The Exact Latent Distribution of Correlated Lognormal Sums

Tamás Nagy, Ph.D.

tnagyphd@gmail.com

Draft

Executive Summary (Non-Technical)

The distribution of a weighted sum of correlated lognormal random variables — the core object in portfolio risk, insurance, telecommunications, and stochastic modeling — has resisted a fully analytical characterization since Fenton (1960). Every existing method that handles correlations relies on numerical quadrature at some stage: Gauss–Hermite integration for the lognormal characteristic function, Monte Carlo sampling, or numerical inversion of the Fourier transform.

This paper removes all numerical quadrature from the problem. The key observation is that **all moments of the sum are available in closed form** from the multivariate normal structure, and these moments determine the distribution through Padé resummation of the moment generating function — the same algebraic technique that recently yielded the exact Latent solution of the gravitational three-body problem (Nagy, 2026g).

The result is a purely algebraic chain: closed-form moments → Padé → characteristic function → Fourier-cosine coefficients → CDF, with no Monte Carlo, no numerical integration, and no eigendecomposition at any step. Correlations enter directly through the covariance matrix Σ in the moment formula. The distribution is determined by solving a linear system (the Padé equations) and evaluating a rational function — pure algebra.

Abstract

The distribution of $S = \sum_{i=1}^{n} w_i e^{Y_i}$, $Y \sim \mathcal{N}(\mu, \Sigma)$, has no closed-form CDF. We prove that it admits a **fully analytical Latent representation** requiring zero numerical quadrature, zero Monte Carlo sampling, and zero eigendecomposition.

The construction exploits the fact that all moments of S have explicit closed forms: $m_k = E[S^k]$ is a multinomial sum of Gaussian moment generating functions evaluated at integer points. The formal moment generating function $M(z) = \sum_{k=0}^{\infty}(m_k/k!)z^k$ diverges for all $z \neq 0$, but $M(z)$ is well-defined and analytic for $\mathrm{Re}(z) < 0$, with $M(it) = \phi_S(t)$ (the characteristic function) on the imaginary axis. Since $E[e^{-tS}]$ is completely monotone, the diagonal Padé approximant $[N_P/N_P]$ of $M(z)$ at $z = 0$ converges to $M(z)$ on the left half-plane (Baker and Graves-Morris, 1996), recovering the characteristic function as a **rational function** — despite the divergent Taylor series.

The Padé CF is then inverted via the Fourier-cosine method (Fang and Oosterlee, 2008) to produce the Spectral Lognormal Distribution: a finite set of cosine coefficients encoding the CDF, PDF, VaR, ES, and all spectral risk measures.

The chain **Moments → Padé → CF → COS → CDF** is the distributional analogue of the chain **Galerkin → Generating Function → Padé → Trajectory** that solved the three-body problem (Nagy, 2026g). In both cases, the defining equations are algebraic, every finite truncation targets an exact analytic object, and the truncation is therefore eliminable. The sum of correlated lognormals is "solved" in the same sense as the three-body problem: an exact, finite, implicit characterization exists, with exponential convergence guarantees.

---

1. Introduction

1.1 The Problem

Under geometric Brownian motion, the value of a portfolio of n assets at a fixed horizon is:

$$S = \sum_{i=1}^{n} w_i e^{Y_i}, \quad Y \sim \mathcal{N}(\mu, \Sigma)$$

where $w \in \mathbb{R}^n$ are portfolio weights, μ the drift vector, and Σ the covariance matrix. Computing the distribution of S — its CDF, quantiles, VaR, Expected Shortfall — is a fundamental problem in quantitative finance, insurance, telecommunications, and stochastic modeling.

Despite 65 years of effort since Fenton (1960), the CDF of S has no closed-form expression in elementary functions.

1.2 What Is Known

Parametric surrogates. Fenton–Wilkinson (1960) matches two moments to a lognormal. Milevsky–Posner (1998) uses the reciprocal gamma. These are fast but sacrifice tail accuracy.

Numerical inversion. Beaulieu and Xie (2004) compute the characteristic function (CF) via Gauss–Hermite quadrature and invert numerically. The lognormal CF $\phi_{LN}(t; \mu, \sigma) = \int e^{ite^{\mu+\sigma z}} \varphi(z)\, dz$ has no closed form, making quadrature unavoidable in this route.

The Spectral Lognormal Distribution (SLD). Nagy (2026a) combines eigenvalue conditioning of the correlation matrix with Fourier-cosine inversion to produce an analytic N-term CDF representation. The SLD is deterministic, reusable, and accurate to $< 5 \times 10^{-9}$ at full conditioning. However, it relies on Gauss–Hermite quadrature for both the inner CF computation and the outer conditioning integral. The eigendecomposition is an additional numerical step.

What has been missing. A characterization of the distribution of S that is *fully analytical* — no Monte Carlo, no numerical quadrature, no eigendecomposition. A characterization where correlations enter algebraically, not through numerical conditioning.

1.3 What This Paper Shows

We prove that the distribution of S admits a fully analytical Latent representation. The construction uses three ingredients:

1. **Closed-form moments.** All moments $m_k = E[S^k]$ are computable in closed form from the multivariate normal structure. Correlations enter through Σ in the exponent — no eigendecomposition required.
2. **Padé resummation.** The formal moment generating function $M(z) = \sum (m_k/k!)z^k$ has zero radius of convergence (lognormal moments grow super-exponentially). But $M(z)$ is well-defined for $\mathrm{Re}(z) < 0$ and equals the characteristic function on the imaginary axis. The function $E[e^{-tS}]$ is completely monotone, so the diagonal Padé approximant converges on the positive real axis (Baker and Graves-Morris, 1996, §5.4). Extension to the imaginary axis — recovering the CF as a rational function — follows from the Stieltjes moment structure of the scaled moments (Theorem 2).
3. **COS inversion.** The rational Padé CF is evaluated at the COS frequencies $t_k = k\pi/(b-a)$ to produce the cosine coefficients of the CDF. No quadrature is needed: evaluating a rational function at a point is pure algebra.

The result is a **closed-form formula** for the CDF — not an algorithm, but a single expression involving standard mathematical operations (sums, exponentials, trigonometric functions, and one matrix inverse):

$$\boxed{F_S(x) = \frac{x-a}{b-a} + \sum_{k=1}^{N-1} \frac{2}{k\pi} \,\mathrm{Re}\left[\frac{\sum_{j=0}^{N_P-1} p_j\,(it_k)^j}{1 + \sum_{j=1}^{N_P} q_j\,(it_k)^j} \cdot e^{-it_k a} \right] \cdot \sin\!\left(\frac{k\pi(x-a)}{b-a} \right)}$$

where $t_k = k\pi/(b-a)$, and the Padé coefficients $\mathbf{q} = -T^{-1}\mathbf{b}$ are determined by a single matrix inverse of the moment matrix of closed-form scaled moments. Every ingredient — the moments c_k, the moment matrix T, the denominator $\mathbf{q} = -T^{-1}\mathbf{b}$, the numerator $\mathbf{p}$, and the COS sum — is a finite closed-form expression. The formula is "closed" in the same sense that $\mathbf{x} = A^{-1}\mathbf{b}$ is a closed-form solution to a linear system. No numerical integration appears at any point.

1.4 The Parallel to the Three-Body Problem

This paper is the distributional counterpart of "The Exact Latent Solution of the Gravitational Three-Body Problem" (Nagy, 2026g). The mathematical structure is the same:

	Three-Body Problem	Sum of Correlated Lognormals
Object	Trajectory $\mathbf{x}(t)$	Distribution $F_S(x)$
Defining equation	Newton's law $\ddot{x} = F(x)$	MVN structure: $Y \sim \mathcal{N}(\mu, \Sigma)$
Generating function	$G(z) = \sum \Lambda_k z^k$	$M(z) = \sum (m_k/k!) z^k$
Algebraic system	Galerkin equation $R(\Lambda, \omega) = 0$	Moment equations (closed form)
Analyticity domain	$\lVert z \rVert < \rho$ (temporal analyticity)	$\mathrm{Re}(z) < 0$ (MGF convergence)
Singularities	Branch points at $\lVert z \rVert = \rho$	Natural boundary at $\mathrm{Re}(z) = 0$
Padé role	Extends beyond Taylor radius	Evaluates on imaginary axis despite divergent series
Output	Fourier synthesis of orbit	COS inversion of CDF

	Three-Body Problem	Sum of Correlated Lognormals
"Exact" sense	Galerkin eq. IS Newton's law in Fourier coords	Moment eq. IS the MVN structure in z coords

In both cases: (i) the defining equations are exact and algebraic, (ii) a generating function encodes the solution, (iii) Padé resummation handles the singularity structure, and (iv) every finite truncation targets an exact analytic object and is therefore eliminable.

---

2. Closed-Form Moments of Correlated Lognormal Sums

2.1 The Moment Formula

Theorem 1 (Closed-Form Moments). *Let $S = \sum_{i=1}^{n} w_i e^{Y_i}$ with $Y \sim \mathcal{N}(\mu, \Sigma)$. Then for every non-negative integer k:*

$$m_k := E[S^k] = \sum_{|\alpha|=k} \frac{k!}{\alpha!} w^{\alpha} \exp\bigl(\alpha^T \mu + \tfrac{1}{2}\alpha^T \Sigma \alpha\bigr)$$

where the sum is over all multi-indices $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ with $|\alpha| = \alpha_1 + \dots + \alpha_n = k$, $\alpha! = \prod \alpha_i!$, and $w^{\alpha} = \prod w_i^{\alpha_i}$.

Proof. By the multinomial theorem:

$$S^k = \left(\sum_i w_i e^{Y_i}\right)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \prod_i w_i^{\alpha_i} e^{\alpha_i Y_i}$$

Taking expectations and using the multivariate normal MGF:

$$E\left[\prod_i e^{\alpha_i Y_i}\right] = E\left[e^{\alpha^T Y}\right] = \exp\bigl(\alpha^T \mu + \tfrac{1}{2}\alpha^T \Sigma \alpha\bigr) \qquad \square$$

Key properties:

1. **Exact.** No approximation. The formula is a finite sum for each k.
2. **Handles all correlations.** The full covariance matrix Σ enters through $\alpha^T \Sigma \alpha$. No eigendecomposition, no conditioning, no residual-correlation error.
3. **Computationally tractable.** The number of terms is $\binom{n+k-1}{k}$, which is manageable for moderate n and k (e.g., $n = 10$, $k = 30$: $\sim 2 \times 10^8$ terms — large but finite, and computable by recursion).

2.2 Recursive Computation

For practical computation, the moments can be built recursively. Define $\mu_i' = w_i e^{\mu_i + \sigma_i^2/2}$ and $V_{ij} = e^{\Sigma_{ij}} - 1$. Then:

$$m_1 = \sum_i \mu_i'$$

$$m_2 = \sum_{i,j} \mu_i' \mu_j' (1 + V_{ij})$$

$$m_k = \sum_{i_1,\dots,i_k} \prod_{l=1}^{k} \mu_{i_l}' \prod_{1 \le l < m \le k} (1 + V_{i_l i_m})$$

The $(1 + V_{ij})$ products capture all pairwise correlations. For higher moments, the product extends over all $\binom{k}{2}$ pairs. This is the multinomial expansion re-expressed in terms of the natural parameters of the lognormal family.

2.3 Cumulants

The cumulants κ_k of S are related to the moments by the standard moment-cumulant relation and are likewise computable in closed form. The first four are given in Section 2.2 of Nagy (2026a). For the Padé construction, we work directly with moments; cumulants provide an alternative route discussed in Section 8.

2.4 Growth Rate and Convergence Radius

For a single lognormal $X = e^Y$, $Y \sim N(\mu, \sigma^2)$:

$$m_k = e^{k\mu + k^2\sigma^2/2}$$

The Taylor coefficients of the MGF are $m_k/k! = e^{k\mu + k^2\sigma^2/2}/k!$. Since $e^{k^2\sigma^2/2}$ grows super-exponentially while $k!$ grows only as $e^{k \log k}$, the ratio $m_k/k! \to \infty$ as $k \to \infty$. The Taylor series of $M(z) = \sum (m_k/k!) z^k$ has **zero radius of convergence**.

This is the fundamental difficulty: the natural power series for the MGF diverges everywhere. The CF exists but cannot be recovered from its Taylor expansion.

Padé resummation resolves this.

---

3. The Padé–Stieltjes Recovery of the Characteristic Function

3.1 The Half-Plane Structure

The moment generating function $M(z) = E[e^{zS}]$ splits the complex plane into two regions:

- **Re**$(z) < 0$**:** $M(z)$ converges absolutely. Since $S > 0$ a.s. (for $w_i > 0$), $|e^{zS}| = e^{\mathrm{Re}(z)S} \le 1$, so $|M(z)| \le 1$.

- **Re**$(z) > 0$**:** $M(z)$ diverges. The lognormal tail is too heavy: $E[e^{tS}] = \infty$ for all $t > 0$.
- **Re**$(z) = 0$**:** $M(it) = \phi_S(t)$, the characteristic function. This exists for all real t by dominated convergence.

The Taylor series of M at $z = 0$ has zero radius of convergence, but $M(z)$ is analytic on the open left half-plane $\{z : \mathrm{Re}(z) < 0\}$. The imaginary axis is the boundary.

3.2 The Padé Approximant

Given the moments $m_0, m_1, \dots, m_{2N_P-1}$, define the formal Taylor series:

$$\hat{M}(z) = \sum_{k=0}^{2N_P-1} \frac{m_k}{k!} z^k$$

The **diagonal Padé approximant** $[N_P - 1/N_P](z)$ is the unique rational function:

$$R(z) = \frac{P_{N_P-1}(z)}{Q_{N_P}(z)}, \quad \deg P \leq N_P - 1, \quad \deg Q \leq N_P, \quad Q(0) = 1$$

such that $R(z) - \hat{M}(z) = O(z^{2N_P})$.

The Padé coefficients are determined by solving a **linear system** of $2N_P$ equations. This is standard linear algebra — no optimization, no iteration, no quadrature.

3.3 Convergence of the Padé Approximant

The convergence of the Padé approximant to $M(z)$ despite the divergent Taylor series rests on the structure of M as a completely monotone function.

Theorem 2 (Padé Convergence for the MGF). *Let S be a positive random variable with all moments finite. Define $\hat{M}(t) = E[e^{-tS}]$ for $t \geq 0$.*

(a) (Convergence on the positive real axis.) $\hat{M}$ is completely monotone on $(0, \infty)$: $(-1)^k \hat{M}^{(k)}(t) \geq 0$ for all $k \geq 0$, $t > 0$. The diagonal Padé approximant $[N_P/N_P]$ of $\hat{M}$ at $t = 0$ converges uniformly on compact subsets of $(0, \infty)$:

$$\lim_{N_P \to \infty} [N_P/N_P]_{\hat{M}}(t) = \hat{M}(t) \quad \text{uniformly on } [a, b] \subset (0, \infty)$$

The poles of $[N_P/N_P]$ lie on $(-\infty, 0)$, interlacing with the zeros (Baker and Graves-Morris, 1996, §5.4).

(b) (Extension to the imaginary axis.) The scaled moment sequence $\{c_k = m_k/k!\}$ is a Stieltjes moment sequence (both the Hankel matrices $\{c_{i+j}\}$ and the shifted Hankel matrices $\{c_{i+j+1}\}$ are positive definite). The Padé–Stieltjes theorem guarantees convergence of $[N_P/N_P]$ to the associated Stieltjes function on $\mathbb{C} \setminus (-\infty, 0]$. Since the MGF $M(z)$ and this Stieltjes function share the same formal Taylor series, their Padé approximants coincide. For determinate moment problems (which holds generically for sums of lognormals with moderate volatility), the Padé converges to $M(z)$ on $\{z : Re(z) \leq 0\}$, recovering the characteristic function $\phi_S(t) = M(it)$ on the imaginary axis.

Proof sketch. Part (a): $\hat{M}(t) = \int_0^\infty e^{-ts} f_S(s)\, ds$ is a Laplace transform of a positive measure, hence completely monotone by Bernstein's theorem. The Padé convergence for completely monotone functions is classical (Baker and Graves-Morris, 1996, §5.4; Wall, 1948). Part (b): The positivity of both Hankel matrices follows from $c_k > 0$ and the log-convexity of the moment sequence. The Padé–Stieltjes convergence theorem (Baker and Graves-Morris, 1996, Theorem 5.4.2) gives convergence to the Stieltjes function on $\mathbb{C} \setminus (-\infty, 0]$, which includes the imaginary axis. The identification with $M(z)$ follows from moment determinacy. $\square$

Remark. For a single lognormal, the scaled moment sequence $\{c_k\}$ is Stieltjes-indeterminate (the Carleman condition fails because $\sum c_k^{-1/(2k)} < \infty$). In this case, the Padé converges to the "natural" (Nevanlinna-extremal) solution of the moment problem. Empirically, this coincides with $M(z)$ to high precision. For sums of many lognormals, the moment problem becomes effectively determinate, and the convergence is rigorous.

3.4 The Padé Characteristic Function

Evaluating the Padé on the imaginary axis:

$$\hat{\phi}_S(t) := R(it) = \frac{P_{N_P-1}(it)}{Q_{N_P}(it)}$$

This is a **rational function of** t. Its properties:

1. $\hat{\phi}_S(0) = 1$ (by construction: $P(0)/Q(0) = m_0 = 1$).
2. $\hat{\phi}_S(-t) = \overline{\hat{\phi}_S(t)}$ (since S is real, the CF is Hermitian).
3. $\hat{\phi}_S(t)$ is bounded on $\mathbb{R}$ (since it converges to $\phi_S(t)$ which is bounded by 1).
4. Every evaluation of $\hat{\phi}_S$ at a point costs $O(N_P)$ operations — polynomial evaluation.

3.5 The Closed-Form CDF Formula

The complete CDF is a single closed-form expression. We write it out explicitly.

Definition. Given (w, μ, Σ) and truncation parameters N_P, N, a, b:

(a) Scaled moments (closed-form sums of exponentials):

$$c_k = \frac{1}{k!} \sum_{|\alpha|=k} \frac{k!}{\alpha!}\, w^\alpha \exp\!\left(\alpha^T \mu + \tfrac{1}{2}\alpha^T \Sigma \alpha\right), \qquad k = 0, \dots, 2N_P - 1$$

(b) Moment matrix (Toeplitz system assembled from the scaled moments):

$$T \in \mathbb{R}^{N_P \times N_P}, \quad T_{ij} = c_{N_P-1+i-j}, \quad i, j = 0, \dots, N_P - 1$$
$$\mathbf{b} = (c_{N_P}, c_{N_P+1}, \dots, c_{2N_P-1})^T$$

This is equivalent to the classical Hankel matrix of the moment sequence under column reversal. Both forms yield the same Padé coefficients.

(c) Padé denominator (one matrix inverse):

$$\mathbf{q} = -T^{-1}\mathbf{b} \in \mathbb{R}^{N_P}, \qquad Q(z) = 1 + q_1 z + \cdots + q_{N_P} z^{N_P}$$

(d) Padé numerator (polynomial multiplication):

$$p_j = c_j + \sum_{l=1}^{\min(j,N_P)} q_l\, c_{j-l}, \qquad j = 0, \ldots, N_P - 1$$

(e) CDF (finite trigonometric sum):

$$\boxed{F_S(x) = \frac{x-a}{b-a} + \sum_{k=1}^{N-1} \frac{2}{k\pi}\, \mathrm{Re}\left[\frac{\sum_{j=0}^{N_P-1} p_j\,(it_k)^j}{1 + \sum_{j=1}^{N_P} q_j\,(it_k)^j} \cdot e^{-it_k a}\right] \cdot \sin\left(\frac{k\pi(x-a)}{b-a}\right)}$$

where $t_k = k\pi/(b-a)$.

This is one formula. Its ingredients are: (a) exponentials and finite sums (the moments), (b) a matrix inverse (the Padé system), (c) polynomial evaluation and trigonometric functions. Each is a standard closed-form operation. The matrix inverse T^{-1} is no less "closed-form" than $\sqrt{2}$ being the solution of $x^2 = 2$ — it is a finite algebraic operation with an explicit formula (Cramer's rule, LU decomposition, etc.).

3.6 Why This Is Fully Analytical

Step	Operation	Type
1. Moments	Evaluate m_k via Theorem 1	Closed-form algebra
2. Padé system	Solve $2N_P \times 2N_P$ linear system for P, Q	Linear algebra
3. CF evaluation	Compute $R(it_k)$ at COS frequencies	Rational function evaluation
4. COS inversion	$A_k = (2/(b-a))\mathrm{Re}[R(it_k)e^{-it_k a}]$	Arithmetic
5. CDF	$F(x) = (A_0/2)(x-a)/(b-a) + \sum_k (A_k/k\pi)\sin(\cdots)$	Finite sum

No numerical integration appears at any step. The only numerical operation beyond arithmetic is the Padé linear solve, which is $O(N_P^3)$ — exact in exact arithmetic, and conditioned by the moment ratios in finite precision.

---

4. The COS Inversion

4.1 From Padé CF to COS Coefficients

The Fourier-cosine coefficients of the CDF are computed exactly as in the standard COS method (Fang and Oosterlee, 2008), but with the Padé CF replacing the quadrature-computed CF:

$$A_k = \frac{2}{b-a}\mathrm{Re}\left[\hat{\phi}_S\left(\frac{k\pi}{b-a}\right)\exp\left(-i\frac{k\pi a}{b-a}\right)\right], \quad k = 0, 1, \dots, N-1$$

Each A_k is a single evaluation of the rational function $R(z)$ at $z = ik\pi/(b-a)$, followed by multiplication by a phase factor. The total cost for all N coefficients is $O(N \cdot N_P)$.

4.2 The Spectral Lognormal Distribution (Latent Form)

Definition 1 (Latent Spectral Lognormal Distribution). The *Latent SLD* $\hat{F} = \mathrm{LSLD}(A_0, \dots, A_{N-1}, a, b)$ is defined by:

$$f(x) = \frac{A_0}{2} + \sum_{k=1}^{N-1} A_k \cos\left(\frac{k\pi(x-a)}{b-a}\right)$$

$$F(x) = \frac{A_0}{2}(x-a) + \sum_{k=1}^{N-1} A_k \frac{b-a}{k\pi}\sin\left(\frac{k\pi(x-a)}{b-a}\right)$$

where A_k are computed from the Padé CF as above. The $N+2$ parameters $(A_0, \dots, A_{N-1}, a, b)$ fully determine the distribution.

This is the same functional form as the SLD in Nagy (2026a). What changes is the **extraction method**: the SLD computes A_k via eigenvalue conditioning and Gauss–Hermite quadrature; the Latent SLD computes A_k via Padé resummation of closed-form moments.

4.3 Properties Inherited from the COS Framework

The Latent SLD inherits all properties of the COS representation:

1. **Analytic CDF:** finite sum of sines.
2. **Analytic PDF:** finite sum of cosines.
3. **Analytic quantile function:** root-finding on the sine series.
4. **Analytic ES:** closed-form from coefficients (Proposition 6 in Nagy 2026a).
5. **Spectral risk measures:** deterministic evaluation of $\rho_\phi(S) = -\int_0^1 \phi(p)F^{-1}(p)\,dp$.

---

5. Convergence Analysis

5.1 Error Decomposition

Theorem 3 (Error Bound for the Latent SLD). *The total CDF error satisfies:*

$$\varepsilon_{\text{total}} := \sup_{x\in[a,b]} |F(x) - F_S(x)| \leq \varepsilon_{\text{Padé}}(N_P) + \varepsilon_N + \varepsilon_{[a,b]}$$

where:

Component	Source	Behavior
$\varepsilon_{\text{Padé}}(N_P)$	Padé approximation of CF	$\to 0$ as $N_P \to \infty$ (Theorem 2)
ε_N	COS truncation at N terms	$O(\rho_{\text{COS}}^{-N})$ (Bernstein ellipse, Lemma 1 in Nagy 2026a)
$\varepsilon_{[a,b]}$	Domain truncation	$\leq 5\,P(S \notin [a,b])$

5.2 Comparison with Eigen-COS Error Budget

The SLD (Eigen-COS) has a six-component error:

$$\varepsilon_{\text{SLD}} \leq \varepsilon_N + \varepsilon_{\text{GH}} + \varepsilon_{\text{outer}} + \varepsilon_{[a,b]} + \varepsilon_{\text{res}} + \varepsilon_{\text{fp}}$$

The Latent SLD **eliminates three components**:

Eliminated	Why
ε_{GH} (inner quadrature)	No Gauss–Hermite quadrature for CF
$\varepsilon_{\text{outer}}$ (outer conditioning integral)	No conditioning — correlations enter through moments
ε_{res} (residual correlation)	Full Σ in moment formula — no eigendecomposition, no residual

These three were the dominant error sources in the SLD for moderate conditioning ($K < n-1$). The Latent SLD replaces them with a single Padé error that converges to zero.

5.3 Padé Convergence Rate

The rate of Padé convergence depends on the singularity structure of $M(z)$ in the complex plane. For the MGF of a sum of lognormals:

- $M(z)$ is analytic for $\text{Re}(z) < 0$.
- The nearest singularity to the evaluation points $z = it_k$ is at $\text{Re}(z) = 0^+$.
- The Padé convergence rate is related to the density of singularities on the boundary (Theorem 2).

For practical purposes, the Padé convergence is rapid: $N_P = 15$–30 moments typically suffice for CDF accuracy better than 10^{-6} in the moderate-volatility regime ($\sigma \leq 0.5$).

5.4 The Eliminability Argument

Following the structure of the 3-body solution (Nagy, 2026g, Section 6.1):

Step	Approximation	Exact object	Why eliminable
Moments $m_0, \dots, m_{2N_P-1}$	Finite number of moments	All moments $\{m_k\}_{k=0}^{\infty}$	Moments are exact (Theorem 1)
Padé $[N_P/N_P]$	Rational approximation of CF	$\phi_S(t) = M(it)$	Padé converges (Theorem 2)

Step	Approximation	Exact object	Why eliminable
COS truncation at N	Finite cosine sum	$F_S(x)$	Exponential convergence (Lemma 1, Nagy 2026a)
Domain $[a, b]$	Finite support	$\mathbb{R}_+$	Tail mass $\to 0$ as $[a, b]$ widens

Every approximation targets an exact analytic object. The moments are exact. The Padé converges to the CF. The COS converges to the CDF. The domain contains the distribution to arbitrary precision. Therefore, the CDF of S is determined by the algebraic moment equations (Theorem 1) and an evaluation formula (Padé + COS), in the same sense that the three-body trajectory is determined by the Galerkin equation and the generating-function evaluation.

Theorem 4 (Exact Latent Representation). *The CDF of $S = \sum w_i e^{Y_i}$, $Y \sim \mathcal{N}(\mu, \Sigma)$, is exactly determined by:*

1. *The moment equation (Theorem 1), which is the multivariate normal structure rewritten in moment coordinates.*
2. *The Padé–COS evaluation chain, where each step converges to the exact object.*

In particular, the distribution is characterized by a finite algebraic system (the moment formula) and a convergent evaluation scheme (Padé + COS). No numerical integration is involved in the characterization.

---

6. The Latent Interpretation

6.1 The Latent of a Distribution

In the framework of Nagy (2026e), the **Latent** of a distribution is the basis-free element of $\mathcal{H}$ that encodes all distributional information. For the sum of correlated lognormals:

- **The Latent** Λ: the abstract object encoding the full distribution of S.
- **Coordinates in cosine basis**: the COS coefficients $A_0, \ldots, A_{N-1}$.
- **Extraction method**: Padé resummation of closed-form moments (this paper) or Eigen-COS with quadrature (Nagy, 2026a).

Both methods produce coordinates for the **same** Latent. The Latent itself is independent of the extraction method — it depends only on (w, μ, Σ).

6.2 The Generating Function Analogy

In the three-body problem, the generating function $G(z) = \sum \Lambda_k z^k$ encodes the trajectory. Its singularity structure determines the convergence rate.

For the lognormal sum, the moment generating function $M(z) = \sum (m_k/k!) z^k$ plays the same role. It encodes the distribution. Its singularity structure (the natural boundary at $\mathrm{Re}(z) = 0$) determines the Padé convergence rate.

Property	3-Body $G(z)$	Lognormal Sum $M(z)$
Domain of analyticity	$\|z\| < \rho$ (disk)	$\mathrm{Re}(z) < 0$ (half-plane)
Coefficients	Fourier Λ_k: decay as ρ^{-k}	Moment $m_k/k!$: grow super-exponentially
Taylor convergence	Converges in disk	**Diverges** everywhere
Padé convergence	Extends beyond disk to singularities	Recovers function on imaginary axis
Physical meaning	Trajectory in Fourier coordinates	Distribution in moment coordinates
Defining equation	Galerkin $R(\Lambda, \omega) = 0$	Moment equation (Theorem 1)

The fundamental difference: the 3-body generating function has a positive radius of convergence (the trajectory IS analytic in time), while the lognormal MGF has zero radius (lognormal moments grow too fast). Padé handles both — extending a convergent series for the 3-body problem, and resumming a divergent series for lognormals.

6.3 The Rational Latent Theorem Applied

The Rational Latent Theorem (Theorem 8 in Nagy 2026e, 2026g) states: if the generating function $G(z)$ is meromorphic with M poles, the Padé $[L/M]$ converges at rate R^{-N} where $R > \rho$ is determined by the next singularity beyond the poles.

For the lognormal sum: the "generating function" $M(z)$ has a natural boundary at $\mathrm{Re}(z) = 0$ (not isolated poles). The Padé still converges (by the Stieltjes theorem), but the rate is determined by the smoothness of the density f_S on $[0, \infty)$ rather than by isolated singularities.

The Latent size for the lognormal sum is:

$$N_{\text{Latent}} = 2N_P + N + 2$$

where N_P is the Padé order (number of moments used), N is the COS order, and 2 accounts for the domain bounds a, b.

---

7. Computational Details

7.1 Algorithm

Algorithm 1 (Latent SLD).

Input: Weights $w \in \mathbb{R}^n$, drift $\mu \in \mathbb{R}^n$, covariance $\Sigma \in \mathbb{R}^{n \times n}$, Padé order N_P, COS order N, domain multiplier L.

1. **Moments.** Compute $m_0, m_1, \ldots, m_{2N_P - 1}$ via Theorem 1.
2. **Domain.** Set $a = \max(0, m_1 - L\sqrt{m_2 - m_1^2})$, $b = m_1 + L\sqrt{m_2 - m_1^2}$.
3. **Padé.** Form $c_k = m_k/k!$ for $k = 0, \ldots, 2N_P - 1$. Solve the Padé $[N_P - 1/N_P]$ linear system to obtain polynomials P, Q.

4. **COS coefficients.** For $k = 0, \dots, N-1$: set $t_k = k\pi/(b-a)$, evaluate $\hat{\phi}(t_k) = P(it_k)/Q(it_k)$, compute $A_k = (2/(b-a))\mathrm{Re}[\hat{\phi}(t_k)e^{-it_k a}]$.
5. **Output.** Return $\mathrm{LSLD}(A_0, \dots, A_{N-1}, a, b)$.

Complexity: $O\Big(\binom{n+2N_P-2}{2N_P-1} + N_P^3 + N \cdot N_P\Big)$.

The first term is the multinomial moment computation (dominant for large n). The N_P^3 term is the Padé linear solve. The $N \cdot N_P$ term is the N rational-function evaluations.

7.2 Numerical Stability

The moments m_k grow as e^{ck^2} for large k. To maintain numerical stability:

1. **Work with scaled Taylor coefficients.** Set $c_k = m_k/k!$ rather than m_k directly.
2. **Use extended precision for the Padé solve.** Python's mpmath or Julia's ArbFloat provide arbitrary-precision arithmetic. The Padé system is small ($2N_P \times 2N_P$, typically 30×30), so the cost of extended precision is negligible.
3. **Moderate N_P.** For $\sigma \leq 0.5$, $N_P = 15$–20 suffices and c_k stays within double precision. For $\sigma > 0.5$, use extended precision with $N_P = 20$–30.

7.3 Practical Recommendations

Regime	$\sigma_{\max}$	N_P	Precision	Expected accuracy
Low volatility	≤ 0.3	15	Double (float64)	$< 10^{-8}$
Moderate	0.3–0.5	20	Double	$< 10^{-6}$
High volatility	0.5–0.9	25–30	Extended (mpmath)	$< 10^{-4}$
Crypto-like	> 0.9	30+	Extended	Case-dependent

---

8. Discussion

8.1 What Is New vs. the Spectral Lognormal Distribution

The SLD (Nagy, 2026a) and the Latent SLD (this paper) produce the **same output** — COS coefficients encoding the distribution — but differ in their extraction method:

	SLD (Eigen-COS)	Latent SLD (Padé)
Correlation handling	Eigendecomposition + conditioning	Direct through Σ in moments
CF computation	Gauss–Hermite quadrature (numerical)	Padé of moment series (algebraic)
Residual correlation error	$\varepsilon_{\text{res}} > 0$ for $K < n-1$	**Zero** (full Σ in moments)
Inner quadrature error	$\varepsilon_{\text{GH}} \sim 10^{-13}$	**Zero** (no quadrature)
Outer quadrature error	$\varepsilon_{\text{outer}}$ (depends on K, n_q)	**Zero** (no conditioning)
Padé error	None	$\varepsilon_{\text{Padé}}(N_P) \to 0$

	SLD (Eigen-COS)	Latent SLD (Padé)
Numerical operations	Eigendecomposition, two quadratures	Moment sums, one linear solve
Scales with	$O(n^3 + n_q^K N n\, n_{\mathrm{gh}})$	$O(\binom{n+2N_P}{2N_P} + N_P^3 + N N_P)$
Fully analytical?	No (quadrature)	**Yes**

The Latent SLD is **simpler in concept** (no eigendecomposition, no conditioning, no quadrature) but **more demanding in moment computation** (multinomial sums grow combinatorially with n and k). The SLD is better suited for large n with few dominant eigenfactors; the Latent SLD is better suited for moderate n where the full correlation structure matters and analytical purity is required.

8.2 The Moment Determinacy Question

A well-known subtlety: the lognormal distribution is **moment-indeterminate** — multiple distributions share the same moments (Heyde, 1963). Does this affect the Padé construction?

No, for the following reason. The Padé approximant is built from the Taylor coefficients of $M(z)$ at $z = 0$. The function $M(z)$ IS unique (it is the specific MGF of the specific random variable S defined by the model). The Padé converges to THIS $M(z)$ by the Stieltjes theorem, regardless of whether other distributions have the same moments. The moment indeterminacy means that the moments alone do not logically determine the distribution; but the Padé–Stieltjes convergence theorem says that the specific analytic function $M(z)$ IS recovered by the Padé — and this function does determine the distribution uniquely (via inverse Laplace transform).

In practice: for the SUM of multiple correlated lognormals (rather than a single lognormal), the moment problem is generically determinate. Sums tend to be better determined by their moments than individual heavy-tailed components.

8.3 The Cumulant Alternative

An alternative to the direct moment Padé is the **cumulant Padé**:

1. Compute cumulants $\kappa_1, \ldots, \kappa_{2N_P}$ from moments.
2. Form the cumulant generating function $K(z) = \sum(\kappa_k/k!)z^k$.
3. Padé resummation: $\hat{K}(z) = P(z)/Q(z)$.
4. CF via exponentiation: $\hat{\phi}(t) = e^{\hat{K}(it)}$.

Advantages: exp(rational) automatically satisfies $|\hat{\phi}(0)| = 1$ and has better positivity properties. Disadvantages: the CF is no longer rational (harder to analyze theoretically). In practice, both routes perform comparably. The direct moment Padé is simpler for theoretical analysis; the cumulant route may have better numerical conditioning.

8.4 Relation to the Latent Framework

In the Latent framework (Nagy, 2026e):

- The **Latent Theorem** guarantees: every smooth system with analyticity parameter $\rho > 1$ has a finite representation of size $N = \Theta(\log(1/\varepsilon)/\log\rho)$.

- The CDF of S is smooth for $x > 0$ (it is a convolution of lognormal densities). Its analyticity parameter ρ_{COS} determines the COS convergence rate.
- The **Rational Latent Theorem** governs the Padé step: the MGF $M(z)$ is the "generating function" of the distribution, and Padé resummation converges at a rate determined by its singularity structure.
- The **grade** of the Latent is 1 (it characterizes a distribution, not a generator or a multi-body interaction).

The complete Latent for the sum of correlated lognormals is therefore a **grade-1 Latent** extracted by the Padé–COS chain, with coordinates $(A_0, \dots, A_{N-1})$ in the cosine basis on $[a, b]$.

8.5 Limitations

1. **Moment computation scales combinatorially.** $\binom{n+k-1}{k}$ terms per moment for n assets and moment order k. For $n = 50$, $k = 30$: $\binom{79}{30} \approx 10^{20}$ terms. This is infeasible by direct enumeration for large n. Recursive and tensor-network methods can reduce this, but the Eigen-COS route (with its $O(n_q^K N n\, n_{\text{gh}})$ cost) may be more practical for large portfolios.

2. **Numerical precision for high volatility.** Moments grow as e^{ck^2}, causing overflow for $\sigma > 0.5$ at moderate N_P. Extended precision is needed.

3. **The Padé system can be ill-conditioned.** The moment matrix in the Padé equations has condition number related to the moment ratios. For well-separated volatilities, this can be large. Regularization or the SVD-based Padé (Gonnet et al., 2013) mitigates this.

4. **Not a new first integral.** As with the 3-body solution, this is a **reformulation**, not new information. The distribution of S is already fully determined by (w, μ, Σ); what changes is the form — algebraic rather than quadrature-based.

---

9. Conclusion

The distribution of a weighted sum of correlated lognormal random variables admits a fully analytical Latent representation. The construction is:

$$\boxed{(w, \mu, \Sigma) \;\to\; \{m_k\}_{\text{closed form}} \;\to\; [N_P/N_P]_{\text{Padé}} \;\to\; \{A_k\}_{\text{COS}} \;\to\; F_S(x)}$$

No Monte Carlo. No numerical quadrature. No eigendecomposition. Correlations enter algebraically through Σ in the closed-form moment formula. The CDF is a single closed-form expression (Section 3.5) involving exponentials, one matrix inverse, and trigonometric functions.

The mathematical structure parallels the exact Latent solution of the three-body problem (Nagy, 2026g): a generating function (the MGF) is defined by algebraic equations (the moment formula), Padé resummation — realized as a matrix inverse of the moment matrix — recovers the function on the evaluation domain (the imaginary axis), and Fourier inversion produces the observable (the CDF). In both cases, every approximation targets an exact analytic object and is therefore eliminable.

The sum of correlated lognormals now has a **closed-form CDF** in the standard mathematical sense: a finite expression built from sums, exponentials, one matrix inverse, and trigonometric functions. This is "closed" in the same way that $\mathbf{x} = A^{-1}\mathbf{b}$ is a closed-form solution — the matrix inverse is a finite algebraic operation, not an iterative procedure.

What is genuinely new is the recognition that the Padé step is not an algorithm but a formula ($\mathbf{q} = H^{-1}\mathbf{b}$), and that the moments feeding it are exact. This makes the CDF of the lognormal sum distribution — one of the most studied open problems in applied probability since 1960 — a closed-form expression involving matrix operations, on the same footing as the CDF of the multivariate normal (which requires the matrix inverse Σ^{-1} in its exponent).

---

10. The Smooth Latent Operator (Outlook)

The Padé-COS chain of Sections 3–5 depends on a discrete truncation parameter N_P (the Padé order). For fixed N_P, the CDF formula is a closed-form expression; for variable N_P, the formula changes discontinuously — the Toeplitz matrix changes dimension, coefficients can jump, and the condition number spikes at certain orders. This section outlines how the discrete chain can be embedded in a smooth operator, removing the truncation artifact entirely.

10.1 The Moment Hilbert Space

The scaled moments $\{c_k = m_k/k!\}_{k=0}^{\infty}$ grow as $e^{\alpha k^2}$ where $\alpha = \sigma^2_{\max}/2$. Define a weighted Hilbert space:

$$\mathcal{H}_\beta = \left\{ \{a_k\}_{k=0}^{\infty} : \sum_{k=0}^{\infty} |a_k|^2 e^{-2\beta k^2} < \infty \right\}$$

with inner product $\langle a, b \rangle_\beta = \sum_k a_k \overline{b_k} e^{-2\beta k^2}$. For $\beta > \alpha$, the moment sequence $\{c_k\} \in \mathcal{H}_\beta$.

The **Latent** of the distribution in this space is $\Lambda = \{c_k\} \in \mathcal{H}_\beta$. The map $(w, \mu, \Sigma) \mapsto \Lambda$ is smooth (each c_k is a smooth function of the parameters by Theorem 1).

10.2 Kernel Recovery of the CF

The characteristic function at frequency t is recovered by a kernel evaluation:

$$\hat{\phi}_S(t) = \sum_{k=0}^{\infty} c_k \,(it)^k \, g_\alpha(k) = \langle \Lambda, K_t \rangle_\beta$$

where $g_\alpha(k)$ is a damping function and $K_t \in \mathcal{H}_\beta$ is the evaluation kernel at frequency t. The classical Padé corresponds to a SPECIFIC kernel choice: the Toeplitz-inverse projection onto the first $2N_P$ coordinates.

A smooth generalization replaces the discrete Padé with Tikhonov-regularized recovery:

$$\mathbf{q}(\alpha) = -(T^*_{\lceil\alpha\rceil} T_{\lceil\alpha\rceil} + e^{-\alpha} I)^{-1} T^*_{\lceil\alpha\rceil} \mathbf{b}_{\lceil\alpha\rceil}$$

where $\alpha \in \mathbb{R}_{>0}$ is now a continuous resolution parameter. As α increases, the effective Padé order grows and the regularization weakens. The map $\alpha \mapsto R_\alpha(z)$ is smooth.

10.3 The Adaptive Resolution

The Latent Theorem (Nagy, 2026e, Theorem 3) determines the natural representation size from the analyticity parameter ρ:

$$N^* = \Theta\left(\frac{\log(1/\varepsilon)}{\log \rho}\right)$$

For the Padé chain, ρ is determined by the singularity structure of $M(z)$, which is itself a smooth function of (w, μ, Σ). Defining $\alpha^*(w, \mu, \Sigma) = N^*(\rho(w, \mu, \Sigma))$ gives a **parameter-free smooth operator**:

$$\boxed{\mathcal{L} : (w, \mu, \Sigma) \xrightarrow{\text{moments}} \Lambda \in \mathcal{H}_\beta \xrightarrow{K_{\alpha^*}} \hat{\phi}_S(\cdot) \xrightarrow{\text{COS}} F_S(x)}$$

The entire chain is smooth in (w, μ, Σ), smooth in x, and has no free truncation parameter.

10.4 The Grade-2 Latent

In the Latent framework, the distribution has a grade-1 Latent Λ (encoding *what* the distribution is). The smooth operator introduces a **grade-2 Latent** α^* (encoding *how to optimally extract* the distribution from its moment representation):

$$(\Lambda, \alpha^*) \in \mathcal{H}_\beta \times \mathbb{R}_{>0}$$

The grade-2 Latent α^* is the natural resolution of the problem — the point where adding more moments yields diminishing returns relative to the precision cost. Low volatility ($\sigma \leq 0.3$) gives large α^* (many moments contribute cheaply). High volatility ($\sigma > 0.8$) gives small α^* (few moments contribute reliably).

This construction resolves the numerical difficulty observed for high-volatility lognormal sums: instead of choosing N_P and hoping the Toeplitz matrix is well-conditioned, the adaptive α^* finds the optimal trade-off automatically. The Toeplitz singularity that blocked the Bitcoin test cases (Section 7.3) is handled by the regularization term $e^{-\alpha} I$.

A full development of the smooth Latent operator — including the RKHS construction, convergence analysis, and numerical implementation — is given in a companion paper (Nagy, 2026j).

10.5 Limitation: Smoothing Does Not Overcome the Moment Growth Rate

The Smooth Latent Operator improves robustness (no parameter tuning, no singularities) and provides smooth dependence on inputs. However, it does **not** overcome the fundamental information-theoretic limit imposed by the moment growth rate $|c_k| \sim e^{\sigma^2 k^2/2}$.

The regularization term $e^{-\alpha} I$ prevents blow-up but introduces bias. For high volatility ($\sigma > 0.8$), the optimal α^* is small — correctly diagnosing that few moments carry reliable information — but the resulting CDF approximation is necessarily poor. The grade-2 Latent **diagnoses** the difficulty but cannot **cure** it.

Concretely:

Regime	Discrete Padé	Smooth Operator	Improvement
$\sigma \leq 0.3$	PASS	PASS	None (already works)
$\sigma = 0.5$–0.8	Unstable, tuning needed	Stable, automatic	**Real improvement**
$\sigma > 1.0$	Singular matrix	No singularity, but inaccurate	More robust, but fundamentally limited
Hedge (negative w)	Empirically works	Same	None

The root cause is not the extraction method — it is the **representation itself**. The moments are the wrong Latent. This motivates the grade-3 analysis below.

10.6 The Moments Are the Wrong Latent (Grade-3 Insight)

The moments $\{m_k\} = \{E[S^k]\}$ are the Taylor coefficients of the MGF $M(z) = E[e^{zS}]$. For lognormal sums, $M(z)$ has **zero convergence radius** — it diverges for all $z > 0$. The Padé resummation attempts to recover a function from a divergent Taylor series. This works (under the completely monotone conditions of Theorem 2) but is inherently ill-conditioned because the coefficients carry exponentially less "shape information" relative to their magnitude as k grows.

The critical realization: **the moments are not a Latent of the distribution — they are a Latent of a specific generating function (the MGF).** The MGF is a poor generating function for lognormal sums. Different generating functions have different Latent representations with different conditioning:

Generating function	Latent (coefficients)	Growth rate
MGF: $M(z) = E[e^{zS}]$	Moments $m_k/k!$	$e^{\sigma^2 k^2/2}$ (super-exponential)
Hermite chaos: $S = \sum c_{\mathbf{k}}^H H_{\mathbf{k}}(Z)$	Hermite coefficients $c_{\mathbf{k}}^H$	$\sigma^{\lvert\mathbf{k}\rvert}/\lvert\mathbf{k}\rvert!$ (factorial decay)

The Hermite-chaos expansion is the natural Latent because it matches the **generative structure** of the problem: the underlying variables Y_i are Gaussian, and the Hermite polynomials are the natural basis for $L^2(\mathbb{R}^n, \gamma_n)$.

10.7 The Hermite Latent (Summary)

Let $L = \mathrm{chol}(\Sigma)$ and $Z = L^{-1}(Y - \mu) \sim N(0, I)$. The Hermite-chaos expansion of S is:

$$S = \sum_{\mathbf{k} \in \mathbb{N}_0^n} c_{\mathbf{k}}^H \, H_{\mathbf{k}}(Z), \quad c_{\mathbf{k}}^H = \sum_{i=1}^{n} w_i e^{\mu_i + \|\ell_i\|^2/2} \prod_{j=1}^{n} \frac{L_{ij}^{k_j}}{k_j!}$$

The coefficients $c_{\mathbf{k}}^H$ are **closed-form** (finite sum of products) and **factorially decaying** ($|c_{\mathbf{k}}^H| \leq C\sigma_{\max}^{|\mathbf{k}|}/|\mathbf{k}|!$). The Hermite Latent $\Lambda^H = \{c_{\mathbf{k}}^H\}$ lives in ℓ^2 — no Gaussian-weighted Hilbert space needed.

The CF is recovered from the Hermite Latent by **Gauss-Hermite quadrature** — a **finite sum**:

$$\phi_{S_K}(t) = \sum_{\ell=1}^{Q^n} w_\ell \, e^{it \cdot p_K(z_\ell)}$$

where (z_ℓ, w_ℓ) are precomputed quadrature nodes/weights and $p_K(z)$ is the Hermite polynomial. This is closed-form in the same sense as the Padé matrix inverse $\mathbf{q} = H^{-1}\mathbf{b}$: a finite, deterministic algebraic expression. The integrand $|e^{it \cdot p_K(z)}| = 1$ is bounded — no ill-conditioning possible.

The growth comparison at $\sigma = 1.0$:

Degree K	Moment $\lvert c_K \rvert$	Hermite $\lvert c_{\mathbf{k}}^H \rvert$ at $\lvert\mathbf{k}\rvert = K$
10	$\sim 10^{21}$	$\sim 10^{-7}$
20	$\sim 10^{87}$	$\sim 10^{-19}$
30	$\sim 10^{195}$	$\sim 10^{-33}$

Convergence for Bitcoin-class volatilities. The L^2 truncation error $\sim \sigma^{2K}/K!$ converges for ALL σ:

σ	Regime	Moment-Padé	Hermite ($K = 20$) CDF error
0.3	Normal	PASS	$\sim 10^{-27}$
0.8	**Bitcoin**	**FAIL** (singular Toeplitz)	$\sim 10^{-11}$
1.2	**Bitcoin extreme**	**FAIL** (dps=800 insufficient)	$\sim 10^{-8.5}$

The Hermite approach **solves the Bitcoin problem**: at $\sigma = 1.2$, CDF sup-norm error $\sim 10^{-8.5}$ with standard double-precision, no arbitrary-precision needed. The moment-Padé chain fails completely in this regime.

The Hermite-COS Formula. The entire chain collapses into a single explicit expression. Let $p_K(z) = \sum_{|\mathbf{m}| \leq K} c_{\mathbf{m}}^H H_{\mathbf{m}}(z)$ and let (z_ℓ, w_ℓ) be Gauss-Hermite quadrature nodes/weights. Then:

$$\boxed{F_S(x) = \frac{x-a}{b-a} + \sum_{k=1}^{N} \frac{4}{k\pi(b-a)} \sin\left(\frac{k\pi(x-a)}{b-a}\right) \sum_{\ell=1}^{Q^n} w_\ell \cos\left(\frac{k\pi}{b-a}\left[p_K(z_\ell) - a\right]\right)}$$

Two nested finite sums — no integrals, no iterations, no matrix inversions. The COS coefficients $\{A_k\}$ serve as the **Latent-of-the-CDF**: each A_k is an explicit algebraic function of the Hermite Latent $\{c_{\mathbf{m}}^H\}$. The two-layer Latent structure — distribution Latent $\{c_{\mathbf{m}}^H\} \to$ CDF Latent $\{A_k\} \to$ CDF formula — makes the full map $(w, \mu, \Sigma) \to F_S(x)$ explicit and algebraic at every step.

The practical cost: $N \times Q^n$ arithmetic operations. For $n \leq 8$: practical. For $n > 10$: factor-model reduction ($\Sigma = BB^T + D$, $B \in \mathbb{R}^{n \times r}$) reduces to r dimensions.

A full development — including convergence theorems, CF computation methods, the factor-model reduction, and the grade-3 formalism — is given in a companion paper (Nagy, 2026k).

10.8 Graded Latent Hierarchy

The three realizations of this paper and its companions form a hierarchy:

Grade	Object	Encodes	Companion paper
0	Parameters (w, μ, Σ)	Problem definition	—
1	Latent Λ (moments or Hermite)	What the distribution IS	This paper (Sections 3–5)
2	Extraction Latent α^*	How hard to extract from a given representation	Nagy (2026j)
3	Representation Latent $\mathcal{B}^*$	Which representation to use	Nagy (2026k)

The grade-3 Latent is the highest-leverage choice: it determines whether the entire chain is feasible. For lognormal sums, grade-3 selects Hermite over moments; for the three-body problem (Nagy, 2026g), grade-3 selects Fourier over Taylor. The **Representation–Structure Matching** principle: the optimal basis matches the problem's generative structure.

The grade-2 Latent remains valuable even after the grade-3 choice: within the Hermite basis, α^* determines the truncation degree K and any residual regularization. But the grade-3 choice ensures that α^* is large (easy extraction) rather than small (hard).

10.9 Three Routes from Latent to CDF Formula

Given the Hermite Latent $\Lambda^H = \{c_{\mathbf{k}}^H\}$, three routes to an explicit CDF formula were investigated:

Route A (Edgeworth-type perturbation): Divergent. Decompose $S = S_1 + R$ (Gaussian + correction). Taylor-expand $F_S(x) = E[N(z - R/\sigma_S)]$ in powers of R/σ_S. The j-th term involves $E[(R/\sigma_S)^j]$, which by hypercontractivity grows as $(j-1)^{jK/2}$, defeating the $1/j!$ Taylor coefficient for all $K \geq 2$. The 1D Hermite basis $\{He_j(z)\}$ is the wrong basis for the CDF — the divergence is intrinsic to the 1D projection, not to the representation quality.

Route B (Hermite-COS special function $\mathcal{H}^n_{K,N,Q}$): Convergent. Define the CDF via two nested finite sums (Section 10.7). The COS basis replaces the divergent 1D Hermite basis; GH quadrature bridges the multi-dimensional Hermite Latent to the 1D COS coefficients without passing through the divergent 1D moment series. This is the convergent resummation that Route A fails to achieve. Factorially convergent ($\sigma^{2K}/K!$) for all σ.

Route C (Quadratic case $K = 2$): Exact in known functions. For $K = 2$, S_2 is a quadratic form in Gaussians. Eigendecompose: $S_2 = \mu_0' + \sum_j d_j(W_j + \delta_j)^2$ — a weighted sum of non-central χ_1^2 variables. CDF computable by Imhof's (1961) formula (single 1D integral). Accurate only for $\sigma \lesssim 0.2$; inadequate for moderate or high volatility.

Route	Formula type	Convergent?	All σ?
A: Edgeworth	$N(z) + \phi(z) \sum \Delta_p$	No	—
B: Hermite-COS $\mathscr{H}$	Double finite sum	Yes	Yes
C: Quadratic ($K = 2$)	Non-central χ^2 CDF	Exact for S_2	$\sigma \lesssim 0.2$ only

Route B is definitive. Full details: Nagy (2026k), Section 7.4.

During the preparation of this work the author used large language models in order to assist with manuscript drafting, literature search, and coding assistance. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

- Fenton, L. F. (1960). The Sum of Log-Normal Probability Distributions in Scatter Transmission Systems. *IRE Transactions on Communications Systems*, CS-8, 57-67. DOI: 10.1109/tcom.1960.1097606
- Baker, G. A. and P. Graves-Morris (1996). Padé Approximants. *Padé Approximants.*
- Fang, Fang and Oosterlee, Cornelis W. (2008). A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions. *SIAM Journal on Scientific Computing*, 31(2), 826-848. DOI: 10.1137/080718061
- Nagy, T. (2026). The Spectral Lognormal Distribution, The Distribution of Portfolio Value. *Zenodo*. DOI: 10.5281/zenodo.18940756
- Nagy, T. (2026). The Latent: Finite Sufficient Representations of Smooth Systems. *Zenodo*. DOI: 10.5281/zenodo.19101209
- Nagy, T. (2026). The Exact Latent Solution of the Gravitational Three-Body Problem. *Zenodo*. DOI: 10.5281/zenodo.19101229
- Beaulieu, Norman C. and Xie, Qing (2004). An Optimal Lognormal Approximation to Lognormal Sum Distributions. *IEEE Transactions on Vehicular Technology*, 53(2), 479-489. DOI: 10.1109/tvt.2004.823494
- Milevsky, Moshe A. and Posner, Steven E. (1998). Asian Options, the Sum of Lognormals, and the Reciprocal Gamma Distribution. *Journal of Financial and Quantitative Analysis*, 33(3), 409-422. DOI: 10.1142/9789812812599_0007
- Heyde, C.C (1963). On a property of the lognormal distribution. *Heyde, C.C.*, 392-393.
- Acerbi, Carlo (2002). Spectral Measures of Risk: A Coherent Representation of Subjective Risk Aversion. *Journal of Banking & Finance*, 26(7), 1505-1518. DOI: 10.1016/S0378-4266(02)00281-9
- Gonnet, P., Güttel, S., and Trefethen, L.N (2013). Robust Padé approximation via SVD. *Gonnet, P., Güttel, S., and Trefethen, L.N.*, 55(1), 101-117. DOI: 10.1137/110853236
- Schwartz, S. C. and Yeh, Y. S. (1982). On the Distribution Function and Moments of Power Sums with Lognormal Components. *Bell System Technical Journal*, 61, 1441-1462. DOI: 10.1002/j.1538-7305.1982.tb04353.x

- Dufresne, Daniel (2004). The Log-Normal Approximation in Financial and Other Computations. *Advances in Applied Probability*, 36, 747-773. DOI: 10.1239/aap/1093962232
- Trefethen, Lloyd N. (2013). Approximation Theory and Approximation Practice. SIAM. DOI: 10.1137/1.9781611975949
- Wall, H.S (1948). Analytic Theory of Continued Fractions. *Wall, H.S.*.
- Nagy, T. (2026). The Smooth Latent Operator: Parameter-Free Distributional Representations via Kernel Moment Recovery. *Working paper*.
- Nagy, T. (2026). The Hermite Latent: Natural Representations for Sums of Correlated Lognormals. *Zenodo*. DOI: 10.5281/zenodo.19134411
- Cameron, R.H. and Martin, W.T (1947). The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *Cameron, R.H. and Martin, W.T.*, 48(2), 385-392. DOI: 10.2307/1969178
- Wiener, N (1938). The homogeneous chaos. *Wiener, N.*, 60(4), 897-936. DOI: 10.2307/2371268