HACETTEPE UNIVERSITY
FACULTY OF SCIENCE
TURKEY

# HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

A Bimonthly Publication
Volume 45 Issue 1
2016

ISSN 1303 5010

# HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS

Volume 45 Issue 1
February 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 1 (2016)**

ISSN 1303 – 5010

| KÜNYE |
|---|
| YAYININ ADI:<br>HACETTEPE JOURNAL OF MATHEMATICS AND STATISTICS |
| YIL : 2016 SAYI : 45 - 1 AY : Şubat |
| YAYIN SAHİBİNİN ADI : H. Ü. Fen Fakültesi Dekanlığı adına<br>Prof. Dr. Bekir Salih |
| SORUMLU YAZI İŞL. MD. ADI : Prof. Dr. Yücel Tıraş |
| YAYIN İDARE MERKEZİ ADRESİ : H. Ü. Fen Fakültesi Dekanlığı |
| YAYIN İDARE MERKEZİ TEL. : 0 312 297 68 50 |
| YAYININ TÜRÜ : Yaygın |
| BASIMCININ ADI : Hacettepe Üniversitesi Hastaneleri Basımevi. |
| BASIMCININ ADRESİ : 06100 Sıhhıye, ANKARA. |
| BASIMCININ TEL. : 0 312 305 1020 |
| BASIM TARİHİ - YERİ : - ANKARA |

Hacettepe Journal of Mathematics and Statistics

**A Bimonthly Publication – Volume 45 Issue 1 (2016)**

ISSN 1303 – 5010



Published by Hacettepe University
Faculty of Science

## CONTENTS

### Mathematics

## Statistics

# MATHEMATICS

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 1 – 14

# Some generalizations of the class of analytic functions with respect to k-symmetric points

Syed Zakar Hussain Bukhari*, Mohsan Raza [†‡]and Maryam Nazir [§]

**Abstract**

In this article, we introduce a new class $M_s^k[\alpha,\beta,\lambda]$ which generalizes the various classes of analytic functions with respect to k-symmetric points. Naturally, the class $M_s^k[\alpha,\beta,\lambda]$ combines the classes $S_s^k(\alpha,\beta)$ and $C_s^k(\alpha,\beta)$. We also study the coefficient estimates and obtain some inequalities related to the coefficients of functions in these classes. We develop the integral representation, inclusions and convolution conditions for the functions in the class $M_s^k[\alpha,\beta,\lambda]$.

**Keywords:** Convex functions, Starlike functions, Symmetric points, subordination.

*2000 AMS Classification:* 30C45, 30C50.

*Received* : 23.01.2014 *Accepted* : 24.11.2014 *Doi* : 10.15672/HJMS.20164512489

## 1. Introduction and preliminaries

Let $A$ be the class of functions $f$ of the form

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n, \tag{1.1}$$

which are analytic in the open unit disc $E = \{z : |z| < 1\}$. A function $f$ is said to be subordinate to a function $g$ written as $f \prec g$ , if there exists a schawarz function $w$ with $w(0) = 0$ and $|w(z)| < 1$ such that $f(z) = g(w(z))$. In particular if $g$ is univalent in $E$, then $f(0) = g(0)$ and $f(E) \subset g(E)$.

The classes of starlike and convex univalent functions are defined as

*Department of Mathematics, Mirpur University of Science and Technology, Mirpur AJ&K. Email : fatmi@must.edu.pk

[†]Department of Mathematics, Government College University Faisalabad, Pakistan. Email : mohsan976@yahoo.com

[‡]Corresponding Author.

[§]Department of Mathematics, Mirpur University of Science and Technology, Mirpur AJ&K. Email : maryamnazir@must.edu.pk

$$S^* = \left\{ f : f \in A \text{ and } Re\, \frac{zf'(z)}{f(z)} > 0,\ z \in E \right\},$$

$$C = \left\{ f : f \in A \, and \, Re \frac{(zf'(z))'}{f'(z)} > 0, \quad z \in E \right\}.$$

A function which is analytic in the open unit disc $E$ is said to be starlike with respect to the symmetric point if it satisfies

$$\text{Re}\, \frac{(zf'(z))}{f(z) - f(-z)} > 0,\ z \in E.$$

The class $S_s^*$ of starlike functions with respect to symmetric points was introduced and studied by Sakaguchi, see [5]. The class $S_s^k[\alpha, \beta]$ consists of functions $f \in S$ satisfying the inequality

$$\left| \frac{2zf'(z)}{f_k(z)} - 1 \right| < \beta \left| \frac{2\alpha zf'(z)}{f_k(z)} + 1 \right| \ for\ 0 \leqslant \alpha \leqslant 1,\ 0 \leqslant \beta \leqslant 1,\ z \in E,$$

where $f_k$ is defined as

$$f_k(z) = \frac{1}{k} \sum_{m=0}^{k-1} \varkappa^{-m} f(\varkappa^m z) \tag{1.2}$$

with $\varkappa^k = 1$ and $k \geqslant 1$ a fixed positive integer. Similarly the class}$C_s^k[\alpha, \beta]$ is defined by:

$$C_s^k[\alpha, \beta] = \left\{ f \in S \text{ and } \left| \frac{(\mathrm{z}\mathrm{f}'(\mathrm{z}))'}{\mathrm{f}'_\mathrm{k}(\mathrm{z})} - 1 \right| < \beta \left| \frac{\alpha\,(\mathrm{z}\mathrm{f}'(\mathrm{z}))'}{\mathrm{f}'_\mathrm{k}(\mathrm{z})} + 1 \right| \right\},$$

where $0 \leqslant \alpha \leqslant 1,\ 0 < \beta \leqslant 1$ and $f_k$ is given in (1.2) . The classes $S_s^k[\alpha, \beta]$ and $C_s^k[\alpha, \beta]$ were defined by Wang [8] and Gao and Zhou [2] respectively. These classes are further studied in [7, 9]. Keeping in view the above mentioned classes, we define the following subclass of analytic function with respect to k-symmetric point.

**1.1. Definition.** A function $f \in S$ is in the class $M_s^k[\alpha, \beta, \lambda]$ if it satisfies the following condition:

(1.3)

$$\left| (1-\lambda)\frac{zf'(z)}{f_k(z)} + \lambda \frac{(zf'(z))'}{f_k'(z)} - 1 \right| < \beta \left| \alpha \left[ (1-\lambda)\frac{zf'(z)}{f_k(z)} + \lambda \frac{(zf'(z))'}{f_k'(z)} \right] + 1 \right|,$$

where $f_k$ is defined by (1.2), $0 \leqslant \alpha \leqslant 1,\ 0 < \beta \leqslant 1,\ k \geqslant 1$ a fixed positive integer, $\lambda \in \mathbb{R}$ and $z \in E$.

where $f_k$ is defined by ( 1.2), $0 \leqslant \alpha \leqslant 1,\ 0 < \beta \leqslant 1,\ k \geqslant 1$ a fixed positive integer, $\lambda \in$ and $z \in E$.

**Special Cases**

$(i)$. For $k = 2$, the class $M_s^k[\alpha, \beta, \lambda]$ reduces to the class $M[\alpha, \beta, \lambda]$ consisting of univalent functions which satisfy the condition:

$$\left|(1-\lambda)\frac{zf'(z)}{f(z)}+\lambda\frac{(zf'(z))'}{f'(z)}-1\right|<\beta\left|\alpha\left[(1-\lambda)\frac{zf'(z)}{f(z)}+\lambda\frac{(zf'(z))'}{f'(z)}\right]+1\right|,$$

where $0\leqslant\alpha\leqslant1,\ 0<\beta\leqslant1,\ k\geqslant1,\ \lambda\in\mathbb{R}$ and $z\in E$.
$(ii)$. For $\lambda=1$, the class $M_s^k[\alpha,\beta,\lambda]$ yields the class $C_s^k(\alpha,\beta)$ introduced and studied by Wang [8].
$(iii)$. When $\lambda=0$, the class $M_s^k[\alpha,\beta,\lambda]$ produces the class $S_s^k(\alpha,\beta)$ studied by Gao and Zhou [2].
$(iv)$. For $k=2,\ \lambda=1$, we obtain the class $C_s(\alpha,\beta)$.
$(v)$. Taking $k=2,\ \lambda=0,\ M_s^k[\alpha,\beta,\lambda]$ reduces to the class $S_s^*(\alpha,\beta)$, see [6].
$(vi)$. For $k=2,\ \lambda=0,\ \alpha=\beta=1,\ M_s^k[\alpha,\beta,\lambda]$ reduces to the class $S_s^*$ [5].

In the following, we have some useful lemmas.

**1.2. Lemma.** [1] *Suppose that the function $\varphi$ is convex univalent in $E$ with $\varphi(0)=1$ and*

$$\mathrm{Re}\left(\beta\varphi(z)+\gamma\right)>0 \text{ for } \beta,\gamma\in\mathbb{C},\ \mathrm{z}\in\mathrm{E}.$$

*If $p$ is analytic in $E$ with $p(0)=1$, then*

$$p(z)+\frac{zp'(z)}{\beta p(z)+\gamma}\prec\varphi(z)\,implies\,p(z)\prec\varphi(z),\ z\in E.$$

**1.3. Lemma.** [4] *Let $\beta,\gamma\in\mathbb{C}$ and $\varphi$ be a convex, univalent function with*

$$\mathrm{Re}\left(\beta\varphi(z)+\gamma\right)>0,\ z\in E.$$

*Also let $h\in A: h(z)\prec\varphi(z)$. If $p\in P$ and*

$$p(z)+\frac{zp'(z)}{\beta h(z)+\gamma}\prec\varphi, then\,p(z)\prec\varphi(z).$$

**1.4. Lemma.** [6] *Let $G$ be analytic in $E$ and let*

$$\left|\frac{1-G(z)}{1+\eta G(z)}\right|<\delta \tag{1.4}$$

*$z\in E,\ 0\leqslant\eta\leqslant1,\ 0<\delta\leqslant1$ with $G(0)=1$. Then*

$$G(z)=\frac{1-z\varphi(z)}{1+\eta z\varphi(z)}, \tag{1.5}$$

*where $\varphi$ is analytic in $E$ and $|\varphi(z)|\leqslant\delta$ for $z\in E$. Conversely any function $G$ given by (1.5) is analytic in $E$ and satisfies (1.4).*

**1.5. Lemma.** [3] *Let $F$ be analytic and convex in $E$. If $f,\ g\in A$ and $f,\ g\prec F$. Then*

$$\sigma f+(1-\sigma)g\prec F,\quad 0\leqslant\sigma\leqslant1.$$

## 2. Main Results

**2.1. Theorem.** *Let* $F_k(z) = (1-\lambda)\frac{zf'(z)}{f_k(z)} + \lambda\frac{(zf'(z))'}{f_k'(z)}, \lambda \in \mathbb{R},\ k \geqslant 1,\ f$ *and* $f_k$ *defined by (1.1) and 1.2 respectively. Then* $F_k(z) = 1+\sum_{j=2}^{\infty} c_j z^{j-1} + ... \in P$ *for* $c_j = [(1-\lambda(1-j))j + ((1-\lambda)j + \lambda - (1+j))\, d_j]\, a_j$ *, where* $c_j \leqslant 2$.

*Proof.* Here we let

$$F_k(z) = (1-\lambda)\frac{zf'(z)}{f_k(z)} + \lambda\frac{(zf'(z))'}{f_k'(z)}. \tag{2.1}$$

It can easily follows from (1.2) that

$$\begin{aligned} f_k(z) &= \frac{1}{k}\sum_{m=0}^{k-1} \varkappa^{-m} f(\varkappa^m z) = \frac{1}{k}\sum_{m=0}^{k-1} \varkappa^{-m}\left[\varkappa^m z + \sum_{j=2}^{\infty} a_j \left(\varkappa^m z\right)^j\right] \\ &= z + \sum_{l=2}^{\infty} a_{(l-1)k+1} z^{(l-1)k+1} = z + \sum_{j=2}^{\infty} a_j d_j z^j, \end{aligned} \tag{2.2}$$

where $d_j = \frac{1}{k}\sum_{m=0}^{k-1} \varkappa^{(j-1)m} = \left\{\begin{array}{ll} 1, & j = (l-1)k+1 \\ 0, & j \neq (l-1)k+1 \end{array}\right\}, \varkappa^k = 1$. On combining ( 2.1) and ( 2.2) , we have

$$\begin{aligned} F_k(z) &= \frac{(1-\lambda)zf'(z)f_k'(z) + \lambda(f'(z) + zf''(z)))f_k(z)}{f_k(z)f_k'(z)} \\ &= \frac{(1-\lambda)(z + \sum_{j=2}^{\infty} ja_j z^j)(1 + \sum_{j=2}^{\infty} ja_j d_j z^{j-1}) + \lambda(1 + \sum_{j=2}^{\infty} j^2 a_j z^{j-1})(z + \sum_{j=2}^{\infty} a_j d_j z^j)}{(z + \sum_{j=2}^{\infty} a_j d_j z^j)(1 + \sum_{j=2}^{\infty} ja_j d_j z^{j-1})} \\ &= \frac{1 + \sum_{j=2}^{\infty}\left[\left((1-\lambda+j\lambda)\, ja_j + (1-\lambda)ja_j d_j + \lambda a_j d_j\right) z^{j-1} + j^2 a_j^2 d_j z^{2j-2}\right]}{1 + \sum_{j=2}^{\infty}\left[(1+j)\, a_j d_j z^{j-1} + ja_j^2 d_j z^{2j-2}\right]} \\ &= \left[1 + \sum_{j=2}^{\infty}\left[\left((1-\lambda+j\lambda)\, ja_j + (1-\lambda)ja_j d_j + \lambda a_j d_j\right) z^{j-1} + j^2 a_j^2 d_j z^{2j-2}\right]\right] \times \\ &\quad \left[1 - \sum_{j=2}^{\infty}\left[(1+j)\, a_j d_j z^{j-1} + ja_j^2 d_j z^{2j-2}\right] + ...\right] \\ &= 1 + \sum_{j=2}^{\infty}\left[(1-\lambda+j\lambda)\, j + ((1-\lambda)j + \lambda - (1+j))\, d_j\right] a_j z^{j-1} + ... \\ &= 1 + \sum_{j=2}^{\infty} c_j z^{j-1} + .... \end{aligned}$$

Thus $F_k(z) = 1 + \sum\limits_{j=2}^{\infty} c_j z^{j-1} + ... \in P$,

where $c_j = \left[(1 - \lambda + j\lambda)\, j + ((1-\lambda)j + \lambda - (1+j))\, d_j\right] a_j$ such that $|c_j| \leqslant 2$. ∎

**2.2. Theorem.** *Let $0 \leqslant \alpha \leqslant 1$, $0 < \beta \leqslant 1$, $k \geqslant 1$, $\lambda \in \mathbb{R}$, and $z \in E$. Then the function $f \in M_s^k[\alpha, \beta, \lambda]$ if and only if*

$$(2.3) \quad (1-\lambda)\frac{zf'(z)}{f_k(z)} + \lambda \frac{(zf'(z))'}{f_k'(z)} \prec \frac{1+\beta z}{1-\alpha\beta z},$$

*where $f_k$ is given in (1.2).*

*Proof.* Let $f \in M_s^k[\alpha, \beta, \lambda]$. Then from (1.3), we have

$$\left|(1-\lambda)\frac{zf'(z)}{f_k(z)} + \lambda\frac{(zf'(z))'}{f_k'(z)} - 1\right| < \beta\left|\alpha\left[(1-\lambda)\frac{zf'(z)}{f_k(z)} + \lambda\frac{(zf'(z))'}{f_k'(z)}\right] + 1\right|.$$

Taking $F_k$ as defined in ( 2.1) we write

$$|F_k(z) - 1|^2 < \beta^2 \left|\alpha F_k(z) + 1\right|^2$$

or

$$\left(1 - \alpha^2\beta^2\right)|F_k(z)|^2 - 2(1+\alpha\beta^2)\operatorname{Re} F_k(z) < \beta^2 - 1.$$

If $\alpha \neq 1$ or $\beta \neq 1$, then we have

$$|F_k(z)|^2 - 2\left(\frac{1+\alpha\beta^2}{1-\alpha^2\beta^2}\right)\operatorname{Re} F_k(z) + \left(\frac{1+\alpha\beta^2}{1-\alpha^2\beta^2}\right)^2 < \frac{\beta^2-1}{1-\alpha^2\beta^2} + \left(\frac{1+\alpha\beta^2}{1-\alpha^2\beta^2}\right)^2$$

or

$$\left|F_k(z) - \frac{1+\alpha\beta^2}{1-\alpha^2\beta^2}\right|^2 < \frac{\beta^2(1+\alpha)^2}{\left(1-\alpha^2\beta^2\right)^2}.$$

This represents the disk with center at $\frac{1+\alpha\beta^2}{1-\alpha^2\beta^2}$ and radius $\frac{\beta(1+\alpha)}{\left(1-\alpha^2\beta^2\right)}$. Also the function $\omega(z) \prec \varphi(z) = \frac{1+\beta z}{1-\alpha\beta z}$ maps the unit disk onto the disk

$$\left|\omega - \frac{1+\alpha\beta^2}{1-\alpha^2\beta^2}\right| < \frac{\beta(1+\alpha)}{\left(1-\alpha^2\beta^2\right)}.$$

From here we notice that, $F_k(E) \subset \varphi(E)$, $F_k(0) = \varphi(0)$ and $\varphi$ is univalent in $E$. Therefore, we write

$$F_k(z) \prec \varphi(z) = \frac{1+\beta z}{1-\alpha\beta z}.$$

Conversely, let $F_k(z) \prec \frac{1+\beta z}{1-\alpha\beta z}$. Then using subordination, we have

$$(2.4) \quad F_k(z) = \frac{1+\beta\omega(z)}{1-\alpha\beta\omega(z)},$$

where $\omega \in \Omega$. From ( 2.4), we write

$$|F_k(z)-1| = \left|\frac{1+\beta\omega(z)}{1-\alpha\beta\omega(z)}-1\right| = \left|\frac{1+\beta\omega(z)-1+\alpha\beta\omega(z)}{1-\alpha\beta\omega(z)}\right|$$

$$|F_k(z)-1| = \beta\left|\frac{(1+\alpha)\,\omega(z)}{1-\alpha\beta\omega(z)}\right|. \tag{2.5}$$

Also

$$|\alpha F_k(z)+1| = \left|\frac{\alpha+\alpha\beta\omega(z)}{1-\alpha\beta\omega(z)}+1\right| < \beta\left|\frac{(1+\alpha)}{1-\alpha\beta\omega(z)}\right|. \tag{2.6}$$

Using (2.5) in (2.6) , we have

$$|F_k(z)-1| < \beta\,|\alpha F_k(z)+1|\,,\ where\ |\omega(z)|<1\ for\ all\ z\in E.$$

This implies that

$$\left|(1-\lambda)\frac{zf'(z)}{f_k(z)}+\lambda\frac{(zf'(z))'}{f_k'(z)}-1\right| < \beta\left|\alpha\left[(1-\lambda)\frac{zf'(z)}{f_k(z)}+\lambda\frac{(zf'(z))'}{f_k'(z)}\right]+1\right|.$$

Hence, $f \in M_s^k[\alpha,\beta,\lambda]$. ∎

**2.3. Theorem.** *Let $f \in M_s^k[\alpha,\beta,\lambda]$. Then $f_k \in M_s[\alpha,\beta,\lambda]$ and also $f_k \in S_s^k(\alpha,\beta)$.*

*Proof.* Let $f \in M_s^k[\alpha,\beta,\lambda]$. Then by Theorem 2.2, we have

$$(1-\lambda)\frac{zf'(z)}{f_k(z)}+\lambda\frac{(zf'(z))'}{f_k'(z)} \prec \frac{1+\beta z}{1-\alpha\beta z},$$

where $0\leqslant\alpha\leqslant 1,\ 0<\beta\leqslant 1,\ k\geqslant 1$ (is fixed positive integer), $\lambda\in\mathbb{R}$ and $f_k$ is defined by (1.2). Now using subordination, we write

$$(1-\lambda)\frac{zf'(z)}{f_k(z)}+\lambda\frac{(zf'(z))'}{f_k'(z)} = \frac{1+\beta\omega(z)}{1-\alpha\beta\omega(z)},\ \omega(0)=0 \text{ and } |\omega(\mathrm{z})|<1. \tag{2.7}$$

On replacing $z$ by $\varkappa^m z$, where $m=0,1,2,...,k-1$ and $\varkappa^k=1$ in (2.7) , we have

$$(1-\lambda)\frac{\varkappa^m zf'(\varkappa^m z)}{f_k(\varkappa^m z)}+\lambda\frac{f'(\varkappa^m z)+\varkappa^m zf''(\varkappa^m z)}{f_k'(\varkappa^m z)} = \frac{1+\beta\omega(\varkappa^m z)}{1-\alpha\beta\omega(\varkappa^m z)}. \tag{2.8}$$

From (1.2), we write $f_k(\varkappa^m z)=\varkappa^m f_k(z)$ and $f_k'(\varkappa^m z)=f_k'(z)$. These results along with (2.8) yield

$$(1-\lambda)\frac{\varkappa^m zf'(\varkappa^m z)}{\varkappa^m f_k(z)}+\lambda\frac{f'(\varkappa^m z)+\varkappa^m zf''(\varkappa^m z)}{f_k'(z)} = \frac{1+\beta\omega(\varkappa^m z)}{1-\alpha\beta\omega(\varkappa^m z)}.$$

Taking summation from $m=0$ to $k-1$, we write

$$(1-\lambda)\frac{1}{k}\sum_{m=0}^{k-1}\frac{\varkappa^m z f'(\varkappa^m z)}{\varkappa^m f_k(z)}+\frac{\lambda}{f_k'(z)}(\frac{1}{k}\sum_{m=0}^{k-1}f'(\varkappa^m z)+\frac{1}{k}\sum_{m=0}^{k-1}\varkappa^m z f''(\varkappa^m z))$$

$$= \quad \frac{1}{k}\sum_{m=0}^{k-1}\left(\frac{1+\beta\omega\left(\varkappa^m z\right)}{1-\alpha\beta\omega\left(\varkappa^m z\right)}\right)$$

or

$$(1-\lambda)\frac{zf_k'(z)}{f_k(z)}+\lambda\frac{(zf_k'(z))'}{f_k'(z)}=\frac{1}{k}\sum_{m=0}^{k-1}\left(\frac{1+\beta\omega\left(\varkappa^m z\right)}{1-\alpha\beta\omega\left(\varkappa^m z\right)}\right)$$

or

$$(1-\lambda)\frac{zf_k'(z)}{f_k(z)}+\lambda\frac{(zf_k'(z))'}{f_k'(z)}=\frac{1}{k}\sum_{m=0}^{k-1}\left(\frac{1+\beta\omega\left(\varkappa^m z\right)}{1-\alpha\beta\omega\left(\varkappa^m z\right)}\right)\in P[\alpha,\beta],$$

where $P[\alpha,\beta]$ is a convex set and containing function $p(z)\prec\frac{1+\beta z}{1-\alpha\beta z}$. This implies that

$$(1-\lambda)\frac{zf_k'(z)}{f_k(z)}+\lambda\frac{(zf_k'(z))'}{f_k'(z)}\prec\frac{1+\beta z}{1-\alpha\beta z},$$

which implies that $f_k\in M_s[\alpha,\beta,\lambda]$. Now, let $h(z)=\frac{zf_k'(z)}{f_k(z)}$. After some manipulation, we have

$$(1-\lambda)\frac{zf_k'(z)}{f_k(z)}+\lambda\frac{(zf_k'(z))'}{f_k'(z)}=h\left(z\right)+\lambda\frac{zh'\left(z\right)}{h\left(z\right)}. \tag{2.9}$$

This implies that

$$h\left(z\right)+\lambda\frac{zh'\left(z\right)}{h\left(z\right)}\prec\frac{1+\beta z}{1-\alpha\beta z}.$$

Using Lemma 1.2, we obtain

$$h\left(z\right)\prec\frac{1+\beta z}{1-\alpha\beta z}. \tag{2.10}$$

Hence, $f_k\in S_s^k\left(\alpha,\beta\right)$. ∎

**2.4. Theorem.** *Let* $0\leqslant\alpha\leqslant 1$, $0<\beta\leqslant 1$, $k\geqslant 2$ *(fixed positive integer),* $\lambda>0$. *Then* $M_s^k\left(\alpha,\beta,\lambda\right)\subset S_s^k\left(\alpha,\beta\right)\subset S$.

*Proof.* For $f\in M_s^k\left[\alpha,\beta,\lambda\right]$, we have

$$(1-\lambda)\frac{zf'(z)}{f_k(z)}+\lambda\frac{(zf'(z))'}{f_k'(z)}\prec\frac{1+\beta z}{1-\alpha\beta z}.$$

Now, we let

$$p\left(z\right)=\frac{zf'(z)}{f_k(z)}\ \ \mathit{and}\, h(z)=\frac{zf_k'(z)}{f_k(z)},$$

where, $h$ and $p$ satisfy the conditions described in Lemma 1.3. Therefore

$$(2.11)\quad (1-\lambda)\frac{zf'(z)}{f_k(z)}+\lambda\frac{(zf'(z))'}{f_k'(z)}=p(z)+\lambda\frac{zp'(z)}{h(z)}\prec\frac{1+\beta z}{1-\alpha\beta z}.$$

By using ( 2.10) ,(2.11) and Lemma 1.3, we obtain

$$p(z)=\frac{zf'(z)}{f_k(z)}\prec\frac{1+\beta z}{1-\alpha\beta z},$$

which implies that , $f\in S_s^{(k)}(\alpha,\beta)\subset S$ or $M_s^k(\alpha,\beta,\lambda)\subset S_s^{(k)}(\alpha,\beta)\subset S$. ∎

**2.5. Theorem.** *Let* $0\leqslant\alpha\leqslant 1$ *and* $0<\beta\leqslant 1$, $0\leqslant\lambda_1<\lambda_2$. *Then* $M_s^k[\alpha,\beta,\lambda_2]\subset M_s^k[\alpha,\beta,\lambda_1]$.

*Proof.* Suppose that $f\in M_s^k[\alpha,\beta,\lambda_2]$. Then by Theorem 2.2, we have

$$h_1(z)=(1-\lambda_2)\frac{zf'(z)}{f_k(z)}+\lambda_2\frac{(zf'(z))'}{f_k'(z)}\prec\frac{1+\beta z}{1-\alpha\beta z}.$$

Also from Theorem 2.4, we write

$$h_2(z)=\frac{zf'(z)}{f_k(z)}\prec\frac{1+\beta z}{1-\alpha\beta z}.$$

Now

$$\begin{aligned}(1-\lambda_1)\frac{zf'(z)}{f_k(z)}+\lambda_1\frac{(zf'(z))'}{f_k'(z)} &= \left(1-\frac{\lambda_1}{\lambda_2}\right)\frac{zf'(z)}{f_k(z)}+\frac{\lambda_1}{\lambda_2}\left\{(1-\lambda_2)\frac{zf'(z)}{f_k(z)}+\lambda_2\frac{(zf'(z))'}{f_k'(z)}\right\}\\ &= \left(1-\frac{\lambda_1}{\lambda_2}\right)h_2(z)+\frac{\lambda_1}{\lambda_2}h_1(z)\end{aligned}$$

Since $\frac{1+\beta z}{1-\alpha\beta z}$ is convex set, therefore by using Lemma 1.5 we get the required result. ∎

**2.6. Theorem.** *Let* $f\in M_s^k[\alpha,\beta,\lambda]$. *Then*

$$f_k(z)=\left[\frac{1}{\lambda}\int_0^z\frac{1}{u}\left\{u.\exp\sum_{m=0}^{k-1}\int_0^u\frac{(1+\alpha)\beta\omega(t)}{t(k-\alpha\beta\omega(t))}dt\right\}^{\frac{1}{\lambda}}du\right]^{\lambda},$$

*where* $\omega\in\Omega$. *For* $\lambda=0$,

$$f(z)=\int_0^z\frac{f_k(\lambda)}{\lambda}\frac{1+\beta\omega(\lambda)}{1-\alpha\beta\omega(\lambda)}d\lambda.$$

*Proof.* Let $f\in M_s^k[\alpha,\beta,\lambda]$. Then from Theorem 2.2, we have

$$(1-\lambda)\frac{zf'(z)}{f_k(z)}+\lambda\frac{(zf'(z))'}{f_k'(z)}\prec\frac{1+\beta z}{1-\alpha\beta z}.$$

Using subordination, we obtain

$$(2.12)\quad (1-\lambda)\frac{zf'(z)}{f_k(z)}+\lambda\frac{(zf'(z))'}{f_k'(z)}=\frac{1+\beta\omega\left(z\right)}{1-\alpha\beta\omega\left(z\right)},$$

where $\omega$ is analytic in $E$ with $\omega(0)=0$ and $|\omega(z)|<1$. Now replacing $z$ by $\varkappa^m z$, where $m=0,1,2...,k-1$, $\varkappa^k=1$, using (1.2) with $f_k(\varkappa^m z)=\varkappa^m f_k(z)$ and $f_k'(\varkappa^m z)=f_k'(z)$ and then taking summation for $m=0,1,2...,k-1$ in (2.12), we obtain

$$\frac{1-\lambda}{k}\left\{\sum_{m=0}^{k-1}\frac{\varkappa^m zf'(\varkappa^m z)}{\varkappa^m f_k(z)}\right\}+\frac{\lambda}{k}\frac{1}{f_k'(z)}\left\{\sum_{m=0}^{k-1}f'(\varkappa^m z)+\sum_{m=0}^{k-1}\varkappa^m zf''(\varkappa^m z)\right\}$$

$$=\quad\frac{1+\frac{\beta}{k}\sum\limits_{m=0}^{k-1}\omega\left(\varkappa^m z\right)}{1-\frac{\alpha\beta}{k}\sum\limits_{m=0}^{k-1}\omega\left(\varkappa^m z\right)}.$$

This implies that

$$(1-\lambda)\frac{zf_k'(z)}{f_k(z)}+\lambda\frac{(zf_k'(z))'}{f_k'(z)}=\left(k+\beta\sum_{m=0}^{k-1}\omega\left(\varkappa^m z\right)\right)/\left(k-\alpha\beta\sum_{m=0}^{k-1}\omega\left(\varkappa^m z\right)\right)$$

or

$$(1-\lambda)\frac{zf_k'(z)}{f_k(z)}+\lambda\frac{(zf_k'(z))'}{f_k'(z)}-\frac{1}{z}=\left((1+\alpha)\,\beta\sum_{m=0}^{k-1}\omega\left(\varkappa^m z\right)\right)/z\left(k-\alpha\beta\sum_{m=0}^{k-1}\omega\left(\varkappa^m z\right)\right).$$

Integrating from 0 to $z$, we write

$$\log\left\{\frac{(f_k(z))^{(1-\lambda)}(zf_k'(z))^{\lambda}}{z}\right\}=\int\limits_0^z\frac{(1+\alpha)\,\beta\sum\limits_{m=0}^{k-1}\omega\left(\varkappa^m\zeta\right)}{\zeta(k-\alpha\beta\sum\limits_{m=0}^{k-1}\omega\left(\varkappa^m\zeta\right))}d\zeta$$

or

$$(2.13)\quad\left[\frac{zf_k'(z)}{f_k(z)}\right]^{\lambda}f_k(z)=z\exp\sum_{m=0}^{k-1}\int\limits_0^z\frac{(1+\alpha)\,\beta\omega\left(\varkappa^m\zeta\right)}{\zeta\left(k-\alpha\beta\omega\left(\varkappa^m\zeta\right)\right)}d\zeta.$$

Let

$$(2.14)\quad F_k(z)=\left[\frac{zf_k'(z)}{f_k(z)}\right]^{\lambda}f_k(z),\quad F_k(0)=0,\ F_k'(0)=1.$$

Since $f_k$ is $\lambda$-convex and if $\lambda$ is not an integer, then we can select a suitable branch, so that $F_k$ is analytic in $E$. Logarithmic differentiation of (2.14) yields

$$\frac{zF_k'(z)}{F_k(z)}=(1-\lambda)\frac{zf_k'(z)}{f_k(z)}+\lambda\frac{(zf_k'(z))'}{f_k'(z)}.$$

Hence, $F_k$ is starlike in $E$. Now we solve (2.14) for $f_k$ by assuming that $\lambda \neq 0$. (The case when $\lambda = 0$ gives $F_k(z) = f_k(z)$). Formal manipulations leads to the solution

$$f_k(z) = \left[ \frac{1}{\lambda} \int_0^z \frac{[F_k(\zeta)]^{\frac{1}{\lambda}}}{\zeta} d\zeta \right]^{\lambda}. \tag{2.15}$$

By using (2.13) and (2.15), we have

$$\left[ \frac{z f_k'(z)}{f_k(z)} \right]^{\lambda} f_k(z) = z. \exp \sum_{m=0}^{k-1} \int_0^z \frac{(1+\alpha)\,\beta\omega\,(\varkappa^m \zeta)}{\zeta\,(k - \alpha\beta\omega\,(\varkappa^m \zeta))} d\zeta,$$

which implies that

$$f_k(z) = \left[ \frac{1}{\lambda} \int_0^z \frac{1}{u} \left[ u. \exp \left\{ \sum_{m=0}^{k-1} \int_0^u \frac{(1+\alpha)\,\beta\omega\,(t)}{t\,(k - \alpha\beta\omega\,(t))} dt \right\} \right]^{\frac{1}{\lambda}} du \right]^{\lambda}.$$

This is the required integral representation for $f_k$ when $f \in M_s^k[\alpha, \beta, \lambda]$. It can be easily verified that for $\lambda = 0$,

$$f(z) = \int_0^z \frac{f_k(\lambda)}{\lambda} \frac{1 + \beta\omega\,(\lambda)}{1 - \alpha\beta\omega\,(\lambda)} d\lambda.$$

∎

**2.7. Theorem.** *Let $f \in M_s^k[\alpha, \beta, \lambda]$. Then we have*

$$f(z) = \int_0^z \frac{1+c}{\lambda\,[f_k(\lambda)]^c} \int_0^{\varkappa^m \lambda} [f_k(t)]^c\, f_k'(t) \frac{1 + \beta\omega\,(t)}{1 - \alpha\beta\omega\,(t)} dt d\lambda,$$

*where $f_k$ is given in* (1.2), *$c = \frac{1}{\lambda} - 1 : \lambda \neq 0$ and $\omega$ is analytic in $E$ with $\omega(0) = 0$ and $|\omega(z)| < 1$. If $\lambda = 0$, then we have*

$$f(z) = \int_0^z \frac{f_k(\lambda)}{\lambda} \frac{1 + \beta\omega\,(\lambda)}{1 - \alpha\beta\omega\,(\lambda)} d\lambda.$$

*Proof.* Let $f \in M_s^k[\alpha, \beta, \lambda]$. Then

$$(1-\lambda) \frac{z f'(z)}{f_k(z)} + \lambda \frac{(z f'(z))'}{f_k'(z)} = \frac{1 + \beta\omega\,(z)}{1 - \alpha\beta\omega\,(z)}, \tag{2.16}$$

where $\omega \in \Omega$. Multiplying both sides of (2.16) by $\lambda^{-1}\,[f_k(z)]^c\, f_k'(z)$, where $c = \frac{1}{\lambda} - 1 : \lambda \neq 0$, we get

(2.17)

$$c z f'(z)\,[f_k(z)]^{c-1}\, f_k'(z) + [f_k(z)]^c\,(z f'(z))' = (1+c)\,[f_k(z)]^c\, f_k'(z) \frac{1 + \beta\omega\,(z)}{1 - \alpha\beta\omega\,(z)}.$$

The left hand side of (2.17) is the exact differential of $zf'(z)\left[f_k(z)\right]^c$. On integrating both sides of (2.17), we obtain

$$f'(z) = \frac{1+c}{z\left[f_k(z)\right]^c}\int_0^z \left[f_k(\zeta)\right]^c f_k'(\zeta)\frac{1+\beta\omega(\zeta)}{1-\alpha\beta\omega(\zeta)}d\zeta$$

or

$$f'(z) = \frac{1+c}{z\left[f_k(z)\right]^c}\int_0^{\varkappa^m z} \left[f_k(t)\right]^c f_k'(t)\frac{1+\beta\omega(t)}{1-\alpha\beta\omega(t)}dt.$$

This implies that

$$f(z) = \int_0^z \frac{1+c}{\lambda\left[f_k(\lambda)\right]^c}\int\limits_0^{\varkappa^m\lambda} \left[f_k(t)\right]^c f_k'(t)\frac{1+\beta\omega(t)}{1-\alpha\beta\omega(t)}dtd\lambda.$$

If $\lambda = 0$, then we have

$$f'(z) = \frac{f_k(z)}{z}\frac{1+\beta\omega(z)}{1-\alpha\beta\omega(z)},$$

which implies that

$$f(z) = \int_0^z \frac{f_k(\lambda)}{\lambda}\frac{1+\beta\omega(\lambda)}{1-\alpha\beta\omega(\lambda)}d\lambda.$$

▮

**2.8. Theorem.** *Let $f \in M_s^k[\alpha,\beta,\lambda]$. Then*

$$\frac{1}{z}\left\{f(z) * \left(\frac{z}{(1-z)^2} - \frac{1+\beta e^{j\theta}}{(1-\alpha\beta e^{j\theta})}h(z)\right)\right\} \neq 0,$$

*where* $0 \leqslant \theta \leqslant 2\pi,\ 0 \leqslant \alpha \leqslant 1,\ 0 < \beta \leqslant 1,\ \lambda > 0$ *and* $z \in E$.

*Proof.* Let $f \in M_s^k(\alpha,\beta,\lambda)$. Then by using Theorem 2.3, we have

$$f \in S_s^{(k)}(\alpha,\beta),$$

which implies that for $0 \leqslant \theta \leqslant 2\pi$, we write

$$\frac{zf'(z)}{f_k(z)} \prec \frac{1+\beta z}{1-\alpha\beta z}$$

or

$$\frac{zf'(z)}{f_k(z)} \neq \frac{1+\beta e^{j\theta}}{1-\alpha\beta e^{j\theta}}.$$

Therefore

$$\frac{1}{z}\left\{zf'(z) - \left(\frac{1+\beta e^{j\theta}}{1-\alpha\beta e^{j\theta}}\right)f_k(z)\right\} \neq 0. \tag{2.18}$$

For $zf'(z) = f(z) * \frac{z}{(1-z)^2}$ and $f_k(z) = z + \sum_{j=0}^{\infty} a_j c_j z^j = (f * h)(z)$, the inequality (2.18) yields

$$\frac{1}{z}\left\{f(z) * \frac{z}{(1-z)^2} - \left(\frac{1+\beta e^{j\theta}}{1-\alpha\beta e^{j\theta}}\right)(f * h)(z)\right\} \neq 0$$

or

$$\frac{1}{z}\left\{f(z) * \frac{z}{(1-z)^2} - \left(\frac{1+\beta e^{j\theta}}{1-\alpha\beta e^{j\theta}}\right)(f * h)(z)\right\} \neq 0$$

or

$$\frac{1}{z}\left\{f(z) * \left(\frac{z}{(1-z)^2} - \frac{1+\beta e^{j\theta}}{(1-\alpha\beta e^{j\theta})} h(z)\right)\right\} \neq 0.$$

∎

**2.9. Theorem.** *Let* $f(z) = z + \sum_{j=2}^{\infty} a_j z^j$. *If for* $0 \leqslant \alpha \leqslant 1,\ 0 < \beta \leqslant 1,\ k \geqslant 1,$ *and* $\lambda \geqslant 0$, *we have*

(2.19)

$$\sum_{\substack{j=2 \\ j\neq lk+1}}^{\infty} \chi(\alpha,\beta,\lambda,j)\, j\,|a_j| + \sum_{j=1}^{\infty} \chi_1(\alpha,\beta,\lambda,j,k)\,|a_{jk+1}| + \sum_{j=1}^{\infty} \chi_2(\alpha,\beta,\lambda,j,k)\,|a_{jk+1}|^2 < \beta(1+\alpha) - 2$$

*where* $\chi_1(\alpha,\beta,\lambda,j,k = \{(1-\lambda)(jk+1)+\lambda\}(1-\alpha\beta) + (1-\beta)(2+jk)$,
$\chi(\alpha,\beta,\lambda,j) = ((1-\lambda)+\lambda \mathrm{j})(1-\alpha\beta)\, m$ *and*
$\chi_2(\alpha,\beta,\lambda,j,k) = \{(1-\alpha\beta)(jk+1)+(1-\beta)\}(\mathrm{jk}+1)$. *Then* $f \in M_s^k[\alpha,\beta,\lambda]$.

*Proof.* Let $f(z) = z + \sum_{j=2}^{\infty} a_j z^j$ and $f_k$ be given in (1.2) for $z \in E$. Then, we have

$$\begin{aligned}(2.20)\ M &= \left|(1-\lambda)zf'(z)f_k'(z) + \lambda(zf'(z))'f_k(z) - f_k(z)f_k'(z)\right| - \\ &\quad \beta\left|\alpha((1-\lambda)zf'(z)f_k'(z) + \lambda(zf'(z))'f_k(z)) + f_k(z)f_k'(z)\right|.\end{aligned}$$

Also for $\varkappa^k = 1$ and $k > 1$,

(2.21)

$$f_k(z) = \frac{1}{k}\sum_{m=0}^{k-1} \varkappa^{-m} f(\varkappa^m z) = z + \sum_{j=2}^{\infty} a_j d_j z^j,\ \textit{where}\ d_j = \left\{\begin{array}{ll} 1, & j = (l-1)k+1 \\ 0, & j \neq (l-1)k+1 \end{array}\right\}$$

From ( 2.20) , ( 2.21) and for $|z| = r < 1$, we have

$$\begin{aligned} M &\leqslant (1-\lambda)(1-\alpha\beta)\left|zf'(z)f_k'(z)\right| + \lambda(1-\alpha\beta)\left|(zf'(z))'f_k(z)\right| + (1-\beta)\left|f_k(z)f_k'(z)\right| \\ &< (1-\lambda)(1-\alpha\beta) + \lambda(1-\alpha\beta) + (1-\beta) \\ &\quad +(1-\lambda)(1-\alpha\beta)\sum_{j=2}^{\infty} j\left|a_j\right| + \lambda(1-\alpha\beta)\sum_{j=2}^{\infty} j^2\left|a_j\right| \\ &\quad +(1-\lambda)(1-\alpha\beta)\sum_{j=2}^{\infty} j\left|a_j d_j\right| + \lambda(1-\alpha\beta)\sum_{j=2}^{\infty}\left|a_j d_j\right| + (1-\beta)\sum_{j=2}^{\infty}(1+j)\left|a_j d_j\right| \\ &\quad +(1-\lambda)(1-\alpha\beta)\sum_{j=2}^{\infty} j^2\left|a_j\right|\left|a_j d_j\right| + \lambda(1-\alpha\beta)\sum_{j=2}^{\infty} j^2\left|a_j\right|\left|a_j d_j\right| + (1-\beta)\sum_{j=2}^{\infty} j\left|a_j d_j\right|^2 \end{aligned}$$

or

$$\begin{aligned} M &< 2-\beta(\alpha+1) + \sum_{\substack{j=2 \\ j\neq lk+1}}^{\infty}\left((1-\lambda)(1-\alpha\beta)j + \lambda(1-\alpha\beta)j^2\right)\left|a_j\right| + \\ &\quad \sum_{j=1}^{\infty}\left(\{(1-\lambda)(jk+1)+\lambda\}(1-\alpha\beta) + (1-\beta)(2+jk)\right)\left|a_{jk+1}\right| + \\ &\quad \sum_{j=1}^{\infty}\left\{\left[(1-\alpha\beta)(jk+1) + (1-\beta)\right](jk+1)\right\}\left|a_{jk+1}\right|^2. \end{aligned}$$

For $M < 0$, we obtain ( 2.19) . Hence, $f \in M_s^k[\alpha, \beta, \lambda]$. ∎

**2.10. Theorem.** *Let $f$ and $f_k$ be defined by* (1.1) *and* (1.2)*. Suppose that $f \in M_s^k[\alpha, \beta, \lambda]$ and $F_k$ is given by (2.1). Then*

$$\begin{aligned} &\left|(1+(j-1)\lambda)a_j d_j - (\lambda j + (1-\lambda))j a_j\right|^2 \\ < \; &\beta\left|\alpha+1\right|^2 + \beta\sum_{k=2}^{j-1}\left[\left(\alpha^2(1-\lambda)^2\left|1+d_k\right|^2 + \alpha^2\lambda^2\left|k\right|^2 + 2\alpha(1-\lambda)\left|1+\left|d_k\right|\right|\alpha\lambda\left|k\right|\right)\left|k a_k\right|\right] \\ &+2\beta\sum_{k=2}^{j-1}\left(\alpha(1-\lambda)\left|1+d_k\right| + \alpha\lambda\left|k\right|\right)\left|\alpha\lambda+1+k\right|\left|d_k\right|\left|a_k\right| \\ &+\sum_{k=2}^{j-1}\left[\beta\left(\left|\alpha k + d_k\right|^2\left|k\right|^2 + \left|\alpha\lambda+1+k\right|^2\right) - \sum_{k=2}^{j-1}\left|1+(k-1)\lambda\right|^2\right]\left|a_k\right|^2\left|d_k\right|^2 \\ &+\sum_{k=2}^{j-1}\left|\lambda k + (1-\lambda)\right|^2\left|k\right|^2\left|a_k\right|^2 - 2\sum_{k=2}^{j-1}\left|1+(k-1)\lambda\right|\left|\lambda k + (1-\lambda)\right|\left|k\right|\left|a_k\right|^2\left|d_k\right|]. \end{aligned}$$

*For the proof of this theorem, we use Lemma 1.4,* (1.2) *and follow the same lines as in Theorem 2.9.*

## References

[1] Eenigenburg, P., Miller, S.S., Mocanu, P.T., & Read, M.O. On a Briot-Bouquet differential subordination,Rev.Roumanie Math.Pures Appl.29, 567-573, 1964.

[2] Gao, C.-Y., & Zhou, S.-Q. A new subclass of close-to-convex functions, Soochow J. Math. 31, 41–49, 2005.

[3] Liu, M.-S. On a subclass of p-valent close-to-convex functions of order and type, J. Math. Study, 30, 102-104, 1997(in chinese).

[4] Parvatham, R. On Starlike and -Close-to convex functions with respect to n-symmetric points,Indian J.Pure Appl.Math. 16, 1114-1122, 1986.

[5] Sakaguchi, K. On certain univalent mapping, J. Math. Soc. Japan 11, 72–75, 1959.

[6] Sudharsan, T. V., Balasubrahmanyam, P., & Subramanian,.K. G. On functions starlike with respect to symmetric and conjugate points, Taiwanese J. Math. 2, 57–68, 1998.

[7] Tuneski, N. and Darus, M. On functions that are Janowski starlike with respect to N-symmetric points, Hacettepe J. Math. and Stat., 41(2), 271 – 275, 2012.

[8] Wang, Z.G. A new subclass of quasi-convex functions with respect to k-symmetric points, Lobachevskii J. Math. 19, 41-50, 2005.

[9] Wang, Z.-G., Gao, C.-Y., & Yuan, S.-M. On certain subclasses of close-to-convex and quasi-convex functions with respect to k-symmetric points, J. Math. Anal. Appl., 322, 97-106, 2006.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 15 – 21

# Fixed points of ordered $F$-contractions

Gonca Durmaz[*] Gülhan Mınak[†] and Ishak Altun[‡ §]

**Abstract**

In his recent paper, Wardowski [16] introduced the concept of $F$-contraction, which is a proper generalization of ordinary contraction on a complete metric space. Then, some generalizations of $F$-contractions including multivalued case are obtained in [2, 4, 7, 13]. In this paper, by considering both $F$-contractions and fixed point result on ordered metric spaces, we introduce a new concept of ordered $F$-contraction on ordered metric space. Then, we give a fixed point theorem for such mapping. To support our result, we give an example showing that our main theorem is applicable, but both results of Ran and Reurings [12] and Wardowski [16] are not.

**Keywords:** Fixed point, $F$-contraction, ordered $F$-contraction, complete metric space.

*2000 AMS Classification:* 54H25, 47H10.

*Received* : 10.08.2014 *Accepted* : 09.01.2015 *Doi* : 10.15672/HJMS.20164512482

## 1. Introduction and preliminaries

Recently, combining the ideas of Tarski's fixed point theorem on ordered sets and Banach fixed point theorem on a complete metric space, Ran and Reurings [12] obtained a fixed point result on an ordered complete metric space as follows:

[*]Department of Mathematics, Faculty of Science and Arts, Kirikkale University, 71450 Yahsihan, Kirikkale, Turkey, Email: gncmatematik@hotmail.com

[†]Department of Mathematics, Faculty of Science and Arts, Kirikkale University, 71450 Yahsihan, Kirikkale, Turkey, Email: g.minak.28@gmail.com

[‡]Department of Mathematics, Faculty of Science and Arts, Kirikkale University, 71450 Yahsihan, Kirikkale, Turkey, Email: ishakaltun@yahoo.com

[§]Corresponding Author.

**1.1. Theorem.** *Let $(X, \preceq)$ be an ordered set and $d$ be a metric on $X$ such that $(X, d)$ is a complete metric space. Let $T : X \to X$ be a continuous and nondecreasing mapping such that there exists $L \in [0, 1)$ with*

$$d(Tx, Ty) \leq Ld(x, y) \text{ for all } x, y \in X \text{ with } x \preceq y. \tag{1.1}$$

*If there exists $x_0 \in X$ such that $x_0 \preceq Tx_0$, then $T$ has a fixed point.*

In this theorem, the usual contraction of Banach fixed point principle is weakened but at the expense that the operator is monotone. Then many fixed point theorists such as Abbas et al. [1], Agarwal et al. [3], Ćirić et al. [5], Kumam et al. [6], Nashine and Altun [8] and O'Regan and Petruşel [11] focused on this interesting result and obtained a lot of generalizations and variants. For example, taking the regularity of the space, which will be define thereinafter, instead of continuity of $T$, Nieto [9] obtained a parallel result. There are several applications of the theorems in this direction to linear and nonlinear matrix equations, differential equations and integral equations (See for example [10, 12]).

On the other hand, in 2012, one of the most popular fixed point theorem on a complete metric space is given by Wardowski [16]. For this, he introduced the concept of $F$-contraction, which is a proper generalization of ordinary contraction. For the sake of completeness we recall this concept. Let $\mathcal{F}$ be the set of all functions $F : (0, \infty) \to \mathbb{R}$ satisfying the following conditions:

$(F1)$ $F$ is strictly increasing, i.e., for all $\alpha, \beta \in (0, \infty)$ such that $\alpha < \beta, F(\alpha) < F(\beta)$,
$(F2)$ for each sequence $\{a_n\}$ of positive numbers,

$$\lim_{n\to\infty} a_n = 0 \text{ if and only if } \lim_{n\to\infty} F(a_n) = -\infty,$$

$(F3)$ there exists $k \in (0, 1)$ such that $\lim_{\alpha\to 0^+} \alpha^k F(\alpha) = 0$.

Some examples of the functions belonging $\mathcal{F}$ are $F_1(\alpha) = \ln \alpha$, $F_2(\alpha) = \alpha + \ln \alpha$, $F_3(\alpha) = -\frac{1}{\sqrt{\alpha}}$ and $F_4(\alpha) = \ln\left(\alpha^2 + \alpha\right)$.

**1.2. Definition** ([16])**.** Let $(X, d)$ be a metric space and $T : X \to X$ be a mapping. Then $T$ is said to be an $F$-contraction if $F \in \mathcal{F}$ and there exists $\tau > 0$ such that

$$\forall x, y \in X \ [d(Tx, Ty) > 0 \Rightarrow \tau + F(d(Tx, Ty)) \leq F(d(x, y))]. \tag{1.2}$$

If we take $F(\alpha) = \ln \alpha$ in Definition 1.2, the inequality (1.2) turns to

$$d(Tx, Ty) \leq e^{-\tau} d(x, y), \text{ for all } x, y \in X, Tx \neq Ty.$$

It is clear that for $x, y \in X$ such that $Tx = Ty$, the inequality $d(Tx, Ty) \leq e^{-\tau} d(x, y)$ also holds. Therefore $T$ is an ordinary contraction with contractive constant $L = e^{-\tau}$. Therefore every ordinary contraction is also $F$-contraction, but the converse may not be true as shown in the Example 2.5 of [16]. If we choose $F(\alpha) = \alpha + \ln \alpha$, the inequality (1.2) turns to

$$\frac{d(Tx, Ty)}{d(x, y)} e^{d(Tx,Ty) - d(x,y)} \leq e^{-\tau}, \text{ for all } x, y \in X, Tx \neq Ty.$$

In addition, Wardowski concluded that every $F$-contraction $T$ is a contractive mapping, i.e.,

$$d(Tx, Ty) < d(x, y), \text{ for all } x, y \in X, Tx \neq Ty.$$

Thus, every $F$-contraction is a continuous map. Also, Wardowski concluded that if $F_1, F_2 \in \mathcal{F}$ with $F_1(\alpha) \leq F_2(\alpha)$ for all $\alpha > 0$ and $G = F_2 - F_1$ is nondecreasing, then every $F_1$-contraction $T$ is an $F_2$-contraction.

The following theorem is main result of Wardowski [16];

**1.3. Theorem.** *Let $(X, d)$ be a complete metric space and let $T : X \to X$ be an $F$-contraction. Then $T$ has a unique fixed point in $X$.*

Considering Theorem 1.3, some extensions and generalizations are obtained in [2, 4, 7, 13, 14, 15]. The aim of this paper is to introduce the concept of ordered $F$-contractions on ordered metric space, by taking into account the ideas of Wardowski [16] and Ran and Reurings [12].

## 2. The results

Let $(X, \preceq)$ be an ordered set and $d$ be a metric on $X$, then we will say that the tripled $(X, \preceq, d)$ is an ordered metric space. If $(X, d)$ is complete, then $(X, \preceq, d)$ will be called ordered complete metric space. We will say that $X$ is regular, if the ordered metric space $(X, \preceq, d)$ provides the following condition:

$$\begin{cases} \text{If } \{x_n\} \subset X \text{ is a non-decreasing sequence with } x_n \to x \text{ in } X, \\ \text{then } x_n \preceq x \text{ for all } n. \end{cases}$$

**2.1. Definition.** Let $(X, \preceq, d)$ be an ordered metric space and $T : X \to X$ be a mapping. Let

$$Y = \{(x, y) \in X \times X : x \preceq y, d(Tx, Ty) > 0\}.$$

We say that $T$ is an ordered $F$-contraction if $F \in \mathfrak{F}$ and there exists $\tau > 0$ such that

$$\forall (x, y) \in Y \Rightarrow \tau + F(d(Tx, Ty)) \leq F(d(x, y)). \tag{2.1}$$

**2.2. Theorem.** *Let $(X, \preceq, d)$ be an ordered complete metric space and $T : X \to X$ be an ordered $F$-contraction. Let $T$ is nondecreasing mapping and there exists $x_0 \in X$ such that $x_0 \preceq Tx_0$. If $T$ is continuous or $X$ is regular, then $T$ has a fixed point.*

*Proof.* Let $x_0 \in X$ be as mentioned in the hypotheses. We define a sequence $\{x_n\}$ in $X$ such that $x_n = Tx_{n-1} = T^n x_0$ for all $n \in \mathbb{N}$. If there exists $n_0 \in \mathbb{N}$ for which $x_{n_0} = x_{n_0+1}$, then $x_{n_0}$ is a fixed point of $T$ and so the proof is completed. Thus, suppose that for every $n \in \mathbb{N}$, $x_{n+1} \neq x_n$. Since $x_0 \preceq Tx_0$ and $T$ is nondecreasing, we obtain

$$x_0 \preceq x_1 \preceq x_2 \preceq \cdots \preceq x_n \preceq \cdots.$$

Now, since $x_n \preceq x_{n+1}$ and $d(Tx_n, Tx_{n-1}) > 0$ for every $n \in \mathbb{N}$, then $(x_n, x_{n+1}) \in Y$, and so, we can use the inequality (2.1) for the consecutive terms of $\{x_n\}$, then we have

$$F(d(x_{n+1}, x_n)) = F(d(Tx_n, Tx_{n-1})) \leq F(d(x_n, x_{n-1})) - \tau. \tag{2.2}$$

Denote $\gamma_n = d(x_n, x_{n+1})$ for $n \in \mathbb{N}$. Then, from (2.2) we obtain

$$F(\gamma_n) \leq F(\gamma_{n-1}) - \tau \leq F(\gamma_{n-2}) - 2\tau \leq \cdots \leq F(\gamma_0) - n\tau. \tag{2.3}$$

From (2.3), we get $\lim_{n\to\infty} F(\gamma_n) = -\infty$. Thus, from (F2), we have

$$\lim_{n\to\infty} \gamma_n = 0.$$

From (F3), there exists $k \in (0, 1)$ such that

$$\lim_{n\to\infty} \gamma_n^k F(\gamma_n) = 0.$$

By (2.2), the following holds for all $n \in \mathbb{N}$

$$\gamma_n^k F(\gamma_n) - \gamma_n^k F(\gamma_0) \leq -\gamma_n^k n\tau \leq 0. \tag{2.4}$$

Letting $n \to \infty$ in (2.4), we obtain that

$$\lim_{n\to\infty} n\gamma_n^k = 0. \tag{2.5}$$

From (2.5), there exits $n_1 \in \mathbb{N}$ such that $n\gamma_n^k \leq 1$ for all $n \geq n_1$. So, we have

$$\gamma_n \leq \frac{1}{n^{1/k}}, \tag{2.6}$$

for all $n \geq n_1$. In order to show that $\{x_n\}$ is a Cauchy sequence consider $m, n \in \mathbb{N}$ such that $m > n \geq n_1$. Using the triangular inequality for the metric and from (2.6), we have

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \\ &= \gamma_n + \gamma_{n+1} + \cdots + \gamma_{m-1} \\ &= \sum_{i=n}^{m-1} \gamma_i \\ &\leq \sum_{i=n}^{\infty} \gamma_i \\ &\leq \sum_{i=n}^{\infty} \frac{1}{i^{1/k}}. \end{aligned}$$

Passing to limit $n \to \infty$, we get $d(x_n, x_m) \to 0$. This yields that $\{x_n\}$ is a Cauchy sequence in $X$. Since $X$ is complete, the sequence $\{x_n\}$ converges to some point $z \in X$.

Now, if $T$ is continuous, then we have

$$z = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} Tx_n = T \lim_{n\to\infty} x_n = Tz$$

and so $z$ is a fixed point of $T$.

Now suppose $X$ is regular, then $x_n \preceq z$ for all $n \in \mathbb{N}$. We will consider the following two cases:

Case 1. If there exists $n_0 \in \mathbb{N}$ for which $x_{n_0} = z$, then we obtain

$$Tz = Tx_{n_0} = x_{n_0+1} \preceq z.$$

Also, since $x_{n_0} \preceq x_{n_0+1}$, then $z \preceq Tz$ and thus, $z = Tz$.

Case 2. Now, suppose that $x_n \neq z$ for every $n \in \mathbb{N}$ and $d(z, Tz) > 0$. Since $\lim_{n\to\infty} x_n = z$, then there exists $n_1 \in \mathbb{N}$ such that $d(x_{n+1}, Tz) > 0$ and $d(x_n, z) < \frac{d(z,Tz)}{2}$ for all $n \geq n_1$. Note that in this case $(x_n, z) \in Y$. Therefore, by considering (F1), we have, for $n \geq n_1$,

$$\tau + F(d(Tx_n, Tz)) \leq F(d(x_n, z)) \leq F(\frac{d(z, Tz)}{2}),$$

which yields

$$d(x_{n+1}, Tz) \leq \frac{d(z, Tz)}{2}. \tag{2.7}$$

Taking limit as $n \to \infty$, we deduce that

$$d(z, Tz) \leq \frac{d(z, Tz)}{2},$$

a contradiction. Therefore, we conclude that $d(z, Tz) = 0$, i.e. $z = Tz$. □

**2.3. Example.** Let $A = \{\frac{1}{n^2} : n \in \mathbb{N}\} \cup \{0\}$, $B = \{2, 3\}$ and $X = A \cup B$. Define an order relation $\preceq$ on $X$ as

$$x \preceq y \Leftrightarrow [x = y \text{ or } x, y \in A \text{ with } x \leq y],$$

where $\leq$ is usual order. Obviously, $(X, \preceq, d)$ is ordered complete metric space with the usual metric $d$. Let $T : X \to X$ be given by

$$Tx = \begin{cases} \frac{1}{(n+1)^2} & , \quad x = \frac{1}{n^2} \\ x & , \quad x \in \{0, 2, 3\} \end{cases}.$$

It is easy to see that $T$ is nondecreasing. Also, for $x_0 = 0$ we have $x_0 \preceq Tx_0$.

On the other side, taking $F$ with

$$F(\alpha) = \begin{cases} \frac{\ln \alpha}{\sqrt{\alpha}} & , \quad 0 < \alpha < e^2 \\ \alpha - e^2 + \frac{2}{e} & , \quad \alpha \geq e^2 \end{cases}.$$

It is easy to see that the conditions $(F1), (F2)$ and $(F3)$ (for $k = \frac{2}{3}$) are satisfied. We claim that $T$ is an ordered $F$-contraction with $\tau = \ln 2$. To see this, let us consider the following calculations:

It is obvious that

$$\begin{aligned} Y &= \{(x, y) \in X \times X : x \preceq y, d(Tx, Ty) > 0\} \\ &= \{(x, y) \in X \times X : x, y \in A \text{ and } x < y\}. \end{aligned}$$

Therefore, to see (2.1), it is sufficient to show that

$$\forall (x, y) \in Y \Rightarrow \ln 2 + F(d(Tx, Ty)) \leq F(d(x, y))$$

$$\Leftrightarrow x, y \in A \text{ and } x < y \Rightarrow d(Tx, Ty)^{\frac{1}{\sqrt{d(Tx,Ty)}}} d(x, y)^{-\frac{1}{\sqrt{d(x,y)}}} \leq \frac{1}{2} \tag{2.8}$$

$$\Leftrightarrow x, y \in A \text{ and } x < y \Rightarrow |Tx - Ty|^{\frac{1}{\sqrt{|Tx-Ty|}}} |x - y|^{-\frac{1}{\sqrt{|x-y|}}} \leq \frac{1}{2}.$$

Using Example 5 of [7], we can see that (2.8) is true. Also, $T$ is continuous (and $X$ is regular). Therefore, all conditions of Theorem 2.2 are satisfied, and so, $T$ has a fixed point in $X$.

On the other hand, since $0 \preceq \frac{1}{n^2}$ for all $n \in \mathbb{N}$ and

$$\sup_{n \in \mathbb{N}} \frac{d(T0, T\frac{1}{n^2})}{d(0, \frac{1}{n^2})} = \sup_{n \in \mathbb{N}} \frac{n^2}{(n+1)^2} = 1,$$

then Theorem 1.1 , which is main result of [12], is not applicable to this example. Again, since

$$d(T2, T3) = d(2, 3) = 1,$$

then for all $F \in \mathcal{F}$ and $\tau > 0$ we have

$$\tau + F(d(T2, T3)) > F(d(2, 3)).$$

Therefore, Theorem 1.3, which is main result of [16], is not applicable to this example.

**2.4. Remark.** In Theorem 2.2, if we assume the following condition:

(2.9) every pair of elements has a lower bound and upper bound,

then, the fixed point of $T$ is unique. To see this, it is sufficient to show that for every $x \in X$,

$$\lim_{n \to \infty} T^n x = z,$$

where $z$ is the fixed point of $T$ such that

$$z = \lim_{n \to \infty} T^n x_0.$$

For this we will consider the following cases: Let $x \in X$ and $x_0$ be as in Theorem 2.2.

Case 1. If $x \preceq x_0$ or $x_0 \preceq x$, then $T^n x \preceq T^n x_0$ or $T^n x_0 \preceq T^n x$ for all $n \in \mathbb{N}$. If $T^{n_0} x = T^{n_0} x_0$ for some $n_0 \in \mathbb{N}$, then $T^n x \to z$. Now let $T^n x_0 \neq T^n x$ for all $n \in \mathbb{N}$, then $d(T^n x_0, T^n x) > 0$ and so $(T^n x_0, T^n x) \in Y$ for all $n \in \mathbb{N}$. Therefore from (2.1), we have

$$\begin{aligned} F(d(T^n x_0, T^n x)) &\leq F(d(T^{n-1} x_0, T^{n-1} x)) - \tau \\ &\leq F(d(T^{n-2} x_0, T^{n-2} x)) - 2\tau \\ &\vdots \\ &\leq F(d(x_0, x)) - n\tau. \end{aligned} \tag{2.10}$$

Taking into account (F2), from (2.10) we have $\lim_{n\to\infty} d(T^n x_0, T^n x) = 0$, and so, $\lim_{n\to\infty} T^n x_0 = \lim_{n\to\infty} T^n x = z$.

Case 2. If $x \npreceq x_0$ or $x_0 \npreceq x$, then from (2.9), there exist $x_1, x_2 \in X$ such that

$$x_2 \preceq x \preceq x_1 \text{ and } x_2 \preceq x_0 \preceq x_1.$$

Therefore, as in the Case 1, we can show that

$$\lim_{n\to\infty} T^n x_1 = \lim_{n\to\infty} T^n x_2 = \lim_{n\to\infty} T^n x = \lim_{n\to\infty} T^n x_0 = z.$$

**2.5. Remark.** As we can see in the Example 2.3, if the condition (2.9) is not satisfied, then the fixed point of $T$ may not be unique.

**Acknowledgment.** The authors are grateful to the referees because their suggestions contributed to improve the paper.

## References

[1] Abbas, M., Nazir, T. and Radenović, S. *Common fixed points of four maps in partially ordered metric spaces*, Appl. Math. Lett, **24**(9) (2011) 1520-1526.

[2] Acar, Ö., Durmaz, G. and Mınak, G. *Generalized multivalued $F$-contractions on complete metric spaces*, Bulletin of the Iranian Mathematical Society, **40**(6) (2014) 1469-1478.

[3] Agarwal, R. P., El-Gebeily, M. A. and O'Regan, D. *Generalized contractions in partially ordered metric spaces*, Appl. Anal., **87**(1) (2008) 109-116.

[4] Altun, I., Mınak, G. and Dağ, H. *Multivalued $F$-contractions on complete metric spaces*, Jounal of Nonlinear and Complex Analysis, **16**(4) (2015) 658-666.

[5] Ćirić, Lj. B., Abbas, M., Saadati, R. and Hussain, N. *Common fixed points of almost generalized contractive mappings in ordered metric spaces*, Appl. Math. Comput., **217**(12) (2011) 5784-5789.

[6] Kumam, P., Rouzkard, F., Imdad, M. and Gopal, D. *Fixed point theorems on ordered metric spaces through a rational contraction*, Abstr. Appl. Anal., **2013**, Article ID 206515, 9 pages.

[7] Mınak, G., Helvacı A. and Altun, I. *Ćirić type generalized $F$-contractions on complete metric spaces and fixed point results*, Filomat, **28**(6) (2014) 1143-1151.

[8] Nashine, H. K. and Altun, I. *A common fixed point theorem on ordered metric spaces*, Bull. Iranian Math. Soc., **38**(4) (2012) 925-934.

[9] Nieto, J. J. and López, R. R. *Contractive mapping theorems in partially ordered sets and applications to ordinary differential equation*, Order, **22** (2005) 223-239.

[10] Nieto, J. J. and Lopez, R. R. *Existence and uniqueness of fixed point in partially ordered sets and applications to ordinary differential equations*, Acta Math. Sinica (English Ser.), **23**(12) (2007) 2205-2212.

[11] O'Regan, D. and Petruşel, A. *Fixed point theorems for generalized contractions in ordered metric spaces*, J. Math. Anal. Appl., **341** (2008) 1241-1252.

[12] Ran, A. C. M. and Reurings, M. C. B. *A fixed point theorem in partially ordered sets and some application to matrix equations*, Proc. Amer. Math. Soc., **132** (2004) 1435-1443.

[13] Sgroi, M. and Vetro, C. *Multi-valued $F$-contractions and the solution of certain functional and integral equations*, Filomat, **27**(7) (2013) 1259-1268.

[14] Shukla, S. and Radenović, S. *Some common fixed point theorems for $F$-contraction type mappings in* 0*-complete partial metric spaces*, Journal of Mathematics, **2013** Article ID 878730, 7 pages.

[15] Shukla, S., Radenović, S. and Kadelburg, Z. *Some fixed point theorems for ordered $F$-generalized contractions in* 0*-$f$-orbitally complete partial metric spaces*, Theory and Applications of Mathematics and Computer Sciences, **4**(1) (2014) 87-98.

[16] Wardowski, D. *Fixed points of a new type of contractive mappings in complete metric spaces*, Fixed Point Theory Applications, **2012**:94, 6 pages.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 23 – 31

# Existence of symmetric positive solutions for a semipositone problem on time scales

S. Gulsan Topal and Arzu Denk *

**Abstract**

This paper studies the existence of symmetric positive solutions for a second order nonlinear semipositone boundary value problem with integral boundary conditions by applying the Krasnoselskii fixed point theorem. Emphasis is put on the fact that the nonlinear term $f$ may take negative value. An example is presented to demonstrate the application of our main result.

**Keywords:** Positive solution, Symmetric solution, Semipositone problems, Fixed point theorems, Time scales.

*2000 AMS Classification:* 34B15, 39A10.

*Received* : 07.03.2013 *Accepted* : 16.12.2014 *Doi* : 10.15672/HJMS.20164512479

## 1. Introduction

We will be concerned with proving the existence of at least one symmetric positive solution to the semipositone second order nonlinear boundary value problem on a symmetric time scale T given by

$$[g(t)u^{\triangle}(t)]^{\nabla} + \lambda f(t, u(t)) = 0, \qquad t \in (a, b), \tag{1.1}$$

$$\alpha u(a) - \beta \lim_{t\to a^+} g(t)u^{\triangle}(t) = \int_a^b h_1(s)u(s)\nabla s, \tag{1.2}$$

$$\alpha u(b) + \beta \lim_{t\to b^-} g(t)u^{\triangle}(t) = \int_a^b h_2(s)u(s)\nabla s, \tag{1.3}$$

where $\lambda > 0$ is a parameter, $\alpha, \beta > 0$, $\nabla$-differentiable function $g \in C([a, b], (0, \infty))$ is symmetric on $[a, b]$, $h_1, h_2 \in L^1([a, b])$ is nonnegative, symmetric on $[a, b]$ and the continuous function $f : [a, b] \times [0, \infty) \to R$ satisfies $f(b + a - t, u) = f(t, u)$.

*Department of Mathematics, Ege University, 35100 Bornova, Izmir-Turkey.
Email : f.serap.topal@ege.edu.tr ; arzu.denk@hotmail.com

A class of boundary value problems with integral boundary conditions arise naturally in thermal condition problems [4], semiconductor problems [7], and hydrodynamic problems [5]. Such problems include two, three and multi-point boundary conditions and have recently been investigated by many authors [3, 6, 8, 9].

The present work is motivated by recent paper [3]. In this paper, Boucherif considered the following second order boundary value problem with integral boundary conditions

$$x''(t) = f(t, x(t)), \qquad 0 < t < 1, \tag{1.4}$$

$$x(0) - cx'(0) = \int_0^1 g_0(s)x(s)ds, \tag{1.5}$$

$$x(1) - dx'(1) = \int_0^1 g_1(s)x(s)ds, \tag{1.6}$$

where $f : [0, 1] \times R \to R$ is continuous, $g_0, g_1 : [0, 1] \to [0, \infty)$ are continuous and positive, $c$ and $d$ are nonnegative real parameters. The author established some excellent results for the existence of positive solutions to problem $(1.4) - (1.6)$ by using the fixed point theorem in cones.

Throughout this paper T is a symmetric time scale with $a, b$ are points in T. By an interval $(a, b)$, we always mean the intersection of the real interval $(a, b)$ with the given time scale, that is $(a, b) \cap \mathrm{T}$. Other types of intervals are defined similarly. For the details of basic notions connected to time scales we refer to [1, 2].

Now, we present some symmetric definition.

**1.1. Definition.** A time scale T is said to be symmetric if for any given $t \in \mathrm{T}$, we have $b + a - t \in \mathrm{T}$.

**1.2. Definition.** A function $u : \mathrm{T} \to R$ is said to be symmetric on T if for any given $t \in \mathrm{T}$, $u(t) = u(b + a - t)$.

## 2. The Preliminary Lemmas

In this section we collect some preliminary results that will be used in subsequent section.

Throughout the paper we will assume that the following conditions are satisfied:

$(H_1)$ $\alpha, \beta > 0$,

$(H_2)$ $\nabla$-differentiable function $g \in C([a, b], (0, \infty))$ is symmetric on $[a, b]$,

$(H_3)$ the continuous function $f : [a, b] \times [0, \infty) \to R$ is semipositone, i.e., $f(t, u)$ needn't be positive for all $(t, u) \in [a, b] \times [0, \infty)$ and $f(., u)$ is symmetric on $[a, b]$ for all $u \geq 0$,

$(H_4)$ $h_1, h_2 \in L^1([a, b])$ is nonnegative, symmetric on $[a, b]$ and $A > 0$, where $A = \mu + (\beta - K)v_1 - \beta v_2$, $\quad K = \dfrac{\mu}{\alpha}$, $\quad \mu = 2\alpha\beta + \alpha^2 \displaystyle\int_a^b \frac{\Delta r}{g(r)}$, $\quad v_1 = \displaystyle\int_a^b h_1(\tau)\nabla\tau$, $\quad v_2 = \displaystyle\int_a^b h_2(\tau)\nabla\tau$.

The lemmas in this section are based on the boundary value problem

$$-[g(t)u^{\triangle}(t)]^{\nabla} = p(t), \qquad t \in (a, b) \tag{2.1}$$

with boundary conditions $(1.2) - (1.3)$.

To prove the main result, we will employ following lemmas.

**2.1. Lemma.** *Let $(H_1), (H_2)$ hold and $A \neq 0$. Then for any $p \in C([a, b])$, the boundary value problem $(2.1) - (1.2) - (1.3)$ has a unique solution $u$ given by*

$$u(t) = \int_a^b H(t,s)p(s)\nabla s,$$

*where*

$$(2.2) \qquad H(t,s) = G(t,s) + B_1 \int_a^b G(s,\tau)h_1(\tau)\nabla\tau + B_2 \int_a^b G(s,\tau)h_2(\tau)\nabla\tau$$

$$(2.3) \qquad G(t,s) = \frac{1}{\mu}\begin{cases} (\beta + \alpha \int_a^s \frac{\Delta r}{g(r)})(\beta + \alpha \int_t^b \frac{\Delta r}{g(r)}), & a \le s \le t \le b, \\ (\beta + \alpha \int_a^t \frac{\Delta r}{g(r)})(\beta + \alpha \int_s^b \frac{\Delta r}{g(r)}), & a \le t \le s \le b, \end{cases}$$

where $\mu = 2\alpha\beta + \alpha^2 \int_a^b \frac{\Delta r}{g(r)}$, $B_1 = \frac{K-\beta}{A}$, $B_2 = \frac{\beta}{A}$.

**2.2. Lemma.** *Assume that* $(H_1), (H_2)$ *and* $(H_4)$ *hold. Then we have*

*(i)* $H(t,s) > 0$, $\quad G(t,s) > 0$, *for* $t, s \in [a,b]$,

*(ii)* $H(b+a-t, b+a-s) = H(t,s)$, $\quad G(b+a-t, b+a-s) = G(t,s)$, *for* $t, s \in [a,b]$,

*(iii)* $\frac{1}{\mu}\beta^2\gamma \le H(t,s) \le H(s,s) \le \frac{1}{\mu}\gamma D$ *and* $\frac{1}{\mu}\beta^2 \le G(t,s) \le G(s,s) \le \frac{1}{\mu}D$, *for* $t, s \in [a,b]$,

*where* $D = (\beta + \alpha \int_a^b \frac{\Delta r}{g(r)})^2$, $\gamma = 1 + B_1 v_1 + B_2 v_2$.

*Proof.* It is clear that $(i)$ hold. Now we prove that $(ii)$ and $(iii)$ hold. First, we consider $(ii)$. If $t \le s$, then $b+a-t \ge b+a-s$. Using (2.3) and the assumption $(H_2)$, we get

$$\begin{aligned} G(b+a-t, b+a-s) &= \frac{1}{\mu}(\beta + \alpha \int_a^{b+a-s} \frac{\Delta r}{g(r)})(\beta + \alpha \int_{b+a-t}^b \frac{\Delta r}{g(r)}) \\ &= \frac{1}{\mu}(\beta + \alpha \int_b^s \frac{\Delta(b+a-r)}{g(b+a-r)})(\beta + \alpha \int_t^a \frac{\Delta(b+a-r)}{g(b+a-r)}) \\ &= \frac{1}{\mu}(\beta + \alpha \int_s^b \frac{\Delta r}{g(r)})(\beta + \alpha \int_a^t \frac{\Delta r}{g(r)}) = G(t,s). \end{aligned}$$

Similarly, we can prove that $G(b+a-t, b+a-s) = G(t,s)$, for $s \le t$. Thus we have $G(b+a-t, b+a-s) = G(t,s)$, for $t, s \in [a,b]$. Now by (2.2), for $t, s \in [a,b]$, we have

$$\begin{aligned} H(b+a-t, b+a-s) &= G(b+a-t, b+a-s) + B_1 \int_a^b G(b+a-s,\tau)h_1(\tau)\nabla\tau \\ &\quad + B_2 \int_a^b G(b+a-s,\tau)h_2(\tau)\nabla\tau \\ &= G(t,s) + B_1 \int_b^a G(b+a-s, b+a-\tau)h_1(b+a-\tau)\nabla(b+a-\tau) \\ &\quad + B_2 \int_b^a G(b+a-s, b+a-\tau)h_2(b+a-\tau)\nabla(b+a-\tau) \\ &= G(t,s) + B_1 \int_a^b G(s,\tau)h_1(\tau)\nabla\tau + B_2 \int_a^b G(s,\tau)h_2(\tau)\nabla\tau \\ &= H(t,s). \end{aligned}$$

So $(ii)$ is established. Now we show that $(iii)$ holds. In fact, if $t \le s$, from (2.3) and the assumption $(H_2)$, then we get

$$G(t,s)=\frac{1}{\mu}(\beta+\alpha\int_a^t\frac{\Delta r}{g(r)})(\beta+\alpha\int_s^b\frac{\Delta r}{g(r)})\le\frac{1}{\mu}(\beta+\alpha\int_a^s\frac{\Delta r}{g(r)})(\beta+\alpha\int_s^b\frac{\Delta r}{g(r)})$$
$$=G(s,s)$$
$$\le\frac{1}{\mu}(\beta+\alpha\int_a^b\frac{\Delta r}{g(r)})(\beta+\alpha\int_a^b\frac{\Delta r}{g(r)})=\frac{1}{\mu}(\beta+\alpha\int_a^b\frac{\Delta r}{g(r)})^2=\frac{1}{\mu}D.$$

Similarly, we can prove that $G(t,s)\le G(s,s)\le\frac{1}{\mu}D$ for $s\le t$.

Therefore $G(t,s)\le G(s,s)\le\frac{1}{\mu}D$, for $t,s\in[a,b]$. And then, by (2.2), we have

$$H(t,s)=G(t,s)+B_1\int_a^b G(s,\tau)h_1(\tau)\nabla\tau+B_2\int_a^b G(s,\tau)h_2(\tau)\nabla\tau$$
$$\le G(s,s)+B_1\int_a^b G(\tau,\tau)h_1(\tau)\nabla\tau+B_2\int_a^b G(\tau,\tau)h_2(\tau)\nabla\tau$$
$$\le\frac{1}{\mu}D+\frac{1}{\mu}DB_1\int_a^b h_1(\tau)\nabla\tau+\frac{1}{\mu}DB_2\int_a^b h_2(\tau)\nabla\tau=\frac{1}{\mu}D(1+B_1v_1+B_2v_2)$$
$$=\frac{1}{\mu}D\gamma.$$

On the other hand, for $t,s\in[a,b]$, we have

$$G(t,s)\ge\frac{1}{\mu}(\beta+\alpha\int_a^a\frac{\Delta r}{g(r)})(\beta+\alpha\int_b^b\frac{\Delta r}{g(r)})=\frac{1}{\mu}\beta^2.$$

And then, we get

$$H(t,s)=G(t,s)+B_1\int_a^b G(s,\tau)h_1(\tau)\nabla\tau+B_2\int_a^b G(s,\tau)h_2(\tau)\nabla\tau$$
$$\ge\frac{1}{\mu}\beta^2+\frac{1}{\mu}\beta^2B_1\int_a^b h_1(\tau)\nabla\tau+\frac{1}{\mu}\beta^2B_2\int_a^b h_2(\tau)\nabla\tau=\frac{1}{\mu}\beta^2\gamma.$$

Thus for $t,s\in[a,b]$, we have

$$\frac{1}{\mu}\beta^2\gamma\le H(t,s)\le H(s,s)\le\frac{1}{\mu}\gamma D\text{ and }\frac{1}{\mu}\beta^2\le G(t,s)\le G(s,s)\le\frac{1}{\mu}D.$$

This completes the proof.

**2.3. Lemma.** *Let $w$ be the unique positive solution of the boundary value problem*

$$[g(t)u^{\triangle}(t)]^{\nabla}+1=0 \tag{2.4}$$

*with the boundary condition* (1.2) – (1.3). *Then,*

$$w(t)\le C\delta,\quad t\in[a,b],$$

*where*

$$\delta=\frac{\beta^2}{D},\quad C=\frac{b-a}{\mu\beta^2}D^2\gamma \tag{2.5}$$

*Proof.* Using Lemma 2.2, for all $t\in[a,b]$, we have

$$w(t) = \int_a^b H(t,s)\nabla s \le \frac{1}{\mu}\gamma D \int_a^b \nabla s = C\delta.$$

The proof is complete.

Let $E$ denote the Banach space $C[a,b]$ with the norm $\|u\| = \max_{t\in[a,b]} |u(t)|$. Define the cone $P \subset E$ by $P = \{u \in E : u(t)$ is symmetric and $u(t) \ge \delta\|u\|$ for $t \in [a,b]\}$.

To obtain the a positive solution of BVP $(1.1)-(1.3)$, the following fixed point theorem is essential.

**2.4. Theorem.** *Let $E = (E, \|.\|)$ be a Banach space, and let $P \subset E$ be a cone in $B$. Assume $\Omega_1, \Omega_2$ are bounded open subsets of $E$ with $0 \in \Omega_1, \overline{\Omega_1} \subset \Omega_2$, and let*

$S : P \cap (\overline{\Omega_2}\backslash\Omega_1) \to P$

*be a continuous and completely continuous operator such that, either*

*(a)* $\|Su\| \le \|u\|, u \in P \cap \partial\Omega_1,$ *and* $\|Su\| \ge \|u\|, u \in P \cap \partial\Omega_2,$ *or*

*(b)* $\|Su\| \ge \|u\|, u \in P \cap \partial\Omega_1,$ *and* $\|Su\| \le \|u\|, u \in P \cap \partial\Omega_2.$

*Then $S$ has a fixed point in $P \cap (\overline{\Omega_2}\backslash\Omega_1)$.*

## 3. Main Results

In this section, we apply the Krasnoselskii fixed point theorem to obtain the existence of at least one symmetric positive solution for the nonlinear boundary value problem $(1.1) - (1.3)$.

The main result of this paper is following:

**3.1. Theorem.** *Let $(H_1) - (H_4)$ hold. Assume that*

$(C_1)$ *There exists a constant $M > 0$ such that $f(t,u) \ge -M$ for all $(t,u) \in [a,b] \times [0,\infty)$,*

$(C_2)$ *There exist $t_1, t_2 \in (a,b)$ such that*

$$\lim_{u\to\infty} \frac{f(t,u)}{u} = \infty$$

*uniformly on $[t_1, t_2]$,*

$(C_3)$ *$r$ is a given positive real number and the parameter $\lambda$ satisfies*

$$0 < \lambda \le \eta := \min\{\frac{r}{M_1\|w\|}, \frac{r}{2MC}\} \tag{3.1}$$

*where $M_1 = \max\{f(t,u) + M : (t,u) \in [a,b] \times [0,r]\}$.*

*Then the boundary value problem $(1.1) - (1.3)$ has at least one symmetric positive solution $u$ such that $\|u\| \ge \dfrac{r}{2}$.*

*Proof.* Let $x(t) = \lambda M w(t)$, where $w$ is the unique solution of the boundary value problem $(2.4) - (1.2) - (1.3)$.

We shall show that the following boundary value problem

$$[g(t)y^{\triangle}(t)]^{\nabla} + \lambda F(t, y(t) - x(t)) = 0,\ t \in (a,b), \tag{3.2}$$

$$\alpha y(a) - \beta \lim_{t\to a^+} g(t)y^{\triangle}(t) = \int_a^b h_1(s)y(s)\nabla s, \tag{3.3}$$

$$\alpha y(b) + \beta \lim_{t\to b^-} g(t)y^{\triangle}(t) = \int_a^b h_2(s)y(s)\nabla s, \tag{3.4}$$

where

$$F(t,z) = \begin{cases} f(t,z) + M, & z \geq 0, \\ f(t,0) + M, & z \leq 0, \end{cases}$$

has at least one positive solution. Thereafter we shall obtain at least one positive solution for the boundary value problem (1.1) – (1.3).

It is well known that the existence of positive solution to the boundary value problem (3.2) – (3.4) is equivalent to the existence of fixed point of the operator $S$. So we shall seek a fixed point of $S$ in our cone $P$ where the operator $S : E \to E$ is defined by

$$Sy(t) = \lambda \int_a^b H(t,s)F(s, y(s) - x(s))\nabla s, \quad t \in [a,b].$$

First, it is obvious that $S$ is continuous and completely continuous.

Now we shall prove that $S(P) \subseteq P$. Let $y \in P$. Then, using Lemma 2.2, we get for $t \in [a,b]$,

$$Sy(t) = \lambda \int_a^b H(t,s)F(s, y(s) - x(s))\nabla s \leq \frac{\lambda}{\mu}\gamma D \int_a^b F(s, y(s) - x(s))\nabla s,$$

and so

$$\|Sy\| \leq \frac{\lambda}{\mu}\gamma D \int_a^b F(s, y(s) - x(s))\nabla s. \tag{3.5}$$

Now, using Lemma 2.2 and (3.5), we obtain for $t \in [a,b]$,

$$\begin{aligned} Sy(t) &= \lambda \int_a^b H(t,s)F(s, y(s) - x(s))\nabla s \geq \frac{\lambda}{\mu}\beta^2\gamma \int_a^b F(s, y(s) - x(s))\nabla s \\ &= \frac{\lambda}{\mu}\delta\gamma D \int_a^b F(s, y(s) - x(s))\nabla s \geq \delta\|Sy\|. \end{aligned}$$

On the other hand, noticing $y(t), x(t)$ and $f(t,u)$ are symmetric on $[a,b]$, we have

$$\begin{aligned} Sy(b+a-t) &= \lambda \int_a^b H(b+a-t,s)F(s, y(s) - x(s))\nabla s \\ &= \lambda \int_a^b H(b+a-t,s)(f(s, y(s) - x(s)) + M)\nabla s \\ &= \lambda \int_b^a H(b+a-t, b+a-s)(f(s,(y-x)(b+a-s)) + M)\nabla(b+a-s) \\ &= \lambda \int_a^b H(t,s)(f(s,(y-x)(s)) + M)\nabla s \\ &= \lambda \int_a^b H(t,s)F(s,(y-x)(s))\nabla s = Sy(t) \end{aligned}$$

Therefore $Sy$ is symmetric.
So, we get $S(P) \subseteq P$.

Let $\Omega_1 = \{y \in E : \|y\| < r\}$. We shall prove that $\|Sy\| \leq \|y\|$ for $y \in P \cap \partial\Omega_1$. If $y \in P \cap \partial\Omega_1$, then $\|y\| = r$. By definition and (3.1), we find for $t \in [a,b]$,

$$Sy(t) = \lambda \int_a^b H(t,s)F(s, y(s) - x(s))\nabla s \leq \lambda M_1 \int_a^b H(t,s)\nabla s \leq \lambda M_1\|w\| \leq r.$$

Therefore, we get $\|Sy\| \leq r = \|y\|$ for $y \in P \cap \partial\Omega_1$.

Let $K$ be a positive real number such that

$$\frac{1}{2}\lambda K(t_2 - t_1)\delta \frac{1}{\mu}\beta^2\gamma > 1. \tag{3.6}$$

In view of $(C_2)$, there exists $N > 0$ such that for all $z \geq N$ and $t \in [t_1, t_2]$,

$$F(t, z) = f(t, z) + M \geq Kz \tag{3.7}$$

Now, set

$$R = r + \frac{2N}{\delta}. \tag{3.8}$$

Let $\Omega_2 = \{y \in E : \|y\| < R\}$. We shall prove that $\|Sy\| \geq \|y\|$ for $y \in P \cap \partial\Omega_2$. If $y \in P \cap \partial\Omega_2$, then $\|y\| = R$. So from Lemma 2.3 and the fact that $y \in P$ , we get for $t \in [a, b]$,

$$x(t) = \lambda M w(t) \leq \lambda M C \delta \leq \lambda M C \frac{y(t)}{R}.$$

This implies for $t \in [a, b]$,

$$y(t) - x(t) \geq (1 - \frac{\lambda MC}{R})y(t) \geq (1 - \frac{\lambda MC}{R})\delta R,$$

and, from (3.1) and (3.8), we get for $t \in [t_1, t_2]$,

$$y(t) - x(t) \geq \frac{1}{2}R\delta \geq N. \tag{3.9}$$

Thus, by (3.7) and (3.9), we see that for $t \in [t_1, t_2]$,

$$F(t, y(t) - x(t)) \geq K(y(t) - x(t)) \geq \frac{1}{2}KR\delta. \tag{3.10}$$

Considering Lemma 2.2 and (3.10), we get for $t \in [a, b]$,

$$\begin{aligned} Sy(t) &= \lambda \int_a^b H(t, s)F(s, y(s) - x(s))\nabla s \geq \lambda \frac{1}{\mu}\beta^2\gamma \int_{t_1}^{t_2} F(s, y(s) - x(s))\nabla s \\ &\geq \frac{1}{2\mu}\lambda KR\delta\beta^2\gamma \int_{t_1}^{t_2} \nabla s \end{aligned}$$

and so by (3.6),

$$\|Sy\| \geq \frac{1}{2\mu}\lambda KR(t_2 - t_1)\delta\beta^2\gamma \geq R.$$

Therefore, we get $\|Sy\| \geq R = \|y\|$ for $y \in P \cap \partial\Omega_2$.

Then it follows from Theorem 2.1 that S has a fixed point $y \in P$ such that

$$r \leq \|y\| \leq R. \tag{3.11}$$

Moreover, using (3.1), (3.11) and Lemma 2.3, we obtain for $t \in [a, b]$,

$$y(t) \geq \delta\|y\| \geq r\delta \geq 2\lambda MC\delta \geq 2\lambda Mw(t) = 2x(t). \tag{3.12}$$

Hence,

$u(t) = y(t) - x(t) \geq 0, \quad t \in [a, b].$

On the other hand, $u(t)$ is symmetric on $[a, b]$ since $y$ and $x$ are symmetric. Now, we shall prove that $u$ is a positive solution of the boundary value problem (1.1) – (1.3). Since $y$ is a fixed point of the operator $S$,

$Sy(t) = y(t), \quad t \in [a, b],$

or

$$y(t) = Sy(t) = \lambda \int_a^b H(t,s)F(s, y(s) - x(s))\nabla s$$
$$= \lambda \int_a^b H(t,s)(f(s, y(s) - x(s)) + M)\nabla s$$

Noticing that,

$$w(t) = \int_a^b H(t,s)\nabla s$$

we have for $t \in [a, b]$,

$$y(t) = \lambda \int_a^b H(t,s)f(s, y(s) - x(s))\nabla s + \lambda Mw(t),$$

or

$$y(t) - x(t) = \lambda \int_a^b H(t,s)f(s, y(s) - x(s))\nabla s,$$

and hence

$$u(t) = \lambda \int_a^b H(t,s)f(s, u(s))\nabla s.$$

This shows that $u$ is a symmetric positive solution of the boundary value problem of (1.1) – (1.3). In addition, from (3.11) and (3.12), it follows that

$$\|u\| \geq \frac{\|y\|}{2} \geq \frac{r}{2}.$$

**3.2. Example.** Let T $= Z$. Consider the following boundary value problem

$$[\frac{100}{t^2+1}u^{\triangle}(t)]^{\nabla} + \lambda(be^u \cos^2 t - t^2) = 0, \ t \in (-3, 3), \tag{3.13}$$

$$25u(-3) - 5 \lim_{t \to -3^+} \frac{100}{t^2+1}u^{\triangle}(t) = \int_{-3}^{3} u(s)\cosh s \nabla s, \tag{3.14}$$

$$25u(3) + 5 \lim_{t \to 3^-} \frac{100}{t^2+1}u^{\triangle}(t) = \int_{-3}^{3} u(s)\cosh s \nabla s, \tag{3.15}$$

where $b > 0$, $\alpha = 25, \beta = 5, h_1(t) = h_2(t) = \cosh t, g(t) = \dfrac{100}{t^2+1}, f(t, u(t)) = be^u \cos^2 t - t^2$. It is obvious that $f$ satisfies the conditions $(C_2)$ and $(H_3)$.

Now we shall obtain the constants $M$ and $M_1$. Clearly, for all $(t, u) \in [-3, 3] \times [0, \infty)$, we get

$f(t, u) = be^u \cos^2 t - t^2 \geq -t^2 \geq -9$ and so we can choose the constant $M = 9$.

$$M_1 = \max_{(t,u)\in[-3,3]\times[0,r]} be^u \cos^2 t - t^2 + M = be^r + M.$$

It follows from a direct calculation that

$v_1 = v_2 = \displaystyle\int_{-3}^{3} h_1(s)\nabla s \cong 21.5, \mu = 2\alpha\beta + \alpha^2 \int_{-3}^{3} \frac{\Delta r}{g(r)} \cong 406.2,$

$D = (\beta + \alpha \int_{-3}^{3} \dfrac{\Delta r}{g(r)})^2 \cong 126.6, A = \mu + (\beta - K)v_1 - \beta v_2 \cong 56,87,$

$B_1 = \dfrac{K-\beta}{A} \cong 0.198, B_2 = \dfrac{\beta}{A} \cong 0.088,\ \gamma = 1 + B_1 v_1 + B_2 v_2 \cong 7.15,$

$C = \dfrac{6}{\mu\beta^2} D^2 \gamma \cong 67.71.$

Then by Theorem 3.1, we see that the boundary value problem $(3.13) - (3.15)$ has at least one symmetric positive solution $u$ such that $\|u\| \geq \frac{r}{2}$ for any $\lambda \in (0, \eta]$ where $\eta := \min\{\frac{r}{M_1\|w\|}, \frac{r}{2MC}\}$, $r$ is a given positive number and $w$ is the unique positive solution of the boundary value problem $[\frac{100}{t^2+1} u^{\triangle}(t)]^{\nabla} + 1 = 0$ with the boundary condition $(3.14) - (3.15)$.

## References


[1] M. Bohner, A. Peterson, Dynamic Equations on Time Scales: An Introduction with Applications, Birkhauser, Boston, Cambridge, MA (2001).

[2] M. Bohner, A. Peterson, Advances in Dynamic Equations on Time Scales, Birkhauser, Boston, Cambridge, MA (2003).

[3] A. Boucherif, Second-order boundary value problems with integral boundary conditions, Nonlinear Anal. 70 (2009) 364-371.

[4] J. R. Cannon, The solution of the heat equation subject to the specification of energy, Quart. Appl. Math. 21 (2) (1963) 155-160.

[5] R. Yu. Chegis, Numerical solution of a heat conduction problem with an integral boundary condition, Litovsk. Math. Sb. 24 (1984) 209-215.

[6] J. Henderson, H. B. Thompson, Multiple symmetric positive solutions for a second order boundary value problem, Proc. Amer. Math. Soc. 128 (2000) 2373-2379.

[7] N. I. Ionkin, Solution of a boundary value problem in heat conduction theory with nonlocal boundary conditions, Differential Equations 13 (1977) 294-304.

[8] F. Li, Y. Zhang, Multiple symmetric nonnegative solutions of second-order ordinary differential equations, Appl. Math. Lett. 17 (2004) 261-267.

[9] Y. Sun, Optimal existence criteria for symmetric positive solutions to a three-point boundary value problem, Nonlinear Anal. 66 (2007) 1051-1063.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 33 – 41

# On norm-preserving isomorphisms of $L^p(\mu, H)$

B.A. Güntürk[*†], B. Cengiz[‡] and M. Gürdal[§]

**Abstract**

Given an arbitrary positive measure space $(X, \mathcal{A}, \mu)$ and a Hilbert space $H$. In this article we give a new proof for the characterization theorem of the surjective linear isometries of the space $L^p(\mu, H)$ (for $1 \leq p < \infty$, $p \neq 2$) which is essentially different from the existing one, and depends on the $p$-projections of $L^p(\mu, H)$. We generalize the known characterization of the $p$-projections of $L^p(\mu, H)$ for $\sigma$-finite measure to the arbitrary case. They are proved to be the multiplication operations by the characteristic functions of the locally measurable sets, or that of the clopen (closed-open) subsets of the hyperstonean space the measure $\mu$ determines.

**Keywords:** Measure space, Bochner space, perfect measure, hyperstonean space, linear isometries.

*2000 AMS Classification:* 46E40, 28B05, 47B38.

*Received* : 23.12.2014 *Accepted* : 20.01.2015 *Doi* : 10.15672/HJMS.20164512488

## 1. Introduction

The isometric theory of Banach spaces still fascinates some mathematicians. One of the main problems in this area is to characterize the linear isometries on or between these spaces. It might seem to be relatively easy at first glance but, generally speaking, it is indeed very difficult a problem to solve. It would be very unrealistic to expect that there might be a complete solution of this problem. However, for some subclasses there has been a great progress in that direction. For instance, the surjective isometries of $C(X)$ or $L^p$ type classical Banach spaces this problem is solved completely, but the case of into isometries is still far from being settled.

Let $(X, \mathcal{A}, \mu)$ be a positive measure space and $H$ a Hilbert space. For any $1 \leq p < \infty$, $p \neq 2$, the Bochner space $L^p(X, \mathcal{A}, \mu; H)$ will be denoted by $L^p(\mu, H)$, if there is no

[*]Suleyman Demirel University, Department of Mathematics, 32260, Isparta, Turkey.
Email : banugunturk@hotmail.com
[†]Corresponding Author.
[‡]Başkent University, Faculty of Engineering, Bağlıca, Ankara, Turkey.
[§]Suleyman Demirel University, Department of Mathematics, 32260, Isparta, Turkey.

ambiguity about the underlying measurable space. For definitions and properties of these spaces we refer to [6].

A *regular set isomorphisms* on $\mathcal{A}$, defined modulo null sets, means a mapping on $\mathcal{A}$ to $\mathcal{A}$ with (i) $\varphi(A') = \varphi X \setminus \varphi A$ for every $A$ in $\mathcal{A}$, where $A'$ denotes the complement of $A$, (ii) $\varphi(\bigcup A_n) = \bigcup \varphi A_n$ for any sequence $\langle A_n \rangle$ in $\mathcal{A}$ mutually disjoint sets, and (iii) $\mu(\varphi A) = 0$ if, and only if, $\mu(A) = 0$. Any such mapping defines a function $\Phi$ on the set of measurable functions which we call the *induced* map. It is characterized by $\Phi(\chi_A e) = \chi_{\varphi A} e$, $A \in \mathcal{A}$, $e \in H$, where$\chi_A$ denotes the characteristic function of $A$ (see [7, pp. 453-454]).

The characterization of the surjective linear isometries of $L^p$ spaces was started by Banach [1] for the Lebesgue measure $\lambda$ on the closed interval $[0, 1]$. He proved that for every linear isometry $T$ of $L^p(\lambda)$, $1 \leq p < \infty$, $p \neq 2$, there exists a measurable function $\sigma$ of $[0, 1]$ such that for $f \in L^p(\lambda)$

$$(Tf)(x) = h(x)f(\sigma(x)) \text{ a.e. on } [0, 1].$$

If $\phi$ is the regular set isomorphism defined by $\phi(A) = \sigma^{-1}(A)$ on the Borel algebra of $[0, 1]$, then the above representation becomes

$$(1.1) \qquad (Tf)(x) = h(x)\Phi(f)(x)) \text{ a.e. on } [0, 1].$$

In [11], Lamperti proves that for any $\sigma$-finite measure space $(\Omega, \mathcal{A}, \mu)$, the linear isometries of $L^p(\mu)$ onto itself, $1 \leq p < \infty$, $p \neq 2$, are indeed of the above form (1) except that the isomorphism $\phi$ of the $\sigma$-algebra $\mathcal{A}$, need not be defined by a point mapping. Moreover, if the measure $\nu$ is defined by $\nu(A) = \mu[\phi^{-1}(A)]$, $A \in \mathcal{A}$, then

$$(1.2) \qquad |h(x)|^p = d\nu/d\mu \text{ a.e. on } \Omega.$$

In [3], Cambern generalizes this result to the Bochner spaces. He proves that if $(\Omega, \mathcal{A}, \mu)$ is a $\sigma$-finite measure space and $H$ is a separable Hilbert space, then for any linear isometry of $L^p(\mu, H)$ onto itself, $1 \leq p < \infty$, $p \neq 2$, in addition to the maps $h$ and $\Phi$ in Lamperti's characterization now there also exists a weakly measurable operator-valued function $U$ defined on $X$, where $U_x = U(x)$ is an isometry of $H$ onto itself for almost all $x \in X$, such that for $F \in L^p(\mu, H)$,

$$(1.3) \qquad (Tf)(x) = U_x(h(x)\Phi(F)(x)) \text{ a.e. on } X.$$

In [9], Greim and Jamison obtain the same representation for an arbitrary Hilbert space, but the measure is still $\sigma$-finite.

In [5] this characterization is extended to perfect measures. In view of the fact that any arbitrary measure space can be replaced by a perfect one without disturbing the $L^p$ spaces for $1 \leq p < \infty$, perhaps possibly enlarging the $L^\infty$ space [4], for Hilbert spaces the above result is then the most general result one can get for the surjective isometries.

The purpose of this article is two-fold. Our first goal is to obtain a complete description of the *p-projections* of $L^p(\mu, H)$ (which is known if $\mu$ is $\sigma$-finite). We will prove that a $p$-projection of $L^p(\mu, H)$ is the characteristic function of a *locally* measurable set (i.e. its intersection with every set of finite measure is measurable), and of a measurable set if $\mu$ is *decomposable,* in particular *perfect.* (For the definition *of a* decomposable measure see [10, p.317].) Our second goal is to use this result to give a second proof for the characterization of the isometries of $L^p(\mu, H)$ which is fundamentally different from the one given in [5], and we shall also demonstrate that in order to prove this characterization theorem one does not have to replace the given measure space by a perfect one, but this will be possible after we characterize the so-called pseudocharacteristic functions on the $\sigma$-rign of all $\sigma$-finite measurable sets.

## 2. Norm-Preserving Isomorphisms of $L^p(\mu, H)$

Let us recall that a compact Hausdorff space $X$ is called *extremally disconnected* if the closure of every open subset is open, and a nonnegative extended real-valued Borel measure $\mu$[¶] on $X$ is called *perfect* if

(i) every nonempty open set contains a clopen set with finite positive measure,
(ii) every nowhere dense Borel set has measure zero (equivalently, every closed set with empty interior has measure zero).

A *perfect measure space* $(X, \mathcal{B}, \mu)$ will mean that $X$ is an extremally disconnected Hausdorff space, $\mathcal{B}$ is the Borel algebra on $X$ and $\mu$ is a perfect measure. A hyperstonean space is an extremally disconnected compact Hausdorff space on which there is a perfect Borel measure.

Let $F$ be a Banach space, and $1 \leq p < \infty$, $p \neq 2$. A linear mapping $P$ on $F$ is said to be a *p-projection* if $P^2 = P$ and

$$\|x\|^p = \|Px\|^p + \|x - Px\|^p \text{ for all } x \in F.$$

The set $\mathbb{P}_p(F)$ of $p$-projections on $F$ is a complete Boolean algebra and its Stonean space is hyperstonean [2, pp.11, 25-26].

Given a measure space $(S, \mathcal{A}, \mu)$ and $1 \leq p < \infty$, $p \neq 2$. For any measurable set $A$, the mapping $f \longrightarrow f\chi_A, f \in L^p(\mu, H)$, is a $p$-projection on $L^p(\mu, H)$, and if $\mu$ is $\sigma$-finite the converse is also true [8, pp.124-126].

In this section we shall fix a perfect measure space $(\Omega, \mathcal{B}, \mu)$ which may be assumed to have the property that every locally null set is actually null and show that the $p$-projections on the Bochner space $L^p(\mu, H), 1 \leq p < \infty$, $p \neq 2$, are of the above form.

Any Borel subset $B$ of $\Omega$ is equivalent to a clopen subset of $\Omega$ in the sense that there exists a clopen subset $U$ of $\Omega$ such that $B \triangle U = (B \setminus U) \cup (U \setminus B)$ is locally null [2, p.31]. Thus any characteristic function $\chi_A$, with $A$ measurable, equals a.e. to the characteristic function of a clopen set.

**2.1. Theorem.** *For any clopen subset $B$ of $\Omega$, the function $\chi_B : f \longrightarrow f\chi_B$ is a $p$-projection on $L^p(\mu, H)$, and conversely, every $p$-projection on $L^p(\mu, H)$ is of this form.*

*Proof.* Property (i) of $\mu$, together with an application of Zorn's lemma, can be used to prove that there exists a disjoint family $\{\Omega_i : i \in I\}$ of clopen subsets of $\Omega$ with positive *finite* measure such that their union is dense in $\Omega$. Therefore, the closed set $\Omega \setminus \bigcup\limits_i \Omega_i$ has measure zero. From this it follows that

$$L^p(\mu, H) = \sum_i \oplus L^p(\Omega_i, H), \text{ ($p$-direct sum).}$$

Now let $P$ be a $p$-projection on $L^p(\mu, H)$. Then by a theorem in [2, p.20], for each $i \in I$ there exists a $p$-projection $P_i$ on the Banach *subspace* $L^p(\Omega_i, H)$ of $L^p(\mu, H)$ such that $P = \sum\limits_i \oplus P_i$ (direct sum), that is, $P(f) = \sum\limits_i P_i(f_i)$ for all $f \in L^p(\mu, H)$ where each $f_i = f\chi_{\Omega_i}$.

Since for each $i$, $\mu$ is finite on $\Omega_i$, $P_i = \chi_{B_i}$ for some clopen subset $B_i$ contained in $\Omega_i$. Hence $P = \sum\limits_i \oplus P_i = \sum\limits_i \oplus \chi_{B_i} = \chi_B$, where $B = cl(\cup B_i)$. This completes the proof. □

Throughout the section, $T$ will denote a fixed surjective linear isometry on $L^1(\mu, H)$.

**2.2. Theorem.** *(i) For each measurable set $A$, the mapping $P = T\chi_A T^{-1}$ is a $p$-projection on $L^p(\mu, H)$.*

[¶]All measures throughout this paper will be nonnegative.

*(ii) The mapping $\varphi$ defined on the Boolean algebra $\mathcal{K}(\Omega)$ by the equation $T\chi_A T^{-1} = \chi_{\varphi A}$ is an isomorphism of $\mathcal{K}(\Omega)$ onto itself. Moreover, for any sequence $\{A_n\}$ in $\mathcal{K}(\Omega)$, $\varphi(\bigvee_n A_n) = \bigvee_n \varphi(A_n)$ (i.e., $\varphi(cl(\bigcup_n A_n)) = cl(\bigcup_n \varphi(A_n))$.*

*Proof.* (i) Obviously $P^2 = P$ and since $T^{-1}$ is also an isometry and $\chi_A$ is a $p$-projection, for each $f$ in $L^p(\mu, H)$ we have

$$\begin{aligned}\|Pf\|^p + \|f - Pf\|^p &= \left\|T\chi_A T^{-1} f\right\|^p + \left\|TT^{-1}f - T\chi_A T^{-1} f\right\|^p \\ &= \left\|\chi_A T^{-1} f\right\|^p \left\|T^{-1}f - \chi_A T^{-1} f\right\|^p \\ &= \left\|T^{-1} f\right\|^p \\ &= \|f\|^p .\end{aligned}$$

This completes the proof of (i).

(ii) Let $B$ be a clopen subset of $\Omega$, then, $T^{-1}\chi_B T$, being a $p$-projection, equals $\chi_A$ for some $A$ in $\mathcal{K}(\Omega)$. Therefore $\chi_B T = T\chi_A$ which means that $B = \varphi A$, hence $\varphi$ maps $\mathcal{K}(\Omega)$ onto itself.

Now let $A, B$ be any two sets in $\mathcal{K}(\Omega)$. Then

$$T\chi_{A\cap B} = T(\chi_A \chi_B) = \chi_{\varphi B} T\chi_A = \chi_{\varphi A}\chi_{\varphi B} T = \chi_{\varphi A \cap \varphi B} T$$

which implies that $\varphi(A \cap B) = \varphi A \cap \varphi B$.

Next we show that $\varphi(A \cup B) = \varphi A \cup \varphi B$. First let us assume that $A \cap B = \emptyset$. Then,

$$T\chi_{A\cup B} = T(\chi_A + \chi_B) = (\chi_{\varphi A} + \chi_{\varphi B})T = \chi_{\varphi A \cup \varphi B} T$$

from which it follows that $\varphi(A \cup B) = \varphi A \cup \varphi B$ as claimed.

Now let $A, B$ be any two sets in $\mathcal{K}(\Omega)$, then by the preceding result,

$$\begin{aligned}\varphi(A \cup B) &= [\varphi(A \setminus B) \cup \varphi(A \cap B)] \cup [\varphi(A \cap B) \cup \varphi(B \setminus A)] \\ &= \varphi A \cup \varphi B.\end{aligned}$$

For the last claim in the theorem, let $\{A_n\}$ be a sequence in $\mathcal{K}(\Omega)$. first let us assume that they are mutually disjoint. Let $A = \bigcup_n A_n$ and for each $n$, $B_n = \bigcup_{i\geq n+1} A_i$. Then $\bigcap_n B_n = \emptyset$ and for $f \in L^p(\mu, H)$, for each $n$ we have

$$\left\|\chi_A^{(t)} f(t) - \sum_{i=1}^{n} \chi_{A_i}^{(t)} f(t)\right\|^p = \chi_{B_n}(t) \left\|f(t)\right\|^p \leq \left\|f(t)\right\|^p$$

for all $t$ in $\Omega$, and since $\|f(t)\|^p$ is integrable, by the dominated convergence theorem [10, p.172] we obtain

$$\lim_n \left\|\chi_A - \sum_{i=1}^{n} \chi_{A_i} f\right\|^p = \lim_n \int_\Omega \chi_{B_n}(.) \left\|f(.)\right\|^p d\mu = 0$$

which means that $\chi_A f = \sum_{i=1}^{\infty} \chi_{A_i} f$ in $L^p(\mu, H)$.

From this, it follows that the series $\sum_{i=1} \chi_{A_i}$ converges pointwise to $\chi_A$ (as operators on $L^p(\mu, H)$). Similarly, the series $\sum_i \varphi A_i$ converges pointwise to the operator $\chi_{\bigcup_i A_i}$. Therefore, since each open set and its closure differ by a null set,

$$\begin{aligned}T\chi_{clA} &= T\chi_A = T(\sum_n \chi_{A_n}) = \sum_n T\chi_{A_n} \\ &= \sum_n (\chi_{\varphi A_n} T) = (\chi_{\bigcup_n \varphi A_n})T = (\chi_{cl \bigcup_n \varphi A_n})T\end{aligned}$$

which implies that

$$\varphi(clA) = cl(\bigcup_n \varphi A_n). \tag{2.1}$$

Now let $\{A_n\}$ be any sequence in $\mathcal{K}(\Omega)$ and define $C_1 = A_1$, $C_n = A_n \setminus \bigcup_{i=1}^{n-1}$ for all $n \geq 2$. Then, $C_n$ 's are mutually disjoint and $\bigcup_n C_n = \bigcup_n A_n$, and since $\varphi$ maps disjoint sets to disjoint set and $C_n \subset A_n$, $\forall n$, it follows that

$$\bigcup_n \varphi(C_n) \subset \bigcup_n \varphi(A_n) \tag{2.2}$$

On the other hand, since $A_n \subset \bigcup_{i=1}^{n} C_i$ we have $\varphi A_n \subset \bigcup_{i=1}^{n} \varphi C_i \subset \bigcup_{i=1}^{\infty} \varphi C_i$ $\forall n$, therefore $\bigcup_n \varphi A_n \subset \bigcup_n \varphi C_n$. Combining this inclusion with (5) we obtain

$$\bigcup_{n=1}^{\infty} \varphi A_n = \bigcup_{n=1}^{\infty} \varphi C_n. \tag{2.3}$$

Hence from (4) and (6) one obtains

$$\begin{aligned} \varphi(clA) &= \varphi(cl\bigcup_n A_n) = \varphi(cl\bigcup_n C_n) = cl(\bigcup_n \varphi C_n) \\ &= cl(\bigcup_n \varphi C_n) \end{aligned}$$

as claimed. This completes the proof of the theorem. □

An element $f$ of $L^p(\mu, H)$ is an equivalence class rather than a function. The support of each function in this class is equivalent to the same clopen set which we shall call the support of $f$ and denote it by *supp*$(f)$.

The following lemmas will be needed for our next theorem.

**2.3. Lemma.** *For each $f \in L^p(\mu, H)$, $\varphi(supp\ f) = supp(Tf)$.*

*Proof.* Fix $f \in L^p(\mu, H)$, and let $A =$*supp*$(f)$, $S =$*supp*$(Tf)$. Since $Tf = T(\chi_A f) = \chi_{\varphi A} Tf$ we conclude that $S \subset \varphi A$. Now let $B \subset \varphi A \setminus S$ be any clopen set with finite measure, and let $u \in H$, $u \neq 0$. Then there exists a function $g$ in $L^p(\mu, H)$ such that $Tg = \chi_B u$. Since $\chi_B u$ and $Tf$ have disjoint supports and $T^{-1}$ maps functions with disjoint supports to functions with disjoint supports [3, p.12] $g$ and $f$ have disjoint supports. Thus, since $A$ and *supp*$(g)$ are disjoint

$$0 = T(\chi_A g) = \chi_{\varphi A} Tg = \chi_{\varphi A} \chi_B u = \chi_B u$$

which means that $B = \emptyset$. Hence $S = \varphi A$, proving our lemma. □

**2.4. Corollary.** *Each $Y_i = \varphi \Omega_i$ is $\sigma$-finite.*

**2.5. Lemma.** *For each $i \in I$, let $Y_i = \varphi \Omega_i$, and $Y = \bigcup_i Y_i$. Then*

*(i) $Y'$ is a closed null set and $T$ maps $L^p(\Omega_i, H)$ onto $L^p(Y_i, H)$;*
*(ii) $\varphi$ maps clopen sets in $\Omega_i$ onto clopen sets in $Y_i$.*

*Proof.* For $f \in L^p(\Omega_i, H)$,

$$supp(Tf) = \varphi(supp(f)) \subset \varphi(\Omega_i) = Y_i$$

which shows that for each $i$, $T$ maps $L^p(\Omega_i, H)$ into $L^p(Y_i, H)$.

Since the sets $Y_i$ are mutually disjoint,

$$L^p(Y,H) = \sum_i \oplus L^p(Y_i,H).$$

Thus, $T$ maps $L^p(\mu,H) = \sum_i \oplus L^p(\Omega_i,H)$ onto $L^p(Y,H)$, which implies that $Y'$ is a closed null set, and that $T$ maps $L^p(\Omega_i,H)$ onto $L^p(Y_i,H)$ for all $i \in I$. This completes the proof of (i).

For (ii) we fix $i \in I$ and show that each clopen set $B_i$ in $Y_i$ is the image under $\varphi$ of a clopen set in $\Omega_i$. Fix a clopen set $B \subset Y_i$ and let $u \in H$, $u \neq 0$. Then, there exists an $f$ in $L^p(\Omega_i,H)$ such that $Tf = \chi_B u$. By Lemma 2.3,

$$\varphi(supp(f)) = supp(Tf) = supp\chi_B u = B,$$

proving (ii). □

We can show very easily that, for each $i \in I$, the mapping $\mu\circ\varphi^{-1}$ is countably additive on the algebra of all clopen subsets of $Y_i$; that is, for any sequence $\{B_n\}$ of mutually disjoint clopen subsets of $Y_i$ whose union is also a clopen subset of $Y_i$,

$$\mu\circ\varphi^{-1}(\bigcup B_n) = \sum_n \mu\circ\varphi^{-1}(B_n)$$

Thus, since $Y_i$ is $\sigma$-finite, $\mu\circ\varphi^{-1}$ extends uniquely to a perfect regular Borel measure on the Borel algebra of $Y_i$ [12, p. 120]. We will denote this extension also by $\mu\circ\varphi^{-1}$. Then we define a measure $\nu$ and $\mathcal{A}$ by

$$\nu(A) = \sum_i \mu\circ\varphi^{-1}(A\cap\Omega_i),\ \ A \in \mathcal{A}.$$

We have completed almost all but few details of the proof of the following theorem:

**2.6. Theorem.** *There exists a locally strongly measurable operator-valued function $U$ and a measurable scalar-valued function $h$ on $\Omega$ such that for each $x \in \Omega$, $U_x = U(x)$ is an isometry of $H$ onto itself and that for every $f$ in $L^p(\mu,H)$,*

$$(Tf)(x) = U_x(h(x)\Phi(f)(x)) \ \textit{a.e. on } \Omega$$

*where $\Phi$ is the isomorphism of $L^p(\mu,H)$ onto itself induced by $\varphi$. Moreover,*

$$|h|^p = \frac{d\nu}{d\mu},$$

*(the Radon-Nikodým derivative). Conversely, every mapping of the above form is an isometry of $L^p(\mu,H)$ onto itself.*

*Proof.* For each $i \in I$, by a theorem of Greim and Jamison [9, p.513], there exists a strongly measurable function $U^{(i)}$ from $Y_i$ into the set of linear surjective isometries of $H$ and a scalar function $h_i$ on $Y_i$ such that for every $f$ in $L^p(\Omega_i,H)$,

$$(Tf)(y) = U^{(i)}_y(h_i(y)\Phi(f)(y)) \text{ a.e. on } Y_i$$

and furthermore,

$$|h_i|^p = \frac{d(\mu\circ\varphi^{-1})}{d\mu_i}$$

where $\mu_i$ denotes the restriction of $\mu$ to $\mathcal{A}(Y_i)$.

Each measurable set $A$ is equivalent to a unique clopen set $A_c$, and so, we may extend $\varphi$ to a regular set isomorphism from $\mathcal{A}$ onto itself, defined modulo null sets, by the equation

$$\varphi(A) = \varphi(A_c),\ \ A \in \mathcal{A}$$

Since for each $i \in I$, $\varphi$ maps $\mathcal{A}(\Omega_i)$ (the trace of $\mathcal{A}$ on $\Omega_i$), isomorphically (modulo null sets) onto $\mathcal{A}(Y_i)$, the induced mapping $\Phi$ is an isomorphism of $L^p(\Omega_i, H)$ onto $L^p(Y_i, H)$.

Now we let $U = \sum_i U^{(i)}$, $h = \sum_i h_i$ on $Y$, and on $Y'$ we let $U_y = I$ (the identity operator on $H$) and $h(y) = 1$. Obviously $U$ is locally strongly measurable, $h$ is measurable, and for each $f$ in $L^p(\mu, H)$,

$$(Tf)(x) = U_x(h(x)\Phi(f)(x))$$

a.e. on $\Omega$, and furthermore,

$$|h|^p = \frac{d\nu}{d\mu}.$$

This completes the proof of our theorem. □

## 3. The Case of Nonperfect Measures

In the proof of Theorem 1, the isomorphism $\varphi$ on $\mathcal{K}(\Omega)$ played a crucial role, and it was constructed by the help of the $p$-projections on $L^p(\mu, H)$ which were characterized as the characteristic functions of the sets in $\mathcal{K}(\Omega)$.

In this section, instead of a perfect measure space we will work with an *arbitrary* measure space $(X, \mathcal{A}, \mu)$, and prove that a similar construction is possible.

Let $\mathbb{A}_\sigma$ denote the Boolean ring of $\sigma$-finite measurable sets, (two sets $A, B$ are regarded to be the same if their symmetric difference $A \triangle B = (A \setminus B) \cup (B \setminus A)$ is locally null, i.e., its intersection with every set in $\mathcal{A}_\sigma$ has measure zero), with the ring operations $U + V = U \triangle V$ and $U.V = U \cap V$.

Following [2] we shall call a function $\gamma : \mathcal{A}_\sigma \longrightarrow \mathcal{A}_\sigma$ a *pseudocharacteristic function* (PCF) if $\gamma(AB) = A\gamma(B) = B\gamma(A)$ for all $A, B$ in $\mathcal{A}_\sigma$. Clearly, for each $A \in \mathcal{A}$, the mapping $\gamma_A : B \longrightarrow B \cap A$, $B \in \mathcal{A}_\sigma$ is a PCF on $\mathcal{A}_\sigma$, but examples show that the converse is not always true. However, as we shall see soon that something very close to this is true.

It is known that if $1 \leq p < \infty$, $p \neq 2$, and $\gamma$ is a PCF on $\mathcal{A}_\sigma$ then the mapping $P : f \longrightarrow f\chi_{\gamma(S(f))}$ is a $p$-projection on the $L^p$-space $L^p(\mu)$ of scalar-valued measurable functions, where $s(f) = supp(f)$; and conversely, every $p$-projection on $L^p(\mu)$ is of this form [2, p.58]. Greim [8] generalizes this result to $L^p(\mu, H)$ for $\sigma$*-finite* $\mu$, that is, he proves that a mapping $P$ on $L^p(\mu, H)$ is a $p$-projection if and only if $P = \chi_A$ for some measurable set $A$.

In this section the above mentioned representations for the $p$-projections on $L^p(\mu)$ and $L^p(\mu, H)$ (with $\mu$ $\sigma$-finite), will be generalized to the $p$-projections on $L^p(\mu, H)$ for *arbitrary* $\mu$. For this we will need the following:

**3.1. Proposition.** *Let $\gamma$ be a PCF on $\mathcal{A}_\sigma$ (defined modulo null sets). Then*

*(i) $\gamma(C) \subset \gamma(B) \subset B$ for all $B, C$ in $\mathcal{A}_\sigma$ and $C \subset B$. In particular $\gamma(\emptyset) = \emptyset$,*

*(ii) for any sequence $\{B_n\}$ in $\mathcal{A}_\sigma$*

$$\gamma(\bigcup_n B_n) = \bigcup_n \gamma(B_n) \text{ and } \gamma(\bigcap_n B_n) = \bigcap_n \gamma(B_n),$$

*consequently $\gamma$ maps disjoint sets to disjoint sets, and*

*(iii) $\gamma(B \setminus C) = \gamma(B) \setminus \gamma(C)$ for any $B$, $C$ in $\mathcal{A}_\sigma$.*

*Proof.* The proof of (i) is trivial.

(ii) Let $\{B_n\}$ be a sequence in $\mathcal{A}_\sigma$, $B = \bigcup_n B_n$. Then, $\gamma(B_n) = \gamma(B) \cap B_n$ for all $n$, therefore,

$$\gamma(\bigcap_n B_n) = \gamma(B) \cap (\bigcap_n B_n) = \bigcap_n [\gamma(B) \cap B_n] = \bigcap_n \gamma(B_n).$$

(iii) Let $A = \gamma(B \cup C)$, then since $E \subset F$ implies that $\gamma(E) = E \cap \gamma(F)$,

$$\begin{aligned}\gamma(B \setminus C) &= \gamma((B \setminus C) \cap (B \cap C)) = (B \setminus C) \cap \gamma(B \cup C) \\ &= (B \setminus C) \cap A = (B \cap A) \setminus (C \cap A) = \gamma(B) \setminus \gamma(C).\end{aligned}$$

This completes the proof. □

We will call two measurable sets *$\mu$-disjoint* if their intersection has measure zero. One can apply Zorn's lemma to show that in any measure space there exists a maximal family of mutually $\mu$-disjoint measurable sets with strictly positive finite measure. Any such family will be called a *$\mu$-decomposition* for the measure space.

Let $\{F_i : i \in I\}$ be a *$\mu$-decomposition* for the measure space $(X, \mathcal{A}, \mu)$. Then, it can be shown very easily that every $\sigma$-finite measurable set is contained a.e. in the union of a countable subfamily of $\{F_i : i \in I\}$.

Given an arbitrary measure space $(X, \mathcal{A}, \mu)$. We may and will assume that the measure space is *complete* in the sense that every subset of a null set is measurable (i.e., if $A \in \mathcal{A}$, $\mu(A) = 0$ then every subset of $A$ is also in $\mathcal{A}$).

**3.2. Theorem.** *(i) Let $\gamma$ be a PCF on $\mathcal{A}_\sigma$. Then there exists a locally measurable subset $Y$ of $X$ such that*

$$\gamma(B) = B \cap Y \text{ for all } B \in \mathcal{A}_\sigma,$$

*(ii) For each locally measurable subset $Y$ of $X$ the mapping $E_\gamma$ defined by*

$$E_\gamma(f) = f\chi_{\gamma(S(f))},\ f \in L^p(\mu, H)$$

*is an p-projection on $L^p(\mu, H)$, where $S(f)$ =supp$(f)$ and $\gamma(B) = B \cap Y$; and conversely, every p-projection on $L^p(\mu, H)$ is of this form.*

*Proof.* Let $\{F_i : i \in I\}$ be a $\mu$-decomposition of the measure space $(X, \mathcal{A}, \mu)$, and let $B$ be a $\sigma$-finite set. Then, there exist indices $i_1, i_2, \ldots$ in $I$ such that $B \subset F_{i_1} \cup F_{i_2} \cup \ldots$a.e. Let $X_1 = \bigcup_k F_{i_k}$, $B_1 = B \cap X_1$, $B_0 = B \cap (X \setminus X_1)$, and for each $i \in I$ let $Y_i = \gamma(F_i)$, $Y = \bigcup_i Y_i$ and $Z = \gamma(X_1)$. Since $B_0$ is null, $\gamma(B_0)$ is null. Therefore,

$$\begin{aligned}\gamma(B) &= \gamma(B_1) \cup \gamma(B_0) = \gamma(B_1 \cap X_1) \cup \gamma(B_0) \\ &= B_1 \cap Z \text{ a.e.} \\ &= B \cap Z \text{ a.e.}\end{aligned}$$

Now, for each $i$, $B \cap (Y_i \setminus Z) = B \cap \gamma(F_i \setminus X_1) \subset B \cap (F_i \setminus X_1) \subset B \cap \gamma(X \setminus X_1) = B_0$, and taking union over $i$, we obtain $B \cap (Y \setminus Z) \subset B_0$, so $B \cap (Y \setminus Z)$ is null, and since $Y \supset Z$ we have

$$\begin{aligned}\gamma(B) &= (B \cap Z) \cup (B \cap (Y \setminus Z)) \text{ a.e.} \\ &= B \cap Y \text{ a.e.}\end{aligned}$$

which proves (i).

(ii) Now let $E$ be a $p$-projection on $L^p(\mu)$. Then $\exists$ a PCF $\gamma$ on $\mathcal{A}_\sigma$ such that

$$E(f) = f\chi_{\gamma(S(f))},\ f \in L^1(\mu) \tag{3.1}$$

and then by the first part of the theorem, there is a locally measurable set $Y$ such that $\gamma(B) = B \cap Y$ for all $B \in \mathcal{A}_\sigma$.

Now, (7) becomes

$$E(f) = f\chi_{\gamma(S(f))} = f\chi_{S(f)\cap Y} = f\chi_Y \text{ for all } f \in L^1(\mu).$$

Generalization of this result to the Bochner space $L^p(\mu, H)$ is routine. This completes the proof of the theorem. □

**3.3. Corollary.** *For each $L^p$-projection $P$ there exists a locally measurable set $Y$ such that*

(3.2) $\quad P(f) = f\chi_Y,\ f \in L^p(\mu, H),$

*and conversely, for each locally measurable set $Y$, the mapping $P$ defined by* (8) *is an $L^p$-projection.*

Clearly the correspondence between the $L^p$-projections and the locally measurable sets is one-to-one modulo locally null sets.

Now, as in Section 2, using the above characterization of the $L^p$-projections we can define a set isomorphisms $\varphi$ from the $\sigma$-algebra $\mathcal{M}_\iota$ of all locally measurable sets onto itself defined modulo locally null sets.

**3.4. Remark.** A natural question is whether or not the characterization obtained for the linear isometries of the Bochner space $L^p(\mu, H)$ onto itself, $1 \leq p < \infty$, $p \neq 2$, holds, should a Banach space replace $H$ as the range space. In general, the answer is in the negative; however, Sourour [13, p.31] was able to replace $H$ by a suitable Banach space.

## Acknowledgments

This paper is supported by Suleyman Demirel University (SDU-BAP) with Project 3464-D1-13.

## References

[1] Banach, S. *Théoriedes opération linéaires*, Monografje Matematyczne Warsaw, 1932.

[2] Behrends, E.et al, *$L^p$ structure in real Banach spaces*, Lecture notes in Mathematics, 613*,* Springer-Verlag, 1977.

[3] Cambern, M., *The isometries of $L^p(X, K)$*, Pacific J. Math., **55**, 9-17, 1974.

[4] Cengiz, B., *On the duals of Lebesgue-Bochner $L^p$ spaces*, Proc. Amer. Math. Soc., **114**, 923-926, 1992.

[5] Cengiz, B., *The isometries of the Bochner space $L^p(\mu, H)$*, Turk J. Math., **23**, 389-399, 1999.

[6] Diestel, J. and Uhl, J. Jr., *Vector Measures, Math.* Surveys Monogr. No.15, Amer. Math. Soc., Providence, RI, 1977.

[7] Doob, J.L., *Stochastic Processes*, Wiley and Sons, New York, 1953.

[8] Greim, P., *Hilbert Spaces have the Banach Stone for Bochner spaces*, Bull. Austral. Math Soc, 27, 121-128, 1983

[9] Greim, P. And Jamison, J.E., *Hilbert spaces have the strong Banach-Stone property for Bochner spaces*, Math. Z., **196**(4), 511-515, 1987.

[10] Hewitt, E. and Karl Stromberg, 1965, *Real and Abstract Analysis*, Springer- Verlag, New-York.

[11] Lamperti, J., *On the isometries of certain function spaces*, Pacific J. Math., **8**, 459-466, 1958.

[12] Lacey, H. E., *The Isometric Theory of Classical Banach Spaces*, Berlin- Heidelberg, New-York, 1974.

[13] Sourour, A.R., *The isometries of $L^p(\Omega, X)$*, J. Funct. Anal., **30**, 276-285, 1978.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 43 – 48

# On a conjecture of double inequality for the tangent function

Mansour Mahmoud *

**Abstract**

In this paper, we prove the conjecture posed by J.-L. Zhao, Q.-M. Luo, B.-N. Guo and F. Qi (Remarks on inequalities for the tangent function, Hacettepe J. Math. Stat., 41, no. 4, 499-506, 2012) about a sharp double inequality of the tangent function, which is a generalization of the Becker-Stark inequality. Also, the new double inequality is compared with the double inequality presented in [3]

**Keywords:** Tangent function, L'Hôpital's rule, inequality, best constants, concavity.

*2000 AMS Classification:* 26D05, 26D15.

*Received* : 17.04.2014 *Accepted* : 29.01.2015 *Doi* : 10.15672/HJMS.20164512486

## 1. Introduction

In 1955, Stečkin [14] obtained the following one-side inequality for the tangent function

$$\frac{\tan x}{x} > \frac{4/\pi}{\pi - 2x}, \qquad 0 < x < \pi/2 \tag{1.1}$$

where the constant $4/\pi$ is the best possible.

Later in 1978, Becker and Stark [2] presented the following double inequality

$$\frac{8}{\pi^2 - 4x^2} < \frac{\tan x}{x} < \frac{\pi^2}{\pi^2 - 4x^2}, \qquad 0 < x < \pi/2 \tag{1.2}$$

which is a generalization of the Stečkin's inequality (1.1).

*King Abdulaziz University, Faculty of Science, Mathematics Department, P. O. Box 80203, Jeddah 21589, Saudi Arabia.
Email: mansour@mans.edu.eg
Permanent address: Mansour Mahmoud , Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.

In 2003, C.-P. Chen and F. Qi [3] established a double inequality for remainder $r_n(x) = \tan x - S_n(x)$, where $S_n(x)$ is the $n^{th}$ partial sum of the power series of $\tan x$. Their double inequality can be reformulated as [16]:

**1.1. Theorem.** *For $0 < x < \pi/2$ and $n \in \mathbb{N}$, we have*

$$\frac{2^{2(n+1)}(2^{2(n+1)}-1)|B_{2n+2}|}{(2n+2)!}x^{2n}\tan x < \tan x - S_n(x) < \left(\frac{2}{\pi}\right)^{2n} x^{2n}\tan x, \tag{1.3}$$

*where*

$$S_n(x) = \sum_{i=1}^{n} \frac{2^{2i}(2^{2i}-1)|B_i|}{(2i)!}x^{2i-1} \tag{1.4}$$

*and $B_j$'s are the Bernoulli numbers.*

The inequality (1.3) for $n = 1$ and $0 < x < \frac{3}{2}\sqrt{\frac{5(\pi^2-8)}{38}}$ will give us a refinement of the left-hand side of the Becker-Stark inequality (1.2). Also, the inequality (1.3) for $n = 2$ is better than the Djokovic inequality [8]

$$x + \frac{1}{3}x^3 < \tan x < x + \frac{4}{9}x^3, \qquad 0 < x < \pi/6.$$

In 2010, Zhu and Hua [17] established the following general refinement of the Becker-Stark inequality

**1.2. Theorem.** *Let $0 < x < \pi/2$ and a natural number $n \geq 0$. Then*

$$\frac{P_{2n}(x) + \alpha_n x^{2n+2}}{\pi^2 - 4x^2} < \frac{\tan x}{x} < \frac{P_{2n}(x) + \beta_n x^{2n+2}}{\pi^2 - 4x^2}, \tag{1.5}$$

*where $P_{2n}(x) = \sum_{i=0}^{n} a_i x^{2i}$ and*

$$a_i = \pi^2 \mid B_{2i+2} \mid \frac{2^{2(i+1)}(2^{2(i+1)}-1)}{(2i+2)!} - 4 \mid B_{2i} \mid \frac{2^{2i}(2^{2i}-1)}{(2i)!}, \qquad i = 0, 1, 2, \dots .$$

*Furthermore, $\alpha_n = \frac{8 - P_{2n}(\pi/2)}{(\pi/2)^{2n+2}}$ and $\beta_n = \alpha_{n+1}$ are the best constants in (1.5).*

In 2012, Zhao, Luo, Guo and Qi [16] showed that the double inequalities (1.3) and (1.5) are not included in each other, reorganized the proof of (1.3) by using the usual definition of Bernoulli numbers and corrected some errors on [12]. Moreover, they propose a sharp double inequality as a conjecture. In this paper we will prove this conjecture. Further interesting generalizations and applications about inequalities of the tangent function can be found in [4]-[6], [9], [10], [15], [18]-[20] and the references therein.

In our present investigation, we will apply the following monotone form of L'Hôpital's rule [1] (see also, [7], [11], [13] ).

**1.3. Theorem.** *Let $-\infty < a < b < \infty$, and let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, differentiable on $(a, b)$. Let $g'(x) \neq 0$ on $(a, b)$. If $f'(x)/g'(x)$ is increasing (decreasing) on $(a, b)$, then so are*

$$\frac{f(x)-f(a)}{g(x)-g(a)} \quad \text{and} \quad \frac{f(x)-f(b)}{g(x)-g(b)}. \tag{1.6}$$

*If $f'(x)/g'(x)$ is strictly monotone, then the monotonicity in the conclusion is also strict.*

## 2. Main Results.

Consider the following two functions for $n \in \mathbb{N}$ and $x \in \left(0, \frac{\pi}{2}\right)$

$$M_n(x) = \frac{\tan x}{S_n(x)}\left[1 - \left(\frac{2x}{\pi}\right)^{2n}\right] \tag{2.1}$$

and

$$h_n(x) = \frac{\tan x - S_n(x)}{x^{2n}\tan x}. \tag{2.2}$$

Then

$$\frac{1}{M_n(x)} = \frac{1 - x^{2n}h_n(x)}{1 - (2/\pi)^{2n}x^{2n}}. \tag{2.3}$$

Let

$$f_n(x) = \begin{cases} 1 - x^{2n}h_n(x) & 0 < x \le \pi/2, \\ 1 & x = 0 \end{cases}$$

and

$$g_n(x) = 1 - (2/\pi)^{2n}x^{2n}, \qquad 0 \le x \le \pi/2.$$

The functions $f_n, g_n : \left[0, \frac{\pi}{2}\right] \to \mathbb{R}$ are continuous on $\left[0, \frac{\pi}{2}\right]$, differentiable on $\left(0, \frac{\pi}{2}\right)$. Also, $g_n'(x) = -2n(2/\pi)^{2n}x^{2n-1} \neq 0$ on $\left(0, \frac{\pi}{2}\right)$. Now, consider the function

$$G_n(x) = \frac{f_n'(x)}{g_n'(x)} = (\pi/2)^{2n}\left[\frac{x}{2n}h_n'(x) + h_n(x)\right],$$

then

$$G_n'(x) = \frac{(\pi/2)^{2n}}{2n}\left[(2n+1)h_n'(x) + xh_n''(x)\right].$$

But the function $h_n(x)$ is absolutely monotonic on $\left(0, \frac{\pi}{2}\right)$ [16], that is

$$(h_n(x))^{(i)} \ge 0, \qquad \forall\, i \in \mathbb{N};\ x \in \left(0, \frac{\pi}{2}\right).$$

Then

$$G_n'(x) > 0, \qquad x \in \left(0, \frac{\pi}{2}\right)$$

and hence the function $G_n = \frac{f_n'}{g_n'}$ is increasing function on $\left(0, \frac{\pi}{2}\right)$. Using Theorem 1.3, we get that

$$\frac{f_n(x) - f_n(\pi/2)}{g_n(x) - g_n(\pi/2)}$$

is also increasing on $\left(0, \frac{\pi}{2}\right)$. But $f_n(\pi/2) = g_n(\pi/2) = 0$ and hence $\frac{1}{M_n(x)} = \frac{f_n(x)}{g_n(x)}$ is increasing on $\left(0, \frac{\pi}{2}\right)$. Then $M_n(x)$ is decreasing on $\left(0, \frac{\pi}{2}\right)$. Then

$$\lim_{x \to \frac{\pi}{2}^-} M_n(x) < M_n(x) < \lim_{x \to 0^+} M_n(x).$$

Using

$$\lim_{x \to 0^+} M_n(x) = \lim_{x \to 0^+}\frac{\tan x}{S_n(x)} = \lim_{x \to 0^+}\frac{\tan x/x}{1 + \sum_{i=2}^{n}\frac{2^{2i}(2^{2i}-1)|B_i|}{(2i)!}x^{2(i-1)}} = 1$$

and

$$\lim_{x \to \frac{\pi}{2}^-} M_n(x) = \frac{1}{S_n(\pi/2)}\lim_{x \to \frac{\pi}{2}^-}\frac{1 - \left(\frac{2x}{\pi}\right)^{2n}}{\cot x} = \frac{4n/\pi}{S_n(\pi/2)},$$

we get the following result

**2.1. Theorem** ([16], Conjecture 1)**.** *For $0 < x < \pi/2$ and $n \in \mathbb{N}$, we have*

$$\frac{4n/\pi}{S_n(\pi/2)} < \frac{\tan x}{S_n(x)}\left[1-\left(\frac{2x}{\pi}\right)^{2n}\right] < 1. \tag{2.4}$$

*Furthermore,* 1 *and* $\frac{4n/\pi}{S_n(\pi/2)}$ *are the best possible constants in (2.4).*

**2.2. Remark.** If we set $n = 1$ in the inequality (2.4), then we obtain the inequality (1.2) and hence the inequality (2.4) is an extension of Becker-Stark inequality (1.2).

Now we will study the concavity of the function $M_n(x)$. Let us recall that, a function $\varphi$ is concave if every chord lies below the graph of $\varphi$. Let $y_n(x)$ be the line segment with the endpoints $(0, 1)$ and $\left(\frac{\pi}{2}, \frac{4n/\pi}{S_n(\pi/2)}\right)$. Then

$$y_n(x) = \left[\frac{4n/\pi}{S_n(\pi/2)} - 1\right]\frac{2x}{\pi} + 1, \qquad 0 \leq x \leq \pi/2$$

and let

$$H_n(x) = \begin{cases} 1 & x = 0, \\ M_n(x) & 0 < x < \pi/2, \\ \frac{4n/\pi}{S_n(\pi/2)} & x = \pi/2. \end{cases}$$

The functions $H_n, y_n : \left[0, \frac{\pi}{2}\right] \to \mathbb{R}$ are continuous on $\left[0, \frac{\pi}{2}\right]$, differentiable on $\left(0, \frac{\pi}{2}\right)$. Also, $y_n'(x) = \left[\frac{4n/\pi}{S_n(\pi/2)} - 1\right] \neq 0\ \forall x$. Now consider the function

$$T_n(x) = \frac{H_n'(x)}{y_n'(x)} = \frac{\pi}{2\left(\frac{4n/\pi}{S_n(\pi/2)} - 1\right)} H_n'(x).$$

If we assume that $H_n''(x) > 0$, then we get $T_n(x)$ is decreasing function. Using Theorem 1.3, we get that

$$F(x) = \frac{H_n(x) - H_n(0)}{y_n(x) - y_n(0)} = \frac{H_n(x) - 1}{y_n(x) - 1}$$

is also decreasing function on $\left(0, \frac{\pi}{2}\right)$. But

$$\lim_{x\to 0^+} F(x) = \lim_{x\to 0^+} \frac{\pi}{2\left(\frac{4n/\pi}{S_n(\pi/2)} - 1\right)} H_n'(x) = 0$$

and

$$\lim_{x\to\frac{\pi}{2}^-} F(x) = \frac{H_n(\pi/2) - 1}{y_n(\pi/2) - 1} = 1,$$

which is a contradiction since $F(x)$ is decreasing. Then we get the following result

**2.3. Lemma** ([16], Conjecture 1)**.** *For $0 < x < \pi/2$ and $n \in \mathbb{N}$, the function*

$$M_n(x) = \frac{\tan x}{S_n(x)}\left[1-\left(\frac{2x}{\pi}\right)^{2n}\right] \tag{2.5}$$

*is concave.*

## 3. Comparison of Theorems 1.1 and 2.1.

The inequalities in Theorems 1.1 and 1.2 are not included in each other [16]. Now, we will compare the Theorems 1.1 and 2.1 .

The inequality (1.3) can be rewritten in the form

$$\frac{1}{1-\left(\frac{2^{2(n+1)}(2^{2(n+1)}-1)|B_{2(n+1)}|}{(2n+2)!}\right)x^{2n}} < \frac{\tan x}{S_n(x)} < \frac{1}{1-\left(\frac{2}{\pi}\right)^{2n}x^{2n}} \tag{3.1}$$

and the inequality (2.4) can be rewritten in the form

$$\frac{4n}{\pi S_n(\pi/2)\left(1-\left(\frac{2}{\pi}\right)^{2n}x^{2n}\right)} < \frac{\tan x}{S_n(x)} < \frac{1}{1-\left(\frac{2}{\pi}\right)^{2n}x^{2n}}. \tag{3.2}$$

So, the two inequalities (3.1) and (3.2) have the same upper bound. To compare the lower bounds of (3.1) and (3.2), take $n = 1, 2, 3$ in the left-hand side of (3.1), to obtain

$$\begin{array}{ll} n=1 & L_1(x) = \frac{1}{1-\frac{x^2}{3}} \\ n=2 & L_2(x) = \frac{1}{1-\frac{2x^4}{15}} \\ n=3 & L_3(x) = \frac{1}{1-\frac{17x^6}{315}} \end{array}$$

and set $n = 1, 2, 3$ in the left-hand side of (3.2), to get

$$\begin{array}{ll} n=1 & K_1(x) = \frac{8}{\pi^2(1-\frac{4x^2}{\pi^2})} \\ n=2 & K_2(x) = \frac{8}{\pi(\pi/2+\pi^3/24)\left(1-\frac{16x^4}{\pi^4}\right)} \\ n=3 & K_3(x) = \frac{12}{\pi(\pi/2+\pi^3/24+\pi^5/240)\left(1-\frac{64x^6}{\pi^6}\right)}. \end{array}$$

Then

$$\begin{array}{ll} L_1(x) > K_1(x) & if\ 0 < x < \frac{1}{2}\sqrt{3(\pi^2-8)} \\ L_1(x) < K_1(x) & if\ \frac{1}{2}\sqrt{3(\pi^2-8)} < x < \pi/2 \\ L_2(x) > K_2(x) & if\ 0 < x < \frac{1}{2}\sqrt[4]{\frac{5(-192\pi^2+12\pi^4+\pi^6)}{3(20-\pi^2)}} \\ L_2(x) < K_2(x) & if\ \frac{1}{2}\sqrt[4]{\frac{5(-192\pi^2+12\pi^4+\pi^6)}{3(20-\pi^2)}} < x < \pi/2 \\ L_3(x) > K_3(x) & if\ 0 < x < \frac{1}{2}\sqrt[6]{\frac{7(-2880\pi^4+120\pi^6+10\pi^8+\pi^{10})}{10(84+7\pi^2-\pi^4)}} \\ L_3(x) < K_3(x) & if\ \frac{1}{2}\sqrt[6]{\frac{7(-2880\pi^4+120\pi^6+10\pi^8+\pi^{10})}{10(84+7\pi^2-\pi^4)}} < x < \pi/2. \end{array}$$

Hence, the lower bounds of (3.1) and (3.2) are not included in each other. Also, we can conclude that inequality (3.1) is better than inequality (3.2) near the origin and that inequality (3.2) is better than inequality (3.1) near $\pi/2$.

## References


[1] Anderson, G. D., Vamanamurthy, M. K. and Vuorinen, M. *Monotonicity of some functions in Calculus.*
(Available online at https://www.math.auckland.ac.nz/Research/Reports/Series/538.pdf).

[2] Becker, M. and Stark, E. L. *On hierarchy of polynomial inequalities for* $\tan(x)$, Univ. Beograd, Publ. Elektrotehn. Fak. Ser. Mat. Fiz. No. 602-633, 133-138, 1978.

[3] Chen, C.-P. and Qi, F. *A double inequality for remainder of power series of tangent function*, Tamkang Journal of Mathematics **34** (4), 351-355, 2003.

[4] Chen, C.-P. and Cheung, W.-S. *Sharp Cusa and Becker-Stark inequalities.* Journal of Inequalities and Applications 2011 2011:136.

[5] Debnath, L., Mortici, C. and Zhu L. *Refinements of Jordan-Stečkin and Becker-Stark inequalities*, Results Math. DOI: 10.1007/s00025-014- 0405-3

[6] Ge, H.-F. *New sharp bounds for the Bernoulli numbers and refinement of Becker-Stark inequalities*, Journal of Applied Mathematics, Article ID 137507, 7 pages, 2012.

[7] Hardy, G. H., Littlewood, J. E. and Pólya, G. *Inequalities*, Cambridge Univ. Press, 1959.

[8] Kuang, J.-Ch. *Chángyòng Bùděngshì(Applied inequalities)*, $3^{rd}$ ed., Shāndōng Kēxué Jìshù Chūbǎn Shè (Shandong Science and Technology Press), Jinan City, Shandong Province, China, 2004 (Chinese).

[9] Mortici, C. *The natural approach of Wilker-Cusa-Huygens inequalities*, Math. Inequal. Appl. 14, no. 3, 535-541, 2011.

[10] Mortici, C. *A subtly analysis of Wilker inequality*, Applied Mathematics and Computation, 231, 516-520, 2014.

[11] Pinelis, I. *L'Hospital rules for monotonicity and the Wilker-Anglesio inequality*, Amer. Math. Monthly **111** (10), 905-909, 2004.

[12] Qi, F., Niu, D.-W. and Guo, B.-N. *Refinements, generalizations and applications of Jordan's inequality and related problems*, Journal of Inequalities and Applications, Article ID 271923, 52 pages, 2009.

[13] Sándor, J. *Sharp Cusa-Huygens and related inequalities*, Notes on Number Theory and Discrete Mathematics **19** (1), 50-54, 2013.

[14] Stečkin, S. B. *Some remarks on trigonometric polynomials* (Russian). Uspehi Mat. Nauk (N.S.) 10, No. 1 (63), 159-166, 1955.

[15] Sun, Z. J. and Zhu, L. *Simple proofs of the Cusa-Huygens-type and Becker-Stark-type inequakities*, Journal of Mathematical Inequalities, **7** (4), 563-567, 2013.

[16] Zhao, J.-L., Luo, Q.-M., Guo, B.-N. and Qi, F. *Remarks on inequalities for the tangent function*, Hacettepe J. Math. Stat., **41** (4), 499-506, 2012.

[17] Zhu, L. and Hua, J.-K. *Sharpening the Becker-Stark inequalities*, Journal of Inequalities and Applications, Article ID 931275, 2010.

[18] Zhu, L. *Sharp Becker-Stark-type inequalities for Bessel functions*, Journal of Inequalities and Applications, Vol. 2010, Article ID 838740, 4 pages.

[19] Zhu, L. *A refinement of the Becker-Stark inequalities*, Mat. Zametki, **93** (3), 401-406, 2013.

[20] Zhu, L. and Mortici, C. *Inequalities Becker-Stark at extreme points*, arXiv:1312.6276 [math.CA].

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 49 – 55

# Differential identities on Jordan ideals of rings with involution

Abdellah Mamouni* and Lahcen Oukhtite†

**Abstract**

In this paper we investigate generalized derivations satisfying certain differential identities on Jordan ideals of rings with involution and discuss related results. Moreover, we provide examples to show that the assumed restriction cannot be relaxed.

**Keywords:** $*$-prime rings, Jordan ideals, generalized derivations, commutativity.

*2000 AMS Classification:* 16W10, 16W25, 16U80.

*Received* : 21.02.2013 *Accepted* : 19.01.2015 *Doi* : 10.15672/HJMS.20164512481

## 1. Introduction

Throughout this paper, $R$ will represent an associative ring with center $Z(R)$. Recall that $R$ is 2-torsion free if $2x = 0$ yields $x = 0$. The ring $R$ is prime if $aRb = 0$ implies $a = 0$ or $b = 0$. An additive map $* : R \longrightarrow R$ is an involution if $(xy)^* = y^*x^*$ and $(x^*)^* = x$ for all $x, y \in R$. If $R$ admits an involution $*$, then $R$ is $*$-prime if $aRb = aRb^* = 0$ forces $a = 0$ or $b = 0$. It is straightforward to check that a $*$-prime ring is necessarily semiprime, that is $xRx = 0$ forces $x = 0$. Furthermore, every prime ring having an involution $*$ is $*$-prime, but the converse need not be true in general. For example, if $R^o$ denotes the opposite ring of a prime ring $R$, then $R \times R^o$ equipped with the exchange involution $*_{ex}$, defined by $*_{ex}(x, y) = (y, x)$, is $*_{ex}$-prime but not prime. This example shows that every prime ring can be injected in a $*$-prime ring and from this point of view $*$-prime rings constitute a more general class of prime rings.
In all that follows $Sa_*(R) = \{x \in R : x^* = \pm x\}$ will denote the set of symmetric and skew-symmetric elements of $R$. We will write for all $x, y \in R, [x, y] = xy - yx$ and $x \circ y = xy + yx$ for the commutator and anticommutator, respectively. An additive subgroup $U$ of $R$ is a Lie ideal if $[x, r] \in U$ for all $x \in U$ and $r \in R$. An additive subgroup

*Université Moulay Ismail, Faculté des Sciences et Techniques, Département de Mathématiques, BP. 509-Boutalamine, 52000 Errachidia; Maroc.
Email : mamouni_1975@live.fr
†Université Moulay Ismail, Faculté des Sciences et Techniques, Département de Mathématiques, BP. 509-Boutalamine, 52000 Errachidia; Maroc.
Email : oukhtitel@hotmail.com

$J$ of $R$ is a Jordan ideal if $x \circ r \in J$ for all $x \in J$ and $r \in R$. Moreover, if $J^* = J$, then $J$ is called a $*$-Jordan ideal. We shall use without explicit mention the fact that if $J$ is a Jordan ideal of $R$, then $2[R,R]J \subseteq J$ and $2J[R,R] \subseteq J$ ([7], Lemma 1). Moreover, From [1] we have $4jRj \subset J$, $4j^2R \subset J$ and $4Rj^2 \subset J$ for all $j \in J$. An additive mapping $d : R \longrightarrow R$ is a derivation if $d(xy) = d(x)y + xd(y)$ for all $x, y \in R$. Many results in literature indicate how the global structure of a ring $R$ is often tightly connected to the behavior of derivations defined on $R$. More recently several authors consider similar situation in the case the derivation $d$ is replaced by a generalized derivation. More specifically an additive map $F : R \longrightarrow R$ is a generalized derivation if there exists a derivation $d$ such that $F(xy) = F(x)y + xd(y)$ for all $x, y \in R$. Basic examples of generalized derivations are the usual derivations on $R$ and left $R$-module mappings from $R$ into itself. Generalized derivations have been primarily studied on operator algebras. Therefore any investigation from the algebraic point of view might be interesting (see for example [2] and [8]).
Recently many authors have studied commutativity of prime and semiprime rings admitting suitably constrained additive mappings, as automorphisms, derivations, skew derivations and generalized derivations acting on appropriate subsets of the rings. Moreover, many of obtained results extend other ones proven previously just for the action of the considered mapping on the whole ring. In this paper we continue the line of investigation regarding the study of commutativity for rings with involution satisfying certain differential identities involving generalized derivations acting on Jordan ideals.

## 2. Differential commutator identities

In 2002 Rehman [9] established that if a 2-torsion free prime ring admits a generalized derivation $F$ associated with a nonzero derivation such that $F\Big([x,y]\Big) = [x,y]$ (or $F\Big([x,y]\Big) = -[x,y]$) for all $x, y$ in a nonzero square closed Lie ideal $U$, then $U \subset Z(R)$. Quadri et al. [8], without 2-torsion freeness hypothesis, proved that a prime ring must be commutative if it admits a generalized derivation $F$, associated with a nonzero derivation, such that $F\Big([x,y]\Big) = [x,y]$ (or $F\Big([x,y]\Big) = -[x,y]$) for all $x, y$ in a nonzero ideal $I$. Motivated by the above results, in this section we explore the commutativity of a $*$-prime ring $R$ in which the generalized derivation $F$ satisfies similar identities on a $*$-Jordan ideal. We shall conclude this section with an application of our results which extend results of [8] and [9] to Jordan ideals with the additional assumption that the ring R be 2-torsion free.

We begin with the following known results which will be used extensively to prove our theorems.

**1. Lemma.** ([3], Lemma 2) Let $R$ be a 2-torsion free $*$-prime ring and $J$ a nonzero $*$-Jordan ideal of $R$. If $aJb = a^*Jb = 0$ (or $aJb = aJb^* = 0$), then $a = 0$ or $b = 0$.

**2. Lemma.** ([5], Lemma 3) Let $R$ be a 2-torsion free $*$-prime ring and $J$ a nonzero $*$-Jordan ideal. If $d$ is a derivation such that $d(x^2) = 0$ for all $x \in J$, then $d = 0$.

**3. Lemma.** Let $R$ be a 2-torsion free $*$-prime ring and $J$ a nonzero $*$-Jordan ideal of $R$. If $R$ admits a nonzero derivation $d$ such that $[[r,s],y]Jd(y^2) = 0$ for all $r, s \in R$ and $y \in J$, then $J \cap Z(R) \neq \{0\}$.

*Proof.* Assume that $J \cap Z(R) = \{0\}$. We have

(2.1) $\quad [[r,s],y]Jd(y^2) = 0$ for all $y \in J,\ r,s \in R.$

Let $y \in J \cap Sa_*(R)$; then (2.1) implies that

$$[[r,s],y]^* Jd(y^2) = 0$$

and combining this equation with (2.1), then Lemma 1 yields that either $d(y^2) = 0$ or $[[r,s],y] = 0$. Suppose

(2.2) $\quad [[r,s],y] = 0$ for all $r,s \in R$.

Substituting $sy$ for $s$ in (2.2) we get

$$0 = [[r,sy],y] = s[[r,y],y] + [s,y][r,y] + [[r,s],y]y$$

and employing (2.2) we find that

(2.3) $\quad [s,y][r,y] = 0$ for all $r,s \in R$.

Replacing $r$ by $rs$ in (2.3) we get $[s,y]r[s,y] = 0$ and thus

(2.4) $\quad [s,y]R[s,y] = 0$ for all $s \in R$.

In view of semi-primeness of $R$, equation (2.4) assures that $y \in Z(R)$ and thus $y = 0$. Accordingly

(2.5) $\quad d(y^2) = 0$ for all $y \in J \cap Sa_*(R)$.

Let $y \in J$, as $y^* - y,\ y^* + y \in J \cap Sa_*(R)$, then (2.5) forces $d(y^2) = -d((y^*)^2)$. Substituting $y^*$ for $y$ in (2.1) we obtain

$$[[r,s],y^*]Jd(y^2) = 0 \text{ for all } r,s \in R.$$

In particular,

$$[[r^*,s^*],y^*]Jd(y^2) = 0 \text{ for all } r,s \in R$$

which implies that

(2.6) $\quad [[r,s],y]^* Jd(y^2) = 0$ for all $y \in J,\ r,s \in R$.

Combining (2.1) and (2.6), we conclude that $d(y^2) = 0$ or $[[r,s],y] = 0$ which, as above, leads to $d(y^2) = 0$. Consequently,

$$d(y^2) = 0 \text{ for all } y \in J$$

and Lemma 2 assures that $d = 0$ which contradicts our hypothesis. □

**1. Theorem.** Let $R$ be a 2-torsion free $*$-prime ring and $J$ be a nonzero $*$-Jordan ideal of $R$. If $R$ admits a generalized derivation $F$ associated with a nonzero derivation $d$ such that $F\Big([x,y]\Big)\Big) = [x,y]$ for all $x,y \in J$, then $R$ is commutative.

*Proof.* Assume that

(2.7) $\quad F\Big([x,y]\Big) = [x,y]$ for all $x,y \in J$.

Replacing $x$ by $4xy^2$ in (2.7) we get $F\Big([x,y]y^2\Big) = [x,y]y^2$ and thus

(2.8) $\quad [x,y]d(y^2) = 0$ for all $x,y \in J$.

Substituting $2[r,s]x$ for $x$ in (2.8), where $r,s \in R$, we find that $[[r,s],y]xd(y^2) = 0$ and therefore

(2.9) $\quad [[r,s],y]Jd(y^2) = 0$ for all $y \in J$ and $r,s \in R$.

In view of (2.9), application of Lemma 3 assures that $J \cap Z(R) \neq \{0\}$. Replacing $x$ by $4x^2u$ in (2.7), where $0 \neq u \in J \cap Z(R)$, we get

$$(2.10) \quad F\Big([2x^2, y]u\Big) = [2x^2, y]u \quad \text{for all} \;\; x, y \in J.$$

Using (2.7), equation (2.10) yields $[x^2, y]d(u) = 0$ and thus

$$(2.11) \quad [x^2, y]Jd(u) = 0 \quad \text{for all} \;\; x, y \in J.$$

Since $J$ is a $*$-ideal, (2.11) forces

$$(2.12) \quad [x^2, y]^*Jd(u) = 0 \quad \text{for all} \;\; x, y \in J.$$

Combining (2.11) and (2.12), Lemma 1 yields $d(u) = 0$ or $[x^2, y] = 0$ for all $x, y \in J$.
If $[x^2, y] = 0$ for all $x, y \in J$, then $R$ is commutative by proof of Theorem 3 in [4].
If $d(u) = 0$, then replacing $x$ by $4ru^2$ in (2.7) we obtain $F\Big([r, y]u^2\Big) = [r, y]u^2$ and thus

$$\Big(F([r, y]) - [r, y]\Big)u^2 = 0.$$

Accordingly

$$(2.13) \quad \bigg(F\Big([r, y]\Big) - [r, y]\bigg)Ju^2 = 0 \quad \text{for all} \;\; y \in J, r \in R.$$

As $0 \neq u^* \in J \cap Z(R)$, then a similar reasoning as above leads to

$$(2.14) \quad \bigg(F\Big([r, y]\Big) - [r, y]\bigg)J(u^2)^* = 0 \quad \text{for all} \;\; y \in J, r \in R.$$

In view of Lemma 1, (2.13) together with (2.14) forces

$$(2.15) \quad F\Big([r, y]\Big) = [r, y] \quad \text{for all} \;\; y \in J \;\; \text{and} \;\; r \in R.$$

Substituting $ry$ for $r$ in (2.15) we get

$$(2.16) \quad [r, y]d(y) = 0 \quad \text{for all} \;\; y \in J \;\; \text{and} \;\; r \in R.$$

Replacing $r$ by $rs$ in (2.16), where $s \in R$, we obtain $[r, y]sd(y) = 0$ so that

$$(2.17) \quad [r, y]Rd(y) = 0 \quad \text{for all} \;\; y \in J \;\; \text{and} \;\; r \in R.$$

Once again using the proof of Theorem 3 in [4], from equation (2.17) it follows that $R$ is commutative. □

As an application of Theorem 1, the following theorem extends ([9], Theorem 3.3) and ([8], Theorem 2.1 ) to Jordan ideals.

**2. Theorem.** Let $R$ be a 2-torsion free prime ring and $J$ be a nonzero Jordan ideal of $R$. If $R$ admits a generalized derivation $F$ associated with a nonzero derivation $d$ such that $F\Big([x, y]\Big) = [x, y]$ for all $x, y \in J$, then $R$ is commutative.

*Proof.* Assume that $F$ is a generalized derivation associated to a nonzero derivation $d$ such that $F\Big([x, y]\Big) = [x, y]$, for all $x, y \in J$. Let $\mathcal{D}$ be the additive mapping defined on $\mathcal{R} = R \times R^0$ by $\mathcal{D}(x, y) = (d(x), 0)$ and $\mathcal{F}(x, y) = (F(x), y)$. Clearly, $\mathcal{D}$ is a nonzero derivation of $\mathcal{R}$ and $\mathcal{F}$ is a generalized derivation associated with $\mathcal{D}$. Moreover, if we set $\mathcal{J} = J \times J$, then $\mathcal{J}$ is a $*_{ex}$-Jordan ideal of $\mathcal{R}$ and $\mathcal{F}\Big([x, y]\Big) = [x, y]$ for all $x, y \in \mathcal{J}$. Since $\mathcal{R}$

is a $*_{ex}$-prime ring, in view of Theorem 1 we deduce that $\mathcal{R}$ is commutative and a fortiori $R$ is commutative. □

A slight modification in the proof of Theorem 1 yields the following result.

**3. Theorem.** Let $R$ be a 2-torsion free $*$-prime ring and $J$ be a nonzero $*$-Jordan ideal of $R$. If $R$ admits a generalized derivation $F$ associated with a nonzero derivation $d$ such that $F\Big([x,y]\Big) = -[x,y]$ for all $x, y \in J$, then $R$ is commutative.

Reasoning as in the proof of Theorem 2, where $\mathcal{F}(x,y) = (F(x), -y)$, and using Theorem 3 we extend ([9], Theorem 3.4) and ([8], Theorem 2.2 ) to Jordan ideals as follows.

**4. Theorem.** Let $R$ be a 2-torsion free prime ring and $J$ be a nonzero Jordan ideal of $R$. If $R$ admits a generalized derivation $F$ associated with a nonzero derivation $d$ such that $F\Big([x,y]\Big) = -[x,y]$ for all $x, y \in J$, then $R$ is commutative.

## 3. Differential anticommutator identities

It is natural to ask what can we say about the commutativity of $R$ if the commutator in the preceding section is replaced by anticommutator. In this section, we have investigated this problem and proved that the commutativity cannot be characterized by the same conditions on anticommutator.

**5. Theorem.** Let $R$ be a 2-torsion free $*$-prime ring and $J$ be a nonzero $*$-Jordan ideal of $R$. If $R$ admits a generalized derivation $F$ associated with a derivation $d$ such that $F(x \circ y) = x \circ y$ for all $x, y \in J$, then $d = 0$ and $F$ is the identity map.

*Proof.* Assume that

(3.1) $\quad F(x \circ y) = x \circ y \quad$ for all $x, y \in J$.

Replacing $y$ by $4yx^2$ in (3.1) we find that

(3.2) $\quad (x \circ y)d(x^2) = 0 \quad$ for all $x, y \in J$.

Substituting $2[r,s]y$ for $y$ in (3.2), where $r, s \in R$, we obtain $(x \circ (2[r,s]y))d(x^2) = 0$ and thus $[x,[r,s]]yd(x^2) = 0$. Hence

(3.3) $\quad [x,[r,s]]Jd(x^2) = 0 \quad$ for all $x \in J$ and $r, s \in R$.

In view of Lemma 3, equation (3.3) assures that $d = 0$ or $J \cap Z(R) \neq 0$.
If there exists $0 \neq u \in J \cap Z(R)$, then replacing $y$ by $4u^2y$ in (3.1), we get

(3.4) $\quad F(2u^2(x \circ y)) = 2u^2(x \circ y) \quad$ for all $x, y \in J$.

Since by assumption of the theorem $F(2u^2) = F(u \circ u) = 2u^2$, then (3.4) leads to

(3.5) $\quad u^2 d(x \circ y) = 0 \quad$ for all $x, y \in J$.

Using the fact that $u \in Z(R)$, from (3.5) it follows that

(3.6) $\quad u^2 J d(x \circ y) = 0 \quad$ for all $x, y \in J$.

As $0 \neq u^* \in J \cap Z(R)$, a similar reasoning as above yields

(3.7) $\quad (u^2)^* J d(x \circ y) = 0 \quad$ for all $x, y \in J$.

We claim that $u^2 \neq 0$. For contradiction assume that $u^2 = 0$, then $uRu = 0$. Since $R$ is semiprime $u = 0$. This is a contradiction. Thus $u^2 \neq 0$.
Now if we combine (3.6) and (3.7) and apply Lemma 1, we conclude that

$$d(x \circ y) = 0 \quad \text{for all } x, y \in J$$

and a fortiori

$$(3.8) \qquad d(x^2) = 0 \quad \text{for all } \ x \in J.$$

In light of Lemma 2, equation (3.8) forces $d = 0$ hence $F$ is a left multiplier.
From $F(x \circ y) = x \circ y$ it then follows $(F(x) - x)y = -(F(y) - y)x$ and replacing $y$ by $y \circ z$ where $z \in J$ we get

$$(3.9) \qquad (F(x) - x)(y \circ z) = 0 \quad \text{for all } \ x, y, z \in J.$$

Replacing $z$ by $2z[r, s]$ in (3.9) where $r, s \in R$ we obtain

$$(3.10) \qquad (F(x) - x)J[y, [r, s]] = (F(x) - x)J\Big([y, [r, s]]\Big)^* = 0 \quad \text{for all } \ x, y \in J, \ \ r, s \in R.$$

Thus, according to Lemma 1, either $F(x) = x$ for all $x \in J$ or $[y, [r, s]] = 0$ for all $y \in J$ and $r, s \in R$.
Assume that $[y, [r, s]] = 0$ for all $y \in J$ and $r, s \in R$, hence as in (2.2) this implies that

$$[s, y]R[s, y] = 0 \quad \text{for all } \ y \in J, \ \ s \in R.$$

and the semi-primeness of $R$ forces $[s, y] = 0$ so that $J \subseteq Z(R)$. Therefore [[6], Lemma 3] assures that $R$ is a commutative ring in which case, as $F$ is a left multiplier, equation (3.1) implies that $F(x)y = xy$.
In conclusion, in either case (3.1) becomes

$$(3.11) \qquad (F(x) - x)y = 0 \quad \text{for all } \ x, y \in J$$

in such a way that $F(x) = x$ for all $x \in J$. Let $r \in R$ and $x \in J$, from $F(x \circ r) = x \circ r$ it follows that

$$\begin{aligned} xr + rx &= F(xr + rx) \\ &= F(x)r + F(r)x \\ &= xr + F(r)x \end{aligned}$$

so that

$$(F(r) - r)x = 0 \quad \text{for all } \ r \in R, \ x \in J$$

and therefore $F(r) = r$ for all $r \in R$. Hence $F$ is the identity map. □

Using similar arguments as used in the proof of Theorem 2, application of Theorem 5 yields the following result which extends ([9], Theorem 3.7) and ([8], Theorem 2.3) to Jordan ideals in the case of a 2-torsion free ring.

**6. Theorem.** Let $R$ be a 2-torsion free prime ring and $J$ be a nonzero Jordan ideal of $R$. If $R$ admits a generalized derivation associated with a derivation $d$ such that $F(x \circ y) = x \circ y$ for all $x, y \in J$, then $d = 0$ and $F$ is the identity map.

Reasoning as in proof of Theorem 5, we can prove the following.

**7. Theorem.** Let $R$ be a 2-torsion free $*$-prime ring and $J$ be a nonzero $*$-Jordan ideal of $R$. If $R$ admits a generalized derivation $F$ associated with a derivation $d$ such that $F(x \circ y) = -\, x \circ y$ for all $x, y \in J$, then $d = 0$ and $(-F)$ is the identity map.

Similarly, application of Theorem 7 yields the following result which improves ([9], Theorem 3.8) and ([8], Theorem 2.4).

**8. Theorem.** Let $R$ be a 2-torsion free prime ring and $J$ be a nonzero Jordan ideal of $R$. If $R$ admits a generalized derivation $F$ associated with a derivation $d$ such that $F(x \circ y) = -\, x \circ y$ for all $x, y \in J$, then $d = 0$ and $(-F)$ is the identity map.

**1. Example.** Let $R = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}$ and consider $F \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} a & 2b \\ 0 & 0 \end{pmatrix}$. It is straightforward to verify that $F$ is a generalized derivation associated with the non zero derivation $d$ defined by $d \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$. Moreover, if we set $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^* = \begin{pmatrix} c & -b \\ 0 & a \end{pmatrix}$, then $R$ is a non $*$-prime ring. Furthermore, it is easy to verify that $J = \left\{ \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \mid a \in \mathbb{R} \right\}$ is a nonzero $*$-Jordan ideal of $R$ such that

$$F(A \circ B) = A \circ B,\ F(A \circ B) = -\, A \circ B,\ F[A, B] = [A, B],\ F[A, B] = -[A, B]$$

for al $A, B \in R$. Hence in theorems 1, 3, 5, 7 the $*$-primeness hypothesis is crucial.

## References

[1] R. Awtar, *Lie and Jordan structure in prime rings with derivations*, Proc. Amer. Math. Soc. 41 (1973), 67-74.

[2] B. Hvala, *Generalized derivations in rings*, Comm. Algebra 26 (1998), 1147-1166.

[3] L. Oukhtite, *On Jordan ideals and derivations in rings with involution*, Comment. Math. Univ. Carolin. 51 (2010), no. 3, 389-395.

[4] L. Oukhtite and A. Mamouni, *Derivations satisfying certain algebraic identities on Jordan ideals*, Arab. J. Math. 1 (2012), no. 3, 341-346.

[5] L. Oukhtite and A. Mamouni, *Generalized derivations centralizing on Jordan ideals of rings with involution*, Turkish J. Math. 38 (2014), no. 2, 225-232.

[6] L. Oukhtite, *Posner's Second Theorem for Jordan ideals in rings with involution*, Expo. Math., 29 (4), 415-419 (2011).

[7] S. M. A. Zaidi, M. Ashraf, and S. Ali, *On Jordan ideals and left $(\theta, \theta)$-derivations in prime rings*, Int. J. Math. Math. Sci. 2004 (2004), no. 37-40, 1957-1964.

[8] M. A. Quadri, M. S. Khan and N. Rehman, *Generalized derivations and commutativity of prime rings,* Indian J. Pure Appl. Math. 34 (2003), no. 9, 1393-1396.

[9] N. Rehman, *On commutativity of rings with generalized derivations,* Math. J. Okayama Univ. 44 (2002), 43-49.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 57 – 67

# Some properties of $AFG$ and $CTF$ rings

Lixin Mao*

**Abstract**

$R$ is said to be a right $AFG$ ring if the right annihilator of every nonempty subset of $R$ is a finitely generated right ideal. $R$ is called a right $CTF$ ring if every cyclic torsionless right $R$-module embeds in a free module. In this paper, we first give new characterizations of $AFG$ rings and study some closure properties of $AFG$ rings. Then we explore the intimate relationships between $AFG$ rings and $CTF$ rings.

**Keywords:** $AFG$ ring; $CTF$ ring; pseudo-coherent ring; $FP$-injective module; singly projective module.

*2000 AMS Classification:* 16D40, 16P70.

*Received* : 26.12.2013 *Accepted* : 27.11.2014 *Doi* : 10.15672/HJMS.20164512483

## 1. Introduction

In [19], we introduced the concept of $AFG$ rings, which is a generalization of Noetherian rings. $R$ is said to be a *right $AFG$ ring* in case the right annihilator of every nonempty subset of $R$ is a finitely generated right ideal, equivalently, every cyclic torsionless right $R$-module is finitely presented, where a right $R$-module $M$ is called *torsionless* if $M$ embeds in a direct product of copies of $R_R$. The concept of $AFG$ rings is very useful in ring theory. For more details about $AFG$ rings, we refer the reader to [19, 20, 21].

In this paper, we gave some new characterizations of $AFG$ rings and further study some properties of $AFG$ rings, such as closure properties under finite direct products, quotients and localizations. On the other hand, we explore the intimate connections between $AFG$ rings and $CTF$ rings, where a ring $R$ is called *right $CTF$* [27] if every cyclic torsionless right $R$-module embeds in a free module.

The layout of the paper is as follows:

Section 2 is devoted to $AFG$ rings. We first prove that $R$ is a right $AFG$ ring if and only if the dual module $\mathrm{Hom}_R(M, R)$ of any cyclic torsionless left $R$-module $M$ is finitely generated if and only if every cyclic torsionless left $R$-module has a projective preenvelope. It is also shown that $R$ is a right $AFG$ ring if $R$ is a left singly injective left

*Department of Mathematics and Physics, Nanjing Institute of Technology, Nanjing 211167, China.
Email : maolx2@hotmail.com

$CF$ ring. Next we discuss the closure properties of $AFG$ rings. We prove that: (1) $R$ and $S$ are right $AFG$ rings if and only if $R \times S$ is a right $AFG$ ring. (2) If $R$ is a right $AFG$ ring and $I$ is an ideal which is a right annihilator in $R$, then $R/I$ is a right $AFG$ ring. (3) If $R$ is a commutative $AFG$ ring and $S$ a multiplicative subset of $R$ without zero-divisors, then $S^{-1}R$ is also an $AFG$ ring. Finally we give some examples to clarify the relationships among $AFG$ rings, $AC$ rings, $\Pi$-coherent rings and pseudo-coherent rings.

In Section 3, we deal with some properties of $CTF$ rings. For example, it is shown that $R$ is a right $CTF$ ring if the dual module of every cyclic torsionless right $R$-module is $H$-finitely generated, and the converse holds if $R$ is a left $f$-injective ring. Furthermore, we explore the close connections between $AFG$ rings and $CTF$ rings. We prove that: (1) If $R$ is a left $AFG$ ring, then $R$ is a right $CTF$ ring. (2) If $R$ is a right $CTF$ right pseudo-coherent ring, then $R$ is a right $AFG$ ring. (3) $R$ is a left $AFG$ ring if and only if $R$ is a right $CTF$ ring and $lr(S)$ is a finitely generated left ideal for any finite subset $S$ of $R$. (4) $R$ is a two-sided $AFG$ two-sided singly injective ring if and only if $R$ is a two-sided $CTF$ two-sided $FP$-injective ring.

Throughout this paper, $R$ is an associative ring with identity and all modules are unitary. $M_R$ (resp. $_RM$) denotes a right (resp. left) $R$-module. For an $R$-module $M$, the dual module $\mathrm{Hom}_R(M,R)$ is denoted by $M^*$ and the character module $M^+$ is defined by $M^+ = \mathrm{Hom}_{\mathbb{Z}}(M, \mathbb{Q}/\mathbb{Z})$. $E(M)$ denotes the injective envelope of $M$. $M^I$ (resp. $M^{(I)}$) stands for the direct product (resp. direct sum) of copies of $M$ indexed by a set $I$. For a subset $X$ of $R$, the right (resp. left) annihilator of $X$ in $R$ is denoted by $r(X)$ (resp. $l(X)$). We refer to [1, 9, 15, 16, 24, 26] for all undefined notions in this article.

## 2. $AFG$ rings

In [19], the author gave some characterizations of $AFG$ rings. For example, $R$ is a right $AFG$ ring if and only if the dual module $M^*$ of any cyclic left $R$-module $M$ is finitely generated if and only if every cyclic left $R$-module has a projective preenvelope. The following theorem gives an improvement of the above result.

Recall that that a homomorphism $f : M \to P$ is called a *projective preenvelope* of a left $R$-module $M$ [9] if $P$ is projective, and for any homomorphism $g$ from $M$ to any projective left $R$-module $P'$, there exists $h : P \to P'$ such that $g = hf$.

We also recall a right $R$-module $M$ is *$FP$-injective* (or *absolutely pure*) [25, 17] if $\mathrm{Ext}^1_R(N, M) = 0$ for any finitely presented right $R$-module $N$. $M$ is called *$\mathcal{A}$-injective* [18] if $\mathrm{Ext}^1_R(R/I, M) = 0$ for any right annihilator $I$ in $R$.

**2.1. Theorem.** *The following are equivalent for a ring $R$:*

(1) *$R$ is a right $AFG$ ring.*
(2) *The dual module $M^*$ of any cyclic torsionless left $R$-module $M$ is finitely generated.*
(3) *For any cyclic torsionless left $R$-module $A$ and $x \in A$, the additive subgroup $H_{A,x} = \{f(x) : f \in \mathrm{Hom}_R(A, R)\}$ of $R$ is a finitely generated right ideal.*
(4) *Every cyclic torsionless left $R$-module has a projective preenvelope.*
(5) *Every $FP$-injective right $R$-module is $\mathcal{A}$-injective.*

*Proof.* (1) ⇒ (2) and (1) ⇒ (4) are obvious by [19, Theorem 2.3].

(2) ⇒ (1) Let $I$ be any right annihilator in $R$. Then the exact sequence

$$0 \to I \xrightarrow{i} R_R \xrightarrow{f} R/I \to 0$$

of right $R$-modules yields the exact sequence of left $R$-modules

$$0 \to (R/I)^* \xrightarrow{f^*} (R_R)^* \xrightarrow{i^*} I^*.$$

Let $B = \mathrm{im}(i^*)$. Then we get the exact sequence

$$0 \to (R/I)^* \xrightarrow{f^*} (R_R)^* \to B \to 0,$$

which gives rise to the exactness of the sequence

$$0 \to B^* \to (R_R)^{**} \to (R/I)^{**}.$$

By [24, Exercise 2.7, p.27], there exists $\phi : I \to B^*$ such that the following diagram with exact rows commutes.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & I & \longrightarrow & R_R & \longrightarrow & R/I & \longrightarrow & 0 \\ & & \Big\downarrow{\scriptstyle \phi} & & \Big\downarrow{\scriptstyle \sigma_R} & & \Big\downarrow{\scriptstyle \sigma_{R/I}} & & \\ 0 & \longrightarrow & B^* & \longrightarrow & (R_R)^{**} & \longrightarrow & (R/I)^{**}. & & \end{array}$$

Since $\sigma_{R/I}$ is a monomorphism, $I \cong B^*$ by the Five Lemma. Note that $I^*$ is torsionless by [1, Proposition 20.14], so $B$ is a cyclic torsionless left $R$-module. Thus $I \cong B^*$ is finitely generated by (2), which implies that $R$ is a right $AFG$ ring.

(2) $\Rightarrow$ (3) Let $A$ be any cyclic torsionless left $R$-module and $x \in A$. Then there exist $f_1, f_2, \cdots, f_n \in A^*$ such that

$$A^* = f_1R + f_2R + \cdots + f_nR.$$

So $H_{A,x} = \sum_{k=1}^{n} f_k(x)R$ is a finitely generated right ideal.

(3) $\Rightarrow$ (2) Let $A = Rx$ be a cyclic torsionless left $R$-module. Define a right $R$-homomorphism $\beta : A^* \to H_{A,x}$ via $f \mapsto f(x)$. It is clear that $\beta$ is an isomorphism. Thus $A^*$ is a finitely generated right $R$-module by (3).

(4) $\Rightarrow$ (2) Let $M$ be a cyclic torsionless left $R$-module. Then $M$ has a projective preenvelope $f : M \to P$. We may choose $P$ to be finitely generated since $M$ is cyclic. So we get the exact sequence $P^* \to M^* \to 0$. Thus $M^*$ is finitely generated.

(1) $\Rightarrow$ (5) is clear.

(5) $\Rightarrow$ (1) Let $M$ be a cyclic torsionless right $R$-module. Then $\mathrm{Ext}^1_R(M, N) = 0$ for any $FP$-injective right $R$-module $N$ by (5). Therefore $M$ is finitely presented by [8], and so $R$ is a right $AFG$ ring. □

Now we investigate $AFG$ rings in terms of singly projective, singly injective and singly flat modules.

Recall that a left $R$-module $M$ is *singly projective* [2] in case for any cyclic submodule $N$ of $M$, the inclusion map $N \to M$ factors through a free module.

According to [22], a left $R$-module $M$ (resp. right $R$-module $N$) is called *singly injective* (resp. *singly flat*) if $\mathrm{Ext}^1_R(F/C, M) = 0$ (resp. $\mathrm{Tor}^R_1(N, F/C) = 0$) for any cyclic submodule $C$ of any finitely generated free left $R$-module $F$. $R$ is called a *left singly injective ring* if ${}_RR$ is a singly injective left $R$-module.

Recall that $R$ is a *left CF ring* [13] if every cyclic left $R$-module embeds in a free module.

**2.2. Proposition.** *The following are true:*

1. *$R$ is a left singly injective ring if and only if every singly projective left $R$-module is singly injective.*
2. *$R$ is a left $CF$ ring if and only if every singly injective left $R$-module is singly projective.*

(3) *If $R$ is a left singly injective left $CF$ ring, then $R$ is a right $AFG$ ring.*

*Proof.* (1) " $\Rightarrow$ " Let $M$ be a singly projective left $R$-module. For any cyclic submodule $C$ of any finitely generated free left $R$-module $F$ and any homomorphism $f : C \to M$, there exist a finitely generated free left $R$-module $G$, $g : C \to G$ and $h : G \to M$ such that $f = hg$. Note that $G$ is singly injective, and so there exists $\varphi : F \to G$ such that $\varphi\lambda = g$, where $\lambda : C \to F$ is the inclusion. Hence $(h\varphi)\lambda = hg = f$. Thus $M$ is singly injective.

" $\Leftarrow$ " is clear.

(2) " $\Rightarrow$ " Let $M$ be a singly injective left $R$-module. For any cyclic submodule $N$ of $M$, there exists a monomorphism $\gamma : N \to R^n, n \in \mathbb{N}$. Thus there is $\theta : R^n \to M$ such that $\iota = \theta\gamma$, where $\iota : N \to M$ is the inclusion. So $M$ is singly projective.

" $\Leftarrow$ " is obvious by [19, Lemma 3.6].

(3) Let $\{M_i\}_{i\in I}$ be a family of singly projective left $R$-modules. Then each $M_i$ is singly injective by (1) and so $M_i^I$ is singly injective. Thus $M_i^I$ is singly projective by (2). Hence $R$ is a right $AFG$ ring by [19, Theorem 2.3]. □

It is known that any singly projective $R$-module is singly flat for any ring $R$ by [22, Lemma 2.4] and any singly flat $R$-module is singly projective for any commutative domain $R$ by [22, Corollary 2.6]. Here we have the following result.

**2.3. Proposition.** *The following are equivalent for a ring $R$:*

(1) *$R$ is right $AFG$ and every singly flat left $R$-module is singly projective.*
(2) *$N^+$ is singly projective for every singly injective right $R$-module $N$.*
(3) *$M^{++}$ is singly projective for every singly flat left $R$-module $M$.*

*Proof.* (1) $\Rightarrow$ (2) Since $R$ is right $AFG$, $N^+$ is singly flat by [22, Theorem 2.10] for any singly injective right $R$-module $N$. So $N^+$ is singly projective by (1).

(2) $\Rightarrow$ (3) Let $M$ be a singly flat left $R$-module. Then $M^+$ is singly injective by [22, Lemma 2.4]. So $M^{++}$ is singly projective by (2).

(3) $\Rightarrow$ (1) Let $\{M_i\}_{i\in I}$ be a family of singly projective left $R$-modules, then the pure exact sequence

$$0 \to (M_i^+)^{(I)} \to (M_i^+)^I$$

induces the split exact sequence

$$((M_i^+)^I)^+ \to ((M_i^+)^{(I)})^+ \to 0.$$

Thus $((M_i^+)^{(I)})^+$ is isomorphic to a direct summand of $((M_i^+)^I)^+$. Note that

$$((M_i^+)^{(I)})^+ \cong (M_i^{++})^I, ((M_i^+)^I)^+ \cong (M_i^{(I)})^{++}.$$

Thus $(M_i^{++})^I$ is singly projective since $(M_i^{(I)})^{++}$ is singly projective by (3). Also $M_i^I$ is a pure submodule of $(M_i^{++})^I$ by [6, Lemma 1(2)]. Hence $M_i^I$ is singly projective by [2, Proposition 14], and so $R$ is right $AFG$ by [19, Theorem 2.3].

On the other hand, let $M$ be any singly flat left $R$-module, then $M^{++}$ is singly projective by (3). Note that $M$ is a pure submodule of $M^{++}$, and so $M$ is singly projective by [2, Proposition 14]. □

Recall that $R$ is a *left dual ring* if every left ideal is a left annihilator in $R$, equivalently, every cyclic left $R$-module is torsionless.

**2.4. Theorem.** *The following are equivalent for a ring $R$:*

(1) *$R$ is a right $AFG$ left dual ring.*
(2) *$R$ is a right $AFG$ ring and the injective envelope of every simple left $R$-module is singly projective.*

(3) *$R$ is a right $AFG$ ring and the injective envelope of every finitely cogenerated left $R$-module is singly projective.*
(4) *$R$ is a right $AFG$ ring and $(R_R)^+$ is singly projective.*
(5) *Every cyclic left $R$-module has a projective preenvelope which is a monomorphism.*

*Proof.* (1) $\Rightarrow$ (5) holds by [19, Theorem 3.7].

(5) $\Rightarrow$ (4) $R$ is a right $AFG$ ring by [19, Theorem 2.3]. Let $N$ be a cyclic submodule of $(R_R)^+$. Since $N$ embeds in $R^n, n \in \mathbb{N}$ and $(R_R)^+$ is injective, the inclusion $N \to (R_R)^+$ factors through $R^n$. So $(R_R)^+$ is singly projective.

(4) $\Rightarrow$ (2) Let $M$ be a simple left $R$-module. Then there is a monomorphism $E(M) \to ((R_R)^+)^I$. So $E(M)$ is isomorphic to a direct summand of $((R_R)^+)^I$. Since $((R_R)^+)^I$ is singly projective by [19, Theorem 2.3], $E(M)$ is singly projective.

(2) $\Rightarrow$ (1) Let $N$ be a cyclic left $R$-module. It is enough to show that for any $0 \neq m \in N$, there exists $f : N \to R$ such that $f(m) \neq 0$. In fact, there is a maximal submodule $K$ of $Rm$, and so $Rm/K$ is simple. Let $\iota : Rm \to N$ and $i : Rm/K \to E(Rm/K)$ be the inclusions, and $\pi : Rm \to Rm/K$ be the natural map. Then there exists $j : N \to E(Rm/K)$ such that $j\iota = i\pi$. So $j(m) = j\iota(m) = i\pi(m) \neq 0$. On the other hand, since $E(Rm/K)$ is singly projective by (2), there exist $n \in \mathbb{N}$, $g : N \to R^n$ and $h : R^n \to E(Rm/K)$ such that $j = hg$. Therefore $g(m) = (x_1, x_2, \cdots, x_n) \neq 0$. Let $x_i \neq 0$ and $p_i : R^n \to R$ be the $i$th projection. Then $p_i g(m) \neq 0$. So $N$ is torsionless. Thus $R$ is a left dual ring.

(2) $\Leftrightarrow$ (3) By [15, Theorem 9.4.3], a left $R$-module $N$ is finitely cogenerated if and only if $E(N) = E(S_1) \oplus E(S_2) \oplus \cdots \oplus E(S_n)$, where $S_1, S_2, \cdots, S_n$ are simple left $R$-modules. So (2) $\Leftrightarrow$ (3) follows. $\square$

Next we discuss the closure properties of $AFG$ rings.

**2.5. Theorem.** *$R$ and $S$ are right $AFG$ rings if and only if $R \times S$ is a right $AFG$ ring.*

*Proof.* " $\Rightarrow$ " Let $M$ be a cyclic torsionless right $(R \times S)$-module. Then $M$ has a unique decomposition that $M = A \oplus B$, where $A = M(R, 0)$ is a right $R$-module and $B = M(0, S)$ is a right $S$-module via $xr = x(r, 0)$ for $x \in A$, $r \in R$, and $ys = y(0, s)$ for $y \in B$, $s \in S$. It is easy to verify that $A$ is a cyclic torsionless right $R$-module and $B$ is a cyclic torsionless right $S$-module. Thus $A$ is a finitely presented right $R$-module and $B$ is a finitely presented right $S$-module by hypothesis. So there exist two exact sequences $P_1 \to P_0 \to A \to 0$ of right $R$-modules and $Q_1 \to Q_0 \to B \to 0$ of right $S$-modules, where each $P_i$ is a finitely generated projective right $R$-module, and each $Q_i$ is a finitely generated projective right $S$-module.

Regarding the above exact sequences as exact sequences of right $(R \times S)$-modules, we have an exact sequence of right $(R \times S)$-modules

$$P_1 \oplus Q_1 \to P_0 \oplus Q_0 \to A \oplus B \to 0.$$

Note that each $P_i \oplus Q_i$ is a finitely generated projective right $(R \times S)$-module. So $M = A \oplus B$ is a finitely presented right $(R \times S)$-module. Thus $R \times S$ is a right $AFG$ ring.

" $\Leftarrow$ " Let $M$ be a cyclic torsionless right $R$-module. Note that $M$ may be regarded as a cyclic torsionless right $(R \times S)$-module, so $M$ is a finitely presented right $(R \times S)$-module by hypothesis. Thus there exists an exact sequence $P_1 \to P_0 \to M \to 0$ of right $(R \times S)$-modules, where each $P_i$ is a finitely generated projective right $(R \times S)$-module. Let $P_i = A_i \oplus B_i$, where $A_i$ is a right $R$-module and $B_i$ is a right $S$-module, $i = 0, 1$. Then we have the exact sequence $A_1 \to A_0 \to M \to 0$ of right $R$-modules. Note that each $A_i$ is a finitely generated projective right $(R \times S)$-module, and so is a finitely generated projective right $R$-module, whence $M$ is a finitely presented right $R$-module. Thus $R$ is a right $AFG$ ring. Similarly $S$ is a right $AFG$ ring. □

**2.6. Proposition.** *Let $R$ be a right AFG ring and $I$ be an ideal which is a right annihilator in $R$. Then $R/I$ is also a right AFG ring.*

*Proof.* Let $M_{R/I}$ be a cyclic torsionless right $R/I$-module. Then $M_R$ is clearly a cyclic right $R$-module. Note that $R/I$ is a torsionless right $R$-module since $I$ is a right annihilator in $R$. Thus $M_R$ is also a torsionless right $R$-module. So $M_R$ is a finitely presented right $R$-module, i.e., there is an exact sequence of right $R$-modules

$$R^n \to R^m \to M_R \to 0.$$

Then we get the exact sequence of right $R/I$-modules

$$R^n \otimes_R R/I \to R^m \otimes_R R/I \to M \otimes_R R/I \to 0,$$

which yields the exact sequence of right $R/I$-modules

$$(R/I)^n \to (R/I)^m \to \ M_{R/I} \to 0.$$

Hence $M_{R/I}$ is a finitely presented right $R/I$-module. It follows that $R/I$ is a right $AFG$ ring. □

**2.7. Theorem.** *Let $R$ be a commutative AFG ring. If $S$ is a multiplicative subset of $R$ without zero-divisors, then $S^{-1}R$ is also an AFG ring.*

*Proof.* Let $M$ be a cyclic $S^{-1}R$-module. Then there exists a cyclic $R$-submodule $N$ of $M$ such that $S^{-1}N = M$. Since $S$ contains no zero-divisors, we get the exact sequence of $R$-modules

$$0 \to R \to S^{-1}R \to S^{-1}R/R \to 0,$$

which induces the exact sequence

$$0 \to \mathrm{Hom}_R(N, R) \to \mathrm{Hom}_R(N, S^{-1}R) \to \mathrm{Hom}_R(N, S^{-1}R/R).$$

On the other hand, there exists an exact sequence $R \to N \to 0$, which induces the exact sequence

$$0 \to \mathrm{Hom}_R(N, S^{-1}R/R) \to \mathrm{Hom}_R(R, S^{-1}R/R) \cong S^{-1}R/R.$$

Since $S^{-1}(S^{-1}R/R) = 0$, we have $S^{-1}(\mathrm{Hom}_R(N, S^{-1}R/R)) = 0$. Thus

$$\mathrm{Hom}_{S^{-1}R}(M, S^{-1}R) \cong \mathrm{Hom}_{S^{-1}R}(S^{-1}R \otimes_R N, S^{-1}R)$$

$$\cong \mathrm{Hom}_R(N, S^{-1}R) \cong S^{-1}\mathrm{Hom}_R(N, S^{-1}R) \cong S^{-1}\mathrm{Hom}_R(N, R).$$

Since $\mathrm{Hom}_R(N, R)$ is a finitely generated $R$-module by [19, Theorem 2.3], we have $\mathrm{Hom}_{S^{-1}R}(M, S^{-1}R)$ is a finitely generated $S^{-1}R$-module. So $R/I$ is an $AFG$ ring by [19, Theorem 2.3] again. □

At the end of this section, we consider several rings related to $AFG$ rings.

Recall that $R$ is said to be a *right AC ring* [18] if the right annihilator of each nonempty subset of $R$ is a cyclic right ideal. $R$ is called a *right* $\Pi$*-coherent ring* [4] in case every finitely generated torsionless right $R$-module is finitely presented. $R$ is called a *right coherent ring* [5] if every finitely generated right ideal is finitely presented. $R$ is called a *right pseudo-coherent ring* [3] if the right annihilator of each finite subset of $R$ is a finitely generated right ideal.

Obviously, we have the following implications:

$$\begin{array}{ccccc} & & \text{right } \Pi-\text{coherent rings} & \Longrightarrow & \text{right coherent rings} \\ & & \Downarrow & & \Downarrow \\ \text{right } AC \text{ rings} & \Longrightarrow & \text{right } AFG \text{ rings} & \Longrightarrow & \text{right pseudo}-\text{coherent rings.} \end{array}$$

But these are not generally reversible as shown by the following examples.

**2.8. Example.** Let $F$ be a field with an isomorphism $x \mapsto \bar{x}$ from $F$ to a subfield $\bar{F} \neq F$. Let $R$ denote the right $F$-space on a basis $\{1,\ \text{c}\}$ where $c^2 = 0$ and $cx = \bar{x}c$ for all $x \in F$. Then by [3, Example] or [28, Example 2.7], $R$ is right Artinian, and so is right $AFG$. But $R$ is not right $AC$. Otherwise, suppose that $R$ is a right $AC$ ring. Let $t \neq 0$ be an element of the Jacobson radical $J = Rc = Fc$, then $J \subseteq r(t) \neq R$. Since $R$ is local, $J = r(t)$. Thus $J = aR$ and so $a = bc$ for some $b \in R$. Note that $b$ is a unit since $b \notin J$. So $cR = b^{-1}aR = b^{-1}J = J = Fc$. But $cR = \bar{F}c$, and so $\bar{F}c = Fc$, which contradicts the fact that $\bar{F} \neq F$.

In fact, we have the following result.

**2.9. Proposition.** *$R$ is a right AC ring if and only if $R$ is a right $AFG$ ring and $rl(S)$ is a cyclic right ideal for any finite subset $S$ of $R$.*

*Proof.* " $\Leftarrow$ " Let $r(T)$ be a right annihilator in $R$ for $T \subseteq R$. Then $r(T) = a_1R + a_2R + \cdots + a_nR$. By [1, Proposition 2.15], we have

$$r(T) = rl(r(T)) = rl\{a_1, a_2, \cdots, a_n\}$$

is a cyclic right ideal of $R$. So $R$ is a right $AC$ ring.

" $\Rightarrow$ " is trivial. □

**2.10. Example.** Let $F$ be a field and $R$ the subring of $F^{\mathbb{N}}$ consisting of "sequences" $(a_1, a_2, \cdots) \in F^{\mathbb{N}}$ that are eventually constant. Then $R$ is a commutative von Neumann regular ring (see [16, Example 7.54]) and so is pseudo-coherent.

Let $e_i \in R$ denote the $i^{th}$ unit vector $(0, \cdots, 1, 0, \cdots)$ and $S = \{e_1, e_3, e_5, \cdots\}$. Then $r(S)$ consists of sequences $(a_1, a_2, \cdots)$ that are eventually zero and such that $a_n = 0$ for $n$ odd. Clearly, $r(S)$ is not a finitely generated ideal of $R$. Thus $R$ is not an $AFG$ ring.

Björk proved that $R$ is a right $AFG$ ring if $R$ is a right pseudo-coherent left perfect ring (see [3, Proposition 4.3]).

**2.11. Example.** Let $x, y_1, y_2, \cdots$ be indeterminates over a field $K$, $S = K[x, y_i]$ and $R = K[x^2, x^3, y_i, xy_i]$. Then $R$ is a subring of the commutative domain $S$. Hence $R$ is also a commutative domain, and so is an $AFG$ ring. But $R$ is not a $\Pi$-coherent ring (see [12, p.110]).

It is known that $R$ is a right $\Pi$-coherent ring if and only if every $n \times n$ matrix ring $\mathbb{M}_n(R)$ $(n \geq 1)$ is a right $AFG$ ring (see [20, Corollary 2.5]). Although being right $\Pi$-coherent ring is Morita invariant, it is false for right $AFG$ rings.

## 3. $CTF$ rings

In [27], Xue introduced the concept of right $CTF$ rings. He called a ring $R$ *right $CTF$* if every cyclic torsionless right $R$-module embeds in a free module. This concept is a generalization of right $FGTF$ rings introduced by Faith [11]. Recall that a ring $R$ is *right $FGTF$* if every finitely generated torsionless right $R$-module embeds in a free module.

**3.1. Lemma.** *The following are true:*

(1) *$R$ is a right $CTF$ ring if and only if every right annihilator in $R$ is a right annihilator of a finite subset of $R$.*
(2) *A ring $R$ is right $FGTF$ if and only if every $n \times n$ matrix ring $M_n(R)$ is right $CTF$ for every $n \geq 1$.*

*Proof.* (1) " $\Rightarrow$ " Let $I$ be a right annihilator in $R$. Then there is a monomorphism $f : R/I \to R^n, n \in \mathbb{N}$. Put $f(\bar{1}) = (a_1, a_2, \cdots, a_n)$. It is easy to check that $I = r\{a_1, a_2, \cdots, a_n\}$.

" $\Leftarrow$ " Let $I$ be a right annihilator in $R$. Then $I = r\{b_1, b_2, \cdots, b_n\}$ by hypothesis. Define $g : R/I \to R^n$ by

$$g(\bar{r}) = (b_1 r, b_2 r, \cdots, b_n r).$$

It is easy to verify that $g$ is a monomorphism. So $R$ is a right $CTF$ ring.

(2) follows from (1) and [11, Theorem 1.1]. □

**3.2. Remark.** (1) Although being right $FGTF$ is Morita invariant, being right $CTF$ is not Morita invariant by Lemma 3.1(2).

(2) If $R$ has the a.c.c. on left annihilators, then $R$ is a right $CTF$ ring by Lemma 3.1(1) and [10, Corollary 2].

(3) Clearly, any right $CF$ ring is right $CTF$. But the converse is not true in general.

**3.3. Example.** Let $k$ be a division ring and $V_k$ be a right $k$-vector space of infinite dimension. Let $R = \mathrm{End}(V_k)$. Then $R$ is a right self-injective von Neumann regular ring but not semisimple Artinian (see [16, Example 3.74B]). Note that $R$ is a Baer ring, so $R$ is a right $CTF$ ring. Clearly $R$ is not a right $CF$ ring.

In fact, we have the following easy observation.

**3.4. Proposition.** *$R$ is a right $CF$ ring if and only if $R$ is a right $CTF$ right dual ring.*

Recall that a left $R$-module $M$ is *$H$-finitely generated* [7] if there is a finitely generated submodule $N$ of $M$ such that $(M/N)^* = 0$.

$R$ is called a *left $f$-injective ring* if $\mathrm{Ext}^1_R(R/I, R) = 0$ for any finitely generated left ideal $I$.

**3.5. Theorem.** *If the dual module of every cyclic torsionless right $R$-module is $H$-finitely generated, then $R$ is a right $CTF$ ring. The converse holds if $R$ is a left $f$-injective ring.*

*Proof.* Let $M$ be a cyclic torsionless right $R$-module. Then there exists a finitely generated submodule $N$ of $M^*$ such that $(M^*/N)^* = 0$ by hypothesis.

Let $N = Rf_1 + Rf_2 + \cdots + Rf_n$. Define $\alpha : M \to R^n$ by

$$\alpha(x) = (f_1(x), f_2(x), \cdots, f_n(x)), x \in M.$$

We next prove that $\alpha$ is a monomorphism.

Let $\alpha(x) = 0$, define $\beta : M^*/N \to R$ by

$$\beta(\bar{g}) = g(x), g \in M^*.$$

It is easy to check that $\beta$ is well defined, and so $\beta = 0$. Thus $x \in \bigcap_{g \in M^*} \ker(g)$. Since $M$ is torsionless, we have $x = 0$. So $\alpha$ is a monomorphism and hence $R$ is a right $CTF$ ring.

Conversely, suppose that $R$ is a right $CTF$ ring and $R$ is left $f$-injective. For any cyclic torsionless right $R$-module $M$, there exists an exact sequence $0 \to M \xrightarrow{\gamma} R^n \to L \to 0$. Let $\pi_i : R^n \to R$ be the $i$th projection, $\varphi_i = \pi_i \gamma \in M^*$ and $N = R\varphi_1 + R\varphi_2 + \cdots + R\varphi_n$. We claim that $(M^*/N)^* = 0$. Otherwise, if there exists $0 \neq \xi \in (M^*/N)^*$, then there exists $\theta \in M^*$ such that $\xi(\bar{\theta}) \neq 0$. Write $\lambda : N \to R\theta + N$ and $\iota : R\theta + N \to M^*$ to be the inclusions. Since $M$ is cyclic, there is an exact sequence $R \xrightarrow{\rho} M \to 0$, which induces the exact sequence $0 \to M^* \stackrel{\rho^*}{\to} R^*$. Since $R$ is a left $f$-injective ring, the exact sequence $0 \to N \stackrel{\rho^*\iota\lambda}{\to} R^*$ induces the exact sequence $R^{**} \stackrel{\lambda^*\iota^*\rho^{**}}{\to} N^* \to 0$. Thus $\lambda^*\iota^*\sigma_M\rho = \lambda^*\iota^*\rho^{**}\sigma_R$ is epic, and so $\lambda^*\iota^*\sigma_M$ is epic. We next show that $\lambda^*\iota^*\sigma_M$ is also monic. In fact, if $\lambda^*\iota^*\sigma_M(x) = 0$, then $\sigma_M(x)\iota\lambda = 0$, and so $\sigma_M(x)\iota\lambda(\varphi_i) = 0, i = 1, 2, \cdots, n$. Thus $\varphi_i(x) = 0$, and so $\gamma(x) = 0$. Since $\gamma$ is monic, $x = 0$. Hence $\lambda^*\iota^*\sigma_M$ is an isomorphism.

Similarly, the exact sequence $0 \to R\theta + N \stackrel{\rho^*\iota}{\to} R^*$ induces the exact sequence $R^{**} \stackrel{\iota^*\rho^{**}}{\to} (R\theta + N)^* \to 0$. Then $\iota^*\sigma_M\rho = \iota^*\rho^{**}\sigma_R$ is an epimorphism. So $\iota^*\sigma_M$ is an epimorphism. Also $\iota^*\sigma_M$ is a monomorphism. Thus $\iota^*\sigma_M$ is an isomorphism. Hence $\lambda^* : (R\theta + N)^* \to N^*$ is an isomorphism. Note that the exact sequence

$$0 \to N \stackrel{\lambda}{\to} R\theta + N \to (R\theta + N)/N \to 0$$

induces the exact sequence

$$0 \to ((R\theta + N)/N)^* \to (R\theta + N)^* \stackrel{\lambda^*}{\to} N^*.$$

So $((R\theta + N)/N)^* = 0$. But $\xi|_{(R\theta+N)/N} \neq 0$, a contradiction. Thus $(M^*/N)^* = 0$. Therefore $M^*$ is $H$-finitely generated. □

**3.6. Corollary.** *$R$ is a quasi-Frobenius ring if and only if $R$ is a two-sided dual ring and the dual module of every cyclic right $R$-module is $H$-finitely generated.*

*Proof.* It follows from Theorem 3.5 and [13, Theorem 2.1]. □

Next we consider the relationships between $AFG$ rings and $CTF$ rings.

**3.7. Lemma.** *The following are true:*

(1) *If $R$ is a left $AFG$ ring, then $R$ is a right $CTF$ ring.*
(2) *If $R$ is a right $CTF$ right pseudo-coherent ring, then $R$ is a right $AFG$ ring.*

*Proof.* (1) By Theorem 2.1, the dual module of every cyclic torsionless right $R$-module is finitely generated and so is $H$-finitely generated. Thus $R$ is a right $CTF$ ring by Theorem 3.5.

(2) is clear by Lemma 3.1(1). □

In general, a right or left $CTF$ ring need not be a left $AFG$ ring.

**3.8. Example.** Let $K$ be a field with a subfield $L$ such that $\dim_L K = \infty$, and there exists a field isomorphism $\varphi : K \to L$ (for instance, $K = \mathbb{Q}(x_1, x_2, x_3, \cdots), L = \mathbb{Q}(x_2, x_3, \cdots)$). Let $R = K \times K$ with multiplication

$$(x, y)(x', y') = (xx', \varphi(x)y' + yx'), x, y, x', y' \in K.$$

Then it is easy to see that $R$ has exactly three right ideals: 0, $R$ and $(0, K)$. Therefore $R$ has the a.c.c and the d.c.c on right annihilators and so has the a.c.c. on left annihilators. Thus $R$ is a two-sided $CTF$ ring by Remark 3.2(2).

On the other hand, let $a = (0, 1) \in R$. Then $l(a)$ is not finitely generated (see [16, Example 4.46 (e)]). Thus $R$ is not a left $AFG$ ring.

However we have the following result.

**3.9. Proposition.** *Let $R$ be a two-sided pseudo-coherent ring. Then the following are equivalent:*

(1) *$R$ is a left $AFG$ ring.*
(2) *$R$ is a right $AFG$ ring.*
(3) *$R$ is a left $CTF$ ring.*
(4) *$R$ is a right $CTF$ ring.*

*Proof.* (1) ⇒ (4) and (2) ⇒ (3) follow from Lemma 3.7(1).
(4) ⇒ (2) and (3) ⇒ (1) hold by Lemma 3.7(2). □

**3.10. Corollary.** *The following are true for a ring $R$:*

(1) *$R$ is a two-sided $AFG$ ring if and only if $R$ is a two-sided $CTF$ two-sided pseudo-coherent ring.*
(2) *$R$ is a two-sided $\Pi$-coherent ring if and only if $R$ is a two-sided $FGTF$ two-sided coherent ring.*

*Proof.* (1) is an immediate consequence of Proposition 3.9.
(2) follows from (1), Lemma 3.1(2) and [20, Corollary 2.5]. □

Recall that $R$ is a *right $FP$-injective ring* if $R_R$ is an $FP$-injective right $R$-module. Clearly, any right $FP$-injective ring is right singly injective.

**3.11. Proposition.** *The following are true:*

(1) *$R$ is a left $AFG$ ring if and only if $R$ is a right $CTF$ ring and $lr(S)$ is a finitely generated left ideal for any finite subset $S$ of $R$.*
(2) *A right singly injective ring $R$ is left $AFG$ if and only if $R$ is right $CTF$.*
(3) [27, Corollary 3.4] *A right $FP$-injective ring $R$ is left $\Pi$-coherent if and only if $R$ is right $FGTF$.*

*Proof.* (1) By Lemma 3.7(1), it is enough to show the sufficiency.

Let $l(T)$ be a left annihilator in $R$ for $T \subseteq R$. By Lemma 3.1(1), $rl(T) = r(S)$ for a finite subset $S$ of $R$. So by [1, Proposition 2.15], $l(T) = lrl(T) = lr(S)$ is a finitely generated left ideal. Hence $R$ is a left $AFG$ ring.

(2) For any finite subset $S = \{r_1, r_2, \cdots, r_n\}$ of $R$, $Rr_1 + Rr_2 + \cdots + Rr_n = l(T)$ for some $T \subseteq R$ by [22, Proposition 2.8] since $R$ is a right singly injective ring. So

$$lr(S) = lr(Rr_1 + Rr_2 + \cdots + Rr_n) = lrl(T) = l(T)$$

is a finitely generated left ideal. Thus the result holds by (1).

(3) By [23, Theorem 5.41 and Corollary 5.42], $R$ is a right $FP$-injective ring if and only if every $n \times n$ matrix ring $M_n(R)$ is right singly injective for every $n \geq 1$. So (3) follows from (2), Lemma 3.1(2) and [20, Corollary 2.5]. □

**3.12. Corollary.** *The following are equivalent for a ring $R$:*

(1) *$R$ is a two-sided $AFG$ two-sided singly injective ring.*
(2) *$R$ is a two-sided $AFG$ two-sided $FP$-injective ring.*
(3) *$R$ is a two-sided $CTF$ two-sided $FP$-injective ring.*

*Proof.* (1) ⇒ (2) We first prove that $R$ is a right coherent ring. Let $I$ and $J$ be two finitely generated right ideals of $R$. Then $I = r(X)$ and $J = r(Y)$ for some finitely generated left ideals $X$ and $Y$ of $R$ by [22, Proposition 2.8] and Proposition 3.11. Thus $I \cap J = r(X + Y)$ is finitely generated. Also $r(a)$ is finitely generated for any $a \in R$. So $R$ is a right coherent ring by [5, Theorem 2.2].

On the other hand, $l(I \cap J) = l(r(X) \cap r(Y)) = l(r(X + Y)) = X + Y = l(I) + l(J)$. Thus $R$ is a right $f$-injective ring by [14, Theorem 1]. So $R$ is a right $FP$-injective ring by [25, Lemma 3.1]. Similarly, $R$ is a left $FP$-injective ring.

(2) ⇒ (3) ⇒ (1) follow from Proposition 3.11. □

**ACKNOWLEDGEMENTS**

This research was supported by NSFC (No. 11371187, 11171149), Jiangsu 333 Project, Jiangsu Six Major Talents Peak Project. The author would like to thank the referee for the very helpful comments and suggestions.

## References


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

[2] G. Azumaya, Finite splitness and finite projectivity, *J. Algebra* **106** (1987), 114-134.

[3] J.E. Björk, Rings satisfying certain chain conditions, *J. Reine Angew Math.* **245** (1970), 63-73.

[4] V. Camillo, Coherence for polynomial rings, *J. Algebra* **132** (1990), 72-76.

[5] S.U. Chase, Direct products of modules, *Trans. Amer. Math. Soc.* **97** (1960), 457-473.

[6] T.J. Cheatham and D.R. Stone, Flat and projective character modules, *Proc. Amer. Math. Soc.* **81** (1981), 175-177.

[7] R.R. Colby, Rings which have flat injective modules, *J. Algebra* **35** (1975), 239-252.

[8] E.E. Enochs, A note on absolutely pure modules, *Canad. Math. Bull.* **19** (1976), 361-362.

[9] E.E. Enochs and O.M.G. Jenda, *Relative Homological Algebra*; Walter de Gruyter: Berlin-New York, 2000.

[10] C. Faith, Rings with ascending chain condition on annihilators, *Nagoya Math. J.* **27** (1966), 179-191.

[11] C. Faith, Embedding torsionless modules in projectives, *Publ. Mat.* **34** (1990), 379-387.

[12] S. Glaz, *Commutative Coherent Rings*; Lecture Notes in Math. 1371, Springer-Verlag: New York, 1989.

[13] J.L.Gómez Pardo, Embedding cyclic and torsionfree modules in free modules, *Arch. Math.* **44** (1985), 503-510.

[14] M. Ikeda and T. Nakayama, On some characteristic properties of quasi-Frobenius and regular rings, *Proc. Amer. Math. Soc.* **5** (1954), 15-19.

[15] F. Kasch, *Modules and Rings*; Academic Press: London-New York, 1982.

[16] T.Y. Lam, *Lectures on Modules and Rings*; Springer-Verlag: New York-Heidelberg-Berlin, 1999.

[17] B. Madox, Absolutely pure modules, *Proc. Amer. Math. Soc.* **18** (1967), 155-158.

[18] L.X. Mao, Rings close to Baer, *Indian J. Pure Appl. Math.* **38** (2007), 129-142.

[19] L.X. Mao, A generalization of Noetherian rings, *Taiwanese J. Math.* **12** (2008), 501-512.

[20] L.X. Mao, Baer endomorphism rings and envelopes, *J. Algebra Appl.* **9** (2010), 365-381.

[21] L.X. Mao, Properties of $P$-coherent and Baer modules, *Period. Math. Hungar.* **60** (2010), 97-114.

[22] L.X. Mao and N.Q. Ding, New characterizations of pseudo-coherent rings, *Forum Math.* **22** (2010), 993-1008.

[23] W.K. Nicholson and M.F. Yousif, *Quasi-Frobenius Rings*; Cambridge Tracts in Math. 158, Cambridge University Press, 2003.

[24] J.J. Rotman, *An Introduction to Homological Algebra*; Academic Press: New York, 1979.

[25] B. Stenström, Coherent rings and $FP$-injective modules, *J. London Math. Soc.* **2** (1970), 323-329.

[26] R. Wisbauer, *Foundations of Module and Ring Theory*; Gordon and Breach, 1991.

[27] W. Xue, Rings related to quasi-Frobenius rings, *Algebra Colloq.* **4** (1998), 471-480.

[28] H.Y. Zhu and N.Q. Ding, Generalized morphic rings and their applications, *Comm. Algebra* **35** (2007), 2820-2837.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 69 – 76

# Comparison of near sets by means of a chain of features

A. Fatih Özcan* and Nurettin Bağırmaz†‡

**Abstract**

If the number of features of objects in a perceptual system, is large, then the objects can be known better and comparable. In this paper basically, we form a chain of feature sets that describe objects and then by means of this chain of feature sets, we investigate the nearness of sets and near sets in a perceptual system.

**Keywords:** Near set, Feature chain, Indiscernibility , Nearness.

*2000 AMS Classification:* 03E75, 03E99, 03E02

*Received* : 03.01.2013 *Accepted* : 15.01.2015 *Doi* : 10.15672/HJMS.20164512480

## 1. Introduction

Near sets were introduced by J.F. Peters [11], which are indeed a form of generalization of rough sets proposed by Z. Pawlak [6]. The algebraic properties of near sets are described in [9]. Recent work has considered near soft sets [20], soft nearness approximation spaces [4], near groups [3], isometries in proximity spaces [18], and applications of near sets [17,19]. The fundamental idea of near set theory is object description and classification according to perceptual knowledge. It is supposed that perceptual knowledge about objects is always given with respect to probe functions, i.e., real-valued functions which represent features of a physical object. Some well known examples of probe functions are the colour, size or weight of an object [1,2,9-16,21].

*Inonu University, Science and Art Faculty, Departmant of Mathematics, Maltya, Turkey, Email: abdullah.ozcan@inonu.edu.tr
†Mardin Artuklu University, Vocational Higher Schools of Mardin, Mardin, Turkey, Email: nurettinbagirmaz@artuklu.edu.tr
‡Corresponding Author.

## 2. Preliminiaries

In this section, we present the basic definitions of near set theory [9,11]. More detailed explanations related to near sets and rough sets can be found in [1,2,9-16,21] and [5-8], respectively.

**2.1. Definition.** [9] (*Perceptual Object*) A perceptual object is something perceivable that has its origin in the physical world.

**2.2. Definition.** [9] (*Probe Function*) A probe function is a real-valued function representing a feature of a perceptual object.Simple examples of probe functions are the colour, size or weight of an object.

**2.3. Definition.** [9] (*Perceptual System*) A perceptual system $\langle O, F\rangle$ consists of a non-empty set $O$ of sample perceptual objects and a non-empty set $F$ of real-valued functions $\phi \in F$ such that $\phi : O \to \mathbb{R}$.

**2.4. Definition.** [9] (*Object Description*) Let $\langle O, F\rangle$ be a perceptual system, and let $B \subseteq F$ be a set of probe functions. Then, the description of a perceptual object $x \in O$ is a feature vector given by

$$\phi_B(x) = (\phi_1(x), \phi_2(x), ..., \phi_i(x), ..., \phi_l(x))$$

where $l$ is the length of the vector $\phi_B$, and each $\phi_i(x)$ in $\phi_B(x)$ is a probe function value that is part of the description of the object $x \in O$ .

**2.5. Definition.** $[2, 6]$ (*Indiscernibility relation*) Let $\langle O, F\rangle$ be a perceptual system. For every $B \subseteq F$ the indiscernibility relation $\sim_B$ is defined as follows:

$$\sim_B = \{(x, y) \in O \times O \mid \forall \phi_i \in B, \phi_i\,(x) = \phi_i\,(y)\}\,.$$

If $B = \{\phi\}$ for some $\phi \in F$, instead of $\sim_{\{\phi\}}$ we write $\sim_\phi$ .

The indiscernibility relation $\sim_B$ is an equivalence relation on object descriptions.

**2.6. Lemma.** [9] *Let $\langle O, F\rangle$ be a perceptual system. For every $B \subseteq F$,*

$$\sim_B = \bigcap_{\phi \in B} \sim_\phi \,.$$

**2.7. Definition.** (*Equivalence Class*) Let $\langle O, F\rangle$ be a perceptual system and let $x \in O$ . For a set $B \subseteq F$ an equivalence class is defined as $x_{/\sim_B} = \{y \in O \mid y \sim_B x\}$ .

**2.8. Definition.** (*Quotient Set*)Let $\langle O, F\rangle$ be a perceptual system.For a set $B \subseteq F$ a quotient set is defined as

$$O_{/\sim_B} = \{x_{/\sim_B} \mid\ x \in O\}\,.$$

**2.9. Definition.** [9] Let $\langle O, F\rangle$ be a perceptual system. Then

$$\prod (O, F) := \bigcup_{B \subseteq F} O_{/\sim_B},$$

i.e., $\prod(O, F)$ is the family of equivalence classes of all indiscernibility relations determined by a perceptual information system $\langle O, F\rangle$ .

**2.10. Definition.** [9] (*Nearness relation*). Let $\langle O, F\rangle$ be a perceptual system and let $X, Y \subseteq O$. A set $X$ is near to a set $Y$ within the perceptual system $\langle O, F\rangle$ $(X \bowtie_F Y)$ iff there are $F_1, F_2 \subseteq F$ and $f \in F$ and there are $A \in O_{/\sim_{F_1}}, B \in O_{/\sim_{F_2}}, C \in O_{/\sim_f}$ such that $A \subseteq X, B \subseteq Y$ ve $A, B \subseteq C$.If a perceptual system is understood, then we say briefly that a set $X$ is near to a set $Y$.

**2.11. Definition.** [9] (*Perceptual near sets*) Let $\langle O, F\rangle$ be a perceptual system and let $X \subseteq O$. A set $X$ is a perceptual near set iff there is $Y \subseteq O$ such that $X \bowtie_F Y$. The family of near sets of a perceptual system $\langle O, F\rangle$ is denoted by $Near_F(O)$.

**2.12. Example.** Let $\langle O, F\rangle$ be a perceptual system such that $O = \{x_1, x_2, ..., x_6\}$, $F = \{\phi_1, \phi_2\}, \phi_1(x_1) = \phi_1(x_2) = \phi_1(x_3), \phi_1(x_4) = \phi_1(x_5) = \phi_1(x_6), \phi_1(x_1) \neq \phi_1(x_4)$ and $\phi_2(x_1) = \phi_2(x_2), \phi_2(x_3) = \phi_2(x_4), \phi_2(x_5) = \phi_2(x_6), \phi_2(x_1) \neq \phi_2(x_4) \neq \phi_2(x_5)$.

Thus $O_{/\sim_{\phi_1}} = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}\}$, $O_{/\sim_{\phi_2}} = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}\}$, $O_{/\sim_{\{\phi_1,\phi_2\}}} = \{\{x_1, x_2\}, \{x_3\}, \{x_4\}, \{x_5, x_6\}\}$.

Let $X = \{x_1, x_2, x_3, x_5\}, Y = \{x_2, x_4, x_5, x_6\}$. Thus there are $A = \{x_4\} \in O_{/\sim_{\{\phi_1,\phi_2\}}}$, $B = \{x_5, x_6\} \in O_{/\sim_{\phi_2}} C = (A \cup B) \in O_{/\sim_{\phi_1}}$ such that $A \subseteq X, B \subseteq Y$. Therefore $X \bowtie_F Y$.

**2.13. Proposition.** [9] *Let $\langle O, F\rangle$ be a perceptual system, $B \subseteq F$ and $x_{/\sim_B} \in O_{/\sim_B}$, where $|x_{/\sim_B}| \geq 2$. All elements belonging to a class $x_{/\sim_B}$ are near each other.*

**2.14. Proposition.** [9] *Let $\langle O, F\rangle$ be a perceptual system. For any $B \subseteq F$, every equivalence class of an indiscernibility relation $\sim_B$ is a near set .*

## 3. Some New Properties of Near Sets

In this section, we give some new propositions which are related to some propositions in [9].

**3.1. Proposition.** [9] *Let $\langle O, F\rangle$ be a perceptual system. For every $X \subseteq O$, the following conditions are equivalent:*

(1) $X \in Near_F(O)$,
(2) there is $A \in \prod(O, F)$ such that $A \subseteq X$,
(3) there is $A \in O_{/\sim_F}$ such that $A \subseteq X$ .

**3.2. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system and $X, Y \subseteq O$ . Then*

$$X \bowtie_F Y \Rightarrow X, Y \in Near_F(O) .$$

*Proof.* Let $X \bowtie_F Y$. From Definition 2.11, there are $A, B \in \prod(O, F)$ such that $A \subseteq X, B \subseteq Y$.Thus, from Proposition 3.1, $X, Y \in Near_F(O)$. □

**3.3. Remark.** From Proposition 3.2, two near sets may not be near to each other. We can see this in the following example.

**3.4. Example.** Let $\langle O, F\rangle$ be a perceptual system such that $O = \{x_1, x_2, ..., x_6\}$ ,for simplicity $F = (\phi)$ *and* $\phi(x_2) = \phi(x_3), \phi(x_4) = \phi(x_5) = \phi(x_6), \phi(x_1) \neq \phi(x_2) \neq \phi(x_4)$ .Thus $O_{/\sim_\phi} = \{\{x_1\}, \{x_2, x_3\}, \{x_4, x_5, x_6\}\}$. Let $X = \{x_1, x_2\}, Y = \{x_2, x_3, x_6\}$ .There are $A = \{x_1\} \in O_{/\sim\phi}, B = \{x_2, x_3\} \in O_{/\sim\phi}$ such that $A \subseteq X, B \subseteq Y$, so $X, Y \in Near_F(O)$. But there is no $C \in O_{/\sim\phi}$ such that $A, B \subseteq C$.Therefore $X$ and $Y$ are not near to each other.

**3.5. Proposition.** [9] *Let $\langle O, F\rangle$ be a perceptual system and $X, Y \subseteq O$ . Then*

$$X, Y \in Near_F(O) \Rightarrow X \cup Y \in Near_F(O),$$

i.e., the family of near sets of a perceptual system $\langle O, F\rangle$ is closed for the union of sets.

**3.6. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system and $X, Y \subseteq O$ . Then*

$$X \bowtie_F Y \Rightarrow X \cup Y \in Near_F(O).$$

*Proof.* It is clear from Proposition 3.2 and Proposition 3.5 . □

**3.7. Proposition.** [9] *Let $\langle O,F\rangle$ be a . Then*

$$X \in \prod(O,F) \Rightarrow X \bowtie_F X,$$

i.e., the relation $\bowtie_F$ is reflexive within the family $\prod(O,F)$ .

**3.8. Proposition.** *Let $\langle O,F\rangle$ be a perceptual system. Then*

$$X \bowtie_F X \;\Leftrightarrow \text{there is } A \in \prod(O,F) \text{ such that } \; A \subseteq X.$$

That is, a set $X \subseteq O$ to be near to itself need not be a equivalence class or need not be a union of equivalence classes. But at least it has to contain an equivalence class.

*Proof.* It is clear. □

**3.9. Proposition.** [9] *Let $\langle O,F\rangle$ be a perceptual system . For any $X,Y \subseteq O$, if there is $A \in \prod(O,F)$ such that $A \subseteq X \cap Y$, then $X \bowtie_F Y$.*

**3.10. Proposition.** *Let $\langle O,F\rangle$ be a perceptual system and let $X,Y \subseteq O$ and $F$ is a singleton set. Then*

$$X \bowtie_F Y \Leftrightarrow \text{there is } A \in \prod(O,F) \text{ such that } \; A \subseteq X \cap Y.$$

*Proof.* It is enough to prove the implication ($\Rightarrow$). From Definition 2.10, there are $A \in O_{/\sim_F}, B \in O_{/\sim_F}, C \in O_{/\sim_F}$ such that $A \subseteq X, B \subseteq Y$ and $A, B \subseteq C$. Since $F$ is a singleton set and $A, B \subseteq C$, then $A = B = C$. Therefore $A \subseteq X \cap Y$. □

**3.11. Proposition.** [9] *Let $\langle O,F\rangle$ be a perceptual system and let $X,Y,Z \subseteq O$. Then the following conditions hold:*

$$(1)\; X \bowtie_F Y \;\&Y \subseteq Z \Rightarrow X \bowtie_F Z,$$
$$(2)\; X \subseteq Y \;\&\; X \bowtie_F Z \Rightarrow Y \bowtie_F Z\ .$$

**3.12. Proposition.** *Let $\langle O,F\rangle$ be a perceptual system and $A_1,A_2,B_1,B_2 \subseteq O$.Then the following conditions hold:*

(1) $A_1 \bowtie_F A_2 \;\&B_1 \bowtie_F B_2 \Rightarrow (A_1 \cup B_1) \bowtie_F (A_2 \cup B_2)$ or $(A_1 \cup B_2) \bowtie_F (A_2 \cup B_1)$ ,
(2) $(A_1 \cap A_2) \bowtie_F (B_1 \cap B_2) \Rightarrow A_1 \bowtie_F B_1$ or $A_1 \bowtie_F B_2$ or $A_2 \bowtie_F B_1$ or $A_2 \bowtie_F B_2$.

*Proof.* Let $\langle O,F\rangle$ be a perceptual system and let $A_1,A_2,B_1,B_2 \subseteq O$.

Case ( 1). Let $A_1 \bowtie_F A_2$ and $B_1 \bowtie_F B_2$. So $A_1 \bowtie_F A_2$ , $A_2 \subseteq A_2 \cup B_2$ and $B_1 \bowtie_F B_2, B_2 \subseteq (A_2 \cup B_2)$ then from Proposition 3.11 (1) $A_1 \bowtie_F (A_2 \cup B_2)$ and $B_1 \bowtie_F (A_2 \cup B_2)$ . Since $A_1 \bowtie_F (A_2 \cup B_2)$ and $B_1 \bowtie_F (A_2 \cup B_2)$, $(A_1 \cup B_1) \bowtie_F (A_2 \cup B_2)$. Similarly it can be shown that $(A_1 \cup B_2) \bowtie_F (A_2 \cup B_1)$.

Case ( 2). Let $(A_1 \cap A_2) \bowtie_F (B_1 \cap B_2)$ . Since $(A_1 \cap A_2) \subseteq A_1$ and from Proposition 3.11 (2) $A_1 \bowtie_F (B_1 \cap B_2)$ . Since $A_1 \bowtie_F (B_1 \cap B_2)$ and from Proposition 3.11 (1), then $A_1 \bowtie_F B_1$. Similarly it can be shown that $A_2 \bowtie_F B_1$ or $A_2 \bowtie_F B_1$ or $A_2 \bowtie_F B_2$. □

The fact that the reverse of the implication reversed in Proposition 3.12 (1) does not hold is shown by example . Similarly it can be shown that the Proposition 3.12 (2) does not hold always.

**3.13. Example.** Let $\langle O,F\rangle$ be a perceptual system such that $O = \{x_1, x_2, ..., x_8\}$ ,so $O_{/\sim_F} = \{\{x_1,x_2,x_3\}, \{x_4,x_5\}, \{x_6,x_7,x_8\}\}$ .Let $A_1 = \{x_2,x_3,x_4\}$ , $A_2 = \{x_1,x_2,x_3,x_5\}$ , $B_1 = \{x_1,x_3,x_4,x_7\}$ , $B_2 = \{x_2,x_4,x_6,x_8\}$ ,so $A_1 \cup B_1 = \{x_1,x_2,x_3,x_4,x_7\}$ and $A_2 \cup B_2 = \{x_1,x_2,x_3,x_4,x_5,x_6,x_8\}$ . Since $\{x_1,x_2,x_3\} \in O_{/\sim_F}$ and $\{x_1,x_2,x_3\} \subseteq A_1 \cup B_1, A_2 \cup B_2$ $A_1 \cup B_1 \bowtie_F A_2 \cup B_2$.But there is no $X_{/\sim_F} \in O_{/\sim_F}, Y_{/\sim_F} \in O_{/\sim_F}, Z_{/\sim_F} \in O_{/\sim_F}$ such that $X_{/\sim_F} \subseteq A_1$, $Y_{/\sim_F} \subseteq A_2$ and $X, Y \subseteq Z$. Therefore, from Definition 2.10, $A_1$ and $A_2$ are not near to each other. For same reason, $B_1$ and $B_2$ are not near to each other.

## 4. Chain of Features, Nearness and Near Sets

In this section basically, a nested chain of probe functions (features) is formed and corresponding indiscernibility relation, nearness relation and near sets in $\langle O, F\rangle$ perceptual system are investigated.

**4.1. Definition.** *Let $\langle O, F\rangle$ be a perceptual system. Then*

$\prod(O, \sim_F) := \{\sim_B \mid B \subseteq F\}$,
i.e. $\prod(O, \sim_F)$ is the family of indiscernibility relations of all probe functions determined by a perceptual information system $\langle O, F\rangle$.

**4.2. Lemma.** *Let $\langle O, F\rangle$ be a perceptual system, $\prod(O, F)$ is the family of equivalence classes of all indiscernibility relations and $\prod(O, \sim_F)$ is the family of indiscernibility relations of all probe functions. Then for all $B \subseteq F$, the function*

$$f : \prod(O, \sim_F) \to \prod(O, F)$$
$$\sim_B \mapsto O_{/\sim_B}$$

is one-to-one and onto.

**4.3. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system and $F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$. Then for all $B_i \subseteq F$ , $1 \leq j, i \leq n$,*

$$B_j \subseteq B_i \Leftrightarrow \sim_{B_i} \subseteq \sim_{B_j} .$$

*Proof.* Let $B_i \subseteq F$ ,$B_j \subseteq B_i$, $1 \leq j, i \leq n$ . Since $\bigcap_{\phi \in B_j} \sim_\phi \subseteq \bigcap_{\phi \in B_i} \sim_\phi$ and, from Lemma 2.6, $\sim_{B_i} \subseteq \sim_{B_j}$ . □

**4.4. Corollary.** *Let $\langle O, F\rangle$ be a perceptual system and $F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$. Then for all $B_i \subseteq F$ ,$B_j \subseteq B_i$, $1 \leq j, i \leq n$ ,*

$$\sim_{B_i} \subseteq \sim_{B_j} \Leftrightarrow \bigcap_{\phi \in B_i} \sim_\phi \subseteq \bigcap_{\phi \in Bj} \sim_\phi .$$

**4.5. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system , $F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$ and $B_i \subseteq F$ ,$B_j \subseteq B_i$, $1 \leq j, i \leq n$ . Then*

$\sim_{B_i} \subseteq \sim_{B_j} \Rightarrow$ For all $A \in O_{/\sim_{B_i}}$ there is a unique $C \in O_{/\sim_{B_j}}$ such that $A \subseteq C$ .

*Proof.* Let $\sim_{B_i} \subseteq \sim_{B_j}$, $x \in O, A = x_{/\sim_{B_i}}$ and $C = x_{/\sim_{B_j}}$ .Since $\sim_{B_i} \subseteq \sim_{B_j}$, then $x_{/\sim_{B_i}} \subseteq x_{/\sim_{B_j}}$. □

**4.6. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system , $X \subseteq O, F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$ and $B_i \subseteq F$ ,$B_j \subseteq B_i$, $1 \leq j, i \leq n$ . Then the following conditions hold:*

(1) $\prod\left(O, \sim_{B_j}\right) \subseteq \prod(O, \sim_{B_i})$,
(2) $\prod(O, B_j) \subseteq \prod(O, B_i)$.

*Proof.* Let $\langle O, F\rangle$ be a perceptual system , $X \subseteq O$, $F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$ and $B_i \subseteq F$ ,$B_j \subseteq B_i$, $1 \leq j, i \leq n$ .
(1) Since $B \subseteq B_j$ then $B \subseteq B_i$ . Thus from Definition 4.1 $\prod\left(O, \sim_{B_j}\right) \subseteq \prod(O, \sim_{B_i})$.
(2) Since $B_j \subseteq B_i$, from Definition 2.9 $\prod(O, B_j) \subseteq \prod(O, B_i)$. □

**4.7. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system ,$F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$ and $B_i \subseteq F$ ,$B_j \subseteq B_i$, $1 \leq j, i \leq n$ . Then*

$$Near_{B_j}(O) \subseteq Near_{B_i}(O).$$

*Proof.* Let $X \subseteq O$ and $X \in Near_{B_j}(O)$. Since $X \in Near_{B_j}(O)$ there is $A \in \prod(O, \sim_{B_j})$ such that $A \subseteq X$. From Proposition 4.6 (1) $A \in \prod(O, \sim_{B_i})$. Therefore $X \in Near_{B_i}(O)$. □

The fact that the reverse of the implication reversed in Proposition 4.7 does not hold. We can see this in the next example.

**4.8. Example.** Let $\langle O, F\rangle$ be perceptual system in Example 2.12. Thus $O = \{x_1, x_2, ..., x_6\}$, $F = \{\phi_1, \phi_2\}$. Recall also that $O_{/\sim_{\phi_1}} = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}\}$, $O_{/\sim_{\phi_2}} = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}\}$, $O_{/\sim_{\{\phi_1,\phi_2\}}} = \{\{x_1, x_2\}, \{x_3\}, \{x_4\}, \{x_5, x_6\}\}$. Let $X \subseteq O$, $B_1, B_2 \subseteq F$ be defined as: $X = \{x_1, x_2, x_4\}$, $B_1 = \{\phi_1\}$, $B_2 = \{\phi_1, \phi_2\}$. Since $\{x_1, x_2\} \in O_{/\sim_{\{\phi_1,\phi_2\}}}$ and $\{x_1, x_2\} \subseteq X$, then $X \in Near_{B_2}(O)$. But there is no $A \in O_{/\sim_{\phi_1}}$ such that $A \subseteq X$, therefore $X \notin Near_{B_1}(O)$.

**4.9. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system ,$F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$, $X, Y \subseteq O$ and $B_i \subseteq F$ ,$B_j \subseteq B_i$, $1 \leq j, i \leq n$ .Then*

$X \bowtie_{B_j} Y \Rightarrow X \bowtie_{B_i} Y.$

*Proof.* Let $X \bowtie_{B_j} Y$. From Definition 2.10 there are $A, B, C \in \prod(O, B_j)$ such that $A \subseteq X, B \subseteq Y$ *and* $A, B \subseteq C$. Since $A, B, C \in \prod(O, B_j)$, then from Proposition 4.6 (2) $A, B, C \in \prod(O, B_i)$. Again from Definition 2.10, $X \bowtie_{B_i} Y$. □

**4.10. Definition.** Let $\langle O, F\rangle$ be a perceptual system ,$X, Y \subseteq O, F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$ and $B_i \subseteq F$. Then the expression

$X \bowtie_{\sim_{B_i}} Y$ means that: A set $X$ is near to a set $Y$ within the perceptual system $\langle O, F\rangle$ only for the $\sim_{B_i}$ relation.

**4.11. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system , $X, Y \subseteq O, F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$ and $B_i \subseteq F$ ,$B_j \subseteq B_i$, $1 \leq j, i \leq n$ . Then*

$$X \bowtie_{\sim_{B_j}} Y \Rightarrow X \bowtie_{\sim_{B_i}} Y.$$

*Proof.* Let $X \bowtie_{\sim_{B_j}} Y$. From Proposition 3.10 and Proposition 3.1, respectively, then $X \cap Y \in Near_{B_j}(O)$. Thus from Proposition 4.7, $X \cap Y \in Near_{B_i}(O)$. Therefore, from Proposition 3.10, then $X \bowtie_{\sim_{B_i}} Y$. □

**4.12. Example.** Let $\langle O, F\rangle$ be perceptual system in the Example 2.12. Recall also that $O_{/\sim_{\phi_2}} = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}\}$, $O_{/\sim_{\{\phi_1,\phi_2\}}} = \{\{x_1, x_2\}, \{x_3\}, \{x_4\}, \{x_5, x_6\}\}$. Let sets $X, Y \subseteq O$, $B_1, B_2 \subseteq F$ be defined as: $X = \{x_2, x_3, x_4\}, Y = \{x_3, x_4, x_6\}$ $B_1 = \{\phi_2\}, B_2 = \{\phi_1, \phi_2\}$. Since $\{x_3, x_4\} \in O_{/\sim_{\{\phi_2\}}}$ and $\{x_3, x_4\} \subseteq X, Y$ then $X \bowtie_{\sim_{B_1}} Y$. Since $\{x_4\} \in O_{/\sim_{\{\phi_1,\phi_2\}}}$ and $\{x_4\} \subseteq \{x_3, x_4\} \subseteq X, Y$ then $X \bowtie_{\sim_{B_2}} Y$.

**4.13. Definition.** Let $\langle O, F\rangle$ be a perceptual system and $F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$.

(4.1) $\quad B_1 \subseteq B_2 \subseteq ... \subseteq B_n$

Then the ascending subsets (4.1) is called as a chain of probe function sets or briefly a feature sets chain.

From Proposition 4.6, we can give following proposition.

**4.14. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system , $F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$ and $B_1 \subseteq B_2 \subseteq ... \subseteq B_n$ be a feature chain. Then the followings hold:*

$$(1)\ \prod(O,\sim_{B_1}) \subseteq \prod(O,\sim_{B_2}) \subseteq ... \subseteq \prod(O,\sim_F)$$
$$(2)\ \prod(O,B_1) \subseteq \prod(O,B_2) \subseteq ... \subseteq \prod(O,F).$$

**4.15. Definition.** Let $\langle O, F\rangle$ be a perceptual system and $F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$.

(4.2) $\quad \bowtie_{B_1} \subseteq \bowtie_{B_2} \subseteq ... \subseteq \bowtie_F$

The relation (4.2) corresponding to (4.1) is called as chain of a perceptual nearness or briefly nearness chain.

From Proposition 4.7 and Proposition 4.9 we can give following proposition.

**4.16. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system ,$F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$, $X, Y \subseteq O$ and $\bowtie_{B_1} \subseteq \bowtie_{B_2} \subseteq ... \subseteq \bowtie_F$ a nearness chain .Then the following conditions hold:*

$$(1)\ X \bowtie_{B_1} Y \Rightarrow X \bowtie_{B_2} Y \Rightarrow ... \Rightarrow X \bowtie_F Y$$
$$(2)\ Near_{B_1}(O) \subseteq Near_{B_2}(O) \subseteq ... \subseteq Near_F(O).$$

**4.17. Definition.** Let $\langle O, F\rangle$ be a perceptual system and $F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$.

(4.3) $\quad \sim_F \subseteq \sim_{B_{n-1}} \subseteq ... \subseteq \sim_{B_1}$

The relation (4.3) corresponding to (4.1) is called a chain of indiscernibility relations or briefly indiscernibility chain.

**4.18. Remark.** By using Definition 4.15 and Definition 4.17, we obtain $\bowtie_{\sim_{B_1}} \subseteq \bowtie_{\sim_{B_2}} \subseteq ... \subseteq \bowtie_{\sim_F}$. In fact, more than one indiscernibility chain can be formed. We can imagine this indiscernibility chain as a tree, i.e., a branching model which is formed by trunk, branch, thinner branch and so on, respectively. So we get a tree which has $n$ features in the trunk and 1 feature in the thinnest branch.

From Proposition 4.11 we can give following proposition.

**4.19. Proposition.** *Let $\langle O, F\rangle$ be a perceptual system , $X, Y \subseteq O, F = B_n = \{\phi_1, \phi_2, ..., \phi_n\}$ and $\bowtie_{\sim_{B_1}} \subseteq \bowtie_{\sim_{B_2}} \subseteq ... \subseteq \bowtie_{\sim_F}$ nearness chain .Then ,*
*$X \bowtie_{\sim_{B_1}} Y \subseteq X \bowtie_{\sim_{B_2}} Y \subseteq ... \subseteq X \bowtie_{\sim_F} Y.$*

**4.20. Remark.** There is a nuance between $X \bowtie_F Y$ and $X \bowtie_{\sim_F} Y$ . $X \bowtie_{\sim_F} Y$ implies that the sets $X$ and $Y$ near to each other with respect to only the $\sim_F$ indiscernebility relation in $\langle O, F\rangle$ perceptual system. However, $X \bowtie_F Y$ implies that the sets $X$ and $Y$ near to each other by means of Definition 2.10.

## References


[1] Henry,C. *Neighbourhoods, Classes and Near Sets*, Applied Mathematical Sciences, Vol. **5**, no. 35,1727-1732, 2011.
[2] Henry,C. *Near Sets: Theory and Aplications*, (Ph.D. Diss., supervisors: J. F. Peters, Dept. Elec. and Comp. Engg., U. of Manitoba, WPG, MB, Canada, 2010).
[3] İnan, E., Öztürk, M., Near groups on nearness approximation spaces, Hacettepe Journal of Mathematics and Statistics **4**, 545–558, 2012.
[4] Öztürk, M.,İnan, E., Soft nearness approximation spaces, Fund. Informaticae (2014) to appear.
[5] Pawlak, Z. and Skowron, A. *Rudiments of rough sets*, Inform. Sci. **177**, 3–27, 2007.
[6] Pawlak, Z. *Rough sets*, Int. J. Comput. Inform. Sci. **11** , 341–356, 1981.

[7] Pawlak, Z. *Rough sets – theoretical aspects of reasoning about data*, (Kluwer Academic Publishers, Dordrecht, 1991).

[8] Pawlak, Z. *Classification of objects by means of attributes*, Polish Acad. Sci., **429,** 1981.

[9] Peters, J. F. and Wasilewsk,P. *Foundations of near sets*, Information Sciences **179 ,** 3091–3109, 2009.

[10] Peters, J. F. *Classification of objects by means of features*, (in: Proceedings of the IEEE Symposium Series on Foundations of Computational Intelligence (IEEE SCCI 2007), Honolulu, Hawaii, 2007).

[11] Peters, J. F. *Near sets, General theory about nearness of objects*, Appl. Math. Sci. **1** (53) , 2029–2609, 2007.

[12] Peters, J. F. and Pal, S. K. *Cantor, Fuzzy, Near , and Rough Sets in Image Analysis*, (In: Pal, S.K. and Peters, J. F.(eds.) Rough Fuzzy Image Analysis: Foundations and Methodologies, CRC Pres ,Taylor and Francis Group, Boca Raton, U.S.A, 2010), 1.1-1.16.

[13] Peters, J.F. *Near sets, special theory about nearness of objects*, Fundam. Inform. **75** (1–4), 407–433, 2007.

[14] Peters, J. F. *Classification of perceptual objects by means of features*, Int. J. Info. Technol. Intell. Comput. **3** (2), 1–35, 2008.

[15] Peters, J. F. *Fuzzy Sets, Near Sets, and Rough Sets for Your Computational Intelligence Toolbo*x, Foundations of Comput. Intel. Vol. **2**, SCI 202, pp. 3-25, 2009.

[16] Peters, J. F. and Ramanna, R. *Feature selection: a near set approach*, (in: ECML & PKDD Workshop on Mining Complex Data, Warsaw, 2007), pp. 1–12.

[17] Peters, J. F., Topology of Digital Images. Visual Pattern Discovery in Proximity Spaces, Intelligent Systems Reference Library, Springer, ISBN 978-3-642-53844-5, vol. **63**, pp. 1-342, 2014.

[18] Peters, J. F., İnan, E., Öztürk, M., Spatial and descriptive isometries in proximity spaces, General Mathematics Notes **21** (2), 125–134, 2014.

[19] Peters, J. F., Naimpally, S., Applications of near sets, Notices of the Amer. Math. Soc. **59** (4), 536–542, DOI: http://dx.doi.org/10.1090/noti817, 2012.

[20] Taşbozan, H., İçen, İ., Bağırmaz, N., Özcan, A., Soft sets and soft topology on nearness approximation spaces, Fund. Informaticae (2014) to appear.

[21] Wolski, W. *Perception and classification. A note on near sets and rough sets*, Foundamenta Informaticae, **101** , 143-155, 2010.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 77 – 89

# $(k, s)$-Riemann-Liouville fractional integral and applications

Mehmet Zeki SARIKAYA[*†], Zoubir DAHMANI[‡], Mehmet Eyüp KIRIS[§] and Farooq AHMAD[¶]

**Abstract**

In this paper, we introduce a new approach on fractional integration, which generalizes the Riemann-Liouville fractional integral. We prove some properties for this new approach. We also establish some new integral inequalities using this new fractional integration.

**Keywords:** Riemann-Liouville fractional integrals, synchronous function, Chebyshev inequality, Hölder inequality.

*2000 AMS Classification:* 26A33, 26A51, 26D15.

*Received* : 21.07.2014 *Accepted* : 24.11.2014 *Doi* : 10.15672/HJMS.20164512484

## 1. Introduction

Fractional calculus and its widely application have recently been paid more and more attentions. For more recent development on fractional calculus, we refer the reader to [7, 12, 15, 16, 19]. There are several known forms of the fractional integrals of which two have been studied extensively for their applications [5, 10, 11, 14, 21]. The first is the Riemann-Liouville fractional integral of $\alpha \geq 0$ for a continuous function f on $[a, b]$ which is defined by

$$J_a^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_a^x (x-t)^{\alpha-1} f(t)dt, \quad \alpha \geq 0, \; a < x \leq b.$$

[*]Department of Mathematics, Faculty of Science and Arts, Düzce University, Düzce, Turkey, Email: sarikayamz@gmail.com

[†]Corresponding Author.

[‡]Laboratory of Pure and Applied Mathematics, UMAB, University of Mostaganem, Algeria, Email: zzdahmani@yahoo.fr

[§]Department of Mathematics, Faculty of Science and Arts, Afyon Kocatepe University, Afyon-TURKEY, Email: mkiris@gmail.com, kiris@aku.edu.tr

[¶]Mathematics Department, Colleege of Science, Majmaah University, Al-Zulfi, Kingdom of Saudi Arabia, Email: farooqgujar@gmail.com, f.ishaq@mu.edu.sa

This integral is motivated by the well known Cauchy formula:

$$\int_a^x dt_1 \int_a^{t_1} dt_2 ... \int_a^{t_{n-1}} f(t_n) dt_n = \frac{1}{\Gamma(n)} \int_a^x (x-t)^{n-1} f(t) dt, n \in \mathbb{N}^*.$$

The second is the Hadamard fractional integral introduced by Hadamard [9]. It is given by:

$$J_a^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_a^x \left(\log \frac{x}{t}\right)^{\alpha-1} f(t) \frac{dt}{t}, \quad \alpha > 0, \; x > a.$$

The Hadamard integral is based on the generalization of the integral

$$\int_a^x \frac{dt_1}{t_1} \int_a^{t_1} \frac{dt_2}{t_2} ... \int_a^{t_{n-1}} \frac{f(t_n)}{t_n} dt_n = \frac{1}{\Gamma(\alpha)} \frac{1}{\Gamma(\alpha)} \int_a^x \left(\log \frac{x}{t}\right)^{n-1} f(t) \frac{dt}{t}$$

for $n \in \mathbb{N}^*$.

In [10], Katugampola gave a new fractional integration which generalizes both the Riemann-Liouville and Hadamard fractional integrals into a single form. This generalization is based on the observation that, for $n \in \mathbb{N}^*$,

$$\int_a^x t_1^s dt_1 \int_a^{t_1} t_2^s dt_2 ... \int_a^{t_{n-1}} t_n^s f(t_n) dt_n = \frac{(s+1)^{1-n}}{\Gamma(\alpha)} \int_a^x \left(x^{s+1} - t^{s+1}\right)^{n-1} t^s f(t) dt,$$

which gives the following fractional version

$${}^s J_a^\alpha f(x) = \frac{(s+1)^{1-n}}{\Gamma(\alpha)} \int_a^x \left(x^{s+1} - t^{s+1}\right)^{\alpha-1} t^s f(t) dt,$$

where $\alpha$ and $s \neq -1$ are real numbers.

Recently, in [6], Diaz and Pariguan have defined new functions called $k$-gamma and $k$-beta functions and the Pochhammer $k$-symbol that is respectively generalization of the classical gamma and beta functions and the classical Pochhammer symbol:

$$\Gamma_k(x) = \lim_{n \to \infty} \frac{n! k^n (nk)^{\frac{x}{k}-1}}{(x)_{n,k}}, \; k > 0,$$

where $(x)_{n,k}$ is the Pochhammer $k$-symbol for factorial function. It has been shown that the Mellin transform of the exponential function $e^{-\frac{t^k}{k}}$ is the $k$-gamma function, explicitly given by

$$\Gamma_k(x) = \int_0^\infty t^{x-1} e^{-\frac{t^k}{k}} dt, x > 0.$$

Clearly, $\Gamma(x) = \lim_{k \to 1} \Gamma_k(x), \; \Gamma_k(x) = k^{\frac{x}{k}-1} \Gamma(\frac{x}{k})$ and $\Gamma_k(x+k) = x\Gamma_k(x)$. Furthermore, $k$-beta function is defined as follows

$$B_k(x,y) = \frac{1}{k} \int_0^1 t^{\frac{x}{k}-1} (1-t)^{\frac{y}{k}-1} dt,$$

so that $B_k(x,y) = \frac{1}{k} B(\frac{x}{k}, \frac{y}{k})$ and $B_k(x,y) = \frac{\Gamma_k(x)\Gamma_k(y)}{\Gamma_k(x+y)}$.

Later, under the above definitions, in [13], Mubeen and Habibullah have introduced the $k$-fractional integral of the Riemann-Liouville type as follows:

$${}_k J^\alpha f(x) = \frac{1}{k\Gamma_k(\alpha)} \int_0^x (x-t)^{\frac{\alpha}{k}-1} f(t) dt, \quad \alpha > 0, \; x > 0.$$

Note that when $k \to 1$, then it reduces to the classical Riemann-liouville fractional integral.

## 2. $(k,s)$-Riemann-Liouville fractional integral

In this section, we present the $(k,s)$ fractional integration which generalizes all of the above Riemann-Liouville fractional integrals as follows:

**2.1. Definition.** Let $f$ be a continuous function on on a the finite real interval $[a,b]$. Then $(k,s)$-Riemann-Liouville fractional integral of $f$ of order $\alpha > 0$ is defined by:

$$ {}_{k}^{s}J_{a}^{\alpha}f(x) := \frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)} \int_a^x \left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1} t^s f(t)dt, \quad x \in [a,b], \tag{2.1}$$

where $k > 0, s \in \mathbb{R}\backslash\{-1\}$.

In the following theorem, we prove that the $(k,s)$ fractional integral is well defined:

**2.2. Theorem.** *Let $f \in L_1[a,b], s \in \mathbb{R}\backslash\{-1\}$ and $k > 0$. Then ${}_{k}^{s}J_{a}^{\alpha}f(x)$ exists for any $x \in [a,b], \alpha > 0$.*

*Proof.* Let $\Delta := [a,b]\times[a,b]$ and $P : \Delta \to \mathbb{R}$ ; $P(x,t) = \left[\left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1} t^s\right]$. It clear to see that $P = P_+ + P_-$, where

$$P_+(x,t) := \begin{cases} \left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1} t^s & , a \le t \le x \le b \\ 0 & , a \le x \le t \le b \end{cases}$$

and

$$P_-(x,t) := \begin{cases} \left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1} x^s & , a \le t \le x \le b \\ 0 & , a \le x \le t \le b. \end{cases}$$

Since $P$ is measurable on $\Delta$, then we can write

$$\int_a^b P(x,t)dt = \int_a^x P(x,t)dt = \int_a^x \left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1} t^s dt = \frac{k}{\alpha}\left(x^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}}.$$

By using the repeated integral, we obtain

$$\begin{aligned} \int_a^b \left(\int_a^b P(x,t)\,|f(x)|\,dt\right)dx &= \int_a^b |f(x)| \left(\int_a^b P(x,t)dt\right)dx \\ &= \frac{k}{\alpha}\int_a^b \left(x^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}} |f(x)|\,dx \\ &\le \frac{k}{\alpha}\left(b^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}} \int_a^b |f(x)|\,dx. \end{aligned}$$

That is

$$\begin{aligned} \int_a^b \left(\int_a^b P(x,t)\,|f(x)|\,dt\right)dx &= \int_a^b |f(x)| \left(\int_a^b P(x,t)dt\right)dx \\ &\le \frac{k}{\alpha}\left(b^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}} \|f(x)\|_{L_1[a,b]} < \infty. \end{aligned}$$

Therefore, the function $Q : \Delta \to \mathbb{R};\ Q(x,t) := P(x,t)f(x)$ is integrable over $\Delta$ by Tonelli's theorem. Hence, by Fubini's theorem $\int_a^b P(x,t)f(x)dx$ is an integrable function on $[a,b]$, as a function of $t \in [a,b]$. That is, ${}_{k}^{s}J_{a}^{\alpha}f(x)$ exists. □

Now, we prove the commutativity and the semigroup properties of the $(k,s)$-Riemann-Liouville fractional integral. We have:

**2.3. Theorem.** *Let $f$ be continuous on $[a,b], k>0$ and $s\in\mathbb{R}\backslash\{-1\}$. Then,*

$$ {}_k^sJ_a^{\alpha}\left[\,{}_k^sJ_a^{\beta}f(x)\right]= {}_k^sJ_a^{\alpha+\beta}f(x)= {}_k^sJ_a^{\beta}\left[\,{}_k^sJ_a^{\alpha}f(x)\right], $$

*for all $\alpha>0,\beta>0,x\in[a,b]$.*

*Proof.* Thanks to Definition 1 and by Dirichlet's formula, we have

$$ {}_k^sJ_a^{\alpha}\left[\,{}_k^sJ_a^{\beta}f(x)\right]=\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^x\left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1}t^s\,{}_k^sJ_a^{\beta}f(t)dt $$

$$ =\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^x\left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1}t^s\left[\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-\tau^{s+1}\right)^{\frac{\beta}{k}-1}\tau^sf(\tau)d\tau\right]dt $$

$$ =\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^x\left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1}t^s\left[\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-\tau^{s+1}\right)^{\frac{\beta}{k}-1}\tau^sf(\tau)d\tau\right]dt. $$

That is

(2.2)

$$ {}_k^sJ_a^{\alpha}\left[\,{}_k^sJ_a^{\beta}f(x)\right]=\frac{(s+1)^{2-\frac{\alpha+\beta}{k}}}{k^2\Gamma_k(\alpha)\Gamma_k(\beta)}\int_a^x\tau^sf(\tau)\left[\int_\tau^x\left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1}t^s\left(t^{s+1}-\tau^{s+1}\right)^{\frac{\beta}{k}-1}dt\right]d\tau. $$

Using the change of variable $y=\left(t^{s+1}-\tau^{s+1}\right)/\left(x^{s+1}-\tau^{s+1}\right)$, we can write

(2.3)

$$ \int_\tau^x\left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1}\left(t^{s+1}-\tau^{s+1}\right)^{\frac{\beta}{k}-1}t^sdt\quad=\frac{\left(x^{s+1}-\tau^{s+1}\right)^{\frac{\alpha+\beta}{k}-1}}{s+1}\int_0^1(1-y)^{\frac{\alpha}{k}-1}y^{\frac{\beta}{k}-1}dy $$

$$ =\frac{\left(x^{s+1}-\tau^{s+1}\right)^{\frac{\alpha+\beta}{k}-1}}{s+1}\int_0^1(1-y)^{\frac{\alpha}{k}-1}y^{\frac{\beta}{k}-1}dy\qquad=\frac{\left(x^{s+1}-\tau^{s+1}\right)^{\frac{\alpha+\beta}{k}-1}}{s+1}kB_k(\alpha,\beta). $$

According to the $k$-beta function and by (2.2) and (2.3), we obtain

$$ \begin{aligned} {}_k^sJ_a^{\alpha}\left[\,{}_k^sJ_a^{\beta}f(x)\right] &= \frac{(s+1)^{1-\frac{\alpha+\beta}{k}}}{k\Gamma_k(\alpha+\beta)}\int_a^x\left(x^{s+1}-\tau^{s+1}\right)^{\frac{\alpha+\beta}{k}-1}\tau^sf(\tau)d\tau \\ &= {}_k^sJ_a^{\alpha+\beta}f(x). \end{aligned} $$

This completes the proof of the Theorem 2.3. □

**2.4. Theorem.** *Let $\alpha,\beta>0,k>0$ and $s\in\mathbb{R}\backslash\{-1\}$. Then, we have*

$$ {}_k^sJ_a^{\alpha}\left[\left(x^{s+1}-a^{s+1}\right)^{\frac{\beta}{k}-1}\right]=\frac{\Gamma_k(\beta)}{(s+1)^{\frac{\alpha}{k}}\Gamma_k(\alpha+\beta)}\left(x^{s+1}-a^{s+1}\right)^{\frac{\alpha+\beta}{k}-1}, \tag{2.4} $$

*where $\Gamma_k$ denotes the $k$-gamma function.*

*Proof.* By Definition 1 and using the change of variable $y = \left(x^{s+1} - t^{s+1}\right) / \left(x^{s+1} - a^{s+1}\right); x \in ]a, b]$, we get

$$\begin{aligned} {}_k^s J_a^\alpha \left[ \left(x^{s+1} - a^{s+1}\right)^{\frac{\beta}{k}-1} \right] &= \frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)} \int_a^x \left(x^{s+1} - t^{s+1}\right)^{\frac{\alpha}{k}-1} t^s \left(t^{s+1} - a^{s+1}\right)^{\frac{\alpha+\beta}{k}-1} dt \\ &= \frac{(s+1)^{-\frac{\alpha}{k}} \left(x^{s+1} - a^{s+1}\right)^{\frac{\alpha+\beta}{k}-1}}{k\Gamma_k(\alpha)} \int_0^1 (1-y)^{\frac{\alpha}{k}-1} y^{\frac{\beta}{k}-1} dy \\ &= \frac{\left(x^{s+1} - a^{s+1}\right)^{\frac{\alpha+\beta}{k}-1}}{(s+1)^{\frac{\alpha}{k}} \Gamma_k(\alpha)} B_k(\alpha, \beta). \end{aligned}$$

The case $a = x$ is trivial. The proof of Theorem 2.4 is complete. □

**2.5. Remark.** $(i:)$ Taking $s = 0, k > 0$ in (2.4), we obtain

$$ {}_k J_a^\alpha \left[ (x-a)^{\frac{\beta}{k}-1} \right] = \frac{\Gamma_k(\beta)}{\Gamma_k(\alpha+\beta)} (x-a)^{\frac{\alpha+\beta}{k}-1}. \tag{2.5}$$

$(ii:)$ The formula (2.4) for $s = 0, k = 1$ becomes

$$J_a^\alpha \left[ (x-a)^{\beta-1} \right] = \frac{\Gamma(\beta)}{\Gamma(\alpha+\beta)} (x-a)^{\alpha+\beta-1}.$$

**2.6. Corollary.** *Let $k > 0$ and $s \in \mathbb{R}\backslash\{-1\}$. Then the formula*

$$ {}_k^s J_a^\alpha(1) = \frac{1}{(s+1)^{\frac{\alpha}{k}} \Gamma_k(\alpha+k)} \left(x^{s+1} - a^{s+1}\right)^{\frac{\alpha}{k}-2} \tag{2.6}$$

*is valid for any $\alpha > 0$.*

**2.7. Remark.** $(a:)$ For $s = 0, k > 0$ in (2.6), we get

$$ {}_k J_a^\alpha(1) = \frac{1}{\Gamma_k(\alpha+k)} (x-a)^{\frac{\alpha}{k}-2}. \tag{2.7}$$

$(b:)$ For $s = 0, k = 1$ we have

$$J_a^\alpha(1) = \frac{1}{\Gamma(\alpha+1)} (x-a)^{\alpha-2}.$$

## 3. Some new $(k, s)$-Riemann-Liouville fractional integral inequalities

Chebyshev inequalities can be represented in $(k, s)$-fractional integral forms as follows:

**3.1. Theorem.** *Let $f$ and $g$ be two synchronous on $[0, \infty)$. Then for all $t > a \geq 0$, $\alpha > 0$, $\beta > 0$, the following inequalities for $(k, s)$-fractional integrals hold:*

$$ {}_k^s J_a^\alpha fg(t) \geq \frac{1}{J_a^\alpha(1)} \, {}_k^s J_a^\alpha f(t) \, {}_k^s J_a^\alpha g(t) \tag{3.1}$$

$$ {}_k^s J_a^\alpha fg(t) \, {}_k^s J_a^\beta(1) + {}_k^s J_a^\beta fg(t) \, {}_k^s J_a^\alpha(1) \geq {}_k^s J_a^\alpha f(t) \, {}_k^s J_a^\beta g(t) + {}_k^s J_a^\alpha g(t) \, {}_k^s J_a^\beta f(t). \tag{3.2}$$

*Proof.* Since the functions $f$ and $g$ are synchronous on $[0, \infty)$, then for all $x, y \geq 0$, we have

$$(f(x) - f(y))(g(x) - g(y)) \geq 0.$$

Therefore

$$f(x) g(x) + f(y) g(y) \geq f(x) g(y) + f(y) g(x). \tag{3.3}$$

Multiplying both sides of (3.3) by $\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s$, then integrating the resulting inequality with respect to $x$ over $(a,t)$, we obtain

$$\begin{aligned} & \frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s f(x)\,g(x)\,dx \\ & +\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s f(y)\,g(y)\,dx \\ \geq\ & \frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s f(x)\,g(y)\,dx \\ & +\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s f(y)\,g(x)\,dx, \end{aligned}$$

i.e.

$$ {}_k^sJ_a^\alpha fg(t)+f(y)\,g(y)\ {}_k^sJ_a^\alpha(1)\geq g(y)\ {}_k^sJ_a^\alpha f(t)+f(y)\ {}_k^sJ_a^\alpha g(t). \tag{3.4}$$

Multiplying both sides of (3.3) by $\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\left(t^{s+1}-y^{s+1}\right)^{\frac{\alpha}{k}-1}y^s$, then integrating the resulting inequality with respect to $y$ over $(a,t)$, we obtain

$$\begin{aligned} & {}_k^sJ_a^\alpha fg(t)\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-y^{s+1}\right)^{\frac{\alpha}{k}-1}y^s dy \\ & +\ {}_k^sJ_a^\alpha(1)\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-y^{s+1}\right)^{\frac{\alpha}{k}-1}y^s f(y)\,g(y)\,dy \\ \geq\ & {}_k^sJ_a^\alpha f(t)\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-y^{s+1}\right)^{\frac{\alpha}{k}-1}y^s g(y)dy \\ & +\ {}_k^sJ_a^\alpha g(t)\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-y^{s+1}\right)^{\frac{\alpha}{k}-1}y^s f(y)\,dy, \end{aligned}$$

that is

$$ {}_k^sJ_a^\alpha fg(t)\geq\frac{1}{J_a^\alpha(1)}\ {}_k^sJ_a^\alpha f(t)\ {}_k^sJ_a^\alpha g(t).$$

The first inequality is thus proved.

Multiplying both sides of (3.3) by $\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\alpha)}\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s$, then integrating the resulting inequality with respect to $y$ over $(a,t)$, we obtain

$$\begin{aligned} & {}_k^sJ_a^\alpha fg(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s dy \\ & +\ {}_k^sJ_a^\alpha(1)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s f(y)\,g(y)\,dy \\ \geq\ & {}_k^sJ_a^\alpha f(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s g(y)dy \\ & +\ {}_k^sJ_a^\alpha g(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\alpha)}\int_a^t \left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s f(y)\,dy, \end{aligned}$$

that is

$$ {}_k^sJ_a^{\alpha} fg(t)\ {}_k^sJ_a^{\beta}(1) + {}_k^sJ_a^{\beta} fg(t)\ {}_k^sJ_a^{\alpha}(1) \geq {}_k^sJ_a^{\alpha} f(t)\ {}_k^sJ_a^{\beta} g(t) + {}_k^sJ_a^{\alpha} g(t)\ {}_k^sJ_a^{\beta} f(t) $$

and the second inequality is proved. The proof is completed. □

**3.2. Theorem.** *Let $f$ and $g$ be two synchronous on $[0,\infty)$, $h \geq 0$. Then for all $t > a \geq 0$, $\alpha > 0$, $\beta > 0$, the following inequalities hold:*

$$ \begin{aligned} & \frac{1}{(s+1)^{\frac{\beta}{k}}\Gamma_k(\beta+k)}\left(t^{s+1}-a^{s+1}\right)^{\frac{\beta}{k}-2}\ {}_k^sJ_a^{\alpha} fgh(t) \\ & + \frac{1}{(s+1)^{\frac{\alpha}{k}}\Gamma_k(\alpha+k)}\left(t^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}-2}\ {}_k^sJ_a^{\beta} fgh(t) \\ \geq\ & {}_k^sJ_a^{\alpha} fh(t)\ {}_k^sJ_a^{\beta} g(t) + {}_k^sJ_a^{\alpha} gh(t)\ {}_k^sJ_a^{\beta} f(t) - {}_k^sJ_a^{\alpha} h(t)\ {}_k^sJ_a^{\beta} fg(t) - {}_k^sJ_a^{\alpha} fg(t)\ {}_k^sJ_a^{\beta} h(t) \\ & + {}_k^sJ_a^{\alpha} f(t)\ {}_k^sJ_a^{\beta} gh(t) + {}_k^sJ_a^{\alpha} g(t)\ {}_k^sJ_a^{\beta} fh(t). \end{aligned} $$

*Proof.* Since the functions $f$ and $g$ are synchronous on $[0,\infty)$ and $h \geq 0$, then for all $x, y \geq 0$, we have

$$ \left(f(x) - f(y)\right)\left(g(x) - g(y)\right)\left(h(x) + h(y)\right) \geq 0. $$

Hence,

$$ \begin{aligned} & f(x)\,g(x)\,h(x) + f(y)\,g(y)\,h(y) \\ \geq\ & f(x)\,g(y)\,h(x) + f(y)\,g(x)\,h(x) - f(y)\,g(y)\,h(x) \\ & - f(x)\,g(x)\,h(y) + f(x)\,g(y)\,h(y) + f(y)\,g(x)\,h(y)\,. \end{aligned} \tag{3.5} $$

Multiplying both sides of (3.5) by $\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s$, then integrating the resulting inequality with respect to $x$ over $(a,t)$, we obtain

$$
\begin{aligned}
& \frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s f(x)\,g(x)\,h(x)\,dx \\
& +f(y)\,g(y)\,h(y)\,\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s dx \\
\geq\ & g(y)\,\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s f(x)\,h(x)\,dx \\
& +f(y)\,\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s g(x)\,h(x)\,dx \\
& -f(y)\,g(y)\,\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s h(x)\,dx \\
& -h(y)\,\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s f(x)\,g(x)\,dx \\
& +g(y)\,h(y)\,\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s f(x)\,dx \\
& +f(y)\,h(y)\,\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\int_a^t\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s g(x)\,dx.
\end{aligned}
$$

That is

$$
\begin{aligned}
(3.6)\qquad & {}_k^sJ_a^\alpha fgh(t)+f(y)\,g(y)\,h(y)\ {}_k^sJ_a^\alpha(1) \\
\geq\ & g(y)\ {}_k^sJ_a^\alpha fh(t)+f(y)\ {}_k^sJ_a^\alpha gh(t)-f(y)\,g(y)\ {}_k^sJ_a^\alpha h(t)-h(y)\ {}_k^sJ_a^\alpha fg(t) \\
& +g(y)\,h(y)\ {}_k^sJ_a^\alpha f(t)+f(y)\,h(y)\ {}_k^sJ_a^\alpha g(t).
\end{aligned}
$$

Multiplying both sides of (3.6) by $\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s$, then integrating the resulting inequality with respect to $y$ over $(a,t)$, we obtain

$$\begin{aligned}
& {}_k^sJ_a^\alpha fgh(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s dy \\
& +{}_k^sJ_a^\alpha(1)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s f(y)\,g(y)\,h(y)dy \\
\geq\ & {}_k^sJ_a^\alpha fh(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s g(y)\,dy \\
& +{}_k^sJ_a^\alpha gh(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s f(y)\,dy \\
& -{}_k^sJ_a^\alpha h(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s f(y)\,g(y)\,dy \\
& -{}_k^sJ_a^\alpha fg(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s h(y)dy \\
& +{}_k^sJ_a^\alpha f(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s g(y)\,h(y)\,dy \\
& +{}_k^sJ_a^\alpha g(t)\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\int_a^t\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s f(y)\,h(y)\,dy.
\end{aligned}$$

Consequently,

$$\begin{aligned}
{}_k^sJ_a^\alpha fgh(t){}_k^sJ_a^\beta(1)+{}_k^s J_a^\alpha(1){}_k^sJ_a^\beta fgh(t) \geq\ & {}_k^sJ_a^\alpha fh(t)\ {}_k^sJ_a^\beta g(t)+\ {}_k^sJ_a^\alpha gh(t)\ {}_k^sJ_a^\beta f(t) \\
& -\ {}_k^sJ_a^\alpha h(t)\ {}_k^sJ_a^\beta fg(t)-\ {}_k^sJ_a^\alpha fg(t)\ {}_k^sJ_a^\beta h(t) \\
& +\ {}_k^sJ_a^\alpha f(t)\ {}_k^sJ_a^\beta gh(t)+\ {}_k^sJ_a^\alpha g(t)\ {}_k^sJ_a^\beta fh(t).
\end{aligned}$$

The proof is thus complete. □

**3.3. Corollary.** *Let $f$ and $g$ be two synchronous on $[0,\infty)$, $h\geq 0$. Then for all $t>a\geq 0$, $\alpha>0$, the following inequalities hold:*

$$\begin{aligned}
& \frac{1}{(s+1)^{\frac{\alpha}{k}}\Gamma_k(\alpha+k)}\left(t^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}-2}\ {}_k^sJ_a^\alpha fgh(t) \\
\geq\ & {}_k^sJ_a^\alpha fh(t)\ {}_k^sJ_a^\alpha g(t)+\ {}_k^sJ_a^\alpha gh(t)\ {}_k^sJ_a^\alpha f(t)-\ {}_k^sJ_a^\alpha h(t)\ {}_k^sJ_a^\alpha fg(t).
\end{aligned}$$

**3.4. Theorem.** *Let $f,g$ and $h$ be three monotonic functions defined on $[0,\infty)$ satisfying the following*

$$(f(x)-f(y))\,(g(x)-g(y))\,(h(x)-h(y))\geq 0$$

*for all* $x, y \in [a,t]$*. Then for all* $t > a \geq 0,\ \alpha > 0,\ \beta > 0,$ *the following inequalities are valid:*

$$\frac{1}{(s+1)^{\frac{\beta}{k}}\Gamma_k(\beta+k)}\left(t^{s+1}-a^{s+1}\right)^{\frac{\beta}{k}-2}\ {}_k^sJ_a^{\alpha}fgh(t)$$
$$-\frac{1}{(s+1)^{\frac{\alpha}{k}}\Gamma_k(\alpha+k)}\left(t^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}-2}\ {}_k^sJ_a^{\beta}fgh(t)$$
$$\geq\ {}_k^sJ_a^{\alpha}fh(t)\ {}_k^sJ_a^{\beta}g(t)+\ {}_k^sJ_a^{\alpha}gh(t)\ {}_k^sJ_a^{\beta}f(t)-\ {}_k^sJ_a^{\alpha}h(t)\ {}_k^sJ_a^{\beta}fg(t)+\ {}_k^sJ_a^{\alpha}fg(t)\ {}_k^sJ_a^{\beta}h(t)$$
$$-{}_k^sJ_a^{\alpha}f(t)\ {}_k^sJ_a^{\beta}gh(t)-\ {}_k^sJ_a^{\alpha}g(t)\ {}_k^sJ_a^{\beta}fh(t).$$

*Proof.* We use the same arguments as in the proof of Theorem 3.2. □

**3.5. Theorem.** *Let* $f$ *and* $g$ *be two functions on* $[0,\infty)$*. Then for all* $t > a \geq 0,\ \alpha > 0,\ \beta > 0,$ *the following inequalities for* $(k,s)$*-fractional integrals hold:*

$$\frac{1}{(s+1)^{\frac{\beta}{k}}\Gamma_k(\beta+k)}\left(t^{s+1}-a^{s+1}\right)^{\frac{\beta}{k}-2}\ {}_k^sJ_a^{\alpha}f^2(t)$$
$$+\frac{1}{(s+1)^{\frac{\alpha}{k}}\Gamma_k(\alpha+k)}\left(t^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}-2}\ {}_k^sJ_a^{\beta}g^2(t)$$
$$\geq\ 2\ {}_k^sJ_a^{\alpha}f(t)\ {}_k^sJ_a^{\beta}g(t) \tag{3.7}$$

$${}_k^sJ_a^{\alpha}f^2(t)\ {}_k^sJ_a^{\beta}g^2(t)+\ {}_k^sJ_a^{\beta}f^2(t)\ {}_k^sJ_a^{\alpha}g^2(t)\geq 2\ {}_k^sJ_a^{\alpha}fg(t)\ {}_k^sJ_a^{\beta}fg(t). \tag{3.8}$$

*Proof.* Since

$$(f(x)-g(y))^2\geq 0,$$

then, we have

$$f^2(x)+g^2(y)\geq 2f(x)g(y). \tag{3.9}$$

Multiplying both sides of (3.9) by $\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}\left(t^{s+1}-x^{s+1}\right)^{\frac{\alpha}{k}-1}x^s$ and $\frac{(s+1)^{1-\frac{\beta}{k}}}{k\Gamma_k(\beta)}\left(t^{s+1}-y^{s+1}\right)^{\frac{\beta}{k}-1}y^s$, then integrating the resulting inequality with respect to $x$ and $y$ over $(a,t)$ respectively, we obtain (3.7).

On the other hand, since

$$(f(x)g(y)-f(y)g(x))^2\geq 0,$$

then, with the same arguments as before, we obtain (3.8). □

**3.6. Corollary.** *Let* $f$ *and* $g$ *be two functions on* $[0,\infty)$*, then for all* $t > a \geq 0,\ \alpha > 0,$ *the following inequalities are valid:*

$$\frac{1}{(s+1)^{\frac{\alpha}{k}}\Gamma_k(\alpha+k)}\left(t^{s+1}-a^{s+1}\right)^{\frac{\alpha}{k}-2}\left[{}_k^sJ_a^{\alpha}f^2(t)+\ {}_k^sJ_a^{\beta}g^2(t)\right]$$
$$\geq\ 2\ {}_k^sJ_a^{\alpha}f(t)\ {}_k^sJ_a^{\alpha}g(t)$$
$${}_k^sJ_a^{\alpha}f^2(t)\ {}_k^sJ_a^{\alpha}g^2(t)\geq\ \left[{}_k^sJ_a^{\alpha}fg(t)\right]^2.$$

**3.7. Theorem.** *Let* $f:\mathbb{R}\to\mathbb{R}$ *with:*

$$\bar{f}(x):=\int_a^x t^s f(t)dt,\quad x>a\geq 0,\ s\in\mathbb{R}\backslash\{-1\}.$$

*Then, for $\alpha \geq k > 0$ we have:*

$$ {}_k^s J_a^\alpha f(x) = \frac{1}{k}\ {}_k^s J_a^{\alpha-k} \bar{f}(x) $$

*Proof.* By definition of the $(k,s)$-fractional integral and by using Dirichlet's formula, we have

$$\begin{aligned} {}_k^s J_a^\alpha \bar{f}(x) &= \frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)} \int_a^x \left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1} t^s \int_a^t u^s f(u)dudt \\ &= \frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)} \int_a^x u^s f(u) \int_u^x \left(x^{s+1}-t^{s+1}\right)^{\frac{\alpha}{k}-1} t^s dtdu \\ &= \frac{(s+1)^{-\frac{\alpha}{k}}}{\Gamma_k(\alpha+k)} \int_a^x \left(x^{s+1}-u^{s+1}\right)^{\frac{\alpha}{k}} u^s f(u)du \\ &= k\ {}_k^s J_a^{\alpha+k} f(x). \end{aligned}$$

This completes the proof of Theorem 3.7. □

We give the generalized Cauchy-Buniakovsky-Schwarz inequality as follows:

**3.8. Lemma.** *Let $f, g, h : [a,b] \to (0,\infty)$ be three functions $0 \leq a < b$. Then*

(3.10)
$$ \left(\int_a^b g^m(t)h^x(t)f(t)dt\right)\left(\int_a^b g^n(t)h^y(t)f(t)dt\right) \geq \left(\int_a^b g^{\frac{m+n}{2}}(t)h^{\frac{x+y}{2}}(t)f(t)dt\right)^2, $$

*where $m, n, x, y$ arbitrary real numbers.*

*Proof.*

$$ \int_a^b \left[\sqrt{g^m(t)h^x(t)f(t)}\sqrt{\int_a^b g^n(t)h^y(t)f(t)dt} - \sqrt{g^n(t)h^y(t)f(t)}\sqrt{\int_a^b g^m(t)h^x(t)f(t)dt}\right]^2 dt \geq 0 $$

$$\begin{aligned} &\int_a^b \Bigg[g^m(t)h^x(t)f(t)\int_a^b g^n(t)h^y(t)f(t)dt + g^n(t)h^y(t)f(t)\int_a^b g^m(t)h^x(t)f(t)dt \\ &-2\, g^{\frac{m+n}{2}}(t)h^{\frac{x+y}{2}}(t)f(t)\sqrt{\int_a^b g^m(t)h^x(t)f(t)dt}\sqrt{\int_a^b g^n(t)h^y(t)f(t)dt}\Bigg] dt \\ \geq\ & 0 \\ & 2\left(\int_a^b g^m(t)h^x(t)f(t)dt\right)\left(\int_a^b g^n(t)h^y(t)f(t)dt\right) \\ \geq\ & 2\left(\int_a^b g^{\frac{m+n}{2}}(t)h^{\frac{x+y}{2}}(t)f(t)dt\right)\sqrt{\int_a^b g^m(t)h^x(t)f(t)dt}\sqrt{\int_a^b g^n(t)h^y(t)f(t)dt} \end{aligned}$$

which gives the desired inequality. □

**3.9. Theorem.** *Let $f \in L_1[a,b]$. Then*

(3.11)
$$ \left({}_k^s J_a^{m(\frac{\alpha}{k}-1)+1} f^r(x)\right)\left({}_k^s J_a^{n(\frac{\alpha}{k}-1)+1} f^p(x)\right) \geq \left({}_k^s J_a^{\frac{m+n}{2}(\frac{\alpha}{k}-1)+1} f^{\frac{r+p}{2}}(x)\right)^2, $$

*for $k, m, n, r, p > 0$ and $\alpha > 1$.*

*Proof.* By taking $g(t) = \left(x^{s+1} - t^{s+1}\right)^{\frac{\alpha}{k}-1}$, $f(t) = \frac{t^s (s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)}$ and $h(t) = f(t)$ in (3.10), we obtain

$$\left(\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)} \int_a^x \left(x^{s+1} - t^{s+1}\right)^{m(\frac{\alpha}{k}-1)} t^s f^r(t)dt\right) \left(\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)} \int_a^x \left(x^{s+1} - t^{s+1}\right)^{n(\frac{\alpha}{k}-1)} t^s f^p(t)dt\right)$$

$$\geq \left(\frac{(s+1)^{1-\frac{\alpha}{k}}}{k\Gamma_k(\alpha)} \int_a^x \left(x^{s+1} - t^{s+1}\right)^{\frac{m+n}{2}(\frac{\alpha}{k}-1)} t^s f^{\frac{r+p}{2}}(t)dt\right)^2$$

which can be written as (3.11). □

**3.10. Remark.** For $k = 1$ in (3.11), we get the following inequalities:

$$\left(J_a^{m(\alpha-1)+1} f^r(x)\right) \left({}_kJ_a^{n(\alpha-1)+1} f^s(x)\right) \geq \left({}_kJ_a^{\frac{m+n}{2}(\alpha-1)+1} f^{\frac{r+s}{2}}(x)\right)^2.$$

## References

[1] Belarbi, S. and Dahmani, Z., *On some new fractional integral inequalities*, J. Ineq. Pure and Appl. Math., 10(3) (2009), Art. 86.

[2] Dahmani, Z., *New inequalities in fractional integrals*, International Journal of Nonlinear Scinece, 9(4) (2010), 493-497.

[3] Dahmani, Z., *On Minkowski and Hermite-Hadamard integral inequalities via fractional integration*, Ann. Funct. Anal. 1(1) (2010), 51-58.

[4] Dahmani, Z., Tabharit, L. and Taf, S., *Some fractional integral inequalities*, Nonl. Sci. Lett. A, 1(2) (2010), 155-160.

[5] Dahmani, Z., Tabharit, L. and Taf, S., *New generalizations of Gruss inequality usin Riemann-Liouville fractional integrals*, Bull. Math. Anal. Appl., 2(3) (2010), 93-99.

[6] Diaz, R. and Pariguan, E., *On hypergeometric functions and Pochhammer $k-$symbol*, Divulg.Math, 15.(2007),179-192.

[7] Gorenflo, R. and Mainardi, F., *Fractional calculus: integral and differential equations of fractional order*, Springer Verlag, Wien (1997), 223-276.

[8] Kilbas, A. A.,Srivastava, H. M. and Trujillo, J. J., *Theory and Applications of Fractional Differential Equations*, North-Holland Mathematics Studies, 204, Elsevier Sci. B.V., Amsterdam, 2006.

[9] Hadamard, J., Essai sur l'étude des fonctions données par leur développement de Taylor,Jour.PureandAppl.Math4(8),1892,101-186.

[10] Katugompola, U.N., *New Approach Generalized Fractional Integral*, Applied Math and Comp. 218(2011),860-865.

[11] Latif, M. A. and Hussain, S., *New inequalities of Ostrowski type for co-ordineted convex functions via fractional integrals*, Journal of Fractional Calculus and Applications,Vol. 2., 2012, No. 9, pp. 1-15.

[12] Miller, S. and Ross, B., *An introduction to the Fractional Calculus and Fractional Differential Equations*, John Wiley & Sons, USA, 1993, p.2.

[13] Mubeen, S. and Habibullah, G.M., $k-$fractional integrals and application, Int. J. Contemp. Math. Sciences, 7(2), 2012, 89-94.

[14] Romero, L. G., Luque, L. L., Dorrego, G. A. and Cerutti, R. A., *On the $k-$Riemann-Liouville Fractional Derivative*, Int. J. Contemp. Math. Sciences, 8(1), 2013, 41-51.

[15] Sarikaya, M.Z. and Ogunmez, H., *On new inequalities via Riemann-Liouville fractional integration*, Abstract and Applied Analysis, Volume 2012, Article ID 428983, 10 pages, doi:10.1155/2012/428983.

[16] Sarikaya, M. Z., Set, E., Yaldiz, H. and Basak, N., *Hermite -Hadamard's inequalities for fractional integrals and related fractional inequalities*, Mathematical and Computer Modelling, 57 (2013) 2403–2407.

[17] Sarikaya, M. Z., *On the Hermite-Hadamard-type inequalities for co-ordinated convex function via fractional integrals*, Integral Transforms and Special Functions, Volume 25, Issue 2, pp:134-147, 2014.

[18] Sarikaya, M. Z., *Ostrowski type inequalities involving the right Caputo fractional derivatives belong to $L_p$*, Facta Universitatis, Series Mathematics and Informatics , Vol. 27 No 2 (2012), 191-197.

[19] Sarikaya, M. Z. and Yaldiz, H., *On weighted Montogomery identities for Riemann-Liouville fractional integrals*, Konuralp Journal of Mathematics, Volume 1 No. 1 pp. 48-53, 2013.

[20] Sarikaya, M. Z., and Karaca, A., *On the k-Riemann-Liouville fractional integral and applications,* International Journal of Statistics and Mathematics, Vol. 1(1), pp. 022-032, August, 2014.

[21] Tunc, M., *On new inequalities for h-convex functions via Riemann-Liouville fractional integration*, Filomat 27:4 (2013), 559–565

[22] Samko, S. G., Kilbas, A. A and Marichev, O. I., *Fractional Integrals and Derivatives Theory and Application*, Gordan and Breach Science, New York, 1993.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 91 – 93

# A comment on "Generating matrix functions for Chebyshev matrix polynomials of the second type"

V. Soler*†

**Abstract**

In this comment we will demonstrate that one of the main formulas given in Ref. [1] is incorrect.

**Keywords:** Erratum, Generating matrix function, Chebyshev matrix polynomials.

*2000 AMS Classification:* 33C25, 15A60.

*Received* : 25.03.2013 *Accepted* : 14.01.2015 *Doi* : 10.15672/HJMS.20164512478

## 1. Introduction and motivation

It is well known that for a class of orthogonal polynomials $\{P_n(x)\}_{n\geq 0}$ the so-called "generating functions" of this class are an useful tool for their study. A generating function is a function $F(x,t)$, analytic for some set $D \subset \mathbb{C}^2$, so that

$$F(x,t) = \sum_{n=0}^{\infty} \alpha_n P_n(x) t^n, (x,t) \in D.$$

For example, $F(x,t) = \exp 2xt - t^2$ is the generating function for Hermite polynomials $\{H_n(x)\}_{n\geq 0}$ because we can write:

$$F(x,t) = \exp\left(2xt - t^2\right) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) t^n, \forall (x,t) \in \mathbb{C}^2.$$

The extension to the matrix framework for the classical families of Jacobi [3], Hermite [2], Gegenbauer [4], Laguerre [5] and Chebyshev [6] polynomials was made in recent years, and properties and applications of different classes for these matrix polynomials have been studied in several papers [7, 8, 10, 11, 9].

*Departamento de Matemática Aplicada. Universitat Politècnica de València. Spain Email: vsoler@dma.upv.es

†Corresponding Author.

In the matrix case, the importance of the generating function is similar to the scalar case, taking into account the possible spectral restrictions (for a matrix $A \in \mathbb{C}^{N\times N}$ we will denote by $\sigma(A)$ the spectrum of the matrix $\sigma(A) = \{z; z \text{ is a eigenvalue of } A\}$). For example:

- Laguerre matrix polynomials. If $A$ is a matrix in $\mathbb{C}^{N\times N}$ such that $-k \notin \sigma(A)$ for every integer $k > 0$, and $\lambda$ is a complex number with $\mathrm{Re}(\lambda) > 0$, the generating function [5] is given by:

$$(1-t)^{-(A+I)} \exp\left(\frac{-\lambda x t}{1-t}\right) = \sum_{n=0}^{\infty} L_n^{(A,\lambda)}(x) t^n, \forall\, x, t \in \mathbb{C}, |t| < 1.$$

- Hermite matrix polynomials. If $A$ is a matrix in $\mathbb{C}^{N\times N}$ such that $\mathrm{Re}(z) > 0, \forall z \in \sigma(A)$, the generating function [2] is given by

$$e^{xt\sqrt{A} - t^2 I} = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x, A) t^n \ , \ (x,t) \in \mathbb{R}^2$$

## 2. The detected mistake. An illustrative example

Recently, in Ref.[1], a generating matrix function for Chebyshev matrix polynomials of the second kind is presented. In theorem 2.1 of [1, p.27], the following formula (2.1) is established:

$$\sum_{n=0}^{\infty} U_n(x, A) t^n = \left(I - \sqrt{2A} x t + t^2 I\right)^{-1} \ , \ |t| < 1, |x| < 1, \quad (2.1)$$

where $I$ denotes the identity matrix of order $N$, matrix $A \in \mathbb{C}^{N\times N}$ satisfies $\mathrm{Re}(\lambda) > 0$ for all eigenvalue $\lambda \in \sigma(A)$ and $\left\|\sqrt{A}\right\| < 1/\sqrt{2}$. This formula (2.1) turns out to be the key for the development of the properties mentioned in the paper [1]. However, we will see that formula (2.1) is incorrect. For this, we only need to show that the matrix function $\left(I - \sqrt{2A} x t + t^2 I\right)$, regarded as a entire function of the complex variables $x$ and $t$, is singular for some values of $x$ and $t$ under the previous hypotheses. For example, we consider $N = 2$, and the matrix

$$A = \begin{pmatrix} \frac{3}{16} + \frac{i}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \in \mathbb{C}^{2\times 2},$$

where $i^2 = -1$. Obviously, $\sigma(A) = \left\{\frac{3}{16} + \frac{i}{4}, \frac{1}{4}\right\}$ and $A$ satisfies condition $\mathrm{Re}(\lambda) > 0$ for all eigenvalue $\lambda \in \sigma(A)$. It is easy to prove that

$$\sqrt{A} = \begin{pmatrix} \frac{1}{2} + \frac{i}{4} & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

which evidently satisfies $\sqrt{A}\sqrt{A} = A$, and condition $\left\|\sqrt{A}\right\| = \sqrt{5}/4 \approx 0.559017 < 1/\sqrt{2} \approx 0.707107$ holds.

It is easy to compute

$$\sqrt{2A} = \begin{pmatrix} \frac{1}{\sqrt{2}} + \frac{i}{2\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix},$$

which evidently satisfies $\sqrt{2A}\sqrt{2A} = 2A$, and then one gets:

$$I_{2\times 2} - \sqrt{2}Axt + t^2 I_{2\times 2} = \begin{pmatrix} 1+t^2-\frac{tx}{\sqrt{2}}-\frac{txi}{2\sqrt{2}} & 0 \\ 0 & 1+t^2-\frac{tx}{\sqrt{2}} \end{pmatrix}.$$

Taking, for example, the values

$$x = \frac{1}{2}, t = \frac{1}{16}\left(-\sqrt{-250+8i} + (2+i)\sqrt{2}\right) \approx 0.160967 - 0.89995i,$$

this choices meet the restrictions outlined ($|x| = \frac{1}{2} < 1, |t| = 0.914232 < 1$), but the term $1 + t^2 - \dfrac{tx}{\sqrt{2}} - \dfrac{txi}{2\sqrt{2}}$ is zero and matrix $I_{2\times 2} - \sqrt{2}Axt + t^2 I_{2\times 2}$ has a column of zeros, thus is singular. Thus (2.1) is meaningless. ‡

Therefore, I ask the authors of Ref. [1] to clarify the domain of choice for the variables $x, t$ in formula (2.1) in order to guarantee the validity of the remaining formulas which are derived from (2.1) and used in the remainder of the aforementioned paper.

## References


[1] Altin, A and Çekim, B. *Generating matrix functions for Chebyshev matrix polynomials of the second type*, Hacettepe Journal of Mathematics and Statistic **41 (1)**, 25-32, 2012.

[2] Jódar L. and Company R. *Hermite matrix polynomials and second order matrix differential equations*, Journal of Approximation Theory and Applications **12 (2)**, 20-30, 1996.

[3] Defez E., Jódar L. and Law A.G. *Jacobi matrix differential equation, polynomial solution and their properties*, Computers and Mathematics with Applications **48**, 789-803, 2004.

[4] Jódar L., Company R. and Navarro E. *Orthogonal matrix polynomials and second order differential equations*, Differential Equations and Dynamical Systems **3(3)**, 269-288, 1995.

[5] Jódar L., Company R. and Navarro E. *Laguerre matrix polynomials and systems of second order differential equations*, Applied Numeric Mathematics **15(1)**, 53-64, 1994.

[6] Defez E. and Jódar L., *Chebyshev matrix polynomials and second order matrix differential equations*, Utilitas Mathematica **61**, 107-123, 2002.

[7] Defez E. and Jódar L., *Some applications of the Hermite matrix polynomials series expansions*, Journal of Computational and Applied Mathematics **99**, 105-117, 1998.

[8] Defez E., Tung M. and Sastre J., *Improvement on the Bound of Hermite Matrix Polynomials*, Linear Algebra And its Applications **434**, 1910-1919, 2011.

[9] Durán A. J. and Grünbaum F. A. *A survey on orthogonal matrix polynomials satisfying second order differential equations*, Journal of Computational and Applied Mathematics **178**, 169-190, 2005.

[10] Sastre J., Ibáñez J., Defez E. and Ruiz P. *Efficient orthogonal matrix polynomial based method for computing matrix exponential*, Applied Mathematics and Computation **217**, 6451-6463, 2011.

[11] Sastre J., Ibáñez J., Defez E. and Ruiz P. *Computing matrix functions solving coupled differential models*, Mathematical and Computer Modelling **50**, 831-839, 2009.


‡

Of course, the choice of these values are not unique. For example, taking the values

$$x = 2i\sqrt{\frac{2}{8+\sqrt{2}}}, t = -i\left(\frac{\sqrt{9+\sqrt{2}}-1}{\sqrt{8+\sqrt{2}}}\right),$$

this choices satisfies the restrictions outlined ($|x| = 0.921835 < 1, |t| = 0.725853 < 1$), but $1 + t^2 - \frac{tx}{\sqrt{2}} = 0$ and then matrix function $I_{2\times 2} - \sqrt{2}Axt + t^2 I_{2\times 2}$ is singular.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 95 – 105

# Annihilator conditions related to the quasi-Baer condition

A. Taherifar*

**Abstract**

We call a ring $R$ an *EGE-ring* if for each $I \trianglelefteq R$, which is generated by a subset of right semicentral idempotents there exists an idempotent $e$ such that $r(I) = eR$. The class $EGE$ includes quasi-Baer, semiperfect rings (hence all local rings) and rings with a complete set of orthogonal primitive idempotents (hence all Noetherian rings) and is closed under direct product, full and upper triangular matrix rings, polynomial extensions (including formal power series, Laurent polynomials, and Laurent series) and is Morita invariant. Also we call $R$ an $AE$-ring if for each $I \trianglelefteq R$, there exists a subset $S \subseteq S_r(R)$ such that $r(I) = r(RSR)$. The class $AE$ includes the principally quasi-Baer ring and is closed under direct products, full and upper triangular matrix rings and is Morita invariant. For a semiprime ring R, it is shown that $R$ is an $EGE$ (resp., $AE$)-ring if and only if the closure of any union of clopen subsets of $Spec(R)$ is open (resp., $Spec(R)$ is an $EZ$-space).

**Keywords:** Quasi-Baer ring, $AE$-ring, $EGE$-ring, $Spec(R)$, Semicentral idempotent, $EZ$-space.

*2000 AMS Classification:* Primary 16D25, 16N60; Secondary 54G05.

*Received* : 11.02.2014 *Accepted* : 20.11.2014 *Doi* : 10.15672/HJMS.20164512485

## 1. Introduction

Throughout this paper, $R$ denotes an associative ring with identity. In this paper, we introduce and investigate the concept of $EGE$ (resp., $AE$)-ring. We call $R$ an $EGE$ (resp., $AE$)-ring, if for any ideal $I$ of $R$ which $I = RSR$, $S \subseteq S_r(R)$ (resp., any ideal $I$ of $R$) there exists an idempotent $e \in R$ (resp., a subset $S \subseteq S_r(R)$) such that $r(I) = eR$ (resp., $r(I) = r(RSR)$), where $r(I)$ (resp., $l(J)$) denotes the right annihilator (resp., left annihilator) of $I$.

In Section 2, we show that any quasi-Baer ring and any ring with a complete set of right (left) triangulating idempotents are $EGE$-ring. Hence semiperfect rings (hence all

*Department of Mathematics, Yasouj University, Yasouj, Iran.
Email : `ataherifar@mail.yu.ac.ir`

local rings) and rings with a complete set of orthogonal primitive idempotents (hence all Noetherian rings) are EGE-ring. We also show that any principally quasi-Baer-ring (hence, biregular rings) is an $AE$-ring. We provide examples of $EGE$ (resp., $AE$)-rings which are not quasi-Baer (resp., principally quasi-Baer )-ring.

In Section 3, we consider the closure of the class of $EGE$ (resp., $AE$)-ring with respect to various ring extensions including matrix, and polynomial extension (including formal power series, Laurent polynomials, and Laurent series). In Theorem 3.3, we obtain a characterization of semicentral idempotents in $\mathbf{M_n(R)}$ (resp., $\mathbf{T_n(R)}$). The $EGE$ (resp., $AE$) property is shown to be Morita invariant in Theorem 3.6.

Topological equivalency of semiprime $EGE$ (resp., $AE$)-ring is the focus of Section 4. In Theorem 4.2, we show that a semiprime ring $R$ is an $EGE$ (resp., $AE$)-ring if and only if the closure of any union of clopen subsets of $Spec(R)$ (i.e., the space of prime ideals of $R$), is open (resp., $Spec(R)$ is an $EZ$-space).

Let $\emptyset \neq X \subseteq R$. Then $X \leq R$ and $X \trianglelefteq R$ denote that $X$ is a right ideal and $X$ is an ideal respectively. For any subset $S$ of $R$, $l(S)$ and $r(S)$ denote the left annihilator and the right annihilator of $S$ in $R$. The ring of $n$-by-$n$ (upper triangular) matrices over $R$ is denoted by $\mathbf{M_n(R)}$ ($\mathbf{T_n(R)}$). We use $R[x]$, $R[[x]]$, $R[x, x^{-1}]$ and $R[[x, x^{-1}]]$ for the ring of polynomials over R, the ring of formal power series over R, the skew Laurent polynomial ring over $R$, and the skew Laurent series ring over R, respectively. A ring $R$ is called (*quasi-*)*Baer* if the left annihilator of every (ideal) nonempty subset of $R$ is generated , as a left ideal, by an idempotent. The (quasi-)Baer conditions are left -right symmetric. It is well known that $R$ is a quasi-Baer if and only if $\mathbf{M_n(R)}$ is quasi-Baer if and only if $\mathbf{T_n(R)}$ is a quasi-Baer ring (see [2], [7], [8], [13] and [18]). An idempotent $e$ of a ring $R$ is called left (resp., right) semicentral if $ae = eae$ (resp., $ea = eae$) for all $a \in R$. It can be easily checked that an idempotent $e$ of $R$ is left (resp., right) semicentral if and only if $eR$ (resp., $Re$) is an ideal. Also note that an idempotent $e$ is left semicentral if and only if $1 - e$ is right semicentral. See [3] and [5], for more detailed account of semicentral idempotents. Thus for a left (resp., right) ideal $I$ of a ring $R$, if $l(I) = Re$ (resp., $r(I) = eR$) with an idempotent $e$, then $e$ is right (resp., left) semicentral, since $Re$ (resp., $eR$) is an ideal. Thus for a left (resp., right) ideal $I$ of a quasi-Baer ring $R$ with $l(I) = Re$ (resp. $r(I) = eR$) for some idempotent $e \in R$, it follows that $e$ is a right (resp., left) semicentral idempotent. We use $S_l(R)$ ($S_r(R)$) to denote the set of left (right) semicentrel idempotents of $R$. For an idempotent $e$ of $R$ if $S_r(R) = \{0, e\}$, then $e$ is called *semicentral reduced*. If 1 is semicentral reduced, then we say $R$ is *semicentral reduced*.

## 2. Preliminary results and examples

**2.1. Definition.** We call $R$ an *EGE-ring*, if for each ideal $I = RSR$, $S \subseteq S_r(R)$, there exists an idempotent $e$ such that $r(I) = eR$. Since for each $S \subseteq S_r(R)$, $r(RSR) = r(RS) = r(SR) = r(S)$, $R$ is an $EGE$-ring if and only if for each $S \subseteq S_r(R)$, there exists an idempotent $e$ such that $r(S) = eR$.

**2.2. Definition.** We call $R$ an *AE-ring*, if for any ideal $I$ of $R$ there exists a subset $S \subseteq S_r(R)$ such that $r(I) = r(RSR) = r(S)$. We know that $I$ is equivalent to $J$ if and only if r(I) = r(J). Then $R$ is an $AE$-ring if an only if every ideal of $R$ is equivalent to one which is generated by a subset of right semicentral idempotents.

**2.3. Lemma.** Let $e_1$ and $e_2$ be two right semicentral idempotents.

(i) $e_1e_2$ is a right semicentral idempotent.
(ii) $(e_1 + e_2 - e_1e_2)$ is a right semicentral idempotent.

(iii) If $S \subseteq S_r(R)$ is finite, then there is a right semicentral idempotent $e$ such that $RSR = ReR =< e >$.

*Proof.* (i) By hypothesis, for any $r \in R$ we have, $e_1e_2r = e_1e_2re_2 = e_1e_2re_1e_2$. On the other hand, $(e_1e_2)^2 = e_1e_2e_1e_2 = e_1e_2^2 = e_1e_2$. Hence $e_1e_2 \in S_r(R)$.

(ii) The routine calculation shows that $(e_1 + e_2 - e_1e_2)^2 = (e_1 + e_2 - e_1e_2)$, and by hypothesis, for any $r \in R$ we have, $(e_1+e_2-e_1e_2)r = e_1r+e_2r-e_1e_2r = e_1re_1+e_2re_2-e_1e_2re_2 = (e_1 + e_2 - e_1e_2)r(e_1 + e_2 - e_1e_2)$. Hence $(e_1 + e_2 - e_1e_2) \in S_r(R)$.

(iii) We use induction. If $S = \{e_1, e_2\}$, then we have $< e_1, e_2 >=< e_1 + e_2 - e_1e_2 >$. By (ii), $e_1 + e_2 - e_1e_2 \in S_r(R)$. Now let the statement is true for $|S| = n$ and let $S = \{e_1, ..., e_n, e_{n+1}\}$. Then we have $< S >=< \{e_1, ..., e_n\} > + < e_{n+1} >$. By hypothesis, there is a right semicentral idempotent $f$ such that $< \{e_1, ..., e_n\} >=< f >$. Hence $< S >=< f + e_{n+1} - fe_{n+1} >$, where by $(ii)$, we have $e = f + e_{n+1} - fe_{n+1} \in S_r(R)$.□

Recall that an ordered set $\{b_1, ..., b_n\}$ of nonzero distinct idempotents in $R$ is called a set of *right triangulating idempotents* of $R$ if all the following hold:

(i) $1 = b_1 + ... + b_n$;
(ii) $b_1 \in S_r(R)$; and
(iii) $b_{k+1} \in S_r(c_kRc_{k+1})$, where $1 = 1 - (b_1 + ... + b_k)$, for $1 \le k \le n$.

Similarly is defined a set of *left triangulating idempotents* of R using (i), $b_1 \in S_l(R)$ and $b_{k+1} \in S_l(c_kRc_k)$. From part (iii) of the above definition, a set of right (left) triangulating idempotents is a set of pairwise orthogonal idempotents.

A set $\{b_1, ..., b_n\}$ of right (left) triangulating idempotents of $R$ is said to be *complete* if each $b_i$ is also semicentral reduced (see [11]).

**2.4. Proposition.** The following statements hold.

(i) Any ring $R$ with finite triangulating dimension (equivalently, $R$ has a complete set of right (left)triangulating idempotents) is an $EGE$-ring.
(ii) A ring $R$ is quasi-Baer if and only if $R$ is $EGE$ and $AE$.

*Proof.* (i) By [5, Theorem 2.9], $R$ has a complete set of right traingulating idempotents if and only if $\{Rb : b \in S_r(R)\}$ is finite. Now let $I = RSR$ be an ideal of $R$ and $S \subseteq S_r(R)$. Then we have $r(I) = r(RS) = r(\{Rb : b \in S\})$. But $\{Rb : b \in S\}$ is finite say $\{Rb_1, ..., Rb_n\}$. Hence $r(I) = r(\{Rb_1, ..., Rb_n\}) = r(\{b_1, ..., b_n\})$. By Lemma 2.3, there exists a right semicentral idempotent $e$ such that $r(I) = r(\{b_1, ..., b_n\}) = r(eR) = r(Re) = (1 - e)R$. Thus $R$ is an $EGE$-ring.

(ii) By definition, any quasi-Baer ring is an $EGE$-ring. If $I$ is an ideal of a quasi-Baer ring $R$, then there is $e \in S_l(R)$ such that $r(I) = eR = r(R(1 - e))$. On the other hand for each $S \subseteq S_r(R)$ we have $r(RS) = r(SR) = r(RSR)$, hence $r(I) = r(RSR)$, where $S = \{1 - e\}$, and $S \subseteq S_r(R)$. Hence $R$ is an $AE$-ring. Conversely, let $I \trianglelefteq R$. Then by hypothesis, there exists a subset $S \subseteq S_r(R)$ such that $r(I) = r(RSR)$. Again by hypothesis, there is an idempotent $e$ such that $r(RSR) = eR$. Thus $r(I) = eR$. □

**2.5. Example.** By Proposition 2.4, all of the rings mentioned in Proposition 2.14 of [5], are $EGE$-rings. Note that this list includes semiperfect rings (hence all local rings, left or right artinian rings) and rings with a complete set of orthogonal primitive idempotents (hence all Noetherian rings) and many more rings.

Recall that, a ring R is *right* (resp., *left*) *principally quasi-Baer* (or simply right (resp., left) *pq*-Baer) if the right (resp., left) annihilator of a principally right (resp/ left) ideal is generated (as a right (resp., left) ideal) by an idempotent (see [9]).

**2.6. Proposition.** The following statements hold.

(i) $R$ is an $EGE$ ring if and only if for each $I \trianglelefteq R$, which is generated by a subset $S \subseteq S_l(R)$, we have $l(I) = Re$, for some idempotent $e \in R$.

(ii) $R$ is an $AE$-ring if and only if for each $a \in R$ there exists a subset $S_a \subseteq S_r(R)$ such that $r(RaR) = r(aR) = r(S_a)$.

(iii) Every right principally quasi-Baer ring is an $AE$-ring.

*Proof.* (i) Let $I = RSR$, where $S \subseteq S_l(R)$. Take $J = RKR$, $K = \{1 - s : s \in S\}$. Then $K \subseteq S_r(R)$. By hypothesis and Lemma 2.3, there is $e \in S_l(R)$ such that $r(J) = r(KR) = r(RK) = eR$. Hence for each $s \in S$, $(1 - s)e = 0$, so $e = se$. Therefore $Re = SRe$. This implies that $l(RSR) = l(RS) = l(SR) = l(SRe) = l(Re) = l(eR) = R(1 - e)$. Similarly we can get the converse.

(ii) By definition, $\Rightarrow$ is evident.

$\Leftarrow$ Now let $I \trianglelefteq R$. We have $r(I) = \bigcap_{a\in I} r(RaR)$. By hypothesis, for each $a \in R$ there exists $S_a \subseteq S_r(R)$ such that $r(RaR) = r(RS_aR)$. Hence $r(I) = \bigcap_{a\in I} r(RS_aR) = r(R(\bigcup_{a\in I} S_a)R)$.

(iii) Let $a \in R$. Then there is an idempotent $e \in R$ such that $r(RaR) = r(aR) = eR = r(R(1-e)) = r((1-e)R) = r(R(1-e)R)$. We know that $1-e$ is a right semicentral idempotent. By $(ii)$, $R$ is an $AE$-ring.

A ring $R$ is called *biregular* if every principal ideal of R is generated by a central idempotent of R (see [8]). Note that a biregular ring is *pq*-Baer. Hence any biregular ring is an $AE$-ring.

Recall from [20] that a topological space $X$ is an *EZ-space* if for every open subset $A$ of $X$ there exists a collection $\{A_\alpha : \alpha \in S\}$ of clopen (i.e., sets that are simultaneously closed and open) subsets of $X$ such that $cl_X A = cl_X(\bigcup_{\alpha\in S} A_\alpha)$. We denote by $C(X)$, the ring of all real-valued continuous functions on a completely regular Hausdorff space $X$. For any $f \in C(X)$, $Z(f) = \{x \in X : f(x) = 0\}$ is called a zero-set. A topological space $X$ is called *extremally disconnected* (resp., *basically disconnected*), if the interior of any open set (resp., the interior of any zero-set) is closed. Clearly any extremally disconnected space is an $EZ$-space, but there exist $EZ$-spaces which are not extremally disconnected (resp., basically disconnected) (see [20]). It is clear that a subset $A$ of $X$ is clopen if and only if $A = Z(f)$ for some idempotent $f \in C(X)$. For terminology and notations, the reader is referred to [15] and [14]. For any subset $A$ of $X$ we denote by $intA$ the interior of $A$ (i.e., the largest open subset of $X$ contained in $A$).

In the following, we provide examples of commutative $AE$ and non-commutative $EGE$ rings which are not quasi-Baer. We need the following lemma which is Corollary 2.2 in [1].

**2.7. Lemma.** For $f, g \in C(X)$, $r(f) = r(g)$ if and only if $intZ(f) = intZ(g)$.

**2.8. Example.** By [20, Theorem 3.7], $C(X)$ is an $AE$-ring if and only if $X$ is an $EZ$-space. On the other hand by [1], we have $C(X)$ is a *pq*-Baer ring if and only if $X$ is a basically disconnected space. So, if $X$ is an $EZ$-space which is not basically disconnected space (e.g., [20, Example 3.4]), then $C(X)$ is an $AE$-ring but is not a *pq*-Baer ring. By Proposition 2.4 (ii), $C(X)$ is not an $EGE$-ring.

**2.9. Example.** The ring $R = \begin{pmatrix} \mathbb{Z} & \mathbb{Z}_2 \\ 0 & \mathbb{Z}_2 \end{pmatrix} = \{\begin{pmatrix} n & a \\ 0 & b \end{pmatrix} : n \in \mathbb{Z}, a, b \in \mathbb{Z}_2\}$ has a finite number of right semicentral idempotents. By Proposition 2.4, $R$ is an $EGE$-ring. But $R$ is not a quasi-Baer ring. If $I = \begin{pmatrix} 0 & \mathbb{Z}_2 \\ 0 & \mathbb{Z}_2 \end{pmatrix}$, then we have $l(I) = \begin{pmatrix} 2\mathbb{Z} & 0 \\ 0 & 0 \end{pmatrix}$, which does not contain any idempotent. By Proposition 2.4 (ii), $R$ is not an $AE$-ring.

**2.10. Theorem.** Let $R = \prod_{x\in X} R_x$ be a direct product of rings.

(i) $R$ is an $EGE$-ring if and only if each $R_x$ is an $EGE$ ring.

(ii) $R$ is an $AE$-ring if and only if each $R_x$ is an $AE$ ring.

*Proof.* (i) Assume that $R$ is an $EGE$-ring. Choose $x \in X$. Let $I_x \trianglelefteq R_x$ and $I_x = < K_x >$, where $K_x \subseteq S_r(R_x)$ and $h_x : R_x \to R$ be the canonical homomorphism. Then $h_x(I_x) \trianglelefteq R$, $h_x(I_x) = < h_x(K_x) >$ and $h_x(K_x) \subseteq S_r(R)$. So there exists an idempotent $e \in R$ such that $r(h_x(I_x)) = eR$. Let $\pi_x : R \to R_x$ denote the canonical projection homomorphism. Then $\pi_x(e)$ is an idempotent in $R_x$ and $r(I_x) = \pi_x(e)R_x$.

Conversely, assume that $R_x$ is an $EGE$-ring for all $x \in X$. Let $I \trianglelefteq R$ and $I = < K >$, $K \subseteq S_r(R)$. Then $I_x = \pi_x(I) = < \pi_x(K) > = < K_x >$. It is easy to see that $K_x \subseteq S_r(R)$ for each $x \in X$. Hence there exists an idempotent $e_x \in R_x$ such that $r(I_x) = e_x R_x$ for each $x \in X$. Let $e = (e_x)_{x\in X}$. Then $e$ is an idempotent in $R$ and $r(I) = eR$.

(ii) Let $R$ be an $AE$-ring. For $x \in X$, suppose that $a_x \in R_x$. Then there is $a \in R$ such that $\pi_x(a) = a_x$. By hypothesis, there exists $S \subseteq S_r(R)$ such that $r(RaR) = r(RSR)$. Now we can see that $r(R_x a_x R_x) = r(R_x S_x R_x)$, where $S_x = \pi_x(S) \subseteq S_r(R_x)$. By Proposition 2.6, $R_x$ is an $AE$-ring. Conversely, suppose that $a \in R$. Then $\pi_x(a) = a_x \in R_x$ for each $x \in X$. By hypothesis, for each $x \in X$ there exists $S_x \subseteq S_r(R_x)$ such that $r(R_x a_x R_x) = r(R_x S_x R_x)$. Now let $S = \prod_{x\in X} S_x$. Then $S \subseteq S_r(R)$ and $r(RaR) = r(RSR)$. By Proposition 2.6, $R$ is an $AE$-ring. □

## 3. Extensions of $EGE$ and $AE$-rings

In this section, we investigate the behavior of the $EGE$ (rep., $AE$)-ring property with respect to various ring extensions including matrix, polynomial, and formal power series. Also semicentral idempotents in $\mathbf{M_n(R)}$ (resp., $\mathbf{T_n(R)}$) are investigated.

The following Lemma is Lemma 3.1 in [4].

**3.1. Lemma.** Let R be a ring and $S = \mathbf{M_n(R)}$.

(i) Then $J \trianglelefteq S$ if and only if $J = \mathbf{M_n(I)}$, for some $I \trianglelefteq R$.
(ii) If $I \trianglelefteq R$, then $r_S(\mathbf{M_n(I)}) = \mathbf{M_n(r_R(I))}$.

**3.2. Lemma.** The following statements hold.

(i) If $R$ is an $EGE$-ring and $e$ is an idempotent, then $eRe$ is an $EGE$-ring.
(ii) If $R$ is an $AE$-ring and $e$ is an idempotent, then $eRe$ is an $AE$-ring.

*Proof.* (i) Let $I \trianglelefteq eRe$ and $I = eReKeRe$, where $K \subseteq S_r(eRe)$. For each $exe \in K$ and $r \in R$, we have $(exe)(re) = (exe)(ere) = (exe)(ere)(exe) = (exe)(re)(exe)$. So $K \subseteq S_r(Re)$. Now let $J = ReKRe$. Then $J \trianglelefteq Re$. By hypothesis and Theorem 2.10, $Re$ is an $EGE$-ring, hence there is an idempotent $f \in Re$ such that $r_{Re}(J) = fRe$. Now we claim that $r_{eRe}(I) = (ef)(eRe)$. For see this, let $exe \in r_{eRe}(eReKeRe)$. Then we have $exe \in r_{eRe}(eKRe) = r_{eRe}(ReKRe)$, so $xe \in r_{Re}(ReKRe)$. This says that $r_{eRe}(I) \subseteq (ef)(eRe)$. Therefore $xe = fse$ for some $s \in R$. But $f = fe$, so $exe = (ef)(ere)$. On the other hand we have $f \in r_{Re}(ReKRe)$. This implies that $Ief = 0$, thus $(ef)(eRe) \subseteq r_{eRe}(I)$.

(ii) Assume that $I \trianglelefteq eRe$. Then $I \leq Re$. By hypothesis and Theorem 2.10, $Re$ is an $AE$-ring. Hence there exists $S \subseteq S_r(Re)$ such that $r_{Re}(I) = r_{Re}(ReSRe)$. We have $eSe(eRe)eSe = eS(Re) = eSReS = eS(eRe)eS$ and for each $s \in S$, $(es)^2 = eses = es^2 = es$. This shows that $eS = eSe \subseteq S_r(eRe)$. Now we claim that $r_{eRe}(I) = r_{eRe}(eRe(eSe)eRe) = r_{eRe}(eReSRe)$. Let $exe \in r_{eRe}(I)$. Then $Iexe = Ixe = 0$. So $xe \in r_{Re}(I) = r_{Re}(ReSRe)$. Therefore $ReSRexe = 0$. This implies that $exe \in r_{eRe}(ReSRe) \subseteq r_{eRe}(eReSRe)$. Now suppose that $exe \in r_{eRe}(eReSRe)$. Then $exe \in r_{eRe}(eSRe) = r_{eRe}(ReSRe)$. Hence $xe \in r_{Re}(ReSRe) = r_{Re}(I)$. Thus $Iexe = Ixe = 0$. This shows that $exe \in r_{eRe}(I)$. □

In the following Theorem, we characterize semicentral idempotents in $\mathbf{M_n(R)}$ and $\mathbf{T_n(R)}$.

**3.3. Theorem.** The following statements hold.

(i) $A = [a_{ij}] \in S_r(\mathbf{M_n(R)})$ if and only if we have;

(a) $a_{11} \in S_r(R)$.

(b) $a_{ij} = a_{ij}a_{11}$ for all $1 \leq i, j \leq n$.

(c) For each $1 \leq i \leq n$, $a_{11}a_{ii} = a_{11}$ and $a_{11}a_{ij} = 0$ for all $1 \leq j \neq i \leq n$.

(ii) $A = [a_{ij}] \in S_r(\mathbf{T_n(R)})$ if and only if we have;

(d) For each $1 \leq i \leq n$, $a_{ii} \in S_r(R)$.

(e) For each $1 \leq i \leq n$, $a_{ki} = a_{ki}a_{ii}$ for all $1 \leq k \leq i$ and $a_{ii}a_{ij} = 0$ for all $i < j \leq n$.

*Proof.* (i)$\Rightarrow$ First we show that (a) holds. Suppose that $r \in R$. Consider $B = [b_{ij}]$, where $b_{11} = r$, and $b_{ij} = 0$ for all $i \neq 1$, $j \neq 1$. Then by hypothesis, $ABA = AB$. This implies that $a_{11}ra_{11} = a_{11}r$, so $a_{11} \in S_r(R)$.

(b) Let $B = [b_{ij}]$, where $b_{j1} = 1$ and $b_{ik} = 0$ for each $i \neq j$ and $k \neq 1$. By hypothesis, $ABA = AB$, so we have $a_{ij}a_{11} = a_{ij}$ for all $1 \leq i, j \leq n$.

(c) For fixed $i$, consider $B = [b_{ij}]$, where $b_{1i} = 1$ and other entries are zero. Then $ABA = AB$ implies that $a_{11}a_{ii} = a_{11}$ and $a_{11}a_{ij} = 0$ for all $1 \leq j \neq i \leq n$.

(i)$\Leftarrow$ $a_{11} \in S_r(R)$ implies that $D = [d_{ij}] \in S_r(\mathbf{M_n(R)})$, where $d_{ii} = a_{11}$ and other entries are zero. On the other hand, by (b) and (c), we can see that $A = AD$ and $DA = D$. Hence, for $B \in \mathbf{M_n(R)}$ we have $ABA = ADBA = ADBDA = ADBD = ADB = AB$. Therefore $A \in S_r(\mathbf{M_n(R)}$.

(ii)$\Rightarrow$ (d) The proof of this part is analogous to that of part (a).

(e) For $B = [b_{ij}]$, where $b_{ii} = 1$ and other entries are zero. We have $ABA = AB$. Therefore $a_{ki} = a_{ki}a_{ii}$ for all $1 \leq k \leq i$ and $a_{ii}a_{ij} = 0$ for all $i < j \leq n$.

(ii)$\Leftarrow$ If $a_{ii} \in S_r(R)$, then $D = [d_{ij}] \in S_r(\mathbf{T_n(R)})$, where $d_{ii} = a_{ii}$ and other entries are zero. On the other hand, by (e), we can see that $A = AD$ and $DA = D$. Hence for $B \in \mathbf{T_n(R)}$, we have $ABA = ADBA = ADBDA = ADBD = ADB = AB$. Therefore $A \in S_r(\mathbf{T_n(R)}$. □

**3.4. Lemma.** If $J \trianglelefteq M_n(R)$ and $J =< S >$, where $S \subseteq S_r(\mathbf{M_n(R)})$, then there is $I \trianglelefteq R$ generated by a subset of right semicentral idempotents of $R$ such that $J = M_n(I)$.

*Proof.* By argument of [16, Theorem 3.1], $J = M_n(I)$, where $I$ is the set of all $(1,1)$-entries of matrices in $J$. Now let $S_{11}$ be the set of all $(1,1)$-entries of matrices in $S$. By Theorem 3.3, $S_{11} \subseteq S_r(R)$, and we can see that $I = RS_{11}R$. □

**3.5. Proposition.** The following statements hold.

(i) $R$ is an $EGE$-ring if and only if $\mathbf{M_n(R)}$ is an $EGE$-ring.

(ii) $R$ is an $AE$-ring if and only if $\mathbf{M_n(R)}$ is an $AE$-ring.

*Proof.* (i) Let $J$ be an ideal of $\mathbf{M_n(R)}$ and $J =< S >$, where $S \subseteq S_r(\mathbf{M_n(R)})$. By Lemma 3.4, there exists $I \trianglelefteq R$, where $I =< S_1 >$ for some $S_1 \subseteq S_r(R)$ and $J = \mathbf{M_n(I)}$. By Lemma 3.1 and hypothesis, we have $r(J) = \mathbf{M_n(r(I))} = \mathbf{M_n(eR)}$ for some idempotent $e$ in $R$. Hence $r(J) = E\mathbf{M_n(R)}$, where in matrix $E$ for each $1 \leq i \leq n$, $E_{ii} = e$ and $E_{ij} = 0$ for all $i \neq j$. Conversely, we have $E\mathbf{M_n(R)E} \simeq \mathbf{R}$, where in matrix $E$, $E_{11} = 1$ and for each $i \neq 1$ and $j \neq 1$, $E_{ij} = 0$. Now by Lemma 3.2, $R$ is an $EGE$-ring.

(ii) Let $J$ be an ideal of $\mathbf{M_n(R)}$. By Lemma 3.1, there is an ideal $I$ of $R$ such that $J = M_n(I)$, and $r(J) = r(\mathbf{M_n(I)}) = \mathbf{M_n(r(I))}$. By hypothesis, there exists $S \subseteq S_r(R)$ such that $r(I) = r(RSR)$. Hence $r(J) = \mathbf{M_n(r(RSR))} = \mathbf{r(M_n(RSR))}$. On the other hand, we can see that $\mathbf{M_n(RSR)} = \mathbf{M_n(R)D_n(S)M_n(R)}$, where $\mathbf{D_n(S)}$ is the set of diagonal matrices over $S$, and $\mathbf{D_n(S)} \subseteq \mathbf{S_r(M_n(R))}$. Thus $r(J) = r(\mathbf{M_n(R)D_n(S)M_n(R)})$. Conversely, by Lemma 3.2, it is obvious. □

**3.6. Theorem.** The following statements hold.

(i) The $EGE$ property is a Morita invariant.
(ii) The $AE$ property is a Morita invariant.

*Proof.* These results are consequences of Lemma 3.2, Proposition 3.5 and [17, Corollary 18.35]. $\square$

**3.7. Theorem.** The following statements hold.

(i) $R$ is an $EGE$-ring if and only if $\mathbf{T_n(R)}$ is an $EGE$-ring.
(ii) $R$ is an $AE$-ring if and only if $\mathbf{T_n(R)}$ is an $AE$-ring.

*Proof.* (i)$\Leftarrow$ Assume that $\mathbf{T_n(R)}$ is an $EGE$-ring. Then we have $E\mathbf{T_n(R)E} \simeq \mathbf{R}$, where in matrix $E$, $E_{11} = 1$ and other entries are zero. Now by Lemma 3.2, $R$ is an $EGE$-ring.

(i)$\Rightarrow$ Let $I$ be an ideal of $T_n(R)$ which is generated by $S = \{A_\alpha : \alpha \in K\} \subseteq S_r(\mathbf{T_n(R)})$. By Theorem 3.3, for each $\alpha \in K$ and $1 \leq i \leq n$, we have $(a_{ii})_\alpha \in S_r(R)$, where $(a_{ii})_\alpha$ is the $(i,i)$-th, entries in $A_\alpha$. Now for each $1 \leq i \leq n$, let $J_i$ be the ideal generated by $\{(a_{ii})_\alpha : \alpha \in K\}$ in $R$. By hypothesis, for each $1 \leq i \leq n$ there is an idempotent $e_i \in R$ such that $r(J_i) = e_iR$. We claim that $r(I) = ET_n(R)$ where for each $1 \leq i \leq n$, $E_{ii} = e_i$ and $E_{ij} = 0$, for all $i \neq j$. By Theorem 3.3, we can see that; for each $\alpha \in K$ there exists a diagonal matrix $D_\alpha$ such that $A_\alpha = A_\alpha D_\alpha$, where $(D_\alpha)_{ii} = (A_\alpha)_{ii}$. So, for each $\alpha \in K$ we have $A_\alpha E = A_\alpha D_\alpha E = 0$. Now let $A \in I$. Then we have $A = \sum_{i=1}^n B_iA_iC_i$, where $A_i \in S$ and $B_i, C_i \in \mathbf{T_n(R)}$. Therefore $AE = (\sum_{i=1}^n B_iA_iC_i)E = \sum_{i=1}^n B_iA_iC_iA_iE = 0$. Hence $E \in r(I)$.

Now suppose that $B \in r(I)$ and $x \in J_i =< (a_{ii})_\alpha : \alpha \in S >$. Then $A \in I$ where $a_{ii} = x$ and other entries are zero. So we have

$$AB = \begin{pmatrix} 0 & 0 & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ xb_{i1} & xb_{i2} & . & . & . & xb_{in} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & 0 \end{pmatrix}_{n\times n} = 0.$$

This equality implies that $b_{ij} \in r_R < (a_{ii})_\alpha : \alpha \in S >$ for each $1 \leq j \leq n$. Hence for fixed i and each $1 \leq j \leq n$ there is $r_{ij} \in R$ such that $b_{ij} = e_ir_{ij}$. Therefore we have

$$B = \begin{pmatrix} e_1r_{11} & e_1r_{12} & . & . & . & e_1r_{1n} \\ 0 & e_2r_{22} & . & . & . & e_2r_{2n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & e_nr_{nn} \end{pmatrix}_{n\times n} = E \times \begin{pmatrix} r_{11} & r_{12} & . & . & . & r_{1n} \\ 0 & r_{22} & . & . & . & r_{2n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & r_{nn} \end{pmatrix}_{n\times n}.$$

Thus $B \in ET_n(R)$.

(ii) Let $I \trianglelefteq T_n(R)$. Then

$$I = \begin{pmatrix} I_{11} & I_{12} & . & . & . & I_{1n} \\ 0 & I_{22} & . & . & . & I_{2n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & I_{nn} \end{pmatrix},$$

where each $I_{ij} \trianglelefteq R$, $I_{ij} = \{0\}$ for all $i > j$, $I_{ij} \subseteq I_{ik}$ for all $k \geq j$, and $I_{hj} \subseteq I_{ij}$ for all $h \geq i$. Therefore

$$r_{T_n(R)}(I) = \begin{pmatrix} r_R(I_{11}) & r_R(I_{11}) & . & . & . & r_R(I_{11}) \\ 0 & r_R(I_{12}) & . & . & . & r_R(I_{12}) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & r_R(I_{1n}) \end{pmatrix}.$$

By hypothesis, for each $1 \leq i, j \leq n$, there exists $S_{ij} \subseteq S_r(R)$ such that $r_R(I_{ij}) = r_R(S_{ij})$. This implies that

$$r_{T_n(R)}(I) = r_{T_n(R)}\left(\begin{pmatrix} S_{11} & 0 & 0 & . & . & 0 \\ 0 & S_{12} & 0 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & S_{1n} \end{pmatrix}\right).$$

On the other hand, it is easy to see that $\begin{pmatrix} S_{11} & 0 & 0 & . & . & 0 \\ 0 & S_{12} & 0 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & S_{1n} \end{pmatrix} \subseteq S_r(T_n(R))$. So we are done. □

We need the following lemma which is Lemma 1.7 in [10].

**3.8. Lemma.** For a ring $R$, let $T$ be $R[x, x^{-1}]$ or $R[[x, x^{-1}]]$. If $e(x) \in S_r(T)$ then $e_0 \in S_r(R)$ where $e_0$ is the constant term of $e(x)$. Moreover, $Te(x) = Te_0$.

Also we need the following lemma which is Proposition 3 in [12].

**3.9. Lemma.** Let $e(x) = \sum_{i=0}^{\infty} e_i x^i$. Then $e(x) \in S_l(R[[x]])$ if and only if

(i) $e_0 \in S_l(R)$.;

(ii) $e_0 r e_i = r e_i$ and $e_i r e_0 = 0$, for all $r \in R$, $i = 1, 2, ...$;

(iii) $\sum_{\substack{i+j=k \\ i,j \geq 1}} e_i r e_j = 0$, for all $r \in R$ and $k \geq 2$.

**3.10. Theorem.** Let R be a ring and $X$ an arbitrary nonempty set of not necessarily commuting indeterminates. Then the following conditions are equivalent:

(i) $R$ is $EGE$;

(ii) $R[X]$ is $EGE$;

(iii) $R[[X]]$ is $EGE$;

(iv) $R[x, x^{-1}]$ is $EGE$;

(v) $R[[x, x^{-1}]]$ is $EGE$.

*Proof.* We will prove the equivalency of (i) and (iv). The equivalency of other cases can be shown similarly, by Lemmas 3.8, 3.9 and [6, Proposition 2.4(iv)]. (i)⇒(iv), let $T = R[x, x^{-1}]$ and $I = TST$, where $S \subseteq S_r(T)$. Let $S_0$ be the set of all constant elements of $S$. Then by Lemma 3.8, $S_0 \subseteq S_r(R)$ and $RS_0R$ is an ideal of $R$. By hypothesis, there exists an idempotent $e \in R$ such that $r_R(RS_0R) = eR$. Now we claim that $r_T(TST) = r_T(S) = eT$. Assume that $e(x) \in S$. Then $e_0 \in S_0$, where $e_0$ is the constant term of $e(x)$. By Lemma 3.8, we have $e(x) = e(x)e_0$, so $e(x)e = e(x)e_0e = 0$. This implies that $eT \subseteq r_T(S)$. Now let $g(x) \in r_T(S)$. For each $f_0 \in S_0$, there exists $f(x) \in S$ such that $f_0$ is the constant term of $f(x)$. By Lemma 3.8, we have $f_0 = f_0 f(x)$.

Therefore $f_0g(x) = f_0f(x)g(x) = 0$. Thus $f_0g_i = 0$, where $g_i$ is the $i$-th coefficient in $g(x)$. Hence $g_i \in r_R(S_0) = eR$. This shows that $g(x) \in eT$.

(iv)⇒(i), let $T = R[x, x^{-1}]$ and $I = RSR$, where $S \subseteq S_r(R)$. Then $r_T(TST) = e(x)T$ for some idempotent $e(x) \in T$ . Since $Se(x) = 0$, it follows that $Se_0 = 0$ and hence $e_0 \in r_R(S) = r_R(I)$, where $e_0$ is the constant term of $e(x)$. Conversely, suppose that $b \in r_R(I)$. Then $b \in r_T(TST)$ and hence $b = e(x)b$. Thus $b = e_0b \in e_0R$. Therefore $r_R(I) = e_0R$. Since $e(x) \in S_r(T)$, it follows that $e_0$ is an idempotent in $R$ by Lemma 3.8. Therefore $R$ is an $EGE$-ring.

## 4. Semiprime $EGE$ (resp., $AE$)-ring

In this section, we show that for a semiprime ring $R$, the $EGE$- condition (resp., $AE$-condition) is equivalent to the closure of any union of clopen subsets of $Spec(R)$ is clopen (resp., $Spec(R)$ is an $EZ$-space).

For any $a \in R$, let $supp(a) = \{P \in \text{Spec(R)}: a \notin P\}$. Shin [19, Lemma 3.1] proved that for any $R$, $\{supp(a) : a \in R\}$ forms a basis of open sets on $Spec(R)$. This topology is called *hull-kernel topology*. We mean of $V(I)$ is the set of $P \in Spec(R)$, where $I \subseteq P$. We use $V(I)(V(a))$ to denote the set of $P \in Spec(R)$, where $I \subseteq P(a \in P)$. Note that $V(I) = \bigcap_{a\in I} V(a)$ (resp., $supp(I) = Spec(R) \setminus V(I)$) and $V(a) = Spec(R) \setminus supp(a)$.

For an open subset $A$ of $Spec(R)$, suppose that $O_A = \{a \in R : A \subseteq V(a)\}$. We can see that $O_A = \bigcap_{P\in A} P$ and $V(O_A) = clA$, where $clA$ is the cluster points of $A$ in $Spec(R)$.

**4.1. Lemma.** Let $R$ be a semiprime ring.

(i) For any $a \in R$, and any ideal $I$ of $R$, $supp(a) \cap supp(I) = supp(Ia)$.
(ii) If $I$ and $J$ are two ideals of $R$, then $r(I) \subseteq r(J)$ if and only if $intV(I) \subseteq intV(J)$.
(iii) $A \subseteq Spec(R)$ is a clopen subset if and only if there exists a central idempotent $e \in R$ such that $A = V(e) = supp(1 - e)$.
(iv) For open subsets $A, B$ of $Spec(R)$, $O_A = O_B$ if and only if $clB = clA$.
(v) For any ideal $I$ of $R$, $r(I) = O_{supp(I)}$.

*Proof.* For the proof of (i), (ii) and (iii) see [4, Lemma 4.2].

(iv) If $O_A = O_B$, then $clA = V(O_A) = V(O_B) = clB$. On the other hand for any subset $A$ of $Spec(R)$ we have $O_{clA} = O_A$, so $clA = clB$ implies that $O_A = O_B$.

(v) If $x \in r(I)$, then $ax = 0$, for all $a \in I$, so $supp(I) \subseteq V(r)$. This shows that $x \in O_{supp(I)}$. Now $x \in O_{supp(I)}$, implies that $supp(I) \subseteq V(x)$. By (i), $supp(Ix) = supp(I) \cap supp(x) = \emptyset$, so $Ix = 0$. This shows that $x \in r(I)$. □

Note that if $A$ is a subset of a topological space $X$, then $X \setminus intA = cl(X \setminus A)$.

**4.2. Theorem.** Let $R$ be a semiprime ring.

(i) $R$ is an $EGE$-ring if and only if the closure of any union of clopen subsets of $X = Spec(R)$ is clopen.
(ii) $R$ is an $AE$-ring if and only if $X = Spec(R)$ is an $EZ$-space.

*Proof.* (i) For each $\alpha \in S$, let $A_\alpha$ be a clopen subset of $X$. Then by Lemma 4.1, for each $\alpha \in S$ there exists a central idempoten $e_\alpha \in R$ (since in a semiprime ring $R$ semicentral idempotents are central) such that $A_\alpha = Supp(e_\alpha)$. Now let $I =< e_\alpha : \alpha \in S >$. By hypothesis, there is an idempotent $e \in R$ such that $r(I) = eR = r(R(1 - e))$. Now by lemma 4.1, $intV(I) = V(1 - e)$. Therefore we have $cl(\bigcup_{\alpha\in S} A_\alpha) = cl(\bigcup_{\alpha\in S} supp(e_\alpha)) = X \setminus int(\bigcap_{\alpha\in S} V(e_\alpha)) = X \setminus intV(I) = X \setminus V(1 - e) = supp(1 - e)$. Hence $cl(\bigcup_{\alpha\in S} A_\alpha)$ is open.

Conversely, let $I =< e_\alpha : \alpha \in S >$, where for each $\alpha \in S$, $e_\alpha$ is a right semicentral idempotent (hence a central idempotent). Then $K = \{V(e_\alpha) : \alpha \in S\}$ is a subset of clopen subsets of $X$. By hypothesis, $intV(I)$ is a clopen subset, because we have,

$$cl(\bigcup_{\alpha\in S} V(1-e_\alpha)) = X\setminus int(\bigcap_{\alpha\in S} V(e_\alpha)) = X\setminus intV(I).$$

Hence by Lemma 4.1, there is an idempotent $e\in R$ such that $intV(I)=V(e)=V(Re)$. Again by Lemma 4.1, $r(I)=r(Re)=(1-e)R$. Thus $R$ is an $EGE$-ring.

(ii) Let $A$ be an open subset of $Spec(R)$. Then there exists a subset $K$ of $R$ such that $A=supp[K]$. Now suppose that $I$ be the ideal generated by $K$ in $R$. Then by hypothesis and Lemma 4.1, there exists a subset $E$ of central idempotents of $R$ such that $O_A=r(I)=r(RER)=O_{supp[E]}$. Therefore, by Lemma 4.1, we have $cl(A)=cl(supp[E])$. Conversely, let $I$ be an ideal of $R$. Then we have $supp(I)$ is an open subset of $Spec(R)$. By hypothesis, there exists a collection $\{A_\alpha : \alpha\in S\}$ of clopen subsets of $Spec(R)$ such that $cl(supp(I))=cl(\bigcup_{\alpha\in S} A_\alpha)$. By Lemma 4.1, for each $\alpha\in S$ there exists an idempotent $e_\alpha$ such that $A_\alpha=supp(e_\alpha)$. Therefore, $cl(supp(I))=cl(\bigcup_{\alpha\in S} supp(e_\alpha))$. Again by Lemma 4.1, we have $r(I)=r(RER)=r(E)$ where $E=\{e_\alpha : \alpha\in S\}$. □

Recall that a ring $R$ is a right $SA$-ring if for each $I,J\trianglelefteq R$ there exists $K\trianglelefteq R$ such that $r(I)+r(J)=r(K)$ (see [4]). By [4, Theorem 4.4], a semiprime ring $R$ is a right SA-ring if and only if the space of prime ideals of $R$ is an extremally disconnected space if and only if $R$ is a quasi-Baer ring. Hence by Proposition 2.4, $R$ is a right SA if and only if $R$ is $EGE$ and AE.

*ACKNOWLEDGEMENTS*

The author wishes to thank the referee for her/his thorough reading of this paper and her/his comments which led to a much improved paper. The author also would like to thank Professor Gary F. Birkenmeier for his encouragement and discussion on this paper, particularly for suggestion which led to an improvement of Proposition 2.4.

## References

[1] Azarpanah, F and Karamzadeh, O.A. S. *Algebraic characterization of some disconnected spaces, Ital. J. Pure Appl. Math.* 12*,* 155–168*,* 2002.

[2] Berberian, S.K. Baer*-rings, Springer Berlin, (1972).

[3] Birkenmeier, G.F. *Idempotents and completely semiprime ideals, Commun. Algebra.* 11*,* 567–580*,* 1983.

[4] Birkenmeier, G.F. Ghirati, M and Taherifar, A. *When is a sum of annihilator ideals an annihilator ideal? Commun. Algebra.* 43*,* 2690-2702*,* 2015.

[5] Birkenmeier, G.F. Heatherly, H.E. Kim, J.Y and Park, J.K. *Triangular matrix representations, Journal of Algebra* 230*,* 558–595*,* 2000.

[6] Birkenmeier, G.F. and Huang, F.-K. Annihilator conditions on polynomials, Commun. Algebra, 29, 2097–2112, 2001.

[7] Birkenmeier, G.F. Kim, J.Y and Park, J.K. *A sheaf representation of quasi-Baer rings, Journal of Pure and Applied Algebra.* 146*,* 209–223*,* 2000.

[8] Birkenmeier, G.F. Kim , J.Y and Park, J.K. *Quasi-Baer ring extensions and biregular rings, Bull. AUSTRAL. Math. Soc.* 16*,* 39–52*,* 2000.

[9] Birkenmeier, G.F. Kim , J.Y and Park, J.K. *Principally Quasi-Baer Rings, Commun. Algebra,* 29*,* 639–660*,* 2001.

[10] Birkenmeier, G.F. Kim , J.Y and Park, J.K. *Polynomial extensions of Baer and quasi-Baer rings, Journal of Pure and Applied Algebra,* 159*,* 25–42*,* 2001.

[11] Birkenmeier, G.F. Kim , J.Y and Park, J.K. *Triangular matrix representations of semiprimary rings, Journal of Algebra and Its Applications* 1(2)*,* 123–131*,* 2002.

[12] Cheng, Y and Huang, F.K. *A note on extensions of principally quasi-baer rings, Taiwanese journal of mathematics.* 12(7)*,* 1721–1731*,* 2008.

[13] Clark, V. *Twisted matrix units semigroup algebra, Duke math. J.* 34*,* 417–424*,* 1967.

[14] Engelking, R. *General Topology, PWN-Polish Sci. Publ,* (1977).

[15] Gillman, L and Jerison, M. *Rings of Continuous Functions, Springer,* (1976).
[16] Lam, T.Y. *A First Course in Non-Commutative Rings, New York. Springer* (1991).
[17] Lam, T.Y. *Lecture on Modules and Rings, Springer, New York* (1999).
[18] Pollinigher, A and Zaks, A. *On Baer and quasi-Baer rings, Duke math. J.* 37*,* 127–138*,* 1970.
[19] Shin, G. *Prime ideals and sheaf representation of a pseudo symmetric ring, Trans. Amer. Math. Soc.* 184*,* 43–60*,* 1973.
[20] Taherifar, A. *Some new classes of topological spaces and annihilator ideals, Topol. Appl.* 165*,* 84–97*,* 2014.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 107 – 120

# ⊕-supplemented modules relative to an ideal

Rachid Tribak*, Yahya Talebi†, Ali Reza Moniri Hamzekolaee‡ and Samira Asgari§

**Abstract**

Let $I$ be an ideal of a ring $R$ and let $M$ be a left $R$-module. A submodule $L$ of $M$ is said to be $\delta$-small in $M$ provided $M \neq L + X$ for any proper submodule $X$ of $M$ with $M/X$ singular. An $R$-module $M$ is called *$I$-⊕-supplemented* if for every submodule $N$ of $M$, there exists a direct summand $K$ of $M$ such that $M = N + K$, $N \cap K \subseteq IK$ and $N \cap K$ is $\delta$-small in $K$. In this paper, we investigate some properties of $I$-⊕-supplemented modules. We also compare $I$-⊕-supplemented modules with ⊕-supplemented modules. The structure of $I$-⊕-supplemented modules and ⊕-$\delta$-supplemented modules over a Dedekind domain is completely determined.

**Keywords:** $\delta$-small submodules, ⊕-supplemented modules, ⊕-$\delta$-supplemented modules, $I$-⊕-supplemented modules

*2000 AMS Classification:* 16D10, 16D70, 16D99

*Received* : 16.02.2014 *Accepted* : 20.05.2014 *Doi* : 10.15672/HJMS.20164512490

## 1. Introduction

All rings considered in this paper will be associative with an identity element and $R$ will always denote a ring. We shall use $J(R)$ to denote the Jacobson radical of $R$. All modules will be unital left $R$-modules. Let $M$ be an $R$-module. A submodule $L$ of $M$ is called *small* ($\delta$-small) in $M$, denoted by $L \ll M$ ($L \ll_\delta M$), if $L + X \neq M$ for any proper submodule $X$ of $M$ ($L + X \neq M$ for any proper submodule $X$ of $M$ with $M/X$ singular). Recall that $M$ is called *⊕-supplemented* (*⊕-$\delta$-supplemented*) if for every submodule $N \leq M$, there exists a direct summand $K$ of $M$ such that $N + K = M$ and $N \cap K \ll K$ ($N \cap K \ll_\delta K$).

*Centre Régional des Métiers de l'Education et de la Formation (CRMEF)-Tanger, Avenue My Abdelaziz, BP 3117 Souani, Tanger, Morocco. Email: tribak12@yahoo.com

†Department of Mathematics, Faculty of Mathematical Sciences, University of Mazandaran, Babolsar, Iran. Email: talebi@umz.ac.ir

‡Department of Mathematics, Faculty of Mathematical Sciences, University of Mazandaran, Babolsar, Iran. Email: a.monirih@umz.ac.ir

§Department of Mathematics, Islamic Azad University, Mobarakeh Branch, Mobarakeh, Iran. Email: samira.asgari.a@gmail.com

In Section 2, we study some special cases of submodules $N$ of a module $M$ for which $N \ll_\delta M$ is equivalent to $N \ll M$.

In Section 3, we introduce the notion of $I$-$\oplus$-supplemented $R$-modules, where $I$ is an ideal of $R$. A module $M$ will be called *$I$-$\oplus$-supplemented* if for every submodule $N$ of $M$, there exists a direct summand $K$ of $M$ such that $M = N + K$, $N \cap K \subseteq IK$ and $N \cap K \ll_\delta K$. We shall compare this notion with the concept of $\oplus$-supplemented modules. Indecomposable $I$-$\oplus$-supplemented modules are characterized.

Section 4 is devoted to the study of some factor modules of an $I$-$\oplus$-supplemented module. Among other results, it is shown that if $M$ is a direct sum of two hollow $I$-$\oplus$-supplemented modules, then any direct summand of $M$ is $I$-$\oplus$-supplemented.

In Section 5, our main results (Theorems 5.4 and 5.13) describe the structure of $I$-$\oplus$-supplemented modules over Dedekind domains. It is also shown that over a Dedekind domain $R$, an $R$-module $M$ is $\oplus$-$\delta$-supplemented if and only if $M$ is $\oplus$-supplemented.

## 2. Some properties of $\delta$-small submodules

We begin with some results presenting some elementary properties of $\delta$-small submodules which will be used in the sequel.

**2.1. Lemma.** ([19, Lemma 1.2]) *Let $N$ be a submodule of a module $M$. The following are equivalent:*

(i) *$N$ is $\delta$-small in $M$;*

(ii) *If $X + N = M$, then $M = X \oplus Y$ for a projective semisimple submodule $Y$ with $Y \leq N$.*

**2.2. Lemma.** (See [19, Lemma 1.3])

(i) *Let $N$ and $K$ be submodules of a module $M$ with $K \subseteq N$. If $N \ll_\delta M$, then $K \ll_\delta M$.*

(ii) *Let $M$ and $M'$ be two modules. If $L \ll_\delta M$ and $f : M \to M'$ is a homomorphism, then $f(L) \ll_\delta M'$. In particular, if $K \ll_\delta M \leq M'$, then $K \ll_\delta M'$.*

(iii) *If $N$ and $L$ are submodules of a module $M$, then $N + L \ll_\delta M$ if and only if $N \ll_\delta M$ and $L \ll_\delta M$.*

(iv) *Let $M_1$ and $M_2$ be two submodules of a module $M$ such that $M = M_1 \oplus M_2$. Let $K_1 \leq M_1$ and $K_2 \leq M_2$. Then $K_1 \oplus K_2 \ll_\delta M_1 \oplus M_2$ if and only if $K_1 \ll_\delta M_1$ and $K_2 \ll_\delta M_2$.*

Let $N$ be a submodule of a module $M$. Recall that $N$ is said to be *DM* in $M$ (or $N$ *decomposes* $M$) if there is a direct summand $D$ of $M$ such that $D \leq N$ and $M = D + X$, whenever $N + X = M$ for a submodule $X$ of $M$ (see [1, Definition 3.1]). Clearly, the following implications hold:

$$(N \ll M) \Rightarrow (N \ll_\delta M) \Rightarrow (N \text{ is } DM \text{ in } M).$$

Next, we exhibit some conditions under which $N \ll_\delta M$ is equivalent to $N \ll M$.

**2.3. Proposition.** *Let $N$ be a proper submodule of an indecomposable module $M$. Then $N$ is DM in $M$ if and only if $N \ll_\delta M$ if and only if $N \ll M$.*

*Proof.* Assume that $N$ is *DM* in $M$. Let $X$ be a submodule of $M$ such that $M = N+X$. Then there exists a direct summand $D$ of $M$ such that $D \leq N$ and $M = D+X$. Since $M$ is indecomposable and $N \neq M$, we have $D = 0$ and $X = M$. Therefore, $N \ll M$. The rest of the proof is immediate. □

The next result was inspired by [16, Proposition 2.3(1)].

**2.4. Proposition.** *Let $N$ be a submodule of a module $M$. Then $N \ll M$ if and only if $N \subseteq Rad(M)$ and $N \ll_\delta M$.*

*Proof.* It is enough to prove the sufficiency. Let $X$ be a submodule of $M$ such that $M = N + X$. Since $N \ll_\delta M$, there exists a projective semisimple submodule $P \leq N$ such that $M = P \oplus X$.

Assume that $P \neq 0$. Then $P$ has a simple direct summand $S$. Since $S \subseteq Rad(M)$, $S \ll M$. Hence $S = 0$, a contradiction. Thus, $P = 0$. It follows that $N \ll M$. □

The following result is a direct consequence of Proposition 2.4.

**2.5. Corollary.** *Let $M$ be a module with $Rad(M) = M$ and let $N$ be a submodule of $M$. Then $N \ll_\delta M$ if and only if $N \ll M$.*

Let $M$ be a module over a commutative integral domain $R$. Let $T(M)$ denote the set of all elements $x \in M$ for which there exists a nonzero element $r \in R$ such that $rx = 0$. It is well known that $T(M)$ is a submodule of $M$. This submodule is called the *torsion submodule* of $M$. If $T(M) = M$, then the module $M$ is said to be a *torsion module*. The module $M$ is said to be *torsion-free* if $T(M) = 0$.

**2.6. Proposition.** *Assume that $R$ is a commutative integral domain. Let $M$ be an $R$-module and $N$ a submodule of $M$ such that $N \subseteq T(M)$. Then $N \ll_\delta M$ if and only if $N \ll M$.*

*Proof.* Assume that $N \ll_\delta M$. Let $X$ be a submodule of $M$ such that $N + X = M$. Then there exists a projective submodule $P \leq N$ such that $P \oplus X = M$. Since $P$ is projective, $P$ is isomorphic to a direct summand of a free $R$-module. Hence, $P$ is torsion-free. But $P$ is a torsion module as $P \subseteq N$. Then $P = 0$ and $X = M$. It follows that $N \ll M$. The converse is obvious. □

Let $N$ and $K$ be submodules of a module $M$. Recall that $K$ is said to be a *supplement* of $N$ in $M$ if $N + K = M$ and $N \cap K \ll K$. Let $M = \oplus_{i \in I} M_i$ be a decomposition of the module $M$. The next example shows that, in general, if $L = \oplus_{i \in I} L_i$ is a submodule of $M$ such that $L_i \ll_\delta M_i$ for each $i \in I$, then $L$ need not be $\delta$-small in $M$.

**2.7. Example.** Let $R$ be a discrete valuation ring with maximal ideal $m$. Let $M = \oplus_{i=1}^{\infty} R/m^i$. By [20, p. 48 The second corollary of Lemma 2.1], $Rad(M)$ does not have a supplement in $M$. Therefore, $Rad(M) = \oplus_{i=1}^{\infty} m/m^i$ is not small in $M$. Applying Proposition 2.6, it follows that $Rad(M)$ is not $\delta$-small in $M$. On the other hand, it is clear that for each $i \geq 1$, $m/m^i \ll R/m^i$.

**2.8. Proposition.** *Let $M = \oplus_{i \in I} M_i$ be a decomposition of a module $M$. Assume that for every submodule $N \leq M$, we have $N = \oplus_{i \in I}(N \cap M_i)$. For each $i$, let $L_i$ be a submodule of $M_i$. The following statements are equivalent:*

(i) *$L_i \ll_\delta M_i$ for every $i \in I$;*

(ii) *$L = \oplus_{i \in I} L_i \ll_\delta M$.*

*Proof.* (i) $\Rightarrow$ (ii) Let $X$ be a submodule of $M$ such that $M = X + L$. By hypothesis, $X = \oplus_{i \in I}(X \cap M_i)$. So, $(X \cap M_i) + L_i = M_i$ for every $i \in I$. By assumption, for every $i \in I$, there exists a semisimple projective submodule $P_i$ of $L_i$ such that $(X \cap M_i) \oplus P_i = M_i$ (see Lemma 2.1). Let $P = \oplus_{i \in I} P_i$. Then $X \oplus P = M$. Note that $P$ is a semisimple projective submodule of $L$. Therefore, $L \ll_\delta M$.

(ii) $\Rightarrow$ (i) By Lemma 2.2(iv). □

## 3. $I$-$\oplus$-supplemented modules

Recall that a module $M$ is called *$\oplus$-supplemented* (*$\oplus$-$\delta$-supplemented*) if for every submodule $N \leq M$, there exists a direct summand $K$ of $M$ such that $N + K = M$ and $N \cap K \ll K$ ($N \cap K \ll_\delta K$).

Recall that a ring $R$ is said to be *semilocal* provided $R/J(R)$ is a semisimple ring.

**3.1. Proposition.** *Let $M$ be a module over a semilocal ring $R$. Then $M$ is $\oplus$-supplemented if and only if for every submodule $N \leq M$, there exists a direct summand $K$ of $M$ such that $M = N + K$, $N \cap K \subseteq J(R)K$ and $N \cap K \ll_\delta K$.*

*Proof.* By Proposition 2.4 and [2, Corollary 15.18]. □

Motivated by the last proposition, we introduce the following notion:

**3.2. Definition.** Let $M$ be an $R$-module and let $I$ be an ideal of $R$. We say that $M$ is $I$-$\oplus$-supplemented, provided for every submodule $N$ of $M$, there exists a direct summand $K$ of $M$ such that $M = N + K$, $N \cap K \subseteq IK$ and $N \cap K \ll_\delta K$.

In this section we investigate some properties of $I$-$\oplus$-supplemented modules.

**3.3. Remark.** (i) It is clear that for every ideal $I$ of $R$, every $I$-$\oplus$-supplemented module is $\oplus$-$\delta$-supplemented.

(ii) Let $M$ be an $R$-module. If $I$ is an ideal of $R$ such that $IM = 0$, then $M$ is $I$-$\oplus$-supplemented if and only if $M$ is semisimple.

Let $M$ be an $R$-module. As in [19], let $\delta(M)$ denote the sum of all $\delta$-small submodules of $M$. In the next proposition we provide a condition under which a $\oplus$-$\delta$-supplemented module is $I$-$\oplus$-supplemented. To prove this result, we need the following elementary lemma.

**3.4. Lemma.** *Let $M$ be an $R$-module and let $I$ be an ideal of $R$. If $K$ is a direct summand of $M$, then we have $IK = K \cap IM$.*

*Proof.* Let $K'$ be a submodule of $M$ such that $M = K \oplus K'$. Then $IM = IK \oplus IK'$. Hence $K \cap IM = IK$. □

**3.5. Proposition.** *Let $M$ be an $R$-module and let $I$ be an ideal of $R$ such that $\delta(M) \subseteq IM$. Then $M$ is $I$-$\oplus$-supplemented if and only if $M$ is $\oplus$-$\delta$-supplemented.*

*Proof.* The necessity is clear. Conversely, suppose that $M$ is $\oplus$-$\delta$-supplemented. Let $N$ be a submodule of $M$. Then there exists a direct summand $K$ of $M$ such that $M = N + K$ and $N \cap K \ll_\delta K$. Note that $IK = K \cap IM$ by Lemma 3.4. Since $\delta(M) \subseteq IM$, we have

$$N \cap K \subseteq \delta(K) \subseteq K \cap \delta(M) \subseteq K \cap IM = IK.$$

Therefore $M$ is $I$-$\oplus$-supplemented. This completes the proof. □

Recall that a nonzero module $M$ is called *hollow* if every proper submodule is small in $M$. The module $M$ is called *local* if it has a proper submodule which contains all other proper submodules. Note that the largest proper submodule of a local module $M$ is $Rad(M)$. It is well known that every hollow module is $\oplus$-supplemented.

**3.6. Example.** (i) It is clear that every semisimple module is $I$-$\oplus$-supplemented for any ideal $I$ of $R$.

(ii) Let $p$ be a prime integer. It is well known that the $\mathbb{Z}$-module $\mathbb{Z}(p^\infty)$ is hollow and injective. It is easily seen that $\mathbb{Z}(p^\infty)$ is $I$-$\oplus$-supplemented for every nonzero ideal $I$ of $\mathbb{Z}$, but $\mathbb{Z}(p^\infty)$ is not 0-$\oplus$-supplemented.

(iii) It is easy to see that every $\oplus$-$\delta$-supplemented module (in particular, every $\oplus$-supplemented module) is $R$-$\oplus$-supplemented (see Proposition 3.5).

**3.7. Proposition.** *Let $M$ be an indecomposable $R$-module and let $I$ be an ideal of $R$. The following conditions are equivalent:*

(i) *$M$ is $I$-$\oplus$-supplemented;*

(ii) *$M$ is hollow with $IM = M$ or $IM = Rad(M)$.*

*Proof.* (i) $\Rightarrow$ (ii) Let $N$ be a proper submodule of $M$. By hypothesis, there exists a direct summand $K$ of $M$ such that $N+K=M$, $N\cap K\subseteq IK$ and $N\cap K\ll_\delta K$. Since $M$ is indecomposable, we have $K=M$. Hence, $N\subseteq IM$ and $N\ll_\delta M$. By Proposition 2.3, we have $N\ll M$. Thus, $M$ is a hollow module. Moreover, note that if $IM\neq M$, then $IM$ contains all other proper submodules of $M$. Hence $M$ is a local module and $IM=Rad(M)$.

(ii) $\Rightarrow$ (i) Let $N$ be a proper submodule of $M$. Then $N+M=M$, $N\cap M=N\subseteq Rad(M)\subseteq IM$ and $N\cap M=N\ll_\delta M$. Therefore, $M$ is $I$-$\oplus$-supplemented. □

It follows from Proposition 3.7 that if $I$ is an ideal of $R$, then every indecomposable $I$-$\oplus$-supplemented $R$-module is $\oplus$-supplemented. Next, we present some examples of $\oplus$-supplemented modules which are not $I$-$\oplus$-supplemented for an ideal $I$ of $R$.

**3.8. Example.** (i) Let $p$ and $q$ be two different prime integers. Consider the local $\mathbb{Z}$-module $M=\mathbb{Z}/\mathbb{Z}p^3$. We have $Rad(M)=\mathbb{Z}p/\mathbb{Z}p^3$. Let $I_1=\mathbb{Z}p$, $I_2=\mathbb{Z}q$ and $I_3=\mathbb{Z}p^2$. Then $I_1M=Rad(M)$, $I_2M=M$ and $I_3M=\mathbb{Z}p^2/\mathbb{Z}p^3$. By Proposition 3.7, $M$ is $I_i$-$\oplus$-supplemented for each $i=1,2$, but not $I_3$-$\oplus$-supplemented. On the other hand, it is clear that $M$ is $\oplus$-supplemented.

(ii) Let $R$ be a discrete valuation ring with maximal ideal $m$. It is well known that the $R$-module ${}_RR$ is $\oplus$-supplemented. Let $I$ be an ideal of $R$. From Proposition 3.7 it follows that ${}_RR$ is $I$-$\oplus$-supplemented if and only if $I=m$ or $I=R$. Therefore, the module ${}_RR$ is not $m^3$-$\oplus$-supplemented.

**3.9. Proposition.** *Let $I$ be an ideal of $R$ and let $M$ be an $R$-module.*

(i) *Assume that for every submodule $N\leq M$, there exists a submodule $K\leq M$ such that $M=N+K$ and $N\cap K\subseteq IM$. Then $M/IM$ is semisimple.*

(ii) *If $M$ is an $I$-$\oplus$-supplemented $R$-module, then $M/IM$ is semisimple.*

*Proof.* (i) Let $N$ be a submodule of $M$ such that $IM\subseteq N$. By assumption, there exists a submodule $K$ of $M$ such that $N+K=M$ and $N\cap K\subseteq IM$. Thus, $(N/IM)+[(K+IM)/IM]=M/IM$. Clearly, we have $N\cap(K+IM)=IM$. So, $N/IM$ is a direct summand of $M/IM$. This completes the proof.

(ii) follows from (i). □

**3.10. Proposition.** *Let $M$ be a module.*

(i) *If $M$ is $\oplus$-$\delta$-supplemented, then $M=M_1\oplus M_2$ such that $Rad(M_1)\ll M_1$ and $Rad(M_2)=M_2$.*

(ii) *If $M$ is $I$-$\oplus$-supplemented, then $M=M_1\oplus M_2$ such that $Rad(M_1)\subseteq IM_1$, $Rad(M_1)\ll M_1$ and $Rad(M_2)=M_2$.*

*Proof.* (i) Since $M$ is $\oplus$-$\delta$-supplemented, there exist submodules $M_1$ and $M_2$ of $M$ such that $M=M_1\oplus M_2$, $Rad(M)+M_1=M$ and $Rad(M)\cap M_1\ll_\delta M_1$. Note that $Rad(M)=Rad(M_1)\oplus Rad(M_2)$. Then $M_1\oplus Rad(M_2)=M$ and $(Rad(M)\cap M_1)\oplus Rad(M_2)=Rad(M)$. Therefore $Rad(M_2)=M_2$ and $Rad(M)\cap M_1=Rad(M_1)$. Moreover, we have $Rad(M_1)\ll M_1$ by Proposition 2.4. This completes the proof.

(ii) This follows by the same method as in (i) and adding the fact that $Rad(M)\cap M_1\subseteq IM_1$. □

Combining Proposition 3.10(ii) and [2, Proposition 5.20(1)], we get the following result.

**3.11. Corollary.** *If $M$ is an $I$-$\oplus$-supplemented module with $Rad(M)\ll M$, then $Rad(M)\subseteq IM$.*

From the last corollary, we conclude that if $I$ is an ideal of a left perfect ring $R$ and $M$ is an $I$-$\oplus$-supplemented $R$-module, then $Rad(M)\subseteq IM$ (see [2, Remark 28.5(3)]).

An $R$-module $M$ is said to be $\delta$-*local* if $\delta(M)\ll_\delta M$ and $\delta(M)$ is a maximal submodule of $M$ (see [4, Definition 3.1]). Next, we give an example of an $R$-$\oplus$-supplemented module which is not $\oplus$-supplemented.

**3.12. Example.** Let $F = \mathbb{Z}/\mathbb{Z}2$ and let $A = F^{\mathbb{N}}$ be the ring of sequences over $F$, whose operations are pointwise multiplication and pointwise addition. Let $R \subseteq A$ be the subring generated by $1_A$ (the unit element of $A$) and all sequences that have only a finite number of nonzero entries. It is shown in [4, p. 318] that the ring $R$ is not semilocal and the $R$-module ${}_RR$ is $\delta$-local. Applying [15, Proposition 3.1], it is easily seen that ${}_RR$ is an $R$-$\oplus$-supplemented module. On the other hand, since the ring $R$ is not semilocal, it is not semiperfect. Hence, the $R$-module ${}_RR$ is not $\oplus$-supplemented by [12, Corollary 4.42].

Next, we present conditions under which an $I$-$\oplus$-supplemented $R$-module is $\oplus$-supplemented.

**3.13. Proposition.** *Let $M$ be an $R$-module with $Rad(M) = M$. Then $M$ is $\oplus$-$\delta$-supplemented if and only if $M$ is $\oplus$-supplemented.*

*Proof.* As $Rad(M) = M$, we have $Rad(K) = K$ for every direct summand $K$ of $M$. The result follows from Corollary 2.5. □

**3.14. Proposition.** *Assume that $R$ is a commutative integral domain and let $M$ be a torsion $R$-module. Then $M$ is $\oplus$-$\delta$-supplemented if and only if $M$ is $\oplus$-supplemented.*

*Proof.* This follows from Proposition 2.6. □

**3.15. Proposition.** *Let $I$ be an ideal of $R$ and let $M$ be an $I$-$\oplus$-supplemented $R$-module. If $IM \subseteq Rad(M)$, then $M$ is $\oplus$-supplemented.*

*Proof.* Let $N$ be a submodule of $M$. By hypothesis, there exists a direct summand $K$ of $M$ such that $M = N + K$, $N \cap K \subseteq IK$ and $N \cap K \ll_\delta K$. Since $IM \subseteq Rad(M)$, we have $IK = K \cap IM \subseteq K \cap Rad(M) = Rad(K)$ by Lemma 3.4 and [5, 20.4(7)]. So $N \cap K \ll K$ by Proposition 2.4. It follows that $M$ is $\oplus$-supplemented. □

**3.16. Corollary.** *Let $I$ be an ideal of $R$ and let $M$ be an $I$-$\oplus$-supplemented $R$-module. Assume that one of the following conditions is satisfied:*

(i) *$I \subseteq J(R)$, or*

(ii) *$R$ is a local ring and $I \neq R$, or*

(iii) *$Rad(M) = M$, or*

(iv) *$R$ is a commutative integral domain and $M$ is a torsion $R$-module.*

*Then $M$ is $\oplus$-supplemented.*

*Proof.* (i) follows from [2, Corollary 15.18] and Proposition 3.15.

(ii) follows from (i).

(iii) follows easily from Proposition 3.13.

(iv) is obvious by Proposition 3.14. □

Next, we focus on when a $\oplus$-supplemented $R$-module is $I$-$\oplus$-supplemented for an ideal $I$ of $R$.

**3.17. Proposition.** *Let $I$ be an ideal of $R$ and let $M$ be a $\oplus$-supplemented $R$-module such that $Rad(M) \subseteq IM$. Then $M$ is $I$-$\oplus$-supplemented.*

*Proof.* Let $N$ be a submodule of $M$. Then there exists a direct summand $K$ of $M$ such that $M = N + K$ and $N \cap K \ll K$. Thus, $N \cap K \ll_\delta K$. Moreover, we have $IK = K \cap IM$ by Lemma 3.4. Since $Rad(M) \subseteq IM$, it follows that

$$Rad(K) \subseteq K \cap Rad(M) \subseteq K \cap IM = IK.$$

Hence, $N \cap K \subseteq IK$. Therefore $M$ is $I$-$\oplus$-supplemented. This completes the proof. □

The next corollary is a direct consequence of Proposition 3.17.

**3.18. Corollary.** *Let $M$ be a $\oplus$-supplemented module such that $IM = M$. Then $M$ is $I$-$\oplus$-supplemented.*

**3.19. Corollary.** *Let $m$ be a maximal ideal of a commutative ring $R$ and let $M$ be an $R$-module. Assume that $I$ is an ideal of $R$ such that $IM = mM$. If $M$ is a $\oplus$-supplemented $R$-module, then $M$ is $I$-$\oplus$-supplemented.*

*Proof.* Note that $Rad(M) \subseteq mM$ by [7, Lemma 3]. The result follows from Proposition 3.17. □

Let $R$ be a commutative integral domain. An $R$-module $M$ is called *divisible* in case $rM = M$ for each nonzero element $r \in R$.

**3.20. Corollary.** *Let $M$ be a divisible module over a commutative integral domain $R$. If $M$ is $\oplus$-supplemented, then $M$ is $I$-$\oplus$-supplemented for every nonzero ideal $I$ of $R$.*

*Proof.* This follows from Corollary 3.18. □

Recall that a ring $R$ is called a *left good* ring if $Rad(M) = J(R)M$ for every $R$-module $M$ (see [18, 23.7]).

**3.21. Corollary.** *Let $M$ be an $R$-module. Suppose further that either*
(i) *$R$ is a left good ring, or*
(ii) *$M$ is a projective module.*
*Then $M$ is $\oplus$-supplemented if and only if $M$ is $J(R)$-$\oplus$-supplemented.*

*Proof.* Note that $Rad(M) = J(R)M$ by [2, Proposition 17.10]. The result follows from Propositions 3.15 and 3.17. □

Combining Lemma 2.2 and the application of the same reasoning of [10, Proposition 3] to $I$-$\oplus$-supplemented modules, we obtain the following theorem.

**3.22. Theorem.** *Let $I$ be an ideal of $R$. Then any finite direct sum of $I$-$\oplus$-supplemented $R$-modules is $I$-$\oplus$-supplemented.*

The next example shows that, in general, a direct sum of $I$-$\oplus$-supplemented modules is not $I$-$\oplus$-supplemented.

**3.23. Example.** Let $p$ be a prime integer. Consider the $\mathbb{Z}$-module $M = \oplus_{i=1}^{\infty}\mathbb{Z}/\mathbb{Z}p^i$. Clearly, $M$ is a torsion module. By [12, Propositions A.7 and A.8], $M$ is not $\oplus$-supplemented. Therefore $M$ is not $(\mathbb{Z}p)$-$\oplus$-supplemented by Corollary 3.16. On the other hand, note that for every $i \geq 1$, $\mathbb{Z}/\mathbb{Z}p^i$ is a $(\mathbb{Z}p)$-$\oplus$-supplemented $\mathbb{Z}$-module by Proposition 3.7.

The next result deals with a special case of a family of $\oplus$-$\delta$-supplemented ($I$-$\oplus$-supplemented) modules $(M_\lambda)_{\lambda\in\Lambda}$ for which $M = \oplus_{\lambda\in\Lambda}M_\lambda$ is $\oplus$-$\delta$-supplemented ($I$-$\oplus$-supplemented).

**3.24. Proposition.** *Let $I$ be an ideal of $R$ and let $M = \oplus_{\lambda\in\Lambda}M_\lambda$ be a direct sum of submodules $M_\lambda$ ($\lambda \in \Lambda$) such that for every submodule $N$ of $M$, we have $N = \oplus_{\lambda\in\Lambda}(N \cap M_\lambda)$. Assume that $M_\lambda$ is $\oplus$-$\delta$-supplemented ($I$-$\oplus$-supplemented) for every $\lambda \in \Lambda$. Then $M$ is $\oplus$-$\delta$-supplemented ($I$-$\oplus$-supplemented).*

*Proof.* Let $N$ be a submodule of $M$. Then $N = \oplus_{\lambda\in\Lambda}(N \cap M_\lambda)$. For every $\lambda \in \Lambda$, there exists a direct summand $K_\lambda$ of $M_\lambda$ such that $(N \cap M_\lambda) + K_\lambda = M_\lambda$, $(N \cap K_\lambda \subseteq IK_\lambda)$ and $N \cap K_\lambda \ll_\delta K_\lambda$. Set $K = \oplus_{\lambda\in\Lambda}K_\lambda$. Clearly, $K$ is a direct summand of $M$ and $N + K = M$. Also, we have $(N \cap K = \oplus_{\lambda\in\Lambda}(N \cap K_\lambda) \subseteq IK)$ and $N \cap K \ll_\delta K$ by Proposition 2.8. This proves the proposition. □

## 4. Homomorphic images of $I$-$\oplus$-supplemented modules

We begin this section by an example showing that the $I$-$\oplus$-supplemented property does not always transfer from a module to each of its factor modules.

**4.1. Example.** Let $F$ be a field. Consider the local ring $R = F[x^2, x^3]/(x^4)$ and let $m$ be the maximal ideal of $R$. Let $n$ be an integer with $n \geq 2$ and let $M = R^{(n)}$. By Proposition 3.7 and Theorem 3.22, $M$ is $m$-$\oplus$-supplemented. Note that $R$ is an artinian local ring which is not a principal ideal ring (see [3, Example on p. 91]). So, there exists a submodule $K$ of $M$ such that the factor module $M/K$ is not $\oplus$-supplemented by [11, Example 2.2]. Therefore $M/K$ is not $m$-$\oplus$-supplemented by Corollary 3.16.

Next, we show that under some conditions, a factor module of an $I$-$\oplus$-supplemented module is $I$-$\oplus$-supplemented.

Recall that a submodule $N$ of a module $M$ is called *fully invariant* if $f(N) \subseteq N$ for every endomorphism $f$ of $M$. A module $M$ is called *distributive* if $(A+B)\cap C = (A\cap C)+(B\cap C)$ for all submodules $A$, $B$, $C$ of $M$ (or equivalently, $(A\cap B)+C = (A+C)\cap(B+C)$ for all submodules $A$, $B$, $C$ of $M$).

Analysis similar to the proofs of [6, Theorems 4.7 and 4.8] yields the following result. We give the first part of its proof for completeness.

**4.2. Proposition.** *Let $I$ be an ideal of $R$ and let $M$ be an $I$-$\oplus$-supplemented module.*

(i) *Let $X \leq M$ be a submodule such that for every direct summand $K$ of $M$, $(X+K)/X$ is a direct summand of $M/X$. Then $M/X$ is $I$-$\oplus$-supplemented.*

(ii) *Let $X \leq M$ be a submodule such that for every decomposition $M = M_1 \oplus M_2$, we have $X = (X\cap M_1)\oplus(X\cap M_2)$. Then $M/X$ is $I$-$\oplus$-supplemented.*

(iii) *If $X$ is a fully invariant submodule of $M$, then $M/X$ is $I$-$\oplus$-supplemented.*

(iv) *If $M$ is a distributive module, then $M/X$ is $I$-$\oplus$-supplemented for every submodule $X$ of $M$.*

*Proof.* (i) Let $N$ be a submodule of $M$ such that $X \subseteq N$. Since $M$ is $I$-$\oplus$-supplemented, there exists a direct summand $K$ of $M$ such that $N+K = M$, $N\cap K \subseteq IK$ and $N\cap K \ll_\delta K$. Therefore $(N/X)+((X+K)/X) = M/X$ and $(N/X)\cap((K+X)/X) = (X+(N\cap K))/X \subseteq ((X+IK)/X) \subseteq I((X+K)/X)$. Consider the natural epimorphism $\pi : K \to (X+K)/X$. Since $N\cap K \ll_\delta K$, we have $\pi(N\cap K) = (X+(N\cap K))/X \ll_\delta (X+K)/X$ by Lemma 2.2(ii). Note that by assumption, $(X+K)/X$ is a direct summand of $M/X$. It follows that $M/X$ is $I$-$\oplus$-supplemented.

(ii), (iii) and (iv) These are consequences of (i). □

The next proposition was inspired by [11, Proposition 2.5].

**4.3. Proposition.** *Let $M$ be an $R$-module and let $I$ be an ideal of $R$. Let $K$ be a fully invariant direct summand of $M$. Then the following assertions are equivalent:*

(i) *$M$ is $I$-$\oplus$-supplemented;*

(ii) *$K$ and $M/K$ are $I$-$\oplus$-supplemented.*

*Proof.* (i) $\Rightarrow$ (ii) Let $L$ be a submodule of $K$. By hypothesis, there exist submodules $A$ and $B$ of $M$ such that $M = A\oplus B$, $M = A+L$, $A\cap L \subseteq IA$ and $A\cap L \ll_\delta A$. Clearly, we have $K = (A\cap K)+L$. Since $K$ is fully invariant in $M$, we have $K = (A\cap K)\oplus(B\cap K)$. Hence, $A\cap K$ is a direct summand of $M$. Thus $I(A\cap K) = (A\cap K)\cap IM$ by Lemma 3.4. It follows that $(A\cap K)\cap L = A\cap L \subseteq (A\cap K)\cap IM = I(A\cap K)$. Since $A\cap L \ll_\delta A$ and $A\cap K$ is a direct summand of $A$, we have $A\cap L \ll_\delta A\cap K$ by Lemma 2.2(iv). Therefore, $K$ is $I$-$\oplus$-supplemented. Moreover, $M/K$ is $I$-$\oplus$-supplemented by Proposition 4.2(iii).

(ii) $\Rightarrow$ (i) This follows from Theorem 3.22. □

Let $I$ be an ideal of $R$. An $R$-module $M$ is called *completely* $I$-$\oplus$-*supplemented* ($\oplus$-*supplemented*) if every direct summand of $M$ is $I$-$\oplus$-supplemented ($\oplus$-*supplemented*). Clearly, semisimple modules are completely $I$-$\oplus$-supplemented. Also, every $I$-$\oplus$-supplemented hollow module is completely $I$-$\oplus$-supplemented. The next result provides another example of completely $I$-$\oplus$-supplemented modules.

Recall that a module $M$ is said to have *finite hollow dimension* $n \in \mathbb{N}$ if there exists a small epimorphism from $M$ to a direct sum of $n$ hollow modules. We denote this by $h.dim(M) = n$. It is well known that a module $M$ is hollow if and only if $h.dim(M) = 1$ (see [5, p. 47 and p. 49]).

**4.4. Proposition.** *Let $M = H_1 \oplus H_2$ be a direct sum of hollow submodules $H_1$ and $H_2$. Then the following statements are equivalent:*

(i) *$H_1$ and $H_2$ are $I$-$\oplus$-supplemented modules;*

(ii) *The module $M$ is completely $I$-$\oplus$-supplemented.*

*Proof.* (i) $\Rightarrow$ (ii) Let $L$ be a nonzero direct summand of $M$. If $L = M$, then $L$ is $I$-$\oplus$-supplemented by Theorem 3.22. Assume that $L \neq M$. Let $K$ be a submodule of $M$ such that $M = L \oplus K$. By [5, 5.4(1)], $h.dim(M) = 2 = h.dim(L) + h.dim(K)$. It follows that $h.dim(L) = 1$ and hence $L$ is a hollow module. Let us prove that $L$ is $I$-$\oplus$-supplemented. To see this, it suffices to show that $IL = L$ or $IL = Rad(L)$ by Proposition 3.7. Since $M$ is $I$-$\oplus$-supplemented, $M/IM \cong (L/IL) \oplus (K/IK)$ is semisimple by Proposition 3.9. As $L$ is a hollow module, $L/IL = 0$ or $L/IL$ is simple. Hence $L = IL$ or $L$ is a local module with maximal submodule $IL$. So $IL = L$ or $IL = Rad(L)$, as required.

(ii) $\Rightarrow$ (i) This is immediate. □

## 5. Modules over Dedekind domains

Our purpose in this section is to determine the structure of all $I$-$\oplus$-supplemented modules and all $\oplus$-$\delta$-supplemented modules over Dedekind domains.

**5.1. Proposition.** *Let $R$ be a Dedekind domain which is not a field. Then the following assertions are equivalent for an injective $R$-module $M$:*

(i) *$M$ is $\oplus$-supplemented;*

(ii) *$M$ is $I$-$\oplus$-supplemented for every nonzero ideal $I$ of $R$;*

(iii) *$M$ is $I$-$\oplus$-supplemented for some nonzero ideal $I$ of $R$;*

(iv) *$M$ is $\oplus$-$\delta$-supplemented.*

*Proof.* (i) $\Rightarrow$ (ii) This follows from Corollary 3.20 since the module $M$ is divisible.

(ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (iv) These are obvious.

(iv) $\Rightarrow$ (i) Since $R$ is a Dedekind domain which is not a field and $M$ is an injective $R$-module, we have $Rad(M) = M$. The result follows from Proposition 3.13. □

Let $R$ be a Dedekind domain which is not a field. If $M$ is an $R$-module, we will denote the sum of all divisible (injective) submodules of $M$ by $d(M)$. It is well known that $d(M)$ is an injective $R$-module. Also, note that if $f$ is an endomorphism of $M$, then $f(d(M))$ is isomorphic to a factor module of $d(M)$. So, $f(d(M))$ is injective as $R$ is a Dedekind domain. Therefore, $f(d(M)) \subseteq d(M)$. It follows that $d(M)$ is a fully invariant submodule of $M$.

**5.2. Proposition.** *Let $R$ be a Dedekind domain which is not a field. Let $I$ be an ideal of $R$ and let $M$ be an $R$-module. Then the following are equivalent:*

(i) *$M$ is $\oplus$-$\delta$-supplemented ($I$-$\oplus$-supplemented);*

(ii) *$M$ can be written as $M = M_1 \oplus M_2$ such that $M_1$ is injective, $Rad(M_2) \ll M_2$ and both of $M_1$ and $M_2$ are $\oplus$-$\delta$-supplemented ($I$-$\oplus$-supplemented) modules.*

*Proof.* (i) $\Rightarrow$ (ii) Let $M_1 = d(M)$ and let $M_2$ be a submodule of $M$ such that $M = M_1 \oplus M_2$. Note that $M_2$ has no submodules $X$ with $Rad(X) = X$. Since $M$ is $\oplus$-$\delta$-supplemented ($I$-$\oplus$-supplemented), $M_1$ and $M_2$ are $\oplus$-$\delta$-supplemented ($I$-$\oplus$-supplemented) by [14, Theorem 2.5] and Proposition 4.3. Moreover, we have $Rad(M_2) \ll M_2$ by Proposition 3.10.

(ii) $\Rightarrow$ (i) This follows by [14, Theorem 2.2] and Theorem 3.22. □

Next, we restrict our investigations about $\oplus$-$\delta$-supplemented modules and $I$-$\oplus$-supplemented modules to the case of modules over discrete valuation rings.

**5.3. Proposition.** *Let $M$ be a module over a discrete valuation ring $R$ and let $I$ be an ideal of $R$. Then $M$ is $\oplus$-$\delta$-supplemented if and only if $M$ is $\oplus$-supplemented. In particular, every $I$-$\oplus$-supplemented $R$-module is $\oplus$-supplemented.*

*Proof.* Assume that $M$ is $\oplus$-$\delta$-supplemented. By Proposition 5.2, $M = M_1 \oplus M_2$ is a direct sum of a $\oplus$-$\delta$-supplemented injective submodule $M_1$ and a submodule $M_2$ with $Rad(M_2) \ll M_2$. By Proposition 5.1, $M_1$ is $\oplus$-supplemented. In addition, $M_2$ is $\oplus$-supplemented by [20, Lemma 2.1] and [12, Proposition A.7]. Therefore, $M$ is $\oplus$-supplemented by [8, Theorem 1.4]. The converse is immediate.

The remaining assertion is obvious. □

Let $P$ be a nonzero prime ideal of a Dedekind domain $R$ and let $n$ be a nonzero natural number. We will use the notation $B_P(1, \dots, n)$ to denote the direct sum of arbitrarily many copies of $R/P$, $R/P^2$, …, $R/P^n$.

The next result provides a structure theorem for modules over a discrete valuation ring.

**5.4. Theorem.** *Assume that $R$ is a discrete valuation ring with maximal ideal $m$, quotient field $K$ and $Q = K/R$. Let $I$ be an ideal of $R$ and let $M$ be an $R$-module.*

(1) *If $I = m$ or $I = R$, then the following are equivalent:*

(i) *$M$ is $I$-$\oplus$-supplemented;*

(ii) *$M$ is $\oplus$-$\delta$-supplemented;*

(iii) *$M$ is $\oplus$-supplemented;*

(iv) *$M \cong R^a \oplus K^b \oplus Q^c \oplus B_m(1, \dots, n)$ for some natural numbers $a$, $b$, $c$ and $n$.*

(2) *If $I \notin \{m, R\}$, then the following are equivalent:*

(i) *$M$ is $I$-$\oplus$-supplemented;*

(ii) *$M \cong K^b \oplus Q^c \oplus (R/m)^{(\Lambda)}$ for some natural numbers $b$ and $c$ and an index set $\Lambda$.*

*Proof.* (1) (i) $\Leftrightarrow$ (iii) By Corollaries 3.18 and 3.19 and Proposition 5.3.

(ii) $\Leftrightarrow$ (iii) By Proposition 5.3.

(iii) $\Leftrightarrow$ (iv) This follows from [12, Proposition A.7].

(2) (i) $\Rightarrow$ (ii) Assume that $M$ is $I$-$\oplus$-supplemented. By Proposition 5.3, $M$ is $\oplus$-supplemented. Applying [12, Proposition A.7], $M \cong R^a \oplus K^b \oplus Q^c \oplus B_m(1, \dots, n)$ for some natural numbers $a$, $b$, $c$ and $n$. Since $M/IM$ is semisimple (see Proposition 3.9) and $I \notin \{m, R\}$, we have $a = 0$ and for each $1 \leq i \leq n$, $R/(I + m^i)$ is semisimple. So, for each $1 \leq i \leq n$, we have $I + m^i = m$ or $I + m^i = R$. Therefore $n = 1$ because $I \subseteq m^2$. It follows that $B_m(1, \dots, n) = B_m(1)$ is semisimple, completing the proof.

(ii) $\Rightarrow$ (i) Note that $K^b \oplus Q^c$ is an injective $\oplus$-supplemented module by [12, Proposition A.7]. The result follows from Propositions 5.1 and 5.2. □

**5.5. Remark.** Let $R$ be a discrete valuation ring with maximal ideal $m$, quotient field $K$ and $Q = K/R$. Let $I$ be an ideal of $R$.

(i) Assume that $I \notin \{m, R\}$. Theorem 5.4(2) and [12, Proposition A.7] provide many examples of $\oplus$-supplemented $R$-modules which are not $I$-$\oplus$-supplemented.

(ii) Note that [11, Corollary 4.5] shows that every $\oplus$-supplemented $R$-module is completely $\oplus$-supplemented.

*Case 1.* Assume that $I \in \{m, R\}$. Then every $I$-$\oplus$-supplemented $R$-module is completely $I$-$\oplus$-supplemented by Theorem 5.4.

*Case 2.* Suppose that $I \notin \{m, R\}$. Let $M$ be an $I$-$\oplus$-supplemented $R$-module. Then $M = K^b \oplus Q^c \oplus (R/m)^{(\Lambda)}$ for some natural numbers $b$ and $c$ and an index set $\Lambda$. Let $N$ and $L$ be submodules of $M$ such that $M = N \oplus L$ and let $d(M)$ be the sum of all injective submodules of $M$. It is clear that $d(M) = d(N) \oplus d(L) = K^b \oplus Q^c$. Then, $d(N) \cong K^{b'} \oplus Q^{c'}$ for some natural numbers $b'$ and $c'$ by [2, Corollary 12.7 and Lemma 25.4]. Therefore, $d(N)$ is $I$-$\oplus$-supplemented by Theorem 5.4. In addition, we have $(R/m)^{(\Lambda)} \cong M/d(M) \cong (N/d(N)) \oplus (L/d(L))$. Hence, $N/d(N)$ is semisimple. Thus, $N/d(N)$ is $I$-$\oplus$-supplemented. Since $d(N)$ is a direct summand of $N$, $N$ is $I$-$\oplus$-supplemented by Theorem 3.22. Consequently, $M$ is completely $I$-$\oplus$-supplemented.

Let $L$ be a submodule of a module $M$. A submodule $K \leq M$ is called a *$\delta$-supplement* of $N$ in $M$ if $M = L + K$ and $L \cap K \ll_\delta K$. The module $M$ is called *$\delta$-supplemented* if every submodule has a $\delta$-supplement in $M$.

Our next goal is to describe $\oplus$-$\delta$-supplemented modules and $I$-$\oplus$-supplemented modules over a nonlocal Dedekind domain $R$. The next proposition shows that every torsion-free $\delta$-supplemented $R$-module is injective. First we prove the following lemma.

**5.6. Lemma.** *Let $L$ be a proper submodule of a module $M$ such that $M/L$ is a cyclic module.*

(i) *If $K$ is a $\delta$-supplement of $L$ in $M$, then $K = P \oplus Rx$, where $P$ is a semisimple projective submodule of $L \cap K$ and $x \in K$. In this case, $Rx$ is a $\delta$-supplement of $L$ in $M$.*

(ii) *If $L$ has a $\delta$-supplement that is a direct summand of $M$, then $L$ has a cyclic $\delta$-supplement that is a direct summand of $M$.*

*Proof.* (i) By assumption, we have $L+K = M$ and $L\cap K \ll_\delta K$. Thus, $M/L \cong K/(L\cap K)$ is cyclic. Let $x \in K$ such that $K = (L \cap K) + Rx$. Since $L \cap K \ll_\delta K$, there exists a semisimple projective submodule $P$ of $L \cap K$ such that $K = P \oplus Rx$ by Lemma 2.1. Note that $L \cap K = L \cap (P \oplus Rx) = P \oplus (L \cap Rx) \ll_\delta P \oplus Rx$. By Lemma 2.2(iv), we have $P \ll_\delta P$ and $L \cap Rx \ll_\delta Rx$. Therefore $P$ is a semisimple projective module by [15, Lemma 2.9]. Also, note that $L + Rx = M$. It follows that $Rx$ is a $\delta$-supplement of $L$ in $M$.

(ii) follows from (i). □

**5.7. Proposition.** *Assume that $R$ is a Dedekind domain which is not local. Let $K$ denote the quotient field of $R$. If $M$ is a $\delta$-supplemented $R$-module, then $M/T(M) \cong K^{(\Lambda)}$ for some index set $\Lambda$.*

*Proof.* Assume that $M$ has a maximal submodule $L$ such that $T(M) \subseteq L$. Since $M$ is $\delta$-supplemented, there exists a cyclic submodule $W$ of $M$ such that $M = L+W$ and $L\cap W \ll_\delta W$ (see Lemma 5.6). Let $A$ be an ideal of $R$ such that $W \cong R/A$. Since $W$ is not contained in $L$, $W$ is not a torsion module. So $A = 0$ and $W \cong {}_RR$. Thus, $W$ is an indecomposable $R$-module. Hence $L\cap W \ll W$ by Proposition 2.3. Since $W/(L \cap W) \cong M/L$, we conclude that $W$ is a local submodule of $M$. This contradicts the fact that $R$ is not a local ring. It follows that $Rad(M/T(M)) = M/T(M)$. Hence, the module $M/T(M)$ is injective. So there exists an index set $\Lambda$ such that $M/T(M) \cong K^{(\Lambda)}$ by [9, Lemma 2.1]. □

**5.8. Proposition.** *Assume that $R$ is a Dedekind domain which is not local. If $M$ is a $\oplus$-$\delta$-supplemented $R$-module with $Rad(M) \ll M$, then $M$ is a torsion module.*

*Proof.* Since $M$ is $\oplus$-$\delta$-supplemented, there exist submodules $A$ and $B$ of $M$ such that $M = A \oplus B = T(M) + B$ and $T(M) \cap B \ll_\delta B$. Since $T(M) = T(A) \oplus T(B)$, we have $M = T(A) \oplus B$ and $T(M) = T(A) \oplus (T(M) \cap B)$. Hence $T(A) = A$ and $T(B) = T(M) \cap B$. So, $T(B) \ll_\delta B$. By Proposition 2.6, we have $T(B) \ll B$. Note that $M/T(M) \cong B/T(B)$ is divisible by Proposition 5.7. It follows that for every nonzero element $r \in R$, we have $rB + T(B) = B$. So, $rB = B$ for every $0 \neq r \in R$. This implies that $B$ is a divisible module, that is, $Rad(B) = B$ (see [9, Lemma 2.1]). But $Rad(B) \ll B$ since $Rad(M) \ll M$. Then $B = 0$ and $M = A$ is a torsion module, as required. □

**5.9. Proposition.** *Assume that $R$ is a nonlocal Dedekind domain. If $M$ is a $\oplus$-$\delta$-supplemented $R$-module, then $M$ is a torsion module.*

*Proof.* By Proposition 5.2, $M = M_1 \oplus M_2$ is a direct sum of $\oplus$-$\delta$-supplemented submodules $M_1$ and $M_2$ such that $Rad(M_1) = M_1$ and $Rad(M_2) \ll M_2$. By Proposition 5.1, $M_1$ is $\oplus$-supplemented. So, $M_1$ is a torsion module by [12, Proposition A.8]. Moreover, $M_2$ is a torsion module by Proposition 5.8. Therefore $M$ is a torsion module, as required. □

**5.10. Corollary.** *Assume that $R$ is a nonlocal Dedekind domain. An $R$-module $M$ is $\oplus$-$\delta$-supplemented if and only if $M$ is $\oplus$-supplemented.*

*Proof.* This follows easily from Propositions 3.14 and 5.9. □

**5.11. Remark.** Combining Proposition 5.3, Corollary 5.10 and [12, Propositions A.7 and A.8], we obtain the structure of $\oplus$-$\delta$-supplemented modules over Dedekind domains.

**5.12. Lemma.** *Assume that $R$ is a Dedekind domain which is not local. Let $P$ be a maximal ideal of $R$ and let $i$ be a nonzero natural number. Then:*

(i) *$I + P = P$ if and only if $I \subseteq P$.*

(ii) *If $i \geq 2$, then $I + P^i = P$ if and only if $I \subseteq P$ and $I \nsubseteq P^2$.*

(iii) *$I + P^i = R$ if and only if $I \nsubseteq P$.*

*Proof.* (i) and (iii) are immediate.

(ii) ($\Rightarrow$) This is obvious.

($\Leftarrow$) By hypothesis, we have $I = PI'$, where $I'$ is an ideal of $R$ which is not contained in $P$ (see [13, Theorem 6.14]). Since $I' + P^{(i-1)} = R$, we see that $PI' + P^i = P$. Hence, $I + P^i = P$. □

Let $M$ be a module over a Dedekind domain $R$ and let $P$ be a nonzero prime ideal of $R$. We will denote by $M_P$ the set $\{x \in M \mid P^n x = 0$ for some integer $n \geq 0\}$ which is called the $P$-primary component of $M$. Note that if $M$ is a torsion $R$-module, then $M$ is a direct sum of its $P$-primary components. Let $K$ be the quotient field of $R$. We will denote by $R(P^\infty)$ the $P$-primary component of the torsion $R$-module $K/R$. It is well known that $R(P^\infty)$ is a hollow module (see [9, Lemma 2.4]).

The next result describes the structure of $I$-$\oplus$-supplemented modules over nonlocal Dedekind domains. Recall that a module $M$ is 0-$\oplus$-supplemented if and only if $M$ is semisimple (see Remark 3.3(ii)).

**5.13. Theorem.** *Assume that $R$ is a nonlocal Dedekind domain. Let $I$ be a nonzero ideal of $R$. Then the following assertions are equivalent for an $R$-module $M$:*

(i) *$M$ is $I$-$\oplus$-supplemented;*

(ii) *$M$ is torsion and every $P$-primary component of $M$ is $I$-$\oplus$-supplemented;*

(iii) *$M$ is torsion and for every nonzero prime ideal $P$ of $R$, there exist natural numbers $a$ and $n$ such that $M_P \cong (R(P^\infty))^a \oplus B_P(1, \ldots, n)$ with $n = 1$ if $I \subseteq P^2$.*

*Proof.* (i) $\Leftrightarrow$ (ii) It is well known that for every nonzero prime ideal $P$ of $R$, $M_P$ is a fully invariant submodule of $M$. The result follows from Propositions 3.24, 4.3 and 5.9.

(ii) $\Rightarrow$ (iii) Let $P$ be a nonzero prime ideal of $R$. Since $M_P$ is $I$-$\oplus$-supplemented, $M_P$ is $\oplus$-supplemented by Corollary 5.10. Thus, there exist natural numbers $a$ and $n$ such that $M_P \cong (R(P^\infty))^a \oplus B_P(1,\ldots,n)$ by [12, Propositions A.7 and A.8]. Let $1 \leq i \leq n$. Since $M/IM$ is semisimple (see Proposition 3.9), $(R/P^i)/((I+P^i)/P^i) \cong R/(I+P^i)$ is semisimple. As $R/P^i$ is a local $R$-module, we have $I+P^i = R$ or $I+P^i = P$. Note that if $I \subseteq P^2$ and $i \geq 2$, then $I+P^i \subseteq P^2$. In this case we have $I+P^i \neq R$ and $I+P^i \neq P$. This shows that $I \subseteq P^2$ forces $n=1$.

(iii) $\Rightarrow$ (ii) Let $P$ be a nonzero prime ideal of $R$. Note that $M_P$ and $(R(P^\infty))^a$ are $\oplus$-supplemented by [12, Propositions A.7 and A.8]. We divide the rest of the proof into three cases:

*Case 1.* Assume that $I \subseteq P^2$. By hypothesis, $n=1$. Therefore $B_P(1,\ldots,n) = B_P(1)$ is semisimple. Hence $M_P \cong (R(P^\infty))^a \oplus B_P(1)$ is $I$-$\oplus$-supplemented (see Proposition 5.1 and Theorem 3.22).

*Case 2.* Suppose that $I \not\subseteq P^2$ and $I \not\subseteq P$. Then, $IM_P = M_P$ by Lemma 5.12(iii). Therefore, $M_P$ is $I$-$\oplus$-supplemented by Corollary 3.18.

*Case 3.* Assume that $I \not\subseteq P^2$ and $I \subseteq P$. In this case we have $IM_P = PM_P$ by Lemma 5.12. Applying Corollary 3.19, we conclude that $M_P$ is $I$-$\oplus$-supplemented. This completes the proof. □

**5.14. Remark.** Let $I$ be an ideal of a nonlocal Dedekind domain $R$. Using Theorem 5.13, [17, Theorem 1] and an analysis similar to that in Remark 5.5, we conclude that every $I$-$\oplus$-supplemented $R$-module is completely $I$-$\oplus$-supplemented.

## References


[1] Alkan, M., Nicholson, W. K. and Özcan, A. Ç. *A generalization of projective covers*, J. Algebra **319**, 4947-4960, 2008.

[2] Anderson, F. W. and Fuller, K. R. *Rings and Categories of Modules* (Springer-Verlag, New-York, 1992).

[3] Atiyah, M. F. and Macdonald, I. G. *Introduction to Commutative Algebra* (Addison-Wesley, London, 1969).

[4] Büyükaşik, E. and Lomp, C. *When δ-semiperfect rings are semiperfect*, Turkish J. Math. **34**, 317-324, 2010.

[5] Clark, J., Lomp, C., Vanaja, N. and Wisbauer, R. *Lifting Modules. Supplements and Projectivity in Module Theory* (Frontiers in Mathematics, Birkhäuser, Basel, 2006).

[6] Ecevit, Ş., Koşan, M. T. and Tribak, R. *Rad-⊕-supplemented modules and cofinitely Rad-⊕-supplemented modules*, Algebra Colloq. **19** (4), 637-648, 2012.

[7] Generalov, A. I. *ω-cohigh purity in the category of modules*, Math. Notes **33**, 402-408, 1983; translation from Mat. Zametki **33**, 785-796, 1983.

[8] Harmanci, A., Keskin, D. and Smith, P. F. *On ⊕-supplemented modules*, Acta Math. Hungar. **83** (1-2), 161-169, 1999.

[9] Hausen, J. *Supplemented modules over Dedekind domains*, Pacific J. Math. **100** (2), 387-402, 1982.

[10] Idelhadj, A. and Tribak, R. *A dual notion of CS-modules generalization*, in: Algebra and Number Theory (Fez) (M. Boulagouaz and J.-P. Tignol, eds.), Lecture Note of Pure and Appl. Math. **208** (Marcel Dekker, New York, 2000), 149-155.

[11] Idelhadj, A. and Tribak, R. *On some properties of ⊕-supplemented modules*, Internat. J. Math. Math. Sci. **69**, 4373-4387, 2003.

[12] Mohamed, S. H. and Müller, B. J. *Continuous and Discrete Modules*, London Math. Soc. Lecture Note Ser. **147** (Cambridge Univ. Press., Cambridge, 1990).

[13] Sharpe, D. W. and Vámos, P. *Injective Modules* Cambridge Tracts in Mathematics and Mathematical Physics **62** (Cambridge University Press, London, 1972).

[14] Talebi, Y. and Talaee, B. *Generalizations of $D_{11}$ and $D_{11}^+$ modules*, Asian-European J. Math. **2** (2), 285-293, 2009.

[15] Tribak, R. *On $\delta$-local modules and amply $\delta$-supplemented modules*, J. Algebra Appl. **12** (2), 1250144 (14 pages), 2013.

[16] Wang, Y. *A generalization of supplemented modules*, arXiv:1108.3381v1 [math.RA].

[17] Warfield Jr., R. B. *A Krull-Schmidt theorem for infinite sums of modules*, Proc. Amer. Math. Soc. **22**, 460-465, 1969.

[18] Wisbauer, R. *Foundations of Module and Ring Theory* (Gordon and Breach, Philadelphia, 1991).

[19] Zhou, Y. *Generalizations of perfect, semiperfect, and semiregular rings*, Algebra Colloq. **7** (3), 305-318, 2000.

[20] Zöschinger, H. *Komplementierte moduln über Dedekindringen*, J. Algebra **29**, 42-56, 1974.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 121 – 128

# Sharp results on linear combination of simple expressions of analytic functions

Nikola Tuneski*†, Teodor Bulboacă‡ and Biljana Jolevska-Tuneska§

**Abstract**

In this paper we use methods from the theory of differential subordinations to study the linear combination $azf''(z) + bf'(z) + c\frac{f(z)}{z}$ and give sharp bounds over the module, the argument and the real part of $\alpha f'(z) + \beta\frac{f(z)}{z}$.

**Keywords:** analytic, univalent, linear combination, differential inequality, differential subordination.

*2000 AMS Classification:* Primary: 30C45, Secondary: 30C80.

*Received* : 10.03.2015 *Accepted* : 17.06.2015 *Doi* : 10.15672/HJMS.20164513114

## 1. Introduction and preliminaries

Let $H(\mathbb{D})$ denote the class of all analytic functions in the open unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and let $\mathcal{A}$ be the subclass of $H(\mathbb{D})$ consisting of those functions that are normalized with the condition $f(0) = f'(0) - 1 = 0$. The functions of $\mathcal{A}$ that are one-to-one are called *normalized univalent functions* (for details see [4]).

A significant part of the theory of univalent functions deals with results over simple expressions of a function $f \in \mathcal{A}$ and its derivatives, such as

$$zf''(z), \quad f'(z) \quad \text{and} \quad \frac{f(z)}{z},$$

and a small part of such results are presented in [6] and [2].

In this paper we will study the differential operator $\mathrm{I}(a, b, c) : \mathcal{A} \to H(\mathbb{D})$ given by

$$\mathrm{I}(a, b, c)[f](z) := azf''(z) + bf'(z) + c\frac{f(z)}{z},$$

*Faculty of Mechanical Engineering, Ss. Cyril and Methodius University in Skopje, Karpoš II b.b., 1000 Skopje, Republic of Macedonia, Email: nikola.tuneski@mf.edu.mk

†Corresponding Author.

‡Faculty of Mathematics and Computer Science, Babeş-Bolyai University, 1, M. Kogălniceanu str., 400084 Cluj-Napoca, Romania, Email: bulboaca@math.ubbcluj.ro

§Faculty of Electrical Engineering and Informational Technologies, Ss. Cyril and Methodius University in Skopje, Karpoš II b.b., 1000 Skopje, Republic of Macedonia,
Email: biljanaj@feit.ukim.edu.mk

which is a linear combination of the above three expressions $zf''(z)$, $f'(z)$ and $f(z)/z$. Our results will provide simple sufficient conditions over the module, the argument and the real part of $\mathrm{I}(a,b,c)[f]$ that imply (in most of the cases) *sharp bound* of the module, the argument and the real part of another differential operator $\mathrm{J}(\alpha,\beta):\mathcal{A}\to H(\mathbb{D})$, given by

$$\mathrm{J}(\alpha,\beta)[f](z) := \alpha f'(z) + \beta\frac{f(z)}{z}.$$

Special choice of the parameters $a$, $b$ and $c$ leads to numerous known results, and we remind some of them:

(i) the special case $a=c=0$ and $b=1$ brings results over the class of functions of bounded turning [11, 12, 13];
(ii) the case $a=1$ and $b=c=0$ was studied in [9];
(iii) the case $b=1$ and $c=0$ was studied in [10];
(iv) the case $a=b=1$ and $c=0$ was studied in [3] and [1].

The sharpness of the results given in this paper closes, and in some cases improves and closes the related problems.

For our study we will use some methods from the theory of differential subordinations. Now we will recall the basic definitions and notion from this theory, that we need to use in our proofs. Valuable references on this topic may be found in [2] and [6].

If $f$ and $g$ are two analytic functions in the unit disk $\mathbb{D}$, then we say that $f$ is *subordinate* to $g$, written $f(z)\prec g(z)$, if there exists a *Schwarz function* $w$ (i.e. $w$ is analytic in $\mathbb{D}$ with $w(0)=0$ and $|w(z)|<1$, $z\in\mathbb{D}$) such that $f(z)=g(w(z))$ for all $z\in\mathbb{D}$. In particular, if $g$ is univalent in $\mathbb{D}$, then $f(z)\prec g(z)$ if and only if $f(0)=g(0)$ and $f(\mathbb{D})\subseteq g(\mathbb{D})$.

The general theory of differential subordinations was introduced by Miller and Mocanu in [7] and [8]. Namely, if $\phi:\mathbb{C}^2\to\mathbb{C}$ is analytic in a domain $D\subset\mathbb{C}$, if $h$ is univalent in $\mathbb{D}$, and if $p$ is analytic in $\mathbb{D}$ with $(p(z),zp'(z))\in D$ for all $z\in\mathbb{D}$, then $p$ is said to satisfy a first-order differential subordination if

$$\phi(p(z),zp'(z))\prec h(z). \tag{1.1}$$

A univalent function $q$ is said to be a *dominant* of the differential subordination (1.1) if $p(z)\prec q(z)$ for all the functions $p$ satisfying (1.1). If $\widetilde{q}$ is a dominant of (1.1) and $\widetilde{q}(z)\prec q(z)$ for all dominants of (1.1), then we say that $\widetilde{q}$ is the *best dominant* of the differential subordination (1.1).

To prove our main results we will use the following well-known lemmas from the theory of first-order differential subordinations:

**1.1. Lemma.** [8] *Let $q$ be univalent in the unit disk $\mathbb{D}$, and let $\theta$ and $\phi$ be analytic in a domain $\mathcal{D}$ containing $q(\mathbb{D})$, with $\phi(w)\neq 0$ for all $w\in q(\mathbb{D})$. Set $Q(z)=zq'(z)\phi(q(z))$, $h(z)=\theta(q(z))+Q(z)$, and suppose that*

(i) *$Q$ is starlike in the unit disk $\mathbb{D}$, and*
(ii) $\operatorname{Re}\frac{zh'(z)}{Q(z)}=\operatorname{Re}\left[\frac{\theta'(q(z))}{\phi(q(z))}+\frac{zQ'(z)}{Q(z)}\right]>0$, *$z\in\mathbb{D}$.*

*If $p$ is analytic in $\mathbb{D}$, with $p(0)=q(0)$, $p(\mathbb{D})\subseteq\mathcal{D}$ and*

$$\theta(p(z))+zp'(z)\phi(p(z))\prec\theta(q(z))+zq'(z)\phi(q(z))=h(z), \tag{1.2}$$

*then $p(z)\prec q(z)$, and $q$ is the best dominant of (1.2).*

**1.2. Lemma.** [7, page 11] *Let $n\geq 0$ be an integer and let $\gamma\in\mathbb{C}$, with $\operatorname{Re}\gamma>-n$. If $f(z)=\sum\limits_{m\geq n}a_mz^m$ is analytic in $\mathbb{D}$ and $F$ is defined by*

$$F(z)=\frac{1}{z^{\gamma}}\int_0^z f(\zeta)\zeta^{\gamma-1}d\zeta=\int_0^1 f(tz)t^{\gamma-1}dt,$$

*then* $F(z)=\sum\limits_{m\ge n}\frac{a_m z^m}{m+\gamma}$ *is analytic in* $\mathbb{D}$.

Now, by using Lemma 1.1 and Lemma 1.2 we will prove

**1.3. Lemma.** *Let* $\alpha,\beta,\gamma\in\mathbb{C}$ *with* $\operatorname{Re}\gamma>-2$*, and let* $q$ *be a univalent function in the unit disk* $\mathbb{D}$*, with* $q(0)=\alpha+\beta$*, satisfying*

$$\operatorname{Re}\left[1+\frac{zq''(z)}{q'(z)}\right]>\max\{0,-1-\operatorname{Re}\gamma\},\ z\in\mathbb{D}. \tag{1.3}$$

(i) *If* $f\in\mathcal{A}$*, then*

$$\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f](z)\prec zq'(z)+(1+\gamma)q(z)=:h(z), \tag{1.4}$$

*implies that*

$$\mathrm{J}(\alpha,\beta)[f](z)\prec q(z). \tag{1.5}$$

(ii) *Moreover, if we suppose, in addition, that*

$$\operatorname{Re}\frac{\beta}{\alpha}>-2,\ \textit{if}\quad \alpha\beta\neq 0,$$

*then the implication given in (i) is sharp. The extremal function* $f_*\in\mathcal{A}$ *that satisfies the subordination* (1.4) *such that* $\mathrm{J}(\alpha,\beta)[f_*]=q$*, is given by*

$$f_*(z)=\begin{cases}\dfrac{zq(z)}{\beta}, & \textit{if}\quad \alpha=0 \quad \textit{and}\quad \beta\neq 0,\\[2mm] z+\displaystyle\sum_{m\ge 1}\frac{q^{(m)}(0)}{(m+1)!\,\alpha}z^{m+1}, & \textit{if}\quad \alpha\neq 0\quad \textit{and}\quad \beta=0,\\[2mm] z+\displaystyle\sum_{m\ge 1}\frac{q^{(m)}(0)}{m!}\frac{z^{m+1}}{\beta+(m+1)\alpha}, & \textit{if}\quad \alpha\beta\neq 0\quad \textit{and}\quad \operatorname{Re}\frac{\beta}{\alpha}>-2.\end{cases}$$

*Proof.* (i) Let define functions $\theta(w)=(1+\gamma)w$ and $\phi(w)=1$, $w\in\mathbb{C}$, that are analytic in the domain $\mathcal{D}=\mathbb{C}$ which contains $q(\mathbb{D})$, and $\phi(w)\neq 0$ for all $w\in q(\mathbb{D})$. Further, the condition (1.3) implies that for the functions $Q(z)=zq'(z)\phi(q(z))=zq'(z)$ and $h(z)=\theta(q(z))+Q(z)=(1+\gamma)q(z)+zq'(z)$ we have

$$\operatorname{Re}\frac{zQ'(z)}{Q(z)}=\operatorname{Re}\left[1+\frac{zq''(z)}{q'(z)}\right]>0,\ z\in\mathbb{D},$$

and

$$\operatorname{Re}\frac{zh'(z)}{Q(z)}=\operatorname{Re}\left[2+\gamma+\frac{zq''(z)}{q'(z)}\right]=1+\operatorname{Re}\gamma+\operatorname{Re}\left[1+\frac{zq''(z)}{q'(z)}\right]>0,\ z\in\mathbb{D}.$$

So, the conditions (i) and (ii) from Lemma 1.1 hold, and moreover, the function $h$ is close-to-convex, hence univalent in $\mathbb{D}$ (see [5]).

Let now choose $p(z)=\mathrm{J}(\alpha,\beta)[f](z)=\alpha f'(z)+\beta\frac{f(z)}{z}$. Then, the function $p$ is analytic in $\mathbb{D}$, with $p(0)=q(0)=\alpha+\beta$, and $p(\mathbb{D})\subseteq\mathcal{D}=\mathbb{C}$. Finally, bearing in mind that

$$\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f](z)=zp'(z)+(1+\gamma)p(z),$$

we obtain that the subordinations (1.2) and (1.4) are equivalent, and the conclusion follows immediately from Lemma 1.1.

(ii) We will prove the second part of our lemma by showing that the differential equation

$$\alpha f'(z)+\beta\frac{f(z)}{z}=q(z) \tag{1.6}$$

has a solution $f_*\in\mathcal{A}$, whenever the assumptions (a) and (b) hold. For this purpose we will divide the analysis in three different cases.

*Case 1.* If $\alpha = 0$ and $\beta \neq 0$, since $q(0) = \alpha + \beta = \beta$, it follows that the equation (1.6) has solution $f_*(z) = \frac{zq(z)}{\beta} \in \mathcal{A}$.

*Case 2.* If $\alpha \neq 0$ and $\beta = 0$, then (1.6) has the analytic solution

$$f_*(z) = \frac{1}{\alpha} \int_0^z q(\zeta) d\zeta = z + \sum_{m \geq 1} \frac{q^{(m)}(0)}{(m+1)!\,\alpha}\, z^{m+1},$$

and since $q(0) = \alpha + \beta = \alpha$ the above function $f_*$ belongs to $\mathcal{A}$.

*Case 3.* Now let analyse the more complex case, assuming that $\alpha\beta \neq 0$. The equation (1.6) is equivalent to

$$f(z) + \frac{1}{\Gamma} z f'(z) = H(z), \quad \text{where} \quad \Gamma := \frac{\beta}{\alpha} \quad \text{and} \quad H(z) := \frac{zq(z)}{\beta}. \tag{1.7}$$

It is easy to see that

$$H(z) = \sum_{m \geq 1} a_m z^m, \ z \in \mathbb{D}, \quad \text{where} \quad a_m = \frac{q^{(m-1)}(0)}{(m-1)!\,\beta}.$$

Since the function $q$ is univalent in $\mathbb{D}$, and thus $q'(z) \neq 0$ for all $z \in \mathbb{D}$, it follows that the second coefficient of the above power expansion of $H$ does not vanish, i.e. $\frac{q'(0)}{\beta} \neq 0$.

Further, according to Lemma 1.2, the differential equation (1.7) has the analytic solution

$$f_*(z) = \frac{\Gamma}{z^\Gamma} \int_0^z H(\zeta) \zeta^{\Gamma - 1} d\zeta = \frac{1}{\alpha z^{\beta/\alpha}} \int_0^z q(\zeta) \zeta^{\beta/\alpha} d\zeta, \ z \in \mathbb{D}, \tag{1.8}$$

whenever $\operatorname{Re}\Gamma > -n$

This means that the solution (1.8) of the differential equation (1.6) is analytic in $\mathbb{D}$ if we assume that

$$\operatorname{Re}\Gamma = \operatorname{Re}\frac{\beta}{\alpha} > -2.$$

Hence, the solution (1.8) of the differential equation (1.6) is an analytic function in $\mathbb{D}$, and has the form

$$f_*(z) = z + \sum_{m \geq 1} \frac{q^{(m)}(0)}{m!} \frac{z^{m+1}}{\beta + (m+1)\alpha}, \ z \in \mathbb{D},$$

that is $f_* \in \mathcal{A}$. □

## 2. Results over the module

In this section we study the module of $\mathrm{I}(a, b, c)[f]$ and receive sharp information about the module of $\mathrm{J}(\alpha, \beta)[f]$.

**2.1. Theorem.** *Let $\alpha, \beta, \gamma \in \mathbb{C}$ with $\beta + 2\alpha \neq 0$, $\operatorname{Re}\gamma > -2$, and let $\delta > 0$. If $f \in \mathcal{A}$, then*

$$\left| \mathrm{I}(\alpha, \alpha + \beta + \alpha\gamma, \beta\gamma)[f](z) - (1+\gamma)(\alpha+\beta) \right| < \delta, \ z \in \mathbb{D}, \tag{2.1}$$

*implies*

$$\left| \mathrm{J}(\alpha, \beta)[f](z) - (\alpha + \beta) \right| < \Delta = \Delta(\gamma, \delta) := \frac{\delta}{|2+\gamma|}, \ z \in \mathbb{D}. \tag{2.2}$$

*This implication is sharp, and the extremal function is*

$$f_*(z) = z + \frac{\delta}{(2+\gamma)(\beta+2\alpha)} z^2. \tag{2.3}$$

*Proof.* If we denote $q(z)=\alpha+\beta+\frac{\delta}{2+\gamma}z$, $z\in\mathbb{D}$, then we have $1+\frac{zq''(z)}{q'(z)}=1$, meaning that (1.3) from Lemma 1.3 holds because of the assumption $\operatorname{Re}\gamma>-2$. Further, the function $h$ defined in (1.4) will be

$$h(z)=(1+\gamma)(\alpha+\beta)+\delta z,\ z\in\mathbb{D},$$

hence the subordination (1.4) is equivalent to (2.1). Therefore, (2.2) follows directly from Lemma 1.3 and the definition of subordination, while a simple computation shows that $f_*$ given by (2.3) is the extremal function. □

If we consider in the above theorem the special case $\alpha=-\beta=1$, then we obtain the first part, while for $\beta=1-\alpha$ we obtain the second part of the next corollary:

**2.2. Corollary.** *Let $\delta>0$ and $\gamma\in\mathbb{C}$, with $\operatorname{Re}\gamma>-2$.*

(i) *If $f\in\mathcal{A}$, then*

$$\left|zf''(z)+\gamma f'(z)-\gamma\frac{f(z)}{z}\right|<\delta,\ z\in\mathbb{D},$$

*implies*

$$\left|f'(z)-\frac{f(z)}{z}\right|<\Delta=\Delta(\gamma,\delta),\ z\in\mathbb{D}.$$

*This implication is sharp, and the extremal function is $f_*(z)=z+\frac{\delta}{2+\gamma}z^2$.*

(ii) *Assuming that $\alpha\in\mathbb{C}\setminus\{-1\}$, if $f\in\mathcal{A}$, then*

$$\left|\alpha zf''(z)+(1+\alpha\gamma)f'(z)+(1-\alpha)\gamma\frac{f(z)}{z}-(1+\gamma)\right|<\delta,\ z\in\mathbb{D},$$

*implies*

$$\left|\alpha f'(z)+(1-\alpha)\frac{f(z)}{z}-1\right|<\Delta=\Delta(\gamma,\delta),\ z\in\mathbb{D}.$$

*This implication is sharp, and the extremal function is $f_*(z)=z+\frac{\delta}{(2+\gamma)(\alpha+1)}z^2$.*

## 3. Results over the argument and the real part

In this section we study the argument of the $\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f]$ and obtain a result for the argument of $\mathrm{J}(\alpha,\beta)[f]$. As an interesting consequence, we receive the corresponding result over the real parts of these operators.

**3.1. Theorem.** *Let $\alpha,\ \beta\in\mathbb{C}$ and $\gamma,\ \lambda\in\mathbb{R}$, such that $\alpha+\beta>\lambda$, $\gamma\geq-1$, and also let $\theta\in(0,1]$.*

(i) *If $f\in\mathcal{A}$, then*

$$\left|\arg\left[\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f](z)-(1+\gamma)\lambda\right]\right|<\Delta,\ z\in\mathbb{D},\tag{3.1}$$

*where*

$$\Delta=\Delta(\theta,\gamma):=\begin{cases}\frac{\theta\pi}{2}+\arctan\frac{\theta}{1+\gamma}, & \textit{if}\quad \gamma>-1,\\ \frac{(\theta+1)\pi}{2}, & \textit{if}\quad \gamma=-1\end{cases}$$

*implies*

$$\left|\arg\mathrm{J}(\alpha,\beta)[f](z)-\lambda\right|<\frac{\theta\pi}{2},\ z\in\mathbb{D}.\tag{3.2}$$

(ii) *Moreover, for the special case* $\gamma = -1$*, if we suppose, in addition, that*

$$\operatorname{Re}\frac{\beta}{\alpha} > -2,\ \textit{if}\quad \alpha\beta \neq 0,$$

*then the implication given in (i) is sharp.*

*Proof.* The assumption $\alpha + \beta > \lambda$, with $\lambda \in \mathbb{R}$, holds if and only if $\operatorname{Im}\alpha = -\operatorname{Im}\beta$ and $\operatorname{Re}\alpha + \operatorname{Re}\beta > \lambda$.

Let consider the function $q(z) = \left(\frac{1+z}{1-z}\right)^{\theta}(\operatorname{Re}\alpha + \operatorname{Re}\beta - \lambda) + \lambda,\ z \in \mathbb{D}$, where the power is taken to its principal value. It follows that $q(0) = \operatorname{Re}\alpha + \operatorname{Re}\beta$,

$$1 + \frac{zq''(z)}{q'(z)} =: H(z) = -1 + \frac{2(1+\theta z)}{1-z^2},\ z \in \mathbb{D},$$

hence

$$H(e^{i\varphi}) = i\frac{\theta + \cos\varphi}{\sin\varphi},\ \varphi \in (-\pi, 0) \cup (0, \pi).$$

Since $H(0) = 1$, we deduce $\operatorname{Re} H(z) > 0$, $z \in \mathbb{D}$, and from $q'(0) = 2\theta\,(\operatorname{Re}\alpha + \operatorname{Re}\beta - \lambda) \neq 0$ we conclude that $q$ is a convex (univalent) function in $\mathbb{D}$.

Thus, the initial assumptions of Lemma 1.3 are satisfied. The proof will be completed if we show that the inequality (3.1) implies the subordination (1.4), and the subordination (1.5) is equivalent to the inequality (3.2).

The function $h$ defined in the subordination (1.4) has the form

$$h(z) = zq'(z) + (1+\gamma)q(z) = (\operatorname{Re}\alpha + \operatorname{Re}\beta - \lambda)\left(\frac{1+z}{1-z}\right)^{\theta}\left(\frac{2\theta z}{1-z^2} + 1 + \gamma\right) + (1+\gamma)\lambda.$$

Even more, since $(h(z) - h(0))/h'(0)$ is a close-to-convex (normalized) function in $\mathbb{D}$, it follows that the function $h$ is univalent.

Now, for $z = e^{i\varphi}$ and $\varphi \in (-\pi, 0) \cup (0, \pi)$, using the fact that $\alpha + \beta > \lambda$ we get

$$\begin{aligned}\arg\left[h(e^{i\varphi}) - (1+\gamma)\lambda\right] &= \arg\left[\left(\tfrac{1+e^{i\varphi}}{1-e^{i\varphi}}\right)^{\theta}\left(\tfrac{2\theta e^{i\varphi}}{1-e^{2i\varphi}} + 1 + \gamma\right)\right] \\ &= \arg\left[\left(i\cot\tfrac{\varphi}{2}\right)^{\theta}\left(\tfrac{i\theta}{\sin\varphi} + 1 + \gamma\right)\right],\end{aligned}$$

hence

$$\arg\left[h(e^{i\varphi}) - (1+\gamma)\lambda\right] = \theta\arg\left(i\cot\frac{\varphi}{2}\right) + \arg\left(\frac{i\theta}{\sin\varphi} + 1 + \gamma\right). \tag{3.3}$$

We will discuss now the following two cases.

*Case 1.* If $\gamma > -1$, from the relation (3.3) we easily deduce that

$$\arg\left[h(e^{i\varphi}) - (1+\gamma)\lambda\right] \geq \frac{\theta\pi}{2} + \arctan\frac{\theta}{1+\gamma},\ \text{if}\ \varphi \in (0, \pi),$$

and

$$\arg\left[h(e^{i\varphi}) - (1+\gamma)\lambda\right] \leq -\left(\frac{\theta\pi}{2} + \arctan\frac{\theta}{1+\gamma}\right),\ \text{if}\ \varphi \in (-\pi, 0),$$

which implies

$$\left|\arg\left[h(e^{i\varphi}) - (1+\gamma)\lambda\right]\right| \geq \frac{\theta\pi}{2} + \arctan\frac{\theta}{1+\gamma},\ \text{for}\ \varphi \in (-\pi, 0) \cup (0, \pi).$$

These inequalities, combined with the fact that $h(0) = (\operatorname{Re}\alpha + \operatorname{Re}\beta)(1+\gamma) > (1+\gamma)\lambda$ and the fact that $h$ is univalent in $\mathbb{D}$, leads to the conclusion that

$$\Omega := \left\{w \in \mathbb{C} : \left|\arg\left[w - (1+\gamma)\lambda\right]\right| < \frac{\theta\pi}{2} + \arctan\frac{\theta}{1+\gamma}\right\} \subset h(\mathbb{D}), \tag{3.4}$$

and from (3.3) we may see that

$$\left|\arg\left[h(e^{i\varphi})-(1+\gamma)\lambda\right]\right|=\frac{\theta\pi}{2}+\arctan\frac{\theta}{1+\gamma}\Leftrightarrow\varphi\in\left\{-\frac{\pi}{2},\frac{\pi}{2}\right\},$$

hence $\Omega\neq h(\mathbb{D})$.

The assumption (3.1) is equivalent to $\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f](\mathbb{D})\subset\Omega$, and according to (3.4) this inclusion implies $\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f](\mathbb{D})\subset h(\mathbb{D})$. Using the fact that $\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f](0)=h(0)=(\operatorname{Re}\alpha+\operatorname{Re}\beta)(1+\gamma)$, this inclusion is equivalent to the subordination (1.4), and by Lemma 1.3 if follows that $\mathrm{J}(\alpha,\beta)[f](z)\prec q(z)$, which is equivalent to our conclusion (3.2).

*Case 2.* If $\gamma=-1$, from the relation (3.3) we similarly deduce that the conclusion is even stronger, that is

$$h(\mathbb{D})=\left\{w\in\mathbb{C}:|\arg w|<\frac{(\theta+1)\pi}{2}\right\}.$$

In this case, the assumption (3.1) is equivalent to $\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f](\mathbb{D})=h(\mathbb{D})$, and from $\mathrm{I}(\alpha,\alpha+\beta+\alpha\gamma,\beta\gamma)[f](0)=h(0)=(\operatorname{Re}\alpha+\operatorname{Re}\beta)(1+\gamma)$, the previous equality between the two sets is equivalent to the subordination (1.4). Thus, from Lemma 1.3 it follows that $\mathrm{J}(\alpha,\beta)[f](z)\prec q(z)$, which is equivalent to our conclusion (3.2).

Notice that, for $\gamma=-1$ the result is sharp. That's because the assumption (3.1) is equivalent to the subordination (1.4), while in the case $\gamma>-1$ the assumption (3.1) is stronger than (1.4). □

By specifying values of some of the parameters in Theorem 3.1 we receive the next results:

**3.2. Corollary.** *Let $\alpha,\ \beta\in\mathbb{C}$ and $\gamma,\ \lambda\in\mathbb{R}$, such that $\alpha+\beta>\lambda$, $\gamma\geq-1$.*

(i) *If $f\in\mathcal{A}$, then*

$$\operatorname{Re}\left[\alpha zf''(z)+(\alpha+\beta+\alpha\gamma)f'(z)+\beta\gamma\frac{f(z)}{z}\right]>(1+\gamma)\lambda,\ z\in\mathbb{D},$$

*implies*

$$\operatorname{Re}\left[\alpha f'(z)+\beta\frac{f(z)}{z}\right]>\lambda,\ z\in\mathbb{D}.$$

*Moreover, for the special case $\gamma=-1$, if we suppose, in addition, that*

$$(a)\quad \operatorname{Re}\frac{\beta}{\alpha}>-2,\ \text{if}\quad \alpha\beta\neq0,$$

*then the implication given in (i) is sharp.*

(ii) *Assuming that $\lambda<0$, if $f\in\mathcal{A}$, then*

$$\operatorname{Re}\left[zf''(z)+\gamma f'(z)-\gamma\frac{f(z)}{z}\right]>(1+\gamma)\lambda,\ z\in\mathbb{D},$$

*implies*

$$\operatorname{Re}\left[f'(z)-\frac{f(z)}{z}\right]>\lambda,\ z\in\mathbb{D}.$$

*Moreover, for the special case $\gamma=-1$ the implication given in (ii) is sharp.*

(iii) *Assuming that and $\lambda<1$, if $f\in\mathcal{A}$, then*

$$\operatorname{Re}\left[\alpha zf''(z)+(1+\alpha\gamma)f'(z)+(1-\alpha)\gamma\frac{f(z)}{z}\right]>(1+\gamma)\lambda,\ z\in\mathbb{D},$$

*implies*

$$\operatorname{Re}\left[\alpha f'(z)+(1-\alpha)\frac{f(z)}{z}\right]>\lambda,\ z\in\mathbb{D}.$$

*Moreover, for the special case* $\gamma = -1$, *if we suppose, in addition, that*

$$(a') \qquad \left|\alpha + \frac{1}{2}\right| > \frac{1}{2} \quad \textit{if} \quad \alpha \notin \{0; 1\},$$

*then the implication given in (iii) is sharp.*

*Proof.* The implication (i) follows directly from Theorem 3.1 for $\theta = 1$ having in mind that $\frac{\pi}{2} + \arctan\frac{1}{1+\gamma} > \frac{\pi}{2}$. The implication (iii) follows from (i) if we choose $\alpha+\beta = 1$, and in that case the assumption (a) is equivalent to (a'). The implication (ii) was obtained from (i) for the special case $\alpha = -\beta = 1$. □

**3.3. Remark.** Taking $\alpha = 1$ and $\beta = \gamma = 0$ in Corollary 3.2(i), we receive that for $\lambda < 1$ and $f \in \mathcal{A}$ the following implication holds:

$$\operatorname{Re}[zf''(z) + f'(z)] > \lambda,\ z \in \mathbb{D} \Rightarrow \operatorname{Re} f'(z) > \lambda,\ z \in \mathbb{D}. \tag{3.5}$$

This improves the result given in Theorem 1(a) from [1] where it was proven that

$$\operatorname{Re}[zf''(z) + f'(z)] > \lambda,\ z \in \mathbb{D} \Rightarrow \operatorname{Re} f'(z) > 1 + 2(1-\lambda)(\log 2 - 1),\ z \in \mathbb{D}. \tag{3.6}$$

Implication (3.5) is stronger than the implication (3.6), since for $\lambda < 1$ we have

$$1 + 2(1-\lambda)(\log 2 - 1) < \lambda.$$

**Acknowledgment.** The work of the second author was entirely supported by the grant given by *Babeş-Bolyai* University dedicated for *Supporting the Excellence Research 2015*.

## References

[1] Ali R. M., On a subclass of starlike functions, *Rocky Mountain J. Math.* 24 (1994), no. 2, 447–451.

[2] Bulboacă T., Differential subordinations and superordinations. New results, *House of Science Book Publ.*, Cluj-Napoca, 2005.

[3] Chichra P. N., New subclasses of the class of close-to-convex functions, *Proc. Amer. Math. Soc.* 62 (1977), 37–43.

[4] Duren P. L., Univalent functions, *Springer-Verlag*, 1983.

[5] Kaplan W., Close to convex schlicht functions, *Michig. Math. J.* 1 (1952), no. 2, 169–185.

[6] Miller S. S., Mocanu P. T., Differential subordinations, Theory and Applications, *Marcel Dekker*, New York-Basel, 2000.

[7] Miller S. S., Mocanu P. T., Differential subordinations and univalent functions, *Michigan Math. J.* 28 (1981), 157–171.

[8] Miller S. S., Mocanu P. T., On some classes of first-order differential subordinations, *Michigan Math. J.* 32 (1985), 185–195.

[9] Obradović M., Simple sufficient conditions for starlikeness, *Mat. Vesnik* 49 (1997), 241–244.

[10] Ponnusamy S., Singh V., Criteria for univalent, starlike and convex functions, *Bull. Belg. Math. Soc. Simon Stevin* 9 (2002), no. 4, 511–531.

[11] Sokol J., Some applications of differential subordinations in the geometric function theory, *J. Inequal. Appl.* 2013, 2013:389, 11 pp.

[12] Tuneski N., Some simple sufficient conditions for starlikeness and convexity, *Appl. Math. Lett.* 22 (2009), 693–697.

[13] Tuneski N., Obradović M., Some properties of certain expression of analytic functions, *Comput. Math. Appl.* 62 (2011), 3438–3445.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 129 – 137

# A generalization of supplemented modules

Yongduo Wang and Dejun Wu*

**Abstract**

Let $M$ be a left module over a ring $R$ and $I$ an ideal of $R$. $M$ is called an $I$-supplemented module (finitely $I$-supplemented module) if for every submodule (finitely generated submodule ) $X$ of $M$, there is a submodule $Y$ of $M$ such that $X + Y = M$, $X \cap Y \subseteq IY$ and $X \cap Y$ is PSD in $Y$. This definition generalizes supplemented modules and $\delta$-supplemented modules. We characterize $I$-semiregular, $I$-semiperfect and $I$-perfect rings which are defined by Yousif and Zhou [12] using $I$-supplemented modules. Some well known results are obtained as corollaries.

**Keywords:** semiregular, semiperfect, supplemented module, small submodule.

*2000 AMS Classification:* 16D10, 16D50.

*Received* : 03.12.2013 *Accepted* : 02.12.2014 *Doi* : 10.15672/HJMS.20164512487

## 1. Introduction and Preliminaries

It is well known that supplemented modules play an important role in characterizing semiperfect, semiregular and perfect rings. Recently, some authors had worked with various extensions of these rings (see for examples [1, 6, 7, 12, 13]). As generalizations of semiregular rings, semiperfect rings and perfect rings, the notions of $I$-semiregular rings, $I$-semiperfect rings and $I$-perfect rings were introduced by Yousif and Zhou [12]. Our purposes of this paper is to chracterize $I$-semiregular rings, $I$-semiperfect rings and $I$-perfect rings by defining $I$-supplemented modules.

Let $R$ be a ring and $I$ an ideal of $R$, $M$ a module and $S \leq M$. $S$ is called *small* in $M$ (notation $S \ll M$) if $M \neq S + T$ for any proper submodule $T$ of $M$. $M$ is said to be singular if $M = Z(M)$, where $Z(M) = \{x \in M : l_R(x) \text{ is essential in } {}_RR\}$. As a proper generalization of small submodules, the concept of $\delta$-small submodules was introduced by Zhou[13]. $N$ is said to be *$\delta$-small* in $M$ if, whenever $N + X = M$, $M/X$ singular, we have $X = M$. $\delta(M) = Rej_M(\wp) = \cap\{N \leq M \mid M/N \in \wp\}$ , where $\wp$ be the class of all singular simple modules. Let $N, L \leq M$. $N$ is called a *supplement* of $L$ in $M$ if

*Department of Applied Mathematics, Lanzhou University of Technology Lanzhou 730050, P. R. China.
Email : ydwang@lut.cn, djwu@lut.cn

$N+L=M$ and $N$ is minimal with respect to this property. Equivalently, $M=N+L$ and $N\cap L\ll N$. $M$ is called *supplemented* if every submodule of $M$ has a supplement in $M$. $M$ is said to be *lifting* if for any submodule $N$ of $M$, there exists a direct summand $K$ of $M$ such that $K\leq N$ and $N/K\ll M/K$, equivalently, for every submodule $N$ of $M$, $M$ has a decomposition with $M=M_1\oplus M_2$, $M_1\leq N$ and $M_2\cap N$ is small in $M_2$. $N$ is called a *$\delta$-supplement* [4] of $L$ if $M=N+L$ and $N\cap L\ll_\delta N$. $M$ is called a *$\delta$-supplemented module* if every submodule of $M$ has a $\delta$-supplement. $M$ is said to be *$\delta$-lifting* [4] if for any submodule $N$ of $M$, there exists a direct summand $K$ of $M$ such that $K\leq N$ and $N/K\ll_\delta M/K$, equivalently, for every submodule $N$ of $M$, $M$ has a decomposition with $M=M_1\oplus M_2$, $M_1\leq N$ and $M_2\cap N$ is $\delta$-small in $M_2$. $M$ is ($\delta$-) *semiregular* [10] if for any cyclic submodule $N$ of $M$, there is a decomposition $M=P\oplus Q$ such that $P\leq N$ and $N\cap Q$ is a ($\delta$-) small submodule of $Q$. An element $m$ of $M$ is called *$I$-semiregular* [1] if there exists a decomposition $M=P\oplus Q$ where $P$ is projective, $P\subseteq Rm$ and $Rm\cap Q\subseteq IM$. $M$ is called *an $I$-semiregular module* if every element of $M$ is $I$-semiregular. $R$ is called *$I$-semiregular* if ${}_RR$ is an $I$-semiregular module. Note that $I$-semiregular rings are left-right symmetric and $R$ is ($\delta-$) semiregular if and only if $R$ is ($\delta({}_RR)-$) $J(R)$-semiregular. $M$ is called ($\delta$)-*semiperfect* [7] if every factor module of $M$ has a projective ($\delta$-)cover. $M$ is called *an $I$-semiperfect module* [7] if for every submodule $K$ of $M$, there is a decomposition $M=A\oplus B$ such that $A$ is projective, $A\subseteq K$ and $K\cap B\subseteq IM$. $R$ is called *$I$-semiperfect* if ${}_RR$ is an $I$-semiperfect module. Note that $R$ is ($\delta-$) semiperfect if and only if $R$ is ($\delta({}_RR)-$) $J(R)$-semiperfect. $R$ is called *a left $I$-perfect ring* [12] if, for any submodule $X$ of a projective module $P$, $X$ has a decomposition $X=A\oplus B$ where $A$ is a direct summand of $P$ and $B\subseteq IP$. By [7, Proposition 2.1], $R$ is a left $I$-perfect ring if and only if every projective module is an $I$-semiperfect module. For other standard definitions we refer to [2, 3, 11].

In this note all rings are associative with identity and all modules are unital left modules unless specified otherwise. Let $R$ be a ring and $M$ a module. We use $Rad(M)$, $Soc(M)$, $Z(M)$ to indicate the Jacobson radical, the socle, the singular submodule of $M$ respectively. $J(R)$ is the radical of $R$ and $I$ is an ideal of $R$.

## 2. PSD submodules and $I$-supplemented modules

In this section, we give some properties of PSD submodules and use PSD submodules to define (finitely) $I$-supplemented modules and $I$-lifting modules which are generalizations of some well-known supplemented modules and lifting modules. Some properties of $I$-supplemented modules are discussed. We begin this section with the following definitions.

**2.1. Definition.** Let $I$ be an ideal of $R$ and $N\leq M$. $N$ is PSD in $M$ if there exists a projective summand $S$ of $M$ such that $S\leq N$ and $M=S\oplus X$ whenever $N+X=M$ for any submodule $X\leq M$. $M$ is PSD for $I$ if any submodule of $IM$ is PSD in $M$. $R$ is a left PSD ring for $I$ if any finitely generated free left $R$-module is PSD for $I$.

**2.2. Lemma.** Let $M$ and $N$ be modules.

(1) If $K$ is PSD in $M$ and $f:M\to N$ is an epimorphism, then $f(K)$ is PSD in $N$.
(2) If $L\leq N\leq M$ and $L$ is PSD in $N$, then $L$ is PSD in $M$.
(3) If $L\leq N\leq M$ and $N$ is PSD in $M$, then $L$ is PSD in $M$.
(4) Let $M=M_1\oplus M_2$. If $N_1$ is PSD in $M_1$ and $N_2$ is PSD in $M_2$, then $N_1\oplus N_2$ is PSD in $M$.
(5) Let $N$ be a direct summand of $M$ and $A\leq N$. Then $A$ is PSD in $M$ if and only if $A$ is PSD in $N$.

*Proof.* (1) Let $f(K) + L = N$ with $L \leq N$. Then $K + f^{-1}(L) = M$. Since $K$ is PSD in $M$, there is a projective summand $H$ of $M$ with $H \leq K$ such that $H \oplus f^{-1}(L) = M$. So $f(H) \oplus L = N, f(H) \subseteq f(K)$. It is easy to see that $f(H) \cong H$ is projective.

(2) Let $M = L + X$ with $X \leq M$. Then $N = L + (N \cap X)$. Since $L$ is PSD in $N$, there is a projective summand $H$ of $N$ with $H \leq L$ such that $N = H \oplus (N \cap X)$, and hence $L = H \oplus (L \cap X)$. So $M = H \oplus X$.

(3) Let $M = L + K$ with $K \leq M$, then $M = N + K$. Since $N$ is PSD in $M$, there is a projective summand $H$ of $M$ with $H \leq N$ such that $M = H \oplus K$, and hence $M/K \cong H$ is projective. Thus the natural epimorphism $f : L \to M/K$ splits and $Kerf = L \cap K$ is a direct summand of $L$. Write $L = (L \cap K) \oplus Q$ with $Q \leq L$, we have $M = Q \oplus K$. The rest is obvious.

(4) Let $M = N_1 \oplus N_2 + L$ with $L \leq M$. Since $N_1$ is PSD in $M_1$, $N_1$ is PSD in $M$. Thus there is a projective summand $S_1$ of $M$ with $S_1 \subseteq N_1$ such that $M = S_1 \oplus (N_2 + L)$. Similarly, there exists a projective summand $S_2$ of $M$ with $S_2 \subseteq N_2$ such that $M = S_1 \oplus S_2 \oplus L$. The rest is obvious.

(5) " $\Rightarrow$ " Since $N$ is a direct summand of $M$, $M = N \oplus K$ for some submodule $K \leq M$. Suppose that $N = A + X$ with $X \leq N$, then $M = A + (X \oplus K)$. Since $A$ is PSD in $M$, there is a projective direct summand $Y$ of $M$ such that $Y \leq A$ and $M = Y \oplus X \oplus K$, and hence $N = N \cap M = X \oplus Y$.

" $\Leftarrow$ " Let $M = A + L$ with $L \leq M$. Then $N = N \cap M = A + N \cap L$. Since $A$ is PSD in $N$, there is a projective summand $K$ of $N$ with $K \leq A$ such that $N = K \oplus (N \cap L)$. It is easy to see that $K \cap L = 0$. Next we only show that $M = K + L$. Let $m \in M$, then $m = a + l$, $a \in A, l \in L$. Since $a = k + s, k \in K, s \in N \cap L$, $m = k + s + l$. Note that $s + l \in L$, so $m \in K + L$, and hence $M = K + L$, as required.

□

**2.3. Proposition.** Let $M$ be a module and $N \leq M$.

(1) $N \ll M$ if and only if $N \subseteq Rad(M)$, $N$ is PSD in $M$.
(2) $N \ll_\delta M$ if and only if $N \subseteq \delta(M)$, $N$ is PSD in $M$.

*Proof.* (1) " $\Rightarrow$ " is clear.

" $\Leftarrow$ " Let $M = N + L$ with $L \leq M$. Since $N$ is PSD in $M$, there is a projective summand $H$ of $M$ with $H \subseteq N \subseteq Rad(M)$ such that $M = H \oplus L$. So $Rad(H) \oplus Rad(L) = Rad(M) = H \oplus Rad(L)$. Thus $Rad(H) = H$. Since $H$ is projective, $H = 0$, and hence $L = M$.

(2) " $\Rightarrow$ " is clear by [13, Lemma 1.2].

" $\Leftarrow$ " Let $M = N + L$ with $L \leq M$. Since $N$ is PSD in $M$, there is a projective summand $H$ of $M$ with $H \subseteq N \subseteq \delta(M)$ such that $M = H \oplus L$. So $\delta(H) \oplus \delta(L) = \delta(M) = H \oplus \delta(L)$. Thus $\delta(H) = H$. Since $H$ is projective, $H$ is semisimple by [7, Proposition 2.13]. Thus $N \ll_\delta M$ by [13, Lemma 1.2].

□

**2.4. Corollary.** Let $M$ be a module. Then

(1) $M$ is ($\delta$-) supplemented if and only if for every submodule $X$ of $M$, there is a submodule $Y$ of $M$ such that $X + Y = M$, $X \cap Y \subseteq (\delta(Y))$ $Rad(Y)$ and $X \cap Y$ is PSD in $Y$.
(2) $M$ is ($\delta$-) lifting if and only if for every submodule $X$ of $M$, there is a decomposition $M = A \oplus B$ such that $A \subseteq X$ and $X \cap B \subseteq (\delta(B))$ $Rad(B)$ and $X \cap B$ is PSD in $B$.

**2.5. Definition.** Let $R$ be a ring and $I$ an ideal of $R$, $M$ a module. $M$ is called an $I$-supplemented module (finitely $I$-supplemented module) if for every submodule (finitely generated submodule ) $X$ of $M$, there is a submodule $Y$ of $M$ such that $X+Y=M$, $X\cap Y\subseteq IY$ and $X\cap Y$ is PSD in $Y$. In this case, we call $Y$ is an $I$-supplement of $X$ in $M$. $M$ is called $I$-lifting if for every submodule $X$ of $M$, there is a decomposition $M=A\oplus B$ such that $A\subseteq X$ and $X\cap B\subseteq IB$ and $X\cap B$ is PSD in $B$.

**2.6. Example.** It is easy to see that a module $M$ is 0-supplemented if and only if $M$ is semisimple, and so the supplemented module $\mathbb{Z}(p^\infty)$ is not 0-supplemented, where $p$ is a prime integer. However, $\mathbb{Z}(p^\infty)$ is $I$-supplemented for every nonzero ideal $I$ of $\mathbb{Z}$.

**2.7. Theorem.** Consider the following statements for a module $M$.

(1) $M$ is a $J(R)$-supplemented module (a $\delta(_RR)$-supplemented module, respectively).
(2) $M$ is a supplemented module ( a $\delta$-supplemented module, respectively).

Then "(1) $\Rightarrow$ (2)", "(2) $\Rightarrow$ (1)" if $M$ is projective or $R$ satisfies $J(R)M=Rad(M)$ $(\delta(_RR)M=\delta(M))$ for any module $M$ over $R$.

*Proof.* "(1) $\Rightarrow$ (2)" By Proposition 2.3.

"(2) $\Rightarrow$ (1)" Let $M$ be a supplemented module. Then for every submodule $X$ of $M$, there is a submodule $Y$ of $M$ such that $X+Y=M$ and $X\cap Y\ll Y$. Since $M$ is projective, $Y$ is a direct summand of $M$, and hence $Y$ is projective. It is clear that $X\cap Y\subseteq Rad(Y)=J(R)Y$ and $X\cap Y$ is PSD in $Y$. (Let $M$ be a $\delta$-supplemented module. Since $M$ is projective, $M$ is $\delta$-lifting. Thus for every submodule $X$ of $M$, there is a direct summand $Y$ of $M$ such that $M=X+Y$ and $X\cap Y\ll_\delta Y$. The rest is obvious.) When $R$ satisfies $J(R)M=Rad(M)$ $(\delta(_RR)M=\delta(M))$ for any module $M$ over $R$, the proof is similar.

□

Similar to the proof of Theorem 2.7, we have the following.

**2.8. Theorem.** Consider the following statements for a module $M$.

(1) $M$ is a finitely $J(R)$-supplemented module (a finitely $\delta(_RR)$-supplemented module, respectively).
(2) $M$ is a finitely supplemented module ( a finitely $\delta$-supplemented module, respectively).

Then "(1) $\Rightarrow$ (2)", "(2) $\Rightarrow$ (1)" if $M$ is projective or $R$ satisfies $J(R)M=Rad(M)$ $(\delta(_RR)M=\delta(M))$ for any module $M$ over $R$.

**2.9. Theorem.** Consider the following statements for a module $M$.

(1) $M$ is a $J(R)$-lifting module (a $\delta(_RR)$-lifting module, respectively).
(2) $M$ is a lifting module ( a $\delta$-lifting module, respectively).

Then "(1) $\Rightarrow$ (2)", "(2) $\Rightarrow$ (1)" if $M$ is projective or $R$ satisfies $J(R)M=Rad(M)$ $(\delta(_RR)M=\delta(M))$ for any module $M$ over $R$.

We know that if a ring $R$ is left $(\delta-)$semiperfect ring, then $(\delta(_RR)M=\delta(M))$ $J(R)M=Rad(M)$ for any module $M$ over $R$. So "(1) $\Leftrightarrow$ (2)" in Theorem 2.7, 2.8 and 2.9 if $R$ is left $(\delta-)$semiperfect ring.

**2.10. Lemma.** Let $M$ be a module and $K,L,H\le M$. If $K$ is an $I$-supplement of $L$ in $M$, $L$ is an $I$-supplement of $H$ in $M$, then $L$ is an $I$-supplement of $K$ in $M$.

*Proof.* Let $M=K+L=L+H$ with $K\cap L\subseteq IK, L\cap H\subseteq IL$ and $K\cap L$ be PSD in $K$, $L\cap H$ be PSD in $L$. We only show that $K\cap L\subseteq IL$ and $K\cap L$ is PSD in $L$. It is easy to see that $K\cap L\subseteq IK\cap L$. Let $l=\Sigma_{i=1}^n p_ik_i\in IK\cap L$, $p_i\in I, k_i\in K$ and $k_i=l_i'+h_i(i=1,2,\cdots,n)$, $l_i'\in L, h_i\in H$. Since $L\cap H\subseteq IL$, $l\in IL$, and hence

$K \cap L \subseteq IL$. Next, we shall prove that $K \cap L$ is PSD in $L$. Let $K \cap L + X = L$ with $X \leq L$, then $M = L + H = K \cap L + X + H$. Since $K \cap L$ is PSD in $K$, $K \cap L$ is PSD in $M$ by Lemma 2.2. Thus there is a projective summand $Y$ of $M$ with $Y \subseteq K \cap L$ such that $M = Y \oplus (X + H)$. Since $L = L \cap M = L \cap (Y \oplus (X + H)) = Y \oplus (X + L \cap H)$ and $L \cap H$ is PSD in $L$, there is a projective summand $Y'$ of $L$ with $Y' \subseteq L \cap H$ such that $L = Y \oplus X \oplus Y'$. Since $L/X \cong Y \oplus Y'$ is projective, the natural epimorphism $f : K \cap L \to L/X$ splits, and hence $Kerf = K \cap X$ is a direct summand of $K \cap L$. Write $K \cap L = (K \cap X) \oplus Q, Q \leq K \cap L$. So $L = Q \oplus X$, as required.

□

**2.11. Lemma.** Let $M$ be a $\pi$-projective module. If $N$ and $K$ are $I$-supplement of each other in $M$, then $N \cap K$ is projective. If in addition $M$ is projective, then $N$ and $K$ are projective.

*Proof.* Let $f : N \oplus K \to N + K = M$ with $(n, k) \mapsto n + k$ for $n \in N, k \in K$. Since $M$ is a $\pi$-projective module, $f$ splits, and so $Kerf = \{(n, -n) | n \in N \cap K\}$ is a direct summand of $N \oplus K$. Write $N \oplus K = Kerf \oplus U, U \cong M$. Since $N \cap K$ is PSD in $N$ and $K$, $Kerf$ is PSD in $N \oplus K$ by Lemma 2.2. Thus there is a projective summand $Y$ of $N \oplus K$ with $Y \subseteq Kerf$ such that $N \oplus K = Y \oplus U$, so $Y = Kerf \cong N \cap K$ is projective. If $M$ is projective, $Y \oplus U$ is projective. So $N$ and $K$ are projective. □

Recall that a pair $(P, f)$ is called a projective $I$-cover of $M$ [9] if $P$ is projective, $f$ is an epimorphism from $P$ to $M$ such that $Kerf \leq IP$, and $Kerf$ is PSD in $P$.

We end this section with the following lemma.

**2.12. Lemma.** Let $M = A + B$. If $M/A$ has a projective $I$-cover, then $B$ contains an $I$-supplement of $A$.

*Proof.* Let $\pi : B \to M/A$ be the canonical homomorphism and $f : P \to M/A$ a projective $I$-cover. Since $P$ is projective, there is a homomorphism $g : P \to B$ such that $\pi g = f$. Thus $M = A + g(P)$ and $A \cap g(P) = g(Kerf)$. Since $Kerf \subseteq IP$ and $Kerf$ is PSD in $P$, $A \cap g(P) \subseteq Ig(P)$ and $A \cap g(P)$ is PSD in $g(P)$ by Lemma 2.2. So $g(P)$ is an $I$-supplement of $A$ contained in $B$. □

## 3. Characterizations of $I$-semiregular, $I$-semiperfect and $I$-perfect rings in terms of $I$-supplemented modules

We shall characterize $I$-semiregular rings, $I$-semiperfect rings and $I$-perfect rings by $I$-supplemented modules in this section. We begin this section with the following.

**3.1. Theorem.** Let $R$ be a ring and $I$ an ideal of $R$, $P$ a projective module. Consider the following conditions:

(1) $P$ is an $I$-supplemented module.
(2) $P$ is an $I$-semiperfect module.

Then (1) $\Rightarrow$ (2), and (2) $\Rightarrow$ (1) if $P$ is PSD for $I$.

*Proof.* "(1) $\Rightarrow$ (2)" Let $P$ be an $I$-supplemented module and $N \leq P$. Then there exists $X \leq P$ such that $P = N + X$, $N \cap X \subseteq IX$ and $N \cap X$ is PSD in $X$. Let $\pi : P \to P/N$ and $\pi |_X : X \to P/N$ be the canonical epimorphisms. Since $P$ is projective, there is a homomorphism $g : P \to X$ such that $\pi |_X \, g = \pi$. We have $P = g(P) + N$ and $X = g(P) + N \cap X$. Since $N \cap X$ is PSD in $X$, there is a projective summand $Y$ of $X$ with $Y \subseteq N \cap X$ such that $X = g(P) \oplus Y$. It is easy to verify that $g(P) \cap N \subseteq Ig(P)$. Since $g(P) \cap N \subseteq N \cap X$ and $N \cap X$ is PSD in $X$, $g(P) \cap N$ is PSD in $X$ by Lemma 2.2, and so $g(P) \cap N$ is PSD in $g(P)$ by Lemma 2.2. Thus $g(P)$ is an $I$-supplement of $N$ in

$P$. Since $P$ is an $I$-supplemented module, $g(P)$ has an $I$-supplement $Q$ in $P$. Thus $g(P)$ is also an $I$-supplement of $Q$ in $P$ by Lemma 2.10, and so $g(P)$ is projective by Lemma 2.11. Since $g(P) \cap N \subseteq Ig(P)$ and $g(P) \cap N$ is PSD in $g(P)$, the canonical epimorphism $g(P) \to P/N$ is a projective $I$-cover of $P/N$. So $P$ is an $I$-semiperfect module by [9, Lemma 2.9].

"(2) $\Rightarrow$ (1)" Let $P$ be an $I$-semiperfect module, then for every submodule $X$ of $P$, there is a decomposition $P = A \oplus Y$ such that $A$ is projective, $A \subseteq X$ and $X \cap Y \subseteq IP$. Thus $P = X + Y$, $X \cap Y \subseteq IY$. Since $P$ is PSD for $I$, $X \cap Y$ is PSD in $Y$ by Lemma 2.2, as desired.

□

By Theorem 3.1, we know that if a module $M$ is projective and PSD for $I$, then $M$ is an $I$-supplemented module if and only if $M$ is $I$-lifting if and only if $M$ is an $I$-semiperfect module.

**3.2. Corollary.** Let $M$ be a projective module with $Rad(M) \ll M$ ($\delta(M) \ll_\delta M$). Then $M$ is a ($\delta$-)supplemented module if and only if $M$ is a ($\delta$-)semiperfect module if and only if $M$ is a ($\delta$-)lifting module.

**3.3. Theorem.** Let $I$ be an ideal of $R$. Consider the following conditions:

(1) Every finitely generated $R$-module is $I$-supplemented.
(2) Every finitely generated projective $R$-module is $I$-supplemented.
(3) Every finitely generated projective $R$-module is $I$-lifting.
(4) ${}_RR$ is $I$-lifting.
(5) ${}_RR$ is $I$-supplemented.
(6) $R$ is $I$-semiperfect.

Then (1) $\Rightarrow$ (2) $\Rightarrow$ (5) $\Rightarrow$ (6) and (3) $\Rightarrow$ (4) $\Rightarrow$ (5) hold; if $R$ is a left PSD ring for $I$, (2) $\Rightarrow$ (3) and (6) $\Rightarrow$ (1) also hold.

*Proof.* "(1) $\Rightarrow$ (2) $\Rightarrow$ (5)" and "(3) $\Rightarrow$ (4) $\Rightarrow$ (5)" are clear.

"(5) $\Rightarrow$ (6)" By Theorem 3.1.

If $R$ is a left PSD ring for $I$, then (2) $\Rightarrow$ (3) is obvious by Theorem 3.1 and [9, Corollary 2.4].

"(6) $\Rightarrow$ (1)" Let $M$ be a finitely generated module and $N \leq M$. Then $M = N + M$ and $M/N$ has a projective $I$-cover by [9, Theorem 2.13], so $M$ contains an $I$-supplement of $N$ by Lemma 2.12. Hence $M$ is $I$-supplemented. □

Let $I = J(R)$ or $\delta({}_RR)$ in Theorem 3.3, since $R$ is a left PSD ring for $I$ and $R$ is ($\delta-$) semiperfect if and only if ($\delta({}_RR)-$)$J(R)$-semiperfect , we use Theorem 2.7 and Theorem 2.9 to obtain the following.

**3.4. Corollary.** ([5, Theorem 4.41]) The following statements are equivalent for a ring $R$.

(1) $R$ is semiperfect.
(2) Every finitely generated $R$-module is supplemented.
(3) Every finitely generated projective $R$-module is supplemented.
(4) Every finitely generated projective $R$-module is lifting.
(5) ${}_RR$ is lifting.
(6) ${}_RR$ is supplemented.

**3.5. Corollary.** ([4, Theorem 3.3]) The following statements are equivalent for a ring $R$.

(1) $R$ is $\delta$-semiperfect.

(2) Every finitely generated $R$-module is $\delta$-supplemented.
(3) Every finitely generated projective $R$-module is $\delta$-supplemented.
(4) Every finitely generated projective $R$-module is $\delta$-lifting.
(5) ${}_RR$ is $\delta$-lifting.
(6) ${}_RR$ is $\delta$-supplemented.

Since if $R$ is $Z({}_RR)$-semiregular, then $Z({}_RR) = J(R) \subseteq \delta({}_RR)$ by [1, Theorem 3.2], and so $R$ is a left PSD ring for $Z({}_RR)$. Thus we have the following result.

**3.6. Corollary.** The following statements are equivalent for a ring $R$.

(1) $R$ is $Z({}_RR)$-semiperfect.
(2) Every finitely generated $R$-module is $Z({}_RR)$-supplemented.
(3) Every finitely generated projective $R$-module is $Z({}_RR)$-supplemented.
(4) Every finitely generated projective $R$-module is $Z({}_RR)$-lifting.
(5) ${}_RR$ is $Z({}_RR)$-lifting.
(6) ${}_RR$ is $Z({}_RR)$-supplemented.

Let $M$ be a projective module, then $Soc(M) = Soc({}_RR)M$. So if $I \leq Soc({}_RR)$, then $IM \subseteq Soc(M)$, and hence $R$ is a left PSD ring for $I$. Thus we have

**3.7. Corollary.** The following statements are equivalent for a ring $R$.

(1) $R$ is $Soc({}_RR)$-semiperfect.
(2) Every finitely generated $R$-module is $Soc({}_RR)$-supplemented.
(3) Every finitely generated projective $R$-module is $Soc({}_RR)$-supplemented.
(4) Every finitely generated projective $R$-module is $Soc({}_RR)$-lifting.
(5) ${}_RR$ is $Soc({}_RR)$-lifting.
(6) ${}_RR$ is $Soc({}_RR)$-supplemented.

**3.8. Theorem.** Let $R$ be a left PSD ring and $I$ an ideal of $R$. Then $R$ is an $I$-semiregular ring if and only if ${}_RR$ is a finitely $I$-supplemented module if and only if $R_R$ is a finitely $I$-supplemented module.

*Proof.* Similar to Theorem 3.3. □

**3.9. Corollary.** ([8, Proposition 19.1]) The following statements are equivalent for a ring $R$.

(1) $R$ is semiregular.
(2) ${}_RR$ is a finitely supplemented module.
(3) $R_R$ is a finitely supplemented module.

**3.10. Corollary.** The following statements are equivalent for a ring $R$.

(1) $R$ is $\delta$-semiregular.
(2) ${}_RR$ is a finitely $\delta$-supplemented module.
(3) $R_R$ is a finitely $\delta$-supplemented module.

**3.11. Corollary.** A ring $R$ is $Soc({}_RR)$-semiregular if and only if ${}_RR$ is a finitely $Soc({}_RR)$-supplemented module if and only if $R_R$ is a finitely $Soc(R_R)$-supplemented module.

**3.12. Corollary.** A ring $R$ is $Z({}_RR)$-semiregular if and only if ${}_RR$ is a finitely $Z({}_RR)$-supplemented module if and only if $R_R$ is a finitely $Z(R_R)$-supplemented module.

Next we use $I$-supplemented modules to characterize $I$-perfect rings.

**3.13. Definition.** A ring $R$ is called a strongly left PSD ring for $I$ if any projective left $R$-module is PSD for $I$.

**3.14. Example.** It is easy to verify that a ring $R$ is perfect if and only if $R$ is a semiperfect ring and a strongly PSD ring for $J(R)$. Let $R = \mathbb{Z}_{(p)}$ (integers localized at the prime $p$). It is well known that $R$ is a commutative, semiperfect ring that is not perfect, and so $R$ is a PSD ring for $J(R)$ that is not a strongly PSD ring for $J(R)$ (Since if $R$ is a strongly PSD ring for $J(R)$, then $R$ is perfect. This is a contradiction.).

**3.15. Theorem.** Let $I$ be an ideal of $R$. Consider the following conditions:

(1) Every $R$-module is $I$-supplemented.
(2) Every projective $R$-module is $I$-supplemented.
(3) Every projective $R$-module is $I$-lifting.
(4) Every free $R$-module is $I$-lifting.
(5) Every free $R$-module is $I$-supplemented.
(6) $R$ is left $I$-perfect.

Then (1) $\Rightarrow$ (2) $\Rightarrow$ (5) $\Rightarrow$ (6) and (3) $\Rightarrow$ (4) $\Rightarrow$ (5) hold; if $R$ is a strongly left PSD ring for $I$, (2) $\Rightarrow$ (3) and (6) $\Rightarrow$ (1) also hold.

*Proof.* "(1) $\Rightarrow$ (2) $\Rightarrow$ (5)" and "(3) $\Rightarrow$ (4) $\Rightarrow$ (5)" are clear.

"(5) $\Rightarrow$ (6)" By Theorem 3.1.

When $R$ is a strongly left PSD ring for $I$, "(2) $\Rightarrow$ (3)" is obvious.

"(6) $\Rightarrow$ (1)" Let $M$ be a module. Then there is a free module $F$ such that $\eta : F \to M$ is epic. Since $F$ is $I$-semiperfect, there is a decomposition $F = F_1 \oplus F_2$ such that $F_1 \subseteq Ker\eta$ and $F_2 \cap Ker\eta \subseteq IF_2$. Since $F$ is PSD for $I$, $F_2 \cap Ker\eta$ is PSD in $F$. By Lemma 2.2, $F_2 \cap Ker\eta$ is PSD in $F_2$, so $\eta|_{F_2} : F_2 \to M$ is a projective $I$-cover of $M$. Thus we prove that an arbitrary module has a projective $I$-cover, and so for $N \leq M$, $M/N$ has a projective $I$-cover. The rest follows by Lemma 2.12. □

Let $I = J(R)$ or $\delta({}_RR)$ in Theorem 3.15. Since $R$ is ($\delta-$)perfect if and only if $R$ is ($\delta({}_RR)-$) $J(R)$-perfect and if a ring $R$ is ($\delta-$)perfect, then for every module $M$, ($\delta(M) \ll_\delta M$) $Rad(M) \ll M$, $R$ is a strongly left PSD ring for $I$. So we have the following.

**3.16. Corollary.** ([5, Theorem 4.41])The following statements are equivalent for a ring $R$.

(1) $R$ is left perfect.
(2) Every $R$-module is supplemented.
(3) Every projective $R$-module is supplemented.
(4) Every projective $R$-module is lifting.
(5) Every free $R$-module is lifting.
(6) Every free $R$-module is supplemented.

**3.17. Corollary.** ([4, Theorem 3.4]) The following statements are equivalent for a ring $R$.

(1) $R$ is left $\delta$-perfect.
(2) Every $R$-module is $\delta$-supplemented.
(3) Every projective $R$-module is $\delta$-supplemented.
(4) Every projective $R$-module is $\delta$-lifting.
(5) Every free $R$-module is $\delta$-lifting.
(6) Every free $R$-module is $\delta$-supplemented.

Since if $I \leq Soc({}_RR)$, then $R$ is a strongly left PSD ring for $I$, and hence we have

**3.18. Corollary.** The following statements are equivalent for a ring $R$.

(1) $R$ is left $Soc({}_RR)$-perfect.

(2) Every $R$-module is $Soc({}_RR)$-supplemented.
(3) Every projective $R$-module is $Soc({}_RR)$-supplemented.
(4) Every projective $R$-module is $Soc({}_RR)$-lifting.
(5) Every free $R$-module is $Soc({}_RR)$-lifting.
(6) Every free $R$-module is $Soc({}_RR)$-supplemented.

Since if $R$ is $Z({}_RR)$-perfect, then $Z({}_RR) = J(R) \subseteq \delta({}_RR)$ by [1, Theorem 3.2], and so we have the following result.

**3.19. Corollary.** The following statements are equivalent for a ring $R$.

(1) $R$ is left $Z({}_RR)$-perfect.
(2) Every $R$-module is $Z({}_RR)$-supplemented.
(3) Every projective $R$-module is $Z({}_RR)$-supplemented.
(4) Every projective $R$-module is $Z({}_RR)$-lifting.
(5) Every free $R$-module is $Z({}_RR)$-lifting.
(6) Every free $R$-module is $Z({}_RR)$-supplemented.

**ACKNOWLEDGMENTS**

The authors wish to thank the referee for very careful reading of the manuscript and also for his/her suggestions. The first author is supported by the Natural Science Foundation of Gansu Province (No. 1310RJZA029) and the fundamental research funds for the universities in Gansu Province. The second author is supported by National Natural Science Foundation of China (No. 11301242).

## References

[1] M. Alkan, A. C. Ozcan, Semiregular modules with respect to a fully invariant submodule, Comm. Algebra 32(11) (2004) 4285-4301.
[2] F. W. Anderson and K. R. Fuller, Rings and Categories of Modules, Springer-Verlag, New York, 1974.
[3] J. Clark, C. Lomp, N. Vanaja and R. Wisbauer, Lifting modules, IN:Supplements and Projectivity in Module Theory Sereis: Frontiers in Mathematics, 2006.
[4] M. T. Kosan, $\delta$-lifting and $\delta$-supplemented modules, Algebra Colloq. 14 (2007) 53-60.
[5] S. H. Mohamed and B. J. Müller, Continuous and Discrete Modules, London Math. Soc.; LNS147, Cambridge Univ. Press: Cambridge, 1990.
[6] W. K. Nicholson, Strong lifting, J. Algebra 285 (2005) 795-818.
[7] A. C. Ozcan, M. Alkan, Semiperfect modules with respect to a preradical, Comm. Algebra 34 (2006) 841-856.
[8] A. Tuganbaev, Rings Close to Regular, Kluwer Academic Publishers, Dordrecht, 2002.
[9] Y. D. Wang, Characterizations of $I$-semiregular and $I$-semiperfect rings, submitted.
[10] Y. D. Wang, Relatively semiregular modules, Algebra Colloq. 13 (2006) 647-654.
[11] R. Wisbauer, Foundations of Module and Ring Theory, Gordon and Breach, Philadelphia, 1991.
[12] M. F. Yousif, Y. Q. Zhou, Semiregular,semiperfect and perfect rings relative to an ideal, Rocky Mountain J. Math. 32 (2002) 1651-1671.
[13] Y. Q. Zhou, Generalizations of perfect, semiperfect, and semiregular rings, Algebra Colloq. 3 (2000) 305-318.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 139 – 151

# Approximations and adjoints for categories of complexes of Gorenstein projective modules

Gang Yang* and Wei Ren†

**Abstract**

In the paper, it is proven that every object in C($R$-GProj) has a special C($R$-Proj)-preenvelope, and then some adjoints in homotopy categories related to Gorenstein projective modules are given, where C($R$-Proj) is the subcategory of complexes of projective $R$-modules, and C($R$-GProj) is the subcategory of complexes of Gorenstein projective $R$-modules.

**Keywords:** (special) preenvelopes, Gorenstein projective modules, DG-projective complexes, adjoints.

*2000 AMS Classification:* 16E05, 18G15, 18G35.

*Received* : 29.05.2014 *Accepted* : 28.12.2014 *Doi* : 10.15672/HJMS.20164512477

## 1. Introduction

Let $R$ be an associative ring, and let $R$-Proj, $R$-Flat, and $R$-GProj be the subcategory of projective, flat, and Gorenstein projective $R$-modules in $R$-Mod, the category of left $R$-modules. If $\mathcal{A}$ is one of the above categories then we use C($\mathcal{A}$) to denote the category of complexes of $R$-modules in $\mathcal{A}$. The category K($\mathcal{A}$) is the homotopy category which has the same objects as C($\mathcal{A}$), and the morphisms are homotopy equivalence classes of morphisms of complexes. It was shown in [15] and [16] that both the inclusions K($R$-Proj) $\to$ K($R$-Flat) and K($R$-Flat) $\to$ K($R$-Mod) have right adjoints. Recently, Diego Bravo, Edgar E. Enochs et. al in [5] showed that some adjoints to inclusion functors may exist if they were given complete cotorsion pairs in the category of complexes. The paper is motivated by the above work to show:

**1.1. Theorem.** Let $R$ be any ring. Then every complex $G \in$ C($R$-GProj) has a special C($R$-Proj)-preenvelope.

*Department of Mathematics, Lanzhou Jiaotong University, Lanzhou 730070, P. R. China. Email : yanggang@mail.lzjtu.cn

†Department of Mathematics, Northwest Normal University, Lanzhou 730070, P. R. China Email : rwemail@163.com

Now suppose that $R$ is quasi-Frobenius. It is well known that the subcategory $R$-GProj is in fact $R$-Mod. Then the categories C($R$-GProj) and C($R$-Mod) are the same, and so Theorem 1.1 says that any complex admits a special C($R$-Proj)-preenvelope. It is natural to ask whether every complex adimits a special DG-projective preenvelope since the class of DG-projective complexes is contained in C($R$-Proj)? We find that the answer is negative in general. In fact, every complex adimits a special DG-projective preenvelope if and only if $R$ has global dimension 0.

Note that the inclusion K($R$-Proj) $\to$ K($R$-Mod) always has a right adjoint ([5, Theorem 4.7]). We are inspired to consider whether there exists a left adjoint to it, and we show the following main result which is based on Theorem 1.1.

**1.2. Theorem.** Let $R$ be any ring. Then the inclusion K($R$-Proj) $\to$ K($R$-GProj) has a left adjoint.

## 2. Preliminaries

Let $\Omega$ be a subcategory of an abelian category $\mathcal{A}$, and $M$ is an object of $\mathcal{A}$. A morphism $f : M \to Q$ is called an $\Omega$-preenvelope of $M$, if $Q \in \Omega$ and the sequence $\mathrm{Hom}(Q, Q') \to \mathrm{Hom}(M, Q') \to 0$ is exact for any $Q' \in \Omega$. If moreover, $g \circ f = f$ implies that $g$ is an automorphism whenever $g \in \mathrm{End}(Q)$, then $f$ is called an $\Omega$-envelope. An $\Omega$-preenvelope $f : M \to Q$ of $M$ is said to be special, if $f$ is injective and $\mathrm{Ext}^1(\mathrm{Coker}(f), Q') = 0$ for any $Q' \in \Omega$. An $\Omega$-precover, an $\Omega$-cover and a special $\Omega$-precover $Q \to M$ are defined dually. See [9, 11] for detail.

Auslander and Reiten [2] and Auslander and Smal$\phi$ [3] use the terminology left and right approximations and minimal left and right approximations for preenvolpes, precovers, envelopes and covers.

A complex $X$ of $R$-modules is a sequence $\cdots \to X_{i+1} \xrightarrow{\delta_{i+1}^X} X_i \xrightarrow{\delta_i^X} X_{i-1} \to \cdots$ of $R$-modules and $R$-homomorphisms such that $\delta_i^X \delta_{i+1}^X = 0$ for all $i \in \mathbb{Z}$. A complex $X$ is said to be acyclic (exact) if $\mathrm{Im}(\delta_{i+1}^X) = \mathrm{Ker}(\delta_i^X)$ for all $i \in \mathbb{Z}$. A complex $X$ is said to be bounded above if $X_i = 0$ holds for $i \gg 0$, bounded below if $X_i = 0$ holds for $i \ll 0$, and bounded if it is bounded above and below, i.e. $X_i = 0$ holds for $|i| \gg 0$. Let $X$ be a complex and let $m$ be an integer. The $m$-fold shift of $X$ is the complex $\Sigma^m X$ given by $(\Sigma^m X)_i = X_{i-m}$ and $\delta_i^{\Sigma^m X} = (-1)^m \delta_{i-m}^X$. Usually, $\Sigma^1 X$ is denoted simply by $\Sigma X$.

Let $X$ and $Y$ be two complexes. We will let $\mathrm{Hom}_R(X, Y)$ denote the complex of $\mathbb{Z}$-modules with $m$th component $\mathrm{Hom}_R(X, Y)_m = \prod_{i \in \mathbb{Z}} \mathrm{Hom}_R(X_i, Y_{i+m})$ and differential $(\delta(g))_i = \delta_{i+m}^Y g_i - (-1)^m g_{i-1} \delta_i^X$ for $g = (g_i)_{i \in \mathbb{Z}} \in \prod_{i \in \mathbb{Z}} \mathrm{Hom}_R(X_i, Y_{i+m})$. By a morphism $f : X \to Y$ we mean a sequence $f_i : X_i \to Y_i$ such that $\delta_i^Y f_i = f_{i-1} \delta_i^X$ for all $i \in \mathbb{Z}$. The mapping cone $\mathrm{Cone}(f)$ of a morphism $f : X \to Y$ is defined as $\mathrm{Cone}(f)_i = Y_i \oplus X_{i-1}$ with $\delta_i^{\mathrm{Cone}(f)} = \begin{pmatrix} \delta_i^Y & f_{i-1} \\ 0 & -\delta_{i-1}^X \end{pmatrix}$.

If $M$ is an $R$-module then we denote the complex $\cdots \to 0 \to M \to 0 \to \cdots$ with $M$ in the $m$th degree by $S^m(M)$, and denote the complex $\cdots \to 0 \to M \xrightarrow{Id} M \to 0 \to \cdots$ with $M$ in the $m-1$ and $m$th degrees by $D^m(M)$. Usually, $S^0(M)$ is denoted simply by $M$. We use $\mathrm{Hom}(X, Y)$ to present the group of all morphisms from $X$ to $Y$. Recall that a complex $P$ is projective if the functor $\mathrm{Hom}(P, -)$ is exact. Equivalently, $P$ is projective if and only if $P$ is acyclic and $\mathrm{Im}(P_{i+1} \to P_i)$ is a projective $R$-module for each $i \in \mathbb{Z}$. For example, if $M$ is a projective $R$-module then each complex $D^m(M)$ is projective. A injective complex is defined dually. Thus C($R$-Mod), the category of complexes of

$R$-modules, has enough projectives and injectives, we can compute right derived functors $\mathrm{Ext}^i(X,Y)$ of $\mathrm{Hom}(-,-)$.

**2.1. Definition.** ([8]) We call an acyclic complex $P = \cdots \to P_2 \to P_1 \to P_0 \to P_{-1} \to P_{-2} \to \cdots$ with all $P_i$ projective a complete projective resolution of an $R$-module $M$, if $M \cong \mathrm{Ker}(P_0 \to P_{-1})$, and $\mathrm{Hom}_R(P,N)$ is acyclic for any projective $R$-module $N$. An $R$-module $M$ is called Gorenstein projective, if there exists a complete projective resolution of $M$.

The dual notions are those of a complete injective resolution and a Gorenstein injective $R$-module.

**2.2. Remark.** (1) The subcategory $R$-GProj is projectively resolving, that is, $R$-GProj contains $R$-Proj, and $F \in R$-GProj if and only if $H \in R$-GProj for any exact sequence $0 \to F \to H \to G \to 0$ with $G \in R$-GProj ([12, Theorem 2.5]).

(2) An $R$-module $M \in R$-GProj with finite projective dimension is projective ([12, Proposition 2.7]).

## 3. The existence of $\mathrm{C}(R$-Proj)-preenvelopes

In this section, we focus on $\mathrm{C}(R$-Proj)-preenvelopes of special complexes over general associative rings. We begin with the following

**3.1. Lemma.** Assume that the following diagram of complexes with exact rows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle\mu} & & \downarrow{\scriptstyle\nu} & & \downarrow{\scriptstyle\omega} & & \\ 0 & \longrightarrow & X & \xrightarrow{p} & Y & \xrightarrow{q} & Z & \longrightarrow & 0 \end{array}$$

is commutative. Then the sequence

$$0 \longrightarrow Cone(\mu) \xrightarrow{\begin{pmatrix} p & 0 \\ 0 & \Sigma f \end{pmatrix}} Cone(\nu) \xrightarrow{\begin{pmatrix} q & 0 \\ 0 & \Sigma g \end{pmatrix}} Cone(\omega) \longrightarrow 0$$

is exact.

*Proof.* It can be checked by standard computation. □

**3.2. Lemma.** Let $G \in \mathrm{C}(R$-GProj) be acyclic and bounded above. If $\mathrm{Hom}_R(G,A)$ is acyclic for any projective $R$-module $A$ then $\mathrm{Ext}^1(G,P) = 0$ for any $P \in \mathrm{C}(R$-Proj).

*Proof.* See [13, Lemma 3.1]. □

**3.3. Definition.** Let $X$ be a complex and let $m$ be an integer. The hard truncation above of $X$ at $m$, denoted $X_{\leqslant m}$, is the complex

$$X_{\leqslant m} = \ 0 \to X_m \xrightarrow{\delta^X_m} X_{m-1} \xrightarrow{\delta^X_{m-1}} X_{m-2} \to \cdots .$$

Similarly, the hard truncation below of $X$ at $m$, denoted $X_{\geqslant m}$, is the complex

$$X_{\geqslant m} = \ \cdots \to X_{m+2} \xrightarrow{\delta^X_{m+2}} X_{m+1} \xrightarrow{\delta^X_{m+1}} X_m \to 0.$$

**3.4. Lemma.** Let $G \in \mathrm{C}(R$-GProj) be bounded above. Then there is an exact sequence $0 \to G \to P \to C \to 0$ such that $P \in \mathrm{C}(R$-Proj) and $C \in \mathrm{C}(R$-GProj) are both bounded above, $C$ is acyclic, and $\mathrm{Hom}_R(C,A)$ is acyclic for any projective $R$-module $A$.

*Proof.* Assume without loss of generality that $G =: \ 0 \to G_0 \to G_{-1} \to G_{-2} \to \cdots$ is a complex of Gorenstein projective $R$-modules with $G_0$ in the 0th degree. If for each $n \geqslant 0$, we let $G(n) = G_{\geqslant -n}$, the hard truncation below of $G$ at $-n$, then $\{(G(n), \alpha_{mn}) | m \geqslant n \geqslant 0\}$ forms a inverse system in $C(R\text{-Mod})$ and $G = \varprojlim G(n)$, where $\alpha_{mn} : G(m) \to G(n)$ is a natural projection for any $m \geqslant n$.

We will show by induction on $n$. For $n = 0$, since $G_0$ is Gorenstein projective, there exists an exact sequence $0 \to G_0 \xrightarrow{f} P_0 \to P_{-1} \to P_{-2} \to \cdots$ with each $P_i$ projective and it remains exact after applying the functor $\mathrm{Hom}_R(-, A)$ for any projective $R$-module $A$. Let $P(0) =: \ 0 \to P_0 \to P_{-1} \to P_{-2} \to \cdots$, and consider the following monomorphism of complexes $\phi(0) : G(0) \to P(0)$.

$$
\begin{array}{ccccccccccc}
G(0) & & 0 & \longrightarrow & G_0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \cdots \\
\downarrow & & & & \downarrow{\scriptstyle f} & & \downarrow & & \downarrow & & \\
P(0) & & 0 & \longrightarrow & P_0 & \longrightarrow & P_{-1} & \longrightarrow & P_{-2} & \longrightarrow & \cdots
\end{array}
$$

Let $C(0) = \mathrm{Coker}(\phi(0))$, that is $C(0) =: \ 0 \to G'_0 \to P_{-1} \to P_{-2} \to \cdots$ with $G'_0 = \mathrm{Coker}(f)$. Clearly, $P(0) \in \mathrm{C}(R\text{-Proj})$ and $C(0) \in \mathrm{C}(R\text{-GProj})$ are both bounded above, $C(0)$ is acyclic, and $\mathrm{Hom}_R(C(0), A)$ is acyclic for any projective $R$-module $A$.

Now for $n \geqslant 0$, suppose that there is a monomorphism $\phi(n) : G(n) \to P(n)$ as follows.

$$
\begin{array}{ccccccccccccccc}
G(n) & & 0 & \longrightarrow & G_0 & \longrightarrow & G_{-1} & \longrightarrow & \cdots & \longrightarrow & G_{-n} & \longrightarrow & 0 & \longrightarrow & \cdots \\
\downarrow & & & & \downarrow{\scriptstyle f_0} & & \downarrow{\scriptstyle f_{-1}} & & & & \downarrow{\scriptstyle f_{-n}} & & \downarrow & & \\
P(n) & & 0 & \longrightarrow & P_0 & \xrightarrow{\delta_0} & P_{-1} & \xrightarrow{\delta_{-1}} & \cdots & \longrightarrow & P_{-n} & \xrightarrow{\delta_{-n}} & P_{-n-1} & \longrightarrow & \cdots
\end{array}
$$

Where $P(n) \in \mathrm{C}(R\text{-Proj})$ and $C(n) = \mathrm{Coker}(\phi(n)) \in \mathrm{C}(R\text{-GProj})$ are both bounded above, $C(n)$ is acyclic, and $\mathrm{Hom}_R(C(n), A)$ is acyclic for any projective $R$-module $A$. Let $G(n+1) =: \ 0 \to G_0 \xrightarrow{d_0} G_{-1} \xrightarrow{d_{-1}} \cdots \to G_{-n} \xrightarrow{d_{-n}} G_{-n-1} \to 0 \to \cdots$, and let $0 \to G_{-n-1} \xrightarrow{g} Q_{-n-1} \to Q_{-n-2} \to Q_{-n-3} \to \cdots$ be an exact sequence with each $Q_i$ projective and it remains exact after applying the functor $\mathrm{Hom}_R(-, A)$ for any projective $R$-module $A$. We denoted by $Q$ the complex $0 \to Q_{-n-1} \to Q_{-n-2} \to Q_{-n-3} \to \cdots$ with $Q_{-n-1}$ in the $(-n-1)$th degree. By the above proof, we have a monomorphism $\iota : S^{-n-1}(G_{-n-1}) \to Q$ such that $\mathrm{Coker}(\iota) \in \mathrm{C}(R\text{-GProj})$ is acyclic and bounded above, and also $\mathrm{Hom}_R(\mathrm{Coker}(\iota), A)$ is acyclic for any projective $R$-module $A$.

Let $\mu : \Sigma^{-1} G(n) \to S^{-n-1}(G_{-n-1})$ be the following morphism

$$
\begin{array}{ccccccccccccccc}
\Sigma^{-1}G(n) & & 0 & \longrightarrow & G_0 & \xrightarrow{-d_0} & \cdots & \longrightarrow & G_{-n+1} & \longrightarrow & G_{-n} & \longrightarrow & 0 & \longrightarrow & \cdots \\
\downarrow & & & & \downarrow & & & & \downarrow & & \downarrow{\scriptstyle d_{-n}} & & \downarrow & & \\
S^{-n-1}(G_{-n-1}) & & 0 & \longrightarrow & 0 & \longrightarrow & \cdots & \longrightarrow & 0 & \longrightarrow & G_{-n-1} & \longrightarrow & 0 & \longrightarrow & \cdots .
\end{array}
$$

Note that the sequence $0 \to \Sigma^{-1} G(n) \to \Sigma^{-1} P(n) \to \Sigma^{-1} C(n) \to 0$ is exact. Then it follows from Lemma 3.2 that the sequence

$$0 \to \mathrm{Hom}(\Sigma^{-1}C(n), Q) \to \mathrm{Hom}(\Sigma^{-1}P(n), Q) \to \mathrm{Hom}(\Sigma^{-1}G(n), Q) \to \mathrm{Ext}^1(\Sigma^{-1}C(n), Q) = 0$$

is exact, and so there exits a morphism $\nu : \Sigma^{-1}P(n) \longrightarrow Q$ such that the following diagram commutes.

$$\begin{array}{ccc} \Sigma^{-1}G(n) & \xrightarrow{\Sigma^{-1}\phi(n)} & \Sigma^{-1}P(n) \\ \downarrow \mu & & \downarrow \nu \\ S^{-n-1}(G_{-n-1}) & \xrightarrow{\iota} & Q \end{array}$$

Thus there exists a morphism $\omega : \Sigma^{-1}C(n) \to \mathrm{Coker}(\iota)$ such that the following diagram with exact rows commutes.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \Sigma^{-1}G(n) & \xrightarrow{\Sigma^{-1}\phi(n)} & \Sigma^{-1}P(n) & \longrightarrow & \Sigma^{-1}C(n) & \longrightarrow & 0 \\ & & \downarrow \mu & & \downarrow \nu & & \downarrow \omega & & \\ 0 & \longrightarrow & S^{-n-1}(G_{-n-1}) & \xrightarrow{\iota} & Q & \longrightarrow & \mathrm{Coker}(\iota) & \longrightarrow & 0 \end{array}$$

By lemma 3.1, the sequence

$$0 \longrightarrow Cone(\mu) \longrightarrow Cone(\nu) \longrightarrow Cone(\omega) \longrightarrow 0$$

is exact. Note that $G(n+1) = Cone(\mu)$. If we put $P(n+1) = Cone(\nu)$ and $C(n+1) = Cone(\omega)$ then we have an exact sequence

$$0 \longrightarrow G(n+1) \xrightarrow{\phi(n+1)} P(n+1) \longrightarrow C(n+1) \longrightarrow 0\ .$$

On one hand, exactness of the sequence

$$0 \longrightarrow Q \longrightarrow P(n+1) \longrightarrow P(n) \longrightarrow 0$$

implies that $P(n+1) \in \mathrm{C}(R\text{-Proj})$ is bounded above since $P(n) \in \mathrm{C}(R\text{-Proj})$ and $Q \in \mathrm{C}(R\text{-Proj})$ are so, and $P(n+1)_{-k} = P(n)_{-k}$ for $0 \leqslant k \leqslant n$. On the other hand, exactness of the sequence

$$0 \longrightarrow \mathrm{Coker}(\iota) \longrightarrow C(n+1) \longrightarrow C(n) \longrightarrow 0$$

implies that $C(n+1) \in \mathrm{C}(R\text{-GProj})$ is bounded above and acyclic with $\mathrm{Hom}_R(C(n+1), A)$ acyclic for any projective $R$-module $A$ since $\mathrm{Coker}(\iota)$ and $C(n)$ are so. Clearly, one has $C(n+1)_{-k} = C(n)_{-k}$ for $0 \leqslant k \leqslant n$.

Note that every morphism $G(n+1) \to G(n)$ is surjective. By [9, Theorem 1.5.13], the sequence

$$0 \longrightarrow G = \varprojlim G(n) \xrightarrow{\varprojlim \phi(n)} \varprojlim P(n) \longrightarrow \varprojlim C(n) \longrightarrow 0$$

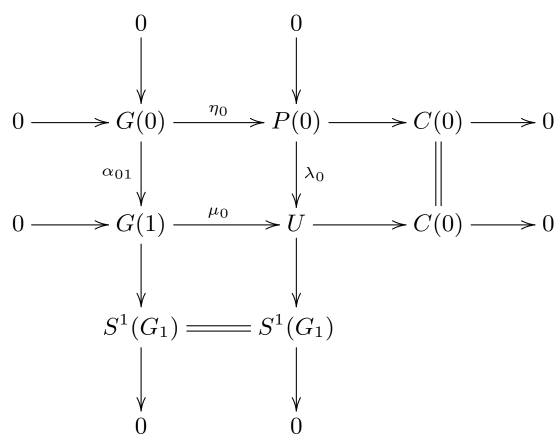
is exact. Let $P = \varprojlim P(n)$, and $C = \varprojlim C(n)$. Then $P_{-k} = \varprojlim P(n)_{-k} = P(k)_{-k}$ for any $k \geqslant 0$ and $P_{-k} = 0$ for any $k \leqslant -1$, $C_{-k} = \varprojlim C(n)_{-k} = C(k)_{-k}$ for any $k \geqslant 0$ and $C_{-k} = 0$ for any $k \leqslant -1$. Thus one can check easily that $P \in \mathrm{C}(R\text{-Proj})$ and $C \in \mathrm{C}(R\text{-GProj})$ are bounded above, $C$ is acyclic, and also $\mathrm{Hom}_R(C, A)$ is acyclic for any projective $R$-module $A$. □

Now we give the following main result which contains Theorem 1.1.

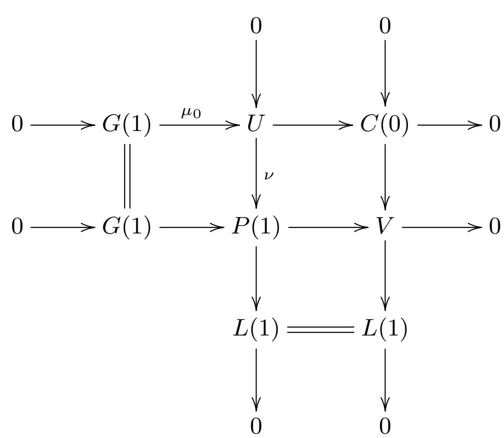
**3.5. Theorem.** Every complex $G \in \mathrm{C}(R\text{-GProj})$ has a special $\mathrm{C}(R\text{-Proj})$-preenvelope $\eta : G \to P$ with $\mathrm{Coker}(\eta) \in \mathrm{C}(R\text{-GProj})$ acyclic.

*Proof.* If we write $G(n) = G_{\leqslant n}$ for each $n \geqslant 0$ then we get that $((G(n)), (\alpha_{mn}))_{n\geqslant 0}$ is a direct system in $C(R\text{-Mod})$ and $\varinjlim G(n) = G$, where $\alpha_{mn} : G(m) \to G(n)$ is a natural injection for any $m \leqslant n$.

By Lemma 3.4, there exists an exact sequence $0 \to G(0) \xrightarrow{\eta_0} P(0) \to C(0) \to 0$ such that $P(0) \in \mathrm{C}(R\text{-Proj})$ and $C(0) \in \mathrm{C}(R\text{-GProj})$ are both bounded above, $C(0)$ is acyclic, and $\mathrm{Hom}_R(C(0), A)$ is acyclic for any projective $R$-module $A$. It follows from lemma 3.2 that $\mathrm{Ext}^1(C(0), Q) = 0$ for any $Q \in \mathrm{C}(R\text{-Proj})$. Thus the monomorphism $\eta_0 : G(0) \to P(0)$ is a special $\mathrm{C}(R\text{-Proj})$-preenvelope of $G(0)$. Consider the push-out diagram of morphisms $\eta_0 : G(0) \to P(0)$ and $\alpha_{01} : G(0) \to G(1)$




Clearly, $U \in \mathrm{C}(R\text{-GProj})$ is bounded above since $P(0)$ and $S^1(G_1)$ are so. By Lemma 3.4 again, we get that there exists an exact sequence $0 \to U \xrightarrow{\nu} P(1) \to L(1) \to 0$ such that $P(1) \in \mathrm{C}(R\text{-Proj})$ and $L(1) \in \mathrm{C}(R\text{-GProj})$ are both bounded above, and $L(1)$ and $\mathrm{Hom}_R(L(1), A)$ are acyclic for any projective $R$-module $A$. Consider the push-out diagram of morphisms $U \to C(0)$ and $\nu : U \to P(1)$




The exactness of the rightmost column implies that $V \in \mathrm{C}(R\text{-GProj})$ is bounded above, $V$ is acyclic, and $\mathrm{Hom}_R(V, A)$ is acyclic for any projective $R$-module $A$. It follows from Lemma 3.2 that the monomorphism $\eta_1 = \nu\mu_0 : G(1) \to P(1)$ is a special $\mathrm{C}(R\text{-Proj})$-preenvelope of $G(1)$. Let $C(1) = V$, and $\beta_{01} = \nu\lambda_0$. Therefore we get, by the construction above, a commutative diagram with exact rows and columns.

$$
\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & G(0) & \xrightarrow{\eta_0} & P(0) & \longrightarrow & C(0) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \alpha_{01}} & & \downarrow{\scriptstyle \beta_{01}} & & \downarrow{\scriptstyle \gamma_{01}} & & \\
0 & \longrightarrow & G(1) & \xrightarrow{\eta_1} & P(1) & \longrightarrow & C(1) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & S^1(G_1) & \longrightarrow & N(1) & \longrightarrow & L(1) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}
$$

Since it is easily seen from the lower row of the above diagram that $N(1) \in \mathrm{C}(R\text{-GProj})$, and the middle column that $N(1)_i$ has finitely projective dimension for each $i \in \mathbb{Z}$, we get by Remark 2.2 that $N(1) \in \mathrm{C}(R\text{-Proj})$.

If we continue this process, then we get a commutative diagram with exact rows as follows

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & G(0) & \xrightarrow{\eta_0} & P(0) & \longrightarrow & C(0) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \alpha_{01}} & & \downarrow{\scriptstyle \beta_{01}} & & \downarrow{\scriptstyle \gamma_{01}} & & \\
0 & \longrightarrow & G(1) & \xrightarrow{\eta_1} & P(1) & \longrightarrow & C(1) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \alpha_{12}} & & \downarrow{\scriptstyle \beta_{12}} & & \downarrow{\scriptstyle \gamma_{12}} & & \\
0 & \longrightarrow & G(2) & \xrightarrow{\eta_2} & P(2) & \longrightarrow & C(2) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \alpha_{23}} & & \downarrow{\scriptstyle \beta_{23}} & & \downarrow{\scriptstyle \gamma_{23}} & & \\
 & & \vdots & & \vdots & & \vdots & &
\end{array}
$$

where each row $0 \longrightarrow G(n) \xrightarrow{\eta_n} P(n) \longrightarrow C(n) \longrightarrow 0$ satisfies that $P(n) \in \mathrm{C}(R\text{-Proj})$ and $C(n) \in \mathrm{C}(R\text{-GProj})$ are both bounded above, $C(n)$ is acyclic, and $\mathrm{Hom}_R(C(n), A)$ is acyclic for any projective $R$-module $A$. In particular, by Lemma 3.2, the monomorphism $\eta_n : G(n) \to P(n)$ is a special $\mathrm{C}(R\text{-Proj})$-preenvelope of $G(n)$ for each $n \geqslant 0$. Also each row $0 \longrightarrow G(n) \xrightarrow{\eta_n} P(n) \longrightarrow C(n) \longrightarrow 0$ has the property that the following diagram with exact rows and columns is commutative.

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & G(n) & \xrightarrow{\eta_n} & P(n) & \longrightarrow & C(n) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \alpha_{n,n+1}} & & \downarrow{\scriptstyle \beta_{n,n+1}} & & \downarrow{\scriptstyle \gamma_{n,n+1}} & & \\
0 & \longrightarrow & G(n+1) & \xrightarrow{\eta_{n+1}} & P(n+1) & \longrightarrow & C(n+1) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & S^{n+1}(G_{n+1}) & \longrightarrow & N(n+1) & \longrightarrow & L(n+1) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

Where $N(n+1) \in \mathrm{C}(R\text{-Proj})$ and $L(n+1) \in \mathrm{C}(R\text{-GProj})$ are both bounded above, $L(n+1)$ is acyclic, $\mathrm{Hom}_R(L(n+1), A)$ is acyclic for any projective $R$-module $A$ and for each $n \geqslant 0$. By Lemma 3.2, one has $\mathrm{Ext}^1(L(n+1), Q) = 0$ for any $Q \in \mathrm{C}(R\text{-Proj})$ and for each $n \geqslant 0$. Clearly, $((P(n)), (\beta_{mn}))_{n\geqslant 0}$ forms a continuous direct systems of monomorphisms in $\mathrm{C}(R\text{-Proj})$ such that $\mathrm{Coker}(\beta_{n,n+1}) = N(n+1) \in \mathrm{C}(R\text{-Proj})$, and we have $\varinjlim P(n) \in \mathrm{C}(R\text{-Proj})$ since $\mathrm{C}(R\text{-Proj})$ is closed under direct transfinite extension. Again since $((C(n)), (\gamma_{mn}))_{n\geqslant 0}$ forms a continuous direct systems of monomorphisms in $\mathrm{C}(R\text{-GProj})$ such that $\mathrm{Coker}(\gamma_{n,n+1}) = L(n+1) \in \mathrm{C}(R\text{-GProj})$, we get that $\varinjlim C(n) \in \mathrm{C}(R\text{-GProj})$ since $R$-GProj is closed under direct transfinite extension [7, Theorem 3.2]. Note that each $C(n)$ is acyclic and the class of acyclic complexes is a left side of a cotorsion pair [10], we get that $\varinjlim C(n)$ is acyclic by [6, Theorem 1.2]. In fact, the monomorphism $\eta : \varinjlim G(n) \to \varinjlim P(n)$, $\eta = \varinjlim \eta_n$, is a special $\mathrm{C}(R\text{-Proj})$-preenvelope of $\varinjlim G(n) = G$. To show this we need only to prove $\mathrm{Ext}^1(\varinjlim C(n), Q) = 0$ for any $Q \in \mathrm{C}(R\text{-Proj})$, but the latter is easily seen by [6, Theorem 1.5] and by the above construction. This completes the proof. □

**3.6. Remark.** The above special $\mathrm{C}(R\text{-Proj})$-preenvelope $\eta : G \to P$ of $G$ is a homology isomorphism since $\eta$ is monomorphic and $\mathrm{Coker}(\eta)$ is acyclic.

Recall from [4] that a complex $P$ is called DG-projective if each $P_i$ is projective and if $\mathrm{Hom}_R(P, E)$ is an acyclic complex of abelian groups for any acyclic complex $E$. Let $R$ be a quasi-Frobenius ring, that is, An $R$-module $M$ is projective if and only if it is injective. Then it is easily seen by Theorem 3.5 that every complex of left $R$-modules has a special $\mathrm{C}(R\text{-Proj})$ preenvelope since every left $R$-module is Gorenstein projective, so it is natural to ask whether every complex of left $R$-modules has a special DG-projective preenvelope, and we have the following result.

**3.7. Proposition.** Let $R$ be a quasi-Frobenius ring. Then every complex of $R$-modules has a special DG-projective preenvelope of and only if $l.gl.dim(R) = 0$.

*Proof.* For the necessity. Suppose $l.gl.dim(R) > 0$ and let $M$ be a non-projective $R$-module. If $S^0(M) \to P$ is a special DG-projective preenvelope (which is injective), then there is an induced morphism $S^0(M) \to P_{\leqslant 0}$. Since the sequence $0 \to P_{\leqslant 0} \to P \to P_{\geqslant 1} \to 0$ is exact and $P_{\geqslant 1}$ and $P$ are DG-projective, it follows that the subcomplex $P_{\leqslant 0}$ is DG-projective. Thus one can check easily that $S^0(M) \to P_{\leqslant 0}$ is a DG-projective preenvelope of $S^0(M)$. In fact, let $K_0 = \mathrm{Coker}(M \to P_0)$. Then $\mathrm{Ext}^1(X, T) = 0$ for any DG-projective complex $T$ since $S^0(M) \to P$ is a special DG-projective preenvelope, where $X = \operatorname{Coker}(S^0(M) \to P) =: \cdots \to P_2 \to P_1 \xrightarrow{0} K_0 \to P_{-1} \to P_{-2} \to \cdots$. But it is easily seen that $K =: 0 \to K_0 \to P_{-1} \to P_{-2} \to \cdots$ is a direct summand of $X$, and so $\operatorname{Ext}^1(K, T) = 0$ for any DG-projective complex $T$. This shows that $S^0(M) \to P_{\leqslant 0}$ is a special DG-projective preenvelope of $S^0(M)$.

Let $0 \to M \to Q_0 \to Q_{-1} \to Q_{-2} \to \cdots$ be a right minimal projective (injective) resolution of $M$, that is to say, $M \to Q_0$ and each $L_{-i+1} \to Q_{-i}$ are projective envelopes of $M$ and $L_{-i+1}$ for $i > 0$, respectively, where $L_0 = \operatorname{Coker}(M \to Q_0)$, and $L_{-i} = \operatorname{Coker}(L_{-i+1} \to Q_{-i})$. Denote the complex $0 \to L_0 \to Q_{-1} \to Q_{-2} \to \cdots$ by $L$ with $L_0$ in the 0th degree. Then we have a morphism $S^0(M) \to S^0(Q_0)$ with $S^0(Q_0)$ DG-projective. Thus there is a commutative diagram

$$
\begin{array}{ccc}
S^0(M) & \longrightarrow & P_{\leqslant 0} \\
\| & & \downarrow \\
S^0(M) & \longrightarrow & S^0(Q_0)
\end{array}
$$

In particular, its commutative square frame in the 0th degree implies that there exists a morphism of $R$-modules $K_0 \to L_0$ such that the following diagram with the bottom row exact is commutative.

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & M & \longrightarrow & P_0 & \longrightarrow & K_0 & \longrightarrow & 0 \\
 & & \| & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & M & \longrightarrow & Q_0 & \longrightarrow & L_0 & \longrightarrow & 0
\end{array}
$$

Now consider the diagram

$$
\begin{array}{cccccccccccc}
K & & 0 & \longrightarrow & K_0 & \longrightarrow & P_{-1} & \longrightarrow & P_{-2} & \longrightarrow & \cdots \\
\downarrow & & & & \downarrow & & \downarrow & & \downarrow & & \\
S^0(L_0) & & 0 & \longrightarrow & L_0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \cdots \\
\uparrow & & & & \| & & \uparrow & & \uparrow & & \\
L_{\geqslant -1} & & 0 & \longrightarrow & L_0 & \longrightarrow & Q_{-1} & \longrightarrow & 0 & \longrightarrow & \cdots
\end{array}
$$

Since the subcomplex $0 \to Q_{-1} \to 0$ of $0 \to L_0 \to Q_{-1} \to 0$ is DG-projective and since $\operatorname{Ext}^1(K, T) = 0$ for any DG-projective complex $T$, we can lift the morphism $K \to S^0(L_0)$ to a morphism $K \to L_{\geqslant -1}$. Then consider the morphism $K \to L_{\geqslant -1}$ and the exact sequence $0 \to S^{-2}(Q_{-2}) \to L_{\geqslant -2} \to L_{\geqslant -1} \to 0$, for the same reason, we can lift the morphism $K \to L_{\geqslant -1}$ to a morphism $K \to L_{\geqslant -2}$. Repeating the procedure, we see that there is a commutative diagram

$$
\begin{array}{cccccccccccccc}
K & & 0 & \longrightarrow & K_0 & \longrightarrow & P_{-1} & \longrightarrow & P_{-2} & \longrightarrow & P_{-3} & \longrightarrow & \cdots \\
\downarrow & & & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
L & & 0 & \longrightarrow & L_0 & \longrightarrow & Q_{-1} & \longrightarrow & Q_{-2} & \longrightarrow & Q_{-3} & \longrightarrow & \cdots
\end{array}
$$

On the other hand, since $0 \to M \to Q_0 \to Q_{-1} \to Q_{-2} \to \cdots$ is a right minimal projective (injective) resolution of $M$, there exist morphisms $Q_i \to P_i$ such that the diagram

$$\begin{array}{ccccccccccccc}
0 & \longrightarrow & M & \longrightarrow & Q_0 & \longrightarrow & Q_{-1} & \longrightarrow & Q_{-2} & \longrightarrow & Q_{-3} & \longrightarrow & \cdots \\
 & & \Vert & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & M & \longrightarrow & P_0 & \longrightarrow & P_{-1} & \longrightarrow & P_{-2} & \longrightarrow & P_{-3} & \longrightarrow & \cdots
\end{array}$$

is commutative, this induces a commutative diagram

$$\begin{array}{cccccccccccc}
L & \quad 0 & \longrightarrow & L_0 & \longrightarrow & Q_{-1} & \longrightarrow & Q_{-2} & \longrightarrow & Q_{-3} & \longrightarrow & \cdots . \\
\downarrow & & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
K & \quad 0 & \longrightarrow & K_0 & \longrightarrow & P_{-1} & \longrightarrow & P_{-2} & \longrightarrow & P_{-3} & \longrightarrow & \cdots
\end{array}$$

But $0 \to L_0 \to Q_{-1} \to Q_{-2} \to \cdots$ is a minimal projective resolution of $L_0$, so one can check easily that $L$ is isomorphic to a direct summand of $K$, and so $0 \to Q_{-1} \to Q_{-2} \to Q_{-3} \to \cdots$ is a direct summand of $P_{\leqslant -1}$. It follows that $P_{\leqslant -1}$ is DG-projective since $S^0(P_0)$ and $P_{\leqslant 0}$ in the exact sequence $0 \to P_{\leqslant -1} \to P_{\leqslant 0} \to S^0(P_0) \to 0$ are so, hence $0 \to Q_{-1} \to Q_{-2} \to Q_{-3} \to \cdots$ is DG-projective and of course then $0 \to Q_0 \to Q_{-1} \to Q_{-2} \to \cdots$ is DG-projective. Now assembling the (left) projective resolution $\cdots \to Q_2 \to Q_1 \to M \to 0$ and the complex $0 \to M \to Q_0 \to Q_{-1} \to Q_{-2} \to \cdots$, one gets an exact sequence $0 \to Q_{\leqslant 0} \to Q \to Q_{\geqslant 1} \to 0$ with $Q_{\leqslant 0}$ and $Q_{\geqslant 1}$ DG-projective, where $Q =: \cdots \to Q_2 \to Q_1 \to Q_0 \to Q_{-1} \to Q_{-2} \to \cdots$. Thus $Q$ is clearly DG-projective. But this complex is acyclic, and so it is a projective complex by [10, Proposition 3.7]. This contradicts to the fact that $M$ is a non-projective $R$-module. Hence $l.gl.dim(R) = 0$.

The sufficiency is trivial. □

## 4. Adjoints to inclusion functors

We have mentioned in the introduction that the inclusion $\mathrm{K}(R\text{-Proj}) \to \mathrm{K}(R\text{-Mod})$ always has a right adjoint ([5, Theorem 4.7]). We are inspired to consider whether there exists a left adjoint to it in this section.

**4.1. Definition.** Let $\mathcal{D}$ be a triangulated category, and let $\mathcal{C}$ be a full subcategory of $\mathcal{D}$. The subcategory is said to be thick if it is a triangulated subcategory, and if every direct summand of any object of $\mathcal{C}$ is in $\mathcal{C}$.

The following result is dual to [16, Proposition 1.4], we give its proof for completeness.

**4.2. Proposition.** Let $\mathcal{T}$ be a triangulated category, and $\mathcal{S}$ a thick subcategory of $\mathcal{T}$. Assume further that

(1) Every object $T \in \mathcal{T}$ admits an $\mathcal{S}$-preenvelope.
(2) Every idempotent in $\mathcal{T}$ splits.

Then the inclusion $\rho : \mathcal{S} \to \mathcal{T}$ has a left adjoint.

*Proof.* Let $T$ be an object in $\mathcal{T}$. In the following we will show that there exists a morphism $g : T \to S$ with $S \in \mathcal{S}$ such that every other morphism $T \to \overline{S}$ with $\overline{S} \in \mathcal{S}$ must factor uniquely through $g$. Firstly, we choose an $\mathcal{S}$-preenvelope $f : T \to \widetilde{S}$ which must exist by hypothesis, every morphism $T \to \overline{S}$, $\overline{S} \in \mathcal{S}$ clearly factors through $f$, but not necessarily uniquely. We will show next that we can choose a direct summand $S$ of $\widetilde{S}$ for which the factorization is unique.

Complete $f : T \to \widetilde{S}$ to a triangle $T \xrightarrow{f} \widetilde{S} \xrightarrow{a} X \longrightarrow \Sigma T$ and then choose an $\mathcal{S}$-preenvelope $b : X \to S'$. Again complete $ba : \widetilde{S} \to S'$ to a triangle

$S'' \xrightarrow{c} \widetilde{S} \xrightarrow{ba} S' \longrightarrow \Sigma S''$ and then we get a morphism of triangles:

$$\begin{array}{ccccccc} T & \xrightarrow{f} & \widetilde{S} & \xrightarrow{a} & X & \longrightarrow & \Sigma T \\ {\scriptstyle d}\downarrow & & {\scriptstyle 1}\downarrow & & {\scriptstyle b}\downarrow & & \downarrow{\scriptstyle \Sigma d} \\ S'' & \xrightarrow{c} & \widetilde{S} & \xrightarrow{ba} & S' & \longrightarrow & \Sigma S'' \end{array} \qquad (*)$$

We get that $S'' \in \mathcal{S}$ since $\widetilde{S}$ and $S'$ are in the thick subcategory $\mathcal{S}$. Since $f$ is an $\mathcal{S}$-preenvelope, the morphism $d : T \to S''$ can be factored as $d = \widetilde{c}f$ with $\widetilde{c}$ a morphism $\widetilde{S} \to S''$. Now let $e = c\widetilde{c} : \widetilde{S} \to \widetilde{S}$ be the composite $\widetilde{S} \xrightarrow{\widetilde{c}} S'' \xrightarrow{c} \widetilde{S}$. Then the diagram $(*)$ implies that $f = cd = c\widetilde{c}f = ef$. To obtain the desired summand $S$ of $\widetilde{S}$, we need the following more steps.

- If the composite $T \xrightarrow{f} \widetilde{S} \xrightarrow{\rho} \overline{S}$ vanishes for some morphism $\rho : \widetilde{S} \to \overline{S}$ with $\overline{S} \in \mathcal{S}$, then so does the composite $\widetilde{S} \xrightarrow{e} \widetilde{S} \xrightarrow{\rho} \overline{S}$.

Let $\rho$ satisfy $\rho f = 0$ as above. Then we have the following morphism of triangles:

$$\begin{array}{ccccccc} T & \xrightarrow{f} & \widetilde{S} & \xrightarrow{a} & X & \longrightarrow & \Sigma T \\ {\scriptstyle 0}\downarrow & & {\scriptstyle \rho}\downarrow & & {\scriptstyle a'}\downarrow & & \downarrow{\scriptstyle 0} \\ 0 & \longrightarrow & \overline{S} & \xrightarrow{=} & \overline{S} & \longrightarrow & \Sigma 0 \end{array}$$

This shows $\rho = a'a$. Since $b : X \to S'$ is an $\mathcal{S}$-preenvelope of $X$, there exists a morphism $b' : S' \to \overline{S}$ such that $a' = b'b$. Thus by the diagram $(*)$ we get that $\rho e = a'ae = b'bae = b'bac\widetilde{c} = b'(bac)\widetilde{c} = 0$.

- Note that $f = ef$, i.e., $(1 - e)f = 0$, it follows from above that the morphism $e : \widetilde{S} \to \widetilde{S}$ is an idempotent, that is, $e^2 = e$.

By the hypothesis that any idempotent in $\mathcal{T}$ splits, the morphism $e : \widetilde{S} \to \widetilde{S}$ has a factorization $\widetilde{S} \xrightarrow{u} S \xrightarrow{v} \widetilde{S}$ with $uv$ being the identity $1_S : S \to S$. We get that $S$ must belong to $\mathcal{S}$ since it is a direct summand of $\widetilde{S}$ and the subcategory $\mathcal{S}$ is thick. Now we assert:

- Let $e = vu$ be a splitting as above, and $g : T \to S$ be the composite $T \xrightarrow{f} \widetilde{S} \xrightarrow{u} S$. Then $g$ has the property that any morphism $T \to \overline{S}$, $\overline{S} \in \mathcal{S}$ factors uniquely through $g$.

It remains to show the last assertion. Suppose that we are given a morphism $h : T \to \overline{S}$ with $\overline{S} \in \mathcal{S}$. Because $f : T \to \widetilde{S}$ is an $\mathcal{S}$-preenvelope the map $h$ must factor as $T \xrightarrow{f} \widetilde{S} \xrightarrow{\sigma} \overline{S}$. Now observe

$$h = \sigma f = \sigma ef = \sigma vuf = (\sigma v)(uf) = (\sigma v)g,$$

and we have factored $h$ through $g$. It remains to show the uniqueness. Suppose $\tau : S \to \overline{S}$ is such that the composite $\tau g = \tau uf$ vanishes. By above proof we have $\tau ue = 0$. Note that $e = vu$, we have $\tau uvu = 0$, and of cause $\tau uvuv = 0$. But $uv = 1$, we conclude that $\tau = 0$, as desired. $\square$

The categories K($R$-Proj), K($R$-GProj) and K($R$-Mod) have coproducts, hence idempotents split by [14, Proposition 1.6.8]. It is clear that K($R$-Proj) is a thick subcategory of either K($R$-GProj) or K($R$-Mod). Now we give the main result in this section.

**4.3. Theorem.** Let $R$ be any ring. Then the inclusion K($R$-Proj) $\to$ K($R$-GProj) has a left adjoint.

*Proof.* It follows from Proposition 4.2 and Theorem 3.5. □

**4.4. Corollary.** Let $R$ be any ring. Then the composition functor $J\widetilde{I} : \mathrm{K}(R\text{-GProj}) \to \mathrm{K}(R\text{-Mod})$ has a right adjoint, where $I : \mathrm{K}(R\text{-Proj}) \to \mathrm{K}(R\text{-GProj})$ and $J : \mathrm{K}(R\text{-Proj}) \to \mathrm{K}(R\text{-Mod})$ are the inclusions, and $\widetilde{I}$ is a left adjoint to $I$.

*Proof.* By [5, Theorem 4.7], the inclusion $J : \mathrm{K}(R\text{-Proj}) \to \mathrm{K}(R\text{-Mod})$ has a right adjoint $\widehat{J}$. Since we have isomorphisms for any $G \in \mathrm{K}(R\text{-GProj})$ and $M \in \mathrm{K}(R\text{-Mod})$

$$\mathrm{Hom}_{K(R\text{-Mod})}(J\widetilde{I}G, M) \cong \mathrm{Hom}_{K(R\text{-Proj})}(\widetilde{I}G, \widehat{J}M) \cong \mathrm{Hom}_{K(R\text{-GProj})}(G, I\widehat{J}M),$$

it follows that $I\widehat{J}$ is a right adjoint to $J\widetilde{I} : \mathrm{K}(R\text{-GProj}) \to \mathrm{K}(R\text{-Mod})$. □

At the end of this section we give adjoints to inclusion functors over special rings.

**4.5. Proposition.** If $R$ is left perfect and right coherent, then the inclusions of $\mathrm{K}(R\text{-Proj})$, into either of the categories $\mathrm{K}(R\text{-GProj})$ and $\mathrm{K}(R\text{-Mod})$, have left adjoints.

*Proof.* By [1, Proposition 3.5], a ring $R$ is left perfect and right coherent if and only if every left $R$-module has a projective preenvelope. Thus it follows from [17, Theorem 4.2] that every complex in $\mathrm{K}(R\text{-GProj})$ or $\mathrm{K}(R\text{-Mod})$ admits a $\mathrm{K}(R\text{-Proj})$-preenvelope since every flat $R$-module is projective under the hypothesis. Now the result follows from Proposition 4.2. □

## Acknowledgements

The authors express their gratitude to the referee for his (her) helpful comments and constructive suggestions (in particular, for the proof of Proposition 3.7), which makes the present paper more readable. This work is supported by NSF of China (Nos. 11561039, 11401476, 11301240), and NSF of Gansu Province of China (Nos. 1506RJZA075, 145RJZA079).

## References

[1] J. Asensio Mayor and J. Martinez Hernandez, On flat and projective envelopes, J. Algebra (1993) **160**: 434-440.

[2] M. Auslander and I. Reiten, Applications of contravariantly finite subcategories, Adv. Math. (1991) **86**: 111-152.

[3] M. Auslander and S. O. Smalø, Preprojective modules over Artin algebra, J. Algebra (1980) **66**: 61-122.

[4] L. L. Avramov and H.-B. Foxby, Homological dimensions of unbounded complexes, J. Pure Appl. Algebra (1991) **71**: 129-155.

[5] D. Bravo, E. E. Enochs, A. C. Iacob, O. M. G. Jenda, and J. Rada, Cotorsion pairs in C(R-Mod), Rocky Mountain J. Math. (2012) **42**(6): 1759-2103.

[6] P. C. Eklof, Homological algebra and set theory, Trans. Amer. Math. Soc. (1977) **227**: 207-225.

[7] E. E. Enochs, A. Iacob, and O. M. G. Jenda, Closure under transfinite extensions, Illinois J. Math. (2007) **51**(2): 561-569.

[8] E. E. Enochs and O. M. G. Jenda, Gorenstein injective and projective modules, Math. Zeit. (1995) **220**(4): 611-633.

[9] E. E. Enochs, O. M. G. Jenda, Relative homological algebra, in: de Gruyter Expositions in Mathematics, Vol 30, Walter de Gruyter and Co. Berlin 2000.

[10] E. E. Enochs, O. M. G. Jenda, and J. Z. Xu, Orthogonality in the category of complexes, Math. J. Okayama Univ. (1996) **38**(1): 25-46.

[11] R. Göbel, and J. Trlifaj, Approximations and endmorphism algebras of modules, Walter De Gruyter, Berlin, 2006.

[12] H. Holm, Gorenstein homological dimensions, J. Pure Appl. Algebra (2004) **189**: 167-193.

[13] L. Liang, N. Q. Ding, and G. Yang, Covers and envelopes by #-$\mathcal{F}$ complexes, Comm. Algebra (2011) **39**: 3253-3277.

[14] A. Neeman, Triangulated categories, Annals of Mathematics Syudies 148, Princeton Univ. Press, 2001.

[15] A. Neeman, The homotopy category of flat modules, and Grothendieck duality, Invent. Math. (2008) **174**: 255-308.

[16] A. Neeman, Some adjoints in homotopy categories, Annals of Mathematics (2010) **171**: 2143-2155.

[17] G. Yang, Z. K. Liu, and L. Liang, On Gorenstein flat preenelopes of complexes, Rend Sem. Mat. Univ. Padova (2013) **129**: 171-187.

# STATISTICS

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 153 – 164

# Robust model selection criteria for robust $S$ and $LTS$ estimators

Meral Çetin *

**Abstract**

Outliers and multi-collinearity often have large influence in the model/variable selection process in linear regression analysis. To investigate this combined problem of multi-collinearity and outliers, we studied and compared Liu-type S (liuS-estimators) and Liu-type Least Trimmed Squares (liuLTS) estimators as robust model selection criteria. Therefore, the main goal of this study is to select subsets of independent variables which explain dependent variables in the presence of multi-collinearity, outliers and possible departures from the normality assumption of the error distribution in regression analysis using these models.

**Keywords:** Liu-estimator, robust-Liu estimator, M-estimator, robust cp, robust Tp, robust model selection.

*2000 AMS Classification:* 62F35.

*Received* : 14.02.2014 *Accepted* : 29.12.2014 *Doi* : 10.15672/HJMS.2015609964

## 1. Introduction

Traditional variable selection methods are based on classical estimators and tests which depend on normality assumption of errors. Even though many robust alternatives to the traditional model selection methods have been offered in the past 30 years, the associated variable selection problem has been somewhat neglected. For instance, in regression analysis, Mallows's Cp (Mallows, 1973) is a powerful selection procedure. But, since the Cp statistics is based on least squares estimation, it is very sensitive to outliers and other departures from the normality assumption on the error distribution. The need for robust selection procedures is obvious, because using Mallow's Cp variable selection method cannot estimate and select parameters robustly. Ronchetti (1985) and Ronchetti et. al. (1997) proposed and investigated the properties of a robust version of Akaike's Information Criterion (AIC). Hampel (1983) suggested a modified version of it. Hurvich and Tsai(1990) compared several model selection procedures for L1 regression. Ronchetti

*Hacettepe University, Department of Statistics, Beytepe-Ankara/TURKEY
Email: meral@hacettepe.edu.tr

and Staudte (1994) proposed a robust version of Mallows's Cp. Sommer and Huggins (1996) proposed a robust Tp criterion based on Wald Statistics.

Consider the linear regression model

$$(1) \qquad \boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{Y}$ is an $n \times 1$ response vector; $\boldsymbol{X}$ is an $n \times p$ full rank matrix of predictors; $\boldsymbol{\beta}$ is an $p$ vector of unknown parameters; $\boldsymbol{\epsilon}$, is an error vector with mean 0 and variance $\sigma^2 I$. For convenience, it is assumed that the $\boldsymbol{X}$ variables are standardized so that $\boldsymbol{X}'\boldsymbol{X}$ has the form of correlation matrix.

Multi-collinearity and outliers are two main problems in regression methods. To cope with multi-collinearity, some techniques are proposed. Ridge regression estimator is one of the most widely used estimators to overcome multi-collinearity. Ridge regression estimator is defined as

$$(2) \qquad \hat{\beta}_r(k) = (X'X + kI)^{-1}X'X\hat{\beta}_{OLS}$$

where $k > 0$ is the shrinkage parameter. Since $\hat{\beta}_R(k)$ is sensitive to outliers in the y-direction, an alternative robust ridge M-estimator has been proposed by Sivapulle (1991). Since $\hat{\beta}_R(k)$ is a complicated function of $k$, Liu (1993) proposes a new biased estimator for $\beta$. Liu estimator

$$(3) \qquad \hat{\beta}_L(d) = (X'X + I)^{-1}(X'X + dI)\hat{\beta}_{OLS}$$

is obtained by shrinking the ordinary least squares (OLS) estimator using the matrix $(X'X+I)^{-1}(X'X+dI)$ Where $0 < d < 1$ is a shrinking parameter. Since OLS is used in Liu estimator, the presence of outliers in $y$ direction may affect $\hat{\beta}_L(d)$. To overcome this problem, Arslan and Billor (2000) proposed an alternative class of Liu-type M-estimators (LM) which is defined as:

$$(4) \qquad \hat{\beta}_{LM}(d) = (X'X + I)^{-1}(X'X + dI)\hat{\beta}_M$$

LM estimator is obtained by shrinking an M-estimator ($\hat{\beta}_M$) instead of the OLS estimator using the matrix $(X'X + I)^{-1}(X'X + dI)$. The main objective of this proposed estimator is to decrease the effects of the simultaneous occurrence of multicollinearity and outliers in the data set.

Let $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_p$ be the eigenvalues of $\boldsymbol{XX}'$ and $q_1, q_2, \ldots, q_p$ be the corresponding eigenvectors. Let $\boldsymbol{\Lambda} = diag(\lambda_1, \lambda_2, \ldots, \lambda_p)$ and $\boldsymbol{P} = (q_1, q_2, \ldots, q_p)$ such that $\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{P\Lambda P}'$. The regression model can be written in the canonical form by

$$\boldsymbol{Y} = \beta_0\boldsymbol{1} + \boldsymbol{C\alpha} + \epsilon$$

where $\boldsymbol{C} = \boldsymbol{XP}$ and $\boldsymbol{\alpha} = \boldsymbol{P}'\boldsymbol{\beta}$.

Then, the LM-estimator of $\alpha$, $\hat{\alpha}_{LM}(d)$, becomes

$$(5) \qquad \hat{\alpha}_{LM}(d) = (\Lambda + I)^{-1}(\Lambda + d\boldsymbol{I})\hat{\alpha_M}$$

This estimator is resistant to the combined problem of multicollinearity and outliers in the y direction (Arslan and Billor, 2000).

In order to obtain $\hat{\alpha}_{LM}(d)$ we used the robust choice of d given in equation (5). Robust $d$ value is

$$(6) \qquad \hat{d}_M = 1 - \hat{A}^2\left[\sum_{i=1}^{p}\frac{1}{\lambda_i(\lambda_i+1)} / \sum_{i=1}^{p}\frac{\hat{\alpha_{Mi}}^2}{(\lambda_i+1)^2}\right]$$

where $\hat{A}^2$ is

$$\hat{A}^2 = s^2(n-p)^{-1}\sum_{i=1}^{n}\left[\Psi(r_i/s)\right]^2 / \left[\frac{1}{n}\sum_{i=1}^{n}\left[\Psi^{'}(r_i/s)\right]\right]^2 \tag{7}$$

(Arslan and Billor, 2000).

## 2. Model Selection Estimators

### 2.1. Robust Cp Criteria.

Mallow's Cp (Mallows, 1973) is a powerful technique for model selection in regression. Since the Cp is based on OLS estimation, it is sensitive to outliers and other departures from the normality assumption on the error distribution.
Ronchetti and Staudte (1994) define a robust version of Cp as follows:

$$RC_p = \frac{W_p}{\hat{\sigma}^2} - (U_p - V_p) \tag{8}$$

where $W_p = \sum_i \hat{w}_i^2 r_i^2 = \sum_i \hat{w}_i^2 (y_i - \hat{y}_i)^2$, $w_i$ is a weight for i. th observation, and $\hat{\sigma}^2$ is a robust and consistent estimator of $\hat{\sigma}$ in the full model given by $\hat{\sigma}^2 = W_{full}/U_{full}$. $W_{full}$, is the weighted residual sum of squares for full model. The constants $U_p = \sum_i var(\hat{w}_i r_i)$ and $V_p = \sum_i var(\hat{w}_i x_i^T(\hat{\beta}_p - \beta))$ are computed assuming that the subsets are correct and $\sigma = 1$.
In robust Cp (RCP) criterion used by Ronchetti and Staudte (1994), Up-Vp value is constant for all models. In our study, the value of Up-Vp changes according to each subset. Moreover, Ronchetti and Staudte (1994) have used weighting least squares (WLS) while computing the estimates. However, we use Huber-type estimation and Huber weights, instead of WLS. So, Up-Vp is

$$U_p - V_p \sim nE\|\eta\|^2 - 2tr(\boldsymbol{N}\boldsymbol{M}^{-1}) + tr(\boldsymbol{L}\boldsymbol{M}^{-1}\boldsymbol{Q}\boldsymbol{M}^{-1}) \tag{9}$$

(see Ronchetti and Staudte,1994).
where $E\|\eta\|^2 = \sum_{1\le i\le n} \eta^2(x_i, \epsilon_i)$, $\boldsymbol{N} = E[\eta^2\eta^{'}\boldsymbol{x}\boldsymbol{x}^{\prime}]$ and $\boldsymbol{L} = E[w^{'}\epsilon(w^{'}\epsilon + 4w)\boldsymbol{x}\boldsymbol{x}^{\prime}]$.
Mallows's Cp and RCP are useful tools for model selection in regression. However, they have several disadvantages. First they are difficult to generalize to the situations where residuals are less defined. Second, they are computer intensive and their computation, particularly in robust version, can be time consuming as they require fitting of all sub-models (Sommer and Huggins, 1996). Sommer and Huggins (1996) proposed a flexible easily generalized alternative based on the Wald test (see, Wald, 1994) which requires computation of estimates only from the full model. Models, with values of RCP close to Vp or smaller than VP, will be preferred to others.

### 2.2. Robust Tp criteria.

A robust version of Tp (RTp), based on generalized M-estimators of the regression parameters, is defined by

$$RT_p = \hat{\beta}_2^{'}\Sigma_{22}^{-1}\hat{\beta}_2 - k + 2p \tag{10}$$

where $\Sigma_n$ is the covariance matrix,

$$\Sigma_n = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

and

$$\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2) = (\boldsymbol{X}^{'}\boldsymbol{V}\boldsymbol{X})^{-1}\boldsymbol{X}^{'}\boldsymbol{V}\boldsymbol{Y},$$

and $k$ and $p$ are dimensions of full model and submodel, respectively, $\hat{\beta}_1 = (\beta_0, \beta_1, \ldots, \beta_{p-1})$ and $\hat{\beta}_2 = (\beta_p, \ldots, \beta_{k-1})$ (Hampel et all,1986 ). If submodel P is correct, the value of $RT_p$ should be close to $p$ (Sommer and Huggins, 1996).

### 2.3. S and LTS estimators.

Rousseeuw and Yohai (1984) proposed S-estimator which is another high breakdown point estimator having the same asymptotic properties as the M-estimator and used in model selection in linear regression analysis. It has a higher statistical efficiency than LTS estimation even though S and LTS estimates share the same breakdown value. The S-estimator minimizes the sample S-scale of the fitted residuals, while the LTS estimator minimizes the sample root mean square error. To obtain a high breakdown point estimator, which is also $\sqrt{n}$-consistent and asymptotically normal, was the motivation for the S-estimators. The $\rho$ is considered to be a quadratic function. Let $k = E_{\Phi}[\rho]$ where $\Phi$ is the standard normal distribution. For any given sample $\{r_1, r_2, \ldots, r_n\}$ of residuals, an M-estimate of scale $\sigma(r_1, r_2, \ldots, r_n)$ is the solution to

$$ave\{\rho(r_i/\sigma)\}$$

where $ave$ denotes the arithmetic mean over $i = 1, 2, \ldots, n$. For each value of $\beta$, the dispersion of the residuals $r_i = y_i - x_i^T\beta$ can be calculated using the upper equation . Then, the S-estimator $\hat{\beta}$ of $\beta$ be defined as

$$arg \min_{\beta} \sigma(r_1(\beta), r_2(\beta), \ldots, r_n(\beta))$$

and the final scale estimate is $\hat{\sigma} = \sigma(r_1(\hat{\beta}), r_2(\hat{\beta}), \ldots, r_n(\hat{\beta}))$.
The least trimmed squares (LTS) estimate proposed by Rousseeuw (1984) is defined as the p-vector

$$\hat{\Theta}_{LTS} = arg \min_{\Theta} \boldsymbol{Q}_{LTS}(\Theta)$$

where

$$\boldsymbol{Q}_{LTS}(\Theta) = \sum_{i=1}^{h} r_{(i)}^2$$

$r_{(1)}^2 \leq r_{(2)}^2 \leq \cdots \leq r_{(n)}^2$ are the ordered squared residuals $r_i^2 = (y_i - x_i^T\Theta)$, $i = 1, \ldots, n$, and $h$ is defined in range $\frac{n}{2} + 1 \leq h \leq \frac{3n+p+1}{4}$.

### 2.4. Suggested Model Selection Method.

In this study, in order to compute RCP and RTP criteria, we propose to use $\hat{\alpha}_S$ and $\hat{\alpha}_{LTS}$ instead of $\hat{\alpha}_M$ in equation (5), leading to

$$\hat{\alpha}_{pr}(d_M) = \hat{\alpha}_S(d_M) = (\Lambda + I)^{-1}(\Lambda + \hat{d}_M\boldsymbol{I})\hat{\alpha}_S \tag{11}$$

and

$$\hat{\alpha}_{pr}(d_M) = \hat{\alpha}_{LTS}(d_M) = (\Lambda + I)^{-1}(\Lambda + \hat{d}_M\boldsymbol{I})\hat{\alpha}_{LTS}$$

Consequently these estimators are used in (12) and (13) to estimate parameters $\hat{\beta}_{Liu.S}$ and $\hat{\beta}_{Liu.LTS}$ that are used in the calculation of selection criteria RCP and RTP.

$$\hat{\beta}_{Liu.S} = P^{'}\hat{\alpha}_S(d_M) \tag{12}$$

$$\hat{\beta}_{Liu.LTS} = P^{'}\hat{\alpha}_{LTS}(d_M) \tag{13}$$

where $\boldsymbol{P} = (q_1, q_2, \ldots, q_p)$ is the eigenvector matrix, such that $\boldsymbol{X}\boldsymbol{X}' = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'$.

In addition, we propose to use S estimator and LTS (least trimmed square) estimator in (12) for the calculation of selection criteria RCP and RTP, given in (8) and (10), respectively. In this way, $\hat{\boldsymbol{\beta}}$ for Vp in (8) and $\hat{\boldsymbol{\beta}}$ in (10) are obtained by using S estimator ($\hat{\beta}_{Liu.S}$), LTS estimator ($\hat{\beta}_{Liu.LTS}$). Equation (12) and Equation (13) are referred as robust Liu-S and robust Liu-LTS estimator, respectively.

### 2.5. Simulation Study.

In this section, a simulation study was performed to investigate and compare the robust variables selection criteria using S and LTS estimators. First, five independent variables were generated from Uniform distribution $(-1, 1)$. The data were obtained according to the $M1 = \beta = (5, 3, \sqrt{6}, 0, 0)$ and $M2 = \beta = (2\sqrt{5}, 4, \sqrt{3}, 1, 0)$ models. These parameters were obtained by considering $\beta'\beta/\sigma^2$ (non-central signal-to-noise parameter) and $\phi = \sum_i V_j'\beta / \sqrt{\sum_j \beta_i^2 \sum_{ij} V_{ij}^2}$ criteria and also used by Gunst and Mason(1977) and Erar(1982). In order to search of the effects of multicolinearity and outlier together over the robust selection criteria, a powerful linear dependency structure and a divergent-value were formed in the data sets between the x1 and x4; x2 and x3 variables in these models. Moreover, robust-d value is used in the robust Liu.M, robust Liu.S and robust Liu.LTS estimators, which are given equation 12-13. These estimators are used for computation of robust Tp (RTP) selection criteria based on Wald tests and robust Cp (RCP).

Furthermore, another goal of this simulation study is to see the results of RCP and RTP selection criteria with liu-S and liu-LTS estimators and to compare the results of RCP and RTP selection criteria with Liu and robust Liu.M estimators used in the previous studies (Çetin, 2009).

In order to obtain the percentages of subsets of criteria, a program was coded by using S-Plus function in this study. The results of the two models are given in the tables below. The numbers in these tables are shown how many times each subset is selected.

Table 1 shows that, the RTP, which is calculated by using both Liu.S and Liu. LTS estimators selects the real model which includes x1, x2, x3, with a proportion of %82 in a hundred repetitions, and it picked optional models x1, x2, x3, x4 and x1, x2 ,x3, x5 respectively in the proportions of %100 and %98. However, RCP criteria which is calculated by Liu.S and Liu.LTS estimators do not determine any subsets. On the contrary to RCP criteria calculated by Liu estimator gives better results than RTP criteria (Çetin, 2009).

As it can be seen from Table 2, RCP and RTP, both Liu.S and Liu.LTS estimators give the same results in case of multicollinearity. However, RTP criteria tend to choose multivariable models more often.

Table 3 gives the results under the assumption of multicollinearity and outliers together. RCP criteria do not work well when both the multicollinearity and outliers are present in the data. RTP criteria results are similar to those given in Table 2.

If we investigate Table 4,5 and 6, we can say that the results are similar to the result of model 1. Thus, we can say that Liu.s estimators with RCP and RTP criteria do not show

**Table 1.** Proportions of subsets order selected by criteria without outlier and multicollinearity for M1 model

| **With robust Liu.S estimators** | | | **With robust Liu.LTS estimators** | |
|---|---|---|---|---|
| **subsets** | **RCP < VP** | **RTP < P** | **RCP < VP** | **RTP < P** |
| X1 | 0 | 0 | 0 | 0 |
| X2 | 0 | 0 | 0 | 0 |
| X3 | 0 | 0 | 0 | 0 |
| X4 | 0 | 0 | 0 | 0 |
| X5 | 0 | 0 | 0 | 0 |
| X1 x2 | 12 | 0 | 14 | 0 |
| X1 x3 | 0 | 0 | 0 | 0 |
| X1 x4 | 0 | 0 | 0 | 0 |
| X1 x5 | 0 | 0 | 0 | 0 |
| X2 x3 | 0 | 0 | 0 | 0 |
| X2 x4 | 0 | 0 | 0 | 0 |
| X2 x5 | 0 | 0 | 0 | 0 |
| X3 x4 | 0 | 0 | 0 | 0 |
| X3 x5 | 0 | 0 | 0 | 0 |
| X4 x5 | 0 | 0 | 0 | 0 |
| X1 x2 x3 | 8 | 82 | 6 | 86 |
| X1 x2 x4 | 2 | 0 | 0 | 0 |
| X1 x2 x5 | 0 | 0 | 0 | 0 |
| X1 x3 x4 | 2 | 0 | 1 | 0 |
| x1 x3 x5 | 0 | 0 | 0 | 0 |
| x1 x4 x5 | 5 | 0 | 1 | 0 |
| x2 x3 x4 | 0 | 0 | 1 | 0 |
| x2 x3 x5 | 0 | 0 | 0 | 0 |
| x2 x4 x5 | 0 | 0 | 0 | 0 |
| x3 x4 x5 | 0 | 0 | 0 | 0 |
| x1 x2 x3 x4 | 0 | 100 | 0 | 100 |
| x1 x2 x3 x5 | 0 | 98 | 0 | 99 |
| x1 x2 x4 x5 | 6 | 0 | 0 | 0 |
| x1 x3 x4 x5 | 0 | 0 | 0 | 0 |
| x2 x3 x4 x5 | 11 | 0 | 16 | 0 |

any improvements. Liu.lts estimators with RCP and RTP criteria show better results. Under multi-collinearity condition, RTP and RCP criteria selected the true model (1234) but RTP tend to select other four variable models as well. Moreover, Liu.S and Liu.lts estimators with RCP and RTP criteria did not perform well under outliers and multi-collinearity.

**Table 2.** Proportions of subsets order selected by criteria in case of multicollinearity for M1 model

| **With robust Liu.S estimators** | | | **With robust Liu.LTS estimators** | |
|---|---|---|---|---|
| **subsets** | **RCP $<$ VP** | **RTP $<$ P** | **RCP $<$ VP** | **RTP $<$ P** |
| x1 | 0 | 3 | 1 | 4 |
| x2 | 28 | 0 | 23 | 0 |
| x3 | 0 | 0 | 0 | 0 |
| x4 | 0 | 0 | 1 | 0 |
| x5 | 0 | 0 | 0 | 0 |
| x1 x2 | 58 | 21 | 42 | 25 |
| x1 x3 | 76 | 22 | 57 | 24 |
| x1 x4 | 0 | 7 | 0 | 5 |
| x1 x5 | 0 | 33 | 1 | 31 |
| x2 x3 | 32 | 1 | 10 | 1 |
| x2 x4 | 78 | 0 | 59 | 2 |
| x2 x5 | 49 | 1 | 41 | 1 |
| x3 x4 | 36 | 1 | 32 | 1 |
| x3 x5 | 3 | 1 | 6 | 1 |
| x4 x5 | 0 | 7 | 0 | 5 |
| x1 x2 x3 | 52 | 72 | 46 | 43 |
| x1 x2 x4 | 7 | 87 | 25 | 83 |
| x1 x2 x5 | 77 | 73 | 70 | 45 |
| x1 x3 x4 | 57 | 93 | 44 | 91 |
| x1 x3 x5 | 83 | 74 | 72 | 45 |
| x1 x4 x5 | 11 | 78 | 45 | 63 |
| x2 x3 x4 | 84 | 91 | 63 | 90 |
| x2 x3 x5 | 75 | 10 | 66 | 5 |
| x2 x4 x5 | 80 | 88 | 75 | 90 |
| x3 x4 x5 | 87 | 100 | 73 | 100 |
| x1 x2 x3 x4 | 5 | 95 | 5 | 96 |
| x1 x2 x3 x5 | 7 | 62 | 3 | 32 |
| x1 x2 x4 x5 | 1 | 90 | 4 | 92 |
| x1 x3 x4 x5 | 0 | 100 | 0 | 100 |
| x2 x3 x4 x5 | 0 | 100 | 0 | 100 |

**Table 3.** Proportions of subsets order selected by criteria in case of multicollinearity and outliers for M1 model

| **With robust Liu.S estimators** | | | **With robust Liu.LTS estimators** | |
|---|---|---|---|---|
| **subsets** | **RCP < VP** | **RTP < P** | **RCP < VP** | **RTP < P** |
| x1 | 4 | 0 | 3 | 1 |
| x2 | 10 | 0 | 13 | 0 |
| x3 | 3 | 0 | 3 | 0 |
| x4 | 2 | 2 | 6 | 2 |
| x5 | 1 | 0 | 6 | 0 |
| x1 x2 | 10 | 1 | 20 | 2 |
| x1 x3 | 11 | 1 | 13 | 2 |
| x1 x4 | 20 | 4 | 18 | 9 |
| x1 x5 | 13 | 1 | 11 | 1 |
| x2 x3 | 3 | 0 | 9 | 0 |
| x2 x4 | 13 | 11 | 16 | 11 |
| x2 x5 | 3 | 0 | 16 | 0 |
| x3 x4 | 10 | 11 | 13 | 12 |
| x3 x5 | 11 | 0 | 24 | 0 |
| x4 x5 | 10 | 7 | 19 | 9 |
| x1 x2 x3 | 11 | 9 | 8 | 4 |
| x1 x2 x4 | 14 | 75 | 24 | 64 |
| x1 x2 x5 | 10 | 10 | 8 | 4 |
| x1 x3 x4 | 18 | 78 | 31 | 66 |
| x1 x3 x5 | 8 | 10 | 8 | 4 |
| x1 x4 x5 | 18 | 62 | 20 | 40 |
| x2 x3 x4 | 3 | 79 | 3 | 68 |
| x2 x3 x5 | 1 | 0 | 3 | 1 |
| x2 x4 x5 | 4 | 99 | 3 | 93 |
| x3 x4 x5 | 1 | 100 | 0 | 100 |
| x1 x2 x3 x4 | 1 | 76 | 0 | 66 |
| x1 x2 x3 x5 | 7 | 14 | 5 | 5 |
| x1 x2 x4 x5 | 1 | 97 | 4 | 91 |
| x1 x3 x4 x5 | 1 | 100 | 4 | 100 |
| x2 x3 x4 x5 | 4 | 100 | 5 | 100 |

**Table 4.** Proportions of subsets order selected by criteria without outlier and multicollinearity for M2 model

| **With robust Liu.S estimators** | | | **With robust Liu.LTS estimators** | |
|---|---|---|---|---|
| **subsets** | **RCP < VP** | **RTP < P** | **RCP < VP** | **RTP < P** |
| x1 | 33 | 0 | 9 | 1 |
| x2 | 23 | 0 | 0 | 0 |
| x3 | 28 | 0 | 0 | 0 |
| x4 | 45 | 0 | 0 | 0 |
| x5 | 7 | 0 | 2 | 0 |
| x1 x2 | 9 | 0 | 24 | 0 |
| x1 x3 | 8 | 48 | 7 | 0 |
| x1 x4 | 6 | 0 | 5 | 0 |
| x1 x5 | 8 | 0 | 1 | 0 |
| x2 x3 | 9 | 0 | 1 | 0 |
| x2 x4 | 10 | 0 | 3 | 0 |
| x2 x5 | 0 | 0 | 5 | 0 |
| x3 x4 | 9 | 0 | 1 | 0 |
| x3 x5 | 3 | 0 | 1 | 0 |
| x4 x5 | 6 | 0 | 2 | 0 |
| x1 x2 x3 | 9 | 56 | 100 | 67 |
| x1 x2 x4 | 2 | 0 | 0 | 2 |
| x1 x2 x5 | 10 | 0 | 2 | 1 |
| x1 x3 x4 | 8 | 53 | 4 | 2 |
| x1 x3 x5 | 8 | 49 | 1 | 1 |
| x1 x4 x5 | 2 | 0 | 0 | 1 |
| x2 x3 x4 | 93 | 0 | 0 | 1 |
| x2 x3 x5 | 6 | 0 | 0 | 2 |
| x2 x4 x5 | 93 | 0 | 0 | 1 |
| x3 x4 x5 | 1 | 0 | 4 | 3 |
| x1 x2 x3 x4 | 43 | 60 | 85 | 100 |
| x1 x2 x3 x5 | 41 | 35 | 0 | 22 |
| x1 x2 x4 x5 | 60 | 0 | 1 | 12 |
| x1 x3 x4 x5 | 72 | 71 | 1 | 1 |
| x2 x3 x4 x5 | 82 | 0 | 1 | 1 |

**Table 5.** Proportions of subsets order selected by criteria in case of multicollinearity for M2 model

| **With robust Liu.S estimators** | | | **With robust Liu.LTS estimators** | |
|---|---|---|---|---|
| **subsets** | **RCP < VP** | **RTP < P** | **RCP < VP** | **RTP < P** |
| x1 | 53 | 73 | 9 | 1 |
| x2 | 43 | 0 | 43 | 1 |
| x3 | 66 | 0 | 10 | 2 |
| x4 | 45 | 0 | 11 | 1 |
| x5 | 3 | 0 | 0 | 3 |
| x1 x2 | 10 | 31 | 12 | 2 |
| x1 x3 | 10 | 37 | 51 | 2 |
| x1 x4 | 19 | 50 | 10 | 5 |
| x1 x5 | 18 | 33 | 15 | 3 |
| x2 x3 | 9 | 13 | 10 | 1 |
| x2 x4 | 12 | 10 | 59 | 2 |
| x2 x5 | 17 | 12 | 46 | 2 |
| x3 x4 | 9 | 9 | 36 | 10 |
| x3 x5 | 3 | 11 | 9 | 14 |
| x4 x5 | 6 | 17 | 7 | 15 |
| x1 x2 x3 | 100 | 73 | 100 | 56 |
| x1 x2 x4 | 2 | 91 | 51 | 81 |
| x1 x2 x5 | 10 | 73 | 87 | 25 |
| x1 x3 x4 | 89 | 97 | 85 | 29 |
| x1 x3 x5 | 83 | 73 | 98 | 13 |
| x1 x4 x5 | 100 | 87 | 98 | 26 |
| x2 x3 x4 | 93 | 96 | 99 | 39 |
| x2 x3 x5 | 60 | 23 | 56 | 7 |
| x2 x4 x5 | 93 | 97 | 78 | 19 |
| x3 x4 x5 | 50 | 85 | 73 | 32 |
| x1 x2 x3 x4 | 58 | 100 | 85 | 99 |
| x1 x2 x3 x5 | 41 | 68 | 78 | 3 |
| x1 x2 x4 x5 | 50 | 100 | 80 | 12 |
| x1 x3 x4 x5 | 72 | 100 | 80 | 10 |
| x2 x3 x4 x5 | 67 | 100 | 78 | 10 |

**Table 6.** Proportions of subsets order selected by criteria in case of multicollinearity and outliers for M2 model

| **With robust Liu.S estimators** | | | **With robust Liu.LTS estimators** | |
|---|---|---|---|---|
| **subsets** | **RCP < VP** | **RTP < P** | **RCP < VP** | **RTP < P** |
| x1 | 62 | 53 | 8 | 3 |
| x2 | 45 | 20 | 23 | 4 |
| x3 | 68 | 20 | 10 | 20 |
| x4 | 45 | 10 | 10 | 15 |
| x5 | 31 | 10 | 9 | 3 |
| x1 x2 | 10 | 28 | 7 | 2 |
| x1 x3 | 10 | 33 | 7 | 5 |
| x1 x4 | 31 | 52 | 11 | 10 |
| x1 x5 | 36 | 31 | 10 | 3 |
| x2 x3 | 19 | 23 | 14 | 5 |
| x2 x4 | 21 | 10 | 34 | 9 |
| x2 x5 | 17 | 17 | 46 | 6 |
| x3 x4 | 39 | 19 | 54 | 11 |
| x3 x5 | 1 | 11 | 7 | 20 |
| x4 x5 | 13 | 17 | 7 | 15 |
| x1 x2 x3 | 100 | 56 | 23 | 15 |
| x1 x2 x4 | 100 | 100 | 78 | 88 |
| x1 x2 x5 | 100 | 73 | 89 | 25 |
| x1 x3 x4 | 99 | 54 | 62 | 29 |
| x1 x3 x5 | 83 | 53 | 72 | 13 |
| x1 x4 x5 | 100 | 100 | 85 | 46 |
| x2 x3 x4 | 93 | 99 | 66 | 49 |
| x2 x3 x5 | 60 | 20 | 72 | 37 |
| x2 x4 x5 | 93 | 97 | 72 | 99 |
| x3 x4 x5 | 50 | 100 | 72 | 100 |
| x1 x2 x3 x4 | 75 | 100 | 83 | 45 |
| x1 x2 x3 x5 | 79 | 93 | 72 | 5 |
| x1 x2 x4 x5 | 64 | 100 | 74 | 100 |
| x1 x3 x4 x5 | 79 | 100 | 86 | 100 |
| x2 x3 x4 x5 | 89 | 100 | 77 | 100 |

## 3. Conclusion

RTP criteria with Liu.S and Liu.LTS estimators propose the best performance in case of the absence of any violation in the model assumptions. Despite the absence of any distortion in the assumptions, RCP criteria does not select the true model. Under the presence of outliers and multicollinearity, both RTP and RCP with Liu.S and Liu.LTS estimators do not work well. However, RCP criteria with Liu estimator showed better results (Çetin, 2009).

## References


[1] Arslan, O. and Billor, N., *Robust Liu estimator for regression based on an M-estimator*, Journal of Applied Stats., Vol. 27, No.1, 39-47, 2000.

[2] Çetin, M., *Variable Selection Criteria in Robust Regression*, PhD. Thesis, University of Hacettepe, 2000.

[3] Çetin, M., Robust model selection criteria for robust Liu estimator, European Journal of Operation Research" , Vol.199, issue 1,21-24, 2009.

[4] Erar, A., *Variable Selection in Linear Regression Models in case of multicollinearity*, Unpublished Ph.D. Thesis, Ankara,Turkey, 1982.

[5] Gunst, R. F., Mason, R.L., *Advantages of examining multicollinearities in regression analysis*, Biometrics, 33, 249-260, 1977.

[6] Hampel, F. R., *Some Aspects of Model Choice in Robust Statistics*, Proceedings of the 44th Session of the ISI, Madrid, Book 2, 767-771, 1983.

[7] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., and Stahel, W. A., *Robust Statistics: The Approach Based on Influence Functions*, New York, John Wiley, 1986.

[8] Hurvich, C., F. and Tsai, C. L., *Model selection for least absolute deviations regression in small samples*, Statistics and Probab. Letters, 9, 259-265, 1990.

[9] Liu, K., *A new class of baised estimate in linear regression*, Comm. in Stats. A, 22,393-402, 1993.

[10] Mallows, C. L., *Some comment on Cp, Technometrics*, Vol.15, No: 4, 661- 675, 1973.

[11] Ronchetti, E., *Robust model selection in regression*, Statistics and Prob. Letters., 3, 21-23, 1985.

[12] Ronchetti, E. and Staudte, R., *A Robust version of Mallows's Cp*, JASA,Vol. 89, 550-559, 1994.

[13] Ronchetti, E., Field, C., and Blanhard, W., *Robust linear model selection by cross-validation*, JASA, Vol.92, No.439, Theory and methods, 1017-1023, 1997.

[14] Rousseeuw, P. J., *Least Median of Squares Regression*, Journal of the American Statistical Association, 79, 871-880, 1984.

[15] Rousseeuw, P. J. and Yohai, V. J., *Robust Regression by Means of S-estimators, in Robust and Nonlinear Time Series Analysis*, ed. By W.H. Franke, and D. Martin, Springer-Verlag, New York, pp.256-272, 1984.

[16] Sommer, S. and Huggins, R. M., *Variable selection using the Wald test and a robust Cp*, Appl. Statist., 45, 15-29, 1996.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 165 – 180

# Data envelopment analysis approach for discriminating efficient candidates in voting systems by considering the priority of voters

Ali Ebrahimnejad*† and Mohammad Reza Bagherzadeh ‡

**Abstract**

There are different ways to allow the voters to express their preferences on a set of candidates. In the traditional voting systems, it is assumed that the votes of all voters have the same importance and there is no preference between them. In this paper, a new approach is proposed to express the preference of voters on a set of candidates. In the proposed approach voters are classified into several categories with different importance levels in which the vote of a higher category may have a greater importance than that of the lower category. Then two models are introduced to measure the best preference scores of the target candidate from the virtual best candidate and the virtual worst candidate point of view. After that, two obtained preference scores are aggregated together in order to obtain an overall ranking. Finally, two numerical examples are provided for illustration the applications of the proposed approach.

**Keywords:** Data envelopment analysis, Voting system, Preference score, Virtual candidate.

*2000 AMS Classification:* 90Bxx, 90B50.

*Received* : 29.04.2014 *Accepted* : 02.12.2014 *Doi* : 10.15672/HJMS.2015489355

---

*Department of Mathematics, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran, Email: aemarzoun@gmail.com and a.ebrahimnejad@srbiau.ac.ir

†Corresponding Author.

‡Department of Public Management, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran Email: dr.mr.bagherzadeh@gmail.com

## 1. Introduction

There are different ways to allow the voters to express their preferences on a set of candidates. In some voting systems, each voter selects some candidates and ranks them from most to least. Among these systems, the well-known procedures to obtain a social ranking or a winning candidate are scoring rules, where fixed scores are assigned to the different ranks. In this way, the score obtained by each candidate is the weighted sum of the scores receives in different places. The plurality rule (the winner candidate is the one who receives more votes in the first place), the Borda rule (the weight assigned to the first place equals to number of candidates and to the second place is one less than the first place and so on) are the best known instances of scoring rules. In spite of the Borda rule has interesting properties in relation to other scoring rules [5], but the utilization of a fixed scoring vector has weak point that a candidate that is not the winner with the scoring vector imposed initially could be so if another scoring is used. To avoid this problem, Cook and Kress [7] suggested evaluating each candidate with the most favorable scoring vector for him/her. With this purpose, they introduced Data Envelopment Analysis (DEA) in this context. DEA determines the most favorable weights for each candidate. Different candidates utilize different sets of weights to calculate their total scores, which are referred to as the best relative total scores and are all restricted to be less than or equal to one. The candidate with the biggest relative total score of unity is said to be efficient candidate and may be considered as a winner. The principal drawback of this method is very often leads to more than one candidate to be efficient candidate. We can judge that the set of efficient candidates is the top group of candidates, but cannot single out only one winner among them. To avoid this weakness, Cook and Kress [7], proposed to maximize the gap between consecutive weights of the scoring vector. However, Green et al. [15] noticed two important drawbacks of the previous procedure. The first one is that the choice of the intensify functions used in their model is not obvious, and that choice determines the winner. The second one is that for an important class of discrimination intensity functions the previous procedure is equivalent to imposing a common set of scores on all candidates. Therefore, when Cook and Kress's model is used with this class of discrimination intensity functions, the aim pursued by these authors (evaluating each candidate with the most favourable scoring vector for him/her) is not reached.

Due to the drawbacks mentioned above, other procedures to discriminate efficient candidates have appeared in the literature. Green et al. [15] proposed to use the cross-evaluation method, introduced by Sexton et al. [32] to discriminate efficient candidates. Hashimoto [18] used the DEA exclusion method (see Andersen and Petersen [4]) to Cook and Kress's model. Hashimoto's model is useful to discriminate efficient candidates, but it is unstable with respect to inefficient candidates too. Noguchi et al. [28] criticized the choice of discrimination intensity functions in Green et al.'s model. In their model, the weight assigned to a certain rank may be zero and, consequently, the votes granted to that rank are not considered. Furthermore, the weights corresponding to two different ranks may be equal and, therefore, the rank votes lose their meaning. To avoid the previous drawbacks, Noguchi et al. [28] gave a strong ordering constraint condition on weights. Besides the previous condition on the scoring vectors, Noguchi et al. [28] introduced two other modifications in the model of Green et al. [15]. On the one hand, in the cross-evaluation matrix each candidate utilizes the same scoring vector to evaluate each of the remaining candidates. However, Noguchi et al.'s model maintains the problems of Green et al.'s model. Obata and Ishii [29] proposed another model that does not use any information about inefficient candidates. To obtain a fair approach, they used weight vectors of the same size, by normalizing the most favorable weight vectors. But it presents other drawbacks. In their model it is necessary to determine the norm and the discrimination intensity functions to use. If these functions are zero and the $L_\infty$-norm is used, the winning candidate coincides with the one obtained by means of a scoring rule. If $L_\infty$-norm is replaced by $L_1$-norm, the outcome could be considered unfair by some candidates. Foroghi and Tamiz [13] and Foroghi et al. [12] extended and simplified their model with fewer constraints and also used it for ranking inefficient as well as efficient candidates. Llamazares and Pena [26] analyzed the principal ranking methods proposed in the literature to discriminate efficient candidates and by solving several examples showed that none of the previous proposed procedures was fully convincing. In fact, although all the previous methods do not require predetermine the weights subjectively, some of them have a serious drawback: the relative order between two candidates may be altered when the number of first, second, $\ldots$, $k$th ranks obtained by other candidates changes, although there is not any variation in the number of first, second, $\ldots$ , $k$th ranks obtained by both candidates. Thus, Llamazares and Pena [26] proposed a model that allows each candidate to be evaluated with the most favorable weighting vector for him/her and avoids the mentioned drawback. Moreover, in some cases, they found a closed expression for the score assigned with their model to each candidate.

Wang and Chin [39] discriminated efficient candidates by considering their least relative total scores. But the least relative total scores and the best relative total scores are not measured within the same range. The obtained conclusion was not persuasive. They also proposed a model in which the total scores are measured within an interval. The upper bound of the interval was set to be one, but they failed to determine the value of the lower bound for the interval. After that, Wang et al. [40] proposed a method to rank multiple efficient candidates, which often happens in DEA method, by comparing the least relative total scores for each efficient candidate with the best and the least relative total scores measured in the same range.

Wang et al. [42] proposed three new models to assess the weights associated with different ranking places in preference voting and aggregation. Two of them are linear programming models which determine a common set of weights for all the candidates considered and the other is a non-linear programming model that determines the most favorable weights for each candidate. Hadi-Vencheh and Mokhtarian [16] presented three counter examples to show that the three new models developed by Wang et al. [42] for preference voting and aggregation may produce a zero weight for the last ranking place and may sometimes identify two candidates as the winner in some specific situations. After that, Wang et al. [43] presented two modified linear programming models for preference voting and aggregation to avoid the zero weight for the last ranking place. In addition, Hadi-Vencheh [17] proposed two improved DEA models to determine the weights of ranking places that each of them can lead to a stable full ranking for all the candidates considered and avoid the mentioned shortcoming. Wu et al. [45] considered a preferential voting system using DEA game cross efficiency model, in which each candidate is viewed as a player that seeks to maximize its own efficiency, under the condition that the cross efficiencies of each of other DMU's does not deteriorate. Jahanshaloo et al. [22] reviewed ranked voting data and its analysis with DEA and proposed a model based on the ranking of units using common weights. Their model gives one common set of weights that is the most favorable for determining the absolute efficiency of all candidates at the same time. Bystricky [6] investigated different approaches to weighted voting systems based on preferential positions. In addition, other models have appeared in the literature in order to deal with this kind of problems [1, 3, 8, 9, 10, 11, 19, 20, 21, 25, 31, 33, 34, 36, 37, 38].

However, all previous models are based on Cook and Kress's model in which the votes of all voters have equal importance and there is no preference among them. In this paper we generalize the existing models to overcome this shortcoming. In fact, in our proposed model voters are classified into several categories with different importance levels that the vote of a higher category may have a greater importance than that of the lower category. Our main contribution in this paper will be the simplification of the model of Wu [44] (first proposed by Wang and Luo [41]) in DEA efficiency assessment for an overall ranking of candidates. We introduce two models that the first model evaluates candidates from the viewpoint of the best possible preference score and the second model evaluates them from the perspective of the worst possible preference score. The two distinctive scores are combined to form a comprehensive index such that an overall ranking for all the candidates can be obtained.

The rest of this paper is organized as follows. Section 2 gives the traditional voting model proposed by Cook and Kress [7] considering all of voters are in one category. Section 3 gives our model to determine efficient candidates by classifying voters into several groups with different importance levels. Section 4 extends the existing ranking method to discriminate the efficient candidates in terms of our proposed model assumptions. In Section 5 we illustrate our new methodology with two numerical examples. This paper is concluded in Section 6.

## 2. Ranked voting data

In this section we consider ranked voting data such that each voter select $m$ candidate from $n$ ( $n \geq m$ ) candidates $\{A_1, A_1, \ldots,\ A_n\ \}$ and rank them from top to the place $m$, each place associated with a relative important weight $u_i^r$ $(i = 1, 2, \ldots, m)$. In this way, the score obtained by the candidate $A_r$ is $z_r = \sum_{i=1}^{m} u_i^r y_i^r$ where $y_i^r$ is the number vote of place $i$ that candidate $A_r$ occupies and $(u_1^r,\ u_2^r, \ldots,\ u_m^r)$ is the scoring vector used.

**2.1. Remark.** In the DEA framework, many voting models are based on that of Cook and Kress [7], where the input variables $u_i^r$ $(i = 1, 2, \ldots, m)$ are the weights, and these values are real numbers. Thus, the DEA models applied to find these weights as the relative importance of each place are not integer models.

Cook and Kress [7] suggested evaluating each candidate with the most favorable scoring vector for him/her based on DEA models. Their DEA/assurance region (DEA/AR) model is as follows:

$$
\begin{aligned}
z_p^* = \max \quad & \sum_{i=1}^{m} u_i^p y_i^p \\
s.t. \quad & \sum_{i=1}^{m} u_i^p y_i^r \leq 1, && r = 1, \cdots, n, \\
& u_i^p - u_{i+1}^p \geq d(i, \varepsilon), && i = 1, \cdots, m-1, \\
& u_m^p \geq d(m, \varepsilon),
\end{aligned}
\tag{2.1}
$$

In the above model, $d(., \varepsilon)$ is called the discrimination intensity function that is non-negative and monotonically increasing in a non-negative $\varepsilon$ and satisfies $d(., 0) = 0$. The last constraints in (2.1) are called the assurance region constraints and ensure that the votes of the higher place has a greater importance that of the lower place. The model (2.1) is solved for each candidate $p$ $(p = 1, \ldots, n)$. The resulting score $z_p^*$ is the preference score of the candidate $p$. This score is used to rank of all candidates in a voting system that assumes that the votes of all voters have the same importance and there is no preference between them.

In this next section, we extend model (2.1) for situations that voters are classified into several categories with different importance levels in which the vote of a higher category may have a greater importance than that of the lower category.

## 3. An extended model

In this section, we introduce a new approach to allow the voters to express their preferences on a set of candidates by classifying voters into several groups with different importance levels. Suppose that in a ranked voting system, voters are classified into $k$ distinct categories. The voters of each category, select $m$ candidates among $n$ ( $n \geq m$ ) candidates $\{A_1, A_1, \ldots, A_n\}$ and rank them from top to the place $m$. Let $y_{ij}^r$ be the votes of the candidate $r$ being ranked in the place $i$ from the category $j$. In evaluating of the candidate $r$, each place is associated with a relative importance weight $u_i^r$ ($i = 1, 2, \ldots, m$) and each category is associated with a relative importance weight $v_j^r$($j = 1, 2, \ldots, k$). The preference score of candidate $r$ in the place $i$ is equal to $\sum_{j=1}^{k} v_j^r y_{ij}^r$. Thus, the total preference score of candidate $r$ will be $z_r = \sum_{i=1}^{m} (u_i^r \sum_{j=1}^{k} v_j^r y_{ij}^r) = \sum_{i=1}^{m}\sum_{j=1}^{k} u_i^r v_j^r y_{ij}^r$ .

It should be noted that if all categories have the same relative importance weights, then the preference score of candidate $r$ in the place $i$ will be $\sum_{j=1}^{k} y_{ij}^r$, that is exactly the number of votes in place $i$ received by candidate $r$. In this case, the preference score of candidate $r$ is equal to the one which assumes voters are in one category. Thus, this value indicates the real score of each candidate.

However to obtain a total ranking of candidates, we require the weight vectors $u^r = (u_1^r, \ldots, u_m^r)$ and $v^r = (v_1^r, \ldots, v_k^r)$ satisfy the following conditions:

$$\begin{aligned} &u_i^r - u_{i+1}^r \geq \overline{d}(i,\varepsilon), \quad i = 1, \cdots, m-1 \\ &u_m^r \geq \overline{d}(m,\varepsilon) \end{aligned} \tag{3.1}$$

$$\begin{aligned} &v_j^r - v_{j+1}^r \geq \overline{\overline{d}}(j,\varepsilon), \quad j = 1, \cdots, k-1 \\ &v_k^r \geq \overline{\overline{d}}(k,\varepsilon) \end{aligned} \tag{3.2}$$

It needs to point out that the constraints (3.1) are introduced in order that the vote of the higher place may have a greater importance than that of the lower place. In a similar way, the constraints (3.2) are introduced in order that the vote of voters in a higher category has a greater importance than that in a lower category. Hence, the following non-linear model evaluates candidate $p$ with the most favorable weight vectors:

$$\begin{aligned} z_p^* = \max \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} u_i^p\, v_j^p\, y_{ij}^p \\ s.t. \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} u_i^p\, v_j^p\, y_{ij}^r \leq 1, \quad r = 1, 2, \ldots, n, \\ & u_i^p - u_{i+1}^p \geq \overline{d}(i,\varepsilon), \quad i = 1, \cdots, m-1 \\ & u_m^p \geq \overline{d}(m,\varepsilon) \\ & v_j^p - v_{j+1}^p \geq \overline{\overline{d}}(j,\varepsilon), \quad j = 1, \cdots, k-1 \\ & v_m^p \geq \overline{\overline{d}}(m,\varepsilon) \end{aligned} \tag{3.3}$$

To transform the non-linear model (3.3) into an equivalent linear model, let

$$w_{ij}^p = u_i^p v_j^p, \qquad i = 1, 2, \ldots, m \quad , \quad j = 1, 2, \ldots, k. \tag{3.4}$$

Now, we should change the constraint of (3.1) and (3.2) in terms of new transformations such that the priority among places and categories preserves. To this end, we multiply the constraints of (3.1) and (3.2) by $v_j^r (j = 1, 2, \ldots, k)$ and $u_i^r (i = 1, 2, \ldots, m)$, from the right and left, respectively. Thus, we have:

$$u_i^r v_j^r - u_{i+1}^r v_j^r \geq \overline{d}(\,i,\varepsilon),\quad i=1,\cdots,m-1,\; j=1,2,\cdots,k$$
$$u_m^r v_j^r \geq \overline{d}(m,\varepsilon),\; j=1,2,\cdots,k$$

$$u_i^r v_j^r - u_i^r v_{j+1}^r \geq \overline{\overline{d}}(j,\varepsilon),\quad j=1,\cdots,k-1,\; i=1,\cdots,m$$
$$u_i^r v_k^r \geq \overline{\overline{d}}(k,\varepsilon),\; i=1,\cdots,m$$

Thus, we have:

$$w_{ij}^r - w_{i+1,j}^r \geq \overline{d}(i,\varepsilon),\quad i=1,\cdots,m-1,\; j=1,2,\cdots,k \qquad (3.5)$$
$$w_{mj}^r \geq \overline{d}(m,\varepsilon),\; j=1,2,\cdots,k$$

$$w_{ij}^r - w_{i,j+1}^r \geq \overline{\overline{d}}(j,\varepsilon),\quad j=1,\cdots,k-1,\; i=1,\cdots,m \qquad (3.6)$$
$$w_{ik}^r \geq \overline{\overline{d}}(k,\varepsilon),\; i=1,\cdots,m$$

By substituting (3.4)-(3.5) into model (3.3), the following linear model is obtained:

$$
\begin{aligned}
z_p^* = \max \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^p\; y_{ij}^p \\
s.t. \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^p\; y_{ij}^r \leq 1, \qquad r=1,2,\ldots,n, \\
& w_{ij}^p - w_{i+1,j}^p \geq \overline{d}(i,\varepsilon),\quad i=1,\cdots,m-1,\; j=1,\cdots,k \\
& w_{mj}^p \geq \overline{d}(m,\varepsilon),\quad j=1,\cdots,k \\
& w_{ij}^p - w_{i,j+1}^p \geq \overline{\overline{d}}(j,\varepsilon),\quad j=1,\cdots,k-1,\; i=1,\cdots,m \\
& w_{ik}^p \geq \overline{\overline{d}}(m,\varepsilon),\quad i=1,\cdots,m
\end{aligned} \qquad (3.7)
$$

In the next section, we introduce two virtual candidates called virtual best candidate (VBC) and virtual worst candidate (VWC) into voting system. The resultant voting models are referred to as the voting system with VBC and VWC candidates, respectively. The first system evaluates candidates from the viewpoint of the best possible preference score and the second system evaluates them from the perspective of the worst possible preference score. The two distinctive scores are combined to form a comprehensive index called the relative closeness (RC) to the VBC just like the well-known TOPSIS approach in multiple attribute decision making (MADM). The RC index is then used as the evidence of overall scores of each candidate, based on which an overall ranking for all the candidates can be obtained.

## 4. Voting systems with VBC and VWC

In this section we give some models so that a voting analysis based on TOPSIS idea can be performed. To do this, we first explore the concepts of virtual best candidate (VBC) and virtual worst candidate (VWC).

**4.1. Definition.** The virtual best candidate (VBC) is a virtual candidate that receives the most votes in each place among all candidates.

It needs to point out that the VBC may not exist in the voting. But he/she receives the most votes in each place among all $n$ candidates. According to the above definition, we denote by $Y_i^{max} = (y_{i1}^{max},\ldots,y_{ik}^{max})$ the number votes of VBC in place $i$, in which the votes of each category in this place are determined by $y_{ij}^{max} = \max_r\{y_{ij}^r\}$ . In fact, VBC receives the most votes in each place and each category among all candidates and will be ranked in first place in any condition.

**4.2. Definition.** The virtual worst candidate (VWC) is a virtual candidate that receives the least votes in each place among all candidates.

It is also important to note that the VWC may not exist in the voting. But he/she receives the least votes in each place among all $n$ candidates. According to the above definition, we denote by $Y_i^{min} = (y_{i1}^{min}, \ldots, y_{ik}^{min})$ the number votes of VWC in place $i$, in which the votes of each category in this place are determined by $y_{ij}^{min} = \min_r\{y_{ij}^r\}$ . In fact, VBC receives the least votes in each place and each category among all candidates and will be ranked in last place in any condition.

It is obvious that the VBC should be able to achieve the highest/best preference score. The best preference score of VBC denoted as $\phi^*$ is determined by the following model:

$$
\begin{aligned}
\phi^*_{VBC} = \max \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{VBC}\; y_{ij}^{max} \\
s.t. \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{VBC}\; y_{ij}^{r} \le 1, \qquad r = 1,2,\ldots,n, \\
& w_{ij}^{VBC} - w_{i+1,j}^{VBC} \ge \overline{d}(i,\varepsilon), \quad i = 1,\cdots,m-1,\; j = 1,\cdots,k \\
& w_{mj}^{VBC} \ge \overline{d}(m,\varepsilon), \quad j = 1,\cdots,k \\
& w_{ij}^{VBC} - w_{i,j+1}^{VBC} \ge \overline{\overline{d}}(j,\varepsilon), \quad j = 1,\cdots,k-1,\; i = 1,\cdots,m \\
& w_{ik}^{VBC} \ge \overline{\overline{d}}(m,\varepsilon), \quad i = 1,\cdots,m
\end{aligned}
\tag{4.1}
$$

Since the above linear programming model (4.1) may have multiple optima, we utilize the following linear programming model to determine the best preference score of candidate $p$ under the condition that the best possible preference score of the VBC remains unchanged:

$$
\begin{aligned}
z^*_p = \max \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{p}\; y_{ij}^{p} \\
s.t. \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{VBC}\; y_{ij}^{max} = \phi^*_{VBC} \\
& \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{p}\; y_{ij}^{r} \le 1, \qquad r = 1,2,\ldots,n, \\
& w_{ij}^{p} - w_{i+1,j}^{p} \ge \overline{d}(\,i,\varepsilon), \quad i = 1,\cdots,m-1,\; j = 1,\cdots,k \\
& w_{mj}^{p} \ge \overline{d}(\,m,\varepsilon), \quad j = 1,\cdots,k \\
& w_{ij}^{p} - w_{i,j+1}^{p} \ge \overline{\overline{d}}(\,j,\varepsilon), \quad j = 1,\cdots,k-1,\; i = 1,\cdots,m \\
& w_{mj}^{p} \ge \overline{\overline{d}}(\,m,\varepsilon), \quad i = 1,\cdots,m
\end{aligned}
\tag{4.2}
$$

Similar to that in Wu [44], the following model is proposed to compute the worst possible preference score of the VWC:

$$
\begin{aligned}
\varphi^*_{VBC} = \quad & min \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{VWC}\; y_{ij}^{min} \\
s.t. \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{VBC}\; y_{ij}^{max} \ge \gamma, \quad \gamma \in [1, \phi^*_{VBC}] \\
& \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{VWC}\; y_{ij}^{r} \le 1, \qquad r = 1,2,\ldots,n, \\
& w_{ij}^{VWC} - w_{i+1,j}^{VWC} \ge \overline{d}(i,\varepsilon), \quad i = 1,\cdots,m-1,\; j = 1,\cdots,k \\
& w_{mj}^{VWC} \ge \overline{d}(m,\varepsilon), \quad j = 1,\cdots,k \\
& w_{ij}^{VWC} - w_{i,j+1}^{VWC} \ge \overline{\overline{d}}(j,\varepsilon), \quad j = 1,\cdots,k-1,\; i = 1,\cdots,m \\
& w_{ik}^{VWC} \ge \overline{\overline{d}}(m,\varepsilon), \quad i = 1,\cdots,m
\end{aligned}
\tag{4.3}
$$

Model (4.3) aims to minimize the preference score of the VWC while at the same time keeping the preference score of the VBC no less than an appropriate parameter $\gamma$, which might be selected in a range from one and the maximal possible value. Although we note that the selection of the value of $\gamma$ is flexible, we will prove by the Theorem 1 in the case of$\gamma = 1$, the model (4.3) is equivalent to the following model (see also Theorem 1 in [44]):

$$
(4.4)\qquad
\begin{aligned}
\varphi^*_{VBC} = \quad & min \sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{VWC}\; y_{ij}^{min} \\
s.t. \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{VBC}\; y_{ij}^{max} = 1 \\
& w_{ij}^{VWC} - w_{i+1,j}^{VWC} \geq \overline{d}(i,\varepsilon), \quad i = 1,\cdots,m-1, \; j = 1,\cdots,k \\
& w_{mj}^{VWC} \geq \overline{d}(m,\varepsilon), \quad j = 1,\cdots,k \\
& w_{ij}^{VWC} - w_{i,j+1}^{VWC} \geq \overline{\overline{d}}(j,\varepsilon), \quad j = 1,\cdots,k-1, \; i = 1,\cdots,m \\
& w_{ik}^{VWC} \geq \overline{\overline{d}}(m,\varepsilon), \quad i = 1,\cdots,m
\end{aligned}
$$

**4.3. Theorem.** *For $\gamma = 1$ model (4.3) and model (4.4) are equivalent.*

*Proof.* Consider the following model in the case of $\gamma = 1$:

$$
(4.5)\qquad
\begin{aligned}
\varphi^*_{VBC} = \quad & min \sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{VWC}\; y_{ij}^{min} \\
s.t. \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{VBC}\; y_{ij}^{max} \geq 1 \\
& \sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{VWC}\; y_{ij}^{r} \leq 1, \qquad r = 1,2,\ldots,n, \\
& w_{ij}^{VWC} - w_{i+1,j}^{VWC} \geq \overline{d}(i,\varepsilon), \quad i = 1,\cdots,m-1, \; j = 1,\cdots,k \\
& w_{mj}^{VWC} \geq \overline{d}(m,\varepsilon), \quad j = 1,\cdots,k \\
& w_{ij}^{VWC} - w_{i,j+1}^{VWC} \geq \overline{\overline{d}}(j,\varepsilon), \quad j = 1,\cdots,k-1, \; i = 1,\cdots,m \\
& w_{ik}^{VWC} \geq \overline{\overline{d}}(m,\varepsilon), \quad i = 1,\cdots,m
\end{aligned}
$$

Assume that $w_{ij}^{*,VWC}$to be optimal weight of model (4.5). We prove that $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{*,VWC} y_{ij}^{max} = 1$. Assume not, i.e. $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{*,VWC} y_{ij}^{max} = q > 1$ and $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{*,VWC} y_{ij}^{r} \leq 1 (r = 1,2,\ldots,n)$. Now set $w_{ij}^{**,VWC} = \frac{w_{ij}^{*,VWC}}{q}$.

Thus we have $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{**,VWC} y_{ij}^{r} < \sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{*,VWC} y_{ij}^{r} \leq 1 (r = 1,2,\ldots,n)$ and $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{**,VWC} y_{ij}^{max} = 1$. So, we have another feasible solution $w_{ij}^{**,VWC}$ with which the obtained value of objective function $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{**,VWC} y_{ij}^{min}$ is less than the assumed optimal value$\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{*,VWC} y_{ij}^{min}$. This is a contradiction and hence the first constraint is constantly binding in any optimal solution.

Since we have $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{*,VWC} y_{ij}^{max} = 1$ it follows that $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{VWC} y_{ij}^{r} \leq \sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{*,VWC} y_{ij}^{max} = 1$. Hence the constraints $\sum_{i=1}^{m}\sum_{j=1}^{k} \; w_{ij}^{VWC} y_{ij}^{r} \leq$

1 ($r = 1, 2, \ldots, n$) are redundant and can be removed from model (4.5). This completes the proof. □

Following the same logic as before, given the worst efficiency of the VWC, the following linear programming model can be used to determine the worst possible preference score of candidate $p$ under the condition that the worst possible preference score of the VWC stays unchanged:

$$(4.6)\qquad \begin{aligned} \varphi_p^* = \min \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^p\ y_{ij}^p \\ s.t. \quad & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^{VBC}\ y_{ij}^{min} = \varphi_{VWC}^* \\ & \sum_{i=1}^{m}\sum_{j=1}^{k} w_{ij}^p\ y_{ij}^r \leq 1, \qquad r = 1, 2, \ldots, n, \\ & w_{ij}^p - w_{i+1,j}^p \geq \overline{d}(i,\varepsilon), \quad i = 1, \cdots, m-1,\ j = 1, \cdots, k \\ & w_{mj}^p \geq \overline{d}(m,\varepsilon), \quad j = 1, \cdots, k \\ & w_{ij}^p - w_{i,j+1}^p \geq \overline{\overline{d}}(j,\varepsilon), \quad j = 1, \cdots, k-1,\ i = 1, \cdots, m \\ & w_{ik}^p \geq \overline{\overline{d}}(m,\varepsilon), \quad i = 1, \cdots, m \end{aligned}$$

From the above discussion it is known that voting models (4.1) and (4.2) measure the best possible preference scores of VBC and the $n$ real candidates based on VBC, while voting models (4.4)and (4.6) measure the worst possible preference scores of VWC and the $n$ real candidates based on VWC. These two distinctive efficiency assessments may lead to quite different conclusions. Therefore, there is a need to consider them together to give an overall assessment of each candidate. In order to do so, we use the following relative closeness (RC) (Wang and Luo [41]), which is widely used in the TOPSIS approach, a well-known MADM methodology.

$$(4.7)\qquad RC_p = \frac{(\varphi_p^* - \varphi_{VWC}^*)}{(\varphi_p^* - \varphi_{VWC}^*) + (\phi_{VBC}^* - \phi_p^*)}$$

It is obvious that the bigger difference between $\varphi_p^*$and $\varphi_{VWC}^*$ and the smaller difference between $\phi_{WBC}^*$ and $\phi_p^*$ mean the better performance of candidate $p$. So, the bigger $RC_p$ value, the better the performance of candidate $p$. Since the $RC$ index integrates both the best and the worst possible preference scores of each candidate, it thus provides an overall assessment for each candidate, based on which an overall ranking for the $n$ real candidates can be easily obtained.

We are in a position to give the following algorithm for overall ranking of candidates:

***Step 1.*** Solve the problems (4.1) for the VBC to obtain the optimal weights $W_{VBC}^*$ and preference score $\phi_{VBC}^*$, solve the problem (4.2) to compute the $\phi_p^*$, $p = 1, 2, \ldots,\ n$.

***Step 2.*** Solve the problems (4.4) for the VWC to obtain the optimal weights $W_{VWC}^*$and preference score $\varphi_{VWC}^*$, solve the problem (4.6) to compute the $\phi_p^*$, $p = 1, 2, \ldots,\ n$.

***Step 3.*** Calculate the relative closeness $RC_p$ of candidate $p$ using (4.7).

***Step 4.*** Select the winner candidate q according to $RC_q^* = \max_{1\leq p\leq n} RC_p$.

In this paper, it is assumed $\overline{d}_j\ (i,\ \epsilon) = \epsilon\overline{d}_i$ and $\overline{\overline{d}}_i\ (j,\ \epsilon) = \epsilon\overline{\overline{d}}_j$, in which $\epsilon$ is a sufficiently small positive value and, $\overline{d}_i$ and $\overline{\overline{d}}_j$ are the preferred values corresponding to gap $i$ of places and gap $j$ of categories, respectively. Without loss of the generality, throughout this paper, we assume $\overline{d}_i = \overline{\overline{d}}_j = 1$. We note that the choice of discriminating function

**Table 1.** Votes received by six candidates

| Candidates | First Place | Second Place | Third Place | Fourth Place |
|---|---|---|---|---|
| $A_1$ | 3 | 3 | 4 | 3 |
| $A_2$ | 4 | 5 | 5 | 2 |
| $A_3$ | 6 | 2 | 3 | 2 |
| $A_4$ | 6 | 2 | 2 | 6 |
| $A_5$ | 0 | 4 | 3 | 4 |
| $A_6$ | 1 | 4 | 3 | 3 |
| $VBC$ | 6 | 5 | 5 | 6 |
| $VWC$ | 0 | 2 | 2 | 2 |

**Table 2.** The preference scores for the six candidates

| | The best score | | The least score | | RC | | Rank | |
|---|---|---|---|---|---|---|---|---|
| Candidates | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ |
| $A_1$ | 0.8091875 | 0.81246687 | 0.032 | 0.00032 | 0.03433845 | 0.00035543 | 4 | 4 |
| $A_2$ | 1 | 1 | 0.043 | 0.00043 | 0.07699472 | 0.00082606 | 1 | 1 |
| $A_3$ | 0.8196875 | 0.81257188 | 0.038 | 0.00038 | 0.04498756 | 0.0004621 | 3 | 3 |
| $A_4$ | 0.1 | 1 | 0.04 | 0.0004 | 0.07006569 | 0.00074618 | 2 | 2 |
| $A_5$ | 0.6758125 | 0.68738313 | 0.022 | 0.00022 | 0.01416807 | 0.00014542 | 6 | 6 |
| $A_6$ | 0.6803125 | 0.68742813 | 0.025 | 0.000025 | 0.01845772 | 0.00018904 | 5 | 5 |

and also values of $\epsilon$ may be influences the results of models and it is the decision maker concern.

## 5. Numerical examples

In this section, we consider two numerical examples using the proposed method to illustrate its applications and show its capability in expressing preferences of voters on a set of candidates.

**5.1. Example.** We will examine the example taken from Cook and Kress [7], in which 20 voters are asked to rank 4 out of 6 candidates $A_1 \sim A_6$ on a ballot. The votes each candidate receives are shown in Table 1. Also the virtual candidates VBC and VWC are defined in the last two rows of Table 1. Using model (4.1), we obtain the best preference score of VBC as $\phi^*_{VBC} = 1.371625$, and $\phi^*_{VBC} = 1.37496625$ for the values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$, respectively. In a similar way, using model (4.4) we obtain the worst preference score of VWC as $\varphi^*_{VWC} = 0.012$ and $\varphi^*_{VWC} = 0.00012$ for the values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$, respectively. Based upon the optimal weights of models (4.2) and (4.6) we can calculate the best preference score and the worst preference score of each candidate as documented in Table 2. In this case, the final overall ranking order can be achieved using the systematic $RC$ index, whose values for the six candidates are presented in Table 2 for the values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$. From Table 2, the full rank of candidates is as $A_2 \succ A_4 \succ A_3 \succ A_1 \succ A_6 \succ A_5$.

Now, suppose the voters classify into two distinct categories ($C1$ and $C2$) that the vote of the first category has a greater importance than that of the second category. In this case, the voters of each category are asked to rank 4 out 6 previous candidates on a ballot. The votes each candidate receives from each category are presented in Table 3. In addition the virtual candidates VBC and VWC are defined in this table.

Using model (4.1), the best preference scores of VBC are obtained as $\phi^*_{VBC} = 1.93931579$, and $\phi^*_{VBC} = 1.94728789$ for the values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$, respectively. In addition, using model (4.4) we obtain the worst preference score of VWC as $\varphi^*_{VWC} = 0.014$ and $\varphi^*_{VWC} = 0.00014$ for the values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$,

**Table 3.** Votes received by six candidates from two categories

| | First Place | | Second Place | | Third Place | | Fourth Place | |
|---|---|---|---|---|---|---|---|---|
| Candidates | C 1 | C 2 | C 1 | C 2 | C 1 | C 2 | C 1 | C 2 |
| $A_1$ | 2 | 1 | 1 | 2 | 1 | 3 | 2 | 1 |
| $A_2$ | 1 | 3 | 1 | 4 | 3 | 2 | 2 | 0 |
| $A_3$ | 3 | 3 | 1 | 1 | 1 | 2 | 1 | 1 |
| $A_4$ | 1 | 5 | 1 | 1 | 1 | 1 | 1 | 5 |
| $A_5$ | 0 | 0 | 3 | 1 | 1 | 2 | 1 | 3 |
| $A_6$ | 1 | 0 | 1 | 3 | 1 | 2 | 1 | 2 |
| $VBC$ | 3 | 5 | 3 | 4 | 3 | 3 | 2 | 3 |
| $VWC$ | 0 | 0 | 1 | 1 | 1 | 2 | 1 | 0 |

**Table 4.** The preference scores for the six candidates based our approach

| | The best score | | The least score | | RC | | Rank | |
|---|---|---|---|---|---|---|---|---|
| Candidates | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ |
| $A_1$ | 0.91647368 | 0.92100684 | 0.038 | 0.00038 | 0.02292609 | 0.0002338 | 4 | 4 |
| $A_2$ | 1 | 1 | 0.05 | 0.0005 | 0.03691112 | 0.00037989 | 1 | 1 |
| $A_3$ | 1 | 1 | 0.044 | 0.00044 | 0.03094967 | 0.00031659 | 2 | 2 |
| $A_4$ | 0.994 | 0.99994 | 0.044 | 0.00044 | 0.03075927 | 0.00031657 | 3 | 3 |
| $A_5$ | 0.80610526 | 0.81569263 | 0.027 | 0.00027 | 0.01134172 | 0.00011487 | 5 | 5 |
| $A_6$ | 0.73057895 | 0.73677947 | 0.029 | 0.00029 | 0.01225754 | 0.0001239 | 6 | 6 |

respectively. Based upon the optimal weights of models (4.2) and (4.6) we can determine the best preference score and the worst preference score of each candidate as documented in Table 4. Thus, the final overall ranking order can be obtained using the systematic $RC$ index, whose values for the six candidates are presented in Table 4 for the values of $\epsilon$ : $\epsilon = 0.001$ and $\epsilon = 0.00001$. From Table 4, the full rank of candidates is as $A_2 \succ A_3 \succ A_4 \succ A_1 \succ A_6 \succ A_5$. As can be seen from Table 4, our model also identifies the candidate $A_2$ as the first winner when $\epsilon = 0.001$ and $\epsilon = 0.00001$. Moreover, by considering the systematic $RC$ index of candidates $A_3$ and $A_4$, we see the candidate $A_3$ is more efficient than the candidate $A_4$. That is, in our opinion the candidate $A_3$ is the second winner and the candidate $A_4$ is the third winner. Thus, there is a different in rank of second and third winner candidate comparing with models that assume all votes have a same importance. In fact, the ability to identify efficient candidates based on our approach is stronger than the previous approach.

**5.2. Example.** We will examine the example taken from Wang et al. [40], in which 155 voters are asked to rank 4 out of 10 candidates $A \sim J$ on a ballot. The votes each candidate receives are shown in Table 5. In addition, the virtual candidates VBC and VWC are defined in the last two rows of Table 5.

The model (4.1) gives the best preference score of VBC as $\phi^*_{VBC} = 1.28020626$, and $\phi^*_{VBC} = 1.28314864$ under two different values of $\epsilon = 0.001$ and $\epsilon = 0.00001$, respectively. In a similar way, model (4.4) gives the worst preference score of VWC as $\varphi^*_{VWC} = 0.54628571$ and $\varphi^*_{VWC} = 0.53582$ when $\epsilon$ takes 0.001 and 0.00001, respectively. Based upon the optimal weights of models (4.2) and (4.6) we can calculate the best preference score and the worst preference score of each candidate as reported in Table 6. Thus, the total ranking order can be determined using the systematic $RC$ index, whose values for the ten candidates are given in Table 6 when $\epsilon$ takes 0.001 and 0.00001. From Table 6, the full rank of candidates is as $G \succ A \succ E \succ I \succ J \succ C \succ B \succ H \succ D \succ F$.

**Table 5.** Votes received by ten candidates

| Candidates | First Place | Second Place | Third Place | Fourth Place |
|---|---|---|---|---|
| $A$ | 20 | 14 | 13 | 11 |
| $B$ | 14 | 16 | 16 | 17 |
| $C$ | 14 | 14 | 19 | 21 |
| $D$ | 14 | 13 | 22 | 11 |
| $E$ | 19 | 14 | 12 | 19 |
| $F$ | 14 | 13 | 9 | 11 |
| $G$ | 18 | 17 | 15 | 9 |
| $H$ | 14 | 13 | 20 | 20 |
| $I$ | 14 | 20 | 15 | 20 |
| $J$ | 14 | 21 | 14 | 16 |
| $VBC$ | 20 | 21 | 22 | 21 |
| $VWC$ | 14 | 13 | 9 | 9 |

**Table 6.** The preference scores for ten candidates

| | The best score | | The least score | | RC | | Rank | |
|---|---|---|---|---|---|---|---|---|
| Candidates | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ |
| $A$ | 0.99414793 | 0.99408349 | 0.69328571 | 0.67479 | 0.33944619 | 0.3246696 | 2 | 2 |
| $B$ | 0.93709298 | 0.93911734 | 0.57728571 | 0.53613 | 0.0828626 | 0.00090027 | 7 | 7 |
| $C$ | 0.97228487 | 0.97712775 | 0.58128571 | 0.53617 | 0.10206421 | 0.00114241 | 6 | 6 |
| $D$ | 0.94033136 | 0.94668142 | 0.57428571 | 0.5361 | 0.07611283 | 0.00083148 | 9 | 9 |
| $E$ | 1 | 1 | 0.67957143 | 0.65501 | 0.32234172 | 0.29624299 | 3 | 3 |
| $F$ | 0.76114455 | 0.7599445 | 0.54828571 | 0.53584 | 0.00383832 | 0.00003822 | 10 | 10 |
| $G$ | 1 | 1 | 0.7075102 | 0.68896286 | 0.36523167 | 0.35101041 | 1 | 1 |
| $H$ | 0.9654328 | 0.97120124 | 0.57928571 | 0.53615 | 0.09488936 | 0.00105675 | 8 | 8 |
| $I$ | 1 | 1 | 0.59028571 | 0.53626 | 0.13571607 | 0.00155154 | 4 | 4 |
| $J$ | 0.97759172 | 0.97634396 | 0.58728571 | 0.53623 | 0.11931975 | 0.00133457 | 5 | 5 |

**Table 7.** Votes received by six candidates from two categories

| | First Place | | | Second Place | | | Third Place | | | Fourth Place | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Candidates | C 1 | C 2 | C 3 | C 1 | C 2 | C 3 | C 1 | C 2 | C 3 | C 1 | C 2 | C 3 |
| $A$ | 5 | 11 | 4 | 5 | 8 | 1 | 4 | 4 | 5 | 2 | 2 | 7 |
| $B$ | 2 | 9 | 3 | 4 | 9 | 3 | 4 | 8 | 4 | 4 | 11 | 2 |
| $C$ | 2 | 12 | 0 | 3 | 8 | 3 | 11 | 4 | 4 | 5 | 10 | 6 |
| $D$ | 3 | 7 | 4 | 4 | 2 | 7 | 5 | 10 | 7 | 7 | 2 | 2 |
| $E$ | 1 | 13 | 5 | 12 | 1 | 1 | 3 | 1 | 8 | 13 | 2 | 4 |
| $F$ | 1 | 13 | 0 | 3 | 4 | 6 | 2 | 4 | 3 | 2 | 2 | 7 |
| $G$ | 5 | 1 | 12 | 1 | 1 | 15 | 13 | 1 | 1 | 1 | 7 | 1 |
| $H$ | 1 | 4 | 9 | 3 | 1 | 9 | 5 | 11 | 4 | 5 | 5 | 10 |
| $I$ | 1 | 2 | 11 | 5 | 2 | 13 | 3 | 3 | 9 | 16 | 3 | 1 |
| $J$ | 2 | 2 | 10 | 7 | 5 | 9 | 1 | 2 | 11 | 4 | 11 | 1 |
| $VBC$ | 5 | 13 | 12 | 12 | 9 | 15 | 13 | 11 | 11 | 16 | 11 | 10 |
| $VWC$ | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 |

Now we suppose the 155 voters are divided into three categories ($C1$, $C2$ and $C3$) based on their priorities and proficiencies. The votes each candidate receives from each category are shown in Table 7. Also the virtual candidates VBC and VWC are defined in this table.

**Table 8.** The preference scores for the ten candidates based on our approach

| | The best score | | The least score | | RC | | Rank | |
|---|---|---|---|---|---|---|---|---|
| Candidates | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ | $\epsilon = 0.001$ | $\epsilon = 0.00001$ |
| $A$ | 1 | 1 | 0.327 | 0.20127 | 0.19162271 | 0.11214341 | 1 | 1 |
| $B$ | 0.96464852 | 0.96810428 | 0.26055556 | 0.09060556 | 0.143802 | 0.02125852 | 8 | 6 |
| $C$ | 1 | 1 | 0.28533333 | 0.13485333 | 0.1647688 | 0.06008753 | 3 | 3 |
| $D$ | 0.94407853 | 0.95292174 | 0.27133333 | 0.13471333 | 0.14863913 | 0.05757814 | 5 | 4 |
| $E$ | 0.9702937 | 0.06859667 | 0.25966667 | 0.96515685 | 0.14384231 | 0.0016434 | 6 | 8 |
| $F$ | 0.71743712 | 0.06800667 | 0.20066667 | 0.71535221 | 0.08401168 | 0.00090331 | 10 | 10 |
| $G$ | 0.99470197 | 1 | 0.26766667 | 0.11817667 | 0.15218444 | 0.04604364 | 4 | 5 |
| $H$ | 0.966497831 | 0.98307375 | 0.23566667 | 0.06835667 | 0.12680534 | 0.00144938 | 9 | 9 |
| $I$ | 1 | 1 | 0.25466667 | 0.06854667 | 0.14383597 | 0.00165028 | 7 | 7 |
| $J$ | 1 | 1 | 0.28933333 | 0.13489333 | 0.16742395 | 0.06012072 | 2 | 2 |

Using model (4.1), the best preference scores of VBC are obtained as $\phi^*_{VBC} = 2.04761608$, and $\phi^*_{VBC} = 2.06472518$ for the values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$, respectively. In addition, using model (4.4) we obtain the worst preference score of VWC as $\varphi^*_{VWC} = 0.07866667$ and $\varphi^*_{VWC} = 0.06678667$ for the values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$, respectively. Based upon the optimal weights of models (4.2) and (4.6) we can determine the best preference score and the worst preference score of each candidate as documented in Table 8. Then, the final overall ranking order can be obtained using the systematic $RC$ index, whose values for the ten candidates are presented in Table 8 for the values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$. From Table 8, when $\epsilon$ takes 0.001, the full rank of candidates is obtained as $A \succ J \succ C \succ G \succ D \succ E \succ I \succ B \succ H \succ F$.

As can be seen from Table 8, there is a different in total rank based on our approach comparing with that approach which assumes all votes have equal importance. Our method identifies the candidates $A$ as the first winner and the candidate G as the forth winner while that approach identifies the candidate A as the second winner and the candidate G as the first winner. However, different from that approach, our approach considers the priority of voters and so the votes in a higher category have more importance than that in a lower category. Thus, the preference scores are measured in a persuasive way.

It is necessary to notice that as we discussed in the end of Section 4, the value of $\epsilon$ may be influences the order of candidates. This point has been illustrated in Example 5.2. From Table 8, when $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$, candidate B is the eighth winner and sixth winner, candidate D is the fifth winner and fourth winner, candidate E is the sixth winner and eighth winner and candidate D is the fourth winner and fifth winner, respectively. This means that there is a small difference in the rank of candidates B, D, E and G when $\epsilon$ varies. However, as can be seen from Table 8, candidates A, C, F, H, I and G should take the first place, the third place, the tenth place, the ninth place, the seventh place and the second place, respectively under the both values of $\epsilon$. This is, based on two different values of $\epsilon: \ \epsilon = 0.001$ and $\epsilon = 0.00001$, candidates A and F should be the first winner and the last winner, respectively.

## 6. Conclusion

It is often necessary in decision making framework to rank a group of candidates in voting systems. In ranked voting systems, each voter selects a subset of candidates and rank them from most to least preferred and hence the score obtained by each candidate is the weighted sum of the scores receives in different places. The principal drawback of such scoring rules is that they assume the votes of all voters have equal importance and there is no preference among them. In this paper, we generalized the existing scoring rules to overcome the mentioned drawback. The ability to identify efficient candidates of our approach is stronger than the existing scoring rules. We also introduced two models that the first model evaluated candidates from the viewpoint of the best possible preference score and the second model evaluated them from the perspective of the worst possible preference score. The two distinctive scores have been combined to form a comprehensive index such that an overall ranking for all the candidates can be obtained. Finally we illustrated our method with two examples. In addition, the extension of some other ranking methods to rank of decision making units in DEA framework such as the proposed approach by Golam Abri et al. [14] can be interesting for ranking of efficient candidates in voting systems as a research work.

In our opinion, we feel that there are many other ranking methods in DEA and should be considered for voting systems later on. Some of these methods are discussed below.

(1) Ramazani-Tarkhorani et al. [30] obtained a common set of weights (CSW) to create the best efficiency score of a group composed of efficient units in DEA. Development of their method for ranking of efficient candidates in voting systems may also produce interesting results.
(2) Jahanshaloo et al. [23] defined an ideal line determined a CSW for efficient units in DEA and then a new efficiency score obtained and ranked them with it. In the second method, they introduced a special line and then compared all efficient units with it and ranked them. Extending of these two methods can be effective for ranking of effective candidates in voting systems.
(3) Jahanshaloo et al. [24] presented a new super-efficient method to rank all decision-making units using the TOPSIS method. Development of this method for ranking of all candidates in voting systems may also give interesting results.
(4) Amirteimoori and Kordrostami [3] proposed a super-efficiency DEA model to discriminate the performance of efficient decision making units. How to apply this model to develop a more general model with sound mathematical properties in ranking of efficient candidates is a direction for future research.

## Acknowledgements

The authors would like to thank the anonymous reviewers and the editor for their insightful comments and suggestions.

## References


[1] Amin, G.R. and Sadeghi, H. *Application of prioritized aggregation operators in preference voting*, International Journal of Intelligent System **25**, 1027-1034, 2010.

[2] Amirteimoori, A. and Kordrostami, S.*A Euclidean distance-based measure of efficiency in data envelopment analysis*, Optimization: A Journal of Mathematical Programming and Operations Research **59** (7), 985-996, 2010.

[3] Amirteimoori, A. and Kordrostami, S. *An alternative clustering approach: a DEA-based procedure*, Optimization: A Journal of Mathematical Programming and Operations Research **62** (2), 227-240, 2013.

[4] Andersen, P. and Petersen, N.C. *A procedure for ranking efficient units in data envelopment analysis*, Management Science **39** (10), 1261-1264, 1993.

[5] Brams, S.J. and Fishburn, P.C. *Voting procedures*, In: Arrow, K.J., Sen, A.K. and Suzumura, K. (Eds.), Handbook of Social, Choice and Welfare, Vol. 1, Elsevier, Amsterdam, pp 173-236, 2002.

[6] Bystrichy, R. *Different approaches to weighted voting systems based on preferential positions*, Kybernetika **48** (3), 536-549, 2012.

[7] Cook, W.D. and Kress, M. *A data envelopment model for aggregating preference rankings*, Management Science **36** (11) 1302-1310, 1990.

[8] Ebrahimnejad, A. *A new approach for ranking of candidates in voting systems*, OPSEARCH, **49**, 103-115, 2012.

[9] Ebrahimnejad, A. *Ranking of candidates in the preferential voting framework based on a new approach*, Recent Researches in Applied Mathematics and Economics, Proceedings of the 6th International Conference on Applied Mathematics, Simulation, Modelling, March 7-9, Athens, Greece, 78-81, 2012.

[10] Ebrahimnejad, A., Shahverdi, R., Rezai Balf, F. and Hatefi, M. *Finding target unit in FDH model by least distance measure model*, Kybernetika **49** (4), 619-635, 2013.

[11] Foroughi, A.A. and Aouni, B. *New approaches for determining a common set of weights for a voting system*, International Transactions in Operational Research **19**, 521-530, 2012.

[12] Foroughi, A. A, Jones, D.F. and Tamiz, M. *A selection method for a preferential election*, Applied Mathematics and Computation **163** 107-116, 2005.

[13] Foroughi, A.A. and Tamiz, M. *An effective total ranking model for a ranked voting system*, Omega **33**, 491-496, 2005.

[14] Gholam Abri, A., Jahanshahloo G.R., Hosseinzadeh Lotfi F., Shoja, N., and Fallah Jelodar M. *A new method for ranking non-extreme efficient units in data envelopment analysis*, Optimization Letters **7** (2), 309-324, 2013.

[15] Green. R.H., Doyle, J.R. and Cook, W.D. *Preference voting and project ranking using DEA and cross-evaluation*, European Journal of operational Research **90**, 461-472, 1996.

[16] Hadi-Vencheh, A. and Mokhtarian, M. N. *Three new models for preference voting and aggregation: A note*, Journal of the Operational Research Society **60**, 1036-1037, 2009.

[17] Hadi-Vencheh, A. *Two effective total ranking models for preference voting and aggregation*, Mathematical Sciences **8** (115), 1-4, 2014.

[18] Hashimoto, A. *A ranked voting system using a DEA/AR exclusion model: A note*, European Journal of operational Research **97**, 600-604, 1997.

[19] Hosseinzadeh Lotfi, F., Hatami-Marbini, A., Agrell, P.J., Aghayi, N. and Gholami, K. *Allocating fixed resources and setting targets using a common-weights DEA approach*, Computers and Industrial Engineering **64**, 631-640, 2013.

[20] Hosseinzadeh Lotfi, F. and Fallahnejad, R. *A note on A solution method to the problem proposed by Wang in voting systems*, Applied Mathematical Sciences, **5**, 3051-3055, 2011.

[21] Hosseinzadeh Lotfi, F., Rostamy-Malkhalifeh, M., Aghayi, N., Ghelej Beigi, Z. and Gholami, K. *An improved method for ranking alternatives in multiple criteria decision analysis*. Applied Mathematical Modelling **37**, 25-33, 2013.

[22] Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Khanmohammadi, M., Kazemimanesh, M. *A method for discriminating efficient candidates with ranked voting data by common weights*, Mathematical and Computational Applications **17**, 1-8, 2012.

[23] Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Khanmohammadi, M., Kazemimanesh, M. and Rezaie, V. *Ranking of units by positive ideal DMU with common weights*, Expert Systems with Applications **37** 7483-7488, 2010.

[24] Jahanshahloo, G.R., Khodabakhshi, M., Hosseinzadeh Lotfi, F. and Moazami Goudarzi. M.R. *A cross-efficiency model based on super-efficiency for ranking units through the TOPSIS approach and its extension to the interval case*, Mathematical and Computer Modelling **53**, 1946-1955, 2011.

[25] Kordrostami, S., Bakhoda Bijarkani, Z. *A new method to measure efficiency in parallel production systems with shared sourced in natural life*, International Journal of Biomathematics **6** (6), 1-18, 2013.

[26] Llamazares, B. and Pena, T. *Preference aggregation and DEA: An analysis of the methods proposed to discriminate efficient candidates*, European Journal of Operational Research **197**, 714-721, 2009.

[27] Llamazares, B. and Pena, T. *Aggregating preferences rankings with variable weights*, European Journal of Operational Research **230**, 348-355, 2013.

[28] Noguchi, H., Ogawa, M. and Ishii, H. *The appropriate total ranking method using DEA for multiple categorized purposes*, Journal of Computational and Applied Mathematics **146** (1), 155-166, 2002.

[29] Obata, T. and Ishii, H. *A method for discriminating efficient candidates with ranked voting data*, European Journal of Operational Research **151**, 233-237, 2003.

[30] Ramazani-Tarkhorani, S., Khodabakhshi, M., Mehrabian, S. and Nuri-Bahmani, F. *Ranking decision making units with Common Weights in DEA*, Applied Mathematical Modelling **38** (15-16),3890-3896, 2014.

[31] Saati, S., Hatami-Marbini, A., Agrell, P.J., Tavana, M. *A common set of weight approach using an ideal decision making unit in data envelopment analysis*, Journal of Industrial and Management Optimization **8** (3), 623-637, 2012.

[32] Sexton, T.R., Silkman, R.H. and Hogan, A.J. *Data envelopment Analysis: Critique and extensions*, In: Silkman R H (1986) (Ed.), Measuring Efficiency: An Assessment of Data Envelopment Analysis, Jossey-bass, San Francisco, pp 73-105, 1986.

[33] Shokouhi, A.H., Hatami-Marbini, A., Tavana, M., Saati, S. *A robust optimization approach for imprecise data envelopment analysis*, Computers and Industrial Engineering **59**,387-397, 2010.

[34] Soltanifar, M., Ebrahimnejad, A., and Farrokhi, M. M. *Ranking of different ranking models using a voting model and its application in determining efficient candidates*, International Journal of Society Systems Science., **2**, 375-389, 2010.

[35] Stein, W.E., Mizzi, P.J. and Pfaffenberger, R.C. *A stochastic dominance analysis of ranked voting systems with scoring*, European Journal of Operational Research **74** (1), 78-85, 1994.

[36] Tavana, M., Hatami-Marbini, A. *A group AHP-TOPSIS framework for human space flight mission planning at NASA*, Expert Systems with Applications **38**, 13588-13603, 2011.

[37] Tavana, M., Khalili-Damaghani, K., Sadi-Nezhad, S. *A fuzzy group data envelopment analysis model for high-technology project selection: A case study at NASA*, Computers and Industrial Engineering **66** (1), 10-23, 2013.

[38] Tavana, M., Mirzagholtabar, H., Mirhedayatian, S.M., Farzipoor Saen, R., Azadi, M. *A new network epsilon-based DEA model for supply chain performance evaluation*, Computers and Industrial Engineering **66** (2), 501-513, 2013.

[39] Wang, Y.M. and Chin, K.S. *Discriminating DEA efficient candidates by considering their least relative total scores*, Journal of Computational and Applied Mathematics **206**, 209-215, 2007.

[40] Wang, N., Yi, R. and Liu, D. *A solution method to the problem proposed by Wang in voting systems*, Journal of Computational and Applied Mathematics **221**, 106-113, 2008.

[41] Wang, Y.M. and Luo, Y. *DEA efficiency assessment using ideal and anti-ideal decision making units*, Applied Mathematics and Computation **173** (2), 902-915, 2006.

[42] Wang, Y.M., Chin, K.S. and Yang, J.B. *Three new models for preference voting and aggregation*, Journal of the Operational Research Society **58**, 1389-1393, 2007.

[43] Wang, Y.M., Chin, K.S. and Yang, J.B, *Improved linear programming models for preference voting and aggregation: Reply to Hadi-Vencheh and Mokhtarian*, Journal of the Operational Research Society **58**, 1037-1038, 2009.

[44] Wu, D. *A note on DEA efficiency assessment using ideal point: An improvement of Wang and Luo's model*, Applied Mathematics and Computation **183**, 819-830, 2006.

[45] Wu, J., Liang, L. and Zha, Y. *Preference voting and ranking using DEA game cross efficiency model*, Journal of the Operations Research Society of Japan **52**, 105-111, 2009.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 181 – 193

# Reliability analysis under constant-stress partially accelerated life tests using hybrid censored data from Weibull distribution

Ali A. Ismail* †

**Abstract**

This article discusses the estimation of Weibull distribution parameters based on hybrid censored data under constant-stress partially accelerated test model. Two estimation methods; maximum likelihood (ML) and percentile bootstrap (PB) are used to make statistical inference on the Weibull distribution parameters and the acceleration factor. The mean square errors of the estimators are calculated to evaluate their performances through a Monte Carlo simulation study. Moreover, the confidence intervals lengths (CILs) and their associated coverage probabilities (CPs) are obtained. Finally, to demonstrate the proposed methodology, an arithmetic example is given.

**Keywords:** Statistics; reliability; percentile bootstrap; confidence interval; coverage rate; hybrid censoring; mean square error.

*2000 AMS Classification:* 62N01, 62N05.

*Received* : 16.09.2014 *Accepted* : 28.01.2015 *Doi* : 10.15672/HJMS.20157211089

## 1. Introduction

The ordinary life testing methods of high reliability products usually need a long period to gain sufficient failure data required to do inferences. So, to perform reliability analysis, accelerated life tests (ALTs) are the most common ways to measure such products' life. Under such test settings, products are tested at higher-than-usual levels of stress to induce failures rapidly and economically. Applying ALTs depends on a life-stress relationship. The parameters of life can be estimated via this relationship by using the failure data obtained under accelerated conditions. However, in some cases such a relationship can't be known or assumed. Thus, ALTs can't be applied and the partially accelerated life tests (PALTs) come to be a good appliance to implement the needed life tests.

*King Saud University, College of Science, Department of Statistics and Operations Research, P. O. BOX 2455, Riyadh 11451, Saudi Arabia. E-mail: ismail@ksu.edu.sa

†Cairo University, Faculty of Economics & Political Science, Department of Statistics, Giza 12613, Egypt. E-mail: ali.ismail@cu.edu.eg

The stress can be used in different techniques, frequently applied techniques are constant-stress and step-stress. Under step-stress PALTs, a test unit is first run at use condition and, if it does not fail for a definite time, then it is run at accelerated condition until failure happens or the observation is stopped. But the constant-stress PALTs run each unit at either use condition or accelerated condition only, i.e. each unit is run at a constant-stress level until the test is finished. Accelerated stresses include higher than normal temperature, power, pressure, load, etc., for more details see Nelson [28].

In this article, we deal with hybrid censored constant-stress PALTs when the lifetime of test unit follows Weibull distribution. PALTs have been considered under Type-I and Type-II censoring schemes by numerous authors, for example, see Goel [18], DeGroot and Goel [12], Bhattcharyya and Soejoeti [9], Bai and Chung [7], Bai et al. [8], Abdel-Ghaly et al. [1], Abdel-Ghaly et al. [2], Abdel-Ghaly et al. [3], Abdel-Ghani [4], Abdel-Ghani [5], Ismail [21], Aly and Ismail [6], Ismail and Sarhan [25], Ismail et al. [20] and Ismail [23].

In general, accelerated tests are frequently ended before all items fail. The estimates from the censored data are less precise than those from complete data. However, this is more than offset by the reduced test time and cost. The most used censoring schemes are Type-I and Type-II censoring. Consider n units placed on life test. In traditional Type-I censoring, the experiment lasts up to a pre-specified time $C_1$. Any failures that happen after that time are not witnessed. The end point $C_1$ of the experiment is supposed to be s-independent of the failure times. But in traditional Type-II censoring, the experimenter finishes the experiment after a pre-identified number of units $R \leq n$ fail. In this situation, only the lowest lifetimes are noticed. In Type-I censoring, the number of failures witnessed is random and the endpoint of the experiment is fixed. But in failure-censoring $R$ is fixed and the termination time is random. Several previous works have considered the reliability analysis using the traditional time- and failure- censoring schemes under different life distributions, for more details one can see Cohen [11].

Concerning hybrid censoring scheme it can be applied as follows. Consider a life testing experiment in which n units are placed on test concurrently. Failure times are noticed and the test is finished either at a pre-specified time $C_1$ or based on a pre-determined number R of failures acquired by a time; say $C_2$ whichever is happened first. Such a combination of Type-I and Type-II censoring schemes is identified as hybrid censoring scheme. So, sampling according to the hybrid censoring scheme is finished at $min\,(C_1, C_2)$. It is noted that the traditional time- and failure- censoring schemes can be found as special cases of hybrid censoring scheme by taking $R = n$ and $C_1 = \infty$, respectively. The most important benefit of applying hybrid censoring scheme is that it preserves the probable experiment time and cost. Several authors have discussed the statistical inference problem about the parameters for sampling schemes Type-I and Type-II censoring. In this work the estimation of parameters is studied under constant-stress partially accelerated life tests (CSPALTs) with hybrid censored data supposing Weibull distribution. It is also supposed that the failed items are not exchanged.

Although the hybrid censoring scheme is applicable, most of preceding works under PALTs were studied using the usual time- and failure- censoring schemes and no consideration has been provided in examining hybrid censored data. All papers prepared under hybrid censoring were correlated with ordinary or fully accelerated tests, see, for example, Fairbanks et al. [17], Draper and Guttman [14], Chen and Bhattacharyya [10], Ebrahimi [15], Gupta and Kundu [19], Kundu [26], Xie [31], Park and Balakrishnan [29] and Zhang et al. [32]. Recently, only two papers made by Ismail [22] and Ismail [24] have considered the hybrid censoring scheme under step-stress PALTs.

The rest of the paper is structured as follows. In Section 2 the model and the hybrid censored data are designated. The maximum likelihood (ML) and percentile bootstrap

(PB) estimations of the CSPALTs model parameters are considered in Section 3. Section 4 covers the simulation results. Section 5 presents an illustrative example. Conclusion is yielded in Section 6.

## 2. Model description

In this study, it is assumed that the lifetime of a test unit say $X$ under normal condition has Weibull distribution with probability density function (PDF) given by

$$f(x;\beta,\eta)=\frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1}e^{-(x/\eta)^{\beta}},x>0,\beta>0,\eta>0, \tag{2.1}$$

In fact, Weibull distribution has high flexibility compared to other distributions. Its failure rate function can be increasing, decreasing and constant according to the value of the shape parameter. For more information, see Dimitri [13].

The survival function of this distribution is given by

$$R(x)=e^{-(x/\eta)^{\beta}}, \tag{2.2}$$

.

The matching failure rate function is

$$h(x)=\frac{\beta}{\eta}(\frac{x}{\eta})^{\beta-1}. \tag{2.3}$$

Constant-stress PALTs can be processed according to the following steps and assumptions.

(1) $n_1$ units randomly selected among $n$ test units sampled are assigned to run under normal stress and $n_2$ $(=n-n_1)$ items are allotted to run under severe stress.
(2) Each item is tested until the censoring times $C_1$ or $C_2$ is realized whichever is smaller or the item fails.
(3) The lifetimes $X_i$, $i=1,...,n_1$ of units consigned to normal (use) stress, are i.i.d. r.v.'s.
(4) The lifetimes $Y_j$, $j=1,...,n_2$ of units assigned to severe stress, are i.i.d r.v.'s.

Now, for a unit subjected to accelerated condition, the PDF of its lifetime say $Y$ is provided by

$$f(y;\lambda,\beta,\eta)=\frac{\lambda\beta}{\eta}(\frac{\lambda y}{\eta})^{\beta-1}e^{-(\lambda y/\eta)^{\beta}},y>0,\lambda>1,\beta>0,\eta>0, \tag{2.4}$$

where $Y=\lambda^{-1}X$ and $\lambda$ is the acceleration factor.

Because the test in Type-I censoring is finished when a pre-specified time $C_1$ is attained and in failure-censoring the test is ended based on a pre-defined number $R$ of failures gained by a time $C_2$; say. Accordingly, the failure times $x_{(??)}\leq ...\leq x_{(}n_{u)}\leq C_1$ (or $C_2$) and $y_{(??)}\leq ...\leq y_{(}n_{a)}\leq C_1$ (or $C_2$) are ordered failure times at use and accelerated conditions respectively, where $n_u$ $(<n_1)$ and $n_a$ $(<n_2)$ are the numbers of items failed at use and accelerated conditions, respectively.

Under hybrid censoring scheme, supposing that $R$ and $C_1$ are predetermined, we can observe the following data.

If $C_1 < C_2$, the sample is $x_{(??)} \leq ... \leq x_{(n_u)} \leq C_1$ and $y_{(??)} \leq ... \leq y_{(n_a)} \leq C_1$.

If not, the sample is $x_{(??)} \leq ... \leq x_{(n_u)} \leq C_2$ and $y_{(??)} \leq ... \leq y_{(n_a)} \leq C_2$.

## 3. Estimation process

Here, the maximum likelihood estimates (MLEs) of the CSPALTs model parameters under hybrid censoring as well as their confidence limits are considered.

**3.1. ML point estimation.** Now, let us define the indicator functions: $\delta_{ui} \equiv \mathrm{I}(X_i \leq C_1$ (or $C_2$)) and $\delta_{aj} \equiv \mathrm{I}(Y_j \leq C_1$ (or $C_2$)). Then the total likelihood function for $(x_1;\delta_{u1}, \ldots, x_{n1};\delta_{un1}, y_1;\delta_{a1}, \ldots, y_{n2};\delta_{an2})$ under CSPALTs is given by

$$\begin{aligned} L(\underline{x}, \underline{y}|\lambda, \beta, \eta) &= \prod_{i=1}^{n_1} L_{ui}(x_i, \delta_{ui}|\beta, \eta) \,.\, \prod_{j=1}^{n_2} L_{aj}(y_j, \delta_{aj}|\lambda, \beta, \eta) \\ &= \prod_{i=1}^{n_1} [\frac{\beta}{\eta}\left(\frac{x_i}{\eta}\right)^{\beta-1} \exp\{-\,(x_i/\eta)^{\beta}\}]^{\delta_{ui}} [\exp\{-\,(C_1^{\delta_{1R}}\, C_2^{\delta_{2R}}/\eta)^{\beta}\}]^{\bar{\delta}_{ui}} \\ \times & \prod_{j=1}^{n_2} [\frac{\lambda\beta}{\eta}\left(\frac{\lambda y_j}{\eta}\right)^{\beta-1} \exp\{-\,(\lambda y_j/\eta)^{\beta}\}]^{\delta_{aj}} [\exp\{-\,(\lambda\, C_1^{\delta_{1R}}\, C_2^{\delta_{2R}}/\eta)^{\beta}\}]^{\bar{\delta}_{aj}}, \end{aligned}$$

(3.1)

where,
$L_{ui}$ and $L_{ai}$ denote the contributions of the items $i$, $i = 1, \ldots, n_1$ and $j$, $j = 1, \ldots, n_2$ to the total likelihood function under use and accelerated conditions, respectively; and
$\bar{\delta}_{ui} = 1 - \delta_{ui}$, $\bar{\delta}_{aj} = 1 - \delta_{aj}$, $\delta_{1R}$=1 if $C_2 > C_1$ and 0 otherwise, and $\delta_{2R}$=1 if $C_2 < C_1$ and 0 otherwise.

The value of $\hat{\eta}$ can be found by

$$\hat{\eta} = \left\{\frac{\psi}{n_u + n_a}\right\}^{\frac{1}{\beta}}, \tag{3.2}$$

where

$$\psi = \sum_{i=1}^{n_1} \delta_{ui}\, x_i^{\beta} + \lambda^{\beta} + \sum_{j=1}^{n_2} \delta_{aj}\, y_j^{\beta} + (C_1^{\delta_{1R}}\, C_2^{\delta_{2R}})^{\beta}(n_1 - n_u) + (\lambda\, C_1^{\delta_{1R}}\, C_2^{\delta_{2R}})^{\beta}(n_2 - n_a)\,.$$

Now, we have two ML non-linear equations which can be extracted as follows.

$$\frac{n_a\, \hat{\beta}}{\hat{\lambda}} - [\frac{(n_u + n_a)\, \hat{\beta}\, \hat{\lambda}^{\hat{\beta}-1}}{\psi}]\, [\sum_{j=1}^{n_2} \delta_{aj}\, y_j^{\hat{\beta}} + (C_1^{\delta_{1R}}\, C_2^{\delta_{2R}})^{\hat{\beta}}(n_2 - n_a)] = 0, \tag{3.3}$$

$$\frac{n_u + n_a}{\hat{\beta}} + \sum_{i=1}^{n_1} \delta_{ui} \ln x_i + \sum_{j=1}^{n_2} \delta_{aj} \ln y_j - (n_u + n_a) \ln\left(\frac{\psi}{n_u + n_a}\right)^{1/\hat{\beta}} + n_a \ln\hat{\lambda} + \hat{\beta}\left(\frac{n_u + n_a}{\psi}\right)^{1/\hat{\beta}} = 0. \tag{3.4}$$

From equation (3.3), the value of $\hat{\lambda}$ is easily determined by the following formula.

$$\hat{\lambda} = \left\{ \frac{n_a[\sum_{i=1}^{n_1} \delta_{ui}\, x_i^{\hat{\beta}} + (C_1^{\delta_{1R}}\ C_2^{\delta_{2R}})^{\hat{\beta}}(n_1 - n_u)]}{n_u\ [\sum_{j=1}^{n_2} \delta_{aj}\, y_j^{\hat{\beta}}\ + (C_1^{\delta_{1R}}\ C_2^{\delta_{2R}})^{\hat{\beta}}(n_2 - n_a)\ ]} \right\}^{\frac{1}{\hat{\beta}}}. \tag{3.5}$$

After substituting for $\hat{\lambda}$, the equation (3.4) can be expressed by

$$\begin{aligned} & \frac{n_u + n_a}{\hat{\beta}} + \sum_{i=1}^{n_1} \delta_{ui} \ln x_i + \sum_{j=1}^{n_2} \delta_{aj} \ln y_j \\ & - n_u\, \frac{\sum_{i=1}^{n_1} \delta_{ui}\, x_i^{\hat{\beta}} \ln\ x_i + (C_1^{\delta_{1R}}\ C_2^{\delta_{2R}})^{\hat{\beta}}(n_1 - n_u) \ln\ (C_1^{\delta_{1R}}\ C_2^{\delta_{2R}})}{\sum_{i=1}^{n_1} \delta_{ui}\, x_i^{\hat{\beta}}\ + (C_1^{\delta_{1R}}\ C_2^{\delta_{2R}})^{\hat{\beta}}(n_1 - n_u)} \\ & - n_a\, \frac{\sum_{j=1}^{n_2} \delta_{aj}\, y_j^{\hat{\beta}} \ln\ y_j + (C_1^{\delta_{1R}}\ C_2^{\delta_{2R}})^{\hat{\beta}}(n_2 - n_a) \ln\ (C_1^{\delta_{1R}}\ C_2^{\delta_{2R}})}{\sum_{j=1}^{n_2} \delta_{aj}\, y_j^{\hat{\beta}}\ + (C_1^{\delta_{1R}}\ C_2^{\delta_{2R}})^{\hat{\beta}}(n_2 - n_a)} = 0. \end{aligned} \tag{3.6}$$

To get the value of $\hat{\beta}$, the Newton-Raphson method is utilized to solve the non-linear equation (3.6), numerically. Consequently, based on the value of $\hat{\beta}$, the values of $\hat{\eta}$ and $\hat{\lambda}$ can be simply determined from (3.2) and (3.5) respectively.

**3.2. ML interval estimation.** In this subsection, the approximate confidence bounds of the parameters are obtained based on the asymptotic distribution of the MLEs of the elements of the vector of unknown parameters $\Omega = (\beta,\ \eta,\ \lambda)$. It is known that the asymptotic distribution of the MLEs of $\Omega$ is given by; see Miller [27],

$$((\hat{\beta}-\beta),\ (\hat{\eta}-\eta),\ (\hat{\lambda}-\lambda)) \rightarrow\ N\ (0,\ \mathbf{I}^{-1}(\beta,\ \eta,\ \lambda))$$

where $\mathbf{I}^{-1}(\beta, \eta, \lambda)$ is the variance-covariance matrix of the unknown parameters $\Omega = (\beta,\ \eta,\ \lambda)$. The elements of the $3 \times 3$ matrix $\mathbf{I}^{-1}$ , $I_{ij}\ (\beta,\ \eta,\ \lambda)$, i, j = 1, 2, 3 ; can be approximated by $I_{ij}(\hat{\beta}, \hat{\eta}, \hat{\lambda})$, where

$$\mathrm{I}_{ij}(\hat{\Omega}) = -\frac{\partial^2 lnL(\Omega)}{\partial\Omega_i \partial\Omega_j}\ |_{\Omega=\hat{\Omega}}$$

Thus, the approximate 100(1 - $\gamma$)% two sided confidence intervals of $\beta$, $\eta$ and $\lambda$ are, respectively, yielded by

$$\pm Z_{\gamma/2}\ \sqrt{\mathrm{I}_{11}^{-1}(\hat{\beta}, \hat{\eta}, \hat{\lambda})},\ \ \hat{\eta} \pm Z_{\gamma/2}\ \sqrt{\mathrm{I}_{22}^{-1}(\hat{\beta}, \hat{\eta}, \hat{\lambda})} \text{ and } \hat{\lambda} \pm Z_{\gamma/2}\ \sqrt{\mathrm{I}_{33}^{-1}(\hat{\beta}, \hat{\eta}, \hat{\lambda})}.$$

where $Z_{\gamma/2}$ is the upper $(\gamma/2)$th percentile of a standard normal distribution.

**3.3. Percentile bootstrap estimation.** In this section, we use a parametric bootstrap method to construct CIs for the unknown parameters $\beta$, $\eta$ and $\lambda$. The bootstrap is a re-sampling technique for statistical inference. It is frequently used to estimate CIs. Also, it can be used to estimate bias and variance of an estimator. It has the advantage of computational ease especially for large sample sizes. We present the percentile bootstrap CIs (PBCIs) proposed by Efron [16]. The following steps can be proceeded to obtain bootstrap samples for the proposed method.

(1) Using the original hybrid censored sample, $x_{(??)} \leq ... \leq x_{(n_u)} \leq C_1$ and $y_{(??)} \leq ... \leq y_{(n_a)} \leq C_1$ if $C_1 < C_2$ or $x_{(??)} \leq ... \leq x_{(n_u)} \leq C_2$ and $y_{(??)} \leq ... \leq y_{(n_a)} \leq C_2$ if $C_2 < C_1$, obtain $\hat{\beta}, \hat{\eta}$ and $\hat{\lambda}$.

(2) Using the values of $n_1$ and $n_2$, generate two independent samples of sizes $n_1$ and $n_2$ from Weibull distribution, $\underline{x}^* = (x_1^* < x_2^* < ... < x_{n_1}^*)$ and $\underline{y}^* = (y_1^* < y_2^* < ... < y_{n_2}^*)$.

(3) As in step 1 based on $\underline{x}^*$ and $\underline{y}^*$ compute the bootstrap sample estimates of $\hat{\beta}, \hat{\eta}$ and $\hat{\lambda}$ say, $\hat{\beta}^*, \hat{\eta}^*$ and $\hat{\lambda}^*$.

(4) Repeat the above steps 2 and 3 $N$ (=10,000) times representing $N$ different bootstrap samples.

(5) Arrange all $\hat{\beta}^*, \hat{\eta}^*$ and $\hat{\lambda}^*$in an ascending order to obtain the bootstrap sample $\hat{\varphi}_\ell^{*[1]}, \hat{\varphi}_\ell^{*[2]}, ..., \hat{\varphi}_\ell^{*[N]}$, $\ell = 1, 2, 3$, where $\varphi_1^* = \beta^*, \varphi_2^* = \eta^*$ and $\varphi_3^* = \lambda^*$.

To obtain PBCIs, let $G(z) = P(\hat{\varphi}_\ell^* \leq z)$ be CDF of $\hat{\varphi}_\ell^*$. Define $\hat{\varphi}_{\ell boot}^* = G^{-1}(z)$ for given $z$. The approximate bootstrap100$(1-\gamma)$%CI of $\hat{\varphi}_\ell^*$ is given by
$(^*_{\ell boot}(\frac{\gamma}{2}),\ \hat{\varphi}_{\ell boot}^*(1-\frac{\gamma}{2}))$.

## 4. Simulation studies

In this section simulation studies are made to evaluate the performances of the MLEs in terms of their mean square errors (MSEs) for various choices of $n$, $R$ and $C_1$ values. Also, the 95 % asymptotic confidence bounds based on the asymptotic distribution of the MLEs are constructed and their lengths are computed and presented with the associated coverage probabilities (CPs). For different hybrid censored data sets, the average values of the MSEs, confidence interval lengths (CILs) and CPs are calculated using 10,000 replications and the results are given in Tables 1-6. In each Table, the odd rows represent the results of the ML estimation for $\beta$, $\eta$ and $\lambda$, respectively, while the even ones denote the results of the percentile bootstrap estimation (between brackets) for the three parameters respectively.

From Tables 1-6 some notes can be discovered concerning the two approaches as follows.

(1) For fixed $n$ and $R$, the MSEs decrease as $C_1$ increases.
(2) For fixed $n$ and $C_1$, the MSEs decrease as $R$ increases.
(3) For fixed $R$ and $C_1$, the MSEs decrease as $n$ increases.
(4) For fixed $R$ and $C_1$, the CILs decrease as $n$ increases.
(5) For fixed $n$ and $C_1$, the CILs decrease as $R$ increases.
(6) For fixed $n$ and $R$, the CILs decrease as $C_1$ increases.

Also, we observed that the computed CPs of the confidence bounds for each parameter are very close to the nominal level as $n$ increases. The same pattern is noticed as $R$ or $C_1$ increases. That is, the procedure is successfully working.

Now, when we compare between the two methods of estimation, it is observed that for relatively small and moderate sample sizes, percentile bootstrap method works better than the ML method. It provides smaller MSEs, narrower CILs with closest CPs to the nominal level. The method of bootstrap is recommended to use even for large samples for computational ease and high precision.

Moreover, point and 95 % confidence interval estimations for the survival function at mission times 3, 5, 7 and 10 are obtained using the two methods of estimation. The estimations of the true survival are calculated via the following expressions:

(x)=exp{$-$ $(x/\hat{\eta})^{\hat{\beta}}$}, for items run under use condition,

or

(y)=exp{$-$ $(\hat{\lambda}y/\hat{\eta})^{\hat{\beta}}$}, for items run under accelerated condition.

As Soliman [30] shows, "estimation of the reliability function of some equipment is one of the main problems of reliability theory. In most practical applications and life-test experiments, the distributions with positive domain, e.g., Weibull, Burr-XII, Pareto, Beta, and Rayleigh, are quite appropriate models".

The estimation results of the true survival function are introduced in Tables 7 and 8. It can be observed that the percentile bootstrap method gives reliability estimations better than the ML method.

**Table 1:** The results of MSEs, CILs and CPs using the methods of ML and percentile bootstrap (between brackets), respectively, with true parameter values set at $\beta$ = 1.5, $\eta$ = 2 and $\lambda$ = 2.5 when $C_1$ = 20 and $n$ = 25 ($n_1$=12, $n_2$=13).

| R = 10 | R = 15 | R = 20 |
|---|---|---|
| 0.038, 1.235, 0.947 | 0.017, 0.992, 0.948 | 0.007, 0.851, 0.952 |
| (0.026), (1.127), (0.948) | (0.014), (0.842), (0.949) | (0.004), (0.762), (0.951) |
| 0.047, 1.786, 0.945 | 0.023, 1.415, 0.946 | 0.012, 1.194, 0.948 |
| (0.031), (1.514), (0.946) | (0.018), (1.217), (0.947) | (0.009), (0.988), (0949) |
| 0.054, 1.911, 0.943 | 0.037, 1.549, 0.944 | 0.029, 1.352, 0.946 |
| (0.041), (1.817), (0.944) | (0.031), (1.311),(0.946) | (0.018), (1.132), (0.948) |

**Table 2:** The results of MSEs, CILs and CPs using the methods of ML and percentile bootstrap (between brackets), respectively, with true parameter values set at $\beta$ = 1.5, $\eta$ = 2 and $\lambda$ = 2.5 when $C_1$ = 20 and $n$ = 35 ($n_1$=17, $n_2$=18).

| R = 15 | R = 20 | R = 25 |
|---|---|---|
| 0.011, 0.715, 0.953 | 0.006, 0.661, 0.950 | 0.002, 0.526, 0.950 |
| (0.008), (0.689), (0.951) | (0.004), (0.541), (0.950) | (0.001), (0.418), (0.950) |
| 0.019, 1.218, 0.947 | 0.011, 0.907, 0.949 | 0.007, 0.789, 0.951 |
| (0.015), (1.115), (0.948) | (0.007), (0.833), (0.950) | (0.003), (0.640), (0.950) |
| 0.033, 1.355, 0.945 | 0.021, 1.141, 0.948 | 0.013, 0.911, 0.951 |
| (0.026), (1.172), (0.947) | (0.013), (0.917), (0.949) | (0.009), (0.763), (0.950) |

**Table 3:** The results of MSEs, CILs and CPs using the methods of ML and percentile bootstrap (between brackets), respectively, with true parameter values set at $\beta$ = 0.5, $\eta$ = 0.7 and $\lambda$ = 3 when $C_1$ = 30 and $n$ = 25 ($n_1$=12, $n_2$=13).

| R = 10 | R = 15 | R = 20 |
|---|---|---|
| 0.021, 1.015, 0.948 | 0.015, 0.910, 0.949 | 0.005, 0.754, 0.951 |
| (0.016), (0.985), (0.949) | (0.011), (0.852), (0.950) | (0.003), (0.669), (0.950) |
| 0.032, 1.487, 0.946 | 0.019, 1.321, 0.947 | 0.011, 0.988, 0.952 |
| (0.024), (1.311), (0.948) | (0.0014), (1.103), (0.949) | (0.005), (0.901), (0.951) |
| 0.041, 1.802, 0.945 | 0.027, 1.463, 0.945 | 0.018, 1.076, 0.954 |
| (0.036), (1.587), (0.947) | (0.021), (1.298), (0.947) | (0.0012), (0.992), (0.952) |

**Table 4:** The results of MSEs, CILs and CPs using the methods of ML and percentile bootstrap (between brackets), respectively, with true parameter values set at $\beta$ = 0.5, $\eta$ = 0.7 and $\lambda$ = 3 when $C_1$ = 30 and $n$ = 35 ($n_1$=17, $n_2$=18).

| R = 15 | R = 20 | R = 25 |
|---|---|---|
| 0.006, 0.411, 0.950<br>(0.004), (0.392), (0.950)<br>0.010, 0.762, 0.949<br>(0.007), (0.611), (0.949)<br>0.024, 1.117, 0.948<br>(0.019), (0.996), (0.949) | 0.004, 0.286, 0.950<br>(0.002), (0.203), (0.950)<br>0.007, 0.531, 0.951<br>(0.004), (0.498), (0.950)<br>0.012, 0.820, 0.952<br>(0.008), (0.758), (0.951) | 0.001, 0.197, 0.950<br>(0.002), (0.163), (0.950) 0.003,<br>0.312, 0.951<br>(0.002), (0.277), (0.950)<br>0.008, 0.524, 0.951<br>(0.005), (0.469), (0.950) |

**Table 5:** The results of MSEs, CILs and CPs using the methods of ML and percentile bootstrap (between brackets), respectively, with true parameter values set at $\beta$ = 1.5, $\eta$ = 0.7 and $\lambda$ = 3 when $C_1$ = 35 and $n$ = 50 ($n_1$=20, $n_2$=30).

| R = 15 | R = 20 | R = 25 |
|---|---|---|
| 0.003, 0.397, 0.950<br>(0.002), (0.364), (0.950)<br>0.007, 0.748, 0.949<br>(0.006), (0.711), (0.950)<br>0.019, 1.001, 0.949<br>(0.014), (0.964), (0.950) | 0.002, 0.254, 0.950<br>(0.001), (0.226), (0.950)<br>0.005, 0.503, 0.950<br>(0.003), (0.489), (0.950) 0.007,<br>0.611, 0.950<br>(0.004), (0.522), (0.950) | 0.001, 0.182, 0.950<br>(0.001), (0.165), (0.950) 0.002,<br>0.292, 0.950<br>(0.001), (0.203), (0.950) 0.003,<br>0.479, 0.950<br>(0.002), (0.445), (0.950) |

**Table 6:** The results of MSEs, CILs and CPs using the methods of ML and percentile bootstrap (between brackets), respectively, with true parameter values set at $\beta$ = 0.5, $\eta$ = 0.7 and $\lambda$ = 3 when $C_1$ = 35 and $n$ = 50 ($n_1$=20, $n_2$=30).

| R = 15 | R = 20 | R = 25 |
|---|---|---|
| 0.004, 0.402, 0.950<br>(0.003), (0.387), (0.950)<br>0.008, 0.751, 0.949<br>(0.006), (0.620), (0.950)<br>0.020, 1.004, 0.949<br>(0.016), (0.981), (0.950) | 0.003, 0.261, 0.950<br>(0.002), (0.207), (0.950)<br>0.006, 0.508, 0.950<br>(0.004), (0.433), (0.950) 0.009,<br>0.627, 0.951<br>(0.004), (0.489), (0.950) | 0.001, 0.182, 0.950<br>(0.001), (0.147), (0.950) 0.003,<br>0.304, 0.950<br>(0.002), (0.287), (0.950) 0.004,<br>0.481, 0.950<br>(0.003), (0.423), (0.950) |

**Table 7:** Average values of point and interval estimations for the survival function at different mission times 3, 5, 7 and 10 according to the methods of ML and percentile bootstrap (between brackets), respectively, with true parameter values set at $\beta$ = 1.5, $\eta$ = 0.7 and $\lambda$ = 3 when $C_1$ = 35 and $n$ = 50 ($n_1$=20, $n_2$=30).

| R = 15 | R = 20 | R = 25 |
|---|---|---|
| 0.746, 0.688, 0.879<br>(0.797), (0.708, 0.893) | 0.767, 0.722, 0.896<br>(0.812), (0.742, 0.898) | 0.782, 0.744, 0.920<br>(0.831), (0.758, 0.932) |
| 0.731, 0.678, 0.864<br>(0.782), (0.711, 0.883) | 0.752, 0. 694, 0.887<br>(0.801), (0.734, 0.883) | 0.771, 0.721, 0.895<br>(0.822), (0.744, 0.912) |
| 0.694, 0.647, 0.822<br>(0.741), (0.683, 0.852) | 0.746, 0.688, 0.873<br>(0.789), (0.723, 0.876) | 0.766, 0.706, 0.889<br>(0.804), (0.723, 0.896) |
| 0.658, 0.625, 0.794<br>(0.719), (0.667, 0.814) | 0.728, 0.671, 0.868<br>(0.763), (0.692, 0.871) | 0.748, 0.697, 0.874<br>(0.781), (0.711, 0.883) |

**Table 8:** Average values of point and interval estimations for the survival function at different mission times 3, 5, 7 and 10 according to the methods of ML and percentile bootstrap (between brackets), respectively, with true parameter values set at $\beta$ = 0.5, $\eta$ = 0.7 and $\lambda$ = 3 when $C_1$ = 35 and $n$ = 50 ($n_1$=20, $n_2$=30).

| R = 15 | R = 20 | R = 25 |
|---|---|---|
| 0.741, 0.682, 0.873<br>(0.789), (0.702, 0.882) | 0.760, 0.704, 0.882<br>(0.802), (0.728, 0.891) | 0.773, 0.728, 0.887<br>(0.825), (0.746, 0.904) |
| 0.728, 0.673, 0.859<br>(0.776), (0.694, 0.871) | 0.747, 0. 694, 0.870<br>(0.788), (0.709, 0.879) | 0.762, 0.694, 0.877<br>(0.815), (0.738, 0.894) |
| 0.685, 0.638, 0.804<br>(0.733), (0.675, 0.838) | 0.738, 0.676, 0.856<br>(0.778), (0.687, 0.864) | 0.743, 0.682, 0.863<br>(0.791), (0.692, 0.871) |
| 0.643, 0.611, 0.781<br>(0.692), (0.642, 0.798) | 0.713, 0.664, 0.843<br>(0.748), (0.663, 0.832) | 0.728, 0.672, 0.841<br>(0.767), (0.687, 0.849) |

## 5. A demonstrative example

To demonstrate the proposed methodology, a demonstrative example via hybrid censored data set from Weibull distribution is considered. We use $n$ = 75 ($n_1$=25, $n_2$=50), $\beta$ = 2, $\eta$ = 2.5 and $\lambda$ = 3 when $C_1$ = 40 and $R$ = 20. The number of failures observed at use and accelerated conditions are $n_u$=11 and $n_a$=39, respectively, with censored items $n_c$=25. The MSEs associated with the MLEs of the parameters $\beta$, $\eta$ and $\lambda$ are 0.002, 0.003 and 0.005, respectively, while those associated with the percentile bootstrap estimation are respectively 0.001, 0.002 and 0.004. In addition, a 95% CILs of the model parameters $\beta$, $\eta$ and $\lambda$ using the two approaches ML and PB are 0.241, 0.462, 0.581 and 0.212, 0.409, 0.523, respectively. Moreover, the CPs associated with ML and PB are respectively 0.948, 0.947, 0.949 and 0.950, 0.951, 0.950. Finally, the point and interval estimations for the survival function at a mission time 6 according to the methods of ML and PB (between brackets) are, respectively, 0.749, 0.691, 0.883 and (0.795), (0.734, 0.887).

## 6. Conclusion

In this article, the likelihood and percentile bootstrap estimation methods has been applied to the CSPALTS model parameters assuming Weibull distribution under hybrid censoring. The performance of the estimators has been examined in terms of their MSEs via simulation studies for the two methods of estimation. Also, the CILs of the model parameters have been obtained as well as their CPs. It is observed that for small and moderate sample sizes, percentile bootstrap method works better than the approximate method. It provides smaller MSEs, narrower CILs with closest CPs to the nominal level. The method of bootstrap is recommended to use even for large samples for computational ease and high precision. Finally, an illustrative example has been given.

## Acknowledgment

This project was supported by King Saud University, Deanship of Scientific Research, College of Science Research Center.

## References

[1] Abdel-Ghaly, A.A., El-Khodary, E.H. and Ismail, A.A. (2002). Maximum Likelihood Estimation and Optimal Design in Step Partially Accelerated Life Tests for Pareto Distribution Using Type-I Censoring, *Proceedings of the 14th Annual Conference of Statistics and Computer Modeling in Human and Social Sciences*, Faculty of Economics and Political Sciences, Cairo, University, 27-40.

[2] Abdel-Ghaly, A.A., El-Khodary, E.H. and Ismail, A.A. (2003a). Estimation and optimum constant-stress partially accelerated life test plans for Pareto distribution with Type-I censoring, *Proceedings of the 38th Annual Conference of Statistics, Computer Sciences and Operation Research*, ISSR, Cairo University, Egypt, 49-65.

[3] Abdel-Ghaly, A.A., El-Khodary, E.H. and Ismail, A.A. (2003b). Estimation and Optimal Design in Step Partially Accelerated Life Tests for Pareto Distribution using Type-II Censoring, *Proceedings of the 15th annual conference on Statistics and Computer Modeling in Human and Social Sciences*, Faculty of Economics and Political Science, Cairo University, Egypt, 16-29.

[4] Abdel-Ghani, M.M. *Investigation of some lifetime models under partially accelerated life tests*, Ph.D. Thesis, Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt, 1998.

[5] Abdel-Ghani, M.M. *The Estimation Problem of The Log-Logistic Parameters in Step Partially Accelerated Life Tests Using Type-I Censored Data*, National Rev. Social Sci., **41** (2), 1-19, 2004.

[6] Aly, H.M. and Ismail, A.A. *Optimum simple time-step stress plans for partially accelerated life testing with censoring*, Far East J.Theor. Stat. **24** (2), 175-200, 2008.

[7] Bai, D.S. and Chung, S.W. *Optimal design of partially accelerated life tests for the Exponential distribution under type-I censoring*, IEEE Trans. Reliab. **41** (3), 400-406, 1992.

[8] Bai, D.S., Chung, S.W. and Chun, Y.R. *Optimal design of partially accelerated life tests for the lognormal distribution under type-I censoring*, Reliab. Eng. Syst. Safety **40**, 85-92, 1993.

[9] Bhattacharyya, G.K. and Soejoeti, Z.A *tampered failure rate model for step-stress accelerated life test*, Commun. Stat.—Theory Methods **18** (5), 1627-1643, 1989.

[10] Chen, S. and Bhattacharya, G.K. *Exact confidence bounds for an exponential parameter under hybrid censoring*, Commun. Stat.—Theory Methods **17**, 1857-1870, 1988.

[11] Cohen, A.C. *Truncated and Censored Samples: Theory and Applications* (Marcel Dekker, Inc., New York, 1991).

[12] DeGroot, M.H. and Goel, P.K. *Bayesian and optimal design in partially accelerated life testing*, Naval Res. Logist. Quart. **16** (2), 223-235, 1979.

[13] Dimitri, K. *Reliability Engineering Handbook,* Vol. 1 (Prentice Hall, Englewood Cliffs, NJ., 1991).

[14] Draper, N. and Guttman, I. *Bayesian analysis of hybrid life tests with exponential failure times*, Ann. Inst. Stat. Math. **39**, 219-225, 1987.

[15] Ebrahimi, N. *Prediction intervals for future failures in exponential distribution under hybrid censoring*, IEEE Transactions on Reliability **41**, 127-132, 1992.

[16] Efron, B. *The jackknife, the bootstrap and other resampling plans* (Philadelphia: SIAM, 1982).

[17] Fairbanks, K., Madson, R. and Dykstra, R. *A confidence interval for an exponential parameter from a hybrid life test*, J. Am. Stat. Assoc. **77**, 137-140, 1982.

[18] Goel, P.K. Some estimation problems in the study of tampered random variables. Technical Report No. 50, Department of statistics, Carnegie-mellon university, Pittspurgh, Pennsylvania, 1971.

[19] Gupta, R.D. and Kundu, D. *Hybrid censoring schemes with exponential failure distribution*, Commun. Stat.—Theory Methods **27** (12), 3065-3083, 1998.

[20] Ismail, A.A, Abdel-Ghaly, A.A. and El-Khodary, E.H. *Optimum constant-stress life test plans for Pareto distribution under type-I censoring*, J. Stat. Comput. Simul. **81** (12), 1835–1845, 2011.

[21] Ismail, A.A. *The Test Design and Parameter Estimation of Pareto Lifetime Distribution under Partially Accelerated Life Tests*, Ph.D. Thesis, Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt, 2004.

[22] Ismail, A.A. *Estimating the parameters of Weibull distribution and the acceleration factor from hybrid partially accelerated life test*, Appl. Math. Model. **36**, 2920-2925, 2012.

[23] Ismail, A.A. On *designing step-stress partially accelerated life tests under failure-censoring scheme*. Proc. IMechE, Part O: J. Risk Reliab. **227** (6), 662–670, 2013.

[24] Ismail, A.A. *Inference for a step-stress partially accelerated life test model with an adaptive Type-II progressively hybrid censored data from Weibull distribution*, J. Comput. Appl. Math. **260**, 533-542, 2014.

[25] Ismail, A.A. and Sarhan, A.M. *Optimal planning time-step stress life test using progressively type-II censored exponential data*, Int. Math. Forum **4** (40), 1963–1976, 2009.

[26] Kundu, D. *On hybrid censored Weibull distribution*, J. Stat. Plan. Infer. **137**, 2127-2142, 2007.

[27] Miller, R.C. *Survival Analysis*, (Wiley, New York, 1981).

[28] Nelson, W. *Accelerated Life Testing: Statistical Models, Data Analysis and Test Plans*, (Wiley, New York, 1990).

[29] Park, S. and Balakrishnan, N. *On simple calculation of the Fisher information in hybrid censoring schemes*, Stat. Probab. Lett. **79** (10), 1311-1319, 2009.

[30] Soliman, A.A. *Reliability estimation in a generalized life-model with application to the Burr-XII*, IEEE Trans. Reliab. **51** (3), 337-343, 2002.

[31] Xie, Q. Exact inference for exponential step-stress model under different forms of censoring. ETD Paper AAINR28250, Collection for McMaster University, http://digitalcommons.mcmaster.ca/dissertaions/AAINR28250, 2006.

[32] Zhang S., Zhang Y., Chaloner K. and Stapleton J.T. *A copula model for bivariate hybrid censored survival data with application to the MACS study*, Lifetime Data Anal. **16** (2), 231-249, 2010.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 195 – 206

# Probability for transition of business cycle and pricing of options with correlated credit risk

Geonwoo Kim*†

**Abstract**

In this paper we propose the transition probability of business cycle for the pricing of options with credit risk. In order to describe business cycles of markets, the regime switching model is considered. We provide the probability density functions of the occupation time of the high volatility regime via Laplace transforms. Using these functions we derive the analytic valuation formulae for options with correlated credit risk and business cycle. We also illustrate the important properties of options with numerical graphs.

**Keywords:** Business cycle, Option pricing, Credit risk, Occupation time.

*2000 AMS Classification:* 60K20, 91B70.

*Received* : 02.12.2014 *Accepted* : 29.01.2015 *Doi* : 10.15672/HJMS.2018499356

## 1. Introduction

In this paper we study the business cycle model for valuing options with credit risk. It is assumed that the financial event occurs at some time in the market. This should lead to the transition of volatilities of both the underlying stock and the option issuer's asset. The financial events are often modeled by the regime switching model to capture the changes of the market environment by the unanticipated events (see, e.g., Hamilton [8], Bollen [2], Buffington and Ellott [4], Boyle and Draviam [3], Zhang et al. [15], Zhu et al. [16], Elliott et al. [7]). Based on this approach, we model the business cycle by a continuous-time two-state regime switching.

The traditional option pricing based on Black-Scholes model [1] has been used the assumption that options have no default risk. However, there exists the default risk of the option writers in the over-the-counter (OTC) markets. OTC markets have grown rapidly in size in recent years. That is, in the OTC markets, the counterparty default risk is very important and should be considered for pricing of options.

*Department of Mathematical Sciences, Seoul National University, Seoul 151-747, Republic of Korea, Email: geonwoo@snu.ac.kr

†Corresponding Author.

Johnson and Stulz [10] proposed the valuation of options with credit risk, which is called *Vulnerable option*. In their model, the options depend on the liabilities of the option issuers. If the default of the counterparty occurs at the maturity, the option holder takes all assets of the counterparty. Their model also considers the correlation between the option issuer's asset and the underlying asset. Klein [11] developed the result of Johnson and Stulz [10] by allowing for the proportional recovery of nominal claims in default. Klein and Inglis [12] dealt with options with credit risk employing the stochastic interest model of Vasicek [14]. Hui et al. [9] extended a vulnerable option valuation model that incorporates a stochastic default barrier which reflects the expected leverage level of the option issuer. Chang and Hung [5] provided analytic formulae to evaluate vulnerable American options under the assumptions of Klein's model. In the recent study, Shiu. et al [13] proposed a closed-form approximation for valuing European basket warrants with credit risk. However, none of the studies consider options with credit risk under the varying market environment.

The rest of the paper is organized as follows. Section 2 presents the business cycle modeling by using regime switching. In particular, we provide the probability density function of the occupation time of high volatility in a given time period. Section 3 gives the formulae for the arbitrage-free price of options with credit risk as integral under our model. Finally, we provide the numerical examples with various graphs to show the properties of option prices in section 4.

## 2. The model

We assume that a given filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, Q)$ satisfies the usual conditions, where $Q$ presents a risk neutral measure[‡] and the filtration $\{\mathcal{F}_t\}$ is generated by Brownian motions and two independent Poisson point processes. Based on the settings of Klein [11], we model the correlated evolutions of the option issuer's asset value process $V_t$ and the underlying stock process $S_t$ as the following:

$$dS_t = rS_tdt + \sigma_1(t)S_tdW_t^1, \tag{2.1}$$

$$dV_t = rV_tdt + \sigma_2(t)V_tdW_t^2, \tag{2.2}$$

where $r$ is a riskless interest rate, $\sigma_i(t), (i = 1, 2)$ are the time-varying volatilities of each process and $W_t^i, (i = 1, 2)$ are standard Brownian motions under a risk neutral measure $Q$ with correlation $\rho$. Here, we model the business cycle by the volatilities with two regimes.

We refer to two regimes as the high volatility and the low volatility. The high volatility region presents the economic contraction period when the market is stressed by some financial event. On the other hands, the low volatility region presents the economic expansion period, where the market has the stable economic environment. For modeling these, we assume that $\sigma_1(t)$ and $\sigma_2(t)$ are governed by two independent Poisson point processes $\mathcal{P}_1$ and $\mathcal{P}_0$ with a two state continuous-time Markov chain.

Let $\mathcal{P}_1$ and $\mathcal{P}_0$ be two independent Poisson point processes with intensity $\lambda_1$ and $\lambda_0$, respectively. If we are in the high regime, issuer's asset's volatility is $\sigma_1 + \delta_1$. We observe the high volatility Poisson point processes $\mathcal{P}_1$. If we get a signal from this high volatility point processes, issuer's asset's volatility is changed from $\sigma_1 + \delta_1$ to $\sigma_1$. If we are in the low regime, issuer's asset's volatility is $\sigma_1$. We observe the low volatility Poisson point processes $\mathcal{P}_0$. If we get a signal from this low volatility point processes, issuer's asset's volatility is changed from $\sigma_1$ to $\sigma_1 + \delta_1$. Surely, the volatility $\sigma_2(t)$ of underlying stock is

‡Elliott et al. [6] show the existence of an equivalent martingale measure in the regime switching model. So, we can get the risk-neutral valuation under our model.

affected by the same Poisson point processes as well. By $\mathcal{P}_0$ and $\mathcal{P}_1$, the volatility $\sigma_2(t)$ moves between $\sigma_2$ and $\sigma_2 + \delta_2$.

Let $U_t$ be the occupation time in the high regime from 0 to option's maturity $T$. Then it is defined by

$$U_t := \int_0^t \varepsilon(s)ds \tag{2.3}$$

$$\text{with} \quad \varepsilon(t) = \begin{cases} 0 & \text{Economic expansion regime} \quad (\sigma_i(t) = \sigma_i), \\ 1 & \text{Economic contraction regime} \quad (\sigma_i(t) = \sigma_i + \delta_i), \end{cases} \quad i = 1, 2,$$

where $\varepsilon(t)$ is the random variable with two regimes 0 (= *Low volatility*) and 1 (= *High volatility*). The following Proposition gives the probability density function of $U_t$ conditioned on $\varepsilon(0)$.

**2.1. Proposition.** *For a given time $T$, the probability density functions of $U_t$ conditioned on $\varepsilon(0)$ are given by*

$$\begin{aligned} &P(U_t = u|\varepsilon(0) = 1) := f_1(u;T) = e^{-\lambda_1 T}\delta_0(T-u) + \lambda_1 e^{-(\lambda_1-\lambda_0)u-\lambda_0 T} \\ &\times[{}_0F_1(2;\lambda_0\lambda_1 u(T-u))\lambda_0 u + {}_0F_1(1;\lambda_0\lambda_1 u(T-u))], \qquad 0 < u < T \end{aligned} \tag{2.4}$$

$$\begin{aligned} &P(U_t = u|\varepsilon(0) = 0) := f_0(u;T) = e^{-\lambda_0 T}\delta_0(u) + \lambda_0 e^{-(\lambda_1-\lambda_0)u-\lambda_0 T} \\ &\times[{}_0F_1(1;\lambda_0\lambda_1 u(T-u)) + {}_0F_1(1;\lambda_0\lambda_1 u(T-u))\lambda_1 - \lambda_1], \quad 0 < u < T \end{aligned} \tag{2.5}$$

*where ${}_0F_1(a;z)$ is the generalized hypergeometric function defined by*

$${}_0F_1(a;z) = \sum_{n=0}^{\infty} \frac{1}{(a)_n}\frac{z^n}{n!},$$

*with the rising factorial $(a)_0 = 1$ and $(a)_n = a(a+1)\cdots(a+n-1)$. And*

$$\delta_x(y) := \begin{cases} 1, & \text{if} \quad x = y, \\ 0, & \text{if} \quad x \neq y. \end{cases}$$

*Proof.* Let $f_j(u;T)$ be the probability density function of $U_t$ over $[0,T]$. Then, by the Laplace transform,

$$m_j(r;T) := \mathrm{E}[e^{-rU_T}|\varepsilon(0) = j] = \mathcal{L}_r(f_j(\cdot;T)). \tag{2.6}$$

We also consider the two cases $\tau_j > T$ and $\tau_j < T$, where $\tau_j$ is the random time of the leaving state $j$ satisfying $P(\tau_j > t) = e^{-\lambda_j t}$ for each state $j \in \{0,1\}$. We then have

$$m_1(r;T) = e^{-rT}e^{-\lambda_1 T} + \int_0^T e^{-\lambda_1 u}\lambda_1 m_0(r;T-u)e^{-ru}du, \tag{2.7}$$

$$m_0(r;T) = e^{-\lambda_0 T} + \int_0^T e^{-\lambda_0 u}\lambda_0 m_1(r;T-u)du. \tag{2.8}$$

Taking the Laplace transform of the above equations gives

$$\widehat{m}_j(r;s) := \mathcal{L}_s(m_j(r;\cdot)) = \mathcal{L}_s[\mathcal{L}_r(f_j(\cdot;T))(r;\cdot)]. \tag{2.9}$$

Then we have

$$\widehat{m}_1(r;s) = \frac{s+\lambda_0+\lambda_1}{rs + r\lambda_0 + s^2 + s\lambda_1 + s\lambda_0}, \tag{2.10}$$

$$\widehat{m}_0(r;s) = \frac{r+s+\lambda_0+\lambda_1}{rs + r\lambda_0 + s^2 + s\lambda_1 + s\lambda_0}. \tag{2.11}$$

The equation (2.10) is equal to

$$
\begin{aligned}
& \int_0^\infty e^{-rx}\frac{s+\lambda_0+\lambda_1}{s+\lambda_0}e^{-\frac{s(s+\lambda_0+\lambda_1)}{s+\lambda_0}x}dx \\
= & \int_0^\infty e^{-rx}e^{-(s+\lambda_1)x}\left(1+\frac{\lambda_1}{s+\lambda_0}\right)\sum_{n=0}^\infty\left(\frac{\lambda_0\lambda_1 x}{s+\lambda_0}\right)^n\frac{1}{n!}dx \\
= & \int_0^\infty e^{-rx}\int_0^\infty e^{-sy}e^{-\lambda_1 x} \\
& \times\left(\delta_x(y)+\sum_{n=1}^\infty\frac{(x\lambda_0\lambda_1)^n e^{-\lambda_0(y-x)}(y-x)^{n-1}}{n!(n-1)!}1_{\{x<y\}}\right)dydx \\
& +\int_0^\infty e^{-rx}\int_0^\infty e^{-sy}e^{-\lambda_1 x}\sum_{n=0}^\infty\frac{\lambda_1(x\lambda_0\lambda_1)^n e^{-\lambda_0(y-x)}(y-x)^n}{n!n!}1_{\{x<y\}}dydx \\
= & \int_0^\infty e^{-rx}\int_0^\infty e^{-sy}\left(e^{-\lambda_1 x}\delta_x(y)+\lambda_1 e^{-(\lambda_1-\lambda_0)x-\lambda_0 y}\right. \\
& \left.\times\left[{}_0F_1(2;\lambda_0\lambda_1 x(y-x))\lambda_0 x+{}_0F_1(1;\lambda_0\lambda_1 x(y-x))\right]1_{\{x<y\}}\right)dydx.
\end{aligned}
$$

Substituting $(u,T)$ for $(x,y)$ yields the equation (2.4). Similarly, from the equation (2.11), we have

$$
\begin{aligned}
& \int_0^\infty e^{-rx}\left(\frac{\delta_0(x)}{s+\lambda_0}+\frac{\lambda_0(s+\lambda_0+\lambda_1)}{(s+\lambda_0)^2}e^{-\frac{s(s+\lambda_0+\lambda_1)}{s+\lambda_0}x}dx\right)dx \\
= & \int_0^\infty e^{-rx}\delta_0(x)\int_0^\infty e^{-sy}e^{-\lambda_0 y}dydx \\
& +\int_0^\infty e^{-rx}\int_0^\infty e^{-sy}e^{-\lambda_1 x}\sum_{n=0}^\infty\frac{\lambda_0(x\lambda_0\lambda_1)^n e^{-\lambda_0(y-x)}(y-x)^n}{n!n!}1_{\{x<y\}}dydx \\
& +\int_0^\infty e^{-rx}\int_0^\infty e^{-sy}e^{-\lambda_1 x} \\
& \times\left(\sum_{n=0}^\infty\frac{\lambda_0\lambda_1(x\lambda_0\lambda_1)^n e^{-\lambda_0(y-x)}(y-x)^n}{n!n!}-\lambda_0\lambda_1 e^{-\lambda_0(y-x)}\right)1_{\{x<y\}}dydx \\
= & \int_0^\infty e^{-rx}\int_0^\infty e^{-sy}\left(e^{-\lambda_0 y}\delta_0(x)+\lambda_0 e^{-(\lambda_1-\lambda_0)x-\lambda_0 y}\left[{}_0F_1(1;\lambda_0\lambda_1 x(y-x))\right.\right. \\
& \left.\left.+{}_0F_1(1;\lambda_0\lambda_1 x(y-x))\lambda_1-\lambda_1\right]1_{\{x<y\}}\right)dydx.
\end{aligned}
$$

In a same way, substituting $(u,T)$ for $(x,y)$ in above equation completes the proof. $\square$

For given $U_t=u$ we also can obtain the following solutions of equation (2.1) and equation (2.2), respectively,

$$S_t=S_0e^{\left(rt-\frac{1}{2}\eta_1(u,t)+\int_0^t\sigma_1(s)dW_s^1\right)},\ V_t=V_0e^{\left(rt-\frac{1}{2}\eta_2(u,t)+\int_0^t\sigma_2(s)dW_s^2\right)}, \tag{2.12}$$

where $\eta_i(u,t)=\sigma_i^2 t+(2\sigma_i\delta_i+\delta_i^2)u$, $i=1,2$.

In order to handle the above processes, we need to verify the properties of

$$J_1(t):=\int_0^t\sigma_1(s)dW_s^1,\ J_2(t):=\int_0^t\sigma_i(s)dW_s^2.$$

If $U_t$ is known, we can find the properties of $J_i(t)$, $(i=1,2)$. The results are presented by the following lemmas.

**2.2. Lemma.** *Conditioned on $U_t=u\le t$, $J_i(t)$ has the normal distribution with mean 0 and variance $\eta_i(u,t)$, for each $i\in\{1,2\}$.*

*Proof.* Let us consider the decomposition of $J_1(t)$ as

$$J_1(t) = \delta_1 \int_0^t \varepsilon(s) dW_s^1 + \sigma_1 W_t^1 := \delta_1 X_1(t) + \sigma_1 W_t^1.$$

For some $k$, $k\varepsilon(t)$ is a bounded simple function. So, the Novikov condition of $\mathbf{E}[e^{\frac{1}{2}\int_0^t (k\varepsilon(s))^2 ds}]$ is satisfied and $e^{\int_0^t k\varepsilon(s)dW_s^1 - \frac{u^2}{2}\int_0^t \varepsilon(s)^2 ds}$ is a martingale for given $U_t = u$. Therefore,

$$\begin{aligned} \mathbf{E}[e^{\int_0^t k\varepsilon(s)dW_s^1 - \frac{k^2}{2}\int_0^t \varepsilon(s)^2 ds}|U_t = u] &= \mathbf{E}[e^{\int_0^t k\varepsilon(s)dW_s^1 - \frac{k^2}{2}u}|U_t = u] \\ &= \mathbf{E}[e^{kX_1(t) - \frac{k^2}{2}u}|U_t = u] = 1. \end{aligned}$$

For given $U_t = u$, since $\mathbf{E}[e^{kX_1(t)}] = e^{\frac{k^2}{2}u}$, $X_1(t)$ has the normal distribution with mean 0 and variance $u$. We also can calculate the covariance of $X_1(t)$ and $W_t^1$ as following:

$$\begin{aligned} & \mathbf{E}[X_1(t)W_t^1 \big| U_t = u] \\ = \ & \mathbf{E}\left[\lim_{n\to\infty}\sum_{k=1}^{n}\int_{(k-1)t/n}^{kt/n} 1_{\{\varepsilon(s)=1, \frac{(k-1)t}{n}\le s\le \frac{kt}{n}\}} dW_s^1 W_t^1 \Big| U_t = u\right] \\ = \ & \lim_{n\to\infty}\mathbf{E}\left[\left(\sum_{\varepsilon(s)=1, \frac{(k-1)t}{n}\le s\le \frac{kt}{n}} W^1_{\frac{kt}{n}} - W^1_{\frac{(k-1)t}{n}}\right) W_t^1 \Big| U_t = u\right] \\ = \ & \lim_{n\to\infty}\mathbf{E}\left[\left(\sum_{\varepsilon(s)=1, \frac{(k-1)t}{n}\le s\le \frac{kt}{n}} W^1_{\frac{kt}{n}} W_t^1\right)\right. \\ & \qquad \left. - \left(\sum_{\varepsilon(s)=1, \frac{(k-1)t}{n}\le s\le \frac{kt}{n}} W^1_{\frac{(k-1)t}{n}} W_t^1\right) \Big| U_t = u\right] \\ = \ & \lim_{n\to\infty}\mathbf{E}\left[\sum_{\varepsilon(s)=1, \frac{(k-1)t}{n}\le s\le \frac{kt}{n}} \frac{t}{n} \Big| U_t = u\right] \\ = \ & \mathbf{E}\left[\lim_{n\to\infty}\sum_{k=1}^{n}\frac{t}{n} 1_{\{\varepsilon(s)=1, \frac{(k-1)t}{n}\le s\le \frac{kt}{n}\}} \Big| U_t = u\right] = u. \end{aligned}$$

Hence, for given $U_t = u$, $J_1(t)$ has the normal distribution with mean 0 and variance $\sigma_1^2 t + (2\sigma_1\delta_1 + \delta_1^2)u$. In a same way, $J_2(t)$ has the the normal distribution with mean 0 and variance $\sigma_2^2 t + (2\sigma_2\delta_2 + \delta_2^2)u$ as well. □

**2.3. Lemma.** *Conditioned on $U_t = u \le t$, the correlation of $J_1(t)$ and $J_2(t)$ is given by*

$$\rho_{12}(u,t) = \frac{[(\sigma_1\delta_2 + \sigma_2\delta_1 + \delta_1\delta_2)u + \sigma_1\sigma_2 t]\rho}{\sqrt{\eta_1(u,t)\eta_2(u,t)}}.$$

*Proof.* From the decomposition in Lemma 2.2 and $dW_t^1 dW_t^2 = \rho dt$, the covariance $J_1(t)$ and $J_2(t)$ is given by

$$\begin{aligned} \mathrm{Cov}(J_1(t), J_2(t)) &= \mathbf{E}[(\delta_1 X_1(t) + \sigma_1 W_t^1)(\delta_2 X_2(t) + \sigma_2 W_t^2)\big| U_t = u] \\ &= \delta_1\delta_2\mathbf{E}[X_1(t)X_2(t)\big|U_t = u] + \sigma_1\delta_2\mathbf{E}[W_t^1 X_2(t)\big|U_t = u] \\ &\quad + \delta_1\sigma_2\mathbf{E}[X_1(t)W_t^2\big|U_t = u] + \sigma_1\sigma_2\mathbf{E}[W_t^1 W_t^2\big|U_t = u] \\ &= \delta_1\delta_2\rho u + \sigma_1\delta_2\rho u + \delta_1\sigma_2\rho u + \sigma_1\sigma_2\rho t. \end{aligned}$$

Therefore, the correlation of $J_1(t)$ and $J_2(t)$ is obtained by Lemma 2.2. □

## 3. Valuation of options with correlated credit risk

In this section we provide the formula of the European call option with credit risk in a business cycle environment. As in Klein [11], we assume that if default or bankruptcy of the option issuer occurs, the option issuer's asset is immediately liquidated and the scrap value at $T$ is $(1-\alpha)V_T D^{-1}(S_T-K)^+$, where $D$ is a constant value of the option issuer's liabilities and $\alpha$ is a constant showing the ratio of bankruptcy costs of the issuer's asset. We also assume that the option issuer declare default only if $V_T < D$. Then, from the equation (2.12), the discounted expected value of the call option with maturity $T$ is given by

$$C(T) = e^{-rT}\mathbf{E}^Q[(S_T-K)^+ 1_{\{V_T\geq D\}} + (1-\alpha)V_T D^{-1}(S_T-K)^+ 1_{\{V_T<D\}}], \tag{3.1}$$

where $K$ is the strike price and $0 \leq \alpha \leq 1$. From this equation, we now provide the valuation formula for a option with credit risk and business cycle by applying the Girsanov's theorem repeatedly.

For notational simplicity, we rewrite notations as

$$\eta_1(u) := \eta_1(u,T), \eta_2(u) := \eta_2(u,T), \widehat{\rho}(u) := \rho_{12}(u,T), \delta_T := (1-\alpha)V_T D^{-1}.$$

**3.1. Proposition.** *Let $C_j$ be the arbitrage free price of a call option with credit risk and initial state $j$ $(j=0,1)$. Then, the value $C_j(T)$ at time 0 of the option with maturity $T$ is given by*

$$C_j(T) = \int_0^T v(u) f_j(u;T)du + \delta_0(j)e^{-\lambda_0 T}v(0) + \delta_1(j)e^{-\lambda_1 T}v(T), \tag{3.2}$$

*where $f_j(u;T)$ $(j=0,1)$ is defined in Proposition 1. And*

$$\begin{aligned} v(u) \quad = \quad & S_0\Phi_2(a_1(u), a_2(u), \widehat{\rho}(u)) - Ke^{-rT}\Phi_2(b_1(u), b_2(u), \widehat{\rho}(u)) \\ & + S_0\delta_0 e^{rT+\widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}}\Phi_2(c_1(u), c_2(u), -\widehat{\rho}(u)) - K\delta_0\Phi_2(d_1(u), d_2(u), -\widehat{\rho}(u)), \end{aligned}$$

*where $\Phi_2$ is the bivariate standard normal cumulative density function and*

$$\begin{aligned} a_1(u) &= \frac{\ln(S_0/K) + rT + \frac{1}{2}\eta_1(u)}{\sqrt{\eta_1(u)}}, \\ a_2(u) &= \frac{\ln(V_0/D) + rT - \frac{1}{2}\eta_2(u) + \widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}}{\sqrt{\eta_2(u)}}, \\ b_1(u) &= \frac{\ln(S_0/K) + rT - \frac{1}{2}\eta_1(u)}{\sqrt{\eta_1(u)}}, \\ b_2(u) &= \frac{\ln(V_0/D) + rT - \frac{1}{2}\eta_2(u)}{\sqrt{\eta_2(u)}}, \\ c_1(u) &= \frac{\ln(S_0/K) + rT + \frac{1}{2}\eta_1(u) + \widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}}{\sqrt{\eta_1(u)}}, \\ c_2(u) &= -\frac{\ln(V_0/D) + rT + \frac{1}{2}\eta_2(u) + \widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}}{\sqrt{\eta_2(u)}}, \\ d_1(u) &= \frac{\ln(S_0/K) + rT - \frac{1}{2}\eta_1(u) + \widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}}{\sqrt{\eta_1(u)}}, \\ d_2(u) &= -\frac{\ln(V_0/D) + rT + \frac{1}{2}\eta_2(u)}{\sqrt{\eta_2(u)}}. \end{aligned}$$

*Proof.* From equation (3.1), the credit-risky call option value $C_j(T)$ at time 0 with maturity $T$ and an initial state $j$ is given by

$$
\begin{aligned}
C_j(T) &= e^{-rT}\mathbf{E}^Q[\mathbf{E}^Q[(S_T-K)^+1_{\{V_T\geq D\}}+\delta_T(S_T-K)^+1_{\{V_T<D\}}|U_t=u]] \\
&= e^{-rT}\int_0^T \mathbf{E}^Q[(S_T-K)^+(1_{\{V_T\geq D\}}+\delta_T 1_{\{V_T<D\}})|U_t=u]f_j(u;T)du \\
&\quad +\delta_0(j)e^{-(r+\lambda_0)T}\mathbf{E}^Q[(S_T-K)^+(1_{\{V_T\geq D\}}+\delta_T 1_{\{V_T<D\}})|U_t=0] \\
&\quad +\delta_1(j)e^{-(r+\lambda_1)T}\mathbf{E}^Q[(S_T-K)^+(1_{\{V_T\geq D\}}+\delta_T 1_{\{V_T<D\}})|U_t=T].
\end{aligned}
\tag{3.3}
$$

Let us consider the first term of the equation (3.3). For a fixed $u$, the conditional expectation in the integral is divided into four terms as

$$
\begin{aligned}
& e^{-rT}\mathbf{E}^Q[(S_T-K)^+(1_{\{V_T\geq D\}}+\delta_T 1_{\{V_T<D\}})|U_t=u] \\
&= e^{-rT}\mathbf{E}^Q[S_T 1_{\{S_T>K,V_T\geq D\}}|U_T=u]-e^{-rT}\mathbf{E}^Q[K1_{\{S_T>K,V_T\geq D\}}\big|U_T=u] \\
&\quad +e^{-rT}\mathbf{E}^Q[S_T\delta_T 1_{\{S_T>K\}}1_{\{V_T<D\}}|U_T=u]-e^{-rT}\mathbf{E}^Q[K\delta_T 1_{\{S_T>K,V_T<D\}}|U_T=u] \\
&:= I_1-I_2+I_3-I_4.
\end{aligned}
$$

Under the measure $Q$, the first term $I_1$ can be expressed as

$$
\begin{aligned}
I_1 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_0 e^{-\frac{1}{2}\eta_1(u)+\sqrt{\eta_1(u)}z_1}1_{\{S_T>K\}}1_{\{V_T\geq D\}} \\
&\qquad \times\frac{1}{2\pi\sqrt{1-\widehat{\rho}(u)}}e^{-\frac{1}{2(1-\widehat{\rho}(u))}(z_1^2-2\widehat{\rho}(u)z_1z_2+z_2^2)}dz_1dz_2,
\end{aligned}
\tag{3.4}
$$

where $z_1=J_T^1/\sqrt{\eta_1(u)}$ and $z_2=J_T^2/\sqrt{\eta_2(u)}$ are the standard normal variables with correlation $\widehat{\rho}(u)$. Then, by the change of variables with $\widetilde{z}_1=z_1-\sqrt{\eta_1(u)}$, $\widetilde{z}_2=z_2-\widehat{\rho}(u)\sqrt{\eta_1(u)}$, we have

$$
\begin{aligned}
I_1 = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_0 1_{\{S_T>K\}}1_{\{V_T\geq D\}} \\
&\times\frac{1}{2\pi\sqrt{1-\widehat{\rho}(u)}}e^{-\frac{1}{2(1-\widehat{\rho}(u))}(\widetilde{z}_1^2-2\widehat{\rho}(u)\widetilde{z}_1\widetilde{z}_2+\widetilde{z}_2^2)}d\widetilde{z}_1d\widetilde{z}_2.
\end{aligned}
\tag{3.5}
$$

Let $\widetilde{Q}$ be the new equivalent probability measure defined by

$$
\frac{d\widetilde{Q}}{dQ}=\exp\left(\int_0^T\theta(s)dW_s-\frac{1}{2}\int_0^T|\theta(s)|^2ds\right),
\tag{3.6}
$$

where $W$ is vector in $R^2$ and $\theta(s)=(\sigma_1(s),\widehat{\rho}(u)\sigma_1(s))$. Then, by Girsanov's theorem,

$$
\begin{pmatrix} d\widetilde{W}_t^1 \\ d\widetilde{W}_t^2 \end{pmatrix}=\begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix}-\theta(t)dt
$$

is a $R^2$-valued standard Brownian motion under the equivalent measure $\widetilde{Q}$.

We consider the equation (3.5) under the measure $\widetilde{Q}$. Then, by applying Lemma 2.2 and Lemma 2.3, we have

$$\begin{aligned}
I_1 &= \mathbf{E}^{\widetilde{Q}}[S_0 1_{\{S_T>K,V_T>D\}}|U_T=u] \\
&= S_0\widetilde{P}\left(S_0 e^{\left(rT-\frac{1}{2}\eta_1(u)+\int_0^T \sigma_1(s)dW_s^1\right)} > K, V_0 e^{\left(rT-\frac{1}{2}\eta_2(u)+\int_0^T \sigma_2(s)dW_s^2\right)} > D\right) \\
&= S_0\widetilde{P}\left(\widetilde{J}_T^1 > -\left(\ln\frac{S_0}{K}+rT+\frac{1}{2}\eta_1(u)\right),\right. \\
&\qquad\qquad \left.\widetilde{J}_T^2 > -\left(\ln\frac{V_0}{D}+rT-\frac{1}{2}\eta_2(u)+\widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}\right)\right) \\
\text{(3.7)}\quad &= S_0\Phi_2(a_1(u),a_2(u),\widehat{\rho}(u)).
\end{aligned}$$

where $\widetilde{J}_T^1=\int_0^T\sigma_1(s)d\widetilde{W}_s^1$ and $\widetilde{J}_T^2=\int_0^T\sigma_1(s)d\widetilde{W}_s^2$.

In a similar way, without the change of measure, $I_2$ can be found.

For the evaluation $I_3$, we change the variables as $\widetilde{z}_1=z_1-\sqrt{\eta_1(u)}-\widehat{\rho}(u)\sqrt{\eta_2(u)}$, $\widetilde{z}_2=z_2-\sqrt{\eta_2(u)}-\widehat{\rho}(u)\sqrt{\eta_1(u)}$. And, define the equivalent measure by $\frac{d\widetilde{Q}}{dQ}=\exp\left(\int_0^T\theta(s)dW_s-\frac{1}{2}\int_0^T|\theta(s)|^2ds\right)$, where $\theta(s)=(\sigma_1(s)+\widehat{\rho}(u)\sigma_2(s),\sigma_2(s)+\widehat{\rho}(u)\sigma_1(s))$. Then, by Girsanov's theorem, we have

$$\begin{aligned}
&I_3=\mathbf{E}^{\widetilde{Q}}[e^{rT}S_0\delta_0 e^{\widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}}1_{\{S_T>K,V_T<D\}}|U_T=u] \\
&=e^{rT}S_0\delta_0 e^{\widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}} \\
&\qquad\times\widetilde{P}\left(\widetilde{J}_1^T > -\left(\ln\frac{S_0}{K}+rT+\frac{1}{2}\eta_1(u)+\widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}\right),\right. \\
&\qquad\qquad \left.\widetilde{J}_2^T > \left(\ln\frac{V_0}{D}+rT+\frac{1}{2}\eta_2(u)+\widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}\right)\right) \\
\text{(3.8)}\quad &=S_0\delta_0 e^{rT+\widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}}\Phi_2(c_1(u),c_2(u),-\widehat{\rho}(u)).
\end{aligned}$$

Again from the Radon-Nikodym derivative (3.6) that allows the change of probability measure, we change the measure with $\theta(s)=(\widehat{\rho}(u)\sigma_2(s),\sigma_2(s))^T$. Then, under an equivalent measure $\widetilde{Q}$, $I_4$ is evaluated as

$$\begin{aligned}
I_4 &= K\delta_0\mathbf{E}^{\widetilde{Q}}[1_{\{S_T>K,V_T<D\}}|U_T=u] \\
\text{(3.9)}\quad &= K\delta_0\Phi_2(d_1(u),d_2(u),-\widehat{\rho}(u)).
\end{aligned}$$

Also one can obtain the second term and the third term of the equation (3.3) from above results. This completes the proof. □

In a similar way, the following Proposition provides the price of the put option with credit risk.

**3.2. Proposition.** *Let $P_j$ be the arbitrage free price of a put option with credit risk and initial state $j$ ($j=0,1$). Then, the value $P_j(T)$ at time 0 of the option with maturity $T$ is given by*

$$P_j(T)=\int_0^T v(u)f_j(u;T)du+\delta_0(j)e^{-\lambda_0 T}v(0)+\delta_1(j)e^{-\lambda_1 T}v(T), \tag{3.10}$$

*where*

$$\begin{aligned}
v(u) &= -S_0\Phi_2(-a_1(u),a_2(u),-\widehat{\rho}(u))+Ke^{-rT}\Phi_2(-b_1(u),b_2(u),-\widehat{\rho}(u)) \\
&\quad -S_0\delta_0 e^{rT+\widehat{\rho}(u)\sqrt{\eta_1(u)\eta_2(u)}}\Phi_2(-c_1(u),c_2(u),\widehat{\rho}(u))+K\delta_0\Phi_2(-d_1(u),d_2(u),\widehat{\rho}(u)).
\end{aligned}$$

*Here, all parameters are given in Proposition 3.1.*

## 4. Numerical example

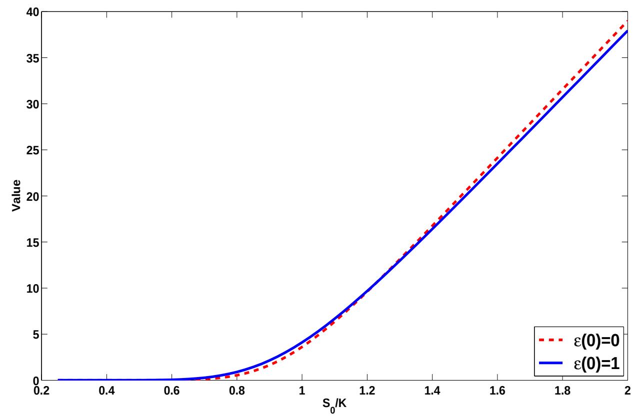



**Figure 1.** Vulnerable call value for different moneyness $(S_0/K)$ and $\varepsilon(0)$

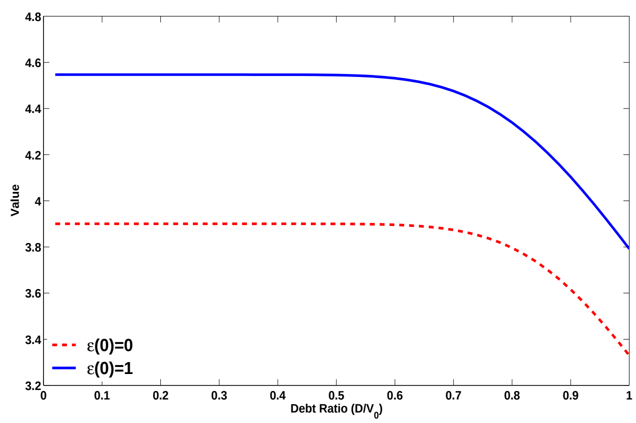



**Figure 2.** Vulnerable call value for different debt ratio $(D/V_0)$ and $\varepsilon(0)$

In the previous section, we provide the option formulae represented as a integral form under our model. In order to calculate these option formulae, we employ the Gauss-Legendre quadrature as a numerical approximation method. Based on the values reported by Boyle and Draviam [3] and Klein and Inglis [12], we use the following parameters unless stated otherwise: $S_0 = K = 40, V_0 = 100, D = 90, r = 0.05, T = 1, \alpha = 0.25, \rho = 0, \sigma_1 = \sigma_2 = 0.15, \delta_1 = \delta_2 = 0.1,$ , $\lambda_0 = \lambda_1 = 1$ and $\varepsilon(0) = 0$.

Fig. 1 illustrates how the prices of vulnerable call option for two initial states change with the moneyness $(S_0/K)$. We can observe that the option with $\varepsilon(0) = 0$ procedure higher prices than the option with $\varepsilon(0) = 1$ in high moneyness region as expected.

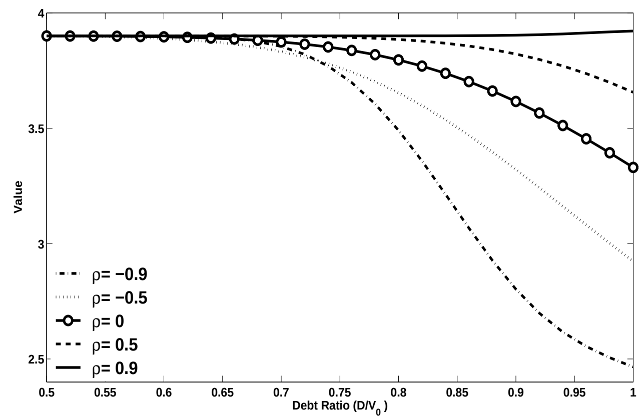



**Figure 3.** Vulnerable call value for different debt ratio $(D/V_0)$ and $\rho$

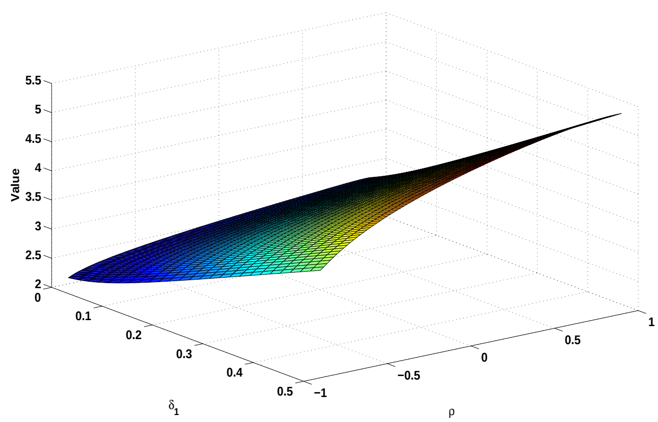



**Figure 4.** Vulnerable call value for different $\delta_1$ and $\rho$

Fig. 2 and Fig. 3 illustrate how the prices change when the debt ratio $(D/V_0)$ vary. Fig. 2 shows decreasing trends of prices for different initial states. Here, the option with $\varepsilon(0) = 0$ has always lower prices than the option with $\varepsilon(0) = 1$. We also can see that the negative correlation $\rho$ between underlying asset and firm value processes leads to lower option prices in Fig. 3.

Fig. 4 and Fig. 5 illustrate the the sensitivities of the options with respect to the shock sizes $\delta_i$, $(i = 1, 2)$ of the volatilities and the correlation. Both Fig. 4 and Fig. 5 show increasing trends of the option prices with respect to the correlation $\rho$. In Fig. 4, the shock size $\delta_1$ of the underlying asset also leads to an increasing trend. In contrast, an decreasing trend of the option values with respect to the shock size $\delta_2$ of the firm value process is found in Fig. 5. In addition, for a negative $\rho$, we can see a sharp decreasing of the option values.

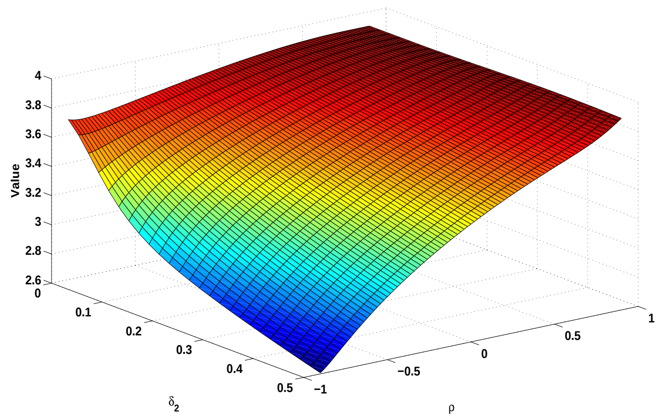



**Figure 5.** Vulnerable call value for different $\delta_2$ and $\rho$

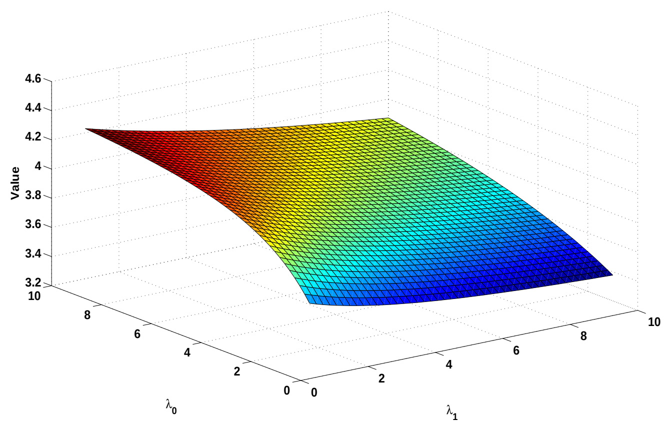



**Figure 6.** Vulnerable call value for different $\lambda_1$ and $\lambda_2$

Finally, Fig. 6 illustrates how the option values have the contrary trends with respect to intensities. Consequently, these results show the changes of the option values when the intensities vary by business cycle.


## Acknowledgement

The author would like to thank the anonymous referees for giving constructive suggestions and useful comments. The research was supported by the National Research Foundation of Korea Grant NRF-2015R1C1A1A02037533 and BK21 PLUS SNU Mathematical Sciences Division.

## References

[1] Black. F and Scholes, M. *The pricing of options and corporate liabilities*, J. Polit. Econ. **81**, 637-654, 1973.

[2] Bollen, N.P.B. *Valuing options in regime-switching models*, J. Deriv. **6**, 38-49, 1998.

[3] Boyle, P. P. and Draviam, T. *Pricing exotic options under regime switching*, Insurance Math. Econom. **40**, 267-282, 2007.

[4] Buffington, J. and Elliott, R. J. *American options with regime switching*, Int. J. Theoret. Appl. **5**, 497-514, 2002.

[5] Chang, L. F. and Huang, M. W. *Valuation of vulnerable American options with correlated credit risk*, Rev. Deriv. Res. **9**, 137-165, 2006.

[6] Elliott, R. J., Chan,L. L. and Siu, T. K. *Option pricing and Esscher transform under regime switching*, Ann. Fin. **1**, 423-432, 2005.

[7] Elliott, R. J., Chan,L. L. and Siu, T. K. *Option valuation under a regime-switching constant elasticity of variance process*, Appl. Math. Comput. **219**, 4434-4443, 2013.

[8] Hamilton, J. D. *A new approach to the economic analysis of nonstationary time series and the business cycle*, Econometrica **57**, 357-384 , 1989.

[9] Hui, C. H., Lo, C. F. and Ku, K. C. *Pricing vulnerable European options with stochastic default barriers*, IMA J. Manag. Math. **18**, 315-329, 2007.

[10] Johnson, H. and Stulz, R. *The pricing of options with default risk*, J. Finance **42**, 267-280, 1987.

[11] Klein, P.C. *Pricing Black-Scholes options with correlated credit risk*, J. Bank. Finance **20**, 1211-1229, 1996.

[12] Klein, P.C. and Inglis, M. *Valuation of European options subject to financial distress and interest rate risk*, J. Deriv. **6**, 44-56, 1999.

[13] Shiu, Y. M., Chou, P. L. and Sheu, J. W. *A closed-form approximation for valuing European basket warrants under credit risk and interest rate risk*, Quant. Finance **8**, 1211-1223, 2013.

[14] Vasicek, O. *A An equilibrium characterization of the term structure*, J. Financ. Econ. **5**, 177-188, 1977.

[15] Zhang, X., Elliott, R. J., Siu, T. K. and Guo, J. *Markovian regime-switching market completion using additional Markov jump assets*, IMA J. Manag. Math. **23**, 283-305, 2012.

[16] Zhu, S. P., Badran, A. and Lu, X. *A new exact solution for pricing European options in a two-state regimeswitching economy*, Comput. Math. Appl. **64**, 2744-2755, 2012.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 207 – 225

# New difference-cum-ratio and exponential type estimators in median ranked set sampling

Nursel Koyuncu*†

**Abstract**

This paper suggests difference-cum-ratio and exponential type estimators of population mean using first or third quartiles and mean of auxiliary variable under median ranked set sampling scheme and we have extended our study in double sampling scheme when the population parameters are unknown. The bias and mean square error of estimators are derived by theoretically both of sampling designs. Empirical studies have been done to demonstrate the efficiency of proposed estimators over the existing estimators. We have found that difference-cum-ratio estimator is always more efficient than regression estimator and both of the estimators are considerable efficient than existing estimators.

**Keywords:** Median ranked set sampling, Exponential estimator, Ratio estimator, Efficiency.

*2000 AMS Classification:* 62D05.

*Received* : 02.07.2014 *Accepted* : 26.01.2015 *Doi* : 10.15672/HJMS.2015509378

## 1. Introduction

The ranked set sampling (RSS) is conducted by selecting n random samples from the population of size n units each, and ranking each unit within each set with respect to variable of interest. Then an actual measurement is taken of the unit with the smallest rank from the first sample. From the second sample an actual measurement is taken from the second smallest rank, and the procedure is continued until the unit with the largest rank is chosen for actual measurement from the n-th sample. Median ranked set sampling (MRSS) as proposed by Muttlak [10] can be formed by selecting n random samples of size n units from the population and rank the units within each sample with respect to variable of interest. Many authors developed and modified this sampling scheme such as Al-Saleh and Al-Omari[2], Jemain and Al-Omari[4] and Jemain et al.[5], Ozturk and Jafari Jozani[11] etc. Recently Al-Omari[1] has introduced modified ratio estimators

*Hacettepe University,Department of Statistics, Beytepe, 06800, Ankara, Turkey Email: nkoyuncu@hacettepe.edu.tr

†Corresponding Author.

in median ranked set sampling. In sampling theory some authors such as Singh and Solanki[12, 13], Singh et al.[14, 15] and Solanki et al.[16] etc. proposed estimators for population parameters using auxiliary information. Bahl and Tuteja [3], Yadav et al. [18], Koyuncu and Kadilar[7], Koyuncu[8], Koyuncu et al.[9] studied the exponential estimators to get more efficient estimators than ratio and regression estimators. In this paper following Koyuncu [8], we have suggested two new estimators of population mean under Al-Omari[1] median ranked set sampling scheme, extended our results to double sampling and we have found that the suggested estimators are considerable efficient than classical ratio estimator and Al-Omari[1] estimator.

**Simple Random Sampling Design**

Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be a bivariate random sample with pdf $f(x, y)$, means $\mu_x$, $\mu_y$, variances $\sigma_x^2$, $\sigma_y^2$ and correlation coefficient $\rho_{xy}$. Assume that the ranking is performed on the auxiliary variable $X$ to estimate the mean of the variable of interest $Y$. Let $(X_{11}, Y_{11}), (X_{12}, Y_{12}), \ldots, (X_{nn}, Y_{nn})$ be $n$ independent bivariate random samples each of size $n$. In this sampling design Al-Omari [1] defined following estimators

$$\hat{\mu}_{YSRS1} = \bar{y}_{SRS}(\frac{\mu_x + q_1}{\bar{x}_{SRS} + q_1}), \qquad \hat{\mu}_{YSRS3} = \bar{y}_{SRS}(\frac{\mu_x + q_3}{\bar{x}_{SRS} + q_3})$$

where $q_1$ and $q_3$ are first and third quartiles of $X$ , respectively.$\bar{x}_{SRS}$ , $\bar{y}_{SRS}$ are sample means of $X$ and $Y$ . Al-Omari [1] rewrite estimators as:

$$\hat{\mu}_{YSRSk} = \bar{y}_{SRS}(\frac{\mu_x + q_k}{\bar{x}_{SRS} + q_k}) \tag{1.1}$$

where $\hat{\mu}_{YSRSk}$ represent $\hat{\mu}_{YSRS1}$ and $\hat{\mu}_{YSRS3}$ for values of $(k = 1, 3)$ . The expression for MSE of $\hat{\mu}_{YSRSk}$ is as follows:

$$MSE(\hat{\mu}_{YSRSk}) = \lambda\sigma_y^2 + \lambda\sigma_x^2(\frac{\mu_y^2}{(\mu_x + q_k)^2} - 2\beta\frac{\mu_y}{(\mu_x + q_k)}) \tag{1.2}$$

where $\lambda = 1/n$, $\beta = \rho_{xy}\sigma_y/\sigma_x$.

**Median Ranked Set Sampling Design**

For the sake of brevity we follow Al-Omari [1]' sampling design and notations. Median ranked set sampling design can be described as in the following steps:

(1) Select $n$ random samples each of size $n$ bivariate units from the population of interest.
(2) The units within each sample are ranked by visual inspection or any other cost free method with respect to a variable of interest.
(3) If $n$ is odd, select the $((n + 1)/2)$ th-smallest ranked unit $X$ together with the associated $Y$ from each set, i.e., the median of each set. If $n$ is even, from the first $n/2$ sets select the $(n/2)$th ranked unit $X$ together with the associated $Y$ and from the other sets select the $((n + 2)/2)$ th ranked unit $X$ together with the associated $Y$ .
(4) The whole process can be repeated $m$ times if needed to obtain a sample of size $nm$ units.

Let $(X_{i(1)}, Y_{i[1]}), (X_{i(2)}, Y_{i[2]}), \ldots, (X_{i(n)}, Y_{i[n]})$ be the order statistics of $X_{i1}, X_{i2}, \ldots, X_{in}$ and the judgement order of $Y_{i1}, Y_{i2}, \ldots, Y_{in}$ $(i = 1, 2, \ldots, n)$, where () and [] indicate that the ranking of $X$ is perfect and ranking of $Y$ has errors. For odd and even sample sizes the units measured using MRSS are denoted by MRSSO and MRSSE, respectively. For odd sample size let $(X_{1(\frac{n+1}{2})}, Y_{1[\frac{n+1}{2}]}), (X_{2(\frac{n+1}{2})}, Y_{2[\frac{n+1}{2}]}), \ldots, (X_{n(\frac{n+1}{2})}, Y_{n[\frac{n+1}{2}]})$ denote the observed units by MRSSO. $\bar{x}_{MRSSO} = \frac{1}{n}\sum_{i=1}^{n} X_{i(\frac{n+1}{2})}$ and $\bar{y}_{MRSSO} = \frac{1}{n}\sum_{i=1}^{n} X_{i[\frac{n+1}{2}]}$ be the sample mean of $X$ and $Y$ respectively.

For even sample size let $(X_{1(\frac{n}{2})}, Y_{1[\frac{n}{2}]}), (X_{2(\frac{n}{2})}, Y_{2[\frac{n}{2}]}), \ldots, (X_{\frac{n}{2}(\frac{n}{2})}, Y_{\frac{n}{2}[\frac{n}{2}]})$, $(X_{\frac{n+2}{2}(\frac{n+2}{2})}, Y_{\frac{n+2}{2}[\frac{n+2}{2}]}), (X_{\frac{n+4}{2}(\frac{n+2}{2})}, Y_{\frac{n+4}{2}[\frac{n+2}{2}]}), \ldots, (X_{n(\frac{n}{2})}, Y_{n[\frac{n}{2}]})$ denote the observed units by MRSSE. $\bar{x}_{MRSSE} = \frac{1}{n}(\sum_{i=1}^{\frac{n}{2}} X_{i(\frac{n}{2})} + \sum_{i=\frac{n+2}{2}}^{n} X_{i(\frac{n+2}{2})})$ and $\bar{y}_{MRSSE} = \frac{1}{n}(\sum_{i=1}^{\frac{n}{2}} Y_{i[\frac{n}{2}]} + \sum_{i=\frac{n+2}{2}}^{n} Y_{i[\frac{n+2}{2}]})$ be the sample mean of $X$ and $Y$ respectively.

To obtain the bias and the mean square error (MSE), let us define

$$\varepsilon_{0(j)} = \frac{\bar{y}_{MRSS(j)} - \mu_y}{\mu_y}, \ \varepsilon_{1(j)} = \frac{\bar{x}_{MRSS(j)} - \mu_x}{\mu_x}, \ \varepsilon_{2(j)} = \frac{s_{yx(j)} - \sigma_{yx(j)}}{\sigma_{yx(j)}},$$

$$\varepsilon_{3(j)} = \frac{s^2_{x(j)} - \sigma^2_{x(j)}}{\sigma^2_{x(j)}}$$

where $j = (E, O)$ denote the sample size even or odd. If sample size $n$ is odd we can write

$$E(\varepsilon^2_{0(O)}) = \frac{1}{n\mu_y^2}\sigma^2_{y[\frac{n+1}{2}]}, \ E(\varepsilon^2_{1(O)}) = \frac{1}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})},$$

$$E(\varepsilon_{0(O)}\varepsilon_{1(O)}) = \frac{1}{n\mu_x\mu_y}\sigma_{xy[\frac{n+1}{2}]}$$

If sample size $n$ is even we can write

$$E(\varepsilon^2_{0(E)}) = \frac{1}{2n\mu_y^2}(\sigma^2_{y[\frac{n}{2}]} + \sigma^2_{y[\frac{n+2}{2}]}), \ E(\varepsilon^2_{1(E)}) = \frac{1}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}),$$

$$E(\varepsilon_{0(E)}\varepsilon_{1(E)}) = \frac{1}{2n\mu_x\mu_y}(\sigma_{yx[\frac{n}{2}]} + \sigma_{yx[\frac{n+2}{2}]})$$

**(i) Al-Omari(2012) estimator** The estimator of population mean proposed by Al-Omari[1] as

$$\hat{\mu}_{YMRRS1} = \bar{y}_{MRSS}(\frac{\mu_x + q_1}{\bar{x}_{MRSS} + q_1}), \hat{\mu}_{YMRRS3} = \bar{y}_{MRSS}(\frac{\mu_x + q_3}{\bar{x}_{MRSS} + q_3})$$

For odd and even sample sizes the estimator can be rewritten as

$$\hat{\mu}_{YMRSSk} = \bar{y}_{MRSS(j)}(\frac{\mu_x + q_k}{\bar{x}_{MRSS(j)} + q_k}) \tag{1.3}$$

To the first degree of approximation the Bias and MSE of $\hat{\mu}_{YMRRSk}$ are respectively given by

$$Bias(\hat{\mu}_{YMRSS(j)}) \cong \begin{cases} \frac{\mu_y\psi}{n\mu_x}(\psi\frac{1}{\mu_x}\sigma_{x(\frac{n+1}{2})^2} - \frac{1}{\mu_y}\sigma_{xy[\frac{n+1}{2}]}) & \text{if } n \text{ is odd} \\ \frac{\mu_y\psi}{2n\mu_x}(\psi\frac{1}{\mu_x}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) \\ -\frac{1}{\mu_y}(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]})) & \text{if } n \text{ is even} \end{cases} \tag{1.4}$$

(1.5)

$$MSE(\hat{\mu}_{YMRSS(j)}) \cong \begin{cases} \frac{1}{n}(\frac{\mu_y^2}{(\mu_x+q_k)^2}\sigma^2_{x(\frac{n+1}{2})} + \sigma^2_{y[\frac{n+1}{2}]} \\ -2\frac{\mu_y}{(\mu_x+q_k)}\sigma_{xy[\frac{n+1}{2}]}) & \text{if } n \text{ is odd} \\ \\ \frac{1}{2n}(\frac{\mu_y^2}{(\mu_x+q_k)^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})})+ \\ (\sigma^2_{y[\frac{n}{2}]} + \sigma^2_{y[\frac{n+2}{2}]}) - 2\frac{\mu_y}{(\mu_x+q_k)}(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]})) & \text{if } n \text{ is even} \end{cases}$$

where $\psi = \frac{\mu_x}{\mu_x+q_k}$.

**(ii) Adapted Regression estimator** We can define regression type estimator in median ranked set sampling given by

(1.6) $$\overline{y}_{reg(j)} = \overline{y}_{MRSS(j)} + b_{yx(j)}(\mu_x - \overline{x}_{MRSS(j)})$$

where

$$b_{xy(j)} \cong \begin{cases} \frac{\rho_{xy[\frac{n+1}{2}]}s_{y[\frac{n+1}{2}]}}{s_{x(\frac{n+1}{2})}} & \text{if } n \text{ is odd} \\ \frac{(\rho_{xy[\frac{n}{2}]}+\rho_{xy[\frac{n+2}{2}]})(s_{y[\frac{n}{2}]}+s_{y[\frac{n+2}{2}]})}{(s_{x(\frac{n}{2})}+s_{x(\frac{n+2}{2})})} & \text{if } n \text{ is even} \end{cases}$$

$$s_{xy[\frac{n+1}{2}]} = \rho_{xy[\frac{n+1}{2}]}s_{x(\frac{n+1}{2})}s_{y[\frac{n+1}{2}]}$$

$$(s_{xy[\frac{n}{2}]} + s_{xy[\frac{n+2}{2}]}) = (\rho_{xy[\frac{n}{2}]} + \rho_{xy[\frac{n+2}{2}]})(s_{x(\frac{n}{2})} + s_{x(\frac{n+2}{2})})(s_{y[\frac{n}{2}]} + s_{y[\frac{n+2}{2}]})$$

(1.7) $$MSE(\overline{y}_{reg(j)}) \cong \begin{cases} \frac{1}{n}\sigma^2_{y[\frac{n+1}{2}]}(1-\rho^2_{xy[\frac{n+1}{2}]}) & \text{if } n \text{ is odd} \\ \frac{1}{2n}(\sigma^2_{y[\frac{n}{2}]} + \sigma^2_{y[\frac{n+2}{2}]})(1-(\rho^2_{xy[\frac{n+1}{2}]} + \rho^2_{xy[\frac{n+2}{2}]})) & \text{if } n \text{ is even} \end{cases}$$

## 2. Suggested estimators in median ranked set sampling

Following Koyuncu [8], we propose difference-cum-ratio estimator estimating the population mean of the study variable in median ranked set sampling as follows:

(2.1) $$\overline{y}_{Nk(M)} = [k_{1(j)}\overline{y}_{MRSS(j)} + k_{2(j)}(\mu_x - \overline{x}_{MRSS(j)})](\frac{\mu_x + q_k}{\overline{x}_{MRSS(j)} + q_k})$$

where $k_{1(j)}$ and $k_{2(j)}$ are determined so as to minimize the MSE of $\overline{y}_{Nk(M)}$ . Expressing $\overline{y}_{Nk(M)}$ in terms of $\varepsilon_{(j)}$'s up to the second degree and extracting $\mu_y$ both sides we have

(2.2) $$\begin{aligned} \overline{y}_{Nk(M)} - \mu_y =& (k_{1(j)}-1)\mu_y + k_{1(j)}\mu_y\varepsilon_{0(j)} - k_{2(j)}\mu_x\varepsilon_{1(j)} - k_{1(j)}\mu_y\psi\varepsilon_{1(j)} \\ &+ k_{1(j)}\mu_y\psi\varepsilon_{0(j)}\varepsilon_{1(j)} + k_{2(j)}\mu_x\psi\varepsilon^2_{1(j)} + k_{1(j)}\mu_y\psi^2\varepsilon^2_{1(j)} \end{aligned}$$

Taking expectation in equation in (2.2), we obtain

(2.3)

$$Bias(\overline{y}_{Nk(M)}) \cong \begin{cases} (k_{1(O)}-1)\mu_y - k_{1(O)}\mu_y\psi\frac{1}{n\mu_y\mu_x}\sigma_{xy[\frac{n+1}{2}]} + (k_{2(O)}\mu_x\psi \\ +k_{1(O)}\mu_y\psi^2)\frac{1}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})} & \text{if } n \text{ is odd} \\ \\ (k_{1(E)}-1)\mu_y - k_{1(E)}\mu_y\psi\frac{1}{2n\mu_y\mu_x}(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]}) \\ +(k_{2(E)}\mu_x\psi + k_{1(E)}\mu_y\psi^2)\frac{1}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) & \text{if } n \text{ is even} \end{cases}$$

Squaring both sides in (2.2), then taking expectation, we obtain the MSE of the estimator $\overline{y}_{Nk(M)}$, as given by

$$(2.4) \quad MSE_{min}(\overline{y}_{Nk(M)}) \cong \begin{cases} (1-\frac{\psi^2}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})}) \\ \times\frac{MSE(\overline{y}_{reg(O)})}{1-\frac{\psi^2}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})}+\frac{1}{\mu_y^2}MSE(\overline{y}_{reg(O)})} & \text{if } n \text{ is odd} \\ (1-\frac{\psi^2}{n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) \\ \times\frac{MSE(\overline{y}_{reg(E)})}{1-\frac{\psi^2}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})}+\sigma^2_{x(\frac{n+2}{2})})+\frac{1}{\mu_y^2}MSE(\overline{y}_{reg(E)})} & \text{if } n \text{ is even} \end{cases}$$

The optimum values of $k_{1(j)}$ and $k_{2(j)}$ for odd and even sample sizes are given respectively

$$k^*_{1(O)} = \frac{1-\frac{\psi^2}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})}}{1-\frac{\psi^2}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})}) + \frac{1}{\mu_y^2}MSE(\overline{y}_{reg(O)})}$$

$$k^*_{1(E)} = \frac{1-\frac{\psi^2}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})})}{1-\frac{\psi^2}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) + \frac{1}{\mu_y^2}MSE(\overline{y}_{reg(E)})}$$

$$k^*_{2(O)} = \frac{\mu_y}{\mu_x}(\psi + \frac{(\frac{1}{\mu_y\mu_x}\sigma_{xy[\frac{n+1}{2}]} - 2\psi\frac{1}{\mu_x^2}\sigma^2_{x(\frac{n+1}{2})})(1-\frac{\psi^2}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})})}{\frac{1}{\mu_x^2}\sigma^2_{x(\frac{n+1}{2})}((1-\frac{\psi^2}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})}) + \frac{1}{\mu_y^2}MSE(\overline{y}_{reg(O)})})$$

$$k^*_{2(E)} = \frac{\mu_y}{\mu_x}(\psi + \frac{(\frac{1}{\mu_y\mu_x}(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]}) - 2\psi\frac{1}{\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}))(1-\frac{\psi^2}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}))}{\frac{1}{\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})})((1-\frac{\psi^2}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})})) + \frac{1}{\mu_y^2}MSE(\overline{y}_{reg(E)})})$$

Note that the optimum choice of the constants involve unknown parameters. These quantities can be guessed quite accurately through pilot sample survey or sample data or experience gathered in due course of time as mentioned in Upadhyaya and Singh[17], and Koyuncu and Kadilar[6].

Secondly following exponential estimator is proposed:

$$(2.5) \quad \overline{y}_{Kk(M)} = [w_{1(j)}\overline{y}_{MRSS(j)} + w_{2(j)}(\frac{\overline{x}_{MRSS(j)}}{\mu_x})^{\gamma}]exp(\frac{\mu_x - \overline{x}_{MRSS(j)}}{\mu_x - \overline{x}_{MRSS(j)} + 2q_k})$$

where $w_{1(j)}$ and $w_{2(j)}$ are determined so as to minimize the MSE of $\overline{y}_{Kk(M)}$.Expressing $\overline{y}_{Kk(M)}$ in terms of $\varepsilon_j$'s up to the second degree and extracting $\mu_y$ both sides, we have

(2.6)

$$\overline{y}_{Kk(M)} - \mu_y = \{w_{1(j)}\mu_y - \mu_y + w_{1(j)}\mu_y\varepsilon_{0(j)} + w_{2(j)} + w_{2(j)}\gamma\varepsilon_{1(j)} + w_{2(j)}\frac{\gamma(\gamma-1)}{2}\varepsilon^2_{1(j)}$$
$$- \frac{1}{2}w_{1(j)}\psi\mu_y\varepsilon_{1(j)} - \frac{1}{2}w_{1(j)}\psi\mu_y\varepsilon_{0(j)}\varepsilon_{1(j)} - \frac{1}{2}w_{2(j)}\psi\varepsilon_{1(j)}$$
$$- \frac{1}{2}w_{2(j)}\psi\gamma\varepsilon^2_{1(j)} + \frac{3}{8}w_{1(j)}\psi^2\mu_y\varepsilon^2_{1(j)} + \frac{3}{8}w_{2(j)}\psi^2\varepsilon^2_{1(j)}\}$$

Taking expectation in equation in (2.6), we obtain

(2.7)

$$Bias(\overline{y}_{Kk(M)}) \cong \begin{cases} w_{1(O)}\mu_y - \mu_y + w_{2(O)} - \frac{w_{1(O)}\psi}{2n\mu_x}\sigma_{xy[\frac{n+1}{2}]} \\ +(\frac{w_{2(O)}}{2}(-\psi\gamma + \frac{3}{4}\psi^2 + \gamma(\gamma-1)) \\ +\frac{3}{8}w_{1(O)}\psi^2\mu_y)\frac{1}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})} & \text{if } n \text{ is odd} \\ \\ w_{1(E)}\mu_y - \mu_y + w_{2(E)} - \frac{w_{1(E)}\psi}{4n\mu_x}(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]}) \\ +(\frac{w_{2(E)}}{2}(-\psi\gamma + \frac{3}{4}\psi^2 + \gamma(\gamma-1)) \\ +\frac{3}{8}w_{1(E)}\psi^2\mu_y)\frac{1}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) & \text{if } n \text{ is even} \end{cases}$$

Squaring both sides in (2.6), then taking expectation, we obtain the MSE of the estimator $\overline{y}_{Kk(M)}$ , as given by

(2.8)

$$MSE(\overline{y}_{Kk(M)}) = \mu_y^2 + w^2_{1(j)}\mu_y^2 + A_j + w^2_{2(j)}B_j + w_{1(j)}\mu_y^2 + D_j + w_{2(j)}\mu_y + G_j + w_{1(j)}w_{2(j)}\mu_yF_j$$

where

$$A_{(j)} = 1 + E(\varepsilon^2_{0(j)}) + \psi^2E(\varepsilon^2_{1(j)}) - 2\psi E(\varepsilon_{0(j)}\varepsilon_{1(j)})$$

$$B_{(j)} = 1 + (\gamma^2 + \psi^2 + \gamma(\gamma-1) - 2\gamma\psi)E(\varepsilon^2_{1(j)})$$

$$D_{(j)} = -2 + \psi E(\varepsilon_{0(j)}\varepsilon_{1(j)}) - \frac{3}{4}\psi^2E(\varepsilon^2_{1(j)})$$

$$G_{(j)} = -2 + (\psi\gamma - \gamma(\gamma-1) - \frac{3}{4}\psi^2)E(\varepsilon^2_{1(j)})$$

$$F_{(j)} = 2 + (\gamma(\gamma-1) - 2\gamma\psi + 2\psi^2)E(\varepsilon^2_{1(j)}) + 2(\gamma-\psi)E(\varepsilon_{0(j)}\varepsilon_{1(j)})$$

Minimization of (2.8) with respect to $w_{1(j)}$ and $w_{2(j)}$ yields its optimum value when

$$(2.9) \qquad w_{1(j)} = \frac{F_{(j)}G_{(j)} - 2B_{(j)}D_{(j)}}{4B_{(j)}A_{(j)} - F^2_{(j)}}, \qquad w_{2(j)} = \mu_y\frac{D_{(j)}F_{(j)} - 2A_{(j)}G_{(j)}}{4B_{(j)}A_{(j)} - F^2_{(j)}}$$

Substituting optimum values of $w_{1(j)}$ and $w_{2(j)}$ in (2.9), we get minimum MSE of $\overline{y}_{Kk(M)}$ as

$$(2.10) \qquad MSE_{min}(\overline{y}_{Kk(M)}) = \mu_y^2[1 - \frac{B_{(j)}D^2_{(j)} + A_{(j)}G^2_{(j)} - D_{(j)}F_{(j)}G_{(j)}}{4B_{(j)}A_{(j)} - F^2_{(j)}}]$$

## 3. Theoretical comparison

Firstly, we compare the MSE of proposed difference-cum-ratio estimator $\overline{y}_{Kk(M)}$ with the MSE of regression estimator $\overline{y}_{Reg(O)}$ when sample size is odd.

$$MSE_{min}(\overline{y}_{Nk(M)}) < MSE(\overline{y}_{Reg(O)})$$

$$(1 - \frac{\psi^2}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})})\frac{MSE(\overline{y}_{reg(O)})}{1 - \frac{\psi^2}{n\mu_x^2}\sigma^2_{x(\frac{n+1}{2})} + \frac{1}{\mu_y^2}MSE(\overline{y}_{reg(O)})} < MSE(\overline{y}_{Reg(O)})$$

$$0 < \frac{1}{\mu_y^2}MSE(\overline{y}_{Reg(O)}) \tag{3.1}$$

From (3.1), we can conclude that $\overline{y}_{Nk(M)}$ is always more efficient that $\overline{y}_{Reg(O)}$. Secondly, we compare the suggested exponential estimator $\overline{y}_{Nk(M)}$ with the regression estimator $\overline{y}_{Reg(E)}$ when sample size $n$ is even.

$$MSE_{min}(\overline{y}_{Nk(M)}) < MSE(\overline{y}_{Reg(E)})$$

$$(1 - \frac{\psi^2}{n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})})\frac{MSE(\overline{y}_{reg(E)})}{1 - \frac{\psi^2}{2n\mu_x^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) + \frac{1}{\mu_y^2}MSE(\overline{y}_{reg(E)})} < MSE(\overline{y}_{Reg(E)})$$

$$0 < \frac{1}{\mu_y^2}MSE(\overline{y}_{Reg(E)}) \tag{3.2}$$

From (3.2), we can conclude that $\overline{y}_{Nk(M)}$ is always more efficient than regression estimator.

## 4. Estimation of population mean when $\mu_x$ is unknown

In practice, when the population mean of auxiliary variable is unknown, double sampling method can be used to estimate $\mu_x$. In this section we assume that mean of auxiliary variable is unavailable. Thus following the procedure outlined in Al-Omari[1], in SRS, a large sample of size $n^{'}$ is selected to estimate $\mu_x$. Then a sub sample of size $n^{''}$ is selected from the target population in order to study the characteristic variable $Y$. In MRSS, simple random sampling is used at first phase and median ranked set sampling is used at second phase where $n^{'} = n^2$ and $n^{''} = n$. Let $\overline{x}^{'}_{SRS(j)}$ and $\overline{x}^{'}_{MRSS(j)}$ be the unbiased sample means of $\mu_x$ obtained using SRS and MRSS, respectively. Al-Omari [1] defined following estimator in double sampling

$$\hat{\mu}^{'}_{YSRSk} = \bar{y}_{SRS}(\frac{\overline{x}^{'}_{SRS} + q_k}{\overline{x}_{SRS} + q_k}) \tag{4.1}$$

In order to obtain the bias and mean square of the estimator in (4.1), let us define

$$e_0 = \frac{\overline{y}_{SRS} - \mu_y}{\mu_y},\ e_1 = \frac{\overline{x}_{SRS} - \mu_x}{\mu_x},\ e_1^{'} = \frac{\overline{x}^{'}_{SRS} - \mu_x}{\mu_x}.$$

Using these notations, the expectations are defined as $E(e_0) = E(e_1) = E(e_1)^{'}$

$$E(e_0^2) = \frac{1}{n^{''}}\frac{\sigma_y^2}{\mu_y^2},\ E(e_1^2) = \frac{1}{n^{''}}\frac{\sigma_x^2}{\mu_x^2},\ E(e_1^{'2}) = \frac{1}{n^{'}}\frac{\sigma_x^2}{\mu_x^2},\ E(e_0e_1) = \frac{1}{n^{''}}\frac{\sigma_{yx}}{\mu_x\mu_y},$$

$$E(e_1 e_1^{'}) = \frac{1}{n^{'}} \frac{\sigma_x^2}{\mu_x^2},\ E(e_0 e_1^{'}) = \frac{1}{n^{'}} \frac{\sigma_{yx}}{\mu_x \mu_y}.$$

Applying the same procedure for double sampling the bias and MSE of $\hat{\mu}_{YSRSk}^{'}$ are obtained respectively as,

$$(4.2)\qquad Bias(\hat{\mu}_{YSRSk}^{'}) = (\frac{1}{n^{''}} - \frac{1}{n^{'}})(\frac{\mu_y}{(\mu_x + q_k)^2}\sigma_x^2 - \frac{1}{\mu_x + q_k}\sigma_{yx})$$

$$(4.3)\qquad MSE(\hat{\mu}_{YSRSk}^{'}) = \frac{1}{n^{''}}\sigma_y^2 + (\frac{1}{n^{''}} - \frac{1}{n^{'}})\sigma_x^2(\frac{\mu_y^2}{(\mu_x + q_k)^2} - 2\beta\frac{\mu_y}{\mu_x + q_k})$$

Secondly Al-Omari [1] defined following estimator in double sampling

$$(4.4)\qquad \hat{\mu}_{YMRSSk}^{'} = \overline{y}_{MRSS(j)}(\frac{\overline{x}_{MRSS(j)}^{'} + q_k}{\overline{x}_{MRSS(j)} + q_k})$$

To obtain the bias and the MSE, let us define

$$\delta_{0(j)} = \frac{\overline{y}_{MRSS(j)} - \mu_y}{\mu_y},\ \delta_{1(j)} = \frac{\overline{x}_{MRSS(j)} - \mu_x}{\mu_x},\ \delta_{1(j)^{'}} = \frac{\overline{x}_{MRSS(j)}^{'} - \mu_x}{\mu_x}.$$

such that $E(\delta_{0(j)}) = E(\delta_{1(j)}) = E(\delta_{1(j)}^{'})$ where $(j) = O, E$ represents the sample size is odd or even. If sample size $n^{''}$ is odd we can write

$$E(\delta_{0(O)}^2) = \frac{1}{n^{''}} \frac{\sigma_{y[\frac{n+1}{2}]}^2}{\mu_y^2},\ E(\delta_{1(O)}^2) = \frac{1}{n^{''}} \frac{\sigma_{x(\frac{n+1}{2})}^2}{\mu_x^2},\ E(\delta_{0(O)}\delta_{1(O)}) = \frac{1}{n^{''}} \frac{\sigma_{xy[\frac{n+1}{2}]}^2}{\mu_x \mu_y},$$

$$E(\delta_{1(O)}^{'2}) = \frac{1}{n^{'}} \frac{\sigma_{x(\frac{n+1}{2})}^2}{\mu_x^2},\ E(\delta_{0(O)}\delta_{1(O)}^{'}) = \frac{1}{n^{'}} \frac{\sigma_{xy(\frac{n+1}{2})}}{\mu_x \mu_y},\ E(\delta_{1(O)}\delta_{1(O)}^{'}) = \frac{1}{n^{'}} \frac{\sigma_{x(\frac{n+1}{2})}^2}{\mu_x^2}.$$

If sample size $n^{''}$ is even we can write

$$E(\delta_{0(E)}^2) = \frac{1}{2n^{''}} \frac{\sigma_{y[\frac{n}{2}]}^2 + \sigma_{y[\frac{n+2}{2}]}^2}{\mu_y^2},\ E(\delta_{1(E)}^2) = \frac{1}{2n^{''}} \frac{\sigma_{x(\frac{n}{2})}^2 + \sigma_{x(\frac{n+2}{2})}^2}{\mu_x^2},$$

$$E(\delta_{0(E)}\delta_{1(E)}) = \frac{1}{2n^{''}} \frac{\sigma_{yx[\frac{n}{2}]} + \sigma_{yx[\frac{n+2}{2}]}}{\mu_x \mu_y},\ E(\delta_{1(E)}^2) = \frac{1}{2n^{''}} \frac{\sigma_{x(\frac{n}{2})}^2 + \sigma_{x(\frac{n+2}{2})}^2}{\mu_x^2},$$

$$E(\delta_{0(E)}\delta_{1(E)}^{'}) = \frac{1}{2n^{'}} \frac{\sigma_{yx[\frac{n}{2}]} + \sigma_{yx[\frac{n+2}{2}]}}{\mu_x \mu_y},\ E(\delta_{1(E)}\delta_{1(E)}^{'}) = \frac{1}{2n^{'}} \frac{\sigma_{x(\frac{n}{2})}^2 + \sigma_{x(\frac{n+2}{2})}^2}{\mu_x^2}.$$

Using the defined expectations the bias and MSE of $\hat{\mu}_{YMRSS(j)}^{'}$ are obtained respectively as,

$$(4.5)\quad Bias(\hat{\mu}_{YMRSS(j)}^{'}) \cong \begin{cases} \frac{\mu_y \psi}{\mu_x}(\frac{1}{n^{''}} - \frac{1}{n^{'}})(\psi \frac{1}{\mu_x}\sigma_{x(\frac{n+1}{2})^2} - \frac{1}{\mu_y}\sigma_{xy[\frac{n+1}{2}]}) & \text{if } n \text{ is odd} \\ \frac{\mu_y \psi}{2\mu_x}(\frac{1}{n^{''}} - \frac{1}{n^{'}})(\psi \frac{1}{\mu_x}(\sigma_{x(\frac{n}{2})}^2 + \sigma_{x(\frac{n+2}{2})}^2) \\ -\frac{1}{\mu_y}(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]})) & \text{if } n \text{ is even} \end{cases}$$

(4.6)

$$MSE(\hat{\mu}^{'}_{YMRSS(j)}) \cong \begin{cases} \frac{1}{n^{''}}\sigma^2_{y[\frac{n+1}{2}]} + (\frac{1}{n^{''}} - \frac{1}{n^{'}})\frac{\mu_y^2}{(\mu_x+q_k)^2}\sigma^2_{x(\frac{n+1}{2})} \\ -2(\frac{1}{n^{''}} - \frac{1}{n^{'}})\frac{\mu_y}{(\mu_x+q_k)}\sigma_{xy[\frac{n+1}{2}]} & \text{if } n \text{ is odd} \\ \\ \frac{1}{2}[\frac{1}{n^{''}}(\sigma^2_{y[\frac{n}{2}]} + \sigma^2_{y[\frac{n+2}{2}]}) + (\frac{1}{n^{''}} - \frac{1}{n^{'}})\frac{\mu_y^2}{(\mu_x+q_k)^2}(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) \\ -2(\frac{1}{n^{''}} - \frac{1}{n^{'}})\frac{\mu_y}{(\mu_x+q_k)}(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]})] & \text{if } n \text{ is even} \end{cases}$$

We have suggested new double sampling estimators as follows:

(4.7) $$\overline{y}^{'}_{reg(j)} = \overline{y}_{MRSS(j)} + b_{yx(j)}(\overline{x}^{'}_{MRSS(j)} - \overline{x}_{MRSS(j)})$$

The MSE of $\overline{y}^{'}_{reg(j)}$ is obtained as,

(4.8)

$$MSE(\overline{y}_{reg(j)}) \cong \begin{cases} \sigma^2_{y[\frac{n+1}{2}]}(\frac{1}{n^{''}} - (\frac{1}{n^{''}} - \frac{1}{n^{'}})\rho^2_{xy[\frac{n+1}{2}]}) & \text{if } n \text{ is odd} \\ \frac{1}{2}(\sigma^2_{y[\frac{n}{2}]} + \sigma^2_{y[\frac{n+2}{2}]})[\frac{1}{n^{''}} - (\frac{1}{n^{''}} - \frac{1}{n^{'}})(\rho^2_{xy[\frac{n+1}{2}]} + \rho^2_{xy[\frac{n+2}{2}]})] & \text{if } n \text{ is even} \end{cases}$$

In double sampling our suggested estimator can be defined as

(4.9) $$\overline{y}^{'}_{Nk(M)} = [k_{1(j)}\overline{y}_{MRSS(j)} + k_{2(j)}(\overline{x}^{'}_{MRSS(j)} - \overline{x}_{MRSS(j)})](\frac{\overline{x}^{'}_{MRSS(j)} + q_k}{\overline{x}_{MRSS(j)} + q_k})$$

The bias and MSE of $\overline{y}^{'}_{Nk(M)}$ are obtained respectively as,

(4.10)

$$Bias(\overline{y}^{'}_{Nk(M)}) \cong \begin{cases} (k_{1(O)} - 1)\mu_y - k_{1(O)}\mu_y\psi\frac{1}{\mu_y\mu_x}(\frac{1}{n^{''}} - \frac{1}{n^{'}})\sigma_{xy[\frac{n+1}{2}]} \\ +(k_{2(O)}\mu_x\psi + k_{1(O)}\mu_y\psi^2)(\frac{1}{n^{''}} - \frac{1}{n^{'}})\frac{1}{\mu_x^2}\sigma^2_{x(\frac{n+1}{2})} & \text{if } n \text{ is odd} \\ \\ (k_{1(E)} - 1)\mu_y - k_{1(E)}\mu_y\psi\frac{1}{2\mu_y\mu_x}(\frac{1}{n^{''}} - \frac{1}{n^{'}})(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]}) \\ +(k_{2(E)}\mu_x\psi + k_{1(E)}\mu_y\psi^2)\frac{1}{2\mu_x^2}(\frac{1}{n^{''}} - \frac{1}{n^{'}})(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) & \text{if } n \text{ is even} \end{cases}$$

(4.11)

$$MSE_{min}(\overline{y}^{'}_{Nk(M)}) \cong \begin{cases} (1 - \frac{\psi^2}{\mu_x^2}(\frac{1}{n^{''}} - \frac{1}{n^{'}})\sigma^2_{x(\frac{n+1}{2})}) \\ \times\frac{MSE(\overline{y}^{'}_{reg(O)})}{1 - \frac{\psi^2}{\mu_x^2}(\frac{1}{n^{''}} - \frac{1}{n^{'}})\sigma^2_{x(\frac{n+1}{2})} + \frac{1}{\mu_y^2}MSE(\overline{y}^{'}_{reg(O)})} & \text{if } n \text{ is odd} \\ \\ (1 - \frac{\psi^2}{\mu_x^2}(\frac{1}{n^{''}} - \frac{1}{n^{'}})(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})})) \\ \times\frac{MSE(\overline{y}^{'}_{reg(E)})}{1 - \frac{\psi^2}{2\mu_x^2}(\frac{1}{n^{''}} - \frac{1}{n^{'}})(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) + \frac{1}{\mu_y^2}MSE(\overline{y}^{'}_{reg(E)})} & \text{if } n \text{ is even} \end{cases}$$

Our second estimator can be defined as in double sampling

(4.12) $$\overline{y}^{'}_{Kk(M)} = [w_{1(j)}\overline{y}_{MRSS(j)} + w_{2(j)}(\frac{\overline{x}_{MRSS(j)}}{\overline{x}^{'}_{MRSS(j)}})^{\gamma}]exp(\frac{\overline{x}^{'}_{MRSS(j)} - \overline{x}_{MRSS(j)}}{\overline{x}_{MRSS(j)} + \overline{x}_{MRSS(j)} + 2q_k})$$

The bias and MSE of $\overline{y}'_{Kk(M)}$ are obtained respectively as,

(4.13)

$$Bias(\overline{y}'_{Kk(M)}) \cong \begin{cases} w_{1(O)}\mu_y - \mu_y + w_{2(O)} - \frac{w_{1(O)}\psi}{2\mu_x}(\frac{1}{n''} - \frac{1}{n'})\sigma_{xy[\frac{n+1}{2}]} \\ +(\frac{w_{2(O)}}{2}(-\psi\gamma + \frac{3}{4}\psi^2 + \gamma(\gamma-1)) \\ +\frac{3}{8}w_{1(O)}\psi^2\mu_y)\frac{1}{\mu_x^2}(\frac{1}{n''} - \frac{1}{n'})\sigma^2_{x(\frac{n+1}{2})} & \text{if } n \text{ is odd} \\ \\ w_{1(E)}\mu_y - \mu_y + w_{2(E)} - \frac{w_{1(E)}\psi}{4\mu_x}(\frac{1}{n''} - \frac{1}{n'})(\sigma_{xy[\frac{n}{2}]} + \sigma_{xy[\frac{n+2}{2}]}) \\ +(\frac{w_{2(E)}}{2}(-\psi\gamma + \frac{3}{4}\psi^2 + \gamma(\gamma-1)) \\ +\frac{3}{8}w_{1(E)}\psi^2\mu_y)\frac{1}{2\mu_x^2}(\frac{1}{n''} - \frac{1}{n'})(\sigma^2_{x(\frac{n}{2})} + \sigma^2_{x(\frac{n+2}{2})}) & \text{if } n \text{ is even} \end{cases}$$

$$(4.14) \quad MSE_{min}(\overline{y}'_{Kk(M)}) = \mu_y^2[1 - \frac{B'_{(j)}D'^2_{(j)} + A'_{(j)}G'^2_{(j)} - D'_{(j)}F'_{(j)}G'_{(j)}}{4B'_{(j)}A'_{(j)} - F'^2_{(j)}}]$$

$$A'_{(j)} = 1 + E(\delta^2_{0(j)}) + \psi^2 E(\delta^2_{1(j)}) - \psi^2 E(\delta'^2_{1(j)}) - 2\psi E(\delta_{0(j)}\delta_{1(j)}) + 2\psi E(\delta_{0(j)}\delta'_{1(j)})$$

$$B'_{(j)} = 1 + (\gamma^2 + \psi^2 + \gamma(\gamma-1) - 2\gamma\psi)(E(\delta^2_{1(j)}) - E(\delta'^2_{1(j)}))$$

$$D'_{(j)} = -2 + \psi E(\delta_{0(j)}\delta_{1(j)}) - \psi E(\delta_{0(j)}\delta'_{1(j)}) - \frac{3}{4}\psi^2 E(\delta^2_{1(j)}) + \frac{3}{4}\psi^2 E(\delta'^2_{1(j)})$$

$$G'_{(j)} = -2 + (\psi\gamma - \gamma(\gamma-1) - \frac{3}{4}\psi^2)(E(\delta^2_{1(j)}) - E(\delta'^2_{1(j)}))$$

$$F'_{(j)} = 2 + (\gamma(\gamma-1) - 2\gamma\psi + 2\psi^2)(E(\delta^2_{1(j)}) - E(\delta'^2_{1(j)})) + 2(\gamma-\psi)(E(\delta_{0(j)}\delta_{1(j)}) - E(\delta_{0(j)}\delta'_{1(j)}))$$

## 5. Simulation study

In this section, we conducted a simulation study to investigate the properties of proposed estimators. . In the simulation study, we consider finite populations of size $N = 10000$ generated from a bivariate normal distribution $N(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho_{xy})$. The samples were generated from a bivariate normal distribution using mvrnorm function in R programme. In the simulation, we considered $\mu_x = 2$, $\mu_y = 4$, $\sigma_x^2 = \sigma_y^2 = 1$ and different values of $\rho_{xy}$. We have computed mean square errors (MSEs) and percent relative efficiencies (PREs) of estimators with respect to $\hat{\mu}_{YSRSk}$ for $n = 3, 4, 5, 6$ on the basis of 60.000 replications using $q_k$ and displayed in Table1 and Table4. When the mean of auxiliary variable is unknown, we used double sampling method to estimate $\mu_x$ and we have calculated MSEs and PREs of estimators given in (4.1)-(4.14). Findings are summarized in Table5 and Table8.

It is observed from all tables, suggested difference-cum-ratio and exponential type estimator performs better than Al-Omari[1] estimator. We can conclude that difference-cum-ratio estimator gives always more efficient results than regression estimator as shown in theoretical comparison section. When we compare difference-cum-ratio and exponential type estimator we can say that exponential type estimator performs better even with the low correlation data sets.

## 6. Conclusion

In this paper we have suggested difference-cum-ratio and exponential type estimator in median ranked set sampling and extended our result to double sampling. We have found that difference-cum-ratio estimator is always more efficient than regression estimator and exponential type is better than difference-cum-ratio estimator. Both of the estimators are considerable efficient than Al-Omari[1] estimator.

**Acknowledgement** I'm thankful to learned referees for their constructive comments.

## References

[1] Al-Omari, A. I. *Ratio estimation of the population mean using auxiliary information in simple random sampling and median ranked set sampling*, Statistics and Probability Letters **82**, 1883-1890, 2012.

[2] Al-Saleh, M.F. and Al-Omari, A.I. *Multistage ranked set sampling*, Journal of Statistical Planning and Inference **102**, 273-286, 2002.

[3] Bahl, S. and Tuteja, R.K. *Ratio and product exponential estimators*, Journal of Information and Optimization Sciences **12** (1), 159-164, 1991.

[4] Jemain, A.A. and Al-Omari, A.I. *Multistage median ranked set samples for estimating the population mean*, Pakistan Journal of Statistics **22**, 195-207, 2006.

[5] Jemain, A.A., Al-Omari, A.I. and Ibrahim, K. *Multistage extreme ranked set sampling for estimating the population mean*, Journal of Statistical Theory and Application **6** (4), 456-471, 2007.

[6] Koyuncu, N. and Kadilar, C. *Family of estimators of population mean using two auxiliary variables in stratified random sampling*, Communications in Statistics: Theory and Methods **38** (14), 2398-2417, 2009.

[7] Koyuncu, N. and Kadilar, C. (2010). *On improvement in estimating population mean in stratified random sampling*, Journal of Applied Statistics **37** (6), 999-1013, 2010.

[8] Koyuncu, N. *Effcient estimators of population mean using auxiliary attributes.* Applied Mathematics Computation **218**, 10900-10905, 2012.

[9] Koyuncu, N., Gupta, S. and Sousa, R. *Exponential type estimators of the mean of a sensitive variable in the presence of non-sensitive auxiliary information*, Communications in Statistics: Simulation and Computation **43**, 1583-1594, 2014.

[10] Muttlak, H.A. *Median ranked set sampling*, Journal of Applied Statistical Science **6**, 245-255, 1997.

[11] Ozturk, O. and Jafari Jozani, M. *Inclusion probabilities in partially rank ordered set sampling*, Computational Statistics and Data Analysis **69**, 122-132, 2014.

[12] Singh, H.P. and Solanki, R.S. *An alternative procedure for estimating the population mean in simple random sampling*, Pak. Jour. Statist. Oper. Res. **8** (2), 213-232, 2012.

[13] Singh, H.P. and Solanki, R.S. *Improved estimation of population mean in simple random sampling using information on auxiliary attribute*, Appl. Math. Com. **218** (15), 7798-7812, 2012.

[14] Singh, H.P., Tailor, R. and Tailor, R. *On ratio and product methods with certain known population parameters of auxiliary variable in sample surveys*, SORT **34** (2), 157-180, 2010.

[15] Singh, S., Singh, H.P. and Upadhyaya, L.N. *Chain ratio and regression type estimators for median estimation in survey sampling*, Statist. Papers **48**, 23-46, 2006.

[16] Solanki R. S., Singh H.P. and Pal S. K. *Improved estimation of finite population mean sample surveys*, J. Adv. Com. **1** (2), 70-78, 2013.

[17] Upadhyaya, L.N., Singh, H.P. *Almost unbiased ratio and product-type estimators of finite population variance in sample surveys*, Statistics in Transition **7** (5), 1087-1096, 2006.

[18] Yadav, R., Upadhyaya, L.N., Singh, H.P. and Chatterjee, S. *Improved separate ratio exponential estimator for population mean using auxiliary information*, Statist. Trans.-new series **12** (2), 401-411, 2011.

**Table 1.** Mean Square Error (MSE) and the Percent Relative Efficiency (PRE) of estimators with respect to $\hat{\mu}_{YSRS1}$ for using $n = 3, 4, 5, 6$ and $q_1$ with positive correlation

| **Correlation** | **Estimator** | **n=3** | **n=4** | **n=5** | **n=6** |
|---|---|---|---|---|---|
| $\rho = 0.99$ | $\hat{\mu}_{YSRS1}$ | 0.02716 | 0.01918 | 0.01473 | 0.01200 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS1(M)}$ | 0.01467 | 0.00602 | 0.00657 | 0.00331 |
| | | 185.20 | 318.43 | 224.29 | 362.87 |
| | $\hat{\mu}_{K1(M)}$ | 0.01134 | 0.00441 | 0.00611 | 0.00276 |
| | | 239.49 | 435.31 | 241.05 | 435.66 |
| | $\hat{\mu}_{N1(M)}$ | **0.00628*** | **0.00236*** | **0.00334*** | **0.00148*** |
| | | **432.73*** | **812.15*** | **440.79*** | **811.72*** |
| $\rho = 0.80$ | $\hat{\mu}_{YSRS1}$ | 0.20222 | 0.14713 | 0.11344 | 0.09326 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS1(M)}$ | 0.12239 | 0.05059 | 0.05039 | 0.02702 |
| | | 165.23 | 290.84 | 225.11 | 345.12 |
| | $\hat{\mu}_{K1(M)}$ | **0.07085*** | **0.03298*** | **0.02971*** | **0.01706*** |
| | | **285.43*** | **446.14*** | **381.86*** | **546.66*** |
| | $\hat{\mu}_{N1(M)}$ | 0.07993 | 0.0344 | 0.03402 | 0.01845 |
| | | 253.00 | 427.70 | 333.44 | 505.59 |
| $\rho = 0.70$ | $\hat{\mu}_{YSRS1}$ | 0.29675 | 0.21575 | 0.16607 | 0.13646 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS1(M)}$ | 0.16636 | 0.07052 | 0.06580 | 0.03650 |
| | | 178.39 | 305.96 | 252.36 | 373.84 |
| | $\hat{\mu}_{K1(M)}$ | **0.07411*** | **0.03588*** | **0.02998*** | **0.01799*** |
| | | **400.41*** | **601.30*** | **553.97*** | **758.65*** |
| | $\hat{\mu}_{N1(M)}$ | 0.10245 | 0.04585 | 0.04203 | 0.0238 |
| | | 289.67 | 470.60 | 395.13 | 573.36 |
| $\rho = 0.50$ | $\hat{\mu}_{YSRS1}$ | 0.47919 | 0.34640 | 0.26697 | 0.22066 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS1(M)}$ | 0.23938 | 0.10748 | 0.09296 | 0.05362 |
| | | 200.18 | 322.29 | 287.17 | 411.51 |
| | $\hat{\mu}_{K1(M)}$ | **0.07394*** | **0.03637*** | **0.02849*** | **0.01778*** |
| | | **648.06*** | **952.42*** | **936.95*** | **1240.76*** |
| | $\hat{\mu}_{N1(M)}$ | 0.13382 | 0.06211 | 0.05219 | 0.03109 |
| | | 358.10 | 557.71 | 511.54 | 709.69 |
| MSE of Estimators | | | | | |
| PRE of Estimators | | | | | |
| ***represent most efficient estimator (having minimum MSE and maximum PRE)** | | | | | |

**Table 2.** (MSE) and (PRE) of estimators with respect to $\hat{\mu}_{YSRS1}$ for using $n = 3, 4, 5, 6$ and $q_1$ with negative correlation

| **Correlation** | **Estimator** | **n=3** | **n=4** | **n=5** | **n=6** |
|---|---|---|---|---|---|
| $\rho = -0.99$ | $\hat{\mu}_{YSRS1}$ | 2.09657 | 1.45252 | 1.10811 | 0.89932 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS1(M)}$ | 0.79973 | 0.37209 | 0.28960 | 0.17612 |
| | | 262.16 | 390.36 | 382.63 | 510.64 |
| | $\hat{\mu}_{K1(M)}$ | **0.00147*** | **0.00059*** | **0.00077*** | **0.00036*** |
| | | **142202.83*** | **246394.52*** | **143164.56*** | **249205.74*** |
| | $\hat{\mu}_{N1(M)}$ | 0.00629 | 0.00236 | 0.00336 | 0.00148 |
| | | 33324.11 | 61423.61 | 33018.28 | 60570.96 |
| $\rho = -0.80$ | $\hat{\mu}_{YSRS1}$ | 185.513 | 129.121 | 0.98902 | 0.80473 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS1(M)}$ | 0.67687 | 0.32443 | 0.24210 | 0.15124 |
| | | 274.07 | 398.00 | 408.52 | 532.10 |
| | $\hat{\mu}_{K1(M)}$ | **0.01921*** | **0.00829*** | **0.00842*** | **0.00454*** |
| | | **9656.02*** | **15573.14*** | **11751.25*** | **17726.36*** |
| | $\hat{\mu}_{N1(M)}$ | 0.08075 | 0.03444 | 0.03443 | 0.01860 |
| | | 2297.51 | 3749.26 | 2872.41 | 4327.54 |
| $\rho = -0.70$ | $\hat{\mu}_{YSRS1}$ | 1.74465 | 1.21628 | 0.93268 | 0.75966 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS1(M)}$ | 0.63102 | 0.30480 | 0.22639 | 0.14191 |
| | | 276.48 | 399.04 | 411.98 | 535.31 |
| | $\hat{\mu}_{K1(M)}$ | **0.02617*** | **0.01163*** | **0.01102*** | **0.00617*** |
| | | **6666.34*** | **10458.08*** | **8460.68*** | **12320.69*** |
| | $\hat{\mu}_{N1(M)}$ | 0.10380 | 0.04589 | 0.04257 | 0.02395 |
| | | 1680.71 | 2650.18 | 2191.01 | 3172.07 |
| $\rho = -0.70$ | $\hat{\mu}_{YSRS1}$ | 1.51489 | 1.06044 | 0.81509 | 0.66543 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS1(M)}$ | 0.55071 | 0.26779 | 0.19972 | 0.12511 |
| | | 275.08 | 396.00 | 408.12 | 531.87 |
| | $\hat{\mu}_{K1(M)}$ | **0.03780*** | **0.01755*** | **0.01524*** | **0.00894*** |
| | | **4007.48*** | **6041.93*** | **5349.97*** | **7445.09*** |
| | $\hat{\mu}_{N1(M)}$ | 0.13334 | 0.06210 | 0.05245 | 0.03112 |
| | | 1136.13 | 1707.74 | 1554.05 | 2138.37 |
| MSE of Estimators | | | | | |
| PRE of Estimators | | | | | |
| ***represent most efficient estimator (having minimum MSE and maximum PRE)** | | | | | |

**Table 3.** (MSE) and (PRE) of estimators with respect to $\hat{\mu}_{YSRS3}$ for using $n = 3, 4, 5, 6$ and $q_3$ with positive correlation

| Correlation | Estimator | n=3 | n=4 | n=5 | n=6 |
|---|---|---|---|---|---|
| $\rho = 0.99$ | $\hat{\mu}_{YSRS3}$ | 0.01302 | 0.00963 | 0.00751 | 0.00625 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS3(M)}$ | 0.00833 | 0.00351 | 0.00402 | 0.00198 |
| | | 156.24 | 274.21 | 186.70 | 315.51 |
| | $\hat{\mu}_{K3(M)}$ | 0.01403 | 0.00545 | 0.00761 | 0.00342 |
| | | 92.82 | 176.83 | 98.71 | 182.94 |
| | $\hat{\mu}_{N3(M)}$ | **0.00613*** | **0.00233*** | **0.00331*** | **0.00147*** |
| | | **212.37*** | **412.55*** | **227.00*** | **425.46*** |
| $\rho = 0.80$ | $\hat{\mu}_{YSRS3}$ | 0.12348 | 0.09280 | 0.07286 | 0.06073 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}_{YRSS3(M)}$ | 0.08088 | 0.03462 | 0.03474 | 0.01869 |
| | | 152.68 | 268.03 | 209.72 | 324.98 |
| | $\hat{\mu}_{K3(M)}$ | 0.08867 | 0.04142 | 0.03733 | 0.02146 |
| | | 139.26 | 224.05 | 195.20 | 283.03 |
| | $\hat{\mu}_{N3(M)}$ | **0.07802*** | **0.03400*** | **0.03371*** | **0.01834*** |
| | | **158.27*** | **272.91*** | **216.14*** | **331.24*** |
| $\rho = 0.70$ | $\hat{\mu}_{YSRS3}$ | 0.18311 | 0.13745 | 0.10786 | 0.08983 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\mu_{YRSS3(M)}$ | 0.11056 | 0.04855 | 0.04559 | 0.02540 |
| | | 165.62 | 283.15 | 236.59 | 353.62 |
| | $\hat{\mu}_{K3(M)}$ | **0.09278*** | **0.04508*** | **0.03768*** | **0.02264*** |
| | | **197.35*** | **304.93*** | **286.25*** | **396.84*** |
| | $\hat{\mu}_{N3(M)}$ | 0.10004 | 0.04533 | 0.04166 | 0.02366 |
| | | 183.04 | 303.22 | 258.89 | 379.64 |
| $\rho = 0.50$ | $\hat{\mu}_{YSRS3}$ | 0.30341 | 0.22668 | 0.17871 | 0.14877 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\mu_{YRSS3(M)}$ | 0.16198 | 0.07396 | 0.06427 | 0.03738 |
| | | 187.32 | 306.48 | 278.08 | 398.02 |
| | $\hat{\mu}_{K3(M)}$ | **0.09301*** | **0.04591*** | **0.03596*** | **0.02248*** |
| | | **326.21*** | **493.78*** | **496.96*** | **661.91*** |
| | $\hat{\mu}_{N3(M)}$ | 0.13094 | 0.06145 | 0.05165 | 0.03088 |
| | | 231.71 | 368.88 | 345.98 | 481.71 |
| MSE of Estimators | | | | | |
| PRE of Estimators | | | | | |
| ***represent most efficient estimator (having minimum MSE and maximum PRE)** | | | | | |

**Table 4.** (MSE) and (PRE) of estimators with respect to $\hat{\mu}_{YSRS3}$ for using $n = 3, 4, 5, 6$ and $q_3$ with negative correlation

| Correlation | Estimator | n=3 | n=4 | n=5 | n=6 |
|---|---|---|---|---|---|
| $\rho = -0.99$ | $\hat{\mu}_{YSRS3}$ | 1.27876 | 0.92691 | 0.72344 | 0.59704 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\mu_{YRSS3(M)}$ | 0.53198 | 0.25563 | 0.19968 | 0.12260 |
| | | 240.38 | 362.59 | 362.30 | 486.98 |
| | $\hat{\mu}_{K3(M)}$ | **0.00170*** | **0.00066*** | **0.00092*** | **0.00041*** |
| | | **75346.28*** | **140653.13*** | **78496.62*** | **143988.13*** |
| | $\hat{\mu}_{N3(M)}$ | 0.00615 | 0.00234 | 0.00332 | 0.00147 |
| | | 20801.50 | 39655.80 | 21769.81 | 40491.47 |
| $\rho = -0.80$ | $\hat{\mu}_{YSRS3}$ | 1.14487 | 0.83140 | 0.65053 | 0.53757 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\mu_{YRSS3(M)}$ | 0.45347 | 0.22367 | 0.16751 | 0.10561 |
| | | 252.47 | 371.70 | 388.35 | 508.99 |
| | $\hat{\mu}_{K3(M)}$ | **0.02409*** | **0.01038*** | **0.01056*** | **0.00568*** |
| | | **4751.97*** | **8012.21*** | **6162.50*** | **9461.94*** |
| | $\hat{\mu}_{N3(M)}$ | 0.07899 | 0.03406 | 0.03409 | 0.01847 |
| | | 1449.33 | 2440.67 | 1908.53 | 2910.45 |
| $\rho = -0.70$ | $\hat{\mu}_{YSRS3}$ | 1.07804 | 0.78365 | 0.61373 | 0.50746 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\mu_{YRSS3(M)}$ | 0.42290 | 0.21001 | 0.15654 | 0.0990 |
| | | 254.92 | 373.14 | 392.05 | 512.48 |
| | $\hat{\mu}_{K3(M)}$ | **0.03289*** | **0.01460*** | **0.01386*** | **0.00774*** |
| | | **3277.86*** | **5366.87*** | **4429.80*** | **6557.28*** |
| | $\hat{\mu}_{N3(M)}$ | 0.10158 | 0.0454 | 0.04214 | 0.02379 |
| | | 1061.33 | 1726.09 | 1456.49 | 2133.32 |
| $\rho = -0.50$ | $\hat{\mu}_{YSRS3}$ | 0.94135 | 0.68590 | 0.53818 | 0.44554 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\mu_{YRSS3(M)}$ | 0.37019 | 0.18468 | 0.13816 | 0.08733 |
| | | 254.29 | 371.40 | 389.52 | 510.17 |
| | $\hat{\mu}_{K3(M)}$ | **0.04753*** | **0.02208*** | **0.01917*** | **0.01124*** |
| | | **1980.46*** | **3106.64*** | **2807.53*** | **3963.50*** |
| | $\hat{\mu}_{N3(M)}$ | 0.13054 | 0.06144 | 0.05192 | 0.03091 |
| | | 721.10 | 1116.30 | 1036.59 | 1441.34 |
| MSE of Estimators | | | | | |
| PRE of Estimators | | | | | |
| ***represent most efficient estimator (having minimum MSE and maximum PRE)** | | | | | |

**Table 5.** (MSE) and (PRE) of estimators with respect to $\hat{\mu}'_{YSRS1}$ for using $n = 3, 4, 5, 6$ and $q_1$ with positive correlation

| Correlation | Estimator | n=3 | n=4 | n=5 | n=6 |
|---|---|---|---|---|---|
| $\rho = 0.99$ | $\hat{\mu}'_{YSRS1}$ | 0.17912 | 0.11756 | 0.08670 | 0.06762 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS1(M)}$ | 0.23023 | 0.11654 | 0.08777 | 0.05530 |
| | | 77.80 | 100.87 | 98.78 | 122.27 |
| | $\hat{\mu}'_{K1(M)}$ | **0.13146*** | **0.06466*** | **0.05121*** | **0.03149*** |
| | | **136.26*** | **181.81*** | **169.30*** | **214.73*** |
| | $\hat{\mu}'_{N1(M)}$ | 0.13819 | 0.06916 | 0.05378 | 0.03343 |
| | | 129.62 | 169.98 | 161.21 | 202.28 |
| $\rho = 0.80$ | $\hat{\mu}'_{YSRS1}$ | 0.23888 | 0.16633 | 0.12742 | 0.10244 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS1(M)}$ | 0.28069 | 0.13827 | 0.10670 | 0.06626 |
| | | 85.10 | 120.30 | 119.42 | 154.60 |
| | $\hat{\mu}'_{K1(M)}$ | **0.09576*** | **0.04582*** | **0.03806*** | **0.02298*** |
| | | **249.45*** | **363.05*** | **334.76*** | **445.74*** |
| | $\hat{\mu}'_{N1(M)}$ | 0.15289 | 0.07387 | 0.05938 | 0.03623 |
| | | 156.25 | 225.17 | 214.58 | 282.76 |
| $\rho = 0.70$ | $\hat{\mu}'_{YSRS1}$ | 0.30028 | 0.21632 | 0.16904 | 0.13807 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS1(M)}$ | 0.32584 | 0.15892 | 0.12276 | 0.07609 |
| | | 92.16 | 136.12 | 137.70 | 181.46 |
| | $\hat{\mu}'_{K1(M)}$ | **0.08211*** | **0.03926*** | **0.03228*** | **0.01962*** |
| | | **365.69*** | **551.02*** | **523.62*** | **703.64*** |
| | $\hat{\mu}'_{N1(M)}$ | 0.16068 | 0.07666 | 0.06180 | 0.03760 |
| | | 186.87 | 282.19 | 273.51 | 367.23 |
| $\rho = 0.50$ | $\hat{\mu}'_{YSRS1}$ | 0.42384 | 0.31654 | 0.25222 | 0.20933 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS1(M)}$ | 0.40172 | 0.19537 | 0.14864 | 0.09282 |
| | | 105.51 | 162.02 | 169.69 | 225.52 |
| | $\hat{\mu}'_{K1(M)}$ | **0.07532*** | **0.03651*** | **0.02847*** | **0.01773*** |
| | | **562.73*** | **867.00*** | **885.88*** | **1180.43*** |
| | $\hat{\mu}'_{N1(M)}$ | 0.16669 | 0.07933 | 0.06271 | 0.03849 |
| | | 254.27 | 399.03 | 402.20 | 543.81 |
| MSE of Estimators | | | | | |
| PRE of Estimators | | | | | |
| ***represent most efficient estimator (having minimum MSE and maximum PRE)** | | | | | |

**Table 6.** (MSE) and (PRE) of estimators with respect to $\hat{\mu}'_{YSRS1}$ for using $n = 3, 4, 5, 6$ and $q_1$ with negative correlation

| **Correlation** | **Estimator** | **n=3** | **n=4** | **n=5** | **n=6** |
|---|---|---|---|---|---|
| $\rho = -0.99$ | $\hat{\mu}'_{YSRS1}$ | 1.44310 | 1.12785 | 0.90938 | 0.76770 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS1(M)}$ | 0.97002 | 0.46323 | 0.34803 | 0.21602 |
| | | 148.77 | 243.48 | 261.29 | 355.38 |
| | $\hat{\mu}'_{K1(M)}$ | **0.03720*** | **0.01674*** | **0.01039*** | **0.00668*** |
| | | **3879.01*** | **6736.19*** | **8751.86*** | **11490.33*** |
| | $\hat{\mu}'_{N1(M)}$ | 0.14743 | 0.07198 | 0.04646 | 0.03019 |
| | | 978.84 | 1566.82 | 1957.45 | 2542.77 |
| $\rho = -0.80$ | $\hat{\mu}'_{YSRS1}$ | 1.30377 | 1.02123 | 0.82264 | 0.69395 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS1(M)}$ | 0.85105 | 0.41594 | 0.30128 | 0.19141 |
| | | 153.20 | 245.52 | 273.05 | 362.54 |
| | $\hat{\mu}'_{K1(M)}$ | **0.05221*** | **0.02310*** | **0.01720*** | **0.01023*** |
| | | **2497.40*** | **4421.60*** | **4783.49*** | **6785.34*** |
| | $\hat{\mu}'_{N1(M)}$ | 0.17864 | 0.08184 | 0.06264 | 0.03715 |
| | | 729.85 | 1247.78 | 1313.19 | 1868.15 |
| $\rho = -0.70$ | $\hat{\mu}'_{YSRS1}$ | 1.23136 | 0.96516 | 0.77731 | 0.65554 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS1(M)}$ | 0.80362 | 0.39541 | 0.28511 | 0.18167 |
| | | 153.23 | 244.09 | 272.64 | 360.85 |
| | $\hat{\mu}'_{K1(M)}$ | **0.05748*** | **0.02561*** | **0.01925*** | **0.01148*** |
| | | **2142.34*** | **3769.17*** | **4037.59*** | **5711.08*** |
| | $\hat{\mu}'_{N1(M)}$ | 0.18225 | 0.08339 | 0.06404 | 0.03800 |
| | | 675.63 | 1157.41 | 1213.80 | 1725.34 |
| $\rho = -0.50$ | $\hat{\mu}'_{YSRS1}$ | 1.09266 | 0.85687 | 0.68971 | 0.58123 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS1(M)}$ | 0.72687 | 0.36005 | 0.25968 | 0.16548 |
| | | 150.32 | 237.99 | 265.60 | 351.23 |
| | $\hat{\mu}'_{K1(M)}$ | **0.06490*** | **0.02952*** | **0.02212*** | **0.01334*** |
| | | **1683.64*** | **2902.88*** | **3117.84*** | **4357.02*** |
| | $\hat{\mu}'_{N1(M)}$ | 0.18003 | 0.08347 | 0.06318 | 0.03790 |
| | | 606.94 | 1026.53 | 1091.65 | 1533.57 |
| MSE of Estimators | | | | | |
| PRE of Estimators | | | | | |
| ***represent most efficient estimator (having minimum MSE and maximum PRE)** | | | | | |

**Table 7.** (MSE) and (PRE) of estimators with respect to $\hat{\mu}'_{YSRS3}$ for using $n = 3, 4, 5, 6$ and $q_3$ with positive correlation

| Correlation | Estimator | n=3 | n=4 | n=5 | n=6 |
|---|---|---|---|---|---|
| $\rho = 0.99$ | $\hat{\mu}'_{YSRS3}$ | 0.15538 | 0.09851 | 0.07115 | 0.05438 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS3(M)}$ | 0.12748 | 0.06367 | 0.04988 | 0.03089 |
| | | 121.88 | 154.74 | 142.64 | 176.06 |
| | $\hat{\mu}'_{K3(M)}$ | 0.17019 | 0.08280 | 0.06496 | 0.03992 |
| | | 91.30 | 118.97 | 109.54 | 136.23 |
| | $\hat{\mu}'_{N3(M)}$ | **0.13735*** | **0.06895*** | **0.05361*** | **0.03335*** |
| | | **113.12*** | **142.89*** | **132.73*** | **163.03*** |
| $\rho = 0.80$ | $\hat{\mu}'_{YSRS3}$ | 0.19485 | 0.13146 | 0.09912 | 0.07844 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS3(M)}$ | 0.1619 | 0.07880 | 0.06320 | 0.03863 |
| | | 120.35 | 166.82 | 156.85 | 203.06 |
| | $\hat{\mu}'_{K3(M)}$ | **0.12187*** | **0.05821*** | **0.04810*** | **0.02907*** |
| | | **159.88*** | **225.85*** | **206.12*** | **269.90*** |
| | $\hat{\mu}'_{N3(M)}$ | 0.15044 | 0.07315 | 0.05893 | 0.03602 |
| | | 129.49 | 179.72 | 168.20 | 217.79 |
| $\rho = 0.70$ | $\hat{\mu}'_{YSRS3}$ | 0.23451 | 0.16460 | 0.12719 | 0.10263 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS3(M)}$ | 0.19195 | 0.09286 | 0.07424 | 0.04542 |
| | | 122.17 | 177.25 | 171.33 | 225.96 |
| | $\hat{\mu}'_{K3(M)}$ | **0.10334*** | **0.04963*** | **0.04074*** | **0.02481*** |
| | | **226.93*** | **331.64*** | **312.17*** | **413.65*** |
| | $\hat{\mu}'_{N3(M)}$ | 0.15691 | 0.07548 | 0.06112 | 0.03727 |
| | | 149.45 | 218.09 | 208.10 | 275.39 |
| $\rho = 0.50$ | $\hat{\mu}'_{YSRS3}$ | 0.31428 | 0.23118 | 0.18340 | 0.15113 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS3(M)}$ | 0.24300 | 0.11807 | 0.09222 | 0.05714 |
| | | 129.33 | 195.81 | 198.88 | 264.49 |
| | $\hat{\mu}'_{K3(M)}$ | **0.09368*** | **0.04586*** | **0.03586*** | **0.02239*** |
| | | **335.49*** | **504.12*** | **511.49*** | **675.08*** |
| | $\hat{\mu}'_{N3(M)}$ | 0.16062 | 0.07734 | 0.06163 | 0.03795 |
| | | 195.66 | 298.92 | 297.59 | 398.21 |
| MSE of Estimators | | | | | |
| PRE of Estimators | | | | | |
| ***represent most efficient estimator (having minimum MSE and maximum PRE)** | | | | | |

**Table 8.** (MSE) and (PRE) of estimators with respect to $\hat{\mu}'_{YSRS3}$ for using $n = 3, 4, 5, 6$ and $q_3$ with negative correlation

| Correlation | Estimator | n=3 | n=4 | n=5 | n=6 |
|---|---|---|---|---|---|
| $\rho = -0.99$ | $\hat{\mu}'_{YSRS3}$ | 0.94752 | 0.75264 | 0.61419 | 0.52321 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS3(M)}$ | 0.61739 | 0.30211 | 0.22966 | 0.14311 |
| | | 153.47 | 249.13 | 267.43 | 365.60 |
| | $\hat{\mu}'_{K3(M)}$ | **0.04589*** | **0.02095*** | **0.01306*** | **0.00842*** |
| | | **2064.80*** | **3592.77*** | **4702.58*** | **6210.80*** |
| | $\hat{\mu}'_{N3(M)}$ | 0.12644 | 0.06595 | 0.04363 | 0.02890 |
| | | 749.41 | 1141.31 | 1407.83 | 1810.45 |
| $\rho = -0.80$ | $\hat{\mu}'_{YSRS3}$ | 0.86080 | 0.68349 | 0.55685 | 0.47373 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS3(M)}$ | 0.53886 | 0.26999 | 0.19753 | 0.12600 |
| | | 159.74 | 253.16 | 281.91 | 375.98 |
| | $\hat{\mu}'_{K3(M)}$ | **0.06459*** | **0.02887*** | **0.02161*** | **0.01287*** |
| | | **1332.75*** | **2367.33*** | **2576.97*** | **3679.86*** |
| | $\hat{\mu}'_{N3(M)}$ | 0.15788 | 0.07594 | 0.05979 | 0.03588 |
| | | 545.22 | 900.09 | 931.41 | 1320.44 |
| $\rho = -0.70$ | $\hat{\mu}'_{YSRS3}$ | 0.81723 | 0.64837 | 0.52778 | 0.44872 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS3(M)}$ | 0.50841 | 0.25640 | 0.18668 | 0.11944 |
| | | 160.74 | 252.88 | 282.73 | 375.70 |
| | $\hat{\mu}'_{K3(M)}$ | **0.07105*** | **0.03197*** | **0.02416*** | **0.01443*** |
| | | **1150.18*** | **2028.15*** | **2184.89*** | **3110.24*** |
| | $\hat{\mu}'_{N3(M)}$ | 0.16234 | 0.07770 | 0.06130 | 0.03678 |
| | | 503.39 | 834.49 | 860.94 | 1219.90 |
| $\rho = -0.50$ | $\hat{\mu}'_{YSRS3}$ | 0.72804 | 0.57617 | 0.46805 | 0.39735 |
| | | 100.00 | 100.00 | 100.00 | 100.00 |
| | $\hat{\mu}'_{YRSS3(M)}$ | 0.45513 | 0.23094 | 0.16833 | 0.10766 |
| | | 159.96 | 249.48 | 278.05 | 369.07 |
| | $\hat{\mu}'_{K3(M)}$ | **0.08034*** | **0.03691*** | **0.02780*** | **0.01679*** |
| | | **906.17*** | **1560.89*** | **1683.76*** | **2366.03*** |
| | $\hat{\mu}'_{N3(M)}$ | 0.16199 | 0.07824 | 0.06071 | 0.03680 |
| | | 449.44 | 736.37 | 770.97 | 1079.70 |
| MSE of Estimators | | | | | |
| PRE of Estimators | | | | | |
| ***represent most efficient estimator (having minimum MSE and maximum PRE)** | | | | | |

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 227 – 244

# Transmuted Dagum distribution: A more flexible and broad shaped hazard function model

Mirza Naveed Shahzad*† and Zahid Asghar‡

## Abstract

In this article, we introduce an extended Dagum distribution, named as transmuted Dagum distribution which can be used for income distribution, actuarial, survival and reliability analysis. The main motivation for generalizing the standard distribution is to provide more flexible distribution to model a variety of data. The extended distribution has been expressed using quadratic rank transmutation map and its tractable properties like moments, moment generating, quantile, reliability and hazard functions are derived. The transmuted Dagum model provides the broader range of hazard behavior than the Dagum model. The densities of its order statistics, generalized TL-moments with its special cases are also studied. The parameters of the new model are estimated by maximum likelihood using Newton-Raphson approach and the information matrix and confidence intervals are also obtained. To illustrate utility and potentiality of the proposed model, it has been applied to rainfall data for the city of Islamabad, Pakistan.

**Keywords:** Transmuted Dagum distribution, Moments, Order statistics, TL-moments, Newton Raphson, Parameter estimation .

*2000 AMS Classification:* 46N30, 47N30, 62G30, 62G32.

*Received* : 12.09.2014 *Accepted* : 05.02.2015 *Doi* : 10.15672/HJMS.2015529452

*Department of Statistics, University of Gujrat, Gujrat, Pakistan.
Email: nvd.shzd@uog.edu.pk
†Corresponding Author.
‡Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
Email: g.zahid@gmail.com

## 1. Introduction

Dagum distribution is widely used for modeling a wide range of data in several fields. It is very worthwhile for analyzing income distribution, actuarial, metrological data and equally preferable for survival analysis. Moreover, it is considered to be the most suitable choice as compared to other three parameter distributions in several cases. It belongs to the generalized Beta distribution and is generated from generalized Beta-II by considering a shape parameter one and referred as inverse Burr distribution. Dagum [5] and Fattorini and Lemmi [13] derived the Dagum distribution independently. Dagum [7] studied the income and income related data by Dagum distributions. Dagum [6] also fitted this distribution on 1978 family incomes data for the United States and proved that its performance is the best among all the models. Bordley, McDonald, and Mantrala [4] also studied the United States family income data by probability distributions along with the Dagum distribution. Bandourian, McDonald, and Turley [2] revealed after the study of 23 countries' income data, the Dagum distribution is the best among the two and three parameter distributions. Quintano and Dagostino [21] studied single-person income distribution of European countries data and found that the Dagum distribution performs better to model the each country data separately. Perez and Alaiz [20] analyzed the personal income data of Spain by Dagum distribution. Alwan, Baharum, and Hassan [1] tried more than fifty distributions to model the reliability of the electrical distribution system and finally the Dagum distribution was considered as the best choice. We have cited very few studies but various other related studies also confirm the better performance of the Dagum distribution. Recently Dagum distribution is found to be quite useful and popular in modeling the skewed data.

Domma, Giordano and Zenga [11] and Domma [8] estimated the parameters of Dagum distribution with censored samples and the right-truncated Dagum distribution by maximum likelihood estimation. McGarvey, et al [17] studied the estimation and skewness test for the Dagum distribution. Shahzad and Asghar [22] estimated the parameter of this distribution by TL-moments. Oluyede and Rajasooriya [18] introduced the Mc-Dagum distribution. Oluyede and Ye [19] presented the class of weighted Dagum and related distributions and Domma and Condino [9] proposed the five parameter beta-Dagum distribution. In this study, we present the transmuted Dagum distribution that is the extension of the Dagum distribution.

Rest of the paper is organized as follows. Section 2 is about the quadratic rank transmutation map, mathematical derivation of the probability density function (pdf) and probability distribution function (cdf) of transmuted Dagum distribution with their graphical presentation. In section 3, $r$th moment and moment generating function are derived and the expression for the coefficient of variation, skewness and kurtosis are also reported. Section 4 is about the quantile function, median and random number generating process for transmuted Dagum distribution. Reliability function, hazard function and their mathematical and graphical presentation are given in Section 5. Section 6 is related to order statistics: the lowest, highest and joint order densities of transmuted Dagum distribution are specified. Section 7 contains the generalized TL-moments and its special cases, such as L-moments, TL-moments, LL-moments, and LH-moments. Methodology for parameter estimation, Newton-Raphson algorithm for maximum likelihood is discussed in Section 8. To compare the suitability of transmuted Dagum distribution with its related distributions, rainfall data is selected and its goodness of fit on empirical data is tested by using likelihood function, AIC, AICC, BIC, KS test, LR test and PP-plots in section 9.

## 2. Transmuted Dagum Distribution

A random variable follows the transmuted distribution, if it satisfies the following relationship that is proposed by Shaw et al. [24] named as quadratic rank transmutation map

$$F(y) = G(y)\big[(1+\lambda) - \lambda G(y)\big] \tag{2.1}$$

Where $G(y)$ is the cdf of the parent distribution and $\lambda$ is the additional parameter that is called transmuted parameter. Due to the transmuted parameter the distribution becomes more flexible distribution to model even the complex data sets.
The pdf of the Dagum (parent) distribution is as

$$g(y;\alpha,\beta,\rho) = \frac{\alpha\rho y^{\alpha\rho-1}}{\beta^{\alpha\rho}\big(1+(y/\beta)^{\alpha}\big)^{\rho+1}}, \quad 0 \le x \le \infty;\ \alpha,\beta,\rho > 0 \tag{2.2}$$

and its cdf is as

$$G(y;\alpha,\beta,\rho) = \big(1+(y/\beta)^{-\alpha}\big)^{-\rho}. \tag{2.3}$$

Where $\alpha$ and $\rho$ are the shape parameters, $\beta$ is the scale parameter and all the three parameters are positive. Now substituting the (2.3) in (2.1), we obtained the cdf of the transmuted Dagum distribution in the following form

$$F(y;\alpha,\beta,\rho,\lambda) = \big(1+(y/\beta)^{-\alpha}\big)^{-\rho}[1+\lambda-\lambda\big(1+(y/\beta)^{-\alpha}\big)^{-\rho}], \tag{2.4}$$

and its respective pdf of transmuted Dagum distribution is given by

$$f(y;\alpha,\beta,\rho,\lambda) = \frac{\alpha\rho y^{2\alpha\rho-1}\big[(1+\lambda)\big(1+(y/\beta)^{-\alpha}\big)^{\rho} - 2\lambda\big]}{\beta^{2\alpha\rho}\big(1+(y/\beta)^{\alpha}\big)^{2\rho+1}}. \tag{2.5}$$

The parameter $\lambda$ has the support $-1 \le y \le 1$ and simply taking $\lambda = 0$ in above pdf and cdf, transmuted distribution reduces to the parent distribution. Dagum distribution due to quadratic rank transmutation map becomes more flexible. The shapes of this density and distribution function assuming various combinations of parameters are illustrated in the Figure 1 and Figure 2, respectively.

## 3. Moments and moments ratio

In this section, main statistical properties such as $r$th moments, mean, variance, and moment generating function for transmuted Dagum distribution are derived and discussed.

**3.1. Theorem.** *Let the random variable $Y$ follows the transmuted Dagum distribution, then its $r$th moment has the following form*

$$E(Y^r) = \beta^r\Gamma\Big(1-\frac{r}{\alpha}\Big)\Big[\frac{(1+\lambda)\Gamma(\rho+\frac{r}{\alpha})}{\Gamma(\rho)} - \frac{\lambda\Gamma(2\rho+\frac{r}{\alpha})}{\Gamma(2\rho)}\Big]. \tag{3.1}$$

*Proof.* Let the $r$th moments is given by

$$m_r^{'} = E(Y^r) = \int_0^{\infty} \frac{\alpha\rho y^{2\alpha\rho+r-1}\big[(1+\lambda)\big(1+(y/\beta)^{-\alpha}\big)^{\rho} - 2\lambda\big]}{\beta^{2\alpha\rho}\big(1+(y/\beta)^{\alpha}\big)^{2\rho+1}}\,\mathrm{d}y$$

$$= \int_0^{\infty} \frac{\alpha\rho y^{2\alpha\rho+r-1}(1+\lambda)\big(1+(y/\beta)^{-\alpha}\big)^{\rho}}{\beta^{2\alpha\rho}\big(1+(y/\beta)^{\alpha}\big)^{2\rho+1}}\,\mathrm{d}y - \int_0^{\infty} \frac{2\lambda\alpha\rho y^{2\alpha\rho+r-1}}{\beta^{2\alpha\rho}\big(1+(y/\beta)^{\alpha}\big)^{2\rho+1}}\,\mathrm{d}y$$

For convenience substitute $x=(y/\beta)^{\alpha}$, hence

$$m_r^{'} = \beta^r(1+\lambda)\int_0^{\infty} x^{\rho+r/\alpha-1}(1+x)^{-(\rho+1)}\mathrm{d}x - 2\alpha\rho\beta^r\int_0^{\infty} x^{\rho+r/\alpha-1}(1+x)^{-(\rho+1)}\mathrm{d}x$$

$$= \beta^r\Big[(1+\lambda)B\Big(\rho+\frac{r}{\alpha},1-\frac{r}{\alpha}\Big) - \lambda B\Big(2\rho+\frac{r}{\alpha},1-\frac{r}{\alpha}\Big)\Big],$$

where $B(.,.)$ is the beta type-II function, defined by

$$B(\theta_1,\theta_2) = \int_0^{\infty} z_1^{\theta}(1+z)^{-(\theta_1+\theta_2)}\mathrm{d}z; \quad \theta_1,\theta_2 > 0$$

after one step simplification, we obtain the result given in (3.1).

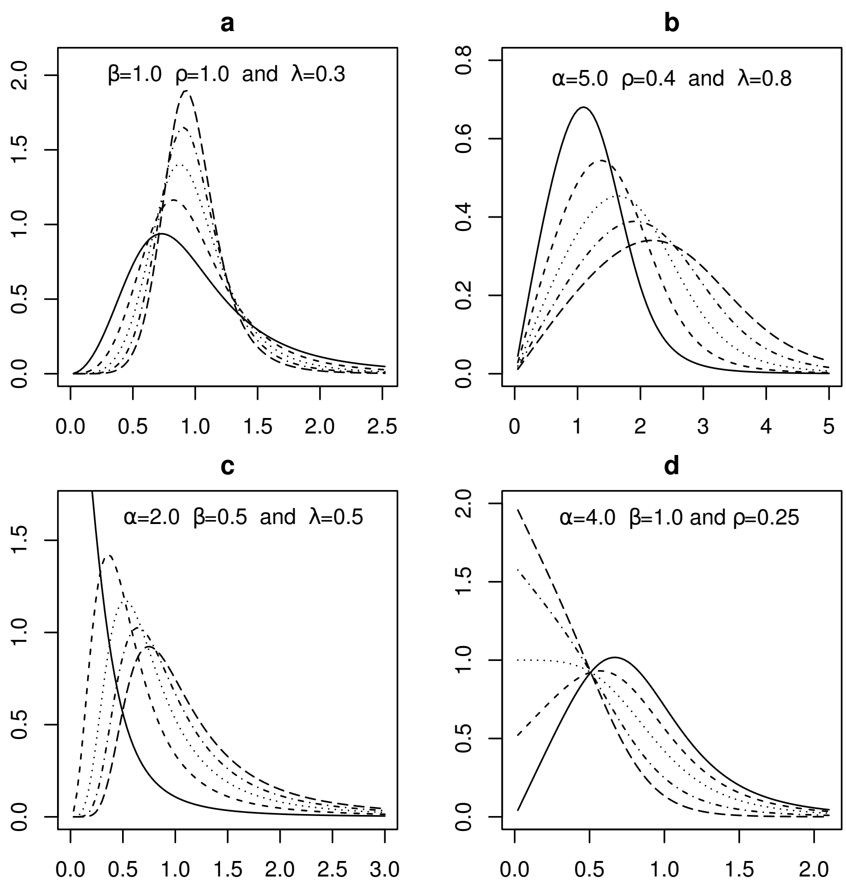



**Figure 1.** The pdf's of various transmuted Dagum distributions for values of parameters: *a)* $\alpha = 3.0[0.5]5.5$; *b)* $\beta = 2.0[0.25]3.25$; *c)* $\rho =$ 0.5, 0.75, 1.0[1.0]4.0; *d)* $\lambda =$ -1.0[0.4]1.0 with solid, dashed, dotted, dotdash and longdash lines, respectively.

In particular, by setting $r = 1$ and $r = 2$ in (3.1), we obtain mean and varianc $(\sigma^2)$ by taking usual steps

$$E(Y) = \beta\Gamma\Big(1-\frac{1}{\alpha}\Big)\Big[\frac{(1+\lambda)\Gamma(\rho+\frac{1}{\alpha})}{\Gamma(\rho)} - \frac{\lambda\Gamma(2\rho+\frac{1}{\alpha})}{\Gamma(2\rho)}\Big] \quad (3.2)$$

and

$$\sigma^2 = \beta^2\Gamma\Big(1-\frac{2}{\alpha}\Big)\big[(1+\lambda)P_{12}-\lambda P_{22}\big] - \beta^2\Big[\Gamma\Big(1-\frac{1}{\alpha}\Big)\Big]^2\big[(1+\lambda)P_{11}-\lambda P_{21}\big], \tag{3.3}$$

where $P_{ir} = \Gamma(i\rho + \frac{r}{\alpha})/\Gamma(i\rho)$.
The following expression can be used to obtain the moment ratios for transmuted Dagum distribution such as Cofficient of vartiation $(CV)$, Skewness $(Sk)$ and Kurtosis $(Kr)$ using $m_r^{'}(r = 1, 2, 3, 4)$.

$$CV = \frac{\sigma}{m_1'},$$

$$Sk = \frac{m_3' - 2m_2'm_1' + 2\left(m_1'\right)^3}{\sigma^3},$$

$$Kr = \frac{m_4' - 4m_3'm_1' + 6m_2'\left(m_1'\right)^2 - 3\left(m_1'\right)^4}{\sigma^4}.$$

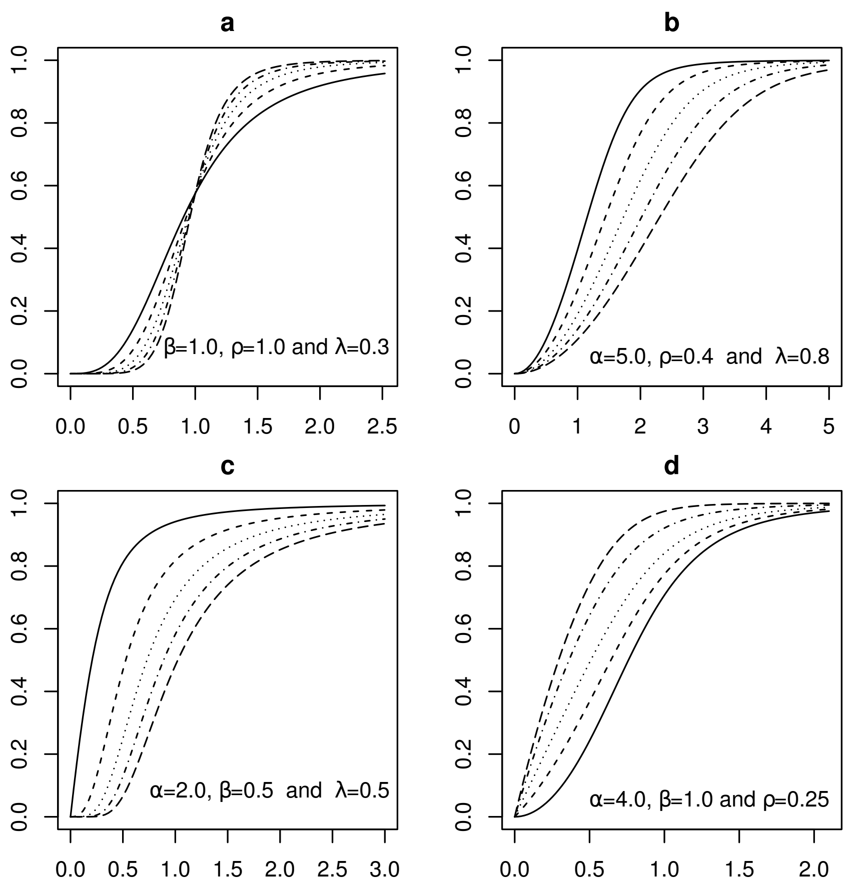



**Figure 2.** The cdf's of various transmuted Dagum distributions for values of parameters: *a*) $\alpha$ = 3.0[0.5]5.5; *b*) $\beta$ = 2.0[0.25]3.25; *c*) $\rho$ = 0.5, 0.75, 1.0[1.0]4.0; *d*) $\lambda$ = -1.0[0.4]1.0 with solid, dashed, dotted, dotdash and longdash lines, respectively.

**3.2. Theorem.** *The moment generating function of $Y$ , $M_y(t)$ when random variable follows the transmuted Dagum distribution is*

$$M_y(t) = \sum_{r=0}^{\infty} \frac{t^r \beta\Gamma(1-1/\alpha)}{r!} \big[(1+\lambda)P_{11} - \lambda P_{21}\big] \tag{3.4}$$

*Proof.* Let the moment generating function for $Y$ is given by

$$\begin{aligned} M_Y(t) = E(e^{ty}) &= \int_0^{\infty} e^{ty} f(y)\mathrm{d}y \\ &= \int_0^{\infty} \left(1 + ty + \frac{t^2y^2}{2!} + ... + \frac{t^ny^n}{n!} + ...\right) f(y)\mathrm{d}y \\ &= \sum_{r=0}^{\infty} \frac{t^r E\left(Y^r\right)}{r!} \\ &= \sum_{r=0}^{\infty} \frac{t^r \beta\Gamma(1-1/\alpha)}{r!} \big[(1+\lambda)P_{11} - \lambda P_{21}\big]. \end{aligned}$$

## 4. Quantile function and random number generation

Hyndman and Fan [16] defined the quantile function for any distribution is in the form

$$Q(q) = F^{-1}(q) = \inf\{y : F(y) \geqslant q\} \quad 0 < q < 1, \tag{4.1}$$

where $F(y)$ is the distribution function. Quantile function divides the ordered data into $q$ equal sized portions. The smallest and largest value of the ordered data corresponds to probability 0 and 1, respectively. The $q$th quantile of transmuted Dagum distribution is obtained using (2.4) and (4.1) is given as

$$Q(q) = \beta\Big[\Big(\frac{1+\lambda+\sqrt{(1+\lambda)^2 - 4\lambda q}}{2q}\Big)^{1/\rho} - 1\Big]^{1/\alpha}. \tag{4.2}$$

Median is the 50th percentile, hence median of transmuted Dagum distribution is obtained from (4.2) as below

$$\text{Median} = \beta\Big[\Big(1+\lambda+\sqrt{1+\lambda^2}\Big)^{1/\rho} - 1\Big]^{1/\alpha}.$$

The expression (4.2) can also be used to find the tertiles, quartiles, quintiles, sextiles, deciles, percentiles and permilles. To generate the random numbers for the transmuted Dagum distribution, let suppose that the $U$ is the standard uniform variate in (4.2) rather than $q$. Then the random variable

$$y = \beta\Big[\Big(\frac{1+\lambda+\sqrt{(1+\lambda)^2 - 4\lambda u}}{2u}\Big)^{1/\rho} - 1\Big]^{1/\alpha} \tag{4.3}$$

follows the transmuted Dagum distribution. Now (4.3) is ready to generate the random number for the distribution, taking $\alpha$, $\beta$, $\rho$ and $\lambda$ known.

## 5. Reliability analysis

The reliability function $R(t)$ gives the probability of surviving of an item at least reach the age of $t$ time. The cdf $F(t)$ and reliability function are reverse of each other as $R(t) + F(t) = 1$. The reliability function for transmuted Dugam distribution is given by

$$\begin{aligned} R(t) = P(T > t) &= \int_t^{\infty} f(t)\mathrm{d}t = 1 - F(t) \\ &= 1 + \big(1 + (t/\beta)^{-\alpha}\big)^{-2\rho}\Big[\lambda - (1+\lambda)\big(1 + (t/\beta)^{-\alpha}\big)^{\rho}\Big]. \end{aligned}$$

With various choices of parametric values the Figure 3 illustrates the reliability function pattern of transmuted Dagum distribution.

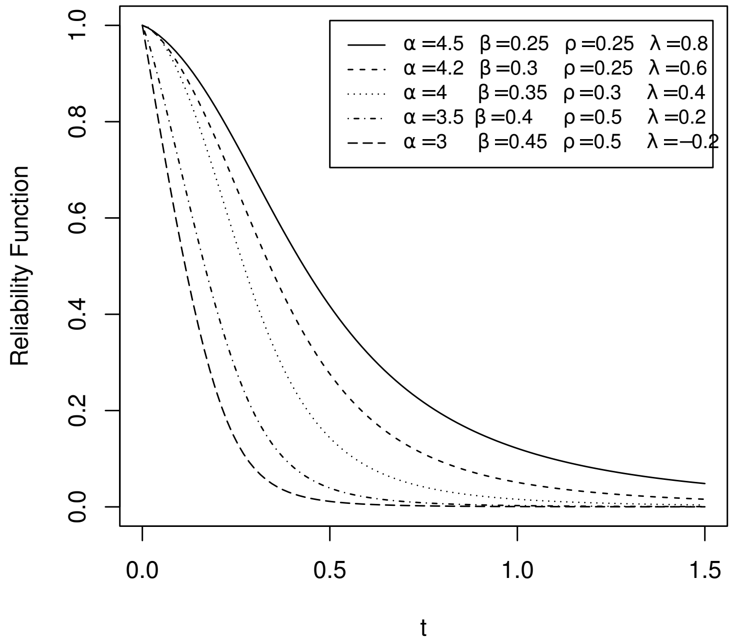



**Figure 3.** The various shapes of reliability function of transmuted Dagum distribution.

An important property of a random variable is the hazard function, it measure the inclination towards failure rate. The probability approaches to failure increases as the value of the hazard function increase. Mathematically, the hazard function and the hazard function of transmuted Dagum distribution is defined as

$$h(t) = \frac{f(t)}{1-F(t)} = \frac{f(t)}{R(t)}$$
$$= \frac{\alpha\rho t^{2\alpha\rho-1}\left[(1+\lambda)\left(1+(t/\beta)^{-\alpha}\right)^{\rho} - 2\lambda\right]}{\beta^{2\alpha\rho}\left(1+(t/\beta)^{\alpha}\right)^{2\rho+1}\left[1+\left(1+(t/\beta)^{-\alpha}\right)^{-2\rho}\left[\lambda-(\lambda+1)\left(1+(t/\beta)^{-\alpha}\right)^{\rho}\right]\right]}.$$

The hazard function of the transmuted Dagum distribution is attractively flexible. Therefore, it is useful and suitable for the real life situations. As in the case of transmuted Dagum distribution when $\lambda = 0$ is the Dagum distribution. Domma [10] and Domma, Giordano and Zenga [11] using Glaser's theorem [14] proved the proposition of the hazard function of the Dagum distribution. So taking these propositions and Glaser's theorem [14], we concentrate on the additional parameter $\lambda$ and find out the following four behaviour of the hazard function on the combinations of parameters.

(1) The hazard function of transmuted Dagum distribution is decreasing if
   (a) $\rho = 2/\alpha - 1$, $\alpha < 2$, $\beta > 0$ and -1$\leq \lambda \leq$1.
   (b) $\alpha\rho = 1$, $\rho < 2/\alpha - 1$, $\alpha < 1$, $\beta > 0$ and -1$\leq \lambda \leq$1
   (c) $\alpha < 1$, $\rho(\alpha^{-1}, 2/\alpha - 1)$, $\beta > 0$ and -1$\leq \lambda \leq$1
(2) It is upside down bathtub (increasing-decreasing) if

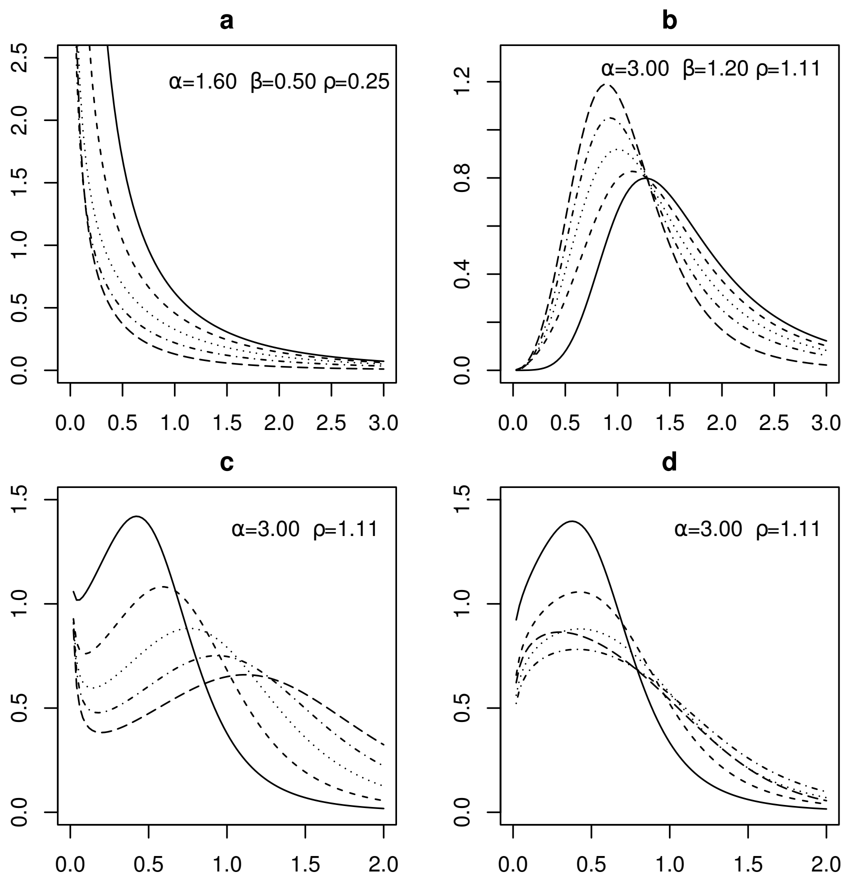



**Figure 4.** The behaviour of the hazard rate function of the transmuted Dagum distributions for various parameters values such as : *a*) $\lambda$ = -1.0[0.5]1.0; *b*) $\lambda$ = -1.0[0.5]1.0; *c*) $\beta$ = 0.75[0.25]1.75, $\lambda$ = -0.8[0.1]-0.4; *d*) $\beta$ = 0.75[0.25]1.75, $\lambda$ = -0.2[0.2]0.8 with solid, dashed, dotted, dotdash and longdash lines, respectively.

   (a) $\alpha\rho > 1$, $\rho \neq 2/\alpha - 1$, $\beta > 0$ and -1$\leq \lambda \leq$1
   (b) $\alpha\rho = 1$, $\rho > 2/\alpha - 1$, $\alpha > 1$, $\beta > 0$ and -1$\leq \lambda \leq$1

(3) It is bathtub and upside down bathtub if
   (a) $\alpha \in (1, 3)$, $\rho \in \left(\frac{3-\alpha}{\alpha+1}, \frac{2}{\alpha} - 1\right)$, $\beta > 0$ and -1$\leq \lambda <$ -0.4
   (b) $\alpha \geq 3$, $\rho \in \left(\frac{2}{\alpha} - 1\right)$, $\beta > 0$ and -1$\leq \lambda <$ -0.4

(4) It is upside down bathtub if
   (a) $\alpha \in (1, 3)$, $\rho \in \left(\frac{3-\alpha}{\alpha+1}, \frac{2}{\alpha} - 1\right)$, $\beta > 0$ and -0.4$\leq \lambda \leq$1
   (b) $\alpha \geq 3$, $\rho \in \left(\frac{2}{\alpha} - 1\right)$, $\beta > 0$ and -0.4$\leq \lambda \leq$1

The graphical pesentation of the behaviour of the hazard rate function for transmuted Dagum distribution is sketched in Figure 4 for various choices of parametric values.

## 6. Order statistics of transmuted Dagum distribution

In probability statistics the distribution of extremes (smallest and/or largest), median and joint order statistics are the most important functions of a random variable. This is only the order statistics that help us to study the peaks of the data to understand the pattern of the extremes. Mathematically the order statistics is defined as, let $Y_1, Y_2, ..., Y_n$ be any real valued random variables and its ordered values denoted as $Y_{(1)} \leq Y_{(2)} \leq ... \leq Y_{(n)}$ then the values $Y_{(1)}, Y_{(2)}, ..., Y_{(n)}$ are the order statistics of random variable .
The density of the $n$th ordered statistics, that follows the transmuted Dagum distribution is derived in the following form

$$\begin{aligned} f_{(n)}(y_{(n)}) &= n[F(y_{(n)})]^{n-1} f_(y_{(n)}) \\ &= \frac{n\alpha\rho\beta^{\alpha}}{y_{(n)}^{\alpha+1}} \Big[(1+\lambda)\Big(1+\Big(\frac{y_{(n)}}{\beta}\Big)^{-\alpha}\Big) - 2\lambda\Big] \sum_{j=0}^{n-1} \binom{n-1}{j} \\ &\times (-\lambda)^j (1+\lambda)^{n-j-1} \Big(1+\Big(\frac{y_{(n)}}{\beta}\Big)^{-\alpha}\Big)^{-\rho(n+j+1)-1}. \end{aligned}$$

Let suppose that the smallest values also follows the transmuted Dagum distribution, then the denity of the smallest order statistic, is obtained as

$$\begin{aligned} f_{(1)}(y_{(1)}) &= n[1-F(y_{(1)})]^{n-1} f(y_{(1)}) \\ &= \frac{n\alpha\rho\beta^{\alpha}}{y_{(1)}^{\alpha+1}} \Big[(1+\lambda)\Big(1+\Big(\frac{y_{(1)}}{\beta}\Big)^{-\alpha}\Big) - 2\lambda\Big] \sum_{i=0}^{n-1}\sum_{j=0}^{i} \binom{i}{j}\binom{n-1}{i} \\ &\times (-1)^{i+j}(-\lambda)^j (1+\lambda)^{i-j} \Big(1+\Big(\frac{y_{(1)}}{\beta}\Big)^{-\alpha}\Big)^{-\rho(i+j+1)-1}. \end{aligned}$$

Generally the pdf of the $r$th order statistics is given by

$$\begin{aligned} f_{(r)}(y_{(r)}) &= \frac{n!}{(r-1)!(n-r)!}[F(y_{(r)})]^{r-1}[1-F(y_{(r)})]^{n-r} f_(y_{(r)}) \\ &= \frac{n!\alpha\rho\beta^{\alpha} y_{(r)}^{-(\alpha+1)}}{(r-1)!(n-r)!} \Big[(1+\lambda)\Big(1+\Big(\frac{y_{(r)}}{\beta}\Big)^{-\alpha}\Big) - 2\lambda\Big] \sum_{i=0}^{n-r}\sum_{j=0}^{r+i-1} \binom{r+i-1}{j} \\ &\times \binom{n-r}{i} (-1)^{i+j}(\lambda)^j (1+\lambda)^{r+i-j-1} \Big(1+\Big(\frac{y_{(r)}}{\beta}\Big)^{-\alpha}\Big)^{-\rho(r+i+j+1)-1}. \end{aligned}$$

Sometimes interest is in the joint pdf such as to find the joint breaking strength of certain equipment, for the transmuted Dagum distribution the pdf of $Y_{(r)}$ and $Y_{(s)}$, when $1 \leq r < s \leq n$ is obtained as

$$\begin{aligned} f_{(r,s)}(u,v) &= \frac{n!}{(r-1)!(s-r-1)!(n-s)!}[F(u)]^{r-1}[F(v)-F(u)]^{s-r-1} \\ &\times [1-F(v)]^{n-s} f(u) f(v) \\ &= \frac{n!}{(r-1)!(s-r-1)!(n-s)!}\Big(\frac{\alpha\beta\rho}{uv}\Big)^2 \Big[(1+\lambda)\Big(1+\Big(\frac{u}{\beta}\Big)^{-\alpha}\Big) - 2\lambda\Big] \\ &\times \sum_{i=0}^{S}\sum_{j=0}^{s-r-1}\sum_{k=0}^{n-s}\sum_{l=0}^{r+j-1} \binom{r+j-1}{l}\binom{n-s}{k}\binom{s-r-1}{j}\binom{S}{j} \\ &\times (-1)^{i+j+k+l}(\lambda)^{i+l}(1+\lambda)^{r+i+j-1}\Big(1+\Big(\frac{u}{\beta}\Big)^{-\alpha}\Big)^{-\rho(r+i+j+1)-1} \\ &\times \Big(1+\Big(\frac{v}{\beta}\Big)^{-\alpha}\Big)^{-\rho(S+i+2)-1}, \end{aligned}$$

where $S = s + k - r - j - 1$.

## 7. Generalized TL-moments of transmuted Dagum distribution

Hosking [15] introduced the L-moments and now these moments are frequently used for extreme value analysis. Elamir and Seheult [12] extended these moments and presented the TL-moments. These moments based on the order statistics used to describe the shape of the probability distribution by evaluating all descriptive statistics including parameter estimation and hypothesis testing. The $r$th generalized TL-moments with $s$ smallest and $t$ largest trimming is defined as follows

$$T_r^{(s,t)} = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E\left(Y_{r+s-k:r+s+t}\right), \tag{7.1}$$

where $T_r^{(s,t)}$ is a linear function of the expectations of the order statistics and $r = 1, 2, ...; t, s = 0, 1, 2....$

The expression for the expected value of the $(r + s - k)$th order statistics of the random sample of size $(r + s + t)$ is as

$$E\left(Y_{r+s-k:r+s+t}\right) = C \int_0^\infty [F(y)]^{r+s-k-1} [1 - F(y)]^{t+k} \, dF(y). \tag{7.2}$$

where $C = \frac{(r+s+t)!}{(r+s-k-1)!(t+k)!}$ and $F$ is the cdf of the transmuted Dagum distribution, and by substitute expression (7.1) into expression (7.2), we obtain $T_r^{(s,t)}$ as

$$T_r^{(s,t)} = \sum_{k=0}^{r-1} \binom{r-1}{k} \frac{C}{r} (-1)^k \int_0^\infty [F(y)]^{r+s-k-1} [1 - F(y)]^{t+k} \, dF(y)$$

Having the cdf and pdf of transmuted Dagum distribution the generalized TL-moments is given by

$$\begin{aligned} T_r^{(s,t)} = {} & \frac{\beta\rho}{r} \sum_{k=0}^{r-1} \sum_{i=0}^{t+k} \sum_{j=0}^{s+r-k+i-1} \binom{s+r-k+i-1}{j} \binom{t+k}{i} \\ & \times \binom{r-1}{k} C(-1)^{i+j+k} \lambda^j (1+\lambda)^{I+s-1} \Gamma(1 - 1/\alpha) \\ & \times \left[(1+\lambda) \frac{\Gamma[\rho(I+s) + 1/\alpha]}{\Gamma[\rho(I+s) + 1]} - 2\lambda \frac{\Gamma[\rho(I+s+1) + 1/\alpha]}{\Gamma[\rho(I+s+1) + 1]}\right], \end{aligned} \tag{7.3}$$

where $I = r - k + i + j$.

This expression of the generalized TL-moments used to obtain its special cases such as L-moments, TL-moments, LH-moments and LL-moments. First two TL-moments $T_1^{(s,t)}$ and $T_2^{(s,t)}$ are used to calculate the location and dispersion of the data, respectively. The ratio of TL-moments $T_{CV}^{(s,t)} = T_2^{(s,t)} \big/ T_1^{(s,t)}$, $T_{Sk}^{(s,t)} = T_3^{(s,t)} \big/ T_2^{(s,t)}$ and $T_{Kr}^{(s,t)} = T_4^{(s,t)} \big/ \lambda_2^{(s,t)}$ are the coefficient of variation, skewness and kurtosis characteristic of the probability distribution, respectively.

**7.1. The TL-moments** $(s = t = 1)$**.** Generally it is possible to trim any number of smallest and largest values from the ordered observation. As a special case, if only one extreme value from both sides $(s = t = 1)$ are trimmed then expression (7.3) becomes the $r$th TL-moments and we get

$$T_r^{(1)} = \sum_{k=0}^{r-1}\sum_{i=0}^{k+1}\sum_{j=0}^{r-k+i} \binom{r-k+i}{j}\binom{k+1}{i}\binom{r-1}{k} \frac{(r+2)!\Gamma(1-1/\alpha)\beta\rho}{r(r-k)!(k+1)!} \times (-1)^{i+j+k}\lambda^j(1+\lambda)^I\Big[(1+\lambda)\frac{\Gamma[\rho(I+1)+1/\alpha]}{\Gamma[\rho(I+1)+1]} - 2\lambda\frac{\Gamma[\rho(I+2)+1/\alpha]}{\Gamma[\rho(I+2)+1]}\Big].$$

**7.2. The L-moments** $(s = t = 0)$**.** When none of the observation is trimmed from the ordered sample, the TL-moments reduced to L-moments and basically L-moments and related moments are due the Hosking [15] methodology. The $r$th L-moments of transmuted Dagum distribution is as

$$T_r^{(0)} = \sum_{k=0}^{r-1}\sum_{i=0}^{k}\sum_{j=0}^{r-k+i-1} \binom{r-k+i-1}{j}\binom{k}{i}\binom{r-1}{k} \frac{(r)!\Gamma(1-1/\alpha)\beta\rho}{r(r-k-1)!(k)!} \times (-1)^{i+j+k}\lambda^j(1+\lambda)^{I-1}\Big[(1+\lambda)\frac{\Gamma[\rho(I)+1/\alpha]}{\Gamma[\rho(I)+1]} - 2\lambda\frac{\Gamma[\rho(I+1)+1/\alpha]}{\Gamma[\rho(I+1)+1]}\Big].$$

**7.3. The LL-moments** $(s = 0, t = t)$**.** LL-moments progressively reflect the characteristics of the lower part of distribution. Bayazit and Onoz [3] introduced these moments and later it became the special case of the TL-moments, when $s = 0$ and $t = t$. Following is the LL-moments

$$T_r^{(0,t)} = \sum_{k=0}^{r-1}\sum_{i=0}^{t+k}\sum_{j=0}^{r-k+i-1} \binom{r-k+i-1}{j}\binom{t+k}{i}\binom{r-1}{k} \times \frac{(r+t)!\beta\rho}{r(r-k-1)!(t+k)!}(-1)^{i+j+k}\lambda^j(1+\lambda)^{I-1}\Gamma(1-1/\alpha) \times \Big[(1+\lambda)\frac{\Gamma[\rho(I)+1/\alpha]}{\Gamma[\rho(I)+1]} - 2\lambda\frac{\Gamma[\rho(I+1)+1/\alpha]}{\Gamma[\rho(I+1)+1]}\Big].$$

**7.4. The LH-moments** $(s = s, t = 0)$**.** LH moments proposed by Wang [26], these moments describe the upper part of the data more precisely. These moments give more weight to the larger values and the theoretical LH-moments for the transmuted Dagum distribution are defined as

$$T_r^{(s,0)} = \sum_{k=0}^{r-1}\sum_{i=0}^{k}\sum_{j=0}^{r+s-k+i-1} \binom{r+s-k+i-1}{j}\binom{k}{i}\binom{r-1}{k} \times \frac{(r+t)!\beta\rho}{r(r-k-1)!(t+k)!}(-1)^{i+j+k}\lambda^j(1+\lambda)^{I-1}\Gamma(1-1/\alpha) \times \Big[(1+\lambda)\frac{\Gamma[\rho(I+s)+1/\alpha]}{\Gamma[\rho(I+s)+1]} - 2\lambda\frac{\Gamma[\rho(I+s+1)+1/\alpha]}{\Gamma[\rho(I+s+1)+1]}\Big].$$

## 8. Parameter estimation

In this section, interest is to estimate the parameters of transmuted Dagum distribution by maximum likelihood estimation. Let $Y_1, Y_2, ..., Y_n$ be i.i.d random variables of transmuted Dagum distribution of size $n$. Then the sample likelihood function and log-likelihood function for this distribution are obtained as follows

$$L(x;.) = \frac{\alpha\rho}{\beta^{2\alpha\rho}} \prod_{i=1}^{n} y_i^{2\alpha\rho-1} \big(1+(y_i/\beta)^{\alpha}\big)^{2\rho+1} \big[(1+\lambda)\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho} - 2\lambda\big] \tag{8.1}$$

and

$$\begin{aligned} \ell(x;.) &= n\ln\alpha + n\ln\rho - 2n\alpha\rho\ln\beta - (2\alpha\rho+1)\sum_{i=1}^{n}\ln\big(1+(y_i/\beta)^{\alpha}\big) \\ &+ (2\alpha\rho+1)\sum_{i=1}^{n}\ln y_i + \sum_{i=1}^{n}\ln\big[(1+\lambda)\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho} - 2\lambda\big], \end{aligned} \tag{8.2}$$

respectively.
To find the parameter estimates, now we take the first order derivatives of (8.2) with respect to parameter ($\alpha, \beta, \rho$ and $\lambda$) and equating them equal to zero, respectively,

$$\begin{aligned} &\frac{n}{\alpha} - 2n\rho\ln\beta + 2\rho\sum_{i=1}^{n}\ln y_i - (2\rho+1)\sum_{i=1}^{n}\frac{(y_i/\beta)^{\alpha}\ln(y_i/\beta)}{\big(1+(y_i/\beta)^{\alpha}\big)} \\ &\qquad - \rho(1+\lambda)\sum_{i=1}^{n}\frac{(y_i/\beta)^{-\alpha}\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho-1}\ln(y_i/\beta)}{\big[(1+\lambda)\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho} - 2\lambda\big]} = 0, \end{aligned}$$

$$\begin{aligned} &-\frac{2n\alpha\rho}{\beta} + \alpha(2\rho+1)\sum_{i=1}^{n}\frac{y_i(y_i/\beta)^{\alpha-1}}{\beta^2\big(1+(y_i/\beta)^{\alpha}\big)} \\ &\qquad - \alpha\lambda(1+\lambda)\sum_{i=1}^{n}\frac{(y_i/\beta)^{-\alpha}\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho-1}\ln(y_i/\beta)}{\big[(1+\lambda)\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho} - 2\lambda\big]} = 0, \end{aligned}$$

$$\begin{aligned} &\frac{n}{\rho} - 2n\alpha\ln\beta - 2\sum_{i=1}^{n}\ln\big(1+(y_i/\beta)^{\alpha}\big) \\ &\qquad + (1+\lambda)\sum_{i=1}^{n}\frac{\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho}\ln\big(1+(y_i/\beta)^{-\alpha}\big)}{\big[(1+\lambda)\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho} - 2\lambda\big]} = 0, \end{aligned}$$

$$\sum_{i=1}^{n}\frac{\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho} - 2}{\big[(1+\lambda)\big(1+(y_i/\beta)^{-\alpha}\big)^{\rho} - 2\lambda\big]} = 0.$$

The exact solution to derive the estimator for unknown parameters is not possible, so the estimates $(\hat{\alpha}, \hat{\beta}, \hat{\rho}, \hat{\lambda})'$ are obtained by solving the above four nonlinear equations simultaneously. This solution of nonlinear system is easier by Newton-Raphson approach. The Newton-Raphson approach used the $j$th element of the gradient and the $(j,k)$th elements of the Hessian matrix and these elements are $g_j = \partial\ell(\theta)/\partial\theta_j$ and $H_{jk} = \partial^2\ell(\theta)/\partial\theta_j\partial\theta_k$, respectively, whereas $j,k$ =1,2,3,4, due to the four parameters of transmuted Dagum distribution. The information matrix, $I(\theta) = I_{jk}(\theta) = -E(H_{jk})$ and then its inverse of matrix $I(\theta)^{-1}$ provides the variances and covariances, diagonal and off diagonal entries, respectively. Asymptotically these estimates of parameters approaches to normality and the z-score are approximately standard normal, which can be used to find the $100(1-r)\%$ two sided confidence interval for $\alpha$, $\beta$, $\rho$ and $\lambda$.

## 9. Application

In this section, the performance of the transmuted Dagum distribution is compared with Dagum distribution and some other related distributions. Monthly maximum precipitation data of Islamabad city is considered for the comparison. Islamabad is the capital city of the Pakistan. The geographical location of this city has Latitude 33.71 and Longitude 73.07 with humid subtropical climate and has five seasons. This area receives heavy rainfall during monsoon season. The data of monthly precipitation retrieved from the Regional Meteorological Center (RMC) Lahore and from Pakistan Metrological Department (PMD) Islamabad. The length of data is 640 recorded from January 1954 to December 2013 excluding some unobserved or unreported months and the summary statistics are given in Table 1 and Table 3.

**Table 1.** Summary Statistics for monthly maximum precipitation data of the Islamabad, Pakistan.

| Length | Average | Minimum | Maximum | $Q_1$ | Median | $Q_3$ | S.D |
|---|---|---|---|---|---|---|---|
| 640 | 86.25 | 0.10 | 641.00 | 20.35 | 49.90 | 101.90 | 94.98 |

In order to compare the transmuted Dagum and its related distribution, we consider criteria like log-likelihood ($\ell$), Akaike information criterion (AIC), Akaike information corrected criterion (AICC), Bayesian information criteria (BIC) and Kolmogoro-Smirnov (KS) goodness of fit test for the data sets. The better distribution have corresponds to smaller $\ell$, AIC, AICC , BIC and KS values. Where

$$\text{AIC}=2\text{k}-2\ell,$$
$$\text{AICC}=\text{AIC}+2k(k+1)/(n-k-1),$$
$$\text{BIC}=2\ell+\text{k}\log(n)$$

and

$$\text{KS}=\max_{i\leq i\leq n}[F(Y_i)-(i-1)/n, i/n-F(Y_i)].$$

Here $k$ is the number of parameters in each distribution, and $n$ is the sample size.
It is better to test the superiority of the transmuted Dagum distribution over the Dagum distribution before analyzing the data. We employed the likelihood ratio (LR) statistic for this purpose. To perform this test the maximized restricted and unrestricted log-likelihoods can be computed under the null and alternative hypothesis
$H_0 : \lambda = 0$ (restricted, Dagum model is true for the data set)
versus
$H_1 : \lambda \neq 0$ (unrestricted, transmuted Dagum model is true for the data set).
The LR statistic for testing the hypothesis is computed by $\omega = 2(\ell(\hat{\theta_0}) - \ell(\hat{\theta_1}))$, where $\hat{\theta_0}$ and $\hat{\theta_1}$ are the maximum likelihood estimates under $H_0$ and $H_1$, respectively. The LR statistic is asymptotically distributed as chi-square $(\chi^2_{v,r})$. The computed value of LR statistic under the hypothesis is $\omega$ =22.74. We may observe that the $\omega > \chi^2_{1,0.05}(3.84)$, so we reject the null hypothesis and found that the transmuted Dagum model is best for the data set.
Variance covariance matrix of the MLEs under the transmuted Dagum distribution is obtained as

$$I(\hat{\theta})^{-1} = \begin{pmatrix} 0.0407 & 2.7555 & -0.0097 & 0.0094 \\ 2.7555 & 1360.2 & -0.5489 & 13.361 \\ -0.0097 & -0.5489 & 0.0028 & 0.0004 \\ 0.0094 & 13.361 & 0.0004 & 0.4924 \end{pmatrix}.$$

**Table 2.** Estimated parameters of the transmuted Dagum and related distributions.

| Model | Estimates | $\ell(.,y)$ | AIC | AICC | BIC | KS |
|---|---|---|---|---|---|---|
| Transmuted Dagum | $\hat{\alpha} = 2.2198$<br>$\hat{\beta} = 132.94$<br>$\hat{\rho} = 0.3981$<br>$\hat{\lambda} = 0.3565$ | 3452.71 | 6913.42 | 6913.48 | 6931.27 | 0.0280 |
| Dagum | $\hat{\alpha} = 2.1302$<br>$\hat{\beta} = 105.49$<br>$\hat{\rho} = 0.5827$ | 3464.08 | 6934.16 | 6934.20 | 6947.544 | 0.1646 |
| Transmuted Pareto | $\hat{a} = 0.1000$<br>$\hat{b} = 0.2374$<br>$\hat{\lambda} = -0.962$ | 4002.16 | 8010.32 | 8010.36 | 8023.70 | 0.3332 |
| Pareto | $\hat{a} = 0.1000$<br>$\hat{b} = 0.1657$ | 4177.59 | 8361.18 | 8361.19 | 8368.10 | 0.4245 |
| Fisk | $\hat{a} = 1.3577$<br>$\hat{b} = 46.452$ | 3476.16 | 6958.32 | 6956.32 | 6965.24 | 0.9718 |
| Inverse Lomax | $\hat{a} = 31.223$<br>$\hat{b} = 1.3335$ | 3508.12 | 7020.24 | 7020.25 | 7029.16 | 0.1209 |

Thus, the variances of the ML estimates are, $var(\hat{\alpha}) = 0.2019$, $var(\hat{\beta}) = 36.8808$, $var(\rho) = 0.0527$ and $var(\hat{\lambda}) = 0.3863$. Therefore, confidence interval for $\alpha$, $\beta$, $\rho$ and $\lambda$ are [1.8240, 2.6156], [60.657, 205.23], [0.2946, 0.5015] and [-0.4007, 1.1136], respectively.

The results in Table 2 indicates that the proposed transmuted Dagum distribution fits well as it has the smallest $\ell(.,y)$ , AIC, AICC and BIC as compared to the Dagum distribution and the others considered distributions. The $KS$ goodness fit test is employed to evaluate the best fitted model for the precipitation data. The calculated value of this test is 0.0280, whereas the tabled critical two-tailed values at 0.05 and 0.01 significance levels are 0.0538 and 0.0644, respectively. According to Sheskin [25], if the value of $KS$ statistic is greater or equal to the critical value then the null hypothesis should be rejected. Thus the null hypothesis cannot be rejected for the transmuted Dagum distribution as the value of the KS-test is not greater or equal to the critical values.

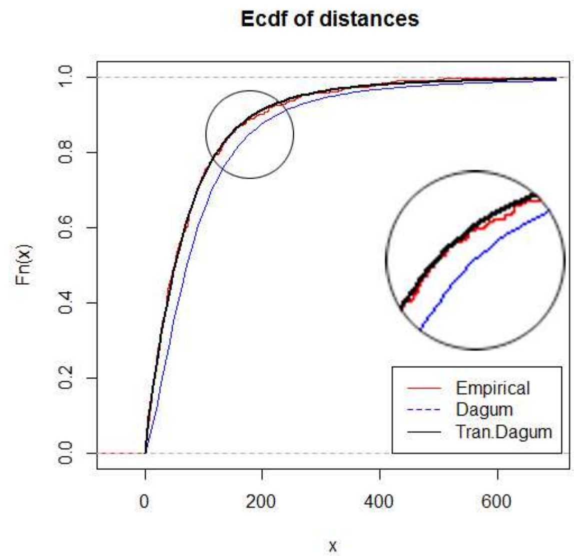



**Figure 5.** Empirical, fitted transmuted Dagum and Dagum cdf of the data set with maximum distance highlight.

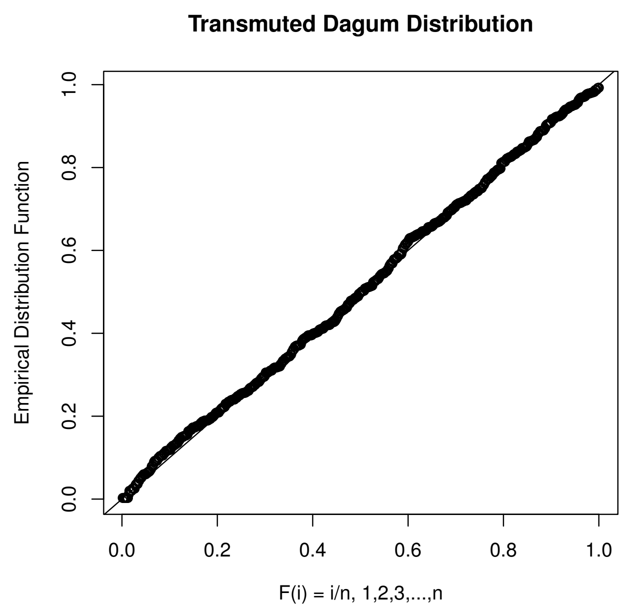



**Figure 6.** PP plots for fitted transmuted Dagum distribution.

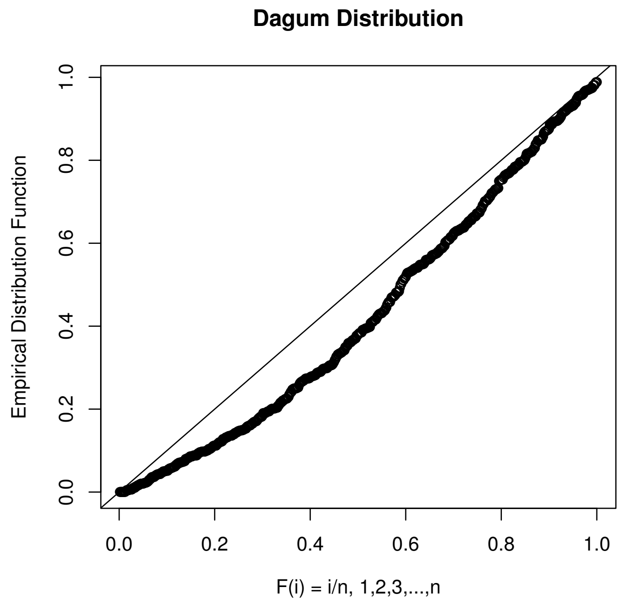



**Figure 7.** PP plots for fitted Dagum distribution

Both empirical cdf and PP-plots also indicate that the transmuted Dagum distribution is better than its competitor Dagum distribution to model the rainfall data. As transmuted Dagum distribution exactly follow the empirical pattern of the data and more closest view showed in circle in Figure 5 and similarly in PP-plot the transmuted Dagum distribution lies almost perfectly on the $45^o$ line. So we conclude that the transmuted Dagum distribution fulfills all the goodness of fit criteria for the data set.

TL-moments evaluate the basic characteristics of data in a better way and show the true picture of the data. First and second moments show the average value and variation in data, respectively. Consistency, symmetry and peakness evaluated by the coefficient of variation $CV$, $Sk$ and $Kr$ using the 2nd, 3rd and 4th moments. These moments and coefficients are calculated and reported in the Table 3 using Islamabad precipitation data set.

**Table 3.** Moments and moment ratios for monthly maximum precipitation data of the Islamabad, Pakistan

| Model | Moments | L-moments | TL-moments | LL-moments | LH-moments |
|---|---|---|---|---|---|
| 1st | 86.2486 | 86.2486 | 114.8730 | 68.2673 | 104.2290 |
| 2nd | 45730.1 | 17.9811 | 24.3523 | 34.9539 | 7.9823 |
| 3rd | $-2.15\times10^6$ | -28.6242 | -15.3700 | -4.0121 | -34.1535 |
| 4th | $-4.24\times10^8$ | -22.6061 | -16.8617 | -16.4118 | -11.8458 |
| $CV$ | 2.4794 | 0.2085 | 0.2120 | 0.5120 | 0.0765 |
| $Sk$ | -1.1691 | -1.5919 | -0.6311 | -0.4115 | -4.2786 |
| $Kr$ | 1.4962 | -1.2572 | -0.6924 | -0.4695 | -1.4840 |

## 10. Conclusion

The transmuted Dagum distribution proposed in this study, is the generalization of the Dagum distribution. This distribution is quite flexible and its application diversities increased due to the fourth transmuted parameter as compared to the standard Dagum distribution. To show the flexibility of new density the plots of the pdf, cdf, reliability function and hazard functions are presented. We derived moments and other basic properties of the proposed distribution. The densities of the lowest, highest, $r$th order statistics, the joint density of the two order statistics and TL-moments are also studied. The parameter estimation is obtained by the maximum likelihood estimation via Newton-Raphson approach. To evaluate its worth five goodness of fit criterion are considered for the selection of most appropriate model among transmuted Dagum, Dagum, transmuted Pareto, Pareto, Fisk and inverse Lomax. On all of these criteria, the results of the application show that transmuted Dagum distribution is superior to the Dagum distribution and other related distribution. Finally, we hope that the proposed model will serve better in income distribution, actuarial, meteorological and survival data analysis.

**Acknowledgement**
We acknowledge the support of the anonymous reviewers, whose constructive comments and suggestions improved the quality of the paper.

## References

[1] Alwan, F.M., Baharum, A. and Hassan, G. S. *Reliability Measurement for Mixed Mode Failures of 33/11 Kilovolt Electric Power Distribution Stations*, PloS one **8** (8), 1-8, 2013.
[2] Bandourian, R., McDonald, J.B. and Turley, R.S. *A Comparison of Parametric Models of Income Distribution across Countries and Over Time*, Estadistica **55**, 135-152, 2003.
[3] Bayazit, M. and Onoz, B. *LL-moments for estimating low flow quantiles*, Hydrological sciences journal **47** (5), 707-720, 2002
[4] Bordley, R. F., McDonald, J.B. and Mantrala, A. *Something new, something old: Parametric models for the size distribution of income*, Journal of Income Distribution **6** (1), 91-103, 1996.
[5] Dagum, C. *A new model of personal income distribution: Specification and estimation*, Economie Applique'e **30**, 413-437, 1977.
[6] Dagum, C. *Income distribution models*, (John Wiley, New York, 1983).
[7] Dagum, C. *The generation and distribution of income, the Lorenz curve and the Gini ratio*, Economie Applique'e **33**, 327-367,1980.
[8] Domma, F. *Asymptotic distribution of the maximum likelihood estimators of the parameters of the right-truncated Dagum distribution*, Communications in Statistics-Simulation and Computation **36** (6), 1187-1199, 2007.
[9] Domma, F. and Condino, F. *The beta-Dagum distribution: definition and properties*, Communications in Statistics-Theory and Methods **42** (22), 4070-4090, 2013.
[10] Domma, F. *L'andamento della Hazard function nel modello di Dagum a tre parametri*, Quaderni di Statistica **4**, 103-114, 2002.
[11] Domma, F., Giordano, S. and Zenga, M. M. *Maximum likelihood estimation in Dagum distribution with censored sample*, Journal of Applied Statistics **38** (12), 2971-2985, 2011.
[12] Elamir, E.A. and Seheult, A.H. *Trimmed L-moments*, Computational Statistics and Data Analysis **43** (3), 299-314, 2003.
[13] Fattorini, L. Lemmi, A. *Proposta di un modello alternativo per lanalisi della distribuzione personale del reddito*, Atti Giornate di Lavoro AIRO **28** (1), 89-117, 1979.
[14] Glaser, R.E. *Bathtub and related failure rate characterizations*, Journal of the American Statistical Association **75** (371), 667-672, 1980.
[15] Hosking, J. R. M. *L-moments: Analysis and estimation of distributions using linear combinations of order statistics*, Journal of the Royal Statistical Society. Series B. Statistical Methodological **52**, 105-124, 1990.

[16] Hyndman. R.J. and Fan, Y. *Sample quantiles in statistical packages*, The American Statistician **50** (4), 361-365, 1996

[17] McGarvey, R.G., Del Castillo, E., Cavalier, T. M. and Lehtihet, E. *Four-parameter beta distribution estimation and skewness test*, Quality and Reliability Engineering International **18** (5), 395-402, 2002.

[18] Oluyede, B.O. and Rajasooriya, S. *The Mc-Dagum Distribution and Its Statistical Properties with Applications*, Asian Journal of Mathematics and Applications **44**, 1-16, 2013.

[19] Oluyede, B.O. and Ye, Y. *Weighted Dagum and related distributions*, Afrika Matematika **25** (4), 1125-1141, 2014.

[20] Perez, C. G. and Alaiz, M. P. *Using the Dagum model to explain changes in personal income distribution*, Applied Economics **43** (28), 4377-4386, 2011.

[21] Quintano, C. and Dagostino, A. *Studying inequality in income distribution of single person households in four developed countries*, Review of Income and Wealth **52** (4), 525-546, 2006.

[22] Shahzad, M. N. and Asghar, Z. *Comparing TL-Moments, L-Moments and Conventional Moments of Dagum Distribution by Simulated data*, Revista Colombiana de Estadistica **36** (1), 79-93, 2013.

[23] Sharma, V. K., Singh, S. K. and Singh, U. *A new upside-down bathtub shaped hazard rate model for survival data analysis*, Applied Mathematics and Computation **239**, 242-253, 2014.

[24] Shaw, W. T. and Buckley, I. R. *[27]. The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map*, arXiv preprint arXiv:0901.0434, 2009.

[25] Sheskin, D. J. *Parametric and Nonparametric Statistical Procedures* (Chapman and Hall/CRC, Bocar Raton, Florida, 2007).

[26] Wang, Q. J. *LH-moments for statistical analysis of extreme events*, Water Resources Research **33** (12), 2841-2848, 1997.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 245 – 265

# The Weibull-power function distribution with applications

M. H. Tahir*†, Morad Alizadeh‡, M. Mansoor§, Gauss M. Cordeiro¶ and M. Zubair‖

**Abstract**

Recently, several attempts have been made to define new models that extend well-known distributions and at the same time provide great flexibility in modelling real data. We propose a new four-parameter model named the *Weibull-power function (WPF)* distribution which exhibits bathtub-shaped hazard rate. Some of its statistical properties are obtained including ordinary and incomplete moments, quantile and generating functions, Rényi and Shannon entropies, reliability and order statistics. The model parameters are estimated by the method of maximum likelihood. A bivariate extension is also proposed. The new distribution can be implemented easily using statistical software packages. We investigate the potential usefulness of the proposed model by means of two real data sets. In fact, the new model provides a better fit to these data than the additive Weibull, modified Weibull, Sarahan-Zaindin modified Weibull and beta-modified Weibull distributions, suggesting that it is a reasonable candidate for modeling survival data.

**Keywords:** Generalized uniform distribution, moments, power function distribution, Weibull-G class.

*2000 AMS Classification:* 60E05, 62E10, 62N05.

*Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan.
Email : mtahir.stat@gmail.com; mht@iub.edu.pk
†Corresponding Author.
‡Department of Statistics, Persian Gulf University of Bushehr, Iran.
Email : moradalizadeh78@gmail.com
§Punjab College, 26-C Shabbir Shaheed Road, Model Town-A, Bahawalpur, Pakistan.
Email : mansoor.abbasi143@gmail.com
¶Department of Statistics, Federal University of Pernambuco, 50740-540, Recife, PE, Brazil.
Email : gauss@de.ufpe.br, gausscordeiro@gmail.com
‖Government Degree College Kahrorpacca, Lodhran, Pakistan.
Email : zubair.stat@yahoo.com

*Received* : 20.10.2014 *Accepted* : 07.12.2014 *Doi* : 10.15672/HJMS.2014428212

## 1. Introduction

A suitable generalized lifetime model is often of interest in the analysis of survival data, as it can provide insight into characteristics of failure times and hazard functions that may not be available with classical models. Four distributions (exponential, Pareto, power and Weibull) are of interest and very attractive in lifetime literature due to their simplicity, easiness and flexible features to model various types of data in different fields. The *power function distribution* (PFD) is a flexible lifetime model which can be obtained from the Pareto model by using a simple transformation $Y = X^{-1}$ [19] and it is also a special case of the beta distribution. Meniconi and Barry [36] discussed the application of the PFD along with other lifetime models, and concluded that the PFD is better than the Weibull, log-normal and exponential models to measure the reliability of electronic components. The PFD can be used to fit the distribution of certain likelihood ratios in statistical tests. If the likelihood ratio (LR) is based on $n$ iid random variables, it is often found that a useful goodness-of-fit can be obtained by letting $(likelihood\ ratio)^{2/n}$ to have a PFD (see [6]). For introduction and statistical properties of the PFD, the reader is referred to Johnson *et al.* [23, 24], Balakrishnan and Nevzorov [13], Kleiber and Kotz [29] and Forbes *et al.* [21]. The estimation of its parameters is discussed in detail by [55, 56, 9]. The estimation of the sample size for parameter estimation is addressed by Kapadia [26]. Ali *et al.* [8] derived the UMVUE of the mean and the right-tail probability of the PFD. Ali and Woo [6] and Ali *et al.* [7] provided inference on reliability and the ratio of variates in the PFD. Sinha *et al.* [51] proposed a preliminary test estimator for a scale parameter of the PFD.

From a Bayesian point of view, the PFD can be used as a prior when there is limited sample information, and especially in cases where the relationship between the variables is known but the data is scarce (possibly due to high cost of collection). The PFD can also be used as prior distribution for the binomial proportion. Saleem *et al.* [45] performed Bayesian analysis of the mixture of PFDs using complete and censored samples. Rehman *et al.* [41] used Bayes estimation and conjugate prior for the PFD. Kifayat *et al.* [28] analyzed this distribution in the Bayesian context using informative and non-informative priors. Zarrin *et al.* [57] discussed the reliability estimation and Bayesian analysis of the system reliability of the PFD.

Several authors have reported characterization of the PFD based on order statistics and records. Rider [44] first derived the distribution of the product and ratio of the order statistics. Govindarajulu [22] gave the characterization of the exponential and PFD. Exact explicit expression for the single and the product moments of order statistics are obtained by Malik [31]. Ahsanullah [2] defined necessary and sufficient conditions based on PFD order statistics. Kabir and Ahsanullah [25] estimated the location and scale of the PFD using linear function of order statistics. Balakrishnan and Joshi [12] derived some recurrence relations for the single and the product moments of order statistics. Moothathu [38, 39] gave characterizations of the PFD through Lorenz curve. The estimation of the PFD parameters based on record values is studied by Ahsanullah [3]. Saran and Singh [47] developed recurrence relations for the marginal and generating functions of generalized order statistics. Saran and Pandey [46] estimated the parameters of the PFD and proposed a characterization based on $k$th record values. The characterization based on the lower generalized order statistics is given in Ahsanullah [4], and Mbah and Ahsanullah [34]. Chang [16] suggested other characterization by independence of records values. Athar and Faizan [10] derived some recurrence relations for single and product moments of lower generalized order statistics. Tavangar [53] gave a characterization based on dual generalized order statistics. Bhatt [14] proposed a characterization based on any arbitrary non-constant function. Recently, Azedine [11] derived single and double moments of the lower record values, and also established recurrence relations for these single and double moments.

Different versions of the PFD are reported in the literature. Some of them are summarized in Table 1, where $\Pi(x)$ denotes its cumulative distribution function (cdf) and $\pi(x)$ denotes its probability probability function (pdf).

Table 1: Some versions of the PFD.

| S.No./Ref. | $\Pi(x)$ | $\pi(x)$ | Range of variable | Parameters |
|---|---|---|---|---|
| 1./ [11] | $x^{\alpha}$ | $\alpha\, x^{\alpha-1}$ | $0 < x < 1$ | $\alpha > 0$ |
| 2./ [5] | $(x/\lambda)^{\alpha}$ | $\alpha\, \lambda^{-\alpha}\, x^{\alpha-1}$ | $0 < x < \lambda$ | $\alpha > 0$ |
| 3./ [18] | $(x\,\beta)^{\alpha}$ | $\alpha\, \beta^{\alpha}\, x^{\alpha-1}$ | $0 < x < \beta^{-1}$ | $\alpha, \beta > 0$ |
| 4./ [10] | $(x/\theta)^{\alpha+1}$ | $(\alpha+1)\, \theta^{-(\alpha+1)}\, x^{\alpha}$ | $0 < x < \theta$ | $\alpha > -1, \theta > 0$ |
| 5./ [47] | $1-(1-x)^{\delta}$ | $\delta\, (1-x)^{\delta-1}$ | $0 < x < 1$ | $\delta > 0$ |
| 6./ [52] | $\left[\frac{x-\theta}{\sigma}\right]^{\nu}$ | $\frac{\nu}{\sigma}\left[\frac{x-\theta}{\sigma}\right]^{\nu-1}$ | $\theta < x < \sigma+\theta$ | $\nu, \sigma > 0$ |
| 7./ [46] | $1-\left[\frac{\beta-x}{\beta-\alpha}\right]^{\gamma}$ | $\frac{\gamma}{\beta-\alpha}\left[\frac{\beta-x}{\beta-\alpha}\right]^{\gamma-1}$ | $\alpha < x < \beta$ | $\gamma > 0$ |
| 8./ [6] | $x^{\left[\frac{\sigma}{1-\sigma}\right]}$ | $\left[\frac{\sigma}{1-\sigma}\right]\, x^{\left[\frac{\sigma}{1-\sigma}\right]-1}$ | $0 < x < 1$ | $0 < \sigma < 1$ |

A random variable $Z$ has the PFD or the generalized uniform distribution (GUD) [40] with two positive parameters $\alpha$ and $\beta$, if its cdf is given by

$$G(x) = \left[\frac{x}{\alpha}\right]^{\beta}, \quad 0 < x < \alpha, \tag{1.1}$$

where $\alpha$ is the scale (threshold) parameter and $\beta$ is the shape parameter. The pdf corresponding to (1.1) reduces to

$$g(x) = \left[\frac{\beta}{\alpha}\right]\left[\frac{x}{\alpha}\right]^{\beta-1}, \quad 0 < x < \alpha, \tag{1.2}$$

The distribution (1.1) has the following special cases:

(i) if $\alpha = 1$, the PFD reduces to standard power distribution,

(ii) if $\alpha = 1$ and $\beta = 1$, it reduces to standard uniform distribution,

(iii) if $\beta = 1$, it gives the rectangular distribution [31, 25],

(iv) if $\beta = 2$, it refers to triangular distribution [31, 25],

(v) if $\beta = 3$, it refers to J-shaped distribution [31, 25],

(vi) if $\alpha = 1$ and $Y = X^{-1}$, then $Y \sim$Pareto$(0, \beta)$ [21],

(vii) if $\alpha = 1$ and $Y = -\log X$, then $Y \sim$Exponential$(\beta^{-1})$ [21],

(viii) if $\alpha = 1$ and $Y = -\log(X^{\beta} - 1)$, then $Y \sim$Logistic$(0, 1)$ [21],

(ix) if $\alpha = 1$ and $Y = [-\log(X^{\beta})]^{1/\gamma}$, then $Y \sim$Weibull$(0, \gamma)$ [21],

(x) if $\alpha = 1$ and $Y = -\log[-b \log X]$, then $Y \sim$Gumbel$(0, 1)$ [21],

(xi) if $\alpha = 1$ and $Y = -b\,[X_1/X_2]$, then $Y \sim$Laplace$(0, 1)$ [21].

Henceforth, let $Z$ be a random variable having the PFD with parameters $\alpha$ and $\beta$, say $Z \sim$PFD$(\alpha, \beta)$. Then, the quantile function (qf) is $G^{-1}(u) = \alpha\, u^{1/\beta}$ (for $0 < u < 1$).

The survival function (sf) $\overline{G}(x)$, hazard rate function (hrf) $\tau(x)$, reversed hazard rate function (rhrf) $r(x)$, cumulative hazard rate function (chrf) $V(x)$ and odd ratio (OR) $G(x)/\overline{G}(x)$ of $Z$ are given by $\overline{G}(x) = 1-(x/\alpha)^{\beta} = \frac{\alpha^{\beta}-x^{\beta}}{\alpha^{\beta}}$, $\tau(x) = \frac{\beta\, x^{\beta-1}}{\alpha^{\beta}-x^{\beta}}$, $r(x) = (\beta/x)$, $V(x) = -\log\left[1-(x/\alpha)^{\beta}\right]$ and OR $= \frac{x^{\beta}}{\alpha^{\beta}-x^{\beta}}$, respectively.

The $n$th moment of $Z$ comes from (1.2) as

$$E(Z^n) = \frac{\alpha^n\,\beta}{\beta+n}. \tag{1.3}$$

The mean and variance of $Z$ are

$$E(Z) = [\alpha\beta/(\beta+1)]$$

and

$$Var(Z) = \left\{\beta\alpha^2/[(\beta+2)(\beta+1)^2]\right\},$$

respectively.

The moment generating function (mgf) of $Z$ becomes

$$M_Z(t) = \frac{\beta\Big[\Gamma(\beta)-\Gamma(\beta,-t\,\alpha)\Big]}{(-t)^{\beta}\,\alpha^{\beta}}, \quad t<0, \tag{1.4}$$

where $\Gamma(a;bx) = b^a \int_x^{\infty} w^{a-1}\,\mathrm{e}^{-\mathrm{b\,w}}\mathrm{dw}$ for $a>0$ and $b>0$ and $\Gamma(\cdot;\cdot)$ is the complementary gamma function.

The $n$th incomplete moment of $Z$ can be expressed as

$$m_{(n,Z)}(x) = \frac{\beta}{\alpha^{\beta}}\,\frac{x^{\beta+n}}{\beta+n}. \tag{1.5}$$

In this paper, we propose an extension of the PFD called the *Weibull power function* (for short "WPF") distribution based on the *Weibull-G* class of distributions defined by Bourguignon *et al.* [15]. Zagrafos and Balakrishnan [58] pioneered a versatile and flexible gamma-G class of distributions based on Stacy's generalized gamma model and record value theory. More recently, Bourguignon *et al.* [15] proposed the *Weibull-G* class of distributions influenced by the gamma-G class. Let $G(x;\Theta)$ and $g(x;\Theta)$ denote the cumulative and density functions of a baseline model with parameter vector $\Theta$ and consider the Weibull cdf $\pi_W(x) = 1-\mathrm{e}^{-a\,x^b}$ (for $x>0$) with scale parameter $a>0$ and shape parameter $b>0$. Bourguignon *et al.* [15] replaced the argument $x$ by $G(x;\Theta)/\overline{G}(x;\Theta)$, where $\overline{G}(x;\Theta) = 1-G(x;\Theta)$, and defined the cdf of their class, say Weibull-G$(a,b,\Theta)$, by

$$F(x) = F(x;a,b,\Theta) = a\,b\int_0^{\left[\frac{G(x;\Theta)}{\overline{G}(x;\Theta)}\right]} x^{b-1}\,\mathrm{e}^{-ax^b}\,dx = 1-\mathrm{e}^{-a\left[\frac{G(x;\Theta)}{\overline{G}(x;\Theta)}\right]^b},\ x\in\Re. \tag{1.6}$$

The Weibull-G class density function becomes

$$f(x) = f(x;a,b,\Theta) = a\,b\,g(x;\Theta)\,\left[\frac{G(x;\Theta)^{b-1}}{\overline{G}(x;\Theta)^{b+1}}\right]\,\mathrm{e}^{-a\left[\frac{G(x;\Theta)}{\overline{G}(x;\Theta)}\right]^b}. \tag{1.7}$$

If $b=1$, it corresponds to the *exponential-G* class. An interpretation of equation (1.6) can be given as follows. Let $Y$ be the lifetime variable having a parent G distribution. Then, the odds that an individual will die at time $x$ is $G(x;\Theta)/\overline{G}(x;\Theta)$. We are interested in modeling the randomness of the odds of death using an appropriate parametric distribution, say $F(x)$. So, we can write

$$F(x) = \Pr(X\le x) = F\Big[\frac{G(x;\Theta)}{\overline{G}(x;\Theta)}\Big].$$

The paper unfolds as follows. In Section 2, we define a new bathtub shaped model called the *Weibull-power function* (WPF) distribution and discuss the shapes of its density and hrf. In Section 3, some of its statistical properties are investigated. In Section 4, Rényi and Shannon entropies are derived and the reliability is determined in Section 5. The density of the order statistics is obtained in Section 6. The model parameters are estimated by maximum likelihood and a simulation study is performed in Section 7. In Section 8, a bivariate extension of the new family is introduced. Applications to two real data sets illustrate the performance of the new model in Section 9. The paper is concluded in Section 10.

## 2. Model definition

Inserting (1.1) in equation (1.6) gives the WPF cdf as

$$F(x) = F(x; a, b, \alpha, \beta) = 1 - \mathrm{e}^{-a\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b}, \quad 0 < x < \alpha, \quad a, b, \alpha, \beta > 0. \tag{2.1}$$

The pdf corresponding to (2.1) is given by

$$f(x) = f(x; a, b, \alpha, \beta) = \frac{a\, b\, \beta\, \alpha^\beta\, x^{\beta\, b-1}}{(\alpha^\beta - x^\beta)^{b+1}}\ \mathrm{e}^{-a\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b}. \tag{2.2}$$

Henceforth, let $X \sim \mathrm{WPF}(a, b, \alpha, \beta)$ be a random variable having pdf (2.2). The sf, hrf, rhrf and chrf of $X$ are given by

$$S(x) = S(x; a, b, \alpha, \beta) = \mathrm{e}^{-a\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b}, \tag{2.3}$$

$$\tau(x) = h(x; a, b, \alpha, \beta) = \frac{a\, b\, \beta\, \alpha^\beta x^{\beta\, b-1}}{(\alpha^\beta - x^\beta)^{b+1}},$$

$$r(x) = r(x; a, b, \alpha, \beta) = \frac{a\, b\, \beta\, \alpha^\beta x^{\beta\, b-1}}{(\alpha^\beta - x^\beta)^{b+1}}\ \frac{\mathrm{e}^{-a\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b}}{\left[1 - \mathrm{e}^{-a\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b}\right]}$$

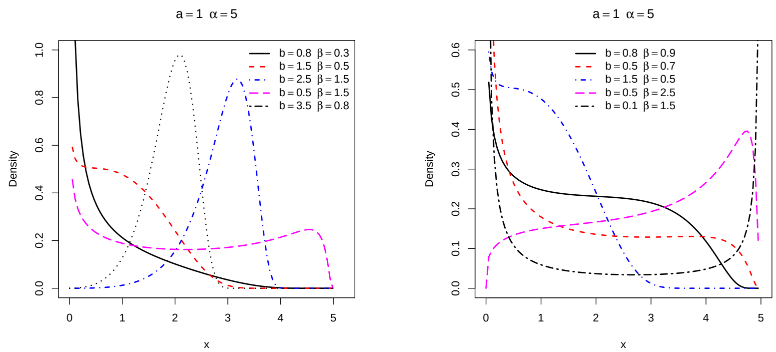
and

$$V(x) = V(x; a, b, \alpha, \beta) = a\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b,$$

respectively.
Figures 1 and 2 display some plots of the pdf and hrf of $X$ for some parameter values. Figure 1 indicates that the WPF pdf has various shapes such as symmetric, right-skewed, left-skewed, reversed-J, S, M and bathtub. Also, Figure 2 indicates that the WPF hrf can have bathtub-shaped, J and U shapes.

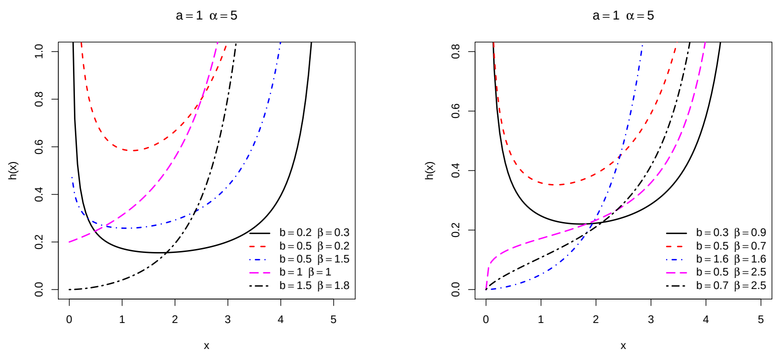
Lemma 2.1 provides some relations of the WPF distribution with the Weibull and exponential distributions.

**2.1. Lemma.** *(Transformation): (a) If a random variable $Y$ follows the Weibull distribution with shape parameter $b$ and scale parameter $a$, then the random variable $X = \alpha\left[\frac{Y}{1+Y}\right]^{1/\beta}$ has the WPF$(a, b, \alpha, \beta)$ distribution.*
*(b) If a random variable $Y$ follows the exponential distribution, then the random variable $X = \alpha\left[\frac{Y^{1/b}}{1+Y^{1/b}}\right]^{1/\beta}$ has the WPF$(a, b, \alpha, \beta)$ distribution.*




(a) (b)

**Figure 1.** Plots of the WPF pdf for some parameters.




(a) (b)

**Figure 2.** Plots of the hazard rate for some parameters.

**2.1. Shape and asymptotics.** The critical points of the density of $X$ are the roots of the equation

$$\frac{b\beta - 1}{x} + \frac{\beta(b+1)x^{\beta-1}}{\alpha^\beta - x^\beta} - \frac{a\, b\, \beta\, \alpha^\beta x^{\beta\, b-1}}{(\alpha^\beta - x^\beta)^{b+1}} = 0. \tag{2.4}$$

The first derivative of the hrf of $X$ is given by

$$\tau'(x) = \frac{x^{b\beta-2}\left\{(\beta+1)x^\beta + (b\,\beta - 1)\alpha^\beta\right\}}{(\alpha^\beta - x^\beta)^{b+2}}. \tag{2.5}$$

The limiting behavior of the pdf and hrf of $X$ are given in the following lemma.

**2.2. Lemma.** *The limits of the pdf and hrf of $X$ when $x \to \alpha^-$ are 0 and $+\infty$. Further, the limits of the pdf and hrf of $X$ when $x \to 0$ are given by*

$$\lim_{x\to 0^+} f(x) = \begin{cases} +\infty & \text{for } b\,\beta < 1; \\ \frac{a}{\alpha} & \text{for } b\,\beta = 1; \\ 0 & \text{for } b\,\beta > 1. \end{cases}$$

$$\lim_{x\to 0^+} \tau(x) = \begin{cases} +\infty & \textit{for } b\,\beta < 1; \\ \frac{a}{\alpha} & \textit{for } b\,\beta = 1; \\ 0 & \textit{for } b\,\beta > 1. \end{cases}$$

The mode of the hrf of $X$ is at $x = 0$ when $\beta\, b \geq 1$ and it occurs at $x = \alpha \left[\frac{1-b\beta}{1+\beta}\right]^{\frac{1}{\beta}}$ when $b\beta < 1$.

**2.3. Theorem.** *The hrf of $X$ is increasing when $b\,\beta \geq 1$ and is bathtub when $b\,\beta < 1$.*

## 3. Mathematical properties

Established algebraic expansions to determine some mathematical properties of the WPF distribution can be more efficient than computing those directly by numerical integration of (2.2), which can be prone to rounding off errors among others. Despite the fact that the cdf and pdf of the WPF distribution require mathematical functions that are widely available in modern statistical packages, frequently analytical and numerical derivations take advantage of certain expansions for its pdf.

**3.1. Quantile function.** The quantile function (qf) of $X$ follows by inverting (2.1) as

$$Q(u) = \alpha \left[ \frac{\left[\frac{-1}{a}\log(1-u)\right]^{\frac{1}{b}}}{1 + \left[\frac{-1}{a}\log(1-u)\right]^{\frac{1}{b}}} \right]^{\frac{1}{\beta}}. \tag{3.1}$$

So, the simulation of the WPF random variable is straightforward. If $U$ is a uniform variate on the unit interval $(0, 1)$, then the random variable $X = Q(U)$ has pdf (2.2).

The analysis of the variability of the the skewness and kurtosis on the shape parameters $\alpha$ and $b$ can be investigated based on quantile measures. The shortcomings of the classical kurtosis measure are well-known. The Bowley skewness [27] based on quartiles is given by

$$B = \frac{Q(3/4) + Q(1/4) - 2\,Q(2/4)}{Q(3/4) - Q(1/4)}.$$

The Moors kurtosis [37] based on octiles is given by

$$M = \frac{Q(3/8) - Q(1/8) + Q(7/8) - Q(5/8)}{Q(6/8) - Q(2/8)}.$$

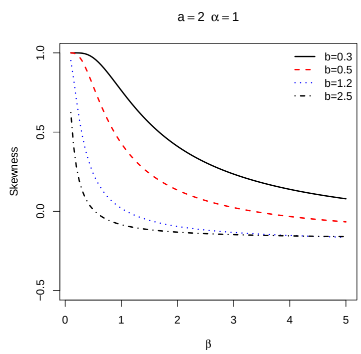
These measures are less sensitive to outliers and they exist even for distributions without moments. In Figure 3, we plot the measures $B$ and $M$ for the WPF distribution. The plots indicate the variability of these measures on the shape parameters $\beta$.

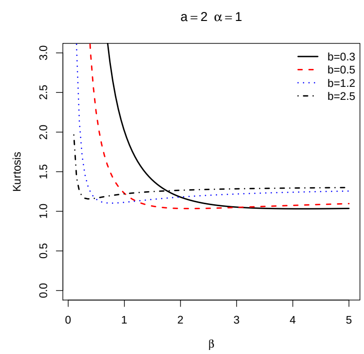
**3.2. Useful expansion.** We use the exponential power series and the expansion

$$\Big[1 - G(x;\Theta)\Big]^{-b} = \sum_{k=0}^{\infty} p_k\, G(x;\Theta)^k,$$

where $p_k = \Gamma(b+k)/[k!\,\Gamma(b)]$. After some algebra, we can easily obtain

$$F(x) = F(x; a, b, \alpha, \beta) = \sum_{\substack{j,k\geq 0 \\ j+k\geq 1}} w_{j,k}\, H(x; \alpha, \beta_{j,k}), \tag{3.2}$$




(a) (b)

**Figure 3.** Skewness (a) and kurtosis (b) plots for WPF distribution based on quantiles.

where $w_{j,k} = (-a)^j \, p_k / j!$, $\beta_{j,k} = (jb + k)\beta$ and $H(x; \alpha, \beta_{j,k})$ is the cdf of the PFD with scale parameter $\alpha$ and shape parameter $\beta_{j,k}$. Let $Z_{j,k}$ be the random variable with cdf $H(x; \alpha, \beta_{j,k})$. By simple differentiation, we can express the pdf of $X$ as

$$f(x) = f(x; a, b, \alpha, \beta) = \sum_{\substack{j,k \geq 0 \\ j+k \geq 1}} w_{j,k} \, h(x; \alpha, \beta_{j,k}), \tag{3.3}$$

where $h(x; \alpha, \beta_{j,k})$ is the pdf of $Z_{j,k}$. Equation (3.3) reveals that the WPF distribution is a mixture of PFDs with the same scale parameter $\alpha$ and different shape parameters. Thus, some WPF mathematical properties can be obtained from those corresponding properties of the PFD.

**3.3. Ordinary and incomplete moments.** The $n$th moment of $X$, say $\mu'_n$ can be expressed from (1.3) and (3.3) as

$$\mu'_n = \alpha^n \sum_{\substack{j,k \geq 0 \\ j+k \geq 1}} \frac{\beta_{j,k} \, w_{j,k}}{\beta_{j,k} + n}. \tag{3.4}$$

Setting $n = 1$ in (3.4), we obtain the mean $\mu'_1 = E(X)$. The central moments ($\mu_n$) and cumulants ($\kappa_n$) of $X$ are obtained from equation (3.4) as

$$\mu_n = \sum_{k=0}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix} (-1)^k \, \mu_1'^k \, \mu'_{n-k} \qquad \text{and} \qquad \kappa_n = \mu'_n - \sum_{k=1}^{n-1} \begin{Bmatrix} n-1 \\ k-1 \end{Bmatrix} \kappa_k \, \mu'_{n-k},$$

respectively, where $\kappa_1 = \mu'_1$ and the notation

$$\begin{Bmatrix} n \\ k \end{Bmatrix}$$

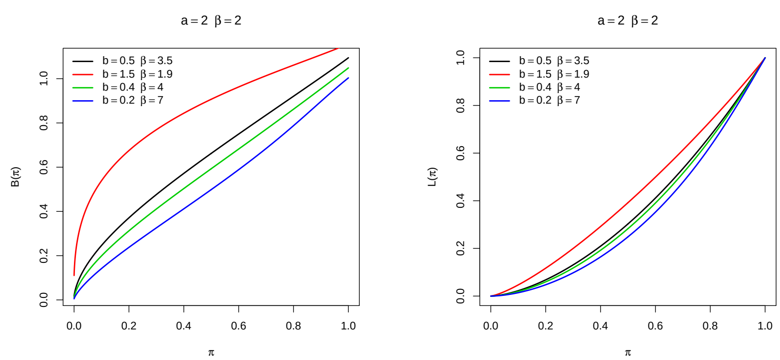
is used to denote the binomial coefficient.

Thus, $\kappa_2 = \mu'_2 - \mu_1'^2$, $\kappa_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2\mu_1'^3$, $\kappa_4 = \mu'_4 - 4\mu'_3\mu'_1 - 3\mu_2'^2 + 12\mu'_2\mu_1'^2 - 6\mu_1'^4$, etc. The skewness and kurtosis can be calculated from the third and fourth standardized cumulants as $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ and $\gamma_2 = \kappa_4/\kappa_2^2$. They are also important to derive Edgeworth expansions for the cdf and pdf of the standardized sum and sample mean of iid random variables having the WPF distribution.

The $n$th incomplete moment of $X$ can be determined from (1.5) and (3.3)

$$m_{(n,X)}(x) = \sum_{\substack{j,k\geq 0\\ j+k\geq 1}} \frac{\beta_{j,k}}{\alpha^{\beta_{j,k}}}\, \frac{x^{\beta_{j,k}+n}}{\beta_{j,k}+n}. \tag{3.5}$$

The main application of the first incomplete moment refers to the Bonferroni and Lorenz curves. These curves are very useful in several fields. For a given probability $\pi$, they are defined by $B(\pi) = m_1(q)/(\pi\, \mu_1')$ and $L(\pi) = m_{(1,X)}(q)/\mu_1'$, respectively, where $m_{(1,X)}(q)$ comes from (3.5) with $r = 1$ and $q = Q(\pi)$ is determined from (3.1).




(a) (b)

**Figure 4.** Plots of the Bonferroni curve (a) and Lorenz curve (b) for the WPF model.

The amount of scatter in a population is measured to some extent by the totality of deviations from the mean and median defined by $\delta_1 = \int_0^\infty |x - \mu_1'|\, f(x)dx$ and $\delta_2(x) = \int_0^\infty |x - M|\, f(x)dx$, respectively, where $\mu_1' = E(X)$ is the mean and $M = Q(0.5)$ is the median. These measures can be expressed as $\delta_1 = 2\mu_1'\, F(\mu_1') - 2m_{(1,X)}(\mu_1')$ and $\delta_2 = \mu_1' - 2m_{(1,X)}(M)$, where $F(\mu_1')$ is given by (2.1) and $m_{(1,X)}(x)$ comes from (3.5) with $n = 1$.

Further applications of the first incomplete moment are related to the mean residual life and mean waiting time given by $s(x; a, b, \alpha, \beta) = [1 - m_{(1,X)}(x)]/S(x) - t$ and $\mu(x; a, b, \alpha, \beta) = t - [m_{(1,X)}(x)/F(x)]$, respectively, where $S(x) = 1 - F(x)$ is obtained from (2.1).

**3.4. Moment generating function.** We obtain the moment generating function (mgf) $M_X(t)$ of $X$ from (3.3) as

$$M(t) \;=\; \sum_{\substack{j,k\geq 0\\ j+k\geq 1}} w_{j,k} \int_0^\alpha \mathrm{e}^{tx}\, h(x; \alpha, \beta_{j,k})\, dx.$$

Based on (1.4), $M(t)$ can be expressed as

$$M(t) \;=\; \sum_{\substack{j,k\geq 0\\ j+k\geq 1}} \frac{w_{j,k}\, \beta_{j,k}}{(-t)^{\beta_{j,k}}\, \alpha^{\beta_{j,k}}} \Big[\Gamma(\beta_{j,k}) - \Gamma(\beta_{j,k}; -t\alpha)\Big],$$

which is the main result of this section.

## 4. Entropies

An entropy is a measure of variation or uncertainty of a random variable $X$. Two popular entropy measures are the Rényi [43] and Shannon [49].

The Rényi entropy of a random variable $X$ with pdf $f(x)$ is defined as

$$I_R(\gamma) = \frac{1}{1-\gamma} \log \left[ \int_0^\infty f^\gamma(x)\, dx \right],$$

for $\gamma > 0$ and $\gamma \neq 1$.

The Shannon entropy of $X$ is defined by $E\{-\log[f(X)]\}$. It is the special case of the Rényi entropy when $\gamma \uparrow 1$. Direct calculation yields

$$\begin{aligned} \mathrm{E}\{-\log[f(X)]\} &= -\log(a\,b\,\beta\,\alpha^\beta) + (1-\beta)\,\mathrm{E}\{\log(X)\} \\ &+ (b+1)\,\mathrm{E}\left[\log(\alpha^\beta - X^\beta)\right] + a\,\mathrm{E}\left[\frac{X^\beta}{\alpha^\beta - X^\beta}\right]^b. \end{aligned}$$

First, we define and compute

$$A(a_1, a_2, a_3; \alpha, \beta, b) = \int_0^\alpha \frac{x^{a_1}}{(\alpha^\beta - x^\beta)^{a_2}}\, \mathrm{e}^{-a_3 \left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b} dx. \tag{4.1}$$

Using the power series and the generalized binomial expansion, and after some algebraic manipulations, we obtain

$$A(a_1, a_2, a_3; \alpha, \beta, b) = \sum_{i,j=0}^\infty \frac{(-1)^{i+j}\, a_3^i\, \alpha^{a_1 - \beta\, a_2}}{[a_1 + \beta\, b\, i + \beta\, j + 1]\; i!} \left\{ \begin{matrix} -a_2 - b\, i \\ j \end{matrix} \right\}.$$

**4.1. Proposition.** *Let $X$ be a random variable with pdf (2.2), then*

$$\mathrm{E}\{\log(X)\} = a\,b\,\beta\,\alpha^\beta \frac{\partial}{\partial t} A(b\beta + t - 1, b+1, a; \alpha, \beta, b)\,|_{t=0},$$

$$\mathrm{E}\left[\log(\alpha^\beta - X^\beta)\right] = a\,b\,\beta\,\alpha^\beta \frac{\partial}{\partial t} A(b\beta - 1, b+1-t, a; \alpha, \beta, b)\,|_{t=0},$$

$$\mathrm{E}\left[\left\{\frac{X^\beta}{\alpha^\beta - X^\beta}\right\}^b\right] = a\,b\,\beta\,\alpha^\beta A(2b\beta - 1, 2b+1, a; \alpha, \beta, b).$$

The simplest formula for the entropy of $X$ is given by

$$\begin{aligned} \mathrm{E}\{-\log[f(X)]\} &= -\log(a\,b\,\beta\,\alpha^\beta) \\ &+ (1-\beta) a\,b\,\beta\,\alpha^\beta \frac{\partial}{\partial t} A(b\beta + t - 1, b+1, a; \alpha, \beta, b)\,|_{t=0} \\ &+ (b+1) a\,b\,\beta\,\alpha^\beta \frac{\partial}{\partial t} A(b\beta - 1, b+1-t, a; \alpha, \beta, b)\,|_{t=0} \\ &+ a^2\,b\,\beta\,\alpha^\beta\, A(2b\beta - 1, 2b+1, a; \alpha, \beta, b). \end{aligned}$$

After some algebraic developments, the Rényi entropy $I_R(\gamma)$ reduces to

$$I_R(\gamma) = \frac{\gamma}{1-\gamma} \log\left[a\,b\,\beta\,\alpha^\beta\right] + \frac{1}{1-\gamma} \log\left\{A\Big[\gamma(\beta\, b - 1), \gamma(b+1), a\gamma; \alpha, \beta, b\Big]\right\}. \tag{4.2}$$

## 5. Reliability

Let $X_1$ and $X_2$ be two continuous and independent WPF random variables with cdfs $F_1(x)$ and $F_2(x)$ and pdfs $f_1(x)$ and $f_2(x)$, respectively. The reliability parameter $R = P(X_1 < X_2)$ is defined by

$$R = P(X_1 < X_2) = \int_0^{\alpha_2} P(X_1 \le X_2 | X_2 = x)\, f_{X_2}(x) dx, \tag{5.1}$$

where $X_1 \sim$WPF$(a_1, b_1, \alpha_1, \beta_1)$ and $X_2 \sim$WPF$(a_2, b_2, \alpha_2, \theta_2)$.

After some algebra, we obtain

$$\begin{aligned}R &= \sum_{\substack{j,k,r,s\geq 0\\ j+k\geq 1, r+s\geq 1}} w^{(1)}_{j,k}\, w^{(2)}_{r,s} \int_0^{\alpha_2} H(x;\alpha_1,\beta^{(1)}_{j,k})\, h(x;\alpha_2,\beta^{(2)}_{r,s}) dx \\ &= \sum_{\substack{j,k,r,s\geq 0\\ j+k\geq 1, r+s\geq 1}} w^{(1)}_{j,k}\, w^{(2)}_{r,s}\, \frac{(b_2 r+s)}{\alpha_2\beta_2} \left[\frac{\alpha_2}{\alpha_1}\right]^{\frac{b_1 j+k}{\beta_1}} \left[\frac{b_1 j+k}{\beta_1}+\frac{b_2 r+s}{\beta_2}\right]^{-1},\end{aligned}$$

where $w^{(1)}_{j,k} = w_{j,k}|_{a=a_1,b=b_1,\beta=\beta_1}$ and $w^{(2)}_{r,s} = w_{r,s}|_{a=a_2,b=b_2,\beta=\beta_2}$.

## 6. Order statistics

Here, we give the density of the $i$th order statistic $X_{i:n}$, $f_{i:n}(x)$ say, in a random sample of size $n$ from the WPF distribution. It is well known that (for $i=1,\ldots,n$)

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i!)}\, f(x)\, F^{i-1}(x)\, \{1-F(x)\}^{n-i}\,. \tag{6.1}$$

Using the binomial expansion, we can rewrite $f_{i:n}(x)$ as

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i!)}\, f(x) \sum_{j=0}^{n-i} (-1)^j \begin{Bmatrix} n-i \\ j \end{Bmatrix} F(x)^{i+j-1}. \tag{6.2}$$

Using (2.2) in (6.2) to compute $F(x)^{i+j-1}$, we obtain

$$\begin{aligned} f_{i:n}(x) &= \frac{n!}{(i-1)!(n-i)!} \sum_{j=0}^{n-i}\sum_{k=0}^{j+i-1} \underbrace{(-1)^{j+k} \begin{Bmatrix} n-i \\ j \end{Bmatrix} \begin{Bmatrix} j+i-1 \\ k \end{Bmatrix}}_{t_{j,k}} \\ &\times \quad \frac{a\, b\, \beta\, \alpha^\beta\, x^{\beta\, b-1}}{(\alpha^\beta - x^\beta)^{b+1}}\ \mathrm{e}^{-a(1+k)\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b}. \end{aligned}$$

The $r$th moment of $X_{i:n}$ can be obtained as

$$\mathrm{E}\,(X^r_{i:n}) = \sum_{j=0}^{n-i}\sum_{k=0}^{j+i-1} t_{j,k}\ A(\beta\, b-1, b+1, a+k; \alpha, \beta), \tag{6.3}$$

where

$$t_{j,k} = \frac{(-1)^{j+k}\, n!}{(i-1)!(n-i)!} \begin{Bmatrix} n-i \\ j \end{Bmatrix} \begin{Bmatrix} j+i-1 \\ k \end{Bmatrix}.$$

After some algebra, the Rényi entropy of $X_{i:n}$ becomes

$$\begin{aligned} I_{R,X_{i:n}}(\gamma) &= \frac{\gamma}{1-\gamma} \log\left[\frac{n!\, a\, b\, \beta\, \alpha^\beta}{(i-1)!(n-i)!}\right] \\ &+ \frac{1}{1-\gamma} \log\Big[\sum_{j,k=0}^{\infty}\sum_{r=0}^{k} t^*_{j,k,r}\ A(\gamma(\beta\, b-1), \gamma(b+1), a(\gamma+r); \alpha, \beta, b)\Big], \end{aligned}$$

where

$$t^*_{j,k,r} = (-1)^{j+k} \begin{Bmatrix} \gamma(n-1) \\ j \end{Bmatrix} \begin{Bmatrix} \gamma(i-1)+j \\ k \end{Bmatrix}.$$

## 7. Estimation

Here, we determine the maximum likelihood estimates (MLEs) of the model parameters of the new family from complete samples only. Let $x_1, \ldots, x_n$ be observed values from the WPF distribution with parameters in $\Theta = (a, b, \beta)$. Then, the total log-likelihood function for $\Theta$ is given by

$$\begin{aligned} \ell_n &= \ell_n(\Theta) = n \log\left[a\, b\, \beta\, \alpha^{\beta}\right] + (\beta\, b - 1)\sum_{i=1}^{n}\log(x_i) \\ &- (b+1)\sum_{i=1}^{n}\log\left(\alpha^{\beta} - x_i^{\beta}\right) - a\sum_{i=1}^{n}\left[\frac{x_i^{\beta}}{\alpha^{\beta} - x_i^{\beta}}\right]^{b}. \end{aligned} \tag{7.1}$$

The log-likelihood function can be maximized either directly by using the SAS (PROC NLMIXED) or the Ox (sub-routine MaxBFGS) program (see [20]), R-language [42] or by solving the nonlinear likelihood equations obtained by differentiating (7.1).

The $\alpha$ is known and we estimate it from the sample maxima. The components of the score function $U_n(\Theta) = (\partial\ell_n/\partial a, \partial\ell_n/\partial b, \partial\ell_n/\partial\beta)^{\top}$ are given by

$$\frac{\partial\ell_n}{\partial a} = \frac{n}{a} - \sum_{i=1}^{n}\left[\frac{x_i^{\beta}}{\alpha^{\beta} - x_i^{\beta}}\right]^{b}$$

$$\frac{\partial\ell_n}{\partial b} = \frac{n}{b} + \beta\sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}\log\left[\alpha^{\beta} - x_i^{\beta}\right] - a\sum_{i=1}^{n}\left[\frac{x_i^{\beta}}{\alpha^{\beta} - x_i^{\beta}}\right]^{b}\log\left[\frac{x_i^{\beta}}{\alpha^{\beta} - x_i^{\beta}}\right]$$

and

$$\begin{aligned} \frac{\partial\ell_n}{\partial\beta} &= \frac{n}{\beta} + n\log(\alpha) + b\sum_{i=1}^{n}\log(x_i) - (b+1)\sum_{i=1}^{n}\left[\frac{\alpha^{\beta}\log(\alpha) - x_i^{\beta}\log(x_i)}{\alpha^{\beta} - x_i^{\beta}}\right] \\ &- a\, b\alpha^{\beta}\sum_{i=1}^{n}\left[\frac{x_i^{b\beta}\log(\frac{x_i}{\alpha})}{(\alpha^{\beta} - x_i^{\beta})^{b+1}}\right]. \end{aligned}$$

Setting these equations to zero and solving them simultaneously yields the MLEs of the three parameters. For interval estimation of the model parameters, we require the $3 \times 3$ observed information matrix $J(\Theta) = \{U_{rs}\}$ (for $r, s = a, b, \beta$), whose elements are listed in Appendix A. Under standard regularity conditions, the multivariate normal $N_3(0, J(\widehat{\Theta})^{-1})$ distribution can be used to construct approximate confidence intervals for the model parameters. Here, $J(\widehat{\Theta})$ is the total observed information matrix evaluated at $\widehat{\Theta}$. Then, the $100(1-\gamma)\%$ confidence intervals for $a$, $b$ and $\beta$ are given by $\hat{a} \pm z_{\alpha^*/2} \times \sqrt{var(\hat{a})}$, $\hat{b} \pm z_{\alpha^*/2} \times \sqrt{var(\hat{b})}$ and $\hat{\beta} \pm z_{\alpha^*/2} \times \sqrt{var(\hat{\beta})}$, respectively, where the $var(\cdot)$'s denote the diagonal elements of $J(\widehat{\Theta})^{-1}$ corresponding to the model parameters, and $z_{\alpha^*/2}$ is the quantile $(1-\alpha^*/2)$ of the standard normal distribution.

**7.1. Simulation study.** To evaluate the performance of the MLEs of the WPF parameters, a simulation study is conducted for a total of twelve parameter combinations and the process in each case is repeated 200 times. Two different sample sizes $n = 100$ and 300 are considered. The MLEs of the parameters and their standard errors are listed in Table 2. In this simulation study, we take $\alpha = 1$. The figures in Table 2 indicate that the MLEs perform well for estimating the model parameters. Further, as the sample size increases, the biases and standard errors of the estimates decrease.

Table 2: MLEs and standard standard errors for some parameter values

| Sample size | Actual values | | | Estimated values | | | Standard errors | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | $a$ | $b$ | $\beta$ | $\tilde{a}$ | $\tilde{b}$ | $\tilde{\beta}$ | $\tilde{a}$ | $\tilde{b}$ | $\tilde{\beta}$ |
| 100 | 0.5 | 0.5 | 1.0 | 0.5124 | 0.5064 | 1.5287 | 0.0138 | 0.0057 | 0.0405 |
| | 0.5 | 1.0 | 1.0 | 0.5928 | 1.0106 | 1.1023 | 0.0361 | 0.0137 | 0.0474 |
| | 0.5 | 1.5 | 2.0 | 0.6161 | 1.4959 | 2.1283 | 0.0399 | 0.0147 | 0.0669 |
| | 1.0 | 1.5 | 2.0 | 1.5224 | 1.4887 | 2.3224 | 0.1386 | 0.0287 | 0.1187 |
| | 1.5 | 1.5 | 2.0 | 1.8829 | 1.5294 | 2.1325 | 0.1543 | 0.0310 | 0.0973 |
| | 2.0 | 1.0 | 1.0 | 2.1982 | 1.0293 | 1.0834 | 0.1271 | 0.0240 | 0.0511 |
| | 2.0 | 0.5 | 1.0 | 1.9921 | 0.5208 | 1.0109 | 0.0525 | 0.0103 | 0.0328 |
| | 2.0 | 0.5 | 2.0 | 1.9807 | 0.5220 | 2.0032 | 0.0509 | 0.0102 | 0.0627 |
| | 2.0 | 0.5 | 1.5 | 1.9977 | 0.5248 | 1.5256 | 0.0539 | 0.0107 | 0.0561 |
| | 2.0 | 0.5 | 0.5 | 1.9794 | 0.5145 | 0.5048 | 0.0475 | 0.0097 | 0.0152 |
| | 2.0 | 1.5 | 0.5 | 2.7821 | 1.5288 | 0.5672 | 0.2529 | 0.0380 | 0.0320 |
| | 2.0 | 2.0 | 0.5 | 2.8568 | 2.0116 | 0.5274 | 0.2984 | 0.0350 | 0.0226 |
| 300 | 0.5 | 0.5 | 1.0 | 0.4999 | 0.5038 | 1.5155 | 0.0046 | 0.0019 | 0.0139 |
| | 0.5 | 1.0 | 1.0 | 0.5301 | 1.0040 | 1.0341 | 0.0105 | 0.0041 | 0.0134 |
| | 0.5 | 1.5 | 2.0 | 0.6161 | 1.4959 | 2.1283 | 0.0230 | 0.0085 | 0.0386 |
| | 1.0 | 1.5 | 2.0 | 1.1401 | 1.5086 | 2.0540 | 0.0454 | 0.0108 | 0.0404 |
| | 1.5 | 1.5 | 2.0 | 1.6533 | 1.5198 | 2.0340 | 0.0565 | 0.0120 | 0.0363 |
| | 1.0 | 1.0 | 2.0 | 1.9977 | 1.0140 | 1.0006 | 0.0384 | 0.0076 | 0.0138 |
| | 2.0 | 0.5 | 1.0 | 1.9912 | 0.5088 | 1.0122 | 0.0178 | 0.0038 | 0.0118 |
| | 2.0 | 0.5 | 2.0 | 2.0012 | 0.5066 | 2.0139 | 0.0160 | 0.0033 | 0.0191 |
| | 2.0 | 0.5 | 1.5 | 2.0412 | 0.4921 | 1.5583 | 0.0153 | 0.0029 | 0.0137 |
| | 2.0 | 0.5 | 0.5 | 2.0110 | 0.5018 | 0.5062 | 0.0170 | 0.0033 | 0.0052 |
| | 2.0 | 1.5 | 0.5 | 2.1140 | 1.5159 | 0.5039 | 0.0722 | 0.0117 | 0.0086 |
| | 2.0 | 2.0 | 0.5 | 2.7353 | 2.0066 | 0.5251 | 0.1319 | 0.0184 | 0.0113 |

## 8. Bivariate extension

Here, we propose an extension of the WPF model using the results of Marshall and Olkin [33].

**8.1. Theorem.** *Let* $X_1 \sim WPF(a_1, b, \alpha, \beta)$, $X_2 \sim WPF(a_2, b, \alpha, \beta))$ *and* $X_3 \sim WPF(a_1, b, \alpha, \beta)$ *be independent random variables.*

*Let* $X = \min\{X_1, X_3\}$ *and* $Y = \min\{X_2, X_3\}$. *Then, the cdf of the bivariate random variable* $(X, Y)$ *is given by*

$$F_{X,Y}(x,y) = 1 - \mathrm{e}^{-a_1\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b - a_2\left[\frac{y^\beta}{\alpha^\beta - y^\beta}\right]^b - a_3\left[\frac{z^\beta}{\alpha^\beta - z^\beta}\right]^b},$$

*where* $z = \max\{x, y\}$.

*The marginal cdf's are given by*

$$F_X(x) \;= 1 - \mathrm{e}^{-(a_1+a_3)\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b}$$

*and*

$$F_Y(y) = 1 - \mathrm{e}^{-(a_2+a_3)\left[\frac{y^\beta}{\alpha^\beta - y^\beta}\right]^b}.$$

The pdf of $(X, Y)$ is given in the Corollary.

**8.2. Corollary.** *Let $X$ and $Y$ defined as in Theorem 8.1,*

$$f_{X,Y}(x,y) = \begin{cases} f_{\mathrm{WPF}}(x\,; a_1, b, \alpha, \beta)\, f_{\mathrm{WPF}}(y\,; a_2 + a_3, b, \alpha, \beta), & \textit{for } x < y; \\ f_{\mathrm{WPF}}(x\,; a_1 + a_3, b, \alpha, \beta)\, f_{\mathrm{WPF}}(y\,; a_2, b, \alpha, \beta), & \textit{for } x > y; \\ \dfrac{a_3}{a_1 \,+\, a_2 \,+\, a_3}\, f_{\mathrm{WPF}}(x\,; a_1 + a_2 + a_3, b, \alpha, \beta), & \textit{for } x = y. \end{cases}$$

*The marginal pdf's are given by*

$$f_X(x) = \frac{(a_1 + a_3)\, b\, \beta\, \alpha^\beta x^{\beta\, b-1}}{(\alpha^\beta - x^\beta)^{b+1}}\, \mathrm{e}^{-a\left[\frac{x^\beta}{\alpha^\beta - x^\beta}\right]^b}$$

*and*

$$f_Y(y) = \frac{(a_2 + a_3)\, b\, \beta\, \alpha^\beta y^{\beta\, b-1}}{(\alpha^\beta - y^\beta)^{b+1}}\, \mathrm{e}^{-a\left[\frac{y^\beta}{\alpha^\beta - y^\beta}\right]^b}.$$

## 9. Applications

In this section, we provide two application to real data in order to illustrate the importance of the WPF distribution. The MLEs of the parameters are determined for the WPF and four other models, and seven goodness-of-fit statistics are computed for checking the adequacy of the all five fitted models.

**9.1. Data set 1: Aarset data.** The first real data set refers to the failure times of 50 items put under a life test. This data set is well-known to exhibit bathtub behavior of the hrf. Aarset [1] first reported these data set which has been analyzed by many authors. The data are: 0.1, 0.2, 1.0, 1.0, 1.0, 1.0, 1.0, 2.0, 3.0, 6.0, 7.0, 11.0, 12.0, 18.0, 18.0, 18.0, 18.0, 18.0, 21.0, 32.0, 36.0, 40.0, 45.0, 45.0, 47.0, 50.0, 55.0, 60.0, 63.0, 63.0, 67.0, 67.0, 67.0, 67.0, 72.0, 75.0, 79.0, 82.0, 82.0, 83.0, 84.0, 84.0, 84.0, 85.0, 85.0, 85.0. 85.0. 85.0. 86.0. 86.0.

**9.2. Data set 2: Device failure times data.** The second real data set refers to 30 devices failure times given in Table 15.1 by Meeker and Escobar [35]. The data are: 275, 13, 147, 23, 181, 30, 65, 10, 300, 173, 106, 300, 300, 212, 300, 300, 300, 2, 261, 293, 88, 247, 28, 143, 300, 23, 300, 80, 245, 266.

We fit the WPF model and other competitive models to both data sets. The other fitted models are: the additive Weibull (AddW) [54], modified-Weibull (MW) [30], Sarhan-Zaindin modified Weibull (SZMW) [48] and beta-modified Weibull (BMW) [50]. Their associated densities are given by:

$$\text{AddW}: \ f_{AddW}(x;\alpha,\beta,\theta,\gamma) = \left(\alpha\,\theta\,x^{\theta-1} + \beta\,\gamma\,x^{\gamma-1}\right) \mathrm{e}^{-\alpha\,x^{\theta} - \beta\,x^{\gamma}}, \quad x>0, \quad \alpha,\beta,\theta,\gamma>0,$$

$$\text{MW}: \ f_{MW}(x;\beta,\gamma,\lambda) = \beta\,(\gamma+\lambda\,x)\ x^{\gamma-1}\,\mathrm{e}^{\lambda\,x}\,\mathrm{e}^{-\beta\ x^{\gamma},\mathrm{e}^{\lambda\,x}},\ x>0,\ \beta,\gamma,\lambda>0,$$

$$\text{SZMW}: \ f_{SZMW}(x;\alpha,\beta,\gamma) = \left(\alpha + \beta\,x^{\gamma-1}\right)\ \mathrm{e}^{-\alpha\,x - \beta\,x^{\gamma}},\ x>0, \quad \alpha,\beta,\gamma>0,$$

$$\text{BMW}: \ f_{BMW}(x;a,b,\alpha,\beta,\lambda) = \tfrac{1}{B(a,b)}\,\alpha\,(\beta+\lambda\,x)\ x^{\beta-1}\,\mathrm{e}^{\lambda\,x}\,\mathrm{e}^{-\alpha\,b\,x^{\beta}} \times \left(1-\mathrm{e}^{-\alpha\,x^{\beta}\mathrm{e}^{\lambda x}}\right)^{a-1},\ x>0, \quad a,b,\alpha,\beta,\lambda>0.$$

The required computations are carried out using a script of the R-language [42], the AdequacyModel, written by Pedro Rafael Diniz Marinho, Cícero Rafael Barros Dias and Marcelo Bourguignon [32] which is freely available. In AdequacyModel package, there exists many maximization algorithms like NR (Newton-Raphson), BFGS (Broyden-Fletcher-Goldfarb-Shanno), BHHH(Berndt-Hall-Hall-Hausman), SANN (Simulated-Annealing), NM (Nelder-Mead) and Limited-Memory quasi-Newton code for Bound-constrained optimization (L-BFGS-B). But here, the MLEs are computed using LBFGS-B method.

The measures of goodness of fit including the log-likelihood function evaluated at the MLEs ($\hat{\ell}$), Akaike information criterion (AIC), consistent Akaike information criterion (CAIC), Hannan-Quinn information criterion (HQIC), Bayesian information criterion (BIC), Anderson-Darling ($A^*$) and Cramér–von Mises ($W^*$) to compare the fitted models. The statistics $W^*$ and $A^*$ are well-defined by Chen and Balakrishnan [17]. In general, the smaller the values of these statistics, the better the fit to the data.

Tables 3 and 5 list the MLEs and their corresponding standard errors (in parentheses) of the model parameters. The numerical values of the statistics $\hat{\ell}$, AIC, CAIC, BIC, HQIC, $W^*$ and $A^*$ are listed in Tables 4 and 6.

Table 3: MLEs and their standard errors (in parentheses) for Aarset data.

| Distribution | $a$ | $b$ | $\alpha$ | $\beta$ | $\theta$ | $\gamma$ | $\lambda$ |
|---|---|---|---|---|---|---|---|
| WPF | 0.7347 | 0.3367 | 86.0 | 1.4898 | - | - | - |
| | (0.2096) | (0.0567) | - | (0.4879) | - | - | - |
| AddW | - | - | 0.0020 | 0.0892 | 1.5164 | 0.3454 | - |
| | - | - | (0.0003) | (0.0424) | (0.0523) | (0.1125) | - |
| MW | - | - | - | 0.0624 | - | 0.3550 | 0.0233 |
| | - | - | - | (0.0266) | - | (0.1126) | (0.0048) |
| SZMW | - | - | 0.0186 | 0.0405 | - | 0.3735 | - |
| | - | - | (0.0038) | (0.0311) | - | (0.1886) | - |
| BMW | 0.2589 | 0.1525 | 0.0034 | 1.0819 | - | - | 0.0401 |
| | (0.0704) | (0.0834) | (0.0015) | (0.2928) | - | - | (0.0122) |

Table 4: The statistics $\hat{\ell}$, AIC, CAIC, BIC, HQIC, $A^*$ and $W^*$ for Aarset data.

| Distribution | $\hat{\ell}$ | AIC | CAIC | BIC | HQIC | $A^*$ | $W^*$ |
|---|---|---|---|---|---|---|---|
| WPF | 205.1732 | 416.3464 | 416.8681 | 422.0824 | 418.5307 | 0.380 | 0.046 |
| AddW | 234.2362 | 476.4725 | 477.3614 | 484.1206 | 479.3849 | 2.174 | 0.343 |
| MW | 227.1552 | 460.3105 | 460.8322 | 466.0465 | 462.4948 | 1.604 | 0.234 |
| SZMW | 239.4842 | 484.9684 | 485.4901 | 490.7045 | 487.1527 | 2.799 | 0.454 |
| BMW | 222.0914 | 454.1827 | 455.5464 | 463.7429 | 457.8233 | 1.276 | 0.169 |

Table 5: MLEs and their standard errors (in parentheses) for Aarset data.

| Distribution | $a$ | $b$ | $\alpha$ | $\beta$ | $\theta$ | $\gamma$ | $\lambda$ |
|---|---|---|---|---|---|---|---|
| WPF | 0.7723 | 0.24487 | 300.0 | 2.8736 | - | - | - |
| | (0.2519) | (0.0553) | - | (1.1351) | - | - | - |
| AddW | - | - | 3.4823E-03 | 1.0000E-10 | 1.0936 | 1.2045E-10 | - |
| | - | - | (1.3515E-03) | (1.1991E-06) | (7.6001E-02) | (9.2675E-11) | - |
| MW | - | - | - | 0.0313 | - | 0.3054 | 0.0081 |
| | - | - | - | (0.0240) | - | (0.1678) | (0.0020) |
| SZMW | - | - | 5.6560E-03 | 1.1789E-05 | - | 7.5972E-03 | - |
| | - | - | (1.0088E-03) | (1.1222E-05) | - | (3.0831E-06) | - |
| BMW | 0.3846 | 0.1832 | 0.0029 | 0.8382 | - | - | 0.0110 |
| | (0.1443) | (0.1305) | (0.0012) | (0.2770) | - | - | (0.0045) |

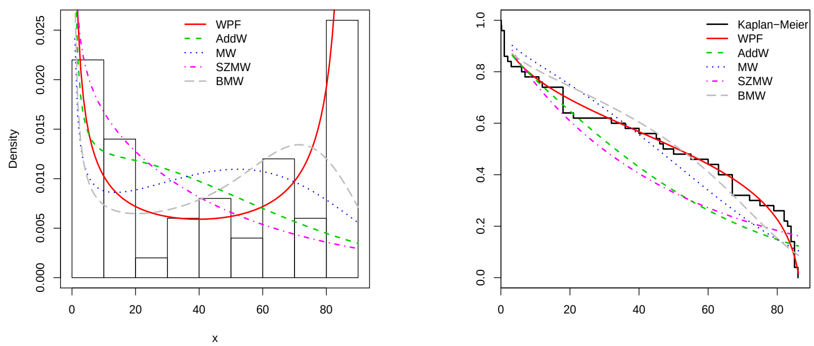
Table 6: The statistics $\hat{\ell}$, AIC, CAIC, BIC, HQIC, $A^*$ and $W^*$ for device failure times data.

| Distribution | $\hat{\ell}$ | AIC | CAIC | BIC | HQIC | $A^*$ | $W^*$ |
|---|---|---|---|---|---|---|---|
| WPF | 152.5768 | 311.1535 | 312.0766 | 315.3571 | 312.4983 | 0.750 | 0.082 |
| AddW | 184.7103 | 377.4206 | 379.0206 | 383.0254 | 379.2136 | 1.872 | 0.314 |
| MW | 178.3303 | 362.6606 | 363.5837 | 366.8642 | 364.0054 | 1.396 | 0.207 |
| SZMW | 185.2905 | 376.5810 | 377.5041 | 380.7846 | 377.9258 | 1.906 | 0.321 |
| BMW | 175.7578 | 361.5157 | 364.0157 | 368.5216 | 363.7569 | 1.262 | 0.182 |

In Tables 4 and 6, we compare the WPF model with the WPF, AddW, MW, SZMW and BMW models. We note that the WPF model gives the lowest values for the $\hat{\ell}$, AIC, CAIC, BIC, HQIC, $A^*$ and $W^*$ statistics for both data sets among the fitted models. So, the WPF model could be chosen as the best model. The histogram of the data sets, and plots the estimated densities and Kaplan-Meier are displayed in Figures 5 and 6. It is clear from Tables 4 and 6 and Figures 5 and 6 that the WPF model provides the best fits to the histogram of these two data sets.

## 10. Concluding remarks

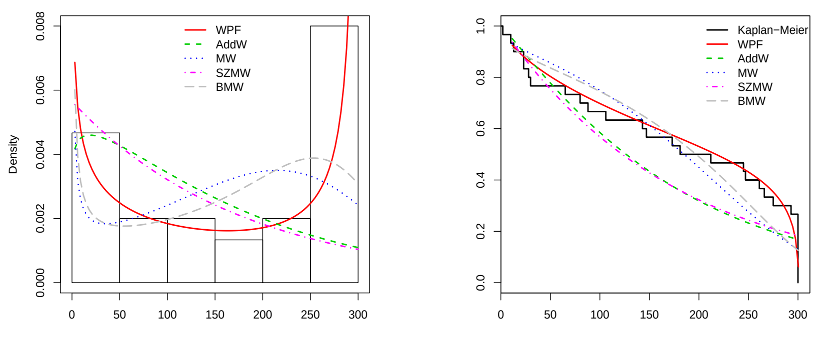
Many new lifetime distributions have been constructed in recent years with a view for better applications in various fields. They usually arise from an adequate transformation of a very-known model. In this paper, we propose a new lifetime model, the Weibull-power function (WPF) distribution, by applying the Weibull-G generator pioneered by Bourguignon *et al.* [15] to the classical power function distribution. We study some of its structural properties including an expansion for the density function and explicit expressions for the ordinary and incomplete moments, generating function, mean deviations, quantile function, entropies, reliability and order statistics. The maximum likelihood method is employed for estimating the model parameters and a simulation study is presented. The WPF model is fitted to two real data sets to illustrate the usefulness of the distribution. It provides consistently a better fit than other competing models. Finally, we hope that the proposed model will attract wider applications in reliability engineering, survival and lifetime data, mortality study and insurance, hydrology, social sciences, economics, among others.




(a) Estimated pdfs

(b) Estimated sfs

**Figure 5.** Plots of the estimated pdfs and sfs for the WPF, AddW, MW, SZMW and BMW models for the data set 1.




(a) Estimated pdfs

(b) Estimated sfs

**Figure 6.** Plots of the estimated pdfs and sfs for the WPF, AddW, MW, SZMW and BMW models for the data set 2.

## Appendix A

The elements of the $3\times 3$ observed information matrix $J(\Theta)=\{U_{rs}\}$ (for $r,s=a,b,\beta$) are given by

$$
\begin{aligned}
U_{aa} &= -\frac{n}{a^2},\\
U_{ab} &= -\sum_{i=1}^{n}\left[\frac{x_i^{\beta}}{\alpha^{\beta}-x_i^{\beta}}\right]^{b}\log\left[\frac{x_i^{\beta}}{\alpha^{\beta}-x_i^{\beta}}\right],\\
U_{a\beta} &= -b\,\alpha^{\beta}\sum_{i=1}^{n}\left[\frac{x_i^{b\,\beta}\,\log\,(x_i/\alpha)}{\left(\alpha^{\beta}-x_i^{\beta}\right)^{b+1}}\right],\\
U_{bb} &= -\frac{n}{b^2}-a\sum_{i=1}^{n}\left[\frac{x_i^{\beta}}{\alpha^{\beta}-x_i^{\beta}}\right]^{b}\left\{\log\left[\frac{x_i^{\beta}}{\alpha^{\beta}-x_i^{\beta}}\right]\right\}^{2},\\
U_{b\beta} &= \sum_{i=1}^{n}\log x_i-\sum_{i=1}^{n}\left[\frac{\alpha^{\beta}\,\log\alpha-x_i^{\beta}\log\,x_i}{\alpha^{\beta}-x_i^{\beta}}\right]\\
&\quad -a\,\alpha^{\beta}\sum_{i=1}^{n}\left[\frac{x_i^{b\,\beta}\log\,(x_i/\alpha)}{(\alpha^{\beta}-x_i^{\beta})^{b+1}}\right]\left[1+b\log(x_i^{\beta}/(\alpha^{\beta}-x_i^{\beta}))\right],\\
U_{\beta\beta} &= -\frac{n}{\beta^2}-(b+1)\sum_{i=1}^{n}\left[\frac{\left(\alpha^{\beta}-x_i^{\beta}\right)\left\{\alpha^{\beta}\,(\log\,\alpha)^2-x_i^{\beta}\,(\log\,x_i)^2\right\}}{(\alpha^{\beta}-x_i^{\beta})^2}\right.\\
&\quad \left.-(\alpha^{\beta}\,\log\,\alpha-x_i^{\beta}\,\log\,x_i)^2\right]\\
&\quad -a\,b\alpha^{\beta}\sum_{i=1}^{n}(\alpha^{\beta}-x_i^{\beta})^{b}\left[\frac{\left(\alpha^{\beta}-x_i^{\beta}\right)x_i^{b\beta}\ \{b\,\log\,x_i+\log\,\alpha\}\,\log\,(x_i/\alpha)}{(\alpha^{\beta}-x_i^{\beta})^{2(b+1)}}\right.\\
&\qquad \left.-\frac{(b+1)\,\alpha^{\beta}x_i^{b\beta}\ \log\,(x_i/\alpha)\left(\alpha^{\beta}\,\log\,\alpha-x_i^{\beta}\log\,x_i\right)}{(\alpha^{\beta}-x_i^{\beta})^{2(b+1)}}\right].
\end{aligned}
$$

## Acknowledgments

The authors would like to thank the Editor and the two referees for careful reading and the comments which greatly improved the paper.

## References

[1] Aarset, M.V. *How to identify bathtub hazard rate*, IEEE Transactions on Reliability **36**, 106–108, 1987.

[2] Ahsanullah, M. *A characterization of the power function distribution*, Communications in Statistics **2**, 259–262, 1973.

[3] Ahsanullah, M. *Estimation of the parameters of a power function distribution by record values*, Pakistan Journal of Statistics **5**, 189–194, 1989.

[4] Ahsanullah, M. *On lower generalized order statistics and a characterization of power function distribution*, Statistical Methods **7**, 16–28, 2005.

[5] Ahsanullah, M., Shakil, M. and Golam Kibria, B.M.G. *A characterization of the power function distribution based on lower records*, ProbStat Forum **6**, 68–72, 2013.

[6] Ali, M.M. and Woo, J. *Inference on reliability $P(Y<X)$ in a power function distribution*, Journal of Statistics and Management Science **8**, 681–686, 2005.

[7] Ali, M.M., Woo, J. and Nadarajah, S. *On the ratio of $X/(X+Y)$ for the power function distribution*, Pakistan Journal of Statistics **21**, 131–138, 2005.

[8] Ali, M.M., Woo, J. and Yoon, G.E. *The UMVUE of mean and right-tail probability in a power function distribution*, Estadistica **52**, 1–10, 2000.

[9] Asghar, Z., Shahzad, M.N. and Shehzad, F. *Parameter estimation of power function distribution with moments*, World Applied Science Journal **26**, 205–212, 2013.

[10] Athar, H. and Faizan, M. *Moments of lower generalized order statistics from power function distribution and its characterization*, International Journal of Statistical Sciences **11**, 125–134, 2011.

[11] Azedine, G. *Characterization of the power function distribution based on lower records*, Applied Mathematical Sciences **7**, 5259–5267, 2013.

[12] Balakrishnan, N. and Joshi, P.C. *Moments of order statistics from doubly truncated power function distribution*, Aligarh Journal of Statististics **1**, 98–105, 1981.

[13] Balakrishnan, N. and Nevzorov, V.B. *A Primer on Statistical Distributions* (Wiley, New York, 2003).

[14] Bhatt, M.B. *Characterization of power function distribution through expectation*, Open Journal of Statistics **3**, 441–443, 2013.

[15] Bourguignon, M., Silva, R.B. and Cordeiro, G. M. *The Weibull–G family of probability distributions*, Journal of Data Science **12**, 53–68, 2014.

[16] Chang, S-K. *Characterizations of the power function distribution by the independence of record values*, Journal of Chungcheong Mathematical Society **20**, 139–146, 2007.

[17] Chen, G. and Balakrishnan, N. *A general purpose approximate goodness-of-fit test*, Journal of Quality Technology **27**, 154–161, 1995.

[18] Cordeiro, G.M. and Brito, R.D.S. *The beta power distribution*, Brazalian Journal of Probability and Statistics **26**, 88–112, 2012.

[19] Dallas, A.C. *Characterization of Pareto and power function distribution*, Annals of Mathematical Statistics **28**, 491–497, 1976.

[20] Doornik, J.A. *Ox 5: An Object-Oriented Matrix Programming Language*, fifth edition (Timberlake Consultants, London, 2007).

[21] Forbes, C., Evans, M., Hastings, N. and Peacock, B. *Statistical Distributions*, fourth edition (Wiley, New York, 2011).

[22] Govindarajulu, Z. *Characterization of exponential and power function distribution*, Scandinavian Acturial Journal **3&4**, 132–136, 1966.

[23] Johnson, N.L., Kotz, S. and Balakrishnan, N. *Continuous Univariate Distributions*, Volume 1, Second edition (Wiley, New York, 1994).

[24] Johnson, N.L., Kotz, S. and Balakrishnan, N. *Continuous Univariate Distributions*, Volume 2, Second edition (Wiley, New York, 1995).

[25] Kabir, A.B.M.L. and Ahsanullah, M. *Estimation of location and scale parameters of a power function distribution by linear functions of order statistics*, Communications in Statistics **3**, 463–467, 1974.

[26] Kapadia, C.H. and Kromer, R.E. *Sample size required to estimate a parameter in the power function distribution*, Work of Statistical and Operations Research **29**, 88–96, 1978.

[27] Kenney, J. and Keeping, E. *Mathematics of Statistics*,Volume 1 (Princeton, 1962).

[28] Kifayat, T., Aslam, M. and Ali, S. *Bayesian inference for the parameters of the power function distribution*, Journal of Reliaboility and Statistical Studies **5**, 54–58, 2012.

[29] Kleiber, C. and Kotz, S. *Statistical Size Distributions in Economics and Actuarial Sciences* (Wiley, New York, 2003).

[30] Lai, C.D. and Xie, M. *A modified Weibull distribution*, IEEE Transactions on Reliability **52**, 33–37, 2003.

[31] Malik, H.J. *Exact moments of order statistics for a power function distribution*, Scandinavian Acturial Journal **1-2**, 64–69, 1967.

[32] Marinho, Cícero, P.R.D., Dias, R.B. and Bourguignon, M. AdequacyModel. http://cran.r-project.org/web/packages/AdequacyModel/AdequacyModel.pdf, 2014.

[33] Marshall, A.W. and Olkin, I. *A multivariate exponential distribution*, Journal Of the American Statistical Association **62**, 30–44, 1967.

[34] Mbah, A. and Ahsanullah, M. *Some characterization of power function distribution based on lower generalized order statistics*, Pakistan Journal of Statistics **23**, 139–146, 2007.

[35] Meeker, W.Q. and Escobar, L.A. *Statistical Methods for Reliability Data* (Wiley, New York, 1998).

[36] Meniconi, M. and Barry, D.M. *The power function distribution: A useful and simple distribution to assess electrical component reliability*, Microeolectronics Reliability **36**, 1207–1212, 1996.

[37] Moors, J.J.A. *A quantile alternative for kurtosis*, The Statistician **37**, 25–32, 1998.

[38] Moothathu, T.S.K. *Characterization of power function and Pareto distribution in terms of their Lorenz curve*, Research Review **3**, 118–123, 1984.

[39] Moothathu, T.S.K. *A characterization of power function distribution through a property of the Lorenz curve*, Sankhyã **B48**, 262–265, 1986.

[40] Proctor, J.W. *Estimation of two generalized curves covering the Pearson system*, Proceedings of The American Statistical Association (Computing Section), pp. 287–289, 1987.

[41] Rahman, H., Roy, M.K. and Baizid, A.R. *Bayes estimation under conjugate prior for the case of power function distribution*, American Journal of Mathematics and Statistics **2**, 44–48, 2012.

[42] R Development Core Team, R. A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2012.

[43] Rényi, A. *On measures of entropy and information*, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability-I (University of California Press, Berkeley, 1961), 547–561.

[44] Rider, P.R. *Distribution of product and quotient of maximum values in samples from a power function population*, Journal of the American Statistical Association **59**, 877–880, 1964.

[45] Saleem, M., Aslam, M. and Economou, P. *On the Bayesian analysis of the mixture of power function distribution using the complete and the censored sample*, Journal of Applied Statistics **37**, 25–40, 2010.

[46] Saran, J. and Pandey, A. *Estimation of parameters of power function distribution and its characterization by the kth record values*, Statistica **64**, 523–536, 2004.

[47] Saran, J. and Singh, A. *Recurrence relations for marginal and joint moment generating functions of generalized order statistics from power function distribution*, Metron **61**, 27–33, 2003.

[48] Sarhan, A.M. and Zaindain, M. *Modified Weibull distribution*, Applied Sciences **11**, 123–136, 2009.

[49] Shannon, C.E. *A mathematical theory of communication*, Bell System Technical Journal **27**, 379–432, 1948.

[50] Silva, G.O., Ortega, E.M.M. and Cordeiro, G.M. *The beta modified Weibull distribution*, Lifetime Data Analysis **16**, 409–430, 2010.

[51] Sinha, S.K., Singh, P., Singh, D.C. and Singh, R. *Preliminary test estimator for scale parameter of the power function distribution*, Journal of Reliability and Statistical Studies **1**, 18–27, 2008.

[52] Sultan, K.S., Childs,,A. and N. Balakrishnan, *Higher order moments of order statistics from the power function distribution and Edgeworth approximate inference*, in: Advances in Stochastic Simulation Methods, N. Balakrishnan, V. B. Melas S. Ermakov (eds.) (Springer, New York, 2000), 245–282.

[53] Tavangar, M. *Power function distribution characterized by dual generalized order statistics*, Journal of the Iranian Statistical Society **10**, 13–27, 2011.

[54] Xie, M., Tang, Y. and Goh, T.N. *A modified Weibull extension with bathtub-shaped failure rate function*, Reliability Engineering and System Safety **76** 279–285, 2002.

[55] Zaka, A. and Akhter, A.S. *Methods for estimating the parameters of power function distribution*, Pakistan Jouranl of Statistics and Operations Research **9**, 213–224, 2013.

[56] Zaka, A., Feroze, N. and Akhter, A.S. *A note on modified estimators for the parameters of power function distribution*, International Journal of Advanced Science and Technology **59**, 71–84, 2013.

[57] Zarrin, S., Saxena, S., Kamal, M. and Arif-ul-Islam. *Reliability computation and Bayesain analysis of system reliability of power function distribution*, International Journal of Advances in Engineering, Science and Technology **2**, 76–86, 2013.

[58] Zografos, K. and Balakrishnan, N. *On families of beta- and generalized gamma-generated distributions and associated inference*. Statistical Methodology **6**, 344–362, 2009.

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 267 – 278

# Inference on $Pr(X > Y)$ Based on Record Values from the Burr Type X Distribution

Bahman Tarvirdizade [*†] and Hossein Kazemzadeh Gharehchobogh[*]

**Abstract**

Our interest is in estimating the stress-strength reliability $Pr(X > Y)$ based on lower record values when X and Y are two independent but not identically distributed Burr type X random variables. The maximum likelihood estimator, Bayes and empirical Bayes estimators using Lindleys approximations, are obtained and their properties are studied. The exact confidence interval, as well as the Bayesian credible sets are obtained. Two examples are presented in order to illustrate the inferences discussed in the previous sections. A Monte Carlo simulation study is conducted to investigate and compare the performance of different types of estimators presented in this paper and to compare them with some bootstrap intervals.

**Keywords:** Likelihood estimation, Bayesian estimation, Burr type X distribution, Record values, Stress-strength reliability, Lindley approximation.

*2000 AMS Classification:* 62F10, 62F15, 62N05.

*Received* : 17.10.2014 *Accepted* : 02.01.2015 *Doi* : 10.15672/HJMS.2015468581

## 1. Introduction

The problem of estimating $R = P(X > Y)$ arises in the context of mechanical reliability of a system with strength X and stress Y and R is chosen as a measure of system reliability. The system fails if and only if, at any time the applied stress is grater than its strength. This type of reliability model is known as the stress-strength model. This problem also arises in situations where X and Y represent lifetimes of two devices and one wants to estimate the probability that one fails before the other. For example, in biometrical studies, the random variable X may represent the remaining lifetime of a patient treated with a certain drug while Y represent the remaining lifetime when treated by another drug. The estimation of stress-strength reliability is very common in the statistical literature. The reader is referred to Kotz et al. [1] for other applications and motivations for the study of the stress-strength reliability.

[*]Department of Statistics, Maku Branch, Islamic Azad University, Maku, Iran.
[†]Corresponding Author. Email: bahtary@gmail.com

Record values arise naturally many real life applications involving data relating to meteorology, hydrology, sports and life-tests. In industry and reliability studies, many products may fail under stress. For example, a wooden beam breaks when sufficient perpendicular force is applied to it, an electronic component ceases to function in an environment of too high temperature, and a battery dies under the stress of time. But the precise breaking stress or failure point varies even among identical items. Hence, in such experiments, measurements may be made sequentially and only values larger (or smaller) than all previous ones are recorded. Data of this type are called record data. Thus, the number of measurements made is considerably smaller than the complete sample size. This measurement saving can be important when the measurements of these experiments are costly if the entire sample was destroyed.

Let $\{X_i,\ i \geq 1\}$ be a sequence of independent and identically distributed (iid) random variables with an absolutely continuous cumulative distribution function (cdf) $F(x)$ and probability density function (pdf) $f(x)$. An observation $X_j$ is called an upper record if its value exceeds all previous observations, i.e. $X_j$ is an upper record if $X_j > X_i$ for every $i < j$. An analogous definition can be given for lower records. Records were first introduced and studied by Chandler [2]. Interested readers may refer to the book by Arnold et al. [3] and the references contained therein.

Burr [4] introduced twelve different forms of cumulative distribution functions for modeling lifetime data or survival data. Among those twelve distribution functions, Burr type X and Burr type XII received the maximum attention. Several aspects of the Burr type X distribution were studied by Sartawi and Abu-Salih [5], Jaheen [6] and Raqab [7]. The cumulative distribution function (cdf) and the probability density function (pdf) of the Burr type X distribution with shape parameter $\theta$, which will be denoted by $Burr(\theta)$, are respectively as follows,

$$F(x) = \left(1 - e^{-x^2}\right)^{\theta}, \quad x > 0, \quad \theta > 0, \tag{1.1}$$

$$f(x) = 2\theta x e^{-x^2}\left(1 - e^{-x^2}\right)^{\theta - 1}, \quad x > 0, \quad \theta > 0. \tag{1.2}$$

The problem of estimating the stress-strength reliability in the Burr type X distribution was considered by Ahmad et al. [8] and Surles and Padgett [9]. Kim and Chung [10] discussed Bayesian estimation of stress-strength reliability from Burr type X model containing spurious observations. We consider the problem of point and interval estimating the stress-strength reliability in the Burr type X distribution based on lower record values. The problem of interval estimating the stress-strength reliability based on record values was considered by Baklizi [11] for the generalized exponential distribution.

The rest of this paper is organized as follows: In Section 2, we discussed likelihood inference for the stress-strength reliability, while in Section 3 we considered Bayesian inference. In Section 4, we presented two numerical examples. A Monte Carlo simulation study is described in Section 5. Finally conclusion of the paper is provided in section 6.

## 2. Likelihood inference

Let X and Y be independent random variables from the Burr type X distribution with the parameters $\theta_1$ and $\theta_2$ respectively. Let $R = \Pr(X > Y)$ be the stress-strength reliability. then,

$$R = \int_0^\infty \int_y^\infty 2\theta_1 x e^{-x^2}\left(1 - e^{-x^2}\right)^{\theta_1 - 1} 2\theta_2 y e^{-y^2}\left(1 - e^{-y^2}\right)^{\theta_2 - 1} dx dy = \frac{\theta_1}{\theta_1 + \theta_2}.$$

Our interest is in estimating $R$ based on lower record values on both variables. Let $\underset{\sim}{r} = (r_1, ..., r_n)$ be a set of lower records from $Burr(\theta_1)$ and let $\underset{\sim}{s} = (s_1, ..., s_m)$ be an independent set of lower records from $Burr(\theta_2)$. The likelihood functions are given by (Ahsanullah [12]),

$$L(\theta_1|\underset{\sim}{r}) = f(r_n)\prod_{i=1}^{n-1}\left(\frac{f(r_i)}{F(r_i)}\right), \quad 0 < r_n < ... < r_1 < \infty,$$

$$L(\theta_2|\underset{\sim}{s}) = g(s_m)\prod_{i=1}^{m-1}\left(\frac{g(s_i)}{G(s_i)}\right), \quad 0 < s_m < ... < s_1 < \infty. \tag{2.1}$$

where $f$ and $F$ are the pdf and cdf of $X \sim Burr(\theta_1)$ respectively and $g$ and $G$ are the pdf and cdf of $Y \sim Burr(\theta_2)$ respectively. Substituting $f$ , $F$ , $g$ and $G$ in the likelihood functions and using Equation(2.1), we obtain

$$L(\theta_1|\underset{\sim}{r}) = (2\theta_1)^n\left(1-e^{-r_n^2}\right)^{\theta_1}\prod_{i=1}^{n}\left(\frac{r_i e^{-r_i^2}}{1-e^{-r_i^2}}\right),$$

$$L(\theta_2|\underset{\sim}{s}) = (2\theta_2)^m\left(1-e^{-s_m^2}\right)^{\theta_2}\prod_{i=1}^{m}\left(\frac{s_i e^{-s_i^2}}{1-e^{-s_i^2}}\right). \tag{2.2}$$

It can be shown that the maximum likelihood estimators (MLE) of $\theta_1$ and $\theta_2$ based on the lower record values are

$$\hat{\theta}_1 = -\frac{n}{\ln(1-e^{-r_n^2})}, \quad \hat{\theta}_2 = -\frac{m}{\ln(1-e^{-s_m^2})}. \tag{2.3}$$

Therefore using the invariance properties of the maximum likelihood estimation, the MLE of R is given by

$$\hat{R} = \frac{\hat{\theta}_1}{\hat{\theta}_1+\hat{\theta}_2}.$$

To study the distribution of $\hat{R}$ we need the distributions of $\hat{\theta}_1$ and $\hat{\theta}_2$. Consider first $\hat{\theta}_1 = -\frac{n}{\ln(1-e^{-r_n^2})}$, the pdf of the $n$th lower record value $R_n$ is given by (Ahsanullah [12]),

$$\begin{aligned} f_{R_n}(r_n) &= \tfrac{1}{(n-1)!}f(r_n)[-\ln F(r_n)]^{n-1} \\ &= \tfrac{2\theta_1^n}{(n-1)!}r_n e^{-r_n^2}\left(1-e^{-r_n^2}\right)^{\theta_1-1}\left(-\ln\left(1-e^{-r_n^2}\right)\right)^{n-1}, 0 < r_n < \infty. \end{aligned} \tag{2.4}$$

Consequently, the pdf of $Z_1 = \hat{\theta}_1$ is given by,

$$f_{Z_1}(z_1) = \frac{(n\theta_1)^n}{(n-1)!z_1^{n+1}}\exp\left(-\frac{n\theta_1}{z_1}\right), \quad z_1 > 0. \tag{2.5}$$

This is recognized as the inverted gamma distribution, i.e., $Z_1 \sim IGamma(n, n\theta_1)$. Similarly, the pdf of $Z_2 = \hat{\theta}_2$ is given by,

$$f_{Z_2}(z_2) = \frac{(m\theta_2)^m}{(m-1)!z_2^{m+1}}\exp\left(-\frac{m\theta_2}{z_2}\right), \quad z_2 > 0. \tag{2.6}$$

Thus $Z_2 \sim IGamma(m, m\theta_2)$. Therefore we can find the pdf of

$$\hat{R} = \frac{\hat{\theta}_1}{\hat{\theta}_1+\hat{\theta}_2} = \frac{Z_1}{Z_1+Z_2} = \frac{1}{1+\frac{Z_2}{Z_1}}.$$

Consider $Z_2/Z_1$. Note that, by the properties of the inverted gamma distribution and its relation with the gamma distribution we have $(n\theta_1/Z_1) \sim Gamma(n, 1)$ and $(n\theta_2/Z_2) \sim Gamma(m, 1)$. Hence $(2n\theta_1/Z_1) \sim \chi^2_{2n}$ and $(2m\theta_2/Z_2) \sim \chi^2_{2m}$. Note that, by the independence of two random quantities we have

$$\frac{(2n\theta_1/2nZ_1)}{(2m\theta_2/2mZ_2)} = \frac{\theta_1 Z_2}{\theta_2 Z_1} \sim F_{(2n,2m)}.$$

Hence, $(Z_2/Z_1) = (\theta_2/\theta_1)F_{(2n,2m)}$, has a scaled F distribution. It follows that the distribution of $\hat{R}$ is that of $\frac{1}{1+(\theta_2/\theta_1)F_{(2n,2m)}}$ which can be obtained using simple transformation techniques. This fact can be used to construct the following $(1-\alpha)\%$ confidence interval for R,

$$\left(\left(1+\frac{z_2}{z_1 F_{\alpha/2,2n,2m}}\right)^{-1},\left(1+\frac{z_2}{z_1 F_{1-\alpha/2,2n,2m}}\right)^{-1}\right). \tag{2.7}$$

## 3. Bayesian inference

Consider the likelihood functions of $\theta_1$ and $\theta_2$ based on the two sets of lower record values from the Burr type X distribution mentioned in Equation (2.2). We have

$$L(\theta_1|\underset{\sim}{r}) \propto \theta_1^n e^{-\theta_1\nu_1(r_n)}, \quad L(\theta_2|\underset{\sim}{s}) \propto \theta_2^m e^{-\theta_2\nu_2(s_m)} \tag{3.1}$$

where $\nu_1(r_n) = -\ln(1-e^{-r_n^2})$ and $\nu_2(s_m) = -\ln(1-e^{-s_m^2})$. These suggest that the conjugate family of prior distributions for $\theta_1$ and $\theta_2$ is the Gamma family of probability distributions,

$$\pi(\theta_1) = \frac{\gamma_1^{\delta_1}\theta_1^{\delta_1-1}e^{-\gamma_1\theta_1}}{\Gamma(\delta_1)}, \theta_1 > 0 \quad \textit{and} \quad \pi(\theta_2) = \frac{\gamma_2^{\delta_2}\theta_2^{\delta_2-1}e^{-\gamma_2\theta_2}}{\Gamma(\delta_2)}, \theta_2 > 0 \tag{3.2}$$

where $\delta_1$, $\gamma_1$, $\delta_2$ and $\gamma_2$ are the parameters of the prior distributions of $\theta_1$ and $\theta_2$ respectively. It can be shown that $(\theta_1|\underset{\sim}{r}) \sim \textit{Gamma}\,(n+\delta_1, \gamma_1+\nu_1(r_n))$ and $(\theta_2|\underset{\sim}{s}) \sim \textit{Gamma}\,(m+\delta_2, \gamma_2+\nu_2(s_m))$. Since the priors $\theta_1$ and $\theta_2$ are independent, then, using standard transformation techniques and after some manipulations, the posterior pdf of $R$ will be

$$f_R(r) = C\frac{r^{n+\delta_1-1}(1-r)^{m+\delta_2-1}}{[r(\gamma_1+\nu_1(r_n))+(1-r)(\gamma_2+\nu_2(s_m))]^{n+m+\delta_1+\delta_2}}, 0 < r < 1 \tag{3.3}$$

where

$$C = \frac{\Gamma(n+m+\delta_1+\delta_2)}{\Gamma(n+\delta_1)\Gamma(m+\delta_2)}(\gamma_1+\nu_1(r_n))^{n+\delta_1}(\gamma_2+\nu_2(s_m))^{m+\delta_2}.$$

The Bayes estimator under squared error loss is the mean of this posterior distribution which can not be computed analytically. Alternatively, using the approximate method of Lindley [13], it can be seen that the approximate Bayes estimator of $R$, say $\tilde{R}_B$, relative to squared error loss function is

$$\tilde{R}_B = \tilde{R}\left(1+\frac{(1-\tilde{R})^2}{n+\delta_1-1}-\frac{\tilde{R}(1-\tilde{R})}{m+\delta_2-1}\right) \tag{3.4}$$

where $\tilde{R} = \frac{\tilde{\theta}_1}{\tilde{\theta}_1+\tilde{\theta}_2}$ and

$$\tilde{\theta}_1 = \left(\frac{n+\delta_1-1}{\gamma_1+\nu_1(r_n)}\right), \quad \tilde{\theta}_2 = \left(\frac{m+\delta_2-1}{\gamma_2+\nu_2(s_m)}\right)$$

are the mode of the posterior densitys $\theta_1$ and $\theta_2$ respectively. On the other hand, it follows from the posterior density $\theta_1$ and $\theta_2$ that $2(\gamma_1+\nu_1(r_n))(\theta_1|\underset{\sim}{r}) \sim \chi^2_{2(n+\delta_1)}$ and $2(\gamma_2+\nu_2(s_m))(\theta_2|\underset{\sim}{s}) \sim \chi^2_{2(m+\delta_2)}$. It follows that $\pi(R|\underset{\sim}{r},\underset{\sim}{s})$, the posterior distribution of $R$ is equal to that of $(1+AW)^{-1}$, where $W \sim F_{2(m+\delta_2),2(n+\delta_1)}$ and $A = \frac{(m+\delta_2)(\gamma_1+\nu_1(r_n))}{(n+\delta_1)(\gamma_2+\nu_2(s_m))}$. Therefore a Bayesian $(1-\alpha)\%$ confidence interval for $R$ is given by,

$$\left((AF_{1-\alpha/2,2(m+\delta_2),2(n+\delta_1)}+1)^{-1},(AF_{\alpha/2,2(m+\delta_2),2(n+\delta_1)}+1)^{-1}\right). \tag{3.5}$$

The case of a noninformative prior can be treated similarly. We consider Jeffereys prior that say, $\pi(\theta_1) \propto \sqrt{|I(\theta_1)|}$ where $I(\theta_1)$ is the Fisher information. This suggest that prior densitys for $\theta_1$ and $\theta_2$ are proportional to $\frac{1}{\theta_1}$ and $\frac{1}{\theta_2}$ respectively. Using direct arguments one can show that $(\theta_1 | \underset{\sim}{r}) \sim Gamma\,(n, \nu_1(r_n))$ and $(\theta_2 | \underset{\sim}{s}) \sim Gamma\,(m, \nu_2(s_m))$. Therefore, it can be seen that the approximate Bayes estimator of $R$ under the Jeffreys prior density, say $\tilde{R}_{JB}$, relative to squared error loss function is

$$\tilde{R}_{JB} = \tilde{R}\left(1 + \frac{(1-\tilde{R})^2}{n-1} - \frac{\tilde{R}(1-\tilde{R})}{m-1}\right) \tag{3.6}$$

where $\tilde{R} = \frac{\tilde{\theta}_1}{\tilde{\theta}_1 + \tilde{\theta}_2}$ and

$$\tilde{\theta}_1 = \left(\frac{n-1}{\nu_1(r_n)}\right), \quad \tilde{\theta}_2 = \left(\frac{m-1}{\nu_2(s_m)}\right).$$

Furthermore, it follows that the posterior distribution of $R$ is equal to that of $(1 + \frac{m\nu_1(r_n)}{n\nu_2(s_m)}W)^{-1}$ where $W \sim F_{2m,2n}$. Therefore a Bayesian $(1-\alpha)\%$ confidence interval for $R$ is given by,

$$\left(\left(\frac{m\nu_1(r_n)}{n\nu_2(s_m)}F_{1-\alpha/2,2m,2n} + 1\right)^{-1}, \left(\frac{m\nu_1(r_n)}{n\nu_2(s_m)}F_{\alpha/2,2m,2n} + 1\right)^{-1}\right). \tag{3.7}$$

Now consider the case when the parameters of prior distributions are themselves unknown. We consider the conjugate prior distributions for $\theta_1$ and $\theta_2$ above when the parameters $\gamma_1$ and $\gamma_2$ are unknown. In the empirical Bayes model, we must estimate them. In order to, we calculate the marginal distribution of lower records, with densitys

$$m(\underset{\sim}{r}\,|\gamma_1) = \int f_{\underset{\sim}{R}}(\underset{\sim}{r}\,|\theta_1)\pi(\theta_1|\gamma_1)d\theta_1, \quad 0 < r_n < ... < r_1 < \infty,$$

$$m(\underset{\sim}{s}\,|\gamma_2) = \int f_{\underset{\sim}{S}}(\underset{\sim}{s}\,|\theta_2)\pi(\theta_2|\gamma_2)d\theta_2, \quad 0 < s_m < ... < s_1 < \infty.$$

Using Equations (2.2) and (3.2), we obtain

$$m(\underset{\sim}{r}\,|\gamma_1) = \frac{\Gamma(n+\delta_1)2^n\gamma_1^{\delta_1}}{\Gamma(\delta_1)(\gamma_1 + \nu_1(r_n))^{n+\delta_1}}\prod_{i=1}^{n}\left(\frac{r_i e^{-r_i^2}}{1-e^{-r_i^2}}\right),$$

$$m(\underset{\sim}{s}\,|\gamma_2) = \frac{\Gamma(m+\delta_2)2^m\gamma_2^{\delta_2}}{\Gamma(\delta_2)(\gamma_2 + \nu_2(s_m))^{m+\delta_2}}\prod_{i=1}^{m}\left(\frac{s_i e^{-s_i^2}}{1-e^{-s_i^2}}\right). \tag{3.8}$$

It can be shown that the maximum likelihood estimators (MLE) of $\gamma_1$ and $\gamma_2$ based on the marginal distributions (3.8) are

$$\hat{\gamma}_1 = \frac{\delta_1\nu_1(r_n)}{n}, \quad \hat{\gamma}_2 = \frac{\delta_2\nu_2(s_m)}{m}. \tag{3.9}$$

Substituting $\hat{\gamma}_1$ and $\hat{\gamma}_2$ into Equation (3.4), the approximate empirical Bayes estimator of $R$, say $\tilde{R}_{EB}$, relative to squared error loss function is given by,

$$\tilde{R}_{EB} = \tilde{R}^*\left(1 + \frac{(1-\tilde{R}^*)^2}{n+\delta_1-1} - \frac{\tilde{R}^*(1-\tilde{R}^*)}{m+\delta_2-1}\right) \tag{3.10}$$

where $\tilde{R}^* = \frac{\tilde{\theta}_1^*}{\tilde{\theta}_1^* + \tilde{\theta}_2^*}$ and

$$\tilde{\theta}_1^* = \left(\frac{n+\delta_1-1}{\nu_1(r_n)\left(1+\frac{\delta_1}{n}\right)}\right), \quad \tilde{\theta}_2^* = \left(\frac{m+\delta_2-1}{\nu_2(s_m)\left(1+\frac{\delta_2}{m}\right)}\right)$$

Furthermore, it can be shown that $(\theta_1|\underset{\sim}{r},\hat{\gamma}_1) \sim Gamma(n+\delta_1,(1+\frac{\delta_1}{n})\nu_1(r_n))$ and $(\theta_2|\underset{\sim}{s},\hat{\gamma}_2) \sim Gamma(m+\delta_2,(1+\frac{\delta_2}{m})\nu_2(s_m))$. It follows that $\pi(R|\underset{\sim}{r},\hat{\gamma}_1,\underset{\sim}{s},\hat{\gamma}_2)$, the empirical posterior distribution of $R$ is equal to that of $(1+\frac{m\nu_1(r_n)}{n\nu_2(s_m)}W)^{-1}$ where $W \sim F_{2(m+\delta_2),2(n+\delta_1)}$. Therefore a Bayesian $(1-\alpha)\%$ confidence interval for $R$ is given by,

$$\left(\left(\frac{m\nu_1(r_n)}{n\nu_2(s_m)}F_{1-\alpha/2,2(m+\delta_2),2(n+\delta_1)}+1\right)^{-1},\left(\frac{m\nu_1(r_n)}{n\nu_2(s_m)}F_{\alpha/2,2(m+\delta_2),2(n+\delta_1)}+1\right)^{-1}\right). \tag{3.11}$$

The construction of highest posterior density (HPD) regions requires finding the set $I=\{\theta:\pi(\theta|\underset{\sim}{r},\underset{\sim}{s})\geq k_\alpha\}$ , where $k_\alpha$ is the largest constant such that $\Pr(\theta\in I)\geq 1-\alpha$. This often requires numerical optimization techniques. Chen and Shao [14] presented a simple Monte Carlo technique to approximate the HPD region.

## 4. Illustrative examples

In this section, two numerical examples are presented to illustrate the inferences discussed in the previous sections.

**Example 4.1** (Real Data Set). We consider a data analysis for two data sets reported by Bennett and Filliben [15]. They have reported minority electron mobility for p-type $Ga_{1-x}Al_xAs$ with seven different values of mole fraction. We use two data sets related to the mole fractions 0.25 and 0.30. These data are given as follows:

*Data Set 1* (belongs to mole fraction 0.25): 3.051, 2.779, 2.604, 2.371, 2.214, 2.045, 1.715, 1.525, 1.296, 1.154, 1.016, 0.7948, 0.7007, 0.6292, 0.6175, 0.6449, 0.8881, 1.115, 1.397, 1.506, 1.528.

*Data Set 2* (belongs to mole fraction 0.30): 2.658, 2.434, 2.288, 2.092, 1.959, 1.814, 1.530, 1.366, 1.165, 1.041, 0.9198, 0.7241, 0.6403, 0.576, 0.5647, 0.5873, 0.8013, 1.002, 1.250, 1.347, 1.368.

We fit the Burr type X distribution to the two data sets separately. We used the Kolmogorov-Smirnov (K-S) tests for each data set to fit the Burr type X model. It is observed that for data sets 1 and 2, the K-S distances are 0.2453 and 0.2026 with the corresponding $p$-values 0.1395 and 0.3110, respectively. Therefore, it is clear that Burr type X model fits well to both the data sets. Moreover, we plot the empirical distribution functions and the fitted distribution functions in Figure 1. This figure show that the empirical and fitted models are very close for each data set.

For the above data, we observe that the first 15 values for both the data sets are the lower record values and the smallest records, $r_n$ and $s_m$, are equal to 0.6175 and 0.5647, respectively. Therefore, we obtain the MLEs of $\theta_1$ and $\theta_2$ as, 13.0576 and 11.5551, respectively. Thus, the MLE of $R$ becomes $\hat{R}=0.5305$. The corresponding 95% confidence interval based on Equation (2.7) is equal to (0.3527,0.7009). To obtain Bayes estimates, we assume $\delta_1=\delta_2=3$ and $\gamma_1=\gamma_2=2$ in Equation (3.4). We obtain $\tilde{\theta}_1=5.3990$, $\tilde{\theta}_2=5.15440$ and $\tilde{R}=0.5116$. Therefore, the approximate Bayes estimator of $R$ becomes $\tilde{R}_B=0.5113$. The corresponding Bayesian 95% confidence interval based on Equation (3.5) is equal to (0.3504,0.6704). So, the approximate Bayes estimator of $R$ based on Equation (3.6) becomes $\tilde{R}_{JB}=0.5294$ and the corresponding Bayesian 95% confidence interval based on Equation (3.7) is equal to (0.3527,0.7009). Finally, using Equation (3.10), we obtain $\tilde{\theta}_1^*=12.3322$, $\tilde{\theta}_2^*=10.9131$ and $\tilde{R}^*=0.5305$. Therefore, the approximate empirical Bayes estimator of $R$ becomes $\tilde{R}_{EB}=0.5269$. The corresponding Bayesian 95% confidence interval based on Equation (3.11) is equal to (0.3678,0.6869).

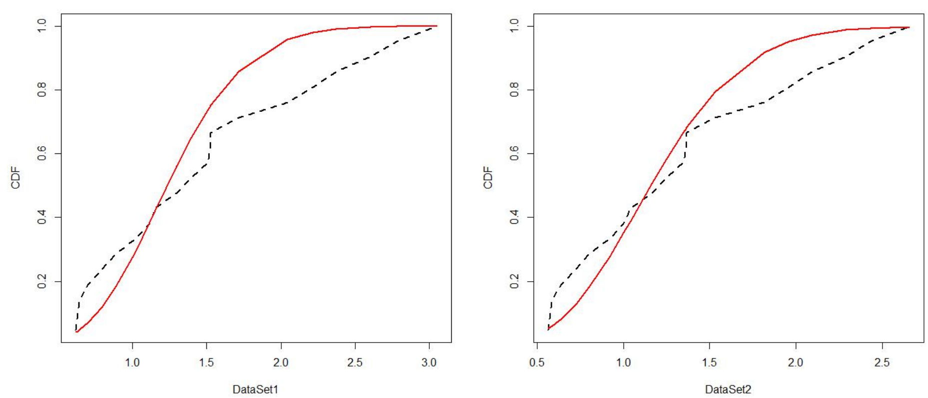



**Figure 1.** The empirical distribution function (dashed) and fitted distribution function for Data Sets 1 and 2.

**Example 4.2** (Simulated Data). We simulate 6 lower record values from $Burr(1.5)$ and 8 lower record values from $Burr(2.5)$. Therefore, $R_{Exact} = 0.375$. The data has been truncated after four decimal places and it has been presented below. The $\underset{\sim}{r}$ lower record values are

$$1.2483, 1.0473, 0.6649, 0.2187, 0.1846, 0.0730,$$

and the corresponding $\underset{\sim}{s}$ lower record values are

$$1.4244, 0.5154, 0.4293, 0.3531, 0.2727, 0.2266, 0.1173, 0.0942.$$

Based on the above data, we obtain the MLEs of $\theta_1$ and $\theta_2$ as, 1.1456 and 1.6916, respectively. Therefore, the MLE of $R$ becomes $\hat{R} = 0.4037$. The corresponding 95% confidence interval based on Equation (2.7) is equal to (0.1768,0.6617). Letting $\delta_1 = \delta_2 = 2$ and $\gamma_1 = \gamma_2 = 4$ in Equation (3.4), we obtain $\tilde{\theta}_1 = 0.7578$, $\tilde{\theta}_2 = 1.0310$ and $\tilde{R} = 0.4236$. Therefore, the approximate Bayes estimator of $R$ becomes $\tilde{R}_B = 0.4321$. The corresponding Bayesian 95% confidence interval based on Equation (3.5) is equal to (0.2199,0.6582). So, the approximate Bayes estimator of $R$ based on Equation (3.6) becomes $\tilde{R}_{JB} = 0.4077$ and the corresponding Bayesian 95% confidence interval based on Equation (3.7) is equal to (0.1768,0.6617). Finally, using Equation (3.10), we obtain $\tilde{\theta}_1^* = 1.0024$, $\tilde{\theta}_2^* = 1.5224$ and $\tilde{R}^* = 0.3970$. Therefore, the approximate empirical Bayes estimator of $R$ becomes $\tilde{R}_{EB} = 0.4070$. The corresponding Bayesian 95% confidence interval based on Equation (3.11) is equal to (0.2016,0.6329).

## 5. A simulation study

In this section, a simulation study is conducted to investigate the performance of different types of estimators presented in this paper and to compare them with some bootstrap intervals. It is important here to note that all inference procedures in this paper depend only on the smallest records, $r_n$ and $s_m$. In the simulation design we used all combinations of $n = 5, 10, 15$ and $m = 5, 10, 15$. We used $\theta_1 = 1$ and $R = 0.1, 0.25, 0.5$. The value of $\theta_2$ is determined by the choice of $\theta_1$ and $R$. In Bayesian simulation, we used

**Table 1.** Simulated biases and mean squared errors (in parentheses) of the estimators

| $n$ | $m$ | $R$ | $ML$ | $Bayes$ | $J.Bayes$ | $E.Bayes$ |
|---|---|---|---|---|---|---|
| 5 | 5 | 0/1 | 0.0149(0.0050) | 0.2724(0.0765) | 0.0320(0.0063) | 0.0247(0.0057) |
| 5 | 5 | 0/25 | 0.0211(0.0162) | 0.1633(0.0296) | 0.0390(0.0157) | 0.0313(0.0158) |
| 5 | 5 | 0/5 | 0.0027(0.0234) | 0.0013(0.0050) | 0.0025(0.0192) | 0.0026(0.0210) |
| 5 | 10 | 0/1 | 0.0179(0.0040) | 0.1858(0.0363) | 0.0252(0.0044) | 0.0234(0.0043) |
| 5 | 10 | 0/25 | 0.0260(0.0129) | 0.0990(0.0125) | 0.0311(0.0120) | 0.0303(0.0124) |
| 5 | 10 | 0/5 | 0.0158(0.0191) | -0.0212(0.0056) | 0.0057(0.0163) | 0.0103(0.0173) |
| 5 | 15 | 0/1 | 0.0176(0.0039) | 0.1377(0.0205) | 0.0218(0.0040) | 0.0212(0.0040) |
| 5 | 15 | 0/25 | 0.0238(0.0115) | 0.0636(0.0065) | 0.0245(0.0105) | 0.0253(0.0109) |
| 5 | 15 | 0/5 | 0.0173(0.0166) | -0.0326(0.0057) | 0.0035(0.0143) | 0.0094(0.0151) |
| 10 | 5 | 0/1 | 0.0055(0.0025) | 0.2931(0.0880) | 0.0229(0.0035) | 0.0156(0.0030) |
| 10 | 5 | 0/25 | 0.0071(0.0104) | 0.1844(0.0369) | 0.0292(0.0108) | 0.0201(0.0105) |
| 10 | 5 | 0/5 | -0.0076(0.0183) | 0.0254(0.0056) | 0.0020(0.0158) | -0.0024(0.0166) |
| 10 | 10 | 0/1 | 0.0071(0.0019) | 0.2028(0.0429) | 0.0150(0.0022) | 0.0130(0.0021) |
| 10 | 10 | 0/25 | 0.0120(0.0069) | 0.1193(0.0169) | 0.0212(0.0070) | 0.0189(0.0069) |
| 10 | 10 | 0/5 | -0.0012(0.0122) | -0.0007(0.0049) | -0.0012(0.0111) | -0.0012(0.0114) |
| 10 | 15 | 0/1 | 0.0079(0.0018) | 0.1531(0.0251) | 0.0128(0.0019) | 0.0120(0.0019) |
| 10 | 15 | 0/25 | 0.0108(0.0061) | 0.0831(0.0095) | 0.0154(0.0060) | 0.0148(0.0060) |
| 10 | 15 | 0/5 | 0.0022(0.0097) | -0.0118(0.0045) | -0.0014(0.0090) | -0.0001(0.0091) |
| 15 | 5 | 0/1 | 0.0035(0.0022) | 0.3020(0.0932) | 0.0210(0.0031) | 0.0137(0.0027) |
| 15 | 5 | 0/25 | 0.0024(0.0092) | 0.1944(0.0405) | 0.0260(0.0098) | 0.0164(0.0094) |
| 15 | 5 | 0/5 | -0.0171(0.0165) | 0.0333(0.0057) | -0.0033(0.0143) | -0.0092(0.0150) |
| 15 | 10 | 0/1 | 0.0058(0.0015) | 0.2115(0.0464) | 0.0139(0.0018) | 0.0119(0.0017) |
| 15 | 10 | 0/25 | 0.0051(0.0061) | 0.1267(0.0188) | 0.0157(0.0062) | 0.0130(0.0062) |
| 15 | 10 | 0/5 | -0.0023(0.0101) | 0.0119(0.0047) | 0.0013(0.0093) | 0.0001(0.0095) |
| 15 | 15 | 0/1 | 0.0053(0.0012) | 0.1609(0.0273) | 0.0105(0.0014) | 0.0096(0.0014) |
| 15 | 15 | 0/25 | 0.0067(0.0046) | 0.0921(0.0109) | 0.0128(0.0047) | 0.0117(0.0047) |
| 15 | 15 | 0/5 | -0.0005(0.0081) | -0.0003(0.0042) | -0.0005(0.0076) | -0.0005(0.0076) |

$\delta_1 = \delta_2 = 3$ and $\gamma_1 = \gamma_2 = 5$ where it is needed. All the results are based on 2000 replications.

First, we compare the performance of point estimators of $R$ in terms of their biases and mean squared errors (MSEs). In order to, we compute the average biases and mean squared errors (MSEs) as

$$Bias = \frac{1}{2000}\sum_{i=1}^{2000}(\hat{R}_i - R), \quad MSE = \frac{1}{2000}\sum_{i=1}^{2000}(\hat{R}_i - R)^2$$

where $\hat{R}$ can be each of the maximum likelihood estimator and the approximate Bayes estimators based on Equations (3.4), (3.6) and (3.10). The results are reported in Table 1.

Next, a simulation study is conducted to investigate and compare the performance of the confidence intervals presented in this paper and some bootstrap intervals in terms of their coverage probability and expected length. There are several bootstrap based intervals discussed in the literature (Efron and Tibshirani [16]). Since all inferences in this paper depend only on the smallest records, therefore we shall use the parametric bootstrap based on the marginal distribution of $R_n$ as given in Equation (2.4). In follows we describe the bootstrapping procedure:

1) Calculate $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{R}$, the maximum likelihood estimators of $\theta_1$, $\theta_2$ and $R$ based on $r_n$ and $s_m$.

2) Generate $r_n^*$ from the distribution given in Equation (2.4) with $\theta_1$ replaced by $\hat{\theta}_1$ and generate $s_m^*$ similarly.

3) Calculate $\hat{\theta}_1^*$, $\hat{\theta}_2^*$ and $\hat{R}^*$ using the $r_n^*$ and $s_m^*$ obtained in step 2.

4) Repeat steps 2 and 3, B times to obtain $\hat{R}_1^*, ..., \hat{R}_B^*$.

Then we can calculate the following bootstrap intervals;

Normal Interval: The simplest $(1-\alpha)$ bootstrap interval is the Normal interval

$$(\hat{R} - z_{1-\alpha/2}\hat{se_{boot}}, \hat{R} + z_{1-\alpha/2}\hat{se_{boot}})$$

where $\hat{se_{boot}}$ is the bootstrap estimate of the standard error based on $\hat{R_1^*}, ..., \hat{R_B^*}$.

Basic Pivotal Interval: The $(1-\alpha)$ bootstrap basic pivotal confidence interval is

$$(2\hat{R} - \hat{r}^*_{(1-\alpha/2)B}, 2\hat{R} - \hat{r}^*_{(\alpha/2)B})$$

where $\hat{r}^*_\beta$ is the $\beta$ quantile of $\hat{R_1^*}, ..., \hat{R_B^*}$.

Percentile Interval: The $(1-\alpha)$ bootstrap percentile interval is defined by

$$(\hat{r}^*_{(1-\alpha/2)B}, \hat{r}^*_{(1-\alpha/2)B})$$

that is, just use the $\alpha/2$ and $1-\alpha/2$ quantiles of the bootstrap sample.

Interested readers may refer to DiCiccio and Efron [17] and the references contained therein to observe more details.

For each generated pair of samples we calculated the following intervals;

1) ML: The interval based on the MLE given in Equation (2.7).

2) Bayes: The interval based on the Bayes estimator given in Equation (3.5).

3) J.B: The interval based on the Bayes estimator given in Equation (3.7).

4) E.B: The interval based on the empirical Bayes estimator given in Equation (3.11).

5) Norm: The normal interval.

6) Basic: The basic pivotal interval.

7) Perc: The percentile interval.

The empirical coverage probability and expected lengths of intervals are obtained by using the 2000 replications. For bootstrap intervals we used 1000 bootstrap samples. The results of our simulations for confidence level $(1-\alpha) = 0.95$ and 0.90 are given in Tables 2 and 3 respectively.

## 6. Conclusion and discussion

Based on simulation results in Table 1, we observe that the biases and the mean squared errors (MSEs) of the estimators are very close, especially for larger sample sizes. It appears that the performance of the MLE and the approximate Bayes estimators based on Equations (3.6) and (3.10) is almost the same in terms of their biases and mean squared errors (MSEs) but the MLE has the better performance for small values of $R$. Furthermore, the approximate Bayes estimators based on Equations (3.4) has the weak performance specially for small values of $R$. Hence, between the point estimators presented in this paper, we recommend to use the MLE.

Based on simulation results in Tables 2 and 3, it appears that the length of the intervals is maximized when $R = 0.5$ and gets shorter and shorter as we move away to the extremes. Increasing the sample size on either variable also results in shorter intervals. The performance of the both basic pivotal interval and percentile interval is similar in terms of expected length but in terms of coverage rate percentile interval has the better performance. The percentile interval appears to be the best among bootstrap intervals. The interval based on the MLE and the interval based on the Bayes estimator given in Equation (3.7) appears to perform almost as well as the percentile interval. The interval based on the Bayes estimator given in Equation (3.5) has the low coverage rate and the long expected length for small values of $R$ since it is dependent on $\gamma_1$ and $\gamma_2$ values. Furthermore, the interval based on the empirical Bayes estimator has the shortest expected length between the other intervals but it has the low coverage rate. It appears that the intervals based on the MLE, the Bayes estimator given in Equation (3.7) and percentile interval simultaneously has the short expected length and very good coverage rate in comparison with the other intervals. Hence, we recommend to use this confidence intervals in all.

**Acknowledgements**

The authors express their sincere thanks to the editor and the referees for valuable suggestions and helpful comments which led to an improvement in the presentation of the paper.

## References

[1] Kotz, S., Lumelskii, Y. and Pensky, M. *The Stress-strength Model and its Generalizations: Theory and Applications*, World Scientific, 2003.

[2] Chandler, K. N. *The distribution and frequency of record values*, Journal of the Royal Statistical Society, **14**, 220–228, 1952.

[3] Arnold, B. C., Balakrishnan, N., and Nagaraja, H.N. *Records*, Wiley, 1998.

[4] Burr, I. W. *Cumulative frequency functions*, Annals of Mathematical Statistics, **13**, 215-232, 1942.

[5] Sartawi, H. A. and Abu-Salih, M. S. *Bayesian prediction bounds for the Burr type X model*, Communication in Statistics: Theory and Methods, **20**, 2307–2330, 1991.

[6] Jaheen, Z. F. *Empirical Bayes estimation of the reliability and failure rate functions of the Burr type X failure model*, Journal of Applied Statistical Science, **3**, 281–288, 1996.

[7] Raqab, M.Z. *Order Statistics from the Burr Type X Model*, Computers and Mathematics with Applications, **36**, 111–120, 1998.

[8] Ahmad, K. E., Fakhry, M. E. and Jaheen, Z. F. *Empirical Bayes estimation of* $P(Y < X)$ *and characterizations of Burr type X model*, Journal of Statistical Planning and Inference, **64**, 297–308, 1997.

[9] Surles, J. G. and Padgett, W. J. *Inference for* $P(Y < X)$ *in the Burr Type X model*, Journal of Applied Statistical Science, **7**, 225–238, 1998.

[10] Kim, C. and Chung, Y. *Bayesian estimation of* $P(Y < X)$ *from Burr-type X model containing spurious observations*, Statistical Papers, **47**, 643–651, 2006.

[11] Baklizi, A. *Likelihood and Bayesian estimation of* $\Pr(X < Y)$ *using lower record values from the generalized exponential distribution*, Computational Statistics and Data Analysis, **52**, 3468–3473, 2008.

[12] Ahsanullah, M. *Record Values*, University Press of America Inc., Lanham, Maryland, USA : Theory and Applications, 2004.

[13] Lindley, D. V. *Approximate Bayesian methods*, Trabajos de Estadistica, **21**, 223–237, 1980.

[14] Chen, Ming-Hui. and Shao, Qi-Man. *Monte Carlo estimation of Bayesian credible and HPD intervals*, Journal of Computational and Graphical Statistics, **8**, 69–92, 1999.

[15] Bennett, H. S. and Filliben, J. J. *A systematic approach for multidimensional, closed-form analytic modeling: minority electron mobilities in* $Ga_{1-x}Al_xAs$ *heterostructures*, Journal of Research of the National Institute of Standards and Technology, **105**, 441–452, 2000.

[16] Efron, B. and Tibshirani, R. J. *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1993.

[17] DiCiccio, T. J. and Efron, B. *Bootstrap confidence intervals*, Statistical Science, **11**, 189–228, 1996.

**Table 2.** Expected lengths and coverage rates (in parentheses) of the confidence intervals with (1-$\alpha$)=0.95

| $n$ | $m$ | $R$ | $ML$ | $Bayes$ | $J.B$ | $E.B$ | $Norm$ | $Basic$ | $Perc$ |
|---|---|---|---|---|---|---|---|---|---|
| 5 | 5 | 0/1 | 0.275(0.947) | 0.437(0.391) | 0.276(0.951) | 0.209(0.878) | 0.282(0.937) | 0.274(0.813) | 0.274(0.947) |
| 5 | 5 | 0/25 | 0.452(0.953) | 0.449(0.901) | 0.448(0.948) | 0.353(0.872) | 0.463(0.905) | 0.451(0.802) | 0.451(0.950) |
| 5 | 5 | 0/5 | 0.538(0.952) | 0.460(0.996) | 0.538(0.952) | 0.433(0.878) | 0.551(0.891) | 0.538(0.792) | 0.538(0.949) |
| 5 | 10 | 0/1 | 0.216(0.953) | 0.344(0.408) | 0.218(0.947) | 0.179(0.882) | 0.253(0.956) | 0.246(0.835) | 0.246(0.949) |
| 5 | 10 | 0/25 | 0.388(0.952) | 0.381(0.985) | 0.388(0.945) | 0.323(0.883) | 0.423(0.933) | 0.415(0.844) | 0.415(0.946) |
| 5 | 10 | 0/5 | 0.481(0.947) | 0.415(0.995) | 0.481(0.949) | 0.399(0.890) | 0.488(0.901) | 0.479(0.828) | 0.479(0.937) |
| 5 | 15 | 0/1 | 0.204(0.939) | 0.291(0.569) | 0.202(0.957) | 0.168(0.906) | 0.249(0.963) | 0.242(0.869) | 0.242(0.927) |
| 5 | 15 | 0/25 | 0.369(0.954) | 0.344(0.966) | 0.368(0.956) | 0.308(0.906) | 0.410(0.941) | 0.403(0.859) | 0.403(0.926) |
| 5 | 15 | 0/5 | 0.461(0.948) | 0.394(0.985) | 0.460(0.952) | 0.382(0.902) | 0.465(0.906) | 0.457(0.847) | 0.457(0.931) |
| 10 | 5 | 0/1 | 0.236(0.950) | 0.405(0.403) | 0.238(0.952) | 0.182(0.900) | 0.210(0.917) | 0.206(0.827) | 0.206(0.946) |
| 10 | 5 | 0/25 | 0.405(0.955) | 0.414(0.726) | 0.403(0.949) | 0.323(0.887) | 0.381(0.905) | 0.374(0.833) | 0.374(0.935) |
| 10 | 5 | 0/5 | 0.482(0.956) | 0.416(0.994) | 0.484(0.960) | 0.402(0.909) | 0.489(0.903) | 0.480(0.834) | 0.480(0.941) |
| 10 | 10 | 0/1 | 0.178(0.942) | 0.316(0.508) | 0.177(0.953) | 0.152(0.914) | 0.180(0.937) | 0.177(0.868) | 0.177(0.940) |
| 10 | 10 | 0/25 | 0.323(0.954) | 0.344(0.864) | 0.324(0.954) | 0.284(0.912) | 0.326(0.928) | 0.322(0.865) | 0.322(0.955) |
| 10 | 10 | 0/5 | 0.405(0.942) | 0.367(0.994) | 0.406(0.950) | 0.358(0.917) | 0.409(0.909) | 0.404(0.865) | 0.404(0.938) |
| 10 | 15 | 0/1 | 0.156(0.947) | 0.263(0.560) | 0.157(0.949) | 0.139(0.927) | 0.166(0.951) | 0.163(0.871) | 0.163(0.944) |
| 10 | 15 | 0/25 | 0.297(0.946) | 0.307(0.938) | 0.296(0.955) | 0.265(0.924) | 0.308(0.932) | 0.305(0.885) | 0.305(0.940) |
| 10 | 15 | 0/5 | 0.375(0.950) | 0.342(0.990) | 0.375(0.953) | 0.336(0.925) | 0.378(0.919) | 0.374(0.881) | 0.374(0.948) |
| 15 | 5 | 0/1 | 0.223(0.946) | 0.389(0.483) | 0.224(0.958) | 0.171(0.903) | 0.189(0.913) | 0.185(0.841) | 0.185(0.942) |
| 15 | 5 | 0/25 | 0.389(0.957) | 0.396(0.750) | 0.385(0.954) | 0.308(0.891) | 0.356(0.893) | 0.349(0.844) | 0.349(0.928) |
| 15 | 5 | 0/5 | 0.460(0.952) | 0.394(0.992) | 0.460(0.955) | 0.382(0.902) | 0.466(0.915) | 0.458(0.855) | 0.458(0.939) |
| 15 | 10 | 0/1 | 0.160(0.950) | 0.299(0.510) | 0.160(0.951) | 0.140(0.927) | 0.154(0.926) | 0.152(0.861) | 0.152(0.944) |
| 15 | 10 | 0/25 | 0.298(0.952) | 0.326(0.781) | 0.301(0.946) | 0.267(0.913) | 0.292(0.920) | 0.289(0.877) | 0.289(0.944) |
| 15 | 10 | 0/5 | 0.375(0.950) | 0.341(0.985) | 0.374(0.942) | 0.335(0.906) | 0.379(0.920) | 0.375(0.884) | 0.375(0.951) |
| 15 | 15 | 0/1 | 0.140(0.946) | 0.246(0.518) | 0.139(0.952) | 0.126(0.920) | 0.142(0.943) | 0.140(0.885) | 0.140(0.946) |
| 15 | 15 | 0/25 | 0.268(0.945) | 0.286(0.854) | 0.267(0.950) | 0.243(0.924) | 0.270(0.937) | 0.267(0.886) | 0.267(0.947) |
| 15 | 15 | 0/5 | 0.339(0.952) | 0.315(0.982) | 0.339(0.951) | 0.310(0.928) | 0.341(0.919) | 0.339(0.886) | 0.339(0.950) |

**Table 3.** Expected lengths and coverage rates (in parentheses) of the confidence intervals with (1-$\alpha$)=0.90

| $n$ | $m$ | $R$ | $ML$ | $Bayes$ | $J.B$ | $E.B$ | $Norm$ | $Basic$ | $Perc$ |
|---|---|---|---|---|---|---|---|---|---|
| 5 | 5 | 0/1 | 0.225(0.888) | 0.371(0.340) | 0.226(0.905) | 0.172(0.813) | 0.236(0.899) | 0.224(0.791) | 0.224(0.889) |
| 5 | 5 | 0/25 | 0.382(0.892) | 0.382(0.716) | 0.378(0.895) | 0.297(0.802) | 0.388(0.857) | 0.382(0.769) | 0.382(0.889) |
| 5 | 5 | 0/5 | 0.461(0.898) | 0.392(0.995) | 0.461(0.910) | 0.368(0.805) | 0.462(0.831) | 0.461(0.753) | 0.461(0.895) |
| 5 | 10 | 0/1 | 0.180(0.904) | 0.291(0.430) | 0.181(0.892) | 0.149(0.814) | 0.212(0.932) | 0.200(0.811) | 0.200(0.896) |
| 5 | 10 | 0/25 | 0.329(0.909) | 0.323(0.924) | 0.329(0.898) | 0.273(0.814) | 0.355(0.897) | 0.348(0.816) | 0.348(0.900) |
| 5 | 10 | 0/5 | 0.411(0.902) | 0.353(0.980) | 0.410(0.899) | 0.339(0.818) | 0.410(0.846) | 0.409(0.781) | 0.409(0.890) |
| 5 | 15 | 0/1 | 0.170(0.891) | 0.245(0.447) | 0.168(0.915) | 0.140(0.833) | 0.209(0.930) | 0.196(0.846) | 0.196(0.881) |
| 5 | 15 | 0/25 | 0.314(0.894) | 0.291(0.974) | 0.312(0.913) | 0.261(0.836) | 0.344(0.888) | 0.338(0.820) | 0.338(0.881) |
| 5 | 15 | 0/5 | 0.392(0.905) | 0.335(0.964) | 0.392(0.911) | 0.324(0.828) | 0.390(0.852) | 0.390(0.802) | 0.390(0.886) |
| 10 | 5 | 0/1 | 0.191(0.911) | 0.343(0.363) | 0.193(0.908) | 0.149(0.831) | 0.176(0.879) | 0.170(0.798) | 0.170(0.904) |
| 10 | 5 | 0/25 | 0.339(0.897) | 0.351(0.495) | 0.337(0.899) | 0.271(0.816) | 0.320(0.856) | 0.317(0.798) | 0.317(0.892) |
| 10 | 5 | 0/5 | 0.411(0.905) | 0.354(0.988) | 0.413(0.923) | 0.341(0.845) | 0.411(0.856) | 0.410(0.792) | 0.410(0.893) |
| 10 | 10 | 0/1 | 0.147(0.890) | 0.266(0.421) | 0.146(0.903) | 0.126(0.859) | 0.151(0.902) | 0.146(0.836) | 0.146(0.892) |
| 10 | 10 | 0/25 | 0.271(0.910) | 0.291(0.715) | 0.272(0.901) | 0.238(0.848) | 0.274(0.880) | 0.271(0.827) | 0.271(0.908) |
| 10 | 10 | 0/5 | 0.343(0.896) | 0.311(0.975) | 0.344(0.907) | 0.303(0.856) | 0.343(0.860) | 0.343(0.818) | 0.343(0.898) |
| 10 | 15 | 0/1 | 0.130(0.898) | 0.221(0.419) | 0.130(0.906) | 0.116(0.864) | 0.139(0.919) | 0.135(0.844) | 0.135(0.899) |
| 10 | 15 | 0/25 | 0.250(0.893) | 0.259(0.835) | 0.250(0.904) | 0.223(0.864) | 0.259(0.886) | 0.256(0.840) | 0.256(0.886) |
| 10 | 15 | 0/5 | 0.318(0.903) | 0.289(0.970) | 0.318(0.906) | 0.284(0.859) | 0.317(0.871) | 0.317(0.840) | 0.317(0.895) |
| 15 | 5 | 0/1 | 0.180(0.905) | 0.329(0.491) | 0.181(0.909) | 0.140(0.825) | 0.158(0.881) | 0.154(0.820) | 0.154(0.900) |
| 15 | 5 | 0/25 | 0.325(0.900) | 0.336(0.543) | 0.321(0.900) | 0.257(0.825) | 0.299(0.853) | 0.297(0.810) | 0.297(0.884) |
| 15 | 5 | 0/5 | 0.392(0.909) | 0.335(0.967) | 0.392(0.911) | 0.324(0.828) | 0.391(0.862) | 0.391(0.809) | 0.391(0.895) |
| 15 | 10 | 0/1 | 0.131(0.903) | 0.251(0.501) | 0.132(0.909) | 0.115(0.863) | 0.129(0.884) | 0.126(0.834) | 0.126(0.897) |
| 15 | 10 | 0/25 | 0.250(0.902) | 0.275(0.546) | 0.252(0.896) | 0.224(0.852) | 0.245(0.883) | 0.243(0.841) | 0.243(0.901) |
| 15 | 10 | 0/5 | 0.318(0.901) | 0.289(0.959) | 0.317(0.883) | 0.283(0.844) | 0.318(0.871) | 0.318(0.841) | 0.318(0.897) |
| 15 | 15 | 0/1 | 0.116(0.893) | 0.207(0.553) | 0.115(0.893) | 0.104(0.854) | 0.119(0.900) | 0.116(0.851) | 0.116(0.890) |
| 15 | 15 | 0/25 | 0.225(0.898) | 0.241(0.726) | 0.224(0.894) | 0.204(0.864) | 0.226(0.877) | 0.225(0.840) | 0.225(0.898) |
| 15 | 15 | 0/5 | 0.287(0.892) | 0.266(0.957) | 0.287(0.900) | 0.262(0.865) | 0.286(0.865) | 0.287(0.836) | 0.287(0.892) |

Hacettepe Journal of Mathematics and Statistics
Volume 45 (1) (2016), 279 – 296

# Bayesian analysis for semiparametric mixed-effects double regression models

Dengke Xu*†‡, Zhongzhan Zhang§ and Liucang Wu¶

**Abstract**

In recent years, based on jointly modeling the mean and variance, double regression models are widely used in practice. In order to assess the effects of continuous covariates or of time scales in a flexible way, a class of semiparametric mixed-effects double regression models(SMMEDRMs) is considered, in which we model the variance of the mixed effects directly as a function of the explanatory variables. In this paper, we propose a fully Bayesian inference for SMMEDRMs on the basis of B-spline estimates of nonparametric components. A computational efficient MCMC method which combines the Gibbs sampler and Metropolis-Hastings algorithm is implemented to simultaneously obtain the Bayesian estimates of unknown parameters and the smoothing function, as well as their standard deviation estimates. Finally, some simulation studies and a real example are used to illustrate the proposed methodology.

**Keywords:** Bayesian analysis, Semiparametric mixed-effects double regression models, Gibbs sampler, Metropolis-Hastings algorithm, B-spline.

*2000 AMS Classification:* 62F15, 62G05.

*Received* : 20.05.2014 *Accepted* : 13.12.2014 *Doi* : 10.15672/HJMS.2014458271

*Department of Statistics, Yunnan University, Kunming, People's Republic of China
†Department of Statistics, Zhejiang Agriculture and Forestry University, Hangzhou, People's Republic of China, Email: xudengke1983@163.com
‡Corresponding Author.
§College of Applied Sciences, Beijing University of Technology, Beijing, People's Republic of China, Email: zzhang@bjut.edu.cn
¶Faculty of Science, Kunming University of Science and Technology, Kunming, People's Republic of China, Email: wuliucang@163.com

## 1. Introduction

Many different approaches have been suggested to the problem of flexibly modeling of the mean. In statistical literature, compared with that of the mean, modeling of the variance has often been neglected. In many applications, particularly in the econometric area and industrial quality improvement experiments, modeling the variance will be of direct interest in its own right, to identify the source of variability in the observations, such as Taguchi-type experiments for robust design. On the other hand, modeling the variance itself may be of scientific interest. Thus, modeling of the variance can be as important as that of the mean. Furthermore, it is well known that efficient estimation of mean parameters in regression depends on correct modeling of the variance. The loss of efficiency may be substantial using constant variance models when the variance is varying. In addition, modeling of the variance is also necessary to obtain correct standard errors and confidence intervals, as well as for many other applications such as prediction and so on. Recently, the joint mean and variance models have been receiving a lot of attention. For example, Aitkin [1] provided maximum likelihood (ML) estimation for a joint mean and variance model and applied it to the commonly cited Minitab tree data. Xie et al. [22] investigated the score tests for homogeneity of a scalar parameter and a skewness parameter in skew-normal nonlinear regression models, which are included in the variance. Wu and Li [21] proposed a unified variable selection procedure which can simultaneously select significant variables in mean and dispersion models of the inverse Gaussian distribution. Zhao et al. [25] considered the issue of variable selection for beta regression models with varying dispersion, in which both the mean and the dispersion depend upon predictor variables. Wu [20] investigated the simultaneous variable selection in joint location and scale models of the skew-t-normal distribution when the dataset under consideration involves heavy tail and asymmetric outcomes. The similar works can be also seen from [12, 13, 24] and so on. On the other hand, semiparametric mixed models are useful extensions to linear mixed models and provide a flexible framework for analyzing longitudinal data. Many authors have studied semiparametric mixed models for longitudinal data (e.g., Ni et al. [15]). But, there is little work about the case in which the variance is additionally modelled. Therefore, in this paper we are interested in jointly modelling mean and variance of semiparametric mixed models.

Bayesian inference for the semiparametric mixed-effects models and the joint mean and variance models have also receiving a lot of attention in recent years. For example, Cepeda and Gamerman [2] summarized the Bayesian approach for modeling variance heterogeneity in normal regression analysis. Chen [3] proposed a fully Bayesian inference for semiparametric mixed-effects models of zero-inflated count data based on a data augmentation scheme that reflects both random effects of covariates and mixture of zero-inflated distribution. Chen and Tang [4] developed a Bayesian procedure for analyzing semiparametric reproductive dispersion mixed-effects models on the basis of P-spline estimates of nonparametric components. Lin and Wang [14] presented a fully Bayesian approach to multivariate regression models whose mean vector and scale covariance matrix are modelled jointly for analyzing longitudinal data. Tang and Duan [18] proposed a semiparametric Bayesian approach to generalized partial linear mixed models for longitudinal data. Xu and Zhang [23] proposed a fully Bayesian inference for semiparametric joint mean and variance models on the basis of B-spline approximations of nonparametric components. However, to the best of our knowledge, there is little work done for Bayesian analysis of semiparametric mixed-effects double regression models with longitudinal data, in which we model the variance of the mixed effects directly as a function of the explanatory variables.

On the other hand, various methods are available for fitting the semiparametric models, such as, the kernel smoothing method and the spline method. See for example,[5, 19, 26] and so on. Recently, the B-spline method is widely used to fit semiparametric models because of its advantages. Firstly, it does not need to estimate the nonparametric component of model point by point, that is, instead of concerning the local quality, the global quality is taken into consideration, which lead to the reduction of the computational complexity. Secondly, there are no boundary effects so that the splines can fit polynomial data exactly. Thirdly, the B-spline base functions have bounded supports and are numerically stable (Schumaker [17]).

Therefore, in this paper we extend the Bayesian methodology proposed in [2, 23] to fit semiparametric mixed-effects double regression models. Hence, a semiparametric Bayesian approach to SMMEDRMs is developed based on the B-spline approximation of nonparametric function and the hybrid algorithm combining the Gibbs sampler and Metropolis-Hastings algorithm in this article.

The outline of the paper is as follows. In Section 2 we first describe semiparametric mixed-effects double regression models. A Bayesian procedure based on a data augmentation scheme, Gibbs sampler and the Metropolis-Hastings algorithm for obtaining estimates is developed in Section 3. The full conditional distributions for implementing the sampling-based methods are also derived. To illustrate the proposed methodology, results obtained from some simulation studies are presented in Section 4. We further illustrate the proposed methodology through an analysis of the CD4 data in Section 5. The article is concluded with a brief discussion in Section 6.

## 2. Semiparametric Mixed-Effects Double Regression Models

Suppose that there are $n$ independent subjects and the $ith$ subject has $m_i$ repeated measurements. Specifically, denote the response vector $Y_i = (Y_{i1}, \cdots, Y_{im_i})^T$ for the $ith$ subject, $i = 1, \cdots, n$, which are observed at time $t_i = (t_{i1}, \cdots, t_{im_i})^T$. We assume that the response is normally distributed as $Y_{ij}|(X_{ij}, v_i, t_{ij}) \sim N(\mu_{ij}, \sigma^2)$. Here, the superscript T denotes the transposed of a vector (or matrix).

In this paper we consider

$$\left\{ \begin{array}{l} \mu_{ij} = X_{ij}^T\beta + v_i + g(t_{ij}), \\ i = 1, 2, \cdots, n, \\ j = 1, 2, \cdots, m_i, \end{array} \right. \tag{2.1}$$

where $t_{ij}$ is a univariate observed covariate, $g(\cdot)$ is an unknown smooth function in the mean model, $v_i$ is a random effect with $v_i \sim N(0, \sigma_i^2)$. Furthermore, if we have variance heterogeneity of the random effect, it is convenient to assume an explicit variance modeling related to some explanatory variables, that is:

$$\sigma_i^2 = h(Z_i, \gamma), \tag{2.2}$$

where $Z_i = (Z_{i1}, \cdots, Z_{iq})^T$ is the observation of explanatory variables associated with the variance of $v_i$ and $\gamma = (\gamma_1, \cdots, \gamma_q)^T$ is a $q \times 1$ vector of regression coefficients in the variance model. Furthermore, we let $Z = (Z_1, Z_2, \cdots, Z_n)^T$. In addition, $h(\cdot, \cdot) > 0$ is a known function. Here two specific forms of $h(\cdot, \cdot)$ are usually taken to model varying variance: (i) log-linear model: $h(Z_i, \gamma) = \exp(\sum_{j=1}^q Z_{ij}\gamma_j)$; (ii) power product model: $h(Z_i, \gamma) = \prod_{j=1}^q Z_{ij}^{\gamma_j} = \exp(\sum_{j=1}^q \gamma_j \log Z_{ij})$. Of course, (ii) requires that the $Z_{ij}$ is strictly positive, while no such restriction is needed for (i). In practice, one may make a choice of the variance weight $h(\cdot, \cdot)$, even a choice of the explanatory variables $Z_i$, according to the domain knowledge or modeling convenience. Therefore, in this article we consider the following semiparametric mixed-effects double regression models (SMMEDRMs):

$$
\begin{cases} Y_{ij} = X_{ij}^T\beta + v_i + g(t_{ij}) + \varepsilon_{ij}, \\ \varepsilon_{ij} \sim N(0, \sigma^2), \\ v_i|Z_i \sim N(0, \sigma_i^2), \\ \sigma_i^2 = h(Z_i, \gamma), \\ i = 1, 2, \cdots, n, \\ j = 1, 2, \cdots, m_i, \end{cases} \tag{2.3}
$$

based on the independent observations $(Y_{ij}, X_{ij}, Z_i, t_{ij}), i = 1, 2, \cdots, n, j = 1, 2, \cdots, m_i$.

## 3. Bayesian Analysis of SMMEDRMs

**3.1. B-splines for the Nonparametric Function.** Without loss of generality, we assume that the covariate $t_{ij}$ is valued on $[0, 1]$. Let $T = (t_1^T, t_2^T, \cdots, t_n^T)^T$. From the model (2.3), we obtain the likelihood function

$$
\begin{aligned} L(\beta, \gamma, \phi^2, v|Y, X, Z, T) &= \prod_{i=1}^{n} \left\{ f(v_i|Z_i, \gamma) \prod_{j=1}^{m_i} f(Y_{ij}|X_{ij}, v_i, t_{ij}, \beta) \right\} \\ &\propto \left\{ \prod_{i=1}^{n} \sigma_i \right\}^{-1} (\phi^2)^{\frac{N}{2}} \exp\left\{ -\frac{\phi^2}{2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} (Y_{ij} - X_{ij}^T\beta - v_i - g(t_{ij}))^2 - \sum_{i=1}^{n} \frac{v_i^2}{2\sigma_i^2} \right\}, \end{aligned} \tag{3.1}
$$

where $\phi^2 = 1/\sigma^2$, $N = \sum_{i=1}^{n} m_i$, $v = (v_1, \cdots, v_n)^T$, $Y = (Y_1^T, \cdots, Y_n^T)^T$, $X = (X_1^T, \cdots, X_n^T)^T$, $X_i = (X_{i1}, \cdots, X_{im_i})^T$.

Since $g(\cdot)$ is nonparametric, (3.1) is not yet ready for optimization. So, we first use B-splines to approximate the nonparametric function $g(\cdot)$. Any computational algorithm developed for generalized linear models (GLM) can be used for fitting a semiparametric extension of GLM, since one can treat a nonparametric function as a linear function with the basis functions as covariates. For simplicity, let $0 = s_0 < s_1 < \cdots < s_{k_n} < s_{k_n+1} = 1$ be a partition of the interval [0, 1]. Using $\{s_i\}$ as the internal knots, we have $K = k_n + M$ normalized B-spline basis functions of order $M$ that form a basis for the linear spline space. Selection of knots is generally an important aspect of spline smoothing. In this paper, similar to He et al. [10], the number of internal knots is taken to be the integer part of $N^{1/5}$. Thus $g(t)$ is approximated by $\pi^T(t)\alpha$ , where $\pi(t) = (\pi_1(t), ..., \pi_K(t))^T$ is the vector of basis functions and $\alpha \in R^K$. With this notation, the mean model in (2.3) can be linearized as

$$
\mu_{ij} = x_{ij}^T\beta + v_i + \pi^T(t_{ij})\alpha. \tag{3.2}
$$

Hence, based on (3.2), the likelihood function (3.1) can be rewritten as follows:

$$
\begin{aligned} L(\beta, \alpha, \gamma, \phi^2, v|Y, X, Z, T) &= \prod_{i=1}^{n} \left\{ f(v_i|Z_i, \gamma) \prod_{j=1}^{m_i} f(Y_{ij}|X_{ij}, v_i, t_{ij}, \beta) \right\} \\ \propto \left\{ \prod_{i=1}^{n} \sigma_i \right\}^{-1} (\phi^2)^{\frac{N}{2}} &\exp\left\{ -\frac{\phi^2}{2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} (Y_{ij} - X_{ij}^T\beta - v_i - \pi^T(t_{ij})\alpha)^2 - \sum_{i=1}^{n} \frac{v_i^2}{2\sigma_i^2} \right\}. \end{aligned} \tag{3.3}
$$

**3.2. Prior Density of Parameters.** To implement a Bayesian approach to estimate the parameters of the models (2.3), we need to specify a prior distribution for the parameters involved. For simplicity, we suppose that $\beta, \alpha$ and $\gamma$ are independent and normally distributed in prior as $\beta|\phi^2 \sim N(\beta_0, \phi^{-2}b_\beta)$, $\alpha \sim N(\alpha_0, \tau^2 I_K)$ and $\gamma \sim N(\gamma_0, B_\gamma)$, where the hyperparameters $\beta_0, \alpha_0, \gamma_0$, $b_\beta$ and $B_\gamma$ are assumed known, and $\tau^2$ is assumed to be distributed as Gamma$(a_\tau, b_\tau)$ with density function

$$p(\tau^2|a_\tau, b_\tau) \propto (\tau^2)^{a_\tau - 1} \exp\left(-b_\tau \tau^2\right),$$

where $a_\tau$ and $b_\tau$ are known positive constants. In addition, we also suppose that $\phi^2$ is distributed in prior as Gamma$(a_{\phi^2}, b_{\phi^2})$, where $a_{\phi^2}$ and $b_{\phi^2}$ are known positive constants.

**3.3. Gibbs Sampling and Conditional Distribution.** Let $\theta = (\beta, \alpha, \gamma, \phi^2)$, $B_i = (\pi(t_{i1}), \cdots, \pi(t_{im_i}))^T$ and $B = (B_1^T, \cdots, B_n^T)^T$. Based on (3.3), we can sample from joint posterior distribution $p(\theta, v|Y, X, Z, T)$ by Gibbs sampling along the following process.

**Step 1.** Setting initial values of parameters as $\theta^{(0)} = (\beta^{(0)}, \alpha^{(0)}, \gamma^{(0)}, \phi^{2(0)})$.

**Step 2.** Based on $\theta^{(l)} = (\beta^{(l)}, \alpha^{(l)}, \gamma^{(l)}, \phi^{2(l)})$, compute $\Sigma^{(l)} = \text{diag}\{h(Z_1, \gamma^{(l)}), \cdots, h(Z_n, \gamma^{(l)})\}$, $\tilde{v}_i^{(l)} = v_i^{(l)} \bigotimes 1_{m_i}$ and $\tilde{v}^{(l)} = ((\tilde{v}_1^{(l)})^T, \cdots, (\tilde{v}_n^{(l)})^T)^T$.

**Step 3.** Based on $\theta^{(l)} = (\beta^{(l)}, \alpha^{(l)}, \gamma^{(l)}, \phi^{2(l)})$, sample $\theta^{(l+1)} = (\beta^{(l+1)}, \alpha^{(l+1)}, \gamma^{(l+1)}, \phi^{2(l+1)})$, $v^{(l+1)}$ and $\tau^{2(l+1)}$ as follows:

• Sampling $\phi^{2(l+1)}$:

(3.4)

$$p(\phi^2|Y, X, v, \beta, \gamma, \alpha) \approx (\phi^2)^{\frac{N+p}{2} + a_{\phi^2} - 1} \exp\Bigg\{ -\phi^2 \Bigg[\frac{1}{2}(Y - X\beta^{(l)} - \tilde{v}^{(l)} - B\alpha^{(l)})^T (Y - X\beta^{(l)} - \tilde{v}^{(l)} - B\alpha^{(l)}) + \tfrac{1}{2}(\beta^{(l)} - \beta_0)^T(\beta^{(l)} - \beta_0) + b_{\phi^2}\Bigg]\Bigg\}.$$

• Sampling $\tau^{2(l+1)}$:

$$p(\tau^2|\alpha) \propto (\tau^2)^{-\frac{K}{2} - a_\tau - 1} \exp\left\{-\frac{(\alpha^{(l)} - \alpha_0)^T(\alpha^{(l)} - \alpha_0) + 2b_\tau}{2\tau^2}\right\}. \tag{3.5}$$

• Sampling $\alpha^{(l+1)}$:

$$p(\alpha|Y, X, Z, T, \beta, \gamma, \tau^2, \phi^2) \propto \exp\left\{-\frac{1}{2}(\alpha - \alpha_0^*)^T b_\alpha^{*\,-1}(\alpha - \alpha_0^*)\right\}, \tag{3.6}$$

where $\alpha_0^* = b_\alpha^*(\tau^{2(l+1)^{-1}} I_K \alpha_0 + \phi^{2(l+1)} B^T(Y - X\beta^{(l)} - \tilde{v}^{(l)}))$ and $b_\alpha^* = (\tau^{2(l+1)^{-1}} I_K + \phi^{2(l+1)} B^T B)^{-1}$, $I_K$ is the identity matrix.

• Sampling $\beta^{(l+1)}$:

$$p(\beta|Y, X, Z, T, \alpha, \gamma, \phi^2) \propto \exp\left\{-\frac{1}{2}(\beta - \beta_0^*)^T b_\beta^{*\,-1}(\beta - \beta_0^*)\right\}, \tag{3.7}$$

where $\beta_0^* = b_\beta^*((\phi^{-2(l+1)} b_\beta)^{-1}\beta_0 + \phi^{2(l+1)} X^T(Y - \tilde{v}^{(l)} - B\alpha^{(l+1)}))$ and $b_\beta^* = ((\phi^{-2(l+1)} b_\beta)^{-1} + \phi^{2(l+1)} X^T X)^{-1}$.

• Sampling $v^{(l+1)}$:

$$\begin{aligned} p(v|Y, X, T, Z, \beta, \gamma, \phi^2) \quad \propto \quad & \exp\Bigg\{ -\frac{\phi^{2(l+1)}}{2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} (Y_{ij} - X_{ij}^T \beta^{(l+1)} \\ & -\pi(t_{ij})^T \alpha^{(l+1)} - v_i)^2 - \sum_{i=1}^{n} \frac{v_i^2}{2\sigma_i^{2(l)}} \Bigg\}, \end{aligned} \tag{3.8}$$

where $\sigma_i^{2(l)} = h(Z_i, \gamma^{(l)})$.

• Sampling $\gamma^{(l+1)}$:

(3.9)

$$p(\gamma|Y,X,Z,\beta,\phi^2) \propto |\Sigma_1|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}v^{(l+1)^T}\Sigma_1 v^{(l+1)} - \frac{1}{2}(\gamma-\gamma_0)^T B_\gamma^{-1}(\gamma-\gamma_0)\right\}.$$

Here, $\Sigma_1 = \text{diag}\{h(Z_1,\gamma),\cdots,h(Z_n,\gamma)\}$.

**Step 4.** Repeating Steps 2 and 3.

Then, we can generate sample series $(\beta^{(t)},\alpha^{(t)},\gamma^{(t)},\phi^{2(t)},\tau^{2(t)}), t = 1,2,\cdots$ by the above program. It is easily seen from (3.4), (3.5), (3.6) and (3.7) that conditional distributions $p(\tau^2|\alpha), p(\alpha|Y,X,Z,T,\beta,\gamma,\tau^2,\phi^2), p(\beta|Y,X,Z,T,\alpha,\gamma,\phi^2)$ and $p(\phi^2|Y,X,T,v,\beta,\gamma,\alpha)$ are some familiar distributions, such as the Gamma and normal distributions. Sampling observations from these standard distributions is straightforward and fast. But conditional distributions $p(v|Y,X,Z,T,\beta,\gamma,\phi^2)$ and $p(\gamma|Y,X,Z,\beta,\phi^2)$ are some unfamiliar and rather complicated, thus drawing observations from the distributions are rather difficult. Hence, the commonly used Metropolis-Hastings algorithm is employed to sample observations from them. To this end, we choose normal distribution $N(v^{(l)},\sigma_v^2\Omega_v^{-1})$ and $N(\gamma^{(l)},\sigma_\gamma^2\Omega_\gamma^{-1})$ as the proposal distribution [11, 16], where $\sigma_v^2$ and $\sigma_\gamma^2$ are chosen such that the average acceptance rate is about between 0.25 and 0.45 (Gelman et al. [8]), and take

$$\Omega_v = E\left(-\frac{\partial^2 \log p(v|Y,X,T,Z,\beta^{(l+1)},\gamma^{(l)},\phi^{2(l+1)})}{\partial v \partial v^T}\right),$$

$$\Omega_\gamma = E\left(-\frac{\partial^2 \log p(\gamma|Y,X,Z,\beta^{(l+1)},\phi^{2(l+1)})}{\partial \gamma \partial \gamma^T}\right).$$

The Metropolis-Hastings algorithm is implemented as follows: at the $(l+1)$th iteration with the current value $v^{(l)}$, $\gamma^{(l)}$, new candidates $v^*$ and $\gamma^*$ are generated from $N(v^{(l)},\sigma_v^2\Omega_v^{-1})$, $N(\gamma^{(l)},\sigma_\gamma^2\Omega_\gamma^{-1})$ and are accepted respectively with probability

$$\min\left\{1, \frac{p(v^*|Y,X,Z,\beta,\gamma,\phi^2)}{p(v^{(l)}|Y,X,Z,\beta,\gamma,\phi^2)}\right\}$$

and

$$\min\left\{1, \frac{p(\gamma^*|Y,X,Z,\beta,\phi^2)}{p(\gamma^{(l)}|Y,X,Z,\beta,\phi^2)}\right\}.$$

**3.4. Bayesian Inference.** Observations generated from the above proposed computational procedure are used to obtain Bayesian estimates of parameters $\beta,\alpha,\gamma$ and $\phi^2$ and their standard deviations.

Let $\{\theta^{(j)} = (\beta^{(j)},\alpha^{(j)},\gamma^{(j)},\phi^{2(j)}) : j = 1,2,\cdots,J\}$ be the observations of $(\beta,\alpha,\gamma,\phi^2)$ generated from the joint conditional distribution $p(\beta,\alpha,\gamma,\phi^2|Y,X,Z,T)$ via the proposed hybrid algorithm. The Bayesian estimates of $\beta,\alpha,\gamma$ and $\phi^2$ are given as:

$$\hat{\beta} = \frac{1}{J}\sum_{j=1}^J \beta^{(j)}, \quad \hat{\alpha} = \frac{1}{J}\sum_{j=1}^J \alpha^{(j)},$$

$$\hat{\gamma} = \frac{1}{J}\sum_{j=1}^J \gamma^{(j)}, \quad \hat{\phi}^2 = \frac{1}{J}\sum_{j=1}^J \phi^{2(j)}.$$

As is shown by Geyer [9], $\hat{\theta} = (\hat{\beta},\hat{\alpha},\hat{\gamma},\hat{\phi}^2)$ is a consistent estimate of the corresponding posterior mean vector as $J$ goes to infinity. Similarly, a consistent estimate of the posterior covariance matrix $\mathrm{Var}(\theta|Y,X,Z,T)$ can be obtained via the sample covariance matrix of the observations $\{\theta^{(j)} : j = 1, 2, \cdots, J\}$, that is

$$\widehat{\mathrm{Var}}(\theta|Y,X,Z,T) = (J-1)^{-1}\sum_{j=1}^{J}(\theta^{(j)}-\hat{\theta})(\theta^{(j)}-\hat{\theta})^T.$$

Thus, the posterior standard deviations for the components can be obtained from the diagonal elements of the matrix.

## 4. Simulation Studies

In this section, some simulation studies are used to illustrate various aspects of the proposed Bayesian method. In the following simulations, $\sigma^2 = 0.5$ and the structure of the mean model is $\mu_{ij} = X_{ij}^T\beta + v_i + 0.5\sin(2\pi t_{ij}), i = 1, 2, \cdots, n, j = 1, 2, \cdots, m$, where $m = 4$, $t_{ij}$ follows uniform distribution $U(0,1)$, $X_{ij}$ is a $3\times 1$ vector with elements independently sampled from normal distribution $N(0,1)$, and $\beta = (1, -0.8, 1)^T$. The structure of the variance model of the random effect $v_i$ will be taken to be different models in the following examples.

To investigate sensitivity of Bayesian estimates to prior inputs, we consider the following three types of hyperparameter values for unknown parameters $\beta, \alpha, \gamma, \tau^2, \phi^2$:

Type I: $\beta_0 = (1, -0.8, 1)^T, b_\beta = I_3, \gamma_0 = (1, -0.5)^T, B_\gamma = I_2,\ a_\tau = 1, b_\tau = 1, a_{\phi^2} = 1, b_{\phi^2} = 1$. This can be regarded as a situation with good prior information.

Type II: $\beta_0 = 1.5\times(1, -0.8, 1)^T, b_\beta = I_3, \gamma_0 = 1.5\times(1, -0.5)^T, B_\gamma = I_2,\ a_\tau = 1, b_\tau = 1, a_{\phi^2} = 1, b_{\phi^2} = 1$. This can be regarded as a situation with inaccurate prior information.

Type III: $\beta_0 = (0, 0, 0)^T, b_\beta = I_3, \gamma_0 = (0, 0)^T, B_\gamma = I_2,\ a_\tau = 1, b_\tau = 1, a_{\phi^2} = 1, b_{\phi^2} = 1$. These hyperparameter values represent a situation with noninformative prior information.

For the above various settings, the preceding proposed hybrid algorithm combining the Gibbs sampler and the Metropolis-Hastings algorithm is used to evaluate the Bayesian estimates of unknown parameters and the smoothing function. In the following simulations, we use the cubic B-splines. Different sample sizes are employed in the simulations to show the effect of sample sizes. For each setting, 100 replications are carried out. For each data set generated in a replication, the convergence of the MCMC sampler is checked by estimated potential scale reduction (EPSR) value [7], and we observe that in all runs, the EPSR values are less than 1.2 after 4000 iterations. Observations are collected after 4000 iterations with $J = 4000$ in producing the Bayesian estimates for each replication.

### 4.1. Example 1: Comparisons for different prior inputs and sample sizes.

In this example, we take the log-linear model as the structure of the variance model of the random effect $v_i$,

$$\log(\sigma_i^2) = Z_i^T\gamma, \quad i = 1, 2, \cdots, n$$

with $\gamma = (1, -0.5)^T$ and $Z_i$ is a $2\times 1$ vector with elements generated randomly from normal distribution $N(0,1)$, n is the sample size ranging from n=30, 50, 100, 150. The summary of the simulation results for parameters is presented in Tables 1 and 2. To investigate accuracy of estimate of function $g(t)$, we plot the true value of function $g(t)$ against its estimates for three types of prior inputs under different sample sizes in Figures 1-4.

**Table 1.** Bayesian estimates of parameters under different priors when $n = 30$ and $n = 50$ in Example 1

| | | $n = 30$ | | | $n = 50$ | | |
|---|---|---|---|---|---|---|---|
| Type | Parameters | BIAS | RMS | SD | BIAS | RMS | SD |
| I | $\beta_1$ | 0.0103 | 0.0769 | 0.0748 | 0.0028 | 0.0553 | 0.0577 |
| | $\beta_2$ | 0.0041 | 0.0693 | 0.0743 | 0.0028 | 0.0542 | 0.0577 |
| | $\beta_3$ | 0.0100 | 0.0731 | 0.0743 | 0.0055 | 0.0559 | 0.0573 |
| | $\gamma_1$ | 0.0711 | 0.3301 | 0.3532 | 0.0093 | 0.2727 | 0.2649 |
| | $\gamma_2$ | 0.0154 | 0.3218 | 0.3520 | 0.0141 | 0.2160 | 0.2576 |
| | $\sigma^2$ | 0.0058 | 0.0704 | 0.0769 | 0.0065 | 0.0553 | 0.0588 |
| II | $\beta_1$ | 0.0002 | 0.0799 | 0.0747 | 0.0037 | 0.0660 | 0.0579 |
| | $\beta_2$ | 0.0019 | 0.0752 | 0.0749 | 0.0029 | 0.0627 | 0.0585 |
| | $\beta_3$ | 0.0102 | 0.0717 | 0.0741 | 0.0038 | 0.0606 | 0.0579 |
| | $\gamma_1$ | 0.0744 | 0.3294 | 0.3448 | 0.0227 | 0.2620 | 0.2727 |
| | $\gamma_2$ | 0.0271 | 0.3318 | 0.3453 | 0.0526 | 0.2770 | 0.2641 |
| | $\sigma^2$ | 0.0028 | 0.0695 | 0.0764 | 0.0107 | 0.0603 | 0.0595 |
| III | $\beta_1$ | 0.0181 | 0.0708 | 0.0775 | 0.0009 | 0.0613 | 0.0578 |
| | $\beta_2$ | 0.0163 | 0.0778 | 0.0765 | 0.0029 | 0.0481 | 0.0581 |
| | $\beta_3$ | 0.0139 | 0.0786 | 0.0778 | 0.0053 | 0.0567 | 0.0580 |
| | $\gamma_1$ | 0.1190 | 0.3099 | 0.3403 | 0.0139 | 0.2387 | 0.2518 |
| | $\gamma_2$ | 0.0358 | 0.2739 | 0.3264 | 0.0132 | 0.2323 | 0.2467 |
| | $\sigma^2$ | 0.0409 | 0.0851 | 0.0818 | 0.0194 | 0.0595 | 0.0605 |

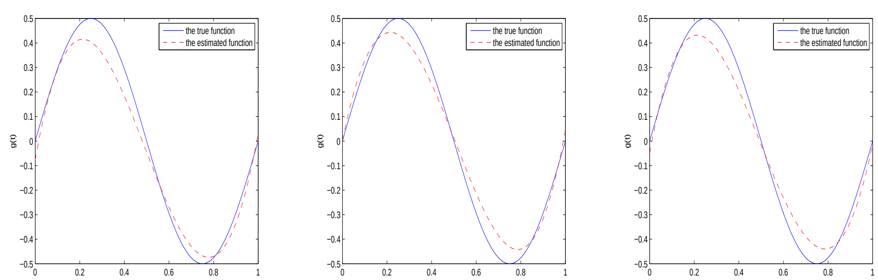



Figure 1. The average of the estimates versus the true value of $g(t)$ under three priors in Example 1: type I (left panel), type II (middle panel) and type III( right panel) when $n = 30$.

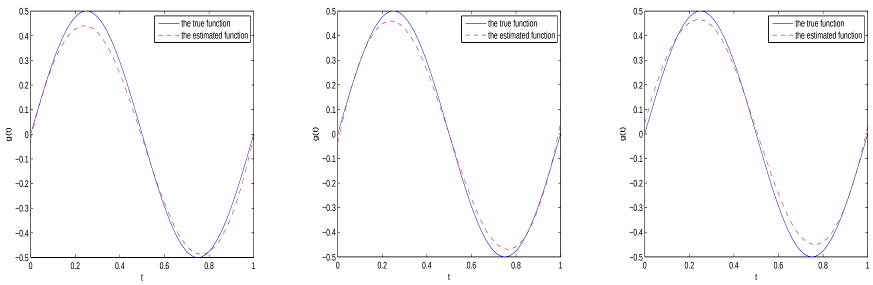



Figure 2. The average of the estimates versus the true value of $g(t)$ under three priors in Example 1: type I (left panel), type II (middle panel) and type III( right panel) when $n = 50$.

**Table 2.** Bayesian estimates of parameters under different priors when $n = 100$ and $n = 150$ in Example 1

| Type | Parameters | $n = 100$ | | | $n = 150$ | | |
|---|---|---|---|---|---|---|---|
| | | BIAS | RMS | SD | BIAS | RMS | SD |
| I | $\beta_1$ | 0.0028 | 0.0394 | 0.0398 | 0.0019 | 0.0349 | 0.0323 |
| | $\beta_2$ | 0.0005 | 0.0393 | 0.0398 | 0.0017 | 0.0340 | 0.0325 |
| | $\beta_3$ | 0.0005 | 0.0394 | 0.0393 | 0.0012 | 0.0346 | 0.0323 |
| | $\gamma_1$ | 0.0074 | 0.2050 | 0.1809 | 0.0290 | 0.1384 | 0.1449 |
| | $\gamma_2$ | 0.0059 | 0.1642 | 0.1723 | 0.0057 | 0.1465 | 0.1394 |
| | $\sigma^2$ | 0.0026 | 0.0405 | 0.0403 | 0.0036 | 0.0324 | 0.0332 |
| II | $\beta_1$ | 0.0012 | 0.0450 | 0.0400 | 0.0005 | 0.0341 | 0.0323 |
| | $\beta_2$ | 0.0055 | 0.0368 | 0.0398 | 0.0013 | 0.0319 | 0.0324 |
| | $\beta_3$ | 0.0001 | 0.0389 | 0.0398 | 0.0014 | 0.0351 | 0.0323 |
| | $\gamma_1$ | 0.0016 | 0.1507 | 0.1743 | 0.0281 | 0.1449 | 0.1434 |
| | $\gamma_2$ | 0.0303 | 0.1777 | 0.1729 | 0.0113 | 0.1583 | 0.1366 |
| | $\sigma^2$ | 0.0037 | 0.0409 | 0.0409 | 0.0020 | 0.0302 | 0.0328 |
| III | $\beta_1$ | 0.0002 | 0.0379 | 0.0402 | 0.0036 | 0.0345 | 0.0322 |
| | $\beta_2$ | 0.0008 | 0.0427 | 0.0401 | 0.0009 | 0.0323 | 0.0326 |
| | $\beta_3$ | 0.0000 | 0.0398 | 0.0399 | 0.0028 | 0.0348 | 0.0323 |
| | $\gamma_1$ | 0.0587 | 0.1764 | 0.1729 | 0.0018 | 0.1398 | 0.1400 |
| | $\gamma_2$ | 0.0318 | 0.1857 | 0.1734 | 0.0016 | 0.1531 | 0.1367 |
| | $\sigma^2$ | 0.0089 | 0.0380 | 0.0414 | 0.0060 | 0.0302 | 0.0332 |

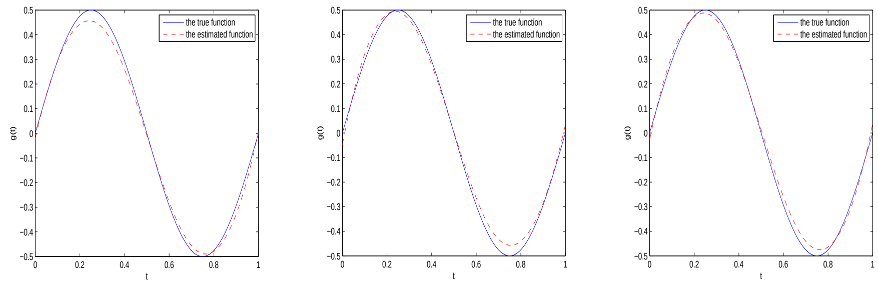



Figure 3. The average of the estimates versus the true value of $g(t)$ under three priors in Example 1: type I (left panel), type II (middle panel) and type III (right panel) when $n = 100$.

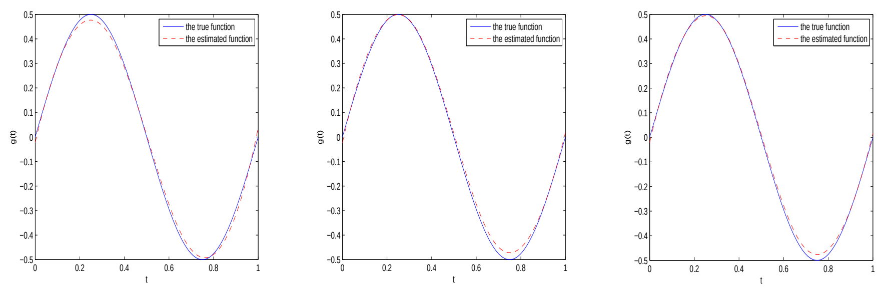



Figure 4. The average of the estimates versus the true value of $g(t)$ under three priors in Example 1: type I (left panel), type II (middle panel) and type III (right panel) when $n = 150$.

In Tables 1 and 2, “BIAS" denotes the absolute difference between the true value and the average of the Bayesian estimates of the parameters based on 100 replications, “SD" denotes the average of the estimated posterior standard deviation obtained from the formula in Section 3.4, and “RMS" denotes the root of mean square errors of the Bayesian estimates based on 100 replications. From Tables 1-2, we can make the following observations:

(i) the Bayesian estimates are reasonably accurate regardless of prior inputs in the sense of bias values of the estimates and their RMS values and SD values;

(ii) the estimates are mild sensitive to prior inputs for smaller sample size, but the infection clear away rapidly as the sample size goes large;

(iii) the estimates become better as the sample size increases, especially for the estimates of the parameters in the variance model.

Examination of Figures 1-4 shows that the shapes of the estimated nonparametric function are very close to the corresponding true line regardless of prior inputs. All in all, all the above findings show that the preceding proposed estimation procedures can well recover the true information in SMMEDRMs.

**4.2. Example 2: Comparisons for different prior inputs and the different number of internal knots.** To investigate the sensitivity of the Bayesian estimate for $g(t)$ to the selection of the number of internal knots, we consider the other two different choices of $K$ in this example, i.e. $K_1 = \lfloor K_0/1.5 \rfloor$ and $K_2 = \lfloor 1.5K_0 \rfloor$, where $K_0$ is the optimal number of interior knots and $\lfloor u \rfloor$ denotes the largest integer not greater than $u$. To save space, here we only present the results of Bayesian estimates in Table 3 and Figures 5-6 for $n = 50$ under different choices of $K$.

**Table 3.** Bayesian estimates of parameters for different choices of $K$ when $n = 50$ in Example 2

| | | $(n=50, K_1)$ | | | $(n=50, K_2)$ | | |
|---|---|---|---|---|---|---|---|
| Type | Parameters | BIAS | RMS | SD | BIAS | RMS | SD |
| I | $\beta_1$ | 0.0091 | 0.0543 | 0.0572 | 0.0031 | 0.0606 | 0.0575 |
| | $\beta_2$ | 0.0054 | 0.0612 | 0.0572 | 0.0022 | 0.0612 | 0.0573 |
| | $\beta_3$ | 0.0001 | 0.0570 | 0.0577 | 0.0030 | 0.0522 | 0.0572 |
| | $\gamma_1$ | 0.0198 | 0.2432 | 0.2624 | 0.0314 | 0.2589 | 0.2631 |
| | $\gamma_2$ | 0.0145 | 0.2677 | 0.2579 | 0.0244 | 0.2377 | 0.2509 |
| | $\sigma^2$ | 0.0102 | 0.0605 | 0.0598 | 0.0053 | 0.0643 | 0.0589 |
| II | $\beta_1$ | 0.0052 | 0.0551 | 0.0570 | 0.0161 | 0.0566 | 0.0572 |
| | $\beta_2$ | 0.0117 | 0.0523 | 0.0576 | 0.0020 | 0.0585 | 0.0573 |
| | $\beta_3$ | 0.0023 | 0.0608 | 0.0567 | 0.0003 | 0.0540 | 0.0575 |
| | $\gamma_1$ | 0.0719 | 0.2630 | 0.2681 | 0.0913 | 0.2571 | 0.2711 |
| | $\gamma_2$ | 0.0467 | 0.2382 | 0.2529 | 0.0398 | 0.2303 | 0.2565 |
| | $\sigma^2$ | 0.0077 | 0.0674 | 0.0586 | 0.0016 | 0.0590 | 0.0587 |
| III | $\beta_1$ | 0.0092 | 0.0578 | 0.0585 | 0.0030 | 0.0568 | 0.0581 |
| | $\beta_2$ | 0.0121 | 0.0575 | 0.0580 | 0.0042 | 0.0584 | 0.0581 |
| | $\beta_3$ | 0.0014 | 0.0566 | 0.0586 | 0.0024 | 0.0524 | 0.0585 |
| | $\gamma_1$ | 0.0476 | 0.2737 | 0.2540 | 0.0669 | 0.2323 | 0.2510 |
| | $\gamma_2$ | 0.0267 | 0.2554 | 0.2634 | 0.0643 | 0.2392 | 0.2507 |
| | $\sigma^2$ | 0.0249 | 0.0589 | 0.0610 | 0.0190 | 0.0624 | 0.0604 |

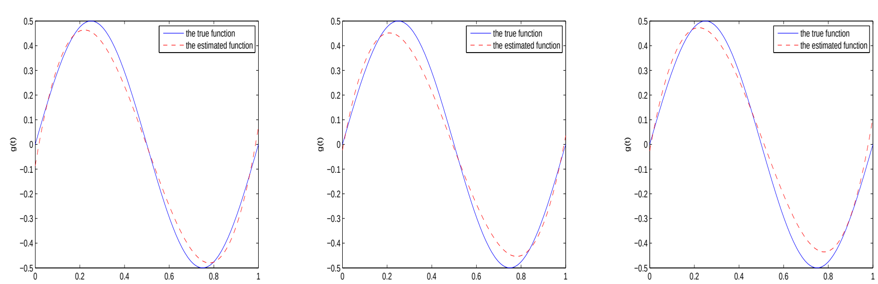



Figure 5. The average of the estimates versus the true value of $g(t)$ under three priors in Example 2: type I (left panel), type II (middle panel) and type III (right panel) for $n = 50$ and $K_1$.

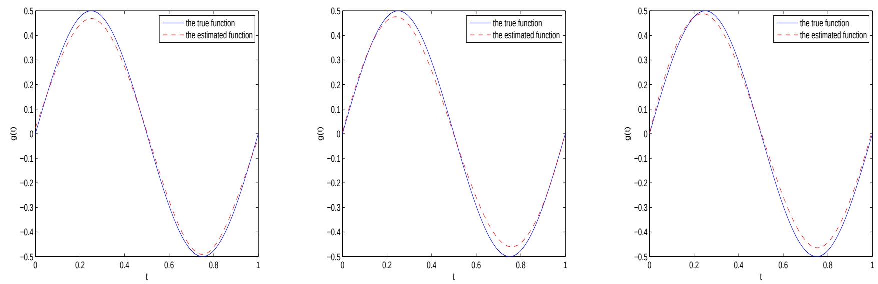



Figure 6. The average of the estimates versus the true value of $g(t)$ under three priors in Example 2: type I (left panel), type II (middle panel) and type III (right panel) for $n = 50$ and $K_2$.

By viewing Table 3 and comparing the results with Tables 1-2, we can see that the Bayesian estimates are reasonably accurate regardless of the values of $K$ in the sense of their SD values and RMS values. From Figures 5-6, we can obtain that the shapes of the estimated nonparametric function are very similar to those in Figure 2. Therefore, the Bayesian estimates for parameter estimates and the nonparametric function $g(t)$ are not very sensitive to the selection of the number of internal knots.

**4.3. Example 3: Comparisons for different prior inputs and different variance model.** To investigate the sensitivity of the proposed Bayesian method to the structure of the variance model in SMMEDRMs, we consider the other common structure of the variance model of the random effect $v_i$ (i.e. power product model), which can be defined as

$$\sigma_i^2 = \prod_{j=1}^{q} Z_{ij}^{\gamma_j}$$

with $\gamma = (1, -0.5)^T$ and $Z_i$ is a $2 \times 1$ vector with elements generated randomly from uniform distribution $U(0, 2)$. The simulation results for the parameters and the nonparametric function are reported in Table 4 and Figures 7-8.

The results in Table 4 show that with using power product model as the variance structure, which is different with the variance model in example 1, the proposed Bayesian method also has the desired performance, which is substantively similar to the results in example 1.

In addition, to consider the effect of variance structure misspecification on parameter estimates, here we do some simulations with n=50 and n=100 under Type I. The main measurements for comparison are differences between the fitted mean parameters $\hat{\beta}$ and the true mean parameters $\beta$, the fitted variances $\hat{\sigma}_i^2 (i = 1, 2, \cdots, n)$ to the true variances $\sigma_i^2 (i = 1, 2, \cdots, n)$, and the fitted error variance $\hat{\sigma}^2$ to the true error variance $\sigma^2$. In particular, we define three relative errors:

$$RERR(\hat{\beta}) = \left| \frac{\sum_{j=1}^{p} (\hat{\beta}_j - \beta_j)}{\sum_{j=1}^{p} \beta_j} \right|; RERR(\hat{\sigma}_i^2) = \left| \frac{\sum_{i=1}^{n} (\hat{\sigma}_i^2 - \sigma_i^2)}{\sum_{i=1}^{n} \sigma_i^2} \right|; RERR(\hat{\sigma}^2) = \left| \frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2} \right|.$$

Here variance structure misspecification means we use the variance structure in example 1 to model the variance of random effect. The results are reported in Table 5. From Table 5 we can find that when the true variance structure follows power product model, the errors in estimating $\hat{\beta}$, $\hat{\sigma}_i^2$ and $\hat{\sigma}^2$ increase when incorrectly modeling the variance using log-linear model. However, for this simulation study, variance model misspecification seems to affect the fitted results not larger, especially for the mean parameters and the error variance.

**Table 4.** Bayesian estimates of parameters under different priors in Example 3

| | | $n = 50$ | | | $n = 100$ | | |
|---|---|---|---|---|---|---|---|
| Type | Parameters | BIAS | RMS | SD | BIAS | RMS | SD |
| I | $\beta_1$ | 0.0005 | 0.0566 | 0.0579 | 0.0025 | 0.0411 | 0.0399 |
| | $\beta_2$ | 0.0029 | 0.0530 | 0.0572 | 0.0014 | 0.0431 | 0.0395 |
| | $\beta_3$ | 0.0000 | 0.0494 | 0.0576 | 0.0055 | 0.0421 | 0.0398 |
| | $\gamma_1$ | 0.0614 | 0.3186 | 0.3712 | 0.0368 | 0.2915 | 0.2398 |
| | $\gamma_2$ | 0.0287 | 0.2471 | 0.2383 | 0.0001 | 0.1540 | 0.1611 |
| | $\sigma^2$ | 0.0071 | 0.0592 | 0.0587 | 0.0021 | 0.0487 | 0.0405 |
| II | $\beta_1$ | 0.0024 | 0.0572 | 0.0580 | 0.0004 | 0.0411 | 0.0399 |
| | $\beta_2$ | 0.0063 | 0.0539 | 0.0574 | 0.0017 | 0.0432 | 0.0396 |
| | $\beta_3$ | 0.0033 | 0.0494 | 0.0576 | 0.0067 | 0.0402 | 0.0398 |
| | $\gamma_1$ | 0.1226 | 0.3470 | 0.3748 | 0.0722 | 0.3187 | 0.2375 |
| | $\gamma_2$ | 0.0507 | 0.2515 | 0.2422 | 0.0062 | 0.1550 | 0.1623 |
| | $\sigma^2$ | 0.0116 | 0.0597 | 0.0593 | 0.0032 | 0.0483 | 0.0405 |
| III | $\beta_1$ | 0.0042 | 0.0589 | 0.0585 | 0.0050 | 0.0432 | 0.0405 |
| | $\beta_2$ | 0.0005 | 0.0563 | 0.0580 | 0.0110 | 0.0469 | 0.0402 |
| | $\beta_3$ | 0.0085 | 0.0512 | 0.0583 | 0.0068 | 0.0479 | 0.0401 |
| | $\gamma_1$ | 0.0583 | 0.2874 | 0.3503 | 0.0493 | 0.2521 | 0.2416 |
| | $\gamma_2$ | 0.0130 | 0.2423 | 0.2365 | 0.0663 | 0.1766 | 0.1581 |
| | $\sigma^2$ | 0.0221 | 0.0627 | 0.0606 | 0.0154 | 0.0462 | 0.0416 |

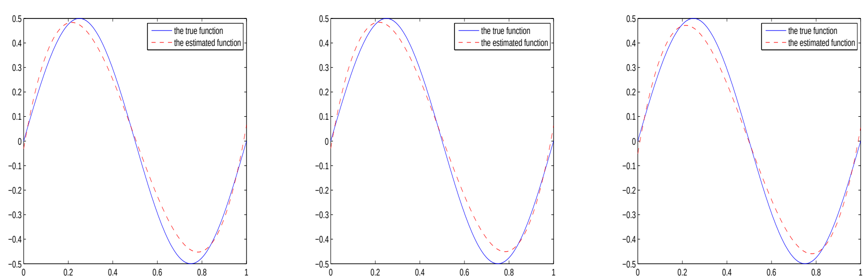



Figure 7. The average of the estimates versus the true value of $g(t)$ under three priors in Example 3: type I (left panel), type II (middle panel) and type III( right panel) when $n = 50$.

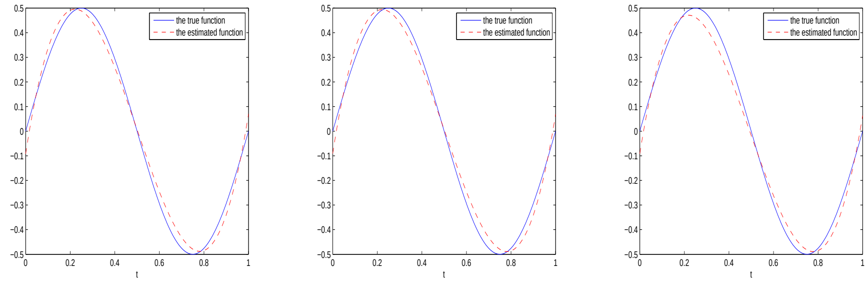



Figure 8. The average of the estimates versus the true value of $g(t)$ under three priors in Example 3: type I (left panel), type II (middle panel) and type III (right panel) when $n = 100$.

**Table 5.** Average of relative errors using different variance structures and sample size under Type I in Example 3

| | | $n = 50$ | $n = 100$ |
|---|---|---|---|
| correct specification | $RERR(\hat{\beta})$ | 0.0048 | 0.0013 |
| | $RERR(\hat{\sigma}_i^2)$ | 0.9226 | 0.6868 |
| | $RERR(\hat{\sigma}^2)$ | 0.0068 | 0.0003 |
| misspecification | $RERR(\hat{\beta})$ | 0.0067 | 0.0024 |
| | $RERR(\hat{\sigma}_i^2)$ | 2.4665 | 1.3371 |
| | $RERR(\hat{\sigma}^2)$ | 0.0069 | 0.0007 |

## 5. Application to Real Data

In this section, we illustrate the proposed method through analysis of a data set from the MultiCenter AIDS Cohort study. The dataset contains the human immunodeficiency virus (HIV) status of 283 homosexual men who were infected with HIV during a follow-up period between 1984 and 1991. This dataset has been used by many authors to illustrate semiparametric linear regression models, such as [6, 26]. The objective of their analysis is to describe the trend of the mean CD4 percentage depletion over time and evaluate the effects of smoking, the pre-HIV infection CD4 percentage, and age at HIV infection on the mean CD4 percentage after infection. This motivates us to use the semiparametric models for this dataset.

Let Y be the individual's CD4 percentage, $X_1$ be the smoking status:(1 for a smoker and 0 for a nonsmoker), $X_2$ be the centered age at HIV infection, $X_3$ be the centered preCD4 percentage. To model jointly the mean for the CD4 cell data and the variance of random effect in the model, we use the following semiparametric mixed-effects double regression models:

$$\begin{cases} Y_{ij} = \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + v_i + g(t_{ij}) + \varepsilon_{ij}, \\ \varepsilon_{ij} \sim N(0, \sigma^2), \\ v_i \sim N(0, \sigma_i^2), \\ \log(\sigma_i^2) = \gamma_1 Z_{1i} + \gamma_2 Z_{2i}, \\ i = 1, 2, \cdots, 283, \end{cases}$$

where $Z_1 = X_1, Z_2 = X_3$, $g(t)$, the baseline CD4 percentage, represents the mean CD4 percentage t years after the infection.

The preceding proposed hybrid algorithm is used to obtain Bayesian estimates of $\beta$'s, $\gamma$'s and $\sigma^2$, where we use noninformative prior information for all unknown parameters. In the Metropolis-Hastings algorithm, we set $\sigma_\gamma^2 = 1.8$ and $\sigma_v^2 = 0.015$ in their corresponding proposal distributions, which give approximate acceptance rates 43.76% and 31.37%. To test the convergence of the algorithm, plot of the EPSR values for all the unknown parameters against iterations is presented in Figure 9, which indicates that the algorithm converges about 5000 iterations because EPSR values of all unknown parameters are less than 1.2 about 5000 iterations. We calculate Bayesian estimates (EST), standard deviation estimates (SD) of the Bayesian estimates of $\beta$'s, $\gamma$'s and $\sigma^2$. Results are given in Table 6, which indicate that $X_3$ has significant impact on the mean of $Y$ and is somehow consistent with the results of variable selection seen in Fan and Li [6]. In addition, the curve of the estimated baseline function is shown in Figure 10. From Figure 10, we find that the mean baseline CD4 percentage decreases very quickly at the beginning of HIV infection, and the rate of decrease somewhat slows down four years after infection. The findings basically agree with that which was discovered by the local linear fitting method of Fan and Li [6].

**Table 6.** The real example: Bayesian estimates and their standard deviations

| Parameter | EST | SD |
|---|---|---|
| $\beta_1$ | 0.4431 | 0.5775 |
| $\beta_2$ | -0.1955 | 0.2648 |
| $\beta_3$ | 3.2706 | 0.2696 |
| $\gamma_1$ | 2.0399 | 0.3593 |
| $\gamma_2$ | 2.3333 | 0.2186 |
| $\sigma^2$ | 80.9993 | 2.7830 |

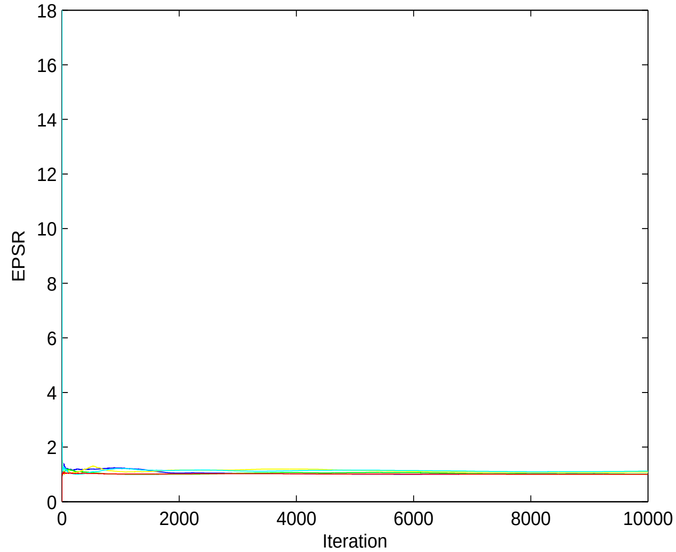



Figure 9. EPSR values of all parameters against iterations in the real example

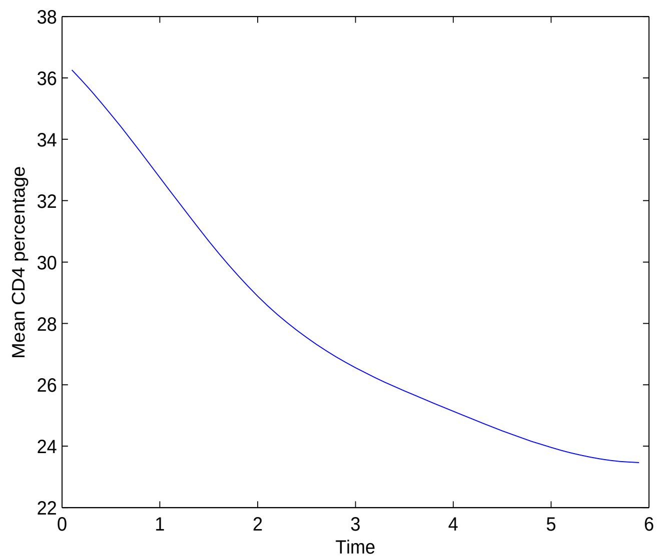



Figure 10. Application to AIDS data. The Bayesian estimate of the mean CD4 percentage $g(t)$. The solid line represents the estimated function.

## 6. Conclusion and Discussion

In this article, based on jointly modeling the mean and variance, we propose semi-parametric mixed-effects double regression models, in which we model the variance of the mixed effects directly as a function of the explanatory variables. Then we extend the Bayesian methodology proposed in [2, 23] to fit SMMEDRMs. A fully Bayesian approach is developed to analyze this models via B-spline estimate of the nonparametric part by combining the Gibbs sampler and Metropolis-Hastings algorithm. Some simulation studies and a real data are used to show the efficiency of the proposed Bayesian approach. The results show that the developed Bayesian method is highly efficient and computationally fast. A possible extension of the current model is being considered when covariates are missing under different missingness mechanisms.

## Acknowledgements

The authors would like to thank the editor and the referees for helpful comments, which lead to improvement of an earlier version of the paper. This work is supported by grants from the National Natural Science Foundation of China (11301485, 11271039, 11261025, 11526188), Collaborative Innovation Center on Capital Social Construction and Social Management (006000546615539), China Postdoctoral Science Foundation funded project (2014M562348), National Statistical Science Research Project(2015LY70), and Zhejiang Provincial Natural Science Foundation of China(LQ15A010008).

## References

[1] Aitkin, M. *Modelling variance heterogeneity in normal regression using GLIM*, Applied Statistics **36**, 332-339, 1987.

[2] Cepeda, E. and Gamerman, D. *Bayesian modeling of variance heterogeneity in normal regression models*, Brazilian Journal of Probability and Statistics **14**, 207-221, 2001.

[3] Chen, X. D. *Bayesian analysis of semiparametric mixed-effects models for zero-inflated count data*, Communications in Statistics−Theory and Methods **38**, 1815-1833, 2009.

[4] Chen, X. D. and Tang, N. S. *Bayesian analysis of semiparametric reproductive dispersion mixed-effects models*, Computational Statistics & Data Analysis **54**, 2145-2158, 2010.

[5] Fan, J. and Gijbels, I. *Local polynomial modelling and its application*, Chapman and Hall, New York, 1996.

[6] Fan, J. and Li, R. *New estimation and model selection procedures for semiparametric modeling in longitudinal data analysis*, Journal of the American Statistical Association **99**, 710-723, 2004.

[7] Gelman, A. *Inference and monitoring convergence in markov chain monte carlo in practice*, Chapman and Hall, London, 1996.

[8] Gelman, A., Roberts, G. O. and Gilks, W. R. *Efficient metropolis jumping rules*, In:J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, (Eds), *Bayesian statistics*, vol. 5. Oxford University Press, Oxford, 599-607, 1995.

[9] Geyer, C. J. *Practical markov chain monte carlo*, Statistical Science **7**, 473-511, 1992.

[10] He, X., Fung, W. K. and Zhu, Z. Y. *Robust estimation in a generalized partial linear model for clustered data*, Journal of the American Statistical Association **100**, 1176-1184, 2005.

[11] Lee, S. Y. and Zhu, H. T. *Statistical analysis of nonlinear structural equation models with continuous and polytomous data*, British Journal of Mathematical and Statistical Psychology **53**, 209-232, 2000.

[12] Li, A. P., Chen, Z. X. and Xie, F. C. *Diagnostic analysis for heterogeneous log-Birnbaum-Saunders regression models*, Statistics & Probability Letters **82**, 1690-1698, 2012.

[13] Lin, J. G. and Wei, B. C. *Testing for heteroscedasticity in nonlinear regression models*, Communications in Statistics−Theory and Methods **32**, 171-192, 2003.

[14] Lin, T. I. and Wang, W. L. *Bayesian inference in joint modelling of location and scale parameters of the t distribution for longitudinal data*, Journal of Statistical Planning and Inference **141**, 1543-1553, 2011.

[15] Ni, X., Zhang, D. and Zhang, H. H. *Variable selection for semiparametric mixed models in longitudinal studies*, Biometrics **66**, 79-88, 2010.

[16] Roberts, G. *Markov chain concepts related to sampling algorithm*, Chapman and Hall, London, 1996.

[17] Schumaker, L. L. *Spline function*, Wiley, New York, 1981.

[18] Tang, N. S. and Duan, X. D. 2012. *A semiparametric bayesian approach to generalized partial linear mixed models for longitudinal data*, Computational Statistics & Data Analysis **56**, 4348-4365.

[19] Tian, R. Q. and Xue, L. G. *Variable selection for semiparametric errors-in-variables regression model with longitudinal data*, Journal of Statistical Computation and Simulation **84**, 1654-1669, 2014.

[20] Wu, L. C. *Variable selection in joint location and scale models of the skew-t-normal distribution*, Communications in Statistics −Simulation and Computation **43**, 615-630, 2014.

[21] Wu, L. C. and Li, H. Q. *Variable selection for joint mean and dispersion models of the inverse Gaussian distribution*, Metrika **75**, 795-808, 2012.

[22] Xie, F. C., Wei, B. C. and Lin, J. G. *Homogeneity diagnostics for skew-normal nonlinear regression models*, Statistics & Probability Letters **79**, 821-827, 2009.

[23] Xu, D. K. and Zhang, Z. Z. *A semiparametric Bayesian approach to joint mean and variance models*, Statistics & Probability Letters **83**, 1624-1631, 2013.

[24] Xu, D. K., Zhang, Z. Z. and Du, J. *Skew-normal semiparametric varying coefficient model and score test*, Journal of Statistical Computation and Simulation **85**, 216-234, 2015.

[25] Zhao, W. H., Zhang, R. Q., Lv, Y. Z. and Liu, J. C. *Variable selection for varying dispersion beta regression model*, Journal of Applied Statistics **41**, 95-108, 2014.

[26] Zhao, P. X. and Xue, L. G. *Variable selection for semiparametric varying coefficient partially linear errors-in-variables models*, Journal of Multivariate Analysis **101**, 1872-1883, 2010.

## *ADVICE TO AUTHORS*

*Hacettepe Journal of Mathematics and Statistics* publishes short to medium length research papers and occasional survey articles written in English. All papers are refereed to international standards.

**Address for Correspondence**

Editorial Office,
Hacettepe Journal of Mathematics and Statistics,
Hacettepe University,
Faculty of Science,
Department of Mathematics,
06532 Beytepe,
Ankara,
Turkey.

**E-mail**: hjms@hacettepe.edu.tr
**Tel**: Editors: + 90 312 297 7898
+ 90 312 297 7850 / 122
Associate Editor: + 90 312 297 7880
**Fax** : + 90 312 299 2017
**Web Page** : http//www.hjms.hacettepe.edu.tr

**Advise to Authors :** The style of articles should be lucid but concise. The text should be preceded by a short descriptive title and an informative abstract of not more than 100 words. Keywords or phrases and the 2010 AMS Classification should also be included.The main body of the text should be divided into numbered sections with appropriate headings. Items should be numbered in the form 2.4. Lemma, 2.5. Definition. These items should be referred to in the text using the form Lemma 2.4., Definition 2.5. Figures and tables should be incorporated in the text. A numbered caption should be placed above them. References should be punctuated according to the following examples, be listed in alphabetical order according to in the (first) author' s surname, be numbered consecutively and referred to in the text by the same number enclosed in square brackets. Only recognized abbreviations of the names of journals should be used.

[1] Banaschewski, B. *Extensions of topological spaces*, Canad. Math. Bull. **7** (1), 1–22, 1964.

[2] Ehrig, H. and Herrlich, H. *The construct PRO of projection spaces: its internal structure*, in: Categorical methods in Computer Science, Lecture Notes in Computer Science **393** (Springer-Verlag, Berlin, 1989), 286–293.

[3] Hurvich, C. M. and Tsai, C. L. *Regression and time series model selection in small samples*, Biometrika **76** (2), 297–307, 1989.

[4] Papoulis, A. *Probability random variables and stochastic process* (McGraw-Hill, 1965).

**INSTRUCTIONS FOR AUTHORS**

**Preparation of Manuscripts :** Manuscripts will be typeset using the LATEX typesetting system. Authors should prepare the article using the HJMS style before submission by e-mail. The style bundle contains the styles for the article and the bibliography as well as example of formatted document as pdf file. The authors should send their paper directly to hjms@hacettepe.edu.tr with the suggestion of area editor and potential reviewers for their submission. The editorial office may not use these suggestions, but this may help to speed up the selection of appropriate reviewers.

**Copyright :** No manuscript should be submitted which has previously been published, or which has been simultaneously submitted for publication elsewhere. The copyright in a published article rests solely with the Faculty of Science of Hacettepe University, and the paper may not be reproduced in whole in part by any means whatsoever without prior written permission.

**Notice :** The content of a paper published in this Journal is the sole responsibility of the author or authors, and its publication does not imply the concurrence of the Editors or the Publisher.