HACETTEPE UNIVERSITY
FACULTY OF SCIENCE
TURKEY

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

A Bimonthly Publication
Volume 45 Issue 2
2016

ISSN 1303 5010

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Volume 45 Issue 2
April 2016

A Peer Reviewed Journal
Published Bimonthly by the
Faculty of Science of Hacettepe University

Abstracted/Indexed in

SCI-EXP, Journal Citation Reports, Mathematical Reviews,
Zentralblatt MATH, Current Index to Statistics,
Statistical Theory & Method Abstracts,
SCOPUS, Tübitak-Ulakbim.

ISSN 1303 5010

This Journal is typeset using LaTeX.

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Cilt 45 Sayı 2 (2016)

ISSN 1303 – 5010

KÜNYE
YAYININ ADI: HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS
YIL : 2016 SAYI : 45 - 2 AY : Nisan
YAYIN SAHİBİNİN ADI : H. Ü. Fen Fakültesi Dekanlığı adına Prof. Dr. Bekir Salih
SORUMLU YAZI İŞL. MD. ADI : Prof. Dr. Yücel Tıraş
YAYIN İDARE MERKEZİ ADRESİ : H. Ü. Fen Fakültesi Dekanlığı
YAYIN İDARE MERKEZİ TEL. : 0 312 297 68 50
YAYININ TÜRÜ : Yaygın
BASIMCININ ADI : Hacettepe Üniversitesi Hastaneleri Basımevi.
BASIMCININ ADRESİ : 06100 Sıhhıye, ANKARA.
BASIMCININ TEL. : 0 312 305 1020
BASIM TARİHİ - YERİ : - ANKARA

Hacettepe Journal of Mathematics and Statistics

A Bimonthly Publication – Volume 45 Issue 2 (2016)

ISSN 1303 – 5010

Published by Hacettepe University
Faculty of Science

CONTENTS

Mathematics

Statistics

MATHEMATICS

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 297 – 310

On M-term approximations of the Nikol'skii - Besov class

G. Akishev *

Abstract

In this paper, we consider a Lebesgue space with a mixed norm of periodic functions of many variables. We obtain the exact estimation of the best M-term approximations of Nikol'skii's and Besov's classes in the Lebesgue space with the mixed norm.

Keywords: Lebesgue space, Nikol'skii - Besov class, approximation.

2000 AMS Classification: 41A10, 41A25.

Received : 03.07.2014 *Accepted* : 25.02.2015 *Doi* : 10.15672/HJMS.20164512492

1. Introduction

Let $\overline{x} = (x_1, ..., x_m) \in \mathbb{T}^m = [0, 2\pi)^m$ and $p_j \in [1, +\infty)$, $j = 1, ..., m$. $L_{\bar{p}}(\mathbb{T}^m)$ denotes the space of Lebesgue measureable functions $f(\bar{x})$ defined on $\mathbb{R}^m$, which have 2π period with respect to each variable such that

$$\|f\|_{\bar{p}} = \left[\int_0^{2\pi}\left[\cdots\left[\int_0^{2\pi}|f(\bar{x})|^{p_1}\,dx_1\right]^{\frac{p_2}{p_1}}\cdots\right]^{\frac{p_m}{p_{m-1}}}dx_m\right]^{\frac{1}{p_m}} < +\infty,$$

where $\overline{p} = (p_1, ..., p_m)$, $1 \leq p_j < +\infty$, $j = 1, ..., m$ (see [18], p. 128, [4], p. 54). In the case $p_1 = ... = p_m = p$, we write $L_p(\mathbb{T}^m)$.

Any function $f \in L_1(\mathbb{T}^m) = L(\mathbb{T}^m)$ can be expanded to the Fourier series

$$\sum_{\overline{n}\in\mathbb{Z}^m} a_{\overline{n}}(f)\, e^{i\langle\overline{n},\overline{x}\rangle},$$

where $\{a_{\overline{n}}(f)\}$ are Fourier coefficients of a function $f \in L_1(\mathbb{T}^m)$ with respect to a multiple trigonometric system $\{e^{i\langle\overline{n},\overline{x}\rangle}\}_{\bar{n}\in\mathbb{Z}^m}$ and $\mathbb{Z}^m$ is the space of points in $\mathbb{R}^m$ with integer coordinates.

For a function $f \in L(\mathbb{T}^m)$ and a number $s \in \mathbb{Z}_+ = \mathbb{N} \cup \{0\}$, let us introduce the notation

$$\delta_0(f, \bar{x}) = a_0(f)$$

*Department of Mathematics and Information Technology, Karaganda State University, Universytetskaya 28 , 100028, Karaganda , Kazakhstan;
Email : akishev_g@mail.ru

and
$$\delta_s(f,\overline{x}) = \sum_{\overline{n}\in\rho(s)} a_{\overline{n}}(f) e^{i\langle \overline{n},\overline{x}\rangle},$$
where $\langle \bar{y},\bar{x}\rangle = \sum_{j=1}^{m} y_j x_j$ and
$$\rho(s) = \left\{ \overline{k} = (k_1,...,k_m) \in \mathbb{Z}^m : \quad [2^{s-1}] \le \max_{j=1,\dots,m} |k_j| < 2^s \right\},$$
where $[a]$ is the integer part of the number a.

Let us consider Nikol'skii's and Besov's classes ([4, 7, 18]). Let $1 < p_j < +\infty$, $j = 1,...,m$, $1 \le \theta \le \infty$, $r > 0$, and
$$H^r_{\overline{p}} = \left\{ f \in L_{\overline{p}}(\mathbb{T}^m) : \sup_{s\in\mathbb{Z}_+} 2^{sr} \|\delta_s(f)\|_{\bar{p}} \le 1 \right\},$$
$$B^r_{\bar{p},\theta} = \Big\{ f \in L_{\overline{p}}(\mathbb{T}^m) : \Big(\sum_{s\in\mathbb{Z}_+} 2^{sr\theta} \|\delta_s(f)\|^{\theta}_{\bar{p}} \Big)^{\frac{1}{\theta}} \le 1 \Big\}.$$

It is known that for $1 \le \theta \le \theta_1 \le \infty$ the following holds
$$B^r_{\bar{p},1} \subset B^r_{\bar{p},\theta} \subset B^r_{\bar{p},\theta_1} \subset B^r_{\bar{p},\infty} = H^r_{\bar{p}}.$$

Let $f \in L_{\bar{p}}(\mathbb{T}^m)$ and $\left\{ \bar{k}^{(j)} \right\}_{j=1}^{M}$ be a system of vectors $\bar{k}^{(j)} = (k_1^{(j)},...,k_m^{(j)})$ with integer coordinates. Consider the quantity
$$e_M(f)_{\bar{p}} = \inf_{\bar{k}^{(j)},b_j} \left\| f - \sum_{j=1}^{M} b_j e^{i\langle \bar{k}^{(j)},\bar{x}\rangle} \right\|_{\bar{p}},$$
where b_j is an arbitrary number. The quantity $e_M(f)_{\bar{p}}$ is called the best M-term approximation of a function $f \in L_{\bar{p}}(\mathbb{T}^m)$. For a given class $F \subset L_{\bar{p}}(\mathbb{T}^m)$ let
$$e_M(F)_{\bar{p}} = \sup_{f\in F} e_M(f)_{\bar{p}}.$$
The best M-term approximation was defined by S.B. Stechkin [22]. Estimations of M-term approximations of different classes were provided by R.S. Ismagilov [13], E.S. Belinsky [6], V.E. Maiorov [17], B.S. Kashin [14], R. DeVore [8], V.N. Temlyakov [23], A.S. Romanyuk [19], Dinh Dung [10], D.B. Bazarkhanov [5], L. Duan [11], M. Hansen and W. Sickel [12], S.A. Stasyuk [20, 21], and others (see bibliography in [1], [2], [8], [21], [23]).

For the case $p_1 = ... = p_m = p$ and $q_1 = ... = q_m = q$, R.A. De Vore and V.N. Temlyakov [9] proved the following theorem.

1.1. Theorem. *(see [9]). Let $1 \le p,q,\theta \le \infty$, $r(p,q) = m\left(\frac{1}{p}-\frac{1}{q}\right)_+$ if $1 \le p \le q \le 2$ or $1 \le q \le p < \infty$ and $r(p,q) = \max\left\{\frac{m}{p},\frac{m}{2}\right\}$ in other cases. Then, for $r > r(p,q)$, the following relation holds*
$$e_M(B^r_{p,\theta})_q \asymp M^{-\frac{r}{m}+\left(\frac{1}{p}-\max\left\{\frac{1}{q},\frac{1}{2}\right\}\right)_+},$$
where $a_+ = \max\{a;0\}$.

Moreover, in the case of $m\left(\frac{1}{p}-\frac{1}{q}\right) < r < \frac{m}{p}$ and $1 < p \le 2 < q < \infty$, S.A. Stasyuk [20, 21] proved that $e_M(B^r_{p,\theta})_q \asymp M^{-\frac{q}{2}\left(\frac{r}{m}-\left(\frac{1}{p}-\frac{1}{q}\right)\right)}$.

The main goal of the present paper is to find the order of the quantity $e_M(F)_{\bar{q}}$ for the class $F = B^r_{\bar{p},\theta}$.

Let us denote by $C(p,q,r,y)$ positive quantities, which depend on the parameters in the parentheses, such that the parameters, in general, are distinct in distinct formulas. $A\,(y) \asymp B\,(y)$ means that there are positive numbers C_1 and C_2 such that $C_1 \cdot A\,(y) \leq B\,(y) \leq C_2 \cdot A\,(y)$.

To prove the main results, we need the following auxiliary results.

1.2. Theorem. *(see [24]). Let* $\bar{n} = (n_1, ..., n_m)$, $n_j \in \mathbb{N}$, $j = 1, ..., m$, *and*

$$T_{\bar{n}}(\bar{x}) = \sum_{|k_j| \leq n_j, j=1,\dots,m} c_{\bar{k}} e^{i\langle \bar{k}, \bar{x}\rangle}.$$

Then, for $1 \leq p_j < q_j \leq \infty$, $j = 1, ...m$, *the following inequality holds*

$$\|T_{\bar{n}}\|_{\bar{q}} \leq 2^m \prod_{j=1}^{m} n_j^{\frac{1}{p_j} - \frac{1}{q_j}} \|T_{\bar{n}}\|_{\bar{p}}.$$

1.3. Theorem. *(see [16]). Let* $p \in (1, \infty)$. *Then there exist positive constants* $C_1(p)$ *and* $C_2(p)$ *such that for each function* $f \in L_p(\mathbb{T}^m)$ *the following estimation is valid*

$$C_1(p)\|f\|_p \leq \left\| \Big(\sum_{s=0}^{\infty} |\delta_s(f)|^2 \Big)^{\frac{1}{2}} \right\|_p \leq C_2(p)\|f\|_p.$$

Let Ω_M be a set containing no more than M vectors $\bar{k} = (k_1, ..., k_m)$ with integer coordinates and $P(\Omega_M, \bar{x})$ be any trigonometric polynomial, which consists of harmonics with "indices" in Ω_M.

1.4. Lemma. *(see [2]). Let* $2 < q_j < +\infty$ *and* $j = 1, ..., m$. *Then, for any trigonometric polynomial* $P(\Omega_N)$ *and for any natural number* $M < N$, *there exists a trigonometric polynomial* $P(\Omega_M)$ *such that the following estimation holds*

$$\|P(\Omega_N) - P(\Omega_M)\|_{\bar{q}} \leq C_1(NM^{-1})^{\frac{1}{2}} \|P(\Omega_N)\|_2,$$

and, moreover, $\Omega_M \subset \Omega_N$.

2. Main results

Let us prove the main results.

2.1. Theorem. *Let* $\bar{p} = (p_1, ..., p_m)$, $\bar{q} = (q_1, ..., q_m)$, $1 < p_j \leq 2 < q_j < \infty$, *and* $1 \leq \theta \leq \infty$.

1. If $\sum_{j=1}^{m} \left(\frac{1}{p_j} - \frac{1}{q_j}\right) < r < \sum_{j=1}^{m} \frac{1}{p_j}$, *then*

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \asymp M^{-\left(2\sum_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r - \sum_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{q_j}\right)\right)}.$$

2. If $r = \sum_{j=1}^{m} \frac{1}{p_j}$, *then*

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \asymp M^{-\frac{1}{2}} (\log(1+M))^{1-\frac{1}{\theta}}.$$

3. If $r > \sum_{j=1}^{m} \frac{1}{p_j}$, *then*

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \asymp M^{-\frac{1}{m}\left(r + \sum_{j=1}^{m}\left(\frac{1}{2} - \frac{1}{p_j}\right)\right)}.$$

Proof. Firstly, we are going to consider the upper bound in the first item. Taking into account the inclusion $B^r_{\bar{p},\theta} \subset H^r_{\bar{p}}$, $1 \le \theta < +\infty$, it suffices to prove it for the class $H^r_{\bar{p}}$.

Let $1 \le p_j < q_j < \infty$ and $\mathbb{N}$ be the set of natural numbers. For a number $M \in \mathbb{N}$ choose a natural number n such that $2^{nm} < M \le 2^{(n+1)m}$. For a function $f \in H^r_{\bar{p}}$, it is known that

$$f(\bar{x}) = \sum_{s=0}^{\infty} \delta_s(f, \bar{x})$$

and

$$\|\delta_s(f)\|_{\bar{p}} \le 2^{-sr}, \quad 1 < p_j < \infty, \quad j = 1, ..., m.$$

We will seek an approximation polynomial $P(\Omega_M, \bar{x})$ in the form

$$P(\Omega_M, \bar{x}) = \sum_{s=0}^{n-1} \delta_s(f, \bar{x}) + \sum_{n \le s < \alpha n} P(\Omega_{N_s}, \bar{x}), \tag{1}$$

where the polynomials $P(\Omega_{N_s}, \bar{x})$ will be constructed for each $\delta_s(f, \bar{x})$ in accordance with Lemma 1.4 and the number $\alpha > 1$ will be chosen during the construction.

Let $\sum\limits_{j=1}^{m} (\frac{1}{p_j} - \frac{1}{q_j}) < r < \sum\limits_{j=1}^{m} \frac{1}{p_j}$. Suppose

$$N_s = \left[2^{nm} 2^{s\left(\sum\limits_{j=1}^{m} \frac{1}{p_j} - r\right)} 2^{-n\alpha\left(\sum\limits_{j=1}^{m} \frac{1}{p_j} - r\right)} \right] + 1,$$

where $[y]$ is the integer part of the number y.

Now we are going to show that the polynomials (1) have no more than M harmonics (in terms of order). By the definition of the number N_s, we have

$$\sum_{s=0}^{n-1} \sharp\{\bar{k} = (k_1, ..., k_m) : [2^{s-1}] \le \max_{j=1,...,m} |k_j| < 2^s\} + \sum_{n \le s < \alpha n} N_s \le C 2^{nm} +$$

$$+ \sum_{n \le s < \alpha n} \left(2^{nm} 2^{s\left(\sum\limits_{j=1}^{m} \frac{1}{p_j} - r\right)} 2^{-n\alpha\left(\sum\limits_{j=1}^{m} \frac{1}{p_j} - r\right)} + 1 \right) \le C 2^{nm} + (\alpha - 1)\, n \le C 2^{nm} \asymp M,$$

where $\sharp A$ denotes the number of elements in the set A.

Next, by the property of the norm, we have

$$\|f - P(\Omega_M)\|_{\bar{q}} \le C \left\| \sum_{n \le s < \alpha n} (\delta_s(f) - P(\Omega_{N_s})) \right\|_{\bar{q}} +$$

$$+ \left\| \sum_{\alpha n \le s < +\infty} \delta_s(f) \right\|_{\bar{q}} = J_1(n) + J_2(n). \tag{2}$$

Let us estimate $J_2(n)$. Applying the inequality of different metrics for trigonometric polynomials (Theorem 1.2), we can obtain

$$J_2(n) \le \sum_{\alpha n \le s < +\infty} \|\delta_s(f)\|_{\bar{q}} \le C \sum_{\alpha n \le s < +\infty} 2^{s \sum\limits_{j=1}^{m} \left(\frac{1}{p_j} - \frac{1}{q_j}\right)} \|\delta_s(f)\|_{\bar{p}}.$$

Therefore, taking into account $f \in H^r_{\bar{p}}$ and $\sum\limits_{j=1}^{m}(\frac{1}{p_j} - \frac{1}{q_j}) < r$, we get

$$J_2(n) \le C \sum_{\alpha n \le s < +\infty} 2^{-s\left(r - \sum\limits_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{q_j}\right)\right)} \le C 2^{-n\alpha\left(r - \sum\limits_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{q_j}\right)\right)}. \tag{3}$$

Let us estimate $J_1(n)$. Using the property of the norm, Lemma 1.4 and the inequality of different metrics (Theorem 1.2), we get

$$J_1(n) \le \sum_{n \le s < \alpha n} \|\delta_s(f) - P(\Omega_{N_s})\|_{\bar{q}} \le C \sum_{n \le s < \alpha n} \left(N_s^{-1} 2^{sm}\right)^{\frac{1}{2}} \|\delta_s(f)\|_2 \le$$

$$\le C \sum_{n \le s < \alpha n} \left(N_s^{-1} 2^{sm}\right)^{\frac{1}{2}} 2^{s\sum\limits_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{2}\right)} \|\delta_s(f)\|_{\bar{p}} \le$$

$$\le C \sum_{n \le s < \alpha n} N_s^{-\frac{1}{2}} 2^{s\sum\limits_{j=1}^{m}\frac{1}{p_j}} 2^{-sr} \le$$

$$\le C 2^{-\frac{nm}{2}} 2^{\frac{n\alpha}{2}\left(\sum\limits_{j=1}^{m}\frac{1}{p_j} - r\right)} \sum_{n \le s < \alpha n} 2^{s\left(\sum\limits_{j=1}^{m}\frac{1}{p_j} - r\right)\frac{1}{2}} \le C 2^{-\frac{nm}{2}} 2^{\frac{n\alpha}{2}\left(\sum\limits_{j=1}^{m}\frac{1}{p_j} - r\right)}. \tag{4}$$

Suppose $\alpha = m\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}$. Then, from the inequality (4), we get

$$J_1(n) \le C 2^{-nm\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r - \sum\limits_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{q_j}\right)\right)} \asymp M^{-\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r - \sum\limits_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{q_j}\right)\right)}. \tag{5}$$

For $\alpha = m\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}$, using the inequality (3) and taking into account $2^{nm} \asymp M$, we obtain

$$J_2(n) \le C M^{-\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r - \sum\limits_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{q_j}\right)\right)}. \tag{6}$$

By (5) and (6), we get from the inequality (2) the following

$$\|f - P(\Omega_M)\|_{\bar{q}} \le C M^{-\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r - \sum\limits_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{q_j}\right)\right)},$$

for any function $f \in H^r_{\bar{p}}$ in the case of $\sum\limits_{j=1}^{m}(\frac{1}{p_j} - \frac{1}{q_j}) < r < \sum\limits_{j=1}^{m}\frac{1}{p_j}$.

From the inclusion $B^r_{\bar{p},\theta} \subset H^r_{\bar{p}}$ and the definition of the M-term approximation, it follows that

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \le C M^{-\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r - \sum\limits_{j=1}^{m}\left(\frac{1}{p_j} - \frac{1}{q_j}\right)\right)}$$

in the case of $\sum\limits_{j=1}^{m}(\frac{1}{p_j} - \frac{1}{q_j}) < r < \sum\limits_{j=1}^{m}\frac{1}{p_j}$.

Let us consider the lower bound. We will use the well-known formula (see [19], p. 79)

$$e_M(f)_{\bar{q}} = \inf_{\Omega_M} \sup_{P \in L^{\perp}_M, \|P\|_{\bar{q}'} \le 1} \left|\int_{\mathbb{T}^m} f(\bar{x})\bar{P}(\bar{x})d\bar{x}\right|, \tag{7}$$

where $\bar{q}' = ({q_1}', ..., {q_m}')$, $\frac{1}{q_j} + \frac{1}{{q_j}'} = 1$, $j = 1, ..., m$, and $L^{\perp}_M$ is the set of functions that are orthogonal to the subspace of trigonometric polynomials with harmonics in the set Ω_M.

Consider the function

$$F_{\bar{q},n}(\bar{x}) = \sum_{\max\limits_{j=1,\ldots,m}|k_j|\le 2^{\left[nm\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\right]}} e^{i\langle\bar{k},\bar{x}\rangle}.$$

Let Ω_M be a set of M vectors with integer coordinates. Suppose

$$g(\bar{x}) = F_{\bar{q},n}(\bar{x}) - \sum_{\bar{k}\in\Omega_M}^{*} e^{i\langle\bar{k},\bar{x}\rangle},$$

where the sum $\sum\limits_{\bar{k}\in\Omega_M}^{*}$ contains those terms in the function $F_{\bar{q},n}(\bar{x})$ with indices only in Ω_M. By the inequality (see [18], p. 88)

$$\Big\|\sum_{\max\limits_{j=1,\ldots,m}|k_j|\le 2^l} e^{i\langle\bar{k},\bar{x}\rangle}\Big\|_{\bar{p}} \le C2^{l\sum\limits_{j=1}^{m}(1-\frac{1}{p_j})} \tag{8}$$

and Perseval's equality for $1 < {q_j}' < 2$, $j = 1, ..., m$, we obtain

$$\|g\|_{\bar{q}'} \le \|F_{\bar{q},n}\|_{\bar{q}'} + (2\pi)^{\sum\limits_{j=1}^{m}(\frac{1}{q_j}-\frac{1}{2})}\Big\|\sum_{\bar{k}\in\Omega_M}^{*} e^{i\langle\bar{k},\bar{x}\rangle}\Big\|_2 \le C(2^{\frac{nm}{2}} + M^{\frac{1}{2}}) \le C2^{\frac{nm}{2}}. \tag{9}$$

Now we consider the function

$$P_1(\bar{x}) = C_2 2^{\left(-\frac{nm}{2}\right)} g(\bar{x}). \tag{10}$$

Then (9) implies that the function P_1 satisfies the assumptions of the formula (7) for some constant $C_2 > 0$.

Consider the function

$$f_1(\bar{x}) = C_3 2^{-nm\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-1\right)\right)} F_{\bar{q},n}(\bar{x}). \tag{11}$$

By the inequality (8), we get

$$\sum_{s=0}^{\infty} 2^{sr}\left\|\delta_s\left(f_1\right)\right\|_{\bar{p}} \le$$

$$\le C2^{-nm\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-1\right)\right)} \sum_{s=0}^{\left[nm\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\right]} 2^{sr}2^{s\sum\limits_{j=1}^{m}\left(1-\frac{1}{p_j}\right)} \le C_3.$$

Hence $C_3^{-1}f_1 \in B^r_{\bar{p},1}$.

For the functions (10) and (11), we have, by the formula (7), the following

$$e_M(f_1)_{\bar{q}} \ge \inf_{\Omega_M}\left|\int_{\mathbb{T}^m} f_1\left(\bar{x}\right)\bar{P}_1\left(\bar{x}\right)d\bar{x}\right| \ge$$

$$\ge C2^{-nm\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-1\right)\right)} 2^{-\frac{nm}{2}}\left(\|F_{\bar{q},n}\|_2^2 - M\right) \ge$$

$$\ge C2^{-nm\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)}. \tag{12}$$

Hence, it follows from (12) by the inclusion $B^r_{\bar{p},1} \subset B^r_{\bar{p},\theta}$ that

$$e_M(f_1)_{\bar{q}} \ge C2^{-nm\left(2\sum\limits_{j=1}^{m}\frac{1}{q_j}\right)^{-1}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)}$$

in the case of $\sum_{j=1}^{m}(\frac{1}{p_j}-\frac{1}{q_j})<r<\sum_{j=1}^{m}\frac{1}{p_j}$. So we have proved the first item.

Now we consider the case $r=\sum_{j=1}^{m}\frac{1}{p_j}$. Let $f\in B^r_{\bar{p},\theta}$. Suppose $\alpha=m\left(2\sum_{j=1}^{m}\frac{1}{q_j}\right)^{-1}$ and

$$N_s=\left[2^{nm}n^{\frac{1}{\theta}-1}\left\|\delta_s(f)\right\|_{\bar{p}}2^{sr}\right]+1.$$

Then, by definition of the numbers N_s and Holder's inequality, we obtain

$$\sum_{s=0}^{n-1}\sharp\{\bar{k}=(k_1,...,k_m):[2^{s-1}]\le\max_{j=1,...,m}|k_j|<2^s\}+\sum_{n\le s<\alpha n}N_s\le$$

$$\le C2^{nm}+(\alpha-1)n+2^{nm}n^{\frac{1}{\theta}-1}((\alpha-1)n)^{\frac{1}{\theta'}}\left(\sum_{s=0}^{\infty}\|\delta_s(f)\|_{\bar{p}}^{\theta}2^{sr\theta}\right)^{\frac{1}{\theta}}\le C2^{nm}\asymp M.$$

Suppose $\beta=\max\{q_1,...,q_m\}$. Then

$$J_1(n)=\left\|\sum_{n\le s<\alpha n}(\delta_s(f)-P(\Omega_{N_s}))\right\|_{\bar{q}}\le C\left\|\sum_{n\le s<\alpha n}(\delta_s(f)-P(\Omega_{N_s}))\right\|_{\beta}.$$

Next, by Theorem 1.3, we have

$$J_1(n)\le C\left\|\left(\sum_{n\le s<\alpha n}|\delta_s(f)-P(\Omega_{N_s})|^2\right)^{\frac{1}{2}}\right\|_{\beta}.$$

Since $\beta>2$, then by applying the property of the norm, Lemma 1.4 and the inequality of different metrics for trigonometric polynomials (see Theorem 1.2), we obtain

$$J_1(n)\le\left(\sum_{n\le s<\alpha n}\|\delta_s(f)-P(\Omega_{N_s})\|_{\beta}^2\right)^{\frac{1}{2}}\le C\left(\sum_{n\le s<\alpha n}N_s^{-1}2^{sm}\|\delta_s(f)\|_2^2\right)^{\frac{1}{2}}\le$$

$$\le C\left(\sum_{n\le s<\alpha n}N_s^{-1}2^{sm}2^{2s\sum_{j=1}^{m}(\frac{1}{p_j}-\frac{1}{2})}\|\delta_s(f)\|_{\bar{p}}^2\right)^{\frac{1}{2}}. \tag{13}$$

Next, since $r=\sum_{j=1}^{m}\frac{1}{p_j}$, we have, by the definition of the numbers N_s and using Holder's inequality, the following

$$J_1(n)\le C(2^{-nm}n^{1-\frac{1}{\theta}})^{\frac{1}{2}}\left(\sum_{n\le s<\alpha n}2^{sr}\|\delta_s(f)\|_{\bar{p}}\right)^{\frac{1}{2}}\le$$

$$\le C(2^{-nm}n^{1-\frac{1}{\theta}})^{\frac{1}{2}}\left(\sum_{n\le s<\alpha n}2^{sr\theta}\|\delta_s(f)\|_{\bar{p}}^{\theta}\right)^{\frac{1}{2\theta}}\left(\sum_{n\le s<\alpha n}1\right)^{\frac{1}{2}(1-\frac{1}{\theta})}$$

$$\le C2^{-\frac{nm}{2}}n^{1-\frac{1}{\theta}}\asymp M^{-\frac{1}{2}}(\log(1+M))^{1-\frac{1}{\theta}}.$$

Thus,

$$J_1(n)\le CM^{-\frac{1}{2}}(\log(1+M))^{1-\frac{1}{\theta}} \tag{14}$$

in the case of $r=\sum_{j=1}^{m}\frac{1}{p_j}$.

To estimate $J_2(n)$, we apply Holder's inequality, and taking into account $r = \sum\limits_{j=1}^{m} \frac{1}{p_j}$ and $\alpha = m\left(2\sum\limits_{j=1}^{m} \frac{1}{q_j}\right)^{-1}$, we obtain

$$J_2(n) \le C \sum_{n\alpha \le s < +\infty} 2^{s\sum\limits_{j=1}^{m}(\frac{1}{p_j}-\frac{1}{q_j})} \|\delta_s(f)\|_{\bar{p}} \le \tag{15}$$

$$\le C\left(\sum_{s=0}^{\infty} 2^{sr\theta} \|\delta_s(f)\|_{\bar{p}}^{\theta}\right)^{\frac{1}{\theta}} \left(\sum_{n\alpha \le s < +\infty} 2^{-s\theta' \sum\limits_{j=1}^{m} \frac{1}{q_j}}\right)^{\frac{1}{\theta'}} \le C2^{-n\alpha\sum\limits_{j=1}^{m}\frac{1}{q_j}} = C2^{-\frac{nm}{2}} \asymp M^{-\frac{1}{2}}.$$

By (14) and (15), the inequality (2) implies that

$$\|f - P(\Omega_M)\|_{\bar{q}} \le CM^{-\frac{1}{2}} (\log(1+M))^{1-\frac{1}{\theta}}$$

in the case of $r = \sum\limits_{j=1}^{m} \frac{1}{p_j}$. It proves the upper bound estimation in the second item.

Let $r > \sum\limits_{j=1}^{m} \frac{1}{p_j}$. Suppose

$$N_s = \left[2^{n\left(r - \sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-1\right)\right)} 2^{-s\left(r-\sum\limits_{j=1}^{m}\frac{1}{p_j}\right)}\right] + 1.$$

Then

$$\sum_{s=0}^{n-1} \sharp\{\bar{k} = (k_1, ..., k_m) : [2^{s-1}] \le \max_{j=1,\ldots,m} |k_j| < 2^s\} + \sum_{n \le s < \alpha n} N_s \le$$
$$\le C2^{nm} + (\alpha - 1)n \le C2^{nm} \le CM.$$

If $f \in H^r_{\bar{p}}$, then, by using the definition of the numbers N_s and $r > \sum\limits_{j=1}^{m} \frac{1}{p_j}$, we obtain from (13) the following

$$J_1(n) \le \left(\sum_{n \le s < \alpha n} N_s^{-1} 2^{sm} 2^{2s\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{2}\right)} \|\delta_s(f)\|_{\bar{p}}^2\right)^{\frac{1}{2}} \le$$

$$\le C2^{-\frac{n}{2}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-1\right)\right)} \left(\sum_{n \le s < \alpha n} 2^{-s\left(r-\sum\limits_{j=1}^{m}\frac{1}{p_j}\right)}\right)^{\frac{1}{2}} \le C2^{-n\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{2}-\frac{1}{p_j}\right)\right)}.$$

Thus,

$$J_1(n) \le CM^{-\frac{1}{m}\left(r+\sum\limits_{j=1}^{m}(\frac{1}{2}-\frac{1}{p_j})\right)} \tag{16}$$

in the case of $r > \sum\limits_{j=1}^{m} \frac{1}{p_j}$.

To estimate $J_2(n)$, we suppose $\alpha = \left(r + \sum\limits_{j=1}^{m}\left(\frac{1}{2} - \frac{1}{p_j}\right)\right)\left(r + \sum\limits_{j=1}^{m}\left(\frac{1}{q_j} - \frac{1}{p_j}\right)\right)^{-1}$ and get

$$J_2(n) \le C \sum_{n\alpha \le s < \infty} 2^{-s\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)} \le C2^{-n\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{2}-\frac{1}{p_j}\right)\right)} \le$$
$$\le CM^{-\frac{1}{m}\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{2}-\frac{1}{p_j}\right)\right)} \tag{17}$$

for a function $f \in H^r_{\bar{p}}$. By (16) and (17), it follows from (2) that

$$\|f - P(\Omega_M)\|_{\bar{q}} \le CM^{-\frac{1}{m}\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{2}-\frac{1}{p_j}\right)\right)}$$

for any function $f \in H^r_{\bar{p}}$ in the case of $r > \sum\limits_{j=1}^{m} \frac{1}{p_j}$.

It follows from $B^r_{\bar{p},\theta} \subset H^r_{\bar{p}}$ that

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \le e_M\left(H^r_{\bar{p}}\right)_{\bar{q}} \le CM^{-\frac{1}{m}\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{2}-\frac{1}{p_j}\right)\right)}$$

in the case of $r > \sum\limits_{j=1}^{m} \frac{1}{p_j}$. It proves the upper bound estimation in the item 3.

Let us consider the lower bound estimation in the case of $r = \sum\limits_{j=1}^{m} \frac{1}{p_j}$. Consider the function

$$g_1(\bar{x}) = \sum_{s=1}^{n} \sum_{\bar{k}\in\rho(s)} \prod_{j=1}^{m} k_j^{-1} \cos k_j x_j. \tag{18}$$

Then

$$\delta_s(g_1, \bar{x}) = \sum_{\bar{k}\in\rho(s)} \prod_{j=1}^{m} k_j^{-1} \cos k_j x_j.$$

It is known that for a function $d_s(\bar{x}) = \sum\limits_{\bar{k}\in\rho(s)} \prod\limits_{j=1}^{m} \cos k_j x_j$ the following relation holds

$$\|d_s\|_{\bar{p}} \asymp 2^{s\sum\limits_{j=1}^{m}(1-\frac{1}{p_j})}, \;\; 1 < p_j < +\infty, \;\; j = 1, ..., m.$$

Therefore, by the Marcinkiewicz theorem on multipliers (see [18]), we have

$$\|\delta_s(g_1)\|_{\bar{p}} \le C2^{-sm}\|d_s\|_{\bar{p}} \le C2^{-s\sum\limits_{j=1}^{m}\frac{1}{p_j}}.$$

Hence, since $r = \sum\limits_{j=1}^{m} \frac{1}{p_j}$, we obtain

$$\left(\sum_{s=0}^{\infty} 2^{sr\theta} \|\delta_s(g_1)\|^{\theta}_{\bar{p}}\right)^{\frac{1}{\theta}} \le C_1 n^{\frac{1}{\theta}}.$$

Therefore, the function $f_2(\bar{x}) = C_1^{-1} n^{-\frac{1}{\theta}} g_1(\bar{x})$ belongs to the class $B^r_{\bar{p},\theta}$, $1 < p_j < +\infty$, $j = 1, ..., m$.

Now, we are going to construct a function P_1, which satisfies the conditions of the formula (7). Let

$$v_1(\bar{x}) = \sum_{s=1}^{n} \sum_{\bar{k}\in\rho(s)} \prod_{j=1}^{m} \cos k_j x_j$$

and Ω_M be an arbitrary set of M vectors $\bar{k} = (k_1, ..., k_m)$ with integer coordinates. Consider the function

$$u_1(\bar{x}) = \sum_{\bar{k}\in\Omega_M}^{*} \prod_{j=1}^{m} \cos k_j x_j$$

which contains only those terms in (18) with indices in Ω_M. Suppose $w_1(\bar{x}) = v_1(\bar{x}) - u_1(\bar{x})$. Then, since $1 < {q_j}' < 2$, $j = 1, ..., m$, we obtain, by Perseval's equality, the following

$$\|w_1\|_{\bar{q}'} \le \|v_1\|_{\bar{q}'} + \|u_1\|_2 \le \|v_1\|_{\bar{q}'} + CM^{\frac{1}{2}}.$$

By the property of the norm and the estimation of the norm of the Dirichlet kernel in the Lebesgue space, we have

$$\|v_1\|_{\bar{q}'} \le \sum_{s=1}^{n} \|\delta_s(v_1)\|_{\bar{q}'} \le$$

$$\le C \sum_{s=1}^{n} 2^{s \sum_{j=1}^{m}\left(1-\frac{1}{q_j{}'}\right)} \le C 2^{n \sum_{j=1}^{m} \frac{1}{q_j}}.$$

Therefore, taking into account $\frac{1}{q_j} < \frac{1}{2}$, $j = 1, ..., m$, we get

$$\|w_1\|_{\bar{q}'} \le C(2^{\frac{nm}{2}} + M^{\frac{1}{2}}) \le C_2 2^{\frac{nm}{2}}.$$

Hence, the function

$$P_1(\bar{x}) = C_2^{-1} 2^{-\frac{nm}{2}} w_1(\bar{x})$$

satisfies the conditions of the formula (7). Then, by substituting the functions f_2 and P_1 into (7) and by orthogonality of the trigonometric system, we obtain

$$e_M\left(f_2\right)_{\bar{q}} \ge C \sum_{n_1 \le s < n} \sum_{\bar{k} \in \rho(s)} \prod_{j=1}^{m} k_j^{-1} 2^{-\frac{nm}{2}} n^{-\frac{1}{\theta}} \ge$$

$$\ge C(\ln 2)^m \sum_{n_1 \le s < n} 2^{-\frac{nm}{2}} n^{-\frac{1}{\theta}} = C(\ln 2)^m 2^{-\frac{nm}{2}} n^{-\frac{1}{\theta}} (n - n_1) \ge$$

$$\ge C(\ln 2)^m 2^{-\frac{nm}{2}} n^{1-\frac{1}{\theta}} \asymp M^{-\frac{1}{2}} (\log(1+M))^{1-\frac{1}{\theta}},$$

where n_1 is a natural number such that $n_1 \le \frac{n}{2}$.

So, for the function $f_2 \in B^r_{\bar{p},\theta}$, it has been proved that

$$e_M\left(f_2\right)_{\bar{q}} \ge C M^{-\frac{1}{2}} (\log(1+M))^{1-\frac{1}{\theta}}$$

in the case of $r = \sum_{j=1}^{m} \frac{1}{p_j}$. Hence

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \ge C M^{-\frac{1}{2}} (\log(1+M))^{1-\frac{1}{\theta}}$$

in the case of $r = \sum_{j=1}^{m} \frac{1}{p_j}$. It proves the lower bound estimation in the second item.

Let us prove the lower bound estimation for the case $r > \sum_{j=1}^{m} \frac{1}{p_j}$. Since in this case an upper bound estimation of the quantity $e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}}$ does not depend on θ and $B^r_{\bar{p},1} \subset B^r_{\bar{p},\theta}$, $1 < \theta \le +\infty$, it suffices to prove the lower bound estimation for $B^r_{\bar{p},1}$.

For a number $M \in \mathbb{N}$, we choose a natural number n such that $2^{nm} < M \le 2^{(n+1)m}$ and $2M \le \sharp\rho(n)$, where $\sharp\rho(n)$ denotes the number of elements in the set $\rho(n)$.

Consider the following function

$$f_3(\bar{x}) = 2^{-n\left(r + \sum_{j=1}^{m}\left(1-\frac{1}{p_j}\right)\right)} \sum_{\bar{k} \in \rho(n)} e^{i\langle \bar{k}, \bar{x}\rangle}.$$

Then $\|\delta_s(f_3)\|_{\bar{p}} = 0$ provided $s \ne n$ and

$$\|\delta_n(f_3)\|_{\bar{p}} = 2^{-n\left(r + \sum_{j=1}^{m}\left(1-\frac{1}{p_j}\right)\right)} \prod_{j=1}^{m} \Big\| \sum_{k_j=2^{n-1}}^{2^n-1} e^{ik_jx_j} \Big\|_{p_j}.$$

By the estimation of the norm of the Dirichlet kernel (see [18], p. 181), we have

$$\Big\|\sum_{k_j=2^{n-1}}^{2^n-1} e^{ik_jx_j}\Big\|_{p_j} \le C2^{n(1-\frac{1}{p_j})},$$

for $p_j \in (1,\infty)$, $j=1,...,m$. Therefore

$$\|\delta_n(f_3)\|_{\bar{p}} \le C2^{-nr}.$$

Hence

$$\sum_{s=0}^{\infty} 2^{sr}\,\|\delta_s(f_3)\|_{\bar{p}} \le C_3,$$

i.e. the function $C_3^{-1}f_3 \in B^r_{\bar{p},1}$. Next, we consider the functions

$$v_2(\bar{x}) = \sum_{\bar{k}\in\rho(n)} e^{i\langle\bar{k},\bar{x}\rangle}$$

and

$$u_2(\bar{x}) = \sum_{\bar{k}\in\rho(n)\cap\Omega_M} e^{i\langle\bar{k},\bar{x}\rangle}.$$

Suppose $w_2(\bar{x}) = v_2(\bar{x}) - u_2(\bar{x})$. By Perseval's equality,

$$\|u_2\|_2 \le M^{\frac{1}{2}}, \quad \|v_2\|_2 \le C2^{\frac{nm}{2}}.$$

From these relations, we obtain, by the properties of the norm, the following

$$\|w_2\|_2 \le \|v_2\|_2 + \|u_2\|_2 \le C_4 2^{\frac{nm}{2}}.$$

Therefore, the function $P_2(\bar{x}) = C_4^{-1}2^{-\frac{nm}{2}}w_2(\bar{x})$ satisfies the conditions of the formula (7). Since $2 < q_j < \infty$, $j=1,...,m$, we have $e_M\left(f_3\right)_2 \le Ce_M\left(f_3\right)_{\bar{q}}$. Now, by the formula (7), we get

$$e_M\left(f_3\right)_{\bar{q}} \ge Ce_M\left(f_3\right)_2 \ge$$

$$\ge C\inf_{\Omega_M}\int_{\mathbb{T}^m} f_3(\bar{x})\bar{P}_2(\bar{x})d\bar{x} =$$

$$= C_2^{-1}2^{-\frac{nm}{2}}2^{-n\left(r+\sum\limits_{j=1}^{m}\left(1-\frac{1}{p_j}\right)\right)}\inf_{\Omega_M}\left[\sharp\rho(n) - \sharp\left(\rho(n)\cap\Omega_M\right)\right] \ge$$

$$\ge C2^{-\frac{nm}{2}}2^{-n\left(r+\sum\limits_{j=1}^{m}\left(1-\frac{1}{p_j}\right)\right)}\left[\sharp\rho(n) - M\right] \ge$$

$$\ge C2^{-\frac{nm}{2}}2^{-n\left(r+\sum\limits_{j=1}^{m}\left(1-\frac{1}{p_j}\right)\right)}\left[\sharp\rho(n) - \frac{\sharp\rho(n)}{2}\right] \ge$$

$$\ge C2^{-\frac{nm}{2}}2^{-n\left(r-\sum\limits_{j=1}^{m}\frac{1}{p_j}\right)}.$$

It follows from the relation $2^{nm} \asymp M$ that

$$e_M\left(f_3\right)_{\bar{q}} \ge CM^{-\frac{1}{m}\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{2}-\frac{1}{p_j}\right)\right)}$$

in the case of $r > \sum\limits_{j=1}^{m}\frac{1}{p_j}$ for the function $C_3^{-1}f_3 \in B^r_{\bar{p},1}$. Hence

$$e_M\left(B^r_{\bar{p},1}\right)_{\bar{q}} \ge CM^{-\frac{1}{m}\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{2}-\frac{1}{p_j}\right)\right)}.$$

Therefore,

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \geq CM^{-\frac{1}{m}\left(r+\sum\limits_{j=1}^{m}\left(\frac{1}{2}-\frac{1}{p_j}\right)\right)}$$

in the case of $r > \sum\limits_{j=1}^{m}\frac{1}{p_j}$. So Theorem 2.1 has been proved.

2.2. Theorem. *Let* $\bar{p}=(p_1,...,p_m)$, $\bar{q}=(q_1,...,q_m)$, $1 < p_j < q_j \leq 2$, *and* $1 \leq \theta \leq +\infty$. *If* $r > \sum\limits_{j=1}^{m}(\frac{1}{p_j}-\frac{1}{q_j})$, *then*

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \asymp M^{-\frac{1}{m}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)}.$$

Proof. For a number $M \in \mathbb{N}$, we choose a natural number n such that $M \asymp 2^{nm}$. By the inequality of distinct metrics (see Theorem 1.2) and by Holder's inequality, we have

$$\|f-\sum_{s=0}^{n}\delta_s(f)\|_{\bar{q}} \leq \sum_{s=n}^{\infty}\|\delta_s(f)\|_{\bar{q}} \leq$$

$$\leq \left[\sum_{s=0}^{\infty}2^{sr\theta}\|\delta_s(f)\|^{\theta}_{\bar{q}}\right]^{\frac{1}{\theta}}\left[\sum_{s=n}^{\infty}2^{s\theta'\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)}\right]^{\frac{1}{\theta'}} \leq$$

$$\leq C2^{n\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)} \leq CM^{-\frac{1}{m}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)}$$

for $f \in B^r_{\bar{p},\theta}$, $\frac{1}{\theta}+\frac{1}{\theta'}=1$. Therefore

$$e_M(f)_{\bar{q}} \leq \|f-\sum_{s=0}^{n}\delta_s(f)\|_{\bar{q}} \leq CM^{-\frac{1}{m}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)}.$$

Hence

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \leq CM^{-\frac{1}{m}\left(r-\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-\frac{1}{q_j}\right)\right)}.$$

It proves the upper bound estimation.

For the lower bound estimation, let us consider the function

$$f_4(\bar{x}) = n^{-r+\sum\limits_{j=1}^{m}\left(\frac{1}{p_j}-1\right)}V_n(\bar{x}),$$

where $V_n(\bar{x})$ is a multiple of the Valle-Poisson sum.

Next, following the proof in [9] (pp. 46-47) and applying Theorem 1.2, we obtain the lower bound estimation of the quantity $e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}}$.

2.3. Theorem. *Let* $\bar{p}=(p_1,...,p_m)$, $\bar{q}=(q_1,...,q_m)$, $2 \leq p_j < q_j < \infty$, $j=1,...,m$, *and* $1 \leq \theta \leq +\infty$. *If* $r > \frac{m}{2}$, *then*

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \asymp M^{-\frac{r}{m}}.$$

Proof. By the inclusion $B^r_{\bar{p},\theta} \subset B^r_{2,\theta} \subset H^r_2$, we have

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \leq e_M\left(B^r_{2,\theta}\right)_{\bar{q}} \leq e_M\left(H^r_2\right)_{\bar{q}}.$$

By Theorem 2.1,

$$e_M\left(H^r_2\right)_{\bar{q}} \leq CM^{-\frac{r}{m}},$$

for $p_j = 2$, $j=1,...,m$. Hence

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \leq CM^{-\frac{r}{m}}.$$

It proves the upper bound estimation.

Let us consider the lower bound estimation. Consider Rudin-Shapiro's polynomial (see [15], p. 155) of the type

$$R_s(x) = \sum_{s=2^{s-1}}^{2^s} \varepsilon_k e^{ikx}, \quad x \in [0, 2\pi], \quad \varepsilon_k = \pm 1.$$

It is known that $\|R_s\|_\infty = \max_{x\in[0,2\pi]} |R_s(x)| \le C 2^{\frac{s}{2}}$ (see [15], p. 155). For a given number M choose a number n such that $M \asymp 2^{nm}$. Now we consider the function

$$f_5(\bar{x}) = 2^{-n(\frac{m}{2}+r)} \sum_{s=1}^{n} \prod_{1}^{m} R_s(x_j).$$

Then, by the continuity, we have $f_5 \in L_{\bar{p}}(\mathbb{T}^m)$ and

$$\sum_{s=0}^{\infty} 2^{s\theta r} \|\delta_s(f_5)\|_{\bar{p}}^{\theta} = 2^{-n(\frac{m}{2}+r)} \sum_{s=1}^{n} 2^{s\theta r} \| \prod_{1}^{m} R_s(x_j)\|_{\bar{p}}^{\theta} \le$$

$$\le 2^{-n(\frac{m}{2}+r)} \sum_{s=1}^{n} 2^{s(\frac{m}{2}+r)\theta} \le C_5.$$

Hence, the function $C_5^{-1} f_5 \in B^r_{\bar{p},\theta}$. Now, we construct a function $P(\bar{x})$, which satisfies the conditions in the formula (7). Suppose

$$v_3(\bar{x}) = \sum_{s=1}^{n} \prod_{1}^{m} R_s(x_j), \quad u_3(\bar{x}) = \sum_{s}^{*} \prod_{1}^{m} R_s(x_j),$$

where the sign $*$ means that the polynomial $u_3(\bar{x})$ contains only those harmonics of v_3, which have indices in Ω_M. Suppose $w_3(\bar{x}) = v_3(\bar{x}) - u_3(\bar{x})$. Then, since $1 < {q_j}' = \frac{q_j}{q_j - 1} < 2$, $j = 1, ..., m$, we have the following (by Perseval's equality)

$$\|w_3\|_{\bar{q}'} \le \|w_3\|_2 \le C_1 2^{\frac{nm}{2}}.$$

Therefore, for the function $P_3(\bar{x}) = C_1^{-1} 2^{-\frac{nm}{2}} w_3(\bar{x})$ the inequality $\|P_3\|_{\bar{q}'} \le 1$ holds. Now, using the formula (7), we obtain

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \ge e_M(f_3)_{\bar{q}} \ge 2^{-n(\frac{m}{2}+r)} 2^{-\frac{nm}{2}} (2^{nm} - M) \ge C 2^{-n(m+r)} 2^{nm} \ge C M^{-\frac{r}{m}}.$$

So

$$e_M\left(B^r_{\bar{p},\theta}\right)_{\bar{q}} \ge C M^{-\frac{r}{m}}.$$

It proves Theorem 2.3.

Remark. In the case $p_j = p$, $q_j = q$, $j = 1, ..., m$, and $r > m(\frac{1}{p} - \frac{1}{q})$, the results of R.A. DeVore and V.N. Temlyakov [9] follow from Theorem 2.1 - 2.3. If $1 < p \le 2 < q < \infty$ and $m(\frac{1}{p} - \frac{1}{q}) < r \le \frac{m}{p}$, the results of S.A. Stasyuk [20, 21] follow from the first and second items of Theorem 2.1. Theorem 2.1 - 2.3 were announced in [3].

Acknowledgements. This work was supported by the Ministry of Education and Science of Republic Kazakhstan (Grant no. 5129GF4) and by the Competitiveness Enhancement Program of the Ural Federal University (Enactment of the Government of the Russian Federation of March 16, 2013 no. 211, agreement no. 02.A03. 21.0006 of August 27, 2013).

References

[1] Akishev, G. *On the exact estimations of the best M-term approximation of Besov's class*, Siberian Electronic Mathematical Reports **7**, 255 – 274, 2010.

[2] Akishev, G. *On the order of the M-term approximation classes in Lorentz spaces*, Matematical Journal. Almaty **11**, (1), 5 - 29, 2011.

[3] Akishev, G. *On the M-term approximation Besov's classes*, International Conference "Theory of approximation of functions and its applications" dedicated to the 70-th anniversary of corresponding member of National Academy of Ukraine, professor A.I. Stepanets (1942 – 2007) May 28 - June 3, Ukraine, Kamianets-Podilsky, p. 12 , 2012.

[4] Amanov, T.I. *Spaces of differentiable functions with dominant mixed derivative*, Alma-Ata, 1976.

[5] Bazarkhanov, D.B. *Estimates for certain approximation characteristics of Nikol'skii-Besov spaces with generalized mixed smoothness*, Doklady Ros. Akade. Nauk **426**, (1), 11 – 14, 2009.

[6] Belinsky, E.S. *Approximation by a 'floating' system of exponents on the classes of smooth periodic functions*, Mat. sb. **132**, (1), 20 – 27, 1987.

[7] Besov, O.V. *Investigation of a family of functional spaces in connection with the theorems of imbedding and extension*, Trudy Mat. Inst. Akad. Nauk SSSR **60**, 42 - 61, 1961.

[8] De Vore, R.A. *Nonlinear approximation*, Acta Numerica **7**, 51 – 150, 1998.

[9] De Vore, R.A. and Temlyakov, V.N. *Nonlinear approximation by trigonometric sums*, Jour. of Fourier Analysis and Applications **2**, (1), 29–48, 1995.

[10] Dinh Dung *On asymptotic order of n-term approximations and non-linear - n widths*, Vietnam Journal Math. **27**, (4), 363 - 367, 1999.

[11] Duan, L. *The best m-term approximations on generalized Besov's classes $MB^{\Omega}_{q,\theta}$ with regard to orthogonal dictionaries*, Journal Approximation Theory **162**, 1964 – 1981, 2010.

[12] Hansen, M. and Sickel, W. *Best m-term approximation and Lizorkin-Triebel spaces*, Journal Approximation Theory **163**, 923 – 954, 2011.

[13] Ismagilov, R.S. *Widths of sets in linear normed spaces and the approximation of functions by trigonometric polynomials*, Uspehi mathem. nauk **29**, (3), 161 – 178, 1974.

[14] Kashin, B.S. *Approximation properties of complete orthnormal systems*, Trudy Mat. Inst. Steklov **172**, 187 – 201, 1985.

[15] Kashin, B.S., and Sahakyan, A.A. *Orthogonal series*, Nauka, Moscow, 1984.

[16] Lizorkin, *P.I. Generalized Holder spaces $B^{(r)}_{p,\theta}$ and their correlation with the Sobolev spaces $L^{(r)}_p$*, Sibirsk. Mat. Zh. **9**, 1127 - 1152, 1968.

[17] Maiorov, V.E. *On linear widths of Sobolev classes and chains of extremal subspaces*, Mat. sb. **113**, (3), 437 – 463, 1980.

[18] Nikol'skii, S. M. *Approximation of classes of functions of several variables and embeding theorems*, Nauka, Moscow, 1977.

[19] Romanyuk, A.S. *On the best M-term trigonometric approximations for the Besov classes of periodic functions of many variables*, Izv. Ros. Akad. Nauk, Ser. Mat. **67**, (2), 61 – 100, 2003.

[20] Stasyuk, S.A. *Best m-term trigonometric approximation for the classes $B^r_{p,\theta}$ of functions of low smoothness*, Ukrain. Mathem. Journal **62**, (1), 114 –122, 2010.

[21] Stasyuk, S.A. *Best m-term trigonometric approximation of periodic functions of several variables from Nikol'skii - Besov classes for small smoothness*, Journal of Approximation Theory **177** , 1 - 16, 2014.

[22] Stechkin, S.B. *On the absolute convergence of orthogonal series*, Doklad. Akadem. Nauk SSSR **102**, (2) , 37 – 40, 1955.

[23] Temlyakov, V. N. *Nonlinear methods approximation*, Foundations of Computational Mathematics **3**, 33–107, 2003.

[24] Uninskii, A.P. *Inequalities for entire functions and trigonometric polynomials with mixed norm*, Proceeding of the All-Union Symposium on embedding theorems, (Baku - 1966) - M.: Nauka, 112 - 118, 1970.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 311 – 335

A generalized operational method for solving integro–partial differential equations based on Jacobi polynomials

Abdollah Borhanifar*† and Khadijeh Sadri‡

Abstract

In this paper, a numerical method is developed for solving linear and nonlinear integro-partial differential equations in terms of the two variables Jacobi polynomials. First, some properties of these polynomials and several theorems are presented then a generalized approach implementing a collocation method in combination with two dimensional operational matrices of Jacobi polynomials is introduced to approximate the solution of some integro–partial differential equations with initial or boundary conditions. Also, it is shown that the resulted approximate solution is the best approximation for the considered problem. The main advantage is to derive the Jacobi operational matrices of integration and product to achieve the best approximation of the two dimensional integro–differential equations. Numerical results are given to confirm the reliability of the proposed method for solving these equations.

Keywords: Best approximation, Collocation method, Integro–partial differential equations, Operational matrix, Shifted Jacobi polynomials.

2000 AMS Classification: 47B36, 14R15, 39B05, 65D15, 15B99, 65R20, 45K005, 47G10, 34K28.

Received : 18.11.2014 *Accepted* : 17.02.2015 *Doi* : 10.15672/HJMS.20164513116

*Department of Mathematics, Faculty of Mathematical Sciences, University of Mohaghegh Ardabili, Ardabil,Iran,
Email: borhani@uma.ac.ir
†Corresponding Author.
‡Department of Mathematics, Faculty of Mathematical Sciences, University of Mohaghegh Ardabili, Ardabil,Iran,
Email: kh.sadri@uma.ac.ir

1. Introduction

Finding the analytical solutions of functional equations has been devoted attention of mathematicians's interest in recent years. Several methods are proposed to achieve this purpose, such as [7, 8, 9, 10, 11, 13]. Many problems in theoretical physics and other sciences lead to integro–partial differential equations. In order to solve these equations, several numerical methods have been proposed such as [21, 22, 25, 26, 28, 29]. The solutions of this kind of equations are often quite complicated. For this reason in many cases, it is required to obtain the approximate solutions. For example the Jacobi collocation method has been applied to solve various differential equations, [3, 14, 15, 17, 18, 19, 20, 23]. Also, Bhrawy and et al have presented a new Legendre spectral collocation method for fractional Burgers equations, [6]. In [16], authors have used Jacobi–Gauss–Lobatto collocation method for the numerical solution of $1 + 1$ nonlinear Schrödinger equation. Also, Bhrawy in [1] has presented an pseudospectral approximation based on Jacobi polynomials for generalized Zakharov system. Two spectral tau algorithms based on Jacobi polynomials have been applied to solve multi–term time–space fractional partial differential equations and time fractional diffusion–Wave equations, [4, 5]. Bhrawy and et al have been presented two different collocation scheme for both temporal and spatial discriminations of mobile–immobile advection–dispersion model (TVFO–MIAD model), [2]. Borhanifar and Sadri have utilized a Jacobi operational collocation method for systems of two dimensional integral equations, [12].

In this paper, the Jacobi polynomials are used as a basis function for solving linear and nonlinear integro–partial differential equations, the numerical solution, $u(x, y)$, is approximated in terms of two variables Jacobi polynomials as $x,\ y\ \in [0, 1]$. In order to realize this aim, the shifted Jacobi polynomials together the collocation technique are used. The Jacobi operational matrices of the integration and product are constructed on the interval $[0, 1]$. The main aim is to improve Jacobi operational matrices to the spectral solution of partial integro-differential equations. For solving the resulted algebraic system, the $(N+\ 1)$ roots of one variable Jacobi polynomials $P^{(\alpha,\beta)}_{N+\ 1}(x)$ and $P^{(\alpha,\beta)}_{N+\ 1}(y)$ are considered in the $x, y-$directions. The domain of two dimensional is represented by a tensor product points $\{x_i\}^N_{i=\ 0}$ and $\{y_j\}^N_{j=\ 0}$ which are roots of $P^{(\alpha,\beta)}_{N+\ 1}(x)$ and $P^{(\alpha,\beta)}_{N+\ 1}(y)$. Each the equations of the algebraic system is collocated in the resulted tensor points $\{(x_i, y_j)\}^N_{i,j=\ 0}$ and is given linear or nonlinear systems of algebraic equations which can be solved using the well–known Newtons iterative method. Thus, the Jacobi coefficients are obtained and the approximate solution is determined.

The remainder of this paper is organized as follows: The Jacobi polynomials, some of their properties and one dimensional operational matrix of integration are introduced in Section 2. Afterwards, some properties of two variables Jacobi polynomials are stated and the operational matrices of integration and product will be extended to two dimensional case in Section 3. In Section 4, the existence and uniqueness of the best approximation are studied and an error estimator is introduced. Section 5 is devoted to applying two dimensional Jacobi operational matrices for solving the partial integro–differential equations. For this purpose, four examples are presented. A conclusion is presented in Section 6.

2. Jacobi polynomials and their operational matrix of integration

The well–known Jacobi polynomials are defined on the interval $z \in [-1, 1]$, constitute an orthogonal system with respect to the weight function $w^{(\alpha,\beta)}(z) = (1 - z)^{\alpha}\ (1 + z)^{\beta}$, and can be determined with the following recurrence formula:

$$P_{i+1}^{(\alpha,\beta)}(z) = A(\alpha,\beta,i)\ P_i^{(\alpha,\beta)}(z) + z\ B(\alpha,\beta,i)\ P_i^{(\alpha,\beta)}(z) - D(\alpha,\beta,i)\ P_{i-1}^{(\alpha,\beta)}(z), \quad i = 1, 2, ..., \tag{2.1}$$

where

$$A(\alpha,\beta,i) = \frac{(2\ i+\alpha+\beta+1)(\alpha^2-\beta^2)}{2(i+1)(i+\alpha+\beta+1)(2\ i+\alpha+\beta)},$$

$$B(\alpha,\beta,i) = \frac{(2\ i+\alpha+\beta+2)(2\ i+\alpha+\beta+1)}{2(i+1)(i+\alpha+\beta+1)},$$

$$D(\alpha,\beta,i) = \frac{(i+\alpha)(i+\beta)(2\ i+\alpha+\beta+2)}{(i+1)(i+\alpha+\beta+1)(2\ i+\alpha+\beta)},$$

and

$$P_0^{(\alpha,\beta)}(z) = 1, \qquad P_1^{(\alpha,\beta)}(z) = \frac{\alpha+\beta+2}{2}z + \frac{\alpha-\beta}{2}.$$

The orthogonality condition of Jacobi polynomials is

$$\int_{-1}^{1} P_j^{(\alpha,\beta)}(z)\ P_k^{(\alpha,\beta)}(z)\ w^{(\alpha,\beta)}(z)\ dz = h_k\ \delta_{jk},$$

where

$$h_k = \frac{2^{\alpha+\beta+1}\Gamma(k+\alpha+1)\Gamma(k+\beta+1)}{(2k+\alpha+\beta+1)k!\ \Gamma(k+\alpha+\beta+1)}.$$

The analytic form of Jacobi polynomials is given by, [27],

$$P_i^{(\alpha,\beta)}(z) = \sum_{k=0}^{i} \frac{(-1)^{(i-k)}\Gamma(i+\beta+1)\Gamma(i+k+\alpha+\beta+1)}{\Gamma(k+\beta+1)\Gamma(i+\alpha+\beta+1)\ (i-k)!\ k!}\ (\frac{1+z}{2})^k, \quad i = 0, 1, \ldots .$$

For practical use of Jacobi polynomials on the interval $x \in [0,1]$, it is necessary to shift the defining domain by means of the following change variable:

$$z = 2\ x - 1, \qquad x \in [0,1]$$

The shifted Jacobi polynomials are generated from the three-term recurrence relation

$$\begin{aligned} P_{i+1}^{(\alpha,\beta)}(x) &= A(\alpha,\beta,i)\ P_i^{(\alpha,\beta)}(x) + (2\ x-1)\ B(\alpha,\beta,i)\ P_i^{(\alpha,\beta)}(x) \\ &\quad - D(\alpha,\beta,i)\ P_{i-1}^{(\alpha,\beta)}(x), \qquad i = 1, 2, ..., \\ P_0^{(\alpha,\beta)}(x) &= 1, \qquad P_1^{(\alpha,\beta)}(x) = \frac{(\alpha+\beta+2)(2\ x-1)}{2} + \frac{\alpha-\beta}{2}. \end{aligned} \tag{2.2}$$

The orthogonality condition and weight function are respectively,

$$\int_0^1 P_i^{(\alpha,\beta)}(x)\ P_j^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx = \theta_i\ \delta_{ij},$$

where

$$\theta_i = \frac{\Gamma(i+\alpha+1)\Gamma(i+\beta+1)}{(2i+\alpha+\beta+1)\ i!\ \Gamma(i+\alpha+\beta+1)},$$

and

$$w^{(\alpha,\beta)}(x) = (1-x)^\alpha\ x^\beta, \qquad x \in [0,1].$$

Also, the analytic form of shifted Jacobi polynomials will be as follows, [27],

$$P_i^{(\alpha,\beta)}(x) = \sum_{k=0}^{i} \frac{(-1)^{(i-k)}\Gamma(i+\beta+1)\Gamma(i+k+\alpha+\beta+1)\ x^k}{\Gamma(k+\beta+1)\ \Gamma(i+\alpha+\beta+1)\ (i-k)!\ k!}, \tag{2.3}$$
$$i = 0, 1, \dots .$$

A continuous function $u(x)$, square integrable in $[0,1]$, can be expressed in terms of shifted Jacobi polynomials as

$$u(x) = \sum_{j=0}^{\infty} C_j\ P_j^{(\alpha,\beta)}(x), \tag{2.4}$$

where the coefficients C_j are given by

$$C_j = \frac{1}{\theta_j} \int_0^1 u(x)\ P_j^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx, \qquad j = 0, 1, \dots .$$

In practice, only the first $(N+1)-$terms shifted Jacobi polynomials are considered. Therefore, one has

$$u_N(x) = \sum_{j=0}^{N} C_j\ P_j^{(\alpha,\beta)}(x) = \Phi^T(x)\ C = C^T\ \Phi(x), \tag{2.5}$$

where the vectors C and $\Phi(x)$ are given by

$$C = [C_0, C_1, ..., C_N]^T, \qquad \Phi(x) = [P_0^{(\alpha,\beta)}(x), P_1^{(\alpha,\beta)}(x), ..., P_N^{(\alpha,\beta)}(x)]^T. \tag{2.6}$$

Some other properties of the shifted Jacobi polynomials are presented as follows.

(1) $P_i^{\ (\alpha,\beta)}(0) = (-1)^i \begin{pmatrix} i+\alpha \\ i \end{pmatrix}$,

(2) $\dfrac{d^i}{dx^i}\ P_n^{(\alpha,\beta)}(x) = \dfrac{\Gamma(n+\alpha+\beta+i+1)}{\Gamma(n+\alpha+\beta+1)}\ P_{n-i}^{(\alpha+i,\beta+i)}(x)$.

2.1. Lemma. *The shifted Jacobi polynomial $P_i^{(\alpha,\beta)}(x)$ can be obtained in the form of:*

$$P_i^{(\alpha,\beta)}(x) = \sum_{k=0}^{i} \gamma_k^{(i)}\ x^k,$$

where $\gamma_k^{(i)}$ are

$$\gamma_k^{(i)} = (-1)^{i-k} \begin{pmatrix} i+k+\alpha+\beta \\ k \end{pmatrix} \begin{pmatrix} i+\alpha \\ i-k \end{pmatrix}.$$

Proof. $\gamma_k^{(i)}$ can be obtained as,

$$\gamma_k^{(i)} = \frac{1}{k!} \frac{d^k}{dx^k}\ P_i^{(\alpha,\beta)}(x)\ |_{x=0}\ .$$

Now, using properties (1) and (2), the lemma can be proved. □

2.2. Lemma. *If $p > \beta - 1$, then*

$$\int_0^1 x^{p-\beta}\ P_n^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx =$$
$$\sum_{l=0}^{n} \frac{(-1)^{n-l}\Gamma(n+\beta+1)\ \Gamma(n+l+\alpha+\beta+1)\ \Gamma(p+l+1)\ \Gamma(\alpha+1)}{\Gamma(l+\beta+1)\ \Gamma(n+\alpha+\beta+1)\ \Gamma(p+l+\alpha+2)\ (n-l)!\ l!}.$$

Proof. For $p-\beta<n$ one has

$$\int_0^1 x^{p-\beta}\ P_n^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx=0.$$

Hence, we suppose $p-\beta\geq n$. From analytic form of shifted Jacobi polynomials, (2.3), one has

$$\begin{aligned}\int_0^1 x^{p-\beta}\ P_n^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx=&\\ &\sum_{l=0}^{n}\frac{(-1)^{n-l}\Gamma(n+\beta+1)\ \Gamma(n+l+\alpha+\beta+1)}{\Gamma(l+\beta+1)\ \Gamma(n+\alpha+\beta+1)\ (n-l)!\ l!}\ B(p+l+1,\alpha+1)\\ &=\sum_{l=0}^{n}\frac{(-1)^{n-l}\Gamma(n+\beta+1)\ \Gamma(n+l+\alpha+\beta+1)\ \Gamma(p+l+1)\ \Gamma(\alpha+1)}{\Gamma(l+\beta+1)\ \Gamma(n+\alpha+\beta+1)\ \Gamma(p+l+\alpha+2)\ (n-l)!\ l!},\end{aligned}$$

where $B(s,t)$ is the Beta function and is defined as

$$B(s,t)=\int_0^1 v^{s-1}(1-v)^{t-1}dv=\frac{\Gamma(s)\ \Gamma(t)}{\Gamma(s+t)}.$$

$\square$

2.3. Lemma. *If $P_j^{(\alpha,\beta)}(x)$ and $P_k^{(\alpha,\beta)}(x)$ are $j-th$ and $k-th$ shifted Jacobi polynomials, respectively, then the product of $P_j^{(\alpha,\beta)}(x)$ and $P_k^{(\alpha,\beta)}(x)$ can be written as*

$$Q_{j+k}^{(\alpha,\beta)}(x)=\sum_{r=0}^{j+k}\lambda_r^{(j,k)}\ x^r.$$

Proof. Defining the $Q_{j+k}^{(\alpha,\beta)}(x)=P_j^{(\alpha,\beta)}(x)\ P_k^{(\alpha,\beta)}(x)$ as a polynomial of degree $j+k$ that can be written as

$$Q_{j+k}^{(\alpha,\beta)}(x)=(\sum_{m=0}^{j}\gamma_m^{(j)}\ x^m)\ (\sum_{n=0}^{k}\gamma_n^{(k)}\ x^n)=\sum_{r=0}^{j+k}\lambda_r^{(j,k)}\ x^r.$$

The relation between coefficients $\lambda_n^{(j,k)}$ with coefficients $\gamma_m^{(j)}$ and $\gamma_m^{(k)}$ will be as follows.

If $j\geq k$:

$$\begin{array}{l}r=0,1,...,j+k,\\ if\ \ r>j\ \ then\\ \quad\lambda_r^{(j,k)}=\sum_{l=r-j}^{k}\gamma_{r-l}^{(j)}\ \gamma_l^{(k)},\\ else\\ \quad r_1=min\{r,k\},\\ \quad\lambda_r^{(j,k)}=\sum_{l=0}^{r_1}\gamma_{r-l}^{(j)}\ \gamma_l^{(k)},\\ end.\end{array}\tag{2.7}$$

If $j<k$:

$$\begin{array}{l}r=0,1,...,j+k,\\ if\ \ r\leq j\ \ then\\ \quad r_1=min\{r,j\},\\ \quad\lambda_r^{(j,k)}=\sum_{l=0}^{r_1}\gamma_{r-l}^{(j)}\ \gamma_l^{(k)},\\ else\\ \quad r_2=min\{r,k\},\\ \quad\lambda_r^{(j,k)}=\sum_{l=r-j}^{r_2}\gamma_{r-l}^{(j)}\ \gamma_l^{(k)},\\ end.\end{array}\tag{2.8}$$

Thus, the coefficients $\lambda_r^{(j,k)}$ is determined. $\square$

2.4. Lemma. *If $P_i^{(\alpha,\beta)}(x)$, $P_j^{(\alpha,\beta)}(x)$ and $P_k^{(\alpha,\beta)}(x)$ are $i-$, $j-$ and $k-th$ shifted Jacobi polynomials, respectively, then*

$$\int_0^1 P_i^{(\alpha,\beta)}(x)\ P_j^{(\alpha,\beta)}(x)\ P_k^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx =$$
$$\sum_{n=0}^{j+k}\sum_{l=0}^{i} \frac{(-1)^{i-l}\lambda_n^{(j,k)}\ \Gamma(i+\beta+1)\ \Gamma(i+l+\alpha+\beta+1)\ \Gamma(n+l+\beta+1)\ \Gamma(\alpha+1)}{\Gamma(l+\beta+1)\ \Gamma(i+\alpha+\beta+1)\ \Gamma(n+l+\alpha+\beta+2)\ (i-l)!\ l!},$$

where $\lambda_n^{(j,k)}$ has been introduced in Lemma 2.3.

Proof. Assuming that $P_j^{(\alpha,\beta)}(x)\ P_k^{(\alpha,\beta)}(x) = Q_{j+k}^{(\alpha,\beta)}(x)$. Using of (2.3), Lemmas 2.2 and 2.3 leads to

$$\begin{aligned}
&\int_0^1 P_i^{(\alpha,\beta)}(x)\ P_j^{(\alpha,\beta)}(x)\ P_k^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx = \\
&\int_0^1 P_i^{(\alpha,\beta)}(x)\ Q_{j+k}^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx \\
&= \sum_{n=0}^{j+k} \lambda_n^{(j,k)} \int_0^1 x^n\ P_i^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx \\
&= \sum_{n=0}^{j+k}\sum_{l=0}^{i} \frac{(-1)^{i-l}\Gamma(i+\beta+1)\ \Gamma(i+l+\alpha+\beta+1)}{\Gamma(l+\beta+1)\ \Gamma(i+\alpha+\beta+1)\ (i-l)!\ l!}\ B(n+l+\beta+1,\alpha+1),
\end{aligned}$$

Thus, desirable result is obtained. $\square$

In performing arithmetic and other operations on the Jacobi bases, we frequently encounter the integration of the vector $\Phi(x)$ defined in (2.6) which is called the operational matrix of the integration. Hence, the matrix relations must be obtained. In this section, this operational matrix will be derived, then it will be extended to two dimensional case in next section. To this end, some useful lemmas and theorems are stated.

2.5. Theorem. *Let $\Phi(x)$ be shifted Jacobi vector in (2.6). Then*

$$\int_0^x \Phi(t)\ dt \simeq P\ \Phi(x),$$

where P is the $(N+1)\times(N+1)$ operational matrix of the integration and is defined by:

$$P = \begin{bmatrix} \Omega(0,0) & \Omega(0,1) & \dots & \Omega(0,N) \\ \Omega(1,0) & \Omega(1,1) & \dots & \Omega(1,N) \\ \vdots & \vdots & \ddots & \vdots \\ \Omega(N,0) & \Omega(N,1) & \dots & \Omega(N,N) \end{bmatrix},$$

where

$$\Omega(i,j) = \sum_{l=0}^{i} \omega_{ijl}, \qquad i,j = 0,1,...,N, \tag{2.9}$$

and ω_{ijl} are given by

$$\omega_{ijl} = \frac{(-1)^{i-l}\ \Gamma(i+\beta+1)\ \Gamma(i+l+\alpha+\beta+1)}{\Gamma(l+\beta+1)\ \Gamma(i+\alpha+\beta+1)\ (l+1)!\ (i-l)!}$$
$$\times \sum_{k=0}^{j} \frac{(-1)^{j-k}\ \Gamma(j+k+\alpha+\beta+1)\ \Gamma(j+\beta+1)\ \Gamma(k+l+\beta+2)\ \Gamma(\alpha+1)}{\theta_j\ \Gamma(k+\beta+1)\ \Gamma(j+\alpha+\beta+1)\ \Gamma(k+l+\alpha+\beta+3)\ k!\ (j-k)!},$$
$$i,j=0,1,...,N, \quad l=0,1,...,i.$$

Proof. Integrating the analytical form of $P_i^{(\alpha,\beta)}(x)$, i.e. (2.3), from 0 to x leads to

$$\int_0^x P_i^{(\alpha,\beta)}(t)\ dt = \sum_{l=0}^{i} \frac{(-1)^{i-l}\ \Gamma(i+\beta+1)\ \Gamma(i+l+\alpha+\beta+1)\ x^{l+1}}{\Gamma(l+\beta+1)\ \Gamma(i+\alpha+\beta+1)\ (l+1)!\ (i-l)!}. \tag{2.10}$$

Now, one can approximate x^{l+1} in terms of shifted Jacobi polynomials as

$$x^{l+1} = \sum_{k=0}^{N} a_{l,j}\ P_j^{(\alpha,\beta)}(x),$$

where

$$a_{l,j} = \frac{1}{\theta_j} \int_0^1 x^{l+1}\ P_j^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx.$$

But according to Lemma 2.2 one has,

$$\int_0^1 x^{l+1}\ P_j^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx =$$
$$\sum_{k=0}^{j} \frac{(-1)^{(j-k)}\ \Gamma(j+k+\alpha+\beta+1)\ \Gamma(j+\beta+1)\ \Gamma(k+l+\beta+2)\ \Gamma(\alpha+1)}{\Gamma(k+\beta+1)\ \Gamma(j+\alpha+\beta+1)\ \Gamma(k+l+\alpha+\beta+3)\ k!\ (j-k)!}.$$

Therefore, (2.10) will be as follows:

$$\begin{aligned}\int_0^x P_i^{(\alpha,\beta)}(t)\ dt &= \sum_{j=0}^{N} \Bigg\{ \sum_{l=0}^{i} \frac{(-1)^{i-l}\ \Gamma(i+\beta+1)\ \Gamma(i+l+\alpha+\beta+1)}{\Gamma(l+\beta+1)\ \Gamma(i+\alpha+\beta+1)\ (l+1)!\ (i-l)!} \\ &\times \sum_{k=0}^{j} \frac{(-1)^{j-k}\ \Gamma(j+k+\alpha+\beta+1)\ \Gamma(j+\beta+1)\ \Gamma(k+l+\beta+2)}{\theta_j\ \Gamma(k+\beta+1)\ \Gamma(j+\alpha+\beta+1)\ \Gamma(k+l+\alpha+\beta+3)\ (j-k)!\ (k)!} \Bigg\} \\ &\times\ P_j^{(\alpha,\beta)}(x) \\ &= \sum_{j=0}^{N}\ \Omega(i,j)\ P_j^{(\alpha,\beta)}(x).\end{aligned} \tag{2.11}$$

where $\Omega(i,j)$ are given in (2.9). Accordingly, rewriting (2.11) as a vector form gives

$$\int_0^x P_i^{(\alpha,\beta)}(t)\ dt = [\Omega(i,0), \Omega(i,1), ..., \Omega(i,N)]\ \Phi(x), \quad i=0,1,..,N.$$

This leads to the desired result. □

3. Two variables Jacobi polynomials and their operational matrices

Now in this section, two variables Jacobi polynomials can be defined by means of one variable Jacobi polynomials as follows:

3.1. Definition. Let $\{P_n^{(\alpha,\beta)}(x)\}_{n=0}^{\infty}$ be the sequence of one variable shifted Jacobi polynomials on $D = [0,1]$. Two variables Jacobi polynomials, $\{P_{m,n}^{(\alpha,\beta)}(x,y)\}_{m,n=0}^{\infty}$ are defined on $D^2 = [0,1] \times [0,1]$ as:

$$P_{m,n}^{(\alpha,\beta)}(x,y) = P_m^{(\alpha,\beta)}(x)\ P_n^{(\alpha,\beta)}(y), \qquad (x,y) \in D^2. \tag{3.1}$$

The family $\{P_{m,n}^{(\alpha,\beta)}(x,y)\}_{m,n=0}^{\infty}$ is orthogonal with weighted function $w^{(\alpha,\beta)}(x,y) = w^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(y)$ on D^2 and forms a basis for $L^2(D^2)$.

3.2. Theorem. *The basis $\{P_{m,n}^{(\alpha,\beta)}(x,y)\}_{m,n=0}^{\infty}$ is orthogonal on D^2.*

Proof. One has

$$\begin{aligned}\int_0^1\int_0^1 P_{m,n}^{(\alpha,\beta)}(x,y)\ P_{k,l}^{(\alpha,\beta)}(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy = \\ \int_0^1 P_m^{(\alpha,\beta)}(x)P_k^{(\alpha,\beta)}(x)w^{(\alpha,\beta)}(x)\ dx \\ \times \int_0^1 P_n^{(\alpha,\beta)}(y)P_l^{(\alpha,\beta)}(y)\ w^{(\alpha,\beta)}(y)\ dy \\ = \begin{cases} \theta_m\ \theta_n, & (m,n) = (k,l), \\ 0, & (m,n) \neq (k,l) \text{ or } m \neq k \text{ or } n \neq l. \end{cases}\end{aligned}$$

□

Similar to one variable case, a two variables continuous function $u(x,y)$ defined over D^2 may be expanded by the two variables Jacobi polynomials as:

$$u(x,y) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} C_{ij}\ P_{i,j}^{(\alpha,\beta)}(x,y), \tag{3.2}$$

where the Jacobi coefficients, C_{ij}, are obtained as:

$$C_{ij} = \frac{1}{\theta_i\theta_j}\int_0^1\int_0^1 P_{i,j}^{(\alpha,\beta)}(x,y)\ u(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy.$$

If the infinite series in (3.2) is truncated up to their $(N+1)-$terms in terms of both two variables x and y then it can be written as:

$$u(x,y) \simeq u_N(x,y) = \sum_{i=0}^{N}\sum_{j=0}^{N} C_{ij}\ P_{i,j}^{(\alpha,\beta)}(x,y) = \Phi^T(x,y)\ C, \tag{3.3}$$

where C and $\Phi(x,y)$ are Jacobi coefficients and Jacobi polynomials vectors, respectively:

$$C = [C_{00}, C_{01}, ..., C_{0N}, C_{10}, ..., C_{1N}, ..., C_{N1}, ..., C_{NN}]^T,$$

$$\begin{aligned}\Phi(x,y) = [&P_{0,0}^{(\alpha,\beta)}(x,y), P_{0,1}^{(\alpha,\beta)}(x,y), ..., P_{0,N}^{(\alpha,\beta)}(x,y), P_{1,0}^{(\alpha,\beta)}(x,y), \\ &..., P_{1,N}^{(\alpha,\beta)}(x,y), ..., P_{N,0}^{(\alpha,\beta)}(x,y), ..., P_{N,N}^{(\alpha,\beta)}(x,y)]^T.\end{aligned} \tag{3.4}$$

A function of four variables, $k(x,y,t,s)$, on D^4 may be approximated based on Jacobi operational matrix as:

$$k(x,y,t,s) \simeq \Phi^T(x,y)\ K\ \Phi(t,s),$$

where $\Phi(x,y)$ is two variables Jacobi vector defined by (3.4) and K is a $(N+1)^2 \times (N+1)^2$ known matrix. Before proceeding, let us represent the partial series (3.3) as following form:

$$S_N(x,y) = \sum_{i=0}^{M} d_i\ Q_i^{(\alpha,\beta)}(x,y), \tag{3.5}$$

where

$$d_i = C_{rs}, \quad Q_i(x,y) = P_{r,s}^{(\alpha,\beta)}(x,y),$$
$$r = \lfloor \frac{i}{N+1} \rfloor, \quad s = i - r(N+1), \quad M = (N+1)^2 - 1.$$

Hence, one has

$$\int_0^1 \int_0^1 Q_k^{(\alpha,\beta)}(x,y)\ Q_l^{(\alpha,\beta)}(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy =$$
$$\int_0^1 P^{(\alpha,\beta)}_{\lfloor \frac{k}{N+1} \rfloor}(x)\ P^{(\alpha,\beta)}_{\lfloor \frac{l}{N+1} \rfloor}(x)\ w^{(\alpha,\beta)}(x)\ dx$$
$$\times \int_0^1 P^{(\alpha,\beta)}_{k-\lfloor \frac{k}{N+1} \rfloor (N+1)}(y)\ P^{(\alpha,\beta)}_{l-\lfloor \frac{l}{N+1} \rfloor (N+1)}(y)\ w^{(\alpha,\beta)}(y)\ dy$$
$$= \theta_{\lfloor \frac{k}{N+1} \rfloor}\ \theta_{k-\lfloor \frac{k}{N+1} \rfloor (N+1)} = \theta_r\ \theta_s, \qquad r = \lfloor \frac{k}{N+1} \rfloor, \quad s = k - r(N+1).$$

3.3. Remark. The relation (3.5) can be rewritten as:

$$S_N(x,y) = \sum_{i=0}^{M} \omega_i\ R_i^{(\alpha,\beta)}(x,y), \tag{3.6}$$

where

$$\omega_i = d_i \sqrt{\theta_r \theta_s}, \quad R_i^{(\alpha,\beta)}(x,y) = \frac{Q_i^{(\alpha,\beta)}(x,y)}{\sqrt{\theta_r \theta_s}}, \qquad r = \lfloor \frac{i}{N+1} \rfloor, \quad s = i - r(N+1).$$

This shows the sequence $\{R_i^{(\alpha,\beta)}(x,y)\}_{i=0}^{M}$ is orthonormal. That is

$$\int_0^1 \int_0^1 R_k^{(\alpha,\beta)}(x,y)\ R_l^{(\alpha,\beta)}(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy = \begin{cases} 1, & k = l, \\ 0, & otherwise \end{cases}$$

Now, the two dimensional operational matrices of integration in x and $y-$direction are defined by following theorem:

3.4. Theorem. *If P is the operational matrix in one dimensional case then the operational matrices of integration in x and $y-$direction are defined as follows.*

$$a) \quad \int_0^x \Phi(t,y)\ dt \simeq P_x\ \Phi(x,y) = (P \otimes I)\ \Phi(x,y),$$

$$b) \quad \int_0^y \Phi(x,s)\ ds \simeq P_y\ \Phi(x,y) = (I \otimes P)\ \Phi(x,y),$$

where P_x and P_y are $(N+1)^2 \times (N+1)^2$ operational matrices of integration in the directions x and y, respectively, I is $(N+1) \times (N+1)$ identity matrix and $\otimes$ denotes the Kronecker product and is defined for two arbitrary matrices A and B as $A \otimes B = (a_{ij}B)$.

Proof. a) Suppose r_j be jth row of matrix P. One has

$$\int_0^x P_j^{(\alpha,\beta)}(t)\ dt = r_j^T \Phi(x).$$

Also, noting the definition of the vector $\Phi(x,y)$ one has

$$\begin{aligned} \Phi(x,y) = [&P_0^{(\alpha,\beta)}(x)P_0^{(\alpha,\beta)}(y), ..., P_0^{(\alpha,\beta)}(x)P_N^{(\alpha,\beta)}(y), \\ &..., P_N^{(\alpha,\beta)}(x)P_0^{(\alpha,\beta)}(y), ..., P_N^{(\alpha,\beta)}(x)P_N^{(\alpha,\beta)}(y)]^T. \end{aligned} \tag{3.7}$$

Integrating of (3.7) from 0 to x yields

$$\begin{aligned}
\int_0^x \Phi(t,y)dt = &[P_0^{(\alpha,\beta)}(y)\int_0^x P_0^{(\alpha,\beta)}(t)dt, ..., P_N^{(\alpha,\beta)}(y)\int_0^x P_0^{(\alpha,\beta)}(t)dt,\\
&..., P_N^{(\alpha,\beta)}(y)\int_0^x P_N^{(\alpha,\beta)}(t)dt]^T\\
&= [r_0.\Phi(x)\ P_0^{(\alpha,\beta)}(y), ..., r_0.\Phi(x)\ P_N^{(\alpha,\beta)}(y), ..., r_N.\Phi(x)\ P_0^{(\alpha,\beta)}(y),\\
&..., r_N.\Phi(x)\ P_N^{(\alpha,\beta)}(y)]^T\\
&= [r_0\ [P_0^{(\alpha,\beta)}(x)\ P_0^{(\alpha,\beta)}(y), ..., P_N^{(\alpha,\beta)}(x)\ P_0^{(\alpha,\beta)}(y)],\\
&..., r_N\ [P_0^{(\alpha,\beta)}(x)\ P_N^{(\alpha,\beta)}(y), ..., P_N^{(\alpha,\beta)}(x)\ P_N^{(\alpha,\beta)}(y)]]^T
\end{aligned}$$

$$= \begin{bmatrix}
P_{00} & 0 & \dots & 0 & P_{01} & 0 & \dots & 0 & \dots & P_{0N} & 0 & \dots & 0\\
0 & P_{00} & \dots & 0 & 0 & P_{01} & \dots & 0 & \dots & 0 & P_{0N} & \dots & 0\\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & 0\\
0 & 0 & \dots & P_{00} & 0 & 0 & \dots & P_{01} & \dots & 0 & 0 & \dots & P_{0N}\\
 & \vdots & & & & \vdots & & & & \vdots & & & \\
P_{N0} & 0 & \dots & 0 & P_{N1} & 0 & \dots & 0 & \dots & P_{NN} & 0 & \dots & 0\\
0 & P_{N0} & \dots & 0 & 0 & P_{N1} & \dots & 0 & \dots & 0 & P_{NN} & \dots & 0\\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\
0 & 0 & \dots & P_{N0} & 0 & 0 & \dots & P_{N1} & \dots & 0 & 0 & \dots & P_{NN}
\end{bmatrix}$$

$$\times \begin{bmatrix}
P_0^{(\alpha,\beta)}(x)\ P_0^{(\alpha,\beta)}(y)\\
P_0^{(\alpha,\beta)}(x)\ P_1^{(\alpha,\beta)}(y)\\
\vdots\\
P_0^{(\alpha,\beta)}(x)\ P_N^{(\alpha,\beta)}(y)\\
\vdots\\
P_N^{(\alpha,\beta)}(x)\ P_0^{(\alpha,\beta)}(y)\\
P_N^{(\alpha,\beta)}(x)\ P_1^{(\alpha,\beta)}(y)\\
\vdots\\
P_N^{(\alpha,\beta)}(x)\ P_N^{(\alpha,\beta)}(y)
\end{bmatrix}$$

$$= \begin{bmatrix}
P_{00}I & P_{01}I & \dots & P_{0N}I\\
P_{10}I & P_{11}I & \dots & P_{1N}I\\
\vdots & \vdots & \ddots & \vdots\\
P_{N0}I & P_{N1}I & \dots & P_{NN}I
\end{bmatrix}
\begin{bmatrix}
P_0^{(\alpha,\beta)}(x)\ \Phi(y)\\
P_1^{(\alpha,\beta)}(x)\ \Phi(y)\\
\vdots\\
P_N^{(\alpha,\beta)}(x)\ \Phi(y)
\end{bmatrix}$$

$$= (P \otimes I)\Phi(x,y).$$

Where P_{ij} denotes $(i,j)-$th entry of the matrix P. The case (b) is proven similarly. □

The following property of the product of two vectors $\Phi(x,y)$ and $\Phi^T(x,y)$ will also be used.

$$\Phi(x,y)\ \Phi^T(x,y)\ C \simeq \tilde{C}\ \Phi(x,y), \tag{3.8}$$

where C and $\tilde{C}$ are a $(N+1)^2 \times 1$ vector and a $(N+1)^2 \times (N+1)^2$ operational matrix of product, respectively.

3.5. Theorem. *The entries of the matrix $\tilde{C}$, in (3.8), are computed as:*

$$\tilde{C}_{m(N+1)+n,k(N+1)+l} = \frac{1}{\theta_k \theta_l} \sum_{r=0}^{N} \sum_{s=0}^{N} C_{rs}\ q_{mrk}\ q_{nsl}, \qquad m,n,k,l = 0,1,...,N.$$

Proof. The left side of equality (3.8) is as follows:

$$\Phi(x,y)\ \Phi^T(x,y)\ C = \begin{bmatrix} \sum_{r=0}^{N} \sum_{s=0}^{N} C_{rs}\ P_{0,0}^{(\alpha,\beta)}(x,y)\ P_{r,s}^{(\alpha,\beta)}(x,y) \\ \sum_{r=0}^{N} \sum_{s=0}^{N} C_{rs}\ P_{0,1}^{(\alpha,\beta)}(x,y)\ P_{r,s}^{(\alpha,\beta)}(x,y) \\ \vdots \\ \sum_{r=0}^{N} \sum_{s=0}^{N} C_{rs}\ P_{N,N}^{(\alpha,\beta)}(x,y)\ P_{r,s}^{(\alpha,\beta)}(x,y) \end{bmatrix}.$$

Consider the $(p+1)$th row of above vector. One puts

$$P_{i,j}^{(\alpha,\beta)}(x,y)\ P_{r,s}^{(\alpha,\beta)}(x,y) = \sum_{k=0}^{N} \sum_{l=0}^{N} u_{kl}\ P_{k,l}^{(\alpha,\beta)}(x,y), \tag{3.9}$$

Multiplying both sides of (3.9) by $P_{m,n}^{(\alpha,\beta)}(x,y),\ m,n = 0,1,...,N$, and integrating the result from 0 to 1 yields:

$$\begin{aligned} &\int_0^1 \int_0^1 P_{i,j}^{(\alpha,\beta)}(x,y)\ P_{r,s}^{(\alpha,\beta)}(x,y)\ P_{m,n}^{(\alpha,\beta)}(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy = \\ &\qquad \sum_{k=0}^{N} \sum_{l=0}^{N} u_{kl} \int_0^1 \int_0^1 P_{k,l}^{(\alpha,\beta)}(x,y)\ P_{m,n}^{(\alpha,\beta)}(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy \\ &\qquad = u_{mn}\ \theta_m\ \theta_n. \end{aligned}$$

Therefore,

$$\begin{aligned} u_{mn} = &\frac{1}{\theta_m\ \theta_n} \int_0^1 P_i^{(\alpha,\beta)}(x)\ P_r^{(\alpha,\beta)}(x)\ P_m^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)dx \\ &\times \int_0^1 P_j^{(\alpha,\beta)}(y)\ P_s^{(\alpha,\beta)}(y)\ P_n^{(\alpha,\beta)}(y)\ w^{(\alpha,\beta)}(y)\ dy. \end{aligned}$$

Now suppose $\int_0^1 P_i^{(\alpha,\beta)}(x)\ P_r^{(\alpha,\beta)}(x)\ P_m^{(\alpha,\beta)}(x)\ w^{(\alpha,\beta)}(x)\ dx = q_{irm}$. one gets

$$u_{mn} = \frac{q_{irm}\ q_{jsn}}{\theta_m\ \theta_n}.$$

Substituting u_{mn} into (3.9) one has:

$$P_{i,j}^{(\alpha,\beta)}(x,y)\ P_{r,s}^{(\alpha,\beta)}(x,y) = \sum_{k=0}^{N} \sum_{l=0}^{N} \frac{q_{irm}\ q_{jsn}}{\theta_m\ \theta_n}\ P_{k,l}^{(\alpha,\beta)}(x,y),$$

Hence each component in the left side of relation (3.8) will be as follows:

$$\begin{aligned} &\sum_{k=0}^{N} \sum_{l=0}^{N} \left\{ \sum_{r=0}^{N} \sum_{s=0}^{N} \frac{C_{rs}\ q_{mrk}\ q_{nsl}}{\theta_k \theta_l} \right\} P_{k,l}^{(\alpha,\beta)}(x,y) \\ &\qquad = \sum_{k=0}^{N} \sum_{l=0}^{N} \tilde{C}_{m(N+1)+n,k(N+1)+l}\ P_{k,l}^{(\alpha,\beta)}(x,y), \\ &\qquad m,n = 0,1,..,N. \end{aligned}$$

Thus, the desirable result is obtained. $\square$

The Next theorem presents the general formula approximating the nonlinear term $v^r(x,y)\ u^s(x,y)$ which may appear in nonlinear equations.

3.6. Theorem. *If c and υ are the $(N+1)^2$ vectors, $\tilde{c}$ and $\tilde{\upsilon}$ are the $(N+1)^2 \times (N+1)$ operational matrices of the product such that*

$$u(x,y) \simeq \Phi^T(x,y)\ c = c^T\ \Phi(x,y), \quad v(x,y) \simeq \Phi^T(x,y)\ \upsilon = \upsilon^T\ \Phi(x,y),$$
$$\Phi(x,y)\ \Phi^T(x,y)\ c \simeq \tilde{c}\ \Phi(x,y)$$

and $\Phi(x,y)\ \Phi^T(x,y)\ \upsilon \simeq \tilde{\upsilon}\ \Phi(x,y)$, then the following proposition is hold:

$$v^r(x,y)\ u^s(x,y) \simeq \upsilon^T\ (\tilde{\upsilon})^{r-1}\ \tilde{B}_{s-1}\ \Phi(x,y), \quad B_{s-1} = (\tilde{c}^T)^{s-1}\ c,$$
$$r, s = 1, 2, \dots\ .$$

Proof. One has

$$u^2(x,y) \simeq (\Phi^T(x,y)\ c)^2 = c^T\ \Phi(x,y)\ \Phi^T(x,y)\ c \simeq c^T\ \tilde{c}\ \Phi(x,y),$$

So, by use of induction, $u^s(x,y)$ will be approximated as

$$u^s(x,y) \simeq c^T\ (\tilde{c})^{s-1}\ \Phi(x,y), \qquad s = 1, 2, \dots\ .$$

To similar way, $v^r(x,y)$ is approximated as

$$v^r(x,y) \simeq \upsilon^T\ (\tilde{\upsilon})^{r-1}\ \Phi(x,y), \qquad r = 1, 2, \dots\ .$$

By using the expressed relations and induction is easily seen,

$$v^r(x,y)\ u^s(x,y) \simeq \upsilon^T\ (\tilde{\upsilon})^{r-1}\ \tilde{B}_{s-1}\ \Phi(x,y), \quad B_{s-1} = (\tilde{c}^T)^{s-1}\ c,$$
$$r, s = 1, 2, \dots\ .$$

□

4. Best approximation and Convergence analysis

In this section, the theorems on existence and uniqueness of best approximation, convergence analysis and error estimation of the proposed method are provided. For this reason, first the space $\mathbb{P}^M$ is considered as follows:

4.1. Definition. The set of all the linear combinations of $R_0^{(\alpha,\beta)}(x,y)$, $R_1^{(\alpha,\beta)}(x,y)$, ..., $R_M^{(\alpha,\beta)}(x,y)$, which $M = (N+1)^2 - 1$, is represented by $\mathbb{P}^M$. On the other hand,

$$\mathbb{P}^M = span\ \{R_0^{(\alpha,\beta)}(x,y), R_1^{(\alpha,\beta)}(x,y), \dots, R_M^{(\alpha,\beta)}(x,y)\}, \tag{4.1}$$

where two variables orthonormal polynomials $R_i^{(\alpha,\beta)}(x,y)$ are introduced by (3.6).

The following lemma is useful to prove the convexity and completeness properties of space $\mathbb{P}^M$.

4.2. Lemma. *There is a number $\eta > 0$ such that for every choice of scalars $\alpha_0, \alpha_1, \dots, \alpha_M$ one has*

$$\|\ \alpha_0\ R_0^{(\alpha,\beta)}(x,y) + \alpha_1\ R_1^{(\alpha,\beta)}(x,y) + \dots + \alpha_M\ R_M^{(\alpha,\beta)}(x,y)\ \| \geq$$
$$\eta\ (|\ \alpha_0\ | + |\ \alpha_1\ | + \dots + |\ \alpha_M\ |).$$

Proof. See [24]. □

4.3. Theorem. *The space $\mathbb{P}^M$, defined by (4.1), is convex and complete.*

Proof. Suppose $v_1(x,y)$ and $v_2(x,y) \in \mathbb{P}^M$. One has for $0 < \lambda < 1$

$$\lambda\ v_1(x,y) + (1-\lambda)\ v_2(x,y) = \sum_{i=0}^{M} (\lambda\ \omega_i^1 + (1-\lambda)\ \omega_i^2)\ R_i^{(\alpha,\beta)}(x,y) \in \mathbb{P}^M.$$

This shows that $\mathbb{P}^M$ is convex.

For proving the completeness property, let us consider Cauchy sequence

$$\{w_n(x,y)\}_{n=0}^{\infty}\ \in \mathbb{P}^M.$$

Then each $w_n(x,y)$ is a unique representation of the form

$$w_n(x,y) = \sum_{i=0}^{M} \lambda_i^{(n)}\ R_i^{(\alpha,\beta)}(x,y).$$

Since $\{w_n(x,y)\}_{n=0}^{\infty}$ is a Cauchy sequence, for every $\varepsilon > 0$ there is a N' such that $\| w_m(x,y) - w_n(x,y) \| < \varepsilon$ where $m, n > N'$. From this and Lemma 4.2, one has for $\eta > 0$

$$\varepsilon > \| w_m(x,y) - w_n(x,y) \| = \| \sum_{i=0}^{M} (\lambda_i^{(m)} - \lambda_i^{(n)})\ R_i^{(\alpha,\beta)}(x,y) \| \geq \eta\ \sum_{i=0}^{M} |\ \lambda_i^{(m)} - \lambda_i^{(n)}\ | \ .$$

Division by $\eta > 0$ gives

$$|\ \lambda_i^{(m)} - \lambda_i^{(n)}\ | \leq\ \sum_{i=0}^{M} |\ \lambda_i^{(m)} - \lambda_i^{(n)}\ | <\ \frac{\varepsilon}{\eta},$$

This shows that each of the $M+1$ sequences $\{\lambda_i^{(n)}\}_{n=0}^{\infty},\ i = 0, 1, ..., M$, is Cauchy in $\mathbb{R}$. Hence it converges. Let λ_i denotes the limit. Using this $M+1$ limits $\lambda_0, \lambda_1, ..., \lambda_M$, one defines

$$\bar{w}(x,y) = \sum_{i=0}^{M} \lambda_i\ R_i^{(\alpha,\beta)}(x,y).$$

Clearly, $\bar{w}(x,y)\ \in\ \mathbb{P}^M$. Furthermore,

$$\begin{aligned} \| w_n(x,y) - \bar{w}(x,y) \| \ &= \| \sum_{i=0}^{M} (\lambda_i^{(n)} - \lambda_i)\ R_i^{(\alpha,\beta)}(x,y) \| \\ &\leq\ \sum_{i=0}^{M} |\ \lambda_i^{(n)} - \lambda_i\ |\ \| R_i^{(\alpha,\beta)}(x,y) \| \ . \end{aligned}$$

On the right, $\lambda_i^{(n)} \to \lambda_i$. Hence $\| w_n(x,y) - \bar{w}(x,y) \| \to 0$, that is, $w_n(x,y) \to \bar{w}(x,y)$. This shows that $\{w_n(x,y)\}_{n=0}^{\infty}$ is convergent in $\mathbb{P}^M$, and the completeness of $\mathbb{P}^M$ is proven. □

4.4. Theorem. *For every given continuous function $u(x,y)$ there exists a unique $u_M(x,y) \in \mathbb{P}^M$ such that*

$$\delta\ =\ \inf_{\tilde{u} \in \mathbb{P}^M} \| u(x,y) - \tilde{u}(x,y) \| = \| u(x,y) - u_M(x,y) \| \ .$$

Proof. Existence. By the definition of an infimum, there is a sequence $\{w_n(x,y)\}_{n=0}^{\infty}$ in $\mathbb{P}^M$ such that $\delta_n \to \delta$ where $\delta_n\ = \| u(x,y) - w_n(x,y) \|$. We show that $\{w_n(x,y)\}_{n=0}^{\infty}$ is Cauchy. Writing $v_n(x,y) = u(x,y) - w_n(x,y)$, one has $\| v_n(x,y) \| = \delta_n$ and

$$\| v_m(x,y) + v_n(x,y) \| = 2\ \| \frac{1}{2}(w_m(x,y) + w_n(x,y)) - u(x,y) \| \geq 2\ \delta,$$

because $\mathbb{P}^M$ is convex, so that $\frac{1}{2}(w_m(x,y)+w_n(x,y)) \in \mathbb{P}^M$. Furthermore, one has $v_m(x,y)-v_n(x,y)=w_n(x,y)-w_m(x,y)$. Hence by the parallelogram equality

$$\begin{aligned}\| w_n(x,y)-w_m(x,y) \|^2 &= \| v_m(x,y)-v_n(x,y) \|^2\\ &= - \| v_m(x,y)+v_n(x,y) \|^2\\ &+ 2\ (\| v_m(x,y) \|^2 + \| v_n(x,y) \|^2)\\ &\leq -(2\ \delta)^2 + 2(\delta_m^2+\delta_n^2) < \varepsilon^2.\end{aligned}$$

This implies that $\{w_n(x,y)\}_{n=0}^{\infty}$ is Cauchy. Since $\mathbb{P}^M$ is complete, $\{w_n(x,y)\}_{n=0}^{\infty}$ converges, say, $w_n(x,y) \to \bar{w}(x,y) \in \mathbb{P}^M$. Since $\bar{w}(x,y) \in \mathbb{P}^M$, one has $\| u(x,y)-\bar{w}(x,y) \| \geq \delta$. Also,

$$\begin{aligned}\| u(x,y)-\bar{w}(x,y) \| &\leq \| u(x,y)-w_n(x,y) \| + \| w_n(x,y)-\bar{w}(x,y) \|\\ &= \delta_n + \| w_n(x,y)-\bar{w}(x,y) \| \to \delta.\end{aligned}$$

This shows that $\| u(x,y)-\bar{w}(x,y) \| = \delta$.

Uniqueness. Let us assume that $\bar{w}(x,y) \in \mathbb{P}^M$ and $w_0(x,y) \in \mathbb{P}^M$ both satisfy

$$\| u(x,y)-\bar{w}(x,y) \| = \delta, \qquad \| u(x,y)-w_0(x,y) \| = \delta.$$

By the parallelogram equality,

$$\begin{aligned}\| \bar{w}(x,y)-w_0(x,y) \|^2 &= \| (\bar{w}(x,y)-u(x,y))-(w_0(x,y)-u(x,y)) \|^2\\ &= 2\ \| \bar{w}(x,y)-u(x,y) \|^2 + 2\ \| w_0(x,y)-u(x,y) \|^2\\ &- \| \bar{w}(x,y)-u(x,y)) + (w_0(x,y)-u(x,y)) \|^2\\ &= 4\ \delta^2 - 4\ \| \frac{1}{2}(\bar{w}(x,y)+w_0(x,y))-u(x,y) \|^2 \leq 0,\end{aligned}$$

because $\frac{1}{2}(\bar{w}(x,y)+w_0(x,y)) \in \mathbb{P}^M$. So that $\bar{w}(x,y)=w_0(x,y)$.

Orthogonality. We assume there be a $0 \neq w_1(x,y) \in \mathbb{P}^M$ such that

$$(z(x,y),w_1(x,y)) = \gamma \neq 0$$

where $z(x,y)=u(x,y)-\bar{w}(x,y)$ and $(.,.)$ denotes the inner product. Furthermore, for any scalar η,

$$\begin{aligned}\| z(x,y)-\eta\ w_1(x,y) \|^2 &= (z(x,y)-\eta\ w_1(x,y), z(x,y)-\eta\ w_1(x,y))\\ &= \| z(x,y) \|^2 - \bar{\eta}\ \gamma - \eta\ (\bar{\gamma} - \bar{\eta}\ \ \| w_1(x,y) \|^2),\end{aligned}$$

Choosing $\bar{\eta} = \frac{\bar{\gamma}}{\| w_1(x,y) \|^2}$ yields

$$\| z(x,y)-\eta\ w_1(x,y) \|^2 = \| z(x,y) \|^2 - \frac{\bar{\gamma}}{\| w_1(x,y) \|^2} = \delta^2 - \frac{\bar{\gamma}}{\| w_1(x,y) \|^2} \geq \delta^2,$$

But this is impossible because one has $\| z(x,y)-\eta\ w_1(x,y) \| \geq \delta$ by the definition of δ. Hence the assumption can not be hold. So $(z(x,y),\tilde{u}(x,y))=0,\ \forall\ \tilde{u}(x,y) \in\ \mathbb{P}^M$.

Now it is shown that $\bar{w}(x,y)=u_M(x,y)$. It was proven that $\bar{w}(x,y)$ is the best approximation for $u(x,y)$. So,

$$\forall j,\ j=0,1,..,M, \qquad (u(x,y)-\bar{w}(x,y),R_j(x,y))=0,$$

where two variables polynomials $R_i^{(\alpha,\beta)}(x,y)$ are introduced by (3.6). One has from this point,

$$\begin{aligned}(\bar{w}(x,y)-u_M(x,y),R_j(x,y)) &= (u(x,y)-(u(x,y)-\bar{w}(x,y))-u_M(x,y),R_j(x,y))\\ &= (u(x,y),R_j(x,y))-(u(x,y)-\bar{w}(x,y),R_j(x,y))\\ &- (u_M(x,y),R_j(x,y))\\ &= \omega_j-\omega_j=0, \qquad j=0,1,...,M,\end{aligned}$$

Therefore $\bar{w}(x,y) - u_M(x,y) = 0$. This shows that $u_M(x,y) = \bar{w}(x,y)$ and proof is completed. □

Two following theorems state the decaying of the Jacobi coefficients and the convergence of the best approximation.

4.5. Theorem. *The Jacobi coefficients ω_i, introduced by (3.6), decay when the number of the terms of the partial sum of the series solution, N, increases.*

Proof. Employing the ω_i and the properties of the orthogonality of $R_i(x,y)$, we have

$$\int_0^1 \int_0^1 S_M(x,y)\ u(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy = \sum_{i=0}^{M} \omega_i \int_0^1 \int_0^1 R_i(x,y)\ u(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy = \sum_{i=0}^{M} \omega_i^2.$$

If $u^2(x,y)\ w^{(\alpha,\beta)}(x,y)$ as well as $u(x,y)\ w^{(\alpha,\beta)}(x,y)$ is integrable, then

$$\begin{aligned}\int_0^1 \int_0^1 [u(x,y) - S_M(x,y)]^2\ w^{(\alpha,\beta)}(x,y)\ dxdy = &\int_0^1 \int_0^1 u^2(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy \\ &- 2 \int_0^1 \int_0^1 u(x,y)\ S_M(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy \\ &+ \int_0^1 \int_0^1 S_M^2(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy \\ &= \int_0^1 \int_0^1 u^2(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy - \sum_{i=0}^{M} \omega_i^2.\end{aligned}$$

Therefore,

$$\sum_{i=0}^{M} \omega_i^2 \ \leq \int_0^1 \int_0^1 u^2(x,y)\ w^{(\alpha,\beta)}(x,y)\ dxdy, \qquad \forall N \in \mathbb{N},$$

Consequently $\sum_{i=0}^{\infty} \omega_i^2$ is convergent and $\lim_{i\to\infty} \omega_i = 0$. □

Theorem 4.5 states the given function $u(x,y)$ may be approximated using only the finite numbers of two variables Jacobi polynomials.

4.6. Theorem. *The series solution (3.3) converges towards $u(x,y)$ in (3.2).*

Proof. Consider the relation (3.6) and define the sequence partial sums

$$\{S_M(x,y)\}_{M=0}^{\infty}$$

as follows,

$$S_M(x,y) = \sum_{i=0}^{M} \omega_i\ R_i(x,y).$$

Let suppose that the $S_M(x,y)$ and $S_L(x,y)$ are partial sums with $M > L$. It is going to prove that $\{S_M(x,y)\}_{M=0}^{\infty}$ is a Cauchy sequence in $\mathbb{P}^M$. For this purpose, it is worked out as follows:

$$\begin{aligned}
\| \sum_{i=L+1}^{M} \omega_i \ R_i(x,y) \ \|^2 &= (\sum_{i=L+1}^{M} \omega_i \ R_i(x,y), \sum_{j=L+1}^{M} \omega_j \ R_j(x,y)) \\
&= \sum_{i=L+1}^{M} \sum_{j=L+1}^{M} \omega_i \ \bar{\omega}_j (R_i(x,y), R_j(x,y)) \\
&= \sum_{i=L+1}^{M} \mid \omega_i \mid^2 .
\end{aligned}$$

That is $\| \ S_M(x,y) - S_L(x,y) \ \|^2 = \sum_{i=L+1}^{M} \mid \omega_i \mid^2$. From Bessel inequality $\sum_{i=0}^{\infty} \mid \omega_i \mid^2$ is convergent and hence $\| \ S_M(x,y) - S_L(x,y) \ \|^2 \leq \varepsilon^2$. This shows $\{S_M(x,y)\}_{M=0}^{\infty}$ is a Cauchy sequence. Since $\mathbb{P}^M$ is complete one has $S_M(x,y) \to S(x,y) \in \mathbb{P}^M$. We show $S(x,y) = u(x,y)$:

$$\begin{aligned}
(S(x,y) - u(x,y), R_j(x,y)) &= (S(x,y), R_j(x,y)) - (u(x,y), R_j(x,y)) \\
&= (\lim_{M\to\infty} S_M(x,y), R_j(x,y)) - (u(x,y), R_j(x,y)) \\
&= \lim_{M\to\infty} (S_M(x,y), R_j(x,y)) - (u(x,y), R_j(x,y)) \\
&= \omega_j - \omega_j = 0, \\
&\Rightarrow \ S(x,y) - u(x,y) = 0, \quad j = 0, 1, ..., M.
\end{aligned}$$

Hence,

$$S(x,y) = \sum_{i=0}^{\infty} \omega_i \ R_i(x,y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} C_{ij} \ P_i^{(\alpha,\beta)}(x) \ P_j^{(\alpha,\beta)}(y) = u(x,y).$$

$\square$

Whenever the solution of a problem is not known, specially in nonlinear phenomena, an error estimator is needed as an essential component of the computational algorithm. To this end, an error estimator for the proposed method is presented in this section.

For simplicity, discussed equations are written in the operator form

$$\mathcal{D}u(x,y) = f(x,y) \tag{4.2}$$

where $\mathcal{D}$ is a integro–partial differential operator. Define the error function as

$$e_N(x,y) = u(x,y) - u_N(x,y).$$

Substituting $u_N(x,y)$ into given equations yields

$$\mathcal{D}u_N(x,y) = f(x,y) + H_N(x,y), \tag{4.3}$$

where $H_N(x,y)$ is a perturbation term. Subtracting (4.3) from (4.2) gives

$$\mathcal{D}e_N(x,y) = -H_N(x,y).$$

Now, it can be proceeded by the same way as the basic problem is solved to get the estimation $e_{N,M}(x,y)$ to the error function $e_N(x,y)$. Note the stated approximations are also substituted in the conditions of the given problem. Subsequently, the conditions of new problem will be homogeneous.

5. Numerical results

In this section, four examples are given to certify the efficiency and accuracy of the proposed method where the maximum absolute and estimate errors are reported for different values of parameters α and β. Also, the absolute and estimate errors are computed at some arbitrary selected points.

5.1. Example. Consider the following linear Volterra–Fredholm integro–partial differential equation.

$$u_{xx}(x,y) + sin(xy)\ u(x,y) = xy\ \sin(xy) - \frac{1}{6}y^3 + \int_0^y \int_0^1 ts\ u_t(t,s)\ dtds, \tag{5.1}$$

with the following conditions,

$$u_x(0,y) = y, \qquad u(0,y) = 0. \tag{5.2}$$

The exact solution of this problem is $u(x,y) = xy$. First, let us consider the following approximation,

$$u_{xx}(x,y) \simeq \Phi^T(x,y)\ C. \tag{5.3}$$

Integrating (5.3) with respect to x from 0 to x, one gets the following approximation for $u_x(x,y)$.

$$u_x(x,y) \simeq \Phi^T(x,y)\ P_x^T\ C + u_x(0,y) \simeq \Phi^T(x,y)\ P_x^T\ C + \Phi^T(x,y)\ V, \tag{5.4}$$

where $u_x(0,y)$ is approximated by $\Phi^T(x,y)\ V$ which V is a $(N+1)\times 1$ known vector. Again, integrating (5.4) with respect to x from 0 to x, one obtains the following approximation for $u(x,y)$,

$$u(x,y) \simeq \Phi^T(x,y)\ (P_x^T)^2\ C + \Phi^T(x,y)\ P_x^T\ V. \tag{5.5}$$

In order to approximate the integral part in the (5.1), the kernel ts is approximated as follows:

$$ts \simeq \Phi^T(x,y)\ K\ \Phi(t,s), \tag{5.6}$$

where K is a $(N+1)^2 \times (N+1)^2$ known matrix and is determined by inner product. Now, the integral part in (5.1) is approximated as:

$$\begin{aligned}\int_0^y \int_0^1 ts\ u_t(t,s)\ dtds &\simeq \int_0^y \int_0^1 \Phi^T(x,y)\ K\ \Phi(t,s)\ \{\Phi^T(t,s)\ P_x^T\ C \\ &\quad + \Phi^T(t,s)\ V\}\ dtds \\ &\simeq \Phi^T(x,y)\ K\{\tilde{V} + \tilde{B}\}P_y\ A,\end{aligned} \tag{5.7}$$

where $\tilde{V}$ is operational matrix of product and its entries are determined in terms of the components of the vector V, $\tilde{B}$ is operational matrix of product corresponding to vector $B = P_x^T\ C$, and $A = \int_0^1 \Phi(t,y)\ dt$. Substituting the approximations (5.3)-(5.7) into (5.1), leads to the following linear algebraic equation.

$$\begin{aligned}&\Phi^T(x,y)\ C + sin(xy)\ \Phi^T(x,y)\ \{(P_x^T)^2\ C + P_x^T\ V\} \\ &\qquad - \Phi^T(x,y)\ K\{\tilde{V} + \tilde{B}\}P_y\ A - xy\ \sin(xy) + \frac{1}{6}y^3 = 0.\end{aligned} \tag{5.8}$$

Setting $N = 3$ and using the roots of $P_4^{(\alpha,\beta)}(x)$ and $P_4^{(\alpha,\beta)}(y)$ in the x and $y-$directions, (5.8) is collocated in 16 inner tensor points for different values of parameters α and β. Hereby, the (5.8) reduces the problem to solve a system of linear algebraic equations and unknown coefficients are obtained for some values of parameters α and β. Table 1 shows the maximum absolute and estimate errors of the approximate solutions for different values of α and β. Table 2 displays different values of the exact and approximate solutions in points $(x,y) = (0.1i, 0.1i), (i = 1, 2, ..., 10)$ for $\alpha = \beta = -1/4$. As can be seen from Tables the results of the solutions obtained by Jacobi polynomials method are almost the same as the results of the exact solutions.

Table 1. Maximum absolute and estimate errors of Example 5.1 for different values α and β

(α,β)	$Error_{Abs}$	$Error_{Est}$	(α,β)	$Error_{Abs}$	$Error_{Est}$
$(0,0)$	5.2715×10^{-21}	4.7132×10^{-21}	$(\frac{1}{4},\frac{1}{4})$	4.9562×10^{-20}	3.8921×10^{-21}
$(1,1)$	2.4802×10^{-20}	2.2000×10^{-20}	$(-\frac{1}{4},-\frac{1}{4})$	5.0940×10^{-19}	2.2476×10^{-19}
$(2,2)$	5.3375×10^{-20}	3.8880×10^{-21}	$(\frac{3}{4},\frac{3}{4})$	9.8549×10^{-19}	4.8000×10^{-19}
$(\frac{1}{2},\frac{1}{2})$	1.5140×10^{-21}	1.5829×10^{-21}	$(\frac{1}{10},\frac{1}{10})$	1.3500×10^{-19}	3.8930×10^{-21}

Table 2. Maximum absolute and estimate errors of Example 5.1 for various values of $\alpha = \beta = -\frac{1}{4}$

(x_i, y_i)	*Exact value*	*Approximate value*	$Error_{Abs}$	$Error_{Est}$
$(0.1, 0.1)$	0.01	0.0099999999999999999918	8.20×10^{-21}	1.6649×10^{-21}
$(0.2, 0.2$	0.04	0.040000000000000000004	4.00×10^{-21}	4.7592×10^{-21}
$(0.3, 0.3)$	0.09	0.090000000000000000008	8.00×10^{-21}	6.7477×10^{-21}
$(0.4, 0.4)$	0.16	0.16000000000000000000	8.00×10^{-21}	5.2854×10^{-21}
$(0.5, 0.5)$	0.25	0.25000000000000000001	1.00×10^{-20}	61.9554×10^{-21}
$(0.6, 0.6)$	0.36	0.35999999999999999999	1.00×10^{-20}	1.7448×10^{-20}
$(0.7, 0.7)$	0.49	0.49000000000000000000	0.00	04.3963×10^{-20}
$(0.8, 0.8)$	0.64	0.64000000000000000000	0.00	8.4713×10^{-20}
$(0.9, 0.9)$	0.84	0.81000000000000000002	2.00×10^{-20}	1.4349×10^{-19}
$(1, 1)$	1.00	1.0000000000000000001	1.00×10^{-19}	2.2476×10^{-19}

5.2. Example. Consider the following linear Volterra integro–partial differential equation.

$$u_x(x,y) + u_y(x,y) = -1 + \exp(x) + \exp(y) + \exp(x+y) + \int_0^x \int_0^y u(t,s)\ ds\ dt, \tag{5.9}$$

with the conditions $u(x,0) = \exp(x)$ and $u(0,y) = \exp(y)$. The exact solution of this problem is $u(x,y) = \exp(x+y)$. Let us consider the following approximation,

$$u_{xy}(x,y) \simeq \Phi^T(x,y)\ C. \tag{5.10}$$

Integrating (5.10) with respect to y from 0 to y, one obtains the following approximation for $u_x(x,y)$.

$$u_x(x,y) \simeq \Phi^T(x,y)\ P_y^T\ C + u_x(x,0) \simeq \Phi^T(x,y)\ P_y^T\ C + \Phi^T(x,y)\ V_1. \tag{5.11}$$

Now, integrating (5.10) with respect to x from 0 to x, one gets the following approximation for $u_y(x,y)$ as follows:

$$u_y(x,y) \simeq \Phi^T(x,y)\ P_x^T\ C + u_y(0,y) \simeq \Phi^T(x,y)\ P_x^T\ C + \Phi^T(x,y)\ V_2. \tag{5.12}$$

Also, by integrating the relation (5.11) with respect to x from 0 to x an approximation yields for $u(x,y)$ as follows:

$$\begin{aligned} u(x,y) &\simeq \Phi^T(x,y)\ P_x^T\ P_y^T\ C + \Phi^T(x,y)\ P_x^T\ V_1 + u(0,y) \\ &\simeq \Phi^T(x,y)\ P_x^T\ P_y^T\ C + \Phi^T(x,y)\ P_x^T\ V_1 + \Phi^T(x,y)\ V_1. \end{aligned} \tag{5.13}$$

The kernel is approximated as follows:

$$1 \simeq \Phi^T(x,y)\ K\ \Phi(t,s). \tag{5.14}$$

Now, the integral part in (5.9) is approximated as:

$$\begin{aligned}\int_0^x \int_0^y u(t,s)\ dtds &\simeq \int_0^y \int_0^1 \Phi^T(x,y)\ K\ \Phi(t,s)\ \Phi^T(x,y)\ \Big\{ P_x^T\ P_y^T\ C \\ &\quad + P_x^T\ V_1 + V_1 \Big\}\ dsdt \\ &\simeq \Phi^T(x,y)\ K\ \tilde{A}\ P_y\ P_x\ \Phi(x,y),\end{aligned} \tag{5.15}$$

where $\tilde{A}$ is operational matrix of product corresponding to vector $A = P_x^T\ P_y^T\ C + P_x^T\ V_1 + V_1$. Substituting the approximations (5.11)-(5.15) into (5.9), leads to the following linear algebraic equation.

$$\begin{aligned}&\Phi^T(x,y)\ P_y^T\ C + \Phi^T(x,y)\ V_1 + \Phi^T(x,y)\ P_x^T\ C + \Phi^T(x,y)\ V_2 \\ &\qquad - \Phi^T(x,y)\ K\ \tilde{A}\ P_y\ P_x\ \Phi(x,y) + 1 - e^x - e^y - e^{x+y} = 0.\end{aligned} \tag{5.16}$$

Setting $N = 7$ and using the roots of $P_8^{(\alpha,\beta)}(x)$ and $P_8^{(\alpha,\beta)}(y)$ in the x and $y-$directions, (5.16) is collocated in 64 inner tensor points for different values of parameters α and β. Hereby, the (5.16) reduces the problem to solve a system of linear algebraic equations and unknown coefficients are obtained for some values of parameters α and β. Table 3 displays the maximum absolute and estimate errors of the approximate solutions for different values of α and β. Table 4 shows different values of the exact and approximate solutions in points $(x,y) = (0.2i, 0.2i), (i = 1, 2, ..., 5)$ for $\alpha = \beta = 1$ and $N = 4, 7, 8$. It can be observed from Table 4 that the errors decrease as N increases. Also,

Table 3. Maximum absolute and estimate errors of Example 5.2 for $N = 7$ and various values of α and β

(α,β)	$Error_{Abs}$	$Error_{Est}$
$(0,0)$	1.7484×10^{-8}	8.5474×10^{-9}
$(\frac{1}{2},\frac{1}{2})$	3.1164×10^{-8}	1.4839×10^{-8}
$(1,1)$	4.6983×10^{-8}	2.0999×10^{-8}
$(-\frac{1}{4},-\frac{1}{4})$	8.2454×10^{-8}	4.8751×10^{-8}
$(\frac{1}{4},\frac{1}{4})$	2.4007×10^{-8}	1.1603×10^{-8}

Table 4. Comparison of the exact and approximate solutions of Example 5.2 for $N = 4, 7, 8$ and $\alpha = \beta = 1$

(x_i, y_i)	u_{Exact}	$Error(u_4)$	$Error(u_7)$	$u_8(x,y)$	$Error(u_8)$
$(0.2, 0.2)$	1.49182469764	2.7289×10^{-5}	1.6429×10^{-9}	1.49182469766	2.2390×10^{-11}
$(0.4, 0.4)$	2.22554092849	3.3508×10^{-5}	4.1350×10^{-10}	2.22554092853	3.5908×10^{-11}
$(0.6, 0.6)$	3.32011692274	3.2674×10^{-5}	3.5768×10^{-10}	3.32011692270	3.8762×10^{-11}
$(0.8, 0.8)$	4.95303242440	5.6798×10^{-5}	3.0065×10^{-9}	4.95303242437	2.7025×10^{-11}
$(1, 1)$	7.38905609893	5.8038×10^{-4}	4.6983×10^{-8}	7.38905609744	1.4893×10^{-9}

5.3. Example. Consider the following nonlinear system of Fredholm integro–partial differential equation.

$$\begin{cases} u(x,y) - v(x,y) + \int_0^1 u(t,y)\ v_t(t,y)\ dt = f_1(x,y), \\ v(x,y) + 3u(x,y) - \int_0^1 u_t(t,y)\ v(t,y)\ dt = f_2(x,y), \end{cases} \tag{5.17}$$

where $f_1(x,y) = x^2\ \cos(y) - y\ \sin(x) - y\ \cos(y)\ (\sin(1) - 2\ \cos(1))$ and $f_2(x,y) = y\ \sin(x) + 3\ x^2\ \cos(y) - 2\ y\ \cos(y)\ (\sin(1) - \cos(1))$ with boundary conditions $u(0,y) = 0$ and $v(0,y) = 0$. The exact solutions of this problem are $u(x,y) = x^2\ \cos(y)$ and

$v(x,y) = y\ \sin(x)$. The following approximations are used for $N=5$,

$$u_x(x,y) \simeq \Phi^T(x,y)\ C_1, \quad u(x,y) \simeq \Phi^T(x,y)\ P_x^T\ C_1 \quad v_x(x,y) \simeq \Phi^T(x,y)\ C_2,$$
$$v(x,y) \simeq \Phi^T(x,y)\ P_x^T\ C_2 \quad 1 \simeq \Phi^T(x,y)\ K\ \Phi(t,y),$$
$$u(t,y)\ v_t(t,y) \simeq A^T\ \Phi(t,y)\ \Phi^T(t,y)\ C_2 \simeq A^T\ \tilde{C_2}\ \Phi(t,y) = \Phi^T(t,y)\ \tilde{C_2}^T\ A,$$
$$u_t(t,y)\ v(t,y) \simeq C_1^T\ \Phi(t,y)\ \Phi^T(t,y)\ B \simeq C_1^T\ \tilde{B}\ \Phi(t,y) = \Phi^T(t,y)\ \tilde{B}^T\ C_1,$$
$$\int_0^1 u(t,y)\ v_t(t,y)\ dt \simeq \Phi(x,y)\ K\ E\ \tilde{C_2}^T\ A,$$
$$\int_0^1 u_t(t,y)\ v(t,y)\ dt \simeq \Phi(x,y)\ K\ E\ \tilde{B}^T\ C_1,$$

where $A = P_x^T\ C_1$, $B = P_x^T\ C_2$, $\tilde{C_2}$ and $\tilde{B}$ are the operational matrices of product corresponding to the vectors C_2 and B, and E is the following matrix:

$$E = \int_0^1 \Phi(t,y)\ \Phi^T(t,y)\ dt.$$

Note for approximating the nonlinear terms $u(t,y)\ v_t(t,y)$ and $u_t(t,y)\ v(t,y)$ has been used the Theorem 3.6. Substituting above approximations into system (5.17), leads to the following nonlinear system of algebraic equations.

$$(5.18)\quad \begin{cases} \Phi^T(x,y)\ P_x^T\ C_1 - \Phi^T(x,y)\ P_x^T\ C_2 + \Phi(x,y)\ K\ E\ \tilde{C_2}^T\ A = f_1(x,y), \\ \Phi^T(x,y)\ P_x^T\ C_2 + 3\ \Phi^T(x,y)\ P_x^T\ C_1 - \Phi(x,y)\ K\ E\ \tilde{B}^T\ C_1 = f_2(x,y), \end{cases}$$

Setting $N=5$ and using the roots of $P_6^{(\alpha,\beta)}(x)$ and $P_6^{(\alpha,\beta)}(y)$ in the x and $y-$directions, each equation of the system (5.18) is collocated in 36 inner tensor points for different values of parameters α and β. Hereby, the system (5.18) reduces the problem to solve a system of nonlinear algebraic equations and 72 unknown coefficients are obtained for some values of parameters α and β by using Newton iterative method. Table 5 shows different values of the exact and approximate solutions in points $(x,y) = (0.1i, 0.1i), (i = 1,2,...,10)$ for $\alpha = \beta = \frac{1}{2}$. Also, in Figure 1 the exact and approximate solutions are compared for the case $\alpha = \beta = 0$. Also, the absolute errors functions obtained by the proposed method are displayed in Figure 1 for $\alpha = \beta = 0$.

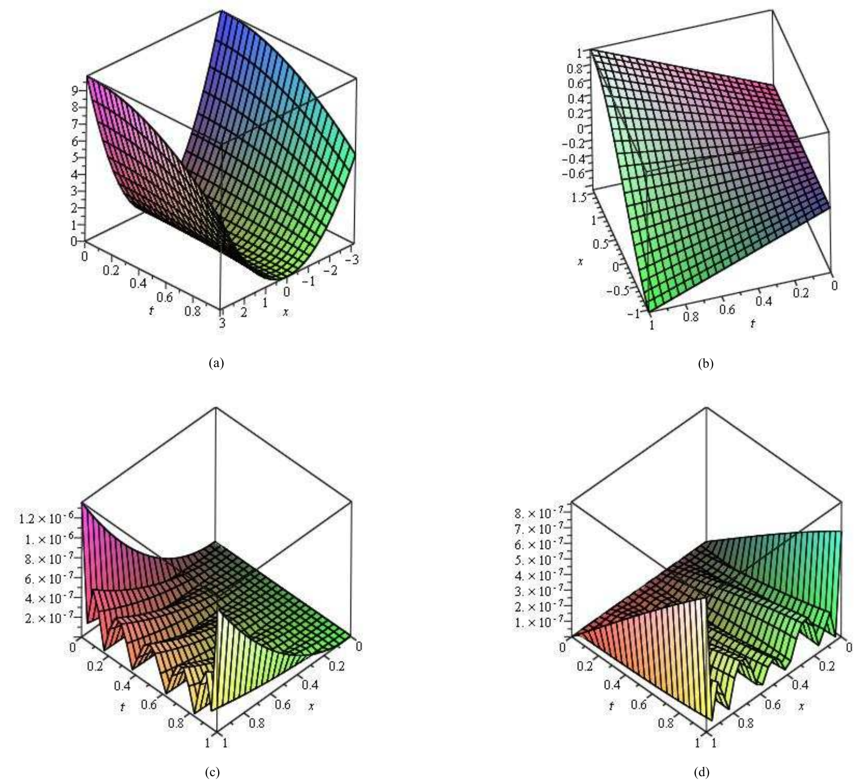
Table 5. Comparison of the exact and approximate solutions of Example 5.3 for $N=5$ and $\alpha = \beta = \frac{1}{2}$

(x_i, y_i)	u_{Exact}	$u_5(x,y)$	$Error_{Abs}$	v_{Exact}	$v_5(x,y)$	$Error_{Abs}$
$(0.1, 0.1)$	0.00995004	0.00995002	2.0419×10^{-8}	0.00998334	0.00998326e	7.8601×10^{-8}
$(0.2, 0.2)$	0.03920266e	0.03920263	2.6117×10^{-8}	0.03973387	0.03973376	1.0216×10^{-7}
$(0.3, 0.3)$	0.08598028e	0.08598027	1.2031×10^{-8}	0.08865606	0.08865595	1.14×10^{-7}
$(0.4, 0.4)$	0.14736975	0.14736970	5.6500×10^{-8}	0.15576734	0.15576711	2.2310×10^{-7}
$(0.5, 0.5)$	0.21939564	0.21939551	1.2727×10^{-7}	0.23971277	0.23971243	3.3863×10^{-7}
$(0.6, 0.6)$	0.29712082	0.29712075	7.1500×10^{-8}	0.33878548	0.33878517	3.1259×10^{-7}
$(0.7, 0.7)$	0.37477267	0.37477277	9.8920×10^{-8}	0.45095238	0.45095218	2.0102×10^{-7}
$(0.8, 0.8)$	0.44589229	0.445892296	2.2183×10^{-9}	0.57388487	0.57388456	3.1067×10^{-7}
$(0.9, 0.9)$	.50350407	0.50350365	4.1437×10^{-7}	0.70499422	0.70499360	6.1633×10^{-7}
$(1, 1)$	0.54030231	0.54030426	1.9550×10^{-6}	0.84147098	0.84147191	9.2444×10^{-7}

5.4. Example. Consider the following nonlinear system Volterra integro–partial differential equation.

$$(5.19)\begin{cases} u_y(x,y) + v(x,y) - \int_0^y \int_0^x t\ \sin(s)\ (u^2(t,s) - v^2(t,s))\ dtds = f_1(x,y), \\ u_y(x,y) + v_y(x,y) + u(x,y) - \int_0^y \int_0^x t\ \cos(s)\ (u(t,s) - v_s(t,s))\ dtds = \\ \qquad f_2(x,y), \end{cases}$$

Figure 1. Comparison of the exact and approximate solutions and their error functions for $\alpha = \beta = 0$ in Example 5.3: Plots of (a) $u_5(x, y)$, (b) $v_5(x, y)$, (c) error function of $u(x, y)$, (d) error function of $v(x, y)$



where $f_1(x, y) = \frac{1}{12}\ (1 + 2\ \cos^3(y) - 3\ \cos(y))\ x^4$ and $f_2(x, y) = x\ (2\ \cos(y) - \sin(y))$ with the conditions $u(x, 0) = x$ and $v(x, 0) = 0$. The exact solutions of this problem are $u(x, y) = x\ \cos(y)$ and $v(x, y) = x\ \sin(y)$. The following approximations are used for $N = 5$,

$$u_y(x, y) \simeq \Phi^T(x, y)\ C_1, \quad v_y(x, y) \simeq \Phi^T(x, y)\ C_2,$$

$$u(x, y) \simeq \Phi^T(x, y)\ P_y^T\ C_1 + u(x, 0) \simeq \Phi^T(x, y)\ P_y^T\ C_1 + \Phi^T(x, y)\ V,$$

$$v(x, y) \simeq \Phi^T(x, y)\ P_y^T\ C_2, \quad t\ sin(s) \simeq \Phi^T(x, y)\ K_1\ \Phi^T(t, s),$$

$$t\ cos(s) \simeq \Phi^T(x, y)\ K_2\ \Phi^T(t, s),$$

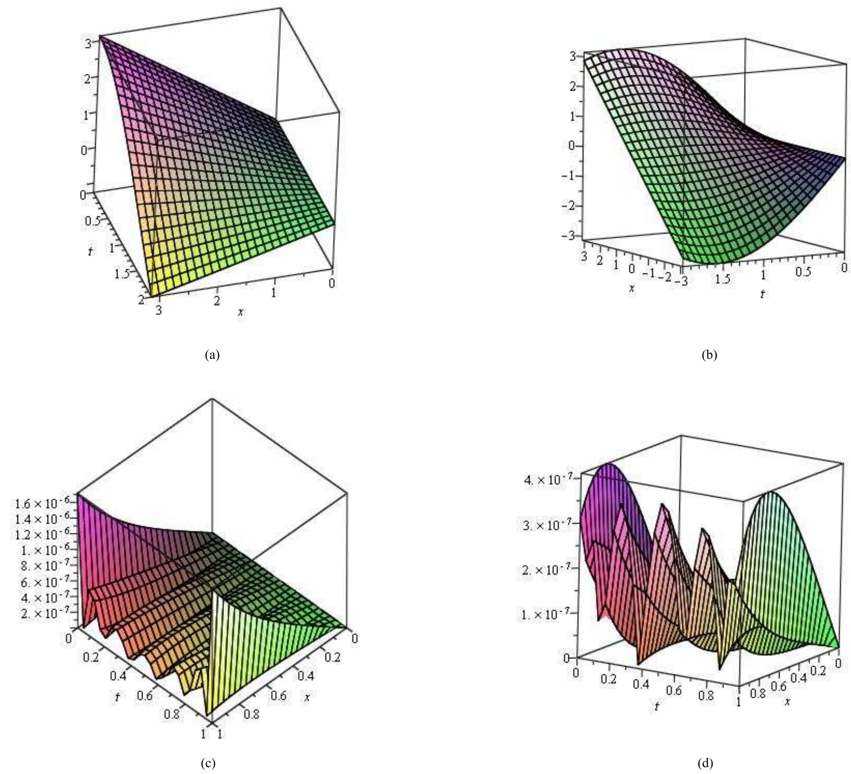
$$u^2(x, y) \simeq A_1^T\ \Phi(x, y)\ \Phi^T(x, y)\ A_1 \simeq A_1^T\ \tilde{A}_1\ \Phi(x, y) = \Phi^T(x, y)\ B_1,$$

$$v^2(x, y) \simeq A_2^T\ \Phi(x, y)\ \Phi^T(x, y)\ A_2 \simeq A_2^T\ \tilde{A}_2\ \Phi(x, y) = \Phi^T(x, y)\ B_2,$$

$$\int_0^y \int_0^x t\ \sin(s)\ (u^2(t, s) - v^2(t, s))\ dtds \simeq \Phi^T(x, y)\ K_1\ (\tilde{B}_1 - \tilde{B}_2)\ P_x\ P_y\ \Phi(x, y),$$

$$\int_0^y \int_0^x t\ \cos(s)\ (u(t, s) - v_s(t, s))\ dtds \simeq \Phi^T(x, y)\ K_2\ (\tilde{A}_1 - \tilde{C}_2)\ P_x\ P_y\ \Phi(x, y),$$

Figure 2. Comparison of the exact and approximate solutions and their error functions for $\alpha = \beta = 0$ in Example 5.4: Plots of (a) $u_5(x, y)$, (b) $v_5(x, y)$, (c) error function of $u(x, y)$, (d) error function of $v(x, y)$

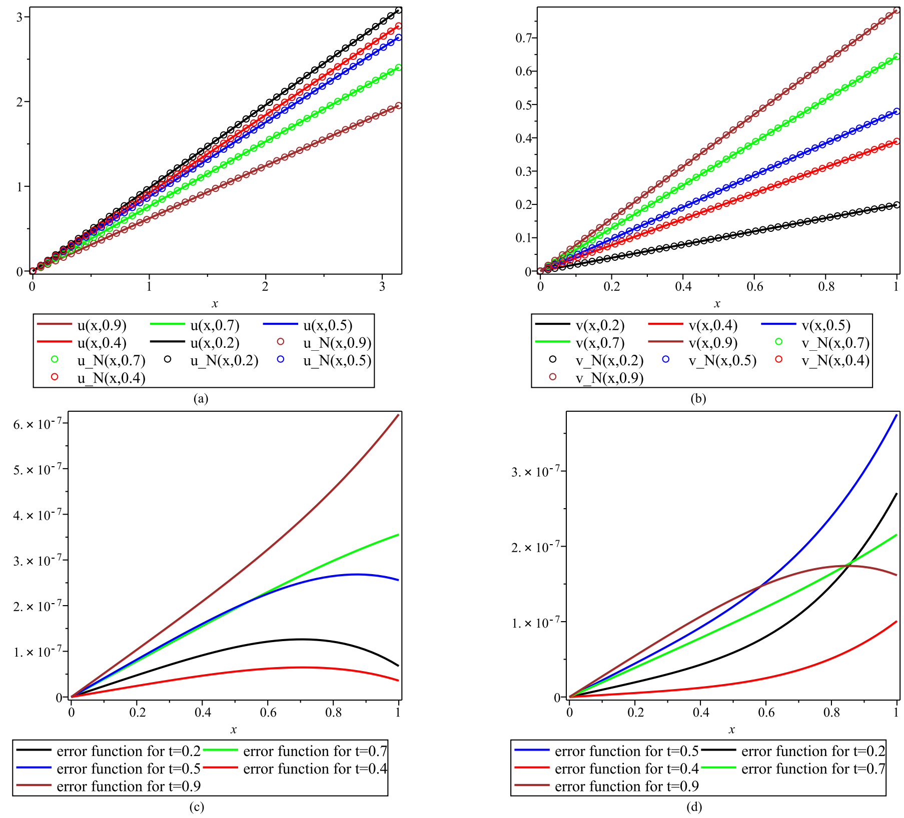


where K_1 and K_2 are known matrices, $A_1 = P_y^T C_1 + V$, $A_2 = P_y^T C_2$, $B_1 = A_1^T \tilde{A}_1$ and $B_2 = A_2^T \tilde{A}_2$. Substituting above approximations into system (5.19), leads to the following nonlinear system of the algebraic equations.

$$
(5.20) \quad \begin{cases} \Phi^T(x, y)\ C_1 + \Phi^T(x, y)\ P_y^T\ C_2 - \Phi^T(x, y)\ K_1\ (\tilde{B_1} - \tilde{B_2}) \\ \ P_x\ P_y\ \Phi(x, y) = f_1(x, y), \\ \Phi^T(x, y)\ C_1 + \Phi^T(x, y)\ C_2 + \Phi^T(x, y)\ (P_y^T\ C_1 + V) - \Phi^T(x, y)\ P_y^T\ C_2 \\ -\Phi^T(x, y)\ K_2\ (\tilde{A_1} - \tilde{C_2})\ P_x\ P_y\ \Phi(x, y) = f_2(x, y). \end{cases}
$$

Setting $N = 5$ and using the roots of $P_6^{(\alpha,\beta)}(x)$ and $P_6^{(\alpha,\beta)}(y)$ in the x and $y-$directions, each equation of the system (5.20) is collocated in 36 inner tensor points for different values of parameters α and β. Hereby, the system (5.20) reduces the problem to solve a system of nonlinear algebraic equations and 72 unknown coefficients are obtained for $\alpha = \beta = 0$ by using Newton iterative method. In Figure 2 the exact and approximate solutions are compared for the case $\alpha = \beta = 0$ and the absolute and estimate errors functions obtained by the proposed method are also displayed in Figure 2 for $\alpha = \beta = 0$. The exact and approximate solutions and their error functions are seen in Figure 3 for various values of $y = 0.2,\ 0.4,\ 0.5,\ 0.7,\ 0.9$.

Figure 3. (a) Comparison of the $u(x,y)$ and $u_6(x,y)$, (b) Comparison of the $v(x,y)$ and $v_6(x,y)$, (c) Plots of absolute error function $u(x,y)$, (d) Plots of absolute error function $v(x,y)$ for $y = 0.2,\ 0.4,\ 0.5,\ 0.7,\ 0.9$ of Example 5.4.



6. Conclusion

In this paper, a computational method based on the generalized collocation method was presented for solving some of linear and nonlinear integro-partial differential equations in terms of two variable Jacobi polynomials, by converting them to a linear or nonlinear system of algebraic equations. The illustrative examples with the satisfactory results were achieved to demonstrate the application of this method. The results indicate the proposed approach can be regarded as the simple approach and those are applicable to the numerical solution of these type of equations. It is predicted that the Jacobi collocation method will be a powerful tools for investigating approximate solutions and even analytic to linear and nonlinear functional equations. For numerical purposes the computer programmes have been written in Maple 13.

Acknowledgment

The authors are deeply grateful to the referee for his careful reading and helpful suggestions on the paper which have improved the paper.

References

[1] Bhrawy, A. H. *An efficient Jacobi pseudospectral approximation for nonlinear complex generalized Zakharov system*, Appl. Math. Comput. **247**, 30–46, 2014.

[2] Bhrawy, A. H., Abdelkawy, M. A., Zaky, M. A., and Baleanu, D. *Numerical simulation of time variable fractional order mobile-immobile advection-dispersion model*, Rom. Rep. Phys. **67** (3), 773–791, 2016.

[3] Bhrawy, A. H., Alofi, A. S., and Ezz–Eldien, S. S. *A quadrature tau method for fractional differential equations with variable coefficients*, Appl. Math. Lett. **24**, 2146–2152, 2011.

[4] Bhrawy, A. H.,Doha, E. H., Baleanu, D., and Ezz–Eldien, S. S. *A spectral tau algorithm based on Jacobi operational matrix for numerical solution of time fractional diffusion–wave equations*, J. Comput. Phys. **293**, Issue C, 142–156, 2015.

[5] Bhrawy, A. H. and Zaky, M. A. *A method based on the Jacobi tau approximation for solving multi–term time–space fractional partial differential equations*, J. Comput. Phys. **281**, 876–895, 2015.

[6] Bhrawy, A. H., Zaky, M. A., and Baleanu, D. *New Numerical Approximations for Space–Time Fractional Burgers' Equations via a Legendre Spectral–Collocation Method*, Rom. Rep. Phys. **67** (2), 340–349, 2015.

[7] Borhanifar, A. and Abazari, R. *Exact solutions for nonlinear Schrödinger equations by differential transformation method*, App. Math. Computing. **35** (1), 37–51, 2011.

[8] Borhanifar, A., Jafari, H., and Karimi, S. A. *New solitary wave solutions for the bad boussinesq and good boussinesq equations*, Numer. Meth. part. Diff. Equ. **25**, 1231–1237, 2009.

[9] Borhanifar, A. and Kabir, M. M. *New periodic and soliton solutions by application of application of exp–function method for nonlinear evolution equation*, Comput. Appl. Math. 229, 158–167, 2009.

[10] Borhanifar, A., Kabir, M. M., and Vahdat, M. *New periodic and soliton wave solutions for the generalized zakharov system and* $(2+1)$*–dimensional nizhnik-novikov-veselov system*, Chaos. Solit. Fract. **42**, 1646–1654, 2009.

[11] Borhanifar, A. and Sadri, Kh. *A new operational approach for numerical solution of generalized functional integro–differential equations*, J. Comput. Appl. Math. **279**, 80–96, 2015.

[12] Borhanifar, A. and Sadri, Kh. *Numerical solution for systems of two dimensional integral equations by using Jacobi operational collocation method*, Sohag. J. Math. **1**, 15–26, 2014.

[13] Borhanifar, A. and zamiri, A. *Application of* $(\frac{G'}{G})-$*expansion method for the Zhiber–Shabat equation and other related equations*, Math. Comput. Model. **54**, 2109–2116, 2011.

[14] Doha, E. H., Abd-Elhameed, W. M., and Youssri, Y. H. *Efficient spectral–Petrov–Galerkin methods for the integrated forms of third–and fifth–order elliptic differential equations using general parameters generalized Jacobi polynomials*, Appl. Math. Comput. **218**, 7727–7740, 2012.

[15] Doha, E. H. and Bhrawy, A. H. *An efficient direct solver for multi dimensional elliptic Robin boundary value problems using a Legendre spectral–Galerkin method*, Comput. Math. Appl. **64** (4), 558– 571, 2012.

[16] Doha, E. H., Bhrawy, A. H., Abdelkawy, M. A., and Van Gorder, R. A. *Jacobi–Gauss–Lobatto collocation method for the numerical solution of* $1+1$ *nonlinear Schrödinger equations*, J. Comput. Phys. **261**, 244–255, 2014.

[17] Doha, E. H., Bhrawy, A. H., and Ezz–Eldien, S. S. *A new Jacobi operational matrix: An application for solving fractional differential equations*, Appl. Math. Model., **36** (10), 4931–4943, 2012.

[18] Doha, E. H., Bhrawy, A. H., and Ezz-Eldien, S. S. *Efficient Chebyshev spectral methods for solving multi–term fractional orders differential equations*, Appl. Math. Model. **35**, 5662–5672, 2011.

[19] Doha, E. H., Bhrawy, A. H., and Hafez, R. M. A *Jacobi–Jacobi dual–Petrov–Galerkin method for third–and fifth–order differential equations*, Math. Com. Model. 2011; **53**, 1820–1832, 2011.

[20] Doha, E. H., Bhrawy, A. H., and Hafez, R. M. *On Shifted Jacobi Spectral Method For High–Order Multi–Point Boundary Value Problems*, Commun. Nonlin. Sci Numer. Simulat. **17** (10), 3802–3810, 2012.

[21] Guezane-Lakoud, A., Bendjazia, N., and Khaldi, R. *Galerkin method applied to telegraph integro-differential equation with a weighted integral condition*, Bound. Val. Prob. **102**, 1–12, 2013.

[22] Hosseini, S. M. and Shahmorad, S. *Numerical solution of a class of integro–differential equations by the Tau method with an error estimation*, Appl. Math. Comput. **136**, 559–570, 2003.

[23] Karimi Vanani, S. and Aminataei, A. *Tau approximate solution of fractional partial differential equations*, Comput. Math. Appl. **62**, 1075–1083, 2011.

[24] Kreyszig, E. *Introduction Functional Analysis with Applications* (USA: Wiley, 1978).

[25] Luo, M., Xu, D., Li, L., and Yang, X. *Quasi wavelet based on numerical method for Volterra integro–differential equations on unbounded spatial domains*, J. Appl. Math. Comput. **227**, 509–517, 2014.

[26] Ma, J. *Finite element methods for partial Volterra integro–differential equations on two–dimensional unbounded spatial domains*, Appl. Math. Comput. **186**, 598–609, 2007.

[27] Szegö, G. *Orthogonal Polynomials* (American Mathematical Society, Providence, Rhode Island, 1939).

[28] Tari, A., Rahimi, M. Y., Shahmorad, S., and Talati, F. *Development of the Tau Method for the Numerical Solution of Two–dimensional Linear Volterra Integro-differential Equations*, Compu. Meth. Appl. Math. **9**, 421–435, 2009.

[29] Tari, A. and Shahmorad, S. *Differential transform method for the system of two dimensional nonlinear Volterra integro–differential equations*, J. Comput. Math. Appl. **61**, 2621–2629, 2011.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 337 – 342

On conditions for univalence of some integral operators

Daniel Breaz* and Virgil Pescar†

Abstract

In this paper, we obtain new univalence conditions for the integral operators $F_{[|\delta|]}(z)$ and $G_{[|\delta|]}(z)$ of analytic functions defined in the open unit disk.

Keywords: Integral operator, univalence, integer part, modulus.

2000 AMS Classification: 30C45.

Received : 07.02.2013 *Accepted* : 02.03.2015 *Doi* : 10.15672/HJMS.20164512496

1. Introduction

Let $\mathcal{A}$ denote the class of functions of the form $f(z) = z + a_2z^2 + ...$ which are analytic in the open unit disc $\mathcal{U} = \{z : |z| < 1\}$. Further, by $\mathcal{S}$ we shall denote the class of all functions in $\mathcal{A}$ which are univalent in $\mathcal{U}$.

Pescar [7], has obtained the following univalence criteria

1.1 Theorem. [7] *Let $\gamma \in \mathbb{C}$, $f \in \mathcal{S}$, $f(z) = z + a_2z^2 + ...$*
If

$$\left| \frac{zf'(z) - f(z)}{zf(z)} \right| \leq 1, \ \forall \, z \in \mathcal{U}$$

and

$$|\gamma| \leq \frac{1}{\max\limits_{|z|\leq 1} \left[\left(1 - |z|^2\right) \cdot |z| \cdot \frac{|z|+|a_2|}{1+|a_2||z|} \right]},$$

then

$$F_\gamma(z) = \int_0^z \left(\frac{f(t)}{t} \right)^\gamma dt$$

*Department of Mathematics, "1 Decembrie 1918" University of Alba Iulia, Faculty of Science, 510009 Alba Iulia, Romania.
Email : `dbreaz@uab.ro`

†Department of Mathematics, "Transilvania" University of Braşov, Faculty of Mathematics and Computer Science, 500091 Braşov, Romania.
Email : `virgilpescar@unitbv.ro`

is in the class $\mathcal{S}$.

1.2 Theorem. **[7]** *Let* $\alpha, \beta, \gamma \in \mathbb{C}$, $f \in \mathcal{S}$, $f(z) = z + a_2 z^2 + \ldots$.
If

$$\left|\frac{zf'(z) - f(z)}{zf(z)}\right| \leq 1, \ \forall\, z \in \mathcal{U},$$

$$\mathrm{Re}\beta \geq \mathrm{Re}\alpha > 0$$

and

$$|\gamma| \leq \frac{1}{\max\limits_{|z|\leq 1}\left[\frac{1-|z|^{2\mathrm{Re}\alpha}}{\mathrm{Re}\alpha} \cdot |z| \cdot \frac{|z|+|a_2|}{1+|a_2||z|}\right]},$$

then

$$G_{\beta,\gamma}(z) = \left[\beta \int_0^z t^{\beta-1}\left(\frac{f(t)}{t}\right)^{\gamma} dt\right]^{\frac{1}{\beta}}$$

is in the class $\mathcal{S}$.

We define the next two integral operators

$$F_{[|\delta|]}(z) = \int_0^z \left(\frac{f_1(t)}{t}\right)^{\alpha_1} \cdot \ldots \cdot \left(\frac{f_{[|\delta|]}(t)}{t}\right)^{\alpha_{[|\delta|]}} dt,$$

where $\delta \in \mathbb{C}$, $|\delta| \notin [0,1)$, $\alpha_i \in \mathbb{C}$, $f_i \in \mathcal{A}$, $i = \overline{1, [|\delta|]}$, $\alpha_1 \cdot \ldots \cdot \alpha_{[|\delta|]} = \delta$ and

$$G_{[|\gamma|]}(z) = \left[\gamma \int_0^z t^{\gamma-1}\left(\frac{f_1(t)}{t}\right)^{\alpha_1} \cdot \ldots \cdot \left(\frac{f_{[|\gamma|]}(t)}{t}\right)^{\alpha_{[|\gamma|]}} dt\right]^{\frac{1}{\gamma}},$$

$\gamma \in \mathbb{C}$, $|\gamma| \notin [0,1)$, $\alpha_i \in \mathbb{C}$, $f_i \in \mathcal{A}$, $i = \overline{1, [|\gamma|]}$, $\alpha_1 \cdot \ldots \cdot \alpha_{[|\gamma|]} = \gamma$.

In this paper, we obtain new univalence conditions for the integral operators $F_{[|\delta|]}(z)$ and $G_{[|\delta|]}(z)$.

2. Preliminary results

In order to derive our main results, we have to recall here the following lemmas:

2.1 Lemma. **[2]** *If the function* f *is regular in unit disk* $\mathcal{U}$, $f(z) = z + a_2 z^2 + \ldots$ *and*

$$\left(1 - |z|^2\right)\left|\frac{zf''(z)}{f'(z)}\right| \leq 1,$$

for all $z \in \mathcal{U}$, *then the function* f *is univalent in* $\mathcal{U}$.

2.2 Lemma. **[5]** *Let* α *be a complex number,* $\mathrm{Re}\alpha > 0$ *and* $f(z) = z + a_2 z^2 + \ldots$ *be a regular function in* $\mathcal{U}$. *If*

$$\frac{1 - |z|^{2\mathrm{Re}\alpha}}{\mathrm{Re}\alpha}\left|\frac{zf''(z)}{f'(z)}\right| \leq 1,$$

for all $z \in \mathcal{U}$, *then for any complex number* β, $\mathrm{Re}\beta \geq \mathrm{Re}\alpha$, *the function*

$$F_\beta(z) = \left[\beta \int_0^z t^{\beta-1} f'(t) dt\right]^{\frac{1}{\beta}}$$

is in the class $\mathcal{S}$.

2.3 Lemma. **[3]** *If the function g is regular in $\mathcal{U}$ and $|g(z)| < 1$ in $\mathcal{U}$, then for all $\xi \in \mathcal{U}$, the following inequalities hold*

$$\left| \frac{g(\xi) - g(z)}{1 - \overline{g(z)} g(\xi)} \right| \leq \left| \frac{\xi - z}{1 - \overline{z}\xi} \right| \quad (2.1)$$

and

$$\left| g'(z) \right| \leq \frac{1 - |g(z)|^2}{1 - |z|^2},$$

the equalities hold in the case $g(z) = \epsilon \frac{z+u}{1+\overline{u}z}$, where $|\epsilon| = 1$ and $|u| < 1$.

2.4 Remark. **[3]** For $z = 0$, from inequality (2.1) we obtain for every $\xi \in \mathcal{U}$,

$$\left| \frac{g(\xi) - g(0)}{1 - \overline{g(0)} g(\xi)} \right| \leq |\xi|$$

and hence,

$$|g(\xi)| \leq \frac{|\xi| + |g(0)|}{1 + \overline{g(0)} g(\xi)}.$$

Considering $g(0) = a$ and $\xi = z$, then

$$|g(z)| \leq \frac{|z| + |a|}{1 + |a|\,|z|},$$

for all $z \in \mathcal{U}$.

3. Main results

3.1 Theorem. *Let $M > 1$, $\delta \in \mathbb{C}$, $|\delta| \notin [0,1)$, $\alpha_i \in \mathbb{C}$, for $i = \overline{1, [|\delta|]}$ and $\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]} = \delta$. If $f_i \in \mathcal{A}$, $f_i(z) = z + a_2^i z^2 + ...$, for $i = \overline{1, [|\delta|]}$ and*

$$\left| \frac{z f_i'(z) - f_i(z)}{z f_i(z)} \right| \leq 1, \ \forall i = \overline{1, [|\delta|]}, \ z \in \mathcal{U}, (3.1)$$

$$\frac{|\alpha_1| + ... + \left|\alpha_{[|\delta|]}\right|}{\left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right|} \leq M, (3.2)$$

$$\left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right| \leq \frac{1}{M \max\limits_{|z| \leq 1} \left[\left(1 - |z|^2\right) \cdot |z| \cdot \frac{|z| + |c|}{1 + |c||z|} \right]}, (3.3)$$

where

$$|c| = \frac{\left| \alpha_1 a_2^1 + ... + \alpha_{[|\delta|]} a_2^{[|\delta|]} \right|}{M \left| \alpha_1 \cdot ... \cdot \alpha_{[|\delta|]} \right|},$$

then

$$F_{[|\delta|]}(z) = \int_0^z \left(\frac{f_1(t)}{t} \right)^{\alpha_1} \cdot ... \cdot \left(\frac{f_{[|\delta|]}(t)}{t} \right)^{\alpha_{[|\delta|]}} dt$$

is in the class $\mathcal{S}$.

Proof. We have $f_i \in \mathcal{A}$, for all $i = \overline{1, [|\delta|]}$ and $\frac{f_i(z)}{z} \neq 0$, for all $i = \overline{1, [|\delta|]}$.

Let g be the function $g(z) = \left(\frac{f_1(z)}{z}\right)^{\alpha_1} \cdot ... \cdot \left(\frac{f_{[|\delta|]}(z)}{z}\right)^{\alpha_{[|\delta|]}}$, $z \in \mathcal{U}$. We have $g(0) = 1$.

Consider the function

$$h(z) = \frac{1}{M \left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right|} \cdot \frac{F''_{[|\delta|]}(z)}{F'_{[|\delta|]}(z)}, z \in \mathcal{U}.$$

The function $h(z)$ has the form:

$$h(z) = \frac{1}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right|} \sum_{i=1}^{[|\delta|]} \alpha_i \frac{zf_i'(z) - f_i(z)}{zf_i(z)}.$$

Also,

$$h(0) = \frac{1}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right|} \sum_{i=1}^{[|\delta|]} \alpha_i a_2^i.$$

By using the relations (3.1) and (3.2) we obtain that $|h(z)| < 1$ and

$$|h(0)| = \frac{\left|\alpha_1 a_2^1 + ... + \alpha_{[|\delta|]} a_2^{[|\delta|]}\right|}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right|} = |c|\,.$$

Applying Remark 2.4 for the function h we obtain

$$\frac{1}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right|} \cdot \left|\frac{F''_{[|\delta|]}(z)}{F'_{[|\delta|]}(z)}\right| \le \frac{|z| + |c|}{1 + |c|\,|z|}, \forall z \in \mathcal{U}$$

and

$$\left|\left(1 - |z|^2\right) \cdot z \cdot \frac{F''_{[|\delta|]}(z)}{F'_{[|\delta|]}(z)}\right| \le M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right| \left(1 - |z|^2\right) \cdot |z| \cdot \frac{|z| + |c|}{1 + |c|\,|z|}, \forall z \in \mathcal{U}. (3.4)$$

Consider the function $H : [0,\ 1] \to \mathbb{R}$ defined by

$$H(x) = (1 - x^2)x\frac{x + |c|}{1 + |c|\,x};\ x = |z|\,.$$

We have

$$H\left(\frac{1}{2}\right) = \frac{3}{8} \cdot \frac{1 + 2\,|c|}{2 + |c|} > 0 \Rightarrow \max_{x \in [0,1]} H(x) > 0.$$

Using this result and from (3.4) we have:

$$\left|\left(1 - |z|^2\right) \cdot z \cdot \frac{F''_{[|\delta|]}(z)}{F'_{[|\delta|]}(z)}\right| \le M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\delta|]}\right| \cdot \max_{|z|<1}\left[\left(1 - |z|^2\right) \cdot |z| \cdot \frac{|z| + |c|}{1 + |c|\,|z|}\right], \forall z \in \mathcal{U}. (3.5)$$

Applying the condition (3.3) in the form (3.5) we obtain that

$$\left(1 - |z|^2\right) \cdot \left|z \cdot \frac{F''_{[|\delta|]}(z)}{F'_{[|\delta|]}(z)}\right| \le 1, \forall z \in \mathcal{U},$$

and from Lemma 2.1 we obtain that $F_{[|\delta|]} \in \mathcal{S}$.

3.2 Theorem. *Let $M > 1$, γ, $\delta \in \mathbb{C}$, $|\gamma| \notin [0,1)$, $\alpha_i \in \mathbb{C}$, for $i = \overline{1, [|\gamma|]}$, $\alpha_1 \cdot ... \cdot \alpha_n = [|\gamma|]$. If $f_i \in \mathcal{A}$, $f_i(z) = z + a_2^i z^2 + ...$, for $i = \overline{1, [|\gamma|]}$ and*

$$\left|\frac{zf_i'(z) - f_i(z)}{zf_i(z)}\right| \le 1,\ \forall i = \overline{1, [|\gamma|]},\ z \in \mathcal{U}, (3.6)$$

$$\frac{|\alpha_1| + ... + \left|\alpha_{[|\gamma|]}\right|}{\left|\alpha_1 \cdot ... \cdot \alpha_{[|\gamma|]}\right|} \leq M, (3.7)$$

$$\mathrm{Re}\gamma \geq \mathrm{Re}\delta > 0,$$

$$\left|\alpha_1 \cdot ... \cdot \alpha_{[|\gamma|]}\right| \leq \frac{1}{M \max\limits_{|z|\leq 1}\left[\left(1-|z|^2\right)\cdot |z| \cdot \frac{|z|+|c|}{1+|c||z|}\right]}, (3.8)$$

where

$$|c| = \frac{\left|\alpha_1 a_2^1 + ... + \alpha_{[|\gamma|]} a_2^{[|\gamma|]}\right|}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\gamma|]}\right|},$$

then

$$G_{[|\gamma|]}(z) = \left[\gamma \int_0^z t^{\gamma-1}\left(\frac{f_1(t)}{t}\right)^{\alpha_1} \cdot ... \cdot \left(\frac{f_{[|\gamma|]}(t)}{t}\right)^{\alpha_{[|\gamma|]}} dt\right]^{\frac{1}{\gamma}}$$

is in the class $\mathcal{S}$.

Proof. We consider the function

$$h(z) = \int_0^z \left(\frac{f_1(t)}{t}\right)^{\alpha_1} \cdot ... \cdot \left(\frac{f_{[|\gamma|]}(t)}{t}\right)^{\alpha_{[|\gamma|]}} dt.$$

Define the function

$$p(z) = \frac{1}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\gamma|]}\right|} \cdot \frac{h''(z)}{h'(z)},\ z \in \mathcal{U}.$$

The function $p(z)$ has the form:

$$p(z) = \frac{1}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\gamma|]}\right|} \sum_{i=1}^{[|\gamma|]} \alpha_i \frac{zf_i'(z) - f_i(z)}{zf_i(z)}.$$

By using the relations (3.6) and (3.7) we obtain $|p(z)| < 1$ and

$$|p(0)| = \frac{\left|\alpha_1 a_2^1 + ... + \alpha_{[|\gamma|]} a_2^{[|\gamma|]}\right|}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\gamma|]}\right|} = |c|.$$

Applying Remark 2.4 for the function h we obtain

$$\frac{1}{M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\gamma|]}\right|} \cdot \left|\frac{h''(z)}{h'(z)}\right| \leq \frac{|z| + |c|}{1 + |c|\,|z|}, \forall z \in \mathcal{U}$$

and

$$\left|\frac{1 - |z|^{2\mathrm{Re}\delta}}{\mathrm{Re}\delta} \cdot z \cdot \frac{h''(z)}{h'(z)}\right| \leq M\left|\alpha_1 \cdot ... \cdot \alpha_{[|\gamma|]}\right| \frac{1 - |z|^{2\mathrm{Re}\delta}}{\mathrm{Re}\delta} \cdot |z| \cdot \frac{|z| + |c|}{1 + |c|\,|z|}, \forall z \in \mathcal{U}. (3.9)$$

Consider the function $Q : [0,\ 1] \to \mathbb{R}$ defined by

$$Q(x) = \frac{1 - x^{2\mathrm{Re}\delta}}{\mathrm{Re}\delta} \cdot x \cdot \frac{x + |c|}{1 + |c|\,x};\ x = |z|.$$

We have $Q\left(\frac{1}{2}\right) > 0 \Rightarrow \max\limits_{x\in[0,1]} Q(x) > 0.$

Using this result in (3.9), we have:

$$\frac{1-|z|^{2\mathrm{Re}\delta}}{\mathrm{Re}\delta}\left|\frac{zh''(z)}{h'(z)}\right| \leq M\left|\alpha_1\cdot ...\cdot\alpha_{[|\gamma|]}\right|\cdot\max_{|z|<1}\left[\frac{1-|z|^{2\mathrm{Re}\delta}}{\mathrm{Re}\delta}\cdot|z|\cdot\frac{|z|+|c|}{1+|c|\,|z|}\right], \forall z\in\mathcal{U}. (3.10)$$

Applying the condition (3.8) in the relation (3.10), we obtain that

$$\frac{1-|z|^{2\mathrm{Re}\delta}}{\mathrm{Re}\delta}\left|\frac{zh''(z)}{h'(z)}\right| \leq 1, \forall z\in\mathcal{U}$$

and from Lemma 2.2, we obtain that $G_{[|\gamma|]}\in\mathcal{S}$.

References

[1] D. Breaz, N. Breaz, *Two integral operators*, Studia Universitatis "Babeş-Bolyai", Mathematica, Cluj-Napoca, No.3, 2002, 13-21.

[2] J. Becker, *Lownersche Differntialgleichung und quasikonform fortsetzbare schlichte Functionen*, J. Reine Angew. Math. 255(1972), 23-43.

[3] G. M. Goluzin, *Geometriceskaia teoria funktii kompleksnogo peremenogo*, Moscova, 1966.

[4] Y.J. Kim, E.P. Merkes, *On an integral of powers of a spirallike function*, Kyungpook Math., J., 12 (1972), 2, 249-253.

[5] N.N. Pascu, *An improvement of Becker's Univalence Criterion*, Proceedings of the Commemorative Session Simion Stoilow, University of Braşov, 1987, 43-48.

[6] N.N. Pascu, V.Pescar, *On the integral operators of Kim-Merkes and Pfaltzgraff*, Mathematica, Univ. "Babeş-Bolyai", Cluj-Napoca, 32 (55), 2(1990), 185-192.

[7] V. Pescar, *On some integral operations which preserve the univalence*, Journal of Mathematics, Vol. XXX(1997), pp. 1-10, Punjab University.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 343 – 353

Semiprime and weakly compressible modules

N. Dehghani * and M. R. Vedadi †

Abstract

An R-module M is called semiprime (resp. weakly compressible) if it is cogenerated by each of its essential submodules (resp. $\mathrm{Hom}_R(M, N)N$ is nonzero for every $0 \neq N \leq M_R$). We carry out a study of weakly compressible (semiprime) modules and show that there exist semiprime modules which are not weakly compressible. Weakly compressible modules with enough critical submodules are characterized in different ways. For certain rings R, including prime hereditary Noetherian rings, it is proved that M_R is weakly compressible (resp. semiprime) if and only if $M \in \mathrm{Cog}(\mathrm{Soc}(M) \oplus R)$ and $M/\mathrm{Soc}(M) \in \mathrm{Cog}(R)$ (resp. $M \in \mathrm{Cog}(\mathrm{Soc}(M) \oplus R)$). These considerations settle two questions, namely Qu 1, and Qu 2, in [6, p 92].

Keywords: Krull dimension, semiprime module, singular semi-Artinian ring, weakly compressible module.

2000 AMS Classification: Primary 16N60; 16P60. Secondary 16D90, 16D80

Received : 21.12.2014 *Accepted* : 15.03.2015 *Doi* : 10.15672/HJMS.20164512503

1. Introduction

Throughout this paper rings will have a nonzero identity, modules will be right and unitary. In [2], a module M_R is called *prime* if $\mathrm{Hom}_R(M, K)N \neq 0$ for all nonzero submodules $K, N \leq M_R$ and it is shown that M_R is prime if and only if it is cogenerated by each of its nonzero submodules. A semiprime notion for modules is then obtained in [4] by setting $K = N$ in the above definition of prime modules. These semiprime modules are precisely weakly compressible modules in the sense of [1]; see for example Theorem 2.5 below. Following [6], a module M_R is called *weakly compressible* if $\mathrm{Hom}_R(M, N)N \neq 0$ for all nonzero $N \leq M_R$. We also call M_R *semiprime* if every essential submodule of M_R cogenerates M_R. In this paper, prime module means the prime module in the sense of

*Department of Mathematics, Isfahan University of Technology, Isfahan, Iran Email: n.dehghani@math.iut.ac.ir
†Department of Mathematics, Isfahan University of Technology, Isfahan, Iran, Corresponding Author. Email: mrvedadi@cc.iut.ac.ir

[2]; see [11, Sections 13, 14] for an excellent reference on the subject. Weakly compressible modules have applied in different situations. For example, in the study of weakly semisimple modules [12] and modules which have semiprime right Goldie endomorphism rings [3, Theorem 2.6]. They have also been appeared in the Cohen-Fishman's question about the semiprimeness of the smash product $A\#H$ when H is a semisimple Hopf algebra and A is a semiprime H-module algebra. In [6, Corollary 7.6] for certain semisimple Hopf algebra H, it is shown that $A\#H$ is a semiprime ring if and only if the $A\#H$-module A is weakly compressible.

In the present work, we carry out a study of weakly compressible (semiprime) modules and show that there are semiprime modules which are not weakly compressible (Examples and Remarks 2.8). Weakly compressible modules with enough critical submodules are characterized in different ways (Theorems 3.4 and 3.7). For certain rings R, including prime hereditary Noetherian rings, it is shown that M_R is weakly compressible (resp. semiprime) if and only if $M \in \mathrm{Cog}(\mathrm{Soc}(M) \oplus R)$ and $M/\mathrm{Soc}(M) \in \mathrm{Cog}(R)$ (resp. $M \in \mathrm{Cog}(\mathrm{Soc}(M) \oplus R)$). Furthermore, if R is a PID then M_R is weakly compressible if and only if $M/\mathrm{Soc}(M) \in \mathrm{Cog}(R)$ (Corollary 4.6). These considerations settle two questions, namely Qu 1, and qu 2, in [6, p 92] where it is asked whether there exists a weakly compressible module M which is not a subdirect product of prime modules. Such a module M cannot satisfy the conditions of Theorem 3.4 or 3.7 or 4.1 because of Remark 4.2. Any unexplained terminology and all the basic results on rings and modules that are used in the sequel can be found in [5] and [7].

2. General properties of weakly compressible modules

In this section, we investigate weakly compressible (semiprime) modules over any ring and show that semiprime modules are not necessarily weakly compressible. We give a characterization of weakly compressible modules and using this we state our main results in the next sections. Let M be an R-module and N be a submodule of M_R. We say that *M is N-weakly compressible* if for each nonzero submodule K of N, there exists an R-homomorphism $f : M \to K$ such that $f(K) \neq 0$. Thus M_R is weakly compressible if and only if M is M-weakly compressible if and only if M is N-weakly compressible for any $0 \neq N \leq M_R$. We use the notation $N \leq_{ess} M$ to denote N is an essential submodule of M. Also, if X and Y are R-modules, then $\cap\{\ker f \mid f : X_R \to Y_R\}$ is denoted by $\mathrm{Rej}(X, Y)$. The module X is cogenerated by Y (write $X \in \mathrm{Cog}(Y)$) if $\mathrm{Rej}(X, Y) = 0$. In the following, some properties of weakly compressible (semiprime) modules are collected.

2.1. Lemma. (a) *Let M be a semiprime R-module. If N is either an essential or fully invariant submodule of M_R, then N is a semiprime R-module.*

(b) *The class of weakly compressible modules is closed under co-products and taking submodules.*

(c) *The class of semiprime modules is closed under products and co-products.*

(d) *Let Λ be a non-empty set. Then M_R is semiprime if and only if $M_R^{(\Lambda)}$ is so.*

(e) *Every weakly compressible module is semiprime.*

(f) *Let M be a nonzero R-module and M_1, M_2 be submodules of M_R such that there is no nonzero R-module X which embeds in M_1 and M_2. Then M is $(M_1 \oplus M_2)$-weakly compressible if and only if M is M_i-weakly compressible for $i = 1, 2$.*

(g) *If M_R is semiprime then $ann_R(M)$ is a semiprime ideal of R.*

(h) *M_R is weakly compressible (resp. semiprime) if and only if $M_{R/I}$ is weakly compressible (resp. semiprime) where $MI = 0$ and $I \rhd R$.*

(i) *If M_R is weakly compressible and N is a fully invariant closed submodule of M_R, then M/N is a weakly compressible R-module.*

Proof. (a) If $N \leq_{ess} M_R$, then it is easy to see that N_R is semiprime. Let N be a fully invariant of M_R and $K \leq_{ess} N$. There exists a submodule L of M_R such that $N \cap L = 0$ and $N \oplus L \leq_{ess} M$. Thus $K \oplus L \leq_{ess} M$. By our assumption $M \in \mathrm{Cog}(K \oplus L)$. Hence there exists an injective homomorphism $\theta : M \to K^I \oplus L^I$ for some set I. Since N is fully invariant of M_R, it is easy to see $\pi\theta(N) = 0$, where $\pi : K^I \oplus L^I \to L^I$ is the natural projection. It follows that $\theta(N)$, and hence N embeds in K^I, proving that N_R is semiprime.
(b) We only prove the co-product case. Let $\{M_i\}_{i\in I}$ be a family of weakly compressible R-modules and N be any nonzero submodule of $\oplus_{i\in I} M_i$. It is easy to verify that there exists subset J of I such that the canonical projection $\pi : \oplus_{i\in I} M_i \to \oplus_{j\in J} M_j =: W$ is one to one on N and $\pi(N) \cap M_j \neq 0$ for each $j \in J$; see also [9, Lemma 2.1]. Because M_j is weakly compressible for each $j \in J$, there are homomorphisms $f_j \in \mathrm{Hom}_R(M_j, \pi(N) \cap M_j)$ such that $f_j(\pi(N) \cap M_j) \neq 0$. Now let $f = \sum_{j\in J} f_j : W \to \pi(N)$ and $\theta = \pi^{-1} f \pi$. Then $\theta : M \to N$ such that $\theta|_N \neq 0$, as desired.
(c) Let N be an essential submodule of product $\prod_{i\in I} M_i$ where each M_i is a semiprime module (the co-product case has a similar proof). Note that for each $i \in I$ we have $(N \cap M_i) \leq_{ess} M_i$. Thus by our assumption, $M_i \in \mathrm{Cog}\ (N)$ for each $i \in I$. It follows that $\prod_{i\in I} M_i \in \mathrm{Cog}(N)$.
(d) The necessity follows by part (c). Conversely, let $M^{(\Lambda)}$ be semiprime and $N \leq_{ess} M$. Then $N^{(\Lambda)} \leq_{ess} M^{(\Lambda)}$. Thus $M^{(\Lambda)} \in \mathrm{Cog}(N^{(\Lambda)})$. This shows that $M \in \mathrm{Cog}(N)$, as desired.
(e) This is obtained by [6, Theorem 5.1(b)].
(f) Just note that if N is a nonzero submodule of $M_1 \oplus M_2$, then by our assumption, either $N \cap M_1 \neq 0$ or $N \cap M_2 \neq 0$.
(g) This follows by [6, Proposition 5.5(viii)].
(h) This has a routine argument.
(i) Let N be a fully invariant closed submodule of M_R. By [5, Proposition 6.32], there exists $K \leq M_R$ such that N is a complement to K in M. It follows that $K \oplus N/N$ is an essential submodule of M/N. Hence, it is enough to show that M/N is $(K \oplus N/N)$-weakly compressible. Now let $(x + N) \in (K \oplus N/N)$ for some nonzero element $x \in K$. Since M_R is weakly compressible, there exists a homomorphism $f : M \to xR$ such that $f(x) \neq 0$. We have $f(N) = 0$ because N is a fully invariant submodule of M. Thus f induces a homomorphism $\bar{f} : M/N \to xR \oplus N/N$ such that $\bar{f}(x + N) \neq 0$. The proof is complete. □

An R-module M is called *torsionless* if it is cogenerated by R. The following result may be already in the literature, but we cannot spot it, we give a proof for the sake of the reader.

2.2. Proposition. *Every torsionless module over a semiprime ring is weakly compressible.*

Proof. Let R be a semiprime ring and M be an R-submodule of R^I for some set I. Suppose that N is a nonzero submodule of M. Thus $\pi_i(N) \neq 0$ for some $i \in I$, where π_i is the canonical projection from R^I to R. Since R is a semiprime ring, $(\pi_i(N))^2 \neq 0$. Hence there exists $x \in N$ such that $x\pi_i(N) \neq 0$. Now let $f = \iota_x \pi$ where $\pi_i|_M = \pi$ and $\iota_x : R \to xR$ is left multiplication by x. Then $f : M \to N$ is a homomorphism such that $f(N) \neq 0$, proving that M_R is weakly compressible. □

2.3. Corollary. *Let R be a ring and $\{I_i\}_{i\in A}$ be a family of semiprime ideals in R. Then $\oplus_{i\in A}(R/I_i)^{\Lambda_i}$ is a weakly compressible R-module, where each Λ_i is a set.*

Proof. This follows by Proposition 2.2 and Lemma 2.1(b),(h). □

2.4. Lemma. *Every nonsingular R-module M contains an essential submodule isomorphic to $\oplus_i I_i$ where each I_i is a right ideal of R.*

Proof. Let x be any nonzero element of M_R. Then $\mathrm{ann}_R(x)$ is not an essential right ideal of R by our assumption on M_R. Thus there exists a nonzero right ideal I_x of R such that $\mathrm{ann}_R(x) \cap I_x = 0$. Note that $I_x \simeq xI_x$. Therefore every nonzero submodule of M contains a nonzero submodule that is isomorphic to a right ideal of R. Now suppose that $\Omega = \{N \leq M_R |$ there is $I \leq R_R$ such that $I \simeq N\}$. If $\{N_\lambda\}_{\lambda\in\Lambda}$ is a maximal independent family of submodules in Ω, then by what we have already proved, $\oplus_{\lambda\in\Lambda} N_\lambda$ is an essential submodule of M_R. □

In [6, Theorem 5.1], it is shown that an R-module M is weakly compressible if and only if $\mathrm{Hom}_R(M, N)^2 \neq 0$ for all nonzero $N \leq M$ if and only if $N\cap$ $\mathrm{Rej}(M, N) = 0$ for any nonzero $N \leq M_R$. In the following we give more equivalent conditions for a nonzero module M to be weakly compressible. We should note that in [1], a module M_R is called "weakly compressible" if for every $0 \neq N \leq M_R$ there exists $f \in \mathrm{Hom}_R(M, N)$ with $f^2 \neq 0$. Such a module M is clearly weakly compressible (in the sense of [6]), but we have been unable to find in the literature a proof to show the converse is true. A proof of this is given below for completeness. Recall that for any R-module M the set $\{m \in M \mid \mathrm{ann}_R(m) \leq_{ess} R_R\}$ is denoted by $\mathrm{Z}(M)$.

2.5. Theorem. *The following conditions are equivalent for a nonzero R-module M.*
(a) *M_R is weakly compressible.*
(b) *For every nonzero $N \leq M$, there exists $f \in Hom_R(M, N)$ such that $f^2 \neq 0$.*
(c) *$N \not\hookrightarrow Rej(M, N)$, for every nonzero $N \leq M_R$.*
(d) *$M_1 \not\hookrightarrow Rej(M, M_2)$ for all nonzero isomorphic R-modules M_1 and M_2.*
(e) *There exists an essential submodule N of M_R such that M is N-weakly compressible.*
(f) *There exists submodule N of M_R such that M is N-weakly compressible and M/N is weakly compressible.*
(g) *There exists a semiprime ideal I of R such that $MI = 0$ and M is $Rej(M, R/I)$-weakly compressible.*
(h) *M is $Z(M)$-weakly compressible and $M/Z_2(M) \in Cog(R/I)$ for some semiprime ideal $I \subseteq ann_R(M)$.*

Proof. (a) ⇒ (b). Let N be a nonzero submodule of M_R and for every $f \in \mathrm{Hom}_R(M, N)$, $f^2 = 0$. It is easy to verify that $fg = -gf$ for all $f, g \in \mathrm{Hom}_R(M, N)$ (note that $(f + g)^2 = 0$). By (a), there exist $f \in \mathrm{Hom}_R(M, N)$ and $g \in \mathrm{Hom}_R(M, f(M))$ such that $f(N) \neq 0$ and $g(f(M)) \neq 0$. Since $gf = -fg$, we have $fg \neq 0$. If follows that $f^2(M) \neq 0$ because $g(M) \subseteq f(M)$. This contradicts our assumption.
(b) ⇒ (c). Let N be any nonzero submodule of M_R. Suppose that there exists an injective homomorphism $\theta : N \to \mathrm{Rej}(M, N)$. Since $N \simeq \theta(N)$, $\mathrm{Rej}(M, \theta(N)) = \mathrm{Rej}(M, N)$. Hence, if $f \in \mathrm{Hom}_R(M, \theta(N))$ then $\mathrm{Im} f \subseteq \mathrm{Rej}(M, \theta(N))$. This shows that $f^2 = 0$ for every $f \in \mathrm{Hom}_R(M, \theta(N))$. This contradicts (b).
(c) ⇒ (d). Just note that if $M_1 \hookrightarrow \mathrm{Rej}(M, M_2)$, then M_1 is isomorphic to a submodule N of M such that $N \hookrightarrow \mathrm{Rej}(M, N)$.
(d) ⇒ (a), (a) ⇔ (e) and (a) ⇒ (f) are clear.
(a)⇒ (g). This is hold because $\mathrm{ann}_R(M)$ is a semiprime ideal of R by Lemma 2.1.
(g) ⇒ (f). Let $N = \mathrm{Rej}(M, R/I)$. Then $M/N \in \mathrm{Cog}(R/I)$. Now apply Proposition 2.2 and Lemma 2.1(h).
(f) ⇒ (a). Suppose (f) holds and K is a nonzero submodule of M_R. We shall show that there exists $g \in \mathrm{Hom}_R(M, K)$ with $g(K) \neq 0$. Now if $K \cap N \neq 0$, then we are done by our assumption on N. If $K \cap N = 0$, then consider the submodule $(N \oplus K)/N$ of M/N. Since M/N is weakly compressible, we can deduce such g exists.

(a) $\Rightarrow$ (h). First note that for any R-module M, we have $M/\mathrm{Z}_2(M)$ is a nonsingular R-module. Let $I = \mathrm{ann}_R(M)$. Thus $M/\mathrm{Z}_2(M) \in \mathrm{Cog}(R/I)$ by Lemmas 2.1(i) and 2.4, the proof is complete.
(h) $\Rightarrow$ (f). Since $\mathrm{Z}(M) \leq_{ess} \mathrm{Z}_2(M)$, it is clear that M_R is also $\mathrm{Z}_2(M)$-weakly compressible. The result is now obtained by Proposition 2.2. □

2.6. Corollary. (a) *If R is a right self injective ring, then M_R is weakly compressible if and only if M_R is $Z(M)$-weakly compressible.*
(b) *If R is a right V-ring (i.e., simple R-modules are injective) and $M/Soc(M)$ is a weakly compressible R-module, then M_R is weakly compressible.*

Proof. (a) Let R be a right self injective ring. For the sufficiency, let N be complement to $\mathrm{Z}(M)$ in M_R. By Theorem 2.5(e), we shall show that M is $\mathrm{Z}(M)\oplus N$-weakly compressible. Since R is right self injective, every nonsingular cyclic R-module is isomorphic to a direct summand of R_R and hence it is an injective R-module. It follows that M is N-weakly compressible. The proof is now completed by Lemma 2.1(f). The converse is clear.
(b) By Theorem 2.5(f). □

2.7. Proposition. *The following statements hold for an extending module M_R.*
(a) *M_R is weakly compressible if and only if $Z_2(M)$ and $M/Z_2(M)$ are weakly compressible R-modules.*
(b) *If $Soc(R_R) \leq_{ess} R_R$, then M_R is semiprime if and only if $Z_2(M)$ and $M/Z_2(M)$ are semiprime R-modules.*

Proof. Let $N = \mathrm{Z}_2(M)$. Since M is extending, it is known that $M \simeq N \oplus M/N$.
(a) Apply Theorem 2.5(f) and note that N is weakly compressible if and only if M is N-weakly compressible.
(b) Since $\mathrm{Soc}(R_R) \leq_{ess} R_R$, it is easy to verify that $\mathrm{Z}(V^\Lambda) = (\mathrm{Z(V)})^\Lambda$ for any R-module V and any set Λ. Now let M_R be semiprime. By Lemma 2.1(a), N_R is semiprime. Suppose that $K/N \leq_{ess} M/N$. Then $K \leq_{ess} M$ and so there exists an injective homomorphism $\theta : M \to K^\Lambda$. Define $\alpha : M/N \to K^\Lambda/N^\Lambda$ by $\alpha(m + N) = \theta(m) + N^\Lambda$. Clearly α is a homomorphism. If $\alpha(m + N) = 0$ then $\theta(m) = \{n_\lambda\}_{\lambda\in\Lambda} \in N^\Lambda$. For each λ, we have $n_\lambda J_\lambda \subseteq \mathrm{Z}(M)$ where $J_\lambda \leq_{ess} R_R$. Thus $\theta(mJ) \subseteq (\mathrm{Z}(M))^\Lambda$ where $\cap_\lambda J_\lambda = J$. By our assumption on R, $J \leq_{ess} R_R$ and $\theta(mJ) \subseteq \mathrm{Z}(K^\Lambda)$. It follows that $mJ \subseteq \mathrm{Z}(M)$ because θ is one to one. Hence $m \in N$, proving that α is injective and so M/N is weakly compressible. □

For every module M_R the intersection of all maximal submodule of M is denoted by $\mathrm{Rad}(M)$. If M does not have maximal submodules, we put $\mathrm{Rad}(M) = M$.

2.8. Examples and Remarks. (a) There are modules N such that $\mathrm{Rad}(N) = 0$ but N is not semiprime. Let P be the set of all prime integer numbers and $p \in P$. Consider the $\mathbb{Z}$-module $N = \{m/p^n \mid m, n \in \mathbb{Z}, n \geq 1\}$. Then for each $q \in P \setminus \{p\}$, qN is a maximal submodule of $N_\mathbb{Z}$. To see this, note that $qN \neq N$ and suppose that K is any submodule of $N_\mathbb{Z}$ such that $qN \subsetneq K$ and $m/p^t \in K \setminus qN$. Hence $(m, q) = 1$. Also, if $a/p^r \in K$ for some $r \geq 1$ and $(a, q) = 1$, then $1/p^r \in K$. It follows that $1/p^n \in K$ for all $n \geq 1$ (take $n \geq t$ or $n \leq t$). Therefore $K = N$ and so qN is a maximal submodule. Clearly $\bigcap_{q\neq p} qN = 0$ and hence Rad $(N) = 0$. Now if $N_\mathbb{Z}$ is semiprime, then $\mathrm{Hom}_\mathbb{Z}(N, \mathbb{Z}) \neq 0$ and since N is uniform, we must have $N \hookrightarrow \mathbb{Z}$, contradiction.

(b) A direct summand of a semiprime module is not necessarily a semiprime module. Assume that P and N are as stated in (a). Let $W = \oplus_{p\in P}\mathbb{Z}_p$ and $L = W \oplus N$. We show that $L_\mathbb{Z}$ is semiprime. Since $\bigcap_{q\neq p} qN = 0$, $N \in \mathrm{Cog}(W)$. Thus from $\mathrm{Soc}(L) = W$, we have $L \in Cog(\mathrm{Soc}(L))$. It follows that L is semiprime as a $\mathbb{Z}$-module because every essential submodule of L contains $\mathrm{Soc}(L)$.

(c) Lemma 2.1(b) and part (b) show that the $\mathbb{Z}$-module L in (b) is semiprime which is not weakly compressible. Furthermore, let R be a commutative regular ring which is not semi-Artinian (for example $R = \prod \mathbb{Z}_2$). Since R is a regular ring, $\mathrm{Rad}(M) = 0$ for all R-modules. Hence every R-module embeds in a semiprime R-module by Lemma 2.1(c). On the other hand, since R is not semi-Artinian, there exists an R-module M which is not weakly compressible by [10, Corollary 3.5]. Now if M embeds in a semiprime R-module L, then L is not weakly compressible by Lemma 2.1(b).

(d) The condition (h) in Theorem 2.5 shows that the study of weakly compressible modules can be reduced to the study of such modules when they are either singular or nonsingular; see Proposition 2.7. However we shall note that, in general, the condition M is $\mathrm{Z}(M)$-weakly compressible is stronger than $\mathrm{Z}(M)$ is a weakly compressible R-module. For example, if $R = \begin{bmatrix} \mathbb{Z} & 0 \\ \mathbb{Z}_2 & \mathbb{Z}_2 \end{bmatrix}$, then $Z(R_R) = \begin{bmatrix} 0 & 0 \\ \mathbb{Z}_2 & 0 \end{bmatrix} =: I$. Thus I_R is weakly compressible, but R is not I-weakly compressible because $\mathrm{Hom}_R(R, I)(I) = 0$.

(e) In view of the condition (c) in Theorem 2.5, we note that the condition $N \hookrightarrow Rej(M, N)$ is weaker than $N \subseteq Rej(M, N)$. For if we consider I as left ideal in R, then $I \simeq \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{Z}_2 \end{bmatrix} =: J$ as left R-modules and $\mathrm{Rej}(R, J) = \mathrm{l.ann}_R(J) = \begin{bmatrix} \mathbb{Z} & 0 \\ \mathbb{Z}_2 & 0 \end{bmatrix}$. Hence $J \hookrightarrow \mathrm{Rej}(R, J)$, but $J \not\subseteq \mathrm{Rej}(R, J)$.

In the following nonsingular weakly compressible modules are characterized and some corollaries are given. For certain module M_R, the condition (c) of Theorem 2.5 is reduced to the ideals of R; see below.

2.9. Proposition. *Let M be a module over a semiprime ring R and $Z(Rej(M, R)) = 0$. Then the following statements are equivalent.*
(a) *M_R is weakly compressible.*
(b) *For all nonzero right ideal I of R, $I \not\hookrightarrow Rej(M, I)$.*
(c) *$M \in Cog(R)$.*

Proof. (a) $\Rightarrow$ (b). By Theorem 2.5(c).
(b) $\Rightarrow$ (c). If $\mathrm{Rej}(M, R)$ is nonzero then by Lemma 2.4, $I \hookrightarrow \mathrm{Rej}(M, R)$ for some nonzero right ideal I of R. It follows that $I \hookrightarrow \mathrm{Rej}(M, I)$, a contradiction. Therefore $\mathrm{Rej}(M, R) = 0$ and so (c) holds.
(c) $\Rightarrow$ (a). By Proposition 2.2. □

2.10. Corollary. *Let M be a nonsingular R-module. Then M_R is weakly compressible if and only if there exists a semiprime ideal $I \subseteq ann_R(M)$ such that $M \in Cog(R/I)$.*

Proof. Note that $\mathrm{Z}(M_{R/I}) \subseteq \mathrm{Z}(M_R)$, for any ideal I of R. The result is now obtained by Proposition 2.9. □

A ring R is called *right (left) duo* ring if every right(left) ideal of R is two sided.

2.11. Corollary. *Let M be a faithful module over a right(left) duo ring R. Then M_R is weakly compressible if and only if M_R is $Z(M)$-weakly compressible, $M/Z(M) \in Cog(R)$ and R is a semiprime ring.*

Proof. It is easy to verify that every semiprime right(left) duo ring must be reduced and hence it is a nonsingular ring [5, Lemma 7.8]. Thus $\mathrm{Z}(M) = \mathrm{Z}_2(M)$. Suppose now M is weakly compressible, then R must be a semiprime ring because M_R is faithful. Also $M/\mathrm{Z}(M)$ is weakly compressible by Lemma 2.1(i), and so $M/\mathrm{Z}(M) \in \mathrm{Cog}(R)$ by Proposition 2.9. The converse is obtained by Theorem 2.5(h). □

3. Weakly compressible modules with enough critical submodules

We are now going to investigate semiprime and weakly compressible modules over rings with Krull dimensions. Let M be an R-module. Following [7, Chapter 6], the *Krull dimension* of M_R, will be denoted by $\mathrm{K.dim}(M)$. Modules with Krull dimensions are known to have *finite uniform dimensions* [7, Lemma 6.2.6]. Let $\alpha \geq 0$ be an ordinal number. A module M_R is called α-*critical* if $\mathrm{K.dim}(M) = \alpha$ and $\mathrm{K.dim}(M/N) < \alpha$ for every nonzero submodule N of M_R. A module is then called *critical* if it is β-critical for some ordinal number β. The submodule $\bigcap\{\ K \leq M_R \mid M/K$ is α-critical $\}$ is denoted by $\mathrm{J}_\alpha(M)$.

3.1. Lemma. *Let M be a semiprime R-module, T be any nonzero submodule of M. If there exist submodules W and N of M such that $N \in Cog(T)$, $T \notin Cog(W)$ and $(N \oplus W) \leq_{ess} M$, then $T \nsubseteq Rej(M,T)$.*

Proof. Since M_R is semiprime, there exists an injective homomorphism $f : M \to N^A \oplus W^A$ for some set A. Since $T \notin \mathrm{Cog}(W)$, $\pi f(T) \neq 0$, where $\pi : N^A \oplus W^A \to N^A$ is natural projection. By our assumption, $N^A \in \mathrm{Cog}(T)$. Hence there exists a homomorphism $\varphi : N^A \to T$ such that $\varphi\pi f(T) \neq 0$, proving that $T \nsubseteq \mathrm{Rej}(M,T)$. □

The following lemma is needed. That is just obtained by the definition of critical submodules.

3.2. Lemma. *Let U and V be critical R-modules and $f : U \to V$ be a nonzero homomorphism. Then either $Kerf = 0$ or $K.dim(V) < K.dim(U)$.*

We say that a module M_R *has enough critical submodules* if every nonzero submodule has a nonzero submodule with Krull dimension (note, modules with Krull dimension have critical submodules).

3.3. Lemma. *Suppose that M_R has enough critical submodules and $\alpha = Min\{K.dim(N) \mid 0 \neq N \leq M_R\}$. If M_R is semiprime, then $N \nsubseteq Rej(M,N)$ for every submodule $N \leq M_R$ with $K.dim(N) = \alpha$.*

Proof. Let $N \leq M_R$ and $\mathrm{K.dim}(N) = \alpha$. By [7, Lemma 6.2.10], there exists a critical submodule $T \leq N$. By choosing of α, T is α-critical. Let $\Lambda = \{T^{'} \leq M_R \mid T^{'} \in Cog(T)\}$, $\{T_i^{'}\}_{i\in I}$ be a maximal independent family of elements in Λ and $N' = \oplus_{i\in I} T_i^{'}$. Since M_R has enough critical submodules, $N' \oplus W \leq_{ess} M_R$ where W is a direct sum of critical submodules. Therefore by Lemma 3.2, $T \notin Cog(W)$ and so $T \nsubseteq \mathrm{Rej}(M,T)$ by Lemma 3.1. The proof is complete. □

A module M_R is called *compressible* if it embeds in every submodule of M. By Lemma 3.2 critical weakly compressible modules are compressible.

3.4. Theorem. *Suppose that M_R has enough critical submodules and $\beta = Sup\{K.dim(N) \mid N$ is a critical submodule of $M_R\}$. Then the following statements are equivalent.*
(a) *M_R is semiprime module and $J_\beta(M) = 0$.*
(b) *M_R embeds in a product of β-critical compressible submodules of M_R.*
(c) *M_R embeds in a product of β-critical compressible R-modules.*
Furthermore, each of the above conditions implies that M_R is weakly compressible.

Proof. (a) $\Rightarrow$ (b). We first show that every critical submodule of M_R is β-critical. Let C be any critical submodule of M_R. By our assumption, $C \nsubseteq \mathrm{J}_\beta(M)$. It follows that there exists a homomorphism f from M_R to a β-critical module T_R such that $f(C) \neq 0$. By Lemma 3.2, f is one to one on C. Thus C_R is β-critical, as desired. Now since M_R has enough critical submodules, $\beta = \mathrm{Min}\{\mathrm{K.dim}(N) \mid 0 \neq N \leq M_R\}$. Therefore M_R is weakly compressible by Lemma 3.3. Hence, every critical submodule of M_R is also weakly compressible as well as compressible. The proof is now complete because M contains an essential submodule that is a direct sum of β-critical compressible submodules.
(b) $\Rightarrow$ (c). This is clear.
(c) $\Rightarrow$ (a). It is easy to see that $\mathrm{J}_\beta(M) = 0$. As we see in the proof of (a) $\Rightarrow$ (b), for every critical submodule C of M_R there exist a β-critical compressible R-module T and homomorphism $\alpha : M \to T$ such that α is one to one on C. Since now T_R is compressible, there exists an injective homomorphism $f : T \to C$. Thus $f\alpha(C) \neq 0$, proving that $C \nsubseteq \mathrm{Rej}(M, C)$. It follows that M_R is weakly compressible, hence semiprime. $\square$

3.5. Remark. Let $R = \mathbb{Z}$, $M = \mathbb{Z}_2 \oplus \mathbb{Z}$ and β be as stated in Theorem 3.4. Then M_R is weakly compressible and $\beta = 1$, but $\mathrm{J}_\beta(M) \neq 0$ because $M \notin \mathrm{Cog}(R)$.

3.6. Lemma. *Suppose that M is an R-module, $\{V_i\}_{i\in I}$ is a family of nonzero submodules of M_R, $\{W_j\}_{j\in J}$ is a family of R-modules and the following conditions (a), (b) hold,*
(a) *For every nonzero submodule N of M_R, there exists $V_i \subseteq N$ for some $i \in I$.*
(b) *For every $i \in I$, there exist $j \in J$ and homomorphism $f : M \to W_j$ such that $Kerf \cap V_i = 0$.*
If M_R has finite uniform dimension, then there exists a finite subset A of J such that M_R embeds in $\oplus_{j\in A} W_j$.

Proof. Let $\Lambda = \{\mathrm{u.dim}(\mathrm{Ker}f) \mid f \in \mathrm{Hom}_R(M, \oplus_{j\in A} W_j)$ and A is a finite set $\}$. By hypothesis Λ is a nonempty set. Let n be the smallest element in Λ, and $f : M \to \oplus_{j\in A} W_j$ such that $\mathrm{u.dim}(\mathrm{Ker}f) = n$. Let $K = \mathrm{Ker}f$. If $K \neq 0$, then by (a), there exists $i \in I$ such that $V_i \subseteq K$ and by (b) there exists a homomorphism $g : M \to W_t$ such that $\mathrm{Ker}g \cap V_i = 0$ for some $t \in J$. Now, define $h : M \to \oplus_{j\in A} W_j \oplus W_t$ by $h(m) = (f(m), g(m))$ for all $m \in M$. It is clear that $\mathrm{Ker}h = \mathrm{K} \cap \mathrm{Ker}g$. Since $\mathrm{Ker}h \cap V_i = 0$, $\mathrm{Ker}h$ is not essential submodule of K. Hence $\mathrm{u.dim}(\ \mathrm{Ker}h) < \mathrm{u.dim}(\ \mathrm{Ker}f)$. This contradicts the choice of f. Therefore $K = 0$ and so M embeds in $\oplus_{j\in A} W_j$, as desired. $\square$

3.7. Theorem. *The following statements are equivalent for a nonzero module M_R.*
(a) *M_R is weakly compressible with finite uniform dimension and $Z(M)$ has Krull dimension.*
(b) *M_R is weakly compressible with finite uniform dimension and $Z(M)$ has enough critical submodules.*
(c) *M_R embeds in a finite direct sum $\oplus_i W_i$ of cyclic compressible submodules of M_R such that each W_i is either uniform nonsingular or critical singular.*
(d) *M_R embeds in $W \oplus V$ such that W_R and V_R are weakly compressible, W is nonsingular with finite uniform dimension and V is singular with Krull dimension.*

Proof. (a) $\Rightarrow$ (b) and (c) $\Rightarrow$ (d) are clear. (d) $\Rightarrow$ (a) is obtained by Lemma 2.1(b) and the fact that modules with Krull dimensions have finite uniform dimensions. We shall show that (b) $\Rightarrow$ (c).
Apply Lemma 3.6 for $\{V_i\}_{ß\in I} = \{W_j\}_{æ\in J=I} = \{C \leq M_R \mid C$ is either uniform nonsingular or critical singular$\}$. By our hypothesis, the condition (a) of Lemma 3.6 holds. Note that every endomorphism of the above submodules C is either injective or zero (Lemma 3.2). Hence, by the weakly compressible condition on M, we have the submodules C

are compressible and the condition (b) of Lemma 3.6 holds. The proof is now complete because any compressible module embeds in each of its cyclic submodule. □

3.8. Corollary. *The following statements are equivalent for a nonzero module M_R.*
(a) *M_R is weakly compressible with Krull dimension.*
(b) *M_R is weakly compressible with finite uniform dimension and it has enough critical submodules.*
(c) *M_R embeds in a finite direct sum of critical compressible submodules of M_R.*
(d) *M_R embeds in a finite direct sum of critical compressible R-modules.*

Proof. This follows by Theorem 3.7. □

The following result is a consequence of Theorem 3.7 which should be compared with Corollary 2.10.

3.9. Corollary. *If M_R is a nonsingular weakly compressible module with finite uniform dimension, then M_R embeds in a finitely generated free R-module.*

Proof. Note that every nonsingular compressible R-module embeds in R (Lemma 2.4). Thus the result is obtained by Theorem 3.7(c). □

4. Weakly compressible modules over singular semi-Artinian rings

In [8, Main Theorem], it is shown that a $\mathbb{Z}$-module M is weakly compressible if and only if $\mathrm{Z}(M)$ is semisimple and $M/\mathrm{Z}(M)$ is torsionless. We conclude the paper with a characterization of weakly compressible (semiprime) modules over certain rings including prime hereditary Noetherian rings. If R is a hereditary Noetherian ring, then by [7, Proposition 5.4.5], *every nonzero singular R-module has a nonzero socle.* We call such rings R *right singular semi-Artinian.*

4.1. Theorem. *Suppose that R is a right singular semi-Artinian ring, M_R is nonzero and $MI = 0$ for some ideal I of R. If M_R is semiprime then $M \in Cog(Soc(M) \oplus R/I)$. The converse holds if I is a prime ideal of R.*

Proof. Since R/I is also a right singular semi-Artinian ring, we can suppose that $I = 0$. Let M_R be semiprime and $\mathrm{Soc}(\mathrm{Z}(M)) \oplus K \leq_{ess} M_R$ where $K \leq M_R$. By our assumption on R, we have $\mathrm{Z}(K) = 0$. Thus $M \in \mathrm{Cog}(\mathrm{Soc}(M) \oplus R)$ by Lemma 2.4.
Conversely, assume that $M \in \mathrm{Cog}(\mathrm{Soc}(M) \oplus R)$ and R is a prime ring. Let N be any essential submodule of M_R. We have $\mathrm{Soc}(\mathrm{Z}(N)) \oplus L \leq_{ess} N$ such that $L \simeq \oplus_{i \in I} I_i$ where each I_i is a right ideal of R. Since $\mathrm{Soc}(M)$ lies in any essential submodule of M_R, we deduce from the hypothesis that $M \in \mathrm{Cog}(\mathrm{Soc}(N) \oplus L \oplus R)$. Now $\mathrm{Rej}(R, L) = \mathrm{ann}_R(L) = 0$ because R is prime ring. It follows that $R \in \mathrm{Cog}(L)$ and hence $M \in \mathrm{Cog}(N)$, proving that M_R is semiprime. □

4.2. Remark. Let R be any ring and M be a nonzero R-module. Then M_R is a subdirect product of prime modules if and only if M is cogenerated by prime modules. Now let M_R be a weakly compressible R-module and $A = \mathrm{ann}_R(M)$. Note that R/A is subdirect product of prime R-modules. Therefore if M_R satisfies the conditions of Theorem 3.4 or 3.7 or 4.1, then M is cogenerated by prime modules and hence it is a subdirect product of prime R-modules. This gives a partially answer to the open problem 1 of [6].

4.3. Proposition. *Let M_R be semiprime and $L \leq M_R$. Then the following statements hold.*
(a) *If $Soc(L)$ is finitely generated then $Soc(L)$ is a direct summand of M. In particular, if M has acc on direct summands, then $Soc(M)$ is a direct summand of M.*
(b) *If every cyclic submodules of L has a finitely generated socle then $Soc(L)$ is a closed*

submodule of M.
(c) *M is Soc(M)-weakly compressible and Soc$(M)\cap$ Rad$(M) = 0$.*

Proof. (a) Suppose that the length of $\mathrm{Soc}(L) = n$. Let T_1 be a simple submodule of L_R, $N = \sum\{T' \leq M_R \mid T' \simeq T_1\}$ and W be a complement to N in M. Then $T_1 \notin \mathrm{Cog}(W)$ and so by Lemma 3.1, $T_1 \nsubseteq Rej(M, T_1)$. Hence there exists a nonzero homomorphism $f : M \to T_1$ such that $f(T_1) \neq 0$. Clearly $\mathrm{Ker}(f)$ is a maximal submodule of M_R. It follows that $M = T_1 \oplus A_1$ where $A_1 \ker f$, and hence $L = T_1 \oplus (L \cap A_1)$. If $\mathrm{Soc}(L \cap A_1) = 0$, then $\mathrm{Soc}(L) = T_1$ and we are done. If not, consider the simple submodule T_2 of $L \cap A_1$. Again we deduce that T_2 is a direct summand of M and hence of A_1. Thus $M = T_1 \oplus T_2 \oplus A_2$ for some $A_2 \leq M$ and $L = T_1 \oplus T_2 \oplus (L \cap A_2)$. Continue to obtain $T_1 \oplus T_2 \oplus ... \oplus T_n$ is a direct summand of M_R, as desired. The last statement is now clear.
(b) If $\mathrm{Soc}(L) \leq_{ess} C$ and $x \in C \leq L$, then $\mathrm{Soc}(xR) \leq_{ess} xR$. Hence by our assumption and (a), we must have $xR = \mathrm{Soc}(xR) \subseteq \mathrm{Soc}(L)$. It follows that $\mathrm{Soc}(L) = C$.
(c) This is obtained by (a) and the fact that every nonzero cyclic submodule in $\mathrm{Rad}(M)$ is a small submodule of M and so cannot be a direct summand. □

4.4. Lemma. *Suppose S and R are two rings, $T = R \oplus S$ and M be a T-module. Then $M = K \oplus L$ where K and L are modules over R and S respectively. In this case, $Z(M_T) = Z(K_R) \oplus Z(L_S)$ and $Soc(M_T) = Soc(K_R) \oplus Soc(L_S)$.*

Proof. Just note that if M is a T-module then $M = Me_1 \oplus Me_2$ where $e_1 = 1_R$ and $e_2 = 1_S$ are central orthogonal idempotents in T such that $e_1 S = e_2 R = 0$. Clearly Me_1 and Me_2 are naturally R-module and S-module respectively. □

4.5. Theorem. *Suppose that M is a nonzero R-module and $MI = 0$ for some semiprime ideal I of R. If $M \in Cog(Soc(M) \oplus R/I)$ and $M/Soc(M) \in Cog(R/I)$, then M_R is weakly compressible. The converse holds if R is a right singular semi-Artinian ring such that every cyclic R-module has a finitely generated socle or acc on direct summands.*

Proof. We may suppose that $I = 0$. Let $N = \mathrm{Soc}(M)$. By Proposition 2.2, M/N is a weakly compressible R-module. Hence by Theorem 2.5(f), we need to show that M is N-weakly compressible. Assume that S is a simple submodule of M, and by hypothesis let $\theta : M \hookrightarrow (N \oplus R)^{\Lambda} =: L$ for some set Λ. Then $\pi_\lambda \theta(S)$ is nonzero for some canonical projection π_λ $(\lambda \in \Lambda)$ on L. Let $U = \pi_\lambda \theta(S)$ and note that $U \simeq S$. Since now any minimal right ideal in a semiprime ring R is a direct summand of R_R, we deduce that there exists R-homomorphism $f : M \to U$ such that $f(S) \neq 0$. It follows $S \nsubseteq \mathrm{Rej}(M, S)$, as desired.
Conversely, suppose that R satisfies the above hypothesis and M_R is weakly compressible. By Theorem 4.1, $M \in \mathrm{Cog}(\mathrm{Soc}(M) \oplus R)$. It remains to show that $M/\mathrm{Soc}(M) \in \mathrm{Cog}(R)$. Since R is assumed to be a semiprime ring, $\mathrm{Soc}(R_R)$ is a direct summand of R by Proposition 4.3(a). It follows that $R \simeq A \oplus B$ where A is a semisimple ring and B is a ring with zero socle. By Lemma 4.4, $M = K \oplus L$ and $\mathrm{Soc}(M) = K \oplus \mathrm{Soc}(L)$. Thus it is enough to show that $L/\mathrm{Soc}(L) \in \mathrm{Cog}(B)$. Now L is a weakly compressible B-module. Since B is a right singular semi-Artinian ring, $\mathrm{Z}(L)$ has an essential socle, and since $\mathrm{Soc}(B_B) = 0$, every simple B-module is singular. Thus $\mathrm{Soc}(L) \leq_{ess} \mathrm{Z}(L)$. On the other hand, if C is a cyclic submodule L_B, then an application of Proposition 4.3(a) for C_B shows that $\mathrm{Soc}(C)$ is a direct summand of C. Hence $\mathrm{Soc}(C)$ is cyclic. It follows that $\mathrm{Soc}(L)$ is a closed submodule of L by Proposition 4.3(b). Therefore $\mathrm{Soc}(L) = \mathrm{Z}(L) = \mathrm{Z}_2(L)$. The proof is now completed by Lemma 2.1(i). □

4.6. Corollary. *Let R be a prime right singular semi-Artinian ring such that cyclic R-modules have finite uniform dimensions. Then the following statements hold for M_R.*
(a) *$M \in Cog(Soc(M) \oplus R)$ and $M/Soc(M) \in Cog(R)$ if and only if M_R is weakly compressible.*
(b) *$M \in Cog(Soc(M) \oplus R)$ if and only if M_R is semiprime.*
(c) *If M_R is semiprime, then either M_R is semisimple or $Z(M) = Soc(M)$.*
(d) *Furthermore, if R is a PID then $M/Soc(M) \in Cog(R)$ if and only if M_R is weakly compressible.*

Proof. (a) and (b). These follow from Theorems 4.1 and 4.5.
(c). By Proposition 4.3(a), $\mathrm{Soc}(R_R)$ is a direct summand of R. Since now R is a prime ring, R is semisimple or $\mathrm{Soc}(R_R) = 0$. If R is a semisimple ring then M_R is semisimple. In case $\mathrm{Soc}(R_R) = 0$, as we see in the proof of Theorem 4.5, $\mathrm{Z}(M) = \mathrm{Soc}(M)$.
(d) The sufficiency holds by part (a). Conversely, let $N = \mathrm{Soc}(M)$ and $M/N \in \mathrm{Cog}(R)$. It follows that $\mathrm{Z}(M) \subseteq N$ and M/N is weakly compressible. Thus we need to show that M is N-weakly compressible. Let S be a simple submodule of M_R and $P = \mathrm{ann}_R(S)$. Let $P = pR$ for some prime element $p \in R$. If $0 \neq x \in S \subseteq MP$, then $x = mp$ for some $m \in M$ and so $p^2R \subseteq \mathrm{ann}_R(m) \subsetneq pR$. Hence $p^2R = \mathrm{ann}_R(m)$. This implies that $mR \subseteq N$, a contradiction. Therefore, $S \cap MP = 0$. Since now $M/MP \simeq S^{(\Lambda)}$ for some set Λ, we can deduce that $S \nsubseteq \mathrm{Rej}(M, S)$. The proof is complete. □

References

[1] Avraamova, O. D. *A generalized density theorem*, in Abelian Groups and Modules, Tomsk Univ. Publ. **8**, 3-16, 1989.

[2] Bican, L., Jambor, P., Kepka, T. and N¡emec, P. *Prime and coprime modules*, Fund. Math. **57**, 33-45, 1980.

[3] Haghany, A. and Vedadi, M. R. *Study of semi-projective retractable modules*, Algebra Colloquium **14** (3), 489-496, 2007.

[4] Jirásko, J. *Notes on generalized prime and coprime modules*, I, Comm. Math. Univ. Carolinae **22** (3), 467-482, 1981.

[5] Lam, T. Y. *Lectures on Modules and Rings*, Grad. Texts in Math. **139**, Springer, New York, 1998.

[6] Lomp, C. *Prime elements in partially ordered groupoids applied to modules and Hopf algebra actions*, J. Algebra Appl. **4** (1) 77-97, 2005.

[7] McConnell, J. C. and Robson, J. C. *Non-commutative Noetherian Rings*, Wiley Interscience, New York, 1987.

[8] Samsonova, I. V. *Weakly compressible abelian groups*, Comm. Moscow. Math. Soc. 187-188, 1993.

[9] Smith, P. F. and Vedadi, M. R. *Submodules of direct sums of compressible modules*, Comm. Algebra. **36**, 3042-3049, 2008

[10] Toloei, Y. and Vedadi, M. R. *On rings whose modules have nonzero homomorphisms to nonzero submodules*, Publ. Mat. **57** (1), 107-122, 2012.

[11] Wisbauer, R. *Modules and Algebras: Bimodule Structure and Group Action on Algebras*, Pitman Monograhs 81, Addison-Wesley Longman, 1996.

[12] Zelmanowitz, J. *Weakly semisimple modules and Density theory*, Comm. Algebra **21** (5), 1785-1808, 1993.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 355 – 365

On the rescaled Riemannian metric of Cheeger-Gromoll type on the cotangent bundle

Aydin Gezer*† and Murat Altunbas‡

Abstract

Let (M, g) be an $n-$dimensional Riemannian manifold and T^*M be its cotangent bundle equipped with a Riemannian metric of Cheeger-Gromoll type which rescale the horizontal part by a positive differentiable function. The main purpose of the present paper is to discuss curvature properties of T^*M and construct almost paracomplex Norden structures on T^*M. We investigate conditions for these structures to be para-Kähler (paraholomorphic) and quasi-para-Kähler. Also, some properties of almost paracomplex Norden structures in context of almost product Riemannian manifolds are presented.

Keywords: Almost paracomplex structure, connection, cotangent bundle, paraholomorphic tensor field, Riemannian metric.

2000 AMS Classification: 53C07, 53C55, 55R10.

Received : 30.01.2014 *Accepted* : 21.01.2015 *Doi* : 10.15672/HJMS.20164515849

1. Introduction

Geometric structures on bundles have been object of much study since the middle of the last century. The natural lifts of the metric g, from a Riemannian manifold (M, g) to its tangent or cotangent bundles, induce new (pseudo) Riemannian structures, with interesting geometric properties. Maybe the best known Riemannian metric ${}^Sg$ on the tangent bundle over Riemannian manifold (M, g) is that introduced by Sasaki in 1958 (see [25]), but in most cases the study of some geometric properties of the tangent bundle endowed with this metric led to the flatness of the base manifold. The metric ${}^Sg$ is called the Sasaki metric. The Sasaki metric ${}^Sg$ has been extensively studied by several authors and in many different contexts. Another Riemannian metric on the tangent bundle TM defined by E. Musso and F. Tricerri [14] is the Cheeger-Gromoll metric ${}^{CG}g$. The metric

*Ataturk University, Faculty of Science, Department of Mathematics, 25240 Erzurum-TURKEY Email: agezer@atauni.edu.tr

†Corresponding Author.

‡Erzincan University, Faculty of Science and Art, Department of Mathematics, 24030 Erzincan-TURKEY Email: maltunbas@erzincan.edu.tr

was defined by J. Cheeger and D. Gromoll [3]; yet, E. Musso and F. Tricerri wrote down its expression, constructed it in a more "comprehensible" way, and gave it the name. In [30], B. V. Zayatuev introduced a Riemannian metric ${}^S\overline{g}$ on the tangent bundle TM given by

$$\begin{aligned} {}^S\overline{g}\left({}^HX, {}^HY\right) &= fg\left(X, Y\right), \\ {}^S\overline{g}\left({}^HX, {}^VY\right) &= {}^S\overline{g}\left({}^VX, {}^HY\right) = 0, \\ {}^S\overline{g}\left({}^VX, {}^VY\right) &= g\left(X, Y\right), \end{aligned}$$

where $f > 0$, $f \in C^\infty(M)$. For $f = 1$, it follows that ${}^S\overline{g} = {}^Sg$. The metric ${}^S\overline{g}$ is called the rescaled Sasaki metric. The authors studied the rescaled Sasaki type metric on the cotangent bundle T^*M over Riemannian manifold (M, g) (see [8]). Also, for rescaled Riemannian metrics on orthonormal frame bundles, see [11].

Let M_{2k} be a $2k$-dimensional differentiable manifold endowed with an almost (para) complex structure φ and a pseudo-Riemannian metric g of signature (k, k) such that $g(\varphi X, Y) = g(X, \varphi Y)$ for arbitrary vector fields X and Y on M_{2k}, i.e. g is pure with respect to φ. The metric g is called Norden metric. Norden metrics are also referred to as anti-Hermitian metrics or B-metrics. They present extensive application in mathematics as well as in theoretical physics. Many authors considered almost (para)complex Norden structures on the tangent, cotangent and tensor bundles [5, 7, 16, 17, 18, 19, 20, 22, 23].

In this paper, firstly, we present curvature tensor of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$. Secondly, we get the conditions under which the cotangent bundle endowed with some paracomplex structures and the rescaled Riemannian metric of Cheeger-Gromoll type ${}^{CG}g_f$ is a paraholomorphic Norden manifold. Finally, for an almost paracomplex manifold to be an specialized almost product manifold, we give some results related to Riemannian almost product structures on the cotangent bundle.

Throughout this paper, all manifolds, tensor fields and connections are always assumed to be differentiable of class C^∞. Also, we denote by $\Im^p_q(M)$ the set of all tensor fields of type (p, q) on M, and by $\Im^p_q(T^*M)$ the corresponding set on the cotangent bundle T^*M. The Einstein summation convention is used, the range of the indices i, j, s being always $\{1, 2, ..., n\}$.

2. Preliminaries

The cotangent bundle of a smooth $n-$dimensional Riemannian manifold may be endowed with a structure of $2n-$dimensional smooth manifold, induced by the structure on the base manifold. If (M, g) is a smooth Riemannian manifold of dimension n, we denote its cotangent bundle by $\pi : T^*M \to M$. A system of local coordinates $\left(U, x^i\right)$, $i = 1, ..., n$ in M induces on T^*M a system of local coordinates $\left(\pi^{-1}\left(U\right), \ x^i, \ x^{\overline{i}} = p_i\right)$, $\overline{i} = n+i = n+1, ..., 2n$, where $x^{\overline{i}} = p_i$ is the components of covectors p in each cotangent space T^*_xM, $x \in U$ with respect to the natural coframe $\left\{dx^i\right\}$.

Let $X = X^i \frac{\partial}{\partial x^i}$ and $\omega = \omega_i dx^i$ be the local expressions in U of a vector field X and a covector (1-form) field ω on M, respectively. Then the vertical lift ${}^V\omega$ of ω and the horizontal lift ${}^HX$ of X are given, with respect to the induced coordinates, by

$$(2.1) \qquad {}^V\omega = \omega_i \partial_{\overline{i}},$$

and

$$(2.2) \qquad {}^HX = X^i\partial_i + p_h\Gamma^h_{ij}X^j\partial_{\overline{i}},$$

where $\partial_i = \frac{\partial}{\partial x^i}$, $\partial_{\overline{i}} = \frac{\partial}{\partial x^{\overline{i}}}$ and Γ_{ij}^h are the coefficients of the Levi-Civita connection ∇ of g.

Let T^*M be the cotangent bundle of a Riemannian manifold (M,g). If the local expression of the metric g is $g = g_{ij}dx^i \otimes dx^j$, then the inverse of the metric g is $g^{-1} = g^{ij}\partial_i \otimes \partial_j$, where g^{ij} are the entries of the inverse matrix of g_{ij}, i.e. $g^{ij}g_{jk} = \delta_k^i$. We define $r^2 = g^{-1}(p,p) = g^{ij}p_ip_j$ and put $\alpha = 1 + r^2$. Then the rescaled Riemannian metric of Cheeger-Gromoll type ${}^{CG}g_f$ is defined on T^*M by the following three equations at $(x,p) \in T^*M$

$$ {}^{CG}g_f\left({}^V\omega, {}^V\theta\right) = \frac{1}{\alpha}(g^{-1}(\omega,\theta) + g^{-1}(\omega,p)\,g^{-1}(\theta,p)), \tag{2.3} $$

$$ {}^{CG}g_f\left({}^V\omega, {}^HY\right) = 0, \tag{2.4} $$

$$ {}^{CG}g_f\left({}^HX, {}^HY\right) = fg(X,Y) \tag{2.5} $$

for any $X, Y \in \Im_0^1(T^*M)$ and $\omega, \theta \in \Im_1^0(T^*M)$, where $f > 0$, $f \in C^\infty(M)$, $g^{-1}(\omega,\theta) = g^{ij}\omega_i\theta_j$.

The Lie bracket operation of vertical and horizontal vector fields on T^*M is given by the formulas

$$ \begin{cases} \left[{}^HX, {}^HY\right] = {}^H[X,Y] + {}^V(p \circ R(X,Y)) \\ \left[{}^HX, {}^V\omega\right] = {}^V(\nabla_X\omega) \\ \left[{}^V\theta, {}^V\omega\right] = 0 \end{cases} \tag{2.6} $$

for any X, $Y \in \Im_0^1(M)$ and θ, $\omega \in \Im_1^0(M)$, where R is the Riemannian curvature of g defined by $R(X,Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]}$ (for details, see [28], p. 238, p. 277).

With the connection ∇ of g on M, we can introduce on each induced coordinate neighborhood $\pi^{-1}(U)$ of T^*M a frame field which allows the tensor calculus to be efficiently done. The adapted frame on $\pi^{-1}(U)$ of T^*M consist of the following $2n$ linearly independent vector fields:

$$ \begin{cases} E_j = \partial_j + p_s\Gamma_{hj}^s\partial_{\overline{h}}, \\ E_{\overline{j}} = \partial_{\overline{j}}. \end{cases} \tag{2.7} $$

We can write the adapted frame as $\{E_\alpha\} = \left\{E_j, E_{\overline{j}}\right\}$. The indices $\alpha, \beta, \gamma, ... = 1, ..., 2n$ indicate the indices with respect to the adapted frame. By the straightforward calculations, we have the lemma below.

2.1. Lemma. *The Lie brackets of the adapted frame of T^*M satisfy the following identities:*

$$ \begin{aligned} [E_i, E_j] &= p_sR_{ijl}{}^sE_{\overline{l}}, \\ \left[E_i, E_{\overline{j}}\right] &= \Gamma_{il}^jE_{\overline{l}}, \\ \left[E_{\overline{i}}, E_{\overline{j}}\right] &= 0 \end{aligned} $$

where $R_{ijl}{}^s$ denote the components of the curvature tensor R of (M,g) ([28], p. 290).

Using (2.1), (2.2) and (2.7), we have

$$ {}^V\omega = \begin{pmatrix} 0 \\ \omega_j \end{pmatrix} \tag{2.8} $$

and

$$ {}^HX = \begin{pmatrix} X^j \\ 0 \end{pmatrix} \tag{2.9} $$

with respect to the adapted frame $\{E_\alpha\}$ (for details, see [28]).

3. The curvature tensor of the rescaled Riemannian metric of Cheeger-Gromoll type

From the equations (2.3)-(2.5), by virtue of (2.8) and (2.9), the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ has components with respect to the adapted frame $\{E_\alpha\}$:

$$(3.1) \qquad {}^{CG}g_f = diag\left(fg_{ij}, \frac{1}{\alpha}(g^{ij} + g^{is}g^{tj}p_s p_t) \right).$$

For the Levi-Civita connection of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ we give the next theorem.

3.1. Theorem. *Let (M, g) be a Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$. Then the corresponding Levi-Civita connection $\widetilde{\nabla}$ satisfies the followings:*

$$(3.2) \qquad \begin{cases} i)\ \widetilde{\nabla}_{E_i} E_j = \{\Gamma^l_{ij} + {}^f A^l_{ij}\} E_l + \dfrac{1}{2} p_s R_{ijl}{}^s E_{\bar{l}}, \\ ii)\ \widetilde{\nabla}_{E_i} E_{\bar{j}} = \dfrac{1}{2f\alpha} p_s R^{l}_{\cdot\ i}{}^{js}_{\ \cdot} E_l - \Gamma^j_{il} E_{\bar{l}}, \\ iii)\ \widetilde{\nabla}_{E_{\bar{i}}} E_j = \dfrac{1}{2f\alpha} p_s R^{l}_{\cdot\ j}{}^{is}_{\ \cdot} E_l, \\ iv)\ \widetilde{\nabla}_{E_{\bar{i}}} E_{\bar{j}} = \{\frac{-1}{\alpha}(p^i \delta^j_l + p^j \delta^i_l) + \frac{\alpha+1}{\alpha^2} g^{ij} p_l + \frac{1}{\alpha^2} p^i p^j p_l\} E_{\bar{l}} \end{cases}$$

with respect to the adapted frame, where ${}^f A^h_{ji}$ is a tensor field of type $(1, 2)$ defined by ${}^f A^h_{ji} = \frac{1}{f}(f_j \delta^h_i + f_i \delta^h_j - f^m_{\cdot} g_{ji})$ and $p^i = g^{it} p_t$, $R^{k}_{\cdot\ j}{}^{is}_{\ \cdot} = g^{kt} g^{im} R_{tjm}{}^s_{\cdot}$.

Proof. The connection $\widetilde{\nabla}$ is characterized by the Koszul formula:

$$\begin{aligned} 2\,{}^{CG}g_f(\widetilde{\nabla}_{\widetilde{X}} \widetilde{Y}, \widetilde{Z}) &= \widetilde{X}({}^{CG}g_f(\widetilde{Y}, \widetilde{Z})) + \widetilde{Y}({}^{CG}g_f(\widetilde{Z}, \widetilde{X})) - \widetilde{Z}({}^{CG}g_f(\widetilde{X}, \widetilde{Y})) \\ -{}^{CG}g_f(\widetilde{X}, [\widetilde{Y}, \widetilde{Z}]) &+ {}^{CG}g_f(\widetilde{Y}, [\widetilde{Z}, \widetilde{X}]) + {}^{CG}g_f(\widetilde{Z}, [\widetilde{X}, \widetilde{Y}]) \end{aligned}$$

for all vector fields $\widetilde{X}, \widetilde{Y}$ and $\widetilde{Z}$ on T^*M. One can verify the Koszul formula for pairs $\widetilde{X} = E_i, E_{\bar{i}}$ and $\widetilde{Y} = E_j, E_{\bar{j}}$ and $\widetilde{Z} = E_k, E_{\bar{k}}$. In calculations, the formulas (2.7), Lemma 2.1 and the first Bianchi identity for R should be applied. We omit standard calculations. □

Let $\tilde{X}$, $\tilde{Y} \in \Im^1_0(T^*M)$. Then the covariant derivative $\tilde{\nabla}_{\tilde{Y}} \tilde{X}$ has components

$$\tilde{\nabla}_{\tilde{Y}} \tilde{X}^\alpha = \tilde{Y}^\gamma E_\gamma \tilde{X}^\alpha + \tilde{\Gamma}^\alpha_{\gamma\beta} \tilde{X}^\beta \tilde{Y}^\gamma$$

with respect to the adapted frame $\{E_\alpha\}$. Using (2.7), (2.8), (2.9) and (3.2), we have the following proposition.

3.2. Proposition. *Let (M, g) be a Riemannian manifold and $\widetilde{\nabla}$ be the Levi-Civita connection of the cotangent bundle T^*M equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$. Then*

$$\begin{aligned} &i)\ \widetilde{\nabla}_{{}^H X} {}^H Y = {}^H\left(\nabla_X Y + {}^f A(X, Y)\right) + \tfrac{1}{2} {}^V\left(p \circ R(X, Y)\right), \\ &ii)\ \widetilde{\nabla}_{{}^H X} {}^V \theta = \tfrac{1}{2f\alpha} {}^H\left(p\left(g^{-1} \circ R(\ , X)\widetilde{\theta}\right)\right) + {}^V\left(\nabla_X \theta\right), \\ &iii)\ \widetilde{\nabla}_{{}^V \omega} {}^H Y = \tfrac{1}{2f\alpha} {}^H\left(p\left(g^{-1} \circ R(\ , Y)\widetilde{\omega}\right)\right), \\ &iv)\ \widetilde{\nabla}_{{}^V \omega} {}^V \theta = -\tfrac{1}{\alpha}\left({}^{CG}g\left({}^V\omega, \gamma\delta\right){}^V\theta + {}^{CG}g_f\left({}^V\theta, \gamma\delta\right){}^V\omega\right) \\ &+ \tfrac{\alpha+1}{\alpha} {}^{CG}g_f\left({}^V\omega, {}^V\theta\right)\gamma\delta - \tfrac{1}{\alpha} {}^{CG}g_f\left({}^V\omega, \gamma\delta\right){}^{CG}g_f\left({}^V\theta, \gamma\delta\right)\gamma\delta \end{aligned}$$

for all $X, Y \in \Im^1_0(M)$, $\omega, \theta \in \Im^0_1(M)$, where $\tilde{\omega} = g^{-1} \circ \omega \in \Im^1_0(M)$, $R(\ , X)\tilde{\omega} \in \Im^1_1(M)$, $g^{-1} \circ R(\ , X)\tilde{\omega} \in \Im^1_0(M)$, R and $\gamma\delta$ denote respectively the curvature tensor of ∇ and the canonical or Liouville vector field on T^*M with the local expression $\gamma\delta = p_i E_{\bar{i}}$ (for $f = 1$, see [1]).

The Riemannian curvature tensor $\widetilde{R}$ of T^*M with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ is obtained from the well-known formula

$$\widetilde{R}\left(\widetilde{X},\widetilde{Y}\right)\widetilde{Z}=\widetilde{\nabla}_{\widetilde{X}}\widetilde{\nabla}_{\widetilde{Y}}\widetilde{Z}-\widetilde{\nabla}_{\widetilde{Y}}\widetilde{\nabla}_{\widetilde{X}}\widetilde{Z}-\widetilde{\nabla}_{[\widetilde{X},\widetilde{Y}]}\widetilde{Z}$$

for all $\widetilde{X},\widetilde{Y},\widetilde{Z}\in\Im_0^1(T^*M)$. Then from Lemma 2.1 and Theorem 3.1, we get the following proposition.

3.3. Proposition. *The components of the curvature tensor $\widetilde{R}$ of the cotangent bundle T^*M with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ are given as follows:*

$$\begin{aligned}
\widetilde{R}(E_l,E_i)E_j &= \{R_{lij}{}^{m}-\frac{1}{2f\alpha}p_tp_aR_{lih}{}^{a}R^{m}{}_{.\,j.}{}^{ht}+\frac{1}{4f\alpha}p_tp_a(R^{m}{}_{.\,l.}{}^{ht}R_{ijh}{}^{a}-R^{m}{}_{.\,i.}{}^{ht}R_{ljh}{}^{a})\\
&\quad+\nabla_l(A_{ij}^m)-\nabla_i(A_{lj}^m)+A_{lh}^mA_{ij}^h-A_{ih}^mA_{lj}^h)\}E_m\\
&\quad+\{\frac{1}{2f}p_t(\nabla_lR_{ijm}{}^{t}-\nabla_iR_{ljm}{}^{t})+\frac{1}{2}p_t(R_{lhm}{}^{t}A_{ij}^h-R_{ihm}{}^{t}A_{lj}^h)\}E_{\bar{m}},\\
\widetilde{R}(E_{\bar{l}},E_i)E_j &= \{\frac{-1}{2f\alpha}p_a\nabla_iR^{m}{}_{.\,j.}{}^{la}+\frac{1}{2f\alpha}p_a(R^{m}{}_{.\,h.}{}^{la}A_{ij}^h-R^{h}{}_{.\,j.}{}^{la}A_{ih}^m+\frac{f_i}{f}R^{m}{}_{.\,j.}{}^{la})\}E_m\\
&\quad+\{\frac{1}{2}R_{ijm}{}^{l}-\frac{1}{4f\alpha}p_tp_aR_{ihm}{}^{t}R^{h}{}_{.\,j.}{}^{la}-\frac{1}{2\alpha}p_ap^lR_{ijm}{}^{a}-\frac{\alpha+1}{2\alpha^2}p_ap_mR_{ij.}{}^{la}\}E_{\bar{m}},\\
\widetilde{R}(E_l,E_{\bar{i}})E_j &= \{\frac{1}{2f\alpha}p_a\nabla_lR^{m}{}_{.\,j.}{}^{ia}+\frac{1}{2f\alpha}p_a(R^{h}{}_{.\,j.}{}^{ia}\ A_{lh}^m-R^{m}{}_{.\,h.}{}^{ia}A_{lj}^h-\frac{f_l}{f}R^{m}{}_{.\,j.}{}^{ia})\}E_m\\
&\quad+\{\frac{-1}{2}R_{ljm}{}^{i}-\frac{1}{4f\alpha}p_tp_aR_{lhm}{}^{a}R^{h}{}_{.\,j.}{}^{it}+\frac{1}{2\alpha}p_ap^iR_{ljm}{}^{a}-\frac{\alpha+1}{2\alpha^2}p_ap_mR_{lj.}{}^{ia}\}E_{\bar{m}},\\
\widetilde{R}(E_{\bar{l}},E_{\bar{i}})E_j &= \{\frac{1}{4f^2\alpha^2}p_tp_a(R^{m}{}_{.\,h.}{}^{la}R^{h}{}_{.\,j.}{}^{it}-R^{m}{}_{.\,h.}{}^{ia}R^{h}{}_{.\,j.}{}^{lt})+\frac{1}{f\alpha}R^{m}{}_{.\,j.}{}^{il})\\
&\quad+\frac{1}{f\alpha^2}p_a(p^iR^{m}{}_{.\,j.}{}^{la}-p^lR^{m}{}_{.\,j.}{}^{ia}\}E_m,\\
\widetilde{R}(E_l,E_i)E_{\bar{j}} &= \{\frac{1}{2f\alpha}p_a(\nabla_lR^{m}{}_{.\,i.}{}^{ja}-\nabla_iR^{m}{}_{.\,l.}{}^{ja})+\frac{1}{2f\alpha}p_a(R^{h}{}_{.\,i.}{}^{ja}A_{lh}^m-R^{h}{}_{.\,l.}{}^{ja}A_{ih}^m\\
&\quad-\frac{f_l}{f}R^{m}{}_{.\,i.}{}^{ja}+\frac{f_i}{f}R^{m}{}_{.\,l.}{}^{ja})\}E_m+\{R_{ilm}{}^{j}+\frac{1}{4f\alpha}p_tp_a(R_{lhm}{}^{t}R^{h}{}_{.\,i.}{}^{ja}\\
&\quad-R_{ihm}{}^{a}R^{h}{}_{.\,l.}{}^{jt})+\frac{1}{\alpha}p_ap^jR_{l\ im}{}^{a}-\frac{\alpha+1}{\alpha^2}p_ap_mR_{li.}{}^{ja}\}E_{\bar{m}},\\
\widetilde{R}(E_{\bar{l}},E_i)E_{\bar{j}} &= \{\frac{1}{2f\alpha}R^{m}{}_{.\,i.}{}^{jl}+\frac{1}{2f\alpha^2}p_a(p^lR^{m}{}_{.\,i.}{}^{ja}+p^iR^{m}{}_{.\,i.}{}^{la})+\frac{1}{4f^2\alpha^2}p_ap_tR^{m}{}_{.\,h.}{}^{la}R^{h}{}_{.\,i.}{}^{jt}\}E_m,\\
\widetilde{R}(E_l,E_{\bar{i}})E_{\bar{j}} &= \{\frac{-1}{2f\alpha}R^{m}{}_{.\,l.}{}^{ji}+\frac{1}{2f\alpha^2}p_a(p^iR^{m}{}_{.\,l.}{}^{ja}+p^jR^{m}{}_{.\,l.}{}^{ia})-\frac{1}{4f^2\alpha^2}p_ap_tR^{m}{}_{.\,h.}{}^{ia}R^{h}{}_{.\,l.}{}^{jt}\}E_m,\\
\widetilde{R}(E_{\bar{l}},E_{\bar{i}})E_{\bar{j}} &= \{\frac{\alpha^2+\alpha+1}{\alpha^3}(g^{ij}\delta_m^l-g^{jl}\delta_m^i)+\frac{\alpha+2}{\alpha^3}(g^{lj}p^ip_m-g^{ij}p_lp_m)\\
&\quad+\frac{\alpha-1}{\alpha^3}(\delta_m^ip^lp^j-\delta_m^lp^ip^j)\}E_{\bar{m}}
\end{aligned}$$

with respect to the adapted frame $\{E_\alpha\}$ (for $f=1$, see [1]*).*

4. Para-Kähler (or paraholomorphic) Norden structures on T^*M

An almost paracomplex manifold is an almost product manifold (M_{2k},φ), $\varphi^2=id$, $\varphi\neq\pm id$, such that the two eigenbundles T^+M_{2k} and T^-M_{2k} associated to the two eigenvalues +1 and -1 of φ, respectively, have the same rank. Note that the dimension of an almost paracomplex manifold is necessarily even. An almost paracomplex Norden manifold (M_{2k},φ,g) is defined to be a real differentiable manifold M_{2k} endowed with an almost paracomplex structure φ and a Riemannian metric g satisfying Nordenian property (or purity condition)

$$g(\varphi X, Y) = g(X, \varphi Y)$$

for any $X, Y \in \Im_0^1(M_{2k})$. The almost paracomplex Norden manifold (M_{2k}, φ, g) is called a paraholomorphic Norden manifold (or a para-Kähler-Norden manifold) such that $\nabla\varphi = 0$, where ∇ is the Levi-Civita connection of g. Also note that $\nabla\varphi = 0$ is equivalent to paraholomorphy of the metric g [21], i.e $\Phi_\varphi g = 0$, where Φ_φ is the Tachibana operator [27]:

$$\begin{aligned}(\Phi_\varphi g)(X, Y, Z) &= (\varphi X)(g(Y, Z)) - X(g(\varphi Y, Z)) \\ &+ g((L_Y\varphi)X, Z) + g(Y, (L_Z\varphi)X)\end{aligned}$$

for any $X, Y, Z \in \Im_0^1(M_{2k})$.

V. Cruceanu defined in [4] an almost paracomplex structure on T^*M as follows:

$$\left\{\begin{array}{l} J({}^HX) = -{}^HX, \\ J({}^V\omega) = {}^V\omega \end{array}\right. \tag{4.1}$$

for any $X \in \Im_0^1(M)$ and $\omega \in \Im_1^0(M)$. One can easily check that the metric ${}^{CG}g_f$ is pure with respect to the almost paracomplex structure J. Hence we state the following theorem.

4.1. Theorem. *Let (M, g) be a Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the paracomplex structure J. Then the triplet $(T^*M, J, {}^{CG}g_f)$ is an almost paracomplex Norden manifold.*

We now give conditions for the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ to be paraholomorphic with respect to the almost paracomplex structure J. Using definition of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the almost paracomplex structure J and by using the fact that ${}^V\omega^V(g^{-1}(\theta, \sigma)) = 0$ and ${}^HX^V(fg(Y, Z)) = {}^V(X(fg(Y, Z)))$ we calculate

$$\begin{aligned}(\Phi_J{}^{CG}g_f)(\tilde{X}, \tilde{Y}, \tilde{Z}) &= (J\tilde{X})({}^{CG}g_f(\tilde{Y}, \tilde{Z})) - \tilde{X}({}^{CG}g_f(J\tilde{Y}, \tilde{Z})) \\ &+ {}^{CG}g_f((L_{\tilde{Y}}J)\tilde{X}, \tilde{Z}) + {}^{CG}g_f(\tilde{Y}, (L_{\tilde{Z}}J)\tilde{X})\end{aligned}$$

for all $\tilde{X}, \tilde{Y}, \tilde{Z} \in \Im_0^1(T^*M)$. For pairs $\tilde{X} = {}^HX, {}^V\omega$, $\widetilde{Y} = {}^HY, {}^V\theta$ and $\widetilde{Z} = {}^HZ, {}^V\sigma$, we get

$$\begin{aligned}(\Phi_J{}^{CG}g_f)({}^HX, {}^V\theta, {}^HZ) &= 2\,{}^{CG}g_f({}^V\theta, {}^V(p \circ R(X, Z)), \\ (\Phi_J{}^{CG}g_f)({}^HX, {}^HY, {}^V\sigma) &= 2\,{}^{CG}g_f({}^V(p \circ R(X, Y), {}^V\sigma)\end{aligned} \tag{4.2}$$

and the others are zero. Therefore, we have the following result.

4.2. Theorem. *Let (M, g) be a Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the paracomplex structure J. Then the triplet $\left(T^*M, J, {}^{CG}g_f\right)$ is a para-Kähler-Norden (paraholomorphic Norden) manifold if and only if (M, g) is flat.*

4.3. Remark. Let (M, g) be a Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$. The diagonal lift ${}^D\gamma$ of $\gamma \in \Im_1^1(M)$ to T^*M is defined by the formulas

$$\begin{aligned}{}^D\gamma^HX &= {}^H(\gamma X), \\ {}^D\gamma^V\omega &= -{}^V(\omega \circ \gamma)\end{aligned}$$

for any $X \in \Im_0^1(M)$ and $\omega \in \Im_1^0(M)$. The diagonal lift ${}^D I$ of the identity tensor field $I \in \Im_1^1(M)$ has the following properties

$$\begin{aligned} {}^D I^H X &= {}^H X \\ {}^D I^V \omega &= -{}^V \omega \end{aligned}$$

and satisfies $({}^D I)^2 = I_{T^*M}$. Thus, ${}^D I$ is an almost paracomplex structure. Also, the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ is pure with respect to ${}^D I$, i.e. the triplet $\left(T^*M, {}^D I, {}^{CG}g_f\right)$ is an almost paracomplex Norden manifold. Finally, by using $\Phi-$operator, we can say that the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ is paraholomorphic with respect to ${}^D I$ if and only if (M, g) is flat.

The following remark follows directly from Proposition 3.3.

4.4. Remark. Let (M, g) be a flat Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$. Then the cotangent bundle $\left(T^*M, {}^{CG}g_f\right)$ is unflat.

As is known, the almost paracomplex Norden structure is a specialized Riemannian almost product structure on a Riemannian manifold. The theory of Riemannian almost product structures was initiated by K. Yano in [29]. The classification of Riemannian almost-product structure with respect to their covariant derivatives is described by A.M. Naveira in [15]. This is the analogue of the classification of almost Hermitian structures by A. Gray and L. Hervella in [10]. Having in mind these results, M. Staikova and K. Gribachev obtained a classification of the Riemannian almost product structures, for which the trace vanishes (see [26]). There are lots of physical applications involving a Riemannian almost product manifold. Now we shall give some applications for almost paracomplex Norden structures in context of almost product Riemannian manifolds.

4.1. Let us recall almost product Riemannian manifolds. If an n-dimensional Riemannian manifold M, endowed with a Riemannian metric g, admits a non-trivial tensor field F of type (1.1) such that

$$F^2 = I$$

and

$$g(FX, Y) = g(X, FY)$$

for all $X, Y \in \Im_0^1(M)$, then F is called an almost product structure and (M, F, g) is called an almost product Riemannian manifold. An almost product Riemannian manifold with integrable almost product F is called a locally product Riemannian manifold. It is known that the integrability of an almost product structure F is equivalent to the vanishing of the Nijenhuis tensor N_F given by

$$N_F(X, Y) = [FX, FY] - F[FX, Y] - F[X, FY] + [X, Y]$$

for all $X, Y \in \Im_0^1(M)$. If F is covariantly constant with respect to the Levi-Civita connection ∇ of g which is equivalent to $\Phi_F g = 0$, then (M, F, g) is called a locally decomposable Riemannian manifold.

Now consider the almost product structure J defined by (4.1) and the Levi-Civita connection $\widetilde{\nabla}$ given by Proposition 3.1. We define a tensor field $\widetilde{S}$ of type $(1, 2)$ on T^*M by

$$\widetilde{S}(\widetilde{X}, \widetilde{Y}) = \frac{1}{2}\{(\widetilde{\nabla}_{J\widetilde{Y}} J)\widetilde{X} + J((\widetilde{\nabla}_{\widetilde{Y}} J)\widetilde{X}) - J((\widetilde{\nabla}_{\widetilde{X}} J)\widetilde{Y})\}$$

for all $\widetilde{X}, \widetilde{Y} \in \Im_0^1(T^*M)$. Then the linear connection

$$\overline{\nabla}_{\widetilde{X}} \widetilde{Y} = \widetilde{\nabla}_{\widetilde{X}} \widetilde{Y} - \widetilde{S}(\widetilde{X}, \widetilde{Y}) \tag{4.3}$$

is an almost product connection on T^*M (for almost product connection, see [12]).

4.5. Theorem. *Let (M,g) be a Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the almost product structure J. Then the almost product connection $\overline{\nabla}$ constructed by the Levi-Civita connection $\widetilde{\nabla}$ of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the almost product structure J is as follows:*

$$(4.4)\qquad \begin{cases} i)\ \overline{\nabla}_{{}^HX}{}^HY = {}^H(\nabla_X Y) + {}^H({}^fA(X,Y)), \\ ii)\ \overline{\nabla}_{{}^HX}{}^V\theta = {}^V(\nabla_X\theta), \\ iii)\ \overline{\nabla}_{{}^V\omega}{}^HY = \dfrac{3}{2f\alpha}\,{}^H\left(p\left(g^{-1}\circ R(\ ,Y)\tilde{\omega}\right)\right), \\ iv)\ \overline{\nabla}_{{}^V\omega}{}^V\theta = -\frac{1}{\alpha}\left({}^{CG}g\left({}^V\omega,\gamma\delta\right){}^V\theta + {}^{CG}g_f\left({}^V\theta,\gamma\delta\right){}^V\omega\right) \\ \qquad\qquad + \frac{\alpha+1}{\alpha}{}^{CG}g_f\left({}^V\omega,{}^V\theta\right)\gamma\delta - \frac{1}{\alpha}{}^{CG}g_f\left({}^V\omega,\gamma\delta\right) \\ \qquad\qquad \times {}^{CG}g_f\left({}^V\theta,\gamma\delta\right)\gamma\delta. \end{cases}$$

Denoting by $\overline{T}$ the torsion tensor of $\overline{\nabla}$, we have from (4.1), (4.3) and (4.4)

$$\begin{aligned} \overline{T}({}^V\omega,{}^V\theta) &= 0, \\ \overline{T}({}^V\omega,{}^HY) &= \frac{3}{2f\alpha}\,{}^H\left(p\left(g^{-1}\circ R(\ ,Y)\tilde{\omega}\right)\right), \\ \overline{T}({}^HX,{}^HY) &= -{}^V(p\circ R(X,Y)). \end{aligned}$$

Hence we have the theorem below.

4.6. Theorem. *Let (M,g) be a Riemannian manifold and let T^*M be its cotangent bundle. Then the almost product connection $\overline{\nabla}$ constructed by the Levi-Civita connection $\widetilde{\nabla}$ of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the almost product structure J is symmetric if and only if (M,g) is flat.*

As is well-known, if there exists a symmetric almost product connection on M then the almost product structure J is integrable [12]. The converse is also true [6]. Thus we get the following conclusion.

4.7. Corollary. *Let (M,g) be a Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the almost product structure J. Then the triplet $(T^*M, J, {}^{CG}g_f)$ is a locally product Riemannian manifold if and only if (M,g) is flat.*

Similarly, let us consider the almost product structure ${}^DI$ and the Levi-Civita connection $\widetilde{\nabla}$ of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$. Another almost product connection can be constructed.

If J is covariantly constant with respect to the Levi-Civita connection $\widetilde{\nabla}$ of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ which is equivalent to $\Phi_J{}^{CG}g_f = 0$, then $(T^*M, J, {}^{CG}g_f)$ is called a locally decomposable Riemannian manifold. In view of Theorem 4.2, we have the following.

4.8. Corollary. *Let (M,g) be a Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the almost product structure J. Then the triplet $(T^*M, J, {}^{CG}g_f)$ is a locally decomposable Riemannian manifold if and only if (M,g) is flat.*

4.2. Let (M_{2k},φ,g) be a non-integrable almost paracomplex manifold with a Norden metric. An almost paracomplex Norden manifold (M_{2k},φ,g) is a quasi-para-Kähler–Norden manifold, if $\underset{X,Y,Z}{\sigma} g((\nabla_X\varphi)Y,Z) = 0$, where σ is the cyclic sum by three arguments [13]. In [24], the authors proved that $\underset{X,Y,Z}{\sigma} g((\nabla_X\varphi)Y,Z) = 0$ is equivalent to

$\underset{X,Y,Z}{\sigma}(\Phi_\varphi g)(X,Y,Z)=0$. We compute

$$A(\tilde{X},\tilde{Y},\tilde{Z})=\underset{\tilde{X},\tilde{Y},\tilde{Z}}{\sigma}(\Phi_J{}^{CG}g_f)(\tilde{X},\tilde{Y},\tilde{Z})$$

for all $\widetilde{X},\widetilde{Y},\widetilde{Z}\in\Im_0^1(T^*M)$. By means of (4.2), we have $A(\tilde{X},\tilde{Y},\tilde{Z})=0$ for all $\widetilde{X},\widetilde{Y},\widetilde{Z}\in\Im_0^1(T^*M)$. Hence we state the following theorem.

4.9. Theorem. *Let (M,g) be a Riemannian manifold and T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the almost paracomplex structure J defined by (4.1). Then the triplet $(T^*M,J,{}^{CG}g_f)$ is a quasi-para-Kähler–Norden manifold.*

O. Gil-Medrano and A.M. Naveira proved that both distributions of the almost product structure on the Riemannian manifold (M,φ,g) are totally geodesic if and only if $\underset{X,Y,Z}{\sigma}g((\nabla_X\varphi)Y,Z)=0$ for any $X,Y,Z\in\Im_0^1(M)$ [9]. As a consequence of Theorem 4.9, we have the following.

4.10. Corollary. *Both distributions of the almost product Riemannian manifold $(T^*M,J,{}^{CG}g_f)$ are totally geodesic.*

4.3. Let F be an almost product structure and ∇ be a linear connection on an n-dimensional Riemannian manifold M. The product conjugate connection $\nabla^{(F)}$ of ∇ is defined by

$$\nabla_X^{(F)}Y=F(\nabla_XFY)$$

for all $X,Y\in\Im_0^1(M)$. If (M,F,g) is an almost product Riemannian manifold, then $(\nabla_X^{(F)}g)(FY,FZ)=(\nabla_Xg)(Y,Z)$, i.e. ∇ is a metric connection with respect to g if and only if $\nabla^{(F)}$ is so. From this, we can say that if ∇ is the Levi-Civita connection of g, then $\nabla^{(F)}$ is a metric connection with respect to g [2].

By the almost product structure J defined by (4.1) and the Levi-Civita connection $\widetilde{\nabla}$ given by Proposition 3.1, we write the product conjugate connection $\widetilde{\nabla}^{(J)}$ of $\widetilde{\nabla}$ as follows:

$$\widetilde{\nabla}_{\widetilde{X}}^{(J)}\widetilde{Y}=J(\widetilde{\nabla}_{\widetilde{X}}J\widetilde{Y})$$

for all $\widetilde{X},\widetilde{Y}\in\Im_0^1(T^*M)$. Also note that $\widetilde{\nabla}^{(J)}$ is a metric connection of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$. The standart calculations give the following theorem.

4.11. Theorem. *Let (M,g) be a Riemannian manifold and let T^*M be its cotangent bundle equipped with the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ and the almost product structure J. Then the product conjugate connection (or metric connection) $\widetilde{\nabla}^{(J)}$ is as follows:*

$$
\begin{aligned}
&i)\ \widetilde{\nabla}_{{}^HX}{}^HY={}^H\left(\nabla_XY+{}^fA(X,Y)\right)-\tfrac{1}{2}{}^V\left(p\circ R\left(X,Y\right)\right),\\
&ii)\ \widetilde{\nabla}_{{}^HX}{}^V\theta=-\tfrac{1}{2f\alpha}{}^H\left(p\left(g^{-1}\circ R\left(\ ,X\right)\widetilde{\theta}\right)\right)+{}^V\left(\nabla_X\theta\right),\\
&iii)\ \widetilde{\nabla}_{{}^V\omega}{}^HY=\tfrac{1}{2f\alpha}{}^H\left(p\left(g^{-1}\circ R\left(\ ,Y\right)\tilde{\omega}\right)\right),\\
&iv)\ \widetilde{\nabla}_{{}^V\omega}{}^V\theta=-\tfrac{1}{\alpha}\left({}^{CG}g\left({}^V\omega,\gamma\delta\right){}^V\theta+{}^{CG}g_f\left({}^V\theta,\gamma\delta\right){}^V\omega\right)\\
&+\tfrac{\alpha+1}{\alpha}{}^{CG}g_f\left({}^V\omega,{}^V\theta\right)\gamma\delta-\tfrac{1}{\alpha}{}^{CG}g_f\left({}^V\omega,\gamma\delta\right){}^{CG}g_f\left({}^V\theta,\gamma\delta\right)\gamma\delta.
\end{aligned}
$$

The relationship between curvature tensors R_∇ and $R_{\nabla^{(F)}}$ of the connections ∇ and $\nabla^{(F)}$ is as follows: $R_{\nabla^{(F)}}(X,Y,Z)=F(R_\nabla(X,Y,FZ)$ for all $X,Y,Z\in\Im_0^1(M)$ [2]. By means of the almost product structure J defined by (4.1) and Proposition 3.3, from $\widetilde{R}_{\widetilde{\nabla}^{(J)}}(\widetilde{X},\widetilde{Y},\widetilde{Z})=J(\widetilde{R}_{\widetilde{\nabla}}(\widetilde{X},\widetilde{Y},J\widetilde{Z})$, components of the curvature tensor $\widetilde{R}_{\widetilde{\nabla}^{(J)}}$ of the product conjugate connection (or metric connection) $\widetilde{\nabla}^{(J)}$ can easily be computed. Finally, using the almost product structure ${}^{D}I$, another metric connection of the rescaled Cheeger-Gromoll type metric ${}^{CG}g_f$ can be constructed.

Acknowledgement. The authors are thankful to the anonymous referees for the valuable suggestions towards the improvement of this manuscript.

References

[1] Agca, F. and Salimov, A. A. *Some notes concerning Cheeger-Gromoll metrics*, Hacet. J. Math. Stat. **42** (5), 533-549, 2013.

[2] Blaga, A. M. and Crasmareanu, M. *The geometry of product conjugate connections*, An. Stiint. Univ. Al. I. Cuza Iasi. Mat. (N.S.), **59** (1), 73-84, 2013.

[3] Cheeger, J. and Gromoll, D. *On the structure of complete manifolds of nonnegative curvature*, Ann. of Math. **96**, 413-443, 1972.

[4] Cruceanu, V. *Une classe de structures géométriques sur le fibré cotangent*, Tensor (N.S.) **53** Commemoration Volume I, 196-201, 1993.

[5] Druta, L. S. *Classes of general natural almost anti-Hermitian structures on the cotangent bundles*, Mediterr. J. Math. **8** (2), 161-179, 2011.

[6] Fujimoto, A. *Theory of G-structures*, Publ. Study Group of Geometry, 1, Tokyo Univ., Tokyo, 1972.

[7] Gezer, A. and Altunbas, M. *Some notes concerning Riemannian metrics of Cheeger-Gromoll type*, J. Math. Anal. Appl. **396** (1), 119-132, 2012.

[8] Gezer, A. and Altunbas, M. *Notes on the rescaled Sasaki type metric on the cotangent bundle*, Acta Math. Sci. Ser. B Engl. Ed. **34** (1), 162-174, 2014.

[9] Gil-Medrano O. and Naveira, A. M. *Some remarks about the Riemannian curvature operator of a Riemannian almost-product manifold*, Rev. Roum. Math. Pures Appl. **30** (18), 647-658, 1985.

[10] Gray, A. and Hervella, L.M. *The sixteen classes of almost Hermitian manifolds and their linear invariants*, Ann. Mat. Pura Appl. IV. Ser. **123**, 35-58, 1980.

[11] Kowalski, O. and Sekizawa, M. *On the geometry of orthonormal frame bundles*, Math. Nachr. **281** (12), 1799-1809, 2008.

[12] de Leon, M. and Rodrigues, P. R. *Methods of Differential Geometry in Analytical Mechanics*, North-Holland Mathematics Studies, 1989.

[13] Manev, M. and Mekerov, D. *On Lie groups as quasi-Kähler manifolds with Killing Norden metric*, Adv. Geom. **8** (3), 343-352, 2008.

[14] Musso, E. and Tricerri, F. *Riemannian Metrics on Tangent Bundles*, Ann. Mat. Pura. Appl. **150** (4), 1-19, 1988.

[15] Naveira, A.M. *A classification of Riemannian almost-product manifolds*, Rend. Mat.Appl. VII. Ser. **3**, 577-592, 1983.

[16] Olszak, Z. *On almost complex structures with Norden metrics on tangent bundles*, Period. Math. Hungar. **51** (2), 59-74, 2005.

[17] Oproiu, V. and Papaghiuc, N. *Einstein quasi-anti-Hermitian structures on the tangent bundle*, An. Ştiinţ. Univ. Al. I. Cuza Iaşi. Mat. (N.S.) **50** (2), 347-360, 2004.

[18] Oproiu, V. and Papaghiuc, N. *Classes of almost anti-Hermitian structures on the tangent bundle of a Riemannian manifold*, An. Ştiinţ. Univ. Al. I. Cuza Iaşi. Mat. (N.S.) **50** (1), 175-190, 2004.

[19] Oproiu, V. and Papaghiuc, N. *Some classes of almost anti-Hermitian structures on the tangent bundle*, Mediterr. J. Math. **1** (3), 269-282, 2004.

[20] Papaghiuc, N. *A locally symmetric complex structure with Norden metric on the tangent bundle of a space form*, An. Ştiinţ. Univ. Al. I. Cuza Iaşi. Mat. (N.S.) **48** (1), 101-112, 2002.

[21] Salimov, A. A., Iscan, M. and Etayo, F. *Paraholomorphic B-manifold and its properties*, Topology Appl. **154** (4), 925-933, 2007.

[22] Salimov, A., Gezer, A. and Iscan, M. *On para-Kähler-Norden structures on the tangent bundles*, Ann. Polon. Math. **103** (3), 247-261, 2012.

[23] Salimov, A. and Gezer, A. *On the geometry of the (1,1) -tensor bundle with Sasaki type metric*, Chin. Ann. Math. Ser. B **32** (3), 369-386, 2011.

[24] Salimov, A. A., Iscan, M. and Akbulut, K. *Notes on para-Norden-Walker 4-manifolds*, Int. J. Geom. Methods Mod. Phys. **7** (8), 1331-1347, 2010.

[25] Sasaki, S. *On the differential geometry of tangent bundles of Riemannian manifolds*, Tohoku Math. J. **10**, 338-358, 1958.

[26] Staikova, M.T. and Gribachev, K. I. *Canonical connections and their conformal invariants on Riemannian almost-product manifolds*, Serdica **18** (3-4), 150-161, 1992.

[27] Yano, K. and Ako, M. *On certain operators associated with tensor field*, Kodai Math. Sem. Rep. **20**, 414-436, 1968.

[28] Yano, K.and Ishihara, S. *Tangent and Cotangent Bundles*, Marcel Dekker, Inc., New York 1973.

[29] Yano, K. *Differential geometry on complex and almost complex spaces*, Pure and Applied Math. vol. 49, New York, Pergamon Press Book, 1965.

[30] Zayatuev, B. V. *On some classes of almost-Hermitian structures on the tangent bundle*, Webs and Quasigroups. T.S.U. 103–106, 2002.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 367 – 371

Continuous dependence of solutions to fourth-order nonlinear wave equation

İpek Güleç*† and Şevket Gür‡

Abstract

This paper gives a priori estimates and continuous dependence of the solutions to fourth-order nonlinear wave equation.

Keywords: Continuous dependence, nonlinear wave equation.

2000 AMS Classification: 35B30, 35L35.

Received : 08.09.2014 *Accepted* : 10.03.2015 *Doi* : 10.15672/HJMS.20164512496

1. Introduction

We consider the following initial boundary value problem

$$u_{tt} - \alpha\Delta u - \beta\Delta u_t - \gamma\Delta u_{tt} = f(u) \tag{1.1}$$

$$u(x,0) = u_0(x), \qquad u_t(x,0) = u_1(x),\ x \in \Omega \tag{1.2}$$

$$u = 0,\ x \in \partial\Omega, t > 0, \tag{1.3}$$

where $\Omega \subset \mathbb{R}^n$ is bounded region with smooth boundary $\partial\Omega$; α,β and γ are positive constants. $f(u)$ is a given nonlinear function which satisfies

$$f \in C^1(R), \left|f'(u)\right| \leq c\left(1 + |u|^{p-1}\right), p \geq 1, (n-2)\, p \leq n \tag{1.4}$$

and

$$\limsup_{u\to\infty} \frac{f(u)}{u} < \alpha\lambda_1 \tag{1.5}$$

where λ_1 is the first eigenvalue of the Laplace operator with the homogeneous Dirichlet boundary condition.

Continuous dependence of solutions on coefficients of equations is a type of structural stability, which reflects the effect of small changes in coefficients of equations on the solutions. Many results of this type can be found in [1].

*Department of Mathematics, Hacettepe University, Ankara, Turkey.
Email : `ipek@hacettepe.edu.tr`
†Corresponding Author.
‡Department of Mathematics, Sakarya University, Sakarya, Turkey.
Email : `sgur@sakarya.edu.tr`

In [2], authors studied asymptotic behaviour of solution to initial value problem of fourth order wave equation with dispersive and dissipative terms by taking coefficients $\alpha = \beta = \gamma = 1$ in (1). They proved that the global strong solution of the problem decays to zero exponentially as $t \to \infty$. The authors Guo-wang Chen and Chang-Shun Hou, in article [3], studied the following initial value problem for a class of fourth order nonlinear wave equations,

$$v_{tt} - a_1 v_{xx} - a_2 v_{xxt} - a_3 v_{xxtt} = f(v_x)_x \quad , x \in R, t > 0$$

$$v(x,0) = v_0(x),\ v_t(x,0) = v_1(x),\ x \in R$$

where a_1, a_2, a_3 are positive constants. They gave also the blow up results for this problem.

In [4], Shang studied the initial boundary value problem

$$\text{(1')} \qquad u_{tt} - \Delta u - \Delta u_t - \Delta u_{tt} = f(u), x \in \Omega, t > 0$$

$$\text{(2')} \qquad u(x,0) = u_0(x), \qquad u_t(x,0) = u_1(x),\ x \in \Omega$$

$$\text{(3')} \qquad u = 0,\ x \in \partial\Omega, t > 0,$$

Under the assumptions that $n = 1,2,3; f \in C^1, f'(u)$ is bounded above and satisfies (i) $|f'(u)| \leq A|u|^p + B, 0 < p < \infty$ if $n = 2$; $0 < p \leq \frac{2}{n-2}$ if $n = 3; u_i(x) \in H^2(\Omega) \cap H_0^1(\Omega)\,(i = 0,1)$, it was proven that problem (1')-(3') admits unique global strong solution u such that $\forall T > 0, u \in W^{2,\infty}\left(0,T; H^2(\Omega) \cap H_0^1(\Omega)\right)$.

In [5], problem (1')-(3') were studied again for all $n \geq 1$. By supposing that $f \in C^1$ and $f'(u)$ is bounded above satisfying (ii)$|f'(u)| \leq A|u|^p + B, 0 < p < \infty$ if $n = 2$; $0 < p \leq \frac{4}{n-2}$ if $n \geq 3, u_i(x) \in H^2(\Omega) \cap H_0^1(\Omega)\,(i = 0,1)$, it was proven that problem (1')-(3') admits unique global strong solution u such that for all $T > 0, u \in W^{2,\infty}\left(0,T; H^2(\Omega) \cap H_0^1(\Omega)\right)$.

In [6], authors studied the spatial behavior of a coupled system of wave-plate type . They got the alternative results of Phragmen-Lindelof type in terms of an area measure of the amplitude in question based on a first-order differential inequality. They also got the spatial decay estimates based on a second-order differential inequality.

The aim of this paper is to prove the continuous dependence of solutions to the problem (1)-(3) on coefficients α, β and γ.

Throughout this paper , we use the notation $\|.\|_p$ for the norm in $L^P(\Omega)$. We use $\|.\|$ instead of $\|.\|_2$.

2. A Priori Estimates

In this section, we obtain a priori estimates for the problem (1)-(3).

2.1. Theorem. *Assume that the conditions (4) and (5) hold. Then for $u_0, u_1 \in H_0^1(\Omega)$ the solution u of problem (1)-(3) satisfies the following estimates:*

$$\text{(2.1)} \qquad \|\nabla u\|^2 + \|\nabla u_t\|^2 \leq D_1$$

and

$$\text{(2.2)} \qquad \int_0^t \|\nabla u_{ss}\|^2 \, ds \leq D_2 t$$

for any $t > 0$. Here $D_1 > 0$ and $D_2 > 0$ depend on initial data and the parameters of (1).

Proof. First, by taking the inner product of (1) by u_t in $L^2(\Omega)$ and integrating by parts, we get

$$\text{(2.3)}\qquad \frac{d}{dt}\left[\frac{1}{2}\|u_t\|^2+\frac{\alpha}{2}\|\nabla u\|^2+\frac{\gamma}{2}\|\nabla u_t\|^2-\int_\Omega F(u)dx\right]+\beta\|\nabla u_t\|^2=0$$

and

$$\text{(2.4)}\qquad E(t)\le E(0)$$

where $F(u)=\int_0^u f(s)\,ds$ and $E(t)=\frac{1}{2}\|u_t\|^2+\frac{\alpha}{2}\|\nabla u\|^2+\frac{\gamma}{2}\|\nabla u_t\|^2-\int_\Omega F(u)dx$. From (5) and definition of limsup we obtain

$$\text{(2.5)}\qquad F(u)\le c+\frac{\alpha\lambda_1}{2}u^2-\frac{\varepsilon}{2}u^2$$

Using (10) and Poincare's inequality from (9) we find (6).
Next we multiply (1) by u_{tt} in $L^2(\Omega)$ to get

$$\text{(2.6)}\qquad \frac{d}{dt}\frac{\beta}{2}\|\nabla u_t\|^2+\gamma\|\nabla u_{tt}\|^2+\|u_{tt}\|^2+\alpha\int_\Omega \nabla u\nabla u_{tt}dx=\int_\Omega f(u)u_{tt}dx$$

Using Cauchy-Schwarz inequality, ε-Cauchy inequality and from (4), we take,

$$\text{(2.7)}\qquad (\gamma\ -\frac{\varepsilon}{2})\|\nabla u_{tt}\|^2+\frac{d}{dt}\frac{\beta}{2}\|\nabla u_t\|^2\le c_2+\frac{|\alpha|^2}{2\varepsilon}\|\nabla u\|^2+\frac{c_1^2}{2}\int_\Omega |u|^{2p}\,dx$$

where c_1, c_2 are constants and ε is sufficiently small and positive. Using Sobolev inequality and (6) we have

$$\text{(2.8)}\qquad \int_\Omega |u|^{2p}\,dx=\|u\|_{2p}^{2p}\le c_3\|\nabla u\|^{2p}\le c_4$$

where c_3 is a Sobolev constant and $c_4=c_4(\alpha,\gamma,u_0,u_1)$. From (12) and (13) we obtain

$$\text{(2.9)}\qquad (\gamma\ -\frac{\varepsilon}{2})\|\nabla u_{tt}\|^2+\frac{d}{dt}\frac{\beta}{2}\|\nabla u_t\|^2\le c_5$$

where c_5 depends on initial data and the parameters of (1). Now, we integrate (14) from (0,t), then we obtain

$$\text{(2.10)}\qquad \int_0^t \|\nabla u_{ss}\|^2\,ds\le c_6 t$$

where c_6 depends on initial data and the parameters of (1). Hence, (7) follows from (15).

3. Continuous Dependence on the Coefficients

In this section, we prove that the solution of the problem (1)-(3) depends continuously on the coefficients α,β and γ in $H^1(\Omega)$.

We consider the problem

$$\text{(3.1)}\qquad u_{tt}-\alpha_1\Delta u-\beta_1\Delta u_t-\gamma_1\Delta u_{tt}=f(u)$$

$$\text{(3.2)}\qquad u(x,0)=0,u_t(x,0)=0$$

$$\text{(3.3)}\qquad u|_{\partial\Omega}=0$$

and

$$\text{(3.4)}\qquad v_{tt}-\alpha_2\Delta v-\beta_2\Delta v_t-\gamma_2\Delta v_{tt}=f(v)$$

(3.5) $\quad v(x,0)=0, v_t(x,0)=0$

(3.6) $\quad v|_{\partial\Omega}=0$

Let us define the difference variables w, α, β and γ by w=u-v, $\alpha=\alpha_1-\alpha_2$, $\beta=\beta_1-\beta_2$ and $\gamma=\gamma_1-\gamma_2$ then w satisfy the following the initial boundary value problem:

(3.7) $\quad w_{tt}-\alpha_1\Delta w-\alpha\Delta v-\beta_1\Delta w_t-\beta\Delta v_t-\gamma_1\Delta w_{tt}-\gamma\Delta v_{tt}=f(u)-f(v)$

(3.8) $\quad w(x,0)=0, w_t(x,0)=0$

(3.9) $\quad w|_{\partial\Omega}=0$

The main result of this section is the following theorem.

3.1. Theorem. *Let w be the solution of the problem (22)-(24). If*

(3.10) $\quad |f(u)-f(v)|\le c_7\left(1+|u|^{p-1}+|v|^{p-1}\right)|u-v|$

holds, then w satisfies the estimate

$$\|w_t\|^2+\|\nabla w\|^2+\|\nabla w_t\|^2\le e^{Mt}K\left[(\alpha_1-\alpha_2)^2+(\beta_1-\beta_2)^2+(\gamma_1-\gamma_2)^2\right]t$$

where M and K are positive constants depending on initial data and the parameters of (1).

Proof. Let us take the inner product of (22) with w_t in $L^2(\Omega)$; we have

$$\frac{d}{dt}\left[\frac{1}{2}\|w_t\|^2+\frac{\alpha_1}{2}\|\nabla w\|^2+\frac{\gamma_1}{2}\|\nabla w_t\|^2\right]+\beta_1\|\nabla w_t\|^2+$$

(3.11) $$\alpha\int\limits_\Omega\nabla v\nabla w_t dx+\beta\int\limits_\Omega\nabla v_t\nabla w_t dx+\gamma\int\limits_\Omega\nabla v_{tt}\nabla w_t dx=\int\limits_\Omega|f(u)-f(v)|\,w_t dx$$

From (26) we obtain

$$\frac{d}{dt}E_1(t)+\beta_1\|\nabla w_t\|^2\le|\alpha|\,\|\nabla w_t\|\,\|\nabla v\|+|\beta|\,\|\nabla w_t\|\,\|\nabla v_t\|+$$

(3.12) $$|\gamma|\,\|\nabla w_t\|\,\|\nabla v_{tt}\|+\int\limits_\Omega|f(u)-f(v)|\,w_t dx$$

where $E_1(t)=\frac{1}{2}\|w_t\|^2+\frac{\alpha_1}{2}\|\nabla w\|^2+\frac{\gamma_1}{2}\|\nabla w_t\|^2$.

Using the Holder, Sobolev, Cauchy-Schwarz inequalities and (25) we obtain the estimate

$$\int\limits_\Omega|f(u)-f(v)|\,w_t dx\le c_7\int\limits_\Omega\left(1+|u|^{p-1}+|v|^{p-1}\right)|w|\,w_t dx$$

$$\le c_8\left(1+\|\nabla u\|^{p-1}+\|\nabla v\|^{p-1}\right)\|w\|_{\frac{2n}{n-2}}\|w_t\|$$

(3.13) $$\le C\left(\|\nabla w\|^2+\|w_t\|^2\right)$$

where c_7, c_8 are constants and $C=C(c_7,c_8)$.Using Cauchy-Schwarz inequality and (28), from (27), we get

$$\frac{d}{dt}E_1(t)+(\beta_1-\varepsilon)\|\nabla w_t\|^2\le\frac{3}{4\varepsilon}|\alpha|^2\|\nabla v\|^2+\frac{3}{4\varepsilon}|\beta|^2\|\nabla v_t\|^2+$$

(3.14) $$\frac{3}{4\varepsilon}|\gamma|^2\|\nabla v_{tt}\|^2+c_9\left(\|\nabla w\|^2+\|w_t\|^2\right)$$

and from (29) we can write

(3.15) $$\frac{d}{dt}E_1(t)\le\frac{3}{4\varepsilon}\left(|\alpha|^2\|\nabla v\|^2+|\beta|^2\|\nabla v_t\|^2+|\gamma|^2\|\nabla v_{tt}\|^2\right)+ME_1(t)$$

where $M=\frac{2C(1+\alpha_1)}{\alpha_1}$. Applying Gronwall's inequality with (6) and (7), we get

(3.16) $$E_1(t)\le e^{Mt}K\left(|\alpha|^2+|\beta|^2+|\gamma|^2\right)t$$

Hence proof is completed.

References

[1] K.A. Ames, B. Straughan, *Non-Standard and Improperly Posed Problems*, in: Mathematics in science and Engineering series, vol.194, Academic Press, San Diego, 1997.

[2] Xu Run-zhang ,Zhao Xi-ren, Shen Ji-hong, *Asymptotic behaviour of solution for fourth order wave equation with dispersive and dissipative terms*, Appl. Math. Mech.-Engl. Ed., 29(2), 259-262, 2008.

[3] Guo-wang Chen, Chang-shun Hou, *Initial value problem for a class of fourth-order nonlinear wave equations*, Appl. Math. Mech.-Engl. Ed. 30(3), 391-401, 2009.

[4] Shang Yadong, *Initial boundary value problem of equation* $u_{tt} - \Delta u - \Delta u_t - \Delta u_{tt} = f(u)$, Acta Mathematicae Applicate Sinica 23(3), 385-393, 2000.

[5] Liu Yacheng, Li Xiaoyuan, *Some remarks on the equation* $u_{tt} - \Delta u - \Delta u_t - \Delta u_{tt} = f(u)$, Journal of Natural Science of Heilongjiang University 21(3),1-6, 2004.

[6] Gusheng Tang, yan Liu and Wenhui Liao, *Spatial Behavior of a Coupled System of Wave-Plate Type*, Abstract and Applied Analysis, Volume 2014, Article ID 853693, 13 pages.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 373 – 379

Null controllability of heat equation with switching controls under Robin's boundary condition

Ali Hamidoğlu*

Abstract

In this paper, we consider the null controllability of 1-d heat equation endowed with Robin's boundary conditions, when the operator $-\frac{d^2}{dx^2}$ has positive eigenvalues and try to find sufficient conditions for building switching controls. In [1], the author developed a first analysis of this problem with Dirichlet's boundary conditions and obtain sufficient conditions for switching controls. We firstly consider 1-d heat system endowed with two controls. Then we try to build switching control strategies guaranteeing that, at each instant of time, only one control is activated.

Keywords: Heat equation, Robin's boundary condition, variational approach, switching control.

2000 AMS Classification: 93B05, 93C20.

Received : 11.12.2014 *Accepted* : 15.04.2015 *Doi* : 10.15672/HJMS.20164512501

1. Introduction

First of all, we define the general problem of controllability in PDEs. Meanly, it consists in investigating whether the solution of the PDE can be driven to a given final target by means of a control. More precisely, the controllability problem may be defined as follows. Consider an evolution system with given a time interval $t \in (0, T)$, initial and final states. We try to find a suitable control such that the solution matches both the initial state at time $t = 0$ and the final one at time $t = T$. This is a type of exact controllability problem. There are other type of controllability problems beside that exact one. For instance, when the final target is achieved to zero, then the system is null controllable or when the set of reachable states (set of final targets) is dense in the space where the evolution system is satisfied, then the system is approximate controllable. These different concepts coincide in finite dimensional space. Because, in finite-dimensional space the only close affine dense subspace is the whole space itself. But this is no longer the case in the context of PDE. Indeed, in infinite- dimensional

*Department of Mathematics, Middle East Technical University, Ankara/Turkey.
Email: `ali.hamidoglu@metu.edu.tr`.

spaces we can easily find strict dense subspaces, while in finite-dimension they do not exist. These are classical problems in control theory and there is a large literature on the topic. We refer for instance to the book by Lee and Marcus [8] for an introduction in the context of finite-dimensional systems. We also refer to the survey article by Russell [5], the articles by Zuazua [3, 4] and to the SIAM Review article and the book of Lions [6, 7] for an introduction to the controllability of PDE, also referred to as Distributed Parameter Systems.

This paper deals with some of new results in null controllability of 1-d heat equation with switching controls under Robin's boundary condition. We firstly consider the 1-d heat equation endowed with two boundary controls and lumped controls under Robin's boundary condition, when the operator $-\frac{d^2}{dx^2}$ has positive eigenvalues respectively, and then we will obtain sufficient conditions for building switching controls. To do this we introduce a new functional based on the adjoint system whose minimizers yield the switching controls. We show that, due to the time analyticity of solutions, under suitable conditions on the location of the controllers, switching control strategies exist in the 1-d heat equation under Robin's boundary condition.

2. Boundary Controls

Consider the heat equation in the space interval $(0, 1)$ with two controls located at the extremes $x = 0, 1$ and satisfying Robin's boundary condition (RBC)

$$\begin{cases} y_t - y_{xx} = 0, & 0 < x < 1, 0 < t < T, \\ y_x(0,t) - a_0 y(0,t) = u_0(t), & 0 < t < T, \\ y_x(1,t) + a_1 y(1,t) = u_1(t), & 0 < t < T, \\ y(x,0) = y^0(x), & 0 < x < 1. \end{cases} \tag{2.1}$$

We consider the problem of null controllability. More precisely, given an initial datum $y_0 \in L^2(0, 1)$ we look for controls $u_0(t), u_1(t) \in L^2(0, T)$ such that $y(x, T) = 0$ and satisfying switching property

$$u_0(t)u_1(t) = 0, \qquad \text{a.e.} \quad t \in (0, T). \tag{2.2}$$

It is worth to mention the fact that whenever a system is controllable, the control can be constructed by minimizing a suitable quadratic functional defined on the class of solutions of the adjoint system (see e.g. [1], [3], [4]).

For φ^0 in $L^2(0, 1)$, we consider the solution $\varphi : [0, 1] \times [0, T] \to C([0, T], L^2(0, 1))$, of the following backward Cauchy linear problem[†]

$$\begin{cases} \varphi_t + \varphi_{xx} = 0, & 0 < x < 1, 0 < t < T, \\ \varphi_x(0,t) - a_0 \varphi(0,t) = 0, & 0 < t < T, \\ \varphi_x(1,t) + a_1 \varphi(1,t) = 0, & 0 < t < T, \\ \varphi(x,T) = \varphi^0(x), & 0 < x < 1. \end{cases} \tag{2.3}$$

This linear system is called the adjoint system corresponding to the 1-d heat equation with Robin's boundary condition. We know that (see e.g. [2]) the Fourier representation of solutions of the adjoint system with positive eigenvalues are of the form:

$$\varphi(x,t) = \sum_{k \geq 1} \beta_k e^{\mu_k^2 (t-T)} \omega_k(x) \tag{2.4}$$

[†]The class of solutions φ of backward Cauchy linear problem is class of smoother functions.

where

$$\omega_k(x) = \cos \mu_k x + \frac{a_0}{\mu_k} \sin \mu_k x. \tag{2.5}$$

Now, for obtaining switching controls, we consider the following quadratic functional (see, e.g. [1]),

$$J_s^\alpha(\varphi^0) = \frac{1}{2}\int_0^T \max\left\{|\varphi(0,t)|^2, |\alpha\varphi(1,t)|^2\right\}dt + \int_0^1 y^0(x)\varphi(x,0)dx \tag{2.6}$$

where $\alpha \in \mathbb{R}$ and minimize (2.6) over the class $\mathcal{H}$ of initial data given by

$$\mathcal{H} = \{\varphi^0 : \int_0^T \Big[|\varphi(0,t)|^2 + |\varphi(1,t)|^2\Big]dt < \infty\}$$

where $\varphi(x,t)$ is the solution of the adjoint system (2.3) associated to the final state φ^0. This space endowed with the canonical norm

$$||\varphi^0||_{\mathcal{H}} = \Big[\int_0^T |\varphi(0,t)|^2 + |\varphi(1,t)|^2 dt\Big]^{\frac{1}{2}}.$$

constitutes a Hilbert space. Let us analyse the positivity of the norm $||\varphi^0||_{\mathcal{H}}$ in space $\mathcal{H}$. Here we will use (2.4), (2.5) as a Fourier representation of the solution of (2.3). Therefore we have

$$\int_0^T |\varphi(0,t)|^2 + |\varphi(1,t)|^2 dt \geq \int_0^T |\varphi(0,t)|^2 dt = \int_0^T \Big|\sum_{k\geq 1} \beta_k e^{\mu_k^2(t-T)}\Big|^2 dt \tag{2.7}$$

Observe that

$$\begin{aligned} ||\varphi(x,0)||^2_{L^2(0,1)} &= \int_0^1 \varphi^2(x,0)dx = \int_0^1 \Big|\sum_{k\geq 1} \beta_k e^{-\mu_k^2 T}\omega_k(x)\Big|^2 dx \\ &\leq C\sum_{k\geq 1} \beta_k^2 e^{-2k^2 T} \end{aligned} \tag{2.8}$$

where $C > 0$, which is independent from $\{\beta_k\}_{k\geq 1}$.
Now, we will give very important lemma on families of real exponentials. This lemma is known as estimates on families of real exponentials (see, e.g. [1],[4]).

2.1. Lemma. *In our case, it is guaranteed that*

$$\int_0^T \Big|\sum_{k\geq 1} \beta_k e^{k^2(t-T)}\Big|^2 dt \geq c_1 \sum_{k\geq 1} e^{-2k^2 T}\beta_k^2$$

for suitable positive constants $c_1 > 0$ which is independent from $\{\beta_k\}_{k\geq 1}$.

By using Lemma 2.1 in (2.7) and comparing with inequality (2.8), we will have the following observability inequality:

$$||\varphi(x,0)||^2_{L^2(0,1)} \leq \hat{C}\int_0^T |\varphi(0,t)|^2 + |\varphi(1,t)|^2 dt \tag{2.9}$$

for positive constant $\hat{C} > 0$ which is independent from $\{\beta_k\}_{k\geq 1}$.
The functional $J_s^\alpha : \mathcal{H} \longrightarrow \mathbb{R}$ is well defined, continuous, and strictly convex.‡ For checking the coercivity property, one should prove that

$$\lim_{||\varphi^0||_{L^2(0,1)}\to\infty} \frac{J_s^\alpha(\varphi^0)}{||\varphi^0||_{L^2(0,1)}} \geq \epsilon.$$

‡This comes from the observability inequality (2.9)

For this to be true, the unique continuation property of the adjoint system (2.3) suffices (see e.g. [1]). Namely,

$$\mu\{t \in (0,T) : |\varphi(0,t)| = |\varphi(1,t)|\} > 0 \Rightarrow \varphi \equiv 0. \tag{2.10}$$

2.2. Lemma. *Assume that*

$$|\alpha| \neq \left[\frac{\mu_k^2 + a_1^2}{\mu_k^2 + a_0^2}\right]^{\frac{1}{2}}, \qquad \forall k \geq 1, \tag{2.11}$$

holds. Then, (2.10) satisfy for solution of the adjoint system (2.3).

Proof. Firstly, assume that $\mu\{t \in (0,T) : |\varphi(0,t)| = |\varphi(1,t)|\} > 0$. We show that under the assumption of (2.11), we have $\varphi \equiv 0$. We know that (see e.g. [2]) positive eigenvalues $\{\mu_k\}_{k\geq 1}$ of adjoint system (2.3) satisfy eigenvalue equation:

$$\tan(\mu_k) = \frac{(a_0 + a_1)\mu_k}{\mu_k^2 - a_0 a_1} \tag{2.12}$$

and also we would have

$$(k-1)\pi < \mu_k < k\pi \quad \text{and} \quad \lim_{k\to\infty} \mu_k - (k-1)\pi = 0, \quad (k = 1, 2, 3, 4, ...).$$

Now assume that $\mu(I) > 0$, using again the Fourier representation of solution (2.4) of (2.3), we have

$$\varphi(0,t) \pm \alpha\varphi(1,t) = \sum_{k\geq 1} \beta_k e^{\mu_k^2(t-T)}(1 \pm \alpha(\cos\mu_k + \frac{a_0}{\mu_k}\sin\mu_k)).$$

The function $\varphi(0,t) \pm \alpha\varphi(1,t)$ is time analytic for $t \leq T$. Consequently, if they vanish for a set of time instants of positive measure, then they vanish for all $t \leq T$. It is then easy to see, by multiplying above identity by the real exponentials $e^{-\eta^2(t-T)}$ successively, starting from $\eta = 1$ and taking limits as $t \to -\infty$ that

$$\beta_k(1 \pm \alpha(\cos\mu_k + \frac{a_0}{\mu_k}\sin\mu_k)) = 0, \qquad \forall k \geq 1.$$

To conclude that $\beta_k = 0$ for all $k \geq 1$, it is sufficient to have that

$$1 \pm \alpha(\cos\mu_k + \frac{a_0}{\mu_k}\sin\mu_k) \neq 0 \tag{2.13}$$

Assume the converse of (2.13), then we have

$$\begin{aligned} \alpha(\cos\mu_k + \frac{a_0}{\mu_k}\sin\mu_k) \pm 1 = 0 \qquad &\iff [\alpha(\cos\mu_k + \frac{a_0}{\mu_k}\sin\mu_k)]^2 = 1 \\ &\iff \alpha^2(1 + 2\frac{a_0}{\mu_k}\tan\mu_k + \frac{a_0^2}{\mu_k^2}\tan^2\mu_k) = 1 + \tan^2\mu_k \end{aligned}$$

Now using eigenvalue equation (2.12) and after some simplification, finally we obtain the following

$$\alpha^2(\mu_k^2 + a_0^2) = (\mu_k^2 + a_1^2).$$

Therefore to obtain the unique continuation property, it is suffice to assume that

$$|\alpha| \neq \left[\frac{\mu_k^2 + a_1^2}{\mu_k^2 + a_0^2}\right]^{\frac{1}{2}}, \qquad \forall k \geq 1.$$

□

Therefore, by using Lemma 2.2, we have that J_s^α admits an unique minimizer $\hat{\varphi}^0 \in \mathcal{H}$. As a result, by using variational approach, we find our switching controls:

$$u_0(t) = -\hat{\varphi}(0,t)1_{S_0}, \qquad u_1(t) = \alpha^2\hat{\varphi}(1,t)1_{S_1} \qquad \text{for } t \in (0,T) \tag{2.14}$$

where

$$S_0 = \{t \in (0,T) : |\varphi(0,t)| > |\alpha\varphi(1,t)|\}$$
$$S_1 = \{t \in (0,T) : |\alpha\varphi(1,t)| > |\varphi(0,t)|\}.$$

At the end, we obtain the following new result

2.3. Theorem. *Given 1-d heat equation (2.1) under Robin's boundary condition, when the operator $-\frac{d^2}{dx^2}$ has positive eigenvalues, to obtain null controls that satisfying switching property (2.2), it is sufficient to assume that α satisfies the following relation*

$$|\alpha| \neq \Big[\frac{\mu_k^2 + a_1^2}{\mu_k^2 + a_0^2}\Big]^{\frac{1}{2}}, \qquad \forall k \geq 1.$$

3. Lumped Controls

Let $f_0(x)$ and $f_1(x)$ be control profiles in $L^2(0,1)$. Consider the following heat equation

$$\begin{cases} y_t - y_{xx} = u_0(t)f_0(x) + u_1(t)f_1(x), & 0 < x < 1, 0 < t < T, \\ y_x(0,t) - a_0 y(0,t) = 0, & 0 < t < T, \\ y_x(1,t) + a_1 y(1,t) = 0, & 0 < t < T, \\ y(x,0) = y^0(x), & 0 < x < 1. \end{cases} \tag{3.1}$$

Here, we would consider the same problem, i.e., given an initial datum $y_0 \in L^2(0,1)$ we are looking for controls $u_0(t), u_1(t) \in L^2(0,T)$ such that null controllability of heat equation holds, i.e, $y(x,T) = 0$ and switching condition satisfies:

$$u_0(t)u_1(t) = 0, \qquad \text{a.e.} \quad t \in (0,T) \tag{3.2}$$

The null control of 1-d heat equation may be computed by minimizing the quadratic functional (see e.g. [1]),

$$\begin{aligned} \hat{J}_s(\varphi^0) \quad &= \quad \frac{1}{2}\int_0^T \max\Big[\Big|\int_0^1 f_0(x)\varphi(x,t)dx\Big|^2, \Big|\int_0^1 f_1(x)\varphi(x,t)dx\Big|^2\Big]dt \\ &\quad - \int_0^T y^0(x)\varphi(x,0)dx \end{aligned}$$

over the class $\tilde{\mathcal{H}}$ of initial data given by

$$\tilde{\mathcal{H}} = \{\varphi^0 : \int_0^T \Big[\Big|\int_0^1 f_0(x)\varphi(x,t)dx\Big|^2 + \Big|\int_0^1 f_1(x)\varphi(x,t)dx\Big|^2\Big]dt < \infty\}$$

which endowed with the canonical norm

$$||\varphi^0||^2_{\tilde{\mathcal{H}}} = \int_0^T \Big[\Big|\int_0^1 f_0(x)\varphi(x,t)dx\Big|^2 + \Big|\int_0^1 f_1(x)\varphi(x,t)dx\Big|^2\Big]dt$$

At first, we will show that $||\cdot||_{\tilde{\mathcal{H}}}$ actually defines norm on $\tilde{\mathcal{H}}$. For this, it is enough to show the positivity of $||\cdot||_{\tilde{\mathcal{H}}}$. Observe that

$$||\varphi^0||^2_{\tilde{\mathcal{H}}} = \int_0^T \Big[\Big|\sum_{k\geq 1}\beta_k e^{\mu_k^2(t-T)}f_{0,k}\Big|^2 + \Big|\sum_{k\geq 1}\beta_k e^{\mu_k^2(tT)}f_{1,k}\Big|^2\Big]dt$$

where

$$f_0(x) = \sum_{k\geq 1} f_{0,k}\omega_k(x), \qquad f_1(x) = \sum_{k\geq 1} f_{1,k}\omega_k(x),$$

and using Lemma 2.1, we then get the following weighted observability inequality

$$||\varphi^0||^2_{\tilde{\mathcal{H}}} \geq c_1 \sum_{k\geq 1} e^{-2\mu_k^2 T}\Big[|f_{0,k}|^2 + |f_{1,k}|^2\Big]\beta_k^2 \tag{3.3}$$

where positive constant c_1 is independent from $\{\beta_k\}_{k\geq 1}$.
In addition, since (3.1) is well posed, the functional $\hat{J}(\varphi^0)$ is obviously continuous in $\tilde{\mathcal{H}}$, the convexity (strictly) of $\hat{J}(\varphi^0)$ comes from the weighted observability inequality (3.3). As we know that (see e.g. [3]) null controllability in time T implies (finite) approximate controllability in time T. This comes form the fact that all the range of the semi-group generated by the heat equation is reachable. Therefore, we first prove the approximate controllability of the heat system in time T under some conditions. For this, we will construct the new functional very similar with previous one $\hat{J}_s$ and with the same coercivity property, allows building approximate controllers: for any $\epsilon > 0$ and any $y^1 \in L^2(0,1)$

$$\begin{aligned}\hat{J}_\epsilon(\varphi^0) \quad &= \quad \frac{1}{2}\int_0^T \max\Big[\Big|\int_0^1 f_0(x)\varphi dx\Big|^2, \Big|\int_0^1 f_1(x)\varphi dx\Big|^2\Big]dt \\ &\quad + \epsilon||(I-\pi_E)\varphi^0||_{L^2(0,1)} + \int_0^1 \varphi^0 y^1 dx - \int_0^1 y^0(x)\varphi(x,0)dx\end{aligned}$$

where E is finite dimensional subspace of $L^2(0,1)$ and π_E denotes the ortogonal projection from $L^2(0,1)$ over E.

Our aim is to build approximate lumped controls that satisfy switching property (3.2). In other words, given $\epsilon > 0$, we try to find (finite) approximate controls u_0^ϵ and u_1^ϵ such that the solution y_ϵ of heat equation satisfies the condition

$$||y_\epsilon(x,T)||_{L^2(0,1)} \leq \epsilon.$$

3.1. Lemma. *Assume that the following unique continuation property*

$$\mu\Big\{t \in (0,T) : \Big|\int_0^1 f_0(x)\varphi(x,t)dx\Big| = \Big|\int_0^1 f_1(x)\varphi(x,t)dx\Big|\Big\} > 0 \Rightarrow \varphi \equiv 0. \tag{3.4}$$

holds. Then the heat system (3.1) is approximate controllable.

For the proof of Lemma 3.1, one should first prove that the functional $\hat{J}_\epsilon$ is coercive in $\tilde{\mathcal{H}}$ which directly comes from the assumption (3.4) and at the end, by using variational approach, one could easily get approximate controls for (3.1) (see e.g. [3]). Therefore, from Lemma 3.1, to get approximate controls, we should prove (3.4). Using (2.4), we have

$$\int_0^1 f_0(x)\varphi(x,t)dx \pm \int_0^1 f_1(x)\varphi(x,t)dx = \sum_{k\geq 1}\beta_k e^{\mu_k^2(t-T)}(f_{1,k} \pm f_{0,k}).$$

The function $\int_0^1 \varphi(x,t)(f_0(x) \pm f_1(x))dx$ is time analytic for $t \leq T$. Consequently, if they vanish for a set of time instants of positive measure, then they vanish for all $t \leq T$. It is then easy to see, by multiplying above identity by the real exponentials $e^{-\eta^2(t-T)}$ successively, starting from $\eta = 1$ and taking limits as $t \to -\infty$, that

$$\beta_k(f_{1,k} \pm f_{0,k}) = 0, \qquad \forall k \geq 1.$$

To conclude that $\beta_k = 0$ for all $k \geq 1$, it is sufficient to assume that

$$f_{1,k} \pm f_{0,k} \neq 0 \qquad \forall k \geq 1. \tag{3.5}$$

Therefore under the condition (3.5), our functional $\hat{J}_\epsilon$ admits an unique minimizer $\hat{\varphi}^0 \in \tilde{\mathcal{H}}$. For every $\epsilon > 0$, by using variational approach, we will obtain approximate switching controls. We would like to say that for each $\epsilon > 0$, we must have the fact that $u_0^\epsilon(t)$ and $u_1^\epsilon(t)$ are uniformly bounded in $L^2(0,T)$. But under the condition on Fourier coefficients of the initial datum y^0

$$(3.6) \qquad \sum_{k\geq 1} \frac{e^{2\mu_k^2 T}}{|f_{1,k}|^2 + |f_{2,k}|^2} |y_k^0|^2 < \infty$$

being satisfied, by using weighted observability inequality (3.3) one could easily prove that $u_0^\epsilon(t)$ and $u_1^\epsilon(t)$ are uniformly bounded in $L^2(0,T)$. Hence at the end, by using variational approach we obtain the following switching controls

$$(3.7) \qquad u_0(t) = -\int_0^1 f_0(x)\hat{\varphi}(x,t)dx, \qquad u_1(t) = 0, \qquad \text{in } S_0$$

$$(3.8) \qquad u_1(t) = -\int_0^1 f_1(x)\hat{\varphi}(x,t)dx, \qquad u_0(t) = 0, \qquad \text{in } S_1$$

where

$$S_0 = \{t \in (0,T) : \Big| \int_0^1 f_0(x)\hat{\varphi}(x,t)dx\Big| > \Big| \int_0^1 f_1(x)\hat{\varphi}(x,t)dx\Big|\}$$
$$S_1 = \{t \in (0,T) : \Big| \int_0^1 f_1(x)\hat{\varphi}(x,t)dx\Big| > \Big| \int_0^1 f_0(x)\hat{\varphi}(x,t)dx\Big|\}.$$

In conclusion, we obtain the following new result

3.2. Theorem. *Assume that $f_0(x)$ and $f_1(x)$ are two control profiles in $L^2(0,1)$ and their Fourier coefficients satisfying (3.5). Let Fourier coefficients of the initial datum y^0 satisfy (3.6). Then, for all $T > 0$, there exist switching controls (3.7) and (3.8) satisfying our switching condition (3.2) and solution of (3.1) satisfies*

$$y(x,T) = 0,$$

i.e, null controllability is satisfied.

Acknowledgements. The author wishes to thank anonymous referee for inspiring comments.

References

[1] E. Zuazua, Switching control, *J. Eur. Math. Soc.* 13, pp. 85–117, 2011.
[2] Walter A. Strauss, *Partial Differential Equations: An Introduction*, Wiley, 1992.
[3] E. Zuazua, Controllability and observability of partial differential equations: Some results and open problems, *In: Handbook of Differential Equations: Evolutionary Equations*, Vol.3, pp. 527–621, 2006.
[4] E. Zuazua, S. Micu, On the controllability of a fractional order parabolic equation, *SIAM J. Control Optim.*, Vol.44, pp. 1950–1972, 2006.
[5] D.L. Russell, Controllability and stabilizability theory for linear partial differential equations. Resent progress and open questions, *SIAM Rev.*, Vol.20, pp. 639–739., 1978.
[6] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlang Berlin Heidelberg, 1971.
[7] J.L. Lions, Exact Controllability, Stabilizability and Perturbations for Distributed Systems", *SIAM Rev.*, Vol. 30, pp. 1–68, 1988.
[8] E.B. Lee, L. Markus, *Foundations of Optimal Control Theory*, John Wiley and Sons, Inc., New York, 1967.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 381 – 390

Hermite-Hadamard type inequalities for harmonically (α, m)-convex functions

İmdat İşcan*†

Abstract

The author introduces the concept of harmonically (α, m)-convex functions and establishes some Hermite-Hadamard type inequalities of these classes of functions.

Keywords: Harmonically (α, m)-convex function, Hermite-Hadamard type inequalities, Hypergeometric function.

2000 AMS Classification: 26D15; 26A51

Received : 23.05.2013 *Accepted* : 06.08.2014 *Doi* : 10.15672/HJMS.20164516901

1. Introduction

Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ be a convex function defined on the interval I of real numbers and $a, b \in I$ with $a < b$. The following double inequality is well known in the literature as Hermite-Hadamard integral inequality

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b f(x)dx \leq \frac{f(a)+f(b)}{2}. \tag{1.1}$$

The class of (α, m)-convex functions was first introduced In [8], and it is defined as follows:

1.1. Definition. The function $f : [0, b] \to \mathbb{R}$, $b > 0$, is said to be (α, m)-convex where $(\alpha, m) \in [0, 1]^2$, if we have

$$f(tx + m(1-t)y) \leq t^\alpha f(x) + m(1-t^\alpha)f(y)$$

for all $x, y \in [0, b]$ and $t \in [0, 1]$.

*Department of Mathematics, Faculty of Arts and Sciences,
Giresun University, 28100, Giresun, Turkey., Email: imdat.iscan@giresun.edu.tr
†Corresponding Author.

It can be easily that for $(\alpha, m) \in \{(0,0), (\alpha,0), (1,0), (1,m), (1,1), (\alpha,1)\}$ one obtains the following classes of functions: increasing, α-starshaped, starshaped, m-convex, convex, α-convex.

Denote by $K_m^\alpha(b)$ the set of all (α, m)-convex functions on $[0,b]$ for which

$f(0) \leq 0$. For recent results and generalizations concerning (α, m)-convex functions (see [2, 4, 5, 6, 8, 9, 10, 11, 12]).

In [7], the author gave definition of harmonically convex functions and established some Hermite-Hadamard type inequalities for harmonically convex functions as follows:

1.2. Definition. Let $I \subset \mathbb{R} \backslash \{0\}$ be a real interval. A function $f : I \to \mathbb{R}$ is said to be harmonically convex, if

$$f\left(\frac{xy}{tx + (1-t)y}\right) \leq tf(y) + (1-t)f(x) \tag{1.2}$$

for all $x, y \in I$ and $t \in [0,1]$. If the inequality in (1.2) is reversed, then f is said to be harmonically concave.

1.3. Theorem. *Let $f : I \subset \mathbb{R} \backslash \{0\} \to \mathbb{R}$ be a harmonically convex function and $a, b \in I$ with $a < b$. If $f \in L[a,b]$ then the following inequalities hold*

$$f\left(\frac{2ab}{a+b}\right) \leq \frac{ab}{b-a} \int_a^b \frac{f(x)}{x^2} dx \leq \frac{f(a)+f(b)}{2}. \tag{1.3}$$

The above inequalities are sharp.

1.4. Theorem. *Let $f : I \subset (0,\infty) \to \mathbb{R}$ be a differentiable function on I°, $a, b \in I$ with $a < b$, and $f' \in L[a,b]$. If $|f'|^q$ is harmonically convex on $[a,b]$ for $q \geq 1$, then*

$$\begin{aligned} &\left| \frac{f(a)+f(b)}{2} - \frac{ab}{b-a} \int_a^b \frac{f(x)}{x^2} dx \right| \\ &\leq \frac{ab\,(b-a)}{2} \lambda_1^{1-\frac{1}{q}} \left[\lambda_2 \left|f'(a)\right|^q + \lambda_3 \left|f'(b)\right|^q\right]^{\frac{1}{q}}, \end{aligned} \tag{1.4}$$

where

$$\begin{aligned} \lambda_1 &= \frac{1}{ab} - \frac{2}{(b-a)^2} \ln\left(\frac{(a+b)^2}{4ab}\right), \\ \lambda_2 &= \frac{-1}{b\,(b-a)} + \frac{3a+b}{(b-a)^3} \ln\left(\frac{(a+b)^2}{4ab}\right), \\ \lambda_3 &= \frac{1}{a\,(b-a)} - \frac{3b+a}{(b-a)^3} \ln\left(\frac{(a+b)^2}{4ab}\right) \\ &= \lambda_1 - \lambda_2. \end{aligned}$$

1.5. Theorem. *Let $f : I \subset (0,\infty) \to \mathbb{R}$ be a differentiable function on I°, $a, b \in I$ with $a < b$, and $f' \in L[a,b]$. If $|f'|^q$ is harmonically convex on $[a,b]$ for $q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$\begin{aligned} &\left| \frac{f(a)+f(b)}{2} - \frac{ab}{b-a} \int_a^b \frac{f(x)}{x^2} dx \right| \\ &\leq \frac{ab\,(b-a)}{2} \left(\frac{1}{p+1}\right)^{\frac{1}{p}} \left(\mu_1 \left|f'(a)\right|^q + \mu_2 \left|f'(b)\right|^q\right)^{\frac{1}{q}}, \end{aligned} \tag{1.5}$$

where

$$\begin{aligned}\mu_1 &= \frac{\left[a^{2-2q}+b^{1-2q}\left[(b-a)(1-2q)-a\right]\right]}{2(b-a)^2(1-q)(1-2q)},\\ \mu_2 &= \frac{\left[b^{2-2q}-a^{1-2q}\left[(b-a)(1-2q)+b\right]\right]}{2(b-a)^2(1-q)(1-2q)}.\end{aligned}$$

In [7], the author gave the following identity for differentiable functions.

1.6. Lemma. *Let $f : I \subset \mathbb{R}\backslash\{0\} \to \mathbb{R}$ be a differentiable function on I° and $a, b \in I$ with $a < b$. If $f' \in L[a,b]$ then*

$$\begin{aligned}&\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\\ &= \frac{ab(b-a)}{2}\int_0^1 \frac{1-2t}{(tb+(1-t)a)^2}f'\left(\frac{ab}{tb+(1-t)a}\right)dt.\end{aligned}$$

The main purpose of this paper is to introduce the concept of harmonically (α, m)-convex functions and establish some new Hermite-Hadamard type inequalities for these classes of functions.

2. Main Results

2.1. Definition. The function $f : (0, b^*] \to \mathbb{R}$, $b^* > 0$, is said to be harmonically (α, m)-convex, where $\alpha \in [0,1]$ and $m \in (0,1]$, if

$$f\left(\frac{mxy}{mty+(1-t)x}\right) = f\left(\left(\frac{t}{x}+\frac{1-t}{my}\right)^{-1}\right) \le t^\alpha f(x) + m(1-t^\alpha)f(y) \tag{2.1}$$

for all $x, y \in (0, b^*]$ and $t \in [0,1]$. If the inequality in (2.1) is reversed, then f is said to be harmonically (α, m)-concave.

2.2. Remark. When $m = \alpha = 1$, the harmonically (α, m)-convex (concave) function defined in Definition 2.1 becomes a harmonically convex (concave) function defined in [7]. Thus, every harmonically convex (concave) function is also harmonically $(1,1)$-convex (concave) function.

The following proposition is obvious.

2.3. Proposition. *Let $f : (0, b^*] \to \mathbb{R}$ be a function.*

a) *if f is (α, m)-convex and nondecreasing function then f is harmonically (α, m)-convex.*
b) *if f is harmonically (α, m)-convex and nonincreasing function then f is (α, m)-convex.*

Proof. For all $t \in [0,1]$, $m \in (0,1]$ and $x, y \in (0, b^*]$ we have

$$t(1-t)(x-my)^2 \ge 0,$$

then the following inequality holds

$$\frac{mxy}{mty+(1-t)x} \le tx + m(1-t)y. \tag{2.2}$$

By the inequality (2.2), the proof is completed. □

2.4. Remark. According to Proposition 2.3, every nondecreasing s-convex function in the first sense (or $(s,1)$-convex function) is also harmonically $(s,1)$-convex function.

2.5. Example. Let $s \in (0,1]$, then the function $f:(0,\infty)\to\mathbb{R}$, $f(x)=x^s$ is a nondecreasing s-convex function in the first sense [3]. According to the above Remark, f is also harmonically $(s,1)$-convex function.

The following result of the Hermite-Hadamard type holds.

2.6. Theorem. *Let $f:(0,\infty)\to\mathbb{R}$ be a harmonically (α,m)-convex function with $\alpha\in[0,1]$ and $m\in(0,1]$. If $0<a<b<\infty$ and $f\in L[a,b]$, then one has the inequality*

$$\frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx \le \min\left\{\frac{f(a)+\alpha m f(\frac{b}{m})}{\alpha+1}, \frac{f(b)+\alpha m f(\frac{a}{m})}{\alpha+1}\right\}. \tag{2.3}$$

Proof. Since $f:(0,\infty)\to\mathbb{R}$ is a harmonically (α,m)-convex function, we have, for all $x,y\in I$

$$f\left(\frac{xy}{ty+(1-t)x}\right) = f\left(\frac{m\frac{y}{m}x}{tm\frac{y}{m}+(1-t)x}\right) \le t^\alpha f(x) + m(1-t^\alpha)f(\frac{y}{m})$$

which gives:

$$f\left(\frac{ab}{tb+(1-t)a}\right) \le t^\alpha f(a) + m(1-t^\alpha)f(\frac{b}{m})$$

and

$$f\left(\frac{ab}{ta+(1-t)b}\right) \le t^\alpha f(b) + m(1-t^\alpha)f(\frac{a}{m})$$

for all $t\in[0,1]$. Integrating on $[0,1]$ we obtain

$$\int_0^1 f\left(\frac{ab}{tb+(1-t)a}\right)dt \le \frac{f(a)+\alpha m f(\frac{b}{m})}{\alpha+1}$$

and

$$\int_0^1 f\left(\frac{ab}{ta+(1-t)b}\right)dt \le \frac{f(b)+\alpha m f(\frac{a}{m})}{\alpha+1}.$$

However,

$$\int_0^1 f\left(\frac{ab}{tb+(1-t)a}\right)dt = \int_0^1 f\left(\frac{ab}{ta+(1-t)b}\right)dt = \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx$$

and the inequality (2.3) is obtained. □

2.7. Remark. If we take $\alpha=m=1$ in Theorem 2.6, then inequality (2.3) becomes the right-hand side of inequality (1.3).

2.8. Corollary. *If we take $m=1$ in Theorem 2.6, then we get*

$$\frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx \le \min\left\{\frac{f(a)+\alpha f(b)}{\alpha+1}, \frac{f(b)+\alpha f(a)}{\alpha+1}\right\} \tag{2.4}$$

2.9. Theorem. *Let $f:I\subset(0,\infty)\to\mathbb{R}$ be a differentiable function on I°, $a,b/m\in I^\circ$ with $a<b$, $m\in(0,1]$ and $f'\in L[a,b]$. If $|f'|^q$ is harmonically (α,m)-convex on $[a,b/m]$ for $q\ge1$, with $\alpha\in[0,1]$, then*

$$\left|\frac{f(a)+f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right|$$

$$\leq \frac{ab\,(b-a)}{2^{2-1/q}}\left[\lambda(\alpha,q;a,b)\left|f'\left(a\right)\right|^{q}+m\mu(\alpha,q;a,b)\left|f'\left(b/m\right)\right|^{q}\right]^{\frac{1}{q}},$$

where

$$\begin{aligned}
\lambda(\alpha,q;a,b) &= \frac{\beta\left(1,\alpha+2\right)}{b^{2q}}.{}_{2}F_{1}\left(2q,1;\alpha+3;1-\frac{a}{b}\right) \\
&\quad -\frac{\beta\left(2,\alpha+1\right)}{b^{2q}}.{}_{2}F_{1}\left(2q,2;\alpha+3;1-\frac{a}{b}\right) \\
&\quad +\frac{2^{2q-\alpha}\beta\left(2,\alpha+1\right)}{\left(a+b\right)^{2q}}.{}_{2}F_{1}\left(2q,2;\alpha+3;1-\frac{2a}{a+b}\right), \\
\mu(\alpha,q;a,b) &= \lambda(0,q;a,b)-\lambda(\alpha,q;a,b),
\end{aligned}$$

β is Euler Beta function defined by

$$\beta\left(x,y\right)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}=\int_{0}^{1}t^{x-1}\left(1-t\right)^{y-1}dt,\quad x,y>0,$$

and ${}_{2}F_{1}$ is hypergeometric function defined by

$${}_{2}F_{1}\left(a,b;c;z\right)=\frac{1}{\beta\left(b,c-b\right)}\int_{0}^{1}t^{b-1}\left(1-t\right)^{c-b-1}\left(1-zt\right)^{-a}dt,\ c>b>0,\ |z|<1$$

(see [1]*).*

Proof. From Lemma 1.6 and using the power mean inequality, we have

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_{a}^{b}\frac{f(x)}{x^{2}}dx\right| \\
\leq\ & \frac{ab\,(b-a)}{2}\int_{0}^{1}\left|\frac{1-2t}{\left(tb+(1-t)a\right)^{2}}\right|\left|f'\left(\frac{ab}{tb+(1-t)a}\right)\right|dt \\
\leq\ & \frac{ab\,(b-a)}{2}\left(\int_{0}^{1}\left|1-2t\right|dt\right)^{1-\frac{1}{q}} \\
&\times\left(\int_{0}^{1}\frac{\left|1-2t\right|}{\left(tb+(1-t)a\right)^{2q}}\left|f'\left(\frac{ab}{tb+(1-t)a}\right)\right|^{q}dt\right)^{\frac{1}{q}}.
\end{aligned}$$

Hence, by harmonically (α,m)-convexity of $|f'|^{q}$ on $[a,b/m]$,we have

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_{a}^{b}\frac{f(x)}{x^{2}}dx\right| \\
\leq\ & \frac{ab\,(b-a)}{2}\left(\frac{1}{2}\right)^{1-\frac{1}{q}}\left(\int_{0}^{1}\frac{\left|1-2t\right|\left[t^{\alpha}\left|f'\left(a\right)\right|^{q}+m(1-t^{\alpha})\left|f'\left(b/m\right)\right|^{q}\right]}{\left(tb+(1-t)a\right)^{2q}}dt\right)^{\frac{1}{q}} \\
\leq\ & \frac{ab\,(b-a)}{2^{2-1/q}}\left[\lambda(\alpha,q;a,b)\left|f'\left(a\right)\right|^{q}+m\left(\lambda(0,q;a,b)-\lambda(\alpha,q;a,b)\right)\left|f'\left(b/m\right)\right|^{q}\right]^{\frac{1}{q}}.
\end{aligned}$$

It is easily check that

$$\int_0^1 \frac{|1-2t|\,t^{\alpha}}{(tb+(1-t)a)^{2q}}dt = 2\int_0^{1/2} \frac{(1-2t)\,t^{\alpha}}{(tb+(1-t)a)^{2q}}dt - \int_0^1 \frac{(1-2t)\,t^{\alpha}}{(tb+(1-t)a)^{2q}}dt$$

$$= \frac{\beta(1,\alpha+2)}{b^{2q}} \cdot {}_2F_1\left(2q,1;\alpha+3;1-\frac{a}{b}\right) - \frac{\beta(2,\alpha+1)}{b^{2q}} \cdot {}_2F_1\left(2q,2;\alpha+3;1-\frac{a}{b}\right)$$
$$+\frac{2^{2q-\alpha}\beta(2,\alpha+1)}{(a+b)^{2q}} \cdot {}_2F_1\left(2q,2;\alpha+3;1-\frac{2a}{a+b}\right) = \lambda(\alpha,q;a,b),$$

$$\int_0^1 \frac{|1-2t|\,(1-t^{\alpha})}{(tb+(1-t)a)^{2q}}dt = \lambda(0,q;a,b) - \lambda(\alpha,q;a,b).$$

This completes the proof. □

If we take $\alpha = m = 1$ in Theorem 2.9 then we get the following a new corrollary for harmonically convex functions:

2.10. Corollary. *Let $f : I \subset (0,\infty) \to \mathbb{R}$ be a differentiable function on I°, $a, b \in I^{\circ}$ with $a < b$ and $f' \in L[a,b]$. If $|f'|^q$ is harmonically convex on $[a,b]$ for $q \geq 1$ then*

$$\left|\frac{f(a)+f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right|$$

$$\leq \frac{ab\,(b-a)}{2^{2-1/q}}\left[\lambda(1,q;a,b)\left|f'(a)\right|^q + \mu(1,q;a,b)\left|f'(b/m)\right|^q\right]^{\frac{1}{q}}.$$

2.11. Corollary. *If we take $m = 1$ in Theorem 2.9 then we get*

$$\left|\frac{f(a)+f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right| \tag{2.5}$$

$$\leq \frac{ab\,(b-a)}{2^{2-1/q}}\left[\lambda(\alpha,q;a,b)\left|f'(a)\right|^q + \mu(\alpha,q;a,b)\left|f'(b)\right|^q\right]^{\frac{1}{q}}.$$

2.12. Theorem. *Let $f : I \subset (0,\infty) \to \mathbb{R}$ be a differentiable function on I°, $a, b/m \in I^{\circ}$ with $a < b$, $m \in (0,1]$ and $f' \in L[a,b]$. If $|f'|^q$ is harmonically (α,m)-convex on $[a,b/m]$ for $q \geq 1$, with $\alpha \in [0,1]$, then*

$$\left|\frac{f(a)+f(b)}{2} - \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right| \leq \frac{ab\,(b-a)}{2} \tag{2.6}$$

$$\times\lambda^{1-\frac{1}{q}}(0,q;a,b)\left[\lambda(\alpha,1;a,b)\left|f'(a)\right|^q + m\mu(\alpha,1;a,b)\left|f'(b/m)\right|^q\right]^{\frac{1}{q}},$$

where λ and μ is defined as in Theorem 2.9.

Proof. From Lemma 1.6, power mean inequality and the harmonically (α, m)-convexity of $|f'|^q$ on $[a, b/m]$,we have,

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right| \\
\le\ & \frac{ab\,(b-a)}{2}\int_0^1 \left|\frac{1-2t}{(tb+(1-t)a)^2}\right|\left|f'\left(\frac{ab}{tb+(1-t)a}\right)\right|dt \\
\le\ & \frac{ab\,(b-a)}{2}\left(\int_0^1 \left|\frac{1-2t}{(tb+(1-t)a)^2}\right|dt\right)^{1-\frac{1}{q}} \\
& \times\left(\int_0^1 \frac{|1-2t|\left[t^\alpha\left|f'\left(a\right)\right|^q+m(1-t^\alpha)\left|f'\left(b/m\right)\right|^q\right]}{(tb+(1-t)a)^2}dt\right)^{\frac{1}{q}} \\
\le\ & \frac{ab\,(b-a)}{2}\lambda^{1-\frac{1}{q}}(0,q;a,b)\left[\lambda(\alpha,1;a,b)\left|f'\left(a\right)\right|^q+m\mu(\alpha,1;a,b)\left|f'\left(b/m\right)\right|^q\right]^{\frac{1}{q}}.
\end{aligned}$$

□

2.13. Remark. If we take $\alpha = m = 1$ in Theorem 2.12 then inequality (2.6) becomes inequality (1.4) of Theorem 1.4.

2.14. Corollary. *If we take $m = 1$ in Theorem 2.12 then we get*

$$\left|\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right| \le \frac{ab\,(b-a)}{2} \tag{2.7}$$

$$\times\lambda^{1-\frac{1}{q}}(0,q;a,b)\left[\lambda(\alpha,1;a,b)\left|f'\left(a\right)\right|^q+\mu(\alpha,1;a,b)\left|f'\left(b\right)\right|^q\right]^{\frac{1}{q}},$$

2.15. Theorem. *Let $f : I \subset (0,\infty) \to \mathbb{R}$ be a differentiable function on I°, $a, b/m \in I^\circ$ with $a < b$, $m \in (0, 1]$, and $f' \in L[a, b]$. If $|f'|^q$ is harmonically (α, m)-convex on $[a, b/m]$ for $q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$, with $\alpha \in [0, 1]$, then*

$$\left|\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}dx\right| \le \frac{ab\,(b-a)}{2}\left(\frac{1}{p+1}\right)^{\frac{1}{p}} \tag{2.8}$$

$$\times\left(\nu(\alpha,q;a,b)\left|f'\left(a\right)\right|^q+m(\nu(0,q;a,b)-\nu(\alpha,q;a,b))\left|f'\left(b/m\right)\right|^q\right)^{\frac{1}{q}}$$

where

$$\nu(\alpha,q;a,b)=\frac{\beta\left(1,\alpha+1\right)}{b^{2q}}.{}_2F_1\left(2q,1;\alpha+2;1-\frac{a}{b}\right).$$

Proof. From Lemma 1.6, Hölder's inequality and the harmonically (α, m)-convexity of $|f'|^q$ on $[a, b/m]$,we have,

$$\begin{aligned}
&\left|\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_a^b\frac{f(x)}{x^2}dx\right| \\
\leq\ & \frac{ab\,(b-a)}{2}\left(\int_0^1|1-2t|^p\,dt\right)^{\frac{1}{p}} \\
&\times\left(\int_0^1\frac{1}{(tb+(1-t)a)^{2q}}\left|f'\left(\frac{ab}{tb+(1-t)a}\right)\right|^q dt\right)^{\frac{1}{q}}
\end{aligned}$$

$$\begin{aligned}
\leq\ & \frac{ab\,(b-a)}{2}\left(\frac{1}{p+1}\right)^{\frac{1}{p}} \\
&\times\left(\int_0^1\frac{t^\alpha\,|f'\,(a)|^q+m(1-t^\alpha)\,|f'\,(b/m)|^q}{(tb+(1-t)a)^{2q}}dt\right)^{\frac{1}{q}} \\
\leq\ & \frac{ab\,(b-a)}{2}\left(\frac{1}{p+1}\right)^{\frac{1}{p}} \\
&\times\left(\nu(\alpha,q;a,b)\left|f'\,(a)\right|^q+m(\nu(0,q;a,b)-\nu(\alpha,q;a,b))\left|f'\,(b/m)\right|^q\right)^{\frac{1}{q}},
\end{aligned}$$

where an easy calculation gives

$$\begin{aligned}
&\int_0^1\frac{t^\alpha}{(tb+(1-t)a)^{2q}}dt \\
=\ & \frac{\beta\,(1,\alpha+1)}{b^{2q}}\cdot{}_2F_1\left(2q,1;\alpha+2;1-\frac{a}{b}\right) \\
=\ & \nu(\alpha,q;a,b)
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^1\frac{1-t^\alpha}{(tb+(1-t)a)^{2q}}dt \\
=\ & \nu(0,q;a,b)-\nu(\alpha,q;a,b).
\end{aligned}$$

This completes the proof. □

2.16. Remark. If we take $\alpha = m = 1$ in Theorem 2.15 then inequality (2.8) becomes inequality (1.5) of Theorem 1.5.

2.17. Corollary. *If we take $m = 1$ in Theorem 2.15 then we get*

$$\left|\frac{f(a)+f(b)}{2}-\frac{ab}{b-a}\int_a^b\frac{f(x)}{x^2}dx\right|\leq\frac{ab\,(b-a)}{2}\left(\frac{1}{p+1}\right)^{\frac{1}{p}} \tag{2.9}$$

$$\times\left(\nu(\alpha,q;a,b)\left|f'\,(a)\right|^q+(\nu(0,q;a,b)-\nu(\alpha,q;a,b))\left|f'\,(b)\right|^q\right)^{\frac{1}{q}}.$$

3. Some applications for special means

Let us recall the following special means of two nonnegative number a, b with $b > a$:

(1) The weighted arithmetic mean

$$A_\alpha (a,b) := \alpha a + (1-\alpha)b,\ \alpha \in [0,1].$$

(2) The arithmetic mean

$$A = A(a,b) := \frac{a+b}{2}.$$

(3) The geometric mean

$$G = G(a,b) := \sqrt{ab}.$$

(4) The harmonic mean

$$H = H(a,b) := \frac{2ab}{a+b}.$$

(5) The p-Logarithmic mean

$$L_p = L_p(a,b) := \left(\frac{b^{p+1}-a^{p+1}}{(p+1)(b-a)}\right)^{\frac{1}{p}},\quad p \in \mathbb{R}\backslash\{-1,0\}.$$

3.1. Proposition. *Let $0 < a < b$. Then we have the following inequality*

$$G^2 L_{\alpha-2}^{\alpha-2} \leq \min\left\{A_{1/(\alpha+1)}(a^\alpha, b^\alpha), A_{1/(\alpha+1)}(b^\alpha, a^\alpha)\right\}.$$

Proof. The assertion follows from the inequality (2.4) in Corollary 2.8, for $f:(0,\infty)\to\mathbb{R},\ f(x)=x^\alpha,\ 0<\alpha<1.$ □

3.2. Proposition. *Let $0 < a < b$, $q \geq 1$ and $0 < \alpha < 1$. Then we have the following inequality*

$$\left|A\left(a^{\frac{\alpha}{q}+1}, b^{\frac{\alpha}{q}+1}\right) - G^2 L_{\frac{\alpha}{q}-1}^{\frac{\alpha}{q}-1}\right|$$
$$\leq \frac{ab(b-a)(\alpha+q)}{q2^{2-1/q}}\left[\lambda(\alpha,q;a,b)a^\alpha + \mu(\alpha,q;a,b)b^\alpha\right]^{\frac{1}{q}}.$$

Proof. The assertion follows from the inequality (2.5) in Corollary 2.11, for $f:(0,\infty)\to\mathbb{R},\ f(x)=x^{\frac{\alpha}{q}+1}/\left(\frac{\alpha}{q}+1\right).$ □

3.3. Proposition. *Let $0 < a < b$, $q \geq 1$ and $0 < \alpha < 1$. Then we have the following inequality*

$$\left|A\left(a^{\frac{\alpha}{q}+1}, b^{\frac{\alpha}{q}+1}\right) - G^2 L_{\frac{\alpha}{q}-1}^{\frac{\alpha}{q}-1}\right|$$
$$\leq \frac{ab(b-a)(\alpha+q)}{2q}\lambda^{1-\frac{1}{q}}(0,q;a,b)\left[\lambda(\alpha,1;a,b)a^\alpha + \mu(\alpha,1;a,b)b^\alpha\right]^{\frac{1}{q}},$$

Proof. The assertion follows from the inequality (2.7) in Corollary 2.14, for $f:(0,\infty)\to\mathbb{R},\ f(x)=x^{\frac{\alpha}{q}+1}/\left(\frac{\alpha}{q}+1\right).$ □

3.4. Proposition. *Let $0 < a < b$, $q > 1$, $1/p + 1/q = 1$ and $0 < \alpha < 1$. Then we have the following inequality*

$$\left|A\left(a^{\frac{\alpha}{q}+1}, b^{\frac{\alpha}{q}+1}\right) - G^2 L_{\frac{\alpha}{q}-1}^{\frac{\alpha}{q}-1}\right|$$
$$\leq \frac{ab(b-a)(\alpha+q)}{2q}\left(\frac{1}{p+1}\right)^{\frac{1}{p}}\left(\nu(\alpha,q;a,b)a^\alpha + (\nu(0,q;a,b)-\nu(\alpha,q;a,b))b^\alpha\right)^{\frac{1}{q}}.$$

Proof. The assertion follows from the inequality (2.9) in Corollary 2.17, for
$f : (0, \infty) \to \mathbb{R},\ f(x) = x^{\frac{\alpha}{q}+1} / \left(\frac{\alpha}{q} + 1\right).$ □

References

[1] Abramowitz M. and Stegun I.A. (Eds.) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1965).

[2] Bakula, M.K., Ozdemir, M. E. and Pecaric, J. *Hadamard type inequalities for m-convex and (α, m)-convex functions*, J. Inequal. Pure Appl. Math. **9** (4), Article 96, p. 12, 2008.

[3] Hudzik, H. and Maligranda, L. *Some remarks on s-convex functions*, Aequationes Math. **48**, 100-111, 1994.

[4] İşcan, İ. *A new generalization of some integral inequalities for (α, m)-convex functions*, Mathematical Sciences **7** (22), 1-8, 2013.

[5] İşcan, İ. *New estimates on generalization of some integral inequalities for (α, m)-convex functions*, Contemp. Anal. Appl. Math. **1** (2), 253-264, 2013.

[6] İşcan, İ. *Hermite-Hadamard type inequalities for functions whose derivatives are (α, m)-convex*, International Journal of Engineering and Applied sciences **2** (3), 69-78, 2013.

[7] İşcan, İ. *Hermite-Hadamard type inequalities for harmonically convex functions*, Hacet. J. Math. Stat. **43**(6), 935-942, 2014.

[8] Miheşan, V. G. *A generalization of the convexity*, Seminer on Functional Equations, Approximation and Convexity, Cluj-Napoca, Romania,1993.

[9] Ozdemir, M. E., Avcı, M. and Kavurmacı, H. *Hermite-Hadamard-type inequalities via (α, m)-convexity*, Comput. Math. Appl. **61,** 2614-2620, 2011.

[10] Ozdemir, M. E., Kavurmacı, H. and Set, E. *Ostrowski's type inequalities for (α, m)-convex functions*, Kyungpook Math. J. **50**, 371-378, 2010.

[11] Ozdemir, M. E., Set, E. and Sarıkaya, M. Z. *Some new Hadamard's type inequalities for co-ordinated m-convex and (α, m)-convex functions*, Hacet. J. Math. Stat. **40** (2), 219-229, 2011.

[12] Set, E., Sardari, M., Ozdemir, M. E. and Rooin, J. *On generalizations of the Hadamard inequality for (α, m)-convex functions,* RGMIA Res. Rep. Coll. **12** (4), Article 4, 2009.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 391 – 416

New existence results for positive solutions of boundary value problems for coupled systems of multi-term fractional differential equations

Yuji Liu* † ‡

Abstract

In this article, we establish some new existence results on positive solutions of a boundary value problem of coupled systems of nonlinear multi-term fractional differential equations. Our analysis rely on the well known fixed point theorems. Numerical examples are given to illustrate the main theorems.

Keywords: Four-point boundary value problem, multi-term fractional differential system, non-Caratheodory function, fixed-point theorem.

2000 AMS Classification: 34A08, 26A33, 39B99, 45G10, 34B37, 34B15, 34B16

Received : 02.12.2014 *Accepted* : 18.02.2015 *Doi* : 10.15672/HJMS.20164512499

1. Introduction

Fractional differential systems have many applications in modeling of physical and chemical processes and in engineering and have been of great interest recently. In its turn, mathematical aspects of studies on fractional differential systems were discussed by many authors, see the text book [6, 13] and papers [1, 7, 9, 11, 14, 15, 16, 20, 21, 22, 23].

In [20], the author studied the existence of positive solutions (continuous on $[0,1]$) of the following $(n-1,1)$-type conjugate boundary value problem for the coupled system of the fractional differential equations

$$(1.1)\qquad \begin{cases} D_{0+}^{\alpha}u+\lambda f(t,v)=0, 0<t<1, \lambda>0,\\ D_{0+}^{\alpha}v+\lambda g(t,u)=0,\\ u^{(i)}(0)=v^{(i)}(0)=0, 0\le i\le n-2,\\ u(1)=v(1)=0,\end{cases}$$

*Department of Mathematics, Guangdong University of Finance and Economics, Guangzhou 510000, P.R.China, Email: `liuyuji888@sohu.com`

†Corresponding Author.

‡Supported by the National Natural Science Foundation of China (No: 11401111), the Natural Science Foundation of Guangdong province (No:S2011010001900) and the Foundation for High-level talents in Guangdong Higher Education Project.

where λ is a parameter, $\alpha \in (n-1, n]$ is a real number and $n \geq 3$, and D^{α}_{0+} is the Riemann-Liouville′s fractional derivative, and f, g are continuous and semipositone.

In [9], the author studied the system of fractional boundary value problems of the form

$$
(1.2) \qquad \begin{cases} D^{\alpha}_{0+}u(t) + \lambda a(t) f(u(t), v(t)) = 0, \quad t \in (0,1), \\ D^{\beta}_{0+}v(t) + \mu b(t) g(u(t), v(t)) = 0, \quad t \in (0,1), \\ u^{(i)}(0) = 0, i = 0, 1, 2, \cdots, n-2, \;\; D^{\gamma}_{0+}u(1) = 0, 1 < \gamma < n-2, \\ v^{(i)}(0) = 0, i = 0, 1, 2, \cdots, n-2, \;\; D^{\gamma}_{0+}v(1) = 0, 1 < \gamma < n-2, \end{cases}
$$

where D_{0+} is the Riemann-Liouville fractional derivative, $n-1 < \alpha, \beta < n$ for $n > 3$ and $n \in N$, a and b are continuous on $[0,1]$, f and g continuous functions defined on R^2. Sufficient conditions for the existence of at least one positive solution (continuous on $[0,1]$) of BVP(1.2) were obtained.

In known literature, $D^{\alpha}_{0+}u(t) + f(t, u(t)) = 0$ is known as a **single term** equation. In certain cases, we find equations containing more than one differential terms. A classical example is the so-called **Bagley Torvik equation**

$$
AD^{2}_{0+}y(x) + BD^{\frac{3}{2}}_{0+}y(x) + Cy(x) = f(x),
$$

where A, B, C are constants and f is a given function. This equation arises from for example the modelling of motion of a rigid plate immersed in a Newtonian fluid. It was originally proposed in [18]. Another example for an application of equations with more than one fractional derivatives is the **Basset equation**

$$
AD^{1}_{0+}y(x) + bD^{n}_{0+}y(x) + cy(x) = f(x), y(0) = y_0,
$$

where $0 < n < 1$. This equation is most frequently, but not exclusively, used with $n = \frac{1}{2}$. It describes the forces that occur when a spherical object sinks in a (relatively dense) incompressible viscous fluid, see [4, 12].

In [17], Su investigated the existence of positive solutions (continuous on $[0,1]$) of the following boundary value problem of nonlinear multi-term fractional differential system

$$
(1.3) \qquad \begin{cases} D^{\alpha}_{0+}u + f(t, v(t), D^{p}_{0+}v(t)) = 0, 0 < t < 1, \\ D^{\beta}_{0+}v + g(t, u(t), D^{q}_{0+}u(t)) = 0, 0 < t < 1, \\ u(0) = 0, \; u(1) = 0, \;\; v(0) = 0, \; v(1) = 0, \end{cases}
$$

where $\alpha, \beta \in (1,2)$, D_{0+} is the Riemann-Liouville's fractional derivative, $0 < p < \beta - 1$, $0 < q < \alpha - 1$, $\gamma\eta^{\alpha-1} < 1$ and $\gamma\eta^{\beta-1} < 1$, $f, g : [0,1] \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ are continuous functions.

In [21], authors studied the existence of multiple positive solutions (continuous on $[0,1]$) of the following boundary value problem of N-dimension nonlinear fractional differential system

$$
(1.4) \qquad \begin{cases} D^{\alpha_1}_{0+}u_1 + f_1(t, u_2(t), D^{\mu_1}_{0+}u_2(t)) = 0, 0 < t < 1, \\ \cdots\cdots\cdots, \\ D^{\alpha_{N-1}}_{0+}u_{N-1} + f_{N-1}(t, u_N(t), D^{\mu_N}_{0+}u_N(t)) = 0, 0 < t < 1, \\ D^{\alpha_N}_{0+}u_N + f_N(t, u_1(t), D^{\mu_N}_{0+}u_1(t)) = 0, 0 < t < 1, \\ u_1(0) = \cdots = u_N(0) = 0, \;\; u_1(1) = \cdots = u_N(1) = 0, \end{cases}
$$

where $\alpha_i \in (1,2)$, D_{0+} is the Riemann-Liouville's fractional derivative, $0 < \mu_{i-1} < \alpha_i - 1$ with $\mu_0 = \mu_N$, $f_i :; [0,1] \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}(i = 1, 2, \cdots, N)$ are continuous functions.

In [1], the authors investigated the existence of positive solutions (continuous on $[0,1]$) of the following boundary value problem of nonlinear multi-term fractional differential system

$$
(1.5)\qquad \begin{cases} D_{0+}^{\alpha}u + f(t, v(t), D_{0+}^{p}v(t)) = 0, 0 < t < 1,\\ D_{0+}^{\beta}v + g(t, u(t), D_{0+}^{q}u(t)) = 0, 0 < t < 1,\\ u(0) = 0,\ u(1) = \gamma u(\eta),\\ v(0) = 0,\ v(1) = \gamma v(\eta), \end{cases}
$$

where $\alpha, \beta \in (1, 2)$, D_{0+} is the Riemann-Liouville's fractional derivative, $0 < p \leq \beta - 1$, $0 < q \leq \alpha - 1$, $\gamma\eta^{\alpha-1} < 1$ and $\gamma\eta^{\beta-1} < 1$, $f, g : [0, 1] \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ are continuous functions.

In [24], authors studied the existence of solutions of the following four-point coupled boundary value problem for nonlinear fractional differential equation

$$
(1.6)\qquad \begin{cases} D_{0+}^{\alpha}u = f(t, u(t), D_{0+}^{\alpha-1}u(t), v(t), D_{0+}^{\beta-1}v(t)), 0 < t < 1,\\ D_{0+}^{\beta}v = g(t, u(t), D_{0+}^{\alpha-1}u(t), v(t), D_{0+}^{\beta-1}v(t)), 0 < t < 1,\\ I_{0+}^{2-\alpha}u(0) = 0,\ u(1) = av(\xi),\ \ I_{0+}^{2-\beta}v(0) = 0,\ v(1) = bu(\eta), \end{cases}
$$

where $1 < \alpha, \beta < 2$, D_{0+}^{*} and I_{0+}^{*} are the standard Riemann-Liouville differentiation and integration, $f, g : [0, 1] \times \mathbb{R}^4 \to R$ are continuous functions, $a, b \in R$, $\xi, \eta \in (0, 1)$ with $ab\xi^{\beta-1}\eta^{\alpha-1} = 1$.

In [8], the existence of positive solutions of the following four-point boundary value problem of multi-term fractional differential system

$$
(1.7)\qquad \begin{cases} D_{0+}^{\alpha}u = f(t, v(t), D_{0+}^{m}v(t)), 0 < t < 1,\\ D_{0+}^{\beta}v = g(t, u(t), D_{0+}^{n}u(t)), 0 < t < 1,\\ u(0) = \gamma u(\xi),\ u(1) = \delta u(\eta),\ \ v(0) = \gamma v(\xi),\ v(1) = \delta v(\eta), \end{cases}
$$

was studied, where $1 < \alpha, \beta < 2$, $0 < m \leq \beta - 1$, $0 < n \leq \alpha - 1$, $\gamma > 0, \delta > 0$, $0 < \xi < \eta < 1$, D_{0+}^{*} is the standard Riemann-Liouville differentiation, $f, g : [0, 1] \times R^4 \to R$ are continuous functions and the following assumption **(A):**

$$
\max\{\delta\eta^{\alpha-1}, \delta\eta^{\alpha-2}\} < 1,\ \max\{\delta\eta^{\beta-1}, \delta\eta^{\beta-2}\} < 1,
$$

$$
\max\{\gamma\xi^{\alpha-1}, \gamma\xi^{\alpha-2}\} < 1,\ \max\{\gamma\xi^{\beta-1}, \gamma\xi^{\beta-2}\} < 1.
$$

In [2], Ahmad and Sivasundaram considered the existence and uniqueness of solutions for the following four-point nonlocal boundary value problem of nonlinear fractional integro-differential equation

$$
\begin{cases} {}^{c}D_{0+}^{q}x(t) = f(t, x(t), (\phi x)(t), (\psi x(t)), 0 < t < 1,\\ x'(0) + \alpha x(\eta_1) = 0,\ \ bx'(1) + x(\eta_2) = 0, \end{cases}
$$

where $1 < q \leq 2, a, b \in [0, 1], 0 < \eta_1 \leq \eta_2 < 1$, ${}^{c}D_{0+}^{q}$ is the Caputo's fractional derivative, $f : [0, 1] \times X \times X \times X \to X$ is continuous, for $\gamma, \delta : [0, 1] \times [0, 1] \to [0, +\infty)$ with

$$
(\phi x)(t) = \int_0^t \gamma(t, s)x(s)ds,\ (\psi x)(t) = \int_0^1 \delta(t, s)x(s)ds.
$$

We remark that the boundary conditions $x'(0) + \alpha x(\eta_1) = 0, bx'(1) + x(\eta_2) = 0$ arise in the study of heat flow problems involving a bar of unit length with two controllers at $t = 0$ and $t = 1$ adding or removing heat according to the temperatures detected by two sensors at $t = \eta_1$ and $t = \eta_2$.

We note firstly that the existence of positive solutions of BVP(1.7) has not been concerned in known papers when the assumption **(A)** does not hold. Secondly, to guarantee the solvability of BVP(1.3) and BVP(1.5) in [1], the assumptions imposed on the nonlinearities are as follows:

$$|f(t,x,y)| \leq a(t) + \epsilon_1|x|^{\rho_1} + \epsilon_2|y|^{\rho_2}, \epsilon_1, \epsilon_2 > 0, 0 < \rho_1, \rho_2 < 1,$$

$$|g(t,x,y)| \leq b(t) + \delta_1|x|^{\sigma_1} + \delta_2|y|^{\sigma_2}, \delta_1, \delta_2 > 0, 0 < \sigma_1, \sigma_2 < 1.$$

While in [17], another assumptions imposed on f, g are as follows:

$$|f(t,x,y)| \leq \epsilon_1|x|^{\rho_1} + \epsilon_2|y|^{\rho_2}, \epsilon_1, \epsilon_2 > 0, \rho_1, \rho_2 > 1,$$

$$|g(t,x,y)| \leq \delta_1|x|^{\sigma_1} + \delta_2|y|^{\sigma_2}, \delta_1, \delta_2 > 0, \sigma_1, \sigma_2 > 1.$$

By carefully checking Example 3.1 in [17], one finds that the solution obtained may be the zero solution. This fact makes these papers far from perfect. Thirdly, it is easy to show that the following problem

$$D_{0^+}^{\frac{7}{3}}x(t) = -t^{-\frac{1}{2}}(1-t)^{-\frac{5}{4}},\ \lim_{t\to 0} t^{\frac{2}{3}}x(t) = 0,\ x(1) = 0$$

has a continuous solution on $[0,1]$

$$x(t) = -\int_0^t \frac{(t-s)^{\frac{4}{3}}}{\Gamma(7/3)} s^{-\frac{1}{2}}(1-s)^{-\frac{5}{4}}ds + t^{\frac{4}{3}}\int_0^1 \frac{(1-s)^{\frac{1}{12}}}{\Gamma(7/3)} s^{-\frac{1}{2}}ds,$$

while $t^{-\frac{1}{2}}(1-t)^{-\frac{5}{4}}$ is not measurable on $(0,1)$. Hence it is interesting to investigate the solvability of mentioned problems with non-Caratheodory functions.

Motivated by above mentioned papers, we discuss the existence of solutions of the following boundary value problem of the multi-term fractional differential system

$$\begin{cases} D_{0^+}^{\alpha}u(t) + p(t)f(t, v(t), D_{0^+}^{n}v(t)) = 0, \quad t \in (0,1), \\ D_{0^+}^{\beta}v(t) + q(t)g(t, u(t), D_{0^+}^{m}u(t)) = 0, \quad t \in (0,1), \\ \lim\limits_{t\to 0} t^{2-\alpha}u(t) - au(\xi) = \int_0^1 \phi_1(t, v(t), D_{0^+}^{n}v(t))dt, \\ u(1) - bu(\eta) = \int_0^1 \psi_1(t, v(t), D_{0^+}^{n}v(t))dt, \\ \lim\limits_{t\to 0} t^{2-\beta}v(t) - cv(\xi) = \int_0^1 \phi_2(t, u(t), D_{0^+}^{m}u(t))dt, \\ v(1) - dv(\eta) = \int_0^1 \psi_2(t, u(t), D_{0^+}^{m}u(t))dt, \end{cases} \tag{1.8}$$

where

(i) $1 < \alpha, \beta \leq 2$, $0 < m \leq \alpha - 1$ and $0 < n \leq \beta - 1$, $D_{0^+}^{*}$ is the standard Riemann-Liouville differentiation of order $* > 0$,

(ii) $0 < \xi \leq \eta < 1$ and $a, b, c, d \geq 0$,

(iii) $\mathbb{R}$ denote the set of real numbers and $\mathbb{R}_+$ the set of nonnegative real numbers, $p, q : (0,1) \to \mathbb{R}_+$, p satisfies that there exist numbers k_1, l_1 such that $k_1 > -1$, $\alpha - m + l_1 > 0$, $2 + k_1 + l_1 > 0$ and $|p(t)| < t^{k_1}(1-t)^{l_1}$ for $t \in (0,1)$, q satisfies that there exist numbers k_2, l_2 such that $k_2 > -1$, $\beta - n + l_2 > 0$, $2 + k_2 + l_2 > 0$ and $|q(t)| < t^{k_2}(1-t)^{l_2}$ for $t \in (0,1)$, with $p(t) \not\equiv 0$ and $q(t) \not\equiv 0$ on $(0,1)$,

(iv) $f, \phi_1, \psi_1 : (0,1) \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$ are $(n, \beta)-$Caratheory functions and $g, \phi_2, \psi_2 : (0,1) \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$ is a $(m, \alpha)-$Caratheory functions with $f(t,0,0) \not\equiv 0$ and $g(t,0,0) \not\equiv 0$ on $(0,1)$.

We obtain the results on positive solutions of BVP(1.8) by using Schauder's fixed point theorem in Banach spaces. A pair of functions (x, y) is called a solution of BVP(1.8) if $x, y \in C^0(0,1]$ and x, y satisfy all equations in (1.8). A pair of functions (x, y) is called a positive solution of BVP(1.8) if $x, y \in C^0(0,1]$ are positive on $(0,1]$ and x, y satisfy all equations in (1.8).

The salient features of the present study are as follows:

(a) the fractional differential equations in (1.8) are multi-term ones and their nonlinearities depend on the lower order fractional derivatives with order greater than $\alpha-1$ and $\beta-1$;

(b) instead of the condition $u(0)=0, v(0)=0$ we consider integral boundary conditions which are more suitable as $D_{0+}^{\alpha}x(t)=0$ with $\alpha\in(1,2)$ implies $x(t)=ct^{\alpha-1}+dt^{\alpha-2}$ and obviously x is not continuous at $t=0$ while $\lim\limits_{t\to 0^+} t^{2-\alpha}x(t)$ exists;

(c) BVP(1.8) is a generalized form of known ones in references, the positive solutions of BVP(1.8) obtained are unbounded (discontinuous at $t=0$) which are different from those ones (continuous on [0,1]) in [1, 21, 20, 9];

(d) since p,q may be un-measurable on $(0,1)$, $p(t)f(t,x,y)$ and $q(t)g(t,x,y)$ may be non-Caratheodory functions (see Example 4.1 in which the nonlinearities are

$$t^{-\frac{1}{10}}(1-t)^{-\frac{21}{20}}f(t,v(t),D_{0+}^{\frac{17}{20}}v(t)),\ \ t^{-\frac{1}{10}}(1-t)^{-\frac{23}{20}}g(t,u(t),D_{0+}^{\frac{4}{5}}u(t))$$

with

$$f(t,u,v)=t^2+b_1tu^{\epsilon_1}+a_1tv^{\delta_1},\ a_1,b_1\geq 0,\ \epsilon_1,\delta_1>0,$$

$$g(t,u,v)=t^5+b_2tu^{\sigma_1}+a_2tv^{\gamma_1},\ a_2,b_2\geq 0,\ \sigma_1,\gamma_1>0).$$

It is easy to see that both $t^{-\frac{1}{10}}(1-t)^{-\frac{21}{20}}$ and $t^{-\frac{1}{10}}(1-t)^{-\frac{23}{20}}$ are not measurable on $(0,1)$. Our results are new and are well illustrated with an example.

(e) The Green's function $G(t,s)$ for the problem $-D_{0+}^{\alpha}x(t))=0$, $\lim\limits_{t\to 0^+} t^{2-\alpha}x(t)-ax(\xi)=0, x(1)-bx(\eta)=0$ is obtained. We proved that $G(t,s)\geq 0$ under some assumptions which are more weaker than **(A)** in [8] and actually generalize Lemma 2.2 in ([10] J. Math. Anal. Appl. 305 (2005) 253-276) for problem $-x''(t)=0, x(0)-ax(\xi)=x(1)-bx(\eta)=0$. See Lemma 2.9.

The remainder of this paper is arranged as follows: in Section 2, we present preliminary results; in Section 3, the main result is presented; and two examples are given in Section 4 to illustrate the main result.

2. Preliminary results

For the convenience of readers, we present here the necessary definitions from fixed point theory and fractional calculus theory.

2.1. Definition. Let X be a Banach space. An operator $T:X\to X$ is completely continuous if it is continuous and maps bounded sets into relatively compact sets [3].

2.2. Definition. The Riemann-Liouville fractional integral of order $\alpha>0$ of a function $f:(0,+\infty)\to R$ is given by $I_{0+}^{\alpha}f(t)=\frac{1}{\Gamma(\alpha)}\int_0^t(t-s)^{\alpha-1}f(s)ds$, provided that the right-hand side exists [13].

2.3. Definition. The Riemann-Liouville fractional derivative of order $\alpha>0$ of a continuous function $f:(0,+\infty)\to R$ is given by $D_{0+}^{\alpha}f(t)=\frac{d^n}{dt^n}\int_0^t\frac{(t-s)^{n-\alpha-1}f(s)}{\Gamma(n-\alpha)}ds$, where $n-1\leq\alpha<n$, provided that the right-hand side exists [13].

2.4. Definition. $h:(0,1)\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is called a $(m,\alpha)-$Carathédory function if it satisfies

(i) $t\to h\left(t,t^{\alpha-2}x,t^{2+m-\alpha}y\right)$ is measurable on $(0,1)$ for all $(x,y)\in\mathbb{R}^2$,

(ii) $(x,y)\to h\left(t,t^{\alpha-2}x,t^{2+m-\alpha}y\right)$ is continuous for a.e. $t\in(0,1)$,

(iii) for each $r>0$, there exists nonnegative function $\phi_r\in L^1(0,1)$ such that $|u|,|v|\leq r$ imply $\left|h\left(t,t^{\alpha-2}x,t^{2+m-\alpha}y\right)\right|\leq\phi_r(t), a.e., t\in(0,1)$.

2.5. Lemma. *Let $n-1 \le \alpha < n$, $u \in C^0(0,b) \bigcap L^1(0,b)$ with $b > 0$. Then*

$$D_{0+}^{\alpha} I_{0+}^{\alpha} u(t) = u(t), \quad I_{0+}^{\alpha} D_{0+}^{\alpha} u(t) = u(t) + C_1 t^{\alpha-1} + C_2 t^{\alpha-2} + \cdots + C_n t^{\alpha-n},$$

where $C_i \in R$, $i = 1, 2, \ldots n$ [13].

Choose

$$X = \left\{ x : (0,1] \to \mathbb{R} \quad \begin{array}{c} x,\ D_{0+}^{m} x \in C^0(0,1] \\ \text{the following limits exist} \\ \lim\limits_{t\to 0} t^{2-\alpha} x(t),\ \lim\limits_{t\to 0} t^{2+m-\alpha} D_{0+}^{m} x(t) \end{array} \right\}$$

with the norm

$$||x|| = \|x\|_X = \max \left\{ \sup_{t\in(0,1]} t^{2-\alpha}|x(t)|,\ \sup_{t\in(0,1]} t^{2+m-\alpha}|D_{0+}^{m} x(t)| \right\}$$

for $x \in X$. It is easy to show that X is a real Banach space.

Choose

$$Y = \left\{ y : (0,1] \to \mathbb{R} \quad \begin{array}{c} y,\ D_{0+}^{n} y \in C^0(0,1] \\ \text{the following limits exist} \\ \lim\limits_{t\to 0} t^{2-\beta} y(t),\ \lim\limits_{t\to 0} t^{2+n-\beta} D_{0+}^{n} y(t) \end{array} \right\}$$

with the norm

$$||y|| = \|y\|_Y = \max \left\{ \sup_{t\in(0,1]} t^{2-\beta}|y(t)|,\ \sup_{t\in(0,1]} t^{2+n-\beta}|D_{0+}^{n} y(t)| \right\}$$

for $y \in Y$. It is easy to show that Y is a real Banach space.

Thus, $(X \times Y, ||\cdot||)$ is Banach space with the norm defined by

$$||(x,y)|| = \max\{||x|| = ||x||_X,\ ||y|| = ||y||_Y\} \text{ for } (x,y) \in X \times Y.$$

For ease expression, we denote $F_{m,x}(t) = F(t, x(t), D_{0+}^{m} x(t))$ for a function $x : (0,1] \to \mathbb{R}$, a number m and a function $F : (0,1) \times \mathbb{R}^2 \to \mathbb{R}$.

Denote

$$\begin{aligned} &\mu_1 = a\xi^{\alpha-1},\ \upsilon_1 = 1 - a\xi^{\alpha-2},\ \omega_1 = 1 - b\eta^{\alpha-1},\ \lambda_1 = 1 - b\eta^{\alpha-2}, \\ &\Delta = \mu_1\lambda_1 + \upsilon_1\omega_1, \\ &\mu_2 = c\xi^{\beta-1},\ \upsilon_2 = 1 - c\xi^{\beta-2},\ \omega_2 = 1 - d\eta^{\beta-1},\ \lambda_2 = 1 - d\eta^{\beta-2}, \\ &\nabla = \mu_2\lambda_2 + \upsilon_2\omega_2. \end{aligned} \tag{2.1}$$

2.6. Lemma. *Suppose that $\Delta \ne 0$ and*

(B0) *$h \in C^0(0,1)$ and there exist $k > -1$ and $l \le 0$ such that $2 + l + k > 0$ and $|h(t)| \le t^k(1-t)^l$ for all $t \in (0,1)$.*

Then $x \in X$ is a solution of problem

$$\begin{cases} D^{\alpha} x(t) + h(t) = 0, 0 < t < 1, \\ \lim\limits_{t\to 0} t^{2-\alpha} x(t) - a x(\xi) = M, \\ x(1) - b x(\eta) = N \end{cases} \tag{2.2}$$

if and only if $x \in X$ satisfies

$$x(t) = \frac{\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2}}{\Delta} N + \frac{\omega_1 t^{\alpha-2}-\lambda_1 t^{\alpha-1}}{\Delta} M$$

$$-\int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds + \frac{\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2}}{\Delta} \int_0^1 \frac{(1-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds \tag{2.3}$$

$$-\frac{b\upsilon_1 t^{\alpha-1}+b\mu_1 t^{\alpha-2}}{\Delta} \int_0^\eta \frac{(\eta-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds + \frac{a\lambda_1 t^{\alpha-1}-a\omega_1 t^{\alpha-2}}{\Delta} \int_0^\xi \frac{(\xi-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds.$$

Proof. From (B0), we have

$$t^{2-\alpha}\left|\int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds\right| \le t^{2-\alpha}\int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} s^k(1-s)^l ds$$

$$\le t^{2-\alpha}\int_0^t \frac{(t-s)^{\alpha+l-1}}{\Gamma(\alpha)} s^k ds = t^{\alpha+l+k} \int 0^1 \frac{(1-w)^{\alpha+l-1}}{\Gamma(\alpha)} w^k dw$$

$$= t^{2+l+k} \frac{\mathbf{B}(\alpha+l,k+1)}{\Gamma(\alpha)} \to 0 \text{ as } t \to 0.$$

Suppose that $x \in X$ is a solution of (2.2). Lemma 2.5 implies that there exist $c_i (i = 1, 2)$ such that

$$x(t) = -\int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds + c_1 t^{\alpha-1} + c_2 t^{\alpha-2}. \tag{2.4}$$

One sees from the boundary conditions in (2.2) that

$$\mu_1 c_1 - \upsilon_1 c_2 = -M + a\int_0^\xi \frac{(\xi-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds,$$

$$\omega_1 c_1 + \lambda_1 c_2 = N + \int_0^1 \frac{(1-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds - b\int_0^\eta \frac{(\eta-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds.$$

It follows that

$$c_1 = \frac{1}{\Delta}\left[\upsilon_1 N - \lambda_1 M + \upsilon_1 \int_0^1 \frac{(1-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds\right.$$

$$\left. -b\upsilon_1 \int_0^\eta \frac{(\eta-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds + a\lambda_1 \int_0^\xi \frac{(\xi-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds\right],$$

$$c_2 = \frac{1}{\Delta}\left[\mu_1 N + \omega_1 M + \mu_1 \int_0^1 \frac{(1-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds\right.$$

$$\left. -b\mu_1 \int_0^\eta \frac{(\eta-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds - a\omega_1 \int_0^\xi \frac{(\xi-s)^{\alpha-1}}{\Gamma(\alpha)} h(s)ds\right].$$

Substitute c_1, c_2 into (2.4), we get (2.3).

On the other hand, if $x \in X$ satisfies (2.3), we can show that $x \in X$ is a solution of BVP(2.2). The proof is completed. $\square$

2.7. Lemma. *Suppose that* $\nabla \neq 0$ *and* **(B0)** *holds. Then* $y \in Y$ *is a solution of problem*

$$(2.5)\qquad \begin{cases} D^{\beta}y(t) + h(t) = 0, 0 < t < 1, \\ \lim\limits_{t\to 0} t^{2-\beta}y(t) - cy(\xi) = M, \\ y(1) - dy(\eta) = N \end{cases}$$

if and only if $y \in Y$ *satisfies*

$$(2.6)\qquad \begin{aligned} y(t) &= \tfrac{\upsilon_2 t^{\beta-1}+\mu_2 t^{\beta-2}}{\nabla}N + \tfrac{\omega_2 t^{\beta-2}-\lambda_2 t^{\beta-1}}{\nabla}M \\ &- \int_0^t \tfrac{(t-s)^{\beta-1}}{\Gamma(\beta)}h(s)ds + \tfrac{\upsilon_2 t^{\beta-1}+\mu_2 t^{\beta-2}}{\nabla}\int_0^1 \tfrac{(1-s)^{\beta-1}}{\Gamma(\beta)}h(s)ds \\ &- \tfrac{d\lambda_2 t^{\beta-1}+d\mu_2 t^{\beta-2}}{\nabla}\int_0^{\eta} \tfrac{(\eta-s)^{\beta-1}}{\Gamma(\beta)}h(s)ds + \tfrac{c\lambda_2 t^{\beta-1}-c\omega_2 t^{\beta-2}}{\nabla}\int_0^{\xi} \tfrac{(\xi-s)^{\beta-1}}{\Gamma(\beta)}h(s)ds. \end{aligned}$$

Proof. The proof is similar to that of Lemma 2.6 and is omitted. □

Define the operator T on $X \times Y$ by $T(x,y)(t) = ((T_1y)(t), (T_2x)(t))$ with

$$\begin{aligned} (T_1y)(t) &= \tfrac{\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2}}{\Delta}\int_0^1 \psi_{1_{n,y}}(s)ds + \tfrac{\omega_1 t^{\alpha-2}-\lambda_1 t^{\alpha-1}}{\Delta}\int_0^1 \phi_{1_{n,y}}(s)ds \\ &- \int_0^t \tfrac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}p(s)f_{n,y}(s)ds + \tfrac{\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2}}{\Delta}\int_0^1 \tfrac{(1-s)^{\alpha-1}}{\Gamma(\alpha)}p(s)f_{n,y}(s)ds \\ &- \tfrac{b\upsilon_1 t^{\alpha-1}+b\mu_1 t^{\alpha-2}}{\Delta}\int_0^{\eta} \tfrac{(\eta-s)^{\alpha-1}}{\Gamma(\alpha)}p(s)f_{n,x}(s)ds \\ &+ \tfrac{a\lambda_1 t^{\alpha-1}-a\omega_1 t^{\alpha-2}}{\Delta}\int_0^{\xi} \tfrac{(\xi-s)^{\alpha-1}}{\Gamma(\alpha)}p(s)f_{n,y}(s)ds \end{aligned}$$

and

$$\begin{aligned} (T_2x)(t) &= \tfrac{\upsilon_2 t^{\beta-1}+\mu_2 t^{\beta-2}}{\nabla}\int_0^1 \psi_{2_{m,x}}(s)ds + \tfrac{\omega_2 t^{\beta-2}-\lambda_2 t^{\beta-1}}{\nabla}\int_0^1 \phi_{2_{m,x}}(s)ds \\ &- \int_0^t \tfrac{(t-s)^{\beta-1}}{\Gamma(\beta)}q(s)g_{m,x}(s)ds + \tfrac{\upsilon_2 t^{\beta-1}+\mu_2 t^{\beta-2}}{\nabla}\int_0^1 \tfrac{(1-s)^{\beta-1}}{\Gamma(\beta)}q(s)g_{m,x}(s)ds \\ &- \tfrac{d\upsilon_2 t^{\beta-1}+d\mu_2 t^{\beta-2}}{\nabla}\int_0^{\eta} \tfrac{(\eta-s)^{\beta-1}}{\Gamma(\beta)}q(s)g_{m,x}(s)ds \\ &+ \tfrac{c\lambda_2 t^{\beta-1}-c\omega_2 t^{\beta-2}}{\nabla}\int_0^{\xi} \tfrac{(\xi-s)^{\beta-1}}{\Gamma(\beta)}q(s)g_{m,x}(s)ds \end{aligned}$$

for $(x,y) \in X \times Y$.

By Lemmas 2.6 and 2.7, we have that $(x,y) \in X \times Y$ is a solution of BVP(1.8) if and only if $(x,y) \in X \times Y$ is a fixed point of T.

2.8. Lemma. *Suppose that (i)-(iv) defined in Section 1 hold,* $\Delta \neq 0$ *and* $\nabla \neq 0$*. Then* $T : X \times Y \to X \times Y$ *is completely continuous.*

Proof. If $0 < m \leq \alpha - 1$ and $0 < n \leq \beta - 1$, we take

$$t^{2-\alpha}(T_1y)(t)|_{t=0} = \lim_{t\to 0} t^{2-\alpha}(T_1y)(t),$$

$$t^{2+m-\alpha}D_{0+}^{m}(T_1y)(t)|_{t=0} = \lim_{t\to 0} t^{2+m-\alpha}D_{0+}^{m}(T_1y)(t),$$

$$t^{2-\beta}(T_2x)(t)|_{t=0} = \lim_{t\to 0} t^{2-\beta}(T_2x)(t),$$

$$t^{2+n-\beta}D_{0+}^{n}(T_2x)(t)|_{t=0} = \lim_{t\to 0} t^{2+n-\beta}D_{0+}^{n}(T_2x)(t),$$

then $t^{2-\alpha}(T_1y)(t)$, $t^{2+m-\alpha}D_{0+}^{m}(T_1y)(t)$ and $t^{2-\beta}(T_2x)(t)$, $t^{2+n-\beta}D_{0+}^{n}(T_2x)(t)$ are continuous on $[0,1]$ for each $(x,y) \in X \times Y$. It is easy to show that T is completely continuous, we refer similar proofs to [1]. The proof is complete. $\square$

Now, we rewrite

$$(T(x,y))(t) = ((T_1y)(t),(T_2x)(t))$$

$$= \Bigg(\tfrac{\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2}}{\Delta} \int_0^1 \psi_{1_{n,y}}(s)ds + \tfrac{\omega_1 t^{\alpha-2}-\lambda_1 t^{\alpha-1}}{\Delta} \int_0^1 \phi_{1_{n,y}}(s)ds$$

$$+ \int_0^1 G(t,s)p(s)f_{n,y}(s)ds,$$

$$\tfrac{\upsilon_2 t^{\beta-1}+\mu_2 t^{\beta-2}}{\nabla} \int_0^1 \psi_{2_{m,x}}(s)ds + \tfrac{\omega_2 t^{\beta-2}-\lambda_2 t^{\beta-1}}{\nabla} \int_0^1 \phi_{2_{m,x}}(s)ds + \int_0^1 H(t,s)g_{m,x}(s)ds \Bigg).$$

Here

$$G(t,s) = \tfrac{1}{\Gamma(\alpha)\Delta} \begin{cases} \begin{array}{l} (\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2})(1-s)^{\alpha-1} \\ +(\lambda_1 a t^{\alpha-1}-\omega_1 a t^{\alpha-2})(\xi-s)^{\alpha-1} \\ -(\upsilon_1 b t^{\alpha-1}+b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1} \\ -\left(\mu_1\lambda_1+\omega_1\upsilon_1\right)(t-s)^{\alpha-1}, \end{array} & 0 \leq s \leq \min\{t,\xi\}, \\ \\ \begin{array}{l} (\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2})(1-s)^{\alpha-1} \\ -(\upsilon_1 b t^{\alpha-1}+b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1} \\ -\left(\mu_1\lambda_1+\omega_1\upsilon_1\right)(t-s)^{\alpha-1}, \end{array} & \xi < s \leq \min\{t,\eta\}, \\ \\ \begin{array}{l} (\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2})(1-s)^{\alpha-1} \\ -(\upsilon_1 b t^{\alpha-1}+b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}, \end{array} & \max\{t,\xi\} < s \leq \eta, \\ \\ \begin{array}{l} (\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2})(1-s)^{\alpha-1} \\ +(\lambda_1 a t^{\alpha-1}-\omega_1 a t^{\alpha-2})(\xi-s)^{\alpha-1} \\ -(\upsilon_1 b t^{\alpha-1}+b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}, \end{array} & t < s \leq \xi, \\ \\ \begin{array}{l} (\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2})(1-s)^{\alpha-1} \\ -\left(\mu_1\lambda_1+\omega_1\upsilon_1\right)(t-s)^{\alpha-1}, \end{array} & \eta < s \leq t, \\ \\ (\upsilon_1 t^{\alpha-1}+\mu_1 t^{\alpha-2})(1-s)^{\alpha-1}, \max\{\eta,t\} < s \leq 1, \end{cases}$$

and

$$H(t,s)=\frac{1}{\Gamma(\beta)\nabla}\begin{cases}\begin{array}{ll}(\upsilon_2t^{\beta-1}+\mu_2t^{\beta-2})(1-s)^{\beta-1} & \\ +(\lambda_2ct^{\beta-1}-\omega_2ct^{\beta-2})(\xi-s)^{\beta-1} & 0\le s\le\min\{t,\xi\},\\ -\left(\mu_2\lambda_2+\omega_2\upsilon_2\right)(t-s)^{\beta-1}, & \\ (\upsilon_2t^{\beta-1}+\mu_2t^{\beta-2})(1-s)^{\beta-1} & \\ -(\upsilon_2dt^{\beta-1}+d\mu_2t^{\beta-2})(\eta-s)^{\beta-1} & \xi<s\le\min\{t,\eta\},\\ -\left(\mu_2\lambda_2+\omega_2\upsilon_2\right)(t-s)^{\beta-1}, & \\ (\upsilon_2t^{\beta-1}+\mu_2t^{\beta-2})(1-s)^{\beta-1} & \\ -(\upsilon_2dt^{\beta-1}+d\mu_2t^{\beta-2})(\eta-s)^{\beta-1}, & \max\{t,\xi\}<s\le\eta,\\ (\upsilon_2t^{\beta-1}+\mu_2t^{\beta-2})(1-s)^{\beta-1} & \\ +(\lambda_2ct^{\beta-1}-\omega_2ct^{\beta-2})(\xi-s)^{\beta-1} & t<s\le\xi,\\ -(\upsilon_2dt^{\beta-1}+d\mu_2t^{\beta-2})(\eta-s)^{\beta-1}, & \\ (\upsilon_2t^{\beta-1}+\mu_2t^{\beta-2})(1-s)^{\beta-1} & \\ -\left(\mu_2\lambda_2+\omega_2\upsilon_2\right)(t-s)^{\beta-1}, & \eta<s\le t,\\ (\upsilon_2t^{\beta-1}+\mu_2t^{\beta-2})(1-s)^{\beta-1},\max\{\eta,t\}<s\le1, & \end{array}\end{cases}$$

2.9. Lemma. *Suppose that*

$$\Delta>0,\ \ 0\le a<\tfrac{1}{\xi^{\alpha-2}(1-\xi)},\ \ 0\le b<\tfrac{1}{\eta^{\alpha-1}},$$

$$\nabla>0,\ \ 0\le c<\tfrac{1}{\xi^{\beta-2}(1-\xi)},\ \ 0\le d<\tfrac{1}{\eta^{\beta-1}}.$$

Then

$$G(t,s)\ge0 \text{ for all } t,s\in(0,1),\ \ H(t,s)\ge0 \text{ for all } t,s\in(0,1). \tag{2.7}$$

Proof. By the definitions of G, we consider six cases:

Case 1. $0\le s\le\min\{t,\xi\}$. Firstly, from $b\eta^{\alpha-1}<1$ and $0\le a\le\frac{1}{\xi^{\alpha-2}(1-\xi)}$, we have

$$\omega_1t^{\alpha-1}-\lambda_1t^{\alpha-2}=t^{\alpha-2}[t-1+b\eta^{\alpha-2}[\eta-t]$$

$$\begin{cases}\le0,\ \ \eta\le t,\\ =t^{\alpha-2}\left[t-1+b\eta^{\alpha-1}\frac{\eta-t}{\eta}\right]<t^{\alpha-2}\left[t-1+\frac{\eta-t}{\eta}\right]=t-1+1-\frac{t}{\eta}\le0,\eta>t,\end{cases}$$

$$\nu_1t^{\alpha-1}+b\mu_1t^{\alpha-2}=t^{\alpha-2}[t-a\xi^{\alpha-2}t+a\xi^{\alpha-1}]$$

$$\begin{cases}\ge0, & a\xi^{\alpha-2}\le1,\\ \ge1-a\xi^{\alpha-2}+a\xi^{\alpha-1}>0, & a\xi^{\alpha-2}>1.\end{cases}$$

It is easy to show that $(t-s)^{\alpha-1} \leq t^{\alpha-1}(1-s)^{\alpha-1}$ for all $0 \leq s \leq t$. Then

$$-\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1} + (a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1}$$
$$-(b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}$$
$$\geq \left[-\triangle_1 t^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2}) + (a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})\xi^{\alpha-1}\right.$$
$$\left.-(b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})\eta^{\alpha-1}\right](1-s)^{\alpha-1}$$
$$= \left[-a\xi^{\alpha-1}(1-b\eta^{\alpha-2})t^{\alpha-1} - (1-a\xi^{\alpha-2})(1-b\eta^{\alpha-1})t^{\alpha-1}\right.$$
$$+((1-a\xi^{\alpha-2})t^{\alpha-1} + a\xi^{\alpha-1}t^{\alpha-2}) + (a(1-b\eta^{\alpha-2})t^{\alpha-1} - a(1-b\eta^{\alpha-1})t^{\alpha-2})\xi^{\alpha-1}$$
$$\left.-(b(1-a\xi^{\alpha-2})t^{\alpha-1} + ab\xi^{\alpha-1}t^{\alpha-2})\eta^{\alpha-1}\right](1-s)^{\alpha-1} = 0.$$

Case 2. $\max\{t,\eta\} < s \leq 1$. We note that $0 \leq a \leq \frac{1}{\xi^{\alpha-2}(1-\xi)}$. Then

$$\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2} = t^{\alpha-1} - a\xi^{\alpha-2}t^{\alpha-1} + a\xi^{\alpha-1}t^{\alpha-2} = t^{\alpha-2}[(1-a\xi^{\alpha-2})t + a\xi^{\alpha-1}]$$
$$\begin{cases} = t^{\alpha-1} + a\xi^{\alpha-2}t^{\alpha-2}(\xi-t) \geq 0, \xi \geq t, \\ \geq 0, \xi < t, a\xi^{\alpha-2} \leq 1, \\ \geq t^{\alpha-2}[1 - a\xi^{\alpha-2} + a\xi^{\alpha-1}] \geq 0, \xi < t, a\xi^{\alpha-2} > 1. \end{cases}$$

Case 3. $\eta < s \leq t$. From $(t-s)^{\alpha-1} \leq t^{\alpha-1}(1-s)^{\alpha-1}$, we have

$$-\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$= -\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$+(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1} - (b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}$$
$$-(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1} + (b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}$$
$$\geq -\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$+(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1} - (b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1} \geq 0.$$

Case 4. $\xi < s \leq t$. We have

$$-\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$-(b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}$$
$$= -\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$+(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1} - (b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}$$
$$+\triangle_1(t-s)^{\alpha-1} + (b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}$$
$$\geq -\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$+(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1} - (b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1} \geq 0.$$

Case 5. $t < s \leq \xi$. We have

$$(\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1} + (a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1}$$
$$-(b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}$$
$$= -\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$+(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1}$$
$$-(b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1} + \triangle_1(t-s)^{\alpha-1}$$
$$\geq -\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$+(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1}$$
$$-(b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1} \geq 0.$$

Case 6. $\max\{t, \xi\} < s \leq \eta$. We have

$$(\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1} - (b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1}$$
$$= -\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$+(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1}$$
$$-(b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1} + \triangle_1(t-s)^{\alpha-1}$$
$$-(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1}$$
$$\geq -\triangle_1(t-s)^{\alpha-1} + (\nu_1 t^{\alpha-1} + \mu_1 t^{\alpha-2})(1-s)^{\alpha-1}$$
$$+(a\omega_1 t^{\alpha-1} - a\lambda_1 t^{\alpha-2})(\xi-s)^{\alpha-1}$$
$$-(b\nu_1 t^{\alpha-1} + b\mu_1 t^{\alpha-2})(\eta-s)^{\alpha-1} \geq 0.$$

We know by the definition of G that $G(t,s) \geq 0$ for all $t, s \in (0,1)$. Similarly we can prove that $H(t,s) \geq 0$ for all $t, s \in (0,1)$ The proof is completed. □

3. Main results

In this section, we prove existence result on solutions of BVP(1.8). Let $\mu_i, \upsilon_i, \omega_i, \lambda_i$ $(i = 1, 2)$ and Δ, ∇ be defined by (2.1). For $\Phi \in L^1(0,1)$, denote $||\Phi||_1 = \int_0^1 |\Phi(s)|ds$. The following assumption will be used in the main theorem.

(B1) there exist nonnegative constants $b_i, a_i (i = 1, 2)$, $B_i, A_i, C_i, D_i (i = 1, 2)$ and $\epsilon_1, \delta_i, \gamma_i$, σ_i $(i = 1, 2)$, $\Phi_i, \Psi_i, \Phi_{i0}, \Psi_{i0} \in L^1(0,1)(i = 1, 2)$ and bounded functions Φ, Ψ

such that

$$\left|f\left(t, \frac{u}{t^{2-\beta}}, \frac{v}{t^{2+n-\beta}}\right) - \Phi(t)\right| \leq b_1|u|^{\epsilon_1} + a_1|v|^{\delta_1}, t \in (0,1), u, v \in \mathbb{R},$$

$$\left|g\left(t, \frac{u}{t^{2-\alpha}}, \frac{v}{t^{2+m-\alpha}}\right) - \Psi(t)\right| \leq b_2|u|^{\sigma_1} + a_2|v|^{\gamma_1}, t \in (0,1), u, v \in \mathbb{R},$$

$$\left|\phi_1\left(t, \frac{u}{t^{2-\beta}}, \frac{v}{t^{2+n-\beta}}\right) - \Phi_{10}(t)\right| \leq \Phi_1(t)[B_1|u|^{\epsilon_1} + A_1|v|^{\delta_1}, t \in (0,1), u, v \in \mathbb{R},$$

$$\left|\psi_1\left(t, \frac{u}{t^{2-\beta}}, \frac{v}{t^{2+n-\beta}}\right) - \Psi_{10}(t)\right| \leq \Psi_1(t)[C_1|u|^{\epsilon_1} + D_1|v|^{\delta_1}], t \in (0,1), u, v \in \mathbb{R},$$

$$\left|\phi_2\left(t, \frac{u}{t^{2-\alpha}}, \frac{v}{t^{2+m-\alpha}}\right) - \Phi_{20}(t)\right| \leq \Phi_2(t)[B_2|u|^{\sigma_1} + A_2|v|^{\gamma_1}], t \in (0,1), u, v \in \mathbb{R},$$

$$\left|\psi_2\left(t, \frac{u}{t^{2-\alpha}}, \frac{v}{t^{2+m-\alpha}}\right) - \Psi_{20}(t)\right| \leq \Psi_2(t)[C_2|u|^{\sigma_1} + D_2|v|^{\gamma_1}], t \in (0,1), u, v \in \mathbb{R}.$$

For ease expression, denote

$$\overline{\Phi}(t) = \frac{(1-a\xi^{\alpha-2})t^{\alpha-1} + a\xi^{\alpha-1}t^{\alpha-2}}{\Delta} \int\limits_0^1 \Psi_{10}(s)ds$$

$$+\frac{(1-b\eta^{\alpha-1})t^{\alpha-2} - (1-b\eta^{\alpha-2})t^{\alpha-1}}{\Delta} \int\limits_0^1 \Phi_{10}(s)ds$$

$$-\int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} p(s)\Phi(s)ds + \frac{(1-a\xi^{\alpha-2})t^{\alpha-1} + a\xi^{\alpha-1}t^{\alpha-2}}{\Delta} \int\limits_0^1 \frac{(1-s)^{\alpha-1}}{\Gamma(\alpha)} p(s)\Phi(s)ds$$

$$-\frac{b(1-a\xi^{\alpha-2})t^{\alpha-1} + ab\xi^{\alpha-1}t^{\alpha-2}}{\Delta} \int\limits_0^\eta \frac{(\eta-s)^{\alpha-1}}{\Gamma(\alpha)} p(s)\Phi(s)ds$$

$$+\frac{a(1-b\eta^{\alpha-2})t^{\alpha-1} - a(1-b\eta^{\alpha-1})t^{\alpha-2}}{\Delta} \int\limits_0^\xi \frac{(\xi-s)^{\alpha-1}}{\Gamma(\alpha)} p(s)\Phi(s))ds$$

$$\overline{\Psi}(t) = \frac{(1-c\xi^{\beta-2})t^{\beta-1} + c\xi^{\beta-1}t^{\beta-2}}{\nabla} \int\limits_0^1 \Psi_{20}(s)ds$$

$$+\frac{(1-d\eta^{\beta-1})t^{\beta-2} - (1-d\eta^{\beta-2})t^{\beta-1}}{\nabla} \int\limits_0^1 \Phi_{20}(s)ds$$

$$-\int_0^t \frac{(t-s)^{\beta-1}}{\Gamma(\beta)} q(s)\Psi(s)ds + \frac{(1-c\xi^{\beta-2})t^{\beta-1} + c\xi^{\beta-1}t^{\beta-2}}{\nabla} \int\limits_0^1 \frac{(1-s)^{\beta-1}}{\Gamma(\beta)} q(s)\Psi(s)ds$$

$$-\frac{d(1-c\xi^{\beta-2})t^{\beta-1} + cd\xi^{\beta-1}t^{\beta-2}}{\nabla} \int\limits_0^\eta \frac{(\eta-s)^{\beta-1}}{\Gamma(\beta)} q(s)\Psi(s)ds$$

$$+\frac{c(1-d\eta^{\beta-2})t^{\beta-1} - c(1-d\eta^{\beta-1})t^{\beta-2}}{\nabla} \int\limits_0^\xi \frac{(\xi-s)^{\beta-1}}{\Gamma(\beta)} q(s)\Psi(s)ds,$$

$$M_1 = \max\Big\{ \tfrac{\upsilon_1+\mu_1}{\Delta}||\Psi_1||_1 C_1 + \tfrac{\omega_1+\lambda_1}{\Delta}||\Phi_1||_1 B_1$$

$$+b_1 \tfrac{[\Delta+(1+b)(\upsilon_1+\mu_1)+a(\lambda_1+\omega_1)]\mathbf{B}(\alpha+l_1,k_1+1)}{\Gamma(\alpha)\Delta},$$

$$\frac{\upsilon_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)} + \mu_1 \frac{\Gamma(\alpha-1)}{\Gamma(\alpha-m-1)}}{\Delta}||\Psi_1||_1 C_1 + \frac{\omega_1 \frac{\Gamma(\alpha-1)}{\Gamma(\alpha-m-1)} + \lambda_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}}{\Delta}||\Phi_1||_1 B_1$$

$$+a_1 \frac{\Big[\upsilon_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)} + \mu_1 \frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|} + \Big(b\upsilon_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)} + b\mu_1 \frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\Big)\eta^{\alpha+k_1+l_1}\Big]\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}$$

$$+ b_1 \frac{\Big(a\lambda_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)} + a\omega_1 \frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\Big)\xi^{\alpha+k_1+l_1}\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta} + \frac{b_1 \mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha-m)}\Bigg\}$$

$$N_1 = \max\Big\{ \tfrac{\upsilon_1+\mu_1}{\Delta}||\Psi_1||_1 D_1 + \tfrac{\omega_1+\lambda_1}{\Delta}||\Phi_1||_1 A_1$$

$$+a_1 \tfrac{[\Delta+(1+b)(\upsilon_1+\mu_1)+a(\lambda_1+\omega_1)]\mathbf{B}(\alpha+l_1,k_1+1)}{\Gamma(\alpha)\Delta},$$

$$\frac{\upsilon_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)} + \mu_1 \frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}}{\Delta}||\Psi_1||_1 D_1 + \frac{\omega_1 \frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|} + \lambda_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}}{\Delta}||\Phi_1||_1 A_1$$

$$+b_1 \frac{\Big[\upsilon_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)} + \mu_1 \frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|} + \Big(b\upsilon_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)} + b\mu_1 \frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\Big)\eta^{\alpha+k_1+l_1}\Big]\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}$$

$$+ a_1 \frac{\Big(a\lambda_1 \frac{\Gamma(\alpha)}{\Gamma(\alpha-m)} + a\omega_1 \frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\Big)\xi^{\alpha+k_1+l_1}\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta} + \frac{a_1 \mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha-m)}\Bigg\},$$

$$M_2 = \max\Big\{ \tfrac{\upsilon_2+\mu_2}{\nabla}||\Psi_2||_1 C_2 + \tfrac{\omega_2+\lambda_2}{\nabla}||\Phi_2||_1 B_2$$

$$+b_2 \tfrac{[\nabla+(1+d)(\upsilon_2+\mu_2)+c(\lambda_2+\omega_2)]\mathbf{B}(\beta+l_2,k_2+1)}{\Gamma(\beta)\nabla},$$

$$\frac{\upsilon_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)} + \mu_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|}}{\nabla}||\Psi_2||_1 C_2 + \frac{\omega_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-m-1)|} + \lambda_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)}}{\nabla}||\Phi_2||_1 B_2$$

$$+a_2 \frac{\Big[\upsilon_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)} + \mu_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|} + \Big(d\upsilon_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)} + d\mu_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|}\Big)\eta^{\beta+k_2+l_2}\Big]\mathbf{B}(\beta-n+l_2,k_2+1)}{\Gamma(\beta)\nabla}$$

$$+ b_2 \frac{\Big(c\lambda_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)} + c\omega_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|}\Big)\xi^{\beta+k_2+l_2}\mathbf{B}(\beta-n+l_2,k_2+1)}{\Gamma(\beta)\nabla} + \frac{b_2 \mathbf{B}(\beta-n+l_2,k_2+1)}{\Gamma(\beta-n)}\Bigg\}$$

and

$$N_2 = \max\Big\{ \tfrac{\upsilon_2+\mu_2}{\nabla}||\Psi_2||_1 D_2 + \tfrac{\omega_2+\lambda_2}{\nabla}||\Phi_2||_1 A_2$$

$$+a_2 \tfrac{[\nabla+(1+d)(\upsilon_2+\mu_2)+c(\lambda_2+\omega_2)]\mathbf{B}(\beta+l_2,k_2+1)}{\Gamma(\beta)\nabla},$$

$$\frac{\upsilon_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)} + \mu_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|}}{\nabla}||\Psi_2||_1 D_2 + \frac{\omega_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|} + \lambda_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)}}{\nabla}||\Phi_2||_1 A_2$$

$$+b_2 \frac{\Big[\upsilon_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)} + \mu_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|} + \Big(d\upsilon_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)} + d\mu_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|}\Big)\eta^{\beta+k_2+l_2}\Big]\mathbf{B}(\beta-n+l_2,k_2+1)}{\Gamma(\beta)\nabla}$$

$$+ a_2 \frac{\Big(c\lambda_2 \frac{\Gamma(\beta)}{\Gamma(\beta-n)} + c\omega_2 \frac{\Gamma(\beta-1)}{|\Gamma(\beta-n-1)|}\Big)\xi^{\beta+k_2+l_2}\mathbf{B}(\beta-n+l_2,k_2+1)}{\Gamma(\beta)\nabla} + \frac{a_2 \mathbf{B}(\beta-n+l_2,k_2+1)}{\Gamma(\beta-n)}\Bigg\}.$$

Denote

$$M = M_1 + N_1,\ N = M_2 + N_2,$$

$$\Phi_0 = \max\{||\overline{\Phi}||_1, 1\},\ \ \Psi_0 = \max\{||\overline{\Psi}||_1, 1\},$$

$$\tau = \max\{\epsilon_1, \delta_1\},\ \ \sigma = \max\{\sigma_1, \gamma_1\}.$$

3.1. Theorem. *Suppose that $\Delta > 0, \nabla > 0$, $b\eta^{\alpha-1} \le 1$, $d\xi^{\alpha-1} \le 1$, (i)-(iv) defined in Section 1 and (B1) hold. Then BVP(1.8) has at least one positive solution if one of the followings is satisfied:*

(I) $\tau\sigma < 1$

(II) $\tau\sigma = 1$ *with* $NM^{1/\sigma} < 1$ *or* $MN^{1/\tau} < 1$

(III) $\tau\sigma > 1$ *with*

$$\frac{M(\tau\sigma-1)\tau\sigma[M\Psi_0+\Phi_0]^{\tau\sigma-1}}{(\tau\sigma-1)^{\tau\sigma}} \le \frac{1}{N^\sigma}\ \ or\ \ \frac{N(\tau\sigma-1)\tau\sigma[N\Phi_0+\Psi_0]^{\tau\sigma-1}}{(\tau\sigma-1)^{\tau\sigma}} \le \frac{1}{M^\tau}.$$

Proof. From Lemmas 2.6 and 2.7, we know that (x, y) is a solution of BVP(1.8) if and only if (x, y) is a fixed point of T. From Lemma 2.8, $T : X \times Y \to X \times Y$ is completely continuous. By Lemma 2.9 and (i)-(iv), (x, y) is a positive solution of BVP(1.8) if and only if (x, y) is a fixed point of T.

To get a fixed point of T, we apply the Schauder's fixed point theorem. We should define an closed convex bounded subset Ω of E such that $T(\Omega) \subseteq \Omega$.

It is easy to see that $\overline{\Phi} \in X, \overline{\Psi} \in Y$. For $r_1 > 0, r_2 > 0$, denote $\Omega = \{(x, y) \in E : ||x - \overline{\Phi}|| \le r_1, ||y - \overline{\Psi}|| \le r_2\}$. For $(x, y) \in \Omega$, we get

$$||x|| \le ||x - \overline{\Phi}|| + ||\overline{\Phi}|| \le r_1 + ||\overline{\Phi}||,\ \ ||y|| \le ||y - \overline{\Psi}|| + ||\overline{\Psi}|| \le r_2 + ||\overline{\Psi}||. \tag{3.1}$$

Furthermore, we have

$$\left|f\left(t, y(t), D^n_{0+}y(t)\right) - \Phi(t)\right| \le b_1|t^{2-\beta}y(t)|^{\epsilon_1} + a_1|t^{2+n-\beta}D^n_{0+}y(t)|^{\delta_1},$$

$$\left|g\left(t, x(t), D^m_{0+}x(t)\right) - \Psi(t)\right| \le b_2|t^{2-\alpha}x(t)|^{\sigma_1} + a_2|t^{2+m-\alpha}D^m_{0+}x(t)|^{\gamma_1}, ,$$

$$\left|\phi_1\left(t, y(t), D^n_{0+}y(t)\right) - \Phi_{10}(t)\right| \le \Phi_1(t)[B_1|t^{2-\beta}y(t)|^{\epsilon_1} + A_1|t^{2+n-\beta}D^n_{0+}y(t)|^{\delta_1},$$

$$\left|\psi_1\left(t, y(t), D^n_{0+}y(t)\right) - \Psi_{10}(t)\right| \le \Psi_1(t)[C_1|t^{2-\beta}y(t)|^{\epsilon_1} + D_1|t^{2+n-\beta}D^n_{0+}y(t)|^{\delta_1}]$$

$$\left|\phi_2\left(t, x(t), D^m_{0+}x(t)\right) - \Phi_{20}(t)\right| \le \Phi_2(t)[B_2|t^{2-\alpha}x(t)|^{\sigma_1} + A_2|t^{2+m-\alpha}D^m_{0+}x(t)|^{\gamma_1}],$$

$$\left|\psi_2\left(t, x(t), D^m_{0+}x(t)\right) - \Psi_{20}(t)\right| \le \Psi_2(t)[C_2|t^{2-\alpha}x(t)|^{\sigma_1} + D_2|t^{2+m-\alpha}D^m_{0+}x(t)|^{\gamma_1}]$$

hold for all $t \in (0, 1)$. It follows that

$$\left|f\left(t, y(t), D^n_{0+}y(t)\right) - \Phi(t)\right| \le b_1||y||^{\epsilon_1} + a_1||y||^{\delta_1}, t \in (0, 1),$$

$$\left|g\left(t, x(t), D^m_{0+}x(t)\right) - \Psi(t)\right| \le b_2||x||^{\sigma_1} + a_2||x||^{\gamma_1}, t \in (0, 1),$$

$$\left|\phi_1\left(t, y(t), D^n_{0+}y(t)\right) - \Phi_{10}(t)\right| \le \Phi_1(t)[B_1||y||^{\epsilon_1} + A_1||y||^{\delta_1}], t \in (0, 1),$$

$$\left|\psi_1\left(t, y(t), D^n_{0+}y(t)\right) - \Psi_{10}(t)\right| \le \Psi_1(t)[C_1||y||^{\epsilon_1} + D_1||y||^{\delta_1}], t \in (0, 1),$$

$$\left|\phi_2\left(t, x(t), D^m_{0+}x(t)\right) - \Phi_{20}(t)\right| \le \Phi_2(t)[B_2||x||^{\sigma_1} + A_2||x||^{\gamma_1}], t \in (0, 1),$$

$$\left|\psi_2\left(t, x(t), D^m_{0+}x(t)\right) - \Psi_{20}(t)\right| \le \Psi_2(t)[C_2||x||^{\sigma_1} + D_2||x||^{\gamma_1}], t \in (0, 1).$$

By the definition of T, we have

$$t^{2-\alpha}|(T_1y)(t)-\overline{\Phi}(t)|$$

$$\leq \tfrac{\upsilon_1+\mu_1}{\Delta}||\Psi_1||_1[C_1||y||^{\epsilon_1}+D_1||y||^{\delta_1}]+\tfrac{\omega_1+\lambda_1}{\Delta}||\Phi_1||_1[B_1||y||^{\epsilon_1}+A_1||y||^{\delta_1}]$$

$$+t^{2-\alpha}\int_0^t \tfrac{(t-s)^{\alpha-1}}{\Gamma(\alpha)}s^{k_1}(1-s)^{l_1}ds[b_1||y||^{\epsilon_1}+a_1||y||^{\delta_1}]$$

$$+\tfrac{\upsilon_1+\mu_1}{\Delta}\int\limits_0^1 \tfrac{(1-s)^{\alpha-1}}{\Gamma(\alpha)}s^{k_1}(1-s)^{l_1}ds[b_1||y||^{\epsilon_1}+a_1||y||^{\delta_1}]$$

$$+\tfrac{b\upsilon_1+b\mu_1}{\Delta}\int\limits_0^\eta \tfrac{(\eta-s)^{\alpha-1}}{\Gamma(\alpha)}s^{k_1}(1-s)^{l_1}ds[b_1||y||^{\epsilon_1}+a_1||y||^{\delta_1}]$$

$$+\tfrac{a\lambda_1+a\omega_1}{\Delta}\int\limits_0^\xi \tfrac{(\xi-s)^{\alpha-1}}{\Gamma(\alpha)}s^{k_1}(1-s)^{l_1}ds[b_1||y||^{\epsilon_1}+a_1||y||^{\delta_1}]$$

$$\leq \Big(\tfrac{\upsilon_1+\mu_1}{\Delta}||\Psi_1||_1C_1+\tfrac{\omega_1+\lambda_1}{\Delta}||\Phi_1||_1B_1$$

$$+b_1\tfrac{[\Delta+(1+b)(\upsilon_1+\mu_1)+a(\lambda_1+\omega_1)]\mathbf{B}(\alpha+l_1,k_1+1)}{\Gamma(\alpha)\Delta}\Big)||y||^{\epsilon_1}$$

$$+\Big(\tfrac{\upsilon_1+\mu_1}{\Delta}||\Psi_1||_1D_1+\tfrac{\omega_1+\lambda_1}{\Delta}||\Phi_1||_1A_1$$

$$+a_1\tfrac{[\Delta+(1+b)(\upsilon_1+\mu_1)+a(\lambda_1+\omega_1)]\mathbf{B}(\alpha+l_1,k_1+1)}{\Gamma(\alpha)\Delta}\Big)||y||^{\delta_1}$$

and similarly we get

$$t^{2+m-\alpha}|D^m_{0^+}(T_1y)(t)-D^m_{0^+}\overline{\Phi}(t)|$$

$$\le \frac{\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{\Gamma(\alpha-m-1)}}{\Delta}||\Psi_1||_1[C_1||y||^{\epsilon_1}+D_1||y||^{\delta_1}]$$

$$+\frac{\omega_1\frac{\Gamma(\alpha-1)}{\Gamma(\alpha-m-1)}+\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}}{\Delta}||\Phi_1||_1[B_1||y||^{\epsilon_1}+A_1||y||^{\delta_1}]$$

$$+t^{2+m-\alpha}\int_0^t\frac{(t-s)^{\alpha-m-1}}{\Gamma(\alpha-m)}s^{k_1}(1-s)^{l_1}ds[b_1||y||^{\epsilon_1}+a_1||y||^{\delta_1}$$

$$+\frac{\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}}{\Delta}\int_0^1\frac{(1-s)^{\alpha-1}}{\Gamma(\alpha)}s^{k_1}(1-s)^{l_1}ds[b_1||y||^{\epsilon_1}+a_1||y||^{\delta_1}$$

$$+\frac{b\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+b\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}}{\Delta}\int_0^\eta\frac{(\eta-s)^{\alpha-1}}{\Gamma(\alpha)}s^{k_1}(1-s)^{l_1}ds[b_1||y||^{\epsilon_1}+a_1||y||^{\delta_1}$$

$$+\frac{a\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+a\omega_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}}{\Delta}\int_0^\xi\frac{(\xi-s)^{\alpha-1}}{\Gamma(\alpha)}s^{k_1}(1-s)^{l_1}ds[b_1||y||^{\epsilon_1}+a_1||y||^{\delta_1}$$

$$\le\Bigg(\frac{\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{\Gamma(\alpha-m-1)}}{\Delta}||\Psi_1||_1C_1+\frac{\omega_1\frac{\Gamma(\alpha-1)}{\Gamma(\alpha-m-1)}+\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}}{\Delta}||\Phi_1||_1B_1$$

$$+\frac{b_1\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha-m)}$$

$$+a_1\frac{\Big[\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}+\Big(b\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+b\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\Big)\eta^{\alpha+k_1+l_1}\Big]\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}$$

$$+\,b_1\frac{\Big(a\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+a\omega_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\Big)\xi^{\alpha+k_1+l_1}\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}\Bigg)||y||^{\epsilon_1}$$

$$+\Bigg(\frac{\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{\Gamma(\alpha-m-1)}}{\Delta}||\Psi_1||_1D_1+\frac{\omega_1\frac{\Gamma(\alpha-1)}{\Gamma(\alpha-m-1)}+\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}}{\Delta}||\Phi_1||_1A_1$$

$$+\frac{a_1\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha-m)}$$

$$+b_1\frac{\Big[\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}+\Big(b\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+b\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\Big)\eta^{\alpha+k_1+l_1}\Big]\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}$$

$$+\,a_1\frac{\Big(a\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+a\omega_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\Big)\xi^{\alpha+k_1+l_1}\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}\Bigg)||y||^{\delta_1}.$$

We get

$$||T_1y-\overline{\Phi}||\le M_1(r_2+||\overline{\Psi}||_1)^{\epsilon_1}+N_1[r_2+||\overline{\Psi}||_1]^{\delta_1}\le M[r_2+\Psi_0]^{\tau}.$$

Similarly we get

$$||T_2x-\overline{\Psi}||\le M_2[r_1+||\overline{\Phi}||]^{\sigma_1}+N_2[+[r_1+||\overline{\Phi}||]^{\gamma_1}\le N[r_1+\Phi_0]^{\sigma}.$$

If there exists $r_1,r_2>0$ such that

$$M[r_2+\Psi_0]^{\tau}\le r_1,\ \ N[r_1+\Phi_0]^{\sigma}\le r_2, \tag{3.2}$$

we let $\Omega = \{(x,y) \in E : ||x-\Phi_1|| \le r_1, ||y-\Phi_2|| \le r_2\}$, then we get $T(\Omega) \subset \Omega$. Hence the Schauder's fixed point theorem implies that T has a fixed point $(x,y) \in \Omega$. So (x,y) is a solution of BVP(1.8).

Now we will prove that (3.2) has positive solution $r_1, r_2 > 0$. We transform (3.2) to the following inequalities:

$$r_2 \le \left(\tfrac{r_1}{M} - \Psi_0\right)^{1/\sigma},\ r_1 \le \left(\tfrac{r_2}{N} - \Phi_0\right)^{1/\tau}.$$

Hence we get

$$N(r_1+\Phi_0)^\tau \le r_2 \le \left(\tfrac{r_1}{M} - \Psi_0\right)^{1/\sigma}$$

or

$$M(r_2+\Psi_0)^\sigma \le r_1 \le \left(\tfrac{r_2}{N} - \Phi_0\right)^{1/\tau}.$$

Case (i) $\sigma\tau < 1$.

It is easy to see that there exists $r_1 > 0$ sufficiently large such that $N(r_1+\Phi_0)^\tau \le \left(\frac{r_1}{M} - \Psi_0\right)^{1/\sigma}$. Then we can choose r_2 satisfying $N(r_1+\Phi_0)^\tau \le r_2 \le \left(\frac{r_1}{M} - \Psi_0\right)^{1/\sigma}$.

Hence (3.2) has positive solution $r_1 > 0, r_2 > 0$. We choose $\Omega = \{(x,y) \in E : ||x-\overline{\Phi}|| \le r_1, ||y-\overline{\Psi}|| \le r_2\}$. Then we get $T(\Omega) \subset \Omega$. Hence the Schauder's fixed point theorem implies that T has a fixed point $(x,y) \in \Omega$. So (x,y) is a positive solution of BVP(1.8).

Case (ii) $\sigma\tau = 1$.

If $NM^{1/\sigma} < 1$, then

$$\lim_{r\to+\infty} \frac{N(r_1+\Phi_0)^\tau}{\left(\frac{r_1}{M} - \Psi_0\right)^{1/\sigma}} = NM^{1/\sigma} < 1.$$

So there exists $r_1 > 0$ sufficiently large such that $N(r_1+\Phi_0)^\tau \le \left(\frac{r_1}{M} - \Psi_0\right)^{1/\sigma}$. Then we can choose r_2 satisfying $N(r_1+\Phi_0)^\tau \le r_2 \le \left(\frac{r_1}{M} - \Psi_0\right)^{1/\sigma}$.

If $MN^{1/\tau} < 1$, then there exists $r_2 > 0$ sufficiently large such that $M(r_2+\Psi_0)^\sigma \le \left(\frac{r_2}{N} - \Phi_0\right)^{1/\tau}$. Then we can choose r_1 satisfying $M(r_2+\Psi_0)^\sigma \le r_1 \le \left(\frac{r_2}{N} - \Phi_0\right)^{1/\tau}$.

Hence (3.2) has positive solution $r_1 > 0, r_2 > 0$. We choose $\Omega = \{(x,y) \in E : ||x-\overline{\Phi}|| \le r_1, ||y-\overline{\Psi}|| \le r_2\}$. Then we get $T(\Omega) \subset \Omega$. Hence the Schauder's fixed point theorem implies that T has a fixed point $(x,y) \in \Omega$. So (x,y) is a positive solution of BVP(1.8).

Case (iii) $\sigma\tau > 1$.

If

$$\frac{M(\tau\sigma-1)\tau\sigma[M\Psi_0+\Phi_0]^{\tau\sigma-1}}{(\tau\sigma-1)^{\tau\sigma}} \le \frac{1}{N^\sigma},$$

then let $r_1 = \frac{\tau\sigma M\Psi_0+\Phi_0}{\tau\sigma-1}$. It is easy to see that $N(r_1+\Phi_0)^\tau \le \left(\frac{r_1}{M} - \Psi_0\right)^{1/\sigma}$. Then we can choose r_2 satisfying $N(r_1+\Phi_0)^\tau \le r_2 \le \left(\frac{r_1}{M} - \Psi_0\right)^{1/\sigma}$.

If

$$\frac{N(\tau\sigma-1)\tau\sigma[N\Phi_0+\Psi_0]^{\tau\sigma-1}}{(\tau\sigma-1)^{\tau\sigma}} \le \frac{1}{M^\tau},$$

then let $r_2 = \frac{\tau\sigma N\Phi_0+\Psi_0}{\tau\sigma-1}$. It is easy to see that $M(r_2+\Psi_0)^\sigma \le \left(\frac{r_2}{N} - \Phi_0\right)^{1/\tau}$. Then we can choose r_1 satisfying $M(r_2+\Psi_0)^\sigma \le r_1 \le \left(\frac{r_2}{N} - \Phi_0\right)^{1/\tau}$.

Hence (3.2) has positive solution $r_1 > 0, r_2 > 0$. We choose $\Omega = \{(x,y) \in E : ||x-\overline{\Phi}|| \le r_1, ||y-\overline{\Psi}|| \le r_2\}$. Then we get $T(\Omega) \subset \Omega$. Hence the Schauder's fixed point theorem implies that T has a fixed point $(x,y) \in \Omega$. So (x,y) is a positive solution of BVP(1.8).

The proof of Theorem 3.1 is complete. $\square$

3.2. Remark. If (B1) holds with $\max\{\epsilon_1,\delta_1\}\max\{\sigma_1,\gamma_1\}\ge 1$, it is easy to see that all known results in [1, 17] can not be applied to establish existence results for solutions of BVP(1.8). It is easy to see that $\lim M=\lim N=0$ for sufficiently small $a_i,b_i,C_i,D_i,A_i,B_i(i=1,2)$. Then

$$NM^{1/\sigma}<1,\quad MN^{1/\tau}<1,$$

$$\frac{M(\tau\sigma-1)\tau\sigma[M\Psi_0+\Phi_0]^{\tau\sigma-1}}{(\tau\sigma-1)^{\tau\sigma}}\le\frac{1}{N^\sigma}\text{ and }\frac{N(\tau\sigma-1)\tau\sigma[N\Phi_0+\Psi_0]^{\tau\sigma-1}}{(\tau\sigma-1)^{\tau\sigma}}\le\frac{1}{M^\tau}.$$

hold for sufficiently small $a_i,b_i,C_i,D_i,A_i,B_i(i=1,2)$. From Theorem 3.1, BVP(1.8) has at least one solution for $\sigma\tau<1$, and for sufficiently small $a_i,b_i,C_i,D_i,A_i,B_i(i=1,2)$ when $\sigma\tau\ge 1$.

4. Numerical examples

In this section, we present two examples for the illustration of our main result (Theorem 3.1).

4.1. Example. We consider the following boundary value problem

$$\begin{cases} D_{0^+}^{\frac{19}{10}}u(t)+t^{-\frac{1}{10}}(1-t)^{-\frac{21}{20}}f(t,v(t),D_{0^+}^{\frac{13}{20}}v(t))=0, & t\in(0,1),\\ D_{0^+}^{\frac{39}{20}}v(t)+t^{-\frac{1}{10}}(1-t)^{-\frac{23}{20}}g(t,u(t),D_{0^+}^{\frac{4}{5}}u(t))=0, & t\in(0,1),\\ \lim\limits_{t\to0}t^{\frac15}u(t)-\frac12u(1/2)=0,\\ u(1)-\frac12u(3/4)=0,\\ \lim\limits_{t\to0}t^{\frac19}v(t)-\frac12v(1/2)=0.\\ v(1)-\frac12v(3/4)=0,\end{cases}\tag{4.1}$$

Then

(i) BVP(4.1) has at least one positive solution if there exists a constant $H>0$ such that

$$|f(t,u,v)-t^2|\le H,\ \ t\in(0,1),u,v\in\mathbb{R},$$

$$|g(t,u,v)-t^5|\le H,\ \ t\in(0,1),u,v\in\mathbb{R}.$$

(ii) BVP(4.1) has at least one positive solution if

$$|f(t,u,v)-t^2|\le b_1t^{\frac{\epsilon_1}{20}}u^{\epsilon_1},\ b_1\ge0,\ \epsilon_1>0,$$

$$|g(t,u,v)-t^5|\le b_2t^{\frac{\sigma_1}{10}}u^{\sigma_1},\ b_2\ge0,\ \sigma_1>0$$

and one of the followings holds:

(a) $\epsilon_1\sigma_1<1$;

(b) $\epsilon_1\sigma_1=1$ with $(38.1089b_1)^{1/\sigma_1}34.0678b_2<1$ or $38.1089b_1(34.0678b_2)^{1/\tau_1}<1$

(c) $\epsilon_1\sigma_1>1$ with

$$\frac{38.1089b_1(\epsilon_1\sigma_1-1)\epsilon_1\sigma_1[38.1089b_1\Psi_0+\Phi_0]^{\epsilon_1\sigma_1-1}}{(\epsilon_1\sigma_1-1)^{\epsilon_1\sigma_1}}(34.0678b_2)^{\sigma_1}\le1$$

or

$$\frac{34.0678b_2(\epsilon_1\sigma_1-1)\epsilon_1\sigma_1[34.0678b_2\Phi_0+\Psi_0]^{\epsilon_1\sigma_1-1}}{(\epsilon_1\sigma_1-1)^{\epsilon_1\sigma_1}}(38.1089b_1)^{\epsilon_1}\le1.$$

(iii) BVP(4.1) has at least one positive solution if

$$f(t,u,v)=t^2+b_1t^{\frac{\epsilon_1}{20}}u^{\epsilon_1}+a_1t^{\frac{7\delta_1}{10}}v^{\delta_1},\ a_1,b_1\ge0,\ \epsilon_1,\delta_1>0,$$

$$g(t,u,v)=t^5+b_2t^{\frac{\sigma_1}{10}}u^{\sigma_1}+a_2t^{\frac{9\gamma_1}{10}}v^{\gamma_1},\ a_2,b_2\ge0,\ \sigma_1,\gamma_1>0.$$

with $a_i,b_i(i=1,2)$ sufficiently small.

Proof. Corresponding to BVP(1.8), we have $\alpha = \frac{19}{10}, \beta = \frac{39}{20}$, $m = \frac{4}{5}$ and $n = \frac{13}{20}$, $\xi = \frac{1}{2}$, $\eta = \frac{3}{4}$, $a = b = c = d = \frac{1}{2}$ and $\phi_i(t,u,v) = \psi_i(t,u,v) \equiv 0 (i = 1,2)$ and $p(t) = t^{-\frac{1}{10}}(1-t)^{-\frac{21}{20}}$, $q(t) = t^{-\frac{1}{10}}(1-t)^{-\frac{23}{20}}$.

It is easy to see that (i)-(iv) hold with $k_1 = -\frac{1}{10} = k_2$, and $l_1 = -\frac{21}{20}$, $l_2 = -\frac{23}{20}$. One sees that $k_1 > -1$, $\alpha - m + l_1 > 0$, $2 + k_1 + l_1 > 0$, $k_2 > -1$, $\beta - n + l_2 > 0$, $2 + k_2 + l_2 > 0$. One sees that both p and q are not integrable on $(0,1)$. Hence (i)-(iv) defined in Section 1 hold.

By direct calculation using Matlab7, we find that

$$\mu_1 = \frac{1}{2}\left(\frac{1}{2}\right)^{\frac{9}{10}} \approx 0.2679, \quad \upsilon_1 = 1 - \frac{1}{2}\sqrt[10]{2} \approx 0.0670,$$

$$\omega_1 = 1 - \frac{1}{2}\left(\frac{3}{4}\right)^{\frac{9}{10}} \approx 0.6141, \quad \lambda_1 = 1 - \frac{1}{2}\sqrt[10]{2} \approx 0.4641,$$

$$\mu_2 = \frac{1}{2}\left(\frac{1}{2}\right)^{\frac{19}{20}} \approx 0.2588, \quad \upsilon_2 = 1 - \frac{1}{2}\sqrt[20]{2} \approx 0.4828,$$

$$\omega_2 = 1 - \frac{1}{2}\left(\frac{3}{4}\right)^{\frac{19}{20}} \approx 0.6196, \quad \lambda_2 = 1 - \frac{1}{2}\sqrt[20]{2} \approx 0.4824,$$

$$\Delta = \frac{1}{2}\left(\frac{1}{2}\right)^{\frac{9}{10}}\left(1 - \frac{1}{2}\sqrt[10]{2}\right) + \left(1 - \frac{1}{2}\sqrt[10]{2}\right)\left(1 - \frac{1}{2}\left(\frac{3}{4}\right)^{\frac{9}{10}}\right) \approx 0.4093,$$

$$\nabla = \frac{1}{2}\left(\frac{1}{2}\right)^{\frac{19}{20}}\left(1 - \frac{1}{2}\sqrt[20]{2}\right) + \left(1 - \frac{1}{2}\sqrt[20]{2}\right)\left(1 - \frac{1}{2}\left(\frac{3}{4}\right)^{\frac{19}{20}}\right) \approx 0.4237.$$

Hence $\Delta > 0, \nabla > 0$, $b\eta^{\alpha-1} \le 1$, $d\xi^{\alpha-1} \le 1$.

Choose $\Phi(t) = t^2$, $\Psi(t) = t^5$. By direct computation, we find that

$$\begin{aligned}
t^{\frac{1}{10}}|\overline{\Phi}(t)| &= t^{\frac{1}{10}}\left| -\int_0^t \frac{(t-s)^{\frac{9}{10}}}{\Gamma(19/10)} s^{\frac{19}{10}}(1-s)^{-\frac{21}{20}}ds + \frac{\nu_1 t^{\frac{9}{10}} + \mu_1 t^{-\frac{1}{10}}}{\Delta}\frac{\mathbf{B}(17/20,29/10)}{\Gamma(19/10)}\right.\\
&- \frac{\nu_1 t^{\frac{9}{10}} + \mu_1 t^{-\frac{1}{10}}}{2\Delta}\int_0^{3/4}\frac{\left(\frac{3}{4}-s\right)^{\frac{9}{10}}}{\Gamma(19/10)} s^{\frac{19}{10}}(1-s)^{-\frac{21}{20}}ds\\
&\left. + \frac{\lambda_1 t^{\frac{9}{10}} - \omega_1 t^{-\frac{1}{10}}}{2\Delta}\int_0^{1/2}\frac{\left(\frac{1}{2}-s\right)^{\frac{9}{10}}}{\Gamma(19/10)} s^{\frac{19}{10}}(1-s)^{-\frac{21}{20}}ds\right|\\
&\le \left(1 + \frac{3}{2}\frac{\nu_1+\mu_1}{\Delta} + \frac{1}{2}\frac{\lambda_1+\omega_1}{\Delta}\right)\frac{\mathbf{B}(17/20,29/10)}{\Gamma(19/10)} \approx 2.3895,
\end{aligned}$$

$$\begin{aligned}
t^{\frac{1}{20}}|\overline{\Psi}(t)| &= t^{\frac{1}{20}}\left| -\int_0^t \frac{(t-s)^{\frac{19}{20}}}{\Gamma(39/20)} s^{\frac{49}{10}}(1-s)^{-\frac{23}{20}}ds + \frac{\nu_2 t^{\frac{19}{20}} + \mu_2 t^{-\frac{1}{20}}}{\nabla}\frac{\mathbf{B}(4/5,59/10)}{\Gamma(39/20)}\right.\\
&- \frac{\nu_2 t^{\frac{19}{20}} + \mu_2 t^{-\frac{1}{20}}}{2\nabla}\int_0^{3/4}\frac{\left(\frac{3}{4}-s\right)^{\frac{19}{20}}}{\Gamma(39/20)} s^{\frac{49}{10}}(1-s)^{-\frac{23}{20}}ds\\
&\left. + \frac{\lambda_2 t^{\frac{19}{20}} - \omega_2 t^{-\frac{1}{20}}}{2\nabla}\int_0^{1/2}\frac{\left(\frac{1}{2}-s\right)^{\frac{19}{20}}}{\Gamma(39/20)} s^{\frac{49}{10}}(1-s)^{-\frac{23}{20}}ds\right|\\
&\le \left(1 + \frac{3}{2}\frac{\nu_2+\mu_2}{\nabla} + \frac{1}{2}\frac{\lambda_2+\omega_2}{\nabla}\right)\frac{\mathbf{B}(4/5,59/10)}{\Gamma(39/20)} \approx 1.4335,
\end{aligned}$$

and

$$t^{\frac{9}{10}}|D_{0^+}^{\frac{4}{5}}\overline{\Phi}(t)| = t^{\frac{9}{10}}\left| -\int_0^t \frac{(t-s)^{\frac{1}{10}}}{\Gamma(11/10)} s^{\frac{19}{10}}(1-s)^{-\frac{21}{20}}ds\right.$$

$$+\frac{\nu_1\frac{\Gamma(19/10)}{\Gamma(11/10)}t^{\frac{1}{10}}+\mu_1\frac{\Gamma(9/10)}{\Gamma(1/10)}t^{-\frac{9}{10}}}{\Delta}\frac{\mathbf{B}(17/20,29/10)}{\Gamma(19/10)}$$

$$-\frac{\nu_1\frac{\Gamma(19/10)}{\Gamma(11/10)}t^{\frac{1}{10}}+\mu_1\frac{\Gamma(9/10)}{\Gamma(1/10)}t^{-\frac{9}{10}}}{2\Delta}\int_0^{3/4}\frac{\left(\frac{3}{4}-s\right)^{\frac{9}{10}}}{\Gamma(19/10)}s^{\frac{19}{10}}(1-s)^{-\frac{21}{20}}ds$$

$$\left.+\frac{\lambda_1\frac{\Gamma(19/10)}{\Gamma(11/10)}t^{\frac{1}{10}}-\omega_1\frac{\Gamma(9/10)}{\Gamma(1/10)}t^{-\frac{9}{10}}}{2\Delta}\int_0^{1/2}\frac{\left(\frac{1}{2}-s\right)^{\frac{9}{10}}}{\Gamma(19/10)}s^{\frac{19}{10}}(1-s)^{-\frac{21}{20}}ds\right|$$

$$\leq \frac{\mathbf{B}(1/20,29/10)}{\Gamma(11/10)}+\left(\frac{3}{2}\frac{\nu_1\frac{\Gamma(19/10)}{\Gamma(11/10)}+\mu_1\frac{\Gamma(9/10)}{\Gamma(1/10)}}{\Delta}+\frac{1}{2}\frac{\lambda_1\frac{\Gamma(19/10)}{\Gamma(11/10)}+\omega_1\frac{\Gamma(9/10)}{\Gamma(1/10)}}{\Delta}\right)\frac{\mathbf{B}(17/20,29/10)}{\Gamma(19/10)}$$

$$\approx 20.7609,$$

$$t^{\frac{7}{10}}|D_{0^+}^{\frac{13}{20}}\overline{\Psi}(t)| = t^{\frac{7}{10}}\left| -\int_0^t \frac{(t-s)^{\frac{3}{10}}}{\Gamma(13/10)} s^{\frac{49}{10}}(1-s)^{-\frac{23}{20}}ds\right.$$

$$+\frac{\nu_2\frac{\Gamma(39/20)}{\Gamma(13)}t^{\frac{3}{10}}+\mu_2\frac{\Gamma(19/20)}{\Gamma(3/10)}t^{-\frac{7}{10}}}{\nabla}\frac{\mathbf{B}(4/5,59/10)}{\Gamma(39/20)}$$

$$-\frac{\nu_2\frac{\Gamma(39/20)}{\Gamma(13)}t^{\frac{3}{10}}+\mu_2\frac{\Gamma(19/20)}{\Gamma(3/10)}t^{-\frac{7}{10}}}{2\nabla}\int_0^{3/4}\frac{\left(\frac{3}{4}-s\right)^{\frac{19}{20}}}{\Gamma(39/20)}s^{\frac{49}{10}}(1-s)^{-\frac{23}{20}}ds$$

$$\left.+\frac{\lambda_2\frac{\Gamma(39/20)}{\Gamma(13/10)}t^{\frac{3}{10}}-\omega_2\frac{\Gamma(19/20)}{\Gamma(3/10)}t^{-\frac{7}{10}}}{2\nabla}\int_0^{1/2}\frac{\left(\frac{1}{2}-s\right)^{\frac{19}{20}}}{\Gamma(39/20)}s^{\frac{49}{10}}(1-s)^{-\frac{23}{20}}ds\right|$$

$$\leq \frac{\mathbf{B}(3/20,59/10)}{\Gamma(13/10)}+\left(\frac{3}{2}\frac{\nu_2\frac{\Gamma(39/20)}{\Gamma(13/10)}+\mu_2\frac{\Gamma(19/20)}{\Gamma(3/10)}}{\nabla}+\frac{1}{2}\frac{\lambda_2\frac{\Gamma(39/20)}{\Gamma(13/10)}+\omega_2\frac{\Gamma(19/20)}{\Gamma(3/10)}}{\nabla}\right)\frac{\mathbf{B}(4/5,59/10)}{\Gamma(39/20)}$$

$$\approx 6.2585.$$

It is easy to see by calculation that

$$||\overline{\Phi}|| = \max\left\{\sup_{t\in(0,1]} t^{\frac{1}{10}}|\overline{\Phi}(t)|, \sup_{t\in(0,1]} t^{\frac{9}{10}}|D_{0^+}^{\frac{4}{5}}\overline{\Phi}(t)|\right\} \leq 20.7609,$$

$$||\overline{\Psi}|| = \max\left\{\sup_{t\in(0,1]} t^{\frac{1}{20}}|\overline{\Psi}(t)|, \sup_{t\in(0,1]} t^{\frac{7}{10}}|D_{0^+}^{\frac{13}{20}}\overline{\Psi}(t)|\right\} \leq 6.2585.$$

One sees that (B1) holds with $A_i = B_i = C_i = D_i = 0 (i = 1, 2)$, $\Phi_{i0}(t) = \Psi_{i0}(t) = 0 (i = 1, 2)$, $\Phi_i(t) = \Psi_i(t) = 0 (i = 1, 2)$, $\Phi(t) = t^2, \Psi(t) = t^5$.

Furthermore, we have

$$M_1 = \max\Big\{b_1 \tfrac{[\Delta+(1+b)(\upsilon_1+\mu_1)+a(\lambda_1+\omega_1)]\mathbf{B}(\alpha+l_1,k_1+1)}{\Gamma(\alpha)\Delta},$$

$$\tfrac{b_1\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha-m)}$$

$$+a_1\tfrac{\left[\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}+\left(b\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+b\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\right)\eta^{\alpha+k_1+l_1}\right]\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}$$

$$+\ b_1\tfrac{\left(a\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+a\omega_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\right)\xi^{\alpha+k_1+l_1}\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}\Big\}$$

$$\le b_1\Big(\tfrac{[\Delta+(1+b)(\upsilon_1+\mu_1)+a(\lambda_1+\omega_1)]\mathbf{B}(\alpha+l_1,k_1+1)}{\Gamma(\alpha)\Delta}+\tfrac{\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha-m)}$$

$$+\ \tfrac{\left(a\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+a\omega_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\right)\xi^{\alpha+k_1+l_1}\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}\Big)$$

$$+a_1\tfrac{\left[\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}+\left(b\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+b\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\right)\eta^{\alpha+k_1+l_1}\right]\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}$$

$$\le a_1\tfrac{\left[0.0670\frac{\Gamma(19/10)}{\Gamma(11/10)}+0.2679\frac{\Gamma(9/10)}{\Gamma(1/10)}+\left(\frac{0.0670}{2}\frac{\Gamma(19/10)}{\Gamma(11/10)}+\frac{0.2679}{2}\frac{\Gamma(9/10)}{\Gamma(1/10)}\right)\left(\frac{3}{4}\right)^{\frac{3}{4}}\right]\mathbf{B}(1/20,9/10)}{0.4093\Gamma(19/10)}$$

$$+b_1\Big(\tfrac{[0.4093+\frac{3}{2}(0.0670+0.2679)+\frac{1}{2}(0.4641+0.6141)]\mathbf{B}(17/20,9/10)}{0.4093\Gamma(19/10)}+\tfrac{\mathbf{B}(1/20,9/10)}{\Gamma(11/10)}$$

$$+\ \tfrac{\left(\frac{0.4641}{2}\frac{\Gamma(19/10)}{\Gamma(11/10)}+\frac{0.6141}{2}\frac{\Gamma(9/10)}{|\Gamma(1/10)|}\right)\left(\frac{1}{2}\right)^{\frac{3}{4}}\mathbf{B}(1/20,9/10)}{0.4093\Gamma(19/10)}\Big)\approx 6.9793a_1+31.0850b_1,$$

and

$$N_1 = \max\Big\{a_1 \tfrac{[\Delta+(1+b)(\upsilon_1+\mu_1)+a(\lambda_1+\omega_1)]\mathbf{B}(\alpha+l_1,k_1+1)}{\Gamma(\alpha)\Delta},$$

$$\tfrac{a_1\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha-m)}$$

$$+b_1\tfrac{\left[\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}+\left(b\upsilon_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+b\mu_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\right)\eta^{\alpha+k_1+l_1}\right]\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}$$

$$+\ a_1\tfrac{\left(a\lambda_1\frac{\Gamma(\alpha)}{\Gamma(\alpha-m)}+a\omega_1\frac{\Gamma(\alpha-1)}{|\Gamma(\alpha-m-1)|}\right)\xi^{\alpha+k_1+l_1}\mathbf{B}(\alpha-m+l_1,k_1+1)}{\Gamma(\alpha)\Delta}\Big\}$$

$$\le b_1\tfrac{\left[0.0670\frac{\Gamma(19/10)}{\Gamma(11/10)}+0.2679\frac{\Gamma(9/10)}{\Gamma(1/10)}+\left(\frac{0.0670}{2}\frac{\Gamma(19/10)}{\Gamma(11/10)}+\frac{0.2679}{2}\frac{\Gamma(9/10)}{\Gamma(1/10)}\right)\left(\frac{3}{4}\right)^{\frac{3}{4}}\right]\mathbf{B}(1/20,9/10)}{0.4093\Gamma(19/10)}$$

$$+a_1\Big(\tfrac{[0.4093+\frac{3}{2}(0.0670+0.2679)+\frac{1}{2}(0.4641+0.6141)]\mathbf{B}(17/20,9/10)}{0.4093\Gamma(19/10)}+\tfrac{\mathbf{B}(1/20,9/10)}{\Gamma(11/10)}$$

$$+\ \tfrac{\left(\frac{0.4641}{2}\frac{\Gamma(19/10)}{\Gamma(11/10)}+\frac{0.6141}{2}\frac{\Gamma(9/10)}{\Gamma(1/10)|}\right)\left(\frac{1}{2}\right)^{\frac{3}{4}}\mathbf{B}(1/20,9/10)}{0.4093\Gamma(19/10)}\Big)\approx 34.1691a_1+7.0239b_1,$$

$$M_2 \le a_2 \frac{\left[0.4828\frac{\Gamma(39/20)}{\Gamma(13/10)}+0.2588\frac{\Gamma(19/20)}{\Gamma(3/10)}+\left(\frac{0.4828}{2}\frac{\Gamma(39/20)}{\Gamma(13/10)}+\frac{0.2588}{2}\frac{\Gamma(19/20)}{\Gamma(3/10)}\right)\left(\frac{3}{4}\right)^{\frac{7}{10}}\right]\mathbf{B}(3/20,9/10)}{0.4237\Gamma(39/20)}$$

$$+b_2\left(\frac{[0.4237+\frac{3}{2}(0.4828+0.2588)+\frac{1}{2}(0.4824+0.6196)]\mathbf{B}(4/5,9/10)}{0.4237\Gamma(39/20)}+\frac{\mathbf{B}(3/10,9/10)}{0.4237\Gamma(13/10)}\right.$$

$$\left.+\frac{\left(\frac{0.4824}{2}\frac{\Gamma(39/20)}{\Gamma(13//10)}+\frac{0.6196}{2}\frac{\Gamma(19/20)}{\Gamma(3/10)}\right)\left(\frac{1}{2}\right)^{\frac{7}{10}}\mathbf{B}(3/20,9/10)}{0.4237\Gamma(39/20)}\right)\approx 14.2808a_2+19.7870b_2,$$

and

$$N_2 \le b_2 \frac{\left[0.4828\frac{\Gamma(39/20)}{\Gamma(13/10)}+0.2588\frac{\Gamma(19/20)}{\Gamma(3/10)}+\left(\frac{0.4828}{2}\frac{\Gamma(39/20)}{\Gamma(13/10)}+\frac{0.2588}{2}\frac{\Gamma(19/20)}{\Gamma(3/10)}\right)\left(\frac{3}{4}\right)^{\frac{7}{10}}\right]\mathbf{B}(3/20,9/10)}{0.4237\Gamma(39/20)}$$

$$+a_2\left(\frac{[0.4237+\frac{3}{2}(0.4828+0.2588)+\frac{1}{2}(0.4824+0.6196)]\mathbf{B}(4/5,9/10)}{0.4237\Gamma(39/20)}+\frac{\mathbf{B}(13/20,9/10)}{\Gamma(13/10)}\right.$$

$$\left.+\frac{\left(\frac{0.4828}{2}\frac{\Gamma(39/20)}{\Gamma(13/10)}+\frac{0.6196}{2}\frac{\Gamma(19/20)}{\Gamma(3/10)}\right)\left(\frac{1}{2}\right)^{\frac{7}{10}}\mathbf{B}(3//20,9/10)}{0.4237\Gamma(39/20)}\right)\approx 12.4878a_2+14.2808b_2.$$

So

$$M = M_1+N_1 \le 41.1484a_1+38.1089b_1, \quad N = M_2+N_2 \le 26.7686a_2+34.0678b_2,$$

$$\Phi_0 = \max\{||\overline{\Phi}||_1, 1\} \le 20.7609, \quad \Psi_0 = \max\{||\overline{\Psi}||_1, 1\} \le 6.2585,$$

$$\tau = \max\{\epsilon_1, \delta_1\}, \quad \sigma = \max\{\sigma_1, \gamma_1\}.$$

(i) If there exists a constant $H > 0$ such that

$$|f(t,u,v) - t^2| \le H, \quad t \in (0,1), u, v \in \mathbb{R},$$

$$|g(t,u,v) - t^5| \le H, \quad t \in (0,1), u, v \in \mathbb{R},$$

then we can choose $\epsilon_1 = \delta_1 = \sigma_1 = \gamma_1 = 0$. So $\tau\sigma = 0 < 1$. Then (i) in Theorem 3.1 implies that BVP(4.1) has at least one positive solution.

(ii) If there exist constants $b_1, b_2 \ge 0$ and $\epsilon_1, \sigma_1 > 0$ such that

$$|f(t,u,v) - t^2| \le b_1 t^{\frac{\epsilon_1}{20}} u^{\epsilon_1},$$

$$|g(t,u,v) - t^5| \le b_2 t^{\frac{\sigma_1}{10}} u^{\sigma_1},$$

we can choose $\delta_1 = \gamma_1 = 0$, $a_1 = a_2 = 0$. Theorem 3.1 implies that BVP(4.1) has at least one solution if one of the followings holds:

(a) $\epsilon_1\sigma_1 < 1$;

(b) $\epsilon_1\sigma_1 = 1$ with $(38.1089b_1)^{1/\sigma_1}34.0678b_2 < 1$ or $38.1089b_1(34.0678b_2)^{1/\tau_1} < 1$

(c) $\epsilon_1\sigma_1 > 1$ with

$$\frac{38.1089b_1(\epsilon_1\sigma_1-1)\epsilon_1\sigma_1[238.5046b_1+20.7069]^{\epsilon_1\sigma_1-1}}{(\epsilon_1\sigma_1-1)^{\epsilon_1\sigma_1}}(34.0678b_2)^{\sigma_1} \le 1$$

or

$$\frac{34.0678b_2(\epsilon_1\sigma_1-1)\epsilon_1\sigma_1[707.2782b_2+6.2585]^{\epsilon_1\sigma_1-1}}{(\epsilon_1\sigma_1-1)^{\epsilon_1\sigma_1}}(38.1089b_1)^{\epsilon_1} \le 1.$$

(iii) If

$$f(t,u,v) = t^2 + b_1 t^{\frac{\epsilon_1}{20}} u^{\epsilon_1} + a_1 t^{\frac{7\delta_1}{10}} v^{\delta_1}, \ a_1, b_1 \ge 0, \ \epsilon_1, \delta_1 > 0,$$

$$g(t,u,v) = t^5 + b_2 t^{\frac{\sigma_1}{10}} u^{\sigma_1} + a_2 t^{\frac{9\gamma_1}{10}} v^{\gamma_1}, \ a_2, b_2 \ge 0, \ \sigma_1, \gamma_1 > 0,$$

then Theorem 3.1 implies that BVP(4.1) has at least one positive solution if one of the followings holds:

(a) $\tau\sigma < 1$

(b) $\tau\sigma = 1$ with $926.7686a_2+34.0678b_2)(41.1484a_1+38.1089b_1)^{1/\sigma} < 1$ or $(41.1484a_1+38.1089b_1)926.7686a_2+34.0678b_2)^{1/\tau} < 1$

(c) $\tau\sigma > 1$ with

$$\frac{(41.1484a_1+38.1089b_1)(\tau\sigma-1)\tau\sigma[6.2585(41.1484a_1+38.1089b_1)+20.7690]^{\tau\sigma-1}}{(\tau\sigma-1)^{\tau\sigma}}$$

$$\leq \frac{1}{926.7686a_2+34.0678b_2)^{\sigma}}$$

$$\text{or } \frac{(26.7686a_2+34.0678b_2)(\tau\sigma-1)\tau\sigma[20.7690(26.7686a_2+34.0678b_2)+6.2585]^{\tau\sigma-1}}{(\tau\sigma-1)^{\tau\sigma}}$$

$$\leq \frac{1}{(41.1484a_1+38.1089b_1)^{\tau}}.$$

□

4.2. Remark. Since both p and q are not measurable on $(0,1)$, we know that all known results in [1, 17] can not be applied to establish existence results for solutions of BVP(4.1). Hence Theorem 3.1 fills a gap not covered by [1, 17].

4.3. Example. We consider the following boundary value problem

$$(4.2)\qquad \begin{cases} D_{0^+}^{\frac{19}{10}}u(t) + t^{-\frac{1}{2}}(1-t)^{-\frac{1}{5}}f(t, v(t), D_{0^+}^{\frac{39}{40}}v(t)) = 0, \quad t \in (0,1), \\ D_{0^+}^{\frac{39}{20}}v(t) + t^{-\frac{1}{2}}(1-t)^{\frac{1}{10}}g(t, u(t), D_{0^+}^{\frac{19}{20}}u(t)) = 0, \quad t \in (0,1), \\ \lim_{t\to 0} t^{\frac{1}{5}}u(t) - \frac{1}{2}u(1/2) = 0, \\ u(1) - \frac{1}{2}u(3/4) = 0, \\ \lim_{t\to 0} t^{\frac{1}{9}}v(t) - \frac{1}{2}v(1/2) = 0, \\ v(1) - \frac{1}{2}v(3/4) = 0, \end{cases}$$

where

$$f(t,u,v) = t^2 + b_1tu^{\epsilon_1} + a_1tv^{\delta_1},\ {}_1, b_1 \geq 0,\ \epsilon_1, \delta_1 > 0,$$

$$g(t,u,v) = 4t^5 + b_2tu^{\sigma_1} + a_2tv^{\gamma_1},\ a_2, b_2 \geq 0,\ \sigma_1, \gamma_1 > 0.$$

Then BVP(4.2) has at least one positive solution for sufficiently small $a_i, b_i (i = 1, 2)$.

Proof. Corresponding to BVP(1.8), we have $\alpha = \frac{19}{10}, \beta = \frac{39}{20}$, $m = \frac{19}{20}$ and $n = \frac{39}{40}$, $a = b = c = d = \frac{1}{2}$ and $\phi_i(t,u,v) = \psi_i(t,u,v) \equiv 0 (i = 1, 2)$ and $p(t) = t^{-\frac{1}{2}}(1-t)^{-\frac{6}{5}}$, $q(t) = t^{-\frac{1}{2}}(1-t)^{\frac{10}{9}}$.

It is easy to see that (i)-(iv) hold with $k_1 = -\frac{1}{10} = k_2$, and $l_1 = -\frac{1}{5}$, $l_2 = -\frac{1}{10}$. One sees that $k_1 > -1$, $\alpha - m + l_1 > 0$, $2 + k_1 + l_1 > 0$, $k_2 > -1$, $\beta - n + l_2 > 0$, $2 + k_2 + l_2 > 0$. One sees $m > \alpha - 1$, $n > \beta - 1$.

Then Theorem 3.1 implies that BVP(4.2) has at least one positive solution if one of the followings is satisfied:

(I) $\tau\sigma < 1$

(II) $\tau\sigma \geq 1$ for sufficiently small $a_i, b_i (i = 1, 2)$. □

Acknowledgments

The author would like to thank the referees and editors for their careful reading and some kindly comments on improving the presentation of this paper.

References

[1] Ahmad, A. and Nieto, J.J. *Existence results for a coupled system of nonlinear fractional differential equations with three-point boundary conditions*, Comput. Math. Appl. **58**, 1838-1843, 2009.

[2] Ahmad, B. and Sivasundaram, S. *On four-point nonlocal boundary value problems of nonlinear integro-differential equations of fractional order*, Appl. Math. Comput. **217**, 480-487, 2010.

[3] Avery R.I., and Peterson, A.C. *Three Positive Fixed Points of Nonlinear Operators on Ordered Banach Spaces*, Comput. Math. Appl. **42**, 313-322, 2001.

[4] Basset, A.B. *On the descent of a sphere in a vicous liquid*, Q. J. Pure Appl. Math. **41**, 369-381, 1910.

[5] Bai, C. and Fang, J. *The existence of a positive solution for a singular coupled system of nonlinear fractional differential equations*, Appl. Math. Comput. **150**, (3), 611-621, 2004.

[6] Caponetto, R., Dongona, G. and Fortuna, L. *Fractional Differential Systems, Modeling and control applications*, World Scientific Series on Nonlinear Science, Ser. A, **72**, World Scientific Publishing Co. Pte. Ltd. 2010.

[7] Duan, J. and Temuer, C. *Solution for system of linear fractional differential equations with constant coefficients*, Journal of Mathematics, **29**, 599-603, 2009.

[8] Gaber, M. and Brikaa, M.G. *Existence results for a coupled system of nonlinear fractional differential equation with four-point boundary conditions*, ISRN Mathematical Analysis, **2011**, Article ID 468346, 14 pages, 2011.

[9] Goodrich, C.S. *Existence of a positive solution to systems of differential equations of fractional order*, Comput. Math. Appl. **62**, 1251-1268, 2011.

[10] Liu, B. *Positive solutions of a nonlinear four-point boundary value problems in Banach spaces*, J. Math. Anal. Appl. **305**, 253-276, 2005.

[11] Liu, L., Zhang, X. and Wu, Y. *On existence of positive solutions of a two-point boundary value problem for a nonlinear singular semipositone system*, Appl. Math. Comput. **192**, 223-232, 2007.

[12] Mainardi, F. *Fraction Calculus: Some basic problems in continuum and statistical machanics*, In: A. Carpinteri, F. Mainardi (eds.) Fratals and Fractional Calculus in Continuum Machanics, (Springer, Vien, 1997), 291-348.

[13] Miller, K.S. and Ross, B. *An Introduction to the Fractional Calculus and Fractional Differential Equation*, (Wiley, New York, 1993.)

[14] Mamchuev, M.O. *Boundary value problem for a system of fractional partial differential equations*, Partial Differential Equations, **44**, 1737-1749, 2008.

[15] Rehman, M. and Khan, R. *A note on boundaryvalueproblems for a coupled system of fractionaldifferential equations*, Comput. Math. Appl. **61**, 2630-2637, 2011.

[16] Wang, J., Xiang, H. and Liu, Z. *Positive Solution to Nonzero Boundary Values Problem for a Coupled System of Nonlinear Fractional Differential Equations*, International Journal of Differential Equations, **2010**, Article ID 186928, 12 pages, 2010.

[17] Su, X. *Boundary value problem for a coupled system of nonlinear fractional differential equations*, Applied Mathematics Letters **22** (1), 64-69, 2009.

[18] Torvik, P.J., and Bagley, R.L. *On the appearance of the fractional derivative in the behavior of real materials*, J. Appl. Mech. **51**, 294-298, 1984.

[19] Trujillo, J.J., Rivero, M. and Bonilla, B. *On a Riemann-Liouville generalized Taylor's formula*, J. Math. Anal. Appl. **231**, 255-265, 1999.

[20] Yuan, C. *Multiple positive solutions for (n-1, 1)-type semipositone conjugate boundary value problems for coupled systems of nonlinear fractional differential equations*, E. J. Qualitative Theory of Diff. Equ. **13** 1-12, 2012.

[21] Yang, A. and Ge, W. *Positive solutions for boundary value problems of N-dimension nonlinear fractional differential systems,* Boundary Value Problems **2008**, article ID 437453, 2008.

[22] Yuan, C., Jiang, D., O′regan, D. and Agarwal, R.P. *Existence and uniqueness of positive solutions of boundary value problems for coupled systems of singular second-order three-point non-linear differential and difference equations*, Appl. Anal. **87** 921-932, 2008.

[23] Yuan, C., Jiang, D., O′regan, D. and Agarwal, R.P. *Multiple positive solutions to systems of nonlinear semipositone fractional differential equations with coupled boundary conditions*, E. J. Qualitative Theory of Diff. Equ. **13** (2012), 1-17, 2012.

[24] Zou, Y., Liu, L. and Cui, Y. *The existence of solutions for four-point coupled boundary value problems of fractional differential equations at resonance*, Abstract and Applied Analysis **2014**, Article ID 314083, 8 pages, 2014.

[25] Ji, D. and Ge, W. *Positive solution for four-point nonlocal boundary value problems of fractional order*, Mathematical Methods in the Applied Sciences **37** (8), 1232-1239, 2014.

[26] Zhao, J., Geng, F. and Ge, W. *Positive solutions to a new kind Sturm-Liouville-like four-point boundary value problem*, Applied Mathematics and Computation **217** (2), 811-819, 2010.

[27] Zhao, X., Chai, C. and Ge, W. *Positive solutions for fractional four-point boundary value problems*, Comm. Nonl. Sci. Numer. Simul. **16** (9), 3665-3672, 2011.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 417 – 427

Oscillation criteria for solutions to nonlinear dynamic equations of higher order

Donal O'Regan* and Taher S. Hassan† ‡§

Abstract

In this paper using some new dynamic inequalities we present some oscillation results for higher order dynamic equation

$$\left\{r_{n-1}(t)\ \phi_{\alpha_{n-1}}\left[(r_{n-2}(t)(...(r_1(t)\phi_{\alpha_1}[x^{\Delta}(t)])^{\Delta}...)^{\Delta})^{\Delta}\right]\right\}^{\Delta} + p\,(t)\,\phi_{\gamma}\,(x\,(g\,(t))) = 0,$$

on an unbounded time scale $\mathbb{T}$. Some new oscillation criteria are obtained using comparison techniques. Some applications illustrating our results are included.

Keywords: Asymptotic behavior, oscillation, higher order, dynamic equations, dynamic inequality, time scales.

2000 AMS Classification: 34K11, 39A10, 39A99.

Received : 05.07.2014 *Accepted* : 12.03.2015 *Doi* : 10.15672/HJMS.20164512495

1. Introduction

This paper considers the oscillatory behavior of the higher order dynamic equation

(1.1)
$$\left\{r_{n-1}(t)\ \phi_{\alpha_{n-1}}\left[(r_{n-2}(t)(...(r_1(t)\phi_{\alpha_1}[x^{\Delta}(t)])^{\Delta}...)^{\Delta})^{\Delta}\right]\right\}^{\Delta} + p\,(t)\,\phi_{\gamma}\,(x\,(g\,(t))) = 0,$$

on an unbounded time scale $\mathbb{T}$, where $\phi_\alpha(u) := |u|^{\alpha-1}\,u$, γ, $\alpha_i > 0$, $i = 1, 2, \ldots, n-1$, r_i, $i = 1, 2, \ldots, n-1$, are positive rd-continuous functions on $\mathbb{T}$, p is a positive rd-continuous function on $\mathbb{T}$, and $g : \mathbb{T} \to \mathbb{T}$ is a rd-continuous function such that $\lim_{t\to\infty} g(t) = \infty$.

*School of Mathematics, Statistics and Applied Mathematics, National University of Ireland, Galway, Ireland, Email: `donal.oregan@nuigalway.ie`

†Department of Mathematics, Faculty of Science, University of Hail, Hail 2440, Saudi Arabia,

‡Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt, Email: `tshassan@mans.edu.eg`

§Corresponding Author.

We recall that a solution x of equation (1.1) is said to be nonoscillatory if there exists $t_0 \in \mathbb{T}$ such that $x(t)x(\sigma(t)) > 0$ for all $t \in [t_0, \infty)_{\mathbb{T}}$; otherwise, it is said to be oscillatory. Equation (1.1) is said to be oscillatory if all its solutions are oscillatory.

In the literature many papers discuss the behavior of solutions for certain classes of dynamic equations; we refer the reader to [1, 3, 5, 9, 11, 12, 13, 15, 18, 19, 20, 21, 23, 24, 25, 26, 27, 29] and the references cited therein. In particular these papers present oscillatory criteria and asymptotic behavior for first, second and third order dynamic equations on time scales and some interesting results were obtained for special cases of (1.1); see [10, 14, 16, 17, 28].

The aim of this paper is to present some new criteria for equation (1.1) . Our approach is to reduce the problem so that specific oscillation results for first, second and third order dynamic equations can be used for the arbitrary higher order case.

The paper will have four sections. In section 2, we state and prove some new dynamic inequalities. Section 3 uses comparison ideas to discuss (1.1). The last section illustrates the main results of our paper.

The theory of time scales was introduced by Stefan Hilger in his Ph. D. Thesis in 1988 in order to unify continuous and discrete analysis, see [22]. A time scale $\mathbb{T}$ is an arbitrary closed subset of the reals, and the cases when this time scale is equal to the reals or to the integers represent the classical theories of differential and of difference equations. Many other interesting time scales exist, and they give rise to many applications (see [6]). This new theory of these so-called "dynamic equations" not only unifies the corresponding theories for the differential equations and difference equations cases, but it also extends these classical cases to cases "in between". That is, we are able to treat the so-called $q-$difference equations when $\mathbb{T} = q^{\mathbb{N}_0} := \{q^n : n \in \mathbb{N}_0 \text{ for } q > 1\}$ and can be applied to different types of time scales like $\mathbb{T} = h\mathbb{N}$, $\mathbb{T} = \mathbb{N}^2$ and $\mathbb{T} = \mathbb{T}_n$ the set of the harmonic numbers. The books on the subject of time scales by Bohner and Peterson [6], [7] summarizes and organizes much of time scale calculus.

For completeness, we recall some concepts on time scales. For $t \in \mathbb{T}$, we define the forward and backward jump operators $\sigma : \mathbb{T} \to \mathbb{T}$ and $\rho : \mathbb{T} \to \mathbb{T}$ by

$$\sigma(t) = \inf\{s \in \mathbb{T} : s > t\}, \ \rho(t) = \sup\{s \in \mathbb{T} : s < t\},$$

where $\inf \emptyset := \sup \mathbb{T}$ and $\sup \emptyset = \inf \mathbb{T}$, where $\emptyset$ denotes the empty set. A point $t \in \mathbb{T}$, $t > \inf \mathbb{T}$, is said to be left-dense if $\rho(t) = t$, right-dense if $t < \sup \mathbb{T}$ and $\sigma(t) = t$, left-scattered if $\rho(t) < t$ and right-scattered if $\sigma(t) > t$. A function $h : \mathbb{T} \to \mathbb{R}$ is said to be right-dense continuous (rd-continuous) provided that h is continuous at right-dense points and at left-dense points in $\mathbb{T}$, left-hand limits exist and are finite. The set of all such rd-continuous functions is denoted by $C_{rd}(\mathbb{T})$. Let $f : \mathbb{T} \to \mathbb{R}$ and let $t \in \mathbb{T}$. If there exists a number $\alpha \in \mathbb{R}$ such that for all $\epsilon > 0$ there exists a neighborhood U of t such that

$$|f(\sigma(t)) - f(s) - \alpha(\sigma(t) - s)| \leq \epsilon |\sigma(t) - s| \qquad \text{for all } s \in U,$$

then f is said to be differentiable at t, and we call α the delta derivative of f at t and denote it by $f^{\Delta}(t)$.

2. Dynamic Inequalities

In this section we state and prove some dynamic inequalities which will be used in the next section. Throughout this paper, we let

$$x \in C^1_{rd}([t_0, \infty)_{\mathbb{T}}, \mathbb{R}) \text{ for some } t_0 \in [0, \infty)_{\mathbb{T}},$$

and

$$x^{[i]} \in C^1_{rd}([t_0, \infty)_{\mathbb{T}}, \mathbb{R}),\ i = 1, ..., n,$$

where

$$x^{[i]}(t) := r_i(t)\ \phi_{\alpha_i}\left[\left(x^{[i-1]}(t)\right)^{\Delta}\right] \text{ with } r_n = \alpha_n = 1 \text{ and } x^{[0]} = x, \tag{2.1}$$

and $\phi_{\alpha_i}(u) := |u|^{\alpha_i - 1}\, u$, $\alpha_i > 0$, $i = 1, \ldots, n-1$, are constants, and $r_i \in C_{rd}([t_0,\infty)_{\mathbb{T}},(0,\infty))$ $i = 1, ..., n-1$, such that

$$\int_{t_0}^{\infty} r_i^{-1/\alpha_i}(s)\Delta s = \infty,\ \ i = 1, ..., n-1. \tag{2.2}$$

2.1. Lemma. *Let*

$$x(t) > 0 \quad \textit{and} \quad x^{[n]}(t) < 0 \tag{2.3}$$

eventually. Then there exists an integer $m \in \{0, \ldots, n\}$ *with* $m+n$ *odd such that*

$$x^{[k]}(t) > 0 \quad \textit{for} \quad k = 0, \ldots, m, \tag{2.4}$$

and

$$(-1)^{m+k}\, x^{[k]}(t) > 0 \quad \textit{for} \quad k = m, ..., n, \tag{2.5}$$

eventually.

Proof. Let

$$x(t) > 0 \quad \text{and} \quad x^{[n]}(t) < 0 \quad \text{for } t \in [t_0,\infty)_{\mathbb{T}}. \tag{2.6}$$

This implies that $x^{[i]}(t),\ i = 1, ..., n-1$, are eventually monotone and hence are of one sign. There are two possibilities:

(a) $x^{[k]}(t)$ and $x^{[k-1]}(t)$ have opposite signs eventually for $k = 1, \ldots, n$; or
(b) there exists a largest $m \in \{1, \ldots, n\}$ such that $x^{[m]}(t)x^{[m-1]}(t) > 0$ eventually.

If (a) holds, then (2.4) and (2.5) hold with $m = 0$ (note that for this case from (2.6) n must be odd).

Assume that (b) holds with $x^{[m]}(t) < 0$ and $x^{[m-1]}(t) < 0$ for $t \geq t_1$, where $t_1 \in [t_0,\infty)_{\mathbb{T}}$. Then

$$\begin{aligned} x^{[m-2]}(t) &= x^{[m-2]}(t_1) + \int_{t_1}^{t} \phi^{-1}_{\alpha_{m-1}}\left[x^{[m-1]}(s)\right] r_{m-1}^{-1/\alpha_{m-1}}(s)\,\Delta s \\ &< x^{[m-2]}(t_1) + \phi^{-1}_{\alpha_{m-1}}\left[x^{[m-1]}(t_1)\right] \int_{t_1}^{t} r_{m-1}^{-1/\alpha_{m-1}}(s)\,\Delta s. \end{aligned}$$

From (2.2) with $i = m-1$, $\lim_{t\to\infty} x^{[m-2]}(t) = -\infty$. Hence $x^{[m-2]}(t) < 0$ eventually. By the same reasoning we see that $x^{[k]}(t) < 0$ eventually for $k = m-2, \ldots, 0$. This contradicts the assumption that $x(t)$ is eventually positive.

Assume that (b) holds with $x^{[m]}(t) > 0$ and $x^{[m-1]}(t) > 0$ eventually. Using an argument similar to the above, we see that $x^{[k]}(t) > 0$ eventually for $k = m-2, \ldots, 0$. Therefore, (2.4) and (2.5) hold with this m (From (2.5) (with $k = n$) we find that $m+n$ is an odd number). □

Let

$$\alpha[h,k] := \begin{cases} \alpha_h \cdots \alpha_k & h \leq k, \\ 1, & h > k, \end{cases}$$

and for a fixed $m \in \{0, \ldots, n-1\}$ and an integer $k \in \{m, ..., n-1\}$, define the functions $R_{i,j}(v, u)$, $j = 0, ..., k$ by the recurrence formula:

$$R_{k,j}(v,u) := \begin{cases} 1, & j = 0, \\ \int_u^v \left[\frac{R_{k,j-1}(v,s)}{r_{k-j+1}(s)}\right]^{1/\alpha_{k-j+1}} \Delta s, & j = 1, ..., k-m+1, \\ \int_u^v \left[\frac{R_{k,j-1}(s,u)}{r_{k-j+1}(s)}\right]^{1/\alpha_{k-j+1}} \Delta s, & j = k-m+2, ..., k. \end{cases}$$

2.2. Lemma. *Assume that (2.2) and (2.3) hold and $m \in \{0, \ldots, n\}$ is given in Lemma 2.1 such that (2.4) and (2.5) hold for $t \geq t_1 \in [t_0, \infty)_{\mathbb{T}}$. Then the following hold for $v \geq u \in [t_1, \infty)_{\mathbb{T}}$:*

(a) *for $j = m, \ldots, k$,*

(2.7) $$(-1)^{m+j}\, x^{[j]}(u) \geq (-1)^{m+k} \phi^{-1}_{\alpha[j+1,k]} \left(x^{[k]}(v)\right) R_{k,k-j}(v,u);$$

(b) *if $m \geq 1$, then for $j = 0, ..., m-1$,*

(2.8) $$x^{[j]}(v) \geq (-1)^{m+k} \phi^{-1}_{\alpha[j+1,k]} \left(x^{[k]}(v)\right) R_{k,k-j}(v,u),$$

where $k \in \{m, ..., n-1\}$.

Proof. (a) From (2.5), we have that $(-1)^{m+k} x^{[k]}$, $k = m, ..., n-1$, are positive, decreasing on $[t_1, \infty)_{\mathbb{T}}$. This shows that (2.7) holds for $j = k$. Then for $v \geq u \in [t_1, \infty)_{\mathbb{T}}$,

$$\begin{aligned} (-1)^{m+k} x^{[k]}(u) &\geq (-1)^{m+k} x^{[k]}(v) \\ &= (-1)^{m+k} \phi^{-1}_{\alpha[k+1,k]} \left(x^{[k]}(v)\right) R_{k,0}(v,u), \end{aligned}$$

which implies

$$(-1)^{m+k} \left(x^{[k-1]}(u)\right)^{\Delta} \geq (-1)^{m+k} \phi^{-1}_{\alpha[k,k]} \left(x^{[k]}(v)\right) \left(\frac{R_{k,0}(v,u)}{r_k(u)}\right)^{1/\alpha_k}.$$

Replacing u by s in the above inequality and then integrating it from u to $v \in [u, \infty)_{\mathbb{T}}$, we obtain that

$$\begin{aligned} & (-1)^{m+k}\, x^{[k-1]}(v) - (-1)^{m+k}\, x^{[k-1]}(u) \\ \geq\; & (-1)^{m+k} \phi^{-1}_{\alpha[k,k]} \left(x^{[k]}(v)\right) \int_u^v \left(\frac{R_{k,0}(v,s)}{r_k(s)}\right)^{1/\alpha_k} \Delta s \\ =\; & (-1)^{m+k} \phi^{-1}_{\alpha[k,k]} \left(x^{[k]}(v)\right) R_{k,1}(v,u). \end{aligned}$$

From (2.5), we obtain

$$\begin{aligned} (-1)^{m+k-1}\, x^{[k-1]}(u) &\geq (-1)^{m+k}\, x^{[k-1]}(v) - (-1)^{m+k}\, x^{[k-1]}(u) \\ &\geq (-1)^{m+k} \phi^{-1}_{\alpha[k,k]} \left(x^{[k]}(v)\right) R_{k,1}(v,u). \end{aligned}$$

This shows that (2.7) holds for $j = k-1$. Assume that (2.7) holds for some $j \in \{m+1, \ldots, k-1\}$. Then for $v \geq u \in [t_1, \infty)_{\mathbb{T}}$,

$$(-1)^{m+j}\, x^{[j]}(u) \geq (-1)^{m+k} \phi^{-1}_{\alpha[j+1,k]} \left(x^{[k]}(v)\right) R_{k,k-j}(v,u),$$

which implies

$$(-1)^{m+j} \left(x^{[j-1]}(u)\right)^{\Delta} \geq (-1)^{m+k} \phi^{-1}_{\alpha[j,k]} \left(x^{[k]}(v)\right) \left(\frac{R_{k,k-j}(v,u)}{r_j(u)}\right)^{1/\alpha_j}.$$

Replacing u by s in the above inequality and then integrating it from u to $v \in [u,\infty)_{\mathbb{T}}$, we have

$$\begin{aligned} & (-1)^{m+j}\, x^{[j-1]}(v) - (-1)^{m+j}\, x^{[j-1]}(u) \\ \geq\ & (-1)^{m+k} \phi^{-1}_{\alpha[j,k]} \left(x^{[k]}\left(v\right)\right) \int_u^v \left(\frac{R_{k,k-j}(v,s)}{r_j\left(s\right)} \right)^{1/\alpha_j} \Delta s \\ =\ & (-1)^{m+k} \phi^{-1}_{\alpha[j,k]} \left(x^{[k]}\left(v\right)\right) R_{k,k-j+1}(v,u). \end{aligned}$$

Then from (2.5), we have

$$(-1)^{m+j-1}\, x^{[j-1]}(u) \geq (-1)^{m+k} \phi^{-1}_{\alpha[j,k]} \left(x^{[k]}\left(v\right)\right) R_{k,k-j+1}(v,u).$$

This shows that (2.7) holds for $j-1$. By induction, (2.7) holds for all $j = m, m+1, \ldots, k$.

(b) From Part (a) we have that for $j = m$

$$x^{[m]}(u) \geq (-1)^{m+k} \phi^{-1}_{\alpha[m+1,k]} \left(x^{[k]}\left(v\right)\right) R_{k,k-m}(v,u),$$

which implies

$$\left(x^{[m-1]}(u) \right)^{\Delta} \geq (-1)^{m+k} \phi^{-1}_{\alpha[m,k]} \left(x^{[k]}\left(v\right)\right) \left(\frac{R_{k,k-m}(v,u)}{r_m(u)} \right)^{1/\alpha_m}.$$

Replacing u by s in the above inequality and then integrating it from u to $v \in [u,\infty)_{\mathbb{T}}$, we have

$$\begin{aligned} x^{[m-1]}(v) &\geq x^{[m-1]}(v) - x^{[m-1]}(u) \\ &\geq (-1)^{m+k} \phi^{-1}_{\alpha[m,k]} \left(x^{[k]}\left(v\right)\right) \int_u^v \left(\frac{R_{k,k-m}(v,s)}{r_m(s)} \right)^{1/\alpha_m} \Delta s \\ &= (-1)^{m+k} \phi^{-1}_{\alpha[m,k]} \left(x^{[k]}\left(v\right)\right) R_{k,k-m+1}(v,u). \end{aligned}$$

This shows that (2.8) holds for $j = m-1$. Assume that (2.8) holds for some $j \in \{1,\ldots,m-1\}$. Then for $v \geq u \in [t_1,\infty)_{\mathbb{T}}$,

$$\left(x^{[j-1]}(v) \right)^{\Delta} \geq (-1)^{m+k} \phi^{-1}_{\alpha[j,k]} \left(x^{[k]}\left(v\right)\right) \left(\frac{R_{k,k-j}(v,u)}{r_j\left(v\right)} \right)^{1/\alpha_j}.$$

Replacing v by s in the above inequality and then integrating it from u to $v \in [u,\infty)_{\mathbb{T}}$, we have

$$\begin{aligned} x^{[j-1]}(v) &\geq x^{[j-1]}(v) - x^{[j-1]}(u) \\ &\geq \int_u^v (-1)^{m+k} \phi^{-1}_{\alpha[j,k]} \left(x^{[k]}\left(s\right)\right) \left(\frac{R_{k,k-j}(s,u)}{r_j\left(s\right)} \right)^{1/\alpha_j} \Delta s \\ &\geq (-1)^{m+k} \phi^{-1}_{\alpha[j,k]} \left(x^{[k]}\left(v\right)\right) \int_u^v \left(\frac{R_{k,k-j}(s,u)}{r_j\left(s\right)} \right)^{1/\alpha_j} \Delta s \\ &= (-1)^{m+k-1} \phi^{-1}_{\alpha[j,k]} \left(x^{[k]}\left(v\right)\right) R_{k,k-j+1}(v,u). \end{aligned}$$

This shows that (2.8) holds for $j-1$. By induction, (2.8) holds for all $j = 0, 1, \ldots, m-1$. $\square$

3. Main Results

In this section we consider the asymptotic behavior of solutions of the nth-order nonlinear dynamic equation (1.1). From (2.1), Eq. (1.1) can be written as

$$x^{[n]}(t) + p\,(t)\,\phi_\gamma\,(x\,(g\,(t))) = 0. \tag{3.1}$$

3.1. Theorem. *Assume that $n \in 2\mathbb{N}$ and (2.2) holds. If for an integer $k \in \{m, ..., n-1\}$,*

$$(-1)^{m+k}\{r_n\phi_{\alpha_n}[(r_{n-1}(...(r_{k+1}\phi_{\alpha_{k+1}}[z^\Delta])^\Delta...)^\Delta)^\Delta]\}(t) + P_k\,(t)\,\phi_{\gamma/\alpha[1,k]}(z(g(t))) \le 0, \tag{3.2}$$

where for sufficiently large $T \in [t_0,\infty)_{\mathbb{T}}$, $P_k(t) := p\,(t)\,R^\gamma_{k,k}(g(t),T)$, has no eventually positive solution, then every solution of Eq. (3.1) is oscillatory.

Proof. Assume that Eq. (3.1) has a nonoscillatory solution $x(t)$. Then without loss of generality, assume that $x\,(g\,(t)) > 0$ on $[t_0,\infty)_{\mathbb{T}}$. From (3.1), we have that for $t \in [t_0,\infty)_{\mathbb{T}}$,

$$x^{[n]}(t) = -p\,(t)\,\phi_\gamma\,(x\,(g\,(t))) < 0.$$

This implies that $x^{[i]}(t)$, $i = 1, 2, ..., n-1$, are eventually monotone and hence are of one sign. It follows from Lemma 2.1 that there exists an odd integer $m \in \{1, \ldots, n\}$ such that (2.4) and (2.5) hold for $t \ge t_1 \in [t_0,\infty)_{\mathbb{T}}$. From Lemma 2.2, Part (b) with $j = 0$, we get for $v \ge u \in [t_1,\infty)_{\mathbb{T}}$,

$$x(v) \ge (-1)^{m+k}\phi^{-1}_{\alpha[1,k]}\left(x^{[k]}\,(v)\right)R_{k,k}(v,u).$$

Setting $v = g(t)$ and $u = t_1$ gives

$$x(g(t)) \ge (-1)^{m+k}\phi^{-1}_{\alpha[1,k]}\left(x^{[k]}\,(g(t))\right)R_{k,k}(g(t),t_1).$$

Therefore (3.1) becomes

$$\begin{aligned} -x^{[n]}(t) &= p\,(t)\,\phi_\gamma\,(x\,(g\,(t))) \\ &\ge p\,(t)\,R^\gamma_{k,k}(g(t),t_1)\ \phi_{\gamma/\alpha[1,k]}\left((-1)^{m+k}x^{[k]}\,(g(t))\right) \\ &= P_k\,(t)\ \phi_{\gamma/\alpha[1,k]}\left((-1)^{m+k}x^{[k]}\,(g(t))\right), \end{aligned}$$

or

$$(-1)^{m+k}\{r_n\phi_{\alpha_n}[(r_{n-1}(...(r_{k+1}\phi_{\alpha_{k+1}}[z^\Delta])^\Delta...)^\Delta)^\Delta]\}(t) + P_k\,(t)\,\phi_{\gamma/\alpha[1,k]}(z(g(t))) \le 0,$$

where $z(t) := (-1)^{m+k}x^{[k]}\,(t) > 0$, for an integer $k \in \{m, ..., n-1\}$. Thus (3.2) has an eventually positive solution, a contradiction. □

3.2. Theorem. *Assume that $n \in 2\mathbb{N}-1$ and (2.2) holds. If (3.2) for an integer $k \in \{m, ..., n-1\}$ has no eventually positive solution and there is a function τ such that $g(t) \le \tau(t)$ on $[t_0,\infty)_{\mathbb{T}}$ and*

$$(-1)^{k}\{r_n\phi_{\alpha_n}[(r_{n-1}(...(r_{k+1}\phi_{\alpha_{k+1}}[z^\Delta])^\Delta...)^\Delta)^\Delta]\}(t) + Q_k(t)\phi_{\gamma/\alpha[1,k]}(z(\tau(t))) \le 0, \tag{3.3}$$

for an integer $k \in \{0, ..., n-1\}$, where $Q_k(t) := p\,(t)\,R^\gamma_{k,k}(\tau(t),g(t))$, has no eventually positive solution, then every solution of Eq. (3.1) is oscillatory.

Proof. Assume that Eq. (3.1) has a nonoscillatory solution $x(t)$. Then without loss of generality, assume that $x(g(t)) > 0$ on $[t_0, \infty)_{\mathbb{T}}$. From (3.1), we have that for $t \in [t_0, \infty)_{\mathbb{T}}$,

$$x^{[n]}(t) = -p(t)\,\phi_\gamma(x(g(t))) < 0. \tag{3.4}$$

This implies that $x^{[i]}(t),\ i = 1, 2, ..., n-1$, are eventually monotone and hence are of one sign. It follows from Lemma 2.1 that there exists an even integer $m \in \{0, \dots, n\}$ such that (2.4) and (2.5) hold for $t \geq t_1 \in [t_0, \infty)_{\mathbb{T}}$.

(i) Assume that $m \geq 1$. Then the same argument as in the proof of Theorem 3.1 leads to a contradiction.

(ii) Assume that $m = 0$. From Lemma 2.2, Part (a) with $j = m = 0$, we get for $v \geq u \in [t_1, \infty)_{\mathbb{T}}$,

$$x(u) \geq (-1)^k \phi^{-1}_{\alpha[1,k]}\left(x^{[k]}(v)\right) R_{k,k}(v, u).$$

Setting $u = g(t)$ and $v = \tau(t)$ gives

$$x(g(t)) \geq \phi^{-1}_{\alpha[1,k]}\left((-1)^k x^{[k]}(\tau(t))\right) R_{k,k}(\tau(t), g(t)).$$

Therefore (3.1) becomes

$$\begin{aligned} -x^{[n]}(t) &= p(t)\,\phi_\gamma(x(g(t))) \\ &\geq p(t)\,R^{\gamma}_{k,k}(\tau(t), g(t))\phi_{\gamma/\alpha[1,k]}\left((-1)^k x^{[k]}(\tau(t))\right) \\ &= Q_k(t)\phi_{\gamma/\alpha[1,k]}\left((-1)^k x^{[k]}(\tau(t))\right), \end{aligned}$$

or

$$(-1)^k\{r_n\phi_{\alpha_n}[(r_{n-1}(...(r_{k+1}\phi_{\alpha_{k+1}}[z^\Delta])^\Delta...)^\Delta)^\Delta]\}(t)$$
$$+Q_k(t)\phi_{\gamma/\alpha[1,k]}(z(\tau(t))) \leq 0,$$

where $z(t) := (-1)^k x^{[k]}(t) > 0$, for an integer $k \in \{0, ..., n-1\}$. Thus (3.3) has an eventually positive solution, a contradiction. □

For further discussion, we introduce the following notation: For any $t \in \mathbb{T}$, define

$$p_j(t) := \begin{cases} p(t), & j = 0, \\ \left[\dfrac{1}{r_{n-j}(t)}\int_t^\infty p_{j-1}(s)\Delta s\right]^{1/\alpha_{n-j}}, & j = 1, 2, \dots, n-1, \end{cases}$$

provided that the improper integrals involved are convergent.

3.3. Theorem. *Assume that $n \in 2\mathbb{N} - 1$, (2.2) and*

$$\int_{t_0}^\infty p_{n-1}(s)\Delta s = \infty, \tag{3.5}$$

hold. If (3.2) for an integer $k \in \{m, ..., n-1\}$ has no eventually positive solution, then every solution of Eq. (3.1) is oscillatory or tends to zero eventually.

Proof. Assume that Eq. (3.1) has a nonoscillatory solution $x(t)$. Then without loss of generality, assume that $x(g(t)) > 0$ on $[t_0, \infty)_{\mathbb{T}}$. It follows from Lemma 2.1 and Theorem 3.2 that there exists an odd integer $m \in \{1, \dots, n\}$ such that (2.4) and (2.5) hold for $t \geq t_1 \in [t_0, \infty)_{\mathbb{T}}$.

(i) Assume that $m \geq 1$. Then the same argument as in the proof of Theorem 3.1 leads to a contradiction.

(ii) Assume that $m = 0$. Since $x^{\Delta} < 0$ eventually, then $\lim_{t\to\infty} x(t) = l \geq 0$. Assume that $l > 0$. Then

$$x(t),\ x(g(t)) > l_1 \quad \text{for } t \geq t_2 \in [t_1, \infty)_{\mathbb{T}}.$$

Integrating (3.1) from t to $v \in [t, \infty)_{\mathbb{T}}$ and using (2.5) we get that

$$\begin{aligned} x^{[n-1]}(t) &\geq -x^{[n-1]}(v) + x^{[n-1]}(t) \\ &= \int_t^v p(s)\,\phi_\gamma(x(g(s)))\,\Delta s \geq c\int_t^v p(s)\,\Delta s. \end{aligned}$$

By taking limits as $v \to \infty$ we have

$$x^{[n-1]}(t) \geq c\int_t^\infty p(s)\,\Delta s = c\int_t^\infty p_0(s)\,\Delta s.$$

Thus

$$\left(x^{[n-2]}(t)\right)^{\Delta} \geq c^{1/\alpha_{n-1}}\left[\frac{1}{r_{n-1}(t)}\int_t^\infty p_0(s)\,\Delta s\right]^{1/\alpha_{n-1}} = c^{1/\alpha_{n-1}}\ p_1(t). \tag{3.6}$$

Integrating the inequality (3.6) from t to $v \in [t, \infty)_{\mathbb{T}}$ and then taking limits as $v \to \infty$ and using the fact $x^{[n-2]} < 0$ eventually, we get

$$\begin{aligned} -x^{[n-2]}(t) &\geq c^{1/\alpha_{n-1}}\int_t^\infty p_1(s)\Delta s \\ &= c^{1/\alpha[n-1,n-1]}\int_t^\infty p_1(s)\Delta s. \end{aligned}$$

Continuing this process, we get

$$-x^{[1]}(t) \geq c^{1/\alpha[2,n-1]}\int_t^\infty p_{n-2}(s)\Delta s,$$

which implies

$$-x^{\Delta}(t) > c^{1/\alpha[1,n-1]}\left[\frac{1}{r_1(t)}\int_t^\infty p_{n-2}(s)\Delta s\right]^{1/\alpha_1} = c^{1/\alpha[1,n-1]}p_{n-1}(t).$$

Again, integrating the above inequality from t_2 to $t \in [t_2, \infty)_{\mathbb{T}}$ and noting that $x > 0$ eventually, we get

$$x(t_2) - x(t) \geq c^{1/\alpha[1,n-1]}\int_{t_2}^t p_{n-1}(s)\Delta s.$$

Using (3.5), we have $\lim_{t\to\infty} x(t) = -\infty$, which contradicts the fact that $x > 0$ on $[t_2, \infty)_{\mathbb{T}}$. Therefore $\lim_{t\to\infty} x(t) = 0$. This completes the proof. □

4. Applications

As direct consequences of Theorems 3.1, 3.2 and 3.3, we obtain the following comparison criteria for Eq. (3.1) when $k = n - 1$.

4.1. Corollary. *Assume that* (2.2) *holds and the first order dynamic inequality*

$$z^{\Delta}(t) + P_{n-1}(t)\,\phi_{\gamma/\alpha[1,n-1]}(z(g(t))) \leq 0, \tag{4.1}$$

where for sufficiently large $T \in [t_0, \infty)_{\mathbb{T}}$, $P_{n-1}(t) := p(t)\,R^{\gamma}_{n-1,n-1}(g(t), T)$, *has no eventually positive solution.*

(i) *If* $n \in 2\mathbb{N}$, *then every solution of Eq.* (3.1) *is oscillatory.*

(ii) *If $n \in 2\mathbb{N}-1$ and the first order dynamic inequality*

$$z^{\Delta}(t) + Q_{n-1}(t)\phi_{\gamma/\alpha[1,n-1]}(z(\tau(t))) \leq 0, \tag{4.2}$$

where $Q_n(t) := p(t)\, R^{\gamma}_{n-1,n-1}(\tau(t), g(t))$, has no eventually positive solution, then every solution of Eq. (3.1) *is oscillatory.*

4.2. Corollary. *Assume that* (2.2) *holds and the first order dynamic inequality* (4.1) *has no eventually positive solution.*

(i) *If $n \in 2\mathbb{N}$, then every solution of Eq.* (3.1) *is oscillatory.*
(ii) *If $n \in 2\mathbb{N}-1$ and* (3.5) *holds, then every solution of Eq.* (3.1) *is oscillatory or tends to zero eventually.*

Using the main results of [29, 5] we get the following oscillation criteria of Eq. (3.1).

4.3. Corollary. *Let $\gamma = \alpha[1, n-1]$, and $g(t) < t$ and $\tau(t) < t$ on $[t_0, \infty)_{\mathbb{T}}$. Assume that* (2.2) *holds and*

$$\limsup_{t\to\infty} \sup_{\lambda \in E_1} \{\lambda\, e_{-\lambda P_{n-1}}(t, g(t))\} < 1, \tag{4.3}$$

where

$$E_1 = \{\lambda : \lambda > 0,\ 1 - \lambda\, P_{n-1}(t)\mu(t) > 0,\ t \in \mathbb{T}\}.$$

(i) *If $n \in 2\mathbb{N}$, then every solution of Eq.* (3.1) *is oscillatory.*
(ii) *If $n \in 2\mathbb{N}-1$ and*

$$\limsup_{t\to\infty} \sup_{\lambda \in E_2} \{\lambda\, e_{-\lambda Q_{n-1}}(t, \tau(t))\} < 1,$$

where

$$E_2 = \{\lambda : \lambda > 0,\ 1 - \lambda\, Q_{n-1}(t)\mu(t) > 0,\ t \in \mathbb{T}\},$$

then every solution of Eq. (3.1) *is oscillatory.*

4.4. Corollary. *Let $\gamma = \alpha[1, n-1]$, and $g(t) < t$ and $\tau(t) < t$ on $[t_0, \infty)_{\mathbb{T}}$. Assume that* (2.2) *and* (4.3) *hold.*

(i) *If $n \in 2\mathbb{N}$, then every solution of Eq.* (3.1) *is oscillatory.*
(ii) *If $n \in 2\mathbb{N}-1$ and* (3.5) *holds, then every solution of Eq.* (3.1) *is oscillatory or tends to zero eventually.*

4.5. Corollary. *Let $\mathbb{T} = \mathbb{R}$, $\gamma = \alpha[1, n-1]$, and $g(t) < t$ and $\tau(t) < t$ on $[t_0, \infty)$. Assume that* (2.2) *holds and*

$$\liminf_{t\to\infty} \int_{g(t)}^{t} P_{n-1}(s)ds > \frac{1}{e}, \tag{4.4}$$

(i) *If $n \in 2\mathbb{N}$, then every solution of Eq.* (3.1) *is oscillatory.*
(ii) *If $n \in 2\mathbb{N}-1$ and*

$$\liminf_{t\to\infty} \int_{\tau(t)}^{t} Q_{n-1}(s)ds > \frac{1}{e}$$

then every solution of Eq. (3.1) *is oscillatory.*

4.6. Corollary. *Let $\mathbb{T} = \mathbb{R}$, $\gamma = \alpha[1, n-1]$, and $g(t) < t$ and $\tau(t) < t$ on $[t_0, \infty)$. Assume that* (2.2) *and* (4.4) *hold.*

(i) *If $n \in 2\mathbb{N}$, then every solution of Eq.* (3.1) *is oscillatory.*
(ii) *If $n \in 2\mathbb{N}-1$ and* (3.5) *holds, then every solution of Eq.* (3.1) *is oscillatory or tends to zero eventually.*

4.7. Corollary. *Let* $\mathbb{T} = \mathbb{Z}$, $\gamma = \alpha\,[1, n-1]\,,$ *and* $g(n) = \tau(n) = n - k$ *for* $n \in \mathbb{Z}$ *and* $k \in \mathbb{N}$. *Assume that* (2.2) *holds and*

$$\liminf_{t\to\infty} \sum_{i=n-k}^{n-1} P_{n-1}(i) > \left(\frac{k}{k+1}\right)^{k+1}, \tag{4.5}$$

(i) *If* $n \in 2\mathbb{N}$, *then every solution of Eq.* (3.1) *is oscillatory.*
(ii) *If* $n \in 2\mathbb{N} - 1$ *and*

$$\sum_{i=n-k}^{n-1} Q_{n-1}(i) > \left(\frac{k}{k+1}\right)^{k+1},$$

then every solution of Eq. (3.1) *is oscillatory.*

4.8. Corollary. *Let* $\mathbb{T} = \mathbb{Z}$, $\gamma = \alpha\,[1, n-1]\,,$ *and* $g(n) = \tau(n) = n - k$ *for* $n \in \mathbb{Z}$ *and* $k \in \mathbb{N}$. *Assume that* (2.2) *and* (4.5) *hold.*

(i) *If* $n \in 2\mathbb{N}$, *then every solution of Eq.* (3.1) *is oscillatory.*
(ii) *If* $n \in 2\mathbb{N} - 1$ *and* (3.5) *holds, then every solution of Eq.* (3.1) *is oscillatory or tends to zero eventually.*

4.9. Remark. (1) For more oscillation criteria, see [4, 5, 8, 29].
(2) When $n = 3$, the result in Corollary 4.3 is related to a problem posed in [3, Remark 3.3] when $\tau(t) < t$ for $t \geq t_0 \in \mathbb{T}$.

Acknowledgement.

(1) This work was supported by Research Deanship of Hail University under grant No. 0150287.
(2) The authors would like to thank the referee for many helpful suggestions.

References

[1] Adıvar, M., Akın E. and Higgins R. *Oscillatory behavior of solutions of third-order delay and advanced dynamic equations*, Journal of Inequalities and Applications **2014**, 2014:95, 16 pp.
[2] Grover, L.K. and Kaur, P. *An improved estimator of the finite population mean in simple random sampling*, Model Assisted Statistics and Applications **6** (1), 47-55, 2011.
[3] Agarwal, R.P., Bohner, M., Li, T. and Zhang, C. *Hille and Nehari type criteria for third-order delay dynamic equations*, Journal of Difference Equations and Applications **19** (10), 1563-1579, 2013.
[4] Binggen, Z., Xinzhou, Y. and Xueyan, L. *Oscillation criteria of certain delay dynamic equations on time scales*, Journal of Difference Equations and Applications **11** (10), 933-946, 2005.
[5] Bohner, M. *Some oscillation criteria for first order delay dynamic equations*, Far East J. Appl. Math. **18** (3), 289-304, 2005.
[6] Bohner, M. and Peterson, A. Dynamic Equations on Time Scales: An Introduction with Applications, Birkhäuser, Boston, 2001.
[7] Bohner, M. and Peterson, A., editors, Advances in Dynamic Equations on Time Scales, Birkhäuser, Boston, 2003.
[8] Bohner, M., Karpuz, B. and Öcalan, Ö. *Iterated oscillation criteria for delay dynamic equations of first order*, Advances in Difference Equations Volume **2008**, Article ID 458687, 12 pp.
[9] Chen, D. *Oscillation and asymptotic behavior of solutions of certain third-order nonlinear delay dynamic equations*, Theoretical Mathematics & Applications **3**, 19-33, 2013.
[10] Erbe, L., Baoguo and Peterson, A. Oscillation of nth order superlinear dynamic equations on time scales, Rocky Mountain Journal of Mathematics **41** (2), 471-491, 2011.

[11] Erbe, L., Hassan, T.S., Peterson, A. and Saker, S.H. *Interval oscillation criteria for forced second-order nonlinear delay dynamic equations with oscillatory potential*, Dynamics of Continuous, Discrete and Impulsive Systems Series A: Mathematical Analysis **17**(4), 533-542, 2010.

[12] Erbe, L., Hassan, T.S. and Peterson, A. *Oscillation of second order functional dynamic equations*, International Journal of Difference Equations **5**(2), 175-193, 2010.

[13] Erbe, L., Karpuz, B. and Peterson, A. *Kamenev-type oscillation criteria for higher order neutral delay dynamic equations*, International Journal Difference Equations **6** (1), 1-16, 2011.

[14] Erbe, L., Mert, R., Peterson, A. and Zafer, A. *Oscillation of even order nonlinear delay dynamic equations on time scales*, Czechoslovak Mathematical Journal **63** (138), 265-279, 2013.

[15] Erbe, L. and Hassan, T.S. *New oscillation criteria for second order sublinear dynamic equations*, Dynamic Systems and Applications **22** (1), 49-63, 2013.

[16] Grace, S.R. and Hassan, T.S. *Oscillation criteria for higher order nonlinear dynamic equations*, Mathematische Nachrichten **287** (14-15), 1659-1673, 2014.

[17] Grace, S.R., Agarwal, R.P. and Zafer, A. *Oscillation of higher order nonlinear dynamic equations on time scales*, Advances in Difference Equations **2012**, 2012:67, 18 pp.

[18] Han, Z., Li, T., Sun S. and Zhang, M. *Oscillation behavior of solutions of third-order nonlinear delay dynamic equations on time scales*, Communications of the Korean Mathematical Society **26** (3), 499-513, 2011.

[19] Hassan, T.S. *Oscillation criteria for second-order nonlinear dynamic equations*, Advances in Difference Equations **2012:171**, 13 pp, 2012.

[20] Hassan, T.S. *Interval oscillation for second order nonlinear differential equations with a damping term*, Serdica Mathematical Journal **34** (4), 715-732, 2008.

[21] Hassan, T.S. *Asymptotic behavior of solutions of second order nonlinear dynamic equations*, Dynamic Systems and Applications **23** (2-3), 179-187, 2014.

[22] Hilger, S. *Analysis on measure chains — a unified approach to continuous and discrete calculus*, Results in Mathematics **18** (1-2), 18-56, 1990.

[23] Karpuz, B. *Unbounded oscillation of higher-order nonlinear delay dynamic equations of neutral type with oscillating coefficients*, Electronic Journal of Qualitative Theory of Differential Equations **2009**, No. 34, 14 pp.

[24] Kiguradze, I. T. *On oscillatory solutions of some ordinary differential equations*, Soviet Mathematics Doklady 144, 33-36, 1962.

[25] Mert, R. *Oscillation of higher order neutral dynamic equations on time scales*, Advances in Difference Equations **2012**, 2012:68, 11 pp.

[26] Saker, S.H. *Oscillation criteria of second-order half-linear dynamic equations on time scales*, Journal of Computational and Applied Mathematics **177** (2), 375-387, 2005.

[27] Sun, Y. and Hassan, T.S. *Comparison criteria for odd order forced nonlinear functional neutral dynamic equations*, Applied Mathematics and Computation, **251**, 387–395, 2015.

[28] Sun, T., Yu, W. and Xi, H. *Oscillatory behavior and comparison for higher order nonlinear dynamic equations on time scales*, Journal of Applied Mathematics & Informatics **30** (1-2), 289-304, 2012.

[29] Zhang, B. G. and Deng, X. *Oscillation of delay differential equations on time scales*, Mathematical and Computer Modeling **36** (11-13), 1307–1318, 2002.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 429 – 446

On second-order linear recurrent functions with period k and proofs to two conjectures of Sroysang

Julius Fergy Tiongson Rabago*

Abstract

Let w be a real-valued function on $\mathbb{R}$ and k be a positive integer. If for every real number $x, w(x+2k) = rw(x+k) + sw(x)$ for some non-negative real numbers r and s, then we call such function a *second-order linear recurrent function with period* k. Similarly, we call a function $w : \mathbb{R} \to \mathbb{R}$ satisfying $w(x+2k) = -rw(x+k) + sw(x)$ an *odd second-order linear recurrent function with period* k. In this work, we present some elementary properties of these type of functions and develop the concept using the notion of f-even and f-odd functions discussed in [9]. We also investigate the products and quotients of these functions and provide in this work a proof of the conjecture of B. Sroysang which he posed in [19]. In fact, we offer here a proof of a more general case of the problem. Consequently, we present findings that confirm recent results in the theory of Fibonacci functions [9] and contribute new results in the development of this topic.

Keywords: Second-order linear recurrent functions with period k, Horadam numbers, generalized Fibonacci sequences, Pell functions with period k, Jacobsthal functions with period k, Sroysang's conjecture.

2000 AMS Classification: Primary: 39A10, Secondary: 11B37.

Received : 21.09.2014 *Accepted* : 06.03.2015 *Doi* : 10.15672/HJMS.20164512497

*Department of Mathematics and Computer Science, College of Science, University of the Philippines Baguio, Governor Pack Road, Baguio City 2600, Benguet, PHILIPPINES
Email: `jtrabago@upd.edu.ph, jfrabago@gmail.com`

1. Introduction

The *Fibonacci numbers* F_n are defined by $F_0 = 0, F_1 = 1$, and for $n \geq 2, F_{n+1} = F_n + F_{n-1}$. Since the introduction of these numbers, different generalizations have been formulated and were extensively studied. In fact, several books were written solely to study the properties of Fibonacci numbers (*see* for instance [5] and [21]). In a book by T. Koshy [11], any sequences G_n, where $G_1 = a, G_2 = b$, and $G_n = G_{n-1} + G_{n-2}, n \geq 3$ is called the *generalized Fibonacci sequence* (GFS). In [10], A. F. Horadam defined a second-order linear recurrence sequence $\{W_n\}$ by the recurrence relation

$$W_0 = a, \quad W_1 = b, \quad W_{n+1} = rW_n + sW_{n-1}, \quad (n \geq 2).$$

The sequence $\{W_n\}$ can be viewed easily as a certain generalization of $\{F_n\}$. It is now known in literature as *Horadam's sequence*. For a good survey paper regarding Horadam numbers, we refer the readers to [12] (see also [13] for a survey update and extensions). The n^{th} Horadam number W_n with initial conditions $W_0 = 0$ and $W_1 = 1$ can be represented by the following Binet's formula:

$$W_n(0, 1; r, s) = \frac{\alpha^n - \beta^n}{\alpha - \beta}, \quad (n \geq 2)$$

where α and β are the roots of the quadratic equation $x^2 - rx - s = 0$, i.e. $\alpha = (r + \sqrt{r^2 + 4s})/2$ and $\beta = (r - \sqrt{r^2 + 4s})$.

In [15], the author presented a formula for solving the missing terms of $\{W_n\}$ given its first term and last term. Another generalization of Fibonacci numbers is the so-called Fibonacci polynomials (*see* [1] and [4] and the references therein). Recently, Fibonacci numbers were involved in the study of difference and differential equations. Particularly, in [20], D. T. Tollu, Y. Yazlik, and N. Taskara investigated the solutions of two special types of the Riccati difference equation

$$x_{n+1} = \frac{1}{1 + x_n} \quad \text{and} \quad y_{n+1} = \frac{1}{-1 + y_n}$$

such that their solutions are associated with Fibonacci numbers. Another interesting investigation, which involves the Fibonnaci numbers, is presented in [6] where A. Hakami found an application of Fibonacci numbers in the study of continued fractions. This work of Hakami has been recently generalized, to some extent, by the author in [17]. Meanwhile, the author and J. B. Bacani consider in [2] the system

$$x_{n+1} = \frac{q}{p + x_n^{\nu}} \quad \text{and} \quad y_{n+1} = \frac{q}{-p + y_n^{\nu}} \quad (p, q \in \mathbb{R}^+ \text{ and } \nu \in \mathbb{N})$$

as a generalization of Tollu et al.'s work [20]. One particular result established in [2] is the solution form of the above system. In fact, it was shown that every solution of the system, for any arbitrary given set of initial values, is expressible in terms of Horadam numbers. In an earlier paper, the author [16], studied homogeneous differential equations of the form

$$w^{(2k)}(x) = rw^{(k)}(x) + sw(x),$$

where $r, s \in \mathbb{R}^+$ and $w^{(k)}$ is the k^{th} derivative of w with respect to x. Intriguingly, it was found that the differential equation has some sort of connection with Horadam numbers.

Other papers dealing with problems involving Fibonacci numbers are what follows. In [7], J. S. Han, H. S. Kim and J. Neggers studied the Fibonacci norm of positive integers, and in [8], they studied Fibonacci sequences in groupoids. Han, Kim, and Neggers also introduced the concept of Fibonacci functions with Fibonacci numbers in [9] which were later on extended by B. Sroysang [19] to Fibonacci functions with period k. In [19], the following generalization of Fibonacci functions was presented.

1.1. Definition ([19])**.** Let $k \in \mathbb{N}$. A function $f\colon \mathbb{R} \to \mathbb{R}$ is called a Fibonacci function with period k if it satisfies the equation

$$(1.1) \qquad f(x+2k) = f(x+k) + f(x), \quad \forall x \in \mathbb{R}.$$

Properties of Fibonacci functions with period k were also presented in [19]. As further generalization of these functions, we define a second-order linear recurrent function with period k as follows:

1.2. Definition. Let k be a positive integer, and r and s be non-negative real numbers. A function $w\colon \mathbb{R} \to \mathbb{R}$ is said to be a second-order linear recurrent function with period k if it satisfies the equation

$$(1.2) \qquad w(x+2k) = rw(x+k) + sw(x), \quad \forall x \in \mathbb{R}.$$

We say that w is a complete second-order linear recurrent function if for any arbitrary r and s, w satisfies (1.2), otherwise it is called conditional.

Now, throughout the rest of this paper, we shall refer to a real-valued function w satisfying equation (1.2) as a recurrent function of order k instead of using the term second-order linear recurrent function with period k for convenience.

Definition (1.2) provides a further generalization of Fibonacci functions with Fibonacci numbers [9] and Fibonacci functions with period k [19]. Our main contribution includes a proof of the following open problems posed by Sroysang [19] (which is in fact a proof of a more general case of the problem):

1.3. Conjecture. *If f is a Fibonacci function with period k, then $f(x+k)/f(x) \to \phi$ as $x \to \infty$.*

1.4. Conjecture. *If f is a Fibonacci function with period k, then $f(x+k)/f(x) \to -\phi$ as $x \to \infty$.*

Here ϕ denotes the well-known golden ratio, i.e. $\phi = (1+\sqrt{5})/2 = 1.6180339\ldots$.

Now the rest of the paper is organized as follows: in Section 2 and Section 3, we give examples and basic properties of recurrent functions with period k and odd recurrent functions with period k, respectively. In Section 4, we develop the notion of these types of recurrent functions using the concept of f-even and f-odd functions discussed in [9]. In Section 5, we study the products of these functions and finally, in Section 6, we investigate the quotients of these functions. The proofs of conjectures (1.3) and (1.4) are also presented in the last section.

2. Recurrent functions with period k

In this section we present some properties of recurrent functions with period k. We begin by defining what we call Pell and Jacobsthal functions. If in equation (1.2), $r = 2$ and $s = 1$ (resp. $r = 1$ and $s = 2$), then we call such function a Pell (resp. Jacobsthal) function. That is, for a given natural number k, a Pell function p with period k satisfies

$$(2.1) \qquad p(x+2k) = 2p(x+k) + p(x), \quad \forall x \in \mathbb{R},$$

and a Jacobsthal function j with period k satisfies

$$(2.2) \qquad j(x+2k) = j(x+k) + 2j(x), \quad \forall x \in \mathbb{R}.$$

2.1. Example. Let $\alpha > 0, k \in \mathbb{N}$, and $w(x) = \alpha^{x/k}$ be a recurrent function with period k. Substituting w in (1.2), we have

$$\alpha^{(x/k)+2} = r\alpha^{(x/k)+1} + s\alpha^{x/k}, \quad \forall x \in \mathbb{R}.$$

So $\alpha^2 - r\alpha - s = 0$ whose roots are $\alpha_{1,2} = (r \pm \sqrt{r^2+4s})/2$. Thus, $w(x) = \alpha^{x/k}$.

The following are special cases of the previous example.

(1) If $(r,s)=(1,1)$, then the function $f(x):=\phi^{x/k}$ is an example of a Fibonacci function with period k.
(2) If $(r,s)=(2,1)$, then $p(x):=\sigma^{x/k}$, where $\sigma=1+\sqrt{2}$ is the well known *silver ratio*, is an example of a Pell function with period k.
(3) If $(r,s)=(1,2)$, then the function $j(x):=2^{x/k}$ is an example of a Jacobsthal function with period k.

2.2. Remark. Clearly, the functions f,p, and j are conditional. Also, any non-zero constant function is conditional but only for positive real numbers r and s such that $r+s=1$. On the other hand, the function $w(x)\equiv 0$ is an example of a complete type. It can be verified directly that any scalar multiple of a recurrent function with period k is again a recurrent function with period k. Furthermore, if a differentiable function w is a recurrent function with period k then so is its derivative w'.

2.3. Proposition. *Let $k\in\mathbb{N}$ and $w\colon\mathbb{R}\to\mathbb{R}$ be a recurrent function with period k. Define $g_t(x)=w(x+t)$ for all $x\in\mathbb{R}$, where $t\in\mathbb{R}$. Then g_t is also a recurrent function with period k.*

2.4. Corollary ([19])**.** *Let $k\in\mathbb{N}$ and $f\colon\mathbb{R}\to\mathbb{R}$ be a Fibonacci function with period k. Define $g_t(x)=f(x+t)$ for all $x\in\mathbb{R}$, where $t\in\mathbb{R}$. Then, g_t is also a Fibonacci function with period k.*

2.5. Example. Let $k\in\mathbb{N}$ and $t\in\mathbb{R}$. Define $g_t:\mathbb{R}\to\mathbb{R}$ by

$$g_t(x)=\alpha^{(x+t)/k},\quad \forall x\in\mathbb{R}, \tag{2.3}$$

then g_t is a recurrent function with period k.

As special cases of the previous example, we have the following:

(1) if $(r,s)=(1,1)$, then we have $g_t(x):=f(x+t)=\phi^{(x+t)/k}$, a Fibonacci function with period k,
(2) if $(r,s)=(2,1)$, then we have $g_t(x):=p(x+t)=\sigma^{(x+t)/k}$, a Pell function with period k,
(3) if $(r,s)=(1,2)$, then we have $g_t(x):=j(x+t)=2^{(x+t)/k}$, a Jacobsthal function with period k.

2.6. Theorem. *Let w be a recurrent function with period k and $\{W_n(0,1;r,s)\}$, or simply $\{W_n\}$, be a Horadam sequence with initial conditions $W_0=0$ and $W_1=1$. Then,*

$$w(x+nk)=W_n w(x+k)+sW_{n-1}w(x),\quad \forall x\in\mathbb{R}, n\in\mathbb{N}. \tag{2.4}$$

Proof. The proof is by induction on n. First, we note that $\alpha+\beta=r, \alpha-\beta=\sqrt{r^2+4s}$, and $\alpha\beta=-s$. We see that the formula obviously holds for $n=1,2$. So we assume (2.4) holds for n and $n+1$ for some natural number $n\geq 2$. Then,

$$\begin{aligned}
w(x+(n+2)k) &= rw(x+(n+1)k)+sw(x+nk)\\
&= r\left(W_{n+1}w(x+k)+sW_n w(x)\right)\\
&\qquad + s\left(W_n w(x+k)+sW_{n-1}w(x)\right)\\
&= \left(rW_{n+1}+sW_n\right)w(x+k)+s\left(rW_n+sW_{n-1}\right)w(x)\\
&= W_{n+2}w(x+k)+sW_{n+1}w(x),\quad \forall x\in\mathbb{R}, n\in\mathbb{N}.
\end{aligned}$$

By induction principle, conclusion follows. $\square$

2.7. Corollary. *Let w be a recurrent function with period k and let $\{W_n\}$ be the sequence of Horadam numbers. Then, $\alpha^n = \alpha W_n + sW_{n-1}$ for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$. In particular, for $r = s = 1$, we have $\phi^n = \phi F_n + F_{n-1}$, where F_n is the n^{th} Fibonacci number, for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$.*

Proof. From Example (2.1) we saw that $w(x) = \alpha^{x/k}$ is a recurrent function with period k, so it satisfies equation (2.4), i.e.

$$\begin{aligned}\alpha^{(x+nk)/k} &= w(x+nk) = W_n w(x+k) + sW_{n-1}w(x)\\ &= \alpha^{(x+k)/k}W_n + sW_{n-1}\alpha^{x/k}, \quad \forall x \in \mathbb{R}, n \in \mathbb{N}.\end{aligned}$$

Upon simplifying, we get

$$\alpha^n = \alpha W_n + sW_{n-1}, \quad \forall n \in \mathbb{N},$$

as desired. Letting $r = s = 1$ in $\alpha = (r + \sqrt{r^2+4s})/2$ we get $\phi^n = \phi F_n + F_{n-1}$. □

2.8. Corollary ([19])**.** *Let f be a Fibonacci function with period k and let $\{F_n\}$ be the sequence of Fibonacci numbers. Then, $f(x+nk) = F_n f(x+k) + F_{n-1}f(x)$ for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$.*

Similar results can be obtained easily for Pell and Jacobsthal functions.

3. Odd recurrent functions with period k

Here we discuss the notion of odd recurrent functions with period k formally defined as follows:

3.1. Definition. Let $k \in \mathbb{N}$. A function $w : \mathbb{R} \to \mathbb{R}$ is said to be an odd recurrent functions with period k if

$$w(x+2k) = -rw(x+k) + sw(x), \quad \forall x \in \mathbb{R}. \tag{3.1}$$

In particular, a function is an odd Fibonacci, odd Pell, and odd Jacobsthal function if it satisfies equation (3.1) with $(r, s) = (1, 1), (2, 1)$, and $(1, 2)$, respectively.

3.2. Example. Let $k \in \mathbb{N}, \tilde{\alpha} > 0$, and $w(x) = \tilde{\alpha}^{x/k}$ be an odd recurrent function with period k. Then, $\tilde{\alpha} = (-p + \sqrt{p^2+4q})/2 = q\alpha^{-1}$. So the function $w(x) = \left(q\alpha^{-1}\right)^{x/k}$ is an odd recurrent function.

Similar to what we remarked for recurrent functions with period k, if a differentiable function w is an odd recurrent function with period k then so is its derivative w'. Furthermore, any function defined by $g_t = w(x+t)$, where w satisfies (3.1) and $t \in \mathbb{R}$, is an odd recurrent function with period k. For example, the function defined by $g_t(x) = \tilde{\alpha}^{(x+t)/k}$ is an odd recurrent function with period k. Of course, the functions $g_t(x) := f(x+t) = \phi^{-(x+t)/k}$ and $g_t(x) := p(x+t) = \sigma^{-(x+t)/k}$ are also an odd Fibonacci and odd Pell function, respectively.

3.3. Theorem. *Let $w\colon \mathbb{R} \to \mathbb{R}$ be an odd recurrent function with period k, and let $\{W_{-n}\} = \{(-1)^{n+1}W_n(0, 1; r, s)\}$, i.e.*

$$W_{-(n+1)} = -rW_{-n} + sW_{-n+1}, \quad \forall n \in \mathbb{N}. \tag{3.2}$$

Then, for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$, we have

$$w(x+nk) = W_{-n}w(x+k) + sW_{-n+1}w(x). \tag{3.3}$$

The above theorem can be proven similarly as in Theorem 4.9 and we leave this to the reader.

3.4. Corollary. *Let w be an odd recurrent function with period k and let $\{W_{-n}\} = \{(-1)^{n+1}W_n\}$, where W_n is the n^{th} Horadam number. Then, $\tilde{\alpha}^n = \tilde{\alpha}W_{-n} + sW_{-n+1}$ for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$. In particular, for $r = 2$ and $s = 1$, we have $\sigma^{-n} = \sigma^{-1}P_n + P_{n-1}$, where P_n is the n^{th} Pell number, for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$.*

3.5. Corollary ([19])**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be an odd Fibonacci function with period k, and let $\{F_{-n}\} = \{(-1)^{n+1}F_n\}$ be a sequence of numbers where F_n is the n^{th} Fibonacci number, i.e.*

$$F_{-(n+1)} = -F_{-n} + F_{-n+1}, \quad \forall n \in \mathbb{N}. \tag{3.4}$$

Then, for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$, we have

$$f(x + nk) = F_{-n}f(x + k) + F_{-n+1}f(x). \tag{3.5}$$

Similar results can be formulated easily for Pell and Jacobsthal functions. Now we develop the concept of recurrent functions with period k using f-even and f-odd functions with period k.

4. f-even and f-odd functions with period k

We start-off this section with the following definition.

4.1. Definition (cf. [19])**.** Let $k \in \mathbb{N}$ and $\varphi \colon \mathbb{R} \to \mathbb{R}$ be such that if $\varphi h \equiv 0$, where $h \colon \mathbb{R} \to \mathbb{R}$ is continuous, then $h \equiv 0$. The map φ is said to be an f-even function with period k (resp. f-odd) function with period k if $\varphi(x+k) = \varphi(x)$ (resp. $\varphi(x+k) = -\varphi(x)$) for any $x \in \mathbb{R}$.

By the above definition, we can see immediately that there is no f-even and f-odd Fibonacci function except possibly when the function is the zero function. In fact, in general, a function $w \colon \mathbb{R} \to \mathbb{R}$ satisfying equation (1.2) is an f-even recurrent function with period k if and only if $r + s = 1$ or $w(x) \equiv 0$ for all $x \in \mathbb{R}$. Similarly, w is an f-odd recurrent function with period k if and only if $B - A = 1$ or $w \equiv 0$ for all $x \in \mathbb{R}$.

We first discuss f-even functions.

4.2. Example. Let $\varphi(x) = \cos(\pi x)$ for all $x \in \mathbb{R}$ and let $h \colon \mathbb{R} \to \mathbb{R}$ be a continuous function such that $(\varphi h)(x) = 0$. For any $x \notin \frac{\pi}{2}\mathbb{Z}$, we have $\varphi(x) \neq 0$, so $h(x) = 0$. Since $\mathbb{R} \backslash \frac{\pi}{2}\mathbb{Z}$ is dense in $\mathbb{R}$ and h is a continuous function, $h(x) = 0$. Now, let k be an even natural number and $x \in \mathbb{R}$. Then,

$$\varphi(x + k) = \cos(\pi(x + k)) = \cos(\pi x)\cos(k\pi) - \sin(\pi x)\sin(k\pi) = \cos(\pi x) = \varphi(x).$$

Hence, $\varphi(x) = \cos(\pi x)$ is an f-even function.

In [19], we have seen that $\varphi(x) = x - \lfloor x \rfloor$ is also an example of f-even functions.

4.3. Theorem. *Let $k \in \mathbb{N}$ and $\varphi \colon \mathbb{R} \to \mathbb{R}$ be an f-even function with period k and let $w \colon \mathbb{R} \to \mathbb{R}$ be a continuous function. Then, w is a recurrent function with period k if and only if φw is a recurrent function with period k.*

Proof. For the necessity part, we assume that w is a recurrent function with period k satisfying equation (1.2) with $r, s \in \mathbb{R}^+$. Then, for any $x \in \mathbb{R}$ we have

$$\begin{aligned}
\varphi(x + 2k)w(x + 2k) &= \varphi(x + k)[rw(x + k) + sw(x)] \\
&= r\varphi(x + k)w(x + k) + s\varphi(x + k)w(x) \\
&= r\varphi(x + k)w(x + k) + s(\varphi w)(x).
\end{aligned}$$

Hence, the product φw is a recurrent function with period k.

Now, for the sufficiency part, we assume that φw is a recurrent function with period k satisfying equation (1.2) with $r, s \in \mathbb{R}^+$. Let $x \in \mathbb{R}$. Then,

$$\begin{aligned}\varphi(x+k)w(x+2k) &= \varphi(x+2k)w(x+2k) = (\varphi w)(x+2k)\\ &= p(\varphi w)(x+k) + q(\varphi w)(x)\\ &= r\varphi(x+k)w(x+k) + s(\varphi w)(x)\\ &= r\varphi(x+k)w(x+k) + s\varphi(x+k)w(x)\\ &= \varphi(x+k)[rw(x+k) + sw(x)].\end{aligned}$$

Thus, $w(x+2k) = rw(x+k) + sw(x)$, and this shows that w is a recurrent function with period k. This completes the proof of the theorem. ☐

4.4. Corollary. *Let $k \in \mathbb{N}$ and $\varphi\colon \mathbb{R} \to \mathbb{R}$ be an f-even function with period k and let $g\colon \mathbb{R} \to \mathbb{R}$ be a continuous function. Then, g is a Fibonacci (resp. Pell and Jacobsthal) function with period k if and only if φg is a Fibonacci (resp. Pell and Jacobsthal) function with period k.*

4.5. Example. Let $k \in \mathbb{N}$ and define $\varphi(x) = \cos(\pi x)$ and $\gamma(x) = x - \lfloor x \rfloor$. Note that φ and γ are examples of f-even functions. Furthermore, recall that the function $w(x) = \alpha^{x/k}$ is a recurrent function with period k. Then, for all $x \in \mathbb{R}$, the products

$$(\varphi w)(x) = \cos(\pi x)\alpha^{x/k} \tag{4.1}$$

and

$$(\gamma w)(x) = (x - \lfloor x \rfloor)\alpha^{x/k} \tag{4.2}$$

are both examples of recurrent functions with period k. Specifically, if $(r, s) = (2, 1)$, then $(\varphi p)(x) = \cos(\pi x)\sigma^{x/k}$ and $(\gamma p)(x) = (x - \lfloor x \rfloor)\sigma^{x/k}$ are both Pell functions with period k.

4.6. Example. Let $k \in \mathbb{N}$ and define

$$\varphi(x) = \begin{cases} 1, & \text{for } x \in \mathbb{Q};\\ -1, & \text{otherwise.}\end{cases}$$

Hence, $\varphi(x+k) = \varphi(x)$ for any $x \in \mathbb{R}$. Also, if $\varphi h \equiv 0$, then $h \equiv 0$ whether or not h is continuous. Thus, φ is an f-even function with period k. We know that $w(x) = (x - \lfloor x \rfloor)\,\alpha^{x/k}$ is a recurrent function with period k. So, by Theorem 4.3, the mapping defined by

$$(\varphi w)(x) = \begin{cases} (x - \lfloor x \rfloor)\,\alpha^{x/k}, & \text{for } x \in \mathbb{Q};\\ (\lfloor x \rfloor - x)\,\alpha^{x/k}, & \text{otherwise,}\end{cases}$$

is also a recurrent function. Specifically, if $(r, s) = (1, 2)$, then $w(x) = 2^{x/k}\,(x - \lfloor x \rfloor)$. So we have,

$$(\varphi w)(x) = \begin{cases} 2^{x/k}\,(x - \lfloor x \rfloor), & \text{for } x \in \mathbb{Q};\\ 2^{x/k}\,(\lfloor x \rfloor - x), & \text{otherwise,}\end{cases}$$

a Jacobsthal function with period k.

4.7. Theorem. *Let $k \in \mathbb{N}, \varphi\colon \mathbb{R} \to \mathbb{R}$ be an f-even function with period k, and let $w\colon \mathbb{R} \to \mathbb{R}$ be a continuous function. Then, w is an odd recurrent function with period k if and only if φw is an odd recurrent function with period k.*

We omit the proof since it is similar on how we prove Theorem 4.3.

4.8. Example. Let k be an even natural number and consider the f-even functions φ and γ. In Example (3.2), we saw that $w(x) = \tilde{\alpha}^{x/k}$ is an odd recurrent function with period k. Then, for all $x \in \mathbb{R}$, the functions $(\varphi w)(x) = \cos(\pi x)\tilde{\alpha}^{x/k}$ and $(\gamma w)(x) = (x - \lfloor x \rfloor)\tilde{\alpha}^{x/k}$ are both odd recurrent functions with period k.

We now discuss f-odd functions. Recall that if $\varphi\colon \mathbb{R} \to \mathbb{R}$ such that for $\varphi h \equiv 0$ and h is continuous, we have $h \equiv 0$. The map φ is said to be an f-odd function with period k if $\varphi(x + k) = -\varphi(x)$ for all $x \in \mathbb{R}$. We have seen in [19] that $\varphi(x) = \sin(\pi x)$ is an example of f-odd function.

4.9. Theorem. *Let $k \in \mathbb{N}, \varphi\colon \mathbb{R} \to \mathbb{R}$ be an f-odd function with period k, and let $w\colon \mathbb{R} \to \mathbb{R}$ be a continuous function. Then, w is a recurrent function with period k if and only if φw is an odd recurrent function with period k.*

Again, we leave the proof to the reader.

4.10. Example. Let k be any odd natural number. Define $\varphi(x) = \sin(\pi x)$ and $w(x) = \tilde{\alpha}^{x/k}$ for all $x \in \mathbb{R}$. Hence, $(\varphi w)(x) = \tilde{\alpha}^{x/k}\sin(\pi x)$. We have seen in our discussion that φ is an f-odd function with period k and w is an odd recurrent function with period k. Hence, by Theorem 4.9, the product φw is a recurrent function with period k. We have the following examples for specific values of r and s.

(1) If $(r, s) = (1, 1)$, then $(\varphi f)(x) = \sin(\pi x)\,(\phi - 1)^{x/k}$ is a Fibonacci function with period k.
(2) If $(r, s) = (2, 1)$, then $(\varphi p)(x) = \sin(\pi x)\,(\sigma - 2)^{x/k}$ is a Pell function with period k.
(3) If $(r, s) = (1, 2)$, then $(\varphi j)(x) = \sin(\pi x)$ is a Jacobsthal function with period k.

5. Products of recurrent functions with period k

In this section, we give conditions so that whenever g and h are any two recurrent functions with period k in $\mathbb{R}$, their product forms another recurrent function with period k.

5.1. Theorem. *Let $k \in \mathbb{N}$ and g and h be two recurrent functions with period k satisfying*

$$g(x + 2k) = Ag(x + k) + Bg(x), \quad \forall x \in \mathbb{R}, \tag{5.1}$$

$$h(x + 2k) = Uh(x + k) + Vh(x), \quad \forall x \in \mathbb{R}, \tag{5.2}$$

respectively, where A, B, U, and V are non-negative real numbers. Suppose the following conditions are satisfied:

(C1) $\mathcal{A} = AU$, $\mathcal{B} = BV$, $AV = BU$,
(C2) *g is an f-even function and h is an f-odd function.*

Then, $w(x) := (gh)(x)$ forms another recurrent function with period k satisfying the equation

$$w(x + 2k) = \mathcal{A}w(x + k) + \mathcal{B}w(x), \quad \forall x \in \mathbb{R}. \tag{5.3}$$

Proof. The proof is straightforward. Suppose g and h are two recurrent functions with period k ($k \in \mathbb{N}$) satisfying equations (5.1) and (5.2), respectively. Furthermore, we suppose that conditions (C1) and (C2) are satisfied. Then,

$$\begin{aligned} w(x+2k) &= (gh)(x+2k) = [Ag(x+k)+Bg(x)][Uh(x+k)+Vh(x)] \\ &= \mathcal{A}w(x+k)+\mathcal{B}w(x)+[AVg(x+k)h(x)+BUg(x)h(x+k)] \\ &= \mathcal{A}w(x+k)+\mathcal{B}w(x)+AVg(x)[h(x)+h(x+k)] \\ &= \mathcal{A}w(x+k)+\mathcal{B}w(x), \quad \forall x \in \mathbb{R}, \end{aligned}$$

proving the theorem. □

5.2. Example. Let k be an odd natural number and $t \in \mathbb{R}^+$. Define $g(x) = x - \lfloor x \rfloor$ and let $A = t/(2t+1)$ and $B = (t+1)/(2t+1)$. Furthermore, define $h(x) = \sin(\pi x)$ and $U = V - 1 = t$. We claim that $w(x) := (gh)(x)$ is a recurrent function with period k satisfying the following equation:

$$w(x+2k) = \mathcal{A}w(x+1)+\mathcal{B}w(x), \quad \forall x \in \mathbb{R}, \tag{5.4}$$

where $\mathcal{A} = AU$ and $\mathcal{B} = BV$. We know that $g(x) = x - \lfloor x \rfloor$ and $h(x) = \sin(\pi x)$ are examples of f-even and f-odd functions with period k, respectively. We first show that g satisfies the equation

$$g(x+2k) = \frac{t}{2t+1}g(x+k)+\frac{t+1}{2t+1}g(x) \quad \forall x \in \mathbb{R}, \tag{5.5}$$

and then show that h satisfies

$$h(x+2k) = th(x+k)+(t+1)h(x), \quad \forall x \in \mathbb{R}. \tag{5.6}$$

We have,

$$\begin{aligned} g(x+2k) &= x+2k-\lfloor x+2k \rfloor = x - \lfloor x \rfloor = \left(\frac{t}{2t+1}+\frac{t+1}{2t+1}\right)(x-\lfloor x \rfloor) \\ &= \frac{t}{2t+1}\left(x+1-\lfloor x+1 \rfloor\right)+\frac{t+1}{2t+1}\left(x+1-\lfloor x+1 \rfloor\right) \\ &= \frac{t}{2t+1}g(x+1)+\frac{t+1}{2t+1}g(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

and

$$\begin{aligned} h(x+2k) &= \sin(\pi(x+2k)) = \sin(\pi x)\cos(2k\pi) \\ &= -t\sin(\pi x)+(t+1)\sin(\pi x) \\ &= t\sin(\pi x)\cos(k\pi)+(t+1)\sin(\pi x) \\ &= t\sin(\pi(x+k))+(t+1)\sin(\pi x) \\ &= th(x+k)+(t+1)h(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

Obviously, $AV = BU$. By Theorem 5.1, it follows that $w(x) = (gh)(x)$ is a recurrent function with period k satisfying equation (5.4).

5.3. Corollary. *Let g and h be two recurrent functions with period k satisfying equations* (5.1) *and* (5.2)*, respectively. Suppose conditions* (C1) *and* (C2) *are satisfied, then $w(x) := (gh)(x)$ is never a Fibonacci function with period k except possibly when $g \equiv 0$ or $h \equiv 0$ for all $x \in \mathbb{R}$.*

Proof. Let g and h be two functions satisfying equations (5.1) and (5.2), respectively and suppose that conditions (C1) and (C2) are satisfied. Hence,

$$\begin{aligned} g(x+k) = g(x+2k) &= Ag(x+k)+Bg(x) \\ &= (A+B)g(x+k), \quad \forall x \in \mathbb{R}, \end{aligned}$$

and

$$\begin{aligned} h(x+k) &= h(x+2k) = Uh(x+k) + Vh(x) \\ &= (U-V)h(x+k), \quad \forall x \in \mathbb{R}. \end{aligned}$$

It follows that $A+B=1$ and $U-V=1$. Since $AV=BU$, we have $AV=(1-B)V=(1+V)B=BU$ which implies that $B=V/(2V+1)$. Letting $V=t\in\mathbb{R}^+$ we get the following equations:

$$g(x+2k) = (t+1)(2t+1)^{-1}g(x+k) + t(2t+1)^{-1}g(x), \tag{5.7}$$

$$h(x+2k) = (t+1)h(x+k) + th(x). \tag{5.8}$$

Hence,

$$\begin{aligned} w(x+2k) &= g(x+2k)h(x+2k) \\ &= \left(\frac{t+1}{2t+1}g(x+k) + \frac{t}{2t+1}g(x)\right)((t+1)h(x+k)+th(x)) \\ &= \frac{(t+1)^2}{2t+1}w(x+k) + \frac{t^2}{2t+1}w(x) + \frac{t(t+1)}{2t+1}g(x)\left(h(x+k)+h(x)\right) \\ &= \frac{(t+1)^2}{2t+1}w(x+k) + \frac{t^2}{2t+1}w(x) \\ &= \mathcal{A}w(x+k) + \mathcal{B}w(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

Suppose w is a Fibonacci function. Then, $\mathcal{A}=\mathcal{B}=1$. This is impossible since

$$\frac{(t+1)^2}{2t+1} = \frac{t^2}{2t+1} + 1 = 2 > 1 = \frac{t^2}{2t+1}.$$

This proves the theorem. □

5.4. Corollary. *Let g and h be two recurrent functions satisfying equation* (5.1) *and* (5.2)*, respectively. Suppose that $AU=2, BV=1, AV=BU$ and condition* (C2) *is satisfied. Then, $w(x):=(gh)(x)$ is a Pell function.*

Proof. The proof follows a similar argument used in the proof of Corollary 5.3 so we omit it. □

5.5. Example. In the proof of Corollary 5.3 we have seen that the product of equations (5.7) and (5.8) forms a recurrent function provided they satisfy conditions (C1) and (C2). If we set $\mathcal{A}=2$ and $\mathcal{B}=1$, then we obtain a Pell function provided we could find a positive real number t such that $t^2-2t-1=0$. The solution to this equation is given by $t=1\pm\sqrt{2}$, so we choose $t=1+\sqrt{2}\in\mathbb{R}^+$. Hence, equations (5.7) and (5.8) become

$$g(x+2k) = \left(\frac{\sigma+1}{2\sigma+1}\right)g(x+k) + \left(\frac{\sigma}{2\sigma+1}\right)g(x) \tag{5.9}$$

and

$$h(x+2k) = (\sigma+1)h(x+k) + \sigma h(x), \tag{5.10}$$

respectively. Now, our goal is to find functions g and h satisfying condition (C2). We can choose $g(x)=x-\lfloor x\rfloor$ and $h(x)=\sin(\pi x)$, and one can check that these equations satisfy equations (5.9) and (5.10). Thus, by Corollary 5.8, $w(x)=(x-\lfloor x\rfloor)\sin(\pi x)$ is a Pell function.

5.6. Corollary. *Let g and h be any two recurrent functions with period k satisfying equation* (5.1) *and* (5.2)*, respectively, such that $AU=1, BV=2, AV=BU$ and condition* (C2) *is satisfied. Then, $w(x):=(gh)(x)$ is a Jacobsthal function.*

5.7. Example. In Example (5.2), we have seen that the function defined by $w(x) := (gh)(x) = (x - \lfloor x \rfloor)\sin(\pi x)$ is a recurrent function with period k satisfying equation (5.4). If we set $\mathcal{A}$ and $\mathcal{B}$ to be in the set of natural numbers such that $\mathcal{B} = \mathcal{A} + 1$, then we have the following

$$\begin{aligned} g(x+2k) &= \theta(2\theta+1)^{-1}g(x+k) + (\theta+1)(2\theta+1)^{-1}g(x), \quad \forall x \in \mathbb{R}, \\ h(x+2k) &= \theta h(x+k) + (\theta+1)\, h(x), \quad \forall x \in \mathbb{R}, \\ w(x+2k) &= \mathcal{A}w(x+k) + (\mathcal{A}+1)w(x), \quad \forall x \in \mathbb{R}, \end{aligned}$$

where $\theta = \mathcal{A} + \sqrt{\mathcal{A}(\mathcal{A}+1)}$. If we let $\mathcal{A} = 1$, then we obtain the following equations,

$$g(x+2k) = \sigma(2\sigma+1)^{-1}g(x+k) + (\sigma+1)(2\sigma+1)^{-1}g(x), \quad \forall x \in \mathbb{R}, \tag{5.11}$$

$$h(x+2k) = \sigma h(x+k) + (\sigma+1)\, h(x), \quad \forall x \in \mathbb{R}, \tag{5.12}$$

$$w(x+2k) = w(x+k) + 2w(x), \quad \forall x \in \mathbb{R}, \tag{5.13}$$

where $\sigma = 1 + \sqrt{2}$ is the well known silver ratio. Suprisingly, equation (5.13)appears to be a Jacobsthal function. Since g and h are f-even and f-odd functions, respectively, we see that the function defined by $w(x) := (gh)(x) = (x - \lfloor x \rfloor)\sin(\pi x)$ with g and h satisfying equation (5.11)and (5.12) is indeed a Jacobsthal function by Corollary 5.6.

5.8. Theorem. *Let g be a recurrent function with period k ($k \in \mathbb{N}$) and h be an odd recurrent function, also, with period k satisfying*

$$g(x+2k) = Ag(x+k) + Bg(x), \quad \forall x \in \mathbb{R}, \tag{5.14}$$

$$h(x+2k) = -Uh(x+k) + Vh(x), \quad \forall x \in \mathbb{R}, \tag{5.15}$$

respectively, where A, B, U, and V are non-negative real numbers. Suppose condition (C1) *is satisfied and*

- (C3) *g and h are both f-even functions, or*
- (C4) *g and h are both f-odd functions.*

Then, $w(x) := (gh)(x)$ is an odd recurrent function with period k satisfying the equation

$$w(x+2k) = -\mathcal{A}w(x+k) + \mathcal{B}w(x). \tag{5.16}$$

Proof. The proof is similar to the proof of Theorem 5.1 so we omit it. □

5.9. Example. Let k be an even natural number and $t \in \mathbb{R}^+$. Consider the functions $g(x) = x - \lfloor x \rfloor$ satisfying equation (5.7) and $h(x) = \cos(\pi x)$. We show that h is an odd recurrent function with period k satisfying the equation

$$h(x+2k) = -th(x+k) + (t+1)h(x). \tag{5.17}$$

We have,

$$\begin{aligned} h(x+2k) &= \cos(\pi(x+2k)) = \cos(\pi x)\cos(2k\pi) \\ &= -t\cos(\pi x) + (t+1)\cos(\pi x) \\ &= -t\cos(\pi x)\cos(k\pi) + (t+1)\cos(\pi x) \\ &= -t\cos(\pi(x+k)) + (t+1)\cos(\pi x) \\ &= -th(x+k) + (t+1)h(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

Clearly, $AV = BU$. Since g and h are both f-even functions, then by Theorem 5.8, $w(x) := (x - \lfloor x \rfloor)\cos(\pi x)$ is an odd recurrent function with period k satisfying the equation given by

$$w(x+2k) = -\left(\frac{t^2}{2t+1}\right)w(x+k) + \left(\frac{(t+1)^2}{2t+1}\right)w(x), \quad \forall x \in \mathbb{R}. \tag{5.18}$$

5.10. Example. Let $k \in \mathbb{N}$ and $A, B, U, V \in \mathbb{R}^+$. Consider f-odd functions g and h satisfying equations (5.14) and (5.15), respectively. Then we have

$$\begin{aligned} g(x+2k) &= Ag(x+k) + Bg(x) = (A-B)g(x+k) \\ &= (B-A)g(x+2k) \end{aligned}$$

and

$$\begin{aligned} h(x+2k) &= -Uh(x+k) + Vh(x) = -(U+V)h(x+k) \\ &= (U+V)h(x+2k). \end{aligned}$$

These imply that $B - A = 1$ and $U + V = 1$. If $AV = BU$, then $AV = (B-1)V = (1-V)B = BU$, which implies that $B = V/(2V-1)$. Hence, we have the following:

$$(5.19) \quad g(x+2k) = (1-V)(2V-1)^{-1} g(x+k) + V(2V-1)^{-1} g(x),$$

$$(5.20) \quad h(x+2k) = -(1-V)h(x+k) + Vh(x).$$

Since $A \in \mathbb{R}^+$, we get $w(x) := (gh)(x)$ satisfying the equation

$$(5.21) \quad w(x+2k) = -\left(\frac{(1-V)^2}{2V-1}\right) w(x+k) + \left(\frac{V^2}{2V-1}\right) w(x), \quad \forall x \in \mathbb{R},$$

is an odd recurrent function with period k if and only if $V \in \left(\frac{1}{2}, 1\right]$.

5.11. Corollary. *Let g be a recurrent function with period k and h be an odd recurrent function, also, with period k satisfying equations* (5.14) *and* (5.15)*. Suppose conditions* (C1)*, and* (C3) *or* (C4) *are satisfied. Then, $w(x) := (gh)(x)$ is never an odd Fibonacci nor an odd Pell function with period k except possibly when $g \equiv 0$ or $h \equiv 0$ for all $x \in \mathbb{R}$.*

F

5.12. Corollary. *Let $g(x)$ be a recurrent function with period k and h be an odd recurrent function, also, with period k satisfying equations* (5.14) *and* (5.15)*. Suppose $AU = 1, BV = 2, AV = BU$, and the functions g and h satisfies condition* (C3) *or* (C4)*, then $w(x) := (gh)(x)$ is an odd Jacobsthal function with period k.*

We leave the verification of Corollary 5.11 and Corollary 5.12 to the reader.

5.13. Example. Let k be an even natural number, $t \in \mathbb{R}^+$. Consider the functions $g(x) = x - \lfloor x \rfloor$ and $h(x) = \cos(\pi x)$. We note that g and h are both f-even functions. We claim that if these two functions satisfy the following equations

$$(5.22) \quad g(x+2k) = \left(\frac{\sigma}{2\sigma+1}\right) g(x+k) + \left(\frac{\sigma+1}{2\sigma+1}\right) g(x), \quad \forall x \in \mathbb{R},$$

and

$$(5.23) \quad h(x+2k) = -\sigma h(x+k) + (\sigma+1)h(x), \quad \forall x \in \mathbb{R},$$

respectively, where σ is the silver ratio, then the product $w(x) := (gh)(x) = (x - \lfloor x \rfloor)\cos(\pi x)$ forms an odd Jacobsthal function with period k which satisfies the equation

$$\begin{aligned} w(x+2k) &= -\left(\frac{\sigma^2}{2\sigma+1}\right) w(x+k) + \left(\frac{(\sigma+1)^2}{2\sigma+1}\right) w(x) \\ &= -w(x+k) + 2w(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

We know that g satisfies equation (5.22) from Example (5.2). Hence, we only need to show that $h(x) = \cos(\pi x)$ satisfies equation (5.23). In fact, we have shown this already in Example (5.9). Thus, by Corollary 5.12, $w(x) = (x - \lfloor x \rfloor)\cos(\pi x)$ is indeed an odd Jacobsthal function with period k.

5.14. Example. In Example (5.10), we have seen that for any arbitrary f-odd functions g and h satisfying equations (5.19) and (5.20) with $U \in \left[0, \frac{1}{2}, 0\right)$, the product $w(x) := (gh)(x)$ is an odd recurrent function with period k. If we let $V = 3 - \sigma \in \left(\frac{1}{2}, 1\right]$, where σ is the silver ratio, then the functions g and h satisfying the following equations

$$\begin{aligned} g(x+2k) &= \left(\frac{\sigma-2}{5-2\sigma}\right) g(x+k) + \left(\frac{3-\sigma}{5-2\sigma}\right) g(x), \quad \forall x \in \mathbb{R}, \\ h(x+2k) &= -(\sigma-2)h(x+k) + (3-\sigma)h(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

whose product is given by $w(x) := (gh)(x)$ satisfying the equation

$$w(x+2k) = -\left(\frac{(\sigma-2)^2}{5-2\sigma}\right) w(x+k) + \left(\frac{(3-\sigma)^2}{5-2\sigma}\right) w(x), \quad \forall x \in \mathbb{R}, \tag{5.24}$$

is an odd Jacobsthal function with period k by Corollary 5.12.

5.15. Theorem. *Let g and h be two functions satisfying equation* (5.1) *and* (5.2)*, respectively. Suppose $\mathcal{A} = AU, \mathcal{B} = BV, AV = -BU$, and condition* (C3) *or* (C4) *is satisfied, then $w(x) := (gh)(x)$ is a recurrent function with period k satisfying equation* (5.16)*.*

Proof. The proof is similar to the proof of Theorem 5.1 so we omit it. □

5.16. Corollary. *Let g and h be two functions satisfying equation* (5.1) *and* (5.2)*, respectively. Suppose $\mathcal{A} = AU, \mathcal{B} = BV, AV = -BU$, and condition* (C3) *or* (C4) *is satisfied, then, $w(x) := (gh)(x)$ si never a (or an odd) Fibonacci nor a (or an odd) Pell function with period k except possibly when $g \equiv 0$ or $h \equiv 0$ for all $x \in \mathbb{R}$. Moreover, if $\mathcal{A} = 1, \mathcal{B} = 2$, and g and h are both f-odd functions, then w is a Jacobsthal function with period k.*

Proof. Let g and h be two functions satisfying equation (5.1) and (5.2), respectively. Suppose $\mathcal{A} = AU, \mathcal{B} = BV, AV = -BU$, and g and h are both f-even functions, then we have the following:

$$\begin{aligned} g(x+2k) &= Ag(x+k) + Bg(x) = (A+B)g(x+k) \\ &= (A+B)g(x+2k), \quad \forall x \in \mathbb{R}, \end{aligned}$$

and

$$\begin{aligned} h(x+2k) &= Uh(x+k) + Vh(x) = (U+V)h(x+k) \\ &= (U+V)h(x+2k), \quad \forall x \in \mathbb{R}. \end{aligned}$$

These imply that $A + B = U + V = 1$. Since $AV = -BU$, we get $B = V/(2V-1)$. Hence,

$$\begin{aligned} g(x+2k) &= (V-1)(2V-1)^{-1} g(x+k) + V(2V-1)^{-1} g(x), \quad \forall x \in \mathbb{R}, \\ h(x+2k) &= (1-V)h(x+k) + Vh(x), \quad \forall x \in \mathbb{R}. \end{aligned}$$

By assumption, $A \in \mathbb{R}^+$, then $w(x) := (gh)(x)$ satisfying the equation

$$w(x+2k) = -\left(\frac{(1-V)^2}{2V-1}\right) w(x+k) + \left(\frac{V^2}{2V-1}\right) w(x), \quad \forall x \in \mathbb{R},$$

is a recurrent function with period k if and only if $V \in \left[0, \frac{1}{2}\right)$. Now, suppose that w is a Fibonacci function with period k, then

$$\frac{V^2}{2V-1} = 1 > 0 > \frac{-V^2 + 2V - 1}{2V-1} = -\frac{(1-V)^2}{2V-1}. \tag{5.25}$$

So, it is impossible that w is a (or an odd) Fibonacci function with period k. Furthermore, it can also be seen in (5.25) that w cannot be a (nor an odd) Pell function with period

k. Similarly, suppose w is a Jacobsthal function with period k, then $V^2/(2V-1) = 2$, which implies that

$$-\frac{(1-V)^2}{2V-1} = \frac{-V^2+2V-1}{2V-1} = -1, \quad \text{or equivalently} \quad V = 2 \pm \sqrt{2}.$$

But $V \in \left[0, \frac{1}{2}\right)$, thus w cannot be a (nor an odd) Jacobsthal function with period k. On the other hand, if g and h are both f-odd functions, then we have

$$\begin{aligned} g(x+2k) &= Ag(x+k) + Bg(x) = (A-B)g(x+k) \\ &= (B-A)g(x+2k), \quad \forall x \in \mathbb{R}, \end{aligned}$$

and

$$\begin{aligned} h(x+2k) &= Uh(x+k) + Vh(x) = (U-V)h(x+k) \\ &= (V-U)h(x+2k), \quad \forall x \in \mathbb{R}. \end{aligned}$$

These imply that $B - A = V - U = 1$. Since $AV = -BU$, we obtain $B = V/(2V-1)$. Hence,

$$\begin{aligned} g(x+2k) &= \left(\frac{1-V}{2V-1}\right) g(x+k) + \left(\frac{V}{2V-1}\right) g(x), \\ h(x+2k) &= (V-1)h(x+k) + Vh(x). \end{aligned}$$

Because $A \in \mathbb{R}^+$, then $w(x) := (gh)(x)$ satisfying the equation

$$w(x+2k) = -\left(\frac{(1-V)^2}{2V-1}\right) w(x+k) + \left(\frac{V^2}{2V-1}\right) w(x), \quad \forall x \in \mathbb{R}, \tag{5.26}$$

is a recurrent function with period k if and only if $V \in \left(\frac{1}{2}, 1\right]$. It can easily be verified that there is no value for $V \in \left(\frac{1}{2}, 1\right]$ such that equation (5.26) is a Fibonacci or a Pell function with period k. However, we can find a value for $V \in \left(\frac{1}{2}, 1\right]$ such that (5.26) is a Jacobsthal function. In particular, we can choose $V = 3 - \sigma \in \left(\frac{1}{2}, 1\right]$ so that we have

$$g(x+2k) = \left(\frac{\sigma-2}{5-2\sigma}\right) g(x+k) + \left(\frac{3-\sigma}{5-2\sigma}\right) g(x), \tag{5.27}$$

$$h(x+2k) = (2-\sigma)h(x+k) + (3-\sigma)h(x). \tag{5.28}$$

Equations (5.27) and (5.28) imply that

$$\begin{aligned} w(x+2k) &= -\left(\frac{(\sigma-2)^2}{5-2\sigma}\right) w(x+k) + \left(\frac{(3-\sigma)^2}{5-2\sigma}\right) w(x) \\ &= -w(x+k) + 2w(x), \quad \forall x \in \mathbb{R}, \end{aligned}$$

is a Jacobsthal function with period k. This verifies the corollary. □

5.17. Theorem. *Let g and h be any two functions satisfying*

$$\begin{aligned} g(x+2k) &= -Ag(x+k) + Bg(x), \\ h(x+2k) &= -Uh(x+k) + Vh(x), \end{aligned}$$

respectively, where $A, B, U,$ and V are non-negative real numbers. Suppose that condition (C1) *is satisfied, and g is an f-even function whereas h is an f-odd function, then $w(x) := (gh)(x)$ forms another recurrent function with period k satisfying equation* (5.3). *Furthermore, if $\mathcal{A} = AU, \mathcal{B} = BV, AV = -BU,$ and condition* (C3) *or* (C4) *is satisfied then, $w(x) := (gh)(x)$ is also a recurrent function with period k satisfying equation* (5.3).

6. Quotients of recurrent functions

Here we discuss the limit of the quotients of recurrent functions with period k and provide proofs to two conjecture of Sroysang [19].

6.1. Theorem. *Let $k \in \mathbb{N}$ and $w\colon \mathbb{R} \to \mathbb{R}$ be a recurrent function with period k. Then, the limit of the quotient $w(x+k)/w(x)$ exists.*

Proof. Let $k, n \in \mathbb{N},\ r, s \in \mathbb{R}^+$, and consider the quotient $Q(x) := w(x+k)/w(x)$, where w is a recurrent function with period k. Then, we have two possibilities: (i) $Q(x) < 0$, and (ii) $Q(x) > 0$. First, suppose that $Q(x) < 0$ then (WLOG), $u := w(x) > 0$ and $v := w(x+k) < 0$. Hence,

$$\begin{aligned}
w(x+2k) &= -rw(x+k) + sw(x)\\
&= -rv + su,\\
w(x+3k) &= rw(x+2k) - sw(x+k) = r(-rv+su) - sv\\
&= -(r^2+s)v + rsu,\\
w(x+4k) &= rw(x+3k) + sw(x+2k) = r(-(r^2+s)v + rsu) + s(-rv+su)\\
&= -(r^3+2rs)v + s(r^2+s)u\\
&\ \ \vdots\\
w(x+nk) &= -W_n v + sW_{n-1}u, \quad \forall n \in \mathbb{N},
\end{aligned}$$

where W_n is the n^{th} Horadam number with initial conditions $W_0 = 0$ and $W_1 = 1$. Let $x' \in \mathbb{R}$. Then, we could find $x \in \mathbb{R}$ and $n \in \mathbb{Z}$ such that $x' = x + nk$. So we have

$$\frac{w(x'+k)}{w(x')} = \frac{w(x+(n+1)k)}{w(x+nk)} = \frac{-W_{n+1}v + sW_n u}{-W_n v + sW_{n-1}u} = \frac{-v\frac{W_{n+1}}{W_n} + su}{-v + su\frac{W_{n-1}}{W_n}}. \tag{6.1}$$

Since $x \to \infty$ as $n \to \infty$, then letting $n \to \infty$ equation (6.1) we get

$$\lim_{x\to\infty} \frac{w(x'+k)}{w(x')} = \lim_{n\to\infty} \frac{-v\frac{W_{n+1}}{W_n} + qu}{-v + qu\frac{W_{n-1}}{W_n}} = \frac{-v\left(\lim_{n\to\infty} \frac{W_{n+1}}{W_n}\right) + qu}{-v + qu\left(\lim_{n\to\infty} \frac{W_{n-1}}{W_n}\right)}.$$

Note that $\beta = \frac{r-\sqrt{r^2+4s}}{2} \in (-1, 0)$. So $\lim_{n\to\infty} \beta^n = 0$. Thus,

$$\lim_{x\to\infty} \frac{w(x'+k)}{w(x')} = \frac{-\alpha v + su}{-v + \alpha^{-1}su} = \alpha,$$

since $\lim_{n\to\infty} \frac{W_{n+1}}{W_n} = \lim_{n\to\infty} \frac{\alpha^{n+1}-\beta^{n+1}}{\alpha^n-\beta^n} = \alpha$. On the other hand, suppose (WLOG), $w(x)$ and $w(x+k)$ are both positive and $x \gg k$. We can express $w(x+k)/w(x)$ as $w(2n+\delta+k)/w(2n+\delta)$ since any non-negative real number x can be written as $x = \delta + 2n$ for some $\delta \in \mathbb{R}$ and $n \in \mathbb{N}$. Now, we claim that

$$\lim_{n\to\infty} \frac{w(2n+\delta+k)}{w(2n+\delta)} = \alpha.$$

We show this by expressing both sides in terms of continued fractions. For the LHS, we have

$$\begin{aligned}
\frac{w(2n+k+\delta)}{w(2n+\delta)} &= \frac{pw(2n+\delta)+qw(2n+\delta-k)}{w(2n+\delta)} = r+s\frac{w(2n+\delta-k)}{w(2n+\delta)} \\
&= r+s\cfrac{1}{r+s\,\cfrac{w(2n+\delta-2k)}{w(2n+\delta-k)}} \\
&= r+s\cfrac{1}{r+s\,\cfrac{1}{r+s\,\cfrac{w(2n+\delta-3k)}{w(2n+\delta-2k)}}} \\
&\;\vdots \\
&= r+s\cfrac{1}{r+s\,\cfrac{1}{r+s\,\cfrac{1}{r+s\,\cfrac{1}{\ddots+s\frac{w(2n+\delta-(2n-1)k)}{w(2n+\delta-2(n-1)k)}}}}}.
\end{aligned}$$

Hence,

$$\frac{w(2n+\delta+k)}{w(2n+\delta)} = r+s\cfrac{1}{r+s\,\cfrac{1}{r+s\,\cfrac{1}{\ddots+s\frac{w(2n+\delta-(2n-1)k)}{w(2n+\delta-2(n-1)k)}}}}. \tag{6.2}$$

For the RHS, we have $\alpha = r + (-r+\sqrt{r^2+4s})/2 = r + s/\alpha$. Thus, we have

$$\alpha = r+s\cfrac{1}{r+s\,\cfrac{1}{r+s\,\cfrac{1}{r+s\,\cfrac{1}{\ddots}}}}.$$

Now, taking the limit of equation (6.2) as $n\to\infty$, we get

$$\lim_{x\to\infty} Q(x) = \lim_{n\to\infty}\frac{w(2n+\delta+k)}{w(2n+\delta)} = r+s\cfrac{1}{r+s\,\cfrac{1}{r+s\,\cfrac{1}{\ddots}}} = \alpha.$$

This proves the theorem. $\square$

6.2. Corollary. *Let $k\in\mathbb{N}$. If f, (resp. p and j) is a Fibonacci (resp. Pell function and Jacobsthal) function with period k, then the limit of the quotient $f(x+1)/f(x)$, (resp. $p(x+1)/p(x)$ and $j(x+1)/j(x)$) exists.*

6.3. Corollary. *Let $k\in\mathbb{N}$ and let w be a recurrent function with period k, then $\lim_{x\to\infty} w(x+k)/w(x)=\alpha$. In particular, if f (resp. p and j) is a Fibonacci (resp. Pell, and Jacobsthal) function with period k, then $\lim_{x\to\infty} f(x+k)/f(x)=\phi$ (resp. $\lim_{x\to\infty} p(x+k)/p(x)=\sigma$ and $\lim_{x\to\infty} j(x+k)/j(x)=2$).*

Proof. Let $n, k \in \mathbb{N}$ and suppose that the quotient $w(x+k)/w(x)$ is positive. Furthermore, assume (WLOG) that $w(x)$ and $w(x+k)$ are both positive, then

$$\frac{w(x+(n+1)k)}{w(x+nk)} = \frac{W_{n+1}w(x+k) + sW_n w(x)}{W_n w(x+k) + sW_{n-1}w(x)} = \frac{\frac{W_{n+1}}{W_n}w(x+k) + sw(x)}{w(x+k) + \frac{W_{n-1}}{W_n}sw(x)}. \tag{6.3}$$

Letting $n \to \infty$ in equation (6.3), we get

$$\lim_{n\to\infty} \frac{w(x+(n+1)k)}{w(x+nk)} = \alpha.$$

If, on the other hand, $w(x+k)/w(x)$ is negative and suppose that (WLOG) $w(x) > 0$ and $w(x+k) < 0$ then, by the proof of Theorem 6.1, we see that $\lim_{x\to\infty} w(x+(n+1)k)/w(x+nk) = \alpha$. This proves the corollary. $\square$

6.4. Remark. In [19, Conjecture 25], Sroysang conjectured that if f is a fibonacci function with period k, then $\lim_{x\to\infty} f(x+k)/f(x) = \phi$. Indeed, this is true by Corollary 6.3.

Sroysang's second conjecture found in [19, Conjecture 26] is also true as stated in the following results.

6.5. Theorem. *Let $k \in \mathbb{N}$ and $w := \mathbb{R} \to \mathbb{R}$ be an odd recurrent function with period k. Then, the limit of the quotient $w(x+k)/w(x)$ exists.*

6.6. Corollary. *Let $k \in \mathbb{N}$. If f, (resp. p and j) is an odd Fibonacci (resp. odd Pell function and odd Jacobsthal) function with period k, then the limit of the quotient $f(x+1)/f(x)$, (resp. $p(x+1)/p(x)$ and $j(x+1)/j(x)$) exists.*

6.7. Corollary. *Let $k \in \mathbb{N}$ and let w be an odd recurrent function with period k, then $\lim_{x\to\infty} w(x+k)/w(x) = -\alpha$. In particular, if f (resp. p and j) is an odd Fibonacci (resp. odd Pell, and odd Jacobsthal) function with period k, then $\lim_{x\to\infty} f(x+k)/f(x) = -\phi$ (resp. $\lim_{x\to\infty} p(x+k)/p(x) = -\sigma$ and $\lim_{x\to\infty} j(x+k)/j(x) = -2$).*

We omit the proof of these results since they can be proven in a similar fashion as in Theorem 6.1 and Corollary 6.3.

7. Summary

We were able to extend successfully the notion of Fibonacci functions [9] and Fibonacci functions with period k [19] by characterizing the concept of second-order linear recurrent functions with period k. We were also able to confirm the conjectures of Sroysang in [19] by proving a more general result about the asymptotic growth rate of Fibonacci functions with period k. In our next investigation [18], we will revisit Sroysang's conjecture and provide another proof using the results presented in [20] by Tollu et al. together with the concept of Cauchy sequences.

Acknowledgement

The author wishes to thank the referee for carefully handling and examining his manuscript.

References

[1] Amdeberhan, T. Chen, X. Moll, V. C. and Sagan, B. E. *Generalized Fibonacci polynomials and Fibonomial coefficients.* Retrieved from `arXiv:1306.6511v2 [math.CO] 26 Jul 2013`.

[2] Bacani, J. B., Rabago, J. F. T. *On two nonlinear difference equations*, submitted.

[3] Chandra, P. and Weisstein, E. W. *Fibonacci Number.* From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/FibonacciNumber.html`

[4] Cook, C. K. Hillman, R. A. and Shannon, A. G. *Some aspects of Fibonacci polynomial congruences*, Ann. Math. Inform. **41**, 211–217, 2013.

[5] Dunlap, R. A. *The golden ratio and Fibonacci numbers* (World Scientific, 1998).

[6] Hakami, A. *An application of Fibonacci sequence on continued fractions*, Int. Math. Forum **10** (2), 69–74, 2015.

[7] Han, J.S. Kim, H. S. and Neggers, J. *The Fibonacci norm of a positive integer n-observations and conjectures*, Int. J. Number Theory **6** (2), 371–385, 2010.

[8] Han, J. S. Kim, H. S. and Neggers, J. *Fibonacci sequences in groupoids*, Adv. Differ. Equ. **2012**, Article 19, 7 pages, 2012.

[9] Han, J. S. Kim, H. S. and Neggers, J. *On Fibonacci functions with Fibonacci numbers*, Adv. Differ. Equ. **2012**, Article 126, 7 pages, 2012.

[10] Horadam, A. F. *Basic properties of certain generalized sequence of numbers*, Fib. Quart. **3**, 161–176, 1965.

[11] Koshy, T. *Fibonacci and Lucas numbers with applications* (Pure and Applied Mathematics, Wiley-Interscience, New York, 2001).

[12] Larcombe, P. J. Bagdasar, O. D. and Fennessey, E. J. *Horadam sequences: a survey* Bulletin of the I.C.A. **67**, 49–72, 2013.

[13] Larcombe, P. J. *Horadam Sequences: a survey update and extension*, submitted.

[14] Ma, Y. and Zhang, T. *On Generalized Fibonacci Polynomials and Bernoulli Numbers*, J. Integer Seq. **8**, Article 05.5.3, 2005.

[15] Rabago, J. F. T. *On solving the second-order linear recurrence sequence*, Int. J. Math. Sci. Comp. **2** (1), 1–2, 2012.

[16] Rabago, J. F. T. *On second-order linear recurrent homogeneous differential equations with period k*, Hacet. J. Math. Stat. **43** (6), 923–933, 2014.

[17] Rabago, J. F. T. *On k-Fibonacci numbers with applications to continued fractions*, Journal of Physics: Conference Series **693** 012005, 2016.

[18] Rabago, J. F. T. *On the closed-form solution of a nonlinear difference equation and another proof to Sroysang's conjecture*, in preparation.

[19] Sroysang, B. *On Fibonacci functions with period k*, Discrete Dyn. Nat. Soc. **2013**, Article ID 418123, 4 pages, 2013.

[20] Tollu, D.T. Yazlik, Y. and Taskara, N. *On the solutions of two special types of Riccati difference equation via Fibonacci numbers*, Adv. Differ. Equ. **2013**, Article 174, 2013.

[21] Vajda, S.A. *Fibonacci and Lucas numbers, and the golden section* (Ellis-Horwood, 1989).

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 447 – 453

μ-paracompact and g_μ-paracompact generalized topological spaces

A. Deb Ray * and Rakesh Bhowmick †

Abstract

This paper defines generalizations of paracompactness on generalized topological spaces (GTS) and establishes that paracompactness, near paracompactness and several other paracompact-like properties follow as special cases, by choosing the GT suitably. Also, the generalizations of locally finite and closure preserving collections in a GTS, have been studied, pointing out their interrelations. Finally, it has been observed that the celebrated theorem of E.Michael in the context of regular paracompact spaces follow as a corollary to a result achieved in this paper.

Keywords: γ_μ-closure, μ-locally finite, g_μ-locally finite, μ-paracompact, g_μ-paracompact, γ_μ-regular.

2000 AMS Classification: 54A05, 54D20.

Received : 23.12.2013 *Accepted* : 27.02.2015 *Doi* : 15672/HJMS.20164512493

1. Introduction & Preliminaries

Paracompactness [2] is a very natural and perhaps the most successful generalization of compactness. Various eminent mathematicians of different times have studied several stronger as well as weaker forms of paracompactness, the most widely investigated one being near paracompactness [5]. The main purpose of this paper is to define a generalization of paracompactness on generalized topological spaces (GTS) which is a wider framework than topological spaces; and establish that by choosing the GT suitably paracompactness as well as near paracompactness follow as special cases. Also, it has been observed that by suitably choosing the generalized topology one may think of various paracompact-like spaces other than the two mentioned above.

*Department of Mathematics,West Bengal State University, Barasat, North 24 Paraganas, Kolkata - 700126, INDIA;
Email : `atasi@hotmail.com`

†Department of Mathematics, West Bengal State University, Barasat, North 24 Parganas, Kolkata - 700126, INDIA;
Email : `r.bhowmick88@gmail.com`

In section 2, we introduce a closure operator γ_μ on a GTS (X, μ) and find certain relationships among the generalized closure operator on (X, μ) and the newly defined one. We have also generalized and studied local finite and closure preserving collections of sets with respect to the GT μ and the operator γ_μ.
In section 3, we define and investigate generalization of paracompactness which we have called μ-paracompactness and g_μ-paracompactness. The celebrated theorem of E.Michael in the context of regular paracompact spaces follow as a corollary to a result achieved in this paper for more general setting what we have called γ_μ-regular g_μ-paracompact GTS.
Let X be a nonempty set and μ be a collection of subsets of X (i.e. $\mu \subseteq \mathcal{P}(X)$). μ is called a generalized topology (briefly GT) [1] on X iff $\phi \in \mu$ and $G_\lambda \in \mu$ for $\lambda \in \Lambda(\neq \phi)$ implies $\cup_{\lambda\in\Lambda} G_\lambda \in \mu$. The pair (X, μ) is called a generalized topological space (briefly GTS). The elements of μ are called μ-open sets and their complements are called μ-closed sets. The generalized closure of a subset S of X, denoted by $c_\mu(S)$, is the intersection of all μ-closed sets containing S. The set of all μ-open sets containing an element $x \in X$ is denoted by $\mu(x)$. The set of all open, δ-open [7] and θ-open [7], subsets of X are denoted respectively by $\tau(X)$ (or τ), $\Delta(X)$ (or Δ) and $\Theta(X)$ (or Θ). In what follows we shall denote the set of all natural numbers, integers and real numbers respectively by $\mathbb{N}, \mathbb{Z}$ and $\mathbb{R}$.

2. Generalized local finite and Generalized closure preserving collection

Before generalizing locally finite and closure preserving collections we introduce a new operater on a GTS (X, μ) and show that such operator actually give rise to a topology on X.

2.1. Definition. Let (X, μ) be a GTS. Then for each $x \in X$ we define
$\mu^*(x) = \{\cap_{i=1}^n W_i : W_i \in \mu(x), \forall\ i = 1, 2, \cdots, n; n \in \mathbb{N}\}$

2.1. Remark. For any $x \in X$, $\mu(x) \subseteq \mu^*(x)$ and $\mu^*(x)$ is closed under finite intersection.

2.2. Definition. Let (X, μ) be a GTS. Then γ_μ-closure of a subset S of X, denoted by $\gamma_\mu(S)$ is defined by
$\gamma_\mu(S) = \{x \in X : V \cap S \neq \phi \text{ for all } V \in \mu^*(x)\}$

The table below shows that how γ_μ-closure operator unifies several closure type operator.

μ	γ_μ
$P(X)$	identity operator
τ	closure operator
Δ	δ-closure operator [7]
Θ	θ-closure operator [7]

In a GTS (X, μ) γ_μ-closure operator satisfies the following properties (i) $\gamma_\mu(\phi) = \phi$, (ii) $S \subseteq X \Rightarrow S \subseteq \gamma_\mu(S) \subseteq c_\mu(S)$ and $\gamma_\mu(\gamma_\mu(S)) = \gamma_\mu(S)$, (iii) $A \subseteq B \subseteq X \Rightarrow \gamma_\mu(A) \subseteq \gamma_\mu(B)$ and $\gamma_\mu(A \cup B) = \gamma_\mu(A) \cup \gamma_\mu(B)$. Clearly γ_μ is a closure operator on X and hence give rise to a topology on X, denoted by μ^* and given by $\mu^* = \{S \subseteq X : \gamma_\mu(X\backslash S) = X\backslash S\}$. The elements of μ^* are called μ^*-open sets and the complements are called μ^*-closed sets. In fact for every $x \in X$, $W \in \mu^*(x)$ is a -open set. From now we may call the elements of $\mu^*(x)$ the open neighbourhoods of x.
In particular, if μ itself is a topology on X then $\mu = \mu^*$. Otherwise μ^* is finer than GT μ.

2.1. Example. Let us consider the set $X = \{a, b, c\}$. Then $\mu = \{\phi, \{a, b\}, \{a, c\}, X\}$ is clearly a GT on X. Let $S = \{b, c\}$. Now for any $V \in \mu(a)$, $V \cap S \neq \phi$ i.e. $a \in c_\mu(S)$, but $a \notin S$. So $c_\mu(S) \neq S$. Therefore S is not a μ-closed set. Again if we take $V_1 = \{a, b\}$ and $V_2 = \{a, c\}$ then $(V_1 \cap V_2) \cap S = \phi$. i.e. $a \notin \gamma_\mu(S)$. This implies that $S = \gamma_\mu(S)$ (using $S \subseteq \gamma_\mu(S)$ and $X = \{a, b, c\}$). Therefore S is a μ^*-closed set.

With the help of μ-open and μ^*-open sets we generalize the known concepts of local finite and closure preserving collections.

2.3. Definition. A family $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ of sets in a GTS (X, μ) is called
(i) μ-locally finite (resp. g_μ-locally finite) if for each $x \in X$ there exists $V \in \mu(x)$ (resp. $V \in \mu^*(x)$) such that V intersects at most finitely many members of $\mathcal{U}$ i.e. $V \cap U_\alpha \neq \phi$ for at most finitely many indices α.
(ii) μ-closure preserving (resp. γ_μ-closure preserving) if for any subcollection $\mathcal{V}$ of $\mathcal{U}$, $c_\mu[\cup\{V : V \in \mathcal{V}\}] = \cup\{c_\mu V : V \in \mathcal{V}\}$ (resp. $\gamma_\mu[\cup\{V : V \in \mathcal{V}\}] = \cup\{\gamma_\mu V : V \in \mathcal{V}\}$).

In general, every μ-locally finite family on a GTS (X, μ) is a g_μ-locally finite family but not conversely.

2.2. Example. Let $X = \mathbb{Z}$. Then $\mu = \{A \subseteq \mathbb{Z} : A \text{ is infinite }\} \cup \{\phi\}$ forms a GT on X. Let us construct $I_n = \{x \in X : x \geq n\}, n \in \mathbb{N}$ and $J_n = \{x \in X : x \leq -n\}, n \in \mathbb{N}$. Now consider the family $\mathcal{U} = \{I_n\} \cup \{J_n\}$. Then for any $x \in X, V \in \mu(x)$ intersects infinitely many members of $\mathcal{U}$. Therefore $\mathcal{U}$ is not a μ-locally finite family. Again for any $x \in X$ if we take $V_1 = \{y \in X : y \geq x\}$ and $V_2 = \{y \in X : y \leq x\}$ then $V_1, V_2 \in \mu(x)$ and $V_1 \cap V_2 (= \{x\}) \in \mu^*(x)$. If $x > 0$ then $V_1 \cap V_2$ intersects only $I_1, I_2, \cdots, I_x$. If $x < 0$ then $V_1 \cap V_2$ intersects only $J_1, J_2, \cdots, J_x$. If $x = 0$ then $V_1 \cap V_2$ intersects no members of $\mathcal{U}$. It follows that $\mathcal{U}$ is a g_μ-locally finite family.

But when we take μ as τ then both coincide with locally finite [2]. Moreover, when we take μ as τ then both of μ-closure preserving and γ_μ-closure preserving property coincide with closure preserving.

2.1. Theorem. If $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ is a μ-locally finite (resp. g_μ-locally finite) family on a GTS (X, μ). Then
(i) any subcollection of $\mathcal{U}$ is also μ-locally finite $(resp. g_\mu$-locally finite).
(ii) $c_\mu \mathcal{U} = \{c_\mu(U) : U \in \mathcal{U}\}$ (resp. $\gamma_\mu \mathcal{U} = \{\gamma_\mu(U) : U \in \mathcal{U}\}$) is also μ-locally finite (resp. g_μ-locally finite).

Proof. (i) Straightforward.
(ii) Let $x \in X$. Then since $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ is μ-locally finite (resp. g_μ-locally finite), there exists $V \in \mu(x)$ (resp. $V \in \mu^*(x)$) such that $V \cap U_\alpha \neq \phi$ for at most finitely many α's. Now we show that $V \cap c_\mu(U_\alpha) \neq \phi$ (resp. $V \cap \gamma_\mu(U_\alpha) \neq \phi$) for at most finitely many α's. Let $y \in V$, then $V \in \mu(y)$ (resp. $V \in \mu^*(y)$) is such that V intersects at most finitely many U_α's. From the definition of $c_\mu(U_\alpha)$ (resp. $\gamma_\mu(U_\alpha)$), $y \in c_\mu(U_\alpha)$ (resp. $\gamma_\mu(U_\alpha)$) for at most finitely many U_α's. This implies that $V \cap c_\mu(U_\alpha) \neq \phi$ (resp. $V \cap \gamma_\mu(U_\alpha) \neq \phi$) for at most finitely many α's, as desired. □

2.2. Theorem. If $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ is a g_μ-locally finite family on a GTS (X, μ), then $\mathcal{U}$ is γ_μ-closure preserving.

Proof. Let $\mathcal{B}$ be any subcollection of $\mathcal{U}$. We show that $\gamma_\mu[\cup\{B : B \in \mathcal{B}\}] = \cup\{\gamma_\mu(B) : B \in \mathcal{B}\}$. Since $\gamma_\mu(B) \subseteq \gamma_\mu[\cup\{B : B \in \mathcal{B}\}]$ for all $B \in \mathcal{B}$, $\cup\{\gamma_\mu(B) : B \in \mathcal{B}\} \subseteq \gamma_\mu[\cup\{B : B \in \mathcal{B}\}]$. Next let $x \notin \cup\{\gamma_\mu(B) : B \in \mathcal{B}\}$ Since $\mathcal{B}$ is a subcollection of a g_μ-locally finite collection $\mathcal{U}$, $\mathcal{B}$ is also g_μ-locally finite and so there exists $V \in \mu^*(x)$ such that V intersects at most finitely many members of $\mathcal{B}$, say $B_1, B_2, \cdots, B_n$. Again since $x \notin \cup\{\gamma_\mu(B) : B \in$

$\mathcal{B}\}$, $x \notin \gamma_\mu(B_i)$ for $i = 1, 2, \cdots, n$ and so there exist $W_i \in \mu^*(x)$ for $i = 1, 2, \cdots, n$ such that $W_i \cap B_i = \phi$. Let $W = V \cap W_1 \cap W_2 \cap \cdots \cap W_n$. Since $V, W_1, W_2, \cdots, W_n \in \mu^*(x)$, $W \in \mu^*(x)$. So we have $W \in \mu^*(x)$ such that $W \cap [\cup\{B : B \in \mathcal{B}\}] = \phi$. This implies that $x \notin \gamma_\mu[\cup\{B : B \in \mathcal{B}\}]$. Therefore $\gamma_\mu[\cup\{B : B \in \mathcal{B}\}] \subseteq \cup\{\gamma_\mu(B) : B \in \mathcal{B}\}$, consequently $\gamma_\mu[\cup\{B : B \in \mathcal{B}\}] = \cup\{\gamma_\mu(B) : B \in \mathcal{B}\}$. □

2.1. Corollary. The arbitrary union of μ^*-closed sets from a g_μ-locally finite family in a GTS (X, μ) is also μ^*-closed.

Let (X, μ) be a GTS. Then a family $\mathcal{U} = \{U_\alpha : \alpha \in \Lambda\}$ on X is said to be a covering of X if $X = \cup_{\alpha\in\Lambda} U_\alpha$. Moreover, if each U_α is μ-open (resp. μ-closed, μ^*-open, μ^*-closed) then $\mathcal{U}$ is called μ-open (resp. μ-closed, μ^*-open, μ^*-closed) covering of X.
Let $\mathcal{U}$ and $\mathcal{V}$ be two covering of X, then $\mathcal{V}$ is said to be subcovering of $\mathcal{U}$ if each member of $\mathcal{V}$ is also a member of $\mathcal{U}$. Moreover if $\mathcal{V}$ contains finite (resp. countable) number of members, then $\mathcal{V}$ is called finite (resp. countable) subcovering of $\mathcal{U}$.
Let (X, μ) be a GTS. Then a family $\mathcal{U} = \{U_\alpha : \alpha \in \Lambda\}$ on X is said to be a point finite covering of X if for each $x \in X$, there exists at most finitely many indices $\alpha \in \Lambda$ such that $x \in A_\alpha$. Moreover, if each member of $\mathcal{U}$ is μ-open then $\mathcal{U}$ is called point finite μ-open covering of X.
Let (X, μ) be a GTS. Let $\mathcal{U}$ and $\mathcal{V}$ be two covering of X, then $\mathcal{V}$ is said to refine (or be a refinement of) $\mathcal{U}$ if for each $V \in \mathcal{V}$ there exists $U \in \mathcal{U}$ such that $V \subseteq U$. We write $\mathcal{V} \prec \mathcal{U}$. If $\mathcal{W} \prec \mathcal{U}$ and $\mathcal{W} \prec \mathcal{V}$ then $\mathcal{W}$ is called common refinement of $\mathcal{U}$ and $\mathcal{V}$.

2.2. Remark. Each subcovering of a covering is a refinement of that covering.

2.3. Theorem. Let (X, μ) be a GTS. Let $\mathcal{A} = \{A_\alpha : \alpha \in A\}$ and $\mathcal{B} = \{B_\beta : \beta \in B\}$ be two covering of X. Then

(1) $\mathcal{A} \wedge \mathcal{B} = \{A_\alpha \cap B_\beta : (\alpha, \beta) \in A \times B\}$ is a covering of X, refining both $\mathcal{A}$ and $\mathcal{B}$. Furthermore if both $\mathcal{A}$ and $\mathcal{B}$ are μ-locally finite (resp.g_μ-locally finite), so also is $\mathcal{A} \wedge \mathcal{B}$.
(2) any common refinement of $\mathcal{A}$ and $\mathcal{B}$ is also a refinement of $\mathcal{A} \wedge \mathcal{B}$.

Proof. Straightforward. □

2.4. Definition. Let (X, μ) be a GTS. A refinement $\{B_\beta : \beta \in \mathcal{B}\}$ of $\{A_\alpha : \alpha \in \mathcal{A}\}$ is called a precise refinement if $\mathcal{A} = \mathcal{B}$ and $B_\alpha \subseteq A_\alpha$, for each α.

2.4. Theorem. Let (X, μ) be a GTS. If a covering $\{A_\alpha : \alpha \in \mathcal{A}\}$ of X has a μ-locally finite (resp. g_μ-locally finite) refinement $\{B_\beta : \beta \in \mathcal{B}\}$ that covers X, then it has a precise μ-locally finite (resp. g_μ-locally finite) refinement $\{C_\alpha : \alpha \in \mathcal{A}\}$ that covers X. Furthermore, if each B_β is μ-open then each C_α can be chosen to be μ-open also.

Proof. Define a map $\phi : \mathcal{B} \to \mathcal{A}$ by assigning each $\beta \in \mathcal{B}$ to some $\alpha \in \mathcal{A}$ such that $B_\beta \subseteq A_\alpha$. For each α, let $C_\alpha = \cup\{B_\beta : \phi(\beta) = \alpha\}$, some C_α may be empty. Clearly $C_\alpha \subseteq A_\alpha$ for each α i.e $\{C_\alpha : \alpha \in \mathcal{A}\}$ is a refinement of $\{A_\alpha : \alpha \in \mathcal{A}\}$. Also since $\{B_\beta : \beta \in \mathcal{B}\}$ is a covering of X, each $\mathcal{B}_\beta$ appears somewhere $\{C_\alpha : \alpha \in \mathcal{A}\}$ and so $\{C_\alpha : \alpha \in \mathcal{A}\}$ is a covering of X. Again since $\{B_\beta : \beta \in \mathcal{B}\}$ is μ-locally finite (resp. g_μ-locally finite), for each $x \in X$ there exists $V \in \mu(x)$ (resp.$V \in \mu^*(x)$) such that V intersects at most finitely many $\mathcal{B}_\beta$'s and consequently finitely many C_α's. This implies that $\{C_\alpha : \alpha \in \mathcal{A}\}$ is μ-locally finite (resp. g_μ-locally finite). Hence the first part follows. For the second part, if each $\mathcal{B}_\beta$ is μ-open then clearly each C_α is also μ-open. □

2.3. Remark. In the above theorem μ-locally finite (resp. g_μ-locally finite) can be replaced by point finite.

2.5. Theorem. Let $\{E_\alpha : \alpha \in \Lambda\}$ be any family of sets on a GTS (X, μ) and $\{B_\beta : \beta \in \mathcal{B}\}$ be any g_μ-locally finite μ^*-closed covering of X. If each B_β intersects at most finitely many sets E_α, then each E_α is contained in a μ^*-open set $U(E_\alpha)$ such that the family $\{U(E_\alpha) : \alpha \in \Lambda\}$ is g_μ-locally finite.

Proof. For each α define $U(E_\alpha) = X \setminus \cup \{B_\beta : B_\beta \cap E_\alpha = \phi\}$. Since, $\{B_\beta\}$is g_μ-locally finite family of μ^*-closed sets $U(E_\alpha)$ is μ^*-open (since, $\cup\{B_\beta : B_\beta \cap E_\alpha = \phi\}$ is μ^*-closed, by corollary 2.1). Also $E_\alpha \subseteq U(E_\alpha)$ (since, $x \notin U(E_\alpha) \Rightarrow \exists\, x \in B_{\beta_0}$ for some $\beta_0 \in \mathcal{B}$ such that $B_{\beta_0} \cap E_\alpha = \phi$. Again $x \in B_{\beta_0}$ and $B_{\beta_0} \cap E_\alpha = \phi \Rightarrow x \notin E_\alpha$. i.e. $x \notin U(E_\alpha) \Rightarrow x \notin E_\alpha$).
We now prove that $\{U(E_\alpha) : \alpha \in \mathcal{A}\}$ is g_μ-locally finite. Since, $\{B_\beta : \beta \in \mathcal{B}\}$ is g_μ-locally finite, for any given $x \in X$ there exists $V \in \mu^*(x)$ such that V intersects at most finitely many B_β's say $B_{\beta_1}, B_{\beta_2}, ..., B_{\beta_n}$. Obviously V contained in $\cup_{i=1}^n B_{\beta_i}$, as $\{B_\beta\}$ forms a covering of X. Since $B_\beta \cap U(E_\alpha) \neq \phi$ iff $B_\beta \cap E_\alpha \neq \phi$ (since, $B_\beta \cap E_\alpha \neq \phi$ iff $B_\beta \nsubseteq \cup\{B_\beta : B_\beta \cap E_\alpha = \phi\}$ iff $B_\beta \cap (X \setminus \cup \{B_\beta : B_\beta \cap E_\alpha = \phi\}) \neq \phi$ iff $B_\beta \cap U(E_\alpha) \neq \phi$) and each $B_{\beta_i}, i = 1, 2, \cdots, n$ intersects at most finitely many E_α, $\cup_{i=1}^n B_{\beta_i}$ intersects at most finitely many $U(E_\alpha)$. Thus we have $V \in \mu^*(x)$ such that V intersects at most finitely many $U(E_\alpha)$ (since,$V \subseteq \cup_{i=1}^n B_{\beta_i}$) and so $\{U(E_\alpha) : \alpha \in \Lambda\}$ is g_μ-locally finite. □

2.2. Corollary. Let $\{E_\alpha : \alpha \in \Lambda\}$ be any family of sets on a GTS (X, μ) with $\mu = \mu^*$ and $\{B_\beta : \beta \in \mathcal{B}\}$ be any μ-locally finite μ-closed covering of X. If each B_β intersects at most finitely many sets E_α, then each E_α is contained in a μ-open set $U(E_\alpha)$ such that the family $\{U(E_\alpha) : \alpha \in \Lambda\}$ is μ-locally finite.

3. μ-paracompactness and g_μ-paracompactness

In this section we define generalized paracompactness to unify the existing concept of paracompact and nearly paracompact spaces. We see that many more paracompact-like properties may also be obtained by choosing the generalized topology suitably.

3.1. Definition. A GTS (X, μ) is said to be μ-paracompact (resp. g_μ-paracompact) if every μ-open covering of X has a μ-locally finite (resp. g_μ-locally finite) μ-open refinement that covers X.

3.1. Remark. g_μ-paracompactness is a generalization of μ-paracompactness, since every μ-paracompact GTS is a g_μ-paracompact GTS, but not conversely in general. If we take μ as τ then both μ-paracompact and g_μ-paracompact coincide with paracompact. If we take μ as Δ then both coincide with nearly paracompact.

3.2. Definition. A GTS (X, μ) is said to be μ-compact [6] (resp. μ-*Lindelöf*) if every μ-open covering of X has a finite (resp. countable) subcovering.

3.2. Remark. In general, every μ-compact GTS (X, μ) is μ-*Lindelöf*, but not conversely.

3.1. Theorem. Let (X, μ) be a GTS. If (X, μ) is μ-compact then it is also μ-paracompact.

Proof. Straightforward. □

The converse of above theorem is not true in general, which follows from the following example:

3.1. Example. Let $X = \mathbb{Z}, \mu=$ discrete topology on X. then $\{\{n\} : n \in \mathbb{Z}\}$ is a μ-open covering of X which has no finite subcover but every μ-open cover of X has a μ-locally finite μ-open refinement $\{\{n\} : n \in \mathbb{Z}\}$ that covers X (since, $\{\{n\} : n \in \mathbb{Z}\}$ is a refinment of every μ-open cover of X).

3.3. Definition. [3] Let (X, μ) be a GTS. Then (X, μ) is said to be μ-regular if for any $x \in X$ and a μ-closed set F not containing x, there exist two disjoint μ-open sets U and V such that $x \in U$ and $F \subseteq V$.

3.4. Definition. Let (X, μ) be a GTS. Then (X, μ) is said to be γ_μ-regular if for any $x \in X$ and a μ-open set U containing x, there exists a μ-open set V contaning x such that $\gamma_\mu(V) \subseteq U$.

3.2. Theorem. Let (X, μ) be a GTS. If (X, μ) is μ-regular then it is also γ_μ-regular.

Proof. Straightforward. $\square$

The converse is not neccesarily true. This is observed in the following example:

3.2. Example. Let $X = \{a, b, c\}$ and $\mu = \{\phi, \{a, b\}, \{b, c\}, \{c, a\}, X\}$. Then μ be a GT on X. Here $c_\mu(U) = X$ and $\gamma_\mu(U) = U$ for every μ-open set containing $x \in X$. It is easy to check that X is γ_μ-regular but not μ-regular.

3.3. Theorem. For any γ_μ-regular GTS (X, μ), (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) hold, where
(1) (X, μ) is g_μ-paracompact.
(2) Every μ-open cover of X has a μ-open refinement that covers X and can be decomposed into at most countable collection of g_μ-locally finite families of μ-open sets.
(3) Each μ-open cover of X has a g_μ-locally finite refinement that cover X.
(4) Each μ-open cover of X has a μ^*-closed g_μ-locally finite refinement that covers X.

Proof. (1) $\Rightarrow$ (2) Straightforward.
(2) $\Rightarrow$ (3) Let $\{U_\beta : \beta \in \mathcal{B}\}$ be any μ-open covering of X. By (2) there exists an μ-open covering $\{V_{n,\alpha} : (n, \alpha) \in \mathbb{N} \times \mathcal{A}\}$, which is a refinement of $\{U_\beta : \beta \in \mathcal{B}\}$, where for each $n_0 \in \mathbb{N}$, the family $\{V_{n_0,\alpha} : \alpha \in \mathcal{A}\}$ is g_μ-locally finite (not necessarily a covering). For each $n \in \mathbb{N}$, let $W_n = \cup_\alpha V_{n,\alpha}$, then $\{W_n, n \in \mathbb{N}\}$ is a μ-open covering of X. For each $i \in \mathbb{N}$ define $A_i = W_i \setminus \cup_{j=1}^{i-1} W_j$. We now show that $\{A_i\}$ is a g_μ-locally finite covering of X. For each $x \in X$, let W_{i_0} is the first member of $\{W_n, n \in \mathbb{N}\}$ such that $x \in W_{i_0}$. Then it is clear that $x \in A_{i_0}$, hence $\{A_i\}$ is a covering of X. Again $W_{i_0} \cap A_i = \phi$ for each $i > i_0$ i.e. we have $W_{i_0} \in \mu(x) \subseteq \mu^*(x)$ such that W_{i_0} intersects at most finitely many members of $\{A_i\}$. Hence $\{A_i\}$ is g_μ-locally finite.
We now show that $\mathcal{K} = \{A_n \cap V_{n,\alpha} : (n, \alpha) \in \mathbb{N} \times \mathcal{A}\}$ is g_μ-locally finite refinement of $\{U_\beta : \beta \in \mathcal{B}\}$ that covers X. For any $A_n \cap V_{n,\alpha} \in \mathcal{K}$, since $\{V_{n,\alpha} : (n, \alpha) \in \mathbb{N} \times \mathcal{A}\}$ is a refinement of $\{U_\beta : \beta \in \mathcal{B}\}$ there exists U_β such that $A_n \cap V_{n,\alpha} \subseteq V_{n,\alpha} \subseteq U_\beta$. Hence $\mathcal{K}$ is a refinement of $\{U_\beta : \beta \in \mathcal{B}\}$. Again $\mathcal{K}$ is obviously a covering of X (since for $x \in X$, $\exists\ A_n$ such that $x \in A_n \Rightarrow x \in W_n = \cup_\alpha V_{n,\alpha} \Rightarrow x \in V_{n,\alpha_0}$ for some $\alpha_0 \in \mathcal{A}$ i.e. $x \in A_n \cap V_{n,\alpha_0}$ for some $(n, \alpha_0) \in \mathbb{N} \times \mathcal{A}$). Next let $x \in X$. Then since $\{A_n : n \in \mathbb{N}\}$ is g_μ-locally finite, there exists $W \in \mu^*(x)$ such that W intersects at most finitely many member of $\{A_n : n \in \mathbb{N}\}$ say, $A_{n_1}, A_{n_2}, \cdots, A_{n_r}$. Again since $\{V_{n_j,\alpha} : \alpha \in \mathcal{A}\}$, (for $j = 1, 2, \cdots, r$) is g_μ-locally finite we have $W_{n_j} \in \mu^*(x)$, (for $j = 1, 2, \cdots, r$) such that W_{n_j} intersects atmost finitely many $V_{n_j,\alpha}$'s. Let $V = W \cap W_{n_1} \cap W_{n_2} \cdots \cap W_{n_r}$ then since $W, W_{n_j} \in \mu^*(x)$, $V \in \mu^*(x)$. So we have $V \in \mu^*(x)$ such that V intersects at most finitely many member of $\mathcal{K}$. Hence $\mathcal{K}$ is g_μ-locally finite. Thus $\mathcal{K}$ is the required g_μ-locally finite refinement of $\{U_\beta : \beta \in \mathcal{B}\}$ that covers X.
(3) $\Rightarrow$ (4) Let $\mathcal{U}$ be a μ-open covering of X. With each $y \in X$, associate a definite $U_y \in \mathcal{U}$ containing it and then since X is γ_μ-regular, there exists a μ-open set V_y such that $y \in V_y \subseteq \gamma_\mu(V_y) \subseteq U_y$. The family $\{V_y : y \in X\}$ is then a μ-open covering and by (2) and theorem 2.4 it has a precise g_μ-locally finite refinement $\{A_y : y \in X\}$. Since $\{\gamma_\mu(A_y) : y \in X\}$ is also g_μ-locally finite (by theorem 2.1) and $\gamma_\mu(A_y) \subseteq \gamma_\mu(V_y) \subseteq U_y$ for each y, $\{\gamma_\mu(A_y) : y \in X\}$ is the desired refinement. $\square$

3.3. Remark. For a g_μ-regular GTS (X, μ), if we take μ as μ^* then all the conditions (1) - (4) stated in the above theorem become equivalent. We have already proved (1)⇒(2), (2)⇒(3) and (3)⇒(4). So if we show (4)⇒(1) then our purpose will be fulfilled.

3.4. Theorem. For any γ_μ-regular GTS (X, μ) with $\mu = \mu^*$, if each μ-open cover of X has a μ^*-closed g_μ-locally finite refinement that covers X then (X, μ) is g_μ-paracompact.

Proof. Let $\mathcal{U}$ be any μ-open covering of X and ξ be any μ^*-closed g_μ-locally finite refinement of it. Since $\mu = \mu^*$, ξ is a μ-closed μ-locally finite refinement. Then for each $x \in X$, there exists a μ-open set V_x containing x such that V_x intersects at most finitely many sets E of ξ. Using the μ-open covering $\{V_x : x \in X\}$, by given hypothesis we get a μ^*-closed g_μ-locally finite and hence a μ-closed μ-locally finite refinement $\mathcal{B}$ that covers X. Since each B of $\mathcal{B}$ intersects at most finitely many sets E of ξ it follows from that we can enlarge each E to an μ-open set $G(E)$ such that $\{G(E)\}$ is μ-locally finite (by corollary 2.2). Associating with each E a single set $U(E) \in \mathcal{U}$ containing E, it is evident that $\{G(E) \cap U(E)\}$ is an μ-open μ-locally finite refiinement of $\mathcal{U}$. □

If we consider a regular topological space (X, τ) and choose in particular the GT as τ then from theorem 3.3 and theorem 3.4 we obtain E.Michael's theorem. On the other hand, if $\mu = \delta$-open sets of (X, τ) then we obtain a characterization parallel to E.Michael's theorem for almost regular nearly paracompact spaces [4].

References

[1] Á.Császár, *Generalized topology, generalized continuity*, Acta Math. Hungar., $96(4)(2002), 351 - 357$.

[2] J.Dieudonné, *Une généralisation des espaces compacts*, J.Math. Pures et. Appl. $23(1944), 65 - 76$.

[3] T.Noiri and B.Roy, *Unification of generalized open sets on Topological spaces*, Acta Math. Hunger., $130(4)(2011), 349 - 357$.

[4] M.N.Mukherjee and A.Debray, *On nearly paracompact spaces via regular even covers*, Matematnykn Bechnk, $50(1998), 23 - 29$.

[5] M.K.Singal and S.P.Arya, *On nearly paracompact spaces.* Matematički Vesnik, $6(21)(1969), 3 - 16$.

[6] R.Shen, *Remarks on Products of generalized topology*, Acta Math Hunger., $124(4)(2009), 363 - 369$.

[7] N.V.Veličko, *H-closed topological spaces*, Amer. Math. Soc. Transl., $78(2)(1968)103 - 118$.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 455 – 471

Distribution of zeros of sublinear dynamic equations with a damping term on time scales

Samir. H. Saker* and Ramy. R. Mahmoud†‡

Abstract

In this paper, for a second order sublinear dynamic equation with a damping term we will study the lower bounds of the distance between zeros of a solution and/or its derivatives and then establish some new criteria for disconjugacy and disfocality. Our results present a slight improvement to some results proved in the litrature. As a special case when $\mathbb{T} = \mathbb{R}$, for a second order linear differential equation, we get some results proved by Brown and Harris as a consequence of our results. The results will be proved by employing the time scales Hölder inequality, the time scales chain rule and some new dynamic Opial-type inequalities designed and proved for this purpose.

Keywords: Disconjugacy, Disfocality, Lyapunov inequality, Opial's inequality, Sublinear dynamic equations, Time Scales.

2000 AMS Classification: 26E70, 34K11, 34C10, 34N05.

Received : 19.04.2014 *Accepted* : 30.03.2015 *Doi* : 10.15672/HJMS.20164512498

1. Introduction

In this paper, we will study the distribution of zeros of solutions of the second-order sublinear dynamic equation with a damping term

$$\left(r(t)\left(y^{\Delta}(t)\right)^{\beta}\right)^{\Delta} + p(t)\left(y^{\Delta}(t)\right)^{\beta} + q(t)\left(y^{\sigma}(t)\right)^{\beta} = 0, \quad \text{on } [a,b]_{\mathbb{T}}, \tag{1.1}$$

on an arbitrary time scale $\mathbb{T}$, where $0 < \beta \leq 1$ is a quotient of odd positive integers, r, p and q are real $rd-$continuous functions defined on $\mathbb{T}$ with $r(t) > 0$. In particular, we will find the lower bounds of the distance between zeros of a solution and/or its derivatives and prove several results related to the problems:

*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura-Egypt, Email: shsaker@mans.edu.eg

†Department of Mathematics, Faculty of Science, Fayoum University, Fayoum-Egypt, Email: rrm00@fayoum.edu.eg

‡Corresponding Author.

(i) obtain lower bounds for the spacing $b-a$ where y is a solution of (1.1) and satisfies
$y(a)=y^{\Delta}(b)=0$, or $y^{\Delta}(a)=y(b)=0$,

(ii) obtain lower bounds for the spacing between consecutive zeros of solutions of (1.1).

By a solution of (1.1) on an interval $\mathbb{I}$, we mean a nontrivial real-valued function $y\in C_{rd}(\mathbb{I})$, which has the property that $r(t)\,y^{\Delta}(t)\in C_{rd}^{1}(\mathbb{I})$ and satisfies (1.1) on $\mathbb{I}$. We say that a solution y of (1.1) has a generalized zero at t if $y(t)=0$ and has a generalized zero in $(t,\sigma(t))$ in case $y(t)\,y^{\sigma}(t)<0$ and $\mu(t)>0$. Equation (1.1) is disconjugate on the interval $[a,b]_{\mathbb{T}}$, if there is no nontrivial solution of (1.1) with two (or more) generalized zeros in $[a,b]_{\mathbb{T}}$. The solution $y(t)$ of (1.1) is said to be oscillatory if it is neither eventually positive nor eventually negative, otherwise it is oscillatory. We say that (1.1) is right disfocal (left disfocal) on $[a,b]_{\mathbb{T}}$ if the solutions of (1.1) such that $y^{\Delta}(a)=0$ ($y^{\Delta}(b)=0$) have no generalized zeros in $[a,b]_{\mathbb{T}}$. We refer the reader to the book [28] for more details about oscillation and nonoscillation theory of dynamic equations on time scales.

We note that, equation (1.1) in its general form covers several different types of differential and difference equations depending on the choice of the time scale $\mathbb{T}$. For example, when $\mathbb{T}=\mathbb{R}$, we have $\sigma(t)=t$, $\mu(t)=0$, $x^{\Delta}(t)=x^{'}(t)$ and (1.1) becomes the second-order sublinear differential equation

$$(1.2)\qquad (r(t)(x^{'}(t))^{\beta})^{'}+p(t)(x^{'}(t))^{\beta}+q(t)x^{\beta}(t)=0.$$

When $\mathbb{T}=\mathbb{Z}$, we have $\sigma(t)=t+1$, $\mu(t)=1$, $x^{\Delta}(t)=\Delta x(t)=x(t+1)-x(t)$ and (1.1) becomes the second-order difference equation

$$(1.3)\qquad \Delta(r(t)\,(\Delta x(t))^{\beta})+p(t)\,(\Delta x(t))^{\beta}+q(t)x^{\beta}(t+1))=0.$$

We present in the sequel some of the results that serve and motivate the contents on this paper. The well known existence results in the literature for disconjugacy is due to C. de la Vallée Poussin [22]. He considered the general n^{th} order linear differential equation

$$(1.4)\qquad x^{(n)}+p_0(t)x^{(n-1)}+\cdots+p_{n-1}(t)x=0,$$

where the coefficients p_i are real continuous functions on an interval $\mathbb{I}=[a,b]$, and proved that if $|p_i(t)|\le q_i$ on $\mathbb{I}$ and the inequality

$$(1.5)\qquad \sum_{i=1}^{n}\frac{q_i\,(b-a)^i}{i!}\le 1,$$

holds, then (1.4) is disconjugate (that is every nontrivial solution of (1.4) has less than n zeros on $\mathbb{I}$, multiple zeros being counted according to their multiplicity).

Lyapunov [17] investigated the best known existence result in the literature for the second order differential equation

$$(1.6)\qquad x^{''}(t)+q(t)x(t)=0,\quad t\in(a,b)\,,$$

and proved that if $x(t)$ is a solution of (1.6) with $x(a)=x(b)=0$ and $q(t)$ is a continuous and nonnegative function on the closed interval $[a,b]$, then

$$(1.7)\qquad \int_a^b q(t)dt>\frac{4}{b-a}.$$

The constant 4 is the best possible and cannot be replaced by a larger number. The inverse of (1.7) gives a sufficient condition for disconjugacy of (1.6). The Lyapunov inequality is very important and has been extended extensively in the study of various properties of ordinary differential equations, for example bounds for eigenvalues, oscillation theory, stability criteria for periodic differential equations, and estimates for intervals of disconjugacy.

Since the appearance of Lyapunov's fundamental paper, there are many improvements and generalizations of (1.7) in several papers and different conditions for the disconjugacy, for the second order differential equation (1.2) and its special cases, have been investigated by many authors. We refer the reader to the papers [12, 16, 19, 21, 30, 32]. A literature review of continuous and discrete Lyapunov-type inequalities and their applications can be found in the survey articles by Brown and Hinton [6], Cheng [8] and Tiryaki [31] and the references cited therein. Hartman in [11, Chap. XI] generalized the classical Lyapunov inequality for the second order linear differential equation

$$\left(r(t)x^{'}(t)\right)^{'}+q(t)x(t)=0, \quad r(t)>0, \tag{1.8}$$

and proved that if $x(a)=x(b)=0$, then

$$\int_a^b q^+(s)ds \geq \frac{4}{\int_a^b r^{-1}(s)ds}, \tag{1.9}$$

where $q^+(t)=\max\{0,q(t)\}$ is the nonnegative part of $q(t)$.

Cohn [9] and Kwong [14] proved that if $x(t)$ is a solution of (1.6) with $x(a)=x^{'}(c)=0$, then

$$\int_a^c (t-a)\,q(t)dt>1,$$

and similarly if $x(t)$ is a solution of (1.6) with $x^{'}(c)=x(b)=0$, then

$$\int_c^b (b-t)\,q(t)dt>1.$$

Harris and Kong [10] proved that if $x(t)$ is a solution of (1.6) with $x(a)=x^{'}(b)=0$, then

$$(b-a)\sup_{a\leq t\leq b}\left|\int_t^b q(s)\Delta s\right|>1, \tag{1.10}$$

and if instead $x^{'}(a)=x(b)=0$, then

$$(b-a)\sup_{a\leq t\leq b}\left|\int_a^t q(s)\Delta s\right|>1. \tag{1.11}$$

Brown and Hinton [7] proved that if $x(t)$ is a solution of (1.6) with $x(a)=x^{'}(b)=0$, then

$$2\int_a^b Q_1^2(t)(t-a)dt>1, \tag{1.12}$$

where $Q_1(t)=\int_t^b q(s)ds$. If instead $x^{'}(a)=x(b)=0$, *then*

$$2\int_a^b Q_2^2(t)(b-t)dt>1, \tag{1.13}$$

where $Q_2(t) = \int_a^t q(s)ds$.

In [29] the author considered the equation (1.2) when $\beta = 1$ and established some criteria for disconjugacy and disfocality of solutions in an interval $\mathbb{I} = [a, b] \subset \mathbb{R}$. He also applied Hardy and Wirtinger type inequalities and established an explicit formula for the lower bound of the first eigenvalue of the eigenvalue problem

$$(1.14) \qquad -\left(x^{'}(t)\right)^{'} - p(t)x^{'}(t) + q(t)x(t) = \lambda x(t), \;\; x(a) = x(b) = 0.$$

For the study of dynamic equations on time scales, Bohner et al. [5] considered the dynamic equation

$$(1.15) \qquad x^{\Delta\Delta}(t) + q(t)x^{\sigma}(t) = 0,$$

and proved a new Lyapunov dynamic inequality on a time scale $\mathbb{T}$, where $q(t)$ is a positive rd-continuous function defined on $\mathbb{T}$. Saker [24], employed some new dynamic Opial type inequalities and established new Lyapunov type inequalities for the equation

$$(1.16) \qquad \left(r(t)\, x^{\Delta}(t)\right)^{\Delta} + q(t)x^{\sigma}(t) = 0, \qquad \text{on } [a, b]_{\mathbb{T}},$$

where r; q are rd-continuous functions satisfy the conditions

$$\int_a^b \frac{1}{r(t)} \Delta t < \infty, \quad \text{and} \quad \int_a^b q(t)\Delta t < \infty.$$

For more results related to these results, we refer the reader to the papers by Karpuz [13] and Saker [23, 27] and the references cited therein.

Following this trend and to develop the study of oscillation of second-order sublinear dynamic equations on time scales, we will prove several results related to the problems $(i) - (ii)$. The rest of the paper is divided into three sections: In Section 2, we present some basic concepts of the time scales calculus and present some dynamic Opial-type inequalities, which are also interesting results in their own right, that will be used in the proof of our main results. In Section 3, we first prove some new generalizations of Opial's inequality on an arbitrary time scale $\mathbb{T}$, then we will employ these inequalities to prove several results related to the problems $(i)-(ii)$ above. In Section 4, we will discuss some special cases of the results. The results yield some conditions for disfocality and disconjugacy for equation (1.1).

2. Preliminaries and Some Opial's Inequalities

In this section, we briefly give some essentials of time scales calculus which are necessary for our results, then we present some dynamic Opial-type inequalities on an arbitrary time scale $\mathbb{T}$.

A time scale $\mathbb{T}$ is an arbitrary nonempty closed subset of the real numbers $\mathbb{R}$. We assume throughout that $\mathbb{T}$ has the topology that it inherits from the standard topology on the real numbers $\mathbb{R}$. The forward jump operator and the backward jump operator are defined by: $\sigma(t) := \inf\{s \in \mathbb{T} : s > t\}$, $\rho(t) := \sup\{s \in \mathbb{T} : s < t\}$, where $\sup \emptyset = \inf \mathbb{T}$. A point $t \in \mathbb{T}$, is said to be left–dense if $\rho(t) = t$ and $t > \inf \mathbb{T}$, is right–dense if $\sigma(t) = t$, is left–scattered if $\rho(t) < t$ and right–scattered if $\sigma(t) > t$. A function $g : \mathbb{T} \to \mathbb{R}$ is said to be right–dense continuous (rd–continuous) provided g is continuous at right–dense points and at left–dense points in $\mathbb{T}$, left hand limits exist and are finite. The set of all such rd–continuous functions is denoted by $C_{rd}(\mathbb{T})$. We denote by $C_{rd}^{(n)}(\mathbb{T})$ the space of all functions $f \in C_{rd}(\mathbb{T})$ such that $f^{\Delta_i} \in C_{rd}(\mathbb{T})$ for $i = 0, 1, 2, ..., n$ for $n \in \mathbb{N}$.

The graininess function μ for a time scale $\mathbb{T}$ is defined by $\mu(t) := \sigma(t) - t \geq 0$, and for any function $f : \mathbb{T} \to \mathbb{R}$ the notation $f^{\sigma}(t)$ denotes $f(\sigma(t))$. We assume that $\sup \mathbb{T} = \infty$, and define the time scale interval $[a,b]_{\mathbb{T}}$ by $[a,b]_{\mathbb{T}} := [a,b] \cap \mathbb{T}$. The three most popular examples of calculus on time scales are differential calculus, difference calculus, and quantum calculus, i.e., when $\mathbb{T} = \mathbb{R}$, $\mathbb{T} = \mathbb{N}$ and $\mathbb{T} = q^{\mathbb{N}_0} = \{q^t : t \in \mathbb{N}_0\}$ where $q > 1$. For more details of time scale analysis we refer the reader to the two books by Bohner and Peterson [2], [3] which summarize and organize much of the time scale calculus. In this paper, we will refer to the (delta) integral which we can define as follows. If $G^{\Delta}(t) = g(t)$, then the Cauchy (delta) integral of g is defined by $\int_a^t g(s)\Delta s := G(t) - G(a)$. It can be shown (see [2]) that if $g \in C_{rd}(\mathbb{T})$, then the Cauchy integral $G(t) := \int_{t_0}^t g(s)\Delta s$ exists, $t_0 \in \mathbb{T}$, and satisfies $G^{\Delta}(t) = g(t)$, $t \in \mathbb{T}$. A simple consequence of Keller's chain rule [2, Theorem 1.90] is given by

$$(x^{\gamma}(t))^{\Delta} = \gamma \int_0^1 [hx^{\sigma}(t) + (1-h)x(t)]^{\gamma-1}\, dh\, x^{\Delta}(t), \tag{2.1}$$

and the integration by parts formula on time scales is given by

$$\int_a^b u(t)v^{\Delta}(t)\Delta t = [u(t)v(t)]_a^b - \int_a^b u^{\Delta}(t)v^{\sigma}(t)\Delta t. \tag{2.2}$$

The Hölder inequality, see [2, Theorem 6.13], on time scales is given by

$$\int_a^b |f(t)g(t)|\Delta t \le \left[\int_a^b |f(t)|^{\gamma}\Delta t\right]^{\frac{1}{\gamma}} \left[\int_a^b |g(t)|^{\nu}\Delta t\right]^{\frac{1}{\nu}}, \tag{2.3}$$

where a, $b \in \mathbb{T}$ and f, $g \in \mathrm{C}_{rd}(\mathbb{I},\mathbb{R})$, $\gamma > 1$ and $\frac{1}{\gamma} + \frac{1}{\nu} = 1$. Throughout the paper, we will assume that the functions in the statements of the theorems are nonnegative and rd-continuous functions and the integrals are assumed to exist.

For completeness, in the following, we recall some of the Opial-type inequalities that serve and motivate the contents of the paper.

In 1960 Opial [20] published an inequality involving integrals of a function and its derivative. Since the discovery of Opial's inequality much work has been done, and many papers which deal with new proofs, various generalizations, extensions and their discrete analogues have been also proved in the literature. The discrete analogy of Opial's inequality has been proved in [15]. In [4] the authors extended the Opial inequality to an arbitrary time scale $\mathbb{T}$ and proved that if $y : [0,h]_{\mathbb{T}} \to \mathbb{R}$ is delta differentiable with $y(0) = 0$, then

$$\int_0^h |y(x) + y^{\sigma}(x)| \left|y^{\Delta}(x)\right| \Delta x \le h \int_0^h \left|y^{\Delta}(x)\right|^2 \Delta x. \tag{2.4}$$

They also proved that if r and q are positive rd-continuous functions on $[0,h]_{\mathbb{T}}$, $\int_0^h \frac{\Delta x}{r(x)} < \infty$, q nonincreasing and $y : [0,h]_{\mathbb{T}} \to \mathbb{R}$ is delta differentiable with $y(0) = 0$, then

$$\int_0^h q^{\sigma}(x) \left|(y(x) + y^{\sigma}(x))\, y^{\Delta}(x)\right| \Delta x \le \int_0^h \frac{\Delta x}{r(x)} \int_0^h r(t)q(x) \left|y^{\Delta}(x)\right|^2 \Delta x. \tag{2.5}$$

In [24] the author proved that if $y : [a,\tau]_{\mathbb{T}} \to \mathbb{R}$ is delta differentiable with $y(a) = 0$, then

$$\int_a^{\tau} s(x)\, |y(x) + y^{\sigma}(x)| \left|y^{\Delta}(x)\right| \Delta x \le K_1(a,\tau) \int_a^{\tau} r(x) \left|y^{\Delta}(x)\right|^2 \Delta x, \tag{2.6}$$

where $s \in C_{rd}([a,\tau]_{\mathbb{T}},\mathbb{R})$ and r be a positive rd-continuous function on $(a,\tau)_{\mathbb{T}}$ such that $\int_a^{\tau} r^{-1}(t)\Delta t < \infty$, and

$$K_1(a,\tau) = \sqrt{2}\left(\int_a^{\tau} \frac{s^2(x)}{r(x)} \left(\int_a^x \frac{\Delta t}{r(t)}\right) \Delta x\right)^{\frac{1}{2}} + \sup_{a \le x \le \tau} \left(\mu(x)\frac{|s(x)|}{r(x)}\right).$$

In [26] the author generalized (2.6) and proved that if $y : [a,\tau]_{\mathbb{T}} \to \mathbb{R}$ is delta differentiable with $y(a) = 0$, then

$$\int_a^\tau s(x)|y(x)+y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x \le H_1(a,\tau)\int_a^\tau r(x)|y^\Delta(x)|^{\lambda+\delta}\Delta x, \tag{2.7}$$

where r, s be nonnegative rd-continuous functions on $[a,\tau]_{\mathbb{T}}$ such that $\int_a^\tau r^{\frac{-1}{\lambda+\delta-1}}(t)\Delta t < \infty$, λ, δ be positive real numbers such that $\lambda \ge 1$ and

$$\begin{aligned} H_1(a,\tau) \quad : \quad &= 2^{\lambda-1}\sup_{a\le x\le\tau}\left(\mu^\lambda(x)\frac{s(x)}{r(x)}\right) + 2^{2\lambda-1}\left(\frac{\delta}{\lambda+\delta}\right)^{\frac{\delta}{\lambda+\delta}} \\ &\times \left(\int_a^\tau \frac{(s(x))^{\frac{\lambda+\delta}{\lambda}}}{(r(x))^{\frac{\delta}{\lambda}}}\left(\int_a^x r^{\frac{-1}{\lambda+\delta-1}}(t)\Delta t\right)^{\lambda+\delta-1}\Delta x\right)^{\frac{\lambda}{\lambda+\delta}}. \end{aligned}$$

In [25] the author proved that if $y : [a,\tau]_{\mathbb{T}} \to \mathbb{R}^+$ is delta differentiable with $y(a) = 0$, then

$$\int_a^\tau s(x)|y(x)+y^\sigma(x)|^p |y^\Delta(x)|^q \Delta x \le K_2(a,\tau)\int_a^\tau r(x)|y^\Delta(x)|^{p+q}\Delta x, \tag{2.8}$$

where p, $q > 0$ such that $p \le 1$, $p+q > 1$, r, s be nonnegative rd-continuous functions such that $\int_a^\tau r^{\frac{-1}{p+q-1}}(t)\Delta t < \infty$ and

$$\begin{aligned} K_2(a,\tau) \quad : \quad &= \sup_{a\le x\le\tau}\left(\mu^p(x)\frac{s(x)}{r(x)}\right) + 2^p\left(\frac{q}{p+q}\right)^{\frac{q}{p+q}} \\ &\times \left(\int_a^\tau \frac{(s(x))^{\frac{p+q}{p}}}{(r(x))^{\frac{q}{p}}}\left(\int_a^x r^{\frac{-1}{p+q-1}}(t)\Delta t\right)^{p+q-1}\Delta x\right)^{\frac{p}{p+q}}. \end{aligned} \tag{2.9}$$

If $[a,\tau]_{\mathbb{T}}$ is replaced by $[\tau,b]_{\mathbb{T}}$, then we get the following result

$$\int_\tau^b s(x)|y(x)+y^\sigma(x)|^p |y^\Delta(x)|^q \Delta x \le K_3(\tau,b)\int_\tau^b r(x)|y^\Delta(x)|^{p+q}\Delta x, \tag{2.10}$$

where

$$\begin{aligned} K_3(\tau,b) \quad : \quad &= \sup_{\tau\le x\le b}\left(\mu^p(x)\frac{s(x)}{r(x)}\right) + 2^p\left(\frac{q}{p+q}\right)^{\frac{q}{p+q}} \\ &\times \left(\int_\tau^b \frac{(s(x))^{\frac{p+q}{p}}}{(r(x))^{\frac{q}{p}}}\left(\int_x^b r^{\frac{-1}{p+q-1}}(t)\Delta t\right)^{p+q-1}\Delta x\right)^{\frac{p}{p+q}}. \end{aligned} \tag{2.11}$$

We assume that there exists $\tau \in (a,b)$ which is the unique solution of the equation

$$K(p,q) = K_2(a,\tau) = K_3(\tau,b) < \infty, \tag{2.12}$$

where $K_2(a,\tau)$ and $K_3(\tau,b)$ are defined as in (2.9) and (2.11). Combining (2.8) and (2.10), we get

$$\int_a^b s(x)|y(x)+y^\sigma(x)|^p |y^\Delta(x)|^q \Delta x \le K(p,q)\int_a^b r(x)|y^\Delta(x)|^{p+q}\Delta x, \tag{2.13}$$

where $y : [a,b]_{\mathbb{T}} \to \mathbb{R}$ is delta differentiable with $y(a) = 0 = y(b)$, $\int_a^b r^{\frac{-1}{p+q-1}}(t)\Delta t < \infty$ and $K(p,q)$ is defined as in (2.12).

3. Main results

In this section, we prove some new Opial-type inequalities on a time scale $\mathbb{T}$ and apply these new inequalities on the second-order sublinear dynamic equation (1.1) to obtain some new Lyapunov-type inequalities related to problems $(i)-(ii)$. Throughout the rest of the paper, we will assume that the functions in the statements of the theorems are nonnegative and rd-continuous functions and the integrals considered are assumed to exist.

3.1. New Opial-type inequalities. Now, we will prove some new Opial type inequalities that will be needed in the proofs of our main results.

3.1. Theorem. *Let $\mathbb{T}$ be a time scale with $a, \tau \in \mathbb{T}$ and λ, δ be positive real numbers such that $\lambda \leq 1$, $\lambda + \delta > 1$, and let r, s be nonnegative rd-continuous functions on $(a,\tau)_{\mathbb{T}}$ such that $\int_a^\tau r^{\frac{-1}{\lambda+\delta-1}}(t)\Delta t < \infty$. If $y : [a,\tau]_{\mathbb{T}} \to \mathbb{R}$ is delta differentiable with $y(a) = 0$, then*

$$\int_a^\tau s(x)|y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x \leq H_1(a,\tau,\lambda,\delta)\int_a^\tau r(x)|y^\Delta(x)|^{\lambda+\delta}\Delta x, \tag{3.1}$$

where

$$\begin{aligned} H_1(a,\tau,\lambda,\delta) \quad : \quad &= \left(\frac{\delta}{\lambda+\delta}\right)^{\frac{\delta}{\lambda+\delta}} \left(\int_a^\tau \frac{(s(x))^{\frac{\lambda+\delta}{\lambda}}}{(r(x))^{\frac{\delta}{\lambda}}} \left(\int_a^x r^{\frac{-1}{\lambda+\delta-1}}(t)\Delta t\right)^{\lambda+\delta-1} \Delta x\right)^{\frac{\lambda}{\lambda+\delta}} \\ &+ \sup_{a \leq x \leq \tau}\left(\mu^\lambda(x)\frac{s(x)}{r(x)}\right), \end{aligned} \tag{3.2}$$

Proof. Since r is nonnegative on $(a,\tau)_{\mathbb{T}}$, it follows from the Hölder inequality with $f(t) = \frac{1}{(r(t))^{\frac{1}{\lambda+\delta}}}$, $g(t) = (r(t))^{\frac{1}{\lambda+\delta}}|y^\Delta(t)|$, $\gamma = \frac{\lambda+\delta}{\lambda+\delta-1}$ and $\beta = \lambda+\delta$, that

$$\begin{aligned} |y(x)| \quad &\leq \quad \int_a^x |y^\Delta(t)|\Delta t = \int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta}}}(r(t))^{\frac{1}{\lambda+\delta}}|y^\Delta(t)|\Delta t \\ &\leq \quad \left(\int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta-1}}}\Delta t\right)^{\frac{\lambda+\delta-1}{\lambda+\delta}} \left(\int_a^x r(t)|y^\Delta(t)|^{\lambda+\delta}\Delta t\right)^{\frac{1}{\lambda+\delta}}. \end{aligned}$$

Then, for $a \leq x \leq \tau$, we can write

$$|y(x)|^\lambda \leq \left(\int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta-1}}}\Delta t\right)^{\lambda(\frac{\lambda+\delta-1}{\lambda+\delta})} \left(\int_a^x r(t)|y^\Delta(t)|^{\lambda+\delta}\Delta t\right)^{\frac{\lambda}{\lambda+\delta}}. \tag{3.3}$$

Now, since $y^\sigma = y + \mu y^\Delta$, by applying inequality (see [18], page 500)

$$2^{r-1}\left|a^r + b^r\right| \leq |a+b|^r \leq \left|a^r + b^r\right|, \quad \text{for} \quad 0 \leq r \leq 1, \tag{3.4}$$

we have that

$$|y^\sigma|^\lambda = |y + \mu y^\Delta|^\lambda \leq |y|^\lambda + \mu^\lambda |y^\Delta|^\lambda. \tag{3.5}$$

Setting

$$z(x) := \int_a^x r(t)|y^\Delta(t)|^{\lambda+\delta}\Delta t, \tag{3.6}$$

we see that $z(a) = 0$, and

$$z^\Delta(x) = r(x)|y^\Delta(x)|^{\lambda+\delta} > 0. \tag{3.7}$$

This gives that

$$|y^{\Delta}(x)|^{\lambda+\delta}=\frac{z^{\Delta}(x)}{r(x)} \quad \text{and} \quad |y^{\Delta}(x)|^{\delta}=\left(\frac{z^{\Delta}(x)}{r(x)}\right)^{\frac{\delta}{\lambda+\delta}}. \tag{3.8}$$

Thus since s is nonnegative on $(a,\tau)_{\mathbb{T}}$, we get from (3.3), (3.5) and (3.8) that

$$\begin{aligned}
s(x)|y^{\sigma}(x)|^{\lambda}|y^{\Delta}(x)|^{\delta} &\leq s(x)|y|^{\lambda}|y^{\Delta}(x)|^{\delta}+s(x)\mu^{\lambda}|y^{\Delta}|^{\lambda+\delta} \\
&\leq s(x)\left(\frac{1}{r(x)}\right)^{\frac{\delta}{\lambda+\delta}} \times \left(\int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta-1}}}\Delta t\right)^{\lambda\left(\frac{\lambda+\delta-1}{\lambda+\delta}\right)} \\
&\quad \times (z(x))^{\frac{\lambda}{\lambda+\delta}}(z^{\Delta}(x))^{\frac{\delta}{\lambda+\delta}}+s(x)\mu^{\lambda}(x)\left(\frac{z^{\Delta}(x)}{r(x)}\right).
\end{aligned}$$

This implies that

$$\begin{aligned}
&\int_a^{\tau} s(x)|y^{\sigma}(x)|^{\lambda}|y^{\Delta}(x)|^{\delta}\Delta x \\
\leq\ &\int_a^{\tau} s(x)\left(\frac{1}{r(x)}\right)^{\frac{\delta}{\lambda+\delta}} \times \left(\int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta-1}}}\Delta t\right)^{\lambda\left(\frac{\lambda+\delta-1}{\lambda+\delta}\right)} \\
&\times (z(x))^{\frac{\lambda}{\lambda+\delta}}(z^{\Delta}(x))^{\frac{\delta}{\lambda+\delta}}\Delta x+\int_a^{\tau}\left(\mu^{\lambda}\frac{s(x)}{r(x)}\right)z^{\Delta}(x)\Delta(x).
\end{aligned}$$

$$\begin{aligned}
\leq\ &\int_a^{\tau} s(x)\left(\frac{1}{r(x)}\right)^{\frac{\delta}{\lambda+\delta}} \times \left(\int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta-1}}}\Delta t\right)^{\lambda\left(\frac{\lambda+\delta-1}{\lambda+\delta}\right)} \\
&\times (z(x))^{\frac{\lambda}{\lambda+\delta}}(z^{\Delta}(x))^{\frac{\delta}{\lambda+\delta}}\Delta x+\sup_{a\leq x\leq\tau}\left(\mu^{\lambda}\frac{s(x)}{r(x)}\right)\int_a^{\tau} z^{\Delta}(x)\Delta(x).
\end{aligned} \tag{3.9}$$

Applying the Hölder inequality (2.3) with indices $(\lambda+\delta)/\lambda$ and $(\lambda+\delta)/\delta$, we have

$$\begin{aligned}
&\int_a^{\tau} s(x)|y^{\sigma}(x)|^{\lambda}|y^{\Delta}(x)|^{\delta}\Delta x \\
\leq\ &\left(\int_a^{\tau} s(x)^{\frac{\lambda+\delta}{\lambda}}\left(\frac{1}{r(x)}\right)^{\frac{\delta}{\lambda}}\left(\int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta-1}}}\Delta t\right)^{(\lambda+\delta-1)}\Delta x\right)^{\frac{\lambda}{\lambda+\delta}} \\
&\times\left(\int_a^{\tau} z^{\frac{\lambda}{\delta}}(x)z^{\Delta}(x)\Delta x\right)^{\frac{\delta}{\lambda+\delta}}+\sup_{a\leq x\leq\tau}\left(\mu^{\lambda}\frac{s(x)}{r(x)}\right)\int_a^{\tau} z^{\Delta}(x)\Delta(x).
\end{aligned} \tag{3.10}$$

From (3.7), and the chain rule (2.1), we get that

$$z^{\frac{\lambda}{\delta}}(x)z^{\Delta}(x)\leq\frac{\delta}{\lambda+\delta}\left(z^{\frac{\lambda+\delta}{\delta}}(x)\right)^{\Delta}. \tag{3.11}$$

Substituting (3.11) into (3.10) and using the fact that $z(a) = 0$, we obtain

$$\begin{aligned}
&\int_a^\tau s(x)|y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x \\
\leq\ & \left(\int_a^\tau s(x)^{\frac{\lambda+\delta}{\lambda}} \left(\frac{1}{r(x)}\right)^{\frac{\delta}{\lambda}} \left(\int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta-1}}} \Delta t\right)^{(\lambda+\delta-1)} \Delta x\right)^{\frac{\lambda}{\lambda+\delta}} \\
&\times \left(\frac{\delta}{\lambda+\delta}\right)^{\frac{\delta}{\lambda+\delta}} \left(\int_a^\tau \left(z^{\frac{\lambda+\delta}{\delta}}(x)\right)^\Delta \Delta t\right)^{\frac{\delta}{\lambda+\delta}} + \sup_{a\leq x\leq\tau} \left(\mu^\lambda \frac{s(x)}{r(x)}\right) \int_a^\tau z^\Delta(x)\Delta(x) \\
=\ & \left(\int_a^\tau s(x)^{\frac{\lambda+\delta}{\lambda}} \left(\frac{1}{r(x)}\right)^{\frac{\delta}{\lambda}} \left(\int_a^x \frac{1}{(r(t))^{\frac{1}{\lambda+\delta-1}}} \Delta t\right)^{(\lambda+\delta-1)} \Delta x\right)^{\frac{\lambda}{\lambda+\delta}} \\
&\times \left(\frac{\delta}{\lambda+\delta}\right)^{\frac{\delta}{\lambda+\delta}} z(\tau) + \sup_{a\leq x\leq\tau} \left(\mu^\lambda \frac{s(x)}{r(x)}\right) z(\tau).
\end{aligned}$$

Using (3.6), we have from the last inequality that

$$\int_a^\tau s(x)|y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x \leq H_1(a,\tau,\lambda,\delta) \int_a^\tau r(x)|y^\Delta(x)|^{\lambda+\delta}\Delta x,$$

which is the required inequality (3.1) with (3.2). This completes the proof. □

Next, we will just state the following theorem, since its proof is the same as that of Theorem 3.1, with $[a,\tau]_{\mathbb{T}}$ replaced by $[\tau,b]_{\mathbb{T}}$.

3.2. Theorem. *Let $\mathbb{T}$ be a time scale with $\tau, b \in \mathbb{T}$ and λ, δ be positive real numbers such that $\lambda \leq 1$, $\lambda+\delta > 1$, and let r, s be nonnegative rd-continuous functions on $(\tau,b)_{\mathbb{T}}$ such that $\int_\tau^b r^{\frac{-1}{\lambda+\delta-1}}(t)\Delta t < \infty$. If $y : [\tau,b]_{\mathbb{T}} \to \mathbb{R}$ is delta differentiable with $y(b) = 0$, then*

$$\int_\tau^b s(x)|y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x \leq H_2(\tau,b,\lambda,\delta) \int_\tau^b r(x)|y^\Delta(x)|^{\lambda+\delta}\Delta x, \tag{3.12}$$

where

$$\begin{aligned}
H_2(\tau,b,\lambda,\delta) \ &= \ \left(\frac{\delta}{\lambda+\delta}\right)^{\frac{\delta}{\lambda+\delta}} \left(\int_\tau^b \frac{(s(x))^{\frac{\lambda+\delta}{\lambda}}}{(r(x))^{\frac{\delta}{\lambda}}} \left(\int_x^b r^{\frac{-1}{\lambda+\delta-1}}(t)\Delta t\right)^{\lambda+\delta-1} \Delta x\right)^{\frac{\lambda}{\lambda+\delta}} \\
&\quad + \sup_{\tau\leq x\leq b} \left(\mu^\lambda(x) \frac{s(x)}{r(x)}\right). \qquad (3.13)
\end{aligned}$$

In the following, we assume that there exists $\tau \in (a,b)_{\mathbb{T}}$ which is the unique solution of the equation

$$H(a,b) = H_1(a,\tau,\lambda,\delta) = H_2(\tau,b,\lambda,\delta) < \infty, \tag{3.14}$$

where $H_1(a,\tau,\lambda,\delta)$ and $H_2(\tau,b,\lambda,\delta)$ are defined as in Theorems 3.1 and 3.2. Note that since

$$\begin{aligned}
\int_a^b s(x)|y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x \ &= \ \int_a^\tau s(x)|y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x \\
&\quad + \int_\tau^b s(x)|y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x,
\end{aligned}$$

then the proof of the following theorem is just a combination of Theorems 3.1 and 3.2 and so, we remove it.

3.3. Theorem. *Let $\mathbb{T}$ be a time scale with $a, b \in \mathbb{T}$ and λ, δ be positive real numbers such that $\lambda \le 1$, $\lambda + \delta > 1$, and let r, s be nonnegative rd-continuous functions on $(a, b)_{\mathbb{T}}$ such that $\int_a^b r^{\frac{-1}{\lambda+\delta-1}}(t)\Delta t < \infty$. If $y : [a, b]_{\mathbb{T}} \to \mathbb{R}$ is delta differentiable with $y(a) = 0 = y(b)$, then*

$$\int_a^b s(x)|y^\sigma(x)|^\lambda |y^\Delta(x)|^\delta \Delta x \le H(a, b) \int_a^b r(x)|y^\Delta(x)|^{\lambda+\delta}\Delta x, \tag{3.15}$$

where $H(a, b)$ is defined as in (3.14).

3.2. New Lyapunov Inequalities. Now, we are ready to prove the results related to problems $(i) - (ii)$. For simplicity, we set the following notations:

$$K_1(a, b, \beta) := \sup_{a \le t \le b}\left(\mu^\beta(t)\frac{Q_1(t)}{r(t)}\right) + 2^\beta \left(\frac{1}{\beta+1}\right)^{\frac{1}{\beta+1}} \left(\int_a^b \frac{|Q_1(t)|^{\frac{\beta+1}{\beta}}}{r^{\frac{1}{\beta}}(t)} R_1(t)\Delta t\right)^{\frac{\beta}{\beta+1}},$$

$$H_1(a, b, \beta) := \left(\frac{\beta}{\beta+1}\right)^{\frac{\beta}{\beta+1}} \left(\int_a^b \frac{(p(t))^{\beta+1}}{(r(t))^\beta} R_1(t)\Delta t\right)^{\frac{1}{1+\beta}} + \sup_{a \le t \le b}\left(\mu(t)\frac{p(t)}{r(t)}\right),$$

where $Q_1(t) = \int_t^b q(s)\Delta s$ and $R_1(t) = \left(\int_a^t r^{-\frac{1}{\beta}}(\theta)\Delta\theta\right)^\beta$,

$$K_2(a, b, \beta) := \sup_{a \le t \le b}\left(\mu^\beta(t)\frac{Q_2(t)}{r(t)}\right) + 2^\beta \left(\frac{1}{\beta+1}\right)^{\frac{1}{\beta+1}} \left(\int_a^b \frac{|Q_2(t)|^{\frac{\beta+1}{\beta}}}{r^{\frac{1}{\beta}}(t)} R_2(t)\Delta t\right)^{\frac{\beta}{\beta+1}},$$

and

$$H_2(a, b, \beta) := \left(\frac{\beta}{\beta+1}\right)^{\frac{\beta}{\beta+1}} \left(\int_a^b \frac{(p(t))^{\beta+1}}{(r(t))^\beta} R_2(t)\Delta t\right)^{\frac{1}{1+\beta}} + \sup_{a \le t \le b}\left(\mu(t)\frac{p(t)}{r(t)}\right),$$

where $Q_2(t) = \int_a^t q(s)\Delta s$ and $R_2(t) = \left(\int_t^b r^{-\frac{1}{\beta}}(\theta)\Delta\theta\right)^\beta$.

3.4. Theorem. *Assume that y is a nontrivial solution of (1.1). If $y(a) = y^\Delta(b) = 0$, then*

$$2^{1-\beta}K_1(a, b, \beta) + H_1(a, b, \beta) \ge 1. \tag{3.16}$$

Proof. Without loss of generality, we may assume that $y(t) > 0$ in $[a, b]_{\mathbb{T}}$. Multiplying (1.1) by y^σ and integrating by parts (see 2.2), we get

$$\begin{aligned}
&\int_a^b \left(r(t)\left(y^\Delta(t)\right)^\beta\right)^\Delta y^\sigma(t)\Delta t + \int_a^b p(t)\left(y^\Delta(t)\right)^\beta y^\sigma(t)\Delta t \\
= \;& r(t)\left(y^\Delta(t)\right)^\beta y(t)\Big|_a^b - \int_a^b r(t)\left(y^\Delta(t)\right)^{\beta+1}\Delta t + \int_a^b p(t)\left(y^\Delta(t)\right)^\beta y^\sigma(t)\Delta t \\
= \;& -\int_a^b q(t)\left(y^\sigma(t)\right)^{\beta+1}\Delta t.
\end{aligned}$$

Using the assumptions that $y(a)=y^{\Delta}(b)=0$ and $Q(t)=\int_t^b q(s)\Delta s$, we have

$$\int_a^b r(t)\left(y^{\Delta}(t)\right)^{\beta+1}\Delta t=\int_a^b p(t)\left(y^{\Delta}(t)\right)^{\beta} y^{\sigma}(t)\,\Delta t-\int_a^b Q^{\Delta}(t)\left(y^{\sigma}(t)\right)^{\beta+1}\Delta t. \tag{3.17}$$

Integrating by parts the term $\int_a^b Q^{\Delta}(t)\left(y^{\sigma}(t)\right)^{\beta+1}\Delta t$, and using the facts that $y(a)=0=Q(b)$, we obtain

$$\int_a^b r(t)\left(y^{\Delta}(t)\right)^{\beta+1}\Delta t=\int_a^b p(t)\left(y^{\Delta}(t)\right)^{\beta} y^{\sigma}(t)\,\Delta t+\int_a^b Q(t)\left(y^{\beta+1}(t)\right)^{\Delta}\Delta t. \tag{3.18}$$

Applying the chain rule formula (2.1) and the inequality (3.4), we see that

$$\begin{aligned}\left|\left(y^{\beta+1}(t)\right)^{\Delta}\right| &\leq (\beta+1)\int_0^1 \left|hy^{\sigma}(t)+(1-h)\,y(t)\right|^{\beta} dh\left|y^{\Delta}(t)\right| \\ &\leq 2^{1-\beta}\left|y^{\sigma}(t)+y(t)\right|^{\beta}\left|y^{\Delta}(t)\right|. \end{aligned}\tag{3.19}$$

This and (3.18) imply that

$$\begin{aligned}\int_a^b r(t)\left|y^{\Delta}(t)\right|^{\beta+1}\Delta t &\leq \int_a^b |p(t)|\left|y^{\sigma}(t)\right|\left|y^{\Delta}(t)\right|^{\beta}\Delta t \\ &\quad +2^{1-\beta}\int_a^b |Q(t)|\left|y^{\sigma}(t)+y(t)\right|^{\beta}\left|y^{\Delta}(t)\right|\Delta t. \end{aligned}\tag{3.20}$$

Applying the inequality (2.7) on the integral $\int_a^b |Q(t)|\left|y^{\sigma}(t)+y(t)\right|^{\beta}\left|y^{\Delta}(t)\right|\Delta t$, with $s=Q,\ p=\beta$, and $q=1$, we have

$$\int_a^b |Q(t)|\left|y^{\sigma}(t)+y(t)\right|^{\beta}\left|y^{\Delta}(t)\right|\Delta t\leq K_1(a,b,\beta)\int_a^b r(t)\left|y^{\Delta}(t)\right|^{\beta+1}\Delta t. \tag{3.21}$$

Applying the inequality (3.1) on the integral $\int_a^b |p(t)|\left|y^{\sigma}(t)\right|\left|y^{\Delta}(t)\right|^{\beta}\Delta t$ with $s=p,\ \lambda=1$, and $\delta=\beta$, we obtain

$$\int_a^b p(t)|y^{\sigma}(t)||y^{\Delta}(t)|^{\beta}\Delta t\leq H_1(a,b,\beta)\int_a^b r(t)|y^{\Delta}(t)|^{\beta+1}\Delta t. \tag{3.22}$$

Substituting (3.21) and (3.22) into (3.20), we get

$$\begin{aligned}\int_a^b r(t)\left|y^{\Delta}(t)\right|^{\beta+1}\Delta t &\leq 2^{1-\beta}K_1(a,b,\beta)\int_a^b r(t)\left|y^{\Delta}(t)\right|^{\beta+1}\Delta t \\ &\quad +H_1(a,b,\beta)\int_a^b r(t)|y^{\Delta}(t)|^{\beta+1}\Delta t. \end{aligned}\tag{3.23}$$

Then, we have from (3.23) after cancelling the term $\int_a^b r(t)\left|y^{\Delta}(t)\right|^{\beta+1}\Delta t$, the desired inequality (3.16). The proof is complete. □

3.5. Remark. Theorem 3.4 gives us a condition for right disfocality of (1.1). In particular, if

$$2^{1-\beta}K_1(a,b,\beta)+H_1(a,b,\beta)<1,$$

then (1.1) is right disfocal in $[a,b]_{\mathbb{T}}$. This means that there is no nontrivial solution of (1.1) in $[a,b]_{\mathbb{T}}$ satisfies $y(a)=y^{\Delta}(b)=0$.

3.6. Theorem. *Assume that y is a nontrivial solution of (1.1). If $y^{\Delta}(a)=y(b)=0$, then*

$$2^{1-\beta}K_2(a,b,\beta)+H_2(a,b,\beta)\geq 1. \tag{3.24}$$

Proof. The proof of (3.24) is similar to (3.16) by employing Opial-type inequalities (2.8) and (3.12) instead of (2.7) and (3.1). The proof is complete. □

3.7. Remark. Theorem 3.6 gives us a condition for left disfocality of (1.1). In particular, if

$$2^{1-\beta}K_2(a,b,\beta)+H_2(a,b,\beta)<1.$$

then (1.1) is left disfocal in $[a,b]_{\mathbb{T}}$. This means that there is no nontrivial solution of (1.1) in $[a,b]_{\mathbb{T}}$ satisfies $y^{\Delta}(a)=y(b)=0$.

In the following, we employ inequalities (2.13) and (3.15) to determine the lower bound for the distance between consecutive zeros of a solution of (1.1).

3.8. Theorem. *Assume that $Q^{\Delta}(t)=q(t)$ and y is a nontrivial solution of (1.1). If $y(a)=y(b)=0$, then*

$$2^{1-\beta}K(a,b)+H(a,b)\geq 1, \tag{3.25}$$

where $K(a,b)$ and $H(a,b)$ are defined as in (2.12) and (3.14), respectively.

Proof. Multiplying (1.1) by y^{σ} and integrating by parts, we get that

$$\int_a^b r(t)\left(y^{\Delta}(t)\right)^{\beta+1}\Delta t=\int_a^b p(t)\left(y^{\Delta}(t)\right)^{\beta}y^{\sigma}(t)\,\Delta t-\int_a^b Q^{\Delta}(t)\left(y^{\sigma}(t)\right)^{\beta+1}\Delta t. \tag{3.26}$$

Using the facts that $y(a)=0=y(b)$, we obtain

$$\begin{aligned}\int_a^b r(t)\left(y^{\Delta}(t)\right)^{\beta+1}\Delta t \quad &\leq \quad \int_a^b p(t)\left(y^{\Delta}(t)\right)^{\beta}y^{\sigma}(t)\,\Delta t\\ &\quad +2^{1-\beta}\int_a^b |Q(t)|\left|y^{\sigma}(t)+y(t)\right|^{\beta}\left|y^{\Delta}(t)\right|\Delta t.\end{aligned} \tag{3.27}$$

Applying the inequality (2.13) on the integral $\int_a^b |Q(t)|\left|y^{\sigma}(t)+y(t)\right|^{\beta}\left|y^{\Delta}(t)\right|\Delta t$, with $s=|Q|,\ \lambda=\beta,\ \delta=1$, we have that

$$\int_a^b |Q(t)|\left|y^{\sigma}(t)+y(t)\right|^{\beta}\left|y^{\Delta}(t)\right|\Delta t\leq K(a,b)\int_a^b r(t)\left|y^{\Delta}(t)\right|^{\beta+1}\Delta t, \tag{3.28}$$

where $K(a,b)$ is defined as in (2.12). Applying the inequality (3.15) on the integral $\int_a^b |p(t)|\,|y^\sigma(t)|\,\left|y^\Delta(t)\right|^\beta \Delta t$ with $s=p,\ \lambda=1,\ \delta=\beta$, we have that

$$\int_a^b p(t)|y^\sigma(t)||y^\Delta(t)|^\beta \Delta t \le H(a,b)\int_a^b r(t)|y^\Delta(t)|^{\beta+1}\Delta t, \tag{3.29}$$

where $H(a,b)$ is defined as in (3.14). Substituting (3.28) and (3.29) into (3.27), we get that

(3.30)

$$\int_a^b r(t)\left|y^\Delta(t)\right|^{\beta+1}\Delta t \le 2^{1-\beta}K(a,b)\int_a^b r(t)\left|y^\Delta(t)\right|^{\beta+1}\Delta t + H(a,b)\int_a^b r(t)|y^\Delta(t)|^{\beta+1}\Delta t.$$

Then, we have from (3.30) after cancelling the term $\int_a^b r(t)\left|y^\Delta(t)\right|^{\beta+1}\Delta t$, that

$$2^{1-\beta}K(a,b)+H(a,b)\ge 1,$$

which is the desired inequality (3.25). The proof is complete. □

3.9. Remark. Theorem 3.8 gives us a condition for disconjugacy of (1.1). In particular, if

$$K(a,b)+H(a,b)<1,$$

then (1.1) is disconjugate in $[a,b]_{\mathbb{T}}$. This means that there is no nontrivial solution of (1.1) in $[a,b]_{\mathbb{T}}$ satisfies $y(a)=y(b)=0$.

4. Applications

In Theorem 3.4 if $\beta=1$, then we have the following result, which improves the obtained result in [23, Corollary 2.2] by removing the additional constant c in the conditions.

4.1. Corollary. *Assume that y is a nontrivial solution of (1.1). If $y(a)=y^\Delta(b)=0$, then*

$$\sup_{a\le t\le b}\frac{1}{r(t)}\left[p(t)\mu(t)+Q_1(t)\mu(t)\right]+\sqrt{2}\left(\int_a^b \frac{|Q_1(t)|^2}{r(t)}R_1(t)\Delta t\right)^{\frac{1}{2}} + \frac{1}{\sqrt{2}}\left(\int_a^b \frac{p^2(t)}{r(t)}R_1(t)\Delta t\right)^{\frac{1}{2}}\ge 1, \tag{4.1}$$

where $Q_1(t)=\int_t^b q(s)\Delta s$ and $R_1(t)=\int_a^t \frac{\Delta\tau}{r(\tau)}$. If $y^\Delta(a)=y(b)=0$, then

$$\sup_{a\le t\le b}\frac{1}{r(t)}\left[p(t)\mu(t)+Q_2(t)\mu(t)\right]+\sqrt{2}\left(\int_a^b \frac{|Q_2(t)|^2}{r(t)}R_2(t)\Delta t\right)^{\frac{1}{2}} + \frac{1}{\sqrt{2}}\left(\int_a^b \frac{p^2(t)}{r(t)}R_2(t)\Delta t\right)^{\frac{1}{2}}\ge 1, \tag{4.2}$$

where $Q_2(t)=\int_a^t q(s)\Delta s$ and $R_2(t)=\int_t^b \frac{\Delta\tau}{r(\tau)}$.

As a special case of Theorem 3.4, when $r(t) = 1$, we obtain the following result, which improves the result that is obtained in [23, Corollary 2.1] by removing the additional constant c in the obtained results.

4.2. Corollary. *Assume that y is a nontrivial solution of (1.1). If $y(a) = y^{\Delta}(b) = 0$, then*

$$\sup_{a\le t\le b}\Big[\mu(t)p(t)+2^{1-\beta}\mu^{\beta}(t)Q(t)\Big]+\frac{2}{(\beta+1)^{\frac{1}{\beta+1}}}\left(\int_a^b |Q(t)|^{\frac{\beta+1}{\beta}}(\tau-a)^{\beta}\,\Delta t\right)^{\frac{\beta}{\beta+1}}$$
$$+\left(\frac{\beta}{\beta+1}\right)^{\frac{\beta}{\beta+1}}\left(\int_a^b |p(t)|^{\beta+1}(\tau-a)^{\beta}\,\Delta t\right)^{\frac{1}{1+\beta}}\ge 1,$$

where $Q(t) = \int_t^b q(s)\Delta s$. If $y^{\Delta}(a) = y(b) = 0$, then

$$\sup_{a\le t\le b}\Big[\mu(t)p(t)+2^{1-\beta}\mu^{\beta}(t)Q(t)\Big]+\frac{2}{(\beta+1)^{\frac{1}{\beta+1}}}\left(\int_a^b |Q(t)|^{\frac{\beta+1}{\beta}}(b-\tau)^{\beta}\,\Delta t\right)^{\frac{\beta}{\beta+1}}$$
$$+\left(\frac{\beta}{\beta+1}\right)^{\frac{\beta}{\beta+1}}\left(\int_a^b |p(t)|^{\beta+1}(b-\tau)^{\beta}\,\Delta t\right)^{\frac{1}{1+\beta}}\ge 1,$$

where $Q(t) = \int_a^t q(s)\Delta s$.

As a special case of Corollary 4.1, when $p(t) = 0$, we have the following results.

4.3. Corollary. *Assume that y is a nontrivial solution of (1.16). If $y(a) = y^{\Delta}(b) = 0$, then*

$$\sup_{a\le t\le b}\frac{1}{r(t)}Q_1(t)\mu(t)+\sqrt{2}\left(\int_a^b \frac{|Q_1(t)|^2}{r(t)}R_1(t)\Delta t\right)^{\frac12}\ge 1, \tag{4.3}$$

where $Q_1(t) = \int_t^b q(s)\Delta s$ and $R_1(t) = \int_a^t \frac{\Delta\tau}{r(\tau)}$. If instead $y^{\Delta}(a) = y(b) = 0$, then

$$\sup_{a\le t\le b}\frac{1}{r(t)}Q_1(t)\mu(t)+\sqrt{2}\left(\int_a^b \frac{|Q_1(t)|^2}{r(t)}R_2(t)\Delta t\right)^{\frac12}\ge 1, \tag{4.4}$$

where $Q_2(t) = \int_a^t q(s)\Delta s$ and $R_2(t) = \int_t^b \frac{\Delta\tau}{r(\tau)}$.

Using the maximum of $|Q_1(t)|$ on $[a,b]_{\mathbb{T}}$ in Corollary 4.3, we get the following results.

4.4. Corollary. *Assume that y is a nontrivial solution of (1.16). If $y(a) = y^{\Delta}(b) = 0$, then*

$$\sup_{a\le t\le b}\frac{1}{r(t)}\left|\int_t^b q(s)\Delta s\right|\mu(t)+\sqrt{2}\max_{a\le t\le b}\left|\int_t^b q(s)\Delta s\right|\left(\int_a^b \frac{R_1(t)}{r(t)}\Delta t\right)^{\frac12}\ge 1, \tag{4.5}$$

where $R_1(t) = \int_a^t \frac{\Delta\tau}{r(\tau)}$. If instead $y^{\Delta}(a) = y(b) = 0$, then

$$\sup_{a\le t\le b}\frac{1}{r(t)}\left|\int_a^t q(s)\Delta s\right|\mu(t)+\sqrt{2}\max_{a\le t\le b}\left|\int_a^t q(s)\Delta s\right|\left(\int_a^b \frac{R_2(t)}{r(t)}\Delta t\right)^{\frac12}\ge 1, \tag{4.6}$$

where $R_2(t) = \int_t^b \frac{\Delta\tau}{r(\tau)}$.

As a special case when $\mathbb{T} = \mathbb{R}$, $\beta = 1$, $r(t) = 1$ and $p(t) = 0$, then $y^\sigma(t) = y(t)$ and equation (1.1) becomes

$$y^{''}(t) + q(t)y(t) = 0. \tag{4.7}$$

Now, the results in Corollary 4.3 reduce to the following results obtained by Brown and Hinton [7].

4.5. Corollary. *Assume that* y *is a solution of the equation (4.7). If* $y(a) = y^{'}(b) = 0$, *then*

$$2\int_a^b Q_1^2(t)(t-a)dt > 1, \tag{4.8}$$

where $Q_1(t) = \int_t^b q(s)ds$. *If instead* $y^{'}(\alpha) = y(\beta) = 0$, *then*

$$2\int_a^b Q_2^2(t)(b-t)dt > 1, \tag{4.9}$$

where $Q_2(t) = \int_a^t q(s)ds$.

As a special case of Corollary 4.4 for the second order differential equation (4.7), we get the following results due to Harris and Kong [10].

4.6. Corollary. *Assume that* y *is a solution of the equation (4.7). If* $y(a) = y^{'}(b) = 0$, *then*

$$(b-a)\sup_{a\le t\le b}\left|\int_t^b q(s)\Delta s\right| > 1. \tag{4.10}$$

If instead $y^{'}(\alpha) = y(\beta) = 0$, *then*

$$(b-a)\sup_{a\le t\le b}\left|\int_a^t q(s)\Delta s\right| > 1. \tag{4.11}$$

Using the maximum of $|Q|$ and $|p|$ on $[a,b]_{\mathbb{T}}$ we have from Corollary 4.2 the following results for the second order difference equation

$$\Delta((\Delta y(t))^\beta + p(t)\,(\Delta y(t))^\beta + q(t)y^\beta(t+1)) = 0, \tag{4.12}$$

where $0 < \beta \le 1$ is a quotient of odd positive integers.

4.7. Corollary. *Assume that* y *is a nontrivial solution of (4.12). If* $y(a) = \Delta y(b) = 0$, *then*

$$\left[\max_{a\le\tau\le b}|p(t)| + 2^{1-\beta}\max_{a\le\tau\le b}\left|\sum_{s=t}^{b-1} q(s)\Delta s\right|\right] + \frac{2(b-a)^\beta}{(\beta+1)}\max_{a\le\tau\le b}\left|\sum_{s=t}^{b-1} q(s)\Delta s\right|$$
$$+\frac{\beta^{\frac{\beta}{\beta+1}}}{\beta+1}(b-a)\max_{a\le\tau\le b}|p(t)| \ge 1.$$

If $\Delta y(a) = y(b) = 0$, *then*

$$\left[\max_{a\le\tau\le b}|p(t)| + 2^{1-\beta}\max_{a\le\tau\le b}\left|\sum_{s=a}^{t-1}q(s)\Delta s\right|\right] + \frac{2(b-a)^{\beta}}{(\beta+1)}\max_{a\le\tau\le b}\left|\sum_{s=a}^{t-1}q(s)\Delta s\right|$$
$$+\frac{\beta^{\frac{\beta}{\beta+1}}}{\beta+1}(b-a)\max_{a\le\tau\le b}|p(t)| \ge 1.$$

References

[1] Agarwal, R.P. and Pang, P.Y.H. *Opial inequalities with Applications in Differential and Difference Equations*, Kluwer, Dordrechet (1995).

[2] Bohner, M. and Peterson, A. *Dynamic Equations on Time Scales: An Introduction with Applications*, Birkhäuser, Boston, Mass, USA (2001).

[3] Bohner, M. and Peterson, A., *Advances in Dynamic Equations on Time Scales*, Birkhäuser, Boston (2003).

[4] Bohner, M. and Kaymakçalan, B. *Opial Inequalities on time scales*, Annal. Polon. Math. **77** (2001), 11-20.

[5] Bohner, M., Clark, S. and Ridenhour, J. *Lyapunov inequalities for time scales*, J. Ineq. Appl. **7** (2002), 61-77.

[6] Brown, R.C. and Hinton, D.B. *Lyapunov inequalities and their applications*, in: T. M. Rassias (Ed.), *Survey on Classical Inequalities*, Kluwer Academic Publishers, Dordrecht, The Netherlands (2000), 1-25.

[7] Brown, R.C. and Hinton, D.B. *Opial's inequality and oscillation of* 2^{nd} *order equations*, Proc. Amer. Math. Soc. **125** (1997), 1123-1129.

[8] Cheng, S. *Lyapunov inequalities for differential and difference equations*, Fasc. Math. **23** (1991) 25–41.

[9] Cohn, J.H.E. *Consecutive zeros of solutions of ordinary second order differential equations*, J. Lond. Math. Soc. **2** (5) (1972), 465-468.

[10] Harris, B. J. and Kong, Q. *On the oscillation of differential equations with an oscillatory coefficient*, Transc. Amer. Math. Soc. **347** (5) (1995), 1831-1839.

[11] Hartman, P. *Ordinary Differential Equations*, Wiley, New York, (1964) and Birkhäuser, Boston (1982).

[12] Hartman, P. and Wintner, A. *On an oscillation criterion of de la Vallée Poussin*, Quart. Appl. Math. **13** (1955), 330-332.

[13] Karpuz, B., Kaymakçalan, B. and Öclan, Ö. *A generalization of Opial's inequality and applications to second order dynamic equations*, Diff. Eqns. Dyn. Sys. **18** (2010), 11-18.

[14] M. K. Kwong, *On Lyapunov's inequality for disfocality*, J. Math. Anal. Appl. **83** (1981), 486-494.

[15] Lasota, A. *A discrete boundary value problem*, Annal. Polon. Math. **20** (1968), 183-190.

[16] Lettenmeyer, F. *Ueber die von einem punktausgehenden Integralkurven einer Differentialgleichung 2.* Ordnung, Deutsche Math. **7** (1944), 56-74.

[17] Lyapunov, A.M. *Probleme général de la stabilité du mouvement*, Ann. Math. Stud. **17** (1947), 203-474.

[18] Mitrinović, D.S., Pečarić, J.E. and Fink, A.M. *Classical and New Inequalities in Analysis*, **vol. 61**, Kluwer Academic Publishers, Dordrecht, The Netherlands, (1993).

[19] Nehari, Z. *On the zeros of solutions of second order linear differential equations*, Amer. J. Math. **76** (1954), 689-697.

[20] Opial, Z. *Sur une inegalite*, Annal. Polon. Math. **8** (1960), 29-32.

[21] Opial, Z. *Sur une inégalité de C. de La Vallée Poussin dans la théorie de l'équation différentielle linéaire du second ordre*, Annal. Polon. Math. **6** (1959), 87-91.

[22] Poussin, C. de la Vallée *Sur l'équation différentielle linéaire du second ordre*, J. Math. Pures Appl. **8** (1929), 125-144.

[23] Saker, S.H. *Applications of Opial inequalities on time scales on dynamic equations with damping terms*, Math. Comp. Model. **58** (2013), 1777-1790.

[24] Saker, S.H. *Opial's type inequalities on time scales and some applications*, Annal. Polon. Math. **104** (2012), 243-260.

[25] Saker, S.H. *New inequalities of Opial's type on time scales and some of their applications*, Disc. Dynam. Nat. Soc. **2012** (2012), 1-23.

[26] Saker, S.H. *Some New Inequalities of Opial's Type on Time Scales*, Abs. Appl. Anal. **2012** (2012), 1-14.

[27] Saker, S.H. *Lyapunov inequalities for half-linear dynamic equations on time scales and disconjugacy*, Dyn. Contin. Discrete Impuls. Syst. Ser. B, Appl. Algorithms **18** (2011), 149-161.

[28] Saker, S.H. *Oscillation Theory of Dynamic Equations on Time Scales: Second and Third Orders*, Lambert Academic Publishing, Germany (2010).

[29] Saker, S.H. *Some new disconjugacy criteria for second order differential equations with a middle term*, Bull. Math. Soc. Sci. Math. Roum. **57** (1) (2014), 109-120.

[30] Szmanda, B. *The distance between the zeros of certain solutions of nth order linear differential equations* (Polish), Fasciculi Math. Nr. **4** (1969), 65-70.

[31] Tiryaki, A. *Recent development of Lyapunov-type inequalities*, Adv. Dyn. Syst. Appl. **5** (2) (2010) 231-248.

[32] Willett, D. *Generalized de la Vallée Poussin disconjugacy tests for linear differential equations*, Canad. Math. Bull. **14** (1971), 419-428.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 473 – 482

The Borel property for 4-dimensional matrices

Emre TAŞ*

Abstract

In 1909 Borel has proved that "Almost all of the sequences of 0's and 1's are Cesàro summable to $\frac{1}{2}$". Then Hill has generalized Borel's result to two dimensional matrices. In this paper we investigate the Borel property for 4-dimensional matrices.

Keywords: Double sequences, Pringsheim convergence, the Borel Property, double sequences of 0's and 1's.

2000 AMS Classification: 40B05, 40C05, 28A12.

Received : 25.12.2014 *Accepted* : 30.03.2015 *Doi* : 10.15672/HJMS.20164512504

1. Introduction

The summability of sequences of 0's and 1's has been studied by various authors ([1], [3], [6], [7], [8], [10]). In 1909 Borel proved that "Almost all of the sequences of 0's and 1's are Cesàro summable to $\frac{1}{2}$". Then Hill [6] has generalized Borel's result to general matrices. We say that the matrix has the Borel property, if a matrix sums almost all of the sequences of 0's and 1's to $\frac{1}{2}$. Establishing a one-to-one correspondence between the interval $(0, 1]$ and the collection of all sequences of 0's and 1's, Hill has given some necessary conditions and also some sufficient conditions for matrices to have the Borel property in [6], [7]. This property has also been examined in [5], [8].

In the present paper we investigate the Borel property for 4-dimensional matrices. In particular we exhibit some necessary and some sufficient conditions for 4-dimensional matrices to have the Borel property.

We first recall some basic notations and results related to double sequences.

A double sequence $s = (s_{ij})$ is said to be Pringsheim convergent (i.e., it is convergent in Pringsheim's sense) to L if for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|s_{ij} - L| < \varepsilon$ whenever $i, j \geq N$ ([2], [11]). In this case L is called the Pringsheim limit of s.

Throughout the paper when there is no confusion, convergence means the Pringsheim convergence.

*Ankara University Faculty of Science Department of Mathematics Tandoğan 06100 Ankara Turkey, Email: `emretas86@hotmail.com`

Let X denote the set of all double sequences of 0's and 1's, that is

$$X = \{x = (x_{jk}) : x_{jk} \in \{0,1\} \text{ for each } j,k \in \mathbb{N}\}.$$

Let $\Re$ be the smallest σ-algebra of subsets of the set X which contains all sets of the form

$$\{x = (x_{jk}) \in X : \ x_{j_1k_1} = a_1, ..., x_{j_nk_n} = a_n\}$$

where each $a_i \in \{0,1\}$ and the pairs $\{(j_ik_i)\}_{i=1}^n$ are pairwise distinct.

There exists a unique probability measure P on the set $\Re$, such that

$$P(\{x = (x_{jk}) \in X : \ x_{j_1k_1} = a_1, ..., x_{j_nk_n} = a_n\}) = \frac{1}{2^n}$$

for all choices of n and all pairwise disjoint pairs $\{(j_ik_i)\}_{i=1}^n$, and all choices of $a_1, ..., a_n$.

Recall that the functions $r_{jk}(x) = 2x_{jk} - 1$, for $x \in X$, are the Rademacher functions (see [4]).

Four dimensional Cesàro matrix $(C,1,1) = \left(c_{jk}^{nm}\right)$ is defined by

$$c_{jk}^{nm} = \begin{cases} \frac{1}{nm} & , \quad 1 \le j \le n \text{ and } 1 \le k \le m \\ 0 & , \quad \textit{otherwise.} \end{cases}$$

It is known that the $(C,1,1)$ matrix is an RH regular, i.e., it sums every bounded convergent sequence to the same limit.

An element x of X is said to be normal ([4]) if for each $\varepsilon > 0$ there is a natural number N_ε such that for $n,m \ge N_\varepsilon$ we have $\left| \frac{1}{nm} \sum\limits_{\substack{j \le n \\ k \le m}} x_{jk} - \frac{1}{2} \right| < \varepsilon$. Let η denote the set of all elements x in X that are normal. This means that normal elements are $(C,1,1)$-summable to $\frac{1}{2}$. It is also proved in [4] that $P(\eta) = 1$. So $(C,1,1)$ method has the Borel property.

It would be appropriate to recall the definition of bounded regularity.

1.1. Definition. Let $\mathcal{A} = \left(a_{jk}^{nm}\right)$ be a 4-dimensional matrix. If the limit

$$\lim_{n,m\to\infty} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} s_{jk} = L$$

exists, the double sequence (s_{jk}) is called $\mathcal{A}$-summable to L and denoted by $s_{jk} \to L$ $(\mathcal{A})$. A matrix $\mathcal{A} = \left(a_{jk}^{nm}\right)$ is bounded regular if every bounded and convergent sequence $s = (s_{jk})$ is $\mathcal{A}$-summable to the same limit and $\mathcal{A}$-means are also bounded [9]. The next corollary characterizes bounded regular matrices.

1.2. Proposition. *$\mathcal{A} = \left(a_{jk}^{nm}\right)$ is bounded regular if and only if*

(i) $\lim\limits_{n,m\to\infty} a_{jk}^{nm} = 0$, $(j,k = 1,2,...)$

(ii) $\lim\limits_{n,m\to\infty} \sum\limits_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} = 1$,

(iii) $\lim\limits_{n,m\to\infty} \sum\limits_{k=1}^{\infty} \left|a_{jk}^{nm}\right| = 0$, $(j = 1,2,...)$

(iv) $\lim\limits_{n,m\to\infty} \sum\limits_{j=1}^{\infty} \left|a_{jk}^{nm}\right| = 0$, $(k = 1,2,...)$

(v) $\sum\limits_{j,k=1,1}^{\infty,\infty} \left|a_{jk}^{nm}\right| \le C < \infty$, $(m,n = 1,2,...)$.

These conditions were first established by Robison [12].

2. The Borel Property

This section is devoted to the Borel property for 4-dimensional matrices.

2.1. Theorem. *If* $\mathcal{A} = \left(a_{jk}^{nm}\right)$ *has the Borel property, then the* $\sum\limits_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm}$ *series converges for each* n, m *and tends to* 1 *as* $n, m \to \infty$.

Proof. Since $\mathcal{A}$ has the Borel property, for almost all $x \in X$, we obtain $\lim\limits_{n,m\to\infty} \sum\limits_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} x_{jk} = \frac{1}{2}$. Indeed $P(E) = 1$ where

$$E = \left\{ x = (x_{jk}) \in X : \ (Ax)_{nm} \to \frac{1}{2} \right\}.$$

Let us define $\overline{x} = (\bar{x}_{jk})$ by

$$\bar{x}_{jk} = \begin{cases} 0 & , \quad x_{jk} = 1 \\ 1 & , \quad x_{jk} = 0 \end{cases} .$$

Let $Y = E \cap \eta$ and $\overline{Y} = \{(\overline{x}_{jk}) : \ x_{jk} \in Y\}$. We get $\overline{Y} = \overline{E} \cap \eta$. Since the mapping $(x_{jk}) \to (\overline{x}_{jk})$ preserves P measure, we obtain $P\left(\overline{Y}\right) = 1$. So $Y \cap \overline{Y} \neq \emptyset$. If $x = (x_{jk}) \in Y \cap \overline{Y}$, then $x \in E$, $x \in \eta$ and $\overline{x} \in E$, $\overline{x} \in \eta$. Since x, $\overline{x} \in E$, it follows that

$$\sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} x_{jk} + \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} \overline{x}_{jk} = \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} \to 1 \quad (n, m \to \infty).$$

This completes the proof. □

2.2. Theorem. *If* $\mathcal{A} = \left(a_{jk}^{nm}\right)$ *has the Borel property, then we have*

$$\sum_{j,k=1,1}^{\infty,\infty} \left(a_{jk}^{nm}\right)^2 < \infty$$

for each $n, m \in \mathbb{N}$.

Proof. Let $r_{jk}(x) = 2x_{jk} - 1$ be the Rademacher functions for double sequences. We have

$$\sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} x_{jk} = \frac{1}{2} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} + \frac{1}{2} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x).$$

Since $\mathcal{A}$ has the Borel property and it follows from Teorem 2.1 that the series $\sum\limits_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x)$ converges for each $n, m \in \mathbb{N}$ and almost all $x \in X$. Furthermore we obtain $\lim\limits_{n,m} \sum\limits_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x) = 0$ for almost all $x \in X$. So $\left(\sum\limits_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x) \right)$ is convergent uniformly on a set D with positive measure for each $n, m \in \mathbb{N}$ with respect to x. Hence for each $n, m \in \mathbb{N}$ and for every $\varepsilon > 0$, there exists $N_1, N_2 \in \mathbb{N}$ such that for $p, \mu > N_1$ and $q, \nu > N_2$

$$\left| \sum_{j,k=1,1}^{p,q} a_{jk}^{nm} r_{jk}(x) - \sum_{j,k=1,1}^{\mu,\nu} a_{jk}^{nm} r_{jk}(x) \right| < \varepsilon.$$

From the last inequality we immediately get

$$\tag{2.1} \begin{aligned} \varepsilon^2 P(D) &> \int_D \left(\sum_{E[\mu,p;\nu,q]} a_{jk}^{nm} r_{jk}(x) \right)^2 dP(x) \\ &= P(D) \sum_{E[\mu,p;\nu,q]} \left(a_{jk}^{nm}\right)^2 + R \end{aligned}$$

where

$$E[\mu, p; \nu, q] = \{(j,k) : \ \mu < j \le p \text{ or } \nu < k \le q\},$$

$$R = 2 \sum_{I[\mu,p;\nu,q]} a_{j_1k_1}^{nm} a_{j_2k_2}^{nm} \int_D r_{j_1k_1}(x) r_{j_2k_2}(x) \, dP(x)$$

and $I[\mu, p; \nu, q] = E[\mu, p; \nu, q] \cap \{(j,k) : \ j_1 \ne j_2 \text{ or } k_1 \ne k_2\}$. On the other hand using the Hölder inequality, we obtain

$$|R| \le 2 \left\{ \sum_{I[\mu,p;\nu,q]} \left(a_{j_1k_1}^{nm} a_{j_2k_2}^{nm}\right)^2 \right\}^{\frac{1}{2}} \left\{ \sum_{I[\mu,p;\nu,q]} \left(\int_D r_{j_1k_1}(x) r_{j_2k_2}(x) \, dP(x) \right)^2 \right\}^{\frac{1}{2}}.$$

Let $v_{j_1k_1j_2k_2}^2 = \left(\int_D r_{j_1k_1}(x) r_{j_2k_2}(x) \, dP(x) \right)^2$. From the Bessel inequality, we get

$$\sum_{\substack{1 \le j_1 < j_2 < \infty \\ 1 \le k_1 < k_2 < \infty}} v_{j_1k_1j_2k_2}^2 \le \int_X (\chi_D(x))^2 \, dP(x) = P(D).$$

For sufficiently large p, q, μ and ν, we have

$$\left\{ \sum_{I[\mu,p;\nu,q]} v_{j_1k_1j_2k_2}^2 \right\}^{\frac{1}{2}} \le \frac{P(D)}{4}.$$

Hence we obtain

$$\begin{aligned} |R| &\le \frac{P(D)}{2} \left\{ \sum_{I[\mu,p;\nu,q]} \left(a_{j_1k_1}^{nm} a_{j_2k_2}^{nm}\right)^2 \right\}^{\frac{1}{2}} \\ &\le \frac{P(D)}{2} \left\{ \sum_{E[\mu,p;\nu,q]} \left(a_{j_1k_1}^{nm} a_{j_2k_2}^{nm}\right)^2 \right\}^{\frac{1}{2}} \\ &\le \frac{P(D)}{2} \sum_{E[\mu,p;\nu,q]} \left(a_{j_1k_1}^{nm}\right)^2. \end{aligned}$$

From (2.1) and last inequality, it follows that

$$\begin{aligned}\varepsilon^2 P(D) &> P(D) \sum_{E[\mu,p;\nu,q]} \left(a_{jk}^{nm}\right)^2 - \frac{P(D)}{2} \sum_{E[\mu,p;\nu,q]} \left(a_{jk}^{nm}\right)^2 \\ &= \frac{P(D)}{2} \sum_{E[\mu,p;\nu,q]} \left(a_{jk}^{nm}\right)^2 .\end{aligned}$$

Also since $P(D) > 0$, we obtain $\sum\limits_{E[\mu,p;\nu,q]} \left(a_{jk}^{nm}\right)^2 < 2\varepsilon^2$. So for each $n, m \in \mathbb{N}$, the series $\left\{ \sum\limits_{j,k=1,1}^{\infty,\infty} \left(a_{jk}^{nm}\right)^2 \right\}$ is convergent. Hence we obtain the result. □

2.3. Theorem. *If $\mathcal{A} = \left(a_{jk}^{nm}\right)$ has the Borel property and satisfies (v), we have*

$$\sum_{j,k=1,1}^{\infty,\infty} \left(a_{jk}^{nm}\right)^2 = o(1), \qquad (n, m \to \infty). \tag{2.2}$$

Proof. Let $\sigma_{nm}(x) = \sum\limits_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x)$. Using the equality

$$\sigma_{nm}^2(x) = \left(\sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x) \right) \left(\sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x) \right)$$

and (v), we can easily obtain

$$\left|\sigma_{nm}^2(x)\right| \le \sum_{j,k=1,1}^{\infty,\infty} \left|a_{jk}^{nm}\right| \sum_{j,k=1,1}^{\infty,\infty} \left|a_{jk}^{nm}\right| < \infty$$

and hence

$$\sigma_{nm}^2(x) = \sum_{\substack{1 \le j_1, j_2 \le \infty \\ 1 \le k_1, k_2 \le \infty}} a_{j_1k_1}^{nm} a_{j_2k_2}^{nm} r_{j_1k_1}(x) r_{j_2k_2}(x)$$

is convergent uniformly almost everywhere. So we have

$$\begin{aligned}\int_X \sigma_{nm}^2(x)\, dP(x) &= \sum_{\substack{1 \le j_1, j_2 \le \infty \\ 1 \le k_1, k_2 \le \infty}} a_{j_1k_1}^{nm} a_{j_2k_2}^{nm} \int_X r_{j_1k_1}(x) r_{j_2k_2}(x)\, dP(x) \\ &= \sum_{j,k=1,1}^{\infty,\infty} \left(a_{jk}^{nm}\right)^2 .\end{aligned} \tag{2.3}$$

Since $\mathcal{A}$ has the Borel property, the uniformly bounded sequence $(\sigma_{nm}(x))$ converges to 0 for almost all x. From (2.3) and the Lebesgue convergence theorem, it follows that $\lim\limits_{n,m\to\infty} \sum\limits_{j,k=1,1}^{\infty,\infty} \left(a_{jk}^{nm}\right)^2 = 0$. This completes the proof. □

Now let us give sufficient conditions for the Borel property. First we consider the following sets

$$D_0(\mathcal{A}) = \left\{x \in X : \ (\mathcal{A}x)_{nm} \text{ diverges}\right\},$$
$$D_1(\mathcal{A}) = \left\{x \in X : \ (\mathcal{A}x)_{nm} \ \textit{converges}\right\},$$
$$D_2(\mathcal{A}) = \left\{x \in X : \ (\mathcal{A}x)_{nm} \to \frac{1}{2} \ (n, m \to \infty)\right\}.$$

We examine the relationship between these sets in the sense of P-measure.

2.4. Theorem. *Let $\mathcal{A} = \left(a_{jk}^{nm}\right)$ be a 4-dimensional bounded regular matrix. The sets $D_1(\mathcal{A})$ and $D_2(\mathcal{A})$ have the same measure and the value is either 0 or 1.*

Proof. Choose an arbitrary $x \in D_1(\mathcal{A})$ (or $D_2(\mathcal{A})$). Let $\widehat{x}$ be a sequence obtained by altering a finite term of x. We have the following equality

$$\begin{aligned}
\sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm}\widehat{x}_{jk} &= \sum_{j,k=1,1}^{j_0,k_0} a_{jk}^{nm}\widehat{x}_{jk} + \sum_{j>j_0 \ \text{veya} \ k>k_0} a_{jk}^{nm}\widehat{x}_{jk} \\
&= \sum_{j,k=1,1}^{j_0,k_0} a_{jk}^{nm}\widehat{x}_{jk} + \sum_{j>j_0 \ \text{veya} \ k>k_0} a_{jk}^{nm}x_{jk}.
\end{aligned}$$

From Proposition 1.2 (i), it follows $\widehat{x} \in D_1(\mathcal{A})$ (or $D_2(\mathcal{A})$). Hence the sets $D_1(\mathcal{A})$ and $D_2(\mathcal{A})$ are homogeneous [14]. Since homogeneous sets have measure 0 or 1 and $D_2(\mathcal{A}) \subset D_1(\mathcal{A})$, the proof will be completed if $P(D_1(\mathcal{A})) = 1$ implies $P(D_2(\mathcal{A})) = 1$. On the other hand we have

$$\lim_{n,m} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm}x_{jk} = \lim_{n,m} \frac{1}{2} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} + \lim_{n,m} \frac{1}{2} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x) \tag{2.4}$$

where $r_{jk}(x) = 2x_{jk} - 1$. If we choose $x \in D_1(\mathcal{A})$, we get $\lim\limits_{n,m} \sum\limits_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x) = h(x)$ for almost all $x \in X$. From (v), interchanging integral and sum we have

$$\begin{aligned}
\int_X h(x)\,dx &= \int_X \left(\lim_{n,m} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x)\right) dP(x) \\
&= \lim_{n,m} \int_X \left(\sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}(x)\right) dP(x) \\
&= \lim_{n,m} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} \left(\int_X r_{jk}(x)\,dP(x)\right) = 0.
\end{aligned}$$

Hence we have $h(x) = 0$ for almost all $x \in X$. Also since first part of the right hand side of (2.4) is $\frac{1}{2}$ we get $x \in D_2(\mathcal{A})$. This completes the proof. □

2.5. Corollary. *Let $\mathcal{A} = \left(a_{jk}^{nm}\right)$ be a 4-dimensional bounded regular matrix. The set $D_0(\mathcal{A})$ has measure 0 or 1.*

2.6. Corollary. *If $\mathcal{A} = \left(a_{jk}^{nm}\right)$ is a 4-dimensional bounded regular matrix sums almost all sequences of 0's and 1's, then the matrix has the Borel property.*

2.7. Theorem. *Let* $\mathcal{A} = \left(a_{jk}^{nm}\right)$ *be a 4-dimensional matrix. If* $P\left(D_1\left(\mathcal{A}\right)\right) = 1$*, then we have*

$$p_{nm} = \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} \text{ converges for each } n,m \text{ and } \lim_{n,m} p_{nm} = p \text{ exists,}$$

$$\mathcal{A}_{nm} = \sum_{j,k=1,1}^{\infty,\infty} \left(a_{jk}^{nm}\right)^2 < \infty \text{ for each } n,m.$$

The proof of the theorem is similar to those of Theorems 2.1 and 2.2, and therefore is omitted.

2.8. Lemma. *If* $\mathcal{A}$ *satisfies condition* (v)*, then we have*

$$\int_X \left|\psi_{nm}\left(x\right)\right|^{2r} dP\left(x\right) \leq \frac{(2r)!}{2^r r!} \left(\mathcal{A}_{nm}\right)^r \tag{2.5}$$

where r *is a positive integer,* $\psi_{nm}\left(x\right) = \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}\left(x\right)$ *and* $\mathcal{A}_{nm} = \sum_{j,k=1,1}^{\infty,\infty} \left(a_{jk}^{nm}\right)^2$.

The proof can be proved using Lemma 1 of [13].

2.9. Theorem. *If* $\mathcal{A} = \left(a_{jk}^{nm}\right)$ *satisfies* (ii), (v) *and the series*

$$\sum_{n,m=1,1}^{\infty,\infty} \left(\sum_{j,k=1,1}^{\infty,\infty} \left(a_{jk}^{nm}\right)^2 \right)^r \tag{2.6}$$

converges for some $r > 0$*, then* $\mathcal{A}$ *has the Borel property.*

Proof. To complete the proof it is sufficient to show that

$$\sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} x_{jk} = \frac{1}{2} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} + \frac{1}{2} \sum_{j,k=1,1}^{\infty,\infty} a_{jk}^{nm} r_{jk}\left(x\right) \tag{2.7}$$

the limit of the right hand side of (2.7) equals $\frac{1}{2}$ for almost all $x \in X$. From Lemma 2.8, the inequality (2.5) holds for every positive integer r. On the other hand since the series in (2.6) converges for some $r > 0$, we easily get

$$\sum_{n,m=1,1}^{\infty,\infty} \int_X \left|\psi_{nm}\left(x\right)\right|^{2r} dP\left(x\right) < \infty.$$

Using the Beppo-Levi theorem, we have $\sum_{n,m=1,1}^{\infty,\infty} \left|\psi_{nm}\left(x\right)\right|^{2r} < \infty$ for almost all $x \in X$. Hence we obtain for almost all $x \in X$ that

$$\lim_{n,m\to\infty} \psi_{nm}\left(x\right) = 0.$$

This completes the proof. □

It is shown in [4] that the 4-dimensional Cesàro matrix method $(C, 1, 1)$ has the Borel property. We can also deduce this result from Theorem 2.9. We have already observed that (2.2) is a necessary condition for the Borel property. We raise the question whether the converse of Theorem 2.3 is true. The answer is no as the following example shows.

Since a 4-dimensional matrix can be considered as a matrix of infinite matrices, we can look at every entry as a matrix.

Consider the 4-dimensional Cesàro matrix, $(C,1,1) = \left(c_{jk}^{nm}\right)$. Now we construct a 4-dimensional matrix $\mathcal{A} = \left(a_{jk}^{nm}\right)$ as follows:

Shift the every column to the right in every possible order as the number of nonzero elements.

For example since there exist two possible order, we have

$$\left(a_{jk}^{11}\right) = \begin{bmatrix} 1 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ ... \end{bmatrix}, \quad \left(a_{jk}^{12}\right) = \begin{bmatrix} 0 & 1 & 0 & ... \\ 0 & 0 & 0 & ... \\ ... \end{bmatrix}.$$

$$\left(a_{jk}^{13}\right) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ ... \end{bmatrix}, \quad \left(a_{jk}^{14}\right) = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ ... \end{bmatrix} ...$$

$$\left(a_{jk}^{17}\right) = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ ... \end{bmatrix}, \quad \left(a_{jk}^{18}\right) = \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & ... \\ ... \end{bmatrix}$$

in the above we have six possible orders. Now let us obtain $\left(a_{jk}^{21}\right)$, ... , $\left(a_{jk}^{26}\right)$.

$$\left(a_{jk}^{21}\right) = \begin{bmatrix} \frac{1}{2} & 0 & ... \\ \frac{1}{2} & 0 & ... \\ 0 & 0 & ... \\ ... \end{bmatrix}, \quad \left(a_{jk}^{22}\right) = \begin{bmatrix} \frac{1}{2} & 0 & 0 & ... \\ 0 & \frac{1}{2} & 0 & ... \\ 0 & 0 & 0 & ... \\ ... \end{bmatrix}, ..., \left(a_{jk}^{26}\right) = \begin{bmatrix} 0 & 0 & 0 & ... \\ \frac{1}{2} & \frac{1}{2} & 0 & ... \\ 0 & 0 & 0 & ... \\ ... \end{bmatrix}.$$

Continuing this procedure we can construct the matrix $\mathcal{A}$.

Observe that the matrix $\mathcal{A}$ constructed above satisfies the condition (2.2).

Now let us consider the sequence $\{x_{jk}\}$ having $(\eta\mu + p)$ times 1 ve $(\eta\mu - p)$ times 0 in the rectangle $(\eta, 2\mu)$.

In the case of $p = 0$, an element of the matrix $\mathcal{A}$ which consists of 0's and $\frac{1}{\eta\mu}$'s sums the sequence $\{x_{jk}\}$ to 0 and the another one sums to 1. Let these terms be (n_0, m_0) and (n_1, m_1) respectively.

If $\left(a_{j,k}^{n_0,m_0}\right)$ containing $\dfrac{1}{\eta\mu}$'s, such that all the 0's of the sequence in the rectangle $(\eta, 2\mu)$ correspond with $\dfrac{1}{\eta\mu}$'s, we have

$$\sum_{j,k} a_{j,k}^{n_0,m_0} x_{jk} = 0.$$

Also if $\left(a_{j,k}^{n_1,m_1}\right)$ containing $\dfrac{1}{\eta\mu}$'s, such that all the 1's of the sequence in the rectangle $(\eta, 2\mu)$ correspond with $\dfrac{1}{\eta\mu}$'s, we have

$$\sum_{j,k} a_{j,k}^{n_1,m_1} x_{jk} = 1.$$

In the case of $p > 0$ there is an entry $\left(a_{j,k}^{n_0,m_0}\right)$ containing $\dfrac{1}{\eta\mu}$'s, such that all the 1's of the sequence in the rectangle $(\eta, 2\mu)$ correspond with $\dfrac{1}{\eta\mu}$'s, we have

$$\sum_{j,k} a_{j,k}^{n_0,m_0} x_{jk} = 1.$$

Also there is another entry $\left(a_{j,k}^{n_1,m_1}\right)$ containing $\dfrac{1}{\eta\mu}$'s, such that all the 0's of the sequence in the rectangle $(\eta, 2\mu)$ correspond with $\dfrac{1}{\eta\mu}$'s, we have

$$\sum_{j,k} a_{j,k}^{n_1,m_1} x_{jk} = \frac{p}{\eta\mu}.$$

In the case of $p < 0$ there is an entry $\left(a_{j,k}^{n_0,m_0}\right)$ containing $\dfrac{1}{\eta\mu}$'s, such that all the 0's of the sequence in the rectangle $(\eta, 2\mu)$ correspond with $\dfrac{1}{\eta\mu}$'s, we have

$$\sum_{j,k} a_{j,k}^{n_0,m_0} x_{jk} = 0.$$

Also there is another entry $\left(a_{j,k}^{n_1,m_1}\right)$ containing $\dfrac{1}{\eta\mu}$'s, such that all the 1's of the sequence in the rectangle $(\eta, 2\mu)$ correspond with $\dfrac{1}{\eta\mu}$'s, we have

$$\sum_{j,k} a_{j,k}^{n_1,m_1} x_{jk} = 1 + \frac{p}{\eta\mu}.$$

In any cases above, the oscillation of the sum $\sum a_{j,k}^{n,m} x_{jk}$ in the inner matrix containing $\frac{1}{\eta\mu}$'s is at least $1 - \frac{|p|}{\eta\mu}$. In order that $\{x_{jk}\}$ is $\mathcal{A}$-summable we necessarrily have $\frac{|p|}{\eta\mu} \to 1$, as $\eta, \mu \to \infty$.

Since almost all double sequences of 0's and 1's is $(C, 1, 1)$-summable to $\frac{1}{2}$, the set of sequences which $\frac{|p|}{\eta\mu}$ tends to 1 has P-measure 1. From this it follows that the set of sequences for which $\frac{|p|}{\eta\mu}$ tends to 1 is of P-measure 0. Therefore, $\mathcal{A}$ does not have the Borel property. That is condition (2.2) can not be sufficient.

References

[1] Borel E., *Les probabilities denombrables et leurs applications arithmetiques*, Rendiconti del Circolo Matematico di Palermo, 27, 247-271, 1909.

[2] Bromwich M.A., *An introduction to the theory of infinite series*, (Macmillan Co., London, 1942).

[3] Connor J., *Almost none of the sequences of* 0*'s and* 1*'s are almost convergent*, Internat. J. Math. Math. Sci. 13, 775-777, 1990.

[4] Crnjac M., Čunjalo F. and Miller H.I., *Subsequence characterizations of statistical convergence of double sequences*, Radovi Math., 12, 163-175, 2004.

[5] Garreau G.A., *A note on the summation of sequences of* 0*'s and* 1*'s*, Annals of Mathematics, 54, 183-185, 1951.

[6] Hill J.D., *Summability of sequences of* 0*'s and* 1*'s*, Annals of Mathematics, 46, 556-562, 1945.

[7] Hill J.D., *The Borel property of summability methods*, Pacific J. Math., 1, 399-409, 1951.

[8] Hill J.D., *Remarks on the Borel property*, Pacific J. Math., 4, 227-242, 1954.

[9] Móricz F. and Rhoades B.E., *Almost convergence of double sequences and strong regularity of summability matrices*, Math. Proc. Camb. Phil. Soc. 104, 283-294, 1988

[10] Parameswaran M.R., *Note on the summability of sequences of zeros and ones*, Proc. Nat. Inst. Sci. India Part A 27, 129-136, 1961.

[11] Pringsheim A., *On the theory of double infinite sequences of numbers. (Zur theorie der zweifach unendlichen zahlenfolgen.)*, Math. Ann., 53, 289-321, 1900.

[12] Robison G.M., *Divergent double sequences and series*, Trans. Amer. Math. Soc. 28, 50-73, 1926.

[13] Tas E. and Orhan C., *Cesàro means of subsequences of double sequences*, submitted.

[14] Visser C., *The law of nought-or-one in the theory of probability*, Studia Mathematica, 7, 143-149, 1938.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 483 – 495

Numerical computation and properties of the two dimensional exponential integrals

Şeyhmus Yardımcı*†, Murat Olgun ‡ and Çağla Can§

Abstract

In this paper, we investigate problem of convergence of the two-dimensional exponential integral (TDEI) functions arising in the study of the radiative transfer in a multi-dimensional medium. In our study, generalized exponential integral function's (GEIF) are expressed with double improper integrals as given in the original expression. Then we study the properties and asymptotic behaviour of the TDEI functions. We also give numerical computations of the values of TDEI functions.

Keywords: Numerical computations, exponential integrals, uniform convergence, multi-dimensional radiative transfer.

2000 AMS Classification: 33F05, 65Z05, 40A10

Received : 26.11.2014 *Accepted* : 27.02.2015 *Doi* : 10.15672/HJMS.20164512505

1. Introduction

The two-dimensional exponential integral (TDEI) functions play an important role in various fields of theoretical physics, quantum chemistry, theory of transport process, theory of fluid flow and theory of radiative transfer in a multi-dimensional medium [6], [7], [11]-[13], [19], [22]. The TDEI functions are especially useful for the study of anisotropic scattering in a two-dimensional medium with a scattering phase function [12], [13], [22]. Breig and Crosbie derived a series expansion and recurrence relations suitable for numerical computation of the one-dimensional exponential integral functions [7]. It is shown that the absorption of solar radiation by the earth's atmosphere is given in terms of first-order exponential integral function. The fundamental integral equation of the radiative transfer of two-dimensional planar media with anisotropic scattering was derived

*Department of Mathematics, Ankara University, 06100 Tandoğan, Ankara, TURKEY, Email : smyardimci@ankara.edu.tr
†Corresponding Author.
‡Department of Mathematics, Ankara University, 06100 Tandoğan, Ankara, TURKEY, Email : olgun@ankara.edu.tr
§Department of Mathematics, Ankara University, 06100 Tandoğan, Ankara, TURKEY, Email : cglcan@hotmail.com

by Crosbie and Dougherty [11]. Note that the TDEI functions are the kernel of that integral equation. The TDEI functions play an important role in the investigation of the two-dimensional radiative transfer in an absorbing-emitting cylindrical medium and determination of the radiative flux [12]. The generalized exponential integral functions are studied in [1]-[5], [10], [17]. In [2] GEIF 's are expressed with the single integrals. In our study, GEIF 's are expressed with double improper integrals as given in the original expression . This depends on the truth that the uniform convergence of integrals gives more precise results. Also in [2] GEIF 's are given In terms of Bessel functions, in the form of series. This GEIF 's approximation gives ruder results compared to ours. This study uses a different methodology from [1], [2], [10], [17] and results are achieved with higher accuracy. The TDEI functions examined in this work are defined as

$$\varepsilon_n(\tau,\beta)=\frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\exp(-r)/r^{n+1}\right]\exp(-i\beta x)\,dxdy \tag{1.1}$$

where $r^2=x^2+y^2+\tau^2$ and $n=1,2,\dots$.
Note that the TDEI functions are two-dimensional analogs of the exponential integral functions [14].

$$E_n(\tau)=\int_{1}^{\infty}t^{-n}\exp(-\tau t)\,dt, \tag{1.2}$$

$n=1,2,\dots$. The exponential integral function (1.2) plays an important role in various fields of theoretical physics, quantum chemistry and theory of transport process [8], [9], [16], [20], [21].
Many properties of the TDEI functions depend on the uniform convergence of the improper integral (1.1). In this paper, we study the problem of convergence of the TDEI functions $\varepsilon_n(\tau,\beta)$. We also investigate the properties, asymptotic behaviour and numerical computation of the TDEI functions.

2. Uniform convergence

Let us consider the improper integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(\tau,\beta,x,y)\,dxdy \tag{2.1}$$

where $(\tau,\beta)\in D\subset\mathbb{R}^2$.

2.1. Definition. [18]. Integral (2.1) is said to be uniformly convergent with respect to $(\tau,\beta)\in D$ if it is convergent for all (τ,β) and if, given any $\varepsilon>0$, there is a sufficiently large number R_0 independent of (τ,β) and such that for any R satisfying the inequality $R>R_0$ there holds the inequality

$$\left|\int\int_{R^2-\omega_R}f(\tau,\beta,x,y)\,dxdy\right|<\varepsilon$$

where ω_R is the ball of radius R with centre of the origin.

2.2. Theorem. [18]. *If for the function $f(\tau,\beta,x,y)$ in question there holds the inequality*

$$|f(\tau,\beta,x,y)|\le\varphi(x,y)\,,\ \ (\tau,\beta)\in D$$

and the improper integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)\,dxdy$$

is convergent then integral (2.1) *convergences uniformly with respect to* (τ,β) *on* D.

For all $\varepsilon>0$ we define the domain

$$D(\varepsilon)=\{\ (\tau,\beta)\in\Omega,\ \tau\in[\varepsilon,\infty),\ \beta\in(-\infty,\infty)\ \},$$

where

$$\Omega=\{\ (\tau,\beta)\in\mathbb{R}^2,\ \tau\in[0,\infty),\ \beta\in(-\infty,\infty)\ \}.$$

2.3. Theorem. *i) The two-dimensional exponential integral function* $\varepsilon_n(\tau,\beta)$ *is uniformly convergent with respect to* (τ,β) *on* $D(\varepsilon)$.
ii) The function $\varepsilon_n(\tau,\beta)$ *is nonuniformly convergent with respect to* (τ,β) *on* Ω.

Proof. i)

$$\varepsilon_n(\tau,\beta)=\frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\exp(-r)/r^{n+1}\right]\exp(-i\beta x)\,dxdy$$

where $r^2=x^2+y^2+\tau^2$.
If we define

$$g(\tau,\beta,x,y)=\left[\exp\left(-\sqrt{x^2+y^2+\tau^2}\right)/\left(\sqrt{x^2+y^2+\tau^2}\right)^{n+1}\right]\exp(-i\beta x)$$

then

$$\varepsilon_n(\tau,\beta)=\frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(\tau,\beta,x,y)\,dxdy.$$

For all $\varepsilon>0$ we have

$$|g(\tau,\beta,x,y)|\le\exp\left(-\sqrt{x^2+y^2+\varepsilon^2}\right)/\varepsilon^n\sqrt{x^2+y^2+\varepsilon^2}$$

and

$$\begin{aligned}&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left(-\sqrt{x^2+y^2+\varepsilon^2}\right)/\varepsilon^n\sqrt{x^2+y^2+\varepsilon^2}dxdy\\ =\ &2\pi\varepsilon^{-n}\exp(-\varepsilon)<\infty.\end{aligned}$$

So from the Theorem 2.1 we find that the function $\varepsilon_n(\tau,\beta)$ is uniformly convergent with respect to (τ,β) on $D(\varepsilon)$.

ii) For all $(\tau,\beta)\in\Omega$ we have

$$\begin{aligned}\varepsilon_n(\tau,\beta)&=\frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\exp(-r)/r^{n+1}\right]\exp(-i\beta x)\,dxdy\\ &>-\frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(-r)/r^{n+1}dxdy.\end{aligned}$$

Let us consider

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\varphi\left(\tau,x,y\right)/r^{n+1}dxdy$$

where $\varphi\left(\tau,x,y\right)=\exp\left(-r\right).$
$\varphi\left(\tau,0,0\right)=\exp\left(-\tau\right)>0$. Then there is a ball ω with centre at the origin and radius so small that $\left|\varphi\left(\tau,x,y\right)\right|>\frac{\left|\varphi(\tau,0,0)\right|}{2}=\frac{\exp(-\tau)}{2}>0$ on that ball.
Therefore

$$\begin{aligned}\left|\int\int\limits_{\omega}\varphi\left(\tau,x,y\right)/r^{n+1}dxdy\right| &= \int\int\limits_{\omega}\left|\varphi\left(\tau,x,y\right)\right|/r^{n+1}dxdy\\ &> \frac{\exp\left(-\tau\right)}{2}\int\int\limits_{\omega}1/r^{n+1}dxdy\to\infty\text{ for }\tau\to 0\end{aligned}$$

Therefore the function $\varepsilon_n\left(\tau,\beta\right)$ is nonuniformly convergent with respect to $\left(\tau,\beta\right)$ on Ω. ∎

3. Properties of the TDEI function

Let $G\in\mathbb{R}_m$ and $D\in\mathbb{R}_n$ where $\mathbb{R}_m$ and $\mathbb{R}_n$ denote m-dimensional and n-dimensional spaces, respectively.
We shall consider an integral of the form

$$F\left(\xi\right)=\int\limits_{D}f\left(\xi,\eta\right)d\eta\qquad\left(\xi\in G\right)\tag{3.1}$$

taken over an unbounded domain D such that it has the point at infinity as its only singularity for any $\xi\in G$.

3.1. Theorem. [18]. *If the function $f\left(\xi,\eta\right)$ is continuous on*

$$G\times D:=\left\{\xi\in G,\ \eta\in D\right\}$$

and if the integral (3.1) *is uniformly convergent with respect to ξ on G, then the function*

$$F\left(\xi\right)=\int\limits_{D}f\left(\xi,\eta\right)d\eta$$

is a continuous function with respect to ξ on G and for all $\xi_0\in G$

$$\lim_{\xi\to\xi_0}\int\limits_{D}f\left(\xi,\eta\right)d\eta=\int\limits_{D}\lim_{\xi\to\xi_0}f\left(\xi,\eta\right)d\eta.$$

3.2. Theorem. [18]. *If the conditions of Theorem 3.1 hold then the function $F\left(\xi\right)$ can be integrated with respect to ξ on G under the integral sign, that is*

$$\int\limits_{G}F\left(\xi\right)d\xi=\int\limits_{G}d\xi\int\limits_{D}f\left(\xi,\eta\right)d\eta=\int\limits_{D}d\eta\int\limits_{G}f\left(\xi,\eta\right)d\xi.$$

3.3. Theorem. *The function $\varepsilon_n\left(\tau,\beta\right)$ is continuous with respect to $\left(\tau,\beta\right)$ on $D\left(\varepsilon\right)$.*

Proof. Let (τ_0, β_0) is any point of $D(\varepsilon)$. In view of Theorems 2.2 and 3.1 we get

$$\begin{aligned}
\lim_{\substack{\tau\to\tau_0\\ \beta\to\beta_0}} \varepsilon_n(\tau,\beta) &= \lim_{\substack{\tau\to\tau_0\\ \beta\to\beta_0}} \frac{\tau^{n-1}}{2\pi}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left[\exp(-r)/r^{n+1}\right]\exp(-i\beta x)\,dxdy \\
&= \frac{\tau_0^{n-1}}{2\pi}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\lim_{\substack{\tau\to\tau_0\\ \beta\to\beta_0}}\left[\exp(-r)/r^{n+1}\right]\exp(-i\beta x)\,dxdy \\
&= \frac{\tau_0^{n-1}}{2\pi}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\frac{\exp\left(-\sqrt{x^2+y^2+\tau_0^2}\right)}{\left(\sqrt{x^2+y^2+\tau_0^2}\right)^{n+1}}\exp(-i\beta_0 x)\,dxdy \\
&= \varepsilon_n(\tau_0,\beta_0).
\end{aligned}$$

Therefore the function $\varepsilon_n(\tau,\beta)$ is continuous on $D(\varepsilon)$. ∎

3.4. Theorem. [18]. *If ξ is a scalar variable running through a closed interval $[a,b]$ and the function $f(\xi,\eta)$ is continuous together with its partial derivative $\frac{\partial f}{\partial \xi}$ on $[a,b]\times\overline{D}$ and if integral (3.1) is convergent while the integral*

$$F_1(\xi) = \int\limits_D \frac{\partial f}{\partial \xi} f(\xi,\eta)\,d\eta$$

is uniformly convergent with respect to $\xi \in [a,b]$ then $F^{'}(\xi) = F_1(\xi)$, that is

$$\frac{\partial f}{\partial \xi}\int\limits_D f(\xi,\eta)\,d\eta = \int\limits_D \frac{\partial f}{\partial \xi} f(\xi,\eta)\,d\eta.$$

3.5. Definition. [23]. Suppose $f : \mathbb{R} \longrightarrow \mathbb{C}$ is a locally integrable function on $\mathbb{R}$.
The function

$$\hat{f}(\beta) := \frac{1}{\sqrt{2\pi}}\int\limits_{\mathbb{R}} f(x)\exp(-i\beta x)\,dx$$

is called *Fourier transform of the function f.*

3.6. Theorem. [18]. *If $f \in L = L(\mathbb{R})$, then $\hat{f}$,as bounded continuous functions, possess the property*

$$\lim_{|\beta|\to\infty}\hat{f}(\beta) = 0$$

3.7. Theorem. $\varepsilon_n(\tau,\beta)$ *satisfies the following asymptotic equations:*

$$\varepsilon_n(\tau,\beta) = o(1), \quad (\tau,\beta)\in D(\varepsilon), \ \tau\to\infty. \tag{3.2}$$

$$\varepsilon_n(\tau,\beta) = o(1), \quad (\tau,\beta)\in D(\varepsilon), \ \beta\to\pm\infty. \tag{3.3}$$

$$\varepsilon_n(\tau,\beta) = E_n(\tau) + o(1), \quad (\tau,\beta)\in D(\varepsilon), \ \beta\to\infty. \tag{3.4}$$

Proof. Using Theorem 2.2 and 3.1, we obtain that, for all $(\tau,\beta)\in D(\varepsilon)$

$$\begin{aligned}\lim_{\tau\to\infty}\varepsilon_n(\tau,\beta) &= \lim_{\tau\to\infty}\frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\exp(-r)/r^{n+1}\right]\exp(-i\beta x)\,dxdy\\ &< \lim_{\tau\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\exp\left(-\sqrt{x^2+y^2+\tau^2}\right)/\tau^2\right]\exp(-i\beta x)\,dxdy\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\lim_{\tau\to\infty}\left[\exp\left(-\sqrt{x^2+y^2+\tau^2}\right)/\tau^2\right]\exp(-i\beta x)\,dxdy\\ &= 0\end{aligned}$$

(3.2) holds.

Let us prove (3.3)

$$\begin{aligned}\varepsilon_n(\tau,\beta) &= \frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\exp(-r)/r^{n+1}\right]\exp(-i\beta x)\,dxdy\\ &< \frac{\tau^{n-1}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left[\exp\left(-\sqrt{x^2+\tau^2}\right)/\left(\sqrt{x^2+\tau^2}\right)^{n+1}\right]\exp(-i\beta x)\,dx\right)dy\\ &= \frac{\tau^{n-1}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(\tau,x)\exp(-i\beta x)\,dx\right)dy\end{aligned}$$

where

$$f(\tau,x)=\exp\left(-\sqrt{x^2+\tau^2}\right)/\left(\sqrt{x^2+\tau^2}\right)^{n+1}.$$

Then from the Definition 3.1 we get

$$\varepsilon_n(\tau,\beta)=\frac{\tau^{n-1}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\hat{f}(\tau,\beta)\,dy \tag{3.5}$$

and due to the following expression

$$\begin{aligned}\int_{-\infty}^{\infty}|f(\tau,x)|\,dx &= \int_{-\infty}^{\infty}\exp\left(-\sqrt{x^2+\tau^2}\right)/\left(\sqrt{x^2+\tau^2}\right)^{n+1}dx\\ &< 2\exp(-\tau)\int_{\tau}^{\infty}1/u^n\sqrt{u^2-\tau^2}du<\infty,\end{aligned}$$

$f\in L(\mathbb{R})$ by virtue of Theorem 3.5

$$\lim_{|\beta|\to\infty}\hat{f}(\tau,\beta)=0. \tag{3.6}$$

From the Theorem 2.2, 3.1 and (3.5), (3.6) we get

$$\begin{aligned}\lim_{|\beta|\to\infty}\varepsilon_n(\tau,\beta) &= \frac{\tau^{n-1}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\lim_{|\beta|\to\infty}\hat{f}(\tau,\beta)\,dy\\ &= 0,\end{aligned}$$

i.e., (3.3) holds.

According to Theorem 2.2, 3.1 we obtain,

$$\begin{aligned}
\lim_{\beta\to 0}\varepsilon_n(\tau,\beta) &= \lim_{\beta\to 0}\frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\exp(-r)/r^{n+1}\right]\exp(-i\beta x)\,dxdy \\
&= \frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\exp(-r)/r^{n+1}\right]\lim_{\beta\to 0}\exp(-i\beta x)\,dxdy \\
&= \frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(-r)/r^{n+1}dxdy \\
&= \frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{1}^{\infty}\exp(-tr)/r^{n}dtdxdy \\
&= \frac{\tau^{n-1}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(\tau,x,y)\,dxdy
\end{aligned}$$

where

$$\phi(\tau,x,y)=\int_{1}^{\infty}\exp(-tr)/r^{n}dt.$$

It is clear that last integral is uniformly convergent with respect to (τ,x,y) on the domain $\left\{\tau\in[\varepsilon,\infty),\ \ (x,y)\in\mathbb{R}^2\right\}.$

So making use of the Theorem 3.2 we have

$$\lim_{\beta\to 0}\varepsilon_n(\tau,\beta)=\frac{\tau^{n-1}}{2\pi}\int_{1}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(t-1)^{n-n}\exp(-tr)/r^{n}dxdydt$$

using the Theorem 2.2, 3.4 and (n-1) times integration by parts yield

$$\begin{aligned}
\lim_{\beta\to 0}\varepsilon_n(\tau,\beta) &= \frac{\tau^{n-1}}{(n-1)!}\frac{1}{2\pi}\int_{1}^{\infty}(t-1)^{n-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(-tr)/rdxdydt \\
&= \frac{\tau^{n-1}}{(n-1)!}\int_{1}^{\infty}(t-1)^{n-1}\exp(-\tau t)/tdt. \qquad (3.7)
\end{aligned}$$

The last term can be written in the following form

$$\begin{aligned}
\int_{1}^{\infty}(t-1)^{n-1}\exp(-\tau t)/tdt &= \int_{1}^{\infty}(t-1)^{n-2}\exp(-\tau t)\,dt-\int_{1}^{\infty}(t-1)^{n-3}\exp(-t\tau)\,dt \\
&\quad+\int_{1}^{\infty}(t-1)^{n-4}\exp(-t\tau)\,dt \\
&\quad-...+(-1)^{n+1}\int_{1}^{\infty}(t-1)\exp(-t\tau)\,dt \\
&\quad+(-1)^{n}\int_{1}^{\infty}(t-1)\exp(-t\tau)/tdt.
\end{aligned}$$

If we use integration by parts for all term we get

$$\begin{aligned}\int_1^\infty (t-1)^{n-1}\exp(-\tau t)/tdt &= [\frac{(n-2)!}{\tau^{n-1}}-\frac{(n-3)!}{\tau^{n-2}}+\frac{(n-4)!}{\tau^{n-3}} \\ &-...+(-1)^n\frac{2!}{\tau^3}+(-1)^{n-1}\frac{1!}{\tau^2}+(-1)^n\frac{0!}{\tau}]\exp(-\tau) \\ &+(-1)^{n-1}\int_1^\infty \exp(-\tau t)/tdt \qquad (3.8)\end{aligned}$$

By means of integration by parts for the right-hand side of (3.8) then

$$\begin{aligned}(-1)^{n-1}\int_1^\infty \exp(-t\tau)/tdt &= (-1)^{n-1}[\frac{0!}{\tau}-\frac{1!}{\tau^2}+\frac{2!}{\tau^3}-...+(-1)^n\frac{(n-4)!}{\tau^{n-3}} \\ &+(-1)^{n-1}\frac{(n-3)!}{\tau^{n-2}}+(-1)^n\frac{(n-2)!}{\tau^{n-1}}]\exp(-\tau) \\ &+\frac{(n-1)!}{\tau^{n-1}}\int_1^\infty \exp(-t\tau)/t^n dt \; label 3.9 \qquad (3.9)\end{aligned}$$

substitution of (3.9) into (3.8) gives

$$\int_1^\infty (t-1)^{n-1}\exp(-\tau t)/tdt = (n-1)!/\tau^{n-1}\int_1^\infty \exp(-\tau t)/t^n dt \qquad (3.10)$$

considering the substitution of (3.10) into (3.7) gives

$$\begin{aligned}\lim_{\beta\to 0}\varepsilon_n(\tau,\beta) &= \int_1^\infty \exp(-\tau t)/t^n dt \\ &= E_n(\tau).\end{aligned}$$

∎

4. Numerical Computation

The numerical computation of TDEI functions have been studied by several authors. Those computation methods consist asymptotic or binomial series for TDEI function which include mass computatious [6]-[9], [11]- [16], [18]-[23]. On the basis of the uniform convergence of $\varepsilon_n(\tau,\beta)$, obtained in this paper we constructed a new simple and an accurate algorithm for the calculation of TDEI function even in a modarate PC. The computations were performed for large values of parameters τ and β. In this paper the TDEI functions were calculated on the Mathematica 8.0 international mathematical software. The comparative examples of computer calculatious for the TDEI functions are given in Tables 1-4. As can be seen from tables, our computational results are in agreement with literature [7]. Also from Tables 1-4 we see that the calculation results of the TDEI functions show good rate of convergence in the range of parameters $\tau \in [10^{-3}, 1]$ and $\beta \in [1, 20]$.

τ	$\beta = 1$	$\beta = 2$	$\beta = 5$	$\beta = 10$	$\beta = 20$
0.001	$A = 6.143314628912719$ $+iE - 1$ $B = 6.1433$	$A = 5.850334088713723$ $-i9.375391894121577E - 10$ $B = 5.8503$	$A = 5.216596350515581$ $-i6.184411880623505E - 17$ $B = 5.2166$	$A = 4.622405662773734$ $+i4.417437057588218E - 18$ $B = 4.6224$	$A = 3.979468204969179$ $+i9.718361526694078E - 17$ $B = 3.9795$
0.005	$A = 4.537901067889449$ $+i1.766974823035287E - 17$ $B = 4.5379$	$A = 4.245006981521351$ $+iE - 1$ $B = 4.2450$	$A = 3.6118092517793587$ $-i1.018351544013959E - 10$ $B = 3.6118$	$A = 3.0192768074691063$ $+i9.055745968055847E - 17$ $B = 3.0193$	$A = 2.381847868201705$ $+i1.987846675914698E - 17$ $B = 2.3818$
0.01	$A = 3.849813982491566$ $i - 3.837162415267503E - 11$ $B = 3.8498$	$A = 3.557142202395409$ $-i4.966478929026777E - 13$ $B = 3.5571$	$A = 2.9252964348968837$ $-i2.298302965135789E - 11$ $B = 2.9253$	$A = 2.3367504277586297$ $-i1.558030050211364E - 12$ $B = 2.3368$	$A = 1.7117877639753523$ $-i3.533949646070574E - 17$ $B = 1.7118$
0.025	$A = 2.9488329356872187$ $+i1.553338671316446E - 10$ $B = 2.9488$	$A = 2.657383190971331$ $+i1.993293375806704E - 12$ $B = 2.6574$	$A = 2.0326700716719333$ $-9.81770526868218E - 12$ $B = 2.0327$	$A = 1.4637881869087266$ $-i1.071335664531752E - 11$ $B = 1.4638$	$A = 8.941114707500399E - 1$ $-i5.300924469105861E - 17$ $B = 8.9411E - 1$
0.05	$A = 2.281453068639485$ $-i5.300924469105861E - 17$ $B = 2.2815$	$A = 1.9933445935560798$ $+i1.046233699931526E - 10$ $B = 1.9933$	$A = 1.3870213973918706$ $-i1.026298209581762E - 11$ $B = 1.3870$	$A = 8.639169226626785E - 1$ $-i4.310311171953727E - 12$ $B = 8.6392E - 1$	$A = 4.026457744554906E - 1$ $-i2.696189610344371E - 22$ $B = 4.0265E - 1$
0.1	$A = 1.64017231138455$ $+iE - 1$ $B = 1.6402$	$A = 1.3617758308830468$ $+i7.0578181485961E - 12$ $B = 1.3618$	$A = 8.043034952598128E - 1$ $-i1.985758332545723E - 12$ $B = 8.0430E - 1$	$A = 3.8435848086903923E - 1$ $+i8.28269448297791E - 17$ $B = 3.8436E - 1$	$A = 1.0713555332477835E - 1$ $+i6.092564529322131E - 12$ $B = 1.0714E - 1$
0.5	$A = 4.2370709749072205E - 1$ $+iE - 1$ $B = 4.2371E - 1$	$A = 2.5067102641438316E - 1$ $+iE - 1$ $B = 2.5067E - 1$	$A = 4.630293776083369E - 2$ $+iE - 1$ $B = 4.6303E - 2$	$A = 3.0239699124587406E - 3$ $+i1.224674029569061E - 16$ $B = 3.0240E - 3$	$A = 1.552042299070299E - 5$ $+iE - 1$ $B = 1.5520E - 5$
1	$A = 1.3554692860259693E - 1$ $+iE - 1$ $B = 1.3555E - 1$	$A = 5.351461855900747E - 2$ $+iE - 1$ $B = 5.3515E - 2$	$A = 2.3951848541335623E - 3$ $+iE - 1$ $B = 2.3952E - 3$	$A = 1.3321816513904781E - 5$ $+i7.07480153754363E - 18$ $B = 1.3322E - 5$	$A = 4.719616145834231E - 10$ $-i9.101430204813521E - 16$ $B = 4.7196E - 10$

Table 1. Value of $\varepsilon_1(\tau, \beta)$

τ	$\beta = 1$	$\beta = 2$	$\beta = 5$	$\beta = 10$	$\beta = 20$
0.001	$A = 9.924434723782973E - 1$ $+i1.413579858428229E - 19$ $B = 9.9244E - 1$	$A = 9.919160966891798E - 1$ $+iE - 1$ $B = 9.9192E - 1$	$A = 9.896973641980313E - 1$ $-i4.915307297290383E - 10$ $B = 9.8970E - 1$	$A = 9.853780518462402E - 1$ $+i1.413579858428229E - 19$ $B = 9.8538E - 1$	$A = 9.761947163307825E - 1$ $+i6.069070914791562E - 11$ $B = 9.7619E - 1$
0.005	$A = 9.702643422545342E - 1$ $+i2.208718528794109E - 20$ $B = 9.7026E - 1$	$A = 9.676568670129494E - 1$ $+i1.413579858428229E - 18$ $B = 9.6766E - 1$	$A = 9.567681149630682E - 1$ $+i7.112741170587211E - 11$ $B = 9.5677E - 1$	$A = 9.35895855570027E - 1$ $-7.619368247065852E - 11$ $B = 9.3590E - 1$	$A = 8.928151526927118E - 1$ $-3.180554681463517E - 18$ $B = 8.9282E - 1$
0.01	$A = 9.474592559549745E - 1$ $-i4.586072717261649E - 15$ $B = 9.4746E - 1$	$A = 9.423160462390637E - 1$ $+i3.533949646070575E - 18$ $B = 9.4232E - 1$	$A = 9.210350252606401E - 1$ $+i3.577803230650029E - 12$ $B = 9.2104E - 1$	$A = 8.810187337876387E - 1$ $+i2.827159716856459E - 18$ $B = 8.8102E - 1$	$A = 8.014083467390319E - 1$ $-i1.464519149540784E - 11$ $B = 8.0141E - 1$
0.025	$A = 8.915415485850164E - 1$ $+i4.417437057588218E - 19$ $B = 8.9154E - 1$	$A = 8.791975192541199E - 1$ $+i3.639194635321662E - 13$ $B = 8.7920E - 1$	$A = 8.294982431226492E - 1$ $+i1.435944127455608E - 12$ $B = 8.2950E - 1$	$A = 7.412356049804204E - 1$ $-i2.8175139356756E - 13$ $B = 7.4124E - 1$	$A = 5.837991466302784E - 1$ $+i1.250068888192043E - 12$ $B = 5.8380E - 1$
0.05	$A = 8.176587748648632E - 1$ $+i2.255574835843579E - 13$ $B = 8.1766E - 1$	$A = 7.945528161372075E - 1$ $+i1.766974823035287E - 18$ $B = 7.9455E - 1$	$A = 7.056034194084425E - 1$ $-i8.834874115176437E - 18$ $B = 7.0560E - 1$	$A = 5.618241448033833E - 1$ $+i4.594134539891746E - 17$ $B = 5.6182E - 1$	$A = 3.472878771358602E - 1$ $-i1.943672305338816E - 17$ $B = 3.4729E - 1$
0.1	$A = 7.041062167438795E - 1$ $-i7.496678717427851E - 13$ $B = 7.0411E - 1$	$A = 6.634519064258186E - 1$ $-i3.909705677063143E - 12$ $B = 6.6345E - 1$	$A = 5.201241100703188E - 1$ $+i1.855323564187051E - 17$ $B = 5.2012E - 1$	$A = 3.27613339905977E - 1$ $+i1.060184893821172E - 17$ $B = 3.2761E - 1$	$A = 1.2428402294804283E - 1$ $+iE - 1$ $B = 1.2428E - 1$
0.5	$A = 2.812151426498713E - 1$ $+iE - 1$ $B = 2.8122E - 1$	$A = 2.015863821445704E - 1$ $+iE - 1$ $B = 2.0159E - 1$	$A = 5.496848551136895E - 2$ $+iE - 1$ $B = 5.4968E - 2$	$A = 5.0600102554520555E - 3$ $-i3.224667578916283E - 15$ $B = 5.0600E - 3$	$A = 3.7076101127100064E - 5$ $+iE - 1$ $B = 3.7076E - 5$
1	$A = 1.0756980583160117E - 1$ $+iE - 1$ $B = 1.0757E - 1$	$A = 5.336330710138579E - 2$ $+iE - 1$ $B = 5.3363E - 2$	$A = 3.7075424200455013E - 3$ $+iE - 1$ $B = 3.7075E - 3$	$A = 2.9869304279352523E - 5$ $-i5.866908592109352E - 19$ $B = 2.9869E - 5$	$A = 1.538334495203213E - 9$ $-i2.63033517911171E - 18$ $B = 1.5383E - 9$

Table 2. Value of $\varepsilon_2(\tau, \beta)$

τ	$\beta = 1$	$\beta = 2$	$\beta = 5$	$\beta = 10$	$\beta = 20$
0.001	$A = 4.990024511109741E-1$ $-i2.261727773485167E-21$ $B = 4.9900E-1$	$A = 4.989982320156957E-1$ $-i1.776096495894175E-12$ $B = 4.9900E-1$	$A = 4.989706876695545E-1$ $-i2.714073328182201E-19$ $B = 4.9897E-1$	$A = 4.9888175736258644E-1$ $-i1.442151075825898E-9$ $B = 4.9888E-1$	$A = 4.9856897323705635E-1$ $+i5.790023100122029E-19$ $B = 4.9857E-1$
0.005	$A = 4.950516902549029E-1$ $-i2.544443745170813E-19$ $B = 4.9505E-1$	$A = 4.949762046326981E-1$ $-i1.90159059838535E-13$ $B = 4.9498E-1$	$A = 4.9449751991327523E-1$ $+i2.261727773485167E-19$ $B = 4.9450E-1$	$A = 4.930242429216283E-1$ $-i1.133479574912619E-13$ $B = 4.9302E-1$	$A = 4.8819820352499177E-1$ $-i7.237528875152536E-18$ $B = 4.8820E-1$
0.01	$A = 4.9018922746454213E-1$ $+i4.229134084646993E-15$ $B = 4.9019E-1$	$A = 4.899385430768269E-1$ $+i8.469661622952918E-15$ $B = 4.8994E-1$	$A = 4.883821102053627E-1$ $+i1.809382218788134E-18$ $B = 4.8838E-1$	$A = 4.837637107484798E-1$ $+iE-1$ $B = 4.8376E-1$	$A = 4.69496901043039E-1$ $-i1.970457721187423E-12$ $B = 4.6950E-1$
0.025	$A = 4.760365743919093E-1$ $+i8.834874115176437E-20$ $B = 4.7604E-1$	$A = 4.748853995162321E-1$ $-i6.626155586382327E-19$ $B = 4.7489E-1$	$A = 4.6805408708589463E-1$ $+i2.827159716856459E-18$ $B = 4.6805E-1$	$A = 4.4938184179542284E-1$ $-i3.180554681463517E-18$ $B = 4.4938E-1$	$A = 3.9923273619928473E-1$ $+2.216955693995738E-12$ $B = 3.9923E-1$
0.05	$A = 4.536942695174573E-1$ $+i4.28843815505793E-12$ $B = 4.5369E-1$	$A = 4.5029816049684535E-1$ $-i8.607344654558792E-12$ $B = 4.5030E-1$	$A = 4.313003524176357E-1$ $+i6.523616104772877E-13$ $B = 4.3130E-1$	$A = 3.849140185306733E-1$ $+i1.711286844512507E-13$ $B = 3.8491E-1$	$A = 2.8307826133448905E-1$ $+i1.753696007240177E-12$ $B = 2.8308E-1$
0.1	$A = 4.129052504262309E-1$ $+i2.260774994155964E-11$ $B = 4.1291E-1$	$A = 4.0381514638243077E-1$ $-i1.210519571178468E-11$ $B = 4.0382E-1$	$A = 3.5787773372208315E-1$ $+iE-1$ $B = 3.5788E-1$	$A = 2.662095412797268E-1$ $+iE-1$ $B = 2.6621E-1$	$A = 1.2689054916664513E-1$ $+iE-1$ $B = 1.2689E-1$
0.5	$A = 2.0269794527722496E-1$ $+iE-1$ $B = 2.0270E-1$	$A = 1.6215841061428077E-1$ $+iE-1$ $B = 1.6216E-1$	$A = 5.843514756064822E-2$ $+iE-1$ $B = 5.8435E-2$	$A = 7.175956423511752E-3$ $-i5.978291420340211E-15$ $B = 7.1760E-3$	$A = 7.267195676212158E-5$ $+iE-1$ $B = 7.2672E-5$
1	$A = 8.726359085360859E-2$ $+iE-1$ $B = 8.7264E-2$	$A = 5.059279476996788E-2$ $+iE-1$ $B = 5.0593E-2$	$A = 4.997178810918784E-3$ $+i1.104359264397054E-18$ $B = 4.9972E-3$	$A = 5.5981429368903135E-5$ $-i3.337882737606331E-19$ $B = 5.5981E-5$	$A = 4.09183571344698E-9$ $+iE-1$ $B = 4.0918E-9$

Table 3. Value of $\varepsilon_3(\tau, \beta)$

τ	$\beta = 1$	$\beta = 2$	$\beta = 5$	$\beta = 10$	$\beta = 20$
0.001	$3.3283365604269766E-1$ $+4.632018480097623E-21$	$3.328331613610311E-1$ $-i1.852807392039049E-20$	$3.3282970307744514E-1$ $+i3.705614784078098E-20$	$3.328173992319921E-1$ $-i5.123697977723025E-14$	$3.327684851036903E-1$ $+i1.02120053559172E-9$
0.005	$3.3084168843793876E-1$ $-i6.33283776575847E-20$	$3.308297036009815E-1$ $+i6.436868196427692E-14$	$3.3074636124698376E-1$ $+iE-1$	$3.3045341230211234E-1$ $+i3.047532523055772E-10$	$3.2932152187417696E-1$ $+i6.712438193978458E-13$
0.01	$3.283668166582222E-1$ $+i2.334533695204764E-14$	$3.2832041450482136E-1$ $-i4.668935305507558E-14$	$3.279999313606849E-1$ $+i7.237528875152537E-19$	$3.2689163817018646E-1$ $-i7.437146200877148E-11$	$3.2275697847261525E-1$ $-i5.587372291617758E-17$
0.025	$3.2104319438498374E-1$ $+i1.467757187226319E-12$	$3.207762091456098E-1$ $-i1.130863886742583E-18$	$3.18970362083023E-1$ $-i2.261727773485167E-18$	$3.130302814359257E-1$ $-i2.705897472175359E-12$	$2.9309972611874663E-1$ $+i4.099130355815903E-11$
0.05	$3.0917216818702864E-1$ $+i2.915508458008224E-19$	$3.0822192734798826E-1$ $+i3.220960448136829E-13$	$3.02023610866854E-1$ $+i2.261727773485167E-18$	$2.832990350153646E-1$ $+i6.43233237426568E-13$	$2.3065645383776914E-1$ $-i9.830608248932432E-13$
0.1	$2.866606684383088E-1$ $+i3.486063214786459E-13$	$2.835510053837311E-1$ $+i7.067899292141149E-20$	$2.6473792976651495E-1$ $-i1.953758586363162E-12$	$2.1692835774847108E-1$ $-i5.654319433712919E-18$	$1.2200667108444715E-1$ $+iE-1$
0.5	$1.5579006769776246E-1$ $+iE-1$	$1.3293249152133196E-1$ $+iE-1$	$5.86308991984218E-2$ $+iE-1$	$9.188975045090655E-3$ $-i1.104359264397054E-18$	$1.2516086082822907E-4$ $+iE-1$
1	$7.252504637283796E-2$ $+iE-1$	$4.690751377329581E-2$ $+iE-1$	$6.173857960619937E-3$ $-i5.521796321985272E-19$	$9.31476402535588E-5$ $+i9.518492990879253E-18$	$9.484194215269618E-9$ $-i2.229293299158516E-21$

Table 4. Value of $\varepsilon_4(\tau, \beta)$

References

[1] Z. Altac, *Integrals involving Bickley and Bessel functions in radiative transfer amd generalized exponential integral functions*, ASME Journal of Heat Transfer, Volume 118, 1996, Pages 789-792.

[2] Z. Altac, *Exact series expansions, recurrence relations, properties and integrals of the generalized exponential integral functions, Journal of Quantitative Spectroscopy and Radiative Transfer*, Volume 104, Issue 3, April 2007, Pages 310-325.

[3] Y. Aygar, E. Bairamov, *Properties of the two-dimensional exponential integral functions*, Journal of Mathematical Chemistry, May 2011, Volume 49, Issue 5, pp 1014-1025.

[4] E. Bairamov, N. Ozalp, *On convergence, asymptotic behaviour and computational algorithm of generalized exponential integral functions arising in a radiative transfer*, Journal of Quantitative Spectroscopy and Radiative Transfer, Volume 113, Issue 14, September 2012, Pages 1818-1825.

[5] E. Bairamov, S. Yardimci, *Uniform convergence and the properties of the exponential and generalized exponential integral functions*, Journal of Quantitative Spectroscopy and Radiative Transfer, Volume 111, Issue 16, November 2010, Pages 2471-2473.

[6] W.F. Breig, A.L. Crosbie, *Two-dimensional radiative equilibrium semi-infinite medium subjected to cosine varying radiation.* J. Quantitative Spectroscopy & Radiative Transfer 13 (1973) 1395–1419.

[7] W.F. Breig, A.L. Crosbie, Numerical computation of a generalized exponential integral function. Math. Comput. 28 (1974) 575–579.

[8] I.M. Busbridge, *The mathematics of radiative transfer.* Cambridge 1960.

[9] S. Chandrasekhar, *Radiative Transfer*, Dover Publications, New York, 1960.

[10] C. Chiccoli, S. Lorenzutta, G. Maino; *Concerning some integrals of the generalized exponential-integral function. Computers & Mathematics with Applications*, Volume 23, Issue 11, June 1992, Pages 13-21.

[11] A.L. Crosbie, R.L. Dougherty, *Two-dimensional in a cylindrical geometry with anisotropic scattering.* J. Quantitative Spectroscopy & Radiative Transfer, 25 (1981) 551–569.

[12] A.L. Crosbie, L.C. Lee, *Relation between multidimensional radiative transfer in cylindrical and rectangular coordinates with anisotropic scattering.* J. Quantitative Spectroscopy & Radiative Transfer, 38 (1987) 231–241.

[13] A.L. Crosbie, L.C. Lee, *Multidimensional radiative transfer: a single integral represantation of anisotropic scattering kernels J. Quantitative Spectroscopy & Radiative Transfer*, 42 (1989) 239–246.

[14] I.S. Gradshteyn, I.M. Ryzhik, *Table of integrals, series and products* Academic Press, NewYork, 2007.

[15] I.I. Guseinov, B.A. Mamedov, *On the evalution of generalized exponential integral functions*, J. Quantitative Spectroscopy & Radiative Transfer, 102 (2006) 251-256.

[16] V. Kourganoff, *Basic Methods in Transfer Problems, Dover Publications*, New York, 1963.

[17] B.A. Mamedov; *Evaluation of generalized exponential integral function using binomial expansion theorems*, Journal of Quantitative Spectroscopy and Radiative Transfer, Volume 94, Issues 3-4, 1 September 2005, Pages 507-514.

[18] S.M. Nikolsky,. *Course of Mathematical Analysis.* Vol. 2, Mir Publishers, Moscow 1977.

[19] R.H. Prabha, R.D.S. Yadav, *Polynomial expressions for Bickley and exponential integral functions.* Ann. Nucl. Energy, 23 (1996) 1021–1025.

[20] V.V. Sobelev, *A Treatise on Radiative Transfer Princeton*, New York, 1963

[21] H.C. Van de Hulst, Multiple Light Scattering, Vol. 1, Academic Press, New York, 1980.

[22] W.W Yuen, C.F. Ho, *Analysis of two-dimensional radiative heat transfer in a gray medium with internal heat generation*, Int. J. Heat Mass Transfer, 28 (1985) 17-23.

[23] V.A. Zorich, *Mathematical Analysis.* Vol. 2, Springer, New York 2004.

STATISTICS

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 497 – 519

Optimal stop-loss reinsurance: a dependence analysis

Anna Castañer* and M. Mercè Claramunt †‡

Abstract

The stop-loss reinsurance is one of the most important reinsurance contracts in the insurance market. From the insurer point of view, it presents an interesting property: it is optimal if the criterion of minimizing the variance of the cost of the insurer is used. The aim of the paper is to contribute to the analysis of the stop-loss contract in one period from the point of view of the insurer and the reinsurer. Firstly, the influence of the parameters of the reinsurance contract on the correlation coefficient between the cost of the insurer and the cost of the reinsurer is studied. Secondly, the optimal stop-loss contract is obtained if the criterion used is the maximization of the joint survival probability of the insurer and the reinsurer in one period.

Keywords: Stop-loss premium, Survival probabilities, Reinsurance.

2000 AMS Classification: 91B30, 60G40.

Received : 11.04.2014 *Accepted* : 22.02.2015 *Doi* : 10.15672/HJMS.2015539472

1. Introduction

An insurance company may decide to sign a reinsurance contract either to assume greater risks or to protect the company. This reinsurance contract transfers part of the risks assumed by the insurer to the reinsurer in exchange of giving also a part of the premiums received from policyholders. Yet, reinsurance is the most important decision that an insurance company has to consider in order to reduce its underwriting risk. Two large groups of reinsurance contracts can be distinguished: the proportional and the non-proportional reinsurance. The proportional reinsurance includes two kinds of reinsurance known as quota-share and surplus. In the former, all the risks are transferred

*Dept. Matemàtica Econòmica, Financera i Actuarial, Universitat de Barcelona, Av. Diagonal, 690, 08034 Barcelona, Spain, Email: acastaner@ub.edu
†Dept. Matemàtica Econòmica, Financera i Actuarial, Universitat de Barcelona, Av. Diagonal, 690, 08034 Barcelona, Spain, Email: mmclaramunt@ub.edu
‡Corresponding Author.

in the same proportion, while in the latter the proportion may vary. As regards the non-proportional reinsurance, the stop-loss and excess-loss contracts stand out. In both cases, the reinsurance offers protection when the aggregate claims exceed a certain agreed level.

The stop-loss reinsurance has been widely studied in the actuarial literature. The actuarial literature on the optimal reinsurance can be classified using as a criterion who takes the decision. A first group may include the works that focus on the insurer point of view and then try to maximize/minimize some measure of the risk of the insurer if he/she signs a reinsurance contract. In a second group we include the papers that search for strategies that are beneficial to the two parts that participate in the contract. Borch, [7], states that "These considerations should remind us that there are two parties to a reinsurance contract, and that this parties have conflicting interests. The optimal contract must then appear as a reasonable compromise between these interests. To me the most promising line of research seems to be the study of contracts, which in different ways can be said to be optimal from the point of view of both parties". Nonetheless, most of the papers on optimal reinsurance have only considered the insurer point of view.

In the first group, a secondary criterion for classification could be the function that is maximized or minimized. Within this literature we find, for instance, the maximization of the expected utility of the insurer's wealth after reinsurance, the minimization of the probability of ruin in the short and in the long run and the minimization of the variance of the retained risk ([46]). Borch [6] (reproduced in [8]) proved that the stop-loss contract is the "most efficient" contract because, for a given net premium, it maximizes the reduction of the variance in the claim distribution of the ceding company. Daykin et al. [17], Gajek et al. [24], and Kaluszka [38] follow a similar line of research.

Gerber [27] uses the expected profit of the insurer in one period as a measure of the profitability, and the adjustment coefficient (closely connected with the probability of ruin) as a measure of security. He assumes normality and considers excess-loss, stop-loss and proportional contracts. Van Wouwe et al. [47] determine the optimal level of excess-loss reinsurance in the case that the ultimate ruin probability is taken as stability criterion. The insurer's survival probability is also considered in [40] and [29], and more recently in [43], [44], [25], [39], [13] and [41]. Guerra and Centeno[28] obtain an optimal reinsurance policy by maximizing the insurer's expected utility.

Several authors have used other kind of measures to find the optimal strategy. Van Heerwaarden et al. [46] use, as optimality criterion, the minimization of the retained risk with respect to the stop-loss order. In turn, Hoøjgaard and Taksar [30] find the optimal proportional dynamic strategy that maximizes the return function of the insurer. And in [31], the previous analysis is extended to include transactions costs. Azcue and Muler [2] consider also a dynamic choice of both the reinsurance policy and the dividend distribution strategy that maximizes the cumulative expected discounted dividend payouts. In [9], [11], [45] and [15] the optimization of a reinsurance contract under the value-at-risk and conditional tail expectation risk measures is conducted. Similarly, Zhu et al. [50] investigate optimal reinsurance strategies for an insurer with multiple lines of business using the multivariate lower-orthant Value-at-Risk. Centeno and Simões [14] provide a good summary of the classical results on optimal reinsurance and a more detailed analysis of recent results (2000-2009).

Balbás et al. [4] use a general risk measure that includes every deviation measure, every expectation bounded risk measure, and most of the coherent, convex or consistent risk measures as particular cases. In [3] the previous analysis is extended to cases where the statistical distribution of claims is not totally known, generating uncertain or ambiguous frameworks. Following the modern studies about distortion risk measures, Cui et al. [16], Zheng and Cui [48], Zheng et al. [49] and Assa [1] use them to find the optimal reinsurance.

In spite of the above comment of Borch [7], the consideration of the interest of both the insurer and the reinsurer has not been really developed until recently. Borch [6] can be considered the first author in adopting this approach to the optimal reinsurance problem. He considers the minimization of the total variance risk for an stop-loss contract. Hürlimann [32] retakes this question and, in [33], obtains also optimal solutions under the total variance risk measure for a partial stop-loss contract.

Cai et al. [10] study the sufficient and necessary conditions for the existence of the optimal reinsurance retentions for the quota-share reinsurance and the stop-loss reinsurance under the expected value reinsurance premium principle, considering as objective function the joint survival probability and the joint profitable probability. For the joint survival probability, in [23], an extension for a combination of quota-share and stop-loss reinsurance contracts is found.

Ignatov et al. [34] and Kaishev and Dimitrova ([37] and [22]) use a different approach to joint optimality criteria. In [34] and [37], they find the parameters of the reinsurance contract that maximize the joint survival probability, when the premiums of the insurer and the reinsurer are fixed, for an excess of loss risk model when the number of claims follows a Poisson process. Salcedo-Sanz et al. [41] solve also this question using evolutionary and swarm intelligence techniques. In [37] the previous analysis is extended to include an optimal split of the premium income between the insurer and the reinsurer, given fixed retention and limiting levels. These two optimization problems are applied, with some numerical examples, to the stop-loss contract over a fixed horizon in [12]. Other optimal problems are added by Dimitrova and Kaishev [22], for an excess-loss, and by Castañer et al. [12], for an stop-loss. In [22], the authors propose a Markowitz type efficient frontier solution to the problem of optimally setting the parameters of reinsurance, so that for a given level of the probability of joint survival the expected profits of the two parties are maximized. Finally, in [12], the optimal split of the total initial reserves between the insurer and the reinsurer that maximizes the joint survival probability is considered.

The objective of this work is to contribute to the analysis of the optimal stop-loss reinsurance in one period, from the joint point of view of the insurer and the reinsurer. The contributions of this paper to the optimal reinsurance can be summarized as follows. First, using total variance risk measure, we add the analysis of the optimal reinsurance for an stop-loss contract with maximum, to the known solutions for the standard stop-loss ([6] and [32]) or the partial standard stop-loss ([33]). We also include the possibility of using the maximization of the correlation coefficient between the insurer's and reinsurer's losses. Second, we consider the maximization of the joint survival probability in one period in an stop-loss with and without maximum, and using the same hypothesis with respect to premiums as in Kaishev and Dimitrova ([37] and [22]), we obtain the optimal parameters of the stop-loss. In addition, in line with [37], we use as a criterion for the calculation of the reinsurer's premiums the maximization of the joint survival probability, given as fixed both the values of the parameters of the reinsurance contract and the initial values of the reserves of the insurer and the reinsurer. In fact, then, we propose a different way of calculating the stop-loss premium that considers not only the losses for the reinsurer, but all the other factors (loss and premium of the insurer and initial capitals of insurer and reinsurer). The solution of these two optimization problems related to the joint survival probability in one period for the stop-loss reinsurance are arguably the main findings of this paper.

The paper is organized as follows. Section 2 analyzes the expression of the covariance and the correlation coefficient and the specific expressions for different distributions of the total cost, considering a stop-loss reinsurance with priority d with and without a maximum m. In Section 3, we find the optimal reinsurance stop-loss if the criterion is the maximization of the variance reduction due to reinsurance. In Section 4, we introduce the probability of joint survival as a measure for the solvency for a reinsurance contract with priority d and reinsurance with d and m. In Section 5, the problem of finding the optimal reinsurance stop-loss if the criterion is the maximization of the joint survival probability is solved. In addition, a number of examples are presented. Section 6 closes the paper offering some final conclusions and remarks.

2. Covariance and correlation between the cost of the insurer and the cost of the reinsurer

In the stop-loss reinsurance contract with priority $d > 0$ the random variable (r.v.) total cost of claims in one period, S, is split between the cost of the insurer, SI, and the cost of the reinsurer, SR, with $S = SI + SR$, $SR = \max\{S - d, 0\}$ and $SI = min\{S, d\}$. The distribution functions of these two r.v., $F_{SI}(s) = P[SI \leq s]$ and $F_{SR}(s) = P[SR \leq s]$, can be calculated from the distribution function of S, $F_S(s) = P[S \leq s]$,

$$F_{SI}(s) = \begin{cases} F_S(s) & \text{if } s < d, \\ 1 & \text{if } s \geq d, \end{cases} \tag{2.1}$$

$$F_{SR}(s) = F_S(s + d). \tag{2.2}$$

The reinsurer can calculate the reinsurance premium with several premium principles. Most of these principles are based on the expectation of the total cost assumed by the reinsurer ([21]). For instance, the net premium principle establishes that the premium is equal to the expectation of the cost. In the actuarial literature, the premium of an stop-loss contract calculated with the net premium principle is called the stop-loss premium. Let us define $\pi(d) = E[SR]$ as the stop-loss premium in a reinsurance stop-loss contract with priority d.

The r.v. cost of the reinsurer SR has the following two ordinary moments[§]:

$$\alpha_1(SR) = E[SR] = \int_d^\infty (s-d) f_S(s) ds = \int_d^\infty (1 - F_S(s)) ds, \tag{2.3}$$

$$\alpha_2(SR) = \int_d^\infty (s-d)^2 f_S(s) ds = 2\int_d^\infty (s-d)(1 - F_S(s)) ds.$$

Hence, the variance is

$$V[SR] = \alpha_2(SR) - \alpha_1^2(SR) = E[SR](-2d - E[SR]) + 2\int_d^\infty s(1 - F_S(s)) ds. \tag{2.4}$$

The expectation and the variance of the insurer cost SI can be calculated from those of S and SR, so:

$$\alpha_1(SI) = E[SI] = E[min(S, d)] = E[S] - E[SR],$$

$$V[SI] = V[S] - V[SR] - 2Cov[SI, SR],$$

[§]In order to obtain the expressions for the first two moments of the cost of the reinsurer, it is necessary to take into account that $-f_S(s)ds = d(1 - F_S(s))$ and then apply integration by parts.

being

$$\begin{aligned} Cov\left[SI,SR\right] &= \int_d^\infty d(s-d)f_S(s)ds - E\left[SR\right]\left(E\left[S\right]-E\left[SR\right]\right) \\ &= E[SR]\left(d-E\left[S\right]+E[SR]\right). \end{aligned} \tag{2.5}$$

The correlation coefficient between SI and SR is

$$r(SI,SR) = \frac{Cov\left[SI,SR\right]}{\sqrt{V\left[SR\right]\left(V\left[S\right]-V\left[SR\right]-2Cov\left[SI,SR\right]\right)}}. \tag{2.6}$$

In addition to the marginal analysis of the cost of the insurer and the reinsurer, we are interested in the bivariate r.v. (SI, SR). In a stop-loss reinsurance contract with priority d, the joint distribution function of the costs of the insurer and the reinsurer in one period is

$$P\left[SI \le x, SR \le y\right] = \begin{cases} P\left[S \le x\right] & \text{if } x < d, \\ P\left[S \le y+d\right] & \text{if } x \ge d > 0. \end{cases} \tag{2.7}$$

This r.v. (SI, SR) is comonotone ([20]) because SI and SR are increasing functions of the risk S. Then, there is a perfect positive dependence between the two marginal r.v. SI and SR and it is granted that the two parts that participate in the exchange of risk (the insurer and the reinsurer) increase their cost when the underlying risk increases. Hence, the correlation coefficient between SI and SR is the maximal one that can be attained between two random variables with the same marginal distributions, but it is not equal to one (this would be the case if one variable could be calculated as a linear function of the other, e.g. in proportional reinsurance)([18]). So, for a fixed d, $r(SI, SR)$ is the maximal one, but it is less than one in absolute value.

The stop-loss reinsurance contract can include a priority d and a maximum m, $m > d > 0$. In this case,

$$SR(d,m) = \min\left\{m-d, \max\left\{S-d, 0\right\}\right\},$$

$$SI(d,m) = \min\left\{S, d\right\} + \max\left\{S-m, 0\right\}.$$

The distribution functions of these two r.v. are

$$F_{SI(d,m)}(s) = \begin{cases} F_S(s) & \text{if } s < d, \\ F_S(s+m-d) & \text{if } s \ge d \end{cases} \tag{2.8}$$

and

$$F_{SR(d,m)}(s) = \begin{cases} F_S(s+d) & \text{if } s < m-d, \\ 1 & \text{if } s \ge m-d. \end{cases} \tag{2.9}$$

Let $\pi(d,m) = E[SR(d,m)]$ be the stop-loss premium, that is the reinsurance premium calculated with the net premium principle. It can be calculated from the premiums of a stop-loss reinsurance with priorities d and m, $\pi(d,m) = \pi(d) - \pi(m)$.

The second ordinary moment $\alpha_2(SR(d,m))$, is

$$\begin{aligned}\alpha_2(SR(d,m)) &= \int_d^m (s-d)^2 f_S(s)ds + \int_m^\infty (m-d)^2 f_S(s)ds \\ &= \int_d^\infty (s-d)^2 f_S(s)ds - \int_m^\infty (s-d)^2 f_S(s)ds + \int_m^\infty (m-d)^2 f_S(s)ds \\ &= \alpha_2(SR(d)) - \int_m^\infty ((s-d)^2-(m-d)^2) f_S(s)ds \\ &= \alpha_2(SR(d)) - \alpha_2(SR(m)) - 2(m-d)\pi(m),\end{aligned}$$

where the last equality follows taking into account that $(s-d)^2-(m-d)^2 = (s-m)^2 + 2(s-m)(m-d)$.
Hence, the variance $V[SR(d,m)]$, is:

$$\begin{aligned}V[SR(d,m)] &= \alpha_2(SR(d,m)) - \alpha_1(SR(d,m))^2 \\ &= \alpha_2(SR(d)) - \alpha_2(SR(m)) - 2(m-d)\pi(m) - (\pi(d)-\pi(m))^2 \\ &= V[SR(d)] - V[SR(m)] + 2\pi(m)(\pi(d)+d-\pi(m)-m).\end{aligned}$$

The covariance between the costs of the insurer and the reinsurer is:

$$\begin{aligned}Cov[SI(d,m),SR(d,m)] &= \int_d^m d(s-d)f_S(s)ds + \int_m^\infty (m-d)(s-m+d)f_S(s)ds \\ &= \int_d^\infty d(s-d)f_S(s)ds \\ &- \int_m^\infty (d(s-d)-(m-d)(s-m+d))f_S(s)ds \\ &= Cov[SI(d),SR(d)] - \int_m^\infty ((s-m)(2d-m))f_S(s)ds \\ &= Cov[SI(d),SR(d)] - (2d-m)\pi(m), \qquad (2.10)\end{aligned}$$

where the last but one equality follows taking into account that $d(s-d)-(m-d)(s-m+d) = (s-m)(2d-m)$.
So, in order to calculate the expectation and the variance of the costs of the insurer and the reinsurer, and the covariance if the stop-loss has a maximum, we only need the expressions of a stop-loss without maximum.

The distribution function of the bivariate r.v. $(SI(d,m),SR(d,m))$ is

$$P[SI(d,m)\le x, SR(d,m)\le y] = \begin{cases} P[S\le x] & \text{if } x<d, \\ P[S\le d] & \text{if } x\ge d \text{ and } y=0, \\ P[S\le y+d] & \text{if } x\ge d \text{ and } 0<y<m-d, \\ P[S\le m] & \text{if } x=d \text{ and } y\ge m-d, \\ P[S\le x+m-d] & \text{if } x>d \text{ and } y\ge m-d. \end{cases} \qquad (2.11)$$

Throughout the paper, we use three approximations for the total cost in a period: gamma with two parameters, translated gamma and normal. The gamma distribution deserves special attention. It has been used in its version of two or three parameters to approximate the distribution of the total cost in a period as an alternative to the exact calculation through convolutions and to other approximations. In several papers ([5], [42], [26]), the accuracy of the translated gamma approximation and the rest of approximations has been quantified. In this sense, [35] uses the translated gamma approximation for the calculation of the stop-loss premium. In order to be self contained and to clarify the formulas that we use, we include in Section 2.1 a summary of the (translated) gamma distribution. Next, we indicate the explicit expressions of $\pi(d)$, $Cov[SI,SR]$ and $V[SR]$, which allow us calculating the coefficient of correlation for three different distributions or approximations for the total cost in a period: gamma with two parameters, translated gamma and normal. As it is a simple calculation, we do not include the processes for obtaining these expressions.

2.1. Statistical summary. The gamma distribution with three parameters (or Pearson Type III) is also known as the translated gamma distribution, with one of its parameters interpreted as follows. If $X \sim Ga(\alpha, \beta, \gamma)$, with $\alpha > 0$, $\beta > 0$ and $\gamma \in \Re$, its density function is

$$f_X(x) = \frac{(x-\gamma)^{\alpha-1} e^{\frac{-(x-\gamma)}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)}, x > \gamma, \tag{2.12}$$

being γ, precisely, the parameter of translation. If $\gamma = 0$, the gamma distribution with two parameters is obtained, $X \sim Ga(\alpha, \beta)$ with $\alpha > 0$ and $\beta > 0$. The standard form of the distribution is obtained if, in addition, $\beta = 1$. Then, $X \sim Ga(\alpha)$, with $\alpha > 0$.

The gamma distribution with three parameters can be calculated through a gamma distribution with two or with one parameter (the standard form). Let $X \sim Ga(\alpha, \beta, \gamma)$, if $Y = (X - \gamma)/\beta$, then, $Y \sim Ga(\alpha)$, and also, $X = Y\beta + \gamma$. If $Z = X - \gamma$, then, $Z \sim Ga(\alpha, \beta)$, and the next relations are met,

$$X = Z + \gamma, Y = \frac{Z}{\beta}.$$

Recall that the moments and measures of X, Y and Z, are related as shown in Table 1.

Table 1. Some characteristics of the gamma distribution

	$Y \sim Ga(\alpha)$	$Z \sim Ga(\alpha, \beta)$	$X \sim Ga(\alpha, \beta, \gamma)$
Mean μ_1	α	$\alpha\beta$	$\alpha\beta + \gamma$
Variance μ_2	α	$\alpha\beta^2$	$\alpha\beta^2$
μ_3	2α	$2\alpha\beta^3$	$2\alpha\beta^3$
Skewness γ_1	$\frac{2}{\sqrt{\alpha}}$	$\frac{2}{\sqrt{\alpha}}$	$\frac{2}{\sqrt{\alpha}}$

The parameters of $X \sim Ga(\alpha, \beta, \gamma)$, can be estimated by the moments' method:

$$\widehat{\alpha} = \frac{4}{\gamma_1^2(X)}, \widehat{\beta} = \frac{\mu_3(X)}{2\mu_2(X)}, \widehat{\gamma} = E[X] - \widehat{\alpha}\widehat{\beta}. \tag{2.13}$$

Taking into account Table 1, a variable $X \sim Ga(\alpha, \beta, \gamma)$, also meets the next relationship with the variable $Y \sim Ga(\alpha)$ (if the parameter α is estimated through the asymmetry of X, as in (2.13)),

$$X = \mu_1(X) + \mu_2^{0.5}(X)\frac{Y - \alpha}{\sqrt{\alpha}}.$$

Then,

$$\begin{aligned} P[X \le x] &= P\left[\mu_1(X) + \mu_2^{0.5}(X)\frac{Y - \alpha}{\sqrt{\alpha}} \le x\right] \\ &= P\left[Y \le \alpha + \sqrt{\alpha}\frac{x - \mu_1(X)}{\mu_2^{0.5}(X)}\right] \\ &= Ga\left(\alpha + \sqrt{\alpha}\frac{x - \mu_1(X)}{\mu_2^{0.5}(X)}; \alpha\right), \end{aligned} \tag{2.14}$$

being $Ga(y; \alpha) = P[Y \le y]$ with $Y \sim Ga(\alpha)$. Or alternatively,

$$\begin{aligned} P[X \le x] &= P[Z + \gamma \le x] = P[Z \le x - \gamma] \\ &= Ga(x - \gamma; \alpha, \beta), \end{aligned} \tag{2.15}$$

being $Ga(z;\alpha,\beta) = P[Z \leq z]$ with $Z \sim Ga(\alpha,\beta)$.

2.2. Gamma distribution (with two parameters). Assume $S \sim Ga(\alpha,\beta)$, with $\alpha > 0$ and $\beta > 0$. The density function and the distribution function are, respectively,

$$\begin{aligned} f_S(s) &= \frac{s^{\alpha-1}e^{-\frac{s}{\beta}}}{\beta^\alpha\Gamma(\alpha)}, \quad s > 0, \\ F_S(s) &= Ga(s;\alpha,\beta), \quad s > 0. \end{aligned}$$

Hence, in this case we have

$$\pi(d) = \alpha\beta\left(1 - Ga(d;\alpha+1,\beta)\right) - d\left(1 - Ga(d;\alpha,\beta)\right),$$

$$\begin{aligned} Cov\left[SI,SR\right] &= \left[\alpha\beta\left(1 - Ga(d;\alpha+1,\beta)\right) - d\left(1 - Ga(d;\alpha,\beta)\right)\right] \\ &\times \left[-\alpha\beta Ga(d;\alpha+1,\beta) + d\ Ga(d;\alpha,\beta)\right] \end{aligned}$$

and

$$\begin{aligned} V\left[SR\right] &= \pi(d)\left(-2d - \pi(d)\right) - d^2\left(1 - Ga(d;\alpha,\beta)\right) \\ &+ (\alpha+1)\alpha\beta^2\left(1 - Ga(d;\alpha+2,\beta)\right). \end{aligned}$$

2.3. Translated gamma distribution. Assume $S \sim Ga(\alpha,\beta,\gamma)$, with $\alpha > 0$, $\beta > 0$ and $\gamma \in \Re$. The density function and the distribution function are, respectively,

$$\begin{aligned} f_S(s) &= \frac{(s-\gamma)^{\alpha-1}e^{-\frac{s-\gamma}{\beta}}}{\beta^\alpha\Gamma(\alpha)}, \quad s > \gamma, \\ F_S(s) &= Ga(s;\alpha,\beta,\gamma), \quad s > \gamma. \end{aligned}$$

For the translated gamma approximation for the distribution of the total cost, we obtain two equivalent expressions for the stop-loss premium depending on the formula used, (2.14) or (2.15). First, from (2.14) we have,

$$\pi(d) = E\left[(S-d)_+\right] \approx \frac{\mu_2^{0.5}(S)}{\sqrt{\alpha}}\left[d' f(d';\alpha) + (\alpha - d')(1 - Ga(d';\alpha))\right], \tag{2.16}$$

being $d' = \alpha + \sqrt{\alpha}\left(\frac{d-\mu_1(S)}{\mu_2^{0.5}(S)}\right)$ and $f(d';\alpha)$, the density function of $Y \sim Ga(\alpha)$ in d'. Second, from (2.15) we have,

$$\begin{aligned} \pi(d) &= E\left[(S-d)_+\right] \approx \alpha\beta\left(1 - Ga(d-\gamma;\alpha+1,\beta)\right) \\ &- (d-\gamma)(1 - Ga(d-\gamma;\alpha,\beta)), \end{aligned} \tag{2.17}$$

Expression (2.16) can be found in [35] as a particular case of the ordinary moments of the cost of the reinsurer.

From (2.4), (2.5) and (2.17) the $Cov[SI,SR]$ can be easily calculated, and the expression of the variance of SR is

$$\begin{aligned} V[SR] &= \pi(d)\left(-2d - \pi(d)\right) + 2\alpha\beta\gamma\left(1 - Ga(d-\gamma;\alpha+1,\beta)\right) \\ &+ (\alpha+1)\alpha\beta^2\left(1 - Ga(d-\gamma;\alpha+2,\beta)\right) + \left(\gamma^2 - d^2\right)\left(1 - Ga(d-\gamma;\alpha,\beta)\right). \end{aligned}$$

2.4. Normal distribution. Assume $S \sim N(\mu,\sigma)$, with $\mu = E[S]$ and $\sigma^2 = V[S] > 0$. The density and distribution functions are, respectively, in terms of the distribution of $N(0,1)$,

$$\begin{aligned} f_S(s) &= \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(s-\mu)^2}{2\sigma^2}}, \\ F_S(s) &= \Phi\left(\frac{s-\mu}{\sigma}\right), \end{aligned}$$

and then,

$$\pi(d) = \sigma\phi\left(\frac{d-\mu}{\sigma}\right) + (\mu - d)\left(1 - \Phi\left(\frac{d-\mu}{\sigma}\right)\right),$$

$$\begin{aligned} Cov\,[SI, SR] \quad &= \quad \left[\sigma\phi\left(\frac{d-\mu}{\sigma}\right) + (\mu - d)\left(1 - \Phi\left(\frac{d-\mu}{\sigma}\right)\right)\right] \\ &\times \quad \left[\sigma\phi\left(\frac{d-\mu}{\sigma}\right) - (\mu - d)\left(\Phi\left(\frac{d-\mu}{\sigma}\right)\right)\right] \end{aligned}$$

and

$$V\,[SR] = -\sigma(d-\mu)\phi\left(\frac{d-\mu}{\sigma}\right) - \pi(d)^2 + \left((\mu - d)^2 + \sigma^2\right)\left(1 - \Phi\left(\frac{d-\mu}{\sigma}\right)\right).$$

3. Optimal reinsurance if the criterion is the maximization of the variance reduction

Following [8], we can choose as measures of risk the variance or the probability of ruin. These two measures have different properties, and correspond to a different idea. If we use the variance we exclusively focus on the randomness of the cost of claims and we disregard the premiums (and then the security loadings) and the initial reserves of the insurer and the reinsurer. These two factors can be also taken into account if we use the probability of ruin as an alternative risk measure.

If our objective is to find an optimal contract, we can not only rely on the insurer's risk measure. We have to keep in mind Borch's statement that there are two parties to a reinsurance contract, and that these parties have conflicting interests. The optimal contract must then appear as a reasonable compromise between the interest of the insurer and the reinsurer and thus, it has to be found undertaking a joint analysis of this two parties.

In this section we perform a first analysis of the optimal reinsurance choosing the variance as measure of risk and maximizing the variance reduction defined as the difference between the variance of the loss and the sum of the variance of the insurer and the reinsurer, $V[S] - (V[SI(d)] + V[SR(d)])$ or $V[S] - (V[SI(d,m)] + V[SR(d,m)])$ if the contract includes a maximum m. By definition, in the first case this difference equals $2Cov[SI(d), SR(d)]$, and in the second case equals $2Cov[SI(d,m), SR(d,m)]$. Then, we choose the reinsurance parameters as those that maximize the covariance between the costs of the insurer and the reinsurer.

We consider first a stop-loss reinsurance contract with priority $d > 0$. The maximization program, from (2.5), is

$$\max_{d}\ \pi(d)\left(d - E[S] + \pi(d)\right) \text{ subject to } 0 < d. \tag{3.1}$$

As the covariance is a continuous function of d and the limits when d tends to 0 and to infinity are zero, the covariance has a maximum for at least one finite, positive value of d ([8]).

3.1. Proposition. The optimal point of program (3.1) is a value of d such that the following conditions are fulfilled:

$$\begin{aligned} \pi(d)\left(2F_S(d) - 1\right) + \left(d - E[S]\right)\left(F_S(d) - 1\right) = 0,& \\ \pi(d) < \frac{2F_S(d)\left(1 - F_S(d)\right)^2}{f_S(d)},& \\ d > 0.& \end{aligned}$$

Proof. The first order condition of optimality is

$$[\pi(d)\,(d - E[S] + \pi(d))]' = 0.$$

Considering that $\pi(d) = \int_d^\infty (1 - F_S(s))\,ds$, this condition is

$$\begin{aligned}\pi(d)F_S(d) &+ (F_S(d) - 1)\,(d - E[S] + \pi(d)) \\ &= \pi(d)\,(2F_S(d) - 1) + (d - E[S])\,(F_S(d) - 1) = 0.\end{aligned} \tag{3.2}$$

The second order condition for the maximization is

$$\begin{aligned}[\pi(d)\,(d - E[S] + \pi(d))]'' = f_S(d)\,(d - E[S] + \pi(d)) \\ + (F_S(d) - 1)\,2f_S(d) < 0.\end{aligned} \tag{3.3}$$

Isolating $(d - E[S] + \pi(d))$ from (3.2) and substituting in (3.3), the condition

$$\pi(d) < \frac{2F_S(d)\,(1 - F_S(d))^2}{f_S(d)}$$

is obtained. □

3.2. Corollary. If S has a symmetric density function, $d = E[S]$ is the only finite point that fulfils condition (3.2).

Proof. If S has a symmetric density function, $F_S(E[S]) = 0.5$, then (3.2) is fulfilled if and only if $d = E[S]$. □

Note: An equivalent expression to (3.2) can be found in Borch (1974) as well as the value of d that fulfils this condition when S follows an exponential distribution.

As an alternative, instead of maximizing the variance reduction in absolute value, we could apply the criterion of maximizing the coefficient of correlation. In this case, the conditions that must fulfil the optimal point are complex but easy to obtain. In order to be concise we only include in the paper (without proof) the necessary conditions.

3.3. Proposition. The optimal point of program

$$\max_d r(SI, SR) = \frac{Cov\,[SI, SR]}{\sqrt{V\,[SR]\,V\,[SI]}} \text{ subject to } 0 < d \tag{3.4}$$

fulfil the necessary condition

$$2\frac{Cov[SI, SR]'}{Cov[SI, SR]} = \frac{V[SI]'}{V[SI]} + \frac{V[SR]'}{V[SR]},$$

being

$$\begin{aligned}Cov[SI, SR]' &= \pi(d)\,(2F_S(d) - 1) + (d - E[S])\,(F_S(d) - 1)\,, \\ V[SI]' &= -2\,(F_S(d) - 1)\,(d - E[S] + \pi(d))\,, \\ V[SR]' &= -2\pi(d)F_S(d).\end{aligned}$$

3.4. Example. We assume that the total cost of a period has the following characteristics: $E\,[S] = 1$, $V\,[S] = 2$ and skewness $\gamma_1(S) = \frac{3}{\sqrt{2}}$. In Table 2 we show the maximum points and the maximum values of the covariance and the coefficient of correlation that are obtained using the gamma, the translated gamma and the normal approximations.

Table 2. Optimal points and maximum values of covariance and correlation coefficient. Stop-loss contract

	d^*	$Cov[SI(d^*), SR(d^*)]$	d^*	$r[SI(d^*), SR(d^*)]$
gamma	2.19654	0.326122	1.3598	0.499926
translated gamma	1.89158	0.324196	1.27352	0.490588
normal	1	0.31831	1	0.466942

If the stop-loss reinsurance contract has also a maximum m, the maximization program, from (2.10) is

$$\max_{d,m} \pi(d)\,(d - E[S] + \pi(d)) - (2d - m)\,\pi(m) \text{ subject to } 0 < d < m \tag{3.5}$$

3.5. Proposition. The optimal point of program (3.5) is a value of $(d, m) \in \Re^2_+$ such that the following conditions are fulfilled:

$$\begin{aligned} \pi(d)F_S(d) + (F_S(d) - 1)\,(d - E[S] + \pi(d)) - 2\pi(m) = 0, \\ \pi(m) - (2d - m)\,(F_S(d) - 1) = 0, \\ f_S(d)\,(2\pi(d) + d - E[S]) + 2F_S(d)\,(F_S(d) - 1) < 0, \end{aligned}$$

$$\begin{aligned} f_S(d)\,(2\pi(d) + d - E[S]) + 2F_S(d)\,(F_S(d) - 1) < \frac{4(F_S(m) - 1)^2}{mf_S(m) + 2\,(F_S(m) - 1)}, \\ 0 < d < m. \end{aligned}$$

Proof. The first order condition of optimality is

$$\begin{cases} \frac{\partial[\pi(d)(d-E[S]+\pi(d))-(2d-m)\pi(m)]}{\partial d} = 0, \\ \frac{\partial[\pi(d)(d-E[S]+\pi(d))-(2d-m)\pi(m)]}{\partial m} = 0. \end{cases}$$

Considering that $\pi(d) = \int_d^\infty (1 - F_S(s))\,ds$, this condition is

$$\begin{cases} \pi(d)F_S(d) + (F_S(d) - 1)\,(d - E[S] + \pi(d)) - 2\pi(m) = 0, \\ \pi(m) - (2d - m)\,(F_S(d) - 1) = 0. \end{cases}$$

The inequalities are obtained applying the second order condition for a maximum. □

Using the values of Example 3.4, we numerically show that there is no solution of the program (3.5), although with the normal distribution the point $(d, m) = (2.38, 3.63)$ is a local optimum that fulfils the conditions included in Proposition 3.5. For illustration, in Figure 1, the covariance for the normal distribution is plotted.

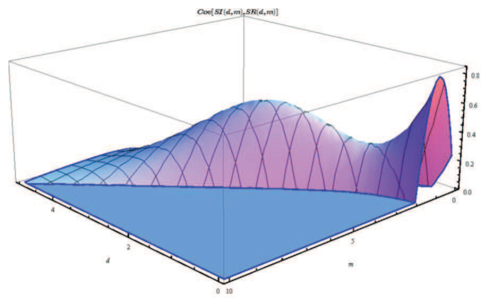


Figure 1. Covariance between the costs of the insurer and the reinsurer with a normal distribution

4. Survival probabilities in one period

The survival probability is one of the most important measures of the solvency of an insurer/reinsurer. The survival probability in one period of an insurer considering only the underwriting risk, can be calculated knowing the distribution of the cost of the insurer, the reserves at the beginning of the period and the premium earned by the insurer to cover the insured risk. If a stop-loss reinsurance contract is agreed, the survival probability of the insurer is obviously different and needs to be calculated again with the new parameters; but, as in this case, if the payment of the claims depends on the two parts, the joint survival probability of insurer and reinsurer is also a quantity of interest.

Let $PT > 0$ be the premium earned by the insurer in the period; let $PR > 0$ be the reinsurer's premium; let $uI \geq 0$ and $uR \geq 0$ be the initial reserves of the insurer and the reinsurer, respectively. It is then possible to incorporate in the model an economic constraint: the reinsurer's premium must be less than the premium earned by the insurer in the period, $0 < PR < PT$.

The survival probability is in fact a particular case of a family of probabilities regarding the technical result at the end of the period. Let $\varphi(u, P, \alpha)$ be the probability that the technical result (initial capital (u) plus earned premiums (P) minus aggregated claims (S)) of an insurer is greater or equal to α,

$$\varphi(u, P, \alpha) = P\left[u + P - S \geq \alpha\right].$$

The technical result has a natural maximum value, $u+P$, that is attained when no claims occur during the period. As $\varphi(u, P, \alpha) = 0$ for $\alpha > u + P$, and $\varphi(u, P, \alpha) = 1$ for $\alpha < 0$, we can consider that $0 \leq \alpha \leq u + P$.

Survival probability $\phi(\cdot)$ is a particular case of $\varphi(\cdot)$ that is obtained considering $\alpha = 0$.

4.1. Stop-loss reinsurance with priority d. Probabilities regarding the technical result of the insurer, $\varphi_I(uI, d, PR, PT, \alpha)$, are

$$\begin{aligned} \varphi_I(uI, d, PR, PT, \alpha) &= P\left[uI + PT - PR - SI \geq \alpha\right] \\ &= P\left[SI \leq uI + PT - PR - \alpha\right] \\ &= F_{SI}(uI + PT - PR - \alpha) \end{aligned}$$

and from (2.1),

$$\varphi_I(uI, d, PR, PT, \alpha) = \begin{cases} F_S(uI + PT - PR - \alpha) & \text{if } uI + PT - PR - \alpha < d, \\ 1 & \text{if } uI + PT - PR - \alpha \geq d. \end{cases} \tag{4.1}$$

The probabilities regarding the technical result of the reinsurer, $\phi_R(uR, d, PR, \alpha)$, are

$$\begin{aligned}\varphi_R(uR, d, PR, \alpha) &= P\,[uR + PR - SR - \alpha \geq 0] = P\,[SR \leq uR + PR - \alpha] \\ &= F_{SR}(uR + PR - \alpha)\end{aligned}$$

and from (2.2),

$$\varphi_R(uR, d, PR, \alpha) = F_S(uR + PR + d - \alpha). \tag{4.2}$$

The joint probabilities regarding the technical result of both the insurer and the reinsurer, $\varphi_{I,R}(uI, uR, d, PR, PT, \alpha_1, \alpha_2)$, are

$$\begin{aligned}&\varphi_{I,R}(uI, uR, d, PR, PT, \alpha_1, \alpha_2) \\ &\quad = P\,[SI \leq uI + PT - PR - \alpha_1, SR \leq uR + PR - \alpha_2]\end{aligned}$$

and from (2.7),

$$\begin{aligned}&\varphi_{I,R}(uI, uR, d, PR, PT, \alpha_1, \alpha_2) \\ &\quad = \begin{cases} F_S(uI + PT - PR - \alpha_1) & \text{if } uI + PT - PR - \alpha_1 < d, \\ F_S(uR + PR + d - \alpha_2) & \text{if } uI + PT - PR - \alpha_1 \geq d. \end{cases}\end{aligned} \tag{4.3}$$

The joint survival probability of the insurer and the reinsurer $\phi_{I,R}(uI, uR, d, PR, PT)$ is obtained when both α_1 and α_2 are equal to zero,

$$\phi_{I,R}(uI, uR, d, PR, PT) = \varphi_{I,R}(uI, uR, d, PR, PT, 0, 0).$$

4.2. Stop-loss reinsurance with priority d and maximum m. The joint probabilities of the insurer, $\varphi_I(uI, d, m, PR, PT, \alpha)$, are

$$\varphi_I(uI, d, m, PR, PT, \alpha) = F_{SI(d,m)}(uI + PT - PR - \alpha)$$

and from (2.8)

$$\begin{aligned}&\varphi_I(uI, d, m, PR, PT, \alpha) \\ &\quad = \begin{cases} F_S(uI + PT - PR - \alpha) & \text{if } uI + PT - PR - \alpha < d, \\ F_S(uI + PT - PR - \alpha + m - d) & \text{if } uI + PT - PR - \alpha \geq d. \end{cases}\end{aligned}$$

The joint probabilities of the reinsurer, $\varphi_R(uR, d, m, PR, \alpha)$, are

$$\varphi_R(uR, d, m, PR, \alpha) = F_{SR(d,m)}(uR + PR - \alpha)$$

and from (2.9)

$$\begin{aligned}&\varphi_R(uR, d, m, PR, \alpha) \\ &\quad = \begin{cases} F_S(uR + PR + d - \alpha) & \text{if } uR + PR - \alpha < m - d, \\ 1 & \text{if } uR + PR - \alpha \geq m - d. \end{cases}\end{aligned}$$

The joint probabilities of the insurer and the reinsurer are

$$\begin{aligned}&\varphi_{I,R}(uI, uR, d, m, PR, PT, \alpha_1, \alpha_2) \\ &\quad = P\,[SI \leq uI + PT - PR - \alpha_1, SR \leq uR + PR - \alpha_2]\end{aligned}$$

and from (2.11)

(4.4) $\varphi_{I,R}(uI, uR, d, m, PR, PT, \alpha_1, \alpha_2)$

$$=\begin{cases} F_S(uI+PT-PR-\alpha_1) & \text{if } uI+PT-PR-\alpha_1<d,\\ F_S(d) & \text{if } uI+PT-PR-\alpha_1\geq d \text{ and } uR+PR-\alpha_2=0,\\ F_S(uR+PR+d-\alpha_2) & \text{if } uI+PT-PR-\alpha_1\geq d \text{ and } 0<uR+PR-\alpha_2<m-d,\\ F_S(m) & \text{if } uI+PT-PR-\alpha_1=d \text{ and } uR+PR-\alpha_2\geq m-d,\\ F_S(uI+PT-PR-\alpha_1+m-d) & \text{if } uI+PT-PR-\alpha_1>d \text{ and } uR+PR-\alpha_2\geq m-d. \end{cases}$$

The joint survival probability of the insurer and the reinsurer is obtained when both α_1 and α_2 are equal to zero,

$$\phi_{I,R}(uI, uR, d, m, PR, PT) = \varphi_{I,R}(uI, uR, d, m, PR, PT, 0, 0).$$

5. Optimal joint survival probability in one period

In this section, we are interested in solving two different optimization problems related with the joint survival probability of the insurer and the reinsurer in one period.

In the first optimization problem, the reinsurance premium is fixed (as it is the total premium PT) and so are the initial values of the reserves of the insurer and the reinsurer. In addition, the parameters of the reinsurance maximize the joint survival probability. This probability is a function of the parameters of the reinsurance, d or d and m. Propositions 5.1 and 5.7 solve this problem. In this case, the insurer has a fixed amount of money available to purchase the reinsurance protection and we look for the most efficient stop-loss contract since it offers the lowest risk (measured by the joint probability of ruin) for this given value of the reinsurer premium. This idea of finding the parameters of the reinsurance contract that maximize the joint survival probability when the premiums of the insurer and the reinsurer are fixed, can also be found in [37] and [22], where the authors consider an excess of loss risk model when the number of claims follows a Poisson process. The assumptions of our model are totally different but, in Proposition 5.1 and 5.7 we consider the same maximization problem.

It is usually considered that PR is a function of the parameters of the stop-loss reinsurance (d, m) and the total cost S. In that instance, the reinsurer would apply for the calculation of the premium some of the usual criteria, for instance, the expected value, variance and standard deviation principles (for more details see [36]). We adopt as a criterion for the calculation of the reinsurer's premiums the maximization of the joint survival probability, given as fixed both the values of the parameters of the reinsurance contract and the initial values of the reserves of the insurer and the reinsurer. Then, in the second optimization problem, the joint survival probability is considered to be a function of the reinsurance premium, PR. Propositions 5.5 and 5.9 tackle this problem.

5.1. Proposition. In a stop-loss reinsurance with priority d, the program

$$\max_d \phi_{I,R}(uI, uR, d, PR, PT) \text{ subject to } 0 < d$$

has as a maximum value $\phi^*_{I,R}(uI, uR, PR, PT) = F_S\,(uI + uR + PT)$, being the optimal point $d^*(uI, uR, PR, PT) = uI + PT - PR$.

Proof. The joint survival probability to be maximized, (4.3), is a step function built with the distribution function of the total cost. Since $F_S(x)$ is increasing in x and $uI + PT - PR < d < uR + PR + d$, for all $d > uI + PT - PR$, $F_S(uI + PT - PR) \leq F_S(uR + PR + uI + PT - PR) = F_S(uR + uI + PT)$, then it is immediate that $\phi^*_{I,R}(uI, uR, PR, PT)$ is attained at $d^*(uI, uR, PR, PT) = uI + PT - PR$. $\square$

5.2. Remark (Proposition 5.1)**.** In Figure 2, we plot the two-step function indicating the argument of the distribution function of the total cost in (4.3), as a function of d.

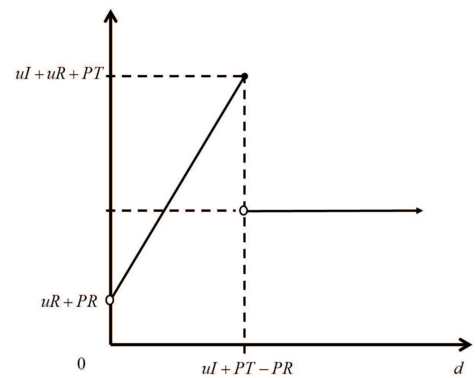


Figure 2. the argument of the distribution function of the total cost in (4.3) as a function of d

5.3. Remark (Proposition 5.1)**.** For this optimal reinsurance, in which the maximum joint survival probability of the insurer and the reinsurer is obtained, the individual survival probability of the insurer (4.1) is $\phi_I(uI, uI + PT - PR, PR, PT) = 1$, whereas the individual survival probability of the reinsurer (4.2) is $\phi_R(uR, uI + PT - PR, PR) = F_S(uI + uR + PT) = \phi^*_{I,R}(uI, uR, PR, PT)$. Hence, the insurer, with this optimal reinsurance, increases his/her individual survival probability (compared to the absence of reinsurance) in $(1 - P[S \leq uI + PT]) > 0$.

5.4. Remark (Proposition 5.1)**.** If the initial capitals of the insurer and the reinsurer are zero, then the maximum joint survival probability is obtained when the priority d is equal to the net premium of the insurer.

5.5. Proposition. In a stop-loss reinsurance with priority d, the program

$$\max_{PR} \phi_{I,R}(uI, uR, d, PR, PT) \text{ subject to } 0 < PR < PT$$

only provides a solution if $uI < d < uI + PT$, being in that case the maximum value $\phi^*_{I,R}(uI, uR, d, PT) = F_S(uI + uR + PT)$, which is reached for $PR^*(uI, uR, d, PT) = uI + PT - d$.

Proof. It is developed in a similar way as in Proposition 5.1. Since $F_S(x)$ is increasing in x, if $d \in (uI, uI + PT)$, for all $0 < PR \leq uI + PT - d$, $F_S(uR + uI + PT - d + d) = F_S(uR + uI + PT) \geq F_S(uR + PR + d)$ and for all $uI + PT - d < PR < PT$, $F_S(uI + uR + PT) > F_S(uI + PT - PR)$. If $d > uI + PT$, for all $0 < PR < PT$, $F_S(uI + PT - PR)$ does not have a maximum. If $d < uI$, for all $0 < PR < PT$, $F_S(uR + PR + d)$ does not have a maximum. Then, the program provides a solution only if $uI < d < uI + PT$ and $\phi^*_{I,R}(uI, uR, d, PT)$ is attained at $PR^*(uI, uR, d, PT) = uI + PT - d$. □

5.6. Remark (Proposition 5.5)**.** In Figure 3, we plot the two-step function indicating the argument of the distribution function of the total in (4.3), as a function of PR when $uI < d < uI + PT$.

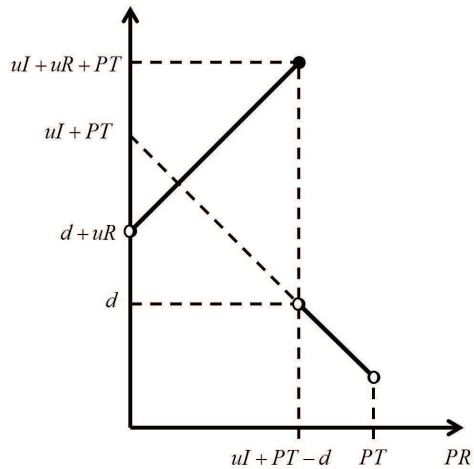


Figure 3. The argument of the distribution function of the total cost in (4.3) as a function of PR when $uI < d < uI + PT$

5.7. Proposition. In a stop-loss reinsurance with priority d and maximum m, the program

$$\max_{(d,m)} \phi_{I,R}(uI, uR, d, m, PR, PT) \text{ subject to } 0 < d < m$$

has a maximum value $\phi^*_{I,R}(uI, uR, PR, PT) = F_S\,(uI + uR + PT)$. This maximum is attained at the non-convex set

$$\big\{(d, m) \in \Re^2_+ \mid d \leq uI + PT - PR \text{ and } \ m = uR + PR + d\big\} \\ \cup \big\{(d, m) \in \Re^2_+ \mid d = uI + PT - PR \text{ and } \ m > uR + PR + d\big\}$$

Proof. The joint survival probability to be maximized now is (4.4), a piecewise function built with the distribution function of the total cost. Since $F_S(x)$ is increasing in x, for all $(d, m) \in \Re^2_+$ such that $d \leq uI + PT - PR$ and $m > uR + PR + d$, $F_S(uR + PR + d) \leq F_S(uR + PR + uI + PT - PR) = F_S(uR + uI + PT)$. For all $(d, m) \in \Re^2_+$ such that $d < uI + PT - PR$ and $m \leq uR + PR + d$, $F_S(uI + PT - PR + m - d) \leq F_S(uI + PT - PR + uR + PR) = F_S(uI + uR + PT)$. Taking into account that $F_S(uI + uR + PT) > F_S(uI + PT - PR)$, the proof is completed. □

5.8. Remark (Proposition 5.7)**.** In Figure 4, we plot the step function indicating the argument of the distribution function of the total cost in (4.4) as a function of d and m and its level curves. For $PT = 1$, $PR = 0.4$ and $uI = uR = 0$, the maximum value is 1 and the set of optimal points are $\{d \leq 0.6 \text{ and } \ m = 0.4 + d\} \cup \{d = 0.6 \text{ and } \ m > 0.4 + d\}$.

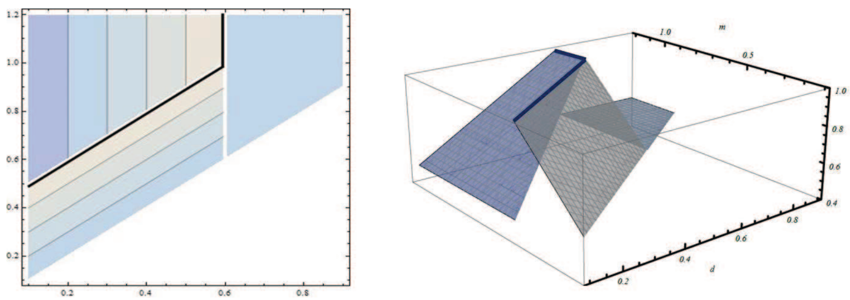


Figure 4. The argument of the distribution function of the total cost in (4.4) as a function of d and m (right graph) and its level curves (left graph) (for $PT = 1$, $PR = 0.4$ and $uI = uR = 0$)

5.9. Proposition. In a stop-loss reinsurance with priority d and maximum m, the program

$$\max_{PR} \phi_{I,R}(uI, uR, d, m, PR, PT) \text{ subject to } 0 < PR < PT$$

only provides solutions if one of the two following conditions is fulfilled: $uI < d < uI + PT$ and $m \geq uI + uR + PT$ (first condition) or $m < uI + uR + PT$ and $PT + uR > m - d > uR$ (second condition).

In that case, the maximum value is $\phi^*_{I,R}(uI, uR, d, m, PT) = F_S(uI + uR + PT)$, being the optimal premiums of the reinsurer

$$PR^*(uI, uR, d, m, PT) = \begin{cases} uI + PT - d & \text{if } uI < d < uI + PT \text{ and } m \geq uI + uR + PT, \\ m - d - uR & \text{if } m < uI + uR + PT \text{ and } PT + uR > m - d > uR. \end{cases}$$

Proof. Taking into account (4.4) and that $0 < PR < PT$, lets first consider the case that $d \in (uI, uI + PT)$. If $uI + PT - d < m - d - uR$, for all $0 < PR \leq uI + PT - d$, $F_S(uR + uI + PT - d + d) = F_S(uR + uI + PT) \geq F_S(uR + PR + d)$ and for all $uI + PT - d < PR < PT$, $F_S(uI + uR + PT) > F_S(uI + PT - PR)$. If $uI + PT - d = m - d - uR$, for all $0 < PR \leq uI + PT - d$, $F_S(m) = F_S(uI + uR + PT) > F_S(uR + PR + d)$ and for all $uI + PT - d < PR < PT$, $F_S(uI + uR + PT) > F_S(uI + PT - PR)$.

Secondly, lets consider that $(m - d) \in (uR, uR + PT)$ and $uI + PT - d > m - d - uR$, for all $0 < PR \leq m - d - uR$, $F_S(uI + PT - m + d + uR + m - d) = F_S(uI + PT + uR) > F_S(uR + PR + d)$ and for all $PR > m - d - uR$, $F_S(uI + uR + PT) > F_S(uI + PT - PR + m - d) > F_S(uI + PT - PR)$.

It is then easy to demonstrate that for all the other possibles values of d and m, the maximum does not exist. □

5.10. Remark (Proposition 5.9)**.** In Figure 5, the argument of the distribution function of the total cost in (4.4) is plotted as a function of PR for the values d and m for which the joint survival probability has a maximum. It can be divided into three cases depending on whether $uI + PT - d$ is less, equal or greater than $m - d - uR$.

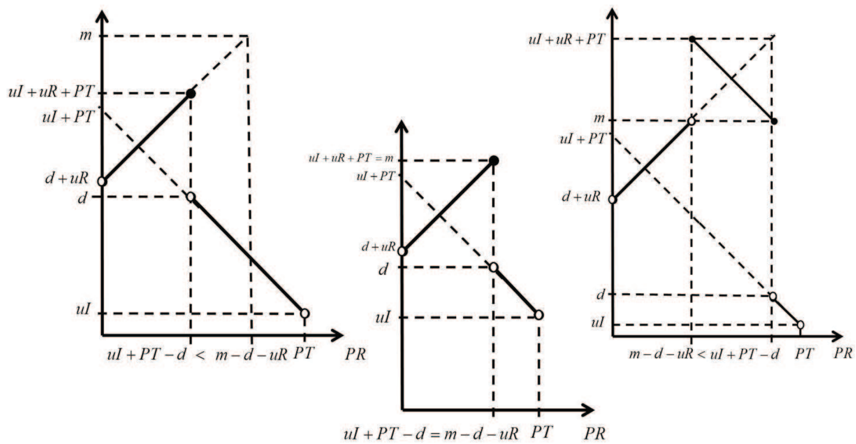


Figure 5. The argument of the distribution function of the total cost in (4.4) as a function of PR when $uI + PT - d \lesseqgtr m - d - uR$. The graph on the left considers $uI + PT - d < m - d - uR$; the graph on the middle considers $uI + PT - d = m - d - uR$ and the graph on the right considers $uI + PT - d > m - d - uR$.

From Propositions 5.1, 5.5, 5.7 and 5.9, the maximum joint survival probability (considering the constraints), when it exists, is equal to

$$F_S\left(uI + uR + PT\right).$$

From the first definition of ruin in a bivariate risk process ([12]), the joint survival probability equals to the minimum between the survival probability of the insurer and the survival probability of the reinsurer, and this is also true at the optimal points. Then, at the optimal points, the survival probability of the insurer or the reinsurer must be equal to $F_S\left(uI + uR + PT\right)$, and the other must be greater than this value. Table 3 includes the values of the survival probability of the insurer and the reinsurer at the points that maximize the joint survival probability.

Table 3. ϕ_I and ϕ_R at the optimal points for the different optimization problems

	Φ_I	Φ_R
$d^* = uI + PT - PR$ (Prop. 5.1)	1	$F_S(uI + uR + PT)$
$PR^* = uI + PT - d$, if $uI < d < uI + PT$ (Prop. 5.5)	1	$F_S(uI + uR + PT)$
$\left\{(d, m) \in \Re_+^2 \mid d \leq uI + PT - PR \text{ and } m = uR + PR + d\right\}$ (Prop. 5.7)	$F_S(uI + uR + PT)$	1
$\left\{(d, m) \in \Re_+^2 \mid d = uI + PT - PR \text{ and } m > uR + PR + d\right\}$ (Prop. 5.7)	$F_S(m)$, $m > uI + uR + PT$	$F_S(uI + uR + PT)$
$PR^* = uI + PT - d$, if $uI<d<uI+PT$ and $m \geq uI+uR+PT$ (Prop. 5.9)	$F_S(m)$, $m > uI + uR + PT$	$F_S(uI + uR + PT)$
$PR^* = m - d - uR$, if $m<uI+uR+PT$ and $PT+uR>m-d>uR$ (Prop. 5.9)	$F_S(uI + uR + PT)$	1

5.11. Example. Using the data for the total cost in Example 3.4, assume first that a stop-loss contract with priority d is agreed and that the initial reserves of the insurer and the reinsurer are zero. The premium fixed by the insurer is 1.8 (so if the criterion is the expected value, the security loading applied by the insurer is 80%). The premium earned by the reinsurer is fixed and equal to $PR = 0.5, ..., 1.5$. In Table 4, we calculate the priority that maximizes the joint survival probability, using Proposition 5.1, and the difference between the premium earned by the reinsurer and the expectation of its cost, $PR - E\left[SR(d^*)\right]$, if the gamma (G), the translated gamma (TG) or the normal approximations (N), are used. In Table 4, we also include the net security premium for the insurer, that is given by $1.8 - PR - E\left[SI(d^*)\right]$. These two quantities included in Table 4, permit us to calculate the security loading of the reinsurer and the insurer (for the insurer it is the net loading) included in the optimal strategy. These security loadings are shown in Table 5. In Table 6, we calculate the maximal joint survival probability (that equals to the survival probability of the reinsurer (Remark 5.3)), and the increase in the survival probability of the insurer if the optimal reinsurance is agreed, when the gamma, the translated gamma or the normal approximations, are used.

Table 4. Priority, security premium for the reinsurer and net security premium for the insurer if the joint survival probability is maximized for several fixed reinsurer's premiums

		$PR - E\left[SR(d^*)\right]$			$1.8 - PR - E\left[SI(d^*)\right]$		
PR	d^*	G	TG	N	G	TG	N
0.5	1.3	0.1013	0.0820	0.0732	0.6987	0.7180	0.7268
0.6	1.2	0.1750	0.1518	0.1302	0.6250	0.6482	0.6698
0.7	1.1	0.2466	0.2195	0.1844	0.5534	0.5805	0.6156
0.8	1	0.3161	0.2847	0.2358	0.4839	0.5153	0.5642
0.9	0.9	0.3831	0.3474	0.2844	0.4169	0.4526	0.5156
1	0.8	0.4474	0.4072	0.3302	0.3526	0.3928	0.4698
1.1	0.7	0.5087	0.4641	0.3732	0.2913	0.3359	0.4268
1.2	0.6	0.5667	0.5178	0.4134	0.2333	0.2822	0.3866
1.3	0.5	0.6209	0.5679	0.4509	0.1791	0.2321	0.3491
1.4	0.4	0.6706	0.6143	0.4858	0.1294	0.1857	0.3142
1.5	0.3	0.7151	0.6565	0.5181	0.0849	0.1435	0.2819

Table 5. Security loadings of the insurer and the reinsurer if the joint survival probability is maximized for several fixed reinsurer's premiums

		$\frac{100(PR-E\left[SR(d^*)\right])}{E\left[SR(d^*)\right]}$			$\frac{100(1.8-PR-E\left[SI(d^*)\right])}{E\left[SI(d^*)\right]}$		
PR	d^*	G	TG	N	G	TG	N
0.5	1.3	25.42	19.61	17.14	116.18	123.38	126.81
0.6	1.2	41.17	33.88	27.71	108.70	117.50	126.34
0.7	1.1	54.40	45.67	35.76	101.24	111.76	127.08
0.8	1	65.31	55.25	41.80	93.78	106.31	129.46
0.9	0.9	74.11	62.85	46.20	86.31	101.18	134.13
1	0.8	80.97	68.70	49.29	78.81	96.44	142.29
1.1	0.7	86.04	72.99	51.34	71.26	92.24	156.26
1.2	0.6	89.49	75.89	52.55	63.61	88.82	181.17
1.3	0.5	91.42	77.57	53.11	55.83	86.63	231.32
1.4	0.4	91.94	78.18	53.14	47.82	86.69	366.30
1.5	0.3	91.12	77.83	52.76	39.44	91.69	1559.80

Table 6. Maximal joint survival probability and the increase in the survival probability of the insurer

	G	TG	N
$\phi^*_{I,R} = \phi_R = F_S(1.8)$	0.8202875	0.7955186	0.7141962
$1 - P\,[S \leq 1.8]$	0.1797125	0.2044814	0.2858038

As it is reflected in Table 6, obviously, the maximal joint survival probability ($\phi^*_{I,R} = \phi_R = F_S(1.8)$) and the increase in the survival probability of the insurer due to the optimal reinsurance ($1 - P\,[S \leq 1.8]$), is always the same and is independent of the specific optimal combination of the reinsurer's premium and priority. Hence, from the point of view of the joint survival probability, the reinsurer survival probability and the insurer survival probability, all the alternative combinations of the reinsurer's premium and priority included in Table 5 are indifferent. The differences in the security loading applied by the reinsurer and the net security loading of the insurer do not modify the optimal survival probabilities.

Assume now that the insurer and the reinsurer have positive initial reserves, and that the reinsurer's premium is 0.5 and the total premium is 1.8. From Proposition 5.1, the optimal priority is $d^* = uI + 1.3$, and the maximum joint survival probability is $F_S(uI + uR + 1.8) = \phi^*_{I,R}$. Table 7 includes the optimal priority and the maximum joint survival probability for several combinations of initial capitals, using the translated gamma approximation.

Table 7. d^* and $\phi^*_{I,R}$ as functions of initial capitals, for $PR = 0.5$ and $PT = 1.8$

uI/uR		0.25	0.5	0.75	1
0.25	d^*	1.55	1.55	1.55	1.55
	$\phi^*_{I,R}$	0.855824	0.8788329	0.8981223	0.9143059
0.5	d^*	1.8	1.8	1.8	1.8
	$\phi^*_{I,R}$	0.8788329	0.8981223	0.9143059	0.9278928
0.75	d^*	2.05	2.05	2.05	2.05
	$\phi^*_{I,R}$	0.8981223	0.9143059	0.9278928	0.9393062
1	d^*	2.3	2.3	2.3	2.3
	$\phi^*_{I,R}$	0.9143059	0.9278928	0.9393062	0.9488984

Table 7 shows that when different combinations of initial capitals are considered for a specific uI, the optimal priority does not vary if uR is increased. This result is due to the fact that d^* does not depend on the initial capital of the reinsurer. However, the joint survival probability does change with increasing values.

6. Concluding remarks

In the stop-loss reinsurance contract, the cost of the claims of both the insurer and the reinsurer are related. This work contributes to the analysis of the optimal stop-loss reinsurance in one period, from the joint point of view of the insurer and the reinsurer and then, incorporating the aforementioned relation.

Several optimal problems with two different objective functions are studied. First, using the total variance risk measure, we analyze the optimal reinsurance parameters (retention and maximum) that maximize the covariance (and also the coefficient of correlation) between the cost of the claims of the insurer and the reinsurer. Second, two optimal problems with the same objective function, the joint survival probability of the insurer and the reinsurer in one period, are solved. The maximum joint survival probability always exists if the reinsurance premium is fixed, and is equal to the probability that the total cost is less than, or equal, to the sum of the total premium and the two initial capitals. This maximum is attained for a unique value of the priority or for a non-convex set of priority and maximum if the reinsurance contract includes a maximum. If we consider that the parameters of the reinsurance contract are fixed, the optimal reinsurance premium and the maximum joint survival probability do not always exist, and in case they exist, the maximum is exactly the same as in the first problem. These findings can be of great help for the insurer and reinsurer in their decision making process.

Acknowledgements

This work was partially supported by MICINN (Spain) grant ECO2010-22065-C03-03. We thank the referees for useful remarks and suggestions. We are also grateful to F. Javier Martínez de Albéniz for helpful comments.

References

[1] Assa, H. *On optimal reinsurance policy with distortion risk measures and premiums*, Insur. Math. Econ. **61**, 70-75, 2015.

[2] Azcue, P. and Muler, N. *Optimal reinsurance and dividend distribution policies in the Cramér-Lundberg model*, Math. Financ. **15** (2), 261-308, 2005.

[3] Balbás, A., Balbás, B., Balbás, R. and Heras, A. *Optimal reinsurance under risk and incertainty*, Insur. Math. Econ. **60**, 61-74, 2015.

[4] Balbás, A., Balbás, R. and Heras, A. *Optimal reinsurance with general risk measures*, Insur. Math. Econ. **44** (3), 374-384, 2009.

[5] Bohman, H. and Esscher, F. *Studies in risk theory with numerical illustrations concerning distribution functions and stop loss premiums. Part I*, Scand. Actuar. J. **1963** (3-4), 173-225, 1963.

[6] Borch, K. *An attempt to determine the optimum amount of stop loss reinsurance*, Transactions of the 16th International Congress of Actuaries **I**, 597-610, 1960.

[7] Borch, K. *The optimal reinsurance treaty*, Astin Bull. **5** (2), 293-297, 1969.

[8] Borch, K. *The mathematical theory of insurance* (Lexington Books, Lexington, 1974).

[9] Cai, J. and Tan, K. S. *Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures*, Astin Bull. **37** (1), 93-112, 2007.

[10] Cai, J., Fang, Y., Li, Z. and Willmot, G. E. *Optimal reciprocal reinsurance treaties under the joint survival probability and the joint profitable probability*, J. Risk Insur. **80** (1), 145-168, 2013.

[11] Cai, J., Tan, K. S., Weng, C. and Zhang, Y. *Optimal reinsurance under VaR and CTE risk measures*, Insur. Math. Econ. **43** (1), 185-196, 2008.

[12] Castañer, A., Claramunt, M. M. and Lefèvre, C. *Survival probabilities in bivariate risk models, with application to reinsurance*, Insur. Math. Econ. **53** (3), 632-642, 2013.

[13] Castañer, A., Claramunt, M. M. and Mármol, M. *Ruin probability and time of ruin with a proportional reinsurance threshold strategy*, Top **20** (3), 614-638, 2012.

[14] Centeno, M. L. and Simões, O. *Optimal reinsurance*, Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat. **103** (2), 387-404, 2009.

[15] Cheung, K. C. *Optimal reinsurance revisited - a geometric approach*, Astin Bull. **40** (1), 221-239, 2010.

[16] Cui, W., Yang, J. and Wu, L. *Optimal reinsurance minimizing the distortion risk measure under general reinsurance premium principles*, Insur. Math. Econ. **53** (1), 74-85, 2013.

[17] Daykin, C. D., Pentikainen, T. and Pesonen, M. *Practical risk theory for actuaries* (Chapman and Hall, London, 1994).

[18] Denuit, M. and Charpentier, A. *Mathématiques de l'assurance non-vie: Tome 1, Principes fondamentaux de théorie du risque* (Economica, Paris, 2004).

[19] Denuit, M. and Varmandele, C. *Optimal reinsurance and stop-loss order*, Insur. Math. Econ. **22** (1), 229-233, 1998.

[20] Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R. and Vyncke, D. *The concept of comonotonicity in actuarial science and finance: applications*, Insur. Math. Econ. **31** (2), 133-161, 2002.

[21] Dickson, D. C. M. *Insurance risk and ruin* (Cambridge University Press, United Kingdom, 2005).

[22] Dimitrova, D. S. and Kaishev, V. K. *Optimal joint survival reinsurance: an efficient frontier approach*, Insur. Math. Econ. **47** (1), 27-35, 2010.

[23] Fang, Y. and Qu, Z. *Optimal combination of quota-share and stop-loss reinsurance treaties under the joint survival probability*, IMA. J. Manag. Math. **25** (1), 89-103, 2014.

[24] Gajek, L. and Zagrodny, D. *Insurer's optimal reinsurance strategies*, Insur. Math. Econ. **27** (1), 105-112, 2000.

[25] Gajek, L. and Zagrodny, D. *Reinsurance arrangements maximizing insurer's survival probability*, J. Risk Insur. **71** (3), 421-435, 2004.

[26] Gendron, M. and Crepeau, H. *On the computation of the aggregate claim distribution when individual claims are inverse gaussian*, Insur. Math. Econ. **8** (3), 251-258, 1989.

[27] Gerber, H. U. *The impact of reinsurance on the insurer's risk*, in: F. De Vylder, M. Goovaerts, J. Haezendonck. Premium calculation in insurance (D. Reidel Publishing Company, Dordrecht, Holland, 1984), 171-181.

[28] Guerra, M. and Centeno, M. D. L. *Optimal reinsurance policy: the adjustment coefficient and the expected utility criteria*, Insur. Math. Econ. **42** (2), 529-539, 2008.

[29] Hesselager, O. *Some results on optimal reinsurance in terms of the adjustment coefficient*, Scand. Actuar. J. **1990** (1), 80-95, 1990.

[30] Hoøjgaard, B. and Taksar, M. *Optimal proportional reinsurance policies for diffusion models*, Scand. Actuar. J. **1998** (2), 166-180, 1998.

[31] Hoøjgaard, B. and Taksar, M. *Optimal proportional reinsurance policies for diffusion models with transaction costs*, Insur. Math. Econ. **22** (1), 41-51, 1998.

[32] Hürlimann, W. *Non-optimality of a linear combination of proportional and non-proportional reinsurance*, Insur. Math. Econ. **24** (3), 219-227, 1999.

[33] Hürlimann, W. *Optimal reinsurance revisited - point of view of cedent an reinsurer*, Astin Bull. **41** (2), 547-574, 2011.

[34] Ignatov, Z. G., Kaishev, V. K. and Krachunov, R. S. *Optimal retention levels, given the joint survival of cedent and reinsurer*, Scand. Actuar. J. **2004** (6), 401-430, 2004.

[35] Kaas, R. *How to (and how not to) compute stop-loss premiums in practice*, Insur. Math. Econ. **13** (3), 241-254, 1993.

[36] Kaas, R., Goovaerts, M. J., Dhaene, J. and Denuit, M. *Modern actuarial risk theory: using R* (Springer, Heidelberg, 2008).

[37] Kaishev, V. K. and Dimitrova, D. S. *Excess of loss reinsurance under joint survival optimality*, Insur. Math. Econ. **39** (3), 376-389, 2006.

[38] Kaluszka, M. *Optimal reinsurance under mean-variance premium principles*, Insur. Math. Econ. **28** (1), 61-67, 2001.

[39] Liang, Z. and Guo, J. *Optimal proportional reinsurance and ruin probability*, Stoch. Model. **1984** (2), 65-90, 2007.

[40] Pesonen, M. I. *Optimal Reinsurance*, Scand. Actuar. J. **1990** (1), 80-95, 1984.

[41] Salcedo-Sanz, S., Carro-Calvo, L., Claramunt, M. M., Castañer, A. and Mármol, M. *Effectively tackling reinsurance problems by using evolutionary and swarm intelligence algorithms*, Risks **2** (2), 132-145, 2014.

[42] Seal, H. L. *Approximations to risk theory's F(x,t) by means of the gamma distribution*, Astin Bull. **9** (1-2), 213-218, 1977.

[43] Schmidli, H. *Optimal proportional reinsurance policies in a dynamic setting*, Scand. Actuar. J. **2001** (1), 55-68, 2001.

[44] Schmidli, H. *On Cramer-Lundberg approximations for ruin probabilities under optimal excess of loss reinsurance*, Working Paper 193. Laboratory of Actuarial Mathematics, University of Copenhagen, 2004.

[45] Tan, K. S., Weng, C. and Zhang, Y. *VaR and CTE criteria for optimal quota-share and stop-loss reinsurance*, N. Am. Actuar. J. **13** (4), 459-482, 2009.

[46] Van Heerwaarden, A. E., Kaas, R. and Goovaerts, M. J. *Optimal reinsurance in relation to ordering of risks*, Insur. Math. Econ. **8** (1), 11-17, 1989.

[47] Van Wouwe, M., De Vylder, F. and Goovaerts, M. *The influence of reinsurance limits on infinite time of ruin probabilities*, in: F. De Vylder, M. Goovaerts, J. Haezendonck. Premium calculation in insurance (D. Reidel Publishing Company, Dordrecht, Holland, 1984), 493-504.

[48] Zheng, Y. and Cui, W. *Optimal reinsurance with premium constraint under distortion risk measures*, Insur. Math. Econ. **59**, 109-120, 2014.

[49] Zheng, Y., Cui, W. and Yang, J. *Optimal reinsurance under distortion risk measures and expected value premium principle for reinsurer*, J. Syst. Sci. Complex. **28** (1), 122-143, 2015.

[50] Zhu, Y., Chi, Y. and Weng, C. *Multivariate reinsurance designs for minimizing an insurer's capital requirement*, Insur. Math. Econ. **59**, 144-155, 2014.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 521 – 540

A data driven runs test to identify first order positive Markovian dependence

Jimmy A. Corzo *† and Myrian E. Vergara‡

Abstract

We propose a data driven test to identify first order positive Markovian dependence in a Bernoulli sequence, based on a combination of two runs tests: a well known runs test for the same purpose conditional on the numbers of ones in the sequence, and a modified runs test independent of the number of ones. We give analytic expressions for the exact distribution of the modified runs test statistic and for its power; also we built an algorithm to calculate it explicitly. To compare the power of the tests, we calculated these for some values of the proportion of ones and the success probability. We show that there are some intervals for the success probability in which the new runs test surpasses the power of the conditional test, and that the data driven test improves the power of the two runs tests, when they are considered separately.

Keywords: Markov-dependent Bernoulli trials, Data driven runs test, Runs distributions, Hypothesis of randomness, Power of a test.

2000 AMS Classification: Primary 62M02; Secondary 60J10

Received : 08.03.2014 *Accepted* : 02.03.2015 *Doi* : 10.15672/HJMS.2015559562

*Department of Statistics, Universidad Nacional de Colombia, Bogota, Colombia, Email: jacorzos@unal.edu.co

†Corresponding Author.

‡Department of Basic Sciences, Universidad de La Salle, Bogota, Colombia, and Department of Statistics, Universidad Nacional de Colombia, Bogota, Colombia Email: mvergara@unisalle.edu.co and mevergaram@unal.edu.co

1. Introduction

Since the pioneering work performed by [1], in which he calculated the power of a test for randomness based on the total number of runs conditioned on the number of successes in a binary sequence versus the first order Markovian dependence alternative, many other works have appeared. For instance, one year later [2] calculated the conditional distribution of the longest success run for a second order Markovian dependence alternative. Later [3] studied the power of the conditional David's test with a parametrization of the transition probabilities. [4] used the total number of runs conditioned on the number of symbols of each type for pattern sequences and calculated critical values for the distribution of the number of runs conditioned on the number of symbols of each type for pattern sequences for randomness tests. [5] proposed a randomness test for the Markovian first order alternative based on the length of the longest run and developed methods of computing the probability of the occurrence of a given success-failure run as a function of the composition of the run, the number n of trials and the probabilities of the possible outcomes at each trial. [6] used the total number of success runs of length greater than k and the total number of success runs of length k as test statistics to test the randomness hypotheses versus three alternatives: First order Markov-dependence, non-systematic unimodal and bimodal clustering and cyclical clustering. By a Monte Carlo study they compared the powers of their tests with the power of two known tests; some based on the total number of success (or failure) runs and others based on the length of the longest success run, and they found that the test based on the number of overlapping success runs of length k is slightly less sensitive than its competitors for success probabilities near to 1. [7] studied a randomness test based on the conditional distribution of the sum of the exact lengths of runs of length greater than k successes. They found that when the type I error must be kept low ($\alpha = 0.01$), their test is more powerful than a test based on the number of runs of exact length k, for the first order Markov-dependence alternative.

Many other researchers have focused their research to calculate explicit expressions for the distributions of runs statistics in many contexts; for example, [5], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18].

We propose a data driven test to identify first order positive Markovian dependence in a Bernoulli sequence, based on a combination of two runs tests: the well known Barton and David's runs test, conditional on the number of ones in the sequence, for the Markovian alternative, and an extension of this test to an unconditional test (on the number of ones). We give analytic expressions for the exact distribution of the original and of the extended runs test statistic and for its power; we built an algorithm and we developed the R code to calculated it explicitly. To compare the power of the tests, we calculate the exact powers of both tests for some values of the proportion of ones and the success probability. We found intervals for the success probability for which the unconditional runs test surpasses the power of the conditional test, and we show by calculating the powers, that the data driven test optimizes the power of the two runs tests, when they are considered separately. Finally, as a bonus product, we developed an algorithm and implemented it in R code to solve a polynomial with matrix coefficients, to find explicitly the distribution of the Markov-Binomial Distribution.

2. Two Runs Tests for Markovian Dependence

Let $\eta_1, \ldots, \eta_N$ be a two state Markov chain and let p be the success probability such that:

$$P(\eta_t = 1) = p, \quad P(\eta_t = 0) = 1 - p, \quad 0 < p < 1 \quad \text{for } t = 1, 2, \ldots, N,$$

and stationary transition probabilities ([3]):

$$\begin{aligned} P_{11} &= P(\eta_t = 1 \mid \eta_{t-1} = 1) = (1-\theta)p + \theta, \\ P_{10} &= P(\eta_t = 0 \mid \eta_{t-1} = 1) = (1-\theta)(1-p), \\ P_{01} &= P(\eta_t = 1 \mid \eta_{t-1} = 0) = (1-\theta)p, \\ P_{00} &= P(\eta_t = 0 \mid \eta_{t-1} = 0) = 1 - (1-\theta)p, \end{aligned} \tag{2.1}$$

where θ is the coefficient of correlation between η_{t-1} and η_t for $t = 2, 3, \ldots, N$.

Although Barton and David gave the bounds ± 1 for θ, they can be improved as follows:

From (2.1), the following is true:

$$\begin{aligned} 0 \le (1-\theta)p + \theta \le 1 \quad \text{or} \quad 0 \le (1-\theta)(1-p) \le 1 \quad \text{implies} \quad -\frac{p}{1-p} \le \theta \le 1, \\ 0 \le (1-\theta)p \le 1 \quad \text{or} \quad 0 \le 1 - (1-\theta)p \le 1 \quad \text{implies} \quad -\frac{1-p}{p} \le \theta \le 1. \end{aligned} \tag{2.2}$$

Now from (2.2) we conclude:

$$\begin{aligned} \text{for} \quad p = 1/2 \quad &\text{it follows that} \quad -1 \le \theta \le 1, \\ \text{for} \quad 0 < p < 1/2 \quad &\text{it follows that} \quad -\frac{p}{1-p} \le \theta \le 1, \\ \text{and for} \quad 1/2 < p < 1 \quad &\text{it follows that} \quad -\frac{1-p}{p} \le \theta \le 1. \end{aligned} \tag{2.3}$$

The conditions (2.3) on θ are represented graphically in Figure 1.

2.1. The Barton-David Test. From now on, we will consider the following test problem:

$$H_0 : \theta = 0 \quad \text{against} \quad H_1 : \theta > 0 \quad \text{(positive Markovian dependence)}$$

Let m be the fixed number of ones (successes), $n = N - m$ the number of zeros (failures), and let R_m be the total number of runs in $\eta_1, \ldots, \eta_N$. [3] gave a conditional (on m) runs test based on R_m, which rejects H_0 in favor of the positive Markovian alternative for few runs. The critical region was justified as follows: under H_0, $\theta = 0$ holds true and hence $\eta_1, \ldots, \eta_N$ is a Bernoulli sequence of independent and identically distributed (i.i.d.) random variables; under H_1, either $\theta > 0$ (positive dependence), which implies $P_{11} > P_{01}$ or $P_{00} > P_{10}$, and then we expect few runs.

Let $\zeta = \sum_{t=1}^{N} \eta_t$ be a random variable denoting the number of ones in the sequence $\eta_1 \ldots, \eta_N$. [3] gave the following expressions for:

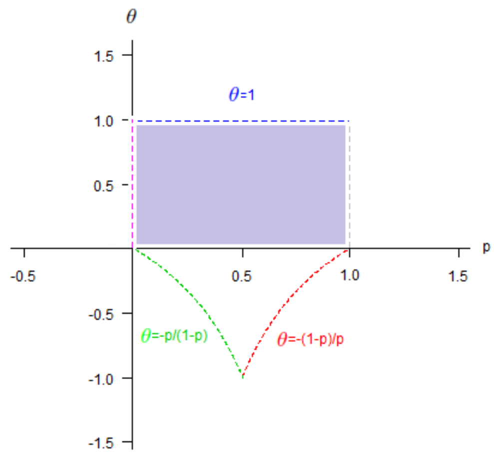


Figure 1. Relation between p and θ

a) the null distribution of R_m

$$P_0(R_m = r \,|\, \zeta = m) = \begin{cases} \dfrac{2\binom{m-1}{\frac{r}{2}-1}\binom{n-1}{\frac{r}{2}-1}}{\binom{N}{m}} & \text{if } r \text{ is even,} \\[2ex] \dfrac{\binom{m-1}{\frac{r-1}{2}}\binom{n-1}{\frac{r-1}{2}-1}+\binom{m-1}{\frac{r-1}{2}-1}\binom{n-1}{\frac{r-1}{2}}}{\binom{N}{m}} & \text{if } r \text{ is odd.} \end{cases} \tag{2.4}$$

b) the power of the conditional R_m test

$$P_\theta(R_m = r \,|\, \zeta = m) = \begin{cases} \frac{2}{S}\frac{1}{1-\theta}\binom{m-1}{\frac{r}{2}-1}\binom{n-1}{\frac{r}{2}-1}\left(\frac{p(1-p)(1-\theta)^2}{(p(1-\theta)+\theta)(1-p(1-\theta))}\right)^{\frac{r}{2}} & \text{if } r \text{ is even,} \\[2ex] \frac{1}{S}\binom{m-1}{\frac{r-1}{2}-1}\binom{n-1}{\frac{r-1}{2}-1}\frac{1}{(p(1-\theta)+\theta)(1-p(1-\theta))}\times & \\ \left(-2p(1-p)-\theta(p^2+(1-p)^2)+\frac{Np(1-p(1-\theta))+n\theta(1-2p)}{\frac{r-1}{2}}\right)\times & \\ \left(\frac{p(1-p)(1-\theta)^2}{(p(1-\theta)+\theta)(1-p(1-\theta))}\right)^{\frac{r-1}{2}} & \text{if } r \text{ is odd,} \end{cases} \tag{2.5}$$

where

$$S = \sum_{k=1}^{n}\left[\frac{p(1-p)(1-\theta)^2}{(p(1-\theta)+\theta)(1-p(1-\theta))}\right]^k \binom{m-1}{k-1}\binom{n-1}{k-1}$$
$$\left[\frac{\theta(1-\theta)k+(1-\theta)\left[Np(1-p)+\theta(Np^2+n(1-2p))\right]}{k(1-\theta)(p(1-\theta)+\theta)(1-p(1-\theta))}\right].$$

2.2. A Modified Runs Test. The conditional R_m test in (2.4) can be modified as follows: as we noted above, under H_0, $\eta_1,\ldots,\eta_N$ is an i.i.d. Bernoulli sequence and hence ζ is Binomial distributed with parameters N and $p = P(\eta_t = 1)$, for $t = 1,\ldots,N$. Now let R be the total number of runs, without taking into account the number of ones in the sequence $\eta_1\ldots,\eta_N$. The modified R test rejects H_0 in favor of the positive Markovian dependence alternative for few runs, with the same arguments as for the R_m test, but now the reject region must be calculated from the unconditional distribution of R, by means of the theorem of total probabilities, as follows:

$$P_0(R = r) = \sum_{m=0}^{N} P_0(R_m = r \,|\, \zeta = m) \binom{N}{m} p^m (1-p)^{N-m},$$

for $r = 1, \ldots, N$, where the conditional distribution of R_m is calculated as in (2.4).

2.2.1. *The Power of the Modified R Test.* We obtain the distribution of the R, under the Markovian alternative H_1 as follows:

$$P_\theta(R = r) = \sum_{m=0}^{N} P_\theta(R_m = r \,|\, \zeta = m) P_\theta(\zeta = m), \tag{2.6}$$

for $r = 1, \ldots, N$, where $P_\theta(R_m = r \,|\, \zeta = m)$ is given in (2.5), and $P_\theta(\zeta = m)$ under H_1 is calculated by means of the probability generating function (pgf) of the Markov-Binomial Distribution ([19]) as a function of the dummy variable s:

$$\begin{aligned} G_N(s) &= \begin{pmatrix} ps & 1-p \end{pmatrix} \begin{pmatrix} ((1-\theta)p+\theta)s & (1-\theta)(1-p) \\ (1-\theta)ps & 1-(1-\theta)p \end{pmatrix}^{N-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} ps & 1-p \end{pmatrix} A^{N-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \end{aligned} \tag{2.7}$$

$N = 1, 2, \ldots,$ for $0 \le s \le 1$, where $A = \begin{pmatrix} ((1-\theta)p+\theta)s & (1-\theta)(1-p) \\ (1-\theta)ps & 1-(1-\theta)p \end{pmatrix}$.

2.2.2. *Algorithm to Calculate the Markov-Binomial Distribution Explicitly.* To calculate the power of the R test in (2.6), we need to extract the coefficients of s in the pgf (2.7) of the Markov-Binomial Distribution, which contains the probability distribution of ζ. For this, the following algorithm is useful:

$$\begin{aligned} A^{N-1} &= \begin{pmatrix} ((1-\theta)p+\theta)s & (1-\theta)(1-p) \\ (1-\theta)ps & 1-(1-\theta)p \end{pmatrix}^{N-1} \\ &= \begin{pmatrix} c_1^{11}s + c_2^{11}s^2 + \cdots + c_{N-1}^{11}s^{N-1} & c_0^{12} + c_1^{12}s + \cdots + c_{N-2}^{12}s^{N-2} \\ c_1^{21}s + c_2^{21}s^2 + \cdots + c_{N-1}^{21}s^{N-1} & c_0^{22} + c_1^{22}s + \cdots + c_{N-2}^{22}s^{N-2} \end{pmatrix} \\ &= \begin{pmatrix} c_1^{11} & 0 \\ c_1^{21} & 0 \end{pmatrix} s + \begin{pmatrix} c_2^{11} & 0 \\ c_2^{21} & 0 \end{pmatrix} s^2 + \cdots + \begin{pmatrix} c_{N-1}^{11} & 0 \\ c_{N-1}^{21} & 0 \end{pmatrix} s^{N-1} + \\ &\qquad \begin{pmatrix} 0 & c_0^{12} \\ 0 & c_0^{22} \end{pmatrix} + \begin{pmatrix} 0 & c_1^{12} \\ 0 & c_1^{22} \end{pmatrix} s + \cdots + \begin{pmatrix} 0 & c_{N-2}^{12} \\ 0 & c_{N-2}^{22} \end{pmatrix} s^{N-2} \\ &= \begin{pmatrix} 0 & c_0^{12} \\ 0 & c_0^{22} \end{pmatrix} + \begin{pmatrix} c_1^{11} & c_1^{12} \\ c_1^{21} & c_1^{22} \end{pmatrix} s + \\ &\qquad \cdots + \begin{pmatrix} c_{N-2}^{11} & c_{N-2}^{12} \\ c_{N-2}^{21} & c_{N-2}^{22} \end{pmatrix} s^{N-2} + \begin{pmatrix} c_{N-1}^{11} & 0 \\ c_{N-1}^{21} & 0 \end{pmatrix} s^{N-1} \\ &= C^{(0)} + C^{(1)}s + \cdots + C^{(N-2)}s^{N-2} + C^{(N-1)}s^{N-1} \\ &= \sum_{m=0}^{N-1} C^{(m)} s^m, \end{aligned} \tag{2.8}$$

where $C^{(m)}$ are matrices with the role of coefficients of the polynomial in s.

For example, for $N = 3$ the coefficients of the polynomial can be calculated as follows: let A be the transition matrix whose first column is multiplied by the auxiliary variable s. Then the power of the matrix in (2.7) can be written as:

$$\begin{aligned}
A^2 &= \begin{pmatrix} ((1-\theta)p+\theta)s & (1-\theta)(1-p) \\ (1-\theta)ps & 1-(1-\theta)p \end{pmatrix}^2 \\
&= \begin{pmatrix} c_1^{11}s + c_2^{11}s^2 & c_0^{12} + c_1^{12}s \\ c_1^{21}s + c_2^{21}s^2 & c_0^{22} + c_1^{22}s \end{pmatrix} \\
&= \begin{pmatrix} 0 & c_0^{12} \\ 0 & c_0^{22} \end{pmatrix} + \begin{pmatrix} c_1^{11} & c_1^{12} \\ c_1^{21} & c_1^{22} \end{pmatrix} s + \begin{pmatrix} c_2^{11} & 0 \\ c_2^{21} & 0 \end{pmatrix} s^2 \\
&= C^{(0)} + C^{(1)}s + C^{(2)}s^2,
\end{aligned}$$

where

$$\begin{aligned}
c_0^{12} &= (1-(1-\theta)p)(1-\theta)(1-p) & c_2^{11} &= ((1-\theta)p+\theta)^2 \\
c_0^{22} &= (1-(1-\theta)p)^2 & c_2^{21} &= ((1-\theta)p+\theta)(1-\theta)p \\
c_1^{11} &= (1-\theta)^2(1-p)p & c_1^{12} &= ((1-\theta)p+\theta)(1-\theta)(1-p) \\
c_1^{21} &= (1-(1-\theta)p)(1-\theta)p & c_1^{22} &= (1-\theta)^2(1-p)p
\end{aligned}$$

Using the polynomial expression for the power of the matrix A introduced in (2.8), the pgf of ζ can be expressed as:

$$\begin{aligned}
G_N(s) &= \begin{pmatrix} ps & 1-p \end{pmatrix} \left[\sum_{m=0}^{N-1} C^{(m)}s^m\right] \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \{\begin{pmatrix} p & 0 \end{pmatrix} s + \begin{pmatrix} 0 & 1-p \end{pmatrix}\} \left[\sum_{m=0}^{N-1} C^{(m)}s^m\right] \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} p & 0 \end{pmatrix} \sum_{m=1}^{N} C^{(m-1)}s^m \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 & 1-p \end{pmatrix} \left[\sum_{m=0}^{N-1} C^{(m)}s^m\right] \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} p & 0 \end{pmatrix} \left[\sum_{m=1}^{N-1} C^{(m-1)}s^m + C^{(N-1)}s^N\right] \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \\
&\qquad \begin{pmatrix} 0 & 1-p \end{pmatrix} \left[C^{(0)} + \sum_{m=1}^{N-1} C^{(m)}s^m\right] \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} p & 0 \end{pmatrix} C^{(N-1)} \begin{pmatrix} 1 \\ 1 \end{pmatrix} s^N + \\
&\qquad \sum_{m=1}^{N-1} \left[\begin{pmatrix} p & 0 \end{pmatrix} C^{(m-1)} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 & 1-p \end{pmatrix} C^{(m)} \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] s^m + \\
&\qquad \begin{pmatrix} 0 & 1-p \end{pmatrix} C^{(0)} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.
\end{aligned} \tag{2.9}$$

2.2.3. *Algorithm to calculate the coefficients $C^{(m)}$.* To obtain the coefficients $C^{(m)}, m = 0, \ldots, N$ in (2.9) we have to decompose the matrix A as follows:

$$\begin{aligned}
A &= \begin{pmatrix} ((1-\theta)p+\theta)s & (1-\theta)(1-p) \\ (1-\theta)ps & 1-(1-\theta)p \end{pmatrix} = \\
&\qquad \begin{pmatrix} 0 & (1-\theta)(1-p) \\ 0 & 1-(1-\theta)p \end{pmatrix} + \begin{pmatrix} ((1-\theta)p+\theta) & 0 \\ (1-\theta)p & 0 \end{pmatrix} s = U + Vs.
\end{aligned}$$

All summands of the $(N-1)$-th power of A can be generated by iterating the following binomial expression:

$$
\begin{aligned}
A^{N-1} &= (U+Vs)^{N-1} \\
&= \underbrace{(UU\dots U)}_{\binom{N-1}{0}\text{ summands}} + \underbrace{(UU\dots UV + UU\dots VU + \dots + VU\dots UU)}_{\binom{N-1}{1}\text{ summands}} s + \\
&\quad \underbrace{(UUU\dots UVV + UU\dots VUV + \dots + VVU\dots UUU)}_{\binom{N-1}{2}\text{ summands}} s^2 + \dots + \\
&\quad \underbrace{(UV\dots VV + VU\dots VV + \dots + VV\dots VU)}_{\binom{N-1}{N-2}\text{ summands}} s^{N-2} + \underbrace{(VV\dots V)}_{\binom{N-1}{N-1}\text{ summands}} s^{N-1}.
\end{aligned}
\tag{2.10}
$$

Comparing the coefficients in (2.8) and (2.10) we obtain:

$$
\begin{aligned}
C^{(0)} &= \underbrace{UU\dots U}_{\binom{N-1}{0}\text{ summands}} \\
C^{(1)} &= \underbrace{UU\dots UV + UU\dots VU + \dots + VU\dots UU}_{\binom{N-1}{1}\text{ summands}} \\
C^{(2)} &= \underbrace{UUU\dots UVV + UU\dots VUV + \dots + VVU\dots UUU}_{\binom{N-1}{2}\text{ summands}} \\
&\vdots \\
C^{(N-2)} &= \underbrace{VV\dots VU + VV\dots UV + \dots + UV\dots VV}_{\binom{N-1}{N-2}\text{ summands}} \\
C^{(N-1)} &= \underbrace{VV\dots V}_{\binom{N-1}{N-1}\text{ summands}}
\end{aligned}
$$

Note that m in $C^{(m)}$ corresponds to the number of times that the matrix V is in the products and it is also the number of ones in the sample. In order to generate all summands, we can iterate over all binary numbers with N bits:

$$\begin{pmatrix} U & U & U & \dots & U & U & U \\ \\ U & U & U & \dots & U & U & V \\ U & U & U & \dots & U & V & U \\ & & & \vdots & & & \\ V & U & U & \dots & U & U & U \\ \\ U & U & U & \dots & U & V & V \\ U & U & U & \dots & V & U & V \\ & & & \vdots & & & \\ V & V & U & \dots & U & U & U \\ \\ \vdots & & & \vdots & & & \vdots \\ \\ V & V & V & \dots & V & V & U \\ V & V & V & \dots & V & U & V \\ & & & \vdots & & & \\ U & V & V & \dots & V & V & V \\ \\ V & V & V & \dots & V & V & V \end{pmatrix}_{2^N\times(N-1)} \longmapsto \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ & & & \vdots & & & \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ \\ 0 & 0 & 0 & \dots & 0 & 1 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 & 1 \\ & & & \vdots & & & \\ 1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \\ \vdots & & & \vdots & & & \vdots \\ \\ 1 & 1 & 1 & \dots & 1 & 1 & 0 \\ 1 & 1 & 1 & \dots & 1 & 0 & 1 \\ & & & \vdots & & & \\ 0 & 1 & 1 & \dots & 1 & 1 & 1 \\ \\ 1 & 1 & 1 & \dots & 1 & 1 & 1 \end{pmatrix}_{2^N\times(N-1)}$$

The first row in the second matrix indicates that in $C^{(0)}$ the matrix U must be multiplied $(N-1)$ times. The following $\binom{N-1}{1}$ rows indicate that for $C^{(1)}$ there are $\binom{N-1}{1}$ summands, each one of them containing the product of $(N-2)$ Us and one V, and so on. These iterations are helpful to identify the summands to calculate $C^{(m)}$.

2.2.4. *Algorithm to Calculate the Power of the R_m Test and of the Modified R Test.* To compare the power of the modified R test with the power of the R_m test, we will calculate it explicitly for some values of θ, p, m and N with the following algorithm:

(1) Calculate the conditional probability distribution of R_m under the alternative as in (2.5), the probability distribution of R under the alternative as in (2.6), and the probability distribution of ζ using (2.9):
   (a) $P(\zeta = N) = \begin{pmatrix} p & 0 \end{pmatrix} C^{(N-1)} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
   (b) $P(\zeta = m) = \left[\begin{pmatrix} p & 0 \end{pmatrix} C^{(m-1)} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 & 1-p \end{pmatrix} C^{(m)} \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]$, for $m = 1, \dots, N-1$
   (c) $P(\zeta = 0) = \begin{pmatrix} 0 & 1-p \end{pmatrix} C^{(0)} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

   where the matrix $C^{(m)}$ for $m = 0, \dots, (N-1)$, is calculated with the algorithm 2.2.3.
(2) Calculate the conditional cumulative distributions of R_m and of R under the alternative.
(3) Fix significance level $\alpha = 0.05$ and find critical values c_m and c such that $P_0(R_m \le c_m \mid \zeta = m) \le \alpha$, and $P_0(R \le c) \le \alpha$ respectively.
(4) Randomize the R_m and R tests such that for $0 < \gamma < 1$ and $0 < \gamma' < 1$

$$\alpha = P_0(R_m \le c_m \mid \zeta = m) + \gamma\, P_0(c_m < R_m \le c_m + 1 \mid \zeta = m)$$

and

$\alpha = P_0(R \leq c) + \gamma' P_0(c < R \leq c+1)$

(5) Calculate the power of the randomized R_m and R tests as follows:

$\pi_{R_m}(\theta) = P_\theta(R_m \leq c_m \,|\, \zeta = m) + \gamma\, P_\theta(c_m < R_m \leq c_m + 1 \,|\, \zeta = m)$.

and

$\pi_R(\theta) = P_\theta(R \leq c) + \gamma' P_\theta(c < R \leq c+1)$.

respectively.

3. A Comparative Power Study and Main Results

We calculated the exact power of the R and R_m tests explicitly, for sample sizes[§] $N = 7(1)22, 30, 40, 50$, for $p = 0.1(0.05)0.9$ and for $\theta = 0(0.05)0.9$. The R_m test was compared with the R test for each value of p, for each N and each m. We have not included the extreme cases $m = 0$ and $m = N$ because for these the power of the R_m test is zero.

We show the main results for $N = 10(10)50$, for $\theta = 0(0.1)0.9$ and for some number of ones in the observed sequence obtained as percents ([N(10%)] and [N(20%)]) of the sample size, to find typical patterns of the powers of the compared tests[¶]. They are in Tables 1 to 15, ordered as follows: Tables 1, 4, 7, 10 and 13 contain the powers of the R test. The other tables are for the R_m test distinguished by the number of ones.

To facilitate the reading of the tables, we built three dimensional graphic illustrations, each containing five graphics denoted by $\pi(p, \theta)$ for each combination of p and θ and the five sample sizes considered. Intersections of the red lines are powers of the R test and the blue ones are powers of the R_m test for fixed values of m.

All figures and all graphics show that the powers of the R test increase with θ as expected, that the powers increase faster for values of p around 0.5 and that the speed of increase is lower when p tends to zero or to one. The power of the R_m test shows small decreases for values of p around 0.5.

In Figure 4, for example, with 10% ones in the observed sequence, for success probabilities p between 0.3 and 0.7, and for sample sizes $N = 10(10)50$, it can be noted that the R test is more powerful than the R_m test. We specially note that for $N = 10$, the same result occurs in a bit larger interval for $0.2 \leq p \leq 0.8$, as can be verified in Tables 1 and 2, with some exceptions for values of θ less than or equal to 0.5, where the power of the R_m test is greater than the power of the R test. The same situation occurs for the powers showed in figure 5 for 20% ones, but now it holds for a smaller interval of values of p: $0.4 \leq p \leq 0.6$.

In general, the interval of values of p for which the power of the R test overtakes the power of the R_m test is smaller when m increases up to 50% ones, the case in which the power of the R_m test overtakes the power of the R test for all values of p. From 50% to

---

[§] $N = a(step)b, c, d, \ldots$ means that N goes from a to b, jumping a "step" each time, and after value b, taking values c,d,...

[¶] All these calculations of this section are available on the web page *http://www.docentes.unal.edu.co/jacorzos/docs/*, document: WebPotencia_R_R_m.pdf.

100% ones, the interval of values of p for which the R test overtakes the R_m test increases.

For a fixed percent of ones, it can be observed that the length of the interval of values of p in which the power of the R test overtakes the power of the R_m test tends to be constant, for all compared sample sizes. It can be noted that the R test seems to be more powerful than the R_m test because the larger the correlation between observations, the fewer and larger are the runs.

In Tables 1, 4, 7, 10 and 13 we can see that the power of the R test increases with N around $p = 1/2$. On the other hand, the power of the R_m test increases with N and with m around $N/2$ when N is even, around $(N-1)/2$ and $(N+1)/2$ when N is odd [‖].

Although these types of Markov chains could seem rare, we highlight the conditions under which they can occur. In Figure 2, side (a), we see that P_{11} increases with θ and p, whilst P_{01} decreases with θ and increases with p. Moreover, in part (b) we see that P_{00} increases with θ and decreases with p, whilst P_{10} decreases with both θ and p. We can also see in part (a), that when θ increases it holds true that $P_{11} > P_{01}$, and P_{11} tends to be much larger than P_{01} when θ tends to one, and that implies few runs and large runs of ones. In part (b), the situation is analogous, but with the zeros instead of the ones.

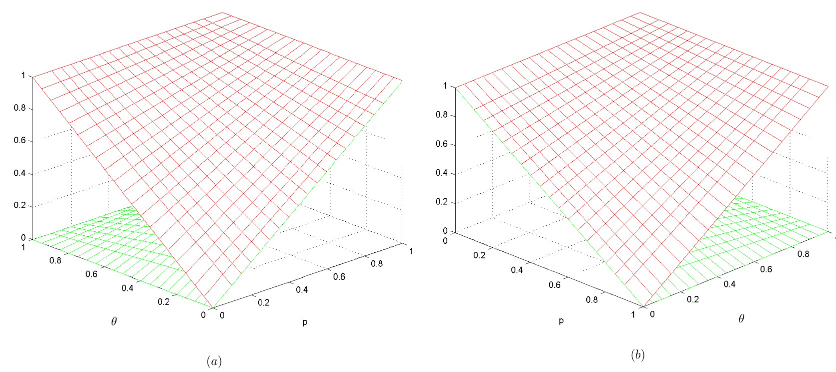


Figure 2. Positive association: (a) Transition Probabilities P_{11} (red) and P_{01}(green), (b) Transition Probabilities P_{00} (red) and P_{10}(green)

A Data Driven Test. As we said, the power of the R test is greater than the power of the R_m test for some intervals of values of the success probability p and for some values of m. This suggests the following selection process of the appropriate test for a fixed level $\alpha = 0.05$:

(1) Calculate the number of ones m in the observed sequence.
(2) Estimate the success probability p, by means of the [20] estimator[**] $\hat{p} = \frac{1}{N}\sum_{i=1}^{N} x_i$.
(3) Choose the data driven test as follows (see Figure 3):

[‖]See list of tables justifying this comments in our web page, in the cited document.

[**]Although $\hat{p}$ is the old well known plug-in estimator of p given by [21], we reference [20] because he shows, among other things, that the maximum likelihood estimator for $(\hat{p}, \hat{\theta})$ obtained under Markovian dependence, is strongly consistent for (p, θ).

Select the R test if:

- $\hat{p} \in [0.25, 0.75]$ and $m \in (0\%N, 13\%N) \cup (87\%N, 100\%N)$
- $\hat{p} \in [0.30, 0.70]$ and $m \in [13\%N, 17\%N) \cup (83\%N, 87\%N]$
- $\hat{p} \in [0.35, 0.65]$ and $m \in [17\%N, 23\%N) \cup (77\%N, 83\%N]$
- $\hat{p} \in [0.40, 0.60]$ and $m \in [23\%N, 33\%N) \cup (67\%N, 77\%N]$
- $\hat{p} \in [0.45, 0.55]$ and $m \in [33\%N, 43\%N) \cup (57\%N, 67\%N]$
- $\hat{p} = 0.50$ and $m \in [43\%N, 57\%N]$

Select the R_m test otherwise.

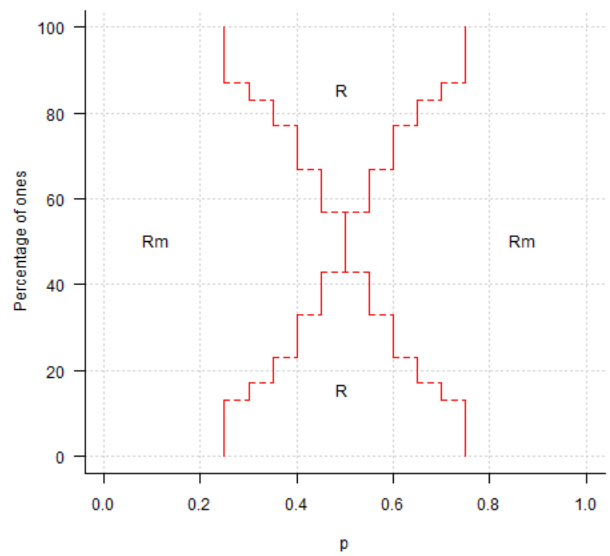


Figure 3. Regions to choose between R and R_m Tests

4. Example

A first order homogeneous Markov Chain (fohMC) can be used to describe the behavior of a buyer as follows[††]: a buyer at a supermarket A switches to buying in a supermarket B on her/his next shopping trip with probability $\lambda > 0$, while she/he switches to supermarket A with probability $\beta > 0$ when her/his last shopping was in supermarket B.

To verify the Markovian assumption of the behavior of the buyer, we simulated first order homogeneous Markov Chains using the Metropolis Hasting Algorithm, for some values of the success probability p and the correlation between successive observations θ. These combinations of values p and θ give values of the transition probabilities λ, β which indicate how Markovian the behavior of the buyer is.

[††]Adapted from [22]

For each combination of (p,θ), $p, \theta = 0.1(0.1)0.9$, we simulate 1000 fohMC and we select those with not all zeros and not all ones, because such stable buyers are not interesting. For the remaining fohMC, we calculate the proportion of rejections of the null hypothesis, and the ratio $(1-\lambda)/\beta$ to find the conditions for which the fohMC is a good assumption for the behavior of the consumer.

The results are in Table 16 for $N = 20$. It can be seen that the empirical powers of the test increase up to 87% for $p = 0.4$ and $\theta = 0.9$. It can be noted also that the ratio $p_{00}/p_{10} = (1-\lambda)/\beta$ grows also with θ, as expected.

The fact that the larger the values of the ratio $(1-\lambda)/\beta$, the greater the empirical power of the runs test, indicates that the behavior of the consumer tends to be most Markovian when the probability of continuing shopping in supermarket A is larger than the probability of switching to supermarket B.

5. Conclusions and Discussion

We have discovered regions of values of m and p where the R test is more powerful than the R_m test, and we have included additional information about p and m to the test statistic, to produce a data driven test which covers the complete $p \times m$ region, and improves the power of the test.

The power of the R test increases with N and especially around $p = 1/2$, while the power of the R_m test increases with N and with m around $N/2$, for N even or $(N-1)/2$ and $(N+1)/2$ for N odd.

Although the proposed data driven test includes information about the length of the runs without being explicit (few runs implies long runs in most cases), a way to improve the power of the test could be to include the length of the runs explicitly, and to use the results of [15] or of [23] about the distribution of the longest runs test.

Acknowledgements: We are thankful to the statistician Gustavo Romero for developing the R code for the algorithm to obtain the coefficients of the pgf.

References

[1] F.N. David. A power function for tests of randomness in a sequence of alternatives. *Biometrika*, 34:335–339, 1947.
[2] G. Bateman. On the power function of the longest run as a test for randomness in a sequence of alternatives. *Biometrika*, 35:97–112, 1948.
[3] D. Barton and F. David. Non-randomness in a sequence of two alternatives. *Biometrika*, 45:253–256, 1958.
[4] P. Shaughnessy. Multiple runs distributions: recurrences and critical values. *Journal of the American Statistical Association*, 76(375):732–736, 1981.
[5] S. Schwager. Run probabilities in sequences of Markov dependent trials. *Journal of the American Statistical Association*, 78:168–175, 1983.
[6] M. Koutras and V. Alexandrou. Non-parametric randomness test based on success runs of fixed length. *Statistic and Probability Letters*, 32:393–404, 1997.
[7] D. Antzoulakos, S. Bersimis, and M. Koutras. On the distribution of the total number of run lengths. *Annals of the Institute of Statistical Mathematics,*, 55(4):865–884, 2003.
[8] M. Koutras. Waiting time distributions associated with runs of fixed length in two-state Markov chains. *Annals of the Institute of Statistical Mathematics*, 49:123–139, 1997.

[9] D.L. Antzoulakos and S. Chadjiconstantinidis. Distributions of numbers of success runs of fixed length in Markov dependent trials. *Annals of the Institute of Statistical Mathematics*, 53(3):599–619, 2001.

[10] J. Fu, W. Lou, Z. Bai, and G. Li. The exact and limiting distributions of the number of successes in success runs within a sequence of Markov-dependent two-state trial. *Annals of the Institute of Statistical Mathematics*, 54:719–730, 2002.

[11] J. Fou, L. Want, and W. Lou. On exact and large deviation approximation for the distribution of the longest run in a sequence of two-state Markov dependent trials. *Journal of Applied Probability*, 40:346–360, 2003.

[12] E. Vaggelatou. On the length of the longest run in a multistate Markov chain. *Statistics and Probability Letters*, 62:211–221, 2003.

[13] L. Savelyev and V. Balakin. The joint distribution of the number of ones and the number of 1-runs in binary markov sequences. *Discrete Math. Appl.*, 14(4):353–372, 2004.

[14] S. Eryilmaz. On the distribution and expectation of success runs in nonhomogeneous Markov dependent trials. *Statistical Papers*, 46:117–128, 2005.

[15] S. Eryilmaz. Some results associated with the longest run statistic in a sequence of Markov dependent trials. *Applied Mathematics and Computation*, 175:119–130, 2006.

[16] S. Demir and S. Erylmaz. Run statistics in a sequence of arbitrarily dependent binary trials. *Statistical Papers. Online first*, 2008.

[17] S. Eryilmaz. Mean success run length. *Journal of the Korean Statistical Society*, 38:65–71, 2009.

[18] S. Eryilmaz and F. Yalcin. Distribution of run statistics in partially exchangeable processes. *Metrica*, 73:293–304, 2011.

[19] A.W.F. Edwards. The meaning of binomial distribution. *Nature*, 186:1074, 1960.

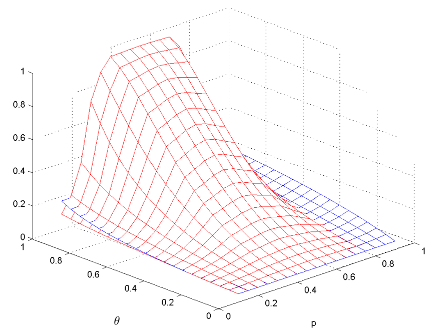
[20] B. Lindqvist. A note on bernoulli trials with dependence. *Scand Journal Statist*, 5:205–208, 1978.

[21] P. Laplace. *Théorie Analytique des Probabilités.* Paris, Ve. Courcier, Paris, seconde edition, 1814.

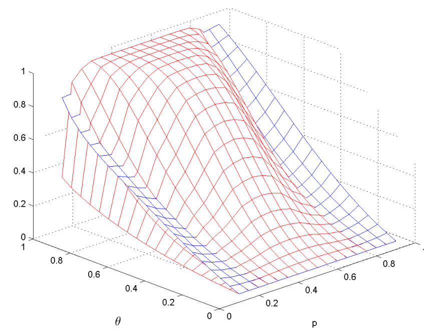
[22] W. Ching and M. Ng. *Markov Chains: Models, Algorithms and Applications.* Springer Science + Business Media, New York, 2006.

[23] F. Beaujean and A. Caldwell. A test statistic for weighted runs. *Journal of Statistical Planning and Inference*, 141:3437–3446, 2011.

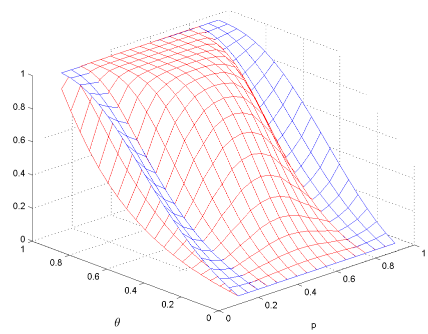
Appendix A. Figures of the Power of the Proposed Test



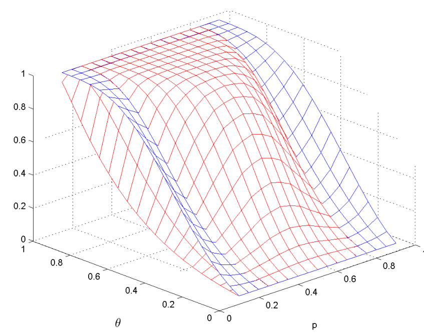
N=10 and m=1



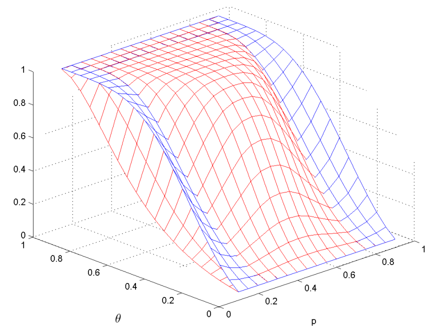
N=20 and m=2



N=30 and m=3

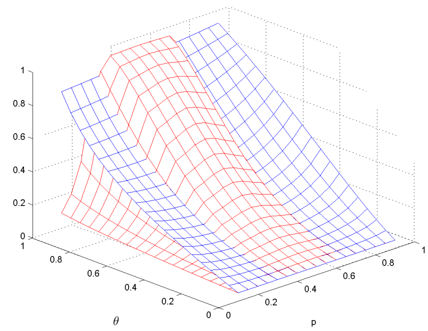


N=40 and m=4

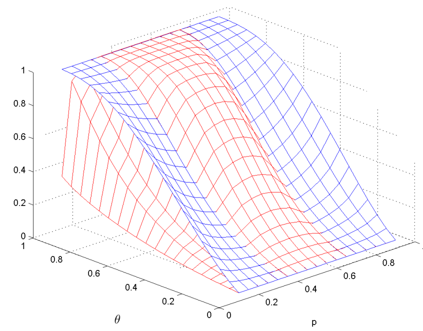


N=50 and m=5

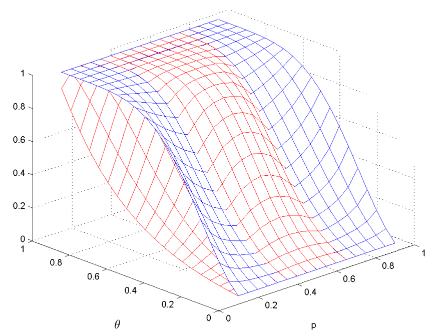
Figure 4. $\pi_R(p,\theta)$ (red) vs. $\pi_{R_m}(p,\theta)$ (blue) for $N = 10, 20, 30, 40, 50$ with 10% ones



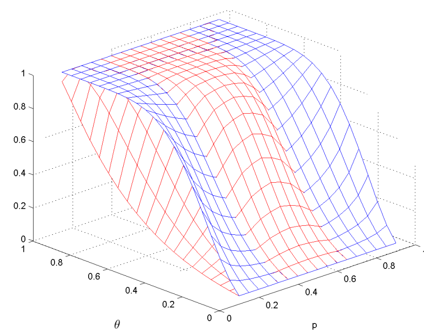
N=10 and m=2



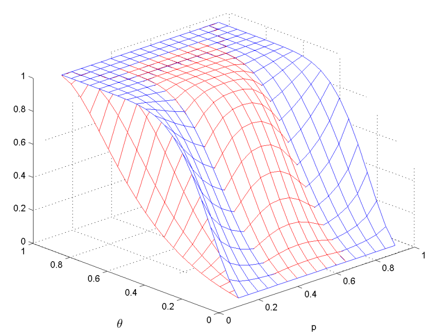
N=20 and m=4



N=30 and m=6



N=40 and m=8



N=50 and m=10

Figure 5. $\pi_R(p,\theta)$ (red) vs. $\pi_{R_m}(p,\theta)$ (blue) for $N = 10, 20, 30, 40, 50$ with 20% ones

Appendix B. Tables of the Power of the Proposed Test

Table 1. Power of the R Test, N=10

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.0552	0.0624	0.0729	0.0815	0.0866	0.0815	0.0729	0.0624	0.0552
0.20	0.0609	0.0776	0.1047	0.1274	0.1404	0.1274	0.1047	0.0776	0.0609
0.30	0.0672	0.0959	0.1480	0.1918	0.2148	0.1918	0.1480	0.0959	0.0672
0.40	0.0740	0.1182	0.2059	0.2780	0.3116	0.2780	0.2059	0.1182	0.0740
0.50	0.0814	0.1453	0.2821	0.3879	0.4305	0.3879	0.2821	0.1453	0.0814
0.60	0.0896	0.1788	0.3797	0.5205	0.5671	0.5205	0.3797	0.1788	0.0896
0.70	0.0990	0.2213	0.5016	0.6692	0.7120	0.6692	0.5016	0.2213	0.0990
0.80	0.1100	0.2774	0.6485	0.8196	0.8495	0.8196	0.6485	0.2774	0.1100
0.90	0.1240	0.3546	0.8176	0.9447	0.9560	0.9447	0.8176	0.3546	0.1240

Table 2. Power of the R_m Test, N=10 and m=1

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.0548	0.0554	0.0562	0.0571	0.0585	0.0605	0.0638	0.0700	0.0864
0.20	0.0605	0.0618	0.0633	0.0654	0.0682	0.0722	0.0786	0.0900	0.1167
0.30	0.0674	0.0694	0.0718	0.0750	0.0793	0.0853	0.0944	0.1100	0.1423
0.40	0.0758	0.0786	0.0820	0.0864	0.0921	0.1000	0.1115	0.1300	0.1643
0.50	0.0864	0.0900	0.0944	0.1000	0.1071	0.1167	0.1300	0.1500	0.1833
0.60	0.1000	0.1045	0.1100	0.1167	0.1250	0.1357	0.1500	0.1700	0.2000
0.70	0.1183	0.1237	0.1300	0.1375	0.1466	0.1577	0.1717	0.1900	0.2147
0.80	0.1441	0.1500	0.1567	0.1643	0.1731	0.1833	0.1955	0.2100	0.2278
0.90	0.1833	0.1885	0.1940	0.2000	0.2065	0.2136	0.2214	0.2300	0.2395

Table 3. Power of the R_m Test, N=10 and m=2

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.0983	0.0816	0.0763	0.0745	0.0749	0.0773	0.0827	0.0947	0.1307
0.20	0.1496	0.1216	0.1114	0.1083	0.1096	0.1153	0.1276	0.1540	0.2255
0.30	0.2042	0.1707	0.1573	0.1535	0.1564	0.1664	0.1869	0.2281	0.3271
0.40	0.2628	0.2298	0.2158	0.2127	0.2183	0.2334	0.2623	0.3165	0.4316
0.50	0.3272	0.3001	0.2887	0.2882	0.2976	0.3182	0.3547	0.4178	0.5361
0.60	0.4008	0.3834	0.3779	0.3822	0.3964	0.4222	0.4637	0.5296	0.6386
0.70	0.4891	0.4836	0.4862	0.4967	0.5158	0.5450	0.5875	0.6484	0.7374
0.80	0.6020	0.6078	0.6184	0.6343	0.6561	0.6848	0.7219	0.7696	0.8312
0.90	0.7586	0.7701	0.7837	0.7994	0.8174	0.8380	0.8616	0.8884	0.9190

Table 4. Power of the R Test, N=20

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.0617	0.0736	0.0871	0.1091	0.1091	0.1049	0.0871	0.0736	0.0617
0.20	0.0759	0.1068	0.1451	0.2086	0.2086	0.1971	0.1451	0.1068	0.0759
0.30	0.0932	0.1528	0.2305	0.3525	0.3525	0.3323	0.2305	0.1528	0.0932
0.40	0.1142	0.2155	0.3482	0.5299	0.5299	0.5034	0.3482	0.2155	0.1142
0.50	0.1397	0.2994	0.4971	0.7126	0.7126	0.6861	0.4971	0.2994	0.1397
0.60	0.1704	0.4097	0.6650	0.8635	0.8635	0.8440	0.6650	0.4097	0.1704
0.70	0.2076	0.5504	0.8255	0.9565	0.9565	0.9470	0.8255	0.5504	0.2076
0.80	0.2531	0.7205	0.9428	0.9933	0.9933	0.9909	0.9428	0.7205	0.2531
0.90	0.3126	0.8988	0.9940	0.9998	0.9998	0.9997	0.9940	0.8988	0.3126

Table 5. Power of the R_m Test, N=20 and m=2

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.1025	0.0815	0.0747	0.0715	0.0717	0.0725	0.0759	0.0840	0.1092
0.20	0.1631	0.1223	0.1079	0.1010	0.1017	0.1035	0.1112	0.1292	0.1818
0.30	0.2302	0.1739	0.1520	0.1412	0.1426	0.1457	0.1587	0.1880	0.2674
0.40	0.3017	0.2372	0.2094	0.1954	0.1979	0.2027	0.2217	0.2627	0.3649
0.50	0.3750	0.3121	0.2819	0.2668	0.2712	0.2780	0.3033	0.3552	0.4723
0.60	0.4489	0.3973	0.3703	0.3582	0.3659	0.3748	0.4062	0.4658	0.5862
0.70	0.5252	0.4922	0.4747	0.4713	0.4838	0.4951	0.5309	0.5925	0.7018
0.80	0.6119	0.6004	0.5971	0.6073	0.6254	0.6385	0.6750	0.7299	0.8133
0.90	0.7351	0.7414	0.7511	0.7728	0.7929	0.8049	0.8334	0.8694	0.9146

Table 6. Power of the R_m Test, N=20 and m=4

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.1496	0.1073	0.0940	0.0879	0.0883	0.0898	0.0961	0.1116	0.1614
0.20	0.2784	0.1912	0.1608	0.1463	0.1475	0.1511	0.1667	0.2034	0.3089
0.30	0.4211	0.3014	0.2544	0.2309	0.2333	0.2396	0.2659	0.3244	0.4702
0.40	0.5656	0.4338	0.3754	0.3449	0.3489	0.3579	0.3938	0.4679	0.6265
0.50	0.7011	0.5794	0.5191	0.4866	0.4922	0.5030	0.5439	0.6213	0.7631
0.60	0.8178	0.7244	0.6735	0.6456	0.6522	0.6629	0.7010	0.7668	0.8700
0.70	0.9078	0.8516	0.8185	0.8007	0.8068	0.8151	0.8425	0.8852	0.9426
0.80	0.9665	0.9437	0.9297	0.9229	0.9267	0.9310	0.9439	0.9619	0.9827
0.90	0.9945	0.9908	0.9885	0.9878	0.9888	0.9897	0.9920	0.9949	0.9979

Table 7. Power of the R Test, N=30

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.0689	0.0826	0.1025	0.1294	0.1294	0.1229	0.1025	0.0826	0.0689
0.20	0.0946	0.1322	0.1922	0.2724	0.2724	0.2537	0.1922	0.1322	0.0946
0.30	0.1294	0.2048	0.3280	0.4739	0.4739	0.4425	0.3280	0.2048	0.1294
0.40	0.1765	0.3058	0.5055	0.6918	0.6918	0.6564	0.5055	0.3058	0.1765
0.50	0.2398	0.4380	0.6985	0.8661	0.8661	0.8395	0.6985	0.4380	0.2398
0.60	0.3249	0.5972	0.8625	0.9623	0.9623	0.9500	0.8625	0.5972	0.3249
0.70	0.4388	0.7666	0.9610	0.9946	0.9946	0.9917	0.9610	0.7666	0.4388
0.80	0.5907	0.9114	0.9954	0.9998	0.9998	0.9996	0.9954	0.9114	0.5907
0.90	0.7892	0.9889	0.9999	1.0000	1.0000	1.0000	0.9999	0.9889	0.7892

Table 8. Power of the R_m Test, N=30 and m=3

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.1127	0.0878	0.0802	0.0772	0.0782	0.0796	0.0851	0.0979	0.1389
0.20	0.1959	0.1420	0.1245	0.1177	0.1204	0.1239	0.1375	0.1685	0.2597
0.30	0.3011	0.2174	0.1884	0.1771	0.1822	0.1886	0.2127	0.2649	0.4016
0.40	0.4271	0.3185	0.2779	0.2619	0.2701	0.2799	0.3155	0.3877	0.5518
0.50	0.5677	0.4474	0.3980	0.3784	0.3897	0.4027	0.4477	0.5322	0.6964
0.60	0.7110	0.5994	0.5485	0.5283	0.5415	0.5561	0.6040	0.6861	0.8215
0.70	0.8395	0.7587	0.7176	0.7013	0.7138	0.7267	0.7670	0.8292	0.9156
0.80	0.9356	0.8963	0.8746	0.8665	0.8747	0.8825	0.9052	0.9365	0.9728
0.90	0.9881	0.9802	0.9757	0.9746	0.9770	0.9790	0.9842	0.9904	0.9964

Table 9. Power of the R_m Test, N=30 and m=6

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.2050	0.1341	0.1128	0.1028	0.1029	0.1047	0.1132	0.1348	0.2057
0.20	0.4221	0.2709	0.2181	0.1918	0.1919	0.1965	0.2182	0.2705	0.4183
0.30	0.6386	0.4498	0.3688	0.3247	0.3245	0.3321	0.3673	0.4461	0.6295
0.40	0.8099	0.6417	0.5515	0.4966	0.4958	0.5053	0.5476	0.6345	0.7992
0.50	0.9193	0.8091	0.7349	0.6837	0.6824	0.6912	0.7294	0.8008	0.9112
0.60	0.9744	0.9235	0.8806	0.8468	0.8455	0.8512	0.8758	0.9175	0.9703
0.70	0.9947	0.9801	0.9650	0.9511	0.9504	0.9527	0.9625	0.9777	0.9935
0.80	0.9995	0.9976	0.9952	0.9927	0.9925	0.9930	0.9947	0.9972	0.9993
0.90	1.0000	1.0000	0.9999	0.9998	0.9998	0.9998	0.9999	0.9999	1.0000

Table 10. Power of the R Test, N=40

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.0715	0.0910	0.1138	0.1403	0.1531	0.1403	0.1138	0.0910	0.0715
0.20	0.1013	0.1576	0.2285	0.3099	0.3464	0.3099	0.2285	0.1576	0.1013
0.30	0.1421	0.2579	0.4014	0.5442	0.5986	0.5442	0.4014	0.2579	0.1421
0.40	0.1975	0.3965	0.6126	0.7743	0.8221	0.7743	0.6126	0.3965	0.1975
0.50	0.2716	0.5674	0.8100	0.9259	0.9499	0.9259	0.8100	0.5674	0.2716
0.60	0.3690	0.7476	0.9394	0.9866	0.9926	0.9866	0.9394	0.7476	0.3690
0.70	0.4947	0.8971	0.9904	0.9990	0.9996	0.9990	0.9904	0.8971	0.4947
0.80	0.6532	0.9793	0.9996	1.0000	1.0000	1.0000	0.9996	0.9793	0.6532
0.90	0.8446	0.9993	1.0000	1.0000	1.0000	1.0000	1.0000	0.9993	0.8446

Table 11. Power of the R_m Test, N=40 and m=4

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.1372	0.1003	0.0893	0.0854	0.0851	0.0878	0.0947	0.1111	0.1637
0.20	0.2647	0.1787	0.1514	0.1414	0.1404	0.1471	0.1644	0.2040	0.3185
0.30	0.4234	0.2912	0.2441	0.2261	0.2241	0.2357	0.2654	0.3300	0.4907
0.40	0.5949	0.4375	0.3727	0.3464	0.3429	0.3592	0.4003	0.4831	0.6574
0.50	0.7548	0.6064	0.5344	0.5026	0.4978	0.5165	0.5625	0.6479	0.7988
0.60	0.8798	0.7728	0.7108	0.6807	0.6753	0.6920	0.7323	0.8007	0.9024
0.70	0.9572	0.9039	0.8667	0.8466	0.8422	0.8526	0.8774	0.9160	0.9646
0.80	0.9912	0.9767	0.9646	0.9572	0.9552	0.9587	0.9671	0.9792	0.9924
0.90	0.9995	0.9984	0.9974	0.9967	0.9965	0.9968	0.9975	0.9985	0.9995

Table 12. Power of the R_m Test, N=40 and m=8

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.2286	0.1478	0.1239	0.1149	0.1131	0.1172	0.1293	0.1592	0.2579
0.20	0.4804	0.3124	0.2530	0.2296	0.2250	0.2359	0.2675	0.3407	0.5350
0.30	0.7150	0.5226	0.4364	0.3993	0.3918	0.4098	0.4594	0.5615	0.7663
0.40	0.8769	0.7300	0.6450	0.6041	0.5957	0.6163	0.6696	0.7651	0.9079
0.50	0.9606	0.8847	0.8278	0.7969	0.7903	0.8066	0.8457	0.9057	0.9731
0.60	0.9916	0.9670	0.9433	0.9287	0.9254	0.9335	0.9513	0.9748	0.9947
0.70	0.9990	0.9949	0.9899	0.9863	0.9855	0.9876	0.9917	0.9964	0.9994
0.80	1.0000	0.9997	0.9994	0.9991	0.9991	0.9992	0.9995	0.9998	1.0000
0.90	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 13. Power of the R Test, N=50

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.0766	0.0992	0.1249	0.1678	0.1678	0.1565	0.1249	0.0992	0.0766
0.20	0.1156	0.1831	0.2645	0.3936	0.3936	0.3618	0.2645	0.1831	0.1156
0.30	0.1716	0.3118	0.4708	0.6719	0.6719	0.6292	0.4708	0.3118	0.1716
0.40	0.2499	0.4846	0.7018	0.8841	0.8841	0.8539	0.7018	0.4846	0.2499
0.50	0.3559	0.6800	0.8839	0.9773	0.9773	0.9668	0.8839	0.6800	0.3559
0.60	0.4926	0.8535	0.9743	0.9981	0.9981	0.9966	0.9743	0.8535	0.4926
0.70	0.6568	0.9603	0.9977	1.0000	1.0000	0.9999	0.9977	0.9603	0.6568
0.80	0.8304	0.9961	1.0000	1.0000	1.0000	1.0000	1.0000	0.9961	0.8304
0.90	0.9659	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9659

Table 14. Power of the R_m Test, N=50 and m=5

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.1703	0.1159	0.0995	0.0918	0.0919	0.0933	0.1000	0.1167	0.1715
0.20	0.3459	0.2221	0.1808	0.1606	0.1607	0.1643	0.1813	0.2225	0.3438
0.30	0.5451	0.3706	0.3020	0.2660	0.2659	0.2722	0.3013	0.3682	0.5366
0.40	0.7298	0.5494	0.4629	0.4131	0.4125	0.4210	0.4598	0.5426	0.7166
0.50	0.8697	0.7303	0.6472	0.5934	0.5921	0.6011	0.6413	0.7202	0.8569
0.60	0.9530	0.8766	0.8192	0.7767	0.7750	0.7820	0.8125	0.8672	0.9449
0.70	0.9891	0.9632	0.9386	0.9173	0.9161	0.9195	0.9342	0.9581	0.9861
0.80	0.9988	0.9949	0.9902	0.9854	0.9850	0.9858	0.9890	0.9938	0.9983
0.90	1.0000	0.9999	0.9997	0.9996	0.9996	0.9996	0.9997	0.9998	1.0000

Table 15. Power of the R_m Test, N=50 and m=10

θ	p=0.10	p =0.20	p=0.30	p=0.40	p=0.50	p=0.60	p=0.70	p=0.80	p=0.90
0.00	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
0.10	0.2674	0.1660	0.1357	0.1241	0.1210	0.1246	0.1371	0.1690	0.2761
0.20	0.5604	0.3630	0.2890	0.2584	0.2503	0.2602	0.2933	0.3717	0.5787
0.30	0.7963	0.5992	0.4998	0.4536	0.4410	0.4569	0.5071	0.6121	0.8142
0.40	0.9293	0.8048	0.7195	0.6740	0.6610	0.6781	0.7279	0.8169	0.9397
0.50	0.9828	0.9326	0.8862	0.8575	0.8490	0.8608	0.8923	0.9396	0.9865
0.60	0.9974	0.9857	0.9713	0.9609	0.9577	0.9624	0.9738	0.9880	0.9982
0.70	0.9998	0.9986	0.9966	0.9949	0.9944	0.9952	0.9970	0.9989	0.9999
0.80	1.0000	1.0000	0.9999	0.9998	0.9998	0.9998	0.9999	1.0000	1.0000
0.90	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Appendix C. Results of the Example

Table 16. Estimated power π (first entry for each value of p) and Ratio $(1-\lambda)/\beta$ second entry for the same value of p), for the simulate Markov Chains

		θ								
	N=20	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
p	0.1	0.04	0.08	0.19	0.25	0.38	0.49	0.61	0.73	0.78
		1.12	1.28	1.48	1.74	2.11	2.67	3.59	5.44	11
	0.2	0.05	0.11	0.19	0.30	0.42	0.55	0.63	0.77	0.84
		1.14	1.31	1.54	1.83	2.25	2.88	3.92	6	12.25
	0.3	0.06	0.12	0.21	0.35	0.47	0.61	0.74	0.78	0.85
		1.16	1.36	1.61	1.95	2.43	3.14	4.33	6.71	13.86
	0.4	0.06	0.14	0.20	0.39	0.51	0.67	0.75	0.82	0.87
		1.19	1.42	1.71	2.11	2.67	3.5	4.89	7.67	16
	0.5	0.06	0.12	0.24	0.37	0.50	0.67	0.75	0.83	0.84
		1.22	1.5	1.86	2.33	3	4	5.67	9	19
	0.6	0.08	0.13	0.21	0.35	0.52	0.68	0.76	0.82	0.85
		1.28	1.63	2.07	2.67	3.5	4.75	6.83	11	23.5
	0.7	0.06	0.13	0.23	0.32	0.51	0.58	0.73	0.78	0.83
		1.37	1.83	2.43	3.22	4.33	6	8.78	14.33	31
	0.8	0.04	0.11	0.21	0.30	0.43	0.58	0.66	0.76	0.81
		1.56	2.25	3.14	4.33	6	8.5	12.67	21	46
	0.9	0.05	0.08	0.16	0.23	0.35	0.51	0.58	0.69	0.81
		2.11	3.5	5.29	7.67	11	16	24.33	41	91

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 541 – 548

Estimation and orthogonal block structure

Sandra S. Ferreira[∗†] , Célia Nunes[‡] , Dário Ferreira[§] , Elsa Moreira[¶] and João Tiago Mexia[‖]

Abstract

Estimators with good behaviors for estimable vectors and variance components are obtained for a class of models that contains the well known models with orthogonal block structure, OBS, see [15], [16] and [1], [2]. The study observations of these estimators uses commutative Jordan Algebras, CJA, and extends the one given for a more restricted class of models, the models with commutative orthogonal block structure, COBS, in which the orthogonal projection matrix on the space spanned by the means vector commute with all variance-covariance matrices, see [7].

Keywords: BLUE, LSE, OBS, UMVUE, Variance components.

2000 AMS Classification: 62J10, C2K10

Received : 09.05.2014 *Accepted* : 17.03.2015 *Doi* : 10.15672/HJMS.2015589756

[∗]Department of Mathematics and Center of Mathematics, University of Beira Interior, Covilhã, Portugal, Email: sandraf@ubi.pt
[†]Corresponding Author.
[‡]Department of Mathematics and Center of Mathematics, University of Beira Interior, Covilhã, Portugal, Email: celian@ubi.pt
[§]Department of Mathematics and Center of Mathematics, University of Beira Interior, Covilhã, Portugal, Email: dario@ubi.pt
[¶]Department of Mathematics and Center of Mathematics and Applications, FCT/UNL, Monte da Caparica, Portugal, Email: efnm@fct.unl.pt
[‖]Department of Mathematics and Center of Mathematics and Applications, FCT/UNL, Monte da Caparica, Portugal, Email: jtm@fct.unl.pt

1. Introduction

Models with orthogonal block structure, OBS, are mixed models with the family $\nu = \left\{\sum_{j=1}^{m} \gamma_j Q_j; \quad \boldsymbol{\gamma} \in \mathbb{R}_{+}^{m}\right\}$, of variance-covariance matrices where the $\boldsymbol{Q}_1, ..., \boldsymbol{Q}_m$ are pairwise orthogonal orthogonal projection matrices, POOPM, summing to the identity matrix, I_n. These designs were introduced by [15] and [16], and continue to play an important part in the theory of randomized block designs, see for instance [1] and [2]. Refer to [9] and [18] for historical developments of the mixed model. The inference for these models is centered on the estimation of treatment contrasts, see [10]. These estimators are obtained from the orthogonal projections of the observation vector, $\boldsymbol{Y}$, on the strata which are the range spaces $\bigtriangledown_1, ..., \bigtriangledown_m$, of the $\boldsymbol{Q}_1, ..., \boldsymbol{Q}_m$. Namely the problem of obtaining estimators from more than one strata has been dealt in detail. Then the weights to be given to each strata have to be estimated, see again [10].
We intend to follow a different approach using commutative Jordan algebras, CJA, to study the algebraic structure of these models. CJA are useful in discussing the algebraic structures of the models in a way that is convenient for deriving estimators both of variance components and estimable vectors through the introduction of sub-vectors. For our purpose it is convenient to write the mixed model as

$$\boldsymbol{Y} = \sum_{i=0}^{w} \boldsymbol{X}_i \boldsymbol{\beta}_i, \tag{1.1}$$

where $\boldsymbol{\beta}_0$ is fixed and $\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_w$ are random independent with null mean vectors and cross covariance matrices as well variance-covariance matrices $\theta_1 \boldsymbol{I}_{g_1}, ..., \theta_m \boldsymbol{I}_{g_w}$. This formulation enables an easy characterization of mixed models with OBS. Then when matrices $\boldsymbol{M}_i = \boldsymbol{X}_i \boldsymbol{X}_i^{'}, i = 1, ..., w$, commute they generate, as we will see, the CJA $\mathcal{A}(\underline{M})$. This is the smallest CJA of symmetric matrices that contains $\underline{M} = \{\boldsymbol{M}_1, ..., \boldsymbol{M}_w\}$. We recall, see [12], that these algebras are linear subspaces constituted by symmetric matrices and containing their squares. We will show that when matrices of $\underline{M}$ commute and constitute a basis for $\mathcal{A}(\underline{M})$ the models has OBS. Then we may use the sub-models $\boldsymbol{Y}_j = A_j \boldsymbol{Y}, j = 1, ..., m$, to obtain estimators for estimable vectors that are BLUE whatever the variance components. Following [21] we say that this estimators are uniformly BLUE, UBLUE. They are quite distinct from the ones for contrasts which are weighted means with estimated weights. Now no weight estimation is required and all estimable vectors may be treated as an unified approach. We point out that estimable contrasts are uni-dimensional estimable vectors so we have a widening of the class of estimable parameters and results that does not depend on weight and, as we shall see, have optimal properties.

Moreover we also obtain, using the sub-models, estimators for variance components which, when quasi-normality is assumed, also have optimal properties.

The role played by the CJA rests on the obtention of the sub-models which have variance-covariance matrices $\gamma_j \boldsymbol{I}_{g_j}$, with $g_j = rank(\boldsymbol{Q}_j), j = 1, ..., m$. The homoscedasticity of these sub-vectors leads to optimal estimators derived from each strata. Then the cross covariance matrices, $\Sigma(\boldsymbol{Y}_j; \boldsymbol{Y}_j^{'})$, are null which are the combinations of estimators derived from different sub-vectors. We will also consider a special class of models with OBS, the commutative orthogonal block structure, COBS, in which $\boldsymbol{T}$, the orthogonal projection matrix on the space Ω spanned by the mean vector commutes with the matrices in principal basis of a CJA $\mathcal{A}$, $pb(\mathcal{A})$. Then, whatever the $\gamma_1, ..., \gamma_m$, the matrix $\boldsymbol{T}$

will commute with

$$(1.2) \qquad \boldsymbol{V} = \sum_{j=1}^{m} \gamma_j \boldsymbol{Q}_j,$$

which, see [23], ensures that whatever the estimable vector $\boldsymbol{\psi}$ it's least square estimator, LSE, is the Best linear unbiased estimator, BLUE. We will say, see [21], that models with COBS have LSE that are UBLUE and show that, for theses models, the LSE are identical with the estimators we obtained for the general case of models with OBS.

2. Commutative Jordan Algebras

We already refer the importance of CJA in these models. We now point out that, see [17], the matrices of $\underline{M}$ commute if and only if they are diagonalized by the same orthogonal matrix $\boldsymbol{P}$. Then $\underline{M}$ will be contained in the CJA $\mathcal{A}(\boldsymbol{P})$ constituted by the matrices diagonalized by $\boldsymbol{P}$, thus $\underline{M}$ is contained in a CJA if and only if it's matrices commute. Since intersecting CJA gives a CJA, the intersection $\mathcal{A}(\underline{M})$ of all CJA containing $\underline{M}$ will be the least CJA containing $\underline{M}$, so we say that it is generated by $\underline{M}$.

[19] showed that any CJA, $\mathcal{A}$, has an unique basis, the $pb(\mathcal{A})$ of $\mathcal{A}$, constituted by POOPM. As stated by [5], Jordan algebras are used to present normal orthogonal models in a canonical form. Moreover:

(1) any family of POOPM is the principal basis of the CJA constituted by their linear combination;
(2) any orthogonal projection matrix, OPM, belonging to a CJA, $\mathcal{A}$, will be sum of matrices in $pb(\mathcal{A})$;
(3) if the matrices in $pb(\mathcal{A}_1)$ are some of matrices in $pb(\mathcal{A}_2)$ we have $\mathcal{A}_1 \subset \mathcal{A}_2$.

We recall that the product of two symmetric matrices is symmetric if they commute, then the product of two OPM that commute will be an OPM since it is symmetric and it is idempotent.

Given an OPM $\boldsymbol{K}$ that commutes with the matrices of $\underline{K} = \{\boldsymbol{K}_1, ..., \boldsymbol{K}_m\} = pb(\mathcal{A})$, the non null matrices $\boldsymbol{K}\boldsymbol{K}_j$ and $\boldsymbol{K}^c\boldsymbol{K}_j, j = 1, ..., m$, with $\boldsymbol{K}^c = \boldsymbol{I}_n - \boldsymbol{K}$, will be POOPM thus constituting the principal basis of a CJA, $\overline{\mathcal{A}}$. We can order the matrices in $pb(\overline{\mathcal{A}})$ so that the first are products by $\boldsymbol{K}$ of matrices in $pb(\mathcal{A})$ and the last $\overline{m} - z$ will be products by $\boldsymbol{K}^c$ also of matrices in $pb(\mathcal{A})$. Clearly we have $\mathcal{A} \subset \overline{\mathcal{A}}$. Those pairs of CJA appear in the theory of models with COBS. Models with this structure was also studied in [11], [5], [6] and [8]. $\mathcal{A}$ is now the CJA with principal basis $\boldsymbol{Q} = \{\boldsymbol{Q}_1, ..., \boldsymbol{Q}_m\}$ when $\nu(\nabla)$ is the family of variance-covariance matrices and $\boldsymbol{T}$ playing the part of $\boldsymbol{K}$.

Let $\boldsymbol{\mu} = \boldsymbol{X}_0\boldsymbol{\beta}_0$ be the mean vector of the model. With $\boldsymbol{Q} = \{\boldsymbol{Q}_1, ..., \boldsymbol{Q}_m\} = pb(\mathcal{A}(\underline{M}))$ let the row vectors of $\boldsymbol{A}_j$ constitute an orthonormal basis for the range space of $\boldsymbol{Q}_j$, $R(\boldsymbol{Q}_j)$, $j = 1, ..., m$, we have

$$(2.1) \qquad \begin{cases} \boldsymbol{A}_j\boldsymbol{A}_j^\top = \boldsymbol{I}_{g_j},\ j = 1, ..., m \\ \boldsymbol{A}_j^\top\boldsymbol{A}_j = \boldsymbol{Q}_j,\ j = 1, ..., m \end{cases},$$

with $g_j = rank(\boldsymbol{Q}_j)$, $j = 1, ..., m$. Let us take $\boldsymbol{X}_{0,j} = \boldsymbol{A}_j\boldsymbol{X}_0$ and represent by $\boldsymbol{P}_j$, $j = 1, ..., m$, and $\boldsymbol{P}_j^c$, $j = 1, ..., m$, the OPM on $\Omega_j = R(\boldsymbol{X}_{0,j})$ and it's orthogonal complement $\Omega_j^\perp$, $j = 1, ..., m$.

2.1. Lemma. *If the model has COBS we have* $\boldsymbol{T}\boldsymbol{Q}_j \neq \boldsymbol{0}_{n\times n}$ *if and only if* $\boldsymbol{X}_{0,j} \neq \boldsymbol{0}_{g_j\times k}$*, assuming* $\boldsymbol{X}_0$ *to be* $n \times k$*, j=1,...,m.*

Proof. We have $\boldsymbol{T}\boldsymbol{Q}_j = \boldsymbol{0}_{n\times n}$ if and only if $R(\boldsymbol{T}\boldsymbol{Q}_j) = \{\boldsymbol{0}_n\}$, so, if the model has COBS, $R(\boldsymbol{T}\boldsymbol{Q}_j) = R(\boldsymbol{Q}_j\boldsymbol{T}) = \boldsymbol{Q}_jR(\boldsymbol{T}) = \boldsymbol{Q}_jR(\boldsymbol{X}_0) = \boldsymbol{A}_j^\top\boldsymbol{A}_jR(\boldsymbol{X}) = \boldsymbol{A}_j^\top R(\boldsymbol{A}_j\boldsymbol{X}) = \boldsymbol{A}_j^\top R(\boldsymbol{X}_{0,j})$ and, since the column vectors of $\boldsymbol{A}_j^\top$ are linearly independent, $R(\boldsymbol{T}\boldsymbol{Q}_j) =$

$\boldsymbol{A}_j^\top R(\boldsymbol{X}_j) = \{\boldsymbol{0}_n\}$ if and only if $R(\boldsymbol{X}_{0,j}) = \{\boldsymbol{0}_n\}$ which is equivalent to $\boldsymbol{X}_{0,j} = \boldsymbol{0}_{g_j \times k}$, $j = 1, ..., m$.

2.2. Corollary. *If the model has COBS we have* $\boldsymbol{TQ}_j \neq \boldsymbol{0}_{n\times n}$ *if and only if* $\boldsymbol{P}_j \neq \boldsymbol{0}_{g_j\times g_j}$, $j = 1, ..., m$.

2.3. Corollary. *If the model has COBS we have* $\boldsymbol{TQ}_j \neq \boldsymbol{0}_{n\times n}$ *if and only if* $\bar{\boldsymbol{Q}}_j = \boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j \neq \boldsymbol{0}_{n\times n}$, $j = 1, ..., m$.

Proof. $\boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j = (\boldsymbol{A}_j^\top \boldsymbol{P}_j)(\boldsymbol{A}_j^\top \boldsymbol{P}_j)^\top$ so, see [20],

$$rank(\boldsymbol{X}_j \boldsymbol{P}_j \boldsymbol{A}_j) = rank(\boldsymbol{A}_j^\top \boldsymbol{P}_j), j = 1, ..., m.$$

Now the column vectors of $\boldsymbol{A}_j^\top$ are linearly independent so $\boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j = \boldsymbol{0}_{n\times n}$. This is $rank(\boldsymbol{A}_j^\top \boldsymbol{P}_j) = rank(\boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j) = 0$ if and only if $\boldsymbol{P}_j = \boldsymbol{0}_{g_j\times g_j}$, $j = 1, ..., m$. Thus, according to Corollary 2.2, $\boldsymbol{TQ}_j \neq \boldsymbol{0}_{n\times n}$ only when $\bar{\boldsymbol{Q}}_j \neq \boldsymbol{0}_{n\times n}$.

2.4. Corollary. *If the model has COBS we can order the* $\boldsymbol{TQ}_1, ..., \boldsymbol{TQ}_m$ *and the*

$$\bar{\boldsymbol{Q}}_1, ..., \bar{\boldsymbol{Q}}_m$$

to have $\boldsymbol{TQ}_j \neq \boldsymbol{0}_{n\times n}$ $[\bar{\boldsymbol{Q}}_j \neq \boldsymbol{0}_{n\times n}]$, *if and only if* $j \leq z$.

2.5. Proposition. *If the model has COBS we have* $\boldsymbol{TQ}_j = \bar{\boldsymbol{Q}}_j$, $j = 1, ..., z$.

Proof. Since $\boldsymbol{TQ}_j$ $[\bar{\boldsymbol{Q}}_j]$, $j = 1, ..., z$, are symmetric and idempotent matrices they are OPM. So we have only to show that $R(\boldsymbol{TQ}_j) = R(\bar{\boldsymbol{Q}}_j)$, $j = 1, ..., z$. Now

$$rank(\boldsymbol{A}_j^\top \boldsymbol{P}_j) = rank(\boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{P}_j \boldsymbol{A}_j) = rank(\boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j) = rank(\bar{\boldsymbol{Q}}_j), j = 1, ..., z,$$

so that

$$R(\bar{\boldsymbol{Q}}_j) = R(\boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j) = R(\boldsymbol{A}_j^\top \boldsymbol{P}_j), j = 1, ..., z,$$

since the first is a subspace of the last set with the same dimension.
Besides this

$$R(\boldsymbol{Q}_j \boldsymbol{TQ}_j) = R(\boldsymbol{Q}_j \boldsymbol{T}) = \boldsymbol{Q}_j R(\boldsymbol{T}) = \boldsymbol{Q} R(\boldsymbol{X}) =$$

$$= \boldsymbol{A}_j^\top \boldsymbol{A}_j R(\boldsymbol{X}) = \boldsymbol{A}_j^\top R(\boldsymbol{A}_j \boldsymbol{X}) = \boldsymbol{A}_j^\top R(\boldsymbol{X}_j) = \boldsymbol{A}_j^\top R(\boldsymbol{P}_j) = R(\boldsymbol{A}_j^\top \boldsymbol{P}_j) = R(\bar{\boldsymbol{Q}}_j),$$

$j = 1, ..., m$, which establish the thesis.

2.6. Corollary. *Putting* $\boldsymbol{T}^c = \boldsymbol{I}_n - \boldsymbol{T}$ *and* $\bar{\boldsymbol{Q}}^{\bullet}{}_j = \boldsymbol{A}_j^\top \boldsymbol{P}_j^c \boldsymbol{A}_j$, $j = z+1, ..., m$, *when the model has COBS we have* $\boldsymbol{T}^c \boldsymbol{Q}_j = \bar{\boldsymbol{Q}}^{\bullet}{}_j$, $j = z+1, ..., m$.

Proof. According to Corollary 2.4 we have $\boldsymbol{T}^c \boldsymbol{Q}_j = \boldsymbol{Q}_j - \boldsymbol{TQ}_j = \boldsymbol{Q}_j$, $j = z+1, ..., m$, as well as $\boldsymbol{A}_j^\top \boldsymbol{P}_j^c \boldsymbol{A}_j = \boldsymbol{A}_j^\top \boldsymbol{A}_j - \boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j = \boldsymbol{Q}_j - \bar{\boldsymbol{Q}}_j = \boldsymbol{Q}_j$, $j = z+1, ..., m$, so the thesis is established.

2.7. Corollary. *When the model has COBS the CJA with principal basis* $\{\boldsymbol{TQ}_1, ..., \boldsymbol{TQ}_z, \boldsymbol{T}^c\boldsymbol{Q}_{z+1}, ..., \boldsymbol{T}^c\boldsymbol{Q}_m\}$ *and* $\{\bar{\boldsymbol{Q}}_1, ..., \bar{\boldsymbol{Q}}_z, \bar{\boldsymbol{Q}}^{\bullet}{}_{z+1}, ..., \bar{\boldsymbol{Q}}^{\bullet}{}_m\}$ *are identical.*

Proof. The result follows from Corollary 2.6 and Proposition 2.5.

2.8. Corollary. *If the model has COBS we have* $\boldsymbol{T} = \sum_{j=1}^{z} \boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j$.

Proof. We have $\boldsymbol{T} = \sum_{j=1}^{z} \boldsymbol{TQ}_j$ so the thesis follows from Proposition 2.5.

3. Mixed Models

We now characterize mixed models with OBS and COBS. If the matrices of $\underline{M} = \{\boldsymbol{M}_1, ..., \boldsymbol{M}_w\}$ commute they will generate a CJA, $\mathcal{A}(\underline{M})$, as we saw in Section 2. With $\boldsymbol{Q} = \{\boldsymbol{Q}_1, ..., \boldsymbol{Q}_m\} = pb(\mathcal{A}(\underline{M}))$ we have $\boldsymbol{M}_i = \sum_{j=1}^{m} b_{i,j}\boldsymbol{Q}_j, i = 1, ..., w$, putting $\boldsymbol{B} = [b_{i,j}]$ and $\psi_i = \{j : b_{i,j} \neq 0\}, i = 1, .., w$, it is easy to see that the OPM on $R(\boldsymbol{M}_i) = R(\boldsymbol{X}_i)$ is $\sum_{j\in\psi_i} \boldsymbol{Q}_j$. Moreover the OPM on $R(\sum_{i=1}^{w} \boldsymbol{M}_i) = R([\boldsymbol{X}_1...\boldsymbol{X}_w])$ will be $\sum_{j=1}^{m} \boldsymbol{Q}_j$. Thus we have $R([\boldsymbol{X}_1...\boldsymbol{X}_w]) = \mathbb{R}^n$ if and only if $\boldsymbol{I}_n = \sum_{j=1}^{m} \boldsymbol{Q}_j$, which is, as we saw, one of the requirements on the POOPM that appear on the variance-covariance matrices of models with OBS. The mixed models will have variance-covariance matrices

$$\boldsymbol{V}(\boldsymbol{\theta}) = \sum_{i=1}^{w} \theta_i \boldsymbol{M}_i = \sum_{i=1}^{w} \theta_i (\sum_{j=1}^{m} b_{i,j}\boldsymbol{Q}_j) = \sum_{j=1}^{m} \gamma_j \boldsymbol{Q}_j, \tag{3.1}$$

where $\gamma_j = \sum_{i=1}^{w} b_{i,j}\theta_i, j = 1, ..., m$, so $\boldsymbol{\gamma} \in R(\boldsymbol{B}^\top)_+$, with ∇_+ the family of vectors of sub-space ∇ with non-negative components.

For the variance-covariance matrices of the model to be all the positive semi-definite matrices given by linear combination of $\boldsymbol{Q}_1, ..., \boldsymbol{Q}_m$ we have to have

$$R(\boldsymbol{B}^\top) = \mathbb{R}^m,$$

this is matrix $\boldsymbol{B}$ must be invertible which occur when and only when $\underline{M}$ is a basis for $\mathcal{A}(\underline{M})$. Then, see [4], the family $\underline{M}$ will be perfect. We now establish

3.1. Proposition. *The mixed model enjoys OBS when $\underline{M}$ is a perfect family and*

$$R([\boldsymbol{X}_1...\boldsymbol{X}_w] = \mathbb{R}^w.$$

Proof. When $R([\boldsymbol{X}_1...\boldsymbol{X}_w]) = \mathbb{R}^w$ but $\underline{M}$ is not perfect we can always complete it adding some random effect terms to the model. We then restrict ourselves to perfect $\underline{M}$ families.

Going over to models with COBS we establish

3.2. Proposition. *$\boldsymbol{T}$ commutes with the matrices of $\underline{M}$ if and only if it commutes with matrices of $\boldsymbol{Q}$.*

Proof. If $\boldsymbol{T}$ commutes with the matrices of $\underline{M}$, the matrices of

$$\boldsymbol{M}^o = \{\boldsymbol{T}, \boldsymbol{M}_1, ..., \boldsymbol{M}_w\}$$

commute so they will generate a CJA $\mathcal{A}(\boldsymbol{M}^o)$ that contains $\boldsymbol{M}^o$, then containing $\boldsymbol{T}$, and the matrices of $\boldsymbol{Q}$ that will commute. Inversely if $\boldsymbol{T}$ commutes with the matrices of $\boldsymbol{Q}$ it commutes with matrices of $\underline{M}$ since $\boldsymbol{M}_i = \sum_{j=1}^{m} b_{i,j}\boldsymbol{Q}_j, i = 1, ..., w$.

3.3. Corollary. *If a model has OBS and $\boldsymbol{T}$ commutes with the matrices of $\underline{M}$ it has COBS.*

4. Estimation

In this section we will use the sub-models $\boldsymbol{Y}_j = \boldsymbol{A}_j\boldsymbol{Y}, j = 1, ..., m$ to obtain estimators for estimable vectors. Taking $\boldsymbol{\mu}_j = \boldsymbol{A}_j\boldsymbol{\mu}, j = 1, ..., m$, where $\boldsymbol{\mu}_j = \boldsymbol{0}_{g_j}$, $j = z+1, ..., m$, a model with generalized OBS, GOBS, has the homoscedastic partition $\boldsymbol{Y} = \sum_{j=1}^{m} \boldsymbol{A}_j^\top \boldsymbol{Y}_j$

where the $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_m$ have mean vectors $\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_m$, and variance-covariance matrices $\gamma_1 \boldsymbol{I}_{g_1}, ..., \gamma_m \boldsymbol{I}_{g_m}$.

Now $\boldsymbol{\psi} = \boldsymbol{G\beta}$ is estimable, see for instance [13], if and only if $\boldsymbol{G} = \boldsymbol{UX}_0$, so that $\boldsymbol{\psi} = \boldsymbol{U\mu} = \sum_{j=1}^{z} \boldsymbol{U}_j \boldsymbol{\mu}_j = \sum_{j=1}^{z} \boldsymbol{\psi}_j$ with $\boldsymbol{U}_j = \boldsymbol{U}\boldsymbol{A}_j^\top$ and $\boldsymbol{\psi}_j = \boldsymbol{U}_j \boldsymbol{\mu}_j$, $j = 1, ..., z$. Now we establish

4.1. Proposition. *$\tilde{\boldsymbol{\psi}} = \sum_{j=1}^{z} \tilde{\boldsymbol{\psi}}_j$, with $\tilde{\boldsymbol{\psi}}_j = \boldsymbol{U}_j \boldsymbol{P}_j \boldsymbol{Y}_j$, $j = 1, ..., z$, is an unbiased estimator of $\boldsymbol{\psi}$, and if $\boldsymbol{\psi}^* = \sum_{j=1}^{z} \boldsymbol{\psi}_j^*$ with $\boldsymbol{\psi}_j^* = \boldsymbol{W}_j \boldsymbol{Y}_j$ is another unbiased estimator of $\boldsymbol{\psi}$, $j = 1, ..., z$, $\boldsymbol{\psi}^*$ is an unbiased estimator of $\boldsymbol{\psi}$, with $\not\Sigma(\tilde{\boldsymbol{\psi}}) \leq \not\Sigma(\boldsymbol{\psi}^*)$ where $\leq$ indicates that $\not\Sigma(\boldsymbol{\psi}^*) - \not\Sigma(\tilde{\boldsymbol{\psi}})$ is positive semi-definite.*

Proof. Since the mean vector of $\boldsymbol{P}_j \boldsymbol{Y}_j$ is $\boldsymbol{P}_j \boldsymbol{\mu}_j = \boldsymbol{\mu}_j$, $j = 1, ..., z$, $\boldsymbol{\psi}^*$ is an unbiased estimator of $\boldsymbol{\psi}$ and it is well known that $\not\Sigma(\tilde{\boldsymbol{\psi}}_j) \leq \not\Sigma(\boldsymbol{\psi}_j^*)$, $j = 1, ..., z$. Now the $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_m$ have null variance-covariance matrices, so

$$\left\{ \begin{array}{l} \not\Sigma(\tilde{\boldsymbol{\psi}}) = \sum_{j=1}^{z} \boldsymbol{U}_j \not\Sigma(\tilde{\boldsymbol{\psi}}_j) \boldsymbol{U}_j^\top = \sum_{j=1}^{z} \not\Sigma(\tilde{\boldsymbol{\psi}}_j) \\ \not\Sigma(\boldsymbol{\psi}^*) = \sum_{j=1}^{z} \boldsymbol{U}_j \not\Sigma(\boldsymbol{\psi}_j^*) \boldsymbol{U}_j^\top = \sum_{j=1}^{z} \not\Sigma(\boldsymbol{\psi}_j^*) \end{array} \right. \tag{4.1}$$

and $\boldsymbol{U}_j \not\Sigma(\tilde{\boldsymbol{\psi}}_j) \boldsymbol{U}_j^\top \leq \boldsymbol{U}_j \not\Sigma(\boldsymbol{\psi}_j^*) \boldsymbol{U}_j^\top$, $j = 1, ..., z$.

4.2. Proposition. *When the model has COBS the $\tilde{\boldsymbol{\psi}}$ are LSE.*

Proof. Since the models enjoys COBS we have $\boldsymbol{T} = \sum_{j=1}^{z} \boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j$ and $\tilde{\boldsymbol{\psi}} = \sum_{j=1}^{z} \tilde{\boldsymbol{\psi}}_j = \sum_{j=1}^{z} \boldsymbol{U}_j \boldsymbol{P}_j \boldsymbol{Y}_j = \boldsymbol{U} \left(\sum_{j=1}^{z} \boldsymbol{A}_j^\top \boldsymbol{P}_j \boldsymbol{A}_j \right) \boldsymbol{Y} = \boldsymbol{UTY} = \boldsymbol{U}\tilde{\boldsymbol{\mu}}$, with $\tilde{\boldsymbol{\mu}}$ the LSE of $\boldsymbol{\mu}$, so the thesis is established.

This result is interesting since in COBS the LSE are UBLUE, being BLUE whatever $\boldsymbol{\theta}$, see [23]. Thus we validate the above proposition showing that our "sub-optimal estimator" is "optimal" when the model enjoys COBS. In the previous phrase "sub-optimal" must be taken in the sense of Proposition 4.1 and "optimal" in the sense of the LSE being UBLUE.

Let us put $q_j = rank(\boldsymbol{P}_j^c)$, $\quad j = 1, ..., m$, as well as $\mathfrak{D} = \{j; \quad q_{j>0}\}$, and

$$\tilde{\gamma}_j = \frac{\boldsymbol{Y}_j^\top \boldsymbol{P}_j^c \boldsymbol{Y}_j}{q_j}, \quad j \in \mathfrak{D}. \tag{4.2}$$

It is also well known that, if $\boldsymbol{\gamma}_j^* = \boldsymbol{I}_j^\top \boldsymbol{W}_j \boldsymbol{Y}_j, \quad j \in \mathfrak{D}$ is a quadratic unbiased estimator of $\boldsymbol{\gamma}_j$, $j \in \mathfrak{D}$, we have $var(\tilde{\boldsymbol{\gamma}}_j) \leq var(\boldsymbol{\gamma}_j^*), \quad j \in \mathfrak{D}$. Let us get the following Proposition. We leave out its proof which can be seen in [17], page 395.

4.3. Proposition. *If $\boldsymbol{Y}$ is quasi-normal we have*

$$var\left(\sum_{j \in \mathfrak{D}} c_j \tilde{\boldsymbol{\gamma}}_j \right) \leq var\left(\sum_{j \in \mathfrak{D}} c_j \boldsymbol{\gamma}_j^* \right). \tag{4.3}$$

5. An application

The mixed model

$$\boldsymbol{Y} = \sum_{i=0}^{w} \boldsymbol{X}_i \boldsymbol{\beta}_i,$$

where $\boldsymbol{\beta}_0$ is fixed and the $\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_w$ are random independent vectors with null mean vector and variance-covariance matrices $\sigma_1^2 \boldsymbol{I}_{g_1}, ..., \sigma_w^2 \boldsymbol{I}_{g_w}$ have GOBS, see for instance [14], when the matrices $\boldsymbol{M}_i = \boldsymbol{X}_i \boldsymbol{X}_i^\top, i = 1, ..., w$ commute.
Namely these matrices will belong to a CJA $\mathcal{A}$, with $pb(\mathcal{A}) = \{\boldsymbol{Q}_1, ..., \boldsymbol{Q}_m\}$, so that $\boldsymbol{M}_i = \sum_{j=1}^{m} b_{i,j} \boldsymbol{Q}_j, i = 1, ..., w$. Note that to consider an extension of OBS we can replace

ν by $\nu(\nabla) = \left\{\sum_{j=1}^{m} \gamma_j \boldsymbol{Q}_j; \quad \boldsymbol{\gamma} \in \nabla_+\right\}$, where ∇_+ is the family of vectors belonging to subspace ∇ with non negative components. Then the model will have GOBS. This application is itself an extension of the one given, see [7], [3] and [14], for models with COBS, and the identity of the two algebras for models with COBS enables us to carry out an unified treatment for models with GOBS.

These models have variance-covariance matrices

$$\boldsymbol{V}(\sigma^2) = \sum_{i=1}^{w} \sigma_i^2 \boldsymbol{M}_i = \sum_{j=1}^{m} \gamma_j \boldsymbol{Q}_j, \tag{5.1}$$

with $\gamma_j = \sum_{i=1}^{w} b_{i,j}\sigma_i^2, j = 1, ..., m$ so that now we have $\boldsymbol{\gamma} \in R(\boldsymbol{B}^\top)_+$, where $\boldsymbol{B} = [b_{i,j}]$.

We point out that for $\boldsymbol{V}(\sigma_1^2) = \boldsymbol{V}(\sigma_2^2)$ implying $\sigma_1^2 = \sigma_2^2$ the matrices $\boldsymbol{M}_1, ..., \boldsymbol{M}_w$ have to be linearly independent. Then the row vectors of $\boldsymbol{B}$ that are the column vectors of $\boldsymbol{B}^\top$, are linearly independent and we have $\boldsymbol{\sigma}^2 = \boldsymbol{B}^{\top+}\boldsymbol{\gamma}$, where $\boldsymbol{A}^+$ indicates MOORE-PENROSE inverse of matrix $\boldsymbol{A}$. Then, if $\boldsymbol{Y}$ is quasi-normal we may apply Proposition 4.3.

6. Final Remarks

Least squares estimators, LSE, have been widely used due to this algebraic structure and to having minimum variance.covariance matrices, under general conditions, whatever the variance components.

Following [21] we may say that, then, the LSE are UBLUE. Now these conditions rest on $\boldsymbol{T}$ commuting with the variance-covariance matrices of the model.

We showed that this commutativity condition was not necessary thus extending the class of models for which we have UBLUE for estimable vectors.We also showed that those UBLUE are LSE when the commutativity condition holds. Thus our results may be considered as an extension of the well known results on UBLUE that are LSE, for instance see [22] and [23].

Besides this we obtain an optimal result for estimators of linear combinations $\sum_{j=1}^{m} c_j\gamma_j$. We point out that in mixed models such as those considered in the application we have $\boldsymbol{\sigma}^2 = (\boldsymbol{B}^\top)^+\boldsymbol{\gamma}$ so we can apply that result to the components of $\widetilde{\boldsymbol{\sigma}}^2 = (\boldsymbol{B}^\top)^+\widetilde{\boldsymbol{\gamma}}$ whenever $\boldsymbol{Y}$ is quasi normal.

Acknowledgements

This work was partially supported by the Center of Mathematics, University of Beira Interior through the project PEst-OE/MAT/UI0212/2011 and CMA through the project PEst-OE/MAT/UI0297/2011.

References

[1] Caliński, T., Kageyama, S. *Block Designs: A Randomization Approach*, Volume I: Analysis, Lect. Notes Stat. **150**, Springer, New York, 2000.

[2] Caliński, T., Kageyama, S. *Block Designs: A Randomization Approach*, Volume II: Design, Lect. Notes Stat., **170**, Springer, New York, 2003.

[3] Carvalho, F., Mexia, J.T., Oliveira, M. *Canonic inference and commutative orthogonal block structure*. Discuss. Math. Probab. Stat., **28**, 171-181, 2008.

[4] Ferreira, S.S., Ferreira, D., Fernandes, C., Mexia, J.T. *Orthogonal Mixed Models and Perfect Families of Symmetric Matrices*, Book of Abstracts, 56-th Session of the International Statistical Institute, ISI, Lisboa, 2007.

[5] Fonseca, M., Mexia, J.T., Zmyślony, R. *Binary operations on Jordan algebras and orthogonal normal models*, Linear Algebra Appl., **417**, 75-86, 2006.

[6] Fonseca, M., Mexia, J.T., Zmyślony, R. *Jordan algebras, generating pivot variables and orthogonal normal models*, J. Interdiscip. Math., **10**(2), 305-326, 2007.

[7] Fonseca, M., Mexia, J.T., Zmyślony, R. *Inference in normal models with commutative orthogonal block structure*. Acta Comment. Univ. Tartu. Math., **12**, 3-16, 2008.

[8] Fonseca, M., Mexia, J.T., Zmyślony, R. *Least squares and generalized least squares in models with orthogonal block structure*, J. Statist. Plann. Inference, **140**(5), 1346-1352, 2010.

[9] Henderson, C. R. *Statistical Methods in Animal Improvement: Historical Overview.* In Gianola, Daniel and Hammond, Keith. Advances in Statistical Methods for Genetic Improvement of Livestock. Springer-Verlag Inc., 2-14, 1990.

[10] Houtman, A.M., Speed, T.P. *Balance in designed experiments with orthogonal block structure*. Ann. Statist., **11**, 1069-1085, 1983.

[11] Jesus, V., Mexia, J.T., Zmyślony, R. *Binary operations and canonical forms for factorial and related models*, Linear Algebra Appl., **430**(10) 2781-2797, 2009.

[12] Jordan, P., Neumann, J. von, Wigner, E. *On an Algebraic Generalization of the Quantum Mechanical Formalism*. Ann. of Math. (Princeton), **35**(1), 29-64, 1934.

[13] Mexia, J.T. *Best linear unbiased estimates, duality of F tests and the Scheffé multiple comparison method in the presence of controlled heteroscedasticity*. Comput. Stat. Data An., **10**(3) 271-281, 1990.

[14] Mexia, J.T., Vaquinhas, R., Fonseca, M., Zmyślony, R. *COBS: Segregation, Matching, Crossing and Nesting.* Latest Trends on Applied Mathematics, Simulation, Modeling, 4th International Conference on Applied Mathematics, Simulation, Modelling (ASM'10) 249-255, 2010.

[15] Nelder, J.A. *The analysis of randomized experiments with orthogonal block structure*. I. Block structure and the null analysis of variance. Proceedings of the Royal Society (London), Series A, **273**, 147-162, 1965a.

[16] Nelder, J.A. *The analysis of randomized experiments with orthogonal block structure*. II. Treatment structure and the general analysis of variance. Proceedings of the Royal Society (London), Series A, **273**, 163-178, 1965b.

[17] Schott, J.R. *Matrix Analysis for Statistics*. John Wiley & Sons, New York, 1997.

[18] Searle, S.R., Casella, G., McCulloch, C.E. *Variance Components*. John Wiley & Sons, New York, 1992.

[19] Seely, J. *Quadratic subspaces and completeness*. Ann. Math. Stat., **42**(2), 710-721, 1971.

[20] Silvey, S.D. *Statistical Inference*. Reprinted. Chapman & Hall, 1975.

[21] VanLeeuwen, D.M., Seely, J.F., Birkes, D.S. *Sufficient conditions for orthogonal designs in mixed linear models*. J. Statist. Plann. Inference, **73**, 373-389, 1998.

[22] Zyskind, G. . *On canonical forms, non-negative covariance matrices and best simple linear least squares estimators in linear models*. Ann. Math. Stat., **38**, 1092-1109, 1967.

[23] Zmyślony, R. *A characterization of best linear unbiased estimators in the general linear model*, Mathematical Statistics and Probability Theory, Proceedings Sixth International Conference, Wisla, Lecture Notes in Statistics, Springer, New York-Berlin, **2**, 365-373, 1978.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 549 – 559

Robust variable selection for mixture linear regression models

Yunlu Jiang[*†]

Abstract

In this paper, we propose a robust variable selection to estimate and select relevant covariates for the finite mixture of linear regression models by assuming that the error terms follow a Laplace distribution to the data after trimming the high leverage points. We introduce a revised Expectation-maximization (EM) algorithm for numerical computation. Simulation studies indicate that the proposed method is robust to both the high leverage points and outliers in the y-direction, and can obtain a consistent variable selection in the case of outliers or heavy-tail error distribution. Finally, we apply the proposed methodology to analyze a real data.

Keywords: Finite mixture of linear regression models, Robustness, EM-algorithm.

2000 AMS Classification: 62G35, 62G35, 62H12.

Received : 03.11.2014 *Accepted* : 02.03.2015 *Doi* : 10.15672/HJMS.2015549560

1. Introduction

Finite mixture of linear regression (FMLR) models provide a very important statistical tool to fit the unobserved heterogeneous relationships. They are extensively used in many research fields, e.g., marketing and social sciences [29, 25], machine learning [12, 13]. A comprehensive review of finite mixture models was given in [20]. It is well-known that the traditional maximum likelihood estimator (MLE) for mixture linear regression models works well when the error term follows a normal distribution. However, the normality based MLE is not robust to outliers in the datasets.

Many robust methodologies were proposed and widely studied for mixture linear regression models in the literature. For instance, [18] and [24] introduced the weighted MLE. [21] proposed the trimmed likelihood estimator. [1] proposed a modified Expectation-maximization (EM)-algorithm by replacing the least squares criterion with a robust

*Department of Statistics, College of Economics, Jinan University, Guangzhou, 510632, China, Email: `tjiangyl@jnu.edu.cn`
†Corresponding Author.

criteria in the M step. [30] and [26] proposed a robust estimation procedure based a t-distribution and a Laplace distribution, respectively.

In many practical applications, there are many covariates involved in the FMLR models. Nevertheless, the number of important ones is usually relatively small. In fact, the problem of variable selection in a FMLR model has received much attention recently. For example, [28] used Akaike information criterion (AIC) and Bayesian information criterion (BIC) to study model choice issues for a class of Poisson mixture models. [15] introduced a penalized likelihood approach for variable selection in FMLR models based on some well-known families such as Gaussian, Poisson, and Binomial distributions, and developed an EM algorithm for numerical computations. [17] proposed a mixture regression LASSO (MR-LASSO) method to penalize both regression coefficients and mixture components simultaneously. [14] gave an overview of the new feature selection methods in FMLR models. [16] studied the issue of variable selection in FMLR models when the number of parameters in the model can increase with the sample size. [5] proposed a penalized likelihood approach to simultaneously select important fixed and random effects in the finite mixtures of linear mixed-effects models. It is very important to note that many of those methods are closely related to the traditional MLE method.

To the best of our knowledge, the robust feature selection for FMLR models has not been well studied. In the linear regression models, the least absolute deviation (LAD)estimator is very important when the error terms follow a heavy-tailed distribution, and has the desired robust properties. In fact, the maximum-likelihood estimator of the regression parameters given a Laplace distributed regression errors is LAD estimator. [26] applied the LAD estimator to a class of FMLR models. In this article, we propose a robust variable selection procedure based on the LAD estimator for FMLR models, and introduce a revised EM-type algorithm for numerical computation. Simulation studies show that the proposed method is robust and can obtain a consistent variable selection when there are outliers in the datasets or the error term follows a heavy-tailed distribution. In addition, the proposed robust variable selection approach works comparably to the traditional penalized likelihood-based method when there are no outliers and the error is normal.

The rest of this paper is organized as follows. In Section 2, we propose a robust variable selection for FMLR models, and introduce a revised EM-algorithm for numerical computation. In Section 3, numerical simulations and a real data analysis are conducted to compare the performance of the proposed method with the existing method. We conclude with some remarks in Section 4.

2. Methodology

Let Z be a latent class variable with $P(Z = i|\mathbf{X} = \mathbf{x}) = \pi_i, i = 1, \cdots, m$, where $\mathbf{x}$ is a q-dimensional vector. Given $Z = i$, suppose that the response Y depends on $\mathbf{X}$ in a linear way

$$Y = \mathbf{X}^T\boldsymbol{\beta}_i + \sigma_i\epsilon_i,$$

where $\boldsymbol{\beta}_i$ is an unknown q-dimensional vectors of regression parameters, σ_i is an unknown positive scalar, and ϵ_i is a random error with density $f_i(\cdot)$ and mean 0, and is independent of $\mathbf{X}$. Then, the density Y given $\mathbf{X}$ is

$$g(y|\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{m} \pi_i \frac{1}{\sigma_i} f_i\left(\frac{y - \mathbf{x}^T\boldsymbol{\beta}_i}{\sigma_i}\right), \tag{2.1}$$

where $\boldsymbol{\theta} = (\pi_1, \boldsymbol{\beta}_1, \sigma_1, \cdots, \pi_m, \boldsymbol{\beta}_m, \sigma_m)^T$.

Suppose that $\mathbf{D}_n = \{(\mathbf{X}_1, Y_1), \cdots, (\mathbf{X}_n, Y_n)\}$ are random observations from the model (2.1). The log-likelihood function is

$$\ell_n(\boldsymbol{\theta}) = \sum_{j=1}^{n} \log \left[\sum_{i=1}^{m} \pi_i \frac{1}{\sigma_i} f_i \left(\frac{Y_j - \mathbf{X}_j^T \boldsymbol{\beta}_i}{\sigma_i} \right) \right].$$

The MLE of $\boldsymbol{\theta}$ is obtained by maximizing the log likelihood function $\ell_n(\boldsymbol{\theta})$.

To simultaneously estimate and select relevant covariates, [6] proposed a unified approach via penalized likelihood. A penalized log-likelihood function is defined as follows:

$$\tilde{\ell}_n(\boldsymbol{\theta}) = \ell_n(\boldsymbol{\theta}) - \sum_{i=1}^{m} \pi_i \left\{ \sum_{k=1}^{q} p_{ni}(\beta_{ik}) \right\}, \tag{2.2}$$

where $p_{ni}(\beta_{ik})$ is nonnegative and nondecreasing functions in $|\beta_{ik}|$. Although there are many methods to deal with the problem of feature selection in finite mixture of linear regression models in the literature, many of those methods are closely related to the least squares method. It is well-known that the least squares estimator is very sensitive to the outliers in the dataset. Next, we will study the robust variable selection for finite mixture of regression models. Similar to the idea proposed by [26], we consider the density function f_i of error term follows a Laplace density function with mean 0 and scale parameter $1/\sqrt{2}$. Then, (2.2) can be written as

$$\hat{\ell}_n(\boldsymbol{\theta}) = \sum_{j=1}^{n} \log \left[\sum_{i=1}^{m} \frac{\pi_i}{\sqrt{2}\sigma_i} \exp\left(-\frac{\sqrt{2}|Y_j - \mathbf{X}_j^T \boldsymbol{\beta}_i|}{\sigma_i} \right) \right] - \sum_{i=1}^{m} \pi_i \left\{ \sum_{k=1}^{q} p_{ni}(\beta_{ik}) \right\}. \tag{2.3}$$

[26] pointed out that the EM algorithm based on the Laplace distribution is robust against outliers along the y-direction, but not for the high leverage points. Therefore, in order to obtain a robust variable selection for both the high leverage points and outliers in the y-direction, we consider a trimmed version of the new method by fitting the new model to the data after trimming the high leverage points. Let $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^T$. For each covariate $\mathbf{X}_j$, we first compute a robust Mahalanobis distance

$$MD_j = (\mathbf{X}_j - m(\mathbf{X}))C(\mathbf{X})^T(\mathbf{X}_j - m(\mathbf{X})),$$

where $m(\mathbf{X})$ and $C(\mathbf{X})$ are robust estimates of location and scatter for $\mathbf{X}$, respectively.

In the literature, there are many robust location and scatter estimators. Those estimators include M-estimator [19], Stahel-Donoho (SD) estimators [27, 4], minimum volume ellipsoid (MVE) [22], S-estimators [3], and minimum covariance determinant (MCD) estimators [2]. Due to the availability of fast MCD algorithm [23], we employ MCD estimators to calculate a robust Mahalanobis distance in this paper. Denote

$$\omega_j = \begin{cases} 1, & \text{if } MD_j \le \chi^2_{q,0.975}, \\ 0, & \text{otherwise.} \end{cases}$$

With such a weight function, the high leverage points are discarded. Then, by taking an adaptive LASSO for the penalty function, the proposed robust variable selection estimator is defined by maximizing the following objective function

$$\bar{\ell}_n(\boldsymbol{\theta}) = \sum_{j=1}^{n} \omega_j \log \left[\sum_{i=1}^{m} \frac{\pi_i}{\sqrt{2}\sigma_i} \exp\left(-\frac{\sqrt{2}|Y_j - \mathbf{X}_j^T \boldsymbol{\beta}_i|}{\sigma_i} \right) \right] - \sum_{i=1}^{m} \pi_i \left\{ \sum_{k=1}^{q} \lambda_{ik} \frac{|\beta_{ik}|}{|\hat{\beta}_{ik}|} \right\}, \tag{2.4}$$

where $\hat{\beta}_{ik}$ is the unpenalized estimator for $\boldsymbol{\beta}$ in (2.4).

2.1. The revised EM algorithm for robust variable selection. If j-th observation $(\mathbf{X}_j, Y_j)$ is from i-th component, we denote $R_{ij} = 1, i = 1, \cdots, m, j = 1, \cdots, n$, otherwise, $R_{ij} = 0$. Assume the complete data set $\{(\mathbf{X}_j, Y_j, R_{ij}), i = 1, \cdots, m, j = 1, \cdots, n\}$ is observed, then, (2.4) can be written as

$$\begin{aligned}
&\bar{\ell}_n(\boldsymbol{\theta}) \\
&= \sum_{j=1}^{n} \omega_j \sum_{i=1}^{m} R_{ij} \log\left[\frac{\pi_i}{\sqrt{2}\sigma_i}\exp\left(-\frac{\sqrt{2}|Y_j - \mathbf{X}_j^T\boldsymbol{\beta}_i|}{\sigma_i}\right)\right] - \sum_{i=1}^{m}\pi_i\left\{\sum_{k=1}^{q}\lambda_{ik}\frac{|\beta_{ik}|}{|\hat{\beta}_{ik}|}\right\} \\
&= \sum_{j=1}^{n}\sum_{i=1}^{m}\omega_j R_{ij}\log\pi_i - \sum_{j=1}^{n}\sum_{i=1}^{m}\omega_j R_{ij}\log(\sqrt{2}\sigma_i) - \sum_{j=1}^{n}\sum_{i=1}^{m}\frac{\omega_j R_{ij}\sqrt{2}|Y_j - \mathbf{X}_j^T\boldsymbol{\beta}_i|}{\sigma_i} \\
&- \sum_{i=1}^{m}\pi_i\left\{\sum_{k=1}^{q}\lambda_{ik}\frac{|\beta_{ik}|}{|\hat{\beta}_{ik}|}\right\}.
\end{aligned}$$

In the following, we introduce the revised EM algorithm to maximize $\bar{\ell}_n(\boldsymbol{\theta})$ iteratively.

(1) Choose an initial value for $\boldsymbol{\theta}$, denote $\boldsymbol{\theta}^{(0)}$.

(2) E-Step. Given the data $\mathbf{D}_n$ and $\boldsymbol{\theta}^{(k)}$, we compute the conditional expectation of the function $\bar{\ell}_n(\boldsymbol{\theta})$ with respect to R_{ij}. The conditional expectation is given as follows:

$$\begin{aligned}
Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)}) &= \sum_{j=1}^{n}\omega_j\sum_{i=1}^{m}\kappa_{ij}^{(k)}\log\pi_i - \sum_{j=1}^{n}\omega_j\sum_{i=1}^{m}\kappa_{ij}^{(k)}\log(\sqrt{2}\sigma_i) \\
&- \sum_{j=1}^{n}\omega_j\sum_{i=1}^{m}\kappa_{ij}^{(k)}\frac{\sqrt{2}|Y_j - \mathbf{X}_j^T\boldsymbol{\beta}_i|}{\sigma_i} - \sum_{i=1}^{m}\pi_i\left\{\sum_{k=1}^{q}\lambda_{ik}\frac{|\beta_{ik}|}{|\hat{\beta}_{ik}|}\right\}.
\end{aligned} \tag{2.5}$$

where

$$\kappa_{ij}^{(k)} = E[R_{ij}|\mathbf{D}_n, \boldsymbol{\theta}^{(k)}] = \frac{\pi_i^{(k)}\sigma_i^{(k)-1}\exp\{-|Y_j - \mathbf{X}_j^T\boldsymbol{\beta}_i^{(k)}|/\sigma_i^{(k)}\}}{\sum_{i=1}^{m}\pi_i^{(k)}\sigma_i^{(k)-1}\exp\{-|Y_j - \mathbf{X}_j^T\boldsymbol{\beta}_i^{(k)}|/\sigma_i^{(k)}\}}.$$

(3) M-step. The M step on the $(k+1)$-th iteration maximizes $Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$ with respect to $\boldsymbol{\theta}$. In the usual EM algorithm, the mixing proportions are updated by

$$\pi_i^{(k+1)} = \frac{\sum_{j=1}^{n}\omega_j\kappa_{ij}^{(k)}}{\sum_{j=1}^{n}\omega_j}, i = 1, \cdots, m,$$

which maximize the leading term of $Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$. This works well in our simulations.

In the following, we consider that the π_k are constant in $Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$, and maximize $Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$ with respect to the other parameters. Since the objective function $Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(k)})$ is not smooth, we maximize the following objective function by the local quadratic approximation [6, 10],

$$\begin{aligned}
&\sum_{j=1}^{n}\omega_j\sum_{i=1}^{m}\kappa_{ij}^{(k)}\log\pi_i - \frac{1}{2}\sum_{j=1}^{n}\omega_j\sum_{i=1}^{m}\kappa_{ij}^{(k)}\log(2\sigma_i^2) \\
&- \sum_{j=1}^{n}\omega_j\sum_{i=1}^{m}\kappa_{ij}^{(k)}\frac{\sqrt{2}(Y_j - \mathbf{X}_j^T\boldsymbol{\beta}_i)^2}{\sigma_i^2}\frac{\sigma_i^{(k)}}{|Y_j - \mathbf{X}_j^T\boldsymbol{\beta}_i^{(k)}|} \\
&- \sum_{i=1}^{m}\pi_i\left\{\sum_{k=1}^{q}\lambda_{ik}\frac{\beta_{ik}^2}{|\hat{\beta}_{ik}||\beta_{ik}^{(k)}|}\right\}.
\end{aligned} \tag{2.6}$$

Then, the regression coefficients are updated by solving the following equations

$$\begin{aligned}&\sum_{j=1}^{n}\omega_j\kappa_{ij}^{(k)}\frac{\partial}{\partial\beta_{it}}\left[\frac{\sqrt{2}(Y_j-\mathbf{X}_j^T\boldsymbol{\beta}_i)^2}{(\sigma_i^{(k)})^2}\frac{\sigma_i^{(k)}}{|Y_j-\mathbf{X}_j^T\boldsymbol{\beta}_i^{(k)}|}\right]\\&+\frac{\partial}{\partial\beta_{it}}\left[\pi_i\left\{\lambda_{it}\frac{\beta_{it}^2}{|\hat{\beta}_{it}||\beta_{it}^{(k)}|}\right\}\right]=0,\end{aligned}\tag{2.7}$$

where $i=1,\cdots,m$, and $t=1,\cdots,q$.

The dispersion parameters are updated by the following expression

$$\sigma_i^{2(k+1)}=\frac{2}{\sum_{j=1}^{n}\omega_j\kappa_{ij}^{(k)}}\sum_{j=1}^{n}\omega_j\kappa_{ij}^{(k)}\frac{\sqrt{2}(Y_j-\mathbf{X}_j^T\boldsymbol{\beta}_i^{(k+1)})^2\sigma_i^{(k)}}{|Y_j-\mathbf{X}_j^T\boldsymbol{\beta}_i^{(k)}|},i=1,\cdots,m.\tag{2.8}$$

(4) Repeat steps 2, 3 until convergence.

Remark 2.1 The above proposed revised EM-algorithm involves in an initial estimator, we select a robust estimation proposed by [26] for the unpenalized FMLR models as an initial estimator, that is, by maximizing the following objective function,

$$\sum_{j=1}^{n}\log\left[\sum_{i=1}^{m}\frac{\pi_i}{\sqrt{2}\sigma_i}\exp\left(-\frac{\sqrt{2}|Y_j-\mathbf{X}_j^T\boldsymbol{\beta}_i|}{\sigma_i}\right)\right].$$

Remark 2.2 To avoid numerical instability of the proposed algorithm due to very small values in the denominator of (2.7) and (2.8), as suggested by [10], we replace $|\beta_{it}^{(k)}|$ and $|Y_j-\mathbf{X}_j^T\boldsymbol{\beta}_i^{(k)}|$ by $|\beta_{it}^{(k)}|+\epsilon$ and $|Y_j-\mathbf{X}_j^T\boldsymbol{\beta}_i^{(k)}|+\epsilon$ for a given small value $\epsilon>0$. In this paper, we take $\epsilon=10^{-6}$.

3. Simulation and Application

3.1. Simulation study. In this section, we will evaluate the finite sample performance of proposed method via simulation studies. To compare the proposed approach with some existing methods, we generate the sample data $(\mathbf{X}_1,Y_1),\cdots,(\mathbf{X}_n,Y_n)$ from the following two-component mixture regression models,

$$Y_i=\begin{cases}\mathbf{X}_i^T\boldsymbol{\beta}_1+\epsilon_1, & \text{if } Z=1,\\ \mathbf{X}_i^T\boldsymbol{\beta}_2+\epsilon_2, & \text{if } Z=2,\end{cases}\quad i=1,\cdots,n\tag{3.1}$$

with $P(Z=1)=\alpha$, $P(Z=2)=1-\alpha$, $\alpha=0.4$. We also simulate $\alpha=0.25$; the outcomes are similar, and thus we do not report the corresponding results here. The sparse regression parameters are

$$\boldsymbol{\beta}_1=(0,\cdots,0,-2.5,-1.5)^T,$$
$$\boldsymbol{\beta}_2=(0,\cdots,0,-2.5,1.5)^T,$$

where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ have eight zero elements. Covariate $\mathbf{X}_i$ follows a multi-normal distribution $N(\mathbf{0},\Sigma)$, where the (i,j)-th element of Σ is $\rho^{|i-j|}$, $\rho=0.5$. The error terms ϵ_1 and ϵ_2 are independent and identically distributed random variables. To study the robustness of proposed method, we consider the following four settings:

(1) The error terms follow a standard normal distribution, $N(0,1)$;

(2) The error terms follow a Student's t-distribution with 2 degrees of freedom, t_2;

(3) The error terms follow a 5% contaminated normal distribution, $CN_{0.05}=0.95N(0,1)+0.05N(10,20^2)$;

(4) The error terms follow a standard normal distribution with 5% high leverage outliers being $\mathbf{X}_1=(50,\cdots,50)^T$, and $Y=100$.

For each setting, we simulate 200 data sets from model (3.1) with sample sizes of $n = 200, 400$, and compare the performance of proposed method (MixregL-MCD) with the penalized likelihood approach (MixregL-ALASSO) [15] and the oracle estimator based on the Laplace error to the data after trimming the high leverage points based on a robust Mahalanobis distance with the MCD estimators. To measure the finite sample performance, we report the proportions of correctly estimated zero coefficients (specificity: S_1) and correctly estimated non-zero coefficients (sensitivity: S_2), and the component-wise median empirical mean squared errors (MEMSE) of the estimators $\hat{\boldsymbol{\beta}}_k, k = 1, 2$. According to [15] and [17], we consider the tuning parameter $\lambda_{ik} = \log(n) \times \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$. In simulation studies, the finite sample performance of $\lambda_{ik} = \log(n) \times 0.2$ is slightly better than that of others. Therefore, we take $\lambda_{ik} = \log(n) \times 0.2$ in all simulation studies and real data applications. Clearly, the choice of tuning parameter is a very important issue, however, we shall not address the problem of how to find the optimal tuning parameter, and will consider the choice of tuning parameter as future work. The simulation results are given in Table 1-4.

From Table 1, we find that when the true distribution of error term is normal and there are no outliers in the dataset, both S_1 and S_2 are around 1 for all three methods. The MEMSE of both methods is close to that of oracle estimator. When there are outliers in the datasets or the error term follows a heavy-tailed distribution, the simulation results clearly show from Table 2 to Table 4 that the proposed method works much better than the MixregL-ALASSO. S_1 and S_2 of the proposed method are higher than those of the MixregL-ALASSO, and our proposed approach has smaller MEMSE than the MixregL-ALASSO. In addition, the performance of proposed method is closer to that of the oracle estimator as the sample size n increases.

Based on the above findings, the proposed method is not sensitive to outliers in the dataset, and has the overall best performance. Thus, we recommend the use of proposed method in practical applications.

In the above simulations, we assume that the number of mixture components is known. However, the order m needs to be estimated based on the dataset in some applications. There are many methods to choose the order m in the literature, e.g., cross-validation (CV), generalized cross-validation (GCV), Akaike information criterion (AIC), and bayesian information criterion (BIC). In this paper, we select the order m by minimizing a following BIC-score

$$BIC(m) = -2l_n(\bar{\boldsymbol{\theta}}_m) + S\log(n),$$

where $\bar{\boldsymbol{\theta}}_m$ is the maximizer of the proposed objective function for a mixture regression model with the order m, and S is the number of nonzero of the estimator $\bar{\boldsymbol{\theta}}_m$.

In the following, we will use simulation studies to illustrate how to select the order. A total of 300 data sets with sample sizes $n = 400$ are generated according to the second setting with true $m = 2$. The simulation result is shown in Figure 1. We can see from Figure 1 that the BIC performs well to select the true order.

3.2. Real data application. In this section, we will apply the proposed methodology to analyze the baseball salaries dataset, which can be downloaded from

`www.amstat.org/publications/jse`.

This dataset contains 337 observations. Of interest are to study the relationships between the salary (measured in thousands of dollars) and the following 16 covariates: batting average (X_1), on-base percentage (X_2), runs (X_3), hits (X_4), doubles (X_5), triples (X_6), home runs (X_7), runs batted in (X_8), walks (X_9), strikeouts (X_{10}), stolen bases(X_{11}) , errors (X_{12}), free agency eligibility (X_{13}), free agent in 1991/2 (X_{14}), arbitration eligibility (X_{15}), and arbitration in 1991/2 (X_{16}). $X_{13}, X_{14}, X_{15}, X_{16}$ are indicators.

Table 1. Simulation results for the first setting

n		Method	S_1	S_2	MEMSE
200	$\boldsymbol{\beta}_1$	MixregL-ALASSO	0.9850	1.0000	0.0087
		MixregL-MCD	1.0000	0.9950	0.0054
		Oracle	1.0000	1.0000	0.0052
	$\boldsymbol{\beta}_2$	MixregL-ALASSO	0.9863	1.0000	0.0027
		MixregL-MCD	1.0000	1.0000	0.0043
		Oracle	1.0000	1.0000	0.0038
400	$\boldsymbol{\beta}_1$	MixregL-ALASSO	0.9950	1.0000	0.0047
		MixregL-MCD	1.0000	1.0000	0.0038
		Oracle	1.0000	1.0000	0.0033
	$\boldsymbol{\beta}_2$	MixregL-ALASSO	1.0000	1.0000	0.0016
		MixregL-MCD	1.0000	1.0000	0.0021
		Oracle	1.0000	1.0000	0.0016

Table 2. Simulation results for the second setting

n		Method	S_1	S_2	MEMSE
200	$\boldsymbol{\beta}_1$	MixregL-ALASSO	0.6925	0.7750	0.2260
		MixregL-MCD	1.0000	0.9850	0.0096
		Oracle	1.0000	1.0000	0.0075
	$\boldsymbol{\beta}_2$	MixregL-ALASSO	0.6775	0.9450	0.0398
		MixregL-MCD	1.0000	0.9900	0.0043
		Oracle	1.0000	1.0000	0.0038
400	$\boldsymbol{\beta}_1$	MixregL-ALASSO	0.8288	0.7100	0.2317
		MixregL-MCD	1.0000	1.0000	0.0055
		Oracle	1.0000	1.0000	0.0039
	$\boldsymbol{\beta}_2$	MixregL-ALASSO	0.8363	0.9500	0.1001
		MixregL-MCD	1.0000	1.0000	0.0026
		Oracle	1.0000	1.0000	0.0019

Table 3. Simulation results for the third setting

n		Method	S_1	S_2	MEMSE
200	$\boldsymbol{\beta}_1$	MixregL-ALASSO	0.8550	0.7600	0.3095
		MixregL-MCD	1.0000	0.9200	0.0097
		Oracle	1.0000	1.0000	0.0065
	$\boldsymbol{\beta}_2$	MixregL-ALASSO	0.8838	0.8750	0.1749
		MixregL-MCD	1.0000	0.9750	0.0064
		Oracle	1.0000	1.0000	0.0039
400	$\boldsymbol{\beta}_1$	MixregL-ALASSO	0.7512	0.7850	0.3382
		MixregL-MCD	1.0000	0.9450	0.0051
		Oracle	1.0000	1.0000	0.0048
	$\boldsymbol{\beta}_2$	MixregL-ALASSO	0.8275	0.9300	0.1486
		MixregL-MCD	1.0000	0.9750	0.0052
		Oracle	1.0000	1.0000	0.0048

We plot a histogram of home runs and stolen bases in Figure 2. Figure 2 indicates that there are unusual points in the dataset. According to the suggestion proposed by [15], we apply the MixregL-ALASSO and MixregL-MCD with $m = 2$ to deal with this dataset. The results are summarized in Table 5. From Table 5, we find that the MixregL-ALASSO

Table 4. Simulation results for the fourth setting

n		Method	S_1	S_2	MEMSE
200	$\boldsymbol{\beta}_1$	MixregL-ALASSO	0.2350	0.9450	0.7180
		MixregL-MCD	1.0000	0.9900	0.0055
		Oracle	1.0000	1.0000	0.0041
	$\boldsymbol{\beta}_2$	MixregL-ALASSO	0.3188	1.0000	0.1766
		MixregL-MCD	1.0000	1.0000	0.0030
		Oracle	1.0000	1.0000	0.0021
400	$\boldsymbol{\beta}_1$	MixregL-ALASSO	0.1075	0.9750	0.6932
		MixregL-MCD	1.0000	1.0000	0.0033
		Oracle	1.0000	1.0000	0.0031
	$\boldsymbol{\beta}_2$	MixregL-ALASSO	0.1862	1.0000	0.1420
		MixregL-MCD	1.0000	1.0000	0.0014
		Oracle	1.0000	1.0000	0.0011

obtains more significant explanatory variables than the MixregL-MCD. However, our proposed method should give the more reasonable model when there are outliers in the dataset.

Table 5. Estimated regression coefficients from the baseball salaries dataset

	Method			
	MixregL-ALASSO		MixregL-MCD	
Variable	Component 1	Component 2	Component 1	Component 2
X_1	6.5864	0.0011	0	0
X_2	10.677	15.388	16.742	19.816
X_3	0	0	0	0
X_4	0	0.0026	0	0
X_5	0	0	0	0
X_6	0	0	0	0
X_7	0	-0.0005	0	0
X_8	0.0063	0.0058	0	0
X_9	-0.0067	-0.0104	0	0
X_{10}	0.0053	0.0066	0	0
X_{11}	0	0	0	0
X_{12}	0	0.0002	0	0
X_{13}	1.7160	1.6827	2.1480	0
X_{14}	-0.0084	-0.0012	0	0
X_{15}	1.4961	1.4833	1.6340	0
X_{16}	0	-0.0001	0	0

4. Discussion

In this article, we proposed a robust variable selection by assuming that the error terms follow a Laplace distribution for FMLR models. We used the revised EM-algorithm to solve the proposed optimization problem. The merits of proposed methodology were illustrated via the simulation studies. According to our simulation studies, the proposed method was robust and possessed a consistent variable selection when there were outliers or the error distribution was heavy-tail.

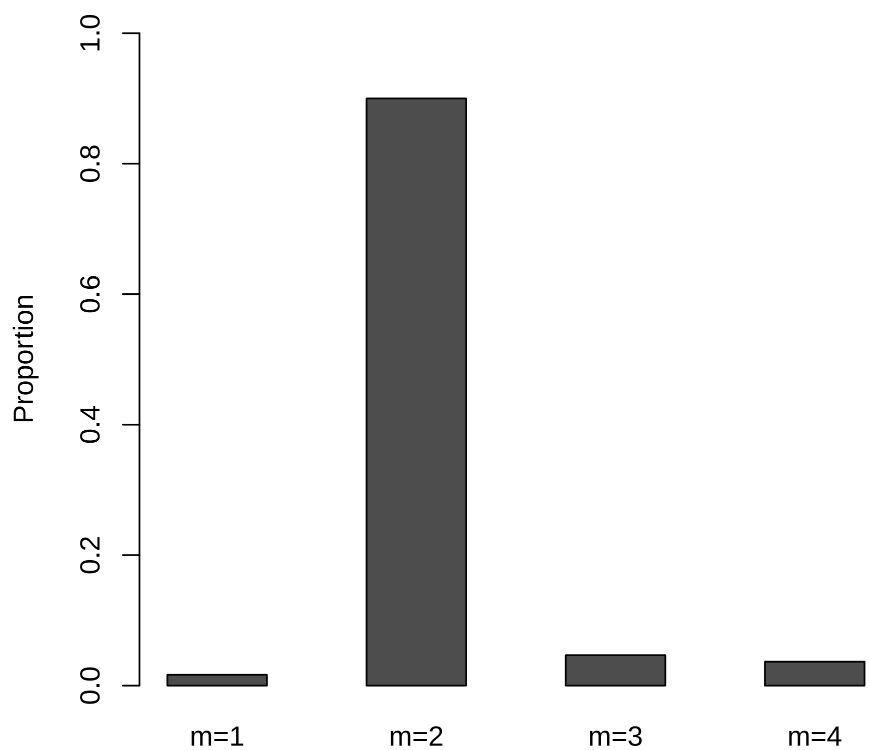
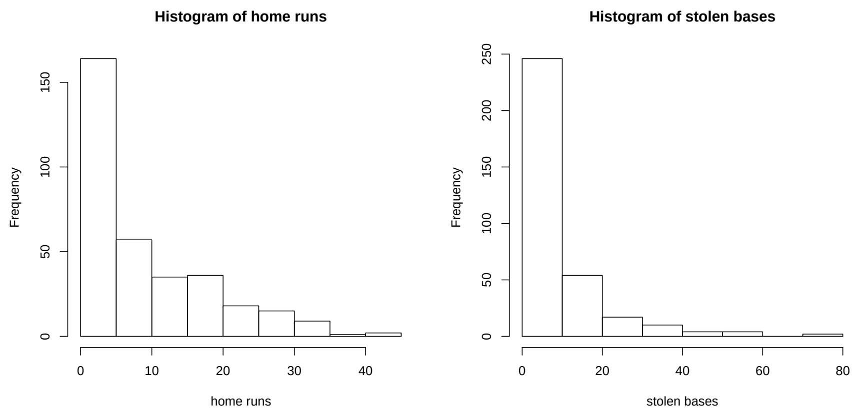


Figure 1. Order selection results based on BIC for the FMLR models with true order $m = 2$

As a variable selection procedure, it is very desirable to enjoy the oracle properties. Therefore, it warrants further effort to investigate the asymptotic properties for the proposed method. Meanwhile, it is very interesting to extent our methodology to nonparametric mixture of regression models [8], a class of semiparametric mixtures of regression models [11, 9], and mixture of gaussian processes [7].

Acknowledgements

Jiang's research is partially supported by the National Natural Science Foundation of China (No. 11301221), and is supported by the Fundamental Research Funds for the Central Universities (No. 11615455).



(a) (b)

Figure 2. Histogram of home runs (a) and stolen bases (b).

References

[1] Bai, X., Yao, W., and Boyer, J. E. *Robust fitting of mixture regression models*, Comput. Stat. Data. An. **56** (7), 2347-2359, 2012.

[2] Croux, C., and Haesbroeck, G. *Influence function and efficiency of the mini- mum covariance determinant scatter matrix estimator*, J. Multivariate Anal. **71** (2), 161-190, 1999.

[3] Davies, P. *Asymptotic Behaviour of S-Estimates of Multivariate Location Param- eters and Dispersion Matrices*, Ann. Statist. **15** (3), 1269-1292, 1987.

[4] Donoho, D. L. *Breakdown properties of multivariate location estimators,, Technical report, Technical report, Harvard University, Boston. URL http://www-stat. stanford. edu/ donoho/Reports/Oldies/BPMLE. pdf*, 1982.

[5] Du, Y., Khalili, A., Neslehova, J. G., and Steele, R. J. *Simultaneous fixed and random effects selection in finite mixture of linear mixed-effects models*, Can. J. Stat. **41** (4), 596-616, 2013.

[6] Fan, J., and Li, R. *Variable selection via nonconcave penalized likelihood and its oracle properties*, J. Amer. Statist. Assoc. **96** (456), 1348-1360, 2001.

[7] Huang, M., Li, R., Wang, H., and Yao, W. *Estimating Mixture of Gaussian Pro- cesses by Kernel Smoothing*, J. Bus. Econ. Stat. **32** (2), 259-270, 2014.

[8] Huang, M., Li, R., and Wang, S. *Nonparametric mixture of regression models*, J. Amer. Statist. Assoc. **108** (503), 929-941, 2013.

[9] Huang, M., and Yao, W. *Mixture of regression models with varying mixing pro- portions: a semiparametric approach*, J. Amer. Statist. Assoc. **107** (498), 711-724, 2012.

[10] Hunter, D. R., and Li, R. *Variable selection using MM algorithms*, Ann. Statist. **33** (4), 1617-1642, 2005.

[11] Hunter, D. R., and Young, D. S. *Semiparametric mixtures of regressions*, J. Nonparametr. Stat. **24** (1), 19-38, 2012.

[12] Jacobs, R. A., Jordan, M. I., Nowlan, S. J., and Hinton, G. E. *Adaptive mixtures of local experts*, Neural. Comput. **3** (1), 79-87, 1991.

[13] Jiang, W., Tanner, M. A. et al. *Hierarchical mixtures-of-experts for exponential fam- ily regression models: approximation and maximum likelihood estimation*, Ann. Statist. **27** (3), 987-1011, 1997.

[14] Khalili, A. *An Overview of the New Feature Selection Methods in Finite Mixture of Regression Models*, J. Iran. Stat. Soc. **10** (2), 201-235, 2011.

[15] Khalili, A., and Chen, J. *Variable Selection in Finite Mixture of Regression Models*, J. Amer. Statist. Assoc. **102** (479), 1025-1038, 2007.

[16] Khalili, A., and Lin, S. *Regularization in finite mixture of regression models with diverging number of parameters*, Biometrics **69** (2), 436-446, 2013.

[17] Luo, R., Wang, H., and Tsai, C.-L.*On Mixture Regression Shrinkage and Selection Via the MR-Lasso*, Int. J. Pure. Ap. Mat. **46**, 403-414, 2008.

[18] Markatou, M. *Mixture models, robustness, and the weighted likelihood methodol- ogy*, Biometrics **56** (2), 483-486, 2000.

[19] Maronna, R. A. et al. *Robust M-Estimators of Multivariate Location and Scatter*, Ann. Statist. **4** (1), 51-67, 1976.

[20] McLachlan, G., and Peel, D. *Finite mixture models*(John Wiley & Sons, 2004).

[21] Neykov, N., Filzmoser, P., Dimova, R., and Neytchev, P. *Robust fitting of mix- tures using the trimmed likelihood estimator*, Comput. Stat. Data. An. **52** (1), 299-308, 2007.

[22] Rousseeuw, P. *Multivariate estimation with high breakdown point*, status: published, 1985.

[23] Rousseeuw, P. J., and Driessen, K. V. *A fast algorithm for the minimum covariance determinant estimator*, Technometrics **41** (3), 212-223, 1990.

[24] Shen, H.-b., Yang, J., and Wang, S.-t. *Outlier detecting in fuzzy switching regres- sion models*, in Artitificial Intelligence: Methodology, Systems, and Applications Springer, 208-215, 2004.

[25] Skrondal, A., and Rabe-Hesketh, S. *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models* (CRC Press, 2004).

[26] Song, W., Yao, W., and Xing, Y. *Robust mixture regression model fitting by Laplace distribution*, Comput. Stat. Data. An. **71**, 128-137, 2014.

[27] Stahel, W. *Robust estimation: Infinitesimal optimality and covariance matrix esti- mators*, Unpublished doctoral dissertation, ETH, Zurich, Switzerland, 1981.

[28] Wang, P., Puterman, M. L., Cockburn, I., and Le, N. *Mixed Poisson regression models with covariate dependent rates*, Biometrics **52** (2), 381-400, 1996.

[29] Wedel, M. *Market segmentation: Conceptual and methodological foundations*(Springer, 2000).

[30] Wei, Y. *Robust mixture regression models using t-distribution*, Master's thesis, 2012.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 561 – 573

Modified Welch test statistic for ANOVA under Weibull distribution

Derya Karagöz[*†‡]

Abstract

A modification to Welch test statistic is proposed to test the equality of population means of various groups under a Weibull distribution. The proposed test statistic is simple and corresponds to the standard Welch test statistic in which the maximum likelihood mean and variance estimators are replaced with robust estimators based on quantile, quantile least square and repeated median. The influence function and breakdown point of these robust estimators are obtained to show their robustness properties. In the simulation study, various experimental designs are considered to evaluate the performance of proposed modified Welch classical ANOVA tests in terms of the type I-errors studies via simulation study.

Keywords: ANOVA, Welch Test Statistic, Robustness of estimators.

2000 AMS Classification: 62K25, 62J10, 62E17

Received : 03.09.2014 *Accepted* : 10.03.2015 *Doi* : 10.15672/HJMS.2015569632

1. Introduction

Analysis of variance (ANOVA) is one of the most used model which can be seen in many fields such as medicine, engineering, agriculture, education, psychology, sociology and biology to investigate the source of the variations. In general, the main interest is in testing the homogeneity of group means using the classical ANOVA which uses F-test statistic. One-way ANOVA is based on assumptions that the normality of the observations and the homogeneity of group variances. If the assumptions of normality and homogeneity of variances are invalid and also outliers are present, classical ANOVA does not give accurate results. Therefore, test statistics based on robust methods should be used instead of the classical ANOVA.

---

[*]Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey, Email: deryacal@hacettepe.edu.tr

[†]Corresponding Author.

[‡]This work forms part of the author's PhD thesis [8]. The author wishes to thank her research supervisor Prof. Dr. Tülay Saraçbaşı for her advice and support throughout.

The one-way ANOVA under the violation of assumptions has been studied extensively. To deal with non-normal data and/or heteroscedastic variances across groups, many alternatives such as Q, Welch, Brown-Forsythe and Modified Brown-Forsythe tests have been developed instead of classical ANOVA. The statistic Q has been extensively studied by many authors under a variety of assumptions. It is one of the most commonly used test statistic for homogeneity at population means in meta-analysis, see for example [5], [12]. [3] showed that under the null hypothesis Q asymptotically follows a Chi-Square distribution. [7] and [13] derived improved approximations to the distribution of Q under the null hypothesis; these approximations are more accurate for small sample sizes of groups. [9] extended the methods of Welch to find approximate distributions to Q under alternative hypotheses. [9] provide approximations for the non-null distributions of their weighted statistics which are found to be useful in obtaining approximations to the power of the Welch test. A number of authors have discussed extensions of the Welch methods based on the use of robust estimators for the population location and scale parameters. Notable among these are the efforts of [14], [15], and the references contained there in. [10] consider three common robust estimators: Huber's proposal two estimator of location and scale, Hampel's M-estimator of location with scale estimated by the median absolute deviation (MAD), and the trimmed mean with scale estimated by the Winsorized standard deviation.

One of the important assumptions of ANOVA is normality. However, in the application this assumption does not work for the real life data modeled by the exponential, Weibull or lognormal distributions especially in reliability, engineering and life science field. The characteristics of these distributions can be explained by Weibull distribution which is also known as Extreme Value Type III minimum distribution. This has made it extremely popular among reliability engineering, biology and medicine. This distribution is the most commonly used distribution for modeling reliability data, because it represents a wide range of asymmetric distributions. Moreover, ANOVA cannot handle censored or interval data because of the non-normality. The simplest possible lifetime distribution is exponential distribution. However, its constant hazard rate is improper and unrealistic in many cases. Gamma distribution is another candidate distribution for lifetimes. Nevertheless, distribution function or survival function of gamma distribution cannot be expressed in a closed form if the shape parameter is not an integer. Since it is in terms of an incomplete gamma function, one needs to obtain the distribution function, survival function or the hazard rate by numerical integration. This makes gamma distribution little bit unpopular compared to the Weibull distribution, which has a nice distribution function, survival function and hazard function [6]. The Weibull distribution was introduced by the Swedish physicist Weibull (1951). He claimed that his distribution applied to a wide range of problems and illustrated this point with seven examples ranging from the strength of steel to the height of adult males in the British Isles [1]. It has been used in many different fields like material science, engineering, physics, chemistry, meteorology,medicine, pharmacy, quality control, biology, geology, geography, economics and business.

This paper proposes a modified Welch test statistic to test the equality of population means of groups by utilizing robust estimators for the means and variances of Weibull distribution with outlier, and evaluates the performance of modified test in terms of the type I-errors via simulation study. The modified test statistic is called robust Welch (RW) test statistic. Since it is obtained by using robust mean and variance estimators based on quantile (Q), quantile least square (QLS) and repeated median ($Rmed$) instead of maximum likelihood. The influence function (IF) and breakdown point (BP) of robust estimators of mean and variance are obtained to show their robustness properties. The behavior of the developed robust test statistic is examined by Monte-Carlo simulation study. In the simulation study, various experimental designs are considered such as balanced and unbalanced sample sizes for k=3,6 groups with homogeneous and heterogeneous variances. The type I errors of the improved robust test statistic and classical ANOVA under the Weibull distribution are obtained.

The remainder of the paper is organized as follows. Section 2 introduces robust modified Welch test statistics. Section 3 gives explicit robust estimators of the mean and variance of Weibull distribution. The most important robustness measures are IF and BP that are derived in Section 4. To show the performance of the considered test statistic, a simulation study and the results are presented in Section 5. The last section summarizes the conclusions of the study.

2. Robust Welch Test Statistic

The Welch test statistic was proposed by [13] as following:

$$W = \frac{\hat{q}_w}{k-1}\{1 + \frac{2(k-2)\hat{A}}{k^2-1}\}^{-1}. \tag{2.1}$$

where

$$\hat{A} = \sum_{i=1}^{k}\left[1-(\hat{w}_i/\sum_{j=1}^{k}\hat{w}_j)\right]^2/v_i \tag{2.2}$$

$$q_w \equiv \sum_{i=1}^{k}\hat{w}_i(\hat{\mu}_i-\hat{\mu}_w)^2 \tag{2.3}$$

and $v_i = n_i - 1$ is the degrees of freedom for i. sample. In (2.2) and (2.3), $\hat{\mu}_i$ is maximum likelihood estimator of the mean for each sample, $\hat{\sigma}_i^2$ is the maximum likelihood estimator of variance and $\hat{w}_i \equiv n_i/\hat{\sigma}_i^2$ is the estimator of weights based on variance estimator. If the appropriate weights are known, the value of estimation is

$$\hat{\mu}_w = \sum_{i=1}^{k}\hat{w}_i\hat{\mu}_i/\sum_{i=1}^{k}\hat{w}_i. \tag{2.4}$$

The Welch test statistic has approximately F_{k-1,v_w} distribution with $k-1$, $v_w = \frac{(k^2-1)}{3\hat{A}}$ degrees of freedom [13].

In this study, a modification to the Welch test statistic is proposed under a Weibull distribution. The test statistic is simple and corresponds to the standard Welch test statistic in which the maximum likelihood mean and variance estimators are replaced with robust estimators based on Q, QLS and $Rmed$. So the robust Welch test statistic is given by

$$RW = \frac{\hat{q}_{rw}}{k-1}\{1 + \frac{2(k-2)\hat{A}_r}{k^2-1}\}^{-1}. \tag{2.5}$$

where

$$\hat{A}_r = \sum_{i=1}^{k}\left[1-(\hat{w}_{ri}/\sum_{j=1}^{k}\hat{w}_{rj})\right]^2/v_i \tag{2.6}$$

$$q_{rw} \equiv \sum_{i=1}^{k}\hat{w}_{ri}(\hat{\mu}_{ri}-\hat{\mu}_{rw})^2 \tag{2.7}$$

and $v_i = n_i - 1$ is the degrees of freedom for i. sample.

In (2.7), $\hat{\mu}_{ri}$ is robust estimator of mean for each sample, $\hat{\sigma}^2_{ri}$ is robust estimator of variance and $\hat{w}_{ri} \equiv n_i/\hat{\sigma}^2_{ri}$ is the robust estimator of weights based on variance estimator. If the appropriate weights are known, the value of robust estimation are

$$\hat{\mu}_{rw} = \sum_{i=1}^{k}\hat{w}_{ri}\hat{\mu}_{ri}/\sum_{i=1}^{k}\hat{w}_{ri}. \tag{2.8}$$

In (2.7) and (2.8), $\hat{\mu}_{ri}$ and $\hat{\sigma}^2_{ri}$ are the robust Q, QLS and *Rmed* estimators of mean and variance for Weibull distribution.

The robust Welch test statistic has approximately $F_{k-1,v_{rw}}$ distribution with $k-1$, $v_{rw} = \frac{(k^2-1)}{3\hat{A}_r}$ degrees of freedom.

3. Robust Estimators of Weibull Distribution

The estimations of mean and variance of Weibull distribution are a basic subject of the paper. The density function $f(x;\lambda,\beta) = \frac{\beta}{\lambda}(x/\lambda)^{\beta-1}\exp[-(x/\lambda)^{\beta}],\ x,\lambda,\beta > 0$. The parameter λ is called a scale parameter. The parameter β is the shape parameter. The mean and variance of a Weibull random variable are the functions of the shape β and scale λ parameters can be expressed as $E(X) = \lambda\Gamma(1+1/\beta)$ and $Var(X) = \lambda^2[\Gamma(1+2/\beta) - \Gamma^2(1+1/\beta)]$. We consider robust estimators which were proposed by [2] to achieve the robust estimates of the mean and variance of this distribution. The estimators proposed by [2] are robust to outliers, but they have the additional advantage of being an explicit function of the data.

In this study we restrict our attention to estimators that have the following set of properties: an explicit formula; a 50% breakdown point and a bounded IF. We propose the robust estimators of mean and variance by considering robust estimators based on Q, QLS and *Rmed*. We also derive their IFs and their breakdown points. The proposed estimators for mean and variance all have a high breakdown point and bounded IF.

In the following Section 3.1 and Section 3.2, quantile and regression estimators are given for robust mean and variance estimators of Weibull distribution.

3.1. Quantile-estimators. The quantile estimators of mean and variance for Weibull distribution are given by

$$\begin{aligned}\hat{\mu}_{\mathrm{WQ}} &= \hat{\lambda}_{\mathrm{Q}}\Gamma(1+1/\hat{\beta}_{\mathrm{Q}}),\\ \hat{\sigma}_{\mathrm{WQ}} &= \hat{\lambda}^2_{\mathrm{Q}}[\Gamma(1+2/\hat{\beta}_{\mathrm{Q}}) - \Gamma^2(1+1/\hat{\beta}_{\mathrm{Q}})]\end{aligned} \tag{3.1}$$

where [2] proposed the quantile estimators of shape and scale parameter
$\hat{\beta}_{\mathrm{Q}} = \frac{1}{\log(\hat{q}_{\alpha_2}/\hat{q}_{\alpha_1})}\log\frac{\log(1-\alpha_2)}{\log(1-\alpha_1)}$, and $\hat{\lambda}_{\mathrm{Q}} = \hat{q}_{\alpha}/[-\log(1-\alpha)]^{1/\hat{\beta}_{\mathrm{Q}}}$.

3.2. Regression estimators. The quantiles of the general log-Weibull distribution in $G^{-1}_{\lambda,\beta}(\alpha) = \beta^{-1}\log(-\log(1-\alpha)) + \log\lambda$ are linearly related to the quantiles of the standard log-Weibull distribution, with intercept $b_0 = \log\lambda$ and slope $b_1 = 1/\beta$. Replacing the theoretical quantiles with their empirical counterparts yields a linear regression equation $y_i = b_0 + b_1 z_i + \varepsilon_i$ where $y_i = \log \hat{q}_{i/(n+1)}$ and $z_i = G^{-1}(i/(n+1))$. [2] considered two robust and explicit regression estimators for b_1 and b_0: the *Quantile Least Squares* and the *Repeated Median* estimators. The corresponding estimates of scale and shape of the Weibull distribution were then directly given by $\hat{\lambda} = \exp(\hat{b}_0)$ and $\hat{\beta} = 1/\hat{b}_1$.

Quantile Least Square: The QLS estimators of mean and variance of Weibull distribution are given by

$$\begin{aligned} \hat{\mu}_{\mathrm{WQLS}} &= \hat{\lambda}_{\mathrm{QLS}}\Gamma(1+1/\hat{\beta}_{\mathrm{QLS}}) \\ \hat{\sigma}_{\mathrm{WQLS}} &= \hat{\lambda}^2_{\mathrm{QLS}}[\Gamma(1+2/\hat{\beta}_{\mathrm{QLS}}) - \Gamma^2(1+1/\hat{\beta}_{\mathrm{QLS}})] \end{aligned} \tag{3.2}$$

where the QLS estimators of shape and scale parameters proposed by [2] :
$\hat{\lambda}_{\mathrm{QLS}} = \exp(\hat{b}_{0\mathrm{QLS}})$ and $\hat{\beta}_{\mathrm{QLS}} = 1/\hat{b}_{1\mathrm{QLS}}$ where $\hat{b}_{0\mathrm{QLS}}$ and $\hat{b}_{1\mathrm{QLS}}$ are QLS regression estimators for b_0 and b_1 (for details see [2]).

Repeated Median: The *Rmed* estimators of mean and variance of Weibull distribution are given by

$$\begin{aligned} \hat{\mu}_{\mathrm{WRmed}} &= \hat{\lambda}_{\mathrm{Rmed}}\Gamma(1+1/\hat{\beta}_{\mathrm{Rmed}}) \\ \hat{\sigma}_{\mathrm{WRmed}} &= \hat{\lambda}^2_{\mathrm{Rmed}}[\Gamma(1+2/\hat{\beta}_{\mathrm{Rmed}}) - \Gamma^2(1+1/\hat{\beta}_{\mathrm{Rmed}})] \end{aligned} \tag{3.3}$$

where the *Rmed* estimators of shape and scale parameters were proposed by [2]: $\hat{\lambda}_{\mathrm{Rmed}} = \exp(\hat{b}_{0\mathrm{Rmed}})$ and $\hat{\beta}_{\mathrm{Rmed}} = 1/\hat{b}_{1\mathrm{Rmed}}$ where $\hat{b}_{0\mathrm{Rmed}}$ and $\hat{b}_{1\mathrm{Rmed}}$ are *Rmed* regression estimators for b_0 and b_1 (for details see [2]).

4. Robustness of estimators

Robustness of estimators can be measured in different ways. The most important robustness measures are IF and BP of the estimators. In this study we derive IF and BP for the proposed estimators of mean and variance for Weibull distribution. The IF acts like the first derivative of functional defined on empirical distributions which we evaluate at the estimator. IF should be bounded to be robust. The breakdown point is a global robustness measure which describes how many percent gross errors are still tolerated before increasingly offensive outliers force our estimator to wander off to infinity. In next Section 4.1 and Section 4.2 we derive their IF and then their breakdown points, respectively.

4.1. IFs for Proposed Robust Estimators. IF gives price information about how to respond to a small amount of distortion at any point. Naturally, the estimators are very sensitive form of the distribution F, too much affected by deterioration in small quantities.

The statistical functionals corresponding with the mean and variance robust estimators are given by

$$\mu_{\mathrm{RE}}(F) = \lambda_{\mathrm{RE}}(F)\Gamma(1+1/\beta_{\mathrm{RE}}(F)) \tag{4.1}$$

$$\sigma_{\mathrm{RE}}(F) = \lambda^2_{\mathrm{RE}}(F)[\Gamma(1+2/\beta_{\mathrm{RE}}(F)) - \Gamma^2(1+1/\beta_{\mathrm{RE}}(F))]. \tag{4.2}$$

The IF of the functional μ_{RE} at the Weibull distribution $F_{\lambda,\beta}$ in (4.1) is given by

$$\begin{aligned} IF(x_0;\mu_{\mathrm{RE}},F_{\lambda,\beta}) &= \frac{\partial}{\partial\varepsilon}(\mu_{\mathrm{RE}}(F_\varepsilon))|_{\varepsilon=0} \\ &= \Gamma(1+1/\hat{\beta}_{\mathrm{RE}})\Big(IF(x_0;\lambda_{\mathrm{RE}},F_{\lambda,\beta}) \\ &- \frac{\hat{\lambda}_{\mathrm{RE}}}{\hat{\beta}^2_{\mathrm{RE}}}\psi(1+1/\hat{\beta}_{\mathrm{RE}})IF(x_0;\beta_{\mathrm{RE}},F_{\lambda,\beta})\Big). \end{aligned} \tag{4.3}$$

The IF of the functional σ_{RE} at the Weibull distribution $F_{\lambda,\beta}$ in (4.2) is given by

$$\begin{aligned} IF(x_0;\sigma_{\mathrm{RE}},F_{\lambda,\beta}) &= \frac{\partial}{\partial\varepsilon}(\sigma_{\mathrm{RE}}(F_\varepsilon))|_{\varepsilon=0} \\ &= 2\hat{\lambda}_{\mathrm{RE}}IF(x_0;\lambda_{\mathrm{RE}},F_{\lambda,\beta})[\Gamma(1+2/\hat{\beta}_{\mathrm{RE}})-\Gamma(1+1/\hat{\beta}_{\mathrm{RE}})^2] \\ &+ 2\frac{\hat{\lambda}^2_{\mathrm{RE}}}{\hat{\beta}^2_{\mathrm{RE}}}IF(x_0;\beta_{\mathrm{RE}},F_{\lambda,\beta})\Big[-\psi(1+2/\hat{\beta}_{\mathrm{RE}})\Gamma(1+2/\hat{\beta}_{\mathrm{RE}}) \\ &+ \Gamma(1+1/\hat{\beta}_{\mathrm{RE}})^2\psi(1+1/\hat{\beta}_{\mathrm{RE}})\Big]. \end{aligned} \tag{4.4}$$

where $\hat{\beta}_{\mathrm{RE}}$ and $\hat{\lambda}_{\mathrm{RE}}$ are the shape and scale robust Q, QLS and $Rmed$ estimators of Weibull parameters. For IFs $IF(x_0;\lambda_{\mathrm{RE}})$ and $IF(x_0;\beta_{\mathrm{RE}})$ in (4.3) and (4.4), see [2].

The IFs for the classic and robust estimators of mean and variance are pictured in Figure 1. It is seen that while the IF of least square (LS) estimator is unbounded function, the IFs for robust estimators are bounded. It should be considered that the IFs of quantile mean and variance estimator are step functions. As a result we can say that the proposed estimators are B-robust which means that their IFs are bounded.

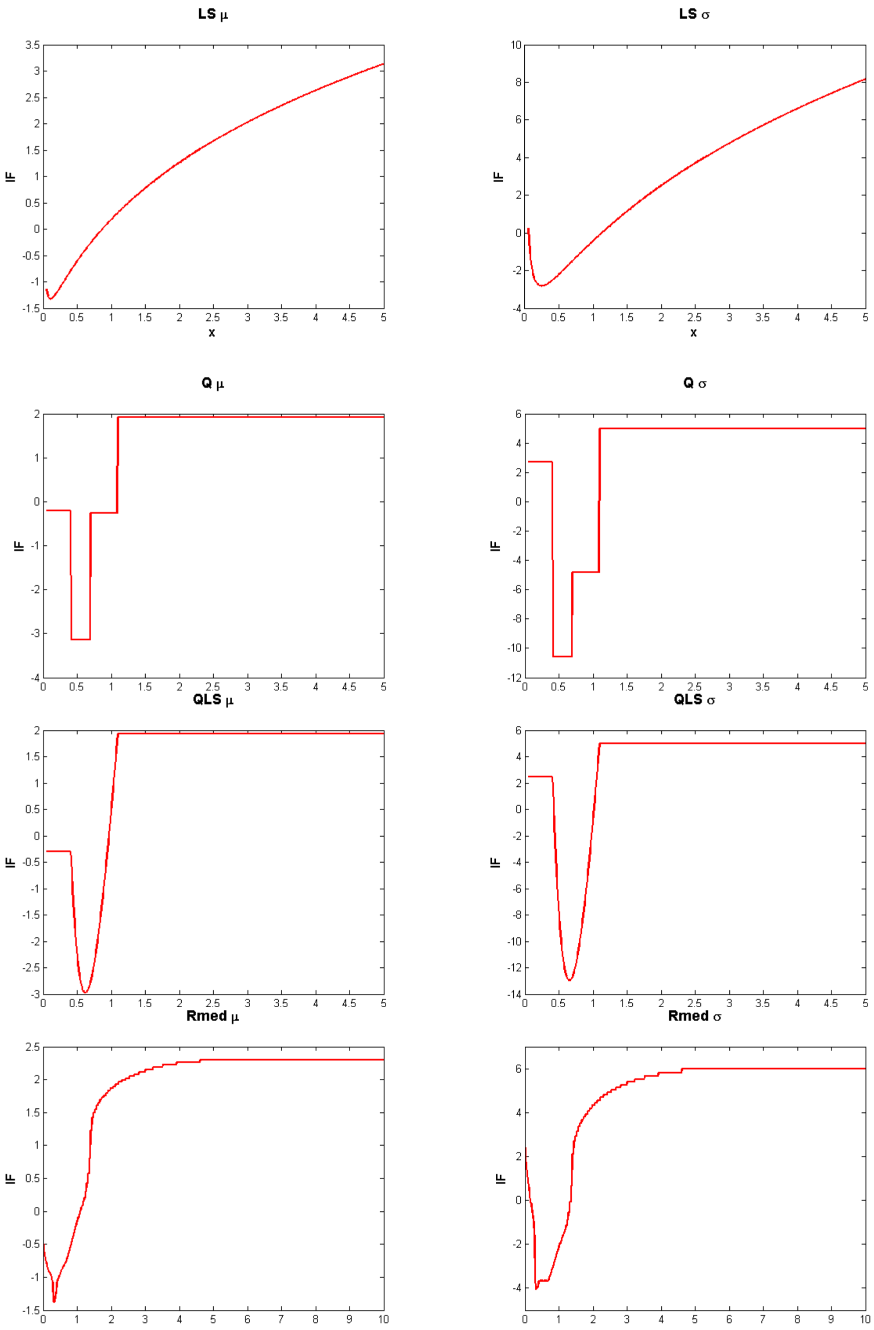
4.2. Breakdown Points of Proposed Robust Estimators. The breakdown point of an estimator is the proportion of incorrect observations an estimator can handle before given an arbitrarily large result. The higher the breakdown point of an estimator, the more robust it is. Instinctively, we can understand that a breakdown point can not exceed %50 because if more than half of the observations are contaminated, it is not possible to distinguish between the underlying distribution and the contaminating distribution. Therefore, the maximum breakdown point is 0.5 and there are estimators which achieve such a breakdown point.

The breakdown points of robust estimators were examined earlier in the previous studies examined by some authors: For linear regression parameters least square estimators: α and repeated median estimators: %50. For shape and scale estimators based on Q, QLS, $Rmed$ methods : [2]. To characterize the robustness of the proposed estimators, we derive their BP, defined as the smallest proportion of observations (for $n \to \infty$) that needs to be replaced with arbitrary values in order for the estimation of λ or β to be arbitrarily close to zero (implosion) or infinity (explosion). To define the breakdown point of the mean and variance we consider the BP of shape and scale estimators of Weibull distribution.

The BP of the mean estimator for Weibull distribution is given by

$$\varepsilon^*_n(\mu,F_n) = min\{\varepsilon^+_n(\mu,F_n),\varepsilon^-_n(\mu,F_n)\}, \tag{4.5}$$

Figure 1. IF of the mean and variance estimators of Weibull Distribution



In 4.5 the explosion BP is

$$\varepsilon_n^+(\mu, F_n) = min\{\frac{m}{n}, m \in 1, \ldots, n | \sup_{F_n'} M(\mu(F_n')) = \infty\}. \tag{4.6}$$

We get $\sup_{F'_n} M(\mu(F'_n)) = \infty$, if $\lambda \to \infty$ or $\gamma(1 + 1/\beta) \to \infty$. For $\lambda \to \infty$ the BP is $\varepsilon^+(\lambda, F)$. For $\gamma(1 + 1/\beta) \to \infty$, if $\beta = -1, -1/2, -1/3, -1/4$. In this condition there is no BP since β is not going to infinity or zero. Therefore the explosion BP is $\varepsilon^+(\mu, F) = (\varepsilon^+(\lambda, F))$.

In 4.5 the implosion BP is

(4.7) $$\varepsilon_n^-(\mu, F_n) = min\{\frac{m}{n}, m \in 1, \ldots, n | \inf_{F'_n} M(\mu(F'_n)) = 0\}.$$

We get $\inf_{F'_n} M(\mu(F'_n)) = 0$, if $\lambda \to 0$ or $\gamma(1+1/\beta) \to 0, \varepsilon^+(\beta, F)$. For $\lambda \to 0$ the BP is $\varepsilon^-(\lambda, F)$. For $\gamma(1+1/\beta) \to 0, \varepsilon^+(\beta, F)$: if $\beta \to \infty$, $\gamma(1+1/\beta) = 1/\beta\gamma(1/\beta) \to 0$. So for $\beta \to \infty$ the BP is $\varepsilon^+(\beta, F)$. The implosion BP is $\varepsilon^-(\mu, F) = (\varepsilon^-(\lambda, F), \varepsilon^+(\beta, F))$. As a result the BP of the mean estimator is given by

$$\begin{aligned} \varepsilon_n^*(\mu, F_n) &= \min\{\varepsilon_n^+(\mu, F_n), \varepsilon_n^-(\mu, F_n)\} \\ &= \min\{\varepsilon^+(\lambda, F), \varepsilon^-(\lambda, F), \varepsilon^+(\beta, F)\} \end{aligned}$$

The BP of variance estimator of Weibull distribution is given by

(4.8) $$\varepsilon_n^*(\sigma, F_n) = min\{\varepsilon_n^+(\sigma, F_n), \varepsilon_n^-(\sigma, F_n)\}.$$

In 4.8 the explosion BP is $\varepsilon_n^+(\sigma, F_n) = min\{\frac{m}{n}, m \in 1, \ldots, n | \sup_{F'_n} M(\sigma(F'_n)) = \infty\}$ We get $\sup_{F'_n} M(\sigma(F'_n)) = \infty$,if $\lambda \to \infty$ or $\gamma(1 + 2/\hat{\beta}) - \gamma(1 + 1/\hat{\beta}) > 0$. For $\lambda \to \infty$ can be obtained if $(\varepsilon_n^+(\lambda, F))$. For $\gamma(1 + 2/\hat{\beta}) - \gamma(1 + 1/\hat{\beta}) > 0$ can be obtained if $\beta > 0$. So the BP is $\varepsilon_n^+(\beta, F)$. Therefore the explosion BP of variance estimator is obtained $\varepsilon^+(\sigma, F) = (\varepsilon^+(\lambda, F), \varepsilon^+(\beta, F))$.

In 4.8 the implosion BP is $\varepsilon_n^+(\sigma, F_n) = min\{\frac{m}{n}, m \in 1, \ldots, n | \inf_{F'_n} M(\sigma(F'_n)) = 0\}$ We get $\inf_{F'_n} M(\sigma(F'_n)) = 0$, if $\lambda \to 0$ or $\beta \to \infty$. For $\lambda \to 0$, the BP is $\varepsilon^-(\lambda, F)$. For $\beta \to \infty$ the BP is$\varepsilon^+(\beta, F)$. Therefore the implosion BP of variance estimator is obtained $\varepsilon^-(\sigma, F) = (\varepsilon^-(\lambda, F), \varepsilon^+(\beta, F))$.

As a result the BP of variance estimator is given by

$$\begin{aligned} \varepsilon_n^*(\sigma, F_n) &= \min\{\varepsilon_n^+(\sigma, F_n), \varepsilon_n^-(\sigma, F_n)\} \\ &= \min\{\varepsilon^+(\lambda, F), \varepsilon^+(\beta, F), \varepsilon^-(\lambda, F)\} \end{aligned}$$

Asymptotic BPs for the robust estimators of mean and variance of Weibull distribution are given by Table 1.

Table 1. Asymptotic BP for the robust estimators for mean and variance of Weibull distribution

Method	$\varepsilon^*(\mu)$	$\varepsilon^*(\sigma)$
Q	$\min(\alpha, 1-\alpha, \alpha_1, \alpha_2 - \alpha_1)$ $\alpha \neq 1 - e^{-1}$ $\min(\alpha, 1-\alpha)$ $\alpha = 1 - e^{-1}$	$\min(\alpha, 1-\alpha, \alpha_1, \alpha_2 - \alpha_1)$ $\alpha \neq 1 - e^{-1}$ $\min(\alpha, 1-\alpha, \alpha_2 - \alpha_1)$ $\alpha = 1 - e^{-1}$
QLS	$\min(\overline{\alpha}, 1 - 2\overline{\alpha})$	$\min(\overline{\alpha}, 1 - 2\overline{\alpha})$
Rmed	%50	%50

5. Simulation Study

The behavior of the robust Welch test statistic is examined according to all three methods with 10,000 repetitions. The type I errors of proposed test statistic is obtained according to robust methods by considering various experimental designs. At the end of the simulation study robust test statistic will be compared in terms of the type I errors, and the comments will be made for experimental designs.

Since the mean and variance of Weibull distribution are functions of the shape and scale parameters, the creation of different combinations depends only on the parameters of the distribution. When scale parameter is one and shape parameter takes different values, the mean and variance do not change much. However, when the scale parameter value is changed, the mean and variance change a lot. Therefore, to generate different experimental design a scale parameter is fixed, it takes $\lambda = 1$ with different values of shape parameter. For example when the shape parameter β is equal to one, then this distribution reduces to the exponential distribution. A model that results in values of probability $prob\{y \geq E(y)\}$ substantially greater or smaller than 0.5 is hardly of any practical interest. For the values of β less than 1.2, $prob\{y \geq E(y)\} < 0.4$ [11]. Moreover, [4] argue that in most applications where a Weibull distribution is applicable β is greater than one. For these reasons, we consider values of $\beta \geq 1.5$. In the simulation study, the value of the shape parameters are selected as in Table 2 with respect to the different experimental designs which we want to create. In this table for equal means, homogeneous variances A1, B1 are used for balanced sample size, C1, D1 are used for unbalanced sample size. For unequal means, heterogenous variances A2, B2 are used for balanced sample size and C2, D2 are used for unbalanced sample size. As we mentioned before to generate data with equal means + homogeneous variances and unequal means + heterogenous variances, we just change the shape parameter of Weibull distribution.

Table 2. Experimental designs for k=3 and k=6

		$k=3$			$k=6$					
A1	n_i	5	5	5	5	5	5	5	5	5
	β	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
A2	n_i	5	5	5	5	5	5	5	5	5
	β	1.5	2	2.5	1.5	2	2.5	1.5	2	2.5
B1	n_i	10	10	10	10	10	10	10	10	10
	β	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
B2	n_i	10	10	10	10	10	10	10	10	10
	β	1.5	2	2.5	1.5	2	2.5	1.5	2	2.5
C1	n_i	5	10	15	5	10	15	5	10	15
	β	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
C2	n_i	5	10	15	5	10	15	5	10	15
	β	1.5	2	2.5	1.5	2	2.5	1.5	2	2.5
D1	n_i	10	20	30	10	20	30	10	20	30
	β	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
D2	n_i	10	20	30	10	20	30	10	20	30
	β	1.5	2	2.5	1.5	2	2.5	1.5	2	2.5

In the simulation study the reference distribution is $W(1,\beta)$ whose characteristic is mentioned in Table 2. By using the proposed estimators Q, QLS and $Rmed$, the type I errors of the robust test statistic and classical ANOVA are obtained with 10,000 repetitions at the significance level of 5%. For quantile methods, $\alpha = 30\%$ is taken. Four different models are discussed below to test the behaviors of the test statistic, when the model is deteriorated and in the presence of outliers:

- Model 1: Clean sample (Reference distribution $W(1,\beta)$),
- Model 2: Dixon model (n-1 observations from $W(1,\beta)$, 1 observation from $W(2,\beta)$),
- Model 3: Mixture model ($0.80W(1,\beta)$ + 0.20 $W(2,\beta/2)$),
- Model 4: Contaminated model (0.80 $W(1,\beta)$ + 0.20(100 Uniform(0,1))).

We obtain $\tau = \frac{\sum_{i=1}^{M} F_{H_i} > F_{T_i}}{M} * 100$ value with respect the classical ANOVA and RW test statistics with 10,000 repetitions. In this equation F_{H_i} indicates the calculated test statistic and F_{T_i} indicates the F table value at the significance of 5% for the ith simulation, so desirable value of τ is to be close $\tau \cong 5$.
By considering combinations of the above-mentioned trial simulation study, the type I errors of test statistic based on three methods is obtained and the results (type I error * 100=$\boldsymbol{\tau}$) are given. The robust test statistic will be compared in terms of type I errors and the comments in detail for each trial will be made.

The results of $\boldsymbol{\tau}$ values for experimental designs with equal means and homogeneous variances are given in Table 3 for $k = 3$. While the classical ANOVA does not deteriorate for clean model (model 1), it badly deteriorates for contaminated model especially for unbalanced sample size. As seen from this table, the results of Q methods are not good. The type I errors of RW test statistic based on $Rmed$ methods are desired level especially for experimental design C1 and D1. RW test statistic based on QLS method can be an alternative only for experimental design C1.

Table 3. The $\boldsymbol{\tau}$ values for k=3, Equal means, homogeneous variances

RE	Model	A1		B1		C1		D1	
		F	**RW**	**F**	**RW**	**F**	**RW**	**F**	**RW**
Q	1	4.80	6.71	5.16	11.8	6.36	11.03	6.37	14.37
	2	3.90	10.55	4.27	13.22	5.53	15.99	6.17	17.68
	3	3.10	9.27	3.61	10.80	4.54	12.98	5.40	15.55
	4	1.60	9.70	0.60	13.90	28.85	11.19	83.61	15.75
QLS	1	5.14	6.91	5.08	14.45	6.23	7.93	6.79	18.42
	2	3.75	2.45	4.23	5.38	5.32	7.89	6.21	8.20
	3	3.11	3.85	3.61	10.40	4.44	5.45	5.58	15.89
	4	1.60	9.70	0.08	18.99	29.85	5.87	82.85	14.91
Rmed	1	4.53	8.96	4.86	6.07	6.64	13.91	6.19	7.00
	2	3.89	7.08	4.46	5.66	5.02	6.86	5.90	7.72
	3	3.26	5.90	3.63	4.43	4.76	6.22	5.37	4.93
	4	0.14	4.51	0.04	4.93	30.25	3.79	82.50	4.88

The results of $\boldsymbol{\tau}$ values for experimental designs with non-equal means and heterogenous variance are given in Table 4 for $k = 3$. For unbalanced sample size classical ANOVA is deteriorate for all methods. Only the RW test statistic based on $Rmed$ has good performance for contaminated model. But RW test statistic does not work for other robust

Table 4. The $\boldsymbol{\tau}$ values for k=3, unequal means, heterogenous variances

RE	Model	A2		B2		C2		D2	
		F	**RW**	**F**	**RW**	**F**	**RW**	**F**	**RW**
Q	1	5.30	7.21	6.35	10.92	12.32	12.46	12.39	14.62
	2	4.74	12.35	4.95	14.32	11.35	18.35	12.05	18.65
	3	3.73	11.35	4.07	12.91	9.21	16.71	9.83	16.35
	4	0.16	9.70	0.06	13.56	32.66	11.37	83.47	14.96
QLS	1	5.52	8.44	5.20	15.99	12.13	16.71	12.55	18.78
	2	4.60	3.76	4.85	5.81	9.04	13.21	12.05	8.93
	3	3.90	5.19	4.05	12.87	8.75	14.08	10.22	17.25
	4	0.17	4.89	0.05	12.91	31.96	12.41	83.74	14.43
Rmed	1	5.89	8.54	5.47	6.61	12.20	9.33	12.63	7.19
	2	4.74	8.23	5.38	5.95	11.25	10.17	12.08	7.98
	3	4.12	6.22	4.41	4.68	8.78	7.77	9.31	8.00
	4	0.15	4.04	0.04	4.29	32.09	4.20	83.37	5.11

methods. For clean, and very few corrupted samples the Type I error level of classical ANOVA is considerably good since the variances are homogeneous for D1 experimental design. However for contaminated model classical ANOVA is deteriorated badly.

The results of $\boldsymbol{\tau}$ values for experimental designs with equal means and homogeneous variances are given in Table 5 for $k = 6$. As seen from the results, the *RW* test statistic based on *Rmed* robust method works well for mixture and contaminated model in only A2 and B2 experimental designs, but the other robust methods do not work.

Table 5. The $\boldsymbol{\tau}$ values for k=6, Equal means, homogeneous variances

RE	Model	A1		B1		C1		D1	
		F	**RW**	**F**	**RW**	**F**	**RW**	**F**	**RW**
Q	1	4.70	10.61	4.92	16.91	5.89	17.73	6.13	23.98
	2	3.90	10.55	4.03	24.89	4.72	27.97	5.43	23.12
	3	3.20	20.47	3.74	22.02	4.07	24.93	4.60	26.45
	4	0.9	22.12	0.77	24.45	61.31	24.07	86.55	28.09
QLS	1	6.03	17.21	4.74	29.80	13.00	35.22	5.65	33.07
	2	3.97	5.15	4.25	9.31	4.53	34.30	5.14	13.73
	3	3.28	9.22	3.28	24.20	4.13	26.12	4.17	28.91
	4	0.13	11.47	0.95	14.19	60.99	28.44	85.85	30.00
Rmed	1	5.10	25.22	5.10	9.50	5.67	27.70	6.18	10.71
	2	3.65	15.03	4.34	9.69	4.96	12.59	5.07	10.65
	3	3.10	11.98	3.39	6.73	4.55	21.27	4.34	7.20
	4	0.14	10.33	0.94	7.10	61.50	7.81	85.69	7.70

The results of $\boldsymbol{\tau}$ values for experimental designs with non-equal means and heterogenous variance are given in Table 6 for $k = 6$. As you see, we can conclude the results for contaminated model for $(k = 6)$ such as that the *RW* test statistic only based on *Rmed* robust method gives desirable results.

To sum up all results, we can say that in the case of contamination proposed robust Welch test statistic can be used for ($k = 3$). When the number of group is small, for contaminated models the Type I errors of RW test statistic has good performance. The number of group is small *Rmed* according to the methods of RW test statistic Type I errors is desirable. The number of group grows, RW test statistic has undergone distortion.

Table 6. The $\boldsymbol{\tau}$ values for k=6, unequal means, heterogenous variances

RE	Model	A2		B2		C2		D1	
		F	**RW**	**F**	**RW**	**F**	**RW**	**F**	**RW**
Q	1	5.94	11.72	5.50	30.62	12.73	19.55	13.12	22.71
	2	4.52	24.98	5.24	29.98	11.25	31.71	12.01	29.81
	3	3.41	22.49	4.09	22.43	9.09	27.71	9.48	26.61
	4	0.90	22.12	0.81	24.45	62.43	24.71	86.52	27.93
QLS	1	4.67	20.41	6.12	31.18	13.00	30.06	13.04	31.90
	2	4.53	6.81	5.07	9.81	11.31	13.87	12.10	12.71
	3	3.37	12.73	3.60	26.59	8.47	26.54	9.77	29.36
	4	0.14	10.97	1.13	27.83	62.78	23.09	86.80	26.67
Rmed	1	5.24	16.41	6.07	10.22	12.61	13.52	12.96	10.31
	2	4.38	16.28	4.93	9.65	10.73	15.45	11.90	11.38
	3	3.62	13.25	3.55	7.71	8.97	11.64	9.27	10.28
	4	0.18	9.46	0.91	6.37	62.51	8.78	86.41	6.69

6. Conclusion

The purpose of this study is to develop test statistic for one-way ANOVA by using robust methods under Weibull distribution with outlier. For this purpose, we propose the robust estimators for mean and variance of Weibull distribution. We also derive not only their BP but also their IFs. The proposed estimators for mean and variance all have a high BP and bounded IF. RW test statistic is obtained by using the estimators based on Q, QLS and *Rmed*. The behavior of the modified robust test statistic is examined by simulation study.

In the simulation study, using various experimental designs, type I errors of the improved robust test statistic and classical ANOVA under Weibull distribution are obtained with respect to three different robust estimators. Balanced and unbalanced sample sizes for k=3,6 groups with homogeneous and heterogeneous variances are considered. Then the simulation results show up: For unbalanced sample size classical ANOVA is deteriorated. When the number of groups is small ($k = 3$), the RW test statistic based on *Rmed* and QLS methods performance is not deteriorated badly. When the number of groups is increasing, especially for contaminated models the proposed RW test based on *Rmed* method gives desirable results. The RW test statistic based on *Rmed* has the best performance for all experimental design especially in contaminated models, while the RW test statistic based on Q does not work. QLS method can be used as an alternative to *Rmed* method.

References

[1] Abernethy, R.B., Breneman, J.E., Medlin, C.H. and Reinman, G.L., 1983, *Weibull Analysis Handbook*, Technical Report, AFWAL-TR-83-207, Air Force Wright Aeronautical Laboratories, Washington, D.C., http://handle.dtic.mil/100.2/ADA143100.

[2] Boudt, K., Caliskan, D. and Croux, C., 2011, *Robust explicit estimators of Weibull parameters*, Metrika, 73, 187-209.

[3] Cochran, W.G., 1937, *Problems arising in the analysis of a series of similar experimentals*, Journal of Royal Statistics Soc.Supp., 4, 102-118.

[4] Cohen, A. C. and Whitten, B. J., 1988, *Parameter Estimation in Reliability and Life Span Models*, New York: Marcel Dekker.

[5] DerSimonian, R. and Laird, N., 1986, *Meta-analysis in clinical trials*, Controlled Clin. Trials, 7, 177-188.

[6] Gupta, R.D. and Kundu, D., 2001, *Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions*, Biometrical Journal 43,1, 117-130

[7] James, G., S., 1951, *The comparison of several groups of observations when the ratios of the population variances are unkown*, Biometrika, 38, 324-329.

[8] Karagöz, D., 2012, *Robust Analaysis of Variance*, PhD thesis: Hacettepe University, Ankara, Turkey.

[9] Kulinskaya, E., Staudte, R.G. and Gao, H., 2003, *Power approximations in testing for unequal means in a One-Way ANOVA weighted for unequal variances*, Communications in Statistics Theory and Methods, 32,2353-2371.

[10] Kulinskaya, E. and Dollinger, M.B., 2007, *Robust weighted One-Way ANOVA: Improved approximation and efficiency*, Journal of Statistical Planning and Inference, 137, 462-472.

[11] Şenoğlu, B., 2005, *Robust 2^k Factorial Design with Weibull Error Distributions*, Journal of Applied Statistics, 32-10, 1051-1066.

[12] Takkouche, B., Cadarso-Suarez, C. and Spiegelman, D., 1999, *Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis*, Amer.J. Epidemiology 150, 206-215.

[13] Welch, B.L., 1951, *On the comparison of several mean values: an alternative approach.* Biometrika, 38, 330-336.

[14] Wilcox, R.R., 1995, *The practical importance of heteroscedastic methods, using trimmed means versus means, and designing simulation studies*, British J. Math.Statist.Psych., 48, 99-114.

[15] Wilcox, R.R., 1997, *Introduction to Robust Estimation and Hypothesis Testing*, Academic Press, New York.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 575 – 591

Some classes of shrinkage estimators in the morgenstern type bivariate exponential distribution using ranked set sampling

Housila P. Singh and Vishal Mehta *

Abstract

This article proposes a class of shrinkage estimators of Morgenstern type bivariate exponential distribution (MTBED) based on concomitants of order statistic in ranked set sampling (RSS). The class of estimators for the parameter is motivated by the work of Jani (1991). The proposed class of shrinkage estimators has smaller mean square error (MSE) than the Chacko and Thomas (2008) estimators and minimum mean squared error (MMSE) estimators for wider range of the parameter. Numerical computations indicate that certain of these estimators substantially improve the usual and minimum mean squared error (MMSE) estimators for value of the parameter near the prior estimate, especially for small sample sizes.

2000 AMS Classification: 62G30, 62H12.

Keywords: Ranked set sampling, Morgenstern type bivariate exponential distribution, Concomitants of order statistic, Minimum mean square error estimator, Shrinkage estimator.

Received : 23.01.2014 *Accepted* : 16.06.2014 *Doi* : 10.15672/HJMS.201611415693

*School of Studies in Statistics , Vikram University ,Ujjain-456010, Madhya Pradesh, India. Email : visdewas@gmail.com Corresponding Author.

1. Introduction

The concept of Ranked set sampling (RSS) was first introduced by McIntyre (1952) as a process of improving the precision of the sample mean as an estimator of the population mean. Ranked set sampling as described in McIntyre (1952) is applicable whenever ranking of a set of sampling units can be done easily by a judgement method (for a detailed discussion on the theory and applications of ranked set sampling see, Chen et al., (2004)). Ranking by judgement method is not recommendable if the judgement method is too crude and is not powerful for ranking by discriminating the units of a moderately large sample. In certain situations, one may prefer exact measurement of some easily measurable variable associated with the study variable rather than ranking the units by a crude judgement method. Suppose the variable of interest say Y, is difficult or much expensive to measure, but an auxiliary variable X correlated with Y is readily measurable and can be ordered exactly. In this case as an alternative to McIntyre (1952) method of RSS, Stokes (1977) used an auxiliary variable for the ranking of the sampling units. If $X_{r(r)}$ is the observation measured on the auxiliary variable X from the unit chosen from the rth set then we write $Y_{r[r]}$ to denote the corresponding measurement made on the study variable Y on this unit, then $Y_{r[r]}, r = 1, 2, ..., n$, form the ranked set sample. Clearly $Y_{r[r]}$ is the concomitant of the rth order statistic arising from the rth sample.

Chacko and Thomas (2008) assumed a Morgenstern type bivariate exponential distribution (MTBED) corresponding to bivariate random variable (X, Y), where X denote the auxiliary variable and Y denote the study variable with probability density function (pdf) as

$$f_{XY}(x,y) = \frac{e^{\frac{-x}{\theta_1}} e^{\frac{-y}{\theta_2}}}{\theta_1\theta_2}[1+\alpha(1-2e^{\frac{-x}{\theta_1}})(1-2e^{\frac{-y}{\theta_2}})]; \tag{1.1}$$

$$x > 0, y > 0, \theta_1 > 0, \theta_2 > 0, -1 \le \alpha \le 1.$$

Stokes (1995) has considered the estimation of parameters of location-scale family of distributions using RSS. Lam et al. (1994, 1995) have obtained the best linear unbiased estimators (BLUEs) of location and scale parameters of exponential distribution and logistic distribution. The Fisher information contained in RSS have been discussed by Chen (2000) and Chen and Bai (2000). Stokes (1980) has considered the method of estimation of correlation coefficient of bivariate normal distribution using RSS. Modarres and Zheng (2004) have considered the problem of estimation of dependence parameter using RSS. Robust estimate of correlation coefficient for bivariate normal distribution have been developed by Zheng and Modarres (2006). Stokes (1977) has suggested the ranked set sample mean as an estimator for the mean of the study variate Y, when an auxiliary variable X is used for ranking the sample units, under the assumption that (X, Y) follows a bivariate normal distribution. Barnett and Moore (1997) have improved the estimator of Stokes (1977) by deriving the BLUE of the mean of the study variate Y, based on ranked set sample obtained on the study variate Y . Al-Saleh and Al-Kadiri (2000) have extended first the usual concept of RSS to double stage ranked set sampling (DSRSS) with an objective of increasing the precision of certain estimators of the population when compared with those obtained based on usual RSS or using random sampling. Al-Saleh and Al-Omari (2002) have further extended DSRSS to multistage ranked set sampling (MSRSS) and shown that there is increase in the precision of estimators obtained based on MSRSS when compared with those based on usual RSS and DSRSS. Al-Saleh (2004) has considered the steady-state RSS.

The remaining plan of the paper is given as follows: In section 2 we have discussed a brief discussion on Chacko and Thomas (2008) estimators in MTBED using RSS. Section 3 dealt with some minimum mean squared error (MMSE) estimators on the lines of Searls (1964), Singh et al. (1973) and Searls and Intarapanich (1990) along with their properties. In section 4 we have proposed some shrinkage estimators of the parameter θ_2 in MTBED on the lines of Jani (1991) and Kourouklis (1994). We have also obtained their biases and mean squared errors (MSEs) and shown theoretically that the shrinkage estimators are superior estimate of θ_2 as compared to Chacko and Thomas (2008) estimators and MMSE estimators. In section 5 we have computed the relative efficiencies of different estimators numerically to evaluate their performance. Section 6 concludes the paper with final comments.

2. Chacko and Thomas (2008) estimators based on ranked set sampling (RSS) in Morgenstern type bivariate exponential distribution (MTBED)

Let (X,Y) be a bivariate random variable which follows a MTBED with pdf defined by (1.1). Let $X_{r(r)}$ be the observation measured on the auxiliary variate X in the rth unit of the RSS and let $Y_{r[r]}$ be the measurement made on the Y variate of the same unit,$r=1,2,...,n$. Then clearly $Y_{r[r]}$ is distributed as the concomitant of rth order statistic of a random sample of size n arising from (1.1). By using the expressions for means and variances of concomitants of order statistics arising from MTBED obtained by Scaria and Nair (1999), the mean and variance of $Y_{r[r]}$ for $-1 \le \alpha \le 1$ are given as

$$E[Y_{r[r]}] = \theta_2\left[1-\frac{\alpha}{2}\left(\frac{n-2r+1}{n+1}\right)\right] = \theta_2\xi_r(say). \tag{2.1}$$

$$Var[Y_{r[r]}] = \theta_2^2\left[1-\frac{\alpha}{2}\left(\frac{n-2r+1}{n+1}\right)-\frac{\alpha^2}{4}\left(\frac{n-2r+1}{n+1}\right)^2\right] = \theta_2^2\delta_r(say). \tag{2.2}$$

Chacko and Thomas (2008) shows ranked set sample mean as

$$t_1 = \theta_2^* = \frac{1}{n}\sum_{r=1}^{n} Y_{r[r]}, \tag{2.3}$$

is an unbiased estimator of θ_2 and its variance is given by

$$Var(t_1) = \frac{\theta_2^2}{n}\left[1-\frac{\alpha^2}{4n}\sum_{r=1}^{n}(\frac{n-2r+1}{n+1})^2\right] = \theta_2^2 V_1, \tag{2.4}$$

where $V_1 = \frac{1}{n}\left[1-\frac{\alpha^2}{4n}\sum_{r=1}^{n}(\frac{n-2r+1}{n+1})^2\right]$.

Chacko and Thomas (2008) further provided a better estimator of θ_2 than that of θ_2^* by deriving the BLUE $\hat{\theta}_2$ of θ_2 provided the parameter α is known as

$$t_2 = \hat{\theta}_2 = \frac{\sum_{r=1}^{n}(\frac{\xi_r}{\delta_r})Y_{r[r]}}{\sum_{r=1}^{n}(\frac{\xi_r^2}{\delta_r})}, \tag{2.5}$$

and

$$Var(t_2) = \frac{\theta_2^2}{\sum_{r=1}^{n}(\frac{\xi_r^2}{\delta_r})} = \theta_2^2 V_2, \tag{2.6}$$

where $V_2 = \frac{1}{\sum_{r=1}^{n}(\frac{\xi_r^2}{\delta_r})}$.
Chacko and Thomas (2008) further obtained BLUE based on single stage unbalanced RSS as

$$t_3 = \hat{\theta}_2^{n(1)} = \frac{1}{n\xi_n}\sum_{i=1}^{n} Y_{[n]i}, \tag{2.7}$$

and

$$Var(t_3) = \frac{\theta_2^2\delta_n}{n[1+\frac{\alpha}{2}](\xi_n)^2} = \theta_2^2 V_3, \tag{2.8}$$

where $V_3 = \frac{\delta_n}{n(\xi_n)^2}$.
Chacko and Thomas (2008) also shows BLUE based on single stage unbalanced steady-state RSS as

$$t_4 = \hat{\theta}_2^{n(\infty)} = \frac{1}{n[1+\frac{\alpha}{2}]}\sum_{i=1}^{n} Y_{n[i]}^{\infty}, \tag{2.9}$$

and

$$Var(t_4) = \frac{\theta_2^2[1+\frac{\alpha}{2}-\frac{\alpha^2}{4}]}{n[1+\frac{\alpha}{2}]^2} = \theta_2^2 V_4, \tag{2.10}$$

where $V_4 = \frac{\left[1+\frac{\alpha}{2}-\frac{\alpha^2}{4}\right]}{n\left[1+\frac{\alpha}{2}\right]^2}$.

3. Minimum mean squared error (MMSE) estimators of the parameter θ_2

The MMSE estimator of the parameter θ_2 based on $t_i's, i = 1, 2, 3, 4$ are

$$T_{im} = \frac{t_i}{(1+V_i)}, \tag{3.1}$$

in the class of estimators $T_i = A_i t_i$, where $A_i's$ are suitably chosen constants such that the MSE of $T_i's$ are minimum.
The biases and MSEs of $T_{im}'s$ are respectively given by

$$B(T_{im}) = -\theta_2\left(\frac{V_i}{(1+V_i)}\right), \tag{3.2}$$

$$MSE(T_{im}) = \theta_2^2\left(\frac{V_i}{(1+V_i)}\right). \tag{3.3}$$

From (2.4), (2.6), (2.8), (2.10) and (3.3) we have that

$$Var(t_i) - MSE(T_{im}) = \frac{\theta_2^2 V_i^2}{(1+V_i)} > 0, i = 1, 2, 3, 4, . \tag{3.4}$$

which shows that $T_{im}'s, i = 1, 2, 3, 4$ are always superior to the Chacko and Thomas (2008) corresponding estimators $t_i's, i = 1, 2, 3, 4$.

4. Suggested class of estimators using θ_{20} as a prior information about θ_2

Inserting $t_i's, i = 1, 2, 3, 4$ in place of sample mean $\bar{X}$ based on simple random sampling (SRS) in Jani's (1991) class of estimators, we define a class of shrinkage estimators of the parameter θ_2

$$T_{i(p)} = \theta_{20} + k_{(p)}(t_i - \theta_{20}), i = 1, 2, 3, 4, \tag{4.1}$$

which is based on ranked set sampling in MTBED, where $k_{(p)} = \frac{\Gamma(n-p)}{n^p \Gamma(n-2p)}, p$ being a non-zero real number.
The biases and MSEs of $T_{i(p)}{}'s$ are respectively given by

$$B(T_{i(p)}) = \theta_2 \phi(1 - k_{(p)}) \tag{4.2}$$

$$MSE(T_{i(p)}) = \theta_2^2[k_{(p)}^2(\phi^2 + V_i) - 2\phi^2 k_{(p)} + \phi^2], \tag{4.3}$$

where $\phi = (\frac{\theta_{20}}{\theta_2} - 1) = (\lambda - 1)$ with $\lambda = (\frac{\theta_{20}}{\theta_2})$.
We now state the following theorems.
Theorem 1 The proposed estimator $T_{i(p)}{}'s, i = 1, 2, 3, 4$ are better than the corresponding unbiased estimators $t_i{}'s, i = 1, 2, 3, 4$ if

$$k_{(p)} < 1, \frac{\theta_{20}}{1 + \sqrt{\frac{(1+k_p)V_i}{(1-k_p)}}} < \theta_2 < \frac{\theta_{20}}{1 - \sqrt{\frac{(1+k_p)V_i}{(1-k_p)}}}. \tag{4.4}$$

Proof
From (2.4),(2.6),(2.8),(2.10) and (4.3) we have that
$MSE(T_{i(p)}) - Var(t_i) = \theta_2^2(1 - k_{(p)})[\phi^2(1 - k_{(p)}) - V_i(1 + k_{(p)})] < 0,$
if
$1 - k_{(p)} > 0, \phi^2 < \frac{(1+k_p)V_i}{(1-k_p)},$
or $k_{(p)} < 1, -\sqrt{\frac{(1+k_p)V_i}{(1-k_p)}} < \phi < \sqrt{\frac{(1+k_p)V_i}{(1-k_p)}},$
or

$$k_{(p)} < 1, \left(1 - \sqrt{\frac{(1 + k_p)V_i}{(1 - k_p)}}\right) < \lambda = \left(\frac{\theta_{20}}{\theta_2}\right) < (1 + \sqrt{\frac{(1 + k_p)V_i}{(1 - k_p)}}\Bigg), \tag{4.5}$$

or $k_{(p)} < 1, \theta_2\left(1 - \sqrt{\frac{(1+k_p)V_i}{(1-k_p)}}\right) < \theta_{20} < \theta_2\left(1 + \sqrt{\frac{(1+k_p)V_i}{(1-k_p)}}\right),$
or $k_{(p)} < 1, \frac{\theta_{20}}{1+\sqrt{\frac{(1+k_p)V_i}{(1-k_p)}}} < \theta_2 < \frac{\theta_{20}}{1-\sqrt{\frac{(1+k_p)V_i}{(1-k_p)}}}.$
Theorem 2 The proposed estimator $T_{i(p)}{}'s, i = 1, 2, 3, 4$ are better than the corresponding MMSE estimators $T_{im}{}'s, i = 1, 2, 3, 4$ if

$$\frac{\theta_{20}}{1 + \sqrt{\frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2}}} < \theta_2 < \frac{\theta_{20}}{1 - \sqrt{\frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2}}}. \tag{4.6}$$

Proof
From (3.3) and (4.3) we have that
$MSE(T_{im}) - MSE(T_{i(p)}) = \theta_2^2(1 + V_i)^{-1}[V_i\big(1 - k_{(p)}^2(1 + V_i)\big)$
$- \phi^2(1 + V_i)(1 - k_{(p)})^2] > 0,$
if
$\phi^2 < \frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2},$
or $-\sqrt{\frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2}} < \phi < \sqrt{\frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2}},$

or

$$(4.7)\quad \left(1-\sqrt{\frac{V_i\big(1-k_{(p)}^2(1+V_i)\big)}{(1+V_i)(1-k_p)^2}}\right)<\lambda<\left(1+\sqrt{\frac{V_i\big(1-k_{(p)}^2(1+V_i)\big)}{(1+V_i)(1-k_p)^2}}\right),$$

or $\theta_2\left(1-\sqrt{\frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2}}\right)<\theta_{20}<\theta_2\left(1+\sqrt{\frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2}}\right),$

or $k_{(p)}<1,\ \frac{\theta_{20}}{\left(1+\sqrt{\frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2}}\right)}<\theta_2<\frac{\theta_{20}}{\left(1-\sqrt{\frac{V_i\left(1-k_{(p)}^2(1+V_i)\right)}{(1+V_i)(1-k_p)^2}}\right)}.$

It can be easily seen that the proposed shrinkage estimators $T_{i(p)}{}'s, i=1,2,3,4$ are better than the corresponding usual estimators $t_i{}'s, i=1,2,3,4$ and corresponding MMSE estimators $T_{im}{}'s, i=1,2,3,4$ for a wider range of θ_2 . The member of the class of estimators $T_{i(p)}{}'s, i=1,2,3,4$ have smaller MSE than $t_i{}'s, i=1,2,3,4$ for all (n,α) and for θ_2 in the neighborhood of θ_{20} . Largest range of dominance of λ is obtained when $p=-1$ with the resulting estimators $T_{i(-1)}=\theta_{20}+k_{(-1)}(t_i-\theta_{20})$ (see Table 3). Thus $T_{i(-1)}=\theta_{20}+k_{(-1)}(t_i-\theta_{20})$ are better than $t_i{}'s$ no matter how much θ_{20} underestimates θ_2 . Roughly speaking, $p's$ with small absolute values give wider neighborhoods of dominance of $T_{i(p)}{}'s$ over $t_i{}'s$ (see Tables 5-6).

Remark: If we have a situation with α unknown, we introduce an estimator (moment type) for α as follows. For MTBED the correlation coefficient between the two variables is given by $\rho=\frac{\alpha}{4}$. If r is the sample correlation coefficient between $X_{i(i)}$ and $Y_{i[i]}, i=1,2,...,n$ then the moment type estimator for α is obtained by equating with the population correlation coefficient ρ and is obtained as [see, Chacko and Thomas (2008)]:

$$\hat{\alpha}=\begin{cases}-1 & \text{if } r<(-1/4)\\ 4r & \text{if } (-1/4)\le r\le(1/4).\\ 1 & \text{if } r>(1/4)\end{cases}$$

5. Relative efficiency

As we have seen on computer screen that the MMSE estimator T_{4m} has the smallest MSE among the estimators $T_{im}{}'s, i=1,2,3,4$, therefore we have made the comparison of the proposed shrinkage estimators with that of T_{4m}. For this purpose we have computed the relative efficiencies of various suggested shrinkage estimators to the MMSE estimator T_{4m} by using following formulae:

$e_1=RE(T_{1(p)},T_{4m})=\frac{V_4}{(1+V_4)[k_{(p)}^2(\phi^2+V_1)-2\phi^2k_{(p)}+\phi^2]};$

$e_2=RE(T_{2(p)},T_{4m})=\frac{V_4}{(1+V_4)[k_{(p)}^2(\phi^2+V_2)-2\phi^2k_{(p)}+\phi^2]};$

$e_3=RE(T_{3(p)},T_{4m})=\frac{V_4}{(1+V_4)[k_{(p)}^2(\phi^2+V_3)-2\phi^2k_{(p)}+\phi^2]};$

$e_4=RE(T_{4(p)},T_{4m})=\frac{V_4}{(1+V_4)[k_{(p)}^2(\phi^2+V_4)-2\phi^2k_{(p)}+\phi^2]}.$

The values of $e_i{}'s, i=1,2,3,4$ for $n=5(5)20, p=\pm1,\pm2, \alpha=0.25(0.25)1.00$ and different values of λ are shown in Table 1.

It is observed from Table 1 that for fixed $(n,\alpha,|p|)$, the values of $e_i{}'s, i=1,2,3,4$ increase as λ increases up to 1, while it decreases if λ goes beyond 1. When the value of λ is 'unity' (i.e. the guessed value θ_{20} coincides with the true value θ_2), the higher gain in efficiency is seen which is expected too. Also higher gain in efficiency is obtained when sample size n is small. In general the higher gain in efficiency are observed by using $T_{4(p)}$ over T_{4m} for all values of $(n,\alpha,|p|,\lambda)$. It follows that $T_{4(p)}$ is the best estimator among the estimators $T_{im}{}'s, i = 1, 2, 3, 4$.

Tables 2-3 depicts the ranges of λ in which the suggested shrinkage estimators $T_{i(p)}{}'s, i = 1, 2, 3, 4$ are better than the corresponding usual unbiased estimators $t_i{}'s, i = 1, 2, 3, 4$ and the corresponding MMSE estimators $T_{im}{}'s, i = 1, 2, 3, 4$.

Tables 2-3 show that the proposed shrinkage estimators $T_{i(p)}{}'s, i = 1, 2, 3, 4$ are better than the corresponding usual unbiased estimators $t_i{}'s, i = 1, 2, 3, 4$ and the corresponding MMSE estimators $T_{im}{}'s, i = 1, 2, 3, 4$ for considerable ranges of λ .

It is further observed from Tables 2-3 that, although the class of estimators $T_{i(-1)}{}'s, i = 1, 2, 3, 4$; has the largest range of dominance, it offers smallest improvement compared with other competitors. The estimator $T_{i(2)}{}'s, i = 1, 2, 3, 4$ and $T_{i(-2)}{}'s, i = 1, 2, 3, 4$; offer large saving in MSE over , MMSE estimator T_{4m} but in a rather small range of λ . Thus it is interesting to mention that there is enough scope of selecting the suggested value θ_{20} of θ_2 to obtain better estimators which are useful in practice.

6. Conclusion

In this paper we have suggested some MMSE estimators and improved shrinkage estimators based on Chacko and Thomas (2008) estimators of the scale parameter θ_2 involved in (1.1) using ranked set sampling. We have obtained the expressions for biases and mean squared errors of the proposed estimators. It has been shown that the suggested estimators based on prior or guessed value θ_{20} are more efficient than those estimators including Chacko and Thomas (2008) estimators which do not utilize the guessed value θ_{20} , for a considerable range of the scale parameter θ_2 . Thus our recommendation is to use the suggested estimators in practice.

Acknowledgement

The authors are highly grateful to the editor and referees for their constructive comments/ suggestions that helped in the improvement of the revised version of the paper.

References

[1] Al-Saleh, M. F. *Steady-state ranked set sampling and parametric inference.* Journal of Statistical Planning and Inference, 123, 83-95, (2004).

[2] Al-Saleh, M. F. and Al-Kadiri, M. (2000).Double ranked set sampling. Statistics and Probability Letters, 48, 205-212.

[3] Al-Saleh, M. F. and Al-Omari, A. (2002).Multistage ranked set sampling. Journal of Statistical Planning and Inference, 102, 273-286.

[4] Barnett, V. and Moore, K. (1997).Best linear unbiased estimates in ranked-set sampling with particular reference to imperfect ordering. Journal of Applied Statistics, 24, 697-710.

[5] Chacko, M. and Thomas, P. Y. (2008). Estimation of parameter of Morgenstern type bivariate exponential distribution by ranked set sampling. Annals of the Institute of Statistical Mathematics, 60,301-318.

[6] Chen, Z. (2000).The efficiency of ranked-set sampling relative to simple random sampling under multi-parameter families. Statistica Sinica, 10, 247-263.

[7] Chen, Z. and Bai, Z. (2000). The optimal ranked set sampling scheme for parametric families. Sankhya Series A, 46, 178-192.

[8] Chen, Z., Bai, Z. and Sinha, B. K. (2004). Lecture notes in statistics, ranked set sampling, theory and applications. New York: Springer.

[9] Jani, P. N. (1991). A class of shrinkage estimators for the scale parameter of the exponential distribution. IEEE Transactions on Reliability, 40, 68-70.

[10] Kourouklis, S. (1994). Estimation in the two-parameter exponential distribution with prior information. IEEE Transactions of Reliability, 43, 3446-450.

[11] Lam, K., Sinha, B. K. and Wu, Z. (1994).Estimation of a two-parameter exponential distribution using ranked set sample. Annals of the Institute of Statistical Mathematics, 46, 723-736.

[12] Lam, K., Sinha, B. K. and Wu, Z. (1995). Estimation of location and scale parameters of a logistic distribution using ranked set sample. In: H. N. Nagaraja, P. K. Sen D. F. Morrison (Eds) Statistical theory and applications: papers in honor of Herbert A. David. New-York: Springer.

[13] McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. Australian Journal of Agricultural Research, 3, 385-390.

[14] Modarres, R. and Zheng, G. (2004). Maximum likelihood estimation of dependence parameter using ranked set sampling. Statistics and Probability Letters, 68, 315-323.

[15] Scaria, J. and Nair, N. U. (1999). On concomitants of order statistics from Morgenstern family. Biometrical Journal, 41, 483-489.

[16] Searls, D. T. and Intarapanich, P. (1990). A note on the estimator for the variance that utilizes the kurtosis. The American Statistician, 44, 295-296.

[17] Searls, D.T. (1964). The utilization of a know coefficient of variation in the estimation procedure. Journal of the American Statistical Association, 59, 1225-1226.

[18] Singh, J., Pandey, B. N. and Hirano, K. (1973). On the utilization of known coefficient of kurtosis in the estimation procedure of variance. Annals of the Institute of Statistical Mathematics, 25, 51-55.

[19] Stokes, S. L. (1977). Ranked set sampling with concomitant variables. Communications in Statistics- Theory and Methods, 6, 1207-1211.

[20] Stokes, S. L. (1980). Inference on the correlation coefficient in bivariate normal populations from ranked set samples. Journal of the American Statistical Association, 75, 989-995.

[21] Stokes, S. L. (1995). Parametric ranked set sampling. Annals of the Institute of Statistical Mathematics, 47, 465-482.

[22] Zheng, G. and Modarres, R. (2006). A robust estimate of correlation coefficient for bivariate normal distribution using ranked set sampling. Journal of Statistical planning and Inference, 136, 298-309.

Table 1 : The values of $e_i^{s\prime}, i = 1, 2, 3, 4$
for $p = -2$:

n	α	$\lambda = 1.00$				$\lambda = 1.20$ and $\lambda = 0.80$				$\lambda = 1.40$ and $\lambda = 0.60$			
		e_1	e_2	e_3	e_4	e_1	e_2	e_3	e_4	e_1	e_2	e_3	e_4
5	0.25	3.7552	3.7553	4.0281	4.2692	2.8696	2.8697	3.0263	3.1604	1.6806	1.6807	1.7332	1.7764
	0.50	3.3568	3.3572	3.7721	4.3556	2.5589	2.5591	2.7933	3.1009	1.4937	1.4937	1.5706	1.6634
	0.75	2.9911	2.9934	3.4495	4.4381	2.2704	2.2717	2.5251	3.0170	1.3178	1.3182	1.3997	1.5388
	1.00	2.6564	2.6647	3.0878	4.5158	2.0039	2.0086	2.2400	2.9068	1.1537	1.1553	1.2283	1.4050
10	0.25	1.9696	1.9691	2.1415	2.2375	1.7493	1.7489	1.8835	1.9574	1.3097	1.3095	1.3835	1.4229
	0.50	1.7487	1.7486	2.0010	2.2617	1.5508	1.5507	1.7460	1.9413	1.1577	1.1577	1.2632	1.3623
	0.75	1.5510	1.5536	1.8181	2.2844	1.3720	1.3741	1.5770	1.9163	1.0192	1.0203	1.1281	1.2917
	1.00	1.3746	1.3832	1.6132	2.3055	1.2115	1.2182	1.3931	1.8808	0.8935	0.8971	0.9885	1.2114
15	0.25	1.5388	1.5379	1.6812	1.7475	1.4273	1.4265	1.5490	1.6051	1.1724	1.1719	1.2533	1.2898
	0.50	1.3628	1.3638	1.5690	1.7604	1.2627	1.2637	1.4379	1.5970	1.0349	1.0355	1.1496	1.2492
	0.75	1.2067	1.2092	1.4217	1.7725	1.1162	1.1183	1.2977	1.5838	0.9112	0.9126	1.0286	1.2005
	1.00	1.0688	1.0776	1.2570	1.7835	0.9861	0.9936	1.1441	1.5646	0.8003	0.8052	0.9013	1.1434
20	0.25	1.3501	1.3492	1.4788	1.5330	1.2780	1.2771	1.3927	1.4407	1.1014	1.1008	1.1856	1.2202
	0.50	1.1942	1.1956	1.3791	1.5416	1.1295	1.1307	1.2936	1.4355	0.9716	0.9725	1.0906	1.1898
	0.75	1.0566	1.0583	1.2478	1.5496	0.9980	0.9996	1.1669	1.4269	0.8557	0.8569	0.9770	1.1528
	1.00	0.9355	0.9443	1.1011	1.5570	0.8819	0.8897	1.0276	1.4140	0.7526	0.7583	0.8562	1.1086

Table 1 : Continued...

for $p = -1$

n	α	$\lambda = 1.00$				$\lambda = 1.20$ and $\lambda = 0.80$				$\lambda = 1.40$ and $\lambda = 0.60$			
		e_1	e_2	e_3	e_4	e_1	e_2	e_3	e_4	e_1	e_2	e_3	e_4
5	0.25	1.0777	1.0777	1.1560	1.2252	1.0691	1.0691	1.1462	1.2141	1.0442	1.0442	1.1175	1.1821
	0.50	0.9634	0.9635	1.0826	1.2500	0.9556	0.9557	1.0728	1.2370	0.9331	0.9332	1.0445	1.1995
	0.75	0.8584	0.8591	0.9900	1.2737	0.8514	0.8520	0.9806	1.2583	0.8310	0.8316	0.9536	1.2142
	1.00	0.7624	0.7647	0.8862	1.2960	0.7559	0.7583	0.8775	1.2776	0.7374	0.7396	0.8526	1.2254
10	0.25	0.9793	0.9791	1.0648	1.1125	0.9754	0.9751	1.0601	1.1074	0.9638	0.9636	1.0465	1.0925
	0.50	0.8695	0.8694	0.9949	1.1245	0.8659	0.8659	0.9903	1.1186	0.8555	0.8555	0.9767	1.1013
	0.75	0.7712	0.7725	0.9040	1.1358	0.7680	0.7693	0.8996	1.1289	0.7585	0.7598	0.8867	1.1087
	1.00	0.6834	0.6877	0.8021	1.1463	0.6805	0.6848	0.7981	1.1381	0.6719	0.6761	0.7863	1.1142
15	0.25	0.9466	0.9460	1.0342	1.0750	0.9440	0.9435	1.0311	1.0717	0.9365	0.9360	1.0222	1.0620
	0.50	0.8383	0.8390	0.9652	1.0829	0.8360	0.8367	0.9621	1.0791	0.8293	0.8299	0.9532	1.0679
	0.75	0.7423	0.7438	0.8745	1.0903	0.7402	0.7418	0.8717	1.0859	0.7341	0.7356	0.8632	1.0728
	1.00	0.6575	0.6629	0.7732	1.0971	0.6556	0.6610	0.7706	1.0919	0.6500	0.6553	0.7629	1.0765
20	0.25	0.9302	0.9295	1.0188	1.0562	0.9283	0.9277	1.0166	1.0538	0.9228	0.9221	1.0099	1.0467
	0.50	0.8227	0.8237	0.9502	1.0621	0.8211	0.8220	0.9479	1.0594	0.8161	0.8170	0.9413	1.0511
	0.75	0.7279	0.7291	0.8597	1.0676	0.7264	0.7276	0.8576	1.0644	0.7219	0.7231	0.8513	1.0547
	1.00	0.6445	0.6506	0.7586	1.0727	0.6432	0.6492	0.7567	1.0689	0.6390	0.6450	0.7510	1.0575

Table 1 : Continued...
for $p = 1$

n	α	$\lambda = 1.00$				$\lambda = 1.20$ and $\lambda = 0.80$				$\lambda = 1.40$ and $\lambda = 0.60$			
		e_1	e_2	e_3	e_4	e_1	e_2	e_3	e_4	e_1	e_2	e_3	e_4
5	0.25	2.0789	2.0789	2.2300	2.3634	1.9086	1.9087	2.0353	2.1458	1.5322	1.5322	1.6128	1.6814
	0.50	1.8584	1.8586	2.0883	2.4113	1.7047	1.7049	1.8962	2.1588	1.3659	1.3660	1.4861	1.6427
	0.75	1.6559	1.6572	1.9097	2.4570	1.5167	1.5178	1.7269	2.1625	1.2113	1.2120	1.3417	1.5907
	1.00	1.4706	1.4752	1.7094	2.5000	1.3441	1.3479	1.5408	2.1552	1.0684	1.0708	1.1891	1.5244
10	0.25	1.2646	1.2643	1.3750	1.4366	1.2336	1.2333	1.3384	1.3967	1.1492	1.1489	1.2396	1.2895
	0.50	1.1228	1.1227	1.2847	1.4521	1.0949	1.0949	1.2484	1.4059	1.0191	1.0190	1.1508	1.2833
	0.75	0.9958	0.9975	1.1673	1.4667	0.9706	0.9722	1.1328	1.4126	0.9020	0.9034	1.0405	1.2719
	1.00	0.8825	0.8881	1.0358	1.4803	0.8595	0.8647	1.0042	1.4165	0.7970	0.8015	0.9199	1.2545
15	0.25	1.1076	1.1070	1.2101	1.2579	1.0920	1.0914	1.1915	1.2378	1.0478	1.0473	1.1391	1.1813
	0.50	0.9809	0.9817	1.1294	1.2672	0.9669	0.9677	1.1109	1.2439	0.9273	0.9280	1.0588	1.1790
	0.75	0.8686	0.8704	1.0233	1.2758	0.8559	0.8577	1.0058	1.2487	0.8200	0.8216	0.9566	1.1737
	1.00	0.7693	0.7757	0.9048	1.2838	0.7577	0.7639	0.8888	1.2518	0.7249	0.7305	0.8440	1.1647
20	0.25	1.0416	1.0409	1.1409	1.1827	1.0314	1.0307	1.1286	1.1696	1.0019	1.0012	1.0933	1.1317
	0.50	0.9213	0.9224	1.0640	1.1894	0.9121	0.9132	1.0518	1.1741	0.8856	0.8866	1.0167	1.1306
	0.75	0.8151	0.8165	0.9626	1.1955	0.8068	0.8081	0.9511	1.1777	0.7828	0.7841	0.9179	1.1273
	1.00	0.7218	0.7285	0.8495	1.2012	0.7141	0.7208	0.8390	1.1802	0.6922	0.6984	0.8088	1.1215

small **Table 1** : Continued...

for $p = 2$

n	α	$\lambda = 1.00$				$\lambda = 1.20$ and $\lambda = 0.80$				$\lambda = 1.40$ and $\lambda = 0.60$			
		e_1	e_2	e_3	e_4	e_1	e_2	e_3	e_4	e_1	e_2	e_3	e_4
5	0.25	116.9370	116.9403	125.4378	132.9438	4.2457	4.2458	4.2562	4.2644	1.0911	1.0911	1.0918	1.0924
	0.50	104.5335	104.5452	117.4651	135.6337	3.7572	3.7572	3.7721	3.7884	0.9653	0.9653	0.9663	0.9674
	0.75	93.1428	93.2150	107.4192	138.2036	3.2909	3.2910	3.3064	3.3292	0.8451	0.8451	0.8461	0.8476
	1.00	82.7206	82.9793	96.1538	140.6250	2.8519	2.8522	2.8657	2.8929	0.7319	0.7319	0.7328	0.7346
10	0.25	4.5882	4.5870	4.9885	5.2121	2.5979	2.5976	2.7216	2.7868	1.1289	1.1288	1.1516	1.1631
	0.50	4.0735	4.0733	4.6612	5.2685	2.2936	2.2935	2.4689	2.6294	0.9925	0.9925	1.0240	1.0506
	0.75	3.6129	3.6191	4.2353	5.3215	2.0148	2.0167	2.1946	2.4542	0.8658	0.8662	0.8974	0.9380
	1.00	3.2020	3.2221	3.7579	5.3706	1.7606	1.7667	1.9166	2.2631	0.7491	0.7502	0.7759	0.8272
15	0.25	2.4172	2.4158	2.6409	2.7451	1.8605	1.8597	1.9903	2.0489	1.1003	1.1001	1.1445	1.1636
	0.50	2.1407	2.1424	2.4646	2.7654	1.6424	1.6434	1.8267	1.9868	0.9671	0.9675	1.0282	1.0771
	0.75	1.8956	1.8995	2.2333	2.7843	1.4464	1.4487	1.6350	1.9121	0.8454	0.8462	0.9065	0.9857
	1.00	1.6789	1.6928	1.9746	2.8017	1.2707	1.2786	1.4331	1.8239	0.7348	0.7374	0.7863	0.8910
20	0.25	1.8246	1.8233	1.9985	2.0718	1.5489	1.5480	1.6724	1.7235	1.0658	1.0653	1.1228	1.1456
	0.50	1.6138	1.6157	1.8638	2.0835	1.3670	1.3684	1.5422	1.6896	0.9371	0.9377	1.0162	1.0782
	0.75	1.4279	1.4303	1.6863	2.0943	1.2050	1.2066	1.3839	1.6473	0.8206	0.8214	0.8999	1.0042
	1.00	1.2643	1.2762	1.4881	2.1042	1.0610	1.0694	1.2143	1.5954	0.7157	0.7195	0.7824	0.9247

Table 2 :The Ranges of λ in which the suggested shrinkage estimators $T'_{i(p)}s, i = 1, 2, 3, 4$ are better than the corresponding usual unbiased estimators $t'_i s, i = 1, 2, 3, 4$ given by Chacko and Thomas (2008):

n	α	$p = -2$				$p = -1$			
		t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4
5	0.25	(0.33,1.67)	(0.33,1.67)	(0.36,1.64)	(0.38,1.62)	(0,2.48)	(0,2.48)	(0,2.43)	(0,2.39)
	0.50	(0.34,1.66)	(0.34,1.66)	(0.38,1.62)	(0.42,1.58)	(0,2.47)	(0,2.47)	(0,2.39)	(0,2.29)
	0.75	(0.34,1.66)	(0.34,1.66)	(0.39,1.61)	(0.46,1.54)	(0,2.46)	(0,2.46)	(0,2.36)	(0,2.20)
	1.00	(0.35,1.65)	(0.35,1.65)	(0.40,1.60)	(0.50,1.50)	(0,2.44)	(0,2.44)	(0,2.34)	(0,2.11)
10	0.25	(0.40,1.60)	(0.40,1.60)	(0.42,1.58)	(0.44,1.56)	(0,2.45)	(0,2.45)	(0,2.39)	(0,2.36)
	0.50	(0.40,1.60)	(0.40,1.60)	(0.44,1.56)	(0.47,1.53)	(0,2.44)	(0,2.44)	(0,2.34)	(0,2.26)
	0.75	(0.41,1.59)	(0.41,1.59)	(0.45,1.55)	(0.51,1.49)	(0,2.42)	(0,2.42)	(0,2.31)	(0,2.17)
	1.00	(0.42,1.58)	(0.42,1.58)	(0.46,1.54)	(0.55,1.45)	(0,2.40)	(0,2.39)	(0,2.29)	(0,2.08)
15	0.25	(0.42,1.58)	(0.42,1.58)	(0.45,1.55)	(0.46,1.54)	(0,2.43)	(0,2.43)	(0,2.37)	(0,2.35)
	0.50	(0.42,1.58)	(0.43,1.57)	(0.46,1.54)	(0.49,1.51)	(0,2.42)	(0,2.42)	(0,2.33)	(0,2.25)
	0.75	(0.43,1.57)	(0.43,1.57)	(0.48,1.52)	(0.53,1.47)	(0,2.41)	(0,2.41)	(0,2.30)	(0,2.16)
	1.00	(0.44,1.56)	(0.44,1.56)	(0.48,1.52)	(0.57,1.43)	(0,2.38)	(0,2.38)	(0,2.28)	(0,2.07)
20	0.25	(0.43,1.57)	(0.43,1.57)	(0.46,1.54)	(0.47,1.53)	(0,2.43)	(0,2.43)	(0,2.36)	(0,2.34)
	0.50	(0.44,1.56)	(0.44,1.56)	(0.48,1.52)	(0.50,1.50)	(0,2.42)	(0,2.42)	(0,2.32)	(0,2.25)
	0.75	(0.44,1.56)	(0.44,1.56)	(0.49,1.51)	(0.54,1.46)	(0,2.40)	(0,2.40)	(0,2.29)	(0,2.16)
	1.00	(0.45,1.55)	(0.46,1.54)	(0.50,1.50)	(0.58,1.42)	(0,2.38)	(0,2.37)	(0,2.27)	(0,2.07)

Table 2 : Continued... :

n	α	$p=1$				$p=2$			
		t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4
5	0.25	(0.11,1.89)	(0.11,1.89)	(0.14,1.86)	(0.16,1.84)	(0.52,1.48)	(0.52,1.48)	(0.53,1.47)	(0.55,1.45)
	0.50	(0.11,1.89)	(0.11,1.89)	(0.16,1.84)	(0.22,1.78)	(0.52,1.48)	(0.52,1.48)	(0.55,1.45)	(0.58,1.42)
	0.75	(0.12,1.88)	(0.12,1.88)	(0.18,1.82)	(0.28,1.72)	(0.52,1.48)	(0.52,1.48)	(0.56,1.44)	(0.61,1.39)
	1.00	(0.13,1.87)	(0.13,1.87)	(0.19,1.81)	(0.33,1.67)	(0.53,1.47)	(0.53,1.47)	(0.56,1.44)	(0.64,1.36)
10	0.25	(0.05,1.95)	(0.05,1.95)	(0.09,1.91)	(0.11,1.89)	(0.51,1.49)	(0.51,1.49)	(0.53,1.47)	(0.54,1.46)
	0.50	(0.06,1.94)	(0.06,1.94)	(0.12,1.88)	(0.17,1.83)	(0.51,1.49)	(0.51,1.49)	(0.54,1.46)	(0.57,1.43)
	0.75	(0.07,1.93)	(0.07,1.93)	(0.14,1.86)	(0.23,1.77)	(0.52,1.49)	(0.52,1.48)	(0.55,1.45)	(0.60,1.40)
	1.00	(0.08,1.92)	(0.09,1.91)	(0.15,1.85)	(0.29,1.71)	(0.52,1.48)	(0.52,1.48)	(0.56,1.44)	(0.63,1.37)
15	0.25	(0.04,1.96)	(0.04,1.96)	(0.08,1.92)	(0.10,1.90)	(0.50,1.50)	(0.50,1.50)	(0.52,1.48)	(0.53,1.47)
	0.50	(0.04,1.96)	(0.04,1.96)	(0.11,1.89)	(0.16,1.84)	(0.50,1.50)	(0.50,1.50)	(0.53,1.47)	(0.56,1.44)
	0.75	(0.05,1.95)	(0.05,1.95)	(0.13,1.87)	(0.22,1.78)	(0.51,1.50)	(0.51,1.49)	(0.54,1.46)	(0.59,1.41)
	1.00	(0.07,1.93)	(0.07,1.93)	(0.14,1.86)	(0.28,1.72)	(0.51,1.49)	(0.51,1.49)	(0.55,1.45)	(0.62,1.38)
20	0.25	(0.03,1.97)	(0.03,1.97)	(0.07,1.93)	(0.09,1.91)	(0.49,1.51)	(0.49,1.51)	(0.51,1.49)	(0.52,1.48)
	0.50	(0.03,1.97)	(0.04,1.96)	(0.10,1.90)	(0.15,1.85)	(0.49,1.51)	(0.49,1.51)	(0.53,1.47)	(0.55,1.45)
	0.75	(0.05,1.95)	(0.05,1.95)	(0.12,1.88)	(0.21,1.79)	(0.50,1.50)	(0.50,1.50)	(0.54,1.46)	(0.59,1.41)
	1.00	(0.06,1.94)	(0.07,1.93)	(0.14,1.86)	(0.27,1.73)	(0.51,1.49)	(0.51,1.49)	(0.55,1.45)	(0.62,1.38)

Table 3 :The Ranges of λ in which the suggested shrinkage estimators $T'_{i(p)}s, i = 1, 2, 3, 4$ are better than the corresponding MMSE estimators $T'_{im}s, i = 1, 2, 3, 4$:

n	α	$p=-2$				$p=-1$			
		T_{1m}	T_{2m}	T_{3m}	T_{4m}	T_{1m}	T_{2m}	T_{3m}	T_{4m}
5	0.25	(0.36,1.64)	(0.36,1.64)	(0.35,1.65)	(0.35,1.65)	(0,2)	(0,2)	(0,2.03)	(0,2.06)
	0.50	(0.36,1.64)	(0.36,1.64)	(0.35,1.65)	(0.34,1.66)	(0,2)	(0,2)	(0,2.05)	(0,2.11)
	0.75	(0.36,1.64)	(0.36,1.64)	(0.36,1.64)	(0.34,1.66)	(0,2)	(0,2)	(0,2.06)	(0,2.15)
	1.00	(0.37,1.63)	(0.37,1.63)	(0.36,1.64)	(0.34,1.66)	(0,2)	(0,2)	(0,2.06)	(0,2.18)
10	0.25	(0.38,1.62)	(0.38,1.62)	(0.38,1.62)	(0.37,1.63)	(0,2)	(0,2)	(0,2.04)	(0,2.06)
	0.50	(0.38,1.62)	(0.38,1.62)	(0.38,1.62)	(0.37,1.63)	(0,2)	(0,2)	(0,2.06)	(0,2.11)
	0.75	(0.39,1.61)	(0.39,1.61)	(0.38,1.62)	(0.37,1.63)	(0,2)	(0,2)	(0,2.07)	(0,2.14)
	1.00	(0.40,1.60)	(0.40,1.60)	(0.39,1.61)	(0.38,1.62)	(0,2)	(0,2)	(0,2.06)	(0,2.17)
15	0.25	(0.39,1.61)	(0.39,1.61)	(0.38,1.62)	(0.38,1.62)	(0,2)	(0,2)	(0,2.04)	(0,2.06)
	0.50	(0.39,1.61)	(0.39,1.61)	(0.38,1.62)	(0.38,1.62)	(0,2)	(0,2)	(0,2.06)	(0,2.10)
	0.75	(0.40,1.60)	(0.40,1.60)	(0.39,1.61)	(0.38,1.62)	(0,2)	(0,2)	(0,2.07)	(0,2.14)
	1.00	(0.41,1.59)	(0.40,1.60)	(0.40,1.60)	(0.39,1.61)	(0,2)	(0,2)	(0,2.06)	(0,2.16)
20	0.25	(0.39,1.61)	(0.39,1.61)	(0.39,1.61)	(0.39,1.61)	(0,2)	(0,2)	(0,2.04)	(0,2.06)
	0.50	(0.39,1.61)	(0.39,1.61)	(0.39,1.61)	(0.38,1.62)	(0,2)	(0,2)	(0,2.06)	(0,2.10)
	0.75	(0.40,1.60)	(0.40,1.60)	(0.39,1.61)	(0.39,1.61)	(0,2)	(0,2)	(0,2.07)	(0,2.14)
	1.00	(0.41,1.59)	(0.41,1.59)	(0.40,1.60)	(0.39,1.61)	(0,2)	(0,2)	(0,2.06)	(0,2.16)

Table 3 : Continued... :

n	α	$p=1$				$p=2$			
		T_{1m}	T_{2m}	T_{3m}	T_{4m}	T_{1m}	T_{2m}	T_{3m}	T_{4m}
5	0.25	(0.23,1.77)	(0.23,1.77)	(0.22,1.78)	(0.22,1.78)	(0.56,1.44)	(0.56,1.44)	(0.56,1.44)	(0.55,1.45)
	0.50	(0.23,1.77)	(0.23,1.77)	(0.22,1.78)	(0.21,1.79)	(0.56,1.44)	(0.56,1.44)	(0.56,1.44)	(0.55,1.45)
	0.75	(0.24,1.76)	(0.24,1.76)	(0.22,1.78)	(0.20,1.80)	(0.56,1.44)	(0.56,1.44)	(0.56,1.44)	(0.55,1.45)
	1.00	(0.25,1.75)	(0.25,1.75)	(0.23,1.77)	(0.20,1.80)	(0.57,1.43)	(0.57,1.43)	(0.56,1.44)	(0.55,1.45)
10	0.25	(0.18,1.82)	(0.18,1.82)	(0.17,1.83)	(0.17,1.83)	(0.53,1.47)	(0.53,1.47)	(0.53,1.47)	(0.53,1.47)
	0.50	(0.18,1.82)	(0.18,1.82)	(0.17,1.83)	(0.16,1.84)	(0.54,1.46)	(0.54,1.46)	(0.53,1.47)	(0.53,1.47)
	0.75	(0.19,1.81)	(0.19,1.81)	(0.17,1.83)	(0.15,1.85)	(0.54,1.46)	(0.54,1.46)	(0.54,1.46)	(0.53,1.47)
	1.00	(0.20,1.80)	(0.20,1.80)	(0.18,1.82)	(0.15,1.85)	(0.55,1.45)	(0.55,1.45)	(0.54,1.46)	(0.54,1.46)
15	0.25	(0.17,1.83)	(0.17,1.83)	(0.15,1.85)	(0.15,1.85)	(0.52,1.48)	(0.52,1.48)	(0.52,1.48)	(0.52,1.48)
	0.50	(0.17,1.83)	(0.17,1.83)	(0.15,1.85)	(0.14,1.86)	(0.52,1.48)	(0.52,1.48)	(0.52,1.48)	(0.52,1.48)
	0.75	(0.18,1.82)	(0.18,1.82)	(0.16,1.84)	(0.14,1.86)	(0.53,1.47)	(0.53,1.47)	(0.52,1.48)	(0.52,1.48)
	1.00	(0.19,1.81)	(0.18,1.82)	(0.17,1.83)	(0.14,1.86)	(0.54,1.46)	(0.53,1.47)	(0.53,1.47)	(0.53,1.47)
20	0.25	(0.16,1.84)	(0.16,1.84)	(0.15,1.85)	(0.14,1.86)	(0.51,1.49)	(0.51,1.49)	(0.51,1.49)	(0.51,1.49)
	0.50	(0.16,1.84)	(0.16,1.84)	(0.14,1.86)	(0.13,1.87)	(0.52,1.48)	(0.52,1.48)	(0.51,1.49)	(0.51,1.49)
	0.75	(0.17,1.83)	(0.17,1.83)	(0.15,1.85)	(0.13,1.87)	(0.52,1.48)	(0.52,1.48)	(0.52,1.48)	(0.51,1.49)
	1.00	(0.18,1.82)	(0.18,1.82)	(0.16,1.84)	(0.13,1.87)	(0.53,1.47)	(0.53,1.47)	(0.52,1.48)	(0.52,1.48)

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 593 – 608

An Improved Bar - Lev, Bobovitch and Boukai randomized response model using moments ratios of scrambling variable

Housila P. Singh* and Tanveer A. Tarray†

Abstract

In this paper, we have suggested a new randomized response model and its properties have been studied. The proposed model is found to be more efficient than the randomized response models studied by Bar – Lev et al. (2004) and Eichhorn and Hayre (1983). The relative efficiency of the proposed model has been studied with respect to the Bar – Lev et al.'s (2004) and Eichhorn and Hayre's (1983) models. Numerical illustrations are also given to support the present study.

Keywords: Randomized response sampling, Estimation of mean, Respondents protection, Sensitive quantitative variable.

2000 AMS Classification: 62D05

Received : 16.06.2014 *Accepted* : 16.02.2015 *Doi* : 10.15672/HJMS.20158912971

1. Introduction

Warner (1965) introduced a randomized response (RR) model to estimate a population proportion for sensitive attribute such as homosexuality, drug addiction or induced abortion. Greenberg et al. (1971) further made an extension of RR technique for quantitative variables. The RR technique has spawned a vast literature which has been reviewed by Fox and Tracy (1986), Chaudhuri and Mukerjee (1988) and scheers (1992). Some more developments are: Kerkvliet (1994), Gupta and Thornton (2002), Singh and Mathur (2005), Bar – Lev et al. (2005), Odumade and Singh (2009), Chaudhuri and Christofides (2013), Singh and Tarray (2013, 2014, 2015), Hussain et al (2015), Tarray and Singh

*School of Studies in Statistics, Vikram University Ujjain - M.P. - India-456010,
†Department of Computer Science and Engineering, Islamic University of Science and Technology – Awantipora – Pulwama – Kashmir – India – 192122,
Email: tanveerstat@gmail.com Corresponding Author.

(2015) and Tarray et al. (2015) etc. Eichhorn and Hayre (1983) suggested a multiplicative model to collect information on sensitive quantitative variables like income, tax evasion, amount of drug used etc. For more examples, the reader is referred to Ahsanullah and Eichhorn (1988). According to Eichhorn and Hayre (1983), each respondent in the sample is requested to report the scrambled response $Z_i = SY_i$, where Y_i is the real value of the sensitive quantitative variable, and S is the scrambling variable whose distribution is assumed to be known. In other words $E(S) = \theta$ and$V(S) = \gamma^2$ are assumed to be known and positive, where E and V denote the expected value and variance over the randomization device. Then an estimator of the population mean μ_y under the simple random sampling with replacement (SRSWR) due to Eichhorn and Hayre (1983) is given by:

$$\text{(1.1)} \qquad \hat{\mu}_{Y(EH)} = \frac{1}{n}\sum_{i=1}^{n}\frac{Z_i}{\theta}$$

with variance

$$\text{(1.2)} \qquad V(\hat{\mu}_{Y(EH)}) = \frac{\mu_Y^2}{n}[C_y^2 + C_\gamma^2(1 + C_y^2)],$$

where $C_\gamma^2 = \dfrac{\gamma^2}{\theta^2}$ and $C_y = \dfrac{\sigma_y}{\mu_Y}$.We shall now discuss a randomized response model studied by Bar – Lev et al. (2004), say the BBB model. The distribution of the responses is given by:

$$\text{(1.3)} \qquad Z_i = \left| \begin{array}{c} Y_iS \text{ with probability } (1-P) \\ Y_i \text{ with probability } P. \end{array} \right.$$

In other words, each respondent is requested to rotate a spinner unobserved by the interviewer, and if the spinner stops in the shaded area, then he/she is requested to report the real response on the sensitive variable, say Y_i; and if the spinner stops in the non shaded area, then the respondent is required to report the scrambled response, say Y_iS, where S is the scrambled variable. Let P be the radial non shaded area of the spinner as shown in Figure 1.
An unbiased estimator of the population mean Y is given by:

$$\text{(1.4)} \qquad \hat{\mu}_{Y(BBB)} = \frac{1}{n[(1-P)\theta + P]}\sum_{i=1}^{n} Z_i$$

with variance under SRSWR sampling given by

$$\text{(1.5)} \qquad V[\hat{\mu}_{Y(BBB)}] = \frac{\mu_Y^2}{n}[C_y^2 + (1 + C_y^2)C_P^2],$$

where

$$\text{(1.6)} \qquad C_P^2 = \frac{(1-P)\theta^2(1 + C_\gamma^2) + P}{[(1-P)\theta + P]^2} - 1.$$

When the coefficient of variation C_y of the study variable is known, Searls (1964) was the first to consider the problem of estimating the population mean μ_y in the absence of scrambled responses. Later on, with known coefficient of variation C_Y of the study variable Y various authors including Khan (1967), Govindarajulu and Sahai (1972), Gleser

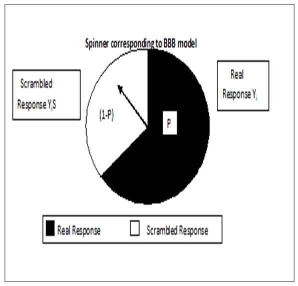


Figure 1. Bar - lev, Bobovitch and Boukai (2004; BBB) randomized response device

and Healy (1976), Sen (1979), Tripathi et al. (1983) and Singh and Katiyar (1988) have considered the problem of estimating the population mean μ_Y of the study variable Y. Sen (1978) was first to use the moments ratios of the study variable Y in estimating the population mean μ_Y . Upadhyaya and Singh (1984) have considered the problem of estimating the population mean μ_Y using moments ratios. Singh and Mathur (2005) and Hussain et al. (2013) have used the coefficient of variation C_Y of the study variable Y at the estimation stage in presence of scrambled responses. Singh and Chen (2009) have used the higher order moments of the scrambling variable at the estimation stage for estimating the proportion of a potentially sensitive attribute in survey sampling.
In this paper we have suggested a new randomized response model and its properties are studied. It has been shown that the resulting (optimum) randomized response model depends on the moments ratios such as C_γ (coefficient of variation), $\beta_{1(S)}$ (coefficient of skewness) and $\beta_{2(S)}$ (coefficient of kurtosis) of the scrambling variable S. We have proved the superiority of the proposed randomized response model over Eichhorn and Hayre (1983) and Bar – Lev et al. (2004) randomized response models both theoretically and empirically.

2. Suggested Randomized Response model

In the proposed randomized response model, we request an individual to rotate a spinner as shown in Figure 2.
In the proposed randomized response model, the distribution of the response is given by

$$Z_i = \left| \begin{array}{c} Y_i[(1-k)S + K\theta(\dfrac{S-\theta}{\gamma})^2] \text{ with probability } (1-P) \\ Y_i \text{ with probability } P \end{array} \right.$$

.

The reported response Z_i can also be expressed as

$$Z_i = \left| \begin{array}{c} Y_i[(1-k)S + K\theta S^{*2}] \text{ with probability } (1-P) \\ Y_i \text{ with probability } P. \end{array} \right. \tag{2.1}$$

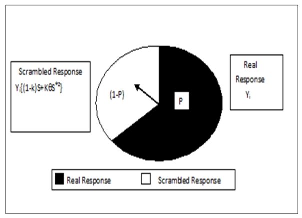


Figure 2. Spinner of the proposed randomized response model

where k is assumed known constant [see Odumade and Singh (2009)] and $S^* = \dfrac{(S-\theta)}{\gamma}$ is the standardized scrambling variable.
In other words, each respondent is requested to rotate a spinner unobserved by the interviewer, and if the spinner stops in the shaded area, then the respondent is requested to report the real response on the sensitive variable, say Y_i; and if the spinner stops in the non shaded area, then the respondent is required to report the scrambled response, say $Y_i[(1-k)S + K\theta S^{*2}]$. Let P be the proportion of the shaded area of the spinner and $(1-P)$ be the non shaded area of the spinner as shown in Figure 2.
For estimating the population mean μ_Y of the real response on the sensitive quantitative variable Y, a simple random and with replacement sample (SRSWR) of n respondents is selected from the population. Then , we have the following theorem.
Theorem 2.1 An unbiased estimator of the population mean μ_Y is given by

$$\hat{\mu}_Y = \frac{\bar{Z}}{[(1-P)\theta + P]} \tag{2.2}$$

Proof- We have from (2.1),
$E(Z_i) = \mu_Y[(1-P)\theta + P]$
Hence, the proposed estimator for μ_Y , based on a random sample of the randomized response; $Z_1, Z_2, ..., Z_n$ is $\hat{\mu}_{Y(HT)} = \dfrac{\bar{Z}}{[(1-P)\theta + P]}$ is unbiased estimator of the population mean μ_Y. Thus the theorem is proved.
The variance of the proposed estimator $\hat{\mu}_{Y(HT)}$ is given in the following theorem.
Theorem 2.2 The variance of $\hat{\mu}_{Y(HT)}$ is given by

$$V(\hat{\mu}_{Y(HT)}) = \frac{\mu_Y^2}{n}\Bigg[C_y^2 + (1+C_y^2)\Bigg\{C_P^2 + \frac{(1-P)\theta^2[k^2(\Delta(S) + (\sqrt{\beta_1(S)} - C_\gamma)^2) - 2kC_\gamma(C_\gamma - \sqrt{\beta_1(S)})]}{[P + \theta(1-P)]^2}\Bigg\}\Bigg] \tag{2.3}$$

where $\Delta(S) = [\beta_2(S) - \beta_1(S) - 1], \beta_2(S) = \dfrac{\mu_4(S)}{\gamma^4}, \beta_1(S) = \dfrac{\mu_3^2(S)}{\gamma^6}$,
$\mu_3(S) = E(S-\theta)^3$ and $\mu_4(S) = E(S-\theta)^4$.

Proof-

$$V(\hat{\mu}_{Y(HT)}) = V(\bar{Z}) = \frac{V(Z_i)}{n[P+\theta(1-P)]^2} \tag{2.4}$$

The variance of Z_i is obtained as follows:
$V(Z_i)$=$E(Z_i^2) - (E(Z_i))^2$

$$\left.\begin{aligned}
&=(1-p)E[(1-k)^2S^2 + \theta^2k^2S^{*4} + 2k(1-k)\theta SS^{*2}] \\
&\quad E(Y_i^2) + PE(Y_i^2) - (E(Z_i))^2 \\
&=\mu_Y^2[(1+C_y^2)[P+(1-P)\theta^2(1+C_\gamma^2)] + (1+C_y^2)\theta^2(1-P) \\
&\quad [k^2(\beta_2(S) - 2C_\gamma\sqrt{\beta_1(S)} + C_\gamma^2 - 1) - 2kC_\gamma(C_\gamma - \sqrt{\beta_1(S)}) \\
&\quad - (P+\theta(1-P))^2]
\end{aligned}\right\}$$

Thus, the variance of $\hat{\mu}_{Y(HT)}$is given by

$$V(\hat{\mu}_{Y(HT)}) = \frac{\mu_Y^2}{n}[C_y^2 + (1+C_y^2)[C_P^2 + \frac{(1-P)\theta^2[k^2[\Delta(S) + (C_\gamma - \sqrt{\beta_1(S)})^2] - 2kC_\gamma(C_\gamma - \sqrt{\beta_1(S)})]}{[P+\theta(1-P)]^2}]$$

which proves the theorem.
Theorem 2.3 The optimum value of k and the minimum variance of $\hat{\mu}_{Y(HT)}$ are respectively given by

$$k_{opt} = \frac{C_\gamma(C_\gamma - \sqrt{\beta_1(S)})}{[\Delta(S) + (C_\gamma - \sqrt{\beta_1(S)})^2]} \tag{2.5}$$

and

$$min.V(\hat{\mu}_{Y(HT)}) = \frac{\mu_Y^2}{n}[C_y^2 + (1+C_y^2)C_P^2 - \frac{(1+C_y^2)\theta^2(1-P)C_\gamma^2(C_\gamma - \sqrt{\beta_1(S)})^2}{[\Delta(S) + (C_\gamma - \sqrt{\beta_1(S)})^2](P+\theta(1-P))^2}] \tag{2.6}$$

$$min.V(\hat{\mu}_{Y(HT)}) = V(\hat{\mu}_{Y(BBB)}) - \frac{\mu_Y^2(1+C_y^2)\theta^2(1-P)C_\gamma^2(C_\gamma - \sqrt{\beta_1(S)})^2}{n[\Delta(S) + (C_\gamma - \sqrt{\beta_1(S)})^2](P+\theta(1-P))^2} \tag{2.7}$$

where $V(\hat{\mu}_{Y(BBB)})$ is given by (1.5).
proof - Differentiating (2.3) with respect to k and equating to zero, we get the optimum value of k as

$$k_{opt} = \frac{C_\gamma(C_\gamma - \sqrt{\beta_1(S)})}{[\Delta(S) + (C_\gamma - \sqrt{\beta_1(S)})^2]}$$

Substitution of k_{opt} in (2.3) yields the minimum variance of $\hat{\mu}_{Y(HT)}$ as given in (2.6) (or (2.7)).
This completes the proof of the theorem.
Now substituting the value of k_{opt} in place of k in (2.1) we get the distribution of the responses as

Taking expectation of(2.8), we have
$E(Z_{oi}) = \mu_Y[P + \theta(1-P)]$.

$$(2.8) \qquad Z_{oi} = \left| \begin{array}{c} Y_i[(1-k_{opt})S + K_{opt}\theta S^{*2}] \text{ with probability } (1-P) \\ \\ Y_i \text{ with probability } P. \end{array} \right.$$

Thus the unbiased estimator of the population mean μ_y based on Z_{oi} is given by

$$(2.9) \qquad \hat{\mu}_{Y(HTO)} = \frac{\bar{Z}_o}{[(1-P)\theta + P]} = \frac{\sum_{i=1}^{n} \frac{\bar{Z}_{oi}}{n}}{[(1-P)\theta + P]}$$

it can be easily shown that the variance of $\hat{\mu}_{Y(HTO)}$ is:

$$(2.10) \qquad V(\hat{\mu}_{Y(HTO)}) = min.V(\hat{\mu}_{Y(HT)})$$

where $min.V(\hat{\mu}_{Y(HT)})$ is given by (2.6) (or(2.7)).
It is well known that $\beta_2(S) > \beta_1(S) + 1$ [Kendal and Stuart (1969)]. Hence the optimum estimator $\hat{\mu}_{Y(HTO)}$ is always more efficient than the Bar – Lev et al.'s (2004) estimator $\hat{\mu}_{Y(BBB)}$ except for population with $\sqrt{\beta_1(S)} = C_\gamma$ for which $\hat{\mu}_{Y(HTO)}$ is as efficient as $\hat{\mu}_{Y(BBB)}$.

3. Efficiency Comparison

(i) Comparison of the proposed optimum estimator $\hat{\mu}_{Y(HTO)}$ (i.e. when the scalar 'k' coincides exactly with that of optimum value k_{opt} of the scalar k) with Bar – Lev et al.'s (2004) estimator$\hat{\mu}_{Y(BBB)}$.
From (1.5) and (2.7), we have

$$V(\hat{\mu}_{Y(BBB)}) - min.V(\hat{\mu}_{Y(HT)})[= V(\hat{\mu}_{Y(HTO)})] =$$

$$(3.1) \qquad \frac{\mu_Y^2(1+C_y^2)\theta^2(1-P)C_\gamma^2(C_\gamma - \sqrt{\beta_1(S)})^2}{n[\Delta(S) + (C_\gamma - \sqrt{\beta_1(S)})^2](P + \theta(1-P))^2} > 0$$

which clearly shows that the proposed optimum estimator $\hat{\mu}_{Y(HTO)}$ is better than the estimator$\hat{\mu}_{Y(BBB)}$ due to Bar – Lev et al. (2004).
Bar – Lev et al. (2004) have proved that for all $P\varepsilon(0,1)$:

$$(3.2) \qquad V(\hat{\mu}_Y = \bar{Y}) < V(\hat{\mu}_{Y(BBB)}) < V(\hat{\mu}_{Y(EH)}).$$

If the distribution of scrambling variables S satisfies

$$(3.3) \qquad 0 < \theta < \frac{2\theta^2(1+C_\gamma^2)}{[1+\theta^2(1+C_\gamma^2)]}$$

where

$$\hat{\mu}_Y = \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}.$$

Thus, under the condition (3.3) and (3.1) we have the following inequality:

$$(3.4) \qquad V(\hat{\mu}_{Y(HTO)}) < V(\hat{\mu}_{Y(BBB)}) < V(\hat{\mu}_{Y(EH)}).$$

It follows from (3.4) that the proposed optimum estimator $\hat{\mu}_{Y(HTO)}$ is more efficient than the Eichhorn and Hayre (1983) estimator $\hat{\mu}_{Y(EH)}$ as long as the condition (3.3) is satisfied.

(ii) Comparison of the proposed estimator $\hat{\mu}_{Y(HT)}$ with Bar – Lev et al.'s (2004) estimator$\hat{\mu}_{Y(BBB)}$ when the value of k does not coincide exactly with its optimum value k_{opt} in (2.5).

From (1.5) and (2.3), we have

$$(3.5)\qquad V(\hat{\mu}_{Y(HT)}) = V(\hat{\mu}_{Y(BBB)}) + \frac{\mu_Y^2(1+C_y^2)\theta^2(1-P)}{n(P+\theta(1-P))^2}[k^2A - 2kB],$$

where
$A = [\Delta(S) + (\sqrt{\beta_1(S)} - C\gamma)^2]$ and $B = C_\gamma(C_\gamma - \sqrt{\beta_1(S)})$
We note that

$$V(\hat{\mu}_{Y(HT)}) - V(\hat{\mu}_{Y(BBB)}) = \frac{\mu_Y^2(1+C_y^2)\theta^2(1-P)}{n(P+\theta(1-P))^2}[k^2A - 2kB]$$

which is negative if

$$k^2A - 2kB < 0$$

i.e. if

$$(3.6)\qquad |k - k_{opt}| < |k_{opt}|$$

i.e. if

$$(3.7)\qquad either 0 < k < 2k_{opt} or 2k_{opt} < k < 0$$

or equivalently,

$$(3.8)\qquad min.(0, 2k_{opt}) < k < max.(0, 2k_{opt}),$$

where $k_{opt} = \dfrac{B}{A}$
Thus, the proposed estimator proposed estimator $(\hat{\mu}_{Y(HT)})$ is more efficient than Bar – Lev et al.'s (2004) estimator $(\hat{\mu}_{Y(BBB)})$ as long as the condition (3.8) is satisfied.
Now, in the following sections we shall discuss our general results in the context of normal and waiting time distributions.

4. Normal Distribution

Let the scrambling variable S have a normal distribution with mean θ and variance γ^2 i.e.$S \sim N(\theta, \gamma^2)$. For this distribution $\sqrt{\beta_1(S)} = 0$ and $\sqrt{\beta_2(S)} = 3 \Rightarrow \Delta(S) = 2$.Thus the optimum value of k_{opt} in (2.5) and the minimum variance (or the variance of the optimum estimator $\hat{\mu}_{Y(HTO)}$ in (2.6)(or(2.7)) respectively reduce to:

$$(4.1)\qquad k_{opt} = \frac{C_\gamma^2}{(2 + C_\gamma^2)}$$

and

$$(4.2)\quad min.V(\hat{\mu}_{Y(HT)}) = \frac{\mu_y^2}{n}[C_y^2 + (1+C_y^2)C_P^2 - \frac{(1+C_y^2)\theta^2(1-P)C_\gamma^4}{(2+C_\gamma^2)(P+\theta(1-P))^2}] = V(\hat{\mu}_{Y(HTO)})$$

Here $\hat{\mu}_{Y(HTO)}$ is defined by

$$(4.3) \qquad \hat{\mu}_{Y(HTO)} = \frac{\sum_{i=1}^{n} Z_{oi}^*}{n}$$

where Z_{oi}^* is defined by

$$(4.4) \qquad Z_{oi}^* = \left| \begin{array}{l} Y_i[\frac{2S}{(2+C_\gamma^2)} + \frac{\theta C_\gamma^2 S^{*2}}{(2+C_\gamma^2)}] \text{ with probability } (1-P) \\ \qquad\qquad Y_i \text{ with probability } P. \end{array} \right.$$

It is interesting to note that the optimum $\hat{\mu}_{Y(HTO)}$ in (4.3) can be used in practice as the coefficient of variation C_γ is known without error.

5. Numerical Illustration using Normal Distribution

To judge the merit of the suggested optimum estimator over Eichhorn and Hayre (1983) estimator $\hat{\mu}_{Y(EH)}$ and the Bar – Lev et al. (2004) estimator $\hat{\mu}_{Y(BBB)}$, we have computed the percent relative efficiency (PRE) of the optimum estimator $\hat{\mu}_{Y(HTO)}$ with respect to the estimators $\hat{\mu}_{Y(BBB)}$ and $\hat{\mu}_{Y(EH)}$ by using the formulae:

$$(5.1) \qquad PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(EH)}) = \frac{[C_y^2 + C_\gamma^2(1+C_y^2)]}{[C_y^2 + C_P^2(1+C_y^2) - A_1]} \times 100.$$

$$(5.2) \qquad PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(BBB)}) = \frac{[C_y^2 + C_P^2(1+C_y^2)]}{[C_y^2 + C_P^2(1+C_y^2) - A_1]} \times 100.$$

for different values of C_y, C_γ, P, θ,
where

$$(5.3) \qquad A_1 = \frac{[(1+C_y^2)\theta^2(1-P)C_\gamma^4]}{[2 + C_\gamma^2[P + \theta(1-P)]^2]} \times 100.$$

Findings are displayed in Tables 1 and 2; and the graphical representation is also given in Figure 3.
The values of $PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(EH)})$ and $PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(BBB)})$ are much greater than 100 as shown by Tables 1 and 2. It follows that the proposed optimum estimator $\hat{\mu}_{Y(HTO)}$ is more efficient than Eichhorn and Hayre's (1983) estimator $\hat{\mu}_{Y(EH)}$ and Bar – Lev et al.'s (2004) estimator $\hat{\mu}_{Y(BBB)}$ with considerable gain in efficiency. These facts can be also seen from Figure 3. Thus, based on our numerical results, the use of the proposed estimator $\hat{\mu}_{Y(HTO)}$ is recommended for its use in practice.

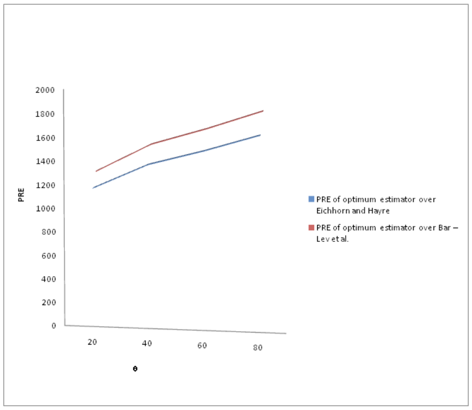


Figure 3. Graphical representation of the suggested optimum estimator over Eichhorn and Hayre (1983) estimator $\hat{\mu}_{Y(EH)}$ and the Bar – Lev et al. (2004) estimator $\hat{\mu}_{Y(BBB)}$.

6. Waiting Time Distribution

We consider the population, where scrambling variable S follows the waiting time distribution (or distribution of intervals between events in a Poisson process) for which

$$(6.1) \qquad f(s) = 1 - exp(-\frac{s}{\theta}), 0 \leq s \leq \propto, \theta > 0$$

so that

$$(6.2) \qquad dF(s) = exp(-\frac{s}{\theta})\frac{ds}{\theta}$$

and $E(S) = \theta, V(S) = \theta^2, \mu_3(S) = 2\theta^3$ and $\mu_4(S) = 9\theta^4$
where $C_\gamma = 1, \sqrt{\beta_1(S)} = 2, \beta_2(S) = 9, \Delta(S) = 4$. Hence, substituting the values of C_γ, $\sqrt{\beta_1(S)}, \beta_2(S) and \Delta(S) in (2.5) and (2.6)$, we have

$$(6.3) \qquad k_{opt} = -\frac{1}{5},$$

and

$$(6.4) \qquad min.V(\hat{\mu}_{Y(HT)}) = \frac{\mu_y^2}{n}[C_y^2 + (1 + C_y^2)C_P^2 - \frac{(1 + C_y^2)\theta^2(1 - P)}{5(P + \theta(1 - P))^2}] = V(\hat{\mu}_{Y(HTO)})$$

Here the optimum estimator $\hat{\mu}_{Y(HTO)}$ is defined by

$$(6.5) \qquad \hat{\mu}_{Y(HTO)} = \frac{\sum_{i=1}^{n} Z_{oi}^{**}}{n}$$

where Z_{oi}^{**} is defined by

with $S^* = \dfrac{(S - \theta)}{\theta}$.

$$(6.6) \qquad Z_{oi}^{**} = \left| \begin{array}{l} Y_i[\frac{6}{5}S - \frac{1}{5}\theta S^{*2}] \text{ with probability } (1-P) \\ \\ \qquad\qquad Y_i \text{ with probability } P. \end{array} \right.$$

Thus in this case we note that the optimum estimator$\hat{\mu}_{Y(HTO)}$ in (6.5) depends on the known quantity θ only.

7. Numerical Illustration using Waiting Time Distribution

To have the tangible idea about the performance of the envisaged optimum estimator $\hat{\mu}_{Y(HTO)}$ over Eichhorn and Hayre (1983) estimator $\hat{\mu}_{Y(EH)}$ and the Bar – Lev et al. (2004) estimator $\hat{\mu}_{Y(BBB)}$, we have computed the percent relative efficiency (PRE) of the optimum estimator $\hat{\mu}_{Y(HTO)}$ with respect to the estimators $\hat{\mu}_{Y(BBB)}$ and $\hat{\mu}_{Y(EH)}$ by using the formulae:

$$(7.1) \qquad PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(EH)}) = \frac{[C_y^2 + C_\gamma^2(1+C_y^2)]}{[C_y^2 + C_P^2(1+C_y^2) - A_2]} \times 100.$$

$$(7.2) \qquad PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(BBB)}) = \frac{[C_y^2 + C_P^2(1+C_y^2)]}{[C_y^2 + C_P^2(1+C_y^2) - A_2]} \times 100.$$

for different values of C_y, C_γ, P, θ,
where

$$(7.3) \qquad A_2 = \frac{[(1+C_y^2)\theta^2(1-P)}{[5[P+\theta(1-P)]^2]} \times 100.$$

Findings are displayed in Tables 3 and 4; and the graphical representation is also given in Figure 4.
Tables 3 and 4 demonstrate that the values of the percent relative efficiency are greater than 100 for all parameter values tabled. This shows the superiority of the optimum estimator $\hat{\mu}_{Y(HTO)}$ over than Eichhorn and Hayre's (1983) estimator $\hat{\mu}_{Y(EH)}$ and Bar – Lev et al.'s (2004) estimator $\hat{\mu}_{Y(BBB)}$. Graphical representation in Figure 4 also depicts the similar inference. Thus, based on our numerical illustrations, our recommendation is to prefer the proposed estimator $\hat{\mu}_{Y(HTO)}$ in practice.

8. Discussion

In this article, we have suggested a new randomized response model and its properties are studied. It has been shown that the resulting (optimum) randomized response model depends on the moments ratios of the scrambling variable S. We have proved the superiority of the proposed randomized response model over Eichhorn and Hayre (1983) and Bar – Lev et al.'s (2004) randomized response models both theoretically and empirically.

Acknowledgements

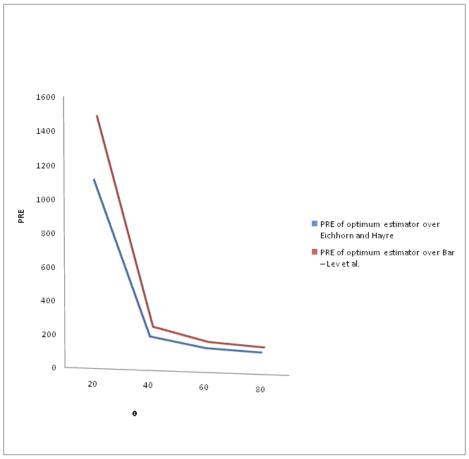


Figure 4. Graphical representation of the suggested optimum estimator over Eichhorn and Hayre (1983) estimator $\hat{\mu}_{Y(EH)}$ and the Bar – Lev et al. (2004) estimator $\hat{\mu}_{Y(BBB)}$.

The authors are thankful to the Editor – in- Chief - Professor Cem- Kadilar, and to the anonymous learned referees for their valuable suggestions regarding improvement of the paper.

References

[1] Ahsanullah M. and Eichhorn B.H. (1988). *On estimation of response from scrambled quantitative data*. Pakistan Journal of Statistics, 4(2), 83-91.

[2] Bar –Lev S.K., Bobovitch E. and Boukai B. (2004). *A note on Randomized response models for quantitative data*. Metrika, 60, 225-250.

[3] Chaudhuri A. and Mukerjee R. (1988). *Randomized Response: Theory and Techniques*. Marcel- Dekker, New York, USA.

[4] Chaudhuri A. and Christofides T. (2013). *Indirect Questioning in Sample Surveys*. DOI 10.1007/978-3-642-36276-7-3, Springer – Verlag Berlin Heidelberg.

[5] Eichhorn B.H. and Hayre, L.S. (1983). *Scrambled randomized response methods for obtaining sensitive quantitative data*. Journal of Statistics Planning and Inference,7,307-316.

[6] Fox J.A. and Tracy P.E. (1986). *Randomized Response: A method of Sensitive Surveys*. Newbury Park, CA: SEGE Publications.

[7] Gjestvang C.R. and Singh S. (2006). *A new randomized response model*. Journal of the Royal Statistical Society, 68, 523-530.

[8] Gjestvang C.R. and Singh S. (2009). *An improved randomized response model: Estimation of mean*, Journal of Applied Statistics, 36(12), 1361-1367.

[9] Gleser L.J. and Healy J.D. (1976). *Estimating the mean of a normal distribution with known coefficient of variation*. Journal of the American Statistical Association, 71,977-981.

[10] Govindaragulu Z. and Sahai H. (1972). *Estimation of the parameters of a normal distribution with known coefficient of variation*. Republic Statistics and Applications and Research, JUSU, 19,86-98.

[11] Gupta S. and Thornton B. (2002). *Circumventing social desirability response bias in personal interview surveys*. American Journal of Mathematics Managment and Sciences, 22, 369-383.

[12] Hussain Z., Hamraz M. and shabbier J. (2013) . *On alternative estimation technique for randomized response model*. Pakistan Journal of Statistics, 29(3), 283-306.

[13] Hussain Z., Mashail, A.M., Bander A., Singh H.P. and Tarray T.A.(2015). *Improved randomized response approaches for additive scrambling models* . Mathematical Population Studiies. (In Press).
[14] Khan RA (1968). *A note on estimating the mean of a normal distribution with known coefficient of variation*. Journal of the American Statistical Association, 63,1039-1041.
[15] Kerkvliet J.(1994). *Estimating a logit model with randomized data: the case of cocaine use.* Austrialian Journal of Statistics,36,9-20.
[16] Mangat N.S. and Singh R. (1990). *An alternative randomized procedure*. Biometrika, 77, 439-442.
[17] Odumade O. and Singh S. (2009). *Improved Bar – lev, Bobovitch and Boukai randomized response models*. Communication in Statistics Theory and Methods, 38(3), 473-502.
[18] Searls D. T. (1964). *The utilization of a known coefficient of variation in the estimation procedure*. Journal of the American Statistical Association, 59, 12225-12226.
[19] Sen A.R. (1978). *Estimation of the mean when the coefficient of variation is known*. Communication in Statistics Theory and Methods, 7(7), 657-672.
[20] Sen A.R. (1979). *Relative efficiency of the estimators of the mean of a normal distribution when coefficient of variation is known*. Biometrical. Journal,21(2), 131-137.
[21] Singh S. and Cheng S.C. (2009). *Utilization of higher order moments of scrambling variables in randomized response sampling*. Journal of Statististics Planning and Inference,139, 3377-3380.
[22] Singh H.P. and Katiyar N.P. (1988). *A generalized class of estimators for common parameters of two normal distributions with known coefficient of variation*. Journal of the Indian Society of the Agricultural Statistics, 40, (2), 127-149.
[23] Singh H.P. and Mathur N. (2005). *Estimation of population mean when coefficient of variation is known using scrambled response technique*. Journal of Statististics Planning and Inference,131 (1), 135-144.
[24] Singh H.P. and Tarray T.A. (2013). *A modified survey technique for estimating the proportion and sensitivity in a dichotomous finite population*. International Journal of the Advanved Science and and Technology Research, 3(6), 459 – 472.
[25] Singh H.P. and Tarray T.A. (2014). *An alternative to stratified Kim and Warde's randomized response model using optimal (Neyman) allocation*. Model Assisted Statistical Applications, 9, 37-62.
[26] Singh H.P. and Tarray T.A. (2015). *An efficient use of moment's ratios of scrambling variables in randomized response sampling*. Communication in Statistics Theory and Methods, (In Press).
[27] Tarray T.A. and Singh H.P. and (2015). *An improved new additive model*. Gazi University Journal of Sciences, (In Press).
[28] Tarray T.A., Singh H.P. and Zaizai Y. (2015). *A stratified optional randomized response model*. Sociological Methods and Research, 1-15.
[29] Tripathi T.P., Maiti P. and Sharma S.D. (1983). *Use of prior information on some parameters in estimating population mean* . Sankhya, 45, A, 3, 372-376.
[30] Upadhyaya L.N. and Singh H.P. (1984) . *On the estimation of the population mean with known coefficient variation*. Biometrical Journal, 26, (8), 915-922.
[31] Warner S.L. (1965). *Randomized response: A survey technique for eliminating evasive answer bias*. Journal of the American Statistical Association, 60, 63-69.

Table 1. The $PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(EH)})$

θ	P	C_γ	C_y	PRE
20.00	0.10	5.00	0.10	1166.50
40.00	0.10	5.55	0.25	1378.32
60.00	0.10	6.00	0.50	1505.49
80.00	0.10	6.50	0.75	1649.02
20.00	0.20	5.00	0.10	1005.22
40.00	0.20	5.55	0.25	1181.50
60.00	0.20	6.00	0.50	1298.17
80.00	0.20	6.50	0.75	1432.54
20.00	0.30	5.00	0.10	857.59
40.00	0.30	5.55	0.25	1000.69
60.00	0.30	6.00	0.50	1104.65
80.00	0.30	6.50	0.75	1227.07
20.00	0.40	5.00	0.10	722.02
40.00	0.40	5.55	0.25	834.03
60.00	0.40	6.00	0.50	923.61
80.00	0.40	6.50	0.75	1031.80
20.00	0.50	5.00	0.10	597.20
40.00	0.50	5.55	0.25	679.95
60.00	0.50	6.00	0.50	753.89
80.00	0.50	6.50	0.75	845.99
20.00	0.60	5.00	0.10	482.12
40.00	0.60	5.55	0.25	537.15
60.00	0.60	6.00	0.50	594.50
80.00	0.60	6.50	0.75	668.99
20.00	0.70	5.00	0.10	376.14
40.00	0.70	5.55	0.25	404.56
60.00	0.70	6.00	0.50	444.59
80.00	0.70	6.50	0.75	500.24
20.00	0.80	5.00	0.10	279.54
40.00	0.80	5.55	0.25	281.53
60.00	0.80	6.00	0.50	303.54
80.00	0.80	6.50	0.75	339.30

Table 2. The $PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(BBB)})$

θ	P	C_γ	C_y	PRE
20.00	0.10	5.00	0.10	1286.41
40.00	0.10	5.55	0.25	1527.41
60.00	0.10	6.00	0.50	1670.15
80.00	0.10	6.50	0.75	1829.85
20.00	0.20	5.00	0.10	1234.45
40.00	0.20	5.55	0.25	1467.05
60.00	0.20	6.00	0.50	1616.15
80.00	0.20	6.50	0.75	1784.75
20.00	0.30	5.00	0.10	1186.85
40.00	0.30	5.55	0.25	1411.58
60.00	0.30	6.00	0.50	1565.74
80.00	0.30	6.50	0.75	1714.95
20.00	0.40	5.00	0.10	11.43.09
40.00	0.40	5.55	0.25	1360.44
60.00	0.40	6.00	0.50	1518.57
80.00	0.40	6.50	0.75	1701.27
20.00	0.50	5.00	0.10	1102.71
40.00	0.50	5.55	0.25	1313.14
60.00	0.50	6.00	0.50	1474.34
80.00	0.50	6.50	0.75	1662.55
20.00	0.60	5.00	0.10	1065.34
40.00	0.60	5.55	0.25	1269.25
60.00	0.60	6.00	0.50	1432.78
80.00	0.60	6.50	0.75	1625.65
20.00	0.70	5.00	0.10	1030.65
40.00	0.70	5.55	0.25	1228.42
60.00	0.70	6.00	0.50	1393.65
80.00	0.70	6.50	0.75	1590.44
20.00	0.80	5.00	0.10	998.36
40.00	0.80	5.55	0.25	1190.34
60.00	0.80	6.00	0.50	1356.73
80.00	0.80	6.50	0.75	1556.80

Table 3. The $PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(EH)})$

θ	P	C_γ	C_y	PRE
20.00	0.05	0.40	0.10	1112.83
40.00	0.05	0.50	0.25	191.56
60.00	0.05	0.60	0.50	133.67
80.00	0.05	0.70	0.75	118.70
20.00	0.06	0.40	0.10	683.47
40.00	0.06	0.50	0.25	179.20
60.00	0.06	0.60	0.50	129.80
80.00	0.06	0.70	0.75	116.41
20.00	0.07	0.40	0.10	490.56
40.00	0.07	0.50	0.25	168.14
60.00	0.07	0.60	0.50	126.07
80.00	0.07	0.70	0.75	114.17
20.00	0.08	0.40	0.10	380.96
40.00	0.08	0.50	0.25	158.18
60.00	0.08	0.60	0.50	122.47
80.00	0.08	0.70	0.75	111.97
20.00	0.09	0.40	0.10	310.27
40.00	0.09	0.50	0.25	149.15
60.00	0.09	0.60	0.50	119.01
80.00	0.09	0.70	0.75	109.80
20.00	0.10	0.40	0.10	260.91
40.00	0.10	0.50	0.25	140.94
60.00	0.10	0.60	0.50	115.66
80.00	0.10	0.70	0.75	107.68
20.00	0.11	0.40	0.10	224.48
40.00	0.11	0.50	0.25	133.44
60.00	0.11	0.60	0.50	112.44
80.00	0.11	0.70	0.75	105.59
20.00	0.12	0.40	0.10	196.49
40.00	0.12	0.50	0.25	126.56
60.00	0.12	0.60	0.50	109.32
80.00	0.12	0.70	0.75	103.54

Table 4. The $PRE(\hat{\mu}_{Y(HTO)}, \hat{\mu}_{Y(BBB)})$

θ	P	C_γ	C_y	PRE
20.00	0.05	0.40	0.10	1471.69
40.00	0.05	0.50	0.25	230.24
60.00	0.05	0.60	0.50	150.17
80.00	0.05	0.70	0.75	129.36
20.00	0.06	0.40	0.10	950.47
40.00	0.06	0.50	0.25	223.07
60.00	0.06	0.60	0.50	149.21
80.00	0.06	0.70	0.75	129.09
20.00	0.07	0.40	0.10	716.29
40.00	0.07	0.50	0.25	216.65
60.00	0.07	0.60	0.50	148.29
80.00	0.07	0.70	0.75	128.83
20.00	0.08	0.40	0.10	583.23
40.00	0.08	0.50	0.25	210.86
60.00	0.08	0.60	0.50	147.41
80.00	0.08	0.70	0.75	128.57
20.00	0.09	0.40	0.10	497.42
40.00	0.09	0.50	0.25	205.62
60.00	0.09	0.60	0.50	146.55
80.00	0.09	0.70	0.75	128.32
20.00	0.10	0.40	0.10	437.49
40.00	0.10	0.50	0.25	200.86
60.00	0.10	0.60	0.50	145.73
80.00	0.10	0.70	0.75	128.07
20.00	0.11	0.40	0.10	393.27
40.00	0.11	0.50	0.25	196.50
60.00	0.11	0.60	0.50	144.93
80.00	0.11	0.70	0.75	127.83
20.00	0.12	0.40	0.10	359.30
40.00	0.12	0.50	0.25	192.51
60.00	0.12	0.60	0.50	144.17
80.00	0.12	0.70	0.75	127.59

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 609 – 628

On the Bayesian analysis of 3-component mixture of exponential distributions under different loss functions

Muhammad Tahir* †, Muhammad Aslam‡ and Zawar Hussain§

Abstract

The memory-less property of the Exponential distribution is a strong reason of its use for testing lifetimes of objects in many lifetime modeling applications. Also, mixture models have extensively been used in survival analysis and reliability studies. This article focuses on the Bayesian analysis of the 3-component mixture of Exponential distributions under type-I right censoring scheme. Taking different non-informative and informative priors, Bayes estimators and posterior risks for the unknown parameters (parameters of component distributions and mixing proportions) are derived under squared error loss function, precautionary loss function and DeGroot loss function. The elicitation of the hyperparameters is also done using prior predictive distribution.The Bayes estimators and posterior risks are looked at as a function of the test termination time. Some important properties and comparisons of the Bayes estimates are presented. Simulated results and real data example are also given to illustrate the study.

Keywords: 3-Component mixture distribution, Non-informative and informative priors, Loss function, Bayes estimators, Posterior risks, Test termination time.

2000 AMS Classification: 62F15, 62N05.

Received : 26.06.2014 *Accepted* : 13.02.2015 *Doi* : 10.15672/HJMS.2015519451

*Department of Statistics, Government College University, Faisalabad, Pakistan.
†Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
Email: tahirqaustat@yahoo.com Corresponding Author.
‡Department of Mathematics and Statistics, Riphah International University, Islamabad, Pakistan.
Email: aslamsdqu@yahoo.com
§Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
Email: zhlangah@yahoo.com

1. Introduction

The Exponential distribution, because of its memory-less property, has many real life applications in testing lifetimes of objects where lifetimes do not depend upon their ages. There are many electronic devices whose failure rate does not depend on their ages, therefore, the Exponential distribution is suitable to model the lifetimes. Generally, in lifetime modeling, population is supposed to be composed of more than one subpopulations mixed by unknown mixing proportions. In our study, we take the data from a population which is characterized by three different members of the Exponential family of distributions. McCullagh (1994) derived some conditions under which quadratic and polynomial Exponential models can be generated as mixtures of the Exponential models. Raqab and Ahsanullah (2001) discussed the location and scale parameters of generalized Exponential distribution based on order statistic. Hebert and Scariano (2005) compared the location estimators for the Exponential mixtures under Pitman's measure of closeness. Ali et al. (2005) studied the Bayes estimators of the Exponential distribution and Abu-Taleb et al. (2007) presented the Bayesian estimation of lifetime parameters of Exponential distributions when survival time and censoring time are both exponentially distributed.

The use of mixture models in situations where data are given only for overall mixture distributions is known as direct application of the mixture models. Li (1983) and Li and Sedransk (1982, 1988) discussed different features of mixture models and defined two types of mixture models. The mixture of the probability density functions from the same family is known as type-I mixture model and type-II mixture model is defined as a mixture of density functions from several families. In this study, the direct application of mixture model (with the unknown component and mixing proportion parameters of the 3-component mixture of Exponential distributions) is considered under type-I mixture modeling.

Due to the development of advanced computational facilities, researchers are now able to find the Bayes estimates, infer and predict about complex systems such as mixture models. With the provision of these computational facilities, the Bayesian technique to analyze a 3-component mixture model has developed the interest among many researchers. The posterior distribution, which is obtained when prior information is combined with likelihood, is the workbench of Bayesian inference. Thus, the prior information, a subjective assessment by an expert before the data are actually gathered, is very important and necessary for Bayesian inference.In this study, the Bayesian analysis of a 3-component mixture of Exponential distributions using the non-informative (uniform and Jeffreys') priors and the informative prior (IP) under squared error loss function (SELF), precautionary loss function (PLF) and DeGroot loss function (DLF) is considered.

There are many fields such as engineering, biological sciences, physical sciences and social sciences where mixture models have been used quite effectively. Most of the researchers worked on the Bayesian analysis of 2-component mixture models. For example, Sinha (1998) used the Bayesian counterpart of the maximum likelihood estimates of the 2-component mixture model considered by Mendenhall and Hader (1958). Saleem and Aslam (2008) discussed the use of the informative and the non-informative priors for Bayesian analysis of the 2-component mixture of Rayleigh distributions. Saleem et al. (2010) presented the Bayesian analysis of the 2-component mixture of Power distributions using the complete and censored data. Kazmi et al. (2012) developed the Bayesian analysis for the 2-component mixture of Maxwell distributions.

In real life applications, most of the times, it is not suitable to continue the testing procedure until failure of the last object under testing. In such situations, censored samples are observed. Censoring is an important and valuable aspect of lifetime applications. A valuable account on censoring is given in Romeu (2004), Gijbels (2010) and Kalbfleisch and Prentice (2011). In this paper, an ordinary type-I right censoring is used with fixed life-test termination time for all objects.

The rest of the paper is organized as follows. In Section 2, the 3-component mixture of Exponential distributions is presented. Posterior distributions using the uniform prior (UP), the Jeffreys' prior (JP) and the informative prior (IP) are derived in Section 3. The Bayes estimators and posterior risks using the UP, the JP and the IP under SELF, PLF and DLF are presented in Sections 4, 5 and 6, respectively. The elicitation of hyperparameters is described in Section 7. The limiting expressions are derived in Section 8. A simulation study and real data example are discussed in Sections 9 and 10, respectively. Finally, the conclusion of the study is given in Section 11.

2. 3-component mixture of exponential distributions

If X is exponentially distributed with parameter θ_m, its probability density function is given as:

$$f_m\left(x;\ \theta_m\right) = \theta_m\ \exp\left(-\theta_m\ x\right),\ x \geq 0,\ \theta_m > 0,\ m = 1,\ 2,\ 3. \tag{2.1}$$

According to Barger (2006) and Strelec and Stehlk (2012), a finite 3-component mixture of Exponential distributions with unknown mixing proportions p_1 and p_2 is defined as:

$$f\left(x\right) = p_1\ f_1\left(x\right) + p_2\ f_2\left(x\right) + \left(1 - p_1 - p_2\right)\ f_3\left(x\right), p_1,\ p_2 \geq 0,\ p_1 + p_2 \leq 1 \tag{2.2}$$

$$\begin{aligned} f\left(x;\ \theta_1,\ \theta_2,\ \theta_3,\ p_1,\ p_2\right) &= p_1\ \theta_1 \exp\left(-\theta_1\ x\right) + p_2\ \theta_2 \exp\left(-\theta_2\ x\right) \\ &+ \left(1 - p_1 - p_2\right)\ \theta_3 \exp\left(-\theta_3\ x\right) \end{aligned} \tag{2.3}$$

As cumulative distribution function of the random variable X is given by:

$$F_m\left(x\right) = 1 - \exp\left(-\theta_m\ x\right),\ m = 1,\ 2,\ 3, \tag{2.4}$$

the cumulative distribution function of 3-component mixture distribution is defined as:

$$F\left(x\right) = p_1\ F_1\left(x\right) + p_2\ F_2\left(x\right) + \left(1 - p_1 - p_2\right)\ F_3\left(x\right) \tag{2.5}$$

$$F\left(x\right) = 1 - p_1\ \exp\left(-\theta_1\ x\right) - p_2\ \exp\left(-\theta_2\ x\right) - \left(1 - p_1 - p_2\right)\ \exp\left(-\theta_3\ x\right) \tag{2.6}$$

3. The posterior distribution using the UP, the JP and the IP

The posterior distributions of parameters given data $\mathbf{x}$ are derived using the UP, the JP and the IP.

3.1. The likelihood function. Suppose n units are used in a life testing experiment with the 3-component mixture modeling. Let r out of n units fail before fixed test termination time t and the remaining $n - r$ units are still working. According to Mendenhall and Hader (1958), there are many practical situations in which the failing objects can be pointed out easily as subset of subpopulation-1, subpopulation-2or subpopulation-3. Out of r units, suppose r_1, r_2 and r_3 units belong to subpopulation-1, subpopulation-2 and subpopulation-3,respectively,such that $r = r_1 + r_2 + r_3$. Now, define x_{lk}, $0 < x_{lk} \leq t$, as the failure time of k^{th} $(k = 1,\ 2, \cdots,\ r_l)$ unit belong to l^{th} $(l = 1,\ 2,\ 3)$ subpopulation. Thus, the likelihood function of the 3-component mixture model for the random sample vector x is given as(cf. Everitt and Hand, 1981):

$$L(\boldsymbol{\psi}\,|\mathbf{x}) \propto \left\{\prod_{k=1}^{r_1} p_1\; f_1(x_{1k})\right\}\left\{\prod_{k=1}^{r_2} p_2\; f_2(x_{2k})\right\}\left\{\prod_{k=1}^{r_3}(1-p_1-p_2)\; f_3(x_{3k})\right\}$$
$$\{1-F(t)\}^{n-r} \tag{3.1}$$

$$\begin{aligned} L(\boldsymbol{\psi}\,|\mathbf{x}) \propto\ & \theta_1^{r_1}\theta_2^{r_2}\theta_3^{r_3}\left[\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\right. \\ & \exp\left\{-\theta_1\left(nt-rt-it+\sum_{k=1}^{r_1}x_{1k}\right)\right\} \\ & \exp\left\{-\theta_2\left(it-jt+\sum_{k=1}^{r_2}x_{2k}\right)\right\}\exp\left\{-\theta_3\left(jt+\sum_{k=1}^{r_3}x_{3k}\right)\right\} \\ & \left. p_1^{n-r-i+r_1}p_2^{i-j+r_2}(1-p_1-p_2)^{j+r_3}\right], \end{aligned} \tag{3.2}$$

where $\boldsymbol{\psi}=(\theta_1,\ \theta_2,\ \theta_3,\ p_1,\ p_2)$ and $\mathbf{x}=(x_{11},...,\ x_{1r_1},\ x_{21},...,\ x_{2r_2},\ x_{31},...,\ x_{3r_3})$.

3.2. The posterior distribution using the UP. When no or little prior information is given, usually, the non-informative prior is assumed to be the UP. Bayes (1763), de Laplace (1820) and Geisser (1984) proposed that one may take the UP for the unknown parameters $\boldsymbol{\psi}=(\theta_1,\ \theta_2,\ \theta_3,\ p_1,\ p_2)$. Following Bayes (1763), de Laplace (1820) and Geisser (1984), UPs over the intervals $(0,\ \infty)$ and $(0,\ 1)$ are taken for the parameters ($\theta_1,\ \theta_2$ and θ_3) of Exponential distributions and for the mixing proportions (p_1 and p_2), respectively. With these settings, joint prior distribution of the parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2, as defined by Saleem (2010), is given by:

$$\pi_1(\boldsymbol{\psi}) \propto 1;\ \theta_1,\ \theta_2,\ \theta_3 > 0,\ p_1,\ p_2 \geq 0,\ p_1+p_2 \leq 1. \tag{3.3}$$

The joint posterior distribution of parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2 given data $\mathbf{x}$, using the UP is defined as:

$$g_1(\boldsymbol{\psi}\,|\mathbf{x}) = \frac{L(\boldsymbol{\psi}\,|\mathbf{x})\,\pi_1(\boldsymbol{\psi})}{\int_{\boldsymbol{\psi}} L(\boldsymbol{\psi}\,|\mathbf{x})\,\pi_1(\boldsymbol{\psi})\,d\boldsymbol{\psi}} \tag{3.4}$$

$$\begin{aligned} g_1(\boldsymbol{\psi}\,|\mathbf{x}) =\ & \frac{1}{E_1\theta_1^{1-A_{11}}\theta_2^{1-A_{21}}\theta_3^{1-A_{31}}}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\exp(-B_{11}\theta_1)\times \\ & \exp(-B_{21}\theta_2)\exp(-B_{31}\theta_3)\,p_1^{A_{01}-1}p_2^{B_{01}-1}(1-p_1-p_2)^{C_{01}-1}, \end{aligned} \tag{3.5}$$

where $A_{11}=r_1+1,\ A_{21}=r_2+1,\ A_{31}=r_3+1,\ B_{11}=nt-rt-it+\sum_{k=1}^{r_1}x_{1k}$, $B_{21}=it-jt+\sum_{k=1}^{r_2}x_{2k}$, $B_{31}=jt+\sum_{k=1}^{r_3}x_{3k}$, $A_{01}=n-r-i+r_1+1, B_{01}=i-j+r_2+1, C_{01}=j+r_3+1,\ E_1=\Gamma(A_{11})\Gamma(A_{21})\Gamma(A_{31})\ \sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{\mathrm{j}}\times$ $B(A_{01},\ B_{01},\ C_{01})\ \mathrm{B}_{11}^{-A_{11}}\ \mathrm{B}_{21}^{-A_{21}}\ \mathrm{B}_{31}^{-A_{31}}$.

3.3. The posterior distribution using the JP. According to Jeffreys' (1946, 1961), Bernardo (1979) and Berger (1985), the JP is defined as $p(\theta_m) \propto \sqrt{|I(\theta_m)|}$, $m=1,\ 2,\ 3$, where $I(\theta_m)=-E\left[\frac{\partial^2 f(x|\theta_m)}{\partial\theta_m^2}\right]$ is the Fisher's information matrix. The prior distributions of the mixing proportions p_1 and p_2 are again taken to be the uniform on over the interval $(0,\ 1)$. The joint prior distribution of parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2 is (cf. Sinha, 1998) given by:

$$\pi_2(\psi) \propto \frac{1}{\theta_1\ \theta_2\ \theta_3},\ \theta_1,\ \theta_2,\ \theta_3 > 0,\ p_1,\ p_2 \geq 0,\ p_1+p_2 \leq 1 \tag{3.6}$$

The joint posterior distribution of parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2 given data $\mathbf{x}$, using the JP is:

$$g_2\left(\boldsymbol{\psi}\,|\mathbf{x}\right)=\frac{L\left(\boldsymbol{\psi}\,|\mathbf{x}\right)\pi_2\left(\boldsymbol{\psi}\right)}{\int_{\boldsymbol{\psi}}L\left(\boldsymbol{\psi}\,|\mathbf{x}\right)\pi_2\left(\boldsymbol{\psi}\right)d\boldsymbol{\psi}} \tag{3.7}$$

$$g_2\left(\boldsymbol{\psi}\,|\mathbf{x}\right) = \frac{\sum\limits_{i=0}^{n-r}\sum\limits_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\exp\left(-B_{12}\theta_1\right)}{E_2\theta_1^{1-A_{12}}\theta_2^{1-A_{22}}\theta_3^{1-A_{32}}}\times \exp\left(-B_{22}\theta_2\right)\exp\left(-B_{32}\theta_3\right)p_1^{A_{02}-1}p_2^{B_{02}-1}(1-p_1-p_2)^{C_{02}-1}, \tag{3.8}$$

where $A_{12}=r_1,\ A_{22}=r_2,\ A_{32}=r_3,\ B_{12}=nt-rt-it+\sum_{k=1}^{r_1}x_{1k},\ B_{22}=it-jt+\sum_{k=1}^{r_2}x_{2k},\ B_{32}=jt+\sum_{k=1}^{r_3}x_{3k}, A_{02}=n-r-i+r_1+1, B_{02}=i-j+r_2+1, C_{02}=j+r_3+1,$ $E_2=\Gamma\left(A_{12}\right)\Gamma\left(A_{22}\right)\Gamma\left(A_{32}\right)\ \sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times\ B\left(A_{02},\ B_{02},\ C_{02}\right)\ \mathrm{B}_{12}^{-A_{12}}\ \mathrm{B}_{22}^{-A_{22}}\ \mathrm{B}_{32}^{-A_{32}}.$

3.4. The posterior distribution using the IP. As an informative prior distribution, we take Gamma distribution for component parameters $\theta_1,\ \theta_2,\ \theta_3$ and bivariate beta distribution for proportion parameters $p_1,\ p_2$, i.e.

$$\pi_4\left(\theta_1;\ a_1,\ b_1\right)=\frac{b_1^{a_1}}{\Gamma\left(a_1\right)}\ \theta_1^{a_1-1}\exp\left(-b_1\ \theta_1\right),\ \theta_1>0,\ a_1,\ b_1>0 \tag{3.9}$$

$$\pi_5\left(\theta_2;\ a_2,\ b_2\right)=\frac{b_2^{a_2}}{\Gamma\left(a_2\right)}\ \theta_2^{a_2-1}\exp\left(-b_2\ \theta_2\right),\ \theta_2>0,\ a_2,\ b_2>0 \tag{3.10}$$

$$\pi_6\left(\theta_3;\ a_3,\ b_3\right)=\frac{b_3^{a_3}}{\Gamma\left(a_3\right)}\ \theta_3^{a_3-1}\exp\left(-b_3\ \theta_3\right),\ \theta_3>0,\ a_3,\ b_3>0 \tag{3.11}$$

$$\pi_7\left(p_1,\ p_2;\ a,\ b,\ c\right)=\frac{1}{B\left(a,\ b,\ c\right)}\ p_1^{a-1}\ p_2^{b-1}\ \left(1-p_1-p_2\right)^{c-1}, \quad p_1,\ \ p_2\geq 0,\ p_1+p_2\leq 1,\ a,\ b,\ c>0. \tag{3.12}$$

So, the joint prior distribution of parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2 using the IP is

$$\pi_3\left(\boldsymbol{\psi}\right)\propto\theta_1^{a_1-1}\exp\left(-b_1\theta_1\right)\theta_2^{a_2-1}\exp\left(-b_2\theta_2\right)\theta_3^{a_3-1}\times \exp\left(-b_3\theta_3\right)p_1^{a-1}p_2^{b-1}\left(1-p_1-p_2\right)^{c-1} \tag{3.13}$$

The joint posterior distribution of parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2 given data $\mathbf{x}$, using the IP is:

$$g_3\left(\boldsymbol{\psi}\,|\mathbf{x}\right)=\frac{L\left(\boldsymbol{\psi}\,|\mathbf{x}\right)\pi_3\left(\boldsymbol{\psi}\right)}{\int_{\boldsymbol{\psi}}L\left(\boldsymbol{\psi}\,|\mathbf{x}\right)\pi_3\left(\boldsymbol{\psi}\right)d\boldsymbol{\psi}} \tag{3.14}$$

$$g_3\left(\boldsymbol{\psi}\,|\mathbf{x}\right) = \frac{\sum\limits_{i=0}^{n-r}\sum\limits_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\exp\left(-B_{13}\theta_1\right)\exp\left(-B_{23}\theta_2\right)}{E_3\theta_1^{1-A_{13}}\theta_2^{1-A_{23}}\theta_3^{1-A_{33}}}\times \exp\left(-B_{33}\theta_3\right)p_1^{A_{03}-1}p_2^{B_{03}-1}(1-p_1-p_2)^{C_{03}-1}, \tag{3.15}$$

where $A_{13}=r_1+a_1,\ A_{23}=r_2+a_2,\ A_{33}=r_3+a_3,\ B_{13}=nt-rt-it+\sum_{k=1}^{r_1}x_{1k}+b_1,$ $B_{23}=it-jt+\sum_{k=1}^{r_2}x_{2k}+b_2,\ B_{33}=jt+\sum_{k=1}^{r_3}x_{3k}+b_3, A_{03}=n-r-i+r_1+a, B_{03}=i-j+r_2+b, C_{03}=j+r_3+c,\ E_3=\Gamma\left(A_{13}\right)\Gamma\left(A_{23}\right)\Gamma\left(A_{33}\right)\sum\limits_{i=0}^{n-r}\sum\limits_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times$ $B\left(A_{03},B_{03},C_{03}\right)\mathrm{B}_{13}^{-A_{13}}\mathrm{B}_{23}^{-A_{23}}\mathrm{B}_{33}^{-A_{33}}.$

4. The Bayes estimators and posterior risks using the UP, the JP and IP under SELF

If $L(\theta,\ d)$ is a loss function then the expected value of the loss function for a given decision with respect to the posterior distribution is posterior risk function and if $\hat{d}$ is a Bayes estimator then $\rho\left(\hat{d}\right)$ is called posterior risk and is given by $\rho\left(\hat{d}\right) = E_{\theta|\mathbf{x}}\left\{L\left(\theta,\ \hat{d}\right)\right\}$. The SELF is suggested by Legendre (1806) and is defined as: $L(\theta, d) = (\theta-d)^2$. The Bayes estimator and posterior risk under SELF are: $\hat{d} = E_{\theta|\mathbf{x}}(\theta)$ and $\rho\left(\hat{d}\right) = E_{\theta|\mathbf{x}}(\theta^2) - \left\{E_{\theta|\mathbf{x}}(\theta)\right\}^2$, respectively. So, the Bayes estimators and posterior risks using the UP, the JP and IP for parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2 under SELF are obtained with their respective marginal posterior distributions as given below:

$$\hat{\theta}_{1v} = \frac{\Gamma(A_{1v}+1)\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\begin{pmatrix} n-r \\ i \end{pmatrix}\begin{pmatrix} i \\ \mathrm{j} \end{pmatrix}\times$$

$$\mathrm{B}_{1v}^{-(A_{1v}+1)}\ \mathrm{B}_{2v}^{-A_{2v}}\ \mathrm{B}_{3v}^{-A_{3v}}\ B(A_{0v},\ C_{0v})\ B(B_{0v},\ A_{0v}+C_{0v}) \tag{4.1}$$

$$\hat{\theta}_{2v} = \frac{\Gamma(A_{1v})\Gamma(A_{2v}+1)\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\begin{pmatrix} n-r \\ i \end{pmatrix}\begin{pmatrix} i \\ \mathrm{j} \end{pmatrix}\times$$

$$\mathrm{B}_{1v}^{-A_{1v}}\ \mathrm{B}_{2v}^{-(A_{2v}+1)}\ \mathrm{B}_{3v}^{-A_{3v}}\ B(A_{0v},\ C_{0v})\ B(B_{0v},\ A_{0v}+C_{0v}) \tag{4.2}$$

$$\hat{\theta}_{3v} = \frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+1)}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\begin{pmatrix} n-r \\ i \end{pmatrix}\begin{pmatrix} i \\ \mathrm{j} \end{pmatrix}\times$$

$$\mathrm{B}_{1v}^{-A_{1v}}\ \mathrm{B}_{2v}^{-A_{2v}}\ \mathrm{B}_{3v}^{-(A_{3v}+1)}\ B(A_{0v},\ C_{0v})\ B(B_{0v},\ A_{0v}+C_{0v}) \tag{4.3}$$

$$\hat{p}_{1v} = \frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\begin{pmatrix} n-r \\ i \end{pmatrix}\begin{pmatrix} i \\ \mathrm{j} \end{pmatrix}\times$$

$$\mathrm{B}_{1v}^{-A_{1v}}\ \mathrm{B}_{2v}^{-A_{2v}}\ \mathrm{B}_{3v}^{-A_{3v}}\ B(B_{0v},\ C_{0v})\ B(A_{0v}+1,\ B_{0v}+C_{0v}) \tag{4.4}$$

$$\hat{p}_{2v} = \frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\begin{pmatrix} n-r \\ i \end{pmatrix}\begin{pmatrix} i \\ \mathrm{j} \end{pmatrix}\times$$

$$\mathrm{B}_{1v}^{-A_{1v}}\ \mathrm{B}_{2v}^{-A_{2v}}\ \mathrm{B}_{3v}^{-A_{3v}}\ B(A_{0v},\ C_{0v})\ B(B_{0v}+1,\ A_{0v}+C_{0v}) \tag{4.5}$$

$$\rho\left(\hat{\theta}_{1v}\right) = \frac{\Gamma(A_{1v}+2)\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times$$

$$\mathrm{B}_{1v}^{-(A_{1v}+2)}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})\,B(B_{0v},A_{0v}+C_{0v}) -$$

$$\left\{\begin{array}{l} \frac{\Gamma(A_{1v}+1)\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum\limits_{i=0}^{n-r}\sum\limits_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-(A_{1v}+1)}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})\,B(B_{0v},A_{0v}+C_{0v}) \end{array}\right\}^2 \tag{4.6}$$

$$\rho\left(\hat{\theta}_{2v}\right) = \frac{\Gamma(A_{1v})\Gamma(A_{2v}+2)\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times$$

$$\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-(A_{2v}+2)}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})\,B(B_{0v},A_{0v}+C_{0v}) -$$

$$\left\{\begin{array}{l} \frac{\Gamma(A_{1v})\Gamma(A_{2v}+1)\Gamma(A_{3v})}{E_v}\sum\limits_{i=0}^{n-r}\sum\limits_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-(A_{2v}+1)}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})\,B(B_{0v},A_{0v}+C_{0v}) \end{array}\right\}^2 \tag{4.7}$$

$$\rho\left(\hat{\theta}_{3v}\right) = \frac{\Gamma\left(A_{1v}\right)\Gamma\left(A_{2v}\right)\Gamma\left(A_{3v}+2\right)}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times$$

$$\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+2)}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)-$$

$$\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+1)}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+1)}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}^2 \tag{4.8}$$

$$\rho\left(\hat{p}_{1v}\right) = \frac{\Gamma\left(A_{1v}\right)\Gamma\left(A_{2v}\right)\Gamma\left(A_{3v}\right)}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times$$

$$\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(B_{0v},C_{0v}\right)B\left(A_{0v}+2,B_{0v}+C_{0v}\right)-$$

$$\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(B_{0v},C_{0v}\right)B\left(A_{0v}+1,B_{0v}+C_{0v}\right)\end{array}\right\}^2 \tag{4.9}$$

$$\rho\left(\hat{p}_{2v}\right) = \frac{\Gamma\left(A_{1v}\right)\Gamma\left(A_{2v}\right)\Gamma\left(A_{3v}\right)}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times$$

$$\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v}+2,A_{0v}+C_{0v}\right)-$$

$$\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v}+1,A_{0v}+C_{0v}\right)\end{array}\right\}^2, \tag{4.10}$$

where $v = 1$ for the UP, $v = 2$ for the JP and $v = 3$ for the IP.

5. The Bayes estimators and posterior risks using the UP, the JP and IP under PLF

Norstrom (1996) discussed an asymmetric PLF and a special case of general class of PLFs is $L(\theta,\ d) = \frac{(\theta-d)^2}{d}$. The Bayes estimator and posterior risk are: $\hat{d} = \left\{E_{\theta|\mathbf{x}}\left(\theta^2\right)\right\}^{\frac{1}{2}}$ and $\rho\left(\hat{d}\right) = 2\left\{E_{\theta|\mathbf{x}}\left(\theta^2\right)\right\}^{\frac{1}{2}} - 2E_{\theta|\mathbf{x}}\left(\theta\right)$, respectively. The respective marginal posterior distributions yield the Bayes estimators and posterior risks using the UP, the JP and the IP for parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2 under PLF as:

$$\hat{\theta}_{1v} = \left\{\begin{array}{l}\frac{\Gamma(A_{1v}+2)\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-(A_{1v}+2)}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}} \tag{5.1}$$

$$\hat{\theta}_{2v} = \left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v}+2)\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-(A_{2v}+2)}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}} \tag{5.2}$$

$$\hat{\theta}_{3v} = \left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+2)}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+2)}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}} \tag{5.3}$$

$$\hat{p}_{1v} = \left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times \\ \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(B_{0v},C_{0v}\right)B\left(A_{0v}+2,B_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}} \tag{5.4}$$

$$\hat{p}_{2v}=\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\times\\ \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v}+2,A_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}} \tag{5.5}$$

$$\begin{aligned}\rho\left(\hat{\theta}_{1v}\right)&=2\left\{\begin{array}{l}\frac{\Gamma(A_{1v}+2)\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-(A_{1v}+2)}\times\\ \mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}}\\ &-2\left\{\begin{array}{l}\frac{\Gamma(A_{1v}+1)\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-(A_{1v}+1)}\times\\ \mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}\end{aligned} \tag{5.6}$$

$$\begin{aligned}\rho\left(\hat{\theta}_{2v}\right)&=2\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v}+2)\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\times\\ \mathrm{B}_{2v}^{-(A_{2v}+2)}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}}\\ &-2\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v}+1)\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\times\\ \mathrm{B}_{2v}^{-(A_{2v}+1)}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}\end{aligned} \tag{5.7}$$

$$\begin{aligned}\rho\left(\hat{\theta}_{3v}\right)&=2\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+2)}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\times\\ \mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+2)}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}}\\ &-2\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+1)}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\times\\ \mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+1)}B\left(A_{0v},C_{0v}\right)B\left(B_{0v},A_{0v}+C_{0v}\right)\end{array}\right\}\end{aligned} \tag{5.8}$$

$$\begin{aligned}\rho\left(\hat{p}_{1v}\right)&=2\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\times\\ \mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(B_{0v},C_{0v}\right)B\left(A_{0v}+2,B_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}}\\ &-2\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\times\\ \mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(B_{0v},C_{0v}\right)B\left(A_{0v}+1,B_{0v}+C_{0v}\right)\end{array}\right\}\end{aligned} \tag{5.9}$$

$$\begin{aligned}\rho\left(\hat{p}_{2v}\right)&=2\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\times\\ \mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v}+2,A_{0v}+C_{0v}\right)\end{array}\right\}^{\frac{1}{2}}\\ &-2\left\{\begin{array}{l}\frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{E_v}\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\times\\ \mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B\left(A_{0v},C_{0v}\right)B\left(B_{0v}+1,A_{0v}+C_{0v}\right)\end{array}\right\}\end{aligned} \tag{5.10}$$

6. The Bayes estimators and posterior risks using the UP, the JP and the IP under DLF

DeGroot (2005) introduced the asymmetric loss function, $L(\theta,\ d)=\left(\frac{\theta-d}{d}\right)^2$, known as DLF.The Bayes estimator and its posterior risk under DLF are: $\hat{d}=\frac{E_{\theta|\mathbf{x}}\left(\theta^2\right)}{E_{\theta|\mathbf{x}}(\theta)}$ and

$\rho\left(\hat{d}\right) = 1 - \frac{\left\{E_{\theta|\mathbf{x}}(\theta)\right\}^2}{E_{\theta|\mathbf{x}}\left(\theta^2\right)}$, respectively. The Bayes estimators and posterior risks using the UP, the JP and the IP for parameters $\theta_1,\ \theta_2,\ \theta_3,\ p_1$ and p_2 under DLF are:

$$\hat{\theta}_{1v} = \frac{\Gamma(A_{1v}+2)\Gamma(A_{2v})\Gamma(A_{3v})}{\Gamma(A_{1v}+1)\Gamma(A_{2v})\Gamma(A_{3v})}\times \frac{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-(A_{1v}+2)}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-(A_{1v}+1)}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})} \tag{6.1}$$

$$\hat{\theta}_{2v} = \frac{\Gamma(A_{1v})\Gamma(A_{2v}+2)\Gamma(A_{3v})}{\Gamma(A_{1v})\Gamma(A_{2v}+1)\Gamma(A_{3v})}\times \frac{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-(A_{2v}+2)}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-(A_{2v}+1)}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})} \tag{6.2}$$

$$\hat{\theta}_{3v} = \frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+2)}{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+1)}\times \frac{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+2)}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+1)}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})} \tag{6.3}$$

$$\hat{p}_{1v} = \frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}\times \frac{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(B_{0v},C_{0v})B(A_{0v}+2,B_{0v}+C_{0v})}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(B_{0v},C_{0v})B(A_{0v}+1,B_{0v}+C_{0v})} \tag{6.4}$$

$$\hat{p}_{2v} = \frac{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})}\times \frac{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v}+2,A_{0v}+C_{0v})}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v}+1,A_{0v}+C_{0v})} \tag{6.5}$$

$$\rho\left(\hat{\theta}_{1v}\right) = 1 - \frac{\{\Gamma(A_{1v}+1)\Gamma(A_{2v})\Gamma(A_{3v})\}^2}{E_v\Gamma(A_{1v}+2)\Gamma(A_{2v})\Gamma(A_{3v})}\times \frac{\left\{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-(A_{1v}+1)}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})\right\}^2}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-(A_{1v}+2)}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})} \tag{6.6}$$

$$\rho\left(\hat{\theta}_{2v}\right) = 1 - \frac{\{\Gamma(A_{1v})\Gamma(A_{2v}+1)\Gamma(A_{3v})\}^2}{E_v\Gamma(A_{1v})\Gamma(A_{2v}+2)\Gamma(A_{3v})}\times \frac{\left\{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-(A_{2v}+1)}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})\right\}^2}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j}\mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-(A_{2v}+2)}\mathrm{B}_{3v}^{-A_{3v}}B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})} \tag{6.7}$$

$$\rho\left(\hat{\theta}_{3v}\right) = 1 - \frac{\{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+1)\}^2}{E_v\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v}+2)} \times \tag{6.8}$$
$$\frac{\left\{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j} \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+1)} B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})\right\}^2}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j} \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-(A_{3v}+2)} B(A_{0v},C_{0v})B(B_{0v},A_{0v}+C_{0v})}$$

$$\rho\left(\hat{p}_{1v}\right) = 1 - \frac{\{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})\}^2}{E_v\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})} \times \tag{6.9}$$
$$\frac{\left\{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j} \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}} B(B_{0v},C_{0v})B(A_{0v}+1,B_{0v}+C_{02})\right\}^2}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j} \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}} B(B_{0v},C_{0v})B(A_{0v}+2,B_{0v}+C_{02})}$$

$$\rho\left(\hat{p}_{2v}\right) = 1 - \frac{\{\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})\}^2}{E_v\Gamma(A_{1v})\Gamma(A_{2v})\Gamma(A_{3v})} \times \tag{6.10}$$
$$\frac{\left\{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j} \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}} B(A_{0v},C_{0v})B(B_{0v}+1,A_{0v}+C_{0v})\right\}^2}{\sum_{i=0}^{n-r}\sum_{j=0}^{i}\binom{n-r}{i}\binom{i}{j} \mathrm{B}_{1v}^{-A_{1v}}\mathrm{B}_{2v}^{-A_{2v}}\mathrm{B}_{3v}^{-A_{3v}} B(A_{0v},C_{0v})B(B_{0v}+2,A_{0v}+C_{0v})}.$$

7. Elicitation of hyperparameters

Elicitation is a tool used to quantify a person's belief and knowledge about the parameter(s) of interest. In Bayesian perspective, elicitation, most often, arises as a method for specifying the prior distribution of the random parameter(s). Elicitation is simply the quantification of prior knowledge about the random parameter(s) so that this can then be combined with the likelihood to obtain posterior distribution for further statistical analysis. Elicitation has remained a challenging problem for the statistician.Authors who have discussed this problem include Kadane et al. (1980), Birch and Bartollucci (1983), Chaloner and Duncan (1983), Gavasakar (1988), Al-Awadhi and Gartwaite (1998), Aslam (2003), Hahn (2006), Saleem and Aslam (2008) and references cited therein. In this study, we adopted a method based on predictive probabilities, given by Aslam (2003).

For eliciting the hyperparameters, prior predictive distribution (PPD) is used. The PPD for a random variable X is:

$$p\left(x\right) = \int_{\boldsymbol{\psi}} p\left(x\left|\boldsymbol{\psi}\right.\right)\pi_3\left(\boldsymbol{\psi}\right)d\boldsymbol{\psi} \tag{7.1}$$

$$p\left(x\right) = \frac{1}{(a+b+c)}\left[\frac{a\;a_1\;b_1^{a_1}}{(b_1+x)^{a_1+1}} + \frac{b\;a_2\;b_2^{a_2}}{(b_2+x)^{a_2+1}} + \frac{c\;a_3\;b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]. \tag{7.2}$$

We choose the prior predictive probabilities, satisfying the laws of probability, to elicit the hyperparameters of the prior density. By following these laws of probability, some minor inconsistencies may arise which are expected to be ignorable. Using the prior predictive distribution given in (7.2) we consider nine intervals (0, 1), (1, 2), (2, 3), (3, 4), (4, 5), (5, 6), (6, 7), (7, 8) and (8, 9) with probabilities 0.57, 0.20, 0.10, 0.05, 0.02, 0.015, 0.01, 0.005 and 0.003, respectively, given as expert opinion. The following nine equations are derived from the given information using the (7.2) as:

$$\frac{1}{(a+b+c)}\int_0^1\left[\frac{a\;a_1\;b_1^{a_1}}{(b_1+x)^{a_1+1}} + \frac{b\;a_2\;b_2^{a_2}}{(b_2+x)^{a_2+1}} + \frac{c\;a_3\;b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx = 0.57 \tag{7.3}$$

$$(7.4)\qquad \frac{1}{(a+b+c)}\int_1^2\left[\frac{a\ a_1\ b_1^{a_1}}{(b_1+x)^{a_1+1}}+\frac{b\ a_2\ b_2^{a_2}}{(b_2+x)^{a_2+1}}+\frac{c\ a_3\ b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx=0.20$$

$$(7.5)\qquad \frac{1}{(a+b+c)}\int_2^3\left[\frac{a\ a_1\ b_1^{a_1}}{(b_1+x)^{a_1+1}}+\frac{b\ a_2\ b_2^{a_2}}{(b_2+x)^{a_2+1}}+\frac{c\ a_3\ b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx=0.10$$

$$(7.6)\qquad \frac{1}{(a+b+c)}\int_3^4\left[\frac{a\ a_1\ b_1^{a_1}}{(b_1+x)^{a_1+1}}+\frac{b\ a_2\ b_2^{a_2}}{(b_2+x)^{a_2+1}}+\frac{c\ a_3\ b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx=0.05$$

$$(7.7)\qquad \frac{1}{(a+b+c)}\int_4^5\left[\frac{a\ a_1\ b_1^{a_1}}{(b_1+x)^{a_1+1}}+\frac{b\ a_2\ b_2^{a_2}}{(b_2+x)^{a_2+1}}+\frac{c\ a_3\ b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx=0.02$$

$$(7.8)\qquad \frac{1}{(a+b+c)}\int_5^6\left[\frac{a\ a_1\ b_1^{a_1}}{(b_1+x)^{a_1+1}}+\frac{b\ a_2\ b_2^{a_2}}{(b_2+x)^{a_2+1}}+\frac{c\ a_3\ b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx=0.015$$

$$(7.9)\qquad \frac{1}{(a+b+c)}\int_6^7\left[\frac{a\ a_1\ b_1^{a_1}}{(b_1+x)^{a_1+1}}+\frac{b\ a_2\ b_2^{a_2}}{(b_2+x)^{a_2+1}}+\frac{c\ a_3\ b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx=0.01$$

$$(7.10)\qquad \frac{1}{(a+b+c)}\int_7^8\left[\frac{a\ a_1\ b_1^{a_1}}{(b_1+x)^{a_1+1}}+\frac{b\ a_2\ b_2^{a_2}}{(b_2+x)^{a_2+1}}+\frac{c\ a_3\ b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx=0.005$$

$$(7.11)\qquad \frac{1}{(a+b+c)}\int_8^9\left[\frac{a\ a_1\ b_1^{a_1}}{(b_1+x)^{a_1+1}}+\frac{b\ a_2\ b_2^{a_2}}{(b_2+x)^{a_2+1}}+\frac{c\ a_3\ b_3^{a_3}}{(b_3+x)^{a_3+1}}\right]dx=0.003$$

The above nine equations (7.3-7.11) are solved simultaneously by using Mathematica software for eliciting the hyperparameters (a_1, b_1, a_2, b_2, a_3, b_3, a, b, c). Through this criteria, the values of the hyperparameters are obtained as (3.8330, 3.7310, 3.3570, 3.1360, 2.9030, 2.7330, 3.0280, 0.6995, 2.7350).

8. The limiting expressions

When t tends to ∞, r tends to n and r_l tends to n_l, $l = 1, 2, 3$, then all the values which are censored become uncensored in our analysis. So, the information contained in the sample is increased. Consequently, the posterior risks of the Bayes estimates diminish. The efficiency of the Bayes estimates is increased because all the values are incorporated in our sample. The limiting (complete sample) expressions for Bayes estimators and posterior risks using the UP, the JP and the IP under SELF, PLF and DLF are given in the Tables 1-6.

Table 1. Limiting Expressions for the Bayes Estimators as $t \to \infty$ using the UP, the JP and the IP under SELF

	Bayes Estimators		
Parameters	UP	JP	IP
θ_1	$\frac{n_1+1}{\sum_{k=1}^{n_1} x_{1k}}$	$\frac{n_1}{\sum_{k=1}^{n_1} x_{1k}}$	$\frac{n_1+a_1}{\sum_{k=1}^{n_1} x_{1k}+b_1}$
θ_2	$\frac{n_2+1}{\sum_{k=1}^{n_2} x_{2k}}$	$\frac{n_2}{\sum_{k=1}^{n_2} x_{2k}}$	$\frac{n_2+a_2}{\sum_{k=1}^{n_2} x_{2k}+b_2}$
θ_3	$\frac{n_3+1}{\sum_{k=1}^{n_3} x_{3k}}$	$\frac{n_3}{\sum_{k=1}^{n_3} x_{3k}}$	$\frac{n_3+a_3}{\sum_{k=1}^{n_3} x_{3k}+b_3}$
p_1	$\frac{n_1+1}{n+3}$	$\frac{n_1+1}{n+3}$	$\frac{n_1+a}{n+a+b+c}$
p_2	$\frac{n_2+1}{n+3}$	$\frac{n_2+1}{n+3}$	$\frac{n_2+b}{n+a+b+c}$

Table 2. Limiting Expressions for the Posterior Risks as $t \to \infty$ using the UP, the JP and the IP under SELF

	Posterior Risks		
Parameters	UP	JP	IP
θ_1	$\frac{n_1+1}{\left(\sum_{k=1}^{n_1} x_{1k}\right)^2}$	$\frac{n_1}{\left(\sum_{k=1}^{n_1} x_{1k}\right)^2}$	$\frac{n_1+a_1}{\left(\sum_{k=1}^{n_1} x_{1k}+b_1\right)^2}$
θ_2	$\frac{n_2+1}{\left(\sum_{k=1}^{n_2} x_{2k}\right)^2}$	$\frac{n_2}{\left(\sum_{k=1}^{n_2} x_{2k}\right)^2}$	$\frac{n_2+a_2}{\left(\sum_{k=1}^{n_2} x_{2k}+b_2\right)^2}$
θ_3	$\frac{n_3+1}{\left(\sum_{k=1}^{n_3} x_{3k}\right)^2}$	$\frac{n_3}{\left(\sum_{k=1}^{n_3} x_{3k}\right)^2}$	$\frac{n_3+a_3}{\left(\sum_{k=1}^{n_3} x_{3k}+b_3\right)^2}$
p_1	$\frac{(n_1+1)(n_2+n_3+2)}{(n+3)^2(n+4)}$	$\frac{(n_1+1)(n_2+n_3+2)}{(n+3)^2(n+4)}$	$\frac{(n_1+a)(n_2+n_3+b+c)}{(n+a+b+c)^2(n+a+b+c+1)}$
p_2	$\frac{(n_2+1)(n_1+n_3+2)}{(n+3)^2(n+4)}$	$\frac{(n_2+1)(n_1+n_3+2)}{(n+3)^2(n+4)}$	$\frac{(n_2+b)(n_1+n_3+a+c)}{(n+a+b+c)^2(n+a+b+c+1)}$

Table 3. Limiting expressions for the Bayes estimators as $t \to \infty$ using the UP, the JP and the IP under PLF

	Bayes Estimators		
Parameters	UP	JP	IP
θ_1	$\frac{(n_1+1)^{1/2}(n_1+2)^{1/2}}{\left(\sum_{k=1}^{n_1} x_{1k}\right)^{1/2}}$	$\frac{(n_1)^{1/2}(n_1+1)^{1/2}}{\left(\sum_{k=1}^{n_1} x_{1k}\right)^{1/2}}$	$\frac{(n_1+a_1)^{1/2}(n_1+a_1+1)^{1/2}}{\left(\sum_{k=1}^{n_1} x_{1k}+b_1\right)^{1/2}}$
θ_2	$\frac{(n_2+1)^{1/2}(n_2+2)^{1/2}}{\left(\sum_{k=1}^{n_2} x_{2k}\right)^{1/2}}$	$\frac{(n_2)^{1/2}(n_2+1)^{1/2}}{\left(\sum_{k=1}^{n_2} x_{2k}\right)^{1/2}}$	$\frac{(n_2+a_2)^{1/2}(n_2+a_2+1)^{1/2}}{\left(\sum_{k=1}^{n_2} x_{2k}+b_2\right)^{1/2}}$
θ_3	$\frac{(n_3+1)^{1/2}(n_3+2)^{1/2}}{\left(\sum_{k=1}^{n_3} x_{3k}\right)^{1/2}}$	$\frac{(n_3)^{1/2}(n_3+1)^{1/2}}{\left(\sum_{k=1}^{n_3} x_{3k}\right)^{1/2}}$	$\frac{(n_3+a_3)^{1/2}(n_3+a_3+1)^{1/2}}{\left(\sum_{k=1}^{n_3} x_{3k}+b_3\right)^{1/2}}$
p_1	$\frac{(n_1+1)^{1/2}(n_1+2)^{1/2}}{(n+3)^{1/2}(n+4)^{1/2}}$	$\frac{(n_1+1)^{1/2}(n_1+2)^{1/2}}{(n+3)^{1/2}(n+4)^{1/2}}$	$\frac{(n_1+a)^{1/2}(n_1+a+1)^{1/2}}{(n+a+b+c)^{1/2}(n+a+b+c+1)^{1/2}}$
p_2	$\frac{(n_2+1)^{1/2}(n_2+2)^{1/2}}{(n+3)^{1/2}(n+4)^{1/2}}$	$\frac{(n_2+1)^{1/2}(n_2+2)^{1/2}}{(n+3)^{1/2}(n+4)^{1/2}}$	$\frac{(n_2+b)^{1/2}(n_2+b+1)^{1/2}}{(n+a+b+c)^{1/2}(n+a+b+c+1)^{1/2}}$

Table 4. Limiting expressions for the posterior risks as $t \to \infty$ using the UP, the JP and the IP under PLF

	Posterior Risks		
Parameters	UP	JP	IP
θ_1	$\frac{2(n_1+1)}{\sum_{k=1}^{n_1} x_{1k}}\left\{\frac{(n_1+2)^{1/2}}{(n_1+1)^{1/2}}-1\right\}$	$\frac{2n_1}{\sum_{k=1}^{n_1} x_{1k}}\left\{\frac{(n_1+1)^{1/2}}{(n_1)^{1/2}}-1\right\}$	$\frac{2(n_1+a_1)}{\left(\sum_{k=1}^{n_1} x_{1k}+b_1\right)}\left\{\frac{(n_1+a_1+1)^{1/2}}{(n_1+a_1)^{1/2}}-1\right\}$
θ_2	$\frac{2(n_2+1)}{\sum_{k=1}^{n_2} x_{2k}}\left\{\frac{(n_2+2)^{1/2}}{(n_2+1)^{1/2}}-1\right\}$	$\frac{2n_2}{\sum_{k=1}^{n_2} x_{2k}}\left\{\frac{(n_2+1)^{1/2}}{(n_2)^{1/2}}-1\right\}$	$\frac{2(n_2+a_2)}{\left(\sum_{k=1}^{n_2} x_{2k}+b_2\right)}\left\{\frac{(n_2+a_2+1)^{1/2}}{(n_2+a_2)^{1/2}}-1\right\}$
θ_3	$\frac{2(n_3+1)}{\sum_{k=1}^{n_3} x_{3k}}\left\{\frac{(n_3+2)^{1/2}}{(n_3+1)^{1/2}}-1\right\}$	$\frac{2n_3}{\sum_{k=1}^{n_3} x_{3k}}\left\{\frac{(n_3+1)^{1/2}}{(n_3)^{1/2}}-1\right\}$	$\frac{2(n_3+a_3)}{\left(\sum_{k=1}^{n_3} x_{3k}+b_3\right)}\left\{\frac{(n_3+a_3+1)^{1/2}}{(n_3+a_3)^{1/2}}-1\right\}$
p_1	$\frac{2(n_1+1)}{(n+3)}\left\{\frac{\frac{(n_1+2)^{1/2}}{(n_1+1)^{1/2}}}{\frac{(n+4)^{1/2}}{(n+3)^{1/2}}}-1\right\}$	$\frac{2(n_1+1)}{(n+3)}\left\{\frac{\frac{(n_1+2)^{1/2}}{(n_1+1)^{1/2}}}{\frac{(n+4)^{1/2}}{(n+3)^{1/2}}}-1\right\}$	$\frac{2(n_1+a)}{(n+a+b+c)}\left\{\frac{\frac{(n_1+a+1)^{1/2}}{(n_1+a)^{1/2}}}{\frac{(n+a+b+c+1)^{1/2}}{(n+a+b+c)^{1/2}}}-1\right\}$
p_2	$\frac{2(n_2+1)}{(n+3)}\left\{\frac{\frac{(n_2+2)^{1/2}}{(n_2+1)^{1/2}}}{\frac{(n+4)^{1/2}}{(n+3)^{1/2}}}-1\right\}$	$\frac{2(n_2+1)}{(n+3)}\left\{\frac{\frac{(n_2+2)^{1/2}}{(n_2+1)^{1/2}}}{\frac{(n+4)^{1/2}}{(n+3)^{1/2}}}-1\right\}$	$\frac{2(n_2+b)}{(n+a+b+c)}\left\{\frac{\frac{(n_2+b+1)^{1/2}}{(n_2+b)^{1/2}}}{\frac{(n+a+b+c+1)^{1/2}}{(n+a+b+c)^{1/2}}}-1\right\}$

9. Simulation study

Simulation study is a flexible methodology to illustrate the properties of the Bayes estimates of the 3-component mixture of Exponential distributions using the UP, the JP and the IP under SELF, PLF and DLF in terms of different sample sizes and test termination times. The samples of different sizes $n = 30,\ 100,\ 200$ are generated from the 3-component mixture of Exponential distributions for each choice of the vector of the parameters $(\theta_1,\ \theta_2,\ \theta_3,\ p_1,\ p_2) = \{(4,\ 3,\ 2,\ 0.5,\ 0.3),\ (3,\ 3,\ 3,\ 0.4,\ 0.4),\ (2,\ 3,\ 4,\ 0.3,\ 0.5)\}$.

Table 5. Limiting expressions for the Bayes estimators as $t \to \infty$ using the UP, the JP and the IP under DLF

Parameters	Bayes Estimators		
	UP	JP	IP
θ_1	$\frac{n_1+2}{\sum_{k=1}^{n_1} x_{1k}}$	$\frac{n_1+1}{\sum_{k=1}^{n_1} x_{1k}}$	$\frac{n_1+a_1+1}{\sum_{k=1}^{n_1} x_{1k}+b_1}$
θ_2	$\frac{n_2+2}{\sum_{k=1}^{n_2} x_{2k}}$	$\frac{n_2+1}{\sum_{k=1}^{n_2} x_{2k}}$	$\frac{n_2+a_2+1}{\sum_{k=1}^{n_2} x_{2k}+b_2}$
θ_3	$\frac{n_3+2}{\sum_{k=1}^{n_3} x_{3k}}$	$\frac{n_3+1}{\sum_{k=1}^{n_3} x_{3k}}$	$\frac{n_3+a_3+1}{\sum_{k=1}^{n_3} x_{3k}+b_3}$
p_1	$\frac{n_1+2}{n+4}$	$\frac{n_1+2}{n+4}$	$\frac{n_1+a+1}{n+a+b+c+1}$
p_2	$\frac{n_2+2}{n+4}$	$\frac{n_2+2}{n+4}$	$\frac{n_2+b+1}{n+a+b+c+1}$

Table 6. Limiting expressions for the posterior risks as $t \to \infty$ using the UP, the JP and the IP under DLF

	Posterior Risks		
Parameters	UP	JP	IP
θ_1	$\frac{1}{n_1+2}$	$\frac{1}{n_1+1}$	$\frac{1}{n_1+a_1+1}$
θ_2	$\frac{1}{n_2+2}$	$\frac{1}{n_2+1}$	$\frac{1}{n_2+a_2+1}$
θ_3	$\frac{1}{n_3+2}$	$\frac{1}{n_3+1}$	$\frac{1}{n_3+a_3+1}$
p_1	$\frac{(n_2+n_3+2)}{(n_1+2)(n+3)}$	$\frac{(n_2+n_3+2)}{(n_1+2)(n+3)}$	$\frac{(n_2+n_3+b+c)}{(n_1+a+1)(n+a+b+c)}$
p_2	$\frac{(n_1+n_3+2)}{(n_2+2)(n+3)}$	$\frac{(n_1+n_3+2)}{(n_2+2)(n+3)}$	$\frac{(n_1+n_3+a+c)}{(n_2+b+1)(n+a+b+c)}$

The simulation is repeated 1000 times and the results are then averaged. Sample of sizes $p_1 n$, $p_2 n$ and $(1-p_1-p_2)\,n$ are chosen randomly from first component density $f_1\,(x;\ \theta_1)$, second component density $f_2\,(x;\ \theta_2)$ and third component density $f_3\,(x;\ \theta_3)$, respectively. To check the impact of test termination time on Bayes estimates, we estimate the parameters of the 3-component mixture of Exponential distributions based on a sample censored at fixed test termination times $t = 0.5,\ 0.8$. The observations which are greater than test termination time t are taken as censored. Only failures can be considered as members of subpopulation-1, subpopulation-2 or subpopulation-3 of the 3-component mixture of Exponential distributions. For the sake of brevity, simulated results only for $n = 30, 100, 200$ and $(\theta_1, \theta_2, \theta_3, p_1, p_2) = (4,\ 3,\ 2,\ 0.5,\ 0.3)$ are presented in the Tables 8-10 (see appendix). The simulated results for $(\theta_1,\ \theta_2,\ \theta_3,\ p_1,\ p_2) = \{(3,\ 3,\ 3,\ 0.4,\ 0.4)\,,\ (2,\ 3,\ 4,\ 0.3,\ 0.5)\}$ are available with the first author and can be obtained on demand.

From Tables 8-10 (see appendix), it can be seen that differences of Bayes estimates of component and proportion parameters from assumed parameters reduce with an increase in sample size at different test termination times and same is the case with large test termination time as compared to small test termination time for different sample sizes.Also, if $\theta_1 > \theta_2 > \theta_3$ and $p_1 > p_2$, first and second component parameters and second proportion parameter using the IP under SELF, PLF and DLF are under-estimated but third component and first proportion parameters are over-estimated at different sample sizes and test termination times with a few exceptions.By using the IP under SELF, PLF and DLF, three component parameters and second proportion parameter are under-estimated, however,first proportion parameter is over-estimated with a few exceptions in case of $\theta_1 = \theta_2 = \theta_3$ and $p_1 = p_2$. Also, if $\theta_1 < \theta_2 < \theta_3$ and $p_1 < p_2$, third component and second proportion parameters using the IP under SELF, PLF and DLF are under-estimated but there is a mixed pattern (over-estimation or under-estimation) for first and second component and first proportion parameters using the IP. Similarly, the component parameters using the UP and the JP under SELF, PLF and DLF are over-estimated but there is a mix pattern (under-estimation or over-estimation) for proportion parameters using the UP and the JP under SELF, PLF and DLF at different sample sizes and test termination times.

It is, also, clear from the Tables 8-10 that for a fixed test termination time, the posterior risks of the Bayes estimates, using the UP, the JP and the IP under SELF, PLF and DLF, reduce with an increase in sample size. On the other hand, for all priors, loss functions and sample sizes considered in this study, posterior risks decrease with an increase in test termination time.The posterior risks using the IP are smaller than the posterior risks using the UP and the JP for different sample sizes and test termination times.Also, the posterior risks using the JP are smaller than that using the UP for different sample sizes and test termination times. It is also observed that in estimating the component parameters θ_1, θ_2 and θ_3, posterior risks are smaller under DLF than under SELF and PLF at different sample sizes and test termination times considered in this study. However, for estimating the mixing proportions, SELF yields smaller posterior risks than SELF and DLF, at different sample sizes and test termination times. Thus, DLF is more suitable for estimating component parameters and SELF is a preferable choice for estimating proportion parameters p_1 and p_2.

10. Real data example

Davis (1952) reported a mixture data on lifetimes (in thousand hours) of many components used in aircraft sets. To illustrate the proposed methodology, we take the data on three components, namely, Transmitter Tube, Combination of Transformers and Combination of Relays. It is unknown that which component (Tubes, Transformers and Relays) fails until a failure (of a radar set) occurs at or before the test termination time $t = 0.4$. The total number of tests are conducted 702 times.For test termination time $t = 0.4$, the data are summarized as below. $n = 702$, $r_1 = 310$, $r_2 = 148$, $r_3 = 181$, $r = 639$, $n-r = 63$, $\sum_{k=1}^{r_1} x_{1k} = 36.875$, $\sum_{k=1}^{r_2} x_{2k} = 22.90$, $\sum_{k=1}^{r_3} x_{3k} = 19.125$. Since $n - r = 63$, we have almost 9 percent censored sample. Thus, this is a type-I right censored data. Bayes estimates and their posterior risks using the UP, the JP and the IP under SELF, PLF and DLF are showcased in Table 7 given below.

From the Table 10, it is noticed that results obtained through real data are compatible with simulation results, however, there are some exceptions which can be attributed to using large data set. The Table 10also reveals that the performance of the IP is best. In addition, results are relatively more precise under the JP than the UP.It is also observed that DLF (SELF) performance better than PLF and SELF (PLF and DLF) for estimating component (proportion) parameters.

11. Conclusion

The importance and application of the 3-component mixture models in real life problems is undeniable.An extensive simulation study is performed to compare and highlight some important and interesting properties of the Bayes estimates of a 3-component mixture of Exponential distributions using the UP, the JP and the IP under SELF, PLF and DLF. The simulation results revealed that an increase in sample size and/or test termination time produced improved (in terms of closeness)and reliable (in terms of posterior risk) Bayes estimates. It is concluded that with an increase in sample size and/or test termination time, the posterior risks decrease.To estimate component as well as proportion parameters, priors can be ordered with respect to their performance as: IP $<$ JP $<$ UP. The ordering of loss functions depends upon the parameters being estimated.

Table 7. Bayes estimates (BEs) and posterior risks (PRs) using the UP, the JP and the IP under SELF, PLF and DLF with Davis (1952) mixture data

Prior	Loss Function		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{p}_1$	$\hat{p}_2$
UP	SELF	BE	6.916945	4.026699	8.372263	0.470263	0.262025
		PR	**0.288346**	**0.194104**	**0.639500**	**0.000426**	**0.000360**
	PLF	BE	6.937758	4.050730	8.410368	0.470716	0.262711
		PR	**0.041624**	**0.048061**	**0.076210**	**0.000905**	**0.001371**
	DLF	BE	6.958632	4.074904	8.448647	0.471169	0.263398
		PR	**0.005991**	**0.011830**	**0.009041**	**0.001922**	**0.005212**
JP	SELF	BE	6.900167	3.999295	8.313222	0.470132	0.262032
		PR	**0.286064**	**0.191420**	**0.635543**	**0.000425**	**0.000359**
	PLF	BE	6.920864	4.023155	8.351360	0.470584	0.262716
		PR	**0.041396**	**0.047721**	**0.076275**	**0.000903**	**0.001368**
	DLF	BE	6.941624	4.047158	8.389672	0.471036	0.263402
		PR	**0.005972**	**0.011826**	**0.009112**	**0.001919**	**0.005201**
IP	SELF	BE	6.339530	3.948387	7.209497	0.473607	0.253837
		PR	**0.212304**	**0.168152**	**0.469845**	**0.000416**	**0.000341**
	PLF	BE	6.356253	3.969624	7.242009	0.474045	0.254509
		PR	**0.033445**	**0.042473**	**0.065024**	**0.000878**	**0.001343**
	DLF	BE	6.373019	3.990975	7.274667	0.474485	0.255182
		PR	**0.005255**	**0.010671**	**0.008958**	**0.001850**	**0.005271**

Specifically, for estimating component parameters, ordering of loss functions is: DLF < PLF < SELF, while it changes to SELF < PLF < DLF when proportion parameters are being estimated. The results obtained through real data coincide with the simulated results. Finally,it can be concluded that for a Bayesian analysis of mixture data, the IP paired with SELF and the IP paired with DLF are preferable choices for estimating proportion and component parameters, respectively.

References

[1] Abu-Taleb, A. A., Smadi, M. M. and Alawneh, A. J. *Bayes estimation of the lifetime parameters for the Exponential distribution*, Journal of Mathematics and Statistics **3**(3), 106-108, 2007.

[2] Al-Awadhi, S. A. and Gartwaite, P. H. *An elicitation method for multivariate normal distributions*, Communication in Statistics, Part A-Theory and Methods **27**, 1123-1142, 1998.

[3] Ali, M. M., Woo, J. and Nadarajah, S. *Bayes estimators of the Exponential distribution*, Journal of Statistics and Management Systems **8**(1), 53-58, 2005.

[4] Aslam, M. *An Application of Prior Predictive Distribution to Elicit the Prior Density*, Journal of Statistical Theory and Applications **2**, 70-83, 2003.

[5] Bayes, T. *An essay towards solving a problem in the doctrine of chances*, Philosophical Transactions of the Royal Society of London **53**, 370-418, 1763.

[6] Barger, K. J. *Mixtures of Exponential distributions to describe the distribution of Poisson means in estimating the number of unobserved* classes, M. Sc. Thesis, Cornell University, 2006.

[7] Berger, J. O. *Statistical decision theory and Bayesian analysis*(New York: Springer-Verlag, 1985).

[8] Bernardo, J. M. *Reference posterior distributions for Bayesian inference*, Journal of the Royal Statistical Society, Series B (Methodological), **41**(2), 113-147, 1979.

[9] Birch, R. and Bartolucci, A. A. *Determination of the hyperparameters of a prior probability model in survival analysis*, Computer Programs in Biomedicine **17**, 89-94, 1983.

[10] Chaloner, K. M. and Duncan, G. T. *Assessment of beta prior distribution: PM elicitation*, The Statistician **32**, 174-180, 1983.

[11] Davis, D. J. *An analysis of some failure data*, Journal of the American Statistical Association **47**, 113-150, 1952.

[12] DeGroot, M. H. *Optimal statistical decision*(McGraw-Hill, 2005).

[13] de Laplace, P. S. *Theorie analytique des probabilities (*Paris: Gautheir-Villars, 1820).

[14] Everitt, B. and Hand, D. J. *Finite mixture distribution* (New York: Chapman and Hall, 1981).

[15] Gavasakar, U. *A comparison of two elicitation methods for a prior distribution for a binomial parameter*, Management Science **34**, 784-790, 1988.

[16] Geisser, S. *On prior distributions for binary trials*, The American Statistician **38**(4), 244-247, 1984.

[17] Gijbels, I. *Censored data*, Wiley Interdisciplinary Reviews: Computational Statistics **2**(2), 178-188, 2010.

[18] Hahn, E. D. *Re-examining informative prior elicitation through the lens of Markovchain Monte Carlo methods*, Journal of the Royal Statistical Society **169**, Series A, 37-48, 2006.

[19] Hebert, J. L. and Scariano, S. M. *Comparing location estimators for Exponential mixtures under Pitman's measure of closeness*, Communications in Statistics- Theory and Methods **33**(1), 29-46, 2005.

[20] Jeffreys, H. *An invariant form for the prior probability in estimation problems*, Proceeding of the Royal Society of London, Series A, Mathematical and Physical Sciences **186**(1007), 453-461, 1946.

[21] Jeffreys, H. *Theory of Probability*(Oxford, UK: Claredon Press, 1961).

[22] Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. and Peter, S. C. *Interactive elicitation of opinion for a normal linear model*, Journal of the American Statistical Association **75**, 845-854, 1980.

[23] Kalbfleisch, J. D. and Prentice, R. L. *The Statistical analysis of failure time data*(New York: John Wiley & Sons, 2011).

[24] Kazmi, S. M. A., Aslam, M. and Ali, S. *On the Bayesian estimation for two-component mixture of Maxwell distribution assuming type-I censored data*, International Journal of Applied Science and Technology **2**(1), 197-218, 2012.

[25] Legendre, A. M. Nouvelles methodes pour la determination des orbites des cometes: Appendice sur la Methode des Moindres Carres (Paris: Gautheir-Villars, 1806)

[26] Li, L. A. *Decomposition theorems, conditional probability, and finite mixtures distributions*, Thesis, State University, New York, Albany, 1983.

[27] Li, L. A. and Sedransk, N. *Inference about the presence of a mixture*, Technical Report, State University, New York, Albany, 1982.

[28] Li, L. A. and Sedransk, N. *Mixtures of distributions: A topological approach*, The annals of Statistics **16**(4), 1623-1634, 1988.

[29] McCullagh, P. *Exponential mixtures and quadratic Exponential families*, Biometrika **81**(4), 721-729, 1994.

[30] Mendenhall, W. and Hader, R. J. *Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data*, Biometrika **45**(3-4), 504-520, 1958.

[31] Norstrom, J. G. *The use of precautionary loss function in risk analysis*, Reliability, IEEE Transactions on **45**(3), 400-403, 1996.

[32] Raqab, M. M. and Ahsanullah, M. *Estimation of the location and scale parameters of generalized Exponential distribution based on order statistic*, Journal of Statistical Computation and Simulation **69**(2), 109-123, 2001.

[33] Romeu, L. J. *Censored data*, Strategic Arms Reduction Treaty **11**(3), 1-8, 2004.

[34] Saleem. M. *Bayesian analysis of mixture distributions*. Ph. D. Dissertation, Dept. of Statistics, Quaid-i-Azam University, Islamabad, Pakistan, 2010.

[35] Saleem, M. and Aslam, M. *On prior selection for the mixture of Raleigh distribution using predictive intervals*, Pakistan Journal of Statistics **24**(1), 21-35, 2008.

[36] Saleem, M. and Aslam, M. *Bayesian analysis of the two-component mixture of the Rayleigh distribution assuming the uniform and the Jeffreys' priors*, Journal of Applied Statistical Science **16**(4), 493-502, 2009.

[37] Saleem, M., Aslam, M. and Economus, P. *On the Bayesian analysis of the mixture of Power distribution using the complete and censored sample*, Journal of Applied Statistics **37**(1), 25-40, 2010.

[38] Sinha, S. K. *Bayesian estimation* (New Delhi: New Age International (p) limited, Publisher, 1998).

[39] Strelec, L. and Stehlk, M. *On simulation of exact tests in Rayleigh and Normal families*, AIP (American Institute of Physics) Conference Proceedings **1479**, 2012.

Appendix

Table 8. Bayes Estimates (BEs) and Posterior Risks (PRs) of 3-Component Mixture of an Exponential Distribution using the UP under SELF, PLF and DLF with $\theta_1 = 4,\ \theta_2 = 3,\ \theta_3 = 2,\ p_1 = 0.5,\ p_2 = 0.3$ and $t = 0.5,\ 0.8$

t	n	Loss Functions		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{p}_1$	$\hat{p}_2$
0.5	30	SELF	BE	5.03127	4.81717	14.4058	0.493360	0.305990
			PR	**3.78733**	**7.04383**	**393.721**	**0.010409**	**0.009351**
		PLF	BE	5.18096	5.34651	7.10057	0.506304	0.321826
			PR	**0.65794**	**1.14487**	**2.26392**	**0.021169**	**0.030100**
		DLF	BE	5.69559	6.08516	32.9960	0.517202	0.336103
			PR	**0.12627**	**0.21448**	**0.31318**	**0.042015**	**0.094862**
	100	SELF	BE	4.32408	3.48119	3.18176	0.501051	0.306898
			PR	**1.06177**	**1.47960**	**2.49490**	**0.004021**	**0.003927**
		PLF	BE	4.43269	3.75565	3.47246	0.505030	0.311397
			PR	**0.23514**	**0.39934**	**0.66400**	**0.008038**	**0.012548**
		DLF	BE	4.51580	3.99458	3.89480	0.510777	0.317893
			PR	**0.05348**	**0.10395**	**0.18747**	**0.015990**	**0.039946**
	200	SELF	BE	4.15969	3.30336	2.65355	0.500567	0.304289
			PR	**0.54772**	**0.77122**	**1.06800**	**0.002233**	**0.002223**
		PLF	BE	4.18300	3.38534	2.92080	0.504513	0.307870
			PR	**0.12781**	**0.22162**	**0.37151**	**0.004457**	**0.007263**
		DLF	BE	4.29189	3.53659	3.03098	0.506274	0.309894
			PR	**0.03027**	**0.06450**	**0.12266**	**0.008705**	**0.023440**
0.8	30	SELF	BE	4.60708	4.06730	3.95910	0.491784	0.305857
			PR	**2.06164**	**3.05704**	**6.71197**	**0.008108**	**0.007047**
		PLF	BE	4.78976	4.41100	4.44132	0.499841	0.317388
			PR	**0.41090**	**0.64347**	**1.02895**	**0.016344**	**0.022641**
		DLF	BE	5.04824	4.71186	4.92032	0.50771	0.328947
			PR	**0.08470**	**0.14425**	**0.21952**	**0.032500**	**0.070789**
	100	SELF	BE	4.19187	3.34385	2.65896	0.498415	0.303129
			PR	**0.57004**	**0.73496**	**0.89049**	**0.002707**	**0.002426**
		PLF	BE	4.27433	3.41190	2.72112	0.499986	0.307765
			PR	**0.13347**	**0.20684**	**0.29326**	**0.005432**	**0.007970**
		DLF	BE	4.34097	3.53208	2.93405	0.503415	0.311406
			PR	**0.03143**	**0.06083**	**0.10652**	**0.010834**	**0.025890**
	200	SELF	BE	4.08408	3.13722	2.34034	0.499107	0.302543
			PR	**0.28609**	**0.35794**	**0.38474**	**0.001398**	**0.001282**
		PLF	BE	4.14270	3.17768	2.43561	0.500479	0.305603
			PR	**0.06892**	**0.10915**	**0.15321**	**0.002780**	**0.004201**
		DLF	BE	4.13182	3.26269	2.52608	0.502583	0.306298
			PR	**0.01687**	**0.03470**	**0.06245**	**0.005568**	**0.013850**

Table 9. Bayes Estimates (BEs) and Posterior Risks (PRs) of 3-Component Mixture of an Exponential Distribution using the JP under SELF, PLF and DLF with $\theta_1 = 4,\ \theta_2 = 3,\ \theta_3 = 2,\ p_1 = 0.5,\ p_2 = 0.3$ and $t = 0.5,\ 0.8$

t	n	Loss Functions		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{p}_1$	$\hat{p}_2$
0.5	30	SELF	BE	4.68438	4.14788	4.14810	0.486255	0.307003
			PR	**3.28607**	**5.84447**	**112.213**	**0.010047**	**0.009113**
		PLF	BE	5.06787	4.50278	4.97240	0.492724	0.320652
			PR	**0.63799**	**1.05847**	**2.13408**	**0.020695**	**0.029650**
		DLF	BE	5.43172	5.13024	11.0016	0.505493	0.336744
			PR	**0.12292**	**0.22330**	**0.37228**	**0.041414**	**0.090771**
	100	SELF	BE	4.28903	3.44976	2.65577	0.496809	0.303804
			PR	**1.00298**	**1.45326**	**1.90329**	**0.003884**	**0.003721**
		PLF	BE	4.38009	3.58409	3.01029	0.503270	0.309635
			PR	**0.22362**	**0.37814**	**0.60703**	**0.007759**	**0.012199**
		DLF	BE	4.48458	3.75851	3.37003	0.504503	0.316679
			PR	**0.05133**	**0.10363**	**0.19409**	**0.015604**	**0.039164**
	200	SELF	BE	4.17641	3.21003	2.41831	0.497047	0.303295
			PR	**0.53203**	**0.72122**	**0.90182**	**0.002145**	**0.002173**
		PLF	BE	4.19861	3.37741	2.64199	0.502472	0.305365
			PR	**0.12219**	**0.21241**	**0.33145**	**0.004263**	**0.006915**
		DLF	BE	4.25922	3.47091	2.83717	0.503951	0.310409
			PR	**0.02966**	**0.06356**	**0.12632**	**0.008578**	**0.022924**
0.8	30	SELF	BE	4.39979	3.52249	3.04594	0.486904	0.306358
			PR	**1.94160**	**2.57549**	**4.29295**	**0.008017**	**0.007075**
		PLF	BE	4.60389	3.91436	3.79902	0.495912	0.316711
			PR	**0.40382**	**0.62240**	**0.98201**	**0.016328**	**0.022631**
		DLF	BE	4.82216	4.34275	3.93999	0.504094	0.327098
			PR	**0.08604**	**0.15228**	**0.25101**	**0.032569**	**0.070579**
	100	SELF	BE	4.13305	3.21866	2.38451	0.498135	0.302131
			PR	**0.55222**	**0.69769**	**0.76144**	**0.002695**	**0.002421**
		PLF	BE	4.20130	3.26838	2.54904	0.499784	0.307532
			PR	**0.13079**	**0.20230**	**0.28690**	**0.005401**	**0.007958**
		DLF	BE	4.22792	3.43384	2.74990	0.502684	0.310692
			PR	**0.03143**	**0.06125**	**0.10953**	**0.010812**	**0.025775**
	200	SELF	BE	4.06307	3.10621	2.24022	0.498726	0.302046
			PR	**0.28071**	**0.35039**	**0.35378**	**0.001388**	**0.001271**
		PLF	BE	4.09626	3.13286	2.34996	0.499902	0.305688
			PR	**0.06796**	**0.10693**	**0.14942**	**0.002775**	**0.004169**
		DLF	BE	4.14565	3.22203	2.39826	0.501271	0.306324
			PR	**0.01668**	**0.03474**	**0.06291**	**0.005545**	**0.013783**

Table 10. Bayes Estimates (BEs) and Posterior Risks (PRs) of 3-Component Mixture of an Exponential Distribution using the IP under SELF, PLF and DLF with $\theta_1 = 4,\ \theta_2 = 3,\ \theta_3 = 2,\ p_1 = 0.5,\ p_2 = 0.3$ and $t = 0.5,\ 0.8$

t	n	Loss Functions		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{p}_1$	$\hat{p}_2$
0.5	30	SELF	BE	2.33666	1.99205	1.56456	0.516529	0.253366
			PR	**0.38446**	**0.44742**	**0.42467**	**0.008301**	**0.006275**
		PLF	BE	2.40593	2.12002	1.67158	0.52332	0.269693
			PR	**0.16074**	**0.21795**	**0.26020**	**0.016032**	**0.024229**
		DLF	BE	2.50499	2.22293	1.82387	0.530591	0.282341
			PR	**0.06619**	**0.10097**	**0.15199**	**0.030634**	**0.089499**
	100	SELF	BE	3.12388	2.56172	1.92370	0.513419	0.284786
			PR	**0.32993**	**0.38752**	**0.37916**	**0.003298**	**0.002739**
		PLF	BE	3.17347	2.64326	2.03329	0.518166	0.288056
			PR	**0.10266**	**0.14824**	**0.18994**	**0.006375**	**0.009540**
		DLF	BE	3.23605	2.73451	2.14476	0.520769	0.295013
			PR	**0.03215**	**0.05510**	**0.09357**	**0.012292**	**0.032628**
	200	SELF	BE	3.47498	2.77357	2.06363	0.511406	0.292558
			PR	**0.25846**	**0.30721**	**0.31254**	**0.001865**	**0.001619**
		PLF	BE	3.49013	2.80368	2.16324	0.512843	0.295125
			PR	**0.07334**	**0.10910**	**0.14811**	**0.003661**	**0.005535**
		DLF	BE	3.54183	2.88947	2.22229	0.514936	0.297619
			PR	**0.02086**	**0.03829**	**0.06791**	**0.007115**	**0.018739**
0.8	30	SELF	BE	2.41149	2.03909	1.63540	0.505104	0.263843
			PR	**0.37034**	**0.41121**	**0.40135**	**0.007235**	**0.005650**
		PLF	BE	2.50768	2.14343	1.76339	0.51284	0.27436
			PR	**0.15051**	**0.19540**	**0.23654**	**0.014208**	**0.021037**
		DLF	BE	2.59032	2.26105	1.88480	0.519925	0.284365
			PR	**0.05938**	**0.08976**	**0.13082**	**0.027575**	**0.070114**
	100	SELF	BE	3.21605	2.57804	1.93565	0.504645	0.288391
			PR	**0.27556**	**0.30674**	**0.29392**	**0.002599**	**0.002186**
		PLF	BE	3.26907	2.61318	2.02791	0.508064	0.293030
			PR	**0.08505**	**0.11580**	**0.14777**	**0.005135**	**0.007558**
		DLF	BE	3.33053	2.77935	2.14013	0.510514	0.296837
			PR	**0.02567**	**0.04355**	**0.07129**	**0.010059**	**0.025568**
	200	SELF	BE	3.53859	2.78925	2.06133	0.504253	0.295431
			PR	**0.18994**	**0.21303**	**0.20664**	**0.001364**	**0.001188**
		PLF	BE	3.57238	2.82378	2.10400	0.505779	0.296186
			PR	**0.05289**	**0.07625**	**0.09668**	**0.002695**	**0.004006**
		DLF	BE	3.58655	2.89477	2.13775	0.507708	0.298415
			PR	**0.01471**	**0.02687**	**0.04567**	**0.005307**	**0.013426**

Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 629 – 647

A new Weibull-G family of distributions

M.H.Tahir*†, M. Zubair‡ , M. Mansoor§ , Gauss M. Cordeiro¶ , Morad Alizadeh‖
and G. G. Hamedani**

Abstract

Statistical analysis of lifetime data is an important topic in reliability engineering, biomedical and social sciences and others. We introduce a new generator based on the Weibull random variable called the new Weibull-G family. We study some of its mathematical properties. Its density function can be symmetrical, left-skewed, right-skewed, bathtub and reversed-J shaped, and has increasing, decreasing, bathtub, upside-down bathtub, J, reversed-J and S shaped hazard rates. Some special models are presented. We obtain explicit expressions for the ordinary and incomplete moments, quantile and generating functions, Rényi entropy, order statistics and reliability. Three useful characterizations based on truncated moments are also proposed for the new family. The method of maximum likelihood is used to estimate the model parameters. We illustrate the importance of the family by means of two applications to real data sets.

Keywords: Generating function, hazard function, moment, reliability function, Rényi entropy, Weibull distribution.

2000 AMS Classification: 60E05; 62E10; 62N05.

Received : 17.05.2014 *Accepted* : 11.03.2015 *Doi* : 10.15672/HJMS.2015579686

*Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan, Email: mtahir.stat@gmail.com, mht@iub.edu.pk

†Corresponding Author.

‡Department of Statistics, Government Degree College Khairpur Tamewali, Pakistan, Email: zubair.stat@yahoo.com

§Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan, Email: mansoor.abbasi143@gmail.com

¶Department of Statistics, Federal University of Pernambuco, 50740-540, Recife, PE, Brazil, Email: gauss@de.ufpe.br, gausscordeiro@gmail.com

‖Department of Statistics, Ferdowsi University of Mashhad, Mashhad, P.O. Box. 91775-1159, Iran, Email: moradalizadeh78@gmail.com

**Department of Mathematics, Statistics and Computer Science, Marquette University, WI 53201-1881, Milwaukee, USA, Email: g.hamedani@mu.edu

1. Introduction

Broadly speaking, there has been an increased interest in defining new generators for univariate continuous families of distributions by introducing one or more additional shape parameter(s) to the baseline distribution. This induction of parameter(s) has been proved useful in exploring tail properties and also for improving the goodness-of-fit of the family under study. The well-known generators are the following: beta-G by Eugene *et al.* [18], Kumaraswamy-G (Kw-G) by Cordeiro and de Castro [12], McDonald-G (Mc-G) by Alexander *et al.* [1], gamma-G type 1 by Zografos and Balakrishanan [29] and Amini *et al.* [7], gamma-G type 2 by Ristić and Balakrishanan [26] and Amini *et al.* [7], odd exponentiated generalized (odd exp-G) by Cordeiro *et al.* [14], transformed-transformer (T-X) (Weibull-X and gamma-X) by Alzaatreh *et al.* [4], exponentiated T-X by Alzaghal *et al.* [6], odd Weibull-G by Bourguignon *et al.* [8], exponentiated half-logistic by Cordeiro *et al.* [11], T-X{Y}-quantile based approach by Aljarrah *et al.* [3], T-R{Y} by Alzaatreh *et al.* [5], Lomax-G by Cordeiro *et al.* [15], logistic-X by Tahir *et al.* [28] and Kumaraswamy odd log-logistic-G by Alizadeh *et al.* [2].

Let $r(t)$ be the probability density function (pdf) of a random variable $T \in [a, b]$ for $-\infty < a < b < \infty$ and let $F(x)$ be the cumulative distribution function (cdf) of a random variable X such that the link function $W(\cdot) : [0, 1] \longrightarrow [a, b]$ satisfies the two conditions: (i) $W(\cdot)$ is differentiable and monotonically non-decreasing, and (ii) $W(0) \to a$ and $W(1) \to b$. If the interval $[a, b]$ is half-open or open, we replace $W(0)$ and/or $W(1)$ for $\lim_{t\to 0^+} W(t) \to a$ and $\lim_{t\to 1^-} W(t) \to b$.

Recently, Alzaatreh *et al.* [4] defined the *T-X family* of distributions by

$$(1.1) \qquad F(x) = \int_a^{W[G(x)]} r(t)\, dt,$$

where $W[G(x)]$ satisfies the conditions (i) and (ii). If $T \in (0, \infty)$, X is a continuous random variable and $W[G(x)] = -\log[1 - G(x)]$, then the pdf corresponding to (1.1) is given by

$$(1.2) \qquad f(x) = \frac{g(x)}{1 - G(x)}\, r(-\log[1 - G(x)]) = h_g(x)\, r[H_g(x)],$$

where $h_g(x)$ and $H_g(x)$ are the hazard and cumulative hazard functions associated to $g(x)$, respectively.

The Weibull distribution is one of the most popular and widely used model for failure time in life-testing and reliability theory. However, a drawback of this distribution as far as lifetime analysis is concerned is the monotonic behaviour of its hazard date function (hrf). In real life applications, empirical hazard rate curves often exhibit non-monotonic shapes such as a bathtub, upside-down bathtub (unimodal) and others. So, there is a genuine desire to search for some generalizations or modifications of the Weibull distribution that can provide more flexibility in lifetime modeling.

If a random variable T has the Weibull distribution with scale parameter $\alpha > 0$ and shape parameter $\beta > 0$, then its cdf and pdf are, respectively, given by

$$F_W(t) = 1 - \mathrm{e}^{-\alpha t^\beta}, \quad t > 0$$

and

$$(1.3) \qquad f_W(t) = \alpha\, \beta\, t^{\beta-1}\, \mathrm{e}^{-\alpha t^\beta}, \quad t > 0.$$

In the recent literature, four Weibull based generators have appeared, namely: the beta Weibull-G by Cordeiro *et al.* [16], the Weibull-X by Alzaatreh *et al.* [4], the Weibull-G by Bourguignon *et al.* [8] and the exponentiated Weibull-X by Alzaghal *et al.* [6].

If $r(t)$ follows (1.3) and setting $W[G(x)] = -\log[1-G(x)]$ in (1.1), Alzaatreh *et al.* [4] defined the cdf of the Weibull-X family by

$$(1.4) \qquad F(x) = \alpha\,\beta \int_0^{-\log[1-G(x)]} x^{\beta-1}\,\mathrm{e}^{-\alpha x^{\beta}}\,dt = 1 - \mathrm{e}^{-\alpha\{-\log[1-G(x)]\}^{\beta}}.$$

The pdf corresponding to (1.4) is

$$(1.5) \qquad f(x) = \alpha\,\beta\,\frac{g(x)}{1-G(x)}\,\mathrm{e}^{-\alpha(-\log[1-G(x)])^{\beta}}\,\{-\log[1-G(x)]\}^{\beta-1}.$$

Zagrafos and Balakrishnan [29] pioneered a versatile and flexible gamma-G class of distributions based on Stacy's generalized gamma distribution and record value theory. More recently, Bourguignon *et al.* [8] proposed the *Weibull-G* family of distributions influenced by the Zografos-Balakrishnan-G class. Bourguignon *et al.* [8] replaced the argument x by $G(x;\Theta)/\overline{G}(x;\Theta)$, where $\overline{G}(x;\Theta) = 1 - G(x;\Theta)$, and defined the cdf of their class (for $\alpha > 0$ and $\beta > 0$), say Weibull-G(α,β,Θ), by

$$(1.6) \qquad F(x;\alpha,\beta,\Theta) = \alpha\,\beta \int_0^{\left[\frac{G(x;\Theta)}{\overline{G}(x;\Theta)}\right]} x^{\beta-1}\,\mathrm{e}^{-\alpha x^{\beta}}\,dx = 1 - \mathrm{e}^{-\alpha\left[\frac{G(x;\Theta)}{\overline{G}(x;\Theta)}\right]^{\beta}}, \ x \in \Re.$$

The Weibull-G family density function becomes

$$(1.7) \qquad f(x;\alpha,\beta,\Theta) = \alpha\,\beta\,g(x;\Theta)\left[\frac{G(x;\Theta)^{\beta-1}}{\overline{G}(x;\Theta)^{\beta+1}}\right]\mathrm{e}^{-\alpha\left[\frac{G(x;\Theta)}{\overline{G}(x;\Theta)}\right]^{\beta}},$$

where $G(x;\Theta)$ and $g(x;\Theta)$ are the cdf and pdf of any baseline distribution that depend on a parameter vector Θ.

In this paper, we propose a class of distributions called the *new Weibull-G* ("NWG" for short) family, which is flexible because of the hazard rate shapes: constant, increasing, decreasing, bathtub, upside-down bathtub, J, reversed-J and S. The paper unfolds as follows. In Section 2, we define the new family. Six special models are presented in Section 3. The forms of the density and hazard rate functions are described analytically in Section 4. In Section 5, we obtain explicit expressions for the quantile function (qf), ordinary and incomplete moments, generating function and entropies. In Sections 6 and 7, we investigate the order statistics and the reliability. Section 8 refers to some characterizations of the NWG family. In Section 9, the parameters of the new family are estimated by the method of maximum likelihood. In Section 10, we illustrate its performance by means of two applications to real data sets. The paper is concluded in Section 11.

2. The new family

In equation (1.1), let $a = 0$, $r(t)$ be as in (1.3) and $W[G(x)] = -\log[G(x;\boldsymbol{\xi})]$. Then, we define the cdf of the NWG family by

$$(2.1) \qquad F(x;\alpha,\beta,\boldsymbol{\xi}) = 1 - \int_0^{-\log[G(x;\boldsymbol{\xi})]} \alpha\,\beta\,t^{\beta-1}\,\mathrm{e}^{-\alpha t^{\beta}}\,dt = \mathrm{e}^{-\alpha\{-\log[G(x;\boldsymbol{\xi})]\}^{\beta}}.$$

Hereafter, a random variable X with cdf (2.1) is denoted by $X \sim \mathrm{NWG}(\alpha,\beta,\boldsymbol{\xi})$. We can motivate equation (2.1) based on linearization of the baseline cdf $G(x;\boldsymbol{\xi})$ as follows. Let Y be a Weibull random variable with scale parameter $\alpha > 0$ and shape parameter $\beta > 0$. The extreme value random variable V can be defined as minus the log of the Weibull random variable, say $V = -\log(Y)$. It gives the limiting distribution for the smallest or largest values in samples drawn from a variety of distributions. The NWG random variable having cdf (2.1) can be derived by

$$P(X \geq x) = P\Big(Y \geq -\log[G(x;\boldsymbol{\xi})]\Big) = P\Big(V \leq -\log\{-\log[G(x;\boldsymbol{\xi})]\}\Big),$$

where $-\log\{-\log[G(x;\boldsymbol{\xi})]\}$ is a simple linearization of the baseline cdf.

Then, the pdf of X reduces to

$$f(x;\alpha,\beta,\boldsymbol{\xi}) = \alpha\beta\, \frac{g(x;\boldsymbol{\xi})}{G(x;\boldsymbol{\xi})}\, \Big\{-\log\left[G(x;\boldsymbol{\xi})\right]\Big\}^{\beta-1}\, \mathrm{e}^{-\alpha\{-\log[G(x;\boldsymbol{\xi})]\}^{\beta}}, \tag{2.2}$$

where $g(x;\boldsymbol{\xi})$ is the parent pdf. Further, we can omit sometimes the dependence on the vector $\boldsymbol{\xi}$ of the parameters and write simply $G(x) = G(x;\boldsymbol{\xi})$ and $g(x) = g(x;\boldsymbol{\xi})$. Equation (2.2) will be most tractable when the cdf $G(x)$ and pdf $g(x)$ have simple analytic expressions.

The quantile function (qf) of X is obtained by inverting (2.1). We have

$$X = Q(u) = Q_G\left(\mathrm{e}^{-t(u)};\boldsymbol{\xi}\right),$$

where $Q_G(\cdot;\cdot) = G^{-1}(\cdot;\cdot)$ is the baseline qf and $t(u) = [\log(1/u)^{1/\alpha}]^{1/\beta}$. Then, if U has a uniform distribution on $(0,1)$, $X = Q(U)$ follows the NWG$(\alpha,\beta,\boldsymbol{\xi})$ family.

Let $h(x;\boldsymbol{\xi})$ be the hrf of the parent G. The hrf $h(x;\alpha,\beta,\boldsymbol{\xi})$ of X is given by

$$h(x;\alpha,\beta,\boldsymbol{\xi}) = \frac{\alpha\beta\frac{g(x;\boldsymbol{\xi})}{G(x;\boldsymbol{\xi})}\ \{-\log\left[G(x;\boldsymbol{\xi})\right]\}^{\beta-1}\ \mathrm{e}^{-\alpha\{-\log\left[G(x;\boldsymbol{\xi})\right]\}^{\beta}}}{1-\mathrm{e}^{-\alpha\{-\log\left[G(x;\boldsymbol{\xi})\right]\}^{\beta}}}.$$

3. Special models

In this section, we provide six special models of the NWG distributions. Suppose that the parent distribution is uniform on the interval $(0,\theta)$, $\theta > 0$. We have $g(x;\theta) = 1/\theta$, $0 < x < \theta < \infty$, and $G(x;\theta) = x/\theta$, and then the Weibull-uniform (WU) cdf is given by

$$F_{WU}(x;\alpha,\beta,\theta) = \mathrm{e}^{-\alpha\left[-\log\left(\frac{x}{\theta}\right)\right]^{\beta}}, \quad 0 < x < \theta < \infty \quad \alpha,\beta,\theta > 0.$$

Now, take the parent distribution as Weibull with pdf and cdf given by $g(x) = \lambda\,\gamma\, x^{\gamma-1}\,\mathrm{e}^{-\lambda\, x^{\gamma}}$ and $G(x) = 1-\mathrm{e}^{-\lambda\, x^{\gamma}}$ for $\lambda,\gamma > 0$. Then, the Weibull-Weibull (WW) cdf becomes

$$F_{WW}(x;\alpha,\beta,\lambda,\gamma) = \mathrm{e}^{-\alpha\left[-\log\left(1-\mathrm{e}^{-\lambda\, x^{\gamma}}\right)\right]^{\beta}}, \quad x > 0, \quad \alpha,\beta,\lambda,\gamma > 0.$$

For $\gamma = 1$ and $\gamma = 2$, we obtain as special cases the Weibull-exponential (WE) and Weibull-Rayleigh (WR) distributions, respectively.

For the Weibull-logistic (WLo) distribution, we have $g(x) = \lambda\,\mathrm{e}^{-\lambda\, x}\,\left(1+\mathrm{e}^{-\lambda\, x}\right)^{-2}$ and $G(x) = \left(1+\mathrm{e}^{-\lambda\, x}\right)^{-1}$. Then, the WLo cdf reduces to

$$F_{WLo}(x;\alpha,\beta,\lambda) = \mathrm{e}^{-\alpha\left\{-\log\left[\left(1+\mathrm{e}^{-\lambda\, x}\right)^{-1}\right]\right\}^{\beta}}, \quad x > 0, \quad \alpha,\beta,\lambda > 0.$$

Consider the parent log-logistic distribution with parameters $s > 0$ and $c > 0$ given by $g(x;s,c) = c\, s^{-c}\, x^{c-1}\, \left[1+\left(\frac{x}{s}\right)^{c}\right]^{-2}$ and $G(x;s,c) = 1-\left[1+\left(\frac{x}{s}\right)^{c}\right]^{-1}$.

Then, the Weibull-log-logistic (WLL) cdf becomes

$$F_{WLL}(x;\alpha,\beta,s,c) = \mathrm{e}^{-\alpha\left[-\log\left\{1-\left[1+\left(\frac{x}{s}\right)^{c}\right]^{-1}\right\}\right]^{\beta}}, \quad x > 0, \quad \alpha,\beta,s,c > 0.$$

We take the parent Burr XII distribution with pdf and cdf given by $g(x) = c\,k\,s^{-c}\,x^{c-1}\,[1+(x/s)^c]^{-k-1}$ and $G(x) = 1-[1+(x/s)^c]^{-k}$. Then, the Weibull-Burr XII (WBXII) cdf reduces to

$$F_{WBXII}(x;\alpha,\beta,s,c,k) = \mathrm{e}^{-\alpha\left[-\log\left\{1-\left[1+\left(\frac{x}{s}\right)^c\right]^{-k}\right\}\right]^{\beta}}, \quad x>0, \quad \alpha,\beta,s,c,k>0.$$

For $c=1$ and $k=1$, we obtain as a special case the Weibull-Lomax (WLx) distribution.

Finally, if we consider the baseline normal distribution, the pdf and cdf are $g(x;\mu,\sigma) = \sigma^{-1}\,\phi\,[(x-\mu)/\sigma]$ and $G(x;\mu,\sigma) = \Phi\,[(x-\mu)/\sigma]$. Then, the Weibull-normal (WN) cdf becomes (for $x \in \Re$)

$$F_{WN}(x;\alpha,\beta,\mu,\sigma) = \mathrm{e}^{-\alpha\left[-\log\left\{\Phi\left(\frac{x-\mu}{\sigma}\right)\right\}\right]^{\beta}}, \quad x\in\Re, \quad \alpha,\beta,\sigma>0, \mu\in\Re.$$

The density of the new family can be symmetrical, left-skewed, right-skewed, bathtub and reversed-J shaped, and has constant, increasing, decreasing, bathtub, upside-down bathtub, J, reversed J and S shaped hazard rates. In Figures 1 and 2, we display some plots of the pdf and hrf of (a) WU, (b) WW, (c) WLL, (d) WLo, (e) WBXII and (f) WN distributions for selected parameter values. Figure 1 indicates that the NWG family generates distributions with various shapes such as symmetrical, left-skewed, right-skewed, bathtub and reversed-J. Also, Figure 2 reveals that this family can produce flexible hazard rate shapes such as increasing, decreasing, bathtub, upside-down bathtub, J, reversed-J and S. This fact implies that the NWG family can be very useful to fit different data sets with various shapes.

4. Shapes of the pdf and hrf

The shapes of the density and hazard rate functions can be described analytically. The critical points of the NWG density are the roots of the equation:

$$\frac{g'(x)}{g(x)} + \frac{g(x)}{G(x)} + \frac{(1-\beta)g(x)}{G(x)\log[G(x)]} + \frac{\alpha\beta\,g(x)\,\{-\log[G(x)]\}^{\beta-1}}{G(x)} = 0. \tag{4.1}$$

The critical points of $h(x)$ are obtained from the equation

$$\begin{aligned}&\frac{g'(x)}{g(x)} + \frac{g(x)}{G(x)} + \frac{(1-\beta)g(x)}{G(x)\log[G(x)]} + \frac{\alpha\beta\,g(x)\,\{-\log[G(x)]\}^{\beta-1}}{G(x)}\\ &+\frac{\alpha\beta\,g(x)\,\{-\log[G(x)]\}^{\beta-1}\,\mathrm{e}^{-\alpha\{-\log[\mathrm{G(x)}]\}^{\beta}}}{G(x)\left[1-\mathrm{e}^{-\alpha\{-\log[\mathrm{G(x)}]\}^{\beta}}\right]} = 0.\end{aligned} \tag{4.2}$$

By using most symbolic computation software platforms, we can examine equations (4.1) and (4.2) to determine the local maximums and minimums and inflexion points.

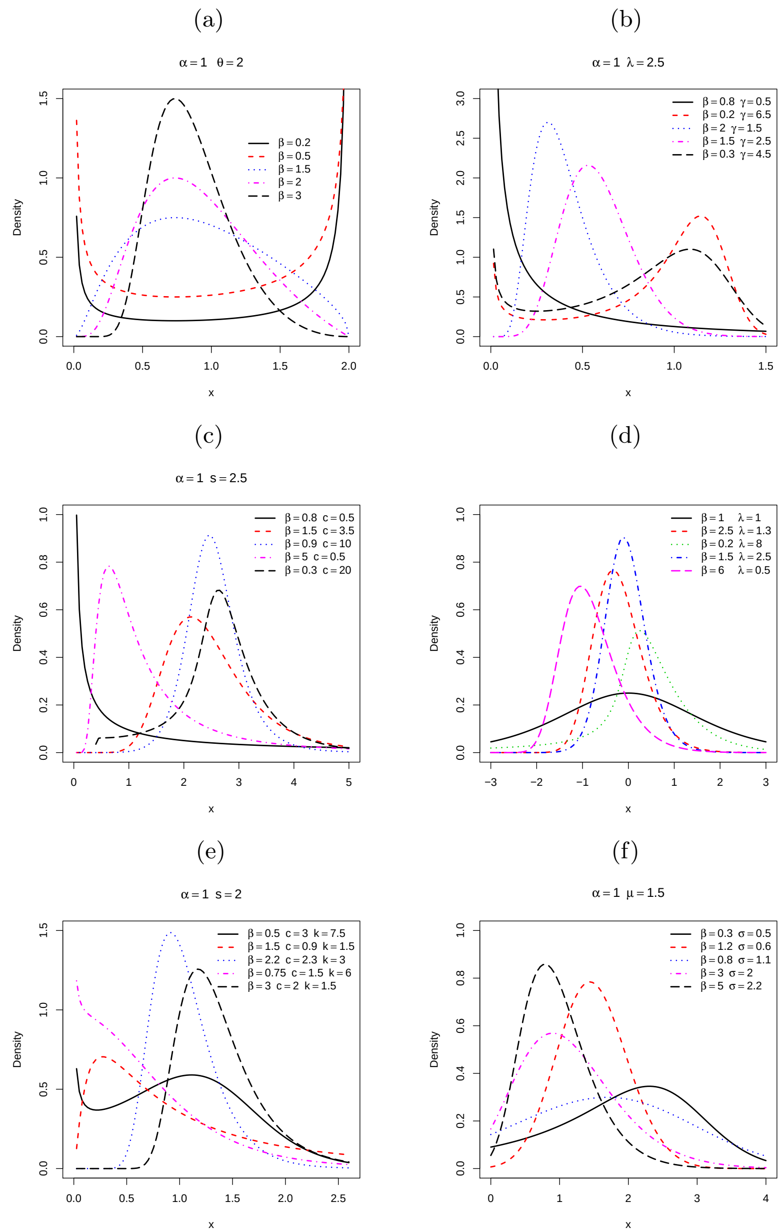


Figure 1. Plots of the (a) WU (b) WW (c) WLL (d) WLo (e) WBXII and (f) WN densities.

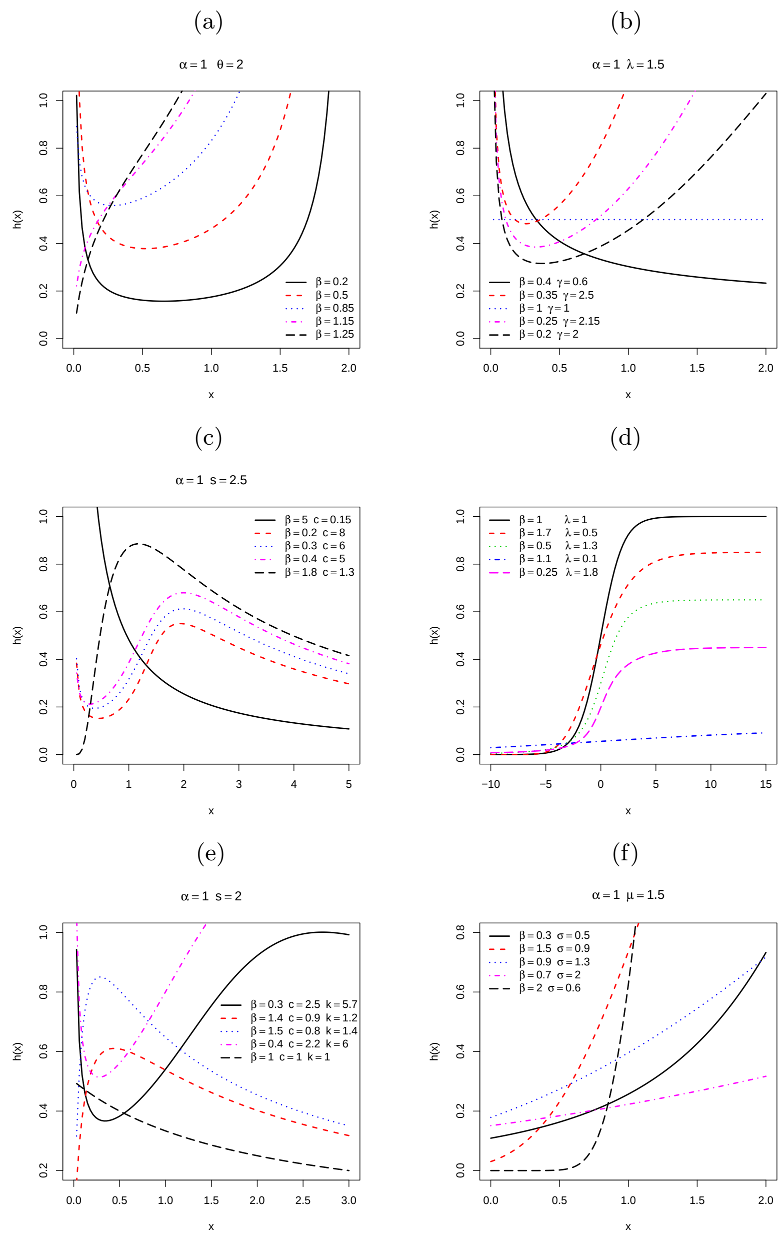


Figure 2. Plots of the (a) WU (b) WW (c) WLL (d) WLo (e) WBXII and (f) WN hazard rates.

5. Mathematical properties

The formulae derived throughout the paper can be easily handled in analytical softwares such as Maple and Mathematica which have the ability to deal with analytic expressions of formidable size and complexity. Established algebraic expansions to determine some mathematical properties of the NWG family can be more efficient than computing those directly by numerical integration of its density function, which can be prone to rounding off errors among others. The infinity limit in these sums can be substituted by a large positive integer such as 20 or 30 for most practical purposes. Here, we provide some mathematical properties of X.

5.1. Expansion for the NWG cdf. Let $A = \mathrm{e}^{-\alpha\{\,-\log[G(x;\boldsymbol{\xi})]\,\}^{\beta}}$. Then, using a power series expansion for A, we can write (2.2) as

$$F(x;\alpha,\beta,\boldsymbol{\xi}) = \sum_{i=0}^{\infty} \frac{(-1)^i\,\alpha^i}{i!}\,\{\,-\log[G(x;\boldsymbol{\xi})]\,\}^{i\,\beta}\,. \tag{5.1}$$

The following formula holds for $i \geq 1$
(http:// functions.wolfram.com/ ElementaryFunctions/Log/06/01/04/03/),
and then we can write

$$\begin{aligned}\{\,-\log[G(x;\boldsymbol{\xi})]\,\}^{i\,\beta} \quad = \quad & \sum_{k,l=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{j+k+l}\,i\,\beta}{(i\beta-j)} \binom{k-i\,\beta}{k}\binom{k}{j}\binom{\beta\,i+k}{l} \\ & \times \quad p_{j,k}\;[G(x;\boldsymbol{\xi})]^{l},\end{aligned}$$

where (for $j \geq 0$) $p_{j,0} = 1$ and (for $k = 1, 2, \ldots$)

$$p_{j,k} = k^{-1}\sum_{m=1}^{k} \frac{(-1)^m[m(j+1)-k]}{(m+1)}\;p_{j,k-m}.$$

By inserting the above power series in equation (5.1) gives

$$F(x;\alpha,\beta,\boldsymbol{\xi}) = \sum_{l=0}^{\infty} b_l\,G(x;\boldsymbol{\xi})^l = \sum_{l=0}^{\infty} b_l\,H_l(x;\boldsymbol{\xi}), \tag{5.2}$$

where $H_l(x;\boldsymbol{\xi}) = G(x;\boldsymbol{\xi})^l$ (for $l \geq 1$), is the exponentiated-G (exp-G) density function with power parameter l, $H_0(x;\boldsymbol{\xi}) = 1$,

$$b_l = \sum_{i,k=0}^{\infty}\sum_{j=0}^{k} \frac{(-1)^{i+j+k+l}\,i\,\beta}{i!\,(i\,\beta-j)} \binom{k-i\,\beta}{k}\binom{k}{j}\binom{i\,\beta+k}{l}\;p_{j,k}.$$

We can write the NWG family density as a mixture of exp-G densities

$$f(x;\alpha,\beta,\boldsymbol{\xi}) = \sum_{l=0}^{\infty} b_{l+1}\,h_{l+1}(x;\boldsymbol{\xi}), \tag{5.3}$$

where $h_{l+1}(x;\boldsymbol{\xi}) = (l+1)\,g(x;\boldsymbol{\xi})\,G(x;\boldsymbol{\xi})^l$ is the exp-G density function with power parameter $l+1$.

Thus, some mathematical properties of the proposed family can be derived from (5.3) and those of exp-G properties. For example, the ordinary and incomplete moments and moment generating function (mgf) of X can be obtained from those exp-G quantities. Some mathematical properties of the exp-G distributions are studied by [20, 21, 23] and others.

5.2. Moments. Let Y_l be a random variable having the exp-G density function $h_{l+1}(x)$. A first formula for the nth moment of X follows from (5.3) as

$$E(X^n) = \sum_{l=0}^{\infty} b_{l+1}\, E(Y_l^n). \tag{5.4}$$

Expressions for moments of several exp-G distributions are given by Nadarajah and Kotz [23], which can be used to obtain $E(X^n)$.

A second formula for $E(X^n)$ can be written from (5.4) in terms of the G qf as

$$E(X^n) = \sum_{l=0}^{\infty} (l+1)\, b_{l+1}\, \tau_{n,l}, \tag{5.5}$$

where $\tau_{n,l} = \int_{-\infty}^{\infty} x^n\, G(x)^l\, g(x)\, dx = \int_0^1 Q_G(u)^n\, u^l\, du$.

Cordeiro and Nadarajah [13] obtained $\tau_{n,l}$ for some well-known distributions such as normal, beta, gamma and Weibull, which can be used to determine the NWG moments.

For empirical purposes, the shapes of many distributions can be usefully described by what we call the incomplete moments. These types of moments play an important role for determining Lorenz and Bonferroni curves.

The nth incomplete moment of X is obtained as

$$m_n(y) = \sum_{l=0}^{\infty} (l+1)\, b_{l+1} \int_0^{G(y)} Q_G(u)^n\, u^l\, du. \tag{5.6}$$

The last integral can be computed for most G distributions. Equations (5.4)-(5.6) are the main results of this section.

5.3. Generating function. Let $M_X(t) = E(\mathrm{e}^{\mathrm{t\,X}})$ be the mgf of X. Then, we can write

$$M_X(t) = \sum_{l=0}^{\infty} b_{l+1}\, M_l(t), \tag{5.7}$$

where $M_l(t)$ is the mgf of Y_l. Hence, $M_X(t)$ can be determined from the exp-G generating function.

A second formula for $M_X(t)$ can be expressed as

$$M_X(t) = \sum_{l=0}^{\infty} (l+1)\, b_{l+1}\, \rho(t,l), \tag{5.8}$$

where $\rho(t,l) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{t\,x}}\, \mathrm{G(x)^l\, g(x)\, dx} = \int_0^1 \mathrm{e}^{\mathrm{t\,Q_G(u)}}\, \mathrm{u^l\, du}$.

We can obtain the mgfs of several distributions directly from equation (5.8).

5.4. Rényi entropy. Entropy has wide application in science, engineering and probability theory, and has been used in various situations as a measure of variation of the uncertainty. The entropy of a random variable X is a measure of variation of uncertainty. Here, we derive explicit expressions for the Rényi entropy [25] of the NWG family. The Shannon entropy [27] of a random variable X is defined by $E\{-\log[f(X)]\}$. It is the special case of the Rényi entropy when $\gamma \uparrow 1$.

The Rényi entropy is defined by

$$I_R(\delta) = \frac{1}{1-\delta} \log[I(\delta)],$$

where $I(\delta) = \int_{-\infty}^{\infty} f^{\delta}(x)\, dx, \delta > 0$ and $\delta \neq 1$.

Let us consider

$$f^{\delta}(x) = (\alpha\,\beta)^{\delta}\, g^{\delta}(x)\; G^{-\delta}(x)\, \{-\log[G(x)]\}^{\delta(\beta-1)}\, \mathrm{e}^{-\delta\alpha\{-\log[G(x)]\}^{\beta}}.$$

Expanding the exponential function in power series and then expanding the power of $\{-\log[G(x)]\}$ as in Section 5.1, we obtain

$$\begin{aligned} f^{\delta}(x) = (\alpha\,\beta)^{\delta} \sum_{i,k=0}^{\infty} \frac{(-1)^{i}\,(\alpha\,\delta)^{i}\,[\delta(\beta-1)+i\beta]}{i!} \binom{k-\delta(\beta-1)-i\beta}{k} \\ \times \sum_{j=0}^{\infty} \frac{(-1)^{j+k}\, p_{j,k}\binom{k}{j}}{[\delta(\beta-1)+i\beta-j]}\;\; g^{\delta}(x)\; G^{\delta}(x)\,[1-G(x)]^{k-\delta(\beta-1)-i\beta}, \end{aligned}$$

where the constants $p_{j,k}$ are given in Section 5.1.

Further, using the binomial expansion in the last equation, we can write

$$f^{\delta}(x) = \sum_{l=0}^{\infty} S_l\; g^{\delta}(x)\; G^{\delta+l}(x),$$

where

$$\begin{aligned} S_l \quad = \quad & \sum_{i,j,k=0}^{\infty} \frac{(-1)^{i+j+k+l}\,(\alpha\,\delta)^{i}\,[\delta(\beta-1)+i\beta]}{i!\,[\delta(\beta-1)+i\beta-j]} \\ & \times\; p_{j,k-m}\binom{k}{j}\binom{k-\delta(\beta-1)-i\beta}{k}\binom{k+\delta(\beta-1)+i\beta}{l}. \end{aligned}$$

Hence, the Rényi entropy reduces to

$$I_R(\delta) = \frac{1}{1-\delta}\log\left[\sum_{l=0}^{\infty} S_l \int_{-\infty}^{\infty} g^{\delta}(x)\, G^{\delta+l}(x)\; dx\right].$$

6. Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1,\ldots,X_n$ be observed values of a sample from the NWG family of distributions. We can write the density of the ith order statistic, say $X_{i:n}$, as

$$f_{i:n}(x) = \frac{n!}{(i-1)!\,(n-i)!} \sum_{j=0}^{n-i}(-1)^{j}\binom{n-i}{j} f(x)\, F(x)^{j+i-1}.$$

Following similar algebraic developments of Nadarajah *et al.* [22], we can write

$$f_{i:n}(x) = \sum_{r,k=0}^{\infty} m_{r,k}\, h_{r+k+1}(x), \tag{6.1}$$

where $h_{r+k+1}(x)$ is the exp-G density function with power parameter $r+k+1$,

$$m_{r,k} = \frac{n!\,(r+1)\,(i-1)!\,b_{r+1}}{(r+k+1)} \sum_{j=0}^{n-i} \frac{(-1)^{j}\, f_{j+i-1,k}}{(n-i-j)!\,j!},$$

and b_k is defined in equation (5.2). Here, the quantities $f_{j+i-1,k}$ are obtained recursively by $f_{j+i-1,0} = b_0^{j+i-1}$ and (for $k \geq 1$)

$$f_{j+i-1,k} = (k\, b_0)^{-1} \sum_{m=1}^{k} [m(j+i)-k]\, b_m\, f_{j+i-1,k-m}.$$

Based on the expansion (6.1), we can obtain some mathematical properties (ordinary and incomplete moments, generating function, etc.) for the NWG order statistics from those exp-G properties.

7. Reliability

We derive the reliability $R = P(X_2 < X_1)$ when $X_1 \sim \mathrm{NWG}(\alpha_1, \beta_1, \boldsymbol{\xi_1})$ and $X_2 \sim \mathrm{NWG}(\alpha_2, \beta_2, \boldsymbol{\xi_2})$ are independent random variables with a positive support. It has many applications especially in engineering concepts. Let $f_i(x)$ and $F_i(x)$ denote the pdf and cdf of X_i for $i = 1, 2$. By using the mixture representations for $F_2(x)$ and $f_1(x)$ given in Section 5.1, we obtain

$$R = \sum_{k,s=0}^{\infty} b_k^{(1)}\; b_{s+1}^{(2)}\; R_{k,s+1},$$

where $b_l^{(1)}$ and $b_{s+1}^{(2)}$ are given in these representations and

$$R_{k,s+1} = \int_0^{\infty} H_k(x; \alpha_1, \beta_1, \boldsymbol{\xi_1})\; h_{s+1}(x; \alpha_2, \beta_2, \boldsymbol{\xi_2})\, dx.$$

If $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$, then

$$R = \sum_{k,s=0}^{\infty} \frac{(s+1)}{(s+k+1)}\, b_k^{(1)}\, b_{s+1}^{(2)}.$$

Finally, if $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$ and $\boldsymbol{\xi_1} = \boldsymbol{\xi_2}$, then $R = 1/2$ as expected.

8. Characterizations of the NWG family

Various characterizations of distributions have been established in many different directions. In this section, three characterizations of the NWG family are presented based on: (i) a simple relationship between two truncated moments; (ii) a single function of the random variable, and (iii) the hazard function.

8.1. Characterization based on truncated moments. Here, we present a characterization of the NWG family in terms of a simple relationship between two truncated moments. The characterization results employ an interesting result due to Glänzel [19] (Theorem 1, below). It has the advantage that the cdf F is not required to have a closed-form and is given in terms of an integral whose integrand depends on the solution of a first order differential equation, which can serve as a bridge between probability and differential equation.

8.1. Theorem. *Let $(\Omega, \Sigma, \mathbf{P})$ be a given probability space and let $H = [a, b]$ be an interval for some $a < b$ ($a = -\infty, b = \infty$ might as well be allowed). Let $X : \Omega \to H$ be a continuous random variable with distribution function $F(x)$ and let q_1 and q_2 be two real functions defined on H such that*

$$\mathbb{E}\left[q_1(X) \mid X \geq x\right] = \mathbb{E}\left[q_2(X) \mid X \geq x\right]\, \eta(x), \quad x \in H,$$

is defined with some real function η. Assume that q_1, $q_2 \in C^1(H)$, $\eta \in C^2(H)$ and $G(x)$ is twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $q_2\eta = q_1$ has no real solution in the interior of H. Then G is uniquely determined by the functions q_1, q_2 and η, particularly

$$F(x) = \int_a^x C \left| \frac{\eta'(u)}{\eta(u)\, q_2(u) - q_1(u)} \right| \mathrm{e}^{-s(u)}\, du\,,$$

where the function s is a solution of the differential equation $s' = \frac{\eta' \, q_2}{\eta \, q_2 - q_1}$ and C is a constant, chosen to make $\int_H dF = 1$.

8.2. Proposition. *Let $X : \Omega \to (0, \infty)$ be a continuous random variable and let $q_2 (x) = \mathrm{e}^{\alpha\{-\log[\mathrm{G}(\mathrm{x};\xi)]\}^{\beta}}$ and $q_1 (x) = q_2 (x) \{-\log [G (x; \xi)]\}$ for $x > 0$. The pdf of X is (2.2) if and only if the function η defined in Theorem 1 has the form*

$$\eta (x) = \frac{\beta}{\beta + 1} \{-\log [G (x; \xi)]\}, \quad x > 0.$$

Proof. Let X have density (2.2), then

$$\big[1 - F (x)\big] \, \mathbb{E} [q_2 (X) \mid X \geq x] = \alpha \{-\log [G (x; \xi)]\}^{\beta}, \quad x > 0,$$

$$\big[1 - F (x)\big] \, \mathbb{E} [q_1 (X) \mid X \geq x] = \frac{\alpha\beta}{\beta + 1} \big\{ -\log [G (x; \xi)] \big\}^{\beta+1}, \quad x > 0,$$

and then

$$\eta (x) \; q_2 (x) - q_1 (x) = -\frac{1}{\beta + 1} q_2 (x) \Big\{ -\log [G (x; \xi)] \Big\} < 0 \quad \text{for } x > 0.$$

Conversely, if η is given as above, then

$$s'(x) = \frac{\eta'(x) \, q_2(x)}{\eta(x) \, q_2(x) - q_1(x)} = \beta \Big[\frac{g (x; \xi)}{G (x; \xi)} \Big] \Big\{ -\log [G (x; \xi)] \Big\}^{-1}, \quad x > 0,$$

and hence

$$s(x) = -\beta \log \Big(\{-\log [G (x; \xi)]\} \Big), \quad x > 0,$$

or

$$\mathrm{e}^{-s(x)} = \Big\{ -\log [G (x; \xi)] \Big\}^{\beta}, \quad x > 0.$$

Now, in view of Theorem 1, X has density function (2.2). $\square$

8.3. Corollary. *Let $X : \Omega \to (0, \infty)$ be a continuous random variable and let $q_2 (x)$ be as in Proposition 1. The pdf of X is (2.2) if and only if there exist functions q_1 and η defined in Theorem 1 satisfying the differential equation*

$$\frac{\eta'(x) \, q_2(x)}{\eta(x) \, q_2(x) - q_1(x)} = \beta \Big[\frac{g (x; \xi)}{G (x; \xi)} \Big] \big\{ -\log [G (x; \xi)] \big\}^{-1}, \quad x > 0.$$

Remark 1. **(a)** The general solution of the differential equation in Corollary 1 is obtained as follows:

$$\eta'(x) - \beta \Big[\frac{g (x; \xi)}{G (x; \xi)} \Big] \{-\log [G (x; \xi)]\}^{-1} = -\beta \, q_1(x) \, \Big[\tfrac{g(x;\xi)}{G(x;\xi)} \Big] \Big\{ -\log [G (x; \xi)] \Big\}^{-1} [q_2(x)]^{-1},$$

or

$$\frac{d}{dx} \big[\{-\log [G (x; \xi)]\}^{\beta} \, \eta(x) \big] = -\beta \, q_1(x) \, \Big[\tfrac{g(x;\xi)}{G(x;\xi)} \Big] \, \big\{ -\log [G (x; \xi)] \big\}^{\beta-1} [q_2(x)]^{-1}.$$

From the above equation, we obtain

$$\eta(x) = \big\{ -\log [G (x; \xi)] \big\}^{-\beta} \times \Big[-\int \beta \, q_1(x) \, \Big[\frac{g (x; \xi)}{G (x; \xi)} \Big] \Big\{ -\log [G (x; \xi)] \Big\}^{\beta-1} [q_2(x)]^{-1} \, dx + D \Big],$$

where D is a constant. One set of appropriate functions is given in Proposition 1 with $D = 0$.

(b) Clearly there are other triplets of functions (q_2, q_1, η) satisfying the conditions of Theorem 1. We present one such triplet in Proposition 1.

8.2. Characterization based on single function of the random variable. Here, we employ a single function ψ of X and state characterization results in terms of $\psi(X)$.

8.4. Proposition. *Let $X : \Omega \to (0, \infty)$ be a continuous random variable with cdf $F(x)$. Let $\psi(x)$ be a differentiable function on $(0, \infty)$ with $\lim_{x\to\infty} \psi(x) = 1$. Then for $\delta \neq 1$,*

$$\mathbb{E}[\psi(X) \,|X < x] = \delta\,\psi(x), \quad x \in (0, \infty)$$

if and only if

$$\psi(x) = F(x)^{\frac{1}{\delta}-1}, \quad x \in (0, \infty).$$

Proof. The proof is straightforward. □

Remark 2. For $\psi(x) = \mathrm{e}^{-\{-\log[G(x;\xi)]\}^{\beta}}$, $x \in (0, \infty)$ and $\delta = \frac{\alpha}{\alpha+1}$, Proposition 2 provides a cdf $F(x)$ given by (2.1).

8.3. Characterizations based on the hazard function. The hrf $h_F(x)$ of a twice differentiable distribution function $F(x)$ and pdf $f(x)$ satisfies the first order differential equation

$$\frac{h_F'(x)}{h_F(x)} - h_F(x) = q(x),$$

where $q(y)$ is an appropriate integrable function. Although this differential equation has an obvious form since

$$\frac{f'(x)}{f(x)} = \frac{h_F'(x)}{h_F(x)} - h_F(x), \tag{8.1}$$

for many univariate continuous distributions (8.1) seems to be the only differential equation in terms of the hrf. The goal of the characterization based on the hazard function is to establish a differential equation in terms of the hrf, which has a simple form as possible and is not of the trivial form (8.1). For some general families of distributions this may not be possible.

8.5. Proposition. *Let $X : \Omega \to (0, \infty)$ be a continuous random variable. The random variable X has pdf (2.2) (for $\beta = 1$) and $G(x) = (1 - \mathrm{e}^{-\lambda x})$ if and only if its hazard function $h_F(x)$ satisfies the differential equation*

$$\begin{aligned} h_F'(x) + \lambda\left(1 - \mathrm{e}^{-\lambda x}\right)^{-1} h_F(x) = {} & \frac{\alpha^2 \lambda^2 \mathrm{e}^{-2\lambda x}}{(1 - \mathrm{e}^{-\lambda x})^2}\, \mathrm{e}^{\alpha \log\left(1-\mathrm{e}^{-\lambda x}\right)} \\ & \times \left[1 - \mathrm{e}^{\alpha \log\left(1-\mathrm{e}^{-\lambda x}\right)}\right]^{-2}, \end{aligned} \tag{8.2}$$

with initial condition $h_F(0) = 0$.

Proof. The $f(x)$ has pdf (2.2), then clearly (8.2) holds. Now, if (8.2) holds, then

$$\frac{d}{dx}\left\{\mathrm{e}^{\lambda x}\left(1 - \mathrm{e}^{-\lambda x}\right) h_F(x)\right\} = \alpha\,\lambda\,\frac{d}{dx}\left\{\left(1 - \mathrm{e}^{\alpha \log\left[1-\mathrm{e}^{-\lambda x}\right]}\right)^{-1} + C\right\},$$

where C is an appropriate constant. Letting $C = -1$, we obtain from the above equation

$$h_F(x) = \frac{\alpha\lambda\,\mathrm{e}^{-\lambda x}}{(1 - e^{-\lambda x})}\left\{\mathrm{e}^{\alpha \log(1-\mathrm{e}^{-\lambda x})}\left[1 - \mathrm{e}^{\alpha \log(1-e^{-\lambda x})}\right]^{-1}\right\}.$$

Integrating both sides of the last equation from 0 to x, we arrive at

$$-\log\left[1 - F(x)\right] = -\log\left[1 - \mathrm{e}^{\alpha \log\left(1-e^{-\lambda x}\right)}\right],$$

from which, we obtain

$$1 - F(x) = 1 - \mathrm{e}^{\alpha \log(1-\mathrm{e}^{-\lambda x})}, \quad x \geq 0.$$

□

9. Estimation

We consider the estimation of the unknown parameters of the NWG family of distributions by the method of maximum likelihood. Let $x_1, \ldots, x_n$ be a sample of size n from the NWG family given by (2.2). The log-likelihood function for the vector of parameters $\boldsymbol{\Theta} = (\alpha, \beta, \boldsymbol{\xi})^\top$ can be expressed as

$$\begin{aligned} \ell(\boldsymbol{\Theta}) \quad = \quad & n \log \alpha + n \log \beta + \sum_{i=1}^{n} \log\, [g(x, \boldsymbol{\xi})] - \sum_{i=1}^{n} \log\, \{G(x, \boldsymbol{\xi})\} \\ & + (\beta - 1) \sum_{i=1}^{n} \log\, \{-\log [G(x, \boldsymbol{\xi})]\} - \alpha \sum_{i=1}^{n} \{-\log [G(x, \boldsymbol{\xi})]\}^{\beta}. \end{aligned}$$

The components of the score vector $U(\boldsymbol{\Theta})$ are given by

$$\begin{aligned} U_\alpha \quad &= \quad \frac{n}{\alpha} - \sum_{i=1}^{n} \{-\log [G(x, \boldsymbol{\xi})]\}^{\beta}, \\ U_\beta \quad &= \quad \frac{n}{\beta} - \sum_{i=1}^{n} \Big[\log\, \{-\log [G(x, \boldsymbol{\xi})]\} \Big] \\ & \quad\quad - \alpha \sum_{i=1}^{n} \Big[\{-\log [G(x, \boldsymbol{\xi})]\}^{\beta}\; \log\; \{-\log [G(x, \boldsymbol{\xi})]\} \Big], \\ U_{\boldsymbol{\xi}_k} \quad &= \quad \sum_{i=1}^{n} \left[\frac{\left(\frac{\partial\, g(x, \boldsymbol{\xi})}{\partial\, \boldsymbol{\xi}_k}\right)}{g(x, \boldsymbol{\xi})} \right] - \sum_{i=1}^{n} \left[\frac{\left(\frac{\partial\, G(x, \boldsymbol{\xi})}{\partial\, \boldsymbol{\xi}_k}\right)}{G(x, \boldsymbol{\xi})} \right] \\ & \quad\quad - (\beta - 1) \sum_{i=1}^{n} \left[\frac{\left(\frac{\partial\, G(x, \boldsymbol{\xi})}{\partial\, \boldsymbol{\xi}_k}\right)}{\{-\log [G(x, \boldsymbol{\xi})]\}\; G(x, \boldsymbol{\xi})} \right] \\ & \quad\quad + \alpha\beta \left[\sum_{i=1}^{n} \frac{\{-\log [G(x, \boldsymbol{\xi})]\}^{\beta-1}\; \left(\frac{\partial\, G(x, \boldsymbol{\xi})}{\partial\, \boldsymbol{\xi}_k}\right)}{G(x, \boldsymbol{\xi})} \right]. \end{aligned}$$

Setting U_α, U_β and $U_{\boldsymbol{\xi}_k}$ equal to zero and solving numerically these equations simultaneously yields the maximum likelihood estimates (MLEs) $\widehat{\boldsymbol{\Theta}} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\boldsymbol{\xi}})^\intercal$. The estimates can be obtained using the R language.

For interval estimation and hypothesis tests, we can use standard likelihood techniques based the observed information matrix, which can be obtained from the authors upon request.

10. Applications

We provide two applications to real life data sets to prove the flexibility of the Weibull-log-logistic (WLL) and Weibull-Weibull (WW) models presented in Section 3. The MLEs of the model parameters and the goodness-of-fit statistics are calculated for the WLL and WW models, and other competitive models. We compare these models with other Weibull-G models under the same baseline distribution, namely the WLL (BSC-WLL, ALF-WLL) and WW (BSC-WW, ALF-WW) models based on Bourguignon *et al.* (2014)'s generator $G(x)/[1-G(x)]$ and Alzaatreh *et al.* (2013)'s generator $-\log[1-G(x)]$. We note that the BSC-LL, ALF-LL, BSC-WW and ALF-WW models are not known in the literature. Further, we also compare the gamma exponentiated-exponential (GEE) (Ristić and Balakrishnan [26]) and exponential-exponential geometric (EEG) (Rezaei et al. [24]) models with the proposed and other competitive models. The density functions of the GEE and EEG distributions are, respectively, given by (for $x > 0$)

$$\begin{aligned} f_{GEE}(x;\lambda,\alpha,\theta) &= \frac{\alpha\,\theta}{\Gamma(\lambda)}\,\mathrm{e}^{-\theta\,x}\left[1-\mathrm{e}^{-\theta\,x}\right]^{\alpha-1}\left[-\alpha\log\left(1-\mathrm{e}^{-\theta\,x}\right)\right]^{\lambda-1}, \\ &\qquad \lambda,\alpha,\theta>0, \\ f_{EEG}(x;p,\alpha,\theta) &= \frac{\alpha\,\theta(1-p)\,\mathrm{e}^{-\theta\,x}}{[1-\mathrm{e}^{-\theta\,x}]^{\alpha-1}\,[1-p+p\,(1-\mathrm{e}^{-\theta\,x})^{\alpha}]^2}, \\ &\qquad 0<p<1,\quad \alpha,\theta>0. \end{aligned}$$

The first real data represents the breaking strength of 100 yarn reported by Duncan [17]. The second real data set corresponds to the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli reported by Bjerkedal [9].

The measures of goodness-of-fit including the log-likelihood function evaluated at the MLEs ($\hat{\ell}$), Akaike information criterion (AIC), Anderson-Darling (A^*), Cram´er-von Mises (W^*) and Kolmogorov-Smirnov (K-S), are calculated to compare the fitted models. The statistics A^* and W^* are described by Chen and Balakrishnan [10]. In general, the smaller the values of these statistics, the better the fit to the data. The required computations are carried out using the R software.

Table 1: MLEs and their standard errors (in parentheses) for the first data set.

Distribution	β	c	s	λ	α	θ	p
WLL	0.6612	25.5915	97.7523	-	-	-	-
	(0.1395)	(6.2313)	(1.0425)	-	-	-	-
BSC-WLL	4.7898	1.5601	105.0254	-	-	-	-
	(195.4617)	(63.6652)	(1.4938)	-	-	-	-
ALF-WLL	0.6528	25.9621	99.6537	-	-	-	-
	(0.1423)	(6.4490)	(1.0920)	-	-	-	-
GEE	-	-	-	20.4987	78.3734	0.0150	-
	-	-	-	(5.4222)	(11.2681)	(0.0022)	-
EEG	-	-	-	-	38.9807	0.0198	0.9974
	-	-	-	-	(5.8133)	(0.0015)	(0.0004)

Table 2: MLEs and their standard errors (in parentheses) for the second data set.

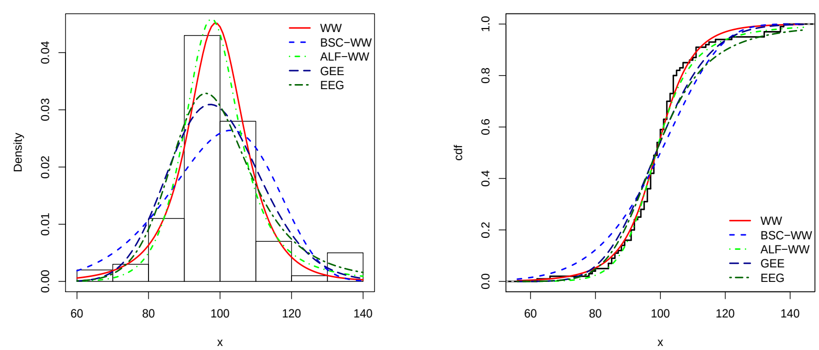
Distribution	β	γ	λ	α	θ	p
WW	2.6594	0.6933	0.0270	-	-	-
	(0.7129)	(0.1707)	(0.0193)	-	-	-
BSC-WW	11.1576	0.0881	0.4574	-	-	-
	(4.5449)	(0.0355)	(0.0770)	-	-	-
ALF-WW	1.7872	0.7795	0.0255	-	-	-
	(0.7821)	(0.3332)	(0.0400)	-	-	-
GEE	-	-	2.1138	2.6006	0.0083	-
	-	-	(1.3288)	(0.5597)	(0.0048)	-
EEG	-	-	-	2.5890	0.0004	0.9999
	-	-	-	(0.4820)	(0.0041)	(0.1036)

Table 3: The statistics $\hat{\ell}$, AIC, A^*, W^* and K-S for the fitted models to the first data set.

Distribution	$\hat{\ell}$	AIC	A^*	W^*	K-S	p-value (K-S)
WLL	-383.5896	773.1792	0.8402	0.1254	0.0805	0.5354
BSC-WLL	-404.7074	815.4147	4.7296	0.7951	0.1948	0.0010
ALF-WLL	-383.6181	773.2361	0.7432	0.1332	0.0888	0.4091
GEE	-392.7053	791.4106	2.3551	0.3976	0.1423	0.0348
EEG	-390.5435	787.0869	1.4894	0.2676	0.1442	0.0312

Table 4: The statistics $\hat{\ell}$, AIC, A^*, W^* and K-S for the fitted models to the second data set.

Distribution	$\hat{\ell}$	AIC	A^*	W^*	K-S	p-value (K-S)
WW	-390.2338	786.4676	0.7811	0.1427	0.1055	0.3994
BSC-WW	-397.8399	801.6797	2.4764	0.4494	0.1510	0.0749
ALF-WW	-397.1477	800.2953	2.3938	0.4348	0.1465	0.0911
GEE	-393.6235	793.2470	1.7208	0.3150	0.1347	0.1467
EEG	-389.9445	785.8890	0.5789	0.1047	0.0861	0.6282

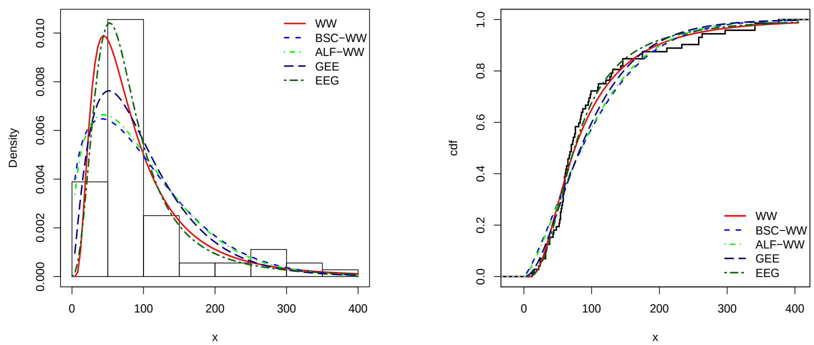


(a) Estimated pdfs

(b) Estimated cdfs

Figure 3. Plots of the estimated pdfs and cdfs of the WLL, BSC-WLL, ALF-WLL, GEE and EEG models.

The MLEs and the corresponding standard errors (in parentheses) of the model parameters are given in Tables 1 and 2. The numerical values of the statistics AIC, A^*, W^* and K-S are listed in Tables 3 and 4. The histograms of the two data sets and the estimated pdfs and cdfs of the proposed and competitive models are displayed in Figures 3 and 4. Based on the figures in Tables 2 and 4, we conclude that the new WLL and WW models provide adequate fits as compared to other Weibull-G models in both applications with small values for AIC, A^*, W^* and K-S, and large p-values. In Application 1, the proposed WLL model is much better than the BSC-WLL, GEE and EEG models, and a good alternative to the ALF-WLL model. In Application 2, the proposed WW model outperforms the BSC-WEW, ALF-WW and GEE models but it is not better than EEG model. Figures 3 and 4 also support the results in Tables 2 and 4.



(a) Estimated pdfs

(b) Estimated cdfs

Figure 4. Plots of the estimated pdfs and cdfs of the WW, BSC-WW, ALF-WW, GEE and EEG models.

11. Concluding remarks

In this paper, we propose and study the *new Weibull-G* (NWG) family. We investigate some of its mathematical properties including an expansion for the density function and explicit expressions for the quantile function, ordinary and incomplete moments, generating function, entropies, reliability and order statistics. Three useful characterizations, based on truncated moments, single function of the random variable and hazard function, are formulated for the NWG family. The advantage of the characterizations given here is that the cumulative distribution is not required to have a closed-form and are given in terms of an integral whose integrand depends on the solution of a first order differential equation. They can serve as a bridge between probability and differential equation. The maximum likelihood method is employed to estimate the model parameters. We fit two special models of the new family to two real data sets to demonstrate the flexibility of the proposed family. These special models can give better fits than other competing models. We hope that the new family and its generated models will attract wider application in areas such as engineering, survival and lifetime data, hydrology, economics, among others.

Acknowledgments

The authors would like to thank the editor and the two referees for careful reading and for their comments which greatly improved the article.

References

[1] Alexander, C., Cordeiro, G.M., Ortega, E.M.M. and Sarabia, J.M. *Generalized beta-generated distributions*, Computational Statistics and Data Analysis **56**, 1880–1897, 2012.

[2] Alizadeh, M., Emadi, M., Doostparast, M., Cordeiro, G.M., Ortega, E.M.M. and Pescim, R.R. *A new family of distributions: the Kumaraswamy odd log-logistic, properties and applications*, Hacettepa Journal of Mathematics and Statistics **44**, 1491–1512, 2015.

[3] Aljarrah, M.A., Lee, C. and Famoye, F. *On generating T-X family of distributions using quantile functions*, Journal of Statistical Distributions and Applications **1**, Article 2, 2014.

[4] Alzaatreh, A., Lee, C. and Famoye, F. *A new method for generating families of continuous distributions*, Metron **71**, 63–79, 2013.

[5] Alzaatreh, A., Lee, C. and Famoye, F. *T-normal family of distributions: A new approach to generalize the normal distribution*, Journal of Statistical Distributions and Applications **1**, Article 16, 2014.

[6] Alzaghal, A., Famoye, F. and Lee, C. *Exponentiated T-X family of distributions with some applications*, International Journal of Probability and Statistics **2**, 31–49, 2013.

[7] Amini, M., MirMostafaee, S.M.T.K. and Ahmadi, J. *Log-gamma-generated families of distributions*, Statistics **48**, 913–932, 2014.

[8] Bourguignon, M., Silva, R.B. and Cordeiro, G.M. *The Weibull–G family of probability distributions*, Journal of Data Science **12**, 53–68, 2014.

[9] Bjerkedal, T. *Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli.* American Journal of Hygiene **72**, 130–148, 1960.

[10] Chen, G. and Balakrishnan, N. *A general purpose approximate goodness-of-fit test*, Journal of Quality Technology **27**, 154–161, 1995.

[11] Cordeiro, G.M., Alizadeh, M. and Ortega, E.M.M. *The exponentiated half-logistic family of distributions: Properties and applications*, Journal of Probability and Statistics Article ID 864396, 21 pages, 2014.

[12] Cordeiro, G.M. and de Castro, M. *A new family of generalized distributions*, Journal of Statistical Computation and Simulation **81**, 883–893, 2011.

[13] Cordeiro, G.M. and Nadarajah, S. *Closed-form expressions for moments of a class of beta generalized distributions*, Brazilian Journal of Probability and Statistics **25**, 14–33, 2011.

[14] Cordeiro, G.M., Ortega, E.M.M. and da Cunha, D.C.C. *The exponentiated generalized class of distributions*, Journal of Data Science **11**, 1–27, 2013.

[15] Cordeiro, G.M., Ortega, E.M.M., Popović, B.V. and Pescim, R.R. *The Lomax generator of distributions: Properties, minification process and regression model*, Applied Mathematics and Computation **247**, 465–486, 2014.

[16] Cordeiro, G.M., Silva, G.O. and Ortega, E.M.M. *The beta extended Weibull distribution*, Journal of Probability and Statistical Science (Taiwan) **10**, 15–40, 2012.

[17] Duncan A.J. *Quality Control and Industrial Statistics*, fourth edition (Irwin- Homewood, 1974).

[18] Eugene, N., Lee, C. and Famoye, F. *Beta-normal distribution and its applications*, Communications in Statistics–Theory and Methods **31**, 497–512, 2002.

[19] Glänzel, W. *A characterization theorem based on truncated moments and its application to some distribution families*, In: Mathematical Statistics and Probability Theory, Volume B, pp. 75–84 (Reidel, Dordrecht, 1987).

[20] Gupta, R. C., Gupta, P. I. and Gupta, R. D. *Modeling failure time data by Lehmann alternatives*, Communications in statistics–Theory and Methods **27**, 887–904, 1998.

[21] Mudholkar, G.S., Srivastava, D.K. and Freimer, M. *The exponentiated Weibull family: A reanalysis of the bus-motor failure data.* Technometrics **37**, 436–445, 1995.

[22] Nadarajah, S., Cordeiro, G.M. and Ortega, E.M.M. *The Zografos-Balakrishnan–G family of distributions: Mathematical properties and applications*, Communications in Statistics–Theory and Methods **44**, 186–215, 2015.

[23] Nadarajah, S. and Kotz, S. *The exponentiated type distributions*, Acta Applicandae Mathematica **92**, 97–111, 2006.

[24] Rezaei, S., Nadarajah, S. and Tahghighnia, N. *A new three-parameter lifetime distribution*, Statistics **47**, 835–860, 2013.

[25] Rényi, A. *On measures of entropy and information*, In: 4th Berkeley Symposium on Mathematical Statistics and Probability **1**, 547–561, 1961.

[26] Ristić, M.M. and Balakrishnan, N. *The gamma-exponentiated exponential distribution*, Journal of Statistical Computation and Simulation **82**, 1191–1206, 2012.

[27] Shannon, C.E. *A mathematical theory of communication*, Bell System Technical Journal **27**, 379–432, 1948.

[28] Tahir, M.H., Cordeiro, G.M., Alzaatreh, A., Mansoor, M. and Zubair, M. *The Logistic-X family of distributions and its applications*, Communications in Statistics–Theory and Methods (to appear) 2016. Doi: 10.1080/03610926.2014.980516.

[29] Zografos K. and Balakrishnan, N. *On families of beta- and generalized gamma-generated distributions and associated inference*, Statistical Methodology **6**, 344–362, 2009.

ADVICE TO AUTHORS

Hacettepe Journal of Mathematics and Statistics publishes short to medium length research papers and occasional survey articles written in English. All papers are refereed to international standards.

Address for Correspondence

Editorial Office,
Hacettepe Journal of Mathematics and Statistics,
Hacettepe University,
Faculty of Science,
Department of Mathematics,
06532 Beytepe,
Ankara,
Turkey.

E-mail: hjms@hacettepe.edu.tr
Tel: Editors: + 90 312 297 7898
+ 90 312 297 7850 / 122
Associate Editor: + 90 312 297 7880
Fax : + 90 312 299 2017
Web Page : http//www.hjms.hacettepe.edu.tr

Advise to Authors : The style of articles should be lucid but concise. The text should be preceded by a short descriptive title and an informative abstract of not more than 100 words. Keywords or phrases and the 2010 AMS Classification should also be included.The main body of the text should be divided into numbered sections with appropriate headings. Items should be numbered in the form 2.4. Lemma, 2.5. Definition. These items should be referred to in the text using the form Lemma 2.4., Definition 2.5. Figures and tables should be incorporated in the text. A numbered caption should be placed above them. References should be punctuated according to the following examples, be listed in alphabetical order according to in the (first) author' s surname, be numbered consecutively and referred to in the text by the same number enclosed in square brackets. Only recognized abbreviations of the names of journals should be used.

[1] Banaschewski, B. *Extensions of topological spaces*, Canad. Math. Bull. **7** (1), 1–22, 1964.

[2] Ehrig, H. and Herrlich, H. *The construct PRO of projection spaces: its internal structure*, in: Categorical methods in Computer Science, Lecture Notes in Computer Science **393** (Springer-Verlag, Berlin, 1989), 286–293.

[3] Hurvich, C. M. and Tsai, C. L. *Regression and time series model selection in small samples*, Biometrika **76** (2), 297–307, 1989.

[4] Papoulis, A. *Probability random variables and stochastic process* (McGraw-Hill, 1965).

INSTRUCTIONS FOR AUTHORS

Preparation of Manuscripts : Manuscripts will be typeset using the LATEX typesetting system. Authors should prepare the article using the HJMS style before submission by e-mail. The style bundle contains the styles for the article and the bibliography as well as example of formatted document as pdf file. The authors should send their paper directly to hjms@hacettepe.edu.tr with the suggestion of area editor and potential reviewers for their submission. The editorial office may not use these suggestions, but this may help to speed up the selection of appropriate reviewers.

Copyright : No manuscript should be submitted which has previously been published, or which has been simultaneously submitted for publication elsewhere. The copyright in a published article rests solely with the Faculty of Science of Hacettepe University, and the paper may not be reproduced in whole in part by any means whatsoever without prior written permission.

Notice : The content of a paper published in this Journal is the sole responsibility of the author or authors, and its publication does not imply the concurrence of the Editors or the Publisher.