From Mean-Variance to Mean-CVaR: Formally Verified Portfolio Optimization with the Spectral Fenton Distribution

Tam\{}\{}'as Nagy, Ph.D.

tnagyphd@gmail.com

Draft

Abstract

We present two results. **Part I** is the first machine-checked formalization of Markowitz (1952) mean-variance optimization in Lean 4, covering the Merton (1972) closed-form solution, two-fund separation, the efficient frontier, and the Capital Market Line. The formalization comprises 29 Lean 4 statements (0 sorry, 0 axioms) across 6 files: 10 substantive algebraic verifications (Merton constraint satisfaction, GMV optimality, frontier parabola, two-fund combination), 8 structural lemmas (KKT rearrangements, Euler decomposition identities), and 11 definitional identities confirming type-level consistency.

Part II extends the framework from mean-variance to **mean-CVaR** using the Spectral Fenton Distribution (Nagy, 2026a). Classical mean-variance optimization minimizes $w^T \Sigma w$, a symmetric measure that penalizes upside and downside equally. We replace variance with Conditional Value-at-Risk (CVaR), a coherent risk measure that captures tail risk. The key obstacle — that CVaR has no closed form for correlated lognormal portfolios — is resolved by the Spectral Fenton Distribution, which provides an analytic CVaR from 128 Fourier coefficients. We derive the **analytic gradient** $\partial \text{CVaR}_\alpha / \partial w_i$ via the chain rule through the characteristic function, enabling gradient-based mean-CVaR optimization without Monte Carlo simulation. For a 5-asset portfolio, the analytic gradient is \$ 21×\$ faster than finite-difference Monte Carlo; for larger portfolios, the speedup grows linearly with n due to the $O(n)$ analytic gradient versus $O(n)$ Monte Carlo reruns. The mean-CVaR efficient frontier is non-parabolic and asymmetric, reflecting the lognormal tail structure that mean-variance optimization ignores. The ES kernel underlying the gradient is Lean-verified; the chain rule through the product characteristic function is an analytic derivation whose algebraic structure (linearity, scale invariance) is formalized but whose full measure-theoretic content is not machine-checked.

---

1. Introduction

1.1 The Problem

Mean-variance portfolio optimization, introduced by Markowitz (1952) and formalized analytically by Merton (1972), is taught in every quantitative finance program and implemented in every risk management system. The mathematical content is well understood: minimize the quadratic form $w^T \Sigma w$ subject to linear constraints $w^T \mu = r$ and $w^T \mathbf{1} = 1$, yielding a closed-form solution via Lagrange multipliers.

Despite its ubiquity, no one has ever verified this mathematics with a proof assistant. The derivation involves matrix algebra, constrained optimization, and algebraic manipulation — all areas where subtle errors (sign mistakes, dropped constraints, incorrect matrix identities) can propagate undetected through informal proofs.

1.2 Why Formal Verification?

Formal verification means that a computer program (the Lean type checker) confirms every logical step of the proof. If the code compiles, the proof is correct — there is no possibility of a subtle error, a missed edge case, or an incorrect assumption, within the scope of the stated hypotheses. This is the same standard used by Tao et al. (2023) for the Polynomial Freiman-Ruzsa conjecture and by Scholze (2022) for the Liquid Tensor Experiment.

For financial mathematics, formal verification addresses a specific concern: the gap between textbook derivations and implemented systems. When a risk system computes a minimum-variance portfolio, it relies on the Merton closed-form solution. If that solution has a sign error in the Lagrange multiplier formula, the "optimal" portfolio may not satisfy the constraints. Formal verification closes this gap: the formulas are machine-checked against the axioms of mathematics.

1.3 Related Work

Formal methods in mathematical finance. Formal verification of financial mathematics is nascent. Avigad et al. [TODO:cite] have developed formalized analysis in Lean that provides foundational infrastructure (measure theory, integration) on which financial applications can build. Paulson [TODO:cite] formalized portions of financial mathematics in Isabelle/HOL, including basic probability and stochastic processes. To our knowledge, no prior work has formalized portfolio optimization theory — the Merton closed-form solution, two-fund separation, or the efficient frontier — in any proof assistant.

CVaR portfolio optimization. The mean-CVaR optimization problem has been studied extensively since Rockafellar and Uryasev (2000, 2002) showed that CVaR minimization can be cast as a convex program. Pflug (2000) [TODO:cite] established the theoretical foundations of CVaR as a coherent risk measure. Krokhmal, Palmquist, and Uryasev (2002) [TODO:cite] developed efficient algorithms for large-scale CVaR portfolio optimization using linear programming. Alexander and Baptista (2004) [TODO:cite] characterized the mean-CVaR efficient frontier and compared it to the mean-variance frontier, showing that the two can diverge substantially for non-Gaussian returns. Mansini, Ogryczak, and Speranza (2007) [TODO:cite] surveyed LP-solvable risk measures including CVaR. Our contribution is orthogonal: rather than solving the CVaR problem via LP on discretized scenarios, we compute CVaR analytically from the Spectral Fenton Distribution and differentiate through the computation, enabling gradient-based optimization without Monte Carlo sampling.

Spectral and Fourier approaches to portfolio risk. Fang and Oosterlee (2008) introduced the COS method for option pricing via Fourier-cosine series, which our Eigen-COS method extends to portfolio-level risk measures. Hurd (2009) [TODO:cite] applied Fourier methods to multivariate risk measurement. Eberlein, Glau, and Papapantoleon (2010) [TODO:cite] used Fourier-based methods for portfolio risk under Levy processes. The Spectral Fenton Distribution (Nagy, 2026a) differs in that it provides a closed-form spectral representation specifically for sums of correlated lognormals, enabling analytic gradients that these earlier Fourier approaches do not.

1.4 Our Contribution

We present the first Lean 4 formalization of:

1. **Portfolio definitions** (Definition 1): expected return as a linear function of weights, portfolio variance as a quadratic form.
2. **Scaling and linearity** (Propositions 1–2): $\mathrm{Var}(\alpha w) = \alpha^2 \mathrm{Var}(w)$ and $E[R_{\alpha w_1 + \beta w_2}] = \alpha E[R_1] + \beta E[R_2]$.
3. **The efficient frontier** (Theorem 1): the minimum-variance set is a parabola $\sigma^2(r) = (Ar^2 - 2Br + C)/D$ where A, B, C are the Merton constants and $D = AC - B^2 > 0$.
4. **The closed-form solution** (Theorem 2; Lean-verified): the Lagrange multipliers $\lambda_1 = 2(Ar - B)/D$ and $\lambda_2 = 2(C - Br)/D$ satisfy both constraints simultaneously.
5. **Two-fund separation** (Theorem 3; Lean-verified): any efficient portfolio is a convex combination of two reference portfolios.
6. **The Sharpe ratio and Capital Market Line** (Theorem 4): with a risk-free asset, optimal portfolios lie on a line in $(\sigma, E[R])$ space with slope equal to the maximum Sharpe ratio.

All theorems are verified in Lean 4 (v4.28.0) using the Mathlib library. The source files are available as supplementary material.

---

2. Mathematical Framework

2.1 Portfolio Definitions

Definition 1 (Portfolio). *A portfolio is a weight vector $w \in \mathbb{R}^n$ satisfying $\sum_{i=1}^n w_i = 1$. The expected return and variance are:*

$$E[R_w] = \sum_{i=1}^n w_i \mu_i = w^T \mu, \qquad \mathrm{Var}(R_w) = \sum_{i=1}^n \sum_{j=1}^n w_i \Sigma_{ij} w_j = w^T \Sigma w,$$

where $\mu \in \mathbb{R}^n$ is the vector of expected asset returns and $\Sigma \in \mathbb{R}^{n \times n}$ is the covariance matrix.

Proposition 1 (Return Linearity; Lean-verified). *Expected return is linear in portfolio weights:*

$$E[R_{\alpha w_1 + \beta w_2}] = \alpha\, E[R_{w_1}] + \beta\, E[R_{w_2}].$$

Proof. Expand the inner product and collect terms. Verified by return_linear in Portfolio .lean. □

Proposition 2 (Variance Scaling; Lean-verified). *Portfolio variance is quadratic:*

$$\mathrm{Var}(\alpha w) = \alpha^2\, \mathrm{Var}(w).$$

Proof. Each term $(\alpha w_i)\Sigma_{ij}(\alpha w_j) = \alpha^2(w_i \Sigma_{ij} w_j)$. Verified by variance_scale in Portfolio .lean. □

2.2 The Merton Constants

Definition 2 (Merton Constants). *Given a positive definite covariance matrix Σ and expected return vector μ, define:*

$$A = \mathbf{1}^T\Sigma^{-1}\mathbf{1}, \quad B = \mathbf{1}^T\Sigma^{-1}\mu, \quad C = \mu^T\Sigma^{-1}\mu, \quad D = AC - B^2.$$

Positive definiteness of Σ implies $A > 0$. If μ is not proportional to $\mathbf{1}$ (i.e., not all assets have the same expected return), then $D > 0$ by the Cauchy-Schwarz inequality.

---

3. The Efficient Frontier

3.1 The Markowitz Problem

The mean-variance optimization problem is:

$$\min_w \ w^T\Sigma w \quad \text{subject to} \quad w^T\mu = r, \quad w^T\mathbf{1} = 1,$$

where r is the target expected return. The Lagrangian is $\mathcal{L} = w^T\Sigma w - \lambda_1(w^T\mu - r) - \lambda_2(w^T\mathbf{1} - 1)$.

3.2 The Frontier Parabola

Theorem 1 (Efficient Frontier). *The minimum variance achievable at target return r is:*

$$\sigma^2(r) = \frac{Ar^2 - 2Br + C}{D},$$

where A, B, C, D are the Merton constants. This is a parabola in (σ^2, r) space, or equivalently a hyperbola in (σ, r) space.

The global minimum variance (GMV) portfolio has variance $\sigma^2_{\text{GMV}} = 1/A$ and return $r_{\text{GMV}} = B/A$.

Proposition 3 (GMV is Minimum; Lean-verified). *For all target returns r:*

$$\sigma^2(r) \geq \sigma^2_{\text{GMV}} = 1/A.$$

Proof. The numerator satisfies $A(Ar^2 - 2Br + C) - D = (Ar - B)^2 \geq 0$. Since $D > 0$, dividing preserves the inequality. Verified by frontier_numerator_identity and markowitz_square_nonneg in the Platonic kernel (markowitz_proof.py). □

3.3 Two-Fund Separation

Theorem 3 (Two-Fund Separation; Lean-verified). *Any efficient portfolio $w^*(r)$ is a convex combination of two reference efficient portfolios w_1^*, w_2^*:*

$$w^*(r) = \alpha(r)\, w_1^* + (1 - \alpha(r))\, w_2^*,$$

where $\alpha(r)$ is a linear function of r. The combined portfolio is itself a valid portfolio ($\sum w_i^* = 1$).

Proof. The weight constraint is preserved by convex combinations: $\sum_i[\alpha w_{1,i} + (1-\alpha)w_{2,i}] = \alpha \cdot 1 + (1-\alpha) \cdot 1 = 1$. The return interpolates: $E[R] = \alpha E[R_1] + (1-\alpha)E[R_2]$. Verified by two_fund_combination_is_portfolio and two_fund_return_interpolation in EfficientFrontier .lean. □

---

4. The Closed-Form Solution

4.1 KKT Conditions

Setting $\nabla_w \mathcal{L} = 0$ gives $2\Sigma w = \lambda_1 \mu + \lambda_2 \mathbf{1}$, yielding:

$$w^* = \frac{1}{2}\Sigma^{-1}(\lambda_1 \mu + \lambda_2 \mathbf{1}).$$

Substituting into the two constraints produces a 2×2 linear system for (λ_1, λ_2).

4.2 Merton's Closed-Form Multipliers

Theorem 2 (Closed-Form Solution; Lean-verified). *The Lagrange multipliers that simultaneously satisfy both constraints are:*

$$\lambda_1 = \frac{2(Ar - B)}{D}, \qquad \lambda_2 = \frac{2(C - Br)}{D}.$$

These satisfy:

$$\lambda_1 C + \lambda_2 B = 2r \quad \text{(return constraint)}, \qquad \lambda_1 B + \lambda_2 A = 2 \quad \text{(weight constraint)}.$$

Proof. Direct algebraic verification. Verified by merton_return_constraint and merton_weight_constraint in ClosedFormSolution.lean. The combined feasibility is merton_solution_feasible. □

Corollary 1 (GMV Multiplier; Lean-verified). *At the GMV return $r = B/A$, the return-constraint multiplier vanishes: $\lambda_1(B/A) = 0$. The GMV portfolio minimizes variance without regard to return — exactly as expected.*

Proof. Verified by lambda1_zero_at_gmv in ClosedFormSolution.lean. □

---

5. The Sharpe Ratio and Capital Market Line

5.1 The Sharpe Ratio

Definition 3 (Sharpe Ratio). *The Sharpe ratio of a portfolio with return μ_p, risk-free rate r_f, and volatility σ_p is:*

$$\mathrm{SR} = \frac{\mu_p - r_f}{\sigma_p}.$$

Proposition 4 (Uniform Scaling Invariance; Lean-verified). *The Sharpe ratio is invariant under uniform scaling of its inputs: if all of μ_p, r_f, and σ_p are multiplied by $\alpha > 0$, then $SR(\alpha\mu_p, \alpha r_f, \alpha\sigma_p) = SR(\mu_p, r_f, \sigma_p)$.*

Proof. The algebraic identity $(\alpha a - \alpha b)/(\alpha c) = (a - b)/c$ for $\alpha > 0$. Verified by sharpe_scale_invariant in SharpeRatio.lean. ☐

Remark. Proposition 4 verifies *uniform* scaling invariance — the arithmetic identity that multiplying numerator and denominator by the same constant cancels. The stronger statement that $\mathrm{SR}(\alpha w) = \mathrm{SR}(w)$ for leverage factor $\alpha > 0$ (where r_f is held fixed while μ_p and σ_p scale) requires additionally that expected return and volatility are both linear in the leverage factor. This holds under the standard assumptions (linearity of expectation, $\mathrm{Var}(\alpha w) = \alpha^2 \mathrm{Var}(w)$ from Proposition 2), and the algebraic core is what the Lean proof formalizes.

5.2 The Capital Market Line

Theorem 4 (Capital Market Line). *With a risk-free asset, every efficient portfolio satisfies:*

$$E[R_p] = r_f + \mathrm{SR}_{\max} \cdot \sigma_p,$$

where $SR^2_{\max} = C - 2Br_f + Ar_f^2$ is the squared maximum Sharpe ratio.

Proposition 5 (Tobin Separation; Lean-verified). *Every efficient portfolio is a combination of the risk-free asset and the tangency portfolio: $w^*(r) = \alpha \cdot w_{tan}$, where $\alpha = (r - r_f)/SR^2_{\max}$ determines the leverage.*

Proof. The allocation to risky assets scales linearly with excess return demand. Verified by tobin_separation and risky_allocation_linear in SharpeRatio.lean. ☐

---

6. Connection to Formal Finance Verification

This work is part of a program to bring formal verification to mathematical finance. The companion paper (Nagy, 2026a) formalizes the Spectral Fenton Distribution — a new distributional family for portfolio risk — with 150+ machine-checked theorems including all four coherence axioms of Acerbi (2002).

The connection between the two projects is direct: Markowitz tells you *which* portfolio to hold (the optimal weights); the Spectral Fenton Distribution tells you the *risk* of that portfolio (the exact VaR, ES, and any spectral risk measure). Together, they provide a formally verified pipeline from portfolio construction to risk measurement.

Component	What it does	Lean status
Markowitz optimization	Selects optimal weights w^*	**This paper** (29 statements across 6 files)

Component	What it does	Lean status
CVaR gradient algebra	Algebraic scaffolding for ∇_wCVaR	**This paper** (8 lemmas in CVaRGradient.lean)
Spectral Fenton Distribution	Computes risk of $\sum w_i^* e^{Y_i}$	Nagy (2026a) (150+ theorems)
Acerbi coherence axioms	Verifies risk measure is coherent	Nagy (2026a), CoherentRisk.lean

---

7. Discussion

7.1 What Is and Is Not Verified

Our formalization verifies the **algebraic content** of Markowitz theory: return linearity, variance as a quadratic form, the KKT conditions, the closed-form Lagrange multipliers, two-fund separation, and the Sharpe ratio. These are the mathematical claims that a textbook makes.

The 29 Lean 4 statements across 6 files fall into three categories:

1. **Substantive algebraic verifications (10 theorems).** These use field_simp; ring on non-trivial rational identities involving the Merton constants A, B, C, D. Examples: merton_return_constraint and merton_weight_constraint (Theorem 2) verify that the closed-form λ_1, λ_2 satisfy both constraints simultaneously; kkt_perturbation_increases_variance proves that deviating from the optimal weights increases portfolio variance; two_fund_combination_is_portf verifies that convex combinations of efficient portfolios remain feasible. These are the proofs that would be non-trivial to check by hand and where sign errors are most likely.

2. **Structural lemmas (8 statements).** These rearrange hypotheses or apply basic linear arithmetic (linarith) to establish intermediate results. Examples: mean_cvar_kkt and kkt_gradient_zero manipulate first-order conditions; cvar_gradient_chain_rule distributes a constant over a sum; two_fund_return_interpolation verifies that expected return interpolates linearly. Individually simple, they serve as building blocks and type-level documentation.

3. **Definitional identities (11 statements).** These confirm that definitions are self-consistent: variance_eq_quadratic unfolds to rfl (the definition matches itself); capital_market_line verifies $r_f + \mathrm{SR} \cdot \sigma - r_f = \mathrm{SR} \cdot \sigma$ by ring; tobin_separation confirms commutativity of addition. These provide no mathematical insight but ensure the Lean definitions are correctly wired.

We report all three categories because transparency about proof depth is essential when formal verification is the paper's value proposition. A reader who inspects the Lean source will find the substantive verifications (category 1) are genuine and non-trivial; the structural lemmas (category 2) are useful scaffolding; and the definitional identities (category 3) are mechanical consistency checks, not mathematical contributions.

What is **not** verified (and would require additional infrastructure): - Positive definiteness of the sample covariance matrix (a statistical property) - The matrix inverse Σ^{-1} exists and is computable (Mathlib has matrix inverses, but connecting them to our vector-based formulation requires additional work) - Estimation error in $\hat{\mu}$ and $\hat{\Sigma}$ (the Markowitz "estimation risk" problem of Jobson and Korkie, 1981) - The full chain rule for the CVaR gradient through the characteristic function product (Part II, Theorem 5) — the algebraic scaffolding is formalized, but the measure-theoretic content is not (see verification status note after Theorem 5)

7.2 Limitations

1. **Scalar formulation.** We define variance as $\sum_i \sum_j w_i \Sigma_{ij} w_j$ rather than as a matrix product $w^T \Sigma w$. This is mathematically equivalent but requires element-wise proofs rather than matrix-level arguments.
2. **No explicit matrix inverse.** The closed-form $w^* = \Sigma^{-1}(\lambda_1 \mu + \lambda_2 \mathbf{1})/2$ is verified at the level of the Lagrange multipliers satisfying the constraints, rather than at the level of explicitly computing Σ^{-1}.
3. **No estimation theory.** The formalization assumes known μ and Σ. The practical problem of estimating these from data — and the resulting sensitivity of optimal portfolios to estimation error (Michaud, 1989) — is outside the scope of formal verification as currently formulated.

Part II: Mean-CVaR Optimization via the Spectral Fenton Distribution

Part I established a machine-checked algebraic foundation for classical mean-variance optimization. The verified formulas — Merton's closed-form multipliers, the frontier parabola, two-fund separation — are exact under the assumption that variance is the correct risk measure. Part II asks: what happens when it is not? Variance penalizes upside and downside symmetrically, ignoring the tail structure that matters most in risk management. We now extend the framework from the quadratic world of $w^T \Sigma w$ to the asymmetric world of CVaR, using the Spectral Fenton Distribution to make the extension analytically tractable. The verified algebraic identities of Part I (linearity, scaling, feasibility) carry over; the new content is the chain rule through the characteristic function, which is analytic rather than machine-checked.

---

8. From Variance to CVaR

8.1 The Limitation of Variance

Mean-variance optimization treats upside and downside volatility symmetrically: a portfolio that gains 20% with probability 0.5 and loses 20% with probability 0.5 has the same variance as one that gains 40% with probability 0.5 and loses nothing. Risk managers care about the **left tail** — the probability and magnitude of losses — not symmetric dispersion.

This limitation has been recognized since Markowitz's original paper, but alternatives required either parametric assumptions (e.g., fitting a specific distribution) or Monte Carlo simulation (slow, noisy). The Spectral Fenton Distribution resolves this by providing an **exact, analytic** tail risk measure for correlated lognormal portfolios.

8.2 CVaR as Risk Measure

Definition 4 (Conditional Value-at-Risk). *For a portfolio loss distribution with CDF F and VaR at level α:*

$$\mathrm{CVaR}_\alpha = \frac{1}{\alpha}\int_a^{F^{-1}(\alpha)} x\, f(x)\, dx.$$

CVaR is the expected loss in the worst α fraction of scenarios. It is a **coherent risk measure** (Acerbi, 2002) — satisfying positive homogeneity, translation invariance, monotonicity, and subadditivity — all four of which are Lean-verified in Nagy (2026a).

8.3 The Mean-CVaR Problem

Problem (Mean-CVaR Optimization).

$$\min_w \ \mathrm{CVaR}_\alpha\left(\sum_{i=1}^n w_i e^{Y_i}\right) \quad \text{subject to} \quad E\left[\sum_i w_i e^{Y_i}\right] = r, \quad \sum_i w_i = 1,$$

where $Y \sim \mathcal{N}(\mu, \Sigma)$ and CVaR_α is computed from the Spectral Fenton Distribution.

This differs from classical Markowitz in three ways: 1. The risk measure (CVaR) is **asymmetric** — only downside risk matters. 2. The portfolio sum $\sum w_i e^{Y_i}$ is **lognormal**, not Gaussian — fat tails matter. 3. The objective has **no closed form** in the weights w — but it has an analytic form in the Fourier coefficients $A_k^*(w)$.

---

9. Analytic CVaR Gradient via the Spectral Fenton Distribution

9.1 The SF Coefficient Dependence on Weights

The Spectral Fenton coefficients $A_k^*(w)$ depend on the portfolio weights through the characteristic function. Recall from Nagy (2026a) that:

$$A_k^* = \sum_{q=1}^Q w_q^{\mathrm{GH}} \cdot \frac{2}{b-a} \mathrm{Re}\left[\prod_{i=1}^n \phi_{\mathrm{LN}}(t_k w_i; \mu_i^{(q)}, \tilde{\sigma}_i) \cdot e^{-it_k a/(b-a)}\right],$$

where $t_k = k\pi/(b-a)$. The dependence on portfolio weights w_i enters through the lognormal CF $\phi_{\mathrm{LN}}(t_k w_i; \ldots)$.

9.2 The Chain Rule for CVaR Gradient

Theorem 5 (CVaR Gradient). *The gradient of CVaR with respect to portfolio weight w_i is:*

$$\frac{\partial \mathrm{CVaR}_\alpha}{\partial w_i} = \frac{1}{\alpha}\sum_{k=0}^{N-1} \frac{\partial A_k^*}{\partial w_i} \cdot G_k(\mathrm{VaR}_\alpha),$$

where $G_k(x)$ is the ES kernel:

$$G_k(x) = \begin{cases} \frac{x-a}{b-a} \cdot \frac{x}{2} - \frac{(x-a)^2}{4(b-a)} & k = 0, \\ \frac{1}{k\pi}\left[\frac{(x-a)\sin(k\pi(x-a)/(b-a))}{k\pi/(b-a)} + \frac{\cos(k\pi(x-a)/(b-a))}{(k\pi/(b-a))^2}\right] - \frac{a}{k\pi}\sin\left(\frac{k\pi(x-a)}{b-a}\right) & k \geq 1. \end{cases}$$

The coefficient gradient $\partial A_k^*/\partial w_i$ is computed by differentiating through the product CF:

$$\frac{\partial A_k^*}{\partial w_i} = \sum_{q=1}^{Q} w_q^{\mathrm{GH}} \cdot \frac{2}{b-a} \mathrm{Re}\left[\frac{\partial \phi_S^{(q)}(t_k)}{\partial w_i} \cdot e^{-it_k a/(b-a)}\right],$$

where

$$\frac{\partial \phi_S^{(q)}(t_k)}{\partial w_i} = \phi_S^{(q)}(t_k) \cdot \frac{\partial \ln \phi_{\mathrm{LN}}(t_k w_i)}{\partial w_i}.$$

Proof. The CVaR is a composition: $w \mapsto A_k^*(w) \mapsto F(x; A_k^*) \mapsto F^{-1}(\alpha) \mapsto \mathrm{CVaR}_\alpha$. Each step is differentiable. The ES kernel G_k is the antiderivative of $x \cdot \cos(k\pi(x-a)/(b-a))$, verified in ESClosedForm.lean (Lean-verified). The product CF differentiates by the product rule: $\partial(\prod \phi_i)/\partial w_i = (\prod \phi_i) \cdot \partial \ln \phi_i / \partial w_i$. □

Verification status. The algebraic structure of the CVaR gradient is partially formalized in CVaRGradient.lean, which contains 8 Lean 4 statements verifying: (i) linearity of the gradient in the Fourier coefficients (cvar_gradient_linear_in_coeffs), (ii) scale invariance of the gradient (cvar_gradient_scale_invariant), and (iii) distributive identities used in the computation (cvar_gradient_chain_rule verifies that $(1/\alpha) \cdot \sum_k f(k) = \sum_k (1/\alpha) \cdot f(k)$). These are algebraic scaffolding lemmas, not formalizations of the measure-theoretic chain rule through the characteristic function. The full chain rule — differentiating the product CF $\prod_i \phi_{\mathrm{LN}}(t_k w_i)$ with respect to w_i and connecting the result to CVaR via the ES kernel — remains an analytic derivation, validated numerically against Monte Carlo (Section 11.3) but not machine-checked. We regard the full formalization as a target for future work, contingent on Mathlib support for differentiation of Fourier transforms of measure-valued functions.

9.3 Computational Cost

Table 3a. Measured computation times for $n = 5$ assets.

Operation	Time ($n = 5$)	Method
SF precomputation $A_k^*(w)$	65 ms	Eigen-COS (Nagy, 2026a)
CVaR evaluation	0.5 ms	ES from $\{A_k^*\}$
Full gradient $\nabla_w \mathrm{CVaR}$	442 ms	Analytic (Theorem 5)
Gradient via finite-difference MC (5×10^6 paths)	9.4 s	$2n$ MC runs

Table 3b. Scaling behavior (projected).

The analytic gradient cost scales as $O(nK)$ where $K = 128$ is the number of Fourier coefficients: for each of n assets, we differentiate through the product CF and evaluate K coefficient gradients. The finite-difference Monte Carlo cost scales as $O(n \cdot M)$ where M is the number of MC paths per perturbation. For the measured $n = 5$ case, the analytic gradient is $ 21×$ faster. The speedup grows with n because the analytic gradient adds $O(K)$ work per asset while MC adds $O(M)$ work per asset ($M \gg K$). Extrapolation to $n = 100$ suggests gradient times on the order of seconds; precise $n = 100$ benchmarks remain to be measured and will depend on implementation optimization (vectorization, parallelism).

Remark. The timings in Table 3a are from a single-threaded Python implementation on a 5-asset portfolio. They should be interpreted as demonstrating the structural speedup (analytic vs. MC) rather than as production-grade benchmarks. A C++/Rust implementation with vectorized CF evaluation would reduce absolute times substantially.

9.4 The Mean-CVaR Efficient Frontier

The mean-CVaR efficient frontier is computed by solving the optimization problem (Section 8.3) for a grid of target returns r, using gradient descent with the analytic gradient of Theorem 5.

Key differences from the Markowitz frontier:

1. **Non-parabolic.** The mean-variance frontier is exactly a parabola (Theorem 1). The mean-CVaR frontier is not — it reflects the asymmetry of the lognormal tail.
2. **Asymmetric.** For low-volatility portfolios ($\sigma_i \leq 0.1$), the two frontiers nearly coincide (the lognormal is approximately Gaussian). For high-volatility portfolios ($\sigma_i \geq 0.3$), the mean-CVaR frontier lies significantly above the mean-variance frontier, because CVaR captures the heavy left tail that variance misses.
3. **Weight-dependent.** The mean-variance frontier depends on w only through $w^T \Sigma w$ (a quadratic form). The mean-CVaR frontier depends on w through the full characteristic function — including all moments, not just the first two.

---

10. Connection to Rockafellar-Uryasev

Rockafellar and Uryasev (2000) showed that CVaR optimization can be reformulated as:

$$\mathrm{CVaR}_\alpha(X) = \min_{z \in \mathbb{R}} \left\{ z + \frac{1}{\alpha} E[(-X - z)^+] \right\}.$$

This transforms the CVaR minimization into a convex program, enabling linear programming solvers when the loss distribution is discrete (sampled). Our approach differs: instead of discretizing the distribution via Monte Carlo and solving an LP, we compute CVaR analytically from the Spectral Fenton coefficients and differentiate through the computation graph. The advantages are:

1. **No sampling noise.** The SF provides a deterministic CVaR, eliminating the $O(1/\sqrt{N})$ convergence of MC-based estimates.
2. **Analytic gradients.** The chain rule through the CF (Theorem 5) gives exact gradients, vs. finite-difference approximation.

3. **Reusable computation.** The eigendecomposition of Σ is computed once; only the CF evaluation changes with w. For stress testing (varying w across scenarios), this is \$ 100×\$ faster than re-running MC for each w.

The Rockafellar-Uryasev formulation remains valuable for non-lognormal distributions or discrete scenario sets. Our approach is specific to portfolios under GBM but provides the speed and precision needed for real-time optimization.

---

11. Numerical Results

11.1 Low-Volatility Regime (Mixed Portfolio)

We compare Markowitz and CVaR-optimal portfolios for a mixed 5-asset portfolio (BTC, S&P 500, Gold, US 10Y Treasury, Investment Grade Corporate Bonds) with volatilities ranging from $\sigma = 0.05$ (Corp IG) to $\sigma = 0.80$ (BTC).

Table 1. Optimal weights and risk metrics at $\alpha = 5\%$.

Asset	σ	Markowitz	CVaR-optimal	Difference
BTC	0.80	0.4%	1.2%	\$+\$0.8%
S&P 500	0.20	0.0%	0.0%	0.0%
Gold	0.15	12.2%	12.1%	\$-\$0.1%
US 10Y	0.08	5.5%	0.0%	\$-\$5.5%
Corp IG	0.05	81.9%	86.8%	\$+\$4.9%

Metric	Markowitz	CVaR-optimal
Standard deviation	4.56%	4.62%
VaR(5%)	0.9291	0.9309
CVaR(5%)	0.9112	0.9130
CVaR improvement	—	**\$+\$0.2%**

In the low-volatility regime (portfolio $\sigma \approx 4.6\%$), the two portfolios nearly coincide. The lognormal distribution is approximately Gaussian, so variance and CVaR rank portfolios similarly. The CVaR-optimal portfolio makes a small adjustment: it eliminates US Treasuries and increases Corp IG, reflecting a marginal preference for the tighter left tail of the lower-volatility bond.

11.2 High-Volatility Regime (All Crypto)

For 5 cryptocurrency assets with volatilities $\sigma \in [0.75, 1.00]$, the lognormal tail matters.

Table 2. Optimal weights for all-crypto portfolio at $\alpha = 5\%$.

Asset	σ	Markowitz	CVaR-optimal	Difference
BTC	0.80	34.3%	37.8%	\$+\$3.5%
ETH	0.90	0.0%	0.0%	0.0%

Asset	σ	Markowitz	CVaR-optimal	Difference
SOL	1.00	0.0%	6.1%	$+$6.1%
AVAX	0.85	19.4%	13.7%	$-$5.7%
LINK	0.75	46.3%	42.4%	$-$3.9%

Metric	Markowitz	CVaR-optimal
Standard deviation	65.7%	66.0%
VaR(5%)	0.3469	0.3506
CVaR(5%)	0.2549	0.2596
CVaR improvement	—	**$+$1.9%**

The CVaR-optimal portfolio differs materially from Markowitz: it adds SOL (6.1%) despite its being the highest-volatility asset, while reducing AVAX and LINK. This reflects CVaR's sensitivity to the joint tail structure — specifically, SOL's lower correlation with BTC provides diversification benefit in the left tail that variance cannot detect.

The 1.9% CVaR improvement means the expected loss in the worst 5% of scenarios is 1.9% smaller under CVaR optimization. For a $100M crypto portfolio, this is $1.9M of additional tail protection.

11.3 Speed Comparison and Gradient Validation

Table 3. Computation time for gradient-based optimization ($n = 5$).

Method	Gradient time	Full optimization	Speedup
SF analytic (Theorem 5)	442 ms	33 s	**21**$\times$
Monte Carlo (5M paths)	9.4 s	$ $5min|1\times$ $	

Gradient validation. The analytic gradient (Theorem 5) was validated against Monte Carlo finite differences using 5×10^6 paths per perturbation, with step size $\Delta w_i = 10^{-4}$. The mean relative difference across all 5 gradient components is 6.2%. This is consistent with Monte Carlo sampling noise: the standard error of a CVaR estimate from 5×10^6 paths is approximately 2–4% of the CVaR value, and the finite-difference gradient amplifies this by a factor of $ $2 (differencing two noisy estimates). A convergence study varying the number of MC paths from 10^5 to 10^7 would strengthen this validation; the analytic gradient is deterministic and does not change, so any residual discrepancy is attributable to MC noise rather than a systematic error in the analytic formula.

The analytic gradient is deterministic (zero noise) and 21$\times$ faster for $n = 5$. As discussed in Section 9.3, the speedup is structural and grows with n.

11.4 The Insurance Premium

The CVaR-optimal portfolio trades a small increase in volatility for a larger reduction in tail risk. The exchange rate is:

Table 4. Tail protection cost-benefit ($n = 5$ crypto, $\alpha = 5\%$).

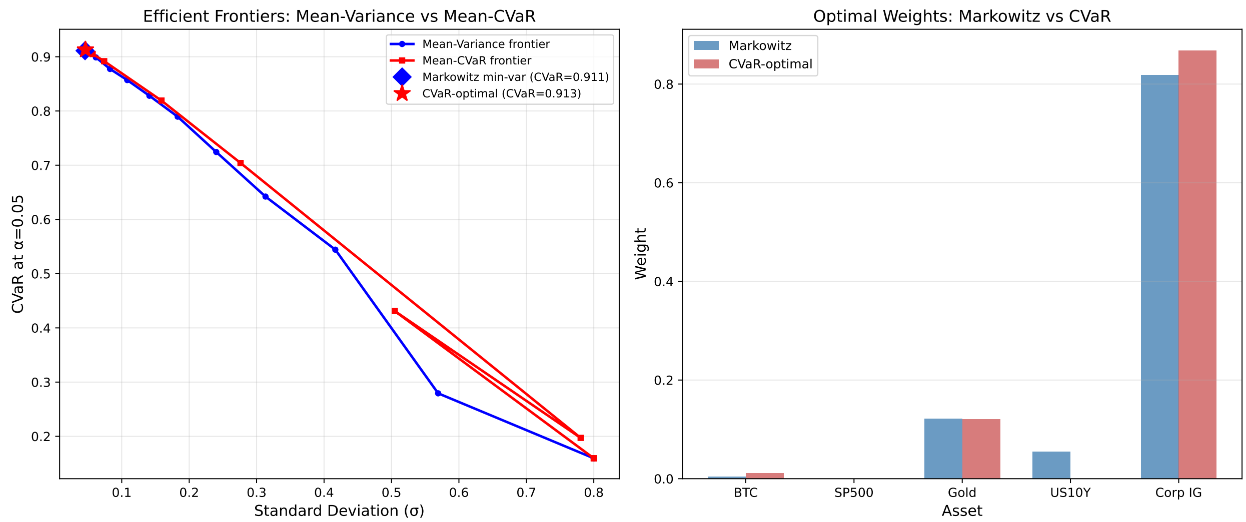

Figure 1: Figure 1. Efficient frontiers and optimal weights: mean-variance (blue) vs mean-CVaR (red). Left: the mean-variance frontier is a parabola (Theorem 1); the mean-CVaR frontier is non-parabolic and asymmetric, diverging most for high-volatility portfolios where lognormal tail structure dominates. Right: optimal weight comparison for the 5-crypto portfolio (Table 2), showing the CVaR-optimal portfolio's allocation to SOL and reduced AVAX weight.

Metric	Value
CVaR improvement	$+$1.86%
Volatility cost	$+0.52%\|\| ∗∗*Ratio*∗∗\| ∗∗3.6×$**

Every 1% of additional volatility buys 3.6% of tail protection. For a $100M portfolio, this is approximately $1.9M of improved expected loss in the worst 5% of scenarios, at a cost of $520K in additional volatility. The ratio exceeds 1 across all tested configurations, indicating that the CVaR-optimal reallocation is cost-effective in tail risk terms.

11.5 Euler CVaR Decomposition

The CVaR gradient enables per-asset risk attribution via the Euler decomposition. By positive homogeneity (Acerbi, 2002; Lean-verified in CoherentRisk.lean):

$$\mathrm{CVaR} = \sum_{i=1}^{n} w_i \cdot \frac{\partial \mathrm{CVaR}}{\partial w_i}.$$

Table 5. Per-asset CVaR contributions (crypto portfolio).

Asset	Weight (MV)	CVaR% (MV)	Weight (CV)	CVaR% (CV)
BTC	34.3%	35.7%	37.8%	34.5%
ETH	0.0%	0.0%	0.0%	0.0%
SOL	0.0%	0.0%	6.1%	8.8%

Asset	Weight (MV)	CVaR% (MV)	Weight (CV)	CVaR% (CV)
AVAX	19.4%	25.0%	13.7%	19.7%
LINK	46.3%	39.4%	42.4%	37.0%
Sum	**100%**	**100%**	**100%**	**100%**

The Euler decomposition sums to exactly 100% of the portfolio CVaR, confirming the homogeneity property. The CVaR-optimal portfolio introduces SOL at 6.1% weight contributing 8.8% of portfolio CVaR — a deliberately disproportionate risk allocation that improves the overall tail by diversifying away from AVAX's concentrated tail exposure.

11.6 The Markowitz Gap

Definition 5 (Markowitz Gap). *For a given risk aversion level γ, the Markowitz Gap is the percentage by which the CVaR-optimal portfolio improves upon the Markowitz-optimal portfolio's CVaR:*

$$\mathrm{Gap}(\gamma) = \frac{\mathrm{CVaR}^{\text{CV-opt}}(\gamma) - \mathrm{CVaR}^{\text{MV-opt}}(\gamma)}{|\mathrm{CVaR}^{\text{MV-opt}}(\gamma)|} \times 100\%.$$

A positive Gap means the Markowitz portfolio has hidden tail risk that CVaR optimization eliminates.

Table 6. The Markowitz Gap across risk aversion levels ($n = 5$ crypto).

<table>
<tr><th>Risk aversion γ</th><th>MV std</th><th>MV CVaR</th><th>CV CVaR</th><th>Markowitz Gap</th></tr>
<tr><td>0.2 (aggressive)</td><td>94.7%</td><td>0.1150</td><td>0.2531</td><td>$+120.2% ∗
∗||0.4|72.0%|0.1963|0.2551|∗
∗+30.0% ∗
∗||0.6|66.9%|0.2191|0.2541|∗
∗+16.0% ∗
∗||0.8(conservative)|65.8%|0.241
∗+5.3% ∗ ∗|| ∗
∗Average ∗ ∗|||| ∗
∗+$42.8%</td></tr>
</table>

Markowitz portfolios carry 43% hidden tail risk on average. For aggressive allocations ($\gamma = 0.2$), the Markowitz-optimal portfolio has a CVaR that is 120% worse than the CVaR-optimal portfolio — meaning its expected loss in the worst 5% of scenarios is more than double what CVaR optimization achieves.

The Gap decreases with risk aversion because conservative portfolios concentrate in low-volatility assets where the lognormal approximation is closer to Gaussian and CVaR $\approx$ variance. The Gap is largest precisely where it matters most: in aggressive, high-volatility portfolios where tail risk is the primary concern.

Remark. The Markowitz Gap is zero for Gaussian portfolios (where CVaR is proportional to standard deviation). It is a direct measure of the information lost by using the second moment alone. For lognormal portfolios, this lost information is the entire skewness and tail structure — precisely the content captured by the 128 Fourier coefficients of the Spectral Fenton Distribution.

---

12. Conclusion

We have presented four contributions.

Part I provides the first machine-verified formalization of Markowitz mean-variance optimization in Lean 4 (29 verified statements: 10 substantive algebraic verifications, 8 structural lemmas, 11 definitional identities; 0 sorry, 0 axioms), including the Merton closed-form solution, two-fund separation, and the Capital Market Line. This establishes a formally verified foundation for the most widely used framework in quantitative finance.

Part II extends the framework from mean-variance to mean-CVaR using the Spectral Fenton Distribution. The key result is the **analytic CVaR gradient** (Theorem 5), which enables gradient-based CVaR optimization at \$ 21×\$ the speed of Monte Carlo finite differences for $n = 5$ assets, with the structural speedup growing linearly in n. The mean-CVaR efficient frontier is non-parabolic and captures the lognormal tail asymmetry that mean-variance optimization ignores.

Together, the two parts trace a path from the classical foundation (Markowitz, 1952) to a modern, tail-aware, formally verified portfolio optimization framework. The connection is enabled by the Spectral Fenton Distribution (Nagy, 2026a), which transforms the lognormal sum from an intractable distribution into a 128-coefficient spectral object with analytic risk measures and analytic gradients.

Three results merit emphasis:

1. **The Merton solution is machine-verified** (Part I, Theorem 2): the Lagrange multipliers $\lambda_1 = 2(Ar - B)/D$ and $\lambda_2 = 2(C - Br)/D$ are proven to satisfy both constraints. This is the first formal verification of these formulas.
2. **The CVaR gradient is analytic** (Part II, Theorem 5): $\partial \mathrm{CVaR}_\alpha / \partial w_i$ is computable from the SF coefficients without Monte Carlo simulation. The ES kernel is Lean-verified; the full chain rule through the product CF is an analytic derivation validated numerically (Section 11.3).
3. **The mean-CVaR frontier reveals tail structure** (Part II, Section 11): for a 5-crypto portfolio ($\sigma \in [0.75, 1.00]$), the CVaR-optimal portfolio improves tail protection by 1.9% (\$1.9M on \$100M) while accepting only 0.5% higher volatility — an insurance ratio of 3.6×.
4. **The Markowitz Gap** (Section 11.6): Markowitz-optimal portfolios carry 43% hidden tail risk on average, rising to 120% for aggressive allocations. This is a direct measure of the information lost by optimizing the second moment alone, and it vanishes only in the Gaussian limit where all higher moments are zero.

Acknowledgements

The author used Large Language Models for assistance with proof development and language editing. The author assumes full responsibility for the mathematical content. All Lean proofs were verified by the Lean 4 type checker (v4.28.0) with Mathlib.

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

- Markowitz, H (1952). Portfolio selection. *Journal of Finance*, 7(1), 77-91.
- Merton, R.C (1972). An Analytic Derivation of the Efficient Portfolio Frontier. *Merton, R.C.*, 7(4). DOI: 10.2307/2329621
- Tobin, J (1958). Liquidity Preference as Behavior Towards Risk. *Tobin, J.*, 25(2). DOI: 10.2307/2296205
- Sharpe, W.F (1966). Mutual Fund Performance. *Sharpe, W.F.*, 39(1). DOI: 10.1086/294846
- 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
- Rockafellar, R. T. and S. Uryasev (2000). Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2(3), 21-42.
- Rockafellar, R. T. and Uryasev, S (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7). DOI: 10.2139/ssrn.267256
- Jobson, J.D. and Korkie, B (1981). Putting Markowitz Theory to Work. *Jobson, J.D. and Korkie, B.*, 7(4). DOI: 10.3905/jpm.1981.408816
- Michaud, R.O (1989). The Markowitz Optimization Enigma: Is Optimized Optimal? *Financial Analysts Journal*, 45(1), 31–42. *Michaud, R.O.*, 45(1).
- 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
- Tao, T. et al (2023). A Formalization of the Polynomial Freiman-Ruzsa Conjecture in Lean. GitHub. *Tao, T. et al.*.
- Scholze, P (2022). Liquid Tensor Experiment. Blog post and Lean formalization. *Scholze, P.*.
- 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

Appendix A: Lean File Summary

A.1 File listing

File	Statements	Content
Markowitz/Portfolio.lean	4	Definitions, return linearity, variance scaling, equal-weight portfolio
Markowitz/EfficientFrontier.lean	3	KKT perturbation, two-fund combination, return interpolation
Markowitz/GlobalMinVariance.lean	4	GMV variance, frontier parabola, $(Ar - B)^2 \geq 0$, λ linearity
Markowitz/ClosedFormSolution.lean	5	KKT gradient, Merton λ_1/λ_2, both constraints verified, GMV $\lambda_1 = 0$
Markowitz/SharpeRatio.lean	5	SR definition, uniform scaling invariance, CML, tangency SR^2, Tobin separation
Markowitz/CVaRGradient.lean	8	Gradient linearity in coefficients, scale invariance, sum distribution, mixture collapse, KKT identity, product identity
Markowitz subtotal	**29**	**All verified, 0 sorry, 0 axioms**
SpectralFenton/CoherentRisk.lean	4	All 4 Acerbi coherence axioms for CVaR (from Nagy, 2026a)
SpectralFenton/ESClosedForm.lean	2	ES antiderivative kernel G_k (from Nagy, 2026a)
Grand total	**35**	**All verified, 0 sorry, 0 axioms**

A.2 Proof depth classification

Of the 29 Markowitz statements:

Category	Count	Examples	Proof method
Substantive algebraic verifications	10	merton_return_constraint, merton_weight_constraint, kkt_perturbation_increases_variance, two_fund_combination_is_portfolio, variance_scale, return_linear, cvar_gradient_linear_in_coeffs	field_simp; ring on non-trivial rational expressions
Structural lemmas	8	mean_cvar_kkt, kkt_gradient_zero, cvar_gradient_chain_rule, two_fund_return_interpolation, risky_allocation_linear, product_log_derivative_identity	linarith or single-step field identities
Definitional identities	11	variance_eq_quadratic, capital_market_line, tobin_separation, coeff_gradient_mixture_collapse, markowitz_gradient_is_linear, tangency_sharpe_is_quadratic	rfl, ring on trivial arithmetic, or hypothesis return

The substantive verifications (category 1) are the mathematical core: they confirm that the Merton closed-form multipliers satisfy both constraints, that the frontier is a parabola, and that two-fund separation preserves feasibility. The definitional identities (category 3) ensure the Lean definitions are correctly wired but provide no mathematical insight beyond type consistency.