State-Contingent Optimality: A Principle for Portfolio Selection

Tony Paul, Nitin

6 August 2025

State-Contingent Optimality: A Principle for Portfolio Selection

Nitin Tony Paul [†]

August 2025

Abstract

This paper explores a normative framework for portfolio selection, the Principle of State-Contingent Optimality (SCO), recasting the classic challenge of finding a single, robust portfolio as a problem in the geometry of distributions. The objective is formulated as minimizing the expected divergence between a portfolio's realized return distribution and a state-dependent, ideal target across all possible market conditions. By employing a metric like the Wasserstein distance, this approach moves beyond simple moments to compare the full shape and character of outcomes, aiming to identify a strategy that is holistically resilient to an uncertain future.

We acknowledge that the principle, in its purest form, rests on profound idealizations: a Platonic target distribution, a knowable state-space, and the validity of ensemble averaging. Rather than treating these as insurmountable barriers, we frame them as explicit signposts for a structured research program. The framework is therefore offered as a theoretical lens, one that cleanly separates the philosophical act of defining investment goals from the mathematical task of achieving them. In doing so, our hope is to provide a more principled way to critique existing methods and guide future inquiry toward truly robust financial solutions.

Keywords: Portfolio Theory, State-Contingent Claims, Stochastic Volatility, Incomplete Markets, Ergodicity, Optimal Transport, Theoretical Finance, Normative Benchmark.
JEL Codes: G11, C02, D52.

[†] Undergraduate student, Model Engineering College, India. Contact: nitintonypaul@gmail.com

Contents

1 Introduction

1.1 Overview

The quest for optimal portfolio allocation was fundamentally transformed by the pioneering work of Harry Markowitz [6], whose **Modern Portfolio Theory (MPT)** laid the groundwork for quantitative investing. While revolutionary for its time—and rightly earning Markowitz a Nobel Prize—MPT offered a relatively static and simplified view of the investment landscape. The real world, however, is far more complex and dynamically evolving.

Subsequent developments in the field, particularly those from the **GARCH** family of models introduced by Bollerslev [4], have made significant strides in addressing these complexities. These models incorporate features such as time-varying volatility, asymmetric effects (e.g., leverage), and regime-switching behavior, bringing us closer to a more realistic understanding of market behavior. Yet, despite these advances, a comprehensive grasp of the market's full dynamics remains elusive—perhaps fundamentally so.

To clearly understand the situation, let us consider a thought experiment that exposes a core challenge for dynamic portfolio theory. Suppose an investor, having set their portfolio weights according to an optimal policy, goes to a party. For the duration of this event, they are disconnected from the market, unable to act upon new information or execute rebalancing trades. It is precisely during this period of inattention that the market regime shifts. This scenario, which we term **The Party Problem**, distills a fundamental question: *What single, static set of weights w would this investor have wished they had chosen beforehand, to be maximally robust against any state of the world they might discover upon their return?*

In light of this challenge, we propose to explore a principle we term *State-Contingent Optimality (SCO)*. SCO seeks to serve as a theoretical benchmark for robust portfolio construction across a continuum of possible market states. The objective is to find a portfolio allocation vector $\mathbf{w}$ that is holistically robust—minimizing its expected divergence from a state-contingent ideal distribution of returns. Rather than treating robustness as a byproduct of other objectives, SCO elevates it to the centerpiece of portfolio design.

1.2 Structure and Flow

This paper is organized as follows. Section 2 reviews foundational work in portfolio optimization, including Modern Portfolio Theory (MVO), risk-based extensions such as CVaR and EVaR, growth-optimal frameworks like Kelly, as well as recent developments in robust optimization and optimal transport. Section 3 outlines the modeling assumptions underlying our framework. Section 4 introduces the core objective function of State-Contingent Optimality (SCO), first in its continuous form, then in its discretized variant, while clarifying key components. Section 5 isolates the single-state formulation and explores a few properties. Section 6 presents an illustrative example using a 3-regime discrete model. Section 7 discusses foundational assumptions and the model's limitations. Section 8 outlines future research directions and unresolved theoretical challenges. Section 9 concludes the paper.

2 Literature Review: Foundations and Motivations

The State-Contingent Optimality (SCO) principle builds upon several distinct but related streams of research in portfolio theory and optimization. It synthesizes insights from classical portfolio selection, risk-centric optimization, robust methods, and the mathematical theory of optimal transport.

2.1 The Classical Paradigm: Mean-Variance Optimization

The bedrock of modern portfolio theory is the seminal work of Markowitz on Mean-Variance Optimization (MVO). This framework revolutionized finance by formalizing the trade-off between risk and return.

Rather than maximizing returns in isolation, Markowitz proposed minimizing portfolio variance for a given level of expected return, or maximizing expected return for a given level of variance. The result is the "efficient frontier," a set of portfolios that are optimal under the mean-variance criterion. The significance of this paradigm cannot be overstated; it provided the first rigorous mathematical language for diversification and established the foundation upon which nearly all subsequent portfolio theory is built.

2.2 Beyond Mean-Variance: Alternative Optimization Principles

While MVO is foundational, its reliance on variance as the sole measure of risk and its assumption of normally distributed returns have led to the development of alternative optimizers targeting more specific principles. These methods often focus on aspects of the return distribution that MVO overlooks.

- The **Kelly Criterion** seeks to maximize the long-run geometric growth rate of capital [1], which is equivalent to maximizing the expected log-utility of wealth. This approach is particularly relevant for long-term investors focused on compound growth.
- **Conditional Value-at-Risk (CVaR)** optimization focuses explicitly on tail risk [2]. Instead of minimizing variance, it seeks to minimize the expected loss in the worst-case scenarios (e.g., the worst 5% of outcomes), providing a direct tool for crash protection.
- Other risk-based approaches, such as **Entropic Value-at-Risk (EVaR)** [3], offer different conceptions of risk and optimality, further enriching the toolkit available to investors.

These optimizers are powerful but typically solve for an optimal portfolio under a single, fixed objective. SCO differs by treating these principles as state-contingent *targets* rather than global objectives.

2.3 Robust Optimization: Acknowledging Input Uncertainty

A major practical challenge for all portfolio optimizers is their sensitivity to input parameters, namely the forecasted means and covariance matrix. Small changes in these inputs can lead to large, often unintuitive, changes in the resulting portfolio weights. **Robust Optimization** confronts this "error maximization" problem directly. Instead of using point estimates for inputs, it assumes these inputs lie within some "uncertainty set" and seeks to find a portfolio that performs well even under the worst-case realization of parameters within that set [10, 9]. This approach marks a philosophical shift towards finding solutions that are robust to model error, a principle central to the motivation for SCO.

2.4 Optimal Transport: The Geometry of Distributions

The heart of the SCO objective function is a metric from the theory of **Optimal Transport (OT)**. Specifically, we employ the **Wasserstein distance**, also known as the Earth Mover's Distance [7]. Unlike traditional statistical divergences (e.g., Kullback-Leibler) that measure differences in distributional shape, the Wasserstein distance quantifies the minimum "cost" or "work" required to transform one probability distribution into another. This cost is defined in terms of both the amount of probability mass moved and the distance it is moved in the underlying space of outcomes. This gives the Wasserstein distance a natural geometric and economic interpretation, making it uniquely suited to measure the "regret" between a portfolio's return distribution and a target ideal.

While historically considered computationally expensive, the practical application of the Wasserstein distance has been revolutionized by the introduction of entropic regularization. The resulting **Sinkhorn algorithm** provides a fast, robust, and differentiable approximation, making it feasible to use optimal transport as the loss function within large-scale optimization problems [8]. This computational breakthrough makes a framework like SCO practically approachable.

3 Model Assumptions

The principle of State-Contingent Optimality rests on three foundational assumptions that define how the market is perceived, how decisions are structured, and how ideals are imposed. These assumptions are not just mathematical conveniences but encode key philosophical commitments about investing under uncertainty.

- **Uncertainty Acknowledgment**
  The investor does not possess precise foresight of future market conditions. Instead, they seek a portfolio that performs robustly across a distribution of plausible future states, conditioned on currently observable market characteristics.
- **State Integrity**
  Each market state S is assumed to maintain its structural properties during evaluation, i.e., states are treated as fixed (non-transitioning) within the optimization context. This allows the framework to consider one state at a time without requiring a full transition model (e.g., Markov chains).
- **Existence of Ideal Portfolios**
  For every state S, there exists a target return distribution $T(S)$, representing the ideal portfolio outcome under that state. This target embodies the investor's principles, risk appetite, and strategic objectives, and serves as the benchmark toward which the actual portfolio distribution is pulled.

4 The Principle of State-Contingent Optimality (SCO)

We begin by defining the core components of the SCO principle. The goal is to find a single, static weight vector $\mathbf{w}$ that is optimal in an aggregate sense over all potential future realities.

4.1 The State-Space (S)

Let S be a state-vector in a high-dimensional, continuous state-space $\mathcal{S} \subseteq \mathbb{R}^d$ representing the fundamental drivers of the market at a given time.

$$S = [\nu, \lambda, \rho, r_f, \dots] \in \mathcal{S} \tag{1}$$

where ν represents a matrix of volatilities, λ a vector of liquidity measures, ρ a matrix of correlations, and r_f the risk-free rate. Crucially, S is understood as a finite-dimensional *projection* of an unobservable, infinite-dimensional true state of the world. The probability measure over this state-space, representing the likelihood of the market materializing in state S, is denoted $p(S)$.

4.2 The Portfolio-Induced Return Distribution ($P(\mathbf{w}|S)$)

A portfolio is defined by a static vector of weights $\mathbf{w} \in \mathbb{R}^N$ over N assets. For any given market state S, this portfolio generates a conditional probability distribution of returns, which we denote $P(\mathbf{w}|S)$. This conditional nature is critical: the return profile of a given allocation is itself a function of the market state. For example, a 60/40 stock-bond portfolio exhibits entirely different distributional properties (mean, variance, skewness, tail behavior) in a high-volatility, inverted-yield-curve state versus a low-volatility, upward-sloping-curve state.

4.3 The State-Contingent Target Distribution ($T(S)$)

We posit the existence of a theoretically optimal target distribution of returns, $T(S)$, for every state S. This represents the return profile an agent with defined preferences *desires* if the market is known to be in state S. The characteristics of $T(S)$, such as target mean, variance, skewness, or tail behavior, are functions of both the agent's preferences and the state itself. Defining $T(S)$ is a problem of *preference elicitation* and is separate from the optimization process. For example, a prospect-theory agent may define $T(S)$ very differently from a log-utility agent, particularly in loss-heavy regions of the state space.

4.4 The Objective Function

The SCO principle seeks to find the single, static set of portfolio weights $\mathbf{w}$ that minimizes the expected true distance between its own conditional distribution and the optimal target distribution, integrated over all possible future states.

$$\min_{\mathbf{w}} \mathbb{E}_S[W(P(\mathbf{w}|S), T(S))] = \int_S W(P(\mathbf{w}|S), T(S))p(S)dS \tag{2}$$

where $p(S)$ is the **probability density function (PDF)** over the state space $\mathcal{S}$ and $W(\cdot,\cdot)$ denotes the **Wasserstein distance**—specifically, the W_p metric from optimal transport theory [7]—which measures the minimal geometric "effort" required to morph one probability distribution into another. Unlike divergences such as Kullback-Leibler (KL) or Jensen-Shannon (JSD), which may become undefined or insensitive when distributions lack overlapping support, Wasserstein distance remains well-behaved and interpretable even in such cases. It defines a true metric on the space of probability measures and is particularly suitable for comparing state-conditioned return distributions in financial contexts.

This focus on a static allocation w is deliberate. It reflects the search for a truly strategic, long-term portfolio that is structurally robust to the unknown future, minimizing the need for constant tactical adjustment and the associated transaction costs, model risk, and parameter instability.

4.5 Discrete Variant: Regime-Based Formulation

While the continuous formulation of SCO provides a coherent framework over a rich market state space, practical implementation often benefits from discretizing the set of possible conditions. In this variant, we consider a finite set of identifiable *regimes*, denoted by $R \in \{R_1, R_2, \ldots, R_K\}$, each representing a meaningful macro-financial condition (e.g., bull, bear, crisis, recovery).

The SCO objective then becomes:

$$\min_{\mathbf{w}} \sum_{r=1}^{K} p(R_r) \cdot W\left(P(\mathbf{w} \mid R_r), T(R_r)\right) \tag{3}$$

where:

- $p(R_r)$ is the empirical or modeled probability of regime R_r,
- $P(\mathbf{w} \mid R_r)$ is the distribution of portfolio returns under weights $\mathbf{w}$ conditioned on regime R_r,
- $T(R_r)$ is the regime-specific ideal (target) return distribution,
- $W(\cdot,\cdot)$ is the Wasserstein distance metric as previously defined.

This regime-based version of SCO is computationally tractable and aligns naturally with empirical market behavior. It enables robust estimation, interpretable diagnostics, and more stable model calibration, particularly when data availability or computation restricts high-dimensional modeling.

5 Single-State Formulation of SCO

5.1 Assumption: Degenerate Market State Distribution

We consider the case where the distribution over market states $p(S)$ collapses to a single deterministic state. That is,

$$p(S) = 1 \quad \text{for a specific } S = S_0.$$

Under this assumption, the SCO objective:

$$\min_w \int_S W(P(w \mid S), T(S)) \cdot p(S)\, dS$$

reduces to:

$$\min_w W\left(P(w \mid S_0), T(S_0)\right) \tag{4}$$

This special case is referred to as the **single-state formulation** of SCO, where the optimal portfolio is found by minimizing the true distance between the portfolio's return distribution and a predefined target distribution, both conditional on the single known market state.

5.2 SCO as a Superset of Portfolio Optimizers

The SCO framework can be shown to encompass traditional portfolio optimization techniques as special cases. The intuition is that since SCO is designed to match an arbitrary target distribution T, it can be configured to replicate any optimizer (e.g., mean-variance, CVaR, Kelly) by setting its target to the optimizer's own distributional output. The following proof formalizes this relationship.

Let Z denote any portfolio optimizer which produces weights w_Z under the forecasted return distribution F. Then define:

$$P(w) := \text{Return distribution under forecast } F \text{ with weights } w$$

$$P(w_Z) := \text{Return distribution under } Z$$

Set the target in single-state formulation as:

$$T := P(w_Z)$$

Then the objective becomes:

$$\min_w W(P(w), P(w_Z))$$

Since W (the Wasserstein distance) is a proper metric (*see Appendix A*) over probability distributions, the minimization problem admits a unique solution when $P(w_Z)$ is fixed and attainable. Therefore,

$$w^\star = w_Z.$$

Hence, **SCO recovers any optimizer** Z as a special case when the target distribution is chosen as the optimizer's own output. This confirms that SCO is a strict superset of all such portfolio optimization techniques.

5.3 Targeting Theoretical Ideal Distributions

In the single-state formulation, we have full flexibility in defining the target distribution T. For instance, one could specify:

- Very low-variance, high-skew distributions (to prefer consistent upside),
- Heavy right-tailed distributions (to favor lottery-like payoffs),
- Compact high-return clusters (to minimize entropy while maximizing mean).

In this formulation, the optimization adjusts the portfolio weights w such that the resulting return distribution $P(w)$ closely approximates the idealized target distribution T. This approach introduces a normative design principle into portfolio construction: *Rather than reacting to the market, we proactively design portfolios to align with the outcomes we wish to realize.*

This enables extremely expressive portfolio construction — going beyond classical tradeoffs of risk and return — and allows experimentation with distributional objectives under full control of shape, tail, skewness, and entropy.

6 An Illustrative Example: 3-Regime Discretization

To render the abstract SCO principle concrete and demonstrate the utility of the tractable, discrete formulation (Eq. 3), we construct a simplified model. Consider a universe with two assets: a single risky asset (e.g., a global stock index) and a risk-free asset. The objective is to find the optimal static allocation, w, to the risky asset.

The process involves four well-defined steps:

1. Regime Definition (The State-Space R)

First, we discretize the continuum of market states into a finite set of meaningful, identifiable regimes. A common and effective approach is to use an observable market indicator like the VIX index. For this illustration, we define three regimes:

- **Regime 1: Bull** (R_1): A low-volatility, risk-on environment, defined as `VIX < 15`.
- **Regime 2: Normal** (R_2): A standard market environment, defined as `15 ≤ VIX < 30`.
- **Regime 3: Crisis** (R_3): A high-volatility, risk-off environment, defined as `VIX ≥ 30`.

2. Regime Probabilities ($p(R_r)$)

Next, we assign a probability to each regime, which can be estimated from historical data. A plausible historical distribution might be:

- $p(R_1) = 0.40$ (40% of the time is in a Bull regime)
- $p(R_2) = 0.50$ (50% of the time is in a Normal regime)
- $p(R_3) = 0.10$ (10% of the time is in a Crisis regime)

3. State-Contingent Target Distributions ($T(R_r)$)

This is the core of the normative approach, where we translate an investor's narrative into a mathematical objective. Suppose the investor's goal is to maintain a relatively stable risk profile across all regimes, de-risking when the market is fragile. This "constant risk" preference can be expressed as a set of target Normal distributions ($\mathcal{N}(\mu_T, \sigma_T^2)$) with a consistent target volatility:

- **Target for Bull (T_1):** $\mathcal{N}(\mu_{T_1} = 0.15, \sigma_{T_1}^2 = 0.12^2)$
- **Target for Normal (T_2):** $\mathcal{N}(\mu_{T_2} = 0.08, \sigma_{T_2}^2 = 0.12^2)$
- **Target for Crisis (T_3):** $\mathcal{N}(\mu_{T_3} = -0.05, \sigma_{T_3}^2 = 0.12^2)$

Here, the investor accepts lower (even negative) expected returns in the crisis regime in exchange for achieving their desired risk profile.

4. Portfolio & Optimization ($P(w|R_r)$)

The actual return distribution of a portfolio with weight w in the risky asset will depend on the empirical properties of that asset in each regime. Let the observed distribution of the risky asset in regime r be $\mathcal{N}(\mu_r, \sigma_r^2)$. The portfolio's distribution is therefore $P(w|R_r) = \mathcal{N}(w\mu_r, (w\sigma_r)^2)$.

The SCO objective is to find the single static weight w that minimizes the expected distance between the portfolio's realized distributions and the investor's target distributions:

$$\min_{w} \sum_{r=1}^{3} p(R_r) \cdot W\Big(\mathcal{N}(w\mu_r, (w\sigma_r)^2)\ ,\ \mathcal{N}(\mu_{T_r}, \sigma_{T_r}^2)\Big) \tag{5}$$

Explicitly expanding the sum for the three regimes, the objective is to find the weight w that minimizes:

$$\begin{aligned} & p(R_1) \cdot W\Big(\mathcal{N}(w\mu_1, (w\sigma_1)^2), \mathcal{N}(\mu_{T_1}, \sigma_{T_1}^2)\Big) + \\ & p(R_2) \cdot W\Big(\mathcal{N}(w\mu_2, (w\sigma_2)^2), \mathcal{N}(\mu_{T_2}, \sigma_{T_2}^2)\Big) + \\ & p(R_3) \cdot W\Big(\mathcal{N}(w\mu_3, (w\sigma_3)^2), \mathcal{N}(\mu_{T_3}, \sigma_{T_3}^2)\Big) \end{aligned} \tag{6}$$

This toy model, while simple, achieves a critical goal: it transforms the abstract, high-level objective of "all-weather" performance into a well-posed and solvable mathematical problem. It directly demonstrates how an investor's narrative and risk preferences can be encoded to find a holistically robust strategic asset allocation, providing a clear roadmap for empirical implementation.

7 Foundational Assumptions and Framework Limitations

The elegance and generality of the SCO formulation rest on several profound assumptions. Acknowledging these is central to understanding the framework's role as a normative benchmark rather than a predictive tool.

7.1 The Exogenous Nature of the Target Distribution $T(S)$

The SCO framework treats the target distribution $T(S)$ as an exogenously given input. This ideal immediately begs the question: *Whose optimum does it represent?* The specific shape of $T(S)$, its mean, variance, skewness, and tail properties, is entirely dependent on the investor's unique utility function, risk posture, and behavioral biases.

The work in behavioral finance, starting with Kahneman and Tversky [5], has shown that a single, universal definition of "rationality" is a fiction. The utility of the SCO framework is that it does not impose one. Instead, it cleanly separates the portfolio selection problem into two distinct parts:

1. The deeply personal, often philosophical problem of defining one's investment goals in the form of a target distribution $T(S)$.
2. The subsequent, purely mathematical problem of finding the portfolio $\mathbf{w}$ that best approximates this goal across all states.

7.2 Assumption of a Knowable State-Space and Measure

The integration over the state-space $\mathcal{S}$ implicitly assumes that this space is fully specified and that its probability measure $p(S)$ is known. This is a powerful idealization, akin to assuming a world where every possible future contingency has a well-defined and knowable likelihood.

In reality, the true market state-space is vast and complex, and the probability measure $p(S)$ is unknowable. Its estimation is one of the central, unsolved challenges in econometrics. The true, continuous SCO framework is therefore theoretically sound only in a world of perfect information. Its practical utility is thus not in its direct application as if it were a perfect model, but in its ability to serve as a theoretical benchmark—a measure of what *could* be achieved if the markets and their probabilities were, in fact, fully understood.

7.3 Ensemble vs. Time Averages in Implementation

By its construction, the SCO objective function is an **ensemble average**. It seeks the portfolio that performs best, on average, across a vast set of "parallel universes," with each universe representing a different possible future state S.

However, in any real-world application, a fundamental tension arises. The true probabilities of these future states, $p(S)$, are unknown. We are forced to estimate them using the single path of history we have observed. Typically, we use the historical frequency of different market regimes as our best guess for their future likelihood.

This creates a direct dependency on the historical data used.

- A **short lookback** period makes the resulting portfolio highly tactical, as its view of the world is shaped heavily by recent market events.
- A **very long lookback** period attempts to find a more stable, strategic average, assuming the long-run past is representative of the long-run future.

Unlike traditional models which often implicitly blend these concepts, the SCO architecture makes this choice explicit. It cleanly separates the step of forming beliefs about the world (choosing $p(S)$ based on historical data) from the subsequent optimization. This forces the modeler to consciously decide how much of history is relevant when constructing a portfolio for the future.

7.4 Computational Complexity and Non-Convexity

The SCO objective function, an expected divergence over a space of distributions, presents significant computational challenges. The function is generally **non-convex** with respect to the portfolio weights $\mathbf{w}$. The Wasserstein distance itself is convex in its distributional arguments, but the mapping from weights $\mathbf{w}$ to the portfolio distribution $P(\mathbf{w}|S)$ is typically non-linear, inducing non-convexity in the overall problem.

This means that standard convex optimization solvers cannot guarantee a globally optimal solution. Finding the optimal $\mathbf{w}$ requires either specialized global optimization algorithms, which are computationally intensive, or the use of iterative, gradient-based methods that may converge to a local minimum. Furthermore, the calculation of the Wasserstein distance within each step of the optimization can be demanding. Practical implementation often relies on approximations, such as the fast but approximate **Sinkhorn algorithm**, to make the problem tractable, especially in high-dimensional settings.

8 Discussion and Future Work

The impracticality of directly solving the full SCO problem does not render it useless. Its value lies in its capacity to potentially structure thought and guide research. The SCO principle serves as a potential benchmark for robust optimization and the limitations discussed above are not roadblocks but signposts pointing to the most critical open questions in finance:

1. **Preference Elicitation for $T(S)$:** How can we map high-level investor narratives ("protect me from crashes") into the functional form of a state-contingent target distribution $T(S)$? Can techniques from inverse optimization or machine learning infer the market-implied $T(S)$ from observed asset prices?
2. **Dimensionality Reduction for S:** Given that the true state-space is intractable, what are the minimal sets of observable state variables that capture the maximum variance in portfolio performance distributions? This is a sophisticated re-framing of the "factor investing" problem.
3. **Non-Ergodic Formulations:** How would the objective function change if we replaced the ensemble integral with a path-wise, time-averaged objective? This would involve optimizing over a space of paths rather than a space of states, a significantly more complex but potentially more realistic problem.

9 Conclusion

This paper has explored the Principle of State-Contingent Optimality (SCO), a normative framework for portfolio selection. We have formulated an objective function based on minimizing the expected distributional divergence between a portfolio's returns and a state-dependent ideal. This approach aims to provide a theoretically robust solution to **The Party Problem**—the challenge of finding a single, static portfolio that can endure periods of strategic inaction. Rather than shying away from the profound assumptions this framework requires, we have centered our analysis on them, arguing that their explicit acknowledgment is a necessary step toward more transparent and principled financial modeling.

The true, continuous form of the SCO framework is presented not as a readily applicable algorithm, but as a theoretical destination. Its primary value lies in the questions it forces us to ask and the coherent structure it provides for future inquiry. The discrete variant of SCO, however, offers a more tangible path toward implementation. By defining this problem with greater structure, our hope is to provide a lens through which we can better judge the partial solutions we build along the way, potentially guiding us toward more robust portfolios.

References

[1] J. L. Kelly Jr. *A New Interpretation of Information Rate*. Bell System Technical Journal, 35(4), 917–926, 1956.

[2] R. T. Rockafellar and S. Uryasev. *Optimization of Conditional Value-at-Risk*. Journal of Risk, 2(3), 21–42, 2000.

[3] A. Ahmadi-Javid. *Entropic Value-at-Risk: A New Coherent Risk Measure*. Journal of Optimization Theory and Applications, 155(3), 1105–1123, 2012.

[4] T. Bollerslev. *Generalized Autoregressive Conditional Heteroskedasticity*. Journal of Econometrics, 31(3), 307–327, 1986.

[5] D. Kahneman and A. Tversky. *Prospect Theory: An Analysis of Decision under Risk*. Econometrica, 47(2), 263–291, 1979.

[6] H. Markowitz. *Portfolio Selection*. The Journal of Finance, 7(1), 77–91, 1952.

[7] C. Villani. *Optimal Transport: Old and New*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 2009.

[8] M. Cuturi. Sinkhorn Distances: Lightspeed Computation of Optimal Transport. In *Advances in Neural Information Processing Systems 26*, pages 2292–2300, 2013.

[9] F. J. Fabozzi, P. N. Kolm, D. A. Pachamanova, and S. M. Focardi. *Robust Portfolio Optimization and Management*. John Wiley & Sons, 2007.

[10] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

A Glossary of Key Terms

Continuum A continuous and unbroken range or spectrum of values or states. In this paper, it refers to the full set of possible market conditions or states, treated as a continuous domain rather than a finite set

State Space ($\mathcal{S}$) The set of all possible market states, each represented by a vector S of observable market factors (e.g. volatilities, correlations, liquidity measures, risk-free rate).

Market State (S) A single element of $\mathcal{S}$, describing the market environment at a point in time. In the discrete variant, a "regime" R_r is a labelled market state.

Portfolio-Induced Return Distribution ($P(w \mid S)$) The probability distribution of portfolio returns under static weights w when the market is in state S.

Target Distribution ($T(S)$) The investor's ideal (Platonic) return distribution for state S, encoding preferences for mean, variance, skew, and tail behavior.

Probability Density Function (PDF) A function that describes the relative likelihood of a continuous random variable taking on a given value. While the probability at any exact point is zero, the area under the PDF over an interval represents the probability of the variable falling within that range. For a random variable X, the PDF $f(x)$ satisfies $P(a \leq X \leq b) = \int_a^b f(x)\,dx$.

VIX The *Volatility Index*, representing the market's expectation of 30-day forward-looking volatility. Often interpreted as a "fear gauge," it is derived from S&P 500 index options and used as a proxy for implied volatility.

Exogenous A variable or shock that originates from outside the system being modeled. In finance, exogenous factors (e.g., macroeconomic news, policy changes) affect asset prices but are not determined within the model.

Convexity A measure of the curvature in the relationship between a bond's price and its yield, or more generally, the property of a function where the line segment between any two points lies above the graph. In portfolio theory, convexity often signals desirable optimization properties (e.g., uniqueness of the optimum).

Divergence A function $D(P\|Q)$ measuring discrepancy between two distributions.

Proper Metric A function $d(P, Q)$ that satisfies the four axioms of a metric space:
(1) Non-negativity
(2) Identity of indiscernibles ($d(P, Q) = 0 \iff P = Q$)
(3) Symmetry ($d(P, Q) = d(Q, P)$)
(4) The triangle inequality ($W_p(P, R) \leq W_p(P, Q) + W_p(Q, R)$)

Regime A discrete classification of market state (e.g. "Bull," "Normal," "Crisis") used to approximate a continuous state-space by a finite sum.

Ergodicity The assumption that an ensemble average (over states) equals a time average (over a single path). In SCO, the objective is an ensemble expectation, which may diverge from individual realized paths if markets are non-ergodic.

Ensemble A collection of models or forecasts, often combined to improve robustness or accuracy. Ensemble methods aggregate predictions from multiple sources (e.g., neural nets, decision trees, or volatility forecasts) to produce a unified estimate.

B Mathematical Formulations of Optimization Frameworks

This appendix provides the formal mathematical definitions for the optimization principles and metrics discussed in the main text.

B.1 Mean-Variance Optimization (MVO)

The classic Markowitz MVO solves the following quadratic programming problem to find the portfolio on the efficient frontier for a given target return μ_p:

$$\begin{aligned} \min_{w} \quad & w^T \Sigma w \\ \text{subject to} \quad & w^T \mu = \mu_p \\ & w^T \mathbf{1} = 1 \end{aligned} \tag{7}$$

where:

- $w \in \mathbb{R}^N$ is the vector of portfolio weights.
- $\Sigma \in \mathbb{R}^{N \times N}$ is the covariance matrix of asset returns.
- $\mu \in \mathbb{R}^N$ is the vector of expected asset returns.
- μ_p is the desired target expected return for the portfolio.
- $\mathbf{1}$ is a vector of ones.

B.2 Conditional Value-at-Risk (CVaR)

CVaR optimization seeks to minimize the expected loss in the tail of the return distribution. For a given confidence level α (e.g., 95%), the problem is to find the portfolio weights w that solve:

$$\min_{w,\zeta} \quad \zeta + \frac{1}{1-\alpha} \int_{r \in \mathbb{R}} \max(-r - \zeta, 0) p(r; w) dr \tag{8}$$

where:

- ζ is the Value-at-Risk (VaR) at the α confidence level, which is optimized jointly with w.
- r represents the portfolio return, whose distribution $p(r; w)$ depends on the weights w.
- The integral calculates the expected loss conditional on the loss exceeding VaR.

This is often solved using a linear programming formulation by discretizing the return distribution into a set of scenarios.

B.3 Kelly Criterion (Log-Utility Formulation)

The Kelly criterion aims to maximize the expected geometric growth rate of wealth. This is equivalent to maximizing the expected logarithm of wealth, which can be expressed as:

$$\max_{w} \quad \mathbb{E}[\log(1 + w^T R)] \tag{9}$$

where:

- w is the vector of portfolio weights.
- $R \in \mathbb{R}^N$ is the random vector of asset returns.
- $\mathbb{E}[\cdot]$ denotes the expectation over the joint distribution of returns R.

Under assumptions of normally distributed returns, this can be approximated by the analytical solution $\arg\max_w (w^T \mu - \frac{1}{2} w^T \Sigma w)$.

B.4 Entropic Value-at-Risk (EVaR)

EVaR is a coherent risk measure derived from Chernoff's inequality. It provides an upper bound on VaR and CVaR and has favorable properties. The EVaR at confidence level $1 - \alpha$ is defined as:

$$\text{EVaR}_{1-\alpha}(X) = \inf_{z>0} \frac{1}{z} \left(\log \left(\frac{M_X(z)}{\alpha} \right) \right) \tag{10}$$

where:

- X is the random portfolio loss (negative return).
- $M_X(z) = \mathbb{E}[e^{zX}]$ is the moment-generating function of the loss.
- z is a positive parameter over which the minimization is performed.

Portfolio optimization with EVaR involves minimizing this value for the portfolio's loss distribution.

B.5 Wasserstein Distance

The p-Wasserstein distance between two probability distributions μ and ν on a space Ω is defined as the solution to an optimal transport problem:

$$W_p(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu, \nu)} \int_{\Omega \times \Omega} d(x, y)^p \, d\gamma(x, y) \right)^{1/p} \tag{11}$$

where:

- $\Gamma(\mu, \nu)$ is the set of all joint distributions (transport plans) $\gamma(x, y)$ whose marginals are μ and ν respectively.
- $d(x, y)$ is a distance metric on the space Ω (e.g., the Euclidean distance for returns).
- The integral represents the total "cost" of moving the mass of μ to match ν according to the plan γ.

B.6 Sinkhorn Distance (Entropic Regularized OT)

The Sinkhorn distance is a computationally efficient approximation of the Wasserstein distance. It adds an entropic regularization term to the optimal transport problem:

$$W_\lambda(\mu, \nu) = \min_{\gamma \in \Gamma(\mu, \nu)} \int_{\Omega \times \Omega} d(x, y) \, d\gamma(x, y) - \lambda H(\gamma) \tag{12}$$

where:

- $H(\gamma) = -\int \gamma(x, y) \log \gamma(x, y) \, dxdy$ is the entropy of the transport plan γ.
- $\lambda > 0$ is the regularization strength. A higher λ leads to a faster but "blurrier" approximation.

This regularized problem can be solved with extreme speed using the iterative Sinkhorn-Knopp algorithm.