HACETTEPE UNIVERSITY
FACULTY OF SCIENCE
TURKEY

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

A Bimonthly Publication
Volume 45 Issue 5
2016

ISSN 1303 5010

HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Volume 45 Issue 5
October 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 5 (2016)

ISSN 1303 – 5010

KÜNYE
YAYININ ADI: HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS
YIL : 2016 SAYI : 45 - 5 AY : Ekim
YAYIN SAHİBİNİN ADI : H. Ü. Fen Fakültesi Dekanlığı adına Prof. Dr. A. Nurdan Özer
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 5 (2016)
ISSN 1303 – 5010

PUBLISHED BY HACETTEPE UNIVERSITY
FACULTY OF SCIENCE

CONTENTS

Mathematics

Statistics

MATHEMATICS

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1321 – 1328

On commutativity of prime gamma rings with derivation

Okan Arslan*† and Hatice Kandamar‡

Abstract

Let M be a weak Nobusawa Γ-ring and γ be a nonzero element of Γ. In this paper, we find a relation between Γ-rings and rings, and give some commutativity conditions on Γ-rings by using this relation. Also, we prove that any Γ-ring M in the sense of Nobusawa with a nonzero element γ in the center of M-ring Γ is γ-prime if and only if M is Γ-prime. As a consequence, we show that the semiprimeness (semisimpleness) of the ring $(M, +, \cdot_\gamma)$ for any $\gamma \in \Gamma$ implies the Γ-semiprimeness(Γ-semisimpleness) of the Γ-ring M.

Keywords: gamma ring, prime Γ-ring, k-derivation, commutativity, γ-radical.

2000 AMS Classification: AMS Primary 16N60; Secondary 16W25, 16Y99

Received : 10.03.2015 *Accepted* : 21.10.2015 *Doi* : 10.15672/HJMS.20164514286

1. Introduction

Let M and Γ be additive Abelian groups. M is said to be a Γ-ring in the sense of Barnes [3] if there exists a mapping $M \times \Gamma \times M \to M$ satisfying these two conditions for all $a, b, c \in M$, $\alpha, \beta \in \Gamma$:

(1) $(a + b)\, \alpha c = a\alpha c + b\alpha c$
$a(\alpha + \beta)c = a\alpha c + a\beta c$
$a\alpha\, (b + c) = a\alpha b + a\alpha c$

(2) $(a\alpha b)\, \beta c = a\alpha\, (b\beta c)$

In addition, if there exists a mapping $\Gamma \times M \times \Gamma \to \Gamma$ such that the following axioms hold for all $a, b, c \in M$, $\alpha, \beta \in \Gamma$:

(3) $(a\alpha b)\, \beta c = a\, (\alpha b\beta)\, c$

*Department of Mathematics, Adnan Menderes University, Aydın, Turkey,
Email: oarslann@icloud.com
†Corresponding Author.
‡Department of Mathematics, Adnan Menderes University, Aydın, Turkey,
Email: hkandamar@adu.edu.tr

(4) $a\alpha b = 0$ for all $a, b \in M$ implies $\alpha = 0$, where $\alpha \in \Gamma$

then M is called a Γ-ring in the sense of Nobusawa [17]. If a Γ-ring M in the sense of Barnes satisfies only the condition (3), then it is called weak Nobusawa Γ-ring [11].

We assume that all gamma rings in this paper are weak Nobusawa gamma ring unless otherwise specified.

Let M be a Γ-ring. M is said to be a Γ-prime gamma ring if $a\Gamma M\Gamma b = 0$ with $a, b \in M$ implies either $a = 0$ or $b = 0$ [15]. M is Γ-simple if $M\Gamma M \neq 0$ and M has no ideals (0) and M itself [15].

$C_M = \{\alpha \in \Gamma \mid \alpha m\beta = \beta m\alpha, \forall m \in M, \beta \in \Gamma\}$ is called the center of M-ring Γ and $C_\gamma = \{c \in M \mid c\gamma m = m\gamma c, \forall m \in M\}$ with $\gamma \in \Gamma$ is called the γ-center of Γ-ring M.

Recall that from [9], an additive mapping $d : M \to M$ is called a derivation on M if $d(a\alpha b) = d(a)\alpha b + a\alpha d(b)$ for all $a, b \in M$ and $\alpha \in \Gamma$. Note that $d = 0$ when d is defined on a prime weak Nobusawa Γ-ring M. So, in this paper we consider k-derivations that has been defined by Kandamar [10] on any gamma ring M.

In this work, we first obtain some commutativity conditions on the γ-prime Γ-ring M with k-derivations and prove that M is γ-prime if and only if M is Γ-prime where γ is a nonzero element in the center of M-ring Γ in the sense of Nobusawa. Then, we also show that if there exists a nonzero element γ in C_M in a Nobusawa Γ-ring M, then (0) is Γ-prime ideal if and only if (0) is γ-prime ideal. Finally, we study the relation between semiprimeness (semisimpleness) of the ring $(M, +, \cdot_\gamma)$ and Γ-semiprimeness (Γ-semisimpleness) of the Γ-ring M where $\gamma \in \Gamma$.

2. Relation between Γ-rings and rings up to γ

We now give some definitions that have been firstly defined by Arslan and Kandamar in [1].

2.1. Definition. Let M be a Γ-ring, γ be a nonzero element of Γ and I be an additive subgroup of M.

(i) M is said to be γ-commutative if $x\gamma y = y\gamma x$ for all $x, y \in M$.

(ii) I is said to be a γ-subring of M if $x\gamma y \in I$ for all $x, y \in I$.

(iii) I is said to be a γ-left ideal(resp. γ-right ideal) of M if $m\gamma a \in I$(resp. $a\gamma m \in I$) for all $m \in M$, $a \in I$. If I is both γ-left and γ-right ideal then I is called a γ-ideal of M.

(iv) I is said to be a γ-prime ideal if $A\gamma B$ implies $A \subseteq I$ or $B \subseteq I$ for any γ-ideals A and B of M.

(v) I is said to be a γ-Lie ideal of M if $[x, m]_\gamma = x\gamma m - m\gamma x \in I$ for all $x \in I$ and $m \in M$.

2.2. Definition. A Γ-ring M is called a γ-prime gamma ring if there exists a nonzero element γ in Γ such that $a\gamma M\gamma b = 0$ with $a, b \in M$ implies either $a = 0$ or $b = 0$.

2.3. Definition. A Γ-ring M is called a γ-simple if $M\gamma M \neq 0$ and M has no γ-ideal besides the (0) and itself.

2.4. Lemma. *Let M be a Γ-ring. Then the following holds:*

(i) *If M is a γ-prime gamma ring, then M is Γ-prime.*

(ii) *If M is a γ-simple gamma ring, then M is Γ-simple.*

Proof. (i) Let M be a γ-prime gamma ring and $a\Gamma M\Gamma b = 0$ for any $a, b \in M$. Therefore, we have $a\gamma M\gamma b = 0$. Since M is a γ-prime gamma ring, we get $a = 0$ or $b = 0$. Hence, the γ-primeness of M implies the Γ-primeness of M.

(ii) It is clear from the definitions of γ-simple and Γ-simple gamma rings.

□

2.5. Proposition. *Let M be a Γ-ring and γ be a nonzero element of Γ. Then the Abelian group M with a binary operation $\cdot_\gamma$ defined by $a \cdot_\gamma b = a\gamma b$ for all $a, b \in M$ is a ring.*

Proof. It is clear from the definition of the gamma ring. □

According to the Proposition 2.5, the Abelian group M can be made into a ring by defining binary operations for all $\gamma \in \Gamma$. We denote this ring by $(M, +, \cdot_\gamma)$.

It is obvious that a γ-ideal of a Γ-ring M is an ideal of the ring $(M, +, \cdot_\gamma)$. Conversely, every ideal of the ring $(M, +, \cdot_\gamma)$ is a γ-ideal of the Γ-ring M. Similarly γ-Lie ideals of the Γ-ring M and Lie ideals of the ring $(M, +, \cdot_\gamma)$ is same. Also, if d is a k-derivation of the Γ-ring M and $k(\gamma) = 0$, then d is a derivation of the ring $(M, +, \cdot_\gamma)$. Thus, we can adapt all of the known results for the ring $(M, +, \cdot_\gamma)$ to the Γ-ring M. For instance, the commutativity of the ring $(M, +, \cdot_\gamma)$ is equal to the γ-commutativity of the Γ-ring M. Similarly one can say the primeness (semiprimeness) of the ring $(M, +, \cdot_\gamma)$ is the same as the γ-primeness (γ-semiprimeness) of the Γ-ring M. We give some results below.

2.6. Theorem. *Let M be a γ-prime gamma ring and d_1, d_2 be nonzero k_1, k_2-derivations of M such that $k_1(\gamma) = k_2(\gamma) = 0$ respectively. If char$M \neq 2$ and $d_1 d_2$ is $k_1 k_2$-derivation of M, then $d_1 = 0$ or $d_2 = 0$.*

Proof. By the hypothesis $d_1 \neq 0$, $d_2 \neq 0$ and $d_1 d_2$ are derivations of the prime ring $(M, +, \cdot_\gamma)$. Also the characteristic of the ring $(M, +, \cdot_\gamma)$ is different from 2. Therefore by [18, Theorem 1] one of the derivations d_1 and d_2 is zero in the ring $(M, +, \cdot_\gamma)$. □

2.7. Corollary. *Let M be a γ-prime gamma ring of characteristic not 2 and d be a 0-derivation of M such that $d^2 = 0$. Then $d = 0$.*

Proof. Let M is a γ-prime gamma ring. Then M is a Γ-prime gamma ring by Lemma 2.4. Since $d^2 = 0$ is a derivation on M, we get $d = 0$ by Theorem 2.6. □

2.8. Theorem. *Let M be a gamma ring and d be a k-derivation of M such that $k(\gamma) = 0$ and $d^3 \neq 0$. Then the γ-subring generated by $d(m)$ for all m in M contains a nonzero γ-ideal of M.*

Proof. Since d is a derivation of the ring $(M, +, \cdot_\gamma)$ and $d^3 \neq 0$, the subring generated by $d(m)$ for all m in M contains a nonzero ideal of $(M, +, \cdot_\gamma)$ by [6, Theorem 1]. Therefore the γ-subring generated by $d(m)$ for all m in M contains a nonzero γ-ideal of M. □

2.9. Corollary. *Let M be a Γ-ring, d be a nonzero 0-derivation on M such that $d^3 \neq 0$. Then, the subring A of M generated by all $d(a\alpha b)$, with $\alpha \in \Gamma$ and $a, b \in M$, contains a nonzero ideal of M.*

Another proof of Corollary 2.9 can be found in [19].

2.10. Theorem. *Let M be a γ-prime gamma ring and d be a nonzero k-derivation of M such that $k(\gamma) = 0$. Then M is γ-commutative if one of the following conditions holds:*

(i) *$[a, d(a)]_\gamma \in C_\gamma$ for all $a \in M$.*
(ii) *char$M \neq 2$ and $[d(M), d(M)]_\gamma \subset C_\gamma$.*
(iii) *char$M \neq 2$ and $d^2(M) \subset C_\gamma$.*
(iv) *d_1, d_2 are nonzero k_1, k_2-derivations of M such that $k_1(\gamma) = k_2(\gamma) = 0$ respectively, char$M \neq 2$ and $d_1 d_2(M) \subset C_\gamma$.*

Proof. (i) By the hypothesis d is a nonzero derivation of the prime ring $(M,+,\cdot_\gamma)$. Since $[a,d(a)]$ is in the center of the ring $(M,+,\cdot_\gamma)$ for all $a\in M$, the ring $(M,+,\cdot_\gamma)$ is commutative by [18, Theorem 2]. Therefore the gamma ring M is γ-commutative since commutativity of $(M,+,\cdot_\gamma)$ requires γ-commutativity of Γ-ring M.

(ii) By the hypothesis d is a nonzero derivation of the prime ring $(M,+,\cdot_\gamma)$, the characteristic of the ring M is different from 2 and $[d(M),d(M)]_\gamma$ is contained in the center of the ring M. Hence M is commutative as a ring by [13, Theorem 2]. Therefore M is γ-commutative.

(iii) By the hypothesis d is a nonzero derivation of the prime ring $(M,+,\cdot_\gamma)$, the characteristic of the ring M is different from 2 and $d^2(M)$ is contained in the center of the ring M. Hence M is commutative as a ring by [13, Theorem 3]. Therefore M is γ-commutative.

(iv) By the hypothesis d_1 and d_2 are nonzero derivations of the prime ring $(M,+,\cdot_\gamma)$. Also the characteristic of the ring $(M,+,\cdot_\gamma)$ is different from 2 and $d_1d_2(M)$ is contained in the center of the ring M. Hence M is commutative as a ring by [13, Theorem 4]. Therefore M is γ-commutative.

□

2.11. Corollary. *Let M be a γ-prime gamma ring for all nonzero elements γ in Γ and d be a nonzero 0-derivation on M. Then M is Γ-commutative if one of the following conditions holds:*

(i) $[a,d(a)]_\gamma\in C_\gamma$ *for all* $a\in M$ *and* $\gamma\in\Gamma$.

(ii) $charM\neq 2$ *and* $[d(M),d(M)]_\gamma\subset C_\gamma$ *for all* $\gamma\in\Gamma$.

(iii) $charM\neq 2$ *and* $d^2(M)\subset C_\gamma$ *for all* $\gamma\in\Gamma$.

(iv) d_1, d_2 *are nonzero 0-derivations of* M, $charM\neq 2$ *and* $d_1d_2(M)\subset C_\gamma$ *for all* $\gamma\in\Gamma$.

2.12. Theorem. *Let M be a γ-prime gamma ring of characteristic not 2 and U be a γ-Lie ideal of M. If $U\nsubseteq C_\gamma$, then there exists a γ-ideal K of M such that $[K,M]_\gamma\subseteq U$ but $[K,M]_\gamma\nsubseteq C_\gamma$.*

Proof. U is a Lie ideal of the prime ring $(M,+,\cdot_\gamma)$ that is not contained in the center of the ring M and the characteristic of the ring M is different from 2 by hypothesis. Hence, there exists an ideal K of $(M,+,\cdot_\gamma)$ such that $[K,M]\subseteq U$ and $[K,M]$ is not contained in the center of the $(M,+,\cdot_\gamma)$ by [4, Lemma 1]. Therefore, there exists an ideal K of Γ-ring M such that $[K,M]_\gamma\subseteq U$ but $[K,M]_\gamma\nsubseteq C_\gamma$. □

2.13. Theorem. *Let M be a γ-prime gamma ring of characteristic not 2 and U be a γ-Lie ideal of M such that $U\nsubseteq C_\gamma$. If d_1, d_2 are nonzero k_1, k_2-derivations of M such that $k_1(\gamma)=k_2(\gamma)=0$ respectively and $d_1d_2(U)=0$, then $d_1=0$ or $d_2=0$.*

Proof. By the hypothesis, d_1 and d_2 are nonzero derivations of the prime ring $(M,+,\cdot_\gamma)$ and U is a Lie ideal of M that is not contained in the center of the ring M. Also the characteristic of the ring $(M,+,\cdot_\gamma)$ is different from 2 and $d_1d_2(U)=0$. Hence $d_1=0$ or $d_2=0$ by [4, Theorem 4]. □

2.14. Theorem. *Let M be a γ-prime gamma ring of characteristic not 2, U be a γ-Lie ideal of M and d be a k-derivation of M such that $k(\gamma)=0$. Then U is contained in the γ-center of M if one of the following conditions holds:*

(i) $d^2(U)=0$.

(ii) $d\neq 0$ *and* $d^2(U)\subset C_\gamma$.

(iii) d_1, d_2 *are nonzero* k_1, k_2*-derivations of* M *such that* $k_1(\gamma) = k_2(\gamma) = 0$ *respectively and* $d_1 d_2(U) \subset C_\gamma$.

Proof. It is similar to the proof of Theorem 2.10. □

2.15. Corollary. *Let* M *be a* γ*-prime gamma ring of characteristic not 2 for all nonzero elements* γ *in* Γ, U *be a* γ*-Lie ideal of* M *and* d *be a* 0*-derivation of* M. *Then* U *is contained in the center of* M *if one of the following conditions holds:*

(i) $d^2(U) = 0$.
(ii) $d \neq 0$ *and* $d^2(U) \subset C_\gamma$ *for all* $\gamma \in \Gamma$.
(iii) d_1, d_2 *are nonzero* 0*-derivations of* M *and* $d_1 d_2(U) \subset C_\gamma$ *for all* $\gamma \in \Gamma$.

2.16. Theorem. *Let* M *be a* γ*-prime gamma ring of characteristic not 2 and* U *be a* γ*-Lie ideal of* M *such that* $U \nsubseteq C_\gamma$. *If* d_1 *and* d_2 *are nonzero* k_1 *and* k_2*-derivations of* M *such that* $k_1(\gamma) = k_2(\gamma) = 0$ *respectively and* $d_1 d_2(U) \subset C_\gamma$, *then* $d_1 = 0$ *or* $d_2 = 0$.

Proof. By the hypothesis d_1 and d_2 are nonzero derivations of the prime ring $(M, +, \cdot_\gamma)$ and U is a Lie ideal of M that is not contained in the center of M. Also the characteristic of the ring $(M, +, \cdot_\gamma)$ is different from 2 and $d_1 d_2(U)$ is contained in the center of M. Hence $d_1 = 0$ or $d_2 = 0$ by [2, Theorem 6]. □

3. γ-Radicals of Gamma Rings

Radicals of Γ-rings has been investigated by a number of authors. Barnes [3] defined prime radicals and proved some properties for gamma rings by methods similar to those of McCoy[16]. Coppage and Luh [5] introduced the notions of Jacobson radical, Levitzki nil radical, nil radical and strongly nilpotent radical for Γ-rings and Barnes' prime radical was studied further. Kyuno [12] also studied prime radicals of gamma rings and showed relations between radicals of gamma rings and radicals of its operator rings.

We define γ-prime radical, strongly γ-nilpotent radical, γ-Levitzki nil radical and γ-Jacobson radical for Γ-rings and show their relations with the radicals of Γ-rings in the literature.

Let M be a gamma ring and $S \subseteq M$. S is said to be a γ-m-system if $S = \emptyset$ or $(a)_\gamma \gamma (b)_\gamma \cap S \neq \emptyset$ for any $a, b \in M$. Here, $(a)_\gamma$ is the set of all elements of the form $ka + m\gamma a + a\gamma x + \sum_{i=1}^{n} u_i \gamma a \gamma v_i$ for $k \in \mathbb{Z}$, $n \in \mathbb{N}$, $m, x, u_i, v_i \in M$.

Proofs of the below results are obvious from the relation given in Section 2. So we omit their proofs.

3.1. Proposition. *Let* M *be a gamma ring and* P *be a* γ*-ideal of* M. *Then* P *is a* γ*-prime ideal if and only if the complement of* P *is a* γ*-*m*-system.*

Let A be a γ-ideal of a Γ-ring M. Then the set of all elements m in M such that every γ-m-system in M which contains m meets A is called γ-prime radical of the γ-ideal A and is denoted by $\mathfrak{B}_\gamma(A)$. γ-prime radical of zero γ-ideal is called γ-prime radical of the Γ-ring M and is denoted by $\mathfrak{B}_\gamma(M)$. In fact, the prime radical of the ring $(M, +, \cdot_\gamma)$ is equal to $\mathfrak{B}_\gamma(M)$.

3.2. Theorem. *If* A *is a* γ*-ideal in the* Γ*-ring* M, *then* $\mathfrak{B}_\gamma(A)$ *coincides with the intersection of all the* γ*-prime ideals in* M *which contain* A.

3.3. Corollary. γ*-prime radical of a* Γ*-ring* M *is the intersection of all the* γ*-prime ideals in* M.

An element a in M is called strongly γ-nilpotent if there exists a positive integer n such that $(a\gamma)^n a = 0$. A subset L of M is called strongly γ-nil if all of the elements in L are strongly γ-nilpotent. A subset S of M is called strongly γ-nilpotent if there exists a positive integer m such that $(S\gamma)^m S = 0$.

The strongly γ-nilpotent radical of M is the sum of all strongly γ-nilpotent ideals of M and is denoted by $\mathfrak{S}_\gamma(M)$.

3.4. Proposition. *If A and B are any strongly γ-nilpotent ideals in a Γ-ring M, then $A+B$ is also a strongly γ-nilpotent ideal in M.*

3.5. Corollary. *The strongly γ-nilpotent radical of a Γ-ring M is a strongly γ-nil ideal in M.*

A subset S of M is called γ-locally nilpotent if for any finite subset F of S there exists a positive integer n such that $(F\gamma)^n F = 0$.

The γ-Levitzki nil radical of M is the sum of all γ-locally nilpotent ideals of M and is denoted by $\mathfrak{L}_\gamma(M)$.

An element a in M is called γ-right quasi regular if there exist $b \in M$ such that $a + b + a\gamma b = 0$. A subset S of M is called γ-right quasi regular if all of the elements in S are γ-right quasi regular.

The γ-Jacobson radical of M is the set of all $a \in M$ such that the principal γ-ideal generated by a is γ-right quasi regular and is denoted by $\mathfrak{J}_\gamma(M)$. In fact, the Jacobson radical of the ring $(M, +, \cdot_\gamma)$ is equal to $\mathfrak{J}_\gamma(M)$.

4. Main Results

Not all of the properties of a ring holds for a gamma ring. For example, let d be a k-derivation of γ-prime gamma ring M of characteristic not 2. If $k(\gamma) \neq 0$, then the hypothesis $d^2 = 0$ does not imply $d = 0$.

4.1. Example. Let $M = \left\{ \begin{pmatrix} a & b & a \\ c & r & c \end{pmatrix} \mid a, b, c, r \in \mathbb{Z} \right\}$, Γ be the set of all 3×2 matrices over $\mathbb{Z}$ and $\gamma = \begin{pmatrix} 0 & 0 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} \in \Gamma$. Then, M is a γ-prime Γ-ring of characteristic not 2. Define $d : M \to M$, $d\begin{pmatrix} a & b & a \\ c & r & c \end{pmatrix} = \begin{pmatrix} -b & 0 & -b \\ -r & 0 & -r \end{pmatrix}$ and $k : \Gamma \to \Gamma$, $k\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ u_{11}+u_{31} & u_{12}+u_{32} \\ 0 & 0 \end{pmatrix}$.

It can be shown that d is a k-derivation and $k(\gamma) \neq 0$. Moreover, it is easy to see that $d \neq 0$ but $d^2 = 0$.

This example also shows that if d is a k-derivation on the Γ-prime gamma ring of characteristic not 2 such that $d^2 = 0$, then d may not be the zero derivation. In such a case, k^2 must be equal to zero as proved in the next theorem.

4.2. Theorem. *Let M be a γ-prime gamma ring in the sense of Nobusawa of characteristic not 2 and d be a k-derivation. If $d^2 = 0$, then either $d = 0$ or $k^2 = 0$.*

Proof. Let $k(\gamma) = 0$. Then, the k-derivation d on M is also a derivation for the ring $(M, +, \cdot_\gamma)$. Therefore, $d = 0$ by [18, Theorem 1]. Now, let $k(\gamma) \neq 0$. By hypothesis we have $d^2(d(x)\beta d(y)) = 0$ for all $x, y \in M$ and $\beta \in \Gamma$. Expanding this we get $d(x)k^2(\beta)d(y) = 0$. Replacing β by $\beta d(z)\alpha$ we have $d(x)k(\beta)d(z)k(\alpha)d(y) = 0$ since char$M \neq 2$. Replacing β by $\beta d(m)\delta$ we get $d(x)k(\beta)d(m) = 0$ since M is Γ-prime

Nobusawa Γ-ring by Lemma 2.4. If we replace x by $d(x)\alpha y$ in the last equation we have $d(x)k(\alpha)y = 0$ or $zk(\beta)d(m) = 0$. If $d(x)k(\alpha)y = 0$, then replacing α by $\alpha mk(\delta)$ we get $d(x)\alpha mk^2(\delta)y = 0$ for all $x, m, y \in M$ and $\alpha, \delta \in \Gamma$. Then, $d = 0$ or $k^2 = 0$ since M is a prime Nobusawa Γ-ring. If we consider the case $zk(\beta)d(m) = 0$, same result can be obtained similarly.

□

4.3. Theorem. *Let M be a Γ-ring in the sense of Nobusawa and γ be a nonzero element of Γ. If $\gamma \in C_M$, then M is γ-prime gamma ring if and only if M is Γ-prime.*

Proof. If M is γ-prime gamma ring then M is Γ-prime by Lemma 2.4. Let M is a Γ-prime gamma ring, $a\gamma M\gamma b = 0$ for any $a, b \in M$ and $a \neq 0$. Then we have $a\Gamma M\gamma M\gamma b = 0$. Since M is a Γ-prime $M\gamma M\gamma b = 0$. Thus $M\gamma M\Gamma b\gamma M = 0$. Hence we get $b = 0$ since M is a Γ-prime gamma ring and $\gamma \in C_M$. Therefore, M is γ-prime. □

4.4. Theorem. *The prime radical of a Γ-ring M is contained in γ-prime radical of M.*

Proof. Let x be an element of $\mathfrak{B}(M)$, the prime radical of M. Suppose that $x \notin \mathfrak{B}_\gamma(M)$. Then, there is a γ-m-system S which contains x such that $0 \notin S$. Therefore, there is an m-system in M which contains x but not contains 0 since S is also an m-system. This contradicts with $x \in \mathfrak{B}(M)$. Hence, if x is an element of $\mathfrak{B}(M)$, then x must be in $\mathfrak{B}_\gamma(M)$. □

4.5. Theorem. *The strongly nilpotent radical of a Γ-ring M is contained in strongly γ-nilpotent radical of M.*

Proof. It is easy to see that a strongly nilpotent ideal of M is also a strongly γ-nilpotent ideal. Therefore, $\mathfrak{S}(M)$, the strongly nilpotent radical of M, is contained in $\mathfrak{S}_\gamma(M)$. □

4.6. Theorem. *The Levitzki nil radical of a Γ-ring M is contained in γ-Levitzki nil radical of M.*

Proof. It is easy to see that a locally nilpotent ideal of M is also a γ-locally nilpotent ideal. Therefore, $\mathfrak{L}(M)$, the Levitzki nil radical of M, is contained in $\mathfrak{L}_\gamma(M)$. □

4.7. Theorem. *The Jacobson radical of a Γ-ring M is contained in γ-Jacobson radical of M.*

Proof. It is easy to see that a right quasi regular element of M is also a γ-right quasi regular. Therefore, $\mathfrak{J}(M)$, the Jacobson radical of M, is contained in $\mathfrak{J}_\gamma(M)$. □

4.8. Corollary. *Let M be a Γ-ring.*

(i) *If the ring $(M, +, \cdot_\gamma)$ for any $\gamma \in \Gamma$ is semiprime, then the Γ-ring M is Γ-semiprime.*
(ii) *If the ring $(M, +, \cdot_\gamma)$ for any $\gamma \in \Gamma$ is semisimple, then the Γ-ring M is Γ-semisimple.*

References

[1] Arslan, O., Kandamar, H., *γ-Lie structures in γ-prime gamma rings with derivations*, J. Algebra Comb. Discrete Appl. **2** (1), 25-37, 2015.
[2] Awtar, R., *Lie and Jordan structure in prime rings with derivations*, Proc. Amer. Math. Soc. **41**, 67-74, 1973.
[3] Barnes, W. E., *On the Γ-rings of Nobusawa*, Pacific J. Math. **18** (3), 411-422, 1966.
[4] Bergen, J., Kerr, J.W., Herstein, I.N., *Lie ideals and derivations of prime rings*, J. Algebra **71**, 259-267, 1981.
[5] Coppage, W.E., Luh, J., *Radicals of gamma rings*, J. Math. Soc. Japan **23** (1), 40-52, 1971.

[6] Herstein, I. N., *A note on derivations*, Canad. Math. Bull. **21** (3), 369-370, 1978.
[7] Herstein, I. N., *A note on derivations II*, Canad. Math. Bull. **22** (4), 509-511, 1979.
[8] Herstein, I. N., *Topics in Ring Theory*, The Univ. of Chicago Press, 132p., Chicago, 1969.
[9] Jing, F. J., *On derivations of* Γ*-rings*, Qu fu Shifan Daxue Xuebeo Ziran Kexue Ban **13** (4), 159-161, 1987.
[10] Kandamar, H., *The k-Derivation of a Gamma-Ring*, Turk. J. Math. **23** (3), 221-229, 2000.
[11] Kyuno, S., *Gamma Rings*, Hadronic Press, Palm Habor, 1991.
[12] Kyuno, S., *On the radicals of gamma rings*, Osaka J. Math. **12**, 639-645, 1975.
[13] Lee, P. H., Lee T. K., *On derivations of prime rings*, Chinese J. Math. **9** (2), 107-110, 1981.
[14] Lee, P. H., Lee T. K., *Lie ideals of prime rings with derivations*, Bull. Inst. Math. Acad. Sinica **11** (1), 75-80, 1983.
[15] Luh, J., *On the theory of simple* Γ*-rings*, Michigan Math. J. **16** (1), 65-75, 1969.
[16] McCoy, N.H., *The Theory of Rings*, Chelsea Publishing Company, New York, 1973.
[17] Nobusawa, N., *On a generalization of the ring theory*, Osaka J. Math. **1**, 81-89, 1964.
[18] Posner, E. C., *Derivations in prime rings*, Proc. Amer. Math. Soc. **8**, 1093-1100, 1957.
[19] Rakhimov, I. S., Dey K. K., Paul A. C., *On commutativity of completely prime gamma-rings*, Malays. J. Math. Sci. **7** (2), 283-295, 2013.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1329 – 1334

Asymptotic behavior of associated primes of certain ext modules

Kamal Bahmanpour*, Reza Naghipour† and Mehdi Sehatkhah‡

Abstract

Let R be a commutative Noetherian ring, I an ideal of R and M a finitely generated R-module. It is shown that, whenever I is principal, then for each integer i the set of associated prime ideals $\mathrm{Ass}_R \mathrm{Ext}^i_R(R/I^n, M)$, $n = 1, 2, \ldots$, becomes independent of n, for large n.

Keywords: Artinian module, Associated prime, Depth, Noetherian module.

2000 AMS Classification: 13E05.

Received : 20.07.2015 *Accepted* : 24.11.2015 *Doi* : 10.15672/HJMS.20164514288

1. Introduction

Let R denote a commutative Noetherian ring (with identity), I an ideal of R, and M a finitely generated R-module. In [7] L.J. Ratliff, Jr., conjectured about the asymptotic behaviour of $\mathrm{Ass}_R R/I^n$ when R is a Noetherian domain. Subsequently, M. Brodmann [1] showed that $\mathrm{Ass}_R M/I^n M$ is ultimately constant for large n. In [6], Melkersson and Schenzel asked whether the sets $\mathrm{Ass}_R \mathrm{Ext}^i_R(R/I^n, M)$ become stable for sufficiently large n. The aim of this paper is to show that, for all $i \geq 0$, the sets of prime ideals $\mathrm{Ass}_R \mathrm{Ext}^i_R(R/I^n, M)$, $n = 1, 2, \ldots$, becomes independent of n, for large n, whenever I is principal, which is an affirmative answer to the above question in the case I is principal. Also, it is shown that, if I is generated by an R-regular sequence and $\mathrm{Ext}^i_R(R/I, M)$ is Artinian, then the set $\cup_{n=1}^{\infty} \mathrm{Ass}_R \mathrm{Ext}^{i+1}_R(R/I^n, M)$ is finite.

For any R-module L, the set $\{\mathfrak{p} \in \mathrm{Ass}_R L \mid \dim R/\mathfrak{p} = \dim L\}$ is denoted by $\mathrm{Assh}_R L$.

*Department of Mathematics, Faculty of Sciences, University of Mohaghegh Ardabili, 56199-11367, Ardabil, Iran.
Email : bahmanpour.k@gmail.com
†Department of Mathematics, University of Tabriz, Tabriz, Iran.
Email : naghipour@ipm.ir
Email : naghipour@tabrizu.ac.ir Corresponding author.
‡Roshdiyeh Higher Education Institution, Tabriz, Iran.
Email : m_sehhat@pnu.ac.ir

2. The Results

2.1. Lemma. Let R be a Noetherian ring, I an ideal of R and M a finitely generated R-module. Then the sequence $\mathrm{Ass}_R\mathrm{Ext}^1_R(R/I^n, M)$ becomes eventually constant, for large n.

Proof. See [4, Corollary 2.3]. □

2.2. Lemma. Let x be an element of the Noetherian ring R. Let M and N be two finitely generated R-modules such that $\mathrm{pd}(N) = t < \infty$. Then for each $i \geq t+2$ and for all large k,

$$\mathrm{Ass}_R\,\mathrm{Ext}^i_R(N/x^kN, M) = \mathrm{Ass}_R\,\mathrm{Ext}^{i-1}_R(N/\Gamma_{Rx}(N), M),$$

and so the sets $\mathrm{Ass}_R\mathrm{Ext}^i_R(N/x^kN, M)$ are eventually constant.

Proof. Suppose that $i \geq t+2$. As, N is finitely generated, it follows that there is an integer n such that

$$\Gamma_{Rx}(N) := \bigcup_{i=0}^{\infty}(0 :_M Rx^i) = (0 :_N x^n) = (0 :_N x^{n+1}) = \cdots .$$

Now we claim that for any $k \geq n$,

$$\mathrm{Ext}^i_R(N/x^kN, M) \cong \mathrm{Ext}^{i-1}_R(N/\Gamma_{Rx}(N), M).$$

To do this, as $(0 :_N x^k) = \Gamma_{Rx}(N)$, it follows that $x^kN \cong N/\Gamma_{Rx}(N)$. Therefore for all $j \geq 0$ we have

$$\mathrm{Ext}^j_R(x^kN, M) \cong \mathrm{Ext}^j_R(N/\Gamma_{Rx}(N), M),$$

for all $k \geq n$. Now the exact sequence

$$0 \longrightarrow x^kN \longrightarrow N \longrightarrow N/x^kN \longrightarrow 0,$$

implies that

$$\mathrm{Ext}^i_R(N/x^kN, M) \cong \mathrm{Ext}^{i-1}_R(x^kN, M) \cong \mathrm{Ext}^{i-1}_R(N/\Gamma_{Rx}(N), M),$$

(Note that $\mathrm{pd}(N) = t$ and $i \geq t+2$.) Hence we have

$$\mathrm{Ass}_R\,\mathrm{Ext}^i_R(N/x^kN, M) = \mathrm{Ass}_R\,\mathrm{Ext}^{i-1}_R(N/\Gamma_{Rx}(N), M),$$

for all $k \geq n$, as required. □

2.3. Theorem. Let R be a Noetherian ring and let x be an element of R. Let M be a finitely generated R-module and i a non-negative integer. Then the sequence

$$\mathrm{Ass}_R\mathrm{Ext}^i_R(R/Rx^k, M),$$

of associated primes is ultimately constant for large k, and if $i \geq 2$, then

$$\mathrm{Ass}_R\,\mathrm{Ext}^i_R(R/Rx^k, M) = \mathrm{Ass}_R\,\mathrm{Ext}^{i-1}_R(R/\Gamma_{Rx}(R), M),$$

for all large k.

Proof. The result follows from Lemmas 2.1 and 2.2. □

2.4. Proposition. Let R be a Noetherian ring and let M, N be tow finitely generated R-modules. Let x be an N-regular element of R. Then, for any given integer $j \geq 0$, the set

$$\bigcup_{n=1}^{\infty} \mathrm{Ass}_R\mathrm{Ext}^j_R(N/x^nN, M),$$

of associated prime ideals, is finite.

Proof. If $j=0$, then

$$\mathrm{Ass}_R\mathrm{Hom}_R(N/x^nN, M) = \mathrm{Ass}_R\mathrm{Hom}_R(N, \mathrm{Hom}_R(R/Rx, M)),$$

and so

$$\bigcup_{n=1}^{\infty} \mathrm{Ass}_R\mathrm{Ext}^0_R(N/x^nN, M)$$

is a finite set. Suppose then that $j \geq 1$, and we use the exact sequence

$$0 \longrightarrow N \xrightarrow{x^n} N \longrightarrow N/x^nN \longrightarrow 0,$$

to obtain the exact sequence

$$\cdots \longrightarrow \mathrm{Ext}^{j-1}_R(N, M) \xrightarrow{x^n} \mathrm{Ext}^{j-1}_R(N, M) \longrightarrow \mathrm{Ext}^j_R(N/x^nN, M)$$
$$\longrightarrow \mathrm{Ext}^j_R(N, M) \xrightarrow{x^n} \mathrm{Ext}^j_R(N, M) \longrightarrow \cdots .$$

Hence we have the following exact sequence,

$$0 \to \mathrm{Ext}^{j-1}_R(N, M)/x^n\mathrm{Ext}^{j-1}_R(N, M) \to \mathrm{Ext}^j_R(N/x^nN, M) \to (0 :_{\mathrm{Ext}^j_R(N,M)} x^n) \to 0.$$

Consequently, it follows from Brodmann's result (see [1]) that the set

$$\bigcup_{n=1}^{\infty} \mathrm{Ass}_R\mathrm{Ext}^j_R(N/x^nN, M)$$

is finite. □

2.5. Lemma. Let R be a Noetherian ring and let M be an R-module. Let N be an Artinian submodule of M. Then

$$\mathrm{Ass}_R M/N \backslash \mathrm{Supp} N = \mathrm{Ass}_R M \backslash \mathrm{Supp} N.$$

Proof. As N is an Artinian R-module, it follows that the set $\mathrm{Supp}N \subseteq \mathrm{Max}R$ is finite. Let $\mathrm{Supp}N = \{\mathfrak{m}_1, \ldots, \mathfrak{m}_n\}$ and $J := \mathfrak{m}_1 \cdots \mathfrak{m}_n$. Then we have

$$\mathrm{Ass}_R M \backslash \mathrm{Supp} N = \mathrm{Ass}_R M/\Gamma_J(M) = \mathrm{Ass}_R M/N \backslash \mathrm{Supp} N,$$

as required. □

Following we let $H^j_I(M)$ denote the j^{th} local cohomology module of M with respect to an ideal I of a Noetherian ring R (cf. [2] and [3]).

2.6. Theorem. Let R be a Noetherian ring and let I be an ideal of R which is generated by an R-regular sequence. Let M be a finitely generated R-module and let i be a non-negative integer such that the R-module $\mathrm{Ext}^i_R(R/I, M)$ is Artinian. Then the set

$$\bigcup_{n=1}^{\infty} \mathrm{Ass}_R\mathrm{Ext}^{i+1}_R(R/I^n, M),$$

is finite. In particular, the set $\mathrm{Ass}_R H^{i+1}_I(M)$ is finite.

Proof. For $n \geq 0$, the exact sequence

$$0 \longrightarrow I^n/I^{n+1} \longrightarrow R/I^{n+1} \longrightarrow R/I^n \to 0$$

induces the exact sequence

$$\mathrm{Ext}^i_R(I^n/I^{n+1}, M) \to \mathrm{Ext}^{i+1}_R(R/I^n, M) \to \mathrm{Ext}^{i+1}_R(R/I^{n+1}, M) \to \mathrm{Ext}^{i+1}_R(I^n/I^{n+1}, M).$$

Since I is generated by an R-regular sequence, by [5, page 125] I^n/I^{n+1} is a finitely generated free R/I-module, and so the sets

$$\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(I^n/I^{n+1},M)=\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I,M), \text{ and}$$
$$\mathrm{Supp}\mathrm{Ext}_R^{i}(I^n/I^{n+1},M)=\mathrm{Supp}\mathrm{Ext}_R^{i}(R/I,M)$$

are finite, (note that the R-module $\mathrm{Ext}_R^i(R/I,M)$ is Artinian). Therefore in view of the above exact sequence and Lemma 2.5, the set

$$\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I^{n+1},M)\backslash\mathrm{Supp}\mathrm{Ext}_R^{i}(R/I,M)$$

is a subset of

$$(\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I^{n},M)\backslash\mathrm{Supp}\mathrm{Ext}_R^{i}(R/I,M))\cup\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I,M),$$

and so the set $\bigcup_{n=1}^{\infty}\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I^n,M)$ is finite, as required. The second assertion follows from the fact that

$$\mathrm{Ass}_R H_I^{i+1}(M)\subseteq\bigcup_{n=1}^{\infty}\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I^n,M).$$

□

2.7. Corollary. Let R be a Noetherian ring and let I be an ideal of R which is generated by an R-regular sequence. Let M be a finitely generated R-module and let i be a non-negative integer such that $\mathrm{Ext}_R^i(R/I,M)=0$. Then the sequence

$$\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I^k,M),$$

of associated primes is increasing and ultimately constant for all large k.

Proof. Since I^k/I^{k+1} is a free R/I-module, it follows that $\mathrm{Ext}_R^i(I^k/I^{k+1},M)=0$, for all $k\geq 1$. Hence the exact sequence

$$0\longrightarrow\mathrm{Ext}_R^{i+1}(R/I^k,M)\longrightarrow\mathrm{Ext}_R^{i+1}(R/I^{k+1},M)\longrightarrow\mathrm{Ext}_R^{i+1}(I^k/I^{k+1},M),$$

implies that

$$\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I^k,M)\subseteq\mathrm{Ass}_R\mathrm{Ext}_R^{i+1}(R/I^{k+1},M).$$

Now the result follows from Theorem 2.6. □

2.8. Lemma. Let $(R,\mathfrak{m})$ be a Noetherian local ring of depth d. Let M be a finitely generated R-module and N an Artinan submodule of M. Then for all $i\leq d-1$,

$$\mathrm{Ext}_R^i(M,R)\cong\mathrm{Ext}_R^i(M/N,R).$$

Proof. The exact sequence

$$0\longrightarrow N\longrightarrow M\longrightarrow M/N\longrightarrow 0$$

induces the exact sequence

$$\mathrm{Ext}_R^{i-1}(N,R)\longrightarrow\mathrm{Ext}_R^{i}(M/N,R)\longrightarrow\mathrm{Ext}_R^{i}(M,R)\longrightarrow\mathrm{Ext}_R^{i}(N,R).$$

As N has finite length and $\operatorname{depth} R=d$, it follows that

$$\mathrm{Ext}_R^{i}(N,R)=0=\mathrm{Ext}_R^{i-1}(N,R).$$

Hence the result follows. □

2.9. Lemma. Let $(R,\mathfrak{m})$ be a local Cohen-Macaulay ring of dimension d and I an ideal of R. Then for any $\mathfrak{p}\in\mathrm{Ass}_R\mathrm{Ext}_R^{\operatorname{grade} I}(R/I,R)$,

$$\operatorname{height}\mathfrak{p}=\operatorname{grade} I.$$

Proof. Let $\operatorname{grade} I = t$. The assertion is clear when $t = 0$. Now suppose that, $t \geq 1$. There exists an R-regular sequence $x_1, \dots, x_t \in I$. As

$$\operatorname{Ext}_R^{\operatorname{grade} I}(R/I, R) \cong \operatorname{Hom}_{R/(x_1,\dots,x_t)}(R/I, R/(x_1, \dots, x_t)),$$

and $R/(x_1, \dots, x_t)$ is a Cohen-Macaulay ring it follows that

$$\operatorname{Ass}_R \operatorname{Ext}_R^{\operatorname{grade} I}(R/I, R) \subseteq \operatorname{Assh}_R R/(x_1, \dots, x_t),$$

that implies for any $\mathfrak{p} \in \operatorname{Ass}_R \operatorname{Ext}_R^{\operatorname{grade} I}(R/I, R)$,

$$\operatorname{height} \mathfrak{p} = \operatorname{grade} I,$$

as required. □

2.10. Theorem. Let $(R, \mathfrak{m})$ be a local Cohen-Macaulay ring of dimension $d \geq 3$. Let I be an ideal of R such that $1 \leq \operatorname{grade} I \leq d - 2$. Then

$$\operatorname{depth} \operatorname{Ext}_R^{\operatorname{grade} I}(R/I, R) \geq 2,$$

and if $\operatorname{grade} I \leq d-3$ then the equality holds if and only if $\mathfrak{m} \in \operatorname{Ass}_R \operatorname{Ext}_R^{1+\operatorname{grade} I}(R/I, R)$.

Proof. Set $t := \operatorname{grade} I$. Let $\Gamma_{\mathfrak{m}}(R/I) := J/I$ for some ideal J of R with $I \subseteq J$. Then it is easy to see that $\mathfrak{m} \notin \operatorname{Ass}_R R/J$ and $\dim R/I = \dim R/J$. Hence as R is a Cohen-Macaulay ring, it follows that $\operatorname{grade} I = \operatorname{grade} J$. Moreover, since J/I has finite length, it follows from Lemma 2.8 that

$$\operatorname{Ext}_R^t(R/I, R) \cong \operatorname{Ext}_R^t(R/J, R) \text{ and } \operatorname{Ext}_R^{t+1}(R/I, R) \cong \operatorname{Ext}_R^{t+1}(R/J, R).$$

Therefore, we may and do replace I with J in the following. Since $\mathfrak{m} \notin \operatorname{Ass}_R R/J$, it follows that there exists an element $x \in R$ such that x is R/J-regular sequence. Then, as $\dim R/(J + Rx) = \dim R/J - 1$ and R is a Cohen-Macaulay ring, it follows that

$$\operatorname{grade}(J + Rx) = \operatorname{grade} J + 1.$$

Now the exact sequence

$$0 \to R/J \xrightarrow{x} R/J \to R/J + Rx \to 0$$

induces the exact sequence

$$0 \to \operatorname{Ext}_R^t(R/J, R) \xrightarrow{x} \operatorname{Ext}_R^t(R/J, R) \to \operatorname{Ext}_R^{t+1}(R/J + Rx, R).$$

Hence

$$\operatorname{Ass}_R \operatorname{Ext}_R^t(R/J, R)/x\operatorname{Ext}_R^t(R/J, R) \subseteq \operatorname{Ass}_R \operatorname{Ext}_R^{t+1}(R/J + Rx, R),$$

and since $1 + \operatorname{grade} J \leq d - 1$, it follows from Lemma 2.9 that

$$\mathfrak{m} \notin \operatorname{Ass}_R \operatorname{Ext}_R^{t+1}(R/J + Rx, R).$$

Now, it easily follows that

$$\operatorname{depth} \operatorname{Ext}_R^t(R/J, R) \geq 2.$$

Now, let $\operatorname{grade} J \leq d - 3$. Then we have the following exact sequence,

$$0 \to \operatorname{Ext}_R^t(R/J, R)/x\operatorname{Ext}_R^t(R/J, R) \to \operatorname{Ext}_R^{t+1}(R/J + Rx, R) \to (0 :_{\operatorname{Ext}_R^{t+1}(R/J,R)} x) \to 0.$$

Since $\operatorname{grade}(J + Rx) = t + 1$ and $t + 1 \leq d - 2$, it follows from the first part that $\operatorname{depth} \operatorname{Ext}_R^{t+1}(R/J + Rx, R) \geq 2$. Therefore it follows from the exact sequence

$$0 \to \operatorname{Hom}_R(R/\mathfrak{m}, (0 :_{\operatorname{Ext}_R^{t+1}(R/J,R)} x)) \to \operatorname{Ext}_R^1(R/\mathfrak{m}, \operatorname{Ext}_R^t(R/J, R)/x\operatorname{Ext}_R^t(R/J, R)) \to 0$$

that $\operatorname{depth} \operatorname{Ext}_R^t(R/J, R) = 2$ if and only if $\operatorname{Hom}_R(R/\mathfrak{m}, (0 :_{\operatorname{Ext}_R^{t+1}(R/J,R)} x)) \neq 0$. Consequently $\operatorname{depth} \operatorname{Ext}_R^t(R/J, R) = 2$ if and only if $\mathfrak{m} \in \operatorname{Ass}_R \operatorname{Ext}_R^{t+1}(R/J, R)$, as required. □

Acknowledgments

The authors are deeply grateful to the referee for his/her valuable suggestions on the article. Also, we would like to thank Professor Hossein Zakeri for his reading of the first draft and useful comments.

References

[1] M. Brodmann, *Asymptotic stability of* $\operatorname{Ass}(M/I^nM)$, Proc. Amer. Math. Soc. **74**(1979), 16-18.

[2] M.P. Brodmann and R.Y. Sharp, *Local cohomology; an algebraic introduction with geometric applications,* Cambridge University Press, Cambridge, 2013.

[3] A. Grothendieck, *Local cohomology,* Notes by R. Hartshorne, Lecture Notes in Math., 862 (Springer, New York, 1966).

[4] K. Khashyarmanesh and S. Salarian, *Asymptotic stability of* $\operatorname{Att}_R \operatorname{Tor}_1^R(R/\mathfrak{a}^n, A)$, Proc. Edin. Math. Soc. **44**(2001), 479-483.

[5] H. Matsumura, *Commutative ring theory*, Cambridge Univ. Press, Cambridge, UK, 1986.

[6] L. Melkersson and P. Schenzel, *Asymptotic prime ideals related to derived functors*, Proc. Amer. Math. Soc.**117** (1993), 935-938.

[7] L.J. Ratliff, Jr., *On the prime divisors of* I^n, n *large*, Michigan Math. J. **23** (1976), 337-352.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1335 – 1342

On rings over which every finitely generated module is a direct sum of cyclic modules

M. Behboodi*† and G. Behboodi Eskandari*

Abstract

In this paper we study (non-commutative) rings R over which every finitely generated left module is a direct sum of cyclic modules (called left FGC-rings). The commutative case was a well-known problem studied and solved in 1970s by various authors. It is shown that a Noetherian local left FGC-ring is either an Artinian principal left ideal ring, or an Artinian principal right ideal ring, or a prime ring over which every two-sided ideal is principal as a left and a right ideal. In particular, it is shown that a Noetherian local duo-ring R is a left FGC-ring if and only if R is a right FGC-ring, if and only if, R is a principal ideal ring.

Keywords: Cyclic modules; FGC-rings; duo-rings; principal ideal rings.

2000 AMS Classification: Primary 16D10, Secondary 16D70, 16P20, 16N60

Received : 01.11.2013 *Accepted* : 03.03.2014 *Doi* : 10.15672/HJMS.2015449425

1. Introduction

The question of which rings have the property that every finitely generated module is a direct sum of cyclic modules has been around for many years. We will call these rings FGC-rings. The problem originated in I. Kaplansky's papers [13] and [14], in which it was shown that a commutative local domain is FGC if and only if it is an almost maximal valuation ring. For several years, this is one of the major open problems in the theory. R. S. Pierce [19] showed that the only commutative FGC-rings among the commutative (von Neumann) regular rings are the finite products of fields. A deep and difficult study was made by Brandal [3], Shores-R. Wiegand [22], S. Wiegand [24], Brandal-R. Wiegand [4] and Vámos [23], leading to a complete solution of the problem in the commutative case. To show that a commutative FGC-ring cannot have an infinite number of minimal

*Department of Mathematical Sciences, Isfahan University of Technology, P.O.Box: 84156-83111, Isfahan, Iran and School of Mathematics, Institute for Research in Fundamental Sciences, P.O.Box: 19395-5746, Tehran, Iran Emails: mbehbood@cc.iut.ac.ir, gh.behboodi@ipm.ir.

†The research of this author was in part supported by a grant from IPM (No. 92130413). This research is partially carried out in the IPM-Isfahan Branch.

prime ideals required the study of topological properties (so-called Zariski and patch topologies). For complete and more leisurely treatment of this subject, see Brandal [2]. It gives a clear and detailed exposition for the reader wanting to study the subject. The main result reads as follows: A commutative ring R is an FGC-ring exactly if it is a finite direct sum of commutative rings of the following kinds: (i) maximal valuation rings; (ii) almost maximal Bézout domains; (iii) so-called torch rings (see [2] or [8] for more details on the torch rings).

The corresponding problem in the non-commutative case is still open; see [21, Appendix B. Dniester Notebook: Unsolved Problems in the Theory of Rings and Modules. Pages 461-516] in which the following problem is considered.

Problem (I. Kaplansky, reported by A. A. Tuganbaev [21, Problem 2.45]): *Describe the rings in which every one-sided ideal is two-sided and over which every finitely generated module can be decomposed as a direct sum of cyclic modules.*

Through this paper, all rings have identity elements and all modules are unital. A *left FGC-ring* is a ring R such that each finitely generated left R-module is a direct sum of cyclic submodules. A *right FGC-ring* is defined similarly, by replacing the word left with right above. A ring R is called a *FGC-ring* if it is a both left and right FGC-ring. Also, a ring R is called *duo-ring* if each one-sided ideal of R is two-sided. Therefore, the Kaplansky problem is: *Describe the FGC-duo-rings.*

In this paper we investigate Noetherian local left FGC-rings (see Section 2). Also, we will present a partial solution to the above problem of Kaplansky (see Section 3).

2. On left FGC-rings

A ring R is *local* in case R has a unique left maximal ideal. An Artinian (resp. Noetherian) ring is a ring which is both left and right Artinian (resp. Noetherian). A *principal ideal ring* is a ring which is both left and right principal ideal ring. Also, for a subset S of ${}_RM$, we denote by $\mathrm{Ann}_R(S)$, the annihilator of S in R. A left R-module M which has a composition series is called a module of finite length. The length of a composition series of ${}_RM$ is said to be the length of ${}_RM$ and denoted by $\mathrm{length}({}_RM)$.

We begin with the following lemma which is an associative, non-commutative version of Brandal [2, Proposition 4.3] for local rings $(R, \mathcal{M})$ with $\mathcal{M}^2 = (0)$. Also, the proof is based on a slight modification of the proof of [1, Theorem 3.1].

2.1. Lemma. *Let $(R, \mathcal{M})$ be a local ring with $\mathcal{M}^2 = (0)$ and ${}_R\mathcal{M} = Ry_1 \oplus \ldots \oplus Ry_t$ such that $t \geq 2$ and each Ry_i is a minimal left ideal of R. If there exist $0 \neq x_1,\ x_2 \in \mathcal{M}$ such that $x_1R \cap x_2R = (0)$, then the left R-module $(R \oplus R)/R(x_1, x_2)$ is not a direct sum of cyclic modules.*

Proof. Since ${}_R\mathcal{M} = Ry_1 \oplus \ldots \oplus Ry_t$ and each Ry_i is a minimal left ideal of R, we conclude that R is of finite composition length and $\mathrm{length}({}_RR) = t + 1$. We put ${}_RG = (R \oplus R)/R(x_1, x_2)$. Since $x_1,\ x_2 \in \mathcal{M}$ and $\mathcal{M}^2 = (0)$, we conclude that $\mathrm{Ann}_R(R(x_1, x_2)) = \mathcal{M}$. Thus $R(x_1, x_2)$ is simple and hence

$$\mathrm{length}({}_RG) = 2 \times \mathrm{length}({}_RR) - \mathrm{length}({}_RR(x_1, x_2)) = 2(t + 1) - 1.$$

We claim that every non-zero cyclic submodule Rz of G has length 1 or $t+1$. If $\mathcal{M}z = 0$, then $\mathrm{length}(Rz) = 1$ since $Rz \simeq R/\mathcal{M}$. Suppose that $\mathcal{M}z \neq 0$, then there exist $c_1, c_2 \in R$ such that $z = (c_1, c_2) + R(x_1, x_2)$. If $c_1, c_2 \in \mathcal{M}$, then $\mathcal{M}z = 0$, since $\mathcal{M}^2 = 0$. Thus without loss of generality, we can assume that $z = (1, c_2) + R(x_1, x_2)$ (since if $c_1 \notin \mathcal{M}$, then c_1 is unit). Now let $r \in \mathrm{Ann}_R(z)$, then $r(1, c_2) = t(x_1, x_2)$ for some $t \in R$. It follows that $r = tx_1$ and $rc_2 = tx_2$. Thus $tx_2 = tx_1c_2$. If $t \notin \mathcal{M}$, then t is unit and so $x_2 = x_1c_2$

that it is contradiction (since $x_1R \cap x_2R = (0)$). Thus $t \in \mathcal{M}$ and so $r = tx_1 = 0$. Therefore, $\mathrm{Ann}_R(z) = 0$ and so $Rz \cong R$. It follows that $\mathrm{length}(Rz) = t + 1$.

Now suppose the assertion of the lemma is false. Then ${}_RG$ is a direct sum of cyclic modules and since ${}_RG$ is of finite length, we have

$$G = Rw_1 \oplus \ldots \oplus Rw_k \oplus Rv_1 \oplus \ldots \oplus Rv_l,$$

where $l, k \geq 0$, and each Rw_i is of length $t+1$ and each Rv_j is of length 1. Clearly $\mathcal{M}\oplus\mathcal{M}$ is not a simple left R-module. Since $R(x_1, x_2)$ is simple, $\mathcal{M}G = (\mathcal{M} \oplus \mathcal{M})/R(x_1, x_2) \neq 0$. It follows that $k \geq 1$. Also, $\mathrm{length}({}_RG) = 2(t+1) - 1 = k(t+1) + l$ and this implies that $k = 1$ and $l = t$. Since $\mathcal{M}v_i = 0$ for each i, $\mathcal{M}G = \mathcal{M}w_1$ and hence

$$G/\mathcal{M}G \simeq Rw_1/\mathcal{M}w_1 \oplus Rv_1 \oplus \ldots \oplus Rv_t.$$

It follows that $\mathrm{length}({}_RG/\mathcal{M}G) = 1 + t$. On the other hand, we have

$$G/\mathcal{M}G \cong R/\mathcal{M} \oplus R/\mathcal{M}$$

and so $\mathrm{length}({}_RG/\mathcal{M}G) = 2$ and so $t = 1$, a contradiction. Thus the left R-module $(R \oplus R)/R(x_1, x_2)$ is not a direct sum of cyclic modules. □

We recall that the socle $\mathrm{soc}({}_RM)$ of a left module M over a ring R is defined to be the sum of all simple submodules of M.

2.2. Theorem. *Let $(R, \mathcal{M})$ be a local ring such that ${}_R\mathcal{M}$ and $\mathcal{M}_R$ are finitely generated. If every left R-module with two generators is a direct sum of cyclic modules, then either $\mathcal{M}$ is a principal left ideal or $\mathcal{M}$ is a principal right ideal.*

Proof. We can assume that $\mathcal{M}$ is not a principal left ideal of R. One can easily see that $\mathcal{M}_R$ is generated by $\{x_1, \cdots, x_n\}$ if and only if $\mathcal{M}/\mathcal{M}^2$ is generated by the set $\{x_1 + \mathcal{M}^2, \cdots, x_n + \mathcal{M}^2\}$ as a right ideal of $R/\mathcal{M}^2$. Thus it suffices to show that $\mathcal{M}/\mathcal{M}^2$ is a principal right ideal of $R/\mathcal{M}^2$. Since every left R-module with two generators is a direct sum of cyclic modules, we conclude that every left $R/\mathcal{M}^2$-module with two generators is also a direct sum of cyclic modules. Therefore, without loss of generality we can assume that $\mathcal{M}^2 = (0)$. It follows that $\mathrm{soc}({}_RR) = \mathrm{soc}(R_R) = \mathcal{M}$. Since ${}_R\mathcal{M}$ is finitely generated, ${}_R\mathcal{M} = Ry_1 \oplus \ldots \oplus Ry_t$ such that $t \geq 2$ and each Ry_i is a minimal left ideal of R. We claim that $\mathcal{M}_R = xR$, for if not, then we can assume that $\mathcal{M}_R = \oplus_{i\in I} x_iR$ where $|I| \geq 2$ and each x_iR is a minimal right ideal of R. We can assume that $\{1, 2\} \subseteq I$ and so $0 \neq x_1$, $x_2 \in \mathcal{M}$ and $x_1R \cap x_2R = (0)$. Now by Lemma 2.1, the left R-module $(R \oplus R)/R(x_1, x_2)$ is not a direct sum of cyclic modules, a contradiction. Thus $\mathcal{M}$ is principal as a right ideal of R. □

A ring whose lattice of left ideals is linearly ordered under inclusion, is called a *left uniserial* ring. A *uniserial ring* is a ring which is both left and right uniserial. Note that left and right uniserial rings are in particular local rings and commutative uniserial rings are also known as valuation rings.

Next, we need the following lemma from [18].

2.3. Lemma. (See Nicholson and Sánchez-Campos [18, Theorem 9]) *For any ring R, the following statements are equivalent:*

(1) *R is local, $J(R) = Rx$ for some $x \in R$ and $x^k = 0$ for some $k \in \mathbb{N}$.*
(2) *There exist $x \in R$ and $k \in \mathbb{N}$ such that $x^{k-1} \neq 0$ and*

$$R \supset Rx \supset \ldots \supset Rx^k = (0)$$

are the only left ideals of R.
(3) *R is left uniserial of finite composition length.*

2.4. Theorem. *Let* $(R, \mathcal{M})$ *be a local ring such that* ${}_R\mathcal{M}$ *and* $\mathcal{M}_R$ *are finitely generated and* $\mathcal{M}^k = (0)$ *for some* $k \in \mathbb{N}$. *If every left* R*-module with two generators is a direct sum of cyclic modules, then either* R *is a left Artinian principal left ideal ring or* R *is a right Artinian principal right ideal ring.*

Proof. Assume that every left R-module with two generators is a direct sum of cyclic modules. Then by Theorem 2.2, either $\mathcal{M}$ is a principal left ideal or $\mathcal{M}$ is a principal right ideal. If $\mathcal{M}$ is a principal left ideal, then by Lemma 2.3, R is a left Artinian principal left ideal ring. Thus we can assume that $\mathcal{M}$ is a principal right ideal. Then by using Lemma 2.3 to the right side, R is a right Artinian principal right ideal ring. □

Next, we need the following lemma from Mohamed H. Fahmy-Susan Fahmy[9]. We note that their definition of a local ring is slightly different than ours; they defined a *local ring* (resp. *scalar local ring*) as a ring R such that it contains a unique maximal ideal $\mathcal{M}$ and $R/\mathcal{M}$ is an Artinian ring (resp. division ring). Thus our definition of a *local ring* and the *scalar local ring* coincide.

2.5. Lemma. (See [9, Theorem 3.2]) *Let* $(R, \mathcal{M})$ *be non-Artinian Noetherian local ring. Then the following conditions are equivalent:*

(1) $\mathcal{M}$ *is principal as a right ideal.*
(2) $\mathcal{M}$ *is principal as a left ideal.*
(3) *Every two-sided ideal of* R *is principal as a left ideal.*
(4) *Every two-sided ideal of* R *is principal as a right ideal. Moreover,* R *is a prime ring.*

Now we are in a position to prove the main theorem of this section.

2.6. Theorem. *Let* $(R, \mathcal{M})$ *be a Noetherian local ring. If every left* R*-module with two generators is a direct sum of cyclic modules, then one of the following holds:*

(1) R *is an Artinian principal left ideal ring.*
(2) R *is an Artinian principal right ideal ring.*
(3) R *is a prime ring and every two-sided ideal of* R *is principal as both left and right ideals.*

Proof. First we assume that R is an Artinian ring. Thus by Theorem 2.4, either R is an Artinian principal left ideal ring or R is an Artinian principal right ideal ring. Now we assume that R is not an Artinian ring. By Theorem 2.2, either $\mathcal{M}$ is a principal left ideal or $\mathcal{M}$ is a principal right ideal. Thus by Lemma 2.5, R is a prime ring and every two-sided ideal of R is principal as both left and right ideals. □

3. A partial solution of Kaplansky's problem on duo-rings

A ring R is said to be *left* (resp. *right*) *hereditary* if every left (resp. right) ideal of R is projective as a left (resp. right) R-module. If R is both left and right hereditary, we say that R is hereditary. Recall that a PID is a domain R in which any left and any right ideal of R is principal. Clearly, any PID is hereditary.

Let R be an hereditary prime ring with quotient ring Q and A be a left R-module. Following Levy [17], we say that $a \in A$ is a *torsion element* if there is a regular element $r \in R$ such that $ra = 0$. Since, by Goldie's theorem, R satisfies the Ore condition, the set of torsion elements of A is a submodule $t(A) \subseteq A$. $A/t(A)$ is evidently torsion free (has no torsion elements).

3.1. Lemma. (Eisenbud-Robson [6, Theorem 2.1]) *Let R be an hereditary Noetherian prime ring, and let A be a finitely generated left R-module. Then $A/t(A)$ is projective and $A \cong t(A) \oplus A/t(A)$.*

A *Dedekind prime ring* [20] is an hereditary Noetherian prime ring with no proper idempotent two-sided ideals (see [7]). Clearly if a duo-ring R is a PID, then R is a Dedekind prime ring.

3.2. Lemma. (Eisenbud-Robson [6, Theorem 3.11]) *Let R be a Dedekind prime ring. Then every finitely generated torsion left R-module A is a direct sum of cyclic modules.*

3.3. Lemma. (Eisenbud-Robson [6, Theorem 2.4]) *Let R be a Dedekind prime ring, and let A be a projective left R-module. Then:*

(1) *If A is finitely generated, then $A \cong F \oplus I$ where F is a finitely generated free module and I is a left ideal of R.*
(2) *If A is not finitely generated, then A is free.*

3.4. Proposition. *Let R be a Dedekind prime ring. If R is a left principal ideal ring, then R is a left FGC-ring.*

Proof. Suppose that A is a finitely generated left R-module. Since R is a Dedekind prime ring, R is Noetherian and so A is also a Noetherian left R-module. Thus by Lemma 3.1, $A/t(A)$ is projective and $A \cong t(A) \oplus A/t(A)$. By Lemma 3.2, $t(A)$ is a direct sum of cyclic modules. Also by Lemma 3.3, $A/t(A) \cong F \oplus I$ where F is a free module and I is a left ideal of R. Since R is a principal left ideal ring, I is a cyclic left R-module, i.e., $A/t(A)$ is a direct sum of cyclic modules. Thus, $A \cong t(A) \oplus A/t(A)$ is a direct sum of cyclic modules. Therefore, R is a left FGC-ring. □

The following proposition is an answer to the question: "What is the structure of FGC Noetherian prime duo-rings?"

3.5. Proposition. (See also Jacobson [11, Page 44, Theorems 18 and 19]) *Let R be a Noetherian prime duo-ring (i.e., R is a Noetherian duo-domain). Then the following statements are equivalents:*

(1) *R is an FGC-ring.*
(2) *R is a left FGC-ring.*
(3) *R is a principal ideal ring.*

The same characterizations also apply for right R-modules.

Proof. (1) ⇒ (2) is clear. (2) ⇒ (3). Suppose that I is an ideal of R. Since I is a direct sum of principal ideals of R and R is a domain, we conclude that I is principal. Thus, R is a principle ideal ring. (3) ⇒ (1) is by Proposition 3.4. □

A *left* (resp., *right*) *Köthe ring* is a ring R such that each left (resp., right) R-module is a direct sum of cyclic submodules. A ring R is called a *Köthe ring* if it is a both left and right Köthe ring. In [16] Köthe proved that an Artinian principal ideal ring is a Köthe ring. Furthermore, a commutative ring R is a Köthe ring if and only if R is an Artinian principal ideal ring (see Cohen and Kaplansky [5]). The corresponding problem in the non-commutative case is still open (see [21, Appendix B, Problem 2.48] or Jain-Srivastava [12, Page 40, Problem 1]. Recently, a generalization of the Köthe-Cohen-Kaplansky theorem is given in [1]. In fact: in [1, Corollary 3.3.], it is shown that if R is a ring in which all idempotents are central, then R is a Köthe ring if and only if R is an Artinian principal ideal ring.

Next, the following theorem is an answer to the question: "What is the structure of FGC Noetherian local duo-rings?"

3.6. Theorem. *Let $(R, \mathcal{M})$ be a Noetherian local duo-ring. Then the following statements are equivalent:*

(1) *R is an FGC-ring.*
(2) *R is a left FGC-ring.*
(3) *Every left R-module with two generators is a direct sum of cyclic modules.*
(4) *Either R is an Artinian principal ideal ring or R is a principal ideal domain.*
(5) *R is a principal ideal ring.*

The same characterizations also apply for right R-modules.

Proof. (1) $\Rightarrow$ (2) $\Rightarrow$ (3) is clear.
(3) $\Rightarrow$ (4). Suppose that every left R-module with two generators is a direct sum of cyclic modules. Thus by Theorem 2.2, $\mathcal{M}$ is principal as both left and right ideals. If R is Artinian, then by Theorem 2.4, R is an Artinian principal ideal ring. If R is not Artinian, then by Lemma 2.5, R is a principal ideal domain.
(4) $\Rightarrow$ (1). If R is an Artinian principal ideal ring, then by the Köthe result, each left, and each right R-module is a direct sum of cyclic modules. Thus R is an FGC-ring. Now assume that R is a principal ideal domain. Then by Proposition 3.5, R is an FGC-ring.
(4) $\Rightarrow$ (5) is clear.
(5) $\Rightarrow$ (4). Assume that R is a principal ideal ring. Then $\mathcal{M}$ is principal as both left and right ideals. If R is Artinian, then by Lemma 2.3, R is an Artinian principal ideal ring. If R is not Artinian, then by Lemma 2.5, R is a principal ideal domain. □

Let $R = \Pi_{i=1}^{n} R_i$ be a finite product of rings R_i. Clearly R is a principal ideal ring if and only if each R_i is a principal ideal ring. On the other hand if R is a left FGC-ring, then each R_i is also a left FGC-ring. Thus as a corollary of Proposition 3.5 and Theorem 3.6, we have the following result.

3.7. Corollary. *Let $R = \Pi_{i=1}^{n} R_i$ be a finite product of Noetherian duo-rings R_i such that each R_i is a domain or a local ring. Then the following statements are equivalent:*

(1) *R is an FGC-ring.*
(2) *R is a left FGC-ring.*
(3) *R is a principal ideal ring.*

The same characterizations also apply for right R-modules.

Next, we need the following lemma from [10] about Artinian duo-rings (its proof is worthwhile even in the commutative case (see [10, Corollary 4] or [15, Lemma 4.2])

3.8. Lemma. *Let R be an Artinian duo-ring. Then R is a finite direct product of Artinian local duo rings.*

Next, we give the following characterizations of an Artinian FGC duo-ring. In fact, on Artinian duo-rings, the notions "FGC" and "Köthe" coincide.

3.9. Theorem. *Let R be an Artinian Duo-ring. Then the following statements are equivalent:*

(1) *R is a left FGC-ring.*
(2) *R is an FGC-ring.*
(3) *Every left R-module with two generators is a direct sum of cyclic modules.*
(4) *R is a left Köthe-ring.*
(5) *R is a Köthe-ring.*
(6) *R is a principal ideal ring.*

The same characterizations also apply for right R-modules.

Proof. Since R is an Artinian duo-ring, by Lemma 3.8, $R = \Pi_{i=1}^{n} R_i$ such that each R_i is an Artinian local duo-ring. Thus by the Köthe result and Corollary 3.7, the proof is complete. □

Acknowledgments. The authors would like to thank the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] Behboodi, M., Ghorbani, A., Moradzadeh-Dehkordi A. and Shojaee, S. H. *On left Köthe rings and a generalization of a Köthe-Cohen-Kaplansky theorem*, Proc. Amer. Math. Soc. **142**., no. 8, 2625-2631, (2014).

[2] Brandal, W. *Commutative Rings Whose Finitely Generated Modules decompose*, Lecture Notes in Mathematics, Vol. **723**. (Springer, 1979).

[3] Brandal, W. *Almost maximal integral domains and finitely generated modules.* Trans. Amer. Math. Soc. **183**., 203-222, (1973).

[4] Brandal, W., and Wiegand, R. *Reduced rings whose finitely generated modules decompose.* Comm. Algebra **6**., (2) 195-201, (1978).

[5] Cohen, I. S. and Kaplansky, I. *Rings for which every module is a direct sum of cyclic modules*, Math. Z. **54**., 97-101, (1951).

[6] Eisenbud, D. and Robson, J. C. *Modules over Dedekind prime rings*, J. Algebra **16**., 67-85 (1970).

[7] Eisenbud, D. and Robson, J. C. *Hereditary Noetherian prime rings*, J. Algebra **16**., 86-104, (1970).

[8] Facchini, A. *On the structure of torch rings.* Rocky Mountain J. Math. **13**., (3) 423-428, (1983).

[9] Fahmy, M. H. and Fahmy, S. *On non-commutative noetherian local rings, non-commutative geometry and particle physics*, Chaos, Solitons and Fractals **14**., 1353-1359, (2002).

[10] Habeb, J. M. *A note on zero commutative and duo rings*, Math. J. Okayama Univ **32**., 73-76, (1990).

[11] Jacobson, N. *The Theorey of Rings*, Amerian Mathematical Society Mathematical Surveys, vol. I, (American Mathematical Society, New York, 1943).

[12] Jain, S. K. and Srivastava, Ashish K. *Rings characterized by their cyclic modules and right ideals: A survey-I.* (See http://euler.slu.edu/ srivastava/articles.html).

[13] Kaplansky, I. *Elementary divisors and modules.* Trans. Amer. Math. Soc. **66**., 464-491, (1949).

[14] Kaplansky, I. *Modules over Dedekind rings and valuation rings.* Trans. Amer. Math. Soc. **72**., 327-340, (1952).

[15] Karamzadeh, N. S. and Karamzadeh, O. A. S. *On artinian modules over Duo rings*, Comm. Algebra **38**., 3521-3531, (2010).

[16] Köthe, G. *Verallgemeinerte abelsche gruppen mit hyperkomplexem operatorenring*, Math. Z. **39**., 31-44, (1935).

[17] Levy, L. *Torsion-Free and divisible modules over non-integral domains*, Canad. J. Math. **15**., 132-151, (1963).

[18] Nicholson, W. K. and Sánchez-Campos, E. *Rings with the dual of the isomorphism theorem*, J. Algebra **271**., (1) 391-406, (2004).

[19] Pierce, R. S. *Modules over commutative regular rings.* Mem. Amer. Math. Soc. **70**., (1967).

[20] Robson, J. C. *Non-commutative Dedekind rings*, J. Algebra **9**., 249-265, (1968).

[21] Sabinin, L., Sbitneva, L. and Shestakov, I. *Non-associative algebra and its applications*, Lecture Notes in Pure and Applied Mathematics **246**., Chapman and Hall/CRC, (2006).

[22] Shores, T. S. and Wiegand, R. *Rings whose finitely generated modules are direct sums of cyclics.* J. Algebra **32**., 152-172, (1974).

[23] Vámos, P. *The decomposition of finitely generated modules and fractionally self-injective rings.* J. London Math. Soc. **16**., 209-220, (1977).

[24] Wiegand, R. and Wiegand, S. *Commutative rings whose finitely generated modules are direct sums of cyclics.* Abelian group theory (Proc. Second New Mexico State Univ. Conf., Las Cruces, N.M., 1976), pp. 406-423. Lecture Notes in Math., Vol. **616**., (Springer, Berlin, 1977).

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1343 – 1354

Fibonacci, and Lucas Pascal triangles

Hacène Belbachir* and László Szalay†‡

Abstract

In this paper, we give explicit formulas for elements of the Fibonacci, and Lucas Pascal triangles. The structure of these objects and Pascal's original triangle coincide. Keeping the rule of addition, we replace both legs of the Pascal triangle by the Fibonacci sequence, and the Lucas sequence, respectively. At the end of the study we describe how to determine such a formula for any binary recurrence $\{G_n\}_{n=0}^{\infty}$ satisfying $G_n = G_{n-1} + G_{n-2}$. Other scattered results are also presented.

Keywords: Generalized Pascal's triangle, Fibonacci sequence, Lucas sequence.

2000 AMS Classification: 11B39, 05B30

Received : 07.07.2015 *Accepted* : 08.12.2015 *Doi* : 10.15672/HJMS.20164515688

1. Introduction

Although a lot is known about the Pascal triangle, its origin is lost in the mist of time. Since the work of Pascal [10] several scholars have contributed with variations, generalizations to this object. An early generalization is due to Raab [11], who introduced the so-called AB-based Pascal triangles. Its structure is identical to the regular Pascal triangle, and the elements are the coefficients of $x^{n-k}y^k$ in the expansion of the polynomial $(Ax + By)^n$. Some variations, for instance the Pascal pyramid, stem from different combinatorial approaches. The Hosoya's triangle [7] is also a triangular arrangement based on the Fibonacci numbers, where each entry is the sum of the two entries above in either the left diagonal or the right diagonal. Koshy [9] gave a description on different Pascal-like triangles which are linked to the sums $\alpha^n + \beta^n$ (and the differences $\alpha^n - \beta^n$), where α and β are the zeros of the characteristic polynomial $x^2 - Ax - B$ of the linear recurrence $G_n = AG_{n-1} + BG_{n-2}$. Sun [12] provided a generalization of the DFF, and $DFFz$ triangles introduced by Ferri et al. [5, 6], respectively. Generally, the cited papers

*USTHB, Faculty of Mathematics, Recits Laboratory, Algiers, Algeria.
Email: hacenebelbachir@gmail.com
†J. Selye University, Department of Mathematics and Informatics, Komarno, Slovakia.
Email: szalay.laszlo@nyme.hu
‡Corresponding Author.

work with elementary considerations, using the properties of binomial coefficients and certain sequences.

In this paper, we keep the arrangement of Pascal's original triangle and, apart from the beginning, the rule of addition, but we vary the sequences located on the legs of the triangle. More precise description will be given later. The main purpose of this work is to give applicable explicit formulas for the elements of the so-called Fibonacci, and Lucas Pascal triangle. Note that the Fibonacci triangle studied by the present paper and Hosoya's triangle do not coincide since the insertion methods are different. Ensley [4] already derived a formula for the elements of Fibonacci Pascal triangle. While his result is given by a weighted sum of certain binomial coefficents, here (in Corollary 6) the exponential and polynomial terms are separated. The principal results are Theorems 1 and 4 in Section 3. As a consequence of these outcomes, we are able to determine analogous formula for the triangle generated by any binary recurrence $\{G_n\}_{n=0}^{\infty}$ satisfying $G_n = G_{n-1} + G_{n-2}$ (Theorem 5). On the other hand, we provide certain interesting arguments on arithmetic triangles, some of them have been justified previously. In fact, such objects are actually very popular, and one can find lot of information about them in the literature (see [2, 3, 4, 12]).

Let $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=0}^{\infty}$ denote two real sequences. There is no importance if $a_0 \neq b_0$, in this case we replace both terms by Ω as an indeterminate object.

The two initial sequences, as it was described by Dil and Mező [3], generate an *infinite matrix* $\mathbf{M} = (M_{k,n})_{k\geq 0, n\geq 0}$ as follows. Put $M_{0,0} = \Omega$, and

$$M_{k,0} = a_k, \qquad M_{0,n} = b_n, \qquad k \geq 1,\ n \geq 1,$$

further let

$$M_{k,n} = M_{k,n-1} + M_{k-1,n}, \qquad kn \neq 0.$$

For $k \geq 1$ and $n \geq 1$ the authors proved the explicit formula

$$M_{k,n} = \sum_{i=1}^{k} \binom{k+n-i-1}{n-1} a_i + \sum_{j=1}^{n} \binom{k+n-j-1}{k-1} b_j. \tag{1.1}$$

A similar approach in constructing a sort of *Generalized Arithmetic Triangle* (in short GAT) was used in [2]. Letting $A,\ B \in \mathbb{R}$, the GAT is structurally identical with Pascal's original triangle (Pascal himself called his object arithmetic triangle) and contains rows numbered by $0,\ 1,\ 2,\ \ldots$ such that the n^{th} row possesses the elements $\left\langle {n \atop k} \right\rangle$ in the positions (say columns) $k = 0,\ 1,\ \ldots,\ n$ as follows.

Let $\left\langle {0 \atop 0} \right\rangle$ be arbitrary denoted by Ω, and for positive integer n put $\left\langle {n \atop 0} \right\rangle = A^n a_n$ and $\left\langle {n \atop n} \right\rangle = B^n b_n$, further for $n \geq 2$ and $1 \leq k \leq n-1$ let

$$\left\langle {n \atop k} \right\rangle = B \left\langle {n-1 \atop k-1} \right\rangle + A \left\langle {n-1 \atop k} \right\rangle.$$

Theorem 1 of [2] admits the direct formula

$$\left\langle {n \atop k} \right\rangle = A^{n-k} B^k \left(\sum_{i=1}^{n-k} \binom{n-1-i}{k-1}_{\Gamma} a_i + \sum_{j=1}^{k} \binom{n-1-j}{n-k-1}_{\Gamma} b_j \right), \tag{1.2}$$

to express $\left\langle {n \atop k} \right\rangle$ in the terms of A, B and the sequences if $1 \leq n$ and $0 \leq k \leq n$. The extension $\binom{\cdot}{\cdot}_{\Gamma}$ of binomial coefficients to arbitrary integers n and k appeared in (1.2) is obtained by the Gamma function (see [1], formula 6.1.21):

$$\binom{n}{k}_{\Gamma} = \lim_{n_1 \to n} \lim_{k_1 \to k} \frac{\Gamma(n_1+1)}{\Gamma(k_1+1) \cdot \Gamma(n_1-k_1+1)}.$$

We really need it since if $k = 0$ or $k = n$ in (1.2), then the lower index of the binomial coefficients is negative. Note that $\binom{n}{k}_\Gamma = \binom{n}{k}$ if k is nonnegative.

Our Generalized Aritmetic Triangle extends Ensley's GAT [4], since here we allow $a_0 \neq b_0$ in the generator sequences, further we also vary the rule of addition by the parameters A and B. In [4] (where $A = B = 1$, and $a_0 = b_0$) the formula

$$(1.3) \qquad \left\langle {n \atop k} \right\rangle = \sum_{i=0}^{n-k} \binom{n-i}{k} \delta_i^a + \sum_{j=0}^{k} \binom{n-j}{k-j} \delta_j^b - \binom{n}{k} a_0$$

was established, where $\delta_i^a = a_i - a_{i-1}$ for positive i, and $\delta_0^a = a_0$ (analogous scheme holds for the sequence $\{b_n\}$). At the first sight (1.3) is strange because it contains a_0, meanwhile the structure says no influence of $a_0 = b_0$ on the triangle. But there is no contradiction, since a short calculation shows that we can exclude a_0 (and δ_0^a, δ_0^b) from (1.3). Moreover it is easy to see that (1.2) and (1.3) are compatible. Indeed, by (1.3) we have

$$\left\langle {n \atop k} \right\rangle = \sum_{i=1}^{n-k} \binom{n-i}{k} (a_i - a_{i-1}) + \binom{n}{k} b_0 + \sum_{j=1}^{k} \binom{n-j}{k-j} (b_j - b_{j-1}),$$

and for $1 \leq k \leq n-1$, via $a_0 = b_0$ it leads to

$$\left\langle {n \atop k} \right\rangle = \sum_{i=1}^{n-k} \binom{n-1-i}{k-1} a_i + \sum_{j=1}^{k} \binom{n-1-j}{k-j} b_j.$$

The next frame collects some relevant Pascal type triangles (or arrays) have been already studied.

$\{a_n\}$	$\{b_n\}$	Reference
1	1	Pascal Triangle (PT)
A^n	B^n	Raab [11] (AB-based PT)
2	1	Hosoya [8] (Asymmetrical PT)
arbitrary	arbitrary	Ensley [4] (GAT)
F_{n+1}	F_{n+1}	Ensley [4] (shifted Fibonacci Triangle)
0	$\frac{1}{n}$	Dil – Mező [3] (Hyperharmonic numbers)
0	F_n	Dil – Mező [3] (Hyper-Fibonacci numbers)
F_{2n-1}	F_{n-1}	Dil – Mező [3]
$A^n a$	$B^n b$	Belbachir – Szalay [2]

Assume now that $A = B = 1$. Then the rectangular shape matrix $\mathbf{M}$ and the triangular shape GAT differ only in their appearance. Indeed, apart from the geometrical display, the identity

$$(1.4) \qquad M_{k,n} = \left\langle {k+n \atop n} \right\rangle$$

transmits them to each other for $k+n \geq 1$. Apparently, for $k \geq 1$ and $n \geq 1$ the formulas (1.2) and (1.1) are equivalent via (1.4). Really, replacing n by $k+n$ and k by n in (1.2) at the same time we arrive at (1.1).

Obviously, with the parameters A and B a kind of Generalized Pascal Array (GPA, given by the matrix $\widetilde{\mathbf{M}}$) can be obtained if we introduce the modification

$$\widetilde{M}_{k,0} = A^k a_k, \qquad \widetilde{M}_{0,n} = B^n b_n, \qquad k \geq 1,\ n \geq 1,$$

and

$$\widetilde{M}_{k,n} = B\widetilde{M}_{k,n-1} + A\widetilde{M}_{k-1,n}, \qquad kn \neq 0.$$

In the sequel we always assume that $A = B = 1$, and in this paper, we basicly investigate the situation, when the sequences $\{a_n\} = \{b_n\}$ are the Fibonacci, or the Lucas sequence. When $a_n = b_n = F_{n+1}$, as a consequence of (1.3), Ensley provided

$$\sum_{i=1}^{n-k} \binom{n-i}{k} F_{i-2} + \sum_{j=1}^{k} \binom{n-j}{k-j} F_{j-2} + \binom{n}{k},$$

for the k^{th} entry in row n, we develope a more informative explicit formula. In this study, we also show a new and simple proof for one of the equivalent formulas (1.2), (1.1) and (1.3). Note that in [4] the author used a combination of two preliminary lemmata, in [3] and in [2] the technique of induction was used. Here we apply an elementary "atomic" observation. Later we will use some preliminary lemmata, which are stated here.

1.1. Lemma. *For arbitrary nonnegative integer n we have*

$$\sum_{i=0}^{n} \binom{n}{i} F_i = F_{2n},$$
$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} F_i = -F_n.$$

Proof. See [9], Theorems 12.5 and 12.6, on pages 157-158. □

We note that in the case of the second statement of Lemma 1.1, some inaccuracy appears in [9].

1.2. Lemma. *Let n be a nonnegative integer. Then*

$$\sum_{i=0}^{n} \binom{n}{i} L_i = L_{2n},$$
$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} L_i = L_n$$

hold.

Proof. See [9], remarks after Theorems 12.5 and 12.6, on pages 157-158. □

1.3. Lemma. *Let n, n_1 and n_2 be nonnegative integers. Then*

$$\sum_{i=0}^{n} \binom{n-i}{n_1} \binom{i}{n_2} = \binom{n+1}{n_1+n_2+1}$$

is valid.

Proof. This is a corollary of the Vandermonde identity. □

2. The "atomic" lemma, the principle of superposition, and some easily computable arithmetic triangles

Recall, that $A = B = 1$. Let $\left\langle{n \atop k}\right\rangle_{\{a_n,b_n\}}$ denote the elements of the Generalized Pascal Triangle (in short GPT) generated by the sequences $\{a_n\}$ and $\{b_n\}$. For every given sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$ and $\{d_n\}$, we can show easily that

$$\left\langle{n \atop k}\right\rangle_{\{a_n,b_n\}} + \left\langle{n \atop k}\right\rangle_{\{c_n,d_n\}} = \left\langle{n \atop k}\right\rangle_{\{a_n+c_n,b_n+d_n\}}.$$

Now we intend to split this effect into elementary parts, and as a consequence we describe a method for determining $\left\langle{n \atop k}\right\rangle$ when the modifying sequences $\{c_n\}$ and $\{d_n\}$ are simple from one sort of point of view. The next lemma describes the elementary situation when an existing GPT is modified by $c \in \mathbb{R}$ at exactly one element of one of the legs. That is, apart from one entry of $\{c_n\}$ or $\{d_n\}$ we assume $c_n = d_n = 0$. Clearly, such a modification can be applied as many times as we need, and the influences of the consecutive modifications can be superposed. At the end of the section we will see, that this approach is not sufficient to handle the case of Fibonacci, and the Lucas triangle.

Assume that there is given a GPT by the sequences $\{a_n\}$ and $\{b_n\}$.

2.1. Lemma. *If one modifies the element $\left\langle{i \atop j}\right\rangle_{\{a_n,b_n\}}$ located on one of the legs of a triangle ($i \geq 1$, further $j = 0$ or $j = i$) by adding $c \in \mathbb{R}$ to, then the only change on the legs is $\left\langle{i \atop j}\right\rangle_{new} = \left\langle{i \atop j}\right\rangle_{\{a_n,b_n\}} + c$. Further, in the inner part of the triangle we find*

(1) *in case of $j = 0$ (left leg)*

$$\left\langle{n \atop k}\right\rangle_{new} = \begin{cases} \left\langle{n \atop k}\right\rangle_{\{a_n,b_n\}} + c\binom{n-i-1}{k-1}, & \text{if } n \geq i+1 \text{ and } 1 \leq k \leq n-i; \\ \left\langle{n \atop k}\right\rangle_{\{a_n,b_n\}}, & \text{otherwise,} \end{cases}$$

(2) *in case of $j = i$ (right leg)*

$$\left\langle{n \atop k}\right\rangle_{new} = \begin{cases} \left\langle{n \atop k}\right\rangle_{\{a_n,b_n\}} + c\binom{n-i-1}{k-i}, & \text{if } n \geq i+1 \text{ and } i \leq k \leq n-1; \\ \left\langle{n \atop k}\right\rangle_{\{a_n,b_n\}}, & \text{otherwise.} \end{cases}$$

Proof. It is obvious by the construction (see Figure 1, as an illustration with $i = 2$, $j = 0$). □

We can build simple GPT's by starting with the empty triangle (any $\left\langle{n \atop k}\right\rangle$ is zero), using element by element of the sequences $\{c_n\}$ and $\{d_n\}$. Note, that the idea is more and less due to Ensley [4], although he started with the classical Pascal triangle, therefore at the end of the procedure he removed that. Observe, that only inserting $c_1, c_2, \ldots, c_{n-k}$, and $d_1, d_2, \ldots, d_k$ has influence on the element $\left\langle{n \atop k}\right\rangle$. Thus, by Lemma 2.1 and the principle of superposition we obtain immediately a (new) proof for the identity (1.2), since $\binom{n}{k}_\Gamma = \binom{n}{k}$ holds if k is nonnegative.

Now we demonstrate the applicabilty of Lemma 2.1 by a few further examples. Locally we use the notations

$$\Lambda_c = \sum_{i=1}^{n-k} \binom{n-1-i}{k-1} c_i \qquad \text{and} \qquad \Lambda_d = \sum_{i=1}^{k} \binom{n-1-i}{n-k-1} d_i.$$

Clearly, we have

$$\left\langle{n \atop k}\right\rangle_{new} = \left\langle{n \atop k}\right\rangle_{old} + \Lambda_c + \Lambda_d.$$

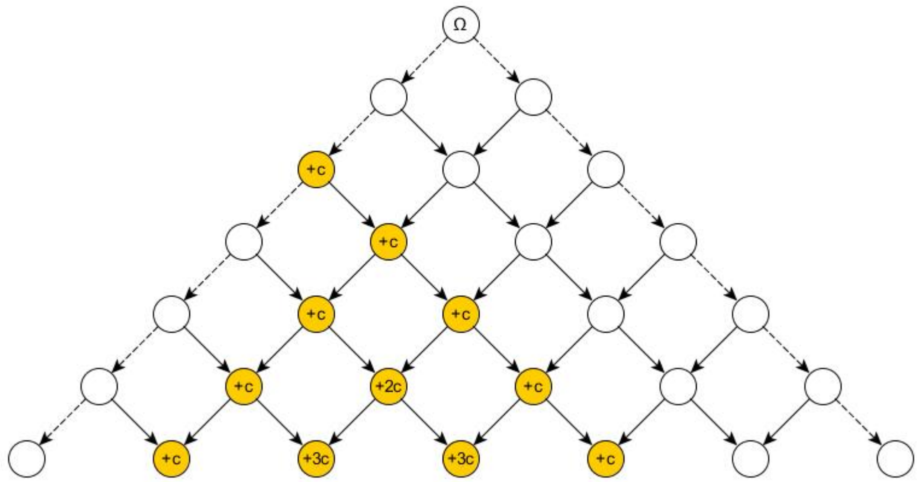


Figure 1. The influence of modification of one leg element

We will also use Lemma 1.3 to calculate the appropriate sums.

(1) Initial triangle: empty; modified by $c_n = d_n = 1$ (case of classical Pascal triangle).

$$\left\langle {n \atop k} \right\rangle_{new} = \sum_{i=1}^{n-k} \binom{n-1-i}{k-1} + \sum_{i=1}^{k} \binom{n-1-i}{n-k-1} = \binom{n-1}{k} + \binom{n-1}{n-k} = \binom{n}{k}.$$

(2) Initial triangle: empty; modified by $c_n = d_n = p(n)$, where $p(x)$ is a given polynomial of degree $d \geq 1$. First we express the polynomial as a linear combination of the binomial coefficients $\binom{x}{i}$, $i = 1, \dots d$. Then we apply Lemma 1.3 to determine the sums appearing as the influence of the left, and the right leg. For instance, put $p(x) = x^2$, so $c_n = d_n = n^2$. Since $x^2 = 2\binom{x}{2} + \binom{x}{1}$, we find

$$\begin{aligned} \Lambda_c &= \sum_{i=1}^{n-k} \binom{n-1-i}{k-1} i^2 = \sum_{i=1}^{n-k} \binom{n-1-i}{k-1} \left(2\binom{i}{2} + \binom{i}{1} \right) \\ &= 2\sum_{i=1}^{n-k} \binom{n-1-i}{k-1}\binom{i}{2} + \sum_{i=1}^{n-k} \binom{n-1-i}{k-1}\binom{i}{1} = 2\binom{n}{k+2} + \binom{n}{k+1}. \end{aligned}$$

Similarly,

$$\Lambda_d = 2\binom{n}{k-2} + \binom{n}{k-1},$$

thus

$$\left\langle {n \atop k} \right\rangle_{new} = 2\binom{n}{k-2} + \binom{n}{k-1} + \binom{n}{k+1} + 2\binom{n}{k+2}.$$

(3) Initial triangle: Pascal triangle; modified by $c_n = 1$ ($n \in \mathbb{N}$, Asymmetrical PT in [8], see Figure 2). Only the changes at $\binom{1}{0}$, $\binom{2}{0}$, ..., $\binom{n-k}{0}$ cause variation at

$\binom{n}{k}$. By Lemma 2.1 we find

$$\left\langle {n \atop k} \right\rangle_{new} = \begin{cases} \binom{n}{k} + \sum_{i=1}^{n-k} \binom{n-i-1}{k-1} = \binom{n}{k} + \binom{n-1}{k}, & \text{if } n \geq 2 \text{ and } 1 \leq k \leq n-1; \\ \binom{n}{n} = 1, & \text{if } k = n, \end{cases}$$

which coincides the result of [8].

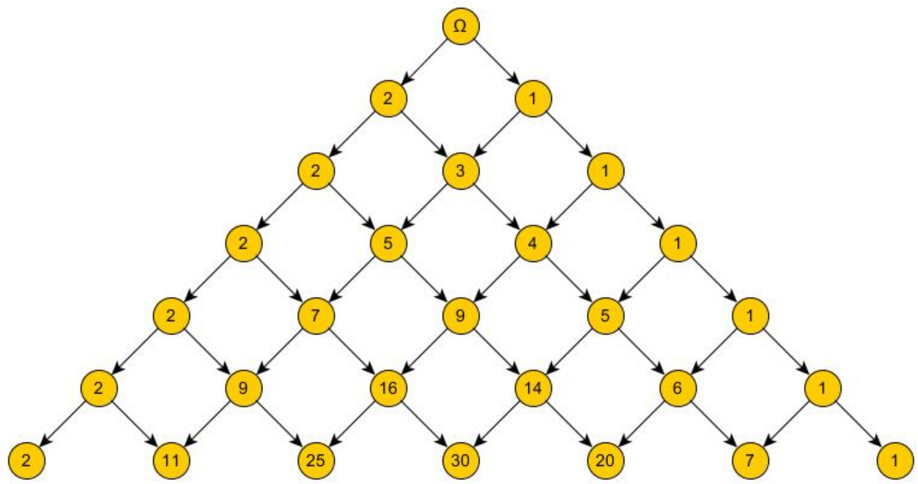


Figure 2. Asymmetric Pascal triangle

Suppose now that we want to modify the empty triangle by the sequences $c_n = d_n = F_n$. Then we face, among others, to the problem of determining the sum

$$\Lambda_c = \sum_{i=1}^{n-k} \binom{n-1-i}{k-1} F_i, \tag{2.1}$$

but unfortunately there is no closed formula to express it. Subsequently, we need something else to describe the Fibonacci Pascal triangle.

3. Fibonacci and Lucas triangles

3.1. Fibonacci triangle. In this part, first we focus on the Fibonacci triangle, which was introduced by Ensley [4]. Recall, that he took $c_n = d_n = F_{n+1}$. Denoting the elements of this triangle by $\left\langle {n \atop k} \right\rangle_{F_{n+1}}$, Ensley showed

$$\left\langle {n \atop k} \right\rangle_{F_{n+1}} = \binom{n}{k} + \sum_{i=1}^{n-k} \binom{n-i}{k} F_{i-2} + \sum_{j=1}^{n-k} \binom{n-j}{k-j} F_{j-2}.$$

At the end of this section we will give a more applicable formula for the elements of this triangle, but now we start with studying the triangle generated by $a_n = b_n = F_n$ (see Figure 3). The main result is the following.

3.1. Theorem. *For any nonnegative integers n and k we have*

$$\left\langle {n \atop k} \right\rangle_{F_n} = F_{n+k} - q_k(n), \tag{3.1}$$

where

$$q_k(x) = 2\sum_{j=0}^{\lfloor k/2\rfloor}\binom{x}{k-2j}F_{2j} \tag{3.2}$$

is a rational polynomial of degree $k-2$ *if* $k \geq 2$, *and* $q_0(x) = q_1(x) = 0$.

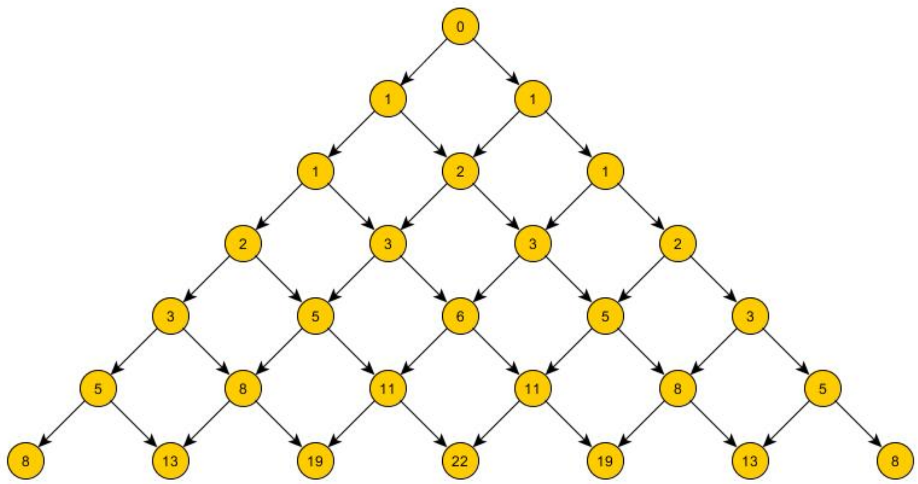


Figure 3. Fibonacci Pascal triangle

For proving Theorem 3.1, we need

3.2. Lemma. *For any nonnegative integer* k, *the specific value* $q_k(k)$ *is given by* $q_k(k) = F_{2k} - F_k$.

Proof.

$$q_k(k) = 2\sum_{j=0}^{\lfloor k/2\rfloor}\binom{k}{k-2j}F_{2j} = 2\sum_{j=0}^{\lfloor k/2\rfloor}\binom{k}{2j}F_{2j} = \sum_{j=0}^{k}\binom{k}{j}F_j + \sum_{j=0}^{k}(-1)^j\binom{k}{j}F_j.$$

It implies, by Lemma 1.1, that $q_k(k) = F_{2k} - F_k$. □

3.3. Lemma. *The equality* $q_k(N+1) - q_k(N) = q_{k-1}(N)$ *fulfils for any nonnegative integers* N *and* k.

Proof. Suppose first that $k = 2\kappa$. Then

$$\begin{aligned}\frac{q_k(N+1)}{2} &= \sum_{j=0}^{\kappa}\binom{N+1}{k-2j}F_{2j} = \sum_{j=0}^{\kappa-1}\left(\binom{N}{k-2j-1} + \binom{N}{k-2j}\right)F_{2j} + \binom{N}{0}F_k \\ &= \sum_{j=0}^{\kappa-1}\binom{N}{k-2j-1}F_{2j} + \sum_{j=0}^{\kappa}\binom{N}{k-2j}F_{2j} = \frac{q_{k-1}(N)}{2} + \frac{q_k(N)}{2}.\end{aligned}$$

If $k = 2\kappa + 1$ is odd, then similarly we get

$$\begin{aligned}\frac{q_k(N+1)}{2} &= \sum_{j=0}^{\kappa} \binom{N+1}{k-2j} F_{2j} = \sum_{j=0}^{\kappa} \left(\binom{N}{k-2j-1} + \binom{N}{k-2j} \right) F_{2j} \\ &= \sum_{j=0}^{\kappa} \binom{N}{k-2j-1} F_{2j} + \sum_{j=0}^{\kappa} \binom{N}{k-2j} F_{2j} = \frac{q_{k-1}(N)}{2} + \frac{q_k(N)}{2}.\end{aligned}$$

□

Now we turn to the proof of Theorem 3.1.

Proof. First we show the statement for the legs of the GPT.

$$\begin{aligned}\left\langle {n \atop 0} \right\rangle_{F_n} &= F_n - q_0(n) = F_n, \\ \left\langle {n \atop n} \right\rangle_{F_n} &= F_{2n} - q_n(n) = F_{2n} - (F_{2n} - F_n) = F_n.\end{aligned}$$

Now assume that $n \geq 2$ and $1 \leq k \leq n-1$. After verifying $\left\langle {2 \atop 1} \right\rangle_{F_n} = F_3 - q_1(2) = 2$, we use the technique of induction. Hence we assume that (3.1) is true for $n \leq N$ $(N \geq 2)$. Applying it, together with Lemma 3.3, we deduce

$$\begin{aligned}\left\langle {N+1 \atop k} \right\rangle_{F_n} &= \left\langle {N \atop k-1} \right\rangle_{F_n} + \left\langle {N \atop k} \right\rangle_{F_n} = (F_{N+k-1} - q_{k-1}(N)) + (F_{N+k} - q_k(N)) \\ &= F_{N+k-1} + F_{N+k} - (q_{k-1}(N) - q_k(N)) = F_{N+k+1} - q_k(N+1).\end{aligned}$$

□

Since the polynomials $q_k(x)$ play crucial role in (3.1), in the next table we give the first few of them explicitly. Recall, that $q_0(x) = q_1(x) = 0$.

k	2	3	4	5	6
$q_k(x)$	2	$2x$	$x^2 - x + 6$	$\frac{1}{3}(x^3 - 3x^2 + 20x)$	$\frac{1}{12}(x^4 - 6x^3 + 47x^2 - 42x + 192)$

The proof of Lemma 3.3 gives a hint how to determine $q_{k-1}(x)$ if one knows $q_k(x)$. The reverse order is more interesting since we know the beginning of the list $q_2(x)$, $q_3(x)$, ... etc. Although we have (3.2) in Theorem 3.1, it may be challenging to know how to generate the next unknown element of the list. Suppose that the polynomial

$$q_{t+1}(x) = b_{t-1}x^{t-1} + b_{t-2}x^{t-2} + \cdots + b_1 x + b_0$$

is known, and we intend to determine the coefficients a_i of the polynomial

$$q_{t+2}(x) = a_t x^t + a_{t-1}x^{t-1} + \cdots + a_1 x + a_0.$$

By Lemma 3.3, we expand the difference $q_{t+2}(x+1)-q_{t+2}(x)$, and compare it to $q_{t+1}(x)$. Hence we must consider the system of equations

$$
\begin{aligned}
b_{t-1} &= \binom{t}{1}a_t,\\
b_{t-2} &= \binom{t}{2}a_t + \binom{t-1}{1}a_{t-1},\\
b_{t-3} &= \binom{t}{3}a_t + \binom{t-1}{2}a_{t-1} + \binom{t-2}{1}a_{t-2},\\
&\vdots\\
b_1 &= \binom{t}{t-1}a_t + \binom{t-1}{t-2}a_{t-1} + \cdots + \binom{2}{1}a_2,\\
b_0 &= \binom{t}{t}a_t + \binom{t-1}{t-1}a_{t-1} + \cdots + \binom{2}{2}a_2 + \binom{2}{2}a_2.
\end{aligned}
$$

From the top of the list of equations to down one can consecutively determine the coefficients $a_t,\ a_{t-1},\ \ldots,\ a_1$. Then the constant term a_0 follows from the equality $q_{t+2}(t+2)=F_{2t+4}-F_{t+2}$.

At the end of this section recall that $\sum_{k=0}^{n}\left\langle{n \atop k}\right\rangle_{F_n}=2^{n+1}-2F_{n+1}$ ([2], after Example 1). Combining the former expression with (3.1), we have the following

3.4. Corollary. *For any nonnegative integer n, the identity*

$$\sum_{k=0}^{n} q_k(n) = F_{2n+2} + F_{n+1} - 2^{n+1}$$

holds.

The nonexistence of closed form for (2.1) was the motivation to work out a different approach for Fibonacci Pascal triangle. For the specific case $n=2k$ we see, that the two sums in (1.3) coincide. This observation, together with Theorem 3.1 implies

3.5. Corollary. *If k is a positive integer, then we get*

$$\sum_{i=1}^{k}\binom{2k-1-i}{k-1}F_i = \frac{F_{3k}-q_k(2k)}{2}.$$

3.2. Lucas Pascal triangle. After studying Fibonacci Pascal triangle, it is natural to consider Lucas Pascal triangle, i.e. when $a_n=b_n=L_n$ is the n^{th} term of the Lucas sequence $\{L_n\}_{n=0}^{\infty}$ (see Figure 4).

Without detailing the proofs (only follow the maintance of Fibonacci PT), we yield the main result and the corresponding lemmata.

3.6. Theorem. *For any nonnegative integers n and k, we have*

$$\left\langle{n \atop k}\right\rangle_{L_n} = L_{n+k} - r_k(n),$$

where

$$r_k(x) = 2\sum_{j=0}^{\lfloor (k-1)/2\rfloor}\binom{x}{k-1-2j}F_{2j+1}$$

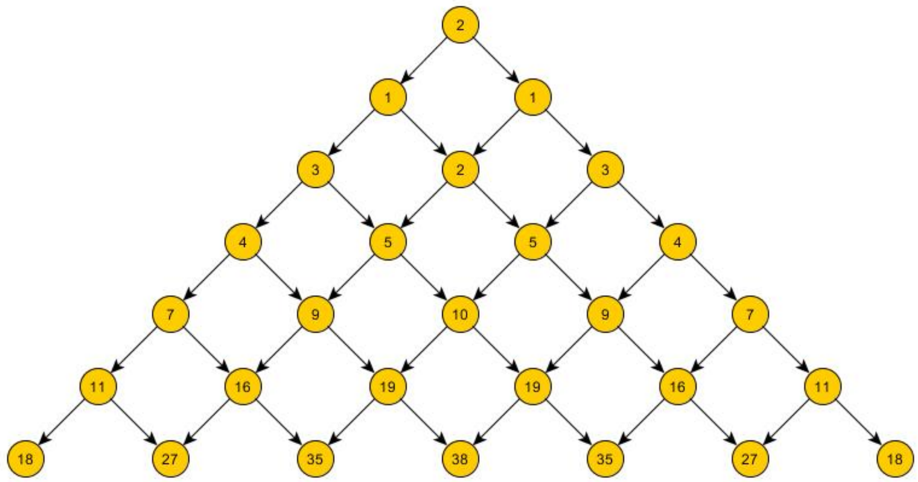


Figure 4. Lucas Pascal triangle

is a rational polinomial of degree $k-1$ *if* $k \geq 1$*, and* $r_0(x) = 0$.

Proof. Without going into details, we follow the method used in the case of $a_n = b_n = F_n$. □

The lemmata we need are the following.

3.7. Lemma. *The equality* $r_k(k) = L_{2k} - L_k$ *holds for any nonnegative integer* k.

3.8. Lemma. *For any nonnegative integers* N *and* k*, we have* $r_k(N+1) - r_k(N) = r_{k-1}(N)$.

The first few polynomials $r_k(x)$ are listed here.

k	1	2	3	4	5
$r_k(x)$	2	$2x$	$x^2 - x + 8$	$\frac{1}{3}(x^3 - 3x^2 + 26x)$	$\frac{1}{12}(x^4 - 6x^3 + 59x^2 - 54x + 264)$

3.3. Ensley's Fibonacci Triangle and a generalization. Now we are ready to handle Ensley's Fibonacci Triangle, when $a_n = b_n = F_{n+1}$, by exploiting the results on Fibonacci and Lucas Triangles.

Assume generally, that the sequence $\{G_n\}$ satisfies the recursive rule

$$G_n = G_{n-1} + G_{n-2} \qquad (n \geq 2) \tag{3.3}$$

with the initial values G_0 and G_1. It is well known, that such a sequence can be given by a linear combination of any two linearly independent recurrences (like Fibonacci and Lucas sequences), which satisfy (3.3). Taking the Fibonacci and Lucas sequences as basis, the solution of the vector equation

$$\begin{bmatrix} F_0 \\ F_1 \end{bmatrix} x + \begin{bmatrix} L_0 \\ L_1 \end{bmatrix} y = \begin{bmatrix} G_0 \\ G_1 \end{bmatrix},$$

(via $F_0 = 0$, $F_1 = 1$, $L_0 = 2$ and $L_1 = 1$) is $y = G_0/2$, $x = (2G_1 - G_0)/2$. That is,

$$G_n = \frac{2G_1 - G_0}{2} F_n + \frac{G_0}{2} L_n,$$

and, by Theorems 3.1 and 3.6, and the principle of superposition we have the following general theorem.

3.9. Theorem. *The GPT generated by $a_n = b_n = G_n$ satisfies*

$$\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle_{G_n} = G_{n+k} - \frac{(2G_1 - G_0)q_k(n) + G_0 r_k(n)}{2}.$$

Specifying $G_n = F_{n+1}$, we have

$$\begin{bmatrix} G_0 \\ G_1 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

hence now $x = y = 1/2$. Thus $F_{n+1} = (F_n + L_n)/2$, and we conclude

3.10. Corollary. *For any nonnegative integers n and k, we conclude*

$$\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle_{F_{n+1}} = F_{n+k+1} - \frac{q_k(n+1) + r_k(n+1)}{2}.$$

Acknowledgements. We thank the referee for valuable comments that greatly improved the quality of this paper.

References

[1] Abramowitz, M. and Stegun, I. A., *Handbook of mathematical functions with formulas, graphs and mathematical tables*, Applied Mathematics Series 55 (10 ed.), New York, USA, 1972.

[2] Belbachir, H. and Szalay, L., *On the arithmetic triangles*, Šiauliai Math. Sem., **9** (17), 15-26, 2014.

[3] Dil, A. and Mező, I., *A symmetric algorithm for hyperharmonic and Fibonacci numbers*, Appl. Math. Comp., **206**, 942-951, 2008.

[4] Ensley, D., *Fibonacci's triangle and other abominations*, in: The Edge of the Universe: Celebrating Ten Years of Maths Horizons, MAA, 287-307, 2006.

[5] Ferri, G., Faccio, M. and D'Amico, A., *A new numerical triangle showing links with Fibonacci numbers*, Fibonacci Quart., **29**, 316-321, 1991.

[6] Ferri, G., Faccio, M. and D'Amico, A., *Fibonacci numbers and ladder network impedance*, Fibonacci Quart., **30**, 62-67, 1992.

[7] Hosoya, H., *Fibonacci triangle*, Fibonacci Quart., **14**, 173-178, 1976.

[8] Hosoya, H., *Pascal's triangle, non-adjacent numbers, and D-dimensional atomic orbitals*, J. Math. Chemistry, **23**, 169-178, 1998.

[9] Koshy, T., *Fibonacci and Lucas Numbers with Applications*, Pure and Applied Mathematics, A Wiley-Interscience Series of Texts, Monographs, and Tracts, 2001.

[10] Pascal, B., *Traité du Triangle Arithmétique, avec quelques autres petitez sur la mesme matiere*, Paris, 1665.

[11] Raab, J. A., *A generalization of the connection between the Fibonacci sequence and Pascal's triangle*, Fibonacci Quart., **1**, 21-31, 1963.

[12] Sun, Y., *Numerical triangles and several classical sequences*, Fibonacci Quart., **43**, 359-370, 2005.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1355 – 1366

Weakly second modules over noncommutative rings

R. Beyranvand*† and F. Rastgoo‡

Abstract

Let R be an arbitrary ring. In this paper we will introduce the concept of a weakly second R-module (a generalization of the second R-module) and we will obtain some related results.

Keywords: Weakly second module; second module; prime module; weakly prime module.

2000 AMS Classification: 16D10; 16D90; 16N60

Received : 18.08.2015 *Accepted* : 06.11.2015 *Doi* : 10.15672/HJMS.20164515691

1. Introduction

Throughout this paper, all rings will have identity elements and all modules will be right unitary. We use the notation "$\subset$" to denote strict inclusion. Unless otherwise stated, R denotes an arbitrary ring with identity element. Let M be an R-module. Then the annihilator of M (in R) is the ideal $\mathrm{ann}_R(M) = \{r \in R \mid Mr = 0\}$. Also for any submodule N of M and any ideal I of R, the submodule $\{x \in M \mid xI \subseteq N\}$ of M is denoted by $(N :_M I)$.

Recall that a nonzero R-module M is *prime* if $\mathrm{ann}_R(M) = \mathrm{ann}_R(N)$ for every nonzero submodule N of M. Also a nonzero R-module M is called *weakly prime* in case $\mathrm{ann}_R(N)$ is a prime ideal of R for every nonzero submodule N of M. By a (*weakly*) *prime submodule* of a module M we mean a submodule N such that the module M/N is (weakly) prime. The notion of prime modules first was introduced by Dauns in [11]. Also in [9], Behboodi and Koohi introduced the notion of weakly prime modules and investigated the properties of this class of modules. More details about prime modules and weakly prime modules can be found in [2, 5, 6, 9].

On the other hand, a nonzero module M is called a *second module* (the dual notion of a prime module) provided $\mathrm{ann}_R(M) = \mathrm{ann}_R(M/N)$ for every proper submodule N of M. This notion was introduced by Yassemi in [15], for modules over commutative rings.

*Department of Mathematics, Lorestan University, Khorramabad, Iran
Email: beyranvand.r@lu.ac.ir

†Corresponding Author.

‡Department of Mathematics, Lorestan University, Khorramabad, Iran,
Email: rastgoo_fahimeh@yahoo.com

Moreover, in [10], the authors generalized second modules from commutative rings to noncommutative setting.

The purpose of this paper is to introduce and study the concept of weakly second modules (the dual of weakly prime modules) over noncommutative rings. A nonzero R-module M is called a *weakly second module* if $\text{ann}_R(M/N)$ is a prime ideal of R for every proper submodule N of M. It is clear that every second module is a weakly second module. By a (*weakly*) *second submodule* of a module we mean a submodule which is also a (weakly) second module. In addition to obtaining some useful information about this class of modules, we investigate which dual of the given results about the weakly prime modules hold for the weakly second ones. For a right R-module M, among other results, we prove the following statements:

• **Theorem 2.3.** ((1), (2), (8)) *M is a weakly second module if and only if for every two ideals I and J of R, $MIJ = MI$ or $MIJ = MJ$, if and only if the set $\{\text{ann}_R(M/N) \mid N$ is a proper submodule of $M\}$ is a chain of prime ideals of R.*

• **Theorem 2.5** and **Theorem 2.6.** *Secondness and weakly secondness are Morita invariant properties.*

• **Proposition 2.7.** *M is weakly second if and only if for every proper submodule K of M, there is a prime ideal I of R contained in $\text{ann}_R(M/K)$ such that M/K cogenerates R/I.*

• **Corollary 2.9.** *If R is a right Artinian ring, then M is a weakly second R-module if and only if M is a homogenous semisimple R-module.*

• **Theorem 2.15.** *If M satisfies the descending chain condition on weakly socle submodules, then every nonzero submodule of M has only a finite number of maximal weakly second submodule.*

• **Proposition 2.17.** *If R satisfies the ascending chain condition on prime ideals, then M has a second submodule if and only if it has a weakly second submodule.*

• **Theorem 3.5.** *Let R be a ring whose two-sided ideals satisfy* ACC. *Then a nonzero R-module M is weakly second if and only if for every two prime ideals I and J of R, $MIJ = MI$ or $MIJ = MJ$.*

• **Proposition 3.7.** *Let R be a prime right Goldie ring. Then every nonzero properly divisible R-module is a weakly second module.*

2. Weakly second modules

We begin this section with the definition of weakly second modules and then some remarks and examples are given.

2.1. Definition. A nonzero right R-module M is called *weakly second* if for every proper submodule N of M, $\text{ann}_R(M/N)$ is a prime ideal of R.

2.2. Examples. (a) The $\mathbb{Z}$-module $\mathbb{Z}_n$ is a weakly second if and only if n is a prime number.

(b) If $n \neq m$, then the $\mathbb{Z}$-module $\mathbb{Z}_n \oplus \mathbb{Z}_m$ is not weakly second.

(c) Let D be a division ring and $V = \oplus_{i=1}^{\infty} e_i D$ be a vector space over D. Set $R = \text{End}(V_D)$ and $T = \{f \in R \mid \text{rank} f < \infty\}$. It is known that R has only three ideals (0), R and T. So T is a maximal ideal and (0) is a prime ideal of R. Now it is easy to check that R as a left R-module is weakly second but is not a second R-module.

(d) Let p and q be two distinct prime numbers. Consider the $\mathbb{Z}$-modules

$$M = < 1/p + \mathbb{Z} > \oplus < 1/q + \mathbb{Z} > \oplus \mathbb{Z}_{p^\infty},$$

and

$$N = < 1/p + \mathbb{Z} > \oplus (0) \oplus \mathbb{Z}_{p^\infty}.$$

It is easily checked that N and M/N are weakly second modules, but M is not a weakly second module.

(e) Recall that a submodule N of an R-module M is *fully invariant* if for every R-endomorphism $f : M \to M$, $f(N) \subseteq N$. A right R-module M is a weakly second module if and only if for every fully invariant proper submodule N of M, $\mathrm{ann}_R(M/N)$ is a prime ideal of R. To see this, let I and J be two ideals of R and L be a proper submodule of M such that $IJ \subseteq \mathrm{ann}_R(M/L)$. Since MIJ is a fully invariant proper submodule of M, $\mathrm{ann}_R(M/MIJ)$ is a prime ideal of R. Now $IJ \subseteq \mathrm{ann}_R(M/MIJ)$ implies that $MI \subseteq MIJ \subseteq L$ or $MJ \subseteq MIJ \subseteq L$, and so $I \subseteq \mathrm{ann}_R(M/L)$ or $J \subseteq \mathrm{ann}_R(M/L)$. Thus $\mathrm{ann}_R(M/L)$ is prime.

In the following theorem, some characterizations of weakly second modules are given.

2.3. Theorem. *For a nonzero right R-module M, the following statements are equivalent:*

(1) *M is a weakly second module;*
(2) *For every two ideals I and J of R, $MIJ = MI$ or $MIJ = MJ$;*
(3) *For every two ideals I and J of R, $J \nsubseteq \mathrm{ann}_R(M/MIJ)$ implies that $MIJ = MI$;*
(4) *For every two ideals I and J of R, $I \nsubseteq \mathrm{ann}_R(M/MIJ)$ implies that $MIJ = MJ$;*
(5) *For every two ideals I and J of R, $\mathrm{ann}_R(M/MIJ) \subset I$ implies that $MIJ = MJ$;*
(6) *For every two ideals I and J of R, $\mathrm{ann}_R(M/MIJ) \subset J$ implies that $MIJ = MI$;*
(7) *Every nonzero quotient of M is weakly second;*
(8) *The set $\{\mathrm{ann}_R(M/N) \mid N$ is a proper submodule of $M\}$ is a chain of prime ideals of R.*

Proof. (1) $\Rightarrow$ (2). Suppose that M is weakly second and I and J are two ideals of R. If $MIJ = M$, then $MIJ = MI$ and $MIJ = MJ$. So suppose that MIJ is a proper submodule of M and hence $\mathrm{ann}_R(M/MIJ)$ is a prime ideal of R. Since $IJ \subseteq \mathrm{ann}_R(M/MIJ)$, $I \subseteq \mathrm{ann}_R(M/MIJ)$ or $J \subseteq \mathrm{ann}_R(M/MIJ)$. Thus $MIJ = MI$ or $MIJ = MJ$.

(2) $\Rightarrow$ (1). Let N be a proper submodule of M and $IJ \subseteq \mathrm{ann}_R(M/N)$ for some two ideals I and J of R. Then $MIJ \subseteq N$ and by the hypothesis $MI \subseteq N$ or $MJ \subseteq N$. Thus $\mathrm{ann}_R(M/N)$ is prime.

(2) $\Leftrightarrow$ (3) and (2) $\Leftrightarrow$ (4) are clear.

(4) $\Rightarrow$ (5). Let I and J be two ideals of R such that $\mathrm{ann}_R(M/MIJ) \subset I$. Then $I \nsubseteq \mathrm{ann}_R(M/MIJ)$ and so $MIJ = MJ$.

(5) $\Rightarrow$ (4). Suppose that I and J are two ideals of R and $I \nsubseteq \mathrm{ann}_R(M/MIJ)$. Then $\mathrm{ann}_R(M/MIJ) \subset I + \mathrm{ann}_R(M/MIJ)$ and we have $\mathrm{ann}_R(M/MIJ) = \mathrm{ann}_R(M/M(I + \mathrm{ann}_R(M/MIJ))J)$. Thus $\mathrm{ann}_R(M/M(I + \mathrm{ann}_R(M/MIJ))J) \subset I + \mathrm{ann}_R(M/MIJ)$ and by (5), $M(I + \mathrm{ann}_R(M/MIJ))J = MJ$. But $M(I + \mathrm{ann}_R(M/MIJ))J \subseteq MIJ$ and so $MIJ = MJ$.

(3) $\Rightarrow$ (6) is similar to (4) $\Rightarrow$ (5) and (6) $\Rightarrow$ (3) is similar to (5) $\Rightarrow$ (4).

(1) $\Rightarrow$ (7) is clear.

(7) $\Rightarrow$ (8). Let N_1 and N_2 be two proper submodules of M, $P = \text{ann}_R(M/N_1)$ and $Q = \text{ann}_R(M/N_2)$. If $P \nsubseteq Q$ and $Q \nsubseteq P$, then there exist two ideals I_1 and I_2 of R such that $I_1 \subseteq P$, $I_2 \subseteq Q$, $I_1 \nsubseteq Q$ and $I_2 \nsubseteq P$. Since $\text{ann}_R(M/(N_1 \cap N_2))$ is a prime ideal of R and $I_1 I_2 \subseteq \text{ann}_R(M/(N_1 \cap N_2))$, $I_1 \subseteq \text{ann}_R(M/(N_1 \cap N_2))$ or $I_2 \subseteq \text{ann}_R(M/(N_1 \cap N_2))$. But $\text{ann}_R(M/(N_1 \cap N_2)) \subseteq \text{ann}_R(M/N_2) = Q$ and $\text{ann}_R(M/(N_1 \cap N_2)) \subseteq \text{ann}_R(M/N_1) = P$. Thus $I_1 \subseteq Q$ or $I_2 \subseteq P$, a contradiction.

(8) $\Rightarrow$ (1). Clear. □

Next, we show that both secondness and weakly secondness are Morita invariant properties.

2.4. Theorem. *Secondness is a Morita invariant property.*

Proof. Let R and S be Morita equivalent rings via an equivalence $F : Mod_R \to Mod_S$. Suppose that M is a second R-module. Let $I = \text{ann}_R(M)$ and $B = \text{ann}_S(F(R/I))$. By [1, Proposition 21.11], R/I is Morita equivalent to S/B. Also by [1, Proposition 21.6], $F(M)$ is faithful as an S/B-module because M is a faithful R/I-module. Thus $B = \text{ann}_S(F(M))$. Now assume that N is a proper S-submodule of $F(M)$. We show that $\text{ann}_S(F(M)) = \text{ann}_S(F(M)/N)$. For a submodule K of M, let $i_{K \leq M} : K \to M$ denote the inclusion monomorphism. Since by [1, Proposition 21.7], the mapping defined by $\Lambda_M : K \to Im\ F(i_{K \leq M})$ is a lattice isomorphism from the lattice of submodules of M onto the lattice of submodules of $F(M)$, there exists $K \leq M$ such that $\Lambda_M(K) = N = Im\ F(i_{K \leq M})$. Since Morita equivalences preserve exactness, $F(M)/Im\ F(i_{K \leq M}) \cong F(M/K)$. Therefore $F(M)/N \cong F(M/K)$ and so $\text{ann}_S(F(M)/N) = \text{ann}_S(F(M/K))$. On the other hand, since M is a second R-module, $I = \text{ann}_R(M) = \text{ann}_R(M/K)$ and so by the first part of the proof, $B = \text{ann}_S(F(M)) = \text{ann}_S(F(M/K))$. Thus $\text{ann}_S(F(M)) = \text{ann}_S(F(M)/N)$ and this implies that $F(M)$ is a second S-module, as desired. □

2.5. Theorem. *Weakly secondness is a Morita invariant property.*

Proof. Let R and S be Morita equivalent rings via an equivalence $F : Mod_R \to Mod_S$. Suppose that M is a weakly second R-module and N is a proper S-submodule of $F(M)$. We show that $\text{ann}_S(F(M)/N)$ is a prime ideal of S. In the notations of the proof of above theorem, there exists $K \leq M$ such that $\Lambda_M(K) = N = Im\ F(i_{K \leq M})$ and $\text{ann}_S(F(M)/N) = \text{ann}_S(F(M/K))$. Let $I = \text{ann}_R(M/K)$. Again by the proof of above theorem, R/I is Morita equivalent to S/B, where $B = \text{ann}_S(F(R/I)) = \text{ann}_S(F(M/K))$. Since M is a weakly second R-module, $I = \text{ann}_R(M/K)$ is a prime ideal of R and hence R/I is a prime ring. By [13, Corollary 18.45], S/B is also a prime ring. Thus $B = \text{ann}_S(F(M/K)) = \text{ann}_S(F(M)/N)$ is a prime ideal of S, as desired. □

2.6. Remark. For a nonzero R-module M, we note that M is a second R-module if and only if for any proper submodule K of M, M/K cogenerates $R/\text{ann}_R(M)$. To see this, suppose that M is second and K is a proper submodule of M. Then $r + \text{ann}_R(M) \to (xr + K)_{x \in M}$ is an R-monomorphism of $R/\text{ann}_R(M)$ into $\prod_{x \in M} M/K$. Thus $R/\text{ann}_R(M)$ is cogenerated by M/K. For the other direction, assume that $f : R/\text{ann}_R(M) \to \prod_{\alpha \in A} M/K$ is an R-monomorphism, where K is a proper submodule of M. Let $f(\overline{1}) = (x_\alpha + K)_{\alpha \in A}$. If $r \in \text{ann}_R(M/K)$, then $f(\overline{r}) = (x_\alpha r + K)_{\alpha \in A} = 0$ and so $\overline{r} = 0$. Thus $r \in \text{ann}_R(M)$. This yields that $\text{ann}_R(M) = \text{ann}_R(M/K)$ and hence M is a second module.

Now for weakly second modules, we have the following statement.

2.7. Proposition. *Let M be a nonzero R-module. Then M is weakly second if and only if for every proper submodule K of M, there is a prime ideal I of R contained in $ann_R(M/K)$ such that M/K cogenerates R/I.*

Proof. First suppose that M is a weakly second module and K is a proper submodule of M. Then $\text{ann}_R(M/K)$ is a prime ideal and clearly M/K cogenerates $R/\text{ann}_R(M/K)$. Conversely, suppose that K is a proper submodule of M and I is a prime ideal of R contained in $\text{ann}_R(M/K)$ such that R/I is cogenerated by M/K. Say $f : R/I \to \prod_{\alpha\in A} M/K$ is an R-monomorphism. Then $\text{ann}_R(M/K) = \text{ann}_R(\prod_{\alpha\in A} M/K) \subseteq \text{ann}_R(R/I) \subseteq I$ and hence $I = \text{ann}_R(M/K)$. □

Recall that a module M is *homogeneous semisimple* if M is a direct sum of pairwise isomorphic simple submodules. Clearly, any homogeneous semisimple module is (weakly) second. We show that the converse is true when R is an Artinian ring. First the following lemma is needed.

2.8. Lemma. *Let R be a ring in which every prime ideal is maximal. For a nonzero R-module M, consider the following statements:*
(1) *M is prime;*
(2) *M is weakly prime;*
(3) *M is second;*
(4) *M is weakly second;*
(5) *M is homogeneous semisimple.*

Then (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4) $\Leftarrow$ (5). *Moreover, if in addition R is a commutative ring, then all five statements are equivalent.*

Proof. (1) $\Rightarrow$ (2), (3) $\Rightarrow$ (4) and (5) $\Rightarrow$ (4) are trivial.
(2) $\Rightarrow$ (3). Suppose that M is weakly prime. Then $\text{ann}_R(M)$ is a prime ideal and hence is a maximal ideal. Thus for any proper submodule N of M, $\text{ann}_R(M) \subseteq \text{ann}_R(M/N) \subset R$ implies that $\text{ann}_R(M) = \text{ann}_R(M/N)$ and so M is a second module.
(4) $\Rightarrow$ (1). Suppose M is a weakly second module. Then $\text{ann}_R(M)$ is a prime ideal and hence is a maximal ideal. Thus $\text{ann}_R(M) \subseteq \text{ann}_R(N) \subset R$ implies that $\text{ann}_R(M) = \text{ann}_R(N)$, for every nonzero submodule N of M. It follows that M is a prime module.
(4) $\Rightarrow$ (5). Suppose that R is a commutative ring and M is a weakly second R-module. Then $\text{ann}_R(M)$ is a prime ideal and by the hypothesis, it is a maximal ideal. Thus R/P is a field where $P = \text{ann}_R(M)$. This implies that M is a homogeneous semisimple as R/P-module and as R-module. □

2.9. Corollary. *Let R be a right Artinian ring. Then for any nonzero right R-module M, the five statements in the previous lemma are all equivalent.*

Proof. We only prove that if M is a weakly second R-module, then it is a homogeneous semisimple R-module. Since M is weakly second, $\text{ann}_R(M)$ is a prime ideal and so R/P is a right Artinian prime ring where $P = \text{ann}_R(M)$. By the Wedderburn-Artin Theorem [12, Theorem 3.5], we conclude that M is a homogeneous semisimple as R/P-module and as R-module. □

2.10. Corollary. *Let R be a commutative von Neumann regular ring and M be an R-module. Then the following statements are equivalent:*

(1) *M is second;*
(2) *M is weakly second;*
(3) *M is homogeneous semisimple.*

Proof. It is well known that every prime ideal in a commutative von Neumann regular ring is a maximal ideal. Now apply Lemma 2.8. □

The next two results were proved for second modules. See [10, Corollary 2.4 and Proposition 3.6].

2.11. Corollary. *Let A be an ideal of a ring R and M be a nonzero right R-module such that $MA = 0$. Then the R-module M is a weakly second module if and only if the R/A-module M is a weakly second module.*

Proof. Suppose first that the R-module M is weakly second and let I and J be two ideals of R containing A. Then $M(I/A)(J/A) = M(IJ + A/A) = MIJ$ and so by Theorem 2.3, $M(I/A)(J/A) = M(I/A)$ or $M(I/A)(J/A) = M(J/A)$. Conversely, suppose that the R/A-module M is a weakly second module. For any two ideals I and J of R, $MIJ = M(IJ+A) = M((IJ+A)/A) = M((I+A)/A)((J+A)/A)$. Using Theorem 2.3, we have $MIJ = M((I + A)/A) = MI$ or $MIJ = M((J + A)/A) = MJ$, as desired. □

Let R be a ring. An ideal A of R is called *right T-nilpotent* if for any sequence $\{a_1, a_2, \ldots\}$ in A, there exists a positive integer n such that $a_n \ldots a_1 = 0$.

2.12. Proposition. *Let A be a right T-nilpotent ideal of a ring R and $\overline{R} = R/A$. Then every nonzero right R-module M has a proper submodule N such that M/N is a weakly second R-module if and only if, every nonzero right $\overline{R}$-module M has a proper submodule N such that M/N is a weakly second $\overline{R}$-module.*

Proof. Suppose first that every nonzero right R-module M has a proper submodule N such that M/N is a weakly second R-module. Let K be a nonzero right $\overline{R}$-module. Then K is a right R-module and by the hypothesis, there exists a proper submodule L of K such that the R-module K/L is weakly second. It follows that the $\overline{R}$-module K/L is a weakly second module. Conversely, suppose that every nonzero right $\overline{R}$-module M has a proper submodule N such that M/N is a weakly second $\overline{R}$-module. Let X be a nonzero right R-module. Then by [1, Lemma 28.3], $X \neq XA$ and so $X/(XA)$ is a nonzero right $\overline{R}$-module. Also by the hypothesis, there exists a proper submodule Y of X containing XA, such that the $\overline{R}$-module X/Y is weakly second. Now by Corollary 2.11, X/Y is a weakly second R-module. □

In [10, Proposition 4.2], it is shown that the union of a chain of second submodules of a module is also second. Here, we show that a similar result holds for a directed set of weakly second submodules of a module.

2.13. Lemma. *Let R be a ring, and let N_i $(i \in I)$ be a directed set of weakly second submodules of a right R-module M. Then $N = \bigcup_{i \in I} N_i$ is a weakly second R-module.*

Proof. Note that N is a nonzero submodule of M. Let A and B be two ideals of R. By Theorem 2.3, it suffices to show that $NAB = NA$ or $NAB = NB$. If there exists $k \in I$ such that for each $i \in I$, $N_j = \sup\{N_i, N_k\}$ satisfies $N_jAB = N_jA$, then for each $i \in I$, we have $N_iA \subseteq N_jA = N_jAB \subseteq NAB$. Thus $NA = NAB$. Now suppose that for every $k \in I$, there exists $i \in I$ such that $N_j = \sup\{N_i, N_k\}$ dose not satisfy $N_jAB = N_jA$. Since N_j is weakly second, $N_jB = N_jAB$. Then for each $i \in I$, we have $N_iB \subseteq N_jB = N_jAB \subseteq NAB$. Thus $NB = NAB$. □

By a maximal weakly second submodule of a module M, we mean a weakly second submodule L of M such that L is not properly contained in another weakly second submodule of M.

2.14. Corollary. *Let M be any nonzero module. Then every weakly second submodule of M contained in a maximal weakly second submodule.*

Proof. By Lemma 2.13 and Zorn's Lemma. □

Let N be a submodule of a right R-module M. We define the *weakly socle* of N as the sum of all weakly second submodules of M contained in N, denoted by $W.soc(N)$. The weakly socle of N is defined to be (0) in case N dose not contain any weakly second submodule. N is said to be a weakly socle submodule of M if $N \neq 0$ and $W.soc(N) = N$.

2.15. Theorem. *Let M be a right R-module. If M satisfies the descending chain condition on weakly socle submodules, then every nonzero submodule of M has only a finite number of maximal weakly second submodule.*

Proof. Suppose that the result is false. Then there exists a nonzero submodule N of M such that it has an infinite number of maximal weakly second submodules. Thus $W.soc(N)$ is a weakly socle submodule of M and $W.soc(N)$ has an infinite number of maximal weakly second submodules. By the assumption, let S be a weakly socle submodule of M chosen minimal such that S has an infinite number of maximal weakly second submodules. If S is weakly second, then every maximal weakly second submodule contained in S is equal to S. Thus S has not an infinite number of maximal weakly second submodules, a contradiction. Therefore S is not weakly second and so there exist two ideals I and J of R and a proper submodule K of M such that $SIJ \subseteq K$, $SI \nsubseteq K$ and $SJ \nsubseteq K$. Thus $S \nsubseteq (K :_S I)$ and $S \nsubseteq (K :_S J)$. Therefore $S \nsubseteq W.soc((K :_S I))$ and $S \nsubseteq W.soc((K :_S J))$. Now we conclude that $W.soc((K :_S I)) \subset S$ and $W.soc((K :_S I)) \subset S$. Let V be a maximal weakly second submodules of M contained in S. Then $VIJ \subseteq SIJ \subseteq K$ and hence $VI \subseteq K$ or $VJ \subseteq K$. Thus $V \subseteq (K :_S I)$ or $V \subseteq (K :_S J)$ so that $V \subseteq W.soc((K :_S I))$ or $V \subseteq W.soc((K :_S J))$. The minimality of S, implies that both $W.soc((K :_S I))$ and $W.soc((K :_S J))$ have only finitely many maximal weakly second submodules. Therefore there is only a finite number of possibilities for the module S, which is a contradiction. □

The following result is immediately obtained.

2.16. Corollary. *Every nonzero Artinian module contains only a finite number of maximal weakly second submodule.*

Clearly, if an R-module has a second submodule, then it has a weakly second submodule. Now, we show that the converse is true when a certain set of ideals of R has the descending chain condition (briefly, DCC). In fact, it is the dual statement of [9, Proposition 5.1].

2.17. Proposition. *Let R be a ring whose prime ideals satisfy DCC and let M be a right R-module. Then M has a second submodule if and only if it has a weakly second submodule.*

Proof. Suppose that N is a weakly second submodule of M. Let $I = \mathrm{ann}_R(N)$. Since I is a prime ideal, R/I is a prime ring and so N is a faithful R/I-module. Without loss of generality, we may assume that R is a prime ring and M is a faithful weakly second module. By Theorem 2.3, the set $T = \{\ \mathrm{ann}_R(M/K) \mid K \text{ is a proper submodule of } M\}$ is a chain of prime ideals. If $T = \{0\}$, then M is a second module and we are through. Thus suppose that the chain T contains a nonzero element. Let $L_0 = \bigcap\{L \subset M \mid 0 \neq \mathrm{ann}_R(M/L) \in T\}$. Clearly, L_0 is a submodule of M. By the hypothesis, assume that P is a minimal among nonzero elements of T. Then $P = \mathrm{ann}_R(M/K)$ for some proper submodule K of M. We claim that $P = \mathrm{ann}_R(M/L_0)$. Since T is a chain, for any proper submodule L of M with $\mathrm{ann}_R(M/L) \neq 0$, we have $\mathrm{ann}_R(M/L) \subseteq \mathrm{ann}_R(M/K)$ or $\mathrm{ann}_R(M/K) \subseteq \mathrm{ann}_R(M/L)$. The minimality of P implies that $P \subseteq \mathrm{ann}_R(M/L)$. Thus $MP \subseteq L$ for any proper submodule L of M with $\mathrm{ann}_R(M/L) \neq 0$. It follows that $MP \subseteq L_0$ and hence $P \subseteq \mathrm{ann}_R(M/L_0)$. By the definition of L_0, we have $P =$

$\text{ann}_R(M/L_0)$. Now we show that L_0 is a second submodule of M. Suppose that I is an ideal of R and $L_0I \neq 0$. Since $MP \subseteq L_0$, $MPI \subseteq L_0I$. But R is a prime ring and so $PI \neq 0$. Thus $\text{ann}_R(M/L_0I) \neq 0$ and by the definition of L_0, we conclude that $L_0 \subseteq L_0I$ and hence $L_0 = L_0I$, as desired. □

3. Further results related to weakly second modules

In this section, we start with some definitions. Then the relationships between weakly second, weakly prime and second modules are investigated. Let R be a ring and M be a right R-module. Then:

(i) M is called a *multiplication module* if for every submodule N of M there exists an ideal I of R such that $N = MI$. This notion was introduced by Baranard in [7].

(ii) M is called a *comultiplication module* if for every submodule N of M there exists an ideal I of R such that $(0 :_M I) = N$. This notion is introduced by Ansari-Toroghy and Farshadifar in [3].

(iii) A submodule N of M is called *secondary submodule* if for every ideal I of R, $NI = N$ or there exists a positive integer n such that $NI^n = 0$.

3.1. Theorem. *Let R be a ring and M be a nonzero right R-module. Then:*

(1) *If M is a multiplication module such that $\text{ann}_R(M)$ is a prime ideal, then it is prime;*
(2) *If M is a comultiplication module such that $\text{ann}_R(M)$ is a prime ideal, then it is second;*
(3) *A submodule N of M is second if and only if it is both a weakly second and a secondary submodule of M;*
(4) *If any two prime ideals of R are comparable, i.e., $I \subseteq J$ or $J \subseteq I$ for every two prime ideals I and J of R, then any sum of weakly second submodules of M is a weakly second submodule of M.*

Proof. (1) Suppose that N is a nonzero submodule of M and $NI = 0$, where I is an ideal of R. Since M is multiplication, $N = MJ$ for some ideal J of R. Then $NI = MJI = 0$ and so $MI = 0$ because $\text{ann}_R(M)$ is prime. Thus M is a prime module.

(2) Suppose that N is a proper submodule of M and $MI \subseteq N$, where I is an ideal of R. Since M is comultiplication, $N = (0 :_M J)$ for some ideal J of R. Then $MI \subseteq (0 :_M J)$ and so $MIJ = 0$. Since $\text{ann}_R(M)$ is prime and $N \neq M$, we have $MI = 0$. Thus $\text{ann}_R(M) = \text{ann}_R(M/N)$ and hence M is a second module.

(3) For one direction, the proof is clear. For the other direction, assume that N is both a secondary and a weakly second submodule of M. Let I be an ideal of R such that $NI \neq 0$ and $NI \neq N$. Since N is secondary, there exists $n \geq 2$ such that $NI^n = 0$. On the other hand, since N is weakly second, we conclude that $NI = 0$, a contradiction.

(4) Let $\{N_i\}_{i\in I}$ be a collection of weakly second submodules of M and $N = \sum_{i\in I} N_i$. Clearly $N \neq 0$. Since for any $i \in I$, N_i is weakly second, $\text{ann}_R(N_i)$ is a prime ideal of R. Also $\text{ann}_R(N) = \text{ann}_R(\sum_{i\in I} N_i) = \bigcap_{i\in I} \text{ann}_R(N_i)$ and since any two prime ideals of R are comparable, $\text{ann}_R(N)$ is a prime ideal of R. To complete the proof, it is enough to show that $(L :_R N)$ is a prime ideal of R, where L is a submodule of M such that $N \nsubseteq L$. Assume for two ideals A and B of R, $NAB \subseteq L$ such that $NA \nsubseteq L$ and $NB \nsubseteq L$. Then there exist $i, j \in I$ such that $N_iA \nsubseteq L$ and $N_jB \nsubseteq L$. This implies that

$$A \nsubseteq \text{ann}_R\left(\tfrac{N_i}{L\cap N_i}\right) \quad \text{and} \quad B \nsubseteq \text{ann}_R\left(\tfrac{N_j}{L\cap N_j}\right). \qquad (*)$$

Now since N_i and N_j are weakly second submodules of M, $\mathrm{ann}_R(\frac{N_i}{L\cap N_i})$ and $\mathrm{ann}_R(\frac{N_j}{L\cap N_j})$ are prime and by the assumption, these are comparable. Without loss of generality, we may assume that

$$\mathrm{ann}_R(\tfrac{N_i}{L\cap N_i}) \subseteq \mathrm{ann}_R(\tfrac{N_j}{L\cap N_j}). \qquad (**)$$

On the other hand, $NAB \subseteq L$ implies that $N_iAB \subseteq L$ and $N_jAB \subseteq L$. Thus $N_iAB \subseteq L\cap N_i$ and $N_jAB \subseteq L\cap N_j$ and so $AB \subseteq \mathrm{ann}_R(\frac{N_i}{L\cap N_i})$ and $AB \subseteq \mathrm{ann}_R(\frac{N_j}{L\cap N_j})$. Since $\mathrm{ann}_R(\frac{N_i}{L\cap N_i})$ and $\mathrm{ann}_R(\frac{N_j}{L\cap N_j})$ are prime, by $(*)$, we have $A \subseteq \mathrm{ann}_R(\frac{N_j}{L\cap N_j})$ and $B \subseteq \mathrm{ann}_R(\frac{N_i}{L\cap N_i})$. Now by $(**)$, $B \subseteq \mathrm{ann}_R(\frac{N_j}{L\cap N_j})$, a contradiction. □

Since every submodule of a comultiplication module is comultiplication and every quotient module of a multiplication module is also multiplication, the following result is immediate.

3.2. Corollary. *Let R be a ring and M be a nonzero right R-module. Then:*
(1) If M is multiplication and N is a weakly prime submodule of M, then N is prime;
(2) If M is comultiplication and N is a weakly second submodule of M, then N is second.

The following result shows that for an R-module M, if R has the ascending chain condition (briefly, ACC) on two-sided ideals, then there exists a factor module of M such that to be second.

3.3. Proposition. *Let R be a ring and M be a nonzero right R-module. If $\mathrm{ann}_R(M/N_0)$ is a maximal member in the family $\{\mathrm{ann}_R(M/N)\}$ where N ranges over all proper submodules of M, then M/N_0 is a second R-module.*

Proof. For any proper submodule K/N_0 of M/N_0, $\mathrm{ann}_R(M/N_0) \subseteq \mathrm{ann}_R(M/K) = \mathrm{ann}_R(\frac{M/N_0}{K/N_0})$ and the maximality of $\mathrm{ann}_R(M/N_0)$, implies that $\mathrm{ann}_R(M/N_0) = \mathrm{ann}_R(\frac{M/N_0}{K/N_0})$. Thus M/N_0 is second R-module. □

3.4. Proposition. *Let R be a ring and M be a nonzero right R-module. If there exists a proper submodule N of M such that M/N is a weakly second module, then $\mathrm{ann}_R(M/N)$ is a maximal member in the collection of ideals I of R such that $MIJ + N \neq MJ + N$ for every ideal J of R with $J \nsubseteq \mathrm{ann}_R(M/N)$.*

Proof. Let $P = \mathrm{ann}_R(M/N)$ and J be an ideal of R such that $J \nsubseteq P$. Then clearly $MPJ + N \neq MJ + N$. Suppose that A is an ideal of R such that $P \subset A$. Then $A \nsubseteq P$ and since M/N is weakly second, $(M/N)A^2 = (M/N)A$. Thus $MA^2 + N = MA + N$ and it follows that P is a maximal member in the stated collection. □

It was seen that a right R-module M is weakly second if and only if $MIJ = MI$ or $MIJ = MJ$ for every two ideals I, J of R. Here, we improve this fact to modules over Noetherian rings.

3.5. Theorem. *Let R be a ring whose two-sided ideals satisfy ACC. Then a nonzero R-module M is weakly second if and only if for every two prime ideals P and Q of R, $MPQ = MP$ or $MPQ = MQ$.*

Proof. For one direction, the proof is clear. For the other direction assume that for every two prime ideals P and Q of R, $MPQ = MP$ or $MPQ = MQ$. First we show that for prime ideals $P_1, \cdots, P_n$ of R, $MP_1\cdots P_n = MP_j$ for some $1 \leq j \leq n$. We proceed by induction on n, the case $n = 2$ being covered by hypothesis. For $n = 3$, $MP_1P_2P_3 = (MP_1P_2)P_3$ and by hypothesis, $MP_1P_2 = MP_1$ or $MP_1P_2 = MP_2$. Thus we have $MP_1P_2P_3 = (MP_1P_2)P_3 = (MP_1)P_3 = MP_1P_3$ or $MP_1P_2P_3 = (MP_1P_2)P_3 =$

$(MP_2)P_3 = MP_2P_3$. This shows that we are reduced to the case $n = 2$. Again by hypothesis, $MP_1P_3 = MP_1$ or $MP_1P_3 = MP_3$ and $MP_2P_3 = MP_2$ or $MP_2P_3 = MP_3$. Therefore $MP_1P_2P_3 = MP_i$ for some $1 \leq i \leq 3$. Now by the induction hypothesis, $MP_1 \cdots P_n = (MP_1 \cdots P_{n-1})P_n = MP_iP_n$ for some $1 \leq i \leq n-1$. Also by the case $n = 2$, we have $MP_1 \cdots P_n = MP_iP_n = MP_j$ for some $1 \leq j \leq n$ and the claim is proved. Now let I and J be two ideals of R. Since R has ACC on two-sided ideals, by [14, Lemma 1], there exist two integers n, m and prime ideals $Q_i(1 \leq i \leq n)$ and $P_j(1 \leq j \leq m)$, such that

$$Q_1 \cdots Q_n \subseteq I \subseteq Q_1 \cap \cdots \cap Q_n$$

and

$$P_1 \cdots P_m \subseteq J \subseteq P_1 \cap \cdots \cap P_m.$$

Thus

$$Q_1 \cdots Q_n P_1 \cdots P_m \subseteq IJ \subseteq (Q_1 \cap \cdots \cap Q_n)(P_1 \cap \cdots \cap P_m),$$

and so

$$MQ_1 \cdots Q_n P_1 \cdots P_m \subseteq MIJ \subseteq M(Q_1 \cap \cdots \cap Q_n)(P_1 \cap \cdots \cap P_m).$$

By the first part of the proof, without loss of generality, we may assume that $MQ_1 \cdots Q_n P_1 \cdots P_m = MQ_s$ for some $1 \leq s \leq n$. Then we have $MQ_s = MQ_1 \cdots Q_n P_1 \cdots P_m \subseteq MIJ \subseteq M(Q_1 \cap \cdots \cap Q_n)(P_1 \cap \cdots \cap P_m) \subseteq MQ_s$ and so $MIJ = MQ_s$. On the other hand, $I \subseteq Q_1 \cap \cdots \cap Q_n$ implies that $MI \subseteq MQ_s$ and hence $MIJ = MI$. □

Following [1, p. 232, ex. 11], a submodule N of a right R-module M is said to be *pure* (in M) provided $N \cap MI = NI$ for every left ideal I of R.

3.6. Proposition. *Let R be a ring and M be a weakly second R-module. Then every pure submodule of M is weakly second.*

Proof. Let N be a nonzero pure submodule of M and I, J be two ideals of R. If $MIJ = MI$, then $NIJ = N \cap MIJ = N \cap MI = NI$. Similarly, $MIJ = MJ$ implies that $NIJ = NJ$. Now by Theorem 2.3, the proof is complete. □

A right R-module X is called *properly divisible* if for every proper submodule Y of X and every regular element c of R, $Y = Yc$.

3.7. Proposition. *Let R be a prime right Goldie ring. Then every nonzero properly divisible R-module is a weakly second module.*

Proof. Let X be a nonzero properly divisible R-module and I, J be two ideals of R. If $XI = X$, then $XIJ = XJ$. Thus suppose that $XI \subset X$ and let $A = \mathrm{ann}_R(XI)$. If $A \neq 0$, then it is an essential ideal of R_R (since R is prime) and by the Goldie's Theorem [13, Theorem 11.13], A contains a regular element c of R. Since X is properly divisible, $XI = XIc = 0$ and so $XIJ = XI$. Now assume that $A = 0$ and d is a regular element of J. Then $XI = XId \subseteq XIJ$ and hence $XI = XIJ$, as desired. □

A right R-module M is called a *semisecond module*, if for any ideal I of R, $MI^2 = MI$, i.e., $\mathrm{ann}_R(M/N)$ is a semiprime ideal of R, for any proper submodule N of M. We conclude the paper with the following result.

3.8. Proposition. *Let R be a ring, M be a right R-module and N be a semisecond submodule of M such that M/N is second. If for any ideal I of R, NI is a weakly prime submodule of M, then M is a semisecond module.*

Proof. Let I be an ideal of R. Since M/N is second, $(M/N)I = M/N$ or $(M/N)I = 0$. If $(M/N)I = M/N$, then $MI + N = M$ and so $MI^2 + NI = MI$. Since N is semisecond, $NI = NI^2$ and hence $MI^2 = MI$. Now suppose that $(M/N)I = 0$. Then $MI + N = N$ implies that $MI^2 + NI = MI^2 + NI^2 = NI$ and so $MI^2 = NI$, because N is semisecond. Since NI is a weakly prime submodule, $MI \subseteq NI$ and we have $MI^2 = MI$. □

References

[1] Anderson, F. W. and Fuller, K. R. *Rings and Categories of Modules,* New York: Springer-Verlag 1974.

[2] Annin, S. *Attached primes over noncommutative rings*, J. Pure Appl. Algebra **212**, 510-521, 2008.

[3] Ansari-Toroghy, H. and Farshadifar, F. *The dual notion of multiplication modules,* Taiwanese Journal of Mathematics **11**, 1189-1201, 2007.

[4] Ansari-Toroghy, H. and Farshadifar, F. *The dual notion of some generalizations of prime submodules*, Comm. Algebra **39**, 2396-2416, 2011.

[5] Azizi, A. *On prime and weakly prime submodules*, Vietnam J. Math. **36**, 315-325, 2008.

[6] Azizi, A. *Weakly prime submodules and prime submodules,* Glasgow Math. J. **48**, 343-346, 2006.

[7] Barnard, A. *Multiplication modules,* Journal of Algebra **71**, 174-178, 1981.

[8] Behboodi, M. *On weakly prime radical of modules and semi-compatible modules,* Acta Math. Hungar. **113**, 243-254, 2006.

[9] Behboodi, M. and Koohi, H. *Weakly prime modules,* Vietnam J. Math. **32**, 185-195, 2004.

[10] Ceken, S. Alkan, M. and Smith, P. F. *Second modules over noncommutative rings,* Comm. Algebra **41**, 83-98, 2013.

[11] Dauns, J. *Prime modules,* J. Reine Angew Math. **298**, 156-181, 1978.

[12] Lam, T. Y. *A first course in noncommutative rings,* New York: Springer-Verlag 1991.

[13] Lam, T. Y. *Lectures on modules and rings,* Graduate Texts in Math. New York Heidelberg-Berlin: Springer-Verlag 1999.

[14] Smith, P. F. *Injective modules and prime ideals,* Comm. Algebra **9**, 989-999, 1981.

[15] Yassemi, S. *The dual notion of prime submodules,* Arch. Math. Brno. **37**, 273-278, 2001.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1367 – 1378

Some starlikeness and convexity properties for two new $p-$valent integral operators

Erhan Deniz*, Esra Deniz and Nizami Mustafa

Abstract

In this paper, we define two new general $p-$valent integral operators in the unit disc $\mathbb{U}$ and obtain the properties of $p-$valent starlikeness and $p-$valent convexity of these integral operators of $p-$valent functions on some classes of $\beta-$uniformly $p-$valent starlike and $\beta-$uniformly $p-$valent convex functions of complex order and type α $(0 \leq \alpha < p)$. As special cases, the properties of $p-$valent starlikeness and $p-$valent convexity of the operators $\int_0^z pt^{p-1}\left(\frac{f(t)}{t^p}\right)^{\delta} dt$ and $\int_0^z pt^{p-1}\left(\frac{g' t)}{pt^{p-1}}\right)^{\delta} dt$ are given.

Keywords: Analytic functions; Integral operators; $\beta-$uniformly $p-$valent starlike and $\beta-$uniformly $p-$valent convex functions; Complex order.

2000 AMS Classification: 30C80, 30C45.

Received : 06.05.2015 *Accepted* : 19.11.2015 *Doi* : 10.15672/HJMS.20164514281

1. Introduction and Preliminaries

Let $\mathcal{A}_p$ denote the class of the form

$$(1.1) \qquad f(z) = z^p + \sum_{k=p+1}^{\infty} a_k z^k \quad (p \in \mathbb{N} = \{1, 2, ..., \}),$$

which are analytic in the open disc $\mathbb{U} = \{z \in \mathbb{C} : \ |z| < 1\}$.

A function $f \in \mathcal{S}_p^*(\gamma, \alpha)$ is $p-$valently starlike of complex order $\gamma\,(\gamma \in \mathbb{C} - \{0\})$ and type $\alpha\,(0 \leq \alpha < p)$, that is, $f \in \mathcal{S}_p^*(\gamma, \alpha)$, if it is satisfies the following condition

$$(1.2) \qquad \operatorname{Re}\left\{p + \frac{1}{\gamma}\left(\frac{zf'(z)}{f(z)} - p\right)\right\} > \alpha \quad (z \in \mathbb{U}).$$

*Department of Mathematics, Faculty of Science and Letters, Kafkas University, 36100, Kars, Turkey.
Email : `edeniz36@gmail.com (Erhan Deniz)`
Email : `edeniz@atauni.edu.tr (Esra Deniz)`
Email : `nizamimustafa@gmail.com (Nizami Mustafa)`

Furthermore, a function $f \in \mathcal{C}_p(\gamma,\alpha)$ is $p-$valently convex of complex order $\gamma\,(\gamma \in \mathbb{C}-\{0\})$ and type $\alpha\,(0 \le \alpha < p)$, that is, $f \in \mathcal{C}_p(\gamma,\alpha)$ if it satisfies the following condition;

$$\text{(1.3)} \qquad \operatorname{Re}\left\{p+\frac{1}{\gamma}\left(1+\frac{zf''(z)}{f'(z)}-p\right)\right\} > \alpha \quad (z \in \mathbb{U}).$$

In particular cases, for $p = 1$ in the classes $\mathcal{S}_p^*(\gamma,\alpha)$ and $\mathcal{C}_p(\gamma,\alpha)$, we obtain the classes $\mathcal{S}^*(\gamma,\alpha)$ and $\mathcal{C}(\gamma,\alpha)$ of starlike functions of complex order $\gamma\,(\gamma \in \mathbb{C}-\{0\})$ and type $\alpha\,(0 \le \alpha < 1)$ and convex functions of complex order $\gamma\,(\gamma \in \mathbb{C}-\{0\})$ and type $\alpha\,(0 \le \alpha < 1)$, respectively, which were introduced and studied by Frasin [15]. Also, for $\alpha = 0$ in the classes $\mathcal{S}_p^*(\gamma,\alpha)$ and $\mathcal{C}_p(\gamma,\alpha)$, we obtain the classes $\mathcal{S}_p^*(\gamma)$ and $\mathcal{C}_p(\gamma)$, which are called $p-$valently starlike of complex order $\gamma\,(\gamma \in \mathbb{C}-\{0\})$ and $p-$valently convex of complex order $\gamma\,(\gamma \in \mathbb{C}-\{0\})$, respectively. Setting $p = 1$ and $\alpha = 0$, we obtain the classess $\mathcal{S}^*(\gamma)$ and $\mathcal{C}(\gamma)$. The class $\mathcal{S}^*(\gamma)$ of starlike functions of complex order $\gamma\,(\gamma \in \mathbb{C}-\{0\})$ was defined by Nasr and Aouf (see [21]) while the class $\mathcal{C}(\gamma)$ of convex functions of complex order $\gamma\,(\gamma \in \mathbb{C}-\{0\})$ was considered earlier by Wiatrowski (see [27]). Note that $\mathcal{S}_p^*(1,\alpha) = \mathcal{S}_p^*(\alpha)$ and $\mathcal{C}_p(1,\alpha) = \mathcal{C}_p(\alpha)$ are, respectively, the classes of $p-$valently starlike and $p-$valently convex functions of order $\alpha\,(0 \le \alpha < p)$ in $\mathbb{U}$. In special cases, $\mathcal{S}_p^*(0) = \mathcal{S}_p^*$ and $\mathcal{C}_p(0) = \mathcal{C}_p$ are, respectively, the familiar classes of $p-$valently starlike and $p-$valently convex functions in $\mathbb{U}$. Also, we note that $\mathcal{S}_1^*(\alpha) = \mathcal{S}^*(\alpha)$ and $\mathcal{C}_1(\alpha) = \mathcal{C}(\alpha)$ are, respectively, the usual classes of starlike and convex functions of order $\alpha\,(0 \le \alpha < 1)$ in $\mathbb{U}$. In special cases, $\mathcal{S}_1^*(0) = \mathcal{S}^*$ and $\mathcal{C}_1 = \mathcal{C}$ are, respectively, the familiar classes of starlike and convex functions in $\mathbb{U}$.

A function $f \in \beta-\mathcal{US}_p(\alpha)$ is $\beta-$uniformly $p-$valently starlike of order α $(0 \le \alpha < p)$, that is, $f \in \beta-\mathcal{US}_p(\alpha)$ if it is satisfies the following condition

$$\text{(1.4)} \qquad \operatorname{Re}\left\{\frac{zf'(z)}{f(z)}\right\} > \beta\left|\frac{zf'(z)}{f(z)}-p\right|+\alpha \quad (\beta \ge 0,\ z \in \mathbb{U}).$$

Furthermore, a function $f \in \beta-\mathcal{UC}_p(\alpha)$ is $\beta-$uniformly $p-$valently convex of order $\alpha\,(0 \le \alpha < p)$, that is, $f \in \beta-\mathcal{UC}_p(\alpha)$ if it satisfies the following condition

$$\text{(1.5)} \qquad \operatorname{Re}\left\{1+\frac{zf''(z)}{f'(z)}\right\} > \beta\left|1+\frac{zf''(z)}{f'(z)}-p\right|+\alpha \quad (\beta \ge 0,\ z \in \mathbb{U}).$$

These classes generalize various other classes which are worthy to mention here. For example $p = 1$, the classes $\beta-\mathcal{US}(\alpha)$ and $\beta-\mathcal{UC}(\alpha)$ introduced by Bharti, Parvatham and Swaminathan (see [2]). Also, the class $\beta-\mathcal{UC}_1(0) = \beta-\mathcal{UCV}$ is the known class of $\beta-$uniformly convex functions [17]. Using the Alexander type relation, we can obtain the class $\beta-\mathcal{US}_p(\alpha)$ in the following way:

$$f \in \beta-\mathcal{UC}_p(\alpha) \Leftrightarrow \frac{zf'}{p} \in \beta-\mathcal{US}_p(\alpha).$$

The class $1-\mathcal{UC}_1(0) = \mathcal{UCV}$ of uniformly convex functions was defined by Goodman [16] while the class $1-\mathcal{US}_1(0) = \mathcal{SP}$ was considered by Rønning [26].

When the classes $\mathcal{S}_p^*(\gamma,\alpha)$ with $\beta-\mathcal{US}_p(\alpha)$ and $\mathcal{C}_p(\gamma,\alpha)$ with $\beta-\mathcal{UC}_p(\alpha)$ are thought together, we define following classes. Let $0 \le \alpha < p$, $\beta \ge 0$ and $\gamma \in \mathbb{C}-\{0\}$. A function $f \in \mathcal{A}_p$ is in the class $\beta-\mathcal{US}_p(\gamma,\alpha)$ if and only if for all $z \in \mathbb{U}$

$$\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{zf'(z)}{f(z)}-p\right)\right\} > \beta\left|\frac{1}{\gamma}\left(\frac{zf'(z)}{f(z)}-p\right)\right|+\alpha$$

and in the class $\beta-\mathcal{UC}_p(\gamma,\alpha)$ if and only if for all $z \in \mathbb{U}$

$$\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{zf''(z)}{f'(z)}+1-p\right)\right\} > \beta\left|\frac{1}{\gamma}\left(\frac{zf''(z)}{f'(z)}+1-p\right)\right|+\alpha.$$

For $f \in \mathcal{A}_p$ given by (1.1) and $g(z)$ given by

$$g(z) = z^p + \sum_{k=p+1}^{\infty} b_k z^k \tag{1.6}$$

their convolution (or Hadamard product), denoted by $(f * g)$, is defined as follows

$$(f * g)(z) = z^p + \sum_{k=p+1}^{\infty} a_k b_k z^k = (g * f)(z) \quad (z \in \mathbb{U}).$$

For a function f in $\mathcal{A}_p$, in [13], the authors defined the *multiplier transformations* $\mathcal{D}^m_{p,\lambda,\mu}$ as follows.

1.1. Definition. Let $f \in \mathcal{A}_p$. For the parameters $\lambda, \mu \in \mathbb{R}$; $0 \le \mu \le \lambda$ and $m \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, define the multiplier transformations $\mathcal{D}^m_{p,\lambda,\mu}$ on $\mathcal{A}_p$ by the following:

$$\begin{aligned}
\mathcal{D}^0_{p,\lambda,\mu} f(z) &= f(z) \\
\mathcal{D}^1_{p,\lambda,\mu} f(z) &= \mathcal{D}_{p,\lambda,\mu} f(z) \\
&= \frac{1}{p}\left[\lambda\mu z^2 f''(z) + (\lambda - \mu + (1-p)\lambda\mu)\, z f'(z) + p(1 - \lambda + \mu) f(z)\right] \\
&\ \ \vdots \\
\mathcal{D}^m_{p,\lambda,\mu} f(z) &= \mathcal{D}_{p,\lambda,\mu}\left(\mathcal{D}^{m-1}_{p,\lambda,\mu}\right)
\end{aligned}$$

for $z \in \mathbb{U}$ and $p \in \mathbb{N} := \{1, 2, ...\}$.

If $f(z)$ is given by (1.1), then from the definition of the multiplier transformations $\mathcal{D}^m_{p,\lambda,\mu} f(z)$, we can easily see that

$$\mathcal{D}^m_{p,\lambda,\mu} f(z) = z^p + \sum_{k=p+1}^{\infty} \Phi^k_p(m, \lambda, \mu) a_k z^k$$

where

$$\Phi^k_p(m, \lambda, \mu) = \left[\frac{(k-p)(\lambda\mu k + \lambda - \mu) + p}{p}\right]^m .$$

By using the operator $\mathcal{D}^m_{p,\lambda,\mu} f(z)$ $(m \in \mathbb{N}_0)$, we introduce the new classes $\beta - \mathcal{US}_p(m, \lambda, \mu, \gamma, \alpha)$ and $\beta - \mathcal{UC}_p(m, \lambda, \mu, \gamma, \alpha)$ as follows:

$$\beta - \mathcal{US}_p(m, \lambda, \mu, \gamma, \alpha) = \left\{f \in \mathcal{A}_p : \ \mathcal{D}^m_{p,\lambda,\mu} f(z) \in \beta - \mathcal{US}_p(\gamma, \alpha)\right\}$$

and

$$\beta - \mathcal{UC}_p(m, \lambda, \mu, \gamma, \alpha) = \left\{f \in \mathcal{A}_p : \ \mathcal{D}^m_{p,\lambda,\mu} f(z) \in \beta - \mathcal{UC}_p(\gamma, \alpha)\right\}$$

where $f \in \mathcal{A}_p$, $0 \le \alpha < p$, $\beta \ge 0$ and $\gamma \in \mathbb{C} - \{0\}$.

We note that by specializing the parameters m, p, γ, β and α in the classes $\beta - \mathcal{US}_p(m, \lambda, \mu, \gamma, \alpha)$ and $\beta - \mathcal{UC}_p(m, \lambda, \mu, \gamma, \alpha)$, these classes are reduced to several well-known subclasses of analytic functions. For example, for $m = 0$ the classes $\beta - \mathcal{US}_p(m, \lambda, \mu, \gamma, \alpha)$ and $\beta - \mathcal{UC}_p(m, \lambda, \mu, \gamma, \alpha)$ are reduced to the classes $\beta - \mathcal{US}_p(\gamma, \alpha)$ and $\beta - \mathcal{UC}_p(\gamma, \alpha)$, respectively. Someone can find more information about these classes in Cağlar [10], Deniz, Orhan and Sokol [11], Deniz, Cağlar and Orhan [12] and Orhan, Deniz and Raducanu [22].

1.2. Definition. Let $l = (l_1, l_2, ..., l_n) \in \mathbb{N}_0^n$, $\delta = (\delta_1, \delta_2, ..., \delta_n) \in \mathbb{R}_+^n$ for all $i = \overline{1, n}$, $n \in \mathbb{N}$. We define the following general integral operators

$$\mathcal{I}^{\delta,\lambda,\mu}_{n,p,l}(f_1, f_2, ..., f_n) : \mathcal{A}_p^n \to \mathcal{A}_p$$

$$\mathcal{I}^{\delta,\lambda,\mu}_{n,p,l}(f_1, f_2, ..., f_n) = \mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z),$$

$$\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z) = \int_0^z pt^{p-1} \prod_{i=1}^{n} \left(\frac{\mathcal{D}^{l_i}_{p,\lambda,\mu} f_i(t)}{t^p} \right)^{\delta_i} dt \tag{1.7}$$

and

$$\mathcal{J}^{\delta,\lambda,\mu}_{n,p,l}(g_1, g_2, ..., g_n) : \mathcal{A}^n_p \to \mathcal{A}_p$$

$$\mathcal{J}^{\delta,\lambda,\mu}_{n,p,l}(g_1, g_2, ..., g_n) = \mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z),$$

$$\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z) = \int_0^z pt^{p-1} \prod_{i=1}^{n} \left(\frac{\left(\mathcal{D}^{l_i}_{p,\lambda,\mu} g_i(t)\right)'}{pt^{p-1}} \right)^{\delta_i} dt \tag{1.8}$$

where $f_i, g_i \in \mathcal{A}_p$ for all $i = \overline{1,n}$ and $\mathcal{D}^{l}_{p,\lambda,\mu}$ is defined in Definition 1.1.

1.3. Remark. We note that if $l_1 = l_2 = ... = l_n = 0$, then the integral operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ is reduced to the operator $F_p(z)$ which was studied by Frasin (see [14]). Upon setting $p = 1$ in the operator (1.7), we can obtain the integral operator $\mathbb{F}_n(z)$ which was studied by Oros G.I. and Oros G.A. (see [23]). For $p = 1$ and $l_1 = l_2 = ... = l_n = 0$ in (1.7), the integral operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ is reduced to the operator $F_m(z)$ which was studied by Breaz D. and Breaz N. (see [6]). Observe that when $p = n = 1$, $l_1 = 0$ and $\delta_1 = \delta$, we obtain the integral operator $I_\delta(f)(z)$ which was studied by Pescar and Owa (see [24]), for $\delta_1 = \delta \in [0, 1]$ special case of the operator $I_\delta(f)(z)$ was studied by Miller, Mocanu and Reade (see [19]). For $p = n = 1$, $l_1 = 0$ and $\delta_1 = 1$ in (1.7), we have Alexander integral operator $I(f)(z)$ in [1].

1.4. Remark. For $l_1 = l_2 = ... = l_n = 0$ in (1.8) the integral operator $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)$ is reduced to the operator $G_p(z)$ which was studied by Frasin (see [14]). For $p = 1$ and $l_1 = l_2 = ... = l_n = 0$ in (1.8), the integral operator $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)$ is reduced to the operator $G_{\delta_1,\delta_2,...,\delta_m}(z)$ which was studied by Breaz D., Owa and Breaz N. (see [8]). If $p = n = 1$, $l_1 = 0$ and $\delta_1 = \delta$, we obtain the integral operator $G(z)$ which was introduced and studied by Pfaltzgraff (see [25]) and Kim and Merkes (see [18]).

In this paper, we consider the integral operators $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ and $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)$ defined by (1.7) and (1.8), respectively, and study their properties on the classes $\beta - \mathcal{US}_p(m, \lambda, \mu, \gamma, \alpha)$ and $\beta - \mathcal{UC}_p(m, \lambda, \mu, \gamma, \alpha)$. As special cases, the order of $p-$valently convexity and $p-$valently starlikeness of the operators $\int_0^z pt^{p-1} \left(\frac{f(t)}{t^p}\right)^\delta dt$ and $\int_0^z pt^{p-1} \left(\frac{g'(t)}{pt^{p-1}}\right)^\delta dt$ are given.

2. Convexity of the integral operators $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ and $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)$

First, in this section we prove a sufficient condition for the integral operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ to be $p-$valently convex of complex order.

2.1. Theorem. *Let $l = (l_1, l_2, ...l_n) \in \mathbb{N}^n_0$, $\delta = (\delta_1, \delta_2, ..., \delta_n) \in \mathbb{R}^n_+$, $0 \le \alpha_i < p$, $\gamma \in \mathbb{C} - \{0\}$ such that $0 < \sum_{i=1}^n \delta_i (p - \alpha_i) \le p$, $\beta_i \ge 0$ and $f_i \in \beta_i - \mathcal{US}_p(l_i, \lambda, \mu, \gamma, \alpha_i)$ for all $i = \overline{1,n}$. Then, the integral operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}$ defined by (1.7) is $p-$valently convex of complex order γ $(\gamma \in \mathbb{C} - \{0\})$ and type $p - \sum_{i=1}^n \delta_i (p - \alpha_i)$, that is, $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l} \in \mathcal{C}_p(\gamma, p - \sum_{i=1}^n \delta_i (p - \alpha_i))$.*

Proof. From the definition (1.7), we observe that $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z) \in \mathcal{A}_p$. On the other hand, it is easy to see that

$$\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]' = pz^{p-1}\prod_{i=1}^{n}\left(\frac{\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i(z)}{z^p}\right)^{\delta_i}. \tag{2.1}$$

Now we differentiate (2.1) logarithmically and we easily obtain

$$p+\frac{1}{\gamma}\left(\frac{z\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]''}{\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]'}+1-p\right) = p+\sum_{i=1}^{n}\delta_i\left(p+\frac{1}{\gamma}\left(\frac{z\left(\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\right)'(z)}{\left(\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\right)(z)}-p\right)\right)-p\sum_{i=1}^{n}\delta_i. \tag{2.2}$$

Then, we calculate the real part of both sides of (2.2) and obtain

$$\begin{aligned}&\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{z\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]''}{\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]'}+1-p\right)\right\}\\ &= \sum_{i=1}^{n}\delta_i\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{z\left(\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\right)'(z)}{\left(\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\right)(z)}-p\right)\right\}-p\sum_{i=1}^{n}\delta_i+p.\end{aligned} \tag{2.3}$$

Since $f_i \in \beta_i - \mathcal{US}_p(l_i,\lambda,\mu,\gamma,\alpha_i))$ for all $i=\overline{1,n}$ from (2.3), we have

$$\begin{aligned}&\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{z\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]''}{\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]'}+1-p\right)\right\}\\ &> \sum_{i=1}^{n}\frac{\delta_i\beta_i}{|\gamma|}\left|\frac{z\left(\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\right)'(z)}{\left(\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\right)(z)}-p\right|+p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right).\end{aligned} \tag{2.4}$$

Because $\sum_{i=1}^{n}\frac{\delta_i\beta_i}{|\gamma|}\left|\frac{z\left(\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\right)'(z)}{\left(\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\right)(z)}-p\right|>0$, from (2.4), we obtain

$$\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{z\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]''}{\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]'}+1-p\right)\right\} > p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right).$$

Therefore, the operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ is $p-$valently convex of complex order $\gamma\,(\gamma\in\mathbb{C}-\{0\})$ and type $p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)$. The proof of Theorem 2.1 is completed. ∎

2.2. Remark.

(1) Letting $\gamma=1$ and $l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.1, we obtain Theorem 2.1 in [14].
(2) Letting $p=1$, $\gamma=1$ and $l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.1, we obtain Theorem 1 in [4].
(3) Letting $p=1$, $\gamma=1$ and $\alpha_i=l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.1, we obtain Theorem 2.5 in [7].
(4) Letting $p=1$, $\beta=0$ and $l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.1, we obtain Theorem 1 in [3].
(5) Letting $p=1$, $\beta=0$, $\alpha_i=\alpha$ and $l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.1, we obtain Theorem 1 in [9].

(6) Letting $p = 1$, $\beta = 0$, $\alpha_i = 0$ and $l_i = 0$ for all $i = \overline{1,n}$ in Theorem 2.1, we obtain Theorem 1 in [5].

Putting $n = 1$, $l_1 = 0$, $\delta_1 = \delta$, $\alpha_1 = \alpha$, $\beta_1 = \beta$ and $f_1 = f$ in Theorem 2.1, we have

2.3. Corollary. *Let $\delta > 0$, $0 \le \alpha < p$, $\beta \ge 0$, $\gamma \in \mathbb{C} - \{0\}$ and $f \in \beta - \mathcal{US}_p(\gamma, \alpha)$. If $\delta \in (0, p \diagup (p-\alpha)]$, then $\int_0^z pt^{p-1}\left(\frac{f(t)}{t^p}\right)^{\delta} dt$ is convex of complex order $\gamma\,(\gamma \in \mathbb{C} - \{0\})$ and type $p - \delta\,(p-\alpha)$ in $\mathbb{U}$.*

2.4. Theorem. *Let $l = (l_1, l_2, ..., l_n) \in \mathbb{N}_0^n$, $\delta = (\delta_1, \delta_2, ..., \delta_n) \in \mathbb{R}_+^n$, $0 \le \alpha_i < p$, $\beta_i \ge 0$, $\gamma \in \mathbb{C} - \{0\}$ and $f_i \in \beta_i - \mathcal{US}_p(l_i, \lambda, \mu, \gamma, \alpha_i)$ for all $i = \overline{1,n}$. If*

$$\left|\frac{z\left(\mathcal{D}^{l_i}_{p,\lambda,\mu} f_i\right)'(z)}{\left(\mathcal{D}^{l_i}_{p,\lambda,\mu} f_i\right)(z)} - p\right| > -\frac{p + \sum_{i=1}^n \delta_i\,(\alpha_i - p)}{\sum_{i=1}^n \frac{\delta_i \beta_i}{|\gamma|}} \tag{2.5}$$

for all $i = \overline{1,n}$, then the integral operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ defined by (1.7) is $p-$valently convex of complex order γ $(\gamma \in \mathbb{C} - \{0\})$.

Proof. From (2.4) and (2.5), we easily get $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ is $p-$valently convex of complex order γ. ∎

From Theorem 2.4, we easily get

2.5. Corollary. *Let $l = (l_1, l_2, ..., l_n) \in \mathbb{N}_0^n$, $\delta = (\delta_1, \delta_2, ..., \delta_n) \in \mathbb{R}_+^n$, $0 \le \alpha_i < p$, $\beta_i \ge 0$, $\gamma \in \mathbb{C} - \{0\}$ and $f_i \in \beta_i - \mathcal{US}_p(l_i, \lambda, \mu, \gamma, \alpha_i)$ for all $i = \overline{1,n}$. If $\mathcal{D}^{l_i}_{p,\lambda,\mu} f_i \in \mathcal{S}^*_p(\sigma)$, where $\sigma = p - \left(p - \sum_{i=1}^n \delta_i\,(p - \alpha_i)\right) \diagup \sum_{i=1}^n \frac{\delta_i \beta_i}{|\gamma|}$; $0 \le \sigma < p$ for all $i = \overline{1,n}$, then the integral operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ is $p-$valently convex of complex order $\gamma\,(\gamma \in \mathbb{C} - \{0\})$.*

Putting $n = 1$, $l_1 = 0$, $\delta_1 = \delta$, $\alpha_1 = \alpha$, $\beta_1 = \beta$ and $f_1 = f$ in Corollary 2.5, we have

2.6. Corollary. *Let $\delta > 0$, $0 \le \alpha < p$, $\beta > 0$, $\gamma \in \mathbb{C} - \{0\}$ and $f \in \mathcal{S}^*_p(\rho)$ where $\rho = [\delta(p\beta + (p-\alpha)\,|\gamma|) - p\,|\gamma|] \diagup \delta\beta$; $0 \le \rho < p$, then the integral operator $\int_0^z pt^{p-1}\left(\frac{f(t)}{t^p}\right)^{\delta} dt$ is $p-$valently convex of complex order $\gamma\,(\gamma \in \mathbb{C} - \{0\})$ in $\mathbb{U}$.*

Next, we give a sufficient condition for the integral operator $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)$ to be $p-$valently convex of complex order.

2.7. Theorem. *Let $l = (l_1, l_2, ...l_n) \in \mathbb{N}_0^n$, $\delta = (\delta_1, \delta_2, ..., \delta_n) \in \mathbb{R}_+^n$, $0 \le \alpha_i < p$, $\gamma \in \mathbb{C} - \{0\}$ such that $0 < \sum_{i=1}^n \delta_i\,(p - \alpha_i) \le p$, $\beta_i \ge 0$ and $g_i \in \beta_i - \mathcal{UC}_p(l_i, \lambda, \mu, \gamma, \alpha_i)$ for all $i = \overline{1,n}$. Then, the integral operator $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}$ defined by (1.8) is $p-$valently convex of complex order $\gamma\,(\gamma \in \mathbb{C} - \{0\})$ and type $p - \sum_{i=1}^n \delta_i\,(p - \alpha_i)$, that is, $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l} \in \mathcal{C}_p(\gamma, p - \sum_{i=1}^n \delta_i\,(p - \alpha_i))$.*

Proof. From the definition (1.8), we observe that $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z) \in \mathcal{A}_p$. On the other hand, it is easy to see that

$$\left[\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]' = pz^{p-1} \prod_{i=1}^n \left(\frac{\left(\mathcal{D}^{l_i}_{p,\lambda,\mu} g_i(z)\right)'}{pz^{p-1}}\right)^{\delta_i}. \tag{2.6}$$

Now, we differentiate (2.6) logarithmically and then do some simple calculations, we have

$$\begin{aligned}(2.7)\qquad &\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{z\left[\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]''}{\left[\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]'}+1-p\right)\right\}\\ &=\sum_{i=1}^{n}\delta_i\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(1+\frac{z\left(\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\right)''(z)}{\left(\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\right)'(z)}-p\right)\right\}-p\sum_{i=1}^{n}\delta_i+p.\end{aligned}$$

Since $g_i\in\beta_i-\mathcal{UC}_p(l_i,\lambda,\mu,\gamma,\alpha_i)$ for all $i=\overline{1,n}$ from (2.7), we have

$$\begin{aligned}(2.8)\qquad &\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{z\left[\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]''}{\left[\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]'}+1-p\right)\right\}\\ &> p-p\sum_{i=1}^{n}\delta_i+\sum_{i=1}^{n}\delta_i\left\{\beta_i\left|\frac{1}{\gamma}\left(\frac{z\left(\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\right)''(z)}{\left(\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\right)'(z)}+1-p\right)\right|+\alpha_i\right\}\\ &= p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)+\sum_{i=1}^{n}\frac{\delta_i\beta_i}{|\gamma|}\left|\frac{z\left(\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\right)''(z)}{\left(\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\right)'(z)}+1-p\right|\\ &> p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right).\end{aligned}$$

Therefore, the operator $\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)$ is $p-$valently convex of complex order $\gamma\,(\gamma\in\mathbb{C}-\{0\})$ and type $p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)$. This evidently completes the proof of Theorem 2.7. ∎

2.8. Remark.

(1) Letting $\gamma=1$ and $l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.7, we obtain Theorem 3.1 in [14].
(2) Letting $p=1$, $\beta=0$ and $l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.7, we obtain Theorem 3 in [3].
(3) Letting $p=1$, $\beta=0$, $\alpha_i=\mu$ and $l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.7, we obtain Theorem 3 in [9].
(4) Letting $p=1$, $\beta=0$, $\alpha_i=0$ and $l_i=0$ for all $i=\overline{1,n}$ in Theorem 2.7, we obtain Theorem 2 in [5].

Putting $n=1$, $l_1=0$, $\delta_1=\delta$, $\alpha_1=\alpha$, $\beta_1=\beta$ and $g_1=g$ in Theorem 2.7, we have

2.9. Corollary. *Let $\delta>0$, $0\leq\alpha<p$, $\beta\geq0$, $\gamma\in\mathbb{C}-\{0\}$ and $g\in\beta-\mathcal{UC}_p(\gamma,\alpha)$. If $\delta\in(0,p/(p-\alpha)]$, then $\int_0^z pt^{p-1}\left(\frac{g't)}{pt^{p-1}}\right)^{\delta}dt$ is $p-$valently convex of complex order $\gamma\,(\gamma\in\mathbb{C}-\{0\})$ and type $p-\delta\,(p-\alpha)$ in $\mathbb{U}$.*

2.10. Theorem. *Let $l=(l_1,l_2,...,l_n)\in\mathbb{N}_0^n$, $\delta=(\delta_1,\delta_2,...,\delta_n)\in\mathbb{R}_+^n$, $0\leq\alpha_i<p$, $\beta_i\geq0$, $\gamma\in\mathbb{C}-\{0\}$ and $g_i\in\beta_i-\mathcal{UC}_p(l_i,\lambda,\mu,\gamma,\alpha_i)$ for all $i=\overline{1,n}$. If*

$$(2.9)\qquad \left|\frac{z\left(\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\right)''(z)}{\left(\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\right)'(z)}+1-p\right|>-\frac{p+\sum_{i=1}^{n}\delta_i\left(\alpha_i-p\right)}{\sum_{i=1}^{n}\frac{\delta_i\beta_i}{|\gamma|}}$$

for all $i=\overline{1,n}$, then the integral operator $\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)$ defined by (1.8) is $p-$valently convex of complex order $\gamma\,(\gamma\in\mathbb{C}-\{0\})$.

Proof. From (2.8) and (2.9), we easily get $\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)$ is $p-$valently convex of complex order γ. ∎

From Theorem 2.10, we easily get

2.11. Corollary. *Let* $l=(l_1,l_2,...,l_n)\in\mathbb{N}_0^n$, $\delta=(\delta_1,\delta_2,...,\delta_n)\in\mathbb{R}_+^n$, $0\le\alpha_i<p$, $\beta_i\ge 0$, $\gamma\in\mathbb{C}-\{0\}$ *and* $g_i\in\beta_i-\mathcal{UC}_p(l_i,\lambda,\mu,\gamma,\alpha_i)$ *for all* $i=\overline{1,n}$. *If* $\mathcal{D}_{p,\lambda,\mu}^{l_i}g_i\in\mathcal{C}_p(\sigma)$, *where* $\sigma=p-\left(p-\sum_{i=1}^n\delta_i\,(p-\alpha_i)\right)\diagup\sum_{i=1}^n\frac{\delta_i\beta_i}{|\gamma|}$; $0\le\sigma<p$ *for all* $i=\overline{1,n}$, *then the integral operator* $\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)$ *is* $p-$*valently convex of complex order* $\gamma\,(\gamma\in\mathbb{C}-\{0\})$.

Putting $n=1$, $l_1=0$, $\delta_1=\delta$, $\alpha_1=\alpha$, $\beta_1=\beta$ and $g_1=g$ in Corollary 2.11, we have

2.12. Corollary. *Let* $\delta>0$, $0\le\alpha<p$, $\beta>0$, $\gamma\in\mathbb{C}-\{0\}$ *and* $g\in\mathcal{C}(\rho)$ *where* $\rho=[\delta(p\beta+(p-\alpha)\,|\gamma|)-p\,|\gamma|]\diagup\delta\beta$; $0\le\rho<p$, *then the integral operator* $\int_0^z pt^{p-1}\left(\frac{g't)}{pt^{p-1}}\right)^\delta dt$ *is convex of complex order* $\gamma\,(\gamma\in\mathbb{C}-\{0\})$ *in* $\mathbb{U}$.

3. Starlikeness of the integral operators $\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)$ and $\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)$

In this section, we will give the sufficient conditions for the integral operators $\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}$ and $\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)$ to be $p-$valently starlike of complex order.
Let

$$H(\mathbb{U})=\{f:\mathbb{U}\to\mathbb{C}:\ f\text{ analytic}\}$$

$$H\,[a,n]=\left\{f\in H(\mathbb{U}):\ f(z)=a+a_nz^n+a_{n+1}z^{n+1}+....,\ \ z\in\mathbb{U},\ a\in\mathbb{C},\ n\in\mathbb{N}_0\right\}.$$

In order to prove our main results, we shall need the following lemma due to S. S. Miller and P. T. Mocanu [20].

3.1. Lemma. *Let the function* $\psi:\mathbb{C}^2\times\mathbb{U}\to\mathbb{U}$ *satisfy*

$$\text{Re }\psi(i\rho,\sigma;z)\le 0$$

for all $\rho,\sigma\in\mathbb{R}$, $n\ge 1$ *with* $\sigma\le-\frac{n}{2}(1+\rho^2)$. *If* $P\in H[1,n]$ *and* $\text{Re }\psi(P(z),zP'(z);z)>0$ *for every* $z\in\mathbb{U}$, *then*

$$\text{Re }P(z)>0.$$

3.2. Lemma. *Let* $n\in\mathbb{N}$, $\kappa\in\mathbb{R}$, $u,v\in\mathbb{C}$ *such that* $\text{Im }v\le 0$, $\text{Re}\,(u-\kappa v)\ge 0$. *Assume the following condition*

$$\text{Re}\left\{P(z)+\frac{zP'(z)}{u-vP(z)}\right\}>\kappa,\quad (z\in\mathbb{U})$$

is satisfy such that $P\in H\,[P(0),n]$, $P(0)\in\mathbb{R}$ *and* $P(0)>\kappa$. *Then,*

$$\text{Re }P(z)>\kappa,\quad (z\in\mathbb{U})\ .$$

Proof. Firstly, we consider the function $R:\mathbb{U}\to\mathbb{C}$,

$$R(z)=\frac{P(z)-\kappa}{P(0)-\kappa}\,.$$

Then, $R(z)\in H\,[1,n]$. Furthermore, since $P(0)-\kappa>0$ and

$$\text{Re}\left\{P(z)+\frac{zP'(z)}{u-vP(z)}\right\}>\kappa,\quad (z\in\mathbb{U})\,,$$

we have

$$\text{Re}\left\{R(z)+\frac{zR'(z)}{u-v\kappa-v\,(P(0)-\kappa)\,R(z)}\right\}>0,\quad (z\in\mathbb{U})\,.$$

Now, we define the function ψ as follows

$$\psi(R(z), zR'(z); z) = R(z) + \frac{zR'(z)}{u - v\kappa - v\,(P(0)-\kappa)\,R(z)}.$$

Thus,

$$\operatorname{Re}\psi(R(z), zR'(z); z) > 0.$$

Now, so then we can use Lemma 3.1, we must show that the following condition

$$\operatorname{Re}\ \psi(i\rho, \sigma; z) \leq 0$$

is satisfied for $\rho \leq 0$, $\sigma \leq -\frac{1+\rho^2}{2}$ and $z \in \mathbb{U}$. Indeed, from hypothesis, we obtain

$$\begin{aligned}
\operatorname{Re}\ \psi(i\rho, \sigma; z) &= \operatorname{Re}\ \frac{\sigma}{u - v\kappa - v\,(P(0)-\kappa)\,\rho i} \\
&= \operatorname{Re}\ \frac{\sigma}{u_1 + iu_2 - (v_1 + iv_2)\,\kappa - (v_1 + iv_2)\,(P(0)-\kappa)\,\rho i} \\
&= \frac{\sigma\,[u_1 - v_1\kappa + v_2\rho\,(P(0)-\kappa)]}{[u_1 - v_1\kappa + v_2\rho\,(P(0)-\kappa)]^2 + [u_2 - v_2\kappa + v_1\rho\,(P(0)-\kappa)]^2} \leq 0.
\end{aligned}$$

Hence, from Lemma 3.1, we get Re $R(z) > 0$. Moreover, from the definition of $R(z)$, we obtain

$$\operatorname{Re}P(z) > \kappa, \quad (z \in \mathbb{U}).$$

∎

Now, we prove the following theorem using Lemma 3.2

3.3. Theorem. *Let* $l = (l_1, l_2, ...l_n) \in \mathbb{N}_0^n$, $\delta = (\delta_1, \delta_2, ..., \delta_n) \in \mathbb{R}_+^n$, $0 \leq \alpha_i < p$, $\gamma \in \mathbb{C} - \{0\}$ *such that* $0 < \sum_{i=1}^n \delta_i\,(p - \alpha_i) \leq p$, $\operatorname{Im}\gamma \geq 0$, $\operatorname{Re}\gamma \leq \frac{p}{\sum_{i=1}^n \delta_i (p-\alpha_i)}$, $\beta_i \geq 0$ *and* $f_i \in \beta_i - \mathcal{US}_p(l_i, \lambda, \mu, \gamma, \alpha_i)$ *for all* $i = \overline{1, n}$. *Then, the integral operator* $\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}$ *defined by (1.7) is* $p-$*valently starlike of complex order* $\gamma\,(\gamma \in \mathbb{C} - \{0\})$ *and type* $p - \sum_{i=1}^n \delta_i\,(p-\alpha_i)$, *that is,* $\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu} \in \mathcal{S}_p^*(\gamma, p - \sum_{i=1}^n \delta_i\,(p-\alpha_i))$.

Proof. We define the analytic function $q: \mathbb{U} \to \mathbb{C}$, $q(0) = p$ as follows

$$q(z) = p + \frac{1}{\gamma}\left(\frac{z\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]'}{\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]} - p\right).$$

Thus, we obtain

$$p + \gamma\,(q(z) - p) = \frac{z\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]'}{\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]}$$

$$\Rightarrow \quad \frac{\gamma z q'(z)}{p(1-\gamma) + \gamma q(z)} = 1 + \frac{z\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]''}{\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]'} - \frac{z\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]'}{\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]}$$

$$\Rightarrow \quad p + \gamma\,(q(z) - p) + \frac{\gamma z q'(z)}{p(1-\gamma) + \gamma q(z)} = 1 + \frac{z\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]''}{\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]'}$$

$$\Rightarrow \quad q(z) + \frac{zq'(z)}{p(1-b) + bq(z)} = p + \frac{1}{\gamma}\left[1 - p + \frac{z\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]''}{\left[\mathcal{F}_{n,p,l}^{\delta,\lambda,\mu}(z)\right]'}\right].$$

When we consider this last equality and the inequality (2.2), we can write

$$q(z)+\frac{zq'(z)}{p(1-\gamma)+\gamma q(z)}=p+\sum_{i=1}^{n}\delta_i\left(p+\frac{1}{\gamma}\left(\frac{z\left(D^{l_i}_{p,\lambda,\mu}f_i\right)'(z)}{D^{l_i}_{p,\lambda,\mu}f_i(z)}-p\right)\right)-p\sum_{i=1}^{n}\delta_i.$$

Similarly to the proof of Theorem 2.1, it can be easly seen that

$$\operatorname{Re}\left\{q(z)+\frac{zq'(z)}{p(1-\gamma)+\gamma q(z)}\right\}>p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right).$$

Here, $q(0)=p>p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)$ and the function q is analytic on $\mathbb{U}$. Also, when we write $\kappa=p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right),\ u=p(1-\gamma)$ and $v=-\gamma$, we find $\operatorname{Im} v\leq 0$ and $\operatorname{Re}(u-\kappa v)\geq 0$. Hence, all the conditions of Lemma 3.1 are satisfied and so

$$\operatorname{Re} q(z)=\operatorname{Re}\left\{p+\frac{1}{\gamma}\left(\frac{z\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]'}{\left[\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)\right]}-p\right)\right\}>p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right).$$

Thus, the proof of the theorem is completed. ∎

Putting $n=1$, $l_1=0$, $\delta_1=\delta$, $\alpha_1=\alpha$, $\beta_1=\beta$ and $f_1=f$ in Theorem 3.3, we have

3.4. Corollary. *Let $\delta>0$, $0\leq\alpha<p$, $\beta\geq 0$, $\gamma\in\mathbb{C}-\{0\}$, $\operatorname{Im}\gamma\geq 0$, $\operatorname{Re}\gamma\leq\frac{p}{\delta(p-\alpha)}$ and $f\in\beta-\mathcal{US}_p(\gamma,\alpha)$. If $\delta\in\left(0,\frac{p}{p-\alpha}\right]$ then $\int_0^z pt^{p-1}\left(\frac{f(t)}{t^p}\right)^{\delta}dt\in\mathcal{S}^*_p(\gamma,p-\delta\left(p-\alpha\right))$.*

From Theorem 3.3, we obtain the following result.

3.5. Theorem. *Let $l=(l_1,l_2,...l_n)\in\mathbb{N}_0^n$, $\delta=(\delta_1,\delta_2,...,\delta_n)\in\mathbb{R}_+^n$, $0\leq\alpha_i<p$, $\gamma\in\mathbb{C}-\{0\}$ such that $0<\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)\leq p$, $\operatorname{Im}\gamma\geq 0$, $\operatorname{Re}\gamma\leq\frac{p}{\sum_{i=1}^{n}\delta_i(p-\alpha_i)}$, $\beta_i\geq 0$ and $f_i\in\beta_i-\mathcal{US}_p(l_i,\lambda,\mu,\gamma,\alpha_i)$ for all $i=\overline{1,n}$. If the inequality (2.5) is satisfied for all $i=\overline{1,n}$, then the integral operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ defined by (1.7) is $p-$valently starlike of complex order $\gamma\,(\gamma\in\mathbb{C}-\{0\})$.*

From Theorem 3.5, we get the following result.

3.6. Corollary. *Let $l=(l_1,l_2,...l_n)\in\mathbb{N}_0^n$, $\delta=(\delta_1,\delta_2,...,\delta_n)\in\mathbb{R}_+^n$, $0\leq\alpha_i<p$, $\gamma\in\mathbb{C}-\{0\}$ such that $0<\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)\leq p$, $\operatorname{Im}\gamma\geq 0$, $\operatorname{Re}\gamma\leq\frac{p}{\sum_{i=1}^{n}\delta_i(p-\alpha_i)}$, $\beta_i\geq 0$ and $f_i\in\beta_i-\mathcal{US}_p(l_i,\lambda,\mu,\gamma,\alpha_i)$ for all $i=\overline{1,n}$. If $\mathcal{D}^{l_i}_{p,\lambda,\mu}f_i\in\mathcal{S}^*_p(\sigma)$, where $\sigma=p-\left(p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)\right)\diagup\sum_{i=1}^{n}\frac{\delta_i\beta_i}{|\gamma|}$; $0\leq\sigma<p$ for all $i=\overline{1,n}$, then the integral operator $\mathcal{F}^{\delta,\lambda,\mu}_{n,p,l}(z)$ is $p-$valently starlike of complex order $\gamma\,(\gamma\in\mathbb{C}-\{0\})$.*

Next, we give a sufficient condition for the integral operator $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)$ to be $p-$valently starlike of complex order.

3.7. Theorem. *Let $l=(l_1,l_2,...l_n)\in\mathbb{N}_0^n$, $\delta=(\delta_1,\delta_2,...,\delta_n)\in\mathbb{R}_+^n$, $0\leq\alpha_i<p$, $\gamma\in\mathbb{C}-\{0\}$ such that $0<\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)\leq p$, $\operatorname{Im}\gamma\geq 0$, $\operatorname{Re}\gamma\leq\frac{p}{\sum_{i=1}^{n}\delta_i(p-\alpha_i)}$, $\beta_i\geq 0$ and $f_i\in\beta_i-\mathcal{UC}_p(l_i,\lambda,\mu,\gamma,\alpha_i)$ for all $i=\overline{1,n}$. Then, the integral operator $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}$ defined by (1.8) is $p-$valently starlike of complex order $\gamma\,(\gamma\in\mathbb{C}-\{0\})$ and type $p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right)$, that is, $\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}\in\mathcal{S}^*_p(\gamma,p-\sum_{i=1}^{n}\delta_i\left(p-\alpha_i\right))$.*

Proof. Let us define the analytic function $q:\mathbb{U}\to\mathbb{C}$ given by

$$q(z)=p+\frac{1}{\gamma}\left(\frac{z\left(\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)\right)'}{\left(\mathcal{G}^{\delta,\lambda,\mu}_{n,p,l}(z)\right)}-p\right).$$

Then, we follow the same steps as in the proof of Theorem 3.3, so we omit the details involved in this case. ∎

Putting $n = 1$, $l_1 = 0$, $\delta_1 = \delta$, $\alpha_1 = \alpha$, $\beta_1 = \beta$ and $g_1 = g$ in Theorem 3.7, we have

3.8. Corollary. *Let* $\delta > 0$, $0 \leq \alpha < p$, $\beta \geq 0$, $\gamma \in \mathbb{C} - \{0\}$, $\operatorname{Im} \gamma \geq 0$, $\operatorname{Re} \gamma \leq \frac{p}{\delta(p-\alpha)}$ *and* $f \in \beta - \mathcal{UC}_p(\gamma, \alpha)$. *If* $\delta \in \left(0, \frac{p}{p-\alpha}\right]$, *then* $\int_0^z pt^{p-1} \left(\frac{g'(t)}{pt^{p-1}}\right)^\delta dt \in \mathcal{S}_p^*(\gamma, p - \delta (p - \alpha))$.

From Theorem 3.7, we obtain the following result.

3.9. Theorem. *Let* $l = (l_1, l_2, ...l_n) \in \mathbb{N}_0^n$, $\delta = (\delta_1, \delta_2, ..., \delta_n) \in \mathbb{R}_+^n$, $0 \leq \alpha_i < p$, $\gamma \in \mathbb{C} - \{0\}$ *such that* $0 < \sum_{i=1}^n \delta_i (p - \alpha_i) \leq p$, $\operatorname{Im} \gamma \geq 0$, $\operatorname{Re} \gamma \leq \frac{p}{\sum_{i=1}^n \delta_i (p-\alpha_i)}$, $\beta_i \geq 0$ *and* $f_i \in \beta_i - \mathcal{US}_p(l_i, \lambda, \mu, \gamma, \alpha_i)$ *for all* $i = \overline{1, n}$. *If the inequality (2.9) is satisfied for all* $i = \overline{1, n}$, *then the integral operator* $\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)$ *defined by (1.8) is* $p-$*valently starlike of complex order* $\gamma\, (\gamma \in \mathbb{C} - \{0\})$.

We obtain the following corollary using Theorem 3.9.

3.10. Corollary. *Let* $l = (l_1, l_2, ...l_n) \in \mathbb{N}_0^n$, $\delta = (\delta_1, \delta_2, ..., \delta_n) \in \mathbb{R}_+^n$, $0 \leq \alpha_i < p$, $\gamma \in \mathbb{C} - \{0\}$ *such that* $0 < \sum_{i=1}^n \delta_i (p - \alpha_i) \leq p$, $\operatorname{Im} \gamma \geq 0$, $\operatorname{Re} \gamma \leq \frac{p}{\sum_{i=1}^n \delta_i (p-\alpha_i)}$, $\beta_i \geq 0$ *and* $f_i \in \beta_i - \mathcal{US}_p(l_i, \lambda, \mu, \gamma, \alpha_i)$ *for all* $i = \overline{1, n}$. *If* $\mathcal{D}_{p,\lambda,\mu}^{l_i} f_i \in \mathcal{C}_p(\sigma)$, *where* $\sigma = p - \left(p - \sum_{i=1}^n \delta_i (p - \alpha_i)\right) \big/ \sum_{i=1}^n \frac{\delta_i \beta_i}{|\gamma|}$; $0 \leq \sigma < p$ *for all* $i = \overline{1, n}$, *then the integral operator* $\mathcal{G}_{n,p,l}^{\delta,\lambda,\mu}(z)$ *is* $p-$*valently starlike of complex order* $\gamma\, (\gamma \in \mathbb{C} - \{0\})$.

References

[1] Alexander, J. W. *Functions which map the interior of the unit circle upon simple regions,* Annals of Mathematics. **17** (1), 12–22, 1915.

[2] Bharti, R., Parvatham, R. and Swaminathan, A. *On subclasses of uniformly convex functions and corresponding class of starlike functions*, Tamkang J. Math. **28** (1), 17-32, 1997.

[3] Bulut, S. *A note on the paper of Breaz and Guney*, J. Math. Ineq. **2**(4), 549-553, 2008.

[4] Breaz, D. *A convexity properties for an integral operator on the classes* $S_p(\alpha)$, Gen. Math. **15** (2-3), 177-183, 2007.

[5] Breaz, D., Aouf, M. K. and Breaz, N. *Some properties for integral operators on some analytic functions with complex order*, Acta. Math. Acad. Paedagog. Nyhazi. **25**, 39-43, 2009.

[6] Breaz, D. and Breaz, N. *Two integral operators*, Stud. Univ. Babes -Bolyai Math. **47** (3), 13-19, 2002.

[7] Breaz, D. and Breaz, N. *Some convexity properties for a general integral operator,* J. Ineq. Pure Appl. Math. **7** (5), Art. 177, 2006.

[8] Breaz, D., Owa, S. and Breaz, N. *A new integral univalent operator*, Acta Univ. Apulensis Math. Inform. **16,** 11-16, 2008.

[9] Breaz, D. and Guney, H. C. *The integral operator on the classes* $S_\alpha^*(b)$ *and* $C_\alpha(b)$, J. Math. Ineq. **2** (1), 97-100, 2008.

[10] Cağlar, M. *The relationships between* $p-$*valent functions and univalent functions,* An. St. Univ. Ovidius Constanta. **23** (1), 65-72, 2015.

[11] Deniz, E., Orhan, H. and Sokol J. *Classes of analytic functions defined by a differential operator related to conic domains*. Ukrainan Mat. J. **67** (9), 1367-1385, 2016.

[12] Deniz, E., Cağlar M. and Orhan, H. *Some convexity properties for two new p-valent integral*, Hacet. J. Math. Stat. **40** (6), 829 – 837, 2011.

[13] Deniz, E. and Orhan, H. *Certain subclasses of multivalent defined by new multiplier transformations*, Arab. J. Sci. Eng. **36**, 1091-1112, 2011.

[14] Frasin, B. A. *Convexity of integral operators of* $p-$*valent functions*, Math. Comput. Model. **51**, 601-605, 2010.

[15] Frasin, B. A. *Family of analytic functions of complex order*, Acta Math. Acad. Paed. Ny. **22,** 179-191, 2006.

[16] Goodman, A.W. *On uniformly convex functions,* Ann. Polon. Math. **56,** 87–92, 1991.

[17] Kanas, S. and Wisniowska, A. *Conic regions and $k-$uniform convexity,* Comput. Appl. Math. **105,** 327–336, 1999.

[18] Kim, Y. J. and Merkes, E. P. *On an integral of powers of a spirallike function,* Kyungpook Math. J. **12,** 249–252, 1972.

[19] Miller, S. S., Mocanu, P. T. and Reade, M. O. *Starlike integral operators,* Pacific J. Math. **79** (1), 157–168, 1978.

[20] S. S. Miller and P. T. Mocanu, *Second order deferential inequalities in the complex plane,* J. Math. Anal. Appl., **65** , 289–305, 1978.

[21] Nasr, M. A. and Aouf, M. K. *Starlike function of complex order,* J. Natur. Sci. Math. **25** (1), 1-12, 1985.

[22] Orhan, H., Deniz, E. and Raducanu, D. *The Fekete–Szego problem for subclasses of analytic functions defined by a differential operator related to conic domains,* Comput. Math. Appl. **59** (1), 283-295, 2010.

[23] Oros, G. I. and Oros, G. *A convexity property for an Integral Operator* $\mathbb{F}_m$, Stud. Univ. Babes -Bolyai Math. **55** (3), 169-177, 2010.

[24] Pescar, V. and Owa, S. *Sufficient conditions for univalence of certain integral operators,* Indian J. Math. **42** (3), 347-35, 2000.

[25] Pfaltzgraff, J. A. *Univalence of the integral of* $(f'(z))^\lambda$, Bull. London Math. Soc. **7** (3), 254–256, 1975.

[26] Rønning, F. *On starlike functions associated with parabolic regions,* Ann. Univ. Mariae Curie-SkÃlodowska Sect. A. **45** (14), 117–122, 1991.

[27] Wiatrowski, P. *The coefficients of a certain family of holomorphic functions,* Zeszyty Nauk. Uniw.Lodz. Nauki Mat. Pryrod. Ser. **3,** 75-85, 1971.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1379 – 1392

Almost conformal Ricci solituons on 3-dimensional trans-Sasakian manifold

Tamalika Dutta*, Nirabhra Basu† and Arindam Bhattacharyya‡

Abstract

In this paper we have shown that if a 3-dimensional trans-Sasakian manifold M admits conformal Ricci soliton (g, V, λ) and if the vector field V is point wise collinear with the unit vector field ξ, then V is a constant multiple of ξ. Similarly we have proved that under the same condition an almost conformal Ricci soliton becomes conformal Ricci soliton. We have also shown that if a 3-dimensional trans-Sasakian manifold admits conformal gradient shrinking Ricci soliton, then the manifold is an Einstein manifold.

Keywords: conformal Ricci soliton, almost conformal Ricci soliton, conformal gradient shrinking Ricci soliton, trans-Sasakian manifold.

2000 AMS Classification: Primary 53C15; Secondary 53A30, 53C25.

Received : 10.09.2015 *Accepted* : 08.12.2015 *Doi* : 10.15672/HJMS.20164514287

1. Introduction

In 1982 Hamilton [9] introduced the concept of Ricci flow and proved its existence. This concept was developed to answer Thurston's geometric conjecture which says that each closed three manifold admits a geometric decomposition. Hamilton also [9] classified all compact manifolds with positive curvature operator in dimension four. The Ricci flow equation is given by

$$\frac{\partial g}{\partial t} = -2S \tag{1.1}$$

---

*First author is supported by DST 'Inspire' fellowship of India. Reference no: IF140748.
Department of Mathematics, Jadavpur University, Kolkata- 700032, India,
Email : tamalika.bagnan@gmail.com

†Department of Mathematics, Bhowanipur Education Society College, Kolkata- 700020, India,
Email : nirabhra.basu@hotmail.com

‡Department of Mathematics, Jadavpur University, Kolkata- 700032, India,
Email : bhattachar1968@yahoo.co.in

on a compact Riemannian manifold M with Riemannian metric g.

A self-similar solution to the Ricci flow [9], [14] is called a Ricci soliton [10] if it moves only by a one parameter family of diffeomorphism and scaling. The Ricci soliton equation is given by

$$\pounds_X g + 2S = 2\lambda g, \tag{1.2}$$

where $\pounds_X$ is the Lie derivative, S is Ricci tensor, g is Riemannian metric, X is a vector field and λ is a scalar. The Ricci soliton is said to be shrinking, steady, and expanding according as λ is positive, zero and negetive respectively.

A. E. Fischer developed the concept of conformal Ricci flow [7] during 2003-04 which is a variation of the classical Ricci flow equation that modifies the unit volume constraint of that equation to a scalar curvature constraint. The conformal Ricci flow on M where M is considered as a smooth closed connected oriented n-manifold is defined by the equation [7]

$$\frac{\partial g}{\partial t} + 2(S + \frac{g}{n}) = -pg \tag{1.3}$$

and $r(g) = -1$,
where p is a scalar non-dynamical field(time dependent scalar field), $r(g)$ is the scalar curvature of the manifold and n is the dimension of manifold.

In 2015, N. Basu and A. Bhattacharyya [2] introduced the notion of conformal Ricci soliton equation as

$$\pounds_X g + 2S = [2\lambda - (p + \frac{2}{n})]g, \tag{1.4}$$

where λ is constant.
The equation is the generalization of the Ricci soliton equation and it also satisfies the conformal Ricci flow equation.

The concept of Ricci almost soliton was first introduced by S. Pigola, M. Rigoli, M. Rimoldi, A. G. Setti in 2010 [12]. R. Sharma has also done excellent work in almost Ricci soliton [13]. A Riemannian manifold (M^n, g) is an almost Ricci soliton [1], if there exist a complete vector field X and a smooth soliton function $\lambda : M^n \to \mathbb{R}$ satisfying,

$$R_{ij} + \frac{1}{2}(X_{ij} + X_{ji}) = \lambda g_{ij},$$

where R_{ij} and $X_{ij} + X_{ji}$ stand for the Ricci tensor and the Lie derivative $\pounds_X g$ in local coordinates respectively. It will be called expanding, steady or shrinking, respectively, if $\lambda < 0, \lambda = 0$ or $\lambda > 0$.

We introduce the notion of almost conformal Ricci soliton by

$$\pounds_X g + 2S = [2\lambda - (p + \frac{2}{n})]g. \tag{1.5}$$

where $\lambda : M^n \to \mathbb{R}$ is a smooth function.

Now a gradient Ricci soliton on a Riemannian manifold (M^n, g_{ij}) is defined by [6]

$$S + \nabla\nabla f = \rho g, \tag{1.6}$$

for some constant ρ and for a smooth function f on M. f is called a potential function of the Ricci soliton and ∇ is the Levi-Civita connection on M. In particular a gradient shrinking Ricci soliton satisfies the equation,

$$S + \nabla\nabla f - \frac{1}{2\tau} g = 0,$$

where $\tau = T - t$ and T is the maximal time of the solution.

Again for conformal Ricci soliton if the vector field is the gradient of a function f, then we call it as a conformal gradient shrinking Ricci soliton [4]. For conformal gradient shrinking Ricci soliton the equation is

$$S + \nabla\nabla f = (\frac{1}{2\tau} - \frac{2}{n} - p)g. \tag{1.7}$$

where $\tau = T - t$ and T is the maximal time of the solution and f is the Ricci potential function.

2. Preliminaries:

Let M be a connected almost contact metric manifold with an almost contact metric structure (ϕ, ξ, η, g) where ϕ is a $(1,1)$ tensor field, ξ is a vector field, η is a 1-form and g is the compatible Riemannian metric such that

$$\phi^2(X) = -X + \eta(X)\xi, \eta(\xi) = 1, \eta \circ \phi = 0, \phi\xi = 0, \tag{2.1}$$

$$g(\phi X, \phi Y) = g(X, Y) - \eta(X)\eta(Y), \tag{2.2}$$

$$g(X, \phi Y) = -g(\phi X, Y), \tag{2.3}$$

$$g(X, \xi) = \eta(X), \tag{2.4}$$

for all vector field $X, Y \in \chi(M)$.

An almost contact metric structure (ϕ, ξ, η, g) on M is called a trans-Sasakian structure [11], if $(M \times R, J, G)$ belongs to the class W_4 [8], where J is the almost complex structure on $M \times R$ defined by $J(X, f\frac{d}{dt}) = (\phi X - f\xi, \eta(X)\frac{d}{dt})$ for all vector fields X on M and smooth functions f on $M \times R$. It can be expressed by the condition [5],

$$(\nabla_X \phi)Y = \alpha(g(X,Y)\xi - \eta(Y)X) + \beta(g(\phi X, Y)\xi - \eta(Y)\phi X), \tag{2.5}$$

for some smooth functions α, β on M and we say that the trans-Sasakian structure is of type (α, β). From the above expression we can write

$$\nabla_X \xi = -\alpha\phi X + \beta(X - \eta(X)\xi), \tag{2.6}$$

$$(\nabla_X \eta)Y = -\alpha g(\phi X, Y) + \beta g(\phi X, \phi Y), \tag{2.7}$$

For a 3-dimensional trans-Sasakian manifold the following relations hold:

$$2\alpha\beta + \xi\alpha = 0, \tag{2.8}$$

$$S(X,\xi) = (2(\alpha^2 - \beta^2) - \xi\beta)\eta(X) - X\beta - (\phi X)\alpha, \tag{2.9}$$

$$\begin{aligned} S(X,Y) &= (\frac{r}{2} + \xi\beta - (\alpha^2 - \beta^2))g(X,Y) - (\frac{r}{2} + \xi\beta - 3(\alpha^2 - \beta^2))\eta(X)\eta(Y) \\ &- (Y\beta + (\phi Y)\alpha)\eta(X) - (X\beta + (\phi X)\alpha)\eta(Y), \end{aligned} \tag{2.10}$$

where S denotes the Ricci tensor of type $(0,2)$, r is the scalar curvature of the manifold M and α, β are smooth functions on M.

For $\alpha, \beta =$ constant the following relations hold:

$$S(X,Y) = (\frac{r}{2} - (\alpha^2 - \beta^2))g(X,Y) - (\frac{r}{2} - 3(\alpha^2 - \beta^2))\eta(X)\eta(Y), \tag{2.11}$$

$$S(X,\xi) = 2(\alpha^2 - \beta^2)\eta(X), \tag{2.12}$$

$$R(X,Y)\xi = (\alpha^2 - \beta^2)(\eta(Y)X - \eta(X)Y), \tag{2.13}$$

$$QX = (\frac{r}{2} - (\alpha^2 - \beta^2))X - (\frac{r}{2} - 3(\alpha^2 - \beta^2))\eta(X)\xi, \tag{2.14}$$

where Q is the Ricci operator given by $S(X,Y) = g(QX,Y)$. Again,

$$\begin{aligned} (\pounds_\xi g)(X,Y) &= (\nabla_\xi g)(X,Y) - \alpha g(\phi X, Y) + 2\beta g(X,Y) - 2\beta\eta(X)\eta(Y) - \alpha g(X, \phi Y) \\ &= 2\beta g(X,Y) - 2\beta\eta(X)\eta(Y) \ [\because g(X,\phi Y) + g(\phi X, Y) = 0]. \end{aligned}$$

Putting the above value in the conformal Ricci soliton equation (1.4) and taking $n = 3$ we get

$$\begin{aligned} S(X,Y) &= \frac{1}{2}[2\lambda - (p+\frac{2}{3})]g(X,Y) - \frac{1}{2}[2\beta g(X,Y) - 2\beta\eta(X)\eta(Y) \\ &= Ag(X,Y) - \beta g(X,Y) + \beta\eta(X)\eta(Y), \end{aligned} \tag{2.15}$$

where $A = \frac{1}{2}[2\lambda - (p+\frac{2}{3})]$.

Hence we can state the following proposition.

Proposition 2.1 : If a 3-dimensional trans-Sasakian manifold admits conformal Ricci soliton (g, ξ, λ), then the manifolds becomes an η-Einstein manifold.

Also,

$$QX = AX - \beta X + \beta\eta(X)\xi. \tag{2.16}$$

Again for almost conformal Ricci soliton

$$\begin{aligned} S(X,Y) &= \lambda g(X,Y) - \frac{1}{2}(p+\frac{2}{3})g(X,Y) - \beta g(X,Y) + \beta\eta(X)\eta(Y) \\ &= (B + \lambda - \beta)g(X,Y) + \beta\eta(X)\eta(Y), \end{aligned}$$

where $B = -\frac{1}{2}(p+\frac{2}{3})$.

Thus we can state the following proposition.

Proposition 2.2 : A 3-dimensional trans-Sasakian manifold admitting almost conformal Ricci soliton (g, ξ, λ) is an η-Einstein manifold.

Example of a 3-dimensional trans-Sasakian manifold:

In this section we construct an example of a 3-dimensional trans-Sasakian manifold.To construct this, we consider the three dimensional manifold $M = \{(x, y, z) \in R^3 : z \neq 0\}$ where (x, y, z) are the standard coordinates in R^3. The vector fields

$$e_1 = e^{-z}(\frac{\partial}{\partial x} - y\frac{\partial}{\partial z}), e_2 = e^{-z}\frac{\partial}{\partial y}, e_3 = \frac{\partial}{\partial z}$$

are linearly independent at each point of M. Let g be the Riemannian metric defined by

$$g(e_1, e_1) = g(e_2, e_2) = g(e_3, e_3) = 1,$$
$$g(e_1, e_2) = g(e_2, e_3) = g(e_3, e_1) = 0.$$

Let η be the 1-form which satisfies the relation

$$\eta(e_3) = 1.$$

Let ϕ be the $(1, 1)$ tensor field defined by $\phi(e_1) = e_2, \phi(e_2) = -e_1, \phi(e_3) = 0$. Then we have

$$\phi^2(Z) = -Z + \eta(Z)e_3,$$

$$g(\phi Z, \phi W) = g(Z, W) - \eta(Z)\eta(W),$$

for any $Z, W \in \chi(M^3)$. Thus for $e_3 = \xi, (\phi, \xi, \eta, g)$ defines an almost contact metric structure on M. Now, after calculating we have

$$[e_1, e_3] = e_1, [e_1, e_2] = ye^{-z}e_2 + e^{-2z}e_3, [e_2, e_3] = e_2.$$

The Riemannian connection ∇ of the metric is given by the Koszul's formula

$$\begin{aligned} 2g(\nabla_X Y, Z) &= Xg(Y, Z) + Yg(Z, X) - Zg(X, Y) \\ &-g(X, [Y, Z]) - g(Y, [X, Z]) + g(Z, [X, Y]). \end{aligned} \tag{2.17}$$

By Koszul's formula we get

$$\begin{aligned} \nabla_{e_1} e_1 &= -e_3, \nabla_{e_2} e_1 = -ye^{-z}e_2 - \frac{1}{2}e^{-2z}e_3, \nabla_{e_3} e_1 = -\frac{1}{2}e^{-2z}e_2, \\ \nabla_{e_1} e_2 &= \frac{1}{2}e^{-2z}e_3, \nabla_{e_2} e_2 = ye^{-z}e_1 - e_3, \nabla_{e_3} e_2 = \frac{1}{2}e^{-2z}e_1, \\ \nabla_{e_1} e_3 &= e_1 - \frac{1}{2}e^{-2z}e_2, \nabla_{e_2} e_3 = \frac{1}{2}e^{-2z}e_1 + e_2, \nabla_{e_3} e_3 = 0. \end{aligned}$$

From the above we have found that $\alpha = \frac{1}{2}e^{-2z}, \beta = 1$ and it can be easily shown that $M^3(\phi, \xi, \eta, g)$ is a trans-Sasakian manifold.

3. Some results for conformal Ricci soliton and almost conformal Ricci soliton on 3-dimensional trans-Sasakian manifold

A conformal Ricci soliton equation on a Riemannian manifold M is defined by

$$\pounds_V g + 2S = [2\lambda - (p + \frac{2}{3})]g,$$

where V is a vector field.

Let V be pointwise co-linear with ξ i.e. $V = \gamma\xi$ where γ is a function on 3-dimensional trans-Sasakian manifold. Then

$$(\pounds_V g + 2S - [2\lambda - (p + \frac{2}{3})]g)(X, Y) = 0,$$

which implies

$$(\pounds_{\gamma\xi} g)(X, Y) + 2S(X, Y) - [2\lambda - (p + \frac{2}{3})]g(X, Y) = 0.$$

Applying the property of Lie derivative and Levi-civita connection we have

$$\begin{aligned} \gamma g(\nabla_X \xi, Y) + (X\gamma)g(\xi, Y) + (Y\gamma)g(\xi, X) + \gamma g(\nabla_Y \xi, X) + 2S(X, Y) \\ -[2\lambda - (p + \frac{2}{3})]g(X, Y) = 0. \end{aligned}$$

Using (2.3) and (2.6) in the above equation we obtain

$$\begin{aligned} &2\beta\gamma g(X,Y) - 2\gamma\beta\eta(X)\eta(Y) + (X\gamma)\eta(Y) + (Y\gamma)\eta(X) \\ &+2S(X,Y) - [2\lambda - (p + \frac{2}{3})]g(X,Y) = 0. \end{aligned} \tag{3.1}$$

Replacing Y by ξ and using (2.12) in (3.1) we get

$$X\gamma + (\xi\gamma)\eta(X) + 2[2(\alpha^2 - \beta^2)\eta(X)] - [2\lambda - (p + \frac{2}{3})]\eta(X) = 0. \tag{3.2}$$

Again putting $X = \xi$ in (3.2) we get

$$\xi\gamma = \frac{1}{2}[2\lambda - (p + \frac{2}{3})] - 2(\alpha^2 - \beta^2). \tag{3.3}$$

Using (3.3) in (3.2) we have

$$\begin{aligned} X\gamma + (\frac{1}{2}[2\lambda - (p + \frac{2}{3})] - 2(\alpha^2 - \beta^2))\eta(X) + 2(2(\alpha^2 - \beta^2)\eta(X)) \\ -[2\lambda - (p + \frac{2}{3})]\eta(X) = 0, \end{aligned}$$

which implies

$$X\gamma = \frac{1}{2}[2\lambda - (p + \frac{2}{3})]\eta(X) - 2(\alpha^2 - \beta^2)\eta(X). \tag{3.4}$$

Applying exterior differentiation in (3.4) and considering λ as constant we have

$$\frac{1}{2}[2\lambda - (p + \frac{2}{3})] - 2(\alpha^2 - \beta^2) = 0, \tag{3.5}$$

(since $d\eta \neq 0$).

Using (3.5) in (3.4) we have

$$X\gamma = 0$$

implies γ is constant.

Hence from (3.1) we have

$$2\beta\gamma g(X,Y) - 2\gamma\beta\eta(X)\eta(Y) + 2S(X,Y) - [2\lambda - (p + \frac{2}{3})]g(X,Y) = 0$$

i.e.

$$S(X,Y) = \frac{1}{2}[2\lambda - (p + \frac{2}{3})]g(X,Y) - \beta\gamma g(X,Y) + \gamma\beta\eta(X)\eta(Y).$$

Putting $X = Y = e_i$ where $\{e_i\}$ is orthonormal basis of the tangent space TM where TM is a tangent bundle of M and summing over i we get,

$$r = \frac{3}{2}[2\lambda - (p + \frac{2}{3})] - 3\beta\gamma + \gamma\beta. \tag{3.6}$$

Now for conformal Ricci soliton $r = -1$, so putting this value in the above equation we get

$$\lambda = \frac{1}{2}p + \frac{2}{3}\beta\gamma. \tag{3.7}$$

So we can state the following theorem:

Theorem 3.1 : A 3-dimensional trans-Sasakian manifold admitting conformal Ricci soliton and if V is point-wise collinear with ξ, then V is a constant multiple of ξ. Also the value of $\lambda = \frac{1}{2}p + \frac{2}{3}\beta\gamma$ provided α, β are constants.

Again for almost conformal Ricci soliton we consider that λ is a smooth function. Then applying exterior derivative in (3.4) we get

$$\frac{1}{2}[2\lambda - (p + \frac{2}{3})] - 2(\alpha^2 - \beta^2) = 0 \tag{3.8}$$

and

$$d\lambda = 0. \tag{3.9}$$

So λ is a constant function and from (3.4) and (3.8) we get γ is constant.

Hence we can conclude the following theorem:

Theorem 3.2 : If a 3-dimensional trans-Sasakian manifold admits almost conformal Ricci soliton and if V is point-wise collinear with ξ, then V is a constant multiple of ξ as well as λ becomes a constant function i.e. almost conformal Ricci soliton becomes conformal Ricci soliton.

Now, from conformal Ricci soliton equation we have

$$(\pounds_\xi g)(X, Y) = 2\beta[g(X, Y) - \eta(X)\eta(Y)].$$

Using (2.11) in the above equation and from (1.4) we have

$$2\beta[g(X, Y) - \eta(X)\eta(Y)] + 2[(\frac{r}{2} - (\alpha^2 - \beta^2))g(X, Y) - (\frac{r}{2} - 3(\alpha^2 - \beta^2))\eta(X)\eta(Y)]$$
$$-[2\lambda - (p + \frac{2}{3})]g(X, Y) = 0.$$

For conformal Ricci soliton we have $r = -1$, so the above equation becomes

$$[2\beta + 2(\frac{-1}{2} - (\alpha^2 - \beta^2)) - (2\lambda - (p + \frac{2}{3}))]g(X, Y)$$
$$-[2\beta + 2(\frac{-1}{2} - 3(\alpha^2 - \beta^2))]\eta(X)\eta(Y) = 0. \tag{3.10}$$

Now taking $X = Y = \xi$ in (3.10) we get

$$2\beta + 2(\frac{-1}{2} - (\alpha^2 - \beta^2)) - (2\lambda - (p + \frac{2}{3})) - 2\beta$$
$$-2(\frac{-1}{2} - 3(\alpha^2 - \beta^2)) = 0,$$

which gives

$$\lambda = \frac{1}{2}[4(\alpha^2 - \beta^2) + (p + \frac{2}{3})].$$

Since $\alpha^2 \neq \beta^2$ so
(1). Suppose $\alpha^2 \geq \beta^2$, then $(\alpha+\beta)(\alpha-\beta) > 0$ which implies α always greater than β. Then $\lambda > 0$ and the conformal Ricci soliton is shrinking.
(2). Suppose $\alpha^2 < \beta^2$ and $(p + \frac{2}{3}) > 4(\alpha^2 - \beta^2)$, then $(\alpha+\beta)(\alpha-\beta) < 0$ which implies α always less than $-\beta$. Then $\lambda > 0$ and the conformal Ricci soliton becomes shrinking.
(3). Suppose $\alpha^2 < \beta^2$ and $(p + \frac{2}{3}) < 4(\alpha^2 - \beta^2)$, then $(\alpha+\beta)(\alpha-\beta) < 0$ which implies α always less than $-\beta$. Then $\lambda < 0$ and the conformal Ricci soliton becomes expanding.

Theorem 3.3 : A 3-dimensional trans-Sasakian manifold admitting a conformal Ricci soliton (g, ξ, λ) satisfies the following relations:
1. For $\alpha > \beta$, the conformal Ricci soliton is shrinking.
2. For $\alpha < -\beta$ and $(p + \frac{2}{3}) > 4(\alpha^2 - \beta^2)$ the conformal Ricci soliton becomes shrinking.
3. For $\alpha < -\beta$ and $(p + \frac{2}{3}) < 4(\alpha^2 - \beta^2)$ the conformal Ricci soliton becomes expanding.

4. Almost conformal gradient shrinking Ricci soliton on 3-dimensional trans-Sasakian manifold

A conformal gradient shrinking Ricci soliton equation is given by

$$S + \nabla\nabla f = (\frac{1}{2\tau} - \frac{2}{3} - p)g. \tag{4.1}$$

This reduces to

$$\nabla_Y Df + QY = (\frac{1}{2\tau} - \frac{2}{3} - p)Y, \tag{4.2}$$

where D is the gradient operator of g.
From (4.2) it follows that

$$\nabla_X \nabla_Y Df + \nabla_X QY = (\frac{1}{2\tau} - \frac{2}{3} - p)\nabla_X Y.$$

Now,

$$\begin{aligned} R(X,Y)Df &= \nabla_X\nabla_Y Df - \nabla_Y\nabla_X Df - \nabla_{[X,Y]}Df \\ &= (\frac{1}{2\tau} - \frac{2}{3} - p)[\nabla_X Y - \nabla_Y X - [X,Y]] - \nabla_X(QY) + \nabla_Y(QX) + Q[X,Y], \end{aligned}$$

where R is the curvature tensor.
Since ∇ is Levi-Civita connection, so from the above equation we get

$$R(X,Y)Df = -\nabla_X(QY) + \nabla_Y(QX) + Q[X,Y] = (\nabla_Y Q)X - (\nabla_X Q)Y. \tag{4.3}$$

Again differentiating equation (2.14) with respect to W and then putting $W = \xi$ we get

$$(\nabla_\xi Q)X = \frac{dr(\xi)}{2}(X - \eta(X)\xi).$$

So

(4.4) $$g((\nabla_\xi Q)X - (\nabla_X Q)\xi, \xi) = g(\frac{dr(\xi)}{2}(X - \eta(X)\xi), \xi) = 0.$$

Putting this value in (4.3) we get

(4.5) $$g(R(\xi, X)Df, \xi) = 0.$$

Again from (2.13) and (4.5) we obtain

$$(\alpha^2 - \beta^2)(g(X, Df) - \eta(X)\eta(Df)) = 0.$$

Since $\alpha^2 \neq \beta^2$, we have from the above equation

$$g(X, Df) = \eta(X)g(Df, \xi)$$

which implies

(4.6) $$Df = (\xi f)\xi.$$

Now from (4.2) we have

$$g(\nabla_Y Df, X) + g(QY, X) = (\frac{1}{2\tau} - \frac{2}{3} - p)g(Y, X)$$

i.e.

$$\begin{aligned} S(X,Y) - (\frac{1}{2\tau} - \frac{2}{3} - p)g(Y,X) &= g(\nabla_Y(\xi f)\xi, X) \\ &= -\alpha(\xi f)g(\phi Y, X) + \beta(\xi f)g(X,Y) \\ &- \beta(\xi f)\eta(Y)\eta(X) + Y(\xi f)\eta(X). \end{aligned} \tag{4.7}$$

Putting $X = \xi$ in (4.7) we get

$$S(X, \xi) - (\frac{1}{2\tau} - \frac{2}{3} - p)\eta(Y) = Y(\xi f).$$

So

(4.8) $$2(\alpha^2 - \beta^2)\eta(Y) - (\frac{1}{2\tau} - \frac{2}{3} - p)\eta(Y) = Y(\xi f).$$

Now from (4.7) and interchanging X, Y we obtain

$$\begin{aligned} S(X,Y) - (\frac{1}{2\tau} - \frac{2}{3} - p)g(Y,X) &= -\alpha(\xi f)g(\phi X, Y) + \beta(\xi f)g(X,Y) \\ &- \beta(\xi f)\eta(Y)\eta(X) + X(\xi f)\eta(Y). \end{aligned} \tag{4.9}$$

Adding (4.7) and (4.9) we get

$$\begin{aligned} 2S(X,Y) - 2(\frac{1}{2\tau} - \frac{2}{3} - p)g(Y,X) &= 2\beta(\xi f)g(X,Y) - 2\beta(\xi f)\eta(Y)\eta(X) \\ &+ (\xi f)(Y\eta(X) + X\eta(Y)). \end{aligned} \tag{4.10}$$

Putting the value of $Y(\xi f)$ in the above equation we get

$$\begin{aligned} QY - (\frac{1}{2\tau} - \frac{2}{3} - p)Y &= \beta(\xi f)Y - \beta(\xi f)\eta(Y)\xi + 2(\alpha^2 - \beta^2)\eta(Y)\xi \\ &- (\frac{1}{2\tau} - \frac{2}{3} - p)\eta(Y)\xi. \end{aligned}$$

Hence from (4.2) we can write

$$\nabla_Y Df = \beta(\xi f)[Y - \eta(Y)\xi] + [2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} - \frac{2}{3} - p)]\eta(Y)\xi.$$

Now,

$$\begin{aligned} R(X,Y)Df &= \nabla_X \nabla_Y Df - \nabla_Y \nabla_X Df - \nabla_{[X,Y]} Df \\ &= \nabla_X(\beta(\xi f)(Y - \eta(Y)\xi) + (2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} - \frac{2}{3} - p))\eta(Y)\xi) \\ &- \nabla_Y(\beta(\xi f)(X - \eta(X)\xi) + (2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} - \frac{2}{3} - p))\eta(X)\xi) \\ &- \nabla_{[X,Y]} Df \\ &= 2(\alpha^2 - \beta^2)[\eta(Y)\nabla_X \xi - \eta(X)\nabla_Y \xi] - \beta(\xi f)[\eta(Y)\nabla_X \xi - \eta(X)\nabla_Y \xi] \\ &- (\frac{1}{2\tau} - \frac{2}{3} - p)[\eta(Y)\nabla_X \xi - \eta(X)\nabla_Y \xi] - \nabla_{[X,Y]} Df \\ &+ \beta(\xi f)[X,Y]. \end{aligned} \tag{4.11}$$

Also

$$\begin{aligned} \nabla_{[X,Y]} Df &= \beta(\xi f)([X,Y] - \eta([X,Y])\xi) + (2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} - \frac{2}{3} - p))\eta([X,Y])\xi \\ &= \beta(\xi f)[X,Y] - \beta(\xi f)\nabla_X \eta(Y)\xi + \beta(\xi f)\xi(\nabla_X \eta)Y + \beta(\xi f)\nabla_Y \eta(X)\xi \\ &- \beta(\xi f)\xi(\nabla_Y \eta)X + [2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} - \frac{2}{3} - p)]\nabla_X \eta(Y)\xi \\ &- [2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} - \frac{2}{3} - p)]\xi(\nabla_X \eta)Y - [2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} \\ &- \frac{2}{3} - p)]\nabla_Y \eta(X)\xi + [2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} - \frac{2}{3} - p)]\xi(\nabla_Y \eta)X. \end{aligned} \tag{4.12}$$

Putting (4.12) in (4.11) and taking inner product with ξ we have

$$2(\alpha^2 - \beta^2) - (\frac{1}{2\tau} - \frac{2}{3} - p) - \beta(\xi f) = 0.$$

From (4.8) we obtain,

$$\beta(\xi f)\eta(Y) = Y(\xi f). \tag{4.13}$$

Using (4.13) in (4.10) we have

$$S(X,Y) - (\frac{1}{2\tau} - \frac{2}{3} - p)g(X,Y) = \beta(\xi f)g(X,Y).$$

After contraction

$$(\xi f) = \frac{-1}{n\beta} - \frac{1}{\beta}(\frac{1}{2\tau} - \frac{2}{3} - p) = C,$$

where C is a constant.

So from (4.6) we get

$$Df = (\xi f)\xi = C\xi \tag{4.14}$$

Therefore

$$g(Df, X) = g(C\xi, X)$$

which gives

$$df(X) = C\eta(X).$$

Applying exterior differentiation on the above relation we get
$Cd\eta = 0$ as $d^2 f(X) = 0$.

So from (4.14) we have found that f is constant as $d\eta = 0$.

Finally from (4.1) we get

$$\begin{aligned} S(X,Y) &= (\frac{1}{2\tau} - \frac{2}{3} - p)g(X,Y) \\ &= 2(\alpha^2 - \beta^2)g(X,Y). \end{aligned}$$

Hence M is an Einstein manifold.

Thus we can conclude the following theorem:

Theorem 4.1 : If a 3-dimensional trans-Sasakian manifold admits conformal gradient shrinking Ricci soliton, then the manifold is an Einstein manifold.

Acknowledgement: Authors are thankful to honorable referee for valuable suggestions to improve the paper.

References

[1] A. Barros, R. Batista, E. Ribeiro Jr, *Compact Almost Ricci Solitons with constant scalar curvature are gradient*, Monatsh Math(2014), Vol. 174, Isu. 1, pp. 29-39.

[2] Nirabhra Basu, Arindam Bhattacharyya, *Conformal Ricci soliton in Kenmotsu manifold*, Global Journal of Advanced Research on Classical and Modern Geometries(2015), Vol. 4, Isu. 1, pp. 15-21.

[3] Bejan, Crasmareanu, *Ricci solitons in manifolds with quasi-constant curvature*, Publ. Math. Debrecen(2011), Vol. 78, Isu. 1, pp. 235-243.

[4] A. Bhattacharyya, N. Basu, *Some curvature identities on Gradient Shrinking Conformal Ricci Soliton*, Analele Stiintifice Ale Universitatii Al.I.Cuza Din Iasi (S.N) Mathematica Tomul LXI(2015), Vol. 61, Isu. 1, pp. 245-252.

[5] D. E. Blair and J. A. Oubina, *Conformal and related changes of metric on the product of two almost contact metric manifolds*, Publ.Mathematiques(1990), Vol. 34, Isu.1, pp. 199-207.

[6] X. Cao, *Compact Gradient Shrinking Ricci Solitons with positive curvature operator*, J. Geom. Anal.(2007), Vol. 17, Isu. 3, pp 425-433.

[7] A. E. Fischer, *An introduction to conformal Ricci flow*, class.Quantum Grav.(2004), Vol. 21, Isu. 3, pp. S171-S218.

[8] A Gray and M. L. Hervella, *The Sixteen classes of almost Hermitian manifolds and their Linear invariants*, Ann.Mat.Pura Appl.(1980), Vol. 123, Isu.4, pp. 35-58.

[9] R. S. Hamilton, *Three Manifold with positive Ricci curvature*, J.Differential Geom.(1982), Vol. 17, Isu.2, pp. 255-306.

[10] R. S. Hamilton, *The Ricci flow on surfaces*, Contemporary Mathematics(1988), Vol. 71, pp. 237-261.

[11] J. A. Oubina, *New Classes of almost Contact metric structures*, Publ.Math.Debrecen(1985), Vol. 32, Isu. 4, pp. 187-193.

[12] S. Pigola, M. Rigoli, M. Rimoldi, A. G. Setti, *Ricci almost solitons*, arXiv:1003.2945v1, (2010).

[13] Ramesh Sharma, *Almost Ricci solitons and K-contact geometry*, Monatsh Math (2014), Vol. 175, Isu. 4, pp. 621-628.

[14] Peter Topping, *Lecture on the Ricci Flow*, Cambridge University Press(2006).

[15] M. M. Tripathi, *Ricci solitons in contact metric manifolds*, arXiv:0801,4222v1, [mathDG], (2008).

[16] M. Turan, U. C. De, A. Yildiz, *Ricci solitons and gradient Ricci solitons in three-dimensional trans-Sasakian manifolds*, Filomat(2012), Vol. 26, Isu. 2, pp. 363-370.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1393 – 1410

Numerical solution of linear integro-differential equation by using modified Haar wavelets

Fernane Khaireddine*† and Ellaggoune Fateh ‡

Abstract

In this paper, we introduce a numerical method for solving linear Fredholm integro-differential equations of the first order. To solve these equations, we consider the equation solution approximately from rationalized Haar (RH) functions.
The numerical solution of a linear integro-differential equation reduces to solving a linear system of algebraic equations. Also, Some numerical examples are presented to illustrate the efficiency of the method.

Keywords: Block-Pulse Functions, Operational matrix, Volterra integral equations, Integro-differential equations, RHFs.

2000 AMS Classification: 65R20, 33D45, 45D05, 45F05

Received : 12.04.2015 *Accepted* : 13.02.2016 *Doi* : 10.15672/HJMS.20164515687

1. Introduction

Some important problems in science and engineering can usually be reduced to a system of integral and integro-differential equations. Integro-differential equations have attracted much attention and solving such equations has been one of the interesting tasks for mathematicians. Several methods have been proposed for numerical solution of these equations (see, e.g., [12]). One technique is the collocation method; of numerous research papers about this approach we cite here ([6], [18]). Since 1991 the wavelet method has been applied to solving integral equations. Various wavelet bases have been employed. In addition to the conventional Daubechies wavelets [12], the Hermite-type trigonometric wavelets [8], linear B-splines [2], Walsh functions [9], Cohen [8] and Fariborzi [10] wavelets have been used. These solutions are often quite complicated, therefore simplifications are welcome. One possibility is to make use of Haar wavelets, which are mathematically the simplest wavelets. For linear integral equations this approach has

*Department of Mathematics, University of 8 May 1945 Guelma, Algeria, Email: kfernane@yahoo.fr
†Fernane Khaireddine.
‡Department of Mathematics, University of 8 May 1945 Guelma, Algeria, Email: fellaggoune@gmail.com

been realized in ([5], [13]). In this paper we examine the rate of convergence of the modified rationalized method using Haar functions for solving Fredholm integro-Differential equations combined with finite difference methods.

Solving the algebraic system obtained by the (RH) functions method allows one to obtain first derivative approximations using a central difference scheme. We apply the proposed method on some test problems to show its accuracy and efficiency. Also, the error evaluation of this method is presented. Before starting, let us recall some definitions.

1.1. Definition. ([5]) The Haar wavelet is the function defined on the real line $\mathbb{R}$ as:

$$H(t)=\begin{cases} 1, & 0\le t\ <\frac{1}{2}, \\ -1, & \frac{1}{2}\le t\ <1, \\ 0, & \text{otherwise.}\end{cases} \tag{1.1}$$

The Haar wavelet $H(t)$ can be used to define a sequence of one-dimensional (RH)functions on $[0,1)$ as follows:

1.2. Definition. ([5]) The (RH)functions $h_n(t)$, for $n=2^i+j$ with $i\in\mathbb{Z}$ and $j=0,1,\ \ldots,\ 2^{i-1}$, are the functions defined on the interval $[0,1)$ as:

$$h_n(t)=H(2^nt-j)_{|[0,1)} \tag{1.2}$$

Also, we define $h_0(t)=1$ for all $t\in[0,1)$.

In Eq.(1.2), are the orthogonal set of rationalized Haar functions and can be defined on the interval $[0,1)$ as [17]:

$$RH(r,\ t)=h_r(t)=\begin{cases} 1, & \text{if } J_1\le t\ <J_{(\frac{1}{2})}, \\ -1, & \text{if } J_{(\frac{1}{2})}\le t\ <J_0, \\ 0, & \text{otherwise.}\end{cases} \tag{1.3}$$

where, $J_u=\dfrac{j-u}{2^i},\ u=0,\ \dfrac{1}{2},1.$

The value of r is defined by two parameters i and j as:

$$r=2^i+j-1,\ i=0,1,2\ldots,\ j=1,2,\ \ldots,\ 2^i$$

$h_0(t)$ is defined for $i=j=0$ and given by:

$$h_0(t)=1,\ 0\le t<1 \tag{1.4}$$

$h_0(t)$ is also included to make this set complete. The orthogonality property is given by:

$$\int_0^1 RH(r,\ t)RH(v,\ t)dt=\begin{cases} 2^{-i}, & r=v \\ 0 & r\neq v\end{cases} \tag{1.5}$$

where

$$v=2^n+m-1,\ n\ =0,1,2,3,\ \ldots,\ m=1,2,3,\ \ldots,\ 2^n$$

2. Function Approximation

Any function $f(t)$ defined over the interval $[0,1)$, which is $\mathbb{L}^2([0,1))$, can be expanded in (RH) functions as([23]);

$$f(t) = \sum_{r=0}^{+\infty} \alpha_r RH(r,\ t),\ r = 0,1,2,\ \ldots.. \tag{2.1}$$

where the (RH) function coefficients α_r are given by:

$$\alpha_r = \frac{\langle f(t), RH(r,t)\rangle}{\langle RH(r,t), RH(r,t)\rangle} = 2^i \int_0^1 f(t)RH(r,\ t)dt,\ r = 0,1,2,\ \ldots \tag{2.2}$$

with $r = 2^j + i - 1,\ i = 0,1,2,3,\ \ldots,\ j = 1,2,3,\ \ldots,\ 2^n$ and $r = 0$ for $i = j = 0$.

Usually, the series expansion of Eq. (2.1) contains infinite terms. If $f(t)$ is piecewise constant by itself, or may be approximated as piecewise constant during each subinterval, then Eq. (2.1) will be terminated at finite terms. Otherwise, it is truncated up to its first m terms as:

$$f(t) \approx \sum_{r=0}^{k-1} a_r RH(r,\ t) = A^T\phi(t) \tag{2.3}$$

where $k = 2^{\alpha+1}$, and $\alpha = 0,1,2,3,\ \ldots$.

The (RH) function coefficients vector $\phi(t)$ and (RH) functions vector $h(t)$ are defined as;

$$A = [a_0, a_1, a_2,\ \ldots, a_{k-1}]^T \tag{2.4}$$

and

$$\phi(t) = [h_0,\ h_1,\ h_3,\ .....,\ h_{k-1}]^T \tag{2.5}$$

where

$$h_r(t) = RH(r,\ t),\ r = 0,1,2,\ \ldots,\ k-1 \tag{2.6}$$

Babolian et al. proved in [21] that:

$$\begin{aligned} \|f(t)\ -\sum_{r=0}^{k-1} a_r RH(r,\ t)\ \|_{\mathbb{L}^2}^2 &== \|\sum_{r=k}^{+\infty} a_r h_r(t)\ \|_{\mathbb{L}^2}^2 \\ &\le \sum_{r=k}^{+\infty} |a_r|^2 \|h_r(t)\ \|_{\mathbb{L}^2}^2 \\ &= \sum_{r=k}^{+\infty} |a_r|^2 \\ &\sim \sum_{r=k}^{+\infty} 2^{-r} \sim 2^{-k} = O(2^{-k}) \le C2^{-k} \end{aligned} \tag{2.7}$$

where C is a constant of integration.

Now, let $k(t,\ s)$ be a function of two independent variables defined for $t \in [0,1)$ and $s \in [0,1)$. Then $k(t,\ s)$ can be expanded in (RH) functions as:

$$k(t,\ s) = \sum_{u=0}^{m-1}\sum_{v=0}^{m-1} h_{uv}h_v(t)h_u(s) \tag{2.8}$$

In Eq. (2.8) h_{uv}, for $u = 0,1,2,..,m-1$ and $v = 0,1,2,...,m-1$, is given as:

$$h_{uv} = 2^{i+q}\int_0^1\int_0^1 k(s,\ t)h_v(t)h_u(s)dtds \tag{2.9}$$

where

$u = 2^i + j,\ i \geq 0$ and $0 \leq j < 2^i v = 2^q + r,\ q \geq 0$ and $0 \leq r < 2^q$ hence we have

$$k(t,\ s) = \phi^T(t)H\phi(s) \tag{2.10}$$

where

$$H = (\hat{\Phi}_{k\times k}^{-1})^T\hat{H}\hat{\Phi}_{k\times k}^{-1} \tag{2.11}$$

with

$$\hat{H} = (h_{uv})_{k\times k}^T \tag{2.12}$$

Where $\hat{H}$ is an $k \times k$ matrix such that:

$$h_{ij} = \frac{\langle RH(i,t), \langle k(t,s), RH(j,s)\rangle\rangle}{\langle RH(i,t), RH(i,t)\rangle\langle RH(j,t), RH(j,t)\rangle} \tag{2.13}$$

Take the Newton-Côtes nodes as:

$$t_i = \frac{2i-1}{2k},\ i = 1,2,..,k \tag{2.14}$$

$$\hat{h}_{lp} = k(\frac{2l-1}{2k},\ \frac{2p-1}{2k}),\ p,\ l = 1,2,....,k. \tag{2.15}$$

2.1. Operational matrix of integration. Discrete Haar functions of order k represented by $2^k \times 2^k$ matrix $\hat{\Phi}_{k\times k}$, in the sequency ordering are given by the following recurrence relation ([23]):

$$\hat{\Phi}_{k\times k} = \left\{ \begin{array}{ll} \hat{\Phi}_{\frac{k}{2}\times\frac{k}{2}} & \otimes[1 \quad 1] \\ I_{\frac{k}{2}\times\frac{k}{2}} & \otimes[1 \quad -1] \end{array} \right\} \tag{2.16}$$

$$\Phi_{1\times 1} = [1] \tag{2.17}$$

where $I_{\frac{k}{2}\times\frac{k}{2}}$ is the identity matrix of dimension k and $\otimes$ is the Kronecker product.

The integration of (RH) functions can be expanded into Haar series with Haar coefficient matrix P as follows:

$$(2.18)\qquad \int_0^1 t\phi(x)dx = P\phi(t)$$

The $k \times k$ square matrix $P = P_k$ is called the operational matrix of integration and is given in [7] as:

$$(2.19)\qquad P_k = \frac{1}{2k}\begin{pmatrix} 2kP_{\frac{k}{2}} & -\hat{\Phi}_{\frac{k}{2}} \\ -\hat{\Phi}_{\frac{k}{2}} & 0 \end{pmatrix}$$

where $\hat{\Phi}_1^{-1} = [1],\ P_1 = [\frac{1}{2}], \hat{\Phi}_k$ is given by Eq. (2.16) and

$$(2.20)\qquad \hat{\Phi}_k^{-1} = \frac{1}{k}\hat{\Phi}_k^T\, diag\left(1, 1, 2, 2, \underbrace{2^2, .., 2^2}_{2^2}, ..., \underbrace{2^{\alpha-1}, .., 2^{\alpha-1}}_{2^{\alpha-1}}\right)$$

Also, the integration of the cross-product of two (RH) function vector is:

$$(2.21)\qquad \int_0^1 \phi(t)\phi^T(t)dt = D$$

where D is a diagonal matrix given by:

$$(2.22)\qquad D = diag\left(1, 1, 2, 2, \underbrace{2^2, .., 2^2}_{2^2}, ..., \underbrace{2^{\alpha-1}, .., 2^{\alpha-1}}_{2^{\alpha-1}}\right)$$

2.2. The product operational matrix. ([23])

Let the product of $\phi(t)$ and $\phi^T(t)$ be called the (RH) product matrix $\psi_{k\times k}(t)$. That is:

$$(2.23)\qquad \phi(t)\phi^T(t) = \psi_{k\times k}(t)$$

The basic multiplication properties of (RH) functions are as:

$$(2.24)\qquad h_0(t)h_i(t) = h_i(t\ ),\ i = 0, 1,\ \ldots,\ m-1$$

and for $i < j$, we have

$$(2.25)\qquad h_i(t)h_j(t) = \begin{cases} h_j(t), & \text{if } h_j \textit{ occurs during the positive } \text{half-}\textit{wave of } h_i \\ -h_j(t), & \text{if } h_j \textit{ occurs during the negative } \text{half-}\textit{wave of } h_i \\ 0 & \text{otherwise.} \end{cases}$$

Also, the square of any (RH) functions is a block-pulse, with magnitude unity during both the positive and negative half-waves of (RH) functions.

For notation simplification, let us define:

$$(2.26)\qquad \hat{\phi}_a(t) = [h_0(t),\ \ldots,\ h_{k/2-1}(t)]^T$$

$$(2.27)\qquad \hat{\phi}_b(t) = [h_{k/2}(t),\ \ldots,\ h_{k-1}(t)]^T$$

The matrix $\psi_{k\times k}(t)$ in Eq. (2.23) can be derived easily as follows from ([7]):

$$\psi_{k\times k}(t) = \begin{bmatrix} \psi_{k/2}(t) & D_{k/2}diag[\hat{\phi}_b(t)] \\ diag[\hat{\phi}_b(t)]D^T_{k/2} & di\alpha g[D^{-1}_{k/2}\hat{\phi}_a(t)] \end{bmatrix} \quad (2.28)$$

where

$$\psi_1(t) = [h_0(t)] \quad (2.29)$$

With the above recursive formulas, we can evaluate $\psi_k(t)$ for any $k = 2^\alpha$, where α is a positive integer. Furthermore, by multiplying the matrix $\psi_k(t)$ in Eq. (2.23) by the vector A in Eq. (2.3) we obtain:

$$\psi_k(t)A = \tilde{A}_k\phi(t) \quad (2.30)$$

Where $\tilde{A}_k$ is a $k \times k$ given by [7]:

$$\tilde{A}_k = \begin{bmatrix} \tilde{A}_{k/2}(t) & D_{k/2}diag[\tilde{c}_b] \\ diag[\tilde{c}_b]D^{-1}_{k/2}(t) & diag[\tilde{c}^T_b D_{k/2}] \end{bmatrix} \quad (2.31)$$

where $C_1 = c_0$, and

$$\tilde{c}_a = [c_0,\ \ldots,\ c_{k/2-1}]^T \quad (2.32)$$

$$\tilde{c}_b = [c_{k/2},\ \ldots,\ c_{k-1}]^T \quad (2.33)$$

3. Application of HAAR wavelet method

3.1. Solution of the Linear Fredholm Integro-Differential Equation. Consider the linear Fredholm integro-differential equation given by:

$$\begin{cases} q(t)y'(t) = \int_0^1 k(t,\ s)y(s)ds\ + r(t)y(t)\ + x(t) \\ y(0) = y_0 \end{cases} \quad (3.1)$$

where the functions $x,\ q,\ r \in \mathbb{L}^2([0,1))$, the kernel $k \in \mathbb{L}^2([0,1) \times [0,1))$ are known and $y(t)$ is the unknown function to be determined.

We approximate $x,\ q,\ r,\ y'$ and k using Haar wavelet space as follows:

$$\begin{cases} y(t)\ = Y^T\phi(t) = \phi^T(t)Y \\ y'(t) = Y'^T\phi(t) = \phi^T(t)Y' \\ y(0) = Y_0^T\phi(t) = \phi^T(t)Y_0 \\ x(t)\ = X^T\phi(t) = \phi^T(t)X \\ k(t,\ s)\ = \psi^T(t)K\psi(s) = \psi^T(s)K^T\psi(t) \\ r(t)\ = R^T\phi(t) = \phi^T(t)R \\ q(t)\ = Q^T\phi(t) = \phi^T(t)Q \end{cases} \quad (3.2)$$

where $\phi(t)$ is given by Eq. (2.5) and Y is an unknown $m \times 1$ vector.

k is a known $m \times m$ dimensional matrix given by Eq. (2.8) and X is a known $m \times 1$ vector given by Eq. (2.3).

Substituting Eq. (3.2) into (3.1) we have:

$$Q^T\phi(t)\phi^T(t)Y' = \int_0^t \phi^T(t)H\phi(s)\phi^T(s)(P^TY'+Y_0)ds + R^T\phi(t)\phi^T(t)(P^TY'+Y_0)+X^T\phi(t) \tag{3.3}$$

we have $\phi(t)\phi^T(t) = \psi_{k\times k}(t)$

$$Q^T\psi_{k\times k}(t)Y' = \int_0^t \phi^T(t)H\psi_{k\times k}(s)(P^TY'+Y_0)ds \tag{3.4}$$

$$+R^T\psi_{k\times k}(t)(P^TY'+Y_0)+X^T\phi(t) \tag{3.5}$$

$$Q^T\psi_{k\times k}(t)Y' = \phi^T(t)H\int_0^t \psi_{k\times k}(s)(P^TY'+Y_0)ds \tag{3.6}$$

$$+R^T\psi_{k\times k}(t)(P^TY'+Y_0)+X^T\phi(t) \tag{3.7}$$

by Eq. (2.21) and by Eq. (2.23), we have $Q^T\psi_{k\times k}(t) = \psi_{k\times k}(t)Q = \widetilde{Q}\phi(t)$

$$\phi^T\widetilde{Q}Y' = \phi^T(t)HD(P^TY'+Y_0)+\phi^T(t)\widetilde{R}(P^TY'+Y_0)+\phi^T(t)X \tag{3.8}$$

or

$$(\widetilde{Q} - HDP^T - RP^T)Y' = HDPY_0+\widetilde{R}Y_0+X \tag{3.9}$$

By solving this linear system we can obtain the vector Y'. Thus,

$$y^{'}(t) = Y'^T\phi(t) = \phi^T(t)Y' \tag{3.10}$$

Eq. (3.9) can be solved for the unknown vector Y'.

The numerical solution y_k is obtained by using finite differences formulas to approximate the first time derivative. In general, the first order derivative of second order error central difference formula can be derived from the Taylor series expansion as follows:

The Algorithm

Step 1:

Put $h = \frac{1}{k}$, $k \in \mathbb{N}$, , $y(0) = y_0$ (initial condition is given)

Step 2:

Set $t_i = ih$, with $t_0 = 0$ and $t_k = 1$, $i = 0, 1, \ldots, k$.

Step 3:

for $i = 1, 2, \ldots, k-1$

$$Y'(t_i) \approx \frac{Y(t_{i+1}) - Y(t_{i-1})}{2h} \tag{3.11}$$

for $i = k$

$$Y'(t_i) \approx \frac{3Y(t_i) - 4Y(t_{i-1}) + Y(t_{i-2})}{2h} \tag{3.12}$$

Use step 1 and step 2, 3 to find the approximate value of y_k. Where $h \approx \frac{1}{k}$ is interval length between nodes.

4. Numerical Examples

In this section, we consider three integro-differential equations. We apply the system of equations in (3.8) and (3.9-3.10). The programs have been provided by **MATLAB 7.8**.

The $\mathbb{L}^2$, $\mathbb{L}^\infty$ error and rate of convergence are defined to be, respectively:

$$e_2 = \|y_k(t) - y_{ex}(t)\|_2 = (\int_0^1 (y_k(t) - y_{ex}(t))^2 dx)^{\frac{1}{2}} \tag{4.1}$$

$$e_\infty = \max_{1 \le i \le 2M} |y_k(t_i) - y_{ex}(t_i)| \tag{4.2}$$

$$\rho_{2,\infty} = \frac{\log[e_{2,\infty}(\frac{k}{2})/e_{2,\infty}(k)]}{\log(2)} \tag{4.3}$$

where $y_{ex}(t)$ is the exact solution and $y_k(t)$ is the approximate solution obtained by Eq. (3.11-3.12).

4.1. Example. Consider the following linear Fredholm integro-differential equation:

$$y'(t) = \int_0^1 e^{ts} y(s)ds + y(t) + \frac{1 - e^{t+1}}{1+t}, \tag{4.4}$$

with initial condition $y(0) = 1$.

The exact solution is as follows: $y(t) = e^t$.

The numerical results are shown in table (1) and in figures (1, 2). Table (1) shows the behaviour of the error for the norm $\mathbb{L}^2$ and norm $\mathbb{L}^\infty$ in function of the parameter of discretization h for different values of k. Note that as h approaches zero, the numerical solution converges to the analytical solution $y(t)$.

k	$\mathbf{L}_2$	$\mathbf{L}_\infty$
8	$8.6353e-002$	$5.4116e-002$
16	$2.9590e-002$	$1.4152e-002$
32	$1.0405e-002$	$3.8126e-003$
64	$3.6778e-003$	$9.9140e-004$
128	$1.3009e-003$	$2.5290e-004$
256	$4.6012e-004$	$6.3875e-005$
512	$1.6271e-004$	$1.6051e-005$
1024	$5.7535e-005$	$4.0231e-006$
2048	$2.0343e-005$	$1.0071e-006$
4096	$7.1925e-006$	$2.5193e-007$
Convergence rate	2.8284	3.9975

Table 1. The errors estimates $\mathbb{L}^2$, $\mathbb{L}^\infty$ and convergence rates $\rho_{2,\infty}$

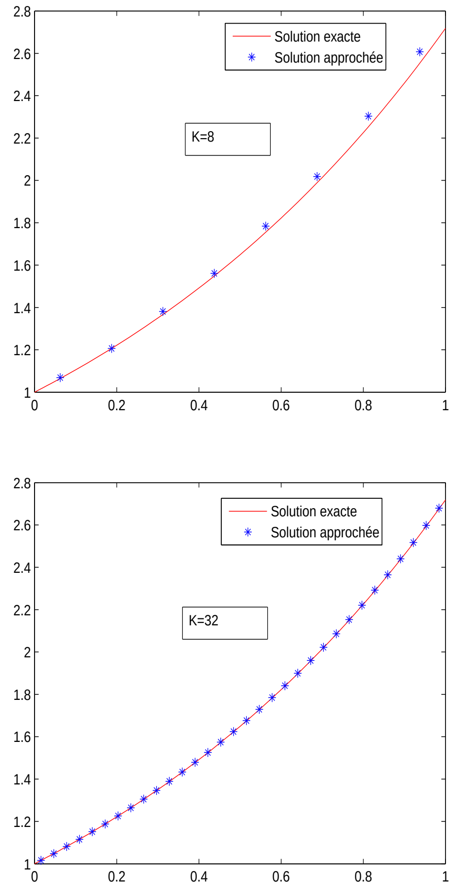


Figure 1. Comparison between approximate solution y_k and exact solution y_{ex}

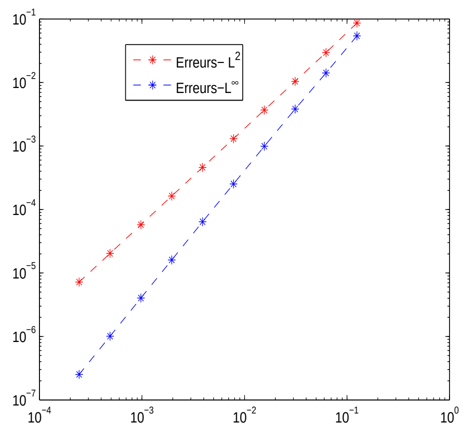


Figure 2. The errors $\mathbb{L}^2$ and $\mathbb{L}^\infty$ with different values of k.

We have also calculated the experimental rate of convergence $\rho_{2,\infty}$. We notice from Table (1) that the convergence rates are 2.8284 in $\mathbb{L}^2$ norm and 3.9975 in norm $\mathbb{L}^\infty$, which is approximately $2\sqrt{2}$ and 4 respectively.

4.2. Example. Consider the following linear Fredholm integro-differential equation:

$$y'(t) = 1 - \frac{1}{3}t + \int_0^1 tsy(s)ds \tag{4.5}$$

with initial condition $y(0) = 0$.

The exact solution is as follows: $y(t) = t$.

The numerical results are shown in table (2) and in figures (3, 4). Table (2) shows the behaviour of the error for the norm $\mathbb{L}^2$ and norm $\mathbb{L}^\infty$ in function of the parameter of discretization h for different values of k. Note that as h approaches zero, the numerical solution converges to the analytical solution $y(t)$.

k	$\mathbf{L}_2$	$\mathbf{L}_\infty$
8	$9.2671e-004$	$6.4523e-004$
16	$4.7431e-004$	$2.1068e-004$
32	$1.1754e-004$	$4.5027e-005$
64	$4.1585e-005$	$1.1443e-005$
128	$1.4705e-005$	$2.8836e-006$
256	$5.1991e-006$	$7.2377e-007$
512	$1.8382e-006$	$1.8130e-007$
1024	$6.4990e-007$	$4.5369e-008$
2048	$2.2977e-007$	$1.1348e-008$
4096	$8.1237e-008$	$2.8376e-009$
Convergence rate	2.8284	3.9992

Table 2. The errors estimates $\mathbb{L}^2$, $\mathbb{L}^\infty$ and convergence rates $\rho_{2,\infty}$

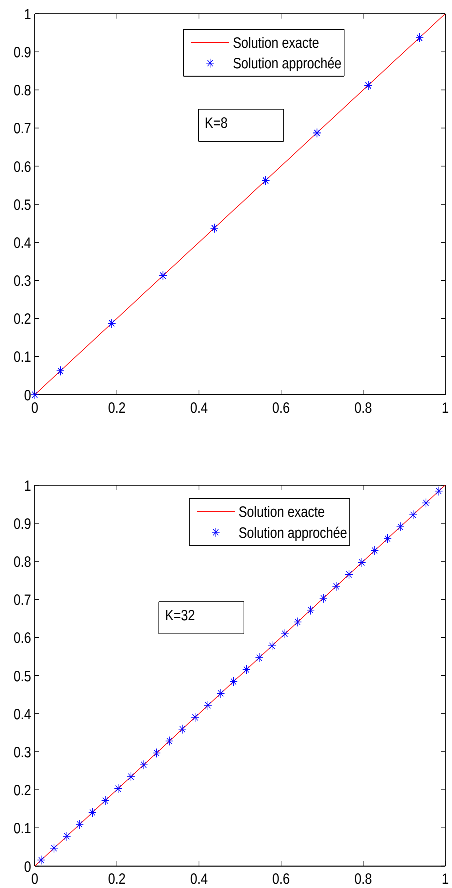


Figure 3. Comparison between approximate solution y_k and exact solution y_{ex}

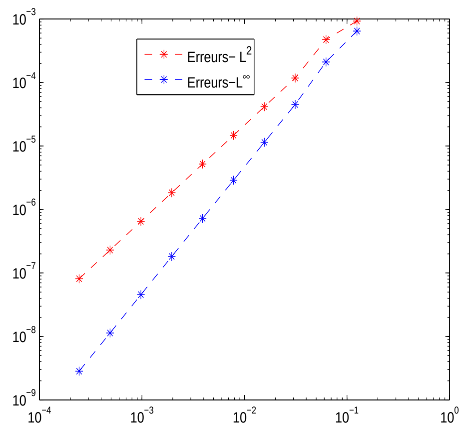


Figure 4. The errors $\mathbb{L}^2$ and $\mathbb{L}^\infty$ with different values of k.

We have also calculated the experimental rate of convergence $\rho_{2,\infty}$. We notice from Table (2) that the convergence rates are 2.8284 in $\mathbb{L}^2$ norm and 3.9975 in norm $\mathbb{L}^\infty$, which is approximately $2\sqrt{2}$ and 4 respectively.

4.3. Example. Consider the following linear Fredholm integro-differential equation:

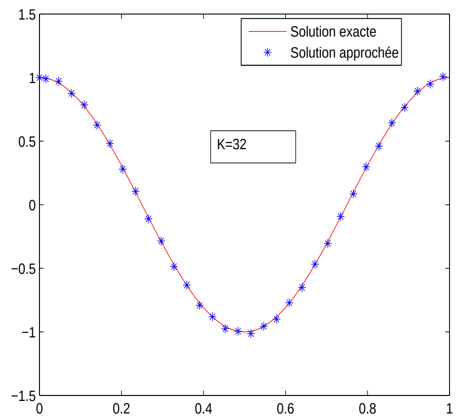
$$y'(t) = \int_0^1 \sin(4\pi t + 2\pi s) y(s) ds + y(t) - \cos(2\pi t) - 2\pi \sin(2\pi t) - \frac{1}{2}\sin(4\pi x), \tag{4.6}$$

with initial condition $y(0) = 1$.

The exact solution is: $y(t) = \cos(2\pi t)$. The numerical results are shown in Table (3) and in figures (5, 6). Table (3) shows the behaviour of the error for the norm $\mathbb{L}^2$ and norm $\mathbb{L}^\infty$ in function of the parameter of discretization h for different values of k. Note that as h approaches zero, the numerical solution converges to the analytical solution $y(t)$.

k	$\mathbf{L}_2$	$\mathbf{L}_\infty$
8	4.9567e-001	3.0091e-001
16	1.7633e-001	7.5816e-002
32	6.2562e-002	1.8990e-002
64	2.2160e-002	4.7500e-003
128	7.8420e-003	1.1877e-003
256	2.7739e-003	2.9696e-004
512	9.8094e-004	7.4244e-005
1024	3.4686e-004	1.8561e-005
2048	1.2264e-004	4.6402e-006
4096	4.3361e-005	1.1601e-006
Convergence rate	2.8284	3.9992

Table 3. The errors estimates $\mathbb{L}^2$, $\mathbb{L}^\infty$ and convergence rates $\rho_{2,\infty}$



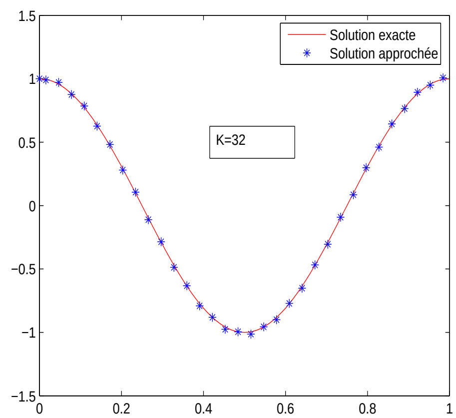


Figure 5. Comparison between approximate solution y_k and exact solution y_{ex}

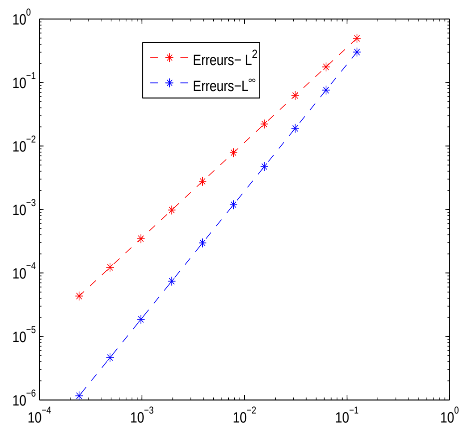


Figure 6. The errors $\mathbb{L}^2$ and $\mathbb{L}^\infty$ with different values of k.

We have also calculated the experimental rate of convergence $\rho_{2,\infty}$. We notice from Table (3) that the convergence rates are 2.8284 in $\mathbb{L}^2$ norm and 3.9975 in norm $\mathbb{L}^\infty$, which is approximately $2\sqrt{2}$ and 4 respectively.

4.4. Example. Consider the Fredholm integral equation of the second kind:

$$y'(t) \quad = \quad \frac{1}{(log2)^2}\int_0^1 (\frac{t}{1+s})y(s)ds + y(t) \; - \frac{1}{2}t \; + \frac{1}{1+t} - log(1+t) \qquad (4.7)$$

with initial condition: $y(0) = 0$. The exact solution is: $y(t) = log(1+t)$. The numerical results are shown in Table (4) and in figures (7, 8). Table (4) shows the behaviour of the error for the norm $\mathbb{L}^2$ and norm $\mathbb{L}^\infty$ in function of the parameter of discretization h for different values of k. Note that as h approaches zero, the numerical solution converges to the analytical solution $y(t)$.

k	$\mathbf{L}_2$	$\mathbf{L}_\infty$
8	1.0262e-002	6.3139e-003
16	4.0064e-003	1.8118e-003
32	1.4848e-003	4.8413e-004
64	5.3714e-004	1.2506e-004
128	1.9207e-004	3.1777e-005
256	6.8289e-005	8.0088e-006
512	2.4211e-005	2.0103e-006
1024	8.5720e-006	5.0359e-007
2048	3.0328e-006	1.2602e-007
4096	1.0726e-006	3.1522e-008
Convergence rate	2.8284	3.9992

Table 4. The errors estimates $\mathbb{L}^2$, $\mathbb{L}^\infty$ and convergence rates $\rho_{2,\infty}$

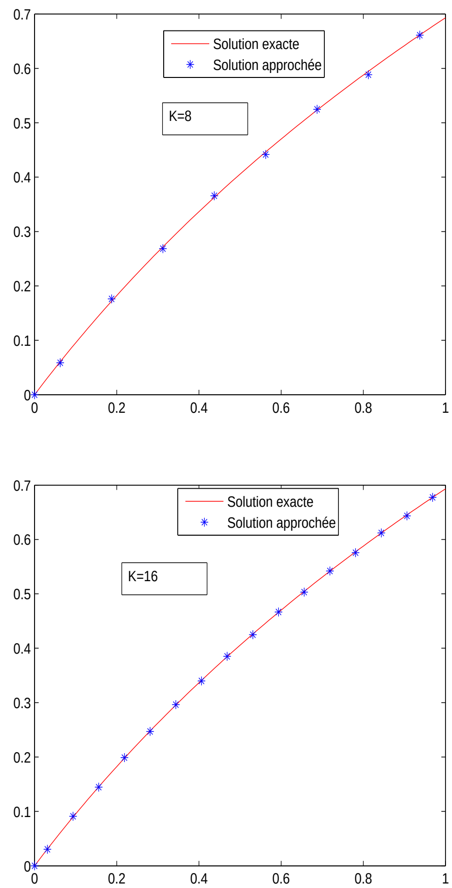


Figure 7. Comparison between approximate solution y_k and exact solution y_{ex}

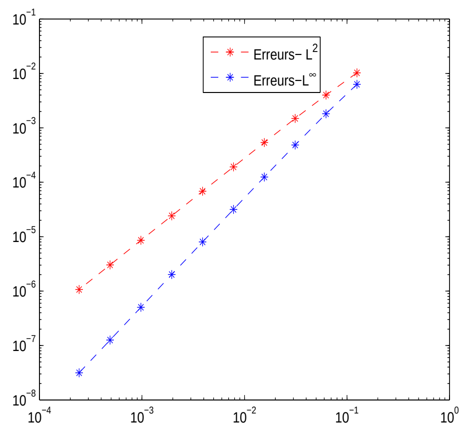


Figure 8. The errors $\mathbb{L}^2$ and $\mathbb{L}^\infty$ with different values of k.

We have also calculated the experimental rate of convergence $\rho_{2,\infty}$. We notice from Table (4) that the convergence rates are 2.8284 in $\mathbb{L}^2$ norm and 3.9975 in norm $\mathbb{L}^\infty$, which is approximately $2\sqrt{2}$ and 4 respectively.

5. CONCLUSION

The proposed method is a powerful procedure for solving linear Fredholm integro-differential. The examples analyzed illustrate the efficiency and reliability of the method presented and show that the method is very simple and effective. The obtained numerical solutions are very accurate, in comparison with the exact solutions. Results also indicate that the convergence rate is fast, and lower order approximations can achieve high accuracy.

ACKNOWLEDGMENT The authors gratefully acknowledge the comments made by anonymous referees.

References

[1] M.A. Abdou, *Fredholm-Volterra integral equation with singular kernel*, Appl. Math. Comput. 137 (2003), pp. 231-243.

[2] A. Ayad, *Spline approximation for first Order fredholm delay integro-differential equations*, International Journal of Computer Mathematics Volume 70, 1999-Issue 3.

[3] E. Babolian, *New direct method to solve nonlinear Volterra-fredholm integral and Integro-differential equations using Operational matrix with block-pulse Functions,* Progress In Electromagnetics Research B, Vol. 8, 59-76, (2008).

[4] E. Babolian and Z. Masouri, *Direct method to solve Volterra integral equation of the rst kind using operational matrix with block-pulse functions*, J. Comput. Appl. Math.,220 (2008), 51-57.

[5] S.BAZM, *A direct rationalized haar functions method to Solve nonlinear two-dimensional fredholm Integral equations*, Acta Universitatis Apulensis, No. 31 (2012), pp. 205-221.

[6] H. Brunner, *Collocation method for Volterra integral and related functional equations, Cambridge Monograph on Applied and Computational Mathematics*, Cambridge University Press, Cambridge, MA, (2004).

[7] C.F. Chen, C.H. Hsiao, *Haar wavelet method for solving lumped and distributed-parameter systems*, IEE Proc. Control Theory Appl. 144 (1997) 87-94.

[8] W.S. Chen, and W. Lin, *Galerkin trigonometric wavelet methods for the natural boundary integral equations*. Appl. Math. Comput., 2005, 121, 7992.

[9] En-Bing Lin, *Wavelet Based Methods for Numerical Solutions of Two Dimensional Integral Equations*, Mathematica Aeterna, Vol. 4, (2014), no. 8, 839-853.

[10] M. A. Fariborzi Araghia, S. Daliri and M. Bahmanpour, *Numerical Solution of Integro-Differential Equation by using Chebyshev Wavelet Operational Matrix of Integration*, International Journal of Mathematical Modelling and Computations Vol. 02, No. 02, (2012), 127-136.

[11] A. Filiz, *Numerical Solution of Parabolic Volterra Integro-Differential Equations via Backward-Euler Scheme*, American Journal of Computational and Applied Mathematics (2013), 3(6): 277-282.

[12] U. Lepik and E. Tamme, *Solution of nonlinear Fredholm integral equations via the Haar wavelet method*, Proc. Estonian Acad. Sci. Phys. Math., (2007), 56, 1, 17-27.

[13] U. Lepik, *Haar wavelet method for nonlinear integro-differential equations*, Applied Mathematics and Computation 176 (2006) 324-333.

[14] F. Mirzaee and S. F. Hoseini, *Solving systems of linear Fredholm integro-differential equations with Fibonacci polynomials*, Ain Shams Engineering Journal (2014) 5, 271-283.

[15] F. Mirzaee and S. F. Hoseini, *Numerical solution of linear Fredholm integral equations system by rationalized Haar functions method*, International Journal of Computer Mathematics Volume 80, 2003-Issue 11.

[16] V. Mishra, H. Kaur, and R.C. Mittal *Haar Wavelet Algorithm For Solving Certain Differential, Integral And Integro-Differential Equations*, Int. J. of Appl. Math and Mech. 8(6): 69-82,(2012).

[17] M. S. Muthuvalu and J. Sulaiman, *Numerical Solutions of Linear Fredholm Integral Equations Using Half-Sweep Arithmetic Mean Method*, International Journal of Applied Mathematics and Computation. Volume 3(1), pp 77-86, (2011).

[18] Y. Ordokhani and M. Razzaghi, *Solution of nonlinear Volterra-Fredholm-Hammerstein integral equa- tions via a collocation method and rationalized Haar functions*, Applied Mathematics Letters 21 (2008), 4-9.

[19] S. Pandit and M. Kumar *Haar Wavelet Approach for Numerical Solution of Two Parameters Singu- larly Perturbed Boundary Value Problems*, Appl. Math. Inf. Sci. 8, No. 6, 2965-2974 (2014).

[20] G.M. Phillips, *Analysis of numerical iterative methods for solving integral and integro-differential equations*, Comput. J., 13 (1970), 297-300.

[21] B. Raftari, *Numerical Solutions of the Linear Volterm Integro-differential Equations*: *Homotopy Perturbation Method and Finite Difference Method*, World Applied Sciences Journa19 (Special Issue of Applied Math): 07-12, (2010).

[22] L. Rahmani, B. Rahimi and M. Mordad, *Numerical Solution of Volterra-Fredholm Integro-Differential Equation by Block Pulse Functions and Operational Matrices*, Gen. Math. Notes, Vol. 4, No. 2, June (2011), pp. 37-48.

[23] M.H. Reihani and Z. Abadi, *Rationalized Haar functions method for solving Fredholm and Volterra integral equations*, Journal of Computational and Applied Mathematics 200 (2007) 12-20.

[24] A. R. Vahidi and T. Damercheli, *A Modiffied ADM for Solving Systems of Linear Fredholm Integral Equations of the Second Kind*, Applied Mathematical Sciences, Vol. 6, (2012), no. 26, 1267-1273.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1411 – 1420

The existence of extremal solutions to nonlinear fractional integro-differential equations with advanced arguments

Neda Khodabakhshi*† and S. Mansour Vaezpour‡

Abstract

This paper deals with the existence of extremal solutions for nonlinear fractional integro-differential equations with advanced arguments. Our analysis rely on monotone iterative method based on upper and lower solutions. Also, we give an illustrative example in order to indicate the validity of our assumptions.

Keywords: Monotone iterative method, Riemann-Liouville fractional derivative, Upper and lower solutions.

2000 AMS Classification: Primary 26A33; Secondary 34K37.

Received : 10.06.2015 *Accepted* : 11.11.2015 *Doi* : 10.15672/HJMS.20164514285

1. Introduction

Fractional calculus is a branch of mathematical analysis, that provides integrals and derivatives of any arbitrary order and due to their multiple applications in many areas of science and engineering has grown extensively. [1, 3, 4, 7, 8, 9, 10, 11, 14, 15, 16, 17]. The monotone iterative method based on upper and lower solutions is a fruitful tools that provides an efficient mechanism to prove the existence results for nonlinear differential problems. We refer the reader to the book [5] and recent papers [2, 6, 12, 13, 18, 19, 20, 21, 22].

As far as we know, few authors consider the existence of extremal solutions for nonlinear Riemann-Liouville fractional integro-differential equations with advanced arguments. So this paper is devoted to study of the following nonlinear boundary value problem:

$$(1.1)\qquad \begin{cases} (D^{\alpha}x(t))' = f(t, x(t), D^{\alpha}x(t), D^{\beta}x(t), Tx(t), Sx(t)), & t \in J := [0, T], \\ D^{\alpha}x(0) = x^*, \ t^{1-\alpha}x(t)|_{t=0} = 0, \ 0 < \beta \leq \alpha \leq 1, \end{cases}$$

*Department of Mathematics and Computer Sciences, Amirkabir University of Technology, Tehran, Iran. Email: khodabakhshi@aut.ac.ir

†Corresponding Author.

‡S. Mansour Vaezpour, Email: vaez@aut.ac.ir

where $f \in C(J \times \mathbb{R}^5, \mathbb{R})$,

$$(Tx)(t) = \int_0^t k(t,s)x(s)ds, \quad (Sx)(t) = \int_0^T h(t,s)x(s)ds,$$

$k(t,s) \in C[D, \mathbb{R}^+],\ h(t,s) \in C[[0,T]^2, \mathbb{R}^+],\ D = \{(t,s) \in \mathbb{R}^2 | 0 \leq s \leq t \leq T\}$ and D^α, D^β are the Riemann-Liouville fractional derivatives.

The innovation of this study is that the nonlinear term f involve unknown function $x(t)$ and it's Riemann-Liouville fractional derivatives with different orders and integral operators Tx, Sx. Therefore, from this point of view, we generalize some recent works. Moreover, with a suitable choice of upper and lower solutions and condition on function f, we obtain the existence of extremal solutions and also present iterative sequences which are convergent to them.

This paper is organized as follows: in section 2, some facts and results about fractional calculus are given, also we consider the existence of the extremal solutions for first order nonlinear differential equation, while in spire of [20] we prove the main result in section 3 and we conclude this paper by considering an example in section 4.

2. Preliminaries and some lemmas

In this section, we present some definitions and results which will be needed later.

2.1. Definition. ([4]) The Riemann-Liouville fractional integral of order $\alpha > 0$ of a function $f : (0, \infty) \to \mathbb{R}$ is defined by

$$I^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s)ds, \quad t > 0,$$

provided that the right-hand side is pointwise defined.

2.2. Definition. ([4]) The Riemann-Liouville fractional derivative of order $\alpha > 0$ of a continuous function $f : (0, \infty) \to \mathbb{R}$ is defined by

$$D^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} (\frac{d}{dt})^n \int_0^t (t-s)^{n-\alpha-1} f(s)ds \quad t > 0,$$

where $n = [\alpha] + 1$, provided that the right-hand side is pointwise defined. In particular, for $\alpha = n$, $D^n f(t) = f^{(n)}(t)$.

1. Remark. The following properties are well known:

$D^\alpha I^\alpha f(t) = f(t),\ \alpha > 0,\ f(t) \in L^1(0, \infty)$,
$D^\beta I^\alpha f(t) = I^{\alpha-\beta} f(t),\ \alpha > \beta > 0,\ f(t) \in L^1(0, \infty)$.

2.1. Lemma. ([4]) Let $Re(\alpha) > 0,\ n = [Re(\alpha)] + 1$ and let $f_{n-\alpha}(t) = I^{n-\alpha} f(t)$ be the fractional integral of order $n - \alpha$. If $f(t) \in L^1(0,T)$ and $f_{n-\alpha} \in AC^n[0,T]$, then we have the following equality

$$I^\alpha D^\alpha f(t) = f(t) - \sum_{i=1}^{n} \frac{f_{n-\alpha}^{(n-i)}(0)}{\Gamma(\alpha - i + 1)} t^{\alpha-i}.$$

2.2. Lemma. The nonlinear fractional differential equation (1.1) is equivalent to the following IVP:

$$(2.1) \qquad \begin{cases} u'(t) = f(t, I^\alpha u(t), u(t), I^{\alpha-\beta} u(t), T_1 u(t), S_1 u(t)), & t \in J, \\ u(0) = x^*,\ 0 < \beta \leq \alpha \leq 1, \end{cases}$$

where

$$T_1u(t) = \int_0^t k_1(t,s)u(s)ds, \quad S_1u(t) = \int_0^T h_1(t,s)u(s)ds,$$

$$k_1(t,s) = \int_s^t \frac{(\tau - s)^{\alpha-1}k(t,\tau)}{\Gamma(\alpha)}d\tau, \quad h_1(t,s) = \int_s^T \frac{(\tau - s)^{\alpha-1}h(t,\tau)}{\Gamma(\alpha)}d\tau.$$

Proof. Take $D^\alpha x(t) = u(t)$ in (1.1), taking into account that $t^{1-\alpha}x(t)|_{t=0} = 0$, we get

$$x(t) = I^\alpha u(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1}u(\tau)d\tau,$$

also

$$\begin{aligned} Tx(t) = T(I^\alpha u(t)) &= \int_0^t k(t,s)(I^\alpha u(t))_{t=s}ds \\ &= \int_0^t k(t,s)\Big(\int_0^s \frac{(s-\tau)^{\alpha-1}u(\tau)}{\Gamma(\alpha)}d\tau\Big)ds \\ &= \int_0^t \Big(\int_\tau^t \frac{(s-\tau)^{\alpha-1}k(t,s)}{\Gamma(\alpha)}ds\Big)u(\tau)d\tau \\ &= \int_0^t \Big(\int_s^t \frac{(\tau-s)^{\alpha-1}k(t,\tau)}{\Gamma(\alpha)}d\tau\Big)u(s)ds \\ &= \int_0^t k_1(t,s)u(s)ds. \end{aligned}$$

The same process can be repeated for S. So the proof is completed. □

Presently, we prove a comparison result for the first order initial value problem (2.1).

2.3. Lemma. Let $w \in C^1(J,\mathbb{R})$ satisfy the relations

$$\begin{cases} w'(t) \geq -KL_\alpha w(t) - Lw(t) - ML_{\alpha-\beta}w(t) - NT_1w(t) - PS_1w(t), \\ w(0) \geq 0, \ 0 < \beta \leq \alpha \leq 1, \end{cases} \tag{2.2}$$

where $K, L, M, N, P \geq 0$ are constants and $L_\alpha w(t) = \int_0^t \frac{(t-s)^{\alpha-1}w(s)}{\Gamma(\alpha)}ds$. If

$$\int_0^T \Big[\frac{Kt^\alpha}{\Gamma(\alpha+1)} + L + \frac{Mt^{\alpha-\beta}}{\Gamma(\alpha-\beta+1)} + N\int_0^t k_1(t,s)ds + P\int_0^T h_1(t,s)ds\Big]dt < 1. \tag{2.3}$$

Then $w(t) \geq 0, \ \forall t \in J$.

Proof. Suppose $w(t) \geq 0$ is not true, then there exists a $t_0 \in (0,T]$ such that $w(t_0) < 0$. Let $\max\{w(t) : 0 \leq t \leq t_0\} = \lambda$, then $\lambda \geq 0$.
If $\lambda = 0$, the proof is similar to Lemma (2.1) of [20].
If $\lambda > 0$, then there exists a $t_1 \in [0,t_0]$ such that $w(t_1) = \lambda > 0$. From (2.2), we have

$$\begin{aligned} w'(t) \geq &-\lambda\Big[\frac{Kt^\alpha}{\Gamma(\alpha+1)} + L + \frac{Mt^{\alpha-\beta}}{\Gamma(\alpha-\beta+1)} \\ &+ N\int_0^t k_1(t,s)ds + P\int_0^T h_1(t,s)ds\Big], \quad \forall t \in [0,t_0]. \end{aligned}$$

Thus, we have

$$\begin{aligned}w(t_0) &= w(t_1) + \int_{t_1}^{t_0} w'(t)dt \\ &\geq \lambda - \lambda \int_0^T \Big[\frac{Kt^\alpha}{\Gamma(\alpha+1)} + L + \frac{Mt^{\alpha-\beta}}{\Gamma(\alpha-\beta+1)} + N\int_0^t k_1(t,s)ds \\ &\qquad + P\int_0^T h_1(t,s)ds\Big]dt \\ &= \lambda\Big(1 - \int_0^T \Big[\frac{Kt^\alpha}{\Gamma(\alpha+1)} + L + \frac{Mt^{\alpha-\beta}}{\Gamma(\alpha-\beta+1)} + N\int_0^t k_1(t,s)ds \\ &\qquad + P\int_0^T h_1(t,s)ds\Big]dt\Big).\end{aligned}$$

Then, by $w(t_0) < 0$, we get

$$\begin{aligned}&\int_0^T \Big[\frac{Kt^\alpha}{\Gamma(\alpha+1)} + L + \frac{Mt^{\alpha-\beta}}{\Gamma(\alpha-\beta+1)} + N\int_0^t k_1(t,s)ds \\ &\quad + P\int_0^T h_1(t,s)ds\Big]dt > 1,\end{aligned}$$

which is contradiction. □

2.4. Lemma. If (2.3) holds. Then the linear problem

$$(2.4)\qquad \begin{cases} u'(t) = g(t) - KI^\alpha u(t) - Lu(t) - ML^{\alpha-\beta}u(t) - NT_1u(t) - PS_1u(t), \\ u(0) = x^*, g \in C(J,\mathbb{R}),\ 0 < \beta \leq \alpha \leq 1, \end{cases}$$

has a unique solution $u^* \in C^1(J,\mathbb{R})$.

Proof. We know that, $u(t) \in C^1(J,\mathbb{R})$ is a solution of (2.4) if and only if $u(t) \in C(J,\mathbb{R})$ is a solution of the following integral equation

$$\begin{aligned}u(t) &= x^* e^{-\int_0^t L ds} + \int_0^t e^{-\int_s^t L d\tau}\Big(g(s) - KI^\alpha u(s) - MI^{\alpha-\beta}u(s) \\ &\qquad - NT_1u(s) - PS_1u(s)\Big)ds \\ &= Au(t).\end{aligned}$$

For any $u, v \in C(J,\mathbb{R})$, we show that A is a contraction operator.

$$
\begin{aligned}
|Au(t)-Av(t)| &= \Big| \int_0^t e^{L(s-t)}\Big[g(s)-KI^{\alpha}u(s)-MI^{\alpha-\beta}u(s)-NT_1u(s)-PS_1u(s)ds\Big] \\
&\quad - \int_0^t e^{L(s-t)}\Big[g(s)-KI^{\alpha}v(s)-MI^{\alpha-\beta}v(s)-NT_1v(s)-PS_1v(s)\Big]ds\Big| \\
&= \Big| \int_0^t e^{L(s-t)}\Big[K(I^{\alpha}(v-u)(s))+M(I^{\alpha-\beta}(v-u)(s))+N(T_1(v-u)(s)) \\
&\qquad + P(S_1(v-u)(s))\Big]ds\Big| \\
&\leq \int_0^T \Big|K(I^{\alpha}(v-u)(s))+M(I^{\alpha-\beta}(v-u)(s))+N(T_1(v-u)(s)) \\
&\qquad + P(S_1(v-u)(s))\Big|ds \\
&\leq \int_0^T \Big[\frac{Ks^{\alpha}}{\Gamma(\alpha+1)}+\frac{Ms^{\alpha-\beta}}{\Gamma(\alpha-\beta+1)}+N\int_0^t k_1(t,s)ds \\
&\qquad + P\int_0^T h_1(t,s)ds\Big]ds||u-v||.
\end{aligned}
$$

Therefore, by condition (2.3), it follows

$$||Au-Av|| < ||u-v||.$$

Thus, by Banach contraction principle A has a unique fixed point u^*, which is unique solution of (2.4). □

2.1. Theorem. Let the following assumptions hold:

- (H_1) There exist $u_0, v_0 \in C^1(J,\mathbb{R})$ satisfying $u_0(t) \leq v_0(t),\ \forall t \in J$,

$$
\begin{cases} u_0'(t) \leq f(t, I^{\alpha}u_0(t), u_0(t), I^{\alpha-\beta}u_0(t), T_1u_0(t), S_1u_0(t)), & t \in J, \\ u_0(0) \leq x^*,\ 0 < \beta \leq \alpha \leq 1, \end{cases} \tag{2.5}
$$

  and v_0 satisfies inverse inequalities of (2.5).
- (H_2) There exist constants $K, L, M, N, P \geq 0$ which satisfy condition (2.3) and

$$
\begin{aligned}
f(t,x,y,z,v,w)-f(t,\bar{x},\bar{y},\bar{z},\bar{v},\bar{w}) &\geq -K(x-\bar{x})-L(y-\bar{y})-M(z-\bar{z}) \\
&\quad -N(u-\bar{u})-P(w-\bar{w}).
\end{aligned}
$$

where $I^{\alpha}u_0(t) \leq \bar{x} \leq x \leq I^{\alpha}v_0(t),\ u_0(t) \leq \bar{y} \leq y \leq v_0(t),\ I^{\alpha-\beta}u_0(t) \leq \bar{z} \leq z \leq I^{\alpha-\beta}v_0(t),\ T_1u_0(t) \leq \bar{v} \leq v \leq T_1v_0(t),\ S_1u_0(t) \leq \bar{w} \leq w \leq S_1v_0(t)\ \forall t \in J$.

Then there exist monotone iterative sequences $\{u_n\}, \{v_n\} \subset [u_0, v_0]$ which converge uniformly to the extremal solutions u^*, v^* of (2.1), respectively, where $\{u_n\}, \{v_n\}$ are defined by

$$u_n(t) = x^* e^{-\int_0^t L ds} + \int_0^t e^{-\int_s^t L d\tau}\Big[f\Big(s, I^\alpha u_{n-1}(s), u_{n-1}(s), I^{\alpha-\beta}u_{n-1}(s), T_1 u_{n-1}(s), S_1 u_{n-1}(s)\Big) - KI^\alpha(u_n - u_{n-1})(s) - L(u_n - u_{n-1})(s) - MI^{\alpha-\beta}(u_n - u_{n-1})(s) - N(T_1(u_n - u_{n-1})(s)) - P(S_1(u_n - u_{n-1})(s))\Big]ds,$$

and

$$v_n(t) = x^* e^{-\int_0^t L ds} + \int_0^t e^{-\int_s^t L d\tau}\Big[f\Big(s, I^\alpha v_{n-1}(s), v_{n-1}(s), I^{\alpha-\beta}v_{n-1}(s), T_1 v_{n-1}(s), S_1 v_{n-1}(s)\Big) - KI^\alpha(v_n - v_{n-1})(s) - L(v_n - v_{n-1})(s) - MI^{\alpha-\beta}(v_n - v_{n-1})(s) - N(T_1(v_n - v_{n-1})(s)) - P(S_1(v_n - v_{n-1})(s))\Big]ds.$$

Also,

$$u_0 \le u_1 \le ... \le u_n \le ... \le u^* \le v^* \le ... \le v_n \le ... \le v_1 \le v_0.$$

Proof. For $\eta \in [u_0, v_0]$, we consider

$$\begin{cases} u'(t) = g_\eta(t) - KI^\alpha u(t) - Lu(t) - MI^{\alpha-\beta}u(t) \\ \quad -N(T_1 u(t)) - P(S_1 u(t)) \\ u(0) = x^*,\ 0 < \beta \le \alpha \le 1, \end{cases} \tag{2.6}$$

where

$$g_\eta(t) = f\Big(t, I^\alpha \eta(t), \eta(t), I^{\alpha-\beta}\eta(t), T_1\eta(t), S_1\eta(t)\Big) + KI^\alpha\eta(t) + L\eta(t) + MI^{\alpha-\beta}\eta(t) + N(T_1\eta(t)) + P(S_1\eta(t)).$$

By Lemma (2.4), we know (2.6) has a unique solution $u \in C^1(J, \mathbb{R})$.
Denote an operator $A : [u_0, v_0] \to C(J, \mathbb{R})$ by $u = A\eta$, then

$$A\eta = x^* e^{-Lt} + \int_0^t e^{L(s-t)}\Big[f\Big(s, I^\alpha\eta(s), \eta(s), I^{\alpha-\beta}\eta(s), T_1\eta(s), S_1\eta(s)\Big) + KI^\alpha\eta(s) + L\eta(s) + MI^{\alpha-\beta}\eta(s) + N(T_1\eta(s)) + P(S_1\eta(s)) - KI^\alpha u(s) - Lu(s) - MI^{\alpha-\beta}u(s) - N(T_1 u(s)) - P(S_1 u(s))\Big]ds.$$

Now, we show that $u_0 \le Au_0,\ Av_0 \le v_0$ and A is nondecreasing.
For the first claim, let $u_1 = Au_0$, $p(t) = u_1(t) - u_0(t)$. we show that $p(t) \ge 0$. By (H_1), we get that

$$\begin{cases} p'(t) \ge -KI^\alpha p(t) - Lp(t) - MI^{\alpha-\beta}p(t) \\ \quad -N(T_1 p(t)) - P(S_1 p(t)), \\ p(0) = u_1(0) - u_0(0) = Au_0(0) - u_0(0) \ge 0. \end{cases}$$

Hence, by Lemma (2.3) $p(t) \ge 0$. Similarly, we can show $Av_0 \le v_0$.
Now, we show that A is nondecresing. Let $u_1 = Au_0$, $v_1 = Av_0$ and $p(t) = v_1(t) - u_1(t)$.

By (H_2), we have

$$\begin{cases} p'(t) \geq -KI^{\alpha}p(t) - L(t) - MI^{\alpha-\beta}p(t) \\ \qquad -N(T_1 p(t)) - P(S_1 p(t)), \\ p(0) = v_1(0) - u_1(0) > 0. \end{cases}$$

So A is nondecreasing.
Next, let $u_n = Au_{n-1}, v_n = Av_{n-1}, n = 1, 2, \ldots$. By the properties of the operator A, we obtain that

$$u_0 \leq u_1 \leq \ldots \leq u_n \leq \ldots \leq u^* \leq v^* \leq \ldots \leq v_n \leq \ldots \leq v_1 \leq v_0.$$

Clearly, u_n, v_n satisfy

$$\begin{cases} u_n'(t) = f(t, I^{\alpha}u_{n-1}, u_{n-1}, I^{\alpha-\beta}u_{n-1}, T_1 u_{n-1}, S_1 u_{n-1}) \\ \qquad -K(I^{\alpha}(u_n - u_{n-1})) - L(u_n - u_{n-1}) - M(I^{\alpha-\beta}(u_n - u_{n-1})) \\ \qquad -N(T_1(u_n - u_{n-1})) - P(S_1(u_n - u_{n-1})), \\ u_n(0) = x^*, \end{cases}$$

$$\begin{cases} v_n'(t) = f(t, I^{\alpha}v_{n-1}, v_{n-1}, I^{\alpha-\beta}v_{n-1}, T_1 v_{n-1}, S_1 v_{n-1}) \\ \qquad -K(I^{\alpha}(v_n - v_{n-1})) - L(v_n - v_{n-1}) - M(I^{\alpha-\beta}(v_n - v_{n-1})) \\ \qquad -N(T_1(v_n - v_{n-1})) - P(S_1(v_n - v_{n-1})), \\ v_n(0) = x^*. \end{cases}$$

The sequences u_n, v_n are uniformly bounded and equicontinuous, so by Arzela-Ascoli Theorem, we find that $\lim_{n\to\infty} u_n(t) = u^*(t)$ and $\lim_{n\to\infty} v_n(t) = v^*(t)$ uniformly on J, and $u^*(t), v^*(t)$ are solutions of (2.1).
Finally, we prove that u^*, v^* are the extremal solutions of (2.1) in $[u_0, v_0]$. Let $w \in [u_0, v_0]$ be any solution of (2.1), then $Aw = w$. By $u_0 \leq w \leq v_0$ and the properties of A, we have

$$u_n \leq w \leq v_n, \quad n = 1, 2, \ldots .$$

Thus, taking limit as $n \to \infty$, we have $u^* \leq w \leq v^*$. That is, u^*, v^* are the extremal solutions of (2.1) in $[u_0, v_0]$.
This completes the proof. □

3. Main result

In this section we prove the existence of extremal solutions of (1.1).
Let $C_{1-\alpha}(J, \mathbb{R}) = \{u \in C(0, T];\ t^{1-\alpha}u \in C(J, \mathbb{R})\}$ and
$DC_{1-\alpha}(J, \mathbb{R}) = \{u \in C_{1-\alpha}(J, \mathbb{R});\ D^{\alpha}u \in C^1(J, \mathbb{R})\}$.

3.1. Theorem. Assume that:
(H_1') There exist $y_0, z_0 \in DC_{1-\alpha}(J, \mathbb{R})$ such that $y_0(t) \leq z_0(t)$ and $D^{\alpha}y_0(t) \leq D^{\alpha}z_0(t)$, are lower and upper solution of (1.1),

$$\begin{cases} (D^{\alpha}y_0(t))' \leq f(t, y_0(t), D^{\alpha}y_0(t), D^{\beta}y_0(t), Ty_0, Sy_0(t)) \\ D^{\alpha}y_0(0) \leq x^*, \quad t^{1-\alpha}y_0(t)|_{t=0} = 0. \end{cases} \tag{3.1}$$

and z_0 satisfies inverse inequalities of (3.1).
(H_2') There exist constants $K, L, M, N, P \geq 0$ which satisfy condition (2.3) such that

$$f(t, x, y, z, u, w) - f(t, \bar{x}, \bar{y}, \bar{z}, \bar{u}, \bar{w}) \geq -K(x - \bar{x}) - L(y - \bar{y}) - M(z - \bar{z}) \\ - N(u - \bar{u}) - P(w - \bar{w}),$$

where $y_0(t) \leq \bar{x} \leq x \leq z_0(t)$, $D^{\alpha}y_0(t) \leq \bar{y} \leq y \leq D^{\alpha}z_0(t)$, $D^{\beta}z_0(t) \leq \bar{z} \leq z \leq D^{\beta}z_0(t)$, $Ty_0(t) \leq \bar{u} \leq u \leq Tz_0(t)$, $Sy_0(t) \leq \bar{w} \leq w \leq Sz_0(t)$.

Then there exist iterative sequences $\{y_n\}$, $\{z_n\}$ which converge uniformly to the extremal solutions y^*, z^* of (1.1), respectively.

Proof. Let $D^\alpha x(t) = u(t)$ in (1.1), then Equation (1.1) is transformed into first order integro-differential equation (2.1). Now, we prove that all the conditions of Theorem (2.1) hold. Let $u_0(t) = D^\alpha y_0(t), v_0(t) = D^\alpha z_0(t)$, we have $u_0(t) \leq v_0(t)$. Also $y_0(t) = I^\alpha u_0(t),\ z_0(t) = I^\alpha v_0(t)$, so by (H_1') u_0, v_0 satisfy (H_1). By (H_2'), it is easy to see that the condition (H_2) holds. Therefore, by Theorem (2.1), we obtain that (2.1) has extremal solutions $u^*, v^* \in C^1(J, \mathbb{R})$ in $[u_0, v_0]$. Let $y^* = I^\alpha u^*,\ z^* = I^\alpha v^*$ so it follows that

$$(3.2) \qquad \begin{cases} D^\alpha y^*(t) = u^*(t) \\ t^{1-\alpha} y^*(t)|_{t=0} = 0 \end{cases}$$

Since u^* satisfies (2.1) and y^* satisfies (3.2), then y^* is a solution of (1.1). Similarly, we can show that z^* is a solution of (1.1). It is easy to show that y^*, z^* are extremal solutions of (1.1). This completes the proof. □

4. Example

Consider the following problem:

$$(4.1) \qquad \begin{cases} (D^{\frac{1}{2}}x(t))' = \frac{-1}{10}x(t) - \frac{1+t}{15}D^{\frac{1}{2}}x(t) - \frac{1+t^2}{20}D^{\frac{1}{4}}x(t) \\ \qquad\qquad - \frac{1+t^3}{30}\int_0^t tsx(s)ds - \frac{1+t^4}{40}\int_0^1 sx(s)ds, \quad t \in [0,1], \\ D^{\frac{1}{2}}x(0) = 0, \quad t^{\frac{1}{2}}x(t)|_{t=0} = 0, \end{cases}$$

where $\alpha = \frac{1}{2}, \beta = \frac{1}{4}$, $k(t,s) = ts,\ h(t,s) = s$. Here,

$$f(t,x,y,z,u,w) = \frac{-1}{10}x - \frac{1+t}{15}y - \frac{1+t^2}{20}z - \frac{1+t^3}{30}u - \frac{1+t^4}{40}w.$$

By easy computation, we have
$K = \frac{1}{10},\ L = \frac{2}{15},\ M = \frac{1}{10},\ N = \frac{1}{15},\ P = \frac{1}{20}$.
Also,

$$\int_0^T \Big[\frac{Kt^\alpha}{\Gamma(\alpha+1)} + L + \frac{Mt^{\alpha-\beta}}{\Gamma(\alpha-\beta+1)} + N\int_0^t k_1(t,s)ds + P\int_0^T h_1(t,s)ds\Big]dt = 0.324 < 1.$$

Now, take $u_0(t) = 0,\ v_0(t) = t^2$. It is easy to see that $u_0,\ v_0$ are lower and upper solution of (4.1). So all the conditions of Theorem (3.1) hold.
Thus there exist iterative sequences $\{u_n\}, \{v_n\}$ which converge uniformly to the extremal solutions u^*, v^* of (4.1), respectively.

Acknowledgement The authors would like to thank the referee for giving useful suggestions and comments for the improvement of this paper.

References

[1] ASME Journal of Computational and Nonlinear Dynamics, Special Issue "Discontinuous and Fractional Dynamical Systems", Guest Editors: J. A. Tenreiro Machado, Albert Luo, vol. 3, Issue 2, 125 pages, April 2008, ISSN: 1555-1415.

[2] Jankowski, T. *Initial value for neutral fractional differential equations involving a Riemann-Liouville derivative*, Appl. Math. Comput. **219** (2013) 7772-7776.

[3] Journal of Vibration and Control, Sage Pub, Special Issue "Mathematical Methods in Engineering", vol. 13, n. 9-10, pp. 1207-1516, Sept. 2007, ISSN: 1077-5463, Guest Editors: K. Tas, J. A. Tenreiro Machado, Dumitru Baleanu.

[4] Kilbas, A.A., Srivastava, H. M. and Trujillo, J. J. *Theory and applications of fractional differential equations*, North-Holland Mathematical Studies, vol. 204, Elsevier Science B.V., Amsterdam, 2006.

[5] Ladde, G.S., Lakshmikantham,V. and Vatsala, A. S. *Monotone Iterative Techniques for Nonlinear Differential Equations*, Pitman, Boston, 1985.

[6] Liu,Z., Sun, J. and Szanto, I *Monotone iterative technique for Riemann-Liouville fractional integro-differential equations with advanced arguments*, Results. Math. **63** (2013), 1277-1287.

[7] Miller, K. S., Ross, B. *An Introduction to the Fractional Calculus and Differential Equations*, Wiley, New York, 1993.

[8] Monje, C. A., Chen, Y., Vinagre, B. M., Xue, D. and Feliu, V. *Fractional-Order Systems and Controls: Fundamentals and Applications*, Springer, New York, NY, USA, 2010.

[9] Oustaloup, A., Sabatier, J., Lanusse, P., Malti, R., Melchior, P. and Moze, M., *An overviewof the CRONE approach in system analysis, modeling and identification, observation and control*, in Proceedings of the 17thWorld Congress, International Federation of Automatic Control (IFAC '08), pp. 14254-14265, Soul, Korea, July 2008.

[10] Petras, I. *Fractional-Order Nonlinear Systems:Modelling, Analysis and Simulation*, Springer, Dordrecht,The Netherlands, 2011.

[11] Podlubny, I. *Fractional Differential Equations*, Academic Press, San Diego, CA, 1999.

[12] McRae, F. A. *Monotone iterative technique and existence results for fractional differential equations*, Nonlinear Anal. **71** (2009) 6093-6096.

[13] Ramirez, J. D., Vatsala, A. S.*Monotone iterative technique for fractional differential equations with periodic boundary conditions*, Opuscula Math. **29** (2009) 289-304.

[14] Tarasov, V. E. *Fractional Dynamics: Applications of Fractional Calculus to Dynamics of Particles*, Fields and Media, Nonlinear Physical Science. Springer, Beijing, Heidelberg (2011).

[15] Tenreiro Machado, J., Kiryakova, V. and Mainardi, F. *A poster about the recent history of fractional calculus*, Fract. Calc. Appl. Anal. **13**, No 3 (2010), 329-334.

[16] Tenreiro-Machado, J. A. , Galhano, A. M. and Trujillo, J. J. *Science Metrics on Fractional Calculus Development Since 1966*, Fract. Calc. Appl. Anal., **16**, No 2 (2013), pp. 479-500.

[17] Uchaikin, V. V. *Fractional Derivatives for Physicists and Engineers*, Volume I Background and Theory; Volume II Applications, Series: Nonlinear Physical Science, Springer Jointly published with Higher Education Press, 2013.

[18] Wang, G. *Monotone iterative technique for boundary value problems of a nonlinear fractional differential equations with deviating arguments*, J. Comput. Appl. Math. **236** (2012) 2425-2430.

[19] Wang, G., Agarwal, R. P. and Cabada, A. *Existence results and the monotone iterative technique for systems of nonlinear fractional differential equations*, Appl. Math. Lett. **25** (2012) 1019-1024.

[20] Wang, G., Baleanu, D. and Zhang, L. *Monotone iterative method for a class of nonlinear fractional differential equations*, Frac. Calc. Appl. Anal. **15** (2012) 244-252.

[21] Wei, Z., Li, G. and Che, J. *Initial value problems for fractional differential equations involving Riemann-Liouville sequential fractional derivative*, J. Math. Anal. Appl. **367** (2010) 260-272.

[22] Zhang, S. *Monotone iterative method for initial value problem involving Riemann-Liouville fractional derivatives*, Nonlinear Anal. **71** (2009) 2087-2093.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1421 – 1434

Quasi-primry submodules satisfying the primeful property I

Hosein Fazaeli Moghimi*† and Mahdi Samiei‡

Abstract

Let R be a commutative ring with identity and M a unital R-module. In this article we extend the notion of quasi-primary ideals to submodules. A proper submodule N of M is called quasi-primary if whenever $rx \in N$ for $r \in R$ and $x \in M$, then $r \in \sqrt{(N:M)}$ or $x \in radN$ where $radN$ is the intersection of all prime submodules of M containing N. Also, we say that a submodule N of M satisfies the primeful property if M/N is a primeful R-module. For a quasi-primary submodule N of M satisfying the primeful property, $\sqrt{(N:M)}$ is a prime ideal of R. For the existence of a module-reduced quasi-primary decomposition, the radical of each term appeared in decomposition must be prime. We provide sufficient conditions, involving the saturation and torsion arguments, to ensure that this property holds as is valid in the ideal case. It is proved that for a submodule N of M over a Dedekind domain R which satisfies the primeful property, N is quasi-primary if and only if $radN$ is prime.

Keywords: Quasi-primary submodule, Primeful property, Prime submodule, Radical of a submodule, Saturation, Torsion.

2000 AMS Classification: 13C13, 13C99

Received : 26.01.2013 *Accepted* : 02.12.2015 *Doi* : 10.15672/HJMS.20164514283

*Department of Mathematics, University of Birjand, P.O.Box 97175-615, Birjand, Iran, Email: `hfazaeli@birjand.ac.ir`

†Corresponding Author.

‡Department of Mathematics, University of Birjand, P.O.Box 97175-615, Birjand, Iran, Email: `mahdisamiei@birjand.ac.ir`

1. Introduction

Throughout this paper all rings are commutative with non-zero identity and all modules are unital. If R is a ring and N a submodule of an R-module M, the ideal $\{r \in R \mid rM \subseteq N\}$ will be denoted by $(N : M)$. Then $ann(M)$, the annihilator of M, is $(0 : M)$. A proper submodule N of M is said to be prime (resp. primary), if $rx \in N$ for $r \in R$ and $x \in M$ implies that either $r \in (N : M)$ (resp. $r \in \sqrt{(N : M)}$) or $x \in N$. In this case, N is called p-prime (resp. p-primary), where $p = (N : M)$ (resp. $p = \sqrt{(N : M)}$) (For more study these notions see for example [3, 13, 14, 16, 17, 19]). The intersection of all prime submodules containing N, denoted $radN$, is called the prime radical of N. Also, N is called a radical submodule if $radN = N$. A proper submodule N of M is called primary-like if $rx \in N$ for $r \in R$ and $x \in M$ implies that $r \in (N : M)$ or $x \in radN$. It is clear that primary-like submodules of R as an R-module and primary ideals of R are the same. Also, N is a prime submodule of M if and only if N is a radical and primary-like submodule of M. The notion of primary-like submodules has been extensively studied by the authors and F. Rashedi in [6].

A proper ideal q of R is said to be quasi-primary if $rs \in q$ for $r, s \in R$ implies $r \in \sqrt{q}$ or $s \in \sqrt{q}$. In particular, q is a quasi-primary ideal of R if and only if $\sqrt{q}$ is a prime ideal of R [7, p.176]. Quasi-primary ideals was first introduced and studied by L. Fuchs [7]. Since primary ideals are quasi-primary, every ideal of a Noetherian ring has a quasi-primary decomposition. Moreover, the uniqueness of the corresponding shortest quasi-primary decompositions of an ideal has been given in [7, Theorem 6]. Here we extend the notion of quasi-primary ideals to submodules. Recall that a proper submodule N of M is quasi-primary if $rx \in N$ for $r \in R$ and $x \in M$ implies that $r \in \sqrt{(N : M)}$ or $x \in radN$. It is clear that primary submodules are quasi-primary. We say that a submodule N of an R-module M satisfies the primeful property if for each prime ideal p of R with $(N : M) \subseteq p$, there exists a prime submodule P containing N such that $(P : M) = p$. If the zero submodule of M satisfies the primeful property, then M is called primeful. For instance finitely generated modules, projective modules over domains and (finite and infinite dimensional) vector spaces are primeful (see [10]). If N satisfies the primeful property, then $\sqrt{(N : M)} = (radN : M)$ [10, Proposition 5.3]. If N is a quasi-primary (primary-like) submodule satisfying the primeful property, then it is easy to verify that $p = \sqrt{(N : M)}$ is a prime ideal of R. In this case, N is called a p-quasi-primary (p-primary like) submodule of M. In [4], Atani and Darani used the term "quasi-primary submodule" in a different way. In fact, they consider a submodule N of an R-module M as a quasi-primary submodule if $\sqrt{(N : M)}$ is a prime ideal of R. Thus a quasi-primary submodule satisfying the primeful property, in the our sense, follows that in [4]. But the converse is not true in general. For example, if $M = \prod_{p\in\Omega} \mathbb{Z}/p\mathbb{Z}$ and $N = \bigoplus_{p\in\Omega} \mathbb{Z}/p\mathbb{Z}$ are $\mathbb{Z}$-modules, where Ω is the set of prime integers, then M is a primeful module and N is a 0-prime submodule of M with $rad(N) = 0$ while N dose not satisfy the primeful property, i.e. M/N is not primeful [10, Example 1 (5)]. Now we give an example of a submodule N such that $\sqrt{(N : M)}$ is a prime ideal while N is not quasi-primary. Consider the $\mathbb{Z}$-module $M = \mathbb{Q} \oplus \mathbb{Z}_p$, where $\mathbb{Q}$ is the additive abelian group of rational numbers and $\mathbb{Z}_p$ is the cyclic group of order p. Then $\mathbb{Q} \oplus 0$ and $0 \oplus \mathbb{Z}_p$ are only prime submodules of M [14, Example 2.6]. Now if $N = 0 \oplus 0$, it is easy to verify that $(N : M) = 0$ and N is not a quasi-primary submodule of M. Also, a quasi-primary submodule does not satisfy the primeful property necessarily. For example, if $M = \mathbb{Z}(p^\infty) \oplus \mathbb{Z}_p$ as a $\mathbb{Z}$-module, where $\mathbb{Z}(p^\infty)$ is the Prüfer group, and $N = 0 \oplus \mathbb{Z}_p$, then $radN = M$ and so N is a quasi-primary submodule of M. But N dose not satisfy the primeful property [14, Example 3.7].

We say that a submodule N of an R-module M has a quasi-primary decomposition if $N = N_1 \cap N_2 \cap \cdots \cap N_t$, where each N_i is a quasi-primary submodule of M. If

$N_i \not\supseteq N_1 \cap \cdots \cap N_{i-1} \cap N_{i+1} \cap \cdots \cap N_t$, for $1 \leq i \leq t$, then the above quasi-primary decomposition is called (1) a reduced quasi-primary decomposition, if the ideals $\sqrt{(N_i : M)}$ are distinct primes; (2) a module-reduced quasi-primary decomposition, if the submodules $radN_i$ are distinct primes; (3) a shortest quasi-primary decomposition, if none of the intersection $(N_{i_1} : M) \cap (N_{i_2} : M) \cap \cdots \cap (N_{i_s} : M)$ $(s > 1)$ is a quasi-primary ideal. In part II, we investigate the existence and uniqueness of these decompositions and relationships between them in different cases. For this purpose we need to some properties and facts about quasi-primary submodules, mostly consideration satisfying the primeful property.

Unlike the ideal case, there are several challenging problems in radical theory of submodules. Finding a good description of $radN$ either in terms of its elements or as some sort of decomposition and splitting the finite intersection of submodules by radical are two examples of them. Some works and methods for characterizing the $radN$ may be found in [1, 12, 15, 16, 18, 19, 20, 21]). One of the main differences between ideal and module cases is that the radical of a quasi-primary submodule is not necessarily prime. In fact, if $R = \mathbb{Z}[x]$, then the submodule $N = R(2, x) + R(x, 0)$ is a quasi-primary submodule of $M = R \oplus R$ whose radical is not prime [19, Theorem 1.9 and Example 1.11]. The mentioned conditions are useful to obtain a module-reduced quasi-primary decomposition from the original one.

In section 2, the behaviuor of quasi-primary submodules (probably satisfying the primeful property) under some operations such as quotient and fraction are considered (Corollary **??** and Theorem 2.14). In this section, it is also shown that q is a quasi-primary ideal of R if and only if qF is a quasi-primary submodule of a free R-module F (Theorem 2.18). In this case $rad(qF)$ is a prime submodule of F. Moreover, it is proved that the radical of every quasi-primary submodule of a free module F over a Noetherian domain R is prime provided that every prime submodule of F contains only finitely many prime submodules (Proposition 2.20).

Let p be a prime ideal of R and N a submodule of M. By the saturation of N with respect to p, we mean the contraction of N_p in M and designate it by $S_p(N)$. It is also known that $S_p(N) = \{x \in M | cx \in N \text{ for some } c \in R\backslash p\}$. Saturations of submodules were investigated in detail in [11] and some results of the study are applied for quasi-primary submodules in section 3. For example, if N is a p-quasi-primary submodule satisfying the primeful property, then we have : (1) $S_p(N)$ is a prime submodule of M if and only if $radN = S_p(N)$ (Theorem 3.3); (2)$S_p(radN) \neq M$ if and only if $radN$ is a prime submodule of M (Theorem 3.9). Also, some other conditions under which the radical of a quasi-primary (probably satisfying the primeful property) is prime have been given in Corollary 3.7, Proposition 3.10, Corollary 3.11 and Theorem 3.14.

The purpose of the section 4 is to discuss about important roles played by torsion submodules in the class of quasi-primary submodules of a module. In Theorem 4.3, it is proved that for a submodule N of a module M over a Dedekind domain R satisfying the primeful property, $radN$ is prime if and only if $M = radN \oplus N'$ for some torsion-free submodule N' of M or $(radN : M) = m$ for some maximal ideal m of R.

In part II, we will characterize the quasi-primary submodules of multiplication modules. Using this, we will fully investigate reduced and module-reduced and shortest quasi-primary decompositions of submodules of multiplication modules. Also, we will give some uniqueness theorems for reduced and module-reduced quasi-primary decompositions of submodules of modules over Noetherian rings.

2. On quasi-primary submodules satisfying the primeful property

In this section, we study basic properties of quasi-primary submodules which probably satisfies the primeful property. In particular we show the affect of some operations on quasi-primary submodules. We start with some elementary results.

2.1. Lemma. *Let M be an R-module. Then the following hold:*

(i) *Any maximal, prime, primary and primary-like submodule is quasi-primary.*
(ii) *Any quasi-primary radical submodule is primary. In particular, if $radN$ is a quasi-primary submodule for a submodule N of M, then $radN$ is primary.*
(iii) *If N is a quasi-primary submodule of M and $(N : M)$ is a radical ideal of R, then N is primary-like.*

2.2. Lemma. *Let M be an R-module. If N is a quasi-primary submodule of M satisfying the primeful property with $p = \sqrt{(N : M)}$, then $radN = rad(N + pM)$.*

Proof. Clearly $radN \subseteq rad(N + pM)$. If P_i is a p_i-prime submodule such that $N \subseteq P_i$, then $p = \sqrt{(N : M)} = (radN : M) \subseteq (P_i : M) = p_i$. Hence $N + pM \subseteq P_i + p_iM \subseteq P_i$. Therefore $rad(N + pM) \subseteq radN$. □

2.3. Theorem. *Let m be a maximal ideal of R and M an R-module. If N is an m-quasi-primary submodule of M satisfying the primeful property, then $radN$ is an m-prime submodule of M. Moreover, $radN = rad(N + mM) = N + mM$.*

Proof. Since N satisfies the primeful property, we have $(radN : M) = \sqrt{(N : M)} = m$ and so $radN$ is an m-prime submodule of M. By Lemma 2.2, $N+mM \subseteq rad(N + mM) = radN$. Sine $radN$ is m-prime, we conclude $m \subseteq (N + mM : M) \subseteq (radN : M) = m$. It follows that $(N + mM : M) = m$. Hence $N + mM$ is a prime submodule containing N. Thus $radN = rad(N + mM) = N + mM$. □

2.4. Proposition. *Let M be an R-module. If N is a quasi-primary submodule of M and L a submodule of M such that $radN \cap radL = rad(N \cap L)$, then $L \subseteq N$ or $N \cap L$ is a quasi-primary submodule of L.*

Proof. Suppose $L \nsubseteq N$. Let $rl \in N \cap L$ for $r \in R \backslash \sqrt{(N \cap L : L)}$ and $l \in L$. Then $rl \in N$ and $r \notin \sqrt{(N : M)}$. Since N is a quasi-primary submodule of M, we have $l \in radN$. Thus $l \in radN \cap radL = rad(N \cap L)$. □

2.5. Corollary. *Let N and K be proper submodules of an R-module M. If N is a quasi-primary submodule of M satisfying the primeful property such that $N \subsetneq K$, then N is also a quasi-primary submodule of K.*

Proof. It follows by applying Proposition 2.4 to N and K. □

2.6. Theorem. *Let N be a proper submodule of a non-zero R-module M. Then the following statements are equivalent:*

(i) *N is a quasi-primary submodule of M;*
(ii) *$\sqrt{(N : K)} = \sqrt{(N : M)}$ for every submodule K of M such that $K \supsetneq radN$.*

Proof. (i)⇒(ii). Let K be any submodule of M such that $K \supsetneq radN$. Then $K/N \subseteq M/N$ and so, $\sqrt{(N : K)} \supseteq \sqrt{(N : M)}$. For the reverse inclusion, let $a \in \sqrt{(N : K)}$. Since $radN \subsetneq K$, we can find an element x of $K \backslash radN$. Then $a^n x \in N$ for some positive integer n. Hence, by (i), $a \in \sqrt{(N : M)}$.
(ii)⇒(i). Suppose $rx \in N$, where $r \in R$ and $x \in M$. Assume $x \notin radN$. Then $radN \subsetneq radN + Rx \subseteq M$. By (ii), $\sqrt{(N : radN + Rx)} = \sqrt{(N : M)}$. Since $rx \in N$, we have

$r(N+Rx)=rN+Rrx\subseteq N$. This shows that $r\in(N:N+Rx)\subseteq\sqrt{(radN:N+Rx)}$. Hence $r\in\sqrt{(N:M)}$, as required. □

2.7. Theorem. *Let $\{N_i : 1\leq i\leq n\}$ be a finite collection of submodules of an R-module M satisfying the primeful property. Then $\cap_{i=1}^n N_i$ satisfies the primeful property and $\sqrt{(\cap_{i=1}^n N_i : M)}=(rad(\cap_{i=1}^n N_i):M)$.*

Proof. Suppose p is a prime ideal of R containing $(\cap_{i=1}^n N_i : M)$. Then $(N_j : M)\subseteq p$, for some $1\leq j\leq n$. Since N_j satisfies the primeful property, there exists a prime submodule P of M containing N_j with $(P:M)=p$. Hence $\cap_{i=1}^n N_i$ satisfies the primeful property and so $\sqrt{(\cap_{i=1}^n N_i : M)}=(rad(\cap_{i=1}^n N_i):M)$. □

The following is a result of Theorem 2.7.

2.8. Corollary. *Let M be an R-module and $\{N_i : i\in I\}$ a collection of quasi-primary submodules of M satisfying the primeful property. Then $(rad(\cap_{i=1}^n N_i):M)=(\cap_{i=1}^n radN_i : M)$.*

It is well-known that for a surjective homomorphism $f:M\to M'$ and a prime submodule N of M containing $Kerf$, $f(N)$ is a prime submodule of M'. It follows that for any submodule N of M, $f(radN)\subseteq radf(N)$. Also if $Kerf\subseteq N$, then $f(radN)=radf(N)$. In particular for every submodule K of M containing N, $rad(K/N)=radK/N$. Analogously we have the following corollaries:

2.9. Theorem. *Let $f:M\to M'$ be a surjective homomorphism. If N' is a quasi-primary submodule of M' such that $f^{-1}(N')$ is containing $Kerf$, then $f^{-1}(N')$ is a quasi-primary submodule of M.*

Proof. Suppose $rm'\in f^{-1}(N')$ and $r\notin\sqrt{(f^{-1}(N'):M)}$. It follows that $rf(m')\in N'$ and $r\notin\sqrt{(N':M')}$. Since N' is a quasi-primary submodule of M', $f(m')\in radN'$; i.e. $f(m')\in P'$ for any prime submodule P' of M' containing N'. Now, let P be a prime submodule of M containing $f^{-1}(N')$. Then $N'=ff^{-1}(N')\subseteq P$. Since $f(P)$ is a prime submodule of M' containing N', we must have $f(m')\in f(P)$. Therefore, there exists an element $x\in P$ such that $m'-x\in Kerf\subseteq P$. Thus $m'\in P$ and so $m'\in rad(f^{-1}(N'))$. □

2.10. Theorem. *Let $f:M\to M'$ be a surjective homomorphism and N a submodule of M. If N is a quasi-primary submodule of M containing $Kerf$, then $f(N)$ is a quasi-primary submodule of M'.*

Proof. Suppose that $rf(x)\in f(N)$ for $r\in R$ and $x\in M$ and $r\notin\sqrt{(f(N):f(M))}$. Hence there exists $n\in N$ such that $rx-n\in Kerf$. Therefor $rx\in N$ and so we have $x\in radN$. Since $f(radN)=rad(f(N))$, we conclude that $f(x)\in rad(f(N))$. □

2.11. Corollary. *Let $f:M\to M'$ be a surjective homomorphism. Then the assignment $N\mapsto f(N)$ defines a one-to-one correspondence between the set of all quasi-primary submodules of M containing $Kerf$ and the set of all quasi-primary submodules N' of M' such that $f^{-1}(N')$ contains $Kerf$.*

From now on, we frequently use the fact that $(radN:M)=\sqrt{(N:M)}$ for a submodule N of M which satisfies the primeful property. Specially it is used in items (ii) and (iii) of the following immediate results.

2.12. Lemma. *Let N be a submodule of an R-module M satisfying the primeful property. Then the following hold:*

(i) *If N is a quasi-primary submodule of M, then $(N : M)$ is a quasi-primary ideal of R.*
(ii) *$radN$ is quasi-primary if and only if $radN$ is primary-like if and only if $radN$ is primary if and only if $radN$ is prime.*
(iii) *If $radN$ is a prime submodule of M, then N is quasi-primary.*

2.13. Theorem. *Let N be a proper submodule of a finitely generated module M over a zero-dimensional ring R. Then N is quasi-primary if and only if there exists a quasi-primary ideal q of R such that $q \subseteq (N : M)$. In particular, N is a quasi-primary submodule of M if and only if $(N : M)$ is a quasi-primary ideal of R.*

Proof. Since M is finitely generated, N satisfies the primeful property, then Lemma 2.12 follows that $(N : M)$ is a quasi-primary ideal. Conversely, let q be a quasi-primary ideal of R such that $q \subseteq (N : M)$. Since M is finitely generated, N is contained in a maximal submodule of M and so $radN \neq M$. Since R is zero-dimensional, $\sqrt{q}$ is a maximal ideal of R and so $\sqrt{q} = \sqrt{(N : M)} = (radN : M)$. Hence $radN$ is a prime submodule of M. Therefore by Lemma 2.12 (iii), N is quasi-primary. □

Let S be a multiplicatively closed subset of R and M an R-module. We denote the ring and module of fractions by $S^{-1}R$ and $S^{-1}M$ respectively.

2.14. Theorem. *Let M be an R-module and N a quasi-primary submodule of M satisfying the primeful property. Let S be a multiplicatively closed subset of R such that $S \cap \sqrt{(N : M)} = \emptyset$. Then $S^{-1}N$ is a quasi-primary submodule of $S^{-1}R$-submodule $S^{-1}M$.*

Proof. It is easy to see that $x/1 \in S^{-1}M \backslash S^{-1}N$ for each $x \in M \backslash radN$ and so $S^{-1}N \neq S^{-1}M$. Suppose $(r/s)(x/t) \in S^{-1}N$ and $r/s \notin \sqrt{(S^{-1}N : S^{-1}M)}$. Since $S^{-1}\sqrt{(N : M)} \subseteq \sqrt{(S^{-1}N : S^{-1}M)}$, then $r \notin \sqrt{(N : M)}$. Thus there exist $u, w \in S$, $y \in N$ such that $wurx = wsty$. It follows $x \in radN$, since N is quasi-primary. Thus $x/t \in S^{-1}radN \subseteq rad(S^{-1}N)$, by [16, Theorems 3.3 and Theorem 3.4]. □

In the following the localization of a ring R and an R-module M at a prime ideal p are denoted by R_p and M_p respectively.

2.15. Theorem. *Let M be an R-module and N a quasi-primary submodule of M satisfying the primeful property. Then $(radN)_p$ is an R_p-prime submodule of M_p where $p = \sqrt{(N : M)} = (radN : M)$. In addition, $radN_p$ is prime and $radN_p = (radN)_p$.*

Proof. By [16, Theorems 3.3 and Theorem 3.4] $(radN)_p \subseteq rad(N_p)$. For the reverse inclusion, it is easy to see that $(radN : M)_p \subseteq ((radN)_p : M_p)$. Since N is quasi-primary, by Lemma 2.12 (i), $(radN : M)_p$ is the unique maximal ideal of R_p. Now we have $(radN : M)_p = ((radN)_p : M_p)$, because $(radN)_p \neq M_p$. Thus $(radN)_p$ is a prime submodule of M_p containing N_p. On the other hand, by [18, Lemma 1.7] $rad(N_p)$ is a prime submodule of M_p containing N_p. Hence $rad(N_p) \subseteq (radN)_p$. □

We remark that if N is a submodule of M satisfying the primeful property, then $radN$ is also satisfies the primeful property. In this case if N is a proper submodule of M, then $radN$ is also proper. Henceforth, we consider $radN \neq M$ when trying to prove $radN$ is prime for a quasi-primary submodule N satisfying the primeful property.

2.16. Proposition. *Let R be a ring and N a quasi-primary submodule of an R-module M satisfying the primeful property. If $\sqrt{(N : M)}$ is a maximal ideal of R, then $radN$ is a prime submodule of M.*

2.17. Proposition. *Let M be an R-module and $\{N_i : i \in I\}$ a collection of submodules of M such that $\sum_{i\in I} N_i$ satisfies the primeful property. Then $\sum_{i\in I} radN_i = M$ if and only if $\sum_{i\in I} N_i = M$.*

Proof. Assume $\sum_{i\in I} radN_i = M$ and $\sum_{i\in I} N_i \neq M$. Then there exists a maximal ideal m of R containing $(\sum_{i\in I} N_i : M)$ and a prime submodule P of M containing $\sum_{i\in I} N_i$ such that $(P : M) = m$. Thus $\sum_{i\in I} radN_i \subseteq P$, a contradiction. The converse is obvious. □

It is well-known that if F is a free R-module and I is an ideal of R, then $(IF : F) = I$ and $rad(IF) = \sqrt{I}F$ [20, Proposition 2.2]. Thus if I is a prime(resp. primary) ideal of R, then IF is prime(resp. primary) submodule of M . Now we give a similar result in the quasi-primary case.

2.18. Theorem. *Let F be a free R-module. Then qF is a quasi-primary submodule of F if and only if q is a quasi-primary ideal of R.*

Proof. Let qF be a quasi-primary submodule of M. Since $(qF : F) = q$, q is a proper ideal of R. Suppose $rs \in q$, for $r \in R$, $s \in R\backslash\sqrt{q}$. Hence $rsF \subseteq qF$ and $s \notin (radqF : F)$, since $radqF = \sqrt{q}F$ [20, Proposition 2.2]. It follows that $r \in \sqrt{qF : F} = \sqrt{q}$. Conversely let q be a quasi-primary ideal of R. Again by $(qF : F) = q$, qF is a proper submodule of F. Suppose $r \notin \sqrt{(qF : F)} = \sqrt{q}$ and $x \notin radqF = \sqrt{q}F$. Hence we have $rx \notin \sqrt{q}F$, since $\sqrt{q}F$ is a prime submodule of F. Thus $rx \notin qF$. □

2.19. Corollary. *Let F be a free R-module. Then the following statements are equivalent.*

(i) *$I = q_1 \cap \cdots \cap q_t$ is a reduced quasi-primary decomposition of the ideal I;*
(ii) *$IF = q_1F \cap \cdots \cap q_tF$ is a reduced quasi-primary decomposition of IF;*
(iii) *$IF = q_1F \cap \cdots \cap q_tF$ is a module-reduced quasi-primary decomposition of IF.*

2.20. Proposition. *If R is a Noetherian domain and F is a free R-module such that every prime submodule of F contains only finitely many prime submodules, then for every non-zero quasi-primary submodule N of F, $radN$ is prime.*

Proof. We first show that R is a one-dimensional ring. Let $0 \subset p' \subseteq p$ be a chain of prime ideals of R. If $p' \neq p$, then there exist infinitely many such prime ideals contained in p [9, p. 144]. It follows from the above argument of Theorem 2.18 that there exist infinitely many prime submodule contained in prime submodule pF, a contradiction. Thus R is a one-dimensional domain. Now, let qF be a non-zero quasi-primary submodule of F. It is clear that $0 \subset q \subseteq \sqrt{(qF : F)}$ and so the proof is completed by Proposition 2.16. □

2.21. Theorem. *Let M be an R-module and N a proper submodule of M. If $N_1, \cdots, N_t$ satisfies the primeful property and N has a reduced quasi-primary decomposition $N = N_1 \cap N_2 \cap \cdots \cap N_t$ such that all the prime ideals associated with N are isolated, then $(N : M) = (N_1 : M) \cap (N_2 : M) \cap \cdots \cap (N_t : M)$ is a reduced quasi-primary decomposition of the ideal $(N : M)$ in R.*

Proof. Suppose not. Since the ideals $\sqrt{(N_i : M)}$ are distinct, we have $(N_i : M) \supseteq \cap_{j\neq i}(N_j : M)$ for some i. Then $\sqrt{(N_i : M)} \supseteq \cap_{j\neq i}\sqrt{(N_j : M)}$. It implies that $\sqrt{(N_i : M)} \supset \sqrt{(N_j : M)}$ for some $i \neq j$, since $\sqrt{(N_i : M)}$ is a prime ideal. The final inclusion contradicts the assumption that $\sqrt{(N_i : M)}$ is an isolated prime ideal of R. □

2.22. Corollary. *Let M be an R-module and N a proper submodule of M. If $N_1, \cdots, N_t$ satisfies the primeful property and N has a reduced quasi-primary decomposition $N = N_1 \cap N_2 \cap \cdots \cap N_t$ such that all the prime ideals associated with N are isolated, then*

(i) *N is quasi-primary if and only if $(N : M)$ is quasi-primary.*
(ii) *N is prime if and only if $(N : M)$ is prime.*

Proof. The necessity of each part is clear. To show sufficiency, let $N = N_1 \cap N_2 \cap \cdots \cap N_t$ be a reduce quasi-primary decomposition of N. By Theorem 2.21 $(N : M) = (N_1 : M) \cap (N_2 : M) \cap \cdots \cap (N_t : M)$ is a reduced quasi-primary decomposition of the ideal $(N : M)$ in R. If $(N : M)$ is quasi-primary, we must have $t = 1$ and so $N = N_1$ is quasi-primary. (ii) is concluded by an analogous argument. □

3. Saturation and radical

Let p be a prime ideal of R and N a submodule of an R-module M. Then $S_p(N) = \{x \in M : cx \in N \text{ for some } c \in R\backslash p\}$ is a submodule of M which is called the saturation of N with respect to p. A submodule N of M is called saturated with respect to p if $S_p(N) = N$. It is easy to verify $S_p(N)$ is a saturated submodule of M with respect to p. In [11], Lu applied the tool of saturation in the context of prime and primary submodules. In this section we develop and use this tool for quasi-primary submodules (probably satisfying the primeful property). In particular, using this, we give some conditions under which the radical of a quasi-primary submodule is prime.

3.1. Lemma. *Let N be a submodule of an R-module M satisfying the primeful property. N is a p-quasi-primary submodule of M if and only if $\sqrt{(N : M)} = p$ is a prime ideal of R and $S_p(N) \subseteq radN$.*

Proof. Suppose N is a p-quasi-primary submodule of M. Since N satisfies the primeful property, it is clear that $\sqrt{(N : M)} = p$ is a prime ideal of R. Let $x \in S_p(N)$. Then $sx \in N$ for some $s \in R \setminus p$. Hence $x \in radN$ and so that $S_p(N) \subseteq radN$.
Assume $\sqrt{(N : M)} = p$ is a prime ideal of R. Let $rx \in N$ and $x \notin radN$. Hence we conclude that $sx \notin N$ for any $s \in R \setminus p$. Thus $r \in p$, as required. □

From now on, we denote the set of all prime ideals of R containing $(N : M)$ by $V(N : M)$.

3.2. Lemma. *Let N be a quasi-primary submodule of an R-module M. Then $S_p(N) \subseteq radN$ for every $p \in V(N : M)$. In particular, if $S_p(N)$ is a prime submodule of M for some $p \in V(N : M)$, then $S_p(N) = radN$.*

Proof. Straightforward. □

3.3. Theorem. *Let N be a p-quasi-primary submodule of an R-module M satisfying the primeful property. $S_p(N)$ is a p-prime submodule of M if and only if $S_p(N) = radN$.*

Proof. Assume that $S_p(N)$ is a p-prime submodule of M. It follows from Lemma 3.2 that $S_p(N) = radN$. Conversely, suppose $S_p(N) = radN$. Let $rx \in S_p(N)$ and $x \notin S_p(N)$. Then $rx \in N$ for some $r \in R \setminus p$. Since N is a p-quasi-primary submodule of M, $r \in \sqrt{(N : M)} = (radN : M) = (S_p(N) : M)$. Thus $S_p(N)$ is a p-prime submodule of M. □

3.4. Lemma. *If a submodule N of an R-module M satisfies the primeful property, then so do $radN$ and $S_p(N)$ for every $p \in V(N : M)$.*

Proof. Suppose p is a prime ideal of R containing $(radN : M)$. Since N satisfies the primeful property and $p \supseteq (N : M)$, there exists a prime submodule P of M containing N such that $(P : M) = p$. It is clear that $P \supseteq radN$ and so $radN$ satisfies the primeful property. For the second part, let p be a prime ideal of R such that $p \supseteq (S_p(N) : M) \supseteq (N : M)$. Then there exists a prime submodule P' of M containing N such that

$(P' : M) = p$. Now, let $x \in S_p(N)$. There exists $s \in R \setminus p$ such that $sx \in N \subseteq P'$. Therefore $x \in P'$. Hence we have $S_p(N) \subseteq P'$, as desired. □

3.5. Theorem. *Let p be a prime ideal of R and N a submodule of an R-module M satisfying the primeful property. Then the following statements are equivalent:*

(i) *$S_p(N)$ is a p-quasi-primary submodule of M;*
(ii) *$(S_p(N) : M)$ is a p-quasi-primary ideal of R;*
(iii) *$\sqrt{(S_p(N) : M)} = (radS_p(N) : M) = p$;*
(iv) *$(S_p(N) : M)$ is a p-primary ideal of R;*
(v) *$S_p(N)$ is a p-primary submodule of M.*

Proof. (i)⇒ (ii)⇒ (iii) is clear by Lemma 3.4.
(iii)⇒ (i). By [11, Result 1(1), page 2658], $S_p(S_p(N)) = S_p(N)$. It implies that $S_p(S_p(N)) \subseteq rad(S_p(N))$ and so $S_p(N)$ is a p-quasi-primary submodule of M by Lemma 3.1.
(iii)⇔ (iv)⇔ (v) is obtained by [11, Theorem 2.3]. □

3.6. Corollary. *Let N be a p-quasi-primary submodule of an R-module M satisfying the primeful property. Then the equivalent conditions in Theorem 3.5 hold.*

Proof. Since N satisfies the primeful property, Lemma 3.4 shows that

$$p = \sqrt{(N : M)} \subseteq \sqrt{(S_p(N) : M)} = (radS_p(N) : M).$$

On the other hand, Lemma 3.1 follows that $(S_p(N) : M) \subseteq p$ and hence $p = (radS_p(N) : M)$. Thus (iii) of Theorem 3.5 holds. □

3.7. Corollary. *Let N be a p-quasi-primary submodule of an R-module M satisfying the primeful property. If $(S_p(N) : M)$ is a radical ideal of R, then $radN$ is a prime submodule of M.*

Proof. It follows from Corollary 3.6 that $S_p(N)$ is a p-primary submodule of M and so $S_p(N)$ is prime, since $(S_p(N) : M)$ is a radical ideal of R. Now the proof is completed by Theorem 3.3. □

3.8. Proposition. *Let N be a p-quasi-primary submodule of an R-module M satisfying the primeful property. Then*

$$\sqrt{S_p(N : M)} = \sqrt{(S_p(N) : M)} = p.$$

In particular, $S_p(N : M)$ and $(S_p(N) : M)$ are p-primary ideals of R.

Proof. Since N is p-quasi-primary, Lemma 3.1 shows that $(S_p(N) : M) \subseteq \sqrt{(N : M)}$. Thus we conclude that $(N : M) \subseteq S_p(N : M) \subseteq (S_p(N) : M) \subseteq \sqrt{(N : M)}$, as required. The second part is clear. □

3.9. Theorem. *Let N be a submodule of an R-module M and p a prime ideal of R such that $p \subseteq (radN : M)$. Then the following statements are equivalent:*

(i) *$S_p(radN) \neq M$;*
(ii) *$(radN : M) = (S_p(radN) : M) = p$;*
(iii) *$S_p(radN)$ is a p-prime submodule of M.*

Further, if N is a p-quasi-primary submodule of M, then the above statements are equivalent to:

(iv) *$radN$ is a p-prime submodule of M.*

Proof. (i)⇒ (ii). By replacing N with $radN$ in [11, Theorem 2.1], we have $(S_p(radN) : M) \subseteq p$. Since $pM \subseteq radN$, we have $p \subseteq (radN : M) \subseteq S_p(radN : M) \subseteq (S_p(radN) : M) \subseteq p$, whence (ii) follows.
(ii)⇒ (iii). Using [11, Theorem 2.3] by replacing N with $radN$.
(iii)⇒ (i) is clearly true.
(iii)⇒ (iv). Let N be a p-quasi-primary submodule of M. It follows from (ii) and (iii) that $S_p(radN)$ is a p-prime submodule of M where $(radN : M) = p$. It follows from Lemma 3.2, $S_p(radN) = radN$. Hence $radN$ is a p-prime submodule of M.
$(iv) \Rightarrow (iii)$ is clear. □

3.10. Proposition. *Let N be a quasi-primary submodule of an R-module M. If $p = (N : M)$ is a prime ideal of R, then $S_p(N) = M$ or $radN$ is a prime submodule of M.*

Proof. Suppose $S_p(N) \neq M$. By [11, Proposition 2.4], $S_p(N)$ is a prime submodule of M. It follows from Lemma 3.2 that $radN$ is a prime submodule of M. □

3.11. Corollary. *Let N be a quasi-primary submodule of an R-module M satisfying the primful property. If $p = (N : M)$ is a prime ideal of R, then $radN$ is a prime submodule of M.*

Proof. Since N satisfies the primeful property, we have $radN \neq M$. Also, it follows from Lemma 3.1 that $S_p(N) \subseteq radN$. Now Proposition 3.10 completes the proof. □

3.12. Proposition. *Let N be a p-quasi-primary submodule of an R-module M satisfying the primeful property. Then $radS_p(N) \subseteq S_p(N + pM) \subseteq S_p(radN)$. In particular, $p = (radS_p(N) : M) = (S_p(N + pM) : M)$.*

Proof. By [11, Theorem 4.3], $S_p(N+pM)$ is a p-prime submodule of M and so $radS_p(N) \subseteq S_p(N + pM)$. Suppose $x \in S_p(N + pM)$. Then $cx \in N + pM$ for some $c \in R\backslash p$. Since $\sqrt{(N : M)} = p$ and $cx \in radN$, we conclude that $x \in S_p(radN)$. Also, we have $p = (radN : M) \subseteq (radS_p(N) : M) \subseteq (S_p(N + pM) : M) = p$, as required. □

The following is a result of [11, Corollary 5.7] and Proposition 3.12.

3.13. Corollary. *Let m be a maximal ideal of R and N a m-quasi-primary submodule of an R-module M satisfying the primeful property. Then $radN = radS_m(N) = S_m(radN) = S_m(N + mM)$.*

3.14. Theorem. *Let R be an Artinian ring and M a module over R. If N is a quasi-primary submodule of M and $p \in V(N : M)$, then the followings hold.*

(i) *$radN$ is a prime submodule of M.*
(ii) *$radS_p(N) = S_p(radN) = S_p(N + pM)$. In particular, $radS_p(N)$ is a prime submodule of M.*

Proof. (i). Since R is an Artinian ring, [2, Theorem 2.16] implies that N satisfies the primeful property. Thus $(N : M)$ is a quasi-primary ideal of R. Since R is zero-dimensional, $\sqrt{(N : M)} = (P : M)$ for all prime submodules P containing N. Hence $p = \sqrt{(N : M)} = (radN : M)$ is a prime ideal of R. Now if $rx \in radN$ and $x \notin radN$, there is a prime submodule P' containing N such that $rx \in P'$ and $x \notin P'$. Thus $r \in (P' : M) = \sqrt{(N : M)} = (radN : M)$ and so $radN$ is prime.
(ii). Suppose $x \in S_p(N + pM)$. Then $cx \in N + pM$ for some $c \in R\backslash p$. Since $\sqrt{(N : M)} = p$, $cx \in radN$ and so $x \in S_p(radN)$. Thus $S_p(N + pM) \subseteq S_p(radN)$. Now if $x \in S_p(radN)$, there exists $c \in R\backslash p$ such that $cx \in P$ and so similar to the process of the proof (i), $x \in P$. Hence we have $x \in radS_p(N)$ and so $S_p(radN) \subseteq radS_p(N)$.

Finally, by [12, Theorem 4.3], $S_p(N+pM)$ is a prime submodule of M and hence $radS_p(N) \subseteq radS_p(N+pM) = S_p(N+pM)$. □

4. Torsioan and radical

Recall that a torsion submodule of a module M over a domain R, denoted by $T(M)$, is the submodule $\{x \in M : ann(x) \neq 0\}$ of M. An R-module M is said to be torsion(resp. torsion-free), if $T(M) = M$(resp. $T(M) = 0$). Compare the following proposition with [8, Lemma 1].

4.1. Lemma. *Let M be an R-module. Let N be a submodule of M satisfying the primeful property. Then $radN$ is a quasi-primary submodule of M if and only if $(N : M)$ is a quasi-primary ideal of R and $T(M/radN) = 0$ as a $R/\sqrt{(N:M)}$-module. In this case $radN$ is a prime submodule of M.*

Proof. Suppose $radN$ is a quasi-primary submodule of M. By Lemma 2.12 (i), $\sqrt{(N:M)} = p$ is a prime ideal of R. If $x + radN \in T(M/radN))$, then $rx \in radN$, for some element $r \in R \setminus p$. Since $radN$ is p-quasi-primary, $x \in radN$ i.e. $T(M/radN) = 0$. Conversely, $\sqrt{(N:M)} = p \neq R$ implies $radN \neq M$. If $rx \in radN$ and $r \notin p$, then $x + radN \in T(M/radN)$ and so $x \in radN$. Thus $radN$ is a quasi-primary submodule of M. In this case $radN$ is prime by Lemma 2.12. □

4.2. Corollary. *Let N be a submodule of an R-module M satisfying the primeful property. If $(N : M)$ is a quasi-primary ideal of R and $T(M/radN) = 0$ as a $R/\sqrt{(N:M)}$-module, then N is a quasi-primary submodule of M.*

Proof. The proof is clear by using Lemma 2.12 and Lemma 4.1. □

4.3. Theorem. *Let R be a Dedekind domain and N a submodule of an R-module M satisfying the primeful property. The following are equivalent:*

(i) *$radN$ is prime;*
(ii) *$M = radN \oplus N'$ for some torsion-free submodule N' of M or $(radN : M) = m$ for some maximal ideal m of R.*

Proof. $(i) \Rightarrow (ii)$. Suppose first that $radN$ is a 0-prime submodule of M. It follows from Lemma 4.1 that $M/radN$ is a torsion-free R-module. It follows from [5, Exercise 19.6(a)] that $M/radN$ is projective and hence $M = radN \oplus N'$ for some submodule N'. Clearly N' is torsion-free. Now, let $radN$ be a prime submodule of M with $(radN : M) \neq 0$. Since R is Dedekind domain, $(radN : M)$ is a maximal ideal of R, as desired.
$(ii) \Rightarrow (i)$. Assume that $M = radN \oplus N'$ for some torsion-free submodule N' of M. Then $M/radN \simeq N'$ follows that $M/radN$ is torsion-free and hence $radN$ is a 0-prime submodule of M by [8, Lemma 1]. On the other hand, it is easy to verify that $radN$ is prime when $(radN : M)$ is a maximal ideal. □

4.4. Theorem. *Let R be a Noetherian domain and M be a non-torsion R-module such that $T(M)$ is contained in only finitely many prime submodules of M. If N is a quasi-primary submodule of M satisfying the primeful property, then $radN$ is prime.*

Proof. We first assume that $(N : M) = 0$. It follows from Corollary 3.11 that $radN$ is a prime submodule of M. Thus we may assume that $(N : M) \neq 0$. If P is a prime submodule containing N, we have the chain $0 = (T(M) : M) \subset \sqrt{(N:M)} \subseteq (P : M)$ of prime ideals of R. If the later containment is proper, by [9, p.144] there are infinitely many prime ideals p with $(N : M) \subset p \subset (P : M)$ and so we have infinitely prime submodules P containing $T(M)$, a contradiction. Hence we have $\sqrt{(N:M)} = (P : M)$, for all prime submodules P containing N. Now if $rx \in radN$ and $x \notin radN$, there is a prime submodule P containing N such that $rx \in P$ and $x \notin P$ and therefore $r \in (P : M) = \sqrt{(N : M)} = (radN : M)$, as required. □

For an R-module M and $x \in M$, we mean that $(N : x)$ is the set $\{r \in R : rx \in N\}$. Now we have the elementary following lemma.

4.5. Lemma. *Let M be an R-module. Then N is a quasi-primary submodule of M if and only if $\sqrt{(N : M)} = \sqrt{(N : x)}$ for all $x \in M \backslash radN$.*

In the following quasi-primary module is considered a module whose the zero submodule is quasi-primary.

4.6. Theorem. *Let M be a quasi-primary and primeful module over a one-dimensional domain R. Then either $\sqrt{ann(M)} = 0$ or $\sqrt{ann(M)} = \sqrt{(N : M)}$ for all proper submodules N of M. In particular, if M is a non-cyclic torsion module, then $\sqrt{(Rx : M)} = \sqrt{ann(x)}$ for all $x \in M \backslash rad0$.*

Proof. Suppose $\sqrt{ann(M)} \neq 0$. Since R is a one-dimensional domain, $\sqrt{ann(M)}$ is a maximal ideal of R. It conclude that $\sqrt{ann(M)} = \sqrt{(N : M)}$ for all proper submodules N. Since 0 is a quasi-primary submodule satisfying the primeful property, $rad0 \neq M$. Now if M is a torsion module, then $\sqrt{ann(M)} \neq 0$. Again since 0 is quasi-primary, $\sqrt{ann(M)} = \sqrt{ann(m)}$ for all $x \in M \backslash rad0$ by Lemma 4.5. Since Rx is a proper submodule for all $x \in M$, by the first part $\sqrt{(Rx : M)} = \sqrt{ann(M)} = \sqrt{ann(x)}$ □

4.7. Theorem. *Let M be a torsion module over a one-dimensional domain R. If M is quasi-primary and primeful, then there exists a prime ideal p of R such that $r \notin p$ implies $rM = M$.*

Proof. Suppose $p = \sqrt{ann(M)}$. If $rM \neq M$, then by Theorem 4.6 $r \in \sqrt{(rM : M)} = \sqrt{ann(M)} = p$. □

4.8. Theorem. *Let M be a quasi-primary primeful and torsion module over a one-dimensional domain R. If $p = \sqrt{ann(M)}$ and M_p is the localization of M at p, then $M/S_p(0) \cong M_p$, an isomorphism of R-modules.*

Proof. Consider the R-module homomorphism $\psi : M \longrightarrow M_p$, given by $m \mapsto m/1$. To show that ψ is an epimorphism, take any $m/s \in M_p$. Since $s \notin p$, $sM = M$ by Theorem 4.8 and so there exists $m' \in M$ such that $m = sm'$. Thus $m/s = sm'/s = m'/1 = \psi(m')$. Also it is easy to verified that the kernel of ψ is $S_p(0)$. Hence $M/S_p(0) \cong M_p$. □

Acknowledgment
The authors express their sincere gratitude to the editor and referee for the careful reading of the original manuscript and useful comments.

References

[1] Behboodi, M. *On the prime radical and bear's lower nilradical of modules*, Acta Math. Hungar., **3**, 293-306, 2009.
[2] Behboodi, M., Aghasi, M. and Sabzevari, M. *A structure sheaf on the spectrum of prime radical modules*, J. Comm. Algebra., Accepted.
[3] Dauns, J. *Prime submodules*, Reine angew. Math., **298**, 156-181, 1978.
[4] Ebrahimi Atani, S. and Yousefian Darani, A., *On quasi-primary submodules*, Chiang Mai J. Sci., **33** (3), 249 - 254, 2006.
[5] Eisenbud, D., *Commutative algebra with a view toward algebraic geometry*, (Springer-Verlag, 1994).

[6] Fazaeli Moghimi, H. and Rashedi, F., *Primary-like submodules satisfying the primeful property*, Transactions on Algebra and its Applications, **1**, 43-54, 2015.
[7] Fuchs, L. *On Quasi-primary ideals*, Acta Sci. Math. (Szeged), 1**1**, 174-183, 1974.
[8] Jenkins, J. and Smith, P. F. *On the prime radical of a module over a commutative Ring*, Comm. Algebra, **20**, 3593-3602, 1992.
[9] Kaplansky, I. *Commutative rings*, (The University of Chicago Press, 1974).
[10] Lu, C.P. *A module whose prime spectrum has the surjective natural map*, Houston J. Math., **33** (1), 125-143, 2007.
[11] Lu, C.P. *Saturations of submodules*, Comm. Algebra, **31** (6), 2655-2673, 2003.
[12] Lu, C.P. *M-radical of submodules in modules II*, Math. Japonica, **35** (5), 991-1001, 1990.
[13] McCasland, R. L. and Moore, M. E. *Prime submodules*, Comm. Algebra, **20** (6), 1803-1817, 1992.
[14] McCasland, R. L., Moore, M. E. and Smith, P. F. *On the spectrum of a module over a commutative ring*, Comm. in Algebra, **25** (1), 79-103, 1997.
[15] McCasland, R. L. and Smith, P. F. *Generalised associated primes and radicals of submodules*, Int. Electron. J. Algebra, **4**, 159-176, 2008.
[16] Moore, M. E. and Smith, S. J. *Prime and radical submodules of modules over commutative rings*, Comm. Alegbra, **30**, 5073-5064, 2002.
[17] Naghipour, R. and Sedghi, M. *Weakly associated primes and primary decomposition of modules over commutative rings*, Acta Math. Hungar. **110** (1-2), 1-12, 2006.
[18] Sharif, H., Sharifi, Y., and Namazi, S. *Rings satisfying the radical formula*, Acta Math. Hungar., **71**(1-2), 103-108, 1996.
[19] Smith, P. F. *Primary modules over commutative rings*, Comm. Algebra, **43** (2001), 103-111.
[20] Pusat-Yilmaz, D. and Smith, P. F. *Radicals of submodules of free modules*, Comm. Algebra, **27** (5), 2253-2266, 1999.
[21] Pusat-Yilmaz, D. and Smith, P. F. *Modules Which Satisfy the Radical Formula*, Acta Math. Hungar., **95** (1-2), 155-167, 2002.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1435 – 1440

Compactness and local compactness of the proximal hyperspace

Marco Rosa*†

Abstract

Compactness and local compactness of the hyperspace endowed with both the Vietoris topology and the Hausdorff metric topology, have been characterized by Costantini, Levi and Pelant. Our aim is to characterize these two properties for the proximal topology, which is related to both of the previous topologies.

Keywords: Hyperspace, Proximal Topology, Local Compactness, Kuratowski Convergence

2000 AMS Classification: Primary 54B20; Secondary 54D30, 54D45

Received : 14.05.2015 *Accepted* : 07.01.2016 *Doi* : 10.15672/HJMS.20164515689

1. Introduction

The first one to characterize compactness of the Vietoris topology on the hyperspace CL(X) of non-empty closed subsets of a topological space X, was Michael in [3]. He also gave a result about local compactness, but it was not correct as remarked in the paper [2]. In that paper, Costantini, Levi and Pelant studied compactness and local compactness of several hyperspace topologies. In particular they characterized compactness and local compactness of CL(X) endowed both with the Vietoris topology τ_V and the Hausdorff metric topology τ_{H_d}.

Following the same spirit and using a similar technique, we characterize compactness and local compactness of CL(X) endowed with the proximal topology $\tau_{\delta(d)}$. We show that both properties are equivalent to compactness of X. The choice of $\tau_{\delta(d)}$ is motivated by the fact that it is deeply connected both to τ_V and τ_{H_d}, because it can be obtained as supremum of the lower Vietoris topology and the upper Hausdorff metric topology, i.e. $\tau_{\delta(d)} = \tau_V^- \vee \tau_{H_d}^+$.

*Dipartimento di Matematica, Informatica ed Economia, Università degli studi della Basilicata, Via dell'Ateneo Lucano 10, 85100 Potenza, Italy. Email: marco.rosa@unibas.it
†Corresponding Author.

2. Preliminaries

Let X be a metrizable space. Given a metric d on X, we denote by D_d the gap between two non-empty closed sets $E, F \in \mathrm{CL}(X)$, defined as:

$$D_d(E,F) = \inf_{x\in E} \inf_{y\in F} d(x,y).$$

Let $x \in X$, we denote by $B_\varepsilon(x)$ the open ball of radius ε and center x. Given $A \in \mathrm{CL}(X)$, we denote by $B_\varepsilon[A]$ the ε-expansion of A, i.e. $B_\varepsilon[A] = \bigcup_{a\in A} B_\varepsilon(a)$. It is easy to check that $B_{\frac{\varepsilon}{2}}[B_{\frac{\varepsilon}{2}}[A]] \subseteq B_\varepsilon[A]$ and $B_{\frac{\varepsilon}{3}}[\overline{B_{\frac{\varepsilon}{3}}[A]}] \subseteq B_\varepsilon[A]$ for every $\varepsilon > 0$ and every $A \in \mathrm{CL}(X)$.

Recall that the Vietoris topology is $\tau_V = \tau_V^- \vee \tau_V^+$, where τ_V^- and τ_V^+ are generated respectively by the collection of all $V^- = \{F \in \mathrm{CL}(X) \mid F \cap V \neq \varnothing\}$ and $W^+ = \{F \in \mathrm{CL}(X) \mid F \subseteq W\}$, when V and W run over all the open subsets of X. The Hausdorff metric topology τ_{H_d} is generated by the Hausdorff distance on $\mathrm{CL}(X)$ induced by d (see for instance [1]). A base for $\tau_{H_d}^+$ is constituted by the collection of all $W^{++} = \{F \in \mathrm{CL}(X) \mid D_d(F, X \setminus W) > 0\} = \{F \in \mathrm{CL}(X) \mid \exists \varepsilon > 0 \ : \ B_\varepsilon[F] \subseteq W\}$, when W runs through the open subsets of X. As recalled before, the proximal topology $\tau_{\delta(d)}$ is the supremum of the lower Vietoris topology and the upper Hausdorff metric topology, i.e. $\tau_{\delta(d)} = \tau_V^- \vee \tau_{H_d}^+$.

Recall that the a net $(C_i)_{i\in I}$ is convergent to C with respect to the Kuratowski convergence if, and only if, it converges with respect to τ_V^- and $C \supseteq \mathrm{Ls}_{i\in I} C_i$ where, denoted by $\mathfrak{U}(x)$ the collection of open neighbourhoods of x,

$$\mathrm{Ls}_{i\in I} C_i = \{x \in X \mid \forall V \in \mathfrak{U}(x) \ \forall i \in I \ \exists j_i \succeq i \ : \ V \cap C_{j_i} \neq \varnothing\} = \bigcap_{i\in I} \overline{\bigcup_{j\succeq i} C_j}.$$

Given a metrizable space X, we denote by $\mathfrak{M}(X)$ the set of all compatible metrics on X.

3. The main result

In the sequel, in the definition of compactness and local compactness we require the space also to be Hausdorff.

Our main result is the following theorem.

3.1. Theorem. *Let X be a metrizable space, let $C \in \mathrm{CL}(X)$ and $d \in \mathfrak{M}(X)$. Then $\mathrm{CL}(X)$ is $\tau_{\delta(d)}$-locally compact at C if, and only if, there exists $\varepsilon > 0$ such that $\overline{B_\varepsilon[C]}$ is compact.*

As a consequence we can characterize both local compactness and compactness of X, using conditions on $\mathrm{CL}(X)$. Moreover we can also characterize compactness of $(\mathrm{CL}(X), \tau_{\delta(d)})$, showing that it is equivalent to local compactness of $(\mathrm{CL}(X), \tau_{\delta(d)})$.

3.2. Corollary. *Let X be a metrizable space. X is locally compact if, and only if, $\mathrm{CL}(X)$ is $\tau_{\delta(d)}$-locally compact at $\{x\}$, for every $x \in X$.*

3.3. Corollary. *Let X be a metrizable space. The following are equivalent:*

(1) *X is compact;*
(2) *$(\mathrm{CL}(X), \tau_{\delta(d)})$ is locally compact;*
(3) *$(\mathrm{CL}(X), \tau_{\delta(d)})$ is locally compact at X.*

Proof. The implications (1) ⇒ (2) and (3) ⇒ (1) follow from Theorem 3.1, while (2) ⇒ (3) is obvious. ■

The following result was proved in [2, Theorem 11].

3.4. Theorem. *Let X be a regular space and let $C \in \mathrm{CL}(X)$. Then $\mathrm{CL}(X)$ is τ_V-locally compact at C if, and only if, there exists an open set $A \subseteq X$ such that $C \subseteq A$ and $\overline{A}$ is compact.*

Combining the previous result and our Theorem 3.1, we obtain as a consequence the equivalence of local compactness and compactness of proximal and Vietoris topologies.

3.5. Corollary. *Let X be a metrizable space. The following are equivalent:*

(1) *X is compact;*
(2) *$(\mathrm{CL}(X), \tau_{\delta(d)})$ is compact;*
(3) *$(\mathrm{CL}(X), \tau_V)$ is compact;*
(4) *$(\mathrm{CL}(X), \tau_{\delta(d)})$ is locally compact;*
(5) *$(\mathrm{CL}(X), \tau_V)$ is locally compact;*

Proof. The equivalence between (1) $\Leftrightarrow$ (3) $\Leftrightarrow$ (5) has been proved in [2, Corollary 13]. If X is compact, then $(\mathrm{CL}(X), \tau_{\delta(d)})$ is compact since $\tau_{\delta(d)} = \tau_V$ and this proves the implication (1) $\Rightarrow$ (2). The implication (2) $\Rightarrow$ (4) is obvious. Finally the equivalence (1) $\Leftrightarrow$ (4) follows from Corollary 3.3. ∎

3.6. Remark. Note that (4) $\Rightarrow$ (5) can be easily proved in a direct way, in order to explicitly use the condition that characterize local compactness of $\tau_{\delta(d)}$ and τ_V. Indeed, if $(\mathrm{CL}(X), \tau_{\delta(d)})$ is locally compact at C, then by Theorem 3.1 there exists $\varepsilon > 0$ such that $\overline{B_\varepsilon[C]}$ is compact. Then $V = B_\varepsilon[C]$ is an open set containing C and with compact closure. Hence $(\mathrm{CL}(X), \tau_V)$ is locally compact at C by Theorem 3.4.

In [2] it has been proved that local compactness of the Vietoris hyperspace is in general a strictly stronger condition than local compactness of the Hausdorff hyperspace. As a consequence of Corollary 3.5, local compactness of the proximal hyperspace has the same behaviour. Finally it has been proved in [1, Theorem 3.2.4] that compactness of the Hausdorff hyperspace is equivalent to compactness of X, and by Corollary 3.5, this is equivalent to compactness of the proximal hyperspace.

4. Proof our main result

To prove our main theorem we need several preliminary results. The following remarks are of easy verification.

4.1. Remark. Let X be a metrizable space and let $d \in \mathfrak{M}(X)$. If C is closed, then C^+ is $\tau_{\delta(d)}$-closed. If K is compact, then $(X \setminus K)^{++} = (X \setminus K)^+$ and therefore $(X \setminus K)^+$ is $\tau_{\delta(d)}$-open and K^- is $\tau_{\delta(d)}$-closed.

Proof. Since $X \setminus C$ is open, $\mathrm{CL}(X) \setminus C^+ = (X \setminus C)^-$ is $\tau_{\delta(d)}$-open.

If K is compact and F is closed, then $F \cap K = \varnothing$ implies $D_d(F, K) > 0$. Therefore $(X \setminus K)^+ = \{F \mid F \subseteq X \setminus K\} = \{F \mid D_d(F, K) > 0\} = (X \setminus K)^{++}$. Moreover $\mathrm{CL}(X) \setminus K^- = (X \setminus K)^+ = (X \setminus K)^{++}$ which is $\tau_{\delta(d)}$-open. ∎

4.2. Remark. Let X be a metrizable space and let $d \in \mathfrak{M}(X)$. Then $(\mathrm{CL}(X), \tau_{\delta(d)})$ is T_2.

Proof. Let $A, C \in \mathrm{CL}(X)$ such that $A \neq C$. We may suppose there exists $a \in A \setminus C$. Since $a \notin \bigcap_{\varepsilon > 0} \overline{B_\varepsilon[C]} = C$, there exists $\varepsilon > 0$ such that $a \in X \setminus \overline{B_\varepsilon[C]}$. Take $\delta > 0$ such that $B_\delta(a) \subseteq X \setminus \overline{B_\varepsilon[C]}$. We claim that $B_\varepsilon[C] \subseteq X \setminus \overline{B_{\frac{\delta}{2}}(a)}$. Indeed let on the contrary $x \in B_\varepsilon[C] \cap \overline{B_{\frac{\delta}{2}}(a)}$. There exists $y \in B_{\frac{\delta}{2}}(x) \cap B_{\frac{\delta}{2}}(a)$, hence $d(x, a) \leq d(x, y) + d(y, a) < \delta$. Then $x \in B_\delta(a) \subseteq X \setminus \overline{B_\varepsilon[C]}$, impossible.

Hence $B_\varepsilon[C] \subseteq X \setminus \overline{B_{\frac{\delta}{2}}(a)}$, that is $C \in (X \setminus \overline{B_{\frac{\delta}{2}}(a)})^{++}$ which is $\tau_{\delta(d)}$-open. Moreover $A \in B_{\frac{\delta}{2}}(a)^-$ which is $\tau_{\delta(d)}$-open and $B_{\frac{\delta}{2}}(a)^- \cap (X \setminus \overline{B_{\frac{\delta}{2}}(a)})^{++} \subseteq B_{\frac{\delta}{2}}(a)^- \cap (X \setminus \overline{B_{\frac{\delta}{2}}(a)})^+ \subseteq B_{\frac{\delta}{2}}(a)^- \cap (X \setminus B_{\frac{\delta}{2}}(a))^+ = \varnothing$. ■

The following result gives a sufficient condition for compactness of a collection $\mathcal{K} \subseteq \mathrm{CL}(X)$.

4.3. Proposition. *Let X be a metrizable space. Let $\mathcal{K} \subseteq \mathrm{CL}(X)$ and $d \in \mathfrak{M}(X)$. If $\mathcal{K}$ is $\tau_{\delta(d)}$-closed, and for every $F \in \mathrm{CL}(X)$, for every $\varepsilon > 0$ and for every open cover $\mathcal{U}$ of $\overline{B_\varepsilon[F]}$, there exists a finite open subcover $\mathcal{F}$ such that*

$$\mathcal{K} \cap \overline{B_\varepsilon[F]}^- \subseteq \bigcup_{U \in \mathcal{F}} U^-,$$

then $\mathcal{K}$ is $\tau_{\delta(d)}$-compact.

Proof. Let $(C_j)_{j\in J}$ be a net in $\mathcal{K}$. By Remark 4.2, we have to prove that it has a convergent subnet. By [1, Theorem 5.2.11], there exists a subnet $(C_{j_i})_{i\in I}$ which is K-convergent to a set $C \in \mathrm{CL}(X) \cup \{\varnothing\}$. Note that since $\{\varnothing\} = \varnothing^{++}$, $\varnothing$ is isolated with respect to $\tau_{\delta(d)}$ and therefore $C \in \mathrm{CL}(X)$. Moreover $(C_{j_i})_{i\in I}$ is τ_V^--convergent to C. We want to prove that $C_{j_i} \to C$ with respect to $\tau_{H_d}^+$, and this would also imply that $C \in \mathcal{K}$ since $\mathcal{K}$ is $\tau_{\delta(d)}$-closed.

On the contrary, suppose there exists W open such that $C \in W^{++}$ but $C_{j_i} \notin W^{++}$ frequently. Since $D_d(C, X \setminus W) > 0$, there exists $\varepsilon > 0$ such that $\overline{B_\varepsilon[X \setminus W]} \subseteq X \setminus C$. Since $C \supseteq \mathrm{Ls}_{i\in I} C_{j_i}$, for every $x \in \overline{B_\varepsilon[X \setminus W]} \subseteq X \setminus C \subseteq X \setminus \mathrm{Ls}_{i\in I} C_{j_i}$, there exist a neighbourhood V_x of x, and a index $i_x \in I$ such that $V_x \cap C_{j_i} = \varnothing$ for every $i \geq i_x$. Since $\overline{B_\varepsilon[X \setminus W]} \subseteq \bigcup_{x \in \overline{B_\varepsilon[X\setminus W]}} V_x$, by hypothesis there exist $x_1, \dots, x_n \in \overline{B_\varepsilon[X \setminus W]}$ such that

$$\mathcal{K} \cap \overline{B_\varepsilon[X \setminus W]}^- \subseteq \bigcup_{k=1}^n V_{x_k}^-.$$

Let $i_0 \in I$ such that $i_0 \geq i_{x_k}$ for $k = 1, \dots, n$. Then for every $i \geq i_0$, $C_{j_i} \cap V_{x_k} = \varnothing$ for every $k = 1, \dots, n$, that is $C_{j_i} \notin \bigcup_{k=1}^n V_{x_k}^-$. Since $C_{j_i} \in \mathcal{K}$, then $C_{j_i} \notin \overline{B_\varepsilon[X \setminus W]}^-$ for every $i \geq i_0$, that is $C_{j_i} \cap \overline{B_\varepsilon[X \setminus W]} = \varnothing$.

On the other hand, $C_{j_i} \notin W^{++}$ frequently, so that there exists $k \geq i_0$, such that for every $\delta > 0$, $B_\delta[C_{j_k}] \cap (X \setminus W) \neq \varnothing$. In particular for $\delta = \varepsilon$, there exist $y \in C_{j_k}$, $z \in X \setminus W$ such that $d(z, y) < \varepsilon$. But then $y \in C_{j_k} \cap B_\varepsilon[X \setminus W] \neq \varnothing$, a contradiction. ■

4.4. Lemma. *Let X be a metrizable space. Let $K \in \mathrm{CL}(X)$ and $d \in \mathfrak{M}(X)$. If K is compact and there exists $\delta > 0$ such that $\overline{B_\delta(x)}$ is compact for every $x \in K$, then $\overline{B_{\frac{\delta}{2}}[K]}$ is compact.*

Proof. Of course $K \subseteq \bigcup_{x\in K} B_{\frac{\delta}{2}}(x)$. Since K is compact there exist $x_1, \dots, x_n \in K$ such that $K \subseteq \bigcup_{i=1}^n B_{\frac{\delta}{2}}(x_i)$. Then $B_{\frac{\delta}{2}}[K] \subseteq \bigcup_{i=1}^n B_\delta(x_i)$ and therefore $\overline{B_{\frac{\delta}{2}}[K]} \subseteq \overline{\bigcup_{i=1}^n B_\delta(x_i)} = \bigcup_{i=1}^n \overline{B_\delta(x_i)}$ which is compact. Hence $\overline{B_{\frac{\delta}{2}}[K]}$ is compact. ■

4.5. Lemma. *Let X be a metrizable space. Let $K \in \mathrm{CL}(X)$ and $d \in \mathfrak{M}(X)$. If K is compact and there exists $\delta > 0$ such that $\overline{B_\delta(x)}$ is compact for every $x \in K$, then $\overline{B_{\frac{\delta}{2}}[K]}^+$ is $\tau_{\delta(d)}$-compact.*

Proof. The set $\overline{B_{\frac{\delta}{2}}[K]}^+$ is $\tau_{\delta(d)}$-closed by Remark 4.1. Let $C \in \mathrm{CL}(X)$, $\varepsilon > 0$ and let $\mathcal{U}$ be an open cover of $\overline{B_\varepsilon[C]}$. By Proposition 4.3, we have to find a finite open subcover $\mathcal{F}$

such that $\overline{B_{\frac{\delta}{2}}[K]}^+ \cap \overline{B_\varepsilon[C]}^- \subseteq \bigcup_{U\in\mathcal{F}} U^-$. If $\overline{B_\varepsilon[C]} \cap \overline{B_{\frac{\delta}{2}}[K]} = \varnothing$ then $\overline{B_\varepsilon[C]}^- \cap \overline{B_{\frac{\delta}{2}}[K]}^+$ is empty. Suppose $\overline{B_\varepsilon[C]} \cap \overline{B_{\frac{\delta}{2}}[K]} \neq \varnothing$. Note that $\overline{B_\varepsilon[C]} \cap \overline{B_{\frac{\delta}{2}}[K]}$ is compact since it is closed and it is contained in $\overline{B_{\frac{\delta}{2}}[K]}$ which is compact by Lemma 4.4. There exist $U_1,\dots,U_n \in \mathcal{U}$ such that $\overline{B_\varepsilon[C]} \cap \overline{B_{\frac{\delta}{2}}[K]} \subseteq \bigcup_{i=1}^n U_i$. Therefore $\overline{B_\varepsilon[C]}^- \cap \overline{B_{\frac{\delta}{2}}[K]}^+ \subseteq (\overline{B_\varepsilon[C]} \cap \overline{B_{\frac{\delta}{2}}[K]})^- \subseteq \bigcup_{i=1}^n U_i^-$. ■

4.6. Proposition. *Let X be a metrizable space and $d \in \mathfrak{M}(X)$. Let $V, V_1, \dots V_k$ be any non-empty set. If $\mathcal{V} = \bigcap_{i=1}^k V_i^- \cap V^+$ is $\tau_{\delta(d)}$-compact and non-empty, then V is compact.*

Proof. Since $\mathcal{V} \neq \varnothing$, for every $i = 1,\dots,k$ we can fix $x_i \in V_i \cap V$.

We first prove that V is closed. Otherwise there should exist $(y_n)_{n\in\mathbb{N}}$ such that $y_n \in V$ and $y_n \to y \in X \smallsetminus V$. Since $\mathcal{V} \neq \varnothing$, for every $i = 1,\dots,k$ we can find $x_i \in V_i \cap V$. For every $n \in \mathbb{N}$ set $C_n = \{x_1,\dots,x_k,y_n\}$ and $C = \{x_1,\dots,x_k,y\}$. Note that $C \notin \mathcal{V}$ because $y \in C \smallsetminus V$. Moreover $C_n \in \mathcal{V}$ for every $n \in \mathbb{N}$. We prove that $C_n \to_{\tau_{\delta(d)}} C$ in order to have a contradiction, since $\mathcal{V}$ is $\tau_{\delta(d)}$-compact and hence closed.

Let $\mathcal{U} = \bigcap_{i=1}^p U_i^- \cap U^{++}$ be a $\tau_{\delta(d)}$-neighbourhood of C. Let $i \in \{1,\dots,p\}$; we distinguish two cases. If there exists $j \in \{1,\dots,k\}$ such that $x_j \in U_i$, then $x_j \in C_n \cap U_i$, hence $C_n \in U_i^-$ for every $n \in \mathbb{N}$. If $x_j \notin U_i$ for every $j = 1,\dots,k$, since $C \cap U_i \neq 0$, then $y \in C \cap U_i$. Take $\varepsilon > 0$ such that $B_\varepsilon(y) \in U_i$. Then $y_n \in B_\varepsilon(y)$ eventually and therefore $C_n \in U_i^-$ eventually. Since $C \in U^{++}$, there exists $\delta > 0$ such that $B_\delta[C] \subseteq U$. Eventually $y_n \in B_{\frac{\delta}{2}}(y)$ and therefore $B_{\frac{\delta}{2}}[C_n] \subseteq B_\delta[C] \subseteq U$, that is $C_n \in U^{++}$. Hence $C_n \in \bigcap_{i=1}^p U_i^- \cap U^{++}$ eventually, that is $C_n \to_{\tau_{\delta(d)}} C$.

We now prove that V is compact. Otherwise, there should exist $(a_n)_{n\in\mathbb{N}}$ in V with no cluster point. That is for every $x \in X$ there exists a neighbourhood V_x of x and $\nu_x \in \mathbb{N}$ such that for every $n \geq \nu_x$, $a_n \notin V_x$. For every $n \in \mathbb{N}$ set $C_n = \{x_1,\dots,x_k,a_n\}$. We will prove that $(C_n)_{n\in\mathbb{N}}$ has no cluster point in order to have a contradiction since $\mathcal{V}$ is compact. Let $C \in \mathrm{CL}(X)$.

- If $C \subseteq \{x_1,\dots,x_k\}$, then $\left(\bigcup_{i=1}^k V_{x_i}\right)^{++}$ is a $\tau_{\delta(d)}$-neighbourhood of C. On the other hand if $\nu \geq \max\{\nu_{x_1},\dots,\nu_{x_k}\}$, then for every $n \geq \nu$, $a_n \notin V_{x_i}$ for every $i = 1,\dots,k$. Then $C_n \notin \left(\bigcup_{i=1}^k V_{x_i}\right)^{++}$ eventually.
- If there exists $x_0 \in C \smallsetminus \{x_1,\dots,x_k\}$, let W be a neighbourhood of x_0 such that $W \cap \{x_1,\dots,x_k\} = \varnothing$. Then $C_n \notin (W \cap V_{x_0})^-$ for every $n \geq \nu_{x_0}$, while $C \in (W \cap V_{x_0})^-$. ■

We are now able to finally prove our main result.

Proof of Theorem 3.1.

$\Rightarrow$) Suppose that $(\mathrm{CL}(X), \tau_{\delta(d)})$ is locally compact at C. There exists a neighbourhood $\mathcal{V} = \bigcap_{i=1}^n V_i^- \cap V^{++}$ of C such that $\overline{\mathcal{V}}$ is $\tau_{\delta(d)}$-compact. For every $i = 1,\dots,n$, let $x_i \in V_i \cap C$ and let $\delta > 0$ such that $B_\delta[C] \subseteq V$. Set $\mathcal{U} = \bigcap_{i=1}^n \{x_i\}^- \cap \overline{B_{\frac{\delta}{3}}[C]}^+$. Note that $\mathcal{U}$ is $\tau_{\delta(d)}$-closed by Remark 4.1. If $T \in \mathcal{U}$, then $B_{\frac{\delta}{3}}[T] \subseteq B_{\frac{\delta}{3}}(\overline{B_{\frac{\delta}{3}}[C]}) \subseteq B_\delta[C] \subseteq V$, and this implies $\mathcal{U} \subseteq \mathcal{V}$. Hence $\mathcal{U}$ is $\tau_{\delta(d)}$-compact being $\tau_{\delta(d)}$-closed and contained in the $\tau_{\delta(d)}$-compact set $\overline{\mathcal{V}}$. By Proposition 4.6, $\overline{B_{\frac{\delta}{3}}[C]}$ is compact.

$\Leftarrow$) Since C is closed and contained in the compact set $\overline{B_\varepsilon[C]}$, it is compact. Moreover for every $x \in C$, $\overline{B_\varepsilon(x)}$ is compact since it is contained in $\overline{B_\varepsilon[C]}$. By Lemma 4.5, $\overline{B_{\frac{\varepsilon}{2}}[C]}^+$ is $\tau_{\delta(d)}$-compact. Moreover $C \in B_{\frac{\varepsilon}{2}}[C]^{++} \subseteq \overline{B_{\frac{\varepsilon}{2}}[C]}^+$. The set $B_{\frac{\varepsilon}{2}}[C]^{++}$ is a $\tau_{\delta(d)}$-neighbourhood of C and its closure in $\mathrm{CL}(X)$ is compact since it is contained in $\overline{B_{\frac{\varepsilon}{2}}[C]}^{+}$, which is $\tau_{\delta(d)}$-compact. Hence $(\mathrm{CL}(X), \tau_{\delta(d)})$ is locally compact at C. ■

References

[1] Beer, G.: *Topologies on closed and closed convex sets*, Mathematics and its Applications, No. 268, Kluwer Academic Publishers Group, Dordrecht, 1993.

[2] Constantini, C.—Levi, S.—Pelant, J.: *Compactness and local compactness in hyperspaces*, Topology Appl. **123**(3) (2002), 573–608.

[3] Michael, E.: *Topologies on spaces of subsets*, Trans. Amer. Math. Soc. **71** (1951), 152–182.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1441 – 1448

On the spectral norms of some special g-circulant matrices

Shou-Qiang Shen *†, Wei-Jun Liu‡ and Jun-Jie He§

Abstract

In the present paper, we give upper and lower bounds for the spectral norm of g-circulant matrix, whose the first row entries are the classical Horadam numbers $U_i^{(a,b)}$. In addition, we also establish an explicit formula of the spectral norm for g-circulant matrix with the first row $([U_0^{(a,b)}]^2, [U_1^{(a,b)}]^2, \cdots, [U_{n-1}^{(a,b)}]^2)$.

Keywords: g-Circulant matrix; Spectral norm; Horadam number; Fibonacci number; Lucas number

2000 AMS Classification: 15A45, 15A60.

Received : 08.01.2015 *Accepted* : 01.03.2016 *Doi* : 10.15672/HJMS.20164515690

1. Introduction and Preliminaries

Many generalizations of the Fibonacci and Lucas sequences have been introduced and studied [1-4]. Here we use the classical Horadam sequence $\{U_n^{(a,b)}\}_{n\in N}$, which is defined in [4]:

$$U_n^{(a,b)} = AU_{n-1}^{(a,b)} + BU_{n-2}^{(a,b)}, \quad U_0^{(a,b)} = a, \quad U_1^{(a,b)} = b, \tag{1.1}$$

where $a, b \in R$ and $A^2 + 4B > 0$. Obviously, if we choose $A = B = 1$ in (1), then the generalized Fibonacci sequence $\{F_n^{(a,b)}\}_{n\in N}$ is obtained. Further more, when $a = 0, b = 1$ and $a = 2, b = 1$, the sequence $\{F_n^{(a,b)}\}_{n\in N}$ reduces to the well-known Fibonacci sequence $\{F_n\}_{n\in N}$ and Lucas sequence $\{L_n\}_{n\in N}$, respectively.

For the Horadam sequence $\{U_n^{(a,b)}\}_{n\in N}$, the following generalization of the Binet's formula of Fibonacci number holds [4]:

$$U_n^{(a,b)} = c_1\alpha^n + c_2\beta^n, \tag{1.2}$$

*Supported by the Key Research Project of Higher school in Henan Province(No. 15A110040).

†School of Mathematics and Statistics, Central South University, 410083 Changsha, China. Corresponding author,
Email : shenshouqiang@126.com

‡School of Mathematics and Statistics, Central South University, 410083 Changsha, China.

§Department of Mathematics, Xinyang Normal University, 464000 Xinyang, China.

where

$$c_1 = \frac{a(A^2+4B)+(2b-aA)\sqrt{A^2+4B}}{2(A^2+4B)}, \tag{1.3}$$
$$c_2 = \frac{a(A^2+4B)-(2b-aA)\sqrt{A^2+4B}}{2(A^2+4B)},$$

$$\alpha = \frac{A+\sqrt{A^2+4B}}{2}, \beta = \frac{A-\sqrt{A^2+4B}}{2}. \tag{1.4}$$

Recently, there has been much interest in investigation of some special matrices. Akbulak and Bozkurt [5] found the lower and upper bounds for the spectral norms of Toeplitz matrices $\mathcal{A} = [F_{i-j}]_{i,j=1}^{n}$ and $\mathcal{B} = [L_{i-j}]_{i,j=1}^{n}$, then Shen [6] generalized these results. Solak [7,8] gave the upper and lower bounds for the spectral norms of circulant matrices whose entries are Fibonacci and Lucas numbers. Then İpek [9] investigated an improved estimation for the spectral norms of these matrices. In addition, there have been several articles focus on the spectral distribution and norms of g-circulant matrices. Bose et al. [10] listed the limiting spectral distribution for a class of g-circulant matrices with heavy tailed input sequence. Zhou and Jiang [11] derived some explicit formulas for the spectral norms of g-circulant matrices whose the first row entries are Fibonacci number, Lucas number and their powers.

Besides, Shen et al. [12] gave some feasible computational formulas for the determinants and inverses of the circulant matrices $\mathcal{A}_n$=Circ$(F_1, F_2, \cdots, F_n)$ and $\mathcal{B}_n$=Circ$(L_1, L_2, \cdots, L_n)$, then Yazlik and Taskara [13] generalized all results from [12]. Stanimirović et al. [4] defined an $n \times n$ Toeplitz matrix $\mathcal{U}_n^{(a,b,s)} = [u_{i,j}^{(a,b,s)}](i, j = 1, 2, \cdots, n)$ of type s, where

$$u_{i,j}^{(a,b,s)} = \begin{cases} U_{i-j+1}^{(a,b)}, & i-j+s \geq 0, \\ 0, & i-j+s < 0. \end{cases} \tag{1.5}$$

then the inverse of the matrix $\mathcal{U}_n^{(a,b,0)}$ was derived, and correlations between the matrix $\mathcal{U}_n^{(a,b,0)}$ and the generalized Pascal matrices of the first and the second kinds were considered. In addition, Shen and He [14] also established an explicit formula of the Moore-Penrose inverse for the matrix $\mathcal{U}_n^{(a,b,-1)}$.

In this paper, let $\mathcal{A}_U$ and $\mathcal{A}_{U^2}$ be two g-circulant matrices, whose the first row entries are $(U_0^{(a,b)}, U_1^{(a,b)}, \cdots, U_{n-1}^{(a,b)})$ and $([U_0^{(a,b)}]^2, [U_1^{(a,b)}]^2, \cdots, [U_{n-1}^{(a,b)}]^2)$, respectively. We give upper and lower bounds for the spectral norm of matrix $\mathcal{A}_U$, and establish an explicit formula of the spectral norm for matrix $\mathcal{A}_{U^2}$, then generalize the main results in [11].

Now we give some preliminaries related to our study. A matrix $\mathcal{A} \in M_n$ is called a g-circulant matrix if it is of the form

$$\mathcal{A} = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-g} & a_{n-g+1} & a_{n-g+2} & \cdots & a_{n-g-1} \\ a_{n-2g} & a_{n-2g+1} & a_{n-2g+2} & \cdots & a_{n-2g-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_g & a_{g+1} & a_{g+2} & \cdots & a_{g-1} \end{pmatrix} \tag{1.6}$$

where g is a nonnegative integer and each of the subscripts is understood to be reduced modulo n. Obviously, when $g = 1$ or $g = n+1$, the g-circulant matrix $\mathcal{A}$ reduces to the standard circulant matrix.

For any $\mathcal{A} = [a_{ij}] \in M_{m,n}$. The well-known Frobenius (or Euclidean) norm of matrix $\mathcal{A}$ is

$$\|\mathcal{A}\|_F = \left[\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right]^{\frac{1}{2}}$$

and also the spectral norm of matrix $\mathcal{A}$ is

$$\|\mathcal{A}\|_2 = \sqrt{\max_{1\le i\le n}\lambda_i(\mathcal{A}^H\mathcal{A})}$$

where $\lambda_i(\mathcal{A}^H\mathcal{A})$ is eigenvalue of $\mathcal{A}^H\mathcal{A}$ and $\mathcal{A}^H$ is conjugate transpose of matrix $\mathcal{A}$. Then the following inequality holds:

$$\frac{1}{\sqrt{n}}\|\mathcal{A}\|_F \le \|\mathcal{A}\|_2 \le \|\mathcal{A}\|_F \tag{1.7}$$

Lemma 1[15] *An $n\times n$ matrix Q_g is unitary if and only if $(n, g) = 1$, where Q_g is a g-circulant matrix with the first row $(1, 0, \cdots, 0)$.*

Lemma 2[15] *$\mathcal{A}$ is a g-circulant matrix with the first row $(a_0, a_1, \cdots, a_{n-1})$ if and only if $\mathcal{A} = Q_gC$, where C is a circulant matrix with the first row $(a_0, a_1, \cdots, a_{n-1})$.*

Lemma 3[16] *Let $\mathcal{A} = [a_{ij}] \in M_n$ is a nonnegative matrix, Then its spectral radius $\rho(\mathcal{A})$ satisfies the following inequality*

$$\min_{1\le i\le n}\sum_{j=1}^{n}a_{ij} \le \rho(\mathcal{A}) \le \max_{1\le i\le n}\sum_{j=1}^{n}a_{ij}. \tag{1.8}$$

Lemma 4 *For the Horadam sequence $\{U_n^{(a,b)}\}_{n\in N}$ satisfying $B \ne -1$ and $B\pm A \ne 1$, the following identity is valid*:

$$\sum_{i=0}^{n-1}[U_i^{(a,b)}]^2 = \frac{M-[U_n^{(a,b)}]^2+B^2[U_{n-1}^{(a,b)}]^2}{(1-B)^2-A^2} + \frac{2(a^2B+abA-b^2)[1-(-B)^n]}{(1+B)(A^2+4B)}, \tag{1.9}$$

where $M = a^2-(aA-b)^2-\frac{2(1+B)(a^2B+abA-b^2)[1-(-B)^n]}{A^2+4B}$.

Proof: From $B\pm A \ne 1$, we get $\alpha \ne \pm 1$ and $\beta \ne \pm 1$, applying identities $U_n^{(a,b)} = c_1\alpha^n + c_2\beta^n$ and $\alpha\beta = -B$, then the following is valid

$$\begin{aligned}\sum_{i=0}^{n-1}[U_i^{(a,b)}]^2 &= \sum_{i=0}^{n-1}(c_1\alpha^i+c_2\beta^i)^2 = c_1^2\sum_{i=0}^{n-1}\alpha^{2i}+c_2^2\sum_{i=0}^{n-1}\beta^{2i}+2c_1c_2\sum_{i=0}^{n-1}(\alpha\beta)^i\\ &= c_1^2\cdot\frac{1-\alpha^{2n}}{1-\alpha^2}+c_2^2\cdot\frac{1-\beta^{2n}}{1-\beta^2}+2c_1c_2\cdot\frac{1-(\alpha\beta)^n}{1-\alpha\beta}\\ &= \frac{c_1^2+c_2^2-(c_2^2\alpha^2+c_1^2\beta^2)-(c_1^2\alpha^{2n}+c_2^2\beta^{2n})+(\alpha\beta)^2(c_1^2\alpha^{2n-2}+c_2^2\beta^{2n-2})}{(1-\alpha^2)(1-\beta^2)}\\ &\quad +2c_1c_2\frac{1-(-B)^n}{1+B}.\end{aligned}$$

By using identities $\alpha+\beta = A$ and $\alpha-\beta = \sqrt{A^2+4B}$, we have

$$c_1 = \frac{a}{2}+\frac{2b-aA}{2\sqrt{A^2+4B}} = \frac{a}{2}+\frac{2b-a(\alpha+\beta)}{2(\alpha-\beta)} = \frac{b-a\beta}{\alpha-\beta},$$

$$c_2 = \frac{a}{2}-\frac{2b-aA}{2\sqrt{A^2+4B}} = \frac{a}{2}-\frac{2b-a(\alpha+\beta)}{2(\alpha-\beta)} = \frac{a\alpha-b}{\alpha-\beta}.$$

So we obtain

$$\begin{aligned} c_2^2\alpha^2 + c_1^2\beta^2 &= (c_2\alpha + c_1\beta)^2 - 2c_1c_2\alpha\beta = \left(\frac{a\alpha - b}{\alpha - \beta}\cdot\alpha + \frac{b - a\beta}{\alpha - \beta}\cdot\beta\right)^2 - 2c_1c_2\alpha\beta \\ &= [a(\alpha+\beta) - b]^2 - 2c_1c_2\alpha\beta = (aA - b)^2 + 2c_1c_2B. \end{aligned}$$

Since $c_1^2 + c_2^2 = a^2 - 2c_1c_2$ and

$$c_1^2\alpha^{2n} + c_2^2\beta^{2n} = (c_1\alpha^n + c_2\beta^n)^2 - 2c_1c_2(\alpha\beta)^n = [U_n^{(a,b)}]^2 - 2c_1c_2(-B)^n,$$

$$(\alpha\beta)^2(c_1^2\alpha^{2n-2} + c_2^2\beta^{2n-2}) = B^2[U_{n-1}^{(a,b)}]^2 - 2c_1c_2(-B)^{n+1}.$$

While $c_1c_2 = \frac{a^2B+abA-b^2}{A^2+4B}$, hence

$$\begin{aligned} \sum_{i=0}^{n-1}[U_i^{(a,b)}]^2 &= \frac{a^2 - (aA-b)^2 - 2c_1c_2(1+B)[1-(-B)^n] - [U_n^{(a,b)}]^2 + B^2[U_{n-1}^{(a,b)}]^2}{(1+\alpha\beta)^2 - (\alpha+\beta)^2} \\ &\quad + 2c_1c_2\frac{1-(-B)^n}{1+B} \\ &= \frac{a^2 - (aA-b)^2 - \frac{2(1+B)(a^2B+abA-b^2)[1-(-B)^n]}{A^2+4B} - [U_n^{(a,b)}]^2 + B^2[U_{n-1}^{(a,b)}]^2}{(1-B)^2 - A^2} \\ &\quad + \frac{2(a^2B+abA-b^2)[1-(-B)^n]}{(1+B)(A^2+4B)}. \end{aligned}$$

Thus the proof is completed. □

2. Main Results

Theorem 1 *Let $\mathcal{A}_U$ be as the matrix in (1.6), with $a_i = U_i^{(a,b)} (i = 0, 1, \cdots, n-1)$ in the first row of $\mathcal{A}_U$. If $B \neq -1$, $B \pm A \neq 1$ and $(n, g) = 1$, then we have*

$$\begin{aligned} \sqrt{\frac{M - [U_n^{(a,b)}]^2 + B^2[U_{n-1}^{(a,b)}]^2}{(1-B)^2 - A^2} + \frac{2(a^2B+abA-b^2)[1-(-B)^n]}{(1+B)(A^2+4B)}} &\leq \|\mathcal{A}_U\|_2 \\ &\leq \frac{1}{\sqrt{A^2+4B}}\left[\frac{|b - a\beta|(1-|\alpha|^n)}{1-|\alpha|} + \frac{|b-a\alpha|(1-|\beta|^n)}{1-|\beta|}\right], \end{aligned}$$

where $\alpha = \frac{A+\sqrt{A^2+4B}}{2}$, $\beta = \frac{A-\sqrt{A^2+4B}}{2}$ and $M = a^2-(aA-b)^2-\frac{2(1+B)(a^2B+abA-b^2)[1-(-B)^n]}{A^2+4B}$.

Proof: Applying the definition of Frobenius norm and formula (1.9), we have

$$\begin{aligned} \|\mathcal{A}_U\|_F^2 &= n\sum_{i=0}^{n-1}[U_i^{(a,b)}]^2 \\ &= n\left(\frac{M - [U_n^{(a,b)}]^2 + B^2[U_{n-1}^{(a,b)}]^2}{(1-B)^2 - A^2} + \frac{2(a^2B+abA-b^2)[1-(-B)^n]}{(1+B)(A^2+4B)}\right), \end{aligned}$$

where $M = a^2 - (aA-b)^2 - \frac{2(1+B)(a^2B+abA-b^2)[1-(-B)^n]}{A^2+4B}$. Hence from (1.7), we obtain

$$\begin{aligned} \|\mathcal{A}_U\|_2 &\geq \frac{1}{\sqrt{n}}\|\mathcal{A}_U\|_F \\ &= \sqrt{\frac{M - [U_n^{(a,b)}]^2 + B^2[U_{n-1}^{(a,b)}]^2}{(1-B)^2 - A^2} + \frac{2(a^2B+abA-b^2)[1-(-B)^n]}{(1+B)(A^2+4B)}}. \end{aligned}$$

On the other hand, using the results from Lemma 1 and Lemma 2, one can verify

$$(\mathcal{A}_U)^H\mathcal{A}_U = (Q_gC)^HQ_gC = C^H(Q_g)^HQ_gC = C^HI_nC = C^HC,$$

where C is a circulant matrix with the first row $(U_0^{(a,b)}, U_1^{(a,b)}, \cdots, U_{n-1}^{(a,b)})$ and I_n is an identity matrix. Hence the spectral norm of matrix $\mathcal{A}_U$ is the same as that of C. Let $f(x) = \sum_{i=0}^{n-1} U_i^{(a,b)} x^i$ be a scalar-valued polynomial, and π_n be an $n \times n$ circulant matrix with the first row $(0, 1, \cdots, 0)$, then we get

$$C = f(\pi_n) = \sum_{i=0}^{n-1} U_i^{(a,b)} \pi_n^i,$$

hence

$$\|\mathcal{A}_U\|_2 = \|C\|_2 = \| \sum_{i=0}^{n-1} U_i^{(a,b)} \pi_n^i \|_2 \leq \sum_{i=0}^{n-1} \|U_i^{(a,b)} \pi_n^i\|_2 \leq \sum_{i=0}^{n-1} |U_i^{(a,b)}| \|\pi_n\|_2^i.$$

Since $\pi_n^H \pi_n = I_n$, then we have

$$\|\pi_n\|_2 = \sqrt{\max_{1 \leq i \leq n} \lambda_i(\pi_n^H \pi_n)} = 1.$$

Note that $|\alpha| \neq 1$ and $|\beta| \neq 1$, hence we obtain

$$\begin{aligned} \|\mathcal{A}_U\|_2 &\leq \sum_{i=0}^{n-1} |U_i^{(a,b)}| = \sum_{i=0}^{n-1} |c_1 \alpha^i + c_2 \beta^i| \leq |c_1| \sum_{i=0}^{n-1} |\alpha|^i + |c_2| \sum_{i=0}^{n-1} |\beta|^i \\ &= |c_1| \frac{1 - |\alpha|^n}{1 - |\alpha|} + |c_2| \frac{1 - |\beta|^n}{1 - |\beta|} = \frac{1}{\alpha - \beta} \left[\frac{|b - a\beta|(1 - |\alpha|^n)}{1 - |\alpha|} + \frac{|b - a\alpha|(1 - |\beta|^n)}{1 - |\beta|} \right] \\ &= \frac{1}{\sqrt{A^2 + 4B}} \left[\frac{|b - a\beta|(1 - |\alpha|^n)}{1 - |\alpha|} + \frac{|b - a\alpha|(1 - |\beta|^n)}{1 - |\beta|} \right], \end{aligned}$$

where $\alpha = \frac{A + \sqrt{A^2 + 4B}}{2}$, $\beta = \frac{A - \sqrt{A^2 + 4B}}{2}$. Thus the proof is completed. □

Example *Let $\mathcal{A}_F$ be a 4-circulant matrix of the order 5 with the first row $(F_0^{(0,-1)}, F_1^{(0,-1)}, \cdots, F_4^{(0,-1)})$, then*

$$\sqrt{15} \leq \|\mathcal{A}_F\|_2 \leq 3 + \frac{12}{\sqrt{5}}.$$

Theorem 2 *Let $\mathcal{A}_{U^2}$ be as (1.6), with $a_i = [U_i^{(a,b)}]^2 (i = 0, 1, \cdots, n-1)$ in the first row of $\mathcal{A}_{U^2}$. If $B \neq -1$, $B \pm A \neq 1$ and $(n, g) = 1$, then we have the following identity*

$$\|\mathcal{A}_{U^2}\|_2 = \frac{M - [U_n^{(a,b)}]^2 + B^2 [U_{n-1}^{(a,b)}]^2}{(1 - B)^2 - A^2} + \frac{2(a^2 B + abA - b^2)[1 - (-B)^n]}{(1 + B)(A^2 + 4B)},$$

where $M = a^2 - (aA - b)^2 - \frac{2(1+B)(a^2 B + abA - b^2)[1 - (-B)^n]}{A^2 + 4B}$.

Proof: Applying the results from Lemma 1 and Lemma 2, the following is valid

$$(\mathcal{A}_{U^2})^H \mathcal{A}_{U^2} = (Q_g C)^H Q_g C = C^H (Q_g)^H Q_g C = C^H I_n C = C^H C,$$

where $C = [c_{ij}] \in M_n$ is a circulant matrix with the first row $([U_0^{(a,b)}]^2, [U_1^{(a,b)}]^2, \cdots, [U_{n-1}^{(a,b)}]^2)$. Hence the spectral norm of matrix $\mathcal{A}_{U^2}$ is the same as that of C.

Since the circulant matrix C is normal, there exists a unitary matrix $V \in M_n$ such that $V^H C V = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$, where λ_i is eigenvalue of C, hence

$$V^H C^H C V = diag(|\lambda_1|^2, |\lambda_2|^2, \cdots, |\lambda_n|^2).$$

Thus, the spectral norm of C is given by its spectral radius. Also since C is nonnegative, its spectral radius $\rho(C)$ satisfies the following inequality:

$$\min_{1 \leq i \leq n} \sum_{j=1}^{n} c_{ij} \leq \rho(C) \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} c_{ij}$$

While

$$\sum_{j=1}^{n} c_{ij} = \sum_{k=0}^{n-1} [U_k^{(a,b)}]^2 = \frac{M - [U_n^{(a,b)}]^2 + B^2 [U_{n-1}^{(a,b)}]^2}{(1-B)^2 - A^2} + \frac{2(a^2 B + abA - b^2)[1-(-B)^n]}{(1+B)(A^2+4B)}$$

for any $i = 1, 2, \cdots, n$, where $M = a^2 - (aA-b)^2 - \frac{2(1+B)(a^2B+abA-b^2)[1-(-B)^n]}{A^2+4B}$. Hence

$$\|\mathcal{A}_{U^2}\|_2 = \|C\|_2 = \frac{M - [U_n^{(a,b)}]^2 + B^2 [U_{n-1}^{(a,b)}]^2}{(1-B)^2 - A^2} + \frac{2(a^2 B + abA - b^2)[1-(-B)^n]}{(1+B)(A^2+4B)}.$$

Thus the proof is completed. □

In the particular case $A = B = 1, a = 0$ and $b = 1$ from Theorem 2, we get the spectral norm for g-circulant matrix with the first row $(F_0^2, F_1^2, \cdots, F_{n-1}^2)$, which is the known result in [11].

Corollary 1 *Let $\mathcal{A}_{F^2}$ be as (1.6), with $a_i = F_i^2 (i = 0, 1, \cdots, n-1)$ in the first row of $\mathcal{A}_{F^2}$. If $(n, g) = 1$, then we have*

$$\|\mathcal{A}_{F^2}\|_2 = F_n F_{n-1}.$$

Proof: We select $A = B = 1, a = 0$ and $b = 1$ in Theorem 2, then the following is valid

$$\|\mathcal{A}_{F^2}\|_2 = F_n^2 - F_{n-1}^2 + (-1)^n,$$

Thus, the proof is completed from the following identity

$$F_n F_{n-1} - (F_n^2 - F_{n-1}^2) = F_{n+1} F_{n-1} - F_n^2 = (-1)^n. \qquad \square$$

In the case $A = B = 1, a = 2$ and $b = 1$ from Theorem 2, we obtain the following result in [11].

Corollary 2 *Let $\mathcal{A}_{L^2}$ be as (1.6), with $a_i = L_i^2 (i = 0, 1, \cdots, n-1)$ in the first row of $\mathcal{A}_{L^2}$. If $(n, g) = 1$, then we have the following identity*

$$\|\mathcal{A}_{L^2}\|_2 = L_n L_{n-1} + 2.$$

Proof: When $A = B = 1, a = 2$ and $b = 1$ in Theorem 2, then we have

$$\|\mathcal{A}_{L^2}\|_2 = L_n^2 - L_{n-1}^2 - 5(-1)^n + 2.$$

On the other hand, applying identities $F_n + L_n = 2F_{n+1}$ and $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$, then we have

$$\begin{aligned} L_n^2 - L_{n-1}^2 &= (2F_{n+1} - F_n)^2 - (2F_n - F_{n-1})^2 \\ &= 4(F_{n+1}^2 - F_{n+1}F_n - F_n^2) + (F_n^2 - F_nF_{n-1} - F_{n-1}^2) + 5F_nF_{n-1} \\ &= 3(-1)^n + 5F_nF_{n-1}, \end{aligned}$$

hence, the following is valid

$$\begin{aligned} L_n^2 - L_{n-1}^2 - L_nL_{n-1} &= 3(-1)^n + 5F_nF_{n-1} - (2F_{n+1} - F_n)(2F_n - F_{n-1}) \\ &= 3(-1)^n + 4(F_nF_{n-1} - F_{n+1}F_n) + 2F_{n+1}F_{n-1} + 2F_n^2 \\ &= 3(-1)^n + 2(F_{n+1}F_{n-1} - F_n^2) \\ &= 5(-1)^n. \end{aligned}$$

Thus the proof is completed. □

3. Numerical tests

In this section, we list the results for Fibonacci and Lucas numbers in Table 1. Employing the formulas in above corollaries, the numerical results demonstrate that the explicit identities of spectral norms of g-circulant matrices hold exactly.

Table 1. Numerical results of $a_i = F_i^2, L_i^2$

n	7					9				
g	2	3	4	5	6	2	4	5	7	8
$\|\mathcal{A}_{F^2}\|_2$	104	104	104	104	104	714	714	714	714	714
$\|\mathcal{A}_{L^2}\|_2$	524	524	524	524	524	3574	3574	3574	3574	3574
F_nF_{n-1}	104	104	104	104	104	714	714	714	714	714
$L_nL_{n-1}+2$	524	524	524	524	524	3574	3574	3574	3574	3574

4. Conclusion

In this paper we introduce the notion of the classical Horadam numbers $U_i^{(a,b)}$, then give upper and lower bounds for the spectral norm of g-circulant matrix, whose the first row entries are $(U_0^{(a,b)}, U_1^{(a,b)}, \cdots, U_{n-1}^{(a,b)})$. In addition, we also establish an explicit formula of the spectral norm for g-circulant matrix with the first row $([U_0^{(a,b)}]^2, [U_1^{(a,b)}]^2, \cdots, [U_{n-1}^{(a,b)}]^2)$. In two particular cases $A = B = 1, a = 0, b = 1$ and $A = B = 1, a = 2, b = 1$, we obtain the known results from [11].

Acknowledgement

The authors thank the referees for their helpful suggestions concerning the presentation of this paper.

References

[1] S. Falcón, Á. Plaza, The k-Fibonacci sequence and the Pascal 2-triangle, Chaos, Solitons Fract. 33(2007) 38–49.

[2] S. Falcón, Á. Plaza, On k-Fibonacci numbers of arithmetic indexes, Appl. Math. Comput. 208(2009) 180–185.

[3] K. Uslu, N. Taskara, H. Kose, The generalized k-Fibonacci and k-Lucas numbers, Ars Combinatoria 99(2011) 25–32.

[4] P. Stanimirović, J. Nikolov, I. Stanimirović, A generalization of Fibonacci and Lucas matrices, Discrete Appl. Math. 156(2008) 2606–2619.

[5] M. Akbulak, D. Bozkurt, On the norms of Toeplitz matrices involving Fibonacci and Lucas numbers, Hacet. J. Math. Stat. 37(2008) 89–95.

[6] S.-Q. Shen, On the norms of Toeplitz matrices involving k-Fibonacci and k-Lucas numbers, Int. J. Contemp. Math. Science 7(2012) 363–368.

[7] S. Solak, On the norms of circulant matrices with the Fibonacci and Lucas numbers, Appl. Math. Comput. 160(2005) 125–132.

[8] S. Solak, Erratum to "On the norms of circulant matrices with the Fibonacci and Lucas numbers" [Appl. Math. Comput. 160(2005) 125–132], Appl. Math. Comput. 190(2007) 1855–1856.

[9] A. İpek, On the spectral norms of circulant matrices with classical Fibonacci and Lucas numbers entries, Appl. Math. Comput. 217(2011) 6011–6012.

[10] A. Bose, S. Guha, R.S. Hazra, K. Saha, Circulant type matrices with heavy tailed entries, Stat. Probab. Lett. 81(2011) 1706–1716.

[11] J. Zhou, Z. Jiang, The spectral norms of g-circulant matrices with classical Fibonacci and Lucas numbers entries, Appl. Math. Comput. 233(2014) 582–587.

[12] S.-Q. Shen, J.-M. Cen, Y. Hao, On the determinants and inverses of circulant matrices with Fibonacci and Lucas numbers, Appl. Math. Comput. 217(2011) 9790–9797.

[13] Y. Yazlik, N. Taskara, On the inverse of circulant matrix via generalized k-Horadam numbers, Appl. Math. Comput. 223(2013) 191–196.

[14] S.-Q. Shen, J.-J. He, Moore-Penrose inverse of generalized Fibonacci matrix and its applications, Int. J. Comput. Math. 93(2016) 1756–1770.

[15] W.T. Stallings, T.L. Boullion, The pseudoinverse of an r-circulant matrix, Proc. AMS 34(1972) 385–388.

[16] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1449 – 1460

Some subclasses of meromorphic functions involving the Hurwitz-Lerch Zeta function

Zhi-Gang Wang*† and Lei Shi‡

Abstract

The main purpose of this paper is to investigate some subclasses of meromorphic functions involving the meromorphic modified version of the familiar Srivastava-Attiya operator. Such results as inclusion relationships, convolution properties, coefficient inequalities, integral-preserving properties, subordination and superordination properties are proved.

Keywords: Analytic function; Meromorphic function; Hurwitz-Lerch Zeta function; Srivastava-Attiya opertor; Differential subordination.

2000 AMS Classification: Primary 30C45; Secondary 30C80.

Received : 01.03.2015 *Accepted* : 23.11.2015 *Doi* : 10.15672/HJMS.20164514284

1. Introduction

Let Σ denote the class of functions of the form

$$f(z) = \frac{1}{z} + \sum_{k=1}^{\infty} a_k z^k, \tag{1.1}$$

which are *analytic* in the *punctured* open unit disk

$$\mathbb{U}^* := \{z : z \in \mathbb{C} \text{ and } 0 < |z| < 1\} =: \mathbb{U}\backslash\{0\}.$$

Let $f,\ g \in \Sigma$, where f is given by (1.1) and g is defined by

$$g(z) = \frac{1}{z} + \sum_{k=1}^{\infty} b_k z^k.$$

*Corresponding Author.

†School of Mathematics and Computing Science, Hunan First Normal University, Changsha 410205, Hunan, People's Republic of China.
Email : wangmath@163.com

‡School of Mathematics and Statistics, Anyang Normal University, Anyang 455000, Henan, People's Republic of China.
Email : shimath@163.com

Then the Hadamard product (or convolution) $f*g$ of the functions f and g is defined by

$$(f*g)(z) := \frac{1}{z} + \sum_{k=1}^{\infty} a_k b_k z^k =: (g*f)(z).$$

Let $\mathcal{P}$ denote the class of functions of the form

$$p(z) = 1 + \sum_{k=1}^{\infty} p_k z^k,$$

which are analytic and convex in $\mathbb{U}$, and satisfy the condition

$$\Re(p(z)) > 0 \quad (z \in \mathbb{U}).$$

For two functions f and g, analytic in $\mathbb{U}$, the function f is said to be subordinate to g in $\mathbb{U}$, or the function g is said to be superordinate to f in $\mathbb{U}$, and write

$$f(z) \prec g(z) \quad (z \in \mathbb{U}),$$

if there exists a Schwarz function ω, which is analytic in $\mathbb{U}$ with

$$\omega(0) = 0 \text{ and } |\omega(z)| < 1 \ (z \in \mathbb{U})$$

such that

$$f(z) = g\big(\omega(z)\big) \quad (z \in \mathbb{U}).$$

Indeed, it is known that

$$f(z) \prec g(z) \ (z \in \mathbb{U}) \Longrightarrow f(0) = g(0) \text{ and } f(\mathbb{U}) \subset g(\mathbb{U}).$$

Furthermore, if the function g is univalent in $\mathbb{U}$, then we have the following equivalence:

$$f(z) \prec g(z) \ (z \in \mathbb{U}) \Longleftrightarrow f(0) = g(0) \text{ and } f(\mathbb{U}) \subset g(\mathbb{U}).$$

The following we recall a general Hurwitz-Lerch Zeta function $\Phi(z,s,a)$ defined by (*cf., e.g.,* [20, p. 121 *et sep.*])

$$\Phi(z,s,a) := \sum_{k=0}^{\infty} \frac{z^k}{(k+a)^s} \tag{1.2}$$

$$(a \in \mathbb{C} \setminus \mathbb{Z}_0^-;\ s \in \mathbb{C} \text{ when } |z| < 1;\ \Re(s) > 1 \text{ when } |z| = 1),$$

where, as usual,

$$\mathbb{Z}_0^- := \mathbb{Z} \setminus \mathbb{N} \quad (\mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\};\ \mathbb{N} := \{1, 2, 3, \ldots\}).$$

Several interesting properties and characteristics of the Hurwitz-Lerch Zeta function $\Phi(z,s,a)$ can be found in the recent investigations by (for example) Choi and Srivastava [1], Ferreira and López [4], Garg *et al.* [5], Lin *et al.* [7], Luo and Srivastava [10], Srivastava *et al.* [21], Ghanim [6] and others.

By making use of the Hurwitz-Lerch Zeta function $\Phi(z,s,a)$, Srivastava and Attiya [19] (see also [8, 9, 14, 17, 22, 23, 24, 27, 28, 29, 30]) recently introduced and investigated the integral operator

$$\mathcal{J}_{s,b}\mathfrak{f}(z) = z + \sum_{k=2}^{\infty} \left(\frac{1+b}{k+b}\right)^s c_k z^k \quad (b \in \mathbb{C} \setminus \mathbb{Z}^-;\ s \in \mathbb{C};\ z \in \mathbb{U}).$$

Motivated essentially by the above-mentioned Srivastava-Attiya operator $\mathcal{J}_{s,b}$, we now introduce the linear operator

$$\mathcal{W}_{s,b}:\ \Sigma \longrightarrow \Sigma$$

defined, in terms of the Hadamard product (or convolution), by

$$\mathcal{W}_{s,b} f(z) := \Theta_{s,b}(z) * f(z) \quad \left(b \in \mathbb{C} \setminus \{\mathbb{Z}_0^- \cup \{1\}\};\ s \in \mathbb{C};\ f \in \Sigma;\ z \in \mathbb{U}^*\right), \tag{1.3}$$

where, for convenience,

$$\Theta_{s,\,b}(z) := (b-1)^s \left[\Phi(z,s,b) - b^{-s} + \frac{1}{z\,(b-1)^s}\right] \quad (z \in \mathbb{U}^*). \tag{1.4}$$

It can easily be seen from (1.1) to (1.4) that

$$\mathcal{W}_{s,\,b}f(z) = \frac{1}{z} + \sum_{k=1}^{\infty} \left(\frac{b-1}{b+k}\right)^s a_k z^k. \tag{1.5}$$

Indeed, the operator $\mathcal{W}_{s,\,b}$ can be defined for $b \in \mathbb{C} \setminus \{\mathbb{Z}^- \cup \{1\}\}$, where

$$\mathcal{W}_{s,\,0}f(z) := \lim_{b\to 0} \{\mathcal{W}_{s,\,b}f(z)\}\,.$$

We observe that

$$\mathcal{W}_{0,\,b}f(z) = f(z), \tag{1.6}$$

and

$$\mathcal{W}_{1,\,\gamma}f(z) = \frac{\gamma-1}{z^\gamma}\int_0^z t^{\gamma-1}f(t)dt \quad (\Re(\gamma) > 1)\,. \tag{1.7}$$

Furthermore, from the definition (1.5), we find that

$$\mathcal{W}_{s+1,\,b}f(z) = \frac{b-1}{z^b}\int_0^z t^{\,b-1}\mathcal{W}_{s,\,b}f(t)dt \quad (\Re(b) > 1)\,. \tag{1.8}$$

Differentiating both sides of (1.8) with respect to z, we get the following useful relationship:

$$z\,(\mathcal{W}_{s+1,\,b}f)'\,(z) = (b-1)\,\mathcal{W}_{s,\,b}f(z) - b\,\mathcal{W}_{s+1,\,b}f(z). \tag{1.9}$$

By using the integral operator (1.5), we now introduce the following subclasses of the class Σ of meromorphic functions.

1.1. Definition. A function $f \in \Sigma$ is said to be in the class $\mathcal{MS}_{s,\,b}(\eta;\phi)$ if it satisfies the subordination

$$\frac{1}{1-\eta}\left(-\frac{z\,(\mathcal{W}_{s,\,b}f)'\,(z)}{\mathcal{W}_{s,\,b}f(z)} - \eta\right) \prec \phi(z) \tag{1.10}$$

$$(s \in \mathbb{C};\ \Re(b) > 1;\ \eta \in [0,1);\ \phi \in \mathcal{P};\ z \in \mathbb{U}).$$

1.2. Definition. A function $f \in \Sigma$ is said to be in the class $\mathcal{MC}_{s,\,b}(\lambda;\phi)$ if it satisfies the condition

$$(1-\lambda)z\,\mathcal{W}_{s+1,\,b}f(z) + \lambda z\,\mathcal{W}_{s,\,b}f(z) \prec \phi(z) \quad (s,\,\lambda \in \mathbb{C};\ \Re(b) > 1;\ \phi \in \mathcal{P};\ z \in \mathbb{U}). \tag{1.11}$$

For some recent investigations on meromorphic functions, see (for example) the earlier works [2, 3, 15, 16, 25, 26, 31] and the references cited therein. In this paper, we aim at deriving the inclusion relationships, convolution properties, coefficient inequalities, integral-preserving properties, subordination and superordination properties for the function classes $\mathcal{MS}_{s,\,b}(\eta;\phi)$ and $\mathcal{MC}_{s,\,b}(\lambda;\phi)$.

2. Preliminary results

The following lemmas will be required in the proof of our main results.

2.1. Lemma. ([11]) *Let $\vartheta, \gamma \in \mathbb{C}$. Suppose that ψ is convex and univalent in $\mathbb{U}$ with*

$$\psi(0) = 1 \text{ and } \Re(\vartheta\psi(z) + \gamma) > 0 \ (z \in \mathbb{U}).$$

If $\mathfrak{p}$ is analytic in $\mathbb{U}$ with $\mathfrak{p}(0) = 1$, then the following subordination

$$\mathfrak{p}(z) + \frac{z\mathfrak{p}'(z)}{\vartheta\mathfrak{p}(z) + \gamma} \prec \psi(z) \quad (z \in \mathbb{U})$$

implies that

$$\mathfrak{p}(z) \prec \psi(z) \quad (z \in \mathbb{U}).$$

2.2. Lemma. *Let $0 \leq \alpha < 1$, $s \in \mathbb{C}$ and $\Re(b) > 1$. Suppose also that the sequence $\{A_k\}_{k=1}^{\infty}$ is defined by*

(2.1)

$$A_1 = (1-\alpha)\left|\frac{b+1}{b-1}\right|^s, \ A_{k+1} = \frac{2(1-\alpha)}{k+2}\left|\frac{b+k+1}{b-1}\right|^s\left(1 + \sum_{m=1}^{k}\left|\frac{b-1}{b+m}\right|^s A_m\right) \ (k \in \mathbb{N}).$$

Then

$$A_k = (1-\alpha)\left|\frac{b+1}{b-1}\right|^s \prod_{j=1}^{k-1}\frac{j-2\alpha+3}{j+2}\left|\frac{b+j+1}{b+j}\right|^s. \tag{2.2}$$

Proof. From (2.1), we find that

$$(k+2)\left|\frac{b-1}{b+k+1}\right|^s A_{k+1} = 2(1-\alpha)\left(1 + \sum_{m=1}^{k}\left|\frac{b-1}{b+m}\right|^s A_m\right), \tag{2.3}$$

and

$$(k+1)\left|\frac{b-1}{b+k}\right|^s A_k = 2(1-\alpha)\left(1 + \sum_{m=1}^{k-1}\left|\frac{b-1}{b+m}\right|^s A_m\right). \tag{2.4}$$

Combining (2.3) and (2.4), we get

$$\frac{A_{k+1}}{A_k} = \frac{k-2\alpha+3}{k+2}\left|\frac{b+k+1}{b+k}\right|^s. \tag{2.5}$$

Thus, for $k \geq 2$, we deduce from (2.5) that

$$A_k = \frac{A_k}{A_{k-1}} \cdot \dots \cdot \frac{A_3}{A_2} \cdot \frac{A_2}{A_1} \cdot A_1 = (1-\alpha)\left|\frac{b+1}{b-1}\right|^s \prod_{j=1}^{k-1}\frac{j-2\alpha+3}{j+2}\left|\frac{b+j+1}{b+j}\right|^s.$$

The proof of Lemma 2.2 is completed. □

2.3. Lemma. ([12]) *Let the function Ω be analytic and convex (univalent) in $\mathbb{U}$ with $\Omega(0) = 1$. Suppose also that the function Θ given by*

$$\Theta(z) = 1 + d_n z^n + d_{n+1}z^{n+1} + \cdots$$

is analytic in $\mathbb{U}$. If

$$\Theta(z) + \frac{z\Theta'(z)}{\zeta} \prec \Omega(z) \quad (\Re(\zeta) > 0;\ \zeta \neq 0;\ z \in \mathbb{U}), \tag{2.6}$$

then

$$\Theta(z) \prec \varpi(z) = \frac{\zeta}{n}z^{-\frac{\zeta}{n}}\int_0^z t^{\frac{\zeta}{n}-1}\Omega(t)dt \prec \Omega(z) \quad (z \in \mathbb{U}),$$

and ϖ is the best dominant of (2.6).

2.4. Lemma. ([18]) *Let q be a convex univalent function in $\mathbb{U}$ and let $\sigma,\ \eta \in \mathbb{C}$ with*

$$\Re\left(1+\frac{zq''(z)}{q'(z)}\right) > \max\left\{0,\ -\Re\left(\frac{\sigma}{\eta}\right)\right\}.$$

If p is analytic in $\mathbb{U}$ and

$$\sigma p\left(z\right)+\eta zp'(z) \prec \sigma q(z)+\eta zq'(z),$$

then $p \prec q$ and q is the best dominant.

Denote by Q the set of all functions f that are analytic and injective on $\overline{\mathbb{U}} - E(f)$, where

$$E(f)=\left\{\varepsilon\in\partial\mathbb{U}:\ \lim_{z\to\varepsilon} f(z)=\infty\right\},$$

and such that $f'(\varepsilon) \neq 0$ for $\varepsilon \in \partial\mathbb{U} - E(f)$. Let $\mathcal{H}(\mathbb{U})$ denote the class of analytic functions in $\mathbb{U}$ and let $\mathcal{H}[a,p]$ denote the subclass of the functions $f \in \mathcal{H}(\mathbb{U})$ of the form:

$$f(z)=a+a_p z^p+a_{p+1}z^{p+1}+\cdots \quad (a\in\mathbb{C};\ p\in\mathbb{N}).$$

2.5. Lemma. ([13]) *Let q be convex univalent in $\mathbb{U}$ and $\kappa \in \mathbb{C}$. Further assume that $\Re(\kappa) > 0$. If*

$$p\in\mathcal{H}[q(0),1]\cap Q,$$

and $p+\kappa zp'$ is univalent in $\mathbb{U}$, then

$$q(z)+\kappa zq'(z) \prec p\left(z\right)+\kappa zp'(z)$$

implies $q \prec p$ and q is the best subordinant.

3. Main results

Firstly, we derive the following inclusion relationship for the function class $\mathcal{MS}_{s,\,b}(\eta;\phi)$.

3.1. Theorem. *Let $0 \leq \eta < 1$ and $\phi \in \mathcal{P}$ with*

$$\Re\left((1-\eta)\phi(z)+\eta-b\right)<0 \quad (z\in\mathbb{U}). \tag{3.1}$$

Then

$$\mathcal{MS}_{s,\,b}(\eta;\phi)\subset\mathcal{MS}_{s+1,\,b}(\eta;\phi). \tag{3.2}$$

Proof. Let $f \in \mathcal{MS}_{s,\,b}(\eta;\phi)$ and suppose that

$$\varphi(z):=\frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s+1,\,b}f\right)'(z)}{\mathcal{W}_{s+1,\,b}f(z)}-\eta\right) \quad (z\in\mathbb{U}). \tag{3.3}$$

Then φ is analytic in $\mathbb{U}$ with $\varphi(0) = 1$. By virtue of (1.9) and (3.3), we get

$$(b-1)\frac{\mathcal{W}_{s,\,b}f(z)}{\mathcal{W}_{s+1,\,b}f(z)}=-(1-\eta)\varphi(z)-\eta+b. \tag{3.4}$$

Differentiating both sides of (3.4) with respect to z logarithmically and using (3.3), we have

$$\frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s,\,b}f\right)'(z)}{\mathcal{W}_{s,\,b}f(z)}-\eta\right)=\varphi(z)+\frac{z\varphi'(z)}{-(1-\eta)\varphi(z)-\eta+b}\prec\phi(z). \tag{3.5}$$

By means of (3.1), an application of Lemma 2.1 to (3.5) yields

$$\varphi(z)=\frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s+1,\,b}f\right)'(z)}{\mathcal{W}_{s+1,\,b}f(z)}-\eta\right)\prec\phi(z),$$

that is $f \in \mathcal{MS}_{s+1,\,b}(\eta;\phi)$, which implies that the assertion (3.2) of Theorem 3.1 holds. □

Next, we derive some convolution properties of the class $\mathcal{MS}_{s,\,b}(\eta;\phi)$.

3.2. Theorem. *Let $f \in \mathcal{MS}_{s,\,b}(\eta;\phi)$. Then*

$$f(z) = \left[z^{-1}\cdot \exp\left((\eta-1)\int_0^z \frac{\phi\left(\omega(\xi)\right)-1}{\xi}d\xi\right)\right] * \left(\frac{1}{z}+\sum_{k=1}^{\infty}\left(\frac{b+k}{b-1}\right)^s z^k\right), \tag{3.6}$$

where ω is analytic in $\mathbb{U}$ with

$$\omega(0)=0 \text{ and } |\omega(z)|<1 \ (z\in\mathbb{U}).$$

Proof. Suppose that $f \in \mathcal{MS}_{s,\,b}(\eta;\phi)$. We find from (1.10) that

$$\frac{z\left(\mathcal{W}_{s,\,b}f\right)'(z)}{\mathcal{W}_{s,\,b}f(z)} = (\eta-1)\phi\left(\omega(z)\right)-\eta, \tag{3.7}$$

where ω is analytic in $\mathbb{U}$ with $\omega(0)=0$ and $|\omega(z)|<1$ $(z\in\mathbb{U})$. From (3.7), we get

$$\frac{\left(\mathcal{W}_{s,\,b}f\right)'(z)}{\mathcal{W}_{s,\,b}f(z)}+\frac{1}{z} = (\eta-1)\frac{\phi\left(\omega(z)\right)-1}{z}, \tag{3.8}$$

which, upon integration, yields

$$\log\left(z\,\mathcal{W}_{s,\,b}f(z)\right) = (\eta-1)\int_0^z \frac{\phi\left(\omega(\xi)\right)-1}{\xi}d\xi. \tag{3.9}$$

It follows from (3.9) that

$$\mathcal{W}_{s,\,b}f(z) = z^{-1}\cdot\exp\left((\eta-1)\int_0^z \frac{\phi\left(\omega(\xi)\right)-1}{\xi}d\xi\right). \tag{3.10}$$

The assertion (3.6) of Theorem 3.2 can directly be derived from (1.5) and (3.10). □

3.3. Theorem. *Let $f \in \Sigma$ and $\phi\in\mathcal{P}$. Then $f \in \mathcal{MS}_{s,\,b}(\eta;\phi)$ if and only if*

(3.11)

$$\frac{1}{z}\left\{f*\left\{-\frac{1}{z}+\sum_{k=1}^{\infty}k\left(\frac{b-1}{b+k}\right)^s z^k-\left[(\eta-1)\phi\left(e^{i\theta}\right)-\eta\right]\left(\frac{1}{z}+\sum_{k=1}^{\infty}\left(\frac{b-1}{b+k}\right)^s z^k\right)\right\}\right\}\neq 0$$
$$(z\in\mathbb{U}^*;\ 0\le\theta<2\pi).$$

Proof. Suppose that $f \in \mathcal{MS}_{s,\,b}(\eta;\phi)$. We know that (1.6) is equivalent to

$$\frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s,\,b}f\right)'(z)}{\mathcal{W}_{s,\,b}f(z)}-\eta\right)\neq\phi\left(e^{i\theta}\right) \quad (z\in\mathbb{U};\ 0\le\theta<2\pi). \tag{3.12}$$

It is easy to see that the condition (3.12) can be written as follows:

$$\frac{1}{z}\left\{z\left(\mathcal{W}_{s,\,b}f\right)'(z)-\left[(\eta-1)\phi\left(e^{i\theta}\right)-\eta\right]\mathcal{W}_{s,\,b}f(z)\right\}\neq 0 \quad (z\in\mathbb{U}^*;\ 0\le\theta<2\pi). \tag{3.13}$$

On the other hand, we find from (1.5) that

$$z\left(\mathcal{W}_{s,\,b}f\right)'(z) = -\frac{1}{z}+\sum_{k=1}^{\infty}k\left(\frac{b-1}{b+k}\right)^s a_k z^k. \tag{3.14}$$

Combining (1.5), (3.13) and (3.14), we get the assertion (3.11) of Theorem 3.3. □

3.4. Theorem. *If $f \in \mathcal{MS}_{s,b}(0;[1+(1-2\alpha)z]/(1-z))$, then*

$$|a_1| \leq (1-\alpha)\left|\frac{b+1}{b-1}\right|^s,$$

and

$$|a_k| \leq (1-\alpha)\left|\frac{b+1}{b-1}\right|^s \prod_{j=1}^{k-1} \frac{j-2\alpha+3}{j+2}\left|\frac{b+j+1}{b+j}\right|^s \quad (k \in \mathbb{N}\backslash\{1\}).$$

Proof. Suppose that

$$h(z) := \frac{-\frac{z\left(\mathcal{W}_{s,b}f\right)'(z)}{\mathcal{W}_{s,b}f(z)} - \alpha}{1-\alpha} = 1 + c_1 z + c_2 z^2 + \cdots. \tag{3.15}$$

It follows from $f \in \mathcal{MS}_{s,b}(0;[1+(1-2\alpha)z]/(1-z))$ that $h \in \mathcal{P}$, and subsequently one has $|c_k| \leq 2$ for $k \in \mathbb{N}$.

By virtue of (3.15), we know that

$$z\left(\mathcal{W}_{s,b}f\right)'(z) = [(\alpha-1)h(z)-\alpha]\mathcal{W}_{s,b}f(z). \tag{3.16}$$

It now follows from (1.5), (3.15) and (3.16) that

(3.17)

$$\frac{1}{z}+\sum_{k=1}^{\infty} k\left(\frac{b-1}{b+k}\right)^s a_k z^k = \left[-1+(\alpha-1)\left(c_1 z + c_2 z^2 + \cdots\right)\right]\left[\frac{1}{z}+\sum_{k=1}^{\infty}\left(\frac{b-1}{b+k}\right)^s a_k z^k\right].$$

By evaluating the coefficients of z^k in both sides of (3.17), we get

$$k\left(\frac{b-1}{b+k}\right)^s a_k = -\left(\frac{b-1}{b+k}\right)^s a_k + (\alpha-1)\left[c_{k+1} + \sum_{l=1}^{k-1} c_l \left(\frac{b-1}{b+k-l}\right)^s a_{k-l}\right]. \tag{3.18}$$

By observing the fact that $|c_k| \leq 2$ for $k \in \mathbb{N}$, we find from (3.18) that

$$|a_k| \leq \frac{2(1-\alpha)}{k+1}\left|\frac{b+k}{b-1}\right|^s \left(1 + \sum_{m=1}^{k-1}\left|\frac{b-1}{b+m}\right|^s |a_m|\right). \tag{3.19}$$

Now, we define the sequence $\{A_k\}_{k=1}^{\infty}$ as follows:

(3.20)

$$A_1 = (1-\alpha)\left|\frac{b+1}{b-1}\right|^s, \;\; A_{k+1} = \frac{2(1-\alpha)}{k+2}\left|\frac{b+k+1}{b-1}\right|^s \left(1 + \sum_{m=1}^{k}\left|\frac{b-1}{b+m}\right|^s A_m\right) \; (k \in \mathbb{N}).$$

In order to prove that

$$|a_k| \leq A_k \quad (k \in \mathbb{N}),$$

we make use of the principle of mathematical induction. By noting that

$$|a_1| \leq A_1 = (1-\alpha)\left|\frac{b+1}{b-1}\right|^s.$$

Therefore, assuming that

$$|a_m| \leq A_m \quad (m = 1,2,3,\cdots,k;\; k \in \mathbb{N}).$$

Combining (3.19) and (3.20), we get

$$\begin{aligned} |a_{k+1}| &\leq \frac{2(1-\alpha)}{k+2}\left|\frac{b+k+1}{b-1}\right|^{s}\left(1+\sum_{m=1}^{k}\left|\frac{b-1}{b+m}\right|^{s}|a_m|\right) \\ &\leq \frac{2(1-\alpha)}{k+2}\left|\frac{b+k+1}{b-1}\right|^{s}\left(1+\sum_{m=1}^{k}\left|\frac{b-1}{b+m}\right|^{s}A_m\right) \\ &= A_{k+1}. \end{aligned}$$

Hence, by the principle of mathematical induction, we have

$$|a_k| \leq A_k \quad (k \in \mathbb{N}) \tag{3.21}$$

as desired.

By virtue of Lemma 2.2 and (3.20), we know that (2.2) holds. Combining (3.21) and (2.2), we readily get the coefficient estimates asserted by Theorem 3.4. □

In what follows, we derive some integral-preserving properties for the class $\mathcal{MS}_{s,\,b}(\eta;\phi)$.

3.5. Theorem. *Let $f \in \mathcal{MS}_{s,\,b}(\eta;\phi)$ with*

$$\Re((1-\eta)\phi(z)+\eta-\mu) < 0 \quad (z \in \mathbb{U};\ \Re(\mu) > 1).$$

Then the integral operator F defined by

$$F(z) := \frac{\mu-1}{z^{\mu}}\int_0^z t^{\mu-1}f(t)dt \quad (z \in \mathbb{U}^*;\ \Re(\mu) > 1) \tag{3.22}$$

belongs to the class $\mathcal{MS}_{s,\,b}(\eta;\phi)$.

Proof. Let $f \in \mathcal{MS}_{s,\,b}(\eta;\phi)$. We then find from (3.22) that

$$z\left(\mathcal{W}_{s,\,b}F\right)'(z) + \mu\mathcal{W}_{s,\,b}F(z) = (\mu-1)\mathcal{W}_{s,\,b}f(z). \tag{3.23}$$

By setting

$$q(z) := \frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s,\,b}F\right)'(z)}{\mathcal{W}_{s,\,b}F(z)} - \eta\right), \tag{3.24}$$

we observe that q is analytic in $\mathbb{U}$ with $q(0) = 1$. It follows from (3.23) and (3.24) that

$$-(1-\eta)q(z) - \eta + \mu = (\mu-1)\frac{\mathcal{W}_{s,\,b}f(z)}{\mathcal{W}_{s,\,b}F(z)}. \tag{3.25}$$

Differentiating both sides of (3.25) with respect to z logarithmically and using (3.24), we get

$$q(z) + \frac{zq'(z)}{-(1-\eta)q(z)-\eta+\mu} = \frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s,\,b}f\right)'(z)}{\mathcal{W}_{s,\,b}f(z)} - \eta\right) \prec \phi(z). \tag{3.26}$$

Since

$$\Re(-(1-\eta)\phi(z)-\eta+\mu) > 0 \quad (z \in \mathbb{U}),$$

by virtue of Lemma 2.1 and (3.26), we obtain

$$\frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s,\,b}F\right)'(z)}{\mathcal{W}_{s,\,b}F(z)} - \eta\right) \prec \phi(z),$$

which implies that the assertion of Theorem 3.5 holds. □

3.6. Theorem. *Let* $f \in \mathcal{MS}_{s,\ b}(\eta;\phi)$ *with*

$$\Re((1-\eta)\delta\,\phi(z)+\eta\,\delta-\mu)<0 \quad (z\in\mathbb{U};\ \delta\neq 0;\ \mu\in\mathbb{C}).$$

Then the function $K \in \Sigma$ *defined by*

$$\text{(3.27)} \qquad \mathcal{W}_{s,\ b}K(z) := \left(\frac{\mu-\delta}{z^{\mu}}\int_0^z t^{\mu-1}\left(\mathcal{W}_{s,\ b}f(t)\right)^{\delta}dt\right)^{1/\delta} \qquad (z\in\mathbb{U}^*;\ \delta\neq 0)$$

belongs to the class $\mathcal{MS}_{s,\ b}(\eta;\phi)$.

Proof. Let $f \in \mathcal{MS}_{s,\ b}(\eta;\phi)$ and suppose that

$$\text{(3.28)} \qquad \varrho(z) := \frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s,\ b}K\right)'(z)}{\mathcal{W}_{s,\ b}K(z)}-\eta\right) \qquad (z\in\mathbb{U}).$$

In view of (3.27) and (3.28), we have

$$\text{(3.29)} \qquad \mu-\eta\,\delta-(1-\eta)\delta\,\varrho(z) = (\mu-\delta)\left(\frac{\mathcal{W}_{s,\ b}f(z)}{\mathcal{W}_{s,\ b}K(z)}\right)^{\delta}.$$

Now, by means of (3.27), (3.28) and (3.29), we obtain

$$\text{(3.30)} \qquad \varrho(z)+\frac{z\varrho'(z)}{\mu-\eta\,\delta-(1-\eta)\delta\,\varrho(z)} = \frac{1}{1-\eta}\left(-\frac{z\left(\mathcal{W}_{s,\ b}f\right)'(z)}{\mathcal{W}_{s,\ b}f(z)}-\eta\right)\prec\phi(z).$$

Since

$$\Re(\mu-\eta\,\delta-(1-\eta)\delta\,\phi(z))>0 \quad (z\in\mathbb{U}),$$

it follows from (3.30) and Lemma 2.1 that $\varrho(z)\prec\phi(z)$, that is $K\in\mathcal{MS}_{s,\ b}(\eta;\phi)$. We thus complete the proof of Theorem 3.6. □

Now, we derive the following subordination property for the class $\mathcal{MC}_{s,\ b}(\lambda;\phi)$.

3.7. Theorem. *Let* $f\in\mathcal{MC}_{s,\ b}(\lambda;\phi)$ *with* $\Re(\lambda/(b-1))>0$. *Then*

$$\text{(3.31)} \qquad z\,\mathcal{W}_{s+1,\ b}f(z)\prec\frac{b-1}{2\lambda}z^{-\frac{b-1}{2\lambda}}\int_0^z t^{\frac{b-1}{2\lambda}-1}\phi(t)dt\prec\phi(z).$$

Proof. Let $f\in\mathcal{MC}_{s,\ b}(\lambda;\phi)$ and suppose that

$$\text{(3.32)} \qquad \mathfrak{h}(z) := z\,\mathcal{W}_{s+1,\ b}f(z) \quad (z\in\mathbb{U}).$$

Then $\mathfrak{h}$ is analytic in $\mathbb{U}$. By virtue of (1.5), (1.11) and (3.32), we find that

$$\text{(3.33)} \qquad \mathfrak{h}(z)+\frac{\lambda}{b-1}z\mathfrak{h}'(z) = (1-\lambda)z\,\mathcal{W}_{s+1,\ b}f(z)+\lambda z\,\mathcal{W}_{s,\ b}f(z)\prec\phi(z).$$

Thus, an application of Lemma 2.3 to (3.33) yields the desired assertion (3.31) of Theorem 3.7. □

3.8. Theorem. *Let* $\lambda_2>\lambda_1\geq 0$. *Then* $\mathcal{MC}_{s,\ b}(\lambda_2;\phi)\subset\mathcal{MC}_{s,\ b}(\lambda_1;\phi)$.

Proof. Suppose that $f\in\mathcal{MC}_{s,\ b}(\lambda_2;\phi)$. It follows that

$$\text{(3.34)} \qquad (1-\lambda_2)z\,\mathcal{W}_{s+1,\ b}f(z)+\lambda_2 z\,\mathcal{W}_{s,\ b}f(z)\prec\phi(z) \quad (z\in\mathbb{U}).$$

Since

$$0\leq\frac{\lambda_1}{\lambda_2}<1$$

and the function ϕ is convex and univalent in $\mathbb{U}$, we deduce from (3.31) and (3.34) that

$$\begin{aligned}&(1-\lambda_1)z\,\mathcal{W}_{s+1,\ b}f(z)+\lambda_1 z\,\mathcal{W}_{s,\ b}f(z)\\&=\frac{\lambda_1}{\lambda_2}\left[(1-\lambda_2)z\,\mathcal{W}_{s+1,\ b}f(z)+\lambda_2 z\,\mathcal{W}_{s,\ b}f(z)\right]+\left(1-\frac{\lambda_1}{\lambda_2}\right)z\,\mathcal{W}_{s+1,\ b}f(z)\prec\phi(z),\end{aligned}$$

which implies that $f \in \mathcal{MC}_{s,\ b}(\lambda_1;\phi)$. The proof of Theorem 3.8 is thus completed. □

3.9. Theorem. *Let $f \in \mathcal{MC}_{s,\ b}(\lambda;\phi)$. If the function $F \in \Sigma$ is defined by (3.22), then*

(3.35) $\quad z\,\mathcal{W}_{s+1,\ b}F(z) \prec \phi(z) \quad (z \in \mathbb{U}).$

Proof. Let $f \in \mathcal{MC}_{s,\ b}(\lambda;\phi)$ and suppose that

(3.36) $\quad \chi(z) := z\,\mathcal{W}_{s+1,\ b}F(z) \quad (z \in \mathbb{U}).$

From (3.22), we find that

(3.37) $\quad z\,(\mathcal{W}_{s+1,\ b}F)'(z) + \mu\,\mathcal{W}_{s+1,\ b}F(z) = (\mu-1)\,\mathcal{W}_{s+1,\ b}f(z).$

By virtue of (3.31), (3.36) and (3.37), we have

(3.38) $$\chi(z) + \frac{1}{\mu-1}z\,\chi'(z) = z\,\mathcal{W}_{s+1,\ b}f(z) \prec \phi(z).$$

Thus, an application of Lemma 2.3 to (3.38), we get the assertion of Theorem 3.9. □

3.10. Theorem. *Let q_1 be univalent in $\mathbb{U}$. Suppose also that q_1 satisfies the condition*

(3.39) $$\Re\left(1+\frac{zq_1''(z)}{q_1'(z)}\right) > \max\left\{0,\ -\Re\left(\frac{b-1}{\lambda}\right)\right\}.$$

If $f \in \Sigma$ satisfies the following subordination

(3.40) $$(1-\lambda)z\,\mathcal{W}_{s+1,\ b}f(z) + \lambda z\,\mathcal{W}_{s,\ b}f(z) \prec q_1(z) + \frac{\lambda}{b-1}zq_1'(z),$$

then

$$z\,\mathcal{W}_{s+1,\ b}f(z) \prec q_1(z),$$

and q_1 is the best dominant.

Proof. Let the function $\mathfrak{h}$ be defined by (3.32). We know that (3.33) holds. Combining (3.33) and (3.40), we find that

(3.41) $$\mathfrak{h}(z) + \frac{\lambda}{b-1}z\mathfrak{h}'(z) \prec q_1(z) + \frac{\lambda}{b-1}zq_1'(z).$$

By Lemma 2.4 and (3.41), we obtain the assertion of Theorem 3.10. □

We now derive the following superordination result for the class $\mathcal{MC}_{s,\ b}(\lambda;\phi)$.

3.11. Theorem. *Let q_2 be convex univalent in $\mathbb{U}$, $\lambda \in \mathbb{C}$ with $\Re(\lambda) > 0$. Also let $z\,\mathcal{W}_{s+1,\ b}\ f(z) \in \mathcal{H}[q_2(0),1] \cap Q$ and $(1-\lambda)z\,\mathcal{W}_{s+1,\ b}\ f(z) + \lambda z\,\mathcal{W}_{s,\ b}f(z)$ be univalent in $\mathbb{U}$. If*

$$q_2(z) + \frac{\lambda}{b-1}zq_2'(z) \prec (1-\lambda)z\,\mathcal{W}_{s+1,\ b}\ f(z) + \lambda z\,\mathcal{W}_{s,\ b}f(z),$$

then

$$q_2(z) \prec z\,\mathcal{W}_{s+1,\ b}f(z),$$

and q_2 is the best subordinant.

Proof. Let the function $\mathfrak{h}$ be defined by (3.32). Then

$$q_2(z) + \frac{\lambda}{b-1}zq_2'(z) \prec (1-\lambda)z\,\mathcal{W}_{s+1,\ b}f(z) + \lambda z\,\mathcal{W}_{s,\ b}f(z) = \mathfrak{h}(z) + \frac{\lambda}{b-1}z\mathfrak{h}'(z).$$

Thus, an application of Lemma 2.5, yields the assertion of Theorem 3.11. □

Finally, combining the above-mentioned subordination and superordination results, we obtain the following sandwich type result.

3.12. Corollary. *Let q_3 be convex univalent and let q_4 be univalent in $\mathbb{U}$, $\lambda \in \mathbb{C}$ with $\Re(\lambda) > 0$. Suppose also that q_4 satisfies the condition*

$$\Re\left(1 + \frac{zq_4''(z)}{q_4'(z)}\right) > \max\left\{0,\ -\Re\left(\frac{b-1}{\lambda}\right)\right\}.$$

If $0 \neq z\,\mathcal{W}_{s+1,\ b}f(z) \in \mathcal{H}[q_3(0), 1] \cap Q$ and $(1-\lambda)z\,\mathcal{W}_{s+1,\ b}f(z) + \lambda z\,\mathcal{W}_{s,\ b}f(z)$ is univalent in $\mathbb{U}$, also

$$q_3(z) + \frac{\lambda}{b-1}zq_3'(z) \prec (1-\lambda)z\,\mathcal{W}_{s+1,\ b}f(z) + \lambda z\,\mathcal{W}_{s,\ b}f(z) \prec q_4(z) + \frac{\lambda}{b-1}zq_4'(z),$$

then

$$q_3(z) \prec z\,\mathcal{W}_{s+1,\ b}f(z) \prec q_4(z),$$

and q_3 and q_4 are, respectively, the best subordinant and the best dominant.

Acknowledgments

The present investigation was supported by the *National Natural Science Foundation* under Grant no. 11301008, the *Natural Science Foundation of Hunan Province* under Grant no. 2016JJ2036, the *Foundation for Excellent Youth Teachers of Colleges and Universities of Henan Province* under Grant no. 2013GGJS-146, the *Foundation of Educational Committee of Henan Province* under Grant no. 17A110014. The authors would like to thank the referees for their valuable comments and suggestions, which essentially improved the quality of this paper.

References

[1] Choi, J., and Srivastava, H. M. *Certain families of series associated with the Hurwitz-Lerch Zeta function*, Appl. Math. Comput. **170** (2005), 399–409.

[2] Dziok, J. *Classes of meromorphic functions associated with conic regions*, Acta Math. Sci. Ser. B Engl. Ed. **32** (2012), 765–774.

[3] Dziok, J. *Classes of multivalent analytic and meromorphic functions with two fixed points*, Fixed Point Theory Appl. **2013**, 2013: 86, pp. 1–18.

[4] Ferreira, C., and López, J. L. *Asymptotic expansions of the Hurwitz-Lerch Zeta function*, J. Math. Anal. Appl. **298** (2004), 210–224.

[5] Garg, M., Jain, K., and Srivastava, H. M. *Some relationships between the generalized Apostol-Bernoulli polynomials and Hurwitz-Lerch Zeta functions*, Integral Transforms Spec. Funct. **17** (2006), 803–815.

[6] Ghanim, F. *A study of a certain subclass of Hurwitz-Lerch-Zeta function related to a linear operator*, Abstr. Appl. Anal. vol. **2013**, Article ID 763756, 7 pages, 2013.

[7] Lin, S.-D., and Srivastava, H. M. *Some families of the Hurwitz-Lerch Zeta functions and associated fractional derivative and other integral representations*, Appl. Math. Comput. **154** (2004), 725–733.

[8] Liu, J.-L. *Sufficient condtions for strongly starlike functions involving the generalized Srivastava-Attiya operator*, Integral Transforms Spec. Funct. **22** (2011), 79–90.

[9] Liu, Z.-H., Wang, Z.-G., Wen, F.-H., and Sun, Y. *Some subclasses of analytic functions involving the generalized Srivastava-Attiya operator*, Hacet. J. Math. Stat. **41** (2012), 421–434.

[10] Luo, Q.-M., and Srivastava, H. M. *Some generalizations of the Apostol-Bernoulli and Apostol-Euler polynomials*, J. Math. Anal. Appl. **308** (2005), 290–302.

[11] Miller S. S., and Mocanu, P. T. *On some classes of first order differential subordinations*, Michigan Math. J. **32** (1985), 185–195.

[12] Miller, S. S., and Mocanu, P. T. *Differential Subordination: Theory and Applications*, in: Series in Pure and Applied Mathematics, Vol. **225**, Marcel Dekker, New York, 2000.

[13] Miller, S. S., and Mocanu, P. T. *Subordinats of differential superordinations*, Complex Var. Theory Appl. **48** (2003), 815–826.

[14] Noor, K. I., and Bukhari, S. Z. H. *Some subclasses of analytic and spiral-like functions of complex order involving the Srivastava-Attiya integral operator*, Integral Transforms Spec. Funct. **21** (2010), 907–916.

[15] Nunokawa, M., and Ahuja, O. P. *On meromorphic starlike and convex functions*, Indian J. Pure Appl. Math. **32** (2001), 1027–1032.

[16] Piejko, K., and Sokól, J. *Subclasses of meromorphic functions associated with the Cho-Kwon-Srivastava operator*, J. Math. Anal. Appl. **337** (2008), 1261–1266.

[17] Raducanu, D., and Srivastava, H. M. *A new class of analytic functions defined by means of a convolution operator involving the Hurwitz-Lerch Zeta function*, Integral Transforms Spec. Funct. **18** (2007), 933–943.

[18] Shanmugam, T. N., Ravichandran, V., and Sivasubramanian, S. *Differential sandwich theorems for subclasses of analytic functions*, Aust. J. Math. Anal. Appl. **3** (2006), Article 8, pp. 1–11, (electronic).

[19] Srivastava, H. M., and Attiya, A. A. *An integral operator associated with the Hurwitz-Lerch Zeta function and differential subordination*, Integral Transforms Spec. Funct. **18** (2007), 207–216.

[20] Srivastava, H. M., and Choi, J. *Series associated with the Zeta and related functions*, Dordrecht, Boston, London: Kluwer Academic Publishers, 2001.

[21] Srivastava, H. M., Luo, M.-J., and Raina, R. K. *New results involving a class of generalized Hurwitz-Lerch Zeta functions and their applications*, Turkish J. Anal. Number Theory **1** (2013), 26–35.

[22] Srivastava, H. M., and Gaboury, S. *A new class of analytic functions defined by means of a generalization of the Srivastava-Attiya operator*, J. Inequal. Appl. **2015**, 2015: 39, pp. 1–15.

[23] Srivastava, H. M., Gaboury, S., and Ghanim, F. *Certain subclasses of meromorphically univalent functions defined by a linear operator associated with the λ-generalized Hurwitz-Lerch Zeta function*, Integral Transforms Spec. Funct. **26** (2015), 258–272.

[24] Srivastava, H. M., Raducanu, D., and Salagean, G. S. *A new class of generalized close-to-starlike functions defined by the Srivastava-Attiya operator*, Acta Math. Sin. (Engl. Ser.) **29** (2013), 833–840.

[25] Srivastava, H. M., Yang, D.-G., and Xu, N. *Some subclasses of meromorphically multivalent functions associated with a linear operator*, Appl. Math. Comput. **195** (2008), 11–23.

[26] Sun, Y., Jiang Y.-P., and Liu, Z.-H. *Some subclasses of meromorphic multivalent functions involving a generalized differential operator*, Hacet. J. Math. Stat. **43** (2014), 435–449.

[27] Sun, Y., Kuang, W.-P., and Wang, Z.-G. *Properties for uniformly starlike and related functions under the Srivastava-Attiya operator*, Appl. Math. Comput. **218** (2011), 3615–3623.

[28] Wang, Z.-G., Li, Q.-G., and Jiang, Y.-P. *Certain subclasses of multivalent analytic functions involving the generalized Srivastava-Attiya operator*, Integral Transforms Spec. Funct. **21** (2010), 221–234.

[29] Wang, Z.-G., Liu, Z.-H., and Sun, Y. *Some properties of the generalized Srivastava-Attiya operator*, Integral Transforms Spec. Funct. **23** (2012), 223–236.

[30] Yuan, S.-M., and Liu, Z.-M. *Some properties of two subclasses of k-fold symmetric functions associated with Srivastava-Attiya operator*, Appl. Math. Comput. **218** (2011), 1136–1141.

[31] Yuan, S.-M., Liu, Z.-M., and Srivastava, H. M. *Some inclusion relationships and integral-preserving properties of certain subclasses of meromorphic functions associated with a family of integral operators*, J. Math. Anal. Appl. **337** (2008), 505–515.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1461 – 1474

Abelian model structures and Ding homological dimensions

Chongqing Wei*, Limin Wang† and Zhongkui Liu‡§

Abstract

Let R be an n-FC ring. For $0 < t \leq n$, we construct a new abelian model structure on R-Mod, called the Ding t-projective (t-injective) model structure. Based on this, we establish a bijective correspondence between dg-t-projective (dg-t-injective) R-complexes and Ding t-projective (t-injective) A-modules under some additional conditions, where $A = R[x]/(x^2)$. This gives a generalized version of the bijective correspondence established in [[14]] between dg-projective (dg-injective) R-complexes and Gorenstein projective (injective) A-modules. Finally, we show that the embedding functors $K(\mathcal{DP}) \longrightarrow K(R\text{-Mod})$ and $K(\mathcal{DI}) \longrightarrow K(R\text{-Mod})$ have right and left adjoints respectively, where $K(\mathcal{DP})$ ($K(\mathcal{DI})$) is the homotopy category of complexes of Ding projective (injective) modules, and $K(R\text{-Mod})$ denotes the homotopy category.

Keywords: model structures; Ding t-projective (injective) modules; dg-t-projective (injective) complexes; adjoint functors.

2000 AMS Classification: 16E10, 18G35, 55U35.

Received : 01.11.2013 *Accepted* : 03.03.2014 *Doi* : 10.15672/HJMS.2015449425

---

*School of Mathematical Sciences, Shanxi University, Taiyuan 030006, Shanxi, People's Republic of China
Email : `wcq-7765363@126.com`

†School of Mathematics, Northwest Normal University, Lanzhou 730070, Gansu, People's Republic of China
Email : `wanglm@nwnu.edu.cn`

‡School of Mathematics, Northwest Normal University, Lanzhou 730070, Gansu, People's Republic of China
Email : `liuzk@nwnu.edu.cn`

§This research was supported by National Natural Science Foundation of China (No.11326061, No.11261050 and No.11201377).

1. Introduction and preliminaries. We study model structures on the categories R-Mod and $\mathrm{Ch}(R)$, where R is an n-FC ring. In [[15]], M. Hovey constructed an abelian model structure on R-Mod where the class of cofibrant objects is given by the class of Gorenstein projective modules, the class of fibrant objects is given by the category R-Mod, and the trivial objects are the left R-modules of finite projective dimension. Dually, there was a model structure on R-Mod with the same trivial objects, the class of cofibrant objects being R-Mod, and the class of fibrant objects being the class of the Gorenstein injective modules. Later in [[13]], J. Gillespie constructed another abelian model structure on R-Mod where the class of cofibrant objects is given by the class of Ding projective modules. Dually, there was a model structure on R-Mod where the class of fibrant objects is given by the class of Ding injective modules.

We construct two new abelian model structures on R-Mod, called the Ding t-projective and Ding t-injective model structures. In the first structure, the class of cofibrant objects is formed by the objects with Ding projective dimension at most t. In the second structure, the class of fibrant objects is given by the class of objects with Ding injective dimension at most t. In order to construct these structures, we use a result known by some authors as the Hovey's Criterion, which allows us to get abelian model structures from compatible and complete cotorsion pairs. In this sense, we prove the completeness of the cotorsion pair cogenerated by the class of Ding t-projective modules. Dually, the cotorsion pair generated by the class of Ding t-injective modules is also complete. These structures have their analogues in the category of chain complexes.

For any ring R, there exists an invertible functor from $\mathrm{Ch}(R)$ to the category of graded $R[x]/(x^2)$-modules. In [[14]], the authors proved that this functor gives rise to a bijective correspondence between the dg-projective complexes over R and the Gorenstein projective $R[x]/(x^2)$-modules. The same also occurred between dg-injective complexes over R and Gorenstein injective $R[x]/(x^2)$-modules. We prove the Ding version of these results.

In the end of this paper, we show that the embedding functors $K(\mathcal{DP}) \longrightarrow K(R\text{-Mod})$ and $K(\mathcal{DI}) \longrightarrow K(R\text{-Mod})$ have right and left adjoints respectively, where $K(\mathcal{DP})$ ($K(\mathcal{DI})$) is the homotopy category with each complex constructed by Ding projective (injective) modules, and $K(R\text{-Mod})$ is the homotopy category.

We next recall some known notions and facts needed in the sequel.

In this paper, R denotes a ring with unity, R-Mod the category of left R-modules, and $\mathrm{Ch}(R)$ the category of complexes of left R-modules. A complex

$$\cdots \xrightarrow{\delta_2} C_1 \xrightarrow{\delta_1} C_0 \xrightarrow{\delta_0} C_{-1} \xrightarrow{\delta_{-1}} \cdots$$

of left R-modules will be denoted (C, δ) or C. Given a left R-module M, we will denote by $D^m(M)$ the complex

$$\cdots \longrightarrow 0 \longrightarrow M \xrightarrow{\mathrm{id}} M \longrightarrow 0 \longrightarrow \cdots$$

with the M in the m and $(m-1)$-th position. Given a complex C, ΣC denotes the complex such that $(\Sigma C)_n = C_{n-1}$ and whose boundary operators are $-\delta^C_{n-1}$.

A homomorphism $\varphi : C \longrightarrow D$ of degree n is a family $(\varphi_i)_{i\in\mathbb{Z}}$ of homomorphisms of R-modules $\varphi_i : C_i \longrightarrow D_{n+i}$. All such homomorphisms form an abelian group, denoted $\mathcal{H}\mathrm{om}_R(C, D)_n$, it is clearly isomorphic to $\prod_{i\in\mathbb{Z}} \mathrm{Hom}_R(C_i, D_{n+i})$. We let $\mathcal{H}\mathrm{om}_R(C, D)$ denote the complex of abelian groups with n-th component $\mathcal{H}\mathrm{om}_R(C, D)_n$ and boundary operator

$$\delta_n((\varphi_i)_{i\in\mathbb{Z}}) = (\delta^D_{n+i}\varphi_i - (-1)^n \varphi_{i-1}\delta^C_i)_{i\in\mathbb{Z}}.$$

A homomorphism $\varphi \in \mathcal{H}\mathrm{om}_R(C, D)_n$ is called a chain map if $\delta(\varphi) = 0$, that is, if $\delta^D_{n+i}\varphi_i = (-1)^n \varphi_{i-1}\delta^C_i$ for all $i \in \mathbb{Z}$. A chain map of degree 0 is called a morphism.

To every complex C we associate the numbers

$$\sup C = \sup\{i \mid C_i \neq 0\}, \quad \inf C = \inf\{i \mid C_i \neq 0\}.$$

The complex C is called bounded above when $\sup C < \infty$, bounded below when $\inf C > -\infty$ and bounded when it is bounded below and above.

For objects C and D of $\mathrm{Ch}(R)$ (R-Mod), $\mathrm{Hom}(C, D)$ ($\mathrm{Hom}_R(C, D)$) is the abelian group of morphisms from C to D in $\mathrm{Ch}(R)$ (R-Mod) and $\mathrm{Ext}^i(C, D)$ ($\mathrm{Ext}^i_R(C, D)$) for $i \geq 1$ will denote the groups we get from the right derived functor of $\mathrm{Hom}(C, D)$ ($\mathrm{Hom}_R(C, D)$).

Let $\mathcal{A}, \mathcal{B}$ be two classes of R-modules. The pair $(\mathcal{A}, \mathcal{B})$ is called a cotorsion pair (also called a cotorsion theory) if $\mathcal{A}^\perp = \mathcal{B}$ and $\mathcal{A} = {}^\perp\mathcal{B}$. Here $\mathcal{A}^\perp$ is the class of R-modules C such that $\mathrm{Ext}^1(A, C) = 0$ for all $A \in \mathcal{A}$, and similarly ${}^\perp\mathcal{B}$ is the class of R-modules C such that $\mathrm{Ext}^1(C, B) = 0$ for all $B \in \mathcal{B}$. A cotorsion pair $(\mathcal{A}, \mathcal{B})$ is said to be hereditary, if whenever $0 \to \widetilde{A} \to A \to \widehat{A} \to 0$ is exact with $A, \widehat{A} \in \mathcal{A}$ then $\widetilde{A}$ is also in $\mathcal{A}$, or equivalently, if $0 \to \widetilde{B} \to B \to \widehat{B} \to 0$ is exact with $\widetilde{B}, B \in \mathcal{B}$ then $\widehat{B}$ is also in $\mathcal{B}$. A cotorsion pair $(\mathcal{A}, \mathcal{B})$ is cogenerated by a set $\mathcal{S} \subseteq \mathcal{A}$ if $\mathcal{B} = \mathcal{S}^\perp$. A cotorsion pair $(\mathcal{A}, \mathcal{B})$ is said to have enough injectives (projectives) [[10]] if for any object M there exists an exact sequence $0 \to M \to B \to A \to 0$ ($0 \to B \to A \to M \to 0$) with $A \in \mathcal{A}$ and $B \in \mathcal{B}$. By [[10], Proposition 1.1.5], a cotorsion pair $(\mathcal{A}, \mathcal{B})$ has enough projectives if and only if it has enough injectives. The cotorsion pair $(\mathcal{A}, \mathcal{B})$ is called complete if it has enough projectives and injectives.

Given a class $\mathcal{B}$ of objects of $\mathrm{Ch}(R)$, a morphism $\phi : X \to B$ is called a $\mathcal{B}$-preenvelope ([[6]]) if $B \in \mathcal{B}$ and $\mathrm{Hom}(B, B') \to \mathrm{Hom}(X, B') \to 0$ is exact for all $B' \in \mathcal{B}$. If, moreover, any $f : B \to B$ such that $f\phi = \phi$ is an automorphism of B then $\phi : X \to B$ is called a $\mathcal{B}$-envelope of X. A complex X is said to have a special $\mathcal{B}$-preenvelope [[9]] if there is an exact sequence $0 \to X \to B \to L \to 0$ with $B \in \mathcal{B}$ and $L \in {}^\perp\mathcal{B}$. (Special) precovers and covers of X are defined dually.

2. Ding t-projective and Ding t-injective model structures. Ding and Chen extended FC rings to n-FC rings [[2], [3]], which are seen to have many properties similar to those of n-Gorenstein rings. Just as a ring is called Gorenstein when it is n-Gorenstein for some nonnegative integer n (a ring R is called n-Gorenstein if it is a left and right Noetherian ring with self injective dimension at most n on both sides for some nonnegative integer n), Gillespie first called a ring Ding-Chen when it is n-FC for some n [[13], Definition 4.1]. An R-module M is called Ding projective if there exists an exact sequence of projective R-modules $\cdots \longrightarrow P_1 \longrightarrow P_0 \longrightarrow P_{-1} \longrightarrow P_{-2} \longrightarrow \cdots$ with $M = \mathrm{Ker}(P_0 \longrightarrow P_{-1})$ and which remains exact after applying $\mathrm{Hom}(-, F)$ for any flat R-module F [[5]]. The class of Ding projective R-modules is denoted by $\mathcal{DP}$. An R-module N is called Ding injective if there exists an exact sequence of injective R-modules $\cdots \longrightarrow I_1 \longrightarrow I_0 \longrightarrow I_{-1} \longrightarrow I_{-2} \longrightarrow \cdots$ with $N = \mathrm{Ker}(I_0 \longrightarrow I_{-1})$ and which remains exact after applying $\mathrm{Hom}(E, -)$ for any FP-injective R-module E [[17]]. The class of Ding injective R-modules is denoted by $\mathcal{DI}$. Note that every Ding injective (respectively, Ding projective) R-module N is Gorenstein injective (respectively, Gorenstein projective), and if R is Gorenstein, then every Gorenstein injective R-module is Ding injective (respectively, Gorenstein projective)[[13]].

From [[13], Theorem 4.2], we know that for a Ding-Chen ring R, the class of all modules with finite flat dimension and the class of all modules with finite FP-injective dimension are the same, and we use $\mathcal{W}_R$ to denote this class throughout this section.

Ding and Mao proved that $({}^\perp\mathcal{W}_R, \mathcal{W}_R)$ forms a complete cotorsion pair when R is a Ding-Chen ring [[4], Theorem 3.8]. Also, $(\mathcal{W}_R, \mathcal{W}_R^\perp)$ forms a complete cotorsion pair when R is a Ding-Chen ring [[16], Theorem 3.4]. Moreover, Gillespie proved that an R-module M is Ding projective if and only if $M \in^{\perp} \mathcal{W}_R$, an R-module N is Ding injective if and only if $N \in \mathcal{W}_R^{\perp}$ [[13], Corollaries 4.5 and 4.6]. So $(\mathcal{DP}, \mathcal{W}_R)$ and $(\mathcal{W}_R, \mathcal{DI})$ are complete hereditary cotorsion pairs (each cogenerated by a set). Hence for every $M \in R$-Mod there exists an epimorphism $D_0 \longrightarrow M$ where D_0 is a Ding projective module. This allows us to construct an exact sequence

$$\cdots \longrightarrow D_k \longrightarrow \cdots \longrightarrow D_1 \longrightarrow D_0 \longrightarrow M \longrightarrow 0,$$

where D_k is a Ding projective module, for every $k \geq 0$. We shall say that this sequence is a *left Ding projective resolution* of M. An R-module M is said to be *Ding t-projective*, if M admits a *left Ding projective resolution* of length at most t (that is, M has Ding projective dimension at most t), where t is a nonnegative integer. Let $\mathcal{DP}_t$ denote the class of Ding t-projective modules. We shall denote by $\mathrm{Dpd}(M)$ the (left) Ding projective dimension of M. Note that $\mathcal{DP}_t = \{M \in R\text{-Mod} : \mathrm{Dpd}(M) \leq t\}$ and that $\mathcal{DP}_0 = \mathcal{DP}$, similarly, we let $\mathcal{P}_t$ denote the class of t-projective R-modules.

Similarly, we can define Ding t-injective modules, and we let $\mathcal{DI}_t$ denote the class of Ding t-injective modules and $\mathcal{I}_t$ the class of t-injective R-modules.

Let

$$\cdots \longrightarrow P_n \xrightarrow{f_n} P_{n-1} \longrightarrow \cdots \longrightarrow P_1 \xrightarrow{f_1} P_0 \xrightarrow{f_0} X \longrightarrow 0, \quad (1)$$

be a projective resolution of X. We shall say that $\mathrm{Im}(f_i)$ is the i-th projective syzygy of X in (1). We shall use the notation $\Omega^i(X)$ for the class of all i-th projective syzygies of X. Dually, given an injective coresolution of X, say

$$0 \longrightarrow X \longrightarrow I^0 \xrightarrow{f^0} I^1 \xrightarrow{f^1} \cdots \longrightarrow I^{n-1} \xrightarrow{f^{n-1}} I^n \longrightarrow \cdots, \quad (2)$$

we shall say that $\mathrm{Ker}(f^i)$ is the i-th injective cosyzygy of X in (2), and we shall use the notation $\Omega^{-i}(X)$ for the class of all i-th injective cosyzygies of X.

We begin with the following result.

1.1. Lemma ([[5], Lemma 3.4]). *Let R be a Ding-Chen ring. Then the following are equivalent:*
(1) *M is Ding t-projective.*
(2) $\mathrm{Ext}_R^i(M, W) = 0$ *for all $i > t$ and for all $W \in \mathcal{W}_R$.*
(3) $\mathrm{Ext}_R^{t+1}(M, W) = 0$ *for all $W \in \mathcal{W}_R$.*
(4) *Every tth Ding projective syzygy of M is Ding projective.*
(5) *Every tth projective syzygy of M is Ding projective.*

1.2. Corollary. *Let R be an n-FC ring. Then for every $0 \leq t \leq n, \mathcal{DP}_t \cap \mathcal{W}_R = \mathcal{P}_t$.*

Proof The inclusion $\mathcal{P}_t \subseteq \mathcal{DP}_t \cap \mathcal{W}_R$ is clear. Now let $M \in \mathcal{DP}_t \cap \mathcal{W}_R$. Then every $G \in \Omega^t(M)$ is in $\mathcal{DP}$ by Lemma 1.1. Since $M \in \mathcal{W}_R$, we have $G \in \mathcal{W}_R$. Then $G \in \mathcal{DP} \cap \mathcal{W}_R = \mathcal{P}_0$ by [[5], Lemma 2.4]. It follows $M \in \mathcal{P}_t$. □

The following results show that $(\mathcal{DP}_t, (\mathcal{DP}_t)^{\perp})$ is a complete cotorsion pair for every $1 \leq t \leq n$.

1.3. Theorem. *Let R be an n-FC ring. $(\mathcal{DP}_t, (\mathcal{DP}_t)^{\perp})$ is a cotorsion pair cogenerated by a set, and so it is complete for every $1 \leq t \leq n$.*

Proof First we prove that $(\mathcal{DP}_t, (\mathcal{DP}_t)^{\perp})$ is a cotorsion pair.

It suffices to show that ${}^{\perp}((\mathcal{DP}_t)^{\perp}) \subseteq \mathcal{DP}_t$. Let $M \in^{\perp} ((\mathcal{DP}_t)^{\perp})$. Consider a left partial projective resolution of M, say $0 \longrightarrow G \longrightarrow P_{t-1} \longrightarrow \cdots \longrightarrow P_1 \longrightarrow P_0 \longrightarrow M \longrightarrow 0$.

By Lemma 1.1, it suffices to show that G is a Ding projective module. Suppose $t = 1$ and let $W \in \mathcal{W}_R$. We have the exact sequence

$$\cdots \longrightarrow \mathrm{Ext}^1_R(P_0, W) \longrightarrow \mathrm{Ext}^1_R(G, W) \longrightarrow \mathrm{Ext}^2_R(M, W) \longrightarrow \cdots,$$

where $\mathrm{Ext}^1_R(P_0, W) = 0$, since P_0 is projective. On the other hand, $\mathrm{Ext}^2_R(M, W) = \mathrm{Ext}^1_R(M, L)$, where $L \in \Omega^{-1}(W)$. We show $L \in (\mathcal{DP}_1)^\perp$. Let $K \in \mathcal{DP}_1$ and consider the short exact sequence $0 \longrightarrow W \longrightarrow I \longrightarrow L \longrightarrow 0$, where I is injective. Then we have an exact sequence

$$\cdots \longrightarrow \mathrm{Ext}^1_R(K, I) \longrightarrow \mathrm{Ext}^1_R(K, L) \longrightarrow \mathrm{Ext}^2_R(K, W) \longrightarrow \cdots,$$

where $\mathrm{Ext}^1_R(K, I) = 0$, since I is injective, and $\mathrm{Ext}^2_R(K, W) = 0$ since $K \in \mathcal{DP}_1$ and $W \in \mathcal{W}_R$. Then $\mathrm{Ext}^1_R(K, L) = 0$ for every $K \in \mathcal{DP}_1$, i.e. $L \in (\mathcal{DP}_1)^\perp$. It follows $\mathrm{Ext}^2_R(M, W) = 0$. Hence $\mathrm{Ext}^1_R(G, W) = 0$ for every $W \in \mathcal{W}_R$, i.e. $G \in \mathcal{DP}$.

Suppose the result is true for every $1 \leq j \leq t-1$. We have an exact sequence

$$0 \longrightarrow G \longrightarrow P_{t-1} \longrightarrow \cdots \longrightarrow P_1 \longrightarrow L \longrightarrow 0,$$

where $L \in \Omega^1(M)$, and a short exact sequence $0 \longrightarrow L \longrightarrow P_0 \longrightarrow M \longrightarrow 0$. Let $K \in (\mathcal{DP}_{t-1})^\perp$. We have $\mathrm{Ext}^1_R(L, K) \cong \mathrm{Ext}^1_R(M, K')$, where $K' \in \Omega^{-1}(K)$. Let $N \in \mathcal{DP}_t$. Then $N' \in \mathcal{DP}_{t-1}$, for every $N' \in \Omega^1(N)$. We have $\mathrm{Ext}^1_R(N, K') \cong \mathrm{Ext}^1_R(N', K) = 0$. So $K' \in (\mathcal{DP}_t)^\perp$. It follows $\mathrm{Ext}^1_R(L, K) \cong \mathrm{Ext}^1_R(M, K') = 0$, for every $K \in (\mathcal{DP}_{t-1})^\perp$. Hence $L \in {}^\perp((\mathcal{DP}_{t-1})^\perp) = \mathcal{DP}_{t-1}$. It follows $M \in \mathcal{DP}_t$.

Now we prove that $(\mathcal{DP}_t, (\mathcal{DP}_t)^\perp)$ is a cotorsion pair cogenerated by a set.

Consider the cogenerating set $\mathcal{U}$ of $(\mathcal{DP}, \mathcal{W}_R)$. On the other hand, it is known that $(\mathcal{P}_t, (\mathcal{P}_t)^\perp)$ is cogenerated by the set $\mathcal{P}_t^{\leq\kappa} := \{M \in \mathcal{P}_t : \mathrm{Card}(M) \leq \kappa\}$, where $\kappa \geq \mathrm{Card}(R)$ is a fixed infinite cardinal number. Set $\mathcal{G}_t := \mathcal{U} \cup \mathcal{P}_t^{\leq\kappa}$. We prove $(\mathcal{DP}_t)^\perp = (\mathcal{G}_t)^\perp$. Since $\mathcal{G}_t \subseteq \mathcal{DP}_t$, we have $(\mathcal{DP}_t)^\perp \subseteq (\mathcal{G}_t)^\perp$. Now let $N \in (\mathcal{G}_t)^\perp$, and consider $M \in \mathcal{DP}_t$. Since $(\mathcal{DP}, \mathcal{W}_R)$ is a complete cotorsion pair, there exists a short exact sequence $0 \longrightarrow M \longrightarrow W \longrightarrow G \longrightarrow 0$, where $W \in \mathcal{W}_R$ and $G \in \mathcal{DP}$. Then $W \in \mathcal{DP}_t \cap \mathcal{W}_R = \mathcal{P}_t$ by Corollary 1.2. We apply the contravariant functor $\mathrm{Ext}(-, N)$ and obtain a long exact sequence

$$\cdots \longrightarrow \mathrm{Ext}^1_R(W, N) \longrightarrow \mathrm{Ext}^1_R(M, N) \longrightarrow \mathrm{Ext}^2_R(G, N) \longrightarrow \cdots.$$

Note that $\mathrm{Ext}^2_R(G, N) = 0$, since $N \in (\mathcal{G}_t)^\perp \subseteq \mathcal{U}^\perp = \mathcal{W}_R$ and $(\mathcal{DP}, \mathcal{W}_R)$ is hereditary. On the other hand, $N \in (\mathcal{P}_t^{\leq\kappa})^\perp = (\mathcal{P}_t)^\perp$ and $W \in \mathcal{P}_t$, so $\mathrm{Ext}^1_R(W, N) = 0$. Hence $\mathrm{Ext}^1_R(M, N) = 0$ for every $M \in \mathcal{DP}_t$, i.e. $N \in (\mathcal{DP}_t)^\perp$. □

This gives the following result.

1.4. Corollary. *Let R be an n-FC ring. Then $(\mathcal{DP}_t)^\perp = \mathcal{W}_R \cap (\mathcal{P}_t)^\perp$ for every $1 \leq t \leq n$.*

Proof Since $\mathcal{DP} \subseteq \mathcal{DP}_t$ and $\mathcal{P}_t \subseteq \mathcal{DP}_t$, we get $(\mathcal{DP}_t)^\perp \subseteq (\mathcal{DP})^\perp = \mathcal{W}_R$ and $(\mathcal{DP}_t)^\perp \subseteq (\mathcal{P}_t)^\perp$, and so $(\mathcal{DP}_t)^\perp \subseteq \mathcal{W}_R \cap (\mathcal{P}_t)^\perp$. Let $N \in \mathcal{W}_R \cap (\mathcal{P}_t)^\perp$. Since $(\mathcal{DP}_t, (\mathcal{DP}_t)^\perp)$ is complete by Theorem 1.3, there exists a short exact sequence $0 \longrightarrow N \longrightarrow K \longrightarrow C \longrightarrow 0$ where $K \in (\mathcal{DP}_t)^\perp$ and $C \in \mathcal{DP}_t$. Since $N, K \in \mathcal{W}_R$, $C \in \mathcal{W}_R$. Then $C \in \mathcal{DP}_t \cap \mathcal{W}_R = \mathcal{P}_t$ by Corollary 1.2 and hence $\mathrm{Ext}^1_R(C, N) = 0$. It follows $K \cong N \oplus C$. Since $(\mathcal{DP}_t)^\perp$ is closed under direct summands and $K \in (\mathcal{DP}_t)^\perp$, we get $N \in (\mathcal{DP}_t)^\perp$. □

1.5. Definition. *Given two cotorsion pairs $(\mathcal{A}, \mathcal{B}')$ and $(\mathcal{A}', \mathcal{B})$ in an abelian category, we shall say that they are compatible if there exists a class of objects $\mathcal{W}$ such that $\mathcal{A}' = \mathcal{A} \cap \mathcal{W}$ and $\mathcal{B}' = \mathcal{B} \cap \mathcal{W}$.*

1.6. Lemma. ***(Hovey's criterion)*** *Let $(\mathcal{A}, \mathcal{B} \cap \mathcal{W})$ and $(\mathcal{A} \cap \mathcal{W}, \mathcal{B})$ be two compatible cotorsion pairs in a bicomplete abelian category $\mathcal{C}$ with enough projective and injective* *objects, where the class $\mathcal{W}$ is thick. Then there exists a unique abelian model structure on $\mathcal{C}$ such that $\mathcal{A}$ is the class of cofibrant objects, $\mathcal{B}$ is the class of fibrant objects, and $\mathcal{W}$ is the class of trivial objects.*

From the above results, there exists a unique abelian model strucutre on R-Mod where $\mathcal{DP}_t$ is the class of cofibrant objects, $(\mathcal{P}_t)^{\perp}$ is the class of fibrant objects, and $\mathcal{W}_R$ is the class of trivial objects. We call this structure the Ding t-projective model structure on R-Mod. Similarly, there is a unique abelian model structure on R-Mod such that ${}^{\perp}(\mathcal{I}_t)$ is the class of cofibrant objects, $\mathcal{DI}_t$ is the class of fibrant objects, and $\mathcal{W}_R$ is the class of trivial objects. We call this structures the Ding t-injective model structure on R-Mod.

We also have the following result.

1.7. Proposition. *Let X be a chain complex bounded below. Then X is Ding t-projective if and only if X_m is a Ding t-projective module for every $m \in \mathbb{Z}$.*

Proof Let X be a Ding t-projective chain complex. Then there exists an exact sequence in $\mathrm{Ch}(R)$

$$0 \longrightarrow D^t \longrightarrow D^{t-1} \longrightarrow \cdots \longrightarrow D^1 \longrightarrow D^0 \longrightarrow X \longrightarrow 0,$$

such that D^i is a Ding projective complex for every $0 \leq i \leq t$. For each $m \in \mathbb{Z}$, we have an exact sequence in R-Mod

$$0 \longrightarrow D^t_m \longrightarrow D^{t-1}_m \longrightarrow \cdots \longrightarrow D^1_m \longrightarrow D^0_m \longrightarrow X_m \longrightarrow 0.$$

Since each D^i is a Ding projective complex, we have that D^i_m is a Ding projective module. So the previous exact sequence turns out to be a right Ding projective resolution of X_m of length t, i.e. $X_m \in \mathcal{DP}_t$.

Now suppose that X_m is a Ding t-projective module for every $m \in \mathbb{Z}$. Consider a partial left projective resolution

$$0 \longrightarrow D^t \longrightarrow P^{t-1} \longrightarrow \cdots \longrightarrow P^1 \longrightarrow P^0 \longrightarrow X \longrightarrow 0.$$

It suffices to show that D^t is a Ding projective chain complex. For each $m \in \mathrm{Z}$, we have a exact sequence

$$0 \longrightarrow D^t_m \longrightarrow P^{t-1}_m \longrightarrow \cdots \longrightarrow P^1_m \longrightarrow P^0_m \longrightarrow X_m \longrightarrow 0.$$

Note that each P^i_m is a projective module. Since $X_m \in \mathcal{DP}_t$, we have $D^t_m \in \Omega^t(X_m) \in \mathcal{DP}$. Hence D^t is a Ding projective complex by [[20], Proposition 3.14]. $\square$

1.8. Definition ([[12], Definition 3.3]). *Let $(\mathcal{A}, \mathcal{B})$ be a cotorsion pair in R-Mod and X an R-complex.*

(1) *X is called an $\mathcal{A}$ complex if it is exact and $\mathrm{Z}_n X \in \mathcal{A}$ for all $n \in \mathbb{Z}$.*

(2) *X is called a $\mathcal{B}$ complex if it is exact and $\mathrm{Z}_n X \in \mathcal{B}$ for all $n \in \mathbb{Z}$.*

(3) *X is called a dg-$\mathcal{A}$ complex if $X_n \in \mathcal{A}$ for each $n \in \mathbb{Z}$, and $\mathcal{H}\mathrm{om}_R(X, B)$ is exact whenever B is a $\mathcal{B}$ complex.*

(4) *X is called a dg-$\mathcal{B}$ complex if $X_n \in \mathcal{B}$ for each $n \in \mathbb{Z}$, and $\mathcal{H}\mathrm{om}_R(A, X)$ is exact whenever A is an $\mathcal{A}$ complex.*

We denote the class of $\mathcal{A}$ complexes by $\widetilde{\mathcal{A}}$ and the class of dg-$\mathcal{A}$ complexes by $dg\widetilde{\mathcal{A}}$. Similarly, the class of $\mathcal{B}$ complexes is denoted by $\widetilde{\mathcal{B}}$ and the class of dg-$\mathcal{B}$ complexes is denoted by $dg\widetilde{\mathcal{B}}$.

As we did in the category R-Mod, we can prove that $(\widetilde{\mathcal{DP}_t}, (\widetilde{\mathcal{DP}_t})^{\perp})$ and $({}^{\perp}(\widetilde{\mathcal{DI}_t}), \widetilde{\mathcal{DI}_t})$ are complete cotorsion pairs. Moreover, we can see that $(\widetilde{\mathcal{DP}_t}, (\widetilde{\mathcal{DP}_t})^{\perp})$ and $(\widetilde{\mathcal{P}}_t, (\widetilde{\mathcal{P}}_t)^{\perp})$ are compatible. So there exists a unique abelian model structure on $\mathrm{Ch}(R)$ such that $\widetilde{\mathcal{DP}_t}$ is the class of cofibrant objects, $(\widetilde{\mathcal{P}}_t)^{\perp}$ is the class of fibrant objects, and $\widetilde{\mathcal{W}}_R$ is the class of trivial objects. Similarly, there is a unique abelian model structure on $\mathrm{Ch}(R)$ such that ${}^{\perp}(\widetilde{\mathcal{I}}_t)$ is the class of cofibrant objects, $\widetilde{\mathcal{DI}}_t$ is the class of fibrant objects, and $\widetilde{\mathcal{W}}_R$ is the class of trivial objects. We call these structures the Ding t-projective model structure and the Ding t-injective model structure on $\mathrm{Ch}(R)$, respectively.

3. Ding homological dimensions over graded rings. A $\mathbb{Z}$-graded ring A is a ring that has a direct sum decomposition into (abelian) additive groups $A = \oplus_{n\in\mathbb{Z}} A_n = \cdots \oplus A_{-1} \oplus A_0 \oplus A_1 \oplus \cdots$ such that the ring multiplication $\cdot$ satisfies $A_m \cdot A_n \subseteq A_{m+n}$, for every $m, n \in \mathbb{Z}$. A graded module is left module over a $\mathbb{Z}$-graded ring A with a direct sum decomposition $M = \oplus_{n\in\mathbb{Z}} M_n$ such that the product $\cdot : A \cdot M \to M$ satisfies $A_m \cdot M_n \subseteq M_{m+n}$, for every $m,\ n \in \mathbb{Z}$.

Given any associative ring with unit R, consider the ring of polynomials $R[x]$ and the ideal (x^2). It is easy to see that the quotient $A := R[x]/(x^2)$ is a $\mathbb{Z}$-graded ring with a direct sum decomposition given by $R[x]/(x^2) = \cdots \oplus 0 \oplus (x) \oplus R \oplus 0 \oplus \cdots$, where the scalars $r \in R$ are the elements of degree 0, and the elements in the ideal (x) form the terms of degree -1. The following we will check that the category A-Mod is isomorphic to the category $\mathrm{Ch}(R)$ of unbounded R-chain complexes. Through this isomorphism, the A-module A corresponds to $D^0(R)$. In particular, we have $\mathrm{Ext}^i_A(-, A) \cong \mathrm{Ext}^i_{\mathrm{Ch}(R)}(-, D^0(R))$.

Now we prove that every A-module can be viewed as a chain complex over R, and vice versa.

Let $\Phi : A\text{-Mod} \longrightarrow \mathrm{Ch}(R)$ be the application defined as follows:
(1) Given a graded A-module $M = \oplus_{n\in\mathbb{Z}} M_n$, note that if $y \in M_n$ then $x \cdot y \in M_{n-1}$, since x has degree -1. Denote by $\Phi(M)_n$ the set M_n endowed with the structure of R-module provided by the graded multiplication. Let $\partial_n : \Phi(M)_n \longrightarrow \Phi(M)_{n-1}$ be the map $y \mapsto x \cdot y$. It is clear that ∂_n is an R-homomorphism. Moreover, $\partial_{n-1} \circ \partial_n(y) = x \cdot (x \cdot y) = x^2 \cdot y = 0 \cdot y = 0$. Then, $\Phi(M) = (\Phi(M)_n, \partial_n)_{n\in\mathrm{Z}}$ is a chain complex over R.
(2) Let $f : M \longrightarrow N$ be a homomorphism of graded A-modules. Then $f(M_n) \subseteq N_n$, for every $n \in \mathrm{Z}$. It follows that $f|_{M_n}$ is an R-homomorphism. Let $\Phi(f)_n := f|_{M_n} : \Phi(M)_n \longrightarrow \Phi(N)_n$. We have $\Phi(f)_{n-1} \circ \partial_n^M(y) = f|_{M_{n-1}}(x \cdot y) = x \cdot f|_{M_n}(y) = \partial_n^N \circ \Phi(f)_n(y)$. So $\Phi(f) = (\Phi(f)_n)_{n\in\mathbb{Z}}$ is a chain map.

Note that $\Phi : A\text{-Mod} \longrightarrow \mathrm{Ch}(R)$ defines a covariant functor. We show that this functor is an isomorphism, by giving an inverse functor $\Psi : \mathrm{Ch}(R) \longrightarrow A\text{-Mod}$.
(1) Let $M = (M_n, \partial_n)_{n\in\mathbb{Z}}$ be a chain complex over R. Let $y \in M_n$ and define the product $r \cdot y = ry \in M_n$ for every $r \in R$, and $x \cdot y = \partial_n(y) \in M_{n-1}$. This gives rise to a graded A-module, that we denote by $\Psi(M) = (\Psi(M)_n)_{n\in\mathbb{Z}}$, where $\Psi(M)_n = M_n$ as sets.
(2) Given a chain map $f : M \longrightarrow N$, we have $x \cdot f(y) = \partial \circ f(y) = f \circ \partial(y) = f(x \cdot y)$. Then f gives rise to a graded A-module homomorphism denoted by $\Phi(f)$.

It is easy to show that $\Psi \circ \Phi = \mathrm{Id}_{A\text{-Mod}}$ and $\Phi \circ \Psi = \mathrm{Id}_{\mathrm{Ch}(R\text{-Mod})}$. It follows that Ψ and Φ map projective (resp., injective, flat) objects into projective (resp., injective, flat) objects. It is also easy to check that both Ψ and Φ preserves exact sequences.

2.1. Definition ([[8]]). *A complex C is called finitely generated if, in case $C = \sum_{i\in I} D^i$, with $D^i \in \mathrm{Ch}(R)$ subcomplexes of C, then there exists a finite subset $J \subset I$ such that $C = \sum_{i\in J} D^i$; A complex C is called finitely presented if C is finitely generated and for every exact sequence of complexes $0 \to K \to L \to C \to 0$ with L finitely generated, K is also finitely generated.*

2.2. Lemma ([[8]]). *An R-complex C is finitely generated if and only if C is bounded and C_n is finitely generated in R-Mod for all $n \in \mathbb{Z}$. A complex C is finitely presented if and only if C is bounded and C_n is finitely presented in R-Mod for all $n \in \mathbb{Z}$.*

It is obvious that Ψ and Φ map finitely presented objects into finitely presented objects by Lemma 2.2. Next, we prove Ψ and Φ map FP-injective objects into FP-injective objects.

2.3. Lemma *Let E be an FP-injective A-module, and Y be an FP-injective R-complex. Then $\Phi(E)$ is an FP-injective R-complex, and $\Psi(Y)$ is an FP-injective A-module.*

Proof We prove the first assertion, the second one can be proven similarly. Let F be a finitely presented A-module. We first prove that $\mathrm{Ext}^i(\Phi(F), \Phi(E)) \cong \mathrm{Ext}^i_A(F, E) = 0$ for every $i \geq 1$. Given a class $[0 \longrightarrow E \longrightarrow Q \longrightarrow F \longrightarrow 0] \in \mathrm{Ext}^1_A(F, E)$, map its representative to the sequence

$$0 \longrightarrow \Phi(E) \longrightarrow \Phi(Q) \longrightarrow \Phi(F) \longrightarrow 0.$$

This sequence is exact since Φ is an exact functor. Also, Φ preserves pullbacks, and hence it preserves Baer sums. It follows

$$[0 \longrightarrow \Phi(E) \longrightarrow \Phi(Q) \longrightarrow \Phi(F) \longrightarrow 0] \in \mathrm{Ext}^1(\Phi(F), \Phi(E)).$$

It is clear that this mapping defines a group isomorphism from $\mathrm{Ext}^1_A(F, E)$ to $\mathrm{Ext}^1(\Phi(F), \Phi(E))$. The same argument works for any $i > 1$. Since $\Phi(F)$ is a finitely presented R-complex, $\Phi(E)$ is an FP-injective R-complex. □

2.4. Proposition *If R is an n-FC ring, then the graded ring $A := R[x]/(x^2)$ is n-FC with weak global dimension ∞.*

Proof Any homogeneous left (resp. right) ideal of A is of the form $I_0 + I_1x$, where I_0 and I_1 are left (resp. right) ideals of R. Let $I_0 + I_1x$ be finitely generated. So I_0, I_1 is finitely generated. Since R is left (right) coherent, I_0, I_1 is finitely presented. Hence $I_0 + I_1x$ is finitely presented. Hence A is left and right coherent. If $M = \oplus_{n\in\mathbb{Z}} M_n$ is a finitely presented A-module, then M_n is a finitely presented R-module for every $n \in \mathbb{Z}$. Since

$$\mathrm{Ext}^i_A(-, A) \cong \mathrm{Ext}^i_{\mathrm{Ch}(R)}(-, D^0(R)),$$

and

$$\mathrm{Hom}_{\mathrm{Ch}(R)}(X, D^0(R)) \cong \mathrm{Hom}_R(X_{-1}, R),$$

where X_{-1} is the degree -1 part of X. Since this functor $(-)_{-1}$ is exact and preserves projectives, we see that

$$\mathrm{Ext}^i_{\mathrm{Ch}(R)}(-, D^0(R)) \cong \mathrm{Ext}^i_R((-)_{-1}, R).$$

In particular, if R is n-FC, so is A.

Since flat chain complexes are exact, any chain complex that is not exact must have infinite flat dimension, so the weak global dimension of A is ∞. □

We say a chain complex X is projective (resp., injective, flat, FP-injective) if it is exact and each cycle Z_nX is projective (resp., injective, flat, FP-injective). We denote these classes of chain complexes by $\widetilde{\mathcal{P}}, \widetilde{\mathcal{I}}, \widetilde{\mathcal{F}}$, and $\widetilde{\mathcal{FI}}$ respectively.

2.5. Lemma *If R is an n-FC ring, then the class of chain complexes with finite FP-injective dimension and the class of chain complexes with finite flat dimension coincide and every exact complex E with cycles of finite flat (FP-injective) dimension has* $\mathrm{fd}(E) \leq n$ *(FP-*$\mathrm{id}(E) \leq n$*).*

Proof From [[19], Theorem 2.26], we know that the class of chain complexes with finite FP-injective (flat) dimension is the class of exact complexes with cycles of bounded FP-injective (flat) dimension. If R is n-FC, then these classes coincide. □

By Proposition 2.4, for an n-FC ring R and $A := R[x]/(x^2)$, the class $\mathcal{W}_A$ must correspond to some collection of chain complexes. Next we will characterize these chain complexes.

2.6. Corollary. *Let R be left and right coherent with finite weak global dimension. Then $\mathcal{W}_A$ corresponds the class of all exact complexes.*

Proof By [[18], Proposition 3.5], [[13], Theorem 4.2] and Lemma 2.5 the conclusion is obvious. □

Recall from [[7]] that a complex P is said to be dg-projective if each P_m is projective and $\mathcal{H}om_R(P, E)$ is exact for any exact complex E. A dg-injective complex is defined dually.

Now we get the following result.

2.7. Proposition. *Suppose R is a ring and let A be the graded ring $R[x]/(x^2)$. Then every dg-projective chain complex over R is a Ding projective A-module. The converse holds if R is left and right coherent and of finite weak global dimension.*

Proof Suppose X is a dg-projective chain complex. We want to show that it is a Ding projective A-module. We first take a projective resolution of X

$$\cdots \longrightarrow P^1 \longrightarrow P^0 \longrightarrow X \longrightarrow 0.$$

Note that since X is dg-projective, the kernel at any spot in the sequence is also dg-projective. Next we use the fact that $(dg\widetilde{\mathcal{P}}, \mathcal{E})$ is complete to find a short exact sequence $0 \longrightarrow X \longrightarrow P_0 \longrightarrow K \longrightarrow 0$ where P_0 is exact and K is dg-projective. But P_0 must also be dg-projective since it is an extension of two dg-projective complexes. Therefore P_0 is a projective complex. Continuing with the same procedure on K we can build a projective coresolution of X as below:

$$0 \longrightarrow X \longrightarrow P_0 \longrightarrow P_1 \longrightarrow P_2 \longrightarrow \cdots.$$

Again the kernel at each spot is dg-projective. Pasting this "right" coresolution together with the "left" resolution above we get an exact sequence

$$\cdots \longrightarrow P^1 \longrightarrow P^0 \longrightarrow P_0 \longrightarrow P_1 \longrightarrow \cdots$$

of projective complexes which satisfies the definition of X being a Ding projective A-module. Indeed since X dg-projective implies $\mathrm{Ext}^1(X, E) = 0$ for any exact chain complex E, we certainly have $\mathrm{Ext}^1(X, F) = 0$ for any flat chain complex F. Therefore applying $\mathrm{Hom}_A(-, F)$ will leave the sequence exact.

Next we let X be a Ding projective A-module and argue that it is a dg-projective R-chain complex, when R is both left and right Coherent and $\mathrm{wD.dim}(R) = n$. Note that by the definition of Ding projective we have $\mathrm{Ext}^i(X, F) = 0$ for all $i > 0$ and flat complexes F. We will be done if we can show that $\mathrm{Ext}^1(X, E) = 0$ for any exact complex E. By Corollary 2.6 $\mathrm{fd}(E) \leq n$, so there exists a finite flat resolution

$$0 \longrightarrow F^n \longrightarrow \cdots \longrightarrow F^1 \longrightarrow F^0 \longrightarrow E \longrightarrow 0.$$

By a dimension shifting argument we see that $\mathrm{Ext}^1(X, E) \cong \mathrm{Ext}^{n+1}(X, F^n) = 0$. □

With a dual proof we get the following.

2.8. Proposition. *Suppose R is a ring and let A be the graded ring $R[x]/(x^2)$. Then every dg-injective chain complex over R is a Ding injective A-module. The converse holds if R is left and right coherent and of finite weak global dimension.*

Now we extend Proposition 2.7 as follows.

2.9. Theorem. *The functor* $\Psi : \mathrm{Ch}(R) \longrightarrow A\text{-Mod}$ *maps dg-t-projective complexes into Ding t-projective A-modules. If R is a left and right coherent ring of finite weak global dimension, then the inverse functor* $\Phi : A\text{-Mod} \longrightarrow \mathrm{Ch}(R)$ *maps Ding t-projective A-modules into dg-t-projective complexes.*

Proof Let $X \in dg\widetilde{\mathcal{P}_t}$. Consider $\Psi(X)$ and a partial left projective resolution

$$0 \longrightarrow G \longrightarrow P_{t-1} \longrightarrow \cdots \longrightarrow P_1 \longrightarrow P_0 \longrightarrow \Psi(X) \longrightarrow 0.$$

We show that G is a Ding projective A-module. Consider the complex $\Phi(G)$ and let E be an exact complex. By the proof of Lemma 2.3 we have $\mathrm{Ext}^1(\Phi(G), E) \cong \mathrm{Ext}^1(X, E')$, where $E' \in \Omega^{-t}(E)$. Note that $E' \in (\widetilde{\mathcal{P}_t})^{\perp}$. In fact, if $Z \in \widetilde{\mathcal{P}_t}$ then $\mathrm{Ext}^1(Z, E') \cong \mathrm{Ext}^{t+1}(Z, E) = 0$. Also, it is easy to check that $E' \in \mathcal{E}$. So $E' \in (\widetilde{\mathcal{P}_t})^{\perp} \cap \mathcal{E} = (dg\widetilde{\mathcal{P}_t})^{\perp}$. It follows $\mathrm{Ext}^1(\Phi(G), E) \cong \mathrm{Ext}^1(X, E') = 0$, for every $E \in \mathcal{E}$. In other words, $\Phi(G)$ is dg-projective, and by Proposition 2.7 we have $G = \Psi(\Phi(G))$ is a Ding projective A-module.

Now suppose that R is a left and right coherent ring of finite weak global dimension. Note that Ψ and Φ define an one-to-one correspondence between the projective objects of $\mathrm{Ch}(R)$ and A-Mod. It follows that Ψ and Φ also define an one-to-one correspondence between t-projective complexes over R and t-projective A-modules. Let $X \in (\widetilde{\mathcal{P}_t})^{\perp}$ and consider $\Psi(X)$. Let M be an t-projective A-module. Then $\Phi(M)$ is a t-projective complex. We have $\mathrm{Ext}^1_A(M, \Psi(X)) \cong \mathrm{Ext}^1(\Phi(M), X) = 0$. It follows $\Psi(X) \in (\mathcal{P}_t)^{\perp}$. Hence, Ψ and Φ give rise to a one-to-one correspondence between $(\widetilde{\mathcal{P}_t})^{\perp}$ and $(\mathcal{P}_t)^{\perp}$. Also, by Corollary 2.6, we have the same correspondence between $\mathcal{E}$ and $\mathcal{W}_A$. Since $(dg\widetilde{\mathcal{P}_t})^{\perp} = (\widetilde{\mathcal{P}_t})^{\perp} \cap \mathcal{E}$ and $\mathcal{W}_A \cap (\mathcal{P}_t)^{\perp} = (\mathcal{DP}_t)^{\perp}$ by Corollary 1.4, we have that a complex Y is in $(dg\widetilde{\mathcal{P}_t})^{\perp}$ if and only if $\Psi(Y)$ is in $(\mathcal{DP}_t)^{\perp}$. Since $dg\widetilde{\mathcal{P}_t} = ^{\perp}((dg\widetilde{\mathcal{P}_t})^{\perp})$ and $\mathcal{DP}_t = ^{\perp}((\mathcal{DP}_t)^{\perp})$, we have that Φ maps Ding t-projective A-modules into dg-t-projective complexes. □

The following result is the dual version of Theorem 2.9.

2.10. Theorem. *The functor* $\Psi : \mathrm{Ch}(R) \longrightarrow A\text{-Mod}$ *maps dg-t-injective complexes into Ding t-injective A-modules. If R is a left and right coherent ring of finite weak global dimension, then the inverse functor* $\Phi : A\text{-Mod} \longrightarrow \mathrm{Ch}(R)$ *maps Ding t-injective A-modules into dg-t-injective complexes.*

4. Adjoint functors. In this section, we show that the embedding functors $K(\mathcal{DP}) \longrightarrow K(R\text{-Mod})$ and $K(\mathcal{DI}) \longrightarrow K(R\text{-Mod})$ have right and left adjoints respectively, where $K(\mathcal{DP})$ ($K(\mathcal{DI})$) is the homotopy category of complexes of Ding projective (injective) modules, and $K(R\text{-Mod})$ denotes the homotopy category. To this end, we will be concerned with the category $\mathrm{Ch}(R)$ and the category $K(R\text{-Mod})$ firstly. These categories have the same objects, and the morphisms in $K(R\text{-Mod})$ are homotopy equivalence classes of chain maps, that is, for objects C and D of $K(R\text{-Mod})$, $\mathrm{Hom}_{K(R\text{-Mod})}(C, D) = \mathrm{H}_0(\mathcal{H}om_R(C, D))$, where $\mathrm{Hom}_{K(R\text{-Mod})}(C, D)$ denotes the abelian group of morphisms from C to D in $K(R\text{-Mod})$. We recall that if $f : C \longrightarrow D$ is a morphism in $\mathrm{Ch}(R)$, then we have the mapping cone $con(f)$ of f. We have that $(con(f))_n = D_n \oplus C_{n-1}$ and the differential d is such that $d(y, x) = (d(y) + f(x), -d(x))$. We have the short exact sequence $0 \longrightarrow D \longrightarrow con(f) \longrightarrow \Sigma C \longrightarrow 0$ where the maps $D \longrightarrow con(f)$ and $con(f) \longrightarrow \Sigma C$ are given by $y \longmapsto (y, 0)$ and $(y, x) \longmapsto x$ respectively. Given $f, g \in \mathrm{Hom}(C, D)$ we will let $f \sim g$ mean that f and g are homotopic. The idea of the next lemma derives from Bravo et al. in [[1]].

3.1. Lemma *Let R be a Ding-Chen ring, X be an R-complex, and $0 \longrightarrow C \longrightarrow D \longrightarrow X \longrightarrow 0$ be exact where $D \in \widetilde{\mathcal{DP}}$, $C \in dg\widetilde{\mathcal{W}_R}$. If $D' \in \widetilde{\mathcal{DP}}$, $f_i \in \mathrm{Hom}(D', X)$ and $g_i \in \mathrm{Hom}(D', D)$ such that*

$$\begin{array}{ccc} & & D' \\ & \swarrow^{g_i} & \downarrow^{f_i} \\ D & \longrightarrow & X \end{array}$$

are commutative for $i = 1, 2$, then $f_1 \sim f_2$ if and only if $g_1 \sim g_2$.

Proof If $g_1 \sim g_2$ then easily $f_1 \sim f_2$. For the converse let $f = f_1 - f_2$ and $g = g_1 - g_2$ we see that we only need show that when $f \sim 0$ we have $g \sim 0$. With such f and g we get the commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & D & \longrightarrow & con(g) & \longrightarrow & \Sigma D' & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & X & \longrightarrow & con(f) & \longrightarrow & \Sigma D' & \longrightarrow & 0. \end{array}$$

Since $f \sim 0$, by [[11], Lemma 2.3.2] we get that the lower short exact sequence splits. A retraction $con(f) \longrightarrow X$ provides us with a commutative diagram

$$\begin{array}{ccc} D & \longrightarrow & con(g) \\ \downarrow & \swarrow & \\ X. & & \end{array}$$

Since $\widetilde{\mathcal{DP}}$ is closed under extensions and suspensions we have $con(g) \in \widetilde{\mathcal{DP}}$. Since $D \longrightarrow X$ is a $\widetilde{\mathcal{DP}}$-precover we get a lifting $con(g) \longrightarrow D$. We now prove that $con(g) \longrightarrow D$ provides a retraction of $D \longrightarrow con(g)$ in $K(R\text{-Mod})$. For this note that the difference of the composition $D \longrightarrow con(g) \longrightarrow D$ and the identity map $\mathrm{id}D$ maps D into the kernel of $D \longrightarrow X$, that is into C. Since $(\widetilde{\mathcal{DP}}, dg\widetilde{\mathcal{W}_R})$ is a complete hereditary cotorsion pair, this difference (as a map into C) is homotopic to 0 by [[11], Lemma 2.3.2]. But then the difference as a map into D is homotopic to 0. So $con(g) \longrightarrow D$ provides a retraction of $D \longrightarrow con(g)$ in $K(R\text{-Mod})$. Next we prove that $con(g) \longrightarrow D$ provides a retraction of $D \longrightarrow con(g)$ in $\mathrm{Ch}(R)$. Let $s : con(g) \longrightarrow D$ (s a morphism in $\mathrm{Ch}(R)$) give a retraction of $D \longrightarrow con(g)$ in $K(R\text{-Mod})$. Let r be the corresponding homotopy, i.e. for $y \in D$ we have $(dr + rd)(y) = y - s(y, 0)$. Define $con(g) \longrightarrow D$ by $(y, x) \mapsto y + rg(x) + s(0, x)$ for $(y, x) \in con(g)$. We can easily prove that this map is a morphism of complexes and it gives the desired retraction. So we get that the short exact sequence $0 \longrightarrow D \longrightarrow con(g) \longrightarrow \Sigma D' \longrightarrow 0$ is split exact in $\mathrm{Ch}(R)$. So by [[11], Lemma 2.3.2] we get that $g \sim 0$. □

3.2. Corollary. *Let R be a Ding-Chen ring, X be an R-complex, and $0 \longrightarrow C \longrightarrow D \longrightarrow X \longrightarrow 0$ be exact where $D \in \widetilde{\mathcal{DP}}$ and $C \in dg\widetilde{\mathcal{W}_R}$. If $D' \in \widetilde{\mathcal{DP}}$, then $\mathrm{Hom}_{K(R\text{-Mod})}(D', D) \longrightarrow \mathrm{Hom}_{K(R\text{-Mod})}(D', X)$ is a bijection.*

Proof We first note that the exact sequence $0 \longrightarrow C \longrightarrow D \longrightarrow X \longrightarrow 0$ gives the exact sequence $\mathrm{Hom}(D', D) \longrightarrow \mathrm{Hom}(D', X) \longrightarrow \mathrm{Ext}^1(D', C) = 0$. So $\mathrm{Hom}(D', D) \longrightarrow \mathrm{Hom}(D', X)$ is surjective. This gives that $\mathrm{Hom}_{K(R\text{-Mod})}(D', D) \longrightarrow \mathrm{Hom}_{K(R\text{-Mod})}(D', X)$ is surjective. Lemma 3.1 guarantees that this function is injective and so bijective. □

This gives the following result.

3.3. Theorem. *Let R be a Ding-Chen ring. Then the embedding $K(\mathcal{DP}) \longrightarrow K(R\text{-Mod})$ has a right adjoint.*

Proof For each $X \in \mathrm{Ch}(R)$, there exists an exact sequence $0 \longrightarrow C \longrightarrow D \longrightarrow X \longrightarrow 0$ in $\mathrm{Ch}(R)$ with $D \in \widetilde{\mathcal{DP}}$ and $C \in dg\widetilde{\mathcal{W}_R}$. We want to define a functor $T : K(R\text{-Mod}) \longrightarrow K(\mathcal{DP})$ so that $T(X) = D$. If $f : X \longrightarrow X'$ is a morphism in $\mathrm{Ch}(R)$ we let $[f]$ represent the corresponding morphism in $K(R\text{-Mod})$. So $[f]$ consists of all $f' : X \longrightarrow X'$ such that $f \sim f'$. We use the following procedure to define $T([f])$. We have the exact sequence $\mathrm{Hom}(D, D') \longrightarrow \mathrm{Hom}(D, X') \longrightarrow \mathrm{Ext}^1(D, C') = 0$. This means that there is a morphism $g \in \mathrm{Hom}(D, D')$ whose image in $\mathrm{Hom}(D, X')$, which is the composition $D \longrightarrow X \longrightarrow X'$. So we have the commutative diagram

$$\begin{array}{ccc} D & \longrightarrow & X \\ {\scriptstyle g}\downarrow & & {\scriptstyle f}\downarrow \\ D' & \longrightarrow & X'. \end{array}$$

For $f' \in [f]$ (so $f \sim f'$) we use the same argument and find a morphism $g' : D \longrightarrow D'$ so that the diagram

$$\begin{array}{ccc} D & \longrightarrow & X \\ {\scriptstyle g'}\downarrow & & {\scriptstyle f'}\downarrow \\ D' & \longrightarrow & X' \end{array}$$

is commutative. Then an application of Lemma 3.1 gives that $g \sim g'$. This means that we can define $T([f])$ to be $[g]$ with f and g as above. Then it can be quickly checked that T is an additive functor. Note that the maps $D \longrightarrow X$ then become maps $T(X) \longrightarrow X$ and give a natural transformation from T to the identity functor on $K(R\text{-Mod})$.

Now we appeal to Corollary 3.2. This Corollary says that $\mathrm{Hom}_{K(R\text{-Mod})}(D', D) \longrightarrow \mathrm{Hom}_{K(R\text{-Mod})}(D', X)$ is a bijection if $D' \in \widetilde{\mathcal{DP}}$ and $0 \longrightarrow C \longrightarrow D \longrightarrow X \longrightarrow 0$ is as above. But $T(X) = D$, so we have the bijection

$$\mathrm{Hom}_{K(R\text{-Mod})}(D', T(X)) \longrightarrow \mathrm{Hom}_{K(R\text{-Mod})}(D', X).$$

From the definition of this map we see that it is natural in D'. From the natural transformation above we see that it is natural in X. So this establishes that T is a right adjoint of the embedding functor $K(\mathcal{DP}) \longrightarrow K(R\text{-Mod})$. □

3.4. Remark *We also have the duals of Lemma 3.1 and the Corollary 3.2. The embedding $K(\mathcal{DI}) \longrightarrow K(R\text{-Mod})$ has a left adjoint.*

Acknowledgements. The authors would like to thank the referee for his/her valuable comments, suggestions and corrections which resulted in a significant improvement of the paper.

References

[1] Bravo D., Enochs E.E., Iacob A.C., Jenda O.M.G., Rada J.: Cotorsion pairs in C(R-Mod), Rocky Mountain J. Math. 42, 1787–1802 (2012).
[2] Ding N.Q., Chen J.L.: The flat dimensions of injective modules, Manuscripta Math. 78, 165–177 (1993).
[3] Ding N.Q., Chen J.L.: Coherent rings with finite self-FP-injective dimension, Comm. Algebra 24, 2963–2980 (1996).
[4] Ding N.Q., Mao L.X.: Relative FP-projective modules, Comm. Algebra 33, 1587–1602 (2005).

[5] Ding N.Q., Li Y.L., Mao L.X.: Strongly Gorenstein flat modules, J. Aust. Math. Soc. 86, 323–338 (2009).
[6] Enochs, E.E.: Injective and flat covers, envelopes and resolvents, Israel J. Mathematics, 39, 189–209 (1981).
[7] Enochs, E.E., Jenda, O.M.G., Xu, J.: Orthogonality in the category of complexes, Math. J. Okayama Univ. 38, 25–46 (1996).
[8] Enochs E.E., García Rozas, J.R.: Flat covers of complexes, J. Algebra 210, 86–102, (1998).
[9] Enochs, E.E., Jenda, O.M.G.: Relative Homological Algebra, de Gruyter Exp. Math., vol. 30, de Gruyter, Berlin, 2000.
[10] Enochs, E.E., Oyonarte, L.: Covers, envelopes and cotorsion theories, Nova Science Publishers, Ins. Hauppauge, New York, 2002.
[11] García Rozas, J.R.: Covers and envelope in the category of complexes of modules. Boca Raton London New York Washington, D.C. 1999.
[12] Gillespie, J.: The flat model structure on Ch(R), Tran. Amer. Math. Soc. 356, 3369–3390 (2004).
[13] Gillespie J.: Model structures on modules over Ding-Chen rings, Homology, Homotopy Appl. 12, 61–73 (2010).
[14] Gillespie J., Hovey M.: Gorenstein model structures and generalized derived categories. Submitted to Proceedings of the Edinburgh Mathematical Society.
[15] Hovey M.: Cotorsion pairs, model category structures and representation theory. Mathematical Zeitschrift. 241, 553–592 (2002).
[16] Mao, L.X., Ding, N.Q.: Envelopes and covers by modules of finite FP-injective and flat dimensions, Comm. Algebra 35, 833–849 (2007).
[17] Mao, L.X., Ding N.Q.: Gorenstein FP-injective and Gorenstein flat modules, J. Algebra Appl. 7, 491–506 (2008).
[18] Stenström, B.: Coherent rings and FP-injective modules, J. London Math. Soc. 2(2), 323–329 (1970).
[19] Wang, Z.P., Liu, Z.K.: FP-injective complexes and FP-injective dimension for complexes, J. Aust. Math. Soc. 91, 163–187 (2011).
[20] Yang, G., Liu, Z.K., Liang, L.: Model structures on categories of complexes over Ding-Chen rings, Comm. Algebra 41, 50–69 (2013).

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1475 – 1484

Integral type contraction and coupled coincidence fixed point theorems for two pairs in G-metric spaces

Akbar Zada*, Rahim Shah† and Tongxing Li‡

Abstract

In this paper, we introduces the idea of integral type contraction with respect to G-metric space and by using the notion of integral type contraction we prove some coupled coincidence fixed point results for two pairs of mapping in G-metric space. Also we give an example as an application point of view.

Keywords: G-metric space; couple coincidence point; common fixed point; integral type contraction.

2000 AMS Classification: 47H09; 47H10.

Received : 12.08.2015 *Accepted* : 04.11.2015 *Doi* : 15672/HJMS.20164514280

1. Introduction

The study of common fixed points of mappings which satisfies certain contractive conditions has been studied by a lot of researchers due to its applications in mathematics. For the study of coincidence point of theory in metric and cone metric spaces we recommend [1, 2, 3, 4, 7, 8, 9, 10, 11, 15, 17, 18]. In 2006 Mustafa and Sims [16], introduced the idea of G-metric space and presented some fixed point theorems in G-metric space. The concept of a coupled coincidence point of mapping was introduced by V. Lakshmikantham [5, 13], they also studied some fixed point theorems in partially ordered metric spaces. In 2010 Shatanawi [19] gave the proof of coupled coincidence fixed point theorems in generalized metric spaces. Also in 2014 Manish Kumar [14] proved a coupled coincidence fixed point theorem in the setting of two pairs of mapping in G-metric space. Moreover

*Department of Mathematics, University of Peshawar, Peshawar, Pakistan.
Email : zadababo@yahoo.com, akbarzada@upesh.edu.pk

†Department of Mathematics, University of Peshawar, Peshawar, Pakistan.
Email : safeer_rahim@yahoo.com

‡LinDa Institute of Shandong Provincial Key Laboratory of Network Based Intelligent Computing, Linyi University, Linyi, Shandong 276005, P. R. China.
School of Informatics, Linyi University, Linyi, Shandong 276005, P. R. China.
Email : litongx2007@163.com

in 2002, Branciari [6] gave the idea of integral type contractive mappings in complete metric spaces and they studied the existence of fixed points for mappings which is defined on complete metric space satisfying integral type contraction. Recently F. Khojasteh et al.[12], gave the idea of integral type contraction in cone metric spaces and proved some fixed point theorems in such spaces. So by using the concept of Branciari [6] of integral type contractive mapping, we presented a coupled coincidence fixed point results of integral type contractive mappings for two pairs in the setting of G-metric spaces. Also we give suitable example that support our main result.

2. Preliminaries

We will need the following definitions and results in this paper.

2.1. Definition. [16] Let Y be a non-empty set and $G : Y \times Y \times Y \to R^+$ is a function that satisfies the following conditions:
(1) $G(a, b, c) = 0$ if $a = b = c$,
(2) $G(a, a, b) > 0$ for all $a, b \in Y$ with $a \neq b$,
(3) $G(a, a, b) \leq G(a, b, c)$, for all $a, b, c \in Y$ with $c \neq b$
(4) $G(a, b, c) = G(a, c, b) = G(b, c, a) = \dots$, symmetry in all variables,
(5) $G(a, b, c) \leq G(a, s, s) + G(s, b, c)$ for all $a, b, c, s \in Y$.
Then the function G is called a generalized metric and the pair (Y, G) is called a G-metric space.

2.2. Example. [16] Let $Y = \{x, y\}$. Define G on $Y \times Y \times Y$ by

$$G(x, x, x) = G(y, y, y) = 0, G(x, x, y) = 1, G(x, y, y) = 2$$

and extend G to $Y \times Y \times Y$ by using the symmetry in the variables. Then it is clear that (Y, G) is a G-metric space.

2.3. Definition. [16] Let (Y, G) be a G-metric space and (a_n) a sequence of points of Y. A point $a \in Y$ is said to be the limit of the sequence (a_n), if $\lim_{n,m\to+\infty} G(a, a_n, a_m) = 0$ and we say that the sequence (a_n) is G-convergent to a.

2.1. Proposition. [16] *Let (Y, G) be a G-metric space. Then the following are equivalent:*
(1) *(a_n) is G-convergent to a.*
(2) *$G(a_n, a_n, a) \to 0$ as $n \to +\infty$.*
(3) *$G(a_n, a, a) \to 0$ as $n \to +\infty$.*
(4) *$G(a_n, a_m, a) \to 0$ as $n, m \to +\infty$.*

2.4. Definition. [15] Let (Y, G) be a G-metric space. A sequence (a_n) is called G-Cauchy if for every $\epsilon > 0$, there is $k \in \mathbf{N}$ such that $G(a_n, a_m, a_l) < \epsilon$, for all $n, m, l \geq k$; that is $G(a_n, a_m, a_l) \to 0$ as $n, m, l \to +\infty$.

2.2. Proposition. [16] *Let (Y, G) be a G-metric space. Then the following are equivalent:*
(1) *The sequence (a_n) is G-Cauchy.*
(2) *For every $\epsilon > 0$, there is $k \in \mathbf{N}$ such that $G(a_n, a_m, a_m) < \epsilon$, for all $n, m \geq k$.*

2.5. Definition. [16] A G-metric space (Y, G) is called G-complete if every G-Cauchy sequence in (Y, G) is G-convergent in (Y, G).

2.6. Definition. [5] An element $(a, b) \in Y \times Y$ is called a coupled coincidence point of the mappings $F : Y \times Y \to Y$ and $g : Y \to Y$ if $F(a, b) = ga$ and $F(b, a) = gb$.

2.7. Definition. [13] Let Y be a non-empty set. Then we say that the mappings $F : Y \times Y \to Y$ and $g : Y \to Y$ are commutative if $gF(a,b) = F(ga, gb)$.

2.8. Definition. [13] An element $(a,b) \in Y \times Y$ is called a coupled fixed point of mapping $F : Y \times Y \to Y$ if $F(a,b) = a$ and $F(b,a) = b$.

In 2002, Branciari in [6] introduced a general contractive condition of integral type as follows.

2.9. Theorem. [6] Let (Y,d) be a complete metric space, $\alpha \in (0,1)$, and $f : Y \to Y$ is a mapping such that for all $x, y \in Y$,

$$\int_0^{d(f(x),f(y))} \phi(t)dt \leq \alpha \int_0^{d(x,y)} \phi(t)dt$$

where $\phi : [0,+\infty) \to [0,+\infty)$ is Lebesgue-integrable mapping which is summable (i.e., with finite integral) on each compact subset of $[0,+\infty)$ such that for each $\epsilon > 0$, $\int_0^\epsilon \phi(t)dt > 0$, then f has a unique fixed point $a \in Y$, such that for each $x \in Y$, $\lim_{n\to\infty} f^n(x) = a$.

We use the above idea of Branciari [6] and presented some coupled coincidence fixed point results of integral type contraction in G-metric space.

3. Main Results

In this section we will prove some fixed point results for two pairs in G-metric space by using integral type contractive mapping. We will start our work with the following important lemma.

3.1. Lemma. *Let (Y,G) be a G-metric space. Suppose $F, S : Y \times Y \to Y$ and $g, h : Y \to Y$ be two mappings such that*

$$(3.1) \qquad \int_0^{G(F(a,b),S(p,q),S(c,r))} \varphi(t)dt \leq k \int_0^{(G(ha,gp,gc)+G(hb,gq,gr))} \varphi(t)dt$$

for all $a, b, c, p, q, r \in Y$ and $\varphi : [0,+\infty) \to [0,+\infty)$ is a Lebesgue integrable mapping which is summable such that for each $\epsilon > 0$, $\int_0^\epsilon \varphi(t)dt > 0$. Assume that (a,b) is coupled coincidence point of the pairs of mappings $\{F,h\}$ and $\{S,g\}$ and $ga = ha$ and $gb = hb$. If $k \in [0, \frac{1}{8})$, then
$S(a,b) = ga = gb = S(b,a)$ and $F(a,b) = ha = hb = F(b,a)$.

Proof. Since (a,b) is a coupled coincidence point of the mappings $\{F,h\}$ and $\{S,g\}$, we have $ha = F(a,b)$, $hb = F(b,a)$ and $ga = S(b,a)$, $gb = S(b,a)$. Suppose $ga \neq gb$. Then by (3.1), we get

$$\begin{aligned} \int_0^{G(ga,gb,gb)} \varphi(t)dt &= \int_0^{G(F(a,b),S(b,a),S(b,a))} \varphi(t)dt \\ &\leq k \int_0^{(G(ha,gb,gb)+G(hb,ga,ga))} \varphi(t)dt \\ &= k \int_0^{(G(ga,gb,gb)+G(gb,ga,ga))} \varphi(t)dt. \end{aligned}$$

Also we have,

$$\begin{aligned}\int_0^{G(gb,ga,ga)}\varphi(t)dt &= \int_0^{G(F(b,a),S(a,b),S(a,b))}\varphi(t)dt\\ &\leq k\int_0^{(G(hb,ga,ga)+G(ha,gb,gb))}\varphi(t)dt\\ &= k\int_0^{(G(gb,ga,ga)+G(ga,gb,gb))}\varphi(t)dt.\end{aligned}$$

Therefore

$$\int_0^{G(ga,gb,gb)}\varphi(t)dt+\int_0^{G(gb,ga,ga)}\varphi(t)dt\leq 2k\int_0^{(G(ga,gb,gb)+G(gb,ga,ga))}\varphi(t)dt.$$

Since $2k<1$, we get

$$\int_0^{G(ga,gb,gb)}\varphi(t)dt+\int_0^{G(gb,ga,ga)}\varphi(t)dt<\int_0^{G(ga,gb,gb)}\varphi(t)dt+\int_0^{G(gb,ga,ga)}\varphi(t)dt$$

which is a contradiction. So $ga=gb$, and hence

$$S(a,b)=ga=gb=S(b,a) \text{ and } F(a,b)=ha=hb=F(b,a).$$

3.1. Theorem. Let (Y,G) be a G-metric space. Let $F,S:Y\times Y\to Y$ and $g,h:Y\to Y$ be two mappings such that

$$\int_0^{G(F(a,b),S(p,q),S(c,r))}\varphi(t)dt\leq k\int_0^{(G(ha,gp,gc)+G(hb,gq,gr))}\varphi(t)dt \tag{3.2}$$

for all $a,b,c,p,q,r\in Y$ and $\varphi:[0,+\infty)\to[0,+\infty)$ is a Lebesgue integrable mapping which is summable such that for each $\epsilon>0$, $\int_0^{\epsilon}\varphi(t)dt>0$. Assume that F,S and g,h satisfy the following conditions:

(i) $F(Y\times Y)\subset g(Y)$ and $S(Y\times Y)\subset h(Y)$

(ii) $g(Y)$ or $h(Y)$ is complete and

(iii) g and h are G-continuous and pairs $\{F,h\}$ and $\{S,g\}$ are commuting mappings.

If $k\in[0,\frac{1}{8})$, then there is a unique $a\in Y$ such that $F(a,a)=S(a,a)=g(a)=h(a)=a$.

Proof. Let $a_0,b_0\in Y$. Since $F(Y\times Y)\subset g(Y)$, choose $a_1,b_1\in Y$ such that $u'_1=ga_1=F(a_0,b_0)$ and $v'_1=gb_1=F(b_0,a_0)$. Again since $S(Y\times Y)\subset h(Y)$, choose $a_2,b_2\in Y$ such that $u'_2=ha_2=S(a_1,b_1)$ and $v'_2=hb_2=S(b_1,a_1)$. Continuing this process, we can construct two sequences (u'_n) and (v'_n) in Y such that $u'_{2n+1}=ga_{2n+1}=F(a_{2n},b_{2n})$, $v'_{2n}=gb_{2n+1}=F(b_{2n+1},a_{2n+1})$ and $u'_{2n+2}=ha_{2n+2}=S(a_{2n+1},b_{2n+1})$, $v'_{2n+2}=hb_{2n+2}=S(b_{2n+1},a_{2n+1})$. For $n\in N$, we have

$$\begin{aligned}\int_0^{G(u'_{2n+1},u'_{2n+2},u'_{2n+2})}\varphi(t)dt &= \int_0^{(G(F(a_{2n},b_{2n}),S(a_{2n+1},b_{2n+1}),S(a_{2n+1},b_{2n+1}))}\varphi(t)dt\\ &\leq k\int_0^{(G(ha_{2n},ga_{2n+1},ga_{2n+1})+G(hb_{2n},gb_{2n+1},gb_{2n+1}))}\varphi(t)dt\\ &= k\int_0^{(G(u'_{2n},u'_{2n+1},u'_{2n+1})+G(v'_{2n},v'_{2n+1},v'_{2n+1}))}\varphi(t)dt. \quad (3.3)\end{aligned}$$

In the same manner

$$\int_0^{G(v'_{2n+1},v'_{2n+2},v'_{2n+2})}\varphi(t)dt \leq k\int_0^{(G(v'_{2n},v'_{2n+1},v'_{2n+1})+G(u'_{2n},u'_{2n+1},u'_{2n+1}))}\varphi(t)dt. \tag{3.4}$$

We have

$$\int_0^{G(u'_{2n+1},u'_{2n+2},u'_{2n+2})}\varphi(t)dt + \int_0^{G(v'_{2n+1},v'_{2n+2},v'_{2n+2})}\varphi(t)dt$$
$$\leq 2k\int_0^{\{G(u'_{2n},u'_{2n+1},u'_{2n+1})+(G(v'_{2n},v'_{2n+1},v'_{2n+1})\}}\varphi(t)dt$$
$$\leq 8k\int_0^{\{G(u'_{2n},u'_{2n+1},u'_{2n+1})+(G(v'_{2n},v'_{2n+1},v'_{2n+1})\}}\varphi(t)dt, \quad (3.5)$$

holds for all $n \in N$, again from

$$\begin{aligned}
\int_0^{G(u'_{2n},u'_{2n+1},u'_{2n+1})}\varphi(t)dt &\leq 2\int_0^{G(u'_{2n+1},u'_{2n},u'_{2n})}\varphi(t)dt\\
&= 2\int_0^{(G(F(a_{2n},b_{2n}),S(a_{2n-1},b_{2n-1}),S(a_{2n-1},b_{2n-1}))}\varphi(t)dt\\
&\leq 2k\int_0^{(G(ha_{2n},ga_{2n-1},ga_{2n-1})+G(hb_{2n},gb_{2n-1},gb_{2n-1}))}\varphi(t)dt\\
&= 2k\int_0^{(G(u'_{2n},u'_{2n-1},u'_{2n-1})+G(v'_{2n},v'_{2n-1},v'_{2n-1}))}\varphi(t)dt\\
&\leq 4k\int_0^{(G(u'_{2n-1},u'_{2n},u'_{2n})+G(v'_{2n-1},v'_{2n},v'_{2n}))}\varphi(t)dt \quad (3.6)
\end{aligned}$$

and

$$\begin{aligned}
\int_0^{G(v'_{2n},v'_{2n+1},v'_{2n+1})}\varphi(t)dt &\leq 2\int_0^{G(v'_{2n+1},v'_{2n},v'_{2n})}\varphi(t)dt\\
&= 2\int_0^{(G(F(b_{2n},a_{2n}),S(b_{2n-1},a_{2n-1}),S(b_{2n-1},a_{2n-1}))}\varphi(t)dt\\
&\leq 2k\int_0^{(G(hb_{2n},gb_{2n-1},gb_{2n-1})+G(ha_{2n},ga_{2n-1},ga_{2n-1}))}\varphi(t)dt\\
&= 2k\int_0^{(G(v'_{2n},v'_{2n-1},v'_{2n-1})+G(u'_{2n},u'_{2n-1},u'_{2n-1}))}\varphi(t)dt\\
&\leq 4k\int_0^{(G(u'_{2n-1},u'_{2n},u'_{2n})+G(v'_{2n-1},v'_{2n},v'_{2n}))}\varphi(t)dt. \quad (3.7)
\end{aligned}$$

We have

$$\int_0^{G(u'_{2n},u'_{2n+1},u'_{2n+1})}\varphi(t)dt + \int_0^{G(v'_{2n},v'_{2n+1},v'_{2n+1})}\varphi(t)dt$$
$$\leq 8k\int_0^{(G(u'_{2n-1},u'_{2n},u'_{2n})+G(v'_{2n-1},v'_{2n},v'_{2n}))}\varphi(t)dt, \quad (3.8)$$

holds for all $n \in N$. Thus, using (3.5) and (3.8) in (3.3), we get

$$\begin{aligned}
\int_0^{G(u'_{2n+1},u'_{2n+2},u'_{2n+2})} \varphi(t)dt &\leq k8k \int_0^{(G(u'_{2n-1},u'_{2n},u'_{2n})+G(v'_{2n-1},v'_{2n},v'_{2n}))} \varphi(t)dt \\
&\leq k(8k)^2 \int_0^{(G(u'_{2n-2},u'_{2n-1},u'_{2n-1})+G(v'_{2n-2},v'_{2n-1},v'_{2n-1}))} \varphi(t)dt \\
&\leq k(8k)^{2n} \int_0^{(G(u'_0,u'_1,u'_1)+G(v'_0,v'_1,v'_1))} \varphi(t)dt \\
&\vdots \\
&\leq (8k)^{2n+1} \int_0^{(G(u'_0,u'_1,u'_1)+G(v'_0,v'_1,v'_1))} \varphi(t)dt,
\end{aligned}$$

and also, using (3.5) and (3.8) in (3.6), we get

$$\begin{aligned}
\int_0^{G(u'_{2n},u'_{2n+1},u'_{2n+1})} \varphi(t)dt &\leq 4k(8k) \int_0^{(G(u'_{2n-2},u'_{2n-1},u'_{2n-2})+G(v'_{2n-1},v'_{2n-1},v'_{2n-1}))} \varphi(t)dt \\
&\vdots \\
&\leq (8k)^{2n} \int_0^{(G(u'_0,u'_1,u_1)+G(v'_0,v'_1,v'_1))} \varphi(t)dt.
\end{aligned}$$

Thus for all $n \in N$, we have

$$\int_0^{G(u'_n,u'_{n+1},u'_{n+1})} \varphi(t)dt \leq (8k)^n \int_0^{(G(u'_0,u'_1,u'_1)+G(v'_0,v'_1,v'_1))} \varphi(t)dt.$$

Let $m, n \in N$ with $m > n$, we have

$$\begin{aligned}
\int_0^{G(u'_n,u'_m,u'_m)} \varphi(t)dt &\leq \int_0^{G(u'_n,u'_{n+1},u'_{n+1})} \varphi(t)dt + \int_0^{G(u'_{n+1},u'_{n+2},u'_{n+2})} \varphi(t)dt \\
&+ \cdots + \int_0^{G(u'_{m-1},u'_m,u'_m)} \varphi(t)dt.
\end{aligned}$$

Since $8k < 1$, we get

$$\begin{aligned}
\int_0^{G(u'_n,u'_m,u'_m)} \varphi(t)dt &\leq \sum_{i=n}^{m-1} (8k)^i \int_0^{(G(u'_0,u'_1,u'_1)+G(v'_0,v'_1,v'_1))} \varphi(t)dt \\
&\leq \frac{(8k)^n}{(1-8k)} \int_0^{(G(u'_0,u'_1,u'_1)+G(v'_0,v'_1,v'_1))} \varphi(t)dt.
\end{aligned}$$

We have

$$\lim_{n,m\to+\infty} G(u'_n, u'_m, u'_m) = 0.$$

Thus (u'_n) is G-Cauchy in $g(Y)$. As $g(Y)$ is G-complete then subsequence $(u'_{2n+1}) = (ga_{2n+1})$ and $(v'_{2n+1}) = (gb_{2n+1})$ are convergent to some $a \in Y$ and $b \in Y$ respectively. As we know that every sequence and subsequence of a G-Cauchy sequence are convergent to the same point. Hence $(u'_{2n}) = (ha_{2n})$ and $(v'_{2n}) = (hb_{2n})$ are also convergent. Since g and h are G-continuous, we have

$$(gga_{2n+1}) \to ga, (hga_{2n+1}) \to ha, (gha_{2n}) \to ga, (hha_{2n}) \to ha$$

and

$$(ggb_{2n+1}) \to gb, (hgb_{2n+1}) \to hb, (ghb_{2n}) \to gb, (hhb_{2n}) \to hb.$$

Since pairs $\{F, h\}$ and $\{S, g\}$ are commutative mappings, we have

$$hga_{2n+1} = hF(a_{2n}, b_{2n}) = F(ha_{2n}, hb_{2n})$$

and

$$gha_{2n} = gS(a_{2n-1}, b_{2n-1}) = S(ga_{2n-1}, ga_{2n-1}).$$

Thus

$$\begin{aligned}\int_0^{G(hga_{2n+1}, gha_{2n}, gha_{2n})} \varphi(t)dt &= \int_0^{G(F(ha_{2n}, hb_{2n}), S(ga_{2n-1}, gb_{2n-1}), S(ga_{2n-1}, gb_{2n-1}))} \varphi(t)dt \\ &\leq k \int_0^{(G(hha_{2n}, gga_{2n-1}, gga_{2n-1}) + G(hhb_{2n}, ggb_{2n-1}, ggb_{2n-1}))} \varphi(t)dt.\end{aligned}$$

Letting $n \to +\infty$, we have

$$\int_0^{G(ha,ga,ga)} \varphi(t)dt = k \int_0^{(G(ha,ga,ga)+G(hb,gb,gb))} \varphi(t)dt.$$

In the same way, we can show that

$$\int_0^{G(hb,gb,gb)} \varphi(t)dt = k \int_0^{(G(hb,gb,gb)+G(ha,ga,ga))} \varphi(t)dt.$$

Thus

$$\int_0^{G(ha,ga,ga)} \varphi(t)dt + \int_0^{G(hb,gb,gb)} \varphi(t)dt = 2k \int_0^{(G(ha,ga,ga)+G(hb,gb,gb))} \varphi(t)dt.$$

Since $2k < 8k < 1$, the last equality happens only if

$$\int_0^{G(ha,ga,ga)} \varphi(t)dt = \int_0^{G(hb,gb,gb)} \varphi(t)dt = 0.$$

Hence $ha = ga$ and $hb = gb$. Again

$$\begin{aligned}\int_0^{G(hga_{2n+1}, S(a,b), S(a,b))} \varphi(t)dt &= \int_0^{G(F(ha_{2n}, hb_{2n}), S(a,b), S(a,b))} \varphi(t)dt \\ &\leq k \int_0^{(G(hha_{2n}, gb, gb) + G(hhb_{2n}, gb, gb))} \varphi(t)dt.\end{aligned}$$

Letting $n \to +\infty$, we have

$$\int_0^{G(ha,S(a,b),S(a,b))} \varphi(t)dt \leq k \int_0^{(G(ha,gb,gb)+G(hb,gb,gb))} \varphi(t)dt = 0.$$

Thus, we get

$$\int_0^{G(ha,S(a,b),S(a,b))} \varphi(t)dt = 0.$$

Which implies that $S(a,b)=ha$. Similarly we can show that $S(b,a)=hb$. By using the same technique, we get

$$\begin{aligned}\int_0^{G(F(a,b),gha_{2n},gha_{2n}))}\varphi(t)dt &= \int_0^{G(F(a,b),S(ga_{2n-1},gb_{2n-1}),S(ga_{2n-1},gb_{2n-1}))}\varphi(t)dt\\ &\leq k\int_0^{(G(ha,gga_{2n-1},gga_{2n-1})+G(hb,ggb_{2n-1},ggb_{2n-1}))}\varphi(t)dt.\end{aligned}$$

Letting $n\to+\infty$, we have

$$\int_0^{G(F(a,b),ga,ga))}\varphi(t)dt = k\int_0^{(G(ha,ga,ga)+G(hb,gb,gb))}\varphi(t)dt = 0.$$

Thus, we get

$$\int_0^{G(F(a,b),ga,ga))}\varphi(t)dt = 0,$$

which means that $F(a,b)=ga$. By using the same method we can show that $F(b,a)=hb$. Hence we get $ga=ha$, $gb=hb$ and $F(a,b)=gb$, $S(a,b)=ha$, $S(b,a)=hb$, by using Lemma 3.1 we have
$F(a,b)=ga=gb=F(b,a)=S(a,b)=ha=hb=S(b,a)$.
Now

$$\begin{aligned}\int_0^{G(ga_{2n+1},ga,ga)}\varphi(t)dt &= \int_0^{G(F(a_{2n},b_{2n}),S(a,b),S(a,b))}\varphi(t)dt\\ &\leq k\int_0^{(G(ha_{2n},ga,ga)+G(hb_{2n},gb,gb))}\varphi(t)dt.\end{aligned}$$

Letting $n\to\infty$, we have,

$$\int_0^{G(a,ga,ga)}\varphi(t)dt = k\int_0^{(G(a,ga,ga)+G(b,gb,gb))}\varphi(t)dt.$$

Similarly, we can show that

$$\int_0^{G(b,gb,gb)}\varphi(t)dt = k\int_0^{(G(b,gb,gb)+G(a,ga,ga))}\varphi(t)dt.$$

Thus

$$\int_0^{G(a,ga,ga)}\varphi(t)dt + \int_0^{G(b,gb,gb)}\varphi(t)dt = 2k\int_0^{(G(a,ga,ga)+G(b,gb,gb))}\varphi(t)dt.$$

Since $2k<8k<1$, the last equality happens only if

$$\int_0^{G(a,ga,ga)}\varphi(t)dt = \int_0^{G(b,gb,gb)}\varphi(t)dt = 0.$$

Hence $a=ga$ and $b=gb$. Thus, we get
$F(a,a)=S(a,a)=ga=ha=a$.
For uniqueness, let $y\in Y$ with $y\neq a$ such that
$F(y,y)=S(y,y)=gy=y$.

Then

$$\begin{aligned}\int_0^{G(a,y,y)} \varphi(t)dt &= \int_0^{G(F(a,a),S(y,y),S(y,y))} \varphi(t)dt \\ &\leq k\int_0^{G(ha,gy,gy)+G(ha,gy,gy)} \varphi(t)dt \\ &= k\int_0^{G(a,y,y)+G(a,y,y)} \varphi(t)dt \\ &= 2k\int_0^{G(a,y,y)} \varphi(t)dt.\end{aligned}$$

Since $2k < 8k < 1$, we get

$$\int_0^{G(a,y,y)} \varphi(t)dt < \int_0^{G(a,y,y)} \varphi(t)dt.$$

Which is a contradiction. Thus F, S, g, h have a unique common fixed point.

3.2. Corollary. Let (Y, G) be a G-metric space. Let $F, S : Y \times Y \to Y$ and $g, h : Y \to Y$ be two mappings such that

$$\int_0^{G(F(a,b),S(p,q),S(p,q))} \varphi(t)dt \leq k\int_0^{(G(ha,gp,gp)+G(hb,gq,gq))} \varphi(t)dt \tag{3.3}$$

for all $a, b, p, q \in Y$ and $\varphi : [0, +\infty) \to [0, +\infty)$ is a Lebesgue integrable mapping which is summable such that for each $\epsilon > 0$, $\int_0^{\epsilon} \varphi(t)dt > 0$. Assume that F, S and g, h satisfy the following conditions:
$(i) F(Y \times Y) \subset g(Y)$ $S(Y \times Y) \subset h(Y)$
$(ii) g(Y)$ or $h(Y)$ is complete,and
(iii) g and h is G-continuous and pairs $\{F, h\}$ and $\{S, g\}$ are of commuting mappings.
If $k \in [0, \frac{1}{8})$, then there is a unique $a \in Y$ such that $F(a, a) = S(a, a) = g(a) = h(a) = a$.

Proof. In Theorem 3.1 by taking $c = p$ and $q = r$.

3.3. Example. Let $Y = [0, 1]$. Define $G : Y \times Y \times Y \to R^+$ by
$G(a, b, c) = |a - b| + |a - c| + |b - c|$
for all $a, b, c \in Y$. Then (Y, G) is a complete G-metric space.
Define mappings $F, S : Y \times Y \to Y$ and $g, h : Y \to Y$ by

$F(a, b) = \frac{1}{36}ab$, $S(a, b) = \frac{1}{144}ab$ and $ga = \frac{1}{4}a$, $ha = \frac{1}{2}a$.

Since $|ab - pq| = |a - p| + |b - q|$ holds for all $a, b, p, q \in Y$.

Then the condition of Theorem (3.1) holds, in fact

$$\begin{aligned}\int_0^{G(F(a,b),S(p,q),S(c,r))} \varphi(t)dt &= \int_0^{(|\frac{1}{36}ab-\frac{1}{144}pq|+|\frac{1}{144}pq-\frac{1}{144}cr|+|\frac{1}{144}cr-\frac{1}{36}ab|)} \varphi(t)dt \\ &\leq \frac{1}{9}\int_0^{\{|\frac{1}{2}a-\frac{1}{4}p|+|\frac{1}{4}p-\frac{1}{4}c|+|\frac{1}{4}c-\frac{1}{2}a|+|\frac{1}{2}b-\frac{1}{4}q|+|\frac{1}{4}q-\frac{1}{4}r|+|\frac{1}{4}r-\frac{1}{2}b|\}} \varphi(t)dt \\ &= \frac{1}{9}\int_0^{G(ha,gp,gc)+G(hb,gq,gr)} \varphi(t)dt\end{aligned}$$

holds for all $a, b, c, p, q, r \in Y$. It is easy to see that F, S, g, h satisfies all the hypothesis of Theorem 3.1. Thus F, S, g, h have a unique common fixed point. Here $F(0, 0) = S(0, 0) = g0 = h0 = 0$.

References

[1] M. Abbas, G. Jungck, *Common fixed point results for noncommuting mappings without continuity in cone metric spaces*, J. Math. Anal. Appl., Vol. **341**, 2008, pp. 416-420.

[2] M. Abbas, B. E. Rhoades, *Common fixed point results for noncommuting mapping without continuity in generalized metric spaces*, Applied Mathematics and Computation, Vol. **215**, 2009, pp. 262-269.

[3] Th. Abdeljawad, *Completion of cone metric spaces*, Hacet. J. Math. Stat., Vol. **39**, 2010, pp. 67-74.

[4] I. Beg, M. Abbas, *Coincidence point and invariant approximation for mappings satisfying generalized weak contractive condition*, Fixed Point Theory. Appl., Article ID 74503, 2006, pp. 1-7.

[5] T. G. Bhaskar, G. V. Lakshmikantham, *Fixed point theorems in partially ordered metric spaces and applications*, Nonlinear Analysis, Vol. **65**, 2006, pp. 1379-1393.

[6] A. Branciari, *A fixed point theorem for mappings satisfying a general contractive condition of integral type*, International Journal of Mathematics and Mathematical Sciences, Vol. **29**, No. 9, 2002, pp. 531-536.

[7] L. G. Haung, X. Zhang, *Cone metric spaces and fixed point theorems of contractive mappings*, J. Math. Anal. Appl., Vol. **332**, 2007, pp. 1468-1476.

[8] G. Jungck, *Commuting maps and fixed points*, Am. Math. Monthly., Vol. **83**, 1976, pp. 261-263.

[9] G. Jungck, *Compatible mappings and common fixed points*, Int. J. Math. Sci., Vol. **9** (4), 1986, pp. 771-779.

[10] G. Jungck, *Common fixed points for commuting and compatible maps on compacta*, Proc. Am. Math. Soc., Vol. **103**, 1988, pp. 977-983.

[11] G. Jungck, N. Hussain, *Compatible maps and invariant approximations*, J. Math. Anal. Appl., Vol. **325** (2), 2007, pp. 1003-1012.

[12] F. Khojasteh, Z. Goodarzi, A. Razani, *Some Fixed Point Theorems of Integral Type Contraction in Cone Metric Spaces*, Fixed Point Theory Appl., Vol. **2010**, Article ID 189684, 13 pages, 2010.

[13] V. Lakshmikantham, Lj. B. Cirić, *Coupled fixed point theorems for nonlinear contractions in partially ordered metric spaces*, Nonlinear Analysis, Vol. **70**, 2009, pp. 4341-4349.

[14] K. Manish, *Coupled coincidence fixed point theorem for two pairs in G-metric spaces*, Int. J. Math. Appl., vol. **3**, 2014, pp. 53-62.

[15] P. P. Murthy, K. Tas, *New common fixed point theorems of Gregus type for R- weakly commuting mappings in 2-metric spaces*, Hacet. J. Math. Stat., Vol. **38**, 2009, pp. 285-291.

[16] Z. Mustafa, B. Sims, *A new approach to generalized metric spaces*, J. Nonlinear Convex Anal., Vol. **7** (2), 2006, pp. 289-297.

[17] R. P. Pant, *Common fixed points of noncommuting mappings*, J. Math. Anal. Appl. Vol. **188**, 1994, pp. 436-440.

[18] V. Popa, M. Mocanu, *Altering distance and common fixed points under implicit rela tions*, Hacet. J. Math. Stat., Vol. **38**, 2009, pp. 329-337.

[19] W. Shatanawi, *Coupled fixed point theorems in generalized metric spaces*, Hecet. J. Math. Stat., Vol. **40** (3), 2011, pp. 441-447.

STATISTICS

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1485 – 1518

Multi-stage multi-objective solid transportation problem for disaster response operation with type-2 triangular fuzzy variables

Abhijit Baidya*† , Uttam Kumar Bera‡ and Manoranjan Maiti§

Abstract

In this paper, for the first time we formulate and solve multi-stage solid transportation problem (MSSTP) to minimize the total cost and time with type-2 fuzzy transportation parameters. During transportation period, loading, unloading cost and time, volume and weight for each item, limitation of volume and weight for each vehicle are normally imprecise and taken into account to formulate the models. To remove the uncertainty of the type-2 fuzzy transportation parameters from objective functions and constraints, we apply CV-Based reduction methods and generalized credibility measure. Disasters are unexpected situations that require significant logistical deployment to transport equipment and humanitarian goods in order to help and provide relief to victims and sometime this transportation is not possible directly from supply point to destination. Again, the availabilities at supply points and requirements at destinations are not known precisely due to disaster. For this reason, we formulate the multi-stage solid transportation problems under uncertainty (type-2 fuzzy). The models are illustrated with a numerical example. Finally, generalized reduced gradient technique (LINGO.13.0 software) is used to solve the models.

Keywords: Multi-stage solid transportation problem, type-2 fuzzy variable, CV-based reduction methods, generalized credibility, goal programming approach.

2000 AMS Classification: 90B06, 90C05, 90C29

*Department of Mathematics, National Institute of Technology, Agartala, Barjala, Jirania,Tripura, India, Pin- 799046, Email: abhijitnita@yahoo.in

†Corresponding Author.

‡Department of Mathematics, National Institute of Technology, Agartala, Barjala, Jirania,Tripura, India, Pin- 799046, Email: bera_uttam@yahoo.co.in

§Department of Applied Mathematics with oceanology and computer programming, Vidyasagar University , Midnapur-721102, W. B. ,India, Email: mmaiti2005@yahoo.co.in

Received : 02.02.2015 *Accepted* : 15.09.2015 *Doi* : 10.15672/HJMS.201611915992

1. Introduction

1.1. Literature review and the main work of the research: The transportation problem originally developed by Hitchcock [7] is one of the most common combinatorial problems involving constraints. The solid transportation problem (STP) was first stated by Shell [8]. Haley [5, 6] developed the solution procedure of a solid transportation problem and made a comparison between the STP and the classical transportation problem. Geoffrion and Graves [19] were the first researchers studied on two-stage distribution problem. After that so many researchers(Pirkul and Jayaraman [20], Heragu [17], Hindi et al. [21], Syarif and Gen [22], Amiri [23], Gen et al. [24]) study two-stage TP. Mahapatra et al. [25] applied fuzzy multi-objective mathematical programming technique on a reliability optimization model. A type-2 fuzzy variable is a map from a fuzzy possibility space to the real number space; it is an appropriate tool for describing type-2 fuzziness. The concept of a type-2 fuzzy set was first proposed in Qin et al. [1] as an extension of an ordinary fuzzy set. Mitchell [10] used the concept of an embedded type-1 fuzzy number. Liang and Mendel [11] proposed the concept of an interval type-2 fuzzy set. Karnik and Mendel [2], Liu [3], Qin et al. [1], Yang et al. [30], Liu et al. [29], Yang et al. [30] worked on type-2 fuzzy set. In the literature, data envelopment analysis (DEA) technology was first proposed in [12]. Sengupta [22] incorporated stochastic input and output variations; Banker [14] incorporated stochastic variables into DEA; Cooper et al. [15] and Land et al. [16] developed a chance-constrained programming to accommodate the stochastic variations in the data. Qin et al. [4] proposes three noble methods of reduction for a type-2 fuzzy variable. Here, we present in tabular form a scenario of literature development made on transportation problem in Table-1.

Table-1: Some remarkable research works on TP/STP

Author(s), Ref.	Objective	Nature	Additional function	Environments	Solution Techniques
Kundu et al.	Multi-objective	STP	Multi-item	fuzzy	LINGO
Yang and Feng	Single-objective	STP	fixed charges	stochastic	Tabu search algorithm
Kundu et al.	Single-objective	TP	Fixed charge	type-2 fuzzy variable	LINGO
Baidya et al.	Single-objective	STP	Safety factor	Fuzzy, Stochastic, Interval	LINGO
Gen et al.	Single-objective	TP	Two-stage	Deterministic	Genetic algorithms
Proposed	Multi-objective	STP	Multi-stage	Triangular type-2 fuzzy	LINGO

In spite of the above developments, there are so many gaps in the literature. Some of these omissions which are used to formulate the model with type-2 triangular fuzzy number are as follows:

- So many ([26], [27], [28], [31], [32], [33],) solid transportation problems exist in the literature to minimize the total transportation cost only but nobody can formulate any STP to minimize the total transportation time, purchasing cost, loading and unloading cost at a time.

- In spite of the above developments, very few can minimized the time objective function which involves total transportation time, loading and unloading time at a time.

- Lots of two stage transportation problems ([19]-[25]) exist in the literature where the transportation cost is minimized. But nobody formulate and solved a multi-stage multi-item multi-objective solid transportation problem to minimize the "total cost" which involves transportation cost, purchasing cost, loading and unloading cost and "total time" which involves transportation time, loading and unloading time.

- Sometimes, the value of the transportation parameters are not known to us precisely but at that time some imprecise data are known to us. For this reason, lots of researchers solved so many transportation problems with fuzzy (triangular fuzzy number, trapezoidal fuzzy number, type-1 fuzzy number, type-2 fuzzy number, interval type-2 fuzzy number etc.) transportation parameters. But nobody solved any multi-stage multi-item multi-objective (cost and time) STP with transportation parameters as type-2 triangular fuzzy number.

- So many STPs are developed in the literature to minimize the total transportation cost and time subjected to the supply constraints, demand constraints and conveyances capacity constraints, budget constraint, safety constraint etc. but nobody formulated any STP subjected to the weights constraints and volumes constraints during transportation.

In this paper, a multi-item multi-stage solid transportation problem is formulated and solved. Type-2 fuzzy theory is an appropriate field for research. To formulate the model, we consider unit transportation cost, time, supplies, demands, conveyances capacities, loading and unloading cost and time, volume and weights for each and every item, volume and weight capacities for each conveyances as type-2 triangular fuzzy variables. The objective functions for the respective transportation model is to minimize the total cost and time. To defuzzify the constraints and objective functions, we apply CV-based reduction method. The goal programming approach is used to solve the multi-objective programming problem. The deterministic problems so obtained are then solved by using the standard optimization solver - LINGO 13.0 software. We have provided numerical examples illustrating the proposed model and techniques. Some sensitivity analyzes for the model are also presented.
The paper is organized as follows. Problem descriptions are includes in the section 2. In section 3 we briefly introduce some fundamental concepts. The assumptions and notations to construct the model are put in the section 4. In section 5, multi-stage multi-objective solid transportation model with type-2 triangular fuzzy variable is formulated. In section 6, we discuss about the methodology and defuzzification method that used to solve the model. A numerical example put in the section 7 is to illustrate the model numerically. The results of solving the model numerically are put in the section 8. A sensitivity analysis of the model is discussed in the section 9. In section 9, we discuss the results obtained by solving the numerical example. The comparisons of the work with the earlier research are discussed in the section 10. The conclusion and future extension of the research work are discussed in the section 11. The references which are used to prepare this manuscript are put in the last. In this work, we formulate and solve a multi-stage multi-item solid transportation problem to minimize the total cost and time under type-2 fuzzy environment. The real life applications of the research work are as follows:

- Basically to provide some relief to the survived peoples in disaster, we developed our MSSTP model. In our paper, we consider all the transportation parameters as type-2 fuzzy variables since after disaster, it is very difficult to define all transportation parameters precisely. Since due to disaster roads, bridges, towers etc. are damaged, thus it is not possible to survive the peoples smoothly with the help of direct transportation network. This is the reason to formulate an n-stage solid transportation model

- To get the permit of a vehicle is a difficult task for the owner of the vehicle. Some vehicles are permitted to driver on a particular state or country. This permit restricts us to carry the goods from one state to another or one country to another. So in the transportation period, it is important to load and unload the goods so many times at the destination centers (the destination centers are lies between supply points and customers).So to overcome these transportation difficulties we can apply this newly developed model.

1.2. Motivations: The motivation for this research dated back to September 2014, the Kashmir region witnessed disastrous floods across majority of its districts caused by torrential rainfall. The Indian administrated Jammu and Kashmir, as well as Azad Kashmir, Gilgit-Baltistan and Punjab in Pakistan, were affected by these floods. By September 24, 2014, nearly 277 people in India and 280 people in Pakistan died due to the floods and more than 1.1 million were affected by the floods. During this period, it is tedious to send the necessary foods to the survived peoples. For this reason, it is important to impose so many destination centers in between supply point and customers.

2. Problem Description:

Disaster (earthquick, flood etc.) is an extra ordinary situation for any country or state and to provide relief to the survived person which is a risky and tedious task to us. But at that time transportation is required to serve the foods, clothes etc. to the peoples. Also due to the disaster, it is not possible to deliver the necessary things directly to the survived people. It required some destination centers in between source point and survived peoples such that the total cost and time should be minimized. This also motivated us to formulate a multi-stage multi-objective solid transportation problem. The exact figure of the survived peoples due to disaster is not known to us exactly. For this reason, the transportation parameters are also remains unknown to us. Since all the transportation parameters are not known to us precisely, so we consider the transportation parameters as type-2 triangular fuzzy variables. In this multi-stage transportation network, destination center for stage-1 is reduced to the supply point for stage-2 and destination center for stage-2 is reduced to the supply point for stage-3 and similarly the destination center for the stage-(n-1) is converted to the supply point to the stage-n. The pictorial representation of the multi-stage solid transportation problem is as follows:

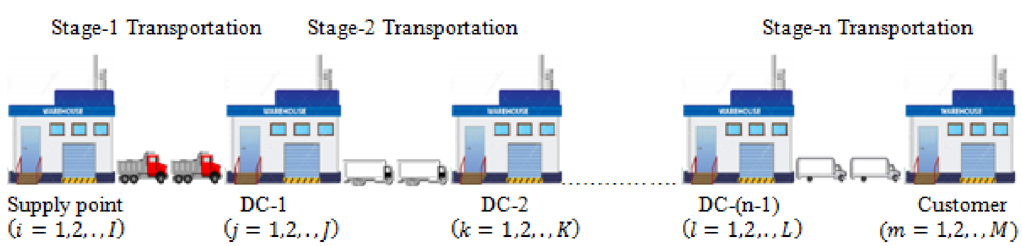


Figure 1. *Multi-stage solid transportation network*

3. Fundamental Concepts

Definition 1. (Regular Fuzzy Variable)
Let Γ be the universe of discourse. An ample field [32] $\mathcal{A}$ on Γ is a class of subsets of Γ that is closed under arbitrary unions, intersections, and complements in Γ.
Let $Pos : \mathcal{A} \to [0,1]$ be a set function on the amplefield Γ. Pos is said to be a possibility measure [32] if it satisfies the following conditions:
(p1) $Pos(\varphi) = 0$ and $Pos(\Gamma) = 1$.
(p2) For any subclass $\{A_i | i \in I\}$ of $\mathcal{A}$ (finite, counter or uncountable), $Pos(\bigcup_{i\in I}(A_i)) = Sup_{i\in I} Pos(A_i)$
The triplet $(\Gamma, \mathcal{A}, Pos)$ is referred to as a possibility space, in which a credibility measure [33] is defined as $Cr(A) = \frac{1}{2}(1 + Pos(A) - Pos(A^c)), A \in \mathcal{A}$.
If $(\Gamma, \mathcal{A}, Pos)$ is a possibility space, then an m-ary regular fuzzy vector $\xi = (\xi_1, \xi_2, ..., \xi_m)$ is defined as a membership map from Γ to the space $[0,1]^m$ in the sense that for every $t = (t_1, t_2, ..., t_m) \in [0,1]^m$, one has
$\{\gamma \in \Gamma | \xi(\gamma) \le t\} = \{\gamma \in \Gamma | \xi_1(\gamma) \le t_1, \xi_2(\gamma) \le t_2, ..., \xi_m(\gamma) \le t_m\} \in \mathcal{A}$
When $m = 1, \xi$ is called a regular fuzzy variable (RFV).

3.1. Critical values for RFVs. .

Definition 2.(Qin et al. [1]) Let ξ be an RFV. Then the optimistic CV of ξ, denoted by $CV^*[\xi]$, defined as $CV^*[\xi] = \underbrace{Sup\{\alpha \wedge Pos\{\xi \ge \alpha\}\}}_{\alpha\in[0,1]}$, while the pessimistic CV of ξ, denoted by $CV_*[\xi]$, is defined as $CV_*[\xi] = \underbrace{Sup\{\alpha \wedge Nec\{\xi \ge \alpha\}\}}_{\alpha\in[0,1]}$.
The CV of ξ , denoted by $CV[\xi]$, is defined as $CV[\xi] = \underbrace{Sup\{\alpha \wedge Cr\{\xi \ge \alpha\}\}}_{\alpha\in[0,1]}$.

Theorem 1. (Qin et al. [1]) Let $\xi = (r_1, r_2, r_3, r_4)$ be a trapezoidal RFV. Then we have

(i) The optimistic CV of ξ is $CV^*[\xi] = \frac{r_4}{1+r_4-r_3}$.
(ii) The pessimistic CV of ξ is $CV_*[\xi] = \frac{r_2}{1+r_2-r_1}$.
(iii) The CV of ξ is $CV[\xi] = \begin{cases} \frac{2r_2-r_1}{1+2(r_2-r_1)}, & \text{if } r_2 > \frac{1}{2} \\ \frac{1}{2}, & \text{if } r_2 \le \frac{1}{2} \le r_3 \\ \frac{r_4}{1+2(r_4-r_3)}, & r_3 \le \frac{1}{2} \end{cases}$

3.2. Methods of reduction for type-2 fuzzy variables (CV-Based Reduction Methods).

Due to the fuzzy membership function of a type-2 fuzzy number, the computation complexity is very high in practical applications. To avoid this difficulty, some defuzzification methods have been proposed in the literature (see [6-8]). In this section, we propose some new methods of reduction for a type-2 fuzzy variable. Compared with the existing methods, the new methods are very much easier to implement when we employ them to build a mathematical model with type-2 fuzzy coefficients.
Let $(\Gamma, \mathcal{A}, \tilde{Pos})$ be a fuzzy possibility space and $\tilde{\xi}$ a type-2 fuzzy variable with a known secondary possibility distribution function $\mu_{\tilde{\xi}}(x)$. To reduce the type-2 fuzziness, one approach is to give a representing value for RFV $\mu_{\tilde{\xi}}(x)$. For this purpose, we suggest employing the CVs of $\tilde{Pos}\{\gamma | \tilde{\xi}(\gamma) = x\}$ as the representing values. This methods the CV-based methods for the type-2 fuzzy variable $\tilde{\xi}$
Theorem 2.(Qin et al. [1]) Let $\tilde{\xi}$ be a type-2 triangular fuzzy variable defined as $\tilde{\xi} = (\tilde{r}_1, \tilde{r}_2, \tilde{r}_3; \theta_l, \theta r)$. Then we have

(i) Using the optimistic CV reduction method, the reduction ξ_1 of $\tilde{\xi}$ has the following possibility distribution:

$$\mu_{\xi_1}(x)=\begin{cases}\frac{(1+\theta_r)(x-r_1)}{(r_2-r_1+\theta_r(x-r_1))}, & \text{if } x\in[r_1,\frac{r_1+r_2}{2}]\\ \frac{(1-\theta_r)x+\theta_r r_2-r_1}{(r_2-r_1+\theta_l(r_2-x))}, & \text{if } x\in[\frac{r_1+r_2}{2},r_2]\\ \frac{(-1+\theta_r)x-\theta_r r_2+r_3}{r_3-r_2+\theta_r(x-r_2)}, & \text{if } x\in[r_2,\frac{r_2+r_3}{2}]\\ \frac{(1+\theta_r)(r_3-x)}{r_3-r_2+\theta_r(r_3-x)}, & \text{if } x\in[\frac{r_2+r_3}{2},r_3]\end{cases}$$

(ii) Using the pessimistic CV reduction method, the reduction ξ_2 of $\tilde{\xi}$ has the following possibility distribution:

$$\mu_{\xi_2}(x)=\begin{cases}\frac{x-r_1}{r_2-r_1+\theta_l(x-r_1)}, & \text{if } x\in[r_1,\frac{r_1+r_2}{2}]\\ \frac{x-r_1}{r_2-r_1+\theta_l(r_2-x)}, & \text{if } x\in[\frac{r_1+r_2}{2},r_2]\\ \frac{r_3-x}{r_3-r_2+\theta_l(x-r_2)}, & \text{if } x\in[r_2,\frac{r_2+r_3}{2}]\\ \frac{r_3-x}{r_3-r_2+\theta_l(r_3-x)}, & \text{if } x\in[\frac{r_2+r_3}{2},r_3]\end{cases}$$

(iii) Using the CV reduction method, the reduction ξ_3 of $\tilde{\xi}$ has the following possibility distribution:

$$\mu_{\xi_3}(x)=\begin{cases}\frac{(1+\theta_r)(x-r_1)}{r_2-r_1+2\theta_r(x-r_1)}, & \text{if } x\in[r_1,\frac{r_1+r_2}{2}]\\ \frac{(1-\theta_l)x+\theta_l r_2-r_1}{r_2-r_1+2\theta_l(r_2-x)}, & \text{if } x\in[\frac{r_1+r_2}{2},r_2]\\ \frac{(-1+\theta_l)x-\theta_l r_2+r_3}{r_3-r_2+2\theta_l(x-r_2)}, & \text{if } x\in[r_2,\frac{r_2+r_3}{2}]\\ \frac{(1+\theta_r)(r_3-x)}{r_3-r_2+2\theta_r(r_3-x)}, & \text{if } x\in[\frac{r_2+r_3}{2},r_3]\end{cases}$$

3.3. Generalized credibility and its properties. Suppose ξ is a general fuzzy variable with the distribution μ. The generalized credibility measure $\tilde{Cr}$ of the event $\{\xi\geq\alpha\}$ is defined by
$\tilde{Cr}(\{\xi\geq\alpha\})=\frac{1}{2}(Sup_{x\in\Re}\mu(x)+Sup_{x\geq r}\mu(x)-Sup_{x<r}\mu(x)), r\in\Re$.
Therefore, if ξ is normalized, it is easy to check that $Cr(\xi\geq\alpha)+Cr(\xi<\alpha)=Sup_{x\in\Re}\mu_\xi(x)=1$; then $\tilde{Cr}$ coincides with the usual credibility measure. The concept of independence for normalized fuzzy variables and its properties were discussed in [35]. In the following, we also need to extend independence to general fuzzy variables. The general fuzzy variables $\xi_1,\xi_2,\xi_2...,\xi_n$ are said to be mutually independent if and only if $\tilde{Cr}\{\xi_i\in B_i, i=1,2,...n\}=Min_{1\leq i\leq n}\tilde{Cr}\{\xi_i\in B_i\}$ for any subsets $B_i, i=1,2,...n$ of $\Re$
Like the α-optimistic value of the normalized fuzzy variable [36], the α-optimistic value of general fuzzy variables can be defined through the generalized credibility measure. Let ξ be a fuzzy variable (not necessary normalized). Then $\xi_{Sup}(\alpha)=Sup\{r|\tilde{Cr}\{\xi\geq r\}\geq\alpha\}, \alpha\in[0,1]$ is called the α-optimistic value of ξ, while $\xi_{inf}=Inf\{r|\tilde{Cr}\{\xi\leq r\}\geq\alpha\}, \alpha\in[0,1]$, is called the $\alpha-$ pessimistic value of ξ.
Theorem 3. (Qin et al. [1]) Let ξ_i be the reduction of the type-2 fuzzy variable$\tilde{\xi_i}=(\tilde{r}_1^i,\tilde{r}_2^i,\tilde{r}_3^i;\theta_{l,i},\theta_{r,i})$ obtained by the CV reduction method for $i=1,2,...n$. Suppose $\xi_1,\xi_2,...,\xi_n$ are mutually independent, and $k_i\geq o$ for $i=1,2,...n$.
Case-I: If $\alpha\in(0,0.25]$, then eqnarray $\tilde{Cr}\{\sum_{i=1}^n k_i\xi_i\leq t\}\geq\alpha$ is equivalent to

$$\sum_{i=1}^{n}\frac{(1-2\alpha+(1-4\alpha)\theta_{r,i})k_ir_1^i+2\alpha k_ir_2^i}{1+(1-4\alpha)\theta_{r,i}}$$

Case-II: If $\alpha\in(0.25,0.50]$, then eqnarray $\tilde{Cr}\{\sum_{i=1}^n k_i\xi_i\leq t\}\geq\alpha$ is equivalent to

$$\sum_{i=1}^{n}\frac{(1-2\alpha)k_ir_1^i+(2\alpha+(4\alpha-1)\theta_{l,i})k_ir_2^i}{1+(1-4\alpha)\theta_{l,i}}$$

Case-III: If $\alpha \in (0.50, 0.75]$, then eqnarray $\tilde{Cr}\{\sum_{i=1}^{n} k_i\xi_i \leq t\} \geq \alpha$ is equivalent to

$$\sum_{i=1}^{n} \frac{(2\alpha-1)k_i r_3^i + (2(1-\alpha)+(3-4\alpha)\theta_{l,i})k_i r_2^i}{1+(3-4\alpha)\theta_{l,i}}$$

Case-IV: If $\alpha \in (0.75, 1]$, then eqnarray $\tilde{Cr}\{\sum_{i=1}^{n} k_i\xi_i \leq t\} \geq \alpha$ is equivalent to

$$\sum_{i=1}^{n} \frac{(2\alpha-1+(4\alpha-3)\theta_{r,i})k_i r_3^i + 2(1-\alpha)k_i r_2^i}{1+(4\alpha-3)\theta_{r,i}}$$

3.4. Goal programming Method. .
The goal programming method is used to solve the multi-objective programming problem (MOPP). A general MOPP is of the following form:

$$(3.1) \begin{cases} \text{Find the values of L decision variables} \, x_1, x_2, ..., x_L \, \text{which minimizes} \\ F(x) = (f_1(x), f_2(x), ..., f_Q(x))^T \\ \text{subject to} \, x \in X \end{cases}$$

Where, $X = \{x = (x_1, x_2, ..., x_L)$ such that $g_t(x) \leq 0, x_l \geq 0, t = 1, 2, ..., T; l = 1, 2, ..., L\}$ and $f_1(x), f_2(x), ..., f_Q(x)$ are $Q(\geq 2)$ objective functions.
The different steps of the goal programming method are as follows:
Step-1: Solve the multi-objective programming problem (1) as a single objective problem using only one objective at a time ignoring the others, and determine the ideal objective vector, say $f_1^{min}, f_2^{min}, ..., f_Q^{min}$.
Step-2: Formulate the following GP problem using the ideal objective vector obtained is Step-1,

$$Min\{\sum_{q=1}^{Q}[(d_q^+)^p + (d_q^-)^p]\}^{\frac{1}{p}}$$

subject to $f_q(x) + d_q^+ - d_q^- = f_q^{min}, d_q^+ \geq 0, d_q^- \geq 0, d_q^+ d_q^- = 0 (q = 1, 2, ..., Q)$, for all $x \in X$.
Step-3: Now, solve the above single objective problem described in Step-2 by GRG method and obtain the compromise solution.

4. Notations and assumptions for the proposed model

(i) $\tilde{C}_{ijk_1q}^1, \tilde{C}_{jkk_2q}^2, \tilde{C}_{lmk_nq}^n$ = Fuzzy unit transportation cost is to transport the q-th item from i-th plant to j-th DC by k_1-th vehicle, j-th plant to k-th DC k_2-th vehicle and l-th plant to m-th customer k_n-th vehicle respectively.

(ii) $\tilde{t}_{ijk_1q}^1, \tilde{t}_{jkk_2q}^2, \tilde{t}_{lmk_nq}^n$ = Fuzzy unit transportation time is to transport the q-th item from i-th plant to j-th DC by k_1-th vehicle, j-th plant to k-th DC k_2-th vehicle and l-th plant to m-th customer k_n-th vehicle respectively.

(iii) $\tilde{x}_{ijk_1q}^1, \tilde{x}_{jkk_2q}^2, \tilde{x}_{lmk_nq}^n, \tilde{x}_{ulk_{(n-1)}q}^n, \tilde{x}_{lmk_nq}^n$ = Unknown quantities which is to be transported from i-th plant to j-th DC of q-th item by k_1-th vehicle for stage-1, j-th plant to k-th DC of q-th item by k_2-th vehicle for stage-2, u-th plant to l-th DC of q-th item by $k_{(n-1)}$-th vehicle for stage-$(n-1)$, and l-th plant to m-th customer of q-th item by k_n-th vehicle for stage-n respectively.

(iv) $\tilde{PC}$ = Fuzzy purchasing cost of q-th item at i-th source.

(v) $\tilde{LO}_i^1, \tilde{LO}_j^2, ..., \tilde{LO}_l^n$ = Fuzzy loading cost at i-th plant of stage-1, j-th plant of stage-2 and l-th plant of stage-n respectively.

(vi) $\tilde{UD}_j^1, \tilde{UD}_k^2, ..., \tilde{UD}_m^n$ =Fuzzy unloading cost at j-th DC of stage-1, k-th DC of stage-2 and m-th customer of stage-n respectively.

(vii) $L\tilde{T}O_i^1, L\tilde{T}O_j^2, ..., L\tilde{T}O_l^n$ = Fuzzy loading time at i-th plant of stage-1, j-th plant of stage-2 and l-th plant of stage-n respectively.

(viii) $U\tilde{T}D_j^1, U\tilde{T}D_k^2, ..., U\tilde{T}D_m^n$ =Fuzzy unloading time at j-th DC of stage-1, k-th DC of stage-2 and m-th customer of stage-n respectively.

(ix) $y_{ijk_1q}^1 = \begin{cases} 1, if\ x_{ijk_1q}^1 > 0 \\ 0,\ otherwise \end{cases}$, $y_{jkk_2q}^2 = \begin{cases} 1, if\ x_{jkk_2q}^2 > 0 \\ 0,\ otherwise \end{cases}$, $y_{lmk_nq}^n = \begin{cases} 1, if\ x_{lmk_nq}^n > 0 \\ 0,\ otherwise \end{cases}$

5. Formulation of solid transportation problem with transportation parameters as type-2 triangular fuzzy variables

Let us consider $'I'$ supply points (or sources), $'J'$ destination centers, K_1 conveyances for stage-1 transportation; $'J'$ supply points (or sources), $'K'$ destination centers, K_2 conveyances for stage-2 transportation; $'U'$ supply points, $'L'$ destination centers, $k_{(n-1)}$ conveyances for stage-$(n-1)$ transportation; $'L'$ supply points (or sources), $'M'$ destination centers, K_n conveyances for stage-n transportation. Also we consider that Q be the number of items which is to be transported from plants to DC by different modes of conveyances.

$$\begin{aligned}(5.1)\qquad Minf_1 = &\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k_1=1}^{K_1}(\tilde{C}_{ijk_1q}^1 + \tilde{LO}_i^1 + \tilde{UD}_j^1 + \tilde{PC}_{iq}^1)x_{ijk_1q}^1 \\ &+\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{k_2=1}^{K_2}(\tilde{C}_{jkk_2q}^2 + \tilde{LO}_j^2 + \tilde{UD}_k^2 + \tilde{PC}_{jq}^2)x_{jkk_2q}^2 + ... \\ &+\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{k_n=1}^{K_n}(\tilde{C}_{lmk_nq}^n + \tilde{LO}_l^n + \tilde{UD}_m^n + \tilde{PC}_{lq}^n)x_{lmk_nq}^n\end{aligned}$$

$$\begin{aligned}(5.2)\qquad Minf_2 = &\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k_1=1}^{K_1}(\tilde{t}_{ijk_1q}^1 + L\tilde{T}O_i^1 + U\tilde{T}D_j^1)y_{ijk_1q} \\ &+\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{k_2=1}^{K_2}(\tilde{t}_{jkk_2q}^2 + L\tilde{T}O_j^2 + U\tilde{T}D_k^2)y_{jkk_2q}^2 + ... \\ &+\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{k_n=1}^{K_n}(\tilde{t}_{lmk_nq}^n + L\tilde{T}O_l^n + U\tilde{T}D_m^n)y_{lmk_nq}^n\end{aligned}$$

$$(5.3)\quad \sum_{j=1}^{J}\sum_{k_1=1}^{K_1} x^1_{ijk_1q} \le \tilde{a}^1_{iq}, i=1,2,...,I; q=1,2,...,Q,$$

$$(5.4)\quad \sum_{i=1}^{I}\sum_{k_1=1}^{K_1} x^1_{ijk_1q} \ge \tilde{b}^1_{jq}, j=1,2,...,J; q=1,2,...,Q,$$

$$(5.5)\quad \sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J} x^1_{ijk_1q} \le \tilde{e}^1_{k_1}, k_1=1,2,...,K_1,$$

$$(5.6)\quad \sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J} \tilde{w}_q x^1_{ijk_1q} \le \tilde{W}^1_{k_1}, k_1=1,2,...,K_1,$$

$$(5.7)\quad \sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J} \tilde{v}_q x^1_{ijk_1q} \le \tilde{V}^1_{k_1}, k_1=1,2,...,K_1,$$

$$(5.8)\quad \sum_{k=1}^{K}\sum_{k_2=1}^{K_2} x^2_{jkk_2q} \le \sum_{i=1}^{I}\sum_{k_1=1}^{K_1} x^1_{ijk_1q}, j=1,2,...,J; q=1,2,...,Q,$$

$$(5.9)\quad \sum_{j=1}^{J}\sum_{k_2=1}^{K_2} x^2_{jkk_2q} \ge \tilde{b}^2_{kq} k=1,2,...K; q=1,2,...,Q,$$

$$(5.10)\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K} x^2_{jkk_2q} \le \tilde{e}^2_{k_2}, k_2=1,2,...,K_2,$$

$$(5.11)\quad \sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K} \tilde{w}_q x^2_{jkk_2q} \le \tilde{W}^2_{k_2}, k_2=1,2,...,K_2,$$

$$(5.12)\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K} \tilde{v}_q x^2_{jkk_2q} \le \tilde{V}^2_{k_2}, k_2=1,2,...,K_2,$$

....

$$(5.13)\quad \sum_{m=1}^{M}\sum_{k_n=1}^{K_n} x^n_{lmk_nq} \le \sum_{u=1}^{U}\sum_{k_{(n-1)}=1}^{K_{(n-1)}} x^{(n-1)}_{ulk_{(n-1)}q}, l=1,2,...,L; q=1,2,...,Q,$$

$$(5.14)\quad \sum_{l=1}^{L}\sum_{k_n=1}^{K_1} x^n_{lmk_nq} \ge \tilde{b}^n_{mq}, m=1,2,...M; q=1,2,...,Q,$$

$$(5.15)\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M} x^n_{lmk_nq} \le \tilde{e}^n_{k_n}, k_n=1,2,...,K_n,$$

$$(5.16)\quad \sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M} \tilde{w}_q x^n_{lmk_nq} \le \tilde{W}^n_{k_n}, k_n=1,2,...,K_n,$$

$$(5.17)\quad \sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M} \tilde{v}_q x^n_{lmk_nq} \le \tilde{V}^n_{k_n}, k_n=1,2,...,K_n,$$

$$x^1_{ijk_1q} \ge 0, x^2_{jkk_2q} \ge 0, ..., x^{(n-1)}_{ulk_{(n-1)}q} \ge 0, x^n_{lmk_nq} \ge 0,$$

for all $i, j, k, u, l, m, q, k_1, k_2, k_{(n-1)}, k_n$.

where, $\tilde{W}^1_{k_1}, \tilde{W}^2_{k_2}, \tilde{W}^n_{k_n}$ are the fuzzy weight capacity of k_1-th vehicle of stage-1, k_2-th vehicle of stage-2, k_n-th vehicle of stage-n. $\tilde{V}^1_{k_1}, \tilde{V}^2_{k_2}, \tilde{V}^n_{k_n}$ are the fuzzy volume capacity of k_1-th vehicle of stage-1, k_2-th vehicle of stage-2, k_n-th vehicle of stage-n. $\tilde{w}_q, \tilde{v}_q$ are the fuzzy weight and volume of the q-th item. Also $\tilde{a}^1_{iq}$ be the fuzzy availabilities of the q-th item at i-th source of stage-1. b^1_{jq}, b^2_{kq}, and b^n_{mq} are the fuzzy demands of q-th item at j-th DC, k-th DC and m-th customer for stage-1, stage-2 and stage-3 respectively. Also, $\tilde{e}^1_{k_1}\tilde{e}^2_{k_2}, \tilde{e}^n_{k_n}$ are the fuzzy conveyances capacities of the k_1-th, k_2-th,k_n-th conveyances for stage-1, stage-2 and stage-n respectively. In this model formulation, we are to minimize two objective functions as total cost and time under supply, demand, conveyance capacity, weight and volume constraints. Here the first summation of the first objective indicates the total cost for stage-1 transportation. Similarly, second and last summation of the first objective function indicates the total cost for stage-2 and stage-n transportation respectively. Also the three summations of second objective denotes the total time in transportation respectively for stage-1, stage-2 and stage-n respectively. We formulate the model in such a way that the goods are loaded at the supply point and it is unloaded at the DC for stage-1 transportation. Since due to disaster, it is not possible to move the vehicle directly to the survived people so after unloading at the first DC it again loaded to another vehicle and goes to the next DC and it is unloaded again in second DC for stage-2. In this way the necessary goods are transported to the survived peoples or customers. For this reason, we impose the loading and unloading cost and time for each stage. Again purchasing cost is also imposed in our model.

6. Methodology and defuzzification technique used to solve the Model

6.1. Methodology. The world has become more complex and almost every important real-world problem involves more than one objective. In such cases, decision makers find imperative to evaluate best possible approximate solution alternatives according to multiple criteria. To solve such multi-objective programming problem we apply goal programming method. Using CV-based reduction method and generalized credibility measure we find the deterministic form of type-2 fuzzy transportation parameters. Finally generalized reduced gradient technique (LINGO 13.0 optimization software) is used to solve the developed model.

6.2. Defuzzification. The deterministic form of the objective functions and constraints obtained by using CV-based reduction method and generalized credibility measure are as follows:

$$
\begin{aligned}
(6.1)\quad & Cr\{(\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k_1=1}^{K_1}(\tilde{C}^1_{ijk_1q} + \tilde{LO}^1_i + \tilde{UD}^1_j + \tilde{PC}^1_{iq})x^1_{ijk_1q} \\
& + \sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{k_2=1}^{K_2}(\tilde{C}^2_{jkk_2q} + \tilde{LO}^2_j + \tilde{UD}^2_k + \tilde{PC}^2_{jq})x^2_{jkk_2q} + ... \\
& + \sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{k_n=1}^{K_n}(\tilde{C}^n_{lmk_nq} + \tilde{LO}^n_l + \tilde{UD}^n_m + \tilde{PC}^n_{lq})x^n_{lmk_nq}) \geq f_1\} \leq \alpha_c
\end{aligned}
$$

$$(6.2)\quad Cr\{(\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k_1=1}^{K_1}(\tilde{t}^1_{ijk_1q}+L\tilde{T}O^1_i+U\tilde{T}D^1_j)y^1_{ijk_1q}$$
$$+\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{k_2=1}^{K_2}(\tilde{t}^2_{jkk_2q}+L\tilde{T}O^2_j+U\tilde{T}D^2_k)y^2_{jkk_2q}+...$$
$$+\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{k_n=1}^{K_n}(\tilde{t}^n_{lmk_nq}+L\tilde{T}O^n_l+U\tilde{T}D^n_m)y^n_{lmk_nq})\geq f_2\}\leq\alpha_t$$

$$(6.3)\quad Cr(\sum_{j=1}^{J}\sum_{k_1=1}^{K_1}x^1_{ijk_1q}\leq\tilde{a}^1_{iq})\geq\alpha_{avail.}, i=1,2,...,I; q=1,2,...,Q,$$

$$(6.4)\quad Cr(\sum_{i=1}^{I}\sum_{k_1=1}^{K_1}x^1_{ijk_1q}\geq\tilde{b}^1_{jq})\geq\alpha_{demand}, j=1,2,...,J; q=1,2,...,Q,$$

$$(6.5)\quad Cr(\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}x^1_{ijk_1q}\leq\tilde{e}^1_{k_1})\geq\alpha_{con.cap.}, k_1=1,2,...,K_1,$$

$$(6.6)\quad Cr(\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}\tilde{w}_qx^1_{ijk_1q}\leq\tilde{W}^1_{k_1})\geq\alpha_{weight}, k_1=1,2,...,K_1,$$

$$(6.7)\quad Cr(\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}\tilde{v}_qx^1_{ijk_1q}\leq\tilde{V}^1_{k_1})\geq\alpha_{volume}, k_1=1,2,...,K_1,$$

$$(6.8)\quad Cr(\sum_{j=1}^{J}\sum_{k_2=1}^{K_2}x^2_{jkk_2q}\geq\tilde{b}^2_{kq})\geq\alpha_{demand}, k=1,2,...K; q=1,2,...,Q,$$

$$(6.9)\quad Cr(\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}x^2_{jkk_2q}\leq\tilde{e}^2_{k_2})\geq\alpha_{con.cap.}, k_2=1,2,...,K_2,$$

$$(6.10)\quad Cr(\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{w}_qx^2_{jkk_2q}\leq\tilde{W}^2_{k_2})\geq\alpha_{weight}, k_2=1,2,...,K_2,$$

$$(6.11)\quad Cr(\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}\tilde{v}_qx^2_{jkk_2q}\leq\tilde{V}^2_{k_2})\geq\alpha_{volume}, k_2=1,2,...,K_2,$$

....

$$(6.12)\quad Cr(\sum_{l=1}^{L}\sum_{k_n=1}^{K_1}x^n_{lmk_nq}\geq\tilde{b}^n_{mq})\geq\alpha_{demand}, m=1,2,...M; q=1,2,...,Q,$$

$$(6.13)\quad Cr(\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}x^n_{lmk_nq}\leq\tilde{e}^n_{k_n})\geq\alpha_{con.cap.}v, k_n=1,2,...,K_n,$$

$$(6.14)\quad Cr(\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}\tilde{w}_qx^n_{lmk_nq}\leq\tilde{W}^n_{k_n})\geq\alpha_{weight}, k_n=1,2,...,K_n,$$

$$(6.15)\quad Cr(\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}\tilde{v}_qx^n_{lmk_nq}\leq\tilde{V}^n_{k_n})\geq\alpha_{volume}, k_n=1,2,...,K_n,$$

Let us consider $\alpha_c, \alpha_t, \alpha_{avail.}, \alpha_{demand}, \alpha_{con.cap.}, \alpha_{weight}, \alpha_{volume}$ be the credibility level for cost, time, availabilities, demands, conveyances capacities, weights, volume respectively for stage-1, stage-2,...,stage-n.
The crisp conversion of the constraints (6.1)-(6.15) are as follows:

$$
\begin{aligned}
Minf_1 = & \sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k_1=1}^{K_1}(S_{\tilde{C}^1_{ijk_1q}} + S_{\tilde{LO}^1_i} + S_{\tilde{UD}^1_j} + S_{\tilde{PC}^1_{iq}})x^1_{ijk_1q} \\
& + \sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{k_2=1}^{K_2}(S_{\tilde{C}^2_{jkk_2q}} + S_{\tilde{LO}^2_j} + S_{\tilde{UD}^2_k} + S_{\tilde{PC}^2_{jq}})x^2_{jkk_2q} + ... \\
& + \sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{k_n=1}^{K_n}(S_{\tilde{C}^n_{lmk_nq}} + S_{\tilde{LO}^n_l} + S_{\tilde{UD}^n_m} + S_{\tilde{PC}^n_{lq}})x^n_{lmk_nq} \\
Minf_2 = & \sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k_1=1}^{K_1}(S_{\tilde{t}^1_{ijk_1q}} + S_{L\tilde{T}O^1_i} + S_{U\tilde{T}D^1_j})y_{ijk_1q} \\
& + \sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{k_2=1}^{K_2}(S_{\tilde{t}^2_{jkk_2q}} + S_{L\tilde{T}O^2_j} + S_{U\tilde{T}D^2_k})y^2_{jkk_2q} + ... \\
& + \sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{k_n=1}^{K_n}(S_{\tilde{t}^n_{lmk_nq}} + S_{L\tilde{T}O^n_l} + S_{U\tilde{T}D^n_m})y^n_{lmk_nq}
\end{aligned}
$$

$$\sum_{j=1}^{J}\sum_{k_1=1}^{K_1} x^1_{ijk_1q} \le S_{\tilde{a}^1_{iq}}, i = 1, 2, ..., I; q = 1, 2, ..., Q,$$

$$\sum_{i=1}^{I}\sum_{k_1=1}^{K_1} x^1_{ijk_1q} \ge S_{\tilde{b}^1_{jq}}, j = 1, 2, ..., J; q = 1, 2, ..., Q,$$

$$\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J} x^1_{ijk_1q} \le S_{\tilde{e}^1_{k_1}}, k_1 = 1, 2, ..., K_1,$$

$$\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J} S_{\tilde{w}_q} x^1_{ijk_1q} \le S_{\tilde{W}^1_{k_1}}, k_1 = 1, 2, ..., K_1,$$

$$\sum_{q=1}^{Q}\sum_{i=1}^{I}\sum_{j=1}^{J} S_{\tilde{v}_q} x^1_{ijk_1q} \le S_{\tilde{V}^1_{k_1}}, k_1 = 1, 2, ..., K_1,$$

$$\sum_{j=1}^{J}\sum_{k_2=1}^{K_2} x^2_{jkk_2q} \geq S_{\tilde{b}^2_{kq}} k = 1,2,...K; q = 1,2,...,Q,$$

$$\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K} x^2_{jkk_2q} \leq S_{\tilde{e}^2_{k_2}}, k_2 = 1,2,...,K_2,$$

$$\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K} S_{\tilde{w}_q} x^2_{jkk_2q} \leq S_{\tilde{W}^2_{k_2}}, k_2 = 1,2,...,K_2,$$

$$\sum_{q=1}^{Q}\sum_{j=1}^{J}\sum_{k=1}^{K} S_{\tilde{v}_q} x^2_{jkk_2q} \leq S_{\tilde{V}^2_{k_2}}, k_2 = 1,2,...,K_2,$$

....

$$\sum_{l=1}^{L}\sum_{k_n=1}^{K_1} x^n_{lmk_nq} \geq S_{\tilde{b}^n_{mq}}, m = 1,2,...M; q = 1,2,...,Q,$$

$$\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M} x^n_{lmk_nq} \leq S_{\tilde{e}^n_{k_n}}, k_n = 1,2,...,K_n,$$

$$\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M} S_{\tilde{w}_q x^n_{lmk_nq}} \leq S_{\tilde{W}^n_{k_n}}, k_n = 1,2,...,K_n,$$

$$\sum_{q=1}^{Q}\sum_{l=1}^{L}\sum_{m=1}^{M} S_{\tilde{v}_q} x^n_{lmk_nq} \leq S_{\tilde{V}^n_{k_n}}, k_n = 1,2,...,K_n,$$

Where $S_{\tilde{C}^1_{ijk_1q}}$, $S_{\tilde{C}^2_{jkk_2q}}$, $S_{\tilde{C}^n_{lmk_nq}}$, $S_{\tilde{t}^1_{ijk_1q}}$, $S_{\tilde{t}^2_{jkk_2q}}$, $S_{\tilde{t}^n_{lmk_nq}}$ $S_{\tilde{LO}^1_i}$, $S_{\tilde{LO}^2_j}$, $S_{\tilde{LO}^n_l}$, $S_{\tilde{UD}^1_j}$, $S_{\tilde{UD}^2_k}$, $S_{\tilde{UD}^n_m}$, $S_{\tilde{PC}^1_{iq}}$, $S_{\tilde{PC}^2_{jq}}$, $S_{\tilde{PC}^n_{lq}}$, $S_{L\tilde{T}O^1_i}$, $S_{L\tilde{T}O^2_j}$, $S_{L\tilde{T}O^n_l}$, $S_{U\tilde{T}D^1_j}$, $S_{U\tilde{T}D^2_k}$, $S_{U\tilde{T}D^n_m}$,, $S_{\tilde{a}^1_{iq}}$, $S_{\tilde{b}^1_{jq}}$, $S_{\tilde{b}^2_{kq}}$, $S_{\tilde{b}^n_{mq}}$, $S_{\tilde{e}^1_{k_1}}$, $S_{\tilde{e}^2_{k_2}}$, $S_{\tilde{e}^n_{k_n}}$, $S_{\tilde{w}_q}$, $S_{\tilde{W}^1_{k_1}}$, $S_{\tilde{v}_q}$, $S_{\tilde{V}^1_{k_1}}$, $S_{\tilde{W}^2_{k_2}}$, $S_{\tilde{V}^2_{k_2}}$, $S_{\tilde{W}^n_{k_n}}$, $S_{\tilde{V}^n_{k_n}}$ are equivalent crisp form of fuzzy parameters respectively and given as follows:

$$S_{\tilde{C}^1_{ijk_1q}} = \begin{cases} \frac{(1-2\alpha+(1-4\alpha_c)\theta_{r,\tilde{C}^1_{ijk_1q}})r_1^{\tilde{C}^1_{ijk_1q}} + 2\alpha_c r_2^{\tilde{C}^1_{ijk_1q}}}{(1+(1-4\alpha_c)\theta_{r,\tilde{C}^1_{ijk_1q}})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{C}^1_{ijk_1q}} + (2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{C}^1_{ijk_1q}})r_2^{\tilde{C}^1_{ijk_1q}}}{(1+(1-4\alpha_c)\theta_{l,\tilde{C}^1_{ijk_1q}})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{C}^1_{ijk_1q}} + (2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{C}^1_{ijk_1q}})r_2^{\tilde{C}^1_{ijk_1q}}}{(1+(3-4\alpha_c)\theta_{l,\tilde{C}^1_{ijk_1q}})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{C}^1_{ijk_1q}})r_3^{\tilde{C}^1_{ijk_1q}} + 2(1-\alpha_c)r_2^{\tilde{C}^1_{ijk_1q}}}{(1+(4\alpha_c-3)\theta_{r,\tilde{C}^1_{ijk_1q}})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{LO}^1_i} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{LO}^1_i})r_1^{\tilde{LO}^1_i} + 2\alpha_c r_2^{\tilde{LO}^1_i}}{(1+(1-4\alpha_c)\theta_{r,\tilde{LO}^1_i})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{LO}^1_i} + (2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{LO}^1_i})r_2^{\tilde{LO}^1_i}}{(1+(1-4\alpha_c)\theta_{l,\tilde{LO}^1_i})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{LO}^1_i} + (2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{LO}^1_i})r_2^{\tilde{LO}^1_i}}{(1+(3-4\alpha_c)\theta_{l,\tilde{LO}^1_i})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{LO}^1_i})r_3^{\tilde{LO}^1_i} + 2(1-\alpha_c)r_2^{\tilde{LO}^1_i}}{(1+(4\alpha_c-3)\theta_{r,\tilde{LO}^1_i})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{UD}_j^1} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{UD}_j^1})r_1^{\tilde{UD}_j^1}+2\alpha_c r_2^{\tilde{UD}_j^1}}{(1+(1-4\alpha_c)\theta_{r,\tilde{UD}_j^1})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{UD}_j^1}+(2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{UD}_j^1})r_2^{\tilde{UD}_j^1}}{(1+(1-4\alpha_c)\theta_{l,\tilde{UD}_j^1})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{UD}_j^1}+(2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{UD}_j^1})r_2^{\tilde{UD}_j^1}}{(1+(3-4\alpha_c)\theta_{l,\tilde{UD}_j^1})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{UD}_j^1})r_3^{\tilde{UD}_j^1}+2(1-\alpha_c)r_2^{\tilde{UD}_j^1}}{(1+(4\alpha_c-3)\theta_{r,\tilde{UD}_j^1})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{PC}_{iq}} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{PC}_{iq}})r_1^{\tilde{PC}_{iq}}+2\alpha_c r_2^{\tilde{PC}_{iq}}}{(1+(1-4\alpha_c)\theta_{r,\tilde{PC}_{iq}})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{PC}_{iq}}+(2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{PC}_{iq}})r_2^{\tilde{PC}_{iq}}}{(1+(1-4\alpha_c)\theta_{l,\tilde{PC}_{iq}})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{PC}_{iq}}+(2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{PC}_{iq}})r_2^{\tilde{PC}_{iq}}}{(1+(3-4\alpha_c)\theta_{l,\tilde{PC}_{iq}})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{PC}_{iq}})r_3^{\tilde{PC}_{iq}}+2(1-\alpha_c)r_2^{\tilde{PC}_{iq}}}{(1+(4\alpha_c-3)\theta_{r,\tilde{PC}_{iq}})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{C}^2_{jkk_2q}} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{C}^2_{jkk_2q}})r_1^{\tilde{C}^2_{jkk_2q}}+2\alpha_c r_2^{\tilde{C}^2_{jkk_2q}}}{(1+(1-4\alpha_c)\theta_{r,\tilde{C}^2_{jkk_2q}})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{C}^2_{jkk_2q}}+(2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{C}^2_{jkk_2q}})r_2^{\tilde{C}^2_{jkk_2q}}}{(1+(1-4\alpha_c)\theta_{l,\tilde{C}^2_{jkk_2q}})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{C}^2_{jkk_2q}}+(2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{C}^2_{jkk_2q}})r_2^{\tilde{C}^2_{jkk_2q}}}{(1+(3-4\alpha_c)\theta_{l,\tilde{C}^2_{jkk_2q}})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{C}^2_{jkk_2q}})r_3^{\tilde{C}^2_{jkk_2q}}+2(1-\alpha_c)r_2^{\tilde{C}^2_{jkk_2q}}}{(1+(4\alpha_c-3)\theta_{r,\tilde{C}^2_{jkk_2q}})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{LO}_j^2} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{LO}_j^2})r_1^{\tilde{LO}_j^2}+2\alpha_c r_2^{\tilde{LO}_j^2}}{(1+(1-4\alpha_c)\theta_{r,\tilde{LO}_j^2})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{LO}_j^2}+(2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{LO}_j^2})r_2^{\tilde{LO}_j^2}}{(1+(1-4\alpha_c)\theta_{l,\tilde{LO}_j^2})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{LO}_j^2}+(2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{LO}_j^2})r_2^{\tilde{LO}_j^2}}{(1+(3-4\alpha_c)\theta_{l,\tilde{LO}_j^2})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{LO}_j^2})r_3^{\tilde{LO}_j^2}+2(1-\alpha_c)r_2^{\tilde{LO}_j^2}}{(1+(4\alpha_c-3)\theta_{r,\tilde{LO}_j^2})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{UD}^2_k} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{UD}^2_k})r_1^{\tilde{UD}^2_k}+2\alpha_c r_2^{\tilde{UD}^2_k}}{(1+(1-4\alpha_c)\theta_{r,\tilde{UD}^2_k})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{UD}^2_k}+(2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{UD}^2_k})r_2^{\tilde{UD}^2_k}}{(1+(1-4\alpha_c)\theta_{l,\tilde{UD}^2_k})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{UD}^2_k}+(2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{UD}^2_k})r_2^{\tilde{UD}^2_k}}{(1+(3-4\alpha_c)\theta_{l,\tilde{UD}^2_k})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{UD}^2_k})r_3^{\tilde{UD}^2_k}+2(1-\alpha_c)r_2^{\tilde{UD}^2_k}}{(1+(4\alpha_c-3)\theta_{r,\tilde{UD}^2_k})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{C}^n_{lmk_nq}} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{C}^n_{lmk_nq}})r_1^{\tilde{C}^n_{lmk_nq}}+2\alpha_c r_2^{\tilde{C}^n_{lmk_nq}}}{(1+(1-4\alpha_c)\theta_{r,\tilde{C}^n_{lmk_nq}})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{C}^n_{lmk_nq}}+(2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{C}^n_{lmk_nq}})r_2^{\tilde{C}^n_{lmk_nq}}}{(1+(1-4\alpha_c)\theta_{l,\tilde{C}^n_{lmk_nq}})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{C}^n_{lmk_nq}}+(2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{C}^n_{lmk_nq}})r_2^{\tilde{C}^n_{lmk_nq}}}{(1+(3-4\alpha_c)\theta_{l,\tilde{C}^n_{lmk_nq}})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{C}^n_{lmk_nq}})r_3^{\tilde{C}^n_{lmk_nq}}+2(1-\alpha_c)r_2^{\tilde{C}^n_{lmk_nq}}}{(1+(4\alpha_c-3)\theta_{r,\tilde{C}^n_{lmk_nq}})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{LO}^n_l} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{LO}^n_l})r_1^{\tilde{LO}^n_l}+2\alpha_c r_2^{\tilde{LO}^n_l}}{(1+(1-4\alpha_c)\theta_{r,\tilde{LO}^n_l})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{LO}^n_l}+(2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{LO}^n_l})r_2^{\tilde{LO}^n_l}}{(1+(1-4\alpha_c)\theta_{l,\tilde{LO}^n_l})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{LO}^n_l}+(2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{LO}^n_l})r_2^{\tilde{LO}^n_l}}{(1+(3-4\alpha_c)\theta_{l,\tilde{LO}^n_l})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{LO}^n_l})r_3^{\tilde{LO}^n_l}+2(1-\alpha_c)r_2^{\tilde{LO}^n_l}}{(1+(4\alpha_c-3)\theta_{r,\tilde{LO}^n_l})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{UD}^n_m} = \begin{cases} \frac{(1-2\alpha_c+(1-4\alpha_c)\theta_{r,\tilde{UD}^n_m})r_1^{\tilde{UD}^n_m}+2\alpha_c r_2^{\tilde{UD}^n_m}}{(1+(1-4\alpha_c)\theta_{r,\tilde{UD}^n_m})}, & \text{if } 0 < \alpha_c \leq 0.25 \\ \frac{(1-2\alpha_c)r_1^{\tilde{UD}^n_m}+(2\alpha_c+(4\alpha_c-1)\theta_{l,\tilde{UD}^n_m})r_2^{\tilde{UD}^n_m}}{(1+(1-4\alpha_c)\theta_{l,\tilde{UD}^n_m})}, & \text{if } 0.25 < \alpha_c \leq 0.50 \\ \frac{(2\alpha_c-1)r_3^{\tilde{UD}^n_m}+(2(1-\alpha_c)+(3-4\alpha_c)\theta_{l,\tilde{UD}^n_m})r_2^{\tilde{UD}^n_m}}{(1+(3-4\alpha_c)\theta_{l,\tilde{UD}^n_m})}, & \text{if } 0.50 < \alpha_c \leq 0.75 \\ \frac{(2\alpha_c-1+(4\alpha_c-3)\theta_{r,\tilde{UD}^n_m})r_3^{\tilde{UD}^n_m}+2(1-\alpha_c)r_2^{\tilde{UD}^n_m}}{(1+(4\alpha_c-3)\theta_{r,\tilde{UD}^n_m})}, & \text{if } 0.75 < \alpha_c \leq 1 \end{cases}$$

$$S_{\tilde{t}^1_{ijk_1q}} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,\tilde{t}^1_{ijk_1q}})r_1^{\tilde{t}^1_{ijk_1q}}+2\alpha_t r_2^{\tilde{t}^1_{ijk_1q}}}{(1+(1-4\alpha_t)\theta_{r,\tilde{t}^1_{ijk_1q}})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{\tilde{t}^1_{ijk_1q}}+(2\alpha_t+(4\alpha_t-1)\theta_{l,\tilde{t}^1_{ijk_1q}})r_2^{\tilde{t}^1_{ijk_1q}}}{(1+(1-4\alpha_t)\theta_{l,\tilde{t}^1_{ijk_1q}})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{\tilde{t}^1_{ijk_1q}}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,\tilde{t}^1_{ijk_1q}})r_2^{\tilde{t}^1_{ijk_1q}}}{(1+(3-4\alpha_t)\theta_{l,\tilde{t}^1_{ijk_1q}})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,\tilde{t}^1_{ijk_1q}})r_3^{\tilde{t}^1_{ijk_1q}}+2(1-\alpha_t)r_2^{\tilde{t}^1_{ijk_1q}}}{(1+(4\alpha_t-3)\theta_{r,\tilde{t}^1_{ijk_1q}})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{L\tilde{T}O_i^1} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,L\tilde{T}O_i^1})r_1^{L\tilde{T}O_i^1}+2\alpha_t r_2^{L\tilde{T}O_i^1}}{(1+(1-4\alpha_t)\theta_{r,L\tilde{T}O_i^1})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{L\tilde{T}O_i^1}+(2\alpha_t+(4\alpha_t-1)\theta_{l,L\tilde{T}O_i^1})r_2^{L\tilde{T}O_i^1}}{(1+(1-4\alpha_t)\theta_{l,L\tilde{T}O_i^1})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{L\tilde{T}O_i^1}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,L\tilde{T}O_i^1})r_2^{L\tilde{T}O_i^1}}{(1+(3-4\alpha_t)\theta_{l,L\tilde{T}O_i^1})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,L\tilde{T}O_i^1})r_3^{L\tilde{T}O_i^1}+2(1-\alpha_t)r_2^{L\tilde{T}O_i^1}}{(1+(4\alpha_t-3)\theta_{r,L\tilde{T}O_i^1})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{U\tilde{T}D_j^1} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,U\tilde{T}D_j^1})r_1^{U\tilde{T}D_j^1}+2\alpha_t r_2^{U\tilde{T}D_j^1}}{(1+(1-4\alpha_t)\theta_{r,U\tilde{T}D_j^1})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{U\tilde{T}D_j^1}+(2\alpha_t+(4\alpha_t-1)\theta_{l,U\tilde{T}D_j^1})r_2^{U\tilde{T}D_j^1}}{(1+(1-4\alpha_t)\theta_{l,U\tilde{T}D_j^1})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{U\tilde{T}D_j^1}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,U\tilde{T}D_j^1})r_2^{U\tilde{T}D_j^1}}{(1+(3-4\alpha_t)\theta_{l,U\tilde{T}D_j^1})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,U\tilde{T}D_j^1})r_3^{U\tilde{T}D_j^1}+2(1-\alpha_t)r_2^{U\tilde{T}D_j^1}}{(1+(4\alpha_t-3)\theta_{r,U\tilde{T}D_j^1})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{\tilde{t}^2_{jkk_2q}} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,\tilde{t}^2_{jkk_2q}})r_1^{\tilde{t}^2_{jkk_2q}}+2\alpha_t r_2^{\tilde{t}^2_{jkk_2q}}}{(1+(1-4\alpha_t)\theta_{r,\tilde{t}^2_{jkk_2q}})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{\tilde{t}^2_{jkk_2q}}+(2\alpha_t+(4\alpha_t-1)\theta_{l,\tilde{t}^2_{jkk_2q}})r_2^{\tilde{t}^2_{jkk_2q}}}{(1+(1-4\alpha_t)\theta_{l,\tilde{t}^2_{jkk_2q}})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{\tilde{t}^2_{jkk_2q}}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,\tilde{t}^2_{jkk_2q}})r_2^{\tilde{t}^2_{jkk_2q}}}{(1+(3-4\alpha_t)\theta_{l,\tilde{t}^2_{jkk_2q}})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,\tilde{t}^2_{jkk_2q}})r_3^{\tilde{t}^2_{jkk_2q}}+2(1-\alpha_t)r_2^{\tilde{t}^2_{jkk_2q}}}{(1+(4\alpha_t-3)\theta_{r,\tilde{t}^2_{jkk_2q}})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{L\tilde{T}O_j^2} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,L\tilde{T}O_j^2})r_1^{L\tilde{T}O_j^2}+2\alpha_t r_2^{L\tilde{T}O_j^2}}{(1+(1-4\alpha_t)\theta_{r,L\tilde{T}O_j^2})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{L\tilde{T}O_j^2}+(2\alpha_t+(4\alpha_t-1)\theta_{l,L\tilde{T}O_j^2})r_2^{L\tilde{T}O_j^2}}{(1+(1-4\alpha_t)\theta_{l,L\tilde{T}O_j^2})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{L\tilde{T}O_j^2}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,L\tilde{T}O_j^2})r_2^{L\tilde{T}O_j^2}}{(1+(3-4\alpha_t)\theta_{l,L\tilde{T}O_j^2})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,L\tilde{T}O_j^2})r_3^{L\tilde{T}O_j^2}+2(1-\alpha_t)r_2^{L\tilde{T}O_j^2}}{(1+(4\alpha_t-3)\theta_{r,L\tilde{T}O_j^2})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{U\tilde{T}D_k^2} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,U\tilde{T}D_k^2})r_1^{U\tilde{T}D_k^2}+2\alpha_t r_2^{U\tilde{T}D_k^2}}{(1+(1-4\alpha_t)\theta_{r,U\tilde{T}D_k^2})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{U\tilde{T}D_k^2}+(2\alpha_t+(4\alpha_t-1)\theta_{l,U\tilde{T}D_k^2})r_2^{U\tilde{T}D_k^2}}{(1+(1-4\alpha_t)\theta_{l,U\tilde{T}D_k^2})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{U\tilde{T}D_k^2}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,U\tilde{T}D_k^2})r_2^{U\tilde{T}D_k^2}}{(1+(3-4\alpha_t)\theta_{l,U\tilde{T}D_k^2})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,U\tilde{T}D_k^2})r_3^{U\tilde{T}D_k^2}+2(1-\alpha_t)r_2^{U\tilde{T}D_k^2}}{(1+(4\alpha_t-3)\theta_{r,U\tilde{T}D_k^2})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{\tilde{l}^n_{lmk_nq}} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,\tilde{l}^n_{lmk_nq}})r_1^{\tilde{l}^n_{lmk_nq}}+2\alpha_t r_2^{\tilde{l}^n_{lmk_nq}}}{(1+(1-4\alpha_t)\theta_{r,\tilde{l}^n_{lmk_nq}})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{\tilde{l}^n_{lmk_nq}}+(2\alpha_t+(4\alpha_t-1)\theta_{l,\tilde{l}^n_{lmk_nq}})r_2^{\tilde{l}^n_{lmk_nq}}}{(1+(1-4\alpha_t)\theta_{l,\tilde{l}^n_{lmk_nq}})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{\tilde{l}^n_{lmk_nq}}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,\tilde{l}^n_{lmk_nq}})r_2^{\tilde{l}^n_{lmk_nq}}}{(1+(3-4\alpha_t)\theta_{l,\tilde{l}^n_{lmk_nq}})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,\tilde{l}^n_{lmk_nq}})r_3^{\tilde{l}^n_{lmk_nq}}+2(1-\alpha_t)r_2^{\tilde{l}^n_{lmk_nq}}}{(1+(4\alpha_t-3)\theta_{r,\tilde{l}^n_{lmk_nq}})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{L\tilde{T}O_l^n} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,L\tilde{T}O_l^n})r_1^{L\tilde{T}O_l^n}+2\alpha_t r_2^{L\tilde{T}O_l^n}}{(1+(1-4\alpha_t)\theta_{r,L\tilde{T}O_l^n})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{L\tilde{T}O_l^n}+(2\alpha_t+(4\alpha_t-1)\theta_{l,L\tilde{T}O_l^n})r_2^{L\tilde{T}O_l^n}}{(1+(1-4\alpha_t)\theta_{l,L\tilde{T}O_l^n})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{L\tilde{T}O_l^n}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,L\tilde{T}O_l^n})r_2^{L\tilde{T}O_l^n}}{(1+(3-4\alpha_t)\theta_{l,L\tilde{T}O_l^n})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,L\tilde{T}O_l^n})r_3^{L\tilde{T}O_l^n}+2(1-\alpha_t)r_2^{L\tilde{T}O_l^n}}{(1+(4\alpha_t-3)\theta_{r,L\tilde{T}O_l^n})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{U\tilde{T}D_m^n} = \begin{cases} \frac{(1-2\alpha_t+(1-4\alpha_t)\theta_{r,L\tilde{T}D_m^n})r_1^{L\tilde{T}D_m^n}+2\alpha_t r_2^{L\tilde{T}D_m^n}}{(1+(1-4\alpha_t)\theta_{r,L\tilde{T}D_m^n})}, & \text{if } 0 < \alpha_t \leq 0.25 \\ \frac{(1-2\alpha_t)r_1^{L\tilde{T}D_m^n}+(2\alpha_t+(4\alpha_t-1)\theta_{l,L\tilde{T}D_m^n})r_2^{L\tilde{T}D_m^n}}{(1+(1-4\alpha_t)\theta_{l,L\tilde{T}D_m^n})}, & \text{if } 0.25 < \alpha_t \leq 0.50 \\ \frac{(2\alpha_t-1)r_3^{L\tilde{T}D_m^n}+(2(1-\alpha_t)+(3-4\alpha_t)\theta_{l,L\tilde{T}D_m^n})r_2^{L\tilde{T}D_m^n}}{(1+(3-4\alpha_t)\theta_{l,L\tilde{T}D_m^n})}, & \text{if } 0.50 < \alpha_t \leq 0.75 \\ \frac{(2\alpha_t-1+(4\alpha_t-3)\theta_{r,L\tilde{T}D_m^n})r_3^{L\tilde{T}D_m^n}+2(1-\alpha_t)r_2^{L\tilde{T}D_m^n}}{(1+(4\alpha_t-3)\theta_{r,L\tilde{T}D_m^n})}, & \text{if } 0.75 < \alpha_t \leq 1 \end{cases}$$

$$S_{\tilde{a}^1_{iq}} = \begin{cases} \frac{(1-2\alpha_{avail.}+(1-4\alpha_{avail.})\theta_{r,\tilde{a}^1_{iq}})r_1^{\tilde{a}^1_{iq}}+2\alpha_{avail.} r_2^{\tilde{a}^1_{iq}}}{(1+(1-4\alpha_{avail.})\theta_{r,\tilde{a}^1_{iq}})}, & \text{if } 0 < \alpha_{avail.} \leq 0.25 \\ \frac{(1-2\alpha_{avail.})r_1^{\tilde{a}^1_{iq}}+(2\alpha_{avail.}+(4\alpha_{avail.}-1)\theta_{l,\tilde{a}^1_{iq}})r_2^{\tilde{a}^1_{iq}}}{(1+(1-4\alpha_{avail.})\theta_{l,\tilde{a}^1_{iq}})}, & \text{if } 0.25 < \alpha_{avail.} \leq 0.50 \\ \frac{(2\alpha_{avail.}-1)r_3^{\tilde{a}^1_{iq}}+(2(1-\alpha_{avail.})+(3-4\alpha_{avail.})\theta_{l,\tilde{a}^1_{iq}})r_2^{\tilde{a}^1_{iq}}}{(1+(3-4\alpha_{avail.})\theta_{l,\tilde{a}^1_{iq}})}, & \text{if } 0.50 < \alpha_{avail.} \leq 0.75 \\ \frac{(2\alpha_{avail.}-1+(4\alpha_{avail.}-3)\theta_{r,\tilde{a}^1_{iq}})r_3^{\tilde{a}^1_{iq}}+2(1-\alpha_{avail.})r_2^{\tilde{a}^1_{iq}}}{(1+(4\alpha_{avail.}-3)\theta_{r,\tilde{a}^1_{iq}})}, & \text{if } 0.75 < \alpha_{avail.} \leq 1 \end{cases}$$

$$S_{\tilde{b}^1_{jq}} = \begin{cases} \frac{(1-2\alpha_{demand}+(1-4\alpha_{demand})\theta_{r,\tilde{b}^1_{jq}})r_1^{\tilde{b}^1_{jq}}+2\alpha_{demand}r_2^{\tilde{b}^1_{jq}}}{(1+(1-4\alpha_{demand})\theta_{r,\tilde{b}^1_{jq}})}, & \text{if } 0 < \alpha_{demand} \leq 0.25 \\ \frac{(1-2\alpha_{demand})r_1^{\tilde{b}^1_{jq}}+(2\alpha_{demand}+(4\alpha_{demand}-1)\theta_{l,\tilde{b}^1_{jq}})r_2^{\tilde{b}^1_{jq}}}{(1+(1-4\alpha_{demand})\theta_{l,\tilde{b}^1_{jq}})}, & \text{if } 0.25 < \alpha_{demand} \leq 0.50 \\ \frac{(2\alpha_{demand}-1)r_3^{\tilde{b}^1_{jq}}+(2(1-\alpha_{demand})+(3-4\alpha_{demand})\theta_{l,\tilde{b}^1_{jq}})r_2^{\tilde{b}^1_{jq}}}{(1+(3-4\alpha_{demand})\theta_{l,\tilde{b}^1_{jq}})}, & \text{if } 0.50 < \alpha_{demand} \leq 0.75 \\ \frac{(2\alpha_{demand}-1+(4\alpha_{demand}-3)\theta_{r,\tilde{b}^1_{jq}})r_3^{\tilde{b}^1_{jq}}+2(1-\alpha_{demand})r_2^{\tilde{b}^1_{jq}}}{(1+(4\alpha_{demand}-3)\theta_{r,\tilde{b}^1_{jq}})}, & \text{if } 0.75 < \alpha_{demand} \leq 1 \end{cases}$$

$$S_{\tilde{e}^1_{k_1}} = \begin{cases} \frac{(1-2\alpha_{con.cap.}+(1-4\alpha_{con.cap.})\theta_{r,tildee^1_{k_1}})r_1^{tildee^1_{k_1}}+2\alpha_{con.cap.}r_2^{tildee^1_{k_1}}}{(1+(1-4\alpha_{con.cap.})\theta_{r,tildee^1_{k_1}})}, & \text{if } 0 < \alpha_{con.cap.} \leq 0.25 \\ \frac{(1-2\alpha_{con.cap.})r_1^{tildee^1_{k_1}}+(2\alpha_{con.cap.}+(4\alpha_{con.cap.}-1)\theta_{l,tildee^1_{k_1}})r_2^{tildee^1_{k_1}}}{(1+(1-4\alpha_{con.cap.})\theta_{l,tildee^1_{k_1}})}, & \text{if } 0.25 < \alpha_{con.cap.} \leq 0.50 \\ \frac{(2\alpha_{con.cap.}-1)r_3^{tildee^1_{k_1}}+(2(1-\alpha_{con.cap.})+(3-4\alpha_{con.cap.})\theta_{l,tildee^1_{k_1}})r_2^{tildee^1_{k_1}}}{(1+(3-4\alpha_{con.cap.})\theta_{l,tildee^1_{k_1}})}, & \text{if } 0.50 < \alpha_{con.cap.} \leq 0.75 \\ \frac{(2\alpha_{con.cap.}-1+(4\alpha_{con.cap.}-3)\theta_{r,tildee^1_{k_1}})r_3^{tildee^1_{k_1}}+2(1-\alpha_{con.cap.})r_2^{tildee^1_{k_1}}}{(1+(4\alpha_{con.cap.}-3)\theta_{r,tildee^1_{k_1}})}, & \text{if } 0.75 < \alpha_{con.cap.} \leq 1 \end{cases}$$

$$S_{\tilde{w}_q} = \begin{cases} \frac{(1-2\alpha_{weight}+(1-4\alpha_{weight})\theta_{r,\tilde{w}_q})r_1^{\tilde{w}_q}+2\alpha_{weight}r_2^{\tilde{w}_q}}{(1+(1-4\alpha_{weight})\theta_{r,\tilde{w}_q})}, & \text{if } 0 < \alpha_{weight} \leq 0.25 \\ \frac{(1-2\alpha_{weight})r_1^{\tilde{w}_q}+(2\alpha_{weight}+(4\alpha_{weight}-1)\theta_{l,\tilde{w}_q})r_2^{\tilde{w}_q}}{(1+(1-4\alpha_{weight})\theta_{l,\tilde{w}_q})}, & \text{if } 0.25 < \alpha_{weight} \leq 0.50 \\ \frac{(2\alpha_{weight}-1)r_3^{\tilde{w}_q}+(2(1-\alpha_{weight})+(3-4\alpha_{weight})\theta_{l,\tilde{w}_q})r_2^{\tilde{w}_q}}{(1+(3-4\alpha_{weight})\theta_{l,\tilde{w}_q})}, & \text{if } 0.50 < \alpha_{weight} \leq 0.75 \\ \frac{(2\alpha_{weight}-1+(4\alpha_{weight}-3)\theta_{r,\tilde{w}_q})r_3^{\tilde{w}_q}+2(1-\alpha_{weight})r_2^{\tilde{w}_q}}{(1+(4\alpha_{weight}-3)\theta_{r,\tilde{w}_q})}, & \text{if } 0.75 < \alpha_{weight} \leq 1 \end{cases}$$

$$S_{\tilde{W}^1_{k_1}} = \begin{cases} \frac{(1-2\alpha_{weight}+(1-4\alpha_{weight})\theta_{r,\tilde{W}^1_{k_1}})r_1^{\tilde{W}^1_{k_1}}+2\alpha_{weight}r_2^{\tilde{W}^1_{k_1}}}{(1+(1-4\alpha_{weight})\theta_{r,\tilde{W}^1_{k_1}})}, & \text{if } 0 < \alpha_{weight} \leq 0.25 \\ \frac{(1-2\alpha_{weight})r_1^{\tilde{W}^1_{k_1}}+(2\alpha_{weight}+(4\alpha_{weight}-1)\theta_{l,\tilde{W}^1_{k_1}})r_2^{\tilde{W}^1_{k_1}}}{(1+(1-4\alpha_{weight})\theta_{l,\tilde{W}^1_{k_1}})}, & \text{if } 0.25 < \alpha_{weight} \leq 0.50 \\ \frac{(2\alpha_{weight}-1)r_3^{\tilde{W}^1_{k_1}}+(2(1-\alpha_{weight})+(3-4\alpha_{weight})\theta_{l,\tilde{W}^1_{k_1}})r_2^{\tilde{W}^1_{k_1}}}{(1+(3-4\alpha_{weight})\theta_{l,\tilde{W}^1_{k_1}})}, & \text{if } 0.50 < \alpha_{weight} \leq 0.75 \\ \frac{(2\alpha_{weight}-1+(4\alpha_{weight}-3)\theta_{r,\tilde{W}^1_{k_1}})r_3^{\tilde{W}^1_{k_1}}+2(1-\alpha_{weight})r_2^{\tilde{W}^1_{k_1}}}{(1+(4\alpha_{weight}-3)\theta_{r,\tilde{W}^1_{k_1}})}, & \text{if } 0.75 < \alpha_{weight} \leq 1 \end{cases}$$

$$S_{\tilde{v}_q} = \begin{cases} \frac{(1-2\alpha_{volume}+(1-4\alpha_{volume})\theta_{r,\tilde{v}_q})r_1^{\tilde{v}_q}+2\alpha_{volume}r_2^{\tilde{v}_q}}{(1+(1-4\alpha_{volume})\theta_{r,\tilde{v}_q})}, & \text{if } 0 < \alpha_{volume} \leq 0.25 \\ \frac{(1-2\alpha_{volume})r_1^{\tilde{v}_q}+(2\alpha_{volume}+(4\alpha_{volume}-1)\theta_{l,\tilde{v}_q})r_2^{\tilde{v}_q}}{(1+(1-4\alpha_{volume})\theta_{l,\tilde{v}_q})}, & \text{if } 0.25 < \alpha_{volume} \leq 0.50 \\ \frac{(2\alpha_{volume}-1)r_3^{\tilde{v}_q}+(2(1-\alpha_{volume})+(3-4\alpha_{volume})\theta_{l,\tilde{v}_q})r_2^{\tilde{v}_q}}{(1+(3-4\alpha_{volume})\theta_{l,\tilde{v}_q})}, & \text{if } 0.50 < \alpha_{volume} \leq 0.75 \\ \frac{(2\alpha_{volume}-1+(4\alpha_{volume}-3)\theta_{r,\tilde{v}_q})r_3^{\tilde{v}_q}+2(1-\alpha_{volume})r_2^{\tilde{v}_q}}{(1+(4\alpha_{volume}-3)\theta_{r,\tilde{v}_q})}, & \text{if } 0.75 < \alpha_{volume} \leq 1 \end{cases}$$

$$S_{\tilde{V}^1_{k_1}} = \begin{cases} \frac{(1-2\alpha_{volume}+(1-4\alpha_{volume})\theta_{r,\tilde{V}^1_{k_1}})r_1^{\tilde{V}^1_{k_1}}+2\alpha_{volume}r_2^{\tilde{V}^1_{k_1}}}{(1+(1-4\alpha_{volume})\theta_{r,\tilde{V}^1_{k_1}})}, & \text{if } 0<\alpha_{volume}\leq 0.25 \\ \frac{(1-2\alpha_{volume})r_1^{\tilde{V}^1_{k_1}}+(2\alpha_{volume}+(4\alpha_{volume}-1)\theta_{l,\tilde{V}^1_{k_1}})r_2^{\tilde{V}^1_{k_1}}}{(1+(1-4\alpha_{volume})\theta_{l,\tilde{V}^1_{k_1}})}, & \text{if } 0.25<\alpha_{volume}\leq 0.50 \\ \frac{(2\alpha_{volume}-1)r_3^{\tilde{V}^1_{k_1}}+(2(1-\alpha_{volume})+(3-4\alpha_{volume})\theta_{l,\tilde{V}^1_{k_1}})r_2^{\tilde{V}^1_{k_1}}}{(1+(3-4\alpha_{volume})\theta_{l,\tilde{V}^1_{k_1}})}, & \text{if } 0.50<\alpha_{volume}\leq 0.75 \\ \frac{(2\alpha_{volume}-1+(4\alpha_{volume}-3)\theta_{r,\tilde{V}^1_{k_1}})r_3^{\tilde{V}^1_{k_1}}+2(1-\alpha_{volume})r_2^{\tilde{V}^1_{k_1}}}{(1+(4\alpha_{volume}-3)\theta_{r,\tilde{V}^1_{k_1}})}, & \text{if } 0.75<\alpha_{volume}\leq 1 \end{cases}$$

$$S_{\tilde{b}^2_{kq}} = \begin{cases} \frac{(1-2\alpha_{demand}+(1-4\alpha_{demand})\theta_{r,\tilde{b}^2_{kq}})r_1^{\tilde{b}^2_{kq}}+2\alpha_{demand}r_2^{\tilde{b}^2_{kq}}}{(1+(1-4\alpha_{demand})\theta_{r,\tilde{b}^2_{kq}})}, & \text{if } 0<\alpha_{demand}\leq 0.25 \\ \frac{(1-2\alpha_{demand})r_1^{\tilde{b}^2_{kq}}+(2\alpha_{demand}+(4\alpha_{demand}-1)\theta_{l,\tilde{b}^2_{kq}})r_2^{\tilde{b}^2_{kq}}}{(1+(1-4\alpha_{demand})\theta_{l,\tilde{b}^2_{kq}})}, & \text{if } 0.25<\alpha_{demand}\leq 0.50 \\ \frac{(2\alpha_{demand}-1)r_3^{\tilde{b}^2_{kq}}+(2(1-\alpha_{demand})+(3-4\alpha_{demand})\theta_{l,\tilde{b}^2_{kq}})r_2^{\tilde{b}^2_{kq}}}{(1+(3-4\alpha_{demand})\theta_{l,\tilde{b}^2_{kq}})}, & \text{if } 0.50<\alpha_{demand}\leq 0.75 \\ \frac{(2\alpha_{demand}-1+(4\alpha_{demand}-3)\theta_{r,\tilde{b}^2_{kq}})r_3^{\tilde{b}^2_{kq}}+2(1-\alpha_{demand})r_2^{\tilde{b}^2_{kq}}}{(1+(4\alpha_{demand}-3)\theta_{r,\tilde{b}^2_{kq}})}, & \text{if } 0.75<\alpha_{demand}\leq 1 \end{cases}$$

$$S_{\tilde{e}^2_{k_2}} = \begin{cases} \frac{(1-2\alpha_{con.cap.}+(1-4\alpha_{con.cap.})\theta_{r,\tilde{e}^2_{k_2}})r_1^{\tilde{e}^2_{k_2}}+2\alpha_{con.cap.}r_2^{\tilde{e}^2_{k_2}}}{(1+(1-4\alpha_{con.cap.})\theta_{r,\tilde{e}^2_{k_2}})}, & \text{if } 0<\alpha_{con.cap.}\leq 0.25 \\ \frac{(1-2\alpha_{con.cap.})r_1^{\tilde{e}^2_{k_2}}+(2\alpha_{con.cap.}+(4\alpha_{con.cap.}-1)\theta_{l,\tilde{e}^2_{k_2}})r_2^{\tilde{e}^2_{k_2}}}{(1+(1-4\alpha_{con.cap.})\theta_{l,\tilde{e}^2_{k_2}})}, & \text{if } 0.25<\alpha_{con.cap.}\leq 0.50 \\ \frac{(2\alpha_{con.cap.}-1)r_3^{\tilde{e}^2_{k_2}}+(2(1-\alpha_{con.cap.})+(3-4\alpha_{con.cap.})\theta_{l,\tilde{e}^2_{k_2}})r_2^{\tilde{e}^2_{k_2}}}{(1+(3-4\alpha_{con.cap.})\theta_{l,\tilde{e}^2_{k_2}\tilde{e}^2_{k_2}})}, & \text{if } 0.50<\alpha_{con.cap.}\leq 0.75 \\ \frac{(2\alpha_{con.cap.}-1+(4\alpha_{con.cap.}-3)\theta_{r,\tilde{e}^2_{k_2}})r_3^{\tilde{e}^2_{k_2}}+2(1-\alpha_{con.cap.})r_2^{\tilde{e}^2_{k_2}}}{(1+(4\alpha_{con.cap.}-3)\theta_{r,\tilde{e}^2_{k_2}})}, & \text{if } 0.75<\alpha_{con.cap.}\leq 1 \end{cases}$$

$$S_{\tilde{W}^2_{k_2}} = \begin{cases} \frac{(1-2\alpha_{weight}+(1-4\alpha_{weight})\theta_{r,\tilde{W}^2_{k_2}})r_1^{\tilde{W}^2_{k_2}}+2\alpha_{weight}r_2^{\tilde{W}^2_{k_2}}}{(1+(1-4\alpha_{weight})\theta_{r,\tilde{W}^2_{k_2}})}, & \text{if } 0<\alpha_{avail.}\leq 0.25 \\ \frac{(1-2\alpha_{weight})r_1^{\tilde{W}^2_{k_2}}+(2\alpha_{weight}+(4\alpha_{weight}-1)\theta_{l,\tilde{W}^2_{k_2}})r_2^{\tilde{W}^2_{k_2}}}{(1+(1-4\alpha_{weight})\theta_{l,\tilde{W}^2_{k_2}})}, & \text{if } 0.25<\alpha_{avail.}\leq 0.50 \\ \frac{(2\alpha_{weight}-1)r_3^{\tilde{W}^2_{k_2}}+(2(1-\alpha_{weight})+(3-4\alpha_{avail.})\theta_{l,\tilde{W}^2_{k_2}})r_2^{\tilde{W}^2_{k_2}}}{(1+(3-4\alpha_{weight})\theta_{l,\tilde{W}^2_{k_2}})}, & \text{if } 0.50<\alpha_{avail.}\leq 0.75 \\ \frac{(2\alpha_{weight}-1+(4\alpha_{weight}-3)\theta_{r,\tilde{W}^2_{k_2}})r_3^{\tilde{W}^2_{k_2}}+2(1-\alpha_{weight})r_2^{\tilde{W}^2_{k_2}}}{(1+(4\alpha_{weight}-3)\theta_{r,\tilde{W}^2_{k_2}})}, & \text{if } 0.75<\alpha_{avail.}\leq 1 \end{cases}$$

$$S_{\tilde{V}^2_{k_2}} = \begin{cases} \frac{(1-2\alpha_{volume}+(1-4\alpha_{volume})\theta_{r,\tilde{V}^2_{k_2}})r_1^{\tilde{V}^2_{k_2}}+2\alpha_{volume}r_2^{\tilde{V}^2_{k_2}}}{(1+(1-4\alpha_{volume})\theta_{r,\tilde{V}^2_{k_2}})}, & \text{if } 0 < \alpha_{volume} \leq 0.25 \\ \frac{(1-2\alpha_{volume})r_1^{\tilde{V}^2_{k_2}}+(2\alpha_{volume}+(4\alpha_{volume}-1)\theta_{l,\tilde{V}^2_{k_2}})r_2^{\tilde{V}^2_{k_2}}}{(1+(1-4\alpha_{volume})\theta_{l,\tilde{V}^2_{k_2}})}, & \text{if } 0.25 < \alpha_{volume} \leq 0.50 \\ \frac{(2\alpha_{volume}-1)r_3^{\tilde{V}^2_{k_2}}+(2(1-\alpha_{volume})+(3-4\alpha_{volume})\theta_{l,\tilde{V}^2_{k_2}})r_2^{\tilde{V}^2_{k_2}}}{(1+(3-4\alpha_{volume})\theta_{l,\tilde{V}^2_{k_2}})}, & \text{if } 0.50 < \alpha_{volume} \leq 0.75 \\ \frac{(2\alpha_{volume}-1+(4\alpha_{volume}-3)\theta_{r,\tilde{V}^2_{k_2}})r_3^{\tilde{V}^2_{k_2}}+2(1-\alpha_{volume})r_2^{\tilde{V}^2_{k_2}}}{(1+(4\alpha_{volume}-3)\theta_{r,\tilde{V}^2_{k_2}})}, & \text{if } 0.75 < \alpha_{volume} \leq 1 \end{cases}$$

$$S_{\tilde{b}^n_{mq}} = \begin{cases} \frac{(1-2\alpha_{demand}+(1-4\alpha_{demand})\theta_{r,\tilde{b}^n_{mq}})r_1^{\tilde{b}^n_{mq}}+2\alpha_{demand}r_2^{\tilde{b}^n_{mq}}}{(1+(1-4\alpha_{demand})\theta_{r,\tilde{b}^n_{mq}})}, & \text{if } 0 < \alpha_{demand} \leq 0.25 \\ \frac{(1-2\alpha_{demand})r_1^{\tilde{b}^n_{mq}}+(2\alpha_{demand}+(4\alpha_{demand}-1)\theta_{l,\tilde{b}^n_{mq}})r_2^{\tilde{b}^n_{mq}}}{(1+(1-4\alpha_{demand})\theta_{l,\tilde{b}^n_{mq}})}, & \text{if } 0.25 < \alpha_{demand} \leq 0.50 \\ \frac{(2\alpha_{demand}-1)r_3^{\tilde{b}^n_{mq}}+(2(1-\alpha_{demand})+(3-4\alpha_{demand})\theta_{l,\tilde{b}^n_{mq}})r_2^{\tilde{b}^n_{mq}}}{(1+(3-4\alpha_{demand})\theta_{l,\tilde{b}^n_{mq}})}, & \text{if } 0.50 < \alpha_{demand} \leq 0.75 \\ \frac{(2\alpha_{demand}-1+(4\alpha_{demand}-3)\theta_{r,\tilde{b}^n_{mq}})r_3^{\tilde{b}^n_{mq}}+2(1-\alpha_{demand})r_2^{\tilde{b}^n_{mq}}}{(1+(4\alpha_{demand}-3)\theta_{r,\tilde{b}^n_{mq}})}, & \text{if } 0.75 < \alpha_{demand} \leq 1 \end{cases}$$

$$S_{\tilde{e}^n_{k_n}} = \begin{cases} \frac{(1-2\alpha_{con.cap.}+(1-4\alpha_{con.cap.})\theta_{r,\tilde{e}^n_{k_n}})r_1^{\tilde{e}^n_{k_n}}+2\alpha_{con.cap.}r_2^{\tilde{e}^n_{k_n}}}{(1+(1-4\alpha_{con.cap.})\theta_{r,\tilde{e}^n_{k_n}})}, & \text{if } 0 < \alpha_{con.cap.} \leq 0.25 \\ \frac{(1-2\alpha_{con.cap.})r_1^{\tilde{e}^n_{k_n}}+(2\alpha_{con.cap.}+(4\alpha_{con.cap.}-1)\theta_{l,\tilde{e}^n_{k_n}})r_2^{\tilde{e}^n_{k_n}}}{(1+(1-4\alpha_{con.cap.})\theta_{l,\tilde{e}^n_{k_n}})}, & \text{if } 0.25 < \alpha_{con.cap.} \leq 0.50 \\ \frac{(2\alpha_{con.cap.}-1)r_3^{\tilde{e}^n_{k_n}}+(2(1-\alpha_{con.cap.})+(3-4\alpha_{con.cap.})\theta_{l,\tilde{e}^n_{k_n}})r_2^{\tilde{e}^n_{k_n}}}{(1+(3-4\alpha_{con.cap.})\theta_{l,\tilde{e}^n_{k_n}})}, & \text{if } 0.50 < \alpha_{con.cap.} \leq 0.75 \\ \frac{(2\alpha_{con.cap.}-1+(4\alpha_{con.cap.}-3)\theta_{r,\tilde{e}^n_{k_n}})r_3^{\tilde{e}^n_{k_n}}+2(1-\alpha_{con.cap.})r_2^{\tilde{e}^n_{k_n}}}{(1+(4\alpha_{con.cap.}-3)\theta_{r,\tilde{e}^n_{k_n}})}, & \text{if } 0.75 < \alpha_{con.cap.} \leq 1 \end{cases}$$

$$S_{\tilde{W}^n_{k_n}} = \begin{cases} \frac{(1-2\alpha_{weight}+(1-4\alpha_{weight})\theta_{r,\tilde{W}^n_{k_n}})r_1^{\tilde{W}^n_{k_n}}+2\alpha_{weight}r_2^{\tilde{W}^n_{k_n}}}{(1+(1-4\alpha_{weight})\theta_{r,\tilde{W}^n_{k_n}})}, & \text{if } 0 < \alpha_{weight} \leq 0.25 \\ \frac{(1-2\alpha_{weight})r_1^{\tilde{W}^n_{k_n}}+(2\alpha_{weight}+(4\alpha_{weight}-1)\theta_{l,\tilde{W}^n_{k_n}})r_2^{\tilde{W}^n_{k_n}}}{(1+(1-4\alpha_{weight})\theta_{l,\tilde{W}^n_{k_n}})}, & \text{if } 0.25 < \alpha_{weight} \leq 0.50 \\ \frac{(2\alpha_{weight}-1)r_3^{\tilde{W}^n_{k_n}}+(2(1-\alpha_{weight})+(3-4\alpha_{weight})\theta_{l,\tilde{W}^n_{k_n}})r_2^{\tilde{W}^n_{k_n}}}{(1+(3-4\alpha_{weight})\theta_{l,\tilde{W}^n_{k_n}})}, & \text{if } 0.50 < \alpha_{weight} \leq 0.75 \\ \frac{(2\alpha_{weight}-1+(4\alpha_{weight}-3)\theta_{r,\tilde{W}^n_{k_n}})r_3^{\tilde{W}^n_{k_n}}+2(1-\alpha_{weight})r_2^{\tilde{W}^n_{k_n}}}{(1+(4\alpha_{weight}-3)\theta_{r,\tilde{W}^n_{k_n}})}, & \text{if } 0.75 < \alpha_{weight} \leq 1 \end{cases}$$

$$S_{\tilde{V}^n_{k_n}} = \begin{cases} \frac{(1-2\alpha_{volume}+(1-4\alpha_{volume})\theta_{r,\tilde{V}^n_{k_n}})r_1^{\tilde{V}^n_{k_n}}+2\alpha_{volume}r_2^{\tilde{V}^n_{k_n}}}{(1+(1-4\alpha_{volume})\theta_{r,\tilde{V}^n_{k_n}})}, & \text{if } 0 < \alpha_{volume} \leq 0.25 \\ \frac{(1-2\alpha_{volume})r_1^{\tilde{V}^n_{k_n}}+(2\alpha_{volume}+(4\alpha_{volume}-1)\theta_{l,\tilde{V}^n_{k_n}})r_2^{\tilde{V}^n_{k_n}}}{(1+(1-4\alpha_{volume})\theta_{l,\tilde{V}^n_{k_n}})}, & \text{if } 0.25 < \alpha_{volume} \leq 0.50 \\ \frac{(2\alpha_{volume}-1)r_3^{\tilde{V}^n_{k_n}}+(2(1-\alpha_{volume})+(3-4\alpha_{volume})\theta_{l,\tilde{V}^n_{k_n}})r_2^{\tilde{V}^n_{k_n}}}{(1+(3-4\alpha_{volume})\theta_{l,\tilde{V}^n_{k_n}})}, & \text{if } 0.50 < \alpha_{volume} \leq 0.75 \\ \frac{(2\alpha_{volume}-1+(4\alpha_{volume}-3)\theta_{r,\tilde{V}^n_{k_n}})r_3^{\tilde{V}^n_{k_n}}+2(1-\alpha_{volume})r_2^{\tilde{V}^n_{k_n}}}{(1+(4\alpha_{volume}-3)\theta_{r,\tilde{V}^n_{k_n}})}, & \text{if } 0.75 < \alpha_{volume} \leq 1 \end{cases}$$

7. Numerical Example

A firm produces two types of food as Bread and Biscuit and stored at two plants which are the supply points of our problem. The goods are delivered to two destination centers (DCs) from these supply points then finally these products are transported to the final destination centers or customers or survived peoples on disaster via the first DCs. That is the transportation happened in two stages. Due to disaster, the requirements, availabilities and other transportation parameters are not known to us precisely. For this reason, we consider all the transportation parameters as type-2 triangular fuzzy numbers. The type-2 triangular fuzzy inputs for unit transportation costs and times, availabilities, demands, conveyances capacities, purchasing cost, loading and unloading cost and time, weights and volumes etc. for stage-1 and stage-2 are as follows:

Type-2 fuzzy unit transportation cost, time for stage-1 and stage-2:

$C^1_{1111} = (11, 12, 14; .4, .6)$, $C^1_{1211} = (12, 13, 14; .2, .3)$, $C^1_{1121} = (11, 13, 14; .2, .3)$, $C^1_{1221} = (12, 14, 16; .1, .2)$,
$C^1_{2111} = (13, 15, 16; .6, .7)$, $C^1_{2211} = (4, 5, 6; .3, .5)$, $C^1_{2121} = (13, 15, 16; .3, 1.2)$, $C^1_{2221} = (14, 16, 17; .2, .5)$,
$C^1_{1112} = (13, 14, 17; .4, .6)$,$C^1_{1212} = (13, 14, 16; .2, .8)$,$C^1_{1122} = (11, 15, 18; .2, .7)$,$C^1_{1222} = (12, 14, 16; .5, 1.2)$,
$C^1_{2112} = (11, 15, 17; .6, .9)$,$C^1_{2212} = (12, 13, 19; .3, .5)$,$C^1_{2122} = (13, 17, 19; .3, .9)$,$C^1_{2222} = (15, 16, 18; .4, .5)$,
$t^1_{1111} = (2, 3, 5; .4, .6)$,$t^1_{1211} = (3, 4, 7; .7, .9)$,$t^1_{1121} = (7, 9, 12; .9, 1)$,$t^1_{1221} = (2, 4, 6; .1, .2)$,
$t^1_{2111} = (7, 10, 13; .6, .9)$,$t^1_{2211} = (5, 7, 8; .8, 1)$,$t^1_{2121} = (4, 5, 8; .8, 1.3)$,$t^1_{2221} = (6, 7, 9; .7, 1.5)$,
$t^1_{1112} = (5, 9, 13; .2, .8)$,$t^1_{1212} = (5, 7, 9; .8, 1.4)$,$t^1_{1122} = (5, 8, 9; .9, 1.9)$,$t^1_{1222} = (4, 5, 6; .5, .7)$,
$t^1_{2112} = (4, 6, 9; .4, .7)$,$t^1_{2212} = (2, 3, 9; .3, .5)$,$t^1_{2122} = (3, 7, 10; .3, .9)$,$t^1_{2222} = (5, 6, 8; .4, 1.5)$,
$C^2_{1111} = (8, 9, 11; .4, .6)$,$C^2_{1211} = (12, 13, 14; .2, 1)$,$C^2_{1121} = (11, 13, 14; .2, 1.3)$,$C^2_{1221} = (13, 15, 16; .1, .7)$,
$C^2_{2111} = (13, 15, 16; .6, 1.9)$,$C^2_{2211} = (14, 15, 16; .3, .5)$,$C^2_{2121} = (13, 15, 16; .3, 1.2)$,$C^2_{2221} = (14, 16, 17; .6, 1.5)$,
$C^2_{1112} = (13, 14, 17; .4, .6)$,$C^2_{1212} = (14, 15, 17; .2, 1)$, $C^2_{1122} = (11, 15, 18; .2, .7)$,$C^2_{1222} = (3, 4, 7; .7, .9)$,
$C^2_{2112} = (11, 15, 17; .6, .9)$,$C^2_{2212} = (12, 17, 19; .3, .5)$,$C^2_{2122} = (16, 17, 19; .3, .9)$,$C^2_{2222} = (5, 7, 8; .8, 1)$,
$t^2_{1111} = (2, 4, 8; .6, .7)$,$t^2_{1211} = (3, 5, 7; .4, .8)$,$t^2_{1121} = (2, 3, 4; .6, 1.7)$,$t^2_{1221} = (2, 4, 6; .1, .2)$,
$t^2_{2111} = (7, 8, 11; .2, 1.1)$,$t^2_{2211} = (3, 7, 9; .8, 1.1)$,$t^2_{2121} = (2, 5, 9; .8, 1.9)$,$t^2_{2221} = (2, 6, 8; .7, 1.5)$,
$t^2_{1112} = (3, 9, 12; .3, .8)$,$t^2_{1212} = (2, 3, 4; .5, 1.1)$,$t^2_{1122} = (3, 4, 11; .7, 1.4)$,$t^2_{1222} = (4, 5, 6; .7, 1.3)$,
$t^2_{2112} = (3, 6, 8; .3, .4)$, $t^2_{2212} = (2, 6, 9; .2, .8)$, $t^2_{2122} = (7, 9, 13; .2, .7)$, $t^2_{2222} = (2, 6, 9; .9, 1.2)$.

Table-2: Type-2 fuzzy availabilities, demands, conveyances capacities, loading and unloading time and cost, weights and volume for stage-1 and stage-2

Availabilities	$\tilde{a}^1_{11} = (60, 66, 67; .2, .5)$,$\tilde{a}^1_{21} = (54, 56, 60; .1, .2)$,$\tilde{a}^1_{12} = (42, 47, 55; .2, .4)$,$\tilde{a}^1_{22} = (47, 53, 55; .5, .6)$
Demands for stage-1	$\tilde{b}^1_{11} = (19, 26, 30; .1, .3)$,$\tilde{b}^1_{21} = (21, 24, 25; .2, .3)$,$\tilde{b}^1_{12} = (20, 21, 22; .7, 2.1)$,$\tilde{b}^1_{22} = (22, 23, 25; .5, 1.2)$
Demands for stage-2	$\tilde{b}^2_{11} = (18, 20, 23; .2, .3)$,$\tilde{b}^2_{21} = (17, 18, 25; .1, .3)$,$\tilde{b}^2_{12} = (15, 16, 17; .1, .4)$,$\tilde{b}^2_{22} = (12, 14, 16; .9, 1.3)$
Conveyances Capacities	$\tilde{e}^1_1 = (52, 54, 56; .2, .3)$,$\tilde{e}^1_2 = (53, 55, 57; .6, .9)$,$\tilde{e}^2_1 = (42, 44, 49; 1.8, 2.3)$,$\tilde{e}^2_2 = (45, 49, 50; .4, .9)$
Loading Cost	$\tilde{LO}^1_1 = (2, 4, 6; .2, .3)$,$\tilde{LO}^1_2 = (5, 6, 7; .2, .3)$,$\tilde{LO}^2_1 = (5, 6, 9; .2, .6)$,$\tilde{LO}^2_2 = (2, 9, 10; .3, .4)$
Unloading Cost	$\tilde{UD}^1_1 = (2, 3, 4; .4, .6)$,$\tilde{UD}^1_2 = (3, 7, 8, .6; .7,)$,$\tilde{UD}^2_1 = (2, 3, 5; .4, .9)$,$\tilde{UD}^2_2 = (6, 8, 9; .6, 1)$
Loading Time	$L\tilde{T}O^1_1 = (3, 7, 9; 1.2, 1.3)$,$L\tilde{T}O^1_2 = (1, 2, 3; 1.1, 1.2)$,$L\tilde{T}O^2_1 = (4, 6, 8; .2, .3)$,$L\tilde{T}O^2_2 = (1, 3, 6; .1, .2)$
Unloading Time	$U\tilde{T}D^1_1 = (3, 7, 10; .3, .8)$,$U\tilde{T}D^1_2 = (2, 9, 11; .1, .4)$,$U\tilde{T}D^2_1 = (2, 4, 5; .1, .5)$,$U\tilde{T}D^2_2 = (4, 5, 6; .5, .6)$
Purchasing Cost	$\tilde{PC}_{11} = (10, 11, 12; 1.2, 1.4)$,$\tilde{PC}_{12} = (13, 15, 16; 1.1, 1.4)$,$\tilde{PC}_{21} = (11, 12, 13; 1.1, 1.2)$,$\tilde{PC}_{22} = (14, 16, 17; .9, 1.3)$
Weights capacity	$\tilde{W}^1_1 = (185, 190, 225; .6, .7)$,$\tilde{W}^1_2 = (288, 320, 400; .7, .8)$, $\tilde{W}^2_1 = (310, 320, 331; .1, .9)$,$\tilde{W}^2_2 = (368, 340, 345; .6, .7)$
Volumes capacity	$\tilde{V}^1_1 = (334, 360, 370; 1.2, 1.3)$,$\tilde{V}^1_2 = (231, 294, 370; .6, .7)$, $\tilde{V}^2_1 = (393, 395, 399; .6, .8)$,$\tilde{V}^2_2 = (353, 354, 356; .8, .9)$
weight,volume of items	$\tilde{w}_1 = (1, 5, 7; .4, .5)$,$\tilde{w}_2 = (1, 4, 5; .7, .9)$,$\tilde{v}_1 = (1, 2, 6; 1.1, 1.3)$,$\tilde{v}_2 = (1, 3, 8; .2, .5)$

8. Results

The fuzzy multi-stage STP is converted to its equivalent crisp problem by using CV-based reduction method and generalized credibility measure. Then using LINGO.13.0 optimization software, we obtain the optimal solution of the deterministic STP. The values of the credibility level for the transportation parameters are sometime lies in the interval (0, 0.25] or (0.25, 0.5] or (0.5, 0.75] or (0.75, 1]. For this reason, we obtain the optimal solution of the newly developed model with the four limitations of the credibility level. A sensitivity analysis is taken into consideration to show the change of the optimal values of the objective functions and the transported amounts with respect to the credibility level of availabilities, demands and conveyances capacities.

Table-3: Changes of optimum cost and transported amount for different credibility levels

Credibility Level	Item-1	Item-2	Item-1	Item-2	Stage-1	Stage-2	Opt. cost	Opt. time
$\alpha_c = 0.07$ $\alpha_t = 0.10$ $\alpha_{avail.} = 0.13$ $\alpha_{demand} = 0.16$ $\alpha_{con.cap.} = 0.19$ $\alpha_{weight} = 0.22$ $\alpha_{volume} = 0.25$	41.92	42.40	35.87	27.71	$x^1_{1111} = 21.02$ $x^1_{2211} = 21.90$ $x^1_{1122} = 20.20$ $x^1_{1222} = 22.20$	$x^2_{1111} = 18.58$ $x^2_{2221} = 17.29$ $x^2_{2112} = 15.28$ $x^2_{2212} = 6.88$ $x^2_{1222} = 5.49$ $x^2_{2222} = 0.06$	3482.99	105.86
$\alpha_c = 0.26$ $\alpha_t = 0.29$ $\alpha_{avail.} = 0.31$ $\alpha_{demand} = 0.33$ $\alpha_{con.cap.} = 0.35$ $\alpha_{weight} = 0.37$ $\alpha_{volume} = 0.40$	51.40	64.08	40.84	41.44	$x^1_{1111} = 13.81$, $x^1_{2211} = 26.19$, $x^1_{1121} = 11.45$, $x^1_{2212} = 14.30$, $x^1_{1122} = 32.68$, $x^1_{1222} = 17.10$,	$x^2_{1111} = 22$, $x^2_{1211} = 3.25$, $x^2_{2211} = 15.59$, $x^2_{2112} = 14.7$, $x^2_{1122} = 2.04$, $x^2_{1222} = 24.7$	5730.85	188.26
$\alpha_c = 0.56$ $\alpha_t = 0.59$ $\alpha_{avail.} = 0.61$ $\alpha_{demand} = 0.63$ $\alpha_{con.cap.} = 0.65$ $\alpha_{weight} = 0.67$ $\alpha_{volume} = 0.69$	40.45	44.62	40.45	30.60	$x^1_{2211} = 24.58$, $x^1_{1121} = 15.87$, $x^1_{1112} = 14.56$, $x^1_{1122} = 6.64$, $x^1_{1222} = 23.42$	$x^2_{1111} = 15.87$, $x^2_{2121} = 4.84$, $x^2_{2221} = 19.74$, $x^2_{1112} = 4.43$, $x^2_{2112} = 9.48$, $x^2_{1212} = 14.36$, $x^2_{1122} = 2.33$.	4912.23	186.70
$\alpha_c = 0.76$ $\alpha_t = 0.79$ $\alpha_{avail.} = 0.83$ $\alpha_{demand} = 0.86$ $\alpha_{con.cap.} = 0.90$ $\alpha_{weight} = 0.95$ $\alpha_{volume} = 0.98$	53.76	46.48	45.53	32.41	$x^1_{2211} = 4.04$, $x^1_{1121} = 29.01$, $x^1_{1221} = 20.71$, $x^1_{1112} = 21.85$, $x^1_{1212} = 17.62$, $x^1_{1222} = 7.01$	$x^2_{1111} = 22.26$ $x^2_{1211} = 6.75$, $x^2_{2211} = 12.57$, $x^2_{2221} = 3.95$, $x^2_{1212} = 6.24$, $x^2_{2122} = 16.76$, $x^2_{1222} = 9.41$	6055.26	284.63

8.1. Particular Case. Let us consider, the credibility level for costs, times, availabilities, demands, conveyances capacities, weights and volume are all equal. i.e., $\alpha_c = \alpha_t = \alpha_{avail.} = \alpha_{demand} = \alpha_{con.cap.} = \alpha_{weight} = \alpha_{volume} = \alpha$, say.

Table-4: Optimal results of the model with same credibility level

Credibility Level	Item-1	Item-2	Item-1	Item-2	Stage-1	Stage-2	Opt. cost	Opt. time
$\alpha = 0.24$	44.74	42.90	36.43	28.38	$x^1_{1111} = 22.32$, $x^1_{2211} = 22.42$, $x^1_{1122} = 20.44$, $x^1_{1222} = 22.46$,	$x^2_{1111} = 18.95$ $x^2_{1111} = 3.37$, $x^2_{1111} = 5.09$, $x^2_{2221} = 9.02$, $x^2_{2112} = 15.47$, $x^2_{1222} = 12.91$	4022.51	115.49
$\alpha = 0.32$	50.60	60.86	40.27	39.02	$x^1_{2211} = 25.70$, $x^1_{1121} = 24.90$, $x^1_{1212} = 19.88$, $x^1_{2212} = 10.19$, $x^1_{1122} = 30.79$,	$x^2_{1111} = 21.61$, $x^2_{1211} = 3.29$, $x^2_{2211} = 15.37$, $x^2_{1212} = 14.24$, $x^2_{2212} = 8.23$, $x^2_{1122} = 16.55$.	7588.72	210.61
$\alpha = 0.72$	42.33	45.24	42.33	31.22	$x^1_{2211} = 25.84$, $x^1_{1121} = 16.49$, $x^1_{1112} = 5.91$, $x^1_{1122} = 15.5$, $x^1_{1222} = 23.83$,	$x^2_{1111} = 16.49$, $x^2_{2111} = 4.8$, $x^2_{2211} = 14.93$, $x^2_{2221} = 6.11$, $x^2_{1112} = 6.61$, $x^2_{2122} = 9.82$, $x^2_{1222} = 14.79$.	5317.32	215.64
$\alpha = 0.80$	53.11	46.07	44.23	31.99	$x^1_{2211} = 8.02$, $x^1_{1121} = 28.49$, $x^1_{1221} = 16.6$, $x^1_{1112} = 21.72$, $x^1_{1212} = 13.12$, $x^1_{1222} = 11.23$,	$x^2_{1111} = 21.87$, $x^2_{1211} = 3.67$, $x^2_{2211} = 18.69$, $x^2_{1212} = 3.4$, $x^2_{2122} = 16.63$, $x^2_{1222} = 11.96$.	6037.55	268.83

8.2. Sensitivity Analysis of the availabilities and demands of the model. We know that the sensitivity analysis is used to analyze the outputs with the given inputs data. For this reason, in Table-5 and -6 we analyze some inputs data and outputs as sensitivity analysis. Basically the minimization of cost objective and time objective in the STP depend on the values of the transportation parameters such as unit transportation costs, times, demands, supplies etc.

Table-5: Sensitivity analysis on availabilities

α_c	α_t	$\alpha_{avail.}$	α_{demand}	$\alpha_{con.cap.}$	α_{weight}	α_{volume}	Opt. cost	Opt. time	Item-1	Item-2
		0.10					3574.57	101.08	35.93	27.78
		0.12					3584.04	100.11	35.99	27.84
0.11	0.15	0.16	0.17	0.20	0.22	0.25	3584.13	91.03	36.00	27.85
		0.19					3587.07	99.96	36.02	27.87
		0.25					3601.32	109.94	36.06	27.92
		0.26					5541.09	201.52	40.27	39.02
		0.29					5600.76	189.59	40.27	39.02
0.26	0.29	0.32	0.32	0.35	0.38	0.40	5556.28	211.09	40.27	39.02
		0.35					5544.85	197.75	40.27	39.02
		0.38					5508.48	163.44	40.27	39.02
		0.53					4821.18	216.28	39.85	30.45
		0.58					4811.73	180.94	40.24	30.56
0.52	0.56	0.63	0.60	0.64	0.68	0.72	4857.09	185.68	40.85	30.72
		0.68					4905.02	195.65	40.46	30.91
		0.73					4938.51	179.64	41.72	31.07
		0.77					6093.96	256.79	45.73	32.46
		0.83					6139.95	259.16	46.11	32.56
0.77	0.83	0.87	0.87	0.91	0.95	0.99	6148.66	277.31	46.49	32.66
		0.92					6173.94	280.47	46.83	32.95
		0.98					6187.11	280.50	46.93	33.04

Table-6: Sensitivity analysis on demands

α_c	α_t	$\alpha_{avail.}$	α_{demand}	$\alpha_{con.cap.}$	α_{weight}	α_{volume}	Opt. cost	Opt. time	Item-1	Item-2
			0.13				3539.57	101.08	35.68	27.54
			0.16				3565.36	100.18	35.87	27.71
0.11	0.15	0.17	0.19	0.19	0.22	0.25	3594.08	102.09	36.06	27.92
			0.22				3626.59	101.59	36.27	28.18
			0.25				3664.37	100.28	36.50	28.50
			0.26				4534.36	184.54	37	29.69
			0.29				4962.18	151.40	38.6	33.82
0.26	0.29	0.32	0.32	0.35	0.38	0.40	5563.87	186.38	40.27	39.02
			0.35				6193.2	201.54	42.01	45.83
			0.38				7072.83	199.66	43.83	55.23
			0.53				4751.82	157.77	38.54	30.12
			0.58				4798.8	164.77	39.47	30.35
0.52	0.56	0.60	0.63	0.64	0.68	0.72	4833.92	195.21	40.44	30.61
			0.68				4913.07	200.54	41.47	30.92
			0.73				5013.48	184.95	42.55	31.32
			0.77				5921.35	255.83	43.5	31.71
			0.83				6049.3	273.49	44.89	32.22
0.77	0.83	0.87	0.87	0.91	0.95	0.99	6213.72	234.95	45.74	32.46
			0.92				6237.77	266.98	46.67	32.69
			0.98				6271.87	233.34	47.69	32.93

8.3. Pictorial representation of the sensitivity analysis. The Pictorial representation of the sensitivity analysis are shown in the figure-2 - figure-17 and those are given below:

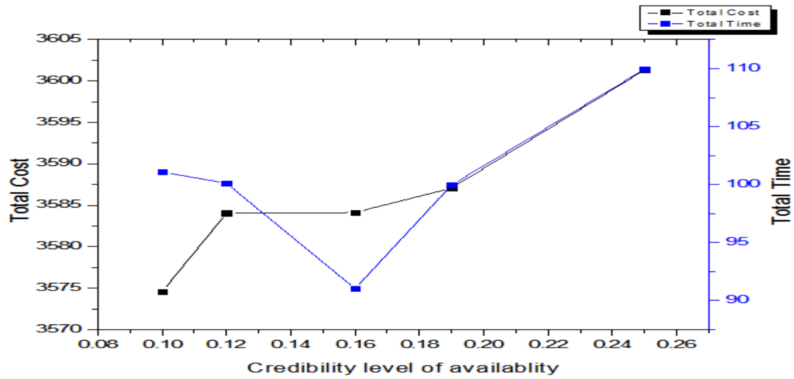


Figure 2. *Change of total optimum cost and time with Credibility level of availability,* $\alpha_{avail.} \in (0, 0.25]$

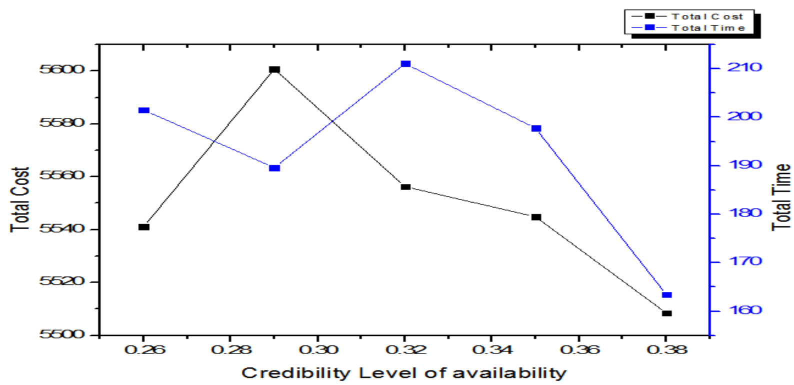


Figure 3. *Change of total optimum cost and time with Credibility level of availability* $\alpha_{avail.} \in (0.25, 0.50]$

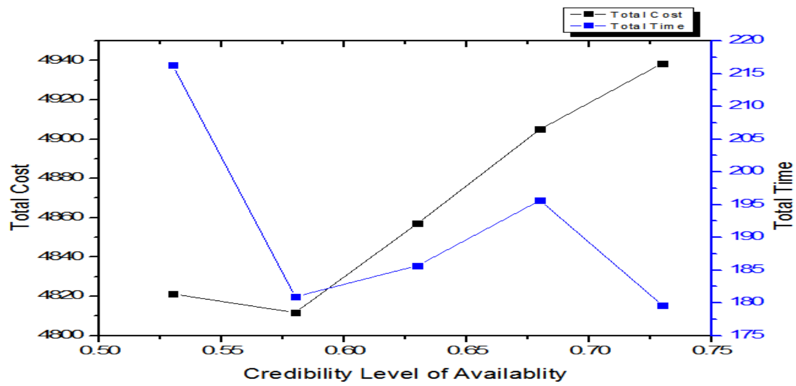


Figure 4. *Change of total optimum cost and time with Credibility level of availability,* $\alpha_{avail.} \in (0.50, 0.75]$

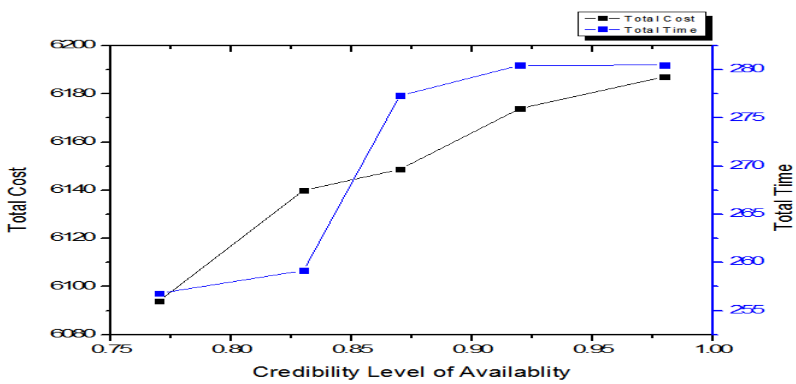


Figure 5. *Change of total optimum cost and time with Credibility level of availability,* $\alpha_{avail.} \in (0.t5, 1]$

9. Discussion

Since the credibility level of the availabilities, demands, conveyances capacities, weights and volumes for each transported item and each vehicle are different, so after taking the variation of each transportation parameters, we obtained lots of results of our STP model where all the transportation parameters are type-2 fuzzy variables and which are discussed below: Following Table-3, we see that the least amount of total cost and time are 3482.99 and 105.86 units respectively and these are obtained when the credibility level of the transportation parameters lies within the interval (0, 0.25]. Again, in Table-4, we put some optimal results which are obtained by taking the credibility level of all the transportation parameters are equal. After careful investigation, we found that the total cost and time are least when the credibility levels are same and it is lies in the interval (0, 0.25]. In Table-5 we have the following:

(i) When credibility level of the availabilities increases with the limit (0, 0.25], then the values of the cost objective also increases and the time objective are sometime increases and decreases. The increase or decrease of the total time within the variation of the credibility level is also significant i.e., if we change the credibility level then the allocations

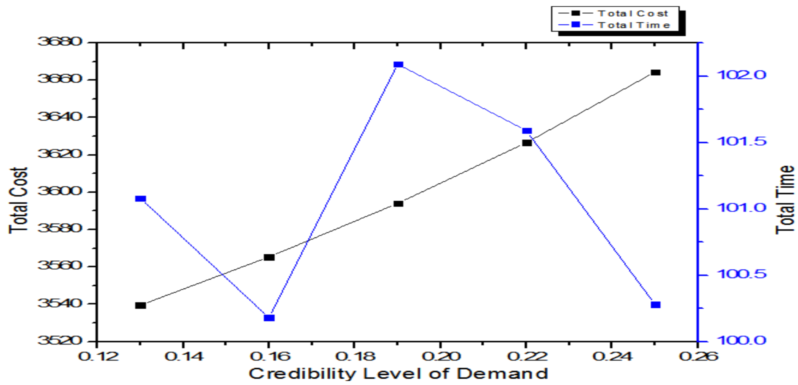


Figure 6. *Change of total optimum cost and time with Credibility level of demand,* $\alpha_{demand} \in (0, 0.25]$

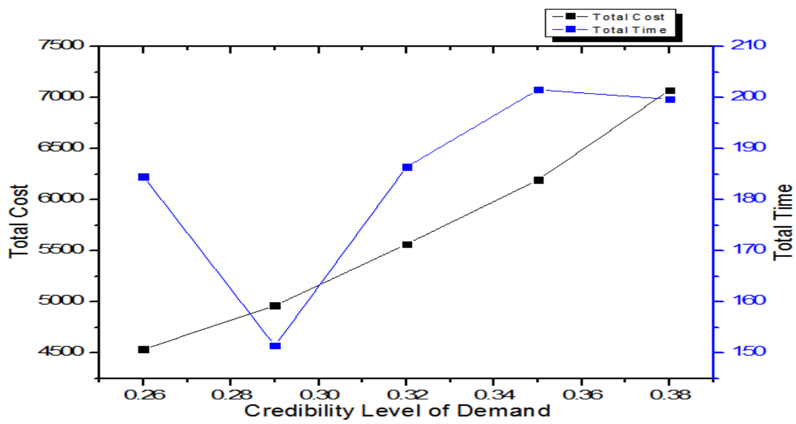


Figure 7. *Change of total optimum cost and time with Credibility level of demand,* $\alpha_{demand} \in (0.25, 0.50]$

are changed and for this reason, it is happening.
(ii) From the third row of the Table-5, we obtained some optimal results of the objectives where the credibility level of availabilities are lies within (0.25, 0.5]. Due to change of credibility level of availabilities, sometimes the value of total cost are increased and sometimes decreased but there is no significant change in time objective. This is found when credibility level increases within the range (0.25, 0.5].
(iii) The value of the cost objective increases when we increase the credibility level of availabilities within the range (0.5, 0.75] but there are some random changes found in the time objective function.
(iv) When credibility level of the availabilities increases within the limit (0.75, 1], then the value of the objectives and transported amounts (item-1 and item-2) are increased.
Again if we can change the value of the credibility level of the demand, then we found some significant changes on the objective functions as well as transported amounts. From Table-6, it is seen that when we increase the credibility level of the demands, then the cost and time objectives are also increases and same type of changes is found on the

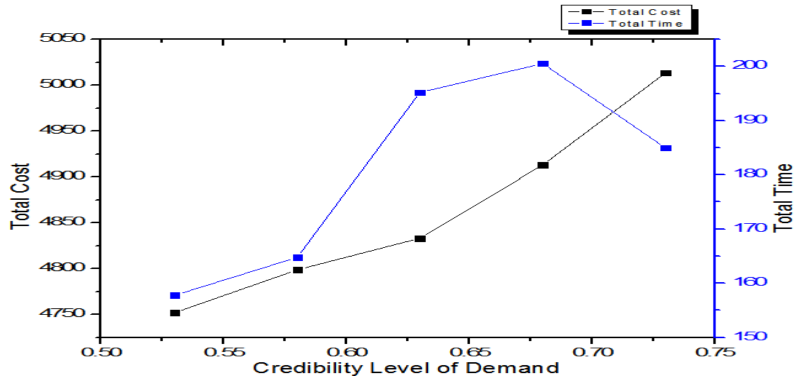


Figure 8. *Change of total optimum cost and time with Credibility level of demand,* $\alpha_{demand} \in (0.50, 0.75]$

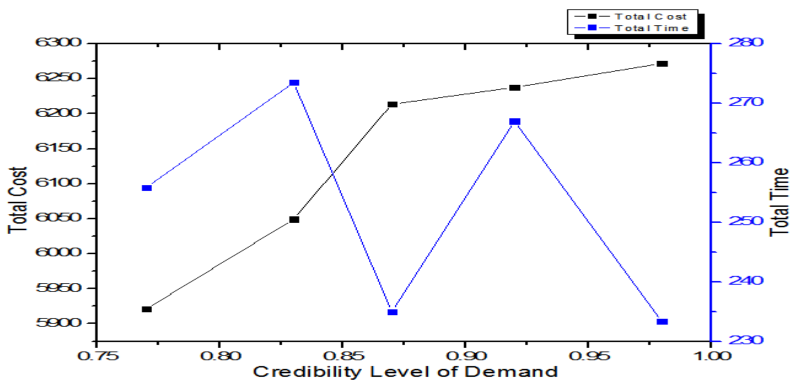


Figure 9. *Change of total optimum cost and time with Credibility level of demand,* $\alpha_{demand} \in (0.75, 1]$

transported amount in the final stages.

10. Comparison with the earlier Research work

Heragu [13] introduced the problem called two stages TP and gave the mathematical model for this problem. The model includes both the inbound and outbound transportation cost and aims to minimize the overall cost. Hindi et al. [12] addressed a two-stage distribution-planning problem. They considered two additional requirements on their problem. First, each customer must be served from a single DC. Second, it must be possible to ascertain the plant origin of each product quantity delivered. A mathematical formulation called PLANWAR presented by Pirkul and Jayaraman [20] to locate a number of sources and destination centers and to design distribution network so that the total operating cost can be minimized. Syarif and Gen [23] considered production/distribution problem formulated as two-stage TP and proposed a hybrid genetic algorithm (GA) for solution. But in our research, we develop a new concept which is totally different from the concept of [20], [13], [12], [23] etc. Here our concept is to supply the commodities from

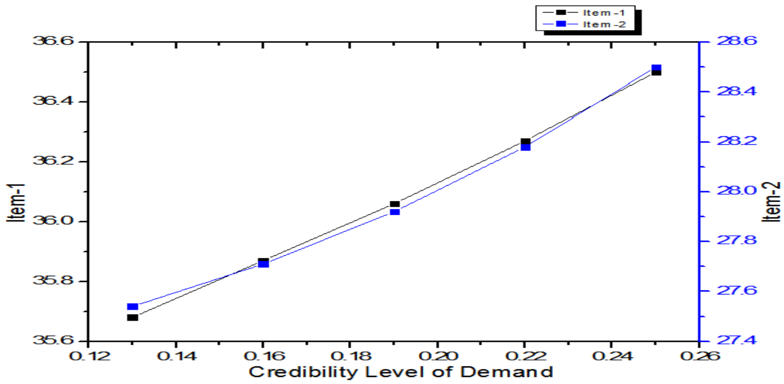


Figure 10. *Change of transported amounts (item-1 and 2) with Credibility level of demand,* $\alpha_{demand} \in (0, 0.25]$

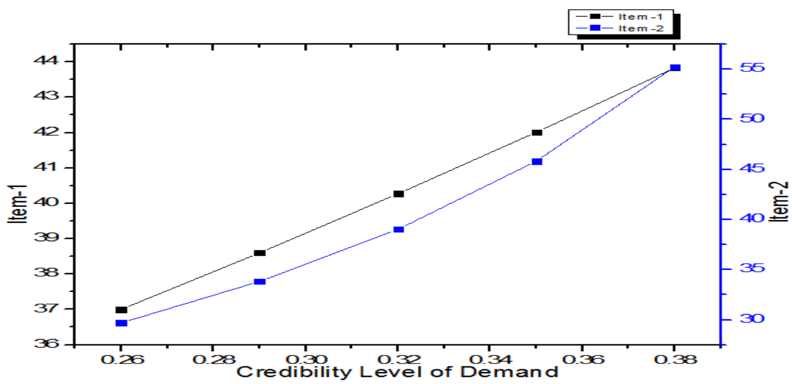


Figure 11. *Change of transported amounts (item-1 and 2) with Credibility level of demand,* $\alpha_{demand} \in (0.25, 0.50]$

sources to destination centers with their requirements in stage-1 and then the transported amounts in stage-1 is converted to the availabilities of the stage-2. The transportation of the stage-2 happened according to requirements of the destination centers of the stage-2 where the availabilities for the stage-2 are the transported amounts for stage-1 and so on for the other stage transportations. So we can't make comparison of our approach to the existing one. But we validate our technique and optimum result by sensitivity analysis.

11. Conclusion and Future Extension of the Research Work

11.1. Conclusion. In this paper, we propose a newly developed STP model under type-2 fuzzy environment. Weight and volume of the transported items and vehicle are more significant in the transportation network. So we add two new additional constraints as weight constraints and volume constraints for each vehicle to handle the STP with different stages. We apply the goal programming method is to solve our multi-objective multi-stage STP since goal programming technique gives the better optimal result of the objective function than the other methods. Here we study four cases of the credibility

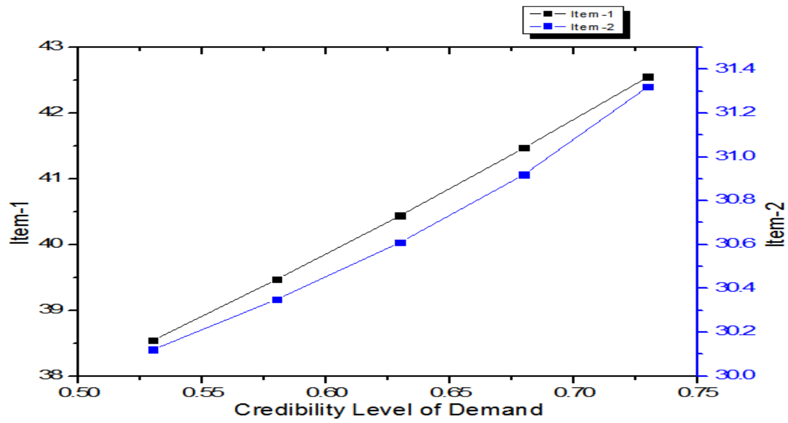


Figure 12. *Change of transported amounts (item-1 and 2) with Credibility level of demand,* $\alpha_{demand} \in (0.50, 0.75]$

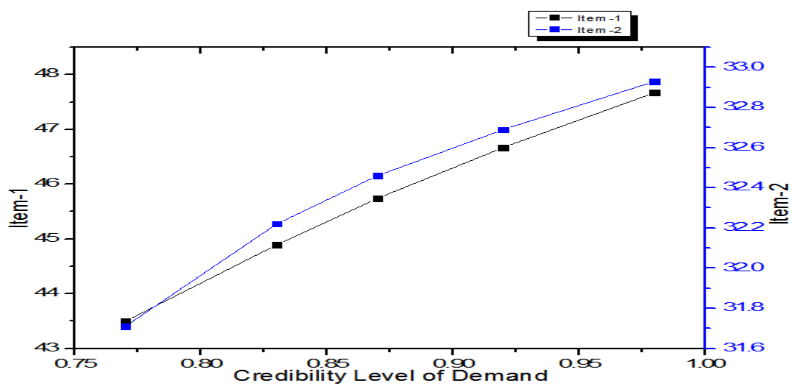


Figure 13. *Change of transported amounts (item-1 and 2) with Credibility level of demand,* $\alpha_{demand} \in (0.75, 1]$

level of the different transportation parameters. Also, after solving the transportation model, we see that the least transportation cost is obtained when the credibility level lies within the range (0, 0.25] and in particular when the credibility level of the transportation parameters are all equal, then a similar type of change is observed in the objective functions. We obtain the optimal solution of the model by using generalized reduced gradient technique (LINGO 13.0 optimization solver) and the results are very effective in real-life sense. So we conclude that, if the credibility levels of the transportation parameters lies within (0, 0.25], then any multi-stage or single stage STP with type-2 fuzzy parameter gives the least value of the objective function.

11.2. Future Extension of the Research Work. The future extensions of our research work are as follows:

• We have formulated the STP model under type-2 fuzzy environment but this model can be developed under fuzzy-rough, fuzzy-random, interval type-2 fuzzy environments

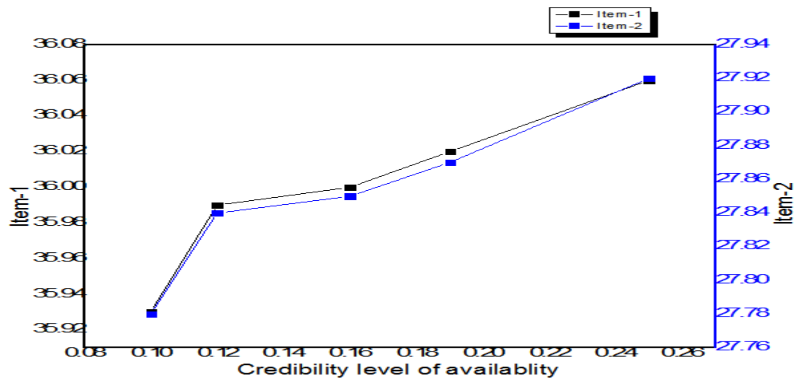


Figure 14. *Change of transported amounts (item-1 and 2) with Credibility level of demand,* $\alpha_{avail.} \in (0, 0.25]$

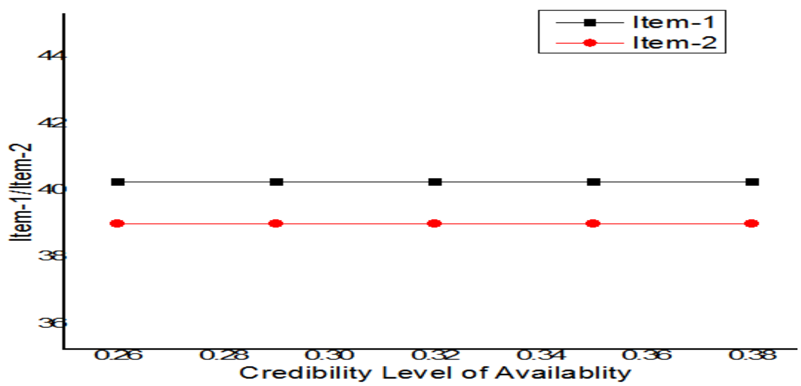


Figure 15. *Change of transported amounts (item-1 and 2) with Credibility level of demand,* $\alpha_{avail.} \in (0.25, 0.50]$

etc.

• In our model we imposed two extra restrictions with the help of weights and volume of each items and vehicles. There is a scope to formulate and solve the model with safety constraints, budget constraint etc.

• In the objective function we considered the unit transportation cost, time, purchasing cost, loading and unloading cost and time etc. but there is a scope to develop the cost objective function of our model with fixed charges, vehicle carrying cost etc.

• In the solution of the imprecise STP model, the transported amounts have been considered as crisp. Hence there is a scope of taking these transported amount as fuzzy also i.e. the models can be formulated as fully fuzzy models.

References

[1] Qin, R., Liu, Y. K. and Liu, Z. Q. "Methods of critical value reduction for type-2 fuzzy variables and their Applications", Journal of Computational and Applied Mathematics, 235, 1454-1481 (2011).

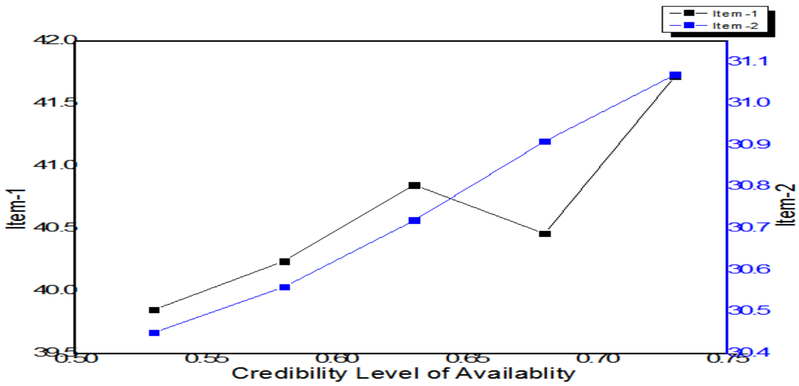


Figure 16. *Change of transported amounts (item-1 and 2) with Credibility level of demand,* $\alpha_{avail.} \in (0.50, 0.75]$

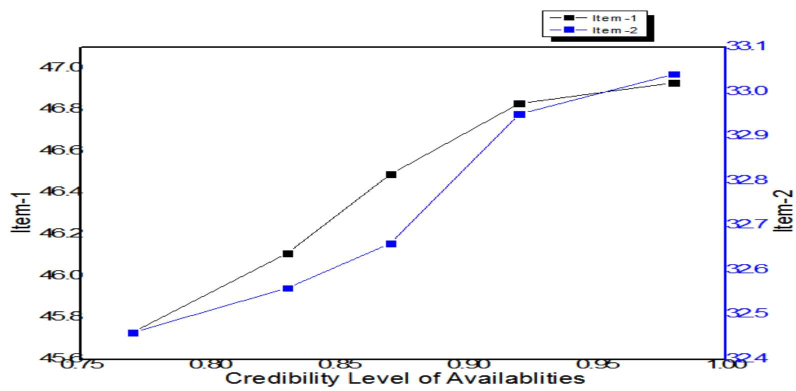


Figure 17. *Change of transported amounts (item-1 and 2) with Credibility level of demand,* $\alpha_{avail.} \in (0.75, 1]$

[2] Karnik, N.N., and Mendel, J.M.,"Centroid of a type-2 fuzzy set", Information Sciences, 132, 195-220 (2001).

[3] Liu, F.,"An efficient centroid type-reduction strategy for general type-2 fuzzy logic system", Information Sciences, 178, 2224-2236(2008).

[4] Qiu, Y., Yang, H., Zhang, Y.-Q. and Zhao, Y.,"Polynomial regression interval-valued fuzzy systems", Soft Computing, 12, 137-145 (2008).

[5] Haley, K.B.,"The solid transportation problem", Operations Research,10, 448-463(1962).

[6] Haley, K.B.,"The multi-index problem", Operations Research,11,368-379(1963).

[7] Hitchcock, F.L.,"The distribution of a product from several sources to numerous localities", Journal of MathematicalPhysics, 20,224-230(1941).

[8] Shell, E.,"Distribution of a product by several properties, in: Directorate of Management Analysis", Proceedings of the Second Symposium in Linear Programming", 2, 615-642, DCS/Comptroller H.Q.U.S.A.F., Washington, DC. (1955).

[9] Zadeh, L.A.,"Concept of a linguistic variable and its application to approximate reasoning I", Information Sciences, 8, 199-249 (1975).

[10] Mitchell, H.,"Ranking type-2 fuzzy numbers", IEEE Transactions on Fuzzy Systems, 14, 327-348 (2006).

[11] Liang, Q. and Mendel, J.M.,"Interval type-2 fuzzy logic systems: theory and design", IEEE Transactions on Fuzzy Systems, 8, 535-549 (2000).

[12] Charnes, A., Cooper, W.W. and Rhodes, E. "Measuring the efficiency of decision making units", European Journal of Operational Research, 6, 429-444 (1978).

[13] Sengupta, J.K.,"Efficiency measurement in stochastic input-output systems", International Journal of System Science, 13, 273-287 (1982).

[14] Banker, R.D.,"Maximum likelihood, consistency and DEA: statistical foundations", Management Science, 39, 1265-1273 (1993).

[15] Cooper, W.W., Huang, Z.M., Li, S.X. and Olesen, O.B., "Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA", Journal of Productivity Analysis, 9, 53-79 (1998).

[16] Land, K.C., Lovell, C.A.K. and Thore, S.,"Chance constrained data envelopment analysis", Managerial and Decision Economics, 14, 541-554 (1993).

[17] Heragu S., "Facilities design", PWS (1997).

[18] Baidya, A., Bera, U. K. and Maiti, M., "Solution of multi-item interval valued solid transportation problem with safety measure using different methods", OPSEARCH, 51(1), 1-22 (2013).

[19] Geoffrion A. M. and Graves G. W., "Multi commodity distribution system design by benders decomposition", Manage Sci, 20, 822-844 (1974).

[20] Pirkul H, Jayaraman V. "A multi-commodity, multi-plant capacitated facility location problem: formulation and efficient heuristic solution", Computer and Operations Research, vol. 25 (10), pp. 869-878(1998).

[21] Hindi K. S., Basta T. and Pienkosz K., "Efficient solution of a multi-commodity, two-stage distribution problem with constraints on assignment of customers to distribution centers", Int Trans Oper Res vol. 5(6), pp. 519-527(1998).

[22] Syarif A. and Gen M. "Double spanning tree-based genetic algorithm for two stage transportation problem", Int J Knowl-Based Intell Eng Syst, 7(4) (2003).

[23] Ali Amiri, "Designing a distribution network in a supply chain system: formulation and efficient solution procedure" Eur J Oper Res, vol. 171, pp. 567-576(2006).

[24] Gen, M., Altiparmak, F. and Lin, L., "A genetic algorithm for two-stage transportation problem using priority-based encoding", OR Spectrum, Vol. 28, pp. 337-354 (2006).

[25] Mahapatra, G. S. and Roy, T. K., "Fuzzy multi-objective mathematical programming on reliability optimization model", Applied Mathematics and Computation, vol. 174, pp. 643-659(2006).

[26] Kundu, P., kar, S. and Maiti, M., "Fixed charge transportation problem with type-2 fuzzy variable", Information Sciences, vol. 255, pp. 170-186 (2014).

[27] Yang, L. and Feng, Y., "A bicriteria solid transportation problem with fixed charges under stochastic environment", Applied Mathematical Modelling, vol.31, pp. 2668-2683 (2007).

[28] Kundu, P., kar, S. and Maiti, M., "Multi-objective multi-item solid transportation problem in fuzzy environment", Applied Mathematical Modelling, vol. 37, pp. 2028-2038 (2012).

[29] Liu, P., Yang, L., Wang, L. and Liu, S., "A solid transportation problem with type-2 fuzzy variables", Applied Soft Computing, vol. 24, pp. 543-558 (2014).

[30] Yang, L., Liu, P., Li, S., Gao, Y. and Ralescu, D.A., "Reduction method of type-2 uncertain variables and their applications to solid transportation problems", Information Sciences, vol. 291, pp. 204-237 (2015).

[31] Giri, P. K., Maiti, M. K. and Maiti, M., "Fully fuzzy fixed charge multi-item solid transportation problem", Applied Soft Computing, vol. 27, pp77-91 (2015).

[32] Yang, L. and Liu, L., "Fuzzy fixed charge solid transportation problem and algorithm", Applied Soft Computing, vol. 7, pp. 879-889 (2007).

[33] Molla-Alizadeh-Zavardehi, S., Sadi Nazhad, S., Tavakkoli-Moghaddam, R. and Yazdani, M., "solving a fuzzy fixed charge solid transportation problem and by metaheuristics", Mathematical and Computer Modelling, vol. 57, pp. 1543-1558 (2013).

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1519 – 1524

Moving towards an optimal sample using VNS algorithm

Priyaranjan Dash*† and Samrat Hore‡

Abstract

In almost all random sampling schemes, we adopt different sampling designs with an objective of obtaining a better representative sample (optimal sample) for the population. Application of different randomization techniques were adopted for providing a supportive basis for this. Now the question arises, whether the final sample selected, on which all our efforts are utilized, from the population is an optimal sample or not? No where we are checking about the *optimality of this sample, i.e.,* whether this sample is the best one or there exists any other sample which is more optimal than the selected one satisfying all the constraints. In all these procedures, we only assume but, nowhere we are establishing a guarantee about the achievement of such a representative sample. The present paper emphasizes on achieving an optimal sample by using *variable neighborhood search (VNS)* technique.

Keywords: Optimal Sample, Representative Sample, Variable Neighborhood Search.

2000 AMS Classification: 62D05

Received : 16.03.2015 *Accepted* : 10.09.2015 *Doi* : 10.15672/HJMS.20158712906

1. Introduction

In any sample survey, we first develop a frame, which emphasizes on specifying the sampled population identical with the target population lacking any kind of ambiguity there on. A sample plays a role of centripetal force in sampling theory literature. An optimum sample is always desirable and fetches attention at all phases because a poor sample ruins the entire effort of the survey whatever attention may be put to other aspects. We put our entire effort in sampling theory to develop methods of sample selection i.e. to get an optimum sample and to draw inferences on the principles of specified precision and minimum cost. In this connection, two rivalry methods of selection

*Department of Statistics, Tripura University, Email: prdashjsp@gmail.com
†Corresponding Author.
‡Department of Statistics, Tripura University, Email: sam.stat724@gmail.com

Table 1. All possible samples selected by SRSWOR scheme

Sample No.	Sample Units ($y_1,\ y_2$)	Sample Mean $\overline{y}$	$(\overline{y}-\overline{Y})^2$	
1	$y_1=5,\ y_2=3$	4	0	(Minimum)
2	$y_1=5,\ y_3=6$	5.5	2.25	
3	$y_1=5,\ y_4=2$	3.5	0.25	
4	$y_1=5,\ y_5=4$	4.5	0.25	
5	$y_2=3,\ y_3=6$	4.5	0.25	
6	$y_2=3,\ y_4=2$	2.5	2.25	
7	$y_2=3,\ y_5=4$	3.5	0.25	
8	$y_3=6,\ y_4=2$	4	0	(Minimum)
9	$y_3=6,\ y_5=4$	5	1	
10	$y_4=2,\ y_5=4$	3	1	

of a sample came into existence: (1) *random selection* and the other one is (2) *purposive (non-random) selection*. Jensen (1926) [10], Gini and Galvani (1929) [3], Neyman (1934) [12] advocated about these methods of selection. But, all of these based on the hope that the sample we get is a representative one. Since, our desire lies on getting an optimum sample (as a proper subset of the target population) whose characteristics $\hat{\Phi}_y = \widetilde{y}(y_1, y_2, \cdots, y_n)$ under study are almost similar with the population characteristic $\Phi_y = \widetilde{Y}(y_1, y_2, \cdots, y_N)$, when we have a sample of size n from the population of size N to infer about the variable y. Unfortunately, an optimum sample does not exist and even if it exists, it is very difficult, even not possible to identify it. In this regard Godambe(1955) [4], Hege (1965) [8], Hanurav (1966) [7] had given significant contributions. The following hypothetical example will clear this idea.

1.1. Example. Consider a finite population with $N = 5$ and $n = 2$. When the population values are known to us (say) $y_1 = 5, y_2 = 3, y_3 = 6, y_4 = 2, y_5 = 4$.So, we can have $\binom{5}{2} = 10$ different possible SRSWOR samples in total. (We are not emphasizing here regarding WR and WOR samples, as both the schemes are indifferent for large samples (Freedman, 1977 [2]). We have to estimate the population mean $\Phi_y = \overline{Y}$. We now calculate the sample means for these as follows. From the Table 1, it indicates that the sample number 1 and 8 are optimum samples and the sample numbers 2 and 6 are poor samples on the basis of the value of the expression $||\hat{\Phi}_y - \Phi_y|| = (\overline{y} - \overline{Y})^2$.
When we use equal probability scheme to select a sample, in that case all the samples are equally likely of being selected. Alternatively, if we use PPSWR sample, sample 2 is more likely to be selected than the others. So, in all the cases, we are not selecting an optimum sample. It emphasis the individual units to be present in the sample for an optimum sample.

The above result encourages to design a sampling scheme, which will guide us at each step of selection of the units for moving towards optimality. However, we have to keep in view about the cost incurred for selecting the sample.

2. An Overview of VNS Algorithm

Mladenović and Hansen (1997) [11] used the variable neighborhood approach for solving the vehicle routing problems. Variable neighborhood search is the systematic change of neighborhood within a possibly randomized local search algorithm yields a simple and effective metaheuristic for combinatorial and global optimization (Hansen and Mladenović, (1999, 2001) [5, 6]). Contrary to the other metaheuristics based on local search methods, VNS does not follow a trajectory but explores increasingly distant neighborhoods of the current solution, and jumps from this solution to a new one, if and only if an improvement has been made. In this way, favorable characteristics of the current solution (e.g., many variables are already at their optimal value), will often be kept and used to obtain promising neighboring solutions. Moreover, a local search routine is applied repeatedly to get from these neighboring solutions to local optima. This kind of VNS algorithm has recently been successfully applied in the field of design of experiments by finding optimum allocation of experimental units with known predictors into two treatment groups (Hore et al., 2014 [9]).

3. An Optimal Sample Using VNS Algorithm

Let x be the auxiliary variable closely related to the study variable y. The corresponding parametric function of interest for x is defined as

$$\Phi_x = \widetilde{X}(x_1, x_2, \cdots, x_N).$$

Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ to be a selected sample of size n over the design $\mathscr{P}$ from N units, gives the sample observation vector for auxiliary variable only. To obtain an optimal sample observation vector over the selected one, an iterative method through neighborhood search algorithm is propsed in this article. Let us denote the neighborhood of α by $\mathcal{N}(\alpha)$ and the neighborhood construction of the corresponding α is derived in Step 3. The proposed algorithm is framed by using the concept of VNS algorithm.
Another sample of size k $(\leq n)$ also has been selected randomly from remaining $(N-n)$ population units gives the sample observation vector

$$\beta = \{\beta_1, \beta_2, \cdots, \beta_k\}$$

with the relation $\beta^{(j)} = \beta^{(j-1)} - \{\beta_j\}$; $\beta^{(0)} = \beta$, $j = 1, 2, \cdots, k$. We repeat the following steps until a *stopping condition* is met.

Step 1: Start with $j = 1$. Select the j^{th} unit from $\beta^{(j-1)}$.

Step 2: Choose the initial sample $\alpha^{(0)}$ and calculate the value of $\hat{\Phi}_x = \Phi_{\alpha^{(0)}}$ and the corresponding value of the objective function

$$V\left(\alpha^{(0)}\right) = ||\Phi_{\alpha^{(0)}} - \Phi_x|| \quad \text{(say)}.$$

Step 3: The neighborhood of $\alpha^{(0)}$ is constructed as

$$\mathcal{N}\left(\alpha^{(0)}\right) = \left\{\alpha^{(0)}_{(i)}, \beta_j : \; i = 1, 2, \cdots, n\right\},$$

where the $\alpha^{(0)}_{(i)}$ is the sample vector of size $(n-1)$ and it is constructed as,

$$\alpha^{(0)}_{(i)} = \alpha^{(0)} - \{\alpha_i\}, \quad i = 1, 2, \cdots, n.$$

Step 4: Consider all the allocations in $\mathcal{N}(\alpha^{(0)})$ and compute $V(\alpha')$, for all $\alpha' \in \mathcal{N}(\alpha^{(0)})$. Find the minimum objective function with respective sample $\alpha^{(0')}$, denoted as

$$\alpha^{(0')} = arg\ min\left\{V(\alpha'), \text{for all } \alpha' \in \mathcal{N}\left(\alpha^{(0)}\right)\right\}$$

Step 5: If $V(\alpha^{(0')}) < V(\alpha^{(0)})$, choose the next improved sample to be $\alpha^{(1)} = \alpha^{(0')}$. Otherwise select $\alpha^{(1)} = \alpha^{(0)}$.

Step 6: *(Moving towards optimality)* Replace $\alpha^{(0)}$ by $\alpha^{(1)}$, j by $j+1$ and start the algorithm again from **Step 1**.

Step 7: *(Stopping Condition)* Continue repeating the above steps until all k units are examined one by one or $\beta^{(k)} = \phi$.

Table 2. Sizes of 15 Large United States Cities (in 1000's) in 1920 (x_i) and 1930 (y_i)

Sl No.	1	2	3	4	5	6	7	8
x_i	76	138	67	29	381	23	37	120
y_i	80	143	67	50	464	48	63	115
Sl No.	9	10	11	12	13	14	15	
x_i	61	387	93	172	78	66	60	
y_i	69	459	104	183	106	86	57	

4. Empirical Illustrations

Table 2 gives the number of inhabitants (in 1000's) of 15 cities of United States in the years 1920 and 1930. Cochran(2011) [1], p. 151-152.

In order to estimate the total number of inhabitants $Y = \sum_{i=1}^{15} y_i$ in these cities in the year 1930, we select an initial sample of 4 cities using SRSWOR scheme. Let the selected cities are 3, 6, 7 and 12. So, we have

$$\alpha^{(0)} = \{\alpha_1 = 67, \alpha_2 = 23, \alpha_3 = 37, \alpha_4 = 172\}.$$

Again we select another sample of size 3 from the remaining $15 - 4 = 11$ cities as 1, 5 and 8. Thus, $\beta = \{\beta_1 = 76, \beta_2 = 381, \beta_3 = 120\}$.

Step 1: Start with $j = 1$. Select the first unit from β as $\beta_1 = 76$.

Step 2: Choose the initial sample $\alpha^{(0)} = \{67, 23, 37, 172\}$ and calculate the value of

$$\begin{aligned} V\left(\alpha^{(0)}\right) &= ||\Phi_{\alpha^{(0)}} - \Phi_x|| = ||\hat{X} - X|| = (N\overline{x} - X)^2 \\ &= (1121.25 - 1788)^2 = 444555.6 \quad \text{(say)}. \end{aligned}$$

Step 3: To find the neighbors of $\alpha^{(0)}$, we consider 4 samples, each of size 3, as

$$\alpha^{(0)}_{(1)} = \{23, 37, 172\},\ \alpha^{(0)}_{(2)} = \{67, 37, 172\},\ \alpha^{(0)}_{(3)} = \{67, 23, 172\},\ \alpha^{(0)}_{(4)} = \{67, 23, 37\}.$$

The neighborhood of $\alpha^{(0)}$ is constructed as

$$\begin{aligned} \mathcal{N}\left(\alpha^{(0)}\right) &= \left\{\left\{\alpha^{(0)}_{(i)}, \beta_j\right\}, \quad i = 1, 2, \cdots, n\right\} \\ &= \{\{23, 37, 172, 76\}, \{67, 37, 172, 76\}, \\ &\qquad \{67, 23, 172, 76\}, \{67, 23, 37, 76\}\}. \end{aligned}$$

Step 4: Here,

$$\begin{aligned} \alpha^{(0')} &= argmin\left\{\left|\left|V\left(\alpha'\right)\right|\right|,\ \forall \alpha^{(0')} \in \mathcal{N}\left(\alpha^{(0)}\right)\right\} \\ &= argmin\left\{(N\overline{x} - X)^2\right\} = \left\{\alpha^{(0)}_{(2)}, \beta_1\right\} = \{67, 37, 172, 76\} \end{aligned}$$

Step 5: Here $V\left(\alpha^{(0')}\right) = 219024 < V\left(\alpha^{(0)}\right)$.

So, the corresponding units $\{3rd,\ 7th,\ 12th,\ 1st\}$ gives a better representation of the population than the initial sample.

Step 6: *(Moving towards optimality)* We replace

$\alpha^{(0)}$ by $\alpha^{(1)} = \{67, 37, 172, 76\}$.

Table 3. Sample values for x and y

Sample units:	u_1	u_3	u_7	u_{12}
y values:	80	67	63	183
x values:	76	67	37	172

Table 4. Estimate of Variance of Different Estimators of Y and their Relative Gain in Efficiency (RGE).

Sample Type	Different Estimators	Sample Type	Est. of Variance	Est. RGE of (a) to (b)
SRSWOR	$\hat{\Phi}_1 = N\overline{y}$	a	178470.4	39.91885
		b	249713.7	
Ratio Estimator	$\hat{\Phi}_2 = \frac{\overline{y}}{\overline{x}}X$	a	11669.76	142.9719
		b	28354.24	
Regression Estimator	$\hat{\Phi}_3 = N\left[\overline{y} + \hat{b}_{yx}\left(\overline{X} - \overline{x}\right)\right]$	a	10256.91	83.44258
		b	18815.54	
Ratio Est. in DS^*	$\hat{\Phi}_4 = \frac{\overline{y}}{\overline{x}}\overline{x}'$	a	80989.51	42.17238
		b	115144.7	
Regression Est. in DS^*	$\hat{\Phi}_5 = N\left[\overline{y} + \hat{b}_{yx}\left(\overline{x}' - \overline{x}\right)\right]$	a	80163.82	37.74642
		b	110422.8	

a. Optimum sample b. Traditional sample *DS: Double Sampling

Step 7: *(Stopping Condition)* Again, proceeding in the previous manner, after two such iterations, we can get $\beta^{(3)} = \phi$ and the corresponding sample units $\{3rd,\ 7th,\ 12th,\ 1st\}$ is the optimum sample as it has the smallest argument.

Here, we get the optimum sample as $s = \{u_3, u_7, u_{12}, u_1\}$. Now, we can only study these units for getting y values. The Table 3 gives the values of x and y for this optimum sample.

If an equivalent two phase sample is selected from this population with $n+k$ units to estimate the unknown population mean of auxiliary variable $\overline{X}$ and a second phase sample of size n units out of $n+k$ units, then in the present example (with $n = 4, k = 3$), observed sample values for x are 67, 23, 37, 172, 76, 381, 120. Table 4 gives the estimated standard errors of different estimators and relative gain in efficiency for estimating population total (Y) in adopting proposed optimal sample to the usual (initial) sample using different estimators under SRSWOR scheme.

5. Conclusion

In all traditional sample survey literature, we are emphasizing on improving the sampling design or the estimators there on by efficiently utilizing the auxiliary information but neglecting the representativeness of the selected sample. The present paper utilizes the readily available auxiliary information in order to get an *improved sample*, viewed by a *better representation of the population*, to estimate the parameters of interest. The proposed procedure provides, by sacrificing a little cost to study the auxiliary variable, a safeguard for arriving at a better representative sample employing *variable neighborhood search (VNS) technique*. It does not require any kind of abstract knowledge about the population values like population correlation coefficient (ρ) between y and x as in case of ratio and regression methods of estimation. The *optimality* of the final selected sample is established by the relative gain in efficiency to the traditional sample, on the basis of a numerical study, shown in Table 4. Therefore, the proposed VNS algorithm for selecting an optimal sample strongly advocates about its better representativeness.

Acknowledgements

The authors are very much thankful to the referees for their valuable timely suggestions for the improvement of the paper.

References

[1] Cochran, W.G. (2011). Sampling Techniques. *Wiley India Pvt. Ltd.*, New Delhi, 3rd edition.

[2] Freedman, D. (1977). *A remark on the difference between sampling with and without replacement.* Journal of the American Statistical Association, **72** (359), 681.

[3] Gini, C. and Galvani, L. (1929). *Di una applicazione del metodo rappresentativo all'ultimo censimento italiano della popolazione* . Annali di Statistica, **6** (4), 1-107.

[4] Godambe, V. P.(1955). *A unified theory of sampling from finite populations.* Journal of the Royal Statistical Society. Series B (Methodological), **17** (2), 269-278.

[5] Hansen, P. and Mladenović, N. (1999). *An introduction to variable neighborhood search in: Metaheuristics, Advances and Trends in Local Search Paradigms for Optimization.* S. Voss et al., eds, Kluwer, Dordrecht.

[6] Hansen, P. and Mladenović, N. (2001). *Variable neighborhood search: Principles and applications.* European Journal of Operational Research, **130** (3), 449-467.

[7] Hanurav, T. V. (1966). *Some aspects of unified sampling theory.* Sankhya, **28**, 175-204.

[8] Hege, V.S. (1965). *Sampling designs which admit uniformly minimum variance unbiased estimators.* Calcutta Statistical Association Bulletin, **14**, 160-162.

[9] Hore, S., Dewanji, A. and Chatterjee, A. (2014) : *Design issues related to allocation of experimental units with known covariates into two treatment groups.* Journal of Statistical Planning and Inference, **155**, 117-126.

[10] Jensen, A. (1926). *Report on representative method in statistics.* Bulletin of the International Statistical Institute, **22** (1), 381-439.

[11] Mladenović, N. and Hansen, P. (1997). *Variable neighborhood search.* Computers & Operations Research, **24**, 1097-1100.

[12] Neyman, J. (1934). *On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection.* Journal of the Royal Statistical Society, **97** (4), 558-625.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1525 – 1540

On the bivariate and multivariate weighted generalized exponential distributions

Indranil Ghosh*[†] and Ayman Alzaatreh[‡]

Abstract

This article proposes a particular member of the weighted bivariate distribution, namely, bivariate weighted generalized exponential distribution. This distribution is obtained via conditioning, starting from three independent generalized exponential distributions with different shape but equal scale parameters. Several structural properties of the proposed bivariate weighted generalized exponential distribution including total positivity of order two, marginal moments, reliability parameter and estimation of the model parameters are studied. A multivariate extension of the proposed model is discussed with some properties. Small simulation experiments have been performed to see the behavior of the maximum likelihood estimators, and one data analysis has been presented for illustrative purposes.

Keywords: Weighted distributions, Total positivity of order 2, Reliability parameter, Multivariate weighted distributions.

2000 AMS Classification: 62E, 62F.

Received : 13.03.2015 *Accepted* : 16.07.2015 *Doi* : 10.15672/HJMS.20157912259

1. Introduction

In recent times, there have been numerous studies on the family of weighted distributions that emerges as a center of attraction in the development of application (see Arellano-Valle and Azzalini (2006) and the references therein). The weighted distribution arises when the density $g(x;\theta_1)$ of the potential observation x gets contaminated so that it is multiplied by some non-negative weight function $w(x;\theta_1,\theta_2)$ involving an additional parameter vector θ_2. Then, the observed data is a random realization from a weighted distribution with density

$$(1.1) \qquad f(x;\theta_1,\theta_2) = \frac{w(x;\theta_1,\theta_2)g(x;\theta_1)}{E\left[w(X;\theta_1,\theta_2)\right]},$$

*University of North Carolina, Wilmington, North Carolina, USA Email: ghoshi@uncw.edu
[†]Corresponding Author.
[‡]Nazarbayev University, Astana, Kazakhstan Email: ayman.alzaatreh@nu.edu.kz

where the expectation in the denominator is just a normalizing constant. An extensive class of weighted distributions are discussed in Rao (1965,1985), Bayarri and DeGroot (1992), Arnold and Beaver (2002), Branco and Dey (2001), Azzalini (1985) and Kim (2005). As elaborated in the articles by Arnold and Nagaraja (1991) as well as in the book by Genton (2004), the application of the weighted distribution extends to the areas of econometrics, astronomy, engineering, medicine as well as psychology. In particular in scenarios where the observed random phenomena can be described by (1.1). Again, if the potential observation x is obtained only from a selected portion of the population of interest, then (1.1) is called a selection model. Weighted distributions, establishing links with selection models obtained from various forms of selection mechanisms are well addressed in the literature; see Genton (2004), Arellano-Valle et al. (2006) and the references therein. The main objective of this study, described here, is to investigate various properties of a class of weighted distributions arising via conditioning where the underlying distributions are independent generalized exponential. Although the class has some resemblance with the selection distributions developed by Arellano-Valle et al. (2006), we are not aware of any detailed exposition of the distributional properties. This lack of detailed exposition motivates the investigation described in this article. This class apart from a theoretical interest, is worthy of investigation from an applied point of view. In the applied view point, the class produce new models that provide us a means to analyze non-normal data such as interval grouped data, screened data and skewed data. We envision a real life scenario as a genesis of the proposed bivariate weighted distribution in a classical stress-strength model context.
Assume a system has two independent components with strengths W_1 and W_2, and suppose that to run the process each component strength has to overcome an outside stress W_0 which is independent of both (W_1 and W_2). If we define $(X, Y) \stackrel{d}{=} ((W_1, W_2)|(\min(W_1, W_2)) > W_0)$ where the W_i's have absolutely continuous distributions, then the resulting joint distribution of (W_1, W_2) is the type of bivariate weighted distribution to be investigated in this paper.

2. The bivariate weighted generalized exponential distribution

Let W_1, W_2 and W_0 be independent random variables with density functions $f_{W_i}(w_i)$, $i = 0, 1, 2$. Define $(X, Y) \stackrel{d}{=} ((W_1, W_2)|W_0 < \min(W_1, W_2))$, then the density function of the corresponding bivariate weighted distribution is given by

$$\begin{aligned} f_{X,Y}(x, y) &= \frac{f_{W_1}(x) f_{W_2}(y) P(W_0 < \min(W_1, W_2)|W_1 = x, W_2 = y)}{P(W_0 < \min(W_1, W_2))} \\ &= \frac{f_{W_1}(x) f_{W_2}(y) F_{W_0}(\min(x, y))}{P(W_0 < \min(W_1, W_2))}. \qquad (2.1) \end{aligned}$$

Indeed the density in (2.1) is a bivariate weighted distribution of (X, Y) with the weight $P(W_0 < \min(W_1, W_2))$. This method was first proposed by Al-Mutairi at el. (2011).

If W_i's for $i = 0, 1, 2$, are identically distributed with common density function $f_W(w)$, then $P(W_0 < \min(W_1, W_2)) = \frac{1}{3}$. Hence, (2.1) reduces to

$$f_{X,Y}(x, y) = 3 f_W(x) f_W(y) F_{W_0}(\min(x, y)). \qquad (2.2)$$

Next, we consider a member of the weighted family in (2.1), the bivariate weighted generalized exponential distribution. The exponentiated exponential distribution (Gupta and Kundu, 2001), known in the literature as the generalized exponential distribution (GED), is a two-parameter right skewed unimodal distribution where the behavior of the density and the hazard functions are quite similar to the density and the hazard functions

of the gamma and Weibull distributions. The generalized exponential distribution can also be used effectively to analyze lifetime data.
Next, if W_i's are independent generalized exponential random variables with parameters (α_i, θ) for $i = 0, 1, 2$. Then the normalizing constant is

$$\begin{aligned} P(W_0 < \min(W_1, W_2)) &= \int_0^\infty \int_0^\infty f_{W_1}(x) f_{W_2}(y) F_{W_0}(\min(x,y)) dxdy \\ &= \int_0^\infty \int_0^\infty \theta^{-2} \alpha_1 e^{-x/\theta} (1 - e^{-x/\theta})^{\alpha_1 - 1} \\ &\quad \times e^{-y/\theta} \alpha_2 (1 - e^{-y/\theta})^{\alpha_2 - 1} (1 - e^{-\min(x,y)/\theta})^{\alpha_0} dxdy \\ &= \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2 + \alpha_0} \left(\frac{1}{\alpha_1 + \alpha_0} + \frac{1}{\alpha_2 + \alpha_0} \right). \end{aligned} \tag{2.3}$$

From (2.1), the density function of the proposed bivariate generalized exponential distribution can be written as

$$\begin{aligned} f_{X,Y}(x,y) &= \theta^{-2} \delta(\alpha_0, \alpha_1, \alpha_2) e^{-(x/\theta + y/\theta)} (1 - e^{-x/\theta})^{\alpha_1 - 1} (1 - e^{-y/\theta})^{\alpha_2 - 1} \\ &\quad \times (1 - e^{-\min(x,y)/\theta})^{\alpha_0} \times I(x > 0, y > 0), \end{aligned} \tag{2.4}$$

where $\delta(\alpha_0, \alpha_1, \alpha_2) = \left\{ \frac{1}{\alpha_1 + \alpha_2 + \alpha_0} \left(\frac{1}{\alpha_1 + \alpha_0} + \frac{1}{\alpha_2 + \alpha_0} \right) \right\}^{-1}$, $\alpha_i > 0$ for $i = 0, 1, 2$ and $\theta > 0$.
A bivariate random variable (X, Y) with the joint p.d.f $f(x, y)$ in (2.4) is said to follow the bivariate weighted generalized exponential distribution with parameters α_0, α_1, α_2 and θ and will be denoted by BWGED$(\alpha_0, \alpha_1, \alpha_2, \theta)$. When $\alpha_0 = \alpha_1 = \alpha_2 = 1$, the BWGED reduces to the bivariate weighted exponential distribution (BWED) with parameters $\lambda_0 = \lambda_1 = \lambda_2 = 1/\theta$ [Al-Mutairi et al., 2011]. Also, when $\alpha_0 \longrightarrow 0$ and $\alpha_1 = \alpha_2 = 1$, the BWGED reduces to the bivariate exponential distribution where X and Y are independent and follow $Exp(\theta)$ distribution.
In Figure 1, various density and contour plots of BWGED density are provided. Figure 1 shows that the joint density function is very flexible in terms of shapes, it can assume various shapes such as strictly decreasing and concave down. The shape of the distribution is strictly decreasing whenever $\alpha_i < 1$, $i = 0, 1, 2$. Also, it appears from the plots that the BWGED density is a unimodal distribution.

The remainder of this paper is organized as follows: In section 3, some properties of the bivariate generalized exponential distribution in (2.4) are discussed. In section 4, some discussion on the multivariate extension of the proposed family is provided. Section 5 deals with the estimation of the bivariate generalized exponential distribution parameters. For illustrative purposes, one data set is studied in section 6. In section 7, some concluding remarks are made regarding the BWGED model.

3. Properties of the bivariate generalized exponential distribution

In this section we discuss various structural properties of the BWGED including moment generating functions, marginal distributions and distributions of the minimum and maximum.

3.1. Moment generating function. The moment generating function of BWGED in (2.4) is

$$M_{X,Y}(t_1, t_2) = E\left(e^{t_1 X + t_2 Y}\right) = I_1 + I_2, \quad \text{say} \tag{3.1}$$

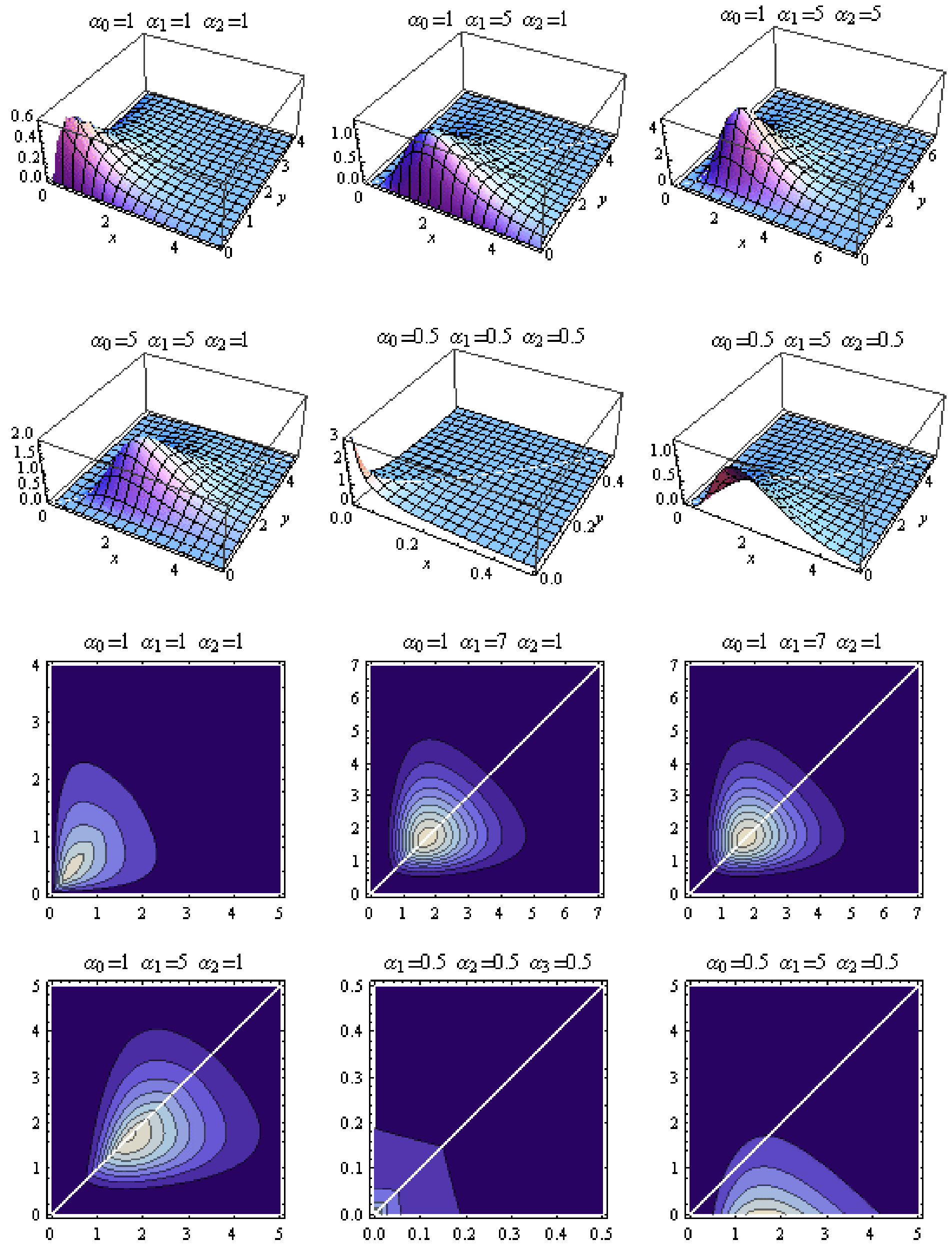


Figure 1. The density and contour plots for various values of α_0, α_1 and α_2.

where

$$I_1 = \int_0^\infty \int_0^x e^{t_1 x + t_2 y} e^{-(x/\theta + y/\theta)} (1 - e^{-x/\theta})^{\alpha_1 - 1} (1 - e^{-y/\theta})^{\alpha_0 + \alpha_2 - 1} dy dx,$$

and

$$I_2 = \int_0^\infty \int_0^y e^{t_1 x + t_2 y} e^{-(x/\theta + y/\theta)} (1 - e^{-x/\theta})^{\alpha_0+\alpha_1-1} (1 - e^{-y/\theta})^{\alpha_2 - 1} dx dy.$$

For I_1, $\int_0^y e^{t_2\theta} e^{-\lambda(y/\theta)}(1-e^{-y/\theta})^{\lambda(\alpha_0+\alpha_2-1)}dx = \theta\, B_{1-e^{-x/\theta}}(\alpha_0+\alpha_2, 1-\theta\, t_2)$, $|t_2| < \theta^{-1}$ and $B_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$, is the incomplete beta function. On using the series representation,
$B_x(a,b) = \sum_{k=0}^\infty \frac{(1-b)_k\, x^{a+k}}{k!\,(a+k)}$ where $(a)_k = a(a-1)\cdots(a-k+1)$,
[http://mathworld.wolfram.com/IncompleteBetaFunction.html], one can show

$$I_1 = \theta^2 \sum_{k=0}^\infty \frac{(\theta\, t_2)_k}{k!\,(\alpha_0+\alpha_2+k)} B(\alpha_0+\alpha_1+\alpha_2+k, 1-\theta\, t_1), \quad |t_1|, |t_2| < \theta^{-1}. \tag{3.2}$$

Similarly,

$$I_2 = \theta^2 \sum_{k=0}^\infty \frac{(\theta\, t_1)_k}{k!\,(\alpha_0+\alpha_1+k)} B(\alpha_0+\alpha_1+\alpha_2+k, 1-\theta\, t_2), \quad |t_1|, |t_2| < \theta^{-1}. \tag{3.3}$$

Substituting (3.2) and (3.3) in (3.1), we get an expression for the joint moment generating function of (X, Y).

3.2. Marginal distributions. From (2.4), the marginal density of X is

$$\begin{aligned} f_X(x) &= \int_0^\infty f_{X,Y}(x,y)dy \\ &= \theta^{-1}\delta(\alpha_0,\alpha_1,\alpha_2)e^{-x/\theta} \\ &\quad \left(\left(\frac{1}{\alpha_2+\alpha_0} - \frac{1}{\alpha_2}\right)(1-e^{-x/\theta})^{\alpha_1+\alpha_2+\alpha_0-1} + \frac{1}{\alpha_2}(1-e^{-x/\theta})^{\alpha_1+\alpha_0-1}\right) \\ &\quad \times I(x>0). \end{aligned} \tag{3.4}$$

Similarly, the marginal density of Y is

$$\begin{aligned} f_Y(y) &= \theta^{-1}\delta(\alpha_0,\alpha_1,\alpha_2)e^{-y/\theta} \\ &\quad \left(\left(\frac{1}{\alpha_1+\alpha_0} - \frac{1}{\alpha_1}\right)(1-e^{-y/\theta})^{\alpha_1+\alpha_2+\alpha_0-1} + \frac{1}{\alpha_1}(1-e^{-y/\theta})^{\alpha_2+\alpha_0-1}\right) \\ &\quad \times I(y>0). \end{aligned} \tag{3.5}$$

Lemma 1. The marginal distributions of X and Y are weighted generalized exponential distributions.

Proof. From (3.4), one can write $f_X(x) = \sum_{i=1}^2 a_i f_{X_i}(x_i)$, where $\sum_{i=1}^2 a_i = 1$, $X_1 \sim GED(\alpha_1+\alpha_2+\alpha_0, \theta)$, $X_2 \sim GED(\alpha_1+\alpha_0,\theta)$, $a_1 = \frac{\delta(\alpha_0,\alpha_1,\alpha_2)}{\alpha_1+\alpha_2+\alpha_0}\left(\frac{1}{\alpha_0+\alpha_2} - \frac{1}{\alpha_2}\right)$ and $a_2 = \frac{\delta(\alpha_0,\alpha_1,\alpha_2)}{(\alpha_0+\alpha_1)\alpha_2}$. Similarly, one can write (3.5) as $f_Y(y) = \sum_{i=1}^2 b_i f_{Y_i}(y_i)$, where $b_1 = \frac{\delta(\alpha_0,\alpha_1,\alpha_2)}{\alpha_1+\alpha_2+\alpha_0}\left(\frac{1}{\alpha_0+\alpha_1} - \frac{1}{\alpha_1}\right)$ and $b_2 = \frac{\delta(\alpha_0,\alpha_1,\alpha_2)}{(\alpha_0+\alpha_1)\alpha_1}$ and $Y_1 \sim GED(\alpha_1+\alpha_2+\alpha_0,\theta)$, $Y_2 \sim GED(\alpha_2+\alpha_0,\theta)$. □

Now, consider the following lemma from Gupta and Kundu (2001).

Lemma 2. If T follows generalized exponential distribution (GED) with parameters (α, λ), then,

(i) $M_T(t) = \alpha B(\alpha, 1 - t/\lambda), \quad |t| < \lambda.$
(ii) $E(T) = (\psi(\alpha+1) - \psi(1))/\lambda$, where $\psi(.)$ is the digamma function.

From Lemma 1, the moment generating function of X and Y, respectively, can be written as

$$M_X(t) = a_1 M_{X_1}(t) + a_2 M_{X_2}(t), \tag{3.6}$$

$$M_Y(t) = b_1 M_{Y_1}(t) + b_2 M_{Y_2}(t), \tag{3.7}$$

where a_1, a_2, b_1 and b_2 are mentioned in the proof of Lemma 1. Here, $X_1, Y_1 \sim GED(\alpha_1 + \alpha_2 + \alpha_0, \theta)$, $X_2 \sim GED(\alpha_1 + \alpha_0, \theta)$ and $Y_2 \sim GED(\alpha_2 + \alpha_0, \theta)$.
Hence, using (3.6), (3.7) and Lemma 2, we get

$$\begin{aligned} M_X(t) &= (\alpha_0 + \alpha_1 + \alpha_2) a_1 B(\alpha_0 + \alpha_1 + \alpha_2, 1 - t/\theta) \\ &\quad + (\alpha_0 + \alpha_1) a_2 B(\alpha_0 + \alpha_1, 1 - t/\theta), |t| < \theta, \end{aligned}$$

$$\begin{aligned} M_Y(t) &= (\alpha_0 + \alpha_1 + \alpha_2) b_1 B(\alpha_0 + \alpha_1 + \alpha_2, 1 - t/\theta) \\ &\quad + (\alpha_0 + \alpha_2) b_2 B(\alpha_0 + \alpha_2, 1 - t/\theta), |t| < \theta, \end{aligned}$$

$$E(X) = a_1\, \theta^{-1}\, \psi(\alpha_0 + \alpha_1 + \alpha_2 + 1) + a_2\, \theta^{-1}\, \psi(\alpha_1 + \alpha_0 + 1) - (a_1 + a_2)\, \theta^{-1}\, \psi(1),$$

and

$$E(Y) = b_1\, \theta^{-1}\, \psi(\alpha_0 + \alpha_1 + \alpha_2 + 1) + b_2\, \theta^{-1}\, \psi(\alpha_2 + \alpha_0 + 1) - (b_1 + b_2)\, \theta^{-1}\, \psi(1).$$

3.3. Distributions of max(X,Y) and min(X,Y). To find the distribution of $Z = \min(X, Y)$, we consider the following: For any $z \in (0, \infty)$

$$\begin{aligned} &P(Z > z) \\ &= \int_z^\infty \int_z^y f(x,y)\,dx\,dy + \int_z^\infty \int_z^x f(x,y)\,dy\,dx \\ &= \frac{\delta(\alpha_0, \alpha_1, \alpha_2)}{(\alpha_1 + \alpha_0)} \left(\frac{1}{\alpha_0 + \alpha_1 + \alpha_2} - \frac{(1 - e^{-z/\theta})^{\alpha_1 + \alpha_0}}{\alpha_2} + \frac{(\alpha_1 + \alpha_0)(1 - e^{-z/\theta})^{\alpha_0 + \alpha_1 + \alpha_2}}{\alpha_2(\alpha_0 + \alpha_1 + \alpha_2)} \right) \\ &+ \frac{\delta(\alpha_0, \alpha_1, \alpha_2)}{(\alpha_2 + \alpha_0)} \left(\frac{1}{\alpha_0 + \alpha_1 + \alpha_2} - \frac{1}{\alpha_1}(1 - e^{-z/\theta})^{\alpha_2 + \alpha_0} + \frac{\alpha_2 + \alpha_0}{\alpha_1(\alpha_0 + \alpha_1 + \alpha_2)}(1 - e^{-z/\theta})^{\alpha_0 + \alpha_1 + \alpha_2} \right). \end{aligned} \tag{3.8}$$

On differentiation (3.8), we get

$$\begin{aligned} f(z) &= \theta^{-1} \delta(\alpha_0, \alpha_1, \alpha_2) e^{-z/\theta} \\ &\times \left(\frac{1}{\alpha_1}(1 - e^{-z/\theta})^{\alpha_2 + \alpha_0 - 1} + \frac{1}{\alpha_2}(1 - e^{-z/\theta})^{\alpha_1 + \alpha_0 - 1} \right. \\ &\left. - \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} \right)(1 - e^{-z/\theta})^{\alpha_0 + \alpha_1 + \alpha_2 - 1} \right) \times I(z > 0). \end{aligned} \tag{3.9}$$

Lemma 3. The distribution of $\min(X, Y)$ is a weighted generalized exponential distribution.

Proof. From (3.9), $f_Z(z) = \sum_{i=1}^{3} c_i f_{Z_i}(z_i)$, where $\sum_{i=1}^{3} c_i = 1$, $Z_1 \sim GED(\alpha_2 + \alpha_0, \theta)$, $Z_2 \sim GED(\alpha_1 + \alpha_0, \theta)$, $Z_3 \sim GED(\alpha_0 + \alpha_1 + \alpha_2, \theta)$ and $c_1 = \frac{\delta(\alpha_0,\alpha_1,\alpha_2)}{\alpha_1(\alpha_2+\alpha_0)}$, $c_2 = \frac{\delta(\alpha_0,\alpha_1,\alpha_2)}{(\alpha_0+\alpha_1)\alpha_2}$ and $c_3 = \frac{\delta(\alpha_0,\alpha_1,\alpha_2)}{\alpha_2+\alpha_1+\alpha_0}\left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2}\right)$. □

For the distribution of $W = \max(X, Y)$, note that for any $w \in (0, \infty)$,

$$\begin{aligned}
\overline{F}_W(w) &= P(W > w) \\
&= P(X > w \quad or \quad Y > w) \\
&= P(X > w) + P(Y > w) - P(X > w \quad and \quad Y > w) \\
&= P(X > w) + P(Y > w) - P(Z > w) \\
&= \overline{F}_X(w) + \overline{F}_Y(w) - \overline{F}_Z(w). \qquad (3.10)
\end{aligned}$$

Differentiating (3.10) with respect to w and using (3.4), (3.5) and (3.9) we get:

$$\begin{aligned}
f_W(w) &= f_X(w) + f_Y(w) - f_Z(w) \\
&= \theta^{-1}(\alpha_0 + \alpha_1 + \alpha_2)e^{-w/\theta}(1 - e^{-w/\theta})^{\alpha_0+\alpha_1+\alpha_2-1} \times I(w > 0). \qquad (3.11)
\end{aligned}$$

From (3.11), $W = \max(X, Y)$, follows the generalized exponential distribution with parameters $\alpha_0 + \alpha_1 + \alpha_2$ and θ. Using equations (3.9), (3.11) and Lemma 2, the moment generating functions and the means of Z and W are:

(i) $M_Z(t) = c_1(\alpha_2 + \alpha_0)B(\alpha_2 + \alpha_0, 1 - t/\theta) + c_2(\alpha_1 + \alpha_0)B(\alpha_1 + \alpha_0, 1 - t/\theta) + c_3(\alpha_0 + \alpha_1 + \alpha_2)B(\alpha_0 + \alpha_1 + \alpha_2, 1 - t/\theta)$, $|t| < \theta$.
$E(Z) = c_1\theta^{-1}\psi(\alpha_2 + \alpha_0 + 1) + c_2\theta^{-1}\psi(\alpha_2 + \alpha_0 + 1) + c_3\theta^{-1}\psi(\alpha_2 + \alpha_1 + \alpha_0 + 1) - (c_1 + c_2 + c_3)\theta^{-1}\psi(1)$.

(ii) $M_W(t) = (\alpha_2 + \alpha_1 + \alpha_0)B(\alpha_2 + \alpha_1 + \alpha_0, 1 - t/\theta)$, $|t| < \theta$.
$E(W) = \theta^{-1}\left(\psi(\alpha_2 + \alpha_1 + \alpha_0 + 1) - \psi(1)\right)$.

3.4. Renyi Entropy. Shannon's (1948), pioneering work, entropy has been used as a major tool in information theory and in almost every branch of science and engineering. One of the main extensions of Shannon entropy was defined by Renyi (1961). This generalized entropy measure is given by

$$I_R(\lambda) = \frac{\log(G(\lambda))}{1 - \lambda}, \quad \lambda > 0, \lambda \neq 1. \qquad (3.12)$$

Where $G(\lambda) = \int_{\mathbb{X}} f^\lambda d\mu$, and μ is a σ-finite measure on $\mathbb{X}$. One can get an expression for the Shannon entropy from (3.12) by taking limit for $\lambda \to 1$.

3.1. Theorem. *The Renyi entropy for the bivariate generalized exponential distribution in (2.4) is* $I_R(\lambda) = (1 - \lambda)^{-1}\log(G(\lambda))$, *where*

(3.13)

$$G(\lambda) = \theta^{2-2\lambda}\delta^\lambda(\alpha_0, \alpha_1, \alpha_2)\sum_{k=0}^{\infty}\left(T^{\lambda,k}_{\alpha_0+\alpha_1} + T^{\lambda,k}_{\alpha_0+\alpha_2}\right)B(\lambda, \lambda(\alpha_2+\alpha_1+\alpha_0-2)+k+2),$$

and $T^{\lambda,k}_x = \frac{(1-\lambda)_k}{k!\,[\lambda(x-1)+k+1]}$.

Proof. From (2.4), we can write

$$G(\lambda) = \theta^{-2\lambda}\delta^{\lambda}(\alpha_0, \alpha_1, \alpha_2) \times (I_1 + I_2), \tag{3.14}$$

where $I_1 = \int_0^{\infty}\int_0^{y} e^{-\lambda(x/\theta+y/\theta)}(1-e^{-x/\theta})^{\lambda(\alpha_0+\alpha_1-1)}(1-e^{-y/\theta})^{\lambda(\alpha_2-1)}dxdy$ and
$I_2 = \int_0^{\infty}\int_0^{x} e^{-\lambda(x/\theta+y/\theta)}(1-e^{-x/\theta})^{\lambda(\alpha_1-1)}(1-e^{-y/\theta})^{\lambda(\alpha_0+\alpha_2-1)}dydx$.
The result in (3.13) follows from (3.14) by using similar approach as in equations (3.2) and (3.3). □

3.5. Stochastic properties. Let t_{11}, t_{12}, t_{21} and t_{22} be real numbers with $0 < t_{11} < t_{12}$ and $0 < t_{21} < t_{22}$. Then (X, Y) has the total positivity of order two (TP_2) property iff

$$f_{X,Y}(t_{11}, t_{21})f_{X,Y}(t_{12}, t_{22}) - f_{X,Y}(t_{12}, t_{21})f_{X,Y}(t_{11}, t_{22}) \geq 0. \tag{3.15}$$

3.2. Theorem. *The bivariate generalized exponential distribution in (2.4) has the* TP_2 *property.*

Proof. Let us consider different cases separately. If $0 < t_{11} < t_{21} < t_{12} < t_{22}$, then for the density function in (2.4), one can easily show that the condition in (3.15) is equivalent to $e^{-t_{21}/\theta} - e^{-t_{12}/\theta} \geq 0$. This inequality holds because $t_{21} < t_{12}$. The other cases can be shown similarly. □

The reliability parameter R is defined as $R = P(X > Y)$, where X and Y are independent random variables. Numerous applications of the reliability parameter have appeared in the literature such as the area of classical stress-strength model and the break down of a system having two components. Other applications of the reliability parameter can be found in Hall (1984) and Weerahandi and Johnson (1992).

3.3. Theorem. *The reliability parameter of the bivariate weighted generalized exponential distribution is*

$$R = \frac{\delta(\alpha_0, \alpha_1, \alpha_2)}{\alpha_1\alpha_2}\left\{\frac{\alpha_1}{\alpha_1+\alpha_2} - \frac{\alpha_0}{\alpha_0+\alpha_2} + \frac{\alpha_0\alpha_2}{\alpha_0+\alpha_1+\alpha_2}\right\}.$$

Proof. Note that $(X, Y) \stackrel{d}{=} [(W_1, W_2)|W_0 < \min(W_1, W_2)]$ where the W_i's are independent and $W_i \sim \mathrm{GED}(\alpha_i, \theta)$ for $i = 0, 1, 2$. Thus,

$$\begin{aligned} P(X > Y) &= P(W_1 > W_2 | W_0 < \min(W_1, W_2)) \\ &= \frac{P(W_0 < W_2 < W_1)}{P(W_0 < \min(W_1, W_2))} \end{aligned} \tag{3.16}$$

By using straightforward integration one can easily show that

$$P(W_0 < W_2 < W_1) = \frac{\alpha_1}{\alpha_1+\alpha_2} - \frac{\alpha_0}{\alpha_0+\alpha_2} + \frac{\alpha_0\alpha_2}{\alpha_0+\alpha_1+\alpha_2}. \tag{3.17}$$

Substituting (2.3) and (3.17) in (3.16), the result follows immediately. □

4. Multivariate weighted generalized exponential distribution

One can obtain a multivariate version of (2.1) by assuming $W_i \sim f_{W_i}(w_i)$ for $i = 0, 1, \cdots, k$ are independent random variables. The resulting multivariate weighted density function is given by

$$f_{X_1,X_2,\ldots X_k}(x_1, x_2, \cdots, x_k) = \frac{\left[\prod_{i=1}^{k} f_{W_i}(x_i)\right] F_{W_0}(\min(x_1, x_2, \ldots x_k))}{P(W_0 < \min(W_1, W_2, \ldots, W_k))}. \tag{4.1}$$

From (4.1), a multivariate extension of the bivariate weighted generalized exponential model in (2.4) is given by

(4.2)

$$f(x_1, x_2, \ldots, x_k) \propto \left(\prod_{i=1}^{k} \frac{\alpha_i}{\theta}\right) e^{(-\sum_{i=1}^{k} \frac{x_i}{\theta})} \left(\prod_{i=1}^{k} (1 - e^{-\frac{x_i}{\theta}})^{\alpha_i - 1}\right) \left(1 - e^{-\frac{x_{1:k}}{\theta}}\right)^{\alpha_i - 1} \times I(\underline{x} > \underline{0}),$$

where $x_{1:k} = \min\{(x_1, x_2, \ldots, x_k\}$.

As a motivation, one can consider the following scenario: suppose that a system consists of k components whose random strengths are denoted by $W_1, W_2, \ldots W_k$ and the random stress is given by W_0. Next, if the system has a series structure then one would be interested to know the distribution of $W_1, W_2, \ldots W_k | W_0 < \min(W_1, W_2, \ldots, W_k)$. In fact the system reliability in that case would be given by $R = P(W_0 < \min(W_1, W_2, \ldots, W_k))$. Next, consider the model in which $Y_1, Y_2, \ldots, Y_j$ are i.i.d. random variables with distribution and density functions G_0 and g_0; $X_1, X_2, \ldots, X_k$ are i.i.d. random variables with distribution and density functions F_0 and f_0 and $Z_1, Z_2, \ldots, Z_\ell$ are i.i.d. random variables with distribution and density functions H_0 and h_0. In this case we have

$$f(x_1, x_2, \ldots, x_k) \propto \left[\prod_{i=1}^{k} f_0(x_i)\right] [G_0(x_{1:k})]^j [1 - H_0(x_{k:k})]^\ell, \tag{4.3}$$

where $x_{k:k} = \max\{(x_1, x_2, \ldots, x_k\}$.

In some specific scenarios it will be possible to evaluate the normalizing constant in (4.3). For example, when the three distributions are generalized exponential, (4.3) reduces to

$$\begin{aligned} f(x_1, x_2, \ldots, x_k) &\propto \left[\alpha_1^k e^{-\theta^{-1}(x_1 + x_2 + \cdots + x_k)}\right] \prod_{i=1}^{k} \left(1 - e^{-x_i/\theta}\right)^{\alpha_1} \\ &\times \left[1 - \left(1 - e^{-x_{k:k}/\theta}\right)^{\alpha_0}\right]^\ell \left(1 - e^{-x_{k:k}/\theta}\right)^{j\,\alpha_2}. \end{aligned} \tag{4.4}$$

To identify the required normalizing constant we must evaluate

$$\begin{aligned} &\int_0^\infty \int_0^\infty \ldots. \int_0^\infty \alpha_1^k e^{-\theta^{-1}(x_1 + x_2 + \cdots + x_k)} \prod_{i=1}^{k} \left(1 - e^{-x_i/\theta}\right)^{\alpha_1} \\ &\times \left[1 - \left(1 - e^{-x_{k:k}/\theta}\right)^{\alpha_0}\right]^\ell \left(\left(1 - e^{-x_{k:k}/\theta}\right)^{\alpha_2}\right)^j dx_1 dx_2 .. dx_k \\ &= \sum_{k_1=0}^{\ell} \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} \binom{\ell}{k_1} \binom{\alpha_0 k_1}{k_2} \binom{\alpha_2 j}{k_3} (-1)^{k_1+k_2+k_3} E\left(e^{-k_1 X_{1:k}/\theta - k_2 X_{k:k}/\theta}\right), \end{aligned} \tag{4.5}$$

where the X_i's have the generalized exponential (α_1, θ) distribution. So we need the joint moment generating function of $(X_{1:k}, X_{k:k})$. Next, the joint distribution of $(X_{1:k}, X_{k:k})$

is

$$\begin{aligned}
&f(x_{1:k}, x_{k:k}) \\
&= \frac{k^3(k-1)\alpha_1^2}{\theta^2} e^{-x_{1:k}/\theta - x_{k:k}/\theta} \left(1 - e^{-x_{k:k}/\theta}\right)^{k\alpha_1 - 1} \left(1 - e^{-x_{1:k}/\theta}\right)^{\alpha_1 - 1} \\
&\left(1 - \left(1 - e^{-x_{1:k}/\theta}\right)^{\alpha_1}\right)^{k-1} \left(\left(1 - e^{-x_{k:k}/\theta}\right)^{\alpha_1} - \left(1 - e^{-x_{1:k}/\theta}\right)^{\alpha_1}\right)^{k-2} \\
&\times I(0 < x_{1:k} < x_{k:k} < \infty)
\end{aligned}$$

Now,

$$\begin{aligned}
&E\left(e^{-X_{1:k}/\theta - X_{k:k}/\theta}\right) \\
&= \int_0^\infty \int_0^{x_{k:k}} \frac{k^3(k-1)\alpha_1^2}{\theta^2} e^{-2x_{1:k}/\theta - 2x_{k:k}/\theta} \left(1 - e^{-x_{k:k}/\theta}\right)^{k\alpha_1 - 1} \left(1 - e^{-x_{1:k}/\theta}\right)^{\alpha_1 - 1} \\
&\times \left(1 - \left(1 - e^{-x_{1:k}/\theta}\right)^{\alpha_1}\right)^{k-1} \left(\left(1 - e^{-x_{k:k}/\theta}\right)^{\alpha_1} - \left(1 - e^{-x_{1:k}/\theta}\right)^{\alpha_1}\right)^{k-2} dx_{1:k} dx_{k:k}, \qquad (4.6)
\end{aligned}$$

which can be written as $\sum_{j=0}^{\infty}(-1)^j \frac{\theta}{2+j}\left(\binom{(2k+1)\alpha_1}{j} + \binom{(2k+1)\alpha_1 - 1}{j}\right)$, after some algebraic simplification. Hence, using (4.6) in (4.5), the normalizing constant corresponds to the distribution in (4.4) is

$$\begin{aligned}
C = \theta \sum_{k_1=0}^{\ell} \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{k_1+k_2+k_3+j}}{2+j} \binom{\ell}{k_1} \binom{\alpha_0 k_1}{k_2} \binom{\alpha_2 j}{k_3} \\
\left(\binom{(2k+1)\alpha_1}{j} + \binom{(2k+1)\alpha_1 - 1}{j}\right),
\end{aligned}$$

Corollary 1. If $(X_1, X_2, ..., X_k)$ has a multivariate weighted generalized exponential distribution in (4.1) with parameters $(\alpha_i, \theta), i = 0, 1, 2, \cdots k$, then the normalizing constant, $C_1 = P(X_1 < \min(X_2, \cdots, X_k)) = \frac{\sum_{i=0}^k \alpha_i}{\prod_{i=1}^k \alpha_i}\left(\sum_{i=1}^k \frac{1}{\alpha_i + \alpha_0}\right)^{-1}$.

Proof. The result follows immediately by using the same logic as in (2.3). □

Corollary 2. If $(X_1, X_2, ..., X_k)$ has a multivariate weighted generalized exponential distribution with parameters $(\alpha_i, \theta), i = 0, 1, 2, \cdots k$, then the distribution of $Z = \min(X_1, X_2, \cdots, X_k)$ has the density

$$\begin{aligned}
f(z) = \theta^{-1} C_1^{-1} e^{-z/\theta} \left(\sum_{i=1}^{k} \frac{1}{\alpha_i}(1 - e^{-z/\theta})^{\alpha_i + \alpha_0 - 1} - \frac{1}{\sum_{i=1}^k \alpha_i}(1 - e^{-z/\theta})^{\sum_{i=0}^k \alpha_i - 1}\right) \\
\times I(z > 0),
\end{aligned}$$

where C_1 is the constant in Corollary 1.

5. Estimation

In this section, we consider the maximum likelihood method to estimate the model parameters of the bivariate generalized exponential distribution in (2.4).

5.1. Maximum likelihood estimation. Assume that a random sample of size n observations are taken from the bivariate density in (2.4), then the corresponding log-likelihood function can be written as

$$\begin{aligned}\ell(\alpha_0,\alpha_1,\alpha_2,\theta) &= -2n\log\theta + n\log(\alpha_0+\alpha_1+\alpha_2) + n\log(\alpha_0+\alpha_1) + n\log(\alpha_0+\alpha_2)\\ &\quad -n\log(2\alpha_0+\alpha_1+\alpha_2) - n\theta^{-1}(\bar{x}+\bar{y}) + (\alpha_1-1)\sum_{i=1}^{n}\log\left(1-e^{-x_i/\theta}\right)\\ &\quad +(\alpha_2-1)\sum_{i=1}^{n}\log\left(1-e^{-y_i/\theta}\right) + \alpha_0\sum_{i=1}^{n}\log\left(1-e^{-\min(x_i,y_i)/\theta}\right).\end{aligned} \tag{5.1}$$

Differentiating (5.1) with respect to α_0, α_1, α_2, and θ we get

$$\frac{\partial}{\partial\alpha_0}\ell(\alpha_0,\alpha_1,\alpha_2,\theta) = \frac{n}{\alpha_0+\alpha_1+\alpha_2} - \frac{2n}{2\alpha_0+\alpha_1+\alpha_2} + \frac{n}{\alpha_0+\alpha_1} + \frac{n}{\alpha_0+\alpha_2} + \sum_{i=1}^{n}\log\left(1-e^{-\min(x_i,y_i)/\theta}\right). \tag{5.2}$$

$$\frac{\partial}{\partial\alpha_1}\ell(\alpha_0,\alpha_1,\alpha_2,\theta) = \frac{n}{\alpha_0+\alpha_1+\alpha_2} - \frac{n}{2\alpha_0+\alpha_1+\alpha_2} + \frac{n}{\alpha_0+\alpha_1} + \sum_{i=1}^{n}\log\left(1-e^{-x_i/\theta}\right). \tag{5.3}$$

$$\frac{\partial}{\partial\alpha_2}\ell(\alpha_0,\alpha_1,\alpha_2,\theta) = \frac{n}{\alpha_0+\alpha_1+\alpha_2} - \frac{n}{2\alpha_0+\alpha_1+\alpha_2} + \frac{n}{\alpha_0+\alpha_2} + \sum_{i=1}^{n}\log\left(1-e^{-y_i/\theta}\right). \tag{5.4}$$

$$\begin{aligned}&\frac{\partial}{\partial\theta}\ell(\alpha_0,\alpha_1,\alpha_2,\theta)\\ &= -2n\theta^{-1} + \theta^{-2}n(\bar{x}+\bar{y}) - (\alpha_1-1)\theta^{-2}\sum_{i=1}^{n}x_i\left(e^{x_i/\theta}-1\right)^{-1}\\ &\quad -(\alpha_2-1)\theta^{-2}\sum_{i=1}^{n}y_i\left(e^{y_i/\theta}-1\right)^{-1} - \alpha_0\theta^{-2}\sum_{i=1}^{n}\min(x_i,y_i)\left(e^{\min(x_i,y_i/\theta)}-1\right)^{-1}.\end{aligned} \tag{5.5}$$

Setting (5.2), (5.3), (5.4) and (5.5) to 0 and solving simultaneously, we get the maximum likelihood estimates for α_0, α_1, α_2 and θ.

If the scale parameter θ is assumed to be known, then setting equations (5.2), (5.3) and (5.4) equal to zero, we get,

$$\frac{1}{\alpha} - \frac{2}{\alpha_0+\alpha} + \frac{1}{\alpha_0+\alpha_1} + \frac{1}{\alpha_0+\alpha_2} = C. \tag{5.6}$$

$$\frac{1}{\alpha} - \frac{1}{\alpha_0+\alpha} + \frac{1}{\alpha_0+\alpha_1} = A. \tag{5.7}$$

$$\frac{1}{\alpha} - \frac{1}{\alpha_0+\alpha} + \frac{1}{\alpha_0+\alpha_2} = B. \tag{5.8}$$

where $A = -\sum_{i=1}^{n}\log\left(1-e^{-X_i/\theta}\right)$, $B = -\sum_{i=1}^{n}\log\left(1-e^{-Y_i/\theta}\right)$,
$C = -\sum_{i=1}^{n}\log\left(1-e^{-\min(X_i,Y_i)/\theta}\right)$ and $\alpha = \alpha_0+\alpha_1+\alpha_2$.

Adding (5.7) and (5.8) and then subtracting from (5.6), we get

$$\alpha = \frac{1}{A + B - C}. \tag{5.9}$$

On using (5.7) and (5.8) and then simplifying, we get

$$\alpha_2 = \alpha - \left(A - B + \frac{1}{\alpha - \alpha_1} \right)^{-1}. \tag{5.10}$$

Therefore, using equations (5.9), (5.10) and the fact that $\alpha_0 = \alpha - \alpha_1 - \alpha_2$, one can easily solve equation (5.6) for α_1. This will increase the calculation efficiency in order to obtain the numerical solution faster. The Fisher information matrix when θ is known, $I(\underline{\delta}) = -E\left(\frac{\partial^2}{\partial\delta_i \partial\delta_j} \log\left(f(X|\underline{\delta})\right)\right) = \{U_{rs}; r, s = \alpha_0, \alpha_1, \alpha_2\}$, can be obtained from equations (30)-(32) as follows:

$$U_{\alpha_0\alpha_0} = n\left(\alpha^{-2} - 4(\alpha_0 + \alpha)^{-2} + (\alpha_0 + \alpha_1)^{-2} + (\alpha_0 + \alpha_2)^{-2}\right).$$

$$U_{\alpha_0\alpha_1} = n\left(\alpha^{-2} - 2(\alpha_0 + \alpha)^{-2} + (\alpha_0 + \alpha_1)^{-2}\right).$$

$$U_{\alpha_0\alpha_2} = n\left(\alpha^{-2} - 2(\alpha_0 + \alpha)^{-2} + (\alpha_0 + \alpha_2)^{-2}\right).$$

$$U_{\alpha_1\alpha_1} = n\left(\alpha^{-2} - (\alpha_0 + \alpha)^{-2} + (\alpha_0 + \alpha_1)^{-2}\right).$$

$$U_{\alpha_1\alpha_2} = n\left(\alpha^{-2} - (\alpha_0 + \alpha)^{-2}\right),$$

$$U_{\alpha_2\alpha_2} = n\left(\alpha^{-2} - (\alpha_0 + \alpha)^{-2} + (\alpha_0 + \alpha_2)^{-2}\right).$$

The Fisher information matrix can be used to obtain interval estimation of the model parameters. Under standard regularity conditions, the multivariate normal $N_3(0, I(\widehat{\underline{\delta}})^{-1})$ distribution can be used to construct approximate confidence intervals for the model parameters. The matrix, $I(\widehat{\underline{\delta}})$ is the Fisher information matrix evaluated at $\widehat{\underline{\delta}}$. Therefore, the $100(1-a)\%$ confidence intervals for α_0, α_1 and α_2 are given by $\hat{\alpha_0} \pm z_{a/2} \times \sqrt{var(\hat{\alpha_0})}$, $\hat{\alpha_1} \pm z_{a/2} \times \sqrt{var(\hat{\alpha_1})}$, and $\hat{\alpha_2} \pm z_{a/2} \times \sqrt{var(\hat{\alpha_2})}$, respectively, where

$$Var(\widehat{\alpha}_0) = \frac{10\alpha_0^4 + 22\alpha_0^3(\alpha_1+\alpha_2) + (\alpha_1+\alpha_2)^4 + \alpha_0^2\left(20\alpha_1^2 + 34\alpha_1\alpha_2 + 20\alpha_2^2\right) + 4\alpha_0\left(2\alpha_1^3 + 5\alpha_1^2\alpha_2 + 5\alpha_1\alpha_2^2 + 2\alpha_2^3\right)}{2n(\alpha_0+\alpha_1)(\alpha_0+\alpha_2)},$$

$$Var(\widehat{\alpha}_1) = \frac{5\alpha_0^3 + 8\alpha_0^2\alpha_1 + 11\alpha_0^2\alpha_2 + 6\alpha_0\alpha_1^2 + 12\alpha_0\alpha_1\alpha_2 + 7\alpha_0\alpha_2^2 + 2\alpha_1^3 + 4\alpha_1^2\alpha_2 + 4\alpha_1\alpha_2^2 + \alpha_2^3}{2n(\alpha_0+\alpha_1)},$$

$$Var(\widehat{\alpha}_2) = \frac{5\alpha_0^3 + 11\alpha_0^2\alpha_1 + 8\alpha_0^2\alpha_2 + 7\alpha_0\alpha_1^2 + 12\alpha_0\alpha_1\alpha_2 + 6\alpha_0\alpha_2^2 + \alpha_1^3 + 4\alpha_1^2\alpha_2 + 4\alpha_1\alpha_2^2 + 2\alpha_2^3}{2n(\alpha_0+\alpha_2)}.$$

5.2. Simulation study. To illustrate the application of the bivariate generalized exponential distribution in (2.4), a small simulation study is conducted. However, in this paper we report only the results for estimation of the model parameters using the maximum likelihood estimation procedure. Bivariate random samples of size 50, 100 and 200 were generated from the density in (2.4) with the following parameter values; Set I: $\alpha_0 = 1$, $\alpha_1 = 5$, $\alpha_2 = 5$ and $\theta = 1$ and Set II: $\alpha_0 = 1$, $\alpha_1 = 4$, $\alpha_2 = 3$ and $\theta = 3$. Since both the conditional distributions of the bivariate density in (2.4), $X|Y$ and $Y|X$, are completely known in closed forms, a Gibbs sampling technique is used to generate bivariate random samples. The simulation is repeated 200 times. The estimated value and the standard deviation of the parameters using the maximum likelihood method are presented in Tables 1 and 2.

Table 1. Parameter estimates and standard deviations for BWGED under set I.

Sample size	$\hat{\alpha_0}$	$\hat{\alpha_1}$	$\hat{\alpha_2}$	$\hat{\theta}$
50	1.2372 (0.3220)	5.0955 (0.8543)	5.0902 (0.7065)	0.9903 (0.0379)
100	1.1616 (0.1662)	4.9123 (0.5579)	5.1261 (0.6607)	0.9767 (0.0217)
200	1.1510 (0.1190)	5.0190 (0.4097)	4.9918 (0.2449)	0.9956 (0.0207)

Table 2. Parameter estimates and standard deviations for the BWGED under set II.

Sample size	$\hat{\alpha_0}$	$\hat{\alpha_1}$	$\hat{\alpha_2}$	$\hat{\theta}$
50	1.1526 (0.4653)	4.1931 (0.5748)	3.5209 (0.7732)	2.8400 (0.2598)
100	1.2204 (0.2971)	3.9411 (0.4501)	3.5043 (0.6746)	2.9127 (0.2081)
200	1.1374 (0.1573)	4.1051 (0.2225)	3.2157 (0.4294)	2.9677 (0.1140)

From Tables 1 and 2, it appears that the maximum likelihood estimation performs quite effectively to estimate the model parameters.

6. Application

In this section, the BWGED is applied to a data set from Al-Mutairi at el. (2011). The data set represents the scores from twenty five first year graduate students in probability and inference classes of a premier Institute in India. For both the courses, Analysis-I is a prerequisite. It is assumed that the knowledge of Analysis-I affects the scores in both the courses. The data set is

$$X : 53, 55, 85, 87, 22, 23, 25, 93, 51, 62, 53, 32, 43, 47, 30, 88, 59, 49, 42, 71, 41, 82, 75, 93, 37.$$

$$Y : 89, 90, 59, 50, 25, 29, 54, 62, 39, 25, 89, 32, 33, 63, 38, 77, 55, 41, 31, 66, 57, 32, 43, 88, 34.$$

We fit the data set to the BWGED and compared the result with the bivariate weighted exponential distribution (Al- Murairi et al., 2011). The maximum likelihood estimates for both models are reported in Table 3. The Kolmogorov-Smirnov test statistic (K-S) for the distribution functions of the marginal X and Y is used to compare the goodness of fit of the BWGED and the bivariate weighted exponential distribution (BWED). The K-S statistics and the p-value for the K-S statistics for the fitted marginal distributions are reported in Tables 3. From Table 3, the p-values indicate that the marginals of the BWGED gives an adequate fit to the data. Figure (2) displays the empirical and the fitted cumulative distribution functions. This figure supports the results in Table 3.

Table 3. Parameter estimates for the scores data

Distribution	BWED	BWGED
Parameter Estimates	$\hat{\lambda_1} = 0.0263$	$\hat{\theta} = 20.9321$
	$\hat{\lambda_2} = 0.0293$	$\hat{\alpha_0} = 10.7633$
	$\hat{\lambda_3} = 0.0005$	$\hat{\alpha_1} = 0.9752$
		$\hat{\alpha_2} = 1 \times 10^{-6}$
K-S for X	0.3290	0.0790
K-S p-value for X	0.2080	0.9977
K-S for Y	0.2250	0.1300
K-S p-value for Y	0.2860	0.7924

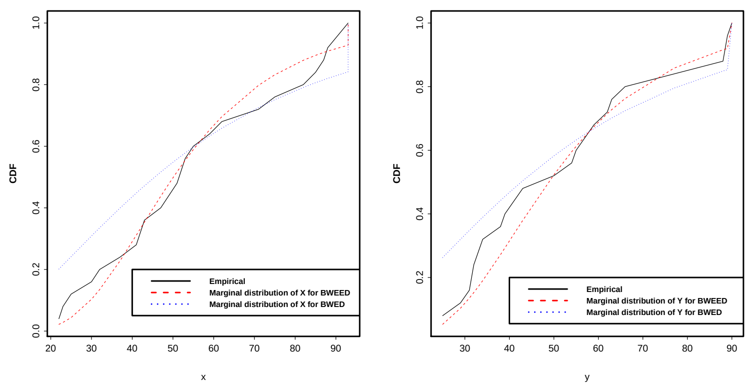


Figure 2. Marginal CDFs for fitted distributions of the scores data

7. Concluding remarks

In this paper, we consider a method for generating bivariate and multivariate generalized exponential distributions. Some structural properties of the bivariate exponentiated-exponential distribution in (2.4) are studied such as marginal distributions, moments, total positivity and parameter estimation. A small simulation study is conducted and the outcome of the simulation study is quite encouraging. Furthermore, one can study general properties for the multivariate generalized exponential distribution in (4.4). Although, in this paper, we focus on the bivariate and multivariate generalized exponential distributions, one can use the techniques in (2.1) and (4.3) to generate different bivariate and multivariate distributions. The analytical tractability of such resulting models is to be investigated before one can explore other properties of the derived model(s).

Acknowledgments
The authors would like to thank the Editor-in-Chief, and the referees for the constructive comments which greatly improved the paper.

References

[1] Al-Mutairi, D. K., Ghitany,M. E., and Kundu, D. *A new bivariate distribution with weighted exponential marginals and its multivariate generalization*, Statistical Papers, **52**, 921-936, 2011.

[2] Arellano-Valle, R.B. and Azzalini, A. *On unification of families of skew-normal distributions*, Scandinavian Journal of Statistics, **33**, 561-574, 2004.

[3] Arnold, B.C. and Beaver, R.J. *Skewed multivariate models related to hidden truncation and or selective reporting.* Test, 11, 7-54, 2002.

[4] Arnold, B.C. and Nagaraja, H.N. *On some properties of bivariate weighted distributions,* Communications in Statistics-Theory and Methods, **20**, 1853-1860, 1991.

[5] Azzalini, A. *A class of distributions which includes the normal ones,* Scandinavian Journal of Statistics, **12**, 171-178, 1985.

[6] Bayarri, M. J., and DeGroot, M. *A BAD view of weighted distributions and selection models. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.)*, Bayesian statistics 4, Oxford: Oxford University Press, 1992.

[7] Branco, M.D. and Dey, D.K. *A general class of multivariate skew-normal distributions*, Journal of Multivariate Analysis, **79**, 99-113, 2001.

[8] Kim, H. J. *On a class of two-piece skew-normal distributions*, Statistics, **39**, 537-553, 2005.

[9] Genton, M.G. *Skew-Elliptical Distributions and their Applications: A Journey beyond Normality,* Chapman and Hall, New York, 2004.

[10] Gupta, R.D. and Kundu, D. *Exponentiated Exponential family; An alternative to Gamma and Weibull' Biometrical Journal*, **33**, 117-130, 2001.

[11] Hall, I.J. *Approximate one-sided tolerance limits for the difference or sum of two independent normal variates*, Journal of Qualitative Technology, **16**,15-19, 1984.

[12] Rao, C.R. *On discrete distributions arising out of methods of ascertainment. In G. P. Patil (Ed.) Classical an contagious discrete distributions*, Calcutta Statistical Publishing Society, 320-333, 1965.

[13] Rao, C.R *Weighted distributions arising out of methods of ascertainment, A Celebration of Statistics*, A.C.Atkinson and S.E. Fienberg, eds., Springer-Verlag, NY, **24**, 543-569, 1985.

[14] Renyi, A *On measures of information and entropy,* Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability , 547-561, 1961.

[15] Weerahandi, S. and Johnson, R.A. *Testing reliability in a stress-strength model when X and Y are normally distributed,* Technometrics, **38**, 83-91, 1992.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1541 – 1552

The Rayleigh paired comparison model with Bayesian analysis

Tanveer Kifayat*† and Muhammad Aslam‡

Abstract

A paired comparison (PC) method is more reliable to rank or compare more than two items/ objects at the same time. It is a well-developed method of ordering attributes or characteristics of a given set of items.The PC model is developed using Rayleigh random variables on the basis of Stern's criteria [17]. The Rayleigh PC model is analyzed in Bayesian framework using non-informative (Jeffreys and Uniform) priors. The Bayesian inference of the developed model is compared with existing the Bradley-Terry model. The preference and predictive probabilities for current and future comparisons are calculated. The posterior probabilities of hypotheses for comparing two parameters are evaluated. The Bayesian 95% credible interval are calculated.Appropriateness of the model is also examined. Graphs of marginal posterior distributions of the parameters are drawn. The Bayesian analysis is performed using real life data sets.

Keywords: Paired Comparisons, Rayleigh Distribution, Non-Informative Prior, Posterior Probability , Predictive Probability.

Received : 17.12.2014 *Accepted* : 14.07.2015 *Doi* : 10.15672/HJMS.20158212610

*Department of Statistics, Quaid-i-Azam University, Islamabad,Pakistan.
Email: tanveerqau@yahoo.com

†Tanveer Kifayat

‡Department of Mathematics and Statistics, Riphah International University, Islamabad, Pakistan.
Email: m.aslam@riphah.edu.pk

1. Introduction

A pair of objects is presented for comparison and two are placed in the relationship preferred or not preferred. If the differences among the objects are distinguishable and fairly apparent then ranking of all objects will be preferable where the objects will be given ranked values depends on preferences. For a detailed discussion on PC method and its usefulness, one is referred to [4], [7] and [9]
A PC model based on two Cauchy random variables has been developed by [1]. The model has been analyzed in the Bayesian frame work using informative, Conjugate and non-informative (Jeffreys and Uniform) priors. The real data set of top five ranked one day international cricket teams is collected for the Bayesian analysis. By study it is concluded that Australia has been ranked on the top. The technique of collecting preference data from judges through binary digits have been highlighted by [2]. The preferred item is denoted by one and zero to the non-preferred item. The Bradley-Terry PC model is used for analysis considering a real data set on ice-cream brands. [3] has been worked on two types of models to use ordinal scales for PC analysis for several parameters. He shows for binary scale that logit transformation for the models simplifies them to the basic Bradley-Terry model. [4] has given the Bayesian analysis of the Bradley-Terry and the Rao-Kupper model. The posterior means of the parameters, posterior probabilities of the hypotheses and predictive probabilities for both the models are included in this study. These results were using the non-informative (Jeffreys and Uniform) priors. [5] have estimated the parameters of the Thurstone-Mosteller PC model by method of maximum likelihood. The Bayesian analysis of the model is carried out using Jeffreys prior. The Binomial discrete logistic model for the relation between sensory and consumer preference have been presented by [8]. It is also concluded that no preference is better to model as a function than considered as ties for the sensory data. The Thurstone-Mosteller model for PCs has been modified by [10] which allows for widely differing proportions of draws. Data relating to games between the 64 greatest chess player of the world is analyzed for the model. [11] has discussed the technique of iterative maximum likelihood estimates algorithms for the generalization of the Bradley Terry model. [13] has presented that PC allow a large number of draws and variability of draw percentages among the players of chess or soccer matches. The results are based on matching the number of home wins, home draws, away wins and away draws for each team with their expected values. Glenn- David model is used for the estimation. [16] have recommended the procedure of lasso that categorized the contestants with similar aptitudes. The standard maximum likelihood method is used for the prediction of rankings. The teams ranking of National football league 2010-2011 and the American college hockey men's division I 2009-2010 have been used for the analysis.

In Section 2, the method of the Rayleigh model development and notations for the model is discussed. The Bradley-Terry model is given in Section 3.The prior distributions and the Bayesian analysis is provided in Sections 4 and 5. Concluding remarks are provided in Section 6.

2. The Rayleigh Paired Comparisons Model

By considering PC experiments of the Rayleigh random variables with same shape parameter and different scale (α_i) parameter, the Rayleigh model is derived on the basis of the Stern's model criteria . The Rayleigh random variables are used to examine wind velocity. The data of MRI images is also Rayleigh distributed. As the Rayleigh distribution can be used in communication theory, so in paired comparison, perception of the preference one object is communicated to the other object in a pair, for this reason, Rayleigh distribution may be considered for PC model. The probability that the preference of T_i over T_j is denoted by $\phi_{i.ij}$ and defined as:

$$\phi_{i.ij} = P(T_j \leq T_i)$$

$$\phi_{i.ij} = \int_0^\infty \int_{t_j}^\infty \frac{t_i}{\alpha_i^2} e^{-\frac{t_i^2}{2\alpha_i^2}} \frac{t_j}{\alpha_j^2} e^{-\frac{t_j^2}{2\alpha_j^2}} dt_i dt_j,$$

$$\phi_{i.ij} = \frac{\alpha_i^2}{\alpha_i^2 + \alpha_j^2} \tag{2.1}$$

and $\phi_{j.ij}$ is the probability that T_j is preferred over T_i and is obtained as:

$$\phi_{j.ij} = 1 - \phi_{i.ij}$$

$$\phi_{j.ij} = \frac{\alpha_j^2}{\alpha_i^2 + \alpha_j^2} \tag{2.2}$$

where α_i; $(i < j) = 1, 2, ...m$ are the treatment parameters. The (2.1) and (2.2) represent the model named as the Rayleigh model for PC.
We define the notations for the model. Let w_{ij} be the random variable associated with the rank of the treatments in the k^{th} repetition of the treatment pair (T_i, T_j), where $(i \neq j; i \geq 1, j \leq m; k = 1, 2, ..., r_{ij})$ and m is the number of observation.

$w_{i.ijk} =$1 or 0 accordingly as treatment T_i is preferred to treatment T_j or not in the k^{th} repetition of comparison.

$w_{j.ijk} =$1 or 0 accordingly as treatment T_j is preferred to treatment T_i or not in the k^{th} repetition of comparison.

$w_{i.ij} = \sum_k w_{i.ijk} =$ the number of times treatment T_i is preferred to treatment T_j.

$w_{j.ij} = \sum_k w_{j.ijk} =$ the number of times treatment T_j is preferred to treatment T_i.

$r_{ij} =$the number of times treatment T_i is compared with treatment T_j.

$r_{ij} = w_{i.ij} + w_{j.ij}$.

The likelihood function of the observed outcomes of the trial $\boldsymbol{w}$ and the parameters $\boldsymbol{\alpha}$=$\alpha_1, \alpha_2, ..., \alpha_m$ is :

$$l(\boldsymbol{w}, \boldsymbol{\alpha}) = \prod_{i<j=1}^{m} \frac{r_{ij}!}{w_{ij}!\ (r_{ij} - w_{ij})!} \frac{\alpha_i^{2w_{i.ij}} \alpha_j^{2w_{j.ij}}}{(\alpha_i^2 + \alpha_j^2)^{w_{i.ij} + w_{j.ij}}} \qquad , \alpha_i > 0 \tag{2.3}$$

A constraint is imposed on parameters of the model i.e., $\sum_{i=1}^m \alpha_i = 1$. This condition confirms that parameters are well defined.

3. The Bradley-Terry Model

The Bradley-Terry model is the basic PC model. [7] proposed a model of PCs assuming the Logistic density instead of the standard normal density using [15] and [18]. The difference between two latent variables (T_i, T_j) has a Logistic density with location parameters ($\ln \alpha_i$, $\ln \alpha_j$). The probability that treatment T_i is preferred to treatment T_j according to Bradley and Terry is given as:

$$\phi_{i.ij} = \frac{\alpha_i}{\alpha_i + \alpha_j} \tag{3.1}$$

$$\phi_{j.ij} = \frac{\alpha_j}{\alpha_i + \alpha_j}, \tag{3.2}$$

where (3.1) and (3.2) is known as the Bradley-Terry model for PC.

4. Non-Informative Prior Distributions

The non-informative (Uniform and Jeffreys) priors are assumed for the Bayesian analysis.

4.1. Uniform Prior. The Bayesian analysis of the unknown parameter using Uniform prior is suggested by [6] and [14]. We use the Uniform U(0,1) as the prior distribution, defined as:

$$p_U(\boldsymbol{\alpha}) \propto 1 \tag{4.1}$$

where α is defined in (2.3) and $\alpha_i > 0$. It is the improper prior.

4.2. Jeffreys Prior. The Jeffreys prior is defined as the density function proportional to partially differentiating twice the log likelihood function and taking the square root of the expected value, i.e.

$$p_J(\boldsymbol{\alpha}) \propto det[I(\boldsymbol{\alpha})]^{\frac{1}{2}} \tag{4.2}$$

$$\text{where } det[I(\boldsymbol{\alpha})] = (-1)^2 \begin{vmatrix} E[\frac{\partial^2 \log l(.)}{\partial \alpha_1^2}] & E[\frac{\partial^2 \log l(.)}{\partial \alpha_1 \alpha_2}] \\ E[\frac{\partial^2 \log l(.)}{\partial \alpha_2 \alpha_1}] & E[\frac{\partial^2 \log l(.)}{\partial \alpha_2^2}] \end{vmatrix}$$

for $m = 3$ and $\alpha_3 = 1 - \alpha_1 - \alpha_2$.

So, the Jeffreys prior for the parameters is derived as:

$$p_J(\boldsymbol{\alpha}) \propto \sqrt{\frac{A_1}{A_2}}$$

$$\begin{aligned} \text{where } A_1 =& 2\alpha_1^6 - 6\alpha_1^5 + 6\alpha_1^5\alpha_2 + 15\alpha_1^4\alpha_2^2 - 14\alpha_1^4\alpha_2 + 7\alpha_1^4 - 4\alpha_1^3 + 12\alpha_1^3\alpha_2 + \\ & 20\alpha_1^3\alpha_2^3 - 28\alpha_1^3\alpha_2^2 - 4\alpha_2\alpha_1^2 + \alpha_1^2 - 28\alpha_1^2\alpha_2^3 + 15\alpha_1^2\alpha_2^4 + 18\alpha_1^2\alpha_2^2 + \\ & 12\alpha_1\alpha_2^3 - 4\alpha_1\alpha_2^2 + 6\alpha_1\alpha_2^5 - 14\alpha_1\alpha_2^4 + \alpha_2^2 + 7\alpha_2^4 - 4\alpha_2^3 - 6\alpha_2^5 + \\ & 2\alpha_2^6 \\ A_2 =& (2\alpha_2^2 + 1 - 2\alpha_1 - 2\alpha_2 + \alpha_1^2 + 2\alpha_1\alpha_2)^2(2\alpha_2^2 + 1 - 2\alpha_1 - 2\alpha_2 + \alpha_2^2 \\ & + 2\alpha_1\alpha_2)^2(\alpha_1^2 + \alpha_2^2)^2 \end{aligned}$$

Maple-15 package is used for the mathematical derivation of Jeffreys prior.

5. Bayesian Analysis of the Model for m=3

The joint posterior distribution of the Rayleigh model parameters given data using the (2.3) and $p(\boldsymbol{\alpha})$ (prior distribution) is:

$$p(\alpha_i, \alpha_j|\boldsymbol{w}) = \frac{1}{K}\, p(\boldsymbol{\alpha}) \prod_{i<j=1}^{m=3} \frac{r_{ij}!}{w_{ij}!\,(r_{ij}-w_{ij})!}\, \frac{\alpha_i^{2w_{i.ij}} \alpha_j^{2w_{j.ij}}}{(\alpha_i^2+\alpha_j^2)^{w_{i.ij}+w_{j.ij}}}$$

where K is the normalizing constant, defined as:

$$K = \int_0^1 \int_0^{1-\alpha_i} p(\boldsymbol{\alpha}) \prod_{i<j=1}^{m=3} \frac{r_{ij}!}{w_{ij}!\,(r_{ij}-w_{ij})!}\, \frac{\alpha_i^{2w_{i.ij}} \alpha_j^{2w_{j.ij}}}{(\alpha_i^2+\alpha_j^2)^{w_{i.ij}+w_{j.ij}}} d\alpha_j d\alpha_i$$

The marginal posterior distribution of the Rayleigh model parameter α_i given data under Uniform prior using the (4.1) and Sec.5.1 is:

$$(5.1) \qquad p(\alpha_i|\boldsymbol{w}) = \frac{1}{K}\int_0^{1-\alpha_i} p_U(\boldsymbol{\alpha}) \prod_{i<j=1}^{m=3} \frac{r_{ij}!}{w_{ij}!\,(r_{ij}-w_{ij})!}\, \frac{\alpha_i^{2w_{i.ij}} \alpha_j^{2w_{j.ij}}}{(\alpha_i^2+\alpha_j^2)^{w_{i.ij}+w_{j.ij}}} d\alpha_j,$$

$$\alpha_i > 0,\ \sum_{i=1}^{m=3} \alpha_i = 1.$$

The marginal posterior distribution of the Rayleigh model parameter α_1 given data under Jeffreys prior using the (4.2) and Sec. 5.2 is:

$$(5.2) \qquad p(\alpha_i|\boldsymbol{w}) = \frac{1}{K}\int_0^{1-\alpha_i} p_J(\boldsymbol{\alpha}) \prod_{i<j=1}^{m=3} \frac{r_{ij}!}{w_{ij}!\,(r_{ij}-w_{ij})!}\, \frac{\alpha_i^{2w_{i.ij}} \alpha_j^{2w_{j.ij}}}{(\alpha_i^2+\alpha_j^2)^{w_{i.ij}+w_{j.ij}}} d\alpha_j,$$

$$\alpha_i > 0,\ \sum_{i=1}^{m=3} \alpha_i = 1.$$

The posterior distribution is not in closed form but can be used numerically using package like SAS.

For illustrative purposes, two real data sets (r_{ij}) of 5 and 30 respondents is collected from the students of the Quaid-i-Azam University Pakistan. These data sets comprise of the three different brands of cigarettes (Benson & Hedges (BH), Marlboro (ML) and Dunhill (DH)) which are commonly used among students. Bayesian analysis for the data sets in Table 1 is carried out using non-informative priors.

Table 1. Data of Cigarette Brands

Pairs	Data 1			Pairs	Data 2		
	$w_{i.ij}$	$w_{j.ij}$	r_{ij}		$w_{i.ij}$	$w_{j.ij}$	r_{ij}
(BH , ML)	1	4	5	(BH , ML)	11	13	24
(BH , DH)	2	3	5	(BH , DH)	12	15	27
(ML , DH)	4	1	5	(ML , DH)	16	11	27

5.1. Posterior Estimates. The posterior means are used as the estimates of the parameters. In the Table 2, the posterior means of the Rayleigh and the Bradley-Terry models are given.

Table 2. Posterior Means under Non-Informative Priors

Parameters	Data 1				Data 2			
	Bradley-Terry		Rayleigh		Bradley-Terry		Rayleigh	
	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys	Uniform
α_1	0.15965	0.18604	0.23427	0.23063	0.28907	0.29027	0.31080	0.31051
α_2	0.63030	0.57543	0.49461	0.49991	0.39659	0.39470	0.36483	0.36523
α_3	0.21005	0.23853	0.27112	0.26946	0.31435	0.31502	0.32437	0.32426

From the Table 2, it is concluded that the cigarette brand Marlboro may be ranked number one among the brands and commonly used by students. Dunhill is ranked number two. Benson & Hedges has the lowest rank. Further observes that ranking of the brands (of cigarettes) have same order under both the models and data sets using non-informative priors.

5.2. Graphs of the Marginal Posterior Distribution. The graphs of the marginal posterior distribution of the Bradley-Terry and the Rayleigh model using both data sets for non-informative priors are drawn below.

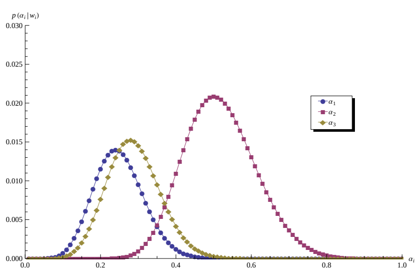


Figure 1. The Marginal Posterior Distributions for α_i of the Rayleigh Model using Uniform Prior (Data-1)

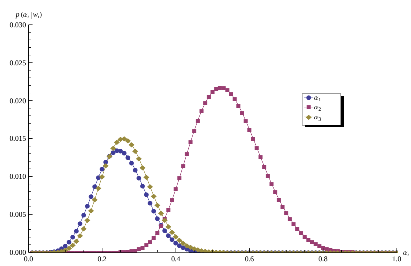


Figure 2. The Marginal Posterior Distributions for α_i of the Rayleigh Model using Jeffreys Prior (Data-1)

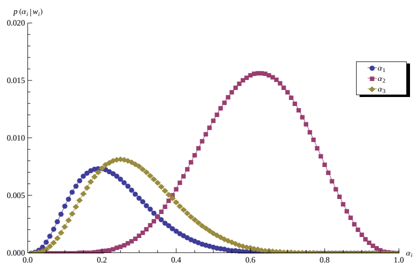


Figure 3. The Marginal Posterior Distributions for α_i of the Bradley-Terry Model using Uniform Prior (Data-1)

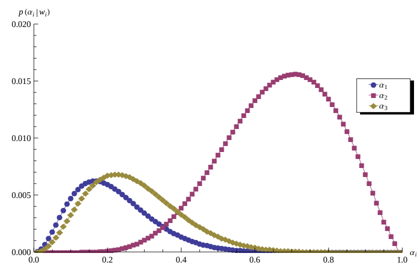


Figure 4. The Marginal Posterior Distributions for α_i of the Bradley-Terry Model using Jeffreys Prior (Data-1)

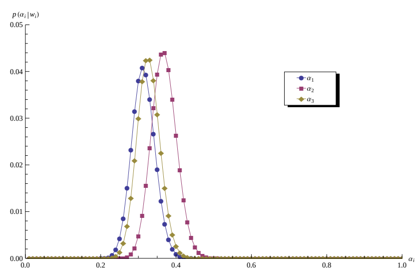


Figure 5. The Marginal Posterior Distributions for α_i of the Rayleigh Model using Uniform Prior (Data-2)

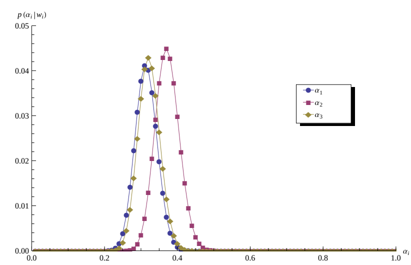


Figure 6. The Marginal Posterior Distributions for α_i of the Rayleigh Model using Jeffreys Prior (Data-2)

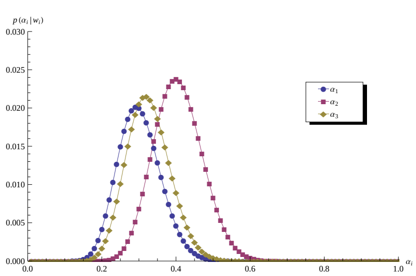


Figure 7. The Marginal Posterior Distributions for α_i of the Bradley-Terry Model using Uniform Prior (Data-2)

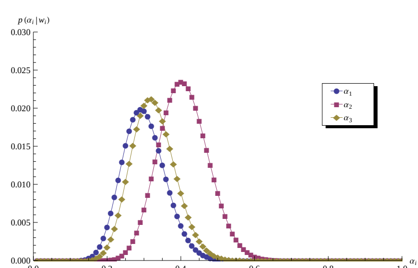


Figure 8. The Marginal Posterior Distributions for α_i of the Bradley-Terry Model using Jeffreys Prior (Data-2)

The Figures 1, 2, 3 and 4 have skewed marginal posterior distributions for the Rayleigh and the Bradley-Terry models under non-informative priors for the data set-1. Where as figures 5, 6, 7 and 8 have symmetrical marginal posterior distributions for the Rayleigh and the Bradley-Terry models under non-informative priors for the data set-2. Due to large data set shows symmetrical graphs.

5.3. Credible Intervals. The 95 % credible intervals are constructed for the Bradley-Terry and the Rayleigh models.

Table 3. 95% Credible Intervals under Non-Informative Priors

Parameters	Data-1			
	Bradley-Terry		Rayleigh	
	Jeffreys	Uniform	Jeffreys	Uniform
α_1	(0.11902, 0.20029)	(0.14547, 0.22662)	(0.20679, 0.26175)	(0.20296, 0.25830)
α_2	(0.56769, 0.69291)	(0.51878, 0.63207)	(0.45733, 0.53189)	(0.46271, 0.53712)
α_3	(0.16432, 0.25577)	(0.19459, 0.28247)	(0.24349, 0.29875)	(0.24151, 0.29741)
Parameters	Data-2			
	Bradley-Terry		Rayleigh	
	Jeffreys	Uniform	Jeffreys	Uniform
α_1	(0.26477, 0.31336)	(0.26623, 0.31431)	(0.29806, 0.32354)	(0.29770, 0.32332)
α_2	(0.37023, 0.42295)	(0.36881, 0.42060)	(0.35191, 0.37776)	(0.35222, 0.37823)
α_3	(0.29121, 0.33748)	(0.29216, 0.33789)	(0.31254, 0.33619)	(0.31236, 0.33616)

From the Table 3, it is observed that 95 % interval are narrower for the data-2. Further more it is concluded that the credible intervals for the Rayleigh model are narrower than the Bradley Terry model under non-informative priors.

5.4. Preference Probability. The term preference probability is used for the superiority of probability of T_i over T_j on some defined attribute or characteristic. Using the posterior means of the Rayleigh and the Bradley-Terry model provided in the Table 2, the preference probabilities are calculated using (2.1), (2.2), (3.1) and (3.2) presented in the Table 4.

Table 4. Preference Probabilities under Non-Informative Priors

$\phi_{i.ij}$	Data-1				Data-2			
	Bradley-Terry		Rayleigh		Bradley-Terry		Rayleigh	
	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys	Uniform
$\phi_{1.12}$	0.20210	0.24431	0.18323	0.17549	0.42159	0.42377	0.42054	0.41955
$\phi_{1.13}$	0.43184	0.43818	0.42747	0.42282	0.47905	0.47956	0.47865	0.47835
$\phi_{2.23}$	0.75004	0.70695	0.76895	0.77487	0.55784	0.55613	0.55850	0.55921

From the Table 4, it is perceived that the preference probabilities implies the same ranking order as the posterior means for both the models and data sets under non-informative priors.

5.5. Predictive Probability. The predictive probabilities is used to predict the future single preference of one treatment T_i over treatment T_j. It is denoted by $P_{i.ij}$ and defined as:

$$P_{i.ij} = \int_{\alpha_i=0}^{1} \int_{\alpha_j=0}^{1-\alpha_i} \phi_{ij}\ p(\alpha_i, \alpha_j|\boldsymbol{w})\ d\alpha_j\ d\alpha_i$$

Table 5. Predictive Probabilities under Non-Informative Priors

$P_{i.ij}$	Data-1				Data-2			
	Bradley-Terry		Rayleigh		Bradley-Terry		Rayleigh	
	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys	Uniform
$P_{1.12}$	0.20875	0.24877	0.20870	0.20078	0.42222	0.42434	0.42222	0.42128
$P_{1.13}$	0.43393	0.43960	0.43388	0.43003	0.47895	0.47945	0.47895	0.47865
$P_{2.23}$	0.74477	0.70352	0.74482	0.75080	0.55721	0.55556	0.55721	0.55788

The predictive probabilities are closed to the preference probabilities and favors the same ranking order for both the models and data sets under non-informative priors.

5.6. Bayesian Hypotheses Testing. In Bayesian analysis, the task of deciding between the hypotheses is conceptually more straightforward. One merely calculates the posterior probabilities and decides between hypotheses accordingly.

$$H_{ij} : \alpha_i \geq \alpha_j \quad Vs. \quad H_{ji} : \alpha_i < \alpha_j,$$

The posterior probability for the hypothesis H_{ij} is:

$$p_{ij} = \int_{\zeta=0}^{1} \int_{\eta=\zeta}^{(1+\zeta)/2} p(\zeta, \eta|w) d\eta d\zeta,$$

The posterior probability for the hypothesis H_{ji} is:

$$q_{ij} = 1 - p_{ij}$$

where $\eta = \alpha_i$ and $\zeta = \alpha_i - \alpha j$.
The decision rule for the hypotheses is based on Bayes factor. It is denoted by 'B' and the most general form of the Bayes factor can be described as follows.:

$$B = \frac{\text{Posterior odd ratios}}{\text{Prior odd ratios}}$$

The central notion of Bayes factor is that prior and posterior information should be combined in a ratio that provides evidence of one model specification over another.
It can be interpreted as the 'odds for H_{ij} to H_{ji} that are given by the data. [12] gives the following typology for comparing H_{ij} Vs. H_{ji}

$$\begin{aligned} B &\geq 1 && \text{support } H_{ij} \\ 10^{-0.5} \leq B &\leq 1 && \text{minimal evidence against } H_{ij} \\ 10^{-1} \leq B &\leq 10^{-0.5} && \text{substantial evidence against } H_{ij} \\ 10^{-2} \leq B &\leq 10^{-1} && \text{strong evidence against } H_{ij} \\ B &\leq 10^{-2} && \text{decisive evidence against } H_{ij} \end{aligned}$$

Table 6. Posterior Probability under Non-Informative Priors

Pairs	Data-1							
	Bradley-Terry				Rayleigh			
	Jeffreys		Uniform		Jeffreys		Uniform	
	p_{ij}	B	p_{ij}	B	p_{ij}	B	p_{ij}	B
$\alpha_1 > \alpha_2$	0.02872	0.02957	0.04139	0.04318	0.02872	0.02957	0.04139	0.04318
$\alpha_1 > \alpha_3$	0.33206	0.49714	0.34256	0.52105	0.33206	0.49714	0.34256	0.52105
$\alpha_2 > \alpha_3$	0.92737	12.76842	0.90968	10.07174	0.92737	12.76842	0.90968	10.07174
Pairs	Data-2							
	Bradley-Terry				Rayleigh			
	Jeffreys		Uniform		Jeffreys		Uniform	
	p_{ij}	B	p_{ij}	B	p_{ij}	B	p_{ij}	B
$\alpha_1 > \alpha_2$	0.15254	0.17999	0.15566	0.18436	0.14174	0.16515	0.14048	0.16344
$\alpha_1 > \alpha_3$	0.37300	0.59490	0.37428	0.59816	0.35536	0.55125	0.35479	0.54988
$\alpha_2 > \alpha_3$	0.75584	3.09567	0.75208	3.03356	0.74108	2.86219	0.74275	2.88727

The Bayes factor in the Table 6 signify substantial evidence against H_{12}, minimal evidence against H_{13} and H_{23} is supported for both the models and data sets under non-informative priors. The preference order of treatments is confirmed through testing of hypotheses.

5.7. Appropriateness of the Model. It is used to compare the discrepancies of the observed preferences among the expected preferences. The Chi-square test is used for the appropriateness of the models. The hypothesis is defined as:
H_0: The model is true for some values of $\boldsymbol{\alpha} = \alpha_0$
H_1: The model is not true for any values of the parameters.
where $\boldsymbol{\alpha}=\alpha_1, \alpha_2, ..., \alpha_m$ is the vector of the unknown parameters, $\alpha_i > 0$.
The χ^2 has the following form:

$$\chi^2 = \sum_{i<j}^{m} \left\{ \frac{(w_{ij} - \hat{w}_{ij})^2}{\hat{w}_{ij}} + \frac{(w_{ji} - \hat{w}_{ji})^2}{\hat{w}_{ji}} \right\}$$

with $(m-1)(m-2)/2$ degrees of freedom [4].
The expected number of preferences are obtained by the following form:

$$\hat{w}_{i.ij} = r_{ij} \frac{\alpha_i^2}{\phi_{ij}} \text{ and } \hat{w}_{j.ij} = r_{ij} \frac{\alpha_j^2}{\phi_{ij}}, \text{where } \phi_{ij} = \alpha_i^2 + \alpha_j^2.$$

w_{ij} and w_{ji} are the observed number of preferences from the data set given in the Table 1.

Table 7. Appropriateness of the Rayleigh Model

	Data-1				Data-2			
	Bradley-Terry		Rayleigh		Bradley-Terry		Rayleigh	
	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys	Uniform
χ^2	0.0873	0.2918	0.0519	0.0495	0.3947	0.3963	0.3945	0.3947
$P-value$	0.2324	0.4109	0.1802	0.1761	0.4702	0.4710	0.4701	0.4702

From the Table 7 , the values of χ^2 for the Rayleigh and the Bradley-Terry model for both data sets under non-informative priors have high P-values as P-values > 0.05 . It is evident from the P-values that both the models have good fit. Furthermore, the Rayleigh model is considered to be better fit for the small data set in the Table 1 than the Bradley-Terry model under both the non-informative priors.

6. Conclusion

A new model for paired comparison is developed, named as the Rayleigh paired comparison model. The Rayleigh paired comparison model is analyzed in the Bayesian framework using non-informative (Uniform and Jeffreys) priors. The results are also compared with the existing Bradley-Terry model. For the analysis, we use the data sets of the preferences of cigarette brands (Benson & Hedges, Marlboro and Dunhill) used by university students. It is noticed that the cigarette brand Marlboro is highly preferred among the students of university. Benson & Hedges is the lowest preferred. The graphs, preference probabilities, predictive probabilities and hypotheses testing also confirm the same preference. The credible intervals for the Rayleigh model are narrower than the Bradley Terry model under non-informative priors. The appropriateness of the models (the Bradley-Terry and the Rayleigh model) through χ^2- statistic suggests that the fit is good but the proposed Rayleigh model is better fit for small data set than the Bradley-Terry as the P-value of χ^2- statistic under the Rayleigh model is smaller than the Bradley-Terry model.

References

[1] Abbas,N. and Aslam, M. *Prioritizing the items through paired comparison models, a Bayesian approach*, Pakistan Journal of Statistics **25** (1), 59-69, 2009.

[2] Abbas, N., Aslam, M. and Hussain, Z. *Refining the judges' assessments of items in paired comparison experiments*, Hacetteppe Journal of Mathematics and Statistics **40** (6), 907-913, 2011.

[3] Agresti, A. *Analysis of ordinal paired comparison data*, Journal of the Royal Statistical Society. Series C (Applied Statistics) **41** (2), 287-297, 1992.

[4] Aslam, M. *Bayesian analysis for paired comparison models allowing ties and not allowing ties*, Pakistan Journal of Statistics - All Series **18** (1), 53-70, 2002.

[5] Aslam, M., Gilani, G. M., Hussain, D. and Ali, S. *Statistical analysis of the TM-model via Bayesian approaches*, Pakistan Journal of Operation Research **8** (4), 849-857, 2012.

[6] Bayes, T. *An essay toward problem solving in the doctrine of chaos*, reprinted in Biometrika, **45** (5), 293-315, 1958.

[7] Bradley, R. A. and Terry, M. E. *The method of paired comparisons*, Biometrika, **39** (3/4), 324-345, 1952.

[8] Buck, D.,Wakeling, I., Greenhoff, K., and Hasted, A. *Predicting paired preferences from sensory data*, Food Quality and Preference, **12** (5), 481-487, 2001.

[9] David, H. A. The Method of Paired Comparisons, 2nd Ed. (Charles Griffin & Comy Ltd., London, 1998).

[10] Henery, R. J. *An extension to the Thurstone-Mosteller model for chess*, Journal of the Royal Statistical Society. Series D (The Statistician), **41** (5), 559-567, 1992.

[11] Hunter, D. R. *Mm algorithms for generalized Bradley-Terry models*, The Annals of Statistics, **32** (1), 384-406, 2004.

[12] Jeffreys, H. Theory of probability, (Oxford University Press, Oxford UK., 1961).

[13] Kuk, A. Y. C. *Modelling paired comparison data with large numbers of draws and large variability of draw percentages among players*, Journal of the Royal Statistical Society. Series D (The Statistician), **44** (4), 523-528, 1995.

[14] Laplace, P. S. *Theorie analytique des probabilites*, Courcier, Paris.

[15] Mosteller, F. *Remarks on the method of paired comparison:I. The least squares solution assuming equal standard deviations and equal correlations*, Phychometrika , **16** (1), 3-9, 1951.

[16] Masarotto, G. and Varin, C. *The ranking lasso and its application to sport tournaments*,The Annals of Applied Statistics, *6* (4), 1949-1970, 2012.

[17] Stern, H. *A continuum of paird comparisons models*, Biometrtsika, *77* (2), 265-273, 1990.

[18] Thurstone, L. L. *A law of comparative judgment*, Psychological review, **34** (4), 273, 1927.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1553 – 1578

A data envelopment analysis based approach for target setting and resource allocation: application in gas companies

Azam Mottaghi*, Ali Ebrahimnejad †, Reza Ezzati ‡ § and Esmail Khorram ¶

Abstract

Resource allocation is an important application of data envelopment analysis and has been investigated by many researchers and managers from both economic and managerial field. In real managerial and economical decisions, situations often occur when extra productions of goods are utilized and the decision maker (DM) would like to determine the numbers of extra products that each unit can produced. In this paper, several methods based on the data envelopment analysis for resource allocation in such situations are introduced that can help the managers to make better decisions. The primary aim of this paper is to allocate resources such that the inefficient decision making units (DMUs) to become efficient as possible. For this aim, firstly several homogeneous units under the control of a central unit are considered and then the efficiency of each unit is determined. In addition, if the production of additional products seems logical, the DM wants to know how much of additional outputs should be produced by each unit such that the total outputs reach to a predetermined level. In this case the proposed algorithms determine quantities of the consumed input and produced output levels for each DMU to obtain the desirable output level. For using the whole power of system, the multi objective programming (MOP) problem has been used. A numerical example is given to show the solution process to improve the clarity of the proposed method. Finally, the real data of a gas company extracted from extant literature are used to demonstrate the proposed method.

*Department of Mathematics Karaj Branch Islamic Azad University Karaj, Iran , Email: az.mottaghi@gmail.com

†Department of Mathematics, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran, Email: a.ebrahimnejad@qaemiau.ac.ir and aemarzoun@gmail.com

‡Department of Mathematics Karaj Branch Islamic Azad University Karaj, Iran , Email: ezati@kiau.ac.ir

§Corresponding Author.

¶Department of Mathematics Amirkabir University of Technologhy Tehran, Iran , Email: eskhor@aut.ac.ir

Keywords: DEA, MOP problem, Efficiency, Resource allocation, Production power.

2000 AMS Classification: 90Bxx, 90B50.

Received : 21.01.2015 *Accepted* : 06.08.2015 *Doi* : 10.15672/HJMS.20158512893

1. Introduction

Data envelopment analysis (DEA) is a non-parametric frontier estimation methodology based on linear programming to measure the relative efficiency of a decision making unit (DMU) and provide DMUs with relative performance assessment on multiple inputs and outputs. It originated from Farell's seminal work [16] , was popularized by Charnes et al. [9], and has gained a wide range of applications measuring comparative efficiency. For instance, Amirteimori et al. [1] proposed a flexible slacks-based measure (FSBM) of efficiency in assessing UK higher education institutions. Ebrahimnejad and Tavana [13] showed used an interactive method for performance assessment in North Atlantic Treaty Organization (NATO) by establishing an equivalent relation between DEA and multi objective programming problem. Ebrahimnejad et al. [12] proposed a DEA model for banking with three stages. Maghbouli et al. [30] used the cooperative and non-cooperative game theories to assess the relative performance of 39 Spanish Airports based on a network DEA model with undesirable factors.

In empirical studies that examines scale economics and productive efficiency within a DEA framework, two of the most frequently used are resource allocation and target setting. Resource allocation is the setting of input- output levels for DMUs when the organization has limited input resources or output possibilities and plays a pivotal role in the management of corporations. Because of this, it has been an interesting topic to both business managers and researchers.

Several researchers have carried out investigations on resource allocation via DEA. Golany et al. [18] presented a five-step procedure for input resource allocation at an organizational level. Their procedure reduces to solving a linear program involving an objective function weighted according to DMU efficiencies. Their procedure does not compute output targets and only distributes input resources guided by current (weighted) DMU efficiencies. Their work uses the additive DEA model of Charnes et al. [8]. Golany and Tamir [19] presented a resource allocation model which simultaneously determines input and output targets based on maximizing total output. Their model is only applied in the case of a single output. For the multiple output case they suggest applying predetermined subjective weights to each output measure. Athanassopoulos [4] presented a goal programming model incorporating ideas from Thanassoulis and Dyson [32]. A significant feature of his work is that DMUs are linked together at the global level with respect to organizational input- output targets. In his model, proportional deviations from current input-output levels for each DMU, as well as proportional deviations from organizational input-output targets, are weighted together in a single linear objective which is to be minimized. In his work, the weights are specified by the DM. Thanassoulis [33] presented a paper dealing with the single input case. He presented a mixed-integer program to simultaneously cluster DMUs into k distinct sets and to determine a marginal resource level (MRL) for each output measure for each such cluster. MRLs were defined as the rate of (input) resource entitlement per unit of output. Once MRLs have been found, a logical numeric basis exists for future input resource allocation by DMs.

Thanassoulis [34] presented a paper concerned with estimating a single set of MRLs that apply to all DMUs in the single input case. He presented both the regression based method and the linear programming based method for such estimation. Both methods use a revised data set produced by replacing observed output levels for all DMUs by output levels that would have rendered each DMU DEA- efficient. Once this has been done, then MRLs (for each output measure) can be found: (a) in the regression based method from the regression coefficients associated with each output measure in the ordinary least-squares linear regression line (b) in the linear programming based method by solving a single linear program designed to ensure that the MRLs obtained will not enable any DMU to attain its DEA-efficient output levels using less resources than DEA suggests is needed for those levels. An efficiency based measure to enable a comparison to be made between alternative sets of MRLs is also presented. (Basso and Peccati [5]) introduced a dynamic programming algorithm to get optimal resource allocation with both minimum and maximum activation levels and fixed costs. Yan et al. [36] discussed a typical inverse optimization problem on the generalized DEA model to identify how to control or adjust the changes in the input and output such that the efficiency index of DMUs concerned is preserved. Beasley [6]developed a resource allocation model aiming to maximize the total efficiency of all DMUs. Lozano and Villa [28] and Lozano et al. [29] introduced the concept of centralized DEA models, which aim at optimizing the combined resource consumption by all units in an organization rather than considering the consumption by each unit separately. Asmild et al. [3] suggested modifying these centralized models to only consider adjustments of previously inefficient unit and showed how this new model formulation relate to a standard DEA model, namely as the analysis of the mean inefficient point. Fang [17] developed a new generalized centralized resource allocation model that extends and generalizes Lozano and Villa 's model [28] and Asmild et al.' s model [3] to a more general case. Hadi-Vencheh et al. [21] used an inverse DEA model for resource allocation in order to estimate increased requirements of the input vector when the output vector is increasing. Amireimoori and Mohaghegh Tabar [2] presented a DEA-based method for allocating fixed resources or costs across a set of decision making units and showed how output targets can be set at the same time as decisions are made about allocating input resources. Bi et al. [7] investigated the resource allocation and target setting for the organization consisting of production units, each of which has several parallel production lines. Wu et al. [35] proposed some new DEA models, which consider both economic and environmental factors in the allocation of a given resource. Li et al. [27] considered the model construction method for resource allocation considering undesirable outputs between different decision making units based on the DEA framework. They proposed some resource allocation models as a multiple objective linear problem which considers the input reduction, desirable output reduction and undesirable output reduction. Hosseinzadeh Lotfi et al. [24] proposed an allocation mechanism that is based on a common dual weights approach. Compared to alternative approaches, their model can be interpreted as providing equal endogenous valuations of the inputs and outputs in the reference set. Du et al. [11] used the cross-efficiency concept in DEA to approach cost and resource allocation problems. Hadi-Vencheh et al.[20] proposed a new method to find how much some inputs/outputs of each decision making unit (DMU) should be reduced such that the total efficiency of all DMUs after reduction being maximized.

It is worth noting that the assumptions that concern the unit's ability to change their input-output mix and efficiency are clearly the key factors affecting the results of the resource allocation. Although many valuable ideas have been proposed concerning these assumptions, the DMUs ability to change their input- output mix and efficiency has not be discussed thoroughly in the literature. In addition, the multiple criteria nature of the resource allocation problem has drawn only limited attention. DEA and multi-objective programming (MOP) can be used as tools in management control and planning. The structures of these two types of models have much in common but DEA is directed to assessing past performances as part of the management control function and MOP to planning further performances Cooper [10]. In order to find the most preferred allocation plan, Korhonen and Syrjanen [26] developed an interactive formal approach based on DEA and MOLP. Their approach concerns the modeling of units abilities to change their production. The authors considered two sets in their model: production possibility set and transformation possibility set. The first set describes all technically feasible production plans while the second describes the units ability to change its production within a planning period. They concluded that their approach can be applied to cases where DM controls only a part of the units.

Nasrabadi et al. [31] presented a model to investigate the resource allocation problem based on efficiency improvement. Their model uses parameters which are not necessarily unique in the case of alternative optimal solution. However, each optimal solution can be applied in the model to achieve performance improvement. This can be a shortcoming of their model, since finding all alternative optimal solutions and solving the model for each one seems unreasonable and time consuming.

Two kinds of factors which often have some relation with each other and play important roles in resource allocation models are economic factors and environmental factors. Economic factors usually refer to the desirable outputs generated in the production process, such as profit. Environmental factors usually refer to the undesirable outputs such as smoke pollution and waste. Jie and Qingxian [25] have proposed some new DEA models which consider not only economic but also environmental factors in the allocation of a given resource.

The purpose of this paper is to develop several algorithms based on MOP and DEA for resource allocation. Here, it is assumed that a central unit simultaneously controls all the units. If the production of additional products seems logical, the DM wants to know how much of additional outputs should be considered for each unit such that the total outputs reach a predetermined level. In the developed algorithms, the unit's abilities to change their production are modeled explicitly. The current input and output values are used to characterize a production possibility set. It is assumed that the units are able to modify their production plan within the production possibility set. These algorithms determine quantities of the consumed input and produced output levels for each DMU, such that the desirable output level is reached. Moreover, the number of efficient units is maximized, simultaneously.

It should be mentioned that in this paper we aim to allocate resources between units of a system which can be efficient but due to different problems such as: lack of proper supervision on the usage of the resources, wasting resources, inefficiency of workers and errors in production line, have become inefficient. These problems can be solved, therefore such units are first detected and then resource allocation is done among them. That is why it was not important to consider how inefficient the units are. The rest of this paper is organized as follows: In Section 2, first some fundamental models and definitions in DEA and MOP are reviewed and then the problem of resource allocation is stated. In Section 3, two algorithms to determine the input-output levels of each unit are introduced such that the maximum production capacity of the system can be used and the number of efficient units is maximized. Section 4, illustrates these algorithms using a numerical example. In Section 5, an empirical example of allocating experts to gas companies is presented to demonstrate the applicability of the proposed framework and exhibit the efficacy of the procedures. The paper ends with the conclusions and future research directions in Section 6.

2. Preliminaries and the statement of the problem

In this section, first a basic DEA model and also the concept of multiobjective linear programming problem are reviewed. Then, the main aim of problem under consideration is stated.

2.1. The input-oriented CCR model. Suppose there is a set of n, DMUs, $\{DMU_j, j = 1, 2, \cdots, n\}$ which produce multiple outputs $y_{rj}(r = 1, 2, \cdots, s)$ by utilizing multiple inputs $x_{ij}(i = 1, 2, \cdots, m)$. Let the inputs and outputs for DMU_j are $x_j = (x_{1j}, x_{2j}, \ldots, x_{mj})^t$ and $y_j = (y_{1j}, y_{2j}, \ldots, y_{sj})^t$, respectively. In addition $x_j \in R^m$, $y_j \in R^s$, $x_j > 0$ and $y_j > 0$, $j = 1, 2, \cdots, n$. We define the set of production possibility set (PPS), as $T = \{(x, y)|\ y \textit{ can be produced by } x\}$ and here we suppose that $T = T_{CCR}$ in which

$$T_{CCR} = \{((x_{1j}, x_{2j} \ldots, x_{mj}), (y_{1j}, y_{2j}, \ldots, y_{sj})) \mid x_{ij} \geq \sum_{j=1}^{n} x_{ij}\lambda_j,$$

$$y_{rj} \leq \sum_{j=1}^{n} y_{rj}\lambda_j,\ i : 1, ..., m;\ r : 1, ..., s; \forall j : \lambda_j \geq 0\}$$

The relative efficiency of the can be obtained by using the following linear programming (LP) model called input-oriented CCR primal model (Charnes et al. [9]):

$$\begin{aligned} &\min_{\lambda,\theta} \quad \theta \\ (2.1) \qquad &s.t \\ &\sum_{j=1}^{n} x_{ij}\lambda_j \leq \theta x_{io}, \quad i : 1, ..., m \\ &\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{ro}, \quad r : 1, ..., s \\ &\lambda_j \geq 0, \quad j = 1, \ldots, n. \end{aligned}$$

Model (2.1) measures the efficiency under a constant return to scale (RTS) assumption of technology. In this model, the vector variable $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_n)$ exhibits the intensity vector variable. The components of this vector represent the contribution of efficient units to constructing the projection point of inefficient units. It should be noted that in this model, the feasible region is non-empty and the optimal value θ_o satisfies $0 < \theta_o \leqslant 1$.

2.1. Definition. The optimal value θ_o^* of model (2.1) is called the efficiency index of DMU_o. If $\theta_o^* = 1$ then DMU_o is called (at least) weakly efficient unit. If $\theta_o^* \neq 1$ then it is called an inefficient unit.

Now, the index sets of efficient and inefficient units are defined as $E = \{j; \theta_j^* = 1\}$ and $I = \{j; \theta_j^* \neq 1\}$, respectively.

2.2. Multi objective programming. Here, we review some fundamentals of MOP problems and Min-ordering method for solving them, which will be used throughout the remainder of this paper (Ehrgott and Galperin, [14]; Ehrgott,[15]; Hosseinzadeh Lotfi et al. [22], [23]).
The MOP problem can be presented as follows:

$$(2.2) \qquad \begin{array}{ll} max & f(x) = (f_1(x), f_2(x), ..., f_p(x)) \\ s.t. & x \in S \end{array}$$

where is the feasible set of the optimization problem (2.2) and $f_k : S \to \mathbb{R}$ for $k = 1, ..., p$ are objective functions. The primary goal in MOP is to find the Pareto optimal (efficient) solutions and to help select the most preferred solution. In fact, a solution represented by a point in the decision variable space is a strictly Pareto optimal solution if it is not possible to move the point within the feasible region to improve an objective function value without deteriorating at least one of the other objectives. In other words, a feasible solution $\overline{x} \in S$ is called a Pareto optimal solution if there is no feasible solution $x \in S$ such that $f(x) \geq f(\overline{x})$ and $f(x) \neq f(\overline{x})$.

To generate any Pareto optimal solution, MOP(2.2) can be written as a following min-ordering problem:

$$(2.3) \qquad \max_{x \epsilon S} \ \min_{1 \leq k \leq p} \ f_k(x)$$

The min-ordering problem (2.3) can be solved as a single objective linear programming problem. If we introduce a variabl φ to the stand for $\min_{1 \leq k \leq p} \quad f_k(x)$ we can rewrite problem 2.3 as follows:

$$(2.4) \qquad \begin{aligned} &\max \ \ \varphi \\ &s.t. \ \ f_k(x) \ \geq \ \varphi; \quad k : 1, ..., p \\ &\qquad x \epsilon S \end{aligned}$$

2.3. Statement of the problem. Now, consider the following question: if the total output of a system wants to increase such that the efficient units remain unchanged, how much of the additional output must be produced by other DMUs?

To answer this question, we suppose the total output of a system, i.e. $Y = \sum_{j=1}^{n} Y_j$ should be increased from Y to β where $\beta \geqslant Y$ and $\beta \neq Y$.
Now, it is required to estimate the output vector Y_j^{new} for every $j \in I$ such that:

$$\sum_{j \in E} Y_j + \sum_{j \in I} Y_j^{new} = \beta$$

In this paper, we try to determine Y_j^{new} for every $j \in I$ such that the total consumed input is minimized by the system.

3. New Algorithms for Resource Allocation

In this section, several algorithms are proposed for solving the resource allocation problem stated in Section 2. In these algorithms, it is tried to determine the different levels of output by use of maximum production power of system. Note that there are systems that do not use their resources efficiently. In such systems, the DM is able to recognize some units as deficient, and if it is possible, change them. In fact we believe that the efficient units use their maximum power, and therefore they will be remained unchanged.

Consider a system where the deficient units are able to alter their input-output composition. Without loss of the generality, assume that that $Card\{I\} = q$, $DMU_1, DMU_2, ...DMU_q \in I$ and $I_o = I \setminus \{o\}$. In continue, several algorithms will be introduced for optimal resource allocation between such units. For achieving this, first the following model is solved to determine the minimum value of k-th output produced by inefficient unit o:

$$
\begin{aligned}
&\min \quad \widehat{y}_{ko}\\
(3.1)\quad &s.t\\
&\sum_{r\in\{1,2,\ldots,s\}\setminus\{k\}} u_r^o y_{ro} + u_k^o \widehat{y}_{ko} = 1\\
&\sum_{i=1}^{m} v_i^o x_{io} = 1,\\
&\sum_{r=1}^{s} u_r^o y_{rj} - \sum_{i=1}^{m} v_i^o x_{ij} \leqslant 0, \quad j \in I_o,\ E\\
&u_r^o \geq \varepsilon, v_i^o \geq \varepsilon, \quad \forall i, r,\\
&\widehat{y}_{kj} \geqslant y_{kj}, \quad \forall j \in I.
\end{aligned}
$$

In this model u_r^o and v_i^o are the variables of the problem. They represent the input and output weighs of the inefficient unit o, respectively. The constraints refer to the condition in which the inefficient unit o and the efficient unit j remain in the respective and the inefficient unit o is transformed into efficient ones.

Model (3.1) is a non-linear MOP problem. To convert this non-linear model into a linear one, let $p_{ko} = u_k^o \widehat{y}_{ko}$ for all $o \in I$. This leads to the following linear MOP problem:

$$
\begin{aligned}
&\min \quad \widehat{y}_{ko}\\
(3.2)\quad &s.t\\
&\sum_{r\in\{1,2,\ldots,s\}\setminus\{k\}} u_r^o y_{ro} + p_{ko} = 1, \quad o \in I\\
&\sum_{i=1}^{m} v_i^o x_{io} = 1, \quad o \in I\\
&\sum_{r=1}^{s} u_r^o y_{rj} - \sum_{i=1}^{m} v_i^o x_{ij} \leqslant 0, \quad j \in I_o,\ E\\
&u_r^o \geq \varepsilon, v_i^o \geq \varepsilon, \forall i, r\\
&\widehat{y}_{kj} \geqslant y_{kj}, \quad \forall j \in I\\
&p_{ko} \geqslant 0.
\end{aligned}
$$

3.1. Theorem. *Models (3.1) and (3.2) are equal to each other.*

Proof. Suppose $(u^o_{r\,r\neq k}, u^o{}_k, v_i{}^o, \widehat{y}_{ko})$ is the optimal solution of model (3.1), then $(u^o_{r\,r\neq k}, v_i^o,\ p_{ko} = u_k^o \widehat{y}_{ko}, \widehat{y}_{ko})$ is a feasible solution of model (3.2). Now we show that this solution is the optimal solution of model (3.2). By contradiction, suppose that $(u^o_{r\,r\neq k}, v_i^o,\ p_{ko} = u_k^o \widehat{y}_{ko}, \widehat{y}_{ko})$ is not the optimal solution of model (3.2) and there exists $(\widetilde{u}_r^o, r \neq k, \widetilde{v}_i^o,\ \widetilde{p}_{ko}, \widehat{\widehat{y}}_{ko})$ as the optimal solution of model (3.2) and $\widehat{\widehat{y}}_{ko} < \widehat{y}_{ko}$. Therefore, the solution $(\widetilde{u}_r^o, r \neq k,\ \widetilde{u}_k^o = \frac{\widetilde{p}_{ko}}{\widehat{\widehat{y}}_{ko}},\ \widetilde{v}_i^o, \widehat{\widehat{y}}_{ko})$ is a feasible solution of model (3.1). Since $\widehat{\widehat{y}}_{ko} < \widehat{y}_{ko}$ is in contrast with the optimality of $(u^o_{r\,r\neq k}, u_k^o, v_i^o, \widehat{y}_{ko})$, thus $(u^o_{r\,r\neq k}, v_i^o,\ p_{ko} = u_k^o \widehat{y}_{ko}, \widehat{y}_{ko})$ is also the optimal solution of model (3.2).

On the other hand, suppose $(u^o_{r\,r\neq k}, v^o_i,\ p_{ko}, \widehat{y}_{ko})$ is the optimal solution of model (3.2); then it is obvious that $(u^o_{r\,r\neq k},\ u^o_k = \frac{p_{ko}}{\widehat{y}_{ko}},\ v^o_i, \widehat{y}_{ko})$ is also a feasible solution of model (3.1). Lets suppose another solution like $(\overline{u}^o_{r\,r\neq k}, \overline{u}^o_k, \overline{v}^o_i, \overline{\widehat{y}}_{ko})$ as the optimal solution of model (3.1) so that $\overline{\widehat{y}}_{ko} < \widehat{y}_{ko}$. Therefore $(\overline{u}^o_{r\,r\neq k}, \overline{u}^o_k, p_{ko} = \overline{u}^o_k\overline{\widehat{y}}_{ko}, \overline{\widehat{y}}_{ko})$ is a feasible solution of model (3.2) and since $\overline{\widehat{y}}_{ko} < \widehat{y}_{ko}$, there is a contradiction with the previous solution. □

In a similar way, the model (3.2) gives the maximum value of output produced by inefficient unit o:

$$\text{(3.3)}\quad \begin{array}{ll} \max & \widehat{y}_{ko} \\ s.t & \\ \multicolumn{2}{l}{Constraints\ of\ model\ (3.2)} \end{array}$$

Suppose that $(\underline{y^*_{1o}}, \underline{y^*_{2o}}, ..., \underline{y^*_{so}})$ and $(\overline{y^*_{1o}}, \overline{y^*_{2o}}, ..., \overline{y^*_{so}})$and for all are optimal solutions of model (3.2) and model (3.3), respectively. We define the following index sets:

$$R_1 = \{r \in \{1,2,...,s\} \mid \beta_r - \sum_{j\in E} y_{rj} \leqslant \sum_{j\in I} \underline{y}^*_{rj}\}$$

$$R_2 = \{r \in \{1,2,...,s\} \mid \sum_{j\in I} \underline{y}^*_{rj} \leqslant \beta_r - \sum_{j\in E} y_{rj} \leqslant \sum_{j\in I} \overline{y}^*_{rj}\}$$

$$R_3 = \{r \in \{1,2,...,s\} \mid \beta_r - \sum_{j\in E} y_{rj} \geqslant \sum_{j\in I} \overline{y}^*_{rj}\}$$

In such situation, two cases may be occur: $R_3 = \phi$ or $R_3 \neq \phi$. These cases are investigated in the following sections, separately.

3.1. Resource allocation when $R_3 = \phi$. If $R_3 = \phi$ then the mentioned system produces the desired level of output without any need to extra resources. In this case, an algorithm is introduced to determine the output levels of each unit, such that the total output is equal to β. In such algorithm, the values of the extra output produced by inefficient units, is determined. The different levels of output are estimated in a way that the number of the inefficient units that are converted to efficient ones will be maximized.

In this algorithm the following symbols are applied:

K: The index set of outputs that amounts of their changes have been determined.

L: The iteration number of algorithm which has been run for the output under consideration.

S^L: The index set of units converted to inefficient unit in the iteration or previous iterations.

$J^i_{r'}$: The index set of units converted to inefficient unit in the iteration during the producing of output .

$O^{r'}$: The index set of units converted to inefficient unit during the producing of output .

$I^{r'}$: The index set of units that are able to change the output.

Algorithm1: Estimation output levels for inefficient units when $R_1 \cup R_2 \neq \phi$

Step 1: Set $K = \phi$ and $L = 1$.

Step 2: Assume that $r' = \arg max_{r \in R_1 \cup R_2 \setminus K}\{|\beta_r - \sum_{j\in E} y_{rj}|\}$.

Set $S^0 = J^0_{r'} = O^{r'} = \phi$. Define $Y^0_{r'} = \sum_{j\in I} \underline{y}^*_{r'j} + \sum_{j\in E} y'^j_r$ and

$$I^{r\prime} = \begin{cases} I; & if\ K = \phi\ or\ K \subset R_2 \\ \bigcup_{r\in K} O^{r\prime}; & if\ K \neq \phi \end{cases}$$

If $r\prime \in R_1$ go to Step 3; otherwise if $r\prime \in R_2$ go to Step 6.

Step 3: Since $r\prime \in R_1$, the level of output $r\prime$ is less than the sum of minimum $r\prime$-th outputs which could be generated by inefficient units. In this case, let $y_{r\prime j} = \underline{y}^*_{r\prime j}$ for $j \in I$ and compute $Y^{L,K}_{r\prime} = Y^{L-1}_{r\prime} - \underline{y}^*_{r\prime k} + y_{r\prime k}$ for every $k \in I^{r\prime}$.

Note that in this case for every $j \in I$, $y_{r\prime j}$ has been replaced with $\underline{y}^*_{r\prime j}$. Thus based on model (3.2) these DMUs are converted to efficient units.

By definition of $Y^{L,K}_{r\prime}$, we attempt to find those DMUs that by substituting $y_{r\prime j}$ instead of $\underline{y}^*_{r\prime j}$, the total output is reached to $\beta_{r\prime}$. If $Y^{L,K}_{r\prime} = \beta_{r\prime}$ then set $S^L = S^{L-1} \cup k$ and go to Step 4, otherwise; go to Step 5.

Step 4: For every $t \in S^L$, the $r\prime$-th output of DMUs is obtained as follows:

$$\widehat{y}_{r\prime j} = \begin{cases} y_{r\prime j}; & j \in E \\ \underline{y}^*_{r\prime j}; & j \in I^{r\prime} \setminus \{\cup_{i=0}^{L-1} J^i_{r\prime}, t\} \\ y_{r\prime j}; & j \in \{\cup_{i=0}^{L-1} J^i_{r\prime}, t\} \\ \underline{y}^*_{r\prime j}; & j \in I \setminus I^{r\prime} \end{cases}$$

Now set $O^{r\prime} := \{\cup_{i=0}^{L-1} J^i_{r\prime}, t\}$ and $K := K \cup \{r\prime\}$. If $K = \{1, 2, ..., s\}$ then stop, otherwise; if $R_1 \cup R_2 \neq \phi$ go to Step 2, else if $R_3 \neq \phi$ run Algorithm 2.
Step 5: Assume that $g_1 = argmin_{k\in I^{r\prime}}\{\mid \beta_{r\prime} - Y^{L,k}_{r\prime} \mid\}$. Consider the following cases:

Case1. $\beta_{r\prime} < Y^{L,g_1}_{r\prime}$: In this case set $Y^L_{r\prime} = Y^{L,g_1}_{r\prime}$ and consider the following sub cases:

Case1.1. $Y^L_{r\prime} < Y^{L-1}_{r\prime}$: In this case let

$$J_{r\prime}^{L} := J_{r\prime}^{L-1} \cup g_1, \quad O^{r\prime} := O^{r\prime} \cup \{\cup_{i=0}^{L-1} J_{r\prime}^{i}, g_1\}, \quad L := L+1, \quad K := K \cup \{r\prime\}$$

If $L > card(\cup_{r\in K} O^{r})$ then set $I^{r\prime} := I \setminus (\cup_{r\in K} O^{r})$ else set $I^{r\prime} := \cup_{r\in K} O^{r}$ and go to Step 3.

Case1.2. $Y_{r\prime}^{L} = Y_{r\prime}^{L-1}$: In this case compute $y_{r\prime j} + \beta_{r\prime} - Y_{r\prime}^{L,j}$ for all $j \in \{\cup_{i=0}^{L} J_{r\prime}^{i}\} \setminus \{\cup_{i=0}^{L-1} J_{r\prime}^{i}\}$.

If all of these values are positive then set $J_{r\prime}^{L} := J_{r\prime}^{L-1} \cup g_1$ and

$$\widehat{y}_{r\prime j} = \begin{cases} y_{r\prime j}; & j \in E \\ \{y_{r\prime j} + \beta_{r\prime} - Y_{r\prime}^{L,j} + (c-1)\underline{y}^{*}_{r\prime j}\} \times \dfrac{1}{c}; & j \in J_{r\prime}^{L} \\ \underline{y}^{*}_{r\prime j}; & j \in I^{r\prime} \setminus \{\cup_{i=0}^{L-1} J_{r\prime}^{i}\} \\ y_{r\prime j}; & j \in \{\cup_{i=0}^{L-1} J_{r\prime}^{i}\} \\ \underline{y}^{*}_{r\prime j}; & j \in I \setminus I^{r\prime} \end{cases}$$

where $c = Card\{J_{r\prime}^{L}\}$. Now set $L := L+1$, $O^{r\prime} := O^{r\prime} \cup \{\cup_{i=0}^{L-1} J_{r\prime}^{i}\}$ and $K := K \cup \{r\prime\}$. If $K = \{1, 2, ..., s\}$ then stop, else go to Step 2.

If all values of $y_{r\prime j} + \beta_{r\prime} - Y_{r\prime}^{L,j}$ are not positive then there exist $h \in J_{r\prime}^{L}$ such that $\{y_{r\prime h} + \beta_{r\prime} - Y_{r\prime}^{L,h} + (c-1)\underline{y}^{*}_{r\prime h}\} \times \dfrac{1}{c} = 0$. In this case put $\widehat{y}_{r\prime h} = y_{r\prime h}$ and remove DMU_h from the system and substitute $\beta_{r\prime}$ and $O_{r\prime}$ by $\beta_{r\prime} - y_{r\prime h}$ and $O_{r\prime} \cup \{h\}$, respectively. Finally run Algorithm1 for the new system.

Case2. $\beta_{r\prime} > Y_{r\prime}^{L,g_1}$: In this case set

$$\widehat{y}_{r\prime j} = \begin{cases} y_{r\prime j}; & j \in E \\ \underline{y}^{*}_{r\prime j}; & j \in I^{r\prime} \setminus \{\cup_{i=0}^{L-1} J_{r\prime}^{i}, g_1\} \\ y_{r\prime g_1} + \beta_{r\prime} - Y_{r\prime}^{L,g_1}; & j = g_1 \\ y_{r\prime j}; & j \in \{\cup_{i=0}^{L-1} J_{r\prime}^{i}\} \\ \underline{y}^{*}_{r\prime j}; & j \in I \setminus I^{r\prime} \end{cases}$$

Now, set $O^{r\prime} := O^{r\prime} \cup \{\cup_{i=0}^{L-1} J_{r\prime}^{i}, g_1\}$ and $K := K \cup \{r\prime\}$. If $K = \{1, 2, ..., s\}$ then stop, otherwise; if $R_1 \cup R_2 \neq \phi$ then set $L = 1$ go to Step 2, else if $R_1 \cup R_2 = \phi$ and $R_3 \neq \phi$ run Algorithm 2.

Step 6: Since $r' \in R_2$ then it is possible to determine the output level of all inefficient units such that they become efficient. To do this, the following model is solved. In this model without loss of the generality, we suppose that $\{o_1, o_2, ..., o_q'\} \in I^{r'}$ and $I_o^{r'} = I^{r'} \setminus \{o\}$.

$$
\begin{aligned}
&\max \quad \{\underline{y}^*_{r'o_1} + t_{r'o_1}(\overline{y}^*_{r'o_1} - \underline{y}^*_{r'o_1}),\ \underline{y}^*_{r'o_2} + t_{r'o_2}(\overline{y}^*_{r'o_2} - \underline{y}^*_{r'o_2}),\ ...,\ \underline{y}^*_{r'o'_q} + \\
&\qquad t_{r'o'_q}(\overline{y}^*_{r'o'_q} - \underline{y}^*_{r'o'_q})\} \\
&s.t \\
&\sum_{r\in K\setminus\{r'\}} u^o_r \widehat{y}_{ro} + u^o_{r'}\underline{y}^*_{r'o} + u^o_{r'} t_{r'o}(\overline{y}^*_{r'o} - \underline{y}^*_{r'o}) = 1; \quad o \in I^{r'} \\
&\sum_{i=1}^{m} v^o_i x_{io} = 1; \quad o \in I^{r'} \\
&\sum_{r\in K\setminus\{r'\}} u^o_r \widehat{y}_{rj} - \sum_{i=1}^{m} v^o_i x_{ij} \leqslant 0; \quad j \in I^{r'}_o \\
&\sum_{r\in K} u^o_r \widehat{y}_{rj} - \sum_{i=1}^{m} v^o_i x_{ij} \leqslant 0, \qquad j \in E \\
&\sum_{r\in I^{r'}} \underline{y}^*_{r'o} + t_{r'o}(\overline{y}^*_{r'o} - \underline{y}^*_{r'o}) + \sum_{j\in E} y_{r'j} = \beta_{r'} \\
&u^o_r \geqslant \varepsilon, \quad v^o_i \geqslant \varepsilon \quad \forall i, r \quad, \forall o \in I^{r'} \\
&\quad 0 \leqslant t_{r'o} \leqslant 1; \quad \forall j \in I^{r'}.
\end{aligned}
\tag{3.4}
$$

In this model $\widehat{y}_{rj}$ and $\widehat{y}_{ro}$ for $r \in K \setminus \{r'\}$, $j \in I^{r'}_o$, $o \in I^{r'}$ are obtained from the previous steps. Note that for some values of $\widehat{y}_{rj}$ and $\widehat{y}_{ro}$ which are not obtained from the previous steps set $\widehat{y}_{rj} := \underline{y}^*_{rj}$ and $\widehat{y}_{ro} := \underline{y}^*_{ro}$.

If $K = \phi$ then set $\widehat{y}_{rj} := \underline{y}^*_{rj}$ and $\widehat{y}_{ro} := \underline{y}^*_{ro}$ for all $j, o \in I$. Also, set $\widehat{y}_{rj} := y_{rj}$ for all $j \in E$. By substituting $P^o_{r'j} = u^o_{r'} t_{r'j}$ the non-linear multi objective model (3.4) is converted to the following linear multi objective model:

$$
\begin{aligned}
&\max \quad \{\underline{y}^*_{r'o_1} + t_{r'o_1}(\overline{y}^*_{r'o_1} - \underline{y}^*_{r'o_1}), \underline{y}^*_{r'o_2} + t_{r'o_2}(\overline{y}^*_{r'o_2} - \underline{y}^*_{r'o_2}), ..., \underline{y}^*_{r'o'_q} + \\
&t_{r'o'_q}(\overline{y}^*_{r'o'_q} - \underline{y}^*_{r'o'_q})\} \\
&s.t \\
&\sum_{r\in K\setminus\{r'\}} u^o_r \widehat{y}_{ro} + u^o_{r'}\underline{y}^*_{r'o} + P^o_{r'o}(\overline{y}^*_{r'o} - \underline{y}^*_{r'o}) = 1; \quad o \in I^{r'} \\
&\sum_{i=1}^{m} v^o_i x_{io} = 1; \quad o \in I^{r'} \\
&\sum_{r\in K\setminus\{r'\}} u^o_r \widehat{y}_{rj} - \sum_{i=1}^{m} v^o_i x_{ij} \leqslant 0; \quad j \in I^{r'}_o
\end{aligned}
\tag{3.5}
$$

$$\sum_{r\in K} u_r^o \widehat{y}_{rj} - \sum_{i=1}^{m} v_i^o x_{ij} \leqslant 0, \qquad j \in E$$
$$\sum_{r\in I^{r'}} \underline{y}^*_{r'o} + t_{r'o}(\overline{y}^*_{r'o} - \underline{y}^*_{r'o}) + \sum_{j\in E} y_{r'j} = \beta_{r'}$$
$$u_r^o \geqslant \varepsilon, \quad v_i^o \geqslant \varepsilon \quad \forall i, r \quad, \forall o \in I^{r'}$$
$$0 \leqslant t_{r'o} \leqslant 1; \quad \forall j \in I^{r'}$$
$$0 \leqslant P^o_{r'j} \leqslant u^o_{r'}; \quad \forall j \in I^{r'}$$

3.2. Theorem. *Models (3.4) and (3.5) are equal to each other.*

Proof. It is similar to the proof of Theorem 3.1. □

Model (3.5) can be rewritten as follows:

$$\max \ \min\{\underline{y}^*_{r'o_1} + t_{r'o_1}(\overline{y}^*_{r'o_1} - \underline{y}^*_{r'o_1}), \underline{y}^*_{r'o_2} + t_{r'o_2}(\overline{y}^*_{r'o_2} - \underline{y}^*_{r'o_2}), ..., \underline{y}^*_{r'o'_q} + t_{r'o'_q}(\overline{y}^*_{r'o'_q} - \underline{y}^*_{r'o'_q})\}$$

(3.6) $\quad$ *s.t Constraints of model* (3.5).

Let us assume

$$t = \min\{\underline{y}^*_{r'o_1} + t_{r'o_1}(\overline{y}^*_{r'o_1} - \underline{y}^*_{r'o_1}), \underline{y}^*_{r'o_2} + t_{r'o_2}(\overline{y}^*_{r'o_2} - \underline{y}^*_{r'o_2}), ..., \underline{y}^*_{r'o'_q} + t_{r'o'_q}(\overline{y}^*_{r'o'_q} - \underline{y}^*_{r'o'_q})\}$$

Thus, model (3.6) can be rewritten as follows:

$$\begin{aligned}
&\max \ t \\
(3.7) \quad &s.t \\
&t \leqslant \underline{y}^*_{r'o_1} + t_{r'o_1}(\overline{y}^*_{r'o_1} - \underline{y}^*_{r'o_1}), \\
&t \leqslant \underline{y}^*_{r'o_2} + t_{r'o_2}(\overline{y}^*_{r'o_2} - \underline{y}^*_{r'o_2}), \\
&\vdots \\
&t \leqslant \underline{y}^*_{r'o'_q} + t_{r'o'_q}(\overline{y}^*_{r'o'_q} - \underline{y}^*_{r'o'_q}) \\
&\textit{Constraints of model } (3.6).
\end{aligned}$$

Assuming that $t^*_{r'o}$ are the optimal solutions of model (3.7), the values of output r' for inefficient units are determined as follows:

$$\widehat{y}_{r'o} = \underline{y}^*_{r'o} + t^*_{r'o}(\overline{y}^*_{r'o} - \underline{y}^*_{r'o}); \quad \forall o \in I$$

Now, set $K := K \cup \{r'\}$. If $K = \{1, 2, ..., s\}$ then stop, otherwise; if $R_1 \cup R_2 \neq \phi$ then set $L = 1$ and go to Step 2, else if$R_1 \cup R_2 = \phi$ and $R_3 \neq \phi$ run Algorithm 2.

3.2. Resource allocation when $R_3 \neq \phi$. If $R_3 \neq \phi$ then the system cannot generate the desired output even if all the inefficient units are converted to efficient units. In such situations, the extra resources must be distributed among the units in order to produce the desired output. In such cases, after the running of Algorithm 1 and finding $\widehat{y}_{rj}$ for $j \in I$ and $r \in R_1 \cup R_2$, there would be $r' \in \{1, 2, ..., s\}$ such that its desired level of output, with the present resources, cannot be generated by the system; even if all of the deficient unit are converted to efficient ones. As a result, Algorithm 2 would be run in order to determine the different levels of input and output for the various units, in a way that it becomes possible to generate the desired level of output.

Algorithm 2: Estimation input-output levels for inefficient units when $R_3 \neq \phi$

Note that in this case, a new unit as DMU_{n+1} is added to system such that its outputs in every level $r \in \{1, 2, ..., s\} \setminus \{\widetilde{r}\}$, is equal to zero but $y_{r,n+1}$ is equal to $\alpha_{\widetilde{r}}$, After solving the correspond model, the minimum inputs that DMU_{n+1} is required to produce $\alpha_{\widetilde{r}}$ are determined. The $\gamma_{i,n+1}$ for $i : 1, 2, \cdots, m$ gives the minimum value of the i -th input consumed by system to produce $\alpha_{\widetilde{r}}$.

Step 1: Assume that

$$\alpha_{\widetilde{r}} = \arg max_{r' \in R_3} \{\beta_{r'} - (\sum_{j \in E} y_{r'j} + \sum_{j \in I} \overline{y}^*_{r'j})\}.$$

It is worth noting that $\alpha_{\widetilde{r}}$ is the maximum value of output that has not been generated yet.

Step 2: Solve the following model in order to determine the minimum required values of systems inputs for generating $\alpha_{\widetilde{r}}$:

$$\begin{aligned}
&\min \quad \{p_{1,n+1}, p_{2,n+1}, ..., p_{m,n+1}\} \\
&s.t \\
&\sum_{r=1}^{s} u_r^{n+1} \widehat{y}_{rj} - \sum_{i=1}^{m} v_i^{n+1} x_{ij} \leqslant 0; \quad j \in E \\
&\sum_{r \in R_1 \cup R_2} u_r^{n+1} \widehat{y}_{rj} + \sum_{r \in R_3} u_r^{n+1} \overline{y}^*_{rj} + \sum_{r \notin \cup_{j=1}^{3} R_j} u_r^{n+1} y_{rj} - \sum v_i^{n+1} x_{ij} \leqslant 0; \quad j \in I \\
&u_r^{n+1} \alpha_{\widetilde{r}} - \sum_{i=1}^{m} p_{i,n+1} = 0 \\
&u_r^{n+1} \geqslant \varepsilon, \quad v_i^{n+1} \geqslant \varepsilon, \quad p_{i,n+1} \geqslant \varepsilon, \quad r : 1, 2, ..., s \quad i : 1, 2, ..., m.
\end{aligned} \tag{3.8}$$

In model (3.8), the values of $\overline{y}^*_{rj}$ and $\widehat{y}^*_{rj}$ are given based on model (3.2) and Algorithm 1, respectively. If Algorithm (1) is not applied set $\widehat{y}_{rj} := y_{rj} (r : 1, 2, ..., s)$. Assuming$(p^*_{1,n+1}, p^*_{2,n+1}, ..., p^*_{m,n+1}, v_1^{n+1*}, v_2^{n+1*}, ..., v_m^{n+1*}, u_1^{n+1*}, u_2^{n+1*}, ..., u_m^{n+1*})$ is a strongly efficient solution of model (3.8), set $\gamma^*_{i,n+1} = \dfrac{{p_{i,n+1}}^*}{v_i^{n+1*}} (i : 1, ..., m)$. In this case the minimum value of the required resources for generating the total output is given by:

$$\Gamma = (\sum_{j=1}^{n} x_{1j} + \gamma^*_{1,n+1}, \sum_{j=1}^{n} x_{2j} + \gamma^*_{2,n+1}, ..., \sum_{j=1}^{n} x_{mj} + \gamma^*_{m,n+1})$$

Step 3: Now we determine how the extra resources $\gamma_{i,n+1}$ for $i : 1, 2, \cdots, m$ should be allocated between the various units and how much extra output should be generated by each unit such that the amount of total output become β . To do this, solve the following model:

$$
\begin{aligned}
& \max \quad \{\theta'_1, \theta'_2, ..., \theta'_\xi\} \\
& \max \quad \{\Delta y_{\tilde{r}1}, \Delta y_{\tilde{r}2}, ..., \Delta y_{\tilde{r}n}\} \\
(3.9) \qquad & s.t \quad \sum_{j=1}^{n} \lambda_j^k \widehat{y}_{rj} \geqslant \widehat{y}_{rk}, \quad r \in R_1 \cup R_2,\ k : 1, 2, ..., n \\
& \sum_{j \in E} \lambda_j^k y_{\tilde{r}j} + \sum_{j \in I} \lambda_j^k \overline{y}^*_{\tilde{r}j} + \xi^k_{\tilde{r}j} \geqslant \overline{y}^*_{\tilde{r}k} + \Delta y_{\tilde{r}k}, \quad k : 1, 2, ..., n \\
& \sum_{j=1}^{n} \lambda_j^k \overline{y}^*_{rj} \geqslant \overline{y}^*_{rk}, \quad r \in R_3 \setminus \tilde{r},\ k : 1, 2, ..., n \\
& \sum_{j=1}^{n} \lambda_j^k y_{rj} \geqslant y_{rk}, \quad r \notin R_1 \cup R_2 \cup R_3,\ k : 1, 2, ..., n \\
& \sum_{j=1}^{n} \lambda_j^k x_{ij} + \delta^k_{ij} \leqslant \theta'_k x_{ik} + \mu_{ik}, \quad k \in \bigcup_{r \in R_1} O^r,\ i : 1, 2, ..., m \\
& \sum_{j=1}^{n} \lambda_j^k x_{ij} + \delta^k_{ij} \leqslant \theta_k (x_{ik} + \Delta x_{ik}), \quad k \in \{1, 2, ..., n\} \setminus \bigcup_{r \in R_1} O^r,\ i : 1, 2, ..., m \\
& \sum_{j=1}^{n} (x_{ij} + \Delta x_{ij}) = \Gamma_{i,n+1}, \quad i : 1, 2, ..., m \\
& \sum_{j=1}^{n} \Delta y_{\tilde{r}j} = \alpha_{\tilde{r}} \\
& \theta'_k \geqslant \theta_k, \quad k \in \bigcup_{r \in R_1} O^r \\
& \xi^k_{\tilde{r}j} \geqslant 0, \quad \delta^k_{ij} \geqslant 0, \mu_{ik} \geqslant 0, \quad \lambda_j^k \geqslant 0, \quad lli, j, k.
\end{aligned}
$$

where $\xi^k_{\tilde{r}j} = \lambda_j^k \Delta y_{\tilde{r}j}$, $\delta^k_{ij} = \lambda_j^k \Delta x_{ij}$ and $\mu_{ik} = \theta'_k \Delta x_{ik}$ for $i : 1, 2, ..., m$ and $k, j : 1, 2, ..., n$. It is worth noting that in model (3.7) $\xi = card(\bigcup_{r \in R_1} O^r)$ if $R_1 \neq \phi$ and $\xi = card(I)$ if $R_1 = \phi$.

In model (3.9), θ_k for $k \in \bigcup_{r \in R_1} O^r$ is the efficiency value of DMU_k given by model (2.1) by putting $y_{rj} := \overline{y}^*_{rj}$ for $r \in R_3$ and $y_{rj} := \widehat{y}_{rj}$ for $r \in R_1 \cup R_2$. Set $R_3 := R_3 \setminus \{\tilde{r}\}$. If $R_3 = \emptyset$then stop; otherwise go to Step 1.

The MOP models are valuable and useful since when under the same constraints they can address several objectives. In some models of the presented paper such as (3.8) and (3.9) where we want to find out the least amount of inputs to produce the intended outputs, it saves us some calculation and time if we solve a m objective problem instead of solving m different problems (for each input).

4. Numerical Example

4.1. Example. Consider the data reported in Table 1 with five DMUs that consume two inputs to produce two outputs.

Table 1. The raw data

DMU	Input 1	Input 2	Output 1	Output 2
1	19	131	150	50
2	27	168	180	72
3	55	255	230	90
4	31	206	152	80
5	50	268	250	100

The efficiency score of each DMU obtained by model (2.1) is reported in Table 2.

Table 2. Efficiency Scores

DMU	1	2	3	4	5
Efficiency	1	1	0.83	0.97	0.87

Thus, we have $E = \{1, 2\}$ and $I = \{3, 4, 5\}$. We first determine the minimum amount of kth output which can be produced by inefficient units according to model (3.2). For example, the corresponding model to find the minimum amount of the first output for DMU_3 is as follows

$$
\begin{aligned}
&\min\ \widehat{y}_{13} \\
(4.1)\qquad &s.t \\
&90u_2^3 + p_{13} = 1; \\
&55v_1^3 + 255v_2^3 = 1; \\
&152u_1^3 + 80u_2^3 - 31v_1^3 - 206v_2^3 \leqslant 0; \\
&250u_1^3 + 100u_2^3 - 50v_1^3 - 268v_2^3 \leqslant 0; \\
&150u_1^3 + 50u_2^3 - 19v_1^3 - 131v_2^3 \leqslant 0; \\
&180u_1^3 + 72u_2^3 - 27v_1^3 - 168v_2^3 \leqslant 0; \\
&u_i^3 \geqslant 0.00001 \quad v_i^3 \geqslant 0.00001;\ i : 1, 2, \\
&\widehat{y}_{13} \geqslant 230;\ p_{13} \geqslant 0
\end{aligned}
$$

The optimal solution of model (4.1) is as follows:

$$
\begin{aligned}
&\widehat{y}_{13} = 291.9847 \quad p_{13}^1 = 1.00000000939939 \\
&u_2^3 = 0.00001 \quad u_1^3 = 0.3424837E - 02 \\
&v_1^3 = 0.00001 \quad v_2^3 = 0.3921569E - 02
\end{aligned}
$$

In addition, we can determine the maximum amount of kth output which can be produced by inefficient units according to model (3.3). For example, the corresponding model to find the maximum amount of the second output for DMU_4 is as follows:

$$(4.2)\quad \begin{aligned} &\max\ \widehat{y}_{24}\\ &s.t\\ &p_{24}+152u_1^4=1;\\ &31v_1^4+206v_2^4=1;\\ &230u_1^4+90u_2^4-55v_1^4-255v_2^4\leqslant 0;\\ &250u_1^4+100u_2^4-50v_1^4-268v_2^4\leqslant 0;\\ &150u_1^4+50u_2^4-19v_1^4-131v_2^4\leqslant 0;\\ &180u_1^4+72u_2^4-27v_1^4-168v_2^4\leqslant 0;\\ &u_i^4\geqslant 0.00001\ \ v_i^4\geqslant 0.00001;\ i:1,2,\\ &\widehat{y}_{24}\geqslant 80; p_{24}\geqslant 0 \end{aligned}$$

The optimal solution of model (4.2) is as follows:

$$\begin{aligned} &\widehat{y}_{24}=132.0593\ \ p_{24}=0.5769581016676\\ &u_2^4=0.4368932E-02\quad u_1^4=0.2783172E-02\\ &v_1^4=0.00001\quad v_2^4=0.4854369E-02 \end{aligned}$$

In a similar way it is possible to find the maximum and minimum amount of the all outputs for other units. The results are presented in Table 3.

Table 3. Maximum and minimum amount of outputs

DMU	Minimum output 1	Maximum output 1	Minimum output 2	Maximum output 2
3	291.985	12750	109.2858	150.5714
4	233.7211	300	82.6667	132.059
5	306.87	7500	114.857	138.519

Now suppose that decision makers want to increase the total output according to following case: The first output of total output is increased to $\beta_1 = 20000$ and the second output of total output is increased to $\beta_2 = 400$.

Now it is required to determine the amount of output produced by each DMU to reach the desired total output. Here, based on maximum and minimum amount of outputs of each DMU, the index sets of R_1, R_2 and R_3 are defined as follows:

$$R_1=\{r\in\{1,2\}\mid \beta_r-\sum_{j\in E}y_{rj}\leqslant\sum_{j\in I}\underline{y}^*_{rj}\}=\{2\}$$

$$R_2=\{r\in\{1,2\}\mid \sum_{j\in I}\underline{y}^*_{rj}\leqslant\beta_r-\sum_{j\in E}y_{rj}\leqslant\sum_{j\in I}\overline{y}^*_{rj}\}=\{1\}$$

$$R_3=\{r\in\{1,2\}\mid \beta_r-\sum_{j\in E}y_{rj}\geqslant\sum_{j\in I}\overline{y}^*_{rj}\}=\phi$$

Since $R_3 = \emptyset$, Algorithm 1 is used for resource allocation. The steps of this algorithm are given as follows:

Step 1: $K = \emptyset$ $L = 1$.

Step 2:

$$r' = \arg max_{r\in\{1,2\}}\{|\beta_r - \sum_{j\in E} y_{rj}|\} = \arg max_{r\in\{1,2\}}\{|20000-330|, |400-122|\} = 1$$

$$S^0 = J_1^0 = O^1 = \emptyset, \quad Y_1^0 = 1162.5761, \; I^1 = \{3, 4, 5\}$$

Since $r' = 1$ then Step 6 is applied.

Step 6: In this step model (4.3) is solved to determine the output level of all inefficient units.

$$
\begin{aligned}
&\max \quad t \\
(4.3) \qquad & s.t \\
& t \leqslant 291.985 + t_{13}(12750 - 291.985) \\
& t \leqslant 233.7211 + t_{14}(300 - 233.7211) \\
& t \leqslant 306.87 + t_{15}(7500 - 306.87) \\
& 291.985u_1^3 + 90u_2^3 + p_{13}^3(12750 - 291.985) = 1 \\
& 233.7211u_1^4 + 80u_2^4 + p_{14}^4(300 - 233.7211) = 1 \\
& 306.87u_1^5 + 100u_2^5 + p_{15}^5(7500 - 306.87) = 1 \\
& 55v_1^3 + 255v_2^3 = 1 \\
& 31v_1^4 + 206v_2^4 = 1 \\
& 50v_1^5 + 268v_2^5 = 1 \\
& 233.7211u_1^3 + 80u_2^3 - 31v_1^3 - 206v_2^3 \leqslant 0 \\
& 306.87u_1^3 + 100u_2^3 - 50v_1^3 - 268v_2^3 \leqslant 0 \\
& 291.985u_1^4 + 90u_2^4 - 55v_1^4 - 255v_2^4 \leqslant 0 \\
& 306.87u_1^4 + 100u_2^4 - 50v_1^4 - 268v_2^4 \leqslant 0 \\
& 291.985u_1^5 + 90u_2^5 - 55v_1^5 - 255v_2^5 \leqslant 0 \\
& 233.7211u_1^5 + 80u_2^5 - 31v_1^5 - 206v_2^5 \leqslant 0 \\
& 150u_1^3 + 50u_2^3 - 19v_1^3 - 131v_2^3 \leqslant 0 \\
& 180u_1^3 + 72u_2^3 - 27v_1^3 - 168v_2^3 \leqslant 0 \\
& 150u_1^4 + 50u_2^4 - 19v_1^4 - 131v_2^4 \leqslant 0 \\
& 180u_1^4 + 72u_2^4 - 27v_1^4 - 168v_2^4 \leqslant 0 \\
& 150u_1^5 + 50u_2^5 - 19v_1^5 - 131v_2^5 \leqslant 0 \\
& 180u_1^5 + 72u_2^5 - 27v_1^5 - 168v_2^5 \leqslant 0 \\
& 291.985 + t_{13}(12750 - 291.985) + 233.7211 + t_{14}(300 - 233.7211) + \\
& 306.87 + t_{15}(7500 - 306.87) + 330 = 20000 \\
& u_i^j \geqslant 0.00001 \quad v_i^j \geqslant 0.00001; \; i : 1, 2, j : 3, 4, 5 \\
& 0 \leqslant t_{1k} \leqslant 1; k : 3, 4, 5 \\
& p_{1k}^k \geqslant 0; k : 3, 4, 5
\end{aligned}
$$

The optimal solution of model (4.3) is as follows:

$t = 300.0000 \quad t_{13} = 0.9470239 \quad t_{14} = 1.000000 \quad t_{15} = 0.9694121$

$u_{13} = 0.8271283E-04 \quad u_{23} = 0.00001 \quad u_{14} = 0.333333E-02 \quad u_{24} = 0.00001$

$u_{15} = 0.1373631E-03 \quad u_{25} = 0.00001 \quad v_{13} = 0.00001 \quad v_{23} = 0.3921569E-02$

$v_{14} = 0.3225806E-01 \quad v_{24} = 0.00001 \quad v_{15} = 0.00001 \quad v_{25} = 0.3731343E-02$

$p_{13}^3 = 0.000078331026 \quad p_{14}^4 = 0.0033333 \quad p_{15}^5 = 0.00013316451233$

According to optimal solution of model (4.3) and according to Equation $\widehat{y}_{r'o} = \underline{y}^*_{r'o} + t^*_{r'o}(\overline{y}^*_{r'o} - \underline{y}^*_{r'o})$; $\forall o \in I$, the output levels of all inefficient units can be determined in this case. The results are presented in Table 4.

Table 4. Resource allocation of case 1

DMU:	1	2	3	4	5
*Output*1:	150	180	12090.0229515585	300	7279.977258873

Now, we set $K = \{1\}$. Since $R_1 \cup R_2 = \{1, 2\} \neq \emptyset$ thus we go to Step 2.

Step 2:

$$r' = \arg max_{r\in\{1,2\}}\{|\beta_r - \sum_{j\in E} y_{rj}|\} = \arg max_{r\in\{2\}}\{|400 - 122|\} = 2$$

$$S^0 = J_2^0 = O^2 = \emptyset, \quad Y_2^0 = 428.8095, \; I^2 = \{3, 4, 5\}$$

Since $r' \in R_1$ thus we go to Step 3.

Step 3: we set $y_{23} = 109.2858$ $y_{24} = 82.6667$, $y_{25} = 114.857$ and compute $Y_2^{1,k} = Y_2^0 - \underline{y}^*_{2k} + y_{2k}$ for every $k \in I^2$. We have:

$$Y_2^{1,3} = 428.8095 - 109.2858 + 90 = 409.5237$$
$$Y_2^{1,4} = 428.8095 - 82.6667 + 80 = 426.1428$$
$$Y_2^{1,5} = 428.8095 - 114.857 + 100 = 413.9525$$

Since $Y_2^{1,k} \neq 400$ for every $k \in I$ thus we go to Step 5.

Step 5: $g_1 = argmin_{k\in I^2}\{| \beta_2 - Y_2^{1,k} |\} = argmin\{9.5237, 26.1428, 13.9525\} = 3$. Since $Y_2^{1,3} = 409.5237 > 400 = \beta_2$ therefore we set $Y_2^1 := 409.5237$ and because of $Y_2^1 := 409.5237 < 428.8095 = Y_2^0$ let:

$$J_2^1 = J_2^0 \cup g_1 = \{3\}, \; O^2 := O^2 \cup \{\bigcup_{i=0}^{L-1} J_2^i, g_1\} = \{3\}, \; L := L + 1, \; K = K \cup \{2\}$$

Since $L = 2 > 0 = card(\bigcup_{r\in\{1\}} O^1)$ then we set $I^2 := I \setminus (r \in \{1\}O^1) = \{3, 4, 5\}$ and go to Step3.

Step 3: By computing $Y_2^{2,k} = Y_2^1 - \underline{y}^*_{2k} + y_{2k}$ for every $k \in I^2$ we have:

$$Y_2^{2,3} = 409.5237 - 109.2858 + 90 = 390.2379$$
$$Y_2^{2,4} = 409.5237 - 82.6667 + 80 = 406.857$$
$$Y_2^{2,5} = 409.5237 - 114.857 + 100 = 394.6667$$

Since $Y_2^{2,k} \neq 400$ for every $k \in I$ thus we go to Step 5.

Step 5: $g_1 = argmin\{9.7621, 6.857, 5.3333\} = 5$.
$Y_2^{2,5} = 394.6667 < 400 = \beta_2$. Therefore we consider Case 2.

$$\widehat{y}_{2j} = \begin{cases} y_{2j}; & j \in E \\ \underline{y}^*_{r\prime j}; & j \in I^2 \setminus \{\cup_{i=0}^1 J_2^i, 5\} \\ y_{25} + \beta_2 - Y_2^{2,5}; & j = 5 \\ y_{2j}; & j \in \{\cup_{i=0}^1 J_2^i\} \\ \\ \underline{y}^*_{2j}; & j \in I \setminus I^2 \end{cases}$$

and set $O^2 := O^2 \cup \{\bigcup_{i=0}^1 J_2^i, 5\}$ and $K = K \cup \{2\}$. Since $K = \{1, 2\}$ then the process is stopped. The results are presented in Table 5.

Table 5. Results of Algorithm 1 for case 1

DMU:	1	2	3	4	5
*Output*2:	50	72	90	82.6667	105.3333

4.2. Example. Application in Gas Companies

In this section, we illustrate the resource allocation discussed in this paper with the analysis of gas companies activity. This example is taken from Amireimoori and Mohaghegh Tabar [2]. The data set consists of 20 gas companies located in 18 regions in Iran. The data for this analysis are derived from operations during 2005. There are six variables from the data set as inputs and outputs in this example. Inputs include capital (x_1), number of staff (x_2), and operational costs (excluding staff costs) (x_3) and outputs include number of subscribers (y_1), length of gas network (y_2) and the sold-out gas income (y_3). Table 6 contains a listing of the original data. In this example, the initial capital, number of the staff and the operation costs are considered as the resources while the number of gas subscribers, the length of gas network and the income from gas distribution are considered as the products.

Suppose that the DM wants to increase the number of subscribers to 9500000. How much of the additional output must be produced by each DMU? We apply the algorithms discussed in this paper to answer this question.
According to model (3.3) the maximum value of the first output which should produce by inefficient units is equal to 572609.81 numbers. Thus we have:

$$\sum_{j \in I} \overline{y}^*_{1j} + \sum_{j \in E} y_{1j} = 572609.81 + 349061 = 921670.81$$

Note that $\beta_1 = 9500000 > 921670.81 = \sum_{j \in I} \overline{y}^*_{1j} + \sum_{j \in E} y_{1j}$ therefore $y_1 \in R_3$.
Thus, Algorithm 2 is used to determine the additional output produced by each DMU

Table 6. Data and efficiency scores for Iranian gas companies

DMU	x_1	x_2	x_3	y_1	y_2	y_3	θ
DMU_1	124313	129	198598	30242	565	61836	1
DMU_2	67545	117	131649	14139	153	46233	0.7106
DMU_3	47208	165	228730	13505	211	42094	0.9015
DMU_4	43494	106	165470	8508	114	44195	0.5977
DMU_5	48308	141	180866	7478	248	45841	1
DMU_6	55959	146	194470	10818	230	136513	1
DMU_7	40605	145	179650	6422	127	70380	0.7044
DMU_8	61402	87	94226	18260	182	36592	1
DMU_9	87950	104	91461	22900	170	47650	1
DMU_{10}	33707	114	88640	3326	85	13410	0.5235
DMU_{11}	100304	254	292995	14780	318	79883	0.6679
DMU_{12}	94286	105	98302	19105	273	32553	1
DMU_{13}	67322	224	287042	15332	241	172316	0.9579
DMU_{14}	102045	104	18082	155514	441	30004	0.9939
DMU_{15}	177430	401	528325	77564	801	201529	1
DMU_{16}	221338	1094	1186905	44136	803	840446	1
DMU_{17}	267806	1079	1323325	27690	251	832616	0.9510
DMU_{18}	160912	444	648685	45882	816	251770	1
DMU_{19}	177214	801	909539	72676	654	341585	1
DMU_{20}	146325	686	545115	19839	177	341585	0.8911

to increase the number of subscribers to 9500000. Since $R_3 = \{1\}$ we have:

$$\alpha_{\tilde{r}} = \arg max_{\tilde{r}\in R_3}\{\beta_{\tilde{r}} - (\sum_{j\in E} y_{\tilde{r}j} + \sum_{j\in I} \overline{y}^*_{\tilde{r}j})\} = 9500000 - 921670.81 = 8578329.19$$

Now, model (3.8) is solved. The minimum value of the required resources for generating 9500000 numbers of subscribers is equal to:

$$\begin{aligned}\Gamma &= (\sum_{j=1}^{20} x_{1j} + \gamma^*_{1,21}, \sum_{j=1}^{20} x_{2j} + \gamma^*_{2,21}, \sum_{j=1}^{20} x_{3j} + \gamma^*_{3,21})\\ &= (2125473 + 20000, 6446 + 35000, 7529507 + 42000)\end{aligned}$$

Regarding the results from the model (3.3), if we want to increase the number of the gas subscribers to 9500000, we need more initial resources. In this case, 20000 must be added to the initial capital, 35000 people should be added to the staff and the operation costs should increase by 42000. Now we solve model (3.9) for determining how the extra resources should be allocated between the various units and how much extra output should be generated by each unit. The obtained results are reported in Table7.

The results of the model show that among the inefficient units, units 17, 20 and 11 have higher positions in terms of both the inputs received and the outputs produced. Units 20, 14 and 10 have exactly the same position in terms of the inputs received and the outputs produced. In fact, the efficiency scores of units 20 and 14 are close to one and thus were able to produce more when they received more. On the other hand, since unit 10 has a low efficiency score, therefore it received fewer inputs and produced fewer outputs. Since

Table 7. Results of resource allocation for Iranian gas companies

DMU	x_1^*	x_2^*	x_3^*	y_1^*	$\sum_{r=2}^{3} \overline{y}_{rj} + y_{1j}^*$	$\sum_{i=1}^{3} x_{ij}^*$
DMU_1	124313	179.521	198598	271767.1	3341168.1	323090.521
DMU_2	68537.485	117	133781.552	111893.18	158279	202436.037
DMU_3	47208	165	230841.324	91358.74	133663.74	278214.324
DMU_4	44486.432	106	167599.944	100001	144310	210392.426
DMU_5	49300.485	141	182998.552	74484.49	120573.49	232440.037
DMU_6	56951.487	146	196602.764	77782.15	214525.15	253700.251
DMU_7	42944.011	145	181848.356	77700.97	148207.97	224937.367
DMU_8	62392.828	87	94226	94800.8	131574.8	156705.8275
DMU_9	88943.086	104	91461	86418.3	134238.3	180508.0855
DMU_{10}	33898.825	2645.827	94160.634	18992.4	32487.4	130705.2855
DMU_{11}	102944.869	32333.56	294488.088	597801.4	678002.4	429766.517
DMU_{12}	95278.485	105	100434.552	86111.49	118937.49	195818.0374
DMU_{13}	68318.133	224	294001.984	93194.32	265751.32	362544.1168
DMU_{14}	103037.485	104	157646.552	94201.23	124646.23	260788.0374
DMU_{15}	178422.485	401	528325	144570.49	346900.49	707148.4854
DMU_{16}	222330.558	1430.849	1189036.775	111330.71	952579.71	1412799.182
DMU_{17}	268788.176	1079	1325500.885	164111.23	996978.23	1595501.936
DMU_{18}	161853	444	651035.936	6896791	7149377	813332.936
DMU_{19}	178206.485	801	911671.552	139682.49	481921.49	1090679.037
DMU_{20}	147317.485	687.241	547247.552	167006.49	508768.49	695252.2784

unit 13 has a higher efficiency score among the inefficient units, it received more inputs compared to inefficient units 4 and 2 but produced less outputs compared to them; thus unit 13 has become efficient.
According to the optimal resource allocation plan, company 17 will receive more resource allocation in comparison with other companies. On the other hand, company 10 will receive the least allocated resource. Also, as Table 7 indicates, the most value of output 1 is set for company 18, whereas the least one is set for company 10. It is to be noted that the company 10 will receive the least allocated resource and so the least target is set for this company. For increasing the number of subscribers to 9500000, the first output of inefficient units will be increased to

$$\overline{y}_{1j}^* + \Delta y_{1j}^*; \quad j \in I = \{2, 3, 4, 7, 10, 11, 13, 14, 17, 20\}$$

where $\overline{y}_{1j}^*$ and Δy_{1j}^* are optimal solutions of model (3.3) and model (3.8), respectively. The efficient units i.e. DMU_j $j \in \{1, 5, 6, 8, 9, 12, 15, 16, 18, 19\}$ increase its first output to Δy_{1j}^* where are obtained from model (3.3).
The values of output 2 and output 3 of the DMUs are not changed but the inputs of all DMUs are increased as followes:

$$(\sum_{j=1}^{20} x_{1j}^*, \sum_{j=1}^{20} x_{2j}^*, \sum_{j=1}^{20} x_{3j}^*) = (2145473, 41446, 7571507)$$

5. Conclusion

In this paper, some algorithms for resource allocation are proposed. These algorithms help the DM in determining the input-output levels of each DMU, when the production of additional products seems to be desirable. DEA and MOP are applied in these algorithms. In fact these algorithms will allocate the resources between the units in a way that their maximum power will be applied for production. As an advantage of the method, it can be mentioned that it can be easily run, does not have complicated calculations and of course it tries to maximize the number of the efficient units. In this study, the data were considered as real, however, the algorithm can be expanded to the situations in which the data are of interval or fuzzy type. These cases will be investigated in future studies.

There are a number of challenges involved in the proposed research. These challenges provide a great deal of fruitful scope for future research. The practicality of this model can be further enhanced by developing the proposed framework into a decision support system to reduce the computation time and effort. Another future research direction, which could be an area of theoretical study, is extending the proposed method under a fuzzy environment.

The proposed algorithm can also be used to solve transportation problems, find the shortest route, obtain the maximum flow in a network, allocate people to jobs and etc. For example in transportation problems, in order to reduce the transfer cost up to a certain amount or to set the profit of transferring goods to a predetermined level, it is possible to consider routes as the DMUs and the goods as the resources allocated to each route. Finally, extending the proposed technique for resource allocation of the two-stage systems is an interesting research work. We hope that our study can inspire others to pursue further research.

Acknowledgements

The authors would like to thank the anonymous reviewers for their helpful and constructive comments that greatly contributed to improving the final version of the paper.

References

[1] Amirteimoori, A., Emrouznejad, A., Khoshandam, L. *Classifying flexible measures in data envelopment analysis: A slack-based measure*, Measurement **46**, 4100-4107, 2013.

[2] Amirteimoori, A. and Tabar, M.M. *Resource allocation and target setting in data envelopment analysis*, Expert Systems with Applications **37** (4), 3036-3039, 2010.

[3] Asmild, M., Paradi, J.C., Pastor, J.T. *Centralized resource allocation BCC models*, Omega **37**, 40-49, 2009.

[4] Athanassopoulos, A.D. *Goal programming and data envelopment analysis (GoDEA) for target-based multi-level planning: Allocating central grants to the Greek local authorities*, European Journal of Operational Research **87**, 535-550, 1995.

[5] Basso, A. and Peccati, L.A. *Optimal resource allocation with minimum activation levels and fixed costs*, European Journal of Operational Research **131** (3), 536-549, 2001.

[6] Beasley, J.E. *Allocating fixed costs and resources via data envelopment analysis*, European Journal of Operational Research **147** (1), 198-216, 2003.

[7] Bi, G., Ding, J., Luo, Y., Liang, L. *Resource allocation and target setting for parallel production system based on DEA*, Applied Mathematical Modelling **35**, 4270-428, 2011.

[8] Charnes, A., Cooper, W.W., Golany, B., Seiford, L., Stutz, J. *Foundations of data envelopment analysis for Pareto Koopmans efficient empirical production functions*, Journal of Econometrics **30**, 91-107, 1985.

[9] Charnes, A., Cooper, W.W. Rhodes, E. *Measuring the efficiency of decision making units*, European Journal of Operational Research **2**(6), 429-444, 1978.

[10] Cooper, W.W. *Origins, uses and relations between goal programming and data envelopment analysis*, In MCDM 2004 whistler, BC Canada 6-11, 2004.

[11] Du, J., Cook, W.D., Liang, L., Zhu. *Fixed cost and resource allocation based on DEA cross-efficiency*, European Journal of Operational Researc, http://dx.doi.org/10.1016/j.ejor.2013.10.002 , 2013.

[12] Ebrahimnejad, A., Tavana M., HosseinzadehLotfi, F., Shahverdi, R., Yousefpour M. *A three-stage Data Envelopment Analysis model with application to banking industry*. Measurement **49**, 308-219, 2014.

[13] Ebrahimnejad, A. and Tavana M. *An Interactive MOLP Method for Identifying Target Units in Output-Oriented DEA Models: The NATO Enlargement Problem*. Measurement **52**, 124-134, 2014.

[14] Ehrgott, M., Galperin, E. *Min-Max formulation of the balance number in multi-objective global optimization*, Computers and Mathematics with applications **44**(7), 899-907, 2002.

[15] Ehrgott, M. *Multicriteria Optimization*. 2nd edn. Springer, Berlin, 2005.

[16] Farell, M. *The measurement of productive efficiency*, Journal of the Royal Statistical Society, Series A (Gernal) **120**, 235-281, 1957.

[17] Fang, L. *A generalized DEA model for centralized resource allocation*, European Journal of Operational Research **228**, 405-412, 2013.

[18] Golany, B. and Phillips, F.Y. and Rousseau, J.J. *Models for improved effectiveness based on DEA efficiency results*, IIE Transactions **25**, 2-10, 1993.

[19] Golany, B., Tamir, E. *Evaluating efficiency-effectiveness equality trade-offs: A data envelopment analysis approach*, Management Science **41**, 1172-1184, 1995.

[20] Hadi-Vencheh, A., Ghelej Beigi, Z., Gholami, K. *On the input/output reduction in efficiency measurement*, Measurement **50**, 244-249, 2014.

[21] Hadi-Vencheh, A., Foroughi A., Soleimani-damaneh, M. *A DEA model for resource allocation*, Economic Modelling **25**(5), 983-993, 2008.

[22] Hosseinzadeh Lotfi F., Jahanshaloo G.R., Ebrahimnejad A, Soltanifar M, Manosourzadeh S.M. *Target setting in the general combined-oriented CCR model using an interactive MOLP method*. Journal of Computational and Applied Mathematics **234**, 1-9, 2010.

[23] Hosseinzadeh Lotfi, F., Jahanshaloo, G.R., Soltanifar, M., Ebrahimnejad, A., Manosourzadeh, S.M. *.Relationship between MOLP and DEA based on output-orientated CCR dual model*. Expert Systems with Applications **37**(6), 4331, 4336, 2010.

[24] Hosseinzadeh Lotfi, F., Hatami-Marbini, A., Agrell, P.J., Aghayi, N., Gholami, K. *Allocating fixed resources and setting targets using a common-weights DEA approach*, Computers and Industrial Engineering **64**, 631-640, 2013.

[25] Jie, W., Qingxiam, A., Shujahat, A., Liang, L. *DEA based resource allocation considering environmental factors*, Mathematical and Computer Modelling **58**, 1128-1137, 2011.

[26] Korhonen, P., Syrjanen, M. *Resource allocation based on efficiency analysis*, Management Science **50**,1134-1144, 2004.

[27] Li, H., Yangc, W., Zhoud, Z., Huange, C. *Resource allocation models construction for the reduction of undesirable outputs based on DEA methods*, Mathematical and Computer Modelling **58**, 913-926, 2013.

[28] Lozano S and Villa G. *Centralized resource allocation using data envelopment analysis*, Journal of Productivity Analysis **22**, 143-161, 2004.

[29] Lozano S, Villa G, Adenso-Diaz B. *Centralized target setting for regional recycling operations using DEA*, OMEGA **32**, 101-110, 2004.

[30] Maghbouli, M., Amirteimoori, A., Kordrostami, S. *Two-stage network structures with undesirable outputs: A DEA based approach* Measurement, 109-118, 2014.

[31] Nasrabadi, N., Dehnokhalaji, A., Kiani, A.N., Korhonen, P., Wallenins, J. *Resource allocation for performance improvement* Annals of Operations Research **196**, 459-468, 2011.

[32] Thanassoulis, E. and Dyson, R.G.*Estimating preferred target input output levels using data envelopment analysis*, European Journal of Operational Research **56**, 80-97, 1992.

[33] Thanassoulis, E. *A data envelopment analysis approach to clustering operating units for resource allocation purposes*, Omega **24**, 463-476, 1996.

[34] Thanassoulis, E.*Estimating efficient marginal resource levels using data envelopment analysis*, Journal of Cost Analysis 29-52, 1998.

[35] Wu, J., An, Q., Ali, S., Liang, L. , *DEA based resource allocation considering environmental factors*, Mathematical and Computer Modelling **58**, 1128-1137, 2013.

[36] Yan, H., Wei, Q.L., Hao, G. *DEA models for resource reallocation and production input/output estimation*, European Journal of Operational Research **136** (1), 19-31, 2002.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1579 – 1604

Generalized Gompertz-power series distributions

Saeed Tahmasebi* and Ali Akbar Jafari †‡

Abstract

In this paper, we introduce the generalized Gompertz-power series class of distributions which is obtained by compounding generalized Gompertz and power series distributions. This compounding procedure follows same way that was previously carried out by [25] and [3] in introducing the compound class of extended Weibull-power series distribution and the Weibull-geometric distribution, respectively. This distribution contains several lifetime models such as generalized Gompertz, generalized Gompertz-geometric, generalized Gompertz-poisson, generalized Gompertz-binomial distribution, and generalized Gompertz-logarithmic distribution as special cases. The hazard rate function of the new class of distributions can be increasing, decreasing and bathtub-shaped. We obtain several properties of this distribution such as its probability density function, Shannon entropy, its mean residual life and failure rate functions, quantiles and moments. The maximum likelihood estimation procedure via a EM-algorithm is presented, and sub-models of the distribution are studied in details.

Keywords: EM algorithm, Generalized Gompertz distribution, Maximum likelihood estimation, Power series distributions.

2000 AMS Classification: 60E05, 62F10.

Received : 02.10.2014 *Accepted* : 16.08.2015 *Doi* : 10.15672/HJMS.20158312681

*Department of Statistics, Persian Gulf University, Bushehr, Iran,
Email: tahmasebi@pgu.ac.ir

†Department of Statistics, Yazd University, Yazd, Iran, Email: aajafari@yazd.ac.ir

‡Corresponding Author.

1. Introduction

The exponential distribution is commonly used in many applied problems, particularly in lifetime data analysis [15]. A generalization of this distribution is the Gompertz distribution. It is a lifetime distribution and is often applied to describe the distribution of adult life spans by actuaries and demographers. The Gompertz distribution is considered for the analysis of survival in some sciences such as biology, gerontology, computer, and marketing science. Recently, [13] defined the generalized exponential distribution and in similar manner, [9] introduced the generalized Gompertz (GG) distribution. A random variable X is said to have a GG distribution denoted by $\mathrm{GG}(\alpha,\beta,\gamma)$, if its cumulative distribution function (cdf) is

$$(1.1)\qquad G(x) = [1-e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}]^{\alpha}, \quad \alpha,\beta>0, \quad \gamma>0; \quad x\geq 0.$$

and the probability density function (pdf) is

$$(1.2)\qquad \mathrm{g}(x) = \alpha\beta e^{\gamma x}e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}[1-e^{\frac{-\beta}{\gamma}(e^{\gamma x}-1)}]^{\alpha-1}.$$

The GG distribution is a flexible distribution that can be skewed to the right and to the left, and the well-known distributions are special cases of this distribution: the generalized exponential proposed by [13] when $\gamma \to 0^+$, the Gompertz distribution when $\alpha = 1$, and the exponential distribution when $\alpha = 1$ and $\gamma \to 0^+$.

In this paper, we compound the generalized Gompertz and power series distributions, and introduce a new class of distribution. This procedure follows similar way that was previously carried out by some authors: The exponential-power series distribution is introduced by [7] which is concluded the exponential- geometric [1, 2], exponential-Poisson [14], and exponential- logarithmic [27] distributions; the Weibull- power series distributions is introduced by [22] and is a generalization of the exponential-power series distribution; the generalized exponential-power series distribution is introduced by [19] which is concluded the Poisson-exponential [5, 18] complementary exponential-geometric [17], and the complementary exponential-power series [10] distributions; linear failure rate-power series distributions [20].

The remainder of our paper is organized as follows: In Section 2, we give the probability density and failure rate functions of the new distribution. Some properties such as quantiles, moments, order statistics, Shannon entropy and mean residual life are given in Section 3. In Section 4, we consider four special cases of this new distribution. We discuss estimation by maximum likelihood and provide an expression for Fisher's information matrix in Section 5. A simulation study is performed in Section 6. An application is given in the Section 7.

2. The generalized Gompertz-power series model

A discrete random variable, N is a member of power series distributions (truncated at zero) if its probability mass function is given by

$$(2.1)\qquad p_n = P(N=n) = \frac{a_n\theta^n}{C(\theta)}, \quad n=1,2,\ldots,$$

where $a_n \geq 0$ depends only on n, $C(\theta)=\sum_{n=1}^{\infty} a_n\theta^n$, and $\theta \in (0,s)$ (s can be ∞) is such that $C(\theta)$ is finite. Table 1 summarizes some particular cases of the truncated (at zero) power series distributions (geometric, Poisson, logarithmic and binomial). Detailed properties of power series distribution can be found in [23]. Here, $C'(\theta)$, $C''(\theta)$ and $C'''(\theta)$ denote the first, second and third derivatives of $C(\theta)$ with respect to θ, respectively.

Table 1. Useful quantities for some power series distributions.

Distribution	a_n	$C(\theta)$	$C'(\theta)$	$C''(\theta)$	$C'''(\theta)$	s
Geometric	1	$\theta(1-\theta)^{-1}$	$(1-\theta)^{-2}$	$2(1-\theta)^{-3}$	$6(1-\theta)^{-4}$	1
Poisson	$n!^{-1}$	$e^{\theta}-1$	e^{θ}	e^{θ}	e^{θ}	∞
Logarithmic	n^{-1}	$-\log(1-\theta)$	$(1-\theta)^{-1}$	$(1-\theta)^{-2}$	$2(1-\theta)^{-3}$	1
Binomial	$\binom{m}{n}$	$(1+\theta)^m-1$	$\frac{m}{(\theta+1)^{1-m}}$	$\frac{m(m-1)}{(\theta+1)^{2-m}}$	$\frac{m(m-1)(k-2)}{(\theta+1)^{3-m}}$	∞

We define generalized Gompertz-Power Series (GGPS) class of distributions denoted as GGPS$(\alpha,\beta,\gamma,\theta)$ with cdf

$$F(x)=\sum_{n=1}^{\infty}\frac{a_n(\theta G(x))^n}{C(\theta)}=\frac{C(\theta G(x))}{C(\theta)}=\frac{C(\theta t^{\alpha})}{C(\theta)},\quad x>0, \tag{2.2}$$

where $t=1-e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}$. The pdf of this distribution is given by

$$f(x)=\frac{\theta\alpha\beta}{C(\theta)}e^{\gamma x}(1-t)t^{\alpha-1}C'\left(\theta t^{\alpha}\right). \tag{2.3}$$

This class of distribution is obtained by compounding the Gompertz distribution and power series class of distributions as follows. Let N be a random variable denoting the number of failure causes which it is a member of power series distributions (truncated at zero). For given N, let $X_1, X_2, \ldots, X_N$ be a independent random sample of size N from a GG(α,β,γ) distribution. Let $X_{(N)}=\max_{1\le i\le N} X_i$. Then, the conditional cdf of $X_{(N)} \mid N=n$ is given by

$$G_{X_{(N)}|N=n}(x)=[1-e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}]^{n\alpha},$$

which has GG$(n\alpha,\beta,\gamma)$ distribution. Hence, we obtain

$$P(X_{(N)}\le x, N=n)=\frac{a_n(\theta G(x))^n}{C(\theta)}=\frac{a_n\theta^n}{C(\theta)}[1-e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}]^{n\alpha}.$$

Therefore, the marginal cdf of $X_{(N)}$ has GGPS distribution. This class of distributions can be applied to reliability problems. Therefore, some of its properties are investigated in the following.

2.1. Proposition. *The pdf of GGPS class can be expressed as infinite linear combination of pdf of order distribution, i.e. it can be written as*

$$f(x)=\sum_{n=1}^{\infty}p_n\ \mathrm{g}_{(n)}(x;n\alpha,\beta,\gamma), \tag{2.4}$$

where $\mathrm{g}_{(n)}(x;n\alpha,\beta,\gamma)$ *is the pdf of* GG$(n\alpha,\beta,\gamma)$.

Proof. Consider $t=1-e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}$. So

$$\begin{aligned} f(x) &= \frac{\theta\alpha\beta}{C(\theta)}e^{\gamma x}(1-t)t^{\alpha-1}C'\left(\theta t^{\alpha}\right)=\frac{\theta\alpha\beta}{C(\theta)}e^{\gamma x}(1-t)t^{\alpha-1}\sum_{n=1}^{\infty}na_n(\theta t^{\alpha})^{n-1} \\ &= \sum_{n=1}^{\infty}\frac{a_n\theta^n}{C(\theta)}n\alpha\beta(1-t)e^{\gamma x}t^{n\alpha-1}=\sum_{n=1}^{\infty}p_n\mathrm{g}_{(n)}(x;n\alpha,\beta,\gamma). \end{aligned}$$

□

2.2. Proposition. *The limiting distribution of* GGPS$(\alpha, \beta, \gamma, \theta)$ *when* $\theta \to 0^+$ *is*

$$\lim_{\theta \to 0^+} F(x) = [1 - e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}]^{c\alpha},$$

which is a GG distribution with parameters $c\alpha$, β, *and* γ, *where* $c = \min\{n \in \mathbb{N} : a_n > 0\}$.

Proof. Consider $t = 1 - e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}$. So

$$\begin{aligned} \lim_{\theta \to 0^+} F(x) &= \lim_{\theta \to 0^+} \frac{C(\lambda t^\alpha)}{C(\theta)} = \lim_{\lambda \to 0^+} \frac{\sum_{n=1}^{\infty} a_n \theta^n t^{n\alpha}}{\sum_{n=1}^{\infty} a_n \theta^n} \\ &= \lim_{\theta \to 0^+} \frac{a_c t^{c\alpha} + \sum_{n=c+1}^{\infty} a_n \theta^{n-c} t^{n\alpha}}{a_c + \sum_{n=c+1}^{\infty} a_n \theta^{n-c}} = t^{c\alpha}. \end{aligned}$$

□

2.3. Proposition. *The limiting distribution of* GGPS$(\alpha, \beta, \gamma, \theta)$ *when* $\gamma \to 0^+$ *is*

$$\lim_{\gamma \to 0^+} F(x) = \frac{C(\theta(1 - e^{-\beta x})^\alpha)}{C(\theta)},$$

i.e. the cdf of the generalized exponential-power series class of distribution introduced by [19].

Proof. When $\gamma \to 0^+$, the generalized Gompertz distribution becomes to generalized exponential distribution. Therefore, proof is obvious. □

2.4. Proposition. *The hazard rate function of the GGPS class of distributions is*

$$h(x) = \frac{\theta\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}C'(\theta t^\alpha)}{C(\theta) - C(\theta t^\alpha)}, \tag{2.5}$$

where $t = 1 - e^{\frac{-\beta}{\gamma}(e^{\gamma x}-1)}$.

Proof. Using (2.2), (2.3) and definition of hazard rate function as $h(x) = f(x)/(1-F(x))$, the proof is obvious. □

2.5. Proposition. *For the pdf in* (2.3), *we have*

$$\lim_{x \to 0^+} f(x) = \begin{cases} \infty & 0 < \alpha < 1 \\ \frac{C'(0)\theta\beta}{C(\theta)} & \alpha = 1 \\ 0 & \alpha > 1, \end{cases} \qquad \lim_{x \to \infty} f(x) = 0.$$

Proof. The proof is a forward calculation using the following limits

$$\lim_{x \to 0^+} t^{\alpha-1} = \begin{cases} \infty & 0 < \alpha < 1 \\ 1 & \alpha = 1 \\ 0 & \alpha > 1, \end{cases} \qquad \lim_{x \to 0^+} t^\alpha = 0, \qquad \lim_{x \to \infty} t = 1.$$

□

2.6. Proposition. *For the hazard rate function in* (2.5), *we have*

$$\lim_{x \to 0^+} h(x) = \begin{cases} \infty & 0 < \alpha < 1 \\ \frac{C'(0)\theta\beta}{C(\theta)} & \alpha = 1 \\ 0 & \alpha > 1, \end{cases} \qquad \lim_{x \to \infty} h(x) = \begin{cases} \infty & \gamma > 0 \\ \beta & \gamma \to 0 \end{cases}$$

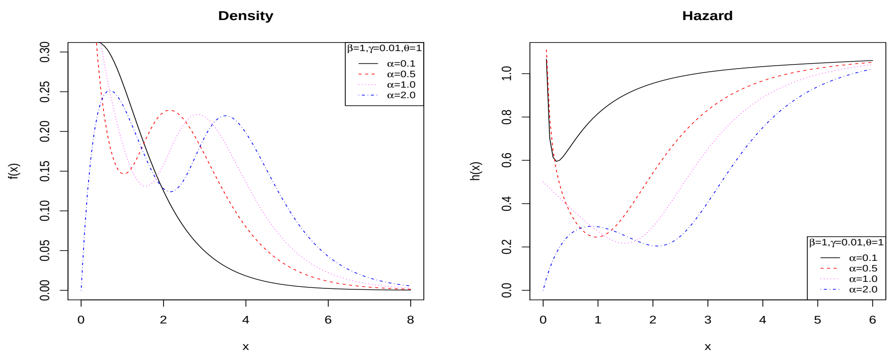


Figure 1. Plots of pdf and hazard rate functions of GGPS with $C(\theta) = \theta + \theta^{20}$.

Proof. Since $\lim_{x\to 0^+}(1 - F(x)) = 1$, we have $\lim_{x\to 0^+} h(x) = \lim_{x\to 0^+} f(x)$.
For $\lim_{x\to\infty} h(x)$, the proof is satisfied using the limits

$$\lim_{x\to\infty} C'(\theta t^{\alpha}) = C'(\theta), \qquad \lim_{x\to\infty} t^{\alpha-1} = 1,$$

$$\lim_{x\to\infty} \frac{e^{\gamma x}(1-t)}{C(\theta) - C(\theta t^{\alpha})} = \lim_{x\to\infty} \frac{e^{\gamma x}(1-t)[\beta e^{\gamma x} - \gamma]}{\theta\beta\alpha C'(\theta) e^{\gamma x}(1-t)} = \begin{cases} \infty & \gamma > 0 \\ \frac{1}{\theta\alpha C'(\theta)} & \gamma \to 0. \end{cases}$$

□

As a example, we consider $C(\theta) = \theta + \theta^{20}$. The plots of pdf and hazard rate function of GGPS for parameters $\beta = 1, \gamma = .01, \theta = 1.0$, and $\alpha = 0.1, 0.5, 1.0, 2.0$ are given in Figure 1. This pdf is bimodal when $\alpha = 2.0$, and the values of modes are 0.7 and 3.51.

3. Statistical properties

In this section, some properties of GGPS distribution such as quantiles, moments, order statistics, Shannon entropy and mean residual life are obtained.

3.1. Quantiles and Moments. The quantile q of GGPS is given by

$$x_q = G^{-1}\left(\frac{C^{-1}(qC(\theta))}{\theta}\right), \qquad 0 < q < 1,$$

where $G^{-1}(y) = \frac{1}{\gamma}\log[1 - \frac{\gamma \log(1-y^{\frac{1}{\alpha}})}{\beta}]$ and $C^{-1}(.)$ is the inverse function of $C(.)$. This result helps in simulating data from the GGPS distribution with generating uniform distribution data.

For checking the consistency of the simulating data set form GGPS distribution, the histogram for a generated data set with size 100 and the exact pdf of GGPS with $C(\theta) = \theta + \theta^{20}$, and parameters $\alpha = 2$, $\beta = 1$, $\gamma = 0.01$, $\theta = 1.0$, are displayed in Figure 2 (left). Also, the empirical cdf and the exact cdf are given in Figure 2 (right).

Consider $X \sim \text{GGPS}(\alpha, \beta, \gamma, \theta)$. Then the Laplace transform of the GGPS class can be expressed as

$$L(s) = E(e^{-sX}) = \sum_{n=1}^{\infty} P(N = n) L_n(s), \tag{3.1}$$

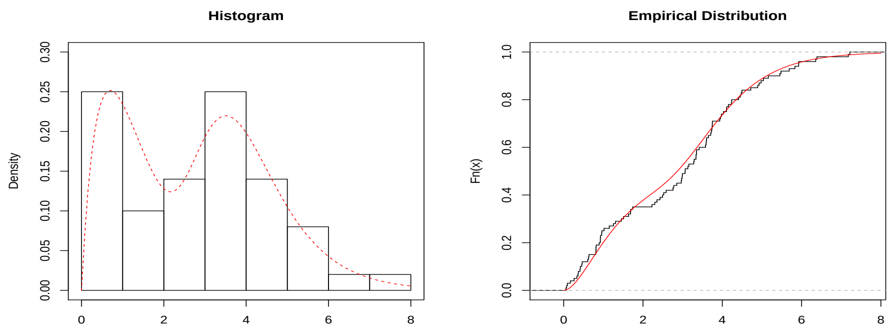


Figure 2. The histogram of a generated data set with the pdf (left) and the empirical cdf with cdf (right) of GGPS distribution.

where $L_n(s)$ is the Laplace transform of $\mathrm{GG}(n\alpha, \beta, \gamma)$ distribution given as

$$\begin{aligned} L_n(s) &= \int_0^{+\infty} e^{-sx} n\alpha\beta e^{\gamma x} e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}[1-e^{\frac{-\beta}{\gamma}(e^{\gamma x}-1)}]^{n\alpha-1}dx \\ &= n\alpha\beta \int_0^{+\infty} e^{(\gamma-s)x} e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)} \\ &\quad \sum_{j=0}^{\infty} \binom{n\alpha-1}{j}(-1)^j e^{\frac{-\beta}{\gamma}j(e^{\gamma x}-1)}dx \\ &= n\alpha\beta \sum_{j=0}^{\infty} \binom{n\alpha-1}{j}(-1)^j e^{\frac{\beta}{\gamma}(j+1)} \\ &\quad \int_0^{+\infty} e^{(\gamma-s)x} e^{\frac{-\beta}{\gamma}(j+1)e^{\gamma x}}dx \\ &= n\alpha\beta \sum_{j=0}^{\infty} \binom{n\alpha-1}{j}(-1)^j e^{\frac{\beta}{\gamma}(j+1)} \\ &\quad \times \int_0^{+\infty} e^{(\gamma-s)x} \sum_{k=0}^{\infty} \frac{(-1)^k(\frac{\beta}{\gamma}(j+1))^k e^{\gamma k x}}{\Gamma(k+1)}dx \\ &= n\alpha\beta \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \binom{n\alpha-1}{j} \frac{(-1)^{j+k} e^{\frac{\beta}{\gamma}(j+1)}[\frac{\beta}{\gamma}(j+1)]^k}{\Gamma(k+1)(s-\gamma-\gamma k)}, \quad s>\gamma. \end{aligned} \tag{3.2}$$

Now, we obtain the moment generating function of GGPS.

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \sum_{n=1}^{\infty} P(N=n)L_n(-t) \\ &= \alpha\beta \sum_{n=1}^{\infty} \frac{a_n\theta^n}{C(\theta)} \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{n\binom{n\alpha-1}{j}(-1)^{j+k+1} e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k}{\Gamma(k+1)(t+\gamma+\gamma k)} \\ &= \alpha\beta E_N[\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{N\binom{N\alpha-1}{j}(-1)^{j+k+1} e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k}{\Gamma(k+1)(t+\gamma+\gamma k)}], \end{aligned} \tag{3.3}$$

where N is a random variable from the power series family with the probability mass function in (2.1) and $E_N[U]$ is expectation of U with respect to random variable N.

We can use $M_X(t)$ to obtain the non-central moments, $\mu_r = E[X^r]$. But from the direct calculation, we have

$$\mu_r = \sum_{n=1}^{\infty} \frac{a_n\theta^n}{C(\theta)} \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{n\alpha\beta\binom{n\alpha-1}{j}(-1)^{j+k+r+1}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k\Gamma(r+1)}{\Gamma(k+1)(\gamma+\gamma k)^{r+1}}$$

$$(3.4) \qquad = \alpha\beta E_N[\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{N\binom{N\alpha-1}{j}(-1)^{j+k+r+1}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k\Gamma(r+1)}{\Gamma(k+1)(\gamma+\gamma k)^{r+1}}].$$

3.1. Proposition. *For non-central moment function in 3.4, we have*

$$\lim_{\theta\to 0^+} \mu_r = E[Y^r],$$

where Y *has* $\mathrm{GG}(c\alpha,\beta,\gamma)$ *and* $c = \min\{n \in \mathbb{N} : a_n > 0\}$.

Proof. If Y has $\mathrm{GG}(c\alpha,\beta,\gamma)$, then

$$E[Y^r] = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{c\alpha\beta\binom{c\alpha-1}{j}(-1)^{j+k+r+1}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k\Gamma(r+1)}{\Gamma(k+1)(\gamma+\gamma k)^{r+1}}.$$

Therefore,

$$\lim_{\theta\to 0^+}\mu_r = \lim_{\theta\to 0^+} \frac{\sum_{n=1}^{\infty} a_n\theta^n E[Y^r]}{\sum_{n=1}^{\infty} a_n\theta^n} = \lim_{\theta\to 0^+} \frac{a_c E[Y^r] + \sum_{n=c+1}^{\infty} a_n\theta^{n-c}E[Y^r]}{a_c + \sum_{n=c+1}^{\infty} a_n\theta^{n-c}} = E[Y^r].$$

□

3.2. Order statistic. Let $X_1, X_2, \ldots, X_n$ be an independent random sample of size n from $\mathrm{GGPS}(\alpha,\beta,\gamma,\theta)$. Then, the pdf of the ith order statistic, say $X_{i:n}$, is given by

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} f(x)[\frac{C(\theta t^\alpha)}{C(\theta)}]^{i-1}[1 - \frac{C(\theta t^\alpha)}{C(\theta)}]^{n-i},$$

where f is the pdf given in (2.3) and $t = 1 - e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}$. Also, the cdf of $X_{i:n}$ is given by

$$F_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!}\sum_{k=0}^{n-i} \frac{(-1)^k\binom{n-i}{k}}{k+i+1}[\frac{C(t^\alpha)}{C(\theta)}]^{k+i}.$$

An analytical expression for rth non-central moment of order statistics $X_{i:n}$ is obtained as

$$E[X_{i:n}^r] = r\sum_{k=n-i+1}^{n} (-1)^{k-n+i-1}\binom{k-1}{n-i}\binom{n}{k}\int_0^{+\infty} x^{r-1}S(x)^k dx$$

$$= r\sum_{k=n-i+1}^{n} \frac{(-1)^{k-n+i-1}}{[C(\theta)]^k}\binom{k-1}{n-i}\binom{n}{k}\int_0^{+\infty} x^{r-1}[C(\theta) - C(\theta t^\alpha)]^k dx,$$

where $S(x) = 1 - F(x)$ is the survival function of GGPS distribution.

3.3. Shannon entropy and mean residual life. If X is a none-negative continuous random variable with pdf f, then Shannon's entropy of X is defined by [24] as

$$H(f) = E[-\log f(X)] = -\int_0^{+\infty} f(x)\log(f(x))dx,$$

and this is usually referred to as the continuous entropy (or differential entropy). An explicit expression of Shannon entropy for GGPS distribution is obtained as

$$\begin{aligned} H(f) &= E\{-\log[\frac{\theta\alpha\beta}{C(\theta)}e^{\gamma X}(e^{-\frac{\beta}{\gamma}(e^{\gamma X}-1)})(1-e^{-\frac{\beta}{\gamma}(e^{\gamma X}-1)})^{\alpha-1} \\ &\quad \times C'\left(\theta(1-e^{-\frac{\beta}{\gamma}(e^{\gamma X}-1)})^{\alpha}\right)]\} \\ &= -\log[\frac{\theta\beta\alpha}{C(\theta)}] - \gamma E(X) + \frac{\beta}{\gamma}E(e^{\gamma X}) - \frac{\beta}{\gamma} \\ &\quad -(\alpha-1)E[\log(1-e^{-\frac{\beta}{\gamma}(e^{\gamma X}-1)})] \\ &\quad -E[\log(C'\left(\theta(1-e^{-\frac{\beta}{\gamma}(e^{\gamma X}-1)})^{\alpha}\right))] \\ &= -\log[\frac{\theta\beta\alpha}{C(\theta)}] - \gamma\mu_1 + \frac{\beta}{\gamma}M_X(\gamma) - \frac{\beta}{\gamma} \\ &\quad -(\alpha-1)\sum_{n=1}^{\infty}P(N=n)\int_0^1 n\alpha t^{n\alpha-1}\log(t)dt \\ &\quad -\sum_{n=1}^{\infty}P(N=n)\int_0^1 nu^{n-1}\log(C'(\theta u))du \\ &= -\log[\frac{\theta\beta\alpha}{C(\theta)}] - \gamma\mu_1 + \frac{\beta}{\gamma}M_X(\gamma) - \frac{\beta}{\gamma} \\ &\quad +\frac{(\alpha-1)}{\alpha}E_N[\frac{1}{N}] - E_N[A(N,\theta)], \end{aligned} \tag{3.5}$$

where $A(N,\theta) = \int_0^1 Nu^{N-1}\log(C'(\theta u))du$, N is a random variable from the power series family with the probability mass function in (2.1), and $E_N[U]$ is expectation of U with respect to random variable N. In reliability theory and survival analysis, X usually denotes a duration such as the lifetime. The residual lifetime of the system when it is still operating at time s, is $X_s = X - s \mid X > s$ which has pdf

$$f(x;s) = \frac{f(x)}{1-F(s)} = \frac{\theta g(x)C'(\theta G(x))}{C(\theta) - C(\theta G(s))}, \quad x \ge s > 0.$$

Also, the mean residual lifetime of X_s is given by

$$\begin{aligned} m(s) = E[X-s|X>s] &= \frac{\int_s^{+\infty}(x-s)f(x)dx}{1-F(s)} \\ &= \frac{\int_s^{+\infty}xf(x)dx}{1-F(s)} - s \\ &= \frac{C(\theta)E_N[Z(s,N)]}{C(\theta) - C(\theta[1-e^{-\frac{\beta}{\gamma}(e^{\gamma s}-1)}]^{\alpha})} - s, \end{aligned}$$

where $Z(s,n) = \int_s^{+\infty} x g_{(n)}(x;n\alpha,\beta,\gamma)dx$, and $g_{(n)}(x;n\alpha,\beta,\gamma)$ is the pdf of GG $(n\alpha,\beta,\gamma)$.

4. Special cases of GGPS distribution

In this Section, we consider four special cases of the GGPS distribution. To simplify, we consider $t = 1 - e^{-\frac{\beta}{\gamma}(e^{\gamma x}-1)}$, $x > 0$, and $A_j = \binom{n\alpha-1}{j}$.

4.1. Generalized Gompertz-geometric distribution. The geometric distribution (truncated at zero) is a special case of power series distributions with $a_n = 1$ and $C(\theta) = \frac{\theta}{1-\theta}$ $(0 < \theta < 1)$. The pdf and hazard rate function of generalized Gompertz-geometric (GGG) distribution is given respectively by

$$f(x) = \frac{(1-\theta)\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}}{(\theta t^{\alpha}-1)^2}, \quad x > 0, \tag{4.1}$$

$$h(x) = \frac{(1-\theta)\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}}{(1-\theta t^{\alpha})(1-t^{\alpha})}, \quad x > 0. \tag{4.2}$$

4.1. Remark. Consider

$$f_M(x) = \frac{\theta^*\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}}{((1-\theta^*)t^{\alpha}-1)^2}, \quad x > 0, \tag{4.3}$$

where $\theta^* = 1-\theta$. Then $f_M(x)$ is pdf for all $\theta^* > 0$ (see [21]). Note that when $\alpha = 1$ and $\gamma \to 0^+$, the pdf of extended exponential geometric (EEG) distribution [1] is concluded from (4.3). The EEG hazard function is monotonically increasing for $\theta^* > 1$; decreasing for $0 < \theta^* < 1$ and constant for $\theta^* = 1$.

4.2. Remark. If $\alpha = \theta^* = 1$, then the pdf in (4.3) becomes the pdf of Gompertz distribution. Note that the hazard rate function of Gompertz distribution is $h(x) = \beta e^{\gamma x}$ which is increasing.

The plots of pdf and hazard rate function of GGG for different values of α, β, γ and θ^* are given in Figure 3.

4.3. Theorem. *Consider the GGG hazard function in* (4.2). *Then, for* $\alpha \geq 1$, *the hazard function is increasing and for* $0 < \alpha < 1$, *is decreasing and bathtub shaped.*

Proof. See Appendix A.1. □

The first and second non-central moments of GGG are given by

$$E(X) = \alpha\beta(1-\theta)\sum_{n=1}^{\infty} n\theta^{n-1}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{A_j(-1)^{j+k}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k}{\Gamma(k+1)(\gamma+\gamma k)^2},$$

$$E(X^2) = 2\alpha\beta(1-\theta)\sum_{n=1}^{\infty} n\theta^{n-1}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{A_j(-1)^{j+k+3}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k}{\Gamma(k+1)(\gamma+\gamma k)^3}.$$

4.2. Generalized Gompertz-Poisson distribution. The Poisson distribution (truncated at zero) is a special case of power series distributions with $a_n = \frac{1}{n!}$ and $C(\theta) = e^{\theta}-1$ $(\theta > 0)$. The pdf and hazard rate function of generalized Gompertz-Poisson (GGP) distribution are given respectively by

$$f(x) = \theta\alpha\beta e^{\gamma x-\theta}(1-t)t^{\alpha-1}e^{\theta t^{\alpha}}, \quad x > 0 \tag{4.4}$$

$$h(x) = \frac{\theta\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}e^{\theta t^{\alpha}}}{e^{\theta}-e^{\theta t^{\alpha}}}, \quad x > 0. \tag{4.5}$$

4.4. Theorem. *Consider the GGP hazard function in (4.5). Then, for* $\alpha \geq 1$, *the hazard function is increasing and for* $0 < \alpha < 1$, *is decreasing and bathtub shaped.*

Proof. See Appendix A.2. □

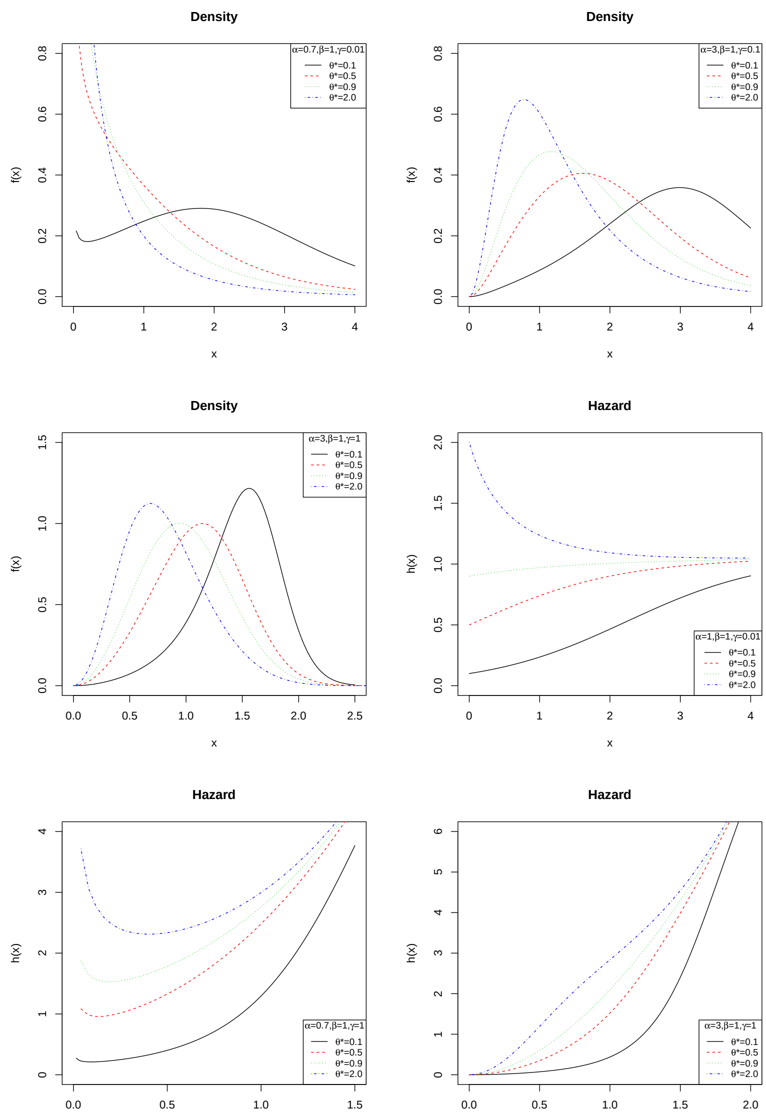


Figure 3. Plots of pdf and hazard rate function of GGG for different values α, β, γ and θ^*.

The first and second non-central moments of GGP can be computed as

$$E(X) = \frac{\alpha\beta}{e^{\theta}-1}\sum_{n=1}^{\infty}\frac{\theta^n}{(n-1)!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{A_j(-1)^{j+k}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k}{\Gamma(k+1)(\gamma+\gamma k)^2},$$

$$E(X^2) = \frac{2\alpha\beta}{e^{\theta}-1}\sum_{n=1}^{\infty}\frac{\theta^n}{(n-1)!}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{A_j(-1)^{j+k+3}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^k}{\Gamma(k+1)(\gamma+\gamma k)^3}.$$

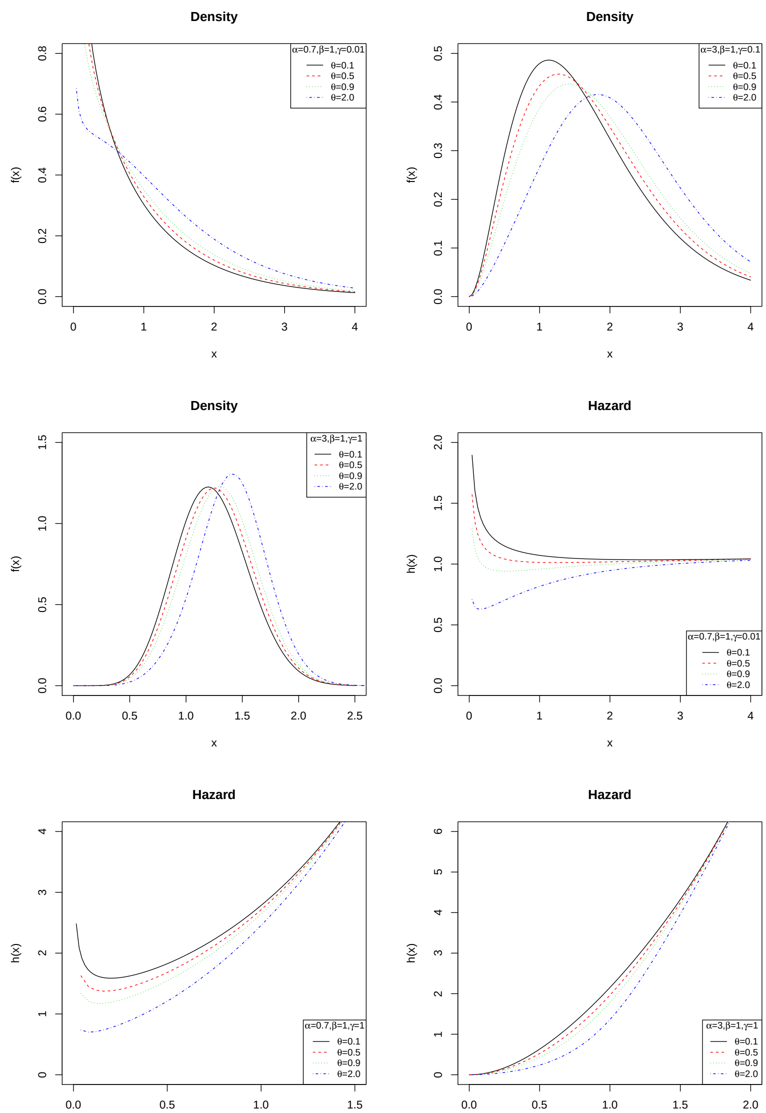


Figure 4. Plots of pdf and hazard rate function of GGP for different values α, β, γ and θ.

The plots of pdf and hazard rate function of GGP for different values of α, β, γ and θ are given in Figure 4.

4.3. Generalized Gompertz-binomial distribution. The binomial distribution (truncated at zero) is a special case of power series distributions with $a_n = \binom{m}{n}$ and $C(\theta) = (\theta+1)^m - 1$ ($\theta > 0$), where m ($n \leq m$) is the number of replicas. The pdf and hazard rate function of generalized Gompertz-binomial (GGB) distribution are given respectively by

$$f(x) \quad = \quad m\theta\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}\frac{(\theta t^{\alpha}+1)^{m-1}}{(\theta+1)^m - 1}, \quad x > 0, \tag{4.6}$$

$$(4.7) \qquad h(x) \quad = \quad \frac{m\theta\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}(\theta t^{\alpha}+1)^{m-1}}{(\theta+1)^{m}-(\theta t^{\alpha}+1)^{m}}, \quad x>0.$$

The plots of pdf and hazard rate function of GGB for $m = 4$, and different values of α, β, γ and θ are given in Figure 5. We can find that the GGP distribution can be obtained as limiting of GGB distribution if $m\theta \to \lambda > 0$, when $m \to \infty$.

4.5. Theorem. *Consider the GGB hazard function in (4.7). Then, for $\alpha \geq 1$, the hazard function is increasing and for $0 < \alpha < 1$, is decreasing and bathtub shaped.*

Proof. The proof is omitted, since $\theta > 0$ and therefore the proof is similar to the proof of Theorem 4.4. □

The first and second non-central moments of GGB are given by

$$E(X) = \frac{\alpha\beta}{(\theta+1)^{m}-1}\sum_{n=1}^{\infty}\theta^{n}n\binom{m}{n}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{A_{j}(-1)^{j+k}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^{k}}{\Gamma(k+1)(\gamma+\gamma k)^{2}},$$

$$E(X^{2}) = \frac{2\alpha\beta}{(\theta+1)^{m}-1}\sum_{n=1}^{\infty}\theta^{n}n\binom{m}{n}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{A_{j}(-1)^{j+k+3}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^{k}}{\Gamma(k+1)(\gamma+\gamma k)^{3}}.$$

4.4. Generalized Gompertz-logarithmic distribution. The logarithmic distribution (truncated at zero) is also a special case of power series distributions with $a_n = \frac{1}{n}$ and $C(\theta) = -\log(1-\theta)$ $(0 < \theta < 1)$. The pdf and hazard rate function of generalized Gompertz-logarithmic (GGL) distribution are given respectively by

$$(4.8) \qquad f(x) \quad = \quad \frac{\theta\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}}{(\theta t^{\alpha}-1)\log(1-\theta)}, \quad x>0,$$

$$(4.9) \qquad h(x) \quad = \quad \frac{\theta\alpha\beta e^{\gamma x}(1-t)t^{\alpha-1}}{(\theta t^{\alpha}-1)\log(\frac{1-\theta}{1-\theta t^{\alpha}})}, \quad x>0.$$

The plots of pdf and hazard rate function of GGL for different values of α, β, γ and θ are given in Figure 6.

4.6. Theorem. *Consider the GGL hazard function in (4.9). Then, for $\alpha \geq 1$, the hazard function is increasing and for $0 < \alpha < 1$, is decreasing and bathtub shaped.*

Proof. The proof is omitted, since $0 < \theta < 1$ and therefore the proof is similar to the proof of Theorem 1. □

The first and second non-central moments of GGL are

$$E(X) = \frac{\alpha\beta}{-\log(1-\theta)}\sum_{n=1}^{\infty}\theta^{n}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{A_{j}(-1)^{j+k}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^{k}}{\Gamma(k+1)(\gamma+\gamma k)^{2}},$$

$$E(X^{2}) = \frac{2\alpha\beta}{-\log(1-\theta)}\sum_{n=1}^{\infty}\theta^{n}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{A_{j}(-1)^{j+k+3}e^{\frac{\beta}{\gamma}(j+1)}(\frac{\beta}{\gamma}(j+1))^{k}}{\Gamma(k+1)(\gamma+\gamma k)^{3}}.$$

5. Estimation and inference

In this section, we will derive the maximum likelihood estimators (MLE) of the unknown parameters $\mathbf{\Theta} = (\alpha, \beta, \gamma, \theta)^{T}$ of the GGPS$(\alpha, \beta, \gamma, \theta)$. Also, asymptotic confidence intervals of these parameters will be derived based on the Fisher information. At the end, we proposed an Expectation-Maximization (EM) algorithm for estimating the parameters.

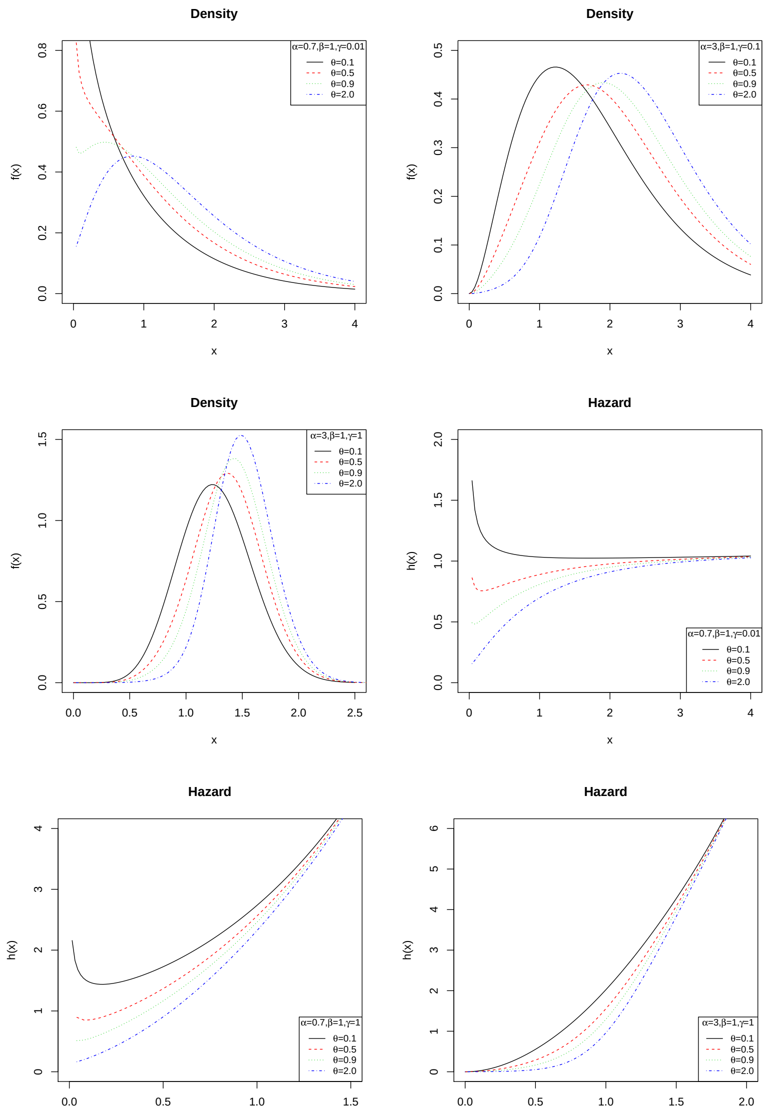


Figure 5. Plots of pdf and hazard rate function of GGB for $m = 5$, and different values α, β, γ and θ.

5.1. MLE for parameters. Let $X_1, \ldots, X_n$ be an independent random sample, with observed values $x_1, \ldots, x_n$ from GGPS$(\alpha, \beta, \gamma, \theta)$ and $\boldsymbol{\Theta} = (\alpha, \beta, \gamma, \theta)^T$ be a parameter vector. The log-likelihood function is given by

$$\begin{aligned} l_n = l_n(\boldsymbol{\Theta}; \boldsymbol{x}) &= n\log(\theta) + n\log(\alpha\beta) + n\gamma\bar{x} \\ &+ \sum_{i=1}^{n}\log(1-t_i) + (\alpha-1)\sum_{i=1}^{n}\log(t_i) \end{aligned}$$

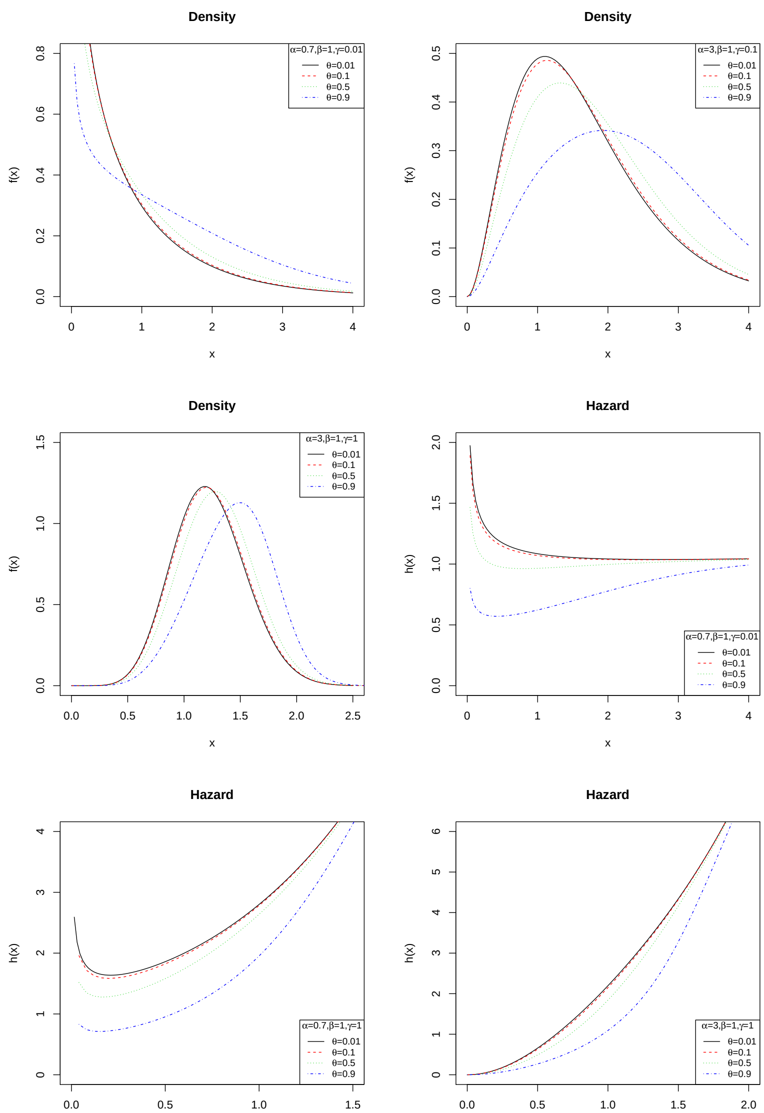


Figure 6. Plots of pdf and hazard rate function of GGL for different values α, β, γ and θ.

$$+\sum_{i=1}^{n}\log(C'(\theta t_i^{\alpha}))-n\log(C(\theta)),$$

where $t_i = 1-e^{-\frac{\beta}{\gamma}(e^{\gamma x_i}-1)}$. Therefore, the score function is given by $U(\boldsymbol{\Theta};\boldsymbol{x}) = (\frac{\partial l_n}{\partial\alpha}, \frac{\partial l_n}{\partial\beta}, \frac{\partial l_n}{\partial\gamma}, \frac{\partial l_n}{\partial\theta})^T$, where

$$(5.1)\qquad \frac{\partial l_n}{\partial\alpha} = \frac{n}{\alpha}+\sum_{i=1}^{n}\log(t_i)+\sum_{i=1}^{n}\frac{\theta t_i^{\alpha}\log(t_i)C''(\theta t_i^{\alpha})}{C'(\theta t_i^{\alpha})},$$

$$\frac{\partial l_n}{\partial \beta} = \frac{n}{\beta} - \frac{1}{\gamma}(\sum_{i=1}^{n} e^{\gamma x_i} - n) + (\alpha-1)\sum_{i=1}^{n}\frac{\frac{\partial t_i}{\partial \beta}}{t_i} + \sum_{i=1}^{n}\frac{\theta\frac{\partial (t_i^\alpha)}{\partial \beta}C''(\theta t_i^\alpha)}{C'(\theta t_i^\alpha)}, \tag{5.2}$$

$$\begin{aligned}\frac{\partial l_n}{\partial \gamma} &= n\bar{x} + \frac{\beta}{\gamma^2}(\sum_{i=1}^{n} e^{\gamma x_i} - n) - \frac{\beta}{\gamma}(\sum_{i=1}^{n} x_i e^{\gamma x_i}) \\ &\quad +(\alpha-1)\sum_{i=1}^{n}\frac{\frac{\partial t_i}{\partial \gamma}}{t_i} + \sum_{i=1}^{n}\frac{\theta\frac{\partial (t_i^\alpha)}{\partial \gamma}C''(\theta t_i^\alpha)}{C'(\theta t_i^\alpha)},\end{aligned} \tag{5.3}$$

$$\frac{\partial l_n}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n}\frac{t_i^\alpha C''(\theta t_i^\alpha)}{C'(\theta t_i^\alpha)} - \frac{nC'(\theta)}{C(\theta)}. \tag{5.4}$$

The MLE of $\boldsymbol{\Theta}$, say $\hat{\boldsymbol{\Theta}}$, is obtained by solving the nonlinear system $U(\boldsymbol{\Theta};\boldsymbol{x}) = \mathbf{0}$. We cannot get an explicit form for this nonlinear system of equations and they can be calculated by using a numerical method, like the Newton method or the bisection method.

For each element of the power series distributions (geometric, Poisson, logarithmic and binomial), we have the following theorems for the MLE of parameters:

5.1. Theorem. *Let* $\mathrm{g}_1(\alpha;\beta,\gamma,\theta,x)$ *denote the function on RHS of the expression in* (5.1)*, where* β*,* γ *and* θ *are the true values of the parameters. Then, for a given* $\beta > 0$*,* $\gamma > 0$ *and* $\theta > 0$*, the roots of* $\mathrm{g}_1(\alpha,\beta;\gamma,\theta,\boldsymbol{x}) = 0$*, lies in the interval*

$$\left(\frac{-n}{\frac{\theta C''(\theta)}{C'(\theta)}+1}(\sum_{i=1}^{n}\log(t_i))^{-1}, -n(\sum_{i=1}^{n}\log(t_i))^{-1})\right),$$

Proof. See Appendix B.1. □

5.2. Theorem. *Let* $\mathrm{g}_2(\beta;\alpha,\gamma,\theta,x)$ *denote the function on RHS of the expression in* (5.3)*, where* α*,* γ *and* θ *are the true values of the parameters. Then, the equation* $\mathrm{g}_2(\beta;\alpha,\gamma,\theta,x) = 0$ *has at least one root.*

Proof. See Appendix B.2. □

5.3. Theorem. *Let* $\mathrm{g}_3(\theta;\alpha,\beta,\gamma,x)$ *denote the function on RHS of the expression in* (5.4) *and* $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$*, where* α*,* β *and* γ *are the true values of the parameters.*
a) The equation $\mathrm{g}_3(\theta;\alpha,\beta,\gamma,x) = 0$ *has at least one root for all GGG, GGP and GGL distributions if* $\sum_{i=1}^{n} t_i^\alpha > \frac{n}{2}$*.*
b) If $\mathrm{g}_3(p;\alpha,\beta,\gamma,x) = \frac{\partial l_n}{\partial p}$*, where* $p = \frac{\theta}{\theta+1}$ *and* $p \in (0,1)$ *then the equation* $\mathrm{g}_3(p;\alpha,\beta,\gamma,x) = 0$ *has at least one root for GGB distribution if* $\sum_{i=1}^{n} t_i^\alpha > \frac{n}{2}$ *and* $\sum_{i=1}^{n} t_i^{-\alpha} > \frac{nm}{m-1}$*.*

Proof. See Appendix B.3. □

Now, we derive asymptotic confidence intervals for the parameters of GGPS distribution. It is well-known that under regularity conditions (see [6], Section 10), the asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})$ is multivariate normal with mean $\mathbf{0}$ and variance-covariance matrix $J_n^{-1}(\boldsymbol{\Theta})$, where $J_n(\boldsymbol{\Theta}) = \lim_{n\to\infty} I_n(\boldsymbol{\Theta})$, and $I_n(\boldsymbol{\Theta})$ is the 4×4 observed information matrix, i.e.

$$I_n(\Theta) = -\begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\beta} & I_{\alpha\gamma} & I_{\alpha\theta} \\ I_{\beta\alpha} & I_{\beta\beta} & I_{\beta\gamma} & I_{\beta\theta} \\ I_{\gamma\alpha} & I_{\gamma\beta} & I_{\gamma\gamma} & I_{\gamma\theta} \\ I_{\theta\alpha} & I_{\theta\beta} & I_{\theta\gamma} & I_{\theta\theta} \end{bmatrix},$$

whose elements are given in Appendix C. Therefore, an $100(1-\eta)$ asymptotic confidence interval for each parameter, $\mathbf{\Theta}_r$, is given by

$$ACI_r = (\hat{\mathbf{\Theta}}_r - Z_{\eta/2}\sqrt{\hat{I}_{rr}}, \hat{\mathbf{\Theta}}_r + Z_{\frac{\eta}{2}}\sqrt{\hat{I}_{rr}}),$$

where $\hat{I}_{rr}$ is the (r,r) diagonal element of $I_n^{-1}(\hat{\mathbf{\Theta}})$ for $r = 1, 2, 3, 4$ and $Z_{\eta/2}$ is the quantile $\frac{\eta}{2}$ of the standard normal distribution.

5.2. EM-algorithm. The traditional methods to obtain the MLE of parameters are numerical methods by solving the equations (5.1)-(5.4), and sensitive to the initial values. Therefore, we develop an Expectation-Maximization (EM) algorithm to obtain the MLE of parameters. It is an iterative method, and is a very powerful tool in handling the incomplete data problem [8].

We define a hypothetical complete-data distribution with a joint pdf in the form

$$g(x,z;\mathbf{\Theta}) = \frac{a_z\theta^z}{C(\theta)} z\alpha\beta e^{\gamma x}(1-t)t^{z\alpha-1},$$

where $t = 1 - e^{\frac{-\beta}{\gamma}(e^{\gamma x}-1)}$, and $\alpha, \beta,\ \gamma,\ \theta > 0,\ x > 0$ and $z \in \mathbb{N}$. Suppose $\mathbf{\Theta}^{(r)} = (\alpha^{(r)}, \beta^{(r)}, \gamma^{(r)}, \theta^{(r)})$ is the current estimate (in the rth iteration) of $\mathbf{\Theta}$. Then, the E-step of an EM cycle requires the expectation of $(Z|X;\mathbf{\Theta}^{(r)})$. The pdf of Z given $X = x$ is given by

$$g(z|x) = \frac{a_z\theta^{z-1}zt^{z\alpha-\alpha}}{C'(\theta t^\alpha)},$$

and since

$$C'(\theta) + \theta C''(\theta) = \sum_{z=1}^{\infty} a_z z\theta^{z-1} + \theta\sum_{z=1}^{\infty} a_z z(z-1)\theta^{z-2} = \sum_{z=1}^{\infty} z^2 a_z\theta^{z-1},$$

the expected value of $Z|X = x$ is obtained as

$$E(Z|X=x) = 1 + \frac{\theta t^\alpha C''(\theta t^\alpha)}{C'(\theta t^\alpha)}. \tag{5.5}$$

By using the MLE over $\mathbf{\Theta}$, with the missing Z's replaced by their conditional expectations given above, the M-step of EM cycle is completed. Therefore, the log-likelihood for the complete-data is

$$\begin{aligned} l_n^*(\boldsymbol{y},\mathbf{\Theta}) \quad \propto \quad & \sum_{i=1}^{n} z_i\log(\theta) + n\log(\alpha\beta) + n\gamma\bar{x} + \sum_{i=1}^{n}\log(1-t_i) \\ & + \sum_{i=1}^{n}(z_i\alpha - 1)\log(t_i) - n\log(C(\theta)), \end{aligned} \tag{5.6}$$

where $\boldsymbol{y} = (\boldsymbol{x};\boldsymbol{z})$, $\boldsymbol{x} = (x_1,\ldots,x_n)$ and $\boldsymbol{z} = (z_1,\ldots,z_n)$. On differentiation of (5.6) with respect to parameters α, β, γ and θ, we obtain the components of the score function, $U(\boldsymbol{y};\mathbf{\Theta}) = (\frac{\partial l_n^*}{\partial\alpha}, \frac{\partial l_n^*}{\partial\beta}, \frac{\partial l_n^*}{\partial\gamma}, \frac{\partial l_n^*}{\partial\theta})^T$, as

$$\begin{aligned} \frac{\partial l_n^*}{\partial\alpha} &= \frac{n}{\alpha} + \sum_{i=1}^{n} z_i\log[1 - e^{\frac{-\beta}{\gamma}(e^{\gamma x_i}-1)}], \\ \frac{\partial l_n^*}{\partial\beta} &= \frac{n}{\beta} - \frac{1}{\gamma}(\sum_{i=1}^{n} e^{\gamma x_i} - n) + \sum_{i=1}^{n}(z_i\alpha - 1)\frac{\frac{1}{\gamma}(e^{\gamma x_i}-1)}{[e^{\frac{\beta}{\gamma}(e^{\gamma x_i}-1)} - 1]}, \\ \frac{\partial l_n^*}{\partial\gamma} &= n\bar{x} + \frac{\beta}{\gamma^2}(\sum_{i=1}^{n} e^{\gamma x_i} - n) - \frac{\beta}{\gamma}(\sum_{i=1}^{n} x_i e^{\gamma x_i}) \end{aligned}$$

$$+\sum_{i=1}^{n}(z_i\alpha-1)\frac{\frac{-\beta}{\gamma^2}(e^{\gamma x_i}-1)+\frac{\beta x_i e^{\gamma x_i}}{\gamma}}{[e^{\frac{\beta}{\gamma}(e^{\gamma x_i}-1)}-1]},$$

$$\frac{\partial l_n^*}{\partial\theta} = \sum_{i=1}^{n}\frac{z_i}{\theta}-n\frac{C'(\theta)}{C(\theta)}.$$

From a nonlinear system of equations $U(\boldsymbol{y};\boldsymbol{\Theta})=0$, we obtain the iterative procedure of the EM-algorithm as

$$\hat{\alpha}^{(j+1)}=\frac{-n}{\sum_{i=1}^{n}\hat{z}_i^{(j)}\log[1-e^{\frac{-\hat{\beta}^{(j)}}{\hat{\gamma}^{(j)}}(e^{\hat{\gamma}^{(j)}x_i}-1)}]},$$

$$\hat{\theta}^{(j+1)}-\frac{C(\hat{\theta}^{(j+1)})}{nC'(\hat{\theta}^{(j+1)})}\sum_{i=1}^{n}\hat{z}_i^{(j)}=0,$$

$$\frac{n}{\hat{\beta}^{(j+1)}}-\frac{1}{\hat{\gamma}^{(j)}}(\sum_{i=1}^{n}e^{\hat{\gamma}^{(j)}x_i}-n)+\sum_{i=1}^{n}(\hat{z}_i\hat{\alpha}^{(j)}-1)\frac{\frac{1}{\hat{\gamma}^{(j)}}(e^{\hat{\gamma}^{(j)}x_i}-1)}{[e^{\frac{\hat{\beta}^{(j+1)}}{\hat{\gamma}^{(j)}}(e^{\hat{\gamma}^{(j)}x_i}-1)}-1]}=0,$$

$$n\bar{x}+\frac{\hat{\beta}^{(j)}}{[\hat{\gamma}^{(j+1)}]^2}(\sum_{i=1}^{n}e^{\hat{\gamma}^{(j+1)}x_i}-n)-\frac{\hat{\beta}^{(j)}}{\hat{\gamma}^{(j+1)}}(\sum_{i=1}^{n}x_i e^{\hat{\gamma}^{(j+1)}x_i})$$

$$+\sum_{i=1}^{n}(\hat{z}_i\hat{\alpha}^{(j)}-1)\frac{\frac{-\hat{\beta}^{(j)}}{[\hat{\gamma}^{(j+1)}]^2}(e^{\hat{\gamma}^{(j+1)}x_i}-1)+\frac{\hat{\beta}^{(j)}x_i e^{\hat{\gamma}^{(j+1)}x_i}}{\hat{\gamma}^{(j+1)}}}{[e^{\frac{\hat{\beta}^{(j)}}{\hat{\gamma}^{(j+1)}}(e^{\hat{\gamma}^{(j+1)}x_i}-1)}-1]}=0,$$

where $\hat{\theta}^{(j+1)}$, $\hat{\beta}^{(j+1)}$ and $\hat{\gamma}^{(j+1)}$ are found numerically. Here, for $i=1,2,\ldots,n$, we have that

$$\hat{z}_i^{(j)}=1+\frac{\theta^{*(j)}C''(\theta^{*(j)})}{C'(\theta^{*(j)})},$$

where $\theta^{*(j)}=\hat{\theta}^{(j)}[1-e^{-\frac{\hat{\beta}^{(j)}}{\hat{\gamma}^{(j)}}(e^{\hat{\gamma}^{(j)}x_i}-1)}]^{\hat{\alpha}^{(j)}}$.

We can use the results of [16] to obtain the standard errors of the estimators from the EM-algorithm. Consider $l_c(\boldsymbol{\Theta};\boldsymbol{x})=E(I_c(\boldsymbol{\Theta};\boldsymbol{y})|\boldsymbol{x})$, where $I_c(\boldsymbol{\Theta};\boldsymbol{y})=-[\frac{\partial U(\boldsymbol{y};\boldsymbol{\Theta})}{\partial\boldsymbol{\Theta}}]$ is the 4×4 observed information matrix.If $l_m(\boldsymbol{\Theta};\boldsymbol{x})=Var[U(\boldsymbol{y};\boldsymbol{\Theta})|\boldsymbol{x}]$, then, we obtain the observed information as

$$I(\hat{\boldsymbol{\Theta}};\boldsymbol{x})=l_c(\hat{\boldsymbol{\Theta}};\boldsymbol{x})-l_m(\hat{\boldsymbol{\Theta}};\boldsymbol{x}).$$

The standard errors of the MLEs of the EM-algorithm are the square root of the diagonal elements of the $I(\hat{\boldsymbol{\Theta}};\boldsymbol{x})$. The computation of these matrices are too long and tedious. Therefore, we did not present the details. Reader can see [19] how to calculate these values.

6. Simulation study

We performed a simulation in order to investigate the proposed estimator of α, β, γ and θ of the proposed EM-scheme. We generated 1000 samples of size n from the GGG distribution with $\beta=1$ and $\gamma=0.1$. Then, the averages of estimators (AE), standard error of estimators (SEE), and averages of standard errors (ASE) of MLEs of the EM-algorithm determined though the Fisher information matrix are calculated. The results are given in Table 2. We can find that
(i) convergence has been achieved in all cases and this emphasizes the numerical stability of the EM-algorithm,
(ii) the differences between the average estimates and the true values are almost small,
(iii) the standard errors of the MLEs decrease when the sample size increases.

Table 2. The average MLEs, standard error of estimators and averages of standard errors for the GGG distribution.

	parameter		AE				SEE				ASE			
n	α	θ	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$
50	0.5	0.2	0.491	0.961	0.149	0.204	0.114	0.338	0.265	0.195	0.173	0.731	0.437	0.782
	0.5	0.5	0.540	0.831	0.182	0.389	0.160	0.337	0.260	0.263	0.210	0.689	0.421	0.817
	0.5	0.8	0.652	0.735	0.154	0.684	0.304	0.377	0.273	0.335	0.309	0.671	0.422	0.896
	1.0	0.2	0.988	0.972	0.129	0.206	0.275	0.319	0.191	0.209	0.356	0.925	0.436	0.939
	1.0	0.5	1.027	0.852	0.147	0.402	0.345	0.352	0.226	0.283	0.408	0.873	0.430	0.902
	1.0	0.8	1.210	0.711	0.178	0.745	0.553	0.365	0.230	0.342	0.568	0.799	0.433	0.898
	2.0	0.2	1.969	0.990	0.084	0.216	0.545	0.305	0.151	0.228	0.766	1.135	0.422	0.902
	2.0	0.5	1.957	0.842	0.113	0.487	0.608	0.334	0.192	0.277	0.820	1.061	0.431	0.963
	2.0	0.8	2.024	0.713	0.161	0.756	0.715	0.396	0.202	0.353	1.143	0.873	0.402	0.973
100	0.5	0.2	0.491	0.977	0.081	0.212	0.084	0.252	0.171	0.179	0.125	0.514	0.283	0.561
	0.5	0.5	0.528	0.883	0.109	0.549	0.124	0.275	0.178	0.247	0.155	0.504	0.275	0.567
	0.5	0.8	0.602	0.793	0.136	0.769	0.215	0.323	0.194	0.299	0.220	0.466	0.259	0.522
	1.0	0.2	0.974	0.997	0.102	0.226	0.195	0.242	0.129	0.206	0.251	0.645	0.280	0.767
	1.0	0.5	1.030	0.875	0.113	0.517	0.262	0.291	0.155	0.270	0.298	0.651	0.295	0.843
	1.0	0.8	1.113	0.899	0.117	0.846	0.412	0.342	0.177	0.331	0.400	0.600	0.287	0.781
	2.0	0.2	1.952	0.995	0.138	0.221	0.424	0.237	0.117	0.209	0.524	0.922	0.321	0.992
	2.0	0.5	2.004	0.885	0.110	0.518	0.493	0.283	0.131	0.274	0.601	0.873	0.321	0.966
	2.0	0.8	2.028	0.981	0.104	0.819	0.605	0.350	0.155	0.339	0.816	0.717	0.289	0.946

7. Real examples

In this Section, we consider two real data sets and fit the Gompertz, GGG, GGP, GGB (with $m = 5$), and GGL distributions. The first data set is negatively skewed, and the second data set is positively skewed, and we show that the proposed distributions fit both positively skewed and negatively skewed data well. For each data, the MLE of parameters (with standard deviations) for the distributions are obtained. To test the goodness-of-fit of the distributions, we calculated the maximized log-likelihood, the Kolmogorov-Smirnov (K-S) statistic with its respective p-value, the AIC (Akaike Information Criterion), AICC (AIC with correction), BIC (Bayesian Information Criterion), CM (Cramer-von Mises statistic) and AD (Anderson-Darling statistic) for the six distributions. Here, the significance level is 0.10. To show that the likelihood equations have a unique solution in the parameters, we plot the profile log-likelihood functions of β, γ, α and θ for the six distributions.

First, we consider the data consisting of the strengths of 1.5 cm glass fibers given in [26] and measured at the National Physical Laboratory, England. This data is also studied by [4] and is given in Table 3.

The results are given in Table 5 and show that the GGG distribution yields the best fit among the GGP, GGB, GGL, GG and Gompertz distributions. Also, the GGG, GGP, and GGB distribution are better than GG distribution. The plots of the pdfs (together with the data histogram) and cdfs in Figure 7 confirm this conclusion. Figures 9 show the profile log-likelihood functions of β, γ, α and θ for the six distributions.

As a second example, we consider a data set from [11], who studied the soil fertility influence and the characterization of the biologic fixation of N_2 for the *Dimorphandra wilsonii rizz growth*. For 128 plants, they made measures of the phosphorus concentration in the leaves. This data is also studied by [25] and is given in Table 4. Figures 10 show the profile log-likelihood functions of β, γ, α and θ for the six distributions.

Table 3. The strengths of glass fibers.

0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2.00, 0.74,
1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11,
1.28, 1.42, 1.50, 1.54, 1.60, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29,
1.48, 1.50, 1.55, 1.61, 1.62, 1.66, 1.70, 1.77, 1.84, 0.84, 1.24, 1.30, 1.48, 1.51,
1.55, 1.61, 1.63, 1.67, 1.70, 1.78, 1.89

Table 4. The phosphorus concentration in the leaves.

0.22, 0.17, 0.11, 0.10, 0.15, 0.06, 0.05, 0.07, 0.12, 0.09, 0.23, 0.25, 0.23, 0.24,
0.20, 0.08, 0.11, 0.12, 0.10, 0.06, 0.20, 0.17, 0.20, 0.11, 0.16, 0.09, 0.10, 0.12,
0.12, 0.10, 0.09, 0.17, 0.19, 0.21, 0.18, 0.26, 0.19, 0.17, 0.18, 0.20, 0.24, 0.19,
0.21, 0.22, 0.17, 0.08, 0.08, 0.06, 0.09, 0.22, 0.23, 0.22, 0.19, 0.27, 0.16, 0.28,
0.11, 0.10, 0.20, 0.12, 0.15, 0.08, 0.12, 0.09, 0.14, 0.07, 0.09, 0.05, 0.06, 0.11,
0.16, 0.20, 0.25, 0.16, 0.13, 0.11, 0.11, 0.11, 0.08, 0.22, 0.11, 0.13, 0.12, 0.15,
0.12, 0.11, 0.11, 0.15, 0.10, 0.15, 0.17, 0.14, 0.12, 0.18, 0.14, 0.18, 0.13, 0.12,
0.14, 0.09, 0.10, 0.13, 0.09, 0.11, 0.11, 0.14, 0.07, 0.07, 0.19, 0.17, 0.18, 0.16,
0.19, 0.15, 0.07, 0.09, 0.17, 0.10, 0.08, 0.15, 0.21, 0.16, 0.08, 0.10, 0.06, 0.08,
0.12, 0.13

The results are given in Table 6. Since the estimation of parameter θ for GGP, GGB, and GGL is close to zero, the estimations of parameters for these distributions are equal to the estimations of parameters for GG distribution. In fact, The limiting distribution of GGPS when $\theta \to 0^+$ is a GG distribution (see Proposition 2.2). Therefore, the value of maximized log-likelihood, $\log(L)$, are equal for these four distributions. The plots of the pdfs (together with the data histogram) and cdfs in Figure 8 confirm these conclusions. Note that the estimations of parameters for GGG distribution are not equal to the estimations of parameters for GG distribution. But the $\log(L)$'s are equal for these distributions. However, from Table 6 also we can conclude that the GG distribution is simpler than other distribution because it has three parameter but GGG, GGP, GGB, and GGL have four parameter. Note that GG is a special case of GGPS family.

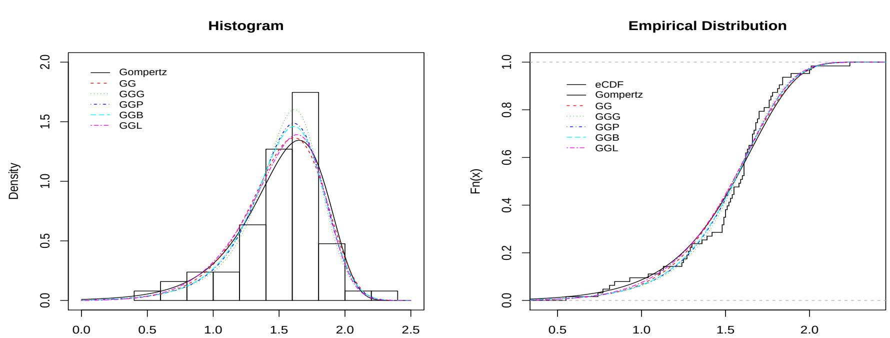


Figure 7. Plots (pdf and cdf) of fitted Gompertz, generalized Gompertz, GGG, GGP, GGB and GGL distributions for the first data set.

Table 5. Parameter estimates (with std.), K-S statistic, p-value, AIC, AICC and BIC for the first data set.

	Distribution					
	Gompertz	GG	GGG	GGP	GGB	GGL
$\hat{\beta}$	0.0088	0.0356	0.7320	0.1404	0.1032	0.1705
$s.e.(\hat{\beta})$	0.0043	0.0402	0.2484	0.1368	0.1039	0.2571
$\hat{\gamma}$	3.6474	2.8834	1.3499	2.1928	2.3489	2.1502
$s.e.(\hat{\gamma})$	0.2992	0.6346	0.3290	0.5867	0.6010	0.7667
$\hat{\alpha}$	—	1.6059	2.1853	1.6205	1.5999	2.2177
$s.e.(\hat{\alpha})$	—	0.6540	1.2470	0.9998	0.9081	1.3905
$\hat{\theta}$	—	—	0.9546	2.6078	0.6558	0.8890
$s.e.(\hat{\theta})$	—	—	0.0556	1.6313	0.5689	0.2467
$-\log(L)$	14.8081	14.1452	12.0529	13.0486	13.2670	13.6398
K-S	0.1268	0.1318	0.0993	0.1131	0.1167	0.1353
p-value	0.2636	0.2239	0.5629	0.3961	0.3570	0.1992
AIC	33.6162	34.2904	32.1059	34.0971	34.5340	35.2796
AICC	33.8162	34.6972	32.7956	34.78678	35.2236	35.9692
BIC	37.9025	40.7198	40.6784	42.6696	43.1065	43.8521
CM	0.1616	0.1564	0.0792	0.1088	0.1172	0.1542
AD	0.9062	0.8864	0.5103	0.6605	0.7012	0.8331

Table 6. Parameter estimates (with std.), K-S statistic, p-value, AIC, AICC and BIC for the second data set.

	Distribution					
	Gompertz	GG	GGG	GGP	GGB	GGL
$\hat{\beta}$	1.3231	13.3618	10.8956	13.3618	13.3618	13.3618
$s.e.(\hat{\beta})$	0.2797	4.5733	8.4255	5.8585	6.3389	7.3125
$\hat{\gamma}$	15.3586	3.1500	4.0158	3.1500	3.1500	3.1500
$s.e.(\hat{\gamma})$	1.3642	2.1865	3.6448	2.4884	2.6095	2.5024
$\hat{\alpha}$	—	6.0906	5.4236	6.0906	6.0906	6.0905
$s.e.(\hat{\alpha})$	—	2.4312	2.8804	2.6246	2.7055	2.8251
$\hat{\theta}$	—	—	-0.3429	1.0×10^{-8}	1.0×10^{-8}	1.0×10^{-8}
$s.e.(\hat{\theta})$	—	—	1.2797	0.8151	0.2441	0.6333
$-\log(L)$	-184.597	-197.133	-197.181	-197.133	-197.133	-197.133
K-S	0.1169	0.0923	0.0898	0.0923	0.0923	0.0923
p-value	0.0602	0.2259	0.2523	0.2259	0.2259	0.2259
AIC	-365.194	-388.265	-386.362	-386.265	-386.265	-386.265
AICC	-365.098	-388.072	-386.0371	-385.940	-385.940	-385.940
BIC	-359.490	-379.709	-374.954	-374.857	-374.857	-374.857
CM	0.3343	0.1379	0.1356	0.1379	0.1379	0.1379
AD	2.3291	0.7730	0.7646	0.7730	0.7730	0.7730

Appendix

A. We demonstrate those parameter intervals for which the hazard function is decreasing, increasing and bathtub shaped, and in order to do so, we follow closely a theorem given

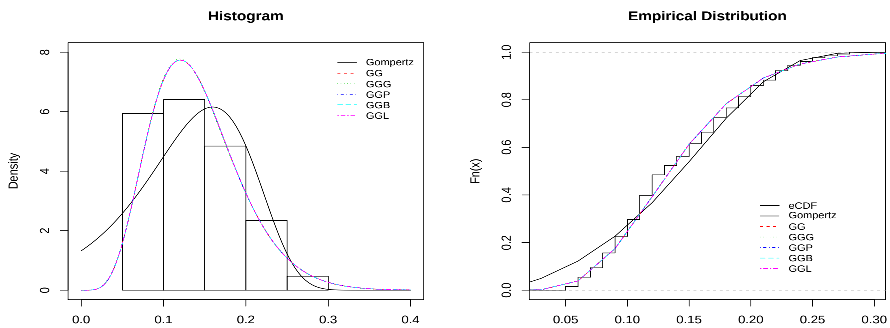


Figure 8. Plots (pdf and cdf) of fitted Gompertz, generalized Gompertz, GGG, GGP, GGB and GGL distributions for the second data set.

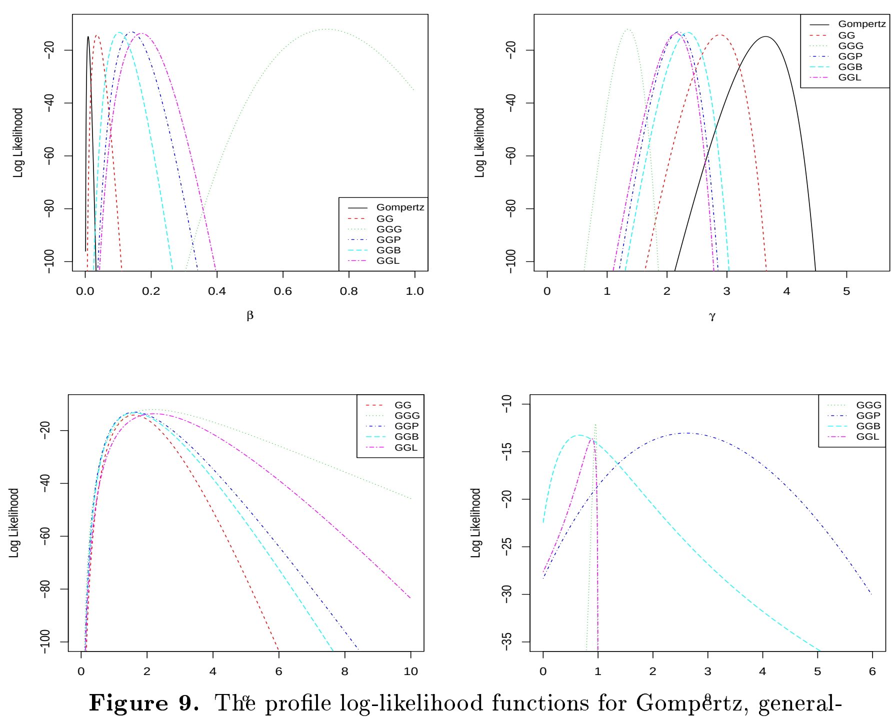


Figure 9. The profile log-likelihood functions for Gompertz, generalized Gompertz, GGG, GGP, GGB and GGL distributions for the first data set.

by [12]. Define the function $\tau(x) = \frac{-f'(x)}{f(x)}$ where $f'(x)$ denotes the first derivative of $f(x)$ in (2.3). To simplify, we consider $u = 1 - \exp(\frac{-\theta}{\gamma}(e^{\gamma x} - 1))$.

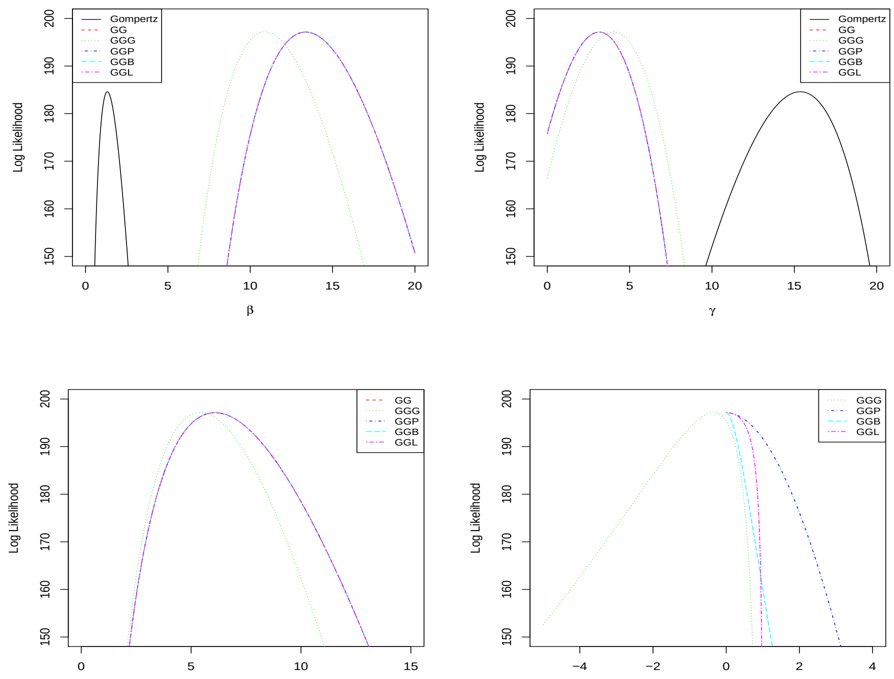


Figure 10. The profile log-likelihood functions for Gompertz, generalized Gompertz, GGG, GGP, GGB and GGL distributions for the second data set.

A.1. Consider the GGG hazard function in (4.2), then we define

$$\tau(u) = \frac{-f'(u)}{f(u)} = \frac{1-\alpha}{u} + \frac{2\alpha\theta u^{\alpha-1}}{1-\theta u^{\alpha}}.$$

If $\alpha \geq 1$, then $\tau'(u) > 0$, and $h(.)$ is an increasing function. If $0 < \alpha < 1$, then

$$\lim_{u\to 0} \tau'(u) = -\infty, \quad \lim_{u\to 1} \tau'(u) = \frac{2\alpha\theta^2}{(1-\theta)^2} + (\alpha-1)(1-\frac{1}{(1-\theta)^2}) > 0.$$

Since the limits have different signs, the equation $\tau'(u) = 0$ has at least one root. Also, we can show that $\tau''(u) > 0$. Therefore, the equation $\tau'(u) = 0$ has one root. Thus the hazard function is decreasing and bathtub shaped in this case.

A.2. The GGP hazard rate is given by $h(u) = \theta\alpha\beta u^{\alpha-1} e^{\theta u^{\alpha}}/(e^{\theta} - e^{\theta u^{\alpha}})$. We define $\eta(u) = \log[h(u)]$. Then, its first derivative is

$$\eta'(u) = \frac{\alpha-1}{u} + \alpha\theta e^{\theta}\frac{u^{\alpha-1}}{e^{\theta} - e^{\theta u^{\alpha}}}.$$

It is clearly for $\alpha \geq 1$, $\eta'(u) > 0$ and $h(u)$ is increasing function. If $0 < \alpha < 1$, then

$$\lim_{u\to 0} \eta'(u) = -\infty, \quad \lim_{u\to 1} \eta'(u) = 0,$$

So the equation $\tau'(u) = 0$ has at least one root. Also, we can show that $\tau''(u) > 0$. It implies that equation $\eta'(u) = 0$ has a one root and the hazard rate increase and bathtub shaped.

B.

B.1. Let $w_1(\alpha) = \sum_{i=1}^{n} \frac{\theta t_i^{\alpha} \log(t_i) C''(\theta t_i^{\alpha})}{C'(\theta t_i^{\alpha})} = \frac{\partial}{\partial \alpha} \sum_{i=1}^{n} \log(C'(\theta t_i^{\alpha}))$. For GGG,

$$w_1(\alpha) = 2\theta \sum_{i=1}^{n} \frac{t_i^{\alpha} \log t_i}{1 - \theta t_i^{\alpha}}, \qquad \frac{\partial w_1(\alpha)}{\partial \alpha} = 2\theta \sum_{i=1}^{n} t_i^{\alpha} [\frac{\log t_i}{1 - \theta t_i^{\alpha}}]^2 > 0.$$

For GGP,

$$w_1(\alpha) = \theta \sum_{i=1}^{n} t_i^{\alpha} \log t_i, \qquad \frac{\partial w_1(\alpha)}{\partial \alpha} = \theta \sum_{i=1}^{n} t_i^{\alpha} [\log t_i]^2 > 0.$$

For GGL,

$$w_1(\alpha) = \theta \sum_{i=1}^{n} \frac{t_i^{\alpha} \log t_i}{1 - \theta t_i^{\alpha}}, \qquad \frac{\partial w_1(\alpha)}{\partial \alpha} = \theta \sum_{i=1}^{n} t_i^{\alpha} [\frac{\log t_i}{1 - \theta t_i^{\alpha}}]^2 > 0.$$

For GGB,

$$w_1(\alpha) = (m-1)\theta \sum_{i=1}^{n} \frac{t_i^{\alpha} \log t_i}{1 + \theta t_i^{\alpha}}, \qquad \frac{\partial w_1(\alpha)}{\partial \alpha} = (m-1)\theta \sum_{i=1}^{n} t_i^{\alpha} [\frac{\log t_i}{1 + \theta t_i^{\alpha}}]^2 > 0.$$

Therefore, $w_1(\alpha)$ is strictly increasing in α and

$$\lim_{\alpha \to 0^+} \mathrm{g}_1(\alpha; \beta, \gamma, \theta, x) = \infty, \qquad \lim_{\alpha \to \infty} \mathrm{g}_1(\alpha; \beta, \gamma, \theta, x) = \sum_{i=1}^{n} \log(t_i).$$

Also,

$$\mathrm{g}_1(\alpha; \beta, \gamma, \theta, x) < \frac{n}{\alpha} + \sum_{i=1}^{n} \log(t_i), \quad \mathrm{g}_1(\alpha; \beta, \gamma, \theta, x) > \frac{n}{\alpha} + (\frac{\theta C''(\theta)}{C'(\theta)} + 1) \sum_{i=1}^{n} \log(t_i).$$

Hence, $\mathrm{g}_1(\alpha; \beta, \gamma, \theta, x) < 0$ when $\frac{n}{\alpha} + \sum_{i=1}^{n} \log(t_i) < 0$, and $\mathrm{g}_1(\alpha; \beta, \gamma, \theta, x) > 0$ when $\frac{n}{\alpha} + (\frac{\theta C''(\theta)}{C'(\theta)} + 1) \sum_{i=1}^{n} \log(t_i) > 0$. The proof is completed.

B.2. It can be easily shown that

$$\lim_{\beta \to 0^+} \mathrm{g}_2(\beta; \alpha, \gamma, \theta, x) = \infty, \qquad \lim_{\beta \to \infty} \mathrm{g}_2(\beta; \alpha, \gamma, \theta, x) = \frac{-1}{\gamma} \sum_{i=1}^{n} (e^{\gamma x_i} - 1).$$

Since the limits have different signs, the equation $\mathrm{g}_2(\beta; \alpha, \gamma, \theta, x) = 0$ has at least one root with respect to β for fixed values α, γ and θ. The proof is completed.

B.3. a) For GGP, it is clear that

$$\lim_{\theta \to 0} \mathrm{g}_3(\theta; \alpha, \beta, \gamma, x) = \sum_{i=1}^{n} t_i^{\alpha} - \frac{n}{2}, \qquad \lim_{\theta \to \infty} \mathrm{g}_3(\theta; \alpha, \beta, \gamma, x) = -\infty.$$

Therefore, the equation $\mathrm{g}_3(\theta; \alpha, \beta, \gamma, x) = 0$ has at least one root for $\theta > 0$, if $\sum_{i=1}^{n} t_i^{\alpha} - \frac{n}{2} > 0$ or $\sum_{i=1}^{n} t_i^{\alpha} > \frac{n}{2}$.
b) For GGG, it is clear that

$$\lim_{\theta \to \infty} \mathrm{g}_3(\theta; \alpha, \beta, \gamma, x) = -\infty, \qquad \lim_{\theta \to 0^+} \mathrm{g}_3(\theta; \alpha, \beta, \gamma, x) = -n + 2 \sum_{i=1}^{n} t_i^{\alpha}.$$

Therefore, the equation $\mathrm{g}_3(\theta, \beta, \gamma, x) = 0$ has at least one root for $0 < \theta < 1$, if $-n + 2\sum_{i=1}^{n} t_i^{\alpha} > 0$ or $\sum_{i=1}^{n} t_i^{\alpha} > \frac{n}{2}$.
For GGL, it is clear that

$$\lim_{\theta \to 0} \mathrm{g}_3(\theta; \alpha, \beta, \gamma, x) = \sum_{i=1}^{n} t_i^{\alpha} - \frac{n}{2}, \qquad \lim_{\theta \to 1} \mathrm{g}_3(\theta; \alpha, \beta, \gamma, x) = -\infty.$$

Therefore, the equation $g_3(\theta;\alpha,\beta,\gamma,x)=0$ has at least one root for $0<\theta<1$, if $\sum_{i=1}^{n} t_i^{\alpha}-\frac{n}{2}>0$ or $\sum_{i=1}^{n} t_i^{\alpha}>\frac{n}{2}$.
For GGB, it is clear that

$$\lim_{p\to 0} g_3(p;\alpha,\beta,\gamma,x)=\sum_{i=1}^{n} t_i^{\alpha}(m-1)-\frac{n(m-1)}{2},$$
$$\lim_{p\to 0} g_3(p;\alpha,\beta,\gamma,x)=\sum_{i=1}^{n}\frac{-m+1+mt_i^{\alpha}}{t_i},$$

Therefore, the equation $g_3(p;\alpha,\beta,\gamma,x)=0$ has at least one root for $0<p<1$, if $\sum_{i=1}^{n} t_i^{\alpha}(m-1)-\frac{n(m-1)}{2}>0$ and $\sum_{i=1}^{n}\frac{-m+1+mt_i^{\alpha}}{t_i^{\alpha}}<0$ or $\sum_{i=1}^{n} t_i^{\alpha}>\frac{n}{2}$ and $\sum_{i=1}^{n} t_i^{-\alpha}>\frac{nm}{1-m}$.

C. Consider $t_i=1-e^{-\frac{\beta}{\gamma}(e^{\gamma x_i}-1)}$. Then, the elements of 4×4 observed information matrix $I_n(\Theta)$ are given by

$$\begin{aligned}
I_{\alpha\alpha} &= \frac{\partial^2 l_n}{\partial\alpha^2}=\frac{-n}{\alpha^2}+\theta\sum_{i=1}^{n} t_i^{\alpha}[\log(t_i)]^2[\frac{C''(\theta t_i^{\alpha})}{C'(\theta t_i^{\alpha})}\\
&\quad +\theta t_i^{\alpha}\frac{C'''(\theta t_i^{\alpha})C'(\theta t_i^{\alpha})-(C''(\theta t_i^{\alpha}))^2}{(C'(\theta t_i^{\alpha}))^2}],\\
I_{\alpha\beta} &= \frac{\partial^2 l_n}{\partial\alpha\partial\beta}=\sum_{i=1}^{n}[\frac{e^{\gamma x_i}-1}{\gamma}]+\frac{\theta}{\gamma}\sum_{i=1}^{n} t_i^{\alpha}[e^{\gamma x_i}-1][(\alpha\log(t_i)+1)\frac{C''(\theta t_i^{\alpha})}{C'(\theta t_i^{\alpha})}\\
&\quad +\alpha\theta t_i^{\alpha}\log(t_i)\frac{C'''(\theta t_i^{\alpha})C'(\theta t_i^{\alpha})-(C''(\theta t_i^{\alpha}))^2}{(C'(\theta t_i^{\alpha}))^2}],\\
I_{\alpha\gamma} &= \frac{\partial^2 l_n}{\partial\alpha\partial\gamma}=\beta\sum_{i=1}^{n}[\frac{e^{\gamma x_i}(\gamma x_i-1)+1}{\gamma^2}]\\
&\quad +\frac{\theta\beta}{\gamma^2}\sum_{i=1}^{n}[e^{\gamma x_i(\gamma x_i-1)}+1][(\alpha\log(t_i)+1)\frac{C''(\theta t_i^{\alpha})}{C'(\theta t_i^{\alpha})}\\
&\quad +\alpha\theta t_i^{\alpha}\log(t_i)\frac{C'''(\theta t_i^{\alpha})C'(\theta t_i^{\alpha})-(C''(\theta t_i^{\alpha}))^2}{(C'(\theta t_i^{\alpha}))^2}],\\
I_{\alpha\theta} &= \frac{\partial^2 l_n}{\partial\alpha\partial\theta}=\sum_{i=1}^{n} t_i^{\alpha}\log(t_i)[\frac{C''(\theta t_i^{\alpha})}{C'(\theta t_i^{\alpha})}+\theta t_i^{\alpha}\frac{C'''(\theta t_i^{\alpha})C'(\theta t_i^{\alpha})-(C''(\theta t_i^{\alpha}))^2}{(C'(\theta t_i^{\alpha}))^2}],\\
I_{\beta\beta} &= \frac{\partial^2 l_n}{\partial\beta^2}=\frac{-n}{\beta^2}+\theta\alpha^2\sum_{i=1}^{n} t_i^{\alpha}[\frac{e^{\gamma x_i}-1}{\gamma}]^2[\frac{C''(\theta t_i^{\alpha})}{C'(\theta t_i^{\alpha})}\\
&\quad +\theta t_i^{\alpha}\frac{C'''(\theta t_i^{\alpha})C'(\theta t_i^{\alpha})-(C''(\theta t_i^{\alpha}))^2}{(C'(\theta t_i^{\alpha}))^2}],\\
I_{\beta\gamma} &= \frac{\partial^2 l_n}{\partial\beta\partial\gamma}=\frac{(\alpha-2)}{\gamma^2}\sum_{i=1}^{n}(e^{\gamma x_i}(\gamma x_i-1)+1)\\
&\quad +\alpha\theta\sum_{i=1}^{n}\frac{t_i}{\gamma^2}(e^{\gamma x_i}(\gamma x_i-1)+1)[\frac{C''(\theta t_i^{\alpha})}{C'(\theta t_i^{\alpha})}\\
&\quad +\frac{\beta^2}{\gamma}(e^{\gamma x_i}-1)\frac{C'''(\theta t_i^{\alpha})C'(\theta t_i^{\alpha})-(C''(\theta t_i^{\alpha}))^2}{(C'(\theta t_i^{\alpha})))^2}]
\end{aligned}$$

$$
\begin{aligned}
I_{\beta\theta} &= \frac{\partial^2 l_n}{\partial\beta\partial\theta} = \sum_{i=1}^{n} t_i^{2\alpha}[\frac{C'''(\theta t_i^\alpha)C'(\theta t_i^\alpha)-(C''(\theta t_i^\alpha))^2}{(C'(\theta t_i^\alpha))^2}],\\
I_{\gamma\gamma} &= \frac{\partial^2 l_n}{\partial\gamma^2} = \frac{2\beta}{\gamma^3}\sum_{i=1}^{n}[e^{\gamma x_i}(\gamma x_i-1)+1]\\
&+(\alpha-1)\beta\sum_{i=1}^{n}[\frac{-2}{\gamma^3}(e^{\gamma x_i}(\gamma x_i-1)+1)+\frac{x_i^2e^{\gamma x_i}}{\gamma^3}]\\
&+\alpha\beta\theta\sum_{i=1}^{n}[\frac{-2}{\gamma^3}(e^{\gamma x_i}(\gamma x_i-1)+1)t_i^\alpha\frac{C''(\theta t_i^\alpha)}{C'(\theta t_i^\alpha)}+\frac{t_i^\alpha x_i^2e^{\gamma x_i}}{\gamma}\frac{C''(\theta t_i^\alpha)}{C'(\theta t_i^\alpha)}\\
&+\frac{\alpha\beta t_i^\alpha}{\gamma^4}(e^{\gamma x_i}(\gamma x_i-1)+1)^2\frac{C''(\theta t_i^\alpha)}{C'(\theta t_i^\alpha)}\\
&+\frac{\alpha\beta t_i^{2\alpha}}{\gamma^4}(e^{\gamma x_i}(\gamma x_i-1)+1)^2\frac{C'''(\theta t_i^\alpha)C'(\theta t_i^\alpha)-(C''(\theta t_i^\alpha))^2}{(C'(\theta t_i^\alpha)))^2}],\\
I_{\theta\gamma} &= \frac{\partial^2 l_n}{\partial\theta\partial\gamma} = \alpha\beta\sum_{i=1}^{n}\frac{t_i^\alpha}{\gamma^2}[e^{\gamma x_i}(\gamma x_i-1)+1][\frac{C''(\theta t_i^\alpha)}{C'(\theta t_i^\alpha)}\\
&+\theta t_i^\alpha\frac{C'''(\theta t_i^\alpha)C'(\theta t_i^\alpha)-(C''(\theta t_i^\alpha))^2}{(C'(\theta t_i^\alpha))^2}],\\
I_{\theta\theta} &= \frac{\partial^2 l_n}{\partial\theta^2} = \frac{-n}{\theta^2}+\sum_{i=1}^{n}t_i^{2\alpha}[\frac{C'''(\theta t_i^\alpha)C'(\theta t_i^\alpha)-(C''(\theta t_i^\alpha))^2}{(C'(\theta t_i^\alpha))^2}]\\
&-n[\frac{C''(\theta)C'(\theta)-(C'(\theta))^2}{(C'(\theta))^2}],
\end{aligned}
$$

Acknowledgements

The authors would like to thank the anonymous referees for many helpful comments and suggestions.

References

[1] Adamidis, K., Dimitrakopoulou, T. and Loukas, S. *On an extension of the exponential-geometric distribution*, Statistics and Probability Letters, **73** (3), 259-269, 2005.

[2] Adamidis, K. and Loukas, S. *A lifetime distribution with decreasing failure rate*, Statistics and Probability Letters, **39** (1), 35-42, 1998.

[3] Barreto-Souza, W., Morais, A. L. and Cordeiro, G. M. *The Weibull-geometric distribution*, Journal of Statistical Computation and Simulation, **81** (5), 645-657, 2011.

[4] Barreto-Souza, W., Santos, A. H. S. and Cordeiro, G. M. *The beta generalized exponential distribution*, Journal of Statistical Computation and Simulation, **80** (2), 159-172, 2010.

[5] Cancho, V. G., Louzada-Neto, F. and Barriga, G. D. C. *The Poisson-exponential lifetime distribution*, Computational Statistics and Data Analysis, **55** (1), 677-686, 2011.

[6] Casella, G. and Berger, R. *Statistical Inference*, Duxbury, Pacific Grove, California, USA, 2001.

[7] Chahkandi, M. and Ganjali, M. *On some lifetime distributions with decreasing failure rate*, Computational Statistics and Data Analysis, **53** (12), 4433-4440, 2009.

[8] Dempster, A. P., Laird, N. M. and Rubin, D. B. *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the Royal Statistical Society, Series B (Methodological), **39** (1), 1-38, 1977.

[9] El-Gohary, A., Alshamrani, A., and Al-Otaibi, A. N. *The generalized Gompertz distribution*, Applied Mathematical Modelling, **37** (1-2), 13-24, 2013.

[10] Flores, J., Borges, P., Cancho, V. G. and Louzada, F. *The complementary exponential power series distribution*, Brazilian Journal of Probability and Statistics, **27** (4), 565-584, 2013.

[11] Fonseca M. and Franca, M. *A influência da fertilidade do solo e caracterização da fixação biológica de N_2 para o crescimento de*, Dimorphandra wilsonii rizz. Master's thesis, Universidade Federal de Minas Gerais, 2007.

[12] Glaser, R. E. *Bathtub and related failure rate characterizations*, Journal of the American Statistical Association, **75** (371), 667-672, 1980.

[13] Gupta, R. D. and Kundu, D. *Generalized exponential distributions*, Australian & New Zealand Journal of Statistics, **41** (2), 173-188, 1999.

[14] Kuş, C. *A new lifetime distribution*, Computational Statistics and Data Analysi, **51** (9), 4497-4509, 2007.

[15] Lawless, J. F. *Statistical Models and Methods for Lifetime Data*, Wiley-Interscience, second edition, 2003.

[16] Louis, T. A. *Finding the observed information matrix when using the EM algorithm*, Journal of the Royal Statistical Society. Series B (Methodological), **44** (2), 226-233, 1982.

[17] Louzada, F., Roman, M. and Cancho, V. G. *The complementary exponential geometric distribution: Model, properties, and a comparison with its counterpart*, Computational Statistics and Data Analysis, **55** (8), 2516-2524, 2011.

[18] Louzada-Neto, F., Cancho, V. G. and Barriga, G. D. C. *The Poisson–exponential distribution: a Bayesian approach*, Journal of Applied Statistics, **38** (6), 1239-1248, 2011.

[19] Mahmoudi, E. and Jafari, A. A. *Generalized exponential-power series distributions*, Computational Statistics and Data Analysis, **56** (12), 4047-4066, 2012.

[20] Mahmoudi, E. and Jafari, A. A. *The compound class of linear failure rate-power series distributions: model, properties and applications*, arXiv preprint arXiv:1402.5282, 2014.

[21] Marshall, A. W. and Olkin, I. *A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families.* Biometrika, **84** (3), 641-652, 1997.

[22] Morais, A. L. and Barreto-Souza, W. *A compound class of Weibull and power series distributions*, Computational Statistics and Data Analysis, **55** (3), 1410-1425, 2011.

[23] Noack, A. *A class of random variables with discrete distributions*, The Annals of Mathematical Statistics, **21** (1), 127-132, 1950.

[24] Shannon, C. *A mathematical theory of communication*, Bell System Technical Journal, **27**, 379-432, 1948.

[25] Silva, R. B., Bourguignon, M., Dias, C. R. B. and Cordeiro, G. M. *The compound class of extended Weibull power series distributions*, Computational Statistics and Data Analysis, **58**, 352-367, 2013.

[26] Smith, R. L. and Naylor, J. C. *A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution*. Applied Statistics, **36** (3), 358-369, 1987.

[27] Tahmasbi, R. and Rezaei, S. *A two-parameter lifetime distribution with decreasing failure rate*, Computational Statistics and Data Analysis, **52**, (8), 3889-3901, 2008.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1605 – 1620

Comparing of some estimation methods for parameters of the Marshall-Olkin generalized exponential distribution under progressive Type-I interval censoring

Hamzeh Torabi*† , Masumeh Forooghi ‡ and Hossein Nadeb §

Abstract

In this paper, we estimate the parameters of the Marshall-Olkin generalized exponential distribution under progressive Type-I interval censoring based on maximum likelihood, moment method and probability plot. A simulation study is conducted to compare these estimates in terms of mean squared errors and biases. Finally, these estimate methods are applied to a real data set based on patients with breast cancer in order to demonstrate the applicabilities.

Keywords: EM algorithm, Generalized exponential distribution, Maximum likelihood estimate, Method of moments, Type-I interval censoring.

2000 AMS Classification: 65C05, 62N01, 62N02.

Received : 06.12.2014 *Accepted* : 28.07.2015 *Doi* : 10.15672/HJMS.20158112561

1. Introduction

Aggarwala [2001] introduced Type-I interval and progressive censoring and developed the statistical inference for the exponential distribution based on progressively Type-I interval censored data. Ng and Wang [2009] introduced the concept of progressive Type-I interval censoring to the Weibull distribution and compared many different estimation methods for two parameters in the Weibull distribution via simulation.

*Department of Statistics, Yazd University, 89175-741, Yazd, IRAN,
Email: htorabi@yazd.ac.ir

†Corresponding Author.

‡Department of Statistics, Yazd University, 89175-741, Yazd, IRAN,
Email: forughi87@gmail.com

§Department of Statistics, Yazd University, 89175-741, Yazd, IRAN,
Email: honadeb@yahoo.com

The generalized exponential (GE) distribution has the following probability density function (pdf)

$$f(t;\alpha,\lambda) = \alpha\lambda(1-e^{-\lambda t})^{\alpha-1}e^{-\lambda t},$$

where $t>0$, $\alpha>0$ and $\lambda>0$. The cumulative distribution and the hazard rate function of the GE distribution are as follows:

$$F(t;\alpha,\lambda) = (1-e^{-\lambda t})^{\alpha},$$

and

$$h(t;\alpha,\lambda) = \frac{\alpha\lambda(1-e^{-\lambda t})^{\alpha-1}e^{-\lambda t}}{1-(1-e^{-\lambda t})^{\alpha}},$$

where $t>0$; The GE distribution was introduced by Gupta and Kunda [2001]. Recently, Chen and Lio [2010] introduced the concept of progressive Type-I interval censoring for the generalized exponential distribution and compared many different estimation methods for the parameters of the distribution via a simulation study.

The Marshall-Olkin generalized exponential (MOGE) distribution was first proposed by Marshall and Olkin [1997] and extensively discussed by Alice and Jose [1999]. The PDF of the MOGE distribution with the parameters λ and α is

$$f(t;\alpha,\lambda) = \frac{\alpha\lambda e^{-\lambda t}}{(1-(1-\alpha)e^{-\lambda t})^2}, \quad t>0, \quad 0<\alpha\leq 1, \lambda>0. \tag{1.1}$$

Also, the distribution function and the hazard rate function of the MOGE distribution are as follows:

$$F(t;\alpha,\lambda) = \frac{1-e^{-\lambda t}}{1-(1-\alpha)e^{-\lambda t}}, \tag{1.2}$$

and

$$h(t;\alpha,\lambda) = \frac{\lambda}{1-(1-\alpha)e^{-\lambda t}},$$

where $t>0$. Note that if $\alpha=1$, the MOGE distribution reduces to the conventional exponential distribution. Plots of the density functions, distribution functions and hazard rate functions for different values of α and λ are given in figures 1, 2 and 3, respectively.

The first two moments and variance of the MOGE distribution are given by

$$E[T] = \frac{\alpha\log(\alpha)}{(\alpha-1)\lambda}, \tag{1.3}$$

$$E[T^2] = \frac{2\alpha\text{PolyLog}[2,1-\alpha]}{(1-\alpha)\lambda^2}, \tag{1.4}$$

$$Var[T] = \frac{-\alpha\Big(\alpha\log(\alpha)^2+2(\alpha-1)\text{PolyLog}[2,1-\alpha]\Big)}{(\alpha-1)^2\lambda^2},$$

where

$$\text{PolyLog}[2,1-\alpha] = \sum_{k=1}^{\infty}\frac{(1-\alpha)^k}{k^2}.$$

In this paper, we study the maximum likelihood estimates, estimates via moment methods and estimates via probability plot for two parameters of the MOGE distribution under the progressive Type-I interval censoring. Section 2 introduces the progressive Type-I interval censoring for the MOGE distribution. In Section 3, some methods for parameters estimation are given. In Section 4, a simulation study is conducted to compare the performances of these estimation methods based on the mean squared error

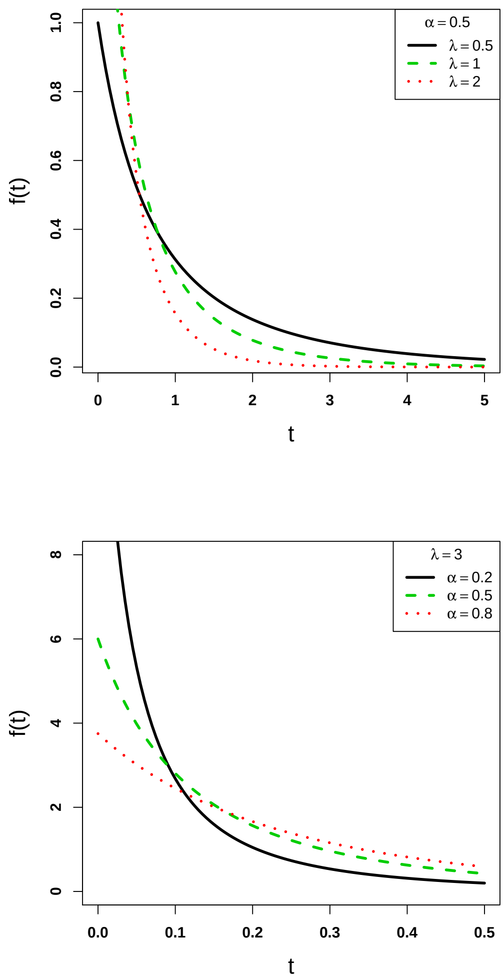
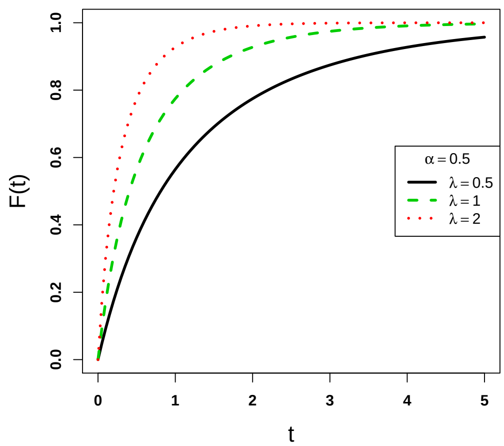


Figure 1. Plots of density functions for different values of α and λ

(MSE) and bias. Finally, a numerical example for a real data set is considered and some discussions and conclusions are given.

2. Progressively Type-I interval censored data

Suppose that n items are placed on a life testing problem simultaneously at time $t_0 = 0$ under inspection at m pre-specified times $t_1 < t_2 < \ldots < t_m$ where t_m is the scheduled time to terminate the experiment. At the ith inspection time, t_i, the number, X_i, of failures within $(t_{i-1}, t_i]$ is recorded and R_i surviving items are randomly removed from the life testing, for $i = 1, \ldots, m$. It is obvious that the number of surviving items at the time t_i is $Y_i = n - \sum_{j=1}^{i} X_j - \sum_{j=1}^{i-1} R_j$. Since Y_i is a random variable and the exact



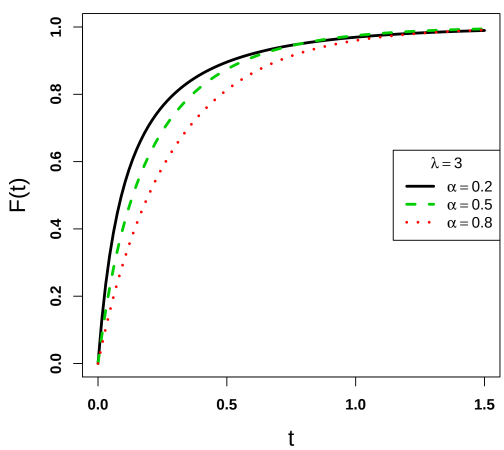


Figure 2. Plots of distributions functions for different values of α and λ

number of items withdrawn should not be greater than Y_i at time schedule t_i, R_i could be determined by the pre-specified percentage of the remaining surviving units at t_i for given $i = 1, 2, \ldots, m$. Also, given pre-specified percentage values, $p_1, \ldots, p_{m-1}$ and $p_m = 1$, for withdrawing at $t_1 < t_2 < \ldots < t_m$, respectively, $R_i = \lfloor p_i y_i \rfloor$ at each inspection time t_i where $i = 1, 2, \ldots, m$. Therefore, a progressively Type-I interval censored sample can be denoted as (X_i, R_i, t_i), $i = 1, 2, \ldots m$, where sample size is $n = \sum_{i=1}^{m}(X_i + R_i)$. Note that if $R_i = 0$, $i = 1, 2, \ldots, m-1$, then the progressively Type-I interval censored sample is a Type-I interval censored sample, $X_1, X_2, \ldots, X_m, X_{m+1} = R_m$.

Let a progressively Type-I interval censored sample be collected as described above, beginning with a random sample of n units with a continuous life time distribution

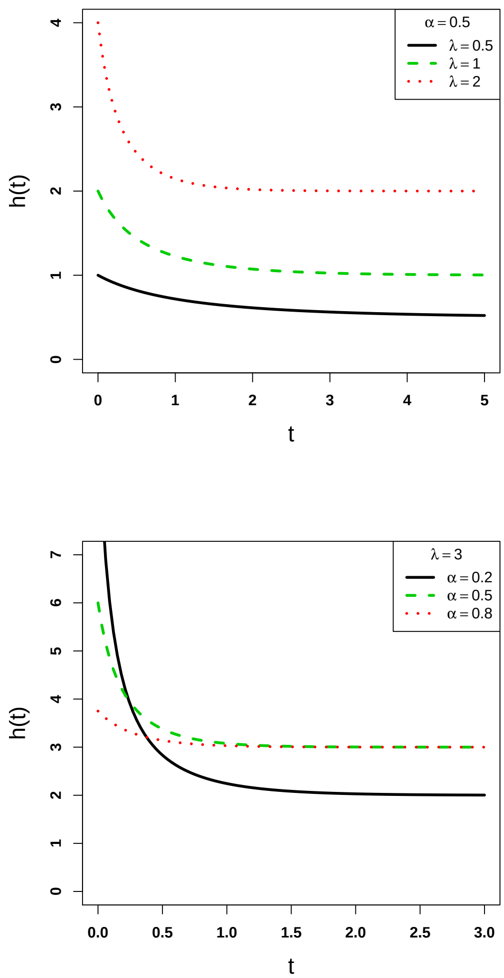


Figure 3. Plots of hazard functions for different values of α and λ

function $F(.;\boldsymbol{\theta})$. Then, based on the observed data, the likelihood function will be as follows:

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^{m} \left[F(t_i;\boldsymbol{\theta}) - F(t_{i-1};\boldsymbol{\theta})\right]^{X_i} \left[1 - F(t_i;\boldsymbol{\theta})\right]^{R_i}.$$

3. Some parameter estimation methods

In this section, we give some estimation methods for the parameters of the MOGE distribution.

3.1. Maximum likelihood estimation. Suppose a progressive Type-I interval censored sample is collected for the MOGE distribution. Using (1.2), the likelihood function is

$$L(\alpha,\lambda) \propto \prod_{i=1}^{m}\left[\frac{1-e^{-\lambda t_i}}{1-(1-\alpha)e^{-\lambda t_i}} - \frac{1-e^{-\lambda t_{i-1}}}{1-(1-\alpha)e^{-\lambda t_{i-1}}}\right]^{X_i}\left[\frac{\alpha e^{-\lambda t_i}}{1-(1-\alpha)e^{-\lambda t_i}}\right]^{R_i},$$

and the log-likelihood function is

$$\begin{aligned}\ell(\alpha,\lambda) \quad &\propto \quad \sum_{i=1}^{m} X_i \log\left[\frac{1-e^{-\lambda t_i}}{1-(1-\alpha)e^{-\lambda t_i}} - \frac{1-e^{-\lambda t_{i-1}}}{1-(1-\alpha)e^{-\lambda t_{i-1}}}\right]\\ &\quad +\sum_{i=1}^{m} R_i \log\left[\frac{\alpha e^{-\lambda t_i}}{1-(1-\alpha)e^{-\lambda t_i}}\right].\end{aligned}$$

Hence, we have the following log-likelihood equations:

$$\begin{cases}\dfrac{\partial \ell(\alpha,\lambda)}{\partial \alpha} = 0,\\ \dfrac{\partial \ell(\alpha,\lambda)}{\partial \lambda} = 0.\end{cases} \tag{3.1}$$

The MLEs of α and λ cannot be obtained in a closed form by solving equations (3.1) and they must be calculated using a numerical method. Since there is no closed form for MLEs, a mid-point approximation and the EM algorithm are introduced as follows for finding the MLEs of α and λ.

3.1.1. *Mid-point approximation method.* The midpoint estimators based on progressively Type-I interval censoring can be obtained by assuming that X_i failures occurred at the center of the interval, $m_i = \frac{t_{i-1}+t_i}{2}$, and the R_i censored items withdrawn at the censoring time t_i. Then log-likelihood function from the MOGE distribution can be specified as follows:

$$\begin{aligned}\log(L^{\circ}) \quad &\propto \quad \sum_{i=1}^{m}[X_i \log(f(m_i;\alpha,\lambda)) + R_i \log(1-F(t_i;\alpha,\lambda))]\\ &= \quad n\log\alpha + \log\lambda \sum_{i=1}^{m} X_i - \lambda \sum_{i=1}^{m}(X_i m_i + R_i t_i)\\ &\quad -2\sum_{i=1}^{m} X_i \log(1-(1-\alpha)e^{-\lambda m_i}) - \sum_{i=1}^{m} R_i \log(1-(1-\alpha)e^{-\lambda t_i}).\end{aligned}$$

Therefore, the maximum likelihood estimate of α, $\hat{\alpha}$, and the maximum likelihood estimate of λ, $\hat{\lambda}$, are the solution of the sequel equations:

$$\frac{n}{\hat{\alpha}} = 2\sum_{i=1}^{m} X_i \frac{e^{-\hat{\lambda} m_i}}{1-(1-\hat{\alpha})e^{-\hat{\lambda} m_i}} + \sum_{i=1}^{m} R_i \frac{e^{-\hat{\lambda} t_i}}{1-(1-\hat{\alpha})e^{-\hat{\lambda} t_i}}, \tag{3.2}$$

and

$$\begin{aligned}\frac{\sum_{i=1}^{m} X_i}{\hat{\lambda}} \quad &= \quad \sum_{i=1}^{m}(X_i m_i + R_i t_i) + 2(1-\hat{\alpha})\sum_{i=1}^{m} X_i m_i \frac{e^{-\hat{\lambda} m_i}}{1-(1-\hat{\alpha})e^{-\hat{\lambda} m_i}}\\ &\quad +(1-\hat{\alpha})\sum_{i=1}^{m} R_i t_i \frac{e^{-\hat{\lambda} t_i}}{1-(1-\hat{\alpha})e^{-\hat{\lambda} t_i}}.\end{aligned} \tag{3.3}$$

There is no closed form for the solutions of (3.2) and (3.3), thus an iterative numerical method is needed to obtain the parameter estimates, i.e., $\hat{\alpha}$ and $\hat{\lambda}$.

3.1.2. *EM algorithm.* The EM algorithm is applicable to obtain the maximum likelihood estimator of the parameters and useful in a variety of incomplete-data problems where algorithms such as the Newton-Raphson method may sometimes be complicated. On each iteration of the EM algorithm, there are two steps called E-step and the M-step: Let $y_{ij}, j = 1, 2, \ldots, X_i$, be the survival times within subinterval $(t_{i-1}, t_i]$ and $z_{ij}, j = 1, 2, \ldots, R_i$, be the survival times for those withdrawn items at t_i for $i = 1, 2, 3, \ldots, m$, then the log-likelihood, $\log(L^*)$, for the complete lifetimes of n items from the MOGE distribution is given as follows:

$$\begin{aligned}
\log(L^*) &\propto \sum_{i=1}^{m}\left[\sum_{j=1}^{X_i}\log(f(y_{ij},\theta)) + \sum_{j=1}^{R_i}\log(f(z_{ij},\theta))\right] \\
&= n(\log\alpha + \log\lambda) - \lambda\sum_{i=1}^{m}\left[\sum_{j=1}^{X_i} y_{ij} + \sum_{j=1}^{R_i} z_{ij}\right] \\
&\quad -2\sum_{i=1}^{m}\left[\sum_{j=1}^{X_i}\log(1-(1-\alpha)e^{-\lambda y_{ij}}) + \sum_{j=1}^{R_i}\log(1-(1-\alpha)e^{-\lambda z_{ij}})\right].
\end{aligned} \tag{3.4}$$

Taking the derivative with respective to α and λ, respectively, on (3.4), likelihood equations are obtained by

$$\frac{n}{\alpha} = 2\sum_{i=1}^{m}\left[\sum_{j=1}^{X_i}\frac{e^{-\lambda y_{ij}}}{(1-(1-\alpha)e^{-\lambda y_{ij}})} + \sum_{j=1}^{R_i}\frac{e^{-\lambda z_{ij}}}{(1-(1-\alpha)e^{-\lambda z_{ij}})}\right],$$

and

$$\begin{aligned}
\frac{n}{\lambda} &= 2\sum_{i=1}^{m}\left[\sum_{j=1}^{X_i}\frac{(1-\alpha)y_{ij}e^{-\lambda y_{ij}}}{(1-(1-\alpha)e^{-\lambda y_{ij}})} + \sum_{j=1}^{R_i}\frac{(1-\alpha)z_{ij}e^{-\lambda z_{ij}}}{(1-(1-\alpha)e^{-\lambda z_{ij}})}\right] \\
&\quad + \sum_{i=1}^{m}\left[\sum_{j=1}^{X_i} y_{ij} + \sum_{j=1}^{R_i} z_{ij}\right].
\end{aligned}$$

The EM-algorithm has the following steps:
Step 1. Given initial estimates of α and λ , say $\alpha^{(0)}$ and $\lambda^{(0)}$;
Step 2. In the kth iteration, the E-step requires to compute

$$E_{1i} = E_{\hat{\alpha}^{(k)},\ \hat{\lambda}^{(k)}}\left[Y \middle| Y \in [t_{i-1}, t_i)\right],$$

$$E_{2i} = E_{\hat{\alpha}^{(k)},\ \hat{\lambda}^{(k)}}\left[Y \middle| Y \in [t_i, \infty)\right],$$

$$E_{3i} = E_{\hat{\alpha}^{(k)},\ \hat{\lambda}^{(k)}}\left[\frac{e^{-\hat{\lambda}^{(k)}Y}}{1-(1-\hat{\alpha}^{(k)})e^{-\hat{\lambda}^{(k)}Y}} \middle| Y \in [t_{i-1}, t_i)\right],$$

$$E_{4i} = E_{\hat{\alpha}^{(k)},\ \hat{\lambda}^{(k)}}\left[\frac{e^{-\hat{\lambda}^{(k)}Y}}{1-(1-\hat{\alpha}^{(k)})e^{-\hat{\lambda}^{(k)}Y}} \middle| Y \in [t_{i-1}, \infty)\right],$$

$$E_{5i} = E_{\hat{\alpha}^{(k)},\ \hat{\lambda}^{(k)}}\left[\frac{Ye^{-\hat{\lambda}^{(k)}Y}}{1-(1-\hat{\alpha}^{(k)})e^{-\hat{\lambda}^{(k)}Y}} \middle| Y \in [t_{i-1}, t_i)\right],$$

and

$$E_{6i} = E_{\hat{\alpha}^{(k)},\ \hat{\lambda}^{(k)}}\left[\frac{Ye^{-\hat{\lambda}^{(k)}Y}}{1-(1-\hat{\alpha}^{(k)})e^{-\hat{\lambda}^{(k)}Y}}\middle| Y \in [t_{i-1}, \infty)\right],$$

where Y is a random variable which has the MOGE distribution density function (1.1).
Step 3. The M-step maximize the likelihood function. Based on the likelihood equations for complete data, we can obtain the estimates

$$\hat{\alpha}^{(k+1)} = \frac{n}{2\sum_{i=1}^{m}\left[\sum_{j=1}^{X_i} E_{3i} + \sum_{j=1}^{R_i} E_{4i}\right]},$$

and

$$\hat{\lambda}^{(k+1)} = \frac{n}{\sum_{i=1}^{m}\left[\sum_{j=1}^{X_i} E_{1i} + 2(1-\hat{\alpha}^{(k+1)})E_{5i} + \sum_{j=1}^{R_i} E_{2i} + 2(1-\hat{\alpha}^{(k+1)})E_{6i}\right]};$$

Step 4. Setting $k = k+1$, the MLEs of α and λ can be obtained by repeating the E-step and M-step until convergence occurs.

Note that numerical integration methods are required to compute the above conditional expectations in Step 2.

3.2. Method of moments. Let Y be a random variable which has the MOGE distribution density function (1.1). The kth moment of a doubly truncated generalized exponential distribution in the interval (a, b) where $0 < a < b$ is given by

$$E_{\alpha,\ \lambda}\left[Y^k \middle| Y \in [a, b)\right] = \frac{\int_a^b y^k f(y; \alpha, \lambda)dy}{F(b; \alpha, \lambda) - F(a; \alpha, \lambda)}$$

Equating the sample moment to the corresponding population moment up to the second order, the following equations can be used to find the estimates of moment method:

$$E[Y] = \frac{1}{n}\left[\sum_{i=1}^{m} X_i E_{\alpha,\lambda}[Y|\, Y \in [t_{i-1}, t_i)] + R_i E_{\alpha,\lambda}[Y|Y \in [t_{i-1}, \infty)]\right],$$

and

$$\begin{aligned} E[Y^2] &= \frac{1}{n}\left[\sum_{i=1}^{m} X_i E_{\alpha,\lambda}[Y^2\big|\, Y \in [t_{i-1}, t_i)]\right] \\ &\quad + \left[\sum_{i=1}^{m} R_i E_{\alpha,\lambda}[Y^2\big|\, Y \in [t_{i-1}, \infty)]\right]. \end{aligned}$$

An iterative procedure can be employed to solve the above equations for α and λ as follows:
Step 1. Consider the initial values of α and λ, say $\hat{\alpha}^{(0)}$ and $\hat{\lambda}^{(0)}$ with $k = 0$;
Step 2. In the $k+1$th iteration,
• we compute $E_{\hat{\alpha}^{(k)},\ \hat{\lambda}^{(k)}}[Y|\, Y \in [t_{i-1}, t_i)]$ and $E_{\hat{\alpha}^{(k)},\ \hat{\lambda}^{(k)}}\left[Y^2\big|\, Y \in [t_{i-1}, t_i)\right]$ and solve the following equation for α, say $\hat{\alpha}^{(k+1)}$:

$$P(\alpha) = \frac{\left[\sum_{i=1}^{m} X_i E_{\alpha,\lambda}[Y|\, Y \in [t_{i-1}, t_i)] + R_i E_{\alpha,\lambda}[Y|\, Y \in [t_{i-1}, \infty)]\right]^2}{n\sum_{i=1}^{m}\left[X_i E_{\alpha,\lambda}[Y^2|Y \in [t_{i-1}, t_i)] + R_i E_{\alpha,\lambda}[Y^2|\, Y \in [t_{i-1}, \infty)]\right]},$$

where using (1.3) and (1.4),

$$P(\alpha) = E^2[Y]/E[Y^2] = \frac{\alpha \log^2 \alpha}{2(1-\alpha)\text{PolyLog}[2, 1-\alpha])}.$$

• The solution for α, say $\hat{\alpha}^{(k+1)}$, is obtained through the following equation:

$$\frac{\hat{\alpha}^{(k+1)} \log(\hat{\alpha}^{(k+1)})}{(\hat{\alpha}^{(k+1)} - 1)\lambda^{(k+1)}} = \frac{1}{n}\Bigg[\sum_{i=1}^{m} X_i E_{\hat{\alpha}^{(k)},\hat{\lambda}^{(k)}}[Y|\, Y \epsilon\, [t_{i-1}, t_i)] + \sum_{i=1}^{m} R_i E_{\hat{\alpha}^{(k)},\hat{\lambda}^{(k)}}[Y|\, Y \in [t_{i-1}, \infty)]\Bigg];$$

Step 3. Checking convergence, if the convergence occurs then the current $\hat{\alpha}^{(k+1)}$ and $\hat{\lambda}^{(k+1)}$ are the estimates of α and λ by the method of moments; otherwise set $k = k+1$ and go to Step 2.

3.3. Estimation based on probability plot. For progressively Type-I interval censored data, (X_i, R_i, t_i), $i = 1, 2, \ldots, m$, of size n, the distribution function at time t_i can be estimated as

$$\hat{F}(t_i) = 1 - \prod_{j=1}^{i}(1 - \hat{p}_j), \quad i = 1, 2, \ldots, m,$$

where

$$\hat{p}_j = \frac{X_j}{n - \sum\limits_{k=0}^{j-1} X_k - \sum\limits_{k=0}^{j-1} R_k}, \quad j = 1, 2, \ldots, m.$$

From (1.2), we have

$$t = -\frac{1}{\lambda} \log \frac{1 - F(t)}{1 - (1-\alpha)F(t)}.$$

If $\hat{F}(t_i)$ is the estimate of $F(t_i)$, then the estimates of α and λ in the MOGE distribution based on probability plot can be obtained by minimizing $\sum\limits_{i=1}^{m}\left[t_i + \frac{1}{\lambda}\log\frac{1-\hat{F}(t_i)}{1-(1-\alpha)\hat{F}(t_i)}\right]^2$ with respect to α and λ.

3.4. Simulation algorithm. In this section, we give A short algorithm for simulating $X_1, X_2, \ldots, X_m$ from a random sample of size n put on life test at time 0 is therefore given below. Let $X_0 = R_0 = 0$; We use the fact that for $i = 1, ..., m$,

$$X_i|\, X_{i-1}, \ldots, X_0, R_{i-1}, \ldots, R_0 \sim Binom\Big(n - \sum_{j=1}^{i-1}(X_j + R_j), \frac{F(t_i) - F(t_{i-1})}{1 - F(t_{i-1})}\Big),$$

and

$$R_i = \left\lfloor p_i\big(n - \sum_{j=1}^{i-1}(X_i + R_i) - X_i\big)\right\rfloor.$$

Hence we can give an algorithm as follows:

Step 1. Set $i = 0$ and let $xsum = rsum = 0$;
Step 2. Next i;
Step 3. If $i = m+1$, exit the algorithm;
Step 4. Generate X_i as a binomial random variable with parameters $(n - xsum - rsum)$

and $\frac{F(t_i)-F(t_{i-1})}{1-F(t_{i-1})}$;

Step 5. Calculate $R_i^{obs} = \left\lfloor p_i(n - \sum_{j=1}^{i-1}(X_i + R_i) - X_i) \right\rfloor$ or $R_i^{obs} = \min(n - xsum - rsum - X_i, R_i)$, depending upon how the censoring scheme is chosen;

Step 6. Set $xsum = xsum + X_i$, $rsum = rsum + R_i^{obs}$;

Step 7. Go to step 2.

3.5. Simulation schemes. Continuing with our exploration of progressive Type-I interval censoring under the MOGE distribution lifetime models, let us consider a numerical example, and discuss some of the issues which arise. We use the values $t_1 = 5.5$, $t_2 = 10.5$, $t_3 = 15.5$, $t_4 = 20.5$, $t_5 = 25.5$, $t_6 = 30.5$, $t_7 = 40.5$, $t_8 = 50.5$ and $t_9 = 60.5$. The lifetime distribution is the MOGE Type with parameters $(\alpha, \lambda) = (0.5, .06)$, where are the simulation input parameters. To compare the performances of the estimation procedures developed in this paper, we consider the following four progressive interval censoring schemes which are similar to the patterns of simulation schemes used in Aggarwala (2001) and also used in Ng and Wang (2009) and Chen and Lio (2010):

$$\begin{aligned}
\boldsymbol{p}_{(1)} &= (.25, .25, .25, .25, .5, .5, .5, .5, 1),\\
\boldsymbol{p}_{(2)} &= (.5, .5, .5, .5, .25, .25, .25, .25, 1),\\
\boldsymbol{p}_{(3)} &= (0, 0, 0, 0, 0, 0, 0, 0, 1),\\
\boldsymbol{p}_{(4)} &= (.25, 0, 0, 0, 0, 0, 0, 0, 1),
\end{aligned}$$

where censoring in $\boldsymbol{p}_{(1)}$ is lighter for the first four intervals and heavier for the next four intervals. The censoring pattern is reversed in $\boldsymbol{p}_{(2)}$. $\boldsymbol{p}_{(3)}$ is the conventional interval censoring where no removals prior to the experiment termination and the censoring in $\boldsymbol{p}_{(4)}$ only occurs at the left-most and the right-most. The initial values of α and λ for iterative progresses of MLE, mid-point approximation, EM algorithm, moment method and probability plot are given the same values, which for each simulation run, is randomly generated.

3.6. Simulation results. The result for the 1000 simulation runs by R software is shown in Table 1 and Table 2 and is graphically illustrated in Figures 4 and 5. As the performances among the four censoring schemes, the third scheme $\boldsymbol{p}_{(3)}$ provides the most precise results as seen from "Bias", "SD" (i.e. the standard deviation) and "MSE" (i.e. the mean squared errors) shown in Table 1 and Table 2 from the dispersions of the boxplots shown in the Figures 1 and 2, then followed by the schemes $\boldsymbol{p}_{(4)}$, $\boldsymbol{p}_{(1)}$ and $\boldsymbol{p}_{(2)}$.

4. Real data analysis

A data set which consists of 118 patients with breast cancer treated at the Sadouqi Hospital of Yazd is used for modelling the MOGE distribution; This data set is explored from [4] and summarized in Table 3. In this table, the first column shows 7 pre-assigned time intervals in years which were determined before the experiment, i.e., $[t_{i-1}, t_i)$, $i = 1, ..., 7$. The second column shows the number of patients who are died in the time intervals, i.e., $X_1, ..., X_7$ and finally, the last column is the number of patients who were dropped out from the study at the right end of each time interval; These dropped patients are known to be survived at the right end of each time interval but no follow up. Hence, the last column in Table 3 provides the values of R_i, $i = 1, \ldots, m = 7$.

Scheme		EM	Midpoint	MLE	MME	probpt
1	Median	0.4813	0.9553	0.5468	0.3978	0.4272
2	Median	0.4307	1.0000	0.5624	0.3858	0.4614
3	Median	0.5036	0.7766	0.5640	0.4640	0.5983
4	Median	0.4985	0.5783	0.3517	0.4316	0.5639
1	Mean	0.5080	0.8411	0.5740	0.3982	0.5193
2	Mean	0.4809	0.9020	0.5883	0.3860	0.5328
3	Mean	0.5237	0.7696	0.6223	0.4706	0.6100
4	Mean	0.5265	0.6433	0.3708	0.4318	0.5912
1	Bias	0.0080	0.3411	0.0740	-0.1018	0.0193
2	Bias	-0.0191	0.4020	0.0883	-0.1140	0.0328
3	Bias	0.0237	0.2696	0.1223	-0.0294	0.1100
4	Bias	0.0265	0.1433	-0.1292	-0.0682	0.0912
1	SD	0.2136	0.2024	0.2777	0.0071	0.3376
2	SD	0.2124	0.1829	0.3110	0.0034	0.3382
3	SD	0.1775	0.2217	0.2683	0.0403	0.2746
4	SD	0.1858	0.2966	0.0979	0.0166	0.2829
1	MSE	0.0457	0.1573	0.0826	0.0104	0.1144
2	MSE	0.0455	0.1950	0.1045	0.0130	0.1155
3	MSE	0.0321	0.1219	0.0869	0.0025	0.0875
4	MSE	0.0352	0.1085	0.0263	0.0049	0.0884

Table 1. Estimates of α from 1000 simulations for the five estimation methods and four simulation schemes.

4.1. Model selection. To select a suitable model for the given data set in Table 3, we start with the MOGE distribution. We will fit the MOGE distribution and statistically test whether the MOGE distribution model can be reduced to the exponential (E) distribution model for the given data set in the Table 3.

Fitting the MOGE to the given data, MLE of (α, λ) is

$$(\hat{\alpha}, \hat{\lambda}) = (0.05785, 0.52959),$$

$-2\log L(MOGE) = 137.4273$ and $AIC(MOGE) = 141.4273$. Then we fit the exponential distribution model to the given data set, MLE of λ is $\hat{\lambda} = 2.99422$, $-2\log L(E) = 152.0508$ and $AIC(E) = 154.0508$. Note that $AIC(MOGE) < AIC(E)$ and also, the log-likelihood ratio statistic is

$$-2\log(\Lambda) = (-2\log L(E)) - (-2\log L(MOGE)) = 14.6235,$$

which is greater than $\chi^2_{0.05}(1) = 3.8415$, hence the MOGE distribution provides a better fit for the data at size 0.05; indeed, the p-value of the test is 0.00013!

Additional model fitting to the GE distribution yields the estimated parameters $(\hat{\alpha}, \hat{\lambda}) = (0.19251, 1.03246)$ and $-2\log L(GE) = 138.1842$, so $AIC(GE) = 142.1842$.

4.2. Conclusion. In this paper, three methods to estimate the parameters of the MOGE distribution under progressive Type-I interval censoring have been developed; These methods were maximum likelihood estimation, estimation of method moments and the estimation based on the probability plot. The simulation study in the case of moderate

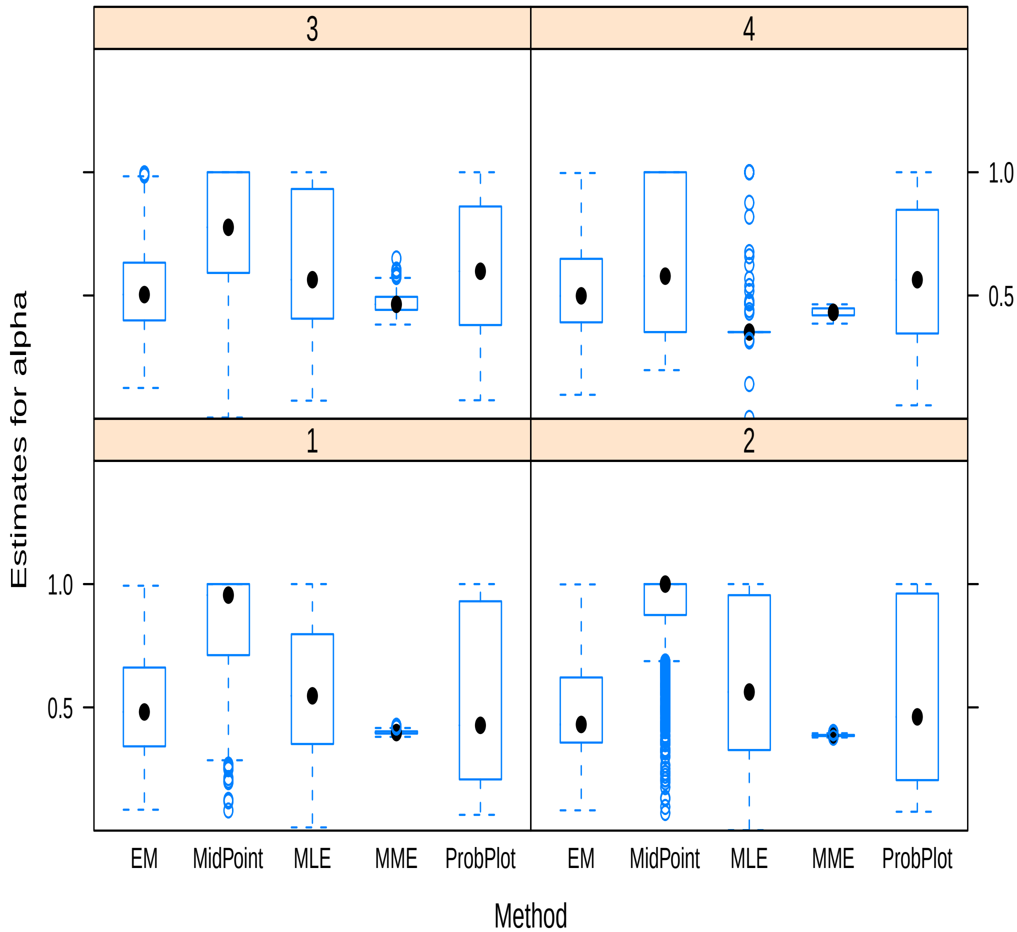


Figure 4. Boxplots for α from 1000 simulations for the five estimation methods and four simulation schemes

large size data set indicated that all these estimators give relatively accurate parameter estimation and the maximum likelihood estimator gives the most precise estimation as summarized in the Table 1 and 2 and Figures 4 and 5. We therefore recommend the "MLE" to be used to estimate the parameters in the MOGE distribution under progressive Type-I interval censoring. In the end of the paper, a real data set based on patients with breast cancer in order to demonstrate the applicabilities was used. Table 4 showed high flexibility of the MOGE distribution to model the data.

Acknowledgment

Scheme		EM	Midpoint	MLE	MME	probpt
1	Median	0.06267	0.08473	0.06333	0.02206	0.05194
2	Median	0.05943	0.09222	0.06375	0.01634	0.05005
3	Median	0.06125	0.06425	0.05359	0.02987	0.06601
4	Median	0.06142	0.04903	0.04344	0.0267	0.06359
1	Mean	0.06625	0.0821	0.06254	0.02236	0.05634
2	Mean	0.07151	0.089	0.08204	0.01659	0.05631
3	Mean	0.06317	0.06403	0.05481	0.03016	0.06446
4	Mean	0.06438	0.05095	0.04309	0.02691	0.06391
1	Bias	0.00625	0.0221	0.00254	-0.03764	-0.00366
2	Bias	0.01151	0.029	0.02204	-0.04341	-0.00369
3	Bias	0.00317	0.00403	-0.00519	-0.02984	0.00446
4	Bias	0.00438	-0.00905	-0.01691	-0.03309	0.00391
1	SD	0.02626	0.01911	0.02387	0.00406	0.0296
2	SD	0.03347	0.03598	0.01973	0.27499	0.0043
3	SD	0.01657	0.01194	0.0136	0.00345	0.01817
4	SD	0.02836	0.01926	0.01341	0.014	0.00374
1	MSE	0.00089	0.00073	0.00085	0.00058	0.00143
2	MSE	0.00113	0.00143	0.00123	0.0761	0.0019
3	MSE	0.00035	0.00028	0.00016	0.00021	9e-04
4	MSE	0.00082	0.00039	0.00026	0.00048	0.00111

Table 2. Estimates of λ from 1000 simulations for the five estimation methods and four simulation schemes.

Intervals	Number of failures	Number of Withdraws
$[0, 0.5)$	99	4
$[0.5, 1.0)$	8	2
$[1.0, 1.5)$	3	0
$[1.5, 2.0)$	1	0
$[2.0, 2.5)$	0	0
$[2.5, 3.0)$	0	0
$[3.0, 3.5)$	0	1

Table 3. Breast cancer survival times.

Distribution	$\hat{\alpha}$	$\hat{\lambda}$	$-2\log\hat{L}$	AIC	BIC	CAIC
E	-	2.99422	152.0508	154.0508	156.8215	154.0853
GE	0.19251	1.03246	138.1842	142.1842	147.7256	142.2885
MOGE	0.05785	0.52959	137.4273	141.4273	146.9687	141.5316

Table 4. Comparison of the E, GE and MOGE distributions

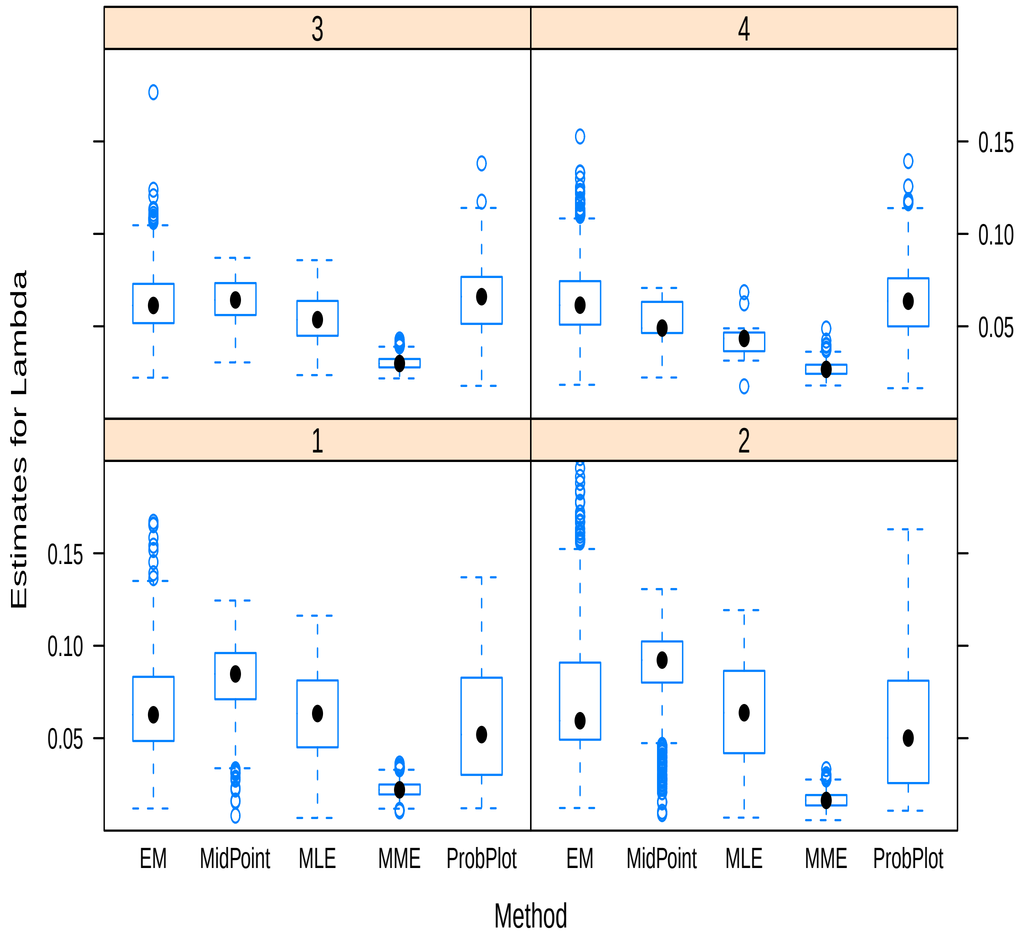


Figure 5. Boxplot for λ from 1000 simulations for the five estimation methods and four simulation schemes

The authors are highly grateful to referees for their valuable comments and suggestions for improving the paper.

References

[1] Aggarwala, R. *Progressively interval censoring Some mathematical results with application to inference*, Communications in Statistics-Theory and Methods, **30**, 1921-1935, 2001.

[2] Alice, T. and Jose, K.K. *On Marshall- Olkin Generalized Exponential Distribution and its Applications.* The Second Annual Conference of the Society of Statistics

and Computer Applications held at St. Thomas College, Pala, Kerala during 28th November to 1st December 1999.

[3] Chen, D.G. and Lio, Y.L. *Parameter estimations for generalized exponential distribution under progressive type-I interval censoring*, Computational Statistics and Data Analysis, **54**, 1581-1591, 2010.

[4] Fallahzadeh, H. *Five Year Survival of Women with Breast Cancer in Sadouqi Hospital of Yazd*. Technical Report, 2014.

[5] Gupta, R.D. and Kundu, D. *Exponentiated exponential distribution: An alternative to gamma and Weibull distributions*, Biometrical Journal, **43**, 117-130, 2001.

[6] Lawless, J. *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, New York, 1982

[7] Marshall, A.W. and Olkin, I. *A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families*, Biometrika, **84**, 641-652, 1997.

[8] Ng, H. and Wang, Z. *Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample*, Journal of Statistical Computation and Simulation, **79**, 145-159, 2009.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1621 – 1640

A Bayesian approach to Cox-Gompertz model

Nihal Ata Tutkun*† and Haydar Demirhan‡

Abstract

Survival analysis has a wide application area from medicine to marketing and Cox model takes an important part in survival analysis. When the distribution of survival data is known or it is appropriate to assume a survival distribution, use of a parametric form of Cox model is employed. In this article, we take into account Cox-Gompertz model from the Bayesian perspective. Considering the difficulties in parameter estimation in classical setting, we propose a simple Bayesian approach for Cox-Gompertz model. We derive full conditional posterior distributions of all parameters in Cox-Gompertz model to run Gibbs sampling. Over an extensive simulation study, estimation accuracies of the classical Cox model and classical and Bayesian settings of Cox-Gompertz model are compared with each other by generating exponential, Weibull, and Gompertz distributed survival data sets. Consequently, if survival data follows Gompertz distribution, most accurate parameter estimates are obtained by the Bayesian setting of Cox-Gompertz model. We also provide a real data analysis to illustrate our approach. In the data analysis, we observe the importance of use of the most accurate model over the survival probabilities of censored observations.

Keywords: Gompertz, Cox model, Gibbs sampling, Bayesian analysis, full conditional, Newton-Raphson, parametric model.

2000 AMS Classification: 62F15, 62N01, 62N02.

Received : 01.12.2014 *Accepted* : 29.07.2015 *Doi* : 10.15672/HJMS.20158012506

*Hacettepe University Department of Statistics Beytepe Ankara Turkey,
Email: nihalata@hacettepe.edu.tr

†Corresponding Author.

‡Hacettepe University Department of Statistics Beytepe Ankara Turkey,
Email: haydarde@hacettepe.edu.tr

1. Introduction

Survival analysis is a class of statistical methods for studying occurrence and timing of events. An event can be defined as development of a disease, response to a treatment, relapse, or death. Therefore, the time from start of a treatment to response, length of remission, and time to death may be taken as a survival time. The most common approach to model covariate effects on survival times is the Cox's semi-parametric regression model, which takes into account the effect of censored observations [5]. In the Cox model, no particular form of probability distribution is assumed for survival times. However, if it is known, parametric models, such as exponential, Weibull, or Gompertz can be applied.

The Cox model is sensitive to the violations of proportional hazards assumption. The form of baseline hazard rate influences the properties of estimators [2]. Because there is no need to assume a particular form of probability distribution for the survival times, the Cox model is more advantageous than the parametric counterparts if baseline hazard is incompatible with a particular distribution. Hazard function is not restricted to a specific functional form; hence, the model has flexibility and widespread applicability. On the other hand, if the assumption of a particular probability distribution is appropriate for data, inferences based on such an assumption will be more precise. In particular, parameter estimates and estimates of quantities such as relative hazards and median survival times will tend to have smaller standard errors than those obtained without a distributional assumption [4]. Based on asymptotic results, Efron [7] and Oakes [29] showed that parametric models lead to more efficient parameter estimates than the Cox model under certain circumstances [28].

Making special assumptions on the distribution of survival times, such as exponential, Weibull, or Gompertz, leads to parametric regression models. Exponential distribution is widely used in survival studies. It plays a role in lifetime studies analogous to normal distribution in other areas of statistics. It is often referred as purely random failure pattern [26]. Although exponential distribution is characterized by a constant hazard function, its constant hazard rate appears to be restrictive in both health and industrial applications [22]. Weibull distribution is a generalization of exponential distribution. It has a hazard function that is monotone increasing, decreasing, or constant. Therefore, it has broader applications. Although use of exponential or Weibull model may be sufficient for a realistic description of various survival time data, other distributions such as Gompertz are required for more precise results. Gompertz distribution is used to describe mortality curves and later modified by Makeham [27] by addition of a constant hazard function. Only exponential, Weibull, and Gompertz models have the assumption of proportional hazards with the Cox model [2]. Because of the functional form of its hazard rate, Gompertz model is more flexible than Weibull model. Also, it allows to asses the influence of independent variables on both parameters of the distribution [3].

Cox-Gompertz model has a wide application area from automobile industry to medicine. Gompertz distribution is commonly used in actuary, reliability, and life testing as a survival time distribution [1]. Firstly, it is used to fit mortality tables by Gompertz [15]. Spickett and Ark [30] fitted the Gompertz distribution to dose-response data of larval tick populations. Grunkemeier et al. [16] used the Gompertz model for the survival times after a surgery for acquired hearth disease. Classical analysis of Gompertz model for cure rate models was given by Gieser, et al. [13]. Willekens [31] provided connections between the Gompertz, the Weibull and other Type-I extreme value distributions. Fabrizio [8] used Gompertz model for cabinet duration times. Klepper [23] used Gompertz distribution to estimate hazard rate models for the length of time for a particular firm stays in the market. Cantner et al. [3] used the approach of Klepper [23] for German automobile industry. Jeong and Fine [19] and Jeong [18] used Gompertz distribution to parameterize cumulative incidence function, which is used to estimate the cumulative probability of locoregional recurrences in the presence of other competing events. Gokovali et al. [14] use Gompertz distribution to analyze the determinants of tourists' length of stay at a destination. Launder and Bender [24] developed adjusted risk difference and number needed to treat measures for use in observational studies with survival time outcomes within the framework of the Cox model taking the distribution of confounders into account. The performance of these estimators is assessed by performing Monte Carlo simulations and is also illustrated by means of data of the Dusseldorf Obesity Mortality Study. Ghitanya et al. [12] studied the maximum-likelihood estimates of the parameters by considering a progressively Type-II censored sample from the Gompertz distribution.

Estimation of parameters of Cox-Gompertz model requires use of numerical techniques such as Newton-Raphson (NR). Because NR method requires only first and second partial derivatives of likelihood function, it is very flexible. However, it is highly sensitive to the initial values, it may require a large number of iterations to converge, and it may converge to a local maximum or may not converge in some cases. NR method gives no insight into the distribution of parameters. Moreover, numerical methods such as NR are asymptotic; hence, standard deviations of parameters are obtained only approximately. These are important disadvantages of the classical setting. Another general disadvantage of the classical setting is that ML estimators need not be finite, so it can occur outside the parameter space. Considering these weaknesses, we propose use of a Bayesian approach for estimation of Cox-Gompertz model.

In survival analysis, Bayesian approaches provide a flexible tool via the Gibbs sampling when the full conditional distributions are found in a closed form. Dellaportas and Smith [6] give a Bayesian approach for proportional hazards model with baseline hazard function of exponential and Weibull distributions. Bayesian approaches to the parametric survival models have some advantages over the classical setting. In the Bayesian setting, inference is exact rather than asymptotic. It provides an entire posterior distribution for each element of the model. However, the classical setting yields a point estimate and a precision estimated via an asymptotic method. In addition, the Bayesian approach would give better estimates of variability than the likelihood analysis [9].

Bayesian approaches to some parametric forms of the Cox model are given by Kim and Ibrahim [21]. They consider Cox-Weibull and extreme value regression models, and suggest use of a uniform prior instead of the Jeffrey's. They also derive sufficient conditions for the existence of posterior moment generating functions and those of the posterior distributions to be proper in the case of Cox-Weibull and extreme value regression models. Kim and Ibrahim [21] give Bayesian estimation procedure for an extreme value type I distribution. In their approach, data is a log-completely observed time or log-censoring time. In this study, however, we consider the Gompertz distribution as the distribution of a completely observed or censoring time without any transformations such as log. Then, we propose a Bayesian approach to Cox-Gompertz model. Although the distributional forms of extreme value and Gompertz distributions are similar, their domains are not the same (see for the distributional forms Bender et al. [2] and Kim and Ibrahim [21]. In fact, there are several distributional forms of Gompertz distribution [20, p.25-26, 81-85]. The one used here can be interpreted as a truncated extreme value type-I distribution. Therefore, we give a Bayesian approach for a different parametric model than the one given by Kim and Ibrahim [21]. In addition, we derive full conditional posterior distributions of the model parameters. Because of not using an approximate method to generate random numbers from the full conditionals, our derivations make the application of Bayesian setting more flexible.

In Section 2 the Cox-Gompertz model is illustrated. In Section 3, Bayesian inference for the Cox-Gompertz model is demonstrated and full conditional distributions are given to derive posterior inferences by the Gibbs sampling. In Section 4, a real data analysis is presented to illustrate our approach. We observe that use of classical Cox model can produce notably different estimates of survival probabilities for censored observations. In Section 5, the simulation study on the comparison of estimation accuracies of Cox, Cox-Gompertz models in the classical setting, and Cox-Gompertz model in the Bayesian setting is given. In Section 6, a short discussion is given.

2. The Cox-Gompertz model

A data set, based on a random sample of size n, consists of $(t_j, \delta_j, \boldsymbol{x}_j)$ for $j = 1, \ldots, n$, where t_j is the time on study for the jth individual, δ_j is the event indicator taking 1 if the event has occurred and 0 otherwise, and $\boldsymbol{x}_j$ is the vector of covariates or risk factors for the jth individual. Hazard function for the Cox model is given as follows:

$$h(t; x) = h_0(t) \exp\{\boldsymbol{X\beta}\}, \tag{1}$$

where $\boldsymbol{X}$ is the design matrix including categorical variables or continuous measurements of each individual, $h_0(t)$ is the baseline hazard function obtained for an individual with $x_{ji} = 0$, and $\boldsymbol{\beta}_{[p\times 1]}$ is a vector of unknown parameters. In the absence of tied observations, complete censored-data likelihood is given as follows:

$$\begin{aligned} L(\boldsymbol{\beta}, h_0(t)) &= \prod_{j=1}^{n} h_0(t_j|\boldsymbol{x}_j)^{\delta_j} S(t_j|\boldsymbol{x}_j) \\ &= \prod_{j=1}^{n} h_0(t_j)^{\delta_j} [\exp\{\boldsymbol{\beta}^{'}\boldsymbol{x}_j\}]^{\delta_j} \exp\big\{-H_0(t_j)\exp\{\boldsymbol{\beta}^{'}\boldsymbol{x}_j\}\big\}, \end{aligned} \tag{2}$$

where $H_0(t)$ is cumulative baseline hazard function and $S(t_j|\boldsymbol{x}_j)$ is survival function [22].

Under Gompertz distribution, the baseline hazard function is defined as follows:

$$h_0(t) = \lambda \exp\{\alpha t\}, \tag{3}$$

where $0 < t \leq \infty$, $\lambda > 0$ is a scale and $-\infty < \alpha < \infty$ is a shape parameter. Cumulative baseline hazard function is as following:

$$H_0(t) = (\lambda/\alpha)[\exp\{\alpha t\} - 1]. \tag{4}$$

Using (3), (4) and the general likelihood function given in (2), likelihood function of the Gompertz model is obtained as following:

$$\begin{aligned} L(h_0(t), \boldsymbol{\beta}|\boldsymbol{t}) \propto \prod_{j=1}^{n} \lambda^{\delta_j} \exp\big\{\delta_j(\alpha t_j + \boldsymbol{\beta}^{'}\boldsymbol{x}_j)\big\} \exp\big\{(\lambda/\alpha)[1 - \exp(\alpha t_j)] \\ \times \exp(\boldsymbol{\beta}^{'}\boldsymbol{x}_j)\big\}. \end{aligned} \tag{5}$$

NR method is a frequently used method to obtain the ML estimates over (5).

3. Bayesian setting for the Cox-Gompertz model

The likelihood function given in (5) is used to obtain a posterior distribution. We consider use of an improper prior distribution to conduct a noninformative Bayesian analysis. Joint prior distribution of $h_0(t)$ and $\boldsymbol{\beta}$ is taken as $p(h_0(t), \boldsymbol{\beta}) \propto constant$. Then the joint posterior distribution of $h_0(t)$ and $\boldsymbol{\beta}$ given the data is found from (5) as follows:

$$p(h_0(t), \boldsymbol{\beta}|\boldsymbol{t}) \propto \lambda^{\sum_{j=1}^{n}\delta_j} \exp\left\{\sum_{j=1}^{n} \delta_j(\alpha t_j + \boldsymbol{\beta}^{'}\boldsymbol{x}_j) + (\lambda/\alpha)\sum_{j=1}^{n}[1 - \exp(\alpha t_j)]\exp(\boldsymbol{\beta}^{'}\boldsymbol{x}_j)\right\}. \tag{6}$$

where $p(h_0(t), \boldsymbol{\beta}|\boldsymbol{t}) = p(\alpha, \lambda, \boldsymbol{\beta}|\boldsymbol{t})$.

Gibbs sampling is employed to draw posterior inferences from the posterior given in (6). Full conditional posterior distributions of α, λ, β_i are required to run the Gibbs sampling. The following full conditionals are obtained:

$$\alpha|\lambda, \boldsymbol{\beta} \sim N\left[\frac{3\sum_{j=1}^{n}(\delta_j t_j - \lambda \frac{t_j^2}{2} e^{\boldsymbol{\beta}'\boldsymbol{x}_j})}{\lambda \sum_{j=1}^{n} t_j^3 e^{\boldsymbol{\beta}'\boldsymbol{x}_j}}, \frac{3}{\lambda \sum_{j=1}^{n} t_j^3 e^{\boldsymbol{\beta}'\boldsymbol{x}_j}}\right], \tag{7}$$

$$\lambda|\alpha, \boldsymbol{\beta} \sim Gamma\left[\sum_{j=1}^{n} \delta_j + 1, \frac{\alpha}{\sum_{j=1}^{n}[\exp(\alpha t_j) - 1]\exp(\boldsymbol{\beta}'\boldsymbol{x}_j)}\right], \tag{8}$$

$$\beta_i|\alpha, \lambda, \boldsymbol{\beta}_{-i} \sim N\left[\frac{s_2}{s_1}, \frac{1}{s_1}\right], \tag{9}$$

where $\boldsymbol{\beta}_{-i}$ contains the regression parameters but β_i, $s_1 = (\lambda/\alpha)\sum_{j=1}^{n} c_j x_{ji}^2 \, [\exp(\alpha t_j) - 1]$, and $s_2 = \sum_{j=1}^{n} \delta_j x_{ji} - (\lambda/\alpha)\sum_{j=1}^{n} c_j x_{ji}[\exp(\alpha t_j) - 1]$. Derivation of all of these full conditionals are given in the Appendices A1-A3. Implementation of Gibbs sampling using these full conditional distributions is straightforward. Number of iterations is determined such that achievement of convergence is ensured. Convergence check can be made by using the potential scale reduction factor, $\widehat{R}$, given by Gelman [11]. If value of $\widehat{R}$ is close to 1 and less than 1.2 then it is concluded that the convergence is achieved for the relevant parameter [11].

The sufficient conditions for the existence of posterior moment generating function of the model parameters and the propriety of the posterior distribution are mentioned by Kim and Ibrahim [21] for the Weibull and extreme value distribution cases. Kim and Ibrahim [21] assume that one of the parameters of hazard function, corresponding to Weibull distribution, is known; and hence, one of the parameters of hazard function in the extreme value distribution case is also assumed to be known. In addition, they note that if these do not assumed, joint posterior distributions are always improper. On the contrary, all of the parameters of the hazard function of the Gompertz distribution that we are working on are random. Thus, the propriety of our joint posterior distribution is uncertain when looked from the perspective of Kim and Ibrahim [21]. Gelfand and Shau [10] state that if a Gibbs sampler is used on the improper joint posterior, it is possible to use obtained iterates to draw inferences on the lower-dimensional proper posteriors. As a result, if full conditionals are proper, foregoing transition density remains valid. When the full conditionals given in (7)- (9) are investigated, it is seen that they are proper if α and λ are both finite. Therefore, we do not need to ascertain propriety of our joint posterior distribution in another way. Instead, we utilize directly the result given by Gelfand and Shau [10] due to the propriety of the full conditionals.

4. A real data example

A popular data set is taken into account to illustrate and discuss our findings. The data is on lung cancer and given by Lawless [25]. The data set is also used by Gelfand and Mallic [9] and Kim and Ibrahim [21]. Gelfand and Mallic [9] used the data set to illustrate their work on Cox model, for which the baseline hazard, the covariate link, and the covariate coefficients are all unknown. Thus, they investigated four models from the Bayesian perspective. Kim and Ibrahim [21] gave the ML and Bayesian estimates using a uniform prior under the Cox-Weibull model by including an intercept term and assuming one of the parameters of the hazard function is known.

The data set consists three covariates that performance status at diagnosis (measure between 0 and 100), age of patients in years, and months from diagnosis to entry into the study. Three of the 40 observations are censored. There are 3 tied observation pairs. One of them includes one censored and one uncensored observations. The censored one and one of the other two tied pairs were discarded from the data set. These tied observations were not noticed by Gelfand and Mallic [9] and Kim and Ibrahim [21]. In addition they do not mention anything about the tied observations. We fit Cox-Gompertz model under the Bayesian setting. In Gibbs sampling, total number of iterations was taken as 2500, and 10 parallel chains were generated. To filtrate the effect of starting values, burn-in period was taken as the first 500 iterations of each chain. Every 25 iterations were recorded to reduce the autocorrelation in each of the chains. Parameter estimates with their estimated standard deviations for the Cox model in classical setting and the Cox-Gompertz model in both of the classical and Bayesian settings, and potential scale reduction factor, corresponding to each parameter are given by Table 1.

Table 1. Classical and Bayesian parameter estimates (estimated standard deviations) over Cox and Cox-Gompertz models, and values of potential scale reduction factor ($\widehat{R}$) values.

	Classical Estimates		Bayesian Estimates	
	Cox Model	Cox-Gompertz Model	Cox-Gompertz Model	$\widehat{R}$
α		0.0003 $(7.42 \cdot 10^{-7})$	-0.0019 $(2.65 \cdot 10^{-7})$	1.003
λ		0.0196$(2.70 \cdot 10^{-7})$	0.0331 $(5.41 \cdot 10^{-3})$	1.001
β_1	-0.0130$(1.13 \cdot 10^{-4})$	-0.0504 $(9.12 \cdot 10^{-5})$	-0.0121 $(2.86 \cdot 10^{-5})$	1.001
β_2	0.0135$(3.16 \cdot 10^{-4})$	0.0351 $(8.79 \cdot 10^{-5})$	-0.0076 $(2.97 \cdot 10^{-5})$	1.001
β_3	-0.0149$(1.42 \cdot 10^{-4})$	0.0219 $(2.65 \cdot 10^{-5})$	-0.0015 $(4.58 \cdot 10^{-7})$	1.004

$\widehat{R}$ values indicate that the convergence is achieved for all of the parameters. Estimated standard deviations given in Table 1 are obtained by using inverse of the Hessian matrix and the generated Gibbs sequence in the classical and Bayesian settings, respectively. It is seen from the Table 1 that estimated standard deviations of the parameters of Cox-Gompertz model are smaller than that of the Cox model in both of the classical and Bayesian settings. ML and the Bayesian estimates are not far from each other. The Bayesian estimates of the covariate coefficients, which are more precise, are closer to that of the classical Cox model.

To investigate which model is more successful in explaining the censoring, we estimate $P(t_{12} > 231|\boldsymbol{x}_{12})$, $P(t_{15} > 103|\boldsymbol{x}_{15})$ and $P(t_{23} > 25|\boldsymbol{x}_{23})$ over the considered models, where $\boldsymbol{x}_{12}$, $\boldsymbol{x}_{15}$ and $\boldsymbol{x}_{23}$ are the observed values of covariates corresponding to the relevant censored observations. The same approach of Gelfand and Mallic [9] is used to calculate the probabilities in the Bayesian case. ML estimates of the Cox-Weibull model given by Kim and Ibrahim [21] are used. The results and product of these probabilities, referred as overall, are given in Table 2.

Benefit of the parametric approach for this data set is clearly seen in the Table 2 that Cox-Gompertz model is better than the classical Cox model in the estimation of censored survival times. Cox-Weibull model is also unsuccessful. This is an example of the case that the baseline hazard is not compatible with the parametric distribution. The Cox-Gompertz model seems to be more successful in the estimation of survival probabilities in both of settings. When the classical and Bayesian settings of Cox-Gompertz model are compared, the probabilities obtained over the classical estimates for the survival times of 25 and 103 are greater than their Bayesian counterparts. However, the case is just the

Table 2. Survival probabilities for censored observations.

	Classical Estimates			Bayesian Estimates
	Cox	Cox-Gompertz	Cox-Weibull	Cox-Gompertz
$P(t_{12} > 231 \mid \boldsymbol{x}_{12})$	< 0.0001	0.0619	< 0.0001	0.1122
$P(t_{15} > 103 \mid \boldsymbol{x}_{15})$	< 0.0001	0.7083	< 0.0001	0.3795
$P(t_{23} > 25 \mid \boldsymbol{x}_{23})$	< 0.0001	0.9095	< 0.0001	0.7883
Overall	< 0.0001	0.0387	< 0.0001	0.0336

reverse for the survival time of 231. Thus, the Bayesian estimates are more successful for longer survival times for the data set of interest. As for the overall performance, the Bayesian and classical estimates of Cox-Gompertz model are similar in estimating the censored survival times.

Plots of posterior marginal distributions of the parameters are given by Figure 1. Most of the probability mass of all marginal posterior densities of the parameters are less or greater than zero. And all of them are nearly symmetric. We can conclude that all of the parameters have statistically significant effects on the survival times.

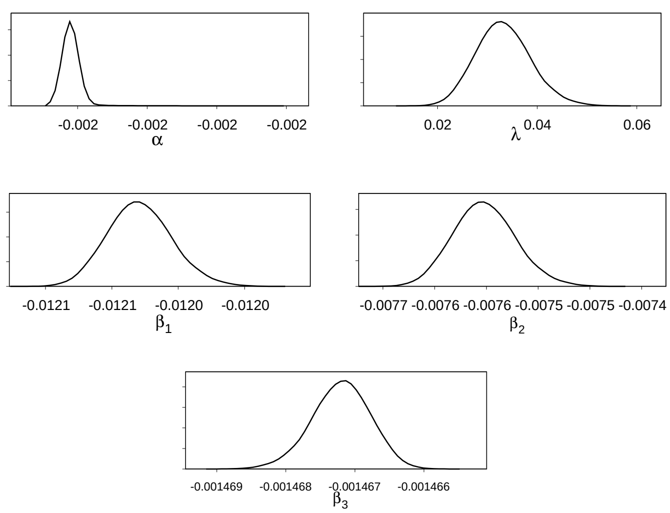


Figure 1. Marginal posterior densities of α, λ and the elements of $\boldsymbol{\beta}$.

5. Simulation study

A simulation study is conducted to investigate the features of our approach and to compare them with classical Cox and Cox-Gompertz models. Two covariates were taken into account. Values of the X_1 is generated from $N(3, 0.1)$ and values of the X_2 is generated from $N(4, 0.5)$. The survival data were generated by using formulas of (10), given by Bender, et al. [2], from the Exponential(λ), Gompertz(α, λ), and Weibull(ν, λ) distributions, respectively.

$$T_j^{\mathrm{E}} = -\frac{\log(U)}{\lambda \exp\{\beta_1 x_{j1} + \beta_2 x_{j2}\}}, T_j^{\mathrm{G}} = (1/\alpha)\log\left[1 - \frac{\alpha \log(U)}{\lambda \exp\{\beta_1 x_{j1} + \beta_2 x_{j2}\}}\right]$$
$$T_j^{\mathrm{W}} = \left[-\frac{\log(U)}{\lambda \exp\{\beta_1 x_{j1} + \beta_2 x_{j2}\}}\right]^{1/\nu}, \tag{10}$$

where U ~Uniform(0,1) and β_i's, $i = 1, 2$, are regression coefficients.

To use a moderate sample size, it is taken as 20. True values of parameters for each survival distribution are given in the third columns of Tables 3-11. Censoring rate is taken as 0 and 0.1, which correspond to cases of no censoring and a moderate rate of censoring, respectively. 1000 independent samples were generated for each of the combinations. Parameter estimates, given by the Tables 3-11 were calculated by averaging the estimates over the generated 1000 samples. Absolute and relative bias, standard deviation and mean square error (MSE) values are reported in Tables 3-11.

It is seen from the Table 3, 4, and 5 that when the survival data are distributed as exponential, parameter estimates and their estimated standard deviations are not affected by the increased censoring for all of three settings. Classical parameter estimates of Cox and Cox-Gompertz models are very different from each other, and estimated standard deviations and MSEs of the parameter estimates of Cox model are smaller than that of Cox-Gompertz model. Absolute and relative biases of parameter estimates of Cox model are smaller than that of Cox-Gompertz model. Thus, Cox model generates better estimates than Cox-Gompertz model in case of exponentially distributed survival data with the classical setting. As for the Bayesian setting, it is interesting that the parameter estimates are similar in all of the cases, in addition their standard deviations and MSEs are close to zero. Absolute biases in the Bayesian setting are somewhat greater than that of Cox model in the classical setting, whereas MSEs are smaller in the Bayesian setting. The cause of this situation is smaller estimated standard deviations of the Bayesian setting. When the classical and Bayesian settings of Cox-Gompertz models are compared, it is seen that absolute and relative biases and the MSEs of the Bayesian setting are smaller than that of the classical setting. As the result, it can be stated the Bayesian approach is neither better nor worse than the classical Cox approach and better than the classical settings of Cox-Gompertz model when the data come from *exponential* distribution. The side effects of the disagreement between the survival distribution and baseline hazard is clearly seen here for the classical settings and obtained smaller variances are neutralized the side effects of the disagreement in the Bayesian setting.

It can be concluded from the Table 6, 7, and 8 that in contrast to the preceding inferences, absolute and relative biases of the parameter estimates obtained over Cox model is greater than that of Cox-Gompertz model for the cases 3 and 4 when the survival data comes from Weibull distribution. In addition, estimated standard deviations and MSEs of the model parameters obtained by Cox model are greater than that of obtained by Cox-Gompertz model for the cases 3 and 4. These situations are just reverse for the case 1. Absolute and relative biases and MSEs of the Bayesian estimates are less than that of Cox and Cox-Gompertz models both. The Bayesian parameter estimates of model parameters are also similar in all of the cases for *Weibull* distributed data. The cause of this can be the conflict between baseline hazard of the Gompertz distribution and the Weibull distributed survival data. When the survival data come from the Weibull distribution, the Bayesian setting is more successful than the classical setting.

When the survival data comes from the Gompertz distribution, see the Tables 9, 10, and 11, the smallest estimated standard deviations are generated by the classical setting of Cox-Gompertz model, whereas the smallest MSEs are given by the Bayesian setting. The classical Cox model produces the worst standard deviations and MSEs among the classical and Bayesian settings of Cox-Gompertz model. This implies that when distribution of the data and underlying baseline hazard agrees, using Cox-Gompertz model is practically reasonable. The smallest absolute biases are seen in the Bayesian setting. Relative biases of the parameter estimates generated by the classical setting of Cox-Gompertz model are greater than that of the Bayesian setting. While the classical Cox model is being affected by the increased censoring, Cox-Gompertz model generates smaller absolute biases for 0 and 0.1 censoring rates in all of the cases. The same inference is valid for the Bayesian approach in the cases 5 and 6. In general, if one has strong information on the distribution of the lifetime data are distributed as *Gompertz*, use of the Bayesian setting for Cox-Gompertz model is a practically reasonable way.

When the overall results are considered, it is concluded that when survival data come from exponential distribution, Cox model in the classical setting gives the best parameter estimates. But if the data come from Weibull distribution, parameter estimates obtained from all of the settings are not sufficient enough. Thus, a Cox-Weibull model can be applied. When the data is distributed as Gompertz, due to the smallest absolute biases and MSEs produced by the Bayesian setting, advantages of the parametric approach over Cox model and advantages of the Bayesian approach over the classical are ascertained.

When the ratio of number of data sets for which NR method were not converged to the total number of the generated data sets is considered, another advantage of the Bayesian approach is clearly seen. Proportion of unconverged iterations for Cox and Cox-Gompertz models are given in Table 12. Cases seen on the first column are the same as the cases defined in the Tables 3, 6, and 9.

It is seen from Table 12 that NR method encounters certain convergence problems for Cox-Gompertz model for exponential and Weibull distributions, because of its dependency to the starting values. Because NR method had not converged in most of the iterations, thus 1000 samples could not be obtained with reasonable number of generations; and hence , some cells of Tables 3 and 6 could not be filled. Convergence of NR method for Cox model under Weibull and Gompertz distributions were less problematic. In general, the Table 12 reflects the problematic dependency of NR method to the starting values for considered models.

6. Discussion

In this article, we consider use of Gibbs sampling to draw posterior inferences for Cox-Gompertz model, when all of the parameters of the hazard function are unknown. We derive required full conditional distributions for all parameters. All of the full conditionals are found to be familiar and proper distributions. Therefore, there is no need to use a random number generation algorithm such as rejection sampling to generate random numbers from full conditionals. This brings in a flexibility to the presented approach.

Main disadvantage of our approach is that if the survival data is not compatible with the Gompertz distribution, it is not as successful as the classical Cox model in the estimation of parameters. This situation is also observed in the simulation study. However, if this is not the case, our approach is more advantageous than Cox model and classical setting of Cox-Gompertz model. It utilizes superiorities of the Bayesian approaches over the classical counterparts, which are mentioned in the Section 1. Because we are treating

Table 3. The ML estimates of parameters of the Cox model over 1000 samples, each was generated from the exponential distribution.

Case	Par. Values		Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
			0						
1	λ	0.071							
	β_1	-0.100		β_1	0.124	0.224	224.401	6.690	44.808
	β_2	-0.200		β_2	-0.104	0.096	47.794	1.337	1.797
			0.1						
2	λ	0.071							
	β_1	-0.100		β_1	0.089	0.189	189.043	7.461	55.700
	β_2	-0.200		β_2	-0.097	0.103	51.286	1.507	2.281
			0						
3	λ	0.071							
	β_1	-1.000		β_1	-1.120	0.120	-11.970	7.227	52.242
	β_2	-0.200		β_2	-0.329	0.129	-64.442	1.516	2.315
			0.1						
4	λ	0.071							
	β_1	-1.000		β_1	-1.026	0.026	-2.645	$> 10^4$	$> 10^4$
	β_2	-0.200		β_2	-0.499	0.299	-149.646	$> 10^4$	$> 10^4$
			0						
5	λ	0.071							
	β_1	0.500		β_1	0.248	0.252	50.452	4.071	16.637
	β_2	-1.000		β_2	-0.577	0.423	42.287	0.836	0.878
			0.1						
6	λ	0.071							
	β_1	0.500		β_1	0.103	0.397	79.368	4.793	23.129
	β_2	-1.000		β_2	-0.603	0.397	39.693	0.985	1.128

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation; MSE: Mean square error.

all parameters of the hazard function as random, our approach is more precise. The convergence problems of the Gibbs sampling are not as much as NR method, as seen in the simulation study.

Gompertz distribution has many application areas, so does the Bayesian approach to Cox-Gompertz model. Moreover, the Bayesian approach makes the application of the Cox-Gompertz model easier, in all of the mentioned areas, because of the superiorities.

Appendix

A1. Derivation of full conditional distribution of α. Full conditional distribution of α given the other parameters is obtained as

$$p(\alpha|\lambda,\boldsymbol{\beta},\boldsymbol{t}) \propto \exp\left\{\sum_{j=1}^{n}\alpha\delta_j t_j + (\lambda/\alpha)\sum_{j=1}^{n}[1-\exp(\alpha t_j)]\exp(\boldsymbol{\beta}^{'}\boldsymbol{x}_j)\right\}. \tag{11}$$

When we use Taylor expansion of $\exp(\alpha t_j)$ at 0, the following is obtained from eq. (11):

$$p(\alpha|\lambda,\boldsymbol{\beta},\boldsymbol{t}) \propto \exp\left\{\sum_{j=1}^{n}\alpha\delta_j t_j + [\lambda/\alpha]\sum_{j=1}^{n}\left[1-(T_m(\alpha,j)+R_m(\alpha,j))\right]\exp(\boldsymbol{\beta}^{'}\boldsymbol{x}_j)\right\} \tag{12}$$

where m is the order of Taylor expansion, $T_m(\alpha,j) = 1+\alpha t_j+(\alpha^2 t_j^2)/2+\cdots+(\alpha^m t_j^m)/m!$, and $R_m(\alpha,j)$ is the reminder term of the Taylor expansion. Because each term is a function of the rv α, to obtain a tractable full conditional distribution, we need to show that the distribution of $T_m(\alpha,j)+R_m(\alpha,j)$ converges to that of T_m as $m\to\infty$. Let $X_m = \alpha^m t_j^m/m!$ be a sequence of rv's for $m = 1,2,\ldots$ and $Y = \alpha t_j$, where $Y \in \mathbb{R}$.

Table 4. The ML estimates of parameters of the Cox-Gompertz model over 1000 samples, each was generated from the exponential distribution.

Case	Par. Values		Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
			0	α	0.022				
1	λ	0.071		λ	0.283	0.211	-295.976	0.061	0.048
	β_1	-0.100		β_1	1.377	1.477	1477.239	4.381	21.371
	β_2	-0.200		β_2	-0.036	0.164	81.815	2.638	6.985
			0.1	α	*	*	*	*	*
2	λ	0.071		λ	*	*	*	*	*
	β_1	-0.100		β_1	*	*	*	*	*
	β_2	-0.200		β_2	*	*	*	*	*
			0	α	0.010				
3	λ	0.071		λ	2.000	1.928	-2700.492	0.178	3.749
	β_1	-1.000		β_1	-6.559	5.559	-555.935	63.643	4081.362
	β_2	-0.200		β_2	-0.001	0.199	99.635	44.223	1955.697
			0.1	α	0.010				
4	λ	0.071		λ	1.997	1.925	-2696.298	0.201	3.746
	β_1	-1.000		β_1	-6.561	5.561	-556.145	64.936	4247.671
	β_2	-0.200		β_2	-0.001	0.199	99.259	45.049	2029.427
			0	α	0.010				
5	λ	0.071		λ	2.111	2.039	-2856.283	0.127	4.175
	β_1	0.500		β_1	-7.004	7.504	1500.855	5628.965	$> 10^4$
	β_2	-1.000		β_2	-0.009	0.991	99.056	4144.209	$> 10^4$
			0.1	α	0.010				
6	λ	0.071		λ	2.084	2.013	-2819.376	0.139	4.072
	β_1	0.500		β_1	-6.923	7.423	1484.535	6196.135	$> 10^4$
	β_2	-1.000		β_2	-0.004	0.996	99.553	4517.634	$> 10^4$

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation.
*: 10^6 data sets had been generated, but the convergence could not be achieved for 1000 of them.
MSE: Mean square error.

For a fixed value of k, suppose $|y| < k$. Then, for all $m > k$ the following result is straightforwardly obtained:

$$|y|^{m-k} < k(k+1)(k+2)\cdots(m-1).$$

Thus,

$$0 < \frac{|y|^m}{m!} \leq \frac{|y|^k m-k}{m!} < \frac{k(k+1)(k+2)\cdots(m-1)}{m!} = \frac{|y|^k}{(k-1)!m}.$$

In terms of rv's, we have the following inequality for all values of α:

$$X_m \leq \frac{|Y|^k}{(k-1)!m}. \tag{13}$$

The definition of convergence in probability to zero is as follows:

$$\lim_{m\to\infty} P(|X_m| < \epsilon) = 1. \tag{14}$$

The inequality in (13) implies that if

$$\lim_{m\to\infty} P\left(\frac{|Y|^k}{(k-1)!m} < \epsilon\right) = 1, \tag{15}$$

then eq. (14) is ensured. Because k is a fixed constant, the limit in (15) is straightforwardly equal to one. Thus,

$$X_m \xrightarrow{p.} 0, \text{ as } m \to \infty; \text{ and hence } S_m = \sum_{i=m}^{\infty} X_i \xrightarrow{p.} 0, \text{ as } m \to \infty.$$

Table 5. The Bayesian estimates of parameters of the Cox-Gompertz model over 1000 samples, each was generated from the exponential distribution.

Case	Par. Values		Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
			0	α	-0.004			0.002	0.000
1	λ	0.071		λ	0.246	0.175	-244.768	0.058	0.034
	β_1	-0.100		β_1	-0.371	0.271	-270.733	0.009	0.073
	β_2	-0.200		β_2	-0.271	0.071	-35.286	0.005	0.005
			0.1	α	-0.005			0.002	0.000
2	λ	0.071		λ	0.214	0.143	-200.176	0.055	0.023
	β_1	-0.100		β_1	-0.371	0.271	-270.686	0.010	0.073
	β_2	-0.200		β_2	-0.270	0.070	-35.242	0.006	0.005
			0	α	-0.002			0.000	0.000
3	λ	0.071		λ	0.016	0.055	77.072	0.005	0.003
	β_1	-1.000		β_1	-0.371	0.629	62.934	0.014	0.396
	β_2	-0.200		β_2	-0.273	0.073	-36.346	0.008	0.005
			0.1	α	-0.002			0.000	0.000
4	λ	0.071		λ	0.015	0.057	79.510	0.005	0.003
	β_1	-1.000		β_1	-0.371	0.629	62.929	0.015	0.396
	β_2	-0.200		β_2	-0.273	0.073	-36.465	0.009	0.005
			0	α	-0.002			0.001	0.000
5	λ	0.071		λ	0.051	0.021	28.831	0.013	0.001
	β_1	0.500		β_1	-0.371	0.871	174.291	0.010	0.760
	β_2	-1.000		β_2	-0.261	0.739	73.900	0.006	0.546
			0.1	α	-0.002			0.001	0.000
6	λ	0.071		λ	0.043	0.028	39.138	0.012	0.001
	β_1	0.500		β_1	-0.372	0.872	174.331	0.011	0.760
	β_2	-1.000		β_2	-0.261	0.739	73.882	0.006	0.546

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation.
MSE: Mean square error.

Let h be a continuous function at zero, if $Y_m \xrightarrow{p.} 0$ as $m \to \infty$ then $h(Y_m) \xrightarrow{p.} h(0)$ as $m \to \infty$ [17, see Theorem 10.2]. Regarding this theorem, if we define $h(S_m)$ as the following:

$$h(S_m) = \exp\left\{ -[\lambda/\alpha] \sum_{j=1}^{n} \exp(\boldsymbol{\beta}' \boldsymbol{x}_j) S_m \right\},$$

then $h(S_m) \xrightarrow{p.} 1$, as $m \to \infty$. This result implies that the reminder term in (12) converges to 1 *in probability*; and hence, it converges to 1 *in distribution*.

Right hand-side of (12) is rewritten as follows:

$$\exp\left\{ \sum_{j=1}^{n} \alpha\delta_j t_j + [\lambda/\alpha] \sum_{j=1}^{n} \exp(\boldsymbol{\beta}' \boldsymbol{x}_j)\big[1 - T_m(\alpha, j)\big] \right\} \cdot h(S_m). \tag{16}$$

Because the value of n is finite, it concludes from the well-known Slutsky's theorem [17, p. 248] that the expression in (16) converges to the following:

$$\exp\left\{ \sum_{j=1}^{n} \alpha\delta_j t_j + [\lambda/\alpha] \sum_{j=1}^{n} \exp(\boldsymbol{\beta}' \boldsymbol{x}_j)\big[1 - T(\alpha, j)\big] \right\} \cdot 1 \tag{17}$$

in distribution as $m \to \infty$. Consequently, the distribution of the remaining expression after the application of Taylor expansion of order m converges to the distribution of original expression in eq. (11). Therefore, it is appropriate to use the Taylor expansion to derive full conditional distribution of α.

Table 6. The ML estimates of parameters of the Cox model over 1000 samples, each was generated from the Weibull distribution.

Case	Par. Values		Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
1	ν	1.500	0						
	λ	0.015							
	β_1	-0.100		β_1	0.987	1.087	1087.272	60.900	3710.049
	β_2	-0.200		β_2	-3.607	3.407	-1703.717	12.245	161.545
2	ν	1.500	0.1						
	λ	0.015							
	β_1	-0.100		β_1	-0.070	0.030	30.043	57.020	3251.321
	β_2	-0.200		β_2	-3.170	2.970	-1484.821	11.487	140.772
3	ν	1.500	0						
	λ	0.015							
	β_1	-1.000		β_1	-5.077	4.077	-407.700	14.652	231.293
	β_2	-0.200		β_2	-1.124	0.924	-462.180	2.975	9.703
4	ν	1.500	0.1						
	λ	0.015							
	β_1	-1.000		β_1	-7.710	6.710	-670.990	19.151	411.769
	β_2	-0.200		β_2	-1.964	1.764	-882.138	3.931	18.568
5	ν	1.500	0						
	λ	0.015							
	β_1	0.500		β_1	1.446	0.946	-189.221	12.754	163.557
	β_2	-1.000		β_2	-2.107	1.107	-110.657	2.568	7.820
6	ν	1.500	0.1						
	λ	0.015							
	β_1	0.500		β_1	2.260	1.760	-351.977	15.216	234.637
	β_2	-1.000		β_2	-2.453	1.453	-145.310	3.094	11.682

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation.
MSE: Mean square error.

We use the third order Taylor expansion of $\exp(\alpha t_j)$ at 0 to obtain $p(\alpha|\lambda, \boldsymbol{\beta}, \boldsymbol{t})$. As the result it is obtained that

$$\begin{aligned} p(\alpha|\lambda,\boldsymbol{\beta},\boldsymbol{t}) &\propto \exp\left\{\sum_{j=1}^{n}\alpha\delta_j t_j + [\lambda/\alpha]\sum_{j=1}^{n}\left[1-(1+\alpha t_j+\frac{\alpha^2 t_j^2}{2}+\frac{\alpha^3 t_j^3}{6})\right]\exp(\boldsymbol{\beta}'\boldsymbol{x}_j)\right\} \\ &\propto \exp\left\{\frac{-1}{2}\left[\frac{\lambda\alpha^2}{3}\sum_{j=1}^{n}t_j^3 e^{\boldsymbol{\beta}'\boldsymbol{x}_j} - 2\alpha\sum_{j=1}^{n}(\delta_j t_j - \lambda\frac{t_j^2}{2}e^{\boldsymbol{\beta}'\boldsymbol{x}_j})\right]\right\} \\ &\propto \exp\left\{\frac{-1}{2}\frac{\lambda\sum_{j=1}^{n}t_j^3 e^{\boldsymbol{\beta}'\boldsymbol{x}_j}}{3}\left[\alpha^2 - 2\alpha\frac{3\sum_{j=1}^{n}(\delta_j t_j - \lambda\frac{t_j^2}{2}e^{\boldsymbol{\beta}'\boldsymbol{x}_j})}{\lambda\sum_{j=1}^{n}t_j^3 e^{\boldsymbol{\beta}'\boldsymbol{x}_j}}\right]\right\} \\ &\propto \exp\left\{\frac{-1}{2\sigma_\alpha^2}(\alpha-\mu_\alpha)^2\right\}. \end{aligned} \tag{18}$$

Then the full conditional distribution of α is obtained normal distribution with mean and variance

$$\mu_\alpha = \frac{3\sum_{j=1}^{n}(\delta_j t_j - \lambda\frac{t_j^2}{2}e^{\boldsymbol{\beta}'\boldsymbol{x}_j})}{\lambda\sum_{j=1}^{n}t_j^3 e^{\boldsymbol{\beta}'\boldsymbol{x}_j}}, \sigma_\alpha^2 = \left[\frac{\lambda\sum_{j=1}^{n}t_j^3 e^{\boldsymbol{\beta}'\boldsymbol{x}_j}}{3}\right]^{-1}, \tag{19}$$

respectively.

To demonstrate appropriateness of the third order Taylor expansion, we consider the mechanism that generates survival times under the Gompertz model. Bender et al. [2] demonstrate that survival times from Gomperts(α, λ) distribution is generated by the transformation of uniformly distributed r.v. U given in eq. (10). We investigate the impact of the value of α on survival times in Gompertz model over eq. (1). Note that in

Table 7. The ML estimates of parameters of the Cox-Gompertz model over 1000 samples, each was generated from the Weibull distribution.

Case	Par. Values		Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
1	ν	1.500	0	α	0.051	1.449	96.579	0.022	2.099
	λ	0.015		λ	0.179	0.164	-1090.546	0.044	0.029
	β_1	-0.100		β_1	1.287	1.387	1387.414	0.688	2.398
	β_2	-0.200		β_2	-0.046	0.154	76.980	0.514	0.288
2	ν	1.500	0.1	α	*	*	*	*	*
	λ	0.015		λ	*	*	*	*	*
	β_1	-0.100		β_1	*	*	*	*	*
	β_2	-0.200		β_2	*	*	*	*	*
3	ν	1.500	0	α	0.010	1.490	99.329	18.455	342.805
	λ	0.015		λ	2.222	2.207	-14713.681	0.136	4.890
	β_1	-1.000		β_1	-7.373	6.373	-637.282	5115.662	26170042.877
	β_2	-0.200		β_2	0.000	0.200	99.910	3801.955	14454860.217
4	ν	1.500	0.1	α	0.010	1.490	99.322	12.018	146.645
	λ	0.015		λ	2.178	2.163	-14417.268	0.147	4.699
	β_1	-1.000		β_1	-7.234	6.234	-623.372	3115.142	9704147.295
	β_2	-0.200		β_2	-0.001	0.199	99.713	2252.882	5075475.292
5	ν	1.500	0	α	*	*	*	*	*
	λ	0.015		λ	*	*	*	*	*
	β_1	0.500		β_1	*	*	*	*	*
	β_2	-1.000		β_2	*	*	*	*	*
6	ν	1.500	0.1	α	*	*	*	*	*
	λ	0.015		λ	*	*	*	*	*
	β_1	0.500		β_1	*	*	*	*	*
	β_2	-1.000		β_2	*	*	*	*	*

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation.
*: 10^6 data sets had been generated, but the convergence could not be achieved for 1000 of them.
MSE: Mean square error.

eq. (1), $\lambda > 0$, $\exp\{\beta_1 x_{j1} + \beta_2 x_{j2}\} > 0$, and $\log(u) < 0$. Because

$$\lim_{\alpha\to-\infty} T_j^{\mathrm{G}} = \lim_{\alpha\to\infty} T_j^{\mathrm{G}} = 0,$$

survival times goes to zero for greater values of α. For smaller values of $\lambda\exp\{\beta_1 x_{j1} + \beta_2 x_{j2}\}$, the value of U should approach to one to make eq. (1) proper. Only for this case the rate of convergence of T_j^{G} to zero decreases; and hence, we can observe reasonable survival times for greater values of α. Due to the decreased range of reasonable values of U, the probability of having such a situation in practice is small. For greater values of $\lambda\exp\{\beta_1 x_{j1} + \beta_2 x_{j2}\}$, any value of U from $(0,1)$ interval makes eq. (1) proper. In this case, values of α close to zero give reasonable survival times. Therefore, the rate of convergence will be very fast due to the small values of α; and hence, use of the third order Taylor expansion is appropriate.

A2. Derivation of full conditional distribution of λ. To derive the $p(\lambda|\alpha,\boldsymbol{\beta},\boldsymbol{t})$, (6) is rewritten by discarding the constants as

$$p(\lambda|\alpha,\boldsymbol{\beta},\boldsymbol{t}) \propto \lambda^{\sum_{j=1}^{n}\delta_j} \exp\Big\{ -(\lambda/\alpha)\sum_{j=1}^{n}[\exp(\alpha t_j)-1]\exp(\boldsymbol{\beta}'\boldsymbol{x}_j)\Big\}. \tag{20}$$

The distribution reached in (14) is gamma with the following shape and scale parameters

$$\sum_{j=1}^{n}\delta_j + 1,\ \alpha\Big[\sum_{j=1}^{n}[\exp(\alpha t_j)-1]\exp(\boldsymbol{\beta}'\boldsymbol{x}_j)\Big]^{-1}. \tag{21}$$

Table 8. The Bayesian estimates of parameters of the Cox-Gompertz model over 1000 samples, each was generated from the Weibull distribution.

Case	Par. Values		Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
1	ν	1.500	0	α	-0.005	1.505	100.329	0.004	2.265
	λ	0.015		λ	0.342	0.327	-2179.422	0.081	0.113
	β_1	-0.100		β_1	-0.371	0.271	-270.731	0.009	0.073
	β_2	-0.200		β_2	-0.272	0.072	-35.797	0.006	0.005
2	ν	1.500	0.1	α	-0.006	1.506	100.426	0.004	2.269
	λ	0.015		λ	0.301	0.286	-1906.160	0.077	0.088
	β_1	-0.100		β_1	-0.371	0.271	-270.931	0.010	0.073
	β_2	-0.200		β_2	-0.271	0.071	-35.532	0.006	0.005
3	ν	1.500	0	α	-0.002	1.502	100.102	0.001	2.255
	λ	0.015		λ	0.044	0.029	-190.718	0.012	0.001
	β_1	-1.000		β_1	-0.370	0.630	63.018	0.010	0.397
	β_2	-0.200		β_2	-0.271	0.071	-35.593	0.007	0.005
4	ν	1.500	0.1	α	-0.002	1.502	100.113	0.001	2.255
	λ	0.015		λ	0.038	0.023	-153.024	0.011	0.001
	β_1	-1.000		β_1	-0.370	0.630	63.024	0.012	0.397
	β_2	-0.200		β_2	-0.271	0.071	-35.465	0.007	0.005
5	ν	1.500	0	α	-0.002	1.502	100.139	0.001	2.256
	λ	0.015		λ	0.106	0.091	-604.296	0.026	0.009
	β_1	0.500		β_1	-0.372	0.872	174.431	0.010	0.761
	β_2	-1.000		β_2	-0.256	0.744	74.389	0.006	0.553
6	ν	1.500	0.1	α	-0.002	1.502	100.159	0.001	2.257
	λ	0.015		λ	0.094	0.079	-525.466	0.025	0.007
	β_1	0.500		β_1	-0.372	0.872	174.458	0.010	0.761
	β_2	-1.000		β_2	-0.256	0.744	74.359	0.006	0.553

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation.
MSE: Mean square error.

A3. Derivation of full conditional distribution of β_i. With the same manner as in Appendix A1, full conditional distribution of a particular regression parameter given the others is obtained by using the Taylor expansion. Then, $p(\beta_i|\beta_{-i},\alpha,\lambda,\boldsymbol{t})$ is obtained by discarding the constants as follows:

$$p(\beta_i|\beta_{-i},\alpha,\lambda,\boldsymbol{t}) \propto \exp\left\{\beta_i\sum_{j=1}^{n}\delta_j x_{ji} + (\lambda/\alpha)\sum_{j=1}^{n}[1-\exp(\alpha t_j)]\exp(x_{ji}\beta_i)c_j\right\}, \tag{22}$$

where $c_j = \exp(\sum_{k=1,k\neq i}^{n}\beta_k x_{jk})$. It is obtained using the second order Taylor expansion of $\exp(x_{ji}\beta_i)$ at 0 that

$$\begin{aligned} p(\beta_i|\beta_{-i},\alpha,\lambda,\boldsymbol{t}) \propto \exp\Bigg\{\beta_i\Big[\sum_{j=1}^{n}\delta_j x_{ji} - (\lambda/\alpha)\sum_{j=1}^{n}c_j x_{ji}[\exp(\alpha t_j)-1]\Big] \\ -[\lambda/(2\alpha)]\beta_i^2\sum_{j=1}^{n}c_j x_{ji}^2[\exp(\alpha t_j)-1]\Bigg\} \end{aligned} \tag{23}$$

by simply arranging (23),

$$p(\beta_i|\beta_{-i},\alpha,\lambda,\boldsymbol{t}) \propto \exp\left\{\frac{-1}{2s_1}\left[\beta_i - s_2/s_1\right]^2\right\}, \tag{24}$$

where $s_1 = (\lambda/\alpha)\sum_{j=1}^{n}c_j x_{ji}^2[\exp(\alpha t_j)-1]$ and $s_2 = \sum_{j=1}^{n}\delta_j x_{ji} - (\lambda/\alpha)\sum_{j=1}^{n}c_j x_{ji}\times[\exp(\alpha t_j)-1]$. Then $p(\beta_i|\beta_{-i},\alpha,\lambda,\boldsymbol{t})$ is approached by the normal distribution with mean s_2/s_1 and variance $1/s_1$.

As for the appropriateness of the second order Taylor expansion, we evaluate the impact of the value of β_i on survival times in Gompertz model as done in Appendix A1. Regarding the second equation in (10), we have the following results for the fixed values

Table 9. The ML estimates of parameters of the Cox model over 1000 samples, each was generated from the Gompertz distribution.

Case	Par. Values		Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
1	α	0.000	0						
	λ	0.010							
	β_1	-0.100		β_1	-0.012	0.088	88.380	6.763	45.746
	β_2	-0.200		β_2	-0.362	0.162	-81.164	1.345	1.834
2	α	0.000	0.1						
	λ	0.010							
	β_1	-0.100		β_1	-0.959	0.859	-859.053	11.972	144.077
	β_2	-0.200		β_2	-0.658	0.458	-228.872	2.328	5.631
3	α	0.000	0						
	λ	0.100							
	β_1	-1.000		β_1	-2.257	1.257	-125.700	9.299	88.055
	β_2	-0.200		β_2	-0.627	0.427	-213.332	1.938	3.940
4	α	0.000	0.1						
	λ	0.100							
	β_1	-1.000		β_1	-4.193	3.193	-319.286	14.321	215.294
	β_2	-0.200		β_2	-1.439	1.239	-619.288	3.076	10.999
5	α	-0.001	0						
	λ	1.000							
	β_1	0.500		β_1	0.254	0.246	49.240	5.368	28.872
	β_2	-1.000		β_2	-0.087	0.913	91.284	1.070	1.979
6	α	-0.001	0.1						
	λ	1.000							
	β_1	0.500		β_1	-0.046	0.546	109.288	6.397	41.216
	β_2	-1.000		β_2	-0.168	0.832	83.214	1.289	2.353

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation.
MSE: Mean square error.

of α, λ, and β_{-i}:

$$\lim_{\beta_i \to -\infty} T_j^{\mathrm{G}} = \infty \text{ and } \lim_{\beta_i \to \infty} T_j^{\mathrm{G}} = 0.$$

Values close the $-\infty$ are unreasonable and greater values give nearly zero survival times. Positive and larger values of β_i correspond to reasonable survival times for very small values of U; hence, probability of occurrence of this situation is small. Accordingly, small values of β_i will correspond to reasonable survival times in practice. Therefore, the rate of convergence will be very fast; and hence, use of the second order Taylor expansion is appropriate.

References

[1] Ananda, M.M., Dalpatadu, R.J. and Singh, A.K. *Adaptive Bayes estimators for parameters of the Gompertz survival model*, Applied Mathematics and Computation, **75** (2), 167-177, 1996.

[2] Bender, R., Augustin, T., Blettner, M. *Generating survival times to simulate Cox proportional hazards models*, Statistics in Medicine, **24**, 1713-1723, 2005.

[3] Cantner, U., DreBler, K., Kruger, J.J. *Firm Survival in the German Automobile Industry*, Empirica, **33**, 49-60, 2006.

[4] Collett, D. *Modelling Survival Data in Medical Research*, Chapman&Hall, UK, 1994.

[5] Cox, D.R. *Regression models and life tables (with discussion)*, Journal of the Royal Statistical Society : Series B, **34**, 187-220, 1972.

[6] Dellaportas, P. and Smith, A.F.M. *Bayesian inferences of generalized linear and proportional hazards models via Gibbs sampling*, Applied Statistics, **42** (3), 443-459, 1993.

[7] Efron, B. *The efficiency of Cox's likelihood function for censored data*", Journal of the American Statistical Association, **72**, 557-565, 1977.

[8] Fabrizio, C. *New Evidence on the Politics and Economics of Multiparty Cabinets Duration*, Scottish Journal of Political Economy, **49** (3), 249-279, 2002.

Table 10. The ML estimates of parameters of the Cox-Gompertz model over 1000 samples, each was generated from the Gompertz distribution.

Case		Par. Values	Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
1	α	0.000	0	α	-0.004	0.004	3854.290	0.000	0.000
	λ	0.010		λ	0.127	0.117	-1170.553	0.000	0.014
	β_1	-0.100		β_1	1.315	1.415	1414.505	0.000	2.001
	β_2	-0.200		β_2	-0.120	0.080	39.925	0.000	0.006
2	α	0.000	0.1	α	-0.004	0.004	4005.207	0.000	0.000
	λ	0.010		λ	0.170	0.160	-1595.591	0.000	0.025
	β_1	-0.100		β_1	1.303	1.403	1403.429	0.000	1.970
	β_2	-0.200		β_2	-0.208	0.008	-3.827	0.000	0.000
3	α	0.000	0	α	-0.004	0.004	395797.522	0.000	0.000
	λ	0.100		λ	0.207	0.107	-106.725	0.000	0.011
	β_1	-1.000		β_1	0.542	1.542	154.200	0.000	2.378
	β_2	-0.200		β_2	-0.132	0.068	33.773	0.000	0.005
4	α	0.000	0.1	α	-0.004	0.004	395723.040	0.000	0.000
	λ	0.100		λ	0.179	0.079	-79.196	0.000	0.006
	β_1	-1.000		β_1	0.709	1.709	170.875	0.000	2.920
	β_2	-0.200		β_2	-0.170	0.030	15.136	0.000	0.001
5	α	-0.001	0	α	-0.003	0.002	-198.367	0.000	0.000
	λ	1.000		λ	0.659	0.341	34.143	0.000	0.117
	β_1	0.500		β_1	2.092	1.592	-318.486	0.000	2.536
	β_2	-1.000		β_2	-1.200	0.200	-19.957	0.000	0.040
6	α	-0.001	0.1	α	-0.007	0.006	-588.075	0.000	0.000
	λ	1.000		λ	0.522	0.478	47.787	0.000	0.228
	β_1	0.500		β_1	2.246	1.746	-349.147	0.000	3.048
	β_2	-1.000		β_2	-1.226	0.226	-22.640	0.000	0.051

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation.
MSE: Mean square error.

[9] Gelfand, A.E. and Mallick, B.K. *Bayesian analysis of proportional hazards built from monotone functions*, Biometrics, **51**, 843-852, 1995.

[10] Gelfand,A.E. and Sahu, S.K. *Identifiability, improper priors, and Gibbs sampling for generalized linear models*, Journal of the American Statistical Association, **94**, 247-253, 1999.

[11] Gelman, A. *Inference and monitoring convergence*, in Markov Chain Monte Carlo (eds. W.R. Gilks, S. Richardson, D.J. Spiegelhalter)", Chapman&Hall/CRC, London, 1996.

[12] Ghitanya, M.E., Alqallafa, F., Balakrishnan, N. *On the likelihood estimation of the parameters of Gompertz distribution based on complete and progressively Type-II censored samples*, Journal of Statistical Computation and Simulation, **84** (8), 1803-1812, 2014.

[13] Gieser, P., Chang, M.N., Rao, P.V., Shuster, J.J. and Pullen, J. *Modelling cure rate using the Gompertz model with covariate information*, Statistics in Medicine, **17**, 831-839, 1998.

[14] Gokovali, U., Bahar, O., Kozak, M. *Determinants of length of stay: A practical use of survival analysis*, Tourism Management, **28**, 736-746, 2007.

[15] Gompertz, B. *On the nature of the function expresssive of the law of human mortality, and on a new method of determining the value of life contingencies*, Phil. Trans. R. Soc., **36**, 513-585, 1825.

[16] Grunkemeier, G.L., Jamieson, W.R.E., Miller, D.C. and Starr, A. *Surgery for Acquired Heart Disease*, J Thorac Cardiovasc Surg, **108**, 709-718, 1994.

[17] Gut, A. *Probability: A Graduate Course*, Springer, New York, 2013.

[18] Jeong, J-H., *Parametric regression on cumulative incidence function*, Biostatistics, **8** (2), 184-196, 2007.

[19] Jeong, J-H. and Fine, J. *Direct parametric inference for the cumulative incidence function*, Journal of the Royal Statistical Society : Series C, **55** (2), 187-200, 2006.

[20] Johnson, N.L., Kotz, S. and Balakrishnan, N. *Continuous Univariate Distributions, Vol. 1.*, Wiley, New-York, 1994.

[21] Kim, S.W. and Ibrahim, J.G. "*On bayesian inference for proportional hazards models using noninformative priors*, Lifetime Data Analysis, **6**, 331-341, 2000.

Table 11. The Bayesian estimates of parameters of the Cox-Gompertz model over 1000 samples, each was generated from the Gompertz distribution.

Case	Par. Values		Cens. Rate	Par.	Estimate	Absolute Bias	Relative Bias	St.Dev.	MSE
1	α	0.000	0	α	-0.003	0.003	2730.118	0.001	0.000
	λ	0.010		λ	0.056	0.046	-458.052	0.011	0.002
	β_1	-0.100		β_1	-0.768	0.668	-668.051	0.017	0.447
	β_2	-0.200		β_2	-0.561	0.361	-180.716	0.010	0.131
2	α	0.000	0.1	α	-0.004	0.004	4442.027	0.001	0.000
	λ	0.010		λ	0.078	0.068	-677.256	0.013	0.005
	β_1	-0.100		β_1	-1.249	1.149	-1149.071	0.023	1.321
	β_2	-0.200		β_2	-0.913	0.713	-356.536	0.014	0.509
3	α	0.000	0	α	-0.004	0.004	355914.065	0.000	0.000
	λ	0.100		λ	0.056	0.044	43.945	0.009	0.002
	β_1	-1.000		β_1	-1.342	0.342	-34.244	0.025	0.118
	β_2	-0.200		β_2	-0.985	0.785	-392.406	0.015	0.616
4	α	0.000	0.1	α	-0.006	0.006	587070.440	0.001	0.000
	λ	0.100		λ	0.080	0.020	19.849	0.011	0.001
	β_1	-1.000		β_1	-2.222	1.222	-122.246	0.035	1.496
	β_2	-0.200		β_2	-1.629	1.429	-714.308	0.021	2.041
5	α	-0.001	0	α	-0.012	0.011	-1070.242	0.006	0.000
	λ	1.000		λ	0.946	0.054	5.370	0.202	0.044
	β_1	0.500		β_1	-0.454	0.954	190.825	0.009	0.910
	β_2	-1.000		β_2	-0.312	0.688	68.771	0.006	0.473
6	α	-0.001	0.1	α	-0.014	0.013	-1309.660	0.006	0.000
	λ	1.000		λ	0.852	0.148	14.843	0.189	0.058
	β_1	0.500		β_1	-0.484	0.984	196.853	0.011	0.969
	β_2	-1.000		β_2	-0.333	0.667	66.712	0.006	0.445

Par. : Parameter; Cens. : Censoring; St. Dev. : Estimated Standard Deviation.
MSE: Mean square error.

Table 12. Table 12: Proportion of unconverged iterations for Cox and Cox-Gompertz models

	Exponential Dist.		Weibull Dist.		Gompertz Dist.	
Case	CM	CGM	CM	CGM	CM	CGM
1	0.190	0.998	0.000	0.998	0.000	0.492
2	0.184	0.999	0.000	0.999	0.002	0.688
3	0.711	0.998	0.000	0.998	0.003	0.708
4	0.769	0.998	0.001	0.999	0.038	0.827
5	0.007	0.998	0.000	0.999	0.000	0.015
6	0.008	0.998	0.000	0.999	0.000	0.075

CM : the Cox-Model; CGM : the Cox-Gompertz Model

[22] Klein, J.P. and Moeschberger M.L. *Survival Analysis : Techniques for Censored and Truncated Data*, Springer, New York, 1997.

[23] Klepper, S. *The Capabilities of New Firms and the Evolution of the U.S. Automobile Industry*, Industrial and Corporate Change, **11**, 645-666, 2002.

[24] Laubender, R.P. and Bender, R. *Estimating adjusted risk difference (RD) and number needed to treat (NNT) measures in the Cox regression model*, Statistics in Medicine, **29**, 851-859, 2010.

[25] Lawless, J.F. *Statistical Models and Methods for Lifetime Data*, Wiley, New York, 1982.

[26] Lee, E.T. and Wang, J.W. *Statistical Methods for Survival Data Analysis*, Wiley, New York, 2003.

[27] Makeham, W.M. *On the law of mortality, and the construction of annuity tables*, Journal of the Institute of Actuaries, **13**, 325-358, 1860.

[28] Nardi, A. and Schemper, M. *Comparing Cox and parametric models in clinical studies*, Statistics in Medicine, **22**, 3597-3610, 2003.

[29] Oakes, D. *The asymptotic information in censored survival data*, Biometrika, **64**, 441-448, 1977.

[30] Spickett, A.M. and van Ark, H. *Fitting the Gompertz function to dose-response data oflarval tick populations*, Onderstepoort J Vet Res., **57** (2), 155-118, 1990.

[31] Willekens, F. *Gompertz in context: the Gompertz and related distributions*, In Forecasting Mortality in Developed Countries - Insights from a Statistical, Demographic and Epidemiological Perspective, edited by E. Tabeau, A. van den Berg Jeths, and C. Heathcote, volume 9 of European Studies of Population, 105-126, Springer, 2002.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (5) (2016), 1641 – 1660

An improved class of estimators for finite population variance

Mazhar Yaqub*† and Javid Shabbir‡

Abstract

We propose an improved class of estimators in estimating the finite population variance, using the auxiliary information. The expressions for the bias and mean squared error of the proposed class of estimators are derived up to the first order of approximation. Some estimators are also derived from a proposed class by allocating the suitable values of known parameters and identified as particular members of the proposed class of estimators. A numerical study is carried out to demonstrate performances of the estimators. It is observed that the proposed class of estimators is more efficient than the usual sample mean estimator, the regression estimator suggested by Isaki (1983), Shabbir and Gupta (2007), Singh and Solanki (2013*b*), Yadav et al. (2013), Yadav and Kadilar (2014) and Singh and Malik (2014) estimators.

Keywords: Auxiliary variable; bias; mean squared error; population variance; percentage relative efficiency.

2000 AMS Classification: Primary 62D05, Secondary 62P20.

Received : 09.05.2014 *Accepted* : 17.04.2015 *Doi* : 10.15672/HJMS.20156310746

1. Introduction

In this article, an improved class of estimators is proposed in estimating the finite population variance under simple random sampling. Various fields of life like genetics, biology and medical studies have been facing the problem in estimating the finite population variance. An agriculturist requires sufficient knowledge of climatic variation to devise appropriate plan for cultivating his crop. A fair understanding of variability is vitally important for better results in different walks of life. Singh et al. [36] and Das and Tripathi [9] have proposed different estimators for finite population variance $\left(S_y^2\right)$.

*Corresponding Author.
†Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
Email : javidshabbir@gmail.com
‡Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.

Chaudhury [8] and Mukhopadhyay [27, 28] have made significant contributions in estimating the finite population variance under super population models. Srivastava and Jhajj [47] and Wu [51], have taken the advantage of correlation between the study and the auxiliary variables in estimating the finite population variance. Isaki [17] has proposed ratio and regression type estimators for population variance. Likewise, Singh [37], Searls and Intrapanich [31], and Prasad and Singh [29, 30] have given some estimators for population variance. Singh and Biradar [39], Garcia and Cebrain [10], Cebrain and Garcia [5], Singh and Joarder [40], Upadhyaya and Singh [49], and Ahmed et al. [2] have paid their attention towards the improved estimation or classes of estimators of S_y^2. AL-Jaraha and Ahmed [3] have given some chain ratio-type as well as chain product-type estimators of S_y^2 using double-sampling scheme. Later on Singh and Singh [41], Upadhyaya et al. [50], Chandra and Singh [7], Arcos et al. [1], Kadilar and Cingi [18, 19, 20, 21, 22, 23], Koyuncu and Kadilar [24, 25], Singh and Vishwakarma [43], Turgut and Cingi [48], Gupta and Shabbir [12, 13, 14], Shabbir and Gupta [33, 34], Grover [11], Singh et al. [38, 42, 44], Yadav et al. [52, 53], Yadav and Kadilar [54], Singh and Solanki [45, 46], and Singh and Malik [35] have paid their attention towards the improved estimation of population variance S_y^2.

Motivated by these studies, the present article focuses on improved class of estimators for S_y^2 using the auxiliary information.

The rest of the article is organized as: Section 2 provides the notations and symbols. Section 3 gives a brief review of some existing estimators of S_y^2. Section 4 gives the expressions for the bias and mean squared error (*MSE*) of the proposed class of estimators. The efficiency comparison of different estimators is shown in Section 5. A numerical study is presented in Section 6. Conclusion is given in Section 7.

2. Notations

Consider a finite population $\Omega = \{1, 2, .., i, .., N\}$ having N units. We draw a sample of size n by using simple random sample without replacement (SRSWOR) sampling scheme from this population. Let y_i and x_i be the values of the study variable (y) and the auxiliary variable (x) respectively. Let $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the sample means, respectively, corresponding to the population means $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ and $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$. Let $s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ and $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ be the sample variances corresponding to population variances $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2$ and $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2$, respectively. Let $\theta_{22} = \frac{\mu_{11}}{\sqrt{\mu_{20}}\sqrt{\mu_{02}}}$, be the covariance between S_y^2 and S_x^2. Let $\beta_{2(y)} = \frac{\mu_{40}}{\mu_{20}^2}$ and $\beta_{2(x)} = \frac{\mu_{04}}{\mu_{02}^2}$ be the population coefficients of kurtosis of y and x, respectively, where $\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^r\left(x_i - \bar{X}\right)^s$ and $\gamma = 1/n$. We ignored the finite population correction(fpc) term because of ease of computation. In order to get biases and *MSEs* of the considered estimators, we use the following relative error terms.

Let $\delta_0 = \frac{s_y^2 - S_y^2}{S_y^2}$ and $\delta_1 = \frac{s_x^2 - S_x^2}{S_x^2}$ such that $E(\delta_i) = 0$ for $i = 0, 1$.

Let $E\left(\delta_0^2\right) = \gamma\left(\beta_{2(y)} - 1\right) \cong V_{20}$, $E\left(\delta_1^2\right) = \gamma\left(\beta_{2(x)} - 1\right) \cong V_{02}$, and $E(\delta_0\delta_1) = \gamma(\theta_{22} - 1) \cong V_{11}$, where $V_{rs} = E\left\{\frac{(s_y^2 - S_y^2)^r (s_x^2 - S_x^2)^s}{(S_y^2)^r (S_x^2)^s}\right\}$.

3. Some existing estimators

We discuss the following estimators.

(i) The variance of the usual unbiased variance estimator $\left(\hat{S}_y^2\right)$, is given by

(3.1) $$Var\left(\hat{S}_y^2\right) = S_y^4 V_{20}.$$

(*ii*) Isaki [17] suggested the following regression estimator for S_y^2, given by

(3.2) $$\hat{S}_{Reg}^2 = s_y^2 + b_{\left(s_y^2, s_x^2\right)}\left(S_x^2 - s_x^2\right),$$

where $b_{\left(s_y^2, s_x^2\right)}$ is the sample regression coefficient whose population regression coefficient is $\beta = \left(S_y^2 V_{11} / S_x^2 V_{02}\right)$. The variance of $\hat{S}_{Reg}^2$, is given by

(3.3) $$Var(\hat{S}_{Reg}^2) \cong S_y^4 V_{20}\left(1 - \rho_{\left(s_y^2, s_x^2\right)}^2\right) = MSE(\hat{S}_{Reg}^2),$$

where $\rho_{\left(s_y^2, s_x^2\right)} = \left(V_{11}/\sqrt{V_{20}V_{02}}\right)$.

(*iii*) Singh et al. [38] considered the following difference type estimator for S_y^2, given by

(3.4) $$\hat{S}_d^2 = k_1 s_y^2 + k_2\left(S_x^2 - s_x^2\right),$$

where k_1 and k_2 are suitably chosen constants.
The bias and minimum *MSE* of $\hat{S}_d^2$, to first order of approximation, at optimum values $k_1^{(opt)} = \frac{V_{02}}{\left(V_{02} + V_{02}V_{20} - V_{11}^2\right)}$ and $k_2^{(opt)} = \frac{S_x^2}{S_y^2}\frac{V_{11}}{\left(V_{02} + V_{02}V_{20} - V_{11}^2\right)}$,
are given by

$$Bias(\hat{S}_d^2) \cong (k_1 - 1)\, S_y^2,$$

and

(3.5) $$MSE_{\min}(\hat{S}_d^2) \cong \frac{Var(\hat{S}_{Reg}^2)}{1 + S_y^{-4} Var(\hat{S}_{Reg}^2)}.$$

(*iv*) Shabbir and Gupta [32] suggested the following ratio-type exponential estimator S_y^2, given by

(3.6) $$\hat{S}_{SG}^2 = \left\{k_3 s_y^2 + k_4\left(S_x^2 - s_x^2\right)\right\}\exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right),$$

where k_3 and k_4 are suitably chosen constants.
The bias of $\hat{S}_{SG}^2$, to first order of approximation, is given by

$$Bias\left(\hat{S}_{SG}^2\right) \cong S_y^2(k_3 - 1) + \frac{3}{8}k_3 S_y^2 V_{02} - \frac{1}{2}k_3 S_y^2 V_{11} + \frac{1}{2}k_4 S_x^2 V_{02}.$$

The minimum *MSE* of $\hat{S}_{SG}^2$, to first order of approximation, at optimum values of $k_3^{(opt)} = \frac{V_{02}}{8}\left(\frac{8 - V_{02}}{V_{02} + V_{20}V_{02} - V_{11}^2}\right)$ and $k_4^{(opt)} = \frac{S_y^2}{8S_x^2}\left(\frac{-4V_{02} + V_{02}^2 + 8V_{11}^2 - V_{11}V_{02} + 4V_{20}V_{02} - 4V_{11}^2}{V_{02} + V_{20}V_{02} - V_{11}^2}\right)$,
is given by

(3.7) $$MSE_{\min}\left(\hat{S}_{SG}^2\right) \cong \frac{S_y^4}{64}\left\{\frac{V_{02}^2 + 16\left(V_{02} - 4\right) S_y^{-4} Var(\hat{S}_{Reg}^2)}{-1 - S_y^{-4} Var(\hat{S}_{Reg}^2)}\right\}.$$

(*v*) Singh and Solanki [46] suggested a difference-in-ratio type estimator for S_y^2, given by

(3.8) $$\hat{S}_{SS}^2 = \left\{k_5 s_y^2 + k_6(S_x^2 - s_x^2)\right\}\left(\frac{aS_x^2 + b}{as_x^2 + b}\right),$$

where k_5 and k_6 are suitably chosen constants and $a\,(\neq 0)$ and b are functions of known parameters of the auxiliary variable x.
The bias of $\hat{S}_{SS}^2$, to first order of approximation, is given by

$$Bias\left(\hat{S}_{SS}^2\right) \cong S_y^2(k_5 - 1) + k_5 S_y^2 \tau^2 V_{02} - k_5 S_y^2 \tau V_{11} + k_6 S_x^2 \tau V_{02},$$

where $\tau = aS_x^2/\left(aS_x^2+b\right)$.
The minimum MSE of $\hat{S}^2_{SS}$, to first order of approximation, at optimum values
$k_5^{(opt)} = \frac{V_{02}(-1+\tau^2 V_{02})}{(-V_{02}-V_{02}V_{20}+\tau^2 V_{02}^2+V_{11}^2)}$ and
4 $k_6^{(opt)} = -\frac{S_y^2}{S_x^2}\left(\frac{-V_{02}\tau+V_{02}^2\tau^3+V_{11}-V_{11}V_{02}\tau^2+\tau V_{02}V_{20}-\tau V_{11}^2}{-V_{02}-V_{02}V_{20}+\tau^2 V_{02}^2+V_{11}^2}\right)$,
is given by

$$(3.9)\qquad MSE_{\min}\left(\hat{S}^2_{SS}\right) \cong S_y^4\left\{\frac{AS_y^{-4}Var(\hat{S}^2_{Reg})}{A+S_y^{-4}Var(\hat{S}^2_{Reg})}\right\},$$

where $A = 1 - V_{02}\tau^2$.

(vi) Yadav et al. [53] suggested a general class of estimators for S_y^2, given by

$$(3.10)\quad \hat{S}^2_{YG} = \left\{k_7 s_y^2 + k_8(S_x^2 - s_x^2)\right\}\left\{\lambda\left(\frac{aS_x^2+b}{as_x^2+b}\right) + (1-\lambda)\exp\left(\frac{a\left(S_x^2-s_x^2\right)}{a\left(S_x^2+s_x^2\right)+2b}\right)\right\},$$

where k_7 and k_8 are suitably chosen constants, λ can takes values 0 or 1 and a, b be the population parameters of the auxiliary variables. Shabbir and Gupta [32] estimator in (3.6) and Singh and Solanki [46] estimator in (3.8) can be generated from (3.10) by substituting the suitable choices of λ, a and b. The bias of $\hat{S}^2_{YG}$, to first order of approximation, is given by

$$Bias\left(\hat{S}^2_{YG}\right) \cong S_y^2\left[(k_7-1) + \left(\frac{3+5\lambda}{8}\right)k_7 V_{02}\tau^2 + \frac{(1+\lambda)\tau}{2}\left\{\left(\frac{S_x^2}{S_y^2}\right)k_8 V_{02} - k_7 V_{11}\right\}\right].$$

The minimum MSE of $\hat{S}^2_{YG}$, to first order of approximation, at optimum values

$$k_7^{(opt)} = \frac{V_{02}}{2}\left\{\frac{8 - V_{02}\tau^2\left(1+3\lambda+4\lambda^2\right)}{4V_{02} - V_{02}^2\lambda\tau^2\left(1+3\lambda\right) + 4V_{02}V_{20} - 4V_{11}^2}\right\},$$

and

$$k_8^{(opt)} = \frac{S_y^2}{2S_x^2}\left[\frac{\left\{\begin{array}{c} 8V_{11} + 3V_{02}^2\lambda\tau^2\left(1+\lambda\right) - 4V_{02}\tau\left(1+\lambda\right) \\ +V_{02}^2\tau^3\left(1+\lambda^3\right) - V_{02}V_{11}\tau^2\left(1+3\lambda+4\lambda^2\right) \\ +4V_{02}V_{20}\tau\left(1+\lambda\right) - 4V_{11}^2\left(1+\lambda\right)\end{array}\right\}}{4V_{02}\left(1-V_{02}\tau^2\lambda^2\right) - V_{02}^2\tau^2\lambda\left(1+\lambda\right) + 4V_{02}V_{20} - 4V_{11}^2}\right],$$

is given by

$$(3.11)\qquad MSE_{\min}\left(\hat{S}^2_{YG}\right) \cong \frac{S_y^4}{16}\left[\frac{\begin{array}{c}\left\{(1+\lambda)^2 - 4\lambda^2\left(1+\lambda-\lambda^2\right)\right\}V_{02}^2\tau^4 \\ +16S_y^{-4}Var(\hat{S}^2_{Reg})\left\{(1+\lambda)^2 V_{02}\tau^2 - 4\right\}\end{array}}{-4+3\lambda^2\tau^2 V_{02} + \lambda\tau^2 V_{02} - 4S_y^{-4}Var(\hat{S}^2_{Reg})}\right].$$

(vii) Yadav and Kadilar [54] suggested two parameters ratio-product-ratio type estimator for S_y^2, given by

$$(3.12)\qquad \hat{S}^2_{YK} = s_y^2\left[\alpha_1\left\{\frac{(1-\beta_1)s_x^2 + \beta_1 S_x^2}{\beta_1 s_x^2 + (1-\beta_1)S_x^2}\right\} + (1-\alpha_1)\left\{\frac{\beta_1 s_x^2 + (1-\beta_1)S_x^2}{(1-\beta_1)s_x^2 + \beta_1 S_x^2}\right\}\right],$$

where α_1 and β_1 are suitably chosen constants.
The bias and MSE of $\hat{S}^2_{YK}$, to first order of approximation, are given by

$$Bias\left(\hat{S}^2_{YK}\right) \cong S_y^2\left\{V_{02}\left(1-\alpha_1-3\beta_1+2\alpha_1\beta_1+2\beta_1^2\right)-V_{11}\left(1-2\alpha_1-2\beta_1+4\alpha_1\beta_1\right)\right\}-S_y^2,$$

and

$$MSE\left(\hat{S}^2_{YK}\right) \cong S_y^4\left\{\begin{array}{c}\left(V_{02}+V_{20}-2V_{11}\right)+16\alpha_1\beta_1V_{02}\left(1-\alpha_1-\beta_1+\alpha_1\beta_1\right)\\ +4V_{11}\left(\alpha_1-\beta_1\right)^2+4V_{02}\left(-\alpha_1-\beta_1+\alpha_1^2+\beta_1^2\right)\end{array}\right\}.$$

Solving above for minimum MSE of $\hat{S}^2_{YK}$, to first order of approximation at $(\alpha_1, \beta_1) = (1/2, 1/2)$, is

$$(3.13) \qquad MSE_{\min}\left(\hat{S}^2_{YK}\right) \cong Var\left(\hat{S}^2_y\right),$$

and at $(\alpha_1, \beta_1) = \left\{\left(V_{02}-V_{11}\right)/2V_{02}, 0\right\}$, we have

$$(3.14) \qquad MSE_{\min}\left(\hat{S}^2_{YK}\right) \cong Var(\hat{S}^2_{Reg}).$$

($viii$) Recently Singh and Malik [35] suggested an improved estimator for S_y^2, given by

$$(3.15) \qquad \hat{S}^2_{SM} = s_y^2\left\{k_9+k_{10}(S_x^2-s_x^2)\right\}\exp\left(\psi_1\frac{(aS_x^2+b)-(as_x^2+b)}{(aS_x^2+b)+(as_x^2+b)}\right),$$

where k_9 and k_{10} are suitably chosen constants. Here ψ_1 is the scalar quantity which takes the values $+1$ and -1 for ratio and product type estimators respectively.
The bias of $\hat{S}^2_{SM}$, to first order of approximation, is given by

$$\begin{aligned}Bias\left(\hat{S}^2_{SM}\right) \cong\ & S_y^2\left(k_9-1\right)+\frac{1}{4}S_y^2k_9\psi_1\tau^2V_{02}+\frac{1}{8}S_y^2k_9\gamma_1^2\tau^2V_{02}\\ &+\frac{1}{2}S_y^2S_x^2k_{10}\psi_1\tau V_{02}-\frac{1}{2}S_y^2S_x^2k_9\psi_1\tau V_{11}-S_y^2S_x^2k_{10}V_{11}.\end{aligned}$$

The minimum MSE of $\hat{S}^2_{SM}$, to first order of approximation, at optimum values

$$k_9^{(opt)} = \frac{1}{4}\left(\frac{-12\psi_1\tau V_{02}V_{11}+3\psi_1^2\tau^2V_{02}^2+16V_{11}^2-8V_{02}-2\psi_1\tau^2V_{02}^2}{\psi_1^2\tau^2V_{02}^2-4\psi_1\tau V_{02}V_{11}+8V_{11}^2-2V_{02}V_{20}-2V_{02}-\psi_1\tau^2V_{02}^2}\right)$$

and $k_{10}^{(opt)} = -\frac{1}{4S_x^2}\left(\frac{-6\psi_1^2\tau^2V_{02}V_{11}+\psi_1^2\tau^3V_{02}^2+8\psi_1\tau V_{11}^2-4\psi_1\tau V_{02}+8V_{11}-8V_{02}V_{11}+4\psi_1\tau V_{02}V_{20}}{\psi_1^2\tau^2V_{02}^2-4\psi_1\tau V_{02}V_{11}+8V_{11}^2-2V_{02}V_{20}-2V_{02}-\psi_1\tau^2V_{02}^2}\right)$, is given by $MSE_{\min}\left(\hat{S}^2_{SM}\right) \cong S_y^4\left\{\frac{}{32\left(\psi_1^2\tau^2V_{02}^2-4\psi_1\tau V_{02}V_{11}+8V_{11}^2-2V_{02}V_{20}-2V_{02}-\psi_1\tau^2V_{02}^2\right)}\right\}$,
or, at $\tau = \psi_1 = 1$, we have

(3.16)

$$MSE_{\min}\left(\hat{S}^2_{SM}\right) \cong \frac{S_y^4}{64}\left[\frac{V_{02}\left\{V_{02}(V_{02}+8V_{11})+16(V_{02}-4)Var(\hat{S}^2_{Reg})+16V_{11}(V_{11}-V_{02})\right\}}{-V_{02}(1+V_{02}+2V_{11})+4V_{11}^2}\right].$$

4. Proposed estimator

Bahl and Tuteja [4] exponential type estimators for population variance $\left(S_y^2\right)$, are given by

$$(4.1) \qquad \hat{S}^2_R = s_y^2\exp\left(\frac{S_x^2-s_x^2}{S_x^2+s_x^2}\right),$$

and

$$(4.2) \qquad \hat{S}^2_{Pr} = s_y^2\exp\left(\frac{s_x^2-S_x^2}{S_x^2+s_x^2}\right).$$

Following Haq and Shabbir [15, 16], the average of the ratio and the product-type exponential estimators given in (4.1) and (4.2) is

$$\hat{S}_A^2 = s_y^2 \frac{1}{2}\left[\exp\left\{\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)\right\} + \exp\left\{\left(\frac{s_x^2 - S_x^2}{S_x^2 + s_x^2}\right)\right\}\right].$$

A generalized form of $\hat{S}_A^2$, is given by

$$\hat{S}_A^2 = s_y^2 \frac{1}{2}\left[\exp\left\{\frac{a\left(S_x^2 - s_x^2\right)}{a\left(s_x^2 + S_x^2\right) + 2b}\right\} + \exp\left\{\frac{a\left(s_x^2 - S_x^2\right)}{a\left(s_x^2 + S_x^2\right) + 2b}\right\}\right], \tag{4.3}$$

where $(a \neq 0)$ and b are functions of known parameters of the auxiliary variable. Motivated by Singh and Solanki [46] and Haq and Shabbir [15, 16], replacing s_y^2 given in (4.3) by $\left(\hat{S}_{SS}^2\right)$ given in (3.8), we propose the following class of estimators for estimating the finite population variance S_y^2, given by

$$\hat{S}_P^2 = \left\{k_{11}s_y^2 + k_{12}(S_x^2 - s_x^2)\right\}\left(\frac{aS_x^2 + b}{as_x^2 + b}\right)\hat{S}_{A_1}^2, \tag{4.4}$$

where k_{11} and k_{12} are suitably chosen constants, and
$\hat{S}_{A_1}^2 = \frac{1}{2}\left[\exp\left\{\frac{a\left(S_x^2 - s_x^2\right)}{a\left(s_x^2 + S_x^2\right) + 2b}\right\} + \exp\left\{\frac{a\left(s_x^2 - S_x^2\right)}{a\left(s_x^2 + S_x^2\right) + 2b}\right\}\right]$.
Expressing (4.4) in term of $\delta' s$ and keeping terms up to power two, we have

$$\hat{S}_P^2 = \left\{k_{11}S_y^2(1+\delta_0) - k_{12}S_x^2\delta_1\right\}\left(1 - \tau\delta_1 + \tau^2\delta_1^2\right)\left(1 + \frac{\tau^2\delta_1^2}{8}\right). \tag{4.5}$$

Solving (4.5), up to first order of approximation, we get

$$\begin{aligned}\hat{S}_P^2 - S_y^2 \cong{}& k_{11}S_y^2 - k_{11}S_y^2\tau\delta_1 + k_{11}S_y^2\delta_0 - k_{12}S_x^2\delta_1 + \frac{9}{8}k_{11}S_y^2\tau^2\delta_1^2 \\ &- k_{11}S_y^2\tau\delta_0\delta_1 + k_{12}S_x^2\tau\delta_1^2 - S_y^2.\end{aligned} \tag{4.6}$$

Using (4.6), the bias and MSE of $\hat{S}_P^2$, to first order of approximation are, respectively given by

$$Bias(\hat{S}_P^2) \cong (k_{11} - 1)\,S_y^2 + \frac{9}{8}k_{11}S_y^2\tau^2V_{02} - k_{11}S_y^2\tau V_{11} + k_{12}S_x^2\tau V_{02}, \tag{4.7}$$

and

$$\begin{aligned}MSE(\hat{S}_P^2) \cong{}& k_{11}^2S_y^4 - 2k_{11}S_y^4 + k_{11}^2S_y^4V_{20} + 2k_{11}S_y^4\tau V_{11} - 2k_{12}S_y^2S_x^2\tau V_{02} \\ &+ k_{12}^2S_x^4V_{02} + 4k_{11}k_{12}S_y^2S_x^2\tau V_{02} - 2k_{11}k_{12}S_y^2S_x^2V_{11} \\ &- 4k_{11}^2S_y^4\tau V_{11} - \frac{9}{4}k_{11}S_y^4\tau^2V_{02} + \frac{13}{4}k_{11}^2S_y^4\tau^2V_{02} + S_y^4.\end{aligned} \tag{4.8}$$

Differentiating (4.8), with respect to k_{11} and k_{12}, we get the optimum values of k_{11} and k_{12} as

$$k_{11}^{(opt)} = \frac{V_{02}}{2}\left(\frac{1 + 7A}{V_{02}^2\tau^2 + 4V_{02}A + 4V_{02}V_{20} - 4V_{11}^2}\right),$$

and

$$k_{12}^{(opt)} = \frac{S_y^2}{2S_x^2}\left(\frac{V_{11} + 7V_{11}A - 8V_{02}\tau A + 8V_{02}V_{20} - 8V_{11}^2}{V_{02}^2\tau^2 + 4V_{02}A + 4V_{02}V_{20} - 4V_{11}^2}\right),$$

where A, is defined earlier.
Substituting the optimum values of k_{11} and k_{12} in (4.8), we get $MSE_{\min}(\hat{S}_P^2)$ as

$$MSE_{\min}(\hat{S}_P^2) \cong \frac{S_y^4}{16}\left\{\frac{64AS_y^{-4}Var(\hat{S}_{Reg}^2) - V_{02}^2\tau^4}{V_{02}\tau^2 + 4A + 4S_y^{-4}Var(\hat{S}_{Reg}^2)}\right\}. \qquad (4.9)$$

4.1. Some members of the proposed class of estimators. Different estimators can be generated from the proposed estimator given in (4.4) by substituting the suitable choices of a, and b. Some generated estimators are listed in Table 1.

Table 1. Some members of proposed class of estimators $\hat{S}_{Pj}^2$ $(j = 1, 2, ..., 6)$

a	b	Estimator
1	0	$\hat{S}_{P1}^2$
N	$-S_x^2$	$\hat{S}_{P2}^2$
N	$-\bar{X}^2$	$\hat{S}_{P3}^2$
n	$-S_x^2$	$\hat{S}_{P4}^2$
n^2	$-\bar{X}^2$	$\hat{S}_{P5}^2$
n^2	$-S_x^2$	$\hat{S}_{P6}^2$

5. Efficiency comparisons

In this section, we compare the propose estimator with existing estimators.
Condition i: By (3.1) and (4.9),

$$Var\left(\hat{S}_y^2\right) - MSE_{\min}\left(\hat{S}_P^2\right) > 0, \qquad \text{if}$$

$$\frac{S_y^4}{16}\left\{\frac{V_{02}\tau^2\left(16V_{20} + V_{02}\tau^2\right) + 64S_y^{-4}Var(\hat{S}_{Reg}^2) + \frac{64AV_{11}^2}{V_{02}}}{4A + V_{02}\tau^2 + 4S_y^{-4}Var(\hat{S}_{Reg}^2)}\right\} > 0.$$

Condition ii: By [(3.3) or (3.14)] and (4.9),

$$\left[Var\left(\hat{S}_{Reg}^2\right) \quad \text{or} \quad MSE_{\min}\left(\hat{S}_{YK}^2\right)\right] - MSE_{\min}\left(\hat{S}_P^2\right) > 0, \qquad \text{if}$$

$$\frac{S_y^4}{16}\left\{\frac{\begin{array}{c}16V_{02}^2\tau^2S_y^{-4}Var(\hat{S}_{yReg}^2) + V_{02}^3\tau^4 \\ +64V_{02}\left(S_y^{-4}Var(\hat{S}_{yReg}^2)\right)^2\end{array}}{1 + 3A + 4S_y^{-4}Var(\hat{S}_{yReg}^2)}\right\} > 0.$$

Condition iii: By (3.5) and (4.9),

$$MSE_{\min}\left(\hat{S}_d^2\right) - MSE_{\min}\left(\hat{S}_P^2\right) > 0, \qquad \text{if}$$

$$\frac{S_y^4\tau^2}{16}\left[\frac{S_y^{-4}Var(\hat{S}_{yReg}^2)\left\{V_{02}\left(16 + V_{02}\tau^2\right) + 64V_{02}S_y^{-4}Var(\hat{S}_{yReg}^2)\right\} + V_{02}^2\tau^2}{\left(1 + S_y^{-4}Var(\hat{S}_{yReg}^2)\right)\left(4 + 4S_y^{-4}Var(\hat{S}_{yReg}^2) + V_{02}\tau\right)}\right] > 0.$$

Condition iv: By (3.7) and (4.9),

$$MSE_{\min}\left(\hat{S}_{SG}^2\right) - MSE_{\min}\left(\hat{S}_P^2\right) > 0, \quad \text{if}$$

$$\frac{S_y^4 V_{02}}{64}\left[\frac{\begin{array}{r} 3V_{02}\tau^2\left\{V_{02}+16S_y^{-4}Var(\hat{S}^2_{Reg})\right\} \\ +4V_{02}\left\{(1+S_y^{-4}Var(\hat{S}^2_{Reg}))(\tau^2+1)(\tau^2-1)\right\} \\ -64S_y^{-4}Var(\hat{S}^2_{Reg})\left\{(1-\tau^2)+(1-4\tau^2)S_y^{-4}Var(\hat{S}^2_{Reg})\right\}\end{array}}{(1+3A+4S_y^{-4}Var(\hat{S}^2_{Reg}))(1+4S_y^{-4}Var(\hat{S}^2_{Reg}))}\right]>0.$$

Condition v: By (3.9) and (4.9),

$$MSE_{\min}\left(\hat{S}^2_{SS}\right)-MSE_{\min}\left(\hat{S}^2_{P}\right)>0, \quad \text{if}$$

$$S_y^4\tau^2V_{02}^2\left\{\frac{\left(AV_{02}\tau^2+(1+15A)\,S_y^{-4}Var(\hat{S}^2_{Reg})\right)}{\left(A+S_y^{-4}Var(\hat{S}^2_{Reg})\right)\left(1+3A+4S_y^{-4}Var(\hat{S}^2_{Reg})\right)}\right\}>0.$$

Condition vi: By (3.11) and (4.9), when using $\lambda=1$.

$$MSE_{\min}\left(\hat{S}^2_{YG}\right)-MSE_{\min}\left(\hat{S}^2_{P}\right)>0, \quad \text{if}$$

$$S_y^4\tau^2V_{02}^2\left\{\frac{\left(AV_{02}\tau^2+(1+15A)\,S_y^{-4}Var(\hat{S}^2_{Reg})\right)}{\left(A+S_y^{-4}Var(\hat{S}^2_{Reg})\right)\left(1+3A+4S_y^{-4}Var(\hat{S}^2_{Reg})\right)}\right\}>0.$$

Condition vii: By (3.16) and (4.9), when using $\tau=\psi_1=1$.

$$MSE_{\min}\left(\hat{S}^2_{SM}\right)-MSE_{\min}\left(\hat{S}^2_{P}\right)>0, \quad \text{if}$$

$$\frac{S_y^4}{32}\left[\frac{\left\{\begin{array}{r} 512V_{11}^2\,(V_{20}-V_{11})+48V_{02}V_{20}\,(1+4V_{02}V_{20})-60V_{02}V_{11}^2\left(V_{02}+V_{11}^2\right) \\ -960V_{02}V_{11}^2S_y^{-4}Var(\hat{S}^2_{Reg})+256V_{20}V_{11}\left\{(V_{02}V_{11})+(V_{02}-V_{11})\,S_y^{-4}Var(\hat{S}^2_{Reg})\right\} \\ 3\left(V_{02}^4+96V_{02}^3V_{11}^2-156\dfrac{V_{11}^4}{V_{02}}\right)\end{array}\right\}}{2\left\{S_y^{-4}Var(\hat{S}^2_{Reg})+(1+4V_{11})-3\frac{V_{11}^2}{V_{02}}\right\}\left\{4S_y^{-4}Var(\hat{S}^2_{Reg})+(4-3V_{02})\right\}}\right]>0.$$

Note that the proposed estimator $(\hat{S}^2_P)$ is more efficient than the existing estimators $\hat{S}^2_i\;(i=y,Reg,d,SG,SS,YG,YK,SM)$, when above conditions are satisfied.

6. Numerical study

In this section, we consider the following data sets for numerical comparisons.

Population I: [Source: Kadilar and Cingi [20]]

Let y = Level of apple production (1 unit=100 tones) and x = Number of apple trees (1 unit=100 trees).

$N=104$, $n=20$, $\gamma=0.05$, $\bar{Y}=6.24064$, $\bar{X}=13929.899$, $S_y=11.670$, $S_x=23026.133$, $\rho_{yx}=0.865$,

$\rho_{(S_y^2,S_x^2)}=0.83675$, $\beta_{2(y)}=16.523$, $\beta_{2(x)}=17.516$, $\theta_{22}=14.398$, $V_{20}=0.77615$, $V_{02}=0.8258$, $V_{11}=0.6699$.

Population II: [Source: Cochran [6], p.34]
Let $y=$ The weekly expenditure on food and $x=$ The weekly family income.
$N=33$, $n=10$, $\gamma=0.1$, $\bar{Y}=27.491$, $\bar{X}=72.547$, $S_y=10.131$, $S_x=10.577$, $\rho_{yx}=0.432738$, $\rho_{(S_y^2,S_x^2)}=-0.474028$, $\beta_{2(y)}=5.38276$, $\beta_{2(x)}=2.015035$, $\theta_{22}=0.000187$, $V_{20}=0.43827$, $V_{02}=0.101503$,
$V_{11}=-0.099981$.

Population III: [Source: Murthy [26], p.399]
Let $y=$ Area of wheat in 1964 and $x=$ Area of wheat in 1963.
$N=80$, $n=10$, $\gamma=0.1$, $\bar{Y}=5182.638$, $\bar{X}=285.125$, $S_y=1835.638$, $S_x=270.429$, $\rho_{yx}=0.988421$,
$\rho_{(S_y^2,S_x^2)}=0.73198$, $\beta_{2(y)}=2.2665$, $\beta_{2(x)}=3.5808$, $\theta_{22}=2.32338$, $V_{20}=0.12665$, $V_{02}=0.25808$, $V_{11}=0.13234$.

We use the following expression for Percentage Relative Efficiency (PRE) and the Absolute Bias (AB).

$$PRE(\hat{S}_y^2,\hat{S}_i^2)=\frac{Var(\hat{S}_y^2)}{MSE_{\min}\left(\hat{S}_i^2\right)\quad \text{or}\quad MSE\left(\hat{S}_i^2\right)},$$

and

$$AB=|Bias\left(\hat{S}_i^2\right)|,\quad \text{for } i{=}Reg,d,SG,SS,YG,YK,SM,P.$$

MSE, PRE and AB values based on Populations I, II and III are given in Tables 2–4.

Table 2. *MSE*, *PRE* and *AB*, values of different estimators for Population I.

a, b		$\hat{S}^2_y$	$\hat{S}^2_{Reg}$	$\hat{S}^2_d$	$\hat{S}^2_{SG}$	$\hat{S}^2_{SS}$	$\hat{S}^2_{YG}$	$\hat{S}^2_{YK}$	$\hat{S}^2_{SM}$	$\hat{S}^2_P$
	MSE	14395.58	4316.32	3501.46	2618.27	1847.80	1847.80	4316.32	3996.51	**903.65**
1, 0	*PRE*	100	333.51	411.13	549.81	779.07	779.07	333.51	360.20	**1593.04**
	AB	—	—	25.71	19.22	13.57	13.57	27.83	29.34	**6.63**
	MSE	14395.58	4316.32	3501.46	2618.27	1939.24	1939.24	4316.32	4017.78	**1006.65**
104, 0.530E+9	*PRE*	100	333.51	411.13	549.81	742.33	742.33	333.51	358.30	**1430.05**
	AB	—	—	25.71	19.22	12.82	12.82	27.83	29.18	**5.82**
	MSE	14395.58	4316.32	3501.46	2618.27	1882.38	1882.38	4316.32	4004.41	**942.17**
104, 0.194E+9	*PRE*	100	333.51	411.13	549.81	764.75	764.75	333.51	359.49	**1527.91**
	AB	—	—	25.71	19.22	13.30	13.30	27.83	29.29	**6.34**
	MSE	14395.58	4316.32	3501.46	2618.27	2239.62	2239.62	4316.32	4097.01	**1372.23**
20, 0.530E+9	*PRE*	100	333.51	411.13	549.81	642.76	642.76	333.51	351.36	**1049.06**
	AB	—	—	25.71	19.22	8.47	8.47	27.83	28.41	**1.67**
	MSE	14395.58	4316.32	3501.46	2618.27	1856.92	1856.92	4316.32	3998.58	**913.76**
400, 0.194E+9	*PRE*	100	333.51	411.13	549.81	775.24	775.24	333.51	360.02	**1575.41**
	AB	—	—	25.71	19.22	13.50	13.50	27.83	29.33	**6.56**
	MSE	14395.58	4316.32	3501.46	2618.27	1872.50	1872.50	4316.32	4002.13	**931.12**
400, 0.530E+9	*PRE*	100	333.51	411.13	549.81	768.78	768.78	333.51	359.69	**1546.05**
	AB	—	—	25.71	19.22	13.38	13.38	27.83	29.30	**6.42**

Bold numbers indicate least values.

Table 3. *MSE*, *PRE* and *AB*, values of different estimators for Population II.

a, b		$\hat{S}_y^2$	$\hat{S}_{Reg}^2$	$\hat{S}_d^2$	$\hat{S}_{SG}^2$	$\hat{S}_{SS}^2$	$\hat{S}_{YG}^2$	$\hat{S}_{YK}^2$	$\hat{S}_{SM}^2$	$\hat{S}_P^2$
1, 0	*MSE*	4616.57	3579.21	2671.46	2602.41	2597.05	2597.05	3579.21	3060.015	**2543.56**
	PRE	100	128.98	172.81	177.396	177.761	177.761	128.98	150.86	**181.50**
	AB	—	—	26.03	25.35	25.30	25.30	**10.03**	29.81	24.78
33, 111.88	*MSE*	4616.57	3579.21	2671.46	2602.41	2601.70	2601.70	3579.21	3059.885	**2551.47**
	PRE	100	128.98	172.81	177.396	177.44	178.11	128.98	150.87	**180.94**
	AB	—	—	26.03	25.35	25.25	25.30	**10.03**	29.815	24.70
33, 5263.08	*MSE*	4616.57	3579.21	2671.46	2602.41	2659.62	2659.62	3579.21	3035.34	**2650.94**
	PRE	100	128.98	172.81	177.396	173.58	173.58	128.98	152.10	**174.15**
	AB	—	—	26.03	25.35	19.67	19.67	**10.02**	22.40	16.11
20, 111.88	*MSE*	4616.57	3579.21	2671.46	2602.41	2610.83	2610.83	3579.21	3059.22	**2567.04**
	PRE	100	128.98	172.81	177.396	176.82	176.82	128.98	150.91	**179.83**
	AB	—	—	26.03	25.35	25.11	25.11	**10.02**	29.80	24.46
400, 5263.08	*MSE*	4616.57	3579.21	2671.46	2602.41	2638.50	2638.50	3579.21	3052.20	**2614.48**
	PRE	100	128.98	172.81	177.396	174.97	174.97	128.98	151.25	**176.5**
	AB	—	—	26.03	25.35	22.75	22.75	**10.02**	29.19	20.63
400, 111.88	*MSE*	4616.57	3579.21	2671.46	2602.41	2598.64	2598.64	3579.21	3059.98	**2546.26**
	PRE	100	128.98	172.81	177.40	177.65	177.65	128.98	150.87	**181.31**
	AB	—	—	26.03	25.35	25.28	25.28	**10.02**	29.81	24.75

Table 4. *MSE*, *PRE* and *AB*, values of different estimators for Population III.

a, b		$\hat{S}_y^2$	$\hat{S}_{Reg}^2$	$\hat{S}_d^2$	$\hat{S}_{SG}^2$	$\hat{S}_{SS}^2$	$\hat{S}_{YG}^2$	$\hat{S}_{YK}^2$	$\hat{S}_{SM}^2$	$\hat{S}_P^2$
	MSE	0.1438E+13	0.667E+12	0.630E+12	0.578E+12	0.618E+12	0.618E+12	0.667E+12	0.586E+12	**0.558E+12**
1, 0	*PRE*	100	215.42	228.09	248.52	232.49	232.49	215.42	245.15	**257.37**
	AB	—	—	187104.52	171720.74	183559.03	183559.03	429103.301	174078.43	**165818.47**
	MSE	0.1438E+13	0.667E+12	0.630E+12	0.578E+12	0.619E+12	0.619E+12	0.667E+12	0.589E+12	**0.561E+12**
80, 73132.06	*PRE*	100	215.42	228.09	248.52	232.32	232.32	215.42	244.05	**256.24**
	AB	—	—	187104.52	171720.74	183438.60	183438.60	429103.301	173252.14	**165044.56**
	MSE	0.1438E+13	0.667E+12	0.630E+12	0.578E+12	0.619E+12	0.619E+12	0.667E+12	0.589E+12	**0.561E+12**
80, 81297.82	*PRE*	100	215.42	228.09	248.52	232.33	232.33	215.42	243.93	**256.12**
	AB	—	—	187104.52	171720.74	183424.74	183424.74	429103.301	173157.49	**164955.42**
	MSE	0.1438E+13	0.667E+12	0.630E+12	0.578E+12	0.621E+12	0.621E+12	0.667E+12	0.505E+12	**0.575E+12**
10, 73132.06	*PRE*	100	215.42	228.09	248.52	231.52	231.52	215.42	237.73	**249.96**
	AB	—	—	187104.52	171720.74	182369.38	182369.38	429103.301	166223.26	**158131.45**
	MSE	0.1438E+13	0.667E+12	0.630E+12	0.578E+12	0.619E+12	0.619E+12	0.667E+12	0.589E+12	**0.561E+12**
100, 81297.82	*PRE*	100	215.42	228.09	248.52	232.37	232.37	215.42	244.17	**256.36**
	AB	—	—	187104.52	171720.74	183452.25	183452.25	429103.301	173345.47	**165132.37**
	MSE	0.1438E+13	0.667E+12	0.630E+12	0.578E+12	0.619E+12	0.619E+12	0.667E+12	0.589E+12	**0.561E+12**
100, 73132.06	*PRE*	100	215.42	228.09	248.52	232.38	232.38	215.42	244.27	**256.45**
	AB	—	—	187104.52	171720.74	183463.21	183463.21	429103.301	173420.45	**165202.84**

7. Conclusion

In this paper, we have suggested an improved class of estimators for estimating the finite population variance using information on the auxiliary variable. Expressions for bias and *MSE* of the proposed class of estimators have been derived to first order of approximation. The proposed class of estimators $\hat{S}_P^2$ is compared with existing estimators both theoretically and numerically. From Tables 2–4, it is observed that the *MSE* of $\hat{S}_P^2$ is smaller as compared to *MSE* of existing estimators for all different choices of a and b considered here. Also bias of $\hat{S}_P^2$ is smaller as compared to all other considered estimators in all three populations except $\hat{S}_{YK}^2$ in population II. Among three populations, the maximum *PRE* gained by proposed estimator is in Population I. So it is preferable to use the estimator $\hat{S}_P^2$.

References

[1] Arcos, A., Rueda, M., Martinez, M. D., Gonzalez S., Roman, Y. *Incorporating the auxiliary information available in variance estimation*, Applied Mathematics and Computation **160(2)**, 387–399, 2005.

[2] Ahmed, M. S., Rahman, M. S., Hossain, M. I. *Some competitive estimators of finite population variance using multivariate auxiliary information*, International Journal of Information and Management Sciences **11(1)**, 49–54, 2000.

[3] AL-Jaraha, J., Ahmed, M. S. *A class of chain estimators of a finite population variance using double-sampling*, International Journal of Information and Management Sciences **13(2)**, 13–18, 2002.

[4] Bahl, S., Tuteja, R. K. *Ratio and product type exponential estimators*, Journal of Information and Optimization Sciences **12**, 159–164, 1991.

[5] Cebrian, A. A., Garcia, M. R. *Variance estimation using auxiliary information– an almost unbiased multivariate ratio estimator*, Metrika **45** , 171–178, 1997.

[6] Cochran, W. G. *Sampling Techniques*: Third Edition. John Wiley & Sons 1977.

[7] Chandra, P., Singh H. P. *A family of estimators for population variance using knowledge of kurtosis of an auxiliary variable in sample survey*, Statistics in Transition **7**, 27–34, 2005.

[8] Chaudhury, A. *On estimating the variance of a finite population*, Metrika **25**, 66–67, 1978.

[9] Das, A. K., Tripathi, T. P. *Use of auxiliary information in estimating the finite population variance*, Sankhya C. **40**, 139–148, 1978.

[10] Garcia, M. R., Cebrian, A. A. *Repeated substitution method-the ratio estimator for population variance*, Metrika **43**, 101–105, 1996.

[11] Grover, L. K. *A correction note on improvement in variance estimation using auxiliary information*, Communication in Statistics –Theory and Methods **39**, 753–764, 2010.

[12] Gupta, S., Shabbir, J. *On the use of transformed auxiliary variables in estimating population mean*, Journal of Statistical Planning and Inference. **137(5)**, 1606–1611, 2007.

[13] Gupta, S., Shabbir, J. *On improvement in estimating the population mean in simple random sampling* Journal of Applied Statistics **35(5)**, 559–566, 2008.

[14] Gupta S., Shabbir J. *Variance estimation in simple random sampling using uxiliary information*, Hacettepe Journal of Mathematics and Statistics **37(1)**, 57 – 67, 2008.

[15] Haq, A., Shabbir, J.*Improved family of ratio estimators in simple and stratified random sampling*, Communication in Statistics – Theory and Methods **42**, 782–799, 2013.

[16] Haq, A., Shabbir, J., Gupta, S. *Improved exponential ratio type estimators in stratified sampling*, Pakistan Journal of Statistics **29(1)**, 13–31, 2013.

[17] Isaki, C. T. *Variance estimation using auxiliary information*, Journal of the American Statistical Association **78**, 117–123, 1983.

[18] Kadilar, C., Cingi, H. *Ratio estimator in stratified sampling*, Biometrical Journal **45(2)**, 218–225, 2003.

[19] Kadilar, C., Cingi, H. *Ratio estimators in simple random sampling*, Applied Mathematics and Computations **151**, 893–902, 2004.

[20] Kadilar, C., Cingi, H. *A new estimator using two auxiliary variables*, Applied Mathematics and Computations **162**, 901–908, (2005).
[21] Kadilar, C., Cingi, H. *Ratio estimators for the population variance in simple and stratified random sampling*, Applied Mathematics and Computations **173(2)**, 1047–1059, 2006.
[22] Kadilar, C., Cingi, H. *Improvement in variance estimation using auxiliary information*, Hacettepe Journal of Mathematics and Statistics **35(1)**, 111–115, 2006.
[23] Kadilar, C., Cingi, H. *Improvement in variance estimation in simple random sampling*, Communication in Statistics –Theory and Methods **36(11)**, 2075–2081, 2007.
[24] Koyuncu, N., Kadilar, C. *Ratio and product estimators in stratified random sampling*, Journal of Statistical Planning and Inference **139(8)**, 2552–2558, 2009.
[25] Koyuncu, N., Kadilar, C. *On improvement in estimating population mean in stratified random sampling*, Journal of Applied Statistics **37(6)**, 999–1013, 2010.
[26] Murthy, M. N. *Sampling Theory and Methods*, Statistical Publishing Society. Calcutta, India, 1967.
[27] Mukhopadhyay, P. *Estimating a finite population variance under a super population model*, Metrika **25**, 115–122, 1978.
[28] Mukhopadhyay, P. *Optimum strategies for estimating the variance of a finite population under a super population model*, Metrika **29**, 143–158, 1982.
[29] Parsad, B., Singh, H. P. *Some improved ratio-type estimators of finite population variance using auxiliary information in sample surveys*, Communication in Statistics -Theory and Methods **19(3)**, 1127–1139, 1990.
[30] Parsad, B., Singh, H. P.*Unbiased estimators of finite population variance using auxiliary information in sample survey* , Communication in Statistics -Theory and Methods **21(5)**, 1367–1376, 1992.
[31] Searl, D. T, Intarapanich, P. *A note on an estimator for the variance that utilizes the kurtosis*, American Statistician **44(4)**, 295–296, 1990.
[32] Shabbir, J., Gupta, S. *On improvement in variance estimation using auxiliary information*, Communication in Statistics -Theory and Methods **36**, 2177–2185, 2007.
[33] Shabbir, J., Gupta, S. *Some estimators of finite population variance of stratified sample mean*, Communication in Statistics -Theory and Methods **39**, 3001–3008, 2010.
[34] Shabbir, J., Gupta, S. *On estimating the finite population mean in simple and stratified random sampling*, Communication in Statistics - Theory and Methods **140(10)**, 2870–2874, 2011.
[35] Singh, R., Malik, S. *Improved estimation of population variance using information on auxiliary attribute in simple random sampling*, Applied Mathematics and Computation **235**, 43–49, 2014.
[36] Singh et al. *On the utilization of known coefficient of kurtosis in estimation procedure of variance*, Annals of the Institute of Statistical Mathematics **25**, 51–55, 1973.
[37] Singh, H. P. *A note on the estimation of variance of sample mean using the knowledge of coefficient of variation in normal population*, Communication in Statistics - Theory and Methods **15(12)**, 3737–3746, 1986.
[38] Singh, H. P., Upadhyaya, L. N., Namjoshi U. D. *Estimation of finite population variance*, Current Science **57**, 1331–1334, 1988.
[39] Singh, H. P., Biradar, R. S. *Estimation of finite population variance using auxiliary information*, Journal of the Indian Society of Statistics and Operation Research $\mathbf{15(1-4)}$, 47–63, 1994.
[40] Singh, S., Joarder, A. H . *Estimation of finite population variance using random nonresponse in survey sampling*, Metrika **47**, 241–249, 1998.
[41] Singh, H. P., Singh, R. *Improved ratio type estimators for variance using auxiliary information.*, Journal of the Indian Social and Agriculture Statistics **54(3)**, 276–287, 2001.
[42] Singh, H. P., Tailor, R. *Use of known correlation coefficient in estimating the finite population mean*, Statistics in Transition **6(4)**, 555–560, 2003.
[43] Singh, H. P., Vishvakarama, G. K. *Some families of estimators of variance of stratified random sample mean using auxiliary information*, Journal of Statistical Theory and Practice **2(1)**, 21–43, 2008.

[44] Singh, H. P., Vishvakarama, G. K. *A general procedure for estimating the population mean in stratified random sampling using auxiliary information*, Metron **68(1)**, 47–65, 2010.

[45] Singh, H. P., Solanki, R. S. *A new procedure for variance estimation in simple random sampling*, Statistical Papers **5**, 479–497, 2013 a.

[46] Singh, H. P., Solanki, R. S. *Improved estimation of finite population variance using auxiliary information*, Communication in Statistics – Theory and Methods **42**, 2718–2730, 2013 b.

[47] Srivastava, S. K., Jhajj, H. S. *A class of estimator using auxiliary information for estimating finite population variance*, Sankhya(C) **42**, 87–96, 1980.

[48] Turgut, Y., Cingi, H. *New generalized estimators for the population variance using auxiliary information* , Hacettepe Journal of Mathematics and Statistics **37(2)**, 177–184, 2008.

[49] Upadhyaya, L. N., Singh, H. P. *An estimator of population variance that utilises the kurtosis of an auxiliary variable in sample survey*, Vikram Math Journal **19**, 14–17, 1999.

[50] Upadhyaya, L. N., Singh, H. P., Singh, S. *A class of estimators for estimating the variance of the ratio estimator*, Journal of the Japan Statistical Society **34(1)**, 47–63, 2004.

[51] Wu, C. F. *Estimation of variance of the ratio estimator.*, Biometrika **69(1)**, 183–189, 1982.

[52] Yadav, C. F., Upadhyaya, L. N., Singh, H. P., Chatterjee, S. *A generalized family of transformed ratio-product estimators for variance in sample surveys*, Communication in Statistics – Theory and Methods **42(10)**, 2013.

[53] Yadav, S. K., Kadilar, C., Shabbir, J., Gupta, S. *Improved family of estimators of population variance in Simple Random Sampling*, Journal of Statistical Theory and Practice **9(2)**, 219–226, 2015.

[54] Yadav, S. K., Kadilar, C. *A two parameter variance estimator using auxiliary information*, Applied Mathematics and Computation **226**, 117–122, 2014.

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.

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.