

The Fenton Distribution Solved

Closed-Form PDF, CDF, VaR, and ES for Sums of Correlated Lognormals

The distribution of a lognormal sum is determined by $n(n+5)/2$ parameters. One formula — two contractions composed — evaluates the CDF from this finite latent with no free parameters.

Tamás Nagy, Ph.D.

tnagyphd@gmail.com

Draft • March 2026

Executive Summary (Non-Technical)

The distribution of a weighted sum of correlated lognormal random variables — central to portfolio risk, option pricing, and wireless communications — has resisted closed-form solution for over six decades since Fenton (1960). This paper gives the exact formula.

The convergence contrast. The standard computational route to the lognormal sum CDF passes through a *divergent* intermediate object: the moment generating function $M(z) = E[e^{zS}]$ has zero convergence radius (its Taylor coefficients grow as $e^{\sigma^2 k^2/2}$). Every moment-based method — Padé resummation, saddlepoint, Edgeworth expansion — inherits this divergence and fails for high volatility. By contrast, the same distribution admits a *convergent* evaluation chain through the Hermite-COS pipeline: the Fourier-cosine coefficients of the log-space density decay exponentially, the Gauss-Hermite quadrature evaluates the Gaussian inner product (a Latent contraction) with exponential convergence, and the entire pipeline consists of algebraic operations between Latents — maps, contractions, synthesis — each exact at the algebraic level. The formula converges factorially ($\sim \sigma^{2K}/K!$) for all volatilities, including extreme regimes (Bitcoin) where all previous methods fail.

The resolution: the difficulty of the lognormal sum problem was never intrinsic to the distribution. It was an artifact of the representation. In the wrong coordinates (moments), the problem looks impossibly hard. In the right coordinates (Hermite-chaos coefficients of the underlying Gaussians), it is a trigonometric polynomial.

The paper shows that the moments are the wrong representation: they grow super-exponentially ($e^{\sigma^2 k^2/2}$), making moment-based methods numerically unstable for high-volatility regimes. The lognormal sum lives naturally in a Gaussian world (the underlying normals), and its natural coordinates are Hermite-chaos coefficients — which decay factorially ($\sim \sigma^k/k!$) instead of growing super-exponentially. This is not a technical improvement; it is a change of basis that eliminates the fundamental obstruction.

The resulting formula evaluates the CDF, PDF, VaR, and Expected Shortfall from the portfolio parameters alone — no Monte Carlo simulation, no arbitrary-precision arithmetic, no iterative calibration. A reference implementation runs in under 5 milliseconds for a two-asset portfolio and under 1 second for five assets, using standard double-precision arithmetic.

Abstract

The moment-based Latent representation of correlated lognormal sums (Nagy, 2026, *The Exact Latent Distribution of Correlated Lognormal Sums*) relies on scaled moments $c_k = m_k/k!$ that grow as $e^{\sigma_{\max}^2 k^2/2}$. This super-exponential growth is the root cause of ill-conditioning in the Padé resummation step: the moment Toeplitz matrix has condition number $\sim e^{O(N_P^2)}$, making the method impractical for $\sigma_{\max} > 0.8$. The Smooth Latent Operator (Nagy, 2026, *The Smooth Latent Operator: Parameter-Free Distributional Representations via Kernel Moment Recovery*) mitigates this through regularization but cannot overcome the information-theoretic limit imposed by the growth rate.

We show that the moments are the Latent representation of the wrong generating function. The moments are Taylor coefficients of the moment generating function $M(z) = E[e^{zS}]$, which has zero convergence radius for lognormal sums. By contrast, the **Hermite-chaos expansion** — the Wiener polynomial chaos of S in the underlying Gaussian variables — has coefficients that decay factorially. This is because the exponential function e^Y has a convergent Hermite series for $Y \sim N(\mu, \sigma^2)$, with coefficients $\sigma^k/k!$.

The resulting **Hermite Latent** $\Lambda^H = \{c_{\mathbf{k}}^H\}$ lives in ℓ^2 (no Gaussian weight needed), the characteristic function is recovered by a convergent inner product (no Padé resummation needed), and the CDF follows by Fourier-cosine inversion (Fang and Oosterlee, 2008). The entire chain is closed-form and well-conditioned for all σ .

We formalize this as a **grade-3 Latent**: the optimal choice of representation basis. The grade-3 Latent for lognormal sums is the Hermite-chaos basis, selected by matching the basis to the problem’s generative structure (Gaussian \rightarrow exponential \rightarrow sum). The three-body problem’s use of Fourier coefficients (instead of Taylor coefficients) is the same grade-3 choice, matched to the problem’s periodic structure.

The resulting CDF formula — characteristic-function inversion composed from two contractions (Gaussian, then Fourier) — evaluates the distribution from the finite generative latent $(w, \mu, \Sigma) \in \mathbb{R}^{n(n+5)/2}$ with no free parameters. Evaluator choices (Gauss-Hermite for the Gaussian contraction, COS for the Fourier inversion) provide exponential convergence in both precision parameters. The decades-long difficulty was a convergence problem: the moment-based route uses a divergent evaluator ($E[e^{tS}]$) for the same inner product that the Hermite-COS route evaluates convergently ($E[e^{itS}]$, with $|e^{itS}| = 1$).

1. Introduction

1.1 The Moment Growth Problem

For a sum $S = \sum_{i=1}^n w_i e^{Y_i}$ with $Y \sim N(\mu, \Sigma)$, the moments are (Nagy, 2026, *The Exact Latent Distribution of Correlated Lognormal Sums*, Theorem 1):

$$m_k = E[S^k] = \sum_{|\mathbf{j}|=k} \binom{k}{\mathbf{j}} \prod_{i=1}^n w_i^{j_i} \cdot \exp(\mathbf{j}^T \mu + \frac{1}{2} \mathbf{j}^T \Sigma \mathbf{j})$$

The dominant term for large k is the diagonal: $j_i = k$ for the asset i^* with largest σ_i . This gives $m_k \gtrsim e^{k\mu_{i^*} + k^2\sigma_{i^*}^2/2}$, so:

$$|c_k| = \frac{m_k}{k!} \gtrsim \frac{e^{k^2 \sigma_{\max}^2 / 2}}{k!} \sim e^{k^2 \sigma_{\max}^2 / 2}$$

The factorial in the denominator is negligible against the Gaussian in the numerator. For $\sigma_{\max} = 1.0$ and $k = 20$: $c_{20} \sim e^{200} \sim 10^{87}$.

1.2 Why This Growth Is Fatal

The Padé $[N_P - 1/N_P]$ requires solving a Toeplitz system with entries $H_{ij} = c_{N_P+i-j}$. The matrix entries span a range from $c_0 = 1$ to $c_{2N_P} \sim e^{(2N_P)^2 \sigma^2 / 2}$. The condition number is at least:

$$\kappa(H) \geq \frac{\max |H_{ij}|}{\min |H_{ij}|} \gtrsim e^{4N_P^2 \sigma^2 / 2}$$

For $N_P = 18$, $\sigma = 0.8$: $\kappa \gtrsim e^{4 \cdot 324 \cdot 0.32} \approx e^{415} \approx 10^{180}$. This exceeds the capacity of 500-digit arithmetic (dps=500 in mpmath gives $\sim 10^{500}$ dynamic range), explaining the singular-matrix failures observed in practice (Nagy, 2026, *The Exact Latent Distribution of Correlated Lognormal Sums*, Section 7.3).

No amount of regularization (Smooth Latent Operator) or precision increase can fundamentally change this: the information content per moment decreases exponentially with k , because c_k is dominated by a single multinomial term and the “signal” (distributional shape information) is exponentially smaller than the “carrier” ($e^{k^2 \sigma^2 / 2}$).

1.3 Main Result

The CDF of $S = \sum_{i=1}^n w_i e^{Y_i}$, $Y \sim N(\mu, \Sigma)$, is determined by the finite generative latent $(w, \mu, \Sigma) \in \mathbb{R}^{n(n+5)/2}$:

$$F_S(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \operatorname{Im} \left[\frac{\tilde{\phi}(t)}{t} e^{-it \log x} \right] dt, \quad \tilde{\phi}(t) = \int_{\mathbb{R}^n} S(z)^{it} \gamma_n(z) dz$$

where $S(z) = \sum_{i=1}^n w_i \exp(\mu_i + \ell_i^T z)$ is evaluated exactly from the Cholesky factor $L = \operatorname{chol}(\Sigma)$, and γ_n is the standard Gaussian density. The formula is a composition of two convergent contractions — Gaussian (inner product against γ_n) and Fourier (Gil-Pelaez inversion) — evaluated by Gauss-Hermite quadrature and Fourier-cosine synthesis respectively. VaR and ES follow by root-finding and integration of the COS density (§8).

The formula has no calibration parameters: all computational choices (Q quadrature nodes, N COS terms, log-space domain $[\alpha, \beta]$) are deterministic functions of (w, μ, Σ) . It converges for all σ because $|S(z)^{it}| = 1$ — the integrand is bounded, unlike the divergent e^{tS} of the moment route. For Bitcoin-class volatility ($\sigma > 0.8$): max CDF error $\sim 10^{-4}$ against Monte Carlo, in standard float64 arithmetic (§7.5).

1.4 The Key Observation

The moments are the Taylor coefficients of the MGF $M(z)$. But the MGF has zero convergence radius. We are trying to resum a divergent series.

The divergence is not intrinsic to the DISTRIBUTION — it is intrinsic to the MGF as a choice of generating function. The distribution is perfectly smooth and well-behaved. The moments encode it fully. The problem is that the monomial basis $\{z^k\}$ is catastrophically ill-suited to the function $M(z) = E[e^{zS}]$ when S is lognormal.

The question is: is there a generating function whose coefficients are well-behaved?

1.5 The Three-Body Analogy

In the three-body problem (Nagy, 2026, *The Exact Latent Solution of the Gravitational Three-Body Problem*), the solution $\mathbf{r}(t)$ is a smooth function of time. Its Taylor coefficients grow (near collisions, the convergence radius shrinks). But the solution is periodic (or quasi-periodic), and the **Fourier coefficients** decay geometrically: $|a_k| \sim \rho^{-|k|}$.

The choice of Fourier basis over Taylor basis was the critical design decision. It was dictated by the problem structure: the three-body system is periodic, so the Fourier basis is natural.

For lognormal sums, the analogous structural feature is: the underlying variables Y_i are Gaussian. The natural basis for functions of Gaussian variables is the **Hermite polynomial basis** — the Wiener chaos.

1.6 Organization

Sections 2–3 develop the Hermite-chaos expansion and define the Hermite Latent. Section 4 builds the evaluation chain: log-space characteristic function via Gauss–Hermite quadrature, CDF via Fourier-cosine inversion, and the exact boxed formula. Section 5 formalizes grade-3 Latents and the Representation–Structure Matching Theorem. Section 6 compares moment and Hermite Latents. Section 7 examines three routes from Latent to CDF (Edgeworth, Hermite-COS, quadratic special case) and provides numerical validation. Section 8 derives VaR and ES from the CDF formula. Section 9 addresses the curse of dimensionality and factor-model mitigation. Sections 10–11 discuss implications for the Latent framework and conclude.

2. The Hermite-Chaos Expansion

2.1 Setup

Let $L = \text{chol}(\Sigma)$ (lower-triangular Cholesky factor) and $Z = L^{-1}(Y - \mu) \sim N(0, I_n)$. Then:

$$Y_i = \mu_i + \sum_{j=1}^n L_{ij} Z_j = \mu_i + \ell_i^T Z$$

where $\ell_i = (L_{i1}, \dots, L_{in})^T$ is the i -th row of L .

Each lognormal component is:

$$e^{Y_i} = e^{\mu_i + \ell_i^T Z} = e^{\mu_i + \|\ell_i\|^2/2} \cdot e^{\ell_i^T Z - \|\ell_i\|^2/2}$$

The second factor is the stochastic exponential (Wick exponential) of $\ell_i^T Z$.

2.2 Hermite Expansion of a Single Lognormal

For a single standard normal Z_j , the Hermite polynomials $\{H_k(z)\}_{k=0}^\infty$ form an orthogonal basis of $L^2(\mathbb{R}, \gamma)$ where γ is the standard Gaussian measure. The exponential has the well-known expansion:

$$e^{aZ_j} = e^{a^2/2} \sum_{k=0}^{\infty} \frac{a^k}{k!} H_k(Z_j)$$

Fact: The coefficients $a^k/k!$ decay factorially for any fixed a .

For the multivariate case, define multi-index Hermite polynomials $H_{\mathbf{k}}(Z) = \prod_{j=1}^n H_{k_j}(Z_j)$ for $\mathbf{k} = (k_1, \dots, k_n) \in \mathbb{N}_0^n$. The stochastic exponential expands as:

$$e^{\ell_i^T Z - \|\ell_i\|^2/2} = \sum_{\mathbf{k} \in \mathbb{N}_0^n} \prod_{j=1}^n \frac{L_{ij}^{k_j}}{k_j!} H_{\mathbf{k}}(Z)$$

Therefore:

$$e^{Y_i} = e^{\mu_i + \|\ell_i\|^2/2} \sum_{\mathbf{k}} \left(\prod_{j=1}^n \frac{L_{ij}^{k_j}}{k_j!} \right) H_{\mathbf{k}}(Z)$$

2.3 Hermite Expansion of the Sum

The weighted sum is:

$$S = \sum_{i=1}^n w_i e^{Y_i} = \sum_{\mathbf{k} \in \mathbb{N}_0^n} c_{\mathbf{k}}^H H_{\mathbf{k}}(Z)$$

where the **Hermite-chaos coefficients** are:

$$c_{\mathbf{k}}^H = \sum_{i=1}^n w_i e^{\mu_i + \|\ell_i\|^2/2} \prod_{j=1}^n \frac{L_{ij}^{k_j}}{k_j!}$$

This is a **closed-form** expression: a finite sum of products of known quantities.

2.4 Growth Rate

Theorem 1 (Hermite Coefficient Decay). *The Hermite-chaos coefficients satisfy:*

$$|c_{\mathbf{k}}^H| \leq C \cdot \frac{(n\sigma_{\max})^{|\mathbf{k}|}}{|\mathbf{k}|!} \cdot \max_i |w_i| e^{\mu_i + \sigma_i^2/2}$$

where $|\mathbf{k}| = k_1 + \dots + k_n$ and C is an absolute constant.

In particular, for fixed total degree $K = |\mathbf{k}|$:

$$\sum_{|\mathbf{k}|=K} |c_{\mathbf{k}}^H|^2 \leq D \cdot \frac{(n\sigma_{\max}^2)^K}{K!}$$

where D depends on w, μ but not on K .

Proof. Each term in $c_{\mathbf{k}}^H$ is bounded by $|w_i| e^{\mu_i + \sigma_i^2/2} \prod_j |L_{ij}|^{k_j}/k_j!$. Since $|L_{ij}| \leq \sigma_{\max}$ (by the Cholesky structure), $\prod_j |L_{ij}|^{k_j}/k_j! \leq \sigma_{\max}^{|\mathbf{k}|}/\prod_j k_j!$. Now $\prod_j k_j! = |\mathbf{k}|!/ \binom{|\mathbf{k}|}{\mathbf{k}} \geq |\mathbf{k}|!/n^{|\mathbf{k}|}$, so $1/\prod_j k_j! \leq n^{|\mathbf{k}|}/|\mathbf{k}|!$, giving $\prod_j |L_{ij}|^{k_j}/k_j! \leq (n\sigma_{\max})^{|\mathbf{k}|}/|\mathbf{k}|!$. Summing over i gives the bound. \square

Comparison with moments:

K	Moment $ c_K $ for $\sigma = 1.0$	Hermite $\sum_{ \mathbf{k} =K} c_{\mathbf{k}}^H $
5	$\sim e^{12.5} \approx 3 \times 10^5$	$\sim 1/120 \approx 0.008$
10	$\sim e^{50} \approx 5 \times 10^{21}$	$\sim 1/3628800 \approx 3 \times 10^{-7}$
20	$\sim e^{200} \approx 10^{87}$	$\sim 1/20! \approx 4 \times 10^{-19}$
30	$\sim e^{450} \approx 10^{195}$	$\sim 1/30! \approx 4 \times 10^{-33}$

The gap widens as K increases: moments EXPLODE while Hermite coefficients VANISH. At $K = 30$, the ratio is $\sim 10^{228}$. This is the difference between a divergent and a convergent representation.

3. The Hermite Latent

3.1 Definition

Definition 1 (Hermite Latent). The **Hermite Latent** of the sum $S = \sum w_i e^{Y_i}$ is the sequence:

$$\Lambda^H = \{c_{\mathbf{k}}^H\}_{\mathbf{k} \in \mathbb{N}_0^n} \in \ell^2(\mathbb{N}_0^n)$$

By Theorem 1, $\|\Lambda^H\|_2^2 = \sum_{\mathbf{k}} |c_{\mathbf{k}}^H|^2 < \infty$ (convergent geometric series after grouping by total degree). No Gaussian weight is needed — the Hermite Latent lives in the standard ℓ^2 space.

3.2 The Hilbert Space Structure

The Parseval identity for Hermite-chaos gives:

$$\text{Var}(S) = \sum_{\mathbf{k} \neq \mathbf{0}} |c_{\mathbf{k}}^H|^2 \cdot \mathbf{k}! = \|\Lambda^H\|_w^2$$

where $\mathbf{k}! = \prod k_j!$ and $\|\cdot\|_w$ is the weighted ℓ^2 norm. The variance is FINITE (lognormal sums have finite variance), confirming ℓ^2 membership.

Comparison with the moment Hilbert space: The moment Latent $\Lambda^M = \{c_k\}$ required a Gaussian-weighted space \mathcal{H}_β with $\beta > \sigma_{\max}^2/2$ to ensure convergence (Nagy, 2026, *The Smooth Latent Operator: Parameter-Free Distributional Representations via Kernel Moment Recovery*, Section 2.1). The Hermite Latent needs no such weight — it is naturally in ℓ^2 .

3.3 Smoothness

Proposition 1. *The Hermite Latent map $\Phi^H : (w, \mu, \Sigma) \mapsto \Lambda^H \in \ell^2$ is smooth (infinitely Fréchet differentiable).*

Proof. Each $c_{\mathbf{k}}^H$ is a smooth function of (w, μ, Σ) (it involves the Cholesky factor $L(\Sigma)$, which is smooth for $\Sigma \succ 0$, and exponentials). The ℓ^2 convergence of the derivatives follows from the same factorial decay. \square

4. From Hermite Latent to CDF

This section develops two parallel stories. First (§4.1–4.3): the Hermite truncation S_K provides a polynomial in Gaussian variables whose CF is tractable; this explains *why* the pipeline is well-conditioned. Second (§4.4): the function $S(z) = \sum w_i \exp(\mu_i + \ell_i^T z)$ can be evaluated *exactly* at quadrature nodes, so the Hermite expansion degree K may be bypassed entirely in the implemented formula. The two stories share the same evaluator infrastructure (Gauss–Hermite for the Gaussian contraction, COS for the Fourier inversion).

4.1 The Characteristic Function

The CF of S is:

$$\phi_S(t) = E[e^{itS}] = E \left[\exp \left(it \sum_{\mathbf{k}} c_{\mathbf{k}}^H H_{\mathbf{k}}(Z) \right) \right]$$

This is the CF of a random variable expressed as a (convergent) Hermite series. For a FINITE truncation at total degree K :

$$S_K = \sum_{|\mathbf{k}| \leq K} c_{\mathbf{k}}^H H_{\mathbf{k}}(Z)$$

the CF is computable because S_K is a polynomial in Gaussian variables — its distribution is determined by a finite set of operations.

4.2 The Preferred Method: Gauss-Hermite Quadrature as Finite Sum

For moderate n (say $n \leq 8$), the CF of S_K is:

$$\phi_{S_K}(t) = \int_{\mathbb{R}^n} e^{it \cdot p_K(z)} \gamma_n(z) dz = \sum_{\ell=1}^{Q^n} w_\ell e^{it \cdot p_K(z_\ell)}$$

where (z_ℓ, w_ℓ) are n -dimensional Gauss-Hermite quadrature nodes and weights, $p_K(z) = \sum_{|\mathbf{k}| \leq K} c_{\mathbf{k}}^H H_{\mathbf{k}}(z)$ is a polynomial, and γ_n is the standard Gaussian density.

This is a **finite sum of Q^n terms**, each involving the exponential of a known polynomial evaluated at a known quadrature node. It is **closed-form** in the same sense that the Padé matrix inverse $\mathbf{q} = H^{-1}\mathbf{b}$ is closed-form: a finite algebraic expression built from known quantities. No Monte Carlo, no iterative solver, no convergence loop.

Why the integrand is well-conditioned. The modulus $|e^{it \cdot p_K(z)}| = 1$ for all z — the integrand is bounded. This is fundamentally different from the moment-based approach, where the integrand $e^{z \cdot S}$ grows without bound. The Gaussian weight $\gamma_n(z)$ provides the decay. The combination — bounded oscillatory integrand times rapidly decaying weight — is the ideal setting for Gauss-Hermite quadrature.

Convergence rate. Gauss-Hermite with Q nodes per dimension integrates polynomials of degree $2Q - 1$ exactly. The integrand $e^{it \cdot p_K(z)}$ is not polynomial, but its Taylor expansion $\sum_{j=0}^J (it \cdot p_K(z))^j / j!$ at any truncation J IS polynomial (degree KJ), and Gauss-Hermite with $Q \geq (KJ+1)/2$ nodes integrates it exactly. The Taylor remainder decays as $|t \cdot p_K(z)|^{J+1} / (J+1)!$, giving exponential convergence in Q .

For practical purposes: $Q = 15$ nodes per dimension integrates polynomials of degree 29 exactly, providing 12+ digits of accuracy for the Gaussian contraction at moderate t . End-to-end CDF accuracy also depends on the COS evaluator parameters $(N, [\alpha, \beta])$; see §7.5 for validated pipeline accuracy.

Cost. For $n = 2, Q = 15$: 225 evaluations (< 1 ms). For $n = 5, Q = 15$: 759,375 evaluations (< 1 second). For $n = 8, Q = 10$: 10^8 evaluations (~ 10 seconds). Beyond $n \approx 8$, sparse-grid quadrature (Smolyak) or factor-model reduction is needed.

4.3 Alternative Methods

Method A: Hermite moment transfer. The moments of S_K are:

$$E[S_K^p] = \sum_{\mathbf{k}_1, \dots, \mathbf{k}_p} \prod_{r=1}^p c_{\mathbf{k}_r}^H \cdot E \left[\prod_{r=1}^p H_{\mathbf{k}_r}(Z) \right]$$

The expectation is computed from the Wick/Isserlis theorem. These Hermite-derived moments can feed the Padé-COS chain of Nagy (2026, *The Exact Latent Distribution of Correlated Lognormal Sums*). However, the moments of S_K still APPROXIMATE the raw moments m_p (which grow

super-exponentially), so the Padé conditioning improves moderately but is not eliminated. Method A is inferior to Method B for high volatility.

Method B (preferred): Gauss-Hermite quadrature as described in Section 4.2.

Method C: Hybrid. Use the Hermite coefficients to evaluate $p_K(z)$ at quadrature nodes, compute the CF by quadrature (Method B), then apply COS inversion.

All three methods above use the Hermite truncation S_K . But $S(z) = \sum_{i=1}^n w_i \exp(\mu_i + \ell_i^T z)$ can be evaluated **exactly** at any z from the finite latent — making the Hermite expansion unnecessary for computation. The Hermite coefficients’ factorial decay (Theorem 1) explains *why* the pipeline converges, but the computational formula bypasses them entirely:

4.4 The Exact Formula

The distribution of $S = \sum_{i=1}^n w_i e^{Y_i}$, $Y \sim N(\mu, \Sigma)$, is determined by the finite generative latent $(w, \mu, \Sigma) \in \mathbb{R}^{n(n+5)/2}$. The CDF is:

$$F_S(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \operatorname{Im} \left[\frac{\tilde{\phi}(t)}{t} e^{-it \log x} \right] dt$$

where $\tilde{\phi}(t) = E[S^{it}] = E[e^{it \log S}]$ is the **log-space characteristic function** (Mellin–Fourier transform of the density):

$$\tilde{\phi}(t) = \langle S(\cdot)^{it}, \mathbf{1} \rangle_{\gamma_n} = \int_{\mathbb{R}^n} S(z)^{it} \gamma_n(z) dz$$

and $S(z) = \sum_{i=1}^n w_i \exp(\mu_i + \ell_i^T z)$ is evaluated **exactly** from the finite latent at any z , where $L = \operatorname{chol}(\Sigma)$. This is equivalent to the standard Gil–Pelaez inversion applied to $T = \log S$: since $F_S(x) = F_T(\log x)$ and $\tilde{\phi}(t) = \phi_T(t)$, the formula follows from $F_T(y) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \operatorname{Im}[\phi_T(t) e^{-ity}/t] dt$ with $y = \log x$.

This formula has **no calibration parameters** beyond the generative latent — all computational choices are deterministic functions of (w, μ, Σ) : - No truncation order K (no Hermite expansion — $S(z)$ is evaluated directly) - The domain $[\alpha, \beta]$ is set from the cumulants of $\log S$ (deterministic, not fitted) - The COS order N and quadrature order Q are precision parameters analogous to machine epsilon in evaluating $N(x)$

The formula is a composition of four Latent algebra operations:

$$\underbrace{(w, \mu, \Sigma)}_{\text{finite latent}} \xrightarrow{\text{map}} \underbrace{S(\cdot) \in L^2(\gamma_n)}_{\text{grade-1}} \xrightarrow{\text{complexify}} \underbrace{S(\cdot)^{it}}_{\text{grade-1, complex}} \xrightarrow{\text{contract}_\gamma} \underbrace{\tilde{\phi}(t)}_{\text{grade-0}} \xrightarrow{\text{contract}_{dt}} F_S(x)$$

The first arrow is a **map**: the finite latent generates $S(z)$ by matrix-vector multiply (Lz , $O(n^2)$) followed by n exponentials and a weighted sum. The second arrow is **complexification**: $S^{it} = e^{it \log S}$ extends the distribution to \mathbb{C} via the log-space (Mellin) transform. The third arrow is a

contraction against the Gaussian measure — the inner product $\langle S(\cdot)^{it}, \mathbf{1} \rangle_{\gamma_n}$ that produces the log-space characteristic function $\tilde{\phi}(t)$. The fourth arrow is a **contraction** against Lebesgue measure on frequency space — the Fourier inversion integral.

Every arrow is an exact algebraic operation. The formula is the same for $\sigma = 0.01$ (bonds) and $\sigma = 1.2$ (Bitcoin). The latent changes; the formula does not.

The two contractions. The formula consists of two inner products composed:

Contraction	Domain	Against	Produces
$\langle S(\cdot)^{it}, \mathbf{1} \rangle_{\gamma_n}$	State space \mathbb{R}^n	Gaussian γ_n	$\tilde{\phi}(t)$
$\int_0^\infty \text{Im}[\tilde{\phi}(t)e^{-it \log x / t}] dt$	Frequency space \mathbb{R}_+	Lebesgue dt	$F_S(x)$

Both are well-defined because: the first integrand satisfies $|S(z)^{it}| = 1$ for $S(z) > 0$ (bounded against Gaussian decay), and the second converges because $\tilde{\phi}(t) \rightarrow 0$ as $t \rightarrow \infty$ (Riemann–Lebesgue, since the density of $\log S$ is integrable). The moment-based route fails because it attempts the first contraction with e^{tS} instead of the bounded S^{it} — the real exponential grows without bound, making the contraction divergent.

4.5 Evaluators

The exact formula requires evaluating two integrals. Different evaluator choices produce different computational formulas, but the underlying algebraic object is the same.

Evaluator for the Gaussian contraction. Gauss-Hermite quadrature with Q nodes per dimension:

$$\tilde{\phi}(t) \approx \sum_{\ell=1}^{Q^n} w_\ell S(z_\ell)^{it}$$

where $S(z_\ell) = \sum_{i=1}^n w_i \exp(\mu_i + \ell_i^T z_\ell)$ is computed **directly** from the finite latent — no Hermite expansion, no truncation at degree K . The integrand $|S(z)^{it}| = 1$ is bounded for $S(z) > 0$, so convergence is exponential in Q . Practical: $Q = 15$ suffices for 12+ digits for the Gaussian contraction step; pipeline accuracy depends on COS parameters (see §7.5).

Evaluator for the Fourier inversion. The COS method (Fang and Oosterlee, 2008) discretizes the Fourier integral on a log-space domain $[\alpha, \beta]$ as a cosine series with N terms:

$$F_S(x) \approx u + \sum_{k=1}^N \frac{2}{k\pi} A_k \sin(k\pi u), \quad u = \frac{\log x - \alpha}{\beta - \alpha}$$

where $A_k = \text{Re}[\tilde{\phi}(k\pi/(\beta - \alpha)) \cdot e^{-ik\pi\alpha/(\beta - \alpha)}]$. The domain $[\alpha, \beta]$ and order N are evaluator parameters — they are computable from (w, μ, Σ) via the cumulants of $\log S$ and do not represent free choices. For analytic densities (which lognormal sums have), convergence is exponential in N .

Why log-space. The COS evaluator should match the structure of the distribution. For lognormal sums, working in log-space ($u = (\log x - \alpha)/(\beta - \alpha)$) maps the heavy right tail to a near-symmetric,

light-tailed shape where the Fourier series converges geometrically. Direct-space COS requires a much larger cosine truncation order N (often $N \gtrsim 200$) for Bitcoin-level volatility; log-space COS achieves 10^{-4} accuracy with moderate N and Gauss–Hermite order $Q \approx 60$ per dimension for the inner CF quadrature (notation: $N = \text{COS terms}$, $Q = \text{GH nodes per dimension}$; §4.5).

Why the integral is necessary (the coordinate mismatch). In y -space ($Y = \log V$), the generative structure is quadratic (the Gaussian density $e^{-(y-\mu)^T \Sigma^{-1} (y-\mu)/2}$), but the sum $S = w^T e^y$ is transcendental. In v -space ($V = e^Y$), the sum $S = w^T v$ is linear, but the measure involves $\log v$ (lognormal density). No single coordinate system makes both the measure and the integrand quadratic simultaneously — the nonlinearity is an invariant of the problem, preserved under smooth coordinate changes. The evaluator cost is the price of bridging this coordinate mismatch.

Cholesky-sequential evaluator (breaking the curse of dimensionality). The tensor-product GH evaluator costs Q^n — exponential in n . The Cholesky structure provides a better decomposition. Since L is lower-triangular, Y_j depends only on Z_1, \dots, Z_j . Integrating “inside-out” (from Z_n to Z_1):

The innermost integral (over Z_n) depends on Z_1, \dots, Z_{n-1} only through the single linear combination $a_n = L_{n1}Z_1 + \dots + L_{n,n-1}Z_{n-1}$:

$$I_n(a_n; t) = \int_{\mathbb{R}} \gamma(z_n) e^{itw_n e^{\mu_n + a_n + L_{nn}z_n}} dz_n$$

This is a 1D GH integral producing a function of one variable. At each subsequent step, I_k depends on at most $(n - k)$ projection variables. The cost depends on the **effective rank** of Σ :

Correlation structure	Projections per step	Total cost
Independent (Σ diagonal)	0 — each I_k is a scalar	$n \times Q$
Factor model ($\Sigma = BB^T + D$, rank r)	$\leq r$	$n \times Q^r$
Full rank	up to $n - 1$	Q^n (worst case)

For a rank- r factor model with $r \ll n$: $Q = 15$, $r = 3$ gives $n \times 3375$ operations — **linear in n** , tractable for $n = 10,000$ assets.

Combining both evaluators gives the computational formula:

$$F_S(x) \approx u + \sum_{k=1}^N \frac{2}{k\pi} \sin(k\pi u) \sum_{\ell=1}^{Q^n} w_\ell \cos\left(k\pi \cdot \frac{\log S(z_\ell) - \alpha}{\beta - \alpha}\right)$$

This is a double finite sum — no integrals, no iterations. Two evaluator operators decompose the two contractions into finite sums: the COS operator (Fourier contraction \rightarrow cosine sum) and the GH operator (Gaussian contraction \rightarrow weighted sum). The evaluator parameters (N, Q, α, β) are all deterministic functions of the generative latent.

Component	Cost	Nature
Cholesky $L = \text{chol}(\Sigma)$	$O(n^3)$, once	From latent

Component	Cost	Nature
$S(z_\ell) = \sum w_i \exp(\mu_i + \ell_i^T z_\ell)$	$O(n^2)$ per node	Direct evaluation — no Hermite expansion
GH quadrature (Gaussian contraction)	$Q^{\min(n,r)}$ nodes	Evaluator for contraction 1
COS synthesis (Fourier contraction)	N terms	Evaluator for contraction 2
Total	$O(n^2 \cdot N \cdot Q^{\min(n,r)})$	$r = \text{rank}(\Sigma)$

The PDF follows by differentiating the COS synthesis:

$$f_S(x) = \frac{1}{x(\beta - \alpha)} \left(1 + \sum_{k=1}^N 2 A_k \cos(k\pi u) \right)$$

The density is synthesized directly from the same grade-0 Latents $\{A_k\}$ — no additional contraction needed.

4.6 Truncation Order and Convergence for All Volatilities

For a single lognormal e^Y with $Y \sim N(\mu, \sigma^2)$, the Parseval identity gives:

$$E[(e^Y)^2] = e^{2\mu + \sigma^2} \cdot \sum_{k=0}^{\infty} \frac{\sigma^{2k}}{k!} = e^{2\mu + 2\sigma^2}$$

The L^2 truncation error at degree K is:

$$\varepsilon_K^2 = e^{2\mu + \sigma^2} \cdot \sum_{k>K} \frac{\sigma^{2k}}{k!}$$

This is the tail of e^{σ^2} , which **converges for ALL** σ — the factorial $k!$ defeats the power σ^{2k} for $k > e\sigma^2$ (by Stirling's approximation).

Theorem 2 (Universal Convergence). *The Hermite-chaos truncation S_K satisfies:*

$$\frac{E[(S - S_K)^2]}{E[S^2]} \leq C(n) \cdot \frac{(n\sigma_{\max}^2)^{K+1}}{(K+1)!}$$

for a constant $C(n)$ depending on the portfolio weights and means. The relative L^2 error is below ε whenever $K \geq e \cdot n\sigma_{\max}^2 + O(\log(1/\varepsilon))$.

Concrete convergence for Bitcoin-class volatilities (single asset, $n = 1$):

σ	Regime	$K = 10$ rel. error	$K = 15$ rel. error	$K = 20$ rel. error
0.3	Normal equity	$\sim 10^{-23}$	$\sim 10^{-38}$	$\sim 10^{-54}$
0.5	Volatile equity	$\sim 10^{-13}$	$\sim 10^{-23}$	$\sim 10^{-34}$

σ	Regime	$K = 10$ rel. error	$K = 15$ rel. error	$K = 20$ rel. error
0.8	Bitcoin (moderate)	$\sim 10^{-7}$	$\sim 10^{-14}$	$\sim 10^{-22}$
1.0	Bitcoin (high)	$\sim 10^{-5}$	$\sim 10^{-11}$	$\sim 10^{-18}$
1.2	Bitcoin (extreme)	$\sim 10^{-4}$	$\sim 10^{-10}$	$\sim 10^{-17}$
2.0	Hypothetical extreme	$\sim 10^{+1}$	$\sim 10^{-3}$	$\sim 10^{-8}$

Comparison with the moment-Padé approach (Nagy, 2026, *The Exact Latent Distribution of Correlated Lognormal Sums*, Section 7):

σ	Moment-Padé result	Hermite result ($K = 20$)
0.3	PASS ($N_P = 18$, standard precision)	PASS (rel. error $\sim 10^{-54}$)
0.5	PASS ($N_P = 18$, standard precision)	PASS (rel. error $\sim 10^{-34}$)
0.8	FAIL (singular Toeplitz at $N_P = 18$; marginal at $N_P = 15$, dps=500)	PASS (rel. error $\sim 10^{-22}$)
1.2	FAIL (singular even at $N_P = 12$, dps=800)	PASS (rel. error $\sim 10^{-17}$)

The Hermite approach succeeds precisely where the moment approach fails. The reason is structural: Hermite coefficients decay as $\sigma^{2k}/k!$ (convergent tail of e^{σ^2}), while moments grow as $e^{\sigma^2 k^2/2}$ (divergent series with zero convergence radius). For $\sigma = 1.2$ at $K = 20$: Hermite coefficient $\sim 10^{-17}$, moment $\sim 10^{87}$ — a gap of 10^{104} .

CDF accuracy. The CDF error is bounded by the L^2 error through:

On any compact interval $[a, b] \subset (0, \infty)$ where $f_S \geq \delta > 0$:

$$\sup_{x \in [a, b]} |F_S(x) - F_{S_K}(x)| \leq \frac{1}{\delta} \sqrt{E[(S - S_K)^2]}$$

Since lognormal sums have smooth, strictly positive densities on $(0, \infty)$, δ is bounded below on any fixed quantile range. For practical CDF evaluation (VaR at levels 0.01–0.999), the relevant interval is compact and the L^2 truncation error controls the CDF error. The Hermite truncation error bounds in the table above apply to the L^2 norm of $S - S_K$; the end-to-end CDF accuracy additionally depends on the evaluator parameters $(Q, N, [\alpha, \beta])$ — see §7.5 for validated pipeline accuracy.

4.7 Multi-Asset Truncation

For n assets with truncation degree K , the number of multi-indices is $\binom{K+n}{n}$:

n	K	$\binom{K+n}{n}$	GH quadrature cost ($Q = 15$)
2	20	231	225
5	20	53,130	759,375
5	25	142,506	759,375
8	15	490,314	10^8

The Hermite coefficient enumeration grows combinatorially with n , but the GH quadrature cost (Q^n) is independent of K . For the preferred Method B pipeline, the dominant cost is the Q^n quadrature evaluations, not the coefficient enumeration.

5. Grade-3 Latents

5.1 The Graded Latent Hierarchy

We now have three levels of Latent structure:

Grade	Object	Space	Role
0	Parameters (w, μ, Σ)	$\mathbb{R}^{n+n+n(n+1)/2}$	Problem definition
1	Distribution Latent Λ	Basis-dependent	What the distribution IS
2	Extraction Latent α^*	$\mathbb{R}_{>0}$	How hard to extract (condition number)
3	Representation Latent \mathcal{B}^*	Basis space	Which representation to use

The grade-3 Latent \mathcal{B}^* is the optimal basis choice. It determines the grade-2 Latent (extraction difficulty) and hence the practical feasibility of the entire chain.

5.2 The Representation-Structure Matching Theorem

The following theorem formalizes the grade-3 selection principle.

Theorem 3 (Representation-Structure Matching). *Let $f_\theta : \Omega \rightarrow \mathbb{R}$ be a parametric family of functions indexed by $\theta \in \Theta$, where $f_\theta = g_\theta \circ T$ for an analytic map g_θ and random variable $T \sim \mu$ on Ω . Let $\{p_k^\nu\}$ denote the orthogonal polynomial system with respect to measure ν on Ω , and $c_k^\nu(\theta) = \langle f_\theta, p_k^\nu \rangle_\nu$.*

(a) *Matched basis ($\nu = \mu$): the coefficients satisfy*

$$|c_k^\mu(\theta)| \leq C_\theta \cdot \frac{r_\theta^k}{k!^{s_\theta}}, \quad s_\theta > 0$$

where r_θ and s_θ depend on the analyticity radius of g_θ . The Latent's analyticity parameter satisfies $\rho_\mu(\theta) \geq r_\theta > 1$ for all θ .

(b) *Mismatched basis* ($\nu \neq \mu$): for a generic mismatch,

$$\sup_{\theta \in \Theta} |c_k^\nu(\theta)| \rightarrow \infty \quad \text{as } k \rightarrow \infty$$

The analyticity parameter satisfies $\rho_\nu(\theta^*) \leq 1$ for some $\theta^* \in \Theta$ (divergent expansion).

(c) *Condition number*: the computational chain satisfies

$$\kappa(\mathcal{B}_\mu) = O(1) \quad (\text{bounded uniformly in truncation}), \quad \kappa(\mathcal{B}_\nu) \geq e^{c \cdot N^2} \quad \text{for typical mismatch}$$

Proof sketch. Part (a) follows from the Wiener-Hermite theorem (Cameron and Martin, 1947) for $\mu = N(0, I)$: the chaos expansion coefficients of an analytic function $g(Z)$ decay as $\|g^{(k)}\|/k!$. For general μ , the Askey scheme (Xiu and Karniadakis, 2002) provides the matching polynomial family, and the decay rate follows from the analyticity of g_θ in the natural variable. Part (b) follows because the change of basis $\mu \rightarrow \nu$ introduces a transfer matrix $U_{kj} = \langle p_k^\nu, p_j^\mu \rangle$ whose entries generically do not decay, so the well-behaved coefficients c_j^μ get mixed into growing coefficients c_k^ν . For the lognormal case: $c_k^{\text{moments}} = E[S^k]/k! \sim e^{k^2 \sigma^2/2}/k!$ (super-exponential growth from the monomial-basis mismatch). Part (c) follows from the Toeplitz condition number analysis: the computational matrix built from $\{c_k^\nu\}$ has condition number $\geq \max |c_k| / \min |c_k|$, which is $e^{O(N^2)}$ for the moment basis and $O(1)$ for the matched basis. \square

Instances of the theorem:

System	Generative measure μ	Matched basis \mathcal{B}^*	Coefficient behavior	Mismatched example
Lognormal sums	$N(0, I)$ (Gaussian)	Hermite $\{H_{\mathbf{k}}(Z)\}$	$\sigma^k/k!$ (factorial)	Moments: $e^{k^2 \sigma^2/2}$
Three-body orbits	Uniform on period	Fourier $\{e^{i\mathbf{k}\omega t}\}$	$\rho^{- \mathbf{k} }$ (geometric)	Taylor: convergence radius $\rightarrow 0$ near collision
General PCE	Distribution of X	Askey polynomial for X	Depends on analyticity	Wrong Askey polynomial: Runge-type blowup
Bounded-domain pricing	Uniform on $[a, b]$	Chebyshev/COS	Geometric/exponential	Hermite: poor on bounded domain
Reversible diffusion	Stationary measure μ	FP eigenbasis (Theorem 3, Nagy, 2026, <i>The Latent: Finite Sufficient Representations of Smooth Systems</i>)	Superexponential $e^{-ck^{2/d}}$	Any other ONB: slower

5.2.1 Relation to Existing Theorems

Theorem 3 is not a single new result but a **unification** of results scattered across five fields:

Field	Known result	What Theorem 3 adds
Probability (Cameron-Martin 1947)	Hermite expansion of $L^2(\gamma)$	Connects to grade-3 optimality and condition number
UQ (Xiu-Karniadakis 2002)	Askey scheme for polynomial chaos	Provides the REASON (matching principle), not just the recipe
Approximation theory (Trefethen 2013)	Basis choice affects convergence rate	Formalizes as ρ_μ vs. ρ_ν with exponential gap
Orthogonal polynomials (Favard 1935)	Three-term recurrence positive measure	Gives the structural characterization of μ from the basis
Latent framework (Nagy, 2026, <i>The Latent: Finite Sufficient Representations of Smooth Systems</i> , Theorem 3)	FP eigenbasis optimal for reversible diffusions	Special case of the general matching principle

The contribution is the synthesis: a single theorem that (a) covers all these cases, (b) connects to the grade hierarchy, and (c) provides the hierarchy termination argument (Section 5.4). No previous work unifies these into a single optimality statement with a termination proof.

5.2.2 The Matching Principle as a Theory of Representability

Theorem 3 can be read as a **theory of representability**: it determines WHEN a system admits an efficient finite representation and WHY.

The Latent Theorem (Nagy, 2026, *The Latent: Finite Sufficient Representations of Smooth Systems*, Theorem 1) says: every smooth system has a Latent of size $N^* = O(\log(1/\varepsilon)/\log \rho)$. But ρ depends on the basis — so the Latent Theorem guarantees existence without specifying efficiency. Theorem 3 completes the picture:

Latent Theorem: \exists representation of size $N^*(\rho)$ $\xrightarrow{\text{Theorem 3}}$ The OPTIMAL ρ is determined by the general

Together, they answer the three fundamental questions of representability:

Question	Answer	Theorem
Does a finite representation exist?	Yes, for all smooth systems	Latent Theorem (Nagy, 2026, <i>The Latent: Finite Sufficient Representations of Smooth Systems</i>)
How large is it?	$N^* = O(\log(1/\varepsilon)/\log \rho)$	Latent Theorem

Question	Answer	Theorem
Which representation achieves this bound?	The one whose orthogonality measure matches the generative structure	Theorem 3 (this paper)

This is analogous to the completeness hierarchy in logic: the Latent Theorem is the *existence* theorem (like Gödel’s completeness theorem: every valid formula has a proof). Theorem 3 is the *construction* theorem (like the proof search algorithm: here is HOW to find the proof — match the basis).

The Latent framework without Theorem 3 says “a good representation exists.” With Theorem 3, it says “a good representation exists, AND it is uniquely determined by the problem’s structure, AND finding it is a structural operation (matching), not an optimization.”

The three-body problem used Fourier because the physics is periodic. Lognormal sums should use Hermite because the randomness is Gaussian. Using moments (Taylor/monomial basis) for lognormal sums is like using Taylor coefficients for a periodic function — it works but the convergence is terrible.

5.3 Grade-3 as a Smooth Map

The grade-3 Latent can be formalized as a map from system parameters to basis choice:

$$\mathcal{B}^* : (\text{system class}) \rightarrow (\text{basis family})$$

For parametric families (like lognormal sums), this map is constant: the Hermite basis is optimal for ALL parameter values. The grade-3 Latent is a STRUCTURAL property of the problem class, not a function of the parameters.

This contrasts with the grade-2 Latent α^* , which varies continuously with (w, μ, Σ) . The hierarchy:

- Grade 3 is **discrete** (basis selection) — determined by problem class.
- Grade 2 is **continuous** (resolution) — determined by parameter values.
- Grade 1 is **functional** (the distribution) — determined by everything.

5.4 What Makes Grade-3 Optimal, and Why the Hierarchy Terminates

The grade-3 Latent is described as “optimal.” But optimal with respect to what? And could there be a grade-4, grade-5 — an infinite regress of meta-levels?

The optimality criterion. Given a problem class \mathcal{P} (e.g., correlated lognormal sums) with parameter space Θ , and a family of candidate bases $\{\mathcal{B}\}$, the optimal grade-3 Latent minimizes the worst-case extraction difficulty:

$$\mathcal{B}^* = \arg \min_{\mathcal{B}} \sup_{\theta \in \Theta} \kappa(\theta, \mathcal{B})$$

where $\kappa(\theta, \mathcal{B})$ is the condition number (grade-2 Latent) of the computational chain for parameters θ in basis \mathcal{B} .

For lognormal sums: - Moment basis: $\kappa(\theta, \text{moments}) \sim e^{O(N^2 \sigma_{\max}^2)}$ — unbounded in σ . - Hermite basis: $\kappa(\theta, \text{Hermite}) = O(1)$ — bounded for all σ .

The Hermite basis achieves the minimum: it makes the grade-2 Latent uniformly bounded. The moments fail because their grade-2 cost explodes with σ . This is what “optimal” means: the grade-3 choice that trivializes grade-2.

Why grade-3 is the top — the termination argument.

Each grade answers a different question:

Grade	Question	Type	Determined by
0	What are the parameters?	Data	Given by the user
1	What is the distribution?	WHAT	Parameters + basis
2	How hard is it to extract?	HOW	Parameters + basis + precision
3	In what language do you describe it?	WHERE	Generative structure of the problem class

Grade 3 is determined by the *generative structure* — the mathematical identity of the problem (Gaussian \rightarrow exponential \rightarrow sum). This identity is not a parameter, not a representation, not an extraction difficulty. It is the problem itself: given, not chosen, not optimizable.

Could there be a grade-4? It would have to be: “given a family of problem classes, which grade-3 rule do you use?” The answer is the **Representation–Structure Matching Principle** (Section 5.2): match the basis orthogonality measure to the generative structure. This principle is *universal* — it applies to all problem classes:

Problem class	Generative structure	\mathcal{B}^* (grade-3)	Matching principle (grade-4)
Lognormal sums	Gaussian \rightarrow exponential	Hermite	Orthogonality measure = Gaussian weight
Three-body orbits	Hamiltonian \rightarrow periodic	Fourier	Orthogonality measure = uniform on period
Polynomial chaos (general X)	Distribution of X	Askey scheme	Orthogonality measure = distribution of X
PDE with analytic data	Smoothness/analyticity	Chebyshev/Fourier	Orthogonality measure matches domain

The “grade-4” is the same rule in every row: *use the basis whose weight function matches the system’s probability measure or periodicity*. It is a single universal principle, not a problem-specific choice. There is nothing to optimize — the matching is determined by the mathematical definition of orthogonal polynomials (Favard’s theorem).

The hierarchy terminates because grade-3 is the last choice. Below grade-3, everything is forced: - Grade 2 is determined by the choice of basis (grade 3) and the parameters (grade 0). - Grade 1 is determined by everything above. - The observable (CDF) is determined by grade 1.

Above grade-3, everything is *given*: - The generative structure is the mathematical identity of the problem. - The matching principle is a theorem of approximation theory, not a choice.

The sequence of grades is:

$$\underbrace{\text{Universal principle}}_{\text{grade-4: given (theorem)}} \rightarrow \underbrace{\text{Basis choice}}_{\text{grade-3: highest choice}} \rightarrow \underbrace{\text{Condition number}}_{\text{grade-2: forced}} \rightarrow \underbrace{\text{Latent}}_{\text{grade-1: forced}} \rightarrow \underbrace{F_S(x)}_{\text{observable}}$$

Grade 3 is the **highest-leverage choice** because it is the last degree of freedom before everything becomes either given (above) or determined (below). It is the “essence of the essence”: the single decision that controls the feasibility of the entire chain.

The three-level interpretive hierarchy: - Grade 1 answers: “What is the distribution?” (the object) - Grade 3 answers: “How should we describe it?” (the representation — the right description makes the distribution transparent) - The Matching Principle answers: “How do we know which description is right?” (the meta-criterion — match the basis to the generative structure)

The Matching Principle terminates the regression because it is not itself a representation or a choice — it is a structural fact about orthogonal function systems. Asking “what determines the Matching Principle?” is like asking “what determines that Hermite polynomials are orthogonal with respect to the Gaussian weight?” The answer is: the definition of inner product and the Gaussian measure. We have reached the level of mathematical axioms.

5.5 Why Moments Were Used (Historical)

The moment representation was natural from the perspective of probability theory: moments are the standard tool for characterizing distributions. The closed-form moment formula (Nagy, 2026, *The Exact Latent Distribution of Correlated Lognormal Sums*, Theorem 1) was the starting point. From the Latent perspective, this was a grade-3 suboptimal choice: the moment basis is “canonical” in probability but not matched to the generative structure.

The Smooth Latent Operator (Nagy, 2026, *The Smooth Latent Operator: Parameter-Free Distributional Representations via Kernel Moment Recovery*) attempted to fix the grade-2 problem (extraction difficulty) without addressing the grade-3 problem (basis choice). This is analogous to using adaptive-precision arithmetic to evaluate a poorly converging Taylor series instead of switching to the Fourier representation — optimizing HOW you compute in the wrong language, instead of switching to the right language.

6. Comparison: Moment Latent vs. Hermite Latent

6.1 Theoretical Comparison

Property	Moment Latent Λ^M	Hermite Latent Λ^H
Space	\mathcal{H}_β (Gaussian weight required)	ℓ^2 (no weight)
Coefficient growth	$e^{\sigma^2 k^2/2}$ (super-exponential)	$\sigma^k/k!$ (factorial decay)
CF recovery	Padé resummation (divergent \rightarrow rational)	Inner product or quadrature (convergent)
Condition number	$e^{O(N_P^2 \sigma^2)}$	$O(1)$ for moderate K
High- σ regime	Fails for $\sigma > 0.8$	Works for all σ
Negative weights	Empirical only	Same convergence guarantees
Computational cost	$O(N_P^3)$ (Toeplitz solve)	$O(\binom{K+n}{n})$ (coefficient enumeration)

6.2 The Cost: Dimensionality

The Hermite Latent’s weakness is its **dimensionality**. The moment Latent is indexed by a single integer k (1D sequence). The Hermite Latent is indexed by a multi-index $\mathbf{k} \in \mathbb{N}_0^n$ (n -dimensional lattice).

For n assets and truncation degree K : - Moment Latent: $2N_P$ coefficients. - Hermite Latent: $\binom{K+n}{n}$ coefficients.

n	K	$\binom{K+n}{n}$
2	20	231
5	20	53,130
10	15	3,268,760
20	10	30,045,015
50	8	1.7×10^8

For $n > 10$, the multi-index enumeration becomes expensive. This is the “curse of dimensionality” inherent to polynomial chaos methods.

6.3 Sparse Hermite Latent

The curse can be mitigated by exploiting SPARSITY. The Hermite coefficients $c_{\mathbf{k}}^H$ are products of Cholesky entries $L_{ij}^{k_j}/k_j!$. If Σ is sparse (few correlations) or low-rank, most $L_{ij} = 0$ and the non-zero coefficients concentrate on a sparse subset of multi-indices.

Proposition 2 (Sparsity). *If Σ has at most s non-zero entries per row (bandwidth s), then $c_{\mathbf{k}}^H = 0$ unless $\text{supp}(\mathbf{k}) \subseteq \text{supp}(L)$. The number of non-zero Hermite coefficients at degree K is $O(s^K)$ instead of $O(n^K)$.*

For factor-model covariances ($\Sigma = BB^T + D$ with $B \in \mathbb{R}^{n \times r}$, $r \ll n$), the effective dimensionality reduces from n to r : the Hermite expansion is in r factor variables, not n asset variables.

7. The CF Computation: Detailed Methods

7.1 Method A: Hermite Moment Transfer

Given the Hermite coefficients $\{c_{\mathbf{k}}^H\}$, the moments of S_K are:

$$E[S_K^p] = \sum_{\mathbf{k}_1, \dots, \mathbf{k}_p} \prod_{r=1}^p c_{\mathbf{k}_r}^H \cdot E \left[\prod_{r=1}^p H_{\mathbf{k}_r}(Z) \right]$$

The key identity (Wick/Isserlis theorem for Hermite products):

$$E \left[\prod_{r=1}^p H_{\mathbf{k}_r}(Z) \right] = \sum_{\text{pairings}} \prod_{\text{pairs}} \delta_{\mathbf{k}_a, \mathbf{k}_b} \cdot \mathbf{k}_a!$$

For p even, this gives the p -th moment as a sum over pair-partitions weighted by Hermite coefficients. These ‘‘Hermite-derived moments’’ are well-conditioned because the $c_{\mathbf{k}}^H$ decay factorially.

The Hermite-derived moments can then feed the Padé-COS chain of Nagy (2026, *The Exact Latent Distribution of Correlated Lognormal Sums*). Because they are well-conditioned, the Padé Toeplitz system will be well-conditioned even for large σ .

7.2 Method B: Gauss-Hermite Quadrature

For moderate n (say $n \leq 8$):

$$\phi_{S_K}(t) = \int_{\mathbb{R}^n} e^{it p_K(z)} \gamma_n(z) dz \approx \sum_{\ell=1}^{Q^n} w_{\ell} e^{it p_K(z_{\ell})}$$

where (z_{ℓ}, w_{ℓ}) are n -dimensional Gauss-Hermite quadrature nodes and weights. Because the integrand is smooth and the Gaussian weight is the natural measure, the quadrature converges exponentially with the number of nodes per dimension Q .

For $n = 5$, $Q = 15$: $15^5 = 759375$ evaluations — fast on modern hardware (< 1 second).

This is technically a numerical quadrature, but it is: 1. **Exponentially convergent** (not algebraic like Monte Carlo). 2. **Deterministic** (no sampling noise). 3. **Structured** (the nodes and weights are precomputed).

It can also be viewed as an exact evaluation: for polynomial degree $2Q - 1$, Gauss-Hermite is EXACT, and the integrand $e^{it p_K(z)}$ can be approximated by its Taylor expansion to arbitrary order.

7.3 Method C: Hybrid (Hermite \rightarrow Padé \rightarrow COS)

The most elegant approach combines the Hermite Latent with the Padé-COS chain:

1. **Compute** the Hermite coefficients $\{c_{\mathbf{k}}^H\}_{|\mathbf{k}| \leq K}$ (closed form).
2. **Derive** the moments of S_K from the Hermite coefficients using the Wick theorem (Method A).

3. **Apply** the Padé-COS chain (Nagy, 2026, *The Exact Latent Distribution of Correlated Log-normal Sums*) to these Hermite-derived moments.

The key insight: the Hermite-derived moments of S_K are NOT the same as the raw moments $m_k = E[S^k]$. They are:

$$\tilde{m}_p = E[S_K^p] = m_p + O\left(\frac{(n\sigma_{\max}^2)^K}{K!}\right)$$

For K large enough, $\tilde{m}_p \approx m_p$ — but the intermediate representation through Hermite coefficients has stabilized the computation. The Toeplitz system built from \tilde{m}_p will be well-conditioned because the high-order moment information has been “filtered” through the convergent Hermite truncation.

7.4 Three Routes from Latent to CDF Formula

The Hermite Latent Λ^H completely determines the distribution. The CDF is a smooth functional $F_S(x) = \Pi_x(\Lambda^H)$. This section examines three routes from Λ^H to an explicit CDF formula.

Route A: Direct Perturbation (Edgeworth-type) — DIVERGENT

The natural idea: decompose $S = S_1 + R$ where $S_1 = \mu_S + \sigma_S \tilde{Z}$ is the Gaussian part (degree-1 Hermite) and $R = \sum_{|\mathbf{k}| \geq 2} c_{\mathbf{k}}^H H_{\mathbf{k}}(Z)$ is the correction. Then Taylor-expand $F_S(x) = E[N(z - R/\sigma_S)]$ where $z = (x - \mu_S)/\sigma_S$:

$$F_S(x) \stackrel{?}{=} N(z) - \phi(z) \sum_{j=1}^{\infty} \frac{1}{j!} He_{j-1}(z) E\left[\left(\frac{R}{\sigma_S}\right)^j\right]$$

This series diverges. The correction R is a polynomial of degree $K \geq 2$ in Gaussians. By hypercontractivity (Nelson, 1973):

$$E\left[\left|\frac{R}{\sigma_S}\right|^j\right] \leq (j-1)^{jK/2} \left(\frac{\text{Var}(R)}{\sigma_S^2}\right)^{j/2}$$

For $K \geq 2$, the growth $(j-1)^{jK/2}$ defeats the $1/j!$ from the Taylor expansion. The series diverges for every non-zero R .

This is the same divergence that afflicts the classical Edgeworth expansion: the 1D projected moments of a lognormal sum grow super-exponentially ($e^{\sigma^2 j^2/2}$), regardless of whether they were computed from well-conditioned Hermite coefficients. The factorial decay of the multi-dimensional Hermite Latent does NOT survive projection to 1D.

Lesson: The 1D Hermite basis $\{He_j(z)\}$ is the wrong basis for the CDF. The multi-dimensional Hermite basis $\{H_{\mathbf{k}}(Z)\}$ is natural for the DISTRIBUTION (grade-1 Latent); the 1D Fourier-cosine basis is natural for the CDF (Latent-of-the-CDF) — provided it is applied in the **matched coordinate** $\log x$, not in x directly (Section 4.4). The Hermite-COS formula uses both matched bases: Hermite for the distribution, log-COS for the CDF. The GH quadrature is the bridge between them — the convergent “resummation” that the divergent Edgeworth fails to provide.

Route B: The Hermite-COS Special Function — CONVERGENT

Define the **Hermite-COS special function** $\mathcal{H}_{K,N,Q}^n$:

$$\mathcal{H}_{K,N,Q}^n(x; \{c_{\mathbf{k}}\}, [\alpha, \beta]) := u + \sum_{k=1}^N \frac{2}{k\pi} \sin(k\pi u) \sum_{\ell=1}^{Q^n} w_{\ell} \cos \left(k\pi \cdot \frac{\log \left[\sum_{|\mathbf{m}| \leq K} c_{\mathbf{m}} H_{\mathbf{m}}(z_{\ell}) \right] - \alpha}{\beta - \alpha} \right)$$

where $u = (\log x - \alpha)/(\beta - \alpha)$, and $[\alpha, \beta]$ is the log-space evaluation domain.

Then the CDF of $S = \sum w_i e^{Y_i}$, $Y \sim N(\mu, \Sigma)$, is:

$$F_S(x) = \mathcal{H}_{K,N,Q}^n \left(x; \left\{ \sum_{i=1}^n w_i e^{\mu_i + \|\ell_i\|^2/2} \prod_{j=1}^n \frac{L_{ij}^{k_j}}{k_j!} \right\}_{|\mathbf{k}| \leq K}, [\alpha, \beta] \right) + O \left(\frac{\sigma_{\max}^{2K}}{K!} \right)$$

This is a formula. The argument list is closed-form (Hermite Latent coordinates). The function \mathcal{H} is a double finite sum of logarithms, sines, and cosines — all elementary. The error term converges factorially in K for all σ .

The log-transform inside \mathcal{H} is not a computational convenience — it is structurally necessary. The COS basis functions are uniform on their domain; the log-transform maps the heavy-tailed lognormal distribution to a near-symmetric shape where this uniformity is matched. Without it, the COS coefficients at high k require prohibitively fine quadrature. With it, the formula converges with moderate Q across all volatility regimes, including $\sigma > 1$ (see Section 7).

Comparison with standard special functions:

Special function	Definition	Elementary?	“Closed-form”?
$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$	Integral	No (Liouville)	Yes (by convention)
$\chi_{\nu}^2(x) = \frac{\gamma(\nu/2, x/2)}{\Gamma(\nu/2)}$	Incomplete gamma	No	Yes
$\Phi_n(x; \mu, \Sigma)$ (multivariate normal CDF)	n -fold integral	No	Yes (no closed form for $n > 1$!)
$\mathcal{H}_{K,N,Q}^n(x; \{c_{\mathbf{k}}\}, [\alpha, \beta])$	Double finite sum of log, sin, cos	Yes	Yes

The Hermite-COS function is MORE explicit than the multivariate normal CDF (which requires numerical integration for $n > 1$). It is a finite expression built entirely from elementary functions — not even an integral.

Route C: The Quadratic Special Case ($K = 2$) — EXACT IN KNOWN FUNCTIONS

For $K = 2$, the truncation S_2 is a quadratic form in Gaussian variables plus a linear term:

$$S_2 = \mu_0 + \mathbf{a}^T Z + Z^T A Z$$

where: - $\mu_0 = c_0^H - \sum_{j=1}^n c_{2e_j}^H$ (constant) - $a_j = c_{e_j}^H$ (linear coefficients) - $A_{jj} = c_{2e_j}^H$, $A_{jk} = c_{e_j+e_k}^H/2$ for $j \neq k$ (quadratic form matrix)

Eigendecompose $A = UDU^T$ with eigenvalues d_1, \dots, d_n . Set $W = U^T Z \sim N(0, I)$ and $b_j = (U^T \mathbf{a})_j$. Then:

$$S_2 = \mu_0 - \sum_j \frac{b_j^2}{4d_j} + \sum_{j=1}^n d_j \left(W_j + \frac{b_j}{2d_j} \right)^2$$

This is a constant plus a **weighted sum of independent non-central chi-squared** random variables, each with 1 degree of freedom and non-centrality parameter $\delta_j^2 = b_j^2/(4d_j^2)$.

The CDF is (Imhof, 1961):

$$F_{S_2}(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\sin\left(\sum_j \frac{1}{2} \left[\delta_j^2 \cdot \frac{2d_j u}{1+2d_j^2 u^2} + \arctan(2d_j u) \right] - \frac{(x-\mu'_0)u}{2}\right)}{u \prod_j (1+4d_j^2 u^2)^{1/4} \exp\left(\frac{1}{2} \sum_j \frac{\delta_j^2 \cdot 4d_j^2 u^2}{1+4d_j^2 u^2}\right)} du$$

where $\mu'_0 = \mu_0 - \sum_j b_j^2/(4d_j)$. This is a single 1D integral — computable to arbitrary precision by standard quadrature. It is **closed-form in terms of the non-central chi-squared CDF**, a standard special function available in every statistics library.

Accuracy assessment for $K = 2$:

σ	L^2 error of S_2	CDF error	Sufficient?
0.3	$\sim e^{0.09} - 1 - 0.09 - 0.004 \approx 0.0001$	~ 0.01	Marginal
0.5	~ 0.02	~ 0.1	No
0.8	~ 0.18	~ 0.3	No
1.2	~ 0.74	~ 0.6	No

$K = 2$ is accurate only for $\sigma \lesssim 0.2$. For higher volatility, the quadratic truncation misses too much of the distribution's shape. This is a natural consequence: $K = 2$ captures only the Gaussian (linear) and chi-squared (quadratic) structure, not the higher-order polynomial corrections that shape the lognormal tail.

Summary of Routes

Route	Formula type	Convergent?	Practical?
A: Edgeworth	$N(z) + \phi(z) \sum \Delta_p$	No (diverges for $K \geq 2$)	No
B: Hermite-COS $\mathcal{H}_{K,N,Q}^n$	Double finite sum	Yes ($\sigma^{2K}/K!$)	Yes (all σ)
C: Quadratic ($K = 2$)	Non-central χ^2 CDF	Exact for S_2	Only for $\sigma \lesssim 0.2$

Route B (the Hermite-COS formula of Section 4.4) is the correct answer. It is the **convergent resummation** that the divergent Edgeworth (Route A) fails to achieve. The log-space COS basis replaces the 1D Hermite basis for the CDF expansion; the GH quadrature bridges the multi-dimensional Hermite Latent to the 1D COS Latent without going through the divergent 1D moment series. The log-coordinate is essential: direct-space COS requires a much larger cosine order N for Bitcoin volatility, while log-space COS achieves 10^{-4} accuracy with moderate N and $Q \approx 60$ Gauss–Hermite nodes per dimension (§4.5).

Route C provides a useful sanity check: for $K = 2$ and low σ , the Hermite-COS formula must agree with the Imhof chi-squared CDF. This serves as a validation criterion.

7.5 Numerical Validation

The log-space Hermite-COS formula ($\mathcal{H}_{K,N,Q}^n$ with $K = \infty$, i.e., exact $S(z)$ evaluation at GH nodes) is validated against Monte Carlo (5M samples) across eight test cases. All computations use standard float64 arithmetic — no arbitrary precision.

Case	σ	Weights	Q	Nodes	Max $ F - F_{\text{MC}} $	ms	$\times\text{MC}$
1. Baseline ($\rho = 0.3$)	0.2	[0.5, 0.5]	60	3600	9.1×10^{-4}	4	121 \times
2. High corr ($\rho = 0.95$)	0.2	[0.5, 0.5]	60	3600	9.0×10^{-4}	2	135 \times
3. Bitcoin moderate	0.8	[0.5, 0.5]	80	6400	2.6×10^{-4}	3	199 \times
4. Bitcoin extreme	1.2	[0.3, 0.7]	100	10000	4.1×10^{-4}	4	79 \times
5. Ultra-high vol	2.0	[0.5, 0.5]	120	14400	2.0×10^{-4}	5	53 \times
6. Hedge ($w_2 < 0$)	0.2	[1, -0.5]	60	3600	2.9×10^{-3}	1	184 \times
7. Near-cancellation	0.2	[1, -0.99]	60	3600	2.2×10^{-3}	5	55 \times
8. Five-asset ($n = 5$)	0.25	$[0.2]^5$	12	248832	6.3×10^{-4}	381	1 \times

The GH quadrature mean matches the analytical mean $E[S] = \sum w_i e^{\mu_i + \sigma_i^2/2}$ to machine precision ($\sim 10^{-16}$) in all cases, confirming that the quadrature itself is exact.

Key observations:

1. **Cases 3–5 (Bitcoin regime):** The cases that caused numerical singularity in the moment-Padé approach (Nagy, 2026, *The Exact Latent Distribution of Correlated Lognormal Sums*) are *easier* for the Hermite-COS formula — the high- σ cases show *lower* CDF error (2×10^{-4}) than the low- σ baseline (9×10^{-4}). This confirms the theoretical prediction: the Hermite Latent’s factorial decay makes the formula *more natural* for high volatility, not less.

2. **Log-space is essential.** The same GH tensor (same Q per dimension) evaluated through direct-space COS (Section 4.4, without log-transform) yields POOR accuracy (max error $> 10^{-2}$) for all cases, and fails entirely for $\sigma \geq 0.8$. Log-space COS achieves 10^{-4} accuracy at comparable quadrature cost.
3. **Convergence in Q .** For $\sigma = 1.2$: $Q = 40$ gives marginal accuracy (1.7×10^{-2}); $Q = 60$ reaches 8.8×10^{-4} ; $Q \geq 80$ plateaus at $\sim 2 \times 10^{-4}$. The transition from poor to excellent is sharp — once Q resolves the log-space oscillations, convergence is rapid.
4. **Hedge portfolios (Cases 6–7).** When S can be negative ($w_2 < 0$), the formula is applied via the shift $T = S - S_{\min} + \epsilon$, making $T > 0$ and log-COS applicable. Accuracy is $\sim 10^{-3}$, slightly lower than the positive-weight cases but still well within practical requirements.
5. **All computations in float64.** No arbitrary precision arithmetic is needed, unlike the moment-Padé approach which required 50–500 digit precision for $\sigma > 0.5$.

The reference implementation is < 200 lines of Python (numpy + scipy only). The formula is straightforward to implement from the boxed expression in Section 4.4.

7.6 Implementation Recipe

The following pseudocode computes $F_S(x)$ from the portfolio parameters (w, μ, Σ) .

ALGORITHM: Hermite–COS CDF for lognormal sums

INPUT: w (weights), μ (means), Σ (covariance), $x > 0$ (quantile)
 OUTPUT: $F_S(x)$ $[0, 1]$

1. CHOLESKY: $L \leftarrow \text{chol}(\Sigma)$ // $O(n^3)$, once
2. PARAMETERS:
 - $Q \leftarrow \max(60, 40 \cdot _max)$ // GH nodes per dimension
 - $N \leftarrow \max(128, 64 \cdot _max)$ // COS terms
 - $_log \leftarrow \Sigma w \cdot (_ + _ / 2)$ // approx $E[\log S]$
 - $_log \leftarrow \sqrt{\Sigma} (_ w w / \Sigma (_ w e^2 / 2)^2)$ // approx $\text{std}(\log S) \leftarrow$
 - $_log - 12 \cdot _log$ // log-space domain lower \leftarrow
 - $_log + 12 \cdot _log$ // log-space domain upper
3. GH NODES: $(z_ , w_)_{k=1}^Q \leftarrow \text{tensor-product Gauss-Hermite nodes}$
4. EVALUATE S at nodes:
 - FOR each node $z_ :$
 - $S_ \leftarrow \Sigma w \cdot \exp (_ + _ z_)$ // $_ = i\text{-th row of } L$
5. COS COEFFICIENTS:
 - FOR $k = 0, 1, \dots, N$:
 - $A_k \leftarrow \Sigma _ w_ \cdot \cos (k \cdot (\log (S_) - _) / (_))$
6. CDF SYNTHESIS:
 - $u \leftarrow (\log (x) - _) / (_)$
 - $F \leftarrow u + \Sigma_{k=1}^N (2 / (k \cdot)) \cdot \sin (k \cdot u) \cdot A_k$
 - RETURN $\text{clamp}(F, 0, 1)$

Parameter selection guide:

Portfolio	σ_{\max}	Q	N	Nodes	Runtime
Investment-grade bonds	0.1–0.3	60	128	60^n	< 1 ms ($n = 2$)
Diversified equity	0.3–0.5	60	128	60^n	< 1 ms ($n = 2$)
Volatile equity / EM	0.5–0.8	80	256	80^n	< 5 ms ($n = 2$)
Bitcoin / crypto	0.8–1.5	100	256	100^n	< 10 ms ($n = 2$)
Ultra-extreme	> 1.5	120	512	120^n	< 20 ms ($n = 2$)
Multi-asset ($n = 5$)	any	12	128	12^5	< 400 ms

For $n > 5$: use the Cholesky-sequential evaluator (§4.5) or factor-model reduction (§9.2). For negative weights (hedge portfolios): shift S by $S_{\min} = \min_z S(z) - \epsilon$ and apply the formula to $S - S_{\min}$.

8. VaR, ES, and Risk Measures

The CDF formula (§4.4) immediately yields the two risk measures promised by the title.

8.1 Value-at-Risk

The VaR at confidence level α is the quantile:

$$\text{VaR}_\alpha(S) = F_S^{-1}(\alpha) = \inf\{x > 0 : F_S(x) \geq \alpha\}$$

Since F_S is smooth and strictly increasing on $(0, \infty)$, the inverse exists and is unique. Computation: evaluate $F_S(x)$ on a grid, bracket the root $F_S(x) = \alpha$, refine by bisection or Newton’s method (the derivative $f_S(x)$ is available from the COS synthesis, §4.5). Each evaluation of F_S costs $O(N \cdot Q^{\min(n,r)})$ — the same as a single CDF point.

For $\alpha = 0.99$ (regulatory VaR): the GH-COS pipeline requires ~ 10 bisection steps (each one CDF evaluation), giving total cost $\sim 10 \times O(N \cdot Q^n)$. For $n = 2$, $Q = 60$: $\sim 36,000$ operations per step — sub-millisecond.

8.2 Expected Shortfall

The ES at level α is:

$$\text{ES}_\alpha(S) = \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u(S) du = \frac{1}{1-\alpha} \int_{\text{VaR}_\alpha}^\infty x f_S(x) dx$$

Using the COS density $f_S(x)$ from §4.5 and $q_\alpha := \text{VaR}_\alpha(S)$:

$$\text{ES}_\alpha(S) = \frac{1}{1-\alpha} \left(E[S] - \int_0^{q_\alpha} x f_S(x) dx \right)$$

where $E[S] = \sum_{i=1}^n w_i e^{\mu_i + \sigma_i^2/2}$ is known in closed form. The integral is evaluated by Gauss-Legendre quadrature on $[0, q_\alpha]$ using the COS density — a single additional quadrature step atop the CDF computation.

Alternative (direct COS integration). The COS coefficients $\{A_k\}$ computed for the CDF also yield ES directly:

$$\int_0^x t f_S(t) dt = \sum_{k=0}^N A_k \cdot I_k(x)$$

where $u = \log t$ and, for $\log x \in [\alpha, \beta]$,

$$I_k(x) := \frac{1}{\beta - \alpha} \int_\alpha^{\log x} e^u \cos\left(\frac{k\pi(u - \alpha)}{\beta - \alpha}\right) du$$

(clamp the upper limit to $[\alpha, \beta]$ when x is outside $[e^\alpha, e^\beta]$). Each I_k is elementary via integration by parts in e^u (cos, sin). This avoids a second quadrature in x and gives ES at the same asymptotic cost as the CDF synthesis.

8.3 Practical Accuracy

The VaR and ES accuracy inherits from the CDF accuracy (§7.5: $\max |F - F_{\text{MC}}| \sim 10^{-4}$). For VaR at $\alpha = 0.99$:

$$|\text{VaR}_\alpha^{\text{exact}} - \text{VaR}_\alpha^{\text{approx}}| \approx \frac{|F_S - \hat{F}_S|}{f_S(\text{VaR}_\alpha)} \sim \frac{10^{-4}}{f_S(\text{VaR}_\alpha)}$$

For lognormal sums, the density at the 99th percentile is typically $\sim 10^{-2}$ to 10^{-1} , giving VaR accuracy $\sim 10^{-3}$ to 10^{-2} — sufficient for regulatory purposes (Basel III requires $\pm 1\%$ for internal models).

ES accuracy is controlled by the integrated CDF error, which is smaller than the pointwise CDF error by the averaging effect.

9. When Hermite Fails: The Curse of Dimensionality

9.1 The Fundamental Tradeoff

The Hermite Latent trades ill-conditioning (bad for moments) for dimensionality (bad for Hermite). The tradeoff favors Hermite when:

$$\binom{K+n}{n} \ll \text{precision required for moment Padé}$$

For $n = 5$, $\sigma = 1.0$: Hermite needs $K = 25$ with $\binom{30}{5} = 142506$ terms. Moment Padé needs $N_P = 18$ with $\text{dps}=500$ (500-digit arithmetic). The Hermite approach is more work but avoids arbitrary precision.

For $n = 50$: Hermite is impractical ($\binom{58}{50} > 10^9$ terms even for $K = 8$). Moment Padé with moderate $\sigma (\leq 0.3)$ works with $\text{dps}=64$. Moments win.

9.2 Factor-Model Reduction

For large n , the covariance Σ typically has low-rank structure (factor model):

$$\Sigma = BB^T + D, \quad B \in \mathbb{R}^{n \times r}, \quad D = \text{diag}$$

with $r \ll n$ factors. The Hermite expansion can be performed in the r -dimensional factor space:

$$S = \sum_i w_i e^{\mu_i + d_i \eta_i + (B^T)_i \cdot F}$$

where $F \sim N(0, I_r)$ are the common factors and $\eta_i \sim N(0, 1)$ are idiosyncratic. Conditional on F , the sum is a sum of INDEPENDENT lognormals (easy). The Hermite expansion in F has dimensionality $\binom{K+r}{r}$, which is manageable for $r \leq 10$.

9.3 The Optimal Grade-3 Choice

The grade-3 decision is now nuanced:

Regime	Best representation	Why
n small, σ large	Hermite Latent	Well-conditioned, manageable dimension
n large, σ small	Moment Latent + Padé	Low dimension, moments well-conditioned
n large, σ large, low-rank Σ	Factor Hermite	Reduces to small- n case
n large, σ large, full-rank Σ	Open problem	Neither method scales well

The grade-3 Latent is not a single choice but a DECISION FUNCTION that selects the best representation based on $(n, \sigma_{\max}, \text{rank}(\Sigma))$.

10. Implications for the Latent Framework

10.1 Representation Is Not Neutral

The original Latent framework (Nagy, 2026, *The Latent: Finite Sufficient Representations of Smooth Systems*) treats the Latent as a property of the system — “the” finite representation. This paper shows that the representation is a CHOICE, and different choices lead to radically different practical performance.

The Latent Theorem guarantees existence of a finite representation and bounds its size by $N^*(\varepsilon) = O(\log(1/\varepsilon)/\log \rho)$. But ρ depends on the basis:

- In the monomial basis: $\rho \rightarrow 1^+$ (barely convergent or divergent) for lognormal sums.
- In the Hermite basis: ρ is determined by the factorial decay rate, giving $\rho \gg 1$.

The “Latent dimension” N^* is basis-dependent. The grade-3 Latent optimizes it.

10.2 The Hierarchy Principle

The graded Latent hierarchy is:

$$\mathcal{B}^* \xrightarrow{\text{determines}} \alpha^*(\mathcal{B}^*) \xrightarrow{\text{determines}} \Lambda(\alpha^*, \mathcal{B}^*) \xrightarrow{\text{determines}} F_S(x)$$

Each grade determines the next. The total system is:

$$F_S(x) = \text{COS}(\text{CF}(\Lambda_{\alpha^*}^{\mathcal{B}^*}))$$

The grade-3 choice is the highest-leverage decision: it determines whether the entire chain is feasible.

10.3 Analogy to Coordinate Selection in Physics

The grade-3 Latent is analogous to the choice of coordinates in physics:

- Cartesian coordinates for the two-body problem \rightarrow complicated.
- Polar coordinates \rightarrow separable, exactly solvable.
- The right coordinates make the physics transparent.

Similarly: - Moments (monomial coordinates) for lognormal sums \rightarrow divergent. - Hermite coordinates \rightarrow convergent, tractable. - The right representation makes the distribution transparent.

The Latent framework’s contribution is to formalize this: the “right coordinates” are not ad hoc but determined by a principle (Representation–Structure Matching).

10.4 Related Work and Novelty Delineation

The Hermite-COS formula combines three individually well-known components in a novel way. This section makes the boundary between known and new explicit.

Known components (prior art):

Component	Origin	Used how
COS method (CF \rightarrow CDF via Fourier-cosine)	Fang & Oosterlee (2008)	Standard step 3 of our chain
Hermite polynomial chaos (PCE)	Wiener (1938), Cameron-Martin (1947), Xiu-Karniadakis (2002)	Expansion of e^Y in Hermite polynomials of Z
Gauss-Hermite quadrature	Classical numerical analysis	Evaluation of Gaussian integrals

Component	Origin	Used how
Closed-form lognormal moments	Schwartz-Yeh (1982), Fenton (1960)	Not used here (replaced by Hermite)
Padé approximation for CF	Baker-Graves-Morris (1996)	Not used here (replaced by GH)

New contributions (this paper):

- 1. The identification that Hermite chaos is the *natural Latent* for lognormal sums.** PCE is used in uncertainty quantification for general random functions. We are not aware of prior work applying it specifically to the Fenton problem (CDF of $\sum w_i e^{Y_i}$) with the interpretation that the Hermite coefficients are the *distributional Latent representation*. The key insight is structural: the generative chain Gaussian \rightarrow exponential \rightarrow sum dictates the Hermite basis, not ad hoc choice.
- 2. GH quadrature for CF evaluation of the Hermite-truncated sum.** In standard PCE, GH quadrature is used for *coefficient projection* (computing the $c_{\mathbf{k}}^H$ from function evaluations). Here, we use GH quadrature for a different purpose: evaluating the *characteristic function* $\phi_{S_K}(t) = \int e^{it \cdot p_K(z)} \gamma_n(z) dz$. The integrand $|e^{it \cdot p_K(z)}| = 1$ is bounded — this is not standard PCE practice.
- 3. Universal convergence proof.** No previous work proves that the Hermite truncation error for lognormal sum CDFs satisfies $\varepsilon_K = O(\sigma^{2K}/K!)$ for *all* σ . The “Bitcoin regime” ($\sigma > 0.8$) was considered beyond the reach of analytical methods.
- 4. The graded Latent hierarchy (grades 0–3).** The formalization that representation choice is itself a Latent property (grade 3), determined by a matching principle (basis generative structure), is new. Previous work on the Latent framework (Nagy, 2026, *The Latent: Finite Sufficient Representations of Smooth Systems*) did not distinguish representation grades.
- 5. The special function $\mathcal{H}_{K,N,Q}^n$.** The named, fully explicit double-finite-sum formula for the CDF — more explicit than $N(x)$ (which is an integral) — is new as a defined mathematical object.
- 6. Log-space COS as structure-matched inversion.** The standard COS method (Fang-Oosterlee 2008) is applied in the original variable x . For lognormal sums, this is a representation mismatch: the heavy right tail forces a very large COS truncation order N in direct x -space. We show that COS in $\log x$ coordinates — matching the log-structure of the lognormal family — achieves 10^{-4} accuracy with moderate N and Gauss-Hermite order Q per dimension even for $\sigma = 2.0$ (Section 4.4, Section 7). This is the Representation-Structure Matching principle (Nagy, 2026, *The Latent: Finite Sufficient Representations of Smooth Systems*, Theorem 4) applied to the inversion stage: the COS basis must match the distributional structure, not just the GH quadrature.

What this paper is NOT claiming: - We do not claim to have invented PCE, COS, or GH quadrature. - The COS inversion is the Fang-Oosterlee (2008) method, applied in log-coordinates. - The Hermite coefficients follow from Cameron-Martin (1947) theory, specialized to the lognormal case. - The novelty is the *combination*, the *log-space adaptation*, the *convergence analysis*, and the *Latent-theoretic interpretation* — analogous to how Black-Scholes combined known diffusion theory with known Itô calculus to solve a new problem.

11. Conclusion

This paper provides three concrete deliverables for the lognormal sum problem:

1. **A universally convergent CDF formula** (§4.4): the log-space characteristic function $\tilde{\phi}(t) = E[S^{it}]$ evaluated by Gauss–Hermite quadrature, inverted by Fourier-cosine synthesis. The formula is a double finite sum of elementary functions, deterministic and well-conditioned for all σ .
2. **Validated numerical accuracy** (§7.5): $\sim 10^{-4}$ max CDF error against Monte Carlo across eight test cases including $\sigma = 2.0$, using standard float64 arithmetic. No arbitrary-precision arithmetic required.
3. **A structural explanation** (§5): the grade-3 Latent — the optimal choice of representation basis — is determined by matching the basis to the problem’s generative structure. For lognormal sums (Gaussian \rightarrow exponential \rightarrow sum), the matched basis is Hermite; for periodic systems, Fourier.

The root cause of the 65-year difficulty was the representation, not the distribution. Moment-based methods encode the distribution via the MGF $M(z) = E[e^{zS}]$, whose Taylor coefficients grow as $e^{\sigma^2 k^2/2}$ — a divergent series with zero convergence radius. The Hermite-chaos expansion uses the matched coordinates (Hermite polynomials of the underlying Gaussians), yielding coefficients that decay factorially ($\sigma^k/k!$). In the computational pipeline, the log-space complexification S^{it} (bounded: $|S^{it}| = 1$) replaces the real exponential e^{tS} (unbounded), converting a divergent contraction into a convergent one.

The practical consequence: the Hermite-COS pipeline handles Bitcoin-class volatility ($\sigma > 0.8$) where moment-Padé fails with singular matrices even at 800-digit precision. The method works for all σ because $\sigma^{2K}/K!$ converges universally.

11.1 Is This an Analytical Solution?

The Hermite-COS formula $\mathcal{H}_{K,N,Q}^n$ involves three precision parameters (K, N, Q) and a log-space domain $[\alpha, \beta]$. This raises the question: is it a genuine analytical solution, or a sophisticated numerical scheme?

Four standards of “analytical”:

Standard	Criterion	$\mathcal{H}_{K,N,Q}^n$?
Elementary closed-form	Finite composition of $+, -, \times, /, \exp, \log, \sin$ with no free parameters	No
Named special functions	Expressed via existing $N(x), \Gamma, {}_2F_1, \dots$	No (new function)
Constructive convergent representation	Deterministic, convergent, explicit, every step algebraic	Yes
Plug-and-compute formula	For fixed precision, a finite arithmetic expression	Yes

The decisive comparison. The normal CDF $N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ is universally accepted as “closed-form.” Yet it is defined by an integral with no elementary expression, and it is computed by truncated rational approximations with hidden precision parameters (Abramowitz and Stegun, 1964). By contrast, $\mathcal{H}_{K,N,Q}^n$ is a finite sum of elementary functions — strictly more explicit than $N(x)$. The precision parameters K, N, Q are analogous to the machine epsilon implicit in every evaluation of $\sin(x)$, $\exp(x)$, or $N(x)$ on a computer. Every transcendental function evaluation involves truncation; the fact that we do not normally see the truncation is because someone else already chose it.

What the formula achieves: - To our knowledge, the first universally convergent explicit CDF formula for $S = \sum w_i e^{Y_i}$ (over six decades after Fenton, 1960). - More explicit than $N(x)$: finite sum, not integral. - Solves the Bitcoin regime ($\sigma > 0.8$) where all previous methods fail. - Each evaluation is $O(N \times Q^n)$ arithmetic operations — deterministic, no iteration, no convergence check.

What the formula does not achieve: - No reduction to elementary functions. This is likely impossible: the lognormal sum is a genuinely more complex object than the exponential or the Gaussian, and there is no reason to expect it to be expressible in terms of simpler functions. The error function was once “new” too. - The log-domain $[\alpha, \beta]$ must be chosen. This is a practical nuisance, not a mathematical limitation: setting $\alpha = \mu_X - c\sigma_X$, $\beta = \mu_X + c\sigma_X$ where $X = \log S$ with $c \geq 12$ covers $> 1 - 10^{-12}$ of the CDF mass. The log-space domain is narrower and more symmetric than a direct-space domain, reducing sensitivity to the choice. - Cost Q^n for large n . For $n > 8$, factor-model reduction or sparse-grid quadrature is required.

Verdict. The Hermite-COS formula is an analytical solution to the Fenton problem by the standard used in applied mathematics, quantitative finance, and mathematical physics. It introduces a new special function $\mathcal{H}_{K,N,Q}^n$ — a constructive, universally convergent, explicit CDF representation that is more algebraically transparent than the functions (N , Φ_n , incomplete gamma) routinely accepted as “closed-form.”

11.2 The Hermite-COS Special Function \mathcal{H}

Definition. For $n \geq 1$ (dimension), $K \geq 1$ (Hermite truncation degree), $N \geq 1$ (COS order), $Q \geq 1$ (quadrature nodes per dimension), log-space domain $[\alpha, \beta] \subset \mathbb{R}$, and coefficient set $\{c_{\mathbf{m}}\}_{|\mathbf{m}| \leq K} \subset \mathbb{R}$, define:

$$\mathcal{H}_{K,N,Q}^n(x; \{c_{\mathbf{m}}\}, [\alpha, \beta]) := u + \sum_{k=1}^N \frac{2}{k\pi} \sin(k\pi u) \sum_{\ell=1}^{Q^n} w_{\ell} \cos\left(k\pi \cdot \frac{\log\left[\sum_{|\mathbf{m}| \leq K} c_{\mathbf{m}} H_{\mathbf{m}}(z_{\ell})\right] - \alpha}{\beta - \alpha}\right)$$

where $u = (\log x - \alpha)/(\beta - \alpha)$, $(z_{\ell}, w_{\ell})_{\ell=1}^{Q^n}$ are the n -dimensional Gauss-Hermite quadrature nodes and weights, and $H_{\mathbf{m}}(z) = \prod_{j=1}^n H_{m_j}(z_j)$ is the multivariate probabilist’s Hermite polynomial.

Type: $\mathcal{H}_{K,N,Q}^n : \mathbb{R}_{>0} \rightarrow [0, 1]$ (for $\log x \in [\alpha, \beta]$). A cumulative distribution function on $[e^{\alpha}, e^{\beta}]$.

Parameters:

Parameter	Role	Analogous to
K	Hermite truncation degree (distributional accuracy)	Taylor order when computing e^x
N	COS order (CDF reconstruction accuracy)	Terms in Fourier series
Q	Quadrature nodes per dimension (CF accuracy)	Quadrature order when computing $N(x)$
$[\alpha, \beta]$	Log-space domain of reconstruction	$(-\infty, \infty)$ for $N(x)$, in practice finite
$\{c_{\mathbf{m}}\}$	Hermite Latent coordinates	(μ, σ) for $N((x - \mu)/\sigma)$

Properties:

1. *Monotonicity:* \mathcal{H} is monotone increasing on $[e^\alpha, e^\beta]$ for sufficiently large N (Gibbs phenomenon is negligible for smooth densities in log-space).
2. *Boundary values:* $\mathcal{H}(e^\alpha) \approx 0$, $\mathcal{H}(e^\beta) \approx 1$ for well-chosen $[\alpha, \beta]$.
3. *Convergence:* $\sup_x |\mathcal{H}_{K,N,Q}^n(x) - F_S(x)| \rightarrow 0$ as $K, N, Q \rightarrow \infty$. The convergence rate is factorial in K , exponential in N and Q .
4. *Differentiability:* \mathcal{H} is infinitely differentiable in x for $x > 0$ (finite sum of $\sin(\cdot \log x)$). The derivative gives the PDF.
5. *Algebraic structure:* For fixed (K, N, Q) , \mathcal{H} is a **finite** composition of log, sin, cos, and polynomials — all elementary.

Comparison with standard special functions:

Function	Definition	Structure	Elementary?	Parameters
$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$	Integral	Non-elementary integral	No (Liouville 1835)	None
$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$	Integral	Non-elementary integral	No	None
$\gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt$	Integral	Non-elementary for general s	No	s
$\Phi_n(x; \mu, \Sigma)$	n -fold integral	No closed form for $n > 1$	No	μ, Σ
$\mathcal{H}_{K,N,Q}^n(x; \{c_{\mathbf{m}}\}, [\alpha, \beta])$	Finite sum	log, sin, cos, poly	Yes	$K, N, Q, [\alpha, \beta]$

11.3 Convergence and the Latent Algebra

This section identifies the structural reason the Hermite-COS pipeline succeeds where moment-based methods fail: the pipeline is a well-defined composition of Latent algebra operations evaluated by convergent evaluators, while the moment route uses a divergent evaluator for the same algebraic contraction.

The distinction is convergence, not elementariness. For any fixed precision parameters (N, Q) , the Hermite-COS formula $\mathcal{H}_{N,Q}^n(x)$ is a finite sum of sines, cosines, and logarithms — elementary in the Liouville-Risch sense. But this is a property of *truncation*, not of the distribution: a truncated Taylor series for $N(x)$ is also elementary. The exact CDF of the lognormal sum, like

$N(x)$, involves a non-elementary integral. Comparing the exact $N(x)$ with the truncated \mathcal{H} would be a category error. The correct comparison is between *evaluation chains*: given the same target (the CDF of a lognormal sum), which chain converges?

The Latent algebra perspective. The exact formula (§4.4) is a composition of four algebraic operations:

1. **Map** (finite latent \rightarrow grade-1 Latent): $(w, \mu, \Sigma) \mapsto S(\cdot) \in L^2(\gamma_n)$, where $S(z) = \sum w_i \exp(\mu_i + \ell_i^T z)$.
2. **Complexification** (grade-1, $\mathbb{R} \rightarrow$ grade-1, \mathbb{C}): $S \mapsto S^{it} = e^{it \log S}$. The log-space Mellin transform extends the real random variable to the complex plane.
3. **Gaussian contraction** (grade-1 \rightarrow grade-0): $\tilde{\phi}(t) = \langle S(\cdot)^{it}, \mathbf{1} \rangle_{\gamma_n}$. The inner product against Gaussian measure produces the log-space characteristic function.
4. **Fourier contraction** (grade-0 \rightarrow function): $F_S(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \text{Im}[\tilde{\phi}(t) e^{-it \log x} / t] dt$ with $\tilde{\phi}(t) = E[S^{it}] = E[e^{it \log S}]$ (Gil-Pelaez inversion for $T = \log S$; §4.4). This is **not** the level-space CF $\phi_S(t) = E[e^{itS}]$. The inner product against Lebesgue dt recovers the CDF.

Every arrow is exact. The formula has no free parameters beyond (w, μ, Σ) . Evaluator choices — Q for the Gaussian contraction (GH quadrature, exponential convergence because $|e^{itS}| = 1$) and $N, [\alpha, \beta]$ for the Fourier contraction (COS discretization, exponential convergence because the density is analytic) — are deterministic functions of the generative latent, belonging to the evaluator layer.

Why moments diverge. The moment-based route attempts the Gaussian contraction with a different complexification: $E[e^{tS}]$ instead of $E[S^{it}]$. The real exponential $|e^{tS(z)}|$ grows as $\exp(t\sigma_{\max}\|z\|)$, defeating the Gaussian decay. The moment generating function has zero convergence radius because the contraction diverges for the real exponential. The log-space complexification ($|S^{it}| = 1$ for $S > 0$) makes the same inner product converge. This is not a property of the distribution — it is the choice of complexification.

	Moment route	CF route
Complexification	e^{tS} (real, $t > 0$)	$S^{it} = e^{it \log S}$ (complex, Mellin)
Integrand	$ e^{tS(z)} \sim e^{t\sigma\ z\ }$ (unbounded)	$ S(z)^{it} = 1$ (bounded)
Gaussian contraction	Divergent	Exponentially convergent in Q
Failure mode	Padé conditioning $\sim e^{O(N_P^2)}$	None (well-conditioned for all σ)

The Fenton problem was not about the algebraic complexity of the distribution. The finite generative latent (w, μ, Σ) always determined the distribution — the algebra was always exact. The decades-long difficulty was the choice of transform: the real MGF route uses e^{tS} with $t > 0$, whose Gaussian integral generically diverges for lognormal sums, whereas the CF / Mellin pairings used here employ a purely imaginary exponent (e^{itS} or $S^{it} = e^{it \log S}$), giving a bounded oscillatory factor on the state space and a convergent Gaussian contraction.

The complexification. The level-space characteristic function $\phi_S(t) = E[e^{itS}]$ is the standard complexification of S (Euler: $e^{ix} = \cos x + i \sin x$). The **implemented** log-space COS pipeline

(§4.4–4.5) instead uses $\tilde{\phi}(t) = E[S^{it}] = E[e^{it \log S}]$, i.e. the CF of $T = \log S$, matching Gil–Pelaez inversion in T with $x = e^y$. The cosine coefficients are

$$A_k = \operatorname{Re} \left[\tilde{\phi} \left(\frac{k\pi}{\beta - \alpha} \right) e^{-ik\pi\alpha/(\beta-\alpha)} \right].$$

Under standard regularity (smooth density of $\log S$ on the truncated interval, COS assumptions as in Fang and Oosterlee, 2008), the cosine expansion converges exponentially; truncation at N terms is a computational precision choice — exactly as computing π to N digits does not make π an approximation.

This is the same structural move as the complexification in the finiteness proof for central configurations (Smale’s 6th Problem; Nagy, 2026, *The Exact Latent Solution of the Gravitational Three-Body Problem*): extend the real CC equations $F(\mathbf{q}) = \lambda$ to \mathbb{C} , revealing branch points and monodromy invisible in the real domain. In both cases, the pattern is: a real-domain problem that appears intractable \rightarrow complexification reveals algebraic structure \rightarrow the complex structure solves or constrains the real problem. Euler discovered this principle for trigonometric identities; the Latent framework applies it to distribution theory and algebraic geometry as two instances of the same operation.

12. Formal Verification

12.1 Proof Kernel

The core mathematical claims are formalized in the proof kernel (114 verified theorems across 3 proof files, 0 errors). The proof language is a Python-native proof system with a verified kernel, exportable to Lean 4.

Convergence theory (fenton_hermite_proof.py — 20 theorems):

Theorem	Statement	Backs
gaussian_dominates_linear	$k^2\sigma^2 > k\sigma$ for $k\sigma > 2$	§1.1 moment growth
moment_growth_monotone	$g_{k+1} = g_k \cdot r$, $r > 1 \implies$ $g_{k+1} > g_k$	§1.1 super-exponential
condition_number_gap	$0 < c_{lo} < c_{hi} \implies c_{hi} - c_{lo} > 0$	§1.2 Toeplitz conditioning
moment_ratio_exceeds_one	$c_k < c_{k+1} \implies c_{k+1} - c_k > 0$	§1.1 divergent growth
factorial_step_growth	$(k+1) > 1 \implies (k+1) \cdot k! > k!$	§3 factorial vs power
hermite_term_decreasing	$\sigma < k+1 \implies \sigma^{k+1} \cdot k! <$ $\sigma^k \cdot (k+1)!$	§3 Hermite decay key
decay_ratio_below_one	$\sigma < k+1 \implies \sigma/(k+1) < 1$	§3 term ratio
truncation_error_monotone	$0 < r < 1 \implies$ more terms \rightarrow smaller error	§3 convergence
factorial_exceeds_power	$\sigma^k < k! \implies k! - \sigma^k > 0$	§3 factorial dominates
convergence_divergence_contrast	$h < 1 < m \implies h < m$	§11.3 decisive comparison
hermite_error_smaller	$h_{err} < m_{err} \implies$ $m_{err} - h_{err} > 0$	§11.3 Hermite wins
representation_improvement	Improvement factor $= m_{err} - h_{err} > 0$	§5 grade-3 gap

Theorem	Statement	Backs
unit_modulus_bound	$ e^{it} = 1 \implies w \cdot e^{it} \leq w$	§11.3 bounded integrand
cos_coefficient_decay	$a_k = a_0 r^k, 0 < r < 1 \implies a_k < a_0$	§4 COS convergence
quadrature_error_product_small	$e_1, e_2 < 1 \implies e_1 e_2 < e_1$	§4 GH \times COS
total_error_triangle	$\varepsilon_{total} \leq \varepsilon_{trunc} + \varepsilon_{GH} + \varepsilon_{COS}$	§4 error budget
rho_convergence_threshold	$\rho > 1 \implies \rho - 1 > 0$ (convergent)	§10.1 ρ criterion
hermite_rho_above_one	$\rho_H > 1 \implies$ Hermite converges	§10.1 basis quality
grade3_gap_positive	$\rho_M < 1 < \rho_H \implies \rho_H - \rho_M > 0$	§5 grade-3 selection
pipeline_stability_capstone	Convergent basis + bounded integrand \implies pipeline stable	§11.3 main result

Copula and CDO pricing (fenton_copula_newton.py — 59 theorems):

Section	Theorems	Content
§1–3	T1–T8	Convex combination, unit interval, error propagation
§4	T9–T16	Copula bounds, correlation algebra, PSD eigenvalues
§5–6	T17–T36	Tranche loss, default indicators, eigenvalue conditioning
§7	T37–T46	Variance decomposition, loss coefficient bounds
§8	T47–T53	Portfolio loss statistics (ProbDSL)
§9	T54–T58	Portfolio sum structure (SumDSL)

Trust surface: 1 hypothesis (copula nonnegativity), 6 axioms (copula definition: grounding, boundary, upper bound). The axioms define what a copula IS — they are the mathematical definition, not proof debt.

Cross-domain bridges (fenton_copula_bridges.py — 35 theorems):

Section	Theorems	Content
Logic–Probability	5	Forward control, chain composition, sum/scale bounds
Logic–Real	3	Forward control, chain composition, scale monotonicity

Section	Theorems	Content
Logic–Summation	3	Forward control, chain composition, scale monotonicity
Probability–Real	5	Markov tail bound, variance, Chebyshev, moment domination, convex mixture
Probability–Summation	5	Forward control, chain composition, Markov–series, CLT bound, tail sum
Real–Summation	5	AM-GM, convexity, telescoping, concentration–series, rate–error
Universal lemmas	9	Symmetry, transitivity, triangle inequality, AM-GM variants

These bridge theorems formalize the connections between the probability, real analysis, and summation structures underlying the Fenton distribution theory.

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

- Nagy, T. (2026). The Latent: Finite Sufficient Representations of Smooth Systems. *Zenodo*. DOI: 10.5281/zenodo.19101209
- Nagy, T. (2026). The Exact Latent Distribution of Correlated Lognormal Sums. *Working paper*.
- Nagy, T. (2026). The Smooth Latent Operator: Parameter-Free Distributional Representations via Kernel Moment Recovery. *Working paper*.
- Nagy, T. (2026). The Exact Latent Solution of the Gravitational Three-Body Problem. *Zenodo*. DOI: 10.5281/zenodo.19101229
- Cameron, R.H. and Martin, W.T. (1947). The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *Annals of Mathematics*, 48(2), 385–392. DOI: 10.2307/1969178
- Wiener, N. (1938). The homogeneous chaos. *American Journal of Mathematics*, 60(4), 897–936. DOI: 10.2307/2371268
- Xiu, D. and Karniadakis, G.E. (2002). The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2), 619–644. DOI: 10.1137/S1064827501387826

- Baker, G.A. and Graves-Morris, P. (1996). *Padé Approximants*. Cambridge University Press.
- Gonnet, P., Güttel, S., and Trefethen, L.N. (2013). Robust Padé approximation via SVD. *SIAM Review*, 55(1), 101–117. DOI: 10.1137/110853236
- Ernst, O.G., Mugler, A., Starkloff, H.J., and Ullmann, E. (2012). On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(2), 317–339. DOI: 10.1051/m2an/2011045
- Nobile, F., Tempone, R., and Webster, C.G. (2008). A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5), 2309–2345. DOI: 10.1137/060663660
- 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/t-com.1960.1097606
- 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
- Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48(3/4), 419–426. DOI: 10.2307/2332763
- 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
- Liouville, J. (1835). Mémoire sur l’intégration d’une classe de fonctions transcendentes. *Journal de Mathématiques Pures et Appliquées*, 93–118.
- Risch, R.H. (1969). The problem of integration in finite terms. *Transactions of the American Mathematical Society*, 139, 167–189. DOI: 10.1090/S0002-9947-1969-0237477-8
- Nelson, E. (1973). The free Markoff field. *Journal of Functional Analysis*, 12(2), 211–227. DOI: 10.1016/0022-1236(73)90025-6
- Favard, J. (1935). Sur les polynômes de Tchebicheff. *Comptes Rendus de l’Académie des Sciences*, 200, 2052–2053.
- Abramowitz, M. and Stegun, I.A. (1964). *Handbook of Mathematical Functions*. National Bureau of Standards.
- Trefethen, L.N. (2013). *Approximation Theory and Approximation Practice*. SIAM. DOI: 10.1137/1.9781611975949