axioms

MDPI

Article

Derivation of Bounds for Majorization Differences by a Novel Method and Its Applications in Information Theory

Abdul Basir [1], Muhammad Adil Khan [1,*], Hidayat Ullah [1], Yahya Almalki [2], Saowaluck Chasreechai [3,4] and Thanin Sitthiwirattham [4,5,*]

[1] Department of Mathematics, University of Peshawar, Peshawar 25000, Pakistan; basirbjr@uop.edu.pk (A.B.); hidayat53@uop.edu.pk (H.U.)
[2] Department of Mathematics, College of Science, King Khalid University, Abha 61413, Saudi Arabia; yalmalki@kku.edu.sa
[3] Department of Mathematics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand; saowaluck.c@sci.kmutnb.ac.th
[4] Research Group for Fractional Calculus Theory and Applications, Science and Technology Research Institute, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand
[5] Mathematics Department, Faculty of Science and Technology, Suan Dusit University, Bangkok 10300, Thailand
* Correspondence: madilkhan@uop.edu.pk (M.A.K.); thanin_sit@dusit.ac.th (T.S.)

Abstract: In the recent era of research developments, mathematical inequalities and their applications perform a very consequential role in different aspects, and they provide an engaging area for research activities. In this paper, we propose a new approach for the improvement of the classical majorization inequality and its weighted versions in a discrete sense. The proposed improvements give several estimates for the majorization differences. Some earlier improvements of the Jensen and Slater inequalities are deduced as direct consequences of the obtained results. We also discuss the conditions under which the main results give better estimates for the majorization differences. Applications of the acquired results are also presented in information theory.

Keywords: convex function; majorization inequality; Slater's inequality; Jensen's inequality; information theory

MSC: 26D15; 26D20

check for updates

Citation: Basir, A.; Khan, M.A.; Ullah, H.; Almalki, Y.; Chasreechai, S.; Sitthiwirattham, T. Derivation of Bounds for Majorization Differences by a Novel Method and Its Applications in Information Theory. *Axioms* **2023**, *12*, 885. https://doi.org/10.3390/axioms12090885

Academic Editor: Behzad Djafari-Rouhani

Received: 9 August 2023
Revised: 14 September 2023
Accepted: 15 September 2023
Published: 16 September 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Mathematical inequalities play a crucial role in various areas of mathematics and have practical applications in many fields such as physics [1], engineering [2], economics [3], epidemiology [4], and statistics [5]. Actually, inequalities provide a way to compare and order numbers, variables, and quantities and help us to determine the relationships between quantities [6–8]. Due to these facts, inequalities are essential in many real-world situations to establish comparisons for making decisions or drawing conclusions [9,10]. Furthermore, inequalities also serve as fundamental tools for solving mathematical problems [11,12]. They are often used to establish conditions or constraints, which help us in formulating and solving equations, optimization problems, as well as systems of equations [7,13]. Inequalities can be used to provide a framework for identifying feasible solutions and finding the best possible outcomes [14–17]. Moreover, inequalities are frequently utilized in mathematical analysis and proof techniques [4,18]. Furthermore, they allow us to establish bounds, estimate limits, and prove the existence or non-existence of certain mathematical objects or properties [19]. In mathematical analysis, calculus, and advanced proof-based mathematics, inequalities are essential elements [20–22]. Mathematical inequalities play a central role to find the best fit solution in optimization problems [7]. In economics, engineering, and operations research, inequalities help us to determine optimal values of variables under some conditions [2,3,18]. Inequalities play a crucial role in decision making processes by

considering trade-offs and constraints [3,23,24]. In computer science and image processing, inequalities are used to develop efficient algorithms for solving and modeling complicated problems [25–27]. Linear programming, integer programming, and combinatorial optimization techniques rely heavily on inequalities to determine feasible solutions and guide search algorithms [5,18,28]. In simple words, mathematical inequalities are vital in various mathematical disciplines and have practical applications in diverse fields [11,29–31]. They provide a framework for comparison, analysis, optimization, decision making, modeling, and proof techniques. Also, they enable us to understand, solve, and make informed choices in complex mathematical and real-world scenarios [9,32–34].

The great potency behind the development of mathematical inequalities is the notion of convexity [24,35–37]. When we are talking about mathematical inequalities and do not account for convex functions, it is not fair to convex functions [18,30,38]. Convex functions are closely related to inequalities, because without convex functions several inequalities would not be possible to prove, such as the Hermite–Hadamard inequality [11], Jensen–Mercer inequality [39], Jensen–Steffensen inequality [18], Favard's inequality [40], and many others. The most inaugurating and interesting inequality for convex functions is the Jensen inequality [41]. This inequality has an intense relationship with convex functions in the sense that it generalizes ordinary convexity. This inequality provides a relationship between the expected value of the image of the convex function and the convex function of the expected value. Here is the statement of Jensen's inequality:

Let ψ be a convex function on $[\xi_1, \xi_2]$ and $\varrho_j \in [\xi_1, \xi_2]$, $\delta_j \geq 0$ for $j = 1, 2, \ldots, m$ with $\delta_* := \sum_{j=1}^{m} \delta_j > 0$. Then,

$$\psi\left(\frac{1}{\delta_*}\sum_{j=1}^{m}\delta_j\varrho_j\right) \leq \frac{1}{\delta_*}\sum_{j=1}^{m}\delta_j\psi(\varrho_j). \tag{1}$$

In the case of the concave function ψ, (1) is valid in the contrary direction. The formula that holds in a continuous form of the Jensen inequality is presented in (1):

Presume that the functions $g_1, g_2 : [c_1, c_2] \to [\xi_1, \xi_2]$ are continuous with $g_1 \geq 0$ and $g_* := \int_{c_1}^{c_2} g_1(\sigma)d\sigma > 0$. Then, the inequality

$$\psi\left(\frac{1}{g_*}\int_{c_1}^{c_2} g_1(\sigma)g_2(\sigma)d\sigma\right) \leq \frac{1}{g_*}\int_{c_1}^{c_2} g_1(\sigma)\psi\big(g_2(\sigma)\big)d\sigma \tag{2}$$

is valid, for every convex function ψ on $[\xi_1, \xi_2]$ such that the integral of $\psi \circ g_2$ exists. For the concave function ψ, (2) will be seen in the reverse sense.

The Jensen inequality has many interesting features, for example, it generalizes the triangular inequality, and also this inequality is the source of many other inequalities such as Hölder, Ky Fan's, Young's, and Hermite–Hadamard inequalities, as well as many others [41,42]. Moreover, this inequality also has a very desirable structure, and due to these facts, a lot of important variants, generalizations, applications, enhancements, and extensions have been provided for the Jensen inequality in different scales with the support of the characteristics and properties of convex functions and its generalizations [43–46]. In 1981, Slater presented a companion inequality to this inequality with the help of a convex function known as Slater's inequality [47]. The Slater inequality can be expressed as follows:

Suppose that $\varrho_j \in (\xi_1, \xi_2)$ and $\delta_j \geq 0$ for $j = 1, 2, \ldots, m$ with $\delta_* := \sum_{j=1}^{m} \delta_j > 0$ and $\psi : (\xi_1, \xi_2) \to \mathbb{R}$ is an increasing convex function such that $\varrho_* := \sum_{j=1}^{m} \delta_j \psi'_+(\varrho_j) \neq 0$. Then,

$$\frac{1}{\delta_*}\sum_{j=1}^{m}\delta_j\psi(\varrho_j) \leq \psi\left(\frac{1}{\varrho_*}\sum_{j=1}^{m}\delta_j\varrho_j\psi'_+(\varrho_j)\right). \tag{3}$$

For the concave function ψ, (3) reverses.

Further in this section, we devote our concentration to majorization. Majorization is a versatile strategy that has been employed to address many kinds of complications, including maximum likelihood estimation, non-negative matrix factorization, image reconstruction, and compressed sensing, among others [12,48]. It provides a powerful framework for solving complex optimization problems by iteratively minimizing simpler surrogate functions [28,49]. Majorization is a concept in mathematics that refers to a partial ordering relationship between vectors or sequences of numbers [50,51]. Given two vectors or sequences, one is said to majorize the other if it is "larger" in a certain sense. More formally, let $\mathbf{k}_1 = (\varrho_1, \varrho_2, \cdots, \varrho_m)$ and $\mathbf{k}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$ be any m-tuples. Then, we say $\mathbf{k}_1$ majorizes $\mathbf{k}_2$, denoted by $\mathbf{k}_1 \succ \mathbf{k}_2$, if we arrange both $\mathbf{k}_1$ and $\mathbf{k}_2$ in decreasing order and the following conditions are satisfied:

- $$\sum_{i=1}^{k} \varrho_j \geq \sum_{i=1}^{k} \sigma_j, \qquad k = 1, 2, \ldots, m-1, \tag{4}$$
- $$\sum_{j=1}^{m} \varrho_j = \sum_{j=1}^{m} \sigma_j. \tag{5}$$

In other words, $\mathbf{k}_1 \succ \mathbf{k}_2$ means that $\mathbf{k}_1$ has larger or equal stochastic dominance than $\mathbf{k}_2$ [52]. It implies that $\mathbf{k}_1$ has a more spread-out distribution than $\mathbf{k}_2$. The majorization ordering has some useful properties. For example, if $\mathbf{k}_1 \succ \mathbf{k}_2$, then any permutation of $\mathbf{k}_1$ majorizes the corresponding permutation of $\mathbf{k}_2$. Additionally, in the theory of convex functions and inequalities, majorization seems very essential. In 1932, Karamata [50] introduced an inequality for the majorized tuples via a convex function, which is known as the majorization inequality. This inequality states the following:

Consider ψ is convex over $[\xi_1, \xi_2]$ and $\mathbf{k}_1 = (\varrho_1, \varrho_2, \cdots, \varrho_m)$ and $\mathbf{k}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$ are any tuples such that $\varrho_j, \sigma_j \in [\xi_1, \xi_2]$ for $j = 1, 2, \cdots, m$. If $\mathbf{k}_1 \succ \mathbf{k}_2$, then

$$\sum_{j=1}^{m} \psi(\varrho_j) \geq \sum_{j=1}^{m} \psi(\sigma_j). \tag{6}$$

If the function ψ is concave, then (6) will reverse. Some years later, Fuchs [53] established the generalized form

$$\sum_{j=1}^{m} \alpha_j \psi(\varrho_j) \geq \sum_{j=1}^{m} \alpha_j \psi(\sigma_j) \tag{7}$$

of majorization inequality (6) by taking arbitrary weights $\mathbf{k}_3 = (\alpha_1, \alpha_2, \cdots, \alpha_m)$ and decreasing tuples $\mathbf{k}_1 = (\varrho_1, \varrho_2, \cdots, \varrho_m)$ and $\mathbf{k}_1 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$ with the conditions that $\sum_{i=1}^{k} \alpha_j \varrho_j \geq \sum_{i=1}^{k} \alpha_j \sigma_j$, $(k = 1, 2, \ldots, m-1)$, and $\sum_{j=1}^{m} \alpha_j \varrho_j = \sum_{j=1}^{m} \alpha_j \sigma_j$ while utilizing the notion of convexity. Inequality (7) is well known as the weighted version or generalized version of the majorization inequality. In 1973, Bullen et al. also proved inequality (7) by taking decreasing tuples with real weights and with relaxed majorization conditions $\sum_{i=1}^{k} \alpha_j \varrho_j \geq \sum_{i=1}^{k} \alpha_j \sigma_j$, $(k = 1, 2, \ldots, m-1, m)$, and keeping the convex function ψ stricter as monotonic. Malagranda et al. [40] obtained inequality (7) for positive weights while keeping the convex function and one tuple relaxed and the second tuple monotonic with the following majorization conditions: $\sum_{i=1}^{k} \alpha_j \varrho_j \geq \sum_{i=1}^{k} \alpha_j \sigma_j$, $(k = 1, 2, \ldots, m-1)$, and $\sum_{j=1}^{m} \alpha_j \varrho_j = \sum_{j=1}^{m} \alpha_j \sigma_j$. In 2004, Dragomir [54] also provided the weighted version of the majorization inequality by taking certain tuples with relaxed majorization conditions. In 2019,

Adil Khan et al. [55] presented the extension of the classical majorization inequality and its weighted versions under different circumstances for convex functions on rectangles. Wu et al. [56] acquired improvements of the majorization-type inequalities for convex functions through Taylor's theorem with a mean value form of remainder. In 2021, Deng et al. [46] refined the Jensen inequality through the theory of majorization by a new method and further explained the importance of the refined inequality by providing its applications in various domains. In 2022, Saeed et al. [57] utilized majorization results and presented refinements of the celebrated integral Jensen inequality. As a direct outcome of the main findings, they have granted modifications to the Hermite–Hadamard and Hölder inequalities. Furthermore, they support their results by providing its applications for the means as well as in information theory.

The main aim of this paper is to provide some estimations for the majorization differences through twice-differentiable functions. The article is designed in the following way:

- In Section 2, we will present main results, which give estimates for the majorization difference. Also, in this section, we will give bounds for the Slater and Jensen differences, which can be deduced from the major main results by performing some particular substitutions.
- In Section 3, we will highlight the conditions under which the main results will give better estimations.
- In Section 4, we will discuss how the main results can be applied in information theory.
- Section 5 is devoted to the concluding remarks on the article.

2. Main Results

This section concerns the main finding, which will provide estimates for the majorization, Jensen, and Slater differences. The principal goal of this section is to find estimates for the majorization differences by taking twice-differentiable functions. The intended estimates can be acquired by applying the notion of convexity, Hölder's inequality, prominent Jensen's inequality, and the famous power mean inequality. The direct consequences of each result will be discussed for the Jensen and Slater differences. Now, we begin this section with the subsequent lemma, which establishes an identity associated with the majorization difference.

Lemma 1. *Assume that* $\boldsymbol{m}_1 = (\varrho_1, \varrho_2, \cdots, \varrho_m)$, $\boldsymbol{m}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m) \in [\xi_1, \xi_2]^m$, *and* $\boldsymbol{m}_3 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_n) \in (0, \infty)^m$. *Further, let* $\psi : [\xi_1, \xi_2] \to \mathbb{R}$ *be a function where* $\psi^{''}$ *exists. Then,*

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) = \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \int_0^1 t\psi^{''}(t\sigma_j + (1-t)\varrho_j)dt - \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)\psi^{'}(\sigma_j). \tag{8}$$

Proof. Without the misfortune of a sweeping statement, assume that $\sigma_j \neq \varrho_j$ for all $j \in \{1, 2, \cdots, m\}$. Utilizing integration by parts, we have

$$\begin{aligned}
&\sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \int_0^1 t\psi^{''}(t\sigma_j + (1-t)\varrho_j)dt \\
&= \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \left(\frac{t}{\sigma_j - \varrho_j} \psi^{'}(t\sigma_j + (1-t)\varrho_j) \Big|_0^1 \right. \\
&\left. - \frac{1}{\sigma_j - \varrho_j} \int_0^1 \psi^{'}(t\sigma_j + (1-t)\varrho_j)dt \right) \\
&= \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \left(\frac{\psi^{'}(\sigma_j)}{\sigma_j - \varrho_j} - \frac{1}{(\sigma_j - \varrho_j)^2} \psi(t\sigma_j + (1-t)\varrho_j) \Big|_0^1 \right)
\end{aligned}$$

$$= \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \left(\frac{\psi^{'}(\sigma_j)}{\sigma_j - \varrho_j} - \frac{1}{(\sigma_j - \varrho_j)^2} \Big(\psi(\sigma_j) - \psi(\varrho_j) \Big) \right)$$
$$= \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j) \psi^{'}(\sigma_j) - \sum_{j=1}^{m} \varsigma_j \Big(\psi(\sigma_j) - \psi(\varrho_j) \Big)$$

which implies that

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) + \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j) \psi^{'}(\sigma_j)$$
$$= \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \int_0^1 t \psi^{''}(t\sigma_j + (1-t)\varrho_j) dt. \tag{9}$$

Clearly, (9) is equivalent to (8). □

In the next theorem, we make use of the Hölder inequality and the concept of a convex function to set up an inequality for the majorization difference.

Theorem 1. *Presume that all predications in Lemma 1 are valid. Moreover, let* $|\psi^{''}|^q$ $(q > 1)$ *be a convex function defined over* $[\xi_1, \xi_2]$. *Then,*

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) \leq - \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j) \psi^{'}(\sigma_j)$$
$$+ \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \left(\frac{(q+1)\left|\psi^{''}(\sigma_j)\right|^q + \left|\psi^{''}(\varrho_j)\right|^q}{(q+1)(q+2)} \right)^{\frac{1}{q}}. \tag{10}$$

Proof. Since $\psi(x) \leq |\psi(x)|$ for all $x \in [\xi_1, \xi_2]$, therefore, from (8), we can write that

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) \leq - \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j) \psi^{'}(\sigma_j)$$
$$+ \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \int_0^1 \left| t \psi^{''}(t\sigma_j + (1-t)\varrho_j) \right| dt. \tag{11}$$

Applying the Hölder inequality to (11), we deduce

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) \leq - \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j) \psi^{'}(\sigma_j)$$
$$+ \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \left(\int_0^1 t^q \left| \psi^{''}(t\sigma_j + (1-t)\varrho_j) \right|^q dt \right)^{\frac{1}{q}}. \tag{12}$$

Now, using the convexity of the function $|\psi^{''}|^q$ in (12), we obtain

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) \leq - \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j) \psi^{'}(\sigma_j)$$
$$+ \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \left(\left|\psi^{''}(\sigma_j)\right|^q \int_0^1 t^{q+1} dt + \left|\psi^{''}(\varrho_j)\right|^q \int_0^1 t^q (1-t) dt \right)^{\frac{1}{q}}. \tag{13}$$

By evaluating integrals in (13), we arrive at (10). □

The subsequent corollary follows from Theorem 1, which provides an estimation for the Jensen difference.

Corollary 1. *Assume that* $\varsigma_j \geq 0$ *and* $\varrho_j \in [\xi_1, \xi_2]$ *for* $j = 1, 2, \ldots, m$ *with* $\varsigma^* := \sum_{j=1}^{m} \varsigma_j > 0$ *and* $\overline{\varrho} := \frac{1}{\varsigma^*} \sum_{j=1}^{m} \varsigma_j \varrho_j$. *If* ψ *is a twice-differentiable function over* $[\xi_1, \xi_2]$ *such that* $|\psi''|^q$ *is convex for* $q > 1$, *then*

$$\frac{1}{\varsigma^*} \sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \psi(\overline{\varrho}) \leq \sum_{j=1}^{m} \varsigma_j (\overline{\varrho} - \varrho_j)^2 \Big(\frac{(q+1)|\psi''(\overline{\varrho})|^q + |\psi''(\varrho_j)|^q}{(q+1)(q+2)} \Big)^{\frac{1}{q}}. \tag{14}$$

Proof. Utilizing $\sigma_j = \overline{\varrho}$ in (10), we obtain (14). □

Remark 1. *Ullah et al. [41] in inequality* (4) *also provided a similar estimate given on the right side of inequality* (14) *for the absolute Jensen difference.*

The preceding corollary estimates the Slater difference as an output of Theorem 1.

Corollary 2. *Presume that* ψ *is a twice-differentiable function over* $[\xi_1, \xi_2]$ *and* $\varsigma_j \geq 0$ *and* $\varrho_j \in [\xi_1, \xi_2]$ *for* $j = 1, 2, \ldots, m$ *with* $\varsigma^* := \sum_{j=1}^{m} \varsigma_j > 0$ *and* $\frac{1}{\varsigma^*} \sum_{j=1}^{m} \varsigma_j \psi'(\varrho_j) \neq 0$. *Further, let* $\overline{\rho} := \frac{1}{\sum_{j=1}^{m} \varsigma_j \psi'(\varrho_j)} \sum_{j=1}^{m} \varsigma_j \varrho_j \psi'(\varrho_j) \in [\xi_1, \xi_2]$ *and* $|\psi''|^q$ *be convex for* $q > 1$. *Then,*

$$\frac{1}{\varsigma^*} \sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \psi(\overline{\rho}) \leq \sum_{j=1}^{m} \varsigma_j (\overline{\rho} - \varrho_j)^2 \Big(\frac{(q+1)|\psi''(\overline{\rho})|^q + |\psi''(\varrho_j)|^q}{(q+1)(q+2)} \Big)^{\frac{1}{q}}. \tag{15}$$

Proof. Inequality (15) can easily be deduced by replacing σ_j with $\overline{\rho}$ in (10). □

Remark 2. *A similar estimate given on the right side of inequality* (15) *was also presented in inequality* (9) *as stated in the article [58] for the absolute Slater difference.*

The theorem that follows offers an inequality for majorization that could possibly be constructed by using the Hölder inequality and the convex function formulation.

Theorem 2. *Presume that the postulates of Theorem 1 are valid. Also, let* $\frac{1}{p} + \frac{1}{q} = 1$. *Then,*

$$\begin{aligned} \sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) &\leq - \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j) \psi'(\sigma_j) \\ &+ \Big(\frac{1}{p+1} \Big)^{\frac{1}{p}} \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \Big(\frac{|\psi''(\sigma_j)|^q + |\psi''(\varrho_j)|^q}{2} \Big)^{\frac{1}{q}}. \end{aligned} \tag{16}$$

Proof. By utilizing the Hölder inequality on the right side of (11), we arrive at

$$\begin{aligned} \sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) &\leq - \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j) \psi'(\sigma_j) \\ &+ \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \Big(\int_0^1 t^p dt \Big)^{\frac{1}{p}} \Big(\int_0^1 \big| \psi''(t\sigma_j + (1-t)\varrho_j) \big|^q dt \Big)^{\frac{1}{q}} \end{aligned}$$

$$= +\left(\frac{1}{p+1}\right)^{\frac{1}{p}} \sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)^2 \left(\int_0^1 \left|\psi''(t\sigma_j + (1-t)\varrho_j)\right|^q dt \right)^{\frac{1}{q}}$$
$$- \sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)\psi'(\sigma_j). \tag{17}$$

To obtain inequality (18), we just use the convex function definition on the right side of (17):

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) \leq - \sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)\psi'(\sigma_j)$$
$$+\left(\frac{1}{p+1}\right)^{\frac{1}{p}} \sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)^2 \left(|\psi''(\sigma)|^q \int_0^1 t dt + |\psi''(\varrho)|^q \int_0^1 (1-t) dt \right)^{\frac{1}{q}}. \tag{18}$$

Now, finding integrals in (18), we obtain (16). □

The subsequent corollary is a direct consequence of Theorem 2, which provides a bound for the Jensen difference.

Corollary 3. *Assume that the conditions of Corollary 1 are satisfied. Moreover, if* $\frac{1}{p} + \frac{1}{q} = 1$, *then*

$$\frac{1}{\varsigma^*} \sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \psi(\overline{\varrho}) \leq \frac{1}{\varsigma^*} \left(\frac{1}{p+1}\right)^{\frac{1}{p}} \sum_{j=1}^{m} \varsigma_j(\overline{\varrho} - \varrho_j)^2 \times \left(\frac{|\psi''(\overline{\varrho})|^q + |\psi''(\varrho_j)|^q}{2} \right)^{\frac{1}{q}}. \tag{19}$$

Proof. To obtain (19), apply inequality (16) for $\sigma_j = \overline{\varrho}$. □

Remark 3. *The same type of estimate stated on the right side of* (19) *is also obtained in inequality* (9) *in [41] for the absolute Jensen difference.*

The below corollary is another consequence of Theorem 2.

Corollary 4. *Assume that the conditions of Corollary 2 are satisfied. Moreover, if* $\frac{1}{p} + \frac{1}{q} = 1$, *then*

$$\frac{1}{\varsigma^*} \sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \psi(\overline{\rho}) \leq \frac{1}{\varsigma^*} \left(\frac{1}{p+1}\right)^{\frac{1}{p}} \sum_{j=1}^{m} \varsigma_j(\overline{\rho} - \varrho_j)^2 \times \left(\frac{|\psi''(\overline{\rho})|^q + |\psi''(\varrho_j)|^q}{2} \right)^{\frac{1}{q}}. \tag{20}$$

Proof. Inequality (20) is simply achieved through substituting $\sigma_j = \overline{\rho}$ into (16). □

Remark 4. *The estimate provided in* (20) *for the Slater difference is also acquired by Khan et al. [58] in inequality* (15) *for the absolute Slater difference.*

Remark 5. *It is important to note that the above results are better with respect to the condition of convexity of* $|\psi''|^q$ *because there are some functions* ϕ *such that* $|\psi|^q$ *or* $|\psi'|^q$ *is not convex while* $|\psi''|^q$ *is convex. For example, if the function* ϕ *is defined by* $\phi(x) = \sqrt{x},\ x > 0$, *then clearly the function* $|\psi|^q$ *is not convex for* $q \in (1,2)$ *but* $|\psi''|^q$ *is convex. Similarly, if the function* ψ *is defined by* $\psi(x) = x^k,\ x > 0, k \in (1, 1 + \frac{1}{q})$, *then clearly the function* $|\psi'|^q$ *is not convex but* $|\psi''|^q$ *is convex.*

The following theorem establishes an inequality for the majorization difference through Hölder's and Jensen's inequalities.

Theorem 3. *Let the conditions of Lemma 1 be satisfied, and further assume that* $|\psi''|$ *is a concave function. Then,*

$$\begin{aligned}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\sum_{j=1}^{m}\varsigma_j\psi(\sigma_j) \quad &\leq \quad \frac{1}{2}\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left|\psi''\left(\frac{2\sigma_j+\varrho_j}{3}\right)\right| \\ &\quad -\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j). \end{aligned} \tag{21}$$

Proof. By using the property of the absolute function, identity (8) can be written as

$$\begin{aligned}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\sum_{j=1}^{m}\varsigma_j\psi(\sigma_j) \quad &\leq \quad \sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\int_0^1|t\psi''(t\sigma_j+(1-t)\varrho_j)|dt \\ &\quad -\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j). \end{aligned} \tag{22}$$

Currently utilizing the Jensen inequality on the right side of (22), we obtain

$$\begin{aligned}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\sum_{j=1}^{m}\varsigma_j\psi(\sigma_j) &\leq -\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j) \\ &\quad +\frac{1}{2}\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left|\psi''\left(\frac{\int_0^1 t\Big(t\sigma_j+(1-t)\varrho_j\Big)dt}{\int_0^1 tdt}\right)\right|. \end{aligned} \tag{23}$$

Simplifying (23), we obtain (21). □

An estimate for the Jensen difference is given in the coming corollary as a consequence of Theorem 3.

Corollary 5. *Assume that* $\varsigma_j\geq 0$ *and* $\varrho_j\in[\xi_1,\xi_2]$ *for* $j=1,2,\ldots,m$ *with* $\varsigma^*:=\sum_{j=1}^{m}\varsigma_j>0$ *and* $\overline{\varrho}:=\frac{1}{\varsigma^*}\sum_{j=1}^{m}\varsigma_j\varrho_j$. *If* ψ *is a twice-differentiable function over* $[\xi_1,\xi_2]$ *such that* $|\psi''|$ *is concave, then*

$$\frac{1}{\varsigma^*}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\psi(\overline{\varrho})\leq\frac{1}{2\varsigma^*}\sum_{j=1}^{m}\varsigma_j(\overline{\varrho}-\varrho_j)^2\left|\psi''\left(\frac{2\overline{\varrho}+\varrho_j}{3}\right)\right|. \tag{24}$$

Proof. Consider $\sigma_j=\overline{\varrho}$ for all $i\in\{1,2,\ldots,n\}$ in inequality (21). We obtain (24). □

Remark 6. *An estimate for the absolute Jensen difference acquired in inequality* (12) *in the article [41] will become alike to the estimate stated on the right side of inequality* (24) *by assuming* $q=1$ *in* (12).

As a consequence of Theorem 3, the below corollary presents a bound for the Slater difference.

Corollary 6. *Assume that* $\varsigma_j\geq 0$ *and* $\varrho_j\in[\xi_1,\xi_2]$ *for* $j=1,2,\ldots,m$ *with* $\varsigma^*:=\sum_{j=1}^{m}\varsigma_j>0$ *and* ψ *is a twice-differentiable function over* $[\xi_1,\xi_2]$ *such that* $|\psi''|$ *is concave. If* $\frac{1}{\varsigma^*}\sum_{j=1}^{m}\varsigma_j\psi'(\varrho_j)\neq 0$ *and* $\overline{\rho}:=\frac{1}{\sum_{j=1}^{m}\varsigma_j\psi'(\varrho_j)}\sum_{j=1}^{m}\varsigma_j\varrho_j\psi'(\varrho_j)\in[\xi_1,\xi_2]$, *then*

$$\frac{1}{\varsigma^*}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\psi(\overline{\rho})\leq\frac{1}{2\varsigma^*}\sum_{j=1}^{m}\varsigma_j(\overline{\rho}-\varrho_j)^2\left|\psi''\left(\frac{2\overline{\rho}+\varrho_j}{3}\right)\right|. \tag{25}$$

Proof. By taking $\sigma_j=\overline{\rho}$ for all $j\in\{1,2,\cdots,m\}$ in (21), we arrive at the desired inequality (25). □

Remark 7. *The estimate provided by Khan et al. [58] in* (13) *for the absolute Slater difference will coincide with the estimate on the right side of* (25) *by just taking* $p=1$ *in* (13).

The next theorem provides another relation for the majorization difference, which can be deduced by utilizing the Hölder inequality and Jensen inequality.

Theorem 4. *Let the conditions of Lemma 1 be satisfied, and further assume that* $|\psi''|^q$ *is a concave function for* $q>1$. *If* $\frac{1}{p}+\frac{1}{q}=1$, *then*

$$\begin{aligned}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\sum_{j=1}^{m}\varsigma_j\psi(\sigma_j)\leq&-\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j)\\&+\left(\frac{1}{p+1}\right)^{\frac{1}{p}}\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left|\psi''\left(\frac{\sigma_j+\varrho_j}{2}\right)\right|.\end{aligned} \tag{26}$$

Proof. By utilizing the Hölder inequality on the right side of (22), we obtain

$$\begin{aligned}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\sum_{j=1}^{m}\varsigma_j\psi(\sigma_j)\leq&-\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j)\\&+\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left(\int_0^1 t^p dt\right)^{\frac{1}{p}}\left(\int_0^1\left|\psi''(t\sigma_j+(1-t)\varrho_j)\right|^q dt\right)^{\frac{1}{q}}\\=&\left(\frac{1}{p+1}\right)^{\frac{1}{p}}\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left(\int_0^1\left|\psi''(t\sigma_j+(1-t)\varrho_j)\right|^q dt\right)^{\frac{1}{q}}\\&-\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j).\end{aligned} \tag{27}$$

Now, applying Jensen's inequality to (27), we acquire

$$\begin{aligned}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\sum_{j=1}^{m}\varsigma_j\psi(\sigma_j)\leq&-\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j)\\&+\left(\frac{1}{p+1}\right)^{\frac{1}{p}}\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left|\psi''\left(\int_0^1\Big(t\sigma_j+(1-t)\varrho_j\Big)dt\right)\right|.\end{aligned} \tag{28}$$

To obtain (26), we find the integral in (28). □

Corollary 7 is a direct outcome of Theorem 4.

Corollary 7. *Assume that* $\varsigma_j\geq 0$ *and* $\varrho_j\in[\xi_1,\xi_2]$ *for* $j=1,2,\ldots,m$ *with* $\varsigma^*:=\sum_{j=1}^{m}\varsigma_j>0$ *and* $\overline{\varrho}:=\frac{1}{\varsigma^*}\sum_{j=1}^{m}\varsigma_j\varrho_j$. *Also, suppose that* ψ *is a twice-differentiable function over* $[\xi_1,\xi_2]$ *such that* $|\psi''|^q$ *is concave for* $q>1$. *If* $\frac{1}{p}+\frac{1}{q}=1$, *then*

$$\frac{1}{\varsigma^*}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\psi(\overline{\varrho})\leq\frac{1}{\varsigma^*}\left(\frac{1}{p+1}\right)^{\frac{1}{p}}\sum_{j=1}^{m}\varsigma_j(\overline{\varrho}-\varrho_j)^2\left|\psi''\left(\frac{\overline{\varrho}+\varrho_j}{2}\right)\right|. \tag{29}$$

Proof. To obtain inequality (29), just replace σ_j with $\overline{\varrho}$ in inequality (26). □

Remark 8. *Ullah et al. [41] in inequality* (14) *acquired the same type of estimate indicated on the right side of* (29) *for the absolute Jensen difference.*

Another consequence of Theorem 4 in terms of an estimate for the Slater difference is in the following corollary.

Corollary 8. *Assume that* $\varsigma_j\geq 0$ *and* $\varrho_j\in[\xi_1,\xi_2]$ *for* $j=1,2,\ldots,m$ *with* $\varsigma^*:=\sum_{j=1}^{m}\varsigma_j>0$ *and* $p,q>1$ *with* $\frac{1}{p}+\frac{1}{q}=1$. *Further, let* ψ *be a twice-differentiable function on* $[\xi_1,\xi_2]$ *such that* $|\psi''|^q$ *is concave,* $\frac{1}{\varsigma^*}\sum_{j=1}^{m}\varsigma_j\psi'(\varrho_j)\neq 0$, *and* $\overline{\rho}:=\frac{1}{\sum_{j=1}^{m}\varsigma_j\psi'(\varrho_j)}\sum_{j=1}^{m}\varsigma_j\varrho_j\psi'(\varrho_j)\in[\xi_1,\xi_2]$. *Then,*

$$\frac{1}{\varsigma^*}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\psi(\overline{\rho})\leq\frac{1}{\varsigma^*}\left(\frac{1}{p+1}\right)^{\frac{1}{p}}\sum_{j=1}^{m}\varsigma_j(\overline{\rho}-\varrho_j)^2\left|\psi''\left(\frac{\overline{\rho}+\varrho_j}{2}\right)\right|. \tag{30}$$

Proof. By utilizing (26) for $\sigma_j=\overline{\rho}$, we obtain (30). □

Remark 9. *An estimate for the Slater difference given on the right side of* (30) *is also produced by Khan et al. [58] in* (19) *for the absolute Slater difference.*

Theorem 5 has been documented by implementing the definition of a convex function and the notable power mean inequality, which provides a bound for the majorization difference.

Theorem 5. *Assuming the presumptions of Theorem 1 are true, then*

$$\begin{aligned}\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\sum_{j=1}^{m}\varsigma_j\psi(\sigma_j)&\leq-\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j)\\&+\left(\frac{1}{2}\right)^{1-\frac{1}{q}}\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left(\frac{2\left|\psi''(\sigma_j)\right|^q+\left|\psi''(\varrho_j)\right|^q}{6}\right)^{\frac{1}{q}}.\end{aligned} \tag{31}$$

Proof. Once the power mean inequality applies to (11), we find

$$\begin{aligned}&\sum_{j=1}^{m}\varsigma_j\psi(\varrho_j)-\sum_{j=1}^{m}\varsigma_j\psi(\sigma_j)\leq-\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j)\\&+\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left(\int_0^1 t\,dt\right)^{1-\frac{1}{q}}\left(\int_0^1 t\left|\psi''(t\sigma_j+(1-t)\varrho_j)\right|^q dt\right)^{\frac{1}{q}}\\&=\left(\frac{1}{2}\right)^{1-\frac{1}{q}}\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)^2\left(\int_0^1 t\left|\psi''(t\sigma_j+(1-t)\varrho_j)\right|^q dt\right)^{\frac{1}{q}}\\&-\sum_{j=1}^{m}\varsigma_j(\sigma_j-\varrho_j)\psi'(\sigma_j).\end{aligned} \tag{32}$$

Utilizing the convexity of the function $|\psi''|^q$ on the right side of (32), we obtain

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j) \leq -\sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)\psi'(\sigma_j)$$
$$+\left(\frac{1}{2}\right)^{1-\frac{1}{q}} \sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)^2 \left(|\psi''(\sigma)|^q \int_0^1 t^2 dt + |\psi''(\varrho)|^q \int_0^1 t(1-t)dt\right)^{\frac{1}{q}}. \tag{33}$$

By calculating integrals in (33), we arrive at (31). □

As an immediate outcome of Theorem 5, the next corollary offers a bound for the Jensen difference.

Corollary 9. *Let the assumptions of Corollary 1 hold. Then,*

$$\frac{1}{\varsigma^*}\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \psi(\overline{\varrho}) \leq \frac{1}{\varsigma^*}\left(\frac{1}{2}\right)^{1-\frac{1}{q}} \sum_{j=1}^{m} \varsigma_j (\overline{\varrho} - \varrho_j)^2 \times \left(\frac{2|\psi''(\overline{\varrho})|^q + |\psi''(\varrho_j)|^q}{6}\right)^{\frac{1}{q}}. \tag{34}$$

Proof. By assuming $\sigma_j = \overline{\varrho}$ in inequality (31), we deduce (34). □

Remark 10. *In inequality* (9)*, Ullah et al. [41] also obtained a similar estimate for the absolute Jensen difference given on the right side of inequality* (34)*.*

The subsequent corollary generates an estimate for the Slater difference as an effect of Theorem 5.

Corollary 10. *Let the assumptions of Corollary 2 hold. Then,*

$$\frac{1}{\varsigma^*}\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \psi(\overline{\rho}) \leq \frac{1}{\varsigma^*}\left(\frac{1}{2}\right)^{1-\frac{1}{q}} \sum_{j=1}^{m} \varsigma_j (\overline{\rho} - \varrho_j)^2 \times \left(\frac{2|\psi''(\overline{\rho})|^q + |\psi''(\varrho_j)|^q}{6}\right)^{\frac{1}{q}}. \tag{35}$$

Proof. To obtain (35), consider $\sigma_j = \overline{\rho}$ in (31). □

Remark 11. *Similarly, a type of estimates for the Slater difference given in* (35) *can be studied in* (21)*, which is stated in the article [58] for the absolute Slater difference.*

3. Discussion about the Betterness of the Main Results

This section discusses the conditions under which the estimates given in Theorem 1 to Theorem 5 will improve.

The results given in Theorem 1 to Theorem 5 will provide good and superior estimates for the majorization difference

$$\sum_{j=1}^{m} \varsigma_j \psi(\varrho_j) - \sum_{j=1}^{m} \varsigma_j \psi(\sigma_j),$$

if

$$\sum_{j=1}^{m} \varsigma_j (\sigma_j - \varrho_j)\varsigma_j \psi'(\sigma_j) \geq 0. \tag{36}$$

Now, we are going to discuss certain conditions under which the term

$$\sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)\varsigma_j)\psi^{'}(\sigma_j) \tag{37}$$

is non-negative.

(1) For $p_j = 1$ for all $j \in \{1, 2, \cdots, m\}$, the term "$\sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)\varsigma_j)\psi^{'}(\sigma_j)$" is non-negative if the function ψ is convex and $\mathbf{m}_2 \succ \mathbf{m}_1$. The details can be found in [50,52].

(2) By utilizing the idea of the proof of the result given in [40], it can be proven that the term "$\sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)\varsigma_j)\psi^{'}(\sigma_j)$" is non-negative by imposing the conditions that ψ is a convex function and

(i) $\mathbf{m}_2$ is a decreasing tuple with

$$\sum_{j=1}^{k} \varsigma_j\varrho_j \leq \sum_{j=1}^{k} \varsigma_j\sigma_j, \qquad k = 1, 2, \ldots, m-1, \tag{38}$$

and

$$\sum_{j=1}^{m} \varsigma_j\varrho_j = \sum_{j=1}^{m} \varsigma_j\sigma_j, \tag{39}$$

OR

(ii) $\mathbf{m}_2$ is an increasing tuple such that

$$\sum_{j=1}^{k} \varsigma_j\varrho_j \geq \sum_{j=1}^{k} \varsigma_j\sigma_j, \qquad k = 1, 2, \ldots, m-1, \tag{40}$$

and

$$\sum_{j=1}^{m} \varsigma_j\varrho_j = \sum_{j=1}^{m} \varsigma_j\sigma_j. \tag{41}$$

(3) If $\mathbf{m}_2$ and $\mathbf{m}_1 - \mathbf{m}_2$ are monotonic in the same sense and satisfying the condition

$$\sum_{j=1}^{m} \varsigma_j\varrho_j = \sum_{j=1}^{m} \varsigma_j\sigma_j, \tag{42}$$

then employing the procedure of the proof of the theorem in [54], one easily show that "$\sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)\varsigma_j)\psi^{'}(\sigma_j)$" is non-negative.

(4) By adopting the method of the proof of the result given in [54] with the assumptions that the function ψ is an increasing convex function and the tuples $\mathbf{m}_2$ and $\mathbf{m}_1 - \mathbf{m}_2$ are monotonic in the parallel direction with

$$\sum_{j=1}^{m} \varsigma_j\varrho_j \geq \sum_{j=1}^{m} \varsigma_j\sigma_j, \tag{43}$$

then we can have $\sum_{j=1}^{m} \varsigma_j(\sigma_j - \varrho_j)\varsigma_j)\psi^{'}(\sigma_j) \geq 0$.

4. Applications in Information Theory

The current section is devoted to the applications of the main results in information theory. The intended applications will provide bounds for the Csiszár and Kullback–Leibler divergences, Shannon entropy, and the Bhattacharyya coefficient.

Definition 1. *Let ψ be real-valued convex function defined on $[a,b]$. Then, for $\boldsymbol{m}_1 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_m)$, $\boldsymbol{m}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m) \in (0,\infty)^m$, the Csiszár divergence is defined by*

$$C_\psi(\boldsymbol{m}_1, \boldsymbol{m}_2) = \sum_{j=1}^{m} \varsigma_j \psi\left(\frac{\sigma_j}{\varsigma_j}\right). \tag{44}$$

Theorem 6. *Assume that $\psi : [\xi_1, \xi_2] \to \mathbb{R}$ is a twice-differentiable function such that $|\psi''|^q$ $(q > 1)$ is convex. Further, let $\boldsymbol{m}_1 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_m)$ and $\boldsymbol{m}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$ be m-tuples and $\boldsymbol{m}_3 = (\varrho_1, \varrho_2, \cdots, \varrho_m)$ be a positive m-tuple. Then,*

$$\begin{aligned} C_\psi(\boldsymbol{m}_1, \boldsymbol{m}_3) - C_\psi(\boldsymbol{m}_1, \boldsymbol{m}_2) \quad \leq \quad & \sum_{j=1}^{m} \frac{(\sigma_j - \varrho_j)^2}{\varsigma_j} \left(\frac{(q+1)\left|\psi''\left(\frac{\sigma_j}{\varsigma_j}\right)\right|^q + \left|\psi''\left(\frac{\varrho_j}{\varsigma_j}\right)\right|^q}{(q+1)(q+2)} \right)^{\frac{1}{q}} \\ & - \sum_{j=1}^{m} (\sigma_j - \varrho_j) \psi'\left(\frac{\sigma_j}{\varsigma_j}\right). \end{aligned} \tag{45}$$

Proof. By utilizing inequality (10) for $\varrho_j = \frac{\varrho_j}{\varsigma_j}$ and $\sigma_j = \frac{\sigma_j}{\varsigma_j}$, we obtain (45). □

Definition 2. *Shannon entropy is defined for any positive probability distribution $\boldsymbol{m}_1 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_m)$ as follows:*

$$\boldsymbol{S}_{\boldsymbol{m}_1} = -\sum_{j=1}^{m} \varsigma_j \log \varsigma_j. \tag{46}$$

Shannon entropy plays a very important role in information theory. Numerous outcomes and results are devoted to Shannon entropy. The basic result of information theory is the statement derived by Shannon that the entropy formula is the only formula that obeys special criteria which are required for a measure of the uncertainty related to the outcome of a random variable. For some important results and applications of Shannon entropy in information theory, we recommend [59].

Corollary 11. *Let $\boldsymbol{m}_1 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_m)$ and $\boldsymbol{m}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$ be any positive tuples with $\sum_{j=1}^{m} \varsigma_j = 1$ and $q > 1$. Then,*

$$\begin{aligned} \sum_{j=1}^{m} \varsigma_j \log\left(\frac{\sigma_j}{\varsigma_j}\right) - \boldsymbol{S}_{\boldsymbol{m}_1} \quad \leq \quad & \sum_{j=1}^{m} \frac{(\sigma_j - 1)^2}{\varsigma_j} \left(\frac{(q+1)\left(\frac{\varsigma_j}{\sigma_j}\right)^{2q} + \varsigma_j^{2q}}{(q+1)(q+2)} \right)^{\frac{1}{q}} \\ & + \sum_{j=1}^{m} (\sigma_j - 1)\left(\frac{\varsigma_j}{\sigma_j}\right). \end{aligned} \tag{47}$$

Proof. Consider $\psi(\varrho) = -\log \varrho$, $\varrho > 0$. Then, $\psi''(\varrho) = \frac{1}{\varrho^2}$ and $\left(\left|\psi''(\varrho)\right|^q\right)'' = 2q(2q+1)\varrho^{-(2q+2)}$. Clearly, ψ'' and $\left(\left|\psi''\right|^q\right)''$ are positive on $(0,\infty)$, which confirms the convexity of ψ and $\left|\psi''\right|^q$. Therefore, applying (45) for $\psi(\varrho) = -\log \varrho$ and $\sigma_j = 1$ for $j = 1, 2, \ldots, m$, we deduce (47). □

Definition 3. *For any positive probability distributions $\boldsymbol{m}_1 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_m)$ and $\boldsymbol{m}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$, the Kullback–Liebler divergence is defined as*

$$\boldsymbol{K}_{(\boldsymbol{m}_1, \boldsymbol{m}_2)} = \sum_{j=1}^{m} \varsigma_j \log\left(\frac{\varsigma_j}{\sigma_j}\right). \tag{48}$$

Corollary 12. *Let* $\boldsymbol{m}_1 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_m)$, $\boldsymbol{m}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$, *and* $\boldsymbol{m}_3 = (\varrho_1, \varrho_2, \cdots, \varrho_m)$ *be positive probability distributions and* $q > 1$. *Then,*

$$\begin{aligned} \boldsymbol{K}_{(\boldsymbol{m}_1,\boldsymbol{m}_3)} - \boldsymbol{K}_{(\boldsymbol{m}_1,\boldsymbol{m}_2)} \quad \leq \quad & \sum_{j=1}^{m} \frac{(\sigma_j - \varrho_j)^2}{\varsigma_j} \left(\frac{(q+1)\left(\frac{\varsigma_j}{\sigma_j}\right)^{2q} + \left(\frac{\varsigma_j}{\varrho_j}\right)^{2q}}{(q+1)(q+2)} \right)^{\frac{1}{q}} \\ & + \sum_{j=1}^{m} (\sigma_j - \varrho_j) \left(\frac{\varsigma_j}{\sigma_j}\right). \end{aligned} \tag{49}$$

Proof. For the specified conditions, the functions $\psi(x) = -\log x$ and $|\psi''|^q = x^{-2q}$ are convex on $(0, \infty)$. Therefore, utilizing (45) for $\psi(x) = -\log x$, we obtain (49). □

Definition 4. *For arbitrary probability distributions* $\boldsymbol{m}_1 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_m)$ *and* $\boldsymbol{m}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$ *with positive entries, the Bhattacharyya coefficient is defined as*

$$\boldsymbol{B}_{(\boldsymbol{m}_1,\boldsymbol{m}_2)} = \sum_{j=1}^{m} \sqrt{\varsigma_j \sigma_j}. \tag{50}$$

Corollary 13. *Let* $\boldsymbol{m}_1 = (\varsigma_1, \varsigma_2, \cdots, \varsigma_m)$, $\boldsymbol{m}_2 = (\sigma_1, \sigma_2, \cdots, \sigma_m)$, *and* $\boldsymbol{m}_3 = (\varrho_1, \varrho_2, \cdots, \varrho_m)$ *be positive probability distributions and* $q > 1$. *Then,*

$$\begin{aligned} \boldsymbol{B}_{(\boldsymbol{m}_1,\boldsymbol{m}_2)} - \boldsymbol{B}_{(\boldsymbol{m}_1,\boldsymbol{m}_3)} \quad \leq \quad & \frac{1}{4} \sum_{j=1}^{m} \frac{(\sigma_j - \varrho_j)^2}{\varsigma_j} \left(\frac{(q+1)\left(\frac{\varsigma_j}{\sigma_j}\right)^{\frac{3q}{2}} + \left(\frac{\varsigma_j}{\varrho_j}\right)^{\frac{3q}{2}}}{(q+1)(q+2)} \right)^{\frac{1}{q}} \\ & + \frac{1}{2} \sum_{j=1}^{m} (\sigma_j - \varrho_j) \sqrt{\frac{\varsigma_j}{\sigma_j}}. \end{aligned} \tag{51}$$

Proof. Consider the function $\psi(x) = -\sqrt{x}$ defined on $(0, \infty)$. Then, $\psi''(x) = \frac{1}{4} x^{-\frac{3}{2}}$ and $\left(|\psi''(x)|^q\right)'' = \left(\frac{1}{4}\right)^q \frac{3q}{2} \left(\frac{3q}{2} - 1\right) x^{-\left(\frac{3q}{2} - 2\right)}$. Clearly, both ψ'' and $\left(|\psi''|^q\right)''$ are positive, which substantiates the convexity of ψ and $|\psi''|^q$. Therefore, using (45) for $\psi(x) = -\sqrt{x}$, we achieve (51). □

Remark 12. *In a similar fashion, the applications of Theorem 2 to Theorem 5 can be provided for the Csiszár divergence and its related cases.*

5. Conclusions

Because of the many potential uses and rich history of mathematical inequalities, this topic has very dynamic characteristics in all areas of science. There are a number of inequalities that have been established for convex functions. The majorization inequality is one of the dominant inequalities, which has also been demonstrated by the support of the convex function. In this paper, we focused on finding bounds for the classical majorization inequality and its weighted forms. The intended bounds have been established by applying the well-known Jensen's inequality, Hölder inequality, power mean inequality, and notion of convexity. Bounds for the Jensen difference and Slater difference are also provided as consequences of the main findings. Moreover, applications of the acquired results are also presented in information theory. The presented applications provided different estimations for the Csiszár and Kullback–Leibler divergences, Bhattacharyya coefficient, and Shannon entropy. The approach and techniques used in this manuscript may be utilized for the integral majorization inequality. Also, the idea may be further generalized for higher-order differentiable functions.

Author Contributions: Conceptualization, A.B., M.A.K., H.U., Y.A., S.C. and T.S.; funding acquisition, S.C. and T.S.; investigation, A.B., M.A.K. and S.C.; methodology, H.U., Y.A. and T.S.; validation, A.B., M.A.K. and H.U.; visualization, Y.A. and S.C.; writing—original draft, A.B., M.A.K., H.U., Y.A. and S.C.; writing—review and editing, A.B., M.A.K., H.U., S.C. and T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science, Research and Innovation Fund (NSRF) and King Mongkut's University of Technology, North Bangkok, with contract no. KMUTNB-FF-66-54.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through a large-group research project under grant number RGP2/366/44. We give thanks to worthy referees for their valuable comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gomez, I.S.; da Costa, B.G.; dos Santos, M.A. Majorization and dynamics of continuous distributions. *Entropy* **2019**, *590*, 21. [CrossRef]
2. Cloud, M.J.; Drachman, B.C.; Lebedev, L.P. *Inequalities with Applications to Engineering*; Springer: Cham, Switzerland; Berlin/Heidelberg, Germany; New York, NY, USA; Dordrecht, The Netherlands; London, UK , 2014.
3. Grinalatt, M.; Linnainmaa, J.T. Jensen's inquality, parameter uncertainty, and multiperiod investment. *Rev. Asset Pricing Stud.* **2001**, *1*, 1–34. [CrossRef]
4. Ullah, H.; Adil Khan, M.; Pečarić, J. New bounds for soft margin estimator via concavity of Gaussian weighting function. *Adv. Differ. Equ.* **2020**, *2020*, 644. [CrossRef]
5. Adil Khan, M.; Sohail, A.; Ullah, H.; Saeed, T. Estimations of the Jensen gap and their applications based on 6–convexity. *Mathematics* **2023**, *11*, 1957. [CrossRef]
6. Horváth, L.; Pečarić, Đ.; Pečarić, J. Estimations of f–and Rényi divergences by using a cyclic refinement of the Jensen's inequality. *Bull. Malays. Math. Sci. Soc.* **2019**, *42*, 933–946. [CrossRef]
7. Borwein, J.; Lewis, A. *Convex Analysis and Nonlinear Optimization, Theory and Examples*; Springer: New York, NY, USA, 2000.
8. Zhao, T.-H.; Wang, M.-K.; Chu, Y.-M. Concavity and bounds involving generalized elliptic integral of the first kind. *J. Math. Inequal.* **2021**, *15*, 701–724. [CrossRef]
9. Marshall, A.W.; Olkin, I.; Arnold, B.C. *Inequalities: Theory of Majorization and Its Applications*, 2nd ed.; Springer Series in Statistics; Springer: New York, NY, USA, 2011.
10. Lakshmikantham, V.; Vatsala, A.S. *Theory of Differential and Integral Inequalities with Initial Time Difference and Applications*; Springer: Berlin/Heidelberg, Germany, 1999.
11. Dragomir, S.S.; Pearce, E.E.M. *Selected Topics on Hermite–Hadamard Inequalities and Applications*; Victoria University: Melbourne, Australia, 2000.
12. Adil Khan, M.; Khalid, S.; Pečarić, J. Refinements of some majorization type inequalities. *J. Math. Inequal.* **2013**, *7*, 73–92. [CrossRef]
13. Lin, M.; Wolkowicz, H. An eigenvalue majorization inequality for positive semidefinite block matrices. *Linear Multilinear Algebra* **2012**, *60*, 1365–1368. [CrossRef]
14. You, X.; Adil Khan, M.; Ullah, H.; Saeed, T. Improvements of Slater's inequality by means of 4-convexity and its applications. *Mathematics* **2022**, *10*, 1274. [CrossRef]
15. Zhao, T.-H.; Wang, M.-K.; Chu, Y.-M. Monotonicity and convexity involving generalized elliptic integral of the first kind. *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM* **2021**, *115*, 1–13. [CrossRef]
16. Chu, H.-H.; Zhao, T.-H.; Chu, Y.-M. Sharp bounds for the Toader mean of order 3 in terms of arithmetic, quadratic and contraharmonic means. *Math. Slovaca* **2020**, *70*, 1097–1112. [CrossRef]
17. Zhao, T.-H.; He, Z.-Y.; Chu, Y.-M. On some refinements for inequalities involving zero-balanced hypergeometric function. *AIMS Math.* **2020**, *5*, 6479–6495. [CrossRef]
18. Pečarić, J.; Persson, L.E.; Tong, Y.L. *Convex Functions, Partial Ordering and Statistical Applications*; Academic Press: New York, NY, USA, 1992.
19. Ullah, H.; Adil Khan, M.; Saeed, T.; Sayed, Z.M.M.M. Some improvements of Jensen's inequality via 4–convexity and applications. *J. Funct. Spaces* **2022**, *2022*, 1–9. [CrossRef]
20. Abdeljawad, T.; Rashid, S.; Hammouch, Z.; Chu, Y.-M. Some new local fractional inequalities associated with generalized (s, m)-onvex functions and applications. *Adv. Differ. Equ.* **2020**, *2020*, 406. [CrossRef]
21. Adeel, M.; Khan, K.A.; Pečarić, Đ.; Pečarić, J. Levinson type inequalities for higher order convex functions via Abel–Gontscharoff interpolation. *Adv. Differ. Equ.* **2019**, *2019*, 430. [CrossRef]

22. Zhao, T.-H. ; Wang, M.-K.; Chu, Y.-M. A sharp double inequality involving generalized complete elliptic integral of the first kind. *AIMS Math.* **2020**, *5*, 4512–4528. [CrossRef]
23. Adeel, M.; Khan, K.A.; Pečarić, Đ.; Pečarić, J. Estimation of f–divergence and Shannon entropy by Levinson type inequalities for higher–order convex functions via Taylor polynomial. *J. Math. Compt. Sci.* **2020**, *21*, 322–334. [CrossRef]
24. Niculescu, C.P.; Persson, L.E. *Convex Functions and Their Applications. A Contemporary Approach*; CMS Books in Mathematics; Springer: New York, NY, USA, 2006; Volume 23.
25. Sajjad, M.; Mehmood, I.; Abbasand, N.; Baik, S.W. Basis pursuit denoising-based image super resolution using a redundant set of atoms. *Signal Image Video Process.* **2016**, *10*, 181–188. [CrossRef]
26. Vasić, Ž. Mijajlović, P.M. On an index set function connected with Jensen inequality. *Univ. Beograd Publ. Elektroteh. Fak. Ser. Mat. Fiz.* **1976**, *544–576*, 110–112.
27. Vespignani, A.; Modelling dynamical processes in complex socio-technical systems. *Nat. Phys.* **2012**, *8*, 32–39. [CrossRef]
28. Kleiner, A.; Moldovanu, B.; Strack, P. Extreme points and majorization: Economic applications. *Econometrica* **2021**, *89*, 1557–1593. [CrossRef]
29. Wang, M.-K.; Hong, M.-Y.; Xu, Y.-F.; Shen, Z.-H.; Chu, Y.-M. Inequalities for generalized trigonometric and hyperbolic functions with one parameter. *J. Math. Inequal.* **2020**, *14*, 1–21. [CrossRef]
30. Rasheed, A.; Hussain, S.; Shah, S.G.A.; Darus, M.; Lodhi, S. Majorization problem for two subclasses of meromorphic functions associated with a convolution operator. *AIMS Math.* **2022**, *5*, 5157–5170. [CrossRef]
31. Kalsoom, H.; Rashid, S.; Idrees, M.; Safdar, F.; Akram, S.; Baleanu, D.; Chu, Y.-M. Post quantum integral inequalities of Hermite-Hadamard-type associated with co-ordinated higherorder generalized strongly pre-invex and quasi-pre-invex mappings. *Symmetry* **2020**, *12*, 443. [CrossRef]
32. Latif, M.A.; Dragomir, S.S.; Momoniat, E. Some Fejer type integral inequalities for geometrically–arithmetically–convex functions with applications. *Filomat* **2018**, *32*, 2193–2206. [CrossRef]
33. Zhao, T.-H.; Shi, L.; Chu, Y.-M. Convexity and concavity of the modified Bessel functions of the first kind with respect to Hölder means. *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM* **2020**, *114*, 1–14. [CrossRef]
34. Ibrahim, M.; Saeed, T.; Algehyne, E.A.; Khan, M.; Chu, Y.-M. The effects of L-shaped heat source in a quarter-tube enclosure filled with MHD nanofuid on heat transfer and irreversibilities, using LBM: Numerical data, optimization using neural network algorithm (ANN). *J. Therm. Anal. Calorim.* **2021**, *144*, 2435–2448. [CrossRef]
35. Chen, S.-B.; Rashid, S.; Noor, M.A.; Hammouch, Z.; Chu, Y.-M. New fractional approaches for n-polynomial P-convexity with applications in special function theory. *Adv. Differ. Equ.* **2020**, *2020*, 543. [CrossRef]
36. Niaz, T.; Khan, K.A.; Pečarić, J. On refinement of Jensen's inequality for 3–convex function at a point. *Turk. J. Inequal.* **2020**, *4*, 70–80.
37. Pečarić, J.E. A companion inequality to Jensen–Steffensen's inequality. *J. Approx. Theory* **1985**, *44*, 289–291. [CrossRef]
38. Ghadimi, S.; Lan, G.; Zhang, H. Generalized uniformly optimal methods for non-linear programming. *J. Sci. Comput.* **2019**, *79*, 1854–1881. [CrossRef]
39. Mercer, A. McD. A Variant of Jensen's Inequality. *J. Inequal. Pure Appl. Math.* **2003**, *4*, 73.
40. Maligranda, L.; Pečarić, J.; Persson, L.E. Weighted Favard and Berwald inequalities. *J. Math. Anal. Appl.* **1995**, *190*, 248–262. [CrossRef]
41. Ullah, H.; Adil Khan, M.; Saeed, T. Determination of bounds for the Jensen gap and its applications. *Mathematics* **2021**, *9*, 3132. [CrossRef]
42. Mićić, J.; Pečarić, J.; Jurica, P. Refined Jensen's operator inequality with condition on spectra. *Oper. Matrices* **2013**, *7*, 293–308.
43. Adil Khan, M.; Ullah, H.; Saeed, T. Some estimations of the Jensen difference and applications. *Math. Meth. Appl. Sci.* **2023**, *46*, 5863–5892. [CrossRef]
44. White, C.C.; Harrington, D.P. Application of Jensen's inequality to adaptive suboptimal design. *J. Optim. Theory Appl.* **1980**, *32*, 89–99. [CrossRef]
45. Viloria, J.M.; Vivas-Cortez, M. Jensen's inequality for convex functions on N–coordinates. *Appl. Math. Inf. Sci.* **2020**, *12*, 931–935. [CrossRef]
46. Deng, Y.; Ullah, H.; Adil Khan, M.; Iqbal, S.; Wu, S. Refinements of Jensen's inequality via majorization results with applications in the information theory. *J. Math.* **2021**, *2021*, 1951799. [CrossRef]
47. Slater, M.L. A companion inequality to Jensen's inequality. *J. Approx. Theory* **1981**, *32*, 160–166. [CrossRef]
48. Goel, A.; Meyerson, A.; Plotkin, S. Approximate majorization and fair online load balancing. In Proceedings of the 12th ACM-SIAM Symposium on Discrete Algorithms, San Francisco, CA, USA, 9–11 January 2000; SIAM: Philadelphia, PA, USA, 2000; pp. 384–390.
49. Barnett, N.S.; Cerone, P.; Dragomir, S.S. Majorisation inequalities for Stieltjes integrals. *Appl. Math. Lett.* **2009**, *22*, 416–421. [CrossRef]
50. Karamata, J. Sur une inégalité relative aux fonctions convexes. *Publ. Math. Univ. Belgrade* **1932**, *1*, 145–148.
51. Siddique, N.; Imran, M.; Khan, K.A.; Pečarić, J. Majorization inequalities via Green functions and Fink's identity with applications to Shannon entropy. *J. Inequal. Appl.* **2020**, *2020*, 192. [CrossRef]
52. Hardy, G.H.; Littlewood, J.E.; Pólya, G. *Inequalities*, 2nd ed.; Cambridge University Press: Cambridge, UK, 1952.
53. Fuchs, L. A new proof of an inequality of Hardy–Littlewood–Pólya. *Mat. Tidsskr. B* **1947**, *13*, 53–54.
54. Dragomir, S.S. Some majorisation type discrete inequalities for convex functions. *J. Math. Inequal. Appl.* **2004**, *7*, 207–216. [CrossRef]
55. Adil Khan, M.; Wu, S.-H.; Ullah, H.; Chu, Y.-M. Discrete majorization type inequalities for convex functions on rectangles. *J. Inequal. Appl.* **2019**, *2019*, 16. [CrossRef]

56. Wu, S.; Adil Khan, M.; Haleemzai, H.U. Refinements of majorization inequality involving convex functions via Taylor's theorem with mean value form of the remainder. *Mathematics* **2019**, *7*, 663. [CrossRef]
57. Saeed, T.; Adil Khan, M.; Ullah, H. Refinements of Jensen's inequality and applications. *AIMS Math.* **2022**, *7*, 5328–5346. [CrossRef]
58. Adil Khan, M.; Ullah, H.; Saeed, T.; Alsulami, H.H.; Sayed, Z.M.M.M.; Alshehri, A.M. Estimations of the slater gap via convexity and its applications in information theory. *Math. Probl. Eng.* **2022**, *2022*, 1750331. [CrossRef]
59. Benish, W.A. A review of the application of information theory to clinical diagnostic testing. *Entropy* **2020**, *22*, 97. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.