entropy

MDPI

Article

Asymptotic Normality and Convergence Rates for Tsallis Entropy Estimators via Stabilization Techniques

Mehmet Sıddık Çadırcı [1] and Martin Singull [2,*]

[1] Department of Statistics, Faculty of Science, Cumhuriyet University, 58140 Sivas, Türkiye; msiddikcadirci@cumhuriyet.edu.tr
[2] Department of Mathematics, Linköping University, 581 83 Linköping, Sweden
* Correspondence: martin.singull@liu.se

Abstract

We study nearest-neighbor-based estimators of Tsallis entropy associated with Poisson and binomial point processes on general metric measure spaces. In this study, by combining existing stabilization methods with the validation of the estimator's local k-nearest-neighbor structure, we investigate nearest-neighbor-based Tsallis entropy estimators under Poisson and binomial distributed input data. Rather than proposing a new second-order Poincaré inequality, this paper details and clearly presents stabilization-based normal approximation bounds for Tsallis-type k-NN functionals. We establish asymptotic normality and derive explicit convergence rates for the Kolmogorov distance. Our analysis avoids explicit score-function decompositions and instead relies on flexible localizations of add-one costs, which simplify the treatment of higher-order terms. Under natural stabilization and moment conditions, the resulting bounds recover the classical normal approximation rates $s^{-1/2}$ and $n^{-1/2}$ and extend corresponding results for Shannon and Rényi entropy estimators. We further illustrate the scope of the framework through examples involving Tsallis entropy functionals, weighted k-NN Shannon entropy estimators. The examples provided highlight the benefits of stabilization-based normal approximations for non-parametric statistical inference in complex spatial and high-dimensional settings.

Keywords: Tsallis entropy; nearest-neighbor estimator; stabilization; Poisson point process; binomial point process; normal approximation; convergence rate

1. Introduction

Let $(\mathcal{X}, \mathcal{F}, d, Q)$ be a metric measure space, where Q is a σ-finite measure and $d\colon \mathcal{X} \times \mathcal{X} \to [0, \infty)$ is a metric. We represent P_s as a Poisson point process on $\mathcal{X}$ with density measure $\lambda_s = sQ$, with $s \geq 1$. When Q is a probability measure, let us also consider the binomial point process $\xi_n = \sum_{i=1}^{n} \delta_{X_i}$; here, $X_1, \ldots, X_n$ denote independent observations with joint distribution Q. This paper examines the asymptotic distributional properties of nearest-neighbor Tsallis entropy estimators based on these point processes and derives quantitative normal approximation limits for the associated stabilizing functionals.

Entropy and associated information measures play a central role in statistics, information theory, machine learning, and the analysis of complex systems [1,2]. This paper aims to develop a new framework for studying non-extensive entropy measures. Many systems of practical interest, such as those encountered in finance, hydrology, turbulence, and networked dynamical systems, exhibit non-extensive behavior that is poorly captured

Check for updates

Academic Editor: Michel Broniatowski

Received: 14 April 2026
Revised: 28 May 2026
Accepted: 30 May 2026
Published: 31 May 2026

Copyright: © 2026 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.

by purely logarithmic entropy measures. Tsallis entropy [3] is a parametric family of generalized entropies that lies between heavy-tailed and compactly supported regimes and has become a standard tool for studying non-equilibrium and long-range-dependent phenomena. New developments in machine learning provide additional motivation to explore generalized entropy measures beyond the classical Shannon framework. In the fields of reinforcement learning and human-guided reinforcement learning (RLHF), entropy-based regularization, which aims to control exploration, prevent early policy convergence, and mitigate over-optimization, is widely used. Furthermore, modern reward and loss distributions can exhibit heavy-tailed behavior; in such cases, classical KL-based information-theoretic tools may be less effective. Recently, studies on the limits of tail-aware information theory for RLHF and stochastic gradient Langevin dynamics have highlighted the importance of heavy-tailed rewards and losses in such settings [4]. In parallel, the Tsallis entropy framework is used to generalize maximum entropy reinforcement learning by introducing an entropic index that controls the exploration behavior of the learned policy [5]. Although this paper does not directly examine RLHF, these developments further encourage the statistical analysis of Tsallis-type entropy estimators under flexible distribution assumptions. Methodologically, nearest neighbor-based estimators form a flexible class of non-parametric entropy estimators that do not involve direct density estimation. Regarding Shannon entropy, this series of studies includes the Kozachenko–Leonenko estimator and its improvements [6], whereas for generalized entropies such as Rényi and Tsallis, several k-NN structures are proposed and evaluated empirically. Specifically, Tsallis-based goodness-of-fit tests using k-NN estimators are developed in [7]; there, the emphasis is on test statistics for multivariate generalized Gaussian and q-Gaussian distributions. Entropy-based tests for generalized Gaussian distributions based on Shannon entropy and the maximum entropy principle have been investigated in [8]. Supplementary Rényi-based goodness-of-fit tests are proposed in [9] for multivariate Student and Pearson type II distributions, based on the principles of maximum Rényi entropy and nearest neighbor Rényi entropy estimators.

Although these studies highlight the practical potential of Shannon, Tsallis, and Rényi entropy estimators for goodness-of-fit testing, their focus is primarily on consistency, mean-square convergence, and calibration of critical values via Monte Carlo methods, as well as consistency for a class of Rényi entropy estimators, examining the empirical behavior of the corresponding test statistics. At the same time, Refs. [7,8] present extensive simulation results for entropy-based tests under generalized Gaussian and related models.

The complex local dependency structure of nearest neighbor-based functionals creates a significant technical challenge by complicating the application of standard central limit theorems. It was first proposed in [10] for geometric functionals of random point sets and further developed in [11–14]; stabilization methods provide a powerful framework for capturing local dependence. In recent work, combining stabilization ideas with Malliavin–Stein methods and second-order Poincaré inequalities yields approximation bounds for a wide class of Poisson and binomial point processes [15–19]. Together, these advances have led to sharp central limit theorems for geometric statistics, such as the volumes, face counts, and Betti numbers of random complexes, as well as for functionals related to random graphs.

Contributions and relation to existing work. This paper does not present a new second-order Poincaré inequality or a new general stabilization theorem. Instead, the paper's contribution lies in adapting existing stabilization and Malliavin–Stein normal approximation tools to the nearest-neighbor Tsallis entropy estimate and in clearly outlining the necessary assumptions and error bounds within this framework. First, we reformulate the Poisson and binomial normal approximation limits using a representation adapted

to the nearest-neighbor entropy functionals. The parameters $\Gamma_1, \ldots, \Gamma_6$ and $\Theta_{K,s}, \Theta_{K,n}$ are explicitly defined; thus, the resulting bounds can be directly controlled via "add-one" and second-order "add-one" costs. Second, we verify the basic stabilizing components of the Tsallis k-NN estimator under standard density assumptions. Specifically, it is shown that the local k-NN score has exponential stabilization properties and satisfies the required uniform L^p-moment control for compact support, bounded density away from zero and infinity, fixed k, fixed entropy index α, and the acceptability condition $p(1-\alpha)_+ < k$ that it satisfies the required uniform L^p-moment control. Third, to clarify the statistical role of the entropy index α, we derive the estimator from the following identity:

$$T_\alpha(f) = \frac{1}{1-\alpha}\Big(\mathbb{E}\{f(X)^{\alpha-1}\} - 1\Big),$$

We derive the estimator and discuss its relationship with bias, consistency, and the existing nearest-neighbor Shannon, Rényi, and Tsallis entropy estimators. Finally, we briefly discuss how this same stabilization perspective relates to weighted k-NN Shannon entropy estimators, Euler characteristic functionals, and minimum spanning tree statistics. The purpose of these examples is to illustrate the scope of the methodology rather than to present new general stabilization results for each case. Therefore, the innovative aspect of this paper is not to replace existing stabilization theory, but to present a validation method specific to the Tsallis model and to provide a transparent, standard approach for estimating nearest-neighbor entropy.

Relation to goodness-of-fit tests. The current results provide a probabilistic complement to entropy-based goodness-of-fit tests for generalized Gaussian and related models. In the study [8], the authors developed k-NN Shannon entropy estimators and tests derived from maximum entropy principles based on generalized Gaussian families, obtained the L^2-consistency of the entropy estimator, and performed extensive simulations on both empirical dimension and power. In [7], Tsallis entropy is employed to generate goodness-of-fit statistics for multivariate generalized Gaussian and q-Gaussian models, revealing that the Tsallis-based estimator performs effectively in scenarios involving heavy tails and non-Gaussian distributions. The Rényi-based procedures in [9] employ nearest-neighbor Rényi entropy estimators and maximum entropy characterizations for multivariate Student and Pearson type II distributions. Our work provides normal approximation results with explicit bounds for Tsallis and associated entropy estimators, providing an abstract probability basis that can be utilized to motivate Gaussian methods and to guide calibration of critical values in such test problems beyond bootstrap or fully simulation-based methods.

The remainder of the paper is organized as follows. Section 2 presents the basic concepts of point processes, incremental operators, and stabilization, and formulates the basic assumptions employed throughout the paper. Section 3 provides general normal approximation outcomes for Poisson and binomial inputs, accompanied by a clear bound outcome. Section 4 extends these results by applying Tsallis entropy estimators and corresponding geometric statistics; we also summarize a simulation strategy and typical summary statistics. Section 5 provides proofs for the principal theorems and results. Section 6 provides a brief overview of potential extensions.

2. Preliminary Information

2.1. Notation

For this article, P_s indicates a Poisson point process with intensity measure $\lambda_s = sQ$, while ξ_n represents a binomial point process consisting of n independent observations with joint distribution Q. The general Poisson and binomial functionals are denoted by F_s and F_n, respectively. The nearest-neighbor Tsallis entropy estimators are designated separately

as $\widehat{T}_{\alpha,s}$ and $\widehat{T}_{\alpha,n}$; here, $\alpha \in (0,\infty) \setminus \{1\}$ is the constant Tsallis entropy index. The constants c_Q and β_Q have been chosen for the metric growth condition on Q. In this section, the fundamental point process framework is recalled, the increment and cost operators are described, and the stabilization and moment assumptions underlying our main results are stated. Keeping the presentation concise, We emphasize the elements necessary for establishing stabilization and the resulting normal approach limits.

2.2. Point Processes and Functionals

Suppose $\mathcal{X}, \mathcal{F}, d, Q$ is a metric measure space with a σ-finite measure Q. Let $\mathcal{X}$ be equipped with the smallest σ-algebra that makes the following mappings

$$m_A \colon \mathcal{N} \to \mathbb{N} \cup \{0,\infty\}, \qquad m_A(M) = M(A), \tag{1}$$

measurable for every $A \in \mathcal{F}$. For a point process η on $\mathcal{X}$, η takes values in $\mathcal{N}$. We denote the space of η's square integrable measurable functionals satisfying the condition $\mathbf{E}[F(\eta)^2] < \infty$ by $F \colon \mathcal{N} \to \mathbb{R}$ as $L^2_\eta(\mathcal{X})$.

Definition 1 (Poisson point process [20]). *A Poisson point process $P(\lambda)$ on $\mathcal{X}$ having intensity measure λ satisfies the following conditions:*

1. *Given any measurable set $B \in \mathcal{F}$, the random variable $P(\lambda)(B)$ exhibits a Poisson distribution having parameter $\lambda(B)$.*
2. *The random variables $P(\lambda)(B_1), \ldots, P(\lambda)(B_m)$ are independent for pairwise disjoint sets $B_1, \ldots, B_m \in \mathcal{F}$.*

Definition 2 (Binomial point process [21]). *Consider a probability measure Q on $\mathcal{X}, \mathcal{F}$ and $n \in \mathbb{N}$. We define the binomial point process ξ_n as follows:*

$$\xi_n = \sum_{i=1}^{n} \delta_{X_i}, \tag{2}$$

where $X_1, \ldots, X_n$ follow an i.i.d. Q law, and δ_x represents the Dirac measure at $x \in \mathcal{X}$.

For two real-valued random variables Y and Z, the Kolmogorov measure is defined as

$$d_K(Y,Z) = \sup_{t\in\mathbb{R}} \big|\mathbb{P}(Y \le t) - \mathbb{P}(Z \le t)\big|. \tag{3}$$

Moreover, we will measure this distance to assess convergence to the standard normal limit.

2.3. Increment Operators and Stabilization

Methods of stabilization are based on the behavior of functionals under local modifications to the base-point configuration. Increment operators formalize this idea.

Definition 3 (Increment operators [16]). *Suppose $F \colon \mathcal{N} \to \mathbb{R}$ is a measurable functional and η is a point process on $\mathcal{X}$. For $x \in \mathcal{X}$, the (first-order) increment operators are as follows:*

$$\Delta_x F(\eta) := F(\eta \cup \{x\}) - F(\eta). \tag{4}$$

For different $x, y \in \mathcal{X}$, the second-order increment operator can be expressed as

$$\Delta_{x,y} F(\eta) := F(\eta \cup \{x,y\}) - F(\eta \cup \{x\}) - F(\eta \cup \{y\}) + F(\eta).$$

We call a function stable if it becomes insensitive to large changes in the configuration.

Definition 4 (Score-based stability [12]). *Consider a measurable score function $f : \mathcal{X} \times \mathcal{N} \to \mathbb{R}$. Then, for $x \in \mathcal{X}$, f is said to be stabilized at x if, for every finite set $A \subset \mathcal{X} \setminus B_x(R_x)$, there exists an almost surely finite random radius $R_x > 0$ such that*

$$f\big(x, (\eta \cap B_x(R_x)) \cup A\big) = f\big(x, \eta \cap B_x(R_x)\big), \tag{5}$$

where $B_x(R_x)$ is the ball of radius R_x centered at x. Score-based stabilization can be useful when it can be written as the sum of the local contributions of the functional. For the present work, it is convenient to rely on the strong stabilization properties expressed directly in terms of $\Delta_x F$ and $\Delta_{x,y} F$, which are more natural for nearest neighbor-based entropy estimators.

Proposition 1. *Suppose $F_s(P_s) = \sum_{x \in P_s} f_s(x, P_s)$ is a functional of a Poisson process P_s having an intensity of sQ. Then, for $x \in \mathcal{X}$,*

$$\Delta_x F_s(P_s) = f_s(x, P_s \cup \{x\}) + \sum_{y \in P_s} \Delta_x f_s(y, P_s). \tag{6}$$

A similar identity stands for functionals of binomial point processes.

Remark 1. *The strong stabilization of F_s suggests that score-based stabilization is reasonable for a reasonable choice of scores; however, the reverse may not hold. Working directly with $\Delta_x F_s$ and $\Delta_{x,y} F_s$ avoids explicit score decompositions and is often more robust for complex functionals such as Tsallis entropy estimators.*

2.4. Assumptions

In this section, the main structural assumptions on the underlying space and the stabilizing functionals considered are stated, adapted from [16,18,22].

Assumption 1 (Regularity of Q). *We assume the measure Q on $(\mathcal{X}, \mathcal{F})$ satisfies a condition of regular growth: there are constants $c_Q > 0$ and $\beta_Q > 1$ such that, for all $x \in \mathcal{X}$, all $r > 0$, and sufficiently small $\epsilon > 0$,*

$$\limsup_{\epsilon \to 0^+} \frac{Q(B_x(r+\epsilon)) - Q(B_x(r))}{\epsilon} \leq c_Q \beta_Q r^{\beta_Q - 1}.$$

Particularly, Q has a diffuse property in the sense that $Q(\{x\}) = 0$ for all $x \in \mathcal{X}$; see [16] [Lemma 5.1(a)].

Assumption 2 (Tail bound on stabilization radius.). *Assume that F_s is a strongly stabilizing function of P_s having a stabilization radius R_x at the point x. Assume there exist constants $C_1, C_2, \gamma > 0$ with the following inequality for all $r \geq 0$:*

$$\mathbb{P}(R_x \geq r) \leq C_1 \exp\big(-C_2 (s^{1/\beta_Q} r)^{\gamma}\big).$$

For binomial processes ξ_n, a similar bound is assumed using n instead of s.

Assumption 3 (Exponential spatial decay of add-one costs.). *Let $K \subseteq \mathcal{X}$ be a measurable set representing the active region of the functional F_s. For $x, y \in \mathcal{X}$, define*

$$d_s(x, K) := \inf_{z \in K} s^{1/\beta_Q} d(x, z), \qquad d_s(x, y) := s^{1/\beta_Q} d(x, y).$$

Assume that there exist constants $C_3, C_4, \delta > 0$ and a number $p \geq 5$ such that, for all $s \geq 1$ and all $x, y \in \mathcal{X}$,

$$\|\Delta_x F_s\|_{L^p} \leq C_3 \exp\left(-C_4 d_s(x, K)^\delta\right), \tag{7}$$

$$\left\|\Delta_{x,y} F_s\right\|_{L^p} \leq C_3 \exp\left(-C_4 \max\{d_s(x, y), d_s(x, K), d_s(y, K)\}^\delta\right). \tag{8}$$

The corresponding binomial version is obtained by replacing s with n and P_s with ξ_n.

Assumption 4 (Uniform moment condition.)**.** *Let $p \geq 5$ be as in Assumption 3. Assume that there exists a constant $M_p < \infty$ such that*

$$\sup_{s \geq 1} \sup_{x, y \in \mathcal{X}} \mathbf{E}\left[|\Delta_x F_s|^p + |\Delta_{x,y} F_s|^p\right] \leq M_p. \tag{9}$$

The same condition is assumed for the binomial functionals F_n, with n replacing s.

Remark 2. *Assumption 3 is not a tail bound with respect to a threshold variable u. Rather, it is a spatial localization condition controlling the size of the first- and second-order add-one costs as the inserted points move away from the active region K. Magnitude control of the add-one costs is handled through the uniform L^p-moment condition in Assumption 4, while the exponential factors in (7) and (8) quantify the spatial decay required for the stabilization-based normal approximation bounds.*

Remark 3. *These assumptions have been standard in the stabilization literature and are satisfied in a wide variety of geometric and topological settings; see [16,18,22]. Within the present context, these assumptions are tailored to ensure that the Tsallis entropy estimators considered in Section 4 allow exponential stabilization radius tails and sufficient moment control for the add-one costs, resulting in explicit normal approximation rates.*

3. Main Results

In this section, we present the normal approximation bounds for stabilizing functionals of Poisson and binomial point processes. Here, N represents a standard normal random variable.

3.1. Poisson Input

Suppose $F \in L^2_{P_s}(\mathcal{X})$ is a square-integrable functional of a Poisson point process P_s having intensity $\lambda = sQ$. Given a measurable subset $A_x \subseteq \mathcal{X}$ depending on x, the following are defined:

$$b_1(x, A_x) := \mathbf{E}\left|\Delta_x F - \Delta_x F(A_x)\right|^4,$$
$$b_2(x, A_x) := \mathbf{E}\left|\Delta_x F(A_x)\right|^4.$$

For a measurable subset $A_x \subset \mathcal{X}$, we define the localized add-one cost by

$$\Delta_x F(A_x) := F\big((\eta \cap A_x) \cup \{x\}\big) - F(\eta \cap A_x).$$

Given a Poisson functional F,

$$\sigma_F^2 := \mathrm{Var}(F) > 0, \qquad \lambda_s := sQ,$$

let us define

$$A_p(x) := \|\Delta_x F\|_{L^p}, \qquad B_p(x, y) := \|\Delta_{x,y} F\|_{L^p}.$$

Quantities appearing in the Poisson normal approximation limit can be written as follows: $\Gamma_i = \Gamma_i(F, \lambda_s)$, $i = 1, \ldots, 6$. They are the normalized first- and second-order Malliavin–Stein error terms related to the add-one and second-order add-one cost operators. Specifically, we note that they depend only on σ_F^2, $A_p(x)$, $B_p(x, y)$, and the density measure λ_s.

Theorem 1 (Normal approximation under Poisson input)**.** *Let F be as above and assume that*

$$\mathbf{E}[F^2] < \infty \qquad \text{and} \qquad \mathbf{E}\Big[\int_{\mathcal{X}} (\Delta_x F)^2 \, \lambda(dx)\Big] < \infty. \tag{10}$$

Then there is an absolute constant $C^* > 0$ such as

$$d_K\Big(\frac{F - \mathbf{E}[F]}{\sqrt{\mathbf{Var}(F)}}, N\Big) \leq \sum_{i=1}^{6} \Gamma_{i,s}(F), \tag{11}$$

where the quantities $\Gamma_{1,s}(F), \ldots, \Gamma_{6,s}(F)$ are the Malliavin–Stein error terms defined explicitly in Appendix A. These quantities are expressed in terms of the first-order difference operator $D_x F$, the second-order difference operator $D^2_{x,y} F$, the variance $\mathbf{Var}(F)$, and the intensity measure $\lambda_s = sQ$.

Remark 4. *Theorem 1 uses adaptable addition cost operators to localize increases without requiring a clear point decomposition. It can be seen as a modified improvement of the second-order Poincaré inequality in [15] and the boundary conditions in [16], which exhibit strong stabilization via an added cost functional.*

3.2. Binomial Input

Now, let us move on to functionals of binomial point processes. Suppose that for $n \geq 2$, ξ_n is a binomial point process guided by Q, and $F_n \in L^2_{\xi_n}(\mathcal{X})$ is a strongly stabilizing functional satisfying Assumptions 2–4.

Theorem 2 (Normal approximation under binomial input)**.** *Assume that F_n is strongly stabilizing with stabilization radius satisfying Assumptions 2–4. In that case, there is a constant $C_0' > 0$, based only on the constants in those assumptions, such that for all $n \geq 2$,*

$$d_K\Big(\frac{F_n - \mathbf{E}[F_n]}{\sqrt{\mathbf{Var}(F_n)}}, N\Big) \leq C_0'\left(\frac{\Theta_{K,n}^{1/2}}{\mathbf{Var}(F_n)} + \frac{\Theta_{K,n}}{\mathbf{Var}(F_n)^{3/2}} + \frac{\Theta_{K,n} + \Theta_{K,n}^{3/2}}{\mathbf{Var}(F_n)^2}\right), \tag{12}$$

where

$$\Theta_{K,n} := n \int_{\mathcal{X}} \exp\Big(-\frac{C_4(p-4)}{4p} \, (d_n(x, K))^2\Big) \, Q(dx)$$

and $d_n(x, K) := \inf_{z \in K} n^{1/\beta_Q} d(x, z)$. Here, β_Q expresses the metric growth parameter of the base measure Q, as described in Assumption 1.

Interpretation of $\Theta_{K,s}$ and $\Theta_{K,n}$

Quantities $\Theta_{K,s}$ and $\Theta_{K,n}$ represent the exponential decay of tail probabilities as a function of stabilization radii and insertion cost magnitudes. Informally, we measure how quickly the function has become insensitive to points away from the corresponding set K. As these terms become uniformly bounded in variance, the resulting Kolmogorov bounds display the optimal rates $s^{-1/2}$ or $n^{-1/2}$.

Remark 5. *The binomial limit in Theorem 2 corresponds to the Poisson case, although it has slightly different exponents because there is no exact binomial analogue of the second-order Poincaré inequality that exists in Poisson spaces; see [17] for the corresponding Berry–Esseen bounds.*

Corollary 1 (Optimal binomial convergence rate). *We assume the setting of Theorem 2 and suppose that $C > 0$ is a constant, so that*

$$\sup_{n\geq 1} \frac{\Theta_{K,n}}{\mathbf{Var}(F_n)} \leq C. \tag{13}$$

Then, there is a constant $C_0'' > 0$, which depends only on C and the assumption constants, such that for all $n \geq 2$,

$$d_K\Big(\frac{F_n - \mathbf{E}[F_n]}{\sqrt{\mathbf{Var}(F_n)}}, N\Big) \leq \frac{C_0''}{\sqrt{\mathbf{Var}(F_n)}}.$$

Particularly, as $Var(F_n)$ increases at least linearly with n, this gives the normal approximation rate of $O(n^{-1/2})$. A similar result holds for the Poisson model under similar stabilization and moment conditions. Particularly, if $\mathbf{Var}(F_s)$ grows linearly with s, then Theorem 1 shows that the convergence rate is of the order $O(s^{-1/2})$.

Theorems 1 and 2 summarize the general Poisson and binomial limits in Table 1.

Table 1. Asymptotic normality and Kolmogorov estimates for general stabilizing processes under the assumptions of Section 2.

Functional	Input Process	Rate at d_K
F	Poisson P_s	$C^* \sum_{i=1}^{6} \Gamma_i$
F_n	Binomial ξ_n	$C_0' \mathcal{R}(\Theta_{K,n}, \mathbf{Var}(F_n))$

Where

$$\mathcal{R}(\Theta_{K,n}, \mathbf{Var}(F_n)) = \frac{\Theta_{K,n}^{1/2}}{\mathbf{Var}(F_n)} + \frac{\Theta_{K,n}}{\mathbf{Var}(F_n)^{3/2}} + \frac{\Theta_{K,n} + \Theta_{K,n}^{3/2}}{\mathbf{Var}(F_n)^2}. \tag{14}$$

The constants C^* and C_0' are only dependent on structural assumptions and moment parameters, as in Theorem 2.

4. Applications and Simulation Strategy

Now, we will explain the scope of Theorems 1 and 2 using several functionals of practical interest, focusing particularly on Tsallis entropy estimators derived from nearest neighbor distances. We will also summarize a simulation strategy for verifying theoretical bounds in concrete settings.

4.1. Tsallis Entropy Estimators Based on Nearest Neighbors

Consider f as a probability density function on $\mathbb{R}^d$. For $\alpha \neq 1$, the α-th-order Tsallis entropy is defined as follows:

$$T_\alpha(f) = \frac{1}{1-\alpha}\left(\int_{\mathbb{R}^d} f(x)^\alpha \, dx - 1\right), \qquad \alpha \neq 1. \tag{15}$$

which is recovered by Shannon entropy as $\alpha \to 1$ [1,3]. Given that $X \sim f$, the integral term can be written as:

$$\int_{\mathbb{R}^d} f(x)^\alpha \, dx = \mathbb{E}\Big[f(X)^{\alpha-1}\Big].$$

Therefore, estimating $T_\alpha(f)$ can be reduced to estimating the expected value of $f(X)^{\alpha-1}$. An example point X_i yields the k-nearest neighbor estimation of the density as follows:

$$\widehat{f}_k(X_i) = \frac{k}{(n-1)\omega_d \rho_{k,i}^d},$$

where $\rho_{k,i}$ represents the distance from X_i to its k-th NN between the remaining observations, and ω_d denotes the volume of the unit sphere in $\mathbb{R}^d$. By substituting $\widehat{f}_k(X_i)$ for $f(X_i)^{\alpha-1}$ and taking the average over the sample, we obtain the estimator for the nearest neighbor Tsallis estimator:

$$\widehat{T}_{\alpha,n} = \frac{1}{1-\alpha}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{k}{(n-1)\omega_d \rho_{k,i}^d}\right)^{\alpha-1} - 1\right].$$

We obtain the Poisson version $\widehat{T}_{\alpha,s}$ in a similar manner by replacing $n-1$ with the density parameter s and summing over the points P_s. Consider a Poisson process P_s on $\mathcal{X} \subseteq \mathbb{R}^d$ with density sQ. Given a constant integer $k \geq 1$, let $\rho_k(x)$ denote the distance between $x \in P_s$ and its k-NN in P_s, and let ω_d be the volume of the unit ball in $\mathbb{R}^d$. Assume

$$N_s := |P_s|, \qquad N_s^+ := \max\{N_s, 1\}.$$

We define the normalized nearest-neighbor Tsallis estimator for the Poisson input as

$$\widehat{T}_{\alpha,s}(P_s) = \frac{1}{1-\alpha}\left[\frac{1}{N_s^+}\sum_{x\in P_s}\left(\frac{k}{s\omega_d \rho_k(x,P_s)^d}\right)^{\alpha-1} - 1\right].$$

Note that N_s^+ is used solely to avoid division by zero when $N_s = 0$. If $Q(S) = 1$, then $N_s \sim \text{Poisson}(s)$, and hence

$$\mathbb{P}(N_s = 0) = e^{-s},$$

so the change is exponentially negligible. The corresponding estimator for the binomial input is

$$\widehat{T}_{\alpha,n}(\xi_n) = \frac{1}{1-\alpha}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{k}{(n-1)\omega_d \rho_{k,i}^d}\right)^{\alpha-1} - 1\right],$$

where $\rho_{k,i}$ is the distance from X_i to its k-th nearest neighbor from the remaining sample points.

4.1.1. Bias and Consistency

The estimator exhibits the usual sources of bias associated with k-NN surrogate estimators. Among these are finite-sample bias in the local density estimate, boundary effects near the support of f, and the sensitivity of the power transformation $u \mapsto u^{\alpha-1}$. Given standard smoothness assumptions on f, a compact support, and a bounded density away from zero and infinity, the k-NN density estimator is locally consistent as $k = k_n \to \infty$ and $k_n/n \to 0$. Therefore, plugging in the estimate of $\mathbb{E}[f(X)^{\alpha-1}]$ leads to a consistent result after controlling for boundary and moment terms. In this paper, we have developed a stabilization-based normal approximation for fixed k and fixed entropy index α. Thus, the goal is not to optimize the bias but to describe the fluctuation behavior of the corresponding local k-NN functional.

4.1.2. Relation to Existing k-NN Entropy Estimators

Our proposed estimator is part of the family of nearest-neighbor pseudo- entropy estimators. Structurally, it is related to the Kozachenko–Leonenko estimator for Shannon entropy and the k-NN estimators for Rényi entropy, but it differs due to the Tsallis power transformation and $(1-\alpha)^{-1}$. As $\alpha \to 1$, the Tsallis functional approaches Shannon

entropy, while values of $\alpha \neq 1$ alter the estimator's sensitivity to low-density and high-density regions. In comparison to Tsallis goodness-of-fit estimators fully calibrated via simulation, this study focuses on the stabilization and normal approximation behavior of the underlying local k-NN functional. Hence, the originality lies not in the introduction of a completely different density estimator, but in the verification of the stabilization and moment conditions required to achieve explicit Gaussian approximation bounds for the Tsallis k-NN statistic.

4.1.3. Choice of the Entropy Index α

In this theoretical analysis, α is treated not as an unknown parameter of the underlying distribution, but as a fixed entropy index. For this reason, normal approximation results are presented based on a predetermined value of α in the interval $\alpha \in (0, \infty) \setminus \{1\}$, and the constants at the limits may depend on this fixed value of α. From a statistical perspective, different choices of α highlight different parts of the distribution: values below and above 1 may change the estimator's sensitivity to tail behavior and local concentration. For this reason, in empirical applications, α is determined through a sensitivity analysis on a finite grid, by minimizing the empirical mean squared error criterion when a reference model is provided, or by selecting the value that yields a stable goodness-of-fit calibration across Monte Carlo repetitions. An entirely model-based estimation theory for α goes beyond the scope of this paper; our goal here is to derive the normal approximation limits based on stabilization for each fixed acceptable entropy index.

We present a theorem that establishes the stabilization and moment properties justifying the application of the general normal approximation results to the Tsallis k-NN estimator.

Proposition 2 (Verification of stabilization and moment conditions for the k-NN Tsallis score)**.** *Suppose that Q satisfies a density f with respect to a compact subset $S \subset \mathbb{R}^d$ and that there exist constants $0 < m < M < \infty$ such that*

$$m \leq f(x) \leq M, \qquad x \in S.$$

Suppose $k \geq 1$ is fixed and let $\alpha \in (0, \infty) \setminus \{1\}$ be fixed. For $x \in P_s$, we define the local k-NN Tsallis score by

$$\psi_{\alpha,s}(x, P_s) = \left(\frac{k}{s\omega_d \rho_k(x, P_s)^d} \right)^{\alpha - 1},$$

where $\rho_k(x, P_s)$ is defined as the distance from x to its k-NN in P_s. We assume that, for some $p > 4$,

$$p(\alpha - 1)_+ < k, \qquad (\alpha - 1)_+ := \max\{\alpha - 1, 0\}.$$

Then the score $\psi_{\alpha,s}$, and therefore the associated estimator $\widehat{T}_{\alpha,s}$, stabilizes exponentially. Moreover, its first- and second-order add-one costs comply with a uniform L^p-moment bound of the following form:

$$\sup_{s \geq 1} \sup_{x,y \in S} \mathbf{E}\left[|\Delta_x \widehat{T}_{\alpha,s}|^p + |\Delta_{x,y} \widehat{T}_{\alpha,s}|^p \right] < \infty.$$

We can say the same for the binomial estimator $\widehat{T}_{\alpha,n}$, where n replaces s.

Proof. Full proof is given in Appendix B. The treatment of the random Poisson denominator $N_s = |P_s|$ is also given in Appendix B. □

Theorem 3 (Normal approximation for Tsallis entropy estimators)**.** *Assume the conditions of Proposition 2. Let there exist constants $C_P, C_B > 0$, depending purely on the structural constants k, d, p, and the fixed entropy index α, such that*

$$d_K\Big(\frac{\widehat{T}_{\alpha,s}(P_s) - \mathbf{E}[\widehat{T}_{\alpha,s}(P_s)]}{\sqrt{\mathbf{Var}(\widehat{T}_{\alpha,s}(P_s))}}, N\Big) \le \frac{C_P}{\sqrt{s}},$$

and

$$d_K\Big(\frac{\widehat{T}_{\alpha,n}(\xi_n) - \mathbf{E}[\widehat{T}_{\alpha,n}(\xi_n)]}{\sqrt{\mathbf{Var}(\widehat{T}_{\alpha,n}(\xi_n))}}, N\Big) \le \frac{C_B}{\sqrt{n}}.$$

Remark 6. *Theorem 3 establishes a unified asymptotic normality consequence for a natural class of nearest-neighbor-based estimators of Tsallis entropy and supplements previous results on the consistency and mean-square approximation of such estimators in the context of goodness-of-fit [7,8]. Under similar stabilization assumptions, the $s^{-1/2}$ and $n^{-1/2}$ rates correspond to the behavior exhibited by Shannon and Rényi entropy estimators and are thus classical central limit scaling classical rates. Thus, the current finding presents a framework similar to the Rényi entropy estimation proposed in [9], but it is supported by stabilization techniques and Poisson/binomial estimation methods.*

4.2. Weighted k-NN Shannon Entropy

To complete, let us briefly recall the weighted k-NN estimator of Shannon entropy recommended in [6]. Assume that $X_1, \dots, X_n$ are independent and identically distributed with density q on $\mathbb{R}^d$. Then the Shannon entropy is

$$H(q) = -\int_{\mathbb{R}^d} q(x) \log q(x)\, dx. \tag{16}$$

Let $\rho_{j,i}$ denote the distance from X_i to its j-th nearest neighbor in the set $\{X_1, \dots, X_n\} \setminus \{X_i\}$. Suppose that ω_d is the volume of the unit ball in $\mathbb{R}^d$ and Ψ is the digamma function. We obtain the estimator for weights $\{w_j\}_{j=1}^k$ satisfying the condition $\sum_{j=1}^k w_j = 1$

$$F_n^{\mathrm{SE}}(\xi_n) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k} w_j \log\left(\frac{(n-1)\omega_d \rho_{j,i}^d}{e^{\Psi(j)}}\right) \tag{17}$$

which has appropriate bias and variance properties under mild regularity conditions.

Theorem 4 (Normal approximation for weighted k-NN Shannon estimator)**.** *Given Assumptions 2–4 and the conditions of regularity from [6] about q and the weights $\{w_j\}$, let there be constants $C_0, \tau > 0$ that, based only on these conditions, are such that for all sufficiently large n,*

$$d_K\Big(\frac{F_n^{\mathrm{SE}}(\xi_n) - H(q)}{\sqrt{\mathbf{Var}(F_n^{\mathrm{SE}}(\xi_n))}}, N\Big) \le C_0\left(\frac{k}{n}\right)^{\tau}. \tag{18}$$

4.3. Geometric Functionals: Euler Characteristic and MST

As a geometric example, suppose we consider a binomial point process ξ_n on $[0,1]^d$ having density q, which is bounded away from zero and infinity on $[0,1]^d$. Denote the Vietoris-Rips or Čech complex with scale $r > 0$ by K_r and write the Euler characteristic as $\chi(K)$. We define

$$F_n^{\mathrm{EC}}(P_n) = \chi(K_r(n^{1/d} P_n)), \qquad F_n^{\mathrm{EC}}(\xi_n) = \chi(K_r(n^{1/d} \xi_n)). \tag{19}$$

Theorem 5 (Normal approach for the Euler characteristic)**.** *Assume that Assumptions 2–4 apply to the corresponding functionals of P_n and ξ_n. Then there exists a constant $C_{\mathrm{EC}} > 0$ with the following property:*

$$d_K\Big(\frac{F_n^{\mathrm{EC}}(\eta_n) - \mathbf{E}[F_n^{\mathrm{EC}}(\eta_n)]}{\sqrt{\mathbf{Var}(F_n^{\mathrm{EC}}(\eta_n))}}, N\Big) \leq \frac{C_{\mathrm{EC}}}{\sqrt{n}}, \tag{20}$$

where η_n denotes either P_n or ξ_n.

To find the normal approach rates for minimal spanning trees, let $V \subset \mathbb{R}^d$ be a finite set of points and let $M(V)$ represent the total length of the minimal spanning tree on V. Given that V is generated by a Poisson point process with intensity measure nQ, stabilization arguments and cost function techniques provide the following normal approach rates.

$$d_K\Big(\frac{F_n^{\mathrm{MST}} - \mathbf{E}[F_n^{\mathrm{MST}}]}{\sqrt{\mathbf{Var}(F_n^{\mathrm{MST}})}}, N\Big) \leq \begin{cases} C_{\mathrm{MST}}\, n^{-\gamma_1}, & d = 2, \\ C_{\mathrm{MST}}\, (\log n)^{-\gamma_2}, & d \geq 3, \end{cases}$$

for appropriate constants $C_{\mathrm{MST}}, \gamma_1, \gamma_2 > 0$ based on the basic distribution and dimension; see [10,14] for corresponding results.

We highlight the different rates achieved for the main functionals discussed in Theorems 1–5. We have summarized the corresponding Kolmogorov estimates in Table 2.

Table 2. The asymptotic normality and Kolmogorov rates for a set of entropy-related and geometric functionals under the stabilization conditions of Section 2.

Functional	Input Process	Target Quantity	Rate in d_K
Tsallis k-NN estimator $F_s^{(\alpha)}$	Poisson P_s	Tsallis entropy	$C_P\, s^{-1/2}$
Tsallis k-NN estimator $F_n^{(\alpha)}$	Binomial ξ_n	Tsallis entropy	$C_B\, n^{-1/2}$
Weighted k-NN estimator F_n^{SE}	Binomial ξ_n	Shannon entropy	$C_0(k/n)^{\tau}$
Euler characteristic F_n^{EC}	Poisson or binomial	Euler characteristic	$C_{\mathrm{EC}}\, n^{-1/2}$
MST total length F_n^{MST}	Poisson P_n	MST length	$C_{\mathrm{MST}} n^{-\gamma_1}$ or $C_{\mathrm{MST}}(\log n)^{-\gamma_2}$

4.4. Monte Carlo Simulation Protocol

To increase the transparency and reproducibility of the numerical component, this paper describes the Monte Carlo protocol used to evaluate the finite-sample behavior of the stabilization-based normal approximation results. The key elements of the design have been outlined in Table 3, while the computational steps are detailed in Algorithm 1.

Table 3. Monte Carlo design utilized to evaluate the finite-sample normal estimation of the nearest-neighbor entropy estimators.

Component	Details
Input models	Binomial samples from generalized Gaussian and Student-t distributions; homogeneous Poisson samples on bounded subsets of $\mathbb{R}^d$.
Dimensions	$d \in \{1, 2, 5\}$.
Sample sizes	$n \in \{200, 500, 1000, 2000\}$.
Entropy indices	Fixed values $\alpha \in \{0.8, 1.0, 1.2\}$, where $\alpha = 1$ is the Shannon limiting case.
Nearest-neighbor parameter	Fixed k, selected identically across sample sizes within a given scenario.
Monte Carlo simulations	$B = 300$ replications for each configuration.
Reported quantities	Empirical mean, empirical variance, scaled variance $n\,\widehat{\mathbf{Var}}$, empirical Kolmogorov distance, and standardized scores Z_n.
Purpose	The aim is to examine variance scaling, approximate Gaussianity, and the finite-sample behavior proposed by the stabilization-based normal approximation bounds.

Algorithm 1 Monte Carlo method for empirical normal approximation

Require: Distributional scenario, dimension d, sample size n, entropy index α, nearest-neighbor parameter k, number of replications B

Ensure: Empirical mean, empirical variance, standardized scores, and empirical Kolmogorov distance

1: **for** $b = 1, \ldots, B$ **do**

2: Generate an independent sample

$$X_1^{(b)}, \ldots, X_n^{(b)}$$

from the specified distribution.

3: Compute the k-nearest-neighbor distances

$$\rho_{k,i}^{(b)}, \qquad i = 1, \ldots, n.$$

4: Compute the Tsallis k-NN estimator

$$\widehat{T}_{\alpha,n}^{(b)}.$$

When required, compute the corresponding weighted Shannon estimator.

5: **end for**

6: Estimate the Monte Carlo mean:

$$\overline{T}_{\alpha,n} = \frac{1}{B} \sum_{b=1}^{B} \widehat{T}_{\alpha,n}^{(b)}.$$

7: Estimate the Monte Carlo variance:

$$\widehat{\mathbf{Var}}(\widehat{T}_{\alpha,n}) = \frac{1}{B-1} \sum_{b=1}^{B} \left(\widehat{T}_{\alpha,n}^{(b)} - \overline{T}_{\alpha,n} \right)^2.$$

8: **for** $b = 1, \ldots, B$ **do**

9: Standardize the estimator:

$$Z_n^{(b)} = \frac{\widehat{T}_{\alpha,n}^{(b)} - \overline{T}_{\alpha,n}}{\sqrt{\widehat{\mathbf{Var}}(\widehat{T}_{\alpha,n})}}.$$

10: **end for**

11: Compute the empirical Kolmogorov distance:

$$\widehat{d}_K = \sup_{t \in \mathbb{R}} \left| \widehat{F}_{Z_n}(t) - \Phi(t) \right|,$$

where $\widehat{F}_{Z_n}$ represents the empirical distribution function of the standardized scores and Φ is the standard normal distribution function.

12: **return** $\overline{T}_{\alpha,n}$, $\widehat{\mathbf{Var}}(\widehat{T}_{\alpha,n})$, $Z_n^{(1)}, \ldots, Z_n^{(B)}$, and $\widehat{d}_K$.

We implemented the simulation described in Table 3 using the Monte Carlo workflow outlined in Algorithm 1. We recorded the empirical variance, scaled variance $n\widehat{\mathbf{Var}}$, empirical Kolmogorov distance $\widehat{d}_K$, and standardized scores for each configuration. As the paper's main contribution is theoretical, the simulation protocol has been used as an illustrative diagnostic of variance scaling and approximate Gaussianity rather than as a separate empirical study. We interpret the graphical summaries in Figures 1 and 2 in accordance with this protocol. In a typical setup, dimensions $d \in \{1, 2, 5\}$, sample sizes $n \in \{200, 500, 1000, 2000\}$, and various fixed values of the entropy index $\alpha \in \{0.8, 1.0, 1.2\}$ are considered. Under the null hypothesis, simulations can be performed from multivariate

generalized Gaussian distributions with different shape parameters. For each configuration, Tsallis, Shannon, and Rényi entropy estimators can be calculated, their empirical means and variances can be estimated, and the standardized estimator can be compared to the Gaussian benchmark across Monte Carlo replications to approximate the Kolmogorov distance to the normal distribution.

The main empirical variables of interest are:

- Bias and mean squared error of the Tsallis and Rényi entropy estimators for various n, d, α;
- Empirical variance and its scaled value in n;
- Empirical Kolmogorov distance between the normalized estimators and a normal distribution;
- Empirical size and power of Tsallis- and Rényi-based goodness-of-fit tests developed from the estimators.

Figure 1 displays the empirical variance of the Tsallis and weighted Shannon k-NN entropy estimators as a function of sample size n for several comparison distributions and sizes on a double logarithmic scale. In all scenarios examined, the variance is approximately linearly decreasing with $\log n$. It has a slope close to -1, consistent with the theoretical estimate $\mathbf{Var}(F_n) = O(n^{-1})$ suggested by the central limit theorems in Section 3. The alignment with the reference $O(n^{-1})$ line is particularly evident for medium and large sample sizes. It is robust with respect to changes in the underlying distribution (light-tailed generalized Gaussian, heavy-tailed generalized Gaussian, and Student-t) and dimension ($d = 1$ and $d = 5$). The Tsallis estimator exhibits slightly higher variance than the Shannon estimator in heavy-tailed and high-dimensional settings. However, both estimators have the same asymptotic slope, suggesting that the stabilization-based normal approximations capture the dominant scaling behavior of the fluctuation magnitude. These empirical findings support the optimality of the $n^{-1/2}$ rates at our Kolmogorov limits and show that theoretical results remain informative even at moderately high dimensions and under significant deviations from the Gaussian distribution.

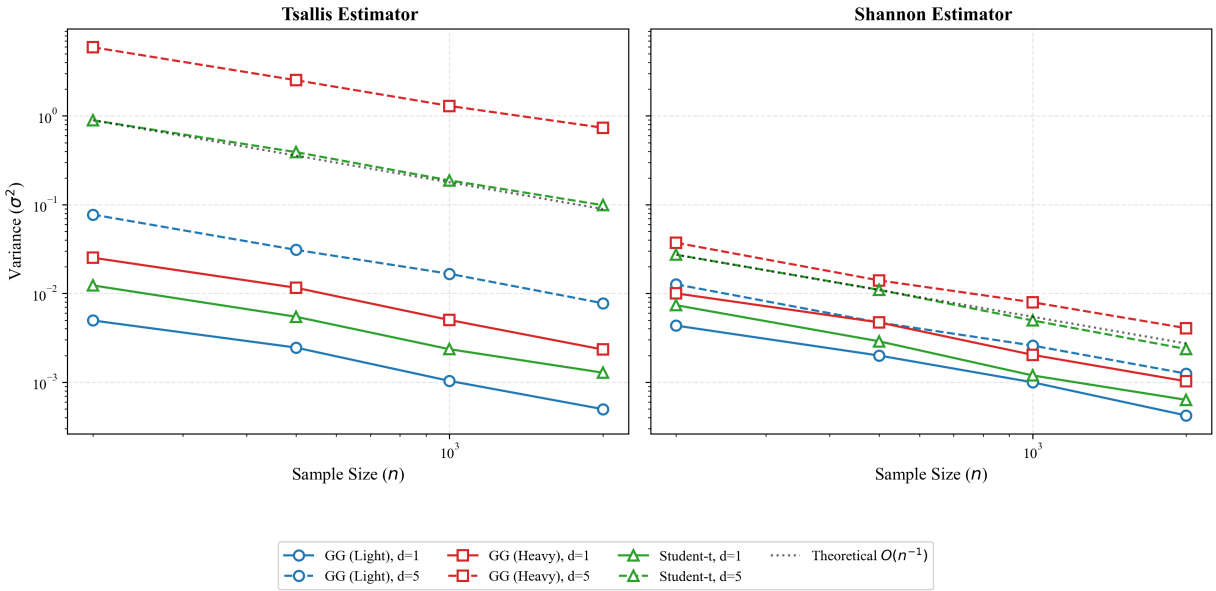


Figure 1. Empirical variance for the Tsallis and the weighted Shannon k-NN entropy estimators, as a function of sample size n on a log-log scale. The results are based on $B = 300$ Monte Carlo replications per setup. You can see a grey dotted reference line with slope -1 that matches the theoretical variance scaling $\mathbf{Var}(\widehat{T}_{\alpha,n}) = O(n^{-1})$. In the panels, they juxtapose light-tailed generalized Gaussian, then heavy-tailed generalized Gaussian, plus Student-t cases, across a few chosen dimensions.

Figure 2 supplies a qualitative evaluation of the asymptotic normality of the Tsallis and weighted Shannon k-NN entropy estimators. We generate $n \in \{200, 500, 1000, 2000\}$ observations for a light-tailed bivariate generalized Gaussian distribution ($d = 2$, shape parameter $\beta = 2.5$) and computed over 300 Monte Carlo replications, which were found to provide stable estimates while keeping computational costs moderate. Monte Carlo replications. We use the following formula to standardize the resulting estimates $\hat{H}$ for each n:

$$Z_n = \frac{\hat{H} - \mathbf{E}[\hat{H}]}{\sqrt{\mathbf{Var}(\hat{H})}}, \tag{21}$$

ensuring that any deviation from the standard normal distribution is reflected in the shape of the standardized sample.

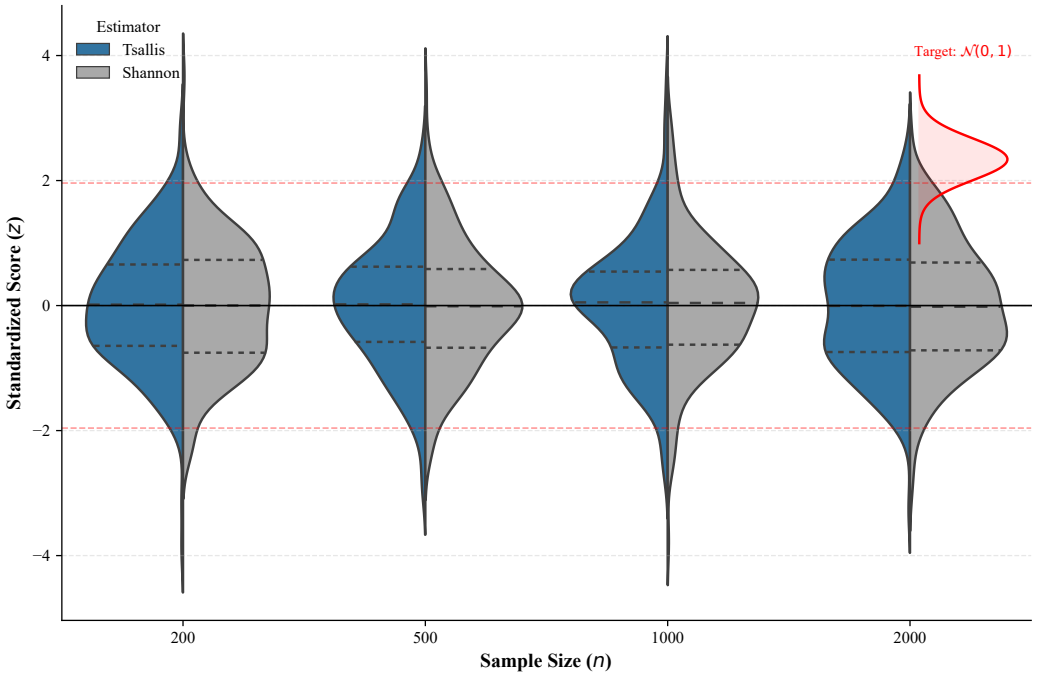


Figure 2. Violin plots of the standardized Tsallis and weighted Shannon k-NN entropy estimators for the bivariate generalized Gaussian model with shape parameter $\beta = 2.5$. From each sample size $n \in \{200, 500, 1000, 2000\}$, standardized scores are calculated from $B = 300$ Monte Carlo simulations using the empirical Monte Carlo mean and variance. The dashed horizontal lines represent the ± 1.96 standard normal quantiles. Increasing symmetry and concentration around zero suggest improved alignment with the Gaussian approximation.

The violin plots in Figure 2 show how the empirical distributions of the standardized Tsallis and Shannon estimators change as the sample size increases. The fact that the distribution's mean is at zero and the reference quantile alignment indicate that the normal approximation holds for every n value examined by the researchers. The $n = 200$ sample shows that both estimators deviate from Gaussianity due to their asymmetrical distributions and additional tail weight. The distribution shapes develop symmetric patterns while their distribution mass begins to concentrate at the origin. The Tsallis estimator displays a dispersion pattern that is very similar to that of the Shannon estimator at all sample sizes considered, indicating that the use of a generalized entropy index (here $\alpha = 1.2$) does not lead to slower convergence or markedly heavier tails. The $O(n^{-1/2})$ convergence rate with theoretical variance scaling indicates that higher n values lead to a decrease in variability. Empirical results supporting the theoretical findings indicate that stabilization-based normal approximations accurately characterize the distributional behavior of estimators. The

visual diagnostics confirm the quantitative Kolmogorov distance results, which match the central limit theorems established in Section 3 because both estimators reach asymptotic normality under the stabilization and moment assumptions established in this study.

5. Proofs of Main Results

The section provides proofs of all main normal approximation results documented in Section 3. The study demonstrates how adaptive add-one cost decompositions interact with second-order Poincaré-type inequalities.

5.1. Poisson Case

The proof of Theorem 1 requires two components, which include the second-order Poincaré inequality for Poisson functionals proved by [15] together with an adaptive decomposition of the increment operators. For convenience, we recall a simplified version of the main inequality.

Theorem 6 (Last–Peccati–Schulte [15])**.** *Suppose F is a square-integrable function of the Poisson point process $P(\lambda)$ with intensity measure λ, satisfying*

$$\mathbf{E}[F^2] < \infty \qquad \textit{and} \qquad \mathbf{E}\Big[\int_{\mathcal{X}} (\Delta_x F)^2 \, \lambda(dx)\Big] < \infty. \tag{22}$$

Then

$$d_K\Big(\frac{F - \mathbf{E}[F]}{\sqrt{\mathbf{Var}(F)}}, N\Big) \le \sum_{i=1}^{6} \gamma_i,$$

where $\gamma_1, \ldots, \gamma_6$ denote explicit integrals involving the first and second derivatives of F.

Proof of Theorem 1. We obtain the proof by applying the second-order Poincaré inequality for Poisson functionals with respect to the localized add-one cost decomposition. Assume that $\sigma_F^2 = \mathbf{Var}(F) > 0$. Let us write, for each $x \in \mathcal{X}$,

$$\Delta_x F = \big(\Delta_x F - \Delta_x F(A_x)\big) + \Delta_x F(A_x).$$

Let the first term measure the localization error, while the second term is the localized add-one cost. Likewise, let the second-order increment $\Delta_{x,y}F$ be decomposed into localized and residual parts based on the same family of localization sets A_x and A_y. The application of the Poisson second-order Poincaré inequality to F yields six Malliavin–Stein error terms that involve products of first- and second-order increments. The localization decomposition separates these terms into contributions governed by

$$b_1(x, A_x) = \mathbf{E}\big|\Delta_x F - \Delta_x F(A_x)\big|^4, \qquad b_2(x, A_x) = \mathbf{E}\big|\Delta_x F(A_x)\big|^4,$$

and analogous quantities that involve $\Delta_{x,y}F$. According to Hölder's inequality and the assumed L^p-moment bounds, each expression is bounded by a normalized integral that involves $\Delta_x F$, $\Delta_{x,y}F$, and the intensity measure λ_s. These normalized integrals represent precisely the quantities $\Gamma_1, \ldots, \Gamma_6$, listed explicitly in Appendix A. We combine the six resulting estimates and absorb universal constants into C^* to produce

$$d_K\Big(\frac{F - \mathbf{E}[F]}{\sqrt{\mathbf{Var}(F)}}, N\Big) \le C^* \sum_{i=1}^{6} \Gamma_i.$$

Given the spatial localization and moment assumptions, the terms Γ_i approach zero as $s \to \infty$, which establishes the asserted normal approximation bound. □

5.2. Binomial Case

Theorem 6 does not have its exact equivalent for the binomial setting, yet [17] proved related inequalities. Theorem 2 results from applying their proof methods to our cost-operator framework. The process for this method follows the same steps as the Poisson method, except that discrete sums replace integrals and suitable combinatorial factors are used.

Proof of Theorem 2. The proof derives from the binomial normal approximations for stabilizing functionals, integrated with the spatial localization and moment assumptions discussed in Section 2. Given $x \in \mathcal{X}$, the add-one cost $\Delta_x F_n$ decomposes into a localized term with a residual term of the form

$$\Delta_x F_n = \big(\Delta_x F_n - \Delta_x F_n(A_x)\big) + \Delta_x F_n(A_x).$$

We apply the same decomposition to the second-order add-one cost $\Delta_{x,y} F_n$. We can express the residual terms in the form of a sum of terms with exponential spatial localization, while the localized terms are governed by the uniform L^p-moment assumption. The binomial Berry–Esseen estimate then yields three types of contributions: a first-order localization term, a second-order interaction term, and a higher-order remaining term. By applying Hölder's inequality and the spatial decay assumption, these contributions become bounded, respectively, by

$$\frac{\Theta_{K,n}^{1/2}}{\mathbf{Var}(F_n)}, \qquad \frac{\Theta_{K,n}}{\mathbf{Var}(F_n)^{3/2}}, \qquad \frac{\Theta_{K,n} + \Theta_{K,n}^{3/2}}{\mathbf{Var}(F_n)^2}.$$

Here $\Theta_{K,n}$ denotes the effective spatial localization quantity described in Theorem 2. Applying these bounds yields

$$d_K\left(\frac{F_n - \mathbf{E}[F_n]}{\sqrt{\mathbf{Var}(F_n)}}, N\right) \le C_0'\left(\frac{\Theta_{K,n}^{1/2}}{\mathbf{Var}(F_n)} + \frac{\Theta_{K,n}}{\mathbf{Var}(F_n)^{3/2}} + \frac{\Theta_{K,n} + \Theta_{K,n}^{3/2}}{\mathbf{Var}(F_n)^2}\right),$$

where C_0' is only based on the constants in the stabilization, spatial decay, and moment assumptions. □

5.3. Proof of Corollary 1

The corollary follows by choosing $A_x = \mathcal{X}$, which simplifies the decomposition so that $b_1 \equiv 0$ and the main contribution comes from b_2. The assumptions imply that $\Theta_{K,n}$ stays under uniform control when compared to $\mathbf{Var}(F_n)$ and Theorem 2 establishes a bound which reduces to the value of $\mathbf{Var}(F_n)^{-1/2}$. The details are similar to those of the Poisson case examined in [16], and therefore we will not include them.

6. Conclusions

The research establishes general stabilization-based processes that researchers can use to prove central limit theorems, providing specific Kolmogorov convergence rates for all functions derived from Poisson and binomial point processes. This method provides first- and second-order growth control through the systematic application of adaptable addition cost operators, eliminating the need for detailed stabilization radius calculations and explicit point-function decompositions.

The framework provides optimal-order normal approximation results for nearest-neighbor Tsallis entropy estimators based on natural regularity and tail and moment conditions applied to the underlying metric measure space. The same methodology recovers and extends existing normal approximation results for weighted k-NN Shannon entropy estimators, Euler characteristics of random geometric complexes, and minimal spanning

tree functionals, highlighting the unifying role of stabilization techniques in geometric probability and information-theoretic estimation.

The results presented here suggest several directions for further research. This work can be extended by testing Tsallis and related entropy estimators in non-Euclidean spaces, such as manifolds and graphs, as these spaces possess unique geometric properties that affect nearest-neighbor relations and stabilization methods. It can be adapted to select the k-NN parameter. It will investigate how data-driven approaches perform in high-dimensional settings, examining their impact on bias and variance during stabilization. The current standard approach should be better combined with entropy-based goodness-of-fit tests and dependency modelling to develop tests and estimators that exhibit strong asymptotic performance in complex, high-dimensional contexts while preserving reliability.

Author Contributions: Conceptualization, M.S.Ç. and M.S.; methodology, M.S.Ç. and M.S.; formal analysis, M.S.Ç.; investigation, M.S.Ç.; writing—original draft preparation, M.S.Ç.; writing—review and editing, M.S.Ç. and M.S.; supervision, M.S.; project administration, M.S.Ç. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The APC was funded by Linköping University.

Data Availability Statement: The code and figure outputs used for the asymptotic analysis of the Tsallis and Shannon entropy estimators are available at: https://github.com/mehmetsiddik/asymptotic-tsallis-entropy-estimators (accessed on 29 May 2026).

Acknowledgments: The authors are grateful to Nikolai Leonenko for his valuable comments and support during the preparation of this work.

Conflicts of Interest: The authors declare no conflicts of interest. The APC funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Malliavin–Stein Quantities Used in the Poisson Bound

This appendix provides an explicit definition of the Malliavin–Stein measures used in Theorem 1. We follow the notation developed by Last, Peccati, and Schulte [15] for second-order Poincaré inequalities for Poisson functionals and the stabilization formulation used for geometric Poisson functionals. Given a Poisson point process P_s on $(\mathcal{X}, \mathcal{F})$ with intensity measure $\lambda_s = sQ$, let $F = F(P_s)$ be a square-integrable Poisson functional satisfying

$$\sigma_s^2 := \mathbf{Var}(F) > 0.$$

Define the first- and second-order difference operators for $x, y \in \mathcal{X}$ by

$$D_x F := F(P_s + \delta_x) - F(P_s),$$

and

$$D^2_{x,y} F := F(P_s + \delta_x + \delta_y) - F(P_s + \delta_x) - F(P_s + \delta_y) + F(P_s).$$

It is assumed that $F \in \operatorname{dom} D$, meaning that

$$\mathbf{E}[F^2] < \infty, \qquad \mathbf{E}\left[\int_{\mathcal{X}} (D_x F)^2 \, \lambda_s(dx)\right] < \infty.$$

To get the Kolmogorov bound, let us define the fifth-moment constants

$$c_{1,s} := \operatorname*{ess\,sup}_{x \in \mathcal{X}} \mathbf{E}\left[|D_x F|^5\right], \qquad c_{2,s} := \operatorname*{ess\,sup}_{x,y \in \mathcal{X}} \mathbf{E}\left[|D^2_{x,y} F|^5\right],$$

where the essential suprema are obtained with respect to λ_s and λ_s^2, respectively. According to Assumption 4, these quantities are finite.

In Theorem 1, the six Malliavin–Stein error terms are described as follows:

$$\Gamma_{1,s}(F) := \frac{2(c_{1,s}c_{2,s})^{1/5}}{\sigma_s^2}\left[\int_{\mathcal{X}^3}\left\{\mathbb{P}(D^2_{x_1,x_3}F\neq 0)\mathbb{P}(D^2_{x_2,x_3}F\neq 0)\right\}^{1/20}\lambda_s^3(d(x_1,x_2,x_3))\right]^{1/2}, \tag{A1}$$

$$\Gamma_{2,s}(F) := \frac{c_{2,s}^{2/5}}{\sigma_s^2}\left[\int_{\mathcal{X}^3}\left\{\mathbb{P}(D^2_{x_1,x_3}F\neq 0)\mathbb{P}(D^2_{x_2,x_3}F\neq 0)\right\}^{1/10}\lambda_s^3(d(x_1,x_2,x_3))\right]^{1/2}, \tag{A2}$$

$$\Gamma_{3,s}(F) := \frac{1}{\sigma_s^3}\int_{\mathcal{X}}\mathbf{E}\left[|D_xF|^3\right]\lambda_s(dx), \tag{A3}$$

$$\Gamma_{4,s}(F) := \frac{c_{1,s}^{3/5}\lambda_s(\mathcal{X})}{\sigma_s^3} + \frac{c_{1,s}^{4/5}\lambda_s(\mathcal{X})^{5/4} + 2c_{1,s}^{4/5}\lambda_s(\mathcal{X})^{3/2}}{\sigma_s^4}, \tag{A4}$$

$$\Gamma_{5,s}(F) := \frac{c_{1,s}^{2/5}\lambda_s(\mathcal{X})^{1/2}}{\sigma_s^2}, \tag{A5}$$

$$\Gamma_{6,s}(F) := \frac{\sqrt{6}(c_{1,s}c_{2,s})^{1/5} + \sqrt{3}c_{2,s}^{2/5}}{\sigma_s^2}\left[\int_{\mathcal{X}^2}\mathbb{P}(D^2_{x_1,x_2}F\neq 0)^{1/10}\lambda_s^2(d(x_1,x_2))\right]^{1/2}. \tag{A6}$$

According to these definitions, the Kolmogorov limit used in Theorem 1 can be expressed as follows:

$$d_K\left(\frac{F-\mathbf{E}[F]}{\sqrt{\mathbf{Var}(F)}}, N\right) \leq \sum_{i=1}^{6}\Gamma_{i,s}(F).$$

The probabilities appearing in $\Gamma_{1,s}(F)$, $\Gamma_{2,s}(F)$, and $\Gamma_{6,s}(F)$ are measures of the second-order difference operators being nonzero. In the stabilization of geometric functionals, these probabilities are determined by the localization radius and diminish exponentially as the included points move farther apart in the rescaled metric. This provides the connection between the abstract Malliavin-Stein limit and the stabilization hypotheses employed in the main text.

Appendix B. Detailed Proof of Proposition 2

For the nearest-neighbor Tsallis score, we offer details concerning stabilization and moment validation. Proof is provided for the Poisson distribution; for the binomial distribution, we apply the standard binomial analog using sample size n in replacement of the Poisson density parameter s.

Step 1: Local dependence of the k-nearest-neighbor score.

The local Tsallis score for $x \in P_s$ is given by:

$$\psi_{\alpha,s}(x,P_s) = \left(\frac{k}{s\omega_d\rho_k(x,P_s)^d}\right)^{\alpha-1},$$

where $\rho_k(x,P_s)$ denotes the distance from x to its k-th nearest neighbor within P_s. Thus, by the configuration of points inside the sphere, the value of $\psi_{\alpha,s}(x,P_s)$ is given by:

$$B_x(\rho_k(x,P_s)).$$

As a result, if the outside configuration of a sphere including the k-nearest neighbors of x is modified, the score remains unchanged. This demonstrates the local dependence of the score.

Step 2: Exponential stabilization.

We take $R_s(x)$ as the stabilization radius for the point x. It is sufficient for the ball $B_x(R_s(x))$ to include the points that determine the k-nearest-neighbor structure of x. For a compact support S under the lower density condition $f \geq m > 0$, for a sufficiently small $r > 0$,

$$Q(B_x(r) \cap S) \geq c_1 r^d.$$

Thus, for a Poisson point processes with density measure sQ,

$$\mathbb{P}\{P_s(B_x(r) \cap S) < k\} \leq \sum_{j=0}^{k-1} \exp\{-sQ(B_x(r) \cap S)\} \frac{(sQ(B_x(r) \cap S))^j}{j!}.$$

By using $Q(B_x(r) \cap S) \geq c_1 r^d$, the right-hand side is restricted to the appropriate constants $C_1, C_2 > 0$ such that

$$C_1 \exp\{-C_2 s r^d\}.$$

Alternatively, by setting $t = s^{1/d} r$,

$$\mathbb{P}\{s^{1/d} R_s(x) > t\} \leq C_1 \exp\{-C_2 t^d\}.$$

Therefore, the stabilization radius exhibits an exponential tail in the rescaled metric. The same argument applies to the binomial point process by using the corresponding binomial lower-tail bound. Consequently, the k-nearest-neighbor Tsallis score stabilizes exponentially.

Step 3: Stabilization of the add-one costs.

Adding a point has a local cost; in fact, adding a point z affects only the scores of those points that have become k nearest neighbors of z. Given a fixed k and a fixed dimension d, the total number of points involved in this process is governed by a deterministic constant depending only on k, d. Thus, the first-order insertion cost can be written as a finite sum of local score differences. Likewise, a second-order addition cost contains only points simultaneously affected in the k-nearest-neighbor structure by both added points. Therefore, both $\Delta_z \widehat{T}_{\alpha,s}$ and $\Delta_{z,w} \widehat{T}_{\alpha,s}$ determine the configuration within a finite combination of stabilizing spheres.

Step 4: Uniform L^p-moment control.

The remaining task is to verify that the first- and second-order "addition" costs possess uniform finite p-th moments in the interval s. Due to the upper density bound $f \leq M$, the local excess density can be uniformly controlled. The possible singular behavior originates from the small k-nearest-neighbor distances. The standard Poisson nearest-neighbor calculation for the k-nearest-neighbor volume yields finite negative moments when the following condition is satisfied:

$$p(1-\alpha)_+ < k.$$

It guarantees that the power term appearing in the Tsallis score has a finite p-th moment when $\alpha < 1$. When $\alpha > 1$, the compactness of S and the exponential tail of the stabilization radius control the correspondingly positive powers of the nearest-neighbor distance.

We obtain the following by combining the finite local interaction property from Step 3 with the moment bound above:

$$\sup_{s\geq 1}\sup_{z,w\in S}\mathbb{E}\Big[|\Delta_z\widehat{T}_{\alpha,s}|^p+|\Delta_{z,w}\widehat{T}_{\alpha,s}|^p\Big]<\infty.$$

Using the same argument, we can apply this to the binomial estimator $\widehat{T}_{\alpha,n}$, where n replaces s. Then, as claimed in Proposition 2, it has exponential stabilization and uniform L^p-moment control.

Step 5: Control of the random Poisson denominator.

It remains to explain the role of the random denominator in the Poisson version of the estimator. Suppose

$$N_s:=|P_s|,\qquad N_s^+:=\max\{N_s,1\},$$

and let

$$S_s(P_s):=\sum_{x\in P_s}\psi_{\alpha,s}(x,P_s),\qquad \overline{S}_s(P_s):=\frac{S_s(P_s)}{N_s^+}.$$

For Q(S)=1, $N_s\sim$ Poisson(s). More generally, $N_s\sim$ Poisson$(sQ(S))$ if the support has a finite Q-measure. For the normalized case $Q(S)=1$, it follows that

$$\mathbb{P}(N_s=0)=e^{-s},$$

and, for every fixed $r>0$,

$$\mathbf{E}\left[\left|\frac{N_s}{s}-1\right|^r\right]=O(s^{-r/2}),\qquad \mathbf{E}\big[(N_s^+)^{-r}\big]=O(s^{-r}).$$

In fact, we can get the second bound by splitting the expectation over $\{N_s\geq s/2\}$ and $\{N_s<s/2\}$, and applying the standard lower-tail concentration bound for Poisson random variables. Given an inserted point z, we have that the first-order add-one cost of the normalized sum satisfies

$$D_z\overline{S}_s=\frac{D_zS_s}{N_s+1}+S_s\left(\frac{1}{N_s+1}-\frac{1}{N_s+1}\right).$$

For the event $\{N_s\geq 1\}$, it can be written as

$$D_z\overline{S}_s=\frac{D_zS_s}{N_s+1}-\frac{S_s}{N_s(N_s+1)}.$$

We can see that the first term is determined by the same local stabilization method used for the unnormalized score, combined with the fact that $N_s+1\asymp s$ occurs with high probability. The second term is a solely normalizing contribution and is controlled by the Poisson distribution of N_s around s, together with the moment bound for the local score sum S_s. The occurrence of $N_s=0$ provides only an exponentially small error of order e^{-s}.

The second-order add-one cost is dealt with in a similar manner by using another difference operator on the above expression. It consists of a combination of normalization expressions containing the following factors, obtained by dividing the local second-order score contributions by N_s+2:

$$\frac{1}{N_s+1'},\qquad \frac{1}{(N_s+1)(N_s+2)},\qquad \frac{1}{N_s(N_s+1)}.$$

Based on the above Poisson concentration limits, all these factors have a $O(s^{-1})$ or $O(s^{-2})$ order of magnitude within L^r, excluding an exponentially small error. As a result, the random denominator does not change the stabilization radius, the uniform L^p-moment control, or the order of the normal approximation limit.

References

1. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
2. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [CrossRef]
3. Tsallis, C. Possible generalization of Boltzmann–Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
4. Zhang, H.; Li, B.; Tian, W.; Sun, Q. Tail-Aware Information-Theoretic Generalization for RLHF and SGLD. *arXiv* **2026**, arXiv:2604.10727. [CrossRef]
5. Lee, K.; Kim, S.; Lim, S.; Choi, S.; Oh, S. Tsallis Reinforcement Learning: A Unified Framework for Maximum Entropy Reinforcement Learning. *arXiv* **2019**, arXiv:1902.00137. [CrossRef]
6. Berrett, T.B.; Samworth, R.J.; Yuan, M. Efficient multivariate entropy estimation via *k*-nearest neighbour distances. *Ann. Stat.* **2019**, *47*, 288–318. [CrossRef]
7. Çadırcı, M.S. Non-parametric goodness-of-fit tests using Tsallis entropy measures. *Entropy* **2025**, *27*, 1210. [CrossRef] [PubMed]
8. Çadırcı, M.S.; Evans, D.; Leonenko, N.; Makogin, V. Entropy-based test for generalized Gaussian distributions. *Comput. Stat. Data Anal.* **2022**, *173*, 107502. [CrossRef]
9. Çadırcı, M.S.; Evans, D.; Leonenko, N.; Makogin, V.; Seleznjev, O. Statistical tests based on Rényi entropy estimation. *arXiv* **2025**, arXiv:2502.08654. [CrossRef]
10. Kesten, H.; Lee, S. The central limit theorem for weighted minimal spanning trees on random points. *Ann. Appl. Probab.* **1996**, *6*, 495–527. [CrossRef]
11. Penrose, M.D.; Yukich, J.E. Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.* **2001**, *11*, 1005–1041. [CrossRef]
12. Penrose, M.D.; Yukich, J.E. Multivariate normal approximation in geometric probability. *Ann. Appl. Probab.* **2005**, *15*, 1677–1700. [CrossRef]
13. Baryshnikov, Y.; Yukich, J.E. Gaussian limits for random measures in geometric probability. *Ann. Appl. Probab.* **2005**, *15*, 213–253. [CrossRef]
14. Chatterjee, S.; Diaconis, P.; Sly, A. Minimal spanning trees and Stein's method. *Probab. Theory Relat. Fields* **2017**, *167*, 849–909. [CrossRef]
15. Last, G.; Peccati, G.; Schulte, M. Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Probab. Theory Relat. Fields* **2016**, *165*, 667–723. [CrossRef]
16. Lachièze-Rey, R.; Schulte, M.; Yukich, J.E. Normal approximation for stabilizing functionals. *Ann. Appl. Probab.* **2019**, *29*, 931–993. [CrossRef]
17. Lachièze-Rey, R.; Peccati, G. New Berry–Esseen bounds for functionals of binomial point processes. *Ann. Appl. Probab.* **2017**, *27*, 1992–2031. [CrossRef]
18. Lachièze-Rey, R.; Peccati, G.; Yang, J. Stabilization and normal approximation via cost operators. *Ann. Appl. Probab.* **2022**, *32*, 1415–1454.
19. Shi, Z.; Balasubramanian, K.; Polonik, W. A flexible approach for normal approximation of geometric and topological statistics. *Bernoulli* **2024**, *30*, 3029–3058. [CrossRef]
20. Kingman, J.F.C. *Poisson Processes*; Clarendon Press: Oxford, UK, 1993.
21. Daley, D.J.; Vere-Jones, D. *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*; Springer: Berlin/Heidelberg, Germany, 2007.
22. Penrose, M.D. Gaussian limits for random geometric measures. *Electron. J. Probab.* **2007**, *12*, 989–1035. [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.