2022 Proceedings of International E-Conference on Mathematical and Statistical Sciences: A Selçuk Meeting

ISBN: 978-625-00-9195-1

https://icomss22.selcuk.edu.tr/

2022 Proceedings of International E-Conference on Mathematical and Statistical Sciences: A Selçuk Meeting

ISBN: 978-625-00-9195-1

Edited by Tuncer Acar

Honorary Chair of ICOMSS'22

Metin Aksoy
Rector of Selçuk University, Konya, Türkiye

Chair of ICOMSS'22

Semahat Küçükkolbaşı
Dean of Faculty of Science of Selçuk University, Konya, Türkiye

Co-Chairs of ICOMSS'22

Tuncer Acar
Department of Mathematics, Faculty of Science, Selçuk University, Konya, Türkiye

İsmail Kınacı
Department of Actuarial Science, Faculty of Science, Selçuk University, Konya, Türkiye

Coşkun Kuş
Department of Statistics, Faculty of Science, Selçuk University, Konya, Türkiye

Scientific Committee

Yunus Akdoğan
Selçuk University, Türkiye

Fazıl Aliyev
Rutgers University, USA

Ali Aral
Kırıkkale University, Türkiye

Barry C. Arnold
University of California, USA

Narayanaswamy Balakrishnan
McMaster University, Canada

Carlo Bardaro
University of Perugia, Italy

Danilo Costarelli
University of Perugia, Italy

Borislav R. Draganov
Sofia University, Bulgaria

Pedro Garrancho
Universidad de Jaén, Spain

Vijay Gupta
Netaji Subhas University of Technology, India

Harun Karslı
Bolu Abant Izzet Baysal University, Türkiye

Kadir Karakaya
Selçuk University, Türkiye

Mehmet Fedai Kaya
Selçuk University, Türkiye

Sathish Kumar
IIT Madras, India

Mohammad Mursaleen
Aligarh Muslim University, India

Haikady N. Nagaraja
Ohio State University, USA

Hon Keung Tony Ng
Bentley University, USA

Fırat Özsaraç
Kırıkkale University, Türkiye

Ioan Raşa
Technical University of Cluj-Napoca, Romania

Buğra Saraçoğlu
Selçuk University, Türkiye

Calogero Vetro
University of Palermo, Italy

Gianluca Vinti
University of Perugia, Italy

Shuo-Jye Wu
Tamkang University, Taiwan

Organizing Committee

Thabet Abdeljawad
Prince Sultan University, Saudi Arabia

Özlem Acar
Selçuk University, Türkiye

Praveen Agarwal
Anand International College of Engineering, India

Akbar Asgharzadeh
University of Mazandaran, Iranian

Kemâl Aydın
Selçuk University, Türkiye

Dan Barbosu
Technical University of Cluj, Romania

Murat Bodur
Konya Technical University, Türkiye

Naim L. Braha
University of Prishtina, Prishtina

Vural Cam
Selçuk University, Türkiye

Naokant Deo
Delhi Technological University, India

Borislav R. Draganov
Sofia University, Bulgaria

Roman Dmytryshyn
Vasyl Stefanyk Precarpathian National University, Ukranie

Naokant Deo
Delhi Technological University, India

Tenzile Erbayram
Selçuk University, Türkiye

Meenu Goyal
Thapar Institute of Engineering and Technology, India

Arun Kajla
Central University of Haryana, India

Fahreddin Kalkan
Selçuk University, Türkiye

İsmail Hakkı Kınalıoğlu
Selçuk University, Türkiye

Sathish Kumar
VNIT Nagpur University, India

R. N. Mohapatra
University of Central Florida, USA

Syed Abdul Mohiuddine
King Abdulaziz University, Saudi Arabia

Muhammed Nazam
Allama Iqbal Open University, Pakistan

Hon Keung Tony Ng
Bentley University, USA

Dilek Söylemez Özden
Selçuk University, Türkiye

Haldun Alpaslan Peker
Selçuk University, Türkiye

Sümeyra Sert
Selçuk University, Türkiye

Demet Sezer
Selçuk University, Türkiye

Honey Sharma
Gulzar Group of Institutions, India

Güner Topcu
Selçuk University, Türkiye

Metin Turgay
Selçuk University, Türkiye

Aynur Yalçıner
Selçuk University, Türkiye

Secretariat

Abdulkadir Eke
Selçuk University, Türkiye

Eşref Erdoğan
Selçuk University, Türkiye

Sadettin Kurşun
Selçuk University, Türkiye

Dilek Özer
Selçuk University, Türkiye

Aybala Sevde Özkapu
Selçuk University, Türkiye

Message from Editor

It's my pleasure to edit Proceedings Book of "International E-Conference on Mathematical and Statistical Sciences: A Selçuk Meeting". Our 2022 conference, which is organized by the Faculty of Science of Selçuk University and supported by Scientific Research Projects Coordinatorship of Selçuk University. By organizing this e-conference, our main aim was the to promote, encourage, and provide a forum for the academic exchange of ideas and recent research works. The conference present new results and future challenges, in a series of virtual keynote lectures and virtual contributed short talks. In our conferences, we provide a forum for mathematicians and statistician to communicate recent research results in the areas of Algebra and Applied Mathematics, Analysis, Geometry and Topology, Actuarial Science, Applied Statistics and Statistical Theory. The conference was only online and there was no registration fee, and only one presentation was allowed for each participants. The all presentation language was English, and submissions were peer-reviewed by at least two referees. This proceedings book also includes the papers which are peer-reviewed by at least two referees.

Thanks.

Assoc. Prof. Dr. Tuncer Acar
Selçuk University
Editor of Proceedings Book of ICOMSS'22

Contents

Generalised Picone Identity and First Eigenvalue for $p(x)$-sub-Laplacian on Stratified Groups

Abimbola Abolarinwa

ABSTRACT. In this paper, we present a generalised variable exponent Picone type identity for horizontal $p(x)$-Laplacian on general stratified Lie groups. As applications we establish some properties of the first Dirichlet eigenvalue of horizontal $p(x)$-sub-Laplacian such uniqueness, simplicity, monotonicity and isolation in variable Lebesgue spaces.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 35B05 35H20, 35P30, 47J10

KEYWORDS: Stratified groups, Picone identity, sub-Laplacian, principal eigenvalue.

1. INTRODUCTION

The study of variable exponent partial differential equations in Euclidean setting is gaining more attention [9, 11, 13, 15, 16, 18]. Models involving $p(x)$-growth condition arise from physical processes such as nonlinear elasticity theory, electrorheological fluids, image processing, etc [2, 3, 24]. It has been observed that $p(x)$-Laplacian is similar in many respect to the classical p-Laplacian (p-constant) but it lacks certain vital properties such as homogeneity. This therefore makes the nonlinearity so much complicated and many of known approaches to p-Laplacian can no longer hold for $p(x)$-Laplacian. It is interesting to consider $p(x)$-Laplacian in the sub-elliptic setting and investigate which of the known results for p-constant hold for variable exponents.

In this paper however, we study the indefinite weighted Dirichlet eigenvalue problem for $p(x)$-sub-Laplacian (5) on $\Omega \subset \mathbb{G}$, $p(x) > 1$, where $\mathbb{G}$ is a stratified Lie group, and discuss some properties of the eigenvalue $\lambda \in \mathbb{R}^+$ and the corresponding eigenfunction $u(x)$ in the context of variable exponent Sobolev spaces. It is well known in the classical setting ($p(x) = p$-constant and $M = \mathbb{R}^n$) that Problem (5) possesses a closed set of nondecreasing sequence of nonnegative eigenvalues $\{\lambda_k\}$ which grows to $+\infty$ as $k \to +\infty$, and that the first nonzero eigevalue is simple and isolated. Due to some complication in the nonlinearities in $p(x)$-Laplacian and inhomogeneity of the corresponding variable exponent norm, some of the results in the classical case may not hold or rather under restrictive assumptions. In [13], the authors studied (5) (with $g(x) = 1$, $\Omega = \mathbb{R}^n$) and showed that the existence of infinitely many eigenvalues and established some sufficient condition for the infimum of the spectrum, called the principal eigenvalue, to be positive. The properties that $\lambda_1 > 0$ is very useful in analysis and applications. Motivated by [13], we are able to assume the existence of $\lambda_1 > 0$ for (5) and proved its uniqueness, monotonicity, simplicity and isolation. The variable exponent Picone type identity is key in all our proofs.

Picone identity is a very useful tool in the study of qualitative properties of solutions of differential equations, and for this, several linear and nonlinear Picone type identities have been derived to handle differential equations of various type. Picone identity was originally developed by Mauro Picone in 1910 to prove Sturm Comparison principle and oscillation theory for a system of differential equations. This identity was later extended to partial differential equation involving Laplacian by Allegretto [4] as follows: for nonnegative differentiable functions u and v with $v \neq 0$, the following formula

$$|\nabla u|^2 + \frac{u^2}{v^2}|\nabla v|^2 - 2\frac{u}{v}\nabla u \nabla v = |\nabla u|^2 - \nabla\left(\frac{u^2}{v}\right)\nabla v \geq 0 \tag{1}$$

holds. It was later generalised to p-Laplacian by Allegretto and Huang [5] which enable them to establish among others, existence and nonexistence of positive solutions, Sturmian comparison principle, Liouville type theorems, Hardy inequalities and some profound results involving p-Laplace equations and systems. Allegretto and Huang's identity reads as follows, for $u \geq 0$, $v > 0$, then

$$|\nabla u|^p + (p-1)\frac{u^p}{v^p}|\nabla v|^p - p\frac{u^{p-1}}{v^{p-1}}|\nabla v|^{p-2}\nabla v \nabla u = R_p(u,v), \tag{2}$$

where

$$R_p(u,v) := |\nabla u|^p - \nabla\left(\frac{u^p}{v^{p-1}}\right)|\nabla v|^{p-2}\nabla v \geq 0.$$

Several extensions and generalisation of Picone identity have been established in order to handle more general elliptic operators. For instance, see [7, 27] for nonlinear versions of (1) and (2), with several applications. For other interesting extension of Picone type identities one can find [21, 22, 23] for general vector fields and p-sub-Laplacian with applications to Grushin plane, Heisenberg group, Stratified Lie groups, [19] (for p-sub-Laplacian on Heisenberg group and applications to Hardy inequalities), [25] for nonlinear Picone identities for p-sub-Laplacian with applications to horizontal Hardy inequalities and weighted eigenvalue problem on Stratified Lie groups.

Allegretto [6] established variable exponent Picone type identity for differentiable functions $v > 0$, $0 \leq u \in C_0^\infty(\Omega)$ and continuous $p(x) > 1$ as follows:

$$\begin{aligned} &\frac{|\nabla u|^{p(x)}}{p(x)} - \nabla\left[\frac{u^{p(x)}}{p(x)v^{p(x)-1}}\right]|\nabla v|^{p(x)-2}\nabla v \\ &\quad = \frac{|\nabla u|^{p(x)}}{p(x)} - \left(\frac{u}{v}\right)^{p(x)-1}|\nabla v|^{p(x)-2}\nabla v\nabla u + \frac{p(x)-1}{p(x)}\left(\frac{u}{v}|\nabla v|\right)^{p(x)} \\ &\qquad + \frac{1}{p(x)}\frac{u^{p(x)}}{v^{p(x)-1}}|\nabla v|^{p(x)-2}\left[\frac{1}{p(x)} - \ln\left(\frac{u}{v}\right)\right]\nabla v\nabla p(x) \geq 0 \end{aligned} \tag{3}$$

on the assumption that $\nabla v \nabla p(x) = 0$. He used the inequality to prove Barta theorem and some other results. Later, Yoshida [26] established similar Picone identities for quasilinear and half-linear elliptic equations involving $p(x)$-Laplacian and pseudo $p(x)$-Laplacian, and consequently developed Sturmian comparison theory. Most recently, Feng and Han [14], motivated by Allegretto [6] proved a modified form of (3) and showed that

$$|\nabla u|^{p(x)} - \nabla\left(\frac{u^{p(x)}}{v^{p(x)-1}}\right)|\nabla v|^{p(x)-2}\nabla v \geq 0 \tag{4}$$

provided $\nabla v \nabla p(x) = 0$ a.e in Ω, with equality if and only if $\nabla(u/v) = 0$ in Ω. They proved monotonicity of principal eigenvalue and a variable exponent Barta inequality for $p(x)$-Laplacian in the the Euclidean setting.

In this paper, we present a new generalised variable exponent Picone type identity for horizontal $p(x)$-Laplacian on stratified Lie groups. Our identity contains some known identities in various setting as will be discussed in Section 2. Consequently, we give several applications to qualitative properties of the principal eigenvalue of $p(x)$-sub-Laplacian such as uniqueness, simplicity, monotonicity and isolatedenss.

2. Preliminaries

In the section we recall some preliminaries which will allow us fix necessary notation. First, we give basics of the stratified group (homogeneous Carnot group). Secondly, we discuss some concepts from the theory of variable Lebesgue and Sobolev spaces, and then present Dirichlet eigenvalue problem for $p(x)$-sub-Laplacian on stratified Lie groups.

2.1. **Stratified groups.** Recall that a stratified Lie algebra $\mathbf{g}$ of step r is a Lie algebra with subspaces $V_1, \cdots, V_r$ satisfying

$$\mathbf{g} = V_1 \oplus \cdots \oplus V_r, \quad [V_1, V_j] = V_{j+1}, \quad j = 1, \cdots, r-1 \text{ and } [V_1, V_r] = 0.$$

If a Lie group $\mathbb{G}$ is Nilpotent, connected and simply connected with a stratified Lie algebra $\mathbf{g}$, then $\mathbb{G}$ is called a stratified Lie group (or homogeneous Carnot group) of step r. In fact, if $N_j = \dim V_j$, using the exponential map we can identify $\mathbb{G}$ with $\mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_r}$ so that each point $\mathbf{g} \in \mathbb{G}$ is identified with a point $x = (x^{(1)}, \cdots, x^{(r)})$ in $\mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_r}$ so that $x^j \in \mathbb{R}^{N_j}$.

The stratified Lie groups have a natural family of anisotropic dilations $\delta_\gamma : \mathbb{G} \to \mathbb{G}$ for every $\gamma > 0$ defined by

$$\delta_\gamma(x) \equiv \delta_\gamma(x', x^{(2)}, \cdots, x^{(r)}) := (\gamma x', \gamma^2 x^{(2)}, \cdots, \gamma^r x^{(r)}),$$

which is an automorphism of $\mathbb{G}$. Here $x' = x^{(1)} \in \mathbb{R}^{N_1}$ and $x^{(k)} \in \mathbb{R}^{N_k}$ for $k = 2, \cdots, r$. The homogeneous dimension Q of the group $\mathbb{G}$ is given by

$$Q := \sum_{k=1}^{r} k \cdot \dim V_k = \sum_{k=1}^{r} k N_k.$$

The Haar measure on $\mathbb{G}$ is denoted by dx, which can be taken to be the Lebesgue measure on $\mathbb{R}_{N_1} \times \cdots \times \mathbb{R}_{N_r}$ Let $X_1, \cdots, X_{N_1}$ be the left invariant vector fields on $\mathbb{G}$ such that $X_k(0) = \frac{\partial}{\partial x_k}\Big|_0$ for $k = 1, \cdots, N_1$. Then, for every $x \in \mathbb{R}_N = \mathbb{R}_{N_1} \times \cdots \times \mathbb{R}_{N_r}$, the Hörmander rank condition

$$\operatorname{rank}(\operatorname{Lie}\{X_1, \cdots, X_{N_1}\}) = N$$

holds, that is, the iterated commutators of $X_1, \cdots, X_{N_1}$ span the Lie algebra of $\mathbb{G}$. The left invariant vector fields X_k has an explicit form given by (see detail in the book [20]),

$$X_k = \frac{\partial}{\partial x'_k} + \sum_{l=2}^{r} \sum_{m=1}^{N_l} a^{(l)}_{k,m}(x', \cdots, x^{(l-1)}) \frac{\partial}{\partial x^{(l)}_m},$$

where $a^{(l)}_{k,m}$ is a homogeneous polynomial function of degree $l-1$, r is the step of $\mathbb{G}$, $x = (x', x^{(2)}, \cdots, , x^{(r)})$, and $x^{(l)} = (x^{(l)}_1, \cdots, x^{(l)}_{N_l})$ are the variables in the l^{th}-stratum.

The canonical (horizontal) sub-Laplacian on $\mathbb{G}$ is defined by

$$\mathscr{L}_X := \sum_{k=1}^{N} X_k^2,$$

which is a left invariant homogeneous second-order hypoelliptic differential operator. It is elliptic if and only if the step of $\mathbb{G}$ is equal to 1. (The hypoellipticity of $\mathscr{L}$ is a special case of Hörmander's sum of square [17]). The Horizontal gradient and divergence on $\mathbb{G}$ are respectively denoted by

$$\nabla_H := (X_1, \cdots, X_{N_1}) \quad \text{and} \quad \operatorname{div}_H w := \nabla_H \cdot w.$$

Let $p : \bar{\Omega} \to \mathbb{R}$ be a continuous function and $p(x) > 1$ for $x \in \bar{\Omega} \subset M$. We define the $p(x)$-sub-Laplacian (horizontal $p(x)$-Laplacian) on $\mathbb{G}$ by the formula

$$\mathscr{L}_{p(x)} u := \operatorname{div}_H(|\nabla_H u|^{p(x)-2} \nabla_H u),$$

where u is a smooth function. If $p(x) = p$ (p=constant), the operator $\mathscr{L}_p u$ becomes the p-sub-Laplacian, $\operatorname{div}_H(|\nabla_H u|^{p-2} \nabla_H u)$. Here the notation

$$|x'| = (x'^2_1 + \cdots + x'^2_{N_1})^{\frac{1}{2}}$$

stands for the Euclidean norm in $\mathbb{R}^{N_1}$.

2.2. **Variable Lebesgue spaces.** In order to discuss generalised solutions, we need some concepts from the theory of variable Lebesgue and Sobolev spaces. Detailed description of these spaces can be found in [8, 10, 12].

Let $\Omega \subset \mathbb{G}$ be an open domain and $\mathcal{E}(\Omega)$ denotes the set of all equivalence classes of Haar measurable real-valued functions defined on Ω being equal almost everywhere. For any positive variable exponent $p(x) : \Omega \to [1, \infty)$ such that

$$1 < p^- := ess \inf_{x \in \bar{\Omega}} p(x) \le p(x) \le p^+ := ess \sup_{x \in \bar{\Omega}} p(x) < \infty,$$

and any $u \in \mathcal{E}(\Omega)$, the identity

$$|u(x)|^{p(x)} = \exp\left[p(x) \log |u(x)|\right]$$

shows that $|u(x)|^{p(x)}$ is measurable. Since it is also nonnegative, the integral $\int_\Omega |u(x)|^{p(x)} dx$ is well defined. Consider the functional (called the ρ-modular) $\rho_{p(\cdot)} : \mathcal{E}(\Omega) \to [0, \infty]$ given by

$$\rho_{p(\cdot)}(u) := \int_{\Omega \setminus \Omega_\infty} |u(x)|^{p(x)} dx + \|u\|_{L^\infty(\Omega_\infty)}.$$

Here we assume that Ω_∞ has zero measure and we stick to

$$\rho_{p(x)}(u) := \int_\Omega |u(x)|^{p(x)} dx.$$

Definition 2.1. *The generalised (variable exponent) Lebesgue space $L^{p(\cdot)}(\Omega)$ is the vector space of measurable function for which the ρ-modular is finite, that is*

$$L^{p(\cdot)}(\Omega) = \left\{ u \in \mathcal{E}(\Omega) : \int_\Omega |u(x)|^{p(\cdot)} dx < \infty \right\}.$$

The space $L^{p(\cdot)}(\Omega)$ is a Banach space equipped with the (Luxemburg) norm

$$\|u\|_{L^{p(\cdot)}(\Omega)} = \inf \left\{ t > 0 : \rho_{p(\cdot)}\left(\frac{u}{t}\right) \le 1 \right\}.$$

The ρ-modular $\rho_{p(\cdot)}(u)$ possesses some useful properties such as (see [8] for details)

(a) $\rho_{p(\cdot)}(u) \ge 0$ with equality if and only if $u = 0$ a.e.
(b) $\rho_{p(\cdot)}(u) \le \rho_{p(\cdot)}(v)$ when $|u| \le |v|$ a.e.
(c) $\rho_{p(\cdot)}(u)$ is convex.
(d) For $\{u_k\} \subset \mathcal{E}(\Omega)$, $\rho_{p(\cdot)}(\liminf_{k\to\infty} |u_k|) \le \liminf_{k\to\infty} \rho_{p(\cdot)}(|u_k|)$ (Fatou's Lemma).

Another basic result is a generalised Hölder's inequality which can be used to define an equivalent norm. If $\frac{1}{p(x)} + \frac{1}{p'(x)} = 1$ a.e. on Ω, then for all $u \in L^{p(\cdot)}(\Omega)$ and $v \in L^{p'(\cdot)}(\Omega)$ we have $uv \in L^1(\Omega)$ and

$$\int_\Omega |u(x)v(x)| dx \le \left(1 + \frac{1}{p^-} - \frac{1}{p^+}\right) \|u\|_{L^{p(\cdot)}(\Omega)} \|v\|_{L^{p('\cdot)}(\Omega)}.$$

Definition 2.2. *The variable exponent Sobolev space $S^{1,p(\cdot)}(\Omega)$ is the vector space of those functions $u \in L^{p(\cdot)}(\Omega)$ for which $\nabla_H u$ is also $L^{p(\cdot)}(\Omega)$. $S^{1,p(\cdot)}(\Omega)$ is also a Banach space with the norm*

$$\|u\|_{S^{1,p(\cdot)}(\Omega)} = \|u\|_{L^{p(\cdot)}(\Omega)} + \|\nabla_H u\|_{L^{p(\cdot)}(\Omega)}.$$

We denote by $\mathring{S}^{1,p(\cdot)}(\Omega)$ the closure of $C_0^\infty(\Omega)$ in $S^{1,p(\cdot)}(\Omega)$ with respect to the norm

$$\|u\|_{\mathring{S}^{1,p(\cdot)}(\Omega)} = \|\nabla_H u\|_{L^{p(\cdot)}(\Omega)}.$$

It can be clearly seen that $L^{p(\cdot)}(\Omega)$, $S^{1,p(\cdot)}(\Omega)$ and $\mathring{S}^{1,p(\cdot)}(\Omega)$ are all separable and reflexive Banach spaces in their respectful norms if $1 < \inf p(x) < \sup p(x) < \infty$ in Ω.

2.3. **Eigenvalue problem for $p(x)$-Laplacian.** Let $\Omega \subset M$ be a bounded domain with smooth boundary $\partial\Omega$. We suppose a continuous function $p : \bar{\Omega} \to \mathbb{R}^+$, $p(x) > 1$ is such that

$$1 < p^- := ess \inf_{x\in\bar{\Omega}} p(x) \le p(x) \le p^+ := ess \sup_{x\in\bar{\Omega}} p(x) < \infty.$$

Now consider the indefinite weighted Dirichlet eigenvalue problem for horizontal $p(x)$- Laplacian

$$\begin{aligned} -\mathscr{L}_{p(x)} u &= \lambda g(x)|u|^{p(x)-2}u, && x \in \Omega, \\ u &> 0, && x \in \Omega, \\ u &= 0, && x \in \partial\Omega, \end{aligned} \tag{5}$$

where Ω is as defined above, $g(x)$ is a positive bounded function and $p : \bar{\Omega} \to [1, \infty)$ is a continuous function for $x \in \bar{\Omega}$.

Definition 2.3. *Let $\lambda \in \mathbb{R}^+$ and $u \in \mathring{S}^{1,p(x)}(\Omega)$, the pair (u, λ) is called a solution of* (5) *if*

$$\int_\Omega |\nabla_H u|^{p(x)-2} \nabla_H u \cdot \nabla_H \phi dx - \lambda \int_\Omega g(x)|u|^{p(x)-2} u\phi dx = 0 \tag{6}$$

for all $\phi \in W_0^{1,p(x)}(\Omega)$. If (u, λ) is a solution of (5)*, we call λ an eigenvalue, and u an eigenfunction corresponding to λ.*

Similarly, by the sup-solution and sub-solution of (5)*, we mean the pair (u, λ) such that*

$$\int_\Omega |\nabla_X u|^{p(x)-2} \nabla_H u \cdot \nabla_H \phi dx - \lambda \int_\Omega g(x)|u|^{p(x)-2} u\phi dx \ge 0 \tag{7}$$

and

$$\int_\Omega |\nabla_X u|^{p(x)-2}\nabla_H u\cdot\nabla_H\phi dx-\lambda\int_\Omega g(x)|u|^{p(x)-2}u\phi dx\leq 0 \tag{8}$$

for all $\phi\in W_0^{1,p(x)}(\Omega)$*, respectively.*

The principal eigenvalue of (5) is denoted by $\lambda_1 := \lambda_1(\Omega)$, and for the solution (u,λ) and $u\neq 0$, we have

$$\lambda_1=\inf_{u\in\overset{\circ}{S}^{1,p(x)}(\Omega)\backslash\{0\}}\frac{\int_\Omega|\nabla_H u|^{p(x)}dx}{\int_\Omega g(x)|u|^{p(x)}dx}.$$

In the case $p(x)=p$(constant), it is well known that $\lambda_1(\Omega)$ given above is the first eigenvalue of p-Laplacian (with $g(x)=1$, $\Omega\subset\mathbb{R}^n$), which must be positive. But this is not true for general $p(x)$ in the sense that λ_1 may be zero [12]. Nevertheless, Fan, Zhang and Zhao in [13] have proved the existence of infinitely many eigenvalues $p(x)$-Laplacian and established sufficient conditions for $\lambda_1(\Omega)>0$. Motivated by [13], we are able to assume the existence of $\lambda_1>0$ in the rest of this paper.

3. Main results: Variable exponent Picone identity and principal eigenvalue

In this section we give the statement the generalised Variable exponent Picone identity, which can be considered as the main result of this paper. The complete proof of the result is not given here since it is almost similar to Theorem 2.3 of our recent preprint [1]. The following is the variable exponent Picone identity.

3.1. Generalised Variable exponent Picone identity.

Theorem 3.1. *Given any domain* Ω *in* $\mathbb{G}$*. Let* $u\geq 0$ *and* $v>0$ *be nonconstant functions differentiable a.e. in* Ω*. Suppose* $p:\bar\Omega\to(0,\infty)$ *is a* C^1*-function for* $p(x)>1$*, and* $f:\mathbb{R}^+\to\mathbb{R}^+$ *is a* C^1*-function satisfying* $f(y)>0$ *and* $f'(y)\geq(p(x)-1)\left(f(y)^{\frac{p(x)-2}{p(x)-1}}\right)$ *for* $y>0$*. Define*

$$R(u,v)=|\nabla_H u|^{p(x)}-\nabla_H\left(\frac{u^{p(x)}}{f(v)}\right)|\nabla_H v|^{p(x)-2}\nabla_H v \tag{9}$$

and

$$\begin{aligned}L(u,v)=&|\nabla_H u|^{p(x)}-\frac{u^{p(x)}\ln u}{f(v)}|\nabla_H v|^{p(x)-2}\nabla_H v\nabla_H p(x)\\&-p(x)\frac{u^{p(x)-1}}{f(v)}|\nabla_H v|^{p(x)-2}\nabla_H v\nabla_H u+\frac{u^{p(x)}f'(v)}{(f(v))^2}|\nabla_H v|^{p(x)}.\end{aligned} \tag{10}$$

Then

(i) $L(u,v)=R(u,v)$.
(ii) $L(u,v)\geq 0$ *if* $\nabla_H v\nabla_H p(x)\equiv 0$*. Moreover* $L(u,v)=0$ *a.e. in* Ω *if and only if* $u=cv$ *a.e. in* Ω *for some constant* $c>0$*.*

Proof. Direct computation gives

$$\nabla_H\left(\frac{u^{p(x)}}{f(v)}\right)=\frac{u^{p(x)}\ln u\nabla_H p(x)+p(x)u^{p(x)-1}\nabla_H u}{f(v)}+\frac{u^{p(x)}f'(v)}{(f(v))^2}\nabla_H v,$$

which when substituted into (9) proves (i) of the theorem. The proof of the rest part of the theorem follows directly from the proof of [1, Theorem 2.3]. □

Remark 1. *Theorem 3.1 generalises many known results:*

(a) In Euclidean setting, if $p(x)=p$ *in (9) and (10) we then recover Picone identity of Bal [7]. If* $f(v)=v^{p(x)-1}$*, we have Picone identity of Allegretto [6] and Feng and Han [14], and that of Allegretto and Huang's [5] for* $p(x)=p$*,* $f(v)=v^{p-1}$*.*
(b) In Subelliptic setting, if p is constant and $f(v)=v^{p-1}$ *in (9) and (10), then we cover Niu, Zhang and Wang [19] (Heisenberg group), Ruzhansky, Sabitbek and Suragan [21] (for general vector fields). Suragan and Yessirkegenov [25] (stratified Lie groups).*

3.2. **Properties of principal eigenvalue of $p(x)$-Laplacian.**

Variable exponent Hardy type inequality.

Proposition 3.2. *Let $\Omega \subset \mathbb{G}$ be an open bounded domain. Suppose that a function $v \in C_0^\infty(\Omega)$ satisfies $\nabla_H v \nabla_H p(x) \equiv 0$ and*

$$\begin{aligned} -\mathscr{L}_p v &= \mu a(x) f(v) && \text{in } \Omega, \\ u &> 0 && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{11}$$

where $f : \mathbb{R}^+ \to \mathbb{R}^+$ is C^1 and satisfies $f'(y) \geq (p(x)-1)\left(f(y)^{\frac{p(x)-2}{p(x)-1}}\right)$, $\mu > 0$ is a constant, $a(x)$ is a positive continuous function. Then there holds

$$\int_\Omega |\nabla_H u|^{p(x)} dx \geq \mu \int_\Omega a(x)|u|^{p(x)} dx$$

for any $0 \leq u \in C_0^1(\Omega)$.

Proof. Since $v > 0$ and solves (11) in Ω, that is, $v \in \mathring{S}^{1,p(x)}(\Omega)$. For a given a $\epsilon > 0$, we set $\phi = \frac{|u|^{p(x)}}{f(v+\epsilon)}$. By the definition of solution (6) we compute

$$\begin{aligned} \mu \int_\Omega a(x) f(v) \frac{|u|^{p(x)}}{f(v+\epsilon)} dx &\leq \int_\Omega |\nabla_H v|^{p-2} \nabla_H v \nabla_X \left(\frac{|u|^{p(x)}}{f(v+\epsilon)}\right) dx \\ &= \int_\Omega |\nabla_H u|^{p(x)} dx - \int_\Omega L(u, v+\epsilon) dx. \end{aligned}$$

Taking the limit as $\epsilon \to 0^+$, applying Fatou's Lemma and Lebesgue dominated convergence theorem respectively on the left hand side and right hand side of the last expression, we obtain

$$0 \leq \int_\Omega |\nabla_H u|^{p(x)} dx - \mu \int_\Omega a(x)|u|^{p(x)} dx$$

since $L(u, v) \geq 0$ a. e. in Ω. This completes the proof. $\square$

Corollary 3.3. *Suppose there exists $\lambda > 0$ and a strictly positive sup-solution of* (5). *Then*

$$\int_\Omega |\nabla_H u|^{p(x)} dx \geq \lambda \int_\Omega g(x)|u|^{p(x)} dx \tag{12}$$

for all $u \in \mathring{S}^{1,p(x)}(\Omega)$.

Proof. Applying Proposition 3.2 by setting $a(x) \equiv g(x)$, $\mu = \lambda$ and $f(v) = |v|^{p(x)-2}v$, then one arrives at the conclusion (11) at once. $\square$

Monotonicity of principal eigenvalue. Here we show the monotonicity of the principal eigenvalue $\lambda_1(\Omega)$ as the function of the set $\Omega \subset \mathbb{G}$.

Proposition 3.4. *Let there exists λ and a strictly positive sup-solution $v \in \mathring{S}^{1,p(x)}(\Omega)$ of* (5). *Then we have*

$$\int_\Omega |\nabla_H u|^{p(x)} dx \geq \lambda \int_\Omega g(x)|u|^{p(x)} dx \tag{13}$$

and

$$\lambda_1(\Omega) \geq \lambda \tag{14}$$

for all $u \in \mathring{S}^1(\Omega)$.

Proof. Suppose there exists $\lambda > 0$, since v is strictly positive sup-solution of (5) in Ω, we have

$$\int_\Omega |\nabla_H v|^{p(x)-2} \nabla_H v \cdot \nabla_H \phi\rangle dx \geq \lambda \int_\Omega g(x)|v|^{p(x)-2} v\phi dx \tag{15}$$

for all $\phi \in \mathring{S}^{1,p(x)}(\Omega)$. For a given small $\epsilon > 0$, setting $\phi = \frac{|u|^{p(x)}}{(v+\epsilon)^{p(x)-1}}$ into (15). Then, following the proof of the Proposition 3.2, we arrive at (13).

Now, let $u_1 \in \mathring{S}^{1,p(x)}(\Omega)$ be the eigenfunction corresponding to the principal eigenvalue $\lambda_1(\Omega)$. We have

$$\int_\Omega |\nabla_H u_1|^{p(x)-2}\nabla_H u_1 \cdot \nabla_H \phi\rangle dx = \lambda_1 \int_\Omega g(x)|u_1|^{p(x)-2}u_1\phi dx \tag{16}$$

for any $\phi \in \mathring{S}^{1,p(x)}(\Omega)$. Choosing $\epsilon > 0$ (small) we can apply the Picone identity as follows (using $\phi = \dfrac{|u_1|^p}{f(v+\epsilon)}$ and (16) with $\phi = u_1$,) we obtain

$$\begin{aligned}
0 &\le \int_\Omega L(u_1, v+\epsilon)dx = \int_\Omega R(u_1, v+\epsilon)dx \\
&= \int_\Omega |\nabla_H u_1|^{p(x)}dx + \int_\Omega \frac{|u_1|^p}{f(v+\epsilon)}\nabla_H(|\nabla_H v|^{p(x)-2}\nabla_H)v dx \\
&\le \lambda_1(\Omega)\int_\Omega g(x)|u_1|^{p(x)}dx - \lambda\int_\Omega g(x)\frac{|u_1|^p}{f(v+\epsilon)}|v|^{p(x)-2}v dx.
\end{aligned}$$

As usual, taking the limit as $\epsilon \to 0^+$, applying Fatou's Lemma and Lebesgue dominated convergence theorem, setting $f(v) = v^{p(x)-1}$, we arrive at

$$0 \le (\lambda_1(\Omega) - \lambda)\int_\Omega g(x)|u_1|^p dx,$$

which implies $\lambda_{1,p}(\Omega) \ge \lambda$.

□

As a corollary to the last proposition, we show strict monotonicity of the principal eigenvalue with respect to domain monotonicity.

Corollary 3.5. *Let $\lambda_1(\Omega) > 0$ be the principal eigenvalue of $\mathscr{L}_p$ on Ω. Suppose $\Omega_1 \subset \Omega_2 \subset \Omega$ and $\Omega_1 \neq \Omega_2$. Then*

$$\lambda_1(\Omega_1) > \lambda_1(\Omega_2)$$

if they both exist.

Proof. Let u_1 and u_2 be positive eigenfunctions corresponding to $\lambda_1(\Omega_1)$ and $\lambda_1(\Omega_2)$, respectively. Clearly with $\phi \in C_0^\infty(\Omega)$, we have by Picone identity that

$$0 \le \int_\Omega L(\phi, u_2)dx = \int_\Omega R(\phi, u_2)dx.$$

Replacing ϕ by u_1 and applying Proposition 3.4 we have

$$\lambda_1(\Omega_1) - \lambda_1(\Omega_2) \ge 0.$$

If we have $\lambda_1(\Omega_1) = \lambda_1(\Omega_2)$, then $L(u_1, u_2) = 0$ a.e. in Ω and thus $u_1 = \alpha u_2$ for some constant $\alpha > 0$. However, this is impossible when $\Omega_1 \subset \Omega_2$ and $\Omega_1 \neq \Omega_2$. □

Uniqueness and simplicity of $\lambda_1(\Omega)$.

Proposition 3.6. *Let there exists $\lambda > 0$ and a strictly positive solution $v \in \mathring{S}^{1,p(x)}(\Omega)$ of* (5). *Then we have*

$$\lambda_1(\Omega) = \lambda.$$

Moreover, let u_1 be the corresponding eigenfunction to $\lambda_1(\Omega)$. Then any other $u \in \mathring{S}^{1,p(x)}(\Omega)$ corresponding to $\lambda_1(\Omega)$ is a constant multiple of u_1.

Proof. Let $u_1 \in \mathring{S}^{1,p(x)}(\Omega)$ be the eigenfunction corresponding to $\lambda_1(\Omega)$ and u be a positive solution of (5). Applying Picone identity by choosing $\epsilon > 0$ (small) as follows:

$$\begin{aligned}
0 &\le \int_\Omega L(u, u_1+\epsilon)dx \\
&= \int_\Omega |\nabla_H u|^{p(x)}dx + \int_\Omega \frac{u^{p(x)}}{f(u_1+\epsilon)}\nabla_H(|\nabla_H u_1|^{p(x)-2}\nabla_H)u_1 dx \\
&= \lambda\int_\Omega g(x)|u|^{p(x)}dx - \lambda_1(\Omega)\int_\Omega g(x)\frac{u^{p(x)}}{(u_1+\epsilon)^{p(x)-1}}|u_1|^{p(x)-2}u_1 dx,
\end{aligned}$$

where we have set $f(u_1+\epsilon) = (u_1+\epsilon)^{p(x)-1}$. Taking the limit as $\epsilon \to 0^+$, applying Fatou's Lemma and Lebesgue dominated convergence theorem, then

$$\lambda_1(\Omega) \leq \lambda.$$

On the other hand by Proposition 3.4, we have

$$\lambda_1(\Omega) \geq \lambda.$$

This therefore implies that $\lambda_1(\Omega) = \lambda$. By this we have proved the uniqueness part.

Now by the hypothesis of the theorem we have for $\phi, \psi \in C_0^\infty(\Omega)$ that

$$\int_\Omega |\nabla_H u|^{p(x)-2}\langle \nabla_H u, \nabla_X \phi\rangle dx = \lambda_{1,p}\int_\Omega g(x)|u|^{p(x)-2}u\phi dx, \tag{17}$$

$$\int_\Omega |\nabla_H u_1|^{p(x)-2}\nabla_H u_1 \cdot \nabla_X \psi dx = \lambda_1 \int_\Omega g(x)|u_1|^{p(x)-2}u_1\psi dx. \tag{18}$$

Taking $\phi = u$ and $\psi = \frac{|u|^p}{(u_1+\epsilon)^{p-1}}$ into (17) and (18), respectively, and sending $\epsilon \to 0^+$, we arrive at

$$\begin{aligned}\int_\Omega |\nabla_H u|^{p(x)}dx &= \lambda_{1,p}\int_\Omega g(x)|u|^{p(x)}dx \\ &= \int_\Omega |\nabla_H u_1|^{p(x)-2}\nabla_H u_1 \nabla_X\Big(\frac{|u|^{p(x)}}{u_1^{p(x)-1}}\Big)dx,\end{aligned}$$

which implies (by choosing $f(u_1) = u_1^{p(x)-2}$)

$$\int_\Omega R(u,u_1)dx = \int_\Omega L(u,u_1)dx = 0$$

and consequently, $\nabla_H(u/v) = 0$, i.e., $u = \alpha u_1$ for some positive constant $\alpha > 0$.

□

The next proposition gives the sign changing nature of any other eigenfunction associated to an eigenvalue other than $\lambda_1(\Omega)$.

Proposition 3.7. *Any eigenfunction v corresponding to an eigenvalue $\lambda \neq \lambda_1(\Omega)$ changes sign.*

Proof. By contradiction we suppose $v > 0$ does not change sign (the case $v \leq 0$ can be handled similarly). Let $\phi > 0$ be an eigenfunction corresponding to $\lambda_1(\Omega)$. Choosing any $\epsilon > 0$ as before, applying Picone identity, we have

$$\begin{aligned}0 &\leq \int_\Omega L(\phi, v+\epsilon)dx \\ &= \int_\Omega \left[|\nabla_H\phi|^{p(x)} - \nabla_H\Big(\frac{\phi^{p(x)}}{f(v+\epsilon)}\Big)|\nabla_H v|^{p(x)-2}\nabla_H v\right]dx \\ &= \int_\Omega |\nabla_H\phi|^{p(x)}dx + \int_\Omega \frac{\phi^{p(x)}}{f(v+\epsilon)}\mathscr{L}_p v dx.\end{aligned}$$

Note that $\frac{\phi^{p(x)}}{(v+\epsilon)^{p(x)-1}}$ is admissible in the weak formulation of (5), so we arrive at

$$0 \leq \lambda_1(\Omega)\int_\Omega g(x)|\phi|^{p(x)}dx - \lambda\int_\Omega \frac{\phi^{p(x)}}{f(v+\epsilon)}g(x)|v|^{p(x)-2}v dx.$$

Setting $f(v+\epsilon) = (v+\epsilon)^{p(x)-1}$ and letting $\epsilon \to 0^+$ in the last inequality as usual we obtain

$$0 \leq (\lambda_1 - \lambda)\int_\Omega g(x)\phi^{p(x)}dx,$$

which is a contradiction. Thus v must change sign.

□

4. Conclusion

A generalised Picone type identity with variable exponent (Theorem 3.1) was presented on general stratified Lie groups. It was also shown that this theorem covers many known results (see Remark 1). The uniqueness, simplicity, monotonicity and isolation of the first Dirichlet eigenvalue of horizontal $p(x)$-sub-Laplacian in variable Lebesgue spaces on the stratified Lie groups were then proved as applications of the generalised Picone type identity. These results were discussed in Section 3. Meanwhile, some basics of the general stratified groups (homogeneous Carnot group) and some concepts from the theory of variable Lebesgue spaces were presented in Section 2.

Acknowledgments

The author gratefully acknowledges the research supports of IMU-Simons African Fellowship Grant and EMS-Simons for African program. This research was carried out during the author's research visit to Ghent Analysis and PDE Centre, Ghent University, Belgium.

References

[1] Abolarinwa A., (2022), Nonlinear variable exponent Picone identity and $p(x)$-sub-Laplacian first eigenvalue for general vector fields, arXiv:2209.05642v1 [math.AP].
[2] Acerbi E., Mingione G., and Seregin G. A., Regularity results for parabolic systems related to a class of non-Newtonian fluids, Ann. Inst. H. Poincaré Anal. Non Lineáire, 21(1) 25–60.
[3] Aboulaich R., Meskine D., and Souissi A., (2008), New diffusion models in image processing, Comput. Math. Appl., 56(4) 874–882.
[4] Allegretto W., (1986), Positive solutions and spectral properties of weakly coupled elliptic systems, J. Math. Anal. Appl., 120, 723–729.
[5] Allegretto W., and Huang Y. X., (1998), A Picone's identity for the p-Laplacian and applications, Nonl. Anal. Th. Meth. Appl., 32(7)819–830.
[6] Allegretto W., (2007), Form estimates for the $p(x)$-Laplacian, Proc. Amer. Math. Soc., 135(7) 2177–2185.
[7] Bal K. , (2013), Generalized Picone's identity and its applications, Electron. J. Diff. Equ., 2013(243), pp. 1–6.
[8] Cruiz-Uribe D, Fiorenza A., Ruzhansky M., and Wirth J., (2014), Variable Lebesgue spaces and Hyperbolic systems, Springer Basel, 2014.
[9] Deng S-G., (2008), Eigenvalues of $p(x)$-Laplacian Dirichlet Steklov problem, J. Math. Anal. Appl., 339(2) 925–937.
[10] Diening L., Harjulehto P., Hästö P., Lê, Út V., and Ruzicka M., (2011), Lebesgue and Sobolev spaces with variable exponents, Springer 2011.
[11] Fan X., (2007), Eigenvalues of the $p(x)$-Laplacian Neumann problems, Nonlinear Anal., 67, 2982–2992.
[12] Fan X., and Zhao D., (2001), On the spaces $L^{p(x)}(\Omega)$ and $W^{m,p(x)}(\Omega)$, J. Math. Anal. Appl., 263(2) 424–446.
[13] Fan X., Zhang Q., and Zhao D., (2005), Eigenvalues of $p(x)$-Laplacian Dirichlet problem, J. Math. Anal. Appl., 302(2), 306–317.
[14] Feng T., and Han J., (2019), A new variable exponent Picone identity and applications, Math. Ineq. Appl., 22(1) 65–75.
[15] Franzina G., and Lindqvist P., (2013), An eigenvalue problem with variable exponents, Nonlinear Anal., 85, 1–16.
[16] Harjulehto P., Hästö P., Lê, Út V., and Nuortio M., (2010), Overview of differential equations with non-standard growth, Nonlinear Anal., 72 4551–4574.
[17] Hörmander H., (1967), Hypoelliptic second-order differential equations, Acta Math., 43, 147–171.
[18] Mihailescu M., and Varga C., (2011), Multiplicity results for some elliptic problems with nonlinear boundary conditions involving variable exponents, Comput. Math. Appl., 62(9) 3464–3471.
[19] Niu P., Zhang H., and Wang Y., (2001), Hardy Type and Rellich Type Inequalities on the Heisenberg Group, Proc. Amer. Math. Soc., 129(12) 3623–3630.
[20] Ruzhansky M., and Suragan D., (2019), Hardy inequalities on homogeneous groups, Progress in Math. Vol. 327, Birkhäuser, 588 pp, (2019).
[21] Ruzhansky M., Sabitbek B., and Suragan D., (2021), Principal frequency of p-versions of sub-Laplacians for general vector fields, Z. Anal. Anwend., 40, 97–109.
[22] Ruzhansky M., Sabitbek B., and Suragan D., (2019), Weighted anisotropic Hardy and Rellich type inequalities for general vector fields, Nonlinear Differ. Eqn. Appl., 26 (13) 2019.
[23] Ruzhansky M., Sabitbek B., and Suragan D., (2019), Weighted L^p-Hardy and L^p-Rellich inequalities with boundary terms on stratified Lie groups, Rev. Mat. Complut., 32(1) 19–35.
[24] Ruzicka M., (2000) Electrorheological Fluids Modeling and Mathematical Theory, Springer-Verlag, Berlin, 2000.
[25] Surugan D., and Yessirkegenov N., (2021), Generalised nonlinear Picone identities for p-sub-Laplacians and p-biharmonic operators and applications, Adv. Op. Th., 6(53) 1–17.
[26] Yoshida N., (2013), Picone identity for quasilinear elliptic equations with $p(x)$-Laplacians and Sturmianian comparison theory, Appl. Math. Comput., 225 (1) 79–91.

[27] Tyagi J., (2013), A nonlinear picone's identity and its applications, Appl. Math. Letters, 26, 624–626.

Department of Mathematics, University of Lagos, Akoka, Lagos State, Nigeria

Email(s): a.abolarinwa1@gmail.com, aabolarinwa@unilag.edu.ng

A New Approach to Fixed Point Theory via Integral Type Mappings on Orthogonal Metric Space

Özlem Acar

Abstract. This study is devoted to investigate the problem whether the existence and uniqueness of integral type contraction mappings on orthogonal metric space.

2020 Mathematics Subject Classifications: Primary 54H25; Secondary, 47H10

Keywords: Fixed point, integral type mapping, orthogonal metric space.

1. Introduction and Preliminaries

The study of fixed point theorems of mappings satisfying contractive conditions of integral type has been a very interesting field of research activity after the establishment of a theorem by A. Branciari [1]. He appointed good integral prescription of the Banach contraction principle such that :

Theorem 1.1. *Let T be a mapping from a complete metric space (M,ρ) into itself , $c\in]0,1[$, and let $T : M \to M$ be a mapping such that for each $x, y \in M$*

$$\int_0^{\rho(Tx,Ty)} \gamma(s)ds \le c \int_0^{\rho(x,y)} \gamma(s)ds$$

$\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ is Lebesque integrable mapping which is summable on each compact subset of $\mathbb{R}^+$,nonnegative and $\int_0^{\epsilon}\gamma(s)ds > 0$ for each $\varepsilon > 0$ then T has a unique fixed point $a \in M$ such that for each $x \in M$, $lim_{n\to+\infty}T^n x = a$.

Since then, many authors have established fixed point theorems for several classes of contractive mappings of integral type.([4, 7, 8]) Especially Liu et. al.[9] extended the result of Brianciari in many different ways.

Recently, for the first time, Gordji et al.[5] expand the literature on metric space by introduced the concept of orthogonality, established the fixed point result. There are several uses for this novel idea of an orthogonal set, as well as numerous forms of orthogonality. Further for more information, we refers the reader to ([2, 3, 6, 10, 11]).

This study is devoted to investigate the problem whether the existence and uniqueness of integral type contraction mappings on orthogonal metric space. To do this we first recall some fundamental definitions and notations of corresponding mappings and space.

Throughout this paper, we assume that $\mathbb{R}^+ = [0,\infty)$ and

$$\Phi_1 \;:\; \left\{ \begin{array}{l} \gamma : \gamma : \mathbb{R}^+ \to \mathbb{R}^+ \text{ is Lebesque integrable,summable on each compact subset of } \mathbb{R}^+ \\ \text{and } \int_0^{\epsilon}\gamma(s)ds > 0 \text{ for each } \epsilon > 0 \end{array} \right\},$$

$$\Phi_2 \;:\; \left\{ \begin{array}{l} \beta : \beta : \mathbb{R}^+ \to \mathbb{R}^+ \text{ satisfies that} \\ \liminf_{n\to\infty}\ \beta(a_n) > 0 \Longleftrightarrow \liminf_{n\to\infty} a_n > 0 \text{ for each } (a_n)_{n\in\mathbb{N}} \subset \mathbb{R}^+ \end{array} \right\},$$

$$\Phi_3 \;:\; \left\{\alpha : \alpha : \mathbb{R}^+ \to \mathbb{R}^+ \text{ is nondecreasing continuous and } \alpha(t) = 0 \iff t = 0\right\},$$

$$\Phi_4 \;:\; \left\{\Delta : \Delta : \mathbb{R}^+ \to [0,1) \text{ satisfy that } \limsup_{s\to t}\Delta(s) < 1 \text{ for each } t > 0\right\}.$$

Lemma 1.2 ([9])**.** *Let $\gamma \in \Phi_1$ and $\{a_n\}_{n\in\mathbb{N}}$ be nonnegative sequence with $\lim_{n\to\infty} a_n = a$. Then*

$$\lim_{n\to\infty}\int_0^{a_n}\gamma(s)ds = \int_0^{a}\gamma(s)ds.$$

Lemma 1.3 ([9]). *Let $\gamma \in \Phi_1$ and $\{a_n\}_{n\in\mathbb{N}}$ be nonnegative sequence. Then*

$$\lim_{n\to\infty} \int_0^{a_n} \gamma(s)ds = 0 \Longleftrightarrow \lim_{n\to\infty} a_n = 0.$$

Lemma 1.4 ([9]). *Let $\beta \in \Phi_2$. Then $\beta(t) > 0 \Longleftrightarrow t > 0$.*

Definition 1.5 ([5]). *Let M be a non-empty set and $\curlywedge$be a binary relation defined on M. If binary relation $\curlywedge$ fulfils the following criteria:*

$$\exists\varsigma_0 (\forall\omega\in M, \omega\curlywedge\varsigma_0) \text{ or } (\forall\vartheta \in M, \varsigma_0 \curlywedge \omega),$$

then pair, $(M, \curlywedge)$ known as an orthogonal set. The element ς_0 is called an orthogonal element. We denote this O-set or orthogonal set by $(M, \curlywedge)$.

Definition 1.6 ([5]). *Let $(M, \curlywedge)$ be an orthogonal set (O-set). Any two elements $\varsigma, \omega \in M$ such that $\varsigma \curlywedge \omega$, then $\varsigma, \omega \in M$ are said to be orthogonally related.*

Definition 1.7 ([5]). *A sequence $\{\varsigma_n\}$ is called an orthogonal sequence (briefly O-sequence) if*

$$(\forall n \in \mathbb{N}, \varsigma_n \curlywedge \varsigma_{n+1}) \text{ or } (\forall n \in \mathbb{N}, \varsigma_{n+1} \curlywedge \varsigma_n).$$

Similarly, a Cauchy sequence $\{\varsigma_n\}$ is said to be a orthogonally Cauchy sequence if

$$(\forall n \in \mathbb{N}, \varsigma_n \curlywedge \varsigma_{n+1}) \text{ or } (\forall n \in \mathbb{N}, \varsigma_{n+1} \curlywedge \varsigma_n).$$

Definition 1.8 ([5]). *Let $(M, \curlywedge)$ be an orthogonal set and ρ be a metric on M. Then $(M, \curlywedge, \rho)$ is called an orthogonal metric space (shortly O-metric space).*

Definition 1.9 ([5]). *Let $(M, \curlywedge, \rho)$ be an orthogonal metric space. Then M is said to be a O-complete if every Cauchy O-sequence is converges in M.*

Definition 1.10 ([5]). *Let $(M, \curlywedge, \rho)$ be an orthogonal metric space. A function $f : M \to M$ is said to be orthogonally continuous ($\curlywedge$-continuous) at ς if for each O-sequence $\{\varsigma_n\}$converging to ς implies $f(\varsigma_n) \to f(\varsigma)$ as $n \to \infty$. Also f is $\curlywedge$-continuous on M if f is $\curlywedge$-continuous at every $\varsigma \in M$.*

Definition 1.11 ([5]). *Let a pair $(M, \curlywedge)$ be an O-set, where $M(\neq \phi)$ be a non-empty set and $\curlywedge$be a binary relation on set M. A mapping $f : M \to M$ is said to be $\curlywedge$-perserving if $f(\varsigma) \curlywedge f(\omega)$ whenever $\varsigma \curlywedge \omega$ and weakly $\curlywedge$-preserving if $f(\varsigma) \curlywedge f(\omega)$ or $f(\omega) \curlywedge f(\varsigma)$ whenever $\varsigma \curlywedge \omega$.*

Definition 1.12 ([11]). *We say that an O-set is a transitive orthogonal set if $\curlywedge$ is transitive.*

Definition 1.13 ([11]). *Let $(M, \curlywedge)$ be an $O-$set. A path of length k in $\curlywedge$ from x to y is a finite sequence $\{z_0, z_1, ..., z_k\} \subset M$ such that*

$$z_0 = x, z_k = y, z_i \curlywedge z_{i+1} \text{ or } z_{i+1} \curlywedge z_i$$

for all $i = 0, 1, ..., k-1$ and also $\lambda(x, y, \curlywedge)$ be denoted as all path of length k in $\curlywedge$ from x to y .

2. Main Results

Definition 2.1. *Let $(M, \curlywedge, \rho)$ be an orthogonal metric space. A mapping $T : X \to X$ is called an orthogonal integral type(B) mapping such that $\forall x, y \in X$ with $x \curlywedge y$*

$$\alpha\left(\int_0^{\rho(Tx,Ty)} \gamma(s)ds\right) \quad \leq \quad \Delta(\rho(x, y))\alpha\left(\int_0^{\rho(x,y)} \gamma(s)ds\right),$$

(1)

where $\alpha \in \Phi_3, \gamma \in \Phi_1, \Delta \in \Phi_4$.

Theorem 2.2. *Let $(M, \perp, \rho)$ be an $0-$complete orthogonal metric space, a_0 is an orthogonal element of M and T be a self mapping on M such that:*
$(i)(M, \perp)$ is a transitive orthogonal set,
$(ii)T$ is $\perp-$preserving,
$(iii)T$ is an orthogonal integral type(B) mapping,
$(iv)T$ is $\perp-$continuous,
Then, T has a unique fixed point in M.

Proof. From the definition of the orthogonality, we have $a_0 \perp T(a_0)$ or $T(a_0) \perp a_0$. Let

$$a_1 := Ta_0, a_2 := Ta_1 = T^2 a_0, \cdots, a_n := Ta_{n-1} = T^n a_0$$

for all $n \in \mathbb{N} \cup \{0\}$.If $a_{n^*} = a_{n^*+1}$ then for some $n^* \in \mathbb{N} \cup \{0\}$, then a_{n^*} is a fixed point of T . So, we suppose that $a_n \neq a_{n+1}$ for all $n \in \mathbb{N} \cup \{0\}$. Thus, we get $\rho(a_n, a_{n+1}) > 0$ for all $n \in \mathbb{N} \cup \{0\}$. Using $\perp-$preserving of T, we obtain

$$a_n \perp a_{n+1} \text{ or } a_{n+1} \perp a_n.$$

Thus $\{a_n\}$ is an $0-$sequence. Set $\varsigma_n = \rho(T^n a_0, T^{n+1} a_0)$ and show that

$$\varsigma_n \leq \varsigma_{n-1}, \forall n \in \mathbb{N}. \tag{2}$$

Suppose that (2) does not hold. It follows that there exists some $n_0 \in \mathbb{N}$ satisfying

$$\varsigma_{n_0} > \varsigma_{n_0-1}. \tag{3}$$

Using (1), (3) and $\alpha \in \Phi_3, \gamma \in \Phi_1, \Delta \in \Phi_4$, we get that

$$\int_0^{\varsigma_{n_0}} \gamma(s)ds > 0$$

and

$$\begin{aligned}
& \alpha\left(\int_0^{\varsigma_{n_0-1}} \gamma(s)ds\right) \\
\leq\ & \alpha\left(\int_0^{\varsigma_{n_0}} \gamma(s)ds\right) \\
=\ & \alpha\left(\int_0^{\rho(T^{n_0}a_0, T^{n_0+1}a_0)} \gamma(s)ds\right) \\
\leq\ & \Delta\left(\rho(T^{n_0-1}a_0, T^{n_0}a_0)\right) \alpha\left(\int_0^{\rho(T^{n_0-1}a_0, T^{n_0}a_0)} \gamma(s)ds\right) \\
=\ & \Delta\left(\varsigma_{n_0-1}\right) \alpha\left(\int_0^{\varsigma_{n_0-1}} \gamma(s)ds\right) \\
\leq\ & \alpha\left(\int_0^{\varsigma_{n_0-1}} \gamma(s)ds\right)
\end{aligned}$$

which is a contradiction, and hence (3) does not hold. Consequently, (2) is true. Notice that the nonnegative sequence $\{\varsigma_n\}_{n\in\mathbb{N}}$ is nonincreasing, which implies that there exists a constant $c \geq 0$ with $\lim_{n\to\infty} \varsigma_n = c$. Suppose that $c > 0$. In light of (1), we infer that

$$\begin{aligned}
& \alpha\left(\int_0^{\varsigma_n} \gamma(s)ds\right) \\
=\ & \alpha\left(\int_0^{\rho(T^n a_0, T^{n+1}a_0)} \gamma(s)ds\right) \\
\leq\ & \Delta\left(\rho(T^{n-1}a_0, T^n a_0)\right) \alpha\left(\int_0^{\rho(T^{n-1}a_0, T^n a_0)} \gamma(s)ds\right) \\
=\ & \Delta\left(\varsigma_{n-1}\right) \alpha\left(\int_0^{\varsigma_{n-1}} \gamma(s)ds\right), \forall n \in \mathbb{N}. \qquad (4)
\end{aligned}$$

Taking upper limit (4) and using Lemma1.2 and $\alpha \in \Phi_3, \gamma \in \Phi_1, \Delta \in \Phi_4$, we know that

$$\begin{aligned}
&\alpha\left(\int_0^c \gamma(s)ds\right)\\
=\ & \limsup_{n\to\infty} \alpha\left(\int_0^{\varsigma_n} \gamma(s)ds\right)\\
\le\ & \limsup_{n\to\infty} \left[\Delta\left(\varsigma_{n-1}\right)\alpha\left(\int_0^{\varsigma_{n-1}} \gamma(s)ds\right)\right]\\
\le\ & \limsup_{n\to\infty} \Delta\left(\varsigma_{n-1}\right).\limsup_{n\to\infty} \alpha\left(\int_0^{\varsigma_{n-1}} \gamma(s)ds\right)\\
<\ & \alpha\left(\int_0^c \gamma(s)ds\right)
\end{aligned}$$

which is a contradiction, so $c = 0$.

Let's show a $\{T^n a_0\}$ is a $O-$Cauchy sequence. Suppose that $\{T^n a_0\}$ is not a $O-$Cauchy sequence which means that there is a constant $\epsilon > 0$ such that for each positive integer k, there are positive integer $m(k)$ and $n(k)$ with $m(k) > n(k) < k$ satisfying

$$\rho(T^{m(k)}a_0, T^{n(k)}a_0) > \epsilon. \tag{5}$$

For each positive integer k, let $m(k)$ denote the least integer exceeding $n(k)$ and satisfying (5). It follows that

$$\rho(T^{m(k)}a_0, T^{n(k)}a_0) > \epsilon$$

and

$$\rho(T^{m(k)-1}a_0, T^{n(k)}a_0) \le \epsilon, \forall k \in \mathbb{N}. \tag{6}$$

Note that

$$\rho(T^{m(k)}a_0, T^{n(k)}a_0) \le \rho(T^{m(k)-1}a_0, T^{n(k)}a_0) + \rho(T^{m(k)-1}a_0, T^{m(k)}a_0), \forall k \in \mathbb{N}.$$

Hence, for all $k \in \mathbb{N}$

$$\begin{aligned}
\left|\rho(T^{m(k)}a_0, T^{n(k)+1}a_0) - \rho(T^{m(k)}a_0, T^{n(k)}a_0)\right| &\le \varsigma_{n(k)},\\
\left|\rho(T^{m(k)+1}a_0, T^{n(k)+1}a_0) - \rho(T^{m(k)}a_0, T^{n(k)+1}a_0)\right| &\le \varsigma_{m(k)},\\
\left|\rho(T^{m(k)+1}a_0, T^{n(k)+1}a_0) - \rho(T^{m(k)+1}a_0, T^{n(k)+2}a_0)\right| &\le \varsigma_{n(k)+1}. \qquad (7)
\end{aligned}$$

From (6) and (7), we obtain

$$\begin{aligned}
\epsilon &= \lim_{k\to\infty} \rho(T^{m(k)}a_0, T^{n(k)}a_0) = \lim_{k\to\infty} \rho(T^{m(k)}a_0, T^{n(k)+1}a_0)\\
&= \lim_{k\to\infty} \rho(T^{m(k)+1}a_0, T^{n(k)+1}a_0) = \lim_{k\to\infty} \rho(T^{m(k)+1}a_0, T^{n(k)+2}a_0). \qquad (8)
\end{aligned}$$

By means of (1), (8), Lemma1.2 and $\alpha \in \Phi_3, \gamma \in \Phi_1, \Delta \in \Phi_4$, we get that

$$\begin{aligned}
&\alpha\left(\int_0^{\epsilon} \gamma(s)ds\right)\\
=\ & \limsup_{k\to\infty} \alpha\left(\int_0^{\rho(T^{m(k)+1}a_0, T^{n(k)+2}a_0)} \gamma(s)ds\right)\\
\le\ & \limsup_{k\to\infty} \left[\Delta\left(\rho(T^{m(k)}a_0, T^{n(k)+1}a_0)\right)\alpha\left(\int_0^{\rho(T^{m(k)}a_0, T^{n(k)+1}a_0)} \gamma(s)ds\right)\right]\\
\le\ & \limsup_{k\to\infty} \Delta\left(\rho(T^{m(k)}a_0, T^{n(k)+1}a_0)\right).\limsup_{k\to\infty} \alpha\left(\int_0^{\rho(T^{m(k)}a_0, T^{n(k)+1}a_0)} \gamma(s)ds\right)\\
<\ & \alpha\left(\int_0^{\epsilon} \gamma(s)ds\right),
\end{aligned}$$

which is a contradiction. Thus $\{T^n a_0\}$ is a $O-$Cauchy sequence. Since M is $O-$complete, then there exists $z^* \in M$ such that $a_n \to z^*$. Since orthogonal continuity of T implies that $Ta_n \to Tz^*$, then

$$Tz^* = T(\lim_{n\to\infty} a_n) = \lim_{n\to\infty} Ta_n = \lim_{n\to\infty} a_{n+1} = z^*$$

so z^* is a fixed point of T.

Now, we can show the uniqueness of the fixed point. Suppose that there exists two distinct fixed point z^* and w^*. Since $\lambda(x, y, \curlywedge)$ is nonempty for all $x, y \in M$, there exists a path $\{z_0, z_1, ..., z_k\}$ of some finite lenght k in $\curlywedge$ from z^* to w^* such that

$$z_0 = z^*, z_k = w^*, z_i \curlywedge z_{i+1} \text{ or } z_{i+1} \curlywedge z_i.$$

Since $(M, \curlywedge)$ transitive orthogonal set, we get $z^* \curlywedge w^*$ or $w^* \curlywedge z^*$. Then, from (1),

$$\begin{aligned}
\alpha\left(\int_0^{\rho(z^*,w^*)} \gamma(s)ds\right) &= \alpha\left(\int_0^{\rho(Tz^*,Tw^*)} \gamma(s)ds\right) \\
&\leq \Delta\left(\rho(z^*,w^*)\right)\alpha\left(\int_0^{\rho(z^*,w^*)} \gamma(s)ds\right) \\
&< \alpha\left(\int_0^{\rho(z^*,w^*)} \gamma(s)ds\right)
\end{aligned}$$

which is a contradiction so z^* is a unique fixed point of T. □

References

[1] Branciari, A: A fixed point theorem for mappings satisfying a general contractive condition of integral type. Int.J. Math. Math. Sci. 29, 531-536 (2002). doi:10.1155/S0161171202007524.

[2] Acar, Ö., Özkapu, A.S., (2022). Multivalued rational type $F-$contraction on orthogonal metric space, Mathematical Foundations of Computing, accepted.

[3] Acar, Ö., Erdoğan, E., Some fixed point results for almost contraction on orthogonal metric space, Creat. Math. Inform., 31 (2022), No. 2,147 - 153.

[4] Altun, I, Türkoglu, D, Rhoades, BE: Fixed points of weakly compatible maps satisfying a general contractive of integral type. Fixed Point Theory Appl. 2007, Article ID 17301 (2007). doi:10.1155/2007/17301.

[5] Gordji M.E., Rameani M., De La SEN M.,Cho Y.J., On orthogonal sets and Banach fixed point theorem. Fixed Point Theory 18, (2017),569–578.

[6] Gordji M.E., Habibi, H., Fixed point theory in generalized orthogonal metric space, Journal of Linear and Topological Algebra (JLTA), 6(3), 251-260.

[7] Jachymski, J: Remarks on contractive conditions of integral type. Nonlinear Anal. 71, 1073-1081 (2009). doi:10.1016/j.na.2008.11.046

[8] Liu, Z., Li, X., Kang, S. M., Cho, SY: Fixed point theorems for mappings satisfying contractive conditions of integral type and applications. Fixed Point Theory Appl. 2011, Article ID 64 (2011). doi:10.1186/1687-1812-2011-64.

[9] Liu,Z., Li, J., Kang, S. M., Fixed point theorems of contractive mappings of integral type, Fixed Point Theory and Applications 2013, 2013:300.

[10] Sawangsup,K., Sintunavarat, W., Cho,Y. J. , Fixed point theorems for orthogonal $F-$contraction mappings on $O-$complete metric spaces, J. Fixed Point Theorey Appl.,(2020)22:10.

[11] Sawangsup,K., Sintunavarat, W., Fixed Point Results for Orthogonal $Z-$Contraction Mappings in $O-$Complete Metric Spaces, International Journal of Applied Physics and Mathematics,Volume 10, Number 1, 2020.

Department of Mathematics, Faculty of Science, Selçuk University, 42003, Konya, Turkey

Email(s): acarozlem@ymail.com

Durrmeyer Type Exponential Sampling Series in Weighted Spaces

Tuncer Acar

ABSTRACT. In this paper, we present a quantitative Voronovskaya type theorem for exponential sampling Durrmeyer operators in logarithmic weighted space of functions in terms of weighted logarithmic modulus of continuity. Such a result allows us to determine a rate of pointwise convergence of the family of exponential sampling Durrmeyer operators and an upper bound for the error of this convergence.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 41A35, 30D10, 94A20, 41A25.

KEYWORDS: Sampling Durrmeyer operators, Weighted approximation, Quantitative Voronovskaya theorem.

1. INTRODUCTION

Very recently, Bardaro et. al [14] introduced a new family of sampling type operators, called exponetial sampling operators, using the continuous function φ satisfying the certain assumptions of approximate identities instead of lin $_c$ function. The generalization of the exponential sampling series is given as

$$(E_w^{\varphi} f)(x) := \sum_{k\in\mathbb{Z}} f\left(e^{k/w}\right)\varphi\left(e^{-k}x^{w}\right), \quad w>0, x\in\mathbb{R}^+, \tag{1}$$

where $f:\mathbb{R}^+\to\mathbb{C}$ is any function for which the series is absolutely convergent. For the generalized exponential sampling series (1), they gave some properties of convergence for log-uniformly continuous and bounded functions on $\mathbb{R}^+$. Also, the operators (1) were studied in Mellin-Lebesgue spaces (see [17]). The series (1) enable to reconstruct a signal which are exponentially spaces by using its values at the nodes $\left(e^{k/w}\right)$ but in practice it is difficult to do. To solve this problem, in [7], Angamuthu and Bajpeyi introduced the Kantorovich form of (1) by using Steklov mean values $w\int_{k/w}^{\frac{k+1}{w}} f(e^u)\,du$ instead of sample values of the form $f\left(e^{k/w}\right)$ defined as

$$(K_w^{\chi} f)(x) := \sum_{k\in\mathbb{Z}} \chi\left(e^{-k}x^{w}\right) w\int_{k/w}^{\frac{k+1}{w}} f(e^u)\,du \tag{2}$$

where $\chi:\mathbb{R}^+\to\mathbb{R}$ is a suitable continuous function and $f:\mathbb{R}^+\to\mathbb{R}$ is locally integrable function for which the series is absolutely convergent. Recently, it has been extensively studied forms of (1) and its Kantorovich type (2), we refer the readers to [2, 8, 9, 10, 11, 12, 35].

In present paper, we investigate the properties of convergence of the generalization of (2) known as exponential sampling Durrmeyer series, in logarithmic weighted spaces first considered in [9]. By replacing Steklov mean (2) with the Mellin convolution operator, exponential sampling Durrmeyer series is defined by

$$(S_w^{\varphi,\chi} f)(x) := \sum_{k\in\mathbb{Z}} \varphi\left(e^{-k}x^{w}\right) w\int_0^{\infty} \chi\left(e^{-k}u^{w}\right) f(u)\frac{du}{u}, \quad w>0, x\in\mathbb{R}^+,$$

where $f:\mathbb{R}^+\to\mathbb{R}$ is integrable function such that for which the above series is absolutely convergent (see [16]).

In this paper, we present a quantitative form of Voronovskaja type theorem using the Mellin-Taylor formula [15, 36] via Mellin derivatives [22].

Here, point out that weighted approximations of generalized sampling series both univariate and bivariate case and its Kantorovich, Durrmeyer forms was studied in [3, 4, 5, 6].

Throughout the paper we shall use the following standart notions: Let $\mathbb{N}, \mathbb{N}_0, \mathbb{R}$ and $\mathbb{R}^+$be the sets of positive integers, nonnegative integers, all real and postive real numbers, respectively.

Let $C(\mathbb{R}^+)$ be the space of all continuous functions defined on $\mathbb{R}^+ \cdot C_B(\mathbb{R}^+)$ is the space of all bounded functions $f \in C(\mathbb{R}^+)$. We say that a function $f \in C_B(\mathbb{R}^+)$ is log-uniformly continuous on $\mathbb{R}^+$, if for any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that $|f(t) - f(x)| < \varepsilon$ whenever $|\log t - \log x| \leq \delta$ for any $t, x \in \mathbb{R}^+$.

For a function $f \in C(\mathbb{R}^+)$ and $n \in \mathbb{N}$, we will say that f belongs to $C^{(n)}(\mathbb{R}^+)$ locally at a point $x \in \mathbb{R}^+$ if there is a neighborhood U of x such that f is $(n-1)$-times continuously differentiable function in U and $f^{(n)}(x)$ exists. By $L_p(\mathbb{R}^+)$ for $1 \leq p < \infty$, we shall denote the space of all Lebesgue measurable and p-integrable complex valued functions on $\mathbb{R}^+$ equipped with the usual norm $\|f\|_p$. Now for $c \in \mathbb{R}$ let us consider the space

$$X_c = \left\{ f : \mathbb{R}^+ \to \mathbb{C} : f(\cdot)(\cdot)^{c-1} \in L_1\left(\mathbb{R}^+\right) \right\}$$

endowed with the norm

$$\|f\|_{X_C} = \left\| f(\cdot)(\cdot)^{c-1} \right\|_1 = \int_0^\infty |f(u)| u^{c-1} du.$$

Let $\varphi \in \mathbb{R}^+ \to \mathbb{R}$ be a continuous function such that the following assumptions are satisfied:
$\varphi 1)$ for every $u \in \mathbb{R}^+$,

$$\sum_{k \in \mathbb{Z}} \varphi(e^{-k} u) = 1,$$

$\varphi 2)$ we have that

$$M_0(\varphi) = \sup_{u \in \mathbb{R}^+} \sum_{k \in \mathbb{Z}} \left| \varphi(e^{-k} u) \right| < \infty$$

and
$\varphi 3)$ for some $v \in \mathbb{N}$,

$$\lim_{r \to \infty} \sum_{\left| k - logu \right| > r} \left| \varphi(e^{-k} u) \right| \left| k - logu \right|^v = 0$$

uniformly with respect to $u \in \mathbb{R}^+$.
We denote by Φ the class of all functions φ satisfying the above assumptions.
Now, let $\chi : \mathbb{R}^+ \to \mathbb{R}$ be a function with the following conditions
$\chi 1)$

$$\int_0^\infty \chi(t) \frac{dt}{t} = 1,$$

$\chi 2)$

$$\tilde{M_0}(\chi) = \int_0^\infty \left| \chi(t) \right| \frac{dt}{t} < \infty$$

and let us denote by Ψ the class of all functions χ satisfynig the above assumptions.
Let $v \in \mathbb{N}_0$. For $x \in \mathbb{R}^+$, we define the algebraic moments of order v of $\varphi \in \Phi$ and $\chi \in \Psi$ as

$$m_v(\varphi, x) := \sum_{k \in \mathbb{Z}} \varphi(e^{-k} x)(k - logx)^v$$

and

$$\tilde{m_v}(\chi) := \int_0^\infty \chi(t) \log^v t \frac{dt}{t}.$$

The absolute moments of order v of $\varphi \in \Phi$ and $\chi \in \Psi$ are defined as

$$M_v(\varphi, x) := \sum_{k \in \mathbb{Z}} \varphi(e^{-k} x)(k - logx)^v$$

and

$$\tilde{M_v}(\chi) := \int_0^\infty \left| \chi(t) \right| \left| \log t \right|^v \frac{dt}{t}.$$

Finally we put $M_v(\varphi) = \sup_{u \in \mathbb{R}^+} M_v(\varphi, x)$.

Lemma 1.1. *Note that, for $\varphi \in \Phi$, if $\mu, v \in \mathbb{N}$ with $\mu < v$, then $M_v(\varphi) < \infty$ implies $M_\mu(\varphi) < \infty$. The same is for the absolute moments of $\chi \in \Psi$ (see [16]).*

Finally, we recall the logarithmic weighted spaces considered in [9]. Considering the weight function $f:\mathbb{R}^+\to\mathbb{R}$, $1+\log^2 x$, the logarithmic weighted space of continuous functions and its natural subspaces are given by

$$B_2\left(\mathbb{R}^+\right):=\left\{f:\mathbb{R}^+\to\mathbb{R}:\exists M>0 \text{ such that } \frac{|f(x)|}{1+\log^2 x}\leq M \text{ for every } x\in\mathbb{R}^+\right\},$$
$$C_2\left(\mathbb{R}^+\right):=C\left(\mathbb{R}^+\right)\cap B_2\left(\mathbb{R}^+\right)$$
$$C_2^*\left(\mathbb{R}^+\right):=\left\{f\in C_2\left(\mathbb{R}^+\right):\lim_{x\to\infty}\frac{|f(x)|}{1+\log^2 x}\in\mathbb{R}^+\right\},$$

respectively. Here, the linear space $B_2\left(\mathbb{R}^+\right)$is normed space with the norm

$$\|f\|_L:=\sup_{x\in\mathbb{R}^+}\frac{|f(x)|}{1+\log^2 x}.$$

Let $\varphi\in\Phi$ and $\chi\in\Psi$. For any $w>0$, and $f:\mathbb{R}^+\to\mathbb{R}$, the exponential sampling Durrmeyer series is defined by

$$\left(S_w^{\varphi,\chi}f\right)(x):=\sum_{k\in\mathbb{Z}}\varphi\left(e^{-k}x^w\right)w\int_0^\infty\chi\left(e^{-k}u^w\right)f(u)\frac{du}{u} \tag{3}$$

for $x\in\mathbb{R}^+$and for any function $f\in\operatorname{dom}S_w^{\varphi,\chi}$, being $\operatorname{dom}S_w^{\varphi},\chi$ the set of all functions f for which the series is absolutely convergent at every x (see [16]).

2. Quantitative Voronovskaya theorem

Here, we study a quantitative Voronovskaya theorem via weighted logarithmic modulus of continuity. For $f\in C_2\left(\mathbb{R}^+\right)$ and $\delta>0$, the weighted logarithmic modulus of continuity is given by

$$\Omega(f,\delta):=\sup_{|\log h|\leq\delta}\frac{|f(hx)-f(x)|}{\left(1+\log^2 h\right)\left(1+\log^2 x\right)} \tag{4}$$

(see [9]).

The weighted logarithmic modulus of continuity (4) satisfies the following some elementary properties.

Lemma 2.1. *([9]) Let $\delta>0$. Then*
1.) for $f\in C_2\left(\mathbb{R}^+\right)$, the quantity $\Omega(f,\delta)$ is finite,
2.) for all $f\in C_2\left(\mathbb{R}^+\right)$and each $\lambda\in\mathbb{R}^+$,

$$\Omega(f,\lambda\delta)\leq 2(1+\lambda)^3\left(1+\delta^2\right)\Omega(f,\delta),$$

3.) for all $f\in C_2\left(\mathbb{R}^+\right)$,

$$|f(u)-f(x)|\leq 16\left(1+\delta^2\right)^2\left(1+\log^2 x\right)\Omega(f,\delta)\left(1+\frac{|\log u-\log x|^5}{\delta^5}\right),\quad u,x>0$$

and
4.) for $f\in C_2^*\left(\mathbb{R}^+\right)$,

$$\lim_{\delta\to 0}\Omega(f,\delta)=0.$$

Let us recall the Mellin-Taylor formula given in [15, 36]. For any $f\in C_B\left(\mathbb{R}^+\right)$ belonging to $C^{(n)}\left(\mathbb{R}^+\right)$ locally at the point $x\in\mathbb{R}^+$, the Mellin-Taylor formula with the Mellin derivatives is defined by

$$f(a)=\sum_{i=0}^n\frac{1}{i!}\Theta^i f(x)(\log a-\log x)^i+R_n(a),\quad a>0,n\in\mathbb{N},$$

where

$$R_n(a)=\frac{\left(\Theta^n f(\xi)-\Theta^n f(x)\right)}{n!}(\log a-\log x)^n$$

is the Lagrange remainder in Mellin-Taylor's formula at the point $x\in\mathbb{R}^+$and ξ is a suitable number between a,x. Now, we need the following estimate presented in [1]:

$$|R_n(a)|\leq\frac{64}{n!}\left(1+\log^2 x\right)\Omega\left(\Theta^n f,\delta\right)\left(|\log a-\log x|^n+\frac{|\log a-\log x|^{n+5}}{\delta^5}\right). \tag{5}$$

Theorem 2.2. *Let $\varphi \in \Phi$ and $\chi \in \Psi$ be such that $M_6(\varphi), \tilde{M_6}(\chi)$ are finite. Suppose that $m_1(\varphi, x) = m_1(\varphi)$ are indepentend of x. If $\Theta f \in C_2^*(\mathbb{R}^+)$, then we have*

$$\begin{aligned}
&\left|w\left[\left(S_w^{\varphi,\chi} f\right)(x) - f(x)\right] - \Theta f(x)\left\{m_1(\varphi) + \tilde{m}_1(\chi)\right\}\right| \\
\leq &\frac{2^{11}}{w}\left(1+\log^2 x\right)\Omega\left(\Theta f, \frac{1}{w}\right)\Big\{M_0(\varphi)\tilde{M}_1(\chi) + M_1(\varphi)\tilde{M}_0(\chi) \\
&+M_0(\varphi)\tilde{M}_6(\chi) + M_6(\varphi)\tilde{M}_0(\chi)\Big\}
\end{aligned}$$

for $x \in \mathbb{R}^+$.

Proof. Considering the Mellin-Taylor formula and using the definition of the operators (3), we can write that

$$\begin{aligned}
\left(S_w^{\varphi,\chi} f\right)(x) = &\sum_{k\in\mathbb{Z}} \varphi\left(e^{-k}x^w\right) w \int_0^\infty \left(\chi e^{-k}u^w\right)\left[f(x) + \Theta f(x)(\log u - \log x)\right]\frac{du}{u} \\
&+ \sum_{k\in\mathbb{Z}} \varphi\left(e^{-k}x^w\right) w \int_0^\infty \left(\chi e^{-k}u^w\right) R_1(u)\frac{du}{u} \\
&:= S_1 + S_2
\end{aligned}$$

for $x \in \mathbb{R}^+$ and $w > 0$. Firstly, we consider S_1. Thanks to assumptions $\varphi 1), \chi 1$) and using the change of variable $e^{-k}u^w = t$, we have

$$\begin{aligned}
S_1 &= f(x) + \frac{\Theta f(x)}{w}\sum_{k\in\mathbb{Z}} \varphi\left(e^{-k}x^w\right)\int_0^\infty \chi(t)\left[\log t + (k - w\log x)\right]\frac{dt}{t} \\
&= f(x) + \frac{\Theta f(x)}{w}\left\{m_1(\varphi) + \tilde{m}_1(\chi)\right\}
\end{aligned}$$

As to S_2, using the inequality (5), we get

$$\begin{aligned}
|S_2| \leq &\sum_{k\in\mathbb{Z}} \left|\varphi\left(e^{-k}x^w\right)\right| \int_0^\infty \left|\chi\left(e^{-k}u^w\right)\right| |R_1(u)|\frac{du}{u} \\
= &\sum_{k\in\mathbb{Z}} \left|\varphi\left(e^{-k}x^w\right)\right| \int_0^\infty |\chi(t)| \left|R_1\left(te^k\right)^{1/w}\right| \frac{dt}{t} \\
\leq &\, 64\left(1+\log^2 x\right)\Omega(\Theta f, \delta)\sum_{k\in\mathbb{Z}} \left|\varphi\left(e^{-k}x^w\right)\right| \int_0^\infty |\chi(t)| \left[\left|\log\left(te^k\right)^{1/w} - \log x\right|\right. \\
&\left. + \frac{\left|\log\left(te^k\right)^{1/w} - \log x\right|^6}{\delta^5}\right]\frac{dt}{t} \\
\leq &64\left(1+\log^2 x\right)\Omega(\Theta f, \delta)\left\{\frac{1}{w}\sum_{k\in\mathcal{Z}} \left|\varphi\left(e^{-k}x^w\right)\right| \int_0^\infty |\chi(t)|\left[|\log t| + |k - w\log x|\right]\frac{dt}{t}\right. \\
&\left. + \frac{2^5}{w^6\delta^5}\sum_{k\in\mathbb{Z}} \left|\varphi\left(e^{-k}x^w\right)\right| \int_0^\infty |\chi(t)|\left[|\log t|^6 + |k - w\log x|^6\right]\frac{dt}{t}\right\} \\
\leq &64\left(1+\log^2 x\right)\Omega(\Theta f, \delta)\left\{\frac{M_0(\varphi)\tilde{M}_1(\chi) + M_1(\varphi)\tilde{M}_0(\chi)}{w}\right. \\
&\left. + \frac{2^5}{w^6\delta^5}\left(M_0(\varphi)\tilde{M}_6(\chi) + M_6(\varphi)\tilde{M}_0(\chi)\right)\right\}.
\end{aligned}$$

Now we set $\delta = w^{-1}$, then we have

$$\begin{aligned}
|S_2| \leq &\frac{2^{11}}{w}\left(1+\log^2 x\right)\Omega\left(\Theta f, \frac{1}{w}\right)\Big\{M_0(\varphi)\tilde{M}_1(\chi) + M_1(\varphi)\tilde{M}_0(\chi) \\
&+M_0(\varphi)\tilde{M}_6(\chi) + M_6(\varphi)\tilde{M}_0(\chi)\Big\}.
\end{aligned}$$

Thus, we obtain which is desired.

□

Corollary 2.3. *From the Theorem 2.2, if we assume that for any* $1 \leq j \leq n-1$, $m_j(\varphi) = \tilde{m}_j(\chi) = 0, m_n(\varphi) \neq 0, \tilde{m}_n(\chi) \neq 0$ *and* $M_{n+5}(\varphi), \tilde{M}_{n+5}(\chi)$ *are finite, we have for that* $\Theta^n f \in C_2^*(\mathbb{R}^+)$ *that*

$$\begin{aligned}
&\left| w\left[\left(S_w^{\varphi,\chi} f\right)(x) - f(x) \right] - \Theta^n f(x) \left\{ m_n(\varphi) + \tilde{m}_n(\chi) \right\} \right| \\
\leq &\frac{2^{n+10}}{w^n n!} \left(1 + \log^2 x\right) \Omega\left(\Theta f, \frac{1}{w^n}\right) \Big\{ M_0(\varphi) \tilde{M}_n(\chi) + M_n(\varphi) \tilde{M}_0(\chi) \\
&+ M_0(\varphi) \tilde{M}_{n+5}(\chi) + M_{n+5}(\varphi) \tilde{M}_0(\chi) \Big\}.
\end{aligned}$$

References

[1] T. Acar, S. Kursun, Pointwise Convergence of Generalized Kantorovich Exponential Sampling Series, Dolomites Res. Notes Approx., Accepted., 2022.

[2] T. Acar, S. Kursun, M. Turgay, Multidimensional Kantorovich modifications of exponential sampling series, Quaestiones Mathematicae, 2022, 1-16, doi.org/10.2989/16073606.2021.1992033.

[3] T. Acar, O. Alagoz, A. Aral, D. Costarelli, M. Turgay, G. Vinti, Convergence of generalized sampling series in weighted spaces, Demonstratio Mathematica, 55 2022, 153-162.

[4] T. Acar, O. Alagoz, A. Aral, D. Costarelli, M. Turgay, G. Vinti, Approximation by sampling Kantorovich series in weighted space of functions, Turkish Journal of Mathematics, 46 (7) 2022, 2663-2676.

[5] T. Acar, M. Turgay, Approximation by Bivariate Generalized Sampling Series in Weighted Spaces of Functions, Dolomites Res. Notes Approx., Accepted., 2022.

[6] O. Alagoz, M. Turgay, T. Acar, M. Parlak, Approximation by sampling Durrmeyer operators in weighted space of functions, Numerical Functional Analysis and Optimization, 43 (10) 2022, 1223-1239.

[7] S. K. Angamuthu, S. Bajpeyi, Direct and inverse result for Kantorovich type exponential sampling series. Results in Mathematics, 75 (3) 2020, doi.org/10.1007/s00025-020-01241-0.

[8] S. K. Angamuthu, S. Bajpeyi, Inverse approximation and GBS of bivariate Kantorovich type sampling series, RACSAM, 114 (2) 2020, doi.org/10.1007/s13398-020-00805-7.

[9] A. Aral, T. Acar, S. Kursun, Generalized Kantorovich forms of exponential sampling series, Analysis and Mathematical Physics, 12 (2) 2022, 1-19.

[10] S. Bajpeyi, S. K. Angamuthu, On approximation by Kantorovich exponential sampling op- erators, Numerical Functional Analysis and Optimization, 42 (9) 2021, 1096-1113.

[11] S. Balsamo, I. Mantellini, On linear combinations of general exponential sampling series, Results in Mathematics, 74 (4) 2019, doi.org/10.1007/s00025-019-1104-x.

[12] C. Bardaro, G. Bevignani, I. Mantellini, M. Seracini, Bivariate generalized exponential sampling series and applications to seismic waves, Constructive Mathematical Analysis, 2 (4) 2019, 153-167.

[13] C. Bardaro, P. L. Butzer, R. L. Stens, G. Vinti, Approximation error of the Whittaker cardinal series in terms of an averaged modulus of smoothness covering discontinuous signals, Journal of Mathematical Analysis and Applications, 316 2006, 269-306.

[14] C. Bardaro, L. Faina, I. Mantellini, A generalization of the exponential sampling series and its approximation properties, Mathematica Slovaca, 67 (6) 2017, 1481-1496.

[15] C. Bardaro, I. Mantellini, A note on the Voronovskaja theorem for Mellin-Fejer convolution operators, Applied Mathematics Letters, 24 2011, 2064-2067.

[16] C. Bardaro, I. Mantellini, On a Durrmeyer-type modification of the exponential sampling series, Rendiconti del Circolo Matematico di Palermo Series, 70 (3) 2021, 1289-1304.

[17] C. Bardaro, I. Mantellini, G. Schmeisser, Exponential sampling series: conergence in Mellin -Lebesgue spaces, Results in Mathematics, 74:119, 2019.

[18] C. Bardaro, G. Vinti, P. L. Butzer, R. Stens, Kantorovich-type generalized sampling series in the setting of Orlicz spaces, Sample. Theory Signal Image Process., 6 (1) 2007, 29-52.

[19] M. Bertero, E. R. Pike, Exponential sampling method for Laplace and other dilationally, in- variant transforms I. Singular system analysis II. Examples in photon correction spectroscopy and Frauenhofer diraction, Inverse Problems, 7 1991, 1-20; 21-41.

[20] P. L. Butzer, M. Dadson, P. J. S. G. Ferreira, J. R. Higgins, G. Schmeisser, R. L. Stens, Seven pivotal theorems of Fourier analysis, signal analysis, numerical analysis and number theory: their interconnections, Bulletin of Mathematical Sciences, 4 (3) 2014, 481-525.

[21] P. L. Butzer, P. J. S. G. Ferreira, J. R. Higgins, G. Schmeisser, R. L. Stens, The sampling theorem, Poisson's summation formula, general Parseval formula, reproducing kernel formula and the Paley-Wiener theorem for bandlimited signals-their interconnections, Applicable Analysis, 90 (3-4) 2011, 431-461.

[22] P. L. Butzer, S. Jansche, A direct approach to the Mellin transform, Journal of Fourier Analysis and Applications, 3 1997, 325-375.

[23] P. L. Butzer, S. Ries, R. L. Stens, Approximation of continuous and discontinuous functions by generalized sampling series, Journal of Approximation Theory, 50 1987, 25-39.

[24] P. L. Butzer, G. Schmeisser, R. L. Stens, An introduction to sampling analysis. In: Marvasti, F. (ed.) Nonuniform Sampling, Theory and Practice, pp. 17-121. Kluwer, New York, 2001.

[25] P. L. Butzer, R. L. Stens, Linear prediction by samples from the past. In: Advanced Topics in Shannon Sampling and Interpolation Theory, New York:Springer Texts Electrical Eng., Springer, pp. 157-183, 1993.

[26] D. Casasent (Ed), Optical data processing, Springer, Berlin, 1978, 241-282.

[27] D. Costarelli, A. M. Minotti, G. Vinti, Approximation of discontinuous signals by sampling Kantorovich series, Journal of Mathematical Analysis and Applications, 450 (2) 2017, 1083-1103.
[28] D. Costarelli, G. Vinti, Order of approximation for sampling Kantorovich operators, Journal of Integral Equations and Applications, 26 (3) 2014, 345-368.
[29] D. Costarelli, G. Vinti, An inverse result of approximation by sampling Kantorovich series, Proceedings of the Edinburgh Mathematical Society, 62 (1) 2019, 265-280.
[30] F. Gori, Sampling in optics In: R. J. Marks II (ed.), Advances topics in Shannon sampling and interpolation theory, Springer, New York, 1993, 37-83.
[31] J. R. Higgins, Sampling theory in Fourier and signal analysis. Foundations. Oxford University Press, Oxford, 1996.
[32] J. R. Higgins, G. Schmeisser, J. J. Voss, The sampling theorem and several equivalent results in analysis, Journal of Computational Analysis and Applications, 2 (4) 2000, 333-371.
[33] A. Kivinukk, G. Tamberg, Interpolating generalized Shannon sampling operators, their norms and approximation properties, Sample. Theory Signal Image Process., 8 (1) 2009, 77-95.
[34] V. A. Kotel'nikov, On the carrying capacity of "ether" and wire in electrocommunications, Material for the First All-Union Conference on the Questions of Communications, Moscow, 1933.
[35] S. Kursun, M. Turgay, O. Alagoz, T. Acar, Approximation properties of multivariate exponential sampling series, Carpathian Mathematical Publications, 13 (3) 2021, 666-675.
[36] R. G. Mamedov, The Mellin transform and approximation theory (in Russian), "Elm", Baku, 1991.
[37] N. Ostrowsky, D. Sornette, P. Parker, E. R. Pike, Exponential sampling method for light scattering polydispersity analysis, Optica Acta, 28 1994, 1059-1070.
[38] C. E. Shannon, Communications in the presence of noise, Proceedings of the IRE, 37 1949, 10-21.
[39] E. T. Whittaker, On the functions, which are represented by expansions of the interpolation theory, Proceedings of the Royal Society Edinburgh, 35 1915, 181-194.
[40] A. I. Zayed, Advances in Shannon's sampling theory, CRC Press, Boca Raton, 1993.

Department of Mathematics, Faculty of Sciences, Selcuk University, Konya, 42003, Turkiye

Email(s): tunceracar@ymail.com

Lipschitz Operator Ideals

Dahmane Achour

Abstract. Farmer and Johnson introduce the Lipschitz p-summing operator ideals between metric spaces. Their work motivated many authors to study different classes of Lipschitz mappings that extend, in some sense, linear operators ideals, leading to the recent notion of Banach Lipschitz operator ideals. In this work we will give some basics of the theory of Lipschitz and two Lipschitz operator ideals, also we introduce the concept of p-nuclear two Lipschitz operator $(1 \leq p \leq \infty)$ with respect to a Banach ideal of two Lipschitz functions. .

2020 Mathematics Subject Classifications: 47L22, 46G25, 47L20.

Keywords: Lipschitz p-summing operator, Lipschitz operator ideal, Two-Lipschitz operator ideal, p-nuclear operator.

1. Introduction

Farmer and Johnson in (*cf.* [7]) have introduced the notion of Lipschitz p-summing operators and proved basic fundamental properties and leave to interested readers a list of open problems (*what results about summing operators have analogues for Lipschitz summing operators?*). Since then, many works have appeared related to this class of Lipschitz mappings and some of the problems have been solved. The theory of ideals of Lipschitz mappings between metric spaces and Banach spaces (Lipschitz-ideals) was initiated by myself et al (*cf.* [1]) as a first step to take to the nonlinear setting the successful theory of ideals of linear operators introduced by Pietsch (*cf.* [12]) (operator ideals). Since then much research has been done in this subject, we mention just a few recent developments: (*cf.* [2, 16, 11]). In 2020 an axiomatic theory of two Lipschitz operator ideals for Banach spaces-valued Lipschitz mappings was given by Hamidi et al. in (*cf.* [11])

Our purpose is to develop a basic theory of the Lipschitz and two Lipschitz operator ideals, also we introduce the concept of p-nuclear two Lipschitz operator $(1 \leq p \leq \infty)$ with respect to a Banach ideal of two Lipschitz functions.

1.1. **Preliminaries.** Let us fix some notation and notions used in this paper. Our notation is standard. X and Y will be pointed metric spaces with a base point denoted by 0 and the metric will be denoted by d. We denote by $B_X = \{x \in X : d(x,0) \leq 1\}$. Also, E and F will stand for Banach spaces over the same field $\mathbb{K}$ (either $\mathbb{R}$ or $\mathbb{C}$) with dual spaces E^* and F^*. A Banach space E will be considered as pointed metric spaces with distinguished point 0 and distance $d(x,x') = \|x - x'\|$. With $Lip_0(X,Y)$ we denote the set of all Lipschitz mappings from X to Y such that maps 0 to 0 and we put

$$Lip(T) = \inf\{C > 0 \colon d(T(x),T(x')) \leq Cd(x,x');\ \forall x,x' \in X\}.$$

In particular, $Lip_0(X,E)$ is the Banach space of all Lipschitz mappings T from X to E that vanish at 0, under the Lipschitz norm $Lip(\cdot)$. When $E = \mathbb{K}$, $Lip_0(X,\mathbb{K})$ is denoted by $X^\#$ and it is called the Lipschitz dual of X. We consider $B_{X^\#}$ endowed with the pointwise topology. Is well known that, with this topology, $B_{X^\#}$ is a compact Hausdorff space. The space of all linear operators from E to F is denoted by $\mathcal{L}(E,F)$ and it is a Banach space with the usual supremum norm. It is clear that $\mathcal{L}(E,F)$ is a subspace of $Lip_0(E,F)$ and, in particular, E^* is a subspace of $E^\#$. Let $p \geq 1$, we write p' the conjugate index of p, that is $1/p + 1/p' = 1$. As usual, when $p = 1, p' = \infty$. For a Banach space E, $\ell_p(E)$ denotes the Banach space of all absolutely p-summable sequences $(x_n)_n$ in E with the norm $\|(x_n)_n\|_p = (\sum_{n=1}^{\infty} \|x_n\|^p)^{\frac{1}{p}}$ and by $\ell_{p,\omega}(E)$ the space of all sequences $(x_n)_n$ in E with the norm $\|(x_n)_n\|_{p,\omega} = \sup_{x^* \in B_{E^*}} \|(x^*(x_n))_n\|_p$.

All the other relevant terminology and preliminaries that we will use are given in corresponding sections. For the theory of Lipschitz mappings to the book of Weaver (*cf.* [15]).

2. Definitions and basic properties

The aim of this section is to introduce the concepts of Lipschitz and two Lipschitz operator ideals. We will follow the spirit of the definition of linear operator ideal explained in the excellent monographs (*cf.* [6, 12]) and the definition of bilinear functional ideal in (*cf.* [13]), for vector-valued bilinear operators it can be found, e.g., in (*cf.* [8, 9]).

We center our study in the case when the Lipschitz operators are consider from a pointed metric spaces to Banach spaces because the theory of Lipschitz operator ideals between a pointed metric space and a Banach space is richer.

2.1. **Lipschitz ideals.** A mapping $T \in Lip_0(X, E)$ has Lipschitz finite dimensional rank if the linear hull of the set $\left\{\frac{T(x)-T(x')}{d(x,x')}, x, x' \in X, x \neq x'\right\}$ is a finite dimensional subspace of E. We denote by $Lip_{0\mathcal{F}}(X, E)$ the set of all Lipschitz finite rank mappings from X to E. Clearly, $Lip_{0\mathcal{F}}(X, E)$ is a linear subspace of $Lip_0(X, E)$. It is proved in(*cf.* [10, Proposition 2.4]) that having finite dimensional rank is equivalent to having Lipschitz finite dimensional rank. This is also equivalent to saying that the linearization T_L of T has finite rank (see ((*cf.* [1])). Such a mapping admits a finite representation $T = \sum_{k=1}^{n} y_k f_k$, with $(y_k)_{k\leq n} \in E, (f_k)_{k\leq n} \in X^{\#}$.

We recall the definition of Lipschitz operator ideal between pointed metric space and Banach space as this notion was introduced by myself et al. in (*cf.* [1]).

Definition 2.1. *A Lipschitz operator ideal $\mathcal{I}_{Lip}$ is a subclass of Lip_0 such that for every pointed metric space X and every Banach space E the components*

$$\mathcal{I}_{Lip}(X, E) := Lip_0(X, E) \cap \mathcal{I}_{Lip}$$

satisfy:
(i) $\mathcal{I}_{Lip}(X, E)$ is a linear subspace of $Lip_0(X, E)$ containing the Lipschitz finite rank operators mappings.
(ii) The ideal property: if $S \in Lip_0(Y, X)$, $T \in \mathcal{I}_{Lip}(X, E)$ and $w \in \mathcal{L}(E, F)$, then the composition wTS is in $\mathcal{I}_{Lip}(Y, F)$.

A Lipschitz operator ideal $\mathcal{I}_{Lip}$ is a normed (Banach) Lipschitz operator ideal if there is $\|.\|_{\mathcal{I}_{Lip}} : \mathcal{I}_{Lip} \longrightarrow [0, +\infty[$ that satisfies
(i') For every pointed metric space X and every Banach space E, the pair $(\mathcal{I}_{Lip}(X, E), \|.\|_{\mathcal{I}_{Lip}})$ is a normed (Banach) space and $Lip(T) \leq \|T\|_{\mathcal{I}_{Lip}}$ for all $T \in \mathcal{I}_{Lip}(X, E)$.
(ii') $\|Id_{\mathbb{K}} : \mathbb{K} \longrightarrow \mathbb{K}, Id_{\mathbb{K}}(\lambda) = \lambda\|_{\mathcal{I}_{Lip}} = 1$.
(iii') If $S \in Lip_0(Y, X)$, $T \in \mathcal{I}_{Lip}(X, E)$ and $w \in \mathcal{L}(E, F)$, then

$$\|wTS\|_{\mathcal{I}_{Lip}} \leq Lip(S) \|T\|_{\mathcal{I}_{Lip}} \|w\|.$$

Remark 1. *a) From conditions (i'), (ii') and (iii') we get for any $v \in E$ and any $g \in X^{\#}$,*

$$\begin{aligned} \|vg\|_{\mathcal{I}_{Lip}} &= \|(v\ Id_{\mathbb{K}}) \circ Id_{\mathbb{K}} \circ g\|_{\mathcal{I}_{Lip}} \leq \|v\ Id_{\mathbb{K}}\| \|Id_{\mathbb{K}}\|_{\mathcal{I}_{Lip}} Lip(g) \\ &= \|v\| Lip(g) = Lip(vg) \leq \|vg\|_{\mathcal{I}_{Lip}}, \end{aligned}$$

that is, $\|vg\|_{\mathcal{I}_{Lip}} = \|v\| Lip(g) = Lip(vg)$.

b) If $g \in X^{\#}$, we have $g \in \mathcal{I}_{Lip}(X, \mathbb{K})$ and $\|g\|_{\mathcal{I}_{Lip}} = Lip(g)$.

Example 2.2. *Lipschitz p-summing operators. Farmer and Johnson introduced the concept of Lipschitz p-summing and Lipschitz p-integral operators (cf. [7]), extending the p-summing and p-integral linear operators to the Lipschitz case. For pointed metric spaces X and Y, a mapping $T \in Lip_0(X, Y)$ is called Lipschitz p-summing, $1 \leq p < \infty$, if there exists a constant $C > 0$ such that regardless of the choice of points $x_1, \ldots, x_n, x'_1, \ldots, x'_n$ in X*

$$\sum_{i=1}^{n} (d(Tx_i, Tx'_i))^p \leq C^p \sup_{f \in B_{X^{\#}}} \sum_{i=1}^{n} |f(x_i) - f(x'_i)|^p. \tag{1}$$

In this case we put $\pi_p^L(T) = \inf\{C : (1)\}$. The set of all Lipschitz p-summing operators from X to Y is denoted by $\Pi_p^L(X, Y)$.

$\cdot$ *(Π_p^L, π_p^L) is a Banach Lipschitz operator ideal.*

Example 2.3. *Strongly Lipschitz p-nuclear operators. Following (cf. [5]), a map $T \in Lip_0(X, E)$ is strongly Lipschitz p-nuclear, $1 \leq p \leq \infty$ if $T = \sum_{n=1}^{\infty} f_n \overline{\otimes} y_n$, where $(Lip(f_n))_n \in$*

ℓ_p and $(y_n)_n \in \ell_{p^*,\omega}(E)$. The set of all such T is denoted by $\mathcal{SN}_p^L(X,E)$ and $\nu_p^L(T)$ denote the infimum of $\|(Lip(f_n))_n\|_p \|(y_n)_n\|_{p,\omega}$ where the infimum is taken over all the strongly Lipschitz p-nuclear representations of T .

$\cdot$ *$(\mathcal{SN}_p^L, \nu_p^L)$ is a Banach Lipschitz operator ideal.*

2.2. **Two-Lipschitz operator ideals.** Let (X, d_X) and (Y, d_Y) be pointed metric spaces and let E be a Banach space, we say that a map $T : X \times Y \longrightarrow E$ is a two-Lipschitz operator if there is a constant $C > 0$ such that for each $x, x' \in X$ and $y, y' \in Y$,

$$\left\|T\left(x,y\right)-T\left(x,y'\right)-T\left(x',y\right)+T\left(x',y'\right)\right\| \leq C.d_X\left(x,x'\right)d_Y\left(y,y'\right). \tag{2}$$

By $BLip_0\left(X,Y;E\right)$ we denote the set of all two-Lipschitz operators from $X \times Y$ to E such that

$$T(x,0) = T(0,y) = 0, \tag{3}$$

for all $x \in X$ and $y \in Y$. For $T \in BLip_0\left(X,Y;E\right)$ we set

$$BLip\left(T\right) = \inf C = \sup_{x\neq x', y\neq y'} \frac{\left\|T\left(x,y\right)-T\left(x,y'\right)-T\left(x',y\right)+T\left(x',y'\right)\right\|}{d_X\left(x,x'\right)d_Y\left(y,y'\right)}. \tag{4}$$

Example 2.4. *If X and Y be Banach spaces. Then every bilinear operator $T : X \times Y \longrightarrow E$ is two-Lipschitz. Moreover, we have $BLip\left(T\right) = \|T\|$. In order to see this, for each $x, x' \in X$ and $y, y' \in Y$,*

$$\begin{aligned}&\left\|T\left(x,y\right)-T\left(x,y'\right)-T\left(x',y\right)+T\left(x',y'\right)\right\| \\ &= \left\|T\left(x-x',y-y'\right)\right\| \leq \|T\|\left\|x-x'\right\|\left\|y-y'\right\|.\end{aligned}$$

Therefore, $BLip\left(T\right) \leq \|T\|$. For the reverse inequality, we will write (4) for $x' = y' = 0$,

$$BLip\left(T\right) \geq \sup_{x\neq 0, y\neq 0} \frac{\left\|T\left(x,y\right)\right\|}{d_X\left(x,0\right)d_Y\left(y,0\right)} = \|T\|.$$

The next theorem and its proof are similar to the Lipschitz case.

Theorem 2.5. *(cf. [11]) 1) Let X, Y, Z, W be pointed metric spaces and let E, F be Banach spaces. If $f \in Lip_0(Z,X)$, $g \in Lip_0(W,Y)$, $T \in BLip_0(X,Y;E)$ and $u \in \mathcal{L}(E,F)$ then $u \circ T \circ (f,g) \in BLip_0(Z,W;F)$, where $(f,g)(z,w) := (f(z), g(w))$, $z \in Z, w \in W$. Moreover,*

$$BLip(u \circ T \circ (f,g)) \leq \|u\|\, BLip(T) Lip(f) Lip(g). \tag{5}$$

2) $BLip_0\left(X,Y;E\right)$ is a Banach space under the norm $BLip\left(\cdot\right)$ defined by (4).

Next we give a simple but crucial example of a two-Lipschitz operator. Let X, Y be pointed metric spaces and let E be Banach space.

Consider non-zero Lipschitz functions $f \in X^{\#}$, $g \in Y^{\#}$ and $e \in E$. Define the mapping $f \cdot g \cdot e : X \times Y \longrightarrow E$ by

$$f \cdot g \cdot e(x,y) = f(x)g(y)e. \tag{6}$$

Then, an easy computation shows that this mapping is two-Lipschitz and

$$BLip(f \cdot g \cdot e) = Lip(f)Lip(g)\left\|e\right\|. \tag{7}$$

Definition 2.6. *We denote by $BLip_{0\mathcal{F}}(X,Y;E)$, the vector subspace of all two-Lipschitz operators generated by the mappings of the special form (6). All elements T of this space are called of finite type. So, any $T \in BLip_{0\mathcal{F}}(X,Y;E)$ admits a finite representation of the form*

$$T = \sum_{i=1}^{n} f_i \cdot g_i \cdot e_i$$

where $(f_i)_{i=1}^n \subset X^{\#}, (g_i)_{i=1}^n \subset Y^{\#}$ and $(e_i)_{i=1}^n \subset E$.

We recall the definition of Banach ideal of two Lipschitz mappings as this notion was introduced by Hamidi et al in (*cf.* [11]).

Definition 2.7. *A two-Lipschitz operator ideal between pointed metric spaces and Banach spaces, $\mathcal{I}_{BLip}$, is a subclass of $BLip_0$ such that for every pointed metric spaces X, Y and every Banach space E the components*

$$\mathcal{I}_{BLip}(X,Y;E) := BLip_0(X,Y;E) \cap \mathcal{I}_{BLip}$$

satisfy:

(i) $\mathcal{I}_{BLip}(X,Y;E)$ *is a vector subspace of* $BLip_0(X,Y;E)$ *containing the two-Lipschitz mappings of finite type.*

(ii) The ideal property: if $f \in Lip_0(Z,X)$, $g \in Lip_0(W,Y)$, $T \in \mathcal{I}_{BLip}(X,Y;E)$ *and* $u \in \mathcal{L}(E,F)$, *then the composition* $u \circ T \circ (f,g)$ *is in* $\mathcal{I}_{BLip}(Z,W;F)$.

A two-Lipschitz operator ideal $\mathcal{I}_{BLip}$ *is a normed (Banach) two-Lipschitz operator ideal if there is* $\|\cdot\|_{\mathcal{I}_{BLip}} : \mathcal{I}_{BLip} \longrightarrow [0,+\infty[$ *that satisfies*

(i') For every pointed metric spaces X,Y *and every Banach space* E *, the pair* $(\mathcal{I}_{BLip}(X,Y;E), \|\cdot\|_{\mathcal{I}_{BLip}})$ *is a normed (Banach) space and* $BLip(T) \leq \|T\|_{\mathcal{I}_{BLip}}$ *for all* $T \in \mathcal{I}_{BLip}(X,Y;E)$.

(ii') $\|Id_{\mathbb{K}^2} : \mathbb{K}\times\mathbb{K} \longrightarrow \mathbb{K} : Id_{\mathbb{K}^2}(\alpha,\beta) = \alpha\beta\|_{\mathcal{I}_{BLip}} = 1$.

(iii') If $f \in Lip_0(Z,X)$, $g \in Lip_0(W,Y)$, $T \in \mathcal{I}_{BLip}(X,Y;E)$ *and* $u \in \mathcal{L}(E,F)$, *the inequality* $\|u \circ T \circ (f,g)\|_{\mathcal{I}_{BLip}} \leq \|u\|\,\|T\|_{\mathcal{I}_{BLip}} Lip(f)Lip(g)$ *holds.*

The two-Lipschitz operator ideal $\mathcal{I}_{BLip}$ is said to be closed if each $\mathcal{I}_{BLip}(X,Y;E)$ is a closed subspace of $BLip(X,Y;E)$ with the norm $BLip(\cdot)$.

Remark 2. *Let* $\mathcal{I}_{BLip}$ *be a normed two-Lipschitz operator ideal,* X,Y *be pointed metric spaces and* E *be Banach space. Then*

$$\|f\cdot g\cdot e\|_{\mathcal{I}_{BLip}} = \|e\|\, Lip(f)Lip(g),$$

for any $f \in X^{\#}$, $g \in Y^{\#}$ *and* $e \in E$.

Proof. Let $f \in X^{\#}$, $g \in Y^{\#}$ and $e \in E$. We can write $f\cdot g\cdot e$ in the following way

$$f\cdot g\cdot e = id_{\mathbb{K}}e \circ Id_{\mathbb{K}^2} \circ (f,g).$$

By (i'),(ii'), (iii') and (7), we obtain directly

$$\begin{aligned}\|f\cdot g\cdot e\|_{\mathcal{I}_{BLip}} &\leq \|id_{\mathbb{K}}e\|\,\|Id_{\mathbb{K}^2}\|_{\mathcal{I}_{BLip}} Lip(f)Lip(g)\\ &= \|e\|\, Lip(f)Lip(g) = BLip(f\cdot g\cdot e)\\ &\leq \|f\cdot g\cdot e\|_{\mathcal{I}_{BLip}},\end{aligned}$$

this gives, $\|f\cdot g\cdot e\|_{\mathcal{I}_{BLip}} = \|e\|\, Lip(f)Lip(g)$. □

The series criterion for operator ideals (*cf.* [6, 9.4], [12, 6.2.3]) and for multi-ideals (*cf.* [4, Satz 2.2.4]) works for Banach two-Lipschitz ideals.

Theorem 2.8. *(Series criterion). Let* $\mathcal{I}_{BLip}$ *be a subclass of the class of all two-Lipschitz operators between metric spaces and Banach spaces endowed with a map* $\|\cdot\|_{\mathcal{I}_{BLip}} : \mathcal{I}_{BLip} \longrightarrow [0,+\infty[$. *Then* $(\mathcal{I}_{BLip}(X,Y;E), \|\cdot\|_{\mathcal{I}_{BLip}})$ *is a Banach ideal if and only if the following conditions are satisfied:*

(1) $Id_{\mathbb{K}^2} \in \mathcal{I}_{BLip}(\mathbb{K},\mathbb{K};\mathbb{K})$ *and* $\|Id_{\mathbb{K}^2}\|_{\mathcal{I}_{BLip}} = 1$.

(2) If $f \in Lip_0(Z,X)$, $g \in Lip_0(W,Y)$, $T \in \mathcal{I}_{BLip}(X,Y;E)$ *and* $u \in \mathcal{L}(E,F)$, *then* $u \circ T \circ (f,g) \in \mathcal{I}_{BLip}(Z,W;F)$ *and* $\|u \circ T \circ (f,g)\|_{\mathcal{I}_{BLip}} \leq \|u\|\,\|T\|_{\mathcal{I}_{BLip}} Lip(f)Lip(g)$.

(3) If $(T_n)_n \subset \mathcal{I}_{BLip}(X,Y;E)$ *such that* $\displaystyle\sum_{n=1}^{\infty} \|T_n\|_{\mathcal{I}_{BLip}} < \infty$, *then* $T := \displaystyle\sum_{n=1}^{\infty} T_n \in \mathcal{I}_{BLip}(X,Y;E)$ *and* $\|T\|_{\mathcal{I}_{BLip}} \leq \displaystyle\sum_{n=1}^{\infty} \|T_n\|_{\mathcal{I}_{BLip}}$.

Proof. First suppose that $(\mathcal{I}_{BLip}, \|\cdot\|_{\mathcal{I}_{BLip}})$ is a Banach two-Lipschitz operator ideal. The map $Id_{\mathbb{K}^2}$ is of finite rank so according to the Definition 2.7 we have

$Id_{\mathbb{K}^2} \in \mathcal{I}_{BLip}(\mathbb{K},\mathbb{K};\mathbb{K})$ and $\|Id_{\mathbb{K}^2}\|_{\mathcal{I}_{BLip}} = 1$. For condition (3), the series $\displaystyle\sum_{n=1}^{\infty} \|T_n\|_{\mathcal{I}_{BLip}}$ is absolutely convergent in the Banach space $\mathcal{I}_{BLip}(X,Y;E)$ then it is convergent and $T \in \mathcal{I}_{BLip}(X,Y;E)$. The continuity of the norm in a normed space gives

$$\begin{aligned}\|T\|_{\mathcal{I}_{BLip}} &= \left\|\sum_{n=1}^{\infty} T_n\right\|_{\mathcal{I}_{BLip}} = \left\|\lim_n \sum_{j=1}^{n} T_j\right\|_{\mathcal{I}_{BLip}} = \lim_n \left\|\sum_{j=1}^{n} T_j\right\|_{\mathcal{I}_{BLip}}\\ &\leq \lim_n \sum_{j=1}^{n} \|T_j\|_{\mathcal{I}_{BLip}} = \sum_{n=1}^{\infty} \|T_n\|_{\mathcal{I}_{BLip}}.\end{aligned}$$

Conversely, we assume that conditions (1), (2) and (3) are satisfied. Let (X, d_X) and (Y, d_Y) be pointed metric spaces and let E be a Banach space. It is clear that if $T = 0$ from $X \times Y$ to E and $\upsilon = 0$ from $\mathbb{K}$ to E, $f \in X^{\#}$ and , $g \in Y^{\#}$ we have

$$T = \upsilon \circ Id_{\mathbb{K}^2} \circ (f, g)$$

and hypothesis (2) gives $T = 0 \in \mathcal{I}_{BLip}(X, Y; E)$ and $\|0\|_{\mathcal{I}_{BLip}} = 0$. Assumption (3) states that if $T_1, T_2 \in \mathcal{I}_{BLip}(X, Y; E)$ then

$$T_1 + T_2 \in \mathcal{I}_{BLip}(X, Y; E) \quad \text{and} \quad \|T_1 + T_2\|_{\mathcal{I}_{BLip}} \leq \|T_1\|_{\mathcal{I}_{BLip}} + \|T_2\|_{\mathcal{I}_{BLip}}$$

Now let $T \in \mathcal{I}_{BLip}(X, Y; E)$ and $\lambda \in \mathbb{K}$. It is clear that the linear operator

$$R_\lambda : E \longrightarrow E, R_\lambda(e) = \lambda e,$$

is continuous with $\|R_\lambda\| = |\lambda|$.
Theorem and Condition (2) gives

$$\lambda T = R_\lambda \circ T \in BLip(X, Y; E), \, \lambda T \in \mathcal{I}_{BLip}(X, Y; E) \quad \text{and} \quad \|\lambda T\|_{\mathcal{I}} \leq |\lambda| \,.\, \|T\|_{\mathcal{I}_{BLip}}$$

On the other hand for $\lambda \neq 0$ we hav

$$\|T\|_{\mathcal{I}_{BLip}} = \left\| \frac{1}{\lambda} \lambda T \right\|_{\mathcal{I}_{BLip}} \leq \frac{1}{|\lambda|} \|\lambda T\|_{\mathcal{I}_{BLip}},$$

Which give $\|\lambda T\|_{\mathcal{I}_{BLip}} \geq |\lambda| \, \|T\|_{\mathcal{I}_{BLip}}$. So

$$\|\lambda T\|_{\mathcal{I}_{BLip}} = |\lambda| \,.\, \|T\|_{\mathcal{I}_{BLip}}$$

For $\lambda = 0$ the equality is obvious. Let $T \in \mathcal{I}_{BLip}(X, Y; E)$ such that $\|T\|_{\mathcal{I}_{BLip}} = 0$. Suppos that $T \neq 0$, then there is $(x, y) \in X \times Y$, $(x, y) \neq (0, 0)$ such that $T(x, y) \neq 0$. By Hahn-Banach theorem there exists a functional $\xi \in E^*$ such that $\xi(T(x, y)) = 1$. The Lipschitz maps defined by

$$S_x : \mathbb{K} \longrightarrow X, \; S_x(\alpha) = x \text{ and } R_y : \mathbb{K} \longrightarrow Y, \; R_y(\beta) = y$$

On the other hand for all non-zero $(\alpha, \beta) \in \mathbb{K} \times \mathbb{K}$ we have

$$\xi \circ T \circ (S_x, R_y)(\alpha, \beta) = \xi(T(S_x(\alpha), R_y(\beta))) = \xi(T(x, y)) = 1 = \frac{1}{\alpha\beta} Id_{\mathbb{K}^2}(\alpha, \beta).$$

So, condition (2) implies that. $\xi \circ T \circ (S_x, R_y) = \frac{1}{\alpha\beta} Id_{\mathbb{K}^2} \in \mathcal{I}_{BLip}(\mathbb{K} \times \mathbb{K}; \mathbb{K})$ and $\frac{1}{|\alpha\beta|} \|Id_{\mathbb{K}^2}\|_{\mathcal{I}_{BLip}} \leq \|\varphi\| \,.\, \|T\|_{\mathcal{I}} \,.\, \|S\| = 0$. Then $\|Id_{\mathbb{K}^2}\|_{\mathcal{I}_{BLip}} = 0$, which contradicts condition (1), so $T = 0$.

Then we have shown that $\mathcal{I}_{BLip}(X, Y; E)$ is a vector subspace of $BLip(X, Y; E)$ and that $\|\cdot\|_{\mathcal{I}_{BLip}}$ is a norm on vector space $\mathcal{I}_{BLip}(X, Y; E)$.

Now we show that $\mathcal{I}_{BLip}(X, Y; E)$ contains two-Lipschitz mappings of finite rank. Let $f \in X^{\#}$, $g \in Y^{\#}$ and $e \in E$. The linear operator $S_e : \mathbb{K} \to E$, $S_e(\lambda) = \lambda e$,is clearly continuous. On the other hand for all $x \in X$ and $y \in Y$ we have

$$\begin{aligned} S_e \circ Id_{\mathbb{K}^2} \circ (f, g)(x, y) &= S_e \circ Id_{\mathbb{K}^2}(f(x), g(y)) = S_e(f(x)g(y)) \\ &= f(x)g(y)e \\ &= f \cdot g \cdot e(x, y), \end{aligned}$$

then we concluded that $f \cdot g \cdot e = S_e \circ Id_{\mathbb{K}^2} \circ (f, g)$. By conditions (1) and (2) we have $f \cdot g \cdot e \in \mathcal{I}_{BLip}(X, Y; E)$. Let $T \in BLip_{0\mathcal{F}}(X, Y; E)$, as $\mathcal{I}_{BLip}(X, Y; E)$ is a vector space, we have

$$T = \sum_{i=1}^{n} f_i \cdot g_i \cdot e_i \in \mathcal{I}_{BLip}(X, Y; E)$$

where $(f_i)_{i=1}^n \subset X^{\#}, (g_i)_{i=1}^n \subset Y^{\#}$ and $(e_i)_{i=1}^n \subset E$.

Finally, the condition (3)) ensures that $(\mathcal{I}_{BLip}, \|\cdot\|_{\mathcal{I}_{BLip}})$ is a Banach space. □

3. Main results

3.1. Nuclear two-Lipschitz operators with respect to a Banach ideal of two-Lipschitz functions. Let m be a natural number, $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$ be a Banach ideal of bounded m-linear functionals, $1 \leq p \leq \infty$ and $E_1, ..., E_m, F$ Banach spaces. Popa in (*cf.* [14]) introduced, $(\mathcal{N}_p^{\mathcal{A}}, \|\cdot\|_{\mathcal{N}_p^{\mathcal{A}}})$, the concept of p-nuclear multilinear operator with respect to a Banach ideal of multilinear functionals (see also (*cf.* [3]) for hyper p-nuclear).

For $1 \leq p \leq \infty$, recall that a bilinear operator $u : E_1 \times E_2 \longrightarrow F$ is $(p, \mathcal{A})$-nuclear if there are $(\psi_n)_n \subset \mathcal{A}(E_1, E_2; \mathbb{K}), (z_n)_n \in \ell_{p^*,\omega}(F)$ such that $(\|\psi_n\|_{\mathcal{A}})_n \in \ell_p$, $(z_n)_n \in \ell_{p^*,\omega}(F)$ and

$$u(x, y) = \sum_{n=1}^{\infty} \psi_n(x, y) z_n, \text{for } (x, y) \in E_1 \times E_2 \tag{8}$$

In this case,

$$\|u\|_{\mathcal{N}_p^{\mathcal{A}}} = \inf \left\{ \|(\|\psi_n\|_{\mathcal{A}})_n\|_p \, \|(z_n)_n\|_{p^*,\omega} \right\},$$

where the infimum is taken over all socalled $(p, \mathcal{A})$-nuclear representations described above. We write $\mathcal{N}_p^{\mathcal{A}}(E_1, E_2; F)$ to denote the space of the all $(p, \mathcal{A})$-nuclear operators from $E_1 \times E_2$ into F . For $p = 1$ we write simply $\mathcal{N}^{\mathcal{A}}$ instead of $\mathcal{N}_1^{\mathcal{A}}$.

Now we are going to construct a new two-Lipschitz $(p, \mathcal{A})$-nuclear operator ideal.

Consider non-zero two Lipschitz function $\psi \in BLip_0(X, Y; \mathbb{K})$ and $z \in F$. Define the mapping $\psi \overline{\otimes} z : X \times Y \longrightarrow F$ by

$$(\psi \overline{\otimes} z)(x, y) = \psi(x, y) z \tag{9}$$

Then, an easy computation shows that this mapping is two-Lipschitz and

$$BLip(\psi \overline{\otimes} z) = BLip(\psi) \, \|z\| \, . \tag{10}$$

Indeed, for any $x, x' \in X$ and $y, y' \in Y$ we have

$$\begin{aligned}
&\|(\psi \overline{\otimes} z)(x, y) - (\psi \overline{\otimes} z)(x, y') - (\psi \overline{\otimes} z)(x', y) + (\psi \overline{\otimes} z)(x', y')\| \\
&= \|\psi(x, y)z - \psi(x, y')z - \psi(x', y)z + \psi(x', y')z\| \\
&= \|(\psi(x, y) - \psi(x, y') - \psi(x', y) + \psi(x', y'))z\| \\
&= \|z\| \, \|\psi(x, y) - \psi(x, y') - \psi(x', y) + \psi(x', y')\| \\
&\leq \|z\| \, BLip(\psi) d_X(x, x') \, d_Y(y, y'),
\end{aligned}$$

and $(\psi \overline{\otimes} z)(x, 0) = \psi(x, 0)z = \psi(0, y)z = (\psi \overline{\otimes} z)(0, y)$.

Then $\psi \overline{\otimes} z \in BLip_0(X, Y; E)$ and

$$BLip(\psi \overline{\otimes} z) \leq BLip(\psi) \, \|z\| \, .$$

For the reverse inequality, let $\varepsilon > 0$ and choose $x_0, x'_0 \in X$ and $y_0, y'_0 \in Y$ such that

$$BLip(\psi) - \varepsilon \leq \frac{|\psi(x_0, y_0) - \psi(x_0, y'_0) - \psi(x'_0, y_0) + \psi(x'_0, y'_0)|}{d_X(x_0, x'_0) \, d_Y(y_0, y'_0)}.$$

Then, we have

$$\begin{aligned}
BLip(\psi \overline{\otimes} z) &\geq \frac{\|\psi \overline{\otimes} z(x_0, y_0) - \psi \overline{\otimes} z(x_0, y'_0) - \psi \overline{\otimes} z(x'_0, y_0) + \psi \overline{\otimes} z(x'_0, y'_0)\|}{d(x_0, x'_0) \, d(y_0, y'_0)} \\
&= \frac{|\psi(x_0, y_0) - \psi(x_0, y'_0) - \psi(x'_0, y_0) + \psi(x'_0, y'_0)|}{d(x_0, x'_0) \, d(y_0, y'_0)} \|z\| \\
&\geq (BLip(\psi) - \varepsilon) \, \|z\| \, .
\end{aligned}$$

Since this holds for every $\varepsilon > 0$ we obtain

$$BLip(\psi) \, \|z\| \leq BLip(\psi \overline{\otimes} z) \, .$$

Definition 3.1. *Let $(\mathcal{I}_{BLip}, \|\cdot\|_{\mathcal{I}_{BLip}})$ be a Banach ideal of two Lipschitz functions, $1 \leq p \leq \infty$, X, Y metric spaces and F Banach space. For a two Lipschitz operator $T : X \times Y \longrightarrow F$ is $(p, \mathcal{I}_{BLip})$-nuclear if there are $(\psi_n)_n \subset \mathcal{I}_{BLip}(X, Y; \mathbb{K}), (z_n)_n \subset F$ such that $(\|\psi_n\|_{\mathcal{I}_{Lip}})_n \in \ell_p$, $(z_n)_n \in \ell_{p^*,\omega}(F)$ and*

$$(11) \qquad T(x,y) = \sum_{n=1}^{\infty} \psi_n(x,y) z_n, \text{for } (x,y) \in X \times Y$$

In this case,

$$\|T\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} = \inf \left\{ \left\| (\|\psi_n\|_{\mathcal{I}_{BLip}})_n \right\|_p \|(z_n)_n\|_{p^*,\omega} \right\},$$

where the infimum is taken over all socalled $(p, \mathcal{I}_{BLip})$-nuclear representations described above. We write $\mathcal{N}_p^{\mathcal{I}_{BLip}}(X,Y;F)$ to denote the space of the all $(p, \mathcal{I}_{BLip})$-nuclear operators from $X \times Y$ into F. For $p = 1$ we write simply $\mathcal{N}^{\mathcal{I}_{BLip}}$ instead of $\mathcal{N}_1^{\mathcal{I}_{BLip}}$. It is obvious that in the Lipschitz case ($T \in Lip_0(X;F)$ and $(\psi_n)_n \subset \mathcal{I}_{Lip}(X;\mathbb{K}) = Lip_0(X;\mathbb{K})$) we get the well-known concept of strongly Lipschitz p-nuclear operator (see *cf.* [5]).

Proposition 3.2. *Let $T \in BLip(X,Y;F)$. Then, $T \in \mathcal{N}_p^{\mathcal{I}_{BLip}}(X,Y;F)$ if and only if there are $(\lambda_n)_n \in \ell_p, (\psi_n)_n \subset \mathcal{I}_{BLip}(X,Y;\mathbb{K})$ with $\sup_n \|\psi_n\|_{\mathcal{I}_{BLip}} < \infty, (z_n)_n \in \ell_{p^*,\omega}(F)$ (respectively $\|\psi_n\|_{\mathcal{I}_{BLip}} \to 0$, $(z_n)_n \in \ell_{1,\omega}(F)$ for $p = \infty$) and*

$$(12) \qquad T(x,y) = \sum_{n=1}^{\infty} \lambda_n \psi_n(x,y) z_n, \text{for } (x,y) \in X \times Y$$

Moreover,

$$\|T\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} = \inf \left\{ \|(\lambda_n)_n\|_p (\sup_n \|\psi_n\|_{\mathcal{I}_{BLip}}) \|(z_n)_n\|_{p^*,\omega} \right\},$$

where the infimum is taken over all sequences $(\lambda_n)_n$, $(\psi_n)_n$, $(z_n)_n$ as above.

Now we state and prove our main result.

Theorem 3.3. *Let $(\mathcal{I}_{BLip}, \|\cdot\|_{\mathcal{I}_{BLip}})$ be a Banach ideal of two Lipschitz functions, $1 \le p \le \infty$, X, Y metric spaces and F Banach space. Then*
$(\mathcal{N}_p^{\mathcal{I}_{BLip}}(X,Y;F), \|\cdot\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}})$ is a Banach ideal of two-Lipschitz operators. Moreover, if $\psi \in BLip(X,Y;\mathbb{K})$, $z \in F$, $z \neq 0$, then $\psi \overline{\otimes} z \in \mathcal{N}_p^{\mathcal{I}_{BLip}}(X,Y;F)$ if and only if $\psi \in \mathcal{I}_{BLip}(X,Y;\mathbb{K})$. In this case, $\|\psi \overline{\otimes} z\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} = BLip(\psi) \|z\|$.

Proof. Let us prove that the conditions in Theorem 2.8 are satisfied. (1) It's clear that $Id_{\mathbb{K}^2} \in \mathcal{N}_p^{\mathcal{I}_{BLip}}(\mathbb{K},\mathbb{K};\mathbb{K})$. Regarding $Id_{\mathbb{K}^2}$ as a representation of itself it follows that $\|Id_{\mathbb{K}^2}\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} \le \|Id_{\mathbb{K}^2}\| = 1$. Assuming that $\|Id_{\mathbb{K}^2}\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} < 1$, there would exist a representation $\sum_{n=1}^{\infty} \psi_n \overline{\otimes} z_n$ of $Id_{\mathbb{K}^2}$ with

$$\left\| (\|\psi_n\|_{\mathcal{I}_{BLip}})_n \right\|_p \|(z_n)_n\|_{p^*,\omega} < 1.$$

By Hölder's inequality,

$$\begin{aligned} 1 = Id_{\mathbb{K}^2}(1,1) = \left| \sum_{n=1}^{\infty} \psi_n(1,1) z_n \right| &\le \sum_{n=1}^{\infty} |\psi_n(1,1) z_n| \\ \le \sum_{n=1}^{\infty} \|\psi_n\| |z_n| \le \sum_{n=1}^{\infty} \|\psi_n\|_{\mathcal{I}_{BLip}} |z_n| \\ \le \left\| (\|\psi_n\|_{\mathcal{I}_{BLip}})_n \right\|_p \|(z_n)_n\|_{p^*,\omega} < 1, \end{aligned}$$

a contradiction that gives $\|Id_{\mathbb{K}^2}\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} = 1$.

(2) Let $(S_j)_j \subset \mathcal{N}_p^{\mathcal{I}_{BLip}}(X,Y;F)$ be such that $\sum_{j=1}^{\infty} \|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} < \infty$. Given $\varepsilon > 0$, for every $j \in \mathbb{N}$ there are sequences$(\psi_{j,n})_n \subset \mathcal{I}_{BLip}(X,Y;\mathbb{K}), (z_{j,n})_n \subset F$ such that $(\|\psi_{j,n}\|_{\mathcal{I}_{Lip}})_n \in \ell_p$, $(z_{j,n})_n \in \ell_{p^*,\omega}(F)$,

$$S_j(x,y) = \sum_{n=1}^{\infty} \psi_{j,n}(x,y) z_{j,n}, \text{for } (x,y) \in X \times Y$$

and $\left\|(\|\psi_{j,n}\|_{\mathcal{I}_{BLip}})_n\right\|_p \|(z_{j,n})_n\|_{p^*,\omega} \leq (1+\varepsilon)\|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}}$. We can assume, for each j, that $\left\|(\|\psi_{j,n}\|_{\mathcal{I}_{BLip}})_n\right\|_p < \left[(1+\varepsilon)\|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}}\right]^{\frac{1}{p}}$ and $\|(z_{j,n})_n\|_{p^*,\omega} < \left[(1+\varepsilon)\|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}}\right]^{\frac{1}{p^*}}$.

From

$$\sum_{j,n=1}^{\infty} \|\psi_{j,n}\|^p_{\mathcal{I}_{BLip}} = \sum_{j=1}^{\infty}\sum_{n=1}^{\infty} \|\psi_{j,n}\|^p_{\mathcal{I}_{BLip}} < \left[(1+\varepsilon)\sum_{j=1}^{\infty}\|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}}\right] < \infty, \tag{13}$$

we conclude that $(\|\psi_{j,n}\|_{\mathcal{I}_{BLip}})_{j,n} \in \ell_p$. For each $\xi \in F^*$ with $\|\xi\| \leq 1$ we have

$$\sum_{j,n=1}^{\infty} |\xi(z_{j,n})|^{p^*} = \sum_{j=1}^{\infty}\left(\sum_{n=1}^{\infty} |\xi(z_{j,n})|^{p^*}\right) \leq \sum_{j=1}^{\infty} \|(z_{j,n})_n\|^{p^*}_{p^*,\omega} < \left[(1+\varepsilon)\sum_{j=1}^{\infty}\|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}}\right] < \infty, \tag{14}$$

Therefore $(z_{j,n})_{j,n} \in \ell_{p^*,\omega}(F)$. so, for all $(x,y) \in X \times Y$, the series

$$\sum_{j,n=1}^{\infty} \psi_{j,n}\overline{\otimes} z_{j,n}(x,y) \tag{15}$$

is absolutely convergent in the Banach space F. Then,

$$\sum_{j,n=1}^{\infty} \psi_{j,n}\overline{\otimes} z_{j,n}(x,y) = \sum_{j=1}^{\infty}\sum_{n=1}^{\infty} \psi_{j,n}(x,y) z_{j,n} = \sum_{j=1}^{\infty} S_j(x,y)$$

defines $S : X \times Y \to F$ and shows that (11) it is a representation of S as in (15), proving that S is two-Lipschitz $(p, \mathcal{I}_{BLip})$-nuclear. From (13) and (14) we get

$$\begin{aligned}
\|S\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} &\leq \left\|(\|\psi_{j,n}\|_{\mathcal{I}_{BLip}})_{j,n}\right\|_p \|(z_{j,n})_{j,n}\|_{p^*,\omega} \\
&\left[(1+\varepsilon)\sum_{j=1}^{\infty}\|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}}\right]^{\frac{1}{p}} \left[(1+\varepsilon)\sum_{j=1}^{\infty}\|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}}\right]^{\frac{1}{p^*}} \\
&\leq (1+\varepsilon)\sum_{j=1}^{\infty}\|S_j\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}}.
\end{aligned}$$

Letting $\varepsilon \to 0$ we obtain the desired inequality.

(3) Let $f \in Lip_0(X_1, X)$, $g \in Lip_0(Y_1, Y)$, $T \in \mathcal{N}_p^{\mathcal{I}_{BLip}}(X, Y; E)$ and $u \in \mathcal{L}(F, G)$, then we can write $T = \sum\limits_{n=1}^{\infty} \psi_n \overline{\otimes} z_n$,with $(\|\psi_n\|_{\mathcal{I}_{Lip}})_n \in \ell_p$, $(z_n)_n \in \ell_{p^*,\omega}(F)$. For all $(x_1, y_1) \in X_1 \times Y_1$, we have

$$\begin{aligned}
u \circ T \circ (f,g)(x_1,y_1) &= u \circ T \circ (f(x_1), g(y_1)) \\
&= u\left[\sum_{n=1}^{\infty} \psi_n(f(x_1), g(y_1)) z_n\right] \\
&= \sum_{n=1}^{\infty} \psi_n(f(x_1), g(y_1)) u(z_n).
\end{aligned}$$

Defining $S_n = \psi_n \circ (f,g)$ and $h_n = u(z_n)$, for every $n \in \mathbb{N}$. It is clear that $(h_n)_n \in \ell_{p^*,\omega}(G)$ and $\|(h_n)_n\|_{p^*,\omega} \leq \|u\| \|(z_n)_n\|_{p^*,\omega}$. As $\psi_n \in \mathcal{I}_{BLip}(X, Y; \mathbb{K})$, then $S_n = \psi_n \circ (f,g) \in \mathcal{I}_{BLip}(X_1, Y_1; \mathbb{K})$ and $\|S_n\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} \leq \|\psi_n\|_{\mathcal{I}_{BLip}} Lip(f) Lip(g) < \infty$. This shows that $(\|S_n\|_{\mathcal{I}_{BLip}})_n \in \ell_p$; hence $\sum\limits_{n=1}^{\infty} S_n \overline{\otimes} h_n$ is a representation of $u \circ T \circ (f,g)$ an is (11).

So $u \circ T \circ (f,g)$ is $(p, \mathcal{I}_{BLip})$-nuclear two-Lipschitz mapping and

$$\begin{aligned}
\|u \circ T \circ (f,g)\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} &\leq \left\|(\|S_n\|_{\mathcal{I}_{BLip}})_n\right\|_p \|(h_n)_n\|_{p^*,\omega} \\
&\leq \|u\| \, Lip(f) Lip(g) \left\|(\|\psi_n\|_{\mathcal{I}_{BLip}})_n\right\|_p \|(z_n)_n\|_{p^*,\omega}.
\end{aligned}$$

Taking the infimum over all $(p, \mathcal{I}_{BLip})$-nuclear representations of T we have

$$\|u \circ T \circ (f, g)\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} \leq \|u\| \, \|T\|_{\mathcal{N}_p^{\mathcal{I}_{BLip}}} Lip(f) Lip(g).$$

$u \circ T \circ (f, g) \in \mathcal{I}_{BLip}(Z, W; F)$ and $\|u \circ T \circ (f, g)\|_{\mathcal{I}_{BLip}} \leq \|u\| \, \|T\|_{\mathcal{I}_{BLip}} Lip(f) Lip(g)$.
For the second part we use the same technique in (*cf.* [14, Proposition 1]). □

Acknowledgments

The author acknowledge with thanks the support of the General Direction of Scientific Research and Technological Development (DGRSDT), Algeria.

References

[1] D. Achour, P. Rueda, E.A. Sánchez-Pérez and R. Yahi, Lipschitz operator ideals and the approximation property. J. Math. Anal. Appl.436 (2016) 217–236.

[2] D. Achour, E. Dahia, P. Turco, The Lipschitz injective hull of Lipschitz operator ideals and applications, Banach Journal of Mathematical Analysis volume 14 (2020) 1241–1257

[3] G. Botelho and E.R.Torres, Hyper-ideals of multilinear operators, Linear Algebra and its Applications 482 (2015) 1–20

[4] H.-A. Braunss, Ideale multilinearer Abbildungen und Räume holomorpher Funktionen, Disserta-tion(A), Pädagogische Hochschule 'Karl Liebknecht', Potsdam, 1984.

[5] D. Chen and B. Zheng, Lipschitz p-integral operators and Lipschitz p-nuclear operators, Nonlinear Analysis 75 (2012) 5270-5282.

[6] A. Defant and K. Floret, Tensor norms and operator ideals. North-Holland Mathematics Studies, 176. North-Holland Publishing Co., Amsterdam, 1993.

[7] J. D. Farmer and W. B. Johnson, Lipschitz p-summing operators, Proc. Amer. Math. Soc. 137, 9 (2009) 2989-2995.

[8] K. Floret, D. García, On ideals of polynomials and multilinear mappings between Banach spaces, Arch. Math. (Basel) 81 (2003) 300–308.

[9] K. Floret, S. Hunfeld, Ultrastability of ideals of homogeneous polynomials and multilinear mappings on Banach spaces, Proc. Amer. Math. Soc. 130 (2002) 1425–1435.

[10] A. Jiménez-Vargas, J. M. Sepulcre and Moisés Villegas-Vallecillos, Lipschitz compact operators. J. Math. Anal. Appl. 415 (2014), no. 2, 889–901.

[11] K. Hamidi , E. Dahiaa, D. Achoura, A. Tallaba, Two-Lipschitz operator ideals, J. Math.Anal.Appl.491 (2020)124346

[12] A. Pietsch, Operator ideals. Deutsch. Verlag Wiss., Berlin, 1978; North-Holland, Amsterdam-London-New York-Tokyo, 1980.

[13] A. Pietsch, Ideals of multilinear functionals, in: Proceedings of the Second International Conference on Operator Algebras, Ideals and Their Applications in Theoretical Physics, in: Teubner-Texte Math., vol.62, 1983, pp.185–199.

[14] D. Popa, Nuclear multilinear operators with respect to a Banach ideal of multilinear functionals, Linear Algebra and its Applications 489 (2016) 93–103

[15] N. Weaver, Lipschitz Algebras, World Scientific Publishing Co., Singapore, 1999.

[16] R. Yahi, Certaines classes d'opérateurs générés par une procédure d'interpolation, Ph.D. thesis, University of Mohamed Boudiaf, M'sila, Algeria, 2017.

Laboratory of Functional Analysis and Geometry of Spaces, University Mohamed Boudiaf of M'sila-Algeria PB 166 M'sila 28000, Algeria

Email(s): dahmane.achour@univ-msila.dz

σ-Derivations on Operators Algebras

Mehsin Jabel Atteya

ABSTRACT. In this paper, we study the results of Tsiu-Kwen Lee [7] from a different point of view. We will study this subject without the condition that maps acting on zero products via using zero commutative ring.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: Primary 16W20, 47B47. Secondary 46L07, 16N60.

KEYWORDS: Elementary operator, generalized skew derivation, prime ring, Martindale quotient ring.

1. Introduction

In 1914, Fraenkels paper which under title On zero divisors and the decomposition of ring, posted the modern definition of abstract ring. One of the significant natural questions of Ring Theory, it specifies the conditions which yield commutativity of the ring.
Historically, the study of derivation was initiated during the 1950s and 1960s. Derivations of rings got a tremendous development in 1957, when Posner [9] established two very striking results in the case of prime rings.
Several authors have studied the relationship between the commutativity of a ring and the behavior of a special mapping on that ring. In particular, there has been considerable interest in centralizing automorphisms and derivations defined on rings (see, e.g., [1],[3], [4], and [9], where further references can be found).
In 2002, Lu and Li proved the following result [8, Theorem 6]: Let B be a standard operator algebra in a Banach space X containing the identity operator I and let $\delta\colon B \to B$ be a linear map such that $\delta(AB) = \delta(A)B + A\delta(B)$ for any pair $A, B \in I$ with $AB = 0$. Then $\delta(AB) = \delta(A)B + A\delta(B) - A\delta(I)B$ for all $A, B \in I$. If in addition $\delta(I) = 0$, then δ is a derivation. In other words, the result says that an additive map on a standard operator algebra is almost a derivation if it satisfies the expansion formula of derivations on pairs of elements with zero product. Since standard operator algebras involve many idempotents,
From this point of view Chebotar, Ke and P.-H.Lee studied maps acting on zero products in the context of prime rings [5]. Tsiu-Kwen Lee [7] posted, let A be a prime ring whose symmetric Martindale quotient ring contains a nontrivial idempotent. Generalized skew derivations of A are characterized by acting on zero products.
Precisely, if $g, \delta\colon A \to A$ are additive maps such that $\sigma(x)g(y) + \delta(x)y = 0$ for all $x, y \in A$ with $xy = 0$, where σ is an automorphism of A, then both g and δ are characterized as specific generalized σ-derivations on a nonzero ideal of A.
Another contribute via Daniel Eremita, Ilja Gogić, and Dijana Ilišević [10]. They suppose R is a semiprime ring with the maximal right ring of quotients Q_{mr}. An additive map $d\colon R \to Q_{mr}$ is called a generalized skew derivation if there exists a ring endomorphism $\sigma\colon R \to R$ and a map $\delta\colon R \to Q_{mr}$ such that $d(xy) = \delta(x)y + \sigma(x)d(y)$ for all $x, y \in R$. If σ is surjective, they determined the structure of generalized skew derivations for which there exists a finite number of elements $a_i, b_i \in Q_{mr}$ such that $d(x) = a_1xb_1 + \ + a_nxb_n$ for all $x \in R$.
Recently, Ilja Gogić [11] states that if A is a ring and $\sigma\colon A \to A$ a ring endomorphism. A generalized skew (or σ-)derivation of A is an additive map $d\colon A \to A$ for which there exists a map $\delta\colon A \to A$ such that $d(xy) = \delta(x)y + \sigma(x)d(y)$ for all $x, y \in A$. Also, if A is a prime C^*-algebra and σ is surjective, he determined the structure of generalized σ-derivations of A that belong to the cb-norm closure of elementary operators $E(A)$ on A, all such maps are of the form $d(x) = bx + axc$ for suitable elements a, b, c of the multiplier algebra $M(A)$. As a consequence, if an epimorphism $\sigma\colon A \to A$ lies in the cb-norm closure of $E(A)$, then σ must be an inner automorphism. Additionally, he observed that these results cannot be extended

even to relatively well-behaved non-prime C^*-algebras like $C(X, \mathbb{M}_2)$.
The purpose of this paper is to obtain some result for Tsiu.-K wen Lee [7] from a different point of view. More precisely, we shall study via using zero commutative ring.

1.1. **Preliminaries.** Throughout this presentation, an additive map $d\colon A \to A$ is called a derivation if the Leibniz's rule $d(xy) = d(x)y + xd(y)$ holds for all $x, y \in A$. Also, let σ be an automorphism of a prime ring A. An additive map $\delta\colon A \to Q_{ml}$ is called a σ-derivation if $\delta(xy) = \sigma(x)\delta(y) + \delta(x)y$ for all $x, y \in A$.
Furthermore, a ring A is called zero commutative if for $a, b \in A, ab = 0$ implies $ba = 0$ (used the term reversible for what is called zero commutative).
Recall that a ring R said to be prime if $aRb = 0$ implies $a = 0$ or $b = 0$, and R is semiprime if $aRa = 0$ implies $a = 0$. Let $n > 1$ be an integer. A ring R is said to be n-torsion free, if $nx = 0$ implies $x = 0$ for all $x \in R$. The symbols $x \circ y$ and $[x, y]$, where $x, y \in R$, stand for the anti-commutator $xy + yx$ and commutator $xy - yx$, respectively. An additive map $g\colon A \to Q_{ml}$ is called a generalized σ-derivation, where σ be an automorphism of A, if there exists an additive map $\delta\colon A \to Q_{ml}$ such that $g(xy) = \sigma(x)g(y) + \delta(x)y$ for all $x, y \in A$. It is clear that δ is uniquely determined by g, which is called the associated additive map of g. It is easy to check that δ is always a σ-derivation (see (*cf.* [8]))
Here we give some well-known results which will be used throughout the next section of the paper.

Lemma 1.1. [12, Theorem 1] *Let $a, b \in S$ such that $axb = bxa$ for all $x \in R$ (hence for all $s \in S$). Then a and b are C-dependent, where $S = RC$, a subring of Q containing R. We shall call C the extended centroid of R and S the central closure of R.*

Lemma 1.2. [7, Lemma 2. 1] *Let U be a nonzero ideal of R and let $f\colon U \to Q_{ml}$ be a left U-module map. Then there exists $q \in Q_{ml}$ such that $f(x) = xq$ for all $x \in U$, where R is a prime ring whose symmetric Martindale quotient ring Q contains a nontrivial idempotent.*

2. Main results

We now state the main result of this paper:

Theorem 2.1. *Let R be a zero commutative prime ring with extended Z and $z \in R$. Suppose that R possesses a nontrivial idempotent and $\kappa\colon R \to R$ is an additive map such that $x\kappa(y) + \kappa(x)y + xzy = 0$ for all $x, y \in R$ and κ acts as a left centralizer of R, then there exist a derivation $d\colon R \to RZ$ and $m \in Z$ such that $\kappa(x) = d(x) + (m - z)x$ for all $x \in R$.*

Proof. We are given that $x\kappa(y) + \kappa(x)y + xzy = 0$ for all $x, y \in R$ and R is a zero commutative prime ring with extended Z, then
$x(\kappa(y) + zy) + \kappa(x)y = 0$. In view of [7, Theorem 1.1] there exist a derivation $d\colon R \to Q, a \in Q_{ml}$ and $m \in Q_{ml}$ such that $\kappa(x) = d(x) + ax$ and $\kappa(x) + cx = d(x) + xm$ for all $x \in R$.
If we select a dense ideals which are a right ideal α and a left ideal λ of R.
It now follows from the above expression that,
$a\alpha \in R$ and $\lambda \in R$. Suppose that $x \in \alpha, w \in R$ and $y \in \lambda$. we obviously see that $xwy \in \alpha \cap \lambda$ which yield $(a + zx), ym \in R$ with $((a + z)x)wy = xw(ym)$. Arguing as Lemma 1.1, we are forced to conclude that
y and ym are $Z-$dependent for $y \in \lambda$. We have only to prove assertion $m \in Z$ which is easy. Thus, we conclude that $a = m-z$. One can see from the lest expression $\kappa(x) = d(x)+(m-z)x$ for all $x \in R$. Consequently, it follows that $d(R) \subset RZ$ which completes the proof of our theorem. $\square$

Corollary 2.2. *Let R be a 2-torsion free prime ring with extended Z and $z \in R$. Suppose that R possesses a nontrivial idempotent and $\kappa\colon R \to R$ is an additive map such that $x\kappa(y) + \kappa(x)y + xzy = 0$ for all $x, y \in R$ and κ acts as a left centralizer of R, then there exist a derivation $d\colon R \to RZ$ and $m \in Z$ such that $\kappa(x) = d(x) + (m - z)x$ for all $x \in R$.*

Proof. By given assumption, we have the additive map $\kappa\colon R \to R$ acts as a left centralizer and satisfying the following relation

$$x\kappa(y) + \kappa(x)y + xzy = 0 \tag{1}$$

We obviously get
$x\kappa(y)+\kappa(xy)+xzy = x(2\kappa(y)+zy)=0$. In light of the primeness of P, we find that

$$(2) \qquad 2\kappa(y) = -zy$$

Using (2) in (1), we arrive to
$2\kappa(x)y = 0$ for all $x, y \in R$. Indicate to R is 2-torsion free with employing the primeness of R, we conclude that

$$(3) \qquad \kappa(x) = 0.$$

In view of [7, Theorem 1.1], there exist a derivation $d\colon R \to RZ$ and $m\mathbb{Z}$, such that

$$(4) \qquad \kappa(x) = d(x) + (m-z)x$$

Using (3) together with (4), we arrive at $d(x) = (z-m)x$ for all $x \in R$ which completes the proof of our corollary. ☐

Theorem 2.3. *Let R be a zero commutative prime ring with extended Z and $z \in R$. Suppose that its symmetric Martindale quotient ring Q contains a nontrivial idempotent. If d and g are two operators of R satisfying $xd(y)+g(x)y = 0$ for all $x, y \in R$ then there exist $a, b, q \in RZ$ such that $d(x) = [q, y] + yb$ and $g(x) = [q, x] + ax$ for all $x, y \in R$.*

Proof. We are given that d and g are elementary operators, there exist finitely many $a_i, b_i, c_i, d_i \in RZ$ such that
$d(x) = \sum_i a_i x b_i$ and $g(x) = \sum_j c_j x d_j$ for all $x \in R$.
By view of [7, Theorem 1.1], there exist a nonzero ideal U of R, a derivation $d\colon R \to Q$ and elements $a \in Q_{mr}, b \in Q_{ml}$ such that
$\sum_i a_i y b_i = d(y) + yb$ and $\sum_j c_j x d_j = d(x) + ax$ for all $x, y \in U$.
We know that d can be uniquely extended to derivation $d\colon Q_{ml} \to Q_{ml}$ and for all $y \in Q_{ml}$, we can verify that

$$(5) \qquad \sum_i a_i y b_i = d(y) + yb$$

Writing $y = 1$ in this relation, we deduce that $b = \sum_i a_i b_i \in RZ$.
An analogous argument proves $a \in RZ$, so $d(RZ) \subseteq RZ$. In view of [13, Lemma 1], we find that d is $X-$inner, that is, there exists $q \in Q$ such that $d(y) = [q, y]$ for all $y \in R$. Consequently, we arrive to
$qy - yq = \sum_i a_i y b_i - yb$. Applying Lemma 1.1, we see that
q lies in the $Z-$linear span of the elements b_i's, b and 1. Then it is easy to show that $q \in RZ + Z$. If we choose $\beta \in Z$, then we find that
$[q+\beta, y] = [q, y]$ for all $y \in R$. Take $q \in RZ$, we get the required result. ☐

Corollary 2.4. *Let R be a zero commutative prime ring with extended Z and $z \in R$. Suppose that its symmetric Martindale quotient ring Q contains a nontrivial idempotent. If d and g are two operators of R satisfying $xd(y)+g(x)y = 0$ for all $x, y \in R$ then there exist $a, b, q \in RZ$ such that $ax = xb$, where g (resp. d) acts as a centralizer map of R.*

Proof. From the hypothesis, we have d and g are two elementary operators of R satisfying the relation $xd(y)+g(x)y = 0$ for all $x, y \in R$. By reason of g acts as a centralizer of R this relation modifies to
$x(d(y) - g(y)) = 0$ for all $x, y \in R$. By using the primeness of R, we conclude that

$$(6) \qquad d(y) = -g(y).$$

According to Theorem 2.1, we find that $d(y) = [q, y] + yb$ and $g(x) = [q, x] + a$. Now, we replace x by y and employing relation (6), we deduce $ay = yb$. Also, using the same techniques as used above with d we get the required result. ☐

Proposition 2.5. *Let R be a prime ring, U be a non-zero ideal of R and $d, g\colon U \to Q_{ml}$ be an additive maps such that $d(x)y+\sigma(x)g(y) = 0$ for all $x, y \in U$, where σ is an automorphism of R then d is a right centralizer of R.*

Proof. In the main relation $d(x)y + \sigma(x)g(y) = 0$, replacing x by zx we find that $d(zx)y + \sigma(zx)g(y) = 0$ for all $x, y, z \in U$. Expanding this expression with using σ is automorphism, we are forced to get

$$d(zx)y + \sigma(z)\sigma(x)g(y) = 0. \tag{7}$$

Left-multiplying the main relation by $\sigma(z)$, we conclude that

$$\sigma(z)d(x)y + \sigma(z)\sigma(x)g(y) = 0. \tag{8}$$

Combining Equations (7) and (8), we arrive to
$(d(zx) - \sigma(z)d(x))y = 0$ for all $x, y, z \in U$. Replacing y by ry where $r \in R$, we find that $(d(zx) - \sigma(z)d(x))ry = 0$ for all $x, y, z \in U, r \in R$. In light of the primeness of R with using the fact that $U \neq 0$, we obviously see that
$d(zx) - \sigma(z)d(x) = 0$. In view of the hypothesis we find that σ is an automorphism of R, we observe that
$d(zx) = td(x)$ for all $x, z, t \in U$. Writing $z = t$, we are forced to conclude that d is a right centralizer mapping. By this result, we complete the proof of our proposition. □

Corollary 2.6. *Let R be a prime ring, U be a non-zero ideal of R and $d, g\colon U \to Q_{ml}$ be an additive maps such that $d(x)y + \sigma(x)g(y) = 0$ for all $x, y \in U$, where σ is an automorphism of Q_{ml} then $d(x) = xc, c \in Q_{ml}$.*

Proof. Applying the same strategy which used to prove Proposition 2.5, we arrive to $d(zx) - \sigma(z)d(x) = 0$ for all $x, z \in U$. We observe σ can be uniquely extended to an automorphism of Q_{ml}. Hence, we consider the map $\mu\colon U \to Q_{ml}$ which defined as
$\mu(x) = \sigma^{-1}(d(x))$ for all $x \in U$. Obviously, μ is a left $U-$module map.
In view of Lemma 1.2, we can write $b \in Q_{ml}$ such that $\mu(x) = xb$. Consequently, $d(x) = \sigma(x)a$ where $a = \sigma(b) \in Q_{ml}$.
Hence, we conclude that $d(x) = \sigma(x)\sigma(b)$. Again, by reason of σ is an automorphism of Q_{ml}. This relation reduces to $d(x) = \sigma(x)c$, such that $c \in Q_{ml}$ which completes the proof of our corollary.

□

3. Conclusion

In this paper, we present a characterization of derivation of a zero commutative prime ring with extended Z. Also, the prime ring has a non-zero ideal and acts as a symmetric Martindale quotient ring Q contains a nontrivial idempotent. According to this characterization, we propose a prime ring possesses a nontrivial idempotent and allowed to satisfy some identities. The results of this article are variable according to the identities.

Acknowledgments

The author is grateful to Al-Mustansiriyah University, the Republic of Iraq and beholden to the reviewer(s) for his/ their accuracy with professionally reading the article.

References

[1] M. J. Atteya, *Generalized derivations of semiprime rings*, LAP Lambert Academic Publishing, Saarbrucken, Germany, 2012.

[2] M. J. Atteya, (2014), Some results on the identity $d(x) = \lambda x + \zeta(x)$, Journal of Mathematical Sciences and Applications, Serie VII, 2(3), 2832.

[3] G. A. Baker, (1959),A new derivation of Newtons identities and their application to the calculation of the eigenvalues of a matrix, Journal of the Society for Industrial and Applied Mathematics, 7, 143148.

[4] H. E. Bell and M. N. Daif, (1995), On derivations and commutativity in prime rings, Acta Math. Hungar., 66, 3

[5] M.A. Chebotar, W.-F. Ke and P.-H. Lee, (2004), Maps characterized by actions on zero products, Pacific J. Math., 216, 217–228.

[6] W. Jing, S. Lu and P. Li, (2002),Characterisations of derivations on some operator algebras, Bull. Austral. Math. Soc., 66, 227-232, MR 1932346 (2003f:47059), Zbl 1035.47019.

[7] Tsiu-Kwen Lee, (2004), Generalized skew derivations characterized by acting on zero products, Pacific J. Math., 216, 293-301.

[8] Tsiu-Kwen Lee and K.-S. Liu, Generalized skew derivations with algebraic values of bounded degree, Houston Journal of Mathematics, University of Houston, Volume 39, No. 3, (2013), 733-740.

[9] E.C. Posner, Derivations in prime rings,(1957), Proc. Amer. Math. Soc., 8, 1093- 1100.

[10] Daniel Eremita, Ilja Gogić, and Dijana Ilišević, (2014), Generalized skew derivations implemented by elementary operators, Algebras and Representation Theory, DOI:10.1007/s10468-013-9429-8.
[11] Ilja Gogić, (2021), The cb-norm approximation of generalized skew derivations by elementary operators,Linear and Multilinear Algebra, 69:8, 1543-1556, DOI: 10.1080/03081087.2019.1632784
[12] W.S. Martindale III, (1969), Prime rings satisfying a generalized polynomial identity, J. Al-gebra, 12, 576584, MR 0238897 (39, 257), Zbl 0175.03102.
[13] V.K. Kharchenko, (1978), Differential identities of prime rings, Algebra i Logika, 17, 220238; Engl. Transl., Algebra and Logic, 17 (1978), 154168, MR 0541758(81f:16025), Zbl 0423.16011.

Department of Mathematics, College of Education,
Al-Mustansiriyah University, Baghdad, Iraq

Email(s): mehsinatteya88@uomustansiriyah.edu.iq

Median-Based Robust Calibration Estimation of Population Mean in Stratified Sampling in the Presence of Outlier

Oluwagbenga Tobi Babatunde [1*] and Abimibola Victoria Oladugba [2]

2020 Mathematics Subject Classifications: 62D05

Keywords: Calibration Estimation, Stratified sampling, Outlier, Media, Robust

1. Introduction

Calibration estimation is a method that utilizes auxiliary information to improve estimate of population parameters and it has received so much attention in literature since the work of ((*cf.* [3]). Calibration estimation involves the minimization of a function subject to one or more constraints ((*cf.* [4]) and was introduced in ((*cf.* [3]). The work of ((*cf.* [3]) has motivated other works in several sampling schemes (see (*cf.* [5, 6, 15, 7, 17, 8])
Under the stratified sampling scheme, several calibration estimators have been proposed using different parameters of the auxiliary variable starting from the work of (*cf.* [18]). The mean and variance were used in (*cf.* [18]), the mean was used in (*cf.* [12, 16]), the mean and coefficient of variation were used in (*cf.* [13, 11]) while coefficient of variation was used in (*cf.* [4]). These parameters of the auxiliary variable are sensitive to outliers (*cf.* [9]).
According to (*cf.* [2]), an observation which is either low or high and differs significantly from other observations is known as an outlier. (*cf.* [10]) observed that the presence of outlying observation may significantly lead to erroneous result. (*cf.* [19]) also noted that the presence of either low or high extreme values decreases the efficiency of an estimator.
Several ratio estimators under stratified sampling scheme have been proposed in literature to deal with the effect of outliers using different parameters of the auxiliary variables. This study proposes a robust calibration estimator that deals with the effect of either low or high outlying observation using median of the auxiliary variable. The proposed estimator was evaluated and compared with five existing classical estimators of (*cf.* [18, 12, 13, 16, 11, 4]) using the absolute bias (*ABS*) and relative root mean square error (*RRMSE*)

2. Definition of Notations

Given a finite population L which is divided into h non-overlapping homogeneous strata. Let Y and X be the study and auxiliary variables respectively. To estimate the mean of Y, a random sample of size n is drawn from Y and X. The sample and population mean of Y are $\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{hi}$ and $\bar{Y}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} y_{hi}$ respectively. The sample mean, variance, coefficient of variation and median of X are $\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}$, $s_{hx}^2 = \frac{\sum_{i=1}^{n_h}(x_{hi}-\bar{x}_h)^2}{n_h-1}$, $c_{hx} = \frac{s_{hx}}{\bar{x}_h}$ and m_h respectively. The population mean, variance, coefficient of variation and median of X are $\bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} x_{hi}$, $S_{hx}^2 = \frac{\sum_{i=1}^{N_h}(x_{hi}-\bar{X}_h)^2}{N_h-1}$, $C_{hx} = \frac{S_{hx}}{\bar{X}_h}$ and M_h respectively.
The population mean of stratified sampling is given by:

$$\bar{y}_{st} = \sum_{h=1}^{J} W_h \bar{y}_h \tag{1}$$

The calibration estimator is obtained by minimizing (2) subject to different calibration constraints.

$$\sum_{h=1}^{J} \frac{(\Omega_h - W_h)^2}{W_h Q_h} \tag{2}$$

where, Ω_h are the calibrated weights and Q_h are weights for obtaining different types of the calibration estimator (*cf.* [1]).

3. Existing Calibration Estimators

Different calibration estimators have been proposed in literature for estimating the population mean of a stratified sampling using one auxiliary variable. Some of the existing calibration estimators are reviewed below.

3.1. **Tracey et al. (2003).** The estimator proposed in (*cf.* [18]) is defined by (3)

$$\bar{y}_T = \sum_{h=1}^{J} W_h \bar{y}_h + \widehat{\delta}_{1(T)} \left(\sum_{h=1}^{J} W_h \left(\bar{X}_h - \bar{x}_h \right) \right) + \widehat{\delta}_{2(T)} \left(\sum_{h=1}^{J} W_h \left(S_{hx}^2 - s_{hx}^2 \right) \right) \tag{3}$$

where,

$$\widehat{\delta}_{1(T)} = \left[\frac{\left(\sum_{h=1}^{J} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h \bar{y}_h\right) - \left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h s_{hx}^2\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{y}_h s_{hx}^2\right)}{\left(\sum_{h=1}^{J} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_{hx}^2\right) - \left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h s_{hx}^2\right)^2} \right]$$

$$\widehat{\delta}_{2(T)} = \left[\frac{\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_{hx}^2\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{y}_h s_{hx}^2\right) - \left(\sum_{h=1}^{J} W_h Q_h s_{hx}^2\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h \bar{y}_h\right)}{\left(\sum_{h=1}^{J} W_h Q_h s_{hx}^4\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_{hx}^2\right) - \left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h s_{hx}^2\right)^2} \right]$$

3.2. **Nidhi et al. (2007).** The estimator introduced in (*cf.* [12]) is defined by (4)

$$\bar{y}_N = \sum_{h=1}^{J} W_h \bar{y}_h + \widehat{\delta}_N \left(\sum_{h=1}^{J} W_h \left(\bar{X}_h - \bar{x}_h \right) \right) \tag{4}$$

where,

$$\widehat{\delta}_N = \left[\frac{\left(\sum_{h=1}^{J} W_h Q_h\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h \bar{y}_h\right) - \left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{y}_h\right)}{\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_{hx}^2\right)\left(\sum_{h=1}^{J} W_h Q_h\right) - \left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h\right)^2} \right]$$

3.3. **Rao et al. (2016).** Two different estimators were proposed in (*cf.* [13]) and defined by (5) and (6) respectively.

$$\bar{y}_{R1} = \sum_{h=1}^{J} W_h \bar{y}_h + \widehat{\delta}_{R1} \left[\sum_{h=1}^{J} W_h \left(\bar{X}_h + C_{hx} \right) - \sum_{h=1}^{J} W_h \left(\bar{x}_h + c_{hx} \right) \right] \tag{5}$$

$$\bar{y}_{R2} = \sum_{h=1}^{J} W_h \bar{y}_h + \widehat{\delta}_{R2} \left[\sum_{h=1}^{J} W_h \left(1 + \bar{X}_h + C_{hx} \right) - \sum_{h=1}^{J} W_h \left(1 + \bar{x}_h + c_{hx} \right) \right] \tag{6}$$

where

$$\widehat{\delta}_{R1} = \left[\frac{\sum_{h=1}^{J} W_h Q_h \bar{y}_h \left(\bar{x}_h + c_{hx} \right)}{\sum_{h=1}^{J} W_h Q_h \left(\bar{x}_h + c_{hx} \right)^2} \right]$$

$$\widehat{\delta}_{R2} = \left[\frac{\sum_{h=1}^{J} W_h Q_h \bar{y}_h \left(1 + \bar{x}_h + c_{hx} \right)}{\sum_{h=1}^{J} W_h Q_h \left(1 + \bar{x}_h + c_{hx} \right)^2} \right]$$

3.4. **Sisodia et al. (2017).** The estimator proposed in (*cf.* [16]) is defined by (7)

$$\bar{y}_S = \sum_{h=1}^{J} W_h \left[\bar{y}_h + \widehat{\delta}_S \left(\bar{X}_h - \bar{x}_h \right) \right] \tag{7}$$

where,

$$\widehat{\delta}_S = \left[\frac{\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h \bar{y}_h\right)\left(\sum_{h=1}^{J} W_h Q_h\right) - \left(\sum_{h=1}^{J} W_h Q_h \bar{y}_h\right)\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h\right)}{\left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h^2\right)\left(\sum_{h=1}^{J} W_h Q_h\right) - \left(\sum_{h=1}^{J} W_h Q_h \bar{x}_h\right)^2} \right]$$

3.5. **Ozqul (2018).** The estimator proposed in (*cf.* [11]) is given by (8)

$$\bar{y}_O = \sum_{h=1}^{J} W_h \left[\bar{y}_h + \widehat{\delta}_{O1} \left(\bar{X}_h - \bar{x}_h \right) + \widehat{\delta}_{O2} \left(C_{hx} - c_{hx} \right) \right] \tag{8}$$

where, $\delta_{O1} = \frac{\gamma_1}{\eta}$ and $\delta_{O2} = \frac{\gamma_2}{\eta}$

$$\begin{aligned}\gamma_1 =& \left(\textstyle\sum_{h=1}^J W_h Q_h \bar{y}_h\right)\left[\left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h c_{hx}\right) - \left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}^2\right)\right] \\ &+ \left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h \bar{y}_h\right)\left[\left(\textstyle\sum_{h=1}^J W_h Q_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}^2\right) - \left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}\right)^2\right] \\ &+ \left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx} \bar{y}_h\right)\left[\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}\right) - \left(\textstyle\sum_{h=1}^J W_h Q_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h c_{hx}\right)\right]\end{aligned}$$

$$\begin{aligned}\gamma_2 =& \left(\textstyle\sum_{h=1}^J W_h Q_h \bar{y}_h\right)\left[\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h c_{hx}\right) - \left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_{hx}^2\right)\right] \\ &+ \left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h \bar{y}_h\right)\left[\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}\right) - \left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h c_{hx}\right)\left(\textstyle\sum_{h=1}^J W_h Q_h\right)\right] \\ &+ \left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx} \bar{y}_h\right)\left[\left(\textstyle\sum_{h=1}^J W_h Q_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_{hx}^2\right) - \left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h\right)^2\right]\end{aligned}$$

$$\begin{aligned}\eta =& \left(\textstyle\sum_{h=1}^J W_h Q_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h^2\right)\left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}^2\right) - \left(\textstyle\sum_{h=1}^J W_h Q_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h c_{hx}\right) \\ &- \left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h^2\right)\left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}\right)^2 - \left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}^2\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h\right)^2 \\ &+ 2\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h\right)\left(\textstyle\sum_{h=1}^J W_h Q_h c_{hx}\right)\left(\textstyle\sum_{h=1}^J W_h Q_h \bar{x}_h c_{hx}\right)\end{aligned}$$

3.6. **Garg and Pachori (2019).** The estimator introduced by (*cf.* [4]) is defined by (9)

$$\bar{y}_G = \sum_{h=1}^{J} W_h \bar{y}_h + \widehat{\delta}_G \left[\sum_{h=1}^{J} W_h \left(C_{hx} - c_{hx} \right) \right] \tag{9}$$

where,

$$\widehat{\delta}_G = \left[\frac{\left(\sum_{h=1}^J W_h Q_h\right)\left(\sum_{h=1}^J W_h Q_h c_{hx} \bar{y}_h\right) - \left(\sum_{h=1}^J W_h Q_h c_{hx}\right)\left(\sum_{h=1}^J W_h Q_h \bar{y}_h\right)}{\left(\sum_{h=1}^J W_h Q_h c_{hx}^2\right)\left(\sum_{h=1}^J W_h Q_h\right) - \left(\sum_{h=1}^J W_h Q_h c_{hx}\right)^2} \right]$$

4. Proposed Calibration Estimator

A robust calibration estimator for dealing with the effect of outlier using the median of the auxiliary variable was proposed and defined by (10)

$$\bar{y}_M = \sum_{h=1}^{J} W_h \left[\bar{y}_h + \widehat{\delta}_M \left(M_h - m_h \right) \right] \tag{10}$$

where,

$$\widehat{\delta}_M = \left[\frac{\left(\sum_{h=1}^J W_h Q_h m_h \bar{y}\right)\left(\sum_{h=1}^J W_h Q_h\right) - \left(\sum_{h=1}^J W_h Q_h \bar{y}_h\right)\left(\sum_{h=1}^J W_h Q_h m_h\right)}{\left(\sum_{h=1}^J W_h Q_h m_h^2\right)\left(\sum_{h=1}^J W_h Q_h\right) - \left(\sum_{h=1}^J W_h Q_h m_h\right)^2} \right]$$

Different values of Q_h were suggested in (*cf.* [4]) as 1, $\frac{1}{\bar{x}}$ and $\frac{1}{c_{hx}}$. These values were adopted to obtained three different types of the proposed estimator.
Type 1: $Q_h = 1$

$$\bar{y}_{M1} = \sum_{h=1}^{J} W_h \left[\bar{y}_h + \widehat{\delta}_{M1} \left(M_h - m_h \right) \right] \tag{11}$$

where,

$$\widehat{\delta}_{M1} = \left[\frac{\left(\sum_{h=1}^J W_h m_h \bar{y}\right)\left(\sum_{h=1}^J W_h\right) - \left(\sum_{h=1}^J W_h \bar{y}_h\right)\left(\sum_{h=1}^J W_h m_h\right)}{\left(\sum_{h=1}^J W_h m_h^2\right)\left(\sum_{h=1}^J W_h\right) - \left(\sum_{h=1}^J W_h m_h\right)^2} \right]$$

Type 2: $Q_h = \frac{1}{\bar{x}}$

$$\bar{y}_{M2} = \sum_{h=1}^{J} W_h \left[\bar{y}_h + \widehat{\delta}_{M2} \left(M_h - m_h \right) \right] \tag{12}$$

where,

$$\widehat{\delta}_{M2}=\left[\frac{\left(\sum_{h=1}^{J}\frac{W_h m_h \bar{y}}{\bar{x}}\right)\left(\sum_{h=1}^{J}\frac{W_h}{\bar{x}}\right)-\left(\sum_{h=1}^{J}\frac{W_h \bar{y}_h}{\bar{x}}\right)\left(\sum_{h=1}^{J}\frac{W_h m_h}{\bar{x}}\right)}{\left(\sum_{h=1}^{J}\frac{W_h m_h^2}{\bar{x}}\right)\left(\sum_{h=1}^{J}\frac{W_h}{\bar{x}}\right)-\left(\sum_{h=1}^{J}\frac{W_h m_h}{\bar{x}}\right)^2}\right]$$

Type 3: $Q_h = \frac{1}{c_{hx}}$

$$\bar{y}_{M3} = \sum_{h=1}^{J} W_h\left[\bar{y}_h + \widehat{\delta}_{M3}\left(M_h - m_h\right)\right] \tag{13}$$

where,

$$\widehat{\delta}_{M3}=\left[\frac{\left(\sum_{h=1}^{J}\frac{W_h m_h \bar{y}}{c_{hx}}\right)\left(\sum_{h=1}^{J}\frac{W_h}{c_{hx}}\right)-\left(\sum_{h=1}^{J}\frac{W_h \bar{y}_h}{c_{hx}}\right)\left(\sum_{h=1}^{J}\frac{W_h m_h}{c_{hx}}\right)}{\left(\sum_{h=1}^{J}\frac{W_h m_h^2}{c_{hx}}\right)\left(\sum_{h=1}^{J}\frac{W_h}{c_{hx}}\right)-\left(\sum_{h=1}^{J}\frac{W_h m_h}{c_{hx}}\right)^2}\right]$$

5. Performance Measures

The performance of the proposed estimator over selected existing estimators was assessed using the absolute bias (*ABS*) and relative root mean square error (*RRMSE*) which are defined in (14) and (15) respectively. An estimator with the least *ABS* and *RRMSE* is judged to have higher efficiency over other estimators. The *ABS* and *RRMSE* of the existing and proposed calibration estimators ($\bar{y}_T, \bar{y}_N, \bar{y}_{R1}, \bar{y}_{R2}, \bar{y}_S, \bar{y}_O, \bar{y}_G, \bar{y}_M$) were obtained and presented in Tables 1 and 2 and Table 3 and 4 respectively.

$$ABS\left(\bar{y}_j\right) = \left|\frac{1}{50000}\sum_{r=1}^{50000}\left(\left(\bar{y}_j\right)_r - \bar{Y}\right)\right| \tag{14}$$

$$RRMSE\left(\bar{y}_j\right) = \sqrt{\frac{1}{50000}\sum_{r=1}^{50000}\left(\left(\left(\bar{y}_j\right)_r - \bar{Y}\right)/\bar{Y}\right)^2} \tag{15}$$

6. Empirical Study

The performance and efficiency of the proposed estimator were assessed using the CO124 population obtained from (*cf.* [14]). The study and auxiliary variables are the 1983 and 1980 populations (in millions) of 95 countries from 3 continents (Africa, Asia, and Europe) respectively. Two different extreme observations classified as low and high were introduced separately to the samples selected from the auxiliary variable using a random number generated from R software. The low extreme values range from 0.01 to 0.1 while the high extreme values range from 1500 to 2000. The results of the *ABS* of the calibration estimators obtained under the different values of $Q_h$for low and high outlying observations are presented in Tables 1 and 2 respectively while the results of the *RRMSE* of the calibration estimators obtained under the different values of $Q_h$for low and high outlying observations are presented in Tables 3 and 4 respectively. From Table 1, we observed that in the presence of low outlying observation, the *ABS* of the proposed estimator were the least under $Q_h = 1$ and $Q_h = \frac{1}{\bar{x}}$ compared to the estimators proposed in (*cf.* [18, 12, 13, 16, 11, 4]) while the *ABS* of the estimator proposed in (*cf.* [4]) was the least for $Q_h = \frac{1}{c_{hx}}$. This indicates that when there is low outlying observation in the auxiliary variable, the proposed calibration estimator will produce an estimate of the population mean which is closest to the true population mean for $Q_h = 1$ and $Q_h = \frac{1}{\bar{x}}$. For the case of high outlying observation, the results in Table 2 show that the *ABS* of the proposed estimator were the least under the three values of Q_h considered compared to the estimators proposed in (*cf.* [18, 12, 13, 16, 11, 4]). This indicates that when there is high outlying observation in the auxiliary variable, the proposed calibration estimator will produce an estimate of the population mean which is closest to the true population mean for all the values of Q_h.

The results in Table 3 showed that the *RRMSE* of the proposed estimator in the presence of low outlying observation under $Q_h = 1$ and $Q_h = \frac{1}{\bar{x}}$ were the least compared to the estimators proposed in (*cf.* [18, 12, 13, 16, 11, 4]) while the *RRMSE* of the estimator proposed in (*cf.* [4]) was the least for $Q_h = \frac{1}{c_{hx}}$. These results indicate that when there is a low outlying observation in the auxiliary variable, the proposed calibration estimator is more efficient for $Q_h = 1$ and $Q_h = \frac{1}{\bar{x}}$. For the case of high outlying observation, the results in Table 4 show that the *RRMSE* of the proposed estimator were the least under the three values of Q_h considered compared to the estimators proposed in (*cf.* [18, 12, 13, 16, 11, 4]). This indicates that when there is high outlying observation in the auxiliary variable, the proposed calibration estimator is more efficient for all the values of Q_h.

Table 1: Absolute Bias of the Calibration Estimators (Low Outlying Observation)

	$Q_h = 1$	$Q_h = \frac{1}{\bar{x}}$	$Q_h = \frac{1}{c_{hx}}$
$ABS(\bar{y}_T)$	2.1460×10^{2}	1.4300×10^{2}	1.7200×10^{2}
$ABS(\bar{y}_N)$	5.2980×10^{-1}	5.1990×10^{-1}	5.3950×10^{-1}
$ABS(\bar{y}_{R1})$	5.4870×10^{-1}	5.4740×10^{-1}	5.4620×10^{-1}
$ABS(\bar{y}_{R2})$	5.4280×10^{-1}	5.4030×10^{-1}	5.4060×10^{-1}
$ABS(\bar{y}_S)$	8.6200×10^{-2}	4.0360×10^{-1}	4.0700×10^{-1}
$ABS(\bar{y}_O)$	3.1460×10^{-1}	3.1470×10^{-1}	3.1450×10^{-1}
$ABS(\bar{y}_G)$	7.5700×10^{-2}	6.4400×10^{-2}	6.1300×10^{-2}
$ABS(\bar{y}_{AS})$	8.5766×10^{-1}	9.9310×10^{-1}	6.8980×10^{-1}
$ABS(\bar{y}_M)$	4.8900×10^{-2}	4.7000×10^{-2}	7.7700×10^{-2}

Table 2: Absolute Bias of the Calibration Estimators (High Outlying Observation)

	$Q_h = 1$	$Q_h = \frac{1}{\bar{x}}$	$Q_h = \frac{1}{c_{hx}}$
$ABS(\bar{y}_T)$	6.6927×10^{0}	4.7614×10^{0}	5.6101×10^{0}
$ABS(\bar{y}_N)$	1.9160×10^{-1}	1.8050×10^{-1}	2.6000×10^{-1}
$ABS(\bar{y}_{R1})$	4.9400×10^{-1}	7.6500×10^{-1}	5.4590×10^{-1}
$ABS(\bar{y}_{R2})$	4.9320×10^{-1}	7.6360×10^{-1}	5.4540×10^{-1}
$ABS(\bar{y}_S)$	8.6200×10^{-2}	4.0360×10^{-1}	4.0700×10^{-1}
$ABS(\bar{y}_O)$	7.5800×10^{-2}	8.1300×10^{-2}	7.8900×10^{-2}
$ABS(\bar{y}_G)$	1.3300×10^{-1}	2.3460×10^{-1}	1.6850×10^{-1}
$ABS(\bar{y}_{AS})$	6.3554×10^{0}	3.9660×10^{0}	4.4710×10^{0}
$ABS(\bar{y}_M)$	4.5000×10^{-3}	6.5000×10^{-3}	9.7000×10^{-3}

Table 3: Relative Root Mean Square Error of the Calibrated Estimators (Low Outlying Observation)

	$Q_h = 1$	$Q_h = \frac{1}{\bar{x}}$	$Q_h = \frac{1}{c_{hx}}$
$RRMSE(\bar{y}_T)$	1.9197×10^{7}	8.5239×10^{6}	1.2441×10^{7}
$RRMSE(\bar{y}_N)$	1.1702×10^{2}	1.1268×10^{2}	1.2134×10^{2}
$RRMSE(\bar{y}_{R1})$	1.2551×10^{2}	1.2492×10^{2}	1.2438×10^{2}
$RRMSE(\bar{y}_{R2})$	1.2285×10^{2}	1.2169×10^{2}	1.2182×10^{2}
$RRMSE(\bar{y}_S)$	6.4136×10^{0}	1.4043×10^{2}	1.4282×10^{2}
$RRMSE(\bar{y}_O)$	4.1249×10^{1}	4.1293×10^{1}	4.1224×10^{1}
$RRMSE(\bar{y}_G)$	2.3895×10^{0}	1.7284×10^{0}	1.5645×10^{0}
$RRMSE(\bar{y}_{AS})$	3.0662×10^{2}	4.1115×10^{2}	1.9835×10^{2}
$RRMSE(\bar{y}_M)$	9.9820×10^{-1}	9.4000×10^{-3}	2.5163×10^{0}

Table 4: Relative Root Mean Square Error of the Calibrated Estimators (High Outlying Observation)

	$Q_h = 1$	$Q_h = \frac{1}{\bar{x}}$	$Q_h = \frac{1}{c_{hx}}$
$RRMSE(\bar{y}_T)$	1.8671×10^{4}	9.4503×10^{3}	1.3119×10^{4}
$RRMSE(\bar{y}_N)$	1.5305×10^{1}	1.3574×10^{1}	2.1290×10^{-1}
$RRMSE(\bar{y}_{R1})$	1.0170×10^{2}	2.4393×10^{2}	1.2421×10^{2}
$RRMSE(\bar{y}_{R2})$	1.0139×10^{2}	2.4303×10^{2}	1.2398×10^{2}
$RRMSE(\bar{y}_S)$	6.4136×10^{0}	1.4043×10^{2}	1.4282×10^{2}
$RRMSE(\bar{y}_O)$	2.3940×10^{0}	2.7552×10^{0}	2.5975×10^{0}
$RRMSE(\bar{y}_G)$	7.3782×10^{0}	2.2937×10^{1}	1.1830×10^{1}
$RRMSE(\bar{y}_{AS})$	1.8671×10^{4}	9.4503×10^{3}	1.3119×10^{4}
$RRMSE(\bar{y}_M)$	8.3000×10^{-3}	1.7500×10^{-2}	3.9000×10^{-2}

7. Conclusion

In this paper, a new improved median-based calibration estimator for the population mean of a stratified sampling in the presence of outlier in the auxiliary variable is introduced. The median is robust to the presence of outliers (*cf.* [9]) and this informed the choice of the median in defining the calibration constraints. The results of the empirical study show that the *ABS* and *RRMSE* of the proposed estimator in the presence of a low outlying observation were the least for $Q_h = 1$ and $Q_h = \frac{1}{\bar{x}}$ compared to the estimators proposed in (*cf.* [18, 12, 13, 16, 11, 4]). This suggests that the proposed estimator is robust and more efficient in the presence of a low outlying observation compared to existing calibration estimators considered in this work. Also, the *ABS* and *RRMSE* of the proposed estimator in the presence of a high outlying observation were the least for all the values of Q_h compared to the estimators proposed in (*cf.* [18, 12, 13, 16, 11, 4]). This suggests that the proposed estimator is robust and more efficient in the presence of a high outlying observation compared to existing calibration estimators considered in this work and it is therefore recommended for estimating the population mean of stratified sampling when there is presence of either low or high outlying observation in the auxiliary variable.

References

[1] Alam, S., and Shabbir, J. (2020), Calibration estimation of mean by using double use of auxiliary information, Communications in Statistics – Simulations and Computation DOI: 10.1080/03610918.2020.1749660

[2] Cousineau, D. and Chartier, S. (2010), Outliers detection and treatment: a review, International Journal of Psychological Research, 3(1), 59-68.

[3] Deville, J. C., and Sarndal, C., -E. (1992), Calibration estimators in survey sampling, Journal of the American Statistical Association 87(418) 376-382.

[4] Garg, N., and Pachori, M. (2019), Use of coefficient of variation in calibration estimation of population mean in stratified sampling, Communications in Statistics - Theory and Methods DOI: 10.1080/03610926.2019.1622729

[5] Hidiroglou, M., and Sarndal, C. (1998), Use of auxiliary Information for two-phase sampling, Survey Methodology, 24: 11-20.

[6] Kim, J. M., Sungur, E. A., and Heo, T. Y. (2007), Calibration approach estimators in stratified sampling, Statistics and Probability Letters 77(1) 99-103.

[7] Koyuncu, N., and Kadilar, C. (2013), Calibration estimator using different distance measures in stratified random sampling, International Journal of Modern Engineering Research, 3(1) 415-419.

[8] Koyuncu, N., and Kadilar, C. (2016). Calibration weighting in stratified random sampling, Communications in Statistics: Simulation and Computation 45(7), 2267-2275.

[9] Leys, C., Ley, C., Klein, O., Bernard, P. and Licata, L. (2013), Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, Journal of Experimental Social Psychology, 49(4), 764-766.

[10] Osborne, J. W. and Overbay, A. (2004), The power of outliers (and why researchers should always check for them), Practical Assessment, Research, and Evaluation, 9(6) https://pareonline.net/getvn.asp?v=9&n=6

[11] Ozgul, N. (2018), New calibration estimator based on two auxiliary variables in stratified sampling, Communication in Statistics – Theory and Methods DOI: 10.1080/03610926.2018.1433852

[12] Nidhi, B. V. S., Sisodia, S., Singh, S. K., and Singh (2017), Calibration approach estimation of the mean in stratified sampling stratified double sampling, Communication in Statistics-Theory and Methods, 46(10), 4932-4942.

[13] Rao, D., Tekabu, T., and Khan, M. (2016), New calibration estimators in stratified sampling, In Computer Science and Engineering (APWC on CSE), 3rd Asia-Pacific World Congress Nadi, Fijion, December 5-6, 66-70. IEEE.

[14] Sarndal, C. -E., Swensson, B. and Wretman, J. (1992), Model assisted survey sampling, New York: Springer-Verlag Publishing. https://doi.org/10.1007/978-1-4612-4378-6

[15] Singh, S., and Arnab, R. (2011), On Calibration of design weights, METRON: International Journal of Statistics, 69(2) 185-205.

[16] Sisodia, B. V. S., Singh, S., and Singh, S. K. (2017), Calibration approach estimation of the mean in stratified sampling and stratified double sampling, Communications in Statistics - Theory and Methods, 46(10), 4932-4942.

[17] Sud, U., Chandra, H. and Gupta, V. (2014), Calibration-based product estimator in single-and two-phase sampling, Journal of Statistical Theory and Practice, 8(1), 1-11.

[18] Tracey, D. S., Singh, S., and Arnab, R. (2003), Note on calibration in stratified and double Sampling, Survey Methodology, 29(1), 99-104.

[19] Zaman, T. and Bulut, H. (2020), Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling, Communication in Statistics-Theory and Methods, 49(14), 3407-3420.

Department of Statistics, University of Nigeria, Nsukka, Enugu-State, Nigeria. [1*] (corresponding author)

Department of Statistics, University of Nigeria, Nsukka, Enugu-State, Nigeria. [2]

Email(s): oluwagbenga.babatunde@unn.edu.ng [1*], abimibola.oladugba@unn.edu.ng [2]

A Note on a Hilfer-Hadamard Fractional Integro-Differential Inclusion

Aurelian Cernea

2020 Mathematics Subject Classifications: 34A60, 34A12, 34A08

Keywords: Differential inclusion, fractional derivative, boundary value problem.

1. Introduction

The recent literature is full of papers devoted to the study of systems governed by fractional order derivatives. The main reason is that the models taking into account fractional derivatives are more realistic than the models realized with classical derivatives (see [3, 8, 12, 13, 14] etc.).

A generalization of both Riemann-Liouville and Caputo fractional derivatives was introduced by Hilfer in [10]. In fact, this derivative interpolates between Riemann-Liouville and Caputo derivatives. Several properties and applications of Hilfer fractional derivative may be found in [11]. The modification of Hilfer fractional derivative resulted in the concept of Hilfer-Hadamard fractional derivative. Similarly, the Hilfer-Hadamard fractional derivative covers the cases of Riemann-Liouville-Hadamard and Caputo-Hadamard fractional derivatives.

In this note we consider the following problem

$$D_{HH}^{\alpha,\beta}x(t) \in F(t,x(t),V(x)(t)) \quad a.e.\ ([1,T]) \tag{1}$$

with boundary conditions of the form

$$x(1)=0, \quad x(T)=\sum_{j=1}^{m}\eta_j x(\xi_j)+\sum_{i=1}^{n}\zeta_i I_H^{\varphi_i}x(\theta_i)+\sum_{k=1}^{r}\lambda_k D_H^{\omega_k}x(\mu_k), \tag{2}$$

where $D_{HH}^{\alpha,\beta}$ is the Hilfer-Hadamard fractional derivative of order $\alpha \in (1,2]$ and type $\beta \in [0,1]$, $\xi_j,\theta_i,\mu_k \in (1,T)$, $\eta_i,\zeta_j,\lambda_k \in \mathbf{R}$, $j=\overline{1,m}, i=\overline{1,n}, k=\overline{1,r}$, I_H^{φ} is the Hadamard fractional integral of order $\varphi>0$, D_H^{ω} is the Hadamard fractional derivative of order $\omega>0$, $F:[1,T]\times\mathbf{R}\times\mathbf{R}\to\mathcal{P}(\mathbf{R})$ is a set-valued map and $V: C([1,T],\mathbf{R})\to C([1,T],\mathbf{R})$ is a nonlinear Volterra integral operator defined by $V(x)(t)=\int_0^t k(t,s,x(s))ds$ with $k(.,.,.):[1,T]\times\mathbf{R}\times\mathbf{R}\to\mathbf{R}$ a given function.

Our study is motivated by a recent paper [1]. Namely, in [1] three existence results for problem (1)-(2) may be found in the case when F does not depends on the last variable. All the results in [1] are proved by using several suitable theorems from fixed point theory.

The goal of this note is to obtain the existence of solutions for problem (1)-(2) in the case when the set-valued map F has nonconvex values but it is assumed to be Lipschitz in the second and third variable. Our result use Filippov's techniques ([9]); namely, the existence of solutions is obtained by starting from a given "quasi" solution. In addition, the result provides an estimate between the "quasi" solution and the solution obtained.

Our result improve an existence theorem in [1] in the case when the right-hand side is Lipschitz in the second variable. Moreover, our result may be viewed as a generalization to the case when the right-hand side contains a nonlinear Volterra integral operator. Even if the method we use here is known in the theory of differential inclusions (e.g., [4, 5, 6, 7] etc.) it is largely ignored by the authors that are dealing with such problems in favor of fixed point approaches, most probably, because it is much easier to handle the applications of classical fixed point theorems.

The paper is organized as follows: in Section 2 we recall some preliminary results that we need in the sequel and in Section 3 we prove our main results.

2. Preliminaries

Let (X,d) be a metric space. Recall that the Pompeiu-Hausdorff distance of the closed subsets $A, B \subset X$ is defined by

$$d_H(A,B) = \max\{d^*(A,B), d^*(B,A)\},\ d^*(A,B) = \sup\{d(a,B); a \in A\},$$

where $d(x,B) = \inf_{y\in B} d(x,y)$.

Let $I = [1,T]$, we denote by $C(I,\mathbf{R})$ the Banach space of all continuous functions from I to $\mathbf{R}$ with the norm $||x(.)||_C = \sup_{t\in I}|x(t)|$ and $L^1(I,\mathbf{R})$ is the Banach space of integrable functions $u(.) : I \to \mathbf{R}$ endowed with the norm $||u(.)||_1 = \int_0^T |u(t)|dt$.

Definition 2.1 (*cf.* [1]). *The Hadamard fractional integral of order $q > 0$ of a Lebesgue integrable function $f : [1,\infty) \to \mathbf{R}$ is defined by*

$$I_H^q f(t) = \frac{1}{\Gamma(q)} \int_1^t \left(\ln\frac{t}{s}\right)^{q-1} \frac{f(s)}{s} ds$$

provided the integral exists and $\Gamma(.)$ is the (Euler's) Gamma function defined by $\Gamma(q) = \int_0^\infty t^{q-1}e^{-t}dt$.

The Hadamard fractional derivative of order $q > 0$ of a function $f : [1,\infty) \to \mathbf{R}$ is defined by

$$D_H^q f(t) = t\frac{d^n}{dt^n}(I_H^{n-q}f)(t),$$

provided the integral exists and $n = [q]+1$, $[q]$ is the integer part of q.

The Hilfer-Hadamard fractional derivative of order $\alpha \in (n-1,n)$, $n \geq 2$ and type $\beta \in [0,1]$ of a function $f \in L^1(I,\mathbf{R})$ is defined by

$$D_{HH}^{\alpha,\beta} f(t) = (I_H^{\beta(n-\alpha)} D_H^\gamma f)(t), \quad \gamma = \alpha + n\beta - \alpha\beta.$$

When $\beta = 0$ the Hilfer-Hadamard fractional derivative gives Hadamard fractional derivative and when $\beta = 1$ the Hilfer-Hadamard fractional derivative gives Caputo-Hadamard fractional derivative.

Lemma 2.2 (*cf.* [1]). *Let $h(.) \in C(I,\mathbf{R})$ and*

$$\Lambda = (\ln T)^{\gamma-1} - \textstyle\sum_{j=1}^m \eta_j(\ln\xi_j)^{\gamma-1} - \sum_{i=1}^n \zeta_i \frac{\Gamma(\gamma)}{\Gamma(\gamma+\varphi_i)}(\ln\theta_i)^{\gamma+\varphi_i-1} - \\ \textstyle\sum_{k=1}^r \lambda_k \frac{\Gamma(\gamma)}{\Gamma(\gamma-\omega_k)}(\ln\mu_k)^{\gamma-\omega_k-1} \neq 0.$$

Then, the solution of problem $D_{HH}^{\alpha,\beta}x(t) = h(t)$ with boundary conditions (2) is given by

$$x(t) = I_H^\alpha h(t) + \tfrac{(\ln T)^{\gamma-1}}{\Lambda}[\textstyle\sum_{j=1}^m \eta_j I_H^\alpha h(\xi_j) + \sum_{i=1}^n \zeta_i I_H^{\alpha+\varphi_i} h(\theta_i) + \\ \textstyle\sum_{k=1}^r \lambda_k I_H^{\alpha-\omega_k} h(\mu_k) - I_H^\alpha h(T)], \quad t \in [1,T]. \tag{3}$$

Definition 2.3. *A function $x(.) \in C(I,\mathbf{R})$ is called a solution of a problem (1)-(2) if there exists $h(.) \in L^1(I,\mathbf{R})$ such that $h(t) \in F(t,x(t),V(x)(t))$ a.e. (I) and $x(.)$ is given by (3).*

Remark 1. *If we denote*

$$G(t,s) = \tfrac{1}{\Gamma(\alpha)}\left(\ln\tfrac{t}{s}\right)^{\alpha-1}\tfrac{\chi_{[1,t]}(s)}{s} + \textstyle\sum_{j=1}^m \tfrac{(\ln t)^{\gamma-1}}{\Lambda}\tfrac{1}{\Gamma(\alpha)}\left(\ln\tfrac{\xi_j}{s}\right)^{\alpha-1}\tfrac{\chi_{[1,\xi_j]}(s)}{s} \\ + \textstyle\sum_{i=1}^n \tfrac{(\ln t)^{\gamma-1}}{\Lambda}\tfrac{\zeta_i}{\Gamma(\alpha+\varphi_i)}\left(\ln\tfrac{\theta_i}{s}\right)^{\alpha+\varphi_i-1}\tfrac{\chi_{[1,\theta_i]}(s)}{s} + \sum_{k=1}^r \tfrac{(\ln t)^{\gamma-1}}{\Lambda}\tfrac{\lambda_k}{\Gamma(\alpha-\omega_k)}\cdot \\ \cdot\left(\ln\tfrac{\mu_k}{s}\right)^{\alpha-\omega_k-1}\tfrac{\chi_{[1,\mu_k]}(s)}{s} - \tfrac{(\ln t)^{\gamma-1}}{\Lambda}\tfrac{1}{\Gamma(\alpha)}\left(\ln\tfrac{T}{s}\right)^{\alpha-1}\tfrac{1}{s}$$

where $\chi_A(.)$ denotes the characteristic function of the set A, then the solution $x(.)$ in (3) may be put as $x(t) = \int_1^T G(t,s)h(s)ds$.

Moreover, if $0 < \omega_k \leq \alpha - 1$, $k = \overline{1,r}$, then for any $t, s \in I$

$$|G(t,s)| \leq \tfrac{(\ln T)^{\alpha-1}}{\Gamma(\alpha)} + \textstyle\sum_{j=1}^m \tfrac{(\ln T)^{\gamma-1}(\ln\xi_j)^{\alpha-1}}{|\Lambda|\Gamma(\alpha)} + \sum_{i=1}^n \tfrac{(\ln T)^{\gamma-1}|\zeta_i|(\ln\theta_i)^{\alpha+\varphi_i-1}}{|\Lambda|\Gamma(\alpha+\varphi_i)} \\ + \textstyle\sum_{k=1}^r \tfrac{(\ln T)^{\gamma-1}|\lambda_k|(\ln\mu_k)^{\alpha-\omega_k-1}}{|\Lambda|\Gamma(\alpha-\omega_k)} + \tfrac{(\ln T)^{\alpha+\gamma-2}}{|\Lambda|\Gamma(\alpha)} =: M$$

Also, we need a variant of Kuratowski and Ryll-Nardzewski selection theorem concerning measurable set-valued maps.

Lemma 2.4 (*cf.* [2]). *Consider X a separable Banach space, B is the closed unit ball in X, $H : I \to \mathcal{P}(X)$ is a set-valued map with nonempty closed values and $g : I \to X, L : I \to \mathbf{R}_+$ are measurable functions. If*

$$H(t) \cap (g(t) + L(t)B) \neq \emptyset \quad a.e.(I),$$

then the set-valued map $t \to H(t) \cap (g(t) + L(t)B)$ has a measurable selection.

3. Main results

In order to prove our results we need the following hypotheses.

Hypothesis H1. i) *$F(.,.) : I \times \mathbf{R} \times \mathbf{R} \to \mathcal{P}(\mathbf{R})$ has nonempty closed values and is $\mathcal{L}(I) \otimes \mathcal{B}(\mathbf{R} \times \mathbf{R})$ measurable.*

ii) *There exists $L(.) \in L^1(I, (0, \infty))$ such that, for almost all $t \in I, F(t,.,.)$ is $L(t)$-Lipschitz in the sense that*

$$d_H(F(t, x_1, y_1), F(t, x_2, y_2)) \leq L(t)(|x_1 - x_2| + |y_1 - y_2|) \quad \forall \ x_1, x_2, y_1, y_2 \in \mathbf{R}.$$

iii) *$k(.,.,.) : I \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ is a function such that $\forall x \in \mathbf{R}$, $(t, s) \to k(t, s, x)$ is measurable.*

iv) $|k(t, s, x) - k(t, s, y)| \leq L(t)|x - y|$ *a.e.* $(t, s) \in I \times I$, $\forall x, y \in \mathbf{R}$.

We use next the following notations

$$M(t) := L(t)(1 + \int_1^t L(u)du), \quad t \in I, \quad K_0 = \int_1^T M(t)dt.$$

Theorem 3.1. *Let $\alpha \in (1, 2]$, $\beta \in [0, 1]$ and $0 < \omega_k \leq \alpha - 1$, $k = \overline{1, r}$. Assume that Hypothesis H1 is satisfied and $MK_0 < 1$. Let $y(.) \in C(I, \mathbf{R})$ be such that $y(1) = 0$, $y(T) = \sum_{j=1}^m \eta_j y(\xi_j) + \sum_{i=1}^n \zeta_i I_H^{\varphi_i} y(\theta_i) + \sum_{k=1}^r \lambda_k D_H^{\omega_k} y(\mu_k)$ and there exists $p(.) \in L^1(I, \mathbf{R}_+)$ with $d(D_{HH}^{\alpha,\beta} y(t), F(t, y(t), V(y)(t))) \leq p(t)$ a.e. (I).*

Then there exists $x(.) : I \to \mathbf{R}$ a solution of problem (1)-(2) satisfying for all $t \in I$

$$|x(t) - y(t)| \leq \frac{M}{1 - MK_0} ||p(.)||_1.$$

Proof. The set-valued map $t \to F(t, y(t), V(y)(t))$ is measurable with closed values and

$$F(t, y(t), V(y)(t)) \cap \{D_{HH}^{\alpha,\beta} y(t) + p(t)[-1, 1]\} \neq \emptyset \quad a.e. \ (I).$$

It follows from Lemma 2.4 that there exists a measurable selection $h_1(t) \in F(t, y(t), V(y)(t))$ *a.e.* (I) such that

$$|h_1(t) - D_{HH}^{\alpha,\beta} y(t)| \leq p(t) \quad a.e. \ (I) \tag{4}$$

Define $x_1(t) = \int_1^T G(t, s)h_1(s)ds$ and one has

$$|x_1(t) - y(t)| \leq M \int_1^T p(t)dt.$$

We construct two sequences $x_n(.) \in C(I, \mathbf{R})$, $h_n(.) \in L^1(I, \mathbf{R})$, $n \geq 1$ with the following properties

$$x_n(t) = \int_1^T G(t, s)h_n(s)ds, \quad t \in I, \tag{5}$$

$$h_n(t) \in F(t, x_{n-1}(t), V(x_{n-1})(t)) \quad a.e. \ (I), \tag{6}$$

$$|h_{n+1}(t) - h_n(t)| \leq L(t)(|x_n(t) - x_{n-1}(t)| + \int_1^t L(s)|x_n(s) - x_{n-1}(s)|ds) \ a.e. \ (I) \tag{7}$$

If this is done, then from (4)-(7) we have for almost all $t \in I$

$$|x_{n+1}(t) - x_n(t)| \leq M(MK_0)^n \int_1^T p(t)dt \quad \forall n \in \mathbf{N}.$$

Indeed, assume that the last inequality is true for $n - 1$ and we prove it for n. One has

$$|x_{n+1}(t) - x_n(t)| \leq \int_1^T |G(t, t_1)|.|h_{n+1}(t_1) - h_n(t_1)|dt_1 \leq$$

$$M\int_1^T L(t_1)[|x_n(t_1)-x_{n-1}(t_1)|+\int_1^{t_1} L(s)|x_n(s)-x_{n-1}(s)|ds]dt_1 \le M$$
$$\int_1^T L(t_1)(1+\int_1^{t_1} L(s)ds)dt_1.M^nK_0^{n-1}\int_1^T p(t)dt = M(MK_0)^n\int_1^T p(t)dt.$$

Therefore $\{x_n(.)\}$ is a Cauchy sequence in the Banach space $C(I,\mathbf{R})$, hence converging uniformly to some $x(.) \in C(I,\mathbf{R})$. Hence, by (7), for almost all $t \in I$, the sequence $\{h_n(t)\}$ is Cauchy in $\mathbf{R}$. Let $h(.)$ be the pointwise limit of $h_n(.)$.

At the same time, one has

$$\begin{aligned}(8)\qquad & |x_n(t)-y(t)| \le |x_1(t)-y(t)| + \textstyle\sum_{i=1}^{n-1}|x_{i+1}(t)-x_i(t)| \le \\ & M\textstyle\int_1^T p(t)dt + \sum_{i=1}^{n-1}(M\int_1^T p(t)dt)(MK_0)^i \le \frac{M\int_1^T p(t)dt}{1-MK_0}.\end{aligned}$$

On the other hand, from (4), (7) and (8) we obtain for almost all $t \in I$

$$\begin{aligned}& |h_n(t)-D_{HH}^{\alpha,\beta}y(t)| \le \textstyle\sum_{i=1}^{n-1}|h_{i+1}(t)-h_i(t)| + |h_1(t)-D_{HH}^{\alpha,\beta}y(t)| \le \\ & L(t)\frac{M\int_1^T p(t)dt}{1-MK_0} + p(t).\end{aligned}$$

Hence the sequence $h_n(.)$ is integrably bounded and therefore $h(.) \in L^1(I,\mathbf{R})$.

Using Lebesgue's dominated convergence theorem and taking the limit in (5), (6) we deduce that $x(.)$ is a solution of (1)-(2) . Finally, passing to the limit in (8) we obtained the desired estimate on $x(.)$.

It remains to construct the sequences $x_n(.), h_n(.)$ with the properties in (5)-(7). The construction will be done by induction.

Since the first step is already realized, assume that for some $N \ge 1$ we already constructed $x_n(.) \in C(I,\mathbf{R})$ and $h_n(.) \in L^1(I,\mathbf{R})$, $n = 1,2,...N$ satisfying (5), (7) for $n = 1,2,...N$ and (6) for $n = 1,2,...N-1$. The set-valued map $t \to F(t, x_N(t), V(x_N)(t))$ is measurable. Moreover, the map $t \to L(t)(|x_N(t)-x_{N-1}(t)| + \int_1^t L(s)|x_N(s)-x_{N-1}(s)|ds)$ is measurable. By the lipschitzianity of $F(t,.)$ we have that for almost all $t \in I$

$$\begin{aligned}& F(t,x_N(t),V(x_N)(t)) \cap \{h_N(t) + L(t)(|x_N(t)-x_{N-1}(t)|+ \\ & \textstyle\int_1^t L(s)|x_N(s)-x_{N-1}(s)|ds)[-1,1]\} \neq \emptyset.\end{aligned}$$

Lemma 2.4 yields that there exist a measurable selection $h_{N+1}(.)$ of $F(.,x_N(.),V(x_N)(.))$ such that for almost all $t \in I$

$$|h_{N+1}(t)-h_N(t)| \le L(t)(|x_N(t)-x_{N-1}(t)| + \int_1^t L(s)|x_N(s)-x_{N-1}(s)|ds).$$

We define $x_{N+1}(.)$ as in (5) with $n = N+1$. Thus $f_{N+1}(.)$ satisfies (6) and (7) and the proof is complete. □

Corollary 3.2. *Let $\alpha \in (1,2]$, $\beta \in [0,1]$ and $0 < \omega_k \le \alpha - 1$, $k = \overline{1,r}$. Assume that Hypothesis H1 is satisfied, $d(0, F(t,0,0) \le L(t)$ a.e. (I) and $MK_0 < 1$. Then there exists $x(.)$ a solution of problem (1)-(2) satisfying for all $t \in I$ $|x(t)| \le \frac{M}{1-MK_0}||L(.)||_1$.*

Proof. It is enough to take $y(.) = 0$ and $p(.) = L(.)$ in Theorem 3.1. □

If F does not depend on the last variable, Hypothesis H1 becames

Hypothesis H2. i) *$F(.,.) : I \times \mathbf{R} \to \mathcal{P}(\mathbf{R})$ has nonempty closed values and is $\mathcal{L}(I) \otimes \mathcal{B}(\mathbf{R})$ measurable.*

ii) *There exists $L(.) \in L^1(I,(0,\infty))$ such that, for almost all $t \in I$, $F(t,.)$ is $L(t)$-Lipschitz in the sense that*

$$d_H(F(t,x_1), F(t,x_2)) \le L(t)|x_1 - x_2| \quad \forall\, x_1, x_2 \in \mathbf{R}.$$

Denote $L_0 = \int_1^T L(t)dt$.

Corollary 3.3. *Let $\alpha \in (1,2]$, $\beta \in [0,1]$ and $0 < \omega_k \le \alpha - 1$, $k = \overline{1,r}$. Assume that Hypothesis H2 is satisfied, $d(0, F(t,0) \le L(t)$ a.e. (I) and $ML_0 < 1$. Then there exists $x(.)$ a solution of the fractional differential inclusion*

$$(9)\qquad D_{HH}^{\alpha,\beta}x(t) \in F(t,x(t)) \quad a.e.\ (I),$$

with boundary conditions (2) satisfying for all $t \in I$

$$|x(t)| \le \frac{ML_0}{1 - ML_0}. \tag{10}$$

Remark 2. *A similar result to the one in Corollary 3.3 may be found in* [1]*; namely, Theorem 7. The proof of Theorem 7 in* [1] *is done by using the set-valued contraction principle. Our approach improves the hypothesis concerning the set-valued map in* [1]*. More exactly, we do not require for the values of F to be compact as in* [1] *and we do not require that the Lipschitz constant of F to be a mapping from $C(I, \mathbf{R})$ as in* [1]*. Moreover, Theorem 7 in* [1] *does not contains a priori bounds for solutions as in (10).*

Example 3.4. *Let us consider the problem*

$$\begin{aligned} &D_{HH}^{\frac{5}{3},\frac{3}{4}}x(t) \in [-(\tfrac{1}{4\sqrt{15}} - \tfrac{1}{16})\tfrac{|x(t)|}{1+|x(t)|}, 0] \cup [0, \\ &(\tfrac{1}{4\sqrt{15}} - \tfrac{1}{16})\tfrac{|\int_1^t x(s)ds|}{1+(\frac{1}{4\sqrt{15}}-\frac{1}{16})|\int_1^t x(s)ds|}], \quad a.e. \ ([1,5]) \end{aligned} \tag{11}$$

with nonlocal integral boundary conditions as in [1]*; namely,*

$$\begin{aligned} &x(1) = 0, \quad x(5) = \tfrac{1}{15}x(\tfrac{5}{4}) + \tfrac{1}{10}x(\tfrac{3}{2}) + \tfrac{2}{15}x(\tfrac{7}{4}) + \tfrac{1}{6}x(2) + \tfrac{1}{18}I_H^{\frac{1}{2}}x(\tfrac{5}{2}) + \\ &\tfrac{1}{9}I_H^1 x(3) + \tfrac{1}{6}I_H^{\frac{3}{2}}x(\tfrac{7}{2}) + \tfrac{1}{28}D_H^{\frac{1}{4}}x(4) + \tfrac{1}{14}D_H^{\frac{2}{3}}x(\tfrac{9}{2}). \end{aligned} \tag{12}$$

In this case, $\alpha = \frac{5}{3}$, $\beta = \frac{3}{4}$, $T = 5$, $m = 4$, $n = 3$, $r = 2$, $\eta_1 = \frac{1}{15}$, $\eta_2 = \frac{1}{10}$, $\eta_3 = \frac{2}{15}$, $\eta_4 = \frac{1}{6}$, $\xi_1 = \frac{5}{4}$, $\xi_2 = \frac{3}{2}$, $\xi_3 = \frac{7}{4}$, $\xi_4 = 2$, $\zeta_1 = \frac{1}{18}$, $\zeta_2 = \frac{1}{9}$, $\zeta_3 = \frac{1}{6}$, $\varphi_1 = \frac{1}{2}$, $\varphi_2 = 1$, $\varphi_3 = \frac{3}{2}$, $\theta_1 = \frac{5}{2}$, $\theta_2 = 3$, $\theta_3 = \frac{7}{2}$, $\lambda_1 = \frac{1}{28}$, $\lambda_2 = \frac{1}{14}$, $\omega_1 = \frac{1}{4}$, $\omega_2 = \frac{2}{3}$, $\mu_1 = 4$, and $\mu_2 = \frac{9}{2}$.

Define $F(.,.) : I \times \mathbf{R} \times \mathbf{R} \to \mathcal{P}(\mathbf{R})$ by

$$F(t,x,y) = [-(\frac{1}{4\sqrt{15}} - \frac{1}{16})\frac{|x|}{1+|x|}, 0] \cup [0, (\frac{1}{4\sqrt{15}} - \frac{1}{16})\frac{|y|}{1+|y|}]$$

and $k(.,.,.) : I \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ by $k(t,s,x) = (\frac{1}{4\sqrt{15}} - \frac{1}{16})x$.

Since

$$\sup\{|u|; \quad u \in F(t,x,y)\} \le (\frac{1}{4\sqrt{15}} - \frac{1}{16}) \quad \forall\, t \in [0,1], \ x, y \in \mathbf{R},$$

$$d_H(F(t,x_1,y_1), F(t,x_2,y_2)) \le (\frac{1}{4\sqrt{15}} - \frac{1}{16})|x_1 - x_2| + (\frac{1}{4\sqrt{15}} - \frac{1}{16})|y_1 - y_2|$$

$\forall\, x_1, x_2, y_1, y_2 \in \mathbf{R}$, in this situation $L(t) \equiv (\frac{1}{4\sqrt{15}} - \frac{1}{16})$ and $K_0 = (\frac{1}{\sqrt{15}} - \frac{1}{4})(\frac{1}{\sqrt{15}} + \frac{3}{4})) < \frac{1}{60}$.

By standard computations (e.g., [1]*) $M \approx 39,4$; therefore, $MK_0 < 1$. So, we may apply Corollary 3.3 in order to obtain the existence of a solution for problem (11)-(12).*

4. Conclusion

In the present paper, we studied a boundary value problem associated to a Hilfer-Hadamard fractional integro-differential inclusion and we established an existence result for this problem when the set-valued map is Lipschitz in the state variables without any assumptions concerning the convexity of the values of the set-valued map. Our approach uses a technique due to Filippov ([9]) instead of an usual application of set-valued fixed point theorems. An illustration of our result is provided by a numerical example. In this way we extended and improved the research in [1]. We underline that this type of result is a basic tool in the study of optimal control problems defined by such kinds of fractional differential inclusions.

References

[1] Ahmad B. and Ntouyas S.K., (2021), Hilfer-Hadamard fractional boundary value problems with nonlocal mixed boundary conditions, Fractal and Fractional, 5 (195), 1-25.

[2] Aubin J.P. and Frankowska H., (1990), *Set-valued Analysis*, Birkhäuser, Basel.

[3] Băleanu D., Diethelm K., Scalas E. and Trujillo J.J., (2012), *Fractional Calculus Models and Numerical Methods*, World Scientific, Singapore.

[4] Cernea A., (2015), Filippov lemma for a class of Hadamard-type fractional differential inclusions, Fract. Calc. Appl. Anal., 18, 163-171.

[5] Cernea A., (2019), On a fractional integro-differential inclusion of Caputo-Katugampola type, Bull. Math. Anal. Appl., 11, 22-27.

[6] Cernea A., (2020) On some fractional integro-differential inclusions with Erdélyi-Kober fractional integral boundary conditions, Mem. Differ. Equa. Math. Phy., 79, 15-26.

[7] Cernea A., (2021), On the solutions of a certain boundary value problem associated to a fractional differential inclusion, Discuss. Math., Differ. Incl., Control Optim., 41, 5-17.

[8] Diethelm K., (2016) *The Analysis of Fractional Differential Equations*, Springer, Berlin.

[9] Filippov, A.F., (1967) Classical solutions of differential equations with multivalued right hand side, SIAM J. Control, 5, 609-621.

[10] Hilfer R., (2000), *Applications of Fractional Calculus in Physics*, World Scientific, Singapore.

[11] Hilfer R., Luchko Y. and Tomovski Z., (2009), Operational method for the solution of fractional differential equations with generalized Riemann-Liouville fractional derivative, Fract. Calc. Appl. Anal., 12, 299-318.

[12] Kilbas A., Srivastava H.M. and Trujillo J.J., (2006) *Theory and Applications of Fractional Differential Equations*, Elsevier, Amsterdam.

[13] Miller K. and Ross B., (1993) *An Introduction to the Fractional Calculus and Differential Equations*, John Wiley, New York.

[14] Podlubny I., (1999) *Fractional Differential Equations*, Academic Press, San Diego.

Faculty of Mathematics and Computer Science, University of Bucharest, Academiei 14, 010014 Bucharest, Romania

Email(s): acernea@fmi.unibuc.ro

$\Psi-$Ideal Convergence of Double Positive Linear Operators of Functions of Two Variables for Analytic P-ideals

Kamil Demirci [1], Fadime Dirik *[2] and Sevda Yıldız [3]

ABSTRACT. In this work, we introduce an extension of Ψ-statistical convergence to the class of all analytic P-ideals for sequences called Ψ-ideal convergence. Also using this convergence, we prove a Korovkin type approximation theorem for double sequences. We compute the rates of Ψ-ideal convergence of positive linear operators.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 40G15, 41A36

KEYWORDS: $\Psi-$statistical convergence, ideal convergence, analytic P-ideals, Korovkin theorem, Berstein-Stancu operator

1. INTRODUCTION AND PRELIMINARIES

The subject of Korovkin-type theory was initiated by Korovkin in 1960 in his pioneering paper [15] then, this theory has been widely studied and it is worthwhile to point out that, it is about approximation to continuous functions by means of positive linear operators (see [1, 15]). Many researchers were considered several extensions in which a more general notion of convergence is used. In particular, the use of statistical convergence has increased more in recent years (also see, [2, 7, 11, 14, 18, 21]). Also, Dirik and Demirci ([9]) have studied the Korovkin-type approximation thorem using the notion of equi-ideal convergence for analytic P-ideals. Then, Bardaro et al. ([3, 4]) have introduced a more general notion of statistical convergence for double sequences of positive linear operators named "triangular $A-$statistical convergence" and they have obtained a Korovkin type approximation theorem in $C(U)$, which is the space of all continuous real functions defined in a compact subset $U \subset \mathbb{R}^2$ and they have introduced a general definition of triangular $A-$statistical convergence using a suitable function $\Psi : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$. After these remarkable works, authors have studied some generalizations of them ([5, 6, 10]).

Our primary interest is to prove Korovkin type approximation theorem by using an extension of $\Psi-$statistical convergence to the class of all analytic P-ideals for sequences called $\Psi-$ideal convergence. We also study the rates of $\Psi-$ideal convergence of positive linear operators.

We begin with some definitions and notations which we will use in the sequel.

Let $\Psi : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ be a fixed function, $K \subset \mathbb{N}^2$ be a nonempty set, and for every $i \in \mathbb{N}$, let $K_i = \{j \in \mathbb{N} : (i,j) \in K, \Psi(i,j) \geq 0\}$. Let $|K_i|$ be the cardinality of K_i then we define $\Psi-$*density* of K by

$$\delta^{\Psi}(K) := \lim_i \frac{1}{i} |K_i|,$$

provided that the limit on the right-hand side exists (see for details, [4]).

Definition 1.1. [4] *The number sequence $x = (x_{i,j})$ is $\Psi-$statistically convergent to L provided that for every $\varepsilon > 0$*

$$\lim_i \frac{1}{i} |K_i(\varepsilon)| = 0,$$

where $K_i(\varepsilon) = \{ j \in \mathbb{N} : \ \Psi(i,j) \geq 0, \ \ |x_{i,j} - L| \geq \varepsilon\}$ and this is denoted by $st^{\Psi}-\lim_i x_{i,j} = L$.

If we take $\Psi(i,j) = i - j \geq 0$, we obtain the notion of triangular statistical convergence. As indicated earlier, triangular statistical convergence and statistical convergence overlap, neither contains the other. Also, if a double sequence, Pringsheim convergent ([19]), then it is $\Psi - A-$statistical convergent. Note that the converse, in general, is not true (see also [3, 4]).

We denote the set of all $\Psi-$statistically convergent sequences by st^{Ψ}.

The concept of convergence of a sequence of real numbers has been extended to statistical convergence by Fast [13]. Then, various kinds of ideal convergence which are extensions of statistical convergence to the class of all analytic P-ideals for sequences of functions have been introduced by Mrożek [17]. We first recall these convergence methods.

By an ideal on $\mathbb{N}$, the set of natural numbers, we mean a family of subsets of $\mathbb{N}$ closed under taking finite unions and subsets of it. If not explicitly said we assume that the ideals are proper ($\neq \mathcal{P}(\mathbb{N})$) and contain all finite sets. By Fin we denote the ideals of all finite sets of $\mathbb{N}$. An ideal $\mathcal{I}$ is a P-ideal if for every sequence $(A_i)_{i\in\mathbb{N}}$ of sets from $\mathcal{I}$ there is an $A \in \mathcal{I}$ such that $A_i \setminus A$ is finite for all i. By identifying subsets of naturals with their characteristic functions, we equip $\mathcal{P}(\mathbb{N})$ with the Contor-space topology and therefore we can assign the topological complexity to the ideals of sets naturals. In particular, an ideal $\mathcal{I}$ is analytic if it is a continuous image of G_δ subsets of the Cantor-space.

A map $\phi : \mathcal{P}(\mathbb{N}) \to [0, \infty]$ is a submeasure on $\mathbb{N}$ if for all $A, B \subset \mathbb{N}$,

$$\begin{aligned}\phi(\emptyset) &= 0,\\ \phi(A) &\leq \phi(A \cup B) \leq \phi(A) + \phi(B).\end{aligned}$$

It is lower semicontinuous if for all $A \subset \mathbb{N}$, we have

$$\phi(A) = \lim_{k\to\infty} \phi(A \cap \{0, 1, 2, ..., k-1\}).$$

For any lower semicontinuous submeasure on $\mathbb{N}$, let $\|A\|_\phi : \mathcal{P}(\mathbb{N}) \to [0, \infty]$ be the submeasure defined by

$$\begin{aligned}\|A\|_\phi &= \limsup_{k\to\infty} \phi(A \setminus \{0, 1, 2, ..., k-1\})\\ &= \lim_{k\to\infty} \phi(A \setminus \{0, 1, 2, ..., k-1\})\end{aligned}$$

where the second equality follows by the monotonicity of ϕ. Let

$$Exh(\phi) = \left\{A \subseteq \mathbb{N} : \|A\|_\phi = 0\right\}.$$

It is clear that $Exh(\phi)$ is an ideal (not necessarily proper) for an arbitrary submeasure ϕ.

Let us introduce the following examples of analytic P-ideals [20], (see [12] for more examples).

- A nontrivial analytic P-ideal is the ideal of sets of statistical density zero, i.e.,

  $$\mathcal{I}_d = \left\{A \subseteq \mathbb{N} : \limsup_{k\to\infty} d_k(A) = 0\right\},$$

  where $d_k(A) = \frac{|A\cap\{0,1,2,...,k-1\}|}{k}$ is the kth partial density of A, where the symbol $|B|$ denotes the cardinality of the set B. If we denote $\phi_d(A) = \sup\limits_k \frac{|A\cap\{0,1,2,...,k-1\}|}{k}$, then $\mathcal{I}_d = Exh(\phi_d)$.
- Let $\mathcal{I}_{\frac{1}{k}} = \left\{A \subseteq \mathbb{N} : \sum\limits_{k\in A} \frac{1}{k+1} < \infty\right\}$. If ϕ is a submeasure defined by $\phi(A) = \sum\limits_{k\in A} \frac{1}{k+1}$, then $\mathcal{I}_{\frac{1}{k}} = Fin\ (\phi)$ where is the ideal which consists of all finite sets of $\mathbb{N}$.

Throughout this paper, let $\mathcal{I}$ be an ideal on $\mathbb{N}$. $\mathcal{I}$ is an analytic P-ideal iff $\mathcal{I} = Exh(\phi)$ for some lower semicontinouos submeasure ϕ on $\mathbb{N}$ and we use the following notations.

Let $C(X)$ be the space of all continuous real valued functions on X, a compact subset of $\mathbb{R}^2$ and let f and $f_{i,j}$ belong to $C(X)$. Let $\Psi : \mathbb{N}\times\mathbb{N} \to \mathbb{R}$ be a fixed function,

$$\begin{aligned}\Gamma_i((x,y),\varepsilon) &:= \{j \in \mathbb{N} : \Psi(i,j) \geq 0,\ |f_{i,j}(x) - f(x)| \geq \varepsilon\},\ (x,y) \in X,\\ \Phi_i(\varepsilon) &:= \left\{j \in \mathbb{N} : \Psi(i,j) \geq 0,\ ||f_{i,j} - f||_{C(X)} \geq \varepsilon\right\},\end{aligned} \tag{1}$$

where $\varepsilon > 0$ and $||f||_{C(X)}$ denotes the usual supremum norm of f in $C(X)$.

Definition 1.2. *$(f_{i,j})$ is said to be pointwise $\Psi-$ ideal convergent to f on X if for every $\varepsilon > 0$ and for each $(x,y) \in X$, $\lim\limits_{k\to\infty} \phi(\Gamma_i((x,y),\varepsilon) \setminus \{0,1,2,...,k-1\}) = 0$ and write $f_{i,j} \to f(\Psi - ideal)$ on X.*

Definition 1.3. *$(f_{i,j})$ is said to be uniform $\Psi-$ideal convergent to f on X if, $\lim\limits_{k\to\infty} \phi(\Phi_i(\varepsilon) \setminus \{0,1,2,...,k-1\}) = 0$, for every $\varepsilon > 0$ and write $f_{i,j} \rightrightarrows f\ (\Psi - ideal)$ on X.*

Remark 1. *If we take $\Psi(i,j) = i - j$ and $I_d = Exh(\phi)$ where*

$$\phi(A) = \sup_{n\in\mathbb{N}} d_n(A),$$

then $\Psi-$ideal convergence is equal to triangular statistical uniform convergence which is given by Çınar in [5].

Using these definitions, the next result is immediately obtained.

Lemma 1.4. *$f_{i,j} \rightrightarrows f$ on X implies $f_{i,j} \rightrightarrows f$ $(\Psi - ideal)$ on X. Furthermore, $f_{i,j} \to f$ on X implies $f_{i,j} \to f$ $(\Psi - ideal)$ on X.*

Example 1.5. *Let $\Psi(i,j) = i - j$ and $\mathcal{I}_d = Exh(\phi)$ where $\phi(A) = \sup_{n\in\mathbb{N}} d_n(A)$, $X = [-1,1] \times [-1,1]$ and h is a function by $h(x,y) = x^2y^2$ for $x, y \in [-1,1]$. For each $i, j \in \mathbb{N}$, define $h_{i,j} \in C(X)$ by*

$$h_{i,j}(x,y) = \begin{cases} 0, & i = j = k^2, \\ \left(xy - \frac{1}{ij}\right)^2, & i = 2k+1,\ j = 2k-1, \\ \frac{x^2y^2}{2}, & i = 2k,\ j = 2(k+1), \\ x^2y^2, & otherwise. \end{cases}$$

Then it is easy show that $(h_{i,j})$ is triangular statistical uniform convergent to h on X with respect to the ideal $\mathcal{I}_d$. But $(h_{i,j})$ is not uniform convergent and statistical uniform convergent to the function h on X.

2. Korovkin-type approximation theorem

In this section, we give a Korovkin-type approximation theorem using the concept of $\Psi-$ideal convergence.

Now we have the following main result.

Theorem 2.1. *Let $(L_{i,j})$ be a double sequence of positive linear operators moving from $C(X)$ into $C(X)$. Then for all $g \in C(X)$,*

$$L_{i,j}(g) \rightrightarrows g \ (\Psi - ideal) \text{ on } X, \tag{2}$$

if and only if

$$L_{i,j}(h_r) \rightrightarrows h_r(\Psi - ideal) \text{ on } X,\ (r = 0,1,2,3), \tag{3}$$

where $h_0(x,y) = 1$, $h_1(x,y) = x$, $h_2(x,y) = y$ and $h_3(x,y) = x^2 + y^2$.

Proof. Since each $h_r \in C(X)$, $(r = 0,1,2,3)$, the implication (2)$\Rightarrow$(3) is obvious. Suppose now that (3) hold. By the continuity of f on X and using the monotonicity and the linearity of the operators $L_{i,j}$, for every $\varepsilon > 0$, we get

$$\begin{aligned} |L_{i,j}(g;x,y) - g(x,y)| &\leq L_{i,j}(|g(u,v) - g(x,y)|;x,y) + M|L_{i,j}(h_0;x,y) - h_0(x,y)| \\ &\leq \left|L_{i,j}\left(\varepsilon + \frac{2M}{\delta^2}\left\{(u-x)^2 + (v-y)^2\right\};x,y\right)\right| \\ &\quad + M|L_{i,j}(h_0;x,y) - h_0(x,y)| \\ &\leq \left(\varepsilon + M + \frac{2M}{\delta^2}\left(K_1^2 + K_2^2\right)\right)|L_{i,j}(h_0;x,y) - h_0(x,y)| \\ &\quad + \frac{4M}{\delta^2}K_1|L_{i,j}(h_1;x,y) - h_1(x,y)| + \frac{4M}{\delta^2}K_2|L_{i,j}(h_2;x,y) - h_2(x,y)| \\ &\quad + \frac{2M}{\delta^2}|L_{i,j}(h_3;x,y) - h_3(x,y)| + \varepsilon \end{aligned}$$

where $M := \|g\|_{C(X)}$, $K_1 := \max|x|$, $K_2 := \max|y|$. Taking supremum over $(x,y) \in X$, we get

$$\|L_{i,j}(g) - g\|_{C(X)} \leq K\left\{\sum_{r=0}^{3}\|L_{i,j}(h_r) - h_r\|_{C(X)}\right\} + \varepsilon. \tag{4}$$

where $K := \max\left\{\varepsilon + M + \frac{2M}{\delta^2}\left(K_1^2 + K_2^2\right), \frac{4M}{\delta^2}K_1, \frac{4M}{\delta^2}K_2, \frac{2M}{\delta^2}\right\}$. Now, for a given $\varepsilon' > 0$, choose $\varepsilon > 0$ such that $\varepsilon < \varepsilon'$. Then, define

$$\begin{aligned}\Phi_i(\varepsilon') \ &: = \left\{j \in \mathbb{N} : \Psi(i,j) \geq 0,\ \left\|L_{i,j}(g) - g\right\|_{C(X)} \geq \varepsilon'\right\},\\ \Phi_i^r(\frac{\varepsilon' - \varepsilon}{4K}) \ &: = \left\{j \in \mathbb{N} : \Psi(i,j) \geq 0,\ \left\|L_{i,j}(h_r) - h_r\right\|_{C(X)} \geq \frac{\varepsilon' - \varepsilon}{4K}\right\},\ r = 0,1,2,3.\end{aligned}$$

Then it is easy to see that $\Phi_i(\varepsilon') \subseteq \bigcup_{r=0}^{3} \Phi_i^r(\frac{\varepsilon'-\varepsilon}{4K})$. Thus, from the monotonicity of ϕ, it follows from (4) that

$$\begin{aligned}\phi\left(\Phi_i(\varepsilon') \setminus k\right) \ &\leq \ \phi\left(\left[\bigcup_{r=0}^{3} \Phi_i^r(\frac{\varepsilon'-\varepsilon}{4K})\right] \setminus \{0,1,2,...,k-1\}\right)\\ &\leq \ \sum_{i=0}^{3} \phi\left(\Phi_i^r(\frac{\varepsilon'-\varepsilon}{4K}) \setminus \{0,1,2,...,k-1\}\right). \qquad (5)\end{aligned}$$

Letting $k \to \infty$, using (3) and considering Definition 1.3, we obtain (2). The proof is complete. □

If we take $\Psi(i,j) = i - j$ and $\mathcal{I}_d = Exh(\phi)$ where $\phi(A) = \sup_{k\in\mathbb{N}} d_k(A)$, then we get the following triangular statistical Korovkin-type approximation theorem which was introduced by Bardaro et al. [4].

Corollary 2.2. [4] *Let $(L_{i,j})$ be a double sequence of positive linear operators moving from $C(X)$ into itself. Then, for all $g \in C(X)$,*

$$st^{\Psi} - \lim_i \left\|L_{i,j}(g) - g\right\|_{C(X)} = 0 \ on\ X$$

if and only if

$$st^{\Psi} - \lim_i \left\|L_{i,j}(h_r) - h_r\right\|_{C(X)} = 0 \ on\ X,$$

where $h_0(x,y) = 1$, $h_1(x,y) = x$, $h_2(x,y) = y$ and $h_3(x,y) = x^2 + y^2$.

3. An Application

We now present an example that works in our new approximation result, but in which the classical case ([22]) and statistical case ([8]) do not. Let $\Psi(i,j) = i - j$ and $\mathcal{I}_d = Exh(\phi)$ where $\phi(A) = \sup_{k\in\mathbb{N}} d_k(A)$ and $X = [0,1]\times[0,1]$. To see this first consider the Bernstein-Stancu operators (see [1])

$$K_{i,j,\alpha,\beta,\gamma,\delta}(g;x,y) = \sum_{s=0}^{i}\sum_{t=0}^{j} g\left(\frac{\alpha+s}{\beta+i}, \frac{\gamma+t}{\delta+j}\right)\binom{i}{s}\binom{j}{t} x^s y^t (1-x)^{i-s}(1-y)^{j-t},$$

where $X = [0,1]\times[0,1]$, $(x,y) \in X$, $f \in C(X)$, $\alpha, \beta, \gamma, \delta$ are fixed real numbers. Also, observe that

$$\begin{aligned}K_{i,j,\alpha,\beta,\gamma,\delta}(1;x,y) &= 1,\\ K_{i,j,\alpha,\beta,\gamma,\delta}(u;x,y) &= \frac{i}{i+\beta}x + \frac{\alpha}{i+\beta},\\ K_{i,j,\alpha,\beta,\gamma,\delta}(v;x,y) &= \frac{j}{j+\delta}y + \frac{\alpha}{j+\delta},\\ K_{i,j,\alpha,\beta,\gamma,\delta}\left(u^2+v^2;x,y\right) &= \frac{i(i-1)}{(i+\beta)^2}x^2 + \frac{(2\alpha+1)i}{(i+\beta)^2}x + \frac{\alpha^2}{(i+\beta)^2}\\ &\quad + \frac{j(j-1)}{(j+\delta)^2}y^2 + \frac{(2\gamma+1)j}{(j+\delta)^2}y + \frac{\gamma^2}{(j+\delta)^2}.\end{aligned}$$

Using this polynomial, we introduce the following positive linear operators on $C(X)$:

$$L_{i,j}(g;x,y) = (1 + x_{i,j})K_{i,j,\alpha,\beta,\gamma,\delta}(g;x,y) \qquad (6)$$

where $x=(x_{i,j})$ be defined

$$x_{i,j}:=\begin{cases} 1, & i=j=k^2, \\ \frac{3k}{7(k+1)}, & i=2k+1,\ j=2k-1, \\ \frac{2k}{5(k+1)}, & i=2k,\ j=2(k+1), \\ 0, & \text{otherwise.} \end{cases}$$

Since $x_{i,j}\to 0$ $(\Psi-ideal)$, we conclude that

$$L_{i,j}(h_r)\rightrightarrows h_r(\Psi-ideal) \text{ on } X,\ (r=0,1,2,3).$$

From Theorem 2.1, we see that

$$L_{i,j}(f)\rightrightarrows f(\Psi-ideal) \text{ on } X.$$

Since x is not ordinary and statistically convergent to zero, the sequence $(L_{i,j})$ given by (6) does not satisfy the conditions of classical ([22]) and statistical ([8]) cases, respectively.

4. Rate of Convergence

In this section, we compute the rates of $\Psi-$ideal convergence of a double sequence of positive linear operators defined on $C(X)$ by means of the modulus of continuity.

Now we give the following definition.

Definition 4.1. *Let (α_k) be a positive non-increasing sequence. The sequence $(f_{i,j})$ is uniform $\Psi-$ideal convergent to f with the rate of $o(\alpha_k)$ if, for each $\varepsilon>0$,*

$$\lim_{k\to\infty}\frac{1}{\alpha_k}\phi\left(\Phi_i(\varepsilon)\setminus\{0,1,2,...,k-1\}\right)=0,$$

where $\Phi_i(\varepsilon)$ is defined by (1). In this case we write $f_{i,j}-f=o(\alpha_k)\,(\Psi-ideal)$ on X.

One of the important element in approximation theory is modulus of continuity $\omega(g;\gamma)$ defined as follows:

$$\omega(g;\gamma):=\sup\left\{|g(u,v)-g(x,y)|:(u,v),(x,y)\in X,\sqrt{(u-x)^2+(v-y)^2}\le\gamma\right\}.$$

where $g\in C(X)$ and $\gamma>0$. Show that $\omega(g;\gamma)$ is an increasing function of γ,

$$\omega(g;\alpha\gamma)\le(1+[\alpha])\,\omega(g;\gamma) \tag{7}$$

for every $\gamma,\delta>0$ (see also [1]).

Using these definitions we obtain the following lemma.

Lemma 4.2. *Let (α_k) and (β_k) be positive non-increasing sequences and $(f_{i,j})$ and $(g_{i,j})$ be sequences of function belonging to $C(X)$. Suppose that $f_{i,j}-f=o(\alpha_k)$ $(\Psi-ideal)$ on X and $g_{i,j}-g=o(\beta_k)$ $(\Psi-ideal)$ on X. Let $c_k=\max\{\alpha_k,\ \beta_k\}$. Then, the following expressions are provided:*

(i) $(f_{i,j}-f)\mp(g_{i,j}-g)=o(c_k)$ $(\Psi-ideal)$ on X,

(ii) $\lambda(f_{i,j}-f)=o(\alpha_k)$ $(\Psi-ideal)$ on X, for any real number λ,

(iii) $\sqrt{|f_{i,j}-f|}=o(\alpha_k)$ $(\Psi-ideal)$ on X.

Proof. (i) Suppose that $f_{i,j}-f=o(\alpha_k)$ $(\Psi-ideal)$ on X and that $g_{i,j}-g=o(\beta_k)$ $(\Psi-ideal)$ on X. Also, for $\varepsilon>0$ and $(x,y)\in X$ define

$$\begin{aligned}\Phi_i(\varepsilon) &:= \left\{j\in\mathbb{N}:\Psi(i,j)\ge 0,\ ||(f_{i,j}-f)\mp(g_{i,j}-g)||_{C(X)}\ge\varepsilon\right\},\\ \Phi_i^1(\frac{\varepsilon}{2}) &:= \left\{j\in\mathbb{N}:\Psi(i,j)\ge 0,\ ||f_{i,j}-f||_{C(X)}\ge\frac{\varepsilon}{2}\right\},\\ \Phi_i^2(\frac{\varepsilon}{2}) &:= \left\{j\in\mathbb{N}:\Psi(i,j)\ge 0,\ ||g_{i,j}-g||_{C(X)}\ge\frac{\varepsilon}{2}\right\}.\end{aligned}$$

Then, observe that

$$\Phi_i(\varepsilon)\subset\Phi_i^1(\frac{\varepsilon}{2})\cup\Phi_i^2(\frac{\varepsilon}{2})$$

which gives

$$\begin{aligned}&\frac{1}{c_k}\phi\left(\Phi_i(\varepsilon)\setminus\{0,1,2,...,k-1\}\right)\\ \le\ &\frac{1}{c_k}\left(\phi\left(\Phi_i^1(\frac{\varepsilon}{2})\setminus\{0,1,2,...,k-1\}\right)+\phi\left(\Phi_i^2(\frac{\varepsilon}{2})\setminus\{0,1,2,...,k-1\}\right)\right).\end{aligned}$$

Since $\gamma_k = \max\{\alpha_k,\ \beta_k\}$, we get

$$\begin{aligned}(8)\qquad & \frac{\phi\left(\Phi_i(\varepsilon)\setminus\{0,1,2,...,k-1\}\right)}{c_k} \\ \leq\ & \frac{\phi\left(\Phi_i^1(\frac{\varepsilon}{2})\setminus\{0,1,2,...,k-1\}\right)}{\alpha_k}+\frac{\phi\left(\Phi_i^2(\frac{\varepsilon}{2})\setminus\{0,1,2,...,k-1\}\right)}{\beta_k}.\end{aligned}$$

Presently, by taking limit as $k\to\infty$ in (8) and by means of the hypothesis, we obtain that

$$\lim_{k\to\infty}\frac{\phi\left(\Phi_i(\varepsilon)\setminus\{0,1,2,...,k-1\}\right)}{c_k}=0.$$

This completes the proof of (i). Since the proofs of (ii) and (iii) can be made similarly to the proof of (i), we omit them. $\square$

Then we have the following result.

Theorem 4.3. *Let (α_k) and (β_k) be positive non-increasing sequences and $(L_{i,j})$ be a sequence of positive linear operators moving from $C(X)$ into $C(X)$. Suppose that the following conditions hold:*

(i) $L_{i,j}(h_0)-h_0=o(\alpha_k)$ $(\Psi-ideal)$ on X,

(ii) $w(g,\alpha_{i,j})=o(\beta_k)$ $(\Psi-ideal)$ on X, where $\alpha_{i,j}=\sqrt{\|L_{i,j}(\varphi)\|_{C(X)}}$ with $\varphi(u,v)=(u-x)^2+(v-y)^2$.

Then we have, for all $g\in C(X)$,

$$L_{i,j}(g)-g=o(\delta_k)(\Psi-ideal)\ on\ X,$$

where $\delta_k=\max\{\alpha_k,\ \beta_k\}$.

Proof. Let $g\in C(X)$ and $(x,y)\in X$ be fixed. As in the proof of the Theorem 4 of [3], it follows from the monotonicity and the linearity of the operators $L_{i,j}$, for any $(i,j)\in\mathbb{N}^2$ and $\delta>0$,

$$\begin{aligned} & \left|L_{i,j}(g;x,y)-g(x,y)\right| \\ \leq\ & \omega(g;\delta)\left|L_{i,j}(h_0;x,y)-h_0(x,y)\right|+\frac{\omega(g;\delta)}{\delta^2}L_{i,j}(\varphi;x,y)+\omega(g;\delta) \\ (9)\qquad & +M\left|L_{i,j}(h_0;x,y)-h_0(x,y)\right|,\end{aligned}$$

where $M:=\|g\|_{C(X)}$. Taking the supremum over $(x,y)\in X$ in (9), for any $\delta>0$, we have,

$$\begin{aligned}\|L_{i,j}(g)-g\|_{C(X)} \leq\ & \omega(g;\delta)\|L_{i,j}(h_0)-h_0\|_{C(X)}+\frac{\omega(g;\delta)}{\delta^2}\|L_{i,j}(\varphi)\|_{C(X)} \\ & +\omega(g;\delta)+M\|L_{i,j}(h_0)-h_0\|_{C(X)}.\end{aligned}$$

If we get $\alpha:=\alpha_{i,j}:=\sqrt{\|L_{i,j}(\varphi)\|}$, then we calculate

$$\|L_{i,j}(g)-g\|_{C(X)}\leq\omega(g;\delta)\|L_{i,j}(h_0)-h_0\|_{C(X)}+2\omega(g;\delta)+M\|L_{i,j}(h_0)-h_0\|_{C(X)}$$

and hence

(10)

$$\|L_{i,j}(g)-g\|_{C(X)}\leq D\left\{\omega(g;\delta)\|L_{i,j}(h_0)-h_0\|_{C(X)}+\omega(g;\delta)+\|L_{i,j}(h_0)-h_0\|_{C(X)}\right\},$$

where $D=\max\{2,M\}$. For a given $r>0$, define the following sets:

$$\begin{aligned}\Phi_i(r) &:= \left\{j\in\mathbb{N}:\Psi(i,j)\geq0,\ \|\ L_{i,j}(g)-g\ \|_{C(X)}\geq r\right\}, \\ \Phi_i^1(\frac{r}{3D}) &:= \left\{j\in\mathbb{N}:\Psi(i,j)\geq0,\ \omega(g;\delta_{i,j})\|\ L_{i,j}(h_0)-h_0\|_{C(X)}\geq\frac{r}{3D}\right\}, \\ \Phi_i^2(\frac{r}{3D}) &:= \left\{j\in\mathbb{N}:\Psi(i,j)\geq0,\ \omega(g;\delta_{i,j})\geq\frac{r}{3D}\right\}, \\ \Phi_i^3(\frac{r}{3D}) &:= \left\{j\in\mathbb{N}:\Psi(i,j)\geq0,\ \|\ L_{i,j}(h_0)-h_0\|_{C(X)}\geq\frac{r}{3D}\right\}.\end{aligned}$$

It follows from (10) that

$$\Phi_i(r)\subset\Phi_i^1(\frac{r}{3D})\cup\Phi_i^2(\frac{r}{3D})\cup\Phi_i^3(\frac{r}{3D}).$$

Also define the sets:

$$\Phi_i^4(\sqrt{\frac{r}{3D}}) \ : \ = \left\{ j \in \mathbb{N} : \Psi(i,j) \geq 0, \ \omega(g;\delta_{i,j}) \geq \sqrt{\frac{r}{3D}} \right\},$$
$$\Phi_i^5(\sqrt{\frac{r}{3D}}) \ : \ = \left\{ j \in \mathbb{N} : \Psi(i,j) \geq 0, \ \|L_{i,j}(h_0) - h_0\|_{C(X)} \geq \sqrt{\frac{r}{3D}} \right\}.$$

Then, observe that $\Phi_i^1(\frac{r}{3D}) \subset \Phi_i^4(\sqrt{\frac{r}{3D}}) \cup \Phi_i^5(\sqrt{\frac{r}{3D}})$. So, we have

$$\Phi_i(r) \subset \Phi_i^2(\frac{r}{3D}) \cup \Phi_i^3(\frac{r}{3D}) \cup \Phi_i^4(\sqrt{\frac{r}{3D}}) \cup \Phi_i^5(\sqrt{\frac{r}{3D}}).$$

Then, from the monotonicity of ϕ, we get

$$\begin{aligned}
& \phi\left(\Phi_i(r) \setminus \{0,1,2,...,k-1\}\right) \\
\leq \ & \phi\left(\left[\Phi_i^2(\frac{r}{3D}) \cup \Phi_i^3(\frac{r}{3D}) \cup \Phi_i^4(\sqrt{\frac{r}{3D}}) \cup \Phi_i^5(\sqrt{\frac{r}{3D}})\right] \setminus \{0,1,2,...,k-1\}\right) \\
\leq \ & \phi\left(\Phi_i^2(\frac{r}{3D}) \setminus \{0,1,2,...,k-1\}\right) + \phi\left(\Phi_i^3(\frac{r}{3D}) \setminus \{0,1,2,...,k-1\}\right) \\
& + \phi\left(\Phi_i^4(\sqrt{\frac{r}{3D}}) \setminus \{0,1,2,...,k-1\}\right) + \phi\left(\Phi_i^5(\sqrt{\frac{r}{3D}}) \setminus \{0,1,2,...,k-1\}\right).
\end{aligned}$$

Then, using the hypotheses (i) and (ii) and considering Lemma 4.2, the proof is complete. □

References

[1] Altomare F. and Campiti M., (1994), *Korovkin-type approximation theory and its applications*, ser. de Gruyter Studies in mathematics, Walter de Gruyter and Co., vol. 17, Berlin.

[2] Anastassiou, G. A. and Duman O., (2011), *Towards intelligent modeling, Statistical approximation theory* (Vol. 14), Springer, Berlin.

[3] Bardaro C., Boccuto A., Demirci K., Mantellini I. and Orhan S., (2015), Triangular $A-$statistical approximation by double sequences of positive linear operators, Results in Mathematics, 68 (3) 271-291.

[4] Bardaro C., Boccuto A., Demirci K., Mantellini I., and Orhan S., (2015), Korovkin-Type Theorems for Modular $\Psi - A-$Statistical Convergence, Journal of Function Spaces, 2015 1-11.

[5] Çınar S., (2021), Triangular $A-$Statistical Relative Uniform Convergence for Double Sequences of Positive Linear Operators, Facta Universitatis, Series: Mathematics and Informatics, 065-077.

[6] Çınar S., Yıldız S. and Demirci K., (2021), Korovkin type approximation via triangular $A-$statistical convergence on an infinite interval, Turkish Journal of Mathematics, 45 (2) 929-942.

[7] Demirci, K. and Dirik F., (2010), Four-dimensional matrix transformation and rate of A-statistical convergence of periodic functions, Mathematical and Computer Modelling, 52 (9-10) 1858-1866.

[8] Demirci K. and Dirik F., (2010), A Korovkin-type approximation theorem for functions of two variables in statistical sense, Turkish Journal of Mathematics, 47 (34) 73-83.

[9] Demirci K. and Dirik F., (2011), Equi-ideal convergence of positive linear operators for analytic P-ideals, Math. Commun., 16 169-178.

[10] Demirci K., Dirik F. and Okçu P., (2017), Approximation in Triangular Statistical Sense to B-Continuous Functions by Positive Linear Operators, Annals of the Alexandru Ioan Cuza University-Mathematics, 63 (3).

[11] Demirci, K. and Orhan S., (2017), Statistical relative approximation on modular spaces, Results in Mathematics, 71 (3) 1167-1184.

[12] Farah I., (2000), Analytic quotients: theory of liftings for quotients over analytic ideals on the integers, Mem. Amer. Math. Soc. 148 XVI+177.

[13] Fast H., (1951), Sur la convergence statistique, Colloq. Math., 2 241-244.

[14] Gadjiev A. D. and Orhan C., (2002), Some approximation theorems via statistical convergence, The Rocky Mountain Journal of Mathematics, 129-138.

[15] Korovkin P.P., (1960), *Linear operators and approximation theory*, Hindustan Publ. Co., New Delhi.

[16] Moricz F., (2004), Statistical convergence of multiple sequences, Arch. Math. (Basel) 81 82-89.

[17] Mrożek N., (2009), Ideal version of Egorov's theorem for analytic P-ideals, J. Math. Anal. Appl., 349 452-458.

[18] Orhan S. and Demirci K., (2014), Statistical $\mathcal{A}$-summation process and Korovkin type approximation theorem on modular spaces, Positivity, 18 (4) 669-686.

[19] Pringsheim A., (1900), Zur theorie der zweifach unendlichen zahlenfolgen, Math. Ann., 53 289-321.

[20] Solecki S., (2005), Local inverses of Borel homomorphisms and analytic P-ideals, Abst. Appl. Anal., 3 207-219.

[21] Ünver M. and Orhan C., (2019), Statistical convergence with respect to power series methods and applications to approximation theory, Numerical Functional Analysis and Optimization, 40 (5) 535-547.

[22] Volkov V.I., (1957), On the convergence of sequences of linear positive operators in the space of two variables, Dokl. Akad. Nauk. SSSR (N.S.) 115, 17-19.

Sinop University Faculty of Sciences and Arts, Department of Mathematics 57000 Sinop, Turkiye [1]
Sinop University Faculty of Sciences and Arts, Department of Mathematics 57000 Sinop, Turkiye [2*] (corresponding author)
Sinop University Faculty of Sciences and Arts, Department of Mathematics 57000 Sinop, Turkiye [3]
Email(s): kamild@sinop.edu.tr [1], fdirik@sinop.edu.tr [2*] (corresponding author), sevdaorhan@sinop.edu.tr [3]

Kantorovich Operators of Order j Based on Pólya Distribution

Vijay Gupta

ABSTRACT. The present article is a study on higher order Kantorovich variant, which are connected with Pólya distribution. We provide some direct estimates for the higher order (j-th order) Kantorovich type integral operators.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 41A25, 41A30

KEYWORDS: Pólya distribution, finite differences, backward difference operator, Kantorovich type integral variant.

1. KANTOROVICH OPERATORS

Convergence is important aspect for problems dealing with linear positive operators in past several years, Gupta and Agarwal in [7] collected and presented many results on this topic. On several other operators, we refer to the recent book [8]. For $\tau > 0$ a parameter and $x \in [0,1]$ the generalization of the well-known Bernstein polynomial was given by Stancu [13] as follows

$$(P_n^{(\tau)} f)(x) = \sum_{i=0}^{n} \binom{n}{i} \frac{x^{[i,-\tau]}(1-x)^{[n-i,-\tau]}}{1^{[n,-\tau]}} f\left(\frac{i}{n}\right), \tag{1}$$

where $u^{[j,-\tau]} = \prod_{r=0}^{j-1}(u + r\tau), j \geq 1;\ \ u^{[0,-\tau]} = 1$. These operators are connected with Pólya distribution and for $\tau = 0$ we get the Bernstein polynomials. Miclăuş [11] studied deeply such operators other special case $\tau = n^{-1}$ was discussed by Lupaş and Lupaş [10].

Very recently Agrawal et al. [2], Baumann [3], Acar et al. [1], Gupta [5] and [9] etc studied some normal and higher order Kantorovich operators. The higher order Pólya-Kantorovich operators for $x \in [0,1]$, was recently proposed in [6] and defined as

$$(\widetilde{K}_n^{(j,\tau)} f)(x) = \frac{n!(n+j)^{j-1}}{(n+j-1)!} \frac{\prod_{l=0}^{j-1}(1+\tau l)}{\tau^j} (\nabla^j \circ P_{n+j}^{(\tau)} \circ F_{\tau,j} f)(x), \tag{2}$$

where

$$(F_{\tau,j} f)(x) = \begin{cases} f(x), & j = 0 \\ \frac{1}{(j-1)!} \displaystyle\int_0^x (x-v)^{j-1} f(v)dv, & j \geq 1. \end{cases}$$

As $P_{n+j}^{(\tau)}$ are convex of order j the operators $\widetilde{K}_n^{(j,\tau)} f$ are linear and positive. For the special case when $j = 1$, these operators reduce to the operators discussed in [6].

2. MOMENTS

Lemma 2.1. *For $j \in \mathbf{N_0}$, $x \in [0,\infty)$ the following equalities hold*

$$\begin{aligned}
(\tilde{K}_n^{(j,\tau)} e_0)(x) &= 1; \\
(\tilde{K}_n^{(j,\tau)} e_1)(x) &= \frac{nx}{(n+j)(1+\tau j)} + \frac{j}{2(n+j)}; \\
(\tilde{K}_n^{(j,\tau)} e_2)(x) &= \frac{n(n-1)x(x+\tau)}{(n+j)^2(1+\tau(j+1))(1+\tau j)} + \frac{(j+1)nx}{(n+j)^2(1+\tau j)} + \frac{j(3j+1)}{12(n+j)^2}
\end{aligned}$$

$$
\begin{aligned}
(\tilde{K}_n^{(j,\tau)}e_3)(x) &= \frac{n(n-1)(n-2)x(x+\tau)(x+2\tau)}{(n+j)^3(1+\tau(j+2))(1+\tau(j+1))(1+\tau j)} \\
&+\frac{3n(n-1)(j+2)x(x+\tau)}{2(n+j)^3(1+\tau(j+1))(1+\tau j)} \\
&+\frac{n(j+1)(3j+4)x}{4(n+j)^3(1+\tau j)}+\frac{j^2(j+1)}{8(n+j)^3} \\
(\tilde{K}_n^{(j,\tau)}e_4)(x) &= \frac{n(n-1)(n-2)(n-3)x(x+\tau)(x+2\tau)(x+3\tau)}{(n+j)^4(1+\tau(j+3))(1+\tau(j+2))(1+\tau(j+1))(1+\tau j)} \\
&+\frac{2n(n-1)(n-2)(j+3)x(x+\tau)(x+2\tau)}{(n+j)^4(1+\tau(j+2))(1+\tau(j+1))(1+\tau j)} \\
&+\frac{n(n-1)(j+2)(3j+7)x(x+\tau)}{2(n+j)^4(1+\tau(j+1))(1+\tau j)} \\
&+\frac{n(j+1)^2(j+2)x}{2(n+j)^3(1+\tau j)}+\frac{j(15j^3+30j^2+5j-2)}{240(n+j)^4}.
\end{aligned}
$$

Proof. By definition of operators, we have

$$
(\tilde{K}_n^{(j,\tau)}e_m)=\frac{m!}{(m+j)!}\frac{n!(n+j)^{j-1}}{(n+j-1)!}\frac{\prod_{l=0}^{j-1}(1+\tau l)}{\tau^j}\nabla^j P_{n+j}^{(\tau)}e_{m+j}. \tag{3}
$$

Using (3) and application of the following identity

$$
\begin{aligned}
\prod_{m=0}^{j-1}(x+\tau m) &= x^j+\frac{j(j-1)}{2}\tau x^{j-1}+\frac{j(j-1)(j-2)(3j-1)}{24}\tau^2x^{j-2} \\
&+\frac{j^2(j-1)^2(j-2)(j-3)}{48}\tau^3x^{j-3} \\
&+\frac{j(j-1)(j-2)(j-3)(j-4)(15j^3-30j^2+5j+2)}{5760}\tau^4x^{j-4} \\
&+\frac{j^2(j-1)^2(j-2)(j-3)(j-4)(j-5)(3j^2-7j-2)}{11520}\tau^5x^{j-5} \\
&+...+(j-1)!\tau^{j-1}x.
\end{aligned}
$$

and applying [6, Lemma 1] and using properties of backward differences we obtain the moments. □

Remark 1. *By using Lemma 2.1, we have*

$$
\begin{aligned}
(\widetilde{K}_n^{(j,n^{-1})}e_0)(x) &= 1, \\
(\widetilde{K}_n^{(j,n^{-1})}e_1)(x) &= \frac{n^2x}{(n+j)^2}+\frac{j}{2(n+j)}, \\
(\widetilde{K}_n^{(j,n^{-1})}e_2)(x) &= \frac{n^2(n-1)x(nx+1)}{(n+j)^3(n+j+1)}+\frac{(j+1)n^2x}{(n+j)^3}+\frac{j(3j+1)}{12(n+j)^2} \\
(\widetilde{K}_n^{(j,n^{-1})}e_3)(x) &= \frac{n^2(n-1)(n-2)x(nx+1)(nx+2)}{(n+j)^4(n+j+2)(n+j+1)}+\frac{3n^2(n-1)(j+2)x(nx+1)}{2(n+j)^4(n+j+1)} \\
&+\frac{n^2(j+1)(3j+4)x}{4(n+j)^4}+\frac{j^2(j+1)}{8(n+j)^3} \\
(\widetilde{K}_n^{(j,n^{-1})}e_4)(x) &= \frac{n^2(n-1)(n-2)(n-3)x(nx+1)(nx+2)(nx+3)}{(n+j)^5(n+j+3)(n+j+2)(n+j+1)} \\
&+\frac{2n^2(n-1)(n-2)(j+3)x(nx+1)(nx+2)}{(n+j)^5(n+j+2)(n+j+1)} \\
&+\frac{n^2(n-1)(j+2)(3j+7)x(nx+1)}{2(n+j)^5(n+j+1)} \\
&+\frac{n^2(j+1)^2(j+2)x}{2(n+j)^4}+\frac{j(15j^3+30j^2+5j-2)}{240(n+j)^4}.
\end{aligned}
$$

Lemma 2.2. *If we denote* $\mu_{n,m}^{(j,n^{-1})}(x) = (\widetilde{K}_n^{(j,n^{-1})}(e_1 - xe_0)^m)$*, then we have*

$$\begin{aligned}
\mu_{n,0}^{(j,n^{-1})}(x) &= 1;\\
\mu_{n,1}^{(j,n^{-1})}(x) &= \frac{j(n+j) - 2jx(2n+j)}{2(n+j)^2};\\
\mu_{n,2}^{(j,n^{-1})}(x) &= \frac{1}{12(n+j)^3(n+j+1)}\Big[(3j^2+j)(n+j)(n+j+1)\\
&\quad + x[24n^3 + 12n^2j - 36nj^3 - 12j^4 - 12j^3 - 24n^2j^2 - 24nj^2]\\
&\quad + x^2[48nj^3 + 48n^2j^2 + 12n^2j + 36nj^2 - 24n^3 + 12j^4 + 12j^3]\Big]\\
&\leq \frac{2}{n+j}\left[x(1-x) + \frac{c_j}{n+j}\right],
\end{aligned}$$

where c_j *is certain constant dependent on* j*.*

The proof of Lemma follows by Remark 1.

3. Convergence

Theorem 3.1. *Let* $f \in C[0,1]$*, then we have*

$$\lim_{n\to\infty} (\widetilde{K}_n^{(j,n^{-1})} f)(x) = f(x),$$

uniformly on $[0,1]$*.*

Proof. Using the well-known Korovkin's theorem and Remark 1, the uniform convergence of $(\widetilde{K}_n^{(j,n^{-1})} f)(x)$ takes place easily. □

Theorem 3.2. *If* f *be a function continuous on* $[0,1]$ *and* f'' *exists at a fixed point* $x \in [0,1]$*, then*

$$\lim_{n\to\infty} n[(\widetilde{K}_n^{(j,n^{-1})} f)(x) - f(x)] \;=\; \left(\frac{j}{2} - 2jx\right) f'(x) + x(1-x)f''(x).$$

Proof. Operating the operators $(\widetilde{K}_n^{(j,n^{-1})} f)$ to the Taylor's formula, we get

$$\begin{aligned}
(\widetilde{K}_n^{(j,n^{-1})} f)(x) - f(x) &= \mu_{n,1}^{(j,n^{-1})}(x) f'(x) + \frac{1}{2}\mu_{n,2}^{(j,n^{-1})}(x) f''(x)\\
&\quad + (\widetilde{K}_n^{(j,n^{-1})} \theta(t,x)(t-x)^2)(x),
\end{aligned}$$

where $\theta \in C[0,1]$ and $\lim_{t\to x} \theta(t,x) = 0$.

Applying Cauchy-Schwarz inequality, we obtain

$$(\widetilde{K}_n^{(j,n^{-1})} \theta(t,x)(t-x)^2)(x) \leq [(\widetilde{K}_n^{(j,n^{-1})} \theta^2(t,x))(x)]^{1/2} [\mu_{n,4}^{(j,n^{-1})}(x)]^{1/2}.$$

Since $\theta^2(x,x) = 0$ and $\theta^2(\cdot,x) \in C[0,1]$, by Theorem 3.1, we obtain

$$\lim_{n\to\infty} (\widetilde{K}_n^{(j,n^{-1})} \left(\theta^2(t,x)\right)(x) = 0$$

uniformly with respect to $x \in [0,1]$. Therefore, from Lemma 2.2, we obtain

$$\lim_{n\to\infty} n(\widetilde{K}_n^{(j,n^{-1})} \left(\theta(t,x)(t-x)^2\right)(x) = 0.$$

The proof of the theorem is immediate by using Lemma 2.2. □

Theorem 3.3. *Let* $f, h \in C^2[0,1]$*, then, for all* $x \in [0,1]$*, we have*

$$\lim_{n\to\infty} n\left[(\widetilde{K}_n^{(j,n^{-1})} fh)(x) - (\widetilde{K}_n^{(j,n^{-1})} f)(x)(\widetilde{K}_n^{(j,n^{-1})} h)(x)\right] = 2x(1-x)f'(x)h'(x).$$

Proof. Obviously, we have

$$\begin{aligned}
& (\widetilde{K}_n^{(j,n^{-1})} fh)(x) - (\widetilde{K}_n^{(j,n^{-1})} f)(x)(\widetilde{K}_n^{(j,n^{-1})} h)(x) \\
= \; & (\widetilde{K}_n^{(j,n^{-1})} fh)(x) - f(x)h(x) - (fh)'(x)\mu_{n,1}^{(j,n^{-1})}(x) - \frac{1}{2}(fh)''(x)\mu_{n,2}^{(j,n^{-1})}(x) \\
& -h(x)\left((\widetilde{K}_n^{(j,n^{-1})} f)(x) - f(x) - f'(x)\mu_{n,1}^{(j,n^{-1})}(x) - \frac{1}{2}f''(x)\mu_{n,2}^{(j,n^{-1})}(x)\right) \\
& -(\widetilde{K}_n^{(j,n^{-1})} f)(x)\left((\widetilde{K}_n^{(j,n^{-1})} h)(x) - h(x) - h'(x)\mu_{n,1}^{(j,n^{-1})}(x) - \frac{1}{2}h''(x)\mu_{n,2}^{(j,n^{-1})}(x)\right) \\
& +\frac{1}{2}\mu_{n,2}^{(j,n^{-1})}(x)\left(f(x)h''(x) + 2f'(x)h'(x) - h''(x)(\widetilde{K}_n^{(j,n^{-1})} f)(x)\right) \\
& +\mu_{n,1}^{(j,n^{-1})}(x)\left(f(x)h'(x) - h'(x)(\widetilde{K}_n^{(j,n^{-1})} f)(x)\right).
\end{aligned}$$

Finally applying Theorem 3.2 and using Lemma 2.2, we get

$$\begin{aligned}
& \lim_{n\to\infty} n\left[(\widetilde{K}_n^{(j,n^{-1})} fh)(x) - (\widetilde{K}_n^{(j,n^{-1})} f)(x)(\widetilde{K}_n^{(j,n^{-1})} h)(x)\right] \\
= \; & \lim_{n\to\infty} nf'(x)h'(x)\mu_{n,2}^{(j,n^{-1})}(x) \\
& + \lim_{n\to\infty} \frac{1}{2}n\, h''(x)\left(f(x) - (\widetilde{K}_n^{(j,n^{-1})} f)(x)\right)\mu_{n,2}^{(j,n^{-1})}(x) \\
& + \lim_{n\to\infty} n\, h'(x)\left(f(x) - (\widetilde{K}_n^{(j,n^{-1})} f)(x)\right)\mu_{n,1}^{(j,n^{-1})}(x) = 2x(1-x)f'(x)h'(x).
\end{aligned}$$

□

Theorem 3.4. *If the function f is bounded on $x \in [0,1]$, then*

$$\|(\widetilde{K}_n^{(j,n^{-1})} f) - f\| \leq \left[1 + (j^4 + j^3 + 5j^2 + 4j + 2)^{1/2}\right]\omega\left(f, n^{-1/2}\right), \; n \geq 1,$$

where $\|\cdot\|$ represent the uniform norm on $[0,1]$ and $\omega(f,\delta)$ is the usual first order moduli of continuity.

Proof. By definition, we have

$$\begin{aligned}
\left|(\widetilde{K}_n^{(j,n^{-1})} f)(x) - f(x)\right| & \leq (\widetilde{K}_n^{(j,n^{-1})} |f(v) - f(x)|)(x) \\
& \leq (\widetilde{K}_n^{(j,n^{-1})} \omega(f, |v-x|))(x).
\end{aligned}$$

Using the following well-known inequality:

$$\omega(f, |v-x|) \leq \left(1 + \sqrt{n}|v-x|\right)\omega\left(f, n^{-1/2}\right),$$

we get

$$\begin{aligned}
\left|(\widetilde{K}_n^{(j,n^{-1})} f)(x) - f(x)\right| & \leq \omega\left(f, n^{-1/2}\right)\left[1 + \sqrt{n}(\widetilde{K}_n^{(j,n^{-1})} |v-x|)(x)\right] \\
& \leq \omega\left(f, n^{-1/2}\right)\left[1 + \sqrt{n}(\mu_{n,2}^{(j,n^{-1})}(x))^{1/2}\right] \\
& \leq \omega\left(f, n^{-1/2}\right)\left[1 + (j^4 + j^3 + 5j^2 + 4j + 2)^{1/2}\right].
\end{aligned}$$

The proof of Theorem 3.4 is completed. □

The Ditzian Totik moduli of order two and corresponding K-functional (see [4]) are given respectively by

$$\omega_2^{\phi}(h,\delta) = \sup_{0<s\leq\sqrt{\delta}} \sup_{x\pm s\phi(x)\in[0,1]} |h(x + s\phi(x)) - 2h(x) + h(x - s\phi(x))|,$$

and

$$K_{2,\phi(x)}(h,\delta) = \inf\{||h-g|| + \delta||\phi^2 g''|| + \delta^2||g''|| : g \in W^2(\phi)\}(\delta > 0)$$

where $\mathrm{W}^2(\phi) = \{g \in C[0,1] : g \in AC_{loc}[0,1], \phi^2 g'' \in C[0,1]\}$ and $\phi(x) = \sqrt{x(1-x)}$. Also, we have

$$K_{2,\phi(x)}(h,\delta) \leq C\omega_2^{\phi}(h,\sqrt{\delta}). \tag{4}$$

The Ditzian Totik moduli of order one is defined as
$$\overrightarrow{\omega}_{\psi}(h,\delta)=\sup_{0<s\leq\delta}\left|h\left(x+\frac{s\psi(x)}{2}\right)-h\left(x-\frac{s\psi(x)}{2}\right)\right|.$$

Theorem 3.5. *Let f be continuous on $[0,1]$, then for $x\in[0,1]$, we have*
$$\begin{aligned}\left|\left|(\widetilde{K}_n^{(j,n^{-1})}f)(x)-f(x)\right|\right| &\leq c_j\omega_2^{\varphi}(f,(n+j)^{-1})\\ &\quad+\overrightarrow{\omega}_{\psi}(f,j(n+j)^{-1})+\omega(f,j(n+j)^{-1}),\end{aligned}$$
where $C>0$ is a constant, $\varphi(x)=\sqrt{x(1-x)}$ and $\psi(x)=\begin{cases}1-4x, & \text{for } x\in[0,1/4]\\ 4x-1, & \text{for } x\in[1/4,1]\end{cases}$.

Proof. We introduce
$$(\widetilde{K}_{j,n}^{*(n^{-1})}f)(x)=(\widetilde{K}_n^{(j,n^{-1})}f)(x)+f(x)-f\left(\frac{n^2x}{(n+j)^2}+\frac{1}{2(n+j)}\right),$$
these modified form preserve linear functions, applying Taylor's formula, we have
$$\begin{aligned}|(\widetilde{K}_{j,n}^{*(n^{-1})}g)(x)-g(x)| &\leq \left(\widetilde{K}_n^{(j,n^{-1})}\left|\int_x^t|t-w|.|g''(w)|dw\right|\right)(x)\\ &\quad+\int_x^{\frac{n^2x}{(n+j)^2}+\frac{1}{2(n+j)}}\left|\frac{n^2x}{(n+j)^2}+\frac{1}{2(n+j)}-w\right||g''(w)|dw.\end{aligned}$$
Take the function $\delta_n^2(x):=\frac{2}{n+j}\left[x(1-x)+\frac{c_j}{n+j}\right]$ concave on $x\in[0,1]$, for $u=\tau x+(1-\tau)t,\tau\in[0,1]$, we get
$$\frac{|t-w|}{\delta_n^2(w)}\leq\frac{|t-x|}{\delta_n^2(x)}.$$
Thus
$$\begin{aligned}|(\widetilde{K}_n^{*(j,n^{-1})}g)(x)-g(x)| &\leq \left(\widetilde{K}_n^{(j,n^{-1})}\left|\int_x^t\frac{|t-w|}{\delta_n^2(w)}dw\right|\right)(x)||\delta_n^2g''||\\ &\quad+\left(\int_x^{\frac{n^2x}{(n+j)^2}+\frac{1}{2(n+j)}}\frac{\left|\frac{n^2x}{(n+j)^2}+\frac{1}{2(n+j)}-w\right|}{\delta_n^2(w)}dw\right)||\delta_n^2g''||\\ &\leq \frac{1}{\delta_n^2(x)}||\delta_n^2g''||\left[\mu_{n,2}^{(j,n^{-1})}(x)+\left(\frac{j(n+j)-2jx(2n+j)}{2(n+j)^2}\right)^2\right].\end{aligned}$$
By Lemma 2.2, we have
$$\begin{aligned}&\mu_{n,2}^{(j,n^{-1})}(x)+\left(\frac{j(n+j)-2jx(2n+j)}{2(n+j)^2}\right)^2\\ &\leq \frac{2}{n+j}\left[x(1-x)+\frac{c_{1,j}}{n+j}\right]+\frac{j^2}{(n+j)^2}\leq\frac{2+j^2}{n+j}\delta_n^2(x).\end{aligned}$$
Thus
$$\begin{aligned}|(\widetilde{K}_n^{*(j,n^{-1})}g)(x)-g(x)| &\leq \frac{2+j^2}{n+j}||\delta_n^2g''||\\ &\leq \frac{2+j^2}{n+j}\left[||\phi^2g''||+\frac{c_{1,j}}{n+j}||g''||\right].\end{aligned}$$
Therefore
$$\begin{aligned}|(\widetilde{K}_n^{(j,n^{-1})}f)(x)-f(x)| &\leq |(\widetilde{K}_n^{*(n^{-1})}f-f)(x)-(f-g)(x)|+|(\widetilde{K}_n^{*(n^{-1})}g)(x)-g(x)|\\ &\quad+|g(x)-f(x)|+\left|f\left(\frac{n^2x}{(n+j)^2}+\frac{1}{2(n+j)}\right)-f(x)\right|\\ &\leq 4||f-g||+\frac{2+j^2}{n+j}\left[||\phi^2g''||+\frac{c_{1,j}}{n+1}||g''||\right]\\ &\quad+\left|f\left(\frac{n^2x}{(n+j)^2}+\frac{1}{2(n+j)}\right)-f(x)\right|.\end{aligned}$$

Taking the infimum on the right hand side over all $g \in \mathrm{W}^2(\varphi)$, we obtain

$$\begin{aligned} |(\widetilde{K}_n^{(j,n^{-1})} f)(x) - f(x)| &\leq c_{2,j} K_{2,\phi}\left(f, (n+j)^{-1}\right) \\ &\quad + \left| f\left(\frac{n^2 x}{(n+j)^2} + \frac{1}{2(n+j)} \right) - f(x) \right|. \end{aligned} \tag{5}$$

Also,

$$\begin{aligned} & \left| f\left(\frac{n^2 x}{(n+j)^2} + \frac{1}{2(n+j)} \right) - f(x) \right| \\ \leq & \left| f\left(x + \frac{j(n+j) - 2jx(2n+j)}{2(n+j)^2} \right) - f\left(x - \frac{j(n+j) - 2jx(2n+j)}{2(n+j)^2} \right) \right| \\ & + \left| f\left(x - \frac{j(n+j) - 2jx(2n+j)}{2(n+j)^2} \right) - f(x) \right| \\ \leq & \overrightarrow{\omega}_{\psi}(f, j(n+j)^{-1}) + \omega(f, j(n+j)^{-1}). \end{aligned} \tag{6}$$

Combining (5) and (6) and using (4) we get the desired result. $\square$

References

[1] Acar T., Aral A. and Rasa I.,(2016), Approximation by k-th order modification of Szász-Mirakyan operators, Studia Scientiarum Mathematicarum Hungarica 53 (3), 379-398.

[2] Agrawal P. N., Ispir N. and Kajla A., (2016) Approximation properties of Lupaş-Kantorovich operators based on Pólya distribution, Rend. Circ. Mat. Palermo 65, 185–208.

[3] Baumann K. , Heilmann M. and Raşa I., (2016) Further results for kth order Kantorovich modification of linking Baskakov type operators, Results Math. 69, 297-315.

[4] Ditzian Z. and Totik V., (1987), Moduli of Smoothness. Springer, New York (1987)

[5] Gupta V., (2021), Higher order Lupaş-Kantorovich operators and finite differences, Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat., RACSAM 115, 100. https://doi.org/10.1007/s13398-021-01034-2

[6] Gupta V., Operators based on Pólya distribution and finite differences, Math. Methods Appl. Sci. DOI: 10.1002/mma.7995

[7] Gupta V. and Agarwal R. P., 2014), Convergence Estimates in Approximation Theory, Springer.

[8] Gupta V. and Rassias M. T., (2021), Computation and Approximation, Ser.: SpringerBriefs in Mathematics, Springer Nature Switzerland AG VIII, 106.

[9] Indrea A., Indrea A. and Pop O.T. (2020), A New class of Kantorovich-type operators, Constr. Math. Anal. 3(3), 116-124.

[10] Lupaş L. and Lupaş A., (1987), Polynomials of binomial type and approximation operators, Studia Univ. Babeş-Bolyai, Mathematica, 32(4), 61–69.

[11] Miclăuş D., (2012), The revision of some results for Bernstein-Stancu type operators, Carpathian J. Math. 28(2), 289–300.

[12] Razi Q., (1989), Approximation of a function by Kantorovich type operators, Mat.Vesnic. 41, 183–192.

[13] Stancu D. D., (1968), Approximation of functions by a new class of linear polynomial operators, Rev. Roum. Math. Pures et Appl. 13, 1173–1194.

Department of Mathematics, Netaji Subhas University of Technology ,Sector 3 Dwarka, New Delhi 110078, India

Email(s): vijaygupta2001@hotmail.com

Spatial Risk Estimation in Tweedie Double Generalized Linear Models

Aritra Halder *[1] Shariq Mohammed [2] Kun Chen [3] and Dipak K. Dey [3]

Abstract: The Tweedie exponential dispersion family constitutes a fairly rich sub-class of the exponential family. In particular, a member, compound Poisson-gamma (CP-g) model has seen extensive use over the past decade for modeling mixed response featuring exact zeros coupled with a continuous gamma tail. This paper proposes a framework to perform residual analysis on CP-g double generalized linear models for spatial uncertainty quantification. Approximations are introduced to the proposed framework to make the procedure scalable, without compromise in accuracy of estimation and model complexity. Proposed framework is applied to quantifying spatial uncertainty in insurance loss costs arising from automobile collision coverage. Scalability is demonstrated by choosing sizable spatial reference domains comprised of groups of states within the United States of America.

2020 Mathematics Subject Classifications: 62P05, 62H11, 62J12, 65K10.

Keywords: double generalized linear models, graph Laplacian, insurance rate-making, spatial analysis, Tweedie models

1. Introduction

Advanced geographic information systems (GIS) are increasingly being used to improve predictive accuracy in a variety of statistical systems. The idea is to incorporate previously untapped spatial information present inherently in recorded data to improve predictive performance of a statistical model. Spatial information can be recorded at many levels, geographical co-ordinates i.e., a longitude-latitude pair (point-referenced), a county/district at observation level, census tract information, a three/five-digit zip code at observation level (areal-referenced). Such information is used to analyze variation in response across a granularity level of interest, for instance, at county/district level, or at zip code level. Along with variation in the response across granularity levels, commonly recorded population covariates, like median age, number of teenage drivers etc. can be potential predictors. Apart from improving predictive accuracy, one may also be interested in identifying areas that behave similarly. This encompasses the idea that neighboring regions show similarity in terms of response; it also includes the possibility of similar response surface characteristics manifesting in non-neighboring locations, due to re-occurring dependence on covariate information. In such a scenario an *ordinal* ranking can be established to identify boundaries derived from variation in response that are different from existing geographic borders. This is referred to as *boundary analysis* in literature [2], [3], [4]. Another related concept which is relevant to context, is the idea of identifying areas that feature rapid changes with respect to a variable of interest. Identifying these zones can also be a part of boundary analysis, and is referred to as wombling [4]. This is also known as “barrier analysis” or “edge detection” in studies of landscape topography, systematic biology, sociology, ecology, and public health.

Spatial information can be incorporated to an existing model in multiple ways. It is both desirable and advantageous to include such information without major changes to an existing modeling structure, by devising a methodology that utilizes already present implementation and builds on it. Consequently, the same unaccounted for spatial information brings in extra variation, that was previously not quantified by the existing model; this excess variation can now be interpreted as *risk* faced by the response. Depending on the nature of response, qualitative characterizations of this risk may vary, for instance when considering the number of road accidents on interstate highways, adverse characterizations of risk would follow. Naturally, the ordinal nature of rankings can then be of considerable interest to an investigator looking to put in place categorically different measures for spatial clusters identified based solely on their "riskiness". Considering the bigger picture, this results in a *nested model*, that consists of structurally dependent and independent components, where proper specification of the structured component aids in explaining excess variation in response.

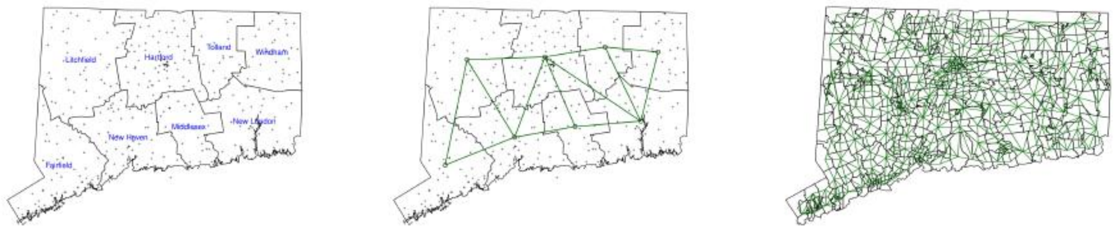
In practice existing modeling implementation encountered most commonly are generalized linear models (GLMs) [39], [41], which provide a very flexible structure for modeling different types of response. While constructing GLMs the response is modeled to follow a probability distribution, in that regard *exponential families* of distributions provide a very general class of choices. The idea of GLMs could be extended to a more general class of models, namely *dispersion models* [38], [40]. The primary reason behind having these elaborate families is to relax unnecessary, restrictive assumptions of normality on the response, in case significant signs of non-normality are evident. GLMs employ a technique called analysis of deviance, which is a generalization of analysis of variance. If deviance is interpreted as a measure of distance between realizations and a location parameter, then based on general functional forms and properties of deviance, dispersion models are divided into broad sub-classes. In this paper we are interested in exploring spatial risk estimation in GLMs for particularly one of those classes, called the Tweedie exponential dispersion models (EDMs) [40], [37], [35].

Tweedie EDMs feature an additional index parameter which classifies different sub-classes of distributions within the family. Commonly known distributions like Poisson, gamma and inverse Gaussian are special cases of Tweedie EDMs (see table 1, [34]). Poisson distributions are obtained if the index parameter, $p = 1$, the Gamma distribution is obtained if $p = 2$, whereas the inverse Gaussian distribution is obtained when $p = 3$. The Poisson distribution is commonly used to model count data, whereas gamma and inverse Gaussian distributions find their use in modeling positive continuous data. However, when $p \in (1, 2)$, we obtain the *compound Poisson-gamma* (CP-g) distribution which is ideal for modeling positive continuous data with exact zeros. Applications include modeling weights of fish species in a single sample or trawl, appearance of exact zeros occurring if a particular species is not caught, otherwise a positive (continuous) weight is recorded [29]. The concept of catch per-unit effort (CPUE) in connection to relative fish stock and abundance is plagued by the zero-catch problem, which is dealt with using these distributions [31]. Other examples include modeling rainfall amount with exact zeros occurring in the case where there is no rainfall, otherwise a positive amount is recorded [33]. Simultaneous modeling of occurrence and size/amount of insurance claims also use CP-g distributions; here exact zeros occur in case of records that show no accidents during a chosen policy-period, whereas a positive claim size is recorded in case of an accident [24], [23]. While modeling insurance claims along with modeling the mean [24], dispersion modeling was also considered [23] to result in simultaneous GLMs for mean and dispersion, which are termed as double generalized linear models (DGLMs). Exact zeros in political contributions or donations (in US dollars) are also modeled similarly [28]. Forest degradation which involves sampling of biomass loss produces continuous data with a large number of exact zeros (no disturbance, implying no loss) which are again analyzed using such distributions [27]. In all of the above examples existing alternatives include removing exact zeros, or adding a small constant to zeros such that "log(0)" problems are avoided when fitting GLMs using a logarithm link function. Apart from applications in varied fields of study, notable developments in methodology include variable selection procedures involving grouped elastic net being developed for Tweedie CP-g models [21], likelihood based Bayesian approaches, that are well-suited alternatives to quasi-likelihood methods in terms of inference for Tweedie compound Poisson *mixed* models have been explored [22]. Machine learning algorithms, like gradient boosting have been used in conjunction with CP-g models to achieve better performance while predicting insurance premiums [25].

Another integral fact that is apparent from examples mentioned above, is the existence of spatial information through location, which may be not have been taken into account and hence associated variation being unaccounted for by implemented models. Presence of such information could help explorations into effects that adjacent regions have on the response. Throughout this paper we use a *first-order adjacency* to account for spatial correlation among locations or regions. The concept of first-order adjacency simply states, if two regions share boundaries then we put an edge between their centroids, and call them neighbors. Naturally, adjacency information has dual representations, a graph as shown in figs. 1 (b) and (c), or a matrix, the order being number of locations, L. We can also see from figures (b) and (c)

TABLE 1. The Tweedie family of distributions for varying values of the index parameter, p, with respective supports S and parameter spaces, Ω.

Tweedie EDMs	p	S	Ω	Examples
Extreme stable	$p < 0$	$\mathbb{R}$	$\mathbb{R}^+$	
Normal	$p = 0$	$\mathbb{R}$	$\mathbb{R}$	–
No EDMs exist	$0 < p < 1$	–	–	–
Discrete	$p = 1$	$\mathbb{N} \cup \{0\}$	$\mathbb{R}^+$	Poisson
Poisson-gamma	$1 < p < 2$	$\mathbb{R}^+ \cup \{0\}$	$\mathbb{R}^+$	–
Gamma	$p = 2$	$\mathbb{R}^+$	$\mathbb{R}^+$	–
Positive stable	$p > 2$	$\mathbb{R}^+$	$\mathbb{R}^+$	inverse Gaussian



(A) (B) (C)

FIGURE 1. Figure showing the (a) 8 counties with zip codes, (b) construction of adjacency based on 8 counties, (c) construction of adjacency based on 282 zip-codes for state of Connecticut.

that depending on a chosen level of granularity, an adjacency graph can have small (county level) to sufficiently large number (zip code level) of edges. Figure 1a shows the 8 counties and zipcodes in the state of Connecticut. Adjacency matrices have been used in multiple situations to capture spatial correlation, some of the earliest and celebrated applications being simultaneous and conditional autoregressive areal models (SAR and CAR respectively) [19], [20]. They are also employed extensively in developing hierarchical Bayesian spatial areal models [18]. Formally stating, if we denote an adjacency matrix by $A = (A_{i_1 i_2})$ of the order $L \times L$ then,

$$A_{i_1 i_2} = \begin{cases} 1 & \text{if } i_2 \in \mathcal{N}(i_1) \\ 0 & \text{otherwise} \end{cases}, \tag{1}$$

where $\mathcal{N}(i_1)$ denotes the set containing neighbors (i.e., locations that share a boundary) of location i_1. In the ensuing discussion we shall only look at adjacency matrices generated by locations that share boundaries, or are first-order neighbors and not locations that are "neighbors" of neighbors (i.e second/higher order).

In section 2 we consider revisiting the formulation, development and necessary details regarding double GLMs (DGLMs) for Tweedie CP-g distributions. Reviewing the choice of an appropriate likelihood function and further characteristics of the Tweedie family of distributions are also discussed in this section. We consider incorporating spatial information as an un-observable fixed effect into the formulated DGLM, where *penalized estimation* of the same fixed effect leads to solving an optimization problem. Section 3 states and discusses necessary conditions for obtaining a solution to this problem. Consequently, the approach we describe essentially fits a DGLM having a penalized spatial fixed effect, where nature of the penalty chosen promotes structural shrinkage. Simulations that demonstrate the efficacy of our approach over existing alternatives are shown in section 4. The penalty we propose can be interpreted as an extension of the ridge penalty [17], naturally making it a baseline for comparing performance. Section 5 describes the nature and results of our real data application, showcasing performance of our approach for characteristically different response surfaces in selective states or, groups of states in USA. Finally, sections 6, 7 and Appendix include conclusions comments about further developments and required proofs, tables and additional figures respectively.

2. Characterizations and DGLMs for compound Poisson distributions

We start this section with some notation that will be used throughout the ensuing discussion. Let $y_{ij}, \mu_{ij} \in \mathbb{R}, \phi_{ij} \in \mathbb{R}^+$ denote the response, mean, dispersion respectively and, $x_{ij}^\top \in \mathbb{R}^{m_1}$, $z_{ij}^\top \in \mathbb{R}^{m_2}$ denote the observed covariate vectors which can be adjusted to include an intercept for the j-th observation at the i-th location, where $i = 1, \ldots, L$ and $j = 1, \ldots, n_i$; with $\sum_{i=1}^{L} n_i = N$ being total number of observations. Notations without subscripts are used to describe model formulations, to avoid cumbersome expressions.

2.1. **Probability density characterizations.** Probability distributions for EDMs, where the response, y can be discrete or continuous depending on the nature of application,

$$f(y;\theta,\phi) = a(y,\phi)\exp\left\{\frac{y\theta-\kappa(\theta)}{\phi}\right\}. \tag{2}$$

In eq. (2) θ is called the canonical parameter, $\kappa(\theta)$ is a known function called the cumulant function, ϕ is the dispersion parameter and $a(y,\phi)$ is a normalizing constant that ensures (2) is a probability function. Eq. (2) is called the canonical form for EDM densities, with other parametrizations being possible. As we can see from table (2) that $a(y,\phi)$ may not always have a closed form (in case of CP-g). For EDMs we have some well-known relations, $E(y) = \mu = \kappa'(\theta)$ and $\mathrm{Var}(y) = \phi\kappa''(\theta)$ [39], [41], [34]. Due to the relationship between θ and μ, $\kappa''(\theta)$ can also be expressed as a function of μ, which is denoted by *variance function* $V(\mu)$; which uniquely corresponds to an exponential dispersion model [38], [34] pgs. 217, [16]. We focus on EDMs with a power variance function $\phi\mu^p$, where $p \in (1,2)$, table 2 shows necessary the details for probability density, or probability mass functions for other members of the Tweedie family. In particular, the probability density function for a CP-g distribution can be expressed in its unit-deviance form as,

$$f(y;\mu,\phi,p) = b(y,\phi,p)\exp\left\{-\frac{1}{\phi}\int_y^\mu \frac{y-u}{V(u)}\mathrm{d}u\right\} = b(y,\phi,p)\exp\left\{-\frac{1}{\phi}\int_y^\mu \frac{y-u}{u^p}\mathrm{d}u\right\}, \tag{3}$$

where $d(y,\mu) = -2\int_y^\mu \frac{y-u}{V(u)}\mathrm{d}u$ is defined as the deviance, i.e. a measure of discrepancy between observation, y and its expected value, μ. An alternative characterization for a known $p \in (1,2)$, random variable Y follows a CP-g distribution [38] if,

$$Y = \begin{cases} 0 & M = 0 \\ \sum\limits_{i=1}^{M} C_i & M > 0 \end{cases}, \qquad M \sim \mathrm{Poisson}(\xi), \qquad C_i \overset{iid}{\sim} \mathrm{Gamma}(\eta,\zeta), \tag{4}$$

Table 2. Commonly used members of the Tweedie family of distributions with their index parameter (p), variance function ($V(\mu)$), cumulant function ($\kappa(\theta)$), canonical parameter (θ), dispersion (ϕ), deviance ($d(y,\mu)$), normalizing constant ($b(y,\phi)$), support (S), and respective parameter spaces for mean (Ω) and the natural parameter (Θ).

Tweedie EDMs	p	$V(\mu)$	$\kappa(\theta)$	θ	ϕ	$d(y,\mu)$	$b(y,\phi,p)$	S	Ω	Θ
Normal	0	1	$\frac{\theta^2}{2}$	μ	σ^2	$(y-\mu)^2$	$\exp\{-(y^2/\phi+\log 2\pi\phi)/2\}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
Poisson	1	μ	$\exp(\theta)$	$\log(\mu)$	1	$2\Big\{y\log\frac{y}{\mu}-(y-\mu)\Big\}$	$\frac{y^{-y}\exp(-y)}{y!}$	$\mathbb{N}\cup\{0\}$	$\mathbb{R}^+$	$\mathbb{R}$
Poisson-gamma	$(1,2)$	μ^p	$\frac{\{(1-p)\theta\}^{(2-p)/(1-p)}}{2-p}$	$\frac{\mu^{1-p}}{1-p}$	ϕ	$2\Big\{\frac{\max(y,0)^{2-p}}{(1-p)(2-p)}-\frac{y\mu^{1-p}}{1-p}+\frac{\mu^{2-p}}{2-p}\Big\}$	–	$\mathbb{R}^+\cup\{0\}$	$\mathbb{R}^+$	$\mathbb{R}^-$
Gamma	2	μ^2	$-\log(-\theta)$	$-\frac{1}{\mu}$	ϕ	$2\Big\{-\log\frac{y}{\mu}+\frac{y-\mu}{\mu}\Big\}$	$\frac{\exp\{-(1+\log\phi)/\phi\}}{y\Gamma(1/\phi)}$	$\mathbb{R}^+$	$\mathbb{R}^+$	$\mathbb{R}$
Inverse Gaussian	3	μ^3	$-\sqrt{-2\theta}$	$-\frac{1}{2\mu^2}$	ϕ	$\frac{(y-\mu)^2}{\mu^2 y}$	$\exp\big(-\{(\phi y)^{-1}+\log(2\pi y^3\phi)\}/2\big)$	$\mathbb{R}^+$	$\mathbb{R}^+$	$\mathbb{R}$

where, M is independent of C_i. Above definition also demonstrates the fact that at $M = 0$, $Y = 0$ with some non-zero probability, followed by $M > 0$ resulting in a sum of gamma random variables producing a skewed continuous distribution on $\mathbb{R}^+$. It can be shown that the two characterization are equivalent by deriving and equating cumulant generating functions for densities in eqs. (3) and (4) [23], [22]. From table 2 and eq. (4) it is evident that $b(y,\phi,p)$

needs to be approximated for obtaining a closed form of the density while characterizing CP-g densities as EDMs. Analogously, evaluating the marginal density of Y in alternative characterization shown in eq. (4) results in an infinite sum representation of $b(y,\phi,p)$ which can be approximated in multiple ways [13], [14].

2.2. **Double generalized linear models.** DGLMs were considered as a further generalization to GLMs for exponential families [15], where the dispersion parameter ϕ was no longer required to be a constant across observations. This resulted in simultaneous GLMs where mean μ and dispersion ϕ both varied across observations as described through the double generalized quasi-likelihood model,

$$g_1(\mu) = x^\top \boldsymbol{\beta}, \qquad \mathrm{var}(y) = \phi V(\mu), \qquad g_2(\phi) = z^\top \boldsymbol{\gamma}. \tag{5}$$

Here $g_1(\cdot)$, $g_2(\cdot)$ are monotonic link functions, $\boldsymbol{\beta} \in \mathbb{R}^{m_1}$, $\boldsymbol{\gamma} \in \mathbb{R}^{m_2}$ are model coefficients and $x^\top \in \mathbb{R}^{m_1}$ and $z^\top \in \mathbb{R}^{m_2}$ are predictor vectors for mean and dispersion GLMs respectively. Based on how we want to explain the response using predictors or covariates, the type of characterization along with choice of likelihood approximation adopted varies. For instance, in likelihood-based variable selection approaches [21], where ϕ is assumed to be constant resulting in a GLM for mean, any desired approximation of $b(y,\phi,p)$, $1 < p < 2$ can be accommodated into the estimation procedure. Whereas in likelihood based methods like [22], the type of approximation used affects the parameter estimation significantly. Alternatively, an approach that allows for a variable dispersion ϕ, similar to the model in eq. (5), would require a particular approximation of $b(y,\phi,p)$, producing extended quasi-likelihoods (EQLs) for CP-g models [23], [11], [12]. We will be working with DGLMs, using marginal likelihood approximations involving Fourier inversion of the characteristic function [14] and joint likelihood for (Y,M) as defined in eq. (4). The joint likelihood for (y,m) is,

$$\alpha = \frac{2-p}{p-1}, \qquad b(y,\phi,p) = \frac{1}{m!\Gamma(m\alpha)y}\left\{\frac{y^\alpha \phi^{-(\alpha+1)}}{(p-1)^\alpha(2-p)}\right\}^m, \tag{6a}$$

where $\Gamma(x) = (x-1)!$ is the Gamma function,

$$f(y,m;\mu,\phi,p) = b(y,\phi,p)\exp\left\{\frac{1}{\phi}t(y,\mu,p)\right\}, \qquad t(y,\mu,p) = y\frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}. \tag{6b}$$

However, if $m = y = 0$, $f(0,0;\mu,\phi,p) = \exp\{-\frac{1}{\phi}t(0,\mu,p)\} = \exp\left\{-\frac{1}{\phi}\frac{\mu^{2-p}}{2-p}\right\}$, which completes the specification [24], [23]. In this paper we extend DGLM in eq. (5) to include an additional fixed effect,

$$g_1(\mu) = x^\top \boldsymbol{\beta} + r^\top \boldsymbol{\alpha}, \qquad \mathrm{var}(y) = \phi V(\mu), \qquad g_2(\phi) = z^\top \boldsymbol{\gamma}, \tag{7}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{L\times 1}$ is the un-observable fixed effect. Particularly we interpret α_i as a spatial effect corresponding to location i, hence r is a $L \times 1$ vector of 0's with 1 in exactly one entry indicating the corresponding index for a location.

In the case of an existing implementation of a DGLM, the quantities $o^{(1)} = x^\top \widehat{\boldsymbol{\beta}}$ and $o^{(2)} = z^\top \widehat{\boldsymbol{\gamma}}$ and p are known. Hence the *negative log-likelihood* is given by,

$$\ell(\boldsymbol{\alpha}) = -\sum_{i=1}^{L}\sum_{j=1}^{n_i} \frac{1}{g_2^{-1}(o_{ij}^{(2)})} t\Big(y_{ij}, g_1^{-1}(o_{ij}^{(1)} + r_{ij}^\top \boldsymbol{\alpha}), p\Big) + c\Big(m_{ij}, y_{ij}, g_2^{-1}(o_{ij}^{(2)}), p\Big) I(y_{ij} > 0), \tag{8}$$

where $c(\cdot)$ is a known function ([13] pg. 6) and $I(\cdot)$ stands for the indicator function. We shall assume that p is known from the existing implementation of DGLM. For most implementations link functions, $g_1(\cdot)$ and $g_2(\cdot)$ are both assumed to be logarithmic, and the covariate vectors/matrices $x^\top$ and $z^\top$ need not necessarily be the same [23]. For a GLM, the *canonical link* function is defined as $g(\cdot) = (\kappa')^{-1}(\cdot)$ such that $E(y) = \kappa'(\theta) = g^{-1}(\theta)$, where $'$ denotes the first derivative of a function. It can be readily seen from eqs. (3), (6b) and table (2) that a logarithmic function is *not* the canonical link for CP-g GLMs.

On further inspection, it can be seen that cross-derivatives of the negative log-likelihood in eq. (6b) have zero expectations,

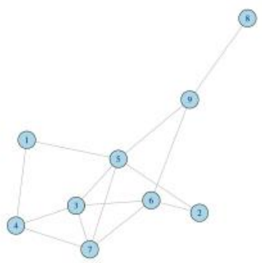
$$E\left(-\frac{\partial^2}{\partial\phi\partial\mu}\log f\right) = E\left(\frac{1}{\phi^2}\frac{y-\mu}{\mu^p}\right) = 0, \qquad E\left(-\frac{\partial^2}{\partial p\partial\mu}\log f\right) = E\left(-\frac{\log(\mu)}{\phi}\frac{y-\mu}{\mu^p}\right) = 0. \tag{9}$$

As a result, off-diagonal elements for μ in the Fisher's information matrix of (μ, ϕ, p) are 0, implying μ is statistically orthogonal to ϕ and p, resulting in $\boldsymbol{\gamma}$, p being independent of $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$. This property insulates the estimation of $\boldsymbol{\alpha}$ from inaccuracies that are associated with using likelihood approximations (EQLs or penalized EQLs) in the existing implementation of a DGLM, for details see [23], pgs. 148 or, [22] pgs. 747, this will also be demonstrated through a sensitivity analysis to p in section 4.

3. Algorithm and computation

3.1. **Graph Laplacian.** Before getting into the optimization problem, we digress briefly to illustrate relevant and related concepts in graph theory. A graph is represented by (V, E), where V is a set of vertices and E is a set of edges, or pairs $(i_1, i_2) \in V$. A graph is said to be un-directed if $(i_1, i_2) \in E \Leftrightarrow (i_2, i_1) \in E$, for example, see figure below in eq. (10); in the notation of eq. (1) $i_2 \in \mathcal{N}(i_1) \Leftrightarrow i_1 \in \mathcal{N}(i_2)$. A diagonal matrix defined as, $D = (D_{i_1 i_1}) = \sum_{i_2} A_{i_1 i_2}$ is defined as a degree matrix, where A is an adjacency matrix as defined in eq. (1). An example is shown below in figure and equation (10).

(10)



$$A = \begin{pmatrix} 0&0&0&1&1&0&0&0&0 \\ 0&0&0&0&1&1&0&0&0 \\ 0&0&0&1&1&1&1&0&0 \\ 1&0&1&0&0&0&1&0&0 \\ 1&1&1&0&0&0&1&0&1 \\ 0&1&1&0&0&0&1&0&1 \\ 0&0&1&1&1&1&0&0&0 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&1&1&0&1&0 \end{pmatrix}, \quad D = \begin{pmatrix} 2&0&0&0&0&0&0&0&0 \\ 0&2&0&0&0&0&0&0&0 \\ 0&0&4&0&0&0&0&0&0 \\ 0&0&0&3&0&0&0&0&0 \\ 0&0&0&0&5&0&0&0&0 \\ 0&0&0&0&0&4&0&0&0 \\ 0&0&0&0&0&0&4&0&0 \\ 0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&3 \end{pmatrix}$$

The *Laplacian* for a graph is defined as $W = D - A$, which is an element-wise difference between degree and adjacency matrices. If we interpret spatial effect as functions, $\boldsymbol{\alpha} : V \to \mathbb{R}$ then the graph Laplacian, W can be equivalently defined as a linear operator,

$$\boldsymbol{\alpha}^\top W \boldsymbol{\alpha} = \frac{1}{2} \sum_{i_1=1}^{L} \sum_{i_2 \in \mathcal{N}(i_1)} (\alpha_{i_1} - \alpha_{i_2})^2, \qquad \forall\, \boldsymbol{\alpha} \in \mathbb{R}^L. \tag{11}$$

As a result, regularization based on graph Laplacian matrices penalize change between adjacent vertices thereby introducing local smoothness [10]. In this paper we work with undirected graphs describing neighborhood structures corresponding to locations, for example, latitude-longitude pairs for counties or zip codes, as shown in fig. 1 on a map, associated adjacency, degree and Laplacian matrices.

(A) (B) (C)

FIGURE 2. Plots showing (a) 896 zip codes for three states Connecticut (CT), Massachusetts (MA) and Rhode Island (RI), locations marked by + (b) the zip code level adjacency matrix, with edges between zip codes within any state colored `darkgreen`, and edges between zip codes across states colored `yellow` (c) zip code level adjacency graphs colored w.r.t states, after the boundary edges have been removed (adjacency graphs for CT, MA and RI are in `darkgreen`, `darkblue` and `violet` respectively).

Sparse matrices consist larger number of zero entries; the adjacency matrix, A in eq. (10) is an example of a sparse matrix. Working with a single state, adjacency matrices for associated neighborhood structures between zip codes are comparatively less *sparse* as opposed to adjacency matrices generated when considering multiple states together. If we consider proportion of zero entries as a measure of sparsity for adjacency matrices, then considering states of Connecticut (CT), Massachusetts (MA) ad Rhode Island (RI) in conjunction the adjacency matrix generated contains 99.4% zeros; individually however, adjacency matrices for CT, MA and RI contain 98.1%, 99.02% and 94.1% of zero entries. We use this sparsity to our advantage by introducing an *approximation* to adjacency for matrices when considering multiple states in our analysis. Let us consider S states in conjunction, if A denotes the associated adjacency matrix, we approximate it by, A_a defined as

$$A = A_a + A_\epsilon, \qquad A_a = \begin{pmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_S \end{pmatrix}, \qquad A_\epsilon = \begin{pmatrix} O & \epsilon_{1,2} & \cdots & \epsilon_{1,S} \\ \epsilon_{1,2} & O & \cdots & \epsilon_{2,S} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{1,S} & \epsilon_{2,S} & \cdots & O \end{pmatrix},$$

where $A_1, \ldots, A_S$ are the adjacency matrices for S states respectively, and ϵ_{k_1,k_2}, with $k_1, k_2 = 1, 2, \ldots, S$, can be interpreted as the cross-state adjacency for pair (k_1, k_2). Figure (2) illustrates this concept for $S = 3$ states viz., CT, MA and RI. This approximation is carried forward to associated degree and Laplacian matrices. Consequently, resulting Laplacian matrices are block diagonal. The computational advantage of having block diagonal Laplacian matrices is immediately evident from nature of solution to the optimization problem described in the following.

3.2. **Optimization problem.** Consider the spatial effect, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_L)^\top$, in this section we primarily focus on solving the minimization problem,

$$\widehat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} F(\boldsymbol{\alpha}), \quad F(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha}) + P(\boldsymbol{\alpha}; \lambda_1, \lambda_2), \quad P(\boldsymbol{\alpha}; \lambda_1, \lambda_2) = \frac{1}{2}\big[\lambda_1 \boldsymbol{\alpha}^\top \boldsymbol{\alpha} + \lambda_2 \boldsymbol{\alpha}^\top W \boldsymbol{\alpha}\big], \tag{12}$$

where W is the graph Laplacian as defined in eq. (11), $\lambda_1, \lambda_2 > 0$ are tuning parameters for the ridge and Laplacian regularization respectively. The penalty function used is similar to the network cohesion penalty proposed by [1]. Under a logarithmic link the negative log-likelihood is,

$$\ell(\boldsymbol{\alpha}) = \sum_i \sum_j \widehat{\phi}_{ij}^{-1} \left[y_{ij} \frac{e^{-(p-1)(\widehat{\mu}_{ij} + r_{ij}^\top \boldsymbol{\alpha})}}{p-1} + \frac{e^{(2-p)(\widehat{\mu}_{ij} + r_{ij}^\top \boldsymbol{\alpha})}}{2-p} \right] - c\Big(y_{ij}, \widehat{\phi}_{ij}, p\Big) I(y_{ij} > 0),$$

where $\widehat{\mu}_{ij}$ and $\widehat{\phi}_{ij}$ are fitted mean and dispersion respectively. Flexibility in structure and advantages of having $P(\boldsymbol{\alpha}, \lambda_1, \lambda_2)$ as a penalty is immediately evident for instances where spatial clustering needs to be introduced. The ridge part penalizes magnitude of estimated spatial effects by regularizing L^2-norm, $||\boldsymbol{\alpha}||_2^2 = \langle \boldsymbol{\alpha}, \boldsymbol{\alpha} \rangle = \boldsymbol{\alpha}^\top \boldsymbol{\alpha}$, while penalty on the Laplacian promotes local neighborhood smoothing on vertices by regularizing induced semi-norm, $||\boldsymbol{\alpha}||_W = \langle \boldsymbol{\alpha}, W\boldsymbol{\alpha} \rangle = \boldsymbol{\alpha}^\top W \boldsymbol{\alpha}$. Solution to optimization problem in eq. (12) is obtained using a majorization descent (MD) algorithm, which utilizes the majorization-minimization (MM) principle, for further details see [7], [8], [6]. Properties and details of proposed MD algorithm are shown below.

Let us denote the $L \times 1$ gradient vector and, $L \times L$ Hessian matrix for the negative log-likelihood $\ell(\boldsymbol{\alpha})$, as $\nabla_1(\boldsymbol{\alpha})$ and $\nabla_2(\boldsymbol{\alpha})$ respectively, which have the following expressions,

$$\nabla_{1,i}(\boldsymbol{\alpha}) = \frac{\partial \ell(\boldsymbol{\alpha})}{\partial \alpha_i} = \sum_j \widehat{\phi}_{ij}^{-1} r_{ij}^\top \Big[- y_{ij} e^{-(p-1)\left(\widehat{\mu}_{ij} + r_{ij}^\top \boldsymbol{\alpha}\right)} + e^{(2-p)\left(\widehat{\mu}_{ij} + r_{ij}^\top \boldsymbol{\alpha}\right)} \Big], \tag{13a}$$

$$\nabla_{2,ii}(\boldsymbol{\alpha}) = \frac{\partial^2 \ell(\boldsymbol{\alpha})}{\partial \alpha_i^2} = \sum_j \widehat{\phi}_{ij}^{-1} r_{ij}^\top \Big[(p-1) y_{ij} e^{-(p-1)\left(\widehat{\mu}_{ij} + r_{ij}^\top \boldsymbol{\alpha}\right)} + (2-p) e^{(2-p)\left(\widehat{\mu}_{ij} + r_{ij}^\top \boldsymbol{\alpha}\right)} \Big] r_{ij}, \tag{13b}$$

where $i = 1, \ldots, L$ and $\nabla_{2,i_1 i_2}(\boldsymbol{\alpha}) = 0$ for all $i_1 \neq i_2$. Hence, $\nabla_2(\boldsymbol{\alpha})$ is a diagonal matrix. If $\boldsymbol{\alpha}^{(t)}$ is the updated estimate of spatial effect from $\boldsymbol{\alpha}$, then we obtain $\boldsymbol{\alpha}^{(t)}$ by solving

$$\arg\min_{\boldsymbol{\alpha}^{(*)}} \ell(\boldsymbol{\alpha}) + (\boldsymbol{\alpha}^{(*)} - \boldsymbol{\alpha})^\top \nabla_1(\boldsymbol{\alpha}) + \frac{1}{2} (\boldsymbol{\alpha}^{(*)} - \boldsymbol{\alpha})^\top \big(\mathrm{I}_L + \nabla_2(\boldsymbol{\alpha})\big) (\boldsymbol{\alpha}^{(*)} - \boldsymbol{\alpha}) + P(\boldsymbol{\alpha}^{(*)}; \lambda_1, \lambda_2), \tag{14}$$

which admits a closed form solution. I_L is the L-dimensional identity matrix. After some algebra it can be shown that,

$$\boldsymbol{\alpha}^{(t)} = \big[(\lambda_1 + 1)\mathrm{I}_L + \lambda_2 W + \nabla_2(\boldsymbol{\alpha}) \big]^{-1} \big\{ (\mathrm{I}_L + \nabla_2(\boldsymbol{\alpha}))\boldsymbol{\alpha} - \nabla_1(\boldsymbol{\alpha}) \big\}. \tag{15}$$

For the estimates $\boldsymbol{\alpha}^{(t)}$ and $\boldsymbol{\alpha}$ we have,

$$\ell(\boldsymbol{\alpha}^{(t)}) \leq \ell(\boldsymbol{\alpha}) + (\boldsymbol{\alpha}^{(t)} - \boldsymbol{\alpha})^\top \nabla_1(\boldsymbol{\alpha}) + \frac{1}{2} (\boldsymbol{\alpha}^{(t)} - \boldsymbol{\alpha})^\top \big(\mathrm{I}_L + \nabla_2(\boldsymbol{\alpha})\big) (\boldsymbol{\alpha}^{(t)} - \boldsymbol{\alpha}) = \mathcal{L}\big(\boldsymbol{\alpha}^{(t)} | \boldsymbol{\alpha}\big). \tag{16}$$

The above inequality in eq. (16) follows from second order Taylor expansion. Therefore,

$$\begin{aligned} F(\boldsymbol{\alpha}^{(t)}) - F(\boldsymbol{\alpha}) &= \ell(\boldsymbol{\alpha}^{(t)}) + \frac{1}{2} \Big[\lambda_1 ||\boldsymbol{\alpha}^{(t)}||_2^2 + \lambda_2 ||\boldsymbol{\alpha}^{(t)}||_W^2 \Big] - \ell(\boldsymbol{\alpha}) - \frac{1}{2} \Big[\lambda_1 ||\boldsymbol{\alpha}||_2^2 + \lambda_2 ||\boldsymbol{\alpha}||_W^2 \Big], \\ &\leq \ell(\boldsymbol{\alpha}) + (\boldsymbol{\alpha}^{(t)} - \boldsymbol{\alpha})^\top \nabla_1(\boldsymbol{\alpha}) + \frac{1}{2} (\boldsymbol{\alpha}^{(t)} - \boldsymbol{\alpha})^\top \big(\mathrm{I}_L + \nabla_2(\boldsymbol{\alpha})\big) (\boldsymbol{\alpha}^{(t)} - \boldsymbol{\alpha}) \\ &\quad + \frac{1}{2} \Big[\lambda_1 ||\boldsymbol{\alpha}^{(t)}||_2^2 + \lambda_2 ||\boldsymbol{\alpha}^{(t)}||_W^2 \Big] - \frac{1}{2} \Big[\lambda_1 ||\boldsymbol{\alpha}||_2^2 + \lambda_2 ||\boldsymbol{\alpha}||_W^2 \Big] - \ell(\boldsymbol{\alpha}), \\ &\leq 0. \end{aligned}$$

The first inequality follows from eq. (16) and the last equality follows from update in eq. (14), for more detailed calculations refer to the Appendix. The algorithm derived above is summarized as algorithm 1.

Algorithm 1 The MD algorithm for estimating penalized spatial effects from a fitted compound Poisson model.

(1) Fit a compound Poisson DGLM without spatial effects $\boldsymbol{\alpha}$, to obtain
  - fitted mean $\widehat{\mu}_{ij}$,
  - fitted dispersion $\widehat{\phi}_{ij}$.

(2) Initialize $\boldsymbol{\alpha}$.

(3) Repeat until $F(\boldsymbol{\alpha})$ converges,
  - Compute $\nabla_1(\boldsymbol{\alpha})$ using eq. (13a)
  - Compute $\nabla_2(\boldsymbol{\alpha})$ using eq. (13b)
  - Compute $\boldsymbol{\alpha}^{(t)}$ using eq. (15)
  - Set $\boldsymbol{\alpha} = \boldsymbol{\alpha}^{(t)}$

Theorem 3.1. *For $\lambda_1 > 0$, the sequence $\{\boldsymbol{\alpha}^{(t)}\}$ produced by Algorithm 1 satisfies,*

$$F\big(\boldsymbol{\alpha}^{(t)}\big) - F\big(\boldsymbol{\alpha}^{(t+1)}\big) \geq \frac{1 + \lambda_1}{2} ||\boldsymbol{\alpha}^{(t)} - \boldsymbol{\alpha}^{(t+1)}||_2^2.$$

Theorem 3.1 shows that the objective function $F(\boldsymbol{\alpha})$ is guaranteed to decrease for all $\lambda_1 > 0$. Proof for theorem 3.1 is postponed to Appendix for sake of brevity.

In algorithm 1, a *convergence criteria* can be selected based on either the objective function, $F(\boldsymbol{\alpha})$ or iterative estimates $\boldsymbol{\alpha}^{(t)}$, $\boldsymbol{\alpha}$, i.e. for an arbitrarily small quantity ϵ, repeat until $F(\boldsymbol{\alpha}) - F(\boldsymbol{\alpha}^{(t)}) < \epsilon$, or equivalently $||\boldsymbol{\alpha}^{(t)} - \boldsymbol{\alpha}||_2^2 < 2\epsilon/(\lambda_1+1)$. An intercept α_0 can be included in the model, as is common practice, it is not penalized. Its estimate can be obtained by direct minimization of the negative log-likelihood at each step. An intercept is interpreted as an overall average spatial effect for all L locations.

The estimated spatial effects with no penalty (un-penalized) and the ridge penalty, which are used as baselines for comparison, can be obtained by following a similar algorithm, only change being $\lambda_1 = \lambda_2 = 0$ and $\lambda_2 = 0$ respectively. Therefore, if $\boldsymbol{\alpha}_0^{(t)}$ and $\boldsymbol{\alpha}_r^{(t)}$ denote their respective solutions, we have

$$\boldsymbol{\alpha}_0^{(t)} = \boldsymbol{\alpha} - \big[\mathrm{I}_L + \nabla_2(\boldsymbol{\alpha})\big]^{-1}\nabla_1(\boldsymbol{\alpha}), \quad \boldsymbol{\alpha}_r^{(t)} = \big[(\lambda_1+1)\mathrm{I}_L + \nabla_2(\boldsymbol{\alpha})\big]^{-1}\big\{(\mathrm{I}_L + \nabla_2(\boldsymbol{\alpha}))\boldsymbol{\alpha} - \nabla_1(\boldsymbol{\alpha})\big\}.$$

When considering estimation of spatial effects for S states together using algorithm 1, the increased dimension of W affects computational complexity adversely. One can alternatively suggest running the algorithm in parallel for individual states, resulting in $2S$ tuning parameters, which is undesirable. In that regard, advantage of the suggested approximation to Laplacian matrices discussed in section 3.1 is immediately evident when applied to solution in eq. (15). It results in,

$$\boldsymbol{\alpha}^{(t)} \approx \big[(\lambda_1+1)\mathrm{I}_L + \lambda_2 W_a + \nabla_2(\boldsymbol{\alpha})\big]^{-1}\big\{(\mathrm{I}_L + \nabla_2(\boldsymbol{\alpha}))\boldsymbol{\alpha} - \nabla_1(\boldsymbol{\alpha})\big\}, \tag{17}$$

where W_a is block diagonal matrix consisting of S blocks, W_k, $k = 1, 2, \ldots, S$. Each block W_k is the *exact* Laplacian for state k and $L = \sum_k L_k$, L_k being the number of locations in state k. Therefore, matrix inverse in eq. (17) can be computed in $O\big(\sum_k L_k^3\big)$ operations (instead of $O(L^3)$), affecting scalability of algorithm 1 significantly by allowing sufficient scope for parallelization while keeping the number of tuning parameters fixed. It is important to note here that, theorem 3.1 still holds for approximate Laplacian, since W_a is still positive semi-definite (p.s.d.) (see Appendix for proof).

4. Simulation

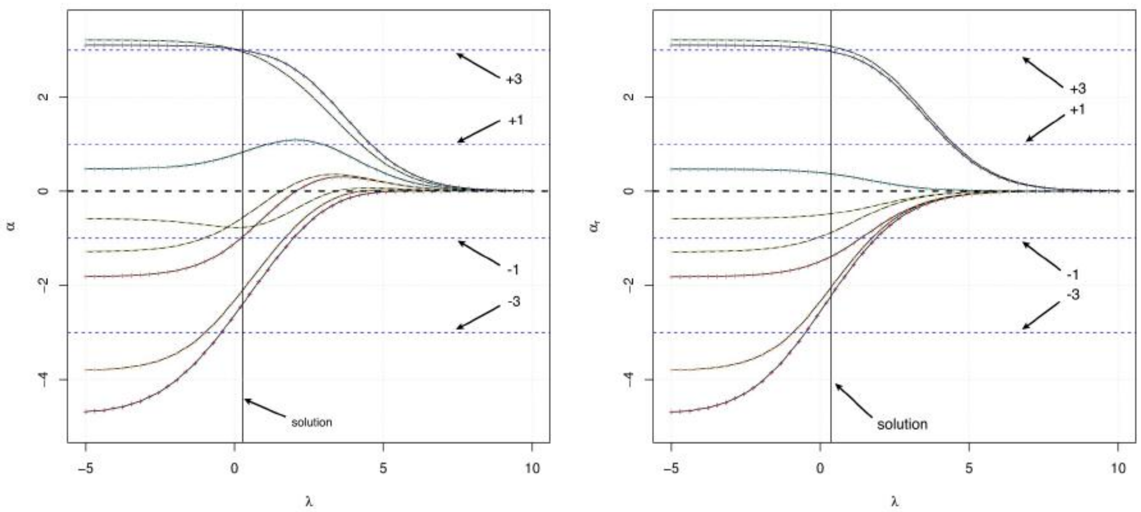
The aims of presented simulation study are,

(1) to assess performance of algorithm 1 under different spatial patterns explained and demonstrated in the ensuing discussion,
(2) to demonstrate and compare performance of algorithm 1
 (i) using exact solution in eq. (15) and,
 (ii) using approximate solution in eq. (17)
 for multiple states.
(3) a sensitivity analysis for estimated spatial effects to the index parameter, p.

In what follows it is important to note that, we use a state (or group of states) only as an example, results shown are in no way indicative of true responses in the region. Their sole purpose is to create an instance that can serve as a test case for the algorithm. We start by describing some error metrics that we will be using for all of the examples listed in this section. Let $\boldsymbol{\alpha}_{(O)}$, $\widehat{\boldsymbol{\alpha}}_0$, $\widehat{\boldsymbol{\alpha}}_r$ and $\widehat{\boldsymbol{\alpha}}$ denote true, un-penalized, ridge and the estimated spatial effect vectors from the algorithm 1 respectively. The response is simulated from a compound Poisson distribution, using likelihood approximations in [13] and [14], made available to use in R-package `tweedie`. We assume a logarithmic link for both GLMs. The index parameter $p = 1.5$ is kept fixed, while the dispersion, ϕ and mean μ are allowed to vary across simulated response. For the entirety of this section simulated data will be split into training and validation sets, with tuning parameters λ_1, λ_2 being estimated using a five-fold leave-one-out cross-validation (LOOC) on the training data by minimizing deviance

$$d(y; \widehat{\mu}, \boldsymbol{\alpha}) = \sum_{i=1}^{L} \sum_{j \,\in\, \mathcal{N}(i)} \frac{2}{\widehat{\phi}_{ij}} \left[y_{ij} e^{-(\widehat{\mu}_{ij} + r_{ij}^\top \boldsymbol{\alpha})/2} + e^{(\widehat{\mu}_{ij} + r_{ij}^\top \boldsymbol{\alpha})/2} \right],$$

over a holdout set in each fold. As a measure of loss, we use error sum of squares (SSE) for estimated spatial effects, whereas prediction error over both training and validation set is measured using a ratio of deviances given by,



(A) (B)

FIGURE 3. Figure showing solution path with λ in logarithmic scale for (a) proposed solutions ($\boldsymbol{\alpha}$) and, (b) ridge solutions ($\boldsymbol{\alpha}_r$). Dashed blue horizontal lines indicate true values, vertical line indicates value of λ at which minimum deviance was obtained.

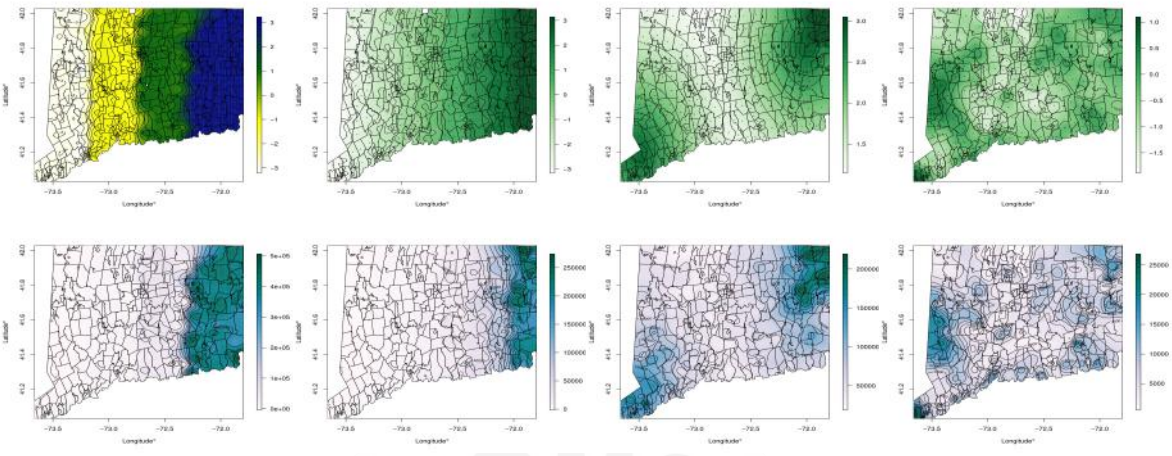
$$SSE = \big|\big|\boldsymbol{\alpha}_{(O)} - \boldsymbol{\alpha}\big|\big|_2^2, \qquad d_r\big(y; \widehat{\mu}, \boldsymbol{\alpha}_{(O)}, \boldsymbol{\alpha}\big) = \frac{d(y, \widehat{\mu}, \boldsymbol{\alpha})}{d\big(y, \widehat{\mu}, \boldsymbol{\alpha}_{(O)}\big)}, \tag{18}$$

respectively, where $\boldsymbol{\alpha}$ can be any one of three estimates, $\widehat{\boldsymbol{\alpha}}_0$, $\widehat{\boldsymbol{\alpha}}_r$ or $\widehat{\boldsymbol{\alpha}}$. Lower SSE and, $d_r(y; \widehat{\mu}, \boldsymbol{\alpha}_{(O)}, \boldsymbol{\alpha})$ closer to 1 are desirable.

Ridge estimates, ($\boldsymbol{\alpha}_r$) are a natural baseline for the proposed estimates. To assess the difference between them it is not enough to just show solutions obtained for a single λ or λ_1, λ_2. To make ease for comparison we use a variant of the penalty shown in eq. (12), $P(\boldsymbol{\alpha}; \lambda) = \lambda(0.4||\boldsymbol{\alpha}||_2 + (1 - 0.4)||\boldsymbol{\alpha}||_W)$. We vary λ on a logarithmic scale in the range $[\lambda_l, \lambda_u]$, shown in figure 3. We choose a small sample size, county level spatial effects for state of CT (8 counties). True values assigned viz., {-3,-1,1,3} are indicated in the figure. For large values of λ, the problem is un-penalized, i.e. $\boldsymbol{\alpha} = \boldsymbol{\alpha}_r = 0$. Starting from a value of 0, estimates under both penalties are obtained by using a "warm-start" strategy, that involves using the estimate, $\boldsymbol{\alpha}^{(t)}$ as a starting value for next iteration $\boldsymbol{\alpha}^{(t+1)}$. This warm-start strategy will be used when working with penalty, $P(\boldsymbol{\alpha}; \lambda_1, \lambda_2)$ in eq. (12). In that scenario, $\lambda_1, \lambda_2 \in [\lambda_{1l}, \lambda_{1u}] \times [\lambda_{2l}, \lambda_{2u}] \subset \mathbb{R}^2$. Differences in solution paths are apparent from fig. 3. The value of λ for which the path of an estimate first hits the true value , i.e. the "hitting-time" occurs much earlier in the proposed estimates. Furthermore, the value of λ at which the solution is obtained under the two penalties, has estimates closer to true values in the case of proposed algorithm indicating lower SSEs between estimated and true spatial effects.

We evaluate the performance of algorithm 1 under four different sample sizes and proportions of zeros in simulated response. Chosen sample sizes for this simulation study vary from 10,000 – 50,000, while proportions of zeros vary in the range 0.15–0.80. The estimated spatial effects are evaluated under each combination of settings to provide a detailed demonstration about efficacy of the proposed algorithm. Under each setting 100 replications are carried out, reported values of error metrics in tables shown are averages over all replications, accompanied by their respective standard deviations. We use four different spatial patterns as examples viz., *block*, *smooth*, *hot-spot* and *structured*; their construction and features will be explained in the following subsection.

For the first part of this simulation we use the state of Connecticut as an instance, second part uses three states as instances viz., Connecticut, Massachusetts and Rhode Island (reason behind such a choice being their adjacent locations). There are 282, 537 and 77 zip codes in the respective state(s). Tables showing relevant results and details are postponed to Appendix B for maintaining continuity.



(A) (B) (C) (D)

FIGURE 4. Spatial plots showing simulated response (bottom row) for (a) block (b) smooth (c) hot-spot and (d) structured spatial patterns (top row), with legends alongside plots showing scales for true spatial effects and resulting simulated response.

4.1. **Examples of different spatial patterns.** In this subsection we describe a comparative study between un-penalized and penalized estimates discussed in section 3.2 under different spatial patterns. In the ensuing examples, construction for each pattern is discussed in detail, followed by comments explaining results at the end.

Under a logarithmic link and index parameter $p = 1.5$, referring to table 2, relationship between the canonical parameter θ and mean μ for compound Poisson is given by $\theta = 2\mu^{-1/2} \Rightarrow \mu = 4/\theta^2$. In all examples shown below, we simulate, $\theta \sim N(-0.16, 0.02^2) \Rightarrow \mu \in (83, 400)$ approx., with $N(\cdot,\cdot)$ denoting a Gaussian distribution. To introduce larger number of zeros in simulated response under each settings described below we use the equivalent relationship between (μ, ϕ, p) and (ξ, η, ζ) for characterizations in eqs. (3) and (4). We have,

$$\mu = \xi\eta\zeta, \qquad \phi = \frac{\xi^{1-p}\cdot(\eta\zeta)^{2-p}}{2-p}.$$

Hence for a fixed mean and index parameter μ, p with $p \in (1, 2)$, increasing dispersion ϕ, would result in increased mean, $\eta\zeta$ of individual gamma components and a decreased rate, ξ for Poisson. A smaller rate for Poisson means higher probability of obtaining exact zeros which is our requirement. Table 8 shows respective reference distributions used for sampling dispersion offsets under each setting to obtain desired proportions of zeros in simulated response with, $U(\cdot,\cdot)$ denoting a uniform distribution.

Example 1 & 2: Block and Smooth

To obtain a block pattern with four blocks, the longitudinal range for a state is divided into four regions. For example, figure 4 column (a), shows the boundaries for four regions. Each region is assigned a fixed spatial effect, in this case $\{-3, -1, 1, 3\}$. Under a logarithmic link it evaluates to a multiplicative effect towards the mean of magnitudes, $\{e^{-3}, e^{-1}, e^{1}, e^{3}\}$. As a result, means for regions with lower magnitudes of spatial effects are expected to have higher number of zeros. Increasing the overall dispersion ϕ introduces more zeros in the simulated response for individual regions, as can be seen from lower plot for column (a) in fig. (4). Finally, for each setting the response is simulated from $y \sim \text{Tweedie}(\mu, \phi, p)$. Figure (4) column (a) shows one such instance. For each replication within a setting, 10×10 grids were chosen for tuning (λ_1, λ_2) in algorithm 1, where $(\lambda_1, \lambda_2) \in [-5, 0] \times [-3, 2]$ in *log-scale*, similarly for the ridge penalty a line search is conducted on 10×1 vector for $\lambda_1 \in [-5, 0]$ in *log-scale*.

Smooth pattern is designed to be a more general version of block patterns having finer divisions for the spatial effect thereby, producing lesser discreteness across their values. Column (b) in figure 4 shows spatial effects smoothly varying in the range $[-3, 3]$ over 10 regions and the resulting simulated response. Grid and line searches for optimizing tuning parameters for ridge and algorithm 1 were conducted over 10×1 line and 10×10 grid, where $\lambda_1 \in [-5, 2]$

and, $(\lambda_1, \lambda_2) \in [-5, 2] \times [-5, 5]$ in *log-scale* respectively. Table 10 shows related results in detail.

Example 3: Hot-spots

A hot spot is defined as location(s) that exhibit higher magnitudes of response, with response magnitudes tapering off with increasing distance from the hot-spot (for further details on hot spots and their detection see [5]). An example is shown in fig. 4 column (c) top plot. For the scope of this simulation we create two hot-spots in Connecticut viz. north-east and south-west corners, being assigned a spatial effect of 3, tapering off with increasing distance (euclidean), $l^{(2)}$ from them via an exponential kernel (i.e. $\exp(-\phi l^{(2)})$). Hence zip codes located in the center (equidistant from both hot spots) are expected to have lower spatial effects. Therefore, simulated response varies accordingly showing higher magnitudes in two hot spots. Associated grids and lines remain same as the "smooth" pattern. Table 11 shows relevant results for this pattern settings.

Example 4: Structured

A structured pattern is created using a distance based covariance kernel to simulate fixed spatial effects from a zero mean multivariate Gaussian. Explicitly, $\boldsymbol{\alpha}_{(O)} \sim N_L(\mathbf{0}, \Sigma)$, where $N_L(\cdot, \cdot)$, $\mathbf{0}$ are the L-dimensional multivariate Gaussian and zero vector respectively, with $\Sigma = \sigma^2 \exp(-\phi l^{(2)})$ being specified using an exponentially structured covariance kernel for the scope of this simulation, with $\sigma^2, \phi = 1$. $l^{(2)}$ is an $L \times L$ euclidean distance matrix with all operations being entry-wise. Figure (4), column (d) (upper plot) shows an example of structured effect and its resulting simulated response. We simulate the spatial fixed effect *once*, to be used across all different combinations of settings for sample sizes and proportion of zeros. Associated grids and lines remain same as the "smooth" and "hot-spot" patterns. Relevant results for different settings in this pattern are shown in table 12.

Comments and Interpretation:

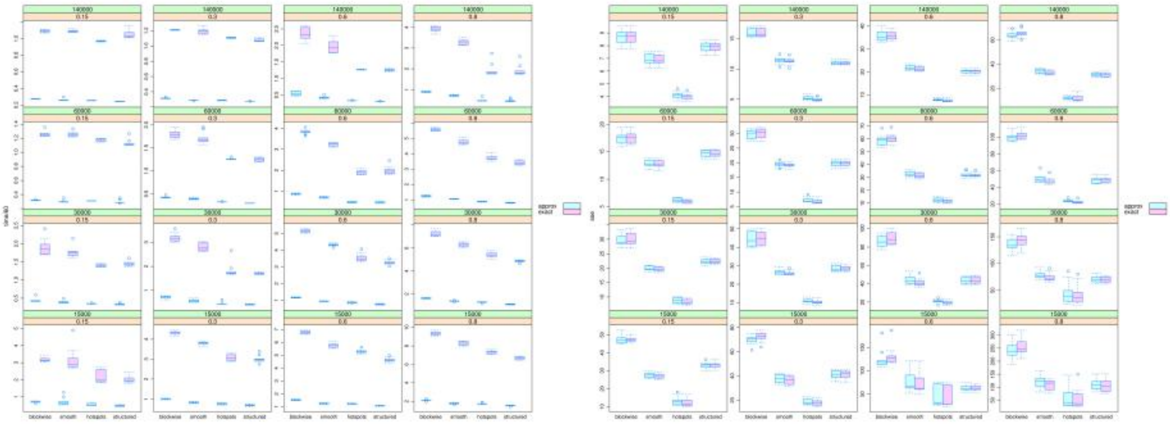
The results in table 8 are to be considered in conjunction with those in tables (9, 10, 11 and 12). With baselines for comparison being ridge and un-penalized estimates, under all different spatial patterns, results *on an average* can be summarized as follows,

(a) estimates from proposed algorithm have comparatively lower SSEs,
(b) if proportion of zeros in simulated response is assumed to be a metric for signal to noise ratio then proposed estimates also show lower SSEs under very low signal to noise ratios (i.e. higher proportion of zeros in response),
(c) under low sample sizes and signal to noise ratios, proposed estimates show low SSEs in comparison to estimates from the ridge penalty,
(d) training and validation deviance ratios, $d_r\left(y; \widehat{\mu}, \boldsymbol{\alpha}_{(O)}, \widehat{\boldsymbol{\alpha}}\right)$ are closer to 1 when compared to corresponding un-penalized and ridge versions.

With regard to above conclusions, flexibility of proposed algorithm over ridge penalty is demonstrated in table 8 from values of regularization parameters obtained under each spatial pattern for the two penalties. Having a function of Laplacian as an additional penalty term primarily allows for superiority in performance under the presence of spatial variation in simulated response. This is evident particularly in simulation results from a hot-spot pattern shown in table (8). Regularization parameters for both ridge and ridge part of proposed penalty show the same values (i.e. no penalty, $\lambda_1 = -5.000$), accompanied by a large value of $\lambda_2 = 2.708$ for Laplacian part of the penalty; on referring to corresponding table 11 we note that estimates from algorithm 1 are superior to ridge estimates.

4.2. **Comparison between algorithms featuring approximate and exact solutions.** This subsection primarily compares performance of two estimates, exact and approximate derived in eqs. (15) and (17) respectively. The metric used to compare them is a trade-off between error sum of squares and time taken to convergence under the same spatial patterns considered in section 4.1. Aim being to demonstrate that proposed approximation in eq. (17) produces estimates with error sum of squares comparable to that of exact estimates in eq. (15), but with significant improvement in time taken. Figure 5 summarizes results for the comparative study.

We considered three adjacent states in the southern New England group for this simulation viz., Connecticut, Massachusetts and Rhode Island. Sample sizes chosen are 15,000, 30,000,



(A) (B)

FIGURE 5. Figures showing (a) time taken in minutes (b) SSE as defined in eq. (18), from 10 replications under each possible combination of sample size and proportion of zeros in simulated response for different spatial patterns. Each figure shows grouped box-plots for approximate and exact estimates respectively under different settings.

60,000 and 140,000, with proportions of zeros in the simulated response varying in 0.15, 0.30, 0.60 and 0.80. The *approximated* adjacency used, was a block diagonal matrix with blocks of the order, 282×282, 537×537 and 77×77 respectively for individual states, as opposed to an exact adjacency matrix of the order 896×896. Associated tuning parameters were estimated over a 10×10 grid, $[-5, 5] \times [-5, 5]$ in logarithm scale, for all different spatial patterns.

As can be readily seen from fig. 5b, approximate estimates show performance comparable to exact counterparts in terms of SSE; furthermore in-sample and out-sample deviance ratios as defined in eq. (18) showed similar findings. Referring to fig. 5a the significant advantage in computational complexity that approximate estimates demonstrate makes it a worthy alternative when considering scalability of algorithm 1. In particular when the synthetic spatial patterns were segmented, for ex. "blockwise" patterns, referring to fig. 5b we can see that approximate estimates prevent unnecessary smoothing, thereby outperforming their exact counterparts (for more details regarding computation, see sec. 7).

4.3. **Sensitivity to choice of index parameter, p.** We elaborate on relevant consequences of properties discussed at the end of section 2.2 (particularly in eq. (9)), by looking further into sensitivity of estimates to the choice of index parameters. Working with offsets from a DGLM for the Tweedie family, where the index parameter, p distinguishes between models, it is desirable and expected for our proposed approach to remain unaffected by erroneous model specification.

Explicitly we consider situations where true value of index parameter, $p \neq 1.5$, however a value of 1.5 was reported (or assumed since missing) for the index parameter. Under orthogonality of mean, μ to index and dispersion parameters, p, ϕ respectively for Tweedie models we expect spatial information in the mean to remain relatively unperturbed. Metrics used for this simulation are (i) SSE and, (ii) L^2-norm (Euclidean) for true against estimated effects. Figure 8 in Appendix C, shows results for the simulation study conducted. Previous simulation shows exact and approximate estimates demonstrating comparable results in terms of SSE, hence conclusions derived for exact can be extended to hold for approximate estimates. As a result, we limit the scope of this simulation to estimating exact spatial effects for state of CT under previously described settings for the index parameter.

Additional simulation settings used consist of the same spatial patterns as described in examples from section 4.1, with sample sizes varying in 10,000, 20,000, 30,000 and 50,000, index parameters varying in 1.3, 1.4, . . . , 1.9 and dispersion parameters are altered accordingly to result in simulated response having 30% zeros under all combinations. Associated tuning parameters were estimated over a 10×10 grid, $[-5, 5] \times [-5, 5]$ in logarithm scale.

Figure 8a shows that using $p = 1.5$ as a reference, significant departures are located from true effects for extreme values of p (ex. $p = 1.9$), which become relatively pronounced under low sample sizes and coarser spatial patterns (ex. block-wise). In fig. 8b norms for estimated and true spatial effects display a similar behavior under all patterns considered. This leads us to conclusively state that under large sample sizes and relatively smooth spatial variations in response, estimates produced by algorithm 1 remain stable w.r.t. model mis-specification.

5. Case Studies

Case studies considered aim to demonstrate the application and performance of proposed algorithm with both, exact and approximate estimates in detecting residual spatial effect while modeling response, i.e. *loss costs per unit insured* related to personal automobile insurance collision claims. Exact estimates using algorithm 1 are obtained for state of CT. While, for approximate estimates we consider two case studies featuring, (i) six states in New England (group of states in the east coast) consisting lower number of zip codes compared to, (ii) three adjacent states in West Coast, having larger number of zip codes, within the United States of America (USA). All of these are subsets of data obtained from a more comprehensive repository named Highway Loss Data Institute (HLDI) maintained by the independent non-profit, Insurance Institute for Highway Safety (IIHS) [47] working towards reducing fatalities arising from motor vehicle crashes. We shall refer to this as the HLDI database.

We briefly describe the HLDI database. It contains data at an individual level. The data contains covariates associated with the individual on,

- accident and model year of the vehicle, ranging from 2000–2015 and 1981 – 2016 respectively,
- risk of the policy having two levels "S", "N",
- age, gender, marital status and gender of partner, where missing values are denoted by 0 and "U" respectively for age and rest of these predictors,
- number of claims, payments (i.e. loss cost), exposure (measured in policy years, eg. 0.5 indicates individual insured for half a year) and deductible limit (categorical with 8 categories).
- 5-digit zip code indicating location, i.e. areally-referenced.

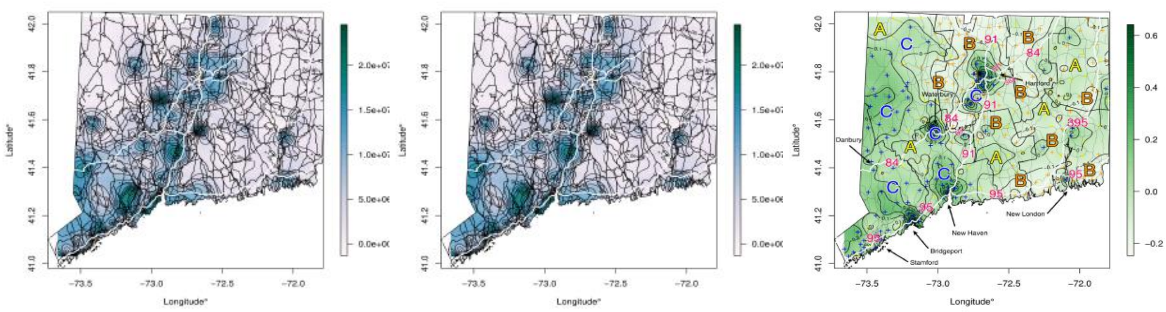
Derived predictors like age categories, vehicle age (accident–model year) in years can be obtained and used in the DGLM.

For all case studies we use implementations and approximations for compound Poisson DGLMs as in [23]. Policy exposures (w_{ij}) are used as weights for associated dispersion parameter ϕ_{ij} in the DGLM, i.e. as $\phi^*_{ij} = \phi_{ij}/w_{ij}$. Response variable used is defined as the ratio of payments to exposure. Fitted mean model consists of deductible, accident year, gender and marital status, while the dispersion model consists of a categorical version of age (consisting of 6 categories) in addition to predictors in the mean model. Index parameter used is $p = 1.6$. We use all of the data to obtain fitted mean and dispersions for the DGLM, which we aim to use as offsets in the process of estimating zip code-level spatial effects. Following which for each case study shown, we then randomly divide the data into training and validation subsets using a 60-40 *stratified* split, with stratification at a zip code level for state(s). Proposed algorithm along with ridge and un-penalized counterparts are fitted on the training sets. We replicate this procedure 20, 10 and 5 times for exact and approximate (two case studies) procedures respectively. While evaluating models we consider predictions in terms of loss costs or payments for all fitted models, after adjusting for exposure. This is done to avoid issues involving misalignment in out-sample predictions.

It is important to note that in our proposed model, a spatial fixed effect for zip-codes is present in the mean model. Consequently, we expect to see substantial improvement in zip-code level aggregated loss costs rather than at individual level. As error metrics we choose deviance at the observation level (as in eqs. 6a, 6b) and mean square error (between observed and predicted loss costs), $MSE = \sum_{i=1}^{L} \left(\sum_{j=1}^{n_i} w_{ij}(y_{ij} - \widehat{y}_{ij}) \right)^2 / L$, at the aggregated level for model comparison. Regarding estimated spatial effects in each replication, we compute associated means, standard deviations and 95% quantiles for each zip-code.

5.1. **Connecticut.**

In HLDI data, the state of Connecticut contains 22,337,318 records on 282 zip codes, with 96.18% of them being exact zeros. Number of claims observed range from 0–8 within given exposure periods. The range of observed losses were 0–190,487 (in US



(A) (B) (C)

FIGURE 6. Spatial plots showing *average* zip-code level out-sample (a) observed losses (b) predicted losses adjusted for spatial effects, (c) estimated spatial effects from algorithm 1 for the state of CT. (*Note:* Plot (a), (b) and (c) shows primary interstate (I- 84, 384, 91, 95, 691, 291, 95, 395), secondary state highways and major cities in CT with bold ("white"), usual lines and arrows respectively. Also in plot (c), across replications, zip codes whose approximate 95% CIs contain zero, both limits are negative and positive are color coded in "yellow" (A), "orange" (B) and "blue" (C) respectively.)

dollars) with mean loss of 167.202, estimated losses, $\widehat{\mu}$ from the DGLM were, 0.002–4023.710 with a mean of 158.337. Estimated dispersions, $\widehat{\phi}$ were 267.045–1038.674.

Estimates from the proposed algorithm 1, along with their ridge and un-penalized counterparts were obtained for 20 replications involving random training and validation splits. Out-sample prediction results for deviance and MSE are shown in table 3. In both observation level and aggregated cases DGLM serves as the baseline. As expected, the improvement at aggregated level is much larger in comparison to the observation level, with an average of 0.33% and 93.10% improvements for predictions adjusted with spatial effects from proposed algorithm respectively. Figure 6a, 6b show resulting zip-code level aggregated observed and predicted loss costs after adjustment with estimated spatial effect from proposed algorithm.

Associated tuning parameters were estimated over a 50×50 grid constructed within $[-5, 5] \times [-5, 5] \in \mathbb{R}^2$ in a logarithmic scale for all replications. Estimated tuning parameters λ_1, λ_2, with standard deviations shown in brackets followed by an approximate 95% interval, for algorithm 1 are (i) $\widehat{\lambda}_1 = 1.37$ (2.39), (0.01, 6.95), (ii) $\widehat{\lambda}_2 = 19.44$ (2.58), (15.72, 23.65). For the ridge variant we have, $\widehat{\lambda}_1 = 37.25$ (4.04), (32.12, 43.62). The overall mean estimated spatial effect across all zip codes and replications was 0.033 (0.167), i.e. a multiplicative effect of $e^{0.033} \approx 1.034$ to loss costs in the state on an average. Figure 6c shows a spatial plot of estimated spatial effects using algorithm 1.

Across replications an approx. 95% confidence interval (CI) is calculated for the spatial effect in each zip-code. Based on them containing zero and algebraic sign of their limits indicating whether the multiplicative effect is identity (if interval contains 0), decreasing (if both limits are negative) and increasing (both limits are positive) these effects are color

TABLE 3. Table showing averaged out-sample results for CT, consisting of deviance and MSE at the observation and zip-code aggregated level (lower is better), accompanied by mean percentage improvements obtained with DGLM as a baseline for un-penalized, ridge and proposed estimates (GL). Standard deviations, first (Q_1) and third (Q_3) for %ge improvements are shown in following rows.

Percentage Improvements	Observation Level – Deviance ($\times 10^5$)				Aggregated Level – MSE ($\times 10^{10}$)			
	DGLM	GL	MLE	Ridge	DGLM	GL	MLE	Ridge
Values	3.743265	3.730861	3.730994	3.730907	122.371867	8.437873	8.675707	8.541663
Mean (%)	–	0.331350	0.327798	0.330115	–	93.100513	92.904780	93.016643
Std. Dev	–	(0.008448)	(0.009071)	(0.008366)	–	(0.764774)	(0.820851)	(0.773495)
Q_1	–	0.319138	0.314905	0.318801	–	91.514987	91.318403	91.554967
Q_3	–	0.346402	0.344539	0.345565	–	94.144875	93.966229	94.066700

coded into three categories which are shown in the spatial plot 6c. For perspective, primary interstate highways and secondary roads passing through and within the state are shown as well. From the southwest corner until the intersection of primary interstates i.e. I-91 and 84 at Hartford, regions marked with "C" having a higher multiplicative spatial effect, in comparison to northwest or the eastern part of CT, which are coded with "A" or "B" indicating the multiplicative factor is 1 or < 1 respectively. Obtaining ordinal spatial risk ranking is rather straightforward, making this a demonstration of boundary analysis in CT, where primary interstate highways, their intersections and big cities serve as defining boundaries for spatial risk.

In comparison to ridge and un-penalized estimates, our proposed estimates show lower standard deviation (average of 7.22% and 15.62% respectively) for zip codes that did not contain 0 in their approx. 95% CI. The number of zip codes with 0 belonging to their approx. CIs, less and more than 0 changed from $\{86, 94, 102\}$ in un-penalized and ridge, to $\{67, 104, 111\}$ in proposed estimates.

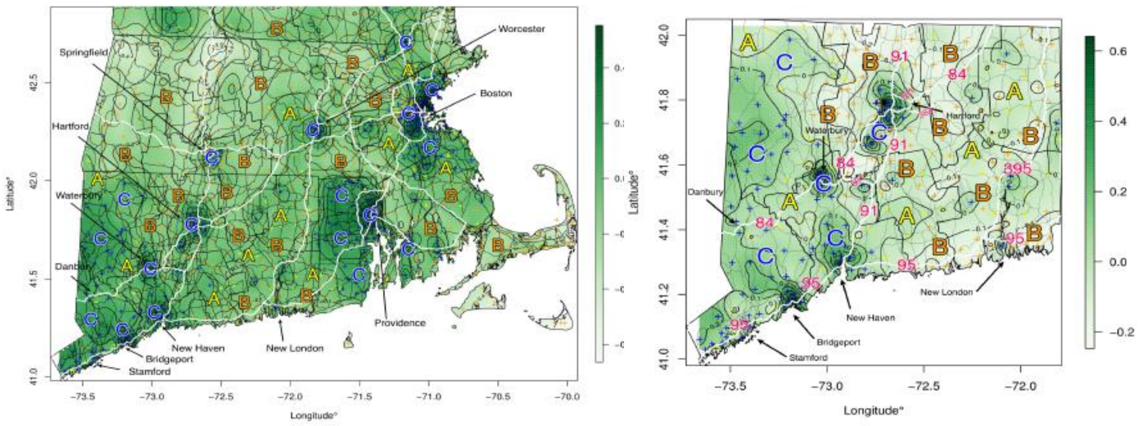
5.2. **New England.** New England is a group of six states in the east coast of USA viz., Connecticut, Rhode Island, Massachusetts, Vermont, New Hampshire and Maine. HLDI data for New England consists of 72,118,625 records on 1831 zip codes. The percentage of exact zeros in the response is 95.27%. Losses (payments) ranged between 0–427,510 with an average loss of 184.23 (in US dollars). Estimated mean, $\widehat{\mu}$ from fitted DGLM ranged from 0.02–26,791.86, with an average of 184.37 (in US dollars) while, estimated dispersions $\widehat{\phi}$ were in the range, 267.04–1038.67.

Spatial effects are estimated using solution in eq. (17) with algorithm 1. We have already observed in simulations comparing exact and approximates solutions, depending on smoothness of unobserved spatial effects performance of the approximate solution is comparable to its exact counterpart with substantial improvement in time complexity. This significantly improves scalability in case studies like the present one with multiple states involved. Tuning parameters are estimated over a 30×30 grid within $[-5, 5] \times [-5, 5]$ in logarithmic scale. Average estimated tuning parameters from 10 replications are shown in table 4 accompanied by percentage of improvement in deviance and MSE at observation and zip-code level aggregation respectively. An improvement of 0.488% and 96.677% is seen at the observation and aggregated levels respectively. Average overall spatial effect estimated across all zip codes is -0.077 (i.e. a multiplicative effect of $e^{-0.077} = 0.926$), having a standard deviation of 0.167. Figure 9, in Appendix C shows a spatially interpolated surface derived from estimated spatial effects. Letters "A", "B" and "C" have the same interpretation as in earlier case study. On closer inspection of the figure following are readily apparent,

- regions in northern New England that are devoid of roads (consequently habitation) show sizable areas with negative spatial effects (marked by"B") that indicate a decreasing multiplicative effect,
- regions in southern New England show more frequently appearing regions with positive spatial effects (marked by"C") that indicate a increasing multiplicative effect.
- regions marked in "A", i.e. neutral spatial regions act as *separating boundaries* between regions marked by "B" and "C".

Table 4. Table showing averaged out-sample results for New England, consisting of deviance and MSE at the observation and zip-code aggregated level (lower is better), accompanied by mean percentage improvements obtained with DGLM as a baseline, compared to proposed *approximate* estimates (GL). Standard deviations, first (Q_1) and third (Q_3) quartile for %ge improvements are shown in following rows.

Percentage Improvements	Observation Level – Deviance ($\times 10^6$)		Aggregated Level – MSE ($\times 10^{10}$)		Tuning Parameters	
	DGLM	GL	DGLM	GL	$\widehat{\lambda}_1$	$\widehat{\lambda}_2$
Values	2.965573	2.951098	93.567022	3.109582		
Mean (%)	–	0.488101	–	96.676626	5.302584	26.466031
Std. Dev	–	(0.004236)	–	(0.667377)		
Q_1	–	0.483620	–	95.233090	(0.938693)	(0.000000)
Q_3	–	0.494925	–	97.029910		



(A) (B)

FIGURE 7. Plots comparing estimated spatial effects obtained from the approximate algorithm for (a) states in southern New England (Connecticut, Rhode Island and Massachusetts) as an inset of fig. 9 and, exact estimates for (b) Connecticut, to demonstrate smoothness in estimated effects across state borders. Symbols and lines have the same description as in fig. 6.

TABLE 5. Table showing summarized aggregated (at city-level) spatial effects for cities shown in fig. 9. Cities having one zip code under them have no estimate of std. deviation, or quartiles (Q_1, Q_3)

Estimated Spatial Effects	State							
	CT		RI	MA		NH	VT	ME
	Hartford	New Haven	Providence	Boston	Worcester	Concord	Montpelier	Augusta
Mean	0.022375	0.068355	0.376642	0.242398	0.198258	-0.137677	-0.110704	-0.055297
Std. Dev	(0.090117)	(0.160785)	(0.241224)	(0.272502)	(0.192553)	(0.037967)	–	–
Q_1	-0.111165	-0.138137	-0.004110	-0.099666	-0.132779	-0.163182	–	–
Q_3	0.121738	0.203204	0.604134	0.712228	0.400212	-0.112172	–	–

- Big cities and intersections of primary interstate highways (ex. I-95, 91, 84, 290 etc.) show positive spatial effects with large magnitudes. For instance, Suffolk (where state capital, Boston is a city) and Norfolk, MA are adjacent counties having spatial effect averaged (std. deviations) at the county level of 0.275 (0.280) and 0.015 (0.123) respectively. They are the only counties in MA having positive aggregated effects. They contain intersections of major interstate highways (viz., I-95, 90 and 93) and cities like Boston and Quincy. Boston and Quincy had aggregated (at city-level) effects (with std. deviations) of 0.242 (0.272) and 0.139 (0.096) respectively.

Table 5 shows summaries for estimated spatial effects aggregated to the city level which support above findings.

The only remaining concern is manifestation of unusual discreteness (ex. large differences in magnitude, or algebraic signs) in estimated spatial effects across state borders. Consequently we compared zip codes lying on the boundary for exact estimated effects in CT, and subset of CT from the approx. estimates, with respect to those in MA and RI. There are {17, 22} zip codes on the CT-MA border and, {9, 10} zip codes in CT-RI border respectively. Presence of Providence and, RI being a state with comparatively smaller area affects magnitudes of estimates on the CT-RI border (means of -0.057 and 0.091 respectively). However, this is not the case for CT-MA border, with Springfield being the only major city, estimates show differences of lower magnitude (means of -0.096 and -0.154 respectively). Figure 7 illustrates this observation visually.

5.3. **West Coast.** Three states namely, California (CA), Oregon (OR) and Washington (WA) compose the west coast. In HLDI, west coast consists of 235,329,963 records on 2775 zip codes. This test case is considered as a demonstration of performance for proposed estimates under large number of zip codes and sample size. The response consisted of 96.27% zeros. Losses range from 0–507,631 (in US dollars). The exposure in policy years range from 0.003 – 14.825 with a mean of 0.612. Mean estimates, $\widehat{\mu}$ from fitted DGLM range from 0.054 – 5,194.105, with estimated dispersion $\widehat{\phi}$ in the range 247.229 – 1,294.303. The mean loss (payment) is 165.980, with an estimated mean loss of 165.721 US dollars from fitted DGLM.

We considered 5 replications, with tuning parameters estimated over a 30×30 grid on $[-5, 5] \times [-5, 5]$ in log-scale. Results obtained are shown in table 6. Average (std. deviation) out-sample improvements of 0.513% (0.004) and 95.557% (0.245) at the observation and aggregated zip-code level were observed respectively over 5 replications. Estimated spatial effects had a mean (std. deviation) of -0.057 (0.222) ($\equiv$ multiplicative effect of $e^{-0.057} \approx 0.945$ to observed losses). Predicted losses (payments) after being adjusted with spatial effects ranged from 0.003–8,806.819 (in US dollars), with an estimated mean of 166.869 US dollars.

Regarding boundary analysis, referring to figure 10 we observe big cities and major intersections of primary interstates show positive spatial effects, for instance the cities shown in table 7. However positive spatial effects (marked with "C") occur more frequently across replication in California along interstate I-5, however Oregon and Washington show spatial effects that are negative (marked by "B") or including zero in ther approx. 95% CIs. Major cities like San Francisco, Los Angeles also form wombling boundaries i.e. show rapid changes in estimated spatial effects.

6. Conclusion

Motivated by lack of methods for spatial estimation that scale well in terms of both processing and memory load, for data involving increasingly large number of locations, we developed exact and approximate (blockwise-exact) methods for penalized estimation of un-observable (fixed) spatial effects, while using predictions from a compound Poisson DGLM as offsets. Both exact and approximate algorithms utilize the majorization descent property which assures convergence. Using offsets paired with a zip-code level adjacency allows for significant improvement in out-sample performance when aggregated at the zip-code (areal) level over the DGLM.

We presented detailed simulations accompanied by motivating case studies to illustrate the performance and features, including scalability of our proposed algorithm under different spatial settings. Methods shown can be applied to any existing GLM implementation after incorporating necessary changes in the link and deviance. With the increasing applicability of compound Poisson models and availability of geo-tagged areal data we feel that proposed algorithms will be of use in analyzing spatial variability in both individual and sizable groups of states.

7. Discussion and future work

We have considered a joint modeling approach to the problem, where inference for the spatial effects are conducted jointly with the mean and dispersion model parameters of the DGLM [48]. However the approach presented here is more scalable owing to the approximate Laplacians considered when modeling the spatial effects over large spatial domains. All

Table 6. Table showing averaged out-sample results for West Coast, consisting of deviance and MSE at the observation and zip-code aggregated level (lower is better), accompanied by mean percentage improvements obtained with DGLM as a baseline, compared to proposed *approximate* estimates (GL). Standard deviations, first (Q_1) and third (Q_3) quartile for %ge improvements are shown in following rows.

Percentage Improvements	Observation Level – Deviance ($\times 10^6$)		Aggregated Level – MSE ($\times 10^{11}$)		Tuning Parameters	
	DGLM	GL	DGLM	GL	$\widehat{\lambda}_1$	$\widehat{\lambda}_2$
Values	7.318123	7.280574	26.64058	1.183930		
Mean (%)	–	0.513094	–	95.557069	4.719600	37.363330
Std. Dev	–	(0.003630)	–	(0.245279)		
Q_1	–	0.508669	–	95.270866	(0.000000)	(0.000000)
Q_3	–	0.517819	–	95.754971		

TABLE 7. Table showing summarized aggregated (at city-level) spatial effects for cities shown in fig. 10. Standard deviations, first and third quartiles (Q_1,Q_3) are shown for multiple zip codes under a city.

Estimated Spatial Effects	State						
	WA	OR	CA				
	Seattle	Portland	Sacramento	San Francisco	San Jose	Los Angeles	San Diego
Mean	-0.044461	-0.106216	0.156509	0.246519	0.119130	0.376755	-0.026721
Std. Dev.	(0.085933)	(0.079656)	(0.094374)	(0.122083)	(0.081754)	(0.102605)	(0.098419)
Q_1	-0.153033	-0.209162	0.014194	0.106287	0.009060	0.187888	-0.149399
Q_3	0.147396	0.039022	0.291680	0.517545	0.305338	0.569905	0.165796

computations shown here were performed on an `Intel (R) Xeon (R) Gold 6150 CPU @ 2.70GHz` with 128GB of RAM. Case studies featuring,

- the state of Connecticut took an average (std. deviation) of 0.684 (0.053) minutes for each of 20 replications,
- New England took an average (std. deviation) of 4.618 (0.095) minutes for each of 10 replications
- West Coast took an average (std. deviation) of 2.087 (0.037) hours for each of 5 replications.

To give brief details regarding scale, HLDI data for different state(s) took disk spaces of 395.52MB, 0.96 and 3.48GB respectively for the three case studies considered. All spatial plots accompanied by computation shown was done using `R` [42]. The library `MBA` [43], was used to spatially interpolate estimates and produce the wombling surface shown in plots using multilevel B-splines. Standard libraries like `sp` [44], [45] and `maptools` [46] were helpful in obtaining shape files necessary for plots.

There is scope for considerable future work. We aim to establish an efficient unified framework for estimating the DGLM parameters simultaneously with spatial effects that can incorporate variable selection methods like elastic net (for example in [21]). Occurrence of exact zeros, as we have seen is primarily controlled by the dispersion model, hence having a spatial effect in the dispersion model would be another future goal.

Acknowledgments

The authors would like to thank Brien Aronov, whose valuable input regarding practical applications, the data and necessary computational tools facilitated this research to a great extent.

References

[1] Li, Tianxi and Levina, Elizaveta and Zhu, Ji and others, (2019) Prediction models for network-linked data, The Annals of Applied Statistics 13(1) 132–164

[2] Banerjee, Sudipto and Gelfand, Alan E and Sirmans, CF, (2003) Directional rates of change under spatial process models, Journal of the American Statistical Association 98(464) 946–954, Taylor & Francis

[3] Lu, Haolan and Carlin, Bradley P, (2005) Bayesian areal wombling for geographical boundary analysis, Geographical Analysis 37(3) 265–285, Wiley Online Library

[4] Womble, William H (1951) Differential systematics. Science 114(2961) 315–322 JSTOR

[5] Cressie, Noel A.C. (1993) *Statistics for spatial data*. Wiley Interscience

[6] Wu, Tong Tong and Lange, Kenneth and others (2010) The MM alternative to EM. Statistical Science 25(4) 492–505, Institute of Mathematical Statistics

[7] Lange, Kenneth (2013) The MM algorithm. Optimization 185–219, Springer

[8] Lange, Kenneth (2016) *MM optimization algorithms*. Optimization 147, SIAM

[9] Barndorff-Nielsen, Ole and Cox, David R (1979) Edgeworth and saddle-point approximations with statistical applications. Journal of the Royal Statistical Society: Series B (Methodological) 41(3) 279–299, Wiley Online Library

[10] Smola, Alexander J and Kondor, Risi (2003) Kernels and regularization on graphs. *Learning theory and kernel machines* 144–158, Springer

[11] Nelder, John A and Pregibon, Daryl (1987) An extended quasi-likelihood function. Biometrika 74(2) 221–232, Oxford University Press

[12] Nelder, JA and Lee, Y (1992) Likelihood, quasi-likelihood and pseudo-likelihood: some comparisons. Journal of the Royal Statistical Society: Series B (Methodological) 54(1) 273–284, Wiley Online Library

[13] Dunn, Peter K and Smyth, Gordon K (2005) Series evaluation of Tweedie exponential dispersion model densities. Statistics and Computing 15(4) 267–280, Springer

[14] Dunn, Peter K and Smyth, Gordon K (2008) Evaluation of Tweedie exponential dispersion model densities by Fourier inversion. Statistics and Computing 18(1) 73–86, Springer

[15] Pregibon, Dary (1984) Review: P. McCullagh, J. A. Nelder, Generalized Linear Models. The Annals of Statistics 12(4) 1589–1596, The Institute of Mathematical Statistics

[16] Barndorff-Nielsen, Ole (2014) Review: P. McCullagh, J. A. Nelder, Generalized Linear Models. Information and exponential families: in statistical theory, John Wiley & Sons

[17] Friedman, Jerome and Hastie, Trevor and Tibshirani, Robert (2001) *The elements of statistical learning* 1 Springer series in statistics, New York

[18] Besag, Julian and York, Jeremy and Mollié, Annie (1991) Bayesian image restoration, with two applications in spatial statistics. Annals of the institute of statistical mathematics 43(1) 1–20 Springer

[19] Whittle, Peter (1954) On stationary processes in the plane. Biometrika, 434–449 JSTOR

[20] Besag, Julian (1974) Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society: Series B (Methodological), 36(2) 192–225, Wiley Online Library

[21] Qian, Wei and Yang, Yi and Zou, Hui (2016) Tweedies compound poisson model with grouped elastic net. Journal of Computational and Graphical Statistics, 25(2) 606–625, Taylor & Francis

[22] Zhang, Yanwei (2013) Likelihood-based and Bayesian methods for Tweedie compound Poisson linear mixed models. Journal of Computational and Graphical Statistics, 23(6) 743–757, Springer

[23] Smyth, Gordon K and Jørgensen, Bent (2002) Fitting Tweedie's compound Poisson model to insurance claims data: Dispersion modelling. ASTIN Bulletin: The Journal of the IAA, 32(1) 143–157, Taylor & Francis

[24] Jørgensen, Bent and Paes De Souza, Marta C (1994) Fitting Tweedie's compound Poisson model to insurance claims data. Scandinavian Actuarial Journal, 1 69–93, Cambridge University Press

[25] Yang, Yi and Qian, Wei and Zou, Hui (2018) Insurance premium prediction via gradient tree-boosted Tweedie compound Poisson models. Journal of Business & Economic Statistics, 36(3) 456–470, Taylor & Francis

[26] Dons, Klaus and Bhattarai, Sushma and Meilby, Henrik and Smith-Hall, Carsten and Panduro, Toke Emil (2016) Indirect approach for estimation of forest degradation in non-intact dry forest: modelling biomass loss with Tweedie distributions. Carbon balance and management, 11(1) 14, Nature Publishing Group

[27] Dons, Klaus and Bhattarai, Sushma and Meilby, Henrik and Smith-Hall, Carsten and Panduro, Toke Emil (2016) Indirect approach for estimation of forest degradation in non-intact dry forest: modelling biomass loss with Tweedie distributions. Carbon balance and management, 11(1) 14, Nature Publishing Group

[28] Lauderdale, Benjamin E (2012) Compound Poissongamma regression models for dollar outcomes that are sometimes zero. Political Analysis, 20(3) 387–399, Cambridge University Press

[29] Foster, Scott D and Bravington, Mark V (2013) A Poisson–gamma model for analysis of ecological non-negative continuous data. Environmental and ecological statistics, 20(4) 533–552, Springer

[30] Shono, Hiroshi (2008) Application of the Tweedie distribution to zero-catch data in CPUE analysis. Fisheries Research, 93(1-2) 154–162, Elsevier

[31] Shono, Hiroshi (2008) Application of the Tweedie distribution to zero-catch data in CPUE analysis. Fisheries Research, 93(1-2) 154–162, Elsevier

[32] Smyth, Gordon K (1996) Regression analysis of quantity data with exact zeros. Proceedings of the Second Australia-Japan Workshop on Stochastic Models in Engineering, Technology and Management, Gold Coast, Australia, 17–19, Citeseer

[33] Dunn, Peter K (2003) Precipitation occurrence and amount can be modeled simultaneously. Faculty of Sciences (Mathematics & Computing) Working Paper Series, University of Southern Queensland

[34] Dunn, Peter K and Smyth, Gordon K (2018) *Generalized linear models with examples in R*. Springer.

[35] Dunn, Peter K and Smyth, Gordon K (2001) Tweedie family densities: methods of evaluation, Proceedings of the 16th International Workshop on Statistical Modelling, Odense, Denmark 2–6. Citeseer.

[36] Tweedie, MCK (1946) The regression of the sample variance on the sample mean, Journal of the London Mathematical Society 1(1) 22–28. Wiley Online Library.

[37] Tweedie, MCK (1984) An index which distinguishes between some important exponential families, Statistics: Applications and new directions: Proc. Indian statistical institute golden Jubilee International conference 579–604.

[38] Jørgensen, Bent (1987) Exponential dispersion models, Journal of the Royal Statistical Society: Series B (Methodological) 49(2) 127–145, Wiley Online Library

[39] Nelder, John Ashworth and Wedderburn, Robert WM (1972) Generalized linear models, Journal of the Royal Statistical Society: Series B (Methodological) 135(3) 370–384, Wiley Online Library

[40] Jørgensen, Bent (1997) *The theory of dispersion models*, CRC Press

[41] McCullagh, Peter (2018) *Generalized linear models*, Routledge

[42] R Core Team (2019) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria

[43] Andrew Finley and Sudipto Banerjee and yvind Hjelle (2017) MBA: Multilevel B-Spline Approximation, R package version 0.0-9, https://CRAN.R-project.org/package=MBA

[44] Edzer J. Pebesma and Roger S. Bivand (2005) Classes and methods for spatial data in R, 5(2) 9–13, November. https://CRAN.R-project.org/doc/Rnews/

[45] Roger S. Bivand and Edzer Pebesma and Virgilio Gomez-Rubio (2013) *Applied spatial data analysis with R, Second edition*, Springer, NY

[46] Roger Bivand and Nicholas Lewin-Koh (2019) maptools: Tools for Handling Spatial Objects, R package version 0.9-5.

[47] IIHS (2019) Insurance Institute for Highway Safety, Accessed: 2019-06-21 www.iihs.org

[48] Halder, A and Mohammed, Shariq and Chen, Kun and Dey, Dipak (2021), Spatial Tweedie exponential dispersion models: an application to insurance rate-making, Scandinavian Actuarial Journal, 10, 1017–1036, Taylor & Francis.

Department of Epidemiology and Biostatistics, Drexel University Dornsife School of Public Health, Floor 5, Nesbitt Hall, 3215 Market Street, Philadelphia, Pennsylvania 19104, USA [1*] (corresponding author)

Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Ave, Boston, Massachusetts 02118, USA [2]

Department of Statistics, University of Connecticut, 215 Glenbrook Rd. U-4120, Storrs, Connecticut 06269-4120, USA [3]

Email(s): aritra.halder@drexel.edu [1*] (corresponding author), shariqm@bu.edu [2]
kun.chen@uconn.edu[3], dipak.dey@uconn.edu[3]

8. Appendix

8.1. Appendix A: Proofs.

Proof of Theorem 3.1:

$\mathcal{L}\big(\boldsymbol{\alpha}^{(t)}\big|\boldsymbol{\alpha}\big)$ in eq. (16) is the majorizing function. For any $\boldsymbol{\delta} \in \mathbb{R}^L$ we have,

$$\begin{aligned}\mathcal{L}\big(\boldsymbol{\alpha}^{(t)}+\boldsymbol{\delta}\big|\boldsymbol{\alpha}\big)-\mathcal{L}\big(\boldsymbol{\alpha}^{(t)}\big|\boldsymbol{\alpha}\big) = &\Big[\ell(\boldsymbol{\alpha})+(\boldsymbol{\alpha}^{(t)}+\boldsymbol{\delta}-\boldsymbol{\alpha})^\top\nabla_1(\boldsymbol{\alpha})+\frac{1}{2}(\boldsymbol{\alpha}^{(t)}+\boldsymbol{\delta}-\boldsymbol{\alpha})^\top\big(\mathrm{I}_L+\nabla_2(\boldsymbol{\alpha})\big)(\boldsymbol{\alpha}^{(t)}+\boldsymbol{\delta}-\boldsymbol{\alpha})\Big]\\ &-\Big[\ell(\boldsymbol{\alpha})+(\boldsymbol{\alpha}^{(t)}-\boldsymbol{\alpha})^\top\nabla_1(\boldsymbol{\alpha})+\frac{1}{2}(\boldsymbol{\alpha}^{(t)}-\boldsymbol{\alpha})^\top\big(\mathrm{I}_L+\nabla_2(\boldsymbol{\alpha})\big)(\boldsymbol{\alpha}^{(t)}-\boldsymbol{\alpha})\Big].\end{aligned}$$

Substituting the solution from eq. (15) we have,

$$\boldsymbol{\delta}^\top\nabla_1(\boldsymbol{\alpha})+\boldsymbol{\delta}^\top\big(\mathrm{I}_L+\nabla_2(\boldsymbol{\alpha})\big)(\boldsymbol{\alpha}^{(t)}-\boldsymbol{\alpha}) = -\boldsymbol{\delta}^\top\Big[\lambda_1\mathrm{I}_L+\lambda_2 W\Big]\boldsymbol{\alpha}^{(t)}.$$

Hence,

$$\mathcal{L}\big(\boldsymbol{\alpha}^{(t)}+\delta\big|\boldsymbol{\alpha}\big)-\mathcal{L}\big(\boldsymbol{\alpha}^{(t)}\big|\boldsymbol{\alpha}\big) = -\boldsymbol{\delta}^\top\Big[\lambda_1\mathrm{I}_L+\lambda_2 W\Big]\boldsymbol{\alpha}^{(t)}+\frac{1}{2}\boldsymbol{\delta}^\top\big(\mathrm{I}_L+\nabla_2(\boldsymbol{\alpha})\big)\boldsymbol{\delta}.$$

Using $\boldsymbol{\delta}=(\boldsymbol{\alpha}-\boldsymbol{\alpha}^{(t)})$, $\mathcal{L}\big(\boldsymbol{\alpha}\big|\boldsymbol{\alpha}\big)+P(\boldsymbol{\alpha};\lambda_1,\lambda_2)=F(\boldsymbol{\alpha})$ and (16) after some algebra we have,

$$F(\boldsymbol{\alpha})-F(\boldsymbol{\alpha}^{(t)}) = \frac{1+\lambda_1}{2}||\boldsymbol{\alpha}-\boldsymbol{\alpha}^{(t)}||_2^2+\frac{1}{2}(\boldsymbol{\alpha}-\boldsymbol{\alpha}^{(t)})^\top\big(\nabla_2(\boldsymbol{\alpha})+\lambda_2 W\big)(\boldsymbol{\alpha}-\boldsymbol{\alpha}^{(t)}).$$

Noting that $\nabla_2(\boldsymbol{\alpha})$ and graph Laplacian W are both p.s.d. matrices we have the result in theorem (3.1).

Proof of positive semi-definiteness of W_a:

For $k=1,\ldots,S$ and arbitrary $\boldsymbol{x}=(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_S)^\top \in \mathbb{R}^L, \boldsymbol{x}_k \in \mathbb{R}^{L_k}$ and $L=\sum_k L_k$, if $\boldsymbol{x}^\top W_k \boldsymbol{x} \geq 0$, i.e. each W_k is assumed to be a p.s.d. matrix, then $\boldsymbol{x}^\top W_a \boldsymbol{x} = \sum_k \boldsymbol{x}_k^\top W_k \boldsymbol{x}_k \geq 0$, i.e W_a is also p.s.d.

8.2. **Appendix B: Tables.**

TABLE 8. Table showing regularization parameters in log-scale for ridge (Ridge) and estimates from algorithm 1 (GL) averaged across *all* replications and settings, with their respective standard deviations shown in brackets below. Also, reference distributions used for simulating dispersion ϕ to obtain differing proportions of zeros in simulated response are shown for each spatial pattern, with $U(\cdot,\cdot)$ denoting a uniform distribution.

Spatial Pattern	Regularization Parameters			Proportion of Zeros			
	Ridge	GL					
	λ_1	λ_1	λ_2	0.15	0.30	0.60	0.80
Block	-1.111 (0.116)	-3.135 (0.050)	0.002 (0.104)	$U(7,12)$	$U(12,30)$	$U(30,140)$	$U(140,400)$
Smooth	-0.333 (0.194)	-4.270 (0.049)	1.377 (0.088)	$U(7,11)$	$U(11,24)$	$U(24,100)$	$U(100,200)$
Hot Spot	-5.000 (0.199)	-5.000 (0.000)	2.708 (5.534)	$U(25,40)$	$U(40,70)$	$U(70,200)$	$U(200,500)$
Structured	-0.333 (0.097)	-1.930 (0.090)	0.774 (0.232)	$U(5,7)$	$U(6,14)$	$U(16,35)$	$U(40,80)$

TABLE 9. Comparative results over 100 replications for the un-penalized (MLE), ridge (Ridge) and estimates from algorithm 1 (GL) for a *block spatial pattern* using metrics described in section 4, are shown for varying proportions of zeros in simulated response and sample sizes. The respective standard deviations are shown in brackets below the value of an error metric.

Sample Size	Prop. of zeros	SSE			Deviance Ratio					
					Training			Validation		
		MLE	Ridge	GL	MLE	Ridge	GL	MLE	Ridge	GL
10000	0.15	20.04 (3.58)	16.42 (2.09)	11.13 (1.20)	0.9973 (0.0002)	0.9975 (0.0002)	0.9978 (0.0003)	1.0029 (0.0005)	1.0026 (0.0005)	1.0021 (0.0004)
	0.30	95.40 (94.56)	33.55 (3.54)	19.18 (1.94)	0.9940 (0.0005)	0.9946 (0.0005)	0.9959 (0.0007)	1.0582 (0.1460)	1.0057 (0.0009)	1.0042 (0.0008)
	0.60	2252.97 (575.81)	98.83 (10.60)	42.69 (5.69)	0.9766 (0.0019)	0.9846 (0.0021)	0.9918 (0.0027)	2.9615 (1.0034)	1.0205 (0.0044)	1.0106 (0.0024)
	0.80	7600.11 (720.21)	274.11 (22.39)	84.41 (13.56)	0.9216 (0.0063)	0.9639 (0.0084)	0.9859 (0.0066)	10.2145 (3.9764)	1.0725 (0.0164)	1.0213 (0.0063)
20000	0.15	9.49 (1.21)	8.68 (0.96)	6.75 (0.71)	0.9986 (0.0001)	0.9987 (0.0001)	0.9988 (0.0001)	1.0015 (0.0003)	1.0014 (0.0003)	1.0013 (0.0002)
	0.30	24.95 (26.72)	17.91 (2.15)	11.92 (1.12)	0.9971 (0.0002)	0.9973 (0.0003)	0.9977 (0.0003)	1.0169 (0.1372)	1.0028 (0.0004)	1.0023 (0.0004)
	0.60	572.46 (283.87)	52.09 (5.66)	27.18 (2.98)	0.9890 (0.0010)	0.9912 (0.0010)	0.9942 (0.0013)	1.5624 (0.4604)	1.0102 (0.0020)	1.0067 (0.0014)
	0.80	4237.76 (595.69)	161.90 (14.27)	58.05 (8.43)	0.9585 (0.0029)	0.9764 (0.0035)	0.9890 (0.0042)	5.4494 (2.0800)	1.0361 (0.0070)	1.0153 (0.0037)
30000	0.15	6.15 (0.86)	5.75 (0.72)	4.81 (0.54)	0.9991 (0.0001)	0.9991 (0.0001)	0.9992 (0.0001)	1.0009 (0.0002)	1.0009 (0.0002)	1.0008 (0.0002)
	0.30	13.47 (1.85)	11.90 (1.25)	8.68 (0.89)	0.9981 (0.0002)	0.9982 (0.0002)	0.9984 (0.0002)	1.0020 (0.0003)	1.0019 (0.0003)	1.0016 (0.0003)
	0.60	150.56 (137.63)	38.19 (3.69)	21.30 (2.16)	0.9927 (0.0006)	0.9936 (0.0006)	0.9954 (0.0008)	1.1470 (0.2768)	1.0069 (0.0012)	1.0049 (0.0009)
	0.80	2748.21 (512.76)	110.68 (9.19)	45.42 (5.22)	0.9731 (0.0022)	0.9831 (0.0024)	0.9910 (0.0031)	3.6249 (1.2627)	1.0231 (0.0053)	1.0114 (0.0030)
50000	0.15	3.57 (0.35)	3.43 (0.32)	3.04 (0.27)	0.9995 (0.0000)	0.9995 (0.0000)	0.9995 (0.0000)	1.0006 (0.0001)	1.0005 (0.0001)	1.0005 (0.0001)
	0.30	7.71 (1.02)	7.16 (0.81)	5.78 (0.64)	0.9988 (0.0001)	0.9989 (0.0001)	0.9990 (0.0001)	1.0012 (0.0002)	1.0012 (0.0002)	1.0010 (0.0002)
	0.60	36.91 (22.96)	24.78 (2.72)	15.20 (1.52)	0.9957 (0.0003)	0.9961 (0.0004)	0.9969 (0.0004)	1.0121 (0.0718)	1.0041 (0.0007)	1.0032 (0.0006)
	0.80	1221.82 (432.43)	69.85 (6.84)	33.67 (3.87)	0.9841 (0.0014)	0.9882 (0.0014)	0.9927 (0.0020)	2.1542 (0.6635)	1.0145 (0.0029)	1.0086 (0.0021)

Table 10. Comparative results over 100 replications for the un-penalized (MLE), ridge (Ridge) and estimates from algorithm 1 (GL) for a *smooth spatial pattern* using metrics described in section 4, are shown for varying proportions of zeros in simulated response and sample sizes. The respective standard deviations are shown in brackets below the value of an error metric.

Sample Size	Prop. of zeros	SSE			Deviance Ratio					
					Training			Validation		
		MLE	Ridge	GL	MLE	Ridge	GL	MLE	Ridge	GL
10000	0.15	18.33 (25.39)	13.92 (1.61)	8.23 (0.71)	0.9970 (0.0002)	0.9972 (0.0003)	0.9985 (0.0006)	1.0088 (0.0564)	1.0030 (0.0005)	1.0020 (0.0004)
	0.30	40.73 (38.74)	25.80 (2.71)	11.97 (1.24)	0.9943 (0.0005)	0.9951 (0.0006)	0.9980 (0.0009)	1.0302 (0.1535)	1.0054 (0.0011)	1.0030 (0.0007)
	0.60	821.07 (331.74)	69.87 (7.64)	22.49 (3.34)	0.9807 (0.0017)	0.9848 (0.0022)	0.9950 (0.0031)	2.0941 (0.9489)	1.0173 (0.0033)	1.0058 (0.0015)
	0.60	4425.71 (760.88)	149.59 (16.00)	29.43 (4.89)	0.9442 (0.0046)	0.9673 (0.0050)	0.9855 (0.0026)	8.5876 (2.7101)	1.0466 (0.0100)	1.0097 (0.0034)
20000	0.15	7.58 (0.84)	7.06 (0.75)	5.12 (0.56)	0.9985 (0.0001)	0.9986 (0.0001)	0.9989 (0.0003)	1.0016 (0.0003)	1.0015 (0.0003)	1.0012 (0.0002)
	0.30	14.63 (2.31)	12.91 (1.43)	7.91 (0.71)	0.9972 (0.0003)	0.9974 (0.0003)	0.9986 (0.0005)	1.0029 (0.0005)	1.0027 (0.0005)	1.0019 (0.0003)
	0.60	127.21 (112.53)	38.58 (3.24)	15.40 (1.64)	0.9909 (0.0006)	0.9923 (0.0011)	0.9974 (0.0014)	1.1263 (0.2419)	1.0088 (0.0014)	1.0040 (0.0008)
	0.80	1636.48 (445.75)	87.83 (9.99)	25.20 (4.75)	0.9731 (0.0022)	0.9806 (0.0029)	0.9930 (0.0040)	3.4035 (1.1488)	1.0237 (0.0039)	1.0072 (0.0023)
30000	0.15	4.93 (0.67)	4.73 (0.55)	3.75 (0.36)	0.9990 (0.0001)	0.9990 (0.0001)	0.9992 (0.0002)	1.0010 (0.0002)	1.0010 (0.0002)	1.0009 (0.0002)
	0.30	9.66 (1.09)	8.87 (0.85)	6.08 (0.55)	0.9981 (0.0001)	0.9982 (0.0001)	0.9987 (0.0004)	1.0020 (0.0003)	1.0019 (0.0003)	1.0015 (0.0003)
	0.60	41.61 (31.37)	27.19 (2.53)	12.53 (1.20)	0.9939 (0.0005)	0.9948 (0.0006)	0.9978 (0.0010)	1.0126 (0.0385)	1.0059 (0.0011)	1.0031 (0.0007)
	0.80	659.54 (309.16)	63.13 (7.70)	20.56 (3.15)	0.9828 (0.0014)	0.9866 (0.0021)	0.9953 (0.0030)	1.9031 (0.7373)	1.0155 (0.0027)	1.0055 (0.0014)
50000	0.15	2.88 (0.35)	2.80 (0.34)	2.41 (0.26)	0.9994 (0.0001)	0.9994 (0.0001)	0.9995 (0.0001)	1.0006 (0.0001)	1.0006 (0.0001)	1.0005 (0.0001)
	0.30	5.43 (0.61)	5.20 (0.55)	4.01 (0.44)	0.9989 (0.0001)	0.9989 (0.0001)	0.9992 (0.0002)	1.0011 (0.0002)	1.0011 (0.0002)	1.0009 (0.0002)
	0.60	19.48 (2.76)	16.45 (1.63)	9.15 (0.79)	0.9964 (0.0003)	0.9968 (0.0004)	0.9982 (0.0008)	1.0040 (0.0007)	1.0036 (0.0006)	1.0023 (0.0005)
	0.80	135.68 (119.62)	41.28 (4.51)	15.93 (1.69)	0.9901 (0.0009)	0.9917 (0.0013)	0.9968 (0.0018)	1.1472 (0.2881)	1.0095 (0.0018)	1.0043 (0.0011)

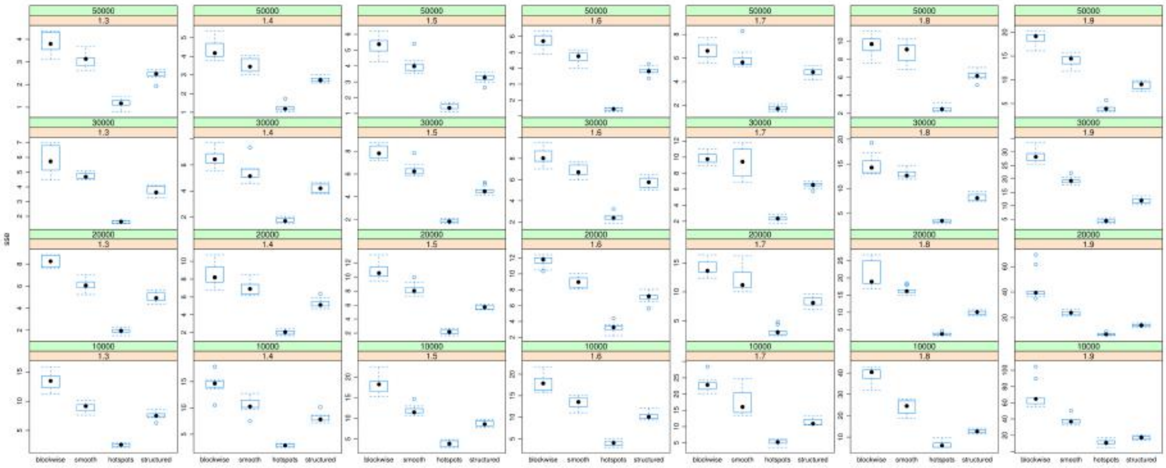
Table 11. Comparative results over 100 replications for the un-penalized (MLE), ridge (Ridge) and estimates from algorithm 1 (GL) for a *hot-spot spatial pattern* using metrics described in section 4, are shown for varying proportions of zeros in simulated response and sample sizes. The respective standard deviations are shown in brackets below the value of an error metric.

Sample Size	Prop. of zeros	SSE			Deviance Ratio					
					Training			Validation		
		MLE	Ridge	GL	MLE	Ridge	GL	MLE	Ridge	GL
10000	0.15	14.27 (1.44)	14.27 (1.44)	2.42 (0.49)	0.9944 (0.0005)	0.9944 (0.0005)	0.9980 (0.0006)	1.0061 (0.0011)	1.0061 (0.0011)	1.0011 (0.0004)
	0.30	24.62 (2.61)	24.67 (2.61)	3.76 (0.90)	0.9906 (0.0009)	0.9906 (0.0009)	0.9967 (0.0017)	1.0100 (0.0019)	1.0100 (0.0019)	1.0018 (0.0007)
	0.60	75.68 (35.23)	73.94 (8.99)	5.53 (0.99)	0.9774 (0.0020)	0.9784 (0.0020)	0.9925 (0.0008)	1.0838 (0.3862)	1.0311 (0.0063)	1.0026 (0.0013)
	0.80	1333.70 (477.50)	266.43 (27.07)	10.49 (3.50)	0.9336 (0.0063)	0.9578 (0.0088)	0.9874 (0.0057)	9.2482 (4.8585)	1.1248 (0.0227)	1.0048 (0.0032)
20000	0.15	6.90 (0.66)	6.91 (0.66)	1.69 (0.22)	0.9972 (0.0002)	0.9972 (0.0002)	0.9985 (0.0002)	1.0030 (0.0005)	1.0030 (0.0005)	1.0008 (0.0003)
	0.30	11.87 (1.16)	11.87 (1.16)	2.16 (0.31)	0.9953 (0.0004)	0.9953 (0.0004)	0.9982 (0.0002)	1.0049 (0.0009)	1.0049 (0.0009)	1.0010 (0.0003)
	0.60	30.31 (3.33)	30.39 (3.36)	4.55 (0.90)	0.9887 (0.0010)	0.9888 (0.0010)	0.9956 (0.0020)	1.0129 (0.0022)	1.0129 (0.0022)	1.0022 (0.0008)
	0.80	142.30 (85.35)	105.68 (14.43)	6.05 (1.27)	0.9701 (0.0028)	0.9722 (0.0028)	0.9919 (0.0010)	1.4604 (1.2308)	1.0477 (0.0096)	1.0032 (0.0015)
30000	0.15	4.56 (0.35)	4.56 (0.35)	1.41 (0.13)	0.9981 (0.0001)	0.9981 (0.0001)	0.9988 (0.0001)	1.0019 (0.0003)	1.0019 (0.0003)	1.0006 (0.0002)
	0.30	7.68 (0.69)	7.69 (0.69)	1.77 (0.21)	0.9969 (0.0003)	0.9969 (0.0003)	0.9985 (0.0002)	1.0032 (0.0005)	1.0032 (0.0005)	1.0008 (0.0002)
	0.60	19.08 (1.84)	19.09 (1.85)	2.98 (0.69)	0.9926 (0.0006)	0.9926 (0.0006)	0.9975 (0.0010)	1.0079 (0.0014)	1.0079 (0.0014)	1.0014 (0.0006)
	0.80	62.30 (25.68)	61.54 (8.42)	5.24 (0.92)	0.9802 (0.0017)	0.9806 (0.0017)	0.9928 (0.0008)	1.0813 (0.5582)	1.0265 (0.0062)	1.0025 (0.0010)
50000	0.15	4.56 (0.25)	4.56 (0.25)	1.41 (0.14)	0.9981 (0.0001)	0.9981 (0.0001)	0.9988 (0.0002)	1.0019 (0.0002)	1.0019 (0.0002)	1.0006 (0.0001)
	0.30	2.70 (0.41)	2.70 (0.41)	1.09 (0.14)	0.9989 (0.0002)	0.9989 (0.0002)	0.9993 (0.0001)	1.0011 (0.0003)	1.0011 (0.0003)	1.0005 (0.0002)
	0.60	10.70 (0.89)	10.70 (0.89)	2.01 (0.29)	0.9957 (0.0003)	0.9957 (0.0003)	0.9982 (0.0002)	1.0046 (0.0008)	1.0046 (0.0008)	1.0009 (0.0003)
	0.80	30.65 (3.20)	30.69 (3.21)	4.32 (0.76)	0.9886 (0.0010)	0.9886 (0.0010)	0.9959 (0.0020)	1.0128 (0.0023)	1.0128 (0.0023)	1.0019 (0.0008)

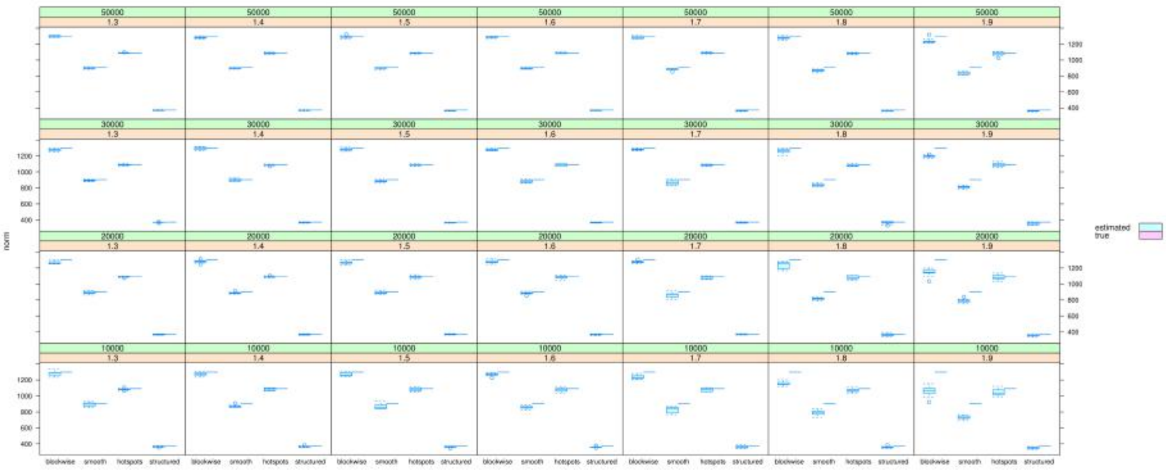
TABLE 12. Comparative results over 100 replications for the un-penalized (MLE), ridge (Ridge) and estimates from algorithm 1 (GL) for a *structured spatial pattern* using metrics described in section 4, are shown for varying proportions of zeros in simulated response and sample sizes. The respective standard deviations are shown in brackets below the value of an error metric.

Sample Size	Prop. of zeros	SSE			Deviance Ratio					
					Training			Validation		
		MLE	Ridge	GL	MLE	Ridge	GL	MLE	Ridge	GL
10000	0.15	14.82 (1.57)	13.21 (1.21)	9.17 (0.90)	0.9945 (0.0005)	0.9950 (0.0005)	0.9964 (0.0011)	1.0058 (0.0011)	1.0052 (0.0009)	1.0037 (0.0008)
	0.30	25.08 (2.49)	20.69 (1.80)	12.19 (1.08)	0.9912 (0.0007)	0.9924 (0.0007)	0.9950 (0.0017)	1.0098 (0.0017)	1.0082 (0.0014)	1.0051 (0.0011)
	0.60	100.39 (50.38)	47.84 (4.27)	19.90 (1.86)	0.9764 (0.0019)	0.9844 (0.0021)	0.9902 (0.0031)	1.0768 (0.1631)	1.0182 (0.0034)	1.0081 (0.0022)
	0.80	1187.51 (418.37)	98.75 (8.15)	28.30 (3.42)	0.9394 (0.0050)	0.9660 (0.0123)	0.9849 (0.0044)	4.7115 (2.4765)	1.0411 (0.0069)	1.0120 (0.0039)
20000	0.15	7.26 (0.57)	6.86 (0.50)	5.50 (0.48)	0.9972 (0.0002)	0.9974 (0.0002)	0.9977 (0.0003)	1.0029 (0.0004)	1.0027 (0.0004)	1.0022 (0.0004)
	0.30	11.88 (1.16)	10.76 (0.96)	7.82 (0.81)	0.9956 (0.0004)	0.9959 (0.0004)	0.9968 (0.0008)	1.0047 (0.0008)	1.0042 (0.0007)	1.0031 (0.0006)
	0.60	33.45 (3.87)	26.21 (2.18)	14.41 (1.17)	0.9885 (0.0010)	0.9905 (0.0009)	0.9941 (0.0023)	1.0131 (0.0026)	1.0104 (0.0020)	1.0058 (0.0014)
	0.80	161.38 (82.12)	55.14 (4.12)	21.28 (2.11)	0.9719 (0.0026)	0.9830 (0.0025)	0.9899 (0.0035)	1.1974 (0.3213)	1.0225 (0.0040)	1.0094 (0.0026)
30000	0.15	4.73 (0.49)	4.56 (0.45)	3.87 (0.35)	0.9982 (0.0002)	0.9982 (0.0002)	0.9984 (0.0002)	1.0019 (0.0003)	1.0018 (0.0003)	1.0015 (0.0002)
	0.30	7.85 (0.77)	7.35 (0.63)	5.80 (0.52)	0.9970 (0.0002)	0.9972 (0.0002)	0.9976 (0.0003)	1.0031 (0.0005)	1.0029 (0.0005)	1.0023 (0.0004)
	0.60	21.15 (2.23)	17.98 (1.68)	11.18 (1.09)	0.9924 (0.0007)	0.9933 (0.0007)	0.9953 (0.0014)	1.0084 (0.0016)	1.0073 (0.0013)	1.0048 (0.0011)
	0.80	65.80 (32.03)	39.20 (3.05)	17.89 (1.76)	0.9815 (0.0016)	0.9865 (0.0014)	0.9918 (0.0027)	1.0387 (0.1055)	1.0156 (0.0028)	1.0076 (0.0020)
50000	0.15	2.80 (0.24)	2.74 (0.24)	2.46 (0.22)	0.9989 (0.0001)	0.9989 (0.0001)	0.9990 (0.0001)	1.0011 (0.0002)	1.0011 (0.0002)	1.0010 (0.0002)
	0.30	4.41 (0.41)	4.25 (0.39)	3.67 (0.35)	0.9983 (0.0001)	0.9984 (0.0001)	0.9985 (0.0002)	1.0017 (0.0003)	1.0017 (0.0003)	1.0015 (0.0003)
	0.60	12.01 (1.06)	10.93 (0.85)	7.98 (0.76)	0.9955 (0.0003)	0.9958 (0.0003)	0.9967 (0.0007)	1.0047 (0.0008)	1.0043 (0.0007)	1.0032 (0.0006)
	0.80	31.25 (3.66)	24.54 (2.14)	13.80 (1.40)	0.9893 (0.0009)	0.9911 (0.0009)	0.9942 (0.0019)	1.0121 (0.0023)	1.0096 (0.0018)	1.0056 (0.0014)

8.3. **Appendix C: Figures.**



(A)



(B)

FIGURE 8. Figures showing (a) SSEs, (b) euclidean norms of estimated (using $p = 1.5$) and true spatial effects for 282 zip codes in CT, for different sample sizes and true values of the index parameter, p under different spatial patterns. Results are obtained from 10 replications under each combination of sample and parameter settings. In both figures (a) and (b) the reference (where the index parameters match, i.e. simulated data and estimated effects have same index parameters) is the column with $p = 1.5$, any significant departure is indicative of inferior performance. Figure (b) shows boxplots of norms for estimates under each pattern with the true norm shown as a horizontal line grouped along with it.

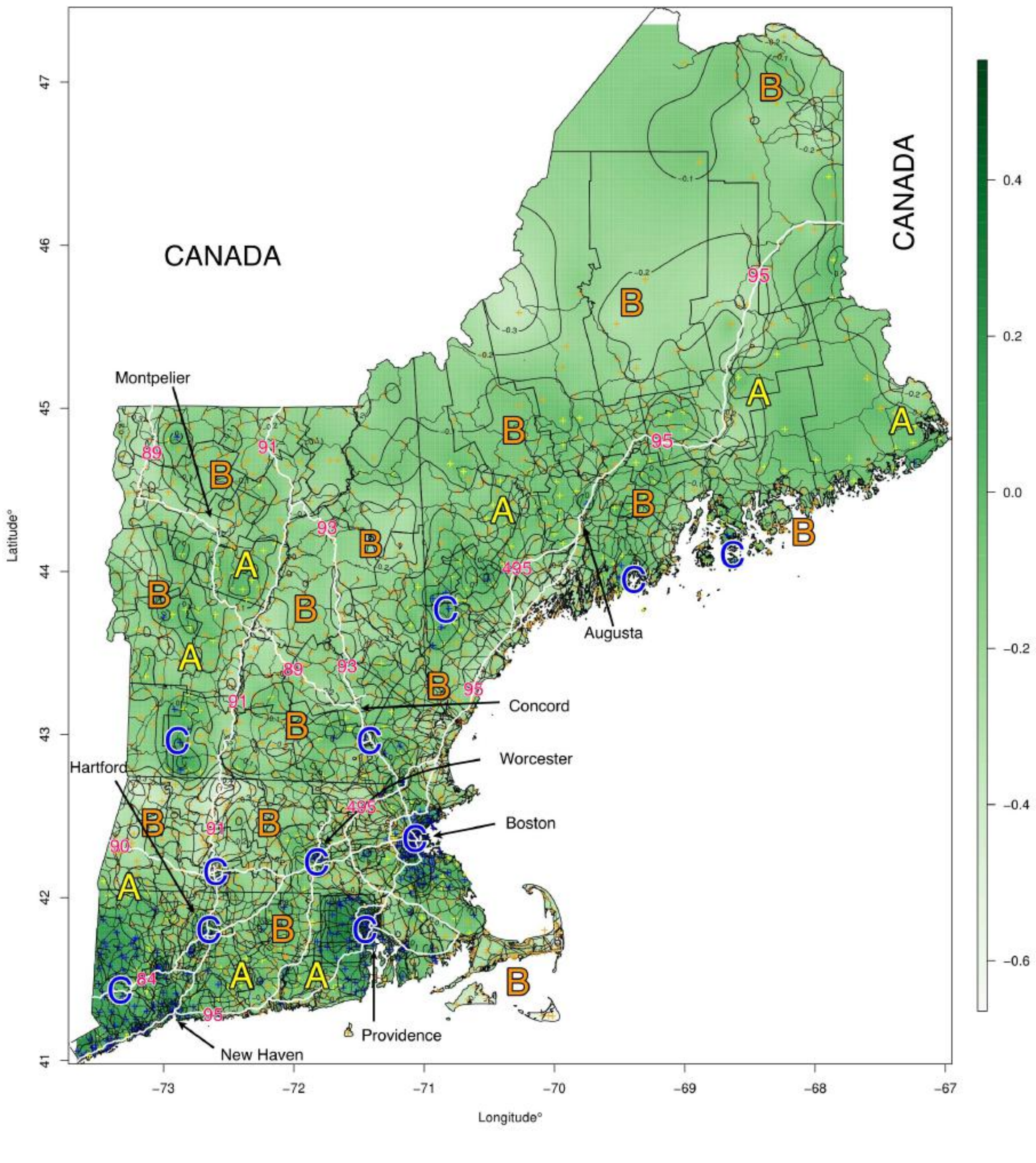
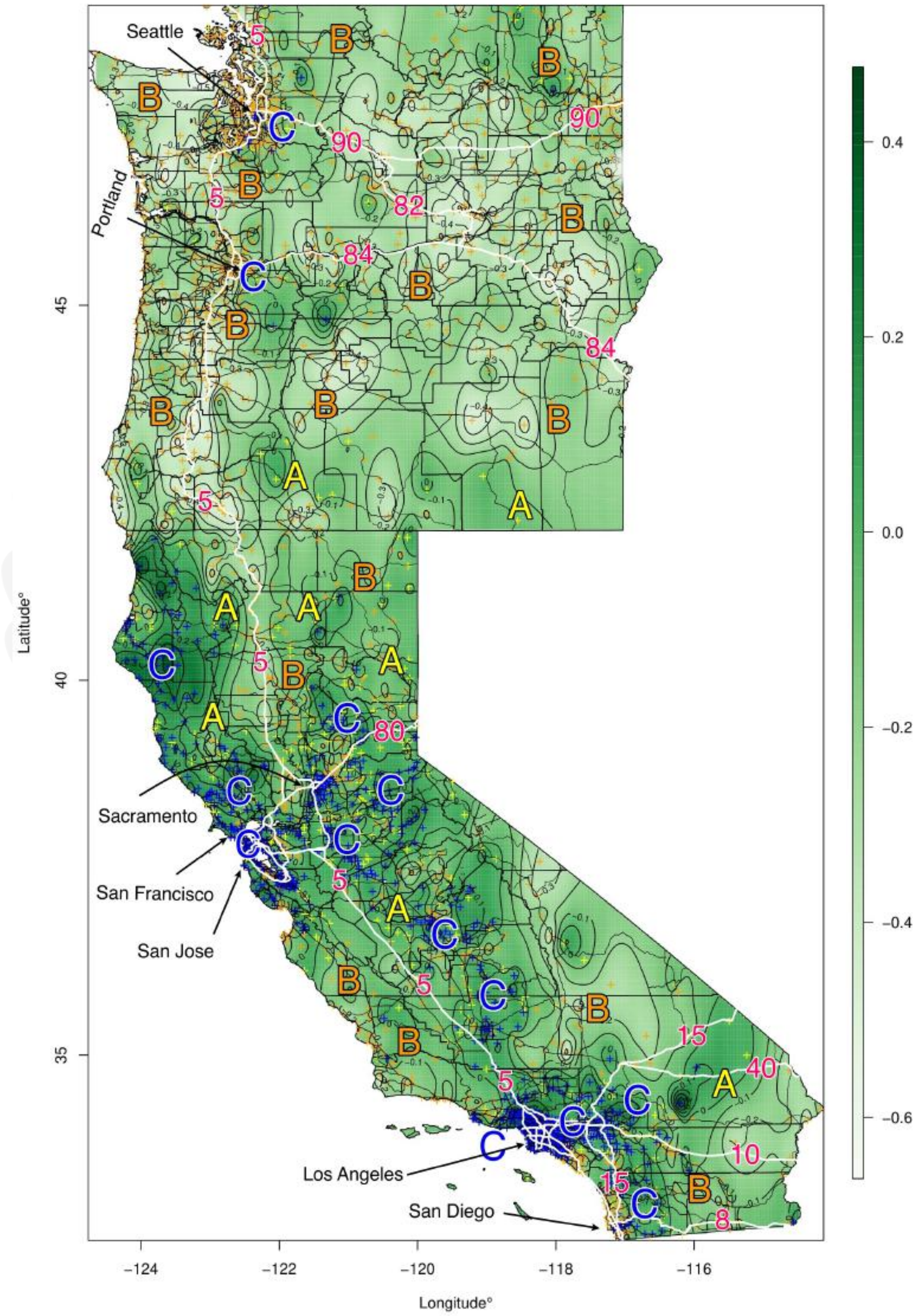


FIGURE 9. Plot showing estimated spatial effects for states in New England (with county borders). It includes all primary (marked in bold "white") and secondary (marked in "black" lines) roads. zip codes (1831 in total) are color coded, ones with 0 in their approx. CIs are "yellow" (marked in "A"), ones below zero are "orange" (sizable regions marked with "B"), and ones above are "blue" (regions marked with "C"). Some major cities are marked with arrows.



(A)

FIGURE 10. Plot showing estimated spatial effects for states in West Coast (with county borders). It includes all primary (marked in bold "white") and secondary (marked in "black" lines) roads. zip codes (2775 in total) are color coded, ones with 0 in their approx. CIs are "yellow" (marked in "A"), ones below zero are "orange" (sizable regions marked with "B"), and ones above are "blue" (regions marked with "C"). Some major cities are marked with arrows.

On Some Well-Posedness Issues of the Inviscid Burgers Equation on Sobolev Spaces

Hasan İnci

Abstract. In this paper we consider for $n \geq 1$ the n dimensional inviscid Burgers equation on the Sobolev space $H^s(\mathbb{R}^n)$, $s > n/2+1$. In Lagrangian coordinates the Burgers equation is "linear". We will use this simple structure of the Burgers equation to study some well-posedness issues for classical solutions. The Burgers equation is locally well-posed in $H^s(\mathbb{R}^n)$ in the "classical" regime $s > n/2+1$. We will show that its solution map, mapping the initial value to the corresponding solution, cannot be continuously extended to $H^{s_c}(\mathbb{R}^n)$ for the borderline case $s_c = n/2+1$. We will further show that the solution map is nowhere locally uniformly continuous in the classical regime $s > n/2+1$. We will also show that there are always solutions which blow up in finite time.

The purpose of this paper is to give some new results, improve known results and to illustrate a Lagrangian approach in the simple setting of the Burgers equation which can be applied to more complicated equations like the incompressible Euler equations.

2020 Mathematics Subject Classifications: 35Q35

Keywords: Inviscid Burgers equation, Well-posedness

1. Introduction

For $n \geq 1$ the initial value problem for the inviscid Burgers equation in $\mathbb{R}^n$ is given by

$$u_t + (u \cdot \nabla)u = 0, \quad u(t=0) = u_0, \tag{1}$$

where $u = (u_1, \dots, u_n) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is the unknown and $u \cdot \nabla$ is the differential operator $\sum_{k=1}^n u_k \partial_k$ acting component wise. The most prominent case of (1) is $n = 1$, i.e.

$$u_t + uu_x = 0, \quad u(t=0) = u_0.$$

This is the simplest nonlinear PDE. It expresses in various settings such as gas dynamics and traffic flow a conservation law.

In this paper we are interested in well-posedness properties of (1). Some of the results we discuss here are known. The novelty in this paper is the use of Lagrangian coordinates of Sobolev type. This simplifies (1) and leads to a strenghtening of previous known results. Another purpose of this paper is to illustrate an approach in the simple setting of the Burgers equation which can be used for more complicated equations like the incompressible Euler equations – see e.g. [3].

The function spaces we are dealing with in this paper are the Sobolev spaces. Recall that for $s \geq 0$ the Sobolev space on $\mathbb{R}^n$ of class s is defined as

$$H^s(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) \mid \|f\|_{H^s} = \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^s |\hat{f}(\xi)|^2 \, d\xi\right)^{1/2} < \infty\},$$

where $\hat{f}$ denotes the Fourier transform of f. The Sobolev space $H^s(\mathbb{R}^n)$ turns out to be the completion of $C_c^\infty(\mathbb{R}^n)$, the space of compactly supported smooth functions on $\mathbb{R}^n$, with respect to the norm $\|\cdot\|_{H^s}$. We denote by $H^s(\mathbb{R}^n;\mathbb{R}^n)$ the vector valued versions of these spaces. The spaces $H^s(\mathbb{R}^n)$ resp. $H^s(\mathbb{R}^n;\mathbb{R}^n)$ for $s \geq 0$ are a scale of Hilbert spaces decreasing in s. For $s > n/2+1$ we have the Sobolev imbedding $H^s(\mathbb{R}^n) \hookrightarrow C_0^1(\mathbb{R}^n)$. Here $C_0^1(\mathbb{R}^n)$ is the space of C^1 functions on $\mathbb{R}^n$ vanishing at infinity together with their derivative endowed with the C^1 norm. For $s > n/2+1$ we say that u is a solution to (1) on $[0,T]$ for some $T > 0$ if $u \in C([0,T]; H^s(\mathbb{R}^n;\mathbb{R}^n))$ and

$$u(t) = u_0 + \int_0^t (u(s) \cdot \nabla)u(s) \, ds, \; 0 \leq t \leq T,$$

as an identity in $H^{s-1}(\mathbb{R}^n;\mathbb{R}^n)$. Note that $H^{s-1}(\mathbb{R}^n)$ is a Banach algebra under pointwise multiplication. We see in particular that for $s > n/2+1$ the equation (1) is satisfied in the classical sence, i.e. with classical derivatives.
In [2] the authors discuss for $s > n/2+1$ the functional space

$$\mathcal{D}^s(\mathbb{R}^n) = \{\varphi : \mathbb{R}^n \to \mathbb{R}^n \mid \varphi - \mathrm{id} \in H^s(\mathbb{R}^n;\mathbb{R}^n),\ \det(d_x\varphi) > 0\ \forall x \in \mathbb{R}^n\},$$

where $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$ is the identity map. Note that due to the Sobolev imbedding $H^s(\mathbb{R}^n) \hookrightarrow C_0^1(\mathbb{R}^n)$ the expression $\det(d_x\varphi)$ is well defined. The functional space $\mathcal{D}^s(\mathbb{R}^n)$ consists of C^1 diffeomorphisms of $\mathbb{R}^n$. Moreover $\mathcal{D}^s(\mathbb{R}^n)$ can be identified via $\varphi \mapsto \varphi - \mathrm{id}$ with an open subset of $H^s(\mathbb{R}^n;\mathbb{R}^n)$ making it a Hilbert manifold. In [2] it was shown that $\mathcal{D}^s(\mathbb{R}^n)$ is a topological group when the group operation is composition of maps. More generally, for $0 \le s' \le s$ the composition map

$$H^{s'}(\mathbb{R}^n) \times \mathcal{D}^s(\mathbb{R}^n) \to H^{s'}(\mathbb{R}^n),\ (f,\varphi) \mapsto f \circ \varphi, \tag{2}$$

is continuous.

2. Local well-posedness of classical solutions

Suppose that $s > n/2+1$ and that $u \in C([0,T];H^s(\mathbb{R}^n;\mathbb{R}^n))$ is a solution to (1) for some $T > 0$. In [3] it was shown that such a u generates a unique flow map $\varphi \in C^1([0,T];\mathcal{D}^s(\mathbb{R}^n))$, i.e. φ satisfies

$$\varphi_t(t) = u(t) \circ \varphi(t),\ 0 \le t \le T,\ \varphi(0) = \mathrm{id}.$$

From this we see that $\mathcal{D}^s(\mathbb{R}^n)$ is the right space for a Lagrangian formulation of (1), i.e. a formulation in terms of the flow map of u. The formulation of (1) in terms of φ turns out to be very easy. Indeed by taking the t derivative of $\varphi_t(t) = u(t) \circ \varphi(t)$ we get

$$\frac{d}{dt}\varphi_t(t) = (u_t(t) + (u(t)\cdot\nabla)u(t)) \circ \varphi(t) = 0.$$

By using the initial values $\varphi(0) = \mathrm{id}$ and $\varphi_t(0) = u_0$ we can solve this explicitly by

$$\varphi(t) = \mathrm{id} + tu_0,\ 0 \le t \le T.$$

With this the solution to (1) is formally given by

$$u(t) = u_0 \circ \varphi(t)^{-1},\ 0 \le t \le T.$$

But due to (2) and the continuity of the inversion map on $\mathcal{D}^s(\mathbb{R}^n)$ this is rigorously true and we immediately get the following well known result.

Theorem 2.1. *Let $n \ge 1$ and $s > n/2+1$. Then the Burgers equation* (1) *is locally well-posed in $H^s(\mathbb{R}^n;\mathbb{R}^n)$.*

Proof. Let $u_0 \in H^s(\mathbb{R}^n;\mathbb{R}^n)$. Then by the Sobolev imbedding $H^s(\mathbb{R}^n;\mathbb{R}^n) \hookrightarrow C_0^1(\mathbb{R}^n;\mathbb{R}^n)$ there is $\delta > 0$ s.t.

$$\varphi(t) = \mathrm{id} + tu_0 \in \mathcal{D}^s(\mathbb{R}^n),\ 0 \le t \le \delta.$$

We define

$$u(t) = u_0 \circ \varphi(t)^{-1},\ 0 \le t \le \delta.$$

From (2) and the continuity of the inversion map we have

$$u \in C([0,\delta];H^s(\mathbb{R}^n;\mathbb{R}^n)).$$

Note that we have for all $x \in \mathbb{R}^n$ and $0 \le t \le \delta$

$$\varphi(t,x) \circ \varphi^{-1}(t,x) = x.$$

Thus by the Implicit Function Theorem we get pointwise

$$\frac{d}{dt}\varphi^{-1}(t,x) = [d\varphi(t,x)]^{-1} \circ \varphi^{-1}(t,x) \cdot \left(\frac{d}{dt}\varphi(t,x)\right) \circ \varphi^{-1}(t,x).$$

Computing $(u(t)\cdot\nabla)u(t) = du(t)\cdot u(t)$ gives pointwise

$$du_0 \circ \varphi(t)^{-1} \cdot [d\varphi(t)]^{-1} \circ \varphi(t)^{-1} \cdot u_0 \circ \varphi(t)^{-1} = \frac{d}{dt}u_0 \circ \varphi(t)^{-1}. \tag{3}$$

Thus we have pointwise for $0 \le t \le \delta$

$$u(t) = u_0 \circ \varphi(t)^{-1} = u_0 + \int_0^t \frac{d}{ds}u_0 \circ \varphi(s)^{-1}\, ds = u_0 + \int_0^t (u(s)\cdot\nabla)u(s)\, ds.$$

But this is an identity in $H^{s-1}(\mathbb{R}^n;\mathbb{R}^n)$ as well, since the integrand is in $C([0,\tau];H^{s-1}(\mathbb{R}^n;\mathbb{R}^n))$ due to the Banach algebra property of $H^{s-1}(\mathbb{R}^n)$ under pointwise multiplication. So u solves (1) on $[0,\delta]$.
Now if u is a solution to (1) on $[0,\delta]$ its flow map φ is necessarily of the form

$$\varphi(t) = \mathrm{id} + tu_0,\ 0 \le t \le \delta,$$

showing uniqueness of solutions.
Finally for a fixed $0 \le t \le \delta$ we know that

$$U \subset H^s(\mathbb{R}^n;\mathbb{R}^n) \to H^s(\mathbb{R}^n;\mathbb{R}^n),\ w_0 \mapsto w_0 \circ (\mathrm{id} + tw_0)^{-1},$$

is continuous. Here U is neighborhood of u_0 small enough to ensure that for $w_0 \in U$ the expression $\mathrm{id} + tw_0$ is invertible in $\mathcal{D}^s(\mathbb{R}^n)$.
So from existence, uniqueness and stability of solutions we conclude that (1) is locally well-posed in $H^s(\mathbb{R}^n;\mathbb{R}^n)$. □

Remark 1. *We see from the proof of Theorem 2.1 that the solution to* (1) *exists as long as* $id + tu_0$ *stays in* $\mathcal{D}^s(\mathbb{R}^n)$.

For a solution u of (1) the scaled quantity

$$u_\lambda(t,x) = \lambda u(\lambda t, x), \quad \lambda > 0,$$

is a solution to

$$u_t + (u \cdot \nabla)u = 0,\ u(t=0) = \lambda u_0.$$

If we fix $s > n/2+1$ and $T > 0$, then due to the scaling property above and Theorem 2.1 the set $U_T^{(s)} \subset H^s(\mathbb{R}^n;\mathbb{R}^n)$ of initial values u_0 for which the solution exists beyond time T is an open star shaped subset with respect to $0 \in H^s(\mathbb{R}^n,\mathbb{R}^n)$. With this the solution map

$$\Phi_T^{(s)} : U_T^{(s)} \subset H^s(\mathbb{R}^n;\mathbb{R}^n) \to H^s(\mathbb{R}^n;\mathbb{R}^n),\ u_0 \mapsto u(T),$$

mapping the initial value u_0 to the time T value of the corresponding solution is well defined and by Theorem 2.1 continuous. From the explicit solution formula

$$\varphi(t) = \mathrm{id} + tu_0$$

we immediately get that for $s > s' > n/2+1$ we have

$$U_T^{(s)} \subset U_T^{(s')}$$

and

$$\Phi_T^{(s)} = \Phi_T^{(s')}\Big|_{U_T^{(s)}}.$$

So for the rest of the paper we will skip the index $s > n/2+1$ and write just U_T resp. Φ_T. In other words the time T solution map is given by

$$\Phi_T : U_T \subset H^s(\mathbb{R}^n;\mathbb{R}^n) \to H^s(\mathbb{R}^n;\mathbb{R}^n),\ u_0 \mapsto u(T). \tag{4}$$

For $T = 1$ we will skip the index T and write U resp. Φ for U_T resp. Φ_T, i.e. we have

$$\Phi : U \subset H^s(\mathbb{R}^n;\mathbb{R}^n) \to H^s(\mathbb{R}^n;\mathbb{R}^n),\ u_0 \mapsto u_0 \circ (\mathrm{id} + u_0)^{-1}.$$

Note that due to the scaling property of (1) we have for $T > 0$

$$\Phi_T(u_0) = \frac{1}{T}\Phi(Tu_0). \tag{5}$$

3. Ill-posedness for $s_c = n/2+1$

Suppose that u is a solution to (1). Then the scaled quantity

$$u^\lambda(t,x) = \frac{1}{\lambda}u(t,\lambda x)$$

for $\lambda > 0$ is a solution to

$$u_t + (u \cdot \nabla)u = 0,\ u(0,x) = \frac{1}{\lambda}u_0(\lambda x).$$

This scaling property makes $s_c = n/2+1$ the "critical" regularity. More precisely we have for $s_c = n/2+1$ the scale invariance

$$\begin{aligned}\|u^\lambda\|^2_{\dot{H}^{s_c}} &= \int_{\mathbb{R}^n} |\xi|^{2s_c} |\widehat{u^\lambda}(\xi)|^2 \, d\xi \\ &= \int_{\mathbb{R}^n} |\xi|^{2s_c} \frac{1}{\lambda^{2n+2}} |\hat{u}(\xi/\lambda)|^2 \, d\xi \\ &= \int_{\mathbb{R}^n} |\xi|^{2s_c} |\hat{u}(\xi)|^2 \, d\xi = \|u\|^2_{\dot{H}^{s_c}}\end{aligned}$$

for the homogeneous Sobolev norm $\|\cdot\|_{\dot{H}^{s_c}}$.
To prove ill-posedness for $s_c = n/2+1$ we will use the following Lemma which tells us that the H^{s_c-1} norm cannot control the L^∞ norm.

Lemma 3.1. *Let $s_c = n/2+1$. Then for every $\varepsilon > 0$ there is $f \in H^{s_c}(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n)$ s.t.*

$$\|f\|_{H^{s_c}} < \varepsilon \text{ and } \min_{x\in\mathbb{R}} \partial_x f(x) = \partial_1 f(0) = -1. \tag{6}$$

Proof. Let $\varepsilon > 0$. For $2 \le A < B < \infty$ we define

$$h_{A,B}(\xi) = \frac{i\xi_1}{|\xi|^{s_c+3/2}\ln(|\xi|)} \cdot \chi_{\{A\le|\xi|\le B\}}(\xi),\ \xi \in \mathbb{R}^n,$$

where $\chi_{\{A\le|\xi|\le B\}}(\xi)$ denotes the characteristic function of the set $\{A \le |\xi| \le B\}$. Note that $h_{A,B}$ is bounded, has a compact support and it satisfies $h_{A,B}(-\xi) = \overline{h_{A,B}(\xi)}$. Thus $h_{A,B}$ is the Fourier transform of some $f_{A,B} \in H^{s_c}(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n)$, i.e. $\hat{f}_{A,B}(\xi) = h_{A,B}(\xi)$. For $f_{A,B}$ we have

$$\|f_{A,B}\|^2_{H^{s_c}} = \int_{\{A\le|\xi|\le B\}} (1+|\xi|^2)^{s_c} \frac{\xi_1^2}{|\xi|^{2s_c+3}\ln^2(|\xi|)}\, d\xi.$$

Since $\int_2^\infty \frac{1}{r\ln^2(r)}\, dr < \infty$ we get that

$$\|f_{A,B}\|_{H^{s_c}} \to 0 \text{ for } A \to \infty.$$

By using the inversion formula for the Fourier transform we have the representation

$$\partial_1 f_{A,B}(x) = \int_{\{A\le|\xi|\le B\}} e^{ix\cdot\xi} \frac{-\xi_1^2}{|\xi|^{s_c+3/2}\ln(|\xi|)}\, d\xi,$$

from which we see that $\partial_1 f_{A,B}(0) = \min_{x\in\mathbb{R}^n} \partial_1 f_{A,B}(x)$. Since $\int_2^\infty \frac{1}{r\ln(r)}\, dr = \infty$ we can make $\partial_1 f_{A,B}(0)$ as negative as we want by choosing B large enough. Thus choosing A large enough and adjusting B we get $\|f_{A,B}\|_{H^{s_c}} < \varepsilon$ and $\min_{x\in\mathbb{R}^n} \partial_1 f_{A,B}(x) = \partial_1 f_{A,B}(0) = -1$. This finishes the proof. □

Our ill-posedness result reads as

Theorem 3.2. *Let $s > n/2+1$, $s_c = n/2+1$ and*

$$\Phi : U \subset H^s(\mathbb{R}^n;\mathbb{R}^n) \to H^s(\mathbb{R}^n;\mathbb{R}^n),\ u_0 \mapsto \Phi(u_0) = u_0 \circ (id + u_0)^{-1}$$

the time $T = 1$ solution map to (1). There is no continuous extension $\tilde{\Phi}$ of Φ to an open neighborhood $\tilde{U} \subset H^{s_c}(\mathbb{R}^n;\mathbb{R}^n)$ of $0 \in H^{s_c}(\mathbb{R}^n;\mathbb{R}^n)$.

Proof. Suppose there is a continuous extension $\tilde{\Phi}$ of Φ to the ball $B_\varepsilon(0) \subset H^{s_c}(\mathbb{R}^n;\mathbb{R}^n)$ for some radius $\varepsilon > 0$, i.e. we have $\tilde{\Phi} : B_\varepsilon(0) \to H^{s_c}(\mathbb{R}^n;\mathbb{R}^n)$ is continuous with $\Phi|_{B_\varepsilon(0)\cap U} = \tilde{\Phi}\big|_{B_\varepsilon(0)\cap U}$. We take $f \in C_c^\infty(\mathbb{R}^n)$ as in Lemma 3.1, i.e. f satisfies (6) and define

$$u_0 = (f, 0, \dots, 0) \in B_\varepsilon(0) \text{ and } \varphi : t \mapsto \mathrm{id} + tu_0.$$

Since $\det(d_x\varphi(t)) = 1 + t\partial_1 f(x)$ we have $tu_0 \in U$ for $0 \le t < 1$. Note that for $0 \le t < 1$

$$(x_1, 0, \dots, 0) \mapsto \varphi(t, x_1, 0, \dots, 0) = (x_1 + tf(x_1, 0, \dots, 0), 0, \dots, 0)$$

is a diffeomorphism from the x_1 axis onto itself. Since $tu_0 \in U$ for $0 \le t < 1$ we have

$$\tilde{\Phi}(tu_0) = tu_0 \circ \varphi(t)^{-1} = (tf \circ \varphi(t)^{-1}, 0, \dots, 0),\ 0 \le t < 1.$$

Applying ∂_1 gives

$$\partial_1\tilde{\Phi}(tu_0) = (\left(\frac{t\partial_1 f}{1+t\partial_1 f}\right)\circ\varphi(t)^{-1}, 0, \ldots, 0),\ 0 \le t < 1.$$

Since $s_c - 1 = n/2$ we know e.g. from [4] that $\partial_1\tilde{\Phi}(tu_0)$ has a trace on the x_1 axis and that this trace belongs to $H^{1/2}(\mathbb{R})$ and hence to $L^2(\mathbb{R})$. So consider

$$\begin{aligned}\| \partial_1\tilde{\Phi}(tu_0)\Big|_{x_1 \text{ axis}} \|_{L^2}^2 &= \int_{\mathbb{R}} \left| \left(\frac{t\partial_1 f(x_1,0,\ldots,0)}{1+t\partial_1 f(x_1,0,\ldots,0)}\right)\circ\varphi(t)^{-1}\right|^2\, dx_1 \\ &= \int_{\mathbb{R}} t^2 \frac{|\partial_1 f(x_1,0,\ldots,0)|^2}{1+t\partial_1 f(x_1,0,\ldots,0)}\, dx_1.\end{aligned}$$

From (6) we conclude that there is $\delta > 0$ s.t.

$$-1 \le \partial_1 f(x_1,0,\ldots,0) \le -1 + cx_1^2,\ -\delta < x_1 < \delta,$$

for some $c > 0$. By making $\delta > 0$ smaller if necessary we can also assume

$$|\partial_1 f(x_1,0,\ldots,0)|^2 \ge 1/2,\ -\delta < x_1 < \delta.$$

Thus we have for $0 \le t < 1$

$$\| \partial_1\tilde{\Phi}(tu_0)\Big|_{x_1 \text{ axis}} \|_{L^2}^2 \ge \int_{-\delta}^{\delta} \frac{1}{2} t^2 \frac{1}{1-t+tcx_1^2}\, dx_1 = t^2 \frac{\arctan(\frac{\sqrt{ct}\delta}{\sqrt{1-t}})}{\sqrt{ct}\sqrt{1-t}}.$$

Therefore

$$\lim_{t\uparrow 1} \| \partial_1\tilde{\Phi}(tu_0)\Big|_{x_1 \text{ axis}} \|_{L^2} = \infty,$$

which contradicts the H^{s_c} continuity of $\tilde{\Phi}$ in u_0 together with the existence of the trace on the x_1 axis in L^2. So there is no continuous extension $\tilde{\Phi}$ of Φ to a neighborhood of 0 in $H^{s_c}(\mathbb{R}^n)$. □

4. Nonuniform dependence for $s > n/2 + 1$

The goal of this section is to prove

Theorem 4.1. *Let $n \ge 1$, $s > n/2 + 1$ and $T > 0$. Then the solution map*

$$\Phi_T : U_T \subset H^s(\mathbb{R}^n;\mathbb{R}^n) \to H^s(\mathbb{R}^n;\mathbb{R}^n),\ u_0 \mapsto u(T),$$

is nowhere locally uniformly continuous.

Remark 2. *Theorem 4.1 tells us that for any nonempty open set $W \subset U_T$ the restriction $\Phi_T|_W$ is not uniformly continuous. In particular Φ_T is nowhere locally Lipschitz, nowhere locally Hölder continuous and nowhere C^1.*

In [5] Kato proves that there is no exponent $0 < \alpha \le 1$ such that the solution map Φ_T is Hölder continuous to the exponent α. Theorem 4.1 is a strengthening of this.
Before we prove Theorem 4.1 we need some technical lemmas about fractional Sobolev spaces.

Lemma 4.2. *Let $s > n/2+1$. Further let $A, B \subset \mathbb{R}^n$ be two disjoint compact subsets. There is a constant $C > 0$ s.t.*

$$\|f\|_{H^s} + \|g\|_{H^s} \le C\|f+g\|_{H^s}$$

for all $f, g \in H^s(\mathbb{R}^n)$ with $\operatorname{supp} f \subset A$, $\operatorname{supp} g \subset B$. *Here* supp *denotes the support of a function.*

Proof. Let $\alpha, \beta \in C_c^\infty(\mathbb{R}^n)$ s.t. the supports of α and β are disjoint and we have

$$\alpha\big|_A \equiv 1,\ \beta\big|_B \equiv 1.$$

Using that multiplication with α resp. β is bounded in $H^s(\mathbb{R}^n)$ we get

$$\|f\|_{H^s} = \|\alpha(f+g)\|_{H^s} \le C\|f+g\|_{H^s}$$

resp.

$$\|g\|_{H^s} = \|\beta(f+g)\|_{H^s} \le C\|f+g\|_{H^s}$$

for some $C > 0$. This finishes the proof. □

Lemma 4.3. *Let $s > n/2 + 1$. Then there is $C > 0$ such that for every $0 < R \le 1$ and balls $B_1, B_2 \subset \mathbb{R}^n$ with radii R and whose centers are $4R$ apart we have the following:*

$$\|f\|_{H^s} + \|g\|_{H^s} \le C\|f + g\|_{H^s}$$

for all $f, g \in H^s(\mathbb{R}^n)$ with $\operatorname{supp} f \subset B_1$, $\operatorname{supp} g \subset B_2$.

Proof. Let $0 < R \le 1$. Without loss of generality we assume that B_1 is the ball of radius R with center $(-2R, 0, \ldots, 0)$ and B_2 is the ball of radius R with center $(2R, 0, \ldots, 0)$. For a fixed compact $K \subset \mathbb{R}^n$ the homogeneous norm $\|\cdot\|_{\dot{H}^s}$ is equivalent to $\|\cdot\|_{H^s}$ for functions with support in K – see [1]. We denote for a function h by h^R its scaling by R, i.e.

$$h^R(x) = h(Rx), \; x \in \mathbb{R}^n.$$

We then have $\|h^R\|_{\dot{H}^s} = R^{s-n/2}\|h\|_{\dot{H}^s}$. Note that f^R has its support in the ball A of radius 1 with center $(-2, 0, \ldots, 0)$ and g^R has its support in the ball B of radius 1 with center $(2, 0, \ldots, 0)$. We have

$$\begin{aligned}\|f\|_{H^s} + \|g\|_{H^s} &\le C(\|f\|_{\dot{H}^s} + \|g\|_{\dot{H}^s}) \\ &= CR^{n/2-s}(\|f^R\|_{\dot{H}^s} + \|g^R\|_{\dot{H}^s}) \\ &\le CR^{n/2-s}\|f^R + g^R\|_{\dot{H}^s} \\ &= C\|f + g\|_{\dot{H}^s} \le C\|f + g\|_{H^s},\end{aligned}$$

where we use that all functions above are supported in the closed ball K with radius 3 around $0 \in \mathbb{R}^n$ and in the second inequality we use Lemma 4.2. This finishes the proof. □

Lemma 4.4. *Let $s > n/2 + 1$ and $\varphi_0 \in \mathcal{D}^s(\mathbb{R}^n)$. Then there is a constant $C > 0$ and a neighborhood $W \subset \mathcal{D}^s(\mathbb{R}^n)$ of φ_0 s.t.*

$$\frac{1}{C}\|f \circ \varphi^{-1}\|_{H^s} \le \|f\|_{H^s} \le C\|f \circ \varphi^{-1}\|_{H^s}$$

for all $f \in H^s(\mathbb{R}^n; \mathbb{R}^n)$ and $\varphi \in W$.

Proof. By the continuity of the composition and inversion in $\mathcal{D}^s(\mathbb{R}^n)$ there is a neighborhood $W_1 \subset \mathcal{D}^s(\mathbb{R}^n)$ of φ_0 and $\delta > 0$ s.t.

$$\|f \circ \varphi^{-1}\|_{H^s} \le 1$$

for all $f \in H^s(\mathbb{R}^n; \mathbb{R}^n)$ with $\|f\|_{H^s} \le \delta$ and $\varphi \in W_1$. By linearity in f we get

$$\|f \circ \varphi^{-1}\|_{H^s} \le \frac{1}{\delta}\|f\|_{H^s}$$

for all $f \in H^s(\mathbb{R}^n; \mathbb{R}^n)$ and $\varphi \in W_1$. This shows the first inequality in the statement of the lemma. The same argument gives a neighborhood $W_2 \subset \mathcal{D}^s(\mathbb{R}^n)$ of φ_0 and a constant $K > 0$ s.t.

$$\|f \circ \varphi\|_{H^s} \le K\|f\|_{H^s}$$

for all $f \in H^s(\mathbb{R}^n; \mathbb{R}^n)$ and $\varphi \in W_2$. By replacing f by $f \circ \varphi^{-1}$ we get the second inequality. Taking $W = W_1 \cap W_2$ finishes the proof. □

Due to (5) Theorem 4.1 will follow from

Lemma 4.5. *The time $T = 1$ solution map*

$$\Phi : U \subset H^s(\mathbb{R}^n; \mathbb{R}^n) \to H^s(\mathbb{R}^n; \mathbb{R}^n), \; u_0 \mapsto \Phi(u_0) = u_0 \circ (id + u_0)^{-1},$$

is nowhere locally uniformly continuous.

Proof. Let $u_0 \in U \subset H^s(\mathbb{R}^n; \mathbb{R}^n)$ with compact support in $\mathbb{R}^n$, i.e. $\operatorname{supp} u_0 \subset \mathbb{R}^n$ is compact and we denote by $\varphi_0 = \mathrm{id} + u_0 \in \mathcal{D}^s(\mathbb{R}^n)$. For $\varphi \in \mathcal{D}^s(\mathbb{R}^n)$ we clearly have

$$\operatorname{supp} u_0 \circ \varphi^{-1} = \varphi(\operatorname{supp} u_0).$$

So by the Sobolev Imbedding Theorem $H^s(\mathbb{R}^n) \hookrightarrow C_0(\mathbb{R}^n)$ we can choose $R_1 > 0$ s.t. the ball $B_{R_1}(u_0) \subset H^s(\mathbb{R}^n; \mathbb{R}^n)$ of radius R_1 around u_0 in $H^s(\mathbb{R}^n; \mathbb{R}^n)$ is contained in U and we have

$$\operatorname{supp} u_0 \circ (\mathrm{id} + w)^{-1} \subset K_1 \subset \mathbb{R}^n, \; \forall w \in B_{R_1}(u_0),$$

for some fixed compact $K_1 \subset \mathbb{R}^n$. We choose $x^* \in \mathbb{R}^n$ s.t. the image of the unit ball $B_1(x^*) \subset \mathbb{R}^n$ of under φ_0 is 2 units away from K_1, i.e. we have

$$\operatorname{dist}(K_1, \varphi_0(B_2(x^*))) \ge 2.$$

Again by the Sobolev Imbedding Theorem we can choose $0 < R_2 \leq R_1$ s.t.

$$\varphi\left(B_1(x^*)\right) \subset K_2,\ \forall \varphi = \mathrm{id} + w \text{ with } w \in B_{R_2}(u_0),$$

for some fixed compact set $K_2 \subset \mathbb{R}^n$ with $\operatorname{dist}(K_1, K_2) \geq 2$. Using the Sobolev Imbedding Theorem $H^s(\mathbb{R}^n) \hookrightarrow C_0^1(\mathbb{R}^n)$ we choose $0 < R_3 \leq R_2$ s.t.

$$|\varphi(x) - \varphi(y)| \leq L|x-y|,\ \forall x, y \in \mathbb{R}^n,$$

for some constant $L \geq 1$ and for all $\varphi = \mathrm{id} + w \in \mathcal{D}^s(\mathbb{R}^n)$ with $w \in B_{R_3}(u_0)$. Finally we choose $0 < R^* \leq R_3$ s.t. the inequality of Lemma 4.4 holds for $W = \mathrm{id} + B_{R^*}(u_0)$ and some constant $C > 0$. Our strategy to prove the lemma is to construct for every $0 < R \leq R^*$ a pair of initial values $(u_0^{(k)})_{k\geq 1}, (\tilde{u}_0^{(k)})_{k \geq 1} \subset B_R(u_0)$ with

$$\lim_{k\to\infty} \|u_0^{(k)} - \tilde{u}_0^{(k)}\|_{H^s} = 0$$

whereas

$$\limsup_{k\to\infty} \|\Phi(u_0^{(k)}) - \Phi(\tilde{u}_0^{(k)})\|_{H^s} > 0.$$

This would implie that $\Phi|_{B_R(u_0)}$ is not uniformly continuous. Since $0 < R \leq R^*$ and $u_0 \in C_c(\mathbb{R}^n; \mathbb{R}^n) \cap U$ are arbitrary this would prove the lemma.
Now let $0 < R \leq R_*$. First we fix $g \in C_c^\infty(\mathbb{R}^n)$ with

$$\operatorname{supp} g \subset B_1(x^*),\ g(x^*) = 1,$$

and a unit vector $e \in \mathbb{R}^n$. We define $g_e = g \cdot e \in C_c^\infty(\mathbb{R}^n; \mathbb{R}^n)$ which will serve as translation in direction of e by one unit on x^*. For every $k \geq 1$ we define the following sequence of radii

$$r_k = \frac{1}{4kL} \leq 1,$$

where $L \geq 1$ is the Lipschitz constant from above. We further choose for every $k \geq 1$ a $w_k \in C_c^\infty(\mathbb{R}^n; \mathbb{R}^n)$ satisfying

$$\operatorname{supp} w_k \subset B_{r_k}(x^*) \text{ and } \|w_k\|_{H^s} = R/2.$$

Note that by construction we have

$$\operatorname{supp} w_k \circ \varphi^{-1} \subset B_{1/4k}(\varphi(x^*))$$

for all $\varphi = \mathrm{id} + w$ with $w \in B_R(u_0)$. With this we define for $k \geq 1$ the sequence of initial values

$$u_0^{(k)} = u_0 + w_k \text{ resp. } \tilde{u}_0^{(k)} = u_0 + w_k + \frac{1}{k} g_e.$$

There is $N \geq 1$ s.t. $u_0^{(k)}, \tilde{u}_0^{(k)} \in B_R(u_0)$ for $k \geq N$. We clearly have

$$\lim_{k\to\infty} \|u_0^{(k)} - \tilde{u}_0^{(k)}\|_{H^s} = 0.$$

We define for $k \geq N$

$$\varphi_k = \mathrm{id} + u_0^{(k)} \text{ resp. } \tilde{\varphi}_k = \mathrm{id} + \tilde{u}_0^{(k)}.$$

Thus for $k \geq N$

$$\Phi(u_0^{(k)}) = u_0 \circ \varphi_k^{-1} + w_k \circ \varphi_k^{-1},\ \Phi(\tilde{u}_0^{(k)}) = u_0 \circ \tilde{\varphi}_k^{-1} + w_k \circ \tilde{\varphi}_k^{-1} + \frac{1}{k} g_e \circ \tilde{\varphi}_k^{-1}.$$

Note that

$$\operatorname{supp} u_0 \circ \varphi_k^{-1},\ \operatorname{supp} u_0 \circ \tilde{\varphi}_k^{-1} \subset K_1$$

and

$$\operatorname{supp} w_k \circ \varphi_k^{-1},\ \operatorname{supp}\left(w_k \circ \tilde{\varphi}_k^{-1} + \frac{1}{k} g_e \circ \tilde{\varphi}_k^{-1}\right) \subset K_2.$$

Thus from Lemma 4.2 we conclude

$$\limsup_{k\to\infty} \|\Phi(u_0^{(k)}) - \Phi(\tilde{u}_0^{(k)})\|_{H^s} \geq C \limsup_{k\to\infty} \|w_k \circ \varphi_k^{-1} - (w_k \circ \tilde{\varphi}_k^{-1} + \frac{1}{k} g_e \circ \tilde{\varphi}_k^{-1})\|_{H^s}$$

for some $C > 0$. For the latter we get by the reverse triangle inequality and Lemma 4.4

$$\begin{aligned}
&\limsup_{k\to\infty} \|w_k \circ \varphi_k^{-1} - (w_k \circ \tilde{\varphi}_k^{-1} + \frac{1}{k} g_e \circ \tilde{\varphi}_k^{-1})\|_{H^s} \\
&\geq \limsup_{k\to\infty} \left(\|w_k \circ \varphi_k^{-1} - w_k \circ \tilde{\varphi}_k^{-1}\|_{H^s} - \|\frac{1}{k} g_e \circ \tilde{\varphi}_k^{-1}\|_{H^s} \right) \\
&\geq \limsup_{k\to\infty} \|w_k \circ \varphi_k^{-1} - w_k \circ \tilde{\varphi}_k^{-1}\|_{H^s}.
\end{aligned}$$

We have by construction $|\varphi_k(x^*) - \tilde{\varphi}_k(x^*)| = 1/k$ and

$$\operatorname{supp} w_k \circ \varphi_k^{-1} \subset B_{1/4k}(\varphi_k(x^*)),\ \operatorname{supp} w_k \circ \tilde{\varphi}_k^{-1} \subset B_{1/4k}(\tilde{\varphi}_k(x^*)).$$

By applying Lemma 4.3 and Lemma 4.4 we get

$$\begin{aligned}
\limsup_{k\to\infty} \|w_k \circ \varphi_k^{-1} - w_k \circ \tilde{\varphi}_k^{-1}\|_{H^s} &\geq C \limsup_{k\to\infty} \left(\|w_k \circ \varphi_k^{-1}\|_{H^s} + \|w_k \circ \tilde{\varphi}_k^{-1}\|_{H^s}\right) \\
&\geq CR
\end{aligned}$$

for some $C > 0$. With this we finally we conclude

$$\limsup_{k\to\infty} \|\Phi(u_0^{(k)}) - \Phi(\tilde{u}_0^{(k)})\|_{H^s} \geq CR$$

for some $C > 0$. This finishes the proof. □

Now we can prove Theorem 4.1

Proof of Theorem 4.1. The proof follows from (5) and Lemma 4.5. □

5. Blow up of classical solutions

The goal of this section is to discuss the existence of blow up of classical solutions to (1). So we assume throughout this section $s > n/2 + 1$. By the explicit solution formula to (1) in Lagrangian coordinates

$$\varphi : t \mapsto \mathrm{id} + t u_0,$$

and the fact that the solution to (1) exists as long $\varphi(t)$ stays in $\mathcal{D}^s(\mathbb{R}^n)$ we can work out "simple" conditions for blow up. For the 1D situation we can explicitly state the blow up time.

Theorem 5.1. *Let $n = 1$ and $u_0 \in H^s(\mathbb{R}) \setminus \{0\}$. Then the solution to (1) blows up in finite time. Moreover the blow up time is*

$$T^*(u_0) = -\frac{1}{\inf_{x\in\mathbb{R}} u_0'(x)} < \infty.$$

Proof. Note that we have the Sobolev Imbedding $H^s(\mathbb{R}) \hookrightarrow C_0^1(\mathbb{R})$. For a nonzero $u_0 \in H^s(\mathbb{R})$ we cannot have $u_0'(x) \geq 0$ for all $x \in \mathbb{R}$. Thus we have

$$\inf_{x\in\mathbb{R}} u_0'(x) = \min_{x\in\mathbb{R}} u_0'(x) < 0.$$

The curve $\varphi : t \mapsto \mathrm{id} + t u_0$ stays in $\mathcal{D}^s(\mathbb{R})$ as long as $\frac{\partial}{\partial x}\varphi(t) = 1 + t u_0'(x) > 0$ for all $x \in \mathbb{R}$. This is violated for the first time for

$$t = T^*(u_0) = -\frac{1}{\inf_{x\in\mathbb{R}} u_0'(x)}.$$

This finishes the proof. □

For the general case we have

Theorem 5.2. *Let $n \geq 2$. Then there is $u_0 \in H^s(\mathbb{R}^n; \mathbb{R}^n)$ s.t. the solution to (1) blows up in finite time.*

Proof. Let $\phi \in C_c^\infty(\mathbb{R}^n)$ with $\phi \neq 0$. We cannot have $\partial_1 \phi(x) \geq 0$ for all $x \in \mathbb{R}^n$. So there is $x^* \in \mathbb{R}^n$ with $\partial_1 \phi(x^*) < 0$. We take

$$u_0 = (\phi, 0, \ldots, 0).$$

The solution in Lagrangian coordinates is $t \mapsto \varphi(t) = \mathrm{id} + t u_0$. Its Jacobian determinant is given by

$$\det(d_x \varphi(t)) = 1 + t \partial_1 \phi(x).$$

Since $\partial_1 \phi(x^*) < 0$ we cannot have $\det(d_x \varphi(t)) > 0$ for all $x \in \mathbb{R}^n$ and $t \geq 0$. Thus the solution with initial value u_0 blows up in finite time. $\square$

Remark 3. *For $n \geq 2$ we could not figure out whether all solutions do blow up or whether there are global solutions.*

References

[1] Bahouri, H.; Chemin, J.-Y.; Danchin, R., *Fourier analysis and nonlinear partial differential equations. Grundlehren der mathematischen Wissenschaften*, 343. Springer, Heidelberg, 2011. xvi+523 pp.

[2] Inci, H.; Kappeler, T.; Topalov, P, *On the regularity of the composition of diffeomorphisms*, Mem. Amer., Math. Soc. 226 (2013), no. 1062.

[3] Inci, H., *On a Lagrangian formulation of the incompressible Euler equation*, J. Partial Differ. Equ. 29 (2016), no. 4, 320–359.

[4] McLean, W., *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge, 2000.

[5] Kato, T., *The Cauchy problem for quasi-linear symmetric hyperbolic systems.*, Arch. Rational Mech. Anal. 58 (1975), no. 3, 181205.

Department of Mathematics, Koç University, İstanbul, Türkiye

Email(s): hinci@ku.edu.tr

Bézier-Baskakov-Durrmeyer Type Operators

Arun Kajla

ABSTRACT. In this study, we construct the Bézier variant of the generalized Baskakov-Durrmeyer type operators. We give a direct approximation theorem in terms of the Ditzian-Totik modulus of smoothness $\omega_{\varphi^\eta}(\zeta, s)(0 \leq \eta \leq 1)$ and the rate of convergence for functions having derivatives of bounded variation.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 26A15, 40A35, 41A25, 41A36

KEYWORDS: Order of approximation, Baskakov operator.

1. INTRODUCTION

In [15], Mihesan considered the following generalized Baskakov operators with a non-negative constant $a \geq 0$, independent of α :

$$(1) \qquad B_\alpha^a(\zeta; \varkappa) = \sum_{\tau=0}^{\infty} W_{\alpha,\tau}^a(\varkappa)\zeta\left(\frac{\tau}{\alpha}\right),$$

where $W_{\alpha,\tau}^a(\varkappa) = e^{-\frac{a\varkappa}{1+\varkappa}} \frac{p_\tau(\alpha, a)}{\tau!} \frac{\varkappa^\tau}{(1+\varkappa)^{\alpha+\tau}}$ such that $\sum_{\tau=0}^{\infty} W_{\alpha,\tau}^a(\varkappa) = 1$ and

$$p_\tau(\alpha, a) = \sum_{i=0}^{\tau} \binom{\tau}{i} (\alpha)_i a^{\tau-i},$$

being the rising factorial given by $(\alpha)_0 = 1$, $(\alpha)_i = \alpha(\alpha+1)\cdots(\alpha+i-1)$, for $i \geq 1$. Wafi and Khatoon [18] gave the rate of convergence, Voronovskaja type theorem and a direct estimate by means of the Ditzian-Totik modulus of smoothness of the operators (1). Erencin and Başcanbaz-Tunca [8] defined a generalization of generalized Baskakov operators and determined the order of approximation by means the usual modulus of smoothness, weighted approximation theorem and elements of Lipschitz class.

Erencin [7] presented the following Durrmeyer type generalization of the operators (1):

$$(2) \qquad \mathcal{G}_\alpha^a(\zeta; \varkappa) = \sum_{\tau=0}^{\infty} W_{\alpha,\tau}^a(\varkappa) \frac{1}{B(\tau+1, \alpha)} \int_0^\infty \frac{s^\tau}{(1+s)^{\alpha+\tau+1}} \zeta(s) ds,$$

where $B(\tau + 1, \alpha)$ is the beta function defined by

$$B(\varkappa, y) = \int_0^\infty \frac{s^{\varkappa-1}}{(1+s)^{\varkappa+y}} ds = \frac{\Gamma(\varkappa)\Gamma(y)}{\Gamma(\varkappa+y)}, \quad \varkappa, y > 0.$$

Erencin [7] studied some direct approximation theorems for these operators, e.g. local approximation, rate of convergence for a Lipschitz type space and weighted approximation theorem. Agrawal et al. [2] discussed simultaneous approximation and rate of approximation of functions having derivatives of bounded variation. Zeng and Chen [20] defined the Bézier variant of Bernstein-Durrmeyer operators and obtained the direct approximation theorem for functions of bounded variation. Bézier type operators were studied by several researchers (cf. [1, 3, 5, 11, 12, 17, 19, 22, 16, 10]).

Let $\mathcal{C}$ denote the class of all Lebesgue measurable functions ζ on $[0, \infty)$ such that

$$\mathcal{C} = \left\{\zeta : \int_0^\infty \frac{|\zeta(s)|}{(1+s)^m} ds < \infty \text{ ,for some positive integer } m\right\}.$$

Clearly, the class $\mathcal{C}$ is bigger than the space $C_B[0, \infty)$.

For $\rho \geq 1$, we now consider the Bézier variant $\mathcal{G}^a_{\alpha,\rho}$ of the operators $\mathcal{G}^a_\alpha$ as follows:

$$\mathcal{G}^a_{\alpha,\rho}(\zeta;\varkappa) = \sum_{\tau=0}^{\infty} \mathcal{P}^{(\rho)}_{\alpha,\tau,a}(\varkappa) \frac{1}{B(\tau+1,\alpha)} \int_0^\infty \frac{s^\tau}{(1+s)^{\alpha+\tau+1}} \zeta(s)ds, \tag{3}$$

where $\mathcal{P}^{(\rho)}_{\alpha,\tau,a}(\varkappa) = \left[\mathcal{S}^a_{\alpha,\tau}(\varkappa)\right]^\rho - \left[\mathcal{S}^a_{\alpha,\tau+1}(\varkappa)\right]^\rho$ with $\mathcal{S}^a_{\alpha,\tau}(\varkappa) = \sum_{v=\tau}^{\infty} W^a_{\alpha,v}(\varkappa)$.
Alternatively we may rewrite the operators (3) as

$$\mathcal{G}^a_{\alpha,\rho}(\zeta;\varkappa) = \int_0^\infty \mathcal{M}_{\alpha,\rho,a}(\varkappa,s)\zeta(s)ds, \ \ \varkappa \in [0,\infty), \tag{4}$$

where

$$\mathcal{M}_{\alpha,\rho,a}(\varkappa,s) = \sum_{\tau=0}^{\infty} \mathcal{P}^{(\rho)}_{\alpha,\tau,a}(\varkappa) \frac{1}{B(\tau+1,\alpha)} \frac{s^\tau}{(1+s)^{\alpha+\tau+1}}.$$

For $\rho = 1$, $\mathcal{G}^a_{\alpha,\rho}$ reduce to $\mathcal{G}^a_\alpha$.

Lemma 1.1. [2] *For the operators $\mathcal{G}^a_\alpha$ we have*

(i) $\mathcal{G}^a_\alpha(s;\varkappa) = \dfrac{1}{(\alpha-1)}\left(\alpha\varkappa + \dfrac{a\varkappa}{(1+\varkappa)} + 1\right), \ \ \alpha > 1,$

(ii) $\mathcal{G}^a_\alpha(s^2;\varkappa) = \dfrac{1}{(\alpha-1)(\alpha-2)}\left((\alpha^2+\alpha)\varkappa^2 + 4\alpha\varkappa + \dfrac{a^2\varkappa^2}{(1+\varkappa^2} + \dfrac{2\alpha a\varkappa)^2}{(1+\varkappa)} + \dfrac{4a\varkappa}{(1+\varkappa)} + 2\right), \ \ \alpha > 2,$

(iii) For each $\varkappa \in (0,\infty)$ and $m \in \mathbb{N}$, $T^a_{\alpha,m}(\varkappa) = a_m(\alpha)\varkappa^m + \alpha^{-1}(p_m(\varkappa,a) + o(1))$,

where $a_m(\alpha) = \dfrac{\prod_{j=0}^{m-1}(\alpha+j)}{\prod_{j=1}^{m}(\alpha-j)}$ and $p_m(\varkappa,a)$ is a rational function of $\varkappa$ depending on the parameters a and m.

Lemma 1.2. [2] *For $m \in \mathbb{N}^0$, the m^{th} order central moment for the operators (2) is defined as*

$$\mu^a_{\alpha,m}(\varkappa) := \mathcal{G}^a_\alpha((s-\varkappa)^m;\varkappa) = \sum_{\tau=0}^{\infty} \frac{W^a_{\alpha,\tau}(\varkappa)}{B(\tau+1,\alpha)} \int_0^\infty \frac{s^\tau}{(1+s)^{\alpha+\tau+1}}(s-\varkappa)^m dt, \ \ \alpha > m.$$

For the function $\mu^a_{\alpha,m}(\varkappa)$, we have $\mu^a_{\alpha,0}(\varkappa) = 1$, and $\mu^a_{\alpha,1}(\varkappa) = \dfrac{(1+\varkappa)}{(\alpha-1)} + \dfrac{a\varkappa}{(1+\varkappa)(\alpha-1)}$ and there holds the following recurrence relation, for all $m \geq 1$:

$$\begin{aligned} \varkappa(1+\varkappa)^2(\mu^a_{\alpha,m}(\varkappa))' &= (\alpha-m-1)(1+\varkappa)\mu^a_{\alpha,m+1}(\varkappa) \\ &\quad - \left\{\Big((2m+1)\varkappa + (m+1)\Big)(1+\varkappa) + a\varkappa\right\}\mu^a_{\alpha,m}(\varkappa) \\ &\quad - 2mx(1+\varkappa)^2\mu^a_{\alpha,m-1}(\varkappa), \ \ \alpha > m+1. \end{aligned}$$

Corollary 1.3. *([2],p. 199) For $\lambda > 2$ and α sufficiently large, we have*

$$\mu^a_{\alpha,2}(\varkappa) \leq \frac{\lambda\varphi^2(\varkappa)}{\alpha}, \tag{5}$$

where $\varphi(\varkappa) = \sqrt{\varkappa(1+\varkappa)}$.

2. Direct Theorem

From [21], we have

$$\begin{aligned} &1 = \mathcal{S}^a_{\alpha,0}(\varkappa) > \mathcal{S}^a_{\alpha,1}(\varkappa) > \cdots > \mathcal{S}^a_{\alpha,\tau}(\varkappa) > \mathcal{S}^a_{\alpha,\tau+1}(\varkappa) > \cdots \\ &0 < \left[\mathcal{S}^a_{\alpha,\tau}(\varkappa)\right]^\rho - \left[\mathcal{S}^a_{\alpha,\tau+1}(\varkappa)\right]^\rho \leq \rho W^a_{\alpha,\tau}(\varkappa), \ \ \rho \geq 1. \end{aligned} \tag{6}$$

To describe our results, we recall the definitions of the Ditizian-Totik modulus $\omega_{\varphi^\eta}(\zeta, s)$ [6]. Let $\varphi(\varkappa) = \sqrt{\varkappa(1+\varkappa)}, 0 \leq \eta \leq 1$,

$$\omega_{\varphi^\eta}(\zeta, s) = \sup_{0<h\leq s} \sup_{\varkappa \pm h\varphi^\eta(\varkappa)/2 \geq 0} \left\{ \left| \zeta\left(\varkappa + \frac{h\varphi^\eta(\varkappa)}{2}\right) - \zeta\left(\varkappa - \frac{h\varphi^\eta(\varkappa)}{2}\right) \right| \right\},$$

and the appropriate Petree's K-functional is defined by

$$\overline{K}_{\varphi^\eta}(\zeta, s) = \inf_{g \in W_\eta} \{||\zeta - g|| + s||\varphi^\eta g'||\},\ s > 0,$$

where $W_\eta = \{g : g \in AC_{loc}, ||\varphi^\eta g'|| < \infty\}$. It is well known that ([6], Theorem 3.1.2) $\overline{K}_{\varphi^\eta}(\zeta, s) \sim \omega_{\varphi^\eta}(\zeta, s)$ which means that there exists a constant $M > 0$ such that

$$M^{-1}\omega_{\varphi^\eta}(\zeta, s) \leq \overline{K}_{\varphi^\eta}(\zeta, s) \leq M\omega_{\varphi^\eta}(\zeta, s). \tag{7}$$

Lemma 2.1. *For $\zeta \in W_\eta, \varphi(\varkappa) = \sqrt{\varkappa(1+\varkappa)}, 0 \leq \eta \leq 1$ and $s, \varkappa > 0$, we have*

$$\left| \int_\varkappa^t \zeta'(u)du \right| \leq 2^\eta \left(\varkappa^{-\eta/2}(1+s)^{-\eta/2} + \varphi^{-\eta}(\varkappa) \right) |s - \varkappa| \left\| \varphi^\eta \zeta' \right\|.$$

Proof. On an application of Hölder's inequality, we may write

$$\begin{aligned} \left| \int_\varkappa^t \zeta'(u)du \right| &\leq \|\varphi^\eta \zeta'\| \left| \int_\varkappa^s \frac{du}{\varphi^\eta(u)} \right| \\ &\leq \left\| \varphi^\eta \zeta' \right\| |s - \varkappa|^{1-\eta} \left| \int_\varkappa^s \frac{du}{\varphi(u)} \right|^\eta. \end{aligned}$$

Since,

$$\left| \int_\varkappa^s \frac{du}{\varphi(u)} \right| \leq \left| \int_\varkappa^s \frac{du}{\sqrt{u}} \right| \left(\frac{1}{\sqrt{1+\varkappa}} + \frac{1}{\sqrt{1+s}} \right)$$

and

$$\left| \int_\varkappa^s \frac{du}{\sqrt{u}} \right| \leq \frac{2|s-\varkappa|}{\sqrt{\varkappa}},$$

we get

$$\begin{aligned} \left| \int_\varkappa^t \zeta'(u)du \right| &\leq \left\| \varphi^\eta \zeta' \right\| |s - \varkappa| \frac{2^\eta}{\varkappa^{\eta/2}} \left| \frac{1}{\sqrt{1+\varkappa}} + \frac{1}{\sqrt{1+s}} \right|^\eta \\ &\leq \left\| \varphi^\eta \zeta' \right\| |s - \varkappa| \frac{2^\eta}{\varkappa^{\eta/2}} \left((1+s)^{-\eta/2} + (1+\varkappa)^{-\eta/2} \right), \end{aligned}$$

on using the inequality $|a_1 + a_2|^r \leq |a_1|^r + |a_2|^r, 0 \leq r \leq 1$. ☐

Lemma 2.2. *We have*

$$\mathcal{G}^a_{\alpha,\rho}((1+s)^{-m}; \varkappa) \leq C_m(1+\varkappa)^{-m}, \tag{8}$$

where C_m is a constant dependent on m.

Proof. For each $\varkappa \in [0, \infty)$, for $\varkappa = 0$, the result holds from (2).

From (6), we have

$$\begin{aligned} \mathcal{G}^a_\alpha((1+s)^{-m}; \varkappa) &= \sum_{\tau=0}^\infty \frac{\mathcal{P}^{(\rho)}_{\alpha,\tau}(\varkappa)}{B(\tau+1, \alpha)} \int_0^\infty \frac{s^\tau}{(1+s)^{\alpha+\tau+m+1}} ds \\ &= \sum_{\tau=0}^\infty \frac{\mathcal{P}^{(\rho)}_{\alpha,\tau}(\varkappa)\Gamma(\alpha+\tau+1)\Gamma(\alpha+m)}{\Gamma(\alpha+m+\tau+1)\Gamma(\alpha)} \\ &\leq \rho(1+\varkappa)^{-m} \sum_{\tau=0}^\infty e^{-\frac{a\varkappa}{1+\varkappa}} \frac{P_\tau(\alpha, a)}{\tau!} \frac{\varkappa^\tau}{(1+\varkappa)^{\alpha+\tau-m}} \frac{\Gamma(\alpha+\tau+1)\Gamma(\alpha+m)}{\Gamma(\alpha+m+\tau+1)\Gamma(\alpha)}. \end{aligned}$$

(9)

By ratio test, the series on the right hand side (9) is convergent. This proves the required result. ☐

Guo et al. [10] obtained the direct, inverse and equivalence theorems for the Baskakov-Bézier operators by using means of the unified Ditzian-Totik modulus of smoothness $\omega_{\varphi^\eta}(\zeta, s)(0 \leq \eta \leq 1)$. Gairola and Agrawal [9] proved the direct and inverse theorem for the Bézier variant of the operators introduced by Gupta and Mohapatra in [13] in terms of Ditzian-Totik modulus of smoothness.
In our next theorem, we obtain the direct approximation theorems for the Bézier-variant of the generalized Baskakov-Durrmeyer operators in terms of the Ditzian-Totik modulus of smoothness.

Theorem 2.3. *For $\zeta \in C_B[0,\infty)$, we have*

$$\left|\mathcal{G}^a_{\alpha,\rho}(\zeta;\varkappa) - \zeta(\varkappa)\right| \leq C\omega_{\varphi^\eta}\left(\zeta, \frac{\varphi^{1-\eta}(\varkappa)}{\sqrt{\alpha}}\right). \tag{10}$$

Proof. By the definition of $\overline{K}_{\varphi^\eta}(\zeta, s)$, for fixed $\alpha, \varkappa, \eta$ we can choose $g = g_{\alpha,\varkappa,\eta} \in W_\eta$ such that

$$||\zeta - g|| + \frac{\varphi^{1-\eta}(\varkappa)}{\sqrt{\alpha}}||\varphi^\eta g'|| \leq \overline{K}_{\varphi^\eta}\left(\zeta, \frac{\varphi^{1-\eta}(\varkappa)}{\sqrt{\alpha}}\right). \tag{11}$$

Since $\mathcal{G}^a_{\alpha,\rho}(1;\varkappa) = 1$, we may write

$$\left|\mathcal{G}^a_{\alpha,\rho}(\zeta;\varkappa) - \zeta(\varkappa)\right| \leq 2||\zeta - g|| + |\mathcal{G}^a_{\alpha,\rho}(g;\varkappa) - g(\varkappa)|. \tag{12}$$

Using the representation $g(s) = g(\varkappa) + \int_\varkappa^s g'(u)du$ and Lemma 2.1, we get

$$\begin{aligned} \left|\mathcal{G}^a_{\alpha,\rho}(g;\varkappa) - g(\varkappa)\right| &= \left|\mathcal{G}^a_{\alpha,\rho}\left(\int_\varkappa^t g'(u)du;\varkappa\right)\right| \\ &\leq 2^\eta||\varphi^\eta g'||\left\{\varphi^{-\eta}(\varkappa)\mathcal{G}^a_{\alpha,\rho}(|s-\varkappa|;\varkappa) + \varkappa^{-\eta/2}\mathcal{G}^a_{\alpha,\rho}\left(\frac{|s-\varkappa|}{(1+s)^{\eta/2}};\varkappa\right)\right\}. \end{aligned} \tag{13}$$

From (6), Corollary 1.3 and Cauchy-Schwarz inequality, we may write

$$\begin{aligned} \mathcal{G}^a_{\alpha,\rho}(|s-\varkappa|;\varkappa) &\leq \left(\mathcal{G}^a_{\alpha,\rho}((s-\varkappa)^2;\varkappa)\right)^{1/2} \\ &\leq \frac{\sqrt{\rho\lambda}\varphi(\varkappa)}{\sqrt{\alpha}}. \end{aligned} \tag{14}$$

Similarly using Lemma 2.2, we get

$$\begin{aligned} \mathcal{G}^a_{\alpha,\rho}\left(\frac{|s-\varkappa|}{(1+s)^{\eta/2}};\varkappa\right) &\leq \rho\mathcal{G}^a_{\alpha}\left(\frac{|s-\varkappa|}{(1+s)^{\eta/2}};\varkappa\right) \\ &\leq \rho\left(\mathcal{G}^a_{\alpha}((s-\varkappa)^2;\varkappa)\right)^{1/2}\left(\mathcal{G}^a_{\alpha}((1+s)^{-\eta};\varkappa)\right)^{1/2} \\ &\leq C_1\frac{\sqrt{\rho\lambda}\varphi(\varkappa)}{\sqrt{\alpha}}(1+\varkappa)^{-\eta/2}. \end{aligned} \tag{15}$$

From (13)-(15), we get

$$\left|\mathcal{G}^a_{\alpha,\rho}(g;\varkappa) - g(\varkappa)\right| \leq C||\varphi^\eta g'||\frac{\varphi^{1-\eta}(\varkappa)}{\sqrt{\alpha}}. \tag{16}$$

Using $\overline{K}_{\varphi^\eta}(\zeta, s) \sim \omega_{\varphi^\eta}(\zeta, s)$, (11), (12) and (16), we obtain (10). □

3. Rate of Convergence

Bojanic and Cheng [4] studied the rate of convergence of Bernstein polynomials for functions with derivatives of bounded variation. Karsli [14] obtained the rate of convergence of the Gamma type operators for functions with derivatives of bounded variation. In our next result, we establish the rate of convergence for functions with derivatives of bounded variation.
Let $\zeta \in DBV_\gamma(0,\infty)$, $\gamma \geq 0$, be the class of all functions defined on $(0,\infty)$, having a derivative of bounded variation on every finite subinterval of $(0,\infty)$ and $|\zeta(s)| \leq Ms^\gamma$, $\forall \ s > 0$.
Since $\zeta \in DBV_\gamma(0,\infty)$, we may write

$$\zeta(\varkappa) = \int_0^\varkappa g(s)ds + \zeta(0),$$

where $g(s)$ is a function of bounded variation on each finite subinterval of $(0,\infty)$.

Lemma 3.1. *Let $\varkappa\in(0,\infty)$, then for $\rho\geq 1,\lambda>2$ and sufficiently large α, we have*

(i) $\beta_{\alpha,\rho,a}(\varkappa,y)=\int_0^y \mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\leq \frac{\rho\lambda}{\alpha}\frac{\varphi^2(\varkappa)}{(\varkappa-y)^2}\,,\ 0\leq y<\varkappa,$

(ii) $1-\beta_{\alpha,\rho,a}(\varkappa,z)=\int_z^\infty \mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\leq \frac{\rho\lambda}{\alpha}\frac{\varphi^2(\varkappa)}{(z-\varkappa)^2},\ \varkappa<z<\infty.$

Proof. (i) From (6) and Corollary 1.3, we have

$$\begin{aligned}\beta_{\alpha,\rho,a}(\varkappa,y) &= \int_0^y \mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\leq \int_0^y\left(\frac{\varkappa-s}{\varkappa-y}\right)^2\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\\ &\leq \mathcal{G}^a_{\alpha,\rho}((s-\varkappa)^2;\varkappa)\ (\varkappa-y)^{-2}\leq\rho\mathcal{G}^a_{\alpha}((s-\varkappa)^2;\varkappa)(\varkappa-y)^{-2}\\ &\leq \rho\frac{\lambda}{\alpha}\frac{\varphi^2(\varkappa)}{(\varkappa-y)^2}.\end{aligned}$$

The proof of (ii) is similar hence we omit it.

Theorem 3.2. *Let $\zeta\in DBV_\gamma(0,\infty),\rho\geq 1$ and let $\upsilon_a^b(\zeta'_\varkappa)$ be the total variation of $\zeta'_\varkappa$ on $[a,b]\subset(0,\infty)$. Then, for every $\varkappa\in(0,\infty)$ and sufficiently large α, we have*

$$\begin{aligned}\left|\mathcal{G}^a_{\alpha,\rho}(\zeta;\varkappa)-\zeta(\varkappa)\right| &\leq \frac{\sqrt{\rho}}{\rho+1}\left|\zeta'(\varkappa+)+\rho\zeta'(\varkappa-)\right|\sqrt{\frac{\lambda}{\alpha}}\varphi(\varkappa)\\ &\quad+\sqrt{\frac{\lambda}{\alpha}}\ \varphi(\varkappa)\frac{\rho^{3/2}}{\rho+1}\left|\zeta'(\varkappa+)-\zeta'(\varkappa-)\right|\\ &\quad+\frac{\rho\lambda(1+\varkappa)}{\alpha}\sum_{\tau=1}^{[\sqrt{\alpha}\,]}\upsilon^{\varkappa}_{\varkappa-\varkappa/\tau}(\zeta'_\varkappa)+\frac{\varkappa}{\sqrt{\alpha}}\upsilon^{\varkappa}_{\varkappa-\varkappa/\sqrt{\alpha}}(\zeta'_\varkappa)\\ &\quad+\frac{\rho\lambda(1+\varkappa)}{\alpha}\sum_{\tau=1}^{[\sqrt{\alpha}\,]}\upsilon_{\varkappa}^{\varkappa+\varkappa/\tau}(\zeta'_\varkappa)+\frac{\varkappa}{\sqrt{\alpha}}\upsilon_{\varkappa}^{\varkappa+\varkappa/\sqrt{\alpha}}(\zeta'_\varkappa),\end{aligned}$$

where $\lambda>2$, and the auxiliary function $\zeta'_\varkappa$ is defined by

$$\zeta'_\varkappa(s)=\begin{cases}\zeta'(s)-\zeta'(\varkappa-), & 0\leq s<\varkappa\\ 0, & s=\varkappa\\ \zeta'(s)-\zeta'(\varkappa+) & \varkappa<s\leq 1\end{cases}.$$

Proof. Since $\int_0^\infty \mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds=\mathcal{G}^a_{\alpha,\rho}(1;\varkappa)=1$, we get

$$\begin{aligned}\mathcal{G}^a_{\alpha,\rho}(\zeta;\varkappa)-\zeta(\varkappa) &= \int_0^\infty[\zeta(s)-\zeta(\varkappa)]\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\\ &= \int_0^\infty\left(\int_\varkappa^s\zeta'(u)du\right)\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds.\end{aligned}\tag{17}$$

From the definition of the function $\zeta'_\varkappa$, for any $\zeta\in DBV_\gamma(0,\infty)$, we may write

$$\begin{aligned}\zeta'(s) &= \frac{1}{\rho+1}\Big(\zeta'(\varkappa+)+\rho\zeta'(\varkappa-)\Big)+\zeta'_\varkappa(s)\\ &\quad+\frac{1}{2}\Big(\zeta'(\varkappa+)-\zeta'(\varkappa-)\Big)\left(sgn(s-\varkappa)+\frac{\rho-1}{\rho+1}\right)\\ &\quad+\delta_\varkappa(s)\left(\zeta'(\varkappa)-\frac{1}{2}\Big(\zeta'(\varkappa+)+\zeta'(\varkappa-)\Big)\right),\end{aligned}\tag{18}$$

where

$$\delta_\varkappa(s)=\begin{cases}1\,, & \varkappa=s\\ 0, & \varkappa\neq s\end{cases}.$$

It is clear that

$$\int_0^\infty \mathcal{M}_{\alpha,\rho,a}(\varkappa,s)\int_{\varkappa}^{s}\left[\zeta'(\varkappa)-\frac{1}{2}\Big(\zeta'(\varkappa+)+\zeta'(\varkappa-)\Big)\right]\delta_\varkappa(u)dudt=0.$$

By (4) and simple computations, we have

$$\begin{aligned}
E_1 &= \int_0^\infty \left(\int_\varkappa^s \frac{1}{\rho+1}\Big(\zeta'(\varkappa+)+\rho\zeta'(\varkappa-)\Big)du\right)\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\\
&= \frac{1}{\rho+1}\Big|\zeta'(\varkappa+)+\rho\zeta'(\varkappa-)\Big|\int_0^\infty |s-\varkappa|\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\\
&\leq \frac{1}{\rho+1}\Big(\zeta'(\varkappa+)+\rho\zeta'(\varkappa-)\Big)\left(\mathcal{G}^a_{\alpha,\rho}((s-\varkappa)^2;\varkappa)\right)^{1/2}\\
&\leq \frac{\sqrt{\rho}}{\rho+1}\Big|\zeta'(\varkappa+)+\rho\zeta'(\varkappa-)\Big|\sqrt{\frac{\lambda}{\alpha}}\varphi(\varkappa) \qquad (19)
\end{aligned}$$

and

$$\begin{aligned}
E_2 &= \int_0^\infty \left(\int_\varkappa^s \frac{1}{2}\Big(\zeta'(\varkappa+)-\zeta'(\varkappa-)\Big)\left(sgn(u-\varkappa)+\frac{\rho-1}{\rho+1}\right)du\right)\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\\
&\leq \frac{\rho}{\rho+1}\Big|\zeta'(\varkappa+)-\zeta'(\varkappa-)\Big|\int_0^\infty |s-\varkappa|\,\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\\
&= \frac{\rho}{\rho+1}\Big|\zeta'(\varkappa+)-\zeta'(\varkappa-)\Big|\mathcal{G}^a_{\alpha,\rho}\left(|s-\varkappa|\,;\varkappa\right)\\
&\leq \frac{\rho}{\rho+1}\Big|\zeta'(\varkappa+)-\zeta'(\varkappa-)\Big|\left(\mathcal{G}^a_{\alpha,\rho}\left((s-\varkappa)^2;\varkappa\right)\right)^{1/2}\\
&\leq \frac{\rho^{3/2}}{\rho+1}\Big|\zeta'(\varkappa+)-\zeta'(\varkappa-)\Big|\sqrt{\frac{\lambda}{\alpha}}\varphi(\varkappa), \qquad (20)
\end{aligned}$$

on an application of Cauchy-Schwarz inequality.

By using Lemma 1.2, Corollary 1.3 and considering (17)-(20) we obtain the following estimate

$$\begin{aligned}
\left|\mathcal{G}^a_{\alpha,\rho}(\zeta;\varkappa)-\zeta(\varkappa)\right| &\leq \left|A_{\alpha,\rho,a}(\zeta'_\varkappa,\varkappa)+B_{\alpha,\rho,a}(\zeta'_\varkappa,\varkappa)\right|\\
&\quad+\frac{\sqrt{\rho}}{\rho+1}\Big|\zeta'(\varkappa+)+\rho\zeta'(\varkappa-)\Big|\sqrt{\frac{\lambda}{\alpha}}\varphi(\varkappa)\\
&\quad+\frac{\rho^{3/2}}{\rho+1}\left|\zeta'(\varkappa+)-\zeta'(\varkappa-)\right|\sqrt{\frac{\lambda}{\alpha}}\;\varphi(\varkappa), \qquad (21)
\end{aligned}$$

where

$$A_{\alpha,\rho,a}(\zeta'_\varkappa,\varkappa)=\int_0^\varkappa\left(\int_\varkappa^s \zeta'_\varkappa(u)du\right)\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds,$$

$$B_{\alpha,\rho,a}(\zeta'_\varkappa,\varkappa)=\int_\varkappa^\infty\left(\int_\varkappa^s \zeta'_\varkappa(u)du\right)\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds.$$

To complete the proof , it remains to estimate the terms $A_{\alpha,\rho,a}(\zeta'_\varkappa,\varkappa)$ and $B_{\alpha,\rho,a}(\zeta'_\varkappa,\varkappa)$. Since $\int_a^b d_s\beta_{\alpha,\rho,a}(\varkappa,s)\leq 1$, for all $[a,b]\subseteq(0,\infty)$, using integration by parts and applying Lemma

3.1 with $y = \varkappa - (\varkappa/\sqrt{\alpha})$, we have

$$\begin{aligned}
\left|A_{\alpha,\rho,a}(\zeta'_{\varkappa},\varkappa)\right| &= \left|\int_0^{\varkappa}\left(\int_{\varkappa}^{s}\zeta'_{\varkappa}(u)du\right)d_s\beta_{\alpha,\rho,a}(\varkappa,s)\right| \\
&= \left|\int_0^{\varkappa}\beta_{\alpha,\rho,a}(\varkappa,s)\zeta'_{\varkappa}(s)ds\right| \\
&\leq \left(\int_0^{y}+\int_y^{\varkappa}\right)\left|\zeta'_{\varkappa}(s)\right|\left|\beta_{\alpha,\rho,a}(\varkappa,s)\right|ds \\
&\leq \rho\frac{\lambda\varphi^2(\varkappa)}{\alpha}\int_0^{y}\upsilon_s^{\varkappa}(\zeta'_{\varkappa})(\varkappa-s)^{-2}ds+\int_y^{\varkappa}\upsilon_s^{\varkappa}(\zeta'_{\varkappa})ds \\
&\leq \rho\frac{\lambda\varphi^2(\varkappa)}{\alpha}\int_0^{y}\upsilon_s^{\varkappa}(\zeta'_{\varkappa})(\varkappa-s)^{-2}ds+\frac{\varkappa}{\sqrt{\alpha}}\upsilon^{\varkappa}_{\varkappa-\varkappa/\sqrt{\alpha}}(\zeta'_{\varkappa}).
\end{aligned}$$

Putting $u = \varkappa/(\varkappa - s)$, we obtain

$$\begin{aligned}
\rho\frac{\lambda\varphi^2(\varkappa)}{\alpha}\int_0^{\varkappa-\varkappa/\sqrt{\alpha}}(\varkappa-s)^{-2}\upsilon_s^{\varkappa}(\zeta'_{\varkappa})ds &= \rho\frac{\lambda(1+\varkappa)}{\alpha}\int_1^{\sqrt{\alpha}}\upsilon^{\varkappa}_{\varkappa-\varkappa/u}(\zeta'_{\varkappa})du \\
&\leq \rho\frac{\lambda(1+\varkappa)}{\alpha}\sum_{\tau=1}^{[\sqrt{\alpha}\,]}\int_{\tau}^{\tau+1}\upsilon^{\varkappa}_{\varkappa-\varkappa/u}(\zeta'_{\varkappa})du \\
&\leq \rho\frac{\lambda(1+\varkappa)}{\alpha}\sum_{\tau=1}^{[\sqrt{\alpha}\,]}\upsilon^{\varkappa}_{\varkappa-\varkappa/\tau}(\zeta'_{\varkappa}).
\end{aligned}$$

Hence we reach the following result

$$\left|A_{\alpha,\rho,a}(\zeta'_{\varkappa},\varkappa)\right| \leq \rho\frac{\lambda(1+\varkappa)}{\alpha}\sum_{\tau=1}^{[\sqrt{\alpha}\,]}\upsilon^{\varkappa}_{\varkappa-\varkappa/\tau}(\zeta'_{\varkappa})+\frac{\varkappa}{\sqrt{\alpha}}\upsilon^{\varkappa}_{\varkappa-\varkappa/\sqrt{\alpha}}(\zeta'_{\varkappa}). \tag{22}$$

Using integration by parts and applying Lemma 3.1 with $z = \varkappa + \varkappa/\sqrt{\alpha}$, we may write

$$\begin{aligned}
|B_{\alpha,\rho,a}(\zeta'_{\varkappa},\varkappa)| &= \left|\int_{\varkappa}^{\infty}\left(\int_{\varkappa}^{s}\zeta'_{\varkappa}(u)du\right)\mathcal{M}_{\alpha,\rho,a}(\varkappa,s)ds\right| \\
&= \left|\int_{\varkappa}^{z}\left(\int_{\varkappa}^{s}\zeta'_{\varkappa}(u)du\right)d_s(1-\beta_{\alpha,\rho,a}(\varkappa,s))+\int_{z}^{\infty}\left(\int_{\varkappa}^{s}\zeta'_{\varkappa}(u)du\right)d_s(1-\beta_{\alpha,\rho,a}(\varkappa,s))\right| \\
&= \left|\left[\left(\int_{\varkappa}^{t}f'_{\varkappa}(u)du\right)(1-\beta_{\alpha,\rho,a}(\varkappa,s))\right]_{\varkappa}^{z}-\int_{\varkappa}^{z}f'_{\varkappa}(s)(1-\beta_{\alpha,\rho,a}(\varkappa,s))ds\right. \\
&\quad\left.+\int_{z}^{\infty}\left(\int_{\varkappa}^{s}\zeta'_{\varkappa}(u)du\right)d_s(1-\beta_{\alpha,\rho,a}(\varkappa,s))\right| \\
&= \left|\left(\int_{\varkappa}^{z}f'_{\varkappa}(u)du\right)(1-\beta_{\alpha,\rho,a}(\varkappa,z))-\int_{\varkappa}^{z}\zeta'_{\varkappa}(s)(1-\beta_{\alpha,\rho,a}(\varkappa,s))ds\right. \\
&\quad\left.+\left[\left(\int_{\varkappa}^{t}f'_{\varkappa}(u)du\right)(1-\beta_{\alpha,\rho,a}(\varkappa,s))\right]_{z}^{\infty}-\int_{z}^{\infty}\zeta'_{\varkappa}(s)(1-\beta_{\alpha,\rho,a}(\varkappa,s))ds\right|
\end{aligned}$$

$$\begin{aligned}
&= \left|\int_{\varkappa}^{z}\zeta'_{\varkappa}(s)(1-\beta_{\alpha,\rho,a}(\varkappa,s))ds+\int_{z}^{\infty}\zeta'_{\varkappa}(s)(1-\beta_{\alpha,\rho,a}(\varkappa,s))ds\right| \\
&< \rho\frac{\lambda\varphi^2(\varkappa)}{\alpha}\int_{z}^{\infty}\upsilon_{\varkappa}^{s}(\zeta')_{\varkappa}(s-\varkappa)^{-2}ds+\int_{\varkappa}^{z}\upsilon_{\varkappa}^{s}(\zeta'_{\varkappa})ds \\
&\leq \rho\frac{\lambda\varphi^2(\varkappa)}{\alpha}\int_{\varkappa+\varkappa/\sqrt{\alpha}}^{\infty}\upsilon_{\varkappa}^{s}(\zeta'_{\varkappa})(s-\varkappa)^{-2}ds+\frac{\varkappa}{\sqrt{\alpha}}\upsilon_{\varkappa}^{\varkappa+\varkappa/\sqrt{\alpha}}(\zeta'_{\varkappa}).
\end{aligned} \tag{23}$$

Putting $u = \varkappa/(s-\varkappa)$, we get

$$\begin{aligned} \rho\frac{\lambda\varphi^2(\varkappa)}{\alpha}\int_{\varkappa+\varkappa/\sqrt{\alpha}}^{\infty} \upsilon_{\varkappa}^{s}(\zeta_{\varkappa}')(s-\varkappa)^{-2}ds &= \rho\frac{\lambda\varphi^2(\varkappa)}{\varkappa\alpha}\int_0^{\sqrt{\alpha}} \upsilon_{\varkappa}^{\varkappa+\varkappa/u}(\zeta_{\varkappa}')du \\ &\leq \rho\frac{\lambda(1+\varkappa)}{\alpha}\sum_{\tau=1}^{[\sqrt{\alpha}\,]}\int_{\tau}^{\tau+1} \upsilon_{\varkappa}^{\varkappa+\varkappa/u}(\zeta_{\varkappa}')du \\ &\leq \rho\frac{\lambda(1+\varkappa)}{\alpha}\sum_{\tau=1}^{[\sqrt{\alpha}\,]} \upsilon_{\varkappa}^{\varkappa+\varkappa/\tau}(\zeta_{\varkappa}') \qquad (24) \end{aligned}$$

Again combining (23)-(24), we get

$$|B_{\alpha,\rho,a}(\zeta_{\varkappa}',\varkappa)| \leq \rho\frac{\lambda(1+\varkappa)}{\alpha}\sum_{\tau=1}^{[\sqrt{\alpha}\,]} \upsilon_{\varkappa}^{\varkappa+\varkappa/\tau}(\zeta_{\varkappa}') + \frac{\varkappa}{\sqrt{\alpha}}\upsilon_{\varkappa}^{\varkappa+\varkappa/\sqrt{\alpha}}(\zeta_{\varkappa}'). \qquad (25)$$

Combining (21), (22) and (25), we get the required result. □

References

[1] U. Abel and V. Gupta, An estimate of the rate of convergence of a Bézier variant of the Baskaokov-Kantorovich operators for bounded variation functions, Demonstratio Math. 36 (1) (2003) 123-136.

[2] P. N. Agrawal, V. Gupta and A. S. Kumar, Generalized Baskakov-Durrmeyer type operators, Rend. Circ. Mat. Palermo, 63 (2014) 193-209.

[3] P. N. Agrawal, N. Ispir and A. Kajla, Approximation properties of Bézier-summation-integral type operators based on Polya-Bernstein functions, Appl. Math. Comput. 259 (2015) 533-539.

[4] R. Bojanic and F. H. Cheng, Rate of convergence of Bernstein polynomials for functions with derivatives of bounded variation, J. Math. Anal. Appl. 141 (1) (1989) 136-151.

[5] G. Chang, Generalized Bernstein-Bézier polynomilas, J. Comput. Math. 1 (4) (1983) 322-327.

[6] Z. Ditzian and V. Totik, Moduli of Smoothness, Springer, New York, (1987).

[7] A. Erencin, Durrmeyer type modification of generalized Baskakov operators, Appl. Math. Comput. 218 (3) (2011) 4384-4390.

[8] A. Erenin and G. Başcanbaz-Tunca, Approximation properties of a class of linear positive operators in weighted spaces, Cr. Acad. Bulg. Sci. 63 (10) (2010) 1397-1404.

[9] A. R. Gairola and P. N. Agrawal, Direct and inverse theorems for the Bézier variant of certain summation-integral type operators, Turk. J. Math. 34 (2010) 221-234.

[10] S. Guo, Q. Qi and G. Liu, The central theorems for Bsakakov-Bézier operators, J. Approx. Theory 147 (2007) 112-124.

[11] V. Gupta, On the Bézier variant of Kantorovich operators, Comput. Math. Anal. 47 (2004) 227-232.

[12] V. Gupta and R. P. Agarwal, Convergence Estimates in Approximation Theory, Springer, (2014).

[13] V. Gupta and R. N. Mohapatra, On the rate of convergence for certain summation-integral type operators, Math. Ineq. Appl. 9 (3) (2006) 465-472.

[14] H. Karsli, Rate of convergence of new Gamma type operators for functions with derivatives of bounded variation, Math. Comput. Modelling 45 (2007) 617-624.

[15] V. Mihesan, Uniform approximation with positive linear operators generated by generalized Baskakov method, Autom. Comput. Appl. Math. 7 (1) (1998) 34-37.

[16] H. M. Srivastava and V. Gupta, Rate of convergence for the Bézier variant of the Bleimann-Butzer-Hahn operators, Appl. Math. Lett. 18 (2005) 849-857.

[17] H. M. Srivastava and X. M. Zeng, Approximation by means of the Szász-Bézier integral operators, Internat. J. Pure Appl. Math. 14 (3) (2004) 283-294.

[18] A. Wafi and S. Khatoon, On the order of approximation of functions by generalized Baskakov operators. Indian J. Pure Appl. Math. 35 (3) (2004) 347-358.

[19] X. M. Zeng, On the rate of convergence of two Bernstein-Bézier type operators for bounded variation functions II, J. Approx. Theory 104 (2000) 330-344.

[20] X. M. Zeng and W. Chen, On the rate of convergence of the generalized Durrmeyer type operators for functions of bounded variation, J. Approx. Theory 102 (2000) 1-12.

[21] X. M. Zeng and V. Gupta, Rate of convergence of two Baskakov-Bézier type operators for locally bounded functions, Comput. Math. Appl. 44 (2002) 1445-1453.

[22] X. M. Zeng and A. Piriou, On the rate of convergence of two Bernstein-Bézier type operators for bounded variation functions, J. Approx. Theory 95 (1998) 369-387.

Department of Mathematics, School of Basic Sciences, Central University of Haryana, Haryana-123031, India

Email(s): arunkajla@cuh.ac.in

Unit Generalized Marshall-Olkin Weibull Distribution: Properties and Applications

Kadir Karakaya

ABSTRACT. In this study, a new unit distribution is introduced. The generalized Marshall-Olkin Weibull distribution is considered a baseline distribution. Some mathematical properties of the new model are discussed. Five estimation methods for the unknown parameters of the model are examined. Moreover, a Monte Carlo simulation study is conducted to evaluate the performance of the estimators. Two practical data sets are provided to demonstrate the efficiency of the new unit distribution.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 60E05, 62E15, 62F10

KEYWORDS: Marshall-Olkin family, Unit distribution, Monte Carlo simulation, Estimation.

1. INTRODUCTION

The beta distribution is the well-known distribution for modeling proportional or percentage data such as mortality rate, recovery rate, proportion of the educational measurements, etc. The Kumaraswamy distribution [8] is another important unit distribution that was proposed as an alternative to the beta distribution. Except for the beta and Kumaraswamy distributions, many new unit distributions are introduced. Some of these can be given as [4], [6], [7] and [11].

The generalized Marshall-Olkin Weibull (GMOW) distribution was introduced by [1]. If the random variable X has the GMOW distribution then the probability density function (pdf) and cumulative distribution function (cdf) are given, respectively, by,

$$g(x;\Xi)=\frac{\beta\theta x^{\beta-1}\left\{(1-\alpha)(1-\lambda)\exp\left(-\theta x^{\beta}\right)+(1+\alpha-\lambda)\exp\left(\theta x^{\beta}\right)+2\lambda-2\right\}}{\left(1-\exp\left(\theta x^{\beta}\right)-\alpha\right)^{2}},\ x>0 \tag{1}$$

and

$$G(x;\Xi)=\frac{\lambda\left(1-\exp\left(-\theta x^{\beta}\right)\right)+(1-\lambda)\left(1-\exp\left(-\theta x^{\beta}\right)\right)^{2}}{\alpha+(1-\alpha)\left(1-\exp\left(-\theta x^{\beta}\right)\right)}, \tag{2}$$

where $\theta>0$ is scale parameter, $\lambda,\alpha\in(0,1)$ and $\beta>0$ are shape parameters and $\Xi=(\theta,\beta,\lambda,\alpha)$ represents the parameter vector. Some distributional properties of the GMOW distribution were examined by [1] and some estimators were also studied for estimating the model parameters.

In this study, a new unit distribution is constructed based on the GMOW distribution. The $\exp(-X)$ transformation is used to generate the new distribution, where the random variable X has the GMOW distribution. The rest of the study is organized as follows: In Section 2, the new model is defined and some of its basic properties are examined. In Section 3, five estimators are discussed to estimate the unknown model parameters. The Monte Carlo simulation study is conducted to assess the performance of the estimators in Section 4. In Section 5, two real data sets are analyzed. Finally, the paper ends with a conclusion in Section 6.

2. UNIT GENERALIZED MARSHALL-OLKIN WEIBULL DISTRIBUTION

The new model is introduced as follows: Let X be a random variable having the GMOW distribution with parameter vector Ξ. The $Y=\exp(-X)$ transformation is used to achieve the new distribution. Then, the pdf and cdf of Y are obtained, respectively, by,

(3)
$$f(y;\Xi)=\frac{\beta\theta\left(-\log\left(y\right)\right)^{\beta-1}\left\{\left(1-\alpha\right)\left(1-\lambda\right)\exp\left(-\theta\left(-\log\left(y\right)\right)^{\beta}\right)+\left(1+\alpha-\lambda\right)\exp\left(\theta\left(-\log\left(y\right)\right)^{\beta}\right)+2\lambda-2\right\}}{y\left(1-\exp\left(\theta\left(-\log\left(y\right)\right)^{\beta}\right)-\alpha\right)^{2}}$$

and

$$F\left(y;\Xi\right)=\frac{1+\alpha-\lambda+\left(\lambda-1\right)\exp\left\{-\theta\left(-\log\left(y\right)\right)^{\beta}\right\}}{\exp\left\{\theta\left(-\log\left(y\right)\right)^{\beta}\right\}+\alpha-1}, \tag{4}$$

where $y\in(0,1)$. The new unit distribution is called as unit generalized Marshall-Olkin Weibull (UGMOW) distribution and is denoted by $UGMOW\left(\Xi\right)$. The hazard rate function (hrf) of the UGMOW distribution is presented by

$$h(y;\Xi)=\frac{\beta\theta\left(-\log\left(y\right)\right)^{\beta-1}\left\{\left(1-\alpha\right)\left(1-\lambda\right)\exp\left(-\theta\left(-\log\left(y\right)\right)^{\beta}\right)+\left(1+\alpha-\lambda\right)\exp\left(\theta\left(-\log\left(y\right)\right)^{\beta}\right)+2\lambda-2\right\}}{y\left(1-\exp\left(\theta\left(-\log\left(y\right)\right)^{\beta}\right)-\alpha\right)^{2}\left\{1-\frac{1+\alpha-\lambda+(\lambda-1)\exp\left\{-\theta(-\log(y))^{\beta}\right\}}{\exp\left\{\theta(-\log(y))^{\beta}\right\}+\alpha-1}\right\}}.$$

The pdf and hrf plots are presented for some parameter choices of Ξ in Figures 1 and 2, respectively. From Figure 1, it is seen that the pdf is decreasing, increasing and unimodal. It is also concluded from Figure 2 that the hrf is increasing and bathtub shaped.

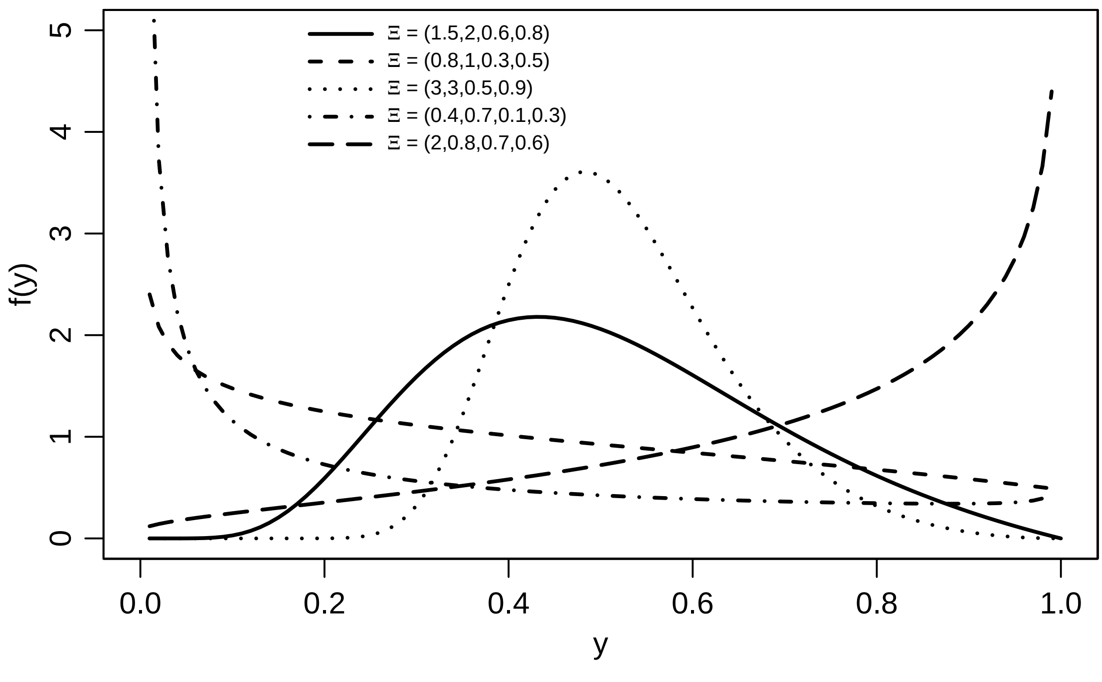


FIGURE 1. The pdf plot of UGMOW distribution for some choices of Ξ

The quantile function of the UGMOW distribution can be obtained by

$$y_{u}\left(\Xi\right)=\exp\left[-\frac{\left\{\log\left(\frac{\alpha-\lambda+1+u-u\alpha+\sqrt{\alpha^{2}(u-1)^{2}-2\alpha(u-1)(1+u-\lambda)+(u+\lambda-1)^{2}}}{2u}\right)\right\}^{1/\beta}}{\theta^{1/\beta}}\right], \tag{5}$$

where $0<u<1$. The median of the UGMOW distribution is obtained by taking $u=0.5$ in Equation (5).

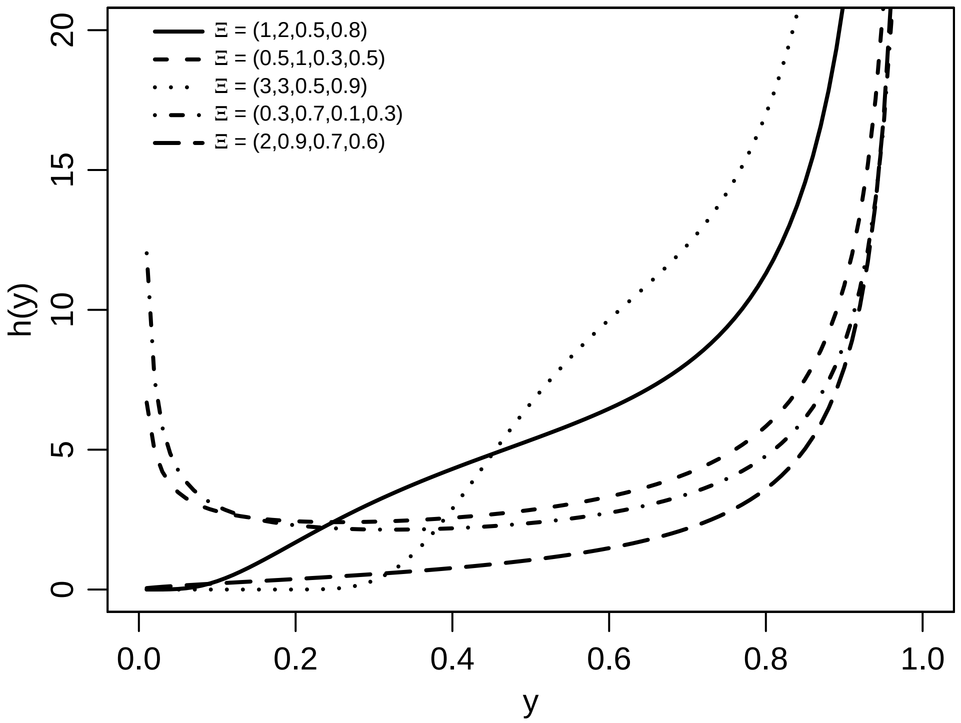


FIGURE 2. The hrf plot of UGMOW distribution for some choices of Ξ

3. Different Estimation Methods

In this section, five estimators for estimating unknown parameters of the $UGMOW\left(\boldsymbol{\Xi}\right)$ distribution are examined. The maximum likelihood, least squares, weighted least squares, Cramér-von Mises and Anderson-Darling estimators are studied. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from UGMOW($\boldsymbol{\Xi}$) distribution and $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ represents the related order statistics. Moreover, $y_{(i)}$ indicates the observed value of $Y_{(i)}$ for $i = 1, 2, \ldots, n$. Then, the log-likelihood function is obtained by

$$\begin{aligned}
\ell\left(\boldsymbol{\Xi}\right) &\approx n\log\left(\beta\theta\right) + \left(\beta-1\right)\sum_{i=1}^{n}\log\left(\left(-\log\left(y_i\right)\right)\right) \\
&+\sum_{i=1}^{n}\log\left\{\left(1-\alpha\right)\left(1-\lambda\right)\exp\left(-\theta\left(-\log\left(y_i\right)\right)^{\beta}\right)+\left(1+\alpha-\lambda\right)\exp\left(\theta\left(-\log\left(y_i\right)\right)^{\beta}\right)+2\lambda-2\right\} \\
&-2\sum_{i=1}^{n}\log\left(1-\exp\left(\theta\left(-\log\left(y_i\right)\right)^{\beta}\right)-\alpha\right).
\end{aligned} \tag{6}$$

Then, the maximum likelihood estimator (MLE) of $\boldsymbol{\Xi}$, say $\widehat{\boldsymbol{\Xi}}_1$, is achieved by

$$\widehat{\boldsymbol{\Xi}}_1 = \arg\max_{\boldsymbol{\Xi}} \ell\left(\boldsymbol{\Xi}\right). \tag{7}$$

Let us give the following functions which help to obtain the other four estimators:

$$Q_{LS}\left(\boldsymbol{\Xi}\right) = \sum_{i=1}^{n}\left(\frac{1+\alpha-\lambda+\left(\lambda-1\right)\exp\left\{-\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}}{\exp\left\{\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}+\alpha-1}-\frac{i}{n+1}\right)^2,$$

$$Q_{WLS}\left(\Xi\right)=$$
$$\sum_{i=1}^{n}\frac{(n+2)(n+1)^2}{i(n-i+1)}\left(\frac{1+\alpha-\lambda+(\lambda-1)\exp\left\{-\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}}{\exp\left\{\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}+\alpha-1}-\frac{i}{n+1}\right)^2,$$

$$Q_{AD}\left(\Xi\right)=$$
$$-n-\frac{1}{n}\sum_{i=1}^{n}\left\{(2i-1)\log\left(\frac{1+\alpha-\lambda+(\lambda-1)\exp\left\{-\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}}{\exp\left\{\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}+\alpha-1}\right)\right\}$$
$$+\frac{1}{n}\sum_{i=1}^{n}\log\left\{1-\frac{1+\alpha-\lambda+(\lambda-1)\exp\left\{-\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}}{\exp\left\{\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}+\alpha-1}\right\},$$

and

$$Q_{CvM}\left(\Xi\right)=\frac{1}{12n}+\sum_{i=1}^{n}\left(\frac{1+\alpha-\lambda+(\lambda-1)\exp\left\{-\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}}{\exp\left\{\theta\left(-\log\left(y_{(i)}\right)\right)^{\beta}\right\}+\alpha-1}-\frac{2i-1}{2n}\right)^2.$$

Then, the least squares estimator (LSE), weighted least squares estimator (WLSE), Anderson Darling estimator (ADE) and the Cramér-von Mises estimator (CvME) of Ξ are obtained, respectively, by

$$\widehat{\Xi}_2=\arg\min_{\Xi}Q_{LS}\left(\Xi\right), \tag{8}$$

$$\widehat{\Xi}_3=\arg\min_{\Xi}Q_{WLS}\left(\Xi\right), \tag{9}$$

$$\widehat{\Xi}_4=\arg\min_{\Xi}Q_{AD}\left(\Xi\right), \tag{10}$$

$$\widehat{\Xi}_5=\arg\min_{\Xi}Q_{CvM}\left(\Xi\right). \tag{11}$$

All optimization problems given in Equations (7)-(11) are solved with the **optim** function in R.

4. Monte Carlo Simulation Experiment

In this section, the bias and mean square errors (MSEs) of all estimators are estimated for UGMOW parameters via Monte Carlo simulation based on 5000 trials. Different sample of sizes are selected in the simulation study. Five parameter settings, such as $\boldsymbol{\Xi_1}=(2,1.5,0.5,0.3)$, $\boldsymbol{\Xi_2}=(1.5,0.5,0.9,0.7)$, $\boldsymbol{\Xi_3}=(0.5,0.5,0.5,0.5)$, $\boldsymbol{\Xi_4}=(1,1,0.4,0.6)$ and $\boldsymbol{\Xi_5}=(5,3,0.4,0.4)$ are considered. Simulation results are reported in Tables 1-2. From Tables 1-2, it can be concluded that the bias and MSEs of all estimates decrease to zero.

5. Real Data Applications

In this section, two real data analyses are conducted to show the ability of the UGMOW distribution. Some competitor unit distributions such as beta (B), Kumaraswamy (Kw), unit-Lindley (UL) [9], unit-Lindley (UL2) [10], the Topp-Leone (TL) [12] and unit Burr-XII (UBXII) [7] are selected for comparing the fitting results. Real data analyses are performed based on the MLE method for all models. The pdfs of all models are given in Table 3.

The MLEs of the distribution parameters, log-likelihood value (ℓ), Akaike's information criteria (AIC), Bayesian information criterion (BIC), consistent AIC (CAIC), Hannan-Quinn information criterion (HQIC), Kolmogorov-Smirnov statistics (KS), Anderson-Darling statistics (AD), Cramér von Mises statistic (CvM) and p-values of statistics (KS p-value, AD p-value and CvM p-value) are given in Tables 5-6 for maximum flood levels data and petroleum reservoirs data, respectively.

The first data set represents the maximum flood levels data and is taken from [3]. For detailed information about the data, one can consult [3] and [5]. In addition, the UGMOW distribution gave better results than all distributions in [5] for the first data set. The first data is given as follow: 0.654, 0.613, 0.315, 0.449, 0.297, 0.402, 0.379, 0.423, 0.379, 0.3235, 0.269, 0.740, 0.418, 0.412, 0.494, 0.416, 0.338, 0.392, 0.484, 0.265.

TABLE 1. Average bias for five parameter settings

		$\widehat{\Xi}_1$				$\widehat{\Xi}_2$				$\widehat{\Xi}_3$				$\widehat{\Xi}_4$				$\widehat{\Xi}_5$			
n	Ξ	θ	β	λ	α	θ	β	λ	α	θ	β	λ	α	θ	β	λ	α	θ	β	λ	α
25		-0.3778	0.9404	0.0733	-0.2489	-0.5372	0.7757	0.0813	-0.2376	-0.5303	0.6550	0.0773	-0.2248	-0.4756	0.7212	0.0924	-0.2307	-0.3496	1.0201	0.0806	-0.2523
50		-0.3510	0.5538	0.0527	-0.2013	-0.4432	0.5747	0.0626	-0.2089	-0.4016	0.4557	0.0466	-0.1901	-0.3916	0.4690	0.0686	-0.1903	-0.3666	0.7098	0.0597	-0.2244
75		-0.2946	0.4252	0.0418	-0.1739	-0.3830	0.4790	0.0507	-0.1905	-0.3484	0.3884	0.0488	-0.1712	-0.3212	0.3774	0.0512	-0.1663	-0.3393	0.5732	0.0475	-0.2049
100	Ξ_1	-0.2708	0.3463	0.0463	-0.1524	-0.3356	0.4106	0.0503	-0.1721	-0.2939	0.3240	0.0394	-0.1541	-0.2891	0.3197	0.0540	-0.1493	-0.3065	0.4823	0.0465	-0.1867
125		-0.2547	0.2884	0.0393	-0.1341	-0.3182	0.3602	0.0405	-0.1600	-0.2817	0.2832	0.0340	-0.1424	-0.2720	0.2758	0.0441	-0.1356	-0.2978	0.4184	0.0361	-0.1736
250		-0.1645	0.1703	0.0349	-0.0825	-0.2202	0.2397	0.0355	-0.1133	-0.1847	0.1820	0.0259	-0.0983	-0.1818	0.1777	0.0399	-0.0882	-0.2149	0.2708	0.0318	-0.1252
500		-0.1147	0.0991	0.0483	-0.0407	-0.1553	0.1498	0.0426	-0.0710	-0.1310	0.1131	0.0422	-0.0553	-0.1255	0.1068	0.0466	-0.0492	-0.1548	0.1664	0.0400	-0.0797
25		-0.6324	0.4144	-0.2454	-0.6461	-0.7395	0.3571	-0.2103	-0.6333	-0.7270	0.3188	-0.2012	-0.6217	-0.6803	0.3396	-0.2065	-0.6314	-0.6577	0.4500	-0.2321	-0.6534
50		-0.5216	0.2474	-0.2590	-0.5908	-0.6315	0.2568	-0.2290	-0.5948	-0.5784	0.2177	-0.2334	-0.5761	-0.5588	0.2205	-0.2278	-0.5806	-0.6080	0.3076	-0.2494	-0.6191
75		-0.4294	0.1874	-0.2669	-0.5507	-0.5408	0.2109	-0.2388	-0.5657	-0.4766	0.1760	-0.2510	-0.5484	-0.4545	0.1718	-0.2502	-0.5458	-0.5323	0.2466	-0.2584	-0.5916
100	Ξ_2	-0.3744	0.1575	-0.2686	-0.5239	-0.4827	0.1853	-0.2454	-0.5465	-0.3996	0.1477	-0.2754	-0.5284	-0.4091	0.1518	-0.2480	-0.5243	-0.4797	0.2128	-0.2628	-0.5706
125		-0.3282	0.1332	-0.2779	-0.5023	-0.4350	0.1659	-0.2574	-0.5289	-0.3797	0.1372	-0.2644	-0.5117	-0.3656	0.1333	-0.2599	-0.5073	-0.4363	0.1883	-0.2726	-0.5525
250		-0.2054	0.0784	-0.2805	-0.4204	-0.3165	0.1121	-0.2617	-0.4651	-0.2592	0.0899	-0.2758	-0.4480	-0.2497	0.0864	-0.2680	-0.4390	-0.3233	0.1246	-0.2746	-0.4870
500		-0.1009	0.0444	-0.2825	-0.3422	-0.2039	0.0755	-0.2737	-0.4021	-0.1503	0.0568	-0.2902	-0.3844	-0.1433	0.0544	-0.2830	-0.3749	-0.2109	0.0823	-0.2824	-0.4190
25		-0.2618	0.2583	-0.0136	-0.4166	-0.2856	0.2616	0.0177	-0.4277	-0.2123	0.1846	-0.0274	-0.3912	-0.2388	0.1976	0.0065	-0.4004	-0.3206	0.3530	0.0194	-0.4499
50		-0.1946	0.1444	-0.0287	-0.3226	-0.2730	0.2249	0.0229	-0.3706	-0.2085	0.1545	-0.0147	-0.3295	-0.1732	0.1261	-0.0353	-0.3160	-0.3102	0.2828	0.0300	-0.4044
75		-0.1483	0.1093	-0.0500	-0.2732	-0.2575	0.2093	0.0032	-0.3601	-0.1545	0.1282	-0.0847	-0.3089	-0.1780	0.1252	-0.0212	-0.2937	-0.2846	0.2490	0.0065	-0.3862
100	Ξ_3	-0.1090	0.0823	-0.0593	-0.2456	-0.2422	0.1840	0.0100	-0.3533	-0.1700	0.1220	-0.0337	-0.3061	-0.1601	0.1065	-0.0170	-0.2841	-0.2637	0.2134	0.0117	-0.3747
125		-0.0887	0.0665	-0.0422	-0.2037	-0.2199	0.1605	0.0266	-0.3134	-0.1346	0.0977	-0.0425	-0.2603	-0.1340	0.0896	-0.0140	-0.2436	-0.2411	0.1853	0.0266	-0.3386
250		-0.0729	0.0472	-0.0362	-0.1573	-0.1437	0.1020	-0.0233	-0.2372	-0.1225	0.0752	-0.0085	-0.2016	-0.1028	0.0631	-0.0310	-0.2006	-0.1607	0.1156	-0.0203	-0.2597
500		-0.0416	0.0242	-0.0355	-0.1120	-0.1227	0.0739	0.0148	-0.1788	-0.0853	0.0472	-0.0094	-0.1523	-0.0791	0.0430	-0.0042	-0.1400	-0.1333	0.0814	0.0120	-0.1983
25		-0.4679	0.7538	0.1101	-0.5175	-0.4668	0.6763	0.1139	-0.4818	-0.4344	0.5688	0.1257	-0.4698	-0.4368	0.5945	0.1241	-0.4952	-0.5163	0.8953	0.1192	-0.5260
50		-0.3654	0.3770	0.0901	-0.4178	-0.3287	0.3788	0.0232	-0.3835	-0.2618	0.2686	0.0220	-0.3461	-0.3137	0.2991	0.0524	-0.3986	-0.3777	0.4901	0.0212	-0.4410
75		-0.3029	0.2802	0.0563	-0.3726	-0.2693	0.3028	-0.0338	-0.3658	-0.2615	0.2497	0.0036	-0.3502	-0.2399	0.2242	-0.0264	-0.3758	-0.3091	0.3771	-0.0345	-0.4144
100	Ξ_4	-0.2651	0.2200	0.0506	-0.3139	-0.1728	0.1904	-0.1048	-0.3073	-0.1701	0.1530	-0.0714	-0.2958	-0.1835	0.1565	-0.0731	-0.3296	-0.2061	0.2421	-0.1117	-0.3598
125		-0.2045	0.1644	0.0317	-0.2606	-0.1074	0.1216	-0.1436	-0.2701	-0.1383	0.1165	-0.0597	-0.2463	-0.1272	0.1049	-0.1011	-0.2890	-0.1375	0.1621	-0.1521	-0.3229
250		-0.0917	0.0651	-0.0246	-0.1652	-0.0563	0.0525	-0.1759	-0.2442	-0.0638	0.0455	-0.1258	-0.2178	-0.0653	0.0460	-0.1233	-0.2228	-0.0785	0.0744	-0.1827	-0.2862
500		-0.0352	0.0243	-0.0368	-0.0987	-0.0176	0.0109	-0.1925	-0.2233	-0.0219	0.0119	-0.1414	-0.1839	-0.0174	0.0088	-0.1451	-0.1840	-0.0332	0.0235	-0.1982	-0.2547
25		1.7978	1.9931	0.1502	-0.3441	-0.2843	1.4478	0.1768	-0.3237	0.2689	1.3891	0.1564	-0.3124	0.4703	1.5146	0.1653	-0.3205	0.5757	1.8201	0.1778	-0.3373
50		0.3689	1.2692	0.1133	-0.2969	-0.1480	1.1958	0.1357	-0.2892	-0.0597	1.0218	0.1069	-0.2715	-0.0117	1.0580	0.1186	-0.2789	0.3510	1.4296	0.1290	-0.3064
75		0.1585	1.0044	0.0882	-0.2621	-0.0040	1.0313	0.0897	-0.2584	-0.0560	0.8596	0.0658	-0.2542	-0.0247	0.8702	0.0795	-0.2501	0.3297	1.2042	0.0863	-0.2778
100	Ξ_5	0.0141	0.8661	0.1038	-0.2349	-0.0519	0.9577	0.1001	-0.2420	-0.1052	0.7658	0.0688	-0.2335	-0.1043	0.7943	0.0983	-0.2297	0.2069	1.0941	0.0932	-0.2623
125		-0.0103	0.7551	0.0973	-0.2153	-0.0114	0.8453	0.0910	-0.2170	-0.0943	0.6952	0.0751	-0.2143	-0.0946	0.7066	0.0963	-0.2099	0.1938	0.9580	0.0836	-0.2389
250		-0.0762	0.5067	0.0836	-0.1569	-0.0101	0.5889	0.0478	-0.1744	-0.0855	0.4935	0.0544	-0.1688	-0.0908	0.4768	0.0612	-0.1624	0.0850	0.6496	0.0428	-0.1920
500		-0.0793	0.3140	0.0689	-0.0925	-0.0262	0.3581	0.0101	-0.1230	-0.0729	0.2946	0.0104	-0.1274	-0.0753	0.2867	0.0183	-0.1200	0.0142	0.3909	0.0062	-0.1385

TABLE 2. Average MSEs for five parameter settings

		$\widehat{\Xi}_1$				$\widehat{\Xi}_2$				$\widehat{\Xi}_3$				$\widehat{\Xi}_4$				$\widehat{\Xi}_5$			
n	Ξ	θ	β	λ	α	θ	β	λ	α	θ	β	λ	α	θ	β	λ	α	θ	β	λ	α
25		0.6603	1.2195	0.0370	0.0672	0.8271	1.0210	0.0536	0.0657	0.8463	0.8179	0.0633	0.0627	0.6878	0.8179	0.0535	0.0589	0.8306	1.4864	0.0437	0.0681
50		0.3351	0.4596	0.0332	0.0503	0.5065	0.5641	0.0472	0.0552	0.4323	0.4056	0.0526	0.0501	0.3861	0.3624	0.0451	0.0449	0.4693	0.7470	0.0385	0.0565
75		0.2355	0.2851	0.0312	0.0417	0.3582	0.3920	0.0423	0.0478	0.2969	0.2780	0.0437	0.0420	0.2660	0.2447	0.0406	0.0385	0.3317	0.4948	0.0348	0.0491
100	Ξ_1	0.1949	0.1978	0.0297	0.0358	0.2900	0.2977	0.0406	0.0436	0.2257	0.2070	0.0412	0.0371	0.2136	0.1800	0.0372	0.0330	0.2713	0.3632	0.0339	0.0438
125		0.1655	0.1472	0.0302	0.0322	0.2381	0.2334	0.0391	0.0407	0.1910	0.1577	0.0390	0.0336	0.1804	0.1391	0.0363	0.0306	0.2234	0.2804	0.0326	0.0403
250		0.0874	0.0620	0.0280	0.0225	0.1320	0.1125	0.0355	0.0314	0.0989	0.0720	0.0335	0.0252	0.0952	0.0643	0.0314	0.0228	0.1272	0.1284	0.0308	0.0297
500		0.0547	0.0293	0.0260	0.0173	0.0812	0.0558	0.0337	0.0248	0.0608	0.0351	0.0281	0.0192	0.0590	0.0319	0.0284	0.0178	0.0794	0.0609	0.0305	0.0231
25		0.5571	0.2166	0.0779	0.4234	0.7113	0.1866	0.0665	0.4114	0.6959	0.1535	0.0658	0.3975	0.6069	0.1550	0.0650	0.4032	0.6377	0.2676	0.0727	0.4298
50		0.3662	0.0791	0.0843	0.3604	0.5029	0.0946	0.0738	0.3687	0.4436	0.0705	0.0807	0.3464	0.4030	0.0654	0.0726	0.3461	0.4711	0.1242	0.0798	0.3894
75		0.2594	0.0469	0.0886	0.3183	0.3785	0.0642	0.0789	0.3379	0.3159	0.0465	0.0893	0.3149	0.2804	0.0412	0.0868	0.3098	0.3627	0.0805	0.0845	0.3584
100	Ξ_2	0.2023	0.0336	0.0898	0.2905	0.3063	0.0493	0.0815	0.3166	0.2358	0.0343	0.1087	0.2933	0.2263	0.0314	0.0811	0.2866	0.2972	0.0601	0.0866	0.3355
125		0.1620	0.0245	0.0958	0.2697	0.2567	0.0398	0.0872	0.3011	0.1991	0.0272	0.0904	0.2767	0.1843	0.0244	0.0868	0.2702	0.2516	0.0476	0.0915	0.3184
250		0.0809	0.0098	0.0981	0.2011	0.1447	0.0191	0.0886	0.2433	0.1011	0.0123	0.0955	0.2198	0.0939	0.0110	0.0901	0.2110	0.1450	0.0219	0.0924	0.2571
500		0.0395	0.0041	0.1010	0.1463	0.0787	0.0096	0.0954	0.1936	0.0493	0.0056	0.1027	0.1701	0.0460	0.0051	0.0983	0.1622	0.0788	0.0106	0.0979	0.2024
25		0.0974	0.0965	0.0227	0.1921	0.1054	0.1163	0.0342	0.1958	0.0883	0.0769	0.0485	0.1803	0.0827	0.0645	0.0365	0.1770	0.1188	0.1759	0.0260	0.2073
50		0.0694	0.0372	0.0275	0.1346	0.1043	0.0844	0.0301	0.1663	0.0781	0.0516	0.0404	0.1449	0.0760	0.0346	0.0438	0.1291	0.1143	0.1130	0.0230	0.1776
75		0.0503	0.0243	0.0349	0.1089	0.0916	0.0678	0.0336	0.1560	0.0725	0.0389	0.0511	0.1357	0.0589	0.0287	0.0428	0.1177	0.0997	0.0854	0.0273	0.1643
100	Ξ_3	0.0444	0.0187	0.0425	0.0994	0.0810	0.0522	0.0359	0.1499	0.0605	0.0308	0.0485	0.1287	0.0505	0.0226	0.0479	0.1114	0.0871	0.0635	0.0303	0.1562
125		0.0402	0.0148	0.0461	0.0840	0.0770	0.0429	0.0421	0.1338	0.0563	0.0244	0.0510	0.1140	0.0487	0.0190	0.0517	0.0982	0.0814	0.0512	0.0366	0.1392
250		0.0306	0.0090	0.0409	0.0664	0.0634	0.0241	0.0407	0.1070	0.0425	0.0138	0.0387	0.0877	0.0383	0.0115	0.0468	0.0812	0.0644	0.0269	0.0390	0.1079
500		0.0197	0.0041	0.0426	0.0520	0.0407	0.0124	0.0393	0.0827	0.0269	0.0067	0.0409	0.0663	0.0250	0.0059	0.0435	0.0610	0.0410	0.0134	0.0367	0.0815
25		0.3032	0.7687	0.0292	0.2829	0.3173	0.7904	0.0452	0.2676	0.2749	0.5860	0.0524	0.2536	0.2642	0.5283	0.0464	0.2598	0.3391	1.1559	0.0416	0.2912
50		0.1880	0.2148	0.0319	0.1968	0.2332	0.3524	0.0498	0.2070	0.1790	0.2060	0.0492	0.1876	0.1628	0.1756	0.0499	0.1875	0.2529	0.4666	0.0515	0.2300
75		0.1439	0.1279	0.0329	0.1677	0.1930	0.2573	0.0448	0.1960	0.1488	0.1599	0.0402	0.1805	0.1283	0.1210	0.0390	0.1739	0.2070	0.3198	0.0486	0.2132
100	Ξ_4	0.1258	0.0916	0.0427	0.1340	0.1636	0.1781	0.0556	0.1662	0.1213	0.1045	0.0456	0.1547	0.1111	0.0857	0.0477	0.1529	0.1721	0.2120	0.0614	0.1833
125		0.0987	0.0644	0.0486	0.1076	0.1315	0.1313	0.0665	0.1478	0.1036	0.0760	0.0549	0.1284	0.0871	0.0596	0.0541	0.1317	0.1365	0.1527	0.0725	0.1628
250		0.0563	0.0261	0.0481	0.0703	0.0852	0.0666	0.0707	0.1317	0.0620	0.0352	0.0560	0.1132	0.0584	0.0318	0.0539	0.1036	0.0858	0.0725	0.0754	0.1409
500		0.0362	0.0139	0.0498	0.0549	0.0502	0.0291	0.0701	0.1213	0.0355	0.0163	0.0524	0.0948	0.0334	0.0147	0.0522	0.0904	0.0495	0.0305	0.0734	0.1263
25		20.4886	5.2146	0.0516	0.1255	3.7788	3.1755	0.0743	0.1171	13.8506	3.4367	0.0797	0.1158	7.7387	3.3701	0.0710	0.1132	4.5186	4.3323	0.0671	0.1213
50		3.4873	2.2394	0.0419	0.0994	2.7480	2.1948	0.0610	0.1026	3.8916	1.9020	0.0607	0.0964	2.7079	1.7328	0.0566	0.0917	3.2226	2.8200	0.0537	0.1065
75		1.4632	1.4392	0.0384	0.0838	2.0277	1.7567	0.0534	0.0953	1.9253	1.3378	0.0544	0.0869	1.4369	1.2246	0.0496	0.0804	2.3326	2.1478	0.0478	0.0968
100	Ξ_5	0.9569	1.0977	0.0438	0.0723	1.6124	1.5104	0.0559	0.0877	1.2732	1.0989	0.0551	0.0786	1.0644	1.0068	0.0530	0.0713	1.8076	1.8039	0.0507	0.0886
125		0.7681	0.8691	0.0429	0.0644	1.4218	1.2663	0.0549	0.0817	1.0649	0.8907	0.0527	0.0714	0.8896	0.8180	0.0519	0.0642	1.5856	1.4835	0.0506	0.0817
250		0.3217	0.4399	0.0418	0.0464	0.6494	0.7580	0.0505	0.0694	0.4044	0.4960	0.0479	0.0577	0.3703	0.4492	0.0476	0.0513	0.6927	0.8435	0.0477	0.0678
500		0.1616	0.2216	0.0462	0.0378	0.2765	0.4227	0.0477	0.0584	0.1798	0.2561	0.0478	0.0492	0.1738	0.2375	0.0491	0.0461	0.2836	0.4492	0.0467	0.0558

TABLE 3. List of the distributions with pdfs and domains of parameters

Distribution	Pdf	Domains of parameters
UGMOW	$\frac{p_1p_2(-\log(y))^{p_2-1}\{(1-p_4)(1-p_3)\exp(-p_1(-\log(y))^{p_2})+(1+p_4-p_3)\exp(p_1(-\log(y))^{p_2})+2p_3-2\}}{y(1-\exp(p_1(-\log(y))^{p_2})-p_4)^2}$,	$p_1, p_2 > 0;\ p_3, p_4 \in (0,1)$
B	$\frac{\Gamma(p_1+p_2)}{\Gamma(p_1)\Gamma(p_2)} y^{p_1-1}(1-y)^{p_2-1}$,	$p_1, p_2 > 0$
K	$p_1p_2y^{p_1-1}(1-y^{p_1})^{p_{2-1}}$,	$p_1, p_2 > 0$
UL	$\frac{p_1^2}{(1+p_1)(1-y)^3\exp\left(\frac{p_1 y}{1-y}\right)}$,	$p_1 > 0$
UL2	$\frac{p_1^2}{(1+p_1)y^3\exp\left(\frac{p_1(1-y)}{y}\right)}$,	$p_1 > 0$
TL	$p_1(2-2y)(2y-y^2)^{p_1-1}$,	$p_1 > 0$
UBXII	$p_1p_2y^{-1}(-\log(y))^{p_1-1}(1+(-\log(y))^{p_2})^{-p_1-1}$,	$p_1, p_2 > 0$

The second data set illustrates the twelve core samples from petroleum reservoirs and is taken from [2]. The second data is presented as follow: 0.0903, 0.2036, 0.2043, 0.2808, 0.1976, 0.3286, 0.1486, 0.1623, 0.2627, 0.1794, 0.3266, 0.2300, 0.1833, 0.1509, 0.2000, 0.1918, 0.1541, 0.4641, 0.1170, 0.1481, 0.1448, 0.1330, 0.2760, 0.4204, 0.1224, 0.2285, 0.1138, 0.2252, 0.1769, 0.2007, 0.1670, 0.2316, 0.2910, 0.3412, 0.4387, 0.2626, 0.1896, 0.1725, 0.2400, 0.3116, 0.1635, 0.1824, 0.1641, 0.1534, 0.1618, 0.2760, 0.2538, 0.2004. Some descriptive statistics, such as mean, variance, skewness, kurtosis, median, minimum and maximum are reported in Table 4 for two data sets.

Table 4. Some descriptive statistics for the maximum flood levels and petroleum reservoirs data

Data	n	Mean	Variance	Skewness	Kurtosis	Median	Minimum	Maximum
Maximum flood levels	20	0.4231	0.0156	1.1558	1.1534	0.4070	0.2650	0.7400
Petroleum reservoirs	48	0.2181	0.0069	1.2073	1.3709	0.1989	0.0903	0.4641

Table 5. The analysis results for the maximum flood levels data

	UGMOW	B	K	UL	UL2	TL	UBXII
ℓ	17.0687	14.0622	12.8662	7.1406	9.1966	7.3674	14.5944
AIC	-26.1374	-24.1245	-21.7324	-12.2812	-16.3932	-12.7348	-25.1889
BIC	-22.1545	-22.1330	-19.7409	-11.2855	-15.3974	-11.7391	-23.1974
CAIC	-23.4708	-23.4186	-21.0265	-12.0590	-16.1709	-12.5126	-24.4830
HQIC	-25.3599	-23.7357	-21.3436	-12.0868	-16.1988	-12.5404	-24.8001
KS	0.1377	0.1988	0.2109	0.3182	0.2964	0.3352	0.1794
AD	0.2475	0.7327	0.9322	2.2801	2.3172	2.8762	0.6002
CvM	0.0456	0.1236	0.1636	0.4348	0.4413	0.5769	0.0894
KS p-value	0.8429	0.4082	0.3359	0.0349	0.0595	0.0224	0.5389
AD p-value	0.9716	0.5302	0.3936	0.0654	0.0625	0.0321	0.6457
CVM p-value	0.9077	0.4847	0.3528	0.0576	0.0553	0.0245	0.6444
p_1	0.8217	6.568	3.3632	1.6205	0.9773	2.2446	1.6501
p_2	4.7739	9.1114	11.7892				4.8034
p_3	0.1194						
p_4	0.0097						

Table 6. The analysis results for the petroleum reservoirs data

	UGMOW	B	K	UL	UL2	TL	UBXII
ℓ	58.4579	55.5993	52.4910	35.3501	43.9113	21.1660	37.2404
AIC	-108.9159	-10.1985	-100.9821	-68.7001	-85.8227	-40.3320	-70.4807
BIC	-101.4310	-103.4561	-97.2397	-66.8289	-83.9515	-38.4608	-66.7383
CAIC	-107.9856	-106.9319	-100.7154	-68.6132	-85.7357	-40.2450	-70.2140
HQIC	-106.0873	-105.7843	-99.5678	-67.9930	-85.1156	-39.6249	-69.0665
KS	0.0748	0.1428	0.1533	0.3237	0.2277	0.3080	0.2592
AD	0.1760	0.7769	1.2891	6.5496	4.5250	10.5685	5.7918
CvM	0.0247	0.1301	0.2060	1.2775	0.8112	2.1463	1.1861
KS p-value	0.9509	0.2820	0.2092	0.0001	0.0138	0.0000	0.0032
AD p-value	0.9957	0.4973	0.2358	0.0006	0.0049	0.0000	0.0012
CVM p-value	0.9909	0.4579	0.2566	0.0005	0.0066	0.0000	0.0008
p_1	0.0480	5.9415	2.7187	4.0491	0.4077	0.9894	0.2305
p_2	5.4501	21.2047	44.6587				9.6042
p_3	0.0929						
p_4	0.0526						

As can be seen from Tables 5-6, the modeling ability of the UGMOW distribution is quite effective. According to many criteria, the UGMOW distribution gives better results than other distributions. Therefore, it can be said that the UGMOW distribution is a proper alternative distribution for modeling the bounded data.

6. Conclusion

In this study, a new unit distribution called as unit generalized Marshall-Olkin Weibull is introduced. Some basic function of the new distribution is presented. The shapes of the probability density function and hazard rate function are illustrated. It is found that the hazard rate function has very flexible shapes and its effective in modeling. Five estimation methods are investigated to estimate the unknown parameters of the new model. The performance of the estimators is evaluated via Monte Carlo simulation. The simulation results show that all estimators gave similar findings. In order to demonstrate the applicability of UGMOW distribution, two real data, such as maximum flood levels data and petroleum reservoirs data, are also examined.

References

[1] Chesneau C., Karakaya K., Bakouch H. S. and Kus C., (2022), An Alternative to the Marshall-Olkin Family of Distributions: Bootstrap, Regression and Applications, Communications on Applied Mathematics and Computation, 4, 1229-1257.

[2] Cordeiro G.M. and dos Santos Brito R., (2012). The beta power distribution, Brazilian journal of probability and statistics, 26(1), 88-112.

[3] Dumonceaux R. and Antle C.E., (1973). Discrimination between the log-normal and the Weibull distributions, Technometrics, 15(4), 923-926.

[4] Gomez-Deniz E., Sordo M.A. and Calderin-Ojeda E., (2014). The Log-Lindley distribution as an alternative to the beta regression model with applications in insurance, Insurance: Mathematics and Economics, 54, 49-57.

[5] Krishna A., Maya R., Chesneau C. and Irshad M.R, (2022). The Unit Teissier Distribution and Its Applications, Mathematical and Computational Applications, 27(1), 12.

[6] Korkmaz M.C., Altun E., Alizadeh M. and El-Morshedy M., (2021). The Log Exponential-Power Distribution: Properties, Estimations and Quantile Regression Model, Mathematics, 9(21), 2634.

[7] Korkmaz M.C. and Chesneau C., (2021). On the unit Burr-XII distribution with the quantile regression modeling and applications, Computational and Applied Mathematics, 40(1), 1-26.

[8] Kumaraswamy P., (1980). A generalized probability density function for double-bounded random processes, Journal of Hydrology, 46(1-2), 79-88.

[9] Mazucheli J., Bapat S.R. and Menezes A.F. B., (2020). A new one-parameter unit-Lindley distribution, Chilean Journal of Statistics (ChJS), 11(1), 53-67.

[10] Mazucheli J., Menezes A.F.B. and Chakraborty S., (2019). On the one parameter unit-Lindley distribution and its associated regression model for proportion data, Journal of Applied Statistics 46(4), 700-714.

[11] Mazucheli J., Menezes A.F.B. and Ghitany M.E., (2018). The unit-Weibull distribution and associated inference, J. Appl. Probab. Stat, 13(2), 1-22.

[12] Topp C.W. and Leone F.C., (1955). A family of J-Shaped frequency functions, Journal of the American Statistical Association, 50, 209–219.

Department of Statistics, Faculty of Science, Selcuk University, Konya, Turkiye

Email(s): kkarakaya@selcuk.edu.tr

Classification of Eye Diseases Based on Retinal Images Using Deep Learning

İsmail Hakkı Kınalıoğlu

ABSTRACT. Deep learning networks are widely used in many fields due to their classification capabilities. One of these fields is Medicine. Especially with the effect of developing medical imaging techniques, deep learning networks are used in the diagnosis of many diseases. In this study, diseases such as Diabetic Retinopathy, Cataract and Glaucoma, which are among the most common eye diseases, are diagnosed by using deep learning networks. For this, seven different convolutional neural network architectures are used. These are RESNET, Densenet121, Mobilenet and EfficientNET-B3,B5,B6,B7 architectures. The study data consists of retinal images of people with related diseases. These images are collected from various web source. The performances of different types of architectures are compared using approximately 5000 retinal images. The most successful performance is achieved with the classification model using Efficient-B6-B7 architectures with an accuracy rate of 92.89%.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 68T01,68T07

KEYWORDS: Classification, Convolutional Neural Networks, Deep Learning, Disease Diagnosis, Eye Disease Diagnosis

1. INTRODUCTION

Deep learning, also known as deep learning networks or deep neural networks in the literature, first emerged in 2006 as a sub-field of machine learning. Simply put, deep learning networks are types of traditional neural networks that have more hidden layers. Until the discovery of deep learning networks, artificial neural networks are not very successful compared to traditional methods, except for a few special problems [16]. Deep learning networks are very successful in areas such as image recognition, voice recognition, natural language processing. For this reason, they are frequently used in solving different types of problems such as classification, clustering, anomaly detection in various fields. One of the fields where deep learning is used most frequently is medicine. The outputs obtained as a result of medical imaging techniques are very suitable materials for the superior capabilities of deep learning networks in image recognition. The success of deep learning networks in classifying these medical images has been demonstrated in many studies.

In this study, eye diseases are diagnosed using various deep learning architectures. Retinal images are used for the diagnosis of diabetic retinopathy, cataract and glaucoma, which are the most common eye diseases among the public. The definitive diagnosis of these diseases is made as a result of expert examination of retinal images. In places where the number of patients is high, the number of specialist personnel can insufficient. Computer-aided diagnostic models are very important in order to minimize the delays caused by this inadequacy. Classification models that diagnose disease with high success using retinal images will ease the burden of specialist health personnel. There are many studies in the literature for this purpose. These studies are generally carried out for the diagnosis of a single disease. The most distinctive feature of this study that distinguishes it from its similar is the ability to diagnose three different types of diseases in a single model.

- **Diabetic Retinopathy**

Diabetes Mellitus is a chronic disease in which blood sugar levels tend to rise due to the inability of the pancreas to produce or secrete enough blood insulin [3]. Diabetic Retinopathy (DR) is a common specific microvascular complication of diabetes. In DR, swelling that occurs as a result of glucose blocking the blood vessels feeding the eye causes serious eye injuries.

FIGURE 1. Eye with diabetic retinopathy

- **Glaucoma**

Glaucoma is a group of eye diseases that impair vision by damaging the intraocular optic nerves [4]. It causes vision loss and blindness and is often incurable [13]. The most common type, open-angle glaucoma often develops without any symptoms and can be difficult to detect without loss of vision. The most common symptoms are high intraocular pressure, optic nerve damage, large cup/disc ratio and visual loss due to these.

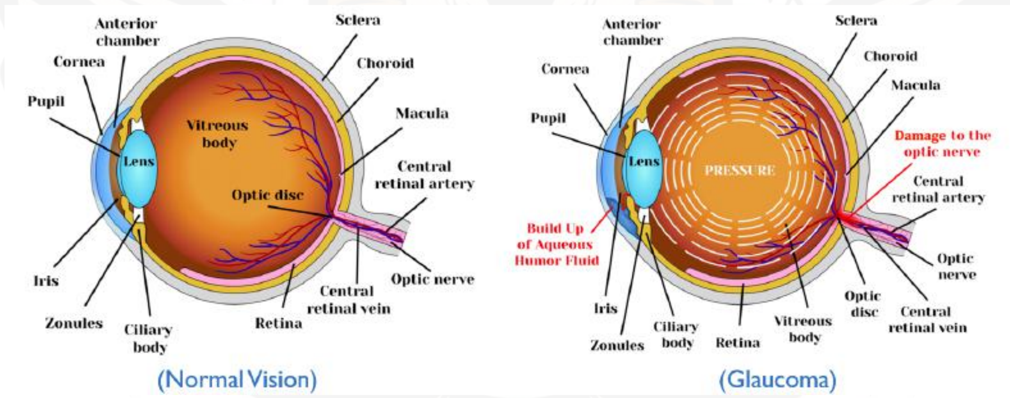


FIGURE 2. Eye with glaucoma

- **Cataract**

Cataract is an eye disease that occurs when the lens in the eye loses its transparency. The task of this lens is to focus the image on the retina. With the loss of transparency, vision loss occurs in the eye. Cataracts usually develop in advanced ages, but there are also congenital types. It can be treated with surgical intervention. The natural intraocular lens is replaced with an artificial lens.

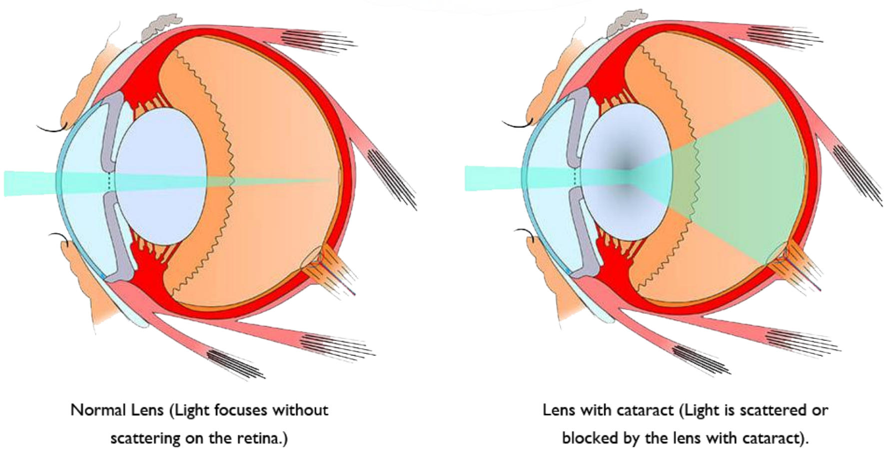


FIGURE 3. Eye with cataract

Some studies in the literature for the diagnosis of eye diseases are given below.

Jain, L. et al. (2018) [5] used a deep learning model to classify retinal images as diseased and healthy. The accuracy of the model they use is between 96.5% and 99.7%.

Nazir, T et al. (2021) [6] proposed a new method consisting of two main steps to automatically diagnose diabetic retinopathy. In the first step, they worked on dataset preparation and feature extraction, and in the second step, they developed a special deep learning-based CenterNet model trained for eye disease classification. They evaluated their models on APTOS and IDRiD datasets and attained average accuracies of 97.93% and 98.10% respectively.

zelik, B. B. & Altan, A. (2021) [9] developed a deep learning-based model for the diagnosis of diabetic retinopathy. The datasets, which include nine hundred retinal images, consist of 5 classes. The classification performance of the proposed model is 97.8%.

Abbas, Q. (2017) [1] developed a hybrid deep learning model for diagnosing glaucoma disease. By testing the model he developed on 1200 retina images, he achieved an average accuracy of 99%.

Uar, M. (2021) [14] used deep learning methods for the diagnosis of glaucoma. He compared VGG16, Inception-V3, EfficientNet, DenseNet, ResNet50 and MobileNet architectures and achieved the highest success rate (96.19%) with the DenseNet architecture.

Aalday, M. F. & nar, A. (2021) [2] compared the success of two different deep learning methods for the diagnosis of cataract disease. These methods are the convolutional neural network (CNN) and the deep residual network (DRN). There is a dataset of 5000 retinal images consisting of 8 classes belonging to different cataract stages. They attained 89% accuracy with the CNN model and 95% accuracy with the DRN model.

2. Material and Method

2.1. Data Collection and Preprocessing. Study data consists of retinal images collected from various web sources. Some of these resources also include datasets previously created for different studies. Some of these datasets are given below.

- IDRID (Indian Diabetic Retinopathy Image Dataset) [11]
- Ocular Disease Recognition [7]
- High-Resolution Fundus (HRF) Image Database [8]
- Retinal Fundus Multi-Disease Image Dataset (RFMID) [10]

Approximately five thousand retinal images are collected labeled by experts in four different types: diabetic retinopathy, glaucoma, cataract, and normal.

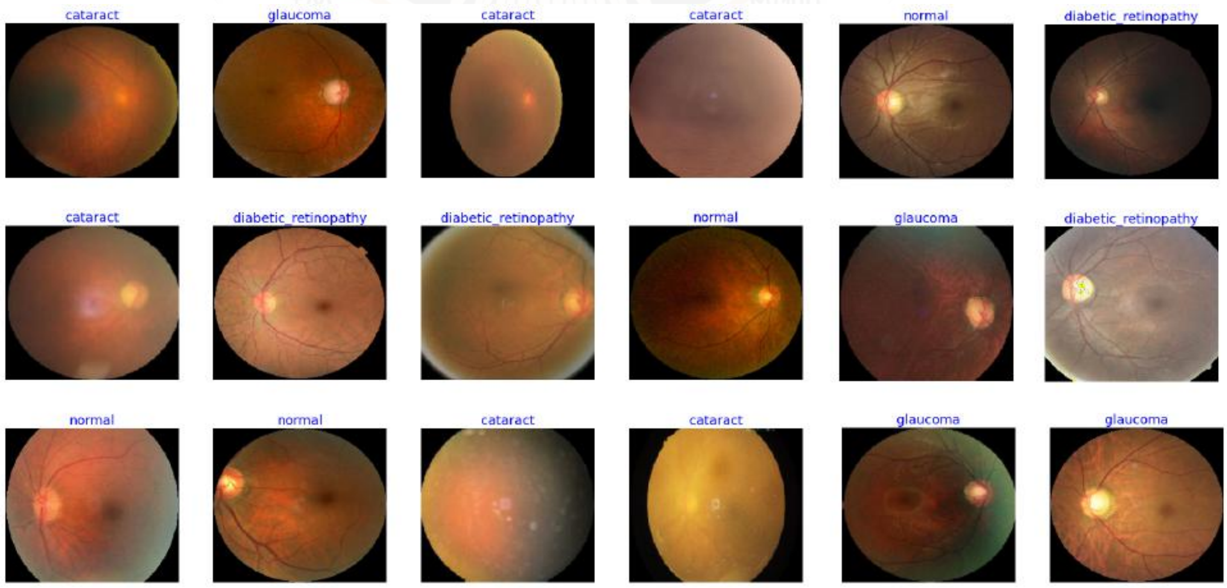


Figure 4. Collected retinal image samples

A total of 2000 images with at least 500 images representing each disease are selected from the collected images. Resizing and background removal operations are performed on these images. Thus, the data set is given its final shape. Then, training, validation and test datasets are created from this dataset.

2.2. **Deep Learning Architectures.** Seven different convolutional neural network models are used in the study. Convolutional neural networks are one of the most used among all deep learning models. It has become quite popular, especially after its successful results in the ImageNet image recognition competition held in 2012 [12]. A CNN consists of three types of layers, these are; convolution, pooling, and fully connected layers. While the convolution and pooling layers perform feature extraction, the fully connected third layer classifies the extracted features [15].

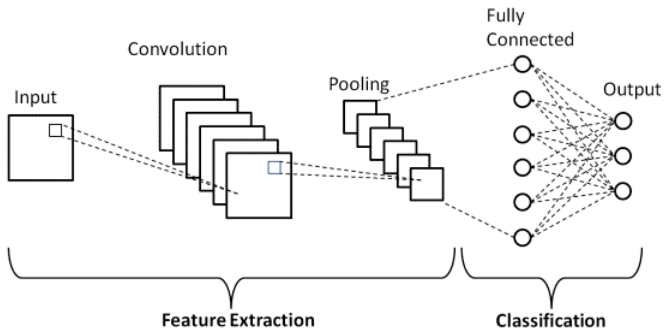


FIGURE 5. Simple a CNN architecture

After the emergence of CNN, it has become one of the most important research topics of many researchers. As a result of these researches, it has continued to be developed continuously. There are CNNs in different architectures in the literature. In this study, Residual Network (ResNet), Densely Connected Convolutional Networks (DenseNe), MobileNet, EfficientNet (EfcN-B3,EfcN-B5,EfcN-B6,EfcN-B7) architectures are used.

- **Residual Network**

Residual Network (ResNET) is a type of convolutional neural network [17] first introduced by Kaiming He et al. in 2015. ResNet uses shortcut links to bypass one or more layers, thus enabling training of extremely deep neural networks. With this feature, it has eliminated the problem of degrading accuracy and vanishing gradient [21].

- **Densely Connected Convolutional Networks**

Densely Connected Convolutional Networks (Densenet) was presented at CVPRC (Computer Vision and Pattern Recognition Conference) in 2017 by Huang et al. and was selected as the best paper [18]. DenseNet is a kind of convolutional neural network architecture [22], consisting of several dense blocks with dense link and transition layers. Each layer takes as input the feature maps of all previous layers [23].

- **MobileNets**

MobileNets are convolutional neural network architectures developed for mobile and embedded vision applications introduced by Google in 2017 [19]. MobileNets has two basic principles: reducing model size by using fewer parameters and reducing complexity by using fewer mathematical operations. Thus, object detection, classification, etc. A very practical convolutional neural network architecture has emerged that can be used for various purposes.

- **EfficientNET**

EfficientNet consists of eight models introduced by Tan and Le in 2019, numbered from simple to complex, from B0 to B7 [20]. EfficientNet uses a technique called composite coefficient to scale models simply and effectively. In most models, width, depth, and resolution dimensions are scaled randomly. In EfficientNet, using unified scaling, each of these three dimensions scales similarly through a fixed set of scaling coefficients.

3. RESULTS

The findings reached with the models created using seven different deep learning architectures are shared in figures 6-10.

RESNET				
	Precision	Recall	F1-Score	ACC
CT	0.9536	0.9231	0.9381	0.9179
DP	1.0000	0.9939	0.9970	
GL	0.8408	0.8742	0.8571	
N	0.8758	0.8758	0.8758	

DENSENET121				
	Precision	Recall	F1-Score	ACC
CT	0.9290	0.9231	0.9260	0.9115
DP	1.0000	0.9939	0.9970	
GL	0.8671	0.8212	0.8435	
N	0.8400	0.9006	0.8735	

MOBILENET				
	Precision	Recall	F1-Score	ACC
CT	0.9416	0.9295	0.9355	0.9163
DP	0.9879	0.9879	0.9879	
GL	0.8725	0.8609	0.8667	
N	0.8606	0.8820	0.8712	

EFFICIENTNET-B3				
	Precision	Recall	F1-Score	ACC
CT	0.9363	0.9423	0.9393	0.9242
DP	1.0000	0.9939	0.9970	
GL	0.8966	0.8606	0.8784	
N	0.8623	0.8944	0.8780	

EFFICIENTNET-B5				
	Precision	Recall	F1-Score	ACC
CT	0.9605	0.9359	0.9481	0.9258
DP	1.0000	0.9879	0.9939	
GL	0.8733	0.8675	0.8704	
N	0.8690	0.9068	0.8875	

EFFICIENTNET-B6				
	Precision	Recall	F1-Score	ACC
CT	0.9733	0.9359	0.9542	0.9289
DP	1.0000	0.9818	0.9908	
GL	0.8766	0.8940	0.8852	
N	0.8683	0.9006	0.8841	

EFFICIENTNET-B7				
	Precision	Recall	F1-Score	ACC
CT	0.9536	0.9231	0.9381	0.9289
DP	1.0000	0.9939	0.9970	
GL	0.8816	0.8874	0.8845	
N	0.8795	0.9068	0.8930	

Figure 6. Classification reports for all models

In the first model, RESNET architecture is used and an average accuracy rate of 91.79% was achieved with this model. The highest accuracy is obtained for DP (ACC:99.70%). Accuracy rates of 93.81% for cataract and 85.71% for glaucoma were achieved.

The average classification success achieved with the model using the Densenet121 architecture is 91.15%. When examined separately for each disease, the accuracy rates are 99.70% for DP, 92.60% for cataracts, and 84.35% for glaucoma.

The model using the Mobilenet architecture achieved a performance between Densenet121 and RESNET (Average ACC: 91.63%, DP: 98.79%, Cataract: 93.55%, Glaucoma: 86.67%).

EfficientNet group architectures are the most successful in terms of performance. Average accuracy rates are 92.42% with EfcN-B3, and 92.58% with EfcN-B5. EfcN-B6 and EfcN-B7 performed better and achieved 92.89% classification success. When analyzed according to diseases, accuracy rates of 99.7% for DP (EfcN-B7), 95.42% for cataract (EfcN-B6), and 88.52% for glaucoma (EfcN-B7) are reached.

RESNET

	CT	DP	GL	N
CT	144	0	7	5
DP	0	164	0	1
GL	5	0	132	14
N	2	0	18	141

DENSENET121

	CT	DP	GL	N
CT	144	0	7	5
DP	0	164	0	1
GL	7	0	124	20
N	4	0	12	145

MOBILENET

	CT	DP	GL	N
CT	145	1	5	5
DP	0	163	0	2
GL	5	0	130	16
N	4	1	14	142

EFFICIENTNET-B3

	CT	DP	GL	N
CT	147	0	3	6
DP	0	164	0	1
GL	5	0	130	16
N	5	0	12	144

EFFICIENTNET-B5

	CT	DP	GL	N
CT	146	0	7	3
DP	0	163	0	2
GL	3	0	131	17
N	3	0	12	146

EFFICIENTNET-B6

	CT	DP	GL	N
CT	146	0	4	6
DP	0	162	1	2
GL	2	0	135	14
N	2	0	14	145

EFFICIENTNET-B7

	CT	DP	GL	N
CT	144	0	6	6
DP	0	164	0	1
GL	4	0	134	13
N	3	0	12	146

CT: Cataract
DP: Diabetic Retinopathy
GL: Glaucom
N: Normal

FIGURE 7. Confusion matrices for all models

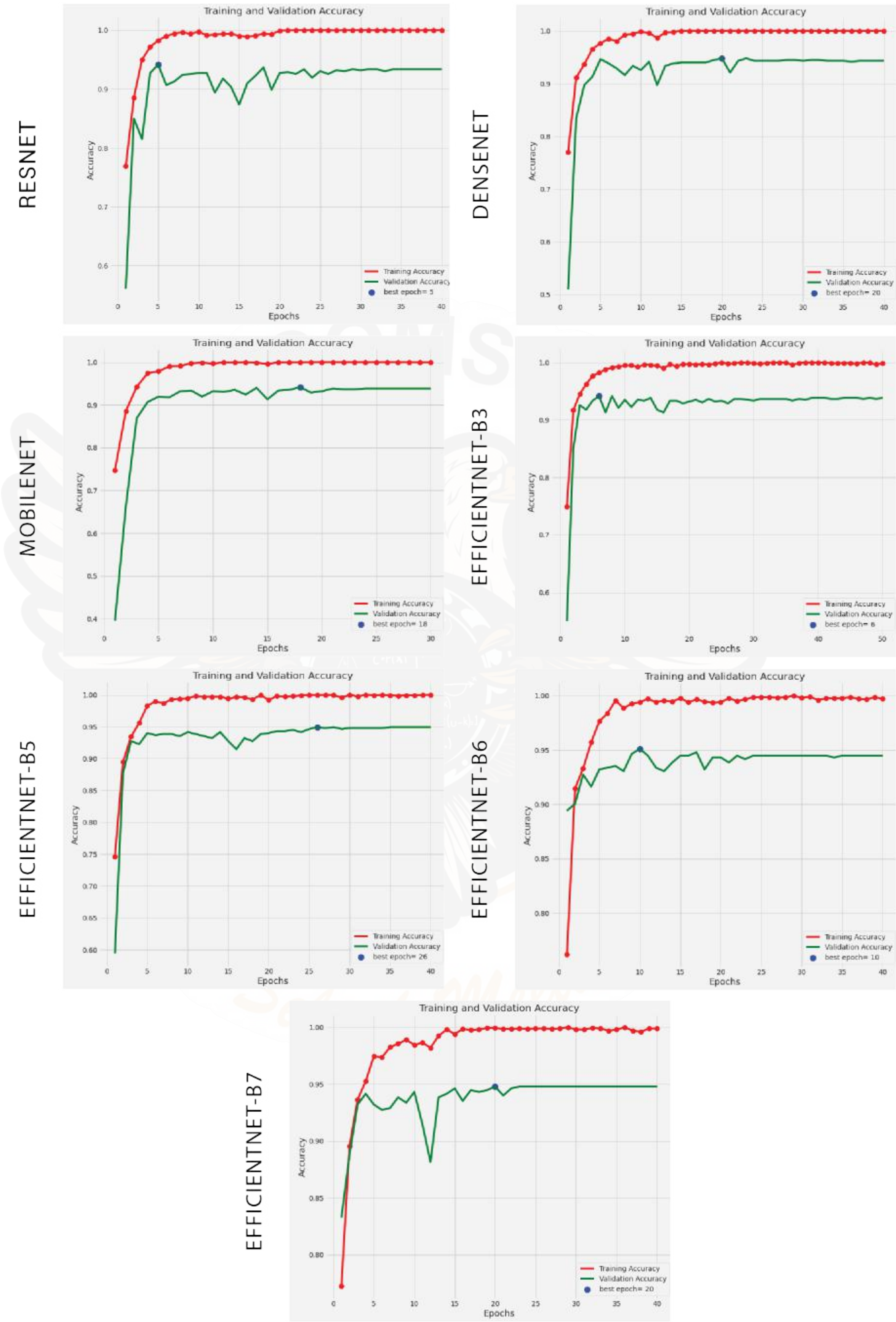


FIGURE 8. Training and validation loss plots for all models

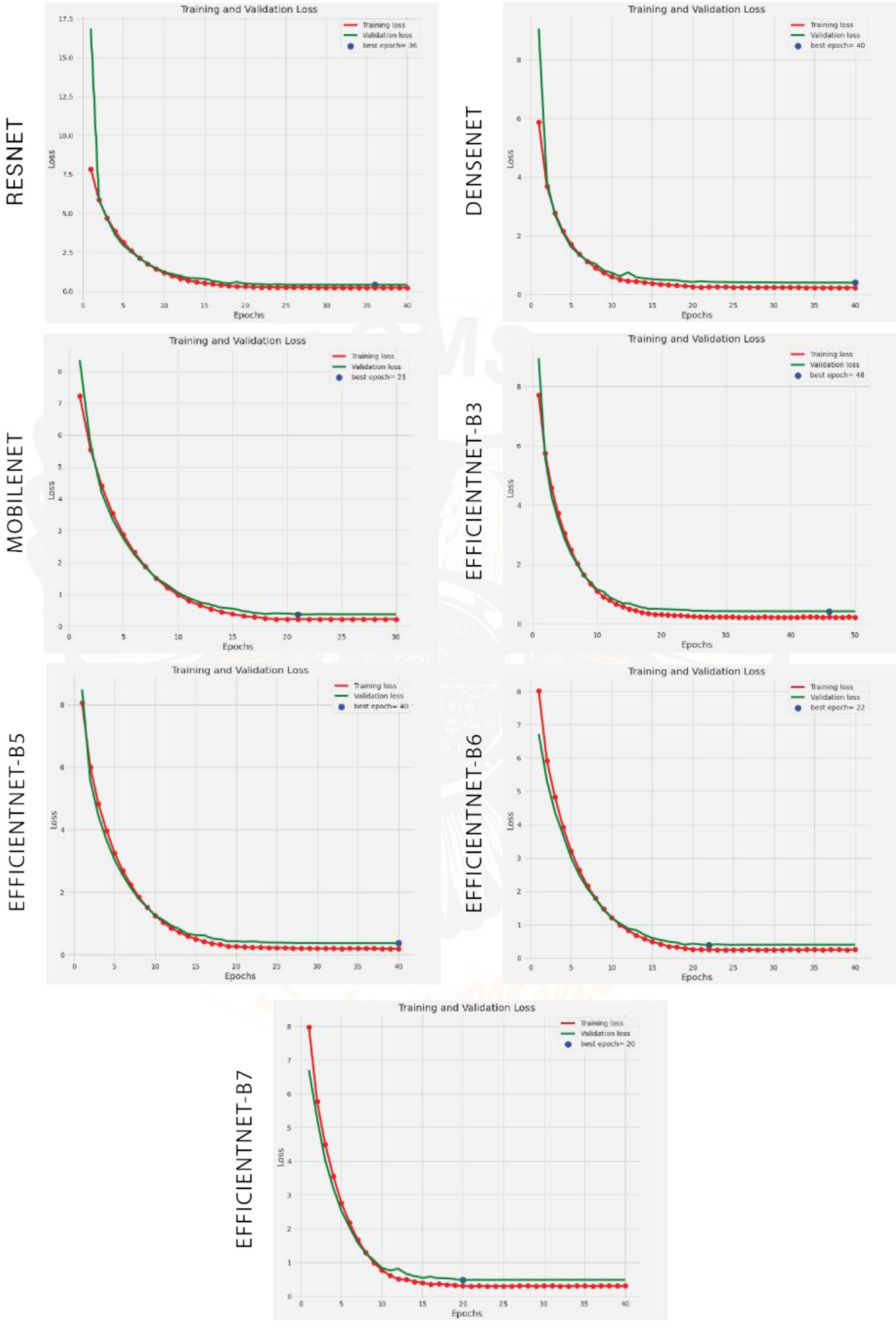


FIGURE 9. Training and validation accuracy plots for all models

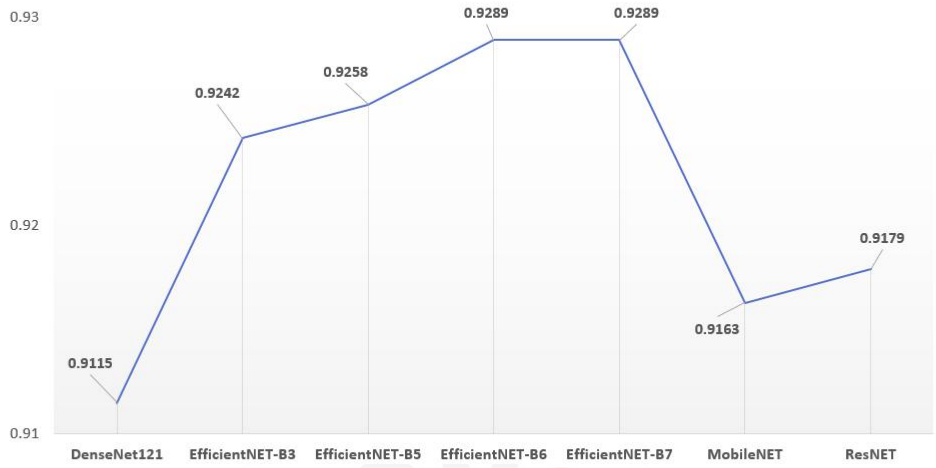


FIGURE 10. Accuracy values for all models

4. Conclusions and Future Studies

In this study, eye diseases such as diabetic retinopathy, cataract and glaucoma are diagnosed using different types of deep learning architectures. For this, retinal images collected from different web sources are used.

According to the findings, EfficientNet has been the most successful deep learning architecture in the diagnosis of eye diseases.. The lowest accuracy is obtained with DenseNET. Nearly one hundred percent success has been achieved with all architectures in the diagnosis of diabetic retinopathy. A relatively lower classification success is achieved in glaucoma compared to other diseases (ACC: 82-89%). Because sometimes there are no symptoms on retinal images in glaucoma. The most successful architecture in the classification of cataract disease is EfficientNet-B3 (ACC: 94.23%). The findings show that a deep learning-based diagnostic model can be quite successful in diagnosing eye diseases.

In future studies, retinal images of patients with different eye diseases will be included in the dataset. Thus, it is aimed to diagnose eye disease more comprehensively. On the other hand, different deep learning architectures can be used or hybrid models can be developed. More successful diagnosis can be made by developing special solutions for glaucoma, which reduces the classification performance. A two-stage classifier would be a good example of these solutions.

References

[1] Abbas, Q. (2017). Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning. International Journal of Advanced Computer Science and Applications, 8(6).

[2] Aalday, M. F. and nar, A. (2021). Derin renme Mimarilerini Kullanarak Katarakt Tespiti. European Journal of Science and Technology, (28), 1428-1433.

[3] Amin, J., Sharif, M., and Yasmin, M. (2016). A review on recent developments for detection of diabetic retinopathy. Scientifica, 2016.

[4] Chai, Y., Liu, H., and Xu, J. (2018). Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. Knowledge-Based Systems, 161, 147-156.

[5] Jain, L., Murthy, H. S., Patel, C., and Bansal, D. (2018, December). Retinal eye disease detection using deep learning. In 2018 Fourteenth International Conference on Information Processing (ICINPRO) (pp. 1-6). IEEE.

[6] Nazir, T., Nawaz, M., Rashid, J., Mahum, R., Masood, M., Mehmood, A., ... and Hussain, A. (2021). Detection of diabetic eye disease from retinal images using a deep learning based CenterNet model. Sensors, 21(16), 5283.

[7] Ocular Disease Recognition Dataset. (2022, 01 Oct.). Access address: https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k

[8] Odstrcilik, J., Kolar, R., Budai, A., Hornegger, J., Jan, J., Gazarek, J., ... and Angelopoulou, E. (2013). Retinal vessel segmentation by improved matched filtering: evaluation on a new highresolution fundus image database. IET Image Processing, 7(4), 373-383.

[9] zelik, Y. B. and Altan, A. (2021). Diyabetik Retinopati Tehisi iin Fundus Grntlerinin Derin renme Tabanl Snflandrlmas. Avrupa Bilim ve Teknoloji Dergisi, (29), 156-167.

[10] Pachade, S., Porwal, P., Thulkar, D., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., ... and Mriaudeau, F. (2021). Retinal fundus multi-disease image dataset (RFMiD): a dataset for multi-disease detection research. Data, 6(2), 14.

[11] Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., and Meriaudeau, F. (2018). Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research. Data, 3(3), 25.

[12] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... and Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. International journal of computer vision, 115(3), 211-252.

[13] Schacknow, P. N., and Samples, J. R. (Eds.). (2010). The glaucoma book: a practical, evidence-based approach to patient care. Springer Science and Business Media.

[14] Uar, M. (2021). Glokom Hastalnn Evriimli Sinir A Mimarileri ile Tespiti. Dokuz Eyll niversitesi Mhendislik Fakltesi Fen ve Mhendislik Dergisi, 23(68), 521-529.

[15] Yamashita, R., Nishio, M., Do, R. K. G., and Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. Insights into imaging, 9(4), 611-629.

[16] Nielsen, M. A. (2015). Neural networks and deep learning (Vol. 25). San Francisco, CA, USA: Determination press.

[17] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[18] Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[19] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

[20] Tan, M., and Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.

[21] Ismael, S. A. A., Mohammed, A., and Hefny, H. (2020). An enhanced deep learning approach for brain cancer MRI images classification using residual networks. Artificial intelligence in medicine, 102, 101779.

[22] Pang, S., Zhang, Y., Ding, M., Wang, X., and Xie, X. (2019). A deep model for lung cancer type identification by densely connected convolutional networks and adaptive boosting. IEEE Access, 8, 4799-4805.

[23] Rubin, J., Parvaneh, S., Rahman, A., Conroy, B., and Babaeizadeh, S. (2018). Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ECG recordings. Journal of electrocardiology, 51(6), S18-S21.

Department of Actuerial Science, Selcuk University, 42250 Konya, Turkey
Email(s): ismailhakkikinali@gmail.com ,

Unbounded Quasi-Normed Convergence in Quasi-Normed Lattices and Some Properties of L_p Spaces for $0 < p < 1$

Enkeleda Zajmi Kotonaj

Abstract. The unbounded norm convergence was first defined by V. Troitsky (2004) in [10] under the name d convergence. The name unbounded norm convergence was introduced by Y. Deng, M. OBrien and V. Troitsky (2017) in [3], where they studied basic properties of unbounded norm convergence.
In this paper, the initial purpose is to extend the concept of unbounded convergence in quasi-normed spaces which is named unbounded quasi - norm convergence and to study some of the basic properties of unbounded quasi-normed convergence in quasi-normed lattices. In assuming that (X, Σ, μ) is finite measurable space, we will discuss also some properties of L_p spaces for $0 < p < 1$ where most important results are the generalizations of dominated convergence theorem and Brzis - Lieb lemma.

2020 Mathematics Subject Classifications: 46A16, 46A40 , 46B40, 46E30, 47B60

Keywords: unbounded quasi-normed convergence, quasi-normed lattices, L_p spaces.

1. Introduction

The unbounded norm convergence was first defined by V. Troitsky (2004) in [10] under the name dconvergence. He studied the relationship between the dconvergence and the measure of non-compactness. The name unbounded norm convergence (shortly un-convergence) was introduced by Y. Deng, M. OBrien and V. Troitsky (2017) in [3], where they studied basic properties of unbounded norm convergence and investigated its relation with unbounded order-and weak convergences. Finally, they showed that un-convergence is topological.
In this paper, the initial purpose is to extend the concept of unbounded convergence in quasi-normed spaces which is named unbounded quasi - norm convergence (shortly uqn-convergence) and to study some of basic properties of unbounded quasi-normed convergence in quasi-normed lattices. The structure of the paper is as follows: We begin with some preliminaries. Our main results are presented in sections 2 and 3. Section 2 contain some basic properties of unbounded quasi-normed convergence in quasi-normed lattices, among which a relation between unbounded order convergence and uqn-convergence in an order continuous quasi-Banach lattice. Similar to the theorem 4.3 in [6] we give a theorem about extended of uqn-convergence of nets from a sublattice to corresponding lattice. Furthermore, if X be an order continuous Banach lattice, then unbounded order convergence of a net implies norm convergence (Proposition 3.7 [5]). In the same way, we have formulated and prove a proposition in case of order continuous quasi-Banach lattice.
Quasi-normed and p-normed spaces are presented at [9] (chapter I.1) and it is stated that, for every quasi-norm function on X we can find an equivalent p-norm function (Aoki-Rolewicz theorem). At the beginning of section 3, using the properties of absolute value and Aoki-Rolewicz theorem we noticed that L_p spaces for $0 < p < 1$ are quasi-normed spaces if same as in the case of L_p spaces for $p > 1$, we equip them with the function $\| \, . \, \|_p$: $L_p \to [0, +\infty)$ such that, $\forall f \in L_p, \| f \|_p = (\int_X \mid f \mid^p d\mu)^{\frac{1}{p}}$. We also notice that these spaces are vector lattices and furthermore, if the measurable space X is finite (definition 2.2 (i), [8]), then L_p spaces for $0 < p < 1$ are quasi-Banach lattices. Started from theorem 2.7 in [7] we have formulated a theorem on convergence of functional series in $L_p(X)$ spaces where $0 < p < 1$ and $0 < \mu(X) < +\infty$. Continuing further, is denoted $L_0(\mu)$ vector lattice of real-valued measurable functions on X modulo almost everywhere (a.e), where (X, Σ, μ) is a measurable space. Unbounded order convergence of a sequence in $L_0(\mu)$ space is equivalent with almost

everywhere convergence (Proposition 3.1 [4]) and we have proof that unbounded order convergence of a sequence in $L_p(X)$ spaces where $0 < p < 1$ and $0 < \mu(X) < +\infty$, is equivalent with uqn-convergence. This enables us to prove the dominated convergence theorem in these spaces. A characterization of regular sublattices via order convergence is given in theorem 1.20 of [1]. Here is another property of regular sublattices in $L_p(X)$ spaces with $0 < p < 1$ and $0 < \mu(X) < +\infty$. We finish this section with Bresiz - Lieb lemma in case of these spaces. This is a generalization of the lemma of the same name in [2].

1.1. **Preliminaries.** A real vector space X which is ordered by some order relation $\leq$ is called a vector lattice if any two elements $x, y \in X$ have a least upper bound denoted by $x \vee y = sup(x, y)$ and a greatest lower bound denoted by $x \wedge y = inf(x, y)$ and the following properties are satisfied:
i) $x \leq y$ implies $x + z \leq y + z$ for all $x, y, z \in X$.
ii) $0 \leq x$ implies $0 \leq tx$ for all $x \in E$ and $t \in R^+$.

Note that:$X_+ = \{x \in X : x \geq 0\}$ is called positive cone in X, $x^+ = x \vee 0, x^- = -x \vee 0, \mid x \mid = x^+ + x^-$.

Let X be a vector space. A function $\parallel . \parallel: X \to [0, +\infty)$ is said to be quasi-norm on X if the following conditions hold:
(1) For every $x \in X$, $\parallel x \parallel \geq 0$ and $\parallel x \parallel = 0 \Leftrightarrow x = 0$.
(2)For every $x \in X$ and for every $\lambda \in R$, $\parallel \lambda x \parallel = |\lambda| \parallel x \parallel$.
(3) For every $x, y \in X$, $\parallel x + y \parallel \leq K(\parallel x \parallel + \parallel y \parallel)$ where $K \geq 1$ is a constant independent of variables x and y.

The smallest of constant K, such that the above conditions hold, is called the modulus of concavity of quasi-norm $\parallel . \parallel$.

If the condition (3) above replaced with (3') exists $0 < p \leq 1$ such that for every $x, y \in X$, $\parallel x + y \parallel^p \leq \parallel x \parallel^p + \parallel y \parallel^p$, then the function $\parallel . \parallel$ is called p - norm.

If the vector space X is equipment with a quasi - norm (p - norm) $\parallel . \parallel$, then $(X, \parallel . \parallel)$ is called quasi - normed (p - normed) space.

A quasi -normed (p - normed) space X is called a quasi-Banach (p - Banach) space if it is complete, which means that a sequence $(x_n)_{n \in \mathbb{N}}$ in X is convergent if and only if $\parallel x_n - x_m \parallel \to 0$ as $m, n \to +\infty$. (see [9])

A quasi-Banach (quasi-normed , p - Banach) space $(X, \parallel . \parallel)$ is called a quasi-Banach lattice (respectively, quasi-normed lattice, p - Banach lattice) if, in addition, X is a vector lattice and $\mid x \mid \leq \mid y \mid$ implies $\parallel x \parallel \leq \parallel y \parallel$ for all $x, y \in X$ ([7] Definition 2.5).

A directed set $(I, \leq)$ consist of a set I with a partial order such that for every pair $\alpha, \beta \in I$, there exists an element $\delta \in I$ with $\delta \geq \alpha$ and $\delta \geq \beta$.
A net $(x_\alpha)_{\alpha \in I}$ in X is a function $I \to X$ sucht that $\alpha \to x_\alpha$ for all $\alpha \in I$ and $(I, \leq)$ is a directed set. A net $(y_\beta)_{\beta \in J}$ is a subnet of the net $(x_\alpha)_{\alpha \in I}$ if $y_\beta = x_{\phi(\beta)}$ for some function $\phi : J \to I$ such that for every $\alpha_0 \in I$, there exists an element $\beta_0 \in J$ for which $\beta > \beta_0$ implies $\phi(\beta) > \alpha_0$.
According to [5], a net $(x_\alpha)_{\alpha \in I}$ in a vector lattice X is said to be order convergent to $x \in X$, written as $x_\alpha \xrightarrow{o} x$ if there exists another net $(y_\beta)_{\beta \in J}$ in X satisfying $y_\beta \downarrow 0$ and for any $\beta \in J$ there exists $\alpha_0 \in I$ such that $\mid x_\alpha - x \mid \leq y_\beta$ for all $\alpha \geq \alpha_0$. A net $(x_\alpha)_{\alpha \in I}$ in a vector lattice X is said to be unbounded order convergent to $x \in X$, written as $x_\alpha \xrightarrow{uo} x$ if $\mid x_\alpha - x \mid \wedge u \xrightarrow{o} 0$ for all $u \in X_+$.

The unbounded norm convergence was first defined by V. Troitsky (2004) in [10]. A net $(x_\alpha)_{\alpha \in I}$ in a normed lattice is said to be unbounded norm convergent to $x \in X$, written as $x_\alpha \xrightarrow{un} x$ if $\parallel \mid x_\alpha - x \mid \wedge u \parallel \to 0$ for all $u \in X_+$.

In the same way we can definition the unbounded quasi-norm convergence. The only difference with the definition above is that $\| \,.\, \|$ is a quasi-norm and written as $x_\alpha \xrightarrow{uqn} x$.

The meaning of order continuous (σ - order continuous) normed lattice is given in [5]. Here we are giving these meanings in quasi-normed lattice:

A quasi-normed lattice $(X, \| \,.\, \|)$ is called order continuous (order σ - continuous) if a net (sequence) $x_\alpha \downarrow 0$ ($x_n \downarrow 0$) in X implies $\| x_\alpha \| \downarrow 0$ ($\| x_n \| \downarrow 0$).

A vector subspace Y of a vector lattice X is said to be a sublattice of X if for each y_1 and y_2 in Y we have $y_1 \vee y_2 \in Y$. A sublattice Y of X is order dense in X if for each $x > 0$ there is $0 < y \in Y$ with $0 < y \leq x$. Y is said to be majorizing in X if for each $x \in X_+$ there exists $y \in Y$ such that $x \leq y$. A sublattice Y of X is called regular if for any subset A of Y, if infA exists in Y, then infA exists in X and the two infinums are equal. We recall a characterization of regular sublattices via order convergence.

For a sublattice Y of a vector lattice X the following statements are equivalent.

1. Y is regular.
2. If (y_α) is a net in Y such that $y_\alpha \downarrow 0$ in Y, then $y_\alpha \downarrow 0$ in X.
3. If (y_α) is a net in Y such that $y_\alpha \xrightarrow{o} y$ in Y, then $y_\alpha \xrightarrow{o} y$ in X ([1] Theorem 1.20).

An interesting characterization of regular sublattices in term of uo-convergence was established in [4].

Let be Y a sublattice of a vector lattice X. The following statements are equivalent:

1. Y is regular.
2. For any net (y_α) in Y, $y_\alpha \xrightarrow{uo} 0$ in Y iff $y_\alpha \xrightarrow{uo} 0$ in X ([4] Theorem 3.2).

A subset Y of X is said to be solid if for $x \in X$ and $y \in Y$ such that $| x | \leq | y |$ it follows that $x \in Y$. A solid vector subspace of a vector lattice is called an ideal. Ideals are automatically vector sublattices since $| x \vee y | \leq | x | + | y |$. The ideal generated by $e \in X$ is called the smallest ideal that includes e. A vector $0 < e \in X$ is called a quasi-interior point if $\overline{I_e} = X$, where I_e denotes the ideal generated by e (see [5]). It can be shown that (see [6]) e is a quasi-interior point iff for every $x \in X_+$ we have $\| x - x \wedge ne \| \to 0$ for $n \to +\infty$.

2. Basic properties of uqn-convergence

The name unbounded norm convergence (shortly un-convergence) was introduced by Y. Deng, M. OBrien and V. Troitsky (2017) in [3], where they studied basic properties of unbounded norm convergence. We are extending the concept of unbounded convergence in quasi-normed spaces which is named unbounded quasi - norm convergence (shortly uqn-convergence). Let's list some of its basic properties first.

The following inequality is routinely used afterwards:

$$(1)\ (x + y) \wedge u \leq x \wedge u + y \wedge u \text{ for all } x, y, u \in X_+.$$

Theorem 2.1. *Let X be a quasi-normed lattice. The following propositions are satisfied:*
i) $x_\alpha \xrightarrow{uqn} x$ *iff* $(x_\alpha - x) \xrightarrow{uqn} 0$.
ii) If $x_\alpha \xrightarrow{uqn} x$, *then* $y_\beta \xrightarrow{uqn} x$ *for every subnet* (y_β) *of* (x_α).
iii) Suppose $x_\alpha \xrightarrow{uqn} x$ *and* $y_\alpha \xrightarrow{uqn} y$. *Then* $(ax_\alpha + by_\alpha) \xrightarrow{uqn} (ax + by)$ *for any* $a, b \in \mathbb{R}$.
iv) If $x_\alpha \xrightarrow{uqn} x$ *and* $x_\alpha \xrightarrow{uqn} y$, *then* $x = y$.
v) If $x_\alpha \xrightarrow{uqn} x$, *then* $| x_\alpha | \xrightarrow{uqn} | x |$.

Proof

Proposition i) and ii) are straightforward.

iii) From properties of vector lattice, we have:

(1) $| (ax_\alpha + by_\alpha) - (ax + by) | \leq | ax_\alpha - ax | + | by_\alpha - by | = | a | \cdot | x_\alpha - x | + | b | \cdot | y_\alpha - y |$

So, we can write:

(2) $| (ax_\alpha + by_\alpha) - (ax + by) | \wedge u \leq (| a | \cdot | x_\alpha - x |) \wedge u + (| b | \cdot | y_\alpha - y |) \wedge u$ *for all* $u \in X_+$.

Because, we can prove that: If $x \leq y$, *then* $\forall z \in X$, $x \wedge z \leq y \wedge z$.

Let see more precisely. Since $x \wedge z = inf(x, z)$, *then* $x \wedge z \leq x$ *and* $x \wedge z \leq z$. *Thus,*

$x \wedge z \leq x \leq y$ and $x \wedge z \leq z$ that imply $x \wedge z \leq y \wedge z$.
From inequality (2) and the above fact, we write:

$$\begin{gathered}\| | (ax_\alpha + by_\alpha) - (ax+by) | \wedge u \| \leq \| (| a | \cdot | x_\alpha - x |) \wedge u + (| b | \cdot | y_\alpha - y |) \wedge u \| \\ \leq K \cdot \| (| a | \cdot | x_\alpha - x |) \wedge u \| + K \cdot \| (| b | \cdot | y_\alpha - y |) \wedge u \| = \\ K \cdot \| (| a | \cdot | x_\alpha - x |) \wedge (| a | \cdot \tfrac{u}{|a|}) \| + K \cdot \| (| b | \cdot | x_\alpha - x |) \wedge (| b | \cdot \tfrac{u}{|b|}) \| = \\ K \cdot | a | \cdot \| | x_\alpha - x | \wedge \tfrac{u}{|a|} \| + K \cdot | b | \cdot \| | y_\alpha - y | \wedge \tfrac{u}{|b|} \| = \\ K \cdot | a | \cdot \| | x_\alpha - x | \wedge u' \| + K \cdot | b | \cdot \| | y_\alpha - y | \wedge u'' \|\end{gathered}$$

where $u \in X_+$, $u' = \frac{u}{|a|}$, $u'' = \frac{u}{|b|}$ and $a, b \neq 0$.
So, $u', u'' \in X_+$ and since $x_\alpha \xrightarrow{uqn} x$ and $y_\alpha \xrightarrow{uqn} y$, we have:
$\| | (ax_\alpha + by_\alpha) - (ax+by) | \wedge u \| \leq K \cdot | a | \cdot \| | x_\alpha - x | \wedge u' \| + K \cdot | b | \cdot \| | y_\alpha - y | \wedge u'' \| \to 0.$
If $a = 0$ or $b = 0$, then

$$\begin{gathered}\| | (ax_\alpha + by_\alpha) - (ax+by) | \wedge u \| = \| | ax_\alpha - ax | \wedge u \| \text{ or} \\ \| | (ax_\alpha + by_\alpha) - (ax+by) | \wedge u \| = \| | by_\alpha - by | \wedge u \| \text{ or} \\ \| | (ax_\alpha + by_\alpha) - (ax+by) | \wedge u \| = \| 0 \| = 0.\end{gathered}$$

Now the proof is clear.
The proof of (iv) and (v) properties is in the same way with lemma 2.1 (iv) and (v) of [3].

It is immediately from definitions of quasi-Banach lattice, uo-convergence and uqn-convergence that: In an order continuous quasi-Banach lattice, the uo-convergence implies uqn-convergence.

Lemma 2.2. *Let X be a quasi-Banach lattice with a quasi-interior point e. Then $x_\alpha \xrightarrow{uqn} 0$ iff $| x_\alpha | \wedge e \xrightarrow{\|\cdot\|} 0$.*

The proof is identically with lemma 2.11 in [3].

In similiar with theorem 4.3 in [6], we can formulate and proof the following theorem:

Theorem 2.3. *Let Y be a sublattice of a quasi-normed lattice X and (y_α) a net in Y such that $y_\alpha \xrightarrow{uqn} 0$ in Y. The following statements hold:*
i) If Y is majorizing in X, then $y_\alpha \xrightarrow{uqn} 0$ in X.
ii) If Y is quasi-norm dense in X, then $y_\alpha \xrightarrow{uqn} 0$ in X.
Proof
i) The proof is immediately from definitions of uqn-convergence and quasi-normed lattice.
ii)Take $u \in X_+$. Since Y is quasi-norm dense in X, there is a $v \in Y_+$ such that $\| u - v \| < \frac{\varepsilon}{2K}$, where K is modulus of concavity of quasi-norm.
By assumption, $y_\alpha \wedge v \xrightarrow{\|\cdot\|} 0$. So, we can find α_0 such that $\| y_\alpha \wedge v \| < \frac{\varepsilon}{2K}$, for every $\alpha \geq \alpha_0$. It follows from $u \leq v + | u - v |$ that $y_\alpha \wedge u \leq y_\alpha \wedge v + | u - v |$. So that $\| y_\alpha \wedge u \| \leq K \cdot \| y_\alpha \wedge v \| + K \cdot \| u - v \| < K \cdot \frac{\varepsilon}{2K} + K \cdot \frac{\varepsilon}{2K} = \varepsilon$, whenever $\alpha \geq \alpha_0$. It follows that $y_\alpha \wedge u \xrightarrow{\|\cdot\|} 0$. Hence $y_\alpha \xrightarrow{uqn} 0$ in X.

In the same way as in corollary 4.6 in [6], we can see that the following corollary is true.

Corollary 2.4. *Let Y be a sublattice of an order continuous quasi-Banach lattice X. If $y_\alpha \xrightarrow{uqn} 0$ in Y, then $y_\alpha \xrightarrow{uqn} 0$ in X.*

3. Some properties of L_p spaces for $0 < p < 1$

The L_p spaces for $0 \leq p < +\infty$, where L_0 is denoted the space of real-valued measurable functions on X modulo almost everywhere (a.e) and L_p for $p > 0$ are the space of p-integrabled functions on X, are vector lattices. In these spaces, the uo-convergence is the same as a.e-convergence (proposition 3.1 and remark 3.4 in [4]).
As we know, the L_p spaces for $1 \leq p < +\infty$ are normed spaces (recall that: the norm function is $\| . \|_p$: $L_p \to [0, +\infty)$ such that, $\forall f \in L_p, \| f \|_p = (\int_X | f |^p \, d\mu)^{\frac{1}{p}}$).

Lemma 3.1. *If $0 < p < 1$, then the above function $\| . \|_p$ is a quasi-norm.*
Proof
First let proof that, $\forall f, g \in L_p, \| f + g \|^p \leq \| f \|^p + \| g \|^p$.

Let x and y be two real numbers. If $x > 0$ and $y < 0$ ($x < 0$ and $y > 0$), then
(i) $\mid x \mid > \mid y \mid$ imply $x > -y$. So, $\mid x + y \mid = x + y < x = \mid x \mid$.
(ii) $\mid x \mid < \mid y \mid$ imply $x < -y$. So, $\mid x + y \mid = -x - y < -y = \mid y \mid$.
(iii) $\mid x \mid = \mid y \mid$ imply $\mid x + y \mid = 0 < \mid x \mid$ or $\mid y \mid$.
So, $\mid x + y \mid \leq \mid x \mid$ or $\mid x + y \mid \leq \mid y \mid$ that imply $\mid x + y \mid^p \leq \mid x \mid^p$ or $\mid x + y \mid^p \leq \mid y \mid^p$. Thus, $\mid x + y \mid^p \leq \mid x \mid^p + \mid y \mid^p$.
Similarly we can prove the same result for $x < 0$ and $y > 0$.
If $x > 0$ and $y > 0$, then the following inequality are equivalent.
(1) $(x + y)^p \leq x^p + y^p$, (2) $x^p(1 + \frac{y}{x})^p \leq x^p(1 + (\frac{y}{x})^p)$ and (3) $(1 + \frac{y}{x})^p - 1 - (\frac{y}{x})^p \leq 0$.
Let see the function $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = (1 + x)^p - 1 - x^p$. This function is decreasin for $x > 0$ that imply $f(x) < f(0) = 0$. So, we proof that te inequality (3) is true.
If $x < 0$ and $y < 0$, then the inequality $(x + y)^p \leq x^p + y^p$ has the form $(-x - y)^p \leq (-x)^p + (-y)^p$. The last inequality is equivalent with $(-x)^p(1 + \frac{y}{x})^p \leq (-x)^p[1 + (\frac{y}{x})^p]$. Since $\frac{y}{x} > 0$, we can continue in the same way with above case (the case:$x > 0$ and $y > 0$).
If $x = 0$ and $y = 0$, then the inequality (1) is clear.
So, we prove that: For every two real numbers x and y, $(x + y)^p \leq x^p + y^p$.
Inequality (1) and properties of Lebesgue integral allow us to write:
For every $f, g \in L_p$, where $0 < p < 1$,
$\| f + g \|_p = (\int_X \mid f + g \mid^p d\mu)^{\frac{1}{p}} \leq [(\int_X \mid f \mid^p d\mu) + (\int_X \mid g \mid^p d\mu)]^{\frac{1}{p}} = (\| f \|_p^p + \| g \|_p^p)^{\frac{1}{p}}$
$\leq \{2max(\| f \|_p^p, \| g \|_p^p)\}^{\frac{1}{p}} = 2^{\frac{1}{p}} \cdot max((\| f \|_p^p)^{\frac{1}{p}} + (\| g \|_p^p)^{\frac{1}{p}}) = 2^{\frac{1}{p}} \cdot (\| f \|_p + \| g \|_p)$.
Thus, the function $\| \,.\, \|_p$ fulfills the third condition of the p-norm and quasi-norm. The fulfillment of the first and second conditions is clear. Therefore the function $\| \,.\, \|_p$ is a p-norm and a quasi-norm simultaneously.

Proposition 3.2. *The function space L_p ($0 < p < +\infty$) is a vector lattice.*
Proof
Every function $f \in L_p(\mu)$ for $0 < p < +\infty$ is μ-measurable and $\int_X \mid f \mid^p d\mu < +\infty$. So for every $f, g \in L_p$, we can write:
If $f \geq g$ and $f, g \geq 0$ or $f > 0$ and $g < 0$, then $sup(f, g) \geq 0$ (the same can be said when $g \geq f$ and $f, g \geq 0$ or $g > 0$ and $f < 0$). So, (4) $sup(f, g)^+ = sup(f^+, g^+)$.
If $f \geq g$ and $f, g < 0$ ($g \geq f$ and $f, g < 0$), then $sup(f, g) < 0$. So, $sup(f, g)^- = -f \leq -g = g^-$ ($sup(f, g)^- = -g \leq -f = f^-$) that imply (5) $sup(f, g)^- = inf(f^-, g^-)$.
So from equations (4) and (5) we have (6) $\mid sup(f, g) \mid = sup(f, g)^+ + sup(f, g)^- = sup(f^+, g^+) + inf(f^-, g^-) \leq f^+ + g^+ + f^- + g^- = (f^+ + f^-) + (g^+ + g^-) = \mid f \mid + \mid g \mid$.
On the other hand, from inequality (6), we have:
for every $p > 0$, $\mid sup(f, g) \mid^p \leq (\mid f \mid + \mid g \mid)^p \leq (2sup(\mid f \mid, \mid g \mid))^p = 2^p \cdot sup(\mid f \mid, \mid g \mid)^p \leq 2^p \cdot (\mid f \mid^p + \mid g \mid^p)$ (7).
Therefore, from inequality (7), we can write:

$$\int_X \mid sup(f, g) \mid^p d\mu \leq 2^p \cdot (\int_X \mid f \mid^p d\mu + \int_X \mid g \mid^p d\mu) < +\infty \ (8)$$

The last inequality implies that $sup(f, g) \in L_p$.
For every $f, g \in L_p$, we can write:
If $f, g \geq 0$ and $f^+ = f \leq g = g^+$ ($g^+ = g \leq f = f^+$), then $0 \leq f = inf(f, g) = inf(f, g)^+$ ($0 \leq g = inf(f, g) = inf(f, g)^+$). So, $inf(f, g)^+ = inf(f^+, g^+)$.
If $f < 0$ or $g < 0$, then $0 = f^+ = inf(f, g)^+$. Whereas, $inf(f^+, g^+) = inf(0, g) = inf(f^+, g^+)$ for $g \geq 0$ or $inf(f^+, g^+) = inf(f, 0) = inf(f^+, g^+)$ for $f \geq 0$ or $inf(f^+, g^+) = inf(0, 0) = inf(f^+, g^+)$ for $f, g < 0$. Thus we have (9) $inf(f, g)^+ = inf(f^+, g^+)$, for every $f, g \in L_p$.
In the same way, we can see that:
The inequality $inf(f, g) < 0$ is true, iff $f < 0$ and $f \leq g$ ($g < 0$ and $g \leq f$). So, $-f = inf(f, g)^- > -g \geq g^-$ ($-g = inf(f, g)^- > -f \geq f^-$) and therefore $inf(f, g)^- = sup(f^-, g^-)$.
The inequality $inf(f, g) \geq 0$ is true, iff $f, g \geq 0$. So, $0 = inf(f, g)^- = sup(0, 0) = sup(f^-, g^-)$.
Thus we have (10) $inf(f, g)^- = sup(f^-, g^-)$, for every $f, g \in L_p$.
From equalities (9) and (10), we can write: $\mid inf(f, g) \mid = inf(f, g)^+ + inf(f, g)^- = inf(f^+, g^+) + sup(f^-, g^-) \leq f^+ + g^+ + f^- + g^- = (f^+ + f^-) + (g^+ + g^-) = \mid f \mid + \mid g \mid$.
The last inequality brings us that: for every $f, g \in L_p$, $inf(f, g) \in L_p$.

Also, conditions (i) and (ii) of vector lattice, hold because every p-integrable function is μ-measurable and $L_0(\mu)$ is a vector lattice.

On the other hand, $| f |<| g |$ bring us $| f |^p<| g |^p$ and from the properties of Lebesgue integral we can write $\| f \|_p \leq \| g \|_p$.

Corollary 3.3. *Every L_p space for $0 < p < 1$ is a quasi-normed lattice.*

Let prove the following proposition:

Proposition 3.4. *If X is a finite measurable space, then the L_p space for $0 < p < 1$ is a quasi-Banach lattice.*

Proof

Let f_n be a Cauchy sequence in L_p space for $0 < p < 1$ and X is a finite measurable space. For every $\varepsilon > 0$, we find a natural number $k(\varepsilon)$ such that:

$$\| f_n - f_m \|_p = (\int_X | f_n - f_m |^p \, d\mu)^{\frac{1}{p}} < \varepsilon, \textit{ for every } m, n > k(\varepsilon).$$

If $| f_n - f_m |^p \geq \frac{\varepsilon^p}{\mu(X)}$, then $(\int_X | f_n - f_m |^p \, d\mu)^{\frac{1}{p}} \geq \frac{\varepsilon}{\mu(X)^{\frac{1}{p}}} \cdot (\int_X d\mu)^{\frac{1}{p}} = \varepsilon$. So, $| f_n - f_m |^p < \frac{\varepsilon^p}{\mu(X)}$, for every $m, n > k(\varepsilon)$. This means that f_n is a Cauchy sequence in real line $\mathbb{R}$.

Let denote $f(x) = \lim_{n\to+\infty} f_n(x)$, for every $x \in X$. Let's prove that the function $f \in L_p$ space.

Let fix a natural number n such that $| f_n - f |< \varepsilon$.

Since $| f | \leq | f_n - f | + | f_n |$, then $| f |^p \leq (| f_n - f | + | f_n |)^p \leq (2max(| f_n - f |, | f_n |))^p \leq 2^p \cdot (| f_n - f |^p + | f_n |^p) < 2^p \cdot (\varepsilon^p + | f_n |^p)$.

Thus $\int_X | f |^p \, d\mu \leq 2^p \cdot (\varepsilon^p \cdot \mu(X) + \int_X | f_n |^p \, d\mu) < +\infty$. This means that $f \in L_p$.

Finally, let's prove that $f(x) = \lim_{n\to+\infty} f_n(x)$ in $f \in L_p$ space (that means, according to the quasi-norm).

We can see that:For every $\varepsilon > 0$, exist a natural number n_0 such that, for every $n > n_0$ we have $\int_X | f_n - f |^p \, d\mu < \varepsilon^p \cdot \int_X d\mu = \varepsilon^p \cdot \mu(X)$.

This imply $\| f_n - f_m \| = (\int_X | f_n - f |^p \, d\mu)^{\frac{1}{p}} < \varepsilon \cdot (\mu(X))^{\frac{1}{p}} = \varepsilon'$. So, $f(x) = \lim_{n\to+\infty} f_n(x)$ in $f \in L_p$ and this complete te proof.

While $| f | \leq | g |$ implies $\| f \| \leq \| g \|$, for all $f, g \in L_p$, the fact that the modulus of concavity of quasi-norm's L_p is $2^{\frac{1}{p}}$ and from theorem 2.7 in [7] we can formulate this theorem.

Theorem 3.5. *For every sequence of functions f_n in L_p, ($0 < p < 1$), where X is finite measurable space ($0 < \mu(X) < +\infty$), such that $\sum_{n=1}^{+\infty} 2^{\frac{n}{p}} \cdot \| f_n \|_p < +\infty$, there exists $f \in L_p$ with $f = \sum_{n=1}^{+\infty} f_n$.*

From properties of Lebesgue integral, we can prove easily that:

Proposition 3.6. *The space L_p ($0 < p < 1$), where X is finite measurable space, is an ordered continuous (ordered σ- continuous) quasi-Banach lattice.*

Another intersting result is given below.

Proposition 3.7. *The sequence of functions f_n in L_p, ($0 < p < 1$), where X is finite measurable space ($0 < \mu(X) < +\infty$), is unbounded order convergent to function $f \in L_0$ if and only if the sequence f_n is unbounded quasi-norm convergent to f.*

Proof

Suppose that $f_n \xrightarrow{uo} f$. For every function $g \geq 0$ we have $\lim_{n\to+\infty}(| f_n - f | \wedge g) = 0$. So, for every $\varepsilon > 0$, exist a natural number n_0 such that for every $n > n_0$, $| f_n - f | \wedge g < \varepsilon$.

From equality $\|| f_n - f | \wedge g \|_p^p = \int_X (| f_n - f | \wedge g)^p d\mu$, we have: $\|| f_n - f | \wedge g \|_p^p < \varepsilon^p \cdot \int_X d\mu = \varepsilon^p \cdot \mu(X) = \varepsilon'$ for every $n > n_0$. Thus, we have prove that $f_n \xrightarrow{uqn} f$.

Let us show the opposite statement. Suppose that $f_n \xrightarrow{uqn} f$. For every $\varepsilon > 0$, exist a natural number n_0 such that for every $n > n_0$, $\|| f_n - f | \wedge g \|_p < \varepsilon$ for every $g > 0$. With similar reasoning as in the propostion 3.4, we conclude that:

For every $\varepsilon > 0$, exist a natural number n_0 such that for every $n > n_0$, $| f_n - f | \wedge g < \varepsilon$ that is equivalent with $f_n \xrightarrow{uo} f$.

Now we can prove the Dominated convergence theorem.

Theorem 3.8. *(Dominated convergence theorem) Let f_n be a sequence of functions in L_p, $(0 < p < 1)$, where X is finite measurable space ($0 < \mu(X) < +\infty$). If f is a measurable function such that $f_n \xrightarrow{uqn} f$ and exists a nonnegative function g in $L_1 = L(\mathbb{R})$ (with $L(\mathbb{R})$ is denoted the Lebesgue integration real valued function space) such that, for every $n \in \mathbb{N}$, $\mid f_n \mid^p \leq g$ a.e for $x \in X$, then the function $f \in L_p$.*

Proof

Since $f_n \xrightarrow{uqn} f$, then $\mid f_n \mid \xrightarrow{uqn} \mid f \mid$ (from theorem 2.1 (v)). The functions $\mid f_n \mid, \mid f \mid$ are measurable. So, the convergence $\mid f_n \mid \xrightarrow{uqn} \mid f \mid$ is equivalent with $\mid f_n \mid \xrightarrow{a.e} \mid f \mid$ (proposition 3.1 in [4]*). Therefore, we can write that $\mid f_n \mid^p \xrightarrow{a.e} \mid f \mid^p$. From Dominated convergence theorem in $L(\mathbb{R})$, we conclude that $\mid f \mid^p \in L(\mathbb{R})$ that means $f \in L_p$.*

Now we are giving some examples of functions that are part of the L_p, $(0 < p < 1)$, where X is finite measurable space ($0 < \mu(X) < +\infty$).

Example 3.9. *Every simple function is p-integrabled.*

Let's see it more precisely:

From definition of simple function, $\varphi(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x)$,where χ_{A_i} is denoted the characteristic function, $a_i \in \mathbb{R}$ and $A_i \in \Sigma$ (with Σ is denoted the Lebesgue measurabled set collection) for every $i \in \{1, 2, ..., n\}$, such that $A_i \cap A_j = \emptyset$ for $i \neq j$ and $\cup_{i=1}^{n} A_i = X$. So, for every $0 < p < 1$, $\mid \varphi(x) \mid^p = \sum_{i=1}^{n} \mid a_i \mid^p \chi_{A_i}(x)$ and it is clear that $\mid \varphi(x) \mid^p$ is a sipmle function. Recall that every simple function is Lebesgue integrabled in a finite measurable space. This complete the proof.

We know from the Lebesgue's integration theory that for every measurable function there is a simple functions sequence that converges a.e at that function. So, every bounded measurable function in a finite measurable space X (for example X = $[a, b]$) satisfy the conditions of Dominated convergence theorem. Thus we proof the following corollary:

Corollary 3.10. *Every bounded measurable function in a finite measurable space X is p-integrabled.*

So we come to the conclusion that: The space L_p, $(0 < p < 1)$, where X is finite measurable space ($0 < \mu(X) < +\infty$), contains all Riemann's integrabled functions as well as all Lebesgue's integrabled functions that are bounded also.

Our last result is the Brezis -Lieb lemma in case of L_p spaces $(0 < p < 1)$, where X is finite measurable space ($0 < \mu(X) < +\infty$).

Let (X, Σ, μ) be a finite measurable space. Let the mapping $j : \mathbb{C} \to \mathbb{C}$ be a continuous function with $j(0) = 0$. In addition, let j satisfy the following hypothesis:

for every sufficiently small $\varepsilon > 0$, there exist two continuous, nonnegative functions φ_ε and ψ_ε such that $\mid j(a + b) - j(a) \mid \leq \varepsilon \cdot \varphi_\varepsilon(a) + \psi_\varepsilon(b)$ for all $a, b \in \mathbb{C}$.

Lemma 3.11. *(Brezis-Lieb lemma in case of L_p spaces $(0 < p < 1)$, where X is finite measurable space)*

Let $f_n = f + g_n$ be a sequence of measurable functions from X to $\mathbb{C}$ such that:

1. $g_n \xrightarrow{a.e} 0$,

2. $j \circ f \in L_p$,

3. $\int_X \mid \varphi_\varepsilon \circ g_n \mid^p d\mu \leq A < +\infty$ for some A independent of ε and n,

4. $\int_X \mid \psi_\varepsilon \circ f \mid^p d\mu < +\infty$ for all $\varepsilon > 0$.

Then $\lim_{n\to+\infty} \int_X \mid j(f + g_n) - j(g_n) - j(f) \mid^p d\mu = 0$.

Proof

Here we reproduce its proof from theorem 2 of [2] *with some additional remarks.*

Fix $\varepsilon > 0$ and take $w_{\varepsilon,n} = [\mid j \circ f_n - j \circ g_n - j \circ f \mid - \varepsilon^{\frac{1}{p}} \cdot (\varphi_\varepsilon \circ g_n)]^+$.

Since $w_{\varepsilon,n} \xrightarrow{a.e} 0$ as $n \to +\infty$, then $\lim_{n\to+\infty} \int_X w_{\varepsilon,n}^p d\mu = 0$. Therefore $\mid j \circ f_n - j \circ g_n - j \circ f \mid^p \leq (w_{\varepsilon,n} + \varepsilon^{\frac{1}{p}} \cdot (\varphi_\varepsilon \circ g_n))^p \leq 2^p \cdot max(w_{\varepsilon,n}, \varepsilon^{\frac{1}{p}} \cdot (\varphi_\varepsilon \circ g_n))^p < 2^p \cdot (w_{\varepsilon,n}^p + \varepsilon \cdot (\varphi_\varepsilon \circ g_n)^p)$ and thus $I_n = \int_X \mid j \circ f_n - j \circ g_n - j \circ f \mid^p d\mu \leq 2^p \cdot (\int_X w_{\varepsilon,n}^p d\mu + \varepsilon \cdot \int_X \mid \varphi_\varepsilon \circ g_n) \mid^p d\mu)$. Consecuently, $\lim_{n\to+\infty} I_n \leq 2^p \cdot (\varepsilon + \varepsilon \cdot A) < 2(\varepsilon + \varepsilon \cdot A) = \varepsilon'$. This complete the proof.

4. Conclusion

As we pointed out above, the space L_p, $(0 < p < 1)$, where X is finite measurable space $(0 < \mu(X) < +\infty$), contains all Riemann's integrabled functions as well as all Lebesgue's integrabled functions that are bounded also. So this space is wider than Riemann's integrabled functions and is included in Lebesgue's integrabled functions. Below we list three of the results achieved.
1.The L_p space for $0 < p < 1$ and X a finite measurable space, is a quasi-Banach lattice.
2.Unbounded order convergence of a sequence in L_p is equivalent with uqn-convergence.
3.Dominated convergence theorem and Brezis-Lieb lemma are also generalized in the case of L_p space for $0 < p < 1$ and X a finite measurable space.

References

[1] Aliprantis C.D. and Burkinshaw O., (2003), Locally solid Riesz spaces with applications to economics, Mathematical Surveys and Monographs, American Mathematical Society, Volume 105 second edition.
[2] *Brézis* H. and Lieb E., (1983), A relation between pointwise convergence of functions and convergence of functionals, Proceedings of the American Mathematical Society, 88(3), 486-490.
[3] Deng Y., OBrien M. and Troitsky V. G., (2017), Unbounded norm convergence in Banach lattices, Positivity, 21 963-974.
[4] Gao N., Troitsky V. G. and Xanthos F., (2017), Uo-convergence and its applications to Cesro means in Banach lattices, Israel Journal of Math., 220 649-689.
[5] Gao N. and Xanthos F. (2014), Unbounded order convergence and application to martingales without probability, J.Math.Anal.Appl, 412(2) 931-947.
[6] *Kandić* M., Marabeh M.A.A. and Troitsky V. G., (2017), Unbounded norm topology in Banach lattices, J.Math.Anal.Appl, 451(1) 259-279.
[7] Kusraev A.G. and Tasoev B.B., (2017), Maximal quasi-normed extension of quasi-normed lattices, Vladikavkaz. Mat. Zh., vol.19 number 3 41-50.
[8] Papadimitrakis M., (2004), Notes on measure theory, University of Crete, Heraklion.
[9] *Pavlović* M. and *Drašković* .Z., (2004), Introduction to function spaces on the disc, *Matematički* institut SANU, Beograd.
[10] Troitsky V. G., (2004), Measures on non-compactness of operators on Banach lattices, Positivity 8(2) 165-178.

University of Tirana, Faculty of Natural Sciences, Bul. Zog 1, Tirana, Albania, ORCID: 0000-0001-9929-3584
Email(s): enkeleida.zajmikotonaj@fshn.edu.al

New Results About Lebesgue-Type Parameters in Greedy Approximation Theory

Pablo M. Berná Larrosa

2020 Mathematics Subject Classifications: 41A65, 46B15.

Keywords: Greedy Algorithm, unconditional bases, Greedy bases

1. Introduction

In 1999, S. V. Konyagin and V. N. Temlyakov introduce one of the most important algorithms in the field of Non-Linear Approximation Theory, the so called Thresholding Greedy Algorithm. To define the algorithm, first of all, we introduce the notation that we need.

Let $(\mathbb{X}, \|\cdot\|)$ be a Banach space over the field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $\mathcal{B} = (\mathbf{x}_n)_{n=1}^{\infty}$ be a semi-normalized Markushevich basis (a *basis* for short), that is,

- $\overline{\text{span}(\mathbf{x}_n : n \in \mathbb{N})} = \mathbb{X}$,
- there exists a unique sequence (called biorthogonal functionals) $\mathcal{B}^* = (\mathbf{x}_n^*)_{n=1}^{\infty}$ such that $\mathbf{x}_n^*(\mathbf{x}_m) = \delta_{n,m}$,
- $\overline{\text{span}(\mathbf{x}_n^* : n \in \mathbb{N})}^{w^*} = \mathbb{X}^*$,
- there exist $c_1, c_2 > 0$ such that
$$0 < c_1 \leq \inf_n \{\|\mathbf{x}_n\|, \|\mathbf{x}_n^*\|\} \leq \sup_n \{\|\mathbf{x}_n\|, \|\mathbf{x}_n^*\|\} \leq c_2 < \infty.$$

Under these conditions, for each $f \in \mathbb{X}$, we have the following series expansion:
$$f \sim \sum_{n=1}^{\infty} \mathbf{x}_n^*(f)\mathbf{x}_n,$$
where $\lim_{n\to+\infty} \mathbf{x}_n^*(f) = 0$. On the other hand, given a finite set $A \subset \mathbb{N}$,
$$1_A[\mathcal{B}, \mathbb{X}] = 1_A := \sum_{n\in A} \mathbf{x}_n, \quad 1_{\varepsilon A}[\mathcal{B}, \mathbb{X}] = 1_{\varepsilon A} := \sum_{n\in A} \varepsilon_n \mathbf{x}_n,$$
$$1_A^* = 1_A[\mathcal{B}^*, \mathbb{X}^*] := \sum_{n\in A} \mathbf{x}_n^*, \quad 1_{\varepsilon A}^* = 1_{\varepsilon A}[\mathcal{B}^*, \mathbb{X}^*] := \sum_{n\in A} \varepsilon_n \mathbf{x}_n^*,$$
where $\varepsilon = (\varepsilon_n)_{n\in A}$ is any sequence such that $|\varepsilon_n| = 1$ for all $n \in A$ (we will use the notation $|\varepsilon| = 1$ for short). Related to this sums, we can define the *fundamental function of* $\mathbb{X}$ as the function
$$\varphi(m) := \sup_{|A|\leq m, |\varepsilon|=1} \|1_{\varepsilon A}\|,$$
and its dual function, that is, the fundamental function of the dual Banach space is
$$\varphi^*(m) := \sup_{|A|\leq m, |\varepsilon|=1} \|1_{\varepsilon A}^*\|$$
Finally, if $a_n \lesssim b_n$ we mean that there exists $C > 0$ such that $a_n \leq C b_n$ for every $n \in \mathbb{N}$.

Then, given a basis $\mathcal{B} = (\mathbf{x}_n)_{n=1}^{\infty}$ in a Banach space $\mathbb{X}$, the Thresholding Greedy Algorithm (TGA) is defined as follows: take the *natural greedy ordering*, that is, a mapping $\rho : \mathbb{N} \to \mathbb{N}$ such that $\text{supp}(f) = \{n \in \mathbb{N} : \mathbf{x}_n^*(f) \neq 0\} \subseteq \rho(\mathbb{N})$ and such that if $j < k$, then either $|\mathbf{x}_{\rho(j)}^*(f)| > |\mathbf{x}_{\rho(k)}^*(f)|$ or $|\mathbf{x}_{\rho(j)}^*(f)| = |\mathbf{x}_{\rho(k)}^*(f)|$ and $\rho(j) < \rho(k)$. Then, the **greedy sum** of order m is given by this order:
$$\mathcal{G}_m(f) := \sum_{n=1}^{m} \mathbf{x}_{\rho(n)}^*(f)\mathbf{x}_{\rho(n)}.$$

Also, this sum is uniquely determined since if $f \in \mathbb{X}$, for each $m \in \mathbb{N}$ there exists a unique set $A_m(f) \subset \mathbb{N}$ such that $A_m(f) = \{\rho(1), \dots, \rho(m)\}$ and $A_m(f)$ is called the **greedy set** of order m of f. The sequence $(\mathcal{G}_m)_{m=1}^{\infty}$ is the (TGA) associated to $\mathcal{B}$ in $\mathbb{X}$.

One natural question when we have an algorithm of approximation is when this algorithm converges. To answer that question, in [9], the authors introduced the notion of quasi-greedy bases: a basis $\mathcal{B}$ is **quasi-greedy** if there is a positive constant C such that

$$\|f - \mathcal{G}_m(f)\| \leq C\|f\|, \ \forall m \in \mathbb{N}, \forall f \in \mathbb{X}.$$

Related to that property, P. Wojtaszczyk proved in [14] that a basis is quasi-greedy if and only if the algorithm converges, that is,

$$\lim_{m \to +\infty} \|f - \mathcal{G}_m(f)\| = 0, \ \forall f \in \mathbb{X}.$$

Hence, quasi-greediness is the minimal condition that guarantees the convergence of the algorithm. Here, our goal is to study how good is $\|f - \mathcal{G}_m(f)\|$ vs $\sigma_m(f)$, where

$$\sigma_m(f) := \inf \left\{ \left\| f - \sum_{n \in A} c_n \mathbf{x}_n \right\| \ : \ |A| = m, \ c_n \in \mathbb{F} \right\}.$$

Concretely, we want to find the smallest $\mathbf{L}_m = \mathbf{L}_m(\mathbb{X}, \mathcal{B})$ such that

$$\sigma_m(f) \leq \|f - \mathcal{G}_m(f)\| \leq \mathbf{L}_m \sigma_m(f), \ \forall f \in \mathbb{X}.$$

The parameter $\mathbf{L}_m$ is called the **Lebesgue-type parameter for the (TGA)**. As a little remark, $\mathbf{L}_m$ is called Lebesgue-type paramter for the (TGA) due to Lebesgue proved in 1909 the following: for any 2π-periodic and continous function f we have

$$\frac{\|f - S_m(f)\|_\infty}{E_m(f)_\infty} \approx \ln(m),$$

where $S_m(f)$ is the m-th partial sum of the Fourier series of f and $E_m(f)_\infty$ is the error of the best approximation of f by trigonometric polynomials of order m in the uniform norm $\|\cdot\|_\infty$. So, we do the same for the (TGA) in a general Banach space and with a general basis.

For $\mathbf{L}_m$ we have two possibilities:

(1) If $\sup_m \mathbf{L}_m = C_g < \infty$, we say that $\mathcal{B}$ is greedy ([9]).
(2) If $\sup_m \mathbf{L}_m = \infty$, we want to study the growth of the parameter $\mathbf{L}_m$.

Respect to the first possibility, due to the complexity to show that a particular basis is greedy by definition, the authors of [9] introduced a nice characterization of these type of bases based on two nice properties.

Definition 1.1. *We say that a basis $\mathcal{B}$ is* ***unconditional*** *if there is $K > 0$ such that*

$$\|f - P_A(f)\| \leq K\|f\|, \ \forall |A| < \infty, \forall f \in \mathbb{X},$$

where $P_A(f) = \sum_{n \in A} \mathbf{x}_n^*(f)\mathbf{x}_n$.

Definition 1.2. *We say that a basis $\mathcal{B}$ is* ***democratic*** *if there is $D > 0$ such that*

$$\|1_A\| \leq D\|1_B\|, \ \forall |A| \leq |B| < \infty.$$

Theorem 1.3 ([9])**.** *A basis $\mathcal{B}$ in a Banach space is greedy if and only if the basis is democratic and unconditional.*

There are several examples of greedy bases. Some of them are the following:

(1) Let $\mathbb{X} = \mathbb{H}$ be a Hilbert space and $\mathcal{B} = (\mathbf{x}_n)_n$ an orthonormal basis. Then $\mathcal{B}$ is greedy with constant $C_g = 1$,
$$\|f - \mathcal{G}_m(f)\| = \sigma_m(f), \ \forall f \in \mathbb{H}, \ \forall m \in \mathbb{N}.$$
(2) The canonical basis in $\mathbb{X} = \ell_p$, $1 \leq p < \infty$, is greedy with constant $C_g = 1$.
(3) The Haar basis is a greedy basis in $L_p[0,1]$ with $1 < p < \infty$ and the constant is 1 if $p = 2$ ([10]).
(4) The trigonometric system is not greedy in $L_p(\mathbb{T})$ with $p \neq 2$ ([11]).

Now, we focus our attention on the second case, that is, when $\sup_m \mathbf{L}_m = \infty$.

The first example in the greedy-literature where this parameter was studied appears in [11], where V. N. Temlyakov proved that the trigonometric system $\mathcal{T} = \{e^{2\pi i x}\}_{k \in \mathbb{Z}}$ in $L^p(\mathbb{T})$, with $1 < p < \infty$, is not greedy with $p \neq 2$ showing the following behaviour: for each $m \in \mathbb{N}$,

$$\mathbf{L}_m \approx m^{\left|\frac{1}{2} - \frac{1}{p}\right|}.$$

The proof of that result involves so many techniques of Functional Analysis, Complex Analysis and Harmonic Analysis. Then, since this year where the author proved this result, many authors have studied different ideas to give a general result for the behaviour of $\mathbf{L}_m$ that works for any basis. In that paper, we will discuss these results and we give a new one where we closed the theory.

1.1. **Preliminaries.** Since 2011, the study of $\mathbf{L}_m$ is focused on the existence of two parameters a_m and b_m such that

$$\max\{a_m, b_m\} \approx \mathbf{L}_m. \tag{1}$$

Of course, these parameters have to be enough natural and efficient to compute. As the characterization of greedy bases is based on the study if the basis is democratic and unconditional, the most natural parameters to analyze $\mathbf{L}_m$ are:

- Democracy parameter:
$$\mu_m := \sup_{|A|\leq|B|\leq m} \frac{\|1_A\|}{\|1_B\|}.$$
- Unconditional parameter:
$$k_m := \sup_{|A|\leq m} \|P_A\|,\ k_m^c := \sup_{|A|\leq m} \|I - P_A\|,$$
where
$$\|P_A\| = \sup_{f\neq 0} \frac{\|P_A(f)\|}{\|f\|},\ \|I - P_A(f)\| = \sup_{f\neq 0} \frac{\|f - P_A(f)\|}{\|f\|}.$$

Of course, $|k_m - k_m^c| \leq 1$ for every $m \in \mathbb{N}$. The first result in the direction (1) was given by V. N. Temlyakov et al. ([13]) establishing the following relation.

Theorem 1.4. *Let $\mathcal{B}$ be a quasi-greedy basis in a Banach space $\mathbb{X}$. Then, for each $m \in \mathbb{N}$,*
$$\mathbf{L}_m \lesssim k_m\mu_m.$$

It is true that this relation also could be completed using the following lower bounds:

$$\max\{k_m, \mu_m\} \lesssim \mathbf{L}_m \lesssim k_m\mu_m, \tag{2}$$

where the lower bounds were proved in [8]. The unique problem with the estimation (2) is in the upper bound since we know the existence of examples of conditional bases that are not democratic. For instance, taking a basis in a Banach space $\mathbb{X}$ such that $\mathcal{B}$ is democratic, quasi-greedy and conditonal, we can construct the space $\mathbb{Y} = \mathbb{X} \oplus c_0$, where in this space, the natural basis is conditional, quasi-greedy but not democratic.

Based on that problem, in 2011, V. N. Temlyakov ([12]) talked in a conference about the necessity to find a natural sequence of some greedy-type parameters which combined linearly with the sequence of $(k_m)_{m=1}^{\infty}$ determines the growth of $\mathbf{L}_m$. Trying to solve this problem, in 2013, G. Garrigós et al. ([8]) proved the following result for quasi-greedy bases.

Theorem 1.5. *Let $\mathcal{B}$ be a quasi-greedy basis in a Banach space $\mathbb{X}$. Then, for each $m \in \mathbb{N}$,*
$$\mathbf{L}_m \approx \max\{\mu_m, k_m\}.$$

This theorem shows a good behavior of $\mathbf{L}_m$ using only k_m and μ_m that are easier parameters than $\mathbf{L}_m$. In fact, that result is optimal: consider $(\mathbf{e}_n)_{n=1}^{\infty}$ the canonical basis of ℓ_1 and define the following vectors:

$$\mathbf{x}_1 = \mathbf{e}_1,\ \ \mathbf{x}_n = \frac{1}{2}(\mathbf{e}_{2n} + \mathbf{e}_{2n+1}),\ n = 1, 2, \dots$$

The collection $\mathcal{B} = (\mathbf{x}_n)_{n=1}^{\infty}$ is the Lindestrauss basis and it is a basis over $\overline{\mathrm{span}(\mathbf{x}_n : n \in \mathbb{N})}$ in ℓ_1. $\mathcal{B}$ is a conditional, democratic and quasi-greedy basis ([7]) and
$$\mathbf{L}_m \approx k_m \approx \ln(m + 1).$$

Another similar result following the same idea is an improvement respect to the previous one proved by S. J. Dilworth et al. ([6]).

Theorem 1.6. *Let $\mathcal{B}$ a quasi-greedy basis in a Banach space $\mathbb{X}$. Then, for each $m \in \mathbb{N}$,*
$$\mathbf{L}_m \approx \max\{\mu_m^d, k_m\},$$
where μ_m^d is the parameter μ_m with the extra condition $A \cap B = \emptyset$.

It is trivial to see the following relation:

$$\mu_m^d \leq \mu_m \leq (\mu_m^d)^2. \tag{3}$$

Thanks to (3), we can affirm that

$$\mathbf{L}_m \approx \max\{\mu_m^d, k_m\},$$

is a better result than

$$\mathbf{L}_m \approx \max\{\mu_m, k_m\}.$$

But the truth is that when $\mathcal{B}$ is a Schauder basis, then $\mu_m \approx \mu_m^d$ ([3]), although in the context of Markushevich bases, we have the following example.

Theorem 1.7 ([3])**.** *There exists a semi-normalized Markushevich (not Schauder) basis in a Banach space such that*

$$\limsup_{m\to+\infty} \frac{\mu_m}{(\mu_m^d)^{2-\varepsilon}} = \infty,\ \forall \varepsilon > 0.$$

Then, we can see that these results follow the philosophy to find some good behavior of $\mathbf{L}_m$, but the unique problem is that the most of these results are given in the context of quasi-greedy bases, that is, bases where the algorithm converges. So, the question raised by V. N. Temlyakov is not answered. In the next section we will see how to give an answer based on the results proved in [1].

2. Main results

In that section, we show one result in the spirit of (1). For that, we need to introduce one new parameter that is the key of the question posed by V. N. Temlyakov in [12].

For each $m \in \mathbb{N}$, we define λ_m as the smallest constant C such that

$$\min_{n\in A} |\mathbf{x}_n^*(f)|\varphi(|A|) \leq C\|f\|,\ \forall f \in \mathbb{X},$$

where $|A| = m$ and A is a greedy set of f. This parameter could be apparently "strange", but have some connections with some greedy-like properties. For instance, one of the most important is the bidemocracy: we say that a basis is bidemocratic if

$$\varphi(m)\varphi^*(m) \lesssim m,\ \forall m \in \mathbb{N}.$$

This property is the main ingredient in the theory to talk about the duality of greedy-like bases. Concretely, in 2003, S. J. Dilworth et al. ([5]) proved that if $\mathcal{B}$ is quasi-greedy (or unconditional), then the basis is bidemocratic if and only if $\mathcal{B}$ and $\mathcal{B}^*$ are quasi-greedy and democratic. The relation between bidemocracy and λ_m is in [1, Section 6].

Another important property that appears in some papers as in [4] is the Property (C) or recentely renamed as restricted quasi-greediness in [2].

Definition 2.1. *Let $\mathcal{B}$ be a basis in a Banach space $\mathbb{X}$. We say that $\mathcal{B}$ is* ***restricted quasi-greedy*** *if there is $C > 0$ such that*

$$\|\mathcal{R}_m(f)\| \leq C\|f\|,\ \forall m \in \mathbb{N}, \forall f \in \mathbb{X},$$

where

$$\mathcal{R}_m(f) = \min_{n\in A_m(f)} |\mathbf{x}_n^*(f)|\|1_{\varepsilon A_m(f)}\|,$$

with $\varepsilon \equiv \{sign(\mathbf{x}_n^*(f))\}_{n\in\mathbb{N}}$ and $A_m(f)$ the greedy set of order m.

With respect to this property, in [1] we can found the following characterization.

Theorem 2.2. *Let $\mathcal{B}$ be a basis in a Banach space $\mathbb{X}$. Then, $\sup_m \lambda_m < \infty$ if and only if the basis is restricted quasi-greedy and democratic.*

Remark 1. *The condition to be restricted quasi-greedy and democratic is so closed to the condition to be almost-greedy (*[1, 5]*).*

Finally, the most important result proved in [1] is the following one closing the question posed by V. N. Temlyakov.

Theorem 2.3. *Let $\mathcal{B}$ be a basis in a Banach space $\mathbb{X}$. Then, for each $m \in \mathbb{N}$,*

$$\mathbf{L}_m \approx \max\{k_m, \lambda_m\}.$$

Remark 2. *The last result is proved also in* [1] *in the general context of quasi-Banach spaces.*

References

[1] F. Albiac, J. L. Ansorena, P. M. Berná, *New parameters and Lebsgue-type estimates in greedy approximation.* Accepted in Forum of Mathematics, Sigma.

[2] F. Albiac, J. L. Ansorena, P. M. Berná, P. Wojtaszczyk, *Greedy approximations for biorthoonal systems ini quasi-Banach spaces*, Dissertatioines Math. (Rozprawy Mat.) **560** (2021), 1-88.

[3] P. M. Berná, O. Blasco, G. Garrigós, E. Hernández and T. Oikhberg, *Lebesgue inequalities for Chebyshev thresholding greedy algorithms*, Rev. Matem. Compl. 33 (2020), 695-722.

[4] P. M. Berná, S. J. Dilworth, D. Kutzarova, T. Oikhberg, B. Wallis, *The weighted property (A) and the greedy algorithm*, J. Approx. Theory **248** (2019), 105300.

[5] S. J. Dilworth, N. J. Kalton, D. Kutzarova, V. N. Temlyakov,*The thresholding greedy algorithm, greedy bases, and duality*, Constr.Approx.**19** (2003), no.4, 575-597.

[6] S. J. Dilworth, D. Kutzarova, T. Oikhberg, *Lebesgue constants for the weak greedy algorithm*, Rev. Matem. Compl. 28 (2) (2015), 393–409.

[7] S. J. Dilworth, D. Mitra, *A conditional quasi-greedy bases of ℓ^1*, Studia Math. 144 (2001), 95-100.

[8] G. Garrigós, E. Hernández and T. Oikhberg, *Lebesgue-type inequalities for quasi-greedy bases*, Constr. Approx. 38 (3) (2013), 447–470.

[9] S.V.Konyagin, V.N.Temlyakov,*A remark on greedy approximation in Banach spaces*, East J. Approx. **5** (1999), 365-379.

[10] V.N.Temlyakov,*The best m-Term Approximation and Greedy Algorithms*, Adv. Comput. Math. **8** (1998), no. 3, 249-265.

[11] V.N.Temlyakov,*Greedy Algorithm and m-Term Trigonometric Approximation*, Constr. Approx. **14** (1998), no. 4, 569-587.

[12] V.N.Temlyakov, Question posed during the *Concentration week on greedy algorithms in Banach spaces and compressed sensing* held on 18-22 July, 2011, at Texas A& M University.

[13] V.N. Temlyakov, M. Yang, P. Ye,*Lebesgue-type inequalities for greedy approximation with respect to quasi-greedy bases*, East J. Approx **17** (2011), 127-138.

[14] P.Wojtaszczyk,*Greedy algorithm for general biorthogonal systems*, J.Approx.Theory **107** (2000), no.2, 293-314.

Departamento de Mtodos Cuantitativos, CUNEF Universidad, 28040 Madrid (Spain)

Email(s): pablo.berna@cunef.edu

Jungck's Fixed Point Theorem for Weakly Compatible Mappings Satisfying Orthogonal F-Contraction

Kübra Özkan

Abstract. In this paper, we proved two generalizations of Jungck's fixed point theorem for a pair of weakly compatible F-contraction mappings in orthogonal metric spaces. With the new results presented in the article, we generalize and enrich methods presented in the literature that we cite.

2020 Mathematics Subject Classifications: 47H10, 54H25

Keywords: Fixed Point Theory, F-Contraction, O-completeness.

1. Introduction

In 1976, Jungck (*cf.* [1]) proved a common fixed point theorem for commuting maps, generalizing the Banach contraction principle. This theorem has many applications. In addition this, Jungck defined a pair of self mappings to be weakly compatible if they commute at their coincidence points. And he established common fixed points for these type mappings (*cf.* [2, 3]).

The notions of orthogonal sets (brieftly, O-sets) and orthogonal metric spaces were introduced by Gordji et al. (*cf.* [4]) in 2017. And they proved Banach's fixed point theorem in that study. In addition this, they discussed the existence of solution of differential equation using their results. For find more details about O-sets and orthogonal metric spaces, the readers are referred to (*cf.* [5, 6, 7, 8, 9, 10, 11, 12, 13, 14]).

In this paper, we proved two generalizations of Jungck's fixed point theorem for a pair of weakly compatible F-contraction mappings in orthogonal metric spaces. With the new results presented in the article, we generalize and enrich methods presented in the literature that we cite.

2. Preliminaries

In 2012, Wardowski introduced the concept of F-contraction as follows:

Definition 2.1 (*cf.* [15, p.2]). *Let $F : \mathbb{R}^+ \to \mathbb{R}$ be a mapping satisfying:*

(F1) F is strictly increasing,

(F2) For each sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ of positive numbers, we have $\lim_{n\to\infty} \alpha_n = 0$ if and only if

$$\lim_{n\to\infty} F(\alpha_n) = -\infty,$$

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

$\mathcal{F}$ is the family of all functions F that satisfy the conditions $(F1)$, $(F2)$ and $(F3)$.

Definition 2.2 (*cf.* [15, p.2]). *Let (X,d) be metric space. A self-mapping T on X is called an F-contraction if there exist $F \in \mathcal{F}$ and $\tau \in \mathbb{R}^+$ such that*

$$d(Tx,Ty) > 0 \Rightarrow \tau + F(d(Tx,Ty)) \le F(d(x,y))$$

for each $x, y \in X$.

Definition 2.3 (*cf.* [4, p.570])**.** *Let X is a nonempty set and $\perp \subseteq X \times X$ be a binary relation. If the relation $\perp$ satisfies the following condition:*

$$\exists x_0 \in X: \ (\forall y, \ y \perp x_0) \ \text{or} \ (\forall y, \ x_0 \perp y)$$

then X is called an orthogonal set (briefly, O-set) and x_0 is called an orthogonal element. We represent this O-set by $(X, \perp)$.

Example 2.4 (*cf.* [4, p.570])**.** *Let $X = \mathbb{Z}$. We define $m \perp n$ if there exists $k \in \mathbb{Z}$ such that $m = kn$. It is obvious that $0 \perp n$ for all $n \in \mathbb{Z}$. So, $(X, \perp)$ is an O-set.*

Example 2.5 (*cf.* [4, p.571])**.** *Let $X = [0, \infty)$. We define $x \perp y$ if $xy \in \{x, y\}$. For orthogonal elements $x_0 = 0$ or $x_0 = 1$, $(X, \perp)$ is an O-set.*

As seen in the above example, x_0 is not necessarily unique.

Definition 2.6 (*cf.* [4, p.572])**.** *Let $(X, \perp)$ be O-set. A sequence $\{x_i\}_{i\in\mathbb{N}}$ is called an orthogonal sequence (briefly, O-sequence) if*

$$(\forall i, x_i \perp x_{i+1}) \ \text{or} \ (\forall i, x_{i+1} \perp x_i).$$

Definition 2.7 (*cf.* [4, p.572])**.** *The triplet $(X, \perp, d)$ is called orthogonal metric space if $(X, \perp)$ is an O-set and (X, d) is a metric space.*

Definition 2.8 (*cf.* [4, p.572])**.** *Let $(X, \perp, d)$ be an orthogonal metric space. The mapping $f : X \to X$ is called orthogonally continuous (or $\perp$-continuous) in $x \in X$ if we get $f(x_i) \to f(a)$ for each O-sequence $\{x_i\}_{i\in\mathbb{N}}$ in X with $x_i \to x$ as $i \to \infty$. And, the mapping f is called $\perp$-continuous on X if f is $\perp$-continuous for all $x \in X$.*

Definition 2.9 (*cf.* [4, p.572])**.** *Let $(X, \perp, d)$ be an orthogonal metric space. X is called orthogonally complete (briefly, O-complete) if every Cauchy O-sequence is convergent.*

Remark 1 (*cf.* [4, p.572])**.** *Every complete metric space is O-complete and the converse is not true.*

Definition 2.10 (*cf.* [4, p.573])**.** *Let $(X, \perp, d)$ be an O-set. A mapping $f : X \to X$ is said to be $\perp$-preserving if $f(x) \perp f(y)$ whenever $x \perp y$. Also, $f : X \to X$ is said to be weakly $\perp$-preserving if $f(x) \perp f(y)$ or $f(y) \perp f(x)$ whenever $x \perp y$.*

Definition 2.11 (*cf.* [14, p.328])**.** *Let $(X, \perp, d)$ be an O-complete metric space, $f : X \to X$ be a mapping. f is called sequential orthogonal mapping, if*

$$fx_n \perp fx_{n+1} \Rightarrow fx_{n+1} \perp fx_{n+2}$$

and

$$fx_{n+1} \perp fx_n \Rightarrow fx_{n+2} \perp fx_{n+1}$$

for all $\{x_n\} \subseteq X$, $n \in \mathbb{N}$.

3. Main results

Definition 3.1. *Let f and g be self mappings of a O-set X. Then f and g are said to be weakly compatible if $fx = gx \Rightarrow fgx = gfx$.*

Definition 3.2. *Let f and g be self mappings of a O-set X. If $w = fx = gx$ for some $x \in X$, then x is called a coincidence point of f and g. And, w is called a point of coincidence of f and g.*

Proposition 3.3. *Let f and g be weakly compatible self mappings of O-set X. If f and g have a unique point of coincidence $w = fx = gx$, then w is the unique common fixed point of f and g.*

Proof. Since $w = fx = gx$ and f, g are weakly compatible, we have

$$fw = fgx = gfx = gw.$$

That is, $fw = gw$ is a point of coincidence of f and g. But w is the only point of coincidence of f and g, so $w = fw = gw$. Moreover, if $z = fz = gz$, then z is a point of coincidence of f and g. From uniqueness, we get $z = w$. Thus, w is the unique common fixed point of f and g. □

Theorem 3.4. *Let $(X, \perp, d)$ be an O-complete metric space. The O-sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ and mappings $f, g : X \to X$ satisfy the following conditions:*
(i) *Let f and g be a weakly compatible mappings and $f(X) \subseteq g(X)$.*
(ii) *Let f be a $\perp-$preserving and sequential orthogonal mapping.*
(iii) *There exists $\exists\lambda \in (0, 1)$ such that*

$$d(fx, fy) > 0 \Rightarrow \tau + F(d(fx, fy)) \leq F(\lambda d(gx, gy))$$

for all $x, y \in X$ with $gx\perp gy$, where $\tau > 0$ and $F \in \mathcal{F}$.
(iv) *If there exist $u, v \in X$ such that $gx_n \to v = gu$, then $gx_n\perp v = gu$ for all $n \in \mathbb{N}^+$.*
(v) *$ga\perp gb$ such that $fa = ga$ and $fb = gb$ for all $a, b \in X$.*
Then f and g have a unique common fixed point.

Proof. Since $(X, \perp)$ is an O-set, there exists $x_0 \in X$ such that

$$(x_0\perp y \quad \text{for all} \quad y \in X) \quad \text{or} \quad (y\perp x_0 \quad \text{for all} \quad y \in X).$$

For orthogonal element $x_0, fx_0, gx_0 \in X$ are well defined. Since $f(X) \subseteq g(X)$, there exists $x_1 \in X$ such that $fx_0 = gx_1$. Then we choose $x_2 \in X$ such that $fx_1 = gx_2$. By proceeding, we get that there exists $x_{n+1} \in X$ such that $fx_n = gx_{n+1}$ for all $n \in \mathbb{N}$. Since x_0 is orthogonal element and f is a $\perp-$preserving and sequential orthogonal mapping, we get

$$x_0\perp x_1 \Rightarrow fx_0\perp fx_1 \Rightarrow gx_1 = fx_0\perp fx_1 = gx_2 \Rightarrow gx_1\perp gx_2$$

or

$$x_1\perp x_0 \Rightarrow fx_1\perp fx_0 \Rightarrow gx_2 = fx_1\perp fx_0 = gx_1 \Rightarrow gx_2\perp gx_1.$$

Continuing this process, we can say that $gx_{n-1}\perp gx_n$ or $gx_n\perp gx_{n-1}$ for $n \in \mathbb{N}$. Then $\{gx_n\}_{n\in\mathbb{N}}$ is an O-sequence. From (iii), we get

$$F(d(gx_{n+1}, gx_n)) = F(d(fx_n, fx_{n-1})) \leq F(\lambda d(gx_n, gx_{n-1})) - \tau \leq F(\lambda d(gx_n, gx_{n-1}))$$

for all $n \in \mathbb{N}$ and $\tau > 0$. Since F is strictly increasing and $\lambda \in (0, 1)$, we get

$$d(gx_{n+1}, gx_n) \leq \lambda d(gx_n, gx_{n-1}) < d(gx_n, gx_{n-1})$$

for all $n \in \mathbb{N}$. We take $\alpha_n = d(gx_{n+1}, gx_n)$ for all $n \in \mathbb{N}$. Then we have $\alpha_n < \alpha_{n-1}$ for all $n \in \mathbb{N}$. Consequently, we get

$$\tau + F(\alpha_n) \leq F(\alpha_{n-1})$$

for all $n \in \mathbb{N}$. Therefore, we get

$$F(\alpha_n) \leq F(\alpha_{n-1}) - \tau \leq \cdots \leq F(\alpha_0) - n\tau \tag{1}$$

for all $n \in \mathbb{N}$. As $n \to \infty$, we get $\lim\limits_{n\to\infty} F(\alpha_n) = -\infty$. By property (F2), we get

$$\lim_{n\to\infty} \alpha_n = 0. \tag{2}$$

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

$$\lim_{n\to\infty} \alpha_n^k F(\alpha_n) = 0. \tag{3}$$

By (1), we get

$$\alpha_n^k F(\alpha_n) - \alpha_n^k F(\alpha_0) \leq -\alpha_n^k n\tau \leq 0 \tag{4}$$

for all $n \in \mathbb{N}$. If we take limit as $n \to \infty$ in (4), using (2) and (3) we get $\lim\limits_{n\to\infty} n\alpha_n^k = 0$. Then there exists $n_1 \in \mathbb{N}$ such that $n\alpha_n^k \leq 1$ for all $n \geq n_1$. So we get

$$\alpha_n \leq \frac{1}{n^{\frac{1}{k}}} \tag{5}$$

for all $n \geq n_1$. Now, we claim that $\{gx_n\}_{n\in\mathbb{N}}$ is a Cauchy O-sequence. We consider $n, m \in \mathbb{N}$ such that $m > n \geq n_1$, then using (5) we get

$$\begin{aligned} d(gx_m, gx_n) &\leq d(gx_m, gx_{m-1}) + d(gx_{m-1}, gx_{m-2}) + \cdots + d(gx_{n+1}, gx_n) \\ &= \alpha_{m-1} + \alpha_{m-2} + \cdots + \alpha_n \\ &= \sum_{i=n}^{m-1} \alpha_i \leq \sum_{i=n}^{\infty} \alpha_i \leq \sum_{i=n}^{\infty} \frac{1}{i^{\frac{1}{k}}}. \end{aligned}$$

Since $\frac{1}{i^{\frac{1}{k}}} < \infty$, it follows that $\{gx_n\}_{n\in\mathbb{N}}$ is a Cauchy O-sequence in X. Since X is O-complete, there exists $v \in X$ such that $gx_n \to v$. Then there exists $\exists u \in X$ such that $gu = v$. From (iii) and (iv), we get

$$F(d(gx_n, fu)) < \tau + F(d(gx_n, fu)) = \tau + F(d(fx_{n-1}, fu)) \leq F(\lambda d(gx_{n-1}, gu))$$

for $n \in \mathbb{N}^+$ and $\tau > 0$. F is a strictly increasing mapping, we have

$$d(gx_n, fu) \leq \lambda d(gx_{n-1}, gu).$$

As $n \to \infty$, we get $\lim_{n\to\infty} d(gx_n, fu) = 0$. On the otherhand, we can say that $\lim_{n\to\infty} d(gx_n, gu) = 0$. From uniqueness of limit, we get $gu = fu$. To show the uniqueness of common fixed point of f and g, let us show that they have a unique point of coincidence. For this, we assume that there exists another point r in X such that $fr = gr$. Then from (iii) and (v), we get

$$F(d(gr, gu)) = F(d(fr, fu)) < \tau + F(d(fr, fu)) \leq F(\lambda d(gr, gu))$$

which is a contradiction. Then we get $d(gr, gu) = 0$. That is, $gr = gu$. Then f and g have a unique point of coincidence in X. From Proposition 3.3, f and g have a unique common fixed point. □

Theorem 3.5. *Let $(X, \perp, d)$ be an O-complete metric space. The O-sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ and mappings $f, g : X \to X$ satisfy the following conditions:*

(i) *Let f and g be a weakly compatible mappings and $f(X) \subseteq g(X)$.*

(ii) *Let f be a $\perp-$preserving and sequential orthogonal mapping.*

(iii) *There exists $\exists\lambda \in \left(0, \frac{1}{2}\right)$ such that*

$$d(fx, fy) > 0 \Rightarrow \tau + F(d(fx, fy)) \leq F(\lambda(d(fx, gx) + d(fy, gy)))$$

for all $x, y \in X$ with $gx \perp gy$, where $\tau > 0$ and $F \in \mathcal{F}$.

(iv) *If there exist $u, v \in X$ such that $gx_n \to v = gu$, then $gx_n \perp v = gu$ for all $n \in \mathbb{N}^+$.*

(v) *$ga \perp gb$ such that $fa = ga$ and $fb = gb$ for all $a, b \in X$.*

Then f and g have a unique common fixed point.

Proof. Since $(X, \perp)$ is an O-set, there exists $x_0 \in X$ such that

$$(x_0 \perp y \quad \text{for all} \quad y \in X) \quad \text{or} \quad (y \perp x_0 \quad \text{for all} \quad y \in X).$$

For orthogonal element $x_0, fx_0, gx_0 \in X$ are well defined. Since $f(X) \subseteq g(X)$, there exists $x_1 \in X$ such that $fx_0 = gx_1$. Then we choose $x_2 \in X$ such that $fx_1 = gx_2$. By proceeding, we get that there exists $x_{n+1} \in X$ such that $fx_n = gx_{n+1}$ for all $n \in \mathbb{N}$. Since x_0 is orthogonal element and f is a $\perp-$preserving and sequential orthogonal mapping, we get

$$x_0 \perp x_1 \Rightarrow fx_0 \perp fx_1 \Rightarrow gx_1 = fx_0 \perp fx_1 = gx_2 \Rightarrow gx_1 \perp gx_2$$

or

$$x_1 \perp x_0 \Rightarrow fx_1 \perp fx_0 \Rightarrow gx_2 = fx_1 \perp fx_0 = gx_1 \Rightarrow gx_2 \perp gx_1.$$

Continuing this process, we can say that $gx_{n-1} \perp gx_n$ or $gx_n \perp gx_{n-1}$ for $n \in \mathbb{N}$. Then $\{gx_n\}_{n\in\mathbb{N}}$ is an O-sequence. From (iii), we get

$$\begin{aligned} F(d(gx_{n+1}, gx_n)) &= F(d(fx_n, fx_{n-1})) \\ &\leq F(\lambda(d(fx_n, gx_n) + d(fx_{n-1}, gx_{n-1}))) - \tau \\ &< F(\lambda(d(gx_{n+1}, gx_n) + d(gx_n, gx_{n-1}))) \end{aligned}$$

for all $n \in \mathbb{N}$ and $\tau > 0$. Since F is strictly increasing, we get

$$d(gx_{n+1}, gx_n) < \lambda(d(gx_{n+1}, gx_n) + d(gx_n, gx_{n-1}))$$

for all $n \in \mathbb{N}$. Then we get

$$d(gx_{n+1}, gx_n) < \frac{\lambda}{1-\lambda} d(gx_n, gx_{n-1})$$

$\lambda \in \left(0, \frac{1}{2}\right)$ for all $n \in \mathbb{N}$. We take $\alpha_n = d(gx_{n+1}, gx_n)$ for all $n \in \mathbb{N}$. Since $\lambda \in \left(0, \frac{1}{2}\right)$, we get $\alpha_n < \alpha_{n-1}$ for all $n \in \mathbb{N}$. Consequently, we get

$$\tau + F(\alpha_n) \leq F(\alpha_{n-1})$$

for all $n \in \mathbb{N}$. Therefore, we get

$$F(\alpha_n) \leq F(\alpha_{n-1}) - \tau \leq \cdots \leq F(\alpha_0) - n\tau \tag{6}$$

for all $n \in \mathbb{N}$. As $n \to \infty$, we get $\lim_{n\to\infty} F(\alpha_n) = -\infty$. By property (F2), we get

$$\lim_{n\to\infty} \alpha_n = 0. \tag{7}$$

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

$$\lim_{n\to\infty} \alpha_n^k F(\alpha_n) = 0. \tag{8}$$

By (6), we get

$$\alpha_n^k F(\alpha_n) - \alpha_n^k F(\alpha_0) \leq -\alpha_n^k n\tau \leq 0 \tag{9}$$

for all $n \in \mathbb{N}$. If we take limit as $n \to \infty$ in (9), using (7) and (8) we get $\lim_{n\to\infty} n\alpha_n^k = 0$. Then there exists $n_1 \in \mathbb{N}$ such that $n\alpha_n^k \leq 1$ for all $n \geq n_1$. So we get

$$\alpha_n \leq \frac{1}{n^{\frac{1}{k}}} \tag{10}$$

for all $n \geq n_1$. Now, we claim that $\{gx_n\}_{n\in\mathbb{N}}$ is a Cauchy O-sequence. We consider $n, m \in \mathbb{N}$ such that $m > n \geq n_1$, then using (10) we get

$$\begin{aligned} d(gx_m, gx_n) &\leq d(gx_m, gx_{m-1}) + d(gx_{m-1}, gx_{m-2}) + \cdots + d(gx_{n+1}, gx_n) \\ &= \alpha_{m-1} + \alpha_{m-2} + \cdots + \alpha_n \\ &= \sum_{i=n}^{m-1} \alpha_i \leq \sum_{i=n}^{\infty} \alpha_i \leq \sum_{i=n}^{\infty} \frac{1}{i^{\frac{1}{k}}}. \end{aligned}$$

Since $\frac{1}{i^{\frac{1}{k}}} < \infty$, it follows that $\{gx_n\}_{n\in\mathbb{N}}$ is a Cauchy O-sequence in X. Since X is O-complete, there exists $v \in X$ such that $gx_n \to v$. Then there exists $\exists u \in X$ such that $gu = v$. From (iii) and (iv), we get

$$F(d(gx_n, fu)) < \tau + F(d(gx_n, fu)) = \tau + F(d(fx_{n-1}, fu)) \leq F(\lambda(d(fx_{n-1}, gx_{n-1}) + d(fu, gu)))$$

for $n \in \mathbb{N}^+$ and $\tau > 0$. F is a strictly increasing mapping, we have

$$\begin{aligned} d(gx_n, fu) &\leq \lambda(d(fx_{n-1}, gx_{n-1}) + d(fu, gu)) \\ &\leq \lambda(d(fx_{n-1}, gx_{n-1}) + d(fu, gx_n) + d(gx_n, gu)). \end{aligned}$$

Since $gx_n \to gu = v$, we get

$$\begin{aligned} d(gx_n, fu) &\leq \lambda(d(fx_{n-1}, gx_{n-1}) + d(fu, gx_n)) \\ &\Rightarrow d(gx_n, fu) \leq \frac{\lambda}{1-\lambda} d(gx_n, gx_{n-1}) \end{aligned}$$

for $\lambda \in \left(0, \frac{1}{2}\right)$ and all $n \in \mathbb{N}^+$. As $n \to \infty$, we get $\lim_{n\to\infty} d(gx_n, fu) = 0$. On the otherhand, we can say that $\lim_{n\to\infty} d(gx_n, gu) = 0$. From uniqueness of limit, we get $gu = fu$. To show the uniqueness of common fixed point of f and g, let us show that they have a unique point of coincidence. For this, we assume that there exists another point r in X such that $fr = gr$. Then from (iii) and (v), we get

$$F(d(gr, gu)) = F(d(fr, fu)) \leq F(\lambda(d(fr, gr) + d(fu, gu))) - \tau < F(\lambda(d(fr, gr) + d(fu, gu)))$$

for $\tau > 0$. Since F is a strictly increasing mapping, we get

$$d(gr, gu) < \lambda(d(fr, gr) + d(fu, gu)) = 0.$$

This is a contradiction. Then we get $d(gr, gu) = 0 \Rightarrow gr = gu$. Then f and g have a unique point of coincidence in X. From Proposition 3.3, f and g have a unique common fixed point. □

References

[1] Jungck G., (1976), Commuting maps and fixed points, Amer. Math. Monthly, 83, 261263.
[2] Jungck G., (1996), Common fixed points for noncontinuous nonself maps on non-metric spaces, Far East J. Math. Sci. (FJMS), 4, 199215.
[3] Jungck G. and Rhoades B.E., (1998), Fixed point for set valued functions without continuity, Indian J. Pure Appl. Math., 29 (3), 227238.
[4] Gordji M.E., Rameani M., De La Sen M. and Cho Y.J., (2017), On orthogonal sets and Banach fixed point theorem, Fixed Point Theory, 18, 569-578.
[5] Acar, Ö., Erdoğan, E., (2022), Some fixed point results for almost contraction on orthogonal metric space, Creat. Math. Inform., 31 (2), 147-153.
[6] Acar, Ö., Özkapu, A.S., (2022), Multivalued rational type F-contraction on orthogonal metric space, Math. Found. Comput., accepted.
[7] Ahmadi Z., Lashkaripour R. and Baghani H., (2018), A fixed point problem with constraint inequalities via a contraction in incomplete metric spaces, Filomat, 32 (9), 3365-3379.
[8] Baghani H., Gordji M.E. and Ramezani M., (2016), Orthogonal sets: The axiom of choice and proof of a fixed point theorem, J. Fixed Point Theory Appl., 18 (3), 465-477.
[9] Baghani H. and Ramezani M., (2019), Coincidence and fixed points for multivalued mappings in incomplete metric spaces with application, Filomat, 33, 13-26.
[10] Baghani H., Agarwal R.P. and Karapnar E., (2019), On coincidence point and fixed point theorems for a general class of multivalued mappings in incomplete metric spaces with an application, Filomat, 33 (14), 4493-4508.
[11] Özkan, K., (2022), Common fixed point theorems in incomplete metric spaces with application, Journal of Nonlinear and Convex Analysis, 23 (5), 891-903.
[12] Sawangsup K., Sintunavarat W. and Cho Y.J., (2020), Fixed point theorems for orthogonal F-contraction mappings on O-complete metric spaces, J. Fixed Point Theory Appl., 22 (1), Article number:10.
[13] Sawangsup,K., Sintunavarat, W., (2020), Fixed Point Results for Orthogonal Z-Contraction Mappings in O-Complete Metric Spaces, Int. J. Appl. Phys. Math., 10 (1).
[14] Sürmelioğlu M. and Güngör Bilgili N., (2019), Ortogonal Koni Metrik Uzaylarda Bazı Ortak Sabit Nokta Teoremleri, SETSCI Conference Proceedings, 4 (1), 326-332.
[15] Wardowski D., (2012), Fixed point theory of a new type of contractive mappings in complete metric spaces, Fixed Point Theory Appl., 2012, Article ID 94 .

Department of Mathematics, Manisa Celal Bayar University, Manisa, Türkiye

Email(s): kubra.ozkan@hotmail.com, kubra.ozkan@cbu.edu.tr

Kashuri Fundo Transform for Solving Chemical Reaction Models

Haldun Alpaslan Peker*[1], Fatma Aybike Çuha[2] and Bilge Peker[3]

ABSTRACT. No event experienced or encountered in the universe we live in does not happen by chance. In the background of these events there are excellent rules or rule systems. By trying to understand and interpret these rules, sciences create different fields of study for themselves and develop themselves. While doing this, they use some models in order to understand and interpret these rules correctly. One of these models is differential equations. The use of these equations in modeling most of the problems in the literature makes it important to reach their solutions. Computations to reach solutions of differential equations can be more complex than computations for algebraic equations. For this reason, different methods have been put forward to reach the solutions of these equations. One of these methods is integral transforms. In this study, we discussed Kashuri Fundo transform, which is one of the integral transforms. In order to show the effectiveness of Kashuri Fundo transform in solving ordinary differential equations, we applied Kashuri Fundo transform to chemical reaction models modeled with ordinary differential equations. With these applications, we have shown that solutions to chemical reaction models can be obtained without the need for complex calculations with the Kashuri Fundo transform. From this point of view, it can be concluded that the Kashuri Fundo transform is a very useful method in terms of its applicability and ease of operation in reaching solutions of ordinary differential equations.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 34A12, 44A15, 65R10

KEYWORDS: integral transform, Kashuri Fundo transform, inverse Kashuri Fundo transform, ordinary differential equation, chemical reaction model

1. INTRODUCTION

Differential equations are used in modeling real life problems in engineering, physics, statistics, economics, chemistry and many other fields. Differential equations contain quantities that vary depending on each other and the rates of change (derivative) of these quantities with respect to each other (*cf.* [22]). Modeling the relationship between derivatives with differential equations not only makes real life problems more understandable, but also makes them easier to plan solutions through these models (*cf.* [21]).

Differential equations form a large and very important branch of modern mathematics (*cf.* [17]). This makes the issue of reaching solutions very important. Computations in solutions of differential equations can be more complex than computations in solutions of algebraic equations. Researchers sought a way to transform differential equations into algebraic equations in order to find a solution to this situation. Until now, many different methods have been put forward to reach solutions of differential equations (*cf.* [1, 9, 16]). One of these methods is integral transforms.

Integral transforms are used in the solutions of initial-value problems, boundary-value problems, differential equations and integral equations that have taken place in fields such as engineering, physics and chemistry. The fact that the application area is so diverse has caused the integral transforms to diversify within themselves (*cf.* [4]). The most well-known integral transforms are Laplace and Fourier transforms.

Integral transforms are based on the purpose of creating an equation that does not include derivatives of the dependent variable by multiplying each term in the equation with an appropriate function called the kernel and taking the integral of each term with respect to the independent variable (*cf.* [3]). In this study, we considered Kashuri Fundo transform, which is one of the integral transforms (*cf.* [12]). We used the Kashuri Fundo transform to arrive at the solution of some chemical reaction models.

In the literature, it is possible to come across many studies on Kashuri Fundo transform. In 2013, Kashuri Fundo transform and homotopy perturbation method were combined and used to solve nonlinear partial differential equations (*cf.* [13]). Also in the same year, the double Kashuri Fundo transform was introduced (*cf.* [15]). In later times, the Kashuri Fundo transform were combined with the expansion coefficients of binomial series, the projected differential transform method and the Adomian decomposition method, and used in various fractional differential equations and nonlinear differential equations (*cf.* [11, 18, 23, 24, 25, 26, 27, 28]). In addition, the solutions of some models encountered in different fields of science were investigated using the Kashuri Fundo transform (*cf.* [2, 5, 6, 7, 8, 10, 14, 19, 20]).

2. Kashuri Fundo Transform

The Kashuri Fundo transform is derived from the classical Fourier integral for solving ordinary and partial differential equations with constant and variable coefficients (*cf.* [12]). The basis of Kashuri Fundo transform, as in other integral transforms, is to facilitate the solution process by converting differential equations into algebraic equations. After the necessary operations are done, the solution of the equation is reached by using the inverse of the Kashuri Fundo transform.

2.1. Definitions related to Kashuri Fundo transform.

Definition 2.1 (*cf.* [12]). *We consider functions in the set F defined as*

$$F = \{f(t) | \exists M, k_1, k_2 > 0, \quad \text{such that} \quad |f(t)| \leqslant Me^{\frac{|t|}{k_i^2}}, \text{if} \quad t \in (-1)^i \times [0, \infty)\}$$

For a function belonging to the set F, the constant M must be finite number. k_1, k_2 may be finite or infinite.

Definition 2.2 (*cf.* [12]). *Kashuri Fundo transform defined on the set F and denoted by $\mathscr{K}(.)$ is defined as,*

$$\mathscr{K}[f(t)](v) = A(v) = \frac{1}{v}\int_0^\infty e^{\frac{-t}{v^2}} f(t)dt, \quad t \geqslant 0, \quad -k_1 < v < k_2 \tag{1}$$

The Kashuri Fundo transform expressed by equation (1) *can also be expressed as,*

$$\mathscr{K}[f(t)](v) = A(v) = v\int_0^\infty e^{-t} f(v^2 t)\, dt$$

Inverse Kashuri Fundo transform is denoted by $\mathscr{K}^{-1}[A(v)] = f(t)$, $t \geq 0$.

Definition 2.3 (*cf.* [12]). *If there exist positive constants T and M such that, $|f(t)| \leq Me^{\frac{-t}{k^2}}$, for all $t \geq T$, the function $f(t)$ is said to have exponential order $\frac{1}{k^2}$.*

2.2. Some theorems related to Kashuri Fundo transform.

Theorem 2.4 (*cf.* [12] **Sufficient conditions for existence of Kashuri Fundo transform**). *If $f(t)$ have exponential order $\frac{1}{k^2}$ and is piecewise continuous on $[0, \infty)$, then $\mathscr{K}[f(t)](v)$ exists for $|v| < k$.*

Theorem 2.5 (*cf.* [12] **Linearity of Kashuri Fundo transform**). *Let the Kashuri Fundo transforms of $f(t)$ and $g(t)$ be exist and c be a scalar. Then,*

$$\mathscr{K}[(f+g)(t)] = \mathscr{K}[f(t)] + \mathscr{K}[g(t)]$$

$$\mathscr{K}[(cf)(t)] = c\mathscr{K}[f(t)]$$

Theorem 2.6 (*cf.* [12] **Kashuri Fundo transform of the derivatives**). *Let's assume that the Kashuri Fundo transform of $f(t)$, denoted by $A(v)$, exists. Then,*

$$\mathscr{K}\left[\frac{df(t)}{dt}\right] = \frac{A(v)}{v^2} - \frac{f(0)}{v} \tag{2}$$

$$\mathscr{K}\left[\frac{d^2f(t)}{dt^2}\right] = \frac{A(v)}{v^4} - \frac{f(0)}{v^3} - \frac{f'(0)}{v} \tag{3}$$

$$\mathscr{K}\left[\frac{d^{(n)}f(t)}{dt^{(n)}}\right] = \frac{A(v)}{v^{2n}} - \sum_{k=0}^{n-1}\frac{f^{(k)}(0)}{v^{2(n-k)-1}} \tag{4}$$

2.3. **Kashuri Fundo transforms of some special functions.** Some functions that are frequently used in models that appear in different fields of science and the Kashuri Fundo transforms of these functions are given below (*cf.* [12, 27, 28]). The functions we encounter in the models in the literature are not limited to those expressed here. We have included the most common functions here.

$f(t)$	$A(v)$
t	v^3
$t^n, \quad n \in \mathbb{Z}$	$n!v^{2n+1}$
e^{ct}	$\frac{v}{1-cv^2}$
$sin(ct)$	$\frac{cv^3}{1+c^2v^4}$
$cos(ct)$	$\frac{v}{1+c^2v^4}$
$sinh(ct)$	$\frac{cv^3}{1-c^2v^4}$
$cosh(ct)$	$\frac{v}{1-c^2v^4}$
$t^\alpha, \quad \alpha \in \mathbb{R}^+$	$\Gamma(\alpha+1)v^{2\alpha+1}$
$\sum_{k=0}^{n} c_k t^k$	$\sum_{k=0}^{n} k!c_k v^{2k+1}$

Table 1. Kashuri Fundo Transform of Some Special Functions

3. Application of Kashuri Fundo transform to chemical reaction models

In this section, we look for solutions by applying Kashuri Fundo transform to chemical reaction models which have an important place in chemistry and are modeled with ordinary differential equations.

Application 3.1 (*cf.* [4, p. 258])**.** The zero-order chemical reaction model satisfies the initial value problem

$$\frac{dc(t)}{dt} = -k_0, \quad t > 0 \tag{5}$$

$$c(0) = c_0 \tag{6}$$

where k_0 is a positive constant and $c(t)$ is the concentration of a reacting substance at time t. Solve the given zero-order chemical reaction model using Kashuri Fundo transform.

Having applied bilaterally the Kashuri Fundo transform to the equation (5), we acquire

$$\mathscr{K}\left[\frac{dc(t)}{dt}\right] = \mathscr{K}[-k_0] \tag{7}$$

Let $\mathscr{K}[c(t)] = A(v)$. Rearranging the equation (7) using the equation (2) and initial condition, we get

$$A(v) = c_0 v - k_0 v^3 \tag{8}$$

Applying the inverse Kashuri Fundo transform to this equation using table 1, we will find the solution of the given initial value problem as

$$c(t) = c_0 - k_0 t \tag{9}$$

Application 3.2 (*cf.* [4, p. 258])**.** The first-order chemical reaction model satisfies the initial value problem

$$\frac{dc(t)}{dt} = -k_1 c(t), \quad k_1 > 0 \tag{10}$$

$$c(0) = c_0 \tag{11}$$

Solve the given first-order chemical reaction model using Kashuri Fundo transform.

Having applied bilaterally the Kashuri Fundo transform to the equation (10), we acquire

$$\mathscr{K}\left[\frac{dc(t)}{dt}\right] = \mathscr{K}\left[-k_1 c(t)\right] \tag{12}$$

Let $\mathscr{K}[c(t)] = A(v)$. Rearranging the equation (12) using the equation (2) and initial condition, we get

$$A(v) = \frac{c_0 v}{1 + k_1 v^2} \tag{13}$$

Applying the inverse Kashuri Fundo transform to this equation using table 1, we will find the solution of the given initial value problem as

$$c(t) = c_0 e^{-k_1 t} \tag{14}$$

Application 3.3 (*cf.* [4, p. 258])**.** Solve the systems of differential equations governing the consecutive chemical reactions of the first order given by

$$\frac{dc_1(t)}{dt} = -k_1 c_1(t), \quad \frac{dc_2(t)}{dt} = k_1 c_1(t) - k_2 c_2(t), \quad \frac{dc_3(t)}{dt} = k_2 c_2(t), \quad t > 0 \tag{15}$$

with the initial conditions

$$c_1(0) = c_0, \quad c_2(0) = c_3(0) = 0 \tag{16}$$

by Kashuri Fundo transform. In the systems (15), $c_1(t)$ is the concentration of a substance A_1 at time t, which breaks down to form a new substance A_2 with concentration $c_2(t)$, and $c_3(t)$ is the concentration of a new element originated from A_2.

We found the solution of the differential equation of function $c_1(t)$ as $c_0 e^{-k_1 t}$ in the previous application. Substituting this solution in the differential equation of the $c_2(t)$ function, we get

$$\frac{dc_2(t)}{dt} = k_1 c_0 e^{-k_1 t} - k_2 c_2(t) \tag{17}$$

Now having applied bilaterally the Kashuri Fundo transform to the equation (17), we acquire

$$\mathscr{K}\left[\frac{dc_2(t)}{dt}\right] = k_1 c_0 \mathscr{K}\left[e^{-k_1 t}\right] - k_2 \mathscr{K}\left[c_2(t)\right] \tag{18}$$

Let $\mathscr{K}[c_2(t)] = A_2(v)$. Rearranging the equation (18) using the equation (2) and table 1, we obtain

$$A_2(v) = k_1 c_0 \frac{v^3}{(1 + k_1 v^2)(1 + k_2 v^2)} \tag{19}$$

If we rearrange this equation so that we can apply the inverse Kashuri Fundo transform, we get

$$A_2(v) = \frac{k_1 c_0}{k_2 - k_1}\left(\frac{v}{1 + k_1 v^2} - \frac{v}{1 + k_2 v^2}\right) \tag{20}$$

Having applied bilaterally the inverse Kashuri Fundo transform to the equation (20) using table 1, we acquire

(21) $$c_2(t) = \frac{k_1 c_0}{k_2 - k_1}\left(e^{-k_1 t} - e^{-k_2 t}\right)$$

We now know the solutions of the differential equation of function $c_2(t)$. If we substitute this solution in the differential equation of the $c_3(t)$ function, we get

(22) $$\frac{dc_3(t)}{dt} = k_2\left(\frac{k_1 c_0}{k_2 - k_1}\left(e^{-k_1 t} - e^{-k_2 t}\right)\right)$$

And now, having applied bilaterally the Kashuri Fundo transform to the equation (22), we acquire

(23) $$\mathscr{K}\left[\frac{dc_3(t)}{dt}\right] = \frac{k_2 k_1 c_0}{k_2 - k_1}\left(\mathscr{K}\left[e^{-k_1 t}\right] - \mathscr{K}\left[e^{-k_2 t}\right]\right)$$

Let $\mathscr{K}[c_3(t)] = A_3(v)$. Rearranging the equation (23) using the equation (2) and table 1, we obtain

(24) $$A_3(v) = \frac{k_2 k_1 c_0}{k_2 - k_1}\left(\frac{v^3}{1 + k_1 v^2} - \frac{v^3}{1 + k_2 v^2}\right)$$

Rearranging the equation (24) so that we can apply the inverse Kashuri Fundo transform, we get

(25) $$A_3(v) = c_0 v - c_0 \frac{v}{1 + k_1 v^2} - \frac{k_1 c_0}{k_2 - k_1}\left(\frac{v}{1 + k_1 v^2} - \frac{v}{1 + k_2 v^2}\right)$$

Having applied bilaterally the inverse Kashuri Fundo transform to this equation using table 1, we acquire

(26) $$c_3(t) = c_0 - c_0 e^{-k_1 t} - \frac{k_1 c_0}{k_2 - k_1}\left(e^{-k_1 t} - e^{-k_2 t}\right)$$

This result can also be expressed as follows, taking into account the solutions so far.

(27) $$c_3(t) = c_0 - c_1(t) - c_2(t)$$

4. Conclusion

Computations made in the solutions of differential equations compared to algebraic equations can be more complex. For this reason, in the solution of differential equations, the use of integral transformations that transform these equations into algebraic equations provides convenience in reaching the solution. In this study, we applied Kashuri Fundo transform to some chemical reaction models in order to show that this transform facilitates the solution of ordinary differential equations. These applications have shown that results can be obtained without the need for complex calculations by using the Kashuri Fundo transform. From this point of view, it can be concluded that the Kashuri Fundo transform is a very useful method in terms of its applicability and ease of operation in the solution of ordinary differential equations.

References

[1] Adomian G., (1988), A review of the decomposition method in applied mathematics, J. Math. Anal. Appl., 135 501-544.

[2] Cuha F. A. and Peker H. A., (2022), Solution of Abel's integral equation by Kashuri Fundo transform, Therm. Sci., 26 (4A) 3003-3010.

[3] Çengel Y. A. and Palm W. J., (2013), *Differential equations for engineering and basic sciences*, İzmir Güven Bookstore, İzmir.

[4] Debnath L. and Bhatta D., (2007), *Integral transforms and their applications*, Taylor and Francis Group, New York.

[5] Dhange D. N., (2020), A new integral transform and its applications in electric circuits and mechanics, JETIR, 7 (11) 80-86.

[6] Dhange N., (2020), A new integral transform for solution of convolution type Volterra integral equation of first kind, IJMTT, 66 (10) 52-57.
[7] Fundo A., Kashuri A. and Liko R., (2016), New integral transform in Caputo type fractional difference operator, Univers. J. Appl. Sci., 4 (1) 7-10.
[8] Güngör N., (2021), Solving convolution type linear Volterra integral equations with Kashuri Fundo transform, J. Abstr. Comput. Math., 6 (2), 1-7.
[9] He H. J., (1999), Variational iteration method - a kind of non-linear analytical technique: some examples, Int. J. Non Linear Mech., 34 (4) 699-708.
[10] Helmi N., Kiftiah M. and Prihandono B., (2016), Solving linear partial differential equations using Artion-Fundo Transformation Method, Buletin Ilmiah Matematika Statistika dan Terapannya, 5 (3) 195-204.
[11] Johansyah M. D., Supriatna A. K., Rusyaman E. and Saputra J., (2022), Solving the economic growth acceleration model with memory effects: an application of combined theorem of Adomian decomposition methods and Kashuri-Fundo transformation methods, Symmetry, 14 (2) 192.
[12] Kashuri A. and Fundo A., (2013), A new integral transform, ATAM, 8 (1) 27-43.
[13] Kashuri A., Fundo A. and Kreku M., (2013), Mixture of a new integral transform and homotopy perturbation method for solving nonlinear partial differential equations, Adv. Pure Math., 3 317-323.
[14] Kashuri A., Fundo A. and Liko R., (2013), New integral transform for solving some fractional differential equations, Int. J. Pure Appl. Math., 103 (4) 675-682.
[15] Kashuri A., Fundo A. and Liko R., (2013), On double new integral transform and double Laplace transform, Eur. Sci. J., 9 (33) 82-90.
[16] Liao S. J., (1992), The proposed homotopy analysis technique for the solution of non-linear problems, Shanghai Jiao Tong University (Ph.D. thesis).
[17] Pala Y., (2013), *Modern applied differential equations*, Nobel Akademik Publishing, Ankara.
[18] Peker H. A. and Cuha F. A., (2022), Application of Kashuri Fundo transform and homotopy perturbation methods to fractional heat transfer and porous media equations, Therm. Sci., 26 (4A) 2877-2884.
[19] Peker H. A., Cuha F. A. and Peker B., (2022), Solving steady heat transfer problems via Kashuri Fundo transform, Therm. Sci., 26 (4A) 3011-3017.
[20] Peker H. A. and Çuha F. A., (in press), Application of Kashuri Fundo transform to decay problem, SDU J. Nat. Appl. Sci.
[21] Ross S. L., (1984), *Differential equations*, John Wiley & Sons, New York.
[22] Sezer M. and Daşcıoğlu A., (2010), *Differential equations I*, Dora Publications, Bursa.
[23] Shah K. B., (2018), Homotopy perturbation new integral transform method for numeric study of space- and time fractional (n+1)-dimensional heat and wave-like equations, Waves Wavelets and Fractals, 4 19-36.
[24] Shah K. and Singh T., (2015), A solution of the Burger's equation arising in the longitudinal dispersion phenomenon in fluid flow through porous media by mixture of new integral transform and homotopy perturbation method, J. Geosci. Environ. Prot., 3 24-30.
[25] Shah K. and Singh T., (2015), The mixture of new integral transform and homotopy perturbation method for solving discontinued problems arising in nanotechnology, Open J. Appl. Sci., 5 688-695.
[26] Shah K., Singh T. and Kılıçman B., (2018), Combination of integral and projected differential transform methods for time-fractional gas dynamics equations, Ain Shams Eng. J., 9 1683-1688.
[27] Subartini B., Sumiati I., Sukono, Riaman and Sulaiman I. M., (2021), Combined Adomian decomposition method with integral transform, Math. Stat., 9 (6) 976-983.
[28] Sumiati I., Sukono and Bon A. T., (2020), Adomian decomposition method and the new integral transform, Proceedings of the 2nd African International Conference on Industrial Engineering and Operations Management, Harare (Zimbabwe) 1882-1887.

Selçuk University, Faculty of Science, Department of Mathematics, Campus, 42130, Konya, Turkiye [1*] (corresponding author)

Selçuk University, Graduate School of Natural and Applied Sciences, Campus, 42130, Konya, Turkiye [2]

Necmettin Erbakan University, Ahmet Keleşoğlu Faculty of Education, Department of Mathematics and Education, Meram, 42090, Konya, Turkiye [3]

Email(s): pekera@gmail.com [1*] (corresponding author), fatmaaybikecuha@gmail.com [2] pekerb@gmail.com [3]

Interpolative Contractions in Convex $b-$Metric Spaces

Isa Yildirim[1], Y. Serdar Sahin[*1]

Abstract. In this paper, some results is given for classes of interpolative Kannan type contraction and interpolative Reich–Rus–Ciri´c type contraction in complete convex $b-$metric spaces.

2020 Mathematics Subject Classifications: 47H10, 54H25

Keywords: Fixed point, interpolative contraction mapping, $b-$metric spaces.

1. Introduction and Preliminaries

b-metric spaces is a generalization of metric spaces. There are many works about such spaces (see [5], [9], [10]). In 2018, Karapinar [7] introduced the well-known fixed point theorem of Kannan under the aspect of interpolation. When the interpolative type contractions in the literature are examined, it is seen that the fixed points of such mappings are obtained by Picard iteration (see [2], [3], [6], [7], [8]). Our aim in this paper is to show that the fixed points of these contractions classes can be found by means of different iterations. It is already known that convex structure is needed for different iteration schemes. So, in this work, we firstly give the definition of the convex b-metric space and we introduce Mann's iteration scheme in such space. We also prove some fixed point theorems for two types of interpolative contraction mapping using Mann's iteration scheme in convex b-metric spaces. For this, the fixed points of these contractions classes are obtained via Mann iteration ([12], [13]) by taking convex metric space instead of metric space.

Now we will give some definitions that we will use in our results, respectively.

Definition 1.1. [11] *Let (C,ρ) be a metric space. A convex structure is a mapping $w : C^2 \times I \to C$ satisfying, for all $\zeta, \eta, r \in C$ and all $\mu \in I$,*

$$\rho(r, w(\zeta, \eta; \mu)) \leq \mu\rho(r, \zeta) + (1 - \mu)\rho(r, \eta)$$

where $I = [0, 1]$. Then (C,ρ,w) is called a convex metric space.

Definition 1.2. [1] *Let (C,ρ) be a b-metric space with constant $s \geq 1$ and let the mapping $w : C^2 \times I \to C$ be a convex structure on (C, ρ). Then (C,ρ_b,w) is said to be a convex b-metric space.*

Throughout the article, we'll call it cbms for the convex b-metric space and ccbms for the complete convex b-metric space.

Definition 1.3. *Let (C,ρ_b,w) be a cbms and $F : C \to C$ be a mapping. The Mann's iteration scheme is defined as follows:*

$$\zeta_{n+1} = w(\zeta_n, F\zeta_n; \mu_n), \qquad n \geq 1,$$

where $\zeta_n \in C$ and $\mu_n \in I$.

Definition 1.4. [7] *Let (C, ρ) be a metric space and $F : C \to C$ be a mapping. The mapping F is called an interpolative Kannan type contraction, if there are constants $\lambda \in [0, 1)$,$\theta \in (0, 1)$ such that*

$$\rho(F\zeta, F\eta) \leq \lambda[\rho(\zeta, F\zeta)]^\theta[\rho(\eta, F\eta)]^{1-\theta},$$

for all $\zeta, \eta \in C \setminus Fix(F)$ where $Fix(F) = \{\zeta \in C : F\zeta = \zeta\}$.

Definition 1.5. [8] *Let (C, ρ) be a metric space and $F : C \to C$ be a mapping. The mapping F is called an interpolative Reich–Rus–Ciri´c type contraction, if there are constants $\lambda \in [0, 1)$ and δ,$\theta \in (0, 1)$ such that*

$$\rho(F\zeta, F\eta) \leq \lambda[\rho(\zeta, \eta)]^\delta[\rho(\zeta, F\zeta)]^\theta[\rho(\eta, F\eta)]^{1-\theta-\delta},$$

for all $\zeta, \eta \in C \setminus Fix(F)$.

2. Main Results

Theorem 2.1. *Let* (C, ρ_b, w) *be a ccbms and* $F : C \to C$ *be an interpolative Kannan type contraction, that is, there exist* $\lambda \in [0, 1)$ *and* $\theta \in (0, 1)$ *such that*

$$\rho_b(F\zeta, F\eta) \leq \lambda[\rho(\zeta, F\zeta)]^{\theta}[\rho(\eta, F\eta)]^{1-\theta}, \tag{1}$$

for all $\zeta, \eta \in C \setminus Fix(F)$ *where* $Fix(F) = \{\zeta \in C : F\zeta = \zeta\}$. *Let the sequence* $\{\zeta_n\}$ *be defined as follows*

$$\zeta_n = w(\zeta_{n-1}, F\zeta_{n-1}; \mu_{n-1}), \tag{2}$$

where $0 \leq \mu_{n-1} \leq b < 1$ *for all* $n \in \mathbb{N}$ *and* $b \in \mathbb{R}^+$. *If* $b + \lambda < \dfrac{1}{s}$ *and* $b < \dfrac{1}{s^2}$ *for all* $n \in \mathbb{N}$, *then* F *has a fixed point which is obtained via the Mann iteration (12).*

Proof. From structure of convex metric space, (1) and (2) we obtain

$$\rho_b(\zeta_n, \zeta_{n+1}) = \rho_b(\zeta_n, w(\zeta_n, F\zeta_n; \mu_n)) \leq (1 - \mu_n)\rho_b(\zeta_n, F\zeta_n), \tag{3}$$

and

$$\begin{aligned} \rho_b(\zeta_n, F\zeta_n) &\leq s\rho_b(\zeta_n, F\zeta_{n-1}) + s\rho_b(F\zeta_{n-1}, F\zeta_n) \\ &= s\rho_b(w(\zeta_{n-1}, F\zeta_{n-1}; \mu_{n-1}), F\zeta_{n-1}) + s\rho_b(F\zeta_{n-1}, F\zeta_n) \\ &\leq s\mu_{n-1}\rho_b(\zeta_{n-1}, F\zeta_{n-1}) + \\ &\quad s\lambda[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\theta}.[\rho_b(\zeta_n, F\zeta_n)]^{1-\theta} \end{aligned} \tag{4}$$

Suppose that $\rho_b(\zeta_{n-1}, F\zeta_{n-1}) < \rho_b(\zeta_n, F\zeta_n)$ for some $n \geq 1$. From (4), we have

$$\rho_b(\zeta_n, F\zeta_n) \leq s\mu_{n-1}\rho_b(\zeta_n, F\zeta_n) + s\lambda[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\theta}.[\rho_b(\zeta_n, F\zeta_n)]^{1-\theta} \tag{5}$$

and

$$(1 - s\mu_{n-1})\rho_b(\zeta_n, F\zeta_n)^{\theta} \leq s\lambda[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\theta}.$$

From above inequality, we obtain

$$\begin{aligned} \rho_b(\zeta_n, F\zeta_n)^{\theta} &\leq \frac{s\lambda}{1 - s\mu_{n-1}}[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\theta} \\ &\leq \frac{s\lambda}{1 - sb}[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\theta}. \end{aligned} \tag{6}$$

Since $b + \lambda < \dfrac{1}{s}$, we obtain that $\rho_b(\zeta_{n-1}, F\zeta_{n-1}) \geq \rho_b(\zeta_n, F\zeta_n)$. But, this implies that a contradiction. Then, we have $\rho_b(\zeta_n, F\zeta_n) \leq \rho_b(\zeta_{n-1}, F\zeta_{n-1})$ for all $n \geq 1$ which implies that $\{\rho_b(\zeta_{n-1}, F\zeta_{n-1})\}$ is a non-increaing sequence. Hence, there exists $v \geq 0$ such that

$$\lim_{n\to\infty} \rho_b(\zeta_{n-1}, \zeta_{n-1}) = v. \tag{7}$$

Now, we will show that $v = 0$. Suppose that $v > 0$. Since $\rho_b(\zeta_n, F\zeta_n) \leq \rho_b(\zeta_{n-1}, F\zeta_{n-1})$ for all $n \geq 1$, using (5), we get

$$\begin{aligned} \rho_b(\zeta_n, F\zeta_n) &\leq \frac{s\lambda}{1 - s\mu_{n-1}}\rho_b(\zeta_{n-1}, F\zeta_{n-1}) \\ &\quad \frac{s\lambda}{1 - sb}\rho_b(\zeta_{n-1}, F\zeta_{n-1}). \end{aligned} \tag{8}$$

Taking limit at the inequality (8), we obtain that

$$v \leq \frac{s\lambda}{1 - sb}v \leq v.$$

This implies that a contradiction. So, we have that $v = 0$. From (3), we also have

$$\rho_b(\zeta_n, \zeta_{n+1}) \leq (1 - \mu_n)\rho_b(\zeta_n, F\zeta_n) \leq \rho_b(\zeta_n, F\zeta_n)$$

which implies that

$$\lim_{n\to\infty} \rho_b(\zeta_n, \zeta_{n+1}) = 0. \tag{9}$$

Now, we need to show that $\{\zeta_n\}$ is a Cauchy sequence. For this, the similar proof method at the proof in [1] will be used. We suppose that $\{\zeta_n\}$ is not a Cauchy sequence. Thus, there exist $\overline{\epsilon}$ and the subsequences $\left\{\zeta_{\theta_{(k)}}\right\}$ and $\left\{\zeta_{\vartheta_{(k)}}\right\}$ of $\{\zeta_n\}$, such that $\theta(k) > \vartheta(k) > k$,

$$\rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)}}) \geq \overline{\epsilon}$$

and

$$\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)}}) < \overline{\epsilon}.$$

Therefore, we have

$$\overline{\epsilon} \leq \rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)}}) \leq s\rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) + s\rho_b(\zeta_{\vartheta_{(k)+1}}, \zeta_{\vartheta_{(k)}}). \tag{10}$$

Using (3) and $\lim_{n\to\infty} \rho_b(\zeta_n, \zeta_{n+1}) = 0$,we obtain

$$\frac{\overline{\epsilon}}{s} \leq \lim_{k\to\infty} \sup \rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}).$$

Also we have

$$\begin{aligned}
\rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) &= \rho_b(w(\zeta_{\theta_{(k)-1}}, F\zeta_{\theta_{(k)-1}}; \mu_{\theta_{(k)-1}}), \zeta_{\vartheta_{(k)+1}}) \\
&\leq \mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}}) + (1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}}) \\
&\leq \mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}}) + s(1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\theta_{(k)-1}}, F\zeta_{\vartheta_{(k)+1}}) \\
&\quad + s(1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\vartheta_{(k)+1}}, \zeta_{\vartheta_{(k)+1}}) \\
&\leq \mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}}) \\
&\quad + s(1-\mu_{\theta_{(k)-1}})\lambda[\rho_b(\zeta_{\theta_{(k)-1}}, F\zeta_{\theta_{(k)-1}}]^{\theta}[\rho_b(\zeta_{\vartheta_{(k)+1}}, F\zeta_{\vartheta_{(k)+1}}]^{1-\theta} \\
&\quad + s(1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\vartheta_{(k)+1}}, \zeta_{\vartheta_{(k)+1}}) \\
&\leq s\mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)}}) + s\mu_{\theta_{(k)-1}}\rho_b(\zeta_{\vartheta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) \\
&\quad + s(1-\mu_{\theta_{(k)-1}})\lambda[\rho_b(\zeta_{\theta_{(k)-1}}, F\zeta_{\theta_{(k)-1}}]^{\theta}[\rho_b(\zeta_{\vartheta_{(k)+1}}, F\zeta_{\vartheta_{(k)+1}}]^{1-\theta} \\
&\quad + s(1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\vartheta_{(k)+1}}, \zeta_{\vartheta_{(k)+1}})
\end{aligned}$$

Taking $\lim\sup$ as $k \to \infty$ on both sides of above inequality and using (7),(9), we obtain

$$\begin{aligned}
\frac{\overline{\epsilon}}{s} &\leq \lim_{k\to\infty} \sup \rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) \\
&\leq s\mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)}}) \\
&\leq sb\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)}}).
\end{aligned}$$

Since $b < \dfrac{1}{s^2}$, we obtain

$$\frac{\overline{\epsilon}}{s} \leq \lim_{k\to\infty} \sup \rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) \leq b\frac{\overline{\epsilon}}{s} < \frac{\overline{\epsilon}}{s}.$$

This is a contradiction. Thus,$\{\zeta_n\}$ is a Cauchy sequence in C. Since (C, ρ_b, w) be a ccbms, there exists $\zeta^* \in C$ such that $\lim_{n\to\infty} \zeta_n = \zeta^* \in C$. Next, we will show that $\zeta^*$is a fixed point of F. Note that

$$\begin{aligned}
\rho_b(\zeta^*, F\zeta^*) &\leq s[\rho_b(\zeta^*, \zeta_n) + \rho_b(\zeta_n, F\zeta^*)] \\
&\leq s\rho_b(\zeta^*, \zeta_n) + s^2[\rho_b(\zeta_n, F\zeta_n) + \rho_b(F\zeta_n, F\zeta^*)] \\
&\leq s\rho_b(\zeta^*, \zeta_n) + s^2\rho_b(\zeta_n, F\zeta_n) + s^2\lambda[\rho_b(\zeta_n, F\zeta_n)]^{\theta}[\rho_b(\zeta^*, F\zeta^*)]^{1-\theta} \\
&\leq s\rho_b(\zeta^*, \zeta_n) + s^2(\frac{s\lambda}{1-sb})^n\rho_b(\zeta_0, F\zeta_0) + \\
&\quad s^2\lambda(\frac{s\lambda}{1-sb})^n[\rho_b(\zeta_0, F\zeta_0)]^{\theta}[\rho_b(\zeta^*, F\zeta^*)]^{1-\theta}.
\end{aligned}$$

If we take limit as $n \to \infty$ on both sides of above inequality ,we deduce that $\rho_b(\zeta^*, F\zeta^*) = 0$ which implies that $F\zeta^* = \zeta^*$. □

At the next theorem, we will give the following result for interpolative Reich–Rus–Ciri´c type contraction in ccbms.

Theorem 2.2. *Let (C, ρ_b, w) be a ccbms and $F : C \to C$ be an interpolative Reich–Rus–Ciri ́c type contraction that is, there exist $\lambda \in [0, 1)$ and $\theta,\delta \in (0, 1)$ such that*

$$\rho(F\zeta, F\eta) \le \lambda[\rho(\zeta, \eta)]^{\delta}[\rho(\zeta, F\zeta)]^{\theta}[\rho(\eta, F\eta)]^{1-\theta-\delta}, \tag{11}$$

for all $\zeta, \eta \in C\setminus\ Fix(F)$ where $Fix(F) = \{\zeta \in C : F\zeta = \zeta\}$. Let $\{\zeta_n\}$ be defined as follows

$$\zeta_n = w(\zeta_{n-1}, \zeta_{n-1}; \mu_{n-1}), \tag{12}$$

where $0 \le \mu_{n-1} \le b < 1$ for all $n \in \mathbb{N}$ and $b \in \mathbb{R}^+$. If $b + \lambda < \frac{1}{s}$ and $b < \frac{1}{s^2}$ for all $n \in \mathbb{N}$, then F has a fixed point which is obtained via the Mann iteration (12).

Proof. Note that, we obtain

$$\rho_b(\zeta_n, \zeta_{n+1}) = \rho_b(\zeta_n, w(\zeta_n, F\zeta_n; \mu_n)) \le (1 - \mu_n)\rho_b(\zeta_n, F\zeta_n), \tag{13}$$

and

$$\begin{aligned}
\rho_b(\zeta_n, F\zeta_n) &\le s\rho_b(\zeta_n, F\zeta_{n-1}) + s\rho_b(\zeta_{n-1}, \zeta_n)\\
&= s\rho_b(w(\zeta_{n-1}, F\zeta_{n-1}; \mu_{n-1}), F\zeta_{n-1}) + s\rho_b(F\zeta_{n-1}, F\zeta_n)\\
&\le s\mu_{n-1}\rho_b(\zeta_{n-1}, F\zeta_{n-1})\\
&\quad +s\lambda[\rho_b(\zeta_{n-1}, \zeta_n)]^{\delta}[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\theta}[\rho_b(\zeta_n, F\zeta_n)]^{1-\delta-\theta}\\
&= s\mu_{n-1}\rho_b(\zeta_{n-1}, F\zeta_{n-1})\\
&\quad +s\lambda[\rho_b(\zeta_{n-1}, w(\zeta_{n-1}, F\zeta_{n-1}; \mu_{n-1}))]^{\delta}\\
&\quad \times[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\theta}[\rho_b(\zeta_n, F\zeta_n)]^{1-\delta-\theta}\\
&\le s\mu_{n-1}\rho_b(\zeta_{n-1}, F\zeta_{n-1})\\
&\quad +s\lambda[(1 - \mu_{n-1})\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\delta}\\
&\quad \times[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\theta}[\rho_b(\zeta_n, F\zeta_n)]^{1-\delta-\theta}.
\end{aligned} \tag{14}$$

Assume that $\rho_b(\zeta_{n-1}, F\zeta_{n-1}) < \rho_b(\zeta_n, F\zeta_n)$ for some $n \ge 1$. Using (14), we have

$$\begin{aligned}
\rho_b(\zeta_n, F\zeta_n) &\le s\mu_{n-1}\rho_b(\zeta_n, \zeta_n) +\\
&\quad s\lambda(1 - \mu_{n-1})^{\delta}[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\delta+\theta}[\rho_b(\zeta_n, F\zeta_n)]^{1-\delta-\theta}
\end{aligned} \tag{15}$$

which implies that

$$\begin{aligned}
\frac{(1 - s\mu_{n-1})\rho_b(\zeta_n, F\zeta_n)}{[\rho_b(\zeta_n, F\zeta_n)]^{1-\delta-\theta}} &\le s\lambda(1 - \mu_{n-1})^{\delta}[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\delta+\theta}\\
\rho_b(\zeta_n, F\zeta_n)^{\delta+\theta} &\le \frac{s\lambda(1 - \mu_{n-1})^{\delta}}{(1 - s\mu_{n-1})}[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\delta+\theta}.
\end{aligned} \tag{16}$$

Since $\dfrac{s\lambda(1 - \mu_{n-1})^{\delta}}{(1 - s\mu_{n-1})} \le \dfrac{s\lambda}{(1 - s\mu_{n-1})} < 1$, from above inequality (16), we obtain

$$\begin{aligned}
\rho_b(\zeta_n, F\zeta_n)^{\delta+\theta} &\le \frac{s\lambda}{(1 - s\mu_{n-1})}[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\delta+\theta}\\
&\quad \frac{s\lambda}{(1 - sb)}[\rho_b(\zeta_{n-1}, F\zeta_{n-1})]^{\delta+\theta}.
\end{aligned} \tag{17}$$

Since $b + \lambda < \frac{1}{s}$, we obtain that $\rho_b(\zeta_{n-1}, \zeta_{n-1}) \ge \rho_b(\zeta_n, F\zeta_n)$. But, this implies that a contradiction. Then, we have $\rho_b(\zeta_n, F\zeta_n) \le \rho_b(\zeta_{n-1}, F\zeta_{n-1})$ for all $n \ge 1$ which implies that $\{\rho_b(\zeta_{n-1}, \zeta_{n-1})\}$ is a non-increaing sequence. Hence, there exists $v \ge 0$ such that

$$\lim_{n\to\infty} \rho_b(\zeta_{n-1}, F\zeta_{n-1}) = v. \tag{18}$$

Now, we will show that $v = 0$. Suppose that $v > 0$. Since $\rho_b(\zeta_n, F\zeta_n) \le \rho_b(\zeta_{n-1}, F\zeta_{n-1})$ for all $n \ge 1$, using (17), we get

$$\begin{aligned}
\rho_b(\zeta_n, F\zeta_n) &\le \frac{s\lambda}{1 - s\mu_{n-1}}\rho_b(\zeta_{n-1}, F\zeta_{n-1})\\
&\quad \frac{s\lambda}{1 - sb}\rho_b(\zeta_{n-1}, F\zeta_{n-1}).
\end{aligned} \tag{19}$$

Taking limit on both sides of the inequality (19), we obtain that

$$v \leq \frac{s\lambda}{1-sb} v \leq v.$$

This implies that a contradiction. So, we have that $v = 0$. From (13), (18) we also have

$$\rho_b(\zeta_n, \zeta_{n+1}) \leq (1-\mu_n)\rho_b(\zeta_n, F\zeta_n) \leq \rho_b(\zeta_n, F\zeta_n)$$

which implies that

$$\lim_{n\to\infty} \rho_b(\zeta_n, \zeta_{n+1}) = 0. \tag{20}$$

Now, we need to show that $\{\zeta_n\}$ is a Cauchy sequence. For this, the similar proof method at the proof in [1] will be used. We suppose that $\{\zeta_n\}$ is not a Cauchy sequence. Thus, there exist $\overline{\epsilon}$ and the subsequences $\left\{\zeta_{\theta_{(k)}}\right\}$ and $\left\{\zeta_{\vartheta_{(k)}}\right\}$ of $\{\zeta_n\}$, such that $\theta(k) > \vartheta(k) > k$,

$$\rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)}}) \geq \overline{\epsilon}$$

and

$$\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)}}) < \overline{\epsilon}. \tag{21}$$

So, we have

$$\overline{\epsilon} \leq \rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)}}) \leq s\rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) + s\rho_b(\zeta_{\vartheta_{(k)+1}}, \zeta_{\vartheta_{(k)}}).$$

From (13) and (20), we get that

$$\frac{\overline{\epsilon}}{s} \leq \lim_{k\to\infty} \sup \rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}).$$

Noticing that

$$\begin{aligned}
\rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) &= \rho_b(w(\zeta_{\theta_{(k)-1}}, F\zeta_{\theta_{(k)-1}}; \mu_{\theta_{(k)-1}}), \zeta_{\vartheta_{(k)+1}}) \\
&\leq \mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}}) + (1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}}) \\
&\leq \mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}}) + s(1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\theta_{(k)-1}}, F\zeta_{\vartheta_{(k)+1}}) \\
&\quad + s(1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\vartheta_{(k)+1}}, \zeta_{\vartheta_{(k)+1}}) \\
&\leq \mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}}) + s(1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\vartheta_{(k)+1}}, \zeta_{\vartheta_{(k)+1}}) \\
&\quad s(1-\mu_{\theta_{(k)-1}})\lambda[\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}})]^{\delta} \times \\
&\quad [\rho_b(\zeta_{\theta_{(k)-1}}, F\zeta_{\theta_{(k)-1}})]^{\theta}[\rho_b(\zeta_{\vartheta_{(k)+1}}, F\zeta_{\vartheta_{(k)+1}})]^{1-\delta-\theta} \\
&\leq s\mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)}}) + s\mu_{\theta_{(k)-1}}\rho_b(\zeta_{\vartheta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) \\
&\quad + s(1-\mu_{\theta_{(k)-1}})\lambda[\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)+1}})]^{\delta} \times \\
&\quad [\rho_b(\zeta_{\theta_{(k)-1}}, F\zeta_{\theta_{(k)-1}})]^{\theta}[\rho_b(\zeta_{\vartheta_{(k)+1}}, F\zeta_{\vartheta_{(k)+1}})]^{1-\delta-\theta} \\
&\quad + s(1-\mu_{\theta_{(k)-1}})\rho_b(F\zeta_{\vartheta_{(k)+1}}, \zeta_{\vartheta_{(k)+1}}).
\end{aligned}$$

Taking $\lim\sup$ as $k \to \infty$ at the above inequality and using (18), (20) and (21) we obtain

$$\begin{aligned}
\frac{\overline{\epsilon}}{s} &\leq \lim_{k\to\infty} \sup \rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) \\
&\leq s\mu_{\theta_{(k)-1}}\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)}}) \\
&\leq sb\rho_b(\zeta_{\theta_{(k)-1}}, \zeta_{\vartheta_{(k)}}).
\end{aligned}$$

Since $b < \dfrac{1}{s^2}$, we obtain

$$\frac{\overline{\epsilon}}{s} \leq \lim_{k\to\infty} \sup \rho_b(\zeta_{\theta_{(k)}}, \zeta_{\vartheta_{(k)+1}}) \leq b\frac{\overline{\epsilon}}{s} < \frac{\overline{\epsilon}}{s}.$$

This is a contradiction. Hence, $\{\zeta_n\}$ is a Cauchy sequence in C. Since (C, ρ_b, w) be a ccbms, there exists $\zeta^* \in C$ such that $\lim_{n\to\infty} \zeta_n = \zeta^* \in C$. Finally, we will show that $\zeta^*$is a fixed point of F. Using (11) and (12), we have

$$\begin{aligned}
\rho_b(\zeta^*, F\zeta^*) &\leq s[\rho_b(\zeta^*, \zeta_n) + \rho_b(\zeta_n, F\zeta^*)] \\
&\leq s\rho_b(\zeta^*, \zeta_n) + s^2[\rho_b(\zeta_n, F\zeta_n) + \rho_b(F\zeta_n, F\zeta^*)] \\
&\leq s\rho_b(\zeta^*, \zeta_n) + s^2\rho_b(\zeta_n, F\zeta_n) \\
&\quad + s^2\lambda[\rho_b(\zeta_n, \zeta^*)]^\delta[\rho_b(\zeta_n, F\zeta_n)]^\theta[\rho_b(\zeta^*, F\zeta^*)]^{1-\delta-\theta} \\
&\leq s\rho_b(\zeta^*, \zeta_n) + s^2(\frac{s\lambda}{1-sb})^n\rho_b(\zeta_0, F\zeta_0) \\
&\quad + s^2\lambda[\rho_b(\zeta_n, \zeta^*)]^\delta[\rho_b(\zeta_n, F\zeta_n)]^\theta[\rho_b(\zeta^*, F\zeta^*)]^{1-\delta-\theta}.
\end{aligned}$$

If we take limit as $n \to \infty$ on both sides of above inequality, we get that $\rho(\zeta^*, F\zeta^*) = 0$ which implies that $F\zeta^* = \zeta^*$. Thus, the proof is complete. $\square$

References

[1] Chen, L., Li, C., Kaczmarek, R., Zhao, Y., Several Fixed Point Theorems in Convex b-Metric Spaces and Applications, Mathematics 2020, 8, 242; doi:10.3390/math8020242.

[2] Errai, Y., Marhrani, E. M., Aamri, M., Some Remarks on Fixed Point Theorems for Interpolative Kannan Contraction, Journal of Function Spaces, Volume 2020, Article ID 2075920, 7 pages.

[3] Gaba, Y. U., Aphane, M., Aydi, H., Interpolative Kannan Contractions in T0-Quasi-Metric Spaces, Journal of Mathematics, Volume 2021, Article ID 6685638, 5 pages.

[4] Kannan R. Some results on fixed points. Bull. Calc. Math. Soc. 1968, 60(1), 71-77.

[5] Karaca, N., Yildirim, I., Fixed Point Theorem for Reich Contraction Mapping in Convex b-Metric Spaces, The Aligarh Bulletin of Mathematics, 2021.

[6] Karapinar, E., Agarwal, R., Aydi, H., Interpolative Reich–Rus–Ciri´ c Type Contractions on Partial Metric Spaces, Mathematics 2018, 6, 256, doi:10.3390/math6110256.

[7] Karapinar, E., Revisiting the Kannan Type contractions via interpolation, Adv.Theory Nonlinear Anal. Appl., 2(2018), 85-87.

[8] Karapinar, E., O. Alqahtani, H. Aydi, On Interpolative Hardy-Rogers Type Contractions, Symmetry 2018, 11, 8; doi:10.3390/sym11010008.

[9] Karapinar, E., A Short Survey on the Recent Fixed Point Results on b-Metric Spaces, Constructive Mathematical Analysis, (2018), No. 1, pp. 15-44.

[10] Kadelburg, Z., Radenović, S., Notes on Some Recent Papers Concerning F-Contractions in b-Metric Spaces, Constructive Mathematical Analysis, 1 (2018), No. 2, pp. 108-112.

[11] Takahashi, W., A convexity in metric space and nonexpansive mappings, Kodai Math. Sem. Rep, 22 (1970) 142-149.

[12] Yildirim, I., Fixed Point Results for Interpolative Type Contractions via Mann Iteration Process, J. Adv. Math. Stud., Vol. 15(2022), No. 1, 35-45.

[13] Yildirim, I., On extended interpolative single and multi-valued F-contractions, Turk J Math, 46, (2022), 688-698.

[14] Zamfirescu, T., Fix point theorems in metric spaces, Arch. Math., 23 (1972), 292–298.

Department of Mathematics, Faculty of Science, Atatürk University, Erzurum 25240, Turkey. [1]

Email(s): isayildirim@atauni.edu.tr, yunussahin96@hotmail.com (corresponding author)

Interval Estimation for the Poisson Regression with Lognormal Unobserved Heterogeneity

Sümeyra Sert

ABSTRACT. The Poisson regression model is one of the widely used model for count data. Considering real world problems and data sets, assumptions are not always provided to perform Poisson regression analysis. Therefore, a need has arisen to consider different models based on Poisson regression model. Considering the studies in the literature, it is seen that the Poisson regression model with log-normal heterogeneity performs well and only credible intervals are obtained for these model parameters. Hence, this study considers the interval estimation for the Poisson regression parameters with lognormal unobserved heterogeneity to fill this gap in the literature. In this regard, the maximum likelihood based, bootstrap and the likelihood ratio confidence intervals are obtained. A numerical example is also provided to illustrate the discussed procedures. The results obtained from this study will provide foreknowledge for the practitioners studying with this topic.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 62, 62F15, 60J310, 65C40, 46N30, 46N70

KEYWORDS: Poisson regression, asymptotic maximum likelihood estimation, likelihood ratio, bootstrap, confidence intervals, unobserved heterogeneity.

1. INTRODUCTION

The Poisson regression model is widely used in many fields such as examining the number of graduate students based on their gender and GPA, number of traffic accidents based on the weather situation and other factors, etc. The Poisson model includes the assumption that the mean is equal to the variance, and hence the model cannot explain the empirical pattern of excessive scatter in the count data. In fact, the usual observation of excess zeros is a direct consequence of unobserved heterogeneity, as shown by Mullahy [1] and Tsionas [2]. As emphasized by Gourieroux [3], the existence of the unobserved heterogeneity can lead to inconsistent estimates and hence misleading interpretations. Many non-linear models are very sensitive to ignoring individual heterogeneity, and therefore, alternative models are preferred for Poisson regression models. Dean et al. [4] described a mixture of Poisson distributions based on the Inverse Gaussian distribution. Tsionas [2] developed a Bayesian inference procedure for the Poisson regression with lognormal unobserved heterogeneity. Recently, Ye et al. [5] proposed a semi-nonparametric Poisson regression model that can accommodate an unobserved heterogeneity unlike the traditional negative binomial model.

The aim of this study is to discuss interval estimation for Poisson regression parameters with lognormal unobserved heterogeneity. In this regard, maximum likelihood estimation (MLE) based confidence intervals, bootstrap confidence intervals and likelihood ratio (LR) confidence intervals are obtained. The performance of these methods are evaluated using interval length criteria.

2. POISSON REGRESSION MODEL WITH LOGNORMAL UNOBSERVED HETEROGENEITY

Let $\mathbf{y} = [y_1, y_2, ..., y_n]^{'}$ be a sample of count data. It is assumed that the sample is generated from a Poisson distribution with

$$p\left(y_i|\lambda_i\right) = \exp\left(-\lambda_i\right)\frac{\lambda_i}{y_i!}, i = 1, ..., n$$

where λ_i follows a lognormal distribution. Given a $k \times 1$ vector of covariates $\mathbf{x}_i$ and a $k \times 1$ parameter vector $\boldsymbol{\beta}$,

(1) $$\log \lambda_i = \mathbf{x}_i^{'}\boldsymbol{\beta} + u_i.$$

Here, $u_i \sim IN\left(0,\sigma^2\right)$. It is clear that when $\sigma = 0$, this model corresponds to the Poisson regression model with no unobserved heterogeneity (Tsionas [2]).

Let the $n \times k$ matrix $\mathbf{X} = \left[\mathbf{x_1'}, \mathbf{x_2'}, ..., \mathbf{x_n'}\right]'$ and the $n \times 1$ vector $\boldsymbol{\lambda} = \left[\lambda_1, \lambda_2, ..., \lambda_n\right]'$. Then, the marginal distribution of the data is given in Tsionas [2] by

$$p\left(y_i|\mathbf{x}_i, \boldsymbol{\beta}, \sigma\right) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} \int_0^\infty \exp\left(-\lambda_i\right) \frac{\lambda_i^{y_i-1}}{y_i!} \exp\left[-\frac{1}{2\sigma^2}\left(\log\lambda_i - \mathbf{x}_i'\beta\right)^2\right] d\lambda_i. \tag{2}$$

Then, the log-likelihood function of $\boldsymbol{\beta}$ and σ is

$$\ell\left(\boldsymbol{\beta}, \sigma\right) = \sum_{i=1}^{n} \log\left\{\left(2\pi\sigma^2\right)^{-\frac{1}{2}} \int_0^\infty \exp\left(-\lambda_i\right) \frac{\lambda_i^{y_i-1}}{y_i!} \exp\left[-\frac{1}{2\sigma^2}\left(\log\lambda_i - \mathbf{x}_i'\beta\right)^2\right] d\lambda_i\right\}. \tag{3}$$

This equation can be maximized to obtain the MLE of β and σ by using **maxLik** function in R.

3. Confidence intervals for the Poisson Regression with Lognormal Unobserved Heterogeneity

Let $\boldsymbol{\beta} = \left(\beta_0, \beta_1, \beta_2\right)^T$ is the regression parameters, $\boldsymbol{x} = \left(x_{i1}, x_{i2}, x_{i3}\right)^T$ are the covariates and $\widehat{\boldsymbol{\theta}} = \left(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{\sigma}\right)$ is the MLE of $\boldsymbol{\theta} = \left(\beta_0, \beta_1, \beta_2, \sigma\right)$. Here, x_{i1} is a $n \times 1$ dimensional vector consisting of 1's. Then, MLE based, bootstrap and likelihood-ratio based confidence intervals are considered and the details of these methods are given in below, respectively.

- An approximate $(1-\alpha)\,\%$ confidence intervals for β_0, β_1, β_2 and σ can be written by

$$\widehat{\beta}_0 \pm z_{1-\alpha/2}\widehat{se}\left(\widehat{\beta}_0\right), \tag{4}$$

$$\widehat{\beta}_1 \pm z_{1-\alpha/2}\widehat{se}\left(\widehat{\beta}_1\right), \tag{5}$$

$$\widehat{\beta}_2 \pm z_{1-\alpha/2}\widehat{se}\left(\widehat{\beta}_2\right), \tag{6}$$

and

$$\widehat{\sigma} \pm z_{1-\alpha/2}\widehat{se}\left(\widehat{\sigma}\right), \tag{7}$$

where $\widehat{se}\left(\widehat{\beta}_0\right)$, $\widehat{se}\left(\widehat{\beta}_1\right)$, $\widehat{se}\left(\widehat{\beta}_2\right)$ and $\widehat{se}\left(\widehat{\sigma}\right)$ can be computed by inverse of the Fisher information matrix.

- Normal bootstrap confidence interval is used following the methodology in Rizzo [6]. The algorithm is given in below.

1- Sample $x_1^*, x_2^*, ..., x_n^*$ with replacement from the original data sample. Let it be the empirical distribution.

2- Obtain the ML estimator of $\boldsymbol{\theta} = \left(\beta_0, \beta_1, \beta_2, \sigma\right)$ from each resample. This bootstrap estimate is denoted by $\hat{\boldsymbol{\theta}}_i^*$ for the i^{th} cycle.

3- Repeat (1,2) n times (n is bootstrap iterations).

The normal bootstrap (N-boot) CI based on $\hat{\boldsymbol{\theta}}^*$ is calculated as

$$CI_{Normal}^{1-\alpha} = \left(2\hat{\boldsymbol{\theta}}^* - \bar{\boldsymbol{\theta}}^* - z_{1-\frac{\alpha}{2}}\sqrt{S\left(\hat{\boldsymbol{\theta}}^*\right)}, 2\hat{\boldsymbol{\theta}}^* - \bar{\boldsymbol{\theta}}^* + z_{1-\frac{\alpha}{2}}\sqrt{S\left(\hat{\boldsymbol{\theta}}^*\right)}\right),$$

where $z_p$is p^{th} the quantile of the standard normal distribution,

$$\bar{\boldsymbol{\theta}}^* = \frac{1}{N_{Boot}} \sum_{i=1}^{N_{Boot}} \hat{\boldsymbol{\theta}}_i^*,$$

and

$$S\left(\hat{\boldsymbol{\theta}}^*\right) = \frac{1}{N_{Boot}-1} \sum_{i=1}^{N_{Boot}} \left(\hat{\boldsymbol{\theta}}_i^* - \bar{\boldsymbol{\theta}}^*\right)^2.$$

- Uncorrected Likelihood Ratio (ULR) confidence intervals are considered by Doganaksoy and Scheme [7] and Doganaksoy [8]. For this study, the likelihood ratio confidence intervals are obtained following Doganaksoy [8]. The likelihood ratio statistic for $\boldsymbol{\theta}$ is defined by
$$W(\boldsymbol{\theta}) = -2\left[\ell\left(\boldsymbol{\theta}, \tilde{\delta}_\theta\right) - \ell\left(\hat{\boldsymbol{\theta}}, \hat{\delta}\right)\right].$$
Here, $\boldsymbol{\theta}$ denotes the scale parameter vector of interest, δ is the vector of p nuisance parameters, $\left(\hat{\boldsymbol{\theta}}, \hat{\delta}\right)$ is the MLE of $(\boldsymbol{\theta}, \delta)$ and $\tilde{\delta}_\theta$ is the constrained MLE of δ under a given value of $\boldsymbol{\theta}$. Also, $\ell(\boldsymbol{\theta}, \delta)$ denotes the log-likelihood function, $\ell\left(\boldsymbol{\theta}, \tilde{\delta}_\theta\right)$ denotes the profile log-likelihood function for $\boldsymbol{\theta}$ and $\ell\left(\hat{\boldsymbol{\theta}}, \hat{\delta}\right)$ denotes the maximized value of the log-likelihood function. For large samples and under true parameter values, $W(\boldsymbol{\theta})$ is approximately distributed as $\chi^2_{(1)}$.
The test statistic Λ can be used for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta$ with critical region $\Lambda > \chi^2_{(1)}(1-\alpha)$. Then, we conclude that the ULR confidence interval for θ readily arises a nice set $\left\{\boldsymbol{\theta} : \Lambda < \chi^2_{(1)}(\alpha)\right\}$. Using this fact, $100(1-\alpha)\%$ ULR CI limits
$$(\theta_L, \theta_U) \tag{8}$$
that satisfy
$$-2\left(\ell(\boldsymbol{\theta}) - \ell\left(\hat{\boldsymbol{\theta}}\right)\right) - \chi^2_{(1)}(1-\alpha) = 0 \tag{9}$$
with $\theta_L < \hat{\theta}$ and $\theta_U > \hat{\theta}$.
It is noticed by Fraser [9] that the ULR and asymptotically normal (AN) CIs are asymptotically equivalent. The ULR CIs are transformation invariant, unlike the AN method. Furthermore, the ULR CIs always produce limits inside of the parameter space, as emphasized also in Sert et al. [10].

4. Numerical Example

In this section, a simulated data set is used to evaluate the performances of above-mentioned methods. In this context, the independent variables X_2 and X_3 are assumed to be $N(0,1)$, and the response variable Y has the pdf in Eq (1). $u_i \sim N(0, \sigma^2)$ and the parameters are prefixed as $\beta_0 = 0.5$, $\beta_1 = 0.5$, $\beta_2 = 0.5$, $\sigma = 1$. For bootstrap sampling, the re-sample size is prefixed as $Nboot = 100$. Also, the number of observations is prefixed $n = 100$. While **maxLik** function is used for the construction of MLE based and bootstrap confidence intervals, **optim** function is used to obtain lower and upper bounds for likelihood ratio based confidence intervals. R program is used for all computations. The results are given in Table (1)-(2).

	β_0	β_1	β_2	σ
ML estimates	0.3296	0.3812	0.4900	0.9343

Table 1. ML estimates for the model parameters

Method	Bound	β_0	β_1	β_2	σ
MLE	Lower	0.0507	0.1182	0.2086	0.7061
	Upper	0.6084	0.6443	0.7714	1.1625
Bootstrap	Lower	0.0737	0.1489	0.2171	0.7195
	Upper	0.6413	0.6187	0.7360	1.2299
LR	Lower	0.0450	0.1128	0.2028	0.3522.
	Upper	0.6141	0.6497	0.7772	1.5164

Table 2. Lower and upper bounds for the model parameters of numerical example

From Table (2), considering simulated data set, it is clear that LR based confidence interval performed best among all methods considered in this study. While there is not much difference between the results of two, one may prefer to use LR confidence intervals in terms of the length criteria instead of bootstrap confidence intervals to avoid long computation time.

5. Conclusion

In this study, interval estimation problem for the Poisson regression with lognormal unobserved heterogeneity is considered. The confidence intervals are constructed using the asymptotic normality of MLE, bootstrap and LR methods. Results indicate that LR method is superior than other methods in terms of interval length. The most important problem encountered during the study is that the program runs very slowly due to integrals in the likelihood function. Another problem should be considered is the lack of numerical integration methods to handle with divergent integrals in R. Further studies are planned to conduct to address these issues.

References

[1] Mullahy J., (1997), Heterogeneity, excess zeros, and the structure of count data models, Journal of applied econometrics, 12 (3) 337-350.

[2] Tsionas E.G., (2010), Bayesian analysis of Poisson regression with lognormal unobserved heterogeneity: With an application to the patent-R&D relationship, Communications in StatisticsTheory and Methods, 39 (10) 1689-1706.

[3] Gourieroux C., and Visser M., (1997), A count data model with unobserved heterogeneity, Journal of Econometrics, 79 (2) 247-268.

[4] Dean C., Lawless J. F., and Willmot G. E., (1989). A mixed poissoninversegaussian regression model, Canadian Journal of Statistics, 17 (2) 171-181.

[5] Ye X., Wang K., Zou Y., and Lord D., (2018). A semi-nonparametric Poisson regression model for analyzing motor vehicle crash data, PloS one, 13 (5) e0197338.

[6] Rizzo M. L., (2019). Statistical computing with R. Chapman and Hall/CRC.

[7] Doganaksoy N., and Schmee J. (2000). Practical Aspects of Corrected Likelihood Ratio Confidence Intervals: A Discussion of JengMeeker and WongWu, Technometrics, 42 (2) 156-159.

[8] Doganaksoy N. (2021). A Simplified Formulation of Likelihood Ratio Confidence Intervals Using a Novel Property, Technometrics, 63 (1) 127-135.

[9] Fraser D A.S. (1976). Probability and Statistics: Theory and Applications North Scituate, MA: Duxbury Press.

[10] Sert S., Abusaif .A., Akgen E., Karakaya K., and Ku C. (2022). Estimation and Prediction for the Half-Normal Distribution based on Progressively Type-II Censored Samples: Accepted-May 2022. REVSTAT-Statistical Journal.

Department of Statistics, Selçuk University, Konya, Turkiye

Email(s): sumeyra.sert@selcuk.edu.tr

The Effect of Reinsurance Contracts on Ruin Probability

Demet Sezer

ABSTRACT. Reinsurance is the transfer of some of the risks of a ceding insurance company to another insurance company. There are various types of reinsurance, both proportional and non-proportional, between ceding and reinsurance companies. Determining which of these reinsurance types will be more effective in reducing the probability of a company's ruin is a rather complex problem. In particular, according to different types of distributions such as symmetrical, left-skewed, right-skewed in terms of claim size distribution, determining which reinsurance contract will be more beneficial for the company in terms of the probability of ruin is a more complex problem. Solving such complex problems with simulation is much more practical than solving them theoretically. In this study, considering different types of claims and reinsurance contracts, the probability of ruin of companies has been tried to be obtained by simulation and it has been tried to determine which reinsurance contract is more suitable for which type of claim size.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 91G05, 62P05

KEYWORDS: Claim size, Reinsurance, Ruin

1. INTRODUCTION

The theoretical foundations of ruin theory (or risk process) were introduced by Lundberg [5]. This work was later republished by Cramer [3]. Risk process (or surplus process) is a model of the accumulation of an insurance company's capital at time t and can be defined as

$$U(t) = u + p(t) - S(t), \ t \geq 0. \tag{1}$$

In Equation (1), $U(t)$ represents insurance company's capital at time t, $p(t)$ represents the capital inflow (incoming cash premiums), u is the initial capital, $S(t)$ represents the capital outflow (outgoing claims) and can be defined as

$$S(t) = \sum_{i=1}^{N(t)} X_i \ , \ t \geq 0, \tag{2}$$

where $N(t)$ is the number of claims up to time t and X_i is the claim amount of i. In a risk model, different light-tailed (suh as exponential, gamma, hyperexponential) and heavy-tailed (such as Weibull, lognormal, Pareto) distributions can be used as the claim size distribution (for more details, see [1]).

Ruin of insurance company occurs when $U(t) < 0$ for some $t > 0$. This can be expressed as follows

$$\text{Ruin}= \{U(t) < 0, \ t > 0\}, \tag{3}$$

and the random variable T, called ruin time, can be defined as

$$T = \inf\{t > 0 : U(t) < 0\}. \tag{4}$$

Accordingly, ruin probability can be defined as

$$\psi(u) = P(\text{ruin} \mid U(0) = u) = P(T < \infty). \tag{5}$$

One can consult [4] for more details. Reinsurance is a type of insurance contract between two insurance companies. Carter [2] summarized the primary functions of reinsurance as follows:

- to protect insurers from underwriting losses which may imperil their solvency;
- to stabilize underwriting results;

- to increase the flexibility of an insurer in the size and types of risk and the volume of business he can underwrite;
- further spread the risk of loss.

Moreover, reinsurance may assist in the financing of insurance operations. There are various types of reinsurance, both proportional and non-proportional, between ceding and reinsurance companies. Determining which of these types of reinsurance will be more effective in reducing the probability of a company going ruin is a complex problem. In this study, considering different types of claims and reinsurance contracts, the probability of ruin of companies has been tried to be obtained by simulation and it has been tried to determine which reinsurance contract is more suitable for which type of claim size.

2. Types of Reinsurance

Reinsurance is the transfer of some of the risks of a ceding insurance company to another insurance company (reinsurance company or reinsurer). Reinsurance transactions are one of the basic elements of insurance. Basically, reinsurance transactions take place in the following three stages:

- The ceding insurance company transfers some of its risks to the reinsurance company.
- The ceding insurance company pays a premium to the reinsurance company for the risk it has transferred.
- In cases where the risk occurs, the compensation is paid by the reinsurance company to ceding insurance company.

The total claim amount of ceding insurance company $\{S(t), t \geq 0\}$ up to time t can be defined as

$$S(t) = \begin{cases} \sum_{i=1}^{N(t)} X_i & , \quad N(t) > 0 \\ 0 & , \quad N(t) = 0 \end{cases} \tag{6}$$

where $N(t), t \geq 0$ is the number of claims up to time t and $X_i, i \in N$ is the individual claim amount.

Considering that the total claim amount in a reinsurance contract is shared between the ceding company and the reinsurer, $S(t)$ can be expressed as follows

$$S(t) = D(t) + R(t) . \tag{7}$$

In Equation (7), $R(t)$ is the amount remaining in the ceding company after reinsurance and $D(t)$ is the amount to be paid by the reinsurance company. For most reinsurance contract, this split is expressed in terms of individual X_i risks as follows:

$$X_i = D_i + R_i. \tag{8}$$

There are various types of reinsurance, both proportional and non-proportional, between ceding and reinsurance companies.

2.1. **Proportional Reinsurance.** In these types of reinsurance contract, insurance premium and claims are shared between the insurance company and the reinsurer at a certain rate.

2.1.1. *Quota Share Reinsurance.* In a quota share reinsurance, R_i and $R(t)$ can be expressed as

$$R_i = aX_i \tag{9}$$

and

$$R(t) = aS(t) \tag{10}$$

where a $(0 < a < 1)$ is the proportionality factor. If the distribution of claim size X is known, the distribution functions of R and D can be obtained as

$$P(R \leq x) = P(aX \leq x) = P\left(X \leq \frac{x}{a}\right) = F_x\left(\frac{x}{a}\right) \tag{11}$$

and

$$P(D \leq x) = P((1-a)X \leq x) = P\left(X \leq \frac{x}{1-a}\right) = F_x\left(\frac{x}{1-a}\right) \tag{12}$$

respectively. Similarly, the distribution functions for the total risk can be can be expressed as

$$P(R(t) \leq x) = P(aS(t) \leq x) = P\left(S(t) \leq \frac{x}{a}\right) \tag{13}$$

and

$$P(D(t) \leq x) = P((1-a)S(t) \leq x) = P\left(S(t) \leq \frac{x}{1-a}\right) \tag{14}$$

2.1.2. *Surplus Share Reinsurance.* In a surplus share reinsurance, R_i and D_i can be expressed as

$$R_i = \left(1 - \frac{M}{Q_i}\right) X_i 1_{\{Q_i > M\}} \tag{15}$$

and

$$D_i = X_i 1_{\{Q_i \leq M\}} + M\frac{X_i}{Q_i} 1_{\{Q_i > M\}} \tag{16}$$

where, M is the retention and Q_i is the sum insured for claim X_i and $1_{\{x \in A\}}$ is the identity function defined as

$$1_{\{x \in A\}} = \begin{cases} 1 & , \quad x \in A \\ 0 & , \quad x \notin A \end{cases} \tag{17}$$

For total claim size, $R(t)$ and $D(t)$ can be expressed as follows

$$R(t) = \sum_{i=1}^{N(t)} R_i \tag{18}$$

and

$$D(t) = \sum_{i=1}^{N(t)} D_i \tag{19}$$

where $N(t)$ is the number of claim up to time t.

2.2. **Non-Proportional Reinsurance.** In these types of reinsurance contracts, if the insurer's loss exceeds a certain amount, the reinsurer is only obliged to pay the excess.

2.2.1. *Excess of Loss Reinsurance.* In a excess of loss reinsurance, $R(t)$ and $D(t)$ can be expressed as

$$R(t) = \sum_{i=1}^{N(t)} (X_i - M)_+ \tag{20}$$

and

$$D(t) = \sum_{i=1}^{N(t)} \min(X_i, M). \tag{21}$$

That is, the reinsurer will pay the excess amount for each claim exceeding the retention share. If the reinsurer will pay a certain upper limit of L, $R(t)$ and $D(t)$ can be defined as given in Equations (22) and (23).

$$R(t) = \sum_{i=1}^{N(t)} \min\left\{(X_i - M)_+, L\right\} \tag{22}$$

and

$$D(t) = \sum_{i=1}^{N(t)} \max\left\{\min(X_i, M), X_i - L\right\} \tag{23}$$

2.2.2. *Stop Loss Reinsurance.* In a stop-loss reinsurance, $R(t)$ can be expressed as

$$R(t) = \sum_{i=1}^{N(t)} (X_i - C)_+ \tag{24}$$

That is, the reinsurer deals with the portion of the total loss that exceeds a certain amount of C. If an upper limit such as d is added to the contract, $R(t)$ can be defined as given in Equation (25).

$$R(t) = \min\left\{\left(\sum_{i=1}^{N(t)} (X_i - C)_+ , d\right)\right\}. \tag{25}$$

3. Simulation Study

In this section, the Monte Carlo simulation (based on 10000 reputations) are conducted to obtain ruin probability of ceding companies in case of different claim size distributions (Pareto, lognormal, Weibull) and different reinsurance contracts (quota share reinsurance, excess of loss reinsurance). Also, simulation results given in Tables 1 - 3 have been obtained for different parameters of the considered claim size distributions and for different cases of the considered reinsurance contracts.

The following situations were taken into consideration while conducting the simulation study:

- while one of the distribution parameters is taken as constant, the other has been determined such that the expected value of the considered claim size distribution is equal to the given mean value.
- the number of claims distribution has been assumed as Poisson with λ parameter. In the simulation study, the results for different λ's (for 2, 5 and 8) were examined.
- the initial capital is taken as $u = 10$.
- the expected value principle has been used as the premium sharing method. Premium loading is taken 0.20 for insurer and 0.10 for ceding insurance company
- two cases have been considered for the quota share reinsurance contract: $a = 0.40$ and $a = 0.50$.
- two cases have been considered for the excess of loss reinsurance contract: $M = 5$ and $M = 8$.

In Table 1, simulation results are given for the case where the claim size distribution is Pareto.

			Ruin Probabilities for Reinsurance Contracts				
λ	Mean claim size	Parameters $(\alpha;\beta)$	No contract	Quota share $(a=0.4)$	Quota share $(a=0.5)$	Excess of loss $(M=5)$	Excess of loss $(M=8)$
2	3	$(4;9)$	0.4527	0.2440	0.1972	0.1390	0.2397
	5	$(4;15)$	0.5580	0.3695	0.3283	0.1328	0.2691
	7	$(4;21)$	0.6098	0.4335	0.3986	0.0924	0.2382
	10	$(4;30)$	0.6642	0.5065	0.4739	0.0661	0.2082
	15	$(4;45)$	0.6973	0.5685	0.5406	0.0469	0.1818
5	3	$(4;9)$	0.5121	0.2740	0.2265	0.1396	0.2445
	5	$(4;15)$	0.5950	0.3862	0.3466	0.1284	0.2664
	7	$(4;21)$	0.6564	0.4609	0.4256	0.1176	0.2680
	10	$(4;30)$	0.7025	0.5322	0.5003	0.0928	0.2537
	15	$(4;45)$	0.7365	0,5897	0.5653	0.0554	0.1933
8	3	$(4;9)$	0.5225	0.2818	0.2401	0.1397	0.2434
	5	$(4;15)$	0.6149	0.3928	0.3502	0.1221	0.2593
	7	$(4;21)$	0.6762	0.4835	0.4442	0.1049	0.2557
	10	$(4;30)$	0.7179	0.5411	0.5095	0.0961	0.2518
	15	$(4;45)$	0.7401	0.6010	0.5786	0.0657	0.2121

TABLE 1. Ruin Probabilities in the case of $X_i \sim \text{Pareto}(\alpha;\beta)$

In Table 2, simulation results are given for the case where the claim size distribution is lognormal.

			Ruin Probabilities for Reinsurance Contracts				
λ	Mean claim size	Parameters $(\mu;\sigma)$	No contract	Quota share $(a=0.4)$	Quota share $(a=0.5)$	Excess of loss $(M=5)$	Excess of loss $(M=8)$
2	3	$(1;0.444)$	0.3056	0.0753	0.0469	0.1644	0.1920
	5	$(1;1.104)$	0.5481	0.3773	0.3370	0.1070	0.2308
	7	$(1;1.375)$	0.5957	0.4631	0.4340	0.1017	0.2403
	10	$(1;1.614)$	0.5970	0.4947	0.4762	0.0865	0.2283
	15	$(1;1.848)$	0.6014	0.5157	0.5016	0.0121	0.0732
5	3	$(1;0.444)$	0.3279	0.0805	0.0516	0.1676	0.1933
	5	$(1;1.104)$	0.5985	0.4033	0.3669	0.1252	0.2592
	7	$(1;1.375)$	0.6633	0.5203	0.4908	0.0816	0.2221
	10	$(1;1.614)$	0.6820	0.5729	0.5543	0.0476	0.1603
	15	$(1;1.848)$	0.6707	0.5886	0.5763	0.0238	0.1103
8	3	$(1;0.444)$	0.3384	0.0827	0.0543	0.1634	0.1914
	5	$(1;1.104)$	0.6272	0.4192	0.3806	0.1340	0.2659
	7	$(1;1.375)$	0.6843	0.5259	0.5005	0.0892	0.2240
	10	$(1;1.614)$	0.7030	0.5885	0.5711	0.0298	0.1204
	15	$(1;1.848)$	0.7012	0.6138	0.6017	0.0094	0.0636

TABLE 2. Ruin Probabilities in the case of $X_i \sim$ lognormal $(\mu;\sigma)$

In Table 3, simulation results are given for the case where the claim size distribution is Weibull.

			Ruin Probabilities for Reinsurance Contracts				
λ	Mean claim size	Parameters $(k;\theta)$	No contract	Quota share $(a=0.4)$	Quota share $(a=0.5)$	Excess of loss $(M=5)$	Excess of loss $(M=8)$
2	3	$(0.7;2.369)$	0.4844	0.2772	0.2271	0.1292	0.2463
	5	$(0.7;3.949)$	0.5873	0.4052	0.3649	0.1078	0.2484
	7	$(0.7;5.529)$	0.6297	0.4660	0.4302	0.0841	0.2356
	10	$(0.7;7.899)$	0.6798	0.5392	0.5039	0.0611	0.2038
	15	$(0.7;11.849)$	0.7073	0.5959	0.5716	0.0501	0.1882
5	3	$(0.7;2.369)$	0.5397	0.3077	0.2614	0.1275	0.2508
	5	$(0.7;3.949)$	0.6337	0.4318	0.3884	0.1197	0.2671
	7	$(0.7;5.529)$	0.6755	0.4979	0.4641	0.0782	0.2172
	10	$(0.7;7.899)$	0.7202	0.5657	0.5383	0.0603	0.1990
	15	$(0.7;11.849)$	0.7447	0.6090	0.5865	0.0448	0.1804
8	3	$(0.7;2.369)$	0.5602	0.3140	0.2633	0.1280	0.2522
	5	$(0.7;3.949)$	0.6461	0.4344	0.3964	0.1269	0.2731
	7	$(0.7;5.529)$	0.6978	0.5174	0.4801	0.0960	0.2477
	10	$(0.7;7.899)$	0.7231	0.5671	0.5355	0.0603	0.1970
	15	$(0.7;11.849)$	0.7552	0.6178	0.5951	0.0375	0.1611

TABLE 3. Ruin Probabilities in the case of $X_i \sim$ Weibull $(k;\theta)$

Tables 1 - 3 show the ruin probabilities for situations with and without a reinsurance contract in the case of different claims number and different distribution parameters. When the tables are examined, the following situations are observed in all three tables:

- in all cases if there is no reinsurance contract, the probability of ruin is highest.
- in the quota share contract, the probability of ruin decreases as the rate (a) increases.
- in the excess of loss contract, the probability of ruin increases as retention (M) increases.
- as the mean claim size increases, the probability of ruin increases in quota share contract and decreases in the excess of loss contract.

4. Conclusion

Reinsurance transactions, which is one of the basic elements of insurance, is the transfer of some of the risks of a ceding insurance company to another insurance company. The reinsurance gives financial strength to companies that insure their risks and also allows the insurance industry to grow and operate on a wider scale. There are different types of reinsurance contracts. In this study, considering different claim size distributions such as Pareto, lognormal and Weibull and different proportional (quota share) and non-proportional (excess of loss) reinsurance contracts, ruin probability of ceding insurance companies is examined by a simulation study. In addition, for all cases discussed in the study, the probability of ruin in the absence of a reinsurance contract has been examined.

References

[1] Asmussen S., (2001), Ruin Probabilities, World Scientific, Singapore.
[2] Carter R. L., (1983), Reinsurance, Springer Science+Business Media Dordrecht.
[3] Cramer H., (1930), On the Mathematical Theory of Risk, Skandia Jubilee Volume,Stockholm.
[4] Dickson D. C. M., (2016), Insurance Risk and Ruin, Cambridge University Press, United Kingdom.
[5] Lundberg F., (1903), Approximerad framstllning af sannolikhetsfunctionen. II. terfrskring af kollektivrisken. Uppsala : Almqvist & Wiksell.

Selcuk University, Faculty of Science, Department of Actuarial Sciences, Konya, Turkiye

Email(s): dsezer@selcuk.edu.tr

Study of Some Approximation Estimates Concerning Convergence of (p,q)- Variant of Linear Positive Operators

Prerna Sharma

ABSTRACT. After studying a number of investigations in the approximation theory on (p,q)-calculus, this paper is a study on (p,q)- Szász-Mirakyan-Beta operators and their approximation properties in weighted space. We construct the operators and derive some lemmas as the auxiliary results. We establish a Voronovskya type asymptotic formula and also present some direct results using modulus of continuity.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 41A25, 41A20

KEYWORDS: (p,q)-calculus, (p,q)-Szász Beta operators, (p,q)-Beta function of second kind, Direct estimates.

1. Introduction

Linear positive operators proposed by Szász [21], for $0 \leq x < \infty$ are given as

$$S_n(f;x) = e^{-nx}\sum_{k=0}^{\infty}\frac{(nx)^k}{k!}f(\frac{k}{n}), f \in C[0,\infty). \tag{1}$$

In the same notions, Gupta and Noor [7] introduced the following sequence of linear positive operators for $0 \leq x < \infty$,

$$S_n(f;x) = \sum_{k=1}^{\infty}s_{n,k}(x)\int_0^{\infty}b_{n,k}(t)f(t)dt + s_{n,0}f(0), \tag{2}$$

where
$s_{n,k}(x) = e^{-nx}\frac{(nx)^k}{k!}, b_{n,k}(t) = \frac{1}{B(n+1,k)}\frac{t^{k-1}}{(1+t)^{n+k+1}}$.

They studied distinct approximation properties for these operators in simultaneous approximation. In order to approximate a function by linear positive operators, application of q-calculus is a very interesting area of research. The first q-analogue of Bernstein polynomials was given by Lupas in [9], after that Phillips [16] introduced another q-analogue of Bernstein polynomials. P. Maheshwari and other authors [10], [12] and [17] studied some approximation properties of certain q-genuine linear positive operators. Many q-generalizations of integral type operators and their approximation behaviors were intensively studied in ([2], [13], [14]).
Post quantum calculus is considered as an extension of quantum calculus. In this direction Acar [1] introduced the (p,q)-analogue of Szász Mirakyan operators. Aral and Gupta ([3] and [4]) gave (p,q)-variant of Szász-Beta operators and also proposed an integral version of these operators by using (p,q)-Beta function of second kind. By the application of (p,q)-Beta function of second kind (p,q)-genuine Baskakov-Durrmeyer operators were defined in [8].
After studying the operators given by the authors ([4], [8]), in the present paper we construct an integral modification of the sequence of linear positive operators defined by Mursaleen et al. [15] by using (p,q)-Beta function of second kind.
P. Maheshwari and M. Abid [11] published a paper on approximation of (p,q) Szasz-Beta-Stancu operators.
We mention some notations, definitions of (p,q)-calculus as follows (for details one can refer [19] and [20]).

The (p,q)-integer is defined as

$$[n]_{p,q} = \frac{p^n - q^n}{p-q},\ n = 0, 1, 2......,\ [0]_{p,q} = 0.$$

The (p,q)-factorial is defined as

$$[n]_{p,q}! = \prod_{m=1}^{n} [m]_{p,q},\ n \geq 1,\ [0]_{p,q}! = 1.$$

The (p,q)-binomial coefficient is given by

$$\left[\begin{array}{c} n \\ m \end{array} \right]_{p,q} = \frac{[n]_{p,q}!}{[n-m]_{p,q}![n]_{p,q}!},\ 0 \leq m \leq n.$$

For $p = 1$, these notations turn to the notation studied for q-analogues. It can easily be verified that $[n]_{p,q} = p^{n-1}[n]_{q/p}$.

Definition 1.1. *The (p,q)-power basis is defined below*

$$\begin{aligned} (x \oplus y)^n_{p,q} &= (x+y)(px+qy)(p^2x+q^2y)..........(p^{n-1}x+q^{n-1}y) \\ (x \ominus y)^n_{p,q} &= (x-y)(px-qy)(p^2x-q^2y)..........(p^{n-1}x-q^{n-1}y). \end{aligned}$$

Definition 1.2. [18] *For $n \geq 0$, the (p,q)-Gamma function is given as*

$$\Gamma_{p,q} n + 1 = \frac{(p \ominus q)^n_{p,q}}{(p-q)^n} = [n]_{p,q}!,\ 0 < q < p.$$

Definition 1.3. *The (p,q)-derivative of the function f is defined as*

$$D_{p,q} f(y) = \frac{f(py) - f(qy)}{(p-q)y},\ y \neq 0$$

and $D_{p,q} f(0) = f'(0)$, provided f is differentiable at zero.

Proposition 1.4. [18] *The (p,q)-integration by parts is defined as*

$$\int_a^b g(px) D_{p,q} h(x) d_{p,q} x = g(b)h(b) - g(a)h(a) - \int_a^b h(qx) D_{p,q} g(x) d_{p,q} x.$$

The (p,q)-Beta function of second kind [3] is given by

$$B_{p,q}(m,n) = \int_0^\infty \frac{x^{m-1}}{(1 \oplus px)^{m+n}_{p,q}} d_{p,q} x,$$

where $m, n \in N$.
The relation between (p,q)-Beta function and (p,q)-Gamma function is given as

$$B_{p,q}(m,n) = q^{\frac{2-m(m-1)}{2}} p^{\frac{-m(m+1)}{2}} \frac{\Gamma_{p,q} m \Gamma_{p,q} n}{\Gamma_{p,q}(m+n)}.$$

The two (p,q)-analogues of exponential functions are defined as

$$e_{p,q}(x) = \sum_{n=0}^{\infty} \frac{p^{\frac{n(n-1)}{2}}}{[n]_{p,q}!} x^n$$

and

$$E_{p,q}(x) = \sum_{n=0}^{\infty} \frac{q^{\frac{n(n-1)}{2}}}{[n]_{p,q}!} x^n,$$

which satisfy the equation $e_{p,q}(x) E_{p,q}(-x) = 1$.
For $p = 1$, we find the q-analogue of exponential functions.

Based on (p,q)-integers Mursaleen et al. [15] introduced (p,q)-Szász Mirakyan operators for $x \in [0,\infty)$ and for $0<q<p\leq 1$, as below

$$S_{n,p,q}(f;x)=\sum_{m=0}^{\infty} s_{n,p,q}^{m} f\left(\frac{[m]_{p,q}}{p^{m-1}[n]_{p,q}}\right), \tag{3}$$

where

$$s_{n,p,q}^{m}=\frac{p^{\frac{m(m-1)}{2}}}{q^{\frac{m(m-1)}{2}}}\frac{([n]_{p,q}x)^m}{[m]_{p,q}!}e_{p,q}(-[n]_{p,q}q^{-m}x).$$

Remark 1. [15] *For the operators defined in (3), for $x\geq 0$ and $n\in N$, the following equalities holds for $0<q<p\leq 1$*

$$S_{n,p,q}(1;x)=1,\ S_{n,p,q}=x,\ S_{n,p,q}(t^2;x)=\frac{x^2}{p}+\frac{x}{[n]_{p,q}},$$

$$S_{n,p,q}(t^3;x)=\frac{x^3}{p^3}+\frac{2p+q}{p^2[n]_{p,q}}x^2+\frac{x}{[n]_{p,q}^2},$$

$$S_{n,p,q}(t^4;x)=\frac{x^4}{p^6}+\frac{3p^2+2pq+q^2}{p^5[n]_{p,q}}x^3+\frac{3p^2+3pq+q^2}{p^3[n]_{p,q}^2}x^2+\frac{x}{[n]_{p,q}^3}.$$

In this paper, we have given some lemmas which are necessary to prove our main results. We obtain local approximation behaviors of the operators given in equation (3) in terms of second order modulus of smoothness and classical modulus of continuity. We also present uniform convergence theorems in terms of weighted approximation for the functions belonging to weighted spaces. In the last sections, we prove the Voronvskaya type asymptotic formula and direct estimates using Peetre's K-functional respectively.

1.1. **Preliminaries and Construction of operators.** For $x\in[0,\infty)$ and for $0<q<p\leq 1$, we present (p,q) Szász-Mirakyan-Beta operators using (p,q)-Beta function of second kind as

$$M_{n,p,q}(f;x)=\sum_{m=1}^{\infty} s_{n,p,q}^{m}\frac{1}{B_{p,q}(m,n+1)}\int_0^{\infty}\frac{t^{m-1}}{(1\oplus pt)_{p,q}^{m+n+1}}f(p^2q^mt)d_{p,q}t+\frac{f(0)}{e_{p,q}([n]_{p,q}x)}. \tag{4}$$

For $p=q=1$, these operators are the same as the operators defined by (2).

Lemma 1.5. *For $x\geq 0$ and $n\in N$, for the operators defined in (4), we have the following moments for $0<q<p\leq 1$*

$$\begin{aligned}
M_{n,p,q}(1;x) &= 1,\\
M_{n,p,q}(t;x) &= x,\\
M_{n,p,q}(t^2;x) &= \frac{[2]_{p,q}x}{pq[n-1]_{p,q}}+\frac{[n]_{p,q}x^2}{p^2[n-1]_{p,q}},\\
M_{n,p,q}(t^3;x) &= \frac{p^2[2]_{p,q}+pq(p+2q)+q^3}{p^3q^3[n-1]_{p,q}[n-2]_{p,q}}x\frac{(p^3+2pq[2]_{p,q}+q^3)[n]_{p,q}}{p^5q^2[n-1]_{p,q}[n-2]_{p,q}}x^2+\\
&\quad \frac{[n]_{p,q}^2}{p^6[n-1]_{p,q}[n-2]_{p,q}}x^3,\\
M_{n,p,q}(t^4;x) &= \frac{p^6+3p^5q+5p^4q^2+5p^3q^3+2p^2q^4+pq^3+q^6}{p^6q^6[n-1]_{p,q}[n-2]_{p,q}[n-3]_{p,q}}x+\\
&\quad \frac{[n]_{p,q}(p^7+3p^6q+4p^5q^2+2p^4q^3+2p^3q^2+p^2q^3+3p^2q^5+3pq^6+q^7)}{p^9q^5[n-1]_{p,q}[n-2]_{p,q}[n-3]_{p,q}}x^2+\\
&\quad \frac{[n]_{p,q}^2(p^3+3p^2q^3+3pq^4+q^5)}{p^{11}q^3[n-1]_{p,q}[n-2]_{p,q}[n-3]_{p,q}}x^3+\frac{[n]_{p,q}^3}{p^{12}[n-1]_{p,q}[n-2]_{p,q}[n-3]_{p,q}}x^4.
\end{aligned}$$

Proof. From the operators (4) and Remark 1, we have

$$\begin{aligned}
M_{n,p,q}(1;x) &= \sum_{m=1}^{\infty} s_{n,p,q}^{m} \frac{1}{B_{p,q}(m,n+1)} \int_{0}^{\infty} \frac{t^{m-1}}{(1\oplus pt)_{p,q}^{m+n+1}} d_{p,q}t + \frac{1}{e_{p,q}([n]_{p,q}x)} \\
&= S_{n,p,q}(1;x) \\
&= 1.
\end{aligned}$$

Next

$$\begin{aligned}
M_{n,p,q}(t;x) &= \sum_{m=1}^{\infty} s_{n,p,q}^{m} \frac{1}{B_{p,q}(m,n+1)} \int_{0}^{\infty} \frac{t^{m-1}}{(1\oplus pt)_{p,q}^{m+n+1}} (p^2q^m t) d_{p,q}t \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} \frac{p^2q^m}{B_{p,q}(m,n+1)} B_{p,q}(m+1,n) \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} p^{-m+1} \frac{[m]_{p,q}}{[n]_{p,q}} \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} \left(\frac{[m]_{p,q}}{p^{m-1}[n]_{p,q}}\right) \\
&= S_{n,p,q}(t;x) \\
&= x.
\end{aligned}$$

Similarly and by using the equality $[m+1]_{p,q} = p^m + q[m]_{p,q}$ and Remark 1, we have

$$\begin{aligned}
M_{n,p,q}(t^2;x) &= \sum_{m=1}^{\infty} s_{n,p,q}^{m} \frac{1}{B_{p,q}(m,n+1)} \int_{0}^{\infty} \frac{t^{m-1}}{(1\oplus pt)_{p,q}^{m+n+1}} (p^4q^{2m}t^2) d_{p,q}t \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} \frac{p^4q^{2m}}{B_{p,q}(m,n+1)} B_{p,q}(m+2,n-1) \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} q^{-1}p^{-2m+1} \frac{[m]_{p,q}[m+1]_{p,q}}{[n-1]_{p,q}[n]_{p,q}} \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} q^{-1}p^{-2m+1} \frac{[m]_{p,q}(p^m+q[m]_{p,q})}{[n-1]_{p,q}[n]_{p,q}} \\
&= \frac{1}{q[n-1]_{p,q}} S_{n,p,q}(t;x) + \frac{[n]_{p,q}}{p[n-1]_{p,q}} S_{n,p,q}(t^2;x) \\
&= \frac{[2]_{p,q}x}{pq[n-1]_{p,q}} + \frac{[n]_{p,q}x^2}{p^2[n-1]_{p,q}}.
\end{aligned}$$

Again using, $[m+1]_{p,q} = p^m + q[m]_{p,q}$, $[m+2]_{p,q} = p^{m+1} + qp^m + q^2[m]_{p,q}$ and Remark 1, we have

$$\begin{aligned}
M_{n,p,q}(t^3;x) &= \sum_{m=1}^{\infty} s_{n,p,q}^{m} \frac{1}{B_{p,q}(m,n+1)} \int_{0}^{\infty} \frac{t^{m-1}}{(1\oplus pt)_{p,q}^{m+n+1}} (p^6q^{3m}t^3) d_{p,q}t \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} \frac{p^6q^{3m}}{B_{p,q}(m,n+1)} B_{p,q}(m+3,n-2) \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} q^{-3}p^{-3m} \frac{[m]_{p,q}[m+1]_{p,q}[m+2]_{p,q}}{[n-2]_{p,q}[n-1]_{p,q}[n]_{p,q}} \\
&= \sum_{m=1}^{\infty} s_{n,p,q}^{m} q^{-3}p^{-3m} \frac{1}{[n-2]_{p,q}[n-1]_{p,q}[n]_{p,q}} \\
&\quad (p^{2m+1}[m]_{p,q} + p^{m+1}q[m]_{p,q}^2 p^{2m}q[m]_{p,q} + 2p^mq^2[m]_{p,q}^2 + q^3[m]_{p,q}^3)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{[n-2]_{p,q}[n-1]_{p,q}}\left(\frac{1}{q^3}+\frac{1}{q^2p}\right)S_{n,p,q}(t;x)+\\
&\quad \frac{[n]_{p,q}}{[n-2]_{p,q}[n-1]_{p,q}}\left(\frac{1}{q^2p}+\frac{2}{qp^2}\right)S_{n,p,q}(t^2;x)+\\
&\quad \frac{[n]_{p,q}^2}{[n-2]_{p,q}[n-1]_{p,q}}\frac{1}{q^3}+\frac{1}{p^3}S_{n,p,q}(t^3;x)\\
&= \frac{p^2[2]_{p,q}+pq(p+2q)+q^3}{p^3q^3[n-1]_{p,q}[n-2]_{p,q}}x\frac{(p^3+2pq[2]_{p,q}+q^3)[n]_{p,q}}{p^5q^2[n-1]_{p,q}[n-2]_{p,q}}x^2+\\
&\quad \frac{[n]_{p,q}^2}{p^6[n-1]_{p,q}[n-2]_{p,q}}x^3.
\end{aligned}
$$

Finally taking, $[m+1]_{p,q}=p^m+q[m]_{p,q}$, $[m+2]_{p,q}=p^{m+1}+qp^m+q^2[m]_{p,q}$, $[m+3]_{p,q}=p^{m+2}+qp^{m+1}+q^2p^m+q^3[m]_{p,q}$ and Remark 1, we have

$$
\begin{aligned}
M_{n,p,q}(t^4;x) &= \sum_{m=1}^{\infty} s_{n,p,q}^m\frac{1}{B_{p,q}(m,n+1)}\int_0^{\infty}\frac{t^{m-1}}{(1\oplus pt)_{p,q}^{m+n+1}}(p^8q^{4m}t^4)d_{p,q}t\\
&= \sum_{m=1}^{\infty} s_{n,p,q}^m\frac{p^8q^{4m}}{B_{p,q}(m,n+1)}B_{p,q}(m+4,n-3)\\
&= \sum_{m=1}^{\infty} s_{n,p,q}^m q^{-6}p^{-4m-2}\frac{[m]_{p,q}[m+1]_{p,q}[m+2]_{p,q}[m+3]_{p,q}}{[n-3]_{p,q}[n-2]_{p,q}[n-1]_{p,q}[n]_{p,q}}\\
&= \frac{p^3+2pq(p+q)+q^3}{p^3q^6[n-3]_{p,q}[n-2]_{p,q}[n-1]_{p,q}}S_{n,p,q}(t;x)+\\
&\quad \frac{[n]_{p,q}(p^3+3p^2q+4pq^2+2q^3)}{p^4q^5[n-3]_{p,q}[n-2]_{p,q}[n-1]_{p,q}}S_{n,p,q}(t^2;x)+\\
&\quad \frac{[n]_{p,q}^2(p^2+2pq+3q^2)}{p^5q^3[n-3]_{p,q}[n-2]_{p,q}[n-1]_{p,q}}S_{n,p,q}(t^3;x)+\\
&\quad \frac{[n]_{p,q}^3}{p^6[n-3]_{p,q}[n-2]_{p,q}[n-1]_{p,q}}S_{n,p,q}(t^4;x)\\
&= \frac{p^6+3p^5q+5p^4q^2+5p^3q^3+2p^2q^4+pq^3+q^6}{p^6q^6[n-1]_{p,q}[n-2]_{p,q}[n-3]_{p,q}}x+\\
&\quad \frac{[n]_{p,q}(p^7+3p^6q+4p^5q^2+2p^4q^3+2p^3q^2+p^2q^3+3p^2q^5+3pq^6+q^7)}{p^9q^5[n-1]_{p,q}[n-2]_{p,q}[n-3]_{p,q}}x^2+\\
&\quad \frac{[n]_{p,q}^2(p^3+3p^2q^3+3pq^4+q^5)}{p^{11}q^3[n-1]_{p,q}[n-2]_{p,q}[n-3]_{p,q}}x^3+\frac{[n]_{p,q}^3}{p^{12}[n-1]_{p,q}[n-2]_{p,q}[n-3]_{p,q}}x^4.
\end{aligned}
$$

□

Corollary 1.6. *Using Lemma 1.5, we immediately have the following formula for the central moments*

$$M_{n,p,q}((t-x);x)=0,$$
$$M_{n,p,q}((t-x)^2;x)=\frac{[2]_{p,q}x}{pq[n-1]_{p,q}}+\left(\frac{[n]_{p,q}}{p^2[n-1]_{p,q}}-1\right)x^2.$$

Remark 2. *As a special case for $p=q=1$, we get the moments of operators discussed in* [7].

2. Main results

In this section, we present the main results of this paper.

2.1. Weighted Approximation.

Remark 3. *Since* $\lim_{n\to\infty}[n]_{p,q}=\frac{1}{p-q}$ *for $q\in(0,1)$ and $p\in(q,1]$, therefore the operators defined in (4) are not in approximation process in weighted space. To study the convergence properties of these operators, we assume that $p=p_n$ and $q=q_n$ satisfy $0<q_n<p_n\leq 1$*

and $p_n \to 1,\ q_n \to 1,\ p_n^n \to 1,\ q_n^n \to 1$ *as* $n \to \infty$. *Here we note that with these assumptions* $\lim_{n\to\infty}[n]_{p,q} = \infty$.

Let us consider the functions in weighted space defined as

(1) $B_2[0,\infty)$ denotes the set of all functions f defined on $[0,\infty)$, such that $|f(x)| \leq M^f(1+x^2)$, where $M^f > 0$ depending only on f.

(2) $C[0,\infty)$ be the set of all continuous functions f defined on $[0,\infty)$.

(3) $C_2[0,\infty)$ denotes the sub space of all continuous functions in $B_2[0,\infty)$.

(4) By $C_2^*[0,\infty)$ represents the subspace of all functions $f \in C_2[0,\infty)$, for which $\lim_{x\to\infty}\frac{f(x)}{1+x^2}$ is finite.

$B_2[0,\infty)$ is a norm linear space with the norm

$$\|f\|_2 = \sup_{x\in[0,\infty)} \frac{|f(x)|}{1+x^2}.$$

Theorem 2.1. *Let* $p = p_n$ *and* $q = q_n$, *such that* $0 < q_n < p_n \leq 1$ *and* $p_n \to 1, q_n \to 1$, $p_n^n \to 1, q_n^n \to 1$ *as* $n \to \infty$. *Then for each* $f \in C_2^*[0,\infty)$, *we have*

$$\lim_{n\to\infty} \|M_{n,p_n,q_n}(f;x) - f\|_2 = 0.$$

Proof. According to [6], it is sufficient to verify the following three conditions

$$\lim_{n\to\infty} \|M_{n,p_n,q_n}(t^\alpha;x) - x^\alpha\|_2 = 0,\ \alpha = 0,1,2. \tag{5}$$

Since $M_{n,p_n,q_n}(1;x) = 1$ and $M_{n,p_n,q_n}(t;x) = x$, hence equation (5) holds for $\alpha = 0,1$. Next by using $[n]_{p_n,q_n} = p_n^{n-1} + q_n[n]_{p_n,q_n}$ and Lemma 1.5, we have for $n > 1$

$$\begin{aligned}
\|M_{n,p_n,q_n}(t^2;x) - x^2|_2 &= \|\left(\frac{[n]_{p_n,q_n}}{p_n^2[n-1]_{p_n,q_n}} - 1\right)x^2 + \frac{[2]_{p_n,q_n}x}{p_n,q_n[n-1]_{p_n,q_n}}\|_2 \\
&\leq \left(\frac{[n]_{p_n,q_n}}{p_n^2[n-1]_{p_n,q_n}} - 1\right)\sup_{x\in[0,\infty)}\frac{x^2}{1+x^2} \\
&+ \frac{[2]_{p_n,q_n}}{p_n,q_n[n-1]_{p_n,q_n}}\sup_{x\in[0,\infty)}\frac{x}{1+x^2} \\
&\leq \left(\frac{[n]_{p_n,q_n}}{p_n^2[n-1]_{p_n,q_n}} - 1\right) + \frac{[2]_{p_n,q_n}}{p_n,q_n[n-1]_{p_n,q_n}} \\
&\leq \frac{p_n^{n-2}}{[n-1]_{p_n,q_n}} + \left(\frac{q_n}{p_n^2} - 1\right) + \frac{[2]_{p_n,q_n}}{p_n,q_n[n-1]_{p_n,q_n}}.
\end{aligned}$$

Which implies

$$\lim_{n\to\infty}\|M_{n,p_n,q_n}(t^2;x) - x^2\|_2 = 0.$$

Therefore the equation (5) holds for $\alpha = 2$.
Hence the proof. □

2.2. **Voronovskaya type Asymptotic formula.**

Theorem 2.2. *Let* $p = p_n$ *and* $q = q_n$ *such that* $0 < q_n < p_n \leq 1$ *and* $p_n \to 1,\ q_n \to 1,$ $p_n^n \to 1,\ q_n^n \to 1$ *as* $n \to \infty$. *Then for any* $f \in C_2^*[0,\infty)$, *such that* $f', f'' \in C_2^*[0,\infty)$, *we have*

$$\lim_{n\to,\infty}[n]_{p_n,q_n}\left(M_{n,p_n,q_n} - f(x)\right) = \frac{f''(x)}{2}[2x + x^2(1+\beta)],$$

where

$$\beta = \lim_{n\to,\infty}[n]_{p_n,q_n}\left(\frac{q_n}{p_n^2} - 1\right)$$

uniformly on any $[0,L],\ L > 0$.

Proof. Let $f, f', f'' \in C_2^*[0,\infty)$ and $x \in [0,\infty)$. Then by Taylor's formula

$$f(t) = f(x) + f'(x)(t-x) + \frac{1}{2}f''(x)(t-x)^2 + \gamma(t,x)(t-x)^2, \tag{6}$$

where $\gamma(t,x)$ is the Peano form of the remainder.
Since $\gamma(\cdot,x) \in C_2^*[0,\infty)$, for sufficiently large n

$$\lim_{t\to x}\gamma(t,x) = 0.$$

Applying the operators (4) to both the sides of (6), we get
$M_{n,p_n,q_n}(f;x) - f(x) = f'(x)M_{n,p_n,q_n}((t-x);x) + \frac{1}{2}f''(x)M_{n,p_n,q_n}((t-x)^2;x)$

$$+M_{n,p_n,q_n}(\gamma(t,x)(t-x)^2;x).$$

By Cauchy-Schwarz-inequality, we have

$$M_{n,p_n,q_n}(\gamma(t,x)(t-x)^2;x) \leq \sqrt{D_{n,p_n,q_n}(\gamma^2(t,x);x)}\sqrt{D_{n,p_n,q_n}((t-x)^4;x)}. \tag{7}$$

We can see that $\gamma^2(x,x) = 0$ and $\gamma^2 \in C_2^*[0,\infty)$.
Using Theorem 2.1, we observe that

$$\lim_{n\to\infty} M_{n,p_n,q_n}(\gamma^2(t,x);x) = \gamma^2(x,x) = 0, \tag{8}$$

uniformly with respect to $x \in [0,L]$.
Therefore from equations (7) and (8), we get

$$\lim_{n\to\infty}[n]_{p_n,q_n} M_{n,p_n,q_n}(\gamma(t,x)(t-x)^2;x) = 0.$$

Now using Corollary 1.6, we have

$$M_{n,p_n,q_n}((t-x);x) = 0.$$

Thus

$$\begin{aligned}
\lim_{n\to\infty}[n]_{p_n,q_n}\left(M_{n,p_n,q_n}(f;x) - f(x)\right) &= \frac{1}{2}f''(x)\lim_{n\to\infty}[n]_{p_n,q_n} M_{n,p_n,q_n}((t-x)^2;x) \\
&= \frac{1}{2}f''(x)\lim_{n\to\infty}[n]_{p_n,q_n}\Bigg[\frac{[2]_{p_n,q_n}}{p_n, q_n, [n-1]_{p_n,q_n}}x + \\
&\quad \left(\frac{[n]_{p_n,q_n}}{p_n^2[n-1]_{p_n,q_n}} - 1\right)x^2\Bigg] \\
&= \frac{1}{2}f''(x)[2x + (1+\beta)x^2].
\end{aligned}$$

Hence proved. □

2.3. **Direct Results.** Let $C_B[0,\infty)$ be the space of all real valued continuous bounded functions f on the interval $[0,\infty)$, endowed with the norm

$$\|f\| = \sup_{x\in[0,\infty)}|f(x)|.$$

The Peetre's K-functional is defined as

$$K_2(f,\delta) = \inf_{g\in C_B^2[0,\infty)}\{\|f-g\| + \delta\|g''\|\},$$

where $C_B^2[0,\infty) = \{g \in C_B[0,\infty) : g', g'' \in C_B[0,\infty)\}$. By [5], there exists an absolute constant $C > 0$ such that

$$K_2(f,\delta) \leq C\omega_2(f,\sqrt{\delta}), \tag{9}$$

where $\delta > 0$ and the second order modulus of smoothness is defined by

$$\omega_2(f,\delta) = \sup_{0<h\leq\delta}\ \sup_{x\in[0,\infty)}|f(x+2h) - 2f(x+h) + f(x)|.$$

The usual modulus of continuity of $f \in C_B[0,\infty)$ is defined by

$$\omega(f,\delta) = \sup_{0<h\leq\delta} \sup_{x\in[0,\infty)} |f(x+h) - f(x)|. \tag{10}$$

Theorem 2.3. *Let $q \in (0,1)$ and $p \in (q,1]$, the operators $M_{n,p,q}$ maps $C_B[0,\infty)$ into $C_B[0,\infty)$ and*

$$\|(M_{n,p,q}(f;x))\| \leq \|f\|.$$

Proof. Using Lemma 1.5, we have

$$\|(M_{n,p,q}(f;x))\| = \sup_{0\leq x<\infty} |M_{n,p,q}(f;x)|.$$

$$\begin{aligned}
|M_{n,p,q}(f;x)| &\leq \sum_{m=1}^{\infty} s_{n,p,q}^{m} \frac{1}{B_{p,q}(m,n+1)} \int_0^{\infty} \frac{t^{m-1}}{(1\oplus pt)_{p,q}^{m+n+1}} |f(p^2q^mt)| d_{p,q}t + \frac{|f(0)|}{e_{p,q}([n]_{p,q}x)} \\
&\leq \sup_{0\leq x<\infty} |f(x)| \sum_{m=0}^{\infty} s_{n,p,q}^{m} \frac{1}{B_{p,q}(m,n+1)} \int_0^{\infty} \frac{t^{m-1}}{(1\oplus pt)_{p,q}^{m+n+1}} d_{p,q}t \\
&= \sup_{0\leq x<\infty} |f(x)| M_{n,p,q}(1;x) \\
&= \|f\|.
\end{aligned}$$

Hence the proof. □

Theorem 2.4. *For $f \in C_B[0,\infty)$ and for every $x \in [0,\infty)$, there exists a constant $C > 0$, such that*

$$|M_{n,p,q}(f;x) - f(x)| \leq C\omega_2\left(f; \sqrt{\frac{[2]_{p,q}}{pq[n-1]_{p,q}} + \left(\frac{[n]_{p,q}}{p^2[n-1]_{p,q}} - 1\right)x^2}\right).$$

where $p,q \in (0,1)$ and $q < p$.

Proof. Let $g \in W^2$, then from the Taylor's expansion, we have

$$g(t) = g(x) + g'(x)(t-x) + \int_x^t (t-u)g''(u)du,\ t \in [0,R],\ R > 0.$$

Now by Corollary 1.6, we get

$$M_{n,p,q}(g;x) = g(x) + M_{n,p,q}\left(\int_x^t (t-u)g''(u)du; x\right)$$

$$\begin{aligned}
|M_{n,p,q}(g;x) - g(x)| &\leq M_{n,p,q}\left(\int_x^t |(t-u)|\ |g''(u)|du; x\right) \\
&\leq M_{n,p,q}((t-x)^2;x)\|g''\|.
\end{aligned}$$

Therefore, we get

$$|M_{n,p,q}(g;x) - g(x)| \leq \|g''\|\left(\frac{[2]_{p,q}}{pq[n-1]_{p,q}} + \left(\frac{[n]_{p,q}}{p^2[n-1]_{p,q}} - 1\right)x^2\right).$$

Using $|M_{n,p,q}(f;x)| \leq \|f\|$, we have

$$\begin{aligned}
|M_{n,p,q}(f;x) - f(x)| &\leq |M_{n,p,q}((f-g);x) - (f-g)(x)| + |M_{n,p,q}(g;x) - g(x)| \\
&\leq \|f-g\| + \|g''\|\left(\frac{[2]_{p,q}}{pq[n-1]_{p,q}} + \left(\frac{[n]_{p,q}}{p^2[n-1]_{p,q}} - 1\right)x^2\right).
\end{aligned}$$

Now taking the infimum on the right hand side over all $g \in W^2$, we get

$$|M_{n,p,q}(f;x) - f(x)| \quad \leq \quad CK_2 \left(f; \sqrt{\frac{[2]_{p,q}}{pq[n-1]_{p,q}} + \left(\frac{[n]_{p,q}}{p^2[n-1]_{p,q}} - 1 \right) x^2} \right).$$

Using equation (9), we get

$$|M_{n,p,q}(f;x) - f(x)| \leq C\omega_2 \left(f; \sqrt{\frac{[2]_{p,q}}{pq[n-1]_{p,q}} + \left(\frac{[n]_{p,q}}{p^2[n-1]_{p,q}} - 1 \right) x^2} \right).$$

Hence the proof is completed. □

3. Conclusion

Here the author has derived the results for (p,q)- Szasz-Mirakyan-Beta operators using (p,q)-Beta function of second kind, similar convergence estimates can be investigated for these type of operators using p,q-Gamma function in the simultaneous approximation.

Acknowledgments

Author is thankful for organisers of the conference and reviewers for accepting her research paper for the publication in the proceedings of ICOMSS'22.

References

[1] Acar T., (2016), (p,q) generalization of Szász-Mirakyan operators, *Math. Method in the Appl. Sci. 39(10), 2685-2695.*

[2] Aral A., Gupta V. and Agarwal R. P., (2013), *Applications of q-Calculus in Operator theory,* Springer.

[3] Aral A. and Gupta V., (2016), (p,q)-type Beta functions of second kind, *Adv.in Operator Theory 1(1), 134-146.*

[4] Aral A. and Gupta V., (2017), (p,q)-variant of Szász Beta operators, *Rev. de la Real Acad. de Cien. Exact. Fis Y Nat. Ser. A Mat 111, 719-733.*

[5] Devore R. A. and Lorentz G. G., (1993), *Constructive Approximation*, Springer Berlin Germany.

[6] Gadžhiev A. D., (1976), Theorem of the type of P. P Korovkin type theorems, *presented at the international conference on the theory of approximation of functions* (Kaluga, 1975). Mat. Zametki 20(5), 781-786.

[7] Gupta V. and Noor M. A., (2006), Convergence of derivatives for certain mixed Szász Beta operators, *J. Math. Anal. Appl. 321(1), 1-9.*

[8] Gupta V. and Rassias T..M, (2016), (p,q)-Genuine Baskakov Derrmeyer operators, *Int. J. Nonlinear. Anal. Appl. 7(2), 69-76.*

[9] Lupas A.,(1987), A q-analogue of the Bernstein operators, *Seminar on Numerical and Statistical Calculus, Univ. of Cluj Napoc 9, 85-92.*

[10] Maheshwari P., (2018), Approximation properties of certain q-genuine Szasz operators, *Comp. Anal. and Ope. theory 12(1), 27-36.*

[11] Maheshwari P. and Abid M., (2019), Approximation by (p,q)-Szasz-Beta-Stancu operators operators, *Arabian J. of Mathematics (2019), DOI 10.1007/s40065-019-0240-1.*

[12] Maheshwari P. and Sharma D., (2012), Approximation by q-Baskakov-Beta-Stancu operators, *Rend. Circ. Mat.Palermo 61, 297-305.*

[13] Mahmudov N. I., (2010) On q-parametric Szász Mirakjan operators, *Mediterr. J. Math. 7(3), 297-311.*

[14] Mahmudov N. I., (2012), Approximation by q-parametric Szász Mirakjan operators, *Abstr. Appl. Anal. (2012), Article ID 751217..*

[15] Mursaleen M., Abied AAH Al. and Alotaibi A., (2017), On (p,q)-Szsz-Mirakyan operators and their approximation properties, *J. Ineq. Appl. 196.*

[16] Phillips G. M., (1997), Bernstein polynomials based on the q-integers, *Ann. Numer. Math. 4, 511-518.*

[17] Prakash O., Sharma D. and Maheshwari P., (2015), Certain Generalized q-operators, *Demonstratio Mathematica, 3(48, 404-412 .*

[18] Sadjang P. N., (2015), On the (p,q)-Gamma and (p,q)-Beta functions, *arXiv 1506. 07394v1.*

[19] Sadjang P. N., (2013), On fundamental theorem of (p,q)- calculus and some (p,q)-Taylor's formulas, *arXiv: 1309. 3934 [Math. QA].*

[20] Sahai V. and Yadav S., (2007), Repersentation of two parametric quantum algebra and (p,q)-special functions, *J. Math. Anal. Appl. 335 , 268-279.*

[21] Szász O., (1950), Generalisation of S. Bernstein's polynomials to the infinite interval, *J. Res. Natl. Bur. Stand. 45, 239-245.*

Department of Basic Science, Sardar Vallabh Bhai Patel University of Agriculture and Technology, Meerut (U.P.), India

Email(s): prernam2002@gmail.com

Asymptotic Properties of Spacings

Alexei Stepanov

ABSTRACT. In this work, the spacings based on order statistics obtained from continuous distributions are discussed. We present distributional results for spacings and a method of classification of distribution tails. By this method, we derive asymptotic results for spacings. We also obtain strong limit results for spacings.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 60, 62

KEYWORDS: order statistics, spacings, limit results.

1. INTRODUCTION

The spacings based on order statistics play an important role in many research areas such as goodness-of-fit tests, reliability analysis, survival analysis and inferential methods. A lot of practical problems and applications are associated with spacings. For example, the problems of using spacings in goodness-of-fit tests have been discussed in [5, 8]. The estimation method based on generalized spacings is known in the literature as a good alternative to the maximum likelihood method and the method of moments; see, for example, [6]. The works [9] and [12] utilize the lowest uniform spacings for testing hypotheses of randomness against alternative hypotheses suggestive of clustering. The papers [2] and [10] also use the lowest uniform spacings for studying scan statistics. For more details on spacings, one may refer to the books [1], [7], [11], [13] and the references contained therein. In the present work, we discuss some asymptotic properties of spacings in general continuous case.

1.1. **Preliminaries.** Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables with continuous cumulative distribution F and

$$X_{1,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$$

be the corresponding order statistics. Then,

$$S_{i,n} = X_{i+1,n} - X_{i,n} \ (i = 1, \ldots, n-1)$$

define the spacings based on these order statistics. Let us also assume that

$$\inf\{x : \ F(x) > 0\} = -\infty \quad \text{and} \quad \sup\{x : \ F(x) < 1\} = \infty.$$

One can raise the following question. Do the spacings tend to zero or infinity in probability or with probability one? We discuss this issue here. Let us start our discussion with analyzing the "middle" spacings $S_{[\alpha n],n}$, where $\alpha \in (0,1)$ and $[y]$ is the integral part of real number y. For continuous distributions, for any $\alpha \in (0,1)$, we have

$$X_{[\alpha n],n} \overset{a.s.}{\to} x_\alpha \quad \text{and} \quad X_{[\alpha n]+1,n} \overset{a.s.}{\to} x_\alpha,$$

where x_α is the α-quantile of F. For spacings in the "middle", we then get that

$$S_{[\alpha n],n} \overset{a.s.}{\to} 0, \quad n \to \infty.$$

For analyzing the limit behavior of lower and upper spacings $S_{k,n}$ and $S_{n-k,n}$, when k is fixed and $n \to \infty$, let us make use of "thickness" of the right and left distribution tails. For $s > 0$, let us consider the limits

$$\lim_{x\to\infty} \frac{1-F(s+x)}{1-F(x)} = \beta(s) \tag{1}$$

and

$$\lim_{x\to-\infty} \frac{F(x-s)}{F(x)} = \gamma(s). \tag{2}$$

The limits in (1) and (2) were previously used for studying the asymptotic behavior of the numbers of observations registered near lower and upper order statistics; see, for example, [3].

We will show in this work that the limit behaviors of $S_{k,n}$ and $S_{n-k,n}$, when k is fixed and $n \to \infty$, depend on the "thickness" of the left and right distribution tails $F(x)$ and $1-F(x)$, respectively. Studying asymptotic properties of upper and lower spacings, we can classify continuous distributions by "thickness" of their tails. Such classification for the right tail is proposed in Definition 1.1.

Definition 1.1. *If the limit in (1) exists for all $s > 0$ and if $\beta(s) = 1$, then we say that the underlying distribution F has a "thick" right tail, if $0 < \beta(s) < 1$, we then say that the distribution F has a "medium" right tail, and if $\beta(s) = 0$, then the distribution F has a "thin" right tail.*

Applying values of γ and using the consimilar argument, we can similarly classify the left tail of F.

2. Main results

In this section, we present our limit results. Upon analyzing the joint distribution of the order statistics $X_{k,n}$ and $X_{k+1,n}$, one can show that the distribution of the spacing $S_{k,n}$ has the form

$$\begin{aligned} F_{S_{k,n}}(s) &= 1 - \frac{n!}{(k-1)!(n-k)!} \\ &\times \int_{\mathbb{R}} F^{k-1}(x)(1-F(x+s))^{n-k} dF(x) \quad (s \geq 0,\ k = 1, \dots, n-1). \end{aligned}$$

Applying the above classification and the last identity, we can derive the following limit result.

Proposition 2.1. *(i) Let the limit in (1) exist for all $s > 0$. Then*

$$P(S_{n-k,n} > s) \to \beta^k(s) \quad (k \geq 1,\ n \to \infty).$$

(ii) Let the limit in (2) exist for all $s > 0$. Then

$$P(S_{k,n} > s) \to \gamma^k(s) \quad (k \geq 1,\ n \to \infty).$$

Corollary 2.2 readily results from Proposition 2.1.

Corollary 2.2. *(i) Let the right tail of F is "thin". Then*

$$S_{n-k,n} \xrightarrow{p} 0 \quad (k \geq 1,\ n \to \infty).$$

Let the right tail of F is "medium". Then

$$ES_{n-k,n} \to \int_0^\infty \beta^k(s) ds \quad (k \geq 1,\ n \to \infty),$$

provided that the integral in the last asymptotic expression exists. Let the right tail of F is "thick". Then

$$S_{n-k,n} \xrightarrow{p} \infty \quad (k \geq 1,\ n \to \infty)$$

and

$$ES_{n-k,n} \to \infty \quad (k \geq 1,\ n \to \infty).$$

(ii) Let the left tail of F is "thin". Then

$$S_{k,n} \xrightarrow{p} 0 \quad (k \geq 1,\ n \to \infty).$$

Let the left tail of F is "medium". Then

$$ES_{k,n} \to \int_0^\infty \gamma^k(s) ds \quad (k \geq 1,\ n \to \infty),$$

provided that the integral in the last asymptotic expression exists. Let the left tail of F is "thick". Then

$$S_{k,n} \xrightarrow{p} \infty \quad (k \geq 1,\ n \to \infty)$$

and

$$ES_{k,n} \to \infty \quad (k \geq 1,\ n \to \infty).$$

Applying further the modern form of Borel-Cantelli lemma, see [4], one can obtain the following strong limit result.

Proposition 2.3. *(i) Let $\beta(s) = 0$ and*

$$\int_{\mathbb{R}} \frac{F(x+s) - F(x-s)}{(1-F(x-s))^2} dF(x) < \infty. \tag{3}$$

Then

$$S_{n-k,n} \overset{a.s.}{\to} 0 \quad (k \geq 1,\ n \to \infty).$$

(ii) Let $\beta(s) = 1$ and (3) holds true, then

$$S_{n-k,n} \overset{a.s.}{\to} \infty \quad (k \geq 1,\ n \to \infty).$$

(iii) Let $\gamma(s) = 0$ and

$$\int_{\mathbb{R}} \frac{F(x+s) - F(x-s)}{F^2(x+s)} dF(x) < \infty. \tag{4}$$

Then

$$S_{k,n} \overset{a.s.}{\to} 0 \quad (k \geq 1,\ n \to \infty).$$

(iiii) Let $\gamma(s) = 0$ and (4) holds true. Then

$$S_{k,n} \overset{a.s.}{\to} \infty \quad (k \geq 1,\ n \to \infty).$$

3. Examples

In final Section 3, we illustrate the above theoretical results by examples.

Example 3.1. *Let*

$$F(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du \quad (x \in \mathbb{R})$$

be the standard normal distribution. The density function of it is symmetric over the origin and all the asymptotic results for $S_{k,n}$ follow from the corresponding asymptotic results for $S_{n-k,n}$. We present here only the results for $S_{n-k,n}$. For $s > 0$, we can show that

$$\lim_{x \to \infty} \frac{1 - \Phi(s+x)}{1 - \Phi(x)} = 0.$$

It follows from Corollary 2.2 that

$$S_{n-k,n} \overset{p}{\to} 0 \quad (k \geq 1,\ n \to \infty).$$

Identity (3) is also valid here. That way,

$$S_{n-k,n} \overset{a.s.}{\to} 0 \quad (k \geq 1,\ n \to \infty).$$

Upon analyzing the expected value of the spacing

$$ES_{n-k,n} = \frac{n!}{(n-k-1)!k!} \int_{\mathbb{R}} \int_{0}^{\infty} \Phi^{n-k-1}(x)(1-\Phi(x+s))^k ds d\Phi(x),$$

we can show here that

$$ES_{n-k,n} \to 0 \quad (k \geq 1,\ n \to \infty).$$

Example 3.2. *Let us take the extreme value distribution*

$$F(x) = e^{-e^{-x}} \quad (x \in \mathbb{R}).$$

Here, we have $\beta(s) = e^{-s}$ and $\gamma(s) = 0$, i.e., the left tail of the distribution F is "medium" and the right one is "thin". It follows from Proposition 2.1 and Corollary 2.2, respectively, that

$$P(S_{n-k,n} \leq s) \to 1 - e^{-ks} \quad (k \geq 1,\ n \to \infty),$$

$$ES_{n-k,n} \to \frac{1}{k}, \quad (k \geq 1,\ n \to \infty)$$

and

$$S_{k,n} \overset{p}{\to} 0, \quad (k \geq 1,\ n \to \infty).$$

Here (4) holds true and then

$$S_{k,n} \overset{a.s.}{\to} 0, \quad (k \geq 1,\ n \to \infty).$$

References

[1] Arnold B.C., Balakrishnan N., Nagaraja H.N., (1992) A first course in order statistics, John Wiley & Sons, New York.

[2] Balakrishnan N., Koutras M.V., (2002) Runs and scans with applications, Wiley, New York.

[3] Balakrishnan N., Stepanov, A., (2005), A note on the number of observations registered near an order statistic, *J. Statist. Plann. Inference*, 134, 1-14.

[4] Balakrishnan N., Stepanov A., (2010), Generalization of Borel-Cantelli lemma, *The Mathematical Scientist*, 35, 61-62.

[5] Beirlant J., Zuijlen M., (1985), The empirical distribution function and strong laws for functions of order statistics of uniform spacings, *J. Multivar. Anal.*, 16, 300-317.

[6] Cheng R.C.H., Amin N.A.K., (1983), Estimating parameters in continuous univariate distributions with a shifted origin, *J. Roy. Statist. Soc., Ser. B*, 45, 394–403.

[7] David H.A., Nagaraja H.N., (2003) Order statistics, 3rd edn. Wiley, Hoboken.

[8] Del Pino G.E., (1979), On the asymptotic distribution of k-spacings with applications to goodness-of-fit tests, *Ann. Statist.*, 7, 1058-1065.

[9] Ederer F., Meyers M.H., Mantel N., (1964), A statistical problem in space and time: Do leukemia cases come in clusters?, *Biometrics*, 20, 626-636.

[10] Glaz J., Balakrishnan N., (eds.) (1999) Scan statistics and applications, Birkhauser, Boston.

[11] Galambos J., (1978) The Asymptotic Theory of Extreme Order Statistics, John Wiley & Sons, NY.

[12] Naus J.I., (1966), Some probabilities, expectations and variances for the side of largest clusters and smallest intervals, *J. Amer. Statist. Ass.*, 61, 1191-1199.

[13] Nevzorov V.B., (2001) Records: Mathematical Theory, American Mathematical Society, Providence, Rhode Island.

Institute of Physics, Mathematics and Information Technology, Immanuel Kant Baltic Federal University, A.Nevskogo 14, Kaliningrad, 236041 Russia

Email(s): alexeistep45@mail.ru

Properties of Hurwitz Stable Matrix Families

Güner Topcu

ABSTRACT. In this work, matrix families which consist of linear and convex combination are given. The $\mathcal{I}_{\mathcal{L}}$ and $\mathcal{I}_{\mathcal{C}}$ intervals of the matrix families are determined so that the linear combination family $\mathcal{L}(A_1, A_2)$ and convex combination family $\mathcal{C}(A_1, A_2)$ are Hurwitz stable. κ^*-Hurwitz stability is examined similarly. Some properties of Hurwitz stable and κ^*-Hurwitz stable matrix families are given.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 34D20, 39A30, 93D05, 49K40

KEYWORDS: Hurwitz stabilty, differential equations, sensitivity.

1. INTRODUCTION

Let $A \in M_N(\mathbb{C})$ and $x(t)$-N dimensional vector and consider the following differential equation system:

$$\frac{d}{dt}x(t) = Ax(t). \tag{1}$$

It is well known that the stability of the system (1), in the literature. According to Lyapunov's theorem a necessary and sufficient condition for matrix A to be asymptotic stable is that the eigenvalues of the matrix A lay in the left open half-plane, that is, $\operatorname{Re}\lambda_i(A) < 0$ for all $i = 1, 2, \ldots, N$, where $\lambda_i(i = 1, 2, \ldots, N)$ stands for the eigenvalues of the coefficient matrix A [1, 2]. In the literature, this is also called spectral criterion. To denote the complex region with $\mathbb{C}$, system (1) is said to be Hurwitz stable if $\sigma(A) \subset C_H = \{z \mid \operatorname{Re}(z) < 0\}$ [8]. Let's also give the family of Hurwitz stable matrices as follows;

$$H_N = \{A \in M_N(\mathbb{C}) \mid \operatorname{Re}\lambda_i(A) < 0 \ \ (i = 1, 2, ..., N)\}.$$

In other words, the eigenvalues of the matrix A formed for the differential equation system (1), fall into the left open half-plane, which determines the Hurwitz stability of this system .

If the locations of these eigenvalues are known approximately, stability analysis of the system can be done with help of many well-known methods. However, in practice it is not so easy to determine the eigenvalue therefore the eigenvalue problem is an ill-posed problem for nonsymmetric matrices [9]. More precisely, small changes in the inputs of the matrices lead to big changes in the eigenvalues. Let's give the following example for this situation [1, 2].

$$A_w(L) = \begin{pmatrix} -20 & 2.5L & & & & \\ 0 & -19 & 3.5L & & 0 & \\ \vdots & & \ddots & \ddots & & \\ \vdots & & & \ddots & \ddots & \\ 0 & 0 & \dots & & -2 & 20.5L \\ w & 0 & 0 & \dots & 0 & -1 \end{pmatrix}$$

If we write $L = 10$ in $w = 1.5L^{-19}$then we get $w = 1.5 \times 10^{-19}$. The matrix $A_0(10)$ is the Hurwitz stable but $A_w(10)$ is non-Hurwitz stable since $0.5 \in \sigma(A)$. Therefore, for the determination of stability, it is more convenient to use the parameters calculated with the help of the solution of a linear algebraic equation which characterizing the stability. With this notion for linear systems, the stability problem is reduced to the problem of the existence of a positive definite solution of Lyapunov matrix equation

$$A^*F + FA + I = 0. \tag{2}$$

According to Lyapunov's theorem, if the Lyapunov matrix equation for the differential system (1), which determines the asymptotic stability of the systems, has a solution

$$F = \int_0^\infty \left(e^{tA}\right)^* e^{tA} dt, \; F = F^* > 0 \tag{3}$$

then system of differential equations (1) is said to be Hurwitz stable [1, 2, 6]. The existence of $F = F^* > 0$ equivalent to having the eigenvalues of the matrix A in the left open half-plane. In other words, the stability of a system means the same as the stability of the coefficient matrix A. However, existence of F does not enough to comment on quality. So we need a parameter which determines the quality of the stability.

Hurwitz stability parameter or simply condition number $\kappa(A)$ for the system (1) is defined as $\kappa(A) = 2\|A\|\|F\| \geq 1$ [3, 4]. Moreover, κ^* be the practical Hurwitz stability parameter of the system (1), where $1 < \kappa^* \in \mathbb{R}$ and the value κ^* chosen by user in view of their problem. If $\kappa(A) \leq \kappa^*$ then the matrix A is κ^*- Hurwitz stable matrix. Otherwise, the matrix A is κ^*-Hurwitz unstable matrix [3, 4, 5].

Let's take Hurwitz stable matrix set as follow,

$$H_N = \{A \in M_N(\mathbb{C}) \mid \kappa(A) < \infty\}.$$

Let's give the matrix families

$$\mathcal{L}(A_1, A_2) = \{A(r_{\mathcal{L}}) = A_1 + r_{\mathcal{L}} A_2 \mid A_1,\; A_2 \in M_N(\mathbb{C})\}$$

and

$$\mathcal{C}(A_1, A_2) = \{A(r_{\mathcal{C}}) = (1 - r_{\mathcal{C}}) A_1 + r_{\mathcal{C}} A_2 \mid A_1,\; A_2 \in M_N(\mathbb{C})\}$$

which consist of linear and convex combination of $A_1,\; A_2 \in M_N(\mathbb{C})$, respectively. Here, for the sake of brevity the indices will be $\mathcal{L}$ and $\mathcal{C}$ instead of $\mathcal{L}(A_1, A_2)$ and $\mathcal{C}(A_1, A_2)$, respectively. If $A_1 \in H_N,\; A_2 \in M_N(\mathbb{C})$ and $r_{\mathcal{L}} \in \mathcal{I}_{\mathcal{L}}$ then the matrix family $\mathcal{L}(A_1, A_2)$ is Hurwitz stable [7]. Similarly, if $A_1 \in H_N,\; A_2 \in M_N(\mathbb{C})$ and $r_{\mathcal{C}} \in \mathcal{I}_{\mathcal{C}}$ then the matrix family $\mathcal{C}(A_1, A_2)$ is Hurwitz stable [7]. These all results known as continuity theorems in the literature. In the remainder of the paper, algebraic properties of $\mathcal{L}(A_1, A_2)$ and $\mathcal{C}(A_1, A_2)$ matrix families introduced. Here these algebraic properties concern about existence, uniqueness, convex combination, arithmetic sum, difference of matrix families $\mathcal{L}(A_1, A_2)$ and $\mathcal{C}(A_1, A_2)$.

2. Properties for Matrix Families

Let's take, $A_1 \in H_N$ and $A_2 \in M_N(\mathbb{C})$ then $\mathcal{L}(A_1, A_2), \mathcal{C}(A_1, A_2) \subset H_N$ with $r_{\mathcal{L}} \in \mathcal{I}_{\mathcal{L}}$ and $r_{\mathcal{C}} \in \mathcal{I}_{\mathcal{C}}$, respectively. Accordingly, the following properties are obtained for the $\mathcal{L}(A_1, A_2)$ and $\mathcal{C}(A_1, A_2)$ matrix families:

(1) $\mathcal{L}(A_1, A_2) \neq \mathcal{L}(A_2, A_1)$.
Let's take $A^1 \in \mathcal{L}(A_1, A_2)$ and $A^2 \in \mathcal{L}(A_2, A_1)$. If we write these equalities, $A^1 = A_1 + rA_2$ and $A^2 = A_2 + rA_1$ are obtained. It is easily seen that, this equality is not satisfied unless $A_1 = A_2$ specifically.

(2) $A \in \mathcal{L}(A_1, A_2)$ there is exist and unique where the interval which $r \in \mathcal{I}_{\mathcal{L}}$.
Indeed, we say the existence of $r \in \mathcal{I}_{\mathcal{L}}$ from the definition of the $A \in \mathcal{L}(A_1, A_2)$ for $A = A_1 + rA_2$. In addition, let be $A = (a_{ij})$,$A_1 = \left(a_{ij}^1\right)$, $A_2 = \left(a_{ij}^2\right)$

$$A \in \mathcal{L}(A_1, A_2) \iff \exists r = \frac{a_{ij} - a_{ij}^1}{a_{ij}^2} \; (i, j = 1, 2, ..., N).$$

On the other hand, if $a_{ij}^2 = 0$ then it means $A_2 = 0$ and automatically get $A = A_1$ i.e. $a_{ij} = a_{ij}^1$.

(3) $A,\; B \in \mathcal{L}(A_1, A_2) \implies \frac{1}{2}(A + B) \in \mathcal{L}(A_1, A_2)$.
Let be $A \in \mathcal{L}(A_1, A_2) \ni A = A_1 + r_1 A_2$ and $B \in \mathcal{L}(A_1, A_2) \ni B = A_1 + r_2 A_2$ for $r_1,\; r_2 \in \mathcal{I}_{\mathcal{L}}$. Since $\frac{r_1 + r_2}{2} = r_3 \in \mathcal{I}_{\mathcal{L}}$ for $r_1,\; r_2 \in \mathcal{I}_{\mathcal{L}}$

$$\frac{1}{2}(A + B) = A_1 + \frac{r_1 + r_2}{2} A_2 = A_1 + r_3 A_2.$$

So we can write $A_1 + r_3 A_2 \in \mathcal{L}(A_1, A_2)$.

(4) $A,\ B\in\mathcal{L}(A_1,A_2) \implies (A-B)\notin\mathcal{L}(A_1,A_2)$.

Indeed, let's take $A=A_1+r_1A_2$ and $B=A_1+r_2A_2$ for $A,\ B\in\mathcal{L}(A_1,A_2)$. Then, we get,

$$A-B=(r_1-r_2)\,A_2\notin\mathcal{L}(A_1,A_2)\,.$$

However, this difference is a member of the $\mathcal{L}(0,A_2)$ matrix family.

(5) The convex sum of matrices A and B is also an element of $\mathcal{L}(A_1,A_2)$ for $A,\ B\in\mathcal{L}(A_1,A_2)$, i.e. $\alpha A+\beta B\in\mathcal{L}(A_1,A_2)$ for $\alpha+\beta=1$ and $\alpha,\ \beta>0$.

Let's take $A=A_1+r_1A_2,\ B=A_1+r_2A_2\in\mathcal{L}(A_1,A_2)$. If we write

$$\alpha A+\beta B=(\alpha+\beta)\,A_1+(\alpha r_1+\beta r_2)\,A_2.$$

We know that $\alpha+\beta=1$ from convex sum and $\alpha r_1+\beta r_2\in\mathcal{I}_{\mathcal{L}}$ from $r_1,r_2\in\mathcal{I}_{\mathcal{L}}$ so we get

$$\alpha A+\beta B\in\mathcal{L}(A_1,A_2)\,.$$

(6) $\mathcal{L}(A_1,A_2)$ Hurwitz stable matrix family then $A_1\in\mathcal{L}(A_1,A_2)$ when $r=0$ also $A_2\notin\mathcal{L}(A_1,A_2)$ for $A_1\neq 0$.

Let's take $A\in\mathcal{L}(A_1,A_2)$ so we have $A=A_1+rA_2$. From this equality, if we specially choose $r=0$ then we get $A=A_1\in\mathcal{L}(A_1,A_2)$ and unless $A_1=0$, A_2 is not element of $\mathcal{L}(A_1,A_2)$.

(7) Let $\mathcal{I}^1_{\mathcal{L}}\subset\mathcal{I}^2_{\mathcal{L}}$ then $\mathcal{L}^1(A_1,A_2)\subset\mathcal{L}^2(A_1,A_2)$.

(8) $\mathcal{C}(A_1,A_2)\neq\mathcal{C}(A_2,A_1)$ for $r\neq 0.5$.

Let's take $A^1\in\mathcal{C}(A_1,A_2)$ and $A^2\in\mathcal{C}(A_2,A_1)$. If written these equalities, $A^1=(1-r)\,A_1+rA_2$ and $A^2=(1-r)\,A_2+rA_1$ are obtained. It is easily seen that, this equality is not satisfied unless $r=0.5$ specifically.

(9) $A\in\mathcal{C}(A_1,A_2)$ there is exist and unique where the interval which $r\in\mathcal{I}_{\mathcal{C}}$.

Indeed, we say the existence of $r\in\mathcal{I}_{\mathcal{C}}$ from the definition of the $A\in\mathcal{C}(A_1,A_2)$ for $A=(1-r)\,A_1+rA_2$. In addition, let be $A=(a_{ij})$,$A_1=\left(a^1_{ij}\right)$, $A_2=\left(a^2_{ij}\right)$

$$A\in\mathcal{C}(A_1,A_2)\iff\exists r=\frac{a_{ij}-a^1_{ij}}{a^2_{ij}-a^1_{ij}}\ \ (i,j=1,2,...,N)\,.$$

On the other hand, if $a^2_{ij}=a^1_{ij}$ then it means $A_2=A_1$ and automatically get $A=A_1$ i.e. $a_{ij}=a^1_{ij}$.

(10) $A,\ B\in\mathcal{C}(A_1,A_2) \implies \frac{1}{2}(A+B)\in\mathcal{C}(A_1,A_2)\,.$

Let be $A\in\mathcal{C}(A_1,A_2)\ni A=(1-r_1)\,A_1+r_1A_2$ and $B\in\mathcal{C}(A_1,A_2)\ni B=(1-r_2)\,A_1+r_2A_2$ for $r_1,\ r_2\in\mathcal{I}_{\mathcal{C}}$. Since $\frac{r_1+r_2}{2}=r_3\in\mathcal{I}_{\mathcal{C}}$ for $r_1,\ r_2\in\mathcal{I}_{\mathcal{C}}$

$$\frac{1}{2}(A+B)=\left(1-\frac{r_1+r_2}{2}\right)A_1+\frac{r_1+r_2}{2}A_2=(1-r_3)\,A_1+r_3A_2.$$

So we can write $(1-r_3)\,A_1+r_3A_2\in\mathcal{C}(A_1,A_2)\,.$

(11) $A,\ B\in\mathcal{C}(A_1,A_2) \implies (A-B)\notin\mathcal{C}(A_1,A_2)$.

Indeed, let's take $A=(1-r_1)\,A_1+r_1A_2$ and $B=(1-r_2)\,A_1+r_2A_2$ for $A,\ B\in\mathcal{C}(A_1,A_2)$. Then, we get,

$$A-B=-(r_1-r_2)\,A_1+(r_1-r_2)\,A_2\notin\mathcal{C}(A_1,A_2)\,.$$

(12) The convex sum of matrices A and B is also an element of $\mathcal{C}(A_1,A_2)$ for $A,\ B\in\mathcal{C}(A_1,A_2)$, i.e. $\alpha A+\beta B\in\mathcal{C}(A_1,A_2)$ for $\alpha+\beta=1$ and $\alpha,\ \beta>0$.

Let's take $A=(1-r_1)\,A_1+r_1A_2,\ B=(1-r_2)\,A_1+r_2A_2\in\mathcal{C}(A_1,A_2)$. If we write

$$\alpha A+\beta B=((\alpha+\beta)-(\alpha r_1+\beta r_2)\,A_1+(\alpha r_1+\beta r_2)\,A_2.$$

We know that $\alpha+\beta=1$ from convex sum and $\alpha r_1+\beta r_2\in\mathcal{I}_{\mathcal{C}}$ from $r_1,r_2\in\mathcal{I}_{\mathcal{C}}$ so we get,

$$\alpha A+\beta B\in\mathcal{C}(A_1,A_2)\,.$$

(13) $\mathcal{C}(A_1,A_2)$ Hurwitz stable matrix family then $A_1\in\mathcal{C}(A_1,A_2)$ when $r=0$ and $A_2\in\mathcal{C}(A_1,A_2)$ if $r=1$.

Let's take $A\in\mathcal{C}(A_1,A_2)$ so we have $A=(1-r)\,A_1+rA_2$. From this equality, if we specially choose $r=0$ then we get $A=A_1\in\mathcal{C}(A_1,A_2)$ and $r=1$ then we get $A=A_2\in\mathcal{C}(A_1,A_2)\,.$

(14) Let $\mathcal{I}^1_{\mathcal{C}}\subset\mathcal{I}^2_{\mathcal{C}}$ then $\mathcal{C}^1(A_1,A_2)\subset\mathcal{C}^2(A_1,A_2)$.

(15) $\mathcal{L}(A_1, A_2)$ and $\mathcal{C}(A_1, A_2)$ Hurwitz stable matrix family then $A_1 \in \mathcal{L} \cap \mathcal{C}$, i.e. $\mathcal{L} \cap \mathcal{C} \neq \emptyset$.
From the above properties it can be seen easily.

Remark 1. *The properties given above are valid for Hurwitz stability of $\mathcal{L}(A_1, A_2)$ and $\mathcal{C}(A_1, A_2)$ matrix families. The set of practical Hurwitz stable matrix is shown with*

$$H_N^* = \{A \in H_N \mid \kappa(A) \leq \kappa^*\}.$$

if $A_1 \in H_N^*$, $A_2 \in M_N(\mathbb{C})$ and $r_{\mathcal{L}} \in \mathcal{I}_{\mathcal{L}}^*$ and $r_{\mathcal{C}} \in \mathcal{I}_{\mathcal{C}}^*$ [7]*, respectively, the above properties are valid for κ^*-Hurwitz stability of matrix families $\mathcal{L}(A_1, A_2)$ and $\mathcal{C}(A_1, A_2)$ matrix families in general.*

3. Conclusion

In this study, new matrix families $\mathcal{L}(A_1, A_2)$ and $\mathcal{C}(A_1, A_2)$ based on linear and convex summation were recalled, respectively. The $\mathcal{I}_{\mathcal{L}}$ and $\mathcal{I}_{\mathcal{C}}$ intervals that make these families Hurwitz stable were reminded. Similarly, $\mathcal{I}_{\mathcal{L}}^*$ and $\mathcal{I}_{\mathcal{C}}^*$ intervals that will provide κ^*-Hurwitz stability of matrix families $\mathcal{L}$ and $\mathcal{C}$ were also recalled. Using these matrix families and their Hurwitz and κ^*-Hurwitz stability intervals, some algebraic properties were given.

References

[1] Bulgak(ov) Y. A., (1995), Matrix computations with guaranteed accuracy in stability theory, Selçuk University.

[2] Bulgak H., (1999), Pseudoeigenvalues, spectral portrait of a matrix and their connections with different criteria of stability, Error Control and Adaptivity in Scientific Computing, pp. 95-124.

[3] Bulgakov H. Ya., (1980), Effectively calculable parameter of stability quality of systems of linear differential equations with constant coefficients, Sibirskii Matematicheskii Zhurnal, vol. 21(3), pp. 32-41.

[4] Bulgakov A. Ya. and Godunov S. K., (1988), Circle dichotomy of the matrix spectrum, Siberia Math. J., vol. 29(5), pp. 59-70.

[5] Bulgak H., (2020), Pseudospectrum and Different Quality of Stability Parameters, Continuum Mechanics, Applied Mathematics and Scientific Computing: Godunov's Legacy, Springer, Cham, pp. 61-67.

[6] Lyapunov, A. M., (1892), The general problem of motion stability, Annals of Mathematics Studies, 17.

[7] Topcu G., Aydın K., Construction of Hurwitz Stability Intervals for Matrix Families (Submitted).

[8] Voicu, M. and Pastravanu, O., (2006), Generalized matrix diagonal stability and linear dynamical systems, Linear Algebra and its Applications, 419, 299-310.

[9] Wilkinson, J. H., 1965. The Algebraic Eigenvalue Problem. Clarendom Press.

Department of Mathematics, Faculty of Sciences, Selcuk University, Konya, 42003, Turkiye

Email(s): guner.ozturk@selcuk.edu.tr

Robustness of Randomization Test of Repeated Measures Design in the Presence of Outlier

Ajali John Obasi[1], Abimibola Victoria Oladugba[1] and Ifeoma Ogochukwu Ude[1*]

ABSTRACT. Randomization test (R-test) is often advocated as an alternative data analysis method when assumptions of most commonly used statistical techniques (such as analysis of variance, regression analysis, t-test, and analysis of covariance etc.) are violated. In this work, the robustness in terms of empirical type-I-error and power (sensitivity) of R-test was evaluated and compared with that of F-test in the analysis of a single factor repeated measures design; when the data are normal, and non-normal with or without outliers. The Monte Carlo approach was used in the simulation study. The results showed that when the data are normal, the R-test was approximately as sensitive and robust as the F-test, while it was more sensitive than the F-test when data had skewed distributions. The R-test was more sensitive and robust than the F-test in the presence of an outlier.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 62-06, 62K25

KEYWORDS: Randomization test, Outlier, Type I error, Robustness, Power and Monte Carlo.

1. INTRODUCTION

Repeated measure designs (RMDs) are one of the most frequently used classes of experimental designs in many of applied fields, e.g., medical, pharmaceutical, agricultural, zoological and biological sciences ([1] and [2]). An advantage of RMDs is that they get rid of the effect of individual differences from the analysis since the same subjects are used in all treatment conditions ([3], [4]). The single factor repeated measure design (RMD) according to [5] is an extension to the dependent t-test (paired t-test), and is traditionally analyzed using the repeated measures analysis of variance (RM ANOVA) based on the F-test statistic with parametric test assumptions of normality, independence and equality of variance. The inability of the F-test to produce a reliable result in the face of failure of data to meet parametric test assumptions is its major disadvantage. In order to overcome this problem, the randomization test was used this work to analyzed single factor RMD.

The randomization test(R-test), also called permutation test, is a way of hypotheses testing about treatment effects in an experiment which is based on rearrangement or randomization of the data without referenced to the distribution of the datasets ([6]). In randomization test, [6] elucidated how a reference distribution which is generated from the data is used in place of a theoretical or standard distribution as a basis for comparison with the actual value of a test statistic from an analysis. [7] discussed methods to analyze RMD based on general approximate permutation test which involved permutation of reduced and modified residuals which removes correlation between the residuals and was found to be superior to the convectional RM ANOVA in terms of maintaining the pre-assigned type-1-error. The ability of a test to make the right decision either by rejecting or accepting of a hypothesis also referred to as power of a test, can be affected by the presence of outliers in the datasets. Outliers are often inevitably seen in datasets of repeated measurements, even when data come from reputable sources and the data collection is carefully executed ([8]).

An outlier is an individual value that is substantially different either larger or smaller from the values obtained for other individuals in a data set ([9]). Outliers often lie far away from other values in a random sample from a population. They may be an indicator of data errors or rare events, and should be investigated carefully to understand why they appear in a sample as highlighted by [10]. The existence of outliers has been recognized and noted for centuries and led to concerns about the disproportionate influence of outliers on statistical analyses like in inflation of type-I-error rates and power reduction in parametric t- and F-tests ([11], [12]). [13] hinted that the adverse effects of outliers which manifests in distortions of statistical significance tests can lead to erroneous conclusions if indications of outliers are not carefully examined and robust test used in analysis of such data. As a remedy to the effects of outliers, several approaches can be used to diminish or lower the impact of outliers such as log transformation and nonparametric statistical ranking ([10]), and Winsorizing ([14]). These approaches highlighted can either have poor power properties as in transformation, loss of information as in use of rank and negative bias in estimation of parameters in the case of winsorizing ([15]), therefore it is necessary to employ the randomization test (which only depends on the use of a permutation technique in executing significance testing) in order to reduce the impact of the outlying observations thus can be said to be a robust test in relation to the effect of outliers.

In this work, the robustness in terms of empirical type-I-error (p-value) and power of R-test was evaluated and compared with that of F-test in the analysis of a single factor repeated measures design; when the data are normal and non-normal with or without outliers.

2. SIMULATION PROCEDURES

Data were simulated from five theoretic distributions. The normal distribution was used to test condition under which normality assumption holds. The skewed distributions used include exponential, lognormal, Chi-square, and Weibull distributions. These represent condition under which the distribution assumption (normality) does not hold. The dataset without outlier is referred to as complete case.

Datasets for comparison were generated via Monte Carlo simulation when the normal and non-normal assumptions are met or not for without and with outlier for various sample sizes of 5, 7 and 9. In the simulation, the experiment was repeated 10,000 times for each distribution under the condition when the dataset is complete, and in the presence of an outlier. The outliers were introduced using Tukey (1977) box plot method of detecting outlier whereby an extreme value beyond the outer fences is a probable. The procedures: the IQR (Inter Quartile Range) is the distance between the lower (Q1) and upper (Q3) quartiles, inner fences are located at a distance 1.5*IQR below Q1 and above Q3 [Q1-1.5*IQR, Q3+1.5*IQR] and outer fences are located at a distance 3*IQR below Q1 and above Q3 [Q1-3*IQR, Q3+3*IQR] then a value between the inner and outer fences is a possible outlier. An extreme value beyond the outer fences is a probable outlier. There is no statistical basis for the reason that Tukey uses 1.5 and 3 regarding the IQR to make inner and outer fences. The Tukey's method as implemented using Sigma Magic software detected four outliers injected into the data in this work. The rate of empirical type-1-error and power for the two tests (R-test and F-test) were computed for the two cases (without and with outlier).

3. THE P-VALUE (EMPIRICAL TYPE-I-ERROR) AND POWER FOR COMPARISON

In statistical hypothesis testing, the p-value or probability value is the probability of obtaining test results at least as extreme as the results actually observed during the test, assuming that the null hypothesis is correct. The percentage of significant test out of 10,000 iterations for a given case was considered as either the rejection or the acceptance rate for that case. The comparison procedures were considered in two cases. For the first case where the null hypothesis is true, the rejection rate of the null hypothesis was regarded as the empirical type-I-error (p-value) for each test. The test that had the closest p-value (empirical type-I-error rate) to the nominal $\alpha = 0.05$ was considered as the more robust of the two. The second case is where the alternative hypothesis true, the rejection rate was considered as the power for each test. The test that had larger power was taken to be more sensitive than the other.

4. RESULTS AND DISCUSSION

Using the procedures described in Section 2, the optimal p-value and the highest power for the normal and skewed distributions for the R- and F-tests were obtained and tabulated in Table 1 for complete case and in the presence of an outlier. The values for the case with outlier are presented in the bracket on the table. From the results without outlier, both tests were roughly sensitive and robust under normal assumption since F-test have a better value ($p = 0.05$) while F-test and R-test were very high at 0.9914 and 0.9805 respectively for the power. For exponential distribution, ($p = 0.0542$) is the optimal p-value for F-test and for

R-test ($p = 0.0581$), while the highest power for F-test and R-test were 0.9028 and 0.9469 respectively which shows that R-test was more powerful than F-test, and more robust too for exponential data. For lognormal distribution, the F-test and R-test respectively had optimal p-values for both tests as 0.0413 and 0.0490, while both tests exhibited power of 0.8689 and 0.8629 respectively for F-test and R-test. For the Chi-square distribution, F-test had optimal p-value of 0.0479 and 0.0524 for R-test. Also, the highest power of the F-test and R-test respectively were 0.6628 and 0.8086. For the Weibull distribution, p-value = 0.0525 for the R-test was more robust and power = 0.8334 was more powerful.

In the presence of an outlier presented in bracket in Table 1, we have under assumption of normal ($p = 0.0824$, power = 0.5869), exponential ($p = 0.0755$, power = 0.6963), lognormal ($p = 0.0792$, power = 0.6876), Chi-square ($p = 0.0843$, power = 0.6628), and Weibull ($p = 0.0818$, power = 0.5994). These results indicated that p-values for F-test under the exponential and lognormal distributions had a decreasing trend but an increasing trend for Chi-square and Weibull distribution. Also, the power for F-test exhibited slight decreasing trend for normal and Weibull, while it increased for exponential Chi-square and lognormal. On the other hand, the R-test was more powerful and robust in all distribution: normal ($p = 0.0588$, power = 0.8209), exponential ($p = 0.0566$, power = 0.8202), lognormal ($p = 0.0512$, power = 0.8094), Chi-square ($p = 0.0516$, power = 0.8180), and Weibull ($p = 0.0536$, power = 0.7933). The value in bracket is p-value and power for with outlier.

Distribution	**F-test**		**R-test**	
	p-value	power	p-value	power
Normal	0.0500(0.0824)	0.9914(0.5869)	0.0429(0.0588)	0.9805(0.8209)
Exponential	0.0542(0.0755)	0.9028(0.6963)	0.0581(0.0566)	0.9469(0.8202)
Lognormal	0.0413(0.0792)	0.8689(0.6876)	0.0490(0.0512)	0.8629(0.8094)
Chi-Square	0.0479(0.0843)	0.6729(0.6628)	0.0524(0.0516)	0.8086(0.8180)
Weibull	0.0525(0.0818)	0.6804(0.5994)	0.0740(0.0536)	0.8334(0.7933)

Table 1. The p-values and power of F-test and R-test without and with outlier

5. CONCLUSION

This work assessed the robustness of R- tests in the presence of outlier for repeated measures design. The empirical type-I-error (p-value) and power of the R-test was evaluated and compared with that of the F-test under normal and skewed distributions. The results showed that when the normality condition was met in a complete case, the randomization test was comparably as robust and sensitive as the F-test. When data had skewed distributions (exponential, lognormal, Chi-square and Weibull), the F-test was less robust and sensitive. Also, in the presence of an outlier, the randomization test was more robust and sensitive than the F-test. In a nutshell, the randomization test was approximately as sensitive as the F-test in RMD when dataset follows normal condition but more sensitive when the dataset is skewed (exponential, Chi-square, lognormal and Weibull). Based on the results obtained in this study we recommend the randomization test to serve as a better option in the presences of outlier.

References

[1] Koc A., (2006). Analysis of repeated milk somatic cell count of Holstein-Friesian cows raised in Mediterranean climatic condition, Journal of Biological Sciences 6 1093-1097.

[2] Gravetter F.J. and Wallnau L.B (2007), Statistics for the behavioral science, Canada: Vicki Knight.

[3] Reed III. J. (2003) Analysis of variance (ANOVA) models in emergency medicine, The Journal of Emergency and Intensive Care Medicine, 7 (2), 21-34.

[4] Howitt D., and Cramer D. (2011) Introduction to Research Methods in Psychology (3rd ed.). Essex: Pearson Education Limited.

[5] Vallejo G., Paula M. F., Livacic-Rojas P. E. and Tuero-Herrero E. (2011) Selecting the best unbalanced repeated measures model, Behavioral Research Methods, 43, 18-36.

[6] Edgington E. S. (1995). Randomization tests. New York: M.

[7] Kherad-Pajouh S. and Renaud O. (2014). A general permutation approach for analyzing repeated measures ANOVA and mixed-model designs ,Computational Statistics and Data Analysis, 21 (5), 42-59.

[8] Dragset I.G. (2009), Analysis of Longitudinal Data with Missing Values.: Methods and Applications in Medical Statistics (Master's thesis, Institutt for matematiske fag).
[9] Liao H., Li,Y., and Brooks G. (2016) Outlier impact and accommodation methods: Multiple comparisons of type I error rates Journal of Modern Applied Statistical Methods, 15 (1), 452-471.
[10] Zimmerman D. W. and Zumbo B. D. (1990). Effect of outliers on the relative power of parametric and nonparametric statistical tests. Perceptual and Motor Skills, 71, 339 -349.
[11] Wilcox R. (1998). How many discoveries have been lost by ignoring modern statistical methods? American Psychologist, 53 (3), 300-314.
[12] Zimmerman D. W. (1994). A note on the influence of outliers on parametric and nonparametric tests. The Journal of General Psychology, 121 (4), 391-401.
[13] Warner R. M. (2012). Applied statistics: From bivariate through multivariate techniques, Thousand Oaks, CA: Sage.
[14] Bakker M., and Wicherts, J. M. (2014), Outlier removal, sum scores, and the inflation of the Type I error rate in independent samples t tests: the power of alternatives and recommendations. Psychological methods, 19(3), 409.
[15] Gleason J. (2013), Comparative power of the ANOVA, randomization ANOVA, And Kruskal-Wallis test (Doctoral Dissertation). Available from Wayne State University Digital Dissertations database.
[16] Tukey J. W. (1977). Exploratory data analysis. Philippines: Addison-Wesely Publishing Company.

Department of Statistics, University of Nigeria, Nsukka, Nigeria [1*] (corresponding author)

Email(s): ude.ifeoma@unn.edu.ng [1*] (corresponding author)

Approximation Properties of Some Nonpositive Kantorovich Type Operators

Bianca Ioana Vasian

Abstract. In this paper we will construct a generalization of Bernstein operators using Kantorovich's method. In this sense we will use a general derivative operator denoted by D^l and its corresponding anti-derivative operator I^l, having the property $D^l \circ I^l = I^l \circ D^l = Id$. We will prove that the convergence on all continuous functions on $[0, 1]$ holds even though the operators constructed this way are not positive.

2020 Mathematics Subject Classifications: 41A10, 41A35.

Keywords: Bernstein operators, Kantorovich operators, approximation, nonpositive linear operators

1. Introduction and Preliminaries

In order to prove Weierstrass's approximation theorem [16], S. N. Bernstein proposed, in paper [3], the following positive and linear operators: let $f \in C[0,1]$:

$$B_n(f,x) = \sum_{k=0}^{n} p_{n,k}(x) f\left(\frac{k}{n}\right),\ x \in [0,1], \tag{1}$$

where $p_{n.k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$, for $0 \leq k \leq n$, and $p_{n.k}(x) = 0$ for $k > n$. It is well-known that these operators can be used to uniformly approximate all continuous functions on $[0, 1]$, and also for the simultaneous approximation: for $f \in C^k[0,1]$ we have $B_n f^{(k)} \to f^{(k)}$, uniformly as $n \to \infty$, see [5, 15].

After the Bernstein operators were introduced, there arose a lot of generalizations of them (for example see [4, 5, 8, 10, 13]). One of these generalizations was given by D.D. Stancu in [14], which introduced a modification of the Bernstein operators (1) depending on two parameters $0 \leq \alpha \leq \beta$: let $f \in C[0,1]$:

$$B_n^{\alpha,\beta}(f,x) = \sum_{k=0}^{n} p_{n,k}(x) f\left(\frac{k+\alpha}{n+\beta}\right),\ x \in [0,1], \tag{2}$$

and proved that $\left\| B_n^{\alpha,\beta} f - f \right\| \to 0$ as $n \to \infty$. Further it has been shown that for $f \in C^k[0,1]$, the simultaneous approximation $B_n^{\alpha,\beta} f^{(k)} \to f^{(k)}$ holds uniformly (see [1]).

Another notable generalization was given by Kantorovich in paper [9], which introduced a modification of Bernstein operators using the following method: let L be an operator, $D = \frac{d}{dx}$ be a differential operator and $If(x) = \int_0^x f(t)\,dt$, the corresponding antiderivative in the sense: $D \circ I = I \circ D = Id$, where Id is the identity operator. In this case, the Kantorovich modification of L is given by

$$K = D \circ L \circ I. \tag{3}$$

Now, if we take $L = B_{n+1}$ in (3), we obtain the Bernstein-Kantorovich operators K_n, studied in detail in [9], defined as $K = D \circ B_{n+1} \circ I$, having the following expression: let $f \in L_1[0,1]$,

$$K_n(f,x) = (n+1) \sum_{k=0}^{n} p_{n,k}(x) \int_{\frac{k}{n+1}}^{\frac{k+1}{n+1}} f(t)\,dt,\ x \in [0,1]. \tag{4}$$

In the case of Bernstein-Kantorovich we have that these operators uniformly approximate all integrable functions on $[0,1]$. After these operators were introduced, they were extensively studied (see, for example [2, 6]).

Another modification using Kantorovich's method was introduced by Gonska H. et. al. in paper [7] where they analyzed the operators obtained by taking the derivative operators $D^k = \frac{d}{dx^k}$, and the antiderivative operators as $I^k f(x) = \int_0^x \frac{(x-t)^{k-1}}{(k-1)!} f(t)\,dt$, and considering the operators $\tilde{K}_n = D^k \circ B_n \circ I^k$ called k-th order Kantorovich operators. They proved the uniform convergence $\tilde{K}_n f \to f$, for $f \in L_1[0,1]$.

In paper [11], Pltnea R. introduced a modification of Bernstein operators using Kantorovich's method by taking the differential operator $D^c f = f' + cf$. He proved that the operators constructed as $K_n^* = D^c \circ B_{n+1} \circ I^c$, where I^c is a corresponding antiderivative operator, can be used to approximate functions on $C[0,1]$ and the operators are not positive. Having in mind these operators we will introduce a new class of Kantorovich type operators which are constructed using Kantorovich's method, but taking a more general differential operator: let $l \in \{1,2,\dots\}$,

$$D^l f = f^{(l)} + a_{l-1} f^{(l-1)} + \cdots + a_1 f' + bf, \tag{5}$$

$a_1, a_2, \dots, a_{l-1}, b \in \mathbb{R}$, and a corresponding antiderivative operator I^l with respect to $D^l \circ I^l = I^l \circ D^l = Id$. This condition leads to:

$$\left(D^l \circ I^l\right)(f) = f,$$

therefore we have:

$$\left(I^l f\right)^{(l)} + a_{l-1}\left(I^l f\right)^{(l-1)} + \cdots + a_1\left(I^l f\right)' + b\left(I^l f\right) = f, \tag{6}$$

which is a linear differential equation of order l with constant coefficients for which we know that $I^l f$ exists but it is not unique and it is of class $C^l[0,1]$. Since there is an infinity of such antiderivatives, one can obtain a unique one by imposing a Cauchy problem for differential equations such as, in a certain point the anti-derivative $I^l f$ and its derivatives up to order $l-1$ should have particular values.

The Kantorovich type operators that will be studied in this paper are of the form

$$K^l = D^l \circ L \circ I^l, \tag{7}$$

where L is an operator.

Further we will consider $L = B_{n+l}$ in (7) and we will prove that these linear operators can approximate continuous functions on $[0,1]$ even though they are not positive operators.

Along the paper we will use the following notions. Denote by e_s the monomials $e_s(t) = t^s$.

Let $\Delta_h f(x) = f(x+h) - f(x)$ be the first finite difference of f with step h. The l-th iterate of Δ is denoted by Δ^l and is defined as follows

$$\Delta_h^l f(x) = \Delta_h\left[\Delta_h^{l-1} f(x)\right]. \tag{8}$$

We can write the finite difference of order l as follows:

$$\Delta_h^l f(x) = \sum_{i=0}^{l} (-1)^{l-i} \binom{l}{i} f(x+ih). \tag{9}$$

Proposition 1.1. *If f is a polynomial of degree $l-1$ then $\Delta_h^l f(x) = 0$.*

Let I be an interval and $x_0, x_1, \dots, x_n \in I$ $n+1$ distinct points of I. Let f be a function defined on I. The divided differences of f on $x_0, x_1, \dots, x_n$ are given by

$$\begin{aligned} f[x_k] &:= f(x_k),\ k = \overline{0,n}, \\ f[x_k, x_{k+1}, \dots, x_{k+p}] &:= \frac{f[x_{k+1}, \dots, x_{k+p}] - f[x_k, x_{k+1}, \dots, x_{k+p-1}]}{x_{k+p} - x_k}, \end{aligned} \tag{10}$$

for $k = \overline{0, n-p},\ j = \overline{0,n}.$

Divided differences have the following properties:

Proposition 1.2. *If f is a polynomial of degree $< n$, then*

$$f[x_0, x_1, \dots, x_n] = 0. \tag{11}$$

Proposition 1.3. *(Mean value theorem for divided differences) If f is n times differentiable, then*

$$f\left[x_0, x_1, \ldots, x_n\right] = \frac{f^{(n)}(\xi)}{n!}, \tag{12}$$

for $\xi \in \left(\min_{k\in\{0,1,\ldots n\}} x_k, \max_{k\in\{0,1,\ldots n\}} x_k\right)$.

Proposition 1.4. *The following relation between finite differences and divided differences holds:*

$$f\left[x, x+h, \ldots, x+lh\right] = \frac{1}{l!h^l}\Delta_h^l f(x). \tag{13}$$

It is well known that if we consider the l-th derivative of Bernstein operators (1), it can be written in terms of finite differences of order l as follows, see [5]:

$$B_n^{(l)}(f,x) = \frac{n!}{(n-l)!}\sum_{j=0}^{n-l}\Delta_{\frac{1}{n}}^l f\left(\frac{j}{n}\right)p_{n-l,j}(x),\ x\in[0,1]. \tag{14}$$

With all the considerations from above, let us define our Kantorovich type operators:

$$K_n^l(f,x) = \left(D^l \circ B_{n+l}\circ I^l\right)(f,x),\ x\in[0,1] \tag{15}$$

which can be written as

$$K_n^l(f,x) = D^l\left(B_{n+l}I^l f\right)(x). \tag{16}$$

For simplicity, let us denote $I^l f := F$

$$\begin{aligned} &K_n^l(f,x) = \\ &= \left[B_{n+l}(F,x)\right]^{(l)} + a_{l-1}\left[B_{n+l}(F,x)\right]^{(l-1)} + \cdots + a_1\left[B_{n+l}(F,x)\right]' + bB_{n+l}(F,x)\\ &= \frac{(n+l)!}{n!}\sum_{j=0}^{n}\Delta_{\frac{1}{n+l}}^l F\left(\frac{j}{n+l}\right)p_{n,j}(x) + \\ &+a_{l-1}\frac{(n+l)!}{(n+1)!}\sum_{j=0}^{n+1}\Delta_{\frac{1}{n+l}}^{l-1}F\left(\frac{j}{n+l}\right)p_{n+1,j}(x) + \\ &+\cdots + a_1\frac{(n+l)!}{(n+l-1)!}\sum_{j=0}^{n+l-1}\Delta_{\frac{1}{n+l}}F\left(\frac{j}{n+l}\right)p_{n+l-1,j}(x) + bB_{n+l}(F,x). \end{aligned} \tag{17}$$

Remark 1. *The operators* (17) *are linear operators.*

Our aim is to prove that

$$\left\|K_n^l f - f\right\| \to 0,\ as\ n\to\infty$$

for all continuous functions on $[0,1]$.

2. Main results

In order to prove our main approximation result we need the following lemma:

Lemma 2.1. *Let F be a C^k function on a compact interval I, then the following convergence holds:*

$$(n+l)^k\Delta_{\frac{1}{n+l}}^k F(x) \to F^{(k)}(x),\ uniformly\ as\ n\to\infty,\ x\in I. \tag{18}$$

Proof. In order to prove the result we will use the following representation of the finite difference:

$$(n+l)^k\Delta_{\frac{1}{n+l}}^k F(x) = (n+l)^k\sum_{i=0}^{k}(-1)^{i+k}\binom{k}{i}F\left(x+\frac{i}{n+l}\right).$$

Now, using the Taylor expansion of order k for the function F in the point x we obtain the following:

$$(19)\qquad (n+l)^k\,\Delta^k_{\frac{1}{n+l}}F(x)=(n+l)^k\sum_{i=0}^{k}(-1)^{i+k}\binom{k}{i}\times$$

$$\times\left\{\sum_{s=0}^{k}\frac{1}{s!}F^{(s)}(x)\left(\frac{i}{n+l}\right)^s+\int\limits_{x}^{x+\frac{i}{n+l}}\frac{\left(x+\frac{i}{n+l}-u\right)^{k-1}}{(k-1)!}\left[F^{(k)}(u)-F^{(k)}(x)\right]du\right\}$$

$$=(n+l)^k\sum_{s=0}^{k}\frac{1}{s!}F^{(s)}(x)\sum_{i=0}^{k}(-1)^{i+k}\binom{k}{i}\left(\frac{i}{n+l}\right)^s+R$$

$$=:T+R.$$

We will treat separately T and R.

$$(20)\qquad T=(n+l)^k\sum_{s=0}^{k}\frac{1}{s!}F^{(s)}(x)\sum_{i=0}^{k}(-1)^{i+k}\binom{k}{i}\left(\frac{i}{n+l}\right)^s$$

$$=(n+l)^k\sum_{s=0}^{k-1}\frac{1}{s!}F^{(s)}(x)\,\Delta^k_{\frac{1}{n+l}}e_s(0)+(n+l)^k\,\frac{1}{k!}F^{(k)}(x)\,\Delta^k_{\frac{1}{n+l}}e_k(0).$$

Using the Proposition 1.1 we have that $\Delta^k_{\frac{1}{n+l}}e_s(0)=0$ for $s\le k-1$.

Now, for the second term we will use the relation between finite differences and divided differences:

$$(21)\qquad \begin{aligned}\Delta^k_{\frac{1}{n+l}}e_k(0)&=k!\frac{1}{(n+l)^k}e_k\left[0,\frac{1}{n+l},\ldots,\frac{k}{n+l}\right]\\&=k!\frac{1}{(n+l)^k}\frac{e_k^{(k)}(\xi)}{k!}=k!\frac{1}{(n+l)^k}\frac{k!}{k!}\\&=k!\frac{1}{(n+l)^k}.\end{aligned}$$

We have obtained that

$$(22)\qquad T=F^{(k)}(x).$$

Our next purpose is to show that $R\to 0$ uniformly. Take

$$(23)\qquad |R|\le$$

$$\le(n+l)^k\sum_{i=0}^{k}\binom{k}{i}\left|\int\limits_{x}^{x+\frac{i}{n+l}}\frac{\left(x+\frac{i}{n+l}-u\right)^{k-1}}{(k-1)!}\left[F^{(k)}(u)-F^{(k)}(x)\right]du\right|$$

$$\le(n+l)^k\sum_{i=0}^{k}\binom{k}{i}\omega\left(F^{(k)},\delta\right)\int\limits_{x}^{x+\frac{i}{n+l}}\frac{\left(x+\frac{i}{n+l}-u\right)^{k-1}}{(k-1)!}\left(1+\frac{(u-x)^2}{\delta^2}\right)du$$

$$=(n+l)^k\sum_{i=0}^{k}\binom{k}{i}\omega\left(F^{(k)},\delta\right)\left[\left(\frac{i}{n+l}\right)^k\frac{1}{k!}+\right.$$

$$\left.+\frac{1}{\delta^2}\int\limits_{x}^{x+\frac{i}{n+l}}\frac{\left(x+\frac{i}{n+l}-u\right)^{k-1}}{(k-1)!}(u-x)^2\,du\right]$$

Treat the integral part separately. For simplicity, denote $\rho := \frac{i}{n+l}$. We have:

$$J = \int_x^{x+\rho} \frac{(x+\rho-u)^{k-1}}{(k-1)!} (u-x)^2 \, du \tag{24}$$

$$= \int_x^{x+\rho} \frac{(x+\rho-u)^{k-1}}{(k-1)!} \left[(x+\rho-u)^2 - 2\rho(x+\rho-u) + \rho^2\right] du$$

$$= \frac{\rho^{k+2}}{(k+2)(k-1)!} - 2\frac{\rho^{k+2}}{(k+1)(k-1)!} + \frac{\rho^{k+2}}{k!}$$

$$\leq C\rho^{k+2},$$

where C is a constant. We get

$$|R| \leq \omega\left(F^{(k)}, \delta\right) \sum_{i=0}^{k} \binom{k}{i} \left[\frac{i^k}{k!} + \frac{i^{k+2}}{\delta^2} C \frac{1}{(n+l)^2}\right]. \tag{25}$$

Taking $\delta = \frac{1}{n+l}$ we obtained that $R \to 0$ uniformly, which proves our result. □

Now, we can state and prove the approximation result.

Theorem 2.2. *Let $f \in C[0,1]$. The following convergence holds:*

$$K_n^l f \to f, \ \textit{uniformly as } n \to \infty. \tag{26}$$

Proof. Let us take into account the expression (17) for the operators $K_n^l f$. As we have proved in Lemma 2.1: $\forall \varepsilon > 0$ there is $n_\varepsilon \in \mathbb{N}$ such that for all $n \geq n_\varepsilon$ we have

$$|(n+l)^k \Delta^k_{\frac{1}{n+l}} F(x) - F^{(k)}(x)| < \varepsilon, \tag{27}$$

uniformly with respect to x, for each $k \in \{0, 1, \dots\}$.

Now, we shall consider the difference:

$$K_n^l f(x) = \frac{(n+l)!}{n!} \frac{1}{(n+l)^l} \sum_{j=0}^{n} (n+l)^l \Delta^l_{\frac{1}{n+l}} F\left(\frac{j}{n+l}\right) p_{n,j}(x) + \tag{28}$$

$$+ a_{l-1} \frac{(n+l)!}{(n+1)!} \frac{1}{(n+l)^{l-1}} \sum_{j=0}^{n+1} (n+l)^{l-1} \Delta^{l-1}_{\frac{1}{n+l}} F\left(\frac{j}{n+l}\right) p_{n+1,j}(x)$$

$$+ \dots + a_1 \frac{(n+l)!}{(n+l-1)!} \sum_{j=0}^{n+l-1} \Delta_{\frac{1}{n+l}} F\left(\frac{j}{n+l}\right) p_{n+l-1,j}(x) + bB_{n+l}(F, x)$$

$$= A_n^0 \sum_{j=0}^{n} (n+l)^l \Delta^l_{\frac{1}{n+l}} F\left(\frac{j}{n+l}\right) p_{n,j}(x) + \tag{29}$$

$$+ a_{l-1} A_n^1 \sum_{j=0}^{n+1} (n+l)^{l-1} \Delta^{l-1}_{\frac{1}{n+l}} F\left(\frac{j}{n+l}\right) p_{n+1,j}(x)$$

$$+ \dots + a_1 A_n^{l-1} \sum_{j=0}^{n+l-1} \Delta_{\frac{1}{n+l}} F\left(\frac{j}{n+l}\right) p_{n+l-1,j}(x) + bB_{n+l}(F, x),$$

where $A_n^k := \frac{(n+l)!}{(n+k)!} \frac{1}{(n+l)^{l-k}}$, $k \in \{0, \dots, l-1\}$.

There exists n_ε^1 such that $\left|A_n^k - 1\right| < \varepsilon$ for $n \geq n_\varepsilon^1$, for $0 \leq k \leq l-1$, and we obtain the following:

$$K_n^l f(x) \to \sum_{j=0}^{n} F^{(l)}\left(\frac{j}{n+l}\right) p_{n,j}(x) + a_{l-1} \sum_{j=0}^{n+1} F^{(l-1)}\left(\frac{j}{n+l}\right) p_{n+1,j}(x) +$$

$$+ \dots + a_1 \sum_{j=0}^{n+l-1} F'\left(\frac{j}{n+l}\right) p_{n+l-1,j}(x) + bB_{n+l}(F, x).$$

Now, we notice that the sums appearing above can be expressed in terms of Bernstein-Stancu operators as follows:

$$\sum_{j=0}^{n+k} F^{(l-k)}\left(\frac{j}{n+l}\right) p_{n+k,j}(x) = B_{n+k}^{0,l-k}\left(F^{(l-k)}, x\right), \ k \le l-1$$

for which we know that $B_{n+k}^{0,l-k}\left(F^{(l-k)}, x\right) \to F^{(l-k)}(x)$ as $n \to \infty$, uniformly with respect to x, and the term $B_{n+l}(F,x) \to F(x)$ since it is the classical Bernstein operator.

$$\begin{aligned} K_n^l f(x) &\to F^{(l)}(x) + a_{l-1}F^{(l-1)}(x) + \cdots + a_1 F'(x) + bF(x) = \\ &= D^l \circ I^l (f(x)) = f(x), \end{aligned}$$

and the convergence above is uniform with respect to x. □

Remark 2. *The operators $K_n^l f(x)$ are not positive as the following example shows:*

Example 2.3. *Let $n=1$ and the differential operator be $D^1 f = f' - f$ which has a fixed corresponding antiderivative $I^1 f(x) = e^x \int_0^x e^{-t} f(t)\,dt$, which will be denoted by $I^1 f(x) := F(x)$ and chosen such that $F(0)=0$. We consider the function*

$$f(t) = e^{-5t}, \ t \in [0,1].$$

Now, we compute K_1^1:

$$K_1^1 f(x) = \left(D^1 \circ B_2 \circ I^1\right)(f(x)) = [B_2(F,x)]' - B_2(F,x) = \tag{30}$$

$$= \left[F(0)(1-x)^2 + 2F\left(\frac{1}{2}\right)x(1-x) + F(1)x^2\right]' - \tag{31}$$

$$\begin{aligned} &- \left[F(0)(1-x)^2 + 2F\left(\frac{1}{2}\right)x(1-x) + F(1)x^2\right] \\ &= 2F\left(\frac{1}{2}\right)\left(1-3x+x^2\right) + F(1)\left(2x-x^2\right) \\ &= 2\left(1-3x+x^2\right)e^{\frac{1}{2}}\int_0^{\frac{1}{2}} e^{-6t}dt + \left(2x-x^2\right)e\int_0^1 e^{-6t}dt \\ &= -\frac{1}{3}\left(1-3x+x^2\right)e^{\frac{1}{2}}\left(e^{-3}-1\right) - \frac{e}{6}(2x-x^2)\left(e^{-6}-1\right). \end{aligned}$$

We take $x=1$ and obtain

$$K_1^1 f(1) = \frac{1}{3}e^{\frac{1}{2}}\left(e^{-3}-1\right) - \frac{e}{6}\left(e^{-6}-1\right) = -7.0288 \times 10^{-2},$$

which holds our remark.

3. Conclusion

The operators considered in this paper constitutes an example of approximation operators that can be used to approximate continuous functions on $[0,1]$ even though they are nonpositive operators.

References

[1] Agratini O., On simultaneous approximation by Stancu-Bernstein operators in Approximation and Optimization, In : Proceedings of ICAOR Romania, Cluj-Napoca, 1996.

[2] Brbosu D. Kantorovich-Stancu type operators. J. Inequal Pure Appl. Math. 2004

[3] Bernstein S. N., Démonstration du théorém de Weierstrass fondeé sur la calcul desprobabilitiés. Commun Soc Math Charkow Sér 2(13), 1912.

[4] Cal J., Valle A. M., A generalization of Bernstein-Kantorovich operators, J. Math. Anal. Appl., 252, 2000.

[5] DeVore R. A., Lorentz G. G., Constructive Approximation, Springer-Verlag Berlin Heidelberg, 1993.

[6] Dhamija M., Deo N., Better Approximation Results by Bernstein-Kantorovich Operators. Lobachevskii J. Math. 2017.

[7] Gonska H., Heilmann M., Rasa I., Kantorovich operators of order k. Numer Funct Anal Optim 2011.

[8] Kajla A., Acar T., Modified α-Bernstein operators with better approximation properties, Ann. Funct. Anal. 10 (4) 570 - 582, 2019.

[9] Kantorovich, L.V. Sur certain developpements suivant les polynomes de la forme de S. Bernstein, I, II. CR Acad. URSS, 1930.

[10] Lorentz, G.G., Bernstein Polynomials, 2nd edn. Chelsea, New York, 1986.

[11] Pltnea R., A note on generalized Bernstein-Kantorovich operators, Bulletin of the Transilvania University of Braşov Vol 6(55), No. 2 - 2013.

[12] Mohiuddine S.A., Acar T., Alotaibi A., Construction of a new family of Bernstein-Kantorovich operators, Math Meth Appl Sci. 2017;111.

[13] Nursel ., Radu V. A., Approximation by generalized Bernstein-Stancu operators, Turkish Journal of Mathematics: Vol. 43, 2019.

[14] Stancu D. D., Asupra unei generalizri a polinoamelor lui Bernstein (in romanian), Stud. Univ. Babes-Bolyai Ser. Math. Phys, 1969.

[15] Totik, V., Approximation by Bernstein Polynomials. American Journal of Mathematics, 116(4), 9951018, 1994.

[16] Weierstrass, K.G.: U die analytische Darstellbarkeit sogenannter licher Funktionen einer reellen Veranderlichen, Sitzungsber. Akad. Berlin 2, 633-639, 1885.

Department of Mathematics and Computer Science, Transilvania University of Braov, Bdul Eroilor 29, 500036 Braov, Romania.

Email(s): bianca.vasian@unitbv.ro

The Inverse Limits of Dicompact Bi-Hausdorff Spaces

Filiz Yıldız

ABSTRACT. Inverse systems and inverse limits are introduced and studied for some categories of dicompact texture spaces, that is the ditopological spaces which are compact in the textural sense as a natural counterpart of the classical compactness. According to that, in this study we especially restrict ourselves to the topological subcategory $\mathbf{ifPDicomp_{T2}}$ consisting of dicompact, bi-Hausdorff ditopological spaces that have plain texturing which is proper, but still quite extensive subclass of textures. In the context of this category, some various properties of inverse limits and the related theorems are discussed. In particular, by taking into consideration the fact that $\mathbf{ifPDicomp_{T2}}$ is the full subcategory of the category **ifPDitop**, whose objects are ditopological plain spaces, we acquired a few topological results about the theory of inverse systems-limits of dicompact texture spaces satisfying a specific condition.

2020 MATHEMATICS SUBJECT CLASSIFICATIONS: 46M40, 54B10, 54B30, 54D30, 18B35, 18A30, 54A05

KEYWORDS: Inverse limit, ditopological plain texture, equalizer, dicompact space, full subcategory, joint topology

1. INTRODUCTION

The theory of inverse systems and inverse limits was handled in (*cf.* [5]) first-time, in the categories of some special textures. Within the framework of this idea, the main purpose of our study is to present some crucial properties of the inverse limits in the context of dicompact bi-Hausdorff texture spaces.

Accordingly, after giving the background information in Section 1 required for this study, the category of dicompact bi-Hausdorff ditopological plain texture spaces and the bicontinuous point functions satisfying a specific condition is introduced and discussed in Section 2 as the full subcategory of the category consisting of ditopological plain texture spaces. Following that the notion of *inverse limit* for the inverse systems in the category of dicompact bi-Hausdorff plain spaces and the other concepts related with the inverse limits of inverse systems in that category are presented in Section 3. Especially, the required theorems and results for the rest of paper are mentioned here in a categorical setting for the plain textures via the notion of joint topology for a ditopology.

Finally, Section 4 as the last part of the paper gives a conclusion about the whole of the work.

1.1. **Preliminaries.** The theory of ditopological textures was introduced by L. M. Brown as a point-based setting for the study of complement-free mathematical concepts. Within the framework of this theory, let us remind the following concepts (*cf.* [2, 3, 4]) will be required for this study:

Texture: If S is a set, a *texturing* $\mathcal{S}$ on S is a subset of $P(S)$ which is a point-separating, complete, completely distributive lattice containing S and $\emptyset$, and for which meet coincides with intersection and finite joins with union. The pair $(S, \mathcal{S})$ is then called a *texture*.

For a texture $(S, \mathcal{S})$, most properties are conveniently defined in terms of the *p-sets* and *q-sets*:

$$P_s = \bigcap\{A \in \mathcal{S} \mid s \in A\}, \quad Q_s = \bigvee\{A \in \mathcal{S} \mid s \notin A\}$$

Plain Texture: The texture $(S, \mathcal{S})$ is called *plain* if $\mathcal{S}$ is closed under arbitrary unions, equivalently if $P_s \not\subseteq Q_s$ for all $s \in S$.

Point Function Between Textures: If $(S, \mathcal{S})$, $(T, \mathcal{T})$ are textures and $\varphi : S \to T$ a point function between the base sets of textures satisfying the compatibility condition $P_s \not\subseteq Q_{s'} \implies P_{\varphi(s)} \not\subseteq Q_{\varphi(s')}$, that is φ is called *ω-preserving.* In addition, the following equalities define the *inverse image* with respect to φ for each $B \in \mathcal{T}$.

$$\varphi^{\leftarrow} B = \bigvee \{P_u \mid \varphi(u) \in B\} = \bigcap \{Q_v \mid \varphi(v) \notin B\}$$

Ditopological Space: Since a texturing $\mathcal{S}$ need not be closed under the operation of taking the set-complement, the notion of topology is replaced by that of *dichotomous topology* or *ditopology*, namely a pair (τ, κ) of subsets of $\mathcal{S}$, where the set of *open sets* τ satisfies

(1) $S, \emptyset \in \tau$,
(2) $G_1, G_2 \in \tau \implies G_1 \cap G_2 \in \tau$ and
(3) $G_i \in \tau$, $i \in I \implies \bigvee_i G_i \in \tau$,

and the set of *closed sets* κ satisfies the dual conditions:

(1) $S, \emptyset \in \kappa$,
(2) $K_1, K_2 \in \kappa \implies K_1 \cup K_2 \in \kappa$ and
(3) $K_i \in \kappa$, $i \in I \implies \bigcap K_i \in \kappa$.

A ditopological texture space with respect to a ditopology (τ, κ) on the texture $(S, \mathcal{S})$ is denoted by $(S, \mathcal{S}, \tau, \kappa)$.

An adequate introduction to the theory of ditopological spaces and the motivation for its study may be obtained from (*cf.* [2, 3, 4]).

Now we will show that we may associate the ditopology (τ, κ) on a plain texture $(S, \mathcal{S})$ with a topology $\mathcal{J}_{\tau\kappa}$ on S, by adapting the notion *appropriate joint topology for a ditopology* described in (*cf.* [3]) to plain case:

Definition 1.1. *(cf.* [4]*) Let $(S, \mathcal{S}, \tau, \kappa) \in \mathrm{Ob}\,\mathbf{ifPDitop}$. We define the joint topology on S in terms of its family $\mathcal{J}^c_{\tau\kappa}$ of closed sets by the condition*

$$W \in \mathcal{J}^c_{\tau\kappa} \iff (s \in S,\ G \in \eta(s), K \in \mu(s) \implies G \cap W \not\subseteq K) \implies s \in W.$$

Here $\eta(s) = \{N \in \mathcal{S} \mid \ P_s \subseteq G \subseteq N \not\subseteq Q_s$ for some $G \in \tau\}$ and $\mu(s) = \{M \in \mathcal{S} \mid \ P_s \not\subseteq M \subseteq K \subseteq Q_s$ for some $K \in \kappa\}$.

Note that the proof of the fact "$\mathcal{J}^c_{\tau\kappa}$ satisfies the closed-set axioms" is straightforward and on passing to the complement this verifies that

(i) *$\{G \subseteq S \mid G \in \tau\} \cup \{S \setminus K \subseteq S \mid K \in \kappa\}$ is a subbase, and*
(ii) *$\{G \cap (S \setminus K) \subseteq S \mid G \in \tau, K \in \kappa\}$ a base*

of open sets for the topology $\mathcal{J}_{\tau\kappa}$ on S.

Remark 1. *(1) From now on, in this work we will use the terms jointly closed (open, dense) for the set which is closed (open, dense) with respect to the appropriate joint topology of the ditopology on space.*
(2) For $(S, \mathcal{S}, \tau, \kappa) \in \mathrm{Ob}\,\mathbf{ifPDitop}$, it is easy to see that $\kappa \subseteq \mathcal{J}^c_{\tau\kappa}$ and $\tau \subseteq \mathcal{J}_{\tau\kappa}$. Also, the family $\tau \cup \kappa^c$ is the subbase for the joint topology $\mathcal{J}_{\tau\kappa}$.

Product Ditopology: Let $(S_j, \mathcal{S}_j, \tau_j, \kappa_j)_{j \in J}$ be ditopological spaces and $(S, \mathcal{S})$ the product of the textures $(S_j, \mathcal{S}_j)_{j \in J}$. Then the ditopology (τ, κ) on $(S, \mathcal{S})$ with subbase $\{\Pi_j^{\leftarrow} G = E(j, G) \mid G \in \tau_j, j \in J\}$ and cosubbase $\{\pi_j^{\leftarrow} K = E(j, K) \mid K \in \kappa_j, j \in J\}$ is called the *product ditopology* on $(S, \mathcal{S})$. In this case, the ditopological space $(S, \mathcal{S}, \tau, \kappa)$ is called the *product of the family* $(S_j, \mathcal{S}_j, \tau_j, \kappa_j)_{j \in J}$, and is denoted by $(\prod_{j \in J} S_j, \bigotimes_{j \in J} \mathcal{S}_j, \bigotimes_{j \in J} \tau_j, \bigotimes_{j \in J} \kappa_j)$.

Bicontinuity: An ω-preserving point function between the ditopological texture spaces, is called *bicontinuous* if the inverse image of every open set is open and the inverse image of every closed set is closed.

The Category **ifPTex** *:* Objects are plain textures and morphisms are ω-preserving point functions.

The Category **ifPDitop** *:* Objects are ditopological plain texture spaces and morphisms are ω-preserving, bicontinuous point functions.

It is easy to verify that the category **ifPDitop** is a topological category over **ifPTex**.

Now we can turn our attention to the theory of inverse systems in the context of textures:

Inverse Systems and Inverse Limits in the Category **ifPTex** The foundations of the theory of inverse systems and their limits that is inverse limits, are handled in a textural context and investigated first-time for the plain textures in (*cf.* [5]).

Specifically, the reader may consult (*cf.* [1]) for all the general concepts relating to category theory.

2. The Category of Dicompact Bi-Hausdorff Spaces

The following notions discussed in (*cf.* [2, 4]) will be required in order to construct a new category.

Definition 2.1. *Let (τ, κ) be a ditopology on the texture $(S, \mathcal{S})$. Then $(S, \mathcal{S}, \tau, \kappa)$ is called*

(i) Compact if whenever $S = \bigvee_{i\in I} G_i$, $G_i \in \tau$, $i \in I$, there is a finite subset J of I with $\bigcup_{j\in J} G_j = S$.
(ii) Cocompact if whenever $\bigcap_{i\in I} F_i = \emptyset$, $F_i \in \kappa$, $i \in I$, there is a finite subset J of I with $\bigcap_{j\in J} F_j = \emptyset$.
(iii) Stable if every $K \in \kappa$ with $K \neq S$ is compact, i.e. whenever $K \subseteq \bigvee_{i\in I} G_i$, $G_i \in \tau$, $i \in I$, there is a finite subset J of I with $K \subseteq \bigcup_{j\in J} G_j$.
(iv) Costable if every $G \in \tau$ with $G \neq \emptyset$ is co-compact, i.e. whenever $\bigcap_{i\in I} F_i \subseteq G$, $F_i \in \kappa$, $i \in I$, there is a finite subset J of I with $\bigcap_{j\in J} F_j \subseteq G$.

We will refer to a ditopological texture space which has all four properties as a dicompact *space.*

Definition 2.2. *A ditopological space $(S, \mathcal{S}, \tau, \kappa)$ is bi-T_2 (bi-Hausdorff) if given $s, s' \in S$ with $Q_s \not\subseteq Q_{s'}$ there exists $H \in \tau, K \in \kappa$ with $H \subseteq K$, $P_s \not\subseteq K$ and $H \not\subseteq Q_{s'}$.*

The Category $\mathbf{ifPDiComp_{T2}}$ *:* Objects are dicompact, bi-Hausdorff plain texture spaces and morphisms are ω-preserving, bicontinuous point functions.

Accordingly, taking into consideration Definition 1.1, we have the following:

Theorem 2.3. *Let $(S, \mathcal{S}, \tau, \kappa) \in \mathrm{Ob}\,\mathbf{ifPDiComp_{T2}}$ and $(S, \mathfrak{J}_{\tau\kappa})$ the corresponding joint topological space. Then $(S, \mathfrak{J}_{\tau\kappa})$ is compact if and only if $(S, \mathcal{S}, \tau, \kappa)$ is dicompact. Moreover, $(S, \mathfrak{J}_{\tau\kappa})$ is Hausdorff.*

Proof. If $(S, \mathcal{S}, \tau, \kappa)$ is an object of $\mathbf{ifPDiComp_{T2}}$ then it is dicompact and bi-Hausdorff. In this case, the corresponding joint topological space $(S, \mathfrak{J}_{\tau\kappa})$ constructed as in Definition 1.1 is compact with the similar proof of (*cf.* [4, Proposition 4.6 (1)]). Additionally, since $(S, \mathcal{S}, \tau, \kappa)$ bi-Hausdorff the space $(S, \mathfrak{J}_{\tau\kappa})$ will be Hausdorff because of (*cf.* [4, Proposition 4.6 (2)]) □

Now we can turn our attention to the following two propositions proved in (*cf.* [5]).

Proposition 2.4. *Let $(S, \mathcal{S})$ be a plain texture and $A \subseteq S$. In this case, the family $\mathcal{S}_A = \{A \cap T \mid T \in \mathcal{S}\}$ defines a texturing on A and in particular, the pair $(A, \mathcal{S}_A)$ is a plain texture.*

Proposition 2.5. *The product of ditopological plain spaces is a ditopological plain space.*

Particularly, as far as the subsets of ditopological plain spaces are concerned, we have also the following statement from :

Proposition 2.6. *Let $(S, \mathcal{S}, \tau, \kappa)$ be a ditopological plain space and $A \subseteq S$. In this case, the families $\mathcal{S}_A = \mathcal{S}|_A = \{A \cap T \mid T \in \mathcal{S}\}$, $\tau_A = \tau|_A = \{A \cap T \mid T \in \tau\}$ and $\kappa_A = \kappa|_A = \{A \cap K \mid K \in \kappa\}$ decribes a texturing, topology and cotopology, respectively on A and in particular, the subspace $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ is a ditopological plain space.*

Additionally, let 's take a glance at the next proposition related to subditopological spaces:

Proposition 2.7. *If $(S, \mathcal{S}, \tau, \kappa)$ is bi-Hausdorff and $A \subseteq S$ then the induced subditopological texture space $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ is bi-Hausdorff.*

Proof. Let $(S, \mathcal{S}, \tau, \kappa)$ be bi-Hausdorff and $A \subseteq S$. Thus in order to prove that the subditopological space $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ is bi-Hausdorff, take $s, s' \in A$ with $Q_s \cap A = Q_s^A \nsubseteq Q_{s'}^A = Q'_s \cap A$. In this case, we have $s, s' \in S$ and $Q_s \nsubseteq Q_{s'}$, so there exists $H \in \tau, K \in \kappa$ with $H \subseteq K$, $P_s \nsubseteq K$ and $H \nsubseteq Q_{s'}$ since $(S, \mathcal{S}, \tau, \kappa)$ is bi-Hausdorff. Hence, $H \cap A \in \tau_A$, $K \cap A \in \kappa_A$ with $H \cap A \subseteq K \cap A$, $P_s^A \nsubseteq K \cap A$ and $H \cap A \nsubseteq Q_{s'}^A$. It means that the subditopological space $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ is bi-Hausdorff. □

According to the above considerations, we are now in a position to present a crucial theorem and a result as follows:

Theorem 2.8. *If $(S, \mathcal{S}, \tau, \kappa)$ is an object of* **ifPDiComp$_{\mathbf{T2}}$** *and $A \subseteq S$ is jointly closed then the induced subditopological space $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ is an object of* **ifPDiComp$_{\mathbf{T2}}$**.

Proof. Suppose that A is jointly closed set in S. In this case, A is closed with respect to the joint topology on S. Also, the joint topological space $(S, \mathcal{J}_{\tau\kappa})$ is compact by Proposition 2.3 (2) and so A is compact with respect to the restricted joint topology $\mathcal{J}_{\tau_A\kappa_A} = \mathcal{J}_{\tau\kappa}|_A$ on A. That is, the joint topological space $(A, \mathcal{J}_{\tau_A\kappa_A})$ of $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ is compact. That is, $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ will be dicompact. Moreover, since $(S, \mathcal{S}, \tau, \kappa)$ is bi-Hausdorff and $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ is a subspace of $(S, \mathcal{S}, \tau, \kappa)$, the subditopological space $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ becomes bi-Hausdorff by virtue of Proposition 2.7. Finally, $(A, \mathcal{S}_A, \tau_A, \kappa_A)$ is an object of **ifPDiComp$_{\mathbf{T2}}$**. □

Corollary 2.9. i) *The category* **ifPDiComp$_{\mathbf{T2}}$** *has products.*
ii) *The category* **ifPDiComp$_{\mathbf{T2}}$** *has equalizers.*

Proof. In view of Proposition 2.5, the part i) is clear. In addition, by using the definition of *equalizer* given in (*cf.* [1, Definition 7.51]), together with Proposition 2.6, the part ii) is seen, easily. □

Incidentally, note here that the fact the category **ifPDiComp$_{\mathbf{T2}}$** has equalizers and products is also trivial since it is a full subcategory of **ifPDitop**.

3. Inverse Limits In The Category ifPDiComp$_{\mathbf{T2}}$

First of all, let us recall the inverse system theory in the category **ifPDitop**:

Definition 3.1. *(cf.* [6]*) Let Λ be a directed set and $(S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha) \in \operatorname{Ob}$* **ifPDitop**, *$\alpha \in \Lambda$. If take the point functions*

$$\varphi_{\alpha\beta} : (S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha) \to (S_\beta, \mathcal{S}_\beta, \tau_\beta, \kappa_\beta)\ ,\ \alpha \geq \beta$$

which are w-preserving and bicontinuous as the morphisms of **ifPDitop**, *the family $\{(S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha), \varphi_{\alpha\beta}\}_{\alpha\geq\beta}$ is called* an inverse system of the ditopological plain spaces *over Λ or called* an inverse system in the category **ifPDitop** over Λ *if the conditions*

i) $\varphi_{\alpha\alpha} = id_{S_\alpha}$
ii) $\varphi_{\beta\gamma} \circ \varphi_{\alpha\beta} = \varphi_{\alpha\gamma}$, *for all* $\alpha \geq \beta \geq \gamma, \alpha, \beta, \gamma \in \Lambda$

are satisfied.

Since the category **ifPDiComp$_{\mathbf{T2}}$** has products and equalizers by Corollary 2.9, each inverse system in **ifPDiComp$_{\mathbf{T2}}$** yields a limit space. It will be manifest by virtue of the considerations given in (*cf.*[1, Sections 12,13]). Thus, we are now in a position to describe an inverse limit space for the inverse system in the category **ifPDiComp$_{\mathbf{T2}}$**, as follows:

Definition 3.2. *Let $\{(S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha), \varphi_{\alpha\beta}\}_{\alpha\geq\beta}$ be an inverse system in the category* **ifPDiComp$_{\mathbf{T2}}$** *over a directed set Λ. Form the product ditopological space*

$$\Big(\prod_\alpha S_\alpha, \bigotimes_\alpha \mathcal{S}_\alpha, \bigotimes_\alpha \tau_\alpha, \bigotimes_\alpha \kappa_\alpha\Big)_{\alpha\in\Lambda}$$

and for each $\alpha \in \Lambda$, let π_α be its projection (point) function onto the α th factor space $(S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha)$. We define a subset of the product $\prod_{\alpha\in\Lambda} S_\alpha$ as follows:

$$I = \{s \in \prod_\alpha S_\alpha \ \mid\ s = \{s_\alpha\}\ ,\ \textit{for each } \alpha, \beta \in \Lambda\ ,\ \alpha \geq \beta\ \Rightarrow\ \pi_\beta(s) = (\varphi_{\alpha\beta} \circ \pi_\alpha)(s)\}$$

Moreover, we have a subtexturing $\mathcal{I} = (\bigotimes_\alpha \mathcal{S}_\alpha) \mid_I$ *and a subditopology* (τ_I, κ_I) *as* $\tau_I = (\bigotimes_\alpha \tau_\alpha) \mid_I$, $\kappa_I = (\bigotimes_\alpha \kappa_\alpha) \mid_I$ *that is, as the restriction of the product ditopology* $(\bigotimes_\alpha \tau_\alpha, \bigotimes_\alpha \kappa_\alpha)_\alpha$ *on the product texture* $(\prod_\alpha S_\alpha, \bigotimes_\alpha \mathcal{S}_\alpha)_\alpha$ *to the set* I.

Therefore, we have a ditopological plain subtexture space $(I, \mathcal{I}, \tau_I, \kappa_I)$, *called* the inverse limit space *of the inverse system* $\{(S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha), \varphi_{\alpha\beta}\}_{\alpha \geq \beta}$ *and usually denoted by the notation*

$$(S_\infty, \mathcal{S}_\infty, \tau_\infty, \kappa_\infty)$$

shortly, where $S_\infty = \lim\limits_{\leftarrow} \{S_\alpha\}_{\alpha \in \Lambda}$.

We are now in a position to present a useful theorem associated with the preservation of the separation axiom bi-Hausdorff in the context of inverse systems and their limits:

Theorem 3.3. *The inverse limit of an inverse system consisting of bi-Hausdorff spaces in the category* **ifPDitop** *is a bi-Hausdorff space as an object of* **ifPDitop**.

Proof. By Proposition 2.5, **ifPDitop** has products and moreover, by (*cf.* [2]) the textural product of bi-Hausdorff ditopological spaces is a bi-Hausdorff space. Thus, the product of objects in the category **ifPDitop** is a bi-Hausdorff space. Also, from Proposition 2.7, it is known that any subditopological space of a bi-Hausdorff ditopological plain space is a bi-Hausdorff plain space. Hence, clearly the inverse limit space will be bi-Hausdorff as a subditopological space of that product. □

Theorem 3.4. *If* $\mathcal{A} = \{(S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha), \varphi_{\alpha\beta}\}_{\alpha \geq \beta}$ *is an inverse system in the category -* $\mathbf{ifPDiComp_{T2}}$ *over the directed set* Λ *then the inverse limit space* $(S_\infty, \mathcal{S}_\infty, \tau_\infty, \kappa_\infty)$ *of the inverse system* $\mathcal{A}$ *is an object of the category* $\mathbf{ifPDiComp_{T2}}$.

Proof. Assume that $(S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha) \in \mathrm{Ob}\, \mathbf{ifPDiComp_{T2}}$, $\alpha \in \Lambda$. In order to show that the space $(S_\infty, \mathcal{S}_\infty, \tau_\infty, \kappa_\infty)$ is jointly closed by using Definition 1.1, take a point $s \in \prod_\alpha S_\alpha$, $s \notin S_\infty$ and for $G \in \eta(s)$, $K \in \mu(s)$ suppose that $G \cap S_\infty \not\subseteq K$. In this case, $\mu_\alpha(s) \neq \varphi_{\beta\alpha}(\mu_\beta(s))$ for $\beta \geq \alpha$ and so $s_\alpha \neq \varphi_{\beta\alpha}(s_\beta)$. It means that $P_{s_\alpha} \neq P_{\varphi_{\beta\alpha}(s_\beta)}$ and since the texture $(S_\alpha, \mathcal{S}_\alpha)$ is plain, we have $Q_{s_\alpha} \neq Q_{\varphi_{\beta\alpha}(s_\beta)}$.

Now suppose that $Q_{s_\alpha} \not\subseteq Q_{\varphi_{\beta\alpha}(s_\beta)}$. Thus, there exists $H_1 \in \tau_\alpha$, $H_2 \in \kappa_\alpha$, $H_1 \subseteq H_2$, $H_1 \not\subseteq Q_{\varphi_{\beta\alpha}(s_\beta)}$ and $P_{s_\alpha} \not\subseteq H_2$ since the space $(S_\alpha, \mathcal{S}_\alpha, \tau_\alpha, \kappa_\alpha)$ is bi-Hausdorff. So $H_1 \in \eta(\varphi_{\beta\alpha}(s_\beta))$ and $H_2 \in \mu(s_\alpha)$.

On the other hand, note that the texture $(S_\beta, \mathcal{S}_\beta)$, $\beta \in \Lambda$ is plain. Hence $\varphi_{\beta\alpha}^{-1} H_1 \in \eta(s_\beta)$ and $\mu_\beta^{-1}[\varphi_{\beta\alpha}^{-1} H_1] = (\varphi_{\beta\alpha} \circ \mu_\beta)^{-1}[H_1] \in \eta(s)$, so $\mu_\alpha^{-1} H_1 \in \eta(s)$ since $\mu_\alpha = \varphi_{\beta\alpha} \circ \mu_\beta$. Also, $\mu_\alpha^{-1} H_2 \in \mu(s)$ as well. Hence the fact $\mu_\alpha^{-1} H_1 \cap S_\infty \not\subseteq \mu_\alpha^{-1} H_2$ is clear by the definition of joint topology given in Definition 1.1. In this case, there exists a point $a \in S_\infty$ such that $\mu_\alpha^{-1} H_1 \cap S_\infty \not\subseteq Q_a$ and $P_a \not\subseteq \mu_\alpha^{-1} H_2$, and so $\mu_\alpha(a) \in H_1$, $\mu_\alpha(a) \notin H_2$. But it contradicts with the fact that $H_1 \subseteq H_2$. In conclusion, $s \in S_\infty$. That is, the closure of the set S_∞ with respect to the joint topology, is itself. It means that the inverse limit space $(S_\infty, \mathcal{S}_\infty, \tau_\infty, \kappa_\infty)$ of $\mathcal{A}$ is jointly closed and thus it is an object of $\mathbf{ifPDiComp_{T2}}$ by Theorem 2.8. □

4. Conclusion

Various aspects of the inverse system-limit theory for ditopological spaces are investigated in the context of dicompact, bi-Hausdorff plain texture spaces and placed them in a categorical setting as the main purpose of this study.

There are considerable difficulties involved in constructing a suitable theory of inverse systems for general dicompact spaces. Hence, in the present context we confined our attention to the inverse systems and inverse limits constructed in the full subcategory $\mathbf{ifPDiComp_{T2}}$ of **ifPDitop**, whose objects are dicompact, bi-Hausdorff spaces which have plain texturing. Finally, We leave as an open problem the task of extending and translating the concepts and also results obtained here to more general categories in the theory of dicompact texture spaces.

Acknowledgements

The author would like to thank the referees for their constructive comments that have helped improve the readability of the paper.

References

[1] Adámek, J., Herrlich, H. and Strecker, G. ,E., (1990) Abstract and Concrete Categories, John Wiley & Sons, Inc. Volume 17, Springer-Verlag.
[2] Brown, L. M., (2005) Quotients of textures and of ditopological texture spaces, Topology Proceedings 29 (2), 337–368.
[3] Yıldız, F., (2006) Spaces of Bicontinuous Real Difunctions and real compactness, (Ph.D. Thesis (in Turkish), Hacettepe University.
[4] Yıldız, F. and Brown, L. M., (2007) Categories of dicompact bi-T_2 texture spaces and a Banach-Stone theorem, Quaestiones Mathematicae 30 (2007), 167–192.
[5] Yıldız, F.,(2016) Inverse systems and inverse limits in the category of plain textures, Topology and Its Applications, 201, 217-234.
[6] Yıldız, F., (2018) Some Categorical Aspects of The Inverse Limits In Ditopological Context, Applied General Topology, 19 (1), 101-127.

Hacettepe University, Faculty of Science, Department of Mathematics,Beytepe-Ankara, Turkey

Email(s): yfiliz@hacettepe.edu.tr