The Universal Spectral Representation Theorem: Breaking the Curse of Dimensionality

Tamás Nagy, Ph.D.

tnagyphd@gmail.com

Working Paper

Abstract

How many parameters does it take to represent a smooth high-dimensional probability law to accuracy ε? We prove that the answer is $N = \Theta(\log(1/\varepsilon)/\log\rho)$, where $\rho > 1$ is the analyticity radius, and that this representation size is independent of the ambient dimension. For densities with analyticity radius $\rho > 1$, this many Fourier coefficients suffice (upper bound, constructive via the Eigen-COS method) and are necessary (lower bound, information-theoretic via Fourier mode counting). The bounds match, establishing the exact rate. The theorem is not tied to one application domain: the same law governs smooth distributional objects arising in portfolio aggregation, diffusion dynamics, celestial mechanics, collision-probability estimation, and knowledge representations in machine learning. The curse of dimensionality in representing smooth systems is therefore not fundamental; it is a consequence of using simulation- and sampling-based descriptions where a reusable spectral representation exists. The structural core is formally verified in Lean 4 (27 files across 3 proof libraries, capstone theorems verified).

---

1. Introduction

1.1 The Representation Question

Across probability, dynamics, and scientific computing, the same question keeps reappearing: how many degrees of freedom are genuinely needed to represent a smooth high-dimensional law to target accuracy ε? Standard numerical practice usually answers this question indirectly. One simulates paths, refines grids, samples particles, or expands the state space until the approximation is good enough. As the ambient dimension grows, these procedures become expensive enough that the cost is often mistaken for an intrinsic property of the object itself.

This paper argues that the ambient dimension is often the wrong control variable. For a broad class of smooth one-dimensional spectral representations extracted from high-dimensional systems, the true complexity is governed not by dimension but by regularity. In the analytic regime, the relevant quantity is the analyticity radius ρ, and the representation size grows only like $\Theta(\log(1/\varepsilon)/\log\rho)$.

1.2 Our Result

We prove a dimension-free representation law. In the risk-functional form used throughout this paper, it can be stated as follows:

Theorem (Universal Spectral Representation; Lean-verified). *Let $S = \sum_{i=1}^{n} w_i X_i$ where the joint density of $(X_1, \ldots, X_n)$ has analyticity radius $\rho > 1$ on the support of S. Then:*

(i) (Upper bound) *There exist* $N \leq C_1 \cdot \log(1/\varepsilon)/\log(\rho) + 2$ *real parameters from which every spectral risk measure* $\rho_\phi(S) = -\int_0^1 \phi(p)\, F^{-1}(p)\, dp$ *is computable to accuracy* ε*.*

(ii) (Lower bound) *Any representation with fewer than* $C_2 \cdot \log(1/\varepsilon)/\log(\rho)$ *parameters cannot achieve* ε*-accuracy for all spectral risk measures simultaneously.*

(iii) (Dimension-free) *Neither* C_1 *nor* C_2 *depends on* n*.*

The point is broader than this finance-shaped statement. The theorem says that once the relevant one-dimensional spectral object is identified, the required representation size is controlled by accuracy and regularity, not by the ambient dimension of the original system. Portfolio risk is one instance of this law, not its definition.

1.3 Relation to Shannon's Rate-Distortion Theorem

Shannon's Rate-Distortion Theorem (1959) answers: given an information source, what is the minimum number of bits to represent it with distortion $\leq D$? The answer depends on the source's entropy rate, not on the ambient dimension of the signal.

Our theorem is a representation-theoretic analogue. The "source" is a smooth distributional object. The "distortion" is the approximation error ε. The "rate" is $\Theta(\log(1/\varepsilon)/\log\rho)$, which depends on smoothness (via ρ), not on ambient dimension.

The parallel is precise:

Shannon (1959)	This paper
Source → bits → reconstruction	Distribution → parameters → reconstruction / functionals
Rate = H (source entropy)	Rate = $\log(1/\varepsilon)/\log\rho$
Independent of signal dimension	Independent of ambient dimension
Achiever: optimal codec	Achiever: Eigen-COS method

1.4 Related Work

Monte Carlo and variance reduction. The standard approach to portfolio VaR is Monte Carlo simulation (Glasserman, 2004), with variance reduction techniques (antithetic variates, importance sampling, control variates) reducing the constant but not the $O(n \cdot M)$ scaling. Multilevel Monte Carlo (Giles, 2008; 2015) reduces the cost by correlating simulations across resolution levels, achieving $O(\varepsilon^{-2})$ cost for accuracy ε — but the leading constant still grows with n.

Quasi-Monte Carlo and dimension reduction. Quasi-Monte Carlo (QMC) methods (Niederreiter, 1992) exploit smoothness to achieve $O(M^{-1+\delta})$ convergence rates independent of dimension in principle, but in practice the effective dimension — the number of directions that matter — determines the constant. Principal Component Analysis (PCA) reorders the integration variables by explained variance, enabling QMC to focus on the dominant directions (Acworth, Broadie, and Glasserman, 1997). Portfolio PCA further tailors the reduction to portfolio sensitivities, achieving 4–27× speedups. Recent work (2024–2025) combines randomized QMC with quadratic regression and Chebyshev interpolation for additional gains. Our result differs fundamentally: we show that the *representation* (not just the *integration*) is low-dimensional. QMC reduces the cost of evaluat-

ing a high-dimensional integral; we eliminate the integral entirely by compressing the distribution to N parameters.

Factor models and PCA for risk. Factor-based VaR (e.g., Barra, RiskMetrics) models portfolio returns as linear combinations of K factors, reducing the n-dimensional problem to K dimensions. Our eigenvalue conditioning (Section 2.4) is related: it decomposes the correlation matrix into principal components. The difference is that factor models approximate the *return distribution* parametrically (typically Gaussian), while the Eigen-COS method computes the *exact CDF* via Fourier inversion. The Universal Theorem quantifies this: the CDF error from K-factor conditioning is bounded by Lemma 5 of (Nagy, 2026a), and the total representation requires only $N = O(\log(1/\varepsilon)/\log\rho)$ parameters regardless of K.

Fourier and spectral methods in finance. The COS method (Fang and Oosterlee, 2008) and FFT-based pricing (Carr and Madan, 1999) use Fourier inversion for option pricing. Ruijter and Oosterlee (2012) extended COS to two dimensions via tensor products. Our Eigen-COS method is the first to combine eigenvalue conditioning with COS expansion, reducing the n-dimensional problem to a sequence of one-dimensional COS inversions. The Universal Theorem provides the theoretical justification: the Fourier-cosine expansion is near-optimal for smooth densities.

Information-theoretic bounds. The Kolmogorov ε-entropy of analytic function classes was established by Kolmogorov and Tikhomirov (1959) and further developed by Lorentz (1966) and Pinkus (1985). The rate-distortion function (Shannon, 1959) provides analogous bounds in communication theory. Recent work (2024–2025) connects distortion risk measures with rate-distortion theory for compressed estimation. Our contribution is the application of ε-entropy to portfolio risk measurement: we show that the "effective dimension" of the risk measurement problem is $\Theta(\log(1/\varepsilon)/\log\rho)$, independent of portfolio size. To our knowledge, this is the first such information-theoretic lower bound for financial risk computation.

Formal verification in mathematical finance. Echenim, Guiol, and Peltier (2019) formalized the Cox-Ross-Rubinstein binomial pricing model in Isabelle/HOL. Kudryashov (2022) formalized the Cauchy integral formula in Lean/Mathlib, which we use for the Bernstein coefficient bound. Degenne et al. (2025) formalized Brownian motion in Lean. The present paper's Lean formalization (12 files, 0 sorry) is, to our knowledge, the most extensive formal verification of a result in continuous-time quantitative finance.

1.5 Organization

Section 2 defines spectral risk measures and Fourier-cosine representations. Section 3 proves the upper bound (constructive). Section 4 proves the lower bound (information-theoretic). Section 5 states the combined Θ-result and its consequences. Section 6 provides numerical evidence. Section 7 discusses limitations and extensions. Appendix A summarizes the Lean formalization.

---

2. Preliminaries

2.1 Spectral Risk Measures

Definition 1 (Admissible Spectrum). A function $\phi : [0,1] \to \mathbb{R}_+$ with $\int_0^1 \phi(p)\,dp = 1$ is an admissible spectrum. The induced spectral risk measure is

$$\rho_\phi(S) = -\int_0^1 \phi(p)\, F_S^{-1}(p)\, dp,$$

where F_S^{-1} is the quantile function of S.

Acerbi (2002) proved that spectral risk measures are the unique class of law-invariant coherent risk measures. Every risk measure used in practice is a special case:

- **VaR at level** α: $\phi(p) = \delta(p - \alpha)$ (point mass)
- **Expected Shortfall at level** α: $\phi(p) = \alpha^{-1}\, \mathbf{1}_{[0,\alpha]}(p)$
- **Exponential spectral**: $\phi(p) = \lambda e^{-\lambda(1-p)}/(1 - e^{-\lambda})$

Acerbi (2002, Section 5) noted: *"The integral is computable only when an explicit analytical expression for the inverse cumulative distribution function is available."* The present theorem shows that for smooth distributions, $N = O(\log(1/\varepsilon)/\log\rho)$ parameters provide such an expression to arbitrary accuracy.

2.2 Fourier-Cosine (COS) Representations

A probability density f on $[a, b]$ admits a Fourier-cosine expansion:

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

with CDF:

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

The $N + 2$ parameters $(A_0, \ldots, A_{N-1}, a, b)$ constitute the spectral representation.

2.3 The Bernstein Ellipse and Coefficient Decay

Definition 2 (Analyticity Radius). For a density f on $[a, b]$ that extends holomorphically to a strip of width $\delta > 0$ around $[a, b]$ in the complex plane, the analyticity radius is

$$\rho = \exp\left(\frac{\pi\delta}{b-a}\right) > 1.$$

Lemma 1 (Bernstein Coefficient Decay). *If f has analyticity radius $\rho > 1$ on $[a, b]$ with $\sup_{\mathcal{E}_\rho} |f(z)| = M_f$, then*

$$|A_k| \le \frac{2\, M_f}{b-a}\, \rho^{-k}.$$

Proof. The argument proceeds in four steps.

Step 1 (Scaling). Map $[a, b]$ to $[-1, 1]$ via $t = 2(x - a)/(b - a) - 1$. The density $g(t) = f(a + (b - a)(t + 1)/2) \cdot (b - a)/2$ inherits the analyticity of f with strip width $\delta' = 2\delta/(b - a)$.

Step 2 (Joukowski map). The map $z = (w + w^{-1})/2$ takes the circle $|w| = \rho$ to the Bernstein ellipse $\mathcal{E}_\rho$ with semi-axes $(\rho + \rho^{-1})/2$ and $(\rho - \rho^{-1})/2$, and foci at ± 1. For $\rho > 1$, the ellipse strictly contains $[-1, 1]$.

Step 3 (Cauchy estimate). The function $h(w) = g((w + w^{-1})/2)$ is holomorphic on the disc $|w| < \rho$. Its Laurent coefficients satisfy

$$|c_k| \leq \frac{1}{2\pi} \oint_{|w|=\rho} \frac{|h(w)|}{|w|^{k+1}} \, |dw| \leq \frac{M_f}{\rho^k},$$

where $M_f = \sup_{|w|=\rho} |h(w)| = \sup_{z \in \mathcal{E}_\rho} |g(z)|$. This step is machine-verified in Mathlib as norm_cauchyPowerSeries_le (Kudryashov, ITP 2022).

Step 4 (Chebyshev-Fourier identification). The k-th Chebyshev polynomial satisfies $T_k(\cos\theta) = \cos(k\theta)$. On the Joukowski image: $T_k(z) = (w^k + w^{-k})/2$ for $z = (w + w^{-1})/2$. Therefore, the Chebyshev expansion coefficient a_k of g equals $c_k + c_{-k}$, and the Fourier-cosine coefficient A_k on $[a, b]$ equals $2a_k/(b - a)$. Combining with Step 3: $|A_k| \leq 2M_f/(b - a) \cdot \rho^{-k}$. □

Remark. Step 3 (the Cauchy estimate) is the only non-elementary component. It is machine-verified in Mathlib. Steps 1, 2, and 4 are elementary algebra and trigonometric identities, partially verified in Lean (JoukowskiMap.lean, BernsteinBridge.lean).

2.4 The Eigen-COS Architecture

The Eigen-COS method (Nagy, 2026a) computes the spectral representation by:

1. **Eigenvalue conditioning**: decompose $C = V \Lambda V^T$ and condition on the top K eigenvectors, making the conditional CF factorizable.
2. **COS inversion**: for each of $Q = n_q^K$ conditioning scenarios, compute N Fourier coefficients from the conditional CF.
3. **Mixture Collapse** (Theorem 1 below): merge the Q conditional representations into a single N-term series.

The key property: step 3 eliminates all dependence on the conditioning structure. The output — N coefficients — contains no trace of K, Q, or the eigenvectors. This is why the representation size is dimension-free.

Remark (Sum density vs. joint density). The Universal Spectral Representation Theorem applies to the **one-dimensional density of the portfolio sum** $S = \sum w_i X_i$, not to the n-dimensional joint density of $(X_1, \ldots, X_n)$. The joint density lives in $\mathbb{R}^n$ and genuinely requires $\Omega(n)$ parameters to describe. The theorem says: for computing **risk measures** (which depend only on the marginal distribution of S), the n-dimensional joint structure is irrelevant after the Eigen-COS reduction. The eigenvalue conditioning "projects" the n-dimensional problem onto a one-dimensional distribution, and it is this one-dimensional object that admits the N-parameter representation.

---

3. The Upper Bound (Constructive)

3.1 The Universal Mixture Collapse

Theorem 1 (Universal Mixture Collapse; Lean-verified). *Let $\{A_{k,q}\}_{k=0}^{N-1}$ be the COS coefficients of Q conditional densities on a shared domain $[a,b]$, with weights $\{w_q\}_{q=1}^{Q}$ satisfying $\sum_q w_q = 1$. Let $\{\phi_k\}_{k=0}^{N-1}$ be any orthogonal basis on $[a,b]$. Then*

$$\sum_{q=1}^{Q} w_q \sum_{k=0}^{N-1} A_{k,q}\,\phi_k(x) = \sum_{k=0}^{N-1} \left(\sum_{q=1}^{Q} w_q A_{k,q} \right) \phi_k(x).$$

The merged coefficients $A_k^* = \sum_q w_q A_{k,q}$ do not depend on the marginal distribution type used to compute $A_{k,q}$.

Proof. Linearity of finite sums. The interchange is valid unconditionally because both sums are finite. □

Remark. The proof uses only the linearity of the Fourier basis — not properties of lognormals, Student-t, or any specific distribution. This makes the Mixture Collapse universal: the Eigen-COS method works for any marginal type whose CF is computable.

3.2 Risk Functional Control

Theorem 2 (Risk Error from Quantile Error; Lean-verified). *Let ϕ be an admissible spectrum and let q_{true}, q_{approx} be two quantile functions with $\sup_p |q_{true}(p) - q_{approx}(p)| \le \delta$. Then*

$$|\rho_\phi(q_{\text{true}}) - \rho_\phi(q_{\text{approx}})| \le \delta.$$

Proof. $|\rho_\phi(q_1) - \rho_\phi(q_2)| = |\sum_i \phi_i (q_1(p_i) - q_2(p_i))| \le \sum_i \phi_i |q_1(p_i) - q_2(p_i)| \le \sum_i \phi_i \cdot \delta = \delta$, using $\sum \phi_i = 1$. □

Corollary (CDF-to-Risk Transfer). *If the CDF approximation satisfies $\sup_x |F(x) - \widehat{F}(x)| \le \varepsilon$, then for any spectral risk measure ρ_ϕ:*

$$|\rho_\phi(F) - \rho_\phi(\widehat{F})| \le \frac{\varepsilon}{\inf_{p \in [\alpha_{\min}, 1]} f_S(F^{-1}(p))},$$

where $\alpha_{\min}$ is the lower support of the spectrum ϕ and f_S is the density. For unimodal densities with mode at m, $f_S(F^{-1}(p)) \ge f_S(a) > 0$ for p bounded away from 0 and 1, so the transfer factor is bounded.

Remark. In practice, the density at the VaR quantile is bounded below by $f_{\min} > 0$ for smooth, unimodal distributions. The transfer factor $1/f_{\min}$ is typically $O(1)$ for standard financial distributions, so the CDF error ε translates to risk error $O(\varepsilon)$ without blowup.

3.3 The Upper Bound Theorem

Theorem 3 (Upper Bound; Lean-verified). *For any portfolio of n assets with joint density having analyticity radius $\rho > 1$, coefficient bound $C_f = 2M_f/(b-a) > 0$, and target accuracy $0 < \varepsilon < C_f$:*

$$N \leq \underbrace{\left\lceil \frac{\log(C_f/(\varepsilon(1-\rho^{-1})))}{\log \rho} \right\rceil}_{=C_1 \cdot \log(1/\varepsilon)/\log \rho + O(1)} + 2$$

parameters suffice to compute every spectral risk measure to accuracy ε. Here $C_1 = 1$ and the $O(1)$ term is $\log(C_f/(1-\rho^{-1}))/\log \rho$, which depends on M_f, $b-a$, and ρ but not on n.

Proof. Compose the chain:

1. **Mixture Collapse** (Theorem 1): the Eigen-COS method produces N merged coefficients A_k^*, regardless of the marginal type and conditioning structure. The number N does not depend on n.
2. **Coefficient Decay** (Lemma 1): $|A_k| \leq C_f \cdot \rho^{-k}$ where $C_f = 2M_f/(b-a)$. The decay rate ρ depends on the density's analyticity, not on n.
3. **Geometric Tail**: the truncation error from using only N terms is

$$\varepsilon_N \leq \frac{C_f \cdot \rho^{-N}}{1-\rho^{-1}}.$$

Setting $\varepsilon_N \leq \varepsilon$ and solving for N gives $N \geq \log(C_f/(\varepsilon(1-\rho^{-1})))/\log \rho$.

4. **Risk Control** (Theorem 2): the CDF error $\varepsilon_N \leq \varepsilon$ implies all spectral risk measures are ε-accurate.

At no step does n appear in the bound on N. The portfolio dimension enters only the computation cost (through K and Q in the Eigen-COS method), not the representation size. □

---

4. The Lower Bound (Information-Theoretic)

4.1 Orthogonal Modes as Independent Dimensions

The Fourier-cosine basis $\{\cos(k\pi(x-a)/(b-a))\}_{k\geq 0}$ on $[a,b]$ is orthogonal. Each coefficient A_k is an independent "dimension" of the density: knowing A_j provides no information about A_k for $j \neq k$.

Definition 3 (Modes Above the Floor). For a density with $|A_k| \leq C_f \cdot \rho^{-k}$, the k-th mode is "above the floor ε" if $C_f \cdot \rho^{-k} > \varepsilon$, i.e., $k < \log(C_f/\varepsilon)/\log \rho$.

Each mode above the floor can take multiple distinguishable values at resolution ε — it carries at least one bit of information.

4.2 The Entropy Lower Bound

Theorem 4 (Entropy Lower Bound; Lean-verified). *Let $N_\varepsilon = \lfloor \log(C_f/\varepsilon)/\log \rho \rfloor$ be the number of modes above the floor. Then:*

(i) All modes $k = 0, 1, \ldots, N_\varepsilon$ satisfy $C_f \cdot \rho^{-k} \geq \varepsilon$.

(ii) Each such mode carries ≥ 1 bit of information (since $C_f \cdot \rho^{-k}/\varepsilon > 1$, there are ≥ 2 distinguishable values).

(iii) The $N_\varepsilon + 1$ modes are independent (orthogonal).

(iv) Therefore, any representation achieving ε-accuracy for all spectral risk measures requires $\geq N_\varepsilon$ parameters.

Proof. We give a detailed proof of each part.

(i) The coefficient bound $C_f \cdot \rho^{-k}$ is monotone decreasing in k because $\rho^{-1} < 1$ (since $\rho > 1$). For any $k \leq N_\varepsilon$: $C_f \cdot \rho^{-k} \geq C_f \cdot \rho^{-N_\varepsilon} > \varepsilon$ by definition of N_ε.

(ii) Consider the k-th mode with bound $B_k = C_f \rho^{-k} > \varepsilon$. The coefficient A_k can take any value in $[-B_k, B_k]$. At resolution ε, there are at least $\lfloor 2B_k/\varepsilon \rfloor \geq 2$ distinguishable values (since $B_k > \varepsilon$ implies $2B_k/\varepsilon > 2$). Each such value corresponds to a different density — the cosine mode $A_k \cos(k\pi(x-a)/(b-a))$ contributes differently to the CDF. Therefore, specifying this mode requires at least $\log_2(2B_k/\varepsilon) \geq 1$ bit.

(iii) The cosine functions $\{\cos(k\pi(x-a)/(b-a))\}_{k \geq 0}$ are orthogonal on $[a, b]$:

$$\int_a^b \cos\left(\frac{j\pi(x-a)}{b-a}\right) \cos\left(\frac{k\pi(x-a)}{b-a}\right) dx = 0 \quad (j \neq k).$$

Consequently, the coefficient A_k is determined by the density's projection onto the k-th mode alone. Changing A_j $(j \neq k)$ does not affect A_k. The $N_\varepsilon + 1$ coefficients $A_0, A_1, \ldots, A_{N_\varepsilon}$ are therefore independent parameters.

(iv) We construct an explicit packing. For each of the $N_\varepsilon + 1$ modes $k = 0, \ldots, N_\varepsilon$, the coefficient A_k ranges over $[-B_k, B_k]$ where $B_k = C_f \rho^{-k} > \varepsilon$. Consider the grid of densities $\{f_{\mathbf{s}}\}_{\mathbf{s} \in \{-1,+1\}^{N_\varepsilon+1}}$ defined by $A_k(\mathbf{s}) = s_k \cdot \varepsilon$ for $k = 0, \ldots, N_\varepsilon$ and $A_k = 0$ for $k > N_\varepsilon$. By orthogonality (part (iii)), for any $\mathbf{s} \neq \mathbf{s}'$, there exists k with $s_k \neq s'_k$, so $|A_k(\mathbf{s}) - A_k(\mathbf{s}')| = 2\varepsilon$. The resulting CDF difference satisfies $\sup_x |F_{\mathbf{s}}(x) - F_{\mathbf{s}'}(x)| \geq 2\varepsilon/(k\pi) > 0$. This gives $2^{N_\varepsilon+1}$ densities that are pairwise ε-separated in sup-CDF norm. Any representation in $\mathbb{R}^N$ with $N < N_\varepsilon$ maps $2^{N_\varepsilon+1}$ points to $\mathbb{R}^N$. By the volumetric argument (the ε-covering number of a bounded subset of $\mathbb{R}^N$ with diameter D is at most $(D/\varepsilon + 1)^N$), when $2^{N_\varepsilon+1} > (D/\varepsilon + 1)^N$ — which holds for $N < N_\varepsilon / \log_2(D/\varepsilon + 1)$ — the pigeonhole principle guarantees two ε-separated densities collide. Since any spectral risk measure ρ_ϕ with spectrum supported on $[0, 1]$ satisfies $|\rho_\phi(f_1) - \rho_\phi(f_2)| \leq \sup_p |q_1(p) - q_2(p)| \leq \sup_x |F_1(x) - F_2(x)|/f_{\min}$, where $f_{\min} > 0$ is the minimum density on the quantile support, the representation cannot achieve ε-accuracy for all spectral risk measures. □

Remark (Relation to Kolmogorov-Tikhomirov). The classical ε-entropy of the class of functions analytic on the Bernstein ellipse $\mathcal{E}_\rho$ with $\sup|f| \leq M$ is $H_\varepsilon(\mathcal{A}(\rho)) = (1 + o(1)) \cdot \log(1/\varepsilon)/\log \rho$ as $\varepsilon \to 0$ (Kolmogorov and Tikhomirov, 1959, Theorem 7; see also Lorentz, 1966, Chapter 10). Our Theorem 4 recovers this rate by a simpler route: counting orthogonal modes above the noise floor, rather than invoking the general metric entropy machinery. The advantage of our approach is that it produces a self-contained Lean-verifiable proof (12 lines of tactics in EntropyLowerBound.lean).

---

5. The Main Theorem

5.1 The Θ-Result

Theorem 5 (Universal Spectral Representation; Lean-verified). *Combining Theorems 3 and 4: for any portfolio of n assets with joint density having analyticity radius $\rho > 1$, the optimal representation size satisfies*

$$N_{\text{optimal}}(\varepsilon, \rho) = \Theta\left(\frac{\log(1/\varepsilon)}{\log \rho}\right).$$

This quantity depends on ε (accuracy) and ρ (smoothness) but not on n (dimension).

Proof. The upper bound (Theorem 3) gives $N \leq C_1 \cdot \log(1/\varepsilon)/\log\rho + 2$. The lower bound (Theorem 4) gives $N \geq C_2 \cdot \log(1/\varepsilon)/\log\rho$. Both bounds are $\Theta(\log(1/\varepsilon)/\log\rho)$ and neither contains n. For $C_2 \leq C_1$ (which holds for the natural constants C_f), the bounds are compatible: $C_2 \cdot L \leq N_{\text{optimal}} \leq C_1 \cdot L + 2$ where $L = \log(1/\varepsilon)/\log\rho$. □

5.2 The Spectral Fenton as First Instance

Corollary 1. *The Spectral Fenton Distribution (Nagy, 2026a) with $N = 128$ is a near-optimal instance of the Universal Representation for lognormal marginals.*

For a 100-asset lognormal portfolio: $n(n+3)/2 = 5{,}150$ input parameters are compressed to $N + 2 = 130$ output parameters — a compression ratio of 39.6× — with all spectral risk measures computable to accuracy $\varepsilon < 10^{-14}$.

5.3 Consequences for Practice

1. **Memory**: $130 \times 8 = 1{,}040$ bytes per portfolio. A desk with 10,000 portfolios stores all distributional information in $ $10 MB.
2. **Speed**: After precomputation ($ $65 ms), each risk query costs $O(N) = O(128)$ operations ($ $0.5 ms). For 100 stress scenarios: 115 ms total vs. 66 seconds for Monte Carlo.
3. **Universality**: The same 130 parameters serve VaR, ES, exponential spectral measures, Wang's distortion measures — every coherent risk measure.
4. **The curse is broken**: For smooth distributions, risk measurement is a $O(\log(1/\varepsilon))$-dimensional problem, regardless of portfolio size.

---

6. Numerical Evidence

6.1 Dimension Independence

Table 1. Representation parameters vs. portfolio size.

n	Input parameters $n(n+3)/2$	SF parameters	Ratio	Note
5	20	130	×6.5 expansion	SF oversized for small n

n	Input parameters $n(n+3)/2$	SF parameters	Ratio	Note
10	65	130	$\times$2.0 expansion	
15	135	130	$\times$1.0 breakeven	Crossover point
20	230	130	1.8$\times$ compression	
50	1,325	130	10.2$\times$ compression	
100	5,150	130	39.6$\times$ compression	
500	125,750	130	967$\times$ compression	
1,000	501,500	130	3,858$\times$ compression	

The SF parameter count (130) is constant. For $n < 15$, the SF representation is larger than the input (no compression); the crossover occurs at $n = 15$ where $n(n+3)/2 = 135 \approx 130$. For $n \geq 20$, compression is genuine and grows quadratically.

Figure 1: Representation size vs portfolio dimension: the SF parameter count (horizontal line at 130) vs the input parameter count $n(n+3)/2$ (parabola). The crossover at $n = 15$ is marked. For $n \geq 20$, genuine compression begins.

Figure 2: The Curse is Broken: the SF representation size (130 parameters) is constant while input parameters grow as $n(n+3)/2$.

6.2 Required N vs Target Accuracy

Table 2. Required N for target accuracy ε, by analyticity radius ρ.

ε	$\rho = 1.05$	$\rho = 1.10$	$\rho = 1.50$	$\rho = 2.00$
10^{-2}	171	81	16	9
10^{-4}	266	130	28	16
10^{-6}	360	178	39	22
10^{-8}	455	226	50	29
10^{-10}	549	275	62	36
10^{-14}	738	371	84	49

Each column is $O(\log(1/\varepsilon))$ — confirming the $\Theta(\log(1/\varepsilon)/\log\rho)$ rate. Smoother distributions (larger ρ) need fewer terms. **No column depends on n.**

Figure 3: Required N vs target accuracy ε for different analyticity radii. All curves are logarithmic in $1/\varepsilon$. The horizontal line $N = 128$ corresponds to the value used in (Nagy, 2026a).

For lognormal portfolios with typical parameters ($\sigma \leq 0.3$, moderate correlation), $\rho \approx 1.3$, giving $N \approx 130$ for $\varepsilon = 10^{-4}$ — consistent with the $N = 128$ used in (Nagy, 2026a).

6.3 Multi-Marginal Verification

Table 3. Monte Carlo VaR(99%) across marginal types ($\sigma = 0.3$, equicorrelation).

n	ρ_{eq}	Lognormal VaR	Student-t_5 VaR	Spread
5	0.0	0.755	0.702	7.1%
5	0.5	0.593	0.488	17.7%
10	0.0	0.832	0.806	3.0%
10	0.5	0.608	0.509	16.4%
50	0.0	0.945	0.948	0.3%
50	0.5	0.621	0.529	14.9%
100	0.0	0.974	0.976	0.3%
100	0.5	0.623	0.531	14.8%

VaR values differ across marginal types (the distributions are genuinely different). But the representation size $N = 128$ works for **all** of them — the marginal type affects the *coefficients*, not the *number of coefficients*.

6.4 Estimated Analyticity Radius

Table 4. Estimated ρ from empirical coefficient decay ($\sigma = 0.3$, $\rho_{\text{eq}} = 0.3$).

Marginal	$n = 5$	$n = 10$	$n = 20$	$n = 50$	$n = 100$
Lognormal	1.31	1.28	1.27	1.32	1.29
Student-t_5	1.00	1.00	1.09	1.00	1.11

The analyticity radius ρ is approximately constant across n for each marginal type — confirming the dimension-free property. Lognormal marginals have $\rho \approx 1.3$ (analytic, fast convergence). Student-t_5 has estimated $\rho \approx 1.0$–1.1, at or near the analyticity boundary. The values $\rho = 1.00$ in Table 4 reflect the empirical decay rate from a finite-N fit; the true analyticity radius for sums of Student-t_5 variables is $\rho > 1$ by the convolution smoothing effect (the sum of independent Student-t variables has a smoother density than any individual marginal), but the excess over 1 is small, requiring $N \gtrsim 500$ terms for $\varepsilon = 10^{-4}$. For such heavy-tailed marginals, the Fejér extension (Section 7.3) provides the appropriate framework: dimension-free representation at polynomial rate $N = O(1/\varepsilon)$, without requiring $\rho > 1$.

Figure 4: Fourier coefficient decay $|A_k|$ vs mode k for portfolios of different sizes ($n = 5, 10, 20, 50, 100$). The decay rate ($\rho \approx 1.3$ for lognormals) is the same for all n, confirming the dimension-free property. All curves reach machine precision (10^{-14}) by $k \approx 120$.

6.5 Worked Example: $n = 5$ vs $n = 100$

To make the dimension-free property concrete:

	$n = 5$ portfolio	$n = 100$ portfolio
Input parameters	$n(n+3)/2 = 20$	$n(n+3)/2 = 5{,}150$
SF parameters	130	130
Compression	0.2× (no compression)	39.6×
Memory	1.0 KB	1.0 KB
VaR query time	$ $0.5 ms	$ $0.5 ms
Precompute time	$ $15 ms	$ $175 ms

Both portfolios use the **same** 130 parameters. The 5-asset and 100-asset portfolios are equally cheap to query — only the precomputation cost differs.

---

7. Discussion

7.1 What Does "Smooth" Mean in Practice?

The analyticity condition ($\rho > 1$) requires the density to extend holomorphically beyond the real axis. Most financial distributions satisfy this:

Distribution	Analytic?	Typical ρ	N for $\varepsilon = 10^{-4}$
Lognormal ($\sigma = 0.3$)	Yes	≈ 1.3	≈ 130
Lognormal ($\sigma = 0.1$)	Yes	≈ 2.0	≈ 50
Normal Inverse Gaussian	Yes	≈ 1.1–1.5	≈ 80–200
Variance-Gamma ($\nu = 5$)	Yes	≈ 1.2	≈ 160
Student-t ($\nu = 10$)	Yes	≈ 1.1	≈ 200
Student-t ($\nu = 3$)	Marginal	≈ 1.01	≈ 2000
Distributions with atoms	No	$\rho = 1$	$\to \infty$

The analyticity radius ρ is primarily determined by the **heaviest tail** in the portfolio. Lognormal marginals ($\sigma \leq 0.3$) are well-behaved ($\rho \approx 1.3$), giving $N \approx 128$ — the value used in (Nagy, 2026a). Student-t with low ν approaches the boundary $\rho \to 1$, where the theorem provides no compression advantage.

7.1.1 How Individual ρ_i Combine

Proposition 6 (Portfolio Analyticity Radius). *If asset i has marginal density with analyticity radius ρ_i on the support of $w_i X_i$, then the portfolio sum $S = \sum w_i X_i$ has analyticity radius*

$$\rho_S \geq \min_i \rho_i.$$

Proof. After eigenvalue conditioning (Section 2.4), the assets are approximately independent. The density of the sum $S = \sum w_i X_i$ of independent random variables is the convolution $f_S = f_{w_1 X_1} * f_{w_2 X_2} * \cdots * f_{w_n X_n}$. The characteristic function of S is the product: $\phi_S(t) = \prod_i \phi_{w_i X_i}(t)$.

For the analyticity: if $f_{w_i X_i}$ extends holomorphically to the strip $\{z \in \mathbb{C} : |\mathrm{Im}(z)| < \delta_i\}$, then the convolution $f_{w_1 X_1} * f_{w_2 X_2}$ extends to $\{z : |\mathrm{Im}(z)| < \min(\delta_1, \delta_2)\}$. This follows from the characteristic function representation: if $\phi_i(t) = \int e^{itx} f_i(x)\, dx$ extends analytically to $|\mathrm{Im}(t)| < \delta_i$ (equivalently, f_i decays as $e^{-\delta_i |x|}$), then the product $\phi_1(t)\phi_2(t)$ extends to $|\mathrm{Im}(t)| < \min(\delta_1, \delta_2)$, since the product of two analytic functions on overlapping strips is analytic on their intersection. The convolution $f_1 * f_2$ has CF $\phi_1 \phi_2$, inheriting the analyticity strip of width $\min(\delta_1, \delta_2)$.

By induction on n: $\delta_S \geq \min_i \delta_i$. Since $\rho_i = \exp(\pi \delta_i / (b_i - a_i))$ and the domain $[a_S, b_S]$ of S satisfies $b_S - a_S \geq \max_i (b_i - a_i)$ (the sum's support is at least as wide), we get $\rho_S = \exp(\pi \delta_S / (b_S - a_S)) \geq \exp(\pi \min_i \delta_i / (b_S - a_S))$. For equal-width domains this gives $\rho_S \geq \min_i \rho_i$; in general, the relationship is $\rho_S \geq \min_i \rho_i^{(b_i - a_i)/(b_S - a_S)}$, which is $\geq \min_i \rho_i$ when $b_i - a_i \leq b_S - a_S$ (always true for a sum). $\square$

Consequence. Adding a single heavy-tailed asset (low ρ_i) to a portfolio reduces the portfolio's analyticity radius to $\min_i \rho_i$. This is the "weakest link" principle: the heaviest tail dominates the convergence rate. A portfolio of 99 lognormals ($\rho \approx 1.3$) plus one Student-t_3 ($\rho \approx 1.01$) has $\rho_S \approx 1.01$, requiring $N \approx 2000$ for $\varepsilon = 10^{-4}$.

7.2 Limitations

1. **Smooth marginals only**: the theorem requires $\rho > 1$. For distributions with jumps or atoms, Monte Carlo remains necessary.
2. **Linear portfolios**: the CF multiplication applies to sums $\sum w_i X_i$. Nonlinear payoffs (options) require payoff-specific COS coefficients.
3. **Known correlation structure**: the Eigen-COS method requires the correlation matrix C. Estimation error in C is not captured by the representation error.
4. **Constants not tight**: the ratio C_1/C_2 between the upper and lower bound constants depends on M_f, $b - a$, and the diameter of the representation range D. For the Spectral Fenton case (lognormal, $\sigma = 0.3$): $C_1 = 1$ and $C_2 \approx 1/\log_2(D/\varepsilon)$, giving $C_1/C_2 \approx \log_2(D/\varepsilon) \approx 50$ for $\varepsilon = 10^{-14}$. The representation is near-optimal in rate ($\Theta(\log(1/\varepsilon))$), with a constant gap of \$ 50×\$ between the proved upper and lower bounds. Tightening the constants is an open problem.

7.3 Beyond Analyticity: The Fejér Extension

The main theorem (Theorem 5) requires $\rho > 1$ (analytic densities). This excludes heavy-tailed distributions near the analyticity boundary (Student-t with low degrees of freedom, distributions with atoms). The Fejér-smoothed variant of the COS expansion extends the dimension-free property to the full regularity spectrum.

Theorem 6 (Fejér Extension; Lean-verified). *For any continuous density on $[a, b]$, the Fejér-smoothed COS expansion with N terms satisfies $\|\sigma_N[f] - f\|_\infty \leq \omega(f, \pi/N)$, where ω is the modulus of continuity. The representation size $N = O(1/\varepsilon)$ is dimension-free.*

The Universal Spectral Representation Theorem thus admits a three-regime generalization:

Marginal regularity	Convergence rate	Parameters for $\varepsilon = 10^{-4}$
Analytic ($\rho > 1$)	$N = \Theta(\log(1/\varepsilon)/\log\rho)$	\$ 130($MainTheorem$)||$Sobolev$(H^s\$, $s > 1/2$)
Continuous only	$N = O(1/\varepsilon)$	\$ \$10,000 (Fejér)

In all three regimes, the representation is **dimension-free** in the portfolio size n. The Fejér weights are machine-verified in Lean 4 (FejerSmoothing/MainTheorem.lean, 10 files, 0 sorry).

7.4 Further Extensions

1. **Gradient-based optimization**: the Fourier coefficients A_k^* are differentiable in w (via the CF chain rule), enabling deterministic mean-risk portfolio optimization. The gradient $\partial \mathrm{VaR}_\alpha / \partial w_i$ is computable from the implicit function theorem applied to the sine series $F(x) = \alpha$, at cost $O(N)$ per component — total $O(nN)$ for the full gradient. This enables mean-CVaR optimization without simulation noise.
2. **Stress testing**: changing the correlation matrix C requires re-running the Eigen-COS pipeline (\$ \$65 ms to \$ \$1 second). However, the eigendecomposition cache is valid for portfolio weight changes (Proposition 5 of Nagy, 2026a: Structure-Scale Separation). A full 50×50 stress grid (correlation shifts $\times$ volatility shifts) is computable in \$ \$2 minutes for a 100-asset portfolio vs. \$ \$5 hours for Monte Carlo.
3. **Spectral risk measure optimization**: the Universal Theorem implies that *choosing* the optimal spectrum ϕ (for a given regulatory or economic objective) is tractable. With the 130-parameter representation, one can evaluate ρ_ϕ for any ϕ in $O(N)$ time, enabling grid search or gradient descent over the spectrum space. This is a new capability: optimizing over the space of coherent risk measures was previously intractable for non-parametric distributions.
4. **Multi-period extension**: for dynamic risk measures (conditional VaR, time-consistent risk), the Universal Theorem applies at each time step. The total representation for a T-period model is $T \cdot N$ parameters — still dimension-free in n, with linear growth in T.
5. **Tighter constants via adaptive basis**: the Theta bound uses the Fourier-cosine basis, which is optimal for smooth functions. For distributions with localized features (e.g., near the VaR quantile), an adaptive basis (wavelets, saddlepoint) might achieve better constants C_1/C_2 while preserving the dimension-free rate.
6. **Sobolev intermediate regime.** For Sobolev-regular densities of order s (H^s, $s > 1/2$), the convergence rate is $N = O(\varepsilon^{-1/(2s-1)})$, interpolating between the analytic and continuous regimes. This is proved in Universal/SobolevDecay.lean.

---

8. Beyond Finance: The Universality of the Spectral Bound

The original formulation (Sections 2–7) proves the $\Theta(\log(1/\varepsilon)/\log\rho)$ bound for portfolio risk measures. Subsequent work (Nagy, 2026c–h) reveals that the same bound governs **any system whose state has a smooth density**. We summarize four domains where the theorem applies without modification.

8.1 Stochastic Differential Equations (The Spectral Generator)

For any Itô diffusion $dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t$, the Fokker–Planck density $p(x,t)$ is analytic when μ and σ are smooth. The spectral generator M — a matrix such that $A(t) = e^{Mt}A(0)$ — is a sufficient statistic for the entire process. All properties (distribution, moments, option prices, Greeks, VaR, ES, stationary distribution, spectral gap, autocorrelation) are computable from M alone. The USRT guarantees that $N = \Theta(\log(1/\varepsilon)/\log\rho)$ modes suffice, with the **same bound** as for static distributions. The extension from vectors (static density) to matrices (generator) to tensors (time-varying or multi-dimensional generator) preserves the dimension-free rate (Nagy, 2026g).

8.2 Celestial Mechanics (The Three-Body Problem)

The Circular Restricted Three-Body Problem (CR3BP) with stochastic perturbation has a density evolving under the Fokker–Planck equation in phase space (x, v). The USRT applies: the orbital uncertainty density requires N modes per dimension regardless of the phase-space dimensionality. For the Earth-Moon system, the spectral generator captures the stationary distribution (where the particle spends time), first passage times (how long to transfer between Lagrange points: $\mathbb{E}[\tau] = -\mathbf{1}^\top M_{\text{killed}}^{-1} A(0)$), and the Routh stability criterion for equilateral points — all from one matrix (Nagy, 2026h). The capstone theorem (14-level Lean gym, graduated) verifies that the spectral completeness holds for the CR3BP.

8.3 Space Debris Conjunction Assessment

NASA's standard collision probability (Gaussian 2D-Pc) underestimates tail probabilities by 2.0–2.4× in LEO when J2 oblateness creates excess kurtosis ($\kappa \approx +1.2$). The USRT guarantees that the non-Gaussian density at the time of closest approach is representable by $N = 64$ spectral coefficients with exponential convergence. The spectral method screens 10,000 conjunctions in 13 seconds — a 77× speedup over Monte Carlo — while providing exact tail probabilities that the Gaussian model misses (Nagy, 2026i).

8.4 Machine Learning (Knowledge Representations)

The eigenspectrum of a trained ML model's prediction function on a data distribution follows the same decay law: the dominant K^* eigenvalues capture the model's learned knowledge, with $K^* = \Theta(\log(n/\sigma^2)/\log\rho)$ where n is the dataset size and σ^2 is the noise. This connects model compression (how small can the model be without losing accuracy?) to the USRT: the answer is K^* spectral coefficients, regardless of the model's parameter count. The spectral information state $\psi_k = (A_k, \sigma_k^2)$ per eigenmode unifies Bayesian and frequentist model selection, with the two frameworks disagreeing only on $O(1/\log\rho) \approx 3$ modes near the truncation boundary (Nagy, 2026j).

8.5 The Unified Statement

The USRT in its general form:

Theorem (Universal Spectral Representation). *For any probability density p on $\mathbb{R}^d$ with analyticity radius $\rho > 1$ per dimension, and any Lipschitz functional Φ of p:*

$$|\Phi[p] - \Phi[p_N]| \le L_\Phi \cdot C \cdot \rho^{-N}$$

where p_N is the N-mode spectral truncation, L_Φ is the Lipschitz constant of Φ, and $N = \Theta(\log(1/\varepsilon)/\log\rho)$ is independent of d, Φ, and the dynamics generating p.

The "Risk" in the original name referred to the first application. The theorem is universal: it applies to any smooth density and any Lipschitz functional — risk measures, option prices, collision probabilities, model predictions, orbital uncertainties.

---

9. Conclusion

The distribution of a smooth system has been treated as a high-dimensional object requiring simulation. We have proved that for smooth distributions, it is low-dimensional: $\Theta(\log(1/\varepsilon)/\log\rho)$ parameters suffice for all spectral risk measures, regardless of portfolio size. The proof combines a constructive upper bound (the generalized Eigen-COS method, backed by Mathlib's machine-verified Cauchy integral formula) with an information-theoretic lower bound (the orthogonal mode counting argument). The bounds match up to constants, establishing the exact rate.

Three aspects of the result merit emphasis. First, the theorem is **constructive**: the Eigen-COS method provides an explicit algorithm that achieves the upper bound, not merely an existence proof. The Spectral Fenton Distribution (Nagy, 2026a) is the first working instance, with $N = 128$ for lognormal portfolios of any size. Second, the theorem is **tight**: the lower bound from Fourier orthogonality matches the upper bound from the Bernstein ellipse, both giving $\Theta(\log(1/\varepsilon)/\log\rho)$. This means the Eigen-COS method is near-optimal — no method can achieve significantly better compression. Third, the theorem is **formally verified**: the structural core (Mixture Collapse, risk control, coefficient decay chain, entropy lower bound, and the Θ-composition) is machine-checked in Lean 4 with 0 unverified axioms and 0 sorry markers.

The practical implication is immediate: for the vast majority of financial distributions used in practice (lognormal, NIG, variance-gamma, moderate Student-t), a portfolio's complete risk profile — VaR, ES, and every coherent risk measure — is encodable in $ $1 KB, regardless of whether the portfolio contains 5 or 5,000 assets. The curse of dimensionality in risk measurement is not fundamental. It is an artifact of simulation-based methods.

Code and Data Availability

The Lean 4 formalization (13 files in LeanProofs/Universal/) is available at https://github.com/tnagy/spectral−fen
Tables 1–4 were generated using the spectral_fenton Python library (https://github.com/tnagy/spectral−fenton);
reproduction scripts are in examples/universal_tables.py.

Acknowledgements

The author used Large Language Models for assistance with Lean 4 formalization and language editing. The author assumes full responsibility for the mathematical content.

Appendix A: Lean 4 Formalization

The following components are machine-verified in Lean 4 (v4.28.0):

Component	File	Status
Universal Mixture Collapse (Thm 1)	GeneralMarginal.lean	Verified
Cauchy Coefficient Estimate	CauchyBound.lean	Verified (Mathlib)
Joukowski Map (disc → ellipse)	JoukowskiMap.lean	Verified
Bernstein Bridge (full chain)	BernsteinBridge.lean	Verified
Coefficient Decay Consequences	CoefficientDecay.lean	Verified
Bernstein Ellipse ($N(\varepsilon)$, dimension-free)	BernsteinEllipse.lean	Verified
Spectral Risk Measure Space (Thm 2)	RiskFunctionalSpace.lean	Verified
Upper Bound (Thm 3)	UpperBound.lean	Verified
Entropy Lower Bound (Thm 4)	EntropyLowerBound.lean	Verified
Kolmogorov Entropy Framework	KolmogorovEntropy.lean	Verified
Capacity Bound + Pigeonhole	CapacityBound.lean	Verified
Lower Bound + Bounds Match	LowerBound.lean	Verified
Main Theorem (Θ-result)	MainTheorem.lean	Verified

0 sorry, 0 unverified axiom dependencies. The Bernstein coefficient bound is backed by Mathlib's norm_cauchyPowerSeries_le + elementary conformal maps. The entropy lower bound is proved from Fourier orthogonality without citing external results. The formalization was verified by lake build (Lean 4.28.0, Mathlib v4.28.0) on March 3, 2026, producing 0 errors across all 85 source files in the LeanProofs/ directory (13 in Universal/, 62 in SpectralFenton/).

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

- Acerbi, Carlo (2002). Spectral Measures of Risk: A Coherent Representation of Subjective Risk Aversion. *Journal of Banking & Finance*, 26(7), 1505-1518. DOI: 10.1016/S0378-4266(02)00281-9

- Acworth, P., Broadie, M., and Glasserman, P (1997). A comparison of some Monte Carlo and quasi-Monte Carlo techniques for option pricing. In *Monte Carlo and Quasi-Monte Carlo Methods 1996*, Springer, 1–18. *Acworth, P., Broadie, M., and Glasserman, P.*.
- Carr, Peter and Madan, Dilip (1999). Option Valuation Using the Fast Fourier. *Journal of Computational Finance*, 2(4), 61-73. DOI: 10.21314/jcf.1999.043
- Degenne, R. et al (2025). Formalization of Brownian motion in Lean. arXiv:2511.20118. *Degenne, R. et al.*.
- Echenim, M., Guiol, H., and Peltier, N (2020). Formalizing the Cox-Ross-Rubinstein pricing of European derivatives in Isabelle/HOL. *Echenim, M., Guiol, H., and Peltier, N.*.
- 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
- Giles, M.B (2008). Multilevel Monte Carlo path simulation. *Giles, M.B.*, 56(3).
- Giles, M.B (2015). Multilevel Monte Carlo methods. *Giles, M.B.*.
- Glasserman, Paul (2003). Monte Carlo Methods in Financial Engineering. Springer.
- Kolmogorov, Andrey N. and Tikhomirov, Vladimir M. (1959). ε-Entropy and ε-Capacity of Sets in Function Spaces. *Uspekhi Matematicheskikh Nauk*, 14(2), 3-86. DOI: 10.1007/978-94-017-2973-4_7
- Kudryashov, Y (2022). Formalizing the divergence theorem and the Cauchy integral formula in Lean. *Kudryashov, Y.*.
- Lorentz, George G (1966). Approximation of Functions. Holt, Rinehart and Winston.
- Nagy, T. (2026). The Exact Latent Distribution of Correlated Lognormal Sums. *Working paper*.
- Nagy, T. (2026). Exact Portfolio VaR Without Monte Carlo: The Eigen-COS Method. *Zenodo*. DOI: 10.5281/zenodo.18910516
- Niederreiter, Harald (1992). Random Number Generation and Quasi-Monte Carlo. SIAM. DOI: 10.1137/1.9781611970081
- Pinkus, Allan (1985). n-Widths in Approximation Theory. Springer. DOI: 10.1007/978-3-642-69894-1
- Ruijter, M. J., & Oosterlee, C. W (2012). Two-dimensional Fourier cosine series expansion method for pricing financial options. *SIAM Journal on Scientific Computing*, 34(5). DOI: 10.1137/120862053
- Shannon, Claude E. (1959). Coding Theorems for a Discrete Source with a Fidelity Criterion. *IRE National Convention Record*, 7(4), 142-163.
- Trefethen, Lloyd N. (2013). Approximation Theory and Approximation Practice. SIAM. DOI: 10.1137/1.9781611975949