hospital environment whose service time is governed by a Markov process. It is possible that the server changes its service speed many times while serving a patient. Here we have studied the order statistics for waiting time distribution where the probability density function of single order statistics $\phi_{i:n}$, cumulative density function of $\Phi_{i:n}$, joint probability density function of $\phi_{i:n}$ and $\phi_{j:n}$, probability density function of extreme order statistics. Also have been considered the moments and recurrence relation of order statistics, the probability density function of sample range and sample median. We derive minimum and maximum order statistics of the service time of patients in the system using first step analysis to obtain an insight on the service process. Further, we use order statistics to compute performance measures such as average queue length and waiting time for severe diseases especially in the outpatient department. This result effectively establishes that as the number of server increases, then the utmost and the minimum waiting time of the patients decreases. Also illustrate the application of the simple Markovian model by using real hospital data.

Keywords: Waiting time, Service time, Patients, Order statistics.

2000 AMS Classification: 60K25

Received: 11.09.2017 *Accepted:* 28.06.2018 *Doi:* 10.15672/HJMS.2018.607

*Department of Statistics, Gauhati University, Guwahati., Email: somadhar7@gmail.com

†Centre for Computational and Numerical Sciences, Institute of Advanced Study in Science and Technology, Guwahati - 781035, India Email: lbmahanta@iasst.gov.in

‡Corresponding Author.

§Department of Statistics, Gauhati University, Guwahati, Email: daskkishore@gmail.com

1. Introduction

The theory of queues has wide applications in the field of health-care management system. The study of queuing systems in hospitals has often been concerned with the busy period and waiting time, because they play a very significant role there. A queuing system is normally described by the patient's entry into a queue, who are then served at a service point by the server (doctor), after which they leave the queue.

Dhar et al.[9] studied the comparison between single and multiple Markovian queuing model in an outpatient department. Also Mahanta et al.[13] proposed a single server queueing model for severe diseases especially in outpatient department. Further consider the infinite server queues with time-varying arrival and departure pattern when the parameters are varying with time derive by Dhar et al.[16].

Order statistics are widely used in applications of statistical models and inference. Both describes random variables which are arranged in order of magnitude. According to Aleem [1], usually the ordered values of independent and identically distributed samples arranged in ascending order of magnitude are known as order statistics. The simplest and most important function of order statistics is the sample cumulative distribution function $F_n(x)$. Suppose X_1, X_2, \dots, X_n are n jointly distributed random variables. Arranging the X 's in increasing order of magnitude, $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ are said to be smallest, second smallest and largest order statistics. Thus $X_{1:n} < X_{2:n}, \dots, < X_{n:n}$. Arnold et al. [2] and David and Nagaraja [7] studied order statistics and functions of these statistics as it plays an important role in wide range of theoretical and practical problems such as characterizations of probability distributions and goodness of fit test, entropy estimation, analysis of censored samples, reliability analysis, quality control and strength of materials. Order statistics arise naturally in many real-life applications involving data relating to life testing studies proved by Shawky[10]. Aleem [1] reported that methods of interpretation based on order statistics are most efficient and are used extensively because of robustness and parsimonious nature. The sample mean and standard deviation provide efficient estimators of the corresponding population parameter under the assumption of normality, but sample range is simpler to use than the sample standard deviation in statistical quality control and the sample median and its deviation furnish more robust estimators when the population have long tail. Extreme (largest and smallest) values statistics, which is an offspring of order statistics, has its importance in hydrology, aeronautics, oceanography, material strength, signal processing and meteorology. Moments of order statistics also plays an important role in the area of quality control testing and reliability. According to David and Nagaraja [8] moments of order statistics can be used to measure the failure rate of reliability and to predict the failure of future events.

A recursive procedure for computing the moments of the busy period for the single-server model can be found in Tarabia [12]. Limit theorems are proved by investigating the extreme values of the maximum queue length, the waiting time and virtual waiting time for different queue models in literature. Serfozo [14] discussed the asymptotic behavior of the maximum value of birth-death processes over large time intervals. Serfozo's results concerned the transient and recurrent birth-death processes and related $M/M/c$ queues. Asmuseen[4] introduced a survey of the present state of extreme value theory for queues and focused on the regenerative properties of queuing systems, which reduced the problem to study the tail of the maximum of the queuing process $X(t)$ during a regenerative cycle, where $X(t)$ is in discrete or continuous time. Artalejo et al.[3] presented an efficient algorithm for computing the distribution for the maximum number of

customers in orbit and in the system during a busy period for the $M/M/c$ retrial queue. The main idea of their algorithm is to reduce the computation of the distribution of the maximum customer number in orbit by computing certain absorption probabilities. For more details of extreme value in queues by Park [15].

In this paper, we studied the maximum and minimum service and waiting time respectively, of the patients who suffer from severe disease especially in public hospital. Here we considered one of the leading public hospitals of the region, *viz.* Pandu P.H.C/F.R.U, Guwahati where it was observed that there was a heavy flow of patients throughout the day. Data was collected from Hospital (*viz.* outpatient department) and from other allied sources. The current chapter will have utility for various practical problems for which the distributions of order statistics play a role and the queuing theory implicit to the health related problems.

2. Formulation of the problem

Let X_1, X_2, \dots, X_n be a random sample from a continuous population with probability density function $\phi(x)$ and cumulative distribution function $\Phi(x)$ and $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ be the order statistics obtained by arranging the random sample in increasing order of magnitude. Then according to David and Nagaraja [7] the probability density function of the i^{th} order statistics $X_{i:n}$, $1 < i < n$ is given by

$$(2.1) \quad \phi_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} [\Phi(x)]^{i-1} [1 - \Phi(x)]^{n-i} \phi(x), -\infty < x < \infty$$

The probability density functions of smallest and largest order statistics are given by Arnold et al.[2] as

$$(2.2) \quad \phi_{1:n} = n[1 - \Phi(x)]^{n-1} \phi(x), -\infty < x < \infty$$

and

$$(2.3) \quad \phi_{n:n} = n[\Phi(x)]^{n-1} \phi(x), -\infty < x < \infty$$

respectively. According to Arnold et al.[2] the cumulative density functions of smallest and largest order statistics are given as

$$(2.4) \quad \Phi_{1:n} = 1 - [1 - \Phi(x)]^n, -\infty < x < \infty$$

and

$$(2.5) \quad \Phi_{n:n} = [\Phi(x)]^n, -\infty < x < \infty$$

respectively.

The p^{th} order moment for the i^{th} order statistics is also given by Arnold et al.[2] as

$$(2.6) \quad \begin{aligned} \mu'(r : n) &= \int_{-\infty}^{\infty} x^p \phi_{r:n}(x) dx \\ &= \frac{n!}{(i-1)!(n-i)!} \int_{-\infty}^{\infty} x^p [\Phi(x)]^{i-1} [1 - \Phi(x)]^{n-i} \\ &\quad \phi(x) dx, -\infty < x < \infty \end{aligned}$$

Assuming $u = X_{i:n}$ and $v = X_{j:n}$ as the i^{th} and j^{th} order statistics, $1 < i < j < n$ from n independent random variable each with probability density function $\phi(x)$, the joint density function of $u = X_{i:n}$ and $v = X_{j:n}$ is given by Arnold et al. [2], as

$$(2.7) \quad \begin{aligned} \phi(u, v : n) &= c'(i, j, n) [\Phi(u)]^{i-1} [\Phi(v) - \Phi(u)]^{i-j-1} [1 - \Phi(v)]^{n-j} \\ &\quad \phi(v) \phi(u), -\infty < u < v < \infty \end{aligned}$$

where $c'(i, j, n) = \frac{n!}{(i-1)!(i-j-1)!(n-i)!}$

David and Nagaraja [7], has given the probability density function of double moment as

$$(2.8) \quad \mu'_{p,q}(i, j : n) = \frac{n!}{(i-1)!(i-j-1)(n-j)!} \int \int_{0 < u < v < \infty} u^p v^q dudv, \quad -\infty < x < \infty$$

Arnold et al.[2] defined the sample range as $W_n = X_{n:n} - X_{1:n}$ and

$$(2.9) \quad \phi_{w_n}(w) = n(n-1) \int_{-\infty}^{\infty} [\Phi(x_1 + w) - \Phi(x_1)]^{n-2} \phi(x_1) \phi(x_1 + w) dx_1, \quad 0 < w < \infty$$

Percentage points of distributions are the most fundamental tools used in test of hypothesis to take decision about various situations of the population based on sample observations and also used to express the difference of risks of probabilities. The percentile points are the point on the measurement scale below which a specified percentage of score falls. In many applications involving these distributions percentage points are required. Bagui[5] defined the percentage points depends on the evaluation of the inverse probability function. In general, percentile points of the distributions have been obtained using approximation, interpolation formula, quadrature formula and by simulation. According to White [11] the p^{th} percentile equation of distribution is given as

$$(2.10) \quad \int_0^x \phi(x) dx = p$$

where p denotes level of significance.

The p^{th} percentile equation of smallest and largest order statistics are given as

$$(2.11) \quad \Phi_{1:n} = \int_0^x \Phi_{1:n} dx = 1 - [1 - \Phi(x)]^n = p$$

and

$$(2.12) \quad \Phi_{n:n} = \int_0^x \Phi_{n:n} dx = [\Phi(x)]^n = p$$

Let $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ be the order statistics of a random variables X_1, X_2, \dots, X_n . Also let $T_{i:n} = X_{i:n} - X_{i-1:n}$, $i = 1, 2, 3, \dots, n$, where T_i represents the difference between each arrival into the system(inter-arrival) of the order statistics $X_{i:n}$ and $X_{i-1:n}$. Then the random variables T_1, T_2, \dots, T_n are called the inter-arrival time between the successive order statistics $X_{1:n}, X_{2:n}, \dots, X_{n:n}$.

Here we consider the sample range R which is denoted by

$$(2.13) \quad R = \sum_{i=2}^n T_{i:n}$$

Moreover, this can be used to construct the interval for the corresponding patients.

2.1. Theorem. *Let T_1, T_2, \dots, T_n be the random sample of size n from a continuous distribution with cumulative density function Φ and the probability density function ϕ . Then the joint distribution of order statistics is given by*

$$(2.14) \quad \Phi_{T_i}(t) = 1 - \int_{-\infty}^{\infty} (i-1) \binom{n}{i-1} \Phi_X^{(i-2)}(y) [1 - \Phi_X(x+t)]^{(n-i+1)} \phi_X(x) dx$$

Proof. We know that,

$$\begin{aligned}
 \Phi_{T_i}(t) &= P(T_i \leq t) \\
 &= P(X_{(i)} - X_{(i-1)} \leq t) \\
 (2.15) \quad &= P(X_{(i)} \leq X_{(i-1)} + t)
 \end{aligned}$$

Let w be the region bounded by $X_{(i)} \leq X_{(i-1)}$ and $X_{(i)} \leq X_{(i-1)} + t$. From equation (2.15),

$$\begin{aligned}
 \Phi_{T_i}(t) &= \int \int_w \phi_{T_i, T_{i-1}}(x, y) dx dy \\
 &= \int_{-\infty}^{\infty} \int_x^{x+t} \frac{n! \phi_X(x) \phi_X(y)}{(i-2)! (n-i)!} [1 - \Phi_X(y)]^{(n-i)} \Phi_X^{(i-2)}(y) \\
 &\quad [\Phi_X(y) - \Phi_X(x)]^{i-i+1-1} dx dy \\
 &= \int_{-\infty}^{\infty} \frac{n! \phi_X(x)}{(i-2)! (n-i)!} \Phi_X^{(i-2)}(y) \left(\int_x^{x+t} \phi_X(y) [1 - \Phi_X(y)]^{(n-i)} dy \right) dx \\
 &= \int_{-\infty}^{\infty} \frac{n! \phi_X(x)}{(i-2)! (n-i)!} \Phi_X^{(i-2)}(y) \left| \frac{-[1 - \Phi_X(y)]^{(n-i+1)}}{(n-i+1)} \right|_x^{x+t} dx \\
 &= \int_{-\infty}^{\infty} \frac{n! \phi_X(x)}{(i-2)! (n-i+1)!} \Phi_X^{(i-2)}(y) \\
 &\quad \left[-1 - \Phi_X(x+t)^{(n-i+1)} + 1 - \Phi_X(x)^{(n-i+1)} \right] dx \\
 &= \int_{-\infty}^{\infty} \frac{n! \phi_X(x)}{(i-2)! (n-i+1)!} \Phi_X^{(i-2)}(y) 1 - \Phi_X(x)^{(n-i+1)} dx \\
 &\quad - \int_{-\infty}^{\infty} \frac{n! \phi_X(x)}{(i-2)! (n-i+1)!} \Phi_X^{(i-2)}(y) 1 - \Phi_X(x+t)^{(n-i+1)} dx
 \end{aligned}$$

by using equation (2.15), since it is the integral of the $(i-1)^{th}$ order statistics over $(-\infty, \infty)$

$$\begin{aligned}
 \Phi_{T_i}(t) &= 1 - \int_{-\infty}^{\infty} \frac{n! \phi_X(x)}{(i-2)! (n-i+1)!} \Phi_X^{(i-2)}(y) 1 - \Phi_X(x+t)^{(n-i+1)} dx \\
 &= 1 - \int_{-\infty}^{\infty} (i-1) \binom{n}{i-1} \Phi_X^{(i-2)}(y) 1 - \Phi_X(x+t)^{(n-i+1)} \phi_X(x) dx
 \end{aligned}$$

□

2.2. Corollary. Let $i = 1$ and $i = n$ in Theorem 2.1

$$\begin{aligned}
 \Phi_{T_1}(t) &= P(T_1 \leq t) \\
 &= P(X_{(1)} \leq t) \\
 &= 1 - [1 - \Phi(t)]^n
 \end{aligned}$$

2.3. Theorem. Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ denote the order statistics of a random sample of size n from a continuous distribution with probability density function $\phi_X(x)$ and a cumulative density function $\Phi_X(x)$. Then the probability density function of the j^{th} order statistics is given by

$$(2.16) \quad \phi_{X_{(j)}}(x) = \frac{n! \phi_X(x)}{(j-1)! (n-j)!} \Phi_X^{(j-1)}(x) 1 - \Phi_X(x)^{(n-j)}$$

Proof. Proof of the theorem 2.3 given in Artalejo et al.[3].

□

2.4. Corollary. Let $\Phi_{T_n}(t) = X_{(n)} - X_{(n-1)}$, $0 < t < \infty$. Then
 $P(T_n \leq t) = 1 - \int_{-\infty}^{\infty} n(n-1)\Phi_X^{(n-2)}(y) [1 - \Phi_X(x+t)]\phi_X(x) dx$

3. Order statistics for waiting time distribution

When a patient wait for service, the two most important characteristics that arises are (i) time spent in the queue and(ii) time spent in the system. Considering the system is in equilibrium, let T_q and T be the amount of time a customer spends in queue and in the system, respectively. However, the waiting time for service (T_q) of an arriving customer is the amount of time required to serve the customers already in the system. The total time in system T is $T_q +$ service time. When there are n customers in the system, since service times are exponential with parameter μ , the total service time of n customers is Erlang with probability density

$$(3.1) \quad \phi_n(x) = e^{-\mu x} \frac{\mu^n x^{n-1}}{(n-1)!}$$

Let $Fq(t) = P(Tq \leq t)$ be the distribution function of the waiting time T_q . Here $\Phi q(0) = P(Tq = 0) = P(Q = 0) = 1 - \rho$. It is noted that because of the memoryless property of the exponential distribution, the remaining service time of the customer in service is also exponential with the same parameter μ .

Let $d\Phi q(t) = P(t < Tq \leq t + dt)$, for $t > 0$, we have

$$\begin{aligned} d\Phi q(t) &= \sum_{n=1}^{\infty} p_n e^{-\mu t} \frac{\mu^n t^{n-1}}{(n-1)!} dt \\ &= (1 - \rho) \sum_{n=1}^{\infty} \rho^n e^{-\mu t} \frac{\mu^n t^{n-1}}{(n-1)!} dt \end{aligned}$$

After simplification it is given by

$$(3.2) \quad = \mu \rho (1 - \rho) e^{-\mu(1-\rho)t} dt$$

Because of the discontinuity at 0 in the distribution of T_q , we get

$$\begin{aligned} \Phi q(t) &= P(Tq = 0) + \int_0^t d\Phi q(t) \\ (3.3) \quad &= 1 - \rho e^{-\mu(1-\rho)t}, \end{aligned}$$

The probability density function of the waiting time in the queue is given by Medhi [17]

$$w_q(t) = \begin{cases} (1 - \rho), & t = 0 \\ \mu \rho (1 - \rho) e^{-\mu(1-\rho)t}, & t > 0 \end{cases}$$

The probability density function of the waiting time in the system is given by Bhat[6]

$$w(x) = \mu(1 - \rho) e^{-\mu(1-\rho)x}, \quad x \geq 0$$

which is a exponential distribution with parameter $\mu(1 - \rho)$

3.1. Derivation of i^{th} order statistics for waiting time distribution. If X_1, X_2, \dots, X_n is a random sample from a continuous population with probability density function $\phi(x)$ and cumulative distribution function $\Phi(x)$ and $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ are the order statistics obtained by arranging the random sample in increasing order of magnitude, then the probability density function of the i^{th} order statistics $X_{i:n}$ for waiting

time distribution using (2.3), $1 < i < n$ is given by

$$\begin{aligned}\phi_{i:n_w}(x) &= \frac{n!}{(i-1)!(n-i)!} [1 - \rho e^{-\mu(1-\rho)x}]^{i-1} \\ &\quad [1 - (1 - \rho e^{-\mu(1-\rho)x})]^{n-i} e^{-\mu x} \frac{\mu^n x^{n-1}}{(n-1)!}\end{aligned}$$

Using binomial expression, the probability density function of i^{th} order statistics $X_{i:n}$, ($1 < i < n$) for waiting time distribution reduces to

$$\begin{aligned}\phi_{i:n_w}(x) &= \frac{n!}{(i-1)!(n-i)!} \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r \rho^r e^{-\mu(1-\rho)r} \\ &\quad \rho^{n-i} e^{-\mu(1-\rho)x(n-i) + \mu x} \frac{\mu^n x^{n-1}}{(n-1)!} \\ &= \frac{n!}{(i-1)!(n-i)!} \rho^{r+n-i} \frac{\mu^n}{(n-1)!} \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r e^{-\mu(1-\rho)r} e^{-i\rho\mu x} x^{n-1} \\ (3.4) \quad &= C(n, r, i) \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r e^{-\mu(1-\rho)r} e^{-i\rho\mu x} x^{n-1}, \quad x > 0, \mu > 0\end{aligned}$$

where

$$C(n, r, i) = \frac{n!}{(i-1)!(n-i)!} \rho^{r+n-i} \frac{\mu^n}{(n-1)!}$$

We observe that

$$\sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r \rho^r e^{-\mu(1-\rho)r} = [1 - \rho e^{-\mu(1-\rho)x}]^{i-1}$$

The cdf of i^{th} order statistics $X_{i:n}$, $1 < i < n$ for waiting time distribution is obtained as

$$\begin{aligned}\Phi_{i:n_w}(x) &= C(n, r, i) \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r e^{-\mu(1-\rho)r} \int_0^x e^{-i\rho\mu x} x^{n-1} dx \\ &= C(n, r, i) \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r e^{-\mu(1-\rho)r} \sum_{j=0}^{\infty} \frac{(-i\rho\mu)^j}{j!} \frac{x^n}{n}, \quad x > 0, \mu > 0\end{aligned}$$

3.2. Derivation of extreme order statistics. The probability density functions of smallest and largest order statistics can be obtained from equation (3.4) by putting $i = 1$ and $i = n$ respectively. The probability density functions of smallest and largest order statistics for waiting time distribution are obtained as

$$(3.5) \quad \phi_{1:n_w}(x) = n e^{-\mu(1-\rho)r} e^{-\rho\mu x} x^{n-1}, \quad x > 0, \mu > 0$$

and

$$(3.6) \quad \phi_{n:n_w}(x) = n \sum_{r=0}^{n-1} \binom{n-1}{r} (-1)^r e^{-\mu(1-\rho)r} e^{-n\rho\mu x} x^{n-1}, \quad x > 0, \mu > 0$$

respectively.

The cumulative density functions of smallest and largest order statistics for waiting time

distribution can be obtained using expression (2.4) and (2.5) as

$$(3.7) \quad \Phi_{1:nw}(x) = ne^{-\mu(1-\rho)^r} e^{-\rho\mu x} x^{n-1}, \quad 0 < x < \infty$$

$$(3.8) \quad \Phi_{n:nw}(x) = 1 - n \sum_{r=0}^n \binom{n}{r} (-1)^r e^{-\mu(1-\rho)^r} e^{-n\rho\mu x} x^{n-1}, \quad 0 < x < \infty$$

3.3. Moments and recurrence relations of i^{th} order statistics for waiting time distribution. Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be independent and identically distributed random sample of size n from a continuous distribution with probability density function $\phi_X(x)$ and a cumulative density function $\Phi_X(x)$ from Waiting time distribution. From the probability density function of i^{th} order statistics for waiting time distribution the p^{th} order moment can be written as

$$(3.9) \quad \begin{aligned} \mu_p'(i:n) &= \frac{n!}{(i-1)!(n-i)!(n-1)!} \frac{\mu^n}{\int_0^\infty x^{n+p-1} [1 - \rho e^{-\mu(1-\rho)x}]^{i-1} \\ &\quad [1 - (1 - \rho e^{-\mu(1-\rho)x})]^{n-i} e^{-\mu x} dx} \\ &= C(n, i) \int_0^\infty x^{n+p-1} [1 - \rho e^{-\mu(1-\rho)x}]^{i-1} \\ &\quad [1 - (1 - \rho e^{-\mu(1-\rho)x})]^{n-i} e^{-\mu x} dx \end{aligned}$$

Here we applied the binomial expression in (3.9), we get

$$\begin{aligned} \mu_p'(i:n) &= C(n, r, i) \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r e^{-\mu(1-\rho)^r} \int_0^\infty e^{-i\rho\mu x} x^{n+p-1} dx \\ &= C(n, r, i) \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r e^{-\mu(1-\rho)^r} \frac{\gamma(n+p)}{(i\mu\rho)^{n+p}} \end{aligned}$$

putting $i = n$, we get the highest order moment which is given by

$$\mu_p'(n:n) = \sum_{r=0}^{n-1} \binom{n-1}{r} (-1)^r e^{-\mu(1-\rho)^r} \frac{\gamma(n+p)}{(n\mu\rho)^{n+p}}$$

The recurrence relation for moments of i^{th} order statistics is given by

$$(3.10) \quad \begin{aligned} \mu_{p+1}'(i:n) &= \sum_{r=0}^{n-1} \binom{n-1}{r} (-1)^r e^{-\mu(1-\rho)^r} \frac{\gamma(n+p+1)}{(n\mu\rho)^{n+p+1}} \\ \implies i\mu\rho\mu_{p+1}'(i:n) &= \mu_p'(i:n) \end{aligned}$$

3.4. Mode for waiting time distribution. The modal value equation of the i^{th} order statistics is

$$\begin{aligned} C(n, r, i) \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r e^{-\mu(1-\rho)^r} \frac{\delta}{\delta x} [e^{-i\rho\mu x} x^{n-1}] &= 0 \\ \implies C(n, r, i) \sum_{r=0}^{i-1} \binom{i-1}{r} (-1)^r e^{-\mu(1-\rho)^r} e^{-i\rho\mu x} x^{n-2} (n-1 - i\rho\mu) &= 0 \end{aligned}$$

3.5. Joint distribution of two order statistics for waiting time distribution. Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ denote the order statistics of a random sample of size n from a continuous distribution with probability density function $\phi_X(x)$ and a cumulative density function $\Phi_X(x)$. Let us assume that $u = X_{i:n}$ and $v = X_{j:n}$ as i^{th} and j^{th} order statistics,

($1 < i < j < n$) from a random sample of size n , each with probability density function $\phi_X(x)$. The joint density function of $u = X_{i:n}$ and $v = X_{j:n}$ is as follows

$$\begin{aligned}
 \phi_w(u, v; n) &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [\Phi(x_r)]^{r-1} [\Phi(x_s) - \Phi(x_r)]^{j-i-1} \\
 &\quad [1 - \Phi(x_s)]^{n-j} \phi(x_r) \phi(x_s) \\
 &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [1 - \rho e^{-\mu(1-\rho)u}]^{i-1} \\
 &\quad [\rho e^{-\mu(1-\rho)u} - \rho e^{-\mu(1-\rho)v}]^{j-i-1} [\rho e^{-\mu(1-\rho)v}]^{n-j} \\
 (3.11) \quad &\quad e^{-\mu u} \frac{\mu^n u^{n-1}}{(n-1)!} e^{-\mu v} \frac{\mu^n v^{n-1}}{(n-1)!}
 \end{aligned}$$

Using Binomial expansion on (3.11), we get

$$\begin{aligned}
 &= C'(i, j; n) \gamma^2 \sum_{\alpha=0}^{j-i-1} \sum_{\beta=0}^{n-j} \binom{j-i-1}{\alpha} \binom{n-j}{\beta} (-1)^{\alpha+\beta} (\rho e^{-\mu(1-\rho)v})^{j-i-1+\alpha+\beta} \\
 &\quad [\rho e^{-\mu(1-\rho)u}]^{\alpha+i-1} e^{-\mu(u+v)} u^{n-1} v^{n-1}, \quad u, v > 0, \quad \mu > 0
 \end{aligned}$$

where

$$C'(i, j; n) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \text{ and } \gamma = \frac{\mu^n}{(n-1)!}$$

4. Derivation of distribution of sample range for waiting time distribution

Let the sample range of the waiting time distribution be defined as

$$R = X_{(n)} - X_{(1)}$$

Also, let

$$X_{(n)} = x \text{ and } X_{(1)} = y \Rightarrow u = x \text{ and } \Rightarrow v = y - u \Rightarrow y = u + v$$

$$J = \begin{vmatrix} \frac{\delta x}{\delta u} & \frac{\delta x}{\delta v} \\ \frac{\delta y}{\delta u} & \frac{\delta y}{\delta v} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1$$

Now using the joint distribution of order statistics

$$\begin{aligned}
 g(u, v) &= \frac{n!}{(1-1)!(n-1-1)!(n-n)!} \Phi^{1-1} [\Phi(u+v) - \Phi(u)]^{n-2} \\
 &\quad [1 - \Phi(u+v)]^{n-n} \phi(u) \phi(u+v) \\
 &= \frac{n!}{(n-2)!} [\Phi(u+v) - \Phi(u)]^{n-2} \phi(u) \phi(u+v)
 \end{aligned}$$

$$\begin{aligned}
g_1(v) &= \int g(u, v) du \\
&= \int_{-\infty}^{\infty} n(n-1) \Phi(u+v) - \Phi(u)]^{n-2} \phi(u) \phi(u+v) \\
&= \int_0^{\infty} n(n-1) \left[\rho [1 - e^{-\mu(1-\rho)(u+v)}] - \rho [1 - e^{-\mu(1-\rho)v}] \right]^{n-2} du \\
&\quad \mu\rho(1-\rho) e^{-\mu(1-\rho)u} \mu\rho(1-\rho) e^{-\mu(1-\rho)(u+v)} \\
&= \int_0^{\infty} n(n-1) \left[\rho [e^{-\mu(1-\rho)u} - e^{-\mu(1-\rho)(u+v)}] \right]^{n-2} \\
&\quad \mu\rho(1-\rho) e^{-\mu(1-\rho)(2u+v)} du \\
&= \int_0^{\infty} n(n-1) [e^{-\mu(1-\rho)u}]^{n-2} \left[\rho [1 - e^{-\mu(1-\rho)v}] \right]^{n-2} \\
&\quad \mu\rho(1-\rho) e^{-\mu(1-\rho)(2u+v)} du \\
&= n(n-1) \mu\rho(1-\rho) e^{-\mu(1-\rho)v} \left[\rho [1 - e^{-\mu(1-\rho)v}] \right]^{n-2} \\
&\quad \int_0^{\infty} [e^{-\mu(1-\rho)nu}] du \\
&= n(n-1) \mu\rho(1-\rho) e^{-\mu(1-\rho)v} \left[\rho [1 - e^{-\mu(1-\rho)v}] \right]^{n-2} \\
&\quad \left| \frac{e^{-\mu(1-\rho)nu}}{-\mu(1-\rho)n} \right|_0^{\infty} du \\
&= \rho^n (n-1) \rho(1-\rho) e^{-\mu(1-\rho)v} \left[\rho [1 - e^{-\mu(1-\rho)v}] \right]^{n-2}
\end{aligned}$$

Therefore the range of the distribution of the waiting time is

$$(4.1) \quad P(R = v) = \rho^n (n-1) \rho(1-\rho) e^{-\mu(1-\rho)v} \left[\rho [1 - e^{-\mu(1-\rho)v}] \right]^{n-2}, \quad -\infty < v < \infty$$

4.1. Theorem. Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ denote the order statistics of a random sample of size n from a continuous distribution with probability density function $\phi_X(x)$ and a cumulative density function $\Phi_X(x)$. Then the probability density function of the range of the waiting time distribution in the system is

$$(4.2) \quad P(R = s) = (n-1) \rho(1-\rho) e^{-\mu(1-\rho)v} \left[\rho [1 - e^{-\mu(1-\rho)v}] \right]^{n-2}, \quad -\infty < v < \infty$$

Proof. As stated as same in equation (4.1). □

5. Derivation of response time distribution

Percentiles are frequently used as indicators of performance in both the public and private hospitals. Percentiles provide information about how a patient or thing relates to a larger group. Relative measures of this type are often extremely valuable to researchers employing statistical techniques.

The formula of the mean response time is given by

Mean number in the system = Arrival rate \times mean response time *i.e.*,

$$(5.1) \quad \begin{aligned} E(n) &= \lambda E(r) \\ \implies F(r) &= \frac{E(n)}{\lambda} = \left(\frac{\rho}{1-\rho} \right) \frac{1}{\lambda} = \frac{\frac{1}{\mu}}{1-\rho} \end{aligned}$$

The cumulative distribution function of the response time is given as

$$(5.2) \quad F(r) = 1 - e^{-r\mu(1-\rho)}$$

The response time is exponentially distributed and q^{th} percentile is

$$(5.3) \quad \begin{aligned} F(r) &= \frac{q}{100} \implies 1 - e^{-r\mu(1-\rho)} = \frac{q}{100} \\ \implies r_q &= \frac{1}{\mu(1-\rho)} \ln \left(\frac{100}{100-q} \right) \end{aligned}$$

The cumulative distribution function of the waiting time is

$$(5.4) \quad F(w) = 1 - \rho e^{-w\mu(1-\rho)}$$

This is a truncated exponential distribution and its q^{th} percentile is given by

$$(5.5) \quad w_q = \frac{1}{\mu(1-\rho)} \ln \left(\frac{100\rho}{100-q} \right)$$

The above formula is applied only if q is greater than $100(1-\rho)$ and all lower percentiles are zero.

$$(5.6) \quad w_q = \max \left\{ 0, \frac{E(w)}{\rho} \ln \left(\frac{100\rho}{100-q} \right) \right\}$$

6. Distribution of sample median

When the sample size is odd, then the probability density function of the sample median is given by

$$\begin{aligned} \phi(x) &= \frac{(2n+1)}{m!^2} [\Phi(z)]^n [1-\Phi(z)]^n \phi(z) \\ &= \frac{(2n+1)}{n!^2} [1 - \rho e^{-\mu(1-\rho)z}]^n [\rho e^{-\mu(1-\rho)z}]^n e^{-\mu z} \frac{\mu^n z^{n-1}}{(n-1)!} \\ &= C(n, \gamma) \sum_{s=0}^r (-1)^s e^{-\mu(s+r)} z^{s+r-1} \end{aligned}$$

When the sample size is even, then the probability density function of the sample median is as follows

$$(6.1) \quad \begin{aligned} \phi(x, y) &= \frac{(2n)}{(n-1)!} [\Phi(x)]^{n-1} [1-\Phi(y)]^{n-1} \phi(x)\phi(y) \\ &= \frac{(2n)}{(n-1)!} [1 - \rho e^{-\mu(1-\rho)x}]^{n-1} [\rho e^{-\mu(1-\rho)y}]^{n-1} e^{-\mu x} \frac{\mu^n x^{n-1}}{(n-1)!} e^{-\mu y} \frac{\mu^n y^{n-1}}{(n-1)!} \\ &= \frac{(2n)}{(n-1)!} [1 - \rho e^{-\mu(1-\rho)x}]^{n-1} [\rho e^{-\mu(1-\rho)y}]^{n-1} e^{-\mu(x+y)} \frac{\mu^n x^{n-1}}{(n-1)!} \frac{\mu^n y^{n-1}}{(n-1)!} \\ &= C(n, \gamma) \sum_{s=0}^r (-1)^s e^{-\mu(x+y)} (xy)^{n-1} \end{aligned}$$

7. Results

Here we evaluate the minimum and maximum waiting time of the patients who are in the queue. The table below gives the minimum number of patients in the system and queue for given number of servers during each interval and it is clear that for both the queue and the system, the waiting time drops measurably from 1st to 5th server, after which the drop is trivial. Hence it is concluded that the maximum and minimum number of patients decrease gradually with the increasing number of servers.

Table 1. The maximum and minimum number of patients served and waiting in the system

<i>Server</i>	$X_n(W)$	$X_1(W_q)$	R
1	5	4.16667	0.83333
2	2.72727	2.27273	0.1388
3	1.82817	1.37363	0.023148
4	1.37125	0.931446	0.003858
5	1.09697	0.671899	0.000643
6	0.914135	0.168631	0.0001071
7	0.783545	0.108038	2.97687E-6
8	0.685602	0.0704791	4.96145E-7
9	0.609424	0.0465906	8.26908E-8
10	0.548481	0.0311082	1.37818E-8

The percentage points of response and waiting times in the system have been presented in Table(2), Table(3), Table(4) and Table(5). From the tables of percentage points of response time in the system, it is clear that for fixed values of p ($p = 0.25, 0.50, 0.75, 0.90, 0.95, 0.99$) and μ . Percentage points remain same as μ increases. On the other hand percentage points decreases as μ increases. Further, the range of smallest and largest order statistics of waiting time distribution and has been presented in Table(1). For chosen values of the parameters and n different values of r_q have been obtained for different significant levels.

Table 2. This table contains percentage of response time in the system when $\mu = 1$

λ	\mathbf{q}					
	0.25	0.50	0.75	0.90	0.95	0.99
2	0.00250	0.00501	0.00753	0.00904	0.00955	0.00995
3	0.00125	0.00251	0.00376	0.00452	0.00477	0.00497
4	0.00083	0.00167	0.00251	0.00301	0.00318	0.00332
5	0.00063	0.00125	0.00188	0.00226	0.00239	0.00249
6	0.00050	0.00100	0.00151	0.00181	0.00191	0.00199
7	0.00042	0.00084	0.00125	0.00151	0.00159	0.00166
8	0.00036	0.00072	0.00108	0.00129	0.00136	0.00142
9	0.00031	0.00063	0.00094	0.00113	0.00119	0.00124
10	0.00028	0.00056	0.00084	0.00100	0.00106	0.00111
11	0.00025	0.00050	0.00075	0.00090	0.00095	0.00099
12	0.00023	0.00046	0.00068	0.00082	0.00087	0.00090
13	0.00021	0.00042	0.0006	3 0.00075	0.00080	0.00083
14	0.00019	0.00039	0.00058	0.00070	0.00073	0.00077
15	0.00018	0.00036	0.00054	0.00065	0.00068	0.00071
16	0.00017	0.00033	0.00050	0.00060	0.00064	0.00066
17	0.00016	0.00031	0.00047	0.00057	0.00060	0.00062
18	0.00015	0.00029	0.00044	0.00053	0.00056	0.00059
19	0.00014	0.00028	0.00042	0.00050	0.00053	0.00055
20	0.00013	0.00026	0.00040	0.00048	0.00050	0.00052

Table 3. This table contains percentage of response time in the system when $\mu = 2$

λ	\mathbf{q}					
	0.25	0.50	0.75	0.90	0.95	0.99
3	0.00250	0.00501	0.00753	0.00904	0.00955	0.00995
4	0.00125	0.00251	0.00376	0.00452	0.00477	0.00497
5	0.00083	0.00167	0.00251	0.00301	0.00318	0.00332
6	0.00063	0.00125	0.00188	0.00226	0.00239	0.00249
7	0.00050	0.00100	0.00151	0.00181	0.00191	0.00199
8	0.00042	0.00084	0.00125	0.00151	0.00159	0.00166
9	0.00036	0.00072	0.00108	0.00129	0.00136	0.00142
10	0.00031	0.00063	0.00094	0.00113	0.00119	0.00124
11	0.00028	0.00056	0.00084	0.00100	0.00106	0.00111
12	0.00025	0.00050	0.00075	0.00090	0.00095	0.00099
13	0.00023	0.00046	0.00068	0.00082	0.00087	0.00090
14	0.00021	0.00042	0.00063	0.00075	0.00080	0.00083
15	0.00019	0.00039	0.00058	0.00070	0.00073	0.00077
16	0.00018	0.00036	0.00054	0.00065	0.00068	0.00071
17	0.00017	0.00033	0.00050	0.00060	0.00064	0.00066
18	0.00016	0.00031	0.00047	0.00057	0.00060	0.00062
19	0.00015	0.00029	0.00044	0.00053	0.00056	0.00059
20	0.00014	0.00028	0.00042	0.00050	0.00053	0.00055
21	0.00013	0.00026	0.00040	0.00048	0.00050	0.00052

Table 4. This table contains percentage of waiting time in the system when $\mu = 1$

λ	q					
	0.25	0.50	0.75	0.90	0.95	0.99
2	0.690644	0.688135	0.685619	0.684106	0.683602	0.683198
3	0.548055	0.5468	0.545542	0.544786	0.544533	0.544331
4	0.461264	0.460427	0.459589	0.459085	0.458916	0.458782
5	0.401734	0.401106	0.400477	0.400099	0.399973	0.399872
6	0.357851	0.357349	0.356846	0.356544	0.356443	0.356362
7	0.323901	0.323483	0.323064	0.322812	0.322727	0.32266
8	0.296705	0.296347	0.295988	0.295772	0.295699	0.295642
9	0.27434	0.274027	0.273712	0.273523	0.27346	0.273409
10	0.255565	0.255286	0.255006	0.254838	0.254782	0.254737
11	0.239539	0.239288	0.239037	0.238885	0.238835	0.238795
12	0.225673	0.225445	0.225216	0.225079	0.225033	0.224996
13	0.213537	0.213328	0.213118	0.212992	0.21295	0.212917
14	0.202812	0.202619	0.202425	0.202309	0.20227	0.202239
15	0.193253	0.193074	0.192894	0.192786	0.19275	0.192721
16	0.184672	0.184505	0.184337	0.184237	0.184203	0.184176
17	0.176919	0.176763	0.176605	0.176511	0.176479	0.176454
18	0.169875	0.169727	0.169579	0.16949	0.16946	0.169437
19	0.163441	0.163301	0.163162	0.163078	0.16305	0.163027
20	0.157538	0.157406	0.157274	0.157194	0.157168	0.157146

Table 5. This table contains percentage of waiting time in the system when $\mu = 2$

λ	q					
	0.25	0.50	0.75	0.90	0.95	0.99
3	0.40296198	0.40045257	0.39793684	0.39642436	0.3959197	0.39551578
4	0.34532203	0.34406732	0.34280946	0.34205322	0.34180088	0.34159892
5	0.30459587	0.3037594	0.30292082	0.30241666	0.30224844	0.3021138
6	0.27402729	0.27339994	0.27277101	0.27239289	0.27226672	0.27216574
7	0.25005197	0.24955009	0.24904694	0.24874444	0.24864351	0.24856273
8	0.23063187	0.23021364	0.22979435	0.22954227	0.22945816	0.22939084
9	0.21451061	0.21415212	0.21379273	0.21357666	0.21350457	0.21344687
10	0.20086685	0.20055317	0.20023871	0.20004965	0.19998656	0.19993607
11	0.18913833	0.18885951	0.18857998	0.18841193	0.18835585	0.18831097
12	0.17892563	0.17867469	0.17842312	0.17827187	0.17822141	0.17818101
13	0.16993628	0.16970815	0.16947945	0.16934195	0.16929607	0.16925935
14	0.16195058	0.16174147	0.16153182	0.16140578	0.16136373	0.16133007
15	0.15479999	0.15460696	0.15441344	0.1542971	0.15425828	0.15422721
16	0.14835274	0.1481735	0.14799381	0.14788577	0.14784972	0.14782087
17	0.1425042	0.14233691	0.14216919	0.14206836	0.14203472	0.14200779
18	0.13717009	0.13701325	0.13685602	0.13676149	0.13672995	0.1367047
19	0.13228169	0.13213407	0.13198609	0.13189712	0.13186743	0.13184367
20	0.12778233	0.12764292	0.12750316	0.12741913	0.12739109	0.12736865
21	0.12362485	0.12349277	0.12336037	0.12328076	0.1232542	0.12323294

Table 6. This table contains percentage of waiting time in the system when $\mu = 3$

λ	q					
	0.25	0.50	0.75	0.90	0.95	0.99
4	0.28517894	0.28266953	0.28015381	0.27864133	0.27813666	0.27773274
5	0.25416125	0.25290654	0.25164868	0.25089244	0.25064011	0.25043815
6	0.23021468	0.22937821	0.22853964	0.22803548	0.22786726	0.22773262
7	0.21119868	0.21057133	0.2099424	0.20956428	0.20943811	0.20933713
8	0.19566522	0.19516334	0.1946602	0.1943577	0.19425677	0.19417598
9	0.18268486	0.18226662	0.18184734	0.18159526	0.18151115	0.18144383
10	0.17163852	0.17128004	0.17092065	0.17070458	0.17063248	0.17057478
11	0.16209748	0.16178381	0.16146934	0.16128028	0.16121272	0.16116671
12	0.15375458	0.15347576	0.15319623	0.15302818	0.15297211	0.15292723
13	0.14638339	0.14613245	0.14588088	0.14572963	0.14567917	0.14563877
14	0.1398129	0.13958477	0.13935607	0.13921857	0.13917269	0.13913597
15	0.13391123	0.13370211	0.13349247	0.13336643	0.13332437	0.13329072
17	0.12857487	0.12838184	0.12818832	0.12807198	0.12803316	0.12800208
18	0.12372128	0.12354204	0.12336234	0.12325431	0.12321826	0.12318941
19	0.11928376	0.11911646	0.11894875	0.11884791	0.11881427	0.11878734
20	0.11520772	0.11505088	0.11489365	0.11479912	0.11476758	0.11474233
21	0.11144805	0.11130044	0.11115245	0.11106348	0.1110338	0.11101004
22	0.10796706	0.10782764	0.10768788	0.10760386	0.10757582	0.10755338

8. Conclusion

The obtained results show that the expected value of the maximum and minimum number of patients decreases gradually with the increasing number of servers. Besides, it is mentioned that the system will be almost empty after the 10th server. That is, the patient will get the service as soon as she or he arrive and will leave the system before the next arrival. This solution accords with the fact that the service rate is greater than the arrival rate. The range of the waiting time R decreases gradually to zero as the number of servers increases.

Acknowledgement

We are sincerely thankful to UGC-BSR Scheme, Government of India for granting us the financial assistance to carry out this research work.

References

- [1] Aleem, M. *Contributions to the theory of order statistics and selection procedures for restricted families of probability distribution*. PhD thesis, Bahauddin Zakariya University, Multan, 1998.
- [2] Arnold, B., Balakrishnan, N., and Nagaraja, H. *A first course in order statistics*. Wiley, New York, 1992.
- [3] Artalejo, J. R., Economou, A., and Lopez-Herrero, M. J. (2007). *Algorithmic analysis of the maximum queue length in a busy period for the M/M/c retrieval queue*. *INFORMS Journal on Computing*, **19**(1), 121-126, 2007.
- [4] Asmussen, S. (1998). *Extreme value theory for queues via cycle maxima*. *Extremes*, **1**(2), 137-168, 1998.
- [5] Bagui, S. C. *CRC Handbook of Percentiles of Non-central T-distributions*. CRC Press, 1993.
- [6] Bhat, U. N. *An introduction to queueing theory: modeling and analysis in applications*. Birkhauser, 2015.

- [7] David, H. and Nagaraja, H. *Order Statistics*. Wiley Series in Probability and Statistics. Wiley, 2004.
- [8] Nadarajah, S. and Pal, M. *Explicit expressions for moments of gamma order statistics*. Bulletin of the Brazilian Mathematical Society, New Series, **39**(1), 45-60, 2008.
- [9] Dhar, S., Das, K. K. and Mahanta, L. B. *Comparative study of waiting and service costs of single and multiple server system: A case study on an outpatient department*. International Journal of Scientific Footprints, **3**(2), 18-30, 2014.
- [10] Shawky, A. and Bakoban, R. *Order statistics from exponentiated gamma distribution and associated inference*. Int. J. Contemp. Math. Sciences, **4**(2), 71-91, 2009.
- [11] White, J. S. *Tables of normal percentile points*. Journal of the American Statistical Association, **65**(330), 635-638, 1970.
- [12] Tarabia, AMK. *A new formula for the transient behaviour of a non-empty M/M/1/8 queue*, Applied mathematics and computation **132** (1), 1-10, 2002.
- [13] Mahanta, L. B., Das, K. K., and Dhar, S. *A queuing model for dealing with patients with severe disease*. Electronic Journal of Applied Statistical Analysis, **9**(2):362-370, 2016.
- [14] Serfozo, Richard F. *Extreme values of birth and death processes and queues*, Stochastic Processes and their Applications, **27**, 291-306, 1987.
- [15] Park, Y. S. *Asymptotic distributions of maximum queue lengths for M/G/1 and GI/M/1 systems*. Journal of the Korean Statistical Society, **24**(1), 19-29, 1994.
- [16] Dhar, S., Das, K. K. and Mahanta, L. B. *An infinite server queueing model with varying arrival and departures rates for health care system*. International Journal of Pure and Applied Mathematics **113**(5), 583-593, 2017.
- [17] Medhi, J. Stochastic models in queueing theory. Academic Press, 2002.

Strong uniform consistency of a kernel conditional quantile estimator for censored and associated data

Wafaa Djelladj* and Abdelkader Tatachak†‡

Abstract

In survival or reliability studies, it is common to deal with data which are not only incomplete but weakly dependent too. Random right-censoring and random left-truncation are two common forms of such data when they are neither independent nor strongly mixing but rather associated. In this paper, we focus on kernel estimation of the conditional quantile function of a strictly stationary associated random variable T given a d -dimensional vector of covariates X , under random right-censoring. As main results, we establish a strong uniform consistency rate for the estimator. Then the finite sample performance of the estimator is illustrated on a simulation study.

Keywords: Associated data, Censored data, Convergence rate, Quantile function, Strong consistency.

2000 AMS Classification: 60G05, 62G20

Received : 23.11.2017 *Accepted :* 28.06.2018 *Doi :* 10.15672/HJMS.2018.609

1. Introduction

Let $\{T_n, n \geq 1\}$ be a strictly stationary sequence of associated random variables (rv's) of interest having an unknown absolutely continuous distribution function (df) F_T . This variable can be considered as a lifetime under biomedical studies. The major characteristic of survival time is the incompleteness.

In survival analysis especially in medical studies, we meet random censorship models which are one of the fundamental assumptions in the theory of survival analysis. Random right censoring is a well-known phenomenon which may be present when observing lifetime data. The lifetime variable may not be completely observable if the patient is

*Lab. MSTD, Department of Probability and Statistics, USTHB, BP 32, El-Alia 16111, Algeria, Email: wjelladj@usthb.dz

†Lab. MSTD, Department of Probability and Statistics, USTHB, BP 32, El-Alia 16111, Algeria, Email: atatachak@usthb.dz

‡Corresponding Author.

still alive at the end of study or is dead for another reason or because of some departures of patients from the testing experimentation. Hence, the available data provide partial information. In this case, the variable of interest T is subject to right censoring by another non-negative rv C . In the sequel, we assume that the censoring lifetimes are independent and identically distributed (iid) and possess an unknown Lipschitz df G . We take in consideration the presence of a strictly stationary and associated covariate \mathbf{X} taking values in \mathbb{R}^d . Under this model, the observable sequence is $\{(Y_i, \delta_i, \mathbf{X}_i), 1 \leq i \leq n\}$, with $Y_i = \min(T_i, C_i)$, $\delta_i = \mathbb{1}_{\{T_i \leq C_i\}}$ and where $\mathbb{1}_A$ denotes the indicator function of the event A .

As usual with random censoring, we assume that the censoring times $\{C_i, 1 \leq i \leq n\}$ are independent of $\{(\mathbf{X}_i, T_i), 1 \leq i \leq n\}$. This means that the censoring mechanism does not depend on the occurring event. Such a condition ensures the identifiability of the model. It is well known that the conditional df $F(\cdot|\mathbf{x})$ of $(T|\mathbf{X} = \mathbf{x})$ is defined by

$$F(t|\mathbf{x}) = \frac{1}{l(\mathbf{x})} \int_{-\infty}^t f(\mathbf{x}, z) dz =: \frac{F_1(\mathbf{x}, t)}{l(\mathbf{x})},$$

where $f(\cdot, \cdot)$ is the joint probability density function (pdf) of (\mathbf{X}, T) , $l(\cdot)$ is the marginal pdf of \mathbf{X} and $F_1(\mathbf{x}, \cdot)$ is the first derivative of the joint df $F(\mathbf{x}, \cdot)$ with respect to \mathbf{x} . The conditional pdf will be denoted by $f(\cdot|\mathbf{x})$. Then, for all fixed $p \in (0, 1)$, the p -th conditional quantile of T given $\mathbf{X} = \mathbf{x}$ is defined by

$$(1.1) \quad \xi_p(\mathbf{x}) := \inf\{t, F(t|\mathbf{x}) \geq p\}.$$

Hence, to get a nonparametric conditional quantile estimator, we clearly have to estimate $F_1(\mathbf{x}, t)$ by the mean of an unbiased kernel estimator and $l(\mathbf{x})$ is estimated by the famous kernel type estimator.

There has been various researches relating to the quantile estimator in view of its interesting properties. The estimator under consideration is renowned for its good description of the data (see Chaudhuri *et al.* [6]) and attracted interest of several authors.

In the complete framework, Samanta [25] established the strong convergence and the asymptotic normality of the kernel conditional quantile in the iid case. Bhattacharya and Gangopadhyay [2] gave a Bahadur-type representation of the conditional quantile and asymptotic models. Moreover, Mehra *et al.* [16] and Xiang [27] gave the almost sure convergence of a kernel type conditional quantile estimator and its asymptotic normality. Honda [12] treated the uniform convergence and asymptotic normality of the conditional quantile using local polynomial fitting approach while Abberger [1] studied quantile smoothing in financial time series.

On the same subject matter and under censoring, Dabrowska [7] established a Bahadur type representation of the quantile regression estimator. Besides, Qin and Wu [24] stated the asymptotic normality of an estimator for a conditional quantile when some auxiliary information is available using the empirical likelihood method and a linear fitting.

The strong representation of the conditional quantile estimator under right censoring and strong mixing condition was stated by Ould Saïd and Sadki [22] while Ould Saïd [20] established its strong uniform convergence rate in the iid case. Recently, Liang and de Uña-Álvarez [15] assessed its strong uniform consistency and asymptotic normality in the α -mixing setting.

Two kinds of dependency are widely used in the literature: mixing (Doukhan [8]) and association (Esary *et al.* [8]). These two concepts are not completely dissociated (see Doukhan and Louhichi [9]). In fact, we can find sequences that are associated but not mixing, associated and mixing, and mixing but not associated. The main advantage of the concept of association compared to mixing is that the conditions of limit theorems are easier to verify: indeed, a covariance is much easier to compute than a mixing coefficient.

Recall that a set of finite family of rv's (T_1, \dots, T_n) are said to be associated if for all non-decreasing functions Ψ_1, Ψ_2

$$Cov(\Psi_1(T_1, \dots, T_n), \Psi_2(T_1, \dots, T_n)) \geq 0,$$

whenever the covariance exists. An infinite family of rv's is associated if any finite sub-family is a set of associated rv's and any independent sequence is associated. In classical statistical inference, the observed rv's of interest are generally assumed to be iid. However, it is more common to have dependent variables in some real life situations. Dependent variables are present in several backgrounds such as medicine, biology and social sciences. Associated rv's are of considerable interest when dealing with reliability problems, percolation theory and some models in statistical mechanics.

The notion of association was firstly introduced by Esary *et al.* [11] mainly for an application in reliability. For more details on the subject we refer the reader to the monographs by Bulinski and Shashkin [3], Oliveira [19] and Prakasa Rao [23].

As far as we know, the problem of drawing nonparametric inference about the conditional quantile function under associated-censored model is not available and this motivates the study we consider here. So, the present paper deals with the almost sure uniform convergence with a rate of the estimator defined in (2.4). The paper is structured as follows: the expression of the studied estimator is presented in Section 2. Section 3 gathers the needed assumptions with some comments. A Simulation study is given in Section 4 while the last section includes the proofs of the main and some auxiliary results.

2. Notations and estimators

Recall that in the complete data case (no censoring), the traditional kernel estimator of $F(t|\mathbf{x})$ is given by

$$(2.1) \quad F_n(t|\mathbf{x}) = \sum_{i=1}^n \omega_{in}(\mathbf{x}) \mathbb{1}_{\{Y_i \leq t\}},$$

where $\omega_{in}(\cdot)$ are measurable functions. These functions called weights were introduced by Nadaraya-Watson in the context of the kernel regression and defined by

$$\omega_{in}(\mathbf{x}) = \frac{K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_{n,1}}\right)}{\sum_{j=1}^n K_d\left(\frac{\mathbf{x} - \mathbf{X}_j}{h_{n,1}}\right)} = \frac{1}{nh_{n,1}^d} \frac{K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_{n,1}}\right)}{l_n(\mathbf{x})},$$

with the convention $0|0 = 0$. Here K_d is a kernel function on \mathbb{R}^d whereas $h_{n,1}$ is a positive sequence of bandwidths tending to 0 along with n and $l_n(\cdot)$ is the Parzen-Rosenblatt kernel estimator of $l(\cdot)$.

In the sequel, we will make use of the Inverse-Probability-of-Censoring Weighted (IPCW) idea of Koul *et al.* [14] to define the weights we will use after, that is

$$(2.2) \quad \omega_{in}(\mathbf{x}) = \frac{1}{nh_{n,1}^d} \frac{\delta_i}{G(Y_i)l_n(\mathbf{x})} K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_{n,1}}\right).$$

It is well known that under right censoring model, the classical empirical distribution does not estimate consistently the df's F_T and G . Therefore, Kaplan and Meier [13] proposed consistent estimators $F_{T,n}$ and G_n for F_T and G , respectively, defined by

$$F_{T,n}(t) = 1 - \prod_{i=1}^n \left[1 - \frac{\delta_{(i)}}{n - i + 1} \right] \mathbb{1}_{\{Y_{(i)} \leq t\}}$$

and

$$G_n(t) = 1 - \prod_{i=1}^n \left[1 - \frac{1 - \delta_{(i)}}{n - i + 1} \right] \mathbb{1}_{\{Y_{(i)} \leq t\}},$$

where $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$ are the order statistics of Y_1, Y_2, \dots, Y_n and $\delta_{(i)}$ is the concomitant of $Y_{(i)}$.

The Kaplan-Meier estimator was studied in depth by many authors. For more details we refer to Stute and Wang [26] for the iid case, Cai [4] under α -mixing condition and Cai and Roussas [5] in the association setting.

Recall that, using the weights defined in (2.2), Ould Saïd [20] established a strong uniform consistency rate for the estimator in (2.1) in the iid case and $d=1$. The smoothed version of $F_n(\cdot|\cdot)$, namely

$$(2.3) \quad F_n(t|\mathbf{x}) =: \frac{F_{1,n}(\mathbf{x}, t)}{l_n(\mathbf{x})} = \frac{\frac{1}{nh_{n,1}^d} \sum_{i=1}^n \frac{\delta_i}{\bar{G}_n(Y_i)} K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_{n,1}} \right) H \left(\frac{t - Y_i}{h_{n,2}} \right)}{\frac{1}{nh_{n,1}^d} \sum_{i=1}^n K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_{n,1}} \right)},$$

was also considered and studied (strong consistency and asymptotic normality) in the iid case by Ould Saïd and Sadki [21]. Here, the bandwidth $h_{n,2}$ is not necessarily equal to $h_{n,1}$ and they will be denoted by $h_1 := h_{n,1}$ and $h_2 := h_{n,2}$.

Note that the estimator in (2.3) is an adapted version of that of Yu and Jones [28] to the censoring case. Originally, this smooth estimate for complete data (without the IPCW $\frac{\delta_i}{\bar{G}_n(Y_i)}$), was proposed and discussed by the last authors mainly to avoid the crossing problem which occurs when using an indicator function instead of a continuous df.

It follows that, in view of (2.3), a natural estimator of (1.1) can be computed by

$$(2.4) \quad \xi_{p,n}(\mathbf{x}) = \inf\{t, F_n(t|\mathbf{x}) \geq p\}.$$

To argue our main results, the following auxiliary pseudo-estimator will be of a great benefit in proving our results

$$(2.5) \quad \tilde{F}_n(t|\mathbf{x}) =: \frac{\tilde{F}_{1,n}(\mathbf{x}, t)}{l_n(\mathbf{x})} = \frac{\frac{1}{nh_1^d} \sum_{i=1}^n \frac{\delta_i}{\bar{G}(Y_i)} K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_1} \right) H \left(\frac{t - Y_i}{h_2} \right)}{\frac{1}{nh_1^d} \sum_{i=1}^n K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_1} \right)}.$$

Note that (2.5) can not be computed since $\bar{G}(\cdot)$ is assumed to be unknown.

3. Assumptions and main results

In the sequel, c stands for a positive constant taking different values and τ will denote a positive real number satisfying $\tau < \tau_F < \tau_G$ where, for any df W , $\tau_W := \sup\{y; W(y) < 1\}$. Define $\Omega_0 = \{\mathbf{x} \in \mathbb{R}^d / l(\mathbf{x}) \geq m_0 := \inf_x l(\mathbf{x}) > 0\}$ and let Ω and \mathcal{C} be compact sets included in Ω_0 and $[0, \tau]$, respectively. The main results will be stated using the following assumptions:

A1. The bandwidths h_1 and h_2 satisfy

- (i) $h_1 \rightarrow 0$, $nh_1^{2\alpha+d(1-\alpha)} \rightarrow +\infty$ and $\frac{\log^5 n}{nh_1^d} \rightarrow 0$ as $n \rightarrow +\infty$,
- (ii) $h_2 \rightarrow 0$ and $nh_1^d h_2 \rightarrow +\infty$ as $n \rightarrow +\infty$;

A2. The kernel K_d is a bounded pdf, compactly supported and satisfies:

- (i) K_d is Hölder continuous of order $\alpha \in (0, 1)$,
- (ii) $\int_{\mathbb{R}^d} u_j K_d(\mathbf{u}) d\mathbf{u} = 0$, for all $j = 1, \dots, d$, where $\mathbf{u} = (u_1, \dots, u_d)^\top$;

A3. The function H in (2.3) is of class \mathcal{C}^1 . Furthermore, its derivative $H^{(1)}$ is assumed to be compactly supported and satisfies the properties of a second order kernel;

A4. The marginal density $l(\cdot)$ is bounded and twice differentiable with:

$$\sup_{\mathbf{x} \in \Omega} \left| \frac{\partial^k l(\mathbf{x})}{\partial x_i \partial x_j^{k-1}}(x) \right| < \infty \text{ for } i, j = 1, \dots, d \text{ and } k = 1, 2;$$

A5. The joint pdf $f(\cdot, \cdot)$ is bounded and twice continuously differentiable;

A6. The joint pdf $l_{i,j}(\cdot, \cdot)$ of $(\mathbf{X}_i, \mathbf{X}_j)$ is bounded;

A7. The joint pdf $f(\cdot, \cdot, \cdot, \cdot)$ of $(\mathbf{X}_i, Y_i, \mathbf{X}_j, Y_j)$ is bounded;

A8. Let us define Λ_{ij} as follows:

$$\Lambda_{ij} := \sum_{k=1}^d \sum_{l=1}^d \text{Cov}(X_i^k, X_j^l) + 2 \sum_{k=1}^d \text{Cov}(X_i^k, Y_j) + \text{Cov}(Y_i, Y_j),$$

with X_i^k the k -th component of \mathbf{X}_i , such that for all $j \geq 1$ and $r > 0$

$$\sup_{i: |j-i| \geq r} \Lambda_{ij} =: \rho(r) \leq \gamma_0 e^{-\gamma r}, \text{ for all } \gamma_0, \gamma > 0;$$

A9. The function $\varsigma(\mathbf{x}) = \int_{\mathbb{R}} \frac{1}{G(v)} f(\mathbf{x}, v) dv$ is bounded, continuously differentiable

$$\text{and } \sup_{\mathbf{x} \in \Omega} \left| \frac{\partial \varsigma}{\partial x_i}(\mathbf{x}) \right| < \infty \text{ for } i = 1, \dots, d.$$

3.1. Remark. Assumption **A1** gives a classical choice of the bandwidths in functional estimation. For the sake of simplicity, many authors consider that $h_1 = h_2$ which is not justified in general. Note that the condition **A1** (ii) implies the first condition in **A1** (i) if $d \geq 2$. For $d = 1$, the comparison is not straightforward and depends upon the order of magnitude of h_2 with respect to h_1^α . Assumption **A2** is quite usual in kernel estimation. Assumptions **A3-A7** are classical in nonparametric estimation under dependency while **A8** is used for covariance calculation under association structure. Furthermore, this assumption gives a progressive trend to asymptotic independence of "past" and "future". Finally, Assumption **A9** is mainly technical.

The first result establishes the rate of convergence of the fluctuation term, that is $\left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] \right|$. This will be done by applying a Bernstein-type inequality stated by Doukhan and Neumann [10] for weakly dependent rv's. The next result in Theorem 3.3 states a uniform almost sure convergence rate of $F_n(t|x)$ toward $F(t|x)$, which will be stated with the help of Theorem 3.2. Then, as an immediate result, the asymptotic behaviour of the kernel conditional quantile estimator will be deduced as presented in Corollary 3.4.

3.2. Theorem. *Suppose that assumptions **A1-A5** and **A7-A9** hold and for n large enough, we have*

$$(3.1) \quad \sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] \right| = O \left(\sqrt{\frac{\log n}{nh_1^d}} \right), \text{ a.s.}$$

3.3. Theorem. *Under the assumptions of Theorem 3.2 and **A6**, for n large enough we have*

$$\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} |F_n(t|\mathbf{x}) - F(t|\mathbf{x})| = O \left\{ (h_1^2 + h_2^2) + n^{-\theta} + \sqrt{\frac{\log n}{nh_1^d}} \right\}, \text{ a.s.}$$

with $0 < \theta < \gamma/(2\gamma + 9 + 3/2\kappa)$ for any $\kappa > 0$.

3.4. Corollary. Under the assumptions of Theorem 3.3, and for each fixed $p \in (0, 1)$ and $x \in \Omega$, if $\inf_{\mathbf{x} \in \Omega} f(\xi_p(\mathbf{x})|\mathbf{x}) > 0$, then for n large enough, we have

$$\sup_{\mathbf{x} \in \Omega} |\xi_{p,n}(\mathbf{x}) - \xi_p(\mathbf{x})| = O \left\{ (h_1^2 + h_2^2) + n^{-\theta} + \sqrt{\frac{\log n}{nh_1^d}} \right\}, \text{ a.s.}$$

3.5. Remark. The uniform positiveness condition on the conditional density in Corollary 3.4 ensures the uniform uniqueness of the conditional quantile. Hence $\forall \varepsilon > 0, \exists \beta > 0, \forall \eta_p : \Omega \rightarrow \mathbb{R}$,

$$\sup_{\mathbf{x} \in \Omega} |\xi_p(\mathbf{x}) - \eta_p(\mathbf{x})| \geq \varepsilon \Rightarrow \sup_{\mathbf{x} \in \Omega} |F(\xi_p(\mathbf{x})|\mathbf{x}) - F(\eta_p(\mathbf{x})|\mathbf{x})| \geq \beta.$$

3.6. Remark. We point out that the rate in Corollary 3.4 depends upon the parameter θ pertaining to the association dependence. In addition, remark that for γ large enough, the parameter θ approaches its upper bound ($\theta=1/2$) and then, the covariances become negligible which in turn permits to compare our rate with those stated in the iid and strong mixing cases.

4. Simulation study

4.1. Description of the models. This part is established with the intention of giving the behaviour of the conditional quantile estimator. For this purpose, we only consider the cases of the conditional mean ($p = 1/2$) and the one dimensional covariate ($d = 1$). The simulation is conducted for different sample sizes and censoring rates (CR). The performance of our estimator is quantified via the Global Mean Square Error (GMSE). The simulated data are obtained as follows:

- Generate $(n + 1)$ iid rv's Z_i from gamma distribution ($Z_i \sim \Gamma(5, 0.5)$);
- Generate n iid rv's ε_i from normal distribution ($\varepsilon_i \sim \mathcal{N}(0, 0.01)$);
- Given Z_i , generate an n associated sequence (X_i, T_i) as follows:
 - a) Linear case

$$\begin{cases} X_i = \exp(Z_{i-1} + Z_{i-2})/2; \\ T_i = 3X_i/2 + 0.45 \varepsilon_i. \end{cases}$$
 - b) Nonlinear case

$$\begin{cases} X_i = \exp(Z_{i-1} + Z_{i-2})/2, \\ T_i = \log(3X_i/2) + 0.45 \varepsilon_i, \end{cases}$$
- Generate n iid rv's C_i from exponential distribution ($C_i \sim \exp(\lambda)$). The parameter λ is adjusted according to the CR 's values;
- Keep the observed data $\{(Y_i := \min\{T_i, C_i\}), X_i, (\delta_i := \mathbb{1}_{\{T_i \leq C_i\}})\}$.

4.1. Remark. In computing the estimators, we use the standardized normal df and a Gaussian kernel for H and K , respectively.

In order to attenuate the boundary effect, we will use optimal local bandwidths. To do so, we first assume that $h_1 = h_2 =: h$, and this bandwidth sweeps the interval $[0.05, 0.8]$. For each model, the process above is repeated $B = 300$ times with fixed values of n and CR . Thus, we compute the conditional quantile estimator along a grid of points in

[1.5, 4]. At the end of the process, we keep the optimal local bandwidth which minimizes the estimating errors by means of the MSE (Mean Square Error) criterion, and then we quantify the $GMSE$. The formula calculating the $GMSE$ is

$$GMSE = \frac{1}{uB} \sum_{\ell=1}^u \sum_{k=1}^B [\xi_{p,n,k}(x_\ell) - \xi_p(x_\ell)]^2,$$

where $\xi_{p,n,k}(x_\ell)$ is the value of $\xi_{p,n}(x_\ell)$ at iteration k and u is the number of equidistant points x_ℓ belonging to [1.5, 4].

To illustrate visually the quality of fit, we will plot the conditional quantile estimator $\xi_{p,n}(x_\ell)$ versus $\xi_p(x_\ell)$.

4.2. Simulation results.

4.2.1. Linear case: Note that under this model, the rv X follows $\Gamma(10, 0.5)$ and the conditional rv $(T|X = x)$ follows $\mathcal{N}(3x/2, 0.0045)$.

To show how is the influence of the censoring rate and the sample size on the quality of fit, we draw curves for different sample sizes $n = 50, 100$ and 300 and $CR = 40\%, 25\%$ and 10% as illustrated by Figures 1, 2 and 3. The corresponding errors with respect to the $GMSE$ are summarized in Table 1.

Table 1. Values of $GMSE$ for $\xi_{p,n}$ with $p = 0.5$

Linear case	$n = 50$	$n = 100$	$n = 300$
$CR = 10\%$	0.0637	0.0245	0.0069
$CR = 25\%$	0.1591	0.0586	0.0113
$CR = 40\%$	0.2465	0.1059	0.0128

4.2. Remark. From Table 1 and the graphs plotted for the linear case, we remark that the quality of fit seems to increase when the CR decreases. The curves reveal also that boundary effects on the right side tend to diminish for large values of n . Of course, the performance is quite acceptable when $n = 50$ and becomes more visible for $n = 300$. It means that the influence of the CR on the quality of fit becomes more and more insignificant along with n .

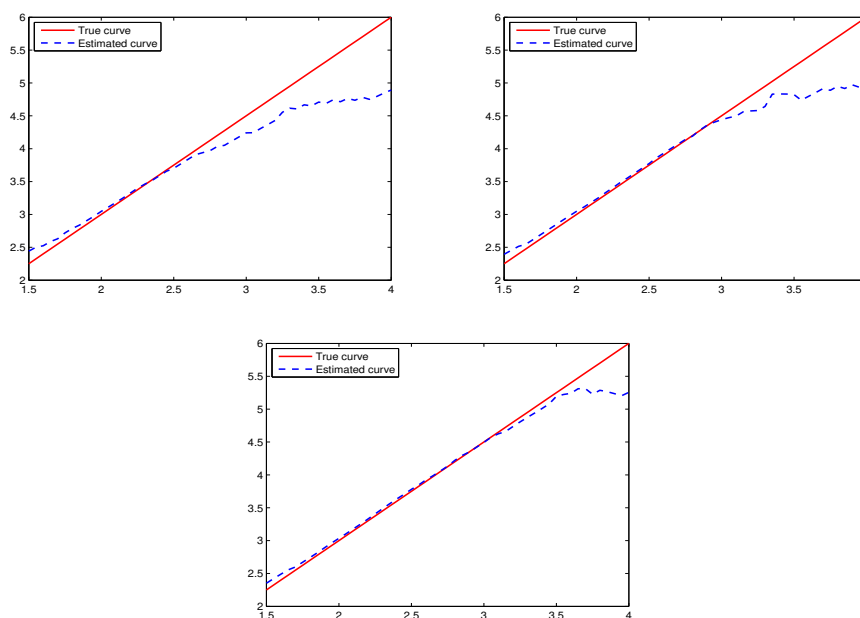


Figure 1. Linear case: $n = 50$ and $CR = 40, 25$ and 10 , respectively

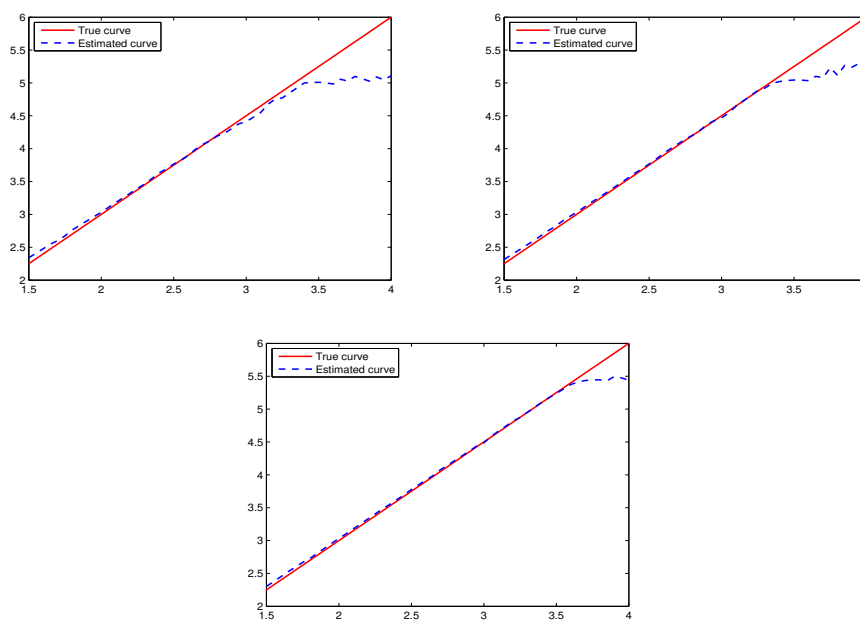


Figure 2. Linear case: $n = 100$ and $CR = 40, 25$ and 10 , respectively

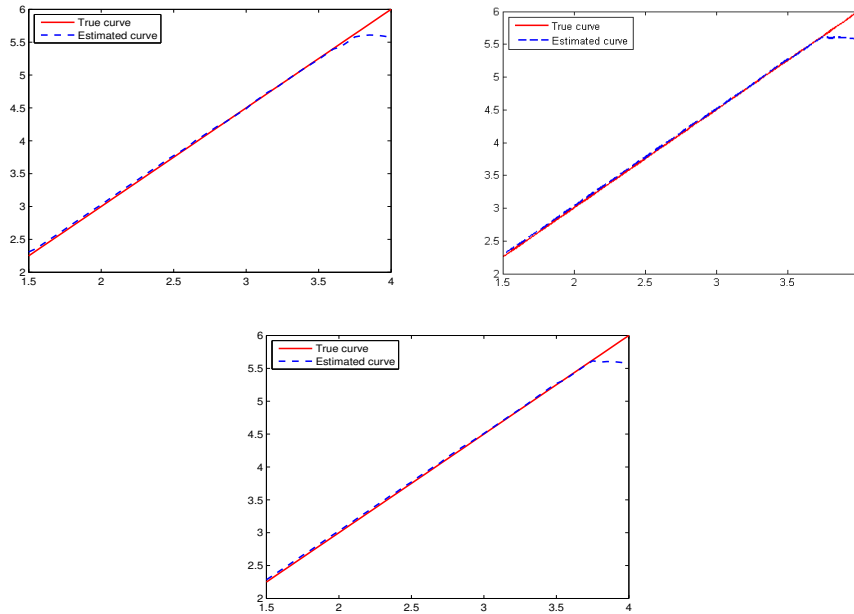


Figure 3. Linear case: $n = 300$ and $CR = 40, 25$ and 10 , respectively

4.2.2. Non-linear case: Note that the rv $(T|X = x)$ follows $\mathcal{N}(\log(3x/2), 0.0045)$ and the choice of the log function permits to preserve the association property by monotonicity.

For the rest we proceed as for the linear case. The $GMSE$'s are summarized in Table 2 and the quality of fit is illustrated through Figures 4,5 and 6.

Table 2. Values of $GMSE$ for $\xi_{p,n}(\cdot)$ with $p = 0.5$

Non-linear case	$n = 50$	$n = 100$	$n = 300$
$CR = 10\%$	24×10^{-3}	15×10^{-3}	5.54×10^{-4}
$CR = 25\%$	69×10^{-3}	25×10^{-3}	8.23×10^{-4}
$CR = 40\%$	11×10^{-2}	51×10^{-3}	16×10^{-3}

4.3. Remark. From Table 2 and the graphs, we observe that the estimator behaves similarly as for the linear case. The quality of fit becomes better along with the sample size which means that the behavior of the estimator remains correct even for large values of CR .

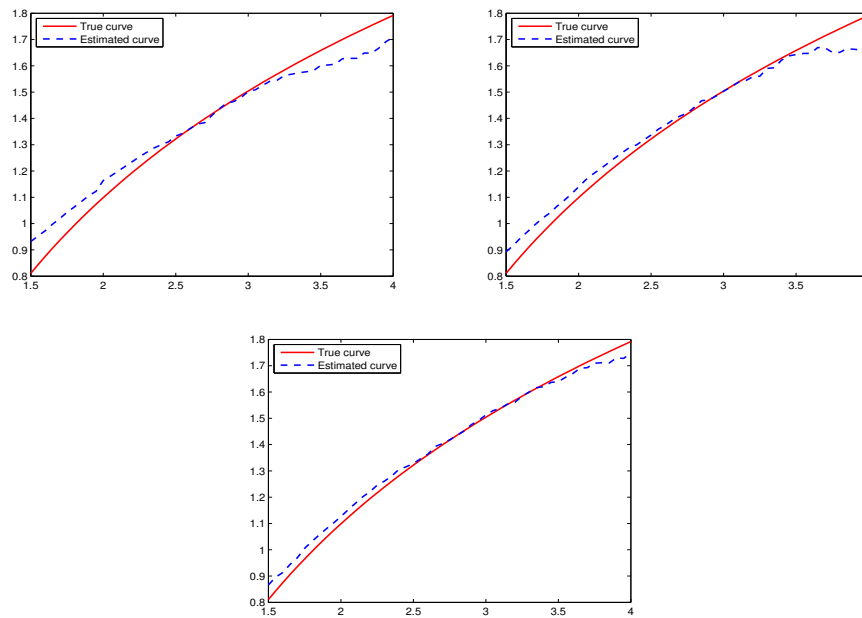


Figure 4. Non linear case: $n = 50$ and $CR = 40, 25$ and 10 , respectively

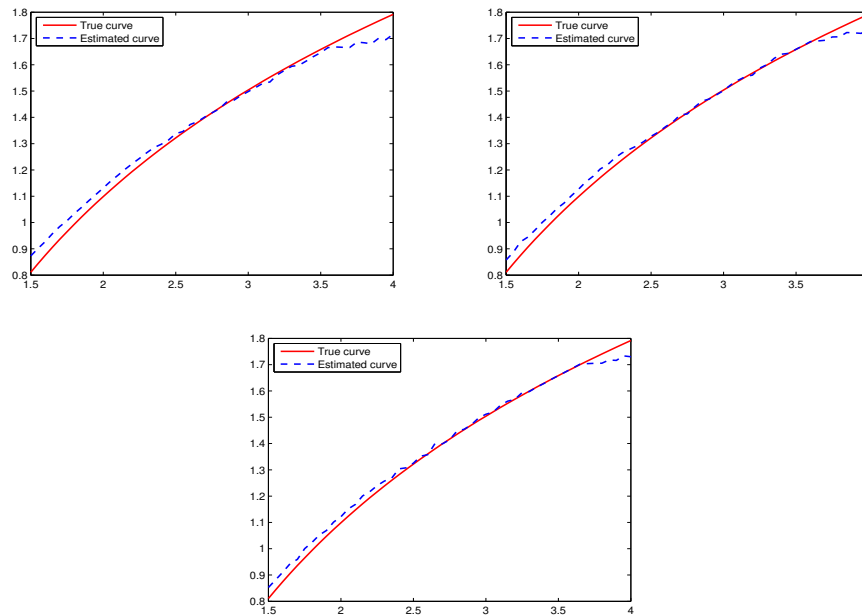


Figure 5. Non linear case: $n = 100$ and $CR = 40, 25$ and 10 , respectively

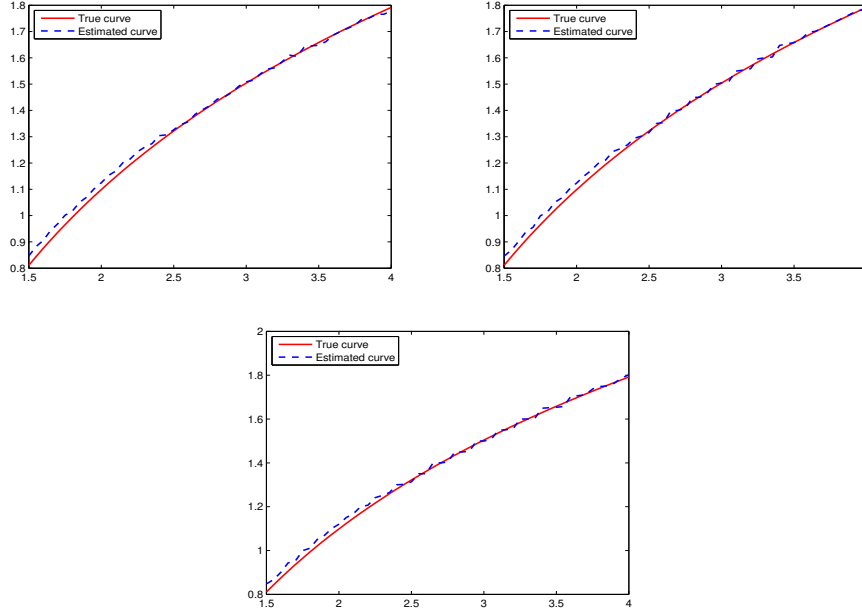


Figure 6. Non linear case: $n = 300$ and $CR = 40, 25$ and 10 , respectively

5. Auxiliary results and proofs

For notational convenience, let us define

$$\begin{aligned} \Delta_i(\mathbf{x}, t) &= \frac{\delta_i}{\bar{G}(Y_i)} K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_1} \right) H \left(\frac{t - Y_i}{h_2} \right) \\ &- \mathbb{E} \left[\frac{\delta_1}{\bar{G}(Y_1)} K_d \left(\frac{\mathbf{x} - \mathbf{X}_1}{h_1} \right) H \left(\frac{t - Y_1}{h_2} \right) \right], \end{aligned}$$

for all $i = 1, \dots, n$. It is easily seen that

$$(5.1) \quad \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] = \frac{1}{nh_1^d} \sum_{i=1}^n \Delta_i(\mathbf{x}, t).$$

The items of the following proposition are similar to the conditions of Theorem 1 in Doukhan and Neumann [10]. Once the conditions are met, it becomes possible to use an exponential inequality to prove Theorem 3.2 related to the fluctuation term.

5.1. Proposition. *Let $\Delta_1(\mathbf{x}, t), \Delta_2(\mathbf{x}, t), \dots, \Delta_n(\mathbf{x}, t)$ be defined as above. Then, there exist constants $M, L_1, L_2, \mu \geq 0, \lambda \geq 0$ and a non-decreasing sequence of real coefficients $(\Upsilon(n))_{n \geq 0}$ so that for all p -tuples (s_1, \dots, s_p) and all q -tuples (v_1, \dots, v_q) with $1 \leq s_1 \leq \dots \leq s_p \leq v_1 \leq \dots \leq v_q \leq n$, we have*

$$\text{a) } \text{Cov} \left(\prod_{i=s_1}^{s_p} \Delta_i(\mathbf{x}, t), \prod_{j=v_1}^{v_q} \Delta_j(\mathbf{x}, t) \right) \leq c^{p+q} h_1^d h_2^{\frac{2}{d+1}} pq \Upsilon(v_1 - s_p),$$

$$\text{b) } \sum_{s=0}^{\infty} (s+1)^{k_0} \Upsilon(s) \leq L_1 L_2^{k_0} (k_0!)^\mu, \forall k_0 \geq 0,$$

$$c) \mathbb{E} [|\Delta_i(\mathbf{x}, t)|^{k_0}] \leq (k_0!)^\lambda M^{k_0}.$$

Proof. Proof of Proposition 5.1 To prove the first item of Proposition 5.1, we need the following lemma:

5.2. Lemma. *Under assumptions A2, A5, A7 and A8, we have*

$$\begin{aligned} \text{i) } Cov \left(\prod_{i=s_1}^{s_p} \Delta_i(\mathbf{x}, t), \prod_{j=v_1}^{v_q} \Delta_j(\mathbf{x}, t) \right) &=: C_1 \leq c^{p+q} h_1^{-2} h_2^{-2} p q \rho (v_1 - s_p), \\ \text{ii) } Cov \left(\prod_{i=s_1}^{s_p} \Delta_i(\mathbf{x}, t), \prod_{j=v_1}^{v_q} \Delta_j(\mathbf{x}, t) \right) &=: C_2 \leq c^{p+q} h_1^{2d} h_2^2. \end{aligned}$$

Proof. Exploiting the definition 5.1, p.88 in Bulinski & Shashkin [3], we recall that the partial Lipschitz constants are defined as follows

$$(5.2) \quad Lip_i(\Phi_m) = \sup_{\substack{z_1, \dots, z_m \\ z_i \neq z'_i, z'_i \in \mathbb{R}}} \frac{|\Phi_m(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_m) - \Phi_m(z_1, \dots, z_{i-1}, z'_i, z_{i+1}, \dots, z_m)|}{|z_i - z'_i|},$$

where $\Phi_m : \mathbb{R}^m \rightarrow \mathbb{R}$ and $Lip(\Phi_m)$ denotes the Lipschitz continuity modulus of Φ_m , viz

$$Lip(\Phi_m) = \sup_{\mathbf{x} \neq \mathbf{y}} \frac{|\Phi_m(\mathbf{x}) - \Phi_m(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|_1},$$

with $\|(z_1, \dots, z_n)\|_1 = |z_1| + \dots + |z_n|$.

To prove part (i) in Lemma 5.2, we use Theorem 5.3, p.89 in (Bulinski and Shashkin [3]). Firstly, we set

$$\Phi_p =: \prod_{i=s_1}^{s_p} \Delta_i \quad \text{and} \quad \Phi_q =: \prod_{j=v_1}^{v_q} \Delta_j.$$

Then, using the fact that K_d , H and G are Lipschitz functions, we have

$$Cov(\Phi_p, \Phi_q) \leq \sum_{i=s_1}^{s_p} \sum_{j=v_1}^{v_q} Lip_i(\Phi_p) Lip_j(\Phi_q) \Lambda_{ij},$$

The definition in (5.2) leads to

$$Lip_i(\Phi_p) \leq \frac{M_0}{h_1 h_2} \left(\frac{2}{\overline{G}(\tau)} \right)^p \|K_d\|_\infty^{p-1}$$

and

$$Lip_j(\Phi_q) \leq \frac{M_0}{h_1 h_2} \left(\frac{2}{\overline{G}(\tau)} \right)^q \|K_d\|_\infty^{q-1},$$

where $M_0 = \max \left\{ h_2 Lip(K) \|K\|_\infty^{d-1}, h_1 \left(Lip(H) + h_2 \frac{Lip(\overline{G})}{\overline{G}(\tau)} \right) \|K_d\|_\infty \right\}$.

Note that the partial Lipschitz constants are obtained as follows

$$\begin{aligned} Lip_i(\Phi_p) &\leq \frac{M_0}{h_1 h_2} \left(\frac{2}{\overline{G}(\tau)} \right)^{p-1} \|K_d\|_\infty^{p-1} \frac{1}{\overline{G}(\tau)} \\ &\leq \frac{M_0}{h_1 h_2} \left(\frac{2}{\overline{G}(\tau)} \right)^p \|K_d\|_\infty^{p-1}. \end{aligned}$$

If Assumption **A8** holds, by stationarity we get

$$\begin{aligned} \text{Cov}(\Phi_p, \Phi_q) &\leq \frac{M_0^2}{h_1^2 h_2^2} \left(\frac{2}{\bar{G}(\tau)} \right)^{p+q} \|K_d\|_\infty^{p+q-2} \sum_{i=s_1}^{s_p} \sum_{j=v_1}^{v_q} \Lambda_{ij} \\ &\leq \frac{c^{p+q}}{h_1^2 h_2^2} pq \rho(v_1 - s_p). \end{aligned}$$

This achieves the proof of (i). In order to prove the second part of Lemma 5.2, we need to calculate the covariance term as shown hereafter by using the fact that

$$\mathbb{E}[\delta_i \delta_j | T_i, T_j] = \mathbb{E} \left[\mathbb{1}_{\{T_i \leq C_i\}} \mathbb{1}_{\{T_j \leq C_j\}} | T_i, T_j \right] = \bar{G}(T_i) \bar{G}(T_j).$$

We also use the following simplified notations

$$K_{d,\mathbf{x},i} := K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_1} \right) \quad \text{and} \quad H_{t,i} := H \left(\frac{t - Y_i}{h_2} \right).$$

Indeed, we have

$$\begin{aligned} &\text{Cov}(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t)) \\ &= \mathbb{E} \left[\frac{\delta_i \delta_j}{\bar{G}(Y_i) \bar{G}(Y_j)} K_{d,\mathbf{x},i} H_{t,i} K_{d,\mathbf{x},j} H_{t,j} \right] \\ &\quad - \mathbb{E} \left[\frac{\delta_i}{\bar{G}(Y_i)} K_{d,\mathbf{x},i} H_{t,i} \right] \times \mathbb{E} \left[\frac{\delta_j}{\bar{G}(Y_j)} K_{d,\mathbf{x},j} H_{t,j} \right] \\ &= \mathbb{E} \left[K_{d,\mathbf{x},i} K_{d,\mathbf{x},j} \mathbb{E} \left(\frac{\delta_i \delta_j}{\bar{G}(Y_i) \bar{G}(Y_j)} H_{t,i} H_{t,j} | \mathbf{X}_i, \mathbf{X}_j \right) \right] \\ &\quad - \mathbb{E} \left[K_{d,\mathbf{x},i} \mathbb{E} \left(\frac{\delta_i}{\bar{G}(Y_i)} H_{t,i} | \mathbf{X}_i \right) \right] \times \mathbb{E} \left[K_{d,\mathbf{x},j} \mathbb{E} \left(\frac{\delta_j}{\bar{G}(Y_j)} H_{t,j} | \mathbf{X}_j \right) \right] \\ &= \mathbb{E} \left[K_{d,\mathbf{x},i} K_{d,\mathbf{x},j} \mathbb{E} \left(H_{t,i} H_{t,j} \frac{\mathbb{E}[\delta_i \delta_j | T_i, T_j]}{\bar{G}(T_i) \bar{G}(T_j)} | \mathbf{X}_i, \mathbf{X}_j \right) \right] \\ &\quad - \mathbb{E} \left[K_{d,\mathbf{x},i} \mathbb{E} \left(H_{t,i} \frac{\mathbb{E}[\delta_i | T_i] | \mathbf{X}_i}{\bar{G}(T_i)} \right) \right] \mathbb{E} \left[K_{d,\mathbf{x},j} \mathbb{E} \left(H_{t,j} \frac{\mathbb{E}[\delta_j | T_j] | \mathbf{X}_j}{\bar{G}(T_j)} \right) \right]. \end{aligned}$$

Then, we get

$$\begin{aligned} &|\text{Cov}(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t))| \\ &\leq \left| \int_{\mathbb{R}^{2d} \times \mathbb{R}^2} K_d \left(\frac{\mathbf{x} - \mathbf{u}}{h_1} \right) H \left(\frac{t - s}{h_2} \right) K_d \left(\frac{\mathbf{x} - \mathbf{r}}{h_1} \right) H \left(\frac{t - v}{h_2} \right) \right. \\ &\quad \times f(\mathbf{u}, s, \mathbf{r}, v) \mathbf{d}\mathbf{u} \mathbf{d}s \mathbf{d}\mathbf{r} \mathbf{d}v \left. \right| \\ &\quad + \left| \int_{\mathbb{R}^{d+1}} K_d \left(\frac{\mathbf{x} - \mathbf{u}}{h_1} \right) H \left(\frac{t - s}{h_2} \right) f(\mathbf{u}, s) \mathbf{d}\mathbf{u} \mathbf{d}s \right. \\ &\quad \times \left. \int_{\mathbb{R}^{d+1}} K_d \left(\frac{\mathbf{x} - \mathbf{r}}{h_1} \right) H \left(\frac{t - v}{h_2} \right) f(\mathbf{r}, v) \mathbf{d}\mathbf{r} \mathbf{d}v \right|. \end{aligned}$$

Moreover, under assumptions **A2**, **A5** and **A7**, using a change of variables we get

$$(5.3) \quad |\text{Cov}(\Delta_i(\mathbf{x}, t), \Delta_j(\mathbf{x}, t))| = O(h_1^{2d} h_2^2).$$

Finally, the second part of Lemma 5.2 follows by simple algebra. \square

We need some auxiliary notations to set up the proof of Proposition 5.1. Impose $\Upsilon(\cdot) = \rho^{\frac{d}{2d+2}}(\cdot)$ and use the upper bounds of Lemma 5.2, namely

$$(5.4) \quad C_1^{\frac{d}{2d+2}} \leq c^{\frac{(p+q)d}{2d+2}} h_1^{\frac{-2d}{2d+2}} h_2^{\frac{-2d}{2d+2}} (pq)^{\frac{d}{2d+2}} \rho^{\frac{d}{2d+2}} (v_1 - s_p),$$

$$(5.5) \quad C_2^{\frac{d+2}{2d+2}} \leq c^{(p+q)\frac{d+2}{2d+2}} h_1^{\frac{2d(d+2)}{2d+2}} h_2^{\frac{2(d+2)}{2d+2}}.$$

Combining (5.4) and (5.5), we get

$$\begin{aligned} C_1^{\frac{d}{2d+2}} C_2^{\frac{d+2}{2d+2}} &\leq c^{p+q} h_1^d h_2^{\frac{2}{d+1}} (pq)^{\frac{d}{2d+2}} \rho^{\frac{d}{2d+2}} (v_1 - s_p) \\ &\leq c^{p+q} h_1^d h_2^{\frac{2}{d+1}} pq \Upsilon(v_1 - s_p). \end{aligned}$$

This inequality concludes the proof of part (a) of Proposition 5.1. Next, under Assumption **A8** and choosing $\lambda = 0$, $\mu = 1$, $L_1 = L_2 = \frac{1}{1 - e^{-\frac{1}{2d+2}}}$, the proofs of the results in (b) and (c) are similar to those used in proving Proposition 8 in (Doukhan and Neumann [10]), then we omit them. The proof of Proposition 5.1 is complete. \square

Proof. Proof of Theorem 3.2 In order to set up the uniform asymptotic expression of the fluctuation term $\left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] \right|$, we apply the triangular inequality and classical techniques to cover compacts. So, Ω can be covered by a finite number $d_{x,n}$ of balls $B_k(\mathbf{x}_k, a_n^d)$ centred at $\mathbf{x}_k = (x_{k,1}, \dots, x_{k,d})$ and \mathcal{C} is split into $d_{t,n}$ subintervals $J_1, \dots, J_{d_{t,n}}$ of lengths b_n , centred at t_ℓ . In other words, for all $\mathbf{x} \in \Omega$, $t \in \mathcal{C}$, there exist integers $k \in \{1, \dots, d_{x,n}\}$ and $\ell \in \{1, \dots, d_{t,n}\}$ such that $\|\mathbf{x} - \mathbf{x}_k\| \leq a_n^d$ and $|t - t_\ell| \leq b_n$, with $a_n^d = (n^{-1} h_1^{2\alpha+d})^{1/2\alpha}$ and $b_n = (n h_1^d)^{-1/2} h_2$. Then, as Ω and \mathcal{C} are bounded, let m_1 and m_2 be positive constants satisfying $d_{x,n} a_n^d \leq m_1$ and $d_{t,n} b_n \leq m_2$.

5.3. Remark. In proving our results we will use Lemma 5.4 stated in Menni and Tatchak [17] (see their Lemma 3) which governs a strong uniform consistency rate of the kernel estimator $l_n(\cdot)$. We recall it hereinafter without proof.

5.4. Lemma. *Under assumptions **A1**, **A2**, **A4**, **A6** and **A8**, for n large enough we have*

$$\sup_{\mathbf{x} \in \Omega} |l_n(\mathbf{x}) - l(\mathbf{x})| = O \left(\max \left\{ \sqrt{\frac{\log n}{n h_1^d}}, h_1^2 \right\} \right) \text{ a.s.}$$

Next, using basic arguments, the left hand side in (3.1) is upper bounded as follows

$$\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] \right| \leq I_{1n} + I'_{1n} + I_{2n} + I'_{2n} + I_{3n},$$

with

$$\begin{aligned} I_{1n} &= \max_{1 \leq k \leq d_{x,n}} \sup_{\mathbf{x} \in B_k} \sup_{t \in \mathcal{C}} \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \tilde{F}_{1,n}(\mathbf{x}_k, t) \right|, \\ I'_{1n} &= \max_{1 \leq k \leq d_{x,n}} \sup_{\mathbf{x} \in B_k} \sup_{t \in \mathcal{C}} \left| \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}_k, t) \right] - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] \right|, \\ I_{2n} &= \max_{1 \leq k \leq d_{x,n}} \max_{1 \leq \ell \leq d_{t,n}} \sup_{t \in J_\ell} \left| \tilde{F}_{1,n}(\mathbf{x}_k, t) - \tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) \right|, \\ I'_{2n} &= \max_{1 \leq k \leq d_{x,n}} \max_{1 \leq \ell \leq d_{t,n}} \sup_{t \in J_\ell} \left| \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) \right] - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] \right|, \\ I_{3n} &= \max_{1 \leq k \leq d_{x,n}} \max_{1 \leq \ell \leq d_{t,n}} \left| \tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) \right] \right| \end{aligned}$$

Concerning I_{1n} and I'_{1n} , we apply the SLLN for associated sequences (see Newman [18]) and Assumption **A2(i)**. We obtain

$$\begin{aligned}
& \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \tilde{F}_{1,n}(\mathbf{x}_k, t) \right| \\
&= \left| \frac{1}{nh_1^d} \sum_{i=1}^n \frac{\delta_i}{\overline{G}(Y_i)} H\left(\frac{t - Y_i}{h_2}\right) \left[K_d\left(\frac{\mathbf{x} - \mathbf{X}_i}{h_1}\right) - K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_1}\right) \right] \right| \\
&\leq \frac{c}{h_1^d \overline{G}(\tau)} \frac{\|\mathbf{x} - \mathbf{x}_k\|^\alpha}{h_1^\alpha} \frac{1}{n} \sum_{i=1}^n \delta_i \\
&\leq \frac{c}{\overline{G}(\tau)} \frac{a_n^{d\alpha}}{h_1^{d+\alpha}} \frac{1}{n} \sum_{i=1}^n \delta_i \\
(5.6) \quad &= O\left(\frac{1}{\sqrt{nh_1^d}}\right).
\end{aligned}$$

To treat the terms I_{2n} and I'_{2n} , we use Assumption **A3** and Lemma 5.4. We get

$$\begin{aligned}
& \left| \tilde{F}_{1,n}(\mathbf{x}_k, t) - \tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) \right| \\
&= \left| \frac{1}{nh_1^d} \sum_{i=1}^n \frac{\delta_i}{\overline{G}(Y_i)} K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_1}\right) \left[H\left(\frac{t - Y_i}{h_2}\right) - H\left(\frac{t_\ell - Y_i}{h_2}\right) \right] \right| \\
&\leq \frac{c}{\overline{G}(\tau)} \frac{|t - t_\ell|}{h_2} \frac{1}{nh_1^d} \sum_{i=1}^n K_d\left(\frac{\mathbf{x}_k - \mathbf{X}_i}{h_1}\right) \\
&\leq \frac{c}{\overline{G}(\tau)} \frac{b_n}{h_2} l_n(\mathbf{x}_k) \\
(5.7) \quad &= O\left(\frac{1}{\sqrt{nh_1^d}}\right).
\end{aligned}$$

We can focus now on upper bounding the term I_{3n} . To do so, we apply an exponential inequality adapted to associated sequences (see, Theorem 1, p.19 in Doukhan and Neumann [10]). Indeed, for all $\varepsilon > 0$, we have

$$(5.8) \quad \mathbb{P}\left(\sum_{i=1}^n \Delta_i(\mathbf{x}_k, t_\ell) \geq \varepsilon\right) \leq \exp\left(-\frac{\varepsilon^2/2}{A_n + B_n^{1/(\mu+\lambda+2)} \varepsilon^{(2\mu+2\lambda+3)/(\mu+\lambda+2)}}\right),$$

where A_n is any number greater than σ_n^2 and

$$\begin{aligned}
& \sigma_n^2 := \left(\sum_{i=1}^n \Delta_i(\mathbf{x}_k, t_\ell)\right), \\
(5.9) \quad & B_n = 2cL_2 \max\left(\frac{2^{4+\mu+\lambda} c n h_1^d h_2^{\frac{2}{d+1}} L_1}{A_n}, 1\right).
\end{aligned}$$

Firstly, we have to calculate σ_n^2 . Indeed $\sigma_n^2 = (nh_1^d)^2 \text{Var} \left(\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) \right)$.
On the other hand, we have

$$\begin{aligned} (nh_1^d)^2 \text{Var} \left(\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) \right) &= n \text{Var} \left(\frac{\delta_1}{\overline{G}(Y_1)} K_{d,\mathbf{x}_k,1} H_{t_\ell,1} \left(\frac{t_\ell - Y_1}{h_2} \right) \right) \\ &+ \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \text{Cov} \left(\frac{\delta_i}{\overline{G}(Y_i)} K_{d,\mathbf{x}_k,i} H_{t_\ell,i}, \frac{\delta_j}{\overline{G}(Y_j)} K_{d,\mathbf{x}_k,j} H_{t_\ell,j} \right) \\ &=: V + S. \end{aligned}$$

Firstly, let us focus on V .

$$\begin{aligned} V &= n \mathbb{E} \left[\frac{\delta_1}{\overline{G}(Y_1)^2} K_d^2 \left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_1} \right) H^2 \left(\frac{t_\ell - Y_1}{h_2} \right) \right] \\ &- n \mathbb{E}^2 \left[\frac{\delta_1}{\overline{G}(Y_1)} K_d \left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_1} \right) H \left(\frac{t_\ell - Y_1}{h_2} \right) \right] \\ &=: n(D_1 - D_2). \end{aligned}$$

Concerning D_1 , we use classical conditional expectation techniques. So, under assumptions **A1(i)**, **A2** and **A9**, a change of variable and a Taylor expansion around \mathbf{x}_k , we get

$$\begin{aligned} D_1 &= \mathbb{E} \left[K_d^2 \left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_1} \right) \mathbb{E} \left[\mathbb{E} \left[\frac{\delta_1}{\overline{G}(Y_1)^2} H^2 \left(\frac{t_\ell - Y_1}{h_2} \right) | T_1 \right] | \mathbf{X}_1 \right] \right] \\ &= \int_{\mathbb{R}^d} K_d^2 \left(\frac{\mathbf{x}_k - \mathbf{u}}{h_1} \right) \mathbb{E} \left[H^2 \left(\frac{t_\ell - T_1}{h_2} \right) \frac{1}{\overline{G}(T_1)} | \mathbf{X}_1 = \mathbf{u} \right] l(\mathbf{u}) d\mathbf{u} \\ &\leq \int_{\mathbb{R}^d} K_d^2 \left(\frac{\mathbf{x}_k - \mathbf{u}}{h_1} \right) \int_{\mathbb{R}} \frac{1}{\overline{G}(v)} f(\mathbf{u}, v) dv d\mathbf{u}, \quad \text{because } \mathbb{H}(\cdot) \text{ is a df;} \\ &= h_1^d \int_{\mathbb{R}^d} K_d^2(\mathbf{z}) \varsigma(\mathbf{x}_k - \mathbf{z}h_1) d\mathbf{z} \\ &= h_1^d \int_{\mathbb{R}^d} \varsigma(\mathbf{x}_k) K_d^2(\mathbf{z}) d\mathbf{z} - h_1^{d+1} \int_{\mathbb{R}^d} K_d^2(\mathbf{z}) \left[z_1 \frac{\partial \varsigma(\mathbf{x}_k^*)}{\partial x_{k,1}} + \dots + z_d \frac{\partial \varsigma(\mathbf{x}_k^*)}{\partial x_{k,d}} \right] d\mathbf{z} \\ &= O(h_1^d). \end{aligned}$$

Here \mathbf{x}_k^* is between $\mathbf{x}_k - \mathbf{z}h_1$ and \mathbf{x}_k . Again, to upper bound D_2 we work similarly as before. Indeed, using a change of variable, Taylor expansion and assumptions **A1(i)**, **A2** and **A4**, we get

$$\begin{aligned} D_2 &= \mathbb{E}^2 \left[K_d \left(\frac{\mathbf{x}_k - \mathbf{X}_1}{h_1} \right) \mathbb{E} \left[\mathbb{E} \left[\frac{\delta_1}{\overline{G}(Y_1)} H \left(\frac{t_\ell - Y_1}{h_2} \right) | T_1 \right] | \mathbf{X}_1 \right] \right] \\ &= O(h_1^{2d}). \end{aligned}$$

Consequently, we get

$$V = O(nh_1^d).$$

Secondly, to evaluate S , we first define

$$B_1 = \{(i, j); 1 \leq |i - j| \leq \eta_n\} \text{ and } B_2 = \{(i, j); \eta_n + 1 \leq |i - j| \leq n - 1\},$$

where $\eta_n = o(n)$. Then, we have

$$\begin{aligned} S &= \sum_{i=1}^n \sum_{B_1} Cov \left(\frac{\delta_i}{\overline{G}(Y_i)} K_{d, \mathbf{x}_k, i} H_{t_l, i}, \frac{\delta_j}{\overline{G}(Y_j)} K_{d, \mathbf{x}_k, j} H_{t_l, j} \right) \\ &+ \sum_{i=1}^n \sum_{B_2} Cov \left(\frac{\delta_i}{\overline{G}(Y_i)} K_{d, \mathbf{x}_k, i} H_{t_l, i}, \frac{\delta_j}{\overline{G}(Y_j)} K_{d, \mathbf{x}_k, j} H_{t_l, j} \right) \\ &=: S_1 + S_2. \end{aligned}$$

From (5.3), it is clear that

$$(5.10) \quad S_1 = n\eta_n O(h_1^{2d} h_2^2) = O(n\eta_n h_1^{2d} h_2^2).$$

Next, the term S_2 will be upper bounded by remarking that result a) in Proposition 5.1 and Assumption **A8** permit to write

$$\begin{aligned} S_2 &\leq \sum_{i=1}^n \sum_{B_2} c^2 h_1^d h_2^{\frac{2}{d+1}} \rho^{\frac{d}{2d+2}} (|i-j|) \\ &\leq n c^2 h_1^d h_2^{\frac{2}{d+1}} \sum_{B_2} \gamma_0^{\frac{d}{2d+2}} e^{-\frac{\gamma|i-j|d}{2d+2}} \\ &\leq n c^2 h_1^d h_2^{\frac{2}{d+1}} \int_{\eta_n}^n e^{-\frac{\gamma u d}{2d+2}} du \\ (5.11) \quad &= O \left(n h_1^d h_2^{\frac{2}{d+1}} e^{-\frac{\gamma \eta_n d}{2d+2}} \right). \end{aligned}$$

So, under Assumption **A1** and taking $\eta_n = O(h_1^{\nu_1 - d} h_2^{\nu_2 - 1})$ with $0 < \nu_1 < d$ and $0 < \nu_2 < 1$, the bounds in (5.10) and (5.11) become of order $o(nh_1^d h_2)$ and $o(nh_1^d h_2^{\frac{2}{d+1}})$, respectively. Consequently

$$\sigma_n^2 = V + S = O(nh_1^d) + o(nh_1^d h_2^{\frac{2}{d+1}}) = O(nh_1^d).$$

Thereby, we get $A_n = O(nh_1^d)$. Next, from (5.9) and choosing μ, λ, L_1 and L_2 as those in the proof of Proposition 5.1, we get $B_n = O(1)$.

Regarding I_{3n} , in view of (5.1), (5.8) and letting $\varepsilon = \varepsilon_0 \sqrt{\frac{\log n}{nh_1^d}}$, we have

$$\begin{aligned}
& \mathbb{P} \left(\max_{1 \leq k \leq d_{x,n}} \max_{1 \leq \ell \leq d_{t,n}} \left| \tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) \right] \right| \geq \varepsilon \right) \\
&= \mathbb{P} \left(\max_{1 \leq k \leq d_{x,n}} \max_{1 \leq \ell \leq d_{t,n}} \left| \sum_{i=1}^n \Delta_i(\mathbf{x}_k, t_\ell) \right| \geq nh_1^d \varepsilon \right) \\
&\leq \sum_{k=1}^{d_{x,n}} \sum_{\ell=1}^{d_{t,n}} \mathbb{P} \left(\left| \sum_{i=1}^n \Delta_i(\mathbf{x}_k, t_\ell) \right| \geq nh_1^d \varepsilon \right) \\
&\leq 2d_{x,n} d_{t,n} \exp \left(\frac{-(nh_1^d)^2 \frac{\varepsilon_0^2}{2} \frac{\log n}{nh_1^d}}{c n h_1^d + \varepsilon_0^{5/3} (nh_1^d)^{5/3} \left(\frac{\log n}{nh_1^d} \right)^{5/6}} \right) \\
&\leq 2 \frac{m_1}{a_n^d} \frac{m_2}{b_n} \exp \left(\frac{-\frac{\varepsilon_0^2}{2} \log n}{c + \varepsilon_0^{5/3} \left(\frac{\log n}{nh_1^d} \right)^{1/6}} \right) \\
&\leq c \left(n^{-1} h_1^{2\alpha+d} \right)^{-1/2\alpha} \left(nh_1^d \right)^{1/2} h_2^{-1} n^{-c\varepsilon_0^2} \\
(5.12) \quad &= c \left(nh_1^{2\alpha+d(1-\alpha)} \right)^{-1/2\alpha} (nh_2)^{-1} n^{-c\varepsilon_0^2 + \frac{1}{\alpha} + \frac{3}{2}}.
\end{aligned}$$

So, under Assumption **A1** and taking $\varepsilon_0^2 > \frac{1}{c} \left(\frac{1}{\alpha} + \frac{5}{2} \right)$, the term in (5.12) is the general term of a convergent series. Then, we have

$$\sum_{n \geq 1} \mathbb{P} \left(\max_{1 \leq k \leq d_{x,n}} \max_{1 \leq \ell \leq d_{t,n}} \left| \tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}_k, t_\ell) \right] \right| \geq \varepsilon_0 \sqrt{\frac{\log n}{nh_1^d}} \right) < \infty.$$

Applying the lemma of Borel-Cantelli, we obtain that

$$(5.13) \quad I_{3n} = O \left(\sqrt{\frac{\log n}{nh_1^d}} \right).$$

Involving (5.6), (5.7) and (5.13), we deduce that

$$\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] \right| = O \left(\sqrt{\frac{\log n}{nh_1^d}} \right).$$

The proof of Theorem 3.2 is achieved. \square

Proof. Proof of Theorem 3.3 First, observe that

$$\begin{aligned}
\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} |F_n(t|\mathbf{x}) - F(t|\mathbf{x})| &\leq \frac{1}{\inf_{\mathbf{x} \in \Omega} (l_n(\mathbf{x}))} \left\{ \sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] - F_1(\mathbf{x}, t) \right| \right. \\
&\quad + \sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| F_{1,n}(\mathbf{x}, t) - \tilde{F}_{1,n}(\mathbf{x}, t) \right| \\
&\quad + \sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| \tilde{F}_{1,n}(\mathbf{x}, t) - \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] \right| \\
&\quad \left. + m_0^{-1} \sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} F_1(\mathbf{x}, t) \sup_{\mathbf{x} \in \Omega} |l_n(\mathbf{x}) - l(\mathbf{x})| \right\}.
\end{aligned}$$

As $m_0 := \inf_{\mathbf{x}} l(\mathbf{x})$, it is easily seen that

$$\begin{aligned} \frac{1}{l_n(\mathbf{x})} &\leq \frac{1}{l(\mathbf{x}) - |l_n(\mathbf{x}) - l(\mathbf{x})|} \\ &\leq \frac{1}{m_0 - \sup_{\mathbf{x} \in \Omega} |l_n(\mathbf{x}) - l(\mathbf{x})|}. \end{aligned}$$

This allows to write

$$\begin{aligned} &\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} |F_n(t|\mathbf{x}) - F(t|\mathbf{x})| \\ &\leq \frac{1}{m_0 - \sup_{\mathbf{x} \in \Omega} |l_n(\mathbf{x}) - l(\mathbf{x})|} \left\{ \vartheta_1 + \vartheta_2 + \vartheta_3 + \vartheta_4 m_0^{-1} \sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} F_1(\mathbf{x}, t) \right\}. \end{aligned}$$

As for the term ϑ_3 , it has been bounded in Theorem 3.2. The following lemmas establish respectively the result of ϑ_1 , ϑ_2 . Then we apply Lemma 5.4 for ϑ_4 .

The bias term ϑ_1 will be stated in Lemma 5.5 by using conditional expectation techniques and a Taylor expansion up to order 2 while Lemma 5.6 deals with bounding ϑ_2 .

5.5. Lemma. *Under assumptions **A1**, **A2**, **A3** and **A5**, for n large enough we have*

$$\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] - F_1(\mathbf{x}, t) \right| = O(h_1^2 + h_2^2), \text{ a.s.}$$

Proof. The following proof does not depend on the dependence structure.

$$\begin{aligned} \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] &= \frac{1}{h_1^d} \mathbb{E} \left[\frac{\delta_1}{\overline{G}(Y_1)} K_d \left(\frac{\mathbf{x} - \mathbf{X}_1}{h_1} \right) H \left(\frac{t - Y_1}{h_2} \right) \right] \\ &= \frac{1}{h_1^d} \mathbb{E} \left[K_d \left(\frac{\mathbf{x} - \mathbf{X}_1}{h_1} \right) \mathbb{E} \left[\frac{\delta_1}{\overline{G}(Y_1)} H \left(\frac{t - Y_1}{h_2} \right) \middle| \mathbf{X}_1 \right] \right]. \end{aligned}$$

We use integration by parts, a change of variable and Assumption **A3**, then we have

$$\begin{aligned} \mathbb{E} \left[\frac{\delta_1}{\overline{G}(Y_1)} H \left(\frac{t - Y_1}{h_2} \right) \middle| \mathbf{X}_1 \right] &= \mathbb{E} \left[\mathbb{E} \left[\frac{\delta_1}{\overline{G}(Y_1)} H \left(\frac{t - Y_1}{h_2} \right) \middle| T_1 \right] \middle| \mathbf{X}_1 \right] \\ &= \mathbb{E} \left[\frac{1}{\overline{G}(T_1)} H \left(\frac{t - T_1}{h_2} \right) \mathbb{E} \left[\mathbb{1}_{\{C_1 \geq T_1\}} \right] \middle| \mathbf{X}_1 \right] \\ &= \mathbb{E} \left[H \left(\frac{t - T_1}{h_2} \right) \middle| \mathbf{X}_1 \right] \\ &= \int_{\mathbb{R}} H \left(\frac{t - y}{h_2} \right) f(y|\mathbf{X}_1) dy \\ &= \int_{\mathbb{R}} H^{(1)}(z) F(t - zh_2|\mathbf{X}_1) dz. \end{aligned}$$

Again, by a change of variable we get

$$\begin{aligned} \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] &= \frac{1}{h_1^d} \mathbb{E} \left[K_d \left(\frac{\mathbf{x} - \mathbf{X}_1}{h_1} \right) \int_{\mathbb{R}} H^{(1)}(z) F(t - zh_2|\mathbf{X}_1) dz \right] \\ &= \int_{\mathbb{R}^d} \frac{1}{h_1^d} K_d \left(\frac{\mathbf{x} - \mathbf{u}}{h_1} \right) \int_{\mathbb{R}} H^{(1)}(z) F(t - zh_2|\mathbf{X}_1 = \mathbf{u}) l(\mathbf{u}) d\mathbf{u} dz \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \frac{1}{h_1^d} K_d \left(\frac{\mathbf{x} - \mathbf{u}}{h_1} \right) H^{(1)}(z) F_1(\mathbf{u}, t - zh_2) d\mathbf{u} dz \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} K_d(\mathbf{r}) H^{(1)}(z) F_1(\mathbf{x} - \mathbf{r}h_1, t - zh_2) d\mathbf{r} dz. \end{aligned}$$

Then, expanding $F_1(\mathbf{x} - \mathbf{r}h_1, t - zh_2)$ up to order 2 around (\mathbf{x}, t) gives

$$\begin{aligned} F_1(\mathbf{x} - \mathbf{r}h_1, t - zh_2) &= F_1(\mathbf{x}, t) \\ &- h_1 \left[r_1 \frac{\partial F_1(\mathbf{x}, t)}{\partial x_1} + \cdots + r_d \frac{\partial F_1(\mathbf{x}, t)}{\partial x_d} \right] - h_2 \left[z \frac{\partial F_1(\mathbf{x}, t)}{\partial t} \right] \\ &+ \frac{h_1^2}{2} \left[r_1^2 \frac{\partial^2 F_1(\mathbf{x}^*, t^*)}{\partial x_1^2} + \cdots + r_d^2 \frac{\partial^2 F_1(\mathbf{x}^*, t^*)}{\partial x_d^2} + 2 \sum_{i,j:i \neq j} r_i r_j \frac{\partial^2 F_1(\mathbf{x}^*, t^*)}{\partial x_i \partial x_j} \right] \\ &+ \frac{h_2^2}{2} \left[z^2 \frac{\partial^2 F_1(\mathbf{x}^*, t^*)}{\partial t^2} \right] + h_1 h_2 \left[r_1 z \frac{\partial^2 F_1(\mathbf{x}^*, t^*)}{\partial x_1 \partial t} + \cdots + r_d z \frac{\partial^2 F_1(\mathbf{x}^*, t^*)}{\partial x_d \partial t} \right]. \end{aligned}$$

Here, (\mathbf{x}^*, t^*) lies between $(\mathbf{x} - \mathbf{r}h_1, t - zh_2)$ and (\mathbf{x}, t) . Finally, assumptions **A1**, **A2**, **A3** and **A5** entail

$$\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| \mathbb{E} \left[\tilde{F}_{1,n}(\mathbf{x}, t) \right] - F_1(\mathbf{x}, t) \right| \leq c(h_1^2 + h_2^2).$$

This provides the desired result. \square

5.6. Lemma. *Under assumptions **A2**, **A4** and **A8**, for n large enough, we have*

$$\sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{C}} \left| F_{1,n}(\mathbf{x}, t) - \tilde{F}_{1,n}(\mathbf{x}, t) \right| = o\left(n^{-\theta}\right), \text{ a.s.}$$

Proof. Firstly, we have

$$\begin{aligned} &|F_{1,n}(\mathbf{x}, t) - \tilde{F}_{1,n}(\mathbf{x}, t)| \\ &= \left| \frac{1}{nh_1^d} \sum_{i=1}^n \delta_i K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_1} \right) H \left(\frac{t - Y_i}{h_2} \right) \left(\frac{1}{\bar{G}_n(Y_i)} - \frac{1}{\bar{G}(Y_i)} \right) \right| \\ &\leq \frac{1}{nh_1^d} \sum_{i=1}^n K_d \left(\frac{\mathbf{x} - \mathbf{X}_i}{h_1} \right) H \left(\frac{t - Y_i}{h_2} \right) \left| \frac{1}{\bar{G}_n(Y_i)} - \frac{1}{\bar{G}(Y_i)} \right| \\ &\leq \frac{l_n(\mathbf{x})}{\bar{G}_n(\tau) \bar{G}(\tau)} \sup_{t \in \mathcal{C}} |\bar{G}_n(t) - \bar{G}(t)|. \end{aligned}$$

Then, following Theorem 1.4 in Cai and Roussas [5] and for n large enough, we easily get

$$(5.14) \quad \sup_{t \in \mathcal{C}} |\bar{G}_n(t) - \bar{G}(t)| = o\left(n^{-\theta}\right), \text{ a.s.}$$

Recall that $0 < \theta < \gamma/(2\gamma + 9 + 3/2\kappa)$ for any real $\kappa > 0$ and γ is referred in Assumption **A8**. Hence, Lemma 5.4 and (5.14) end the proof of Lemma 5.6. \square

To end the proof of Theorem 3.3, it suffices to apply Lemma 5.4 for ϑ_4 . \square

The last step consists in proving the result on the behavior of the conditional quantile estimator $\xi_{p,n}(\mathbf{x})$, stated in Corollary 3.4.

Proof. Proof of Corollary 3.4 It suffices to use the following triangular inequality jointly with basic arguments. Let $\mathbf{x} \in \Omega$, then we have

$$\begin{aligned} |F(\xi_{p,n}(\mathbf{x})|\mathbf{x}) - F(\xi_p(\mathbf{x})|\mathbf{x})| &\leq |F_n(\xi_{p,n}(\mathbf{x})|\mathbf{x}) - F(\xi_{p,n}(\mathbf{x})|\mathbf{x})| \\ &\quad + |F_n(\xi_{p,n}(\mathbf{x})|\mathbf{x}) - F(\xi_p(\mathbf{x})|\mathbf{x})| \\ (5.15) \quad &\leq 2 \sup_{t \in \mathcal{C}} |F_n(t|\mathbf{x}) - F(t|\mathbf{x})|. \end{aligned}$$

So, the first part of Corollary 3.4 is straightforwardly deduced from Theorem 3.3. And, a Taylor expansion of $F(\xi_{p,n}|\mathbf{x})$ in the neighborhood of $\xi_p(\mathbf{x})$ permits to get

$$F(\xi_{p,n}(\mathbf{x})|\mathbf{x}) - F(\xi_p(\mathbf{x})|\mathbf{x}) = (\xi_{p,n}(\mathbf{x}) - \xi_p(\mathbf{x}))f(\xi_p^*(\mathbf{x})|\mathbf{x}),$$

where $\xi_p^*(\cdot)$ is between $\xi_p(\cdot)$ and $\xi_{p,n}(\cdot)$. Thus from (5.15), we obtain

$$\sup_{\mathbf{x} \in \Omega} |\xi_{p,n}(\mathbf{x}) - \xi_p(\mathbf{x})|f(\xi_p^*(\mathbf{x})|\mathbf{x}) \leq 2 \sup_{\mathbf{x} \in \Omega} \sup_{t \in \mathcal{E}} |F_n(t|\mathbf{x}) - F(t|\mathbf{x})|.$$

Note that if the condition in Corollary 3.4 is not checked, one has to consider a higher order-Taylor expansion. Finally, the desired result follows immediately from Assumption **A5** and Theorem 3.3. \square

References

- [1] Abberger, K. *Quantile smoothing in financial time series*, Statist. Papers **38**, 125–148, 1997.
- [2] Bhattacharya, P.K. and Gangopadhyay, A. *Kernel and nearest neighbour estimation of conditional quantile*, Ann. Statist. **18**, 1400–1415, 1990.
- [3] Bulinski, A. and Shashkin, A. *Limit theorems for associated random fields and related systems*, Vol **10**, Advanced series on statistical science & applied probability, World Scientific, Singapore, 2007.
- [4] Cai, Z. *Asymptotic properties of Kaplan-Meier estimator for censored dependent data*, Statist. Probab. Lett. **37**, 381–389, 1998.
- [5] Cai, Z. and Roussas, G.G. *Kaplan-Meier estimator under association*, J. Multivariate Anal. **67**, 318–348, 1998.
- [6] Chaudhuri, P., Doksum, K. and Samarov, A. *On average derivative quantile regression*, Ann. Statist. **25**, 715–744, 1997.
- [7] Dabrowska, D. *Nonparametric quantile regression with censored data*, Sankhya A **54**, 252–259, 1992.
- [8] Doukhan, P. *Mixing: Properties and Examples*, Lect. Notes in Statist. **61**, Springer Verlag, 1994.
- [9] Doukhan, P. and Louhichi, S. *A new weak dependence condition and applications to moment inequalities*, Stochastic Process. Appl. **84**, 313–342, 1999.
- [10] Doukhan, P. and Neumann, M. *Probability and moment inequalities for sums of weakly dependent random variables, with applications*, Stochastic Process. Appl. **117**, 878–903, 2007.
- [11] Esary, J., Proschan, F. and Walkup, D. *Association of random variables with applications*, Ann. Math. Statist. **38**, 1466–1476, 1967.
- [12] Honda, T. *Nonparametric estimation of a conditional quantile for strong mixing processes*, Ann. Inst. Statist. Math. **52**, 459–470, 2000.
- [13] Kaplan, E.M. and Meier, P. *Nonparametric estimation from incomplete observations*, J. Amer. Statist. Assoc. **53**, 457–481, 1958.
- [14] Koul, H., Susarla, V. and Van Ryzin, J. *Regression analysis with randomly right censored data*, Ann. Statist. **9**, 1276–1288, 1981.
- [15] Liang, H.-Y. and de Uña-Álvarez, J. *Asymptotic properties of conditional quantile estimator for censored dependent observations*, Ann. Inst. Statist. Math. **63**, 267–289, 2011.
- [16] Mehra, K.L., Rao M.S. and Upadrasta S.P. *A smooth conditional quantile estimator and related applications of conditional empirical processes*, J. Multivariate Anal. **37**, 151–179, 1991.
- [17] Menni, N. and Tatachak, A. *A note on estimating the conditional expectation under censoring and association: strong uniform consistency*, Statist. Papers, DOI:10.1007/s00362-016-0801-8, 2016.
- [18] Newman, C.M. *Asymptotic independence and limit theorems for positively and negatively dependent random variables*, in: Tong, Y.L. (Ed.) Inequalities in statistics and probability, IMS Lecture Notes-Monograph Series, vol. **5**, Hayward, CA, 127–140, 1984.
- [19] Oliveira, P.E. *Asymptotics for Associated Random Variables*, Springer Verlag, 2012.

- [20] Ould Saïd, E. *A strong uniform convergence rate of kernel conditional quantile estimator under random censorship*, Statist. Probab. Lett. **76**, 579–586, 2006.
- [21] Ould Saïd, E. and Sadki, O. *Prediction via the conditional quantile for right censorship model*, Far East Journal of Theoretical Statistics **25**, 145–179, 2008.
- [22] Ould Saïd, E. and Sadki, O. *Strong approximation of quantile function for strong mixing and censored processes*, Comm. Statist. Theory Methods **34**, 1449–1459, 2005.
- [23] Prakasa Rao, B.L.S. *Associated Sequences, Demimartingales and Nonparametric Inference*, Probability and its Applications, Springer Basel AG, 2012.
- [24] Qin, Y.S. and Wu, Y. *An estimator of a conditional quantile in the presence of auxiliary information*, J. Statist. Plann. Inference **99**, 59–70, 2001.
- [25] Samanta, M. *Nonparametric estimation of conditional quantiles*, Statist. Probab. Lett. **7**, 407–412, 1989.
- [26] Stute, W. and Wang, J.L. *The strong law under random censorship*, Ann. Statist. **21**, 1591–1607, 1993.
- [27] Xiang, X. *A kernel estimator of a conditional quantile*, J. Multivariate Anal. **59**, 206–216, 1996.
- [28] Yu, K. and Jones, M.C. *Local linear quantile regression*, J. Amer. Statist. Assoc. **93**, 228–238, 1998.

Two-stage network DEA with convex hull in intermediate products

Samira Maleki*, Ali Ebrahimnejad^{†‡} and Reza Kazemi Matin[§]

Abstract

In recent Data envelopment analysis (DEA) literature, many researchers have examined systems with a two-stage network structure and its pitfalls. In these studies, two-stage network systems operations for converting inputs into outputs have been performed in two stages, meanwhile the intermediate products were considered as the outputs from the first stage and as inputs to the second stage. This duality in dealing with intermediate products imposes restrictions on the pricing of these products. In this paper by focusing on the convexity axiom, we define a new production possibility set. The main contributions of this paper are fourfold: (1) we propose models for evaluating the overall efficiency measure of decision making units (DMUs) in a two-stage network structure based on the convex hull in intermediate products; (2) we propose a procedure to determine the target unit of each inefficient DMU; (3) we explain how to calculate the divisional efficiency; (4) we demonstrate the feasibility and richness of the obtained solutions in the context of two examples.

Keywords: Data envelopment analysis (DEA), Two-stage network, Intermediate products, Convex hull, Overall efficiency, Projection.

2000 AMS Classification: 90Cxx, 90C05.

Received : 26.05.2017 *Accepted :* 28.06.2018 *Doi :* 10.15672/HJMS.2018.610

*Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran,
Email: samira.maleky@gmail.com

[†]Department of Mathematics, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran,
Emails: aemarzoun@gmail.com and a.ebrahimnejad@qaemiau.ac.ir

[‡]Corresponding Author.

[§]Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran,
Email: rkmatin@kiaiu.ac.ir

1. Introduction

Data envelopment analysis (DEA) is one of the popular and growing methods in evaluating the relative efficiency of a set of similar decision making units (DMUs). In practice, DEA performs in the presence of multi-inputs variables and multi-outputs to evaluate efficiency using a model originally proposed by Charnes et al. [1]. Later, extended DEA models were applied in different contexts (see Cook and Seiford[7]). Initial DEA methods measured the efficiency of systems without any attention to the internal structure of system operations. DEA has many applications for interpreting the productivity of complex economical and engineering systems (Ebrahimnejad et al. [8, 9], Ebrahimnejad and Bagerzadeh [10], Mottaghi et al. [15], Hatami-Marbini et al. [11], Tavana et al. [19])

Over time, researchers had more attention to system operation analysis in order to find the causes of system inefficiency. The first study with a two-stage network structure using DEA was reported by Charnes et al. [2], examining matters related to employment in the army. This two-stage network model was then used by many researchers, such as Lovell et al. [14], Seiford and Zhu [17] and Sexton and Lewis [18].

In recent years, several models have been proposed to improve efficiency measurement in two-stage network systems. Wang et al. [21] introduced a two-stage method with variable returns to scale (VRS), which in each stage considered variables independently and provided intermediate products. Rho and An [16] considered slack variables in a model that provided assessment of DMUs with weak efficiency. Kao and Hwang [12] examined the possibility of decomposition in the system's overall efficiency by considering intermediate products' weights. Tone and Tsutsui [20], using a production possibility set (PPS), introduced models based on the slack variable, and Chen et al. [5] provided a new method for determining efficient projections for inefficient DMUs. Although the main contributions of these models were improvement in measuring efficiency in a two-stage network structure, but they have many problems. For example, Chen et al. [6] reported some of the limitations in efficiency measurement related to the different behaviors occurring in the stages due to using intermediate products. Furthermore, Chen et al. [6] examined the determination of projection, efficient frontiers, and divisional efficiency become challenges in network DEA models. In all these methods, intermediate products were considered in the first stage as outputs and in the second stage as inputs with free disposability. Therefore, considering intermediate products with two different roles is caused problems within the system. In this study, we are going to have a uniform behavior with the intermediate products, when they are applied as inputs (consumer) of the second stage and outputs (products) of the first stage. For this purpose, we introduce a set of separated properties for every stage of the two-stage network, including consideration of the convex hull of the intermediate products. Models based upon this new production possibility set are presented to calculate the overall efficiency and projections. Finally, the new network DEA model is compared with similar methods in evaluating how well they addressed two-stage network limitations.

The rest of the paper is organized into several sections. In Section (2), a brief review of some systems with a two-stage network structure is presented. In Section (3), some properties regarding two-stage network DEA with convex intermediate products are explored, and T_{CHI} is established by accepting these principals. Section (4) presents a new model to compute the overall efficiency score in the proposed PPS. Section (5) explored models determined by the divisional efficiency in T_{CHI} . In Section (6), we propose a method to improve inefficient DMUs and to calculate frontier projections in T_{CHI} . Some examples are illustrated in Section (7). Finally, Section (8), including the main conclusions as well as some interesting future research lines, ends the paper.

2. Two-stage network

Consider Figure 1 that represents a two-stage network structure for each of a set of n DMUs.

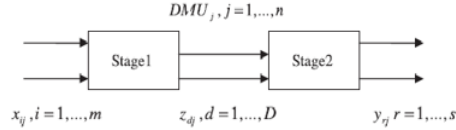


Figure 1. Two-stage process.

We apply Kao and Hwang's model [12] to explain the main concepts. For each DMU_j ($j = 1, 2, \dots, n$), in the first stage, inputs x_{ij} , ($i = 1, 2, \dots, m$) is used to produce a set of D , intermediate products, z_{dj} , ($d = 1, 2, \dots, D$), and in the second stage all outputs of the first stage, namely, z_{dj} , ($d = 1, 2, \dots, D$) is used to produce the final outputs y_{rj} , ($r = 1, 2, \dots, s$). In what follows $x = (x_1, \dots, x_m) \in R_+^m, z = (z_1, \dots, z_d) \in R_+^D, y = (y_1, \dots, y_s) \in R_+^s$, represent the input vector, intermediate products vector, and output vector, respectively.

In conventional models of DEA, two different methods are commonly used to evaluate the efficiency of two-stage systems. The first method calculated efficiency of each division based upon the application of the definition of relative efficiency in a set of DMUs, and the multiplier-based network DEA models are derived according to this method. In the second method, the production possibility set is used for measuring efficiency of each division, and the envelopment models are derived with a two-stage network structure. Network DEA pitfalls were represented by applying different concepts of efficiency in these two different methods. A brief review of these two methods appears in the following subsections.

2.1. The multiplier models with two-stage network structure. The multiplier-based network models are generally applied to calculate overall and divisional efficiency. Of course, one of the limitations of the network DEA models is that the divisional efficiency envelopment models are infeasible in some cases. Kao and Hwang [12], under constant returns to scale (CRS) assumption, calculated stages' efficiency scores separately, then considering a series relationship between stages, they obtained an overall efficiency score by the products of each stages' efficiency. One notable point in their method is that the weights related to intermediate products are equal in both stages ($w_d; d = 1, \dots, D$).

Kao and Hwang [12] proposed the following linear programming (LP) model to evaluate the overall efficiency measure for DMU_o :

$$\begin{aligned}
 E_o = \max \quad & \sum_{r=1}^s u_r y_{ro} \\
 \text{s.t.} \quad & \sum_{r=1}^s u_r y_{rj} - \sum_{d=1}^D w_d z_{dj} \leq 0 \quad j = 1, \dots, n \\
 & \sum_{d=1}^D w_d z_{dj} - \sum_{i=1}^m v_i x_{ij} \leq 0 \quad j = 1, \dots, n
 \end{aligned}$$

$$(2.1) \quad \begin{aligned} & \sum_{i=1}^m v_i x_{io} = 1 \\ & v_i \geq 0, u_r \geq 0, w_d \geq 0, i = 1, \dots, m; r = 1, \dots, s; d = 1, \dots, D \end{aligned}$$

In model (2.1), $u \in R_+^s, v \in R_+^m$ and $w \in R_+^D$ are the associated unknown weights of the output, input and intermediate products, respectively. Kao and Hwang [13] also provided a method for calculating the overall efficiency score of DMU_o under VRS. They introduced following models for calculating the divisional efficiency scores.

$$(2.2) \quad \begin{aligned} T_o^1 = \max & \quad \sum_{d=1}^D \acute{w}_d z_{do} - \acute{w}_0 \\ \text{s.t.} & \quad \sum_{i=1}^m v_i x_{io} = 1 \\ & \quad E_o \sum_{i=1}^m v_i x_{io} - \sum_{r=1}^s u_r y_{ro} = 0 \\ & \quad \sum_{r=1}^s u_r y_{rj} - \sum_{d=1}^D w_d z_{dj} \leq 0 \quad j = 1, \dots, n \\ & \quad \sum_{d=1}^D w_d z_{dj} - \sum_{i=1}^m v_i x_{ij} \leq 0 \quad j = 1, \dots, n \\ & \quad \sum_{d=1}^D \acute{w}_d z_{dj} - \acute{w}_0 - \sum_{i=1}^m v_i x_{ij} \leq 0 \quad j = 1, \dots, n \\ & \quad u_r, v_i, w_p, \acute{w}_p \geq \varepsilon \\ & \quad \acute{w}_0 \text{ free in sign} \end{aligned}$$

$$(2.3) \quad \begin{aligned} T_o^2 = \max & \quad \sum_{r=1}^s u_r y_{ro} \\ \text{s.t.} & \quad \sum_{d=1}^D \acute{w}_d z_{do} + \acute{w}_0 = 1 \\ & \quad E_o \sum_{i=1}^m v_i x_{io} - \sum_{r=1}^s u_r y_{ro} = 0 \\ & \quad \sum_{r=1}^s u_r y_{rj} - \sum_{d=1}^D w_d z_{dj} \leq 0 \quad j = 1, \dots, n \\ & \quad \sum_{d=1}^D w_d z_{dj} - \sum_{i=1}^m v_i x_{ij} \leq 0 \quad j = 1, \dots, n \\ & \quad \sum_{r=1}^s u_r y_{rj} - \sum_{d=1}^D \acute{w}_d z_{dj} - \acute{w}_0 \leq 0 \quad j = 1, \dots, n \\ & \quad u_r, v_i, w_p, \acute{w}_p \geq \varepsilon \\ & \quad \acute{w}_0 \text{ free in sign} \end{aligned}$$

Hence, E_o is the overall efficiency score under CRS calculated from model (2.1). Model (2.2) computes the input-oriented technical efficiency score of the first step and model

(2.3) evaluates the output-oriented efficiency score of the second stage for DMU_o . The overall efficiency score of DMU_o , under VRS, is obtained via the products of the stages' efficiency scores.

Chen et al. [3] calculated the overall efficiency in a two-stage system by use of specific weights in the objective function. They proposed the following model to compute the VRS's overall efficiency in a two-stage system, when DMU_o is under evaluation:

$$\begin{aligned}
(2.4) \quad & \max \quad \sum_{d=1}^D w_d z_{do} + u^1 + \sum_{r=1}^s u_r y_{ro} + u^2 \\
& s.t. \quad \sum_{r=1}^s u_r y_{rj} - \sum_{d=1}^D w_d z_{dj} + u^2 \leq 0 \quad j = 1, \dots, n \\
& \quad \sum_{d=1}^D w_d z_{dj} - \sum_{i=1}^m v_i x_{ij} + u^1 \leq 0 \quad j = 1, \dots, n \\
& \quad \sum_{i=1}^m v_i x_{io} + \sum_{d=1}^D w_d z_{do} = 1 \\
& \quad v_i \geq 0, u_r \geq 0, w_d \geq 0, i = 1, \dots, m; r = 1, \dots, s; d = 1, \dots, D \\
& \quad u^1, u^2 \text{ free in sign}
\end{aligned}$$

As can be seen from Model (2.4), intermediate products' weights are considered the same in both stages of the proposed model.

2.2. Envelopment models with two-stage network structure. DEA models with network structures are used for determining projections on the efficiency frontier. Chen et al. [4] introduced a radial version of the envelopment-based network model to compute the input-oriented CRS overall efficiency for DMU_o as follows:

$$\begin{aligned}
(2.5) \quad & \min \theta \\
& s.t. \quad \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io}, \quad i = 1, 2, \dots, m \\
& \quad \sum_{j=1}^n \lambda_j z_{dj} \geq \tilde{z}_{do}, \quad d = 1, 2, \dots, D \\
& \quad \sum_{j=1}^n \mu_j z_{dj} \leq \tilde{z}_{do}, \quad d = 1, 2, \dots, D \\
& \quad \sum_{j=1}^n \mu_j y_{rj} \geq y_{ro}, \quad r = 1, 2, \dots, s \\
& \quad \lambda_j \geq 0, \mu_j \geq 0 \quad j = 1, 2, \dots, n \\
& \quad \tilde{z}_{do} \geq 0 \quad d = 1, 2, \dots, D
\end{aligned}$$

Model (2.5), is equivalent to the dual of Model (2.1). In Model (2.5), the intermediate products are treated as free links.

Chen et al. [5] applied redundant constraints of Model (2.5), such as $\tilde{z}_{do} \geq 0$ ($d = 1, 2, \dots, D$) in Model (2.5) as unrestricted variables, i.e. \tilde{z}_{do} ($d = 1, 2, \dots, D$) are free in sign. Therefore, in the dual model, both constraints of intermediate products corresponding to these free variables considered as equal constraints. For reformulating Model (2.5) under the assumption of VRS, it is enough to add the convexity constraints ($\sum_{j=1}^n \mu_j = \sum_{j=1}^n \lambda_j = 1$) to model (2.5).

Tone and Tsutsui [20] introduced slacks-based network DEA models by using production possibility set. They explored several models based upon the intermediate products as both fixed links and free links. A version of the input-oriented envelopment-based network model, where the intermediate products referred to as free link cases can be modeled as follows:

$$\begin{aligned}
(2.6) \quad & \max \quad \sum_{i=1}^m \frac{s_i^-}{x_{io}} \\
& s.t. \quad \sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{io}, \quad i = 1, \dots, m \\
& \quad \quad \sum_{j=1}^n \lambda_j z_{dj} \geq \tilde{z}_{do}, \quad d = 1, \dots, D \\
& \quad \quad \sum_{j=1}^n \mu_j z_{dj} \leq \tilde{z}_{do}, \quad d = 1, \dots, D \\
& \quad \quad \sum_{j=1}^n \mu_j y_{rj} - s_r^+ = y_{ro}, \quad r = 1, \dots, s \\
& \quad \quad \lambda_j \geq 0, \mu_j \geq 0, \quad j = 1, \dots, n \\
& \quad \quad \tilde{z}_{do} \geq 0, \quad d = 1, \dots, D
\end{aligned}$$

In general, two different behaviors with intermediate products in most models of a two-stage network can be considered one of their most significant problems while minimum attention has been paid to them. These methods by assigning the same weight variables to the intermediate products in two stages of the multiplier-based network models (for example, $w_d; d = 1, \dots, D$) impose uniform shadow prices to the system and so they have limitations or have less flexibility.

By changing the direction of inequality in the corresponding constraints of intermediate products in the envelopment network models, the model provides the possibility of disposability, which is consumed in the next stage. Therefore, models manage the problem from the outside. In the next section, we examine the impact of uniform behavior on these products for calculating overall efficiency and projections.

3. Two-stage network DEA with convex intermediate products

In this section, we propose separate axioms for each stage in a two-stage network structure. Using these axioms, we form a new production possibility set for a two-stage network DEA with convex intermediate products. In addition, some of the related properties are also presented.

We postulate the following axioms for the production possibility set of the first stage:

- A1** . The observed activities $(x_j, z_j), (j = 1, 2, \dots, n)$ belong to T_1 .
- A2** . Any convex combination of activities in T_1 belongs to T_1 .
- A3** .For an activity (x, z) , in T_1 , any semi positive activity (\bar{x}, z) with $\bar{x} \geq x$ is included in T_1 .

Thus, we define the production possibility set T_1 that satisfies A1-A3 as follows:

$$(3.1) \quad T_1 = \{(x, z) : \sum_j \lambda_j x_j \leq x, \sum_j \lambda_j z_j = z, \sum_j \lambda_j = 1, \lambda_j \geq 0; j = 1, \dots, n\}$$

In PPS T_1 , the variable $\lambda \in R^n$ is the vector of intensity variables of the first stage.

3.1. Theorem. *Technology of the first stage, T_1 , which is defined in set (3.1) is the minimal set that contains all observations and satisfies the axioms of strong disposability of inputs and convexity.*

The proof of Theorem (3.1) is given in the Appendix.

For the PPS of the second stage, we postulate the following axioms:

- B1** . The observed activities $(z_j, y_j), (j = 1, 2, \dots, n)$ belong to T_2 .
- B2** . Any convex combination of activities in T_2 belongs to T_2 .
- B3** . For an activity (z, y) , in T_2 , any semi positive activity (z, \bar{y}) with $y \geq \bar{y}$ is included in T_2 .

Then, we define the production possibility set T_2 that satisfies **B1-B3** as follows:

$$(3.2) \quad T_2 = \left\{ (z, y) : \sum_j \mu_j z_j = z, \sum_j \mu_j y_j \geq y, \sum_j \mu_j = 1, \mu_j \geq 0; j = 1, \dots, n \right\}$$

In PPS T_2 , the variables $\mu \in R^n$ denote the intensity levels of the DMUs for the second stage.

3.2. Theorem. *The second stage technology T_2 defined in (3.2) is the minimal set that contains all observations and satisfies the axioms of strong disposability of outputs and convexity.*

The proof of Theorem (3.2) is similar to the proof of Theorem (3.1) and is omitted.

According to the proposed axioms for each stage, and under the assumption of VRS, we define the overall production possibility set for the two-stage network with convex intermediate products as follows:

$$(3.3) \quad T_{CHI} = \left\{ (x, z, y) : \sum_j \lambda_j x_j \leq x, \sum_j \lambda_j z_j = z, \sum_j \mu_j z_j = z, \right. \\ \left. \sum_j \mu_j y_j \geq y, \sum_j \mu_j = 1, \sum_j \lambda_j = 1, \lambda_j \geq 0, \right. \\ \left. \mu_j \geq 0, \quad j = 1, \dots, n \right\}$$

In fact, notation CHI represents the convex hull of intermediate products.

Note that the intermediate products in T_{CHI} are examined by two separate sets of $\lambda \in R^n$ and $\mu \in R^n$, and thus λ determines the relation between inputs and intermediate products and μ determines the relation between intermediate products and outputs.

The main difference of the technology expressed in (3.3) from the conventional technology of the two-stage network is that the former allows the free disposability for the intermediate products.

In the above technology, the produced output ratio in the first stage is equal to the consumed input ratio to the second stage. Therefore, access to resources became restricted, and the produced possibility set generated by technology set (3.3) becomes a subset of the traditional two-stage network production possibility set.

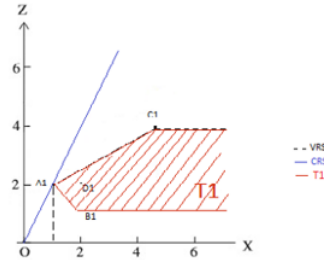
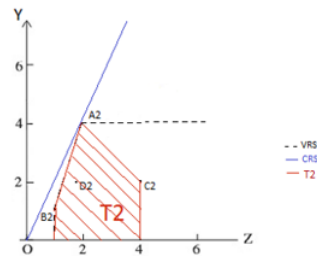
Here, we present an illustrative example to compare the overall efficiency frontier and each stage of the two-stage system, using efficiency frontier conventional technology under the condition of CRS and VRS.

3.3. Example. Consider a system that includes four DMUs. Each DMU has one input, one output, and one intermediate measure. The data set is given in Table 1.

Now, we can show the overall production technology, and PPS of each stage of the two-stage structure as seen in Figure 2, Figure 3, and Figure 4.

Table 1. Data set of example (3.3)

DMU	x	z	y
A	1	2	4
B	2	1	1
C	5	4	2
D	2	2	2

**Figure 2.** Efficient frontier generated by T_1 .**Figure 3.** Efficient frontier generated by T_2 .

In Figure 2, the red lines show the production frontier of stage 1 (T_1). Note that the efficiency frontier has been expanded by the convex hull of observations and the strong disposability in inputs.

In Figure 3, the red lines show the production frontier of stage 2 (T_2), in which the frontier points are obtained with the convex hull of observations and the strong disposability axiom in outputs. The blue lines and black dotted lines on both Figures 2 and 3 represent the efficiency frontier, under the assumptions of CRS and VRS, respectively. Clearly, the production possibility set with the convex hull in intermediate products is a subset of the production possibility set under the assumption of both CRS and VRS.

Figure 4 illustrate the tridimensional network technology showing the convex hull in intermediate products.

4. Introducing a model to determine overall efficiency in T_{CHI}

In this section, we present a new network DEA model to calculate the overall efficiency in T_{CHI} . To do this, we first consider an input-oriented model.

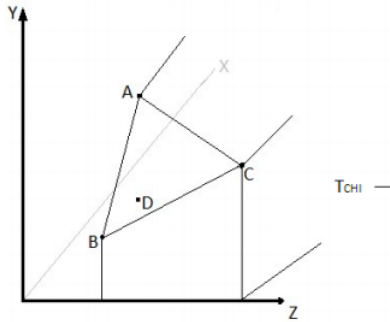


Figure 4. Efficient frontier generated by T_{CHI} .

The new network model suggested to evaluate the overall efficiency of DMU_o in T_{CHI} is given as follows:

$$\begin{aligned}
 \theta_o &= \min \theta \\
 \text{s.t. } & \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io} \quad i = 1, \dots, m \\
 & \sum_{j=1}^n \lambda_j z_{dj} = z_{do} \quad d = 1, \dots, D \\
 & \sum_{j=1}^n \mu_j z_{dj} = z_{do} \quad d = 1, \dots, D \\
 & \sum_{j=1}^n \mu_j y_{rj} \geq y_{ro} \quad r = 1, \dots, s \\
 & \sum_{j=1}^n \mu_j = \sum_{j=1}^n \lambda_j = 1 \\
 (4.1) \quad & \lambda_j \geq 0, \mu_j \geq 0 \quad j = 1, \dots, n
 \end{aligned}$$

Model (4.1) is similar to the VRS two-stage network model of Chen et al. [4] given in model (2.5), with the difference that in the above model the constraints related to intermediate products are considered in a convex set of intermediate data. In fact, outputs of the first stage are exactly equal to the inputs of the second stage and the overall efficiency of DMU_o is evaluated into a set of fixed intermediate products. It can be seen that model (4.1) is always feasible, and $0 < \theta_o^* \leq 1$.

4.1. Definition (Input-oriented overall efficiency). The under evaluation DMU_o is said to be overall input-efficient with intermediate convex products, if the optimal value of model (4.1) is equal to one; namely $\theta_o^* = 1$.

In order to describe model (4.1), we can rewrite it to evaluate the overall efficiency score for DMU_B given in example 3.3 as follows:

$$\begin{aligned}
 \theta_o^B &= \min \theta \\
 \text{s.t. } & \lambda_1 + 2\lambda_2 + 5\lambda_3 + 2\lambda_4 \leq 2\theta \\
 & 2\lambda_1 + \lambda_2 + 4\lambda_3 + 2\lambda_4 = 1 \\
 & \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1 \\
 & 2\mu_1 + \mu_2 + 4\mu_3 + 2\mu_4 = 1 \\
 & 4\mu_1 + \mu_2 + 2\mu_3 + 2\mu_4 \geq 1 \\
 & \mu_1 + \mu_2 + \mu_3 + \mu_4 = 1 \\
 & \lambda_j \geq 0, \mu_j \geq 0 \quad j = 1, \dots, 4
 \end{aligned}$$

By solving this model, the optimal overall efficiency score for DMU_B is achieved as $\theta_o^{B*} = 1$.

Similarly, the new model can evaluate the overall efficiency scores for the input-oriented units in Example 3.3. The results of this calculation are reported in Table 2.

Table 2. Overall efficiency scores for four DMUs in example 3.3

DMU	A	B	C	D
θ_o^*	1	1	1	0.5

Thus, DMUs A, B, and C are efficient. Note that the obtained scores for overall efficiency were in keeping with the level of intermediate products computed at the level z_o .

If we allow intermediate products in model (4.1) to change the convex hull, then further improvements in the optimal solution would be possible. We expressed this improvement in Section 6 as free intermediate products.

The dual of Model (4.1) (multiplier formation) can be presented as follows:

$$\begin{aligned}
 \max \quad & \sum_{r=1}^s u_r y_{ro} + \sum_{d=1}^D w^1_d z_{do} - \sum_{d=1}^D w^2_d z_{do} + u_0 + v_0 \\
 \text{s.t. } \quad & \sum_{d=1}^D w^1_d z_{dj} - \sum_{i=1}^m v_i x_{ij} + v_0 \leq 0 \quad j = 1, \dots, n \\
 & \sum_{r=1}^s u_r y_{rj} - \sum_{d=1}^D w^2_d z_{dj} + u_0 \leq 0 \quad j = 1, \dots, n \\
 & \sum_{i=1}^m v_i x_{io} = 1 \\
 & v_i \geq 0, \quad u_r \geq 0 \quad m; r = 1, \dots, s \\
 (4.2) \quad & u_0, v_0, w^2, w^1 \text{ free in sign}
 \end{aligned}$$

In Model (4.2) w^1 and w^2 denote the weights of the intermediate products in the first stage and second stage, respectively, and v_0, u_0, u, v denotes unknown weights. It is trivial that the objective function value of this model is less than or equal to 1.

Model (4.2) can be rewritten as following fractional model:

$$\begin{aligned}
 \max \quad & \frac{\sum_{r=1}^s u_r y_{ro} + \sum_{d=1}^D (w^1_d - w^2_d) z_{do} + u_0 + v_0}{\sum_{i=1}^m v_i x_{io} = 1} \\
 \text{s.t.} \quad & \frac{\sum_{d=1}^D w^1_d z_{dj} + v_0}{\sum_{i=1}^m v_i x_{io} = 1} \leq 1 \quad j = 1, \dots, n \\
 & \frac{\sum_{r=1}^s u_r y_{rj} + u_0}{\sum_{d=1}^D w^2_d z_{dj}} \leq 1 \quad j = 1, \dots, n \\
 & v_i \geq 0, u_r \geq 0 \quad i = 1, \dots, m; r = 1, \dots, s \\
 (4.3) \quad & u_0, v_0, w^2, w^1 \quad \text{free in sign}
 \end{aligned}$$

The first and second sets of constraints in Model (4.3) show the relative efficiency of input-oriented units for DMU_o in the first stage and the second stage, respectively. It should be noted that in the objective function of Model (4.3), the value $(w^1 - w^2)$ in the numerator gives two different roles for z_o . If this value is non-negative, the performance of z_o as output is more effective; otherwise, it is used as input.

By considering equal sign in the constraints related to intermediate products of Model (4.1), the dual variables corresponding to these constraints, namely w^1 and w^2 have no restriction signs. This means that the intermediate values can be measured with positive, negative, or even zero amounts. Thus, the system is allowed to measure intermediate products with different patterns, without considering their input or output roles. Therefore, the new model does not restrict the pricing of intermediate products.

However, in model (2.1), proposed by Kao and Hwang [12], which is equivalent to the dual of model (2.5) ([4]), same non-negative weights are assigned to two intermediate products constraints. Therefore, the obtained values of the intermediate products in model (2.1) had less flexible and imposed targeted pricing methods on the system. In fact, model (4.2) is somewhat similar to model (2.4) ([3]), with the difference that in model (2.4) the same positive weights are assigned to the intermediate products and considering predetermined weights by the decision-maker in model (2.4) are caused differences in the constraints of the normalized equations of these two models.

It should be noted that model (4.1) can also be used to assess the overall efficiency in output-oriented units. The difference is that we should replace the minimum contraction in inputs with the maximum expansion in outputs.

5. Divisional metric converter Product ID0000019 Efficiency in T_{CHI}

For calculating the divisional efficiency, we use the production possibility sets (3.1) and (3.2). The efficiency scores in input-oriented units for DMU_o in the first stage, and its dual model can be computed by the following models:

$$\begin{aligned}
 \theta_1 = \min \theta \\
 \text{s.t.} \quad & \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io} \quad i = 1, \dots, m \\
 & \sum_{j=1}^n \lambda_j z_{dj} = z_{do} \quad d = 1, \dots, D \\
 (5.1) \quad & \sum_{j=1}^n \lambda_j = 1 \quad \lambda_j \geq 0 \quad j = 1, \dots, n
 \end{aligned}$$

$$\begin{aligned}
D_1 = \max \quad & \sum_{d=1}^D w^1_d z_{do} + v_0 \\
\text{s.t.} \quad & \sum_{d=1}^D w^1_d z_{dj} - \sum_{i=1}^m v_i x_{ij} + v_0 \leq 0 \quad j = 1, \dots, n \\
& \sum_{i=1}^m v_i x_{io} = 1 \\
& v_i \geq 0 \quad i = 1, \dots, m \\
(5.2) \quad & v_0, w^1 \text{ free in sign}
\end{aligned}$$

In addition, the efficiency scores for input-oriented units in the second stage can be obtained by the following models which are dual of each other:

$$\begin{aligned}
\theta_2 = \min \theta \\
\text{s.t.} \quad & \sum_{j=1}^n \mu_j z_{dj} = \theta z_{do} \quad d = 1, \dots, D \\
& \sum_{j=1}^n \mu_j y_{rj} \geq y_{ro} \quad r = 1, \dots, s \\
& \sum_{j=1}^n \mu_j = 1 \\
(5.3) \quad & \mu_j \geq 0 \quad j = 1, \dots, n
\end{aligned}$$

$$\begin{aligned}
D_2 = \max \quad & \sum_{r=1}^s u_r y_{ro} + u_0 \\
\text{s.t.} \quad & \sum_{r=1}^s u_r y_{rj} - \sum_{d=1}^D w^2_d z_{dj} + u_0 \leq 0 \quad j = 1, \dots, n \\
& \sum_{d=1}^D w^2_d z_{do} = 1 \\
& u_r \geq 0 \quad r = 1, \dots, s \\
(5.4) \quad & u_0, w^2 \text{ free in sign}
\end{aligned}$$

The interpretation of the models (5.3) and (5.4) are similar to standard models of DEA. The only difference is using the equality restriction in the output constraints of the first step of model (5.1) and in the input constraints of model (5.3). The dual variables corresponding to these constraints are free in sign, but the objective function values of models (5.1) and (5.3) are between zero and one, so based on the duality theorem, the problem is always bounded.

In general, input-based (output-based) models with a two-stage network structure given based on the concept of convex hull in intermediate products do not give information about divisional efficiency. This challenge may be due to using fixed intermediate products, or not using the optimal intermediate products in computation of the divisional efficiency. Indeed, the efficiency score of the first stage may represent the overall efficiency score.

6. Frontier projection in T_{CHI}

In this section we introduce a new network DEA model that gives the efficient projection of inefficient DMUs.

Note that according to the multiplier-based network DEA models, it is not possible to determine the efficiency frontier or the frontier projection of units under assessment. Thus, it is not possible to determine the amount of saving in inputs while keeping the current outputs and also to determine the amount of maximum increased outputs with fixed input values.

Chen et al. [5] expressed that the resulting projections of the dual model (2.1) fail on the efficiency frontier. Therefore, they proposed model (2.5), which is equivalent to the dual of model (2.1). The key point of their model was to modify data of frontier projections with proper adjustments to the intermediate products. They replaced constraints related to intermediate measures in dual of model (2.1) with two sets of constraints, so that the right side of both constraints were replaced with a set of non-negative variables (\tilde{z}_{do}). This model not also provides the frontier projections for inefficient DMUs, but also gives an overall efficiency score.

In order to determine the frontier projections in T_{CHI} , we replace the amounts on the right side in both intermediate products constraints in model (4.1) with a set of the same non-negative variables. In this case we propose the following linear programming model:

$$\begin{aligned}
 & \min \theta \\
 & s.t. \quad \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io} \quad i = 1, \dots, m \\
 & \quad \quad \sum_{j=1}^n \lambda_j z_{dj} = \tilde{z}_{do} \quad d = 1, \dots, D \\
 & \quad \quad \sum_{j=1}^n \mu_j z_{dj} = \tilde{z}_{do} \quad d = 1, \dots, D \\
 & \quad \quad \sum_{j=1}^n \mu_j y_{rj} \geq y_{ro} \quad r = 1, \dots, s \\
 & \quad \quad \sum_{j=1}^n \mu_j = \sum_{j=1}^n \lambda_j = 1 \\
 & \quad \quad \lambda_j, \mu_j \geq 0 \quad j = 1, \dots, n \\
 & \quad \quad \tilde{z}_{do} \geq 0 \quad d = 1, \dots, D
 \end{aligned}
 \tag{6.1}$$

Note that in the Model (6.1), \tilde{z}_{do} denotes an unknown variable. It indicates an optimal amount of intermediate products, produced in the first stage and consumed in the second stage. The projection point for DMU_o is given based upon optimal solution of Model (6.1) as $(\theta^* x_{io}, \tilde{z}_{do}^*, y_{ro})$.

6.1. Theorem. *The projection point for unit under assessment by model (6.1) is overall input-oriented efficient with convex intermediate products.*

The proof of Theorem (6.1) appears in the Appendix.

The dual of model (6.1) can be expressed as follows:

$$\begin{aligned}
 \max \quad & \sum_{r=1}^s u_r y_{ro} + u_0 + v_0 \\
 \text{s.t.} \quad & \sum_{d=1}^D w_d^1 z_{dj} - \sum_{i=1}^m v_i x_{ij} + v_0 \leq 0 \quad j = 1, \dots, n \\
 & \sum_{r=1}^s u_r y_{rj} - \sum_{d=1}^D w_d^2 z_{dj} + u_0 \leq 0 \quad j = 1, \dots, n \\
 & \sum_{i=1}^m v_i x_{io} = 1 \\
 & \sum_{d=1}^D w_d^1 - \sum_{d=1}^D w_d^2 \geq 0 \\
 & v_i \geq 0, u_r \geq 0 \quad i = 1, \dots, m; r = 1, \dots, s \\
 (6.2) \quad & u_0, v_0, w^2, w^1 \text{ free in sign}
 \end{aligned}$$

The fractional program of model (6.2) can be expressed as:

$$\begin{aligned}
 \max \quad & \frac{\sum_{r=1}^s u_r y_{ro} + u_0 + v_0}{\sum_{i=1}^m v_i x_{io}} \\
 \text{s.t.} \quad & \frac{\sum_{d=1}^D w_d^1 z_{dj} + v_0}{\sum_{i=1}^m v_i x_{ij}} \leq 1 \quad j = 1, \dots, n \\
 & \frac{\sum_{r=1}^s u_r y_{rj} + u_0}{\sum_{d=1}^D w_d^2 z_{dj}} \leq 1 \quad j = 1, \dots, n \\
 & \sum_{d=1}^D w_d^1 - \sum_{d=1}^D w_d^2 \geq 0 \\
 & v_i \geq 0, u_r \geq 0 \quad i = 1, \dots, m; r = 1, \dots, s \\
 (6.3) \quad & u_0, v_0, w^2, w^1 \text{ free in sign}
 \end{aligned}$$

Here, \tilde{z}_{do} imposes the third set of constraints ($\sum_{d=1}^D w_d^1 - \sum_{d=1}^D w_d^2 \geq 0$) to model (6.3). Note that these constraints are not redundant. This means that the cost of intermediate products considered as inputs is smaller or equal than to the cost when the same products were considered as outputs.

In fact, the problem is optimized in such a way that the price of providing intermediate products in the second stage equals, at most, to the price of selling the same product in the first stage.

7. Illustrative examples

In this section, the suggested models are used to assess overall and divisional efficiency scores and to determine frontier projections. In addition, we compare the findings of this study with some other two-stage network models.

First, we consider the data given in Table 1 and solve the models (6.1) and (2.5), which were proposed by Chen et al. [5], under the assumption of VRS. Then, using the results, we calculate the frontier projection units. The results are reported in Table 3.

The results show that the frontier projections determined by model (6.1) are exactly the same those obtained by model (2.5). In addition, when calculating the overall

efficiency score of model (4.1) for DMU_B , due to applying the restriction on the intermediate products, this unit is efficient. However, with permission to change the convex hull of the intermediate products in model (6.1) the possibility of further abatement is created in inputs. Therefore, DMU_B is inefficient under the model (6.1).

Table 3. Frontier projection results for four units in example 3.3

DMU	MODEL (6.1)			MODEL (2.5)		
	$\theta^* \mathbf{x}$	$\tilde{\mathbf{z}}^*$	\mathbf{y}	$\theta^* \mathbf{x}$	$\tilde{\mathbf{z}}^*$	\mathbf{y}
A	1	2	4	1	2	4
B	1	2	1	1	2	1
C	1	2	2	1	2	2
D	1	2	2	1	2	2

To examine divisional efficiency scores in Example 3.3, we use Models (5.2) and (5.4). To compare results, we solve the proposed models by Kao and Hwang [13], under VRS assumption. The results are reported in Tables 4 and 5.

As can be seen in Tables 4 and 5, the efficiency scores of first stage given by the model proposed in this study are greater than or equal to those obtained based on the model proposed by Kao and Hwang [13].

Table 4. Efficiency scores of four units in first stage for example 3.3

DMU	MODEL (5.2)	MODEL (2.2)
A	1	1
B	1	0.5
C	1	1
D	0.5	0.5

Table 5. Efficiency scores of four units in second stage for Example 3.3

DMU	MODEL (5.4)	MODEL (2.3)
A	1	1
B	1	1
C	0.33333	0.5
D	0.66667	0.5

7.1. Example. In this example, we evaluate the overall efficiency and divisional efficiency for a two-stage system using 12 DMUs. Two inputs, three intermediate products, and two outputs using hypothetical values are used in this evaluation. The data set is shown in Table 6.

Table 6. Data set of Example 7.1

DMU	x_1	x_2	z_1	z_2	z_3	y_1	y_2
1	61	5	3	12	4	10	75
2	14	23	15	21	90	3	42
3	25	10	6	50	12	6	6
4	8	4	41	3	6	7	18
5	53	20	18	14	55	90	10
6	22	70	19	9	11	1	12
7	17	55	40	20	15	30	14
8	33	82	91	16	8	25	16
9	25	45	6	10	78	7	38
10	66	19	16	1	10	66	20
11	21	64	9	8	15	22	42
12	30	75	70	12	9	36	18

The calculated overall and divisional efficiency scores of the units using models (4.1), (5.2), and (5.4) are reported in Table 7. The overall efficiency scores in most units are the same and are equal to one. Therefore, most units are efficient and lie on the VRS frontier, when they are evaluated by the new network model. These results are not surprising, as the production possibility set with the convex hull in intermediate products is limited. Significantly, the results shown in Table 7 revealed the equality of the overall efficiency scores with the efficiency scores for the first stage.

Table 7. Overall and divisional efficiency scores results in Example 7.1

DMU	Overall	stage 1	stage 2
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	0.7860	0.7860	1
6	0.9077	0.9077	0.6491
7	1	1	0.4292
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	0.9346
12	0.7816	0.7816	0.9346

By using the optimal solution for model (6.1) for DMU_o , the computed efficient projection of $(\theta^* x_{io}, \tilde{z}_{do}^*, y_{ro})$ revealed the improved activity shown in Table 8. Clearly, with the allowable change in the convex hull of intermediate products, the possibility of further improvement is created in the units.

Table 8. Frontier projection results in Example 7.1

DMU	$\theta^* x_1$	$\theta^* x_2$	\tilde{z}_1^*	\tilde{z}_2^*	\tilde{z}_3^*	y_1	y_2
1	61	5	3	12	4	10	75
2	10.48	17.22	35.87	9.88	6.87	3	42
3	10	4	41	3	6	6	6
4	8	4	41	3	6	7	18
5	41.65	15.72	18	14	55	90	10
6	7.98	25.41	41	3	6	1	12
7	8.48	27.44	39.73	4.22	7.67	30	14
8	8.38	20.82	40.01	3.95	7.31	25	16
9	10.07	18.13	36.72	8.73	6.73	7	38
10	30.03	8.645	23.70	7.75	24.70	66	20
11	10.62	32.38	35.56	10.30	6.93	22	42
12	8.61	21.52	39.40	4.54	8.11	36	18

8. Conclusion

Conventional DEA models with a two-stage network structure utilizing intermediate products, have different behaviors that this duality could impose limiting conditions on the pricing system. In this paper, we proposed uniform behavior using these intermediate products. For this purpose, we introduced a new overall production possibility set under the assumption of VRS, considering the convex hull of intermediate products. In addition, we proposed a network DEA model to assess the overall efficiency score and frontier projections. Then, we explained that the use of equality constraints in the intermediate product models decreased disposability, but due to considering separate and free variables in sign, or w^1 and w^2 within the dual models, the system allowed to price the intermediate products using different methods. Therefore, the proposed method is more flexible than

conventional DEA models in a two-stage network structure. On the other hand, \tilde{z}_{do} , in assessment of the model for frontier projections shows more compatibility with production assumptions. The main reason is that the model is optimized in such a way that the price of providing intermediate products in the second stage equals, at most, the price of selling the same product in the first stage. We examined the assessment methods for divisional efficiency in T_{CHI} , and explored that an assessment of divisional efficiency was not possible, because the efficiency of the first stage may represent the overall efficiency. Some illustrative examples were then applied to explain and compare the results of the approach presented here with those obtained by other methods.

Appendix

8.1. Theorem. *The first stage technology or T_1 , which is defined in set (3.1) is the minimal set that contains all observations and satisfies the axioms of strong disposability of inputs and convexity.*

Proof. Assume technology T satisfies the axioms (A1)-(A3). We show that $T_1 \subseteq T$. Namely, if activity $(x_1, z_1) \in T_1$ satisfies (A2) and (A3) with some vectors $\lambda \in R_+^N$ then $(x_1, z_1) \in T$. Let,

$$(x_1, z_1) \in T_1, \exists \lambda_1, \dots, \lambda_n, \quad \sum_j \lambda_j = 1 : \begin{cases} x_1 \geq \sum_j \lambda_j x_j \\ z_1 = \sum_j \lambda_j z_j \end{cases}$$

Since T , satisfy (A1) then for any $(x_j, z_j) \in T, j = 1, \dots, N$. Also, T satisfies (A2) then we have :

$$\sum_j \lambda_j (x_j, z_j) \in T \Rightarrow \left(\sum_j \lambda_j x_j, \sum_j \lambda_j z_j \right) \in T, \sum_j \lambda_j = 1.$$

Finally, T must satisfy the strong disposability in inputs then, $(x_1, z_1) \in T$. Then, the proof is completed. \square

8.2. Theorem. *The projection point for units under assessment given model (6.1) is overall input-efficient with intermediate convex products.*

Proof. The efficiency of projection point obtained for DMU_o based on model (6.1) namely $(\theta^* x_{io}, \tilde{z}_{do}^*, y_{ro})$, is evaluated by solving model (4.1).

We have :

$$\begin{aligned} & \min \hat{\theta} \\ & s. t. \quad (\hat{\theta}(\theta^* x_{io}), \tilde{z}_o^*, y_o) \in T_{CHI} \end{aligned}$$

We claim, $\hat{\theta}^* = 1$. Suppose not and let, $\hat{\theta}^* < 1$, (contrary hypothesis). Thus ,

$$\exists \hat{\lambda} \geq 0, 1\hat{\lambda} = 1, \hat{\mu} \geq 0, 1\hat{\mu} = 1 \rightarrow \begin{cases} \sum_j \hat{\lambda}_j x_j \leq \hat{\theta}(\theta^* x_o) \\ \sum_j \hat{\lambda}_j z_j = \tilde{z}_o^* = \sum_{j=1}^N \lambda_j^* z_j \\ \sum_j \hat{\mu}_j z_j = \tilde{z}_o^* = \sum_{j=1}^N \mu_j^* z_j \\ \sum_j \hat{\mu}_j y_j \geq y_o \end{cases}$$

Then $(\hat{\theta}\theta^*, \hat{\lambda}, \hat{\mu})$, is a feasible solution of model (4.1).

On the other hand according to contrary hypothesis we have $\hat{\theta}^*\theta^* < \theta^*$. However, θ^* is part of an optimal solution and this is inconsistent with the optimality of θ^* . Thus, $\hat{\theta}^* = 1$ and the proof is completed. \square

References

- [1] Charnes, A. Cooper, W.W. Rhodes, E. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2, 429–444. 1978.
- [2] Charnes, A. Cooper, W.W. Golany, B. Halek, R. Klopp, G. Schmitz, E. Thomas, D., Two phase data envelopment analysis approach to policy evaluation and management of army recruiting activities : Trade offs between Joint Services and Army Advertising. Research Report CCS no. 532, Center for Cybernetic Studies, The University of Texas at Austin, Texas. 1986.
- [3] Chen, Y. Cook, W.D. Li, N. Zhu, J. Additive efficiency decomposition in two stage DEA. *European Journal of Operational Research* 96, 1170–1176. 2009a.
- [4] Chen, Y. Liang, L. Zhu, J. Equivalence in two-stage DEA approaches. *European Journal of Operational Research* 193 (2), 600-604. 2009b.
- [5] Chen, Y. Cook, W.D. Zhu, J. Deriving the DEA frontier for two-stage processes. *European Journal of Operational Research* 202, 138–142. 2010.
- [6] Chen, Y. Cook, W. D. Kao, C. Zhu, J., Network DEA pitfall : Divisional efficiency and frontier projection under general network structures. *European Journal of Operational Research* 226, 507-515. 2013.
- [7] Cook, W. D. Seiford L. M. Data envelopment analysis (DEA) thirty years on. *European Journal of Operational Research* 192, 1-17. 2009.
- [8] Ebrahimnejad, A., Tavana, M., Hosseinzadeh Lotfi, F., Shahverdi R. and Yousefpour, M., A three-stage data envelopment analysis model with application to banking industry, *Measurement*, 49,308-319. 2014a.
- [9] Ebrahimnejad E. and Tavana, M., An interactive MOLP method for identifying target units in Output-Oriented DEA models : The NATO enlargement problem, *Measurement*, 52, 124-134. 2014b.
- [10] Ebrahimnejad, A. , Bagherzadeh, M. R., Data envelopment analysis approach for discriminating efficient candidates in voting systems by considering the priority of voters, *Hacettepe Journal of Mathematics and Statistics*, 49 (1), 165-180. 2016.
- [11] Hatami-Marbini, A. Ebrahimnejad, A. Lozano, S. Fuzzy efficiency measures in data envelopment analysis using lexicographic multiobjective approach, *Computers & Industrial Engineering*, 105, 362–376. 2017.
- [12] Kao, C. Hwang, S.-N. Efficiency decomposition in two-stage data envelopment analysis : an application to non-life insurance companies in Taiwan. *European Journal of Operational Research* 185 (1), 418–429. 2008.
- [13] Kao, C. Hwang, S.-N. Decomposition of technical and scale efficiencies in two-stage production systems. *European Journal of Operational Research* 211,515–519. 2011.
- [14] Lovell, C.A.K. Walters, L.C. Wood, L.L. Stratified models of education production using modified DEA and regression analysis. In : Charnes, A. Liang, L. Cook, W.D. Zhu, J. 2008. *DEA models for two-stage processes : game approach and efficiency decomposition. Naval Research Logistics* 55, 643–653. 1994.
- [15] Mottaghi, A. A. Ebrahimnejad, R. Ezzati, E. Khorram, A data envelopment analysis based approach for target setting and resource allocation : application in gas companies, *Hacettepe Journal of Mathematics and Statistics*, 45(5), 1553 -1578. 2016.
- [16] Rho, s. An, J. Evaluating the efficiency of a two stage production process Using data envelopment analysis. *International Transactions in Operational Research* 14, 395-410. 2007.
- [17] Seiford, L.M. Zhu, J. Profitability and marketability of the top 55 US commercial banks. *Management Science* 45, 1270–1288. 1999.
- [18] Sexton, T.R. Lewis, H.F. Two-stage DEA : An application to major league baseball. *Journal of Productivity Analysis* 19, 227–249. 2003.
- [19] Tavana, M., Ebrahimnejad, A., Santos-Arteaga, F. J., Mansourzadeh, S. M., Kazemi Matin, R. , A hybrid DEA-MOLP model for public school assessment and closure decision in the City of Philadelphia, *Socio-Economic Planning Sciences*, doi.org/10.1016/j.seps.2016.09.003. 2017.
- [20] Tone, K. Tsutsui, M. Network DEA : a slacks-based measure approach. *European Journal of Operational Research* 197 (1), 243–252. 2009.

- [21] Wang, C. H. Gopal, R. Zions. S. Use of data envelopment analysis in assess in information technology impact on firm performance. *Annals of Operation Research*73, 191-213. 1997.

Moments and estimation of reduced Kies distribution based on progressive type-II right censored order statistics

Sanku Dey*, Mazen Nassar[†] and Devendra Kumar^{‡§}

Abstract

Based on progressive type-II censored samples, we first derive the recurrence relations for the single and product moments and then use these results to compute the means and variances of reduced Kies distribution (RKD), a new distribution, recently introduced by [21]. Next, we obtain the maximum likelihood estimators of the unknown parameter and the approximate confidence interval of the RKD. Finally, we consider Bayes estimation under the symmetric and asymmetric loss functions using gamma prior for the shape parameter. We have also derived two-sided Bayes probability interval (TBPI) and the highest posterior density (HPD) credible intervals of this distribution. Monte Carlo simulations are performed to compare the performances of the proposed methods, and a data set has been analyzed for illustrative purposes.

Keywords: Progressive type-II right censored order statistics, Single moments, Product moments, Recurrence relations, Reduced Kies distribution.

2000 AMS Classification: 62F15, 62F10, 62F99

Received : 10.09.2017 *Accepted :* 28.06.2018 *Doi :* 10.15672/HJMS.2018.611

*Department of Statistics, St. Anthony's College, Shillong-793001, Meghalaya, India, Email: sankud66@gmail.com

[†]Department of Statistics, Faculty of Commerce, Zagazig University, Egypt, Email: mezo10011@gmail.com

[‡]Department of Statistics, Central University of Haryana, Haryana, India, Email: devendrastats@gmail.com

[§]Corresponding Author.

1. Introduction

Kies distribution was introduced by [19], in connection with the study of breaking strength of glass. Since then, a little work has taken place related to this distribution in different field of science and technology. An important characteristic of the Kies distribution is that its hazard function is decreasing, increasing and bathtub shaped where Weibull models are inappropriate. [21] considered a special case of the Kies distribution, which they termed as "reduced Kies distribution" and its probability density function (pdf) and cumulative density function (CDF) are given by

$$(1.1) \quad f(x; \beta) = \beta x^{\beta-1} (1-x)^{-\beta-1} e^{-\left(\frac{x}{1-x}\right)^\beta}, \quad 0 < x < 1, \beta > 0$$

and

$$(1.2) \quad F(x; \beta) = 1 - e^{-\left(\frac{x}{1-x}\right)^\beta}, \quad 0 < x < 1, \beta > 0.$$

This distribution can be viewed as a functional form of the Weibull distribution with shape parameter β and it can be useful for modeling data sets with increasing and bathtub shaped hazard rate functions. Simple probability distributions generally do not exhibit bathtub-shaped failure rate, including Weibull, gamma, and log-normal. In most cases, bathtub shaped hazard functions have at least two parameters, whereas reduced Kies distribution has only one parameter which exhibit both increasing and bathtub shaped hazard rate. [21] observed that $RKD(\beta)$ is a better model compared to the Weibull as well as its extended models such as beta Weibull distribution, beta generalised Weibull distribution etc. Interested readers may refer to [22] and [23] for an excellent exposure to the Kies distribution.

In case of complete data, it is necessary to continue the experiment until the last item/product failed. Very often, one may find that quite a number of items have very long lifetimes and the experiment continues for a very long period of time so much so that the results may no longer be of any interest or use. In such situations, it may be desirable to terminate the test prior to failure of all items under test. When test is discontinued prior to failure of all items, resulting observations will be called the censored sample. There exist various types of censored samples including Type-II, progressive Type-II, progressive first-failure censored samples and record values etc.

In this paper, we consider a more general censoring scheme called the progressive type-II right censoring scheme. Progressive type II right censoring is a useful scheme in which a specific fraction of individuals at risk may be removed from the experiment at each of several ordered failure times. There is a large body of literature dealing with progressive type II right censoring. For example, see [8], [10], [9] and [25] and the references therein. A Type-II progressively censored scheme can be expressed as: Suppose that n units are put on life test at time 0 and the experimenter decides before hand the quantity m , the number of failures to be observed. Now at the time of first failure, R_1 units are randomly removed from the remaining $n - 1$ surviving units. At the second failure, R_2 units from the remaining $n - 2 - R_1$ units are randomly removed. The test continues until the m -th failure. At this time, all remaining $R_m = n - m - R_1 - R_2 - \dots - R_{m-1}$ units are removed. In this censoring scheme, R_i and m are previously fixed. The resulting m ordered values, that are obtained as a consequence of this type of censoring are appropriately referred to as progressive Type-II censored ordered statistics. Note that, if $R_1 = R_2 = \dots = R_{m-1} = 0$, so that $R_m = n - m$, this scheme reduces to conventional type II right censoring scheme. Also note that if $R_1 = R_2 = \dots = R_m = 0$, so that $m = n$, the progressively type II censoring scheme reduces to the case of no censoring (ordinary order statistics).

Now in the view of (1.1) and (1.2), we have

$$(1.3) \quad f(x) = \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p x^{\beta+p-1}}{p!} [1 - F(x)],$$

where $(e)_k = e(e+1)\cdots(e+k-1)$ denotes the ascending factorial. This equation will be exploited in order to derive some recurrence relations for the single and product moments of progressive Type-II right censored order statistics from the reduced Kies distribution. If the life times of an item are based on an absolutely continuous distribution function $F(x)$ with probability density function $f(x)$, the joint probability density function of the progressively censored failure times $X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}$, is given by (see [8]).

$$(1.4) \quad \begin{aligned} & f_{X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}}(x_1, x_2, \dots, x_m) \\ &= A(n, m-1) \prod_{i=0}^m f(x_i) [1 - F(x_i)]^{R_i} \\ & \quad -\infty < x_1 < x_2 < \dots < x_m < \infty, \end{aligned}$$

where

$$(1.5) \quad \begin{aligned} & A(n, m-1) \\ &= n(n - R_1 - 1) \cdots (n - R_1 - R_2 - \dots - R_{m-1} - m + 1). \end{aligned}$$

Let X_1, X_2, \dots, X_n be a random sample from the reduced kies distribution with pdf and cdf given in (1.1) and (1.2) respectively. The corresponding progressive type-II right censored order statistics with censoring scheme (R_1, R_2, \dots, R_m) , $m \leq n$ will be

$$X_{1:m:n}^{(R_1, R_2, \dots, R_m)}, X_{2:m:n}^{(R_1, R_2, \dots, R_m)}, \dots, X_{m:m:n}^{(R_1, R_2, \dots, R_m)}.$$

The single moments of the progressive type-II right censored order statistics can be written as follows (see, [8]),

$$(1.6) \quad \begin{aligned} \mu_{i:m:n}^{(R_1, R_2, \dots, R_m)^{(k)}} &= E \left[x_{i:m:n}^{(R_1, R_2, \dots, R_m)^{(k)}} \right] \\ &= A(n, m-1) \int \int \cdots \int_{0 < x_1 < x_2 < \dots < x_m < \infty} x_i^k f(x_1) \\ & \quad \times [1 - F(x_1)]^{R_1} f(x_2) [1 - F(x_2)]^{R_2} f(x_3) [1 - F(x_3)]^{R_3} \cdots f(x_m) \\ & \quad \times [1 - F(x_m)]^{R_m} dx_2 dx_3 \cdots dx_m, \end{aligned}$$

where $f(\cdot)$ and $F(\cdot)$ are given respectively in (1.1), (1.2), and $A(n, m-1)$ as defined in (1.5). When $k = 1$, the superscript in the notation of the mean of the progressive type-II right censored order statistics may be omitted without any confusion.

Recurrence relations for single and product moments for any continuous distribution can be used to compute all means and variances of a distribution. Many authors have obtained the recurrence relation for progressively type-II right censored order statistics for different distributions, see for example [15], [26], [29], [18], [3], [2], [30], [4], [1], [5], [6], [7], [16], [17], [11], [28], [24], [12], [13], [14], [20] and the reference cited in.

The motivation of the paper is three fold: first, we derive recurrence relations for the single and product moments of the corresponding progressive Type-II right censored order statistics. These recurrence relations will allow one for the recursive computation of these moments for all sample sizes and all possible censoring schemes, second is to obtain the maximum likelihood estimators and confidence intervals of the unknown parameter of the model and third is to obtain the Bayes estimator under the symmetric and asymmetric loss functions using gamma prior for the shape parameter and two-sided Bayes

probability interval (TBPI) and the highest posterior density (HPD) credible intervals. The uniqueness of this study comes from the fact that we provide explicit expressions for single and product moments using progressive type-II right censored order statistics along with parameter estimation using frequentist and Bayes.

The outline of this note is as follows: Recurrence relations for single moments of progressive type-II right censored order statistics from RKD are given in section 2. Further, section 3 describes the recurrence relations for product moments of progressive type-II right censored order statistics from RKD. The recurrence algorithm is carried out in section 4. In Section 5, we introduce the maximum likelihood estimation of the unknown parameter along with approximate confidence interval. In Section 6, We consider Bayesian estimation of the unknown parameter along with two-sided Bayes probability interval (TBPI) and the highest posterior density (HPD) credible intervals. A Monte Carlo simulation study is presented in Section 7 to evaluate the performances of the estimation method discussed in Sections 5 and 6. Then, in Section 8, we illustrate the methodology developed in this manuscript and the usefulness of the RKD based on progressive type-II right censored order statistics using a real data example. Finally, some concluding remarks are provided in Section 9.

2. Recurrence relation for single moments

In this section, we establish several new recurrence relations satisfied by the single moments of progressive type-II right censored order statistics from the reduced Kies distribution. These recurrence relations may be used to compute the means and variances of reduced Kies distribution based on progressive type-II right censored order statistics for all sample sizes n and all censoring schemes (R_1, R_2, \dots, R_m) , $m \leq n$.

2.1. Theorem. For $2 \leq m \leq n$ and $k \geq 0$,

$$\begin{aligned} \mu_{1:m:n}^{(R_1, R_2, \dots, R_m)^{(k)}} &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p!(k+\beta+p)} \left\{ (n-R_1-1) \mu_{1:m-1:n}^{(R_1+1+R_2, \dots, R_m)^{(k+\beta+p)}} \right. \\ (2.1) \quad &+ \left. (1+R_1) \mu_{1:m:n}^{(R_1, R_2, \dots, R_m)^{(k+\beta+p)} \right\}. \end{aligned}$$

Proof. From equations (1.5) and (1.6), we have

$$\begin{aligned} \mu_{1:m:n}^{(R_1, R_2, \dots, R_m)^{(k)}} &= A(n, m-1) \int \int \dots \int_{0 < x_1 < x_2 < \dots < x_m < \infty} \\ (2.2) \quad &\times L(x_2) f(x_2) [1-F(x_2)]^{R_2} f(x_3) [1-F(x_3)]^{R_3} \dots f(x_m) \\ &\times [1-F(x_m)]^{R_m} dx_2 dx_3 \dots dx_m, \end{aligned}$$

where

$$\begin{aligned} L(x_2) &= \int_0^{x_2} x_1^k f(x_1) [1-F(x_1)]^{R_1} dx_1 \\ &= \int_0^{x_2} x_1^k \left\{ \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p!} x_1^{\beta+p-1} [1-F(x_1)] \right\} [1-F(x_1)]^{R_1} dx_1 \\ (2.3) \quad &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p!} \int_0^{x_2} x_1^{k+\beta+p-1} [1-F(x_1)]^{R_1+1} dx_1. \end{aligned}$$

Integrating (2.3) by parts, we get after simplification

$$(2.4) \quad = \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (k+\beta+p)} \left\{ [1-F(x_2)]^{R_1+1} x_2^{k+\beta+p} + (R_1+1) \int_0^{x_2} x_1^{k+\beta+p} \right. \\ \left. \times [1-F(x_1)]^{R_1} f(x_1) dx_1 \right\}.$$

Substituting the value of $L(x_2)$ from (2.4) in (2.2) and using (1.6), we have

$$\begin{aligned} \mu_{1:m:n}^{(R_1, R_2, \dots, R_m)^{(k)}} &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (k+\beta+p)} \left\{ \int \int \dots \int x_2^{k+\beta+p} (1-F(x_2))^{R_1+1} f(x_2) \right. \\ &\quad \times (1-F(x_2))^{R_2} \dots f(x_m) (1-F(x_m))^{R_m} \\ &\quad \left. + (1+R_1) \mu_{1:m:n}^{(R_1, R_2, \dots, R_m)^{(k+\beta+p)}} \right\} \\ &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (k+\beta+p)} \left\{ (n-R_1-1) \mu_{1:m-1:n}^{(R_1+1+R_2, \dots, R_m)^{(k+\beta+p)}} \right. \\ &\quad \left. + (1+R_1) \mu_{1:m:n}^{(R_1, R_2, \dots, R_m)^{(k+\beta+p)}} \right\}, \end{aligned}$$

rearranging the above equation gives the required result in (2.1). \square

2.2. Theorem. For $m=1, n=1, 2, \dots$ and $k \geq 0$,

$$(2.5) \quad \mu_{1:1:n}^{(n-1)^{(k)}} = n \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (k+\beta+p)} \mu_{1:1:n}^{(n-1)^{(k+\beta+p)}}.$$

Proof. Similar to the proof of Theorem 2.1. \square

2.3. Remark. We may use the fact that the first progressive Type-II right censored order statistics is the same as the first usual order statistic from a sample of size n , regardless of the censoring scheme employed.

2.4. Theorem. For $2 \leq i \leq m-1, m \leq n$ and $k \geq 0$,

$$(2.6) \quad \begin{aligned} \mu_{i:m:n}^{(R_1, R_2, \dots, R_m)^{(k)}} &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (k+\beta+p)} \left\{ (n-R_1-R_2-\dots-R_i-i) \right. \\ &\quad \times \mu_{i:m-1:n}^{(R_1, R_2, \dots, R_{i-1}, R_i+R_{i+1}+1, R_{i+2}, \dots, R_m)^{(k+\beta+p)}} \\ &\quad + (1+R_i) \mu_{i:m:n}^{(R_1, R_2, \dots, R_m)^{(k+\beta+p)}} \\ &\quad - (n-R_1-R_2-\dots-R_{i-1}-i+1) \\ &\quad \left. \times \mu_{i-1:m-1:n}^{(R_1, R_2, \dots, R_{i-2}, R_{i-1}+R_i+1, R_{i+1}, \dots, R_m)^{(k+\beta+p)}} \right\}. \end{aligned}$$

Proof. Similar to the proof of Theorem 2.1. \square

2.5. Theorem. For $2 \leq m \leq n$, and $k \geq 0$,

$$(2.7) \quad \begin{aligned} \mu_{m:m:n}^{(R_1, R_2, \dots, R_m)^{(k)}} &= \beta (1+R_m) \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (k+\beta+p)} \left\{ \mu_{m:m:n}^{(R_1, R_2, \dots, R_m)^{(k+\beta+p)}} \right. \\ &\quad \left. - \mu_{m-1:m-1:n}^{(R_1, R_2, \dots, R_{m-2}, R_{m-1}+R_m+1, R_{i+1}, \dots, R_m)^{(k+\beta+p)}} \right\}. \end{aligned}$$

Proof. Similar to the proof of Theorem 2.1. \square

2.6. Remark. Using these recurrence relations, we can obtain all the single moments of all progressive Type-II right censored order statistics for all sample sizes and censoring schemes (R_1, R_2, \dots, R_m) in a sample recursive manner.

Deductions: For the special case $R_1 = R_2 = \dots = R_m = 0$ so that $m = n$ in which the progressive censored order statistics become the usual order statistics $X_{1:n}, X_{2:n}, \dots, X_{n:n}$, then

(i) From Eq. (2.1): For $k \geq 0$, we get

$$(2.8) \quad \mu_{1:n}^{(k)} = \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (k+\beta+p)} \left\{ \mu_{1:n}^k + (n-1) \mu_{1:n-1:n}^{(1,0,0,\dots,0)^{(k+\beta+p)}} \right\}.$$

(ii) From Eq. (2.6): For $k \geq 0$, we get

$$(2.9) \quad \begin{aligned} \mu_{i:n}^{(k)} &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (k+\beta+p)} \left\{ \mu_{i:n}^{(k+\beta+p)} + (n-i) \mu_{i:n}^{(k+\beta+p)} \right. \\ &\quad \left. - (n-i+1) \mu_{i-1:n}^{(k+\beta+p)} \right\}. \end{aligned}$$

3. Recurrence relation for product moments

In this section, we establish some recurrence relations for product moments of the progressive type-II right censored order statistics from the reduced Kies distribution. The $(i, j)^{th}$ product moment of the progressive type-II right censored order statistics can be written as

$$(3.1) \quad \begin{aligned} \mu_{i,j:m:n}^{(R_1, R_2, \dots, R_m)} &= E \left[x_{i:m:n}^{(R_1, R_2, \dots, R_m)} x_{j:m:n}^{(R_1, R_2, \dots, R_m)} \right] \\ &= A(n, m-1) \int \int \dots \int_{0 < x_1 < x_2 < \dots < x_m < \infty} x_i x_j f(x_1) [1 - F(x_1)]^{R_1} \\ &\quad \times f(x_2) [1 - F(x_2)]^{R_2} \dots f(x_m) [1 - F(x_m)]^{R_m} dx_1 dx_2 dx_3 \dots dx_m, \end{aligned}$$

where $f(\cdot)$ and $F(\cdot)$ are given respectively in (1.1) and (1.2) and $A(n, m-1)$ is defined in (1.5).

3.1. Theorem. For $1 \leq i < j \leq m-1$ and $m \leq n$,

$$(3.2) \quad \begin{aligned} \mu_{i,j:m:n}^{(R_1, R_2, \dots, R_m)} &= \beta (R_j + 1) \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (\beta+p)} \left\{ \mu_{i,j:m:n}^{(R_1, R_2, \dots, R_m)^{(1, \beta+p)}} \right. \\ &\quad + (n - R_1 - 1 - \dots - R_j - j) \\ &\quad \times \mu_{i,j:m-1:n}^{(R_1, R_2, \dots, R_{j-1}, R_j+R_{j+1}+1, \dots, R_m)^{(1, \beta+p)}} \\ &\quad - (n - R_1 - 1 - \dots - R_{j-1} - j + 1) \\ &\quad \left. \times \mu_{i,j-1:m-1:n}^{(R_1, R_2, \dots, R_{j-1}+R_j+1, \dots, R_m)^{(1, \beta+p)} \right\}. \end{aligned}$$

Proof. Using (1.3) and (1.6), we have

$$(3.3) \quad \begin{aligned} \mu_{i,j:m:n}^{(R_1, R_2, \dots, R_m)} &= A(n, m-1) \int \int \dots \int_{0 < x_1 < \dots < x_{j-1} < x_{j+1} < \dots < x_m < \infty} \\ &\quad \times \left\{ \int_{x_{j-1}}^{x_{j+1}} \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p!} x_j^{\beta+p-1} [1 - F(x_j)]^{R_j+1} dx_j \right\} x_i f(x_1) \\ &\quad \times [1 - F(x_1)]^{R_1} \dots f(x_{j-1}) [1 - F(x_{j-1})]^{R_{j-1}} f(x_{j+1}) [1 - F(x_{j+1})]^{R_{j+1}} \\ &\quad \times \dots f(x_m) [1 - F(x_m)]^{R_m} dx_1 dx_2 \dots dx_{j-1} dx_{j+1} \dots dx_m. \end{aligned}$$

Integrating the innermost integral by parts, we obtain

$$\begin{aligned} & \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p!} \int_{x_{j-1}}^{x_{j+1}} x^{\beta+p-1} [1-F(x_j)]^{R_j+1} dx_j \\ &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (\beta+p)} \left\{ x_{j+1}^{\beta+p} [1-F(x_{j+1})]^{1+R_j} \right. \\ & \quad \left. - x_{j-1}^{\beta+p} [1-F(x_{j-1})]^{1+R_j} + (1+R_j) \right. \\ & \quad \left. \times \int_{x_{j-1}}^{x_{j+1}} [1-F(x_j)]^{R_j} f(x_j) x_j^{\beta+p} dx_j \right\}, \end{aligned}$$

which, when substituted into equation (3.3) and using (3.1), we have

$$\begin{aligned} \mu_{i:m:n}^{(R_1, R_2, \dots, R_m)} &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (\beta+p)} \left\{ (n - R_1 - 1 - \dots - R_j - j) \right. \\ & \quad \times \mu_{i,j:m-1:n}^{(R_1, R_2, \dots, R_{j-1}, R_j+R_{j+1}+1, \dots, R_m)}^{(1, \beta+p)} \\ & \quad - (n - R_1 - 1 - \dots - R_{j-1} - j + 1) \\ & \quad \times \mu_{i,j-1:m-1:n}^{(R_1, R_2, \dots, R_{j-1}+R_j+1, \dots, R_m)}^{(1, \beta+p)} \\ & \quad \left. + (R_j + 1) \mu_{i,j:m:n}^{(R_1, R_2, \dots, R_m)}^{(1, \beta+p)} \right\}. \end{aligned}$$

Upon rearrangement of this equation, we obtain the relation in (3.2). \square

3.2. Theorem. For $1 \leq i \leq m-1$ and $m \leq n$,

$$\begin{aligned} \mu_{i:m:n}^{(R_1, R_2, \dots, R_m)} &= \beta \sum_{p=0}^{\infty} \frac{(\beta+1)_p}{p! (\beta+p)} \left\{ (R_m + 1) \mu_{i,m:m:n}^{(R_1, R_2, \dots, R_m)}^{(1, \beta+p)} \right. \\ & \quad (n - R_1 - 1 - \dots - R_{m-1} - m + 1) \\ (3.4) \quad & \quad \left. \times \mu_{i,m-1:m-1:n}^{(R_1, R_2, \dots, R_{m-1}+R_m+1, \dots, R_m)}^{(1, \beta+p)} \right\}. \end{aligned}$$

Proof. Similar to the proof of Theorem 3.1. \square

3.3. Remark. Using these recurrence relations, we can obtain all the product moments of progressive type-II right censored order statistics for all sample sizes and censoring schemes (R_1, R_1, \dots, R_m) .

4. Recursive algorithm

Using the recurrence relations established in Sections 2 and 3, the means and variances of all progressive type-II right censored order statistics from the reduced Kies distribution can be readily computed as follows:

All the first and second order moments with $m = 1$ for all sample sizes n can be obtained by setting $k = 0$ in equation (2.5) and then again setting $k = 1$ in the same equation. Next using equation (2.1), we can determine all the moments of the form $\mu_{1,2:n}^{(R_1, R_2)}$, $n = 2, 3, \dots$, which can in turn be used again with equation (2.1), to determine all moments of the form $\mu_{1,2:n}^{(R_1, R_2)^2}$, $n = 2, 3, \dots$. Equation (2.7) can then be used to obtain $\mu_{2,2:n}^{(R_1, R_2)}$ for all R_1, R_2 and $n \geq 2$ and these values can be used to obtain all moments of the form $\mu_{1,2:n}^{(R_1, R_2)^2}$ by using equation (2.7) again. Equation (2.1) can now be used again to obtain $\mu_{1,3:n}^{(R_1, R_2, R_3)}$, $\mu_{1,3:n}^{(R_1, R_2, R_3)^2}$ for all n, R_1, R_2 and R_3 and equation (2.7) can be used next to obtain all moments of the form $\mu_{2,3:n}^{(R_1, R_2, R_3)}$, $\mu_{2,3:n}^{(R_1, R_2, R_3)^2}$. Finally, equation (2.7) can be

used to obtain all moments of the form $\mu_{3,3:n}^{(R_1, R_2, R_3)}$, $\mu_{3,3:n}^{(R_1, R_2, R_3)^2}$.

This process can be continued until all desired first and second order moments and hence all variances are obtained.

Table 1. Means of progressively Type-II right censored order statistics.

		$\beta = 2$					
$m \downarrow$	$n \downarrow$	Scheme			Mean		
5	2	(0,3)	0.041723	0.072175			
5	2	(3,0)	0.041723	0.250621			
8	2	(6,0)	0.021305	0.196329			
8	2	(0,6)	0.021305	0.032634			
10	2	(8,0)	0.018253	0.172537			
10	2	(0,8)	0.018253	0.047385			
12	2	(10,0)	0.013252	0.177032			
12	2	(0,10)	0.013252	0.039572			
15	2	(13,0)	0.021425	0.172439			
15	2	(0,13)	0.021425	0.036125			
18	2	(16,0)	0.016234	0.169932			
18	2	(0,16)	0.016234	0.029620			
20	2	(18,0)	0.011772	0.175931			
20	2	(0,18)	0.011772	0.023972			
5	3	(2,0,0)	0.041720	0.113935	0.278143		
5	3	(0,0,2)	0.041720	0.075434	0.276428		
8	3	(5,0,0)	0.029243	0.096352	0.306950		
8	3	(0,0,5)	0.029243	0.040270	0.071275		
10	3	(7,0,0)	0.016306	0.099054	0.343772		
10	3	(0,0,7)	0.016306	0.041004	0.062801		
12	3	(9,0,0)	0.023762	0.091843	0.254153		
12	3	(0,0,9)	0.023762	0.034956	0.063072		
15	3	(12,0,0)	0.021934	0.088032	0.262183		
15	3	(0,0,12)	0.021934	0.031835	0.041152		
18	3	(15,0,0)	0.012201	0.091148	0.260841		
18	3	(0,0,15)	0.012201	0.028103	0.036110		
20	3	(17,0,0)	0.009083	0.090581	0.249481		
20	3	(0,0,17)	0.009083	0.024065	0.038214		
5	4	(1,0,0,0)	0.041042	0.090662	0.168825	0.410550	
5	4	(0,0,0,1)	0.041042	0.071167	0.124023	0.201833	
8	4	(4,0,0,0)	0.023810	0.073911	0.205070	0.301149	
8	4	(0,0,0,4)	0.023810	0.052061	0.074052	0.099701	
10	4	(6,0,0,0)	0.016113	0.071152	0.140319	0.290661	
10	4	(0,0,0,6)	0.016113	0.043860	0.061734	0.081195	
12	4	(8,0,0,0)	0.013762	0.073054	0.142933	0.294821	
12	4	(0,0,0,8)	0.013762	0.039928	0.053028	0.063427	
15	4	(11,0,0,0)	0.031830	0.073663	0.147430	0.290271	
15	4	(0,0,0,11)	0.031830	0.031846	0.042835	0.050728	
18	4	(14,0,0,0)	0.015221	0.062283	0.145327	0.311052	
18	4	(0,0,0,14)	0.015221	0.024201	0.034028	0.040475	
20	4	(16,0,0,0)	0.009117	0.057631	0.132802	0.283144	
20	4	(0,0,0,16)	0.009117	0.024240	0.030329	0.047301	
5	5	(0,0,0,0,0)	0.051683	0.070176	0.129787	0.211832	0.403193
8	5	(3,0,0,0,0)	0.027094	0.063837	0.199263	0.199330	0.398422
8	5	(0,0,0,0,3)	0.027094	0.051054	0.074056	0.104373	0.158304
10	5	(5,0,0,0,0)	0.017884	0.061539	0.154028	0.197502	0.379529
10	5	(0,0,0,0,5)	0.017884	0.043805	0.063183	0.081104	0.103304
12	5	(7,0,0,0,0)	0.020523	0.063056	0.212178	0.193653	0.472983
12	5	(0,0,0,0,7)	0.020523	0.038402	0.050117	0.063062	0.084025
15	5	(10,0,0,0,0)	0.020311	0.057324	0.103851	0.190437	0.411947
15	5	(0,0,0,0,10)	0.020311	0.039542	0.046210	0.051183	0.063826
18	5	(13,0,0,0,0)	0.009104	0.050284	0.108204	0.180852	0.415738
18	5	(0,0,0,0,13)	0.009104	0.021095	0.038431	0.049627	0.050642
20	5	(15,0,0,0,0)	0.008630	0.056418	0.108432	0.186549	0.485429
20	5	(0,0,0,0,15)	0.008630	0.039843	0.038754	0.047652	0.087434

Table 2. Means of progressively Type-II right censored order statistics.

			$\beta = 4$				
$m \downarrow$	$n \downarrow$	Scheme			Mean		
5	2	(0,3)	0.052113	0.213631			
5	2	(3,0)	0.052113	0.310047			
8	2	(6,0)	0.030174	0.301174			
8	2	(0,6)	0.030174	0.069876			
10	2	(8,0)	0.031010	0.301632			
10	2	(0,8)	0.031010	0.051042			
12	2	(10,0)	0.021273	0.314250			
12	2	(0,10)	0.021273	0.051634			
15	2	(13,0)	0.021225	0.302284			
15	2	(0,13)	0.021225	0.041229			
18	2	(16,0)	0.020318	0.280847			
18	2	(0,16)	0.020318	0.031054			
20	2	(18,0)	0.019134	0.341186			
20	2	(0,18)	0.019134	0.032832			
5	3	(2,0,0)	0.049102	0.210431	0.401152		
5	3	(0,0,2)	0.049102	0.201043	0.210116		
8	3	(5,0,0)	0.031028	0.191050	0.401432		
8	3	(0,0,5)	0.031028	0.069032	0.190320		
10	3	(7,0,0)	0.029875	0.178320	0.3956375		
10	3	(0,0,7)	0.029875	0.051023	0.0988527		
12	3	(9,0,0)	0.021029	0.149025	0.3783066		
12	3	(0,0,9)	0.021029	0.054193	0.0710345		
15	3	(12,0,0)	0.017843	0.139852	0.3698548		
15	3	(0,0,12)	0.017843	0.039985	0.0501142		
18	3	(15,0,0)	0.014325	0.129984	0.360012		
18	3	(0,0,15)	0.014325	0.029959	0.042073		
20	3	(17,0,0)	0.013062	0.125080	0.345211		
20	3	(0,0,17)	0.013062	0.024643	0.039853		
5	4	(1,0,0,0)	0.048104	0.132273	0.241104	0.498623	
5	4	(0,0,0,1)	0.048104	0.119984	0.198953	0.311032	
8	4	(4,0,0,0)	0.030218	0.112260	0.238421	0.468432	
8	4	(0,0,0,4)	0.030218	0.069852	0.140542	0.150436	
10	4	(6,0,0,0)	0.029844	0.108873	0.22062	0.450417	
10	4	(0,0,0,6)	0.029844	0.058426	0.080113	0.128421	
12	4	(8,0,0,0)	0.020115	0.097431	0.226542	0.449853	
12	4	(0,0,0,8)	0.020115	0.041854	0.072104	0.090438	
15	4	(11,0,0,0)	0.016420	0.093852	0.215490	0.440462	
15	4	(0,0,0,11)	0.016420	0.032052	0.050842	0.070114	
18	4	(14,0,0,0)	0.013054	0.090114	0.210113	0.439912	
18	4	(0,0,0,14)	0.013054	0.027043	0.041032	0.057083	
20	4	(16,0,0,0)	0.012056	0.089215	0.205890	0.436782	
20	4	(0,0,0,16)	0.012056	0.024054	0.037852	0.051201	
5	5	(0,0,0,0,0)	0.047134	0.110895	0.191013	0.310532	0.531052
8	5	(3,0,0,0,0)	0.029042	0.087552	0.172728	0.283820	0.521684
8	5	(0,0,0,0,3)	0.029042	0.062845	0.110405	0.145656	0.213482
10	5	(5,0,0,0,0)	0.024020	0.081095	0.167046	0.282608	0.514021
10	5	(0,0,0,0,5)	0.024020	0.049720	0.078632	0.120676	0.158226
12	5	(7,0,0,0,0)	0.020132	0.078010	0.163101	0.278743	0.510045
12	5	(0,0,0,0,7)	0.020132	0.041201	0.064322	0.090221	0.127035
15	5	(10,0,0,0,0)	0.016310	0.074135	0.160236	0.274888	0.506021
15	5	(0,0,0,0,10)	0.016310	0.032830	0.050623	0.070108	0.091026
18	5	(13,0,0,0,0)	0.013741	0.071565	0.156662	0.264530	0.503621
18	5	(0,0,0,0,13)	0.013741	0.027345	0.041801	0.057221	0.073743
20	5	(15,0,0,0,0)	0.012454	0.070280	0.155381	0.271032	0.502336
20	5	(0,0,0,0,15)	0.012454	0.024628	0.037478	0.051084	0.065541

Table 3. Variances of progressively Type-II right censored order statistics.

		$\beta = 2$					
$m \downarrow$	$n \downarrow$	Scheme	Variance				
5	2	(0,3)	0.011030	0.006372			
5	2	(3,0)	0.011030	0.054530			
8	2	(6,0)	0.000724	0.063226			
8	2	(0,6)	0.000724	0.002816			
10	2	(8,0)	0.000624	0.055025			
10	2	(0,8)	0.000624	0.001284			
12	2	(10,0)	0.000460	0.054761			
12	2	(0,10)	0.000460	0.000821			
15	2	(13,0)	0.000326	0.054628			
15	2	(0,13)	0.000326	0.000521			
18	2	(16,0)	0.000254	0.054554			
18	2	(0,16)	0.000254	0.000361			
20	2	(18,0)	0.000232	0.054524			
20	2	(0,18)	0.000232	0.000292			
5	3	(2,0,0)	0.003040	0.016404	0.070105		
5	3	(0,0,2)	0.003040	0.005572	0.012217		
8	3	(5,0,0)	0.000845	0.015101	0.068601		
8	3	(0,0,5)	0.000845	0.002016	0.003502		
10	3	(7,0,0)	0.000624	0.014801	0.068201		
10	3	(0,0,7)	0.000624	0.001284	0.002120		
12	3	(9,0,0)	0.000460	0.014635	0.068137		
12	3	(0,0,9)	0.000460	0.000821	0.001437		
15	3	(12,0,0)	0.000326	0.014502	0.068013		
15	3	(0,0,12)	0.000326	0.000521	0.000836		
18	3	(15,0,0)	0.000254	0.014531	0.068130		
18	3	(0,0,15)	0.000254	0.000361	0.000562		
20	3	(17,0,0)	0.000242	0.014418	0.068120		
20	3	(0,0,17)	0.000242	0.000292	0.000436		
5	4	(1,0,0,0)	0.003012	0.008173	0.022348	0.075850	
5	4	(0,0,0,1)	0.003031	0.005572	0.012317	0.025712	
8	4	(4,0,0,0)	0.000924	0.006960	0.022133	0.075435	
8	4	(0,0,0,4)	0.000924	0.002016	0.003502	0.005643	
10	4	(6,0,0,0)	0.000624	0.006368	0.020743	0.074245	
10	4	(0,0,0,6)	0.000624	0.001204	0.002120	0.003212	
12	4	(8,0,0,0)	0.000460	0.006514	0.021472	0.075170	
12	4	(0,0,0,8)	0.000460	0.000824	0.001437	0.002118	
15	4	(11,0,0,0)	0.000326	0.006271	0.020446	0.074148	
15	4	(0,0,0,11)	0.000326	0.000520	0.000836	0.001207	
18	4	(14,0,0,0)	0.000254	0.006218	0.020374	0.073875	
18	4	(0,0,0,14)	0.000254	0.000361	0.000561	0.000816	
20	4	(16,0,0,0)	0.000232	0.006167	0.020342	0.073843	
20	4	(0,0,0,16)	0.000232	0.000372	0.000456	0.000641	
5	5	(0,0,0,0,0)	0.003130	0.005572	0.012031	0.025712	0.079204
8	5	(3,0,0,0,0)	0.000924	0.004281	0.011013	0.024319	0.076112
8	5	(0,0,0,0,3)	0.000924	0.002016	0.003512	0.005643	0.008986
10	5	(5,0,0,0,0)	0.000624	0.004056	0.010801	0.025076	0.078511
10	5	(0,0,0,0,5)	0.000624	0.001284	0.002210	0.003212	0.004718
12	5	(7,0,0,0,0)	0.000860	0.003804	0.009748	0.024124	0.077425
12	5	(0,0,0,0,7)	0.000860	0.000822	0.001437	0.002108	0.002938
15	5	(10,0,0,0,0)	0.000326	0.003672	0.009615	0.023920	0.077311
15	5	(0,0,0,0,10)	0.000326	0.000521	0.000836	0.001217	0.001730
18	5	(13,0,0,0,0)	0.000254	0.003617	0.009542	0.023917	0.077221
18	5	(0,0,0,0,13)	0.000254	0.000361	0.000571	0.000806	0.001158
20	5	(15,0,0,0,0)	0.000234	0.003566	0.009511	0.023686	0.077187
20	5	(0,0,0,0,15)	0.000234	0.000290	0.000456	0.000641	0.000852

Table 4. Variances of progressively Type-II right censored order statistics.

		$\beta = 4$					
$m \downarrow$	$n \downarrow$	Scheme	Variance				
5	2	(0,3)	0.005102	0.011251			
5	2	(3,0)	0.005102	0.113001			
8	2	(6,0)	0.002453	0.110560			
8	2	(0,6)	0.002453	0.003716			
10	2	(8,0)	0.002101	0.110162			
10	2	(0,8)	0.002101	0.002325			
12	2	(10,0)	0.000784	0.110702			
12	2	(0,10)	0.000784	0.001611			
15	2	(13,0)	0.000533	0.110441			
15	2	(0,13)	0.000533	0.001144			
18	2	(16,0)	0.000407	0.110303			
18	2	(0,16)	0.000407	0.000664			
20	2	(18,0)	0.000340	0.110245			
20	2	(0,18)	0.000340	0.000536			
5	3	(2,0,0)	0.005103	0.030213	0.130125		
5	3	(0,0,2)	0.005103	0.011150	0.022372		
8	3	(5,0,0)	0.001652	0.027480	0.135586		
8	3	(0,0,5)	0.001652	0.003716	0.006474		
10	3	(7,0,0)	0.002103	0.027016	0.135022		
10	3	(0,0,7)	0.002103	0.002325	0.003901		
12	3	(9,0,0)	0.000784	0.026610	0.134616		
12	3	(0,0,9)	0.000784	0.001611	0.002612		
15	3	(12,0,0)	0.000533	0.026360	0.134466		
15	3	(0,0,12)	0.000533	0.000964	0.001537		
18	3	(15,0,0)	0.000407	0.026224	0.134330		
18	3	(0,0,15)	0.000407	0.000664	0.001135		
20	3	(17,0,0)	0.000340	0.026165	0.134271		
20	3	(0,0,17)	0.000340	0.000536	0.000845		
5	4	(1,0,0,0)	0.005113	0.016016	0.041042	0.151148	
5	4	(0,0,0,1)	0.005113	0.011151	0.022272	0.047320	
8	4	(4,0,0,0)	0.002453	0.013576	0.038602	0.146708	
8	4	(0,0,0,4)	0.002453	0.003706	0.006476	0.011281	
10	4	(6,0,0,0)	0.002110	0.013012	0.038041	0.146145	
10	4	(0,0,0,6)	0.002110	0.002325	0.003901	0.005933	
12	4	(8,0,0,0)	0.000784	0.012707	0.037733	0.145840	
12	4	(0,0,0,8)	0.000784	0.001611	0.002612	0.003848	
15	4	(11,0,0,0)	0.000533	0.013456	0.037483	0.145611	
15	4	(0,0,0,11)	0.000533	0.000964	0.001637	0.002342	
18	4	(14,0,0,0)	0.000417	0.012321	0.037347	0.145453	
18	4	(0,0,0,14)	0.000417	0.000664	0.001135	0.001581	
20	4	(16,0,0,0)	0.000341	0.012362	0.037288	0.145414	
20	4	(0,0,0,16)	0.000341	0.000536	0.000845	0.001271	
5	5	(0,0,0,0,0)	0.005113	0.011251	0.022372	0.047321	0.155405
8	5	(3,0,0,0,0)	0.002453	0.007911	0.019832	0.044860	0.153065
8	5	(0,0,0,0,3)	0.002453	0.003716	0.006476	0.011281	0.017537
10	5	(5,0,0,0,0)	0.002110	0.007435	0.021260	0.045205	0.161301
10	5	(0,0,0,0,5)	0.002110	0.002325	0.003911	0.005933	0.008733
12	5	(7,0,0,0,0)	0.000788	0.007041	0.019163	0.044121	0.152126
12	5	(0,0,0,0,7)	0.000788	0.001611	0.002612	0.003848	0.005412
15	5	(10,0,0,0,0)	0.000533	0.006812	0.018713	0.043740	0.151845
15	5	(0,0,0,0,10)	0.000533	0.000964	0.001637	0.002332	0.003161
18	5	(13,0,0,0,0)	0.000417	0.006654	0.018577	0.043613	0.151710
18	5	(0,0,0,0,13)	0.000417	0.000664	0.001135	0.001581	0.002111
20	5	(15,0,0,0,0)	0.000341	0.006641	0.018518	0.043545	0.151651
20	5	(0,0,0,0,15)	0.000341	0.000536	0.000845	0.001271	0.001662

5. Parameter Estimation under Progressive Type - II Censored Order Statistics

5.1. Maximum Likelihood Method. Let $X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}$ be a progressively Type-II censored sample from $RKD(\beta)$ with (R_1, R_2, \dots, R_m) being the progressive censoring scheme. The likelihood function based on the progressive censored sample is given by

$$(5.1) \quad \begin{aligned} & f_{X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}}(x_1, x_2, \dots, x_m) \\ &= A(n, m-1) \prod_{i=1}^m f(x_i) [1 - F(x_i)]^{R_i}. \end{aligned}$$

where $f(x)$ and $F(x)$ are given respectively by eqns. (1.1) and (1.2). Substituting eqns. (1.1) and (1.2) into eqn. (5.1), the likelihood function is

$$(5.2) \quad \begin{aligned} L(\mathbf{x}|\beta) &= A(n, m-1) \prod_{i=1}^m \left\{ \beta x_i^{\beta-1} (1-x_i)^{-\beta-1} e^{-\left(\frac{x_i}{1-x_i}\right)^\beta} \right\} \\ &\times \left[e^{-\left(\frac{x_i}{1-x_i}\right)^\beta} \right]^{R_i}. \end{aligned}$$

The log of likelihood function is

$$(5.3) \quad \begin{aligned} \ln L(\mathbf{x}|\beta) &= C + m \ln \beta + (\beta - 1) \sum_{i=1}^m \ln(x_i) - (\beta + 1) \sum_{i=1}^m \ln(1 - x_i) \\ &- \sum_{i=1}^m (1 + R_i) \left(\frac{x_i}{1 - x_i} \right)^\beta, \end{aligned}$$

where $C = \ln \{A(n, m-1)\}$. Upon differentiating (5.3) with respect to β and equating to zero, the resulting equation must be satisfied to obtain MLE of β . The equation is given by

$$(5.4) \quad \begin{aligned} \frac{\partial \ln L(\mathbf{x}|\beta)}{\partial \beta} &= \frac{m}{\beta} + \sum_{i=1}^m \ln(x_i) \\ &- \sum_{i=1}^m \ln(1 - x_i) - \sum_{i=1}^m (1 + R_i) \left(\frac{x_i}{1 - x_i} \right)^\beta \ln \left(\frac{x_i}{1 - x_i} \right) = 0. \end{aligned}$$

Using large sample approximation, the asymptotic distribution of the MLE is $[\sqrt{n}(\hat{\beta} - \beta)] \rightarrow N(0, I^{-1}(\beta))$, where $I^{-1}(\beta)$, is the inverse of the observed information matrix of the unknown parameter. The element of the observed information matrix is

$$\frac{\partial^2 \ln L(\mathbf{x}|\beta)}{\partial \beta^2} = \frac{-m}{\beta^2} - \sum_{i=1}^m (1 + R_i) \left(\frac{x_i}{1 - x_i} \right)^\beta \ln^2 \left(\frac{x_i}{1 - x_i} \right).$$

The approximate $100(1 - \tau)\%$ confidence intervals of the parameters β is

$$\hat{\beta} \pm z_{\tau/2} \sqrt{\text{var}(\hat{\beta})},$$

where $\text{var}(\hat{\beta})$ is obtained from $I^{-1}(\beta)$ and $z_{\tau/2}$ is the upper $(\tau/2)$ th percentile of the standard normal distribution.

5.2. Bayesian Estimation. This section discusses the Bayes procedure to derive the point and interval estimates of the parameter β based on progressively Type-II censored data. In our Bayesian analysis, we have assumed three types of loss functions. In this article, the proposed prior for the parameters β is considered as

$$g(\beta) \propto \beta^{a-1} e^{-b\beta}; \quad \beta, a, b > 0.$$

The posterior distribution of β is obtained after simplification as

$$(5.5) \quad \pi(\beta|\mathbf{x}) = \frac{1}{J_1} \beta^{m+a-1} e^{-\left[b\beta + \sum_{i=1}^m (1+R_i) \left(\frac{x_i}{1-x_i}\right)^\beta\right]} \prod_{i=1}^m x_i^{\beta-1} (1-x_i)^{-\beta-1}.$$

where

$$(5.6) \quad J_1 = \int_0^\infty \beta^{m+a-1} e^{-\left[b\beta + \sum_{i=1}^m (1+R_i) \left(\frac{x_i}{1-x_i}\right)^\beta\right]} \prod_{i=1}^m x_i^{\beta-1} (1-x_i)^{-\beta-1} d\beta.$$

We use three different loss functions to obtain the Bayes estimate of the unknown parameter β .

1- The first loss function is the symmetric squared error (SE) loss function. Using SE loss function, the Bayes estimate of the parameter β , denoted by $\tilde{\beta}_{SE}$, is the posterior mean.

2- The second loss function is the asymmetric LINEX loss function proposed by Varian (1975). Under LINEX loss function, the Bayes estimate of the parameter β , denoted by $\tilde{\beta}_{LINEX}$ is given by

$$(5.7) \quad \tilde{\beta}_{LINEX} = -\frac{1}{v} \ln E(e^{-v\beta}),$$

where $v \neq 0$ is constant.

3- The third loss function is the asymmetric general entropy (GE) loss function. The Bayes estimate of the parameter β , denoted by $\tilde{\beta}_{GE}$, is given by

$$(5.8) \quad \tilde{\beta}_{GE} = [E(\beta^{-c})]^{-1/c},$$

where c is the shape parameter of the loss function, provided that $E(\beta^{-c})$ exists.

From (5.8) and using the squared error loss function, the Bayes estimator of β is given by

$$(5.9) \quad \tilde{\beta}_{SE} = \frac{1}{J_1} \int_0^\infty \beta^{m+a} e^{-\left[b\beta + \sum_{i=1}^m (1+R_i) \left(\frac{x_i}{1-x_i}\right)^\beta\right]} \prod_{i=1}^m x_i^{\beta-1} (1-x_i)^{-\beta-1} d\beta.$$

Similarly, from (5.8) and (5.10), the Bayes estimator of β under LINEX loss function is given by

$$(5.10) \quad \begin{aligned} \tilde{\beta}_{LINEX} &= -\frac{1}{v} \ln \left[\frac{1}{J_1} \int_0^\infty \beta^{m+a-1} e^{-\left[v\beta + b\beta + \sum_{i=1}^m (1+R_i) \left(\frac{x_i}{1-x_i}\right)^\beta\right]} \right. \\ &\quad \left. \times \prod_{i=1}^m x_i^{\beta-1} (1-x_i)^{-\beta-1} d\beta \right]. \end{aligned}$$

Using (5.8) and (5.11), the Bayes estimator of β under GE loss function is given by

$$(5.11) \quad \begin{aligned} \tilde{\beta}_{GE} &= \left[\frac{1}{J_1} \int_0^\infty \beta^{m+a-c-1} e^{-\left[b\beta + \sum_{i=1}^m (1+R_i) \left(\frac{x_i}{1-x_i}\right)^\beta\right]} \right. \\ &\quad \left. \times \prod_{i=1}^m x_i^{\beta-1} (1-x_i)^{-\beta-1} d\beta \right]^{-1/c}. \end{aligned}$$

The Bayesian method of interval estimation is more straightforward than the classical method of confidence intervals. Once the posterior distribution of β has been obtained, a symmetric $100(1-\tau)\%$ two-sided Bayes probability interval (TBPI) of β , denoted by $[\beta_L, \beta_U]$, can be obtained by solving the following two equations (see [27], page 208-209).

$$\int_0^{\beta_L} \pi(\theta|x) d\beta = \frac{\tau}{2}, \quad \int_{\beta_U}^\infty \pi(\beta|x) d\beta = \frac{\tau}{2}$$

for the limits β_L and β_U . Now we compute the highest posterior density (HPD) credible intervals for β . Since $\pi(\beta|x)$ is unimodal, the corresponding $100(1 - \gamma)\%$ HPD credible interval $[H_L^\beta, H_U^\beta]$ can be obtained from the simultaneous solution of the following equations

$$(5.12) \quad P(H_L^\beta < \beta < H_U^\beta) = \int_{H_L^\beta}^{H_U^\beta} \pi(\beta|x) d\beta = 1 - \gamma,$$

and

$$(5.13) \quad \pi(H_L^\beta|x) = \pi(H_U^\beta|x).$$

6. Simulation Study

In this section, a simulation study is conducted to study the behaviour of the ML and Bayes estimates under the different loss function by considering $(n, m) = (30, 5), (30, 10), (45, 5), (45, 15), (60, 10)$ and $(60, 20)$ and different values of the parameter β , where $\beta = 1.5, 3$ in all the cases. We have obtained the ML and Bayes estimates by using the following progressive censoring schemes

- Scheme 1: $R_1 = \dots = R_m = \frac{n-m}{m}$.
- Scheme 2: $R_1 = \dots = R_{m-1} = 1$ and $R_m = n - 2m + 1$.
- Scheme 3: $R_1 = \dots = R_{m-1} = 0$ and $R_m = n - m$.

We use the algorithm introduced by [4] to generate progressively censored reduced Kies samples. We consider three types of priors to obtain the Bayes estimates, Prior 0: $a = b = c = d = 0$, which describes the case of non-informative prior. We use Prior 0 to obtain the Bayes estimates for the two values of the parameter β . Prior 1: $a = 3, b = 2$ to obtain the Bayes estimates for $\beta = 1.5$ and Prior 2: $a = 1.5, b = 0.5$ to obtain the Bayes estimates when $\beta = 3$. It is to be noted that, prior 1 and prior 2 describe the case of informative prior. In each setting, we obtain the MLEs and Bayes estimates under SE, LINEX ($\nu = 0.5$) and GE ($c = 0.5$) loss functions. The process is replicated 1000 times. The average values of the estimates, mean squared errors (MSEs), confidence/credible interval lengths and coverage probabilities are obtained and tabulated.

The average values of the estimates and the corresponding MSEs are displayed in Table 5 for $\beta = 1.5$ and in Table 7 for $\beta = 3$. The average confidence/credible interval lengths and the corresponding coverage probabilities are presented in Table 6 for $\beta = 1.5$ and in Table 8 for $\beta = 3$. From Tables 5-8, it is to be noted that the Bayes estimates under SE loss function under Prior 0 is quite close to the MLEs. In terms of MSEs and confidence/credible interval lengths, the Bayes estimates using the informative priors (i.e. Prior 1 and Prior 2) perform better than those based on the non-informative prior (Prior 0) and the MLEs for two parameter values of β . The Bayes estimates under LINEX loss function have the smallest MSEs for all cases when $\beta = 3$, and in some cases when $\beta = 1.5$. For fixed n , when the number of observed failure m increases, the MSEs and the confidence/credible interval lengths decreases in all cases. Comparing the three censoring schemes, it is clear that the MSEs, confidence/credible interval lengths are smaller for Scheme 1 than Schemes 2 and 3.

7. Real Data Analysis

In this section we analyze a real data set given by [31] and also studied by [23]. The original data consists of 40 observations and it describes the strength of a kind of glass, which were measured by three-point flexural method. From the complete data set, we generate three progressively censored samples from $n = 40$ and $m = 10$ according to the

Table 5. Average values of different estimators and the corresponding MSEs (in parentheses) for $\beta = 1.5$.

(n, m)	Scheme	ML Estimate	Bayes Estimates (Prior 0)			Bayes Estimates (Prior 1)		
			SE	LINEX	GE	SE	LINEX	GE
(30,5)	1	1.565(0.134)	1.561(0.132)	1.529(0.119)	1.502(0.119)	1.543(0.093)	1.517(0.085)	1.507(0.085)
	2	1.565(0.136)	1.562(0.135)	1.529(0.121)	1.501(0.121)	1.544(0.094)	1.517(0.086)	1.508(0.086)
	3	1.559(0.147)	1.567(0.149)	1.531(0.133)	1.500(0.133)	1.545(0.099)	1.515(0.091)	1.512(0.091)
(30,10)	1	1.591(0.139)	1.579(0.135)	1.548(0.120)	1.521(0.119)	1.560(0.095)	1.534(0.086)	1.511(0.085)
	2	1.586(0.137)	1.579(0.135)	1.548(0.121)	1.520(0.119)	1.559(0.095)	1.534(0.086)	1.509(0.086)
	3	1.579(0.132)	1.576(0.131)	1.546(0.117)	1.519(0.117)	1.558(0.093)	1.532(0.084)	1.509(0.084)
(45,5)	1	1.539(0.091)	1.545(0.093)	1.524(0.086)	1.503(0.086)	1.535(0.072)	1.515(0.068)	1.503(0.068)
	2	1.539(0.093)	1.546(0.094)	1.524(0.087)	1.503(0.087)	1.536(0.073)	1.515(0.069)	1.504(0.069)
	3	1.539(0.106)	1.552(0.108)	1.526(0.099)	1.504(0.099)	1.538(0.081)	1.516(0.075)	1.504(0.075)
(45,15)	1	1.561(0.099)	1.553(0.098)	1.533(0.091)	1.515(0.091)	1.543(0.077)	1.526(0.072)	1.511(0.072)
	2	1.557(0.100)	1.553(0.099)	1.532(0.092)	1.514(0.092)	1.544(0.078)	1.525(0.073)	1.508(0.072)
	3	1.549(0.092)	1.547(0.092)	1.528(0.085)	1.511(0.085)	1.538(0.073)	1.520(0.069)	1.505(0.070)
(60,10)	1	1.555(0.071)	1.552(0.071)	1.539(0.068)	1.528(0.068)	1.546(0.059)	1.532(0.057)	1.521(0.058)
	2	1.556(0.073)	1.554(0.073)	1.539(0.068)	1.524(0.067)	1.547(0.062)	1.533(0.059)	1.521(0.061)
	3	1.557(0.081)	1.560(0.082)	1.543(0.076)	1.527(0.075)	1.552(0.067)	1.535(0.063)	1.522(0.067)
(60,20)	1	1.541(0.059)	1.534(0.059)	1.521(0.055)	1.507(0.055)	1.529(0.050)	1.515(0.048)	1.504(0.047)
	2	1.539(0.061)	1.536(0.061)	1.522(0.059)	1.509(0.059)	1.531(0.053)	1.517(0.049)	1.505(0.051)
	3	1.539(0.062)	1.538(0.061)	1.521(0.059)	1.510(0.061)	1.533(0.053)	1.519(0.051)	1.508(0.053)

Table 6. Average confidence interval/credible interval lengths and the coverage percentages (in parentheses) for $\beta = 1.5$.

(n, m)	Scheme	Approximate	Symmetric Credible Interval		HPD Interval	
			Prior 0	Prior 1	Prior 0	Prior 1
			(30,5)	1	1.3212(93.70)	1.3468(95.30)
(30,10)	2	1.3362(93.60)	1.3586(95.10)	1.2586(97.10)	1.3929(95.00)	1.2772(96.60)
	3	1.4082(93.30)	1.4359(94.90)	1.3167(97.10)	1.4653(95.30)	1.3296(96.50)
	1	1.2860(93.00)	1.3187(94.40)	1.2089(97.00)	1.3231(94.90)	1.1930(95.70)
(45,5)	2	1.3016(93.00)	1.3321(95.00)	1.2277(96.60)	1.3391(94.70)	1.2270(96.20)
	3	1.2846(93.40)	1.3092(94.70)	1.2154(96.10)	1.3431(94.80)	1.2312(96.10)
	1	1.1211(94.40)	1.0581(95.00)	1.0208(95.90)	1.0857(92.70)	1.0848(95.40)
(45,15)	2	1.1291(94.40)	1.0694(94.90)	1.0245(95.90)	1.1609(94.30)	1.0920(95.40)
	3	1.2092(94.50)	1.1273(94.70)	1.0667(96.00)	1.2417(94.20)	1.1566(95.70)
	1	1.0509(92.70)	0.8137(94.00)	0.9453(95.20)	1.0652(96.40)	1.0129(97.00)
(60,10)	2	1.0660(92.50)	1.0433(94.40)	0.9671(94.80)	1.0799(96.20)	1.0266(96.80)
	3	1.0496(92.90)	1.0269(94.50)	0.9640(95.30)	1.0881(95.90)	1.0351(96.90)
	1	0.9263(93.50)	0.8173(94.80)	0.8526(95.10)	0.9388(94.00)	0.9022(94.30)
(60,20)	2	0.9377(93.30)	0.9261(94.80)	0.8967(95.40)	0.9507(94.20)	0.9132(94.30)
	3	0.9912(93.20)	1.0175(94.90)	0.9486(95.80)	1.0083(93.90)	0.9630(94.40)
	1	0.9069(94.20)	0.8713(95.60)	0.8707(96.60)	0.9287(94.00)	0.8974(95.30)
(60,20)	2	0.9211(93.90)	0.8881(95.10)	0.8705(96.10)	0.9451(94.50)	0.9096(95.30)
	3	0.9083(93.20)	0.8473(95.00)	0.8108(95.80)	0.9590(93.50)	0.9151(95.30)

three censoring schemes discussed in section (6). The generated progressively censored samples are given in the following table

Scheme	Censored data
1	0.477, 0.502, 0.524, 0.525, 0.529, 0.538, 0.546, 0.555, 0.611, 0.624
2	0.477, 0.502, 0.524, 0.525, 0.529, 0.538, 0.539, 0.546, 0.575, 0.600
3	0.477, 0.502, 0.524, 0.525, 0.529, 0.538, 0.539, 0.546, 0.547, 0.549

The MLE and Bayes estimates and the corresponding confidence/credible intervals are reported in Tables 9 and 10, respectively. The Bayes estimates under SE, LINEX and GE loss functions are obtained based on a non-informative prior, because we have no information about the unknown parameter β . Figure 1 display the posterior distribution

Table 7. Average values of different estimators and the corresponding MSEs (in parentheses) for $\beta = 3$.

(n, m)	Scheme	ML Estimate	Bayes Estimates (Prior 0)			Bayes Estimates (Prior 2)		
			SE	LINEX	GE	SE	LINEX	GE
(30,5)	1	3.166(0.558)	3.159(0.555)	3.033(0.449)	3.039(0.492)	3.137(0.458)	3.023(0.378)	3.026(0.409)
	2	3.167(0.573)	3.163(0.571)	3.034(0.461)	3.039(0.505)	3.139(0.469)	3.022(0.386)	3.026(0.419)
	3	3.168(0.647)	3.183(0.657)	3.039(0.517)	3.047(0.575)	3.153(0.525)	3.024(0.424)	3.029(0.465)
(30,10)	1	3.142(0.529)	3.117(0.514)	3.000(0.421)	3.004(0.463)	3.100(0.423)	2.992(0.353)	2.996(0.383)
	2	3.134(0.523)	3.120(0.515)	3.000(0.421)	3.004(0.463)	3.102(0.423)	2.992(0.352)	2.994(0.383)
	3	3.119(0.465)	3.113(0.460)	2.997(0.383)	3.001(0.416)	3.096(0.387)	2.989(0.327)	2.992(0.353)
(45,5)	1	3.095(0.399)	3.108(0.406)	3.019(0.351)	3.025(0.377)	3.097(0.356)	3.012(0.309)	3.014(0.334)
	2	3.096(0.405)	3.109(0.411)	3.021(0.355)	3.023(0.381)	3.097(0.361)	3.013(0.314)	3.014(0.336)
	3	3.102(0.465)	3.127(0.478)	3.025(0.403)	3.029(0.435)	3.109(0.408)	3.015(0.349)	3.020(0.375)
(45,15)	1	3.091(0.316)	3.075(0.311)	2.999(0.275)	2.999(0.294)	3.068(0.277)	2.996(0.247)	2.995(0.263)
	2	3.084(0.319)	3.074(0.316)	2.997(0.279)	2.999(0.295)	3.068(0.282)	2.992(0.251)	2.995(0.268)
	3	3.074(0.302)	3.070(0.301)	2.995(0.267)	2.996(0.282)	3.025(0.397)	2.991(0.241)	2.993(0.254)
(60,10)	1	3.111(0.289)	3.108(0.299)	3.049(0.265)	3.048(0.294)	3.102(0.273)	3.044(0.244)	3.044(0.274)
	2	3.113(0.297)	3.109(0.299)	3.051(0.267)	3.065(0.418)	3.104(0.272)	3.045(0.246)	3.051(0.276)
	3	3.113(0.326)	3.122(0.330)	3.050(0.290)	3.054(0.313)	3.111(0.298)	3.047(0.265)	3.052(0.288)
(60,20)	1	3.089(0.236)	3.075(0.236)	3.022(0.213)	3.024(0.232)	3.068(0.223)	3.018(0.198)	3.021(0.221)
	2	3.086(0.244)	3.081(0.246)	3.020(0.220)	3.024(0.249)	3.077(0.226)	3.017(0.205)	3.022(0.236)
	3	3.080(0.242)	3.078(0.243)	3.021(0.218)	3.021(0.229)	3.072(0.222)	3.018(0.202)	3.021(0.213)

Table 8. Average confidence interval/credible interval lengths and the coverage percentages (in parentheses) for $\beta = 3$.

(n, m)	Scheme	Approximate	Symmetric Credible Interval		HPD Interval	
			Prior 0	Prior 2	Prior 0	Prior 2
(30,5)	1	2.6396(93.00)	3.1169(95.30)	2.4761(97.00)	3.1710(95.10)	2.9930(95.80)
	2	2.6696(93.00)	3.0261(95.10)	2.4920(96.60)	3.1894(95.30)	3.0440(95.80)
	3	2.8161(92.70)	3.4086(94.60)	2.7255(96.00)	3.3785(95.90)	3.0588(95.90)
(30,10)	1	2.5681(94.00)	3.0935(95.90)	2.2692(96.80)	2.7698(95.90)	2.5529(96.20)
	2	2.6016(94.00)	3.1084(96.30)	2.3806(97.00)	2.7716(95.90)	2.6959(96.10)
	3	2.5697(93.90)	2.9045(96.20)	2.5833(97.10)	2.7845(95.60)	2.7456(96.00)
(45,5)	1	2.2436(93.80)	2.6012(94.20)	1.5559(94.50)	2.2890(96.10)	2.2578(96.80)
	2	2.2591(93.60)	2.6268(94.20)	1.6068(94.60)	2.3033(96.10)	2.3044(96.70)
	3	2.4147(93.50)	2.9079(94.60)	2.0136(95.20)	2.4747(96.60)	2.3717(96.70)
(45,15)	1	2.0987(94.90)	2.2901(96.20)	1.6122(95.90)	2.1663(95.20)	2.0994(95.90)
	2	2.1296(94.50)	2.3078(96.20)	1.6767(95.60)	2.5561(95.80)	2.1478(96.50)
	3	2.1004(94.80)	2.2467(96.60)	1.7197(96.30)	2.6493(96.50)	2.2090(96.90)
(60,10)	1	1.8536(93.40)	2.0157(94.10)	1.3082(93.10)	2.0682(94.70)	1.8823(95.30)
	2	1.8766(93.30)	2.1079(93.90)	1.1414(93.00)	2.0808(95.00)	2.0128(95.40)
	3	1.9843(93.20)	2.2553(94.50)	1.3480(93.70)	2.1085(95.00)	2.0843(95.40)
(60,20)	1	1.8142(93.80)	1.9073(95.50)	1.2506(93.20)	1.8319(95.90)	1.7980(96.10)
	2	1.8419(93.80)	1.9539(95.10)	1.4656(93.70)	1.8590(96.20)	1.8238(96.30)
	3	1.8146(93.60)	1.9355(95.00)	2.9850(94.10)	1.8595(96.10)	1.8355(96.70)

of β for different censoring schemes. From Table 10 and Figure 1, it is observed that the symmetric and HDP intervals are coincide because the posterior distribution of β is approximately symmetric.

Table 9. MLE and Bayes Estimates and the corresponding interval/credible intervals (in parentheses) of β .

Scheme	MLE	Bayes (SE)	Bayes (LINEX)	Bayes (GE)
Complete	1.3447	1.3326	1.3283	1.3227
1	1.1403	1.1117	1.0904	1.0513
2	0.7621	0.7617	0.7519	0.7154
3	1.2731	1.2501	1.2205	1.1748

Table 10. The approximate confidence /symmetric and HPD intervals of β .

Scheme	MLE	Symmetric Credible Interval	HPD Interval
Complete	(1.1055, 1.5839)	(1.0763, 1.5909)	(1.0756, 1.5901)
1	(0.5137, 1.7669)	(0.5968, 1.7097)	(0.5600, 1.6898)
2	(0.2714, 1.2528)	(0.3875, 1.1860)	(0.4073, 1.0785)
3	(0.6987, 1.8475)	(0.6439, 1.9976)	(0.6022, 1.9383)

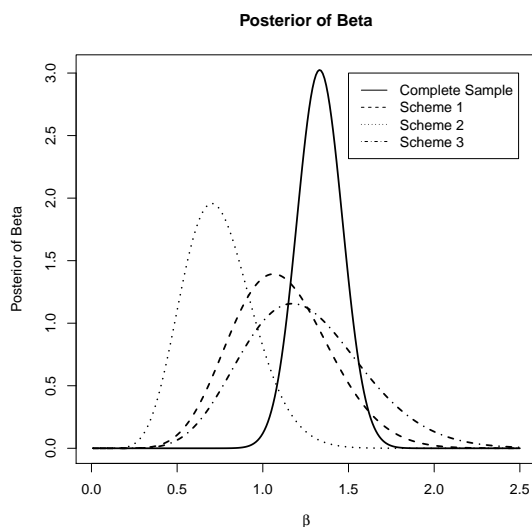
**Figure 1.** Posterior distribution of β for different censoring schemes of [31] data.

Figure 1 shows the effect of censoring schemes on the shape of the posterior distribution of β . It can be seen that the shape of posterior distribution of β based on scheme 2 shows a shift towards the right more than that of complete sample and other censoring schemes. Also, it can be observed from Table 9 that the estimate of β using scheme 2 is the lowest when compared with the complete sample and the other censoring schemes.

8. Concluding Remarks.

In this paper, we have provided explicit expressions and recurrence relations for single and product moments of progressively type-II censored samples of the reduced Kies distribution. We also characterized the distribution by means of recurrence relation. In addition, estimation of unknown parameter of the Reduced Kies distribution has been considered. We have compared the MLEs and different Bayes estimators with respect to the mean squared errors. We have also compared the asymptotic confidence intervals with the two-sided Bayes probability intervals and HPDI obtained from the posterior distribution functions. The simulation study shows that Bayes estimates performs better than the MLEs with regard to mean squared errors. The two-sided Bayes probability

intervals are also of shorter length with competitive coverage percentages of the true parameter than the confidence intervals. Using a real data set, we demonstrated that proposed Bayes estimators perform better than MLE.

Acknowledgements

The authors would like to thank the editor and reviewers for their valuable and very constructive comments, which have greatly improved the contents of the paper.

References

- [1] Aggarwala, R. and Balakrishnan, N. *Recurrence relations for single and product moments of progressively Type-II censored order statistics from a exponential and truncated exponential distribution*, Ann. Inst. Statist. Math., **48**, 757-771, (1996).
- [2] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. *A First Course in Order Statistics*, John Wiley & Sons, New York, (1992).
- [3] Balakrishnan, N. and Malik, H. J. *Order statistics from the linear-exponential distribution, part I: Increasing hazard rate case*, Communications in Statistics, Theory and Methods, **15**, 179-203, (1986).
- [4] Balakrishnan, N. and Sandhu, R. A. *A simple simulational algorithm for generating progressive Type-II censored samples*, Amer. Statist., **49**, 229-230, (1995).
- [5] Balakrishnan, N. and Aggarwala, R. *Recurrence relations for single and product moments of order statistics from a generalized logistic distribution with applications to inference and generalizations to double truncation*, Handbook of Statistics, **17**, 85-126., (1998).
- [6] Balakrishnan, N. and Sultan, K. S. *Recurrence relations for single and product moments of progressively Type-II censored order statistics from a exponential and truncated exponential distribution*, Ann. Inst. Statist. Math., **48**, 757-771, (1998).
- [7] Balakrishnan, N. and Sultan, K. S. *Recurrence relations and identities for moments of order statistics*, In: N. Balakrishnan and C. R. Rao, eds. *Handbook of Statistics*, **16**, 149-228, *Order Statistics: Theory and Methods*, North-Holland, Amsterdam, The Netherlands., (1998).
- [8] Balakrishnan, N. and Aggarwala, R. *Progressive Censoring: Theory, Method and Applications* Birkhauser, Boston., (2000).
- [9] Balakrishnan N, Lin C. T. *On the distribution of a test for exponentiality based on progressively type-II right censored spacings*, J. Stat. Comput. Simul. **73**, 277-283, (2003).
- [10] Balakrishnan, N., Ng, H. K.T. and Kannan, N. *A test of exponentiality based on spacings for progressively type-II censored data*. In: Huber-Carol C et al (eds) *Goodness-of-fit tests and model validity*. Birkhauser, Boston, (2002).
- [11] Balakrishnan, N., Kannan, N., Lin, C. T., and Wu, S. J. S. *Inference for the extreme value distribution under progressive type-II censoring*, Journal of Statistical Computation & Simulation, **74**, 25-45, (2004).
- [12] Balakrishnan, N. *Progressive Censoring Methodology: An Appraisal Test*, **16**, 211-296 (with discussion), (2007).
- [13] Balakrishnan, N., AL-Hussaini, E. K. and Saleh, H. M. *Recurrence relations for moments of progressively censored order statistics from logistic distribution with applications to inference*, Journal of Statistical Planning and Inference, **141**(1), 17-30, (2011).
- [14] Balakrishnan, N. and Saleh, H. M. *Recurrence relations for single and product moments of progressively Type-II censored order statistics from a generalized halflogistic distribution with application to inference*, Journal of Statistical Computation and Simulation , **83**, 1704-1721, (2013).
- [15] Cohen, A.C. *Progressively censored samples in life testing*, Technometrics, **5**, 327-329, (1963).
- [16] David, H. A. and Nagaraja, H. N. *Order Statistics, Third Edition*, John Wiley & Sons, New York. **119**, 171-189, (2003).
- [17] Fernandez, A. J. *On estimating exponential parameters with general type II progressive censoring*, Journal of Statistical Planning and Inference, **121**, 135-147, (2004).

- [18] Joshi, P. C. *Recurrence relations between moments of order statistics from exponential and truncated exponential distributions*, Sankhya Set. B, **39**, 362-371, (1978).
- [19] Kies, J.A. *The strength of glass performance. Naval Research Lab Report No. 5093, Washington, D.C.*, (1958).
- [20] Kumar, D., Dey, S., and Nadarajah, S. *Extended exponential distribution based on order statistics*, Communications in Statistics- Theory and Methods, **46**, 9166-9184, (2017).
- [21] Kumar, C. S and Dharmaja, S. H. S. *On reduced Kies distribution. Collection of Recent Statistical Methods and Applications, (C.S. Kumar, M. Chacko and E.I.A. Sathar, Eds.), 111-123, Department of Statistics, University of Kerala Publishers, Trivandrum.* (2013).
- [22] Kumar, C. S and Dharmaja, S. H. S. *On some properties of Kies distribution*, Metron, **72**, 97-122, (2014).
- [23] Kumar, C. S and Dharmaja, S. H. S. *The Exponentiated Reduced Kies distribution- Properties and Applications*, Communications in Statistics-Theory and Methods, **46(17)**, 8778-8790, (2016).
- [24] Mahmoud, R. M., Sultan, K. S. and Saleh, H. M. *Progressively censored data from the linear exponential distribution, moments and estimation*, Metron, **LXIV(2)**, 199-215, (2006).
- [25] Malik, M.R and Kumar, D. *Relations for moments of progressively Type-II right censored order statistics from Erlang-truncated exponential distribution*, Statistics in Transition New Series, **18**, 651-668, (2017).
- [26] Mann, N. R. *Best linear invariant estimation for Weibull parameters under progressive censoring*, Technometrics, **13**, 521-534, (1971).
- [27] Martz, H., and Waller, R. *Bayesian Reliability Analysis*, New York: John Wiley, (1982).
- [28] Sultan, K. S., Mahmoud, M. R. and Saleh, H. M. *Moments of estimation from progressively censored data of the half logistic distribution*, International Journal of Reliability and Applications, **7(2)**, 187-201, (2006).
- [29] Thomas, D. R. and Wilson, W. M. *Linear order statistic estimation for the two-parameter Weibull and extreme-value distributions from Type-II progressively censored samples*, Technometrics, **14**, 679-691, (1972).
- [30] Viveros, R. and Balakrishnan, N. *Interval estimation of life characteristics from progressively censored data*, Technometrics, **36**, 84-91, (1994).
- [31] Xu, Y., Cheng, L., Zhang, L., Yan, D. and You, C. *Optimization of sample number for Weibull functions of brittle materials strength*, Ceramics International, 27, 239-241, (2001).