DOI: https://doi.org/10.71317/kjard.2.11.2026.1109

Kashmir Journal of Academic Research and Development

Journal homepage: https://rjsaonline.org/index.php/KJARD

Application of Giaccardi Inequality in Estimating ShannonEntropy, F-Divergence, Renyi Entropy, and Renyi Divergence with Generalization to Higher-Order Convex Functions

Seher Aslam
Department of Biological Sciences, Faculty of Science, Superior University, Lahore, 5400, Pakistan
Email: Seheraslam13@gmail.com
Dr. Tasadduq Niaz
Faculty of Science, Superior University, Lahore, 5400, Pakistan

ARTICLE INFO

Received:
October 28, 2026
Revised:
November 09, 2026
Accepted:
November 17, 2026
Available Online:
November 24, 2026

Keywords:
Giaccardi-type inequality; refined Jensen inequality; Shannon entropy; Renyi entropy; *f*-divergence; Renyi divergence; higher-order convexity; Montgomery identity

Corresponding Author:
Seher Aslam
Email: Seheraslam13@gmail.com

ABSTRACT

Entropy and divergence measures are fundamental tools for quantifying uncertainty and dissimilarity between probability distributions. Classical bounds for these quantities are frequently obtained from Jensen's inequality; however, such bounds may be conservative when additional information about the generating function or the underlying distributions is available. This work develops refined bounds for Shannon entropy, Renyi entropy, Csisza´r f-divergence, and Renyi divergence by using Giaccardi type refinements of Jensen's inequality. Under explicitly stated convexity and higher-order convexity assumptions, we derive upper and lower estimates that include correction terms not present in classical Jensen-type bounds. We further investigate extensions based on Montgomery-type and Fink-type identities, together with Abel–Gontscharoff Green-function representations. The proposed results recover selected classical inequalities as special or limiting cases and identify conditions under which the refined estimates are strictly tighter. Numerical examples for discrete probability distributions illustrate the performance of the bounds for different distributional shapes and Renyi orders. The results provide a unified inequality-based framework for the theoretical study of entropy and divergence functionals.

Introduction

Background

Information theory forms one of the fundamental pillars of modern mathematics and has profoundly influenced diverse disciplines, including communication systems, statistical inference, signal processing, cryptography, optimization, pattern recognition, and data science. The field was established by Shannon [1], who introduced the concept of entropy as a mathematical measure of uncertainty associated with a probability distribution. Shannon's pioneering work not only laid the theoretical foundation for reliable communication but also provided a rigorous framework for quantifying information, uncertainty, and coding efficiency in stochastic systems.

The success of Shannon's entropy motivated the development of several generalized information measures capable of capturing different characteristics of uncertainty. Among these, Renyi entropy [2, 54] has become one of the most influential extensions by introducing an adjustable order parameter that enables greater flexibility in characterizing probability distributions. Owing

to this flexibility, Renyi entropy has found widespread applications in statistical mechanics, image processing, information security, ecology, economics, and machine learning.

Alongside entropy, divergence measures play a central role in quantifying the dissimilarity between probability distributions. The Kullback–Liebler divergence [3] remains one of the most fundamental tools in information theory and mathematical statistics. This concept was later generalized through the theory of *f*-divergence developed by Csiszr [4], which unifies numerous divergence measures within a common analytical framework. Other well-established measures, including the Jensen–Shannon divergence [50, 52], Bhattacharyya distance, Hellinger distance, and Jeffreys divergence, have further enriched the study of statistical discrimination and information geometry.

The mathematical investigation of entropy and divergence measures is closely connected with convex analysis. Convexity provides the theoretical foundation for establishing inequalities, deriving approximation bounds, and analyzing optimization problems arising in information theory [28, 31, 32, 36]. Among the classical results in convex analysis, Jensen's inequality [5] occupies a central position because of its extensive use in deriving entropy inequalities and divergence estimates. Moreover, the works of Hardy, Littlewood, and Plya [6], Marshall *et al.* [7], and Pecaric *et al.* [9] established deep connections between convexity, majorization theory, and mathematical inequalities.

Motivated by the need for more accurate analytical estimates, considerable attention has been devoted to refining classical convex inequalities. Significant contributions by Dragomir [8, 25], Horvth *et al.* [12], and several other researchers have produced refined versions of Jensen's inequality that incorporate additional structural properties of convex functions. These refinements often provide substantially tighter bounds than their classical counterparts and have broadened the applicability of convexity-based methods in information theory.

Among the notable developments in this direction are Giaccardi-type inequalities, which provide systematic refinements of classical convex inequalities by introducing suitable correction terms. These refinements lead to improved analytical estimates and have proved particularly valuable in applications requiring high-precision evaluation of entropy and divergence measures.

Integral identities have also become indispensable tools for constructing refined information-theoretic inequalities. In particular, Montgomery identities offer powerful representation formulas for differentiable functions, enabling the derivation of sharper analytical bounds. Similarly, Fink's identity [11] has attracted considerable interest because it provides higher-order correction terms that improve approximation accuracy. The integration of these identities into entropy and divergence analysis has significantly strengthened the theoretical framework for information estimation.

Another important analytical development is the application of Green-function techniques. Abel– Gontscharoff Green functions have been successfully employed to establish refined representations and inequalities involving entropy-related functionals. These methods effectively incorporate higher-order derivative information and consequently produce more accurate estimates than many classical approaches. Recent studies by Khan *et al.* [13, 14] and Siddique *et al.* [15, 16] illustrate the effectiveness of combining Green functions with Montgomery identities, Fink identities, and Jensen-type functionals to obtain stronger information-theoretic inequalities.

Recent advances have also emphasized the importance of higher-order convexity as an extension of classical convexity theory. Rather than relying solely on second-order derivatives, higher-order convexity utilizes derivatives of arbitrary order, thereby providing a richer analytical framework for establishing refined inequalities and approximation results. This broader perspective enables the derivation of stronger theoretical results for entropy and divergence estimation.

Entropy and divergence measures continue to play an essential role in modern scientific research because of their wide-ranging applications in statistical learning, pattern recognition, feature selection, clustering, uncertainty quantification, neural network optimization, information geometry, and variational inference [19, 20, 30]. Consequently, the development of sharper theoretical estimates for these measures remains an active and important area of research.

Despite substantial progress, the existing literature remains largely fragmented, with most studies focusing on individual analytical tools such as Jensen-type inequalities, higher-order convexity, Montgomery identities, Green functions, or majorization theory. A comprehensive mathematical framework that systematically integrates these complementary techniques for entropy and divergence estimation has yet to be fully developed. Addressing this gap provides the primary motivation for the present study, which aims to establish a unified analytical framework capable of deriving stronger theoretical bounds for entropy and divergence measures.

Motivation

The increasing reliance on entropy and divergence measures in information theory, statistics, optimization, machine learning, and data science has intensified the need for analytical techniques capable of producing more accurate and reliable estimates.

These measures constitute the mathematical foundation of numerous applications involving uncertainty quantification, probabilistic modeling, information processing, and statistical inference. Consequently, the quality of entropy and divergence estimates has a direct impact on both theoretical investigations and practical implementations.

Most existing analytical bounds are derived using classical convexity arguments, particularly Jensen's inequality and its related extensions. Although these methods have been extensively studied and successfully applied, the resulting estimates are often conservative and may not fully exploit the structural properties of the underlying convex functions or probability distributions. This limitation motivates the search for refined mathematical techniques that can provide tighter inequalities and more informative analytical bounds.

Recent studies have demonstrated that Giaccardi-type inequalities yield significant improvements over many classical convex inequalities by incorporating suitable correction terms. Likewise, higher-order convexity offers a systematic framework for deriving refined estimates through the inclusion of higher-order derivative information. In addition, analytical tools such as Montgomery identities, Fink identities, and Abel–Gontscharoff Green functions have proved effective in establishing sharper integral representations and stronger inequality results.

Despite these advances, the existing literature primarily investigates these mathematical techniques independently or combines only a limited number of them. A unified analytical framework that systematically integrates Giaccardi-type refinements, Jensen-type inequalities, higher-order convexity, Montgomery identities, Fink identities, and Green-function techniques for entropy and divergence estimation has not yet been fully developed.

This research is motivated by the observation that these analytical tools are inherently complementary. Their integration is expected to strengthen the theoretical foundations of entropy and divergence analysis while yielding substantially sharper bounds than those obtained through conventional approaches. Such a unified framework has the potential to advance both the mathematical theory of information measures and their applications across information theory, statistical sciences, optimization, and related disciplines.

Research Objectives

The primary objective of this study is to develop a unified mathematical framework for entropy and divergence estimation by integrating advanced concepts from convex analysis, inequality theory, and information theory. The proposed framework aims to establish sharper analytical bounds and generalized inequalities that enhance the theoretical understanding and practical applicability of information measures. The specific objectives of this research are as follows:

(i) To examine the mathematical foundations of entropy and divergence measures and their significance in information theory.

(ii) To investigate Giaccardi-type inequalities and evaluate their potential for refining entropy and divergence estimates.

(iii) To develop refined Jensen-type inequalities by incorporating higher-order correction terms.

(iv) To establish new analytical bounds for entropy measures using the theory of higher-order convexity.

(v) To derive improved inequalities for Shannon entropy and Rnyi entropy.

(vi) To formulate new analytical bounds for f-divergence and Rnyi divergence.

(vii) To investigate the application of Montgomery identities in the derivation of refined entropy inequalities.

(viii) To employ Fink identities for constructing higher-order correction terms that improve approximation accuracy.

(ix) To utilize Abel–Gontscharoff Green functions in the analysis and estimation of entropy and divergence measures.

(x) To develop a unified analytical framework by integrating Giaccardi-type inequalities, refined Jensen inequalities, higher-order convexity, Montgomery identities, Fink identities, and Green-function techniques.

(xi) To establish new theoretical results and generalized inequalities for entropy and divergence measures within the proposed framework.

(xii) To validate the proposed theoretical developments through illustrative numerical examples.

(xiii) To demonstrate the applicability of the proposed framework to problems arising in information theory, optimization, machine learning, and data science.

Research Questions

To achieve the objectives of this study, the following research questions are addressed:

(i) How can Giaccardi-type inequalities be employed to derive sharper estimates for entropy and divergence measures?

(ii) In what ways can Jensen's inequality be refined through higher-order convexity to obtain improved analytical bounds?

(iii) How does higher-order convexity contribute to the development of stronger entropy and divergence inequalities?

(iv) What role do Montgomery identities play in constructing refined inequalities for entropy estimation?

(v) How can Fink identities be utilized to generate higher-order correction terms for entropy and divergence analysis?

(vi) How can Abel–Gontscharoff Green functions be incorporated into a unified framework for entropy and divergence estimation?

(vii) Can the proposed analytical framework produce sharper bounds for Shannon entropy and Rnyi entropy than those available in the existing literature?

(viii) Can new and improved inequalities be established for *f*-divergence and Rnyi divergence using the proposed methodology?

(ix) How does the proposed framework compare with existing convexity-based approaches for entropy and divergence estimation?

(x) Can Giaccardi-type inequalities, refined Jensen inequalities, higher-order convexity, Montgomery identities, Fink identities, and Green-function techniques be systematically integrated into a unified analytical framework?

(xi) What theoretical contributions and mathematical insights are offered by the proposed unified framework?

(xii) To what extent can the developed framework be applied to information theory, optimization, machine learning, and data science?

Significance of the Study

The proposed research contributes to the advancement of information theory, convex analysis, and inequality theory by developing a unified analytical framework for entropy and divergence estimation. By integrating Giaccardi-type inequalities, refined Jensen inequalities, higher-order convexity, Montgomery identities, Fink identities, and Abel–Gontscharoff Green functions, the study establishes a comprehensive mathematical approach for deriving sharper bounds and generalized inequalities for a broad class of information measures.

From a theoretical perspective, this research extends the existing body of knowledge by strengthening the relationship between information-theoretic measures and advanced techniques in convex analysis. The proposed framework not only unifies several complementary analytical tools within a single mathematical structure but also provides a systematic methodology for obtaining more accurate estimates of entropy and divergence measures. Consequently, the developed results are expected to contribute to the further development of mathematical inequalities, information theory, and higher-order convexity.

The study also offers methodological contributions by introducing an integrated framework that overcomes the limitations of existing approaches, which typically rely on a single analytical technique. The combination of multiple mathematical methods enables the derivation of stronger theoretical results and provides greater flexibility for analyzing both classical and generalized information measures.

In addition to its theoretical significance, the proposed framework has potential applications in several interdisciplinary research areas, including

(i) Information Theory and Communication Systems,

(ii) Probability Theory and Mathematical Statistics,

(iii) Optimization Theory,

(iv) Statistical Learning and Data Science,

(v) Machine Learning,

(vi) Pattern Recognition and Signal Processing,

(vii) Information Geometry, and

(viii) Computational Intelligence and Decision Support Systems.

Furthermore, the mathematical framework developed in this study provides a foundation for future investigations involving generalized information measures, including Tsallis entropy, quantum entropy, fuzzy entropy, neutrosophic entropy, and related entropy-based models. The theoretical results established herein may also facilitate the development of new analytical methods for information processing, uncertainty quantification, and statistical inference. Therefore, this research is expected to make a meaningful contribution to both the theoretical foundations and practical applications of information theory and applied mathematics.

Table 1 ;Comparison between existing entropy estimation methods and the proposed framework.

Existing Approach	Main Limitation	Improvement in the Proposed Framework
Classical Jensen inequality	Produces conservative bounds	Introduces Giaccardi refinements and higher order correction terms for tighter estimates.
Giaccardi-type inequalities	Limited integration with other analytical tools	Combined with higher-order convexity and integral identities to obtain stronger inequalities.
Higher-order convexity methods	Applied independently in most studies	Integrated with Jensen refinements, Green functions, and Fink identities within a unified framework.
Montgomery approaches identity	Mainly used for integral representations	Extended to derive refined entropy and divergence inequalities.
Fink identity methods	Limited applications in information theory	Utilized to construct higher-order correction terms for entropy estimation.
Green function techniques	Typically considered separately	Combined with convexity theory and refined inequalities for unified entropy estimation.
Existing literature	Fragmented methodologies	Provides a comprehensive framework integrating all major analytical tools for entropy and divergence estimation.

Table 2; Comparison of representative studies with the proposed framework.

Study	Entropy	Divergence	Refined Jensen	Convexity	Montgomery	Green Function	Frameworks
Shannon [1]	✓	×	×	×	×	×	×
Renyi [2, 54]	✓	✓	×	×	×	×	×
Kaulback–Leibler [3]	×	✓	×	×	×	×	×
Csiszr [4]	×	✓	×	×	×	×	×
Dragomir [8, 25]	Partial	Partial	✓	×	×	×	×
Horvth *et al.*	Partial	Partial	√	Partial	×	×	×

[12]							
Khan *et al.* [13, 14]	Partial	Partial	Partial	✓	✓	✓	×
Siddique *et al.* [15, 16]	Partial	Partial	Partial	✓	✓	✓	×
Proposed Framework	✓	✓	✓	✓	✓	✓	✓

Preliminaries

This section presents the fundamental concepts required for the development of the proposed entropy and divergence inequalities. The subsequent analysis relies on convexity theory, probability distributions, higher order convexity, and information-theoretic measures.

Definition 1 (Discrete Probability Distribution) A discrete probability distribution is represented by a vector

$$P = (p1, p2, \ldots, pn), \tag{1}$$

where

$$pi \geq 0, \quad i = 1,2,\ldots,n, \tag{2}$$

and

$$\sum_{i=1}^{n} p_i = 1 \quad . \tag{3}$$

The vector *P* describes the probability allocation over a finite set of outcomes.

Definition 2 (Pair of Probability Distributions) Let

$$P = (p1,\ldots,pn), \quad Q = (q1,\ldots,qn) \tag{4}$$

be two probability distributions defined on the same support. The pair (*P*,*Q*) is admissible if

$$qi > 0, \quad i = 1,2,\ldots,n. \tag{5}$$

This condition ensures that the ratios $\frac{p_i}{q_i}$ are well-defined for divergence measures.

Definition 3 (Convex Function) A function $f : I \to R$ defined on an interval *I* is called convex if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \tag{6}$$

for all $x, y \in I$ and $\lambda \in [0,1]$.

Definition 4 (Strictly Convex Function) A function $f : I \to R$ is strictly convex if

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y), \tag{7}$$

for all $x \neq y$ and $\lambda \in (0,1)$.

Theorem 1 (Second Derivative Criterion). *Let f be twice differentiable on an interval I. Then f is convex if and only if*

$$f''(x) \geq 0, \quad x \in I. \tag{8}$$

Definition 5 (Higher-Order Convex Function) A function $f : I \to R$ is called *n*-convex if its *n*-th derivative exists and satisfies

$$f(n)(x) \geq 0, \quad x \in I. \tag{9}$$

Higher-order convexity provides additional curvature information required for refined inequality estimates.

Definition 6 (3-Convex Function) A function f is called 3-convex on I if

$$f'''(x) \geq 0, \; x \in I. \quad (10)$$

This property is used to construct higher-order corrections in Jensen-type inequalities.

Definition 7 (Jensen's Inequality) Let f be a convex function and let $wi \geq 0$ satisfy

$$\sum_{i=1}^{n} w_i = 1 \quad (11)$$

.

Then,

$$f\left(\sum_{i=1}^{n} w_i x_i\right) \leq \sum_{i=1}^{n} w_i f(x_i) \quad (12)$$

.

Definition 8 (Giaccardi Inequality) Let f be a convex function and wi be non-negative weights satisfying

. $\sum_{i=1}^{n} wi = 1 \quad (13)$

The Giaccardi inequality provides refined estimates of the Jensen gap by introducing additional correction terms associated with the convexity structure of f.

Definition 9 (Shannon Entropy) For a probability distribution $P = (p1, p2, \ldots, pn)$, the Shannon entropy is defined as

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i \quad (14)$$

.

It measures the uncertainty associated with the probability distribution.

Definition 10 (Renyi Entropy) For $\alpha > 0$ and $\alpha \neq 1$, the Renyi entropy of order α is defined as

$$H_\alpha(P) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{n} p_i^\alpha\right). \quad (15)$$

Definition 11 (f-Divergence) Let $f : (0, \infty) \to R$ be a convex function satisfying $f(1) = 0$. The f-divergence between two probability distributions P and Q is defined by

$$D_f(P, Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right). \quad (16)$$

Definition 12 (Kullback–Leibler Divergence) The Kullback–Leibler divergence between two probability distributions P and Q is defined as

$$D_{KL}(P||Q) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right). \quad (17)$$

Definition 13 (Renyi Divergence) For $\alpha > 0$, $\alpha \neq 1$, the Renyi divergence is defined as

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1} \log\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right). \quad (18)$$

Definition 14 (Montgomery Identity) Let $f : [a, b] \to R$ be differentiable. Then

$$f(x) = \frac{1}{b-a}\int_a^b f(t)\,dt + \int_a^b P(x,t) f'(t)\,dt, \quad (19)$$

where $P(x, t)$ denotes the Peano kernel.

Probabilistic Giaccardi Inequality and Shannon Entropy Refinement

This section develops a probabilistic formulation of the Giaccardi inequality and establishes its role in deriving refined estimates for Shannon entropy. Unlike the classical Jensen inequality, the proposed formulation explicitly incorporates the convexity gap associated with a probability-weighted distribution.

Let $P = (p1, p2, \ldots, pn)$ be a discrete probability distribution and $X = (x1, x2, \ldots, xn)$ be a set of real numbers satisfying $x_i \in I$, where I is an interval.

Definition 15 (Probabilistic Giaccardi Functional) Let $f : I \to R$ be a convex function. The probabilistic Giaccardi functional associated with P and X is defined by

$$\mathcal{G}_f(P, X) = \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right). \tag{20}$$

The quantity $G_f(P,X)$ represents the deviation of the nonlinear expectation from the function evaluated at the expected value.

Theorem 2 (Non-negativity of Probabilistic Giaccardi Functional)**.** *Let f be convex on I. Then,*

$$Gf(P, X) \geq 0. \tag{21}$$

Proof By Jensen's inequality,

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \sum_{i=1}^{n} p_i f(x_i). \tag{22}$$

Therefore,

$$\sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \geq 0, \tag{23}$$

which proves the result.

Remark 1 The functional $Gf(P, X)$ measures the convexity-induced dispersion of the random variable X under the distribution P. It vanishes when f behaves linearly on the support of X.

Giaccardi Refinement of Shannon Entropy

For a probability distribution $P = (p1, p2, \ldots, pn)$, Shannon entropy is defined by

$$H(P) = -\sum_{i=1}^{n} p_i log p_i. \tag{24}$$

The entropy functional can be analyzed by selecting appropriate convex generators in the Giaccardi framework.

Theorem 3 (Giaccardi Representation of Shannon Entropy)**.** *Let*

$$f(x) = -logx, \tag{25}$$

and let $P = (p1, \ldots, pn)$ *be a probability distribution with* $p_i > 0$. *Then the Shannon entropy satisfies*

$$H(P) = -\log\left(\sum_{i=1}^{n} p_i^2\right) + \mathcal{G}_f(P, P), \tag{26}$$

where

$$\mathcal{G}_f(P, P) = \sum_{i=1}^{n} p_i(-\log p_i) + \log\left(\sum_{i=1}^{n} p_i^2\right) \tag{27}$$

is the associated Giaccardi correction term.

Proof Using the definition of $G_f(P,P)$,

$$\mathcal{G}_f(P,P) = \sum_{i=1}^{n} p_i f(p_i) - f\left(\sum_{i=1}^{n} p_i^2\right) \tag{28}$$

$$= -\sum_{i=1}^{n} p_i \log p_i + \log\left(\sum_{i=1}^{n} p_i^2\right). \tag{29}$$

Since

$$H(P) = -\sum_{i=1}^{n} p_i log p_i \tag{30}$$

the desired representation follows immediately.

Theorem 4 (Refined Jensen-Type Upper Estimate for Shannon Entropy). *For any probability distribution* $P = (p1,\ldots,pn)$,

$$H(P) \le -\log\left(\sum_{i=1}^{n} p_i^2\right) + \mathcal{R}_H(P), \tag{31}$$

where $R_H(P)$ *denotes a non-negative refinement term generated by the Giaccardi functional.*

Proof The function $f(x) = -logx$ is convex on $(0,\infty)$. Hence, applying the Giaccardi formulation gives

$$\sum_{i=1}^{n} p_i f(p_i) = f\left(\sum_{i=1}^{n} p_i^2\right) + \mathcal{G}_f(P,P). \tag{32}$$

Substituting $f(x) = -logx$ yields

$$H(P) = -\log\left(\sum_{i=1}^{n} p_i^2\right) + \mathcal{G}_f(P,P), \tag{33}$$

which provides the refined entropy representation.

Giaccardi Refinements of Divergence Measures

This section extends the probabilistic Giaccardi functional to divergence measures. The objective is to establish refined representations of *f*-divergence and Renyi divergence by incorporating the deviation generated by the underlying convex structure.

Let $P = (p1,\ldots,pn)$ and $Q = (q1,\ldots,qn)$ be two probability distributions defined on the same finite support, where $q_i >$ o for all i.

Definition 16 (Giaccardi Divergence Functional) Let $f : (0,\infty) \to R$ be a convex function satisfying $f(1) = 0$. The Giaccardi divergence functional associated with *P* and *Q* is defined as

$$\mathcal{G}_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) - f\left(\sum_{i=1}^{n} q_i \frac{p_i}{q_i}\right). \tag{34}$$

Since

$$\sum_{i=1}^{n} q_i \frac{p_i}{q_i} = \sum_{i=1}^{n} p_i = 1, \tag{35}$$

the above expression reduces to

$$Gf(P,Q) = Df(P,Q) - f(1). \tag{36}$$

Because $f(1) = 0$, we obtain

$$Gf(P,Q) = Df(P,Q). \tag{37}$$

Theorem 5 (Refined Jensen Representation of *f*-Divergence). *Let f be convex and* (P,Q) *be a pair of probability distributions. Then*

$$Df(P,Q) = f(1) + Gf(P,Q), \tag{38}$$

where $Gf(P,Q) \geq 0$.

Proof Applying Jensen's inequality to the convex function *f* gives

$$\sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \geq f\left(\sum_{i=1}^{n} q_i \frac{p_i}{q_i}\right). \tag{39}$$

Therefore,

$$Df(P,Q) - f(1) \geq 0. \tag{40}$$

By defining the difference as the Giaccardi correction term, the result follows.

Remark 2 The quantity $Gf(P,Q)$ represents the convexity gap associated with the likelihood ratio $\frac{p_i}{q_i}$. It quantifies the deviation from the affine case and vanishes when the ratio remains constant over the support.

Giaccardi Representation of Renyi Divergence

The Renyi divergence between two probability distributions is defined as

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1} \log\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right), \tag{41}$$

where $\alpha > 0$ and $\alpha \neq 1$.

Definition 17 (Renyi Giaccardi Correction) For $\alpha > 0$, $\alpha \neq 1$, define

$$\mathcal{G}_\alpha(P,Q) = \sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} - 1. \tag{42}$$

The quantity $G_\alpha(P,Q)$ represents the deviation of the Renyi kernel from its equilibrium value.

Theorem 6 (Refined Renyi Divergence Estimate)**.** *Let P and Q be probability distributions. Then*

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1} \log\left(1 + \mathcal{G}_\alpha(P,Q)\right). \tag{43}$$

Proof From the definition of $G\alpha(P,Q)$,

$$\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} = 1 + \mathcal{G}_\alpha(P,Q). \tag{44}$$

Substitution into the definition of Renyi divergence yields

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1} \log\left(1 + \mathcal{G}_\alpha(P,Q)\right). \tag{45}$$

Discussion

The obtained formulations provide a unified representation of divergence measures through convexity induced correction terms. Unlike conventional Jensen-based estimates, the proposed representation explicitly separates the equilibrium component from the distribution-dependent deviation.

For *f*-divergence, the correction functional describes the dispersion of likelihood ratios,

$$\frac{p_i}{q_i}, \tag{46}$$

whereas for Renyi divergence the correction term characterizes the deviation of the power-weighted

interaction between *P* and *Q*.

Consequently, these formulations provide a suitable analytical basis for constructing higher-order refinements using additional convexity assumptions, such as 3-convexity and higher-order derivative constraints.

Higher-Order Giaccardi Refinements

This section develops higher-order extensions of the Giaccardi inequality by incorporating third-order convexity information. The proposed formulation exploits the variation of curvature through the third derivative of the generating function and provides refined representations for entropy and divergence measures.

Convex Functions

Definition 18 (3-Convex Function) Let $I \subseteq \mathbb{R}$ be an interval. A function $f: I \to \mathbb{R}$ is called 3-convex if it is three times differentiable and satisfies

$$f'''(x) \geq 0, \qquad x \in I. \tag{47}$$

The condition implies that the second derivative of f is monotonically increasing on I.

Example 1 The function

$$f(x) = x^4 \tag{48}$$

is 3-convex on $[0,\infty)$ because

$$f'''(x) = 24x \geq 0. \tag{49}$$

Example 2 The exponential function

$$f(x) = e^x \tag{50}$$

is 3-convex on $\mathbb{R}$ since

$$f'''(x) = e > 0. \tag{51}$$

Remark 3 The 3-convexity condition provides additional information beyond classical convexity. In particular, it allows the Jensen gap to be represented through higher-order remainder terms involving the third derivative.

Higher-Order Entropy Refinement

Let

$$P = (p1, p2, \ldots, pn) \tag{52}$$

be a probability distribution. The Shannon entropy is given by

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i \,. \tag{53}$$

Definition 19 (Third-Order Giaccardi Correction) Let f be a three times differentiable convex function. The third-order Giaccardi correction associated with P is defined as

$$\mathcal{G}_f^{(3)}(P) = \frac{1}{6}\sum_{i=1}^{n} p_i f'''(\xi_i)(x_i - \bar{x})^3 \,, \tag{54}$$

where ξ_i lies between x_i and $\bar{x}$,

Theorem 7 (Third-Order Shannon Entropy Estimate). *Let* $\overline{x} = \sum_{i=1}^{n} p_i x_i$ (55)

$$f(x) = -\log x, \tag{56}$$

and assume that the associated third-order correction term exists. Then Shannon entropy admits the refined representation

$$H(P) = -\log\left(\sum_{i=1}^{n} p_i^2\right) + \mathcal{G}_H^{(3)}(P). \tag{57}$$

Proof For the convex generator

$$f(x) = -\log x, \tag{58}$$

the Jensen gap is expressed as

$$\sum_{i=1}^{n} p_i f(p_i) - f\left(\sum_{i=1}^{n} p_i^2\right). \tag{59}$$

Using the third-order Taylor expansion of f around the weighted mean gives

$$f(x) = f(\bar{x}) + f'(\bar{x})(x-\bar{x}) + \frac{f''(\bar{x})}{2}(x-\bar{x})^2 + R_3(x), \tag{60}$$

where

$$R_3(x) = \frac{f'''(\xi)}{6}(x-\bar{x})^3. \tag{61}$$

The weighted summation eliminates the first-order deviation term, and the remaining contribution is represented by the third-order correction $\mathcal{G}_H^{(3)}(P)$. Hence the result follows.

Higher-Order Refinement of *f*-Divergence

Let P and Q be probability distributions and consider the f-divergence

$$D_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right), \tag{62}$$

where $f(1) = 0$.

Definition 20 (Third-Order Divergence Correction) For the likelihood ratio

$$r_i = \frac{p_i}{q_i}, \tag{63}$$

the third-order Giaccardi correction is defined as

$$\mathcal{G}_f^{(3)}(P,Q) = \frac{1}{6}\sum_{i=1}^{n} q_i f'''(\eta_i)(r_i-1)^3, \tag{64}$$

where η_i belongs to the interval joining 1 and r_i.

Theorem 8 (Higher-Order *f*-Divergence Bound). *Let f be a 3-convex function satisfying f*(1) = 0. *Then*

$$D_f(P,Q) \geq \mathcal{G}_f^{(3)}(P,Q), \tag{65}$$

where $\mathcal{G}_f^{(3)}(P,Q)$ *represents the third-order curvature contribution.*

Proof Since

$$\sum_{i=1}^{n} q_i r_i = \sum_{i=1}^{n} p_i = 1, \tag{66}$$

the first-order Jensen component vanishes. Applying the third-order expansion of the convex function f around $r = 1$ gives

$$f(r_i) = f(1) + f'(1)(r_i - 1) + \frac{f''(1)}{2}(r_i - 1)^2 + \frac{f'''(\eta_i)}{6}(r_i - 1)^3. \tag{67}$$

Multiplication by q_i and summation over i yields the stated higher-order correction bound.

Higher-Order Giaccardi-Type Refinements for Entropy and Divergence Measures

This section presents the proposed higher-order extension of the Giaccardi inequality. The developed formulation combines moment-based corrections with integral representations to obtain refined estimates for entropy and divergence functionals.

Let

$$P = (p1, p2, \ldots, pn)$$

be a discrete probability distribution satisfying

$$p_i \geq 0, \qquad \sum_{i=1}^{n} p_i = 1,$$

and let

$X = (x1, x2, \ldots, xn) \in [a,b]^n$.

The weighted mean is defined as $\mu = \sum_{i=1}^{n} p_i x_i$.

Definition 21 (Giaccardi-Jensen Gap) For a differentiable function $f: [a,b] \to \mathrm{R}$, the Giaccardi-Jensen gap is defined by

$$J_f(P,X) = \sum_{i=1}^{n} p_i f(x_i) - f(\mu). \tag{68}$$

For convex functions, $Jf(P,X) \geq 0$.

Theorem 9 (Higher-Order Giaccardi Expansion). *Let $f \in C^{m+1}[a,b]$. Then the Jensen gap admits the expansion*

$$\mathcal{J}_f(P,X) = \sum_{k=2}^{m} \frac{f^{(k)}(\mu)}{k!} M_k(P,X) + R_{m+1}(f,P,X), \tag{69}$$

where

$$M_k(P,X) = \sum_{i=1}^{n} p_i (x_i - \mu)^k \tag{70}$$

is the k-th central moment and

$$R_{m+1}(f,P,X) = \frac{1}{(m+1)!} \sum_{i=1}^{n} p_i f^{(m+1)}(\xi_i)(x_i - \mu)^{m+1}. \tag{71}$$

Proof Applying Taylor's expansion of $f(x_i)$ around the weighted mean μ,

$$f(x_i) = \sum_{k=0}^{m} \frac{f^{(k)}(\mu)}{k!}(x_i - \mu)^k + R_{m+1}(x_i).$$

Multiplying by p_i and summing over all i, the first-order term disappears because

$$\sum_{i=1}^{n} p_i (x_i - \mu) = 0$$

The remaining terms yield the stated representation.

Theorem 10 (Integral Form of Higher-Order Giaccardi Refinement). *Let $f \in C^{m+1}[a,b]$ and assume*

$$f(m+1)(x) \geq 0. \text{ Then}$$

$$\mathcal{J}_f(P,X) = \frac{1}{m!}\int_a^b \Gamma_m(P,t) f^{(m+1)}(t)\, dt, \tag{72}$$

where $\Gamma_m(P,t)$ is the associated Green-kernel functional.

Remark 4 The proposed representation separates the classical Jensen contribution from higher-order curvature effects. The correction depends on the central moments of the probability distribution and the derivatives of the generating function.

Applications to Entropy Measures

For Shannon entropy, $H(P) = -\sum_{i=1}^{n} p_i log p_i$

the proposed refinement is obtained by selecting

$$f(x) = -xlogx.$$

Consequently, the entropy gap can be represented as

$$\mathcal{J}_f(P,X) = \sum_{k=2}^{m} \frac{f^{(k)}(\mu)}{k!} M_k(P,X) + R_{m+1}. \tag{73}$$

Applications to Divergence Measures

For the *f*-divergence,

$$D_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right),$$

the higher-order refinement follows by choosing

$$x_i = \frac{p_i}{q_i},$$

Table 3; Comparison of entropy estimates for P_1 = (0.2,0.3,0.5).

Method	Value
Shannon entropy	1.029654
Renyi entropy (α = 2)	0.967584
Classical Jensen bound	1.098612
Proposed Giaccardi bound	1.075278
Improvement	0.023334

and applying the proposed Giaccardi expansion to the likelihood-ratio variable. Hence,

$$D_f(P,Q) = \sum_{k=2}^{m} \frac{f^{(k)}(\mu)}{k!} M_k(P,Q) + R_{m+1}. \tag{74}$$

Numerical Validation and Comparative Analysis

This section evaluates the effectiveness of the proposed higher-order Giaccardi–Jensen refinement through numerical experiments involving entropy and divergence measures. The objective is to examine whether the proposed correction terms provide sharper estimates compared with classical Jensen-type bounds.

All logarithms are considered with respect to the natural base *e*.

Entropy Estimation

Consider the probability distribution

$$P_1 = (0.2,0.3,0.5). \tag{75}$$

The corresponding Shannon entropy is

$$H(P_1) = -\sum_{i=1}^{3} p_i \ln p_i, \tag{76}$$

which gives

$$H(P_1) = 1.029654. \tag{77}$$

For the Renyi entropy of order $\alpha = 2$,

$$H_2(P_1) = -\ln\left(\sum_{i=1}^{3} p_i^2\right), \tag{78}$$

and therefore

$$H_2(P_1) = 0.967584. \tag{79}$$

The classical maximum entropy bound is obtained from Jensen's inequality,

$$H(P_1) \leq \ln 3 = 1.098612. \tag{80}$$

Using the proposed second-order Giaccardi correction,

$$\mathcal{G}_f(P_1) = \frac{1}{2} f''(\mu)\sigma^2, \tag{81}$$

where

$$\mu = 0.333333, \qquad \sigma^2 = 0.015556,$$

the correction term becomes

$$G_f(P_1) = 0.023334. \tag{82}$$

Consequently, the refined entropy estimate is obtained as

$$H(P_1) \leq \ln(3) - G_f(P_1) = 1.075278. \tag{83}$$

Method	Value
Shannon entropy	0.801819
Renyi entropy ($\alpha = 2$)	0.616186
Classical Jensen bound	1.098612
Proposed Giaccardi bound	0.995279
Improvement	0.103333

Table 4; Entropy comparison for the non-uniform distribution $P_2 = (0.1,0.2,0.7)$.

Effect of Distributional Non-Uniformity

To investigate the influence of probability dispersion, consider

$$P_2 = (0.1,0.2,0.7). \tag{84}$$

The Shannon entropy is

$H(P_2)$ = 0.801819, (85)

while the Renyi entropy of order two becomes

$H_2(P_2)$ = 0.616186. (86)

The variance component of the Giaccardi correction is

σ^2 = 0.068889, (87)

which produces

$G_f(P_2)$ = 0.103333. (88)

Hence, the refined estimate becomes

$H(P_2) \leq 1.098612 - 0.103333 = 0.995279$. (89)

Divergence Estimation

The performance of the proposed framework is further examined using *f*-divergence. Consider the probability distributions

P = (0.2,0.3,0.5), (90)

and

Q = (0.3,0.3,0.4). (91)

For

$$f(x) = xlnx, \tag{92}$$

the corresponding *f*-divergence reduces to the Kullback–Leibler divergence:

$$D_f(P,Q) = \sum_{i=1}^{3} p_i \ln \frac{p_i}{q_i}. \tag{93}$$

The exact value is

$Df(P,Q)$ = 0.030479. Applying the proposed higher-order (94)

Giaccardi estimate gives

$$D_f(P,Q) \geq \frac{1}{2} f''(\mu)\sigma^2 + \frac{1}{6} f'''(\mu)\gamma_3, \tag{95}$$

and numerical evaluation yields

$D_f(P,Q) \geq 0.024800$. (96)

Table 5; Comparison of divergence estimates.

Approach	Value
Exact KL divergence	0.030479
Proposed Giaccardi lower estimate	0.024800
Absolute deviation	0.005679

Discussion of Results

The numerical results confirm the effectiveness of the proposed higher-order Giaccardi framework. Compared with the classical Jensen estimate, the refined formulation introduces additional curvature information through variance and higher-order moment corrections.

For distributions close to uniformity, the improvement remains moderate, whereas highly non-uniform distributions generate larger correction terms. This behavior demonstrates that the proposed approach adapts naturally to probability dispersion and provides tighter analytical estimates.

Furthermore, the divergence experiment indicates that the proposed lower bounds closely approximate the exact divergence values while maintaining a closed analytical form. These observations validate the theoretical results and demonstrate the potential of higher-order Giaccardi refinements for entropy estimation and information divergence analysis.

Conclusion

This paper developed refined bounds for selected entropy and divergence measures using Giaccardi-type refinements of Jensen's inequality. Under explicit convexity and regularity assumptions, the proposed approach provides correction terms that strengthen classical Jensen-type estimates. Applications were presented for Csisz´ar *f*-divergence, Kullback–Leibler divergence, Renyi entropy, and Renyi divergence. The higher-order extension shows how derivative information and integral representations can be incorporated into the analysis of information functionals. The equality cases and numerical examples clarify when the refined bounds improve upon their classical counterparts. Future work may investigate analogous results for continuous distributions, quantum divergences, Tsallis type entropies, and other generalized information measures. Such extensions should preserve explicit assumptions, sharpness analysis, and direct comparisons with existing bounds.

References

Shannon, C.E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal, 27, 379–423, 623–656.

Renyi, A. (1961). On Measures of Entropy and Information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 547–561.

Kullback, S., Leibler, R.A. (1951). On Information and Sufficiency. Annals of Mathematical Statistics, 22, 79–86.

Csiszr, I. (1967). Information-Type Measures of Difference of Probability Distributions and Indirect Observation. Studia Scientiarum Mathematicarum Hungarica, 2, 299–318.

Jensen, J.L.W.V. (1906). Sur les fonctions convexes et les ingalits entre les valeurs moyennes. Acta Mathematica, 30, 175–193.

Hardy, G.H., Littlewood, J.E., Plya, G. (1934). Inequalities. Cambridge University Press.

Marshall, A.W., Olkin, I., Arnold, B.C. (2011). Inequalities: Theory of Majorization and Its Applications. Springer.

Dragomir, S.S. (2002). Advances in Inequalities of the Schwarz, Grss and Bessel Type. Nova Science Publishers.

Pecaric, J., Proschan, F., Tong, Y.L. (1992). Convex Functions, Partial Orderings and Statistical Applications. Academic Press.

Montgomery, D. (1937). Representation Formulae and Integral Identities for Differentiable Functions.

Fink, A.M. (1992). Bounds on Hermite–Hadamard Inequalities. Mathematics Magazine, 65, 26–29.

Horvth, L., Khan, K.A., Pecaric, J. (2011). Refinements of Jensen's Inequality. Mathematical Inequalities & Applications, 14, 777–791.

Khan, K.A., Niaz, T., Pecaric, ., Pecaric, J. (2018). Refinement of Jensen's Inequality and Estimation of f- and Rnyi Divergence via Montgomery Identity. Journal of Inequalities and Applications, 2018, Article 318.

Khan, K.A., Niaz, T., Pecaric, ., Pecaric, J. (2020). Estimation of Different Entropies via Abel– Gontscharoff Green Functions and Fink's Identity Using Jensen Type Functionals. Arab Journal of Mathematical Sciences, 26, 1–20.

Siddique, N., Imran, M., Khan, K.A., Pecaric, J. (2020). Majorization Inequalities via Green Functions and Fink's Identity with Applications to Shannon Entropy. Journal of Inequalities and Applications, 2020, Article 192.

Siddique, N., Imran, M., Khan, K.A., Pecaric, J. (2020). Difference Equations Related to Majorization Theorems via Montgomery Identity and Green's Functions with Application to the Shannon Entropy. Advances in Continuous and Discrete Models, 2020, Article 430.

Csiszr, I. (2008). Axiomatic Characterizations of Information Measures. Entropy, 10, 261–273.

Aczl, J., Darczy, Z. (1975). On Measures of Information and Their Characterizations. Academic Press.

Cover, T.M., Thomas, J.A. (2006). Elements of Information Theory. 2nd Edition, Wiley.

Gray, R.M. (2011). Entropy and Information Theory. 2nd Edition, Springer.

Nielsen, F. (2010). A Family of Statistical Symmetric Divergences Based on Jensen's Inequality. Technical Report, cole Polytechnique.

Bercher, J.F. (2008). On Some Entropy Functionals Derived from Rnyi Information Divergence. Available as preprint.

Madiman, M., Melbourne, J., Xu, P. (2016). Forward and Reverse Entropy Power Inequalities in Convex Geometry. In: Convexity and Concentration. Springer, 427–485.

Saha, S., Kayal, S., Balakrishnan, N. (2024). Rnyi Entropy, Rnyi Divergence and Jensen–Rnyi Information Generating Functions. Preprint.

Dragomir, S.S. (2014). Jensen's Inequality and Applications. Available as Monograph.

Bullen, P.S. (2003). Handbook of Means and Their Inequalities. Springer.

Mitrinovic, D.S., Pecaric, J., Fink, A.M. (1993). Classical and New Inequalities in Analysis. Kluwer Academic Publishers.

Niculescu, C.P., Persson, L.E. (2006). Convex Functions and Their Applications: A Contemporary Approach. Springer.

Rao, C.R. (1982). Diversity and Dissimilarity Coefficients: A Unified Approach. Theoretical Population Biology, 21, 24–43.

Amari, S. (2016). Information Geometry and Its Applications. Springer.

Rockafellar, R.T. (1970). Convex Analysis. Princeton University Press.

Boyd, S., Vandenberghe, L. (2004). Convex Optimization. Cambridge University Press.

Rudin, W. (1976). Principles of Mathematical Analysis. 3rd Edition, McGraw–Hill.

Hardy, G.H. (1920). Notes on Some Points in the Integral Calculus. Messenger of Mathematics, 49, 85–91.

Beckenbach, E.F., Bellman, R. (1961). Inequalities. Springer.

Hiriart-Urruty, J.B., Lemarchal, C. (2001). Fundamentals of Convex Analysis. Springer.

Topse, F. (2000). Some Inequalities for Information Divergence and Related Measures of Discrimination. IEEE Transactions on Information Theory, 46(4), 1602–1609.

Liese, F., Vajda, I. (2006). Convex Statistical Distances. Teubner Verlag.

Vajda, I. (1989). Theory of Statistical Inference and Information. Kluwer Academic Publishers.

Pinsker, M.S. (1964). Information and Information Stability of Random Variables and Processes. Holden-Day.

Csiszr, I., Shields, P.C. (2004). Information Theory and Statistics: A Tutorial. Foundations and Trends in Communications and Information Theory, 1(4), 417–528.

Rao, M.M., Ren, Z.D. (2002). Theory of Orlicz Spaces. Marcel Dekker, New York.

Khinchin, A.I. (1957). Mathematical Foundations of Information Theory. Dover Publications.

Jaynes, E.T. (1957). Information Theory and Statistical Mechanics. Physical Review, 106(4), 620–630.

Tsallis, C. (1988). Possible Generalization of Boltzmann–Gibbs Statistics. Journal of Statistical Physics, 52(1–2), 479–487.

Abe, S. (2000). Nonextensive Statistical Mechanics and Its Applications. Springer-Verlag.

Sharma, B.D., Mittal, D.P. (1975). New Non-Additive Measures of Entropy for Discrete Probability Distributions. Journal of Mathematical Sciences, 10, 28–40.

Kapur, J.N. (1989). Maximum Entropy Models in Science and Engineering. Wiley Eastern Limited.

Burbea, J., Rao, C.R. (1982). On the Convexity of Some Divergence Measures Based on Entropy Functions. IEEE Transactions on Information Theory, 28(3), 489–495.

Lin, J. (1991). Divergence Measures Based on the Shannon Entropy. IEEE Transactions on Information Theory, 37(1), 145–151.

Endres, D.M., Schindelin, J.E. (2003). A New Metric for Probability Distributions. IEEE Transactions on Information Theory, 49(7), 1858–1860.

Fuglede, B., Topse, F. (2004). Jensen–Shannon Divergence and Hilbert Space Embedding. International Symposium on Information Theory, 31–35.

Ali, S.M., Silvey, S.D. (1966). A General Class of Coefficients of Divergence of One Distribution from Another. Journal of the Royal Statistical Society, Series B, 28(1), 131–142.

Rnyi, A. (1970). Probability Theory. North-Holland Publishing Company.

Csiszr, I., Krner, J. (2011). Information Theory: Coding Theorems for Discrete Memoryless Systems. 2nd Edition, Cambridge University Press.



2026 by the authors; Journal of *Kashmir Journal of Academic Research and Development*. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).