Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Semi-Analytic Approximation for Scale-Homogeneous Sums of Random Variables: Lyapunov-Guided Partitioning Between Gaussian Aggregation and Targeted Simulation

bing xiong

bingxiong1211@163.com

Sichuan University

Research Article

Keywords: Berry–Esseen theory, Lyapunov ratio, clustering, semi-analytic simulation

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11612083/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Semi-Analytic Approximation for Scale-Homogeneous Sums of Random Variables: Lyapunov-Guided Partitioning Between Gaussian Aggregation and Targeted Simulation

Bing Xiong

We consider the distribution of a sum of many independent scale-homogeneous positive random variables $X_i \overset{d}{=} k_i X_0$, where full Monte Carlo (MC) simulation is prohibitive and Normal approximation suffers severe tail errors. We propose a semi-analytic approximation that acts as a diagnostic scalpel: using the Lyapunov ratio from Berry–Esseen theory as a heuristic criterion, clustering the variables into a "Good" group, which is aggregated by a closed-form Gaussian with Cornish-Fisher adjustment, and a "Bad" group of eccentric variables, which is retained for targeted MC simulation. Numerical experiments show KS-distance reductions of one to two orders of magnitude versus full Normal approximation, with the 99% and 99.9% quantile errors reduced significantly, while cutting the sampling burden from n variables to |Bad| variables. The framework has broad applications, such as ABS pools, insurance aggregates.

Key words: Berry–Esseen theory, Lyapunov ratio, clustering, semi-analytic simulation

Contents

1. Introduction and related works 3

2. The model 5

2.1 The Berry–Esseen Bound and the Lyapunov Ratio 6

2.2 The Lyapunov Ratio of a Single Variable and of a Group 8

2.3 Closure Property of Good Groups 10

3. The Scale-Homogeneous Case and the Berry Measure 10

3.1 Properties of the Berry Measure 12

3.2 Heuristic Clustering on the Transformed Scale Parameters 13

4. The Semi-Analytic Algorithm 15

5. Error Analysis 16

5.1 Notation and Decomposition of the Error 16

5.2 Finite-Simulation Error 17

5.3 Accuracy of the CF-Adjusted Normal Approximation 17

5.4 Pointwise Variability of a Simulated CDF 18

6. Numerical Experiments 18

6.1 Experimental Setup 18

6.2 Small Volume of Random Variables: First Experiment, n = 50, List 1 18

6.3 Summary and Conclusions 23

6.4 Large Volume of Random Variables: Performance Comparison with FFT and Müller Inspired Method 27

6.5 Discussion on Numerical Experiment 30

7. Conclusion 31

Appendix A: Scale Parameters Used in the Experiments 33

References 34

1. Introduction and related works

Monte Carlo (MC) simulation is the standard approach to computing the distribution of aggregate random variables. Although straightforward and robust, it converges slowly, and the computational burden becomes prohibitive when high accuracy is required. The classical alternatives to Monte Carlo simulation are the Fast Fourier Transform (FFT) .

The FFT approach evaluates the characteristic function of the sum,

$\varphi_S(\mathrm{t})=\prod_{i=1}^{n}\varphi_i(\mathrm{t})$

on a frequency grid and recovers the density by an inverse transform, at a cost of O(nN) characteristic-function evaluations for a grid of size N . Three structural drawbacks are particularly relevant here.

First, the cost scales linearly with n regardless of how close to Gaussian the sum already is: when n is large and the summands are homogeneous, the central limit theorem already suffices. More generally, the method provides no criterion for whether the sum is already sufficiently close to normal

Second, the discretization introduces intertwined numerical errors. The discrete transform periodizes the density, folding tail mass outside the truncation window back into it − the aliasing error studied by Embrechts and Frei (2009) and Schaller and Temnov (2008) − which decays only polynomially for heavy-tailed summands. Since resolution and window width are locked to each other for fixed N , tail coverage and accuracy cannot be refined independently; exponential tilting mitigates aliasing but shifts rather than eliminates the error, distorts the opposite tail, and amplifies floating-point cancellation. When the base distribution is available only empirically, its characteristic function must itself be estimated, with the noise compounding across the n factors. These error sources are entangled in the output and cannot be bounded separately.

A closely related refinement of the central limit theorem is proposed by **Müller (2019)**, who improves the Gaussian approximation by explicitly accounting for the contribution of extreme order statistics. Let $X_1, \dots, X_n$ be i.i.d. random variables with a heavy right tail satisfying a Pareto-type condition

$$1 - F(x) \sim \omega^{1/\xi} x^{-1/\xi}, \quad \omega > 0,$$

for a tail parameter

$$\frac{1}{3} < \xi < 1.$$

Rather than approximating the entire sum by a Gaussian distribution, Müller separates the largest k order statistics from the remaining observations. The sum can therefore be represented schematically as

$$S_n = \sum_{i=1}^{n} X_i \approx \underbrace{\sum_{j=1}^{k} X_{(n-j+1)}}_{\text{extreme component}} + \underbrace{\mathcal{N}(\mu_k, \sigma_k^2)}_{\text{non-extreme component}},$$

where the non-extreme component is approximated by a Gaussian distribution after truncation, while the contribution of the extreme order statistics is approximated using extreme-value asymptotics. This construction can substantially improve upon the ordinary CLT approximation when the assumed tail structure accurately describes the relevant extreme region.

However, Müller's approximation is structurally dependent on a Pareto-like tail model and the associated extreme-value asymptotics. The accuracy of the approximation for the extreme component is therefore inherently tied to how well the empirical tail is represented by the assumed asymptotic form in the finite-sample region of interest. In particular, the theoretical refinement is derived under the relatively restrictive tail condition

$$\frac{1}{3} < \xi < 1,$$

which limits it to a rather narrow class of distributions to which the extreme-value approximation directly applies.

A different classical refinement of the central limit theorem is the Edgeworth expansion and its quantile counterpart, the Cornish–Fisher expansion (Cornish and Fisher, 1937; Fisher and Cornish, 1960). The Edgeworth expansion corrects the Gaussian approximation to the distribution function by explicit polynomial terms in the standardized cumulants of the sum, and the Cornish–Fisher expansion inverts it to deliver corrected quantiles. These corrections are analytic and computationally nearly free. Applied globally to the whole sum, however, they inherit the structural limitation of the plain central limit theorem: when a small subset of summands dominates the higher moments, the expansion order required for accuracy grows, the corrected quantile function loses monotonicity, and the corrected density may even assign negative mass in the tails (Jaschke, 2002).

2. 2. The model

Our algorithm is inspired by Berry–Esseen theory and by the Edgeworth–Cornish–Fisher refinements of the central limit theorem.

For convenience, we assume throughout that $E(X_i) = 0$.

2.1 The Berry–Esseen Bound and the Lyapunov Ratio

Theorem 1 (Berry–Esseen; Petrov, 1995). *Let* $X_1, X_2, \dots, X_n$ *be independent random variables satisfying* $E[(X_i)] = 0$, $E[(X_i^2)] = \sigma_i^2 > 0$, $E(|X_i^3|) = \rho_i < \infty$.

Let $S_n = \frac{X_1+X_2+\cdots+X_n}{\sqrt{\sigma_1^2+\sigma_2^2+\cdots+\sigma_n^2}}$, *with distribution function* $F_n(x)$.

$$Then\ \sup_x |F_n(x) - \Phi(x)| \le C_0 \cdot \psi_0,$$

where $\Phi(x)$ *is the standard normal distribution function,*

$L = \psi_0(\vec{\sigma}, \vec{\rho}) = (\sum_{i=1}^n \sigma_i^2)^{-3/2} \cdot \sum_{i=1}^n \rho_i$, and C_0 is an absolute constant with C_0<0.5583(see Shevtsova ,2013).For convenience, we assume C_0=0.56

The quantity ψ_0 is called the *Lyapunov ratio* (or L-ratio). Since the Kolmogorov distance is shift invariant and scale invariant, for a general independent sum $X = \sum_{i=1}^n X_i$ and its normal approximation $Y \sim N(\mu, Var)$ (where $\mu = \sum E(X_i)], Var = \sum \sigma_i^2$), the upper bound on its KS distance can be written as

$$d_K(X, Y) \le 0.56 \cdot L_ratio, \quad where\ L_ratio = \frac{\sum_{i=1}^n \rho_i}{(\sum_{i=1}^n \sigma_i^2)^{3/2}}, \qquad (1)$$

($\rho_i = E|X_i|^3$ and $\sigma_i^2 = E(X_i{}^2)$. Although the Berry bound 0.56 · L-ratio is a very loose theoretical upper bound, numerical experiments reveal a strong positive relationship between the L-ratio (denoted by L) and the KS distance for random variables belonging to the same "scale-homogeneous group".

$KS \le C \cdot L$. This is *not* $KS \approx c \cdot L$. The bound is conservative:

Table 1. The Berry–Esseen bound versus the actual KS distance for a single variable.

Distribution	L (single)	Actual KS	BE bound (C = 0.56)	Conservativeness
Centered Rayleigh(1)	1.62	0.047	0.91	19 ×
Centered Laplace	2.12	0.062	1.19	119 ×
Centered Exponential	2.41	0.159	1.20	7.5 ×

The pointwise Edgeworth correction is linear in the standardized cumulants, although the KS distance, which is obtained by taking the supremum over x , need not inherit this linearity. Under suitable regularity conditions, the leading Edgeworth correction to the CDF of the standardized sum can be written formally as (Hall, 1992)

$$F_n(x) - \Phi(x) \approx -\frac{\gamma_1}{6}(x^2 - 1)\phi(x) + higher\ order\ terms,$$

where $\gamma_1 = E\,[X^3)/\big(E[X^2)\big)^{3/2}$ are the skewness. The KS distance is the supremum of this pointwise error over all x. The location of the supremum shifts with γ, making the overall relationship nonlinear in a complex way.

The Edgeworth expansion above corrects the distribution function, but risk applications typically require quantiles — the 90%, 95%, 99% and 99.9% quantiles studied in Section 6. Inverting the first-order Edgeworth term yields the Cornish–Fisher correction (see Abramowitz, M. and Stegun, I. A. (1964)): with $z_\alpha = \Phi^{-1}(\alpha)$, the α-quantile of the standardized sum is approximated by $w_\alpha = z_\alpha + \frac{\gamma_1}{6}({z_\alpha}^2 - 1)$. Equivalently, a standard normal draw Z can be transformed directly as

$$\tilde{X} = \mu + \sigma\left[Z + \frac{\gamma_1}{6}(Z^2 - 1)\right] \qquad (2)$$

which reproduces the distribution of the sum up to the first-order skewness correction. We call this distribution the *Normal with Cornish–Fisher adjustment* (CF-adjusted Normal): sampling from it requires only standard normal draws and a fixed quadratic transform. Only the first-order term is retained — for the 'Good groups'(see 2.3) the constructed γ_1 is small, and the higher-order terms contribute little. Indeed,the correction is a local asymptotic refinement and the polynomial is monotone only on a limited range; (see, e.g., Jaschke, 2002).

Note the difference between the skewness γ_1 and the L-ratio: the numerator of the L-ratio is the absolute third moment $E|X - \mu|^3$, whereas that of γ_1 is the signed third moment $E(X - \mu)^3$, standardized by the same power of the variance. Hence $|\gamma_1| \leq L_ratio$ always. The Good-group threshold on the L-ratio therefore automatically supplies an upper bound on the skewness entering the adjustment, $|\gamma_1| \leq \theta$: the same criterion that certifies the Gaussian approximation also certifies the validity of its Cornish–Fisher refinement.

Based on this, we adopt the L-ratio as a key quantity.A group is labeled *Good* when its Lyapunov ratio falls below a prescribed heuristic threshold. This label guides partitioning; it does not by itself certify the accuracy of either the Gaussian or the CF-adjusted approximation. Correspondingly ,if L ratio is greater than or equal to the threshold, the group is regarded as not well approximated even after the Cornish–Fisher adjustment; its distribution is then computed by Monte Carlo simulation, and the group is called a *Bad group*.

2.2 The Lyapunov Ratio of a Single Variable and of a Group

Figure 1. Lyapunov ratio of a single variable versus a group of random variables.

From $\left(E[X^2]\right)^{1/2} \leq (E|X|^3)^{1/3}$ it follows that $E\,|X|^3/\left(E[X^2]\right)^{3/2} \geq 1$, i.e., the L-ratio of a single random variable is at least 1. To visualize this, consider a coordinate system in which the x-axis is $\left(E[X^2]\right)^{3/2}$ and the y-axis is $E|X|^3$. A single random variable corresponds to a point in this coordinate system, and its L-ratio (l) is the slope of the line connecting that point to the origin.

Consider two independent zero-mean groups of random variables:

- Group 1: variance sum $V_1 = \sum_{i\in G_1} \sigma_i^2$, absolute third-moment sum $T_1 = \sum_{i\in G_1} \rho_i$
- Group 2: variance sum $V_2 = \sum_{j\in G_2} \sigma_j^2$,absolute third-moment sum $T_2 = \sum_{j\in G_2} \rho_j$

For the merged group:

$$l_1 = \frac{T_1}{V_1^{3/2}}, l_2 = \frac{T_2}{V_2^{3/2}}; V = V_1 + V_2, T = T_1 + T_2, l = \frac{T}{V^{3/2}}.$$

Substituting $T_1 = l_1 V_1^{3/2}$ and $T_2 = l_2 V_2^{3/2}$:

$$l = \frac{l_1 V_1^{3/2} + l_2 V_2^{3/2}}{(V_1+V_2)^{3/2}} = l_1 w^{3/2} + l_2 (1-w)^{3/2}, \qquad (3)$$

where $w = V_1/(V_1 + V_2) \in (0,1)$ is Group 1's share of the total variance.

For n random variables, the L-ratio $= \frac{\sum_{i=1}^n \rho_i}{\left(\sum_{i=1}^n \sigma_i^2\right)^{3/2}}$

Let $V_i = \sum_{j\in G_i} \sigma_j^2$, $w(i) = V_i/\sum_{i=1}^n V_i$, and $p(i) = w(i)^{3/2}$. The L-ratio after merging is

$$l = \sum_{i=1}^n l\,(i) \cdot p(i), \qquad (4)$$

where $w(i)$ is each group's variance-proportion weight, summing to 1, and $p(i)$ is the pseudo-weight, whose sum satisfies the following bounds.

Take $p = 3/2$, $q = 3$ (satisfying $1/p + 1/q = 1$) and apply Hölder to the sequences $\left(w(i)\right)_{i=1}^N$ and $(1)_{i=1}^N$:

$$\sum w(i) \cdot 1 \leq \left(\sum w(i)^{3/2}\right)^{2/3} \cdot (\sum 1^3)^{1/3}, \text{i.e.} 1 \leq \left(\sum p(i)\right)^{2/3} \cdot N^{1/3},$$

$$\text{hence} \sum_{i=1}^{n} p\,(i) \geq \frac{1}{\sqrt{N}},$$

with equality if and only if all $w(i)^{3/2}$ are equal, i.e., all $w(i) = 1/N$.

Since $w(i)^{3/2} \leq w(i)$ for $w(i) \in [0,1]$,

$$\sum_{i=1}^{N} p\,(i) \leq \sum_{i=1}^{N} w\,(i) = 1,$$

with equality if and only if some $w(i) = 1$ and all the others are 0.

Similarly , the skewness γ of merged n independent random variables is

$$\gamma = \sum_{i=1}^{n} \gamma\,(i) \cdot p(i)$$

2.3 Closure Property of Good Groups

From the upper bound on the pseudo weights in (3) and (4),we have

$l \leq max_i l(i), \qquad l \leq max(l_{G1}\,, l_{G2}\,)$ (5)

If a group of random variables has L-ratio defined in (1) <θ, then define it as Good group, otherwise define it as Bad group. From (5) ,we have proved merging independent good groups preserves the Good-group property. The resulting Lyapunov ratio is no larger than the maximum ratio among the component groups, and it decreases substantially when the variance contributions are sufficiently balanced. By the closure property of Good groups, multiple Good groups can be merged into a single Good group, whose sum is then approximated using the Cornish – Fisher-adjusted Normal distribution introduced in Section 2.1.

3. The Scale-Homogeneous Case and the Berry Measure

Definition 1. We say that a family of random variables belongs to centered *scale-homogeneous class* if there is a Base random variable X_0, with E(X_0)=0 , finite second and third absolute moment. And

$X(i) = k(i) \cdot X_{0,i}$, where $X_{0,i}$ are i.i.d copies of X_0

where k _i are non-zero real numbers

Many common mean-zero parametric families are scale-homogeneous (k>0):

- $k \cdot (Exp(1) - 1) \sim Exp(1/k) - \mathrm{k}$
- $k \cdot Laplace(0,1) \sim Laplace(0, |k|)$

- $k \cdot (Weibull(\beta, 1) - \Gamma(1 + 1/\beta)) \sim Weibull(\beta, k) - \mathrm{k} \cdot \Gamma(1 + 1/\beta)$

In particular, $Weibull(\beta = 1, \lambda)$ is $Exponential(\lambda)$, and $Weibull(\beta = 2, \lambda)$ is the $Rayleigh(\lambda)$ distribution.

Lyapunov ratio and skewness of a single variable (independent of k_i) are :

$$\lambda_0 = \frac{E|X_0|^3}{\left(E((X_0^2)\right)^{3/2}}, \quad \gamma_0 = \frac{E(\mathrm{X}_0)^3}{\left(E((X_0^2)\right)^{3/2}} \qquad (6)$$

The Lyapunov ratio(L-ratio) and skewness for *N scale-homogeneous* variables:

$$\mathrm{L}_{combined} = \lambda_0 \cdot \frac{\sum_{i=1}^{N} |k_i|^3}{\left(\sum_{i=1}^{N} k_i^2\right)^{3/2}}, \ \gamma = \gamma_0 \cdot \frac{\sum_{i=1}^{N} (\mathrm{k_i})^3}{\left(\sum_{i=1}^{N} k_i^2\right)^{3/2}} \qquad (7)$$

We defined *Berry measure as*

$$B(\{k_i\}) = \frac{\sum_{i=1}^{N} |k_i|^3}{\left(\sum_{i=1}^{N} k_i^2\right)^{3/2}} \qquad (8)$$

In the *scale-homogeneous setting ,L-ratio=*$\lambda_0 \times Berry\ measure.$

Also, let

$$\mathrm{S}(\{k_i\}) = \frac{\sum_{i=1}^{N} (\mathrm{k_i})^3}{\left(\sum_{i=1}^{N} k_i^2\right)^{3/2}} \qquad (9)$$

In other words, Skewness of group= $\gamma_0 \times S$

3.1 Properties of the Berry Measure

Berry measure has lower bound follows from Hölder's inequality. Taking $a_i = k_i^2, p = 3/2, q = 3$:

$$\sum a_i \leq \left(\sum a_i^{3/2}\right)^{2/3} \cdot N^{1/3} = \left(\sum |k_i|^3\right)^{2/3} \cdot N^{1/3}.$$

Raising both sides to the power $3/2$:

$$\left(\sum k_i^2\right)^{3/2} \leq \left(\sum |k_i|^3\right) \cdot \sqrt{N}, \quad \text{hence} \frac{\sum |k_i|^3}{\left(\sum k_i^2\right)^{3/2}} \geq \frac{1}{\sqrt{N}}, \qquad (10)$$

with equality if and only if all $|k_i|$ are equal.

For Upper bound, let $S = \sum_i k_i^2 > 0$ (if all $k_i = 0$ both sides are 0 and equality holds). Set $a_i = k_i^2/S$. Then $a_i \geq 0, \sum_i a_i = 1$, and since $0 \leq a_i \leq 1$ we have $a_i^{3/2} \leq a_i$. Therefore:

$$\frac{\sum|k_i|^3}{(\sum k_i^2)^{3/2}} = \frac{\sum(k_i^2)^{3/2}}{S^{3/2}} = \sum\left(\frac{k_i^2}{S}\right)^{3/2} = \sum a_i^{3/2} \leq \sum a_i = 1.$$

To sum up,

Let G be a group of random variables with scale parameters $\{k_i\}$, and write $n = |G|$.

$$B(G) = \frac{\sum_{i\in G}|k_i|^3}{(\sum_{i\in G} k_i^2)^{3/2}}, \qquad \frac{1}{\sqrt{n}} \leq B(G) \leq 1.$$

- **Lower bound:** attained when all k_i are equal.
- **Upper bound:** attained when one k_i dominates.
- **merge inequality:** if $G = G_1 \cup G_2$, then $B(G) \leq max\big(B(G_1), B(G_2)\big)$.

  This further validates the closure property for Good group.
- **Scale invariance:** $B(\{c \cdot k_i\}) = B(\{k_i\})$ for any real number c; B is invariant under global rescaling.

This further shows that only when one random variable's variance dominates the group does the merged Berry measure approach 1.

The Berry measure is invariant under global rescaling, $B(\{c \cdot k_i\}) = B(\{k_i\})$.For a fixed Base random variable the algorithm is therefore sensitive only to relative structural heterogeneity, not to absolute notional amounts. This renders the method naturally suited to scale-homogeneous products — such as ABS loan pools, insurance aggregates, and any portfolio of structurally similar, positively scaled risks.

3.2 Heuristic clustering on the Transformed Scale Parameters

From the bounds on B(G) (10), components with more similar scale parameters k{i} tend to produce a smaller Berry measure, making the corresponding group more likely to satisfy the Normal Edgeworth approximation criterion. Conversely, groups containing widely dispersed scale parameters are more likely to have a large Berry measure and therefore require further treatment.

More precisely, the Berry measure is invariant to rescaling of all scale parameters.

For example,

B(|k|, |10k|)=B(|10k|, |100k|)

This scale-invariance property should also be reflected in the distance measure used for clustering

Thus, applying a clustering algorithm directly to the original scale parameters would make the clustering result dependent on the absolute magnitude of the scales rather than only on their relative differences.

Also, the Berry measure depends on the scale parameters only through their absolute values, the clustering transformation is applied to $|k_i|$. The signs of k_i are retained when computing group skewness and simulating the Bad-group components.

A natural way to retain the scale invariance property is to apply a logarithmic transformation on the absolute value of scale parameter. Under the log transformation,

d(log|k|, log|10k|)=d(|log10k|, log|100k|),

However, in the present application, clustering is not used solely to identify scale-homogeneous groups. We are ultimately interested in accurately approximating the right tail of the sum, and components with substantially larger scale parameters can have a disproportionately large impact on extreme realizations. Therefore, it is desirable for the transformation to assign greater separation to increasingly large scale parameters. In particular, we would like the transformed distance to satisfy

d(f(|k|), f(|10k|))<d(f(|10k|), f(|100k|))<10d(f(|k|), f(|10k|)),

so that a component with scale 100k is more likely to be separated from the 10k group than a component with scale 10k is from the 1k group.

Here we first adopt the transform $f(x) = \sqrt{|x|}$.

Overall, the transformed scale serves two purposes: it reflects the scale-invariant structure of the Berry measure and provides a flexible mechanism for assigning greater clustering separation to large-scale components that are more relevant to right-tail behavior.

More generally, the choice of f is not unique. A broad class of transformations can be considered, as long as, for sufficiently large x,

Log|x|≤f(|x|)<|x|.

Such transformations provide greater separation for large-scale components than the logarithmic transformation while avoiding the strong dependence on absolute scale induced by the original parameter x.

The transformed-scale clustering is a heuristic preprocessing step designed to identify potentially scale-homogeneous groups. The Berry measure is then evaluated on each resulting group, and the final Good/Bad classification is determined by the Lyapunov-ratio threshold.

4. The Semi-Analytic Algorithm

1. For a scale-homogeneous family $X(i) = k(i) \cdot X_{0,i}$, compute the Berry measure and skewness for Base random variable X_0, $\lambda_0 and\ \gamma_0$ as defined in (6) .

2. Let $c(i) = \sqrt{|k(i)|}$ and apply K-means or hierarchical clustering (see, e.g., Hastie, Tibshirani, and Friedman, 2009) to $\{c(i)\}$.

3. Partition into $G_1, G_2, \dots, G_K$; compute each group's Lyapunov ratio via (7)

   L(j)= $\lambda_0 \cdot B\big(\{k(i)|\mathrm{k(i)} \in \mathrm{G_j}\}\big)$

4. If the group's L-ratio L(j)<θ, (θ is a parameter) then classify it as a Good group; otherwise mark it as a Bad group. By the closure property of Good group in (5), merging several Good groups results in a Big Good group G, which is approximated by the Normal with Cornish–Fisher adjustment introduced in Section 2.1: the mean and variance are the sums of the individual means and variances, also along with skewness computed using $\gamma_{\mathrm{G}} = \gamma_0 \cdot S(\{k_i\}|\mathrm{k(i)} \in \mathrm{G})$

   for CF-normal approximation .

   $$\tilde{X} = \mu + \sigma\left[Z + \frac{\gamma_{\mathrm{G}}}{6}(Z^2 - 1)\right]$$

   where μ,σ,γ are mean, standard deviation, skewness for the Big group, Z is drawn from standard normal distribution

5. The Bad group consists of the random variables $\{b_1 \cdot X_0, b_2 \cdot X_0, \dots, b_t \cdot X_0\}$.

6. The sum of all random variables is

   $$S = \tilde{X} + (b_1 \cdot X_0 + b_2 \cdot X_0 + \cdots + b_t \cdot X_0).$$

   Note that $\tilde{X}$ is sampled in closed form from the CF-adjusted Normal described in step 4, the second term requires only independent samples of X_0, multiplied by the corresponding b_i and then summed. Adding the two gives the distribution of S.

5. Error Analysis

5.1 Notation and decomposition of the error

Let:

G denote the true sum of the Good-group components;

B denote the sum of the Bad-group components;

G_{CF} denote the CF-adjusted Normal approximation to G;

S = G + B denote the true total sum;

SCF = G_{CF} + B denote the ideal semi-analytic approximation, in which the distribution of B is preserved exactly;

SCF_M denote the empirical distribution obtained from M independent simulations of SCF.

Throughout this section, assume that the Good-group sum and the Bad-group sum are independent. The notation SCF = G_{CF} + B refers to the idealized semi-analytic distribution; in the implemented algorithm, this distribution is estimated by simulation.

For two cumulative distribution functions (CDFs) F and H, define the Kolmogorov distance by

$$d_K(\mathrm{F}, \mathrm{H}) = \sup |\mathrm{F}(x) - \mathrm{H}(x)|.$$

The convolution inequality states that, for independent sums

$d_K(F_1 * F_2, H_1 * H_2) \le d_K(F_1, H_1) + d_K(F_2, H_2)$. In particular, because the Bad-group distribution is the same in S and S_CF,

$$d_K(\mathrm{S}, \mathrm{SCF}) \le d_K(\mathrm{G}, G_{CF}). \qquad (11)$$

Thus, at the level of the ideal distributions, convolution with the Bad-group sum does not increase the Kolmogorov error arising from the Good-group approximation.

5.2 Finite-simulation error

The implemented semi-analytic distribution is estimated using M independent simulations. By the Dvoretzky–Kiefer–Wolfowitz (DKW) inequality(Massart, P. (1990)), for any δ in (0, 1),

$$\mathrm{P}\{ d_K(\mathrm{SCF}, \mathrm{SCF_M}) > \sqrt{\frac{log(2/\delta)}{2M}} \} \le \delta. \qquad (12)$$

Equivalently, with probability at least 1 − δ, the empirical CDF based on M simulations is within $\sqrt{\frac{log(2/\delta)}{2M}}$ of the true semi-analytic CDF, uniformly over all x.

For example, at 99% confidence (δ = 0.01), the bound is approximately $1.628/\sqrt{M}$. If M = 10^6, this gives a uniform error bound of approximately 0.001628.

Combining (11), (12), and the triangle inequality yields the following overall bound: with probability at least 1 − δ,

$$d_K(\mathrm{S}, \mathrm{SCF_M}) \le d_K(\mathrm{G}, G_{CF}) + \sqrt{\frac{log(2/\delta)}{2M}}. \qquad (13)$$

This decomposition does not require the approximation error and the simulation error to be independent. It follows directly from the triangle inequality. The first term is the distributional approximation error of the Good groups,which will be investigated in 5.3; the second term is the finite-simulation error.

5.3 Accuracy of the CF-adjusted Normal approximation

The Berry–Esseen theorem provides a direct bound for the Kolmogorov distance between a standardized sum and its ordinary Normal approximation. Under the usual independence and finite-third-absolute-moment assumptions, this bound is proportional to the corresponding Lyapunov ratio. However, the Berry–Esseen bound does not directly provide the same guarantee for the CF-adjusted Normal approximation.

The Cornish–Fisher expansion gives a higher-order approximation to quantiles under suitable asymptotic regularity conditions. For a standardized sum, its first-order form can be written as

$$q_\alpha = z_\alpha + \frac{\gamma_1}{6}({z_\alpha}^2 - 1)$$

The first-order Cornish‑Fisher (CF) correction improves upon the ordinary Normal approximation in an asymptotic sense by explicitly accounting for the leading skewness effect. While the ordinary Normal approximation neglects this contribution, resulting in a leading quantile error typically of order $\mathrm{O}(\frac{1}{\sqrt{N}})$, the CF adjustment removes this leading-order term and can reduce the remaining quantile error to $\mathrm{O}(\frac{1}{N})$ under suitable Edgeworth regularity conditions. (Hall, 1992)Thus, the CF-adjusted Normal approximation provides a higher-order asymptotic approximation to the underlying distribution, although a uniformly smaller error for every finite sample size is not guaranteed.

This quantile result does not, by itself, imply that the entire CF-adjusted distribution has a smaller Kolmogorov distance than the ordinary Normal distribution. Accordingly, we do not claim a universal inequality of the form $d_K(\mathrm{G}, G_{CF}) \le d_K(\mathrm{G}, G_N)$, where G_N denotes the matching Normal approximation. Instead, the CF adjustment is motivated by its higher-order quantile correction, while its distributional performance is assessed empirically.

5.4 Pointwise variability of a simulated CDF

For completeness, let F be the CDF of a distribution being simulated and let F_N be its benchmark simulation CDF based on N independent draws. At any fixed x,

$$\mathrm{Var}[F_N(x)] = F(x)[1 - F(x)]/N \le \frac{1}{4N}.$$

$$\text{Therefore, } \mathrm{SD}[F_N(x)] \le \frac{1}{\sqrt{N}}.$$

For $N = 10^7$, this pointwise standard-deviation bound is approximately 1.58×10^{-4}. This is a pointwise statement: it controls the standard deviation at each fixed x, not the supremum of the error. A uniform statement requires a bound such as the DKW inequality(12).

In numerical experiments, we may use a sufficiently large benchmark simulation (for example, $N = 10^7$) as a practical proxy for the true distribution, for simplicity. This is an experimental convention rather than a claim that the empirical benchmark is mathematically identical to the true CDF.

6. Numerical Experiments

6.1 Experimental Setup

We use three different base random variables:

Table 2. The three base random variables used in the experiments.

Base distribution	λ_0	**Shape**
Weibull($\beta = 2$) (Rayleigh)	1.620	Slightly right-skewed
Laplace(1)	2.121	Symmetric but exponential tail
Exponential(1)	2.416	Right-skewed

Both θ and the number of clusters K are parameters that must be tuned to the specific data; there is no universal best value of θ or K. Our experiments find that $\theta = 0.6$ and $K = 3, K = 4$ work well, with both K-means and hierarchical clustering giving good results; below we show only K-means clustering.

The value of K is chosen according to the general shape of how the data clusters. If K is chosen too large, it tends to make many small isolated groups, making the Lyapunov ratio relatively high, so the group tends to be classified as a 'Bad group'**.**

Even if all groups are classified as Good group, we still need to simulate CF-normal approximation .Hence ,practically set M = max($5 \cdot 10^4 \cdot$ (number of simulated variables),10^6).

True distribution simulation in experiment is $M=10^7$

Note that all base variables have been centered.

Hardware setup:

Processor (CPU): AMD Ryzen 7 3700X 8-Core Processor (3.60 GHz base clock)

Memory (RAM): 16.0 GB

Operating System: Windows (64-bit)

No GPU acceleration was used. Wall times are reported for relative comparison only.

6.2 Small volume of Random variables: First Experiment, n = 50, List 1

The scale parameters $k(i)$ are taken from List 1 (Appendix A). Their distribution is shown in Figure 2 .

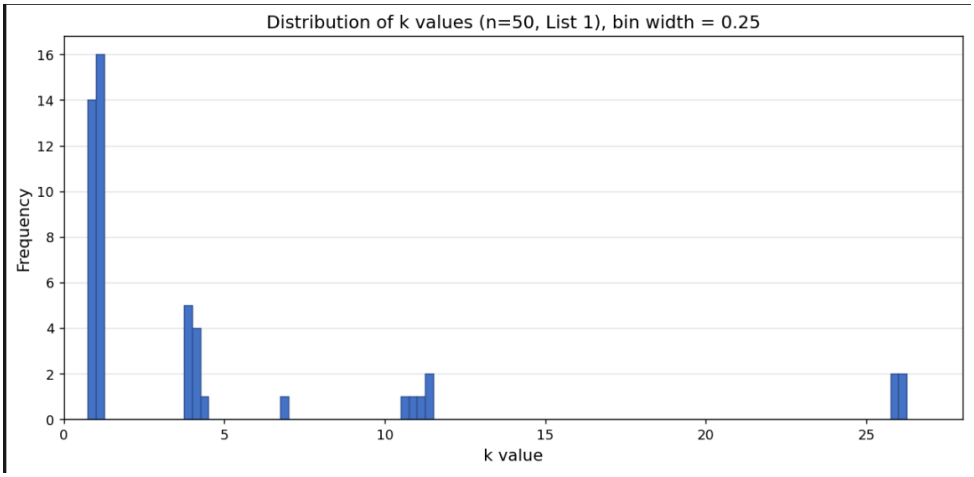


Figure 2. Distribution of $\{k_i\}$ *for* $n = 50$*, List 1.*

Feature of k_i: Trimodal distribution with a few very large isolated $k(i)$**.**

Number of Clusters K=3

Base random variable: Centered Rayleigh(1)

Table 3. Clustering result and KS comparison (base = Centered Rayleigh(1), List 1, K = 3).

Group	Count	Mean k	L-ratio	Variance	Decision (θ = 0.6)

				share	
Cluster 1	30	1.00	0.2980	0.85%	Good
Cluster 2	16	6.38	0.5349	23.00%	Good
Cluster 3	4	26.00	0.8100	76.15%	Bad

Table 4. KS distance comparison (base = Centered Rayleigh(1), List 1, K = 3).

Comparison	KS distance
Overall Normal vs True	0.0158
Semi-approx vs True	0.0022

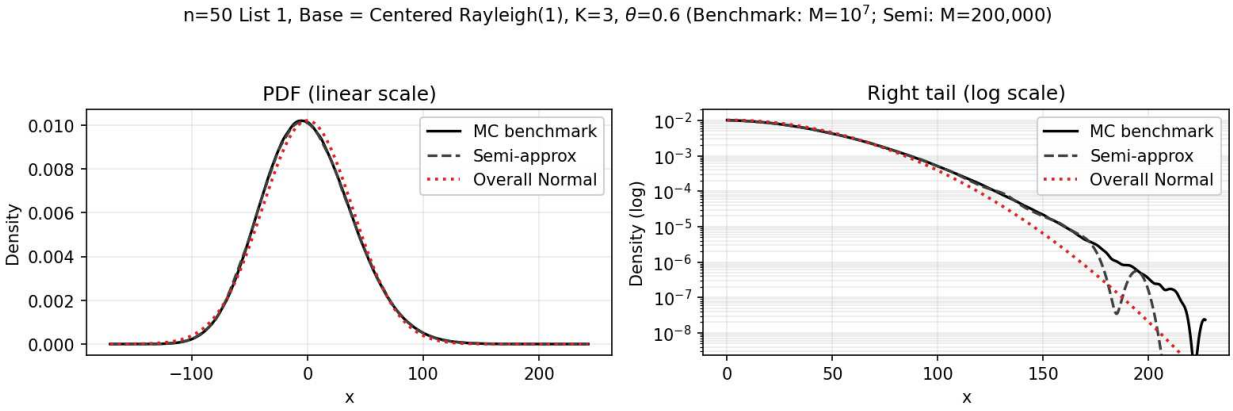


Figure 3. Semi-approximation vs. all Normal vs. true distribution (base = Centered Rayleigh(1), List 1, K = 3).

Table 5. Quantile comparison (base = Centered Rayleigh(1), $n = 50$, List 1).

Quantile	True	Semi-approx	Overall Normal	Semi error	Normal error
90%	51.039	51.136	50.032	0.19%	-1.97%
95%	66.762	66.671	64.215	-0.14%	-3.81%
99%	97.076	96.418	90.821	-0.68%	-6.44%
99.9%	132.320	131.342	120.643	-0.74%	-8.82%

Number of Clusters K=4

Base random variable: Centered Rayleigh(1)

Table 6. Clustering result and KS comparison (base = Centered Rayleigh(1), List 1, K = 4).

Group	Count	Mean k	L-ratio	Variance	Decision (θ = 0.6)

				share	
Cluster 1	30	1.00	0.2980	0.85%	Good
Cluster 2	11	4.25	0.5215	5.80%	Good
Cluster 3	5	11.05	0.7248	17.20%	Bad
Cluster 4	4	26.00	0.8099	76.15%	Bad

Table 7. KS distance comparison (base = Centered Rayleigh(1), List 1, K = 4).

Comparison	KS distance
Overall Normal vs True	0.0158
Semi-approx vs True	0.0013

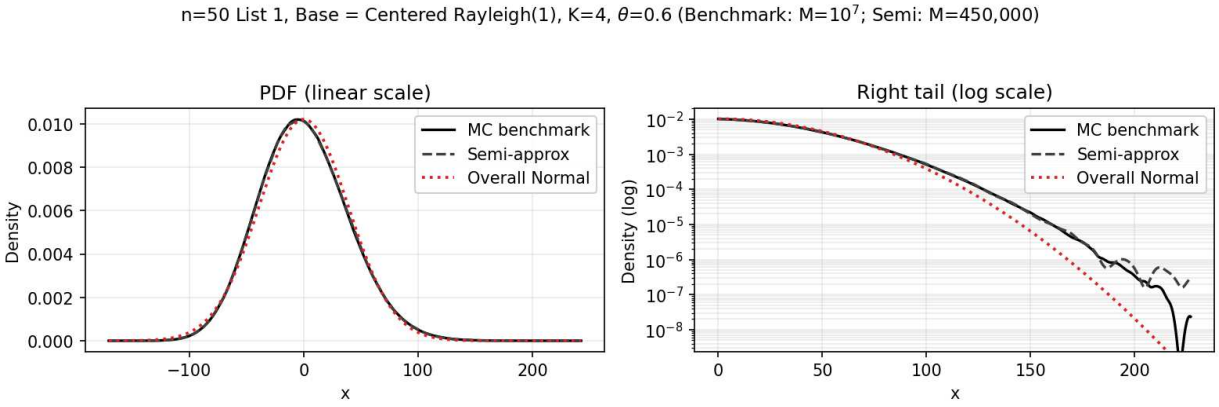


Figure 4. Semi-approximation vs. all Normal vs. true distribution (base = Centered Rayleigh(1), List 1, K = 4).

Table 8. Quantile comparison (base = Centered Rayleigh(1), List 1, K = 4).

Quantile	True	Semi-approx	Overall Normal	Semi error	Normal error
90%	51.039	50.942	50.032	-0.19%	-1.97%
95%	66.762	66.645	64.215	-0.18%	-3.81%
99%	97.076	97.317	90.821	+0.25%	-6.44%
99.9%	132.320	131.991	120.643	-0.25%	-8.82%

Number of Clusters K=3

Base random variable: Centered Laplace(0,1)

Table 9. Clustering result and KS comparison (base = Centered Laplace(0,1), List 1, K = 3).

Group	Count	Mean k	L-ratio	Variance share	Decision (θ = 0.6)
Cluster 1	30	1.00	0.3903	0.85%	Good
Cluster 2	16	6.38	0.7005	23.00%	Bad
Cluster 3	4	26.00	1.0607	76.15%	Bad

Comparison	KS distance
Overall Normal vs True	0.0102
Semi-approx vs True	0.0006

n=50 List 1, Base = Laplace(0,1), K=3, θ=0.6 (Benchmark: M=10^7; Semi: M=1,000,000)

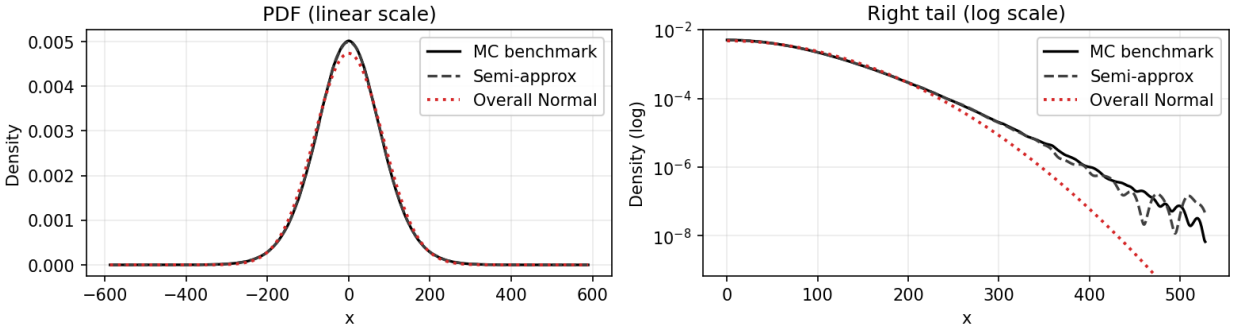


Figure 5. Semi-approximation vs. all Normal vs. true distribution (base = Centered Laplace(0,1), List 1, K = 3).

Table 10. Quantile comparison (base = Centered Laplace(0,1), List 1, K = 3).

Quantile	True	Semi-approx	Overall Normal	Semi error	Normal error
90%	105.452	105.529	108.002	+0.07%	+2.42%
95%	138.021	138.072	138.619	+0.04%	+0.43%
99%	204.277	204.020	196.051	-0.13%	-4.03%
99.9%	287.536	284.791	260.427	-0.95%	-9.43%

Number of Clusters K=4

Base random variable: Centered Laplace(0,1)

Table 11. Clustering result and KS comparison (base = Centered Laplace(0,1), List 1, K = 4).

Group	Count	Mean k	L-ratio	Variance share	Decision (θ = 0.6)
Cluster 1	30	1.00	0.3903	0.85%	Good
Cluster 2	11	4.25	0.6830	5.80%	Bad
Cluster 3	5	11.05	0.9493	17.20%	Bad
Cluster 4	4	26.00	1.0607	76.15%	Bad

Comparison	KS distance
Overall Normal vs True	0.0102
Semi-approx vs True	0.0006

n=50 List 1, Base = Laplace(0,1), K=4, θ=0.6 (Benchmark: M=10^7; Semi: M=1,000,000)

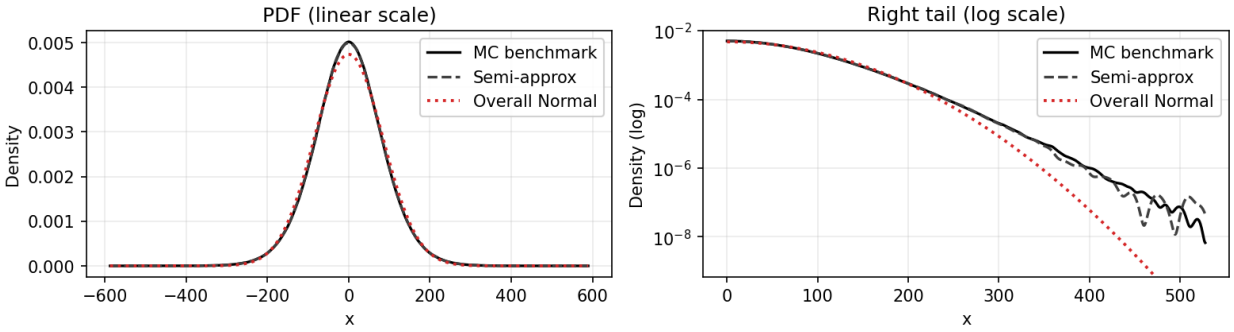


Figure 6. Semi-approximation vs. all Normal vs. true distribution (base = Centered Laplace(0,1), List 1, K = 4).

Table 12. Quantile comparison (base = Centered Laplace(0,1), List 1, K = 4).

Quantile	True	Semi-approx	Overall Normal	Semi error	Normal error
90%	105.452	105.529	108.002	+0.07%	+2.42%
95%	138.021	138.072	138.619	+0.04%	+0.43%
99%	204.277	204.020	196.051	-0.13%	-4.03%
99.9%	287.536	284.791	260.427	-0.95%	-9.43%

Number of Clusters K=3

Base Random variable :Centered Exp(1)

Table 13. Clustering result and KS comparison (base = Centered Exp(1), List 1, K = 3).

Group	Count	Mean k	L-ratio	Variance share	Decision (θ = 0.6)
Cluster 1	30	1.00	0.4442	0.85%	Good
Cluster 2	16	6.38	0.7973	23.00%	Bad
Cluster 3	4	26.00	1.2073	76.15%	Bad

Comparison	KS distance
Overall Normal vs True	0.0485
Semi-approx vs True	0.0006

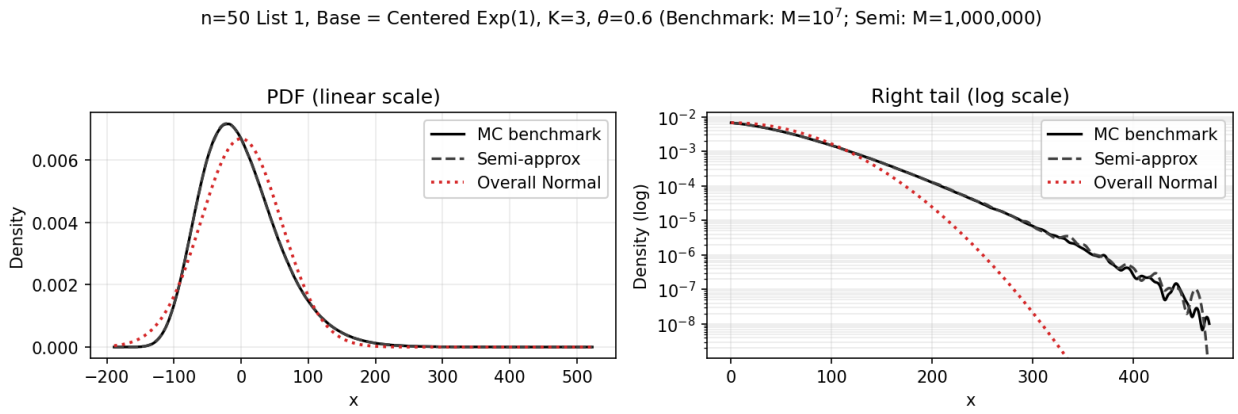


Figure 7. Semi-approximation vs. all Normal vs. true distribution (base = Centered Exp(1), List 1, K = 3).

Table 14. Quantile comparison (base = Centered Exp(1), List 1, K = 3).

Quantile	True	Semi-approx	Overall Normal	Semi error	Normal error
90%	79.029	79.004	76.369	-0.03%	-3.37%
95%	108.529	108.539	98.018	+0.01%	-9.68%
99%	170.389	170.794	138.629	+0.24%	-18.64%
99.9%	250.604	251.537	184.150	+0.37%	-26.52%

Number of Clusters K=4

Base Random variable :Centered Exp(1)

Table 15. Clustering result and KS comparison (base = Centered Exp(1), List 1, K = 4).

Group	Count	Mean k	L-ratio	Variance share	Decision (θ = 0.6)
Cluster 1	30	1.00	0.4442	0.85%	Good
Cluster 2	11	4.25	0.7774	5.80%	Bad
Cluster 3	5	11.05	1.0805	17.20%	Bad
Cluster 4	4	26.00	1.2073	76.15%	Bad

Comparison	KS distance
Overall Normal vs True	0.0485
Semi-approx vs True	0.0006

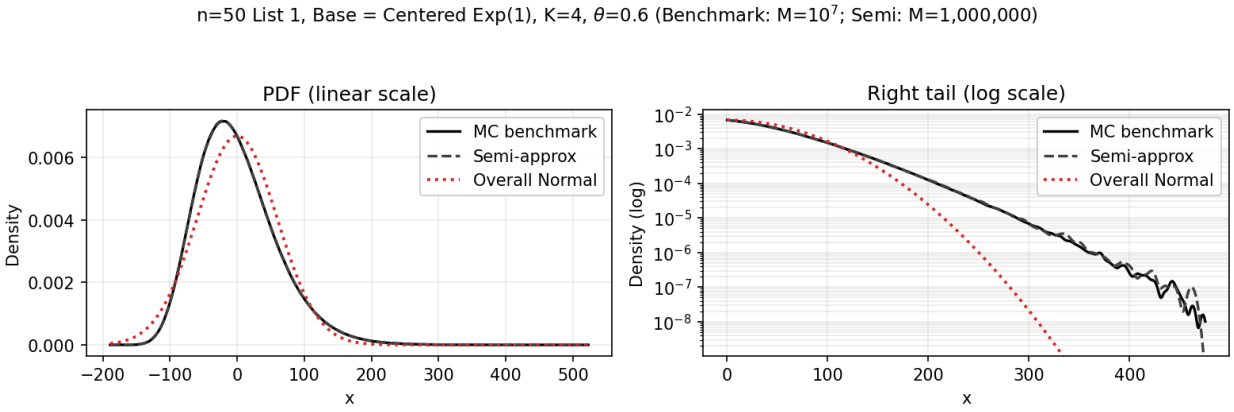


Figure 8. Semi-approximation vs. all Normal vs. true distribution (base = Centered Exp(1), List 1, K = 4).

Table 16. Quantile comparison (base = Centered Exp(1), List 1, K = 4).

Quantile	True	Semi-approx	Overall Normal	Semi error	Normal error
90%	79.029	79.004	76.369	-0.03%	-3.37%
95%	108.529	108.539	98.018	+0.01%	-9.68%
99%	170.389	170.794	138.629	+0.24%	-18.64%
99.9%	250.604	251.537	184.150	+0.37%	-26.52%

6.3 Summary and Conclusions

Table 17. Summary of all six runs ($\theta = 0.6$, true $M = 10^7$).

Base	K	Simulated	KS Semi	KS Normal	Semi 95% err	Normal 95% err	Semi 99.9% err	Normal 99.9% err
Rayleigh(1)	3	4	0.0022	0.0158	-0.14%	-3.81%	-0.74%	-8.82%
Rayleigh(1)	4	9	0.0013	0.0158	-0.18%	-3.81%	-0.25%	-8.82%
Laplace(0,1)	3	20	0.0006	0.0102	0.04%	-0.43%	-0.95%	-9.43%
Laplace(0,1)	4	20	0.0006	0.0102	0.04%	-0.43%	-0.95%	-9.43%
Exp(1)	3	20	0.0006	0.0485	0.01%	-9.68%	+0.37%	-26.52%
Exp(1)	4	20	0.0006	0.0485	0.01%	-9.68%	+0.37%	-26.52%

Across all three base distributions, the semi-analytic approximation outperforms all Normal approximation by large at both KS distance and right tail error.

To compare the proposed semi-analytic method with two alternative approaches−the Fast Fourier Transform (FFT) method and Müller’s extreme-value-theory-based approximation−we consider large collections of independent random variables with heterogeneous scale parameters, $X(i) = k(i) \cdot X_{0,i}$. Throughout the experiments,X_0 are fixed to follow a centered exponential distribution, and k(i) are generated from a mixture of Uniform distribution with seed=20261024 in Python

6.4 Large volume of random variables: Performance comparison with FFT and Müller inspired method

In FFT method ,the important parameter ‘a’ is the half-width of the truncation window in the x -domain: the FFT inversion represents the density on [−a,a] with N equally spaced points, so the grid spacing is h=2a/N and the frequency spacing is $\Delta\omega = \pi/a$. Probability mass outside the window is folded back into it (aliasing), hence a must extend well beyond the quantiles of interest; with N fixed by memory, enlarging a coarsens the resolution h − tail coverage and resolution cannot be refined independently.

The original method proposed by Müller(2019) applies only to heavy-tailed, Pareto-like distributions. However, it can be adapted into a semi-analytic approximation by simulating only the top 10% of random variables with the largest scale parameters and approximating the sum of the remaining variables as Normal.

The top-10% Müller-inspired benchmark is constructed for this experiment, rather than a direct implementation of Müller’s original asymptotic approximation.

Second experiment parameter setting

$$X(i) = k(i) \cdot X_{0,i}$$

k_parameter= (G1: 2500×U(9,11), G2: 10×U(50,100), G3: 300×U(300,400), G4: 200×U(400,600), n =3010)

Base R.V:X_0= Exp(1) centered; seed = 20261024; Number of cluster =4

Simulation number M = max(10^4 × (number of simulated variables),10^6).

Semi-analytic R.Vs cluster table. G2 are clustered as Bad group ,others are all Good

Table 18. Clustering result for the second experiment (n = 3010; G2 classified as the Bad group).

Cluster	# variables	k range	L = λ·B	Decision
C1	2500	[9.0, 11.0]	0.0485	Good
C2	10	[50.9, 98.4]	0.8263	Bad
C3	335	[300.6, 432.3]	0.1337	Good
C4	165	[433.9, 600.0]	0.1907	Good

Comparison between different methodology

Table 19. Comparison with FFT and Müller-inspired methods (second experiment, n = 3010).

Method	Variables simulated	M	Wall time (second)	KS distance	95% error	99% error	99.9% error
Semi-Analytic	10	1,000,000	0.2	0.0011	0.09%	0.06%	0.01%
Muller-inspired (top 10%)	301	3,010,000	9.6	0.0013	-0.43%	-0.59%	-1.24%
FFT (N=2^14, a=4sd)	0	— (4.93e+07 CF evals)	2.8	0.0006	-0.09%	-0.12%	-0.76%
FFT (N=2^17, a=4sd)	0	— (3.95e+08 CF evals)	21.4	0.0004	-0.07%	-0.11%	-0.75%
Overall Normal	0	1,000,000	0.0	0.0066	-1.65%	-3.13%	-4.89%

In this example, both FFT and Müller-inspired are computationally more expensive ,semi-analytic simulates only 1/30 as many random variables as Müller-inspired with remarkable accuracy. This matters in practice, one reason is the number of R.V required for simulation

grows linearly with n, second is when the input CDFs are themselves estimated with error, and these errors accumulate in proportion to the number of simulated variables.

Third experiment parameter setting

k_parameter= (G1: 1000×U(9,11), G2: 1000×U(20,40), G3: 20×U(200,400),n=2020)

X_0= Exp(1) centered; seed = 20261024; Number of cluster =4

Semi-analytic R.Vs cluster table .All clusters are Good

Table 20. Clustering result for the third experiment (n = 2020; all groups classified as Good).

Cluster	# variables	k range	L = λ·B	Decision
C1	1000	[9.0, 11.0]	0.0767	Good
C2	467	[20.0, 29.5]	0.1138	Good
C3	533	[29.5, 40.0]	0.1057	Good
C4	20	[207.8, 392.1]	0.5719	Good

Table 21. Comparison with FFT and Müller-inspired methods (third experiment, n = 2020).

Method	Variables simulated	M (replications)	Wall time (second)	KS distance	95% error	99% error	99.9% error
Semi-Analytic	0	1,000,000 (fully analytic)	0.2	0.0028	0.42%	-0.13%	-0.89%
Muller-inspired (top 10%)	202	2,020,000	4.9	0.0007	-0.18%	-0.47%	-0.39%
FFT (N=2^14, a=4sd)	0	— (3.31e+07 CF evals)	1.8	0.0008	-0.14%	-0.48%	-2.42%
FFT (N=2^17, a=4sd)	0	— (2.65e+08 CF evals)	14.1	0.0007	-0.11%	-0.46%	-2.40%
Overall Normal	0	1,000,000	0.0	0.0161	-3.72%	-7.57%	-11.49%

In this experiment, the semi-analytic method requires no direct simulation of the original summands and achieves reasonably good accuracy in the right tail. Müller-inspired methods achieve good accuracy on right tail but computationally less efficient .The FFT methods require more computation in this experiment, while their 99.9th-percentile errors remain comparatively large

6.5 Discussion on numerical experiment

The three experiment illustrate an important property of semi-analytic approximation: **the suitability of Gaussian aggregation is determined by the relative heterogeneity of the scale parameters within a group, rather than by their absolute magnitudes alone**. In the second experiment, the intermediate-scale group fails the prescribed Berry-measure criterion and is therefore retained for direct simulation, whereas both the small-scale group and the substantially larger but more homogeneous group satisfy the criterion and are approximated by Gaussian distributions with CF adjustment.

Across the configurations examined, the semi-analytic method substantially improves upon the overall Normal approximation in several distributional and upper-quantile error metrics. Its relative performance against FFT and the Müller-inspired benchmark depends on the scale configuration and the accuracy metric considered.

7. Conclusion

This paper has proposed and systematically studied the semi-analytic approximation — a hybrid normal-approximation algorithm for sums of independent scale-homogeneous random variables $S = \sum_i X\,(i)$ with $X(i) = k(i) \cdot X_{0,i}$ — grounded in the Berry–Esseen theorem. The central idea is to use the Berry measure $B(\{k_i\}) = \sum |k_i|^3/(\sum k_i^2)^{3/2}$ as a criterion for non-normality: the variables are grouped by clustering $\{\sqrt{k(i)}\}$, and the L-ratio $\gamma = \lambda_0 \cdot B$ is compared against a threshold θ to identify the variables genuinely responsible for the deviation from normality (Bad groups), which are then simulated exactly, while merging the Good groups is guaranteed to yield a Good group.

The error analysis reveals that Bad groups do not add a separate distributional approximation term when their distribution is preserved exactly in the ideal construction.

Numerical experiments with various sets of $k(i)$ also suggest that the semi-approximation works exceptionally well: both the KS distance and the right-tail quantile errors decrease substantially.

It should also be noted that although the base random variables in our demonstration are all parametric, our algorithm also works for any base random variable with only an empirical distribution: first define the base random variable by setting

$X_0 = (X - E(X))/\sigma(X)$, and then $|k(i)| = \sigma(X_i)$.The sign of k(i) depends on skewness. Once having $k(i)$, we can apply the clustering algorithm as before and Monte Carlo simulation to generate samples for the Bad group.

For real-world applications, suppose that a collection of independent random variables

$\{Y_i\}$ satisfies

$Y_i = \mathrm{u_i} + k(i) \cdot X_{0,i}$, where $X_{0,i}$ are i.i.d copies of mean zero X_0 ,

$\mathrm{u_i}$=E(Y_i),The distribution of X_0 is specified either parametrically or empirically.

We can apply the proposed semi-analytic algorithm to the centered, scale-homogeneous components $\mathrm{X_i} = k(i) \cdot X_{0,i}$ using the corresponding distribution of $\mathrm{X_0}$ to estimate the required moments and to simulate the components assigned to Bad groups. The original aggregate sum can then be recovered by adding the deterministic location shift:

$S=\sum_{i=1}^{n} Y_i \overset{d}{=} \sum_{i=1}^{n} \mathrm{X_i} + \sum_{i=1}^{n} \mathrm{u_i}$

This formulation extends the practical applicability of the framework beyond fully specified parametric models, allowing it to accommodate empirically estimated component distributions while retaining the same scale-based partitioning principle.

Appendix A: Scale Parameters Used in the Experiments

$N = 50$

List 1:

[0.88, 0.95, 1.02, 0.92, 1.08, 0.98, 1.05, 0.90, 1.10, 0.96, 1.12, 0.94, 1.00, 1.06, 0.99, 1.03, 0.91, 1.07, 0.97, 1.04, 1.15, 1.01, 0.93, 1.09, 0.89, 1.11, 1.00, 0.95, 1.02, 0.98, 3.80, 4.20, 3.90, 4.10, 3.85, 4.15, 3.95, 4.05, 3.75, 4.25, 25.80, 26.20, 25.90, 26.10, 10.70, 11.30, 10.85, 11.15, 6.75, 11.25]

References

Durrett.R (2019). *Probability: Theory and Examples*, 5th ed. Cambridge University Press, Cambridge.

V. V. Petrov (1995). *Limit Theorems of Probability Theory*, Oxford Science Publications.

Gibbs, A. L. and Su, F. E (2002). On choosing and bounding probability metrics. *International Statistical Review* 70(3), 419–435.

Hastie, T., Tibshirani, R, and Friedman, J (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, New York.

Shevtsova, Irina (2013). "On the absolute constants in the Berry‑Esseen inequality and its structural and nonuniform improvements".

Informatika i Ee Primeneniya [Informatics and its Applications], 2013, Volume 7, Issue 1, Pages 124 ‑ 125

Embrechts, P. and Frei, M (2009). Panjer recursion versus FFT for compound distributions. *Mathematical Methods in Operations Research* 69, 497–508.

Schaller, P. and Temnov, G (2008). Efficient and precise computation of convolutions: applying FFT to heavy tailed distributions. *Journal of Computational Methods in Applied Mathematics* 8, 187–200.

UlrichK. Müller (2019). Refining the Central Limit Theorem Approximation via Extreme Value Theory. Statistics & Probability Letters,Dec 2019.

Cornish, E. A. and Fisher, R. A (1937). Moments and cumulants in the specification of distributions. *Revue de l'Institut International de Statistique* 5(4), 307–320.

Fisher, R. A. and Cornish, E. A (1960). The percentile points of distributions having known cumulants. *Technometrics* 2(2), 209–225.

Jaschke, S. R (2002). The Cornish–Fisher expansion in the context of delta-gamma-normal approximations. *Journal of Risk* 4(4), 33–52.

Abramowitz, M. and Stegun, I. A (1964). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 935.

Peter Hall (1992).The Bootstrap and Edgeworth Expansion ,15-28.

Massart, P. (1990). "The Tight Constant in the Dvoretzky–Kiefer–Wolfowitz Inequality." The Annals of Probability, 18(3), 1269–1283.

Ethics approval: Not applicable. This article does not contain any studies with human participants or animals performed by the author.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- semiedgeworthn50threebases.txt
- semiedgeworthlarge.py