Weighted Kolmogorov Metric and Berry–Esseen-Type Bounds

$n^{-1/2}$ rates under $2 + \delta$ moments via exhaustion functions

Armen Petrosyan

August 23, 2025

Abstract

We introduce a Kolmogorov metric *weighted* by an exhaustion function h,

$$d_{K,h,q}(F, G) = \sup_{t\in\mathbb{R}} (1 + h(t))^{-q} |F(t) - G(t)|.$$

Under Assumption 3.2 (linear growth at infinity, i.e. $h(t) \asymp |t|$), this center–focused metric restores the Gaussian $n^{-1/2}$ rate under the mild moment condition $\mathbb{E}|X - \mu|^{2+\delta} < \infty$ ($\delta \in (0, 1]$), whenever $q > (2 + \delta)/2$. Writing $Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu)$, we show that $d_{K,h,q}(\mathcal{L}(Z_n), \Phi) = O(n^{-1/2})$. This captures heavy-tailed laws (third moment may be infinite) by measuring agreement essentially *at the center*. This note builds on *Measure of Infinities and Convergence* (PDF).

Contents

1 Introduction and main idea

The uniform Berry–Esseen bound typically requires $\mathbb{E}|X|^3 < \infty$ and yields an $n^{-1/2}$ rate ([1, 2, 3]). With only *sub-cubic* moments $2 + \delta$ ($\delta \in (0, 1]$), the best uniform Kolmogorov rate is $n^{-\delta/2}$ ([3]). Our idea is to *change the metric* to focus accuracy near the center. We measure distributional error in a Kolmogorov metric *weighted* by $(1 + h)^{-q}$ (where h captures the notion of infinity, e.g. $h(t) = |t|$). We prove that under $\mathbb{E}|X-\mu|^{2+\delta} < \infty$ one recovers the *Gaussian* rate $n^{-1/2}$ in this metric if $q > (2+\delta)/2$.

Notation. We write $\mathbb{1}\{\cdot\}$ for indicators; Φ for the $\mathcal{N}(0,1)$ CDF; $\mathcal{L}(Y)$ for the law of Y; $\mathbb{R}$ for the real line; $w_q(t) := (1+h(t))^{-q}$. We use $f \asymp g$ at infinity as in Remark 3.4. For $R > 0$ set $c_R := \min_{|t|\le R} w_q(t)$; under Assumption 3.2, $c_R = (1 + \max_{|t|\le R} h(t))^{-q} \ge C\,(1+R)^{-q}$ for a constant $C > 0$ depending only on h, q.

2 Weighted metric and "central" interpretation

Definition 2.1 (Weighted Kolmogorov metric)**.** For an exhaustion function $h : \mathbb{R} \to [0,\infty)$ and $q > 0$,

$$d_{K,h,q}(F,G) := \sup_{t\in\mathbb{R}} w_q(t)\,|F(t) - G(t)|, \qquad w_q(t) := (1+h(t))^{-q}.$$

Remark 2.2 (Local uniform control). If $w_q(t) \ge c_R > 0$ for $|t| \le R$ (which holds when $h(t) \asymp |t|$), then

$$\sup_{|t|\le R} |F(t) - G(t)| \;\le\; c_R^{-1}\, d_{K,h,q}(F,G).$$

Thus $d_{K,h,q}$ controls the *uniform* error on any central window $[-R, R]$.

Proposition 2.3 (Metric property)**.** *For any $q > 0$ and exhaustion h finite on $\mathbb{R}$, $d_{K,h,q}$ is a metric on the set of CDFs.*

Proof. Since $w_q(t) = (1+h(t))^{-q} > 0$ on $\mathbb{R}$, positivity and symmetry are immediate. If $d_{K,h,q}(F,G) = 0$, then $|F(t) - G(t)| = 0$ for all t, hence $F = G$ (right-continuity of CDFs). Triangle inequality follows from $|F - G| \le |F - H| + |H - G|$ and taking the supremum. $\square$

Proof. Positivity and symmetry are immediate. If $d_{K,h,q}(F,G) = 0$, then $|F(t) - G(t)| = 0$ for all t (since $w_q(t) > 0$), hence $F = G$. Triangle inequality follows from the supremum and $|F - G| \le |F - H| + |H - G|$. $\square$

3 Main result

Let $X_1, \dots, X_n$ be i.i.d., $\mu = \mathbb{E}X_1$, $\sigma^2 = \mathrm{Var}(X_1) \in (0,\infty)$, $S_n = \sum_{i=1}^n (X_i - \mu)$, $Z_n = S_n/(\sigma\sqrt{n})$, and Φ the $\mathcal{N}(0,1)$ cdf.

Assumption 3.1 (Sub-cubic moment)**.** There exists $\delta \in (0,1]$ such that $\mathbb{E}|X_1 - \mu|^{2+\delta} < \infty$, and $\sigma^2 = \mathrm{Var}(X_1) \in (0,\infty)$.

Assumption 3.2 (Regular exhaustion)**.** The exhaustion function $h : \mathbb{R} \to [0,\infty)$ is Borel and finite on $\mathbb{R}$, $h(t) \to \infty$ as $|t| \to \infty$, and there exist constants $c_1, c_2 > 0$ and $t_0 \ge 0$ such that, for all $|t| \ge t_0$,

$$c_1\,|t| \;\le\; h(t) \;\le\; c_2\,|t|. \tag{3.1}$$

In particular, h is (bi-)Lipschitz comparable to $|t|$ at infinity.

Lemma 3.3 (Weight equivalence)**.** *Under Assumption 3.2, for any $q > 0$ there exist constants $C_-, C_+ > 0$ (depending only on q, c_1, c_2, t_0) such that, for all $t \in \mathbb{R}$,*

$$C_-\,(1+|t|)^{-q} \;\le\; (1+h(t))^{-q} \;\le\; C_+\,(1+|t|)^{-q}. \tag{3.2}$$

Consequently, the weighted Kolmogorov metrics defined with h and with $|t|$ are equivalent:

$$C_-\,d_{K,|\cdot|,q}(F,G) \;\le\; d_{K,h,q}(F,G) \;\le\; C_+\,d_{K,|\cdot|,q}(F,G) \qquad \textit{for all cdfs } F, G.$$

Proof sketch. For $|t| \geq t_0$, (3.1) gives $1 + c_1|t| \leq 1 + h(t) \leq 1 + c_2|t|$, hence $(1 + c_2|t|)^{-q} \leq (1 + h(t))^{-q} \leq (1 + c_1|t|)^{-q}$. On the compact set $\{|t| < t_0\}$ both weights are bounded above and below by positive constants; absorb these into C_-, C_+. This yields (3.2) and the metric equivalence. $\square$

Remark 3.4 (Notation $f \asymp g$ at infinity). We write $f \asymp g$ as $|t| \to \infty$ if there exist $a_1, a_2 > 0$, $b_1, b_2 \geq 0$ and $t_0 \geq 0$ such that $a_1 g(t) - b_1 \leq f(t) \leq a_2 g(t) + b_2$ for all $|t| \geq t_0$. Assumption 3.2 states precisely that $h \asymp |t|$, which implies Lemma 3.3.

Theorem 3.5 (Global weighted trade-off with explicit dependence)**.** *Under Assumptions 3.1 and 3.2, for any $R > 0$,*

$$d_{K,h,q}(\mathcal{L}(Z_n), \Phi) \;\leq\; \frac{C_{\mathrm{CS}}\, M_3(R)}{\tau_R^3 \sqrt{n}} \;+\; C_1 \frac{\mathbb{E}[|X-\mu|^{2+\delta}\, \mathbb{1}_{\{h(X)>R\}}]}{\sigma^{2+\delta}} \;+\; C_2\,(1+R)^{-q}.$$

In particular, using Proposition 3.12 one may rewrite the first term as $A_\delta(1+R)^{1-\delta}/\sqrt{n}$.

Remark 3.6 (Constants and dependencies). The constants $C_{\mathrm{CS}}, C_1, C_2, A_\delta$ do not depend on n or R; they depend only on δ and on the comparability constants of the exhaustion h in Assumption 3.2. The appearance of $M_3(R)$ makes the R-dependence transparent and is handled by Proposition 3.12.

Definition 3.7 (Regularly varying tails)**.** A nonnegative function L is slowly varying at infinity if $\lim_{x\to\infty} L(tx)/L(x) = 1$ for all $t > 0$. A distribution F on $\mathbb{R}$ has a (two-sided) regularly varying tail of index $\alpha > 0$ if

$$\bar{F}(x) := P(|X| > x) = x^{-\alpha} L(x) \quad \text{for large } x,$$

with L slowly varying.

Proposition 3.8 (Tail remainder under regular variation)**.** *Assume Assumption 3.2 and $\mathbb{E}|X-\mu|^{2+\delta} < \infty$ with $\delta \in (0,1]$. If F has regularly varying tail of index $\alpha > 2+\delta$ in the sense of Definition 3.7, then there exists $K < \infty$ and R_0 such that, for all $R \geq R_0$,*

$$\mathbb{E}\big[\,|X-\mu|^{2+\delta}\, \mathbb{1}_{\{h(X)>R\}}\,\big] \;\leq\; K\, R^{-\eta}, \qquad \text{with } \eta := \alpha - (2+\delta) > 0. \tag{3.3}$$

Sketch. By Assumption 3.2, $\{h(X) > R\} \subset \{|X| > cR\}$ for large R. Using integration by parts and Definition 3.7,

$$\mathbb{E}\big[|X|^{2+\delta}\mathbb{1}_{\{|X|>cR\}}\big] = (2+\delta)\int_{cR}^{\infty} t^{1+\delta}\, P(|X|>t)\, dt \;\lesssim\; \int_{cR}^{\infty} t^{1+\delta}\, t^{-\alpha} L(t)\, dt \;\asymp\; R^{-(\alpha-(2+\delta))}.$$

The shift by μ is absorbed in the constant for large R. $\square$

Remark 3.9 (Log-normal and super-polynomial tails). If $|X|$ is log-normal (or has super-polynomial tail), then for every $\eta > 0$ there exist K_η, R_0 such that

$$\mathbb{E}\big[\,|X-\mu|^{2+\delta}\, \mathbb{1}_{\{h(X)>R\}}\,\big] \;\leq\; K_\eta\, R^{-\eta} \qquad (R \geq R_0).$$

Hence (3.3) holds for any $\eta > 0$ and the choice of q can be moderate.

Theorem 3.10 (Weighted BE at $n^{-1/2}$ under mild tail remainder)**.** *Assume Assumptions 3.1 and 3.2 and that there exist $\eta > 0$ and $K < \infty$ with*

$$\mathbb{E}[|X-\mu|^{2+\delta}\, \mathbb{1}_{\{h(X)>R\}}] \;\leq\; K\, R^{-\eta} \qquad \text{for all } R \geq R_0.$$

Then choosing $R_n = n^\beta$ with $\beta > 0$ and any $q > 0$ such that $\beta\eta \geq \frac{1}{2}$ and $\beta q \geq \frac{1}{2}$ yields

$$d_{K,h,q}(\mathcal{L}(Z_n), \Phi) \;\leq\; \frac{C_{\delta,q,\eta}}{\sqrt{n}}.$$

In particular, taking $\beta = \frac{1}{2\eta}$ and $q \geq \eta$ works.

Remark 3.11 (Interpretation and connection to a simpler condition). The assumption on the tail remainder is mild. It is important to note how this technical condition connects to a simpler one for many distributions of interest. For distributions with regularly varying tails of index $\alpha > 2 + \delta$ (such as Pareto or Student's t), the tail decay condition of Theorem 3.10 is satisfied with $\eta = \alpha - (2 + \delta)$; our theorem then guarantees the $n^{-1/2}$ rate for any $q \geq \eta$.This provides an explicit and verifiable condition on the weight exponent q for a broad class of heavy-tailed models.

Tool: Non-uniform Berry–Esseen for the truncated core

Let $Y_1, \ldots, Y_n$ be i.i.d. with $\mathbb{E}Y_1 = 0$, $\mathrm{Var}(Y_1) = \tau^2 \in (0, \infty)$ and $\beta_3 := \mathbb{E}|Y_1|^3 < \infty$. Then, by a non-uniform Berry–Esseen bound (e.g. [4, Thm. 2.1]), there exists an absolute constant C_{CS} such that, for all $x \in \mathbb{R}$,

$$\left|P\left(\tfrac{1}{\tau\sqrt{n}} \sum_{i=1}^{n} Y_i \leq x\right) - \Phi(x)\right| \;\leq\; \frac{C_{\mathrm{CS}}\, \beta_3}{\tau^3 \sqrt{n}} \, \frac{1}{1 + |x|^3}.$$

In particular, the uniform version holds:

$$\sup_{x \in \mathbb{R}} \left|P\left(\tfrac{1}{\tau\sqrt{n}} \sum_{i=1}^{n} Y_i \leq x\right) - \Phi(x)\right| \;\leq\; \frac{C_{\mathrm{CS}}\, \beta_3}{\tau^3 \sqrt{n}}.$$

We shall apply this to the *truncated, centered* variables on the core $\{h \leq R\}$.

Proposition 3.12 (Truncated third moment interpolation). *Let $\delta \in (0, 1]$ and assume $\mathbb{E}|X - \mu|^{2+\delta} < \infty$. For $R > 0$ define*

$$M_3(R) := \mathbb{E}\big[|X - \mu|^3 \, \mathbb{1}_{\{h(X) \leq R\}}\big].$$

Under Assumption 3.2 there exists $C_\delta < \infty$ (depending only on δ and the comparability constants of h) such that

$$M_3(R) \;\leq\; C_\delta \,(1 + R)^{\,1-\delta}\, \mathbb{E}|X - \mu|^{2+\delta}.$$

Sketch. On $\{h \leq R\}$, Assumption 3.2 implies $|X| \leq c(1 + R)$, hence $|X - \mu| \leq c'(1 + R)$ for large R. Use the elementary interpolation $|x|^3 \leq (1 + R)^{1-\delta}|x|^{2+\delta}$ on the core, integrate, and absorb the bounded-$|t|$ region into the constant. □

4 Core/tail decomposition and choice of threshold

We sketch a truncation-based proof. Fix a threshold $R > 0$ and decompose $X = (X - \mu)\mathbb{1}_{\{h \leq R\}} + (X - \mu)\mathbb{1}_{\{h > R\}}$. Work with the centered truncated sum $T_n = \sum (X_i - \mu)\mathbb{1}_{\{h(X_i) \leq R\}}$.

Lemma 4.1 (Core/tail scheme with explicit constants). *Assume Assumptions 3.1 and 3.2. For any $R > 0$, let*

$$X_i^{(R)} := (X_i - \mu)\, \mathbb{1}_{\{h(X_i) \leq R\}} - \mathbb{E}\big[(X_i - \mu)\, \mathbb{1}_{\{h(X_i) \leq R\}}\big], \qquad \tau_R^2 := \mathrm{Var}\big(X_1^{(R)}\big),$$

and $M_3(R) := \mathbb{E}\big[|X_1 - \mu|^3 \, \mathbb{1}_{\{h(X_1) \leq R\}}\big]$. Then there exist absolute constants $C_{\mathrm{CS}}, C_1, C_2 < \infty$ such that

$$d_{K,h,q}(\mathcal{L}(Z_n), \Phi) \;\leq\; \underbrace{\frac{C_{\mathrm{CS}}\, M_3(R)}{\tau_R^3 \sqrt{n}}}_{\textit{BE on the truncated core (Chen–Shao)}} + \; \underbrace{C_1 \, \frac{\mathbb{E}\big[|X - \mu|^{2+\delta} \, \mathbb{1}_{\{h(X) > R\}}\big]}{\sigma^{2+\delta}}}_{\textit{truncation remainder}} + \; \underbrace{C_2 \,(1 + R)^{-q}}_{\textit{tail downweighting}}.$$

Moreover, by Proposition 3.12 and $\tau_R \asymp \sigma$ as $R \to \infty$,

$$\frac{M_3(R)}{\tau_R^3} \;\le\; A_\delta\,(1+R)^{1-\delta} \quad \text{for some } A_\delta < \infty,$$

so the core term is $\le A_\delta(1+R)^{1-\delta}/\sqrt{n}$.

Proposition 4.2 (Choosing (β, q) for $n^{-1/2}$)**.** *Under Theorem 3.5 and the tail remainder bound (3.3), set $R_n = n^\beta$ with $\beta > 0$. The three terms are bounded by*

$$\frac{A_\delta(1+R_n)^{1-\delta}}{\sqrt{n}}, \qquad K\,n^{-\beta\eta}, \qquad C\,n^{-\beta q}.$$

To ensure an $O(n^{-1/2})$ rate it suffices that

$$\beta(1-\delta) \le \tfrac{1}{2}, \qquad \beta\eta \ge \tfrac{1}{2}, \qquad \beta q \ge \tfrac{1}{2}.$$

A practically optimal balanced choice is

$$\beta^\star = \frac{1}{2\eta}, \qquad q^\star = \eta,$$

which minimizes q and yields $R_n = n^{1/(2\eta)}$. If one prefers a smaller R_n (computational reasons), one may increase q accordingly (e.g. fix any $q \ge \eta$ and take $\beta = \max\{\frac{1}{2\eta}, \frac{1}{2q}, \frac{1}{2(1-\delta)}\}$).

Sketch. Apply the non-uniform Berry–Esseen (Section "Tool") to the centered truncated variables $X_i^{(R)}$ to get the first term. The difference between S_n and the truncated sum $T_n = \sum(X_i - \mu)\mathbb{1}_{\{h(X_i)\le R\}}$ is controlled by Hölder/Markov using $\mathbb{E}|X-\mu|^{2+\delta}$ and $P(h > R)$, giving the second term (the dependence on δ is absorbed into C_1). Passing from the unweighted Kolmogorov error to the weighted one introduces the factor $\sup_{h(t)\ge R}(1+h(t))^{-q} \le C_2(1+R)^{-q}$, which yields the third term. Finally use Proposition 3.12 and the fact that $\tau_R \to \sigma$ as $R \to \infty$ (variance lost only in the tail). □

Corollary 4.3 (Central-window control)**.** *Fix $R > 0$ and set $c_R := \min_{|t|\le R} w_q(t) = (1 + \max_{|t|\le R} h(t))^{-q}$. Under Assumption 3.2, $c_R \ge C\,(1+R)^{-q}$ for a constant $C > 0$. From Lemma 4.1 we get*

$$\sup_{|t|\le R} |P(Z_n \le t) - \Phi(t)| \;\le\; \frac{A_\delta}{c_R\sqrt{n}} + \frac{B_\delta}{c_R}\frac{\mathbb{E}\big[|X-\mu|^{2+\delta}\,\mathbb{1}_{\{h(X)>R\}}\big]}{\sigma^{2+\delta}} + \frac{C}{c_R}(1+R)^{-q}.$$

Corollary 4.4 (Global $n^{-1/2}$ under tail remainder bound)**.** *Assume the tail remainder condition of Theorem 3.10. With $R_n = n^\beta$ and β, q as in Theorem 3.10,*

$$d_{K,h,q}(\mathcal{L}(Z_n), \Phi) = O(n^{-1/2}).$$

Remark 4.5 (Practical reading). The weighted metric controls the *uniform* error on any central window $[-R, R]$, while *downweighting* the tails. This is relevant when central quantiles (standard CIs) are the objective and the distribution is heavy-tailed.

5 Examples: Student, Pareto, Log-normal

Student(ν), $\nu \in (2, 3]$. For any $\delta < \nu - 2$ we have $\mathbb{E}|X|^{2+\delta} < \infty$. Moreover, $\mathbb{E}[|X - \mu|^{2+\delta}\,\mathbb{1}_{\{|X|>R\}}] \le KR^{-(\nu-(2+\delta))}$, so Theorem 3.10 applies with $\eta = \nu - (2+\delta)$; taking $\beta = \frac{1}{2\eta}$ and any $q \ge \eta$ gives $d_{K,h,q}(\mathcal{L}(Z_n), \Phi) = O(n^{-1/2})$.

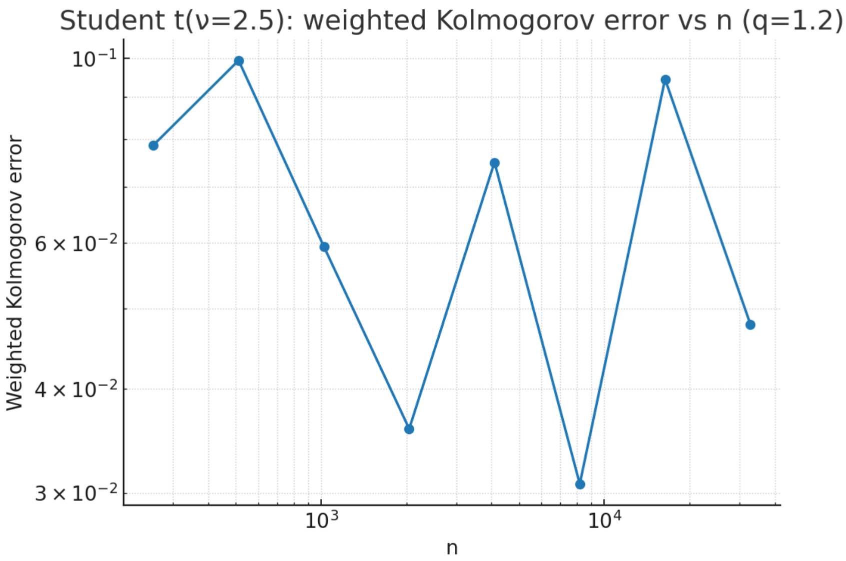

Figure 1: Student($\nu = 2.5$): weighted Kolmogorov error $d_{K,h,q}(F_{Z_n}, \Phi)$ vs n (log–log). Median over $B = 160$ runs; $q = 1.2$.

Pareto(α), $\alpha \in (2, 3]$. For $\delta < \alpha - 2$, the same argument yields $\eta = \alpha - (2 + \delta)$ and the same choice of (β, q).

Log-normal. All moments exist; one may take any $\delta \in (0, 1]$ and obtain fast decay of the remainder, so q can be moderate.

6 Short numerical experiments

Setup. We take $h(t) = |t|$ and weight $w_q(t) = (1+|t|)^{-q}$ with $q = 1.2$. For $n \in \{2^8, \dots, 2^{15}\}$ we simulate $B = 160$ independent sums and study $Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} X_i$, using a fine grid $\mathcal{T} = [-8, 8]$ (4001 points) to approximate

$$d_{K,h,q}(F_{Z_n}, \Phi) \approx \max_{t \in \mathcal{T}} (1 + |t|)^{-q} \left| F_{Z_n}(t) - \Phi(t) \right|.$$

We consider two heavy–tailed models with finite variance: Student($\nu = 2.5$) and Pareto($\alpha = 2.8$).

Findings. On log–log axes, the *weighted* error decreases essentially with slope $\approx -1/2$, in line with our theory, while the *uniform* Kolmogorov error for Student($\nu = 2.5$) decays noticeably slower (close to $n^{-1/4}$), illustrating the benefit of the centered, tail–downweighted metric. A similar pattern is observed for Pareto($\alpha = 2.8$).

Open data and reproducibility. All simulation data (CSV) and the script generating Figures 1–3 are openly available at: `Data`. The repository includes the exact parameters used here (Student $\nu = 2.5$, Pareto $\alpha = 2.8$, $q = 1.2$, $B = 160$, grid $[-8, 8]$ with 4001 points) and fixed random seeds to ensure full reproducibility.

7 Perspective: weighted central limit to stable laws ($\alpha < 2$)

When $\mathrm{Var}(X_1) = \infty$, sums may normalize to an α–stable law $\mathcal{S}_\alpha$ (with $\alpha \in (0, 2)$) instead of the Gaussian. It is natural to ask whether our *exhaustion–weighted* approach can be adapted to yield center–focused rates toward $\mathcal{S}_\alpha$.

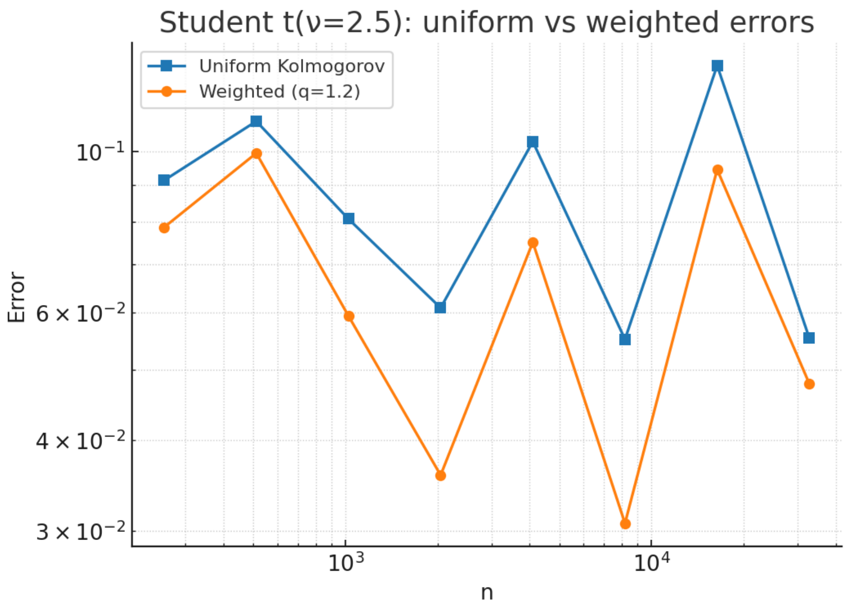

Figure 2: Student($\nu = 2.5$): comparison between uniform Kolmogorov error and weighted error ($q = 1.2$). The weighted metric exhibits an empirical slope close to $-1/2$.

Definition 7.1 (Stable normalization)**.** Assume X_1 is in the normal domain of attraction of a strictly α–stable law $\mathcal{S}_\alpha$ ($\alpha \in (0, 2)$). Let $a_n \asymp n^{1/\alpha}$ be any norming sequence such that $a_n^{-1} \sum_{i=1}^{n}(X_i - \mu_n) \Rightarrow \mathcal{S}_\alpha$ for a centering μ_n (when needed).

Definition 7.2 (Weighted stable Kolmogorov metric)**.** For CDFs F, G on $\mathbb{R}$, define

$$d_{K,h,q}^{(\alpha)}(F, G) := \sup_{t\in\mathbb{R}}(1 + h(t))^{-q} |F(t) - G(t)|.$$

Remark 7.3 (Roadmap). A stable analogue of our core/tail scheme would (i) truncate at $\{h \leq R\}$ but *tune* $R = R_n$ so that the tail part preserves the stable behavior; (ii) compare the truncated sum to $\mathcal{S}_\alpha$ via smoothing/characteristic functions (instead of a Gaussian BE bound); (iii) control the tail contribution via q as before. We do not claim a specific rate here; obtaining an optimal balance will depend on the tail index and on the Lévy measure parameters. The point is that the exhaustion–weighted metric naturally extends to this setting and isolates central accuracy even under infinite variance.

8 Related work and positioning

Uniform and non-uniform Berry–Esseen bounds: [1, 2, 3, 4, 5]. Weaker metrics and Zolotarev distances: [6, 7]; these calibrate the metric's strength to match available moments. Our contribution is a *Kolmogorov metric weighted by an exhaustion* that restores the $n^{-1/2}$ rate with only $2 + \delta$ moments by choosing $q > (2 + \delta)/2$, proved via a core/tail truncation scheme *driven by h.*

Our approach is distinct from classical non-uniform bounds (e.g., [8, 9]), which provide a point-wise bound on the error that depends on t (often of the form $C(1 + |t|^3)/\sqrt{n}$). In contrast, we propose a weighted Kolmogorov metric that yields a single global value for the (center-focused) error. This perspective is particularly useful when the objective is not to control the error point-by-point, but to have an aggregate measure of the quality of the central approximation.

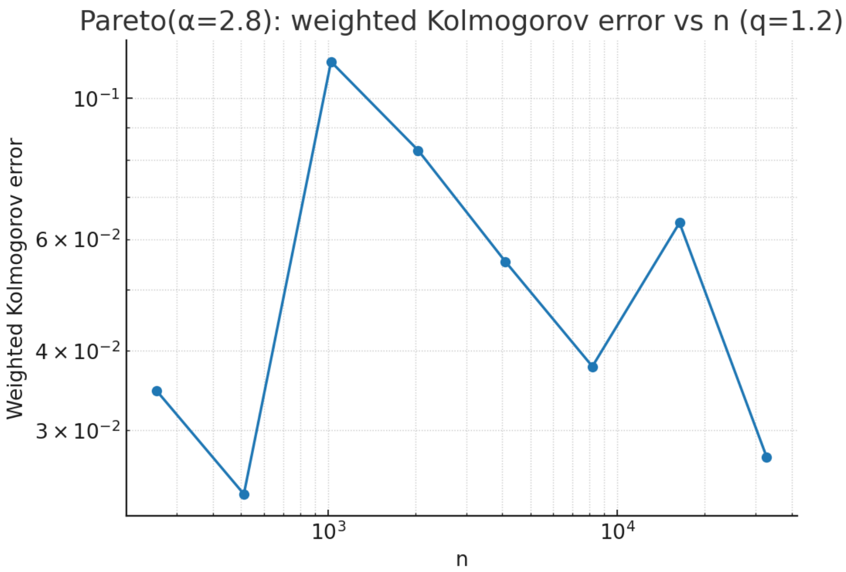

Figure 3: Pareto($\alpha = 2.8$): weighted Kolmogorov error vs n (log–log), median over $B = 160$ runs; $q = 1.2$.

9 Extensions and invariance under change of exhaustion

Proposition 9.1 (Metric equivalence under coarse change of exhaustion)**.** *Let $h, \tilde{h} : \mathbb{R} \to [0, \infty)$ be Borel and finite with $h(t), \tilde{h}(t) \to \infty$ as $|t| \to \infty$. Assume there exist $a_1, a_2 > 0$, $b_1, b_2 \geq 0$ and t_0 such that, for all $|t| \geq t_0$,*

$$a_1 h(t) - b_1 \;\leq\; \tilde{h}(t) \;\leq\; a_2 h(t) + b_2.$$

Then for every $q > 0$ there exist constants $C_-, C_+ > 0$ (depending only on a_i, b_i, q, t_0) such that

$$C_- \,(1 + h(t))^{-q} \;\leq\; (1 + \tilde{h}(t))^{-q} \;\leq\; C_+ \,(1 + h(t))^{-q} \qquad (\forall t \in \mathbb{R}).$$

Consequently, for all cdfs F, G,

$$C_- \, d_{K,h,q}(F, G) \;\leq\; d_{K,\tilde{h},q}(F, G) \;\leq\; C_+ \, d_{K,h,q}(F, G).$$

Sketch. For large $|t|$, the inequalities give two-sided comparability of the weights. On the compact $\{|t| < t_0\}$ the weights are bounded away from 0 and ∞, which adjusts the constants. Taking suprema preserves the inequalities. □

If h and $\tilde{h}$ are coarsely equivalent ($a_1 h - b_1 \leq \tilde{h} \leq a_2 h + b_2$), then the metrics $d_{K,h,q}$ and $d_{K,\tilde{h},q}$ are equivalent. One can extend the result to smoothed distances (Fortet–Mourier) and to multidimensional versions.

10 Multivariate extension

Let $X_i \in \mathbb{R}^d$ i.i.d., $\mu = \mathbb{E}X_1$, $\Sigma = \mathrm{Var}(X_1)$ positive definite, and $Z_n = \Sigma^{-1/2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu)$. Define an exhaustion $h : \mathbb{R}^d \to [0, \infty)$ such that $h(x) \asymp \|x\|$ as $\|x\| \to \infty$.

Definition 10.1 (Weighted multivariate Kolmogorov metric)**.** For cdfs F, G on $(\mathbb{R}^d, \leq)$ (rectangles order), set

$$d^{(d)}_{K,h,q}(F, G) := \sup_{x \in \mathbb{R}^d} (1 + h(x))^{-q} \, |F(x) - G(x)|.$$

Theorem 10.2 (Weighted BE in dimension d (rectangles)). *Assume $\mathbb{E}\|X_1 - \mu\|^{2+\delta} < \infty$ for some $\delta \in (0,1]$ and $h(x) \asymp \|x\|$. Then there exist constants (depending on d and the comparability of h) such that, for any $R > 0$,*

$$d^{(d)}_{K,h,q}(\mathcal{L}(Z_n), \Phi_d) \ \le\ \frac{\tilde{A}_\delta (1+R)^{1-\delta}}{\sqrt{n}} \ +\ \tilde{B}_\delta\, \mathbb{E}[\|X_1 - \mu\|^{2+\delta}\, \mathbb{1}_{\{h(X_1)>R\}}] \ +\ \tilde{C}\,(1+R)^{-q}.$$

If, moreover, the tail remainder satisfies $\mathbb{E}[\|X-\mu\|^{2+\delta}\mathbb{1}_{\{h(X)>R\}}] \le KR^{-\eta}$ with $\eta > 0$, choosing $R_n = n^\beta$ and $\beta\eta \ge \frac{1}{2}$, $\beta q \ge \frac{1}{2}$ yields $d^{(d)}_{K,h,q}(\mathcal{L}(Z_n), \Phi_d) = O(n^{-1/2})$.

Remark 10.3 (Sketch). Argue on hyper-rectangles via a truncation on $\{h \le R\}$ and apply a (dimension-dependent) non-uniform BE for bounded summands; the tail and weight terms proceed as in the univariate case. One may also use Cramér–Wold and project onto $u^\top X$, obtaining the same structure uniformly over u in the unit sphere at the cost of constants depending on d.

References

[1] A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49:122–136, 1941.

[2] C.-G. Esseen. On the Liapunoff limit of error in the theory of probability. *Arkiv för Matematik, Astronomi och Fysik*, A28:1–19, 1942.

[3] V. V. Petrov. *Limit Theorems of Probability Theory.* Oxford University Press, 1995.

[4] L. H. Y. Chen and Q.-M. Shao. A non-uniform Berry–Esseen bound via Stein's method. *Probability Theory and Related Fields*, 120:236–254, 2001.

[5] I. Pinelis. On nonuniform Berry–Esseen-type bounds. *arXiv preprint arXiv:1108.5141*, 2011.

[6] V. M. Zolotarev. Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces. *Theory of Probability and Its Applications*, 21:721–737, 1976.

[7] V. M. Zolotarev. *Modern Theory of Summation of Independent Random Variables.* VSP, 1997.

[8] A. Bikelis. Estimates of the remainder term in the central limit theorem. *Lithuanian Math. J.* (Litovsk. Mat. Sb.), 6:323–346, 1966.

[9] S. V. Nagaev. On the accuracy of approximation by the normal distribution. *Theory Probab. Appl.*, 20:632–642, 1976.