mathematics

MDPI

Article

Managing Risk Across Time: An Intertemporal Spectral Risk Measures Framework for Multi-Period Portfolio Optimization

Chengneng Jin and Jianjun Gao *

School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China; a@stu.sufe.edu.cn
* Correspondence: gao.jianjun@shufe.edu.cn

Abstract: This paper introduces a novel framework for multi-period portfolio optimization that incorporates intertemporal spectral risk measures (ISRMs). The model dynamically manages risk by considering both tail risk, through spectral risk measures, and overall portfolio volatility, through variance, across multiple time periods. This approach allows investors to specify time-varying risk preferences via a spectral function, making it particularly suitable for investors with evolving risk management needs. We develop an efficient solution methodology based on the Progressive Hedging Algorithm (PHA), enhanced with specialized reformulations to handle linkage objectives and constraints inherent in the multi-period setting. We establish the theoretical convergence properties of our algorithm, demonstrating a q-linear convergence rate under mild conditions. Numerical experiments validate the effectiveness of our approach, showing that the intertemporal weighting scheme provides more consistent risk management across the investment horizon compared to terminal-focused strategies. Notably, our approach exhibits superior downside risk protection, as evidenced by improved Sortino and Omega ratios, and generates more balanced wealth distributions with moderate tails. These findings offer valuable insights and practical tools for investors seeking to implement dynamic risk-management strategies that account for both intermediate and terminal objectives.

Keywords: multi-period portfolio optimization; spectral risk measures; Progressive Hedging Algorithm; dynamic risk management

MSC: 91G10; 91G70

check for updates

Academic Editor: Andrea Scozzari

Received: 25 April 2025
Revised: 21 May 2025
Accepted: 23 May 2025
Published: 25 May 2025

Citation: Jin, C.; Gao, J. Managing Risk Across Time: An Intertemporal Spectral Risk Measures Framework for Multi-Period Portfolio Optimization. *Mathematics* **2025**, *13*, 1754. https://doi.org/10.3390/math13111754

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The mean-variance (MV) portfolio selection model, introduced by [1], has been widely utilized in both academic research and investment practice. However, a drawback of the MV formulation is its symmetric nature, which penalizes both gains and losses equally. To address this issue, various risk measures have been introduced into the portfolio optimization framework. The concept of a *coherent risk measure* [2], along with its extension, the *convex risk measure* [3], has laid down fundamental principles for defining more reasonable risk measures. Along with the development of modern risk measures, the mean-risk portfolio optimization models have been extensively studied under the frameworks of static and dynamic portfolio optimizations (e.g., [4–11]). Recently, researchers (e.g., [12–14]) have also demonstrated the successful application of the downside risk-based models in portfolio insurance management. Among these studies, one particular direction focuses on the portfolio optimization model with multiple hybrid risk measures (e.g., [10,11,15–17]). This line of research is motivated by the fact that, in cases where the distribution of portfolio

returns is asymmetrical and non-Gaussian, solely minimizing variance is insufficient for effectively managing downside risk.

This paper adopts a similar approach to previous studies, exploring the dynamic mean-risk portfolio decision model by integrating both the mean-variance (MV) formulation and the *spectral risk measure* (SRM). Building on the foundational work of [18], SRM calculates risk as a weighted average of the quantiles of the return (or wealth) distribution. When the weighting function, also referred to as the spectrum, satisfies certain mild conditions, SRM qualifies as a coherent risk measure, as demonstrated by [9,18]. While SRM represents a specific case within the broader spectrum of risk measures, such as the distortion risk measure or weighted VaR [7,9], it possesses several advantageous properties, including convexity, tractability, and flexibility when selecting the spectrum function. The power of distortion functions to capture nuanced risk preferences is further highlighted by their use in other advanced frameworks, such as distributionally robust optimization under distorted expectations [19]. These attributes make SRM a particularly suitable choice for constructing portfolio optimization models.

In this study, we propose a novel hybrid portfolio optimization model that integrates multiple risk measures, specifically the dynamic SRM-MV portfolio optimization model. The integration of SRM into the dynamic MV framework is significant for several reasons. First, it enhances the ability to shape the distribution of terminal wealth, particularly by strengthening the management of downside risk in the loss domain. Recent research by [20] highlights that the terminal wealth distribution under a dynamic MV policy may exhibit a long tail in the loss domain, underscoring the importance of incorporating downside risk measures to control tail risk. Second, the inclusion of SRM diminishes the dominance of variance in the objective function, thereby alleviating the conservatism associated with variance-based risk management within the gains domain. This hybrid approach not only improves the robustness of portfolio optimization but also aligns more closely with real-world risk management needs.

On the other hand, in the traditional multi-period portfolio selection model, the focus is primarily on the risk and return associated with terminal wealth. However, in real-world investment scenarios, investors are also concerned with the performance of the portfolio during intermediate periods within the investment horizon. To address this, intertemporal risk–reward restrictions have been introduced in the mean-variance (MV) portfolio selection framework. These restrictions focus on the intermediate expected values and variances in the portfolio, thereby enabling investors to impose constraints that help mitigate significant variations in portfolio performance. This approach is particularly beneficial for institutional investors with evolving risk management needs, as it allows them to dynamically manage both tail risk and overall portfolio volatility. The problem of controlling intermediate portfolio behavior, whether in terms of returns or variance, was first explored by [21]. They proposed a solution using the embedding method introduced by [22]. However, the computational complexity of this approach posed challenges for practical implementation. To address this, [23] developed a novel mean-field framework. This framework provides an efficient modeling tool and offers a semi-analytical solution scheme, significantly simplifying the process of tackling such problems. Their work represents a significant advancement in multi-period portfolio optimization, bridging the gap between theoretical models and practical applications. The trend of embedding intertemporal risk controls within multi-period mean-variance frameworks has continued, with models incorporating, for example, intertemporal variances directly in the objective function [24], or featuring intertemporal constraints as explored by [25].

Although some research has studied the portfolio optimization model with hybrid risk measures or the intertemporal risk–reward restriction, the current literature provides

little insight on a model including both of these features. Our work meet this challenge. Our contribution is threefold. First, we introduce a novel formulation that combines SRMs with variance measures across multiple time periods, allowing investors to manage both tail risk and overall portfolio volatility dynamically. The flexibility of our model enables investors to specify time-varying risk preferences through the spectral function, making it particularly suitable for institutional investors with evolving risk management requirements. Second, we develop an efficient solution methodology based on the Progressive Hedging Algorithm (PHA), enhanced with specialized reformulations for handling linkage objectives and constraints. Our theoretical analysis provides strong convergence guarantees, demonstrating that the algorithm achieves a q-linear convergence rate under mild conditions. Third, the numerical results validate the effectiveness of our approach, showing that the intertemporal weighting scheme provides more consistent risk management across the investment horizon compared to terminal-focused strategies. Particularly noteworthy is the superior performance in terms of downside risk protection, as evidenced by improved Sortino and Omega ratios. These findings have significant practical implications for portfolio managers seeking to implement dynamic risk-management strategies that account for both intermediate and terminal objectives.

The remainder of this paper is organized as follows: Review of the Related Literature reviews the relevant literature on portfolio optimization, risk measures, and solution methodologies. Section 2 presents our mathematical formulation of the multi-period portfolio optimization problem incorporating intertemporal spectral risk measures. Section 3 develops our solution methodology based on the Progressive Hedging Algorithm and establishes its theoretical properties. Section 4 provides comprehensive numerical experiments that demonstrate the effectiveness of our approach and compare different weighting schemes. Finally, Section 5 concludes with a discussion of implications and future research directions.

For notational clarity, we adopt the following conventions throughout this paper: bold lowercase letters (e.g., $\boldsymbol{x}$, $\boldsymbol{u}$) denote vectors, while bold uppercase letters (e.g., $\boldsymbol{A}$, $\boldsymbol{B}$) represent matrices. Scalar quantities are indicated by non-bold letters (e.g., α, r). The transpose of a vector or matrix is denoted by a superscript $\top$ (e.g., $\boldsymbol{x}^\top$, $\boldsymbol{A}^\top$). We use $\langle \cdot, \cdot \rangle$ to represent the Frobenius inner product of two matrices, defined as $\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \mathrm{tr}(\boldsymbol{A}^\top \boldsymbol{B})$, where $\mathrm{tr}(\cdot)$ denotes the trace operator. For vectors, $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\top \boldsymbol{y}$ represents the standard inner product. The Euclidean norm of a vector is denoted by $\|\boldsymbol{x}\|$, and the Frobenius norm of a matrix by $\|\boldsymbol{A}\|_F$. We use $\boldsymbol{R}$ to denote the set of real numbers, $\boldsymbol{R}^n$ for n-dimensional real vector space, and $\boldsymbol{R}^{m \times n}$ for the space of real $m \times n$ matrices. The expectation operator is denoted by $\boldsymbol{E}[\cdot]$, and probability by $\boldsymbol{P}(\cdot)$.

Review of the Related Literature

This research is closely aligned with the pursuit of a multiple-mean-risks portfolio optimization model. In the realm of static portfolio optimization, ref. [16] proposed a model integrating variance and Conditional Value-at-Risk (CVaR) as risk measures. By leveraging the parametric representation of CVaR [26], this problem can be reformulated as a convex quadratic programming (QP) problem. Building on this foundation, ref. [17] explored VaR-MV portfolio optimization and demonstrated superior out-of-sample performance compared to equally weighted portfolios and MV-CVaR portfolios. In the context of multi-objective optimization, ref. [27] investigated the inverse fund optimization problem. However, the dynamic version of such portfolio models remains underexplored in the literature. In continuous-time market settings, ref. [15] examined portfolio optimization models that incorporate CVaR or the Safety-First Principle within the mean-variance (MV) framework. Beyond the traditional mean-risk formulation, researchers have also focused on

utility maximization problems with downside risk constraints, such as VaR, the Safety-First Principle, or variance (see, e.g., [14,28–30]).

As for the static SRM-based portfolio optimization model, ref. [18] formulate it as a linear programming problem when representing uncertainty with discrete scenarios. The performance of the SRM-based model is later evaluated by [9], while [31] explore a model incorporating multiple SRMs. Building on these foundations, ref. [32] recently extend these static SRM models to robust formulations and proposed a tractable solution approach. The applicability of SRMs is also being recognized in other complex decision-making paradigms; for example, ref. [33] leverage static SRMs to enhance risk-awareness in distributional reinforcement learning, offering an alternative to more common risk measures like CVaR in that domain. In the realm of continuous-time portfolio optimization based on SRM, our research represents a pioneering effort in addressing such problems. This work also intersects with VaR-based portfolio optimization. The authors of [6] are the first to investigate mean-VaR portfolio optimization in a continuous-time setting. To handle the ill-posed nature of this problem, they introduced an artificial upper bound for wealth. A similar approach has been employed to tackle mean Safety-First portfolio optimization [34] and the Weighted-VaR-based model [7]. Meanwhile, another line of research has sought to extend the concept of static SRM to dynamic SRM [35].

In this work, we employ the Progressive Hedging Algorithm (PHA), introduced by [36], to address our problem. PHA is a powerful iterative algorithm designed for solving large-scale multi-period stochastic optimization problems. By effectively combining scenario decomposition and augmented Lagrangian relaxation, the PHA has demonstrated exceptional performance in handling multi-stage stochastic programs. Its versatility has led to widespread applications across various domains, including resource allocation [37–39], supply chain management [40,41], and financial portfolio optimization [42,43]. One of the standout features of PHA is its ability to handle non-convex problems while providing convergence guarantees, making it an invaluable tool for tackling complex stochastic optimization challenges.

2. Problem Formulation

In this section, we formulate a multi-period portfolio optimization problem that incorporates intertemporal spectral risk measures (ISRMs) to manage risk dynamically over time. The model captures the evolution of an investor's wealth as they allocate assets in a financial market comprising both risk-free and risky assets. The objective is to minimize a combination of SRMs and variances in wealth in each period while satisfying the expected wealth targets.

2.1. Market Model and Wealth Dynamics

We consider a discrete-time financial market with a finite horizon T, divided into T periods indexed by $t = 0, 1, \ldots, T-1$. The market consists of two types of assets. First, a risk-free asset that produces a deterministic return r^f_{t+1} from time t to $t+1$. Second, a set of M risky assets with uncertain returns. The return vector on the risk-free rate for the period from t to $t+1$ is denoted by $\boldsymbol{e}_{t+1} \in \boldsymbol{R}^n$.

The investor begins with an initial wealth $x_0 > 0$. At each time t, the investor decides how to allocate wealth between the risk-free asset and the risky assets. We define x_t as the total wealth at time t, and $\boldsymbol{u}_t = (u_{t,1}, u_{t,2}, \ldots, u_{t,M})^\top$ as the allocation vector at time t, where $u_{t,i}$ is the amount invested in the i-th risky asset. The amount invested in the risk-free asset at time t is implicitly $x_t - \mathbf{1}^\top \boldsymbol{u}_t$, where $\mathbf{1}$ is an n-dimensional vector of ones. The evolution

of the investor's wealth is governed by the portfolio returns and the allocation decisions. The wealth at time $t+1$ is given by:

$$x_{t+1} = r_{t+1}^f(x_t - \mathbf{1}^\top \boldsymbol{u}_t) + \boldsymbol{e}_{t+1}^\top \boldsymbol{u}_t = r_{t+1}^f x_t + \boldsymbol{r}_{t+1}^\top \boldsymbol{u}_t, \quad t = 0, 1, \ldots, T-1. \tag{1}$$

where $\boldsymbol{r}_{t+1} = \boldsymbol{e}_{t+1} - r_{t+1}^f$ is the excess return of risky asset for the period from t to $t+1$. This equation captures the return from the risk-free investment and the excess returns from the risky assets based on the allocations.

2.2. Risk Measures

To quantify the risk of the portfolio, we employ two types of risk measures. First, we utilize variance (Var), a traditional measure that captures the dispersion of wealth around its mean. Second, we use the Spectral Risk Measure (SRM) to manage the downside risk of wealth of portfolio. The SRM is an important risk measure in modern risk measure theory, since it is a coherent, comonotonic additive, and law-invariant risk measure (see, e.g., [18,44,45]). Given a probability space with probability measure $\boldsymbol{P}$, we use $F_t(y) \triangleq \boldsymbol{P}(x_t \leq y)$ and $G_t(s) \triangleq \inf\{y \in \boldsymbol{R} | F_t(y) > s\}$ to denote the cumulative distribution function and the upper quantile function x_t, respectively. Following a similar definition in [18], we define an intertemporal SRM, as follows. (The intertemporal SRM implies that the SRM throughout the entire investment horizon is also taken into account.)

$$\mathrm{SRM}_{\phi_t}(x_t) = -\int_0^1 \phi_t(s) G_t(s)\, d\alpha, \tag{2}$$

where $\phi_t(\cdot)$ is a time-dependent, real-valued, non-negative function, with a non-increasing, right-continuous integrable function defined on the interval $[0,1]$, satisfying $\int_0^1 \phi_t(s)\, ds = 1$ for all t. Within the SRM framework, different weighting functions, $\phi_t(\cdot)$, also known as the spectra, yield distinct risk measures. For instance, if $\phi_t(\cdot)$ is a step function defined as $\phi_t(s) = \frac{1}{\gamma}\mathbf{1}_{\{0 \leq s \leq \gamma\}}$, then (2) becomes $\mathrm{SRM}_{\phi_t}(x_t) = -\frac{1}{\gamma}\int_0^\gamma G_t(s) ds$, which is commonly known as the γ-level Expected Shortfall (ES). Beyond ES, there are alternative methods to specify the spectrum. Drawing upon utility theory, (ref. [46] suggests that utility functions can be used to design suitable spectrum functions in SRM.) it is possible to define exponential and power-function-based spectra, as follows [47]:

$$\text{Exponential spectrum: } \psi(s) \triangleq \frac{k_e \cdot e^{-k_e \cdot s}}{1 - e^{-k_e}} \text{and power spectrum: } \psi(s) \triangleq k_p \cdot s^{k_p - 1},$$

where $k_e \in (0, \infty)$ and $k_p \in (0, 1]$ are the parameters (see, e.g., [47,48]). Figure 1 displays the typical shapes of the exponential-, power-, and step-function-based spectra. One crucial characteristic of these spectral functions is that they assign more weight to smaller probabilities than larger ones.

Armed with these definitions of wealth dynamics and risk measures, we can now formally state the multi-period portfolio optimization problem.

2.3. Multi-Period Portfolio Optimization Model

The investor's objective is to minimize a weighted sum of the SRMs and variances in wealth over all periods, subject to expected wealth targets and the wealth dynamics:

$$\begin{aligned} (\hat{\mathcal{P}}) \qquad \min_{\{\boldsymbol{u}_t \in \mathcal{A}_t\}|_{t=0}^{T-1}} & \sum_{t=1}^{T} \mu_t \mathrm{SRM}_{\phi_t}(x_t) + \sum_{t=1}^{T} \kappa_t \mathrm{Var}(x_t), \\ \text{s.t.} & \boldsymbol{E}[x_t] = d_t, \forall t \in \{1, \ldots, T\} \\ & \{x_t, \boldsymbol{u}_t\} \text{ satisfy} (1), \forall t \in \{1, \ldots, T\}, \end{aligned}$$

where $\mu_t \geq 0$ and $\kappa_t \geq 0$ are the weights assigned to the SRM and variance at time t, respectively, which are used to balance the importance between the variance and SRM. d_t is the target expected wealth at time t, which should be set greater than $x_0 \prod_{t=0}^{t-1} r_t^f$. This ensures that the optimization problem seeks a strategy that outperforms the risk-free investment over the entire horizon. $\mathcal{A}_t$ is the feasible set of allocations $\boldsymbol{u}_t$ at time t. This set is defined by constraints derived from regulatory requirements or investment policies. Such constraints may include the following (the unconstrained problem, where $\boldsymbol{u}_t \in \boldsymbol{R}^n$, can be considered a special case of our more general constrained formulation): (1) Budget constraint: $\mathbf{1}^\top \boldsymbol{u}_t \leq x_t$. (2) No-short-selling constraint: $\boldsymbol{u}_t \geq 0$.

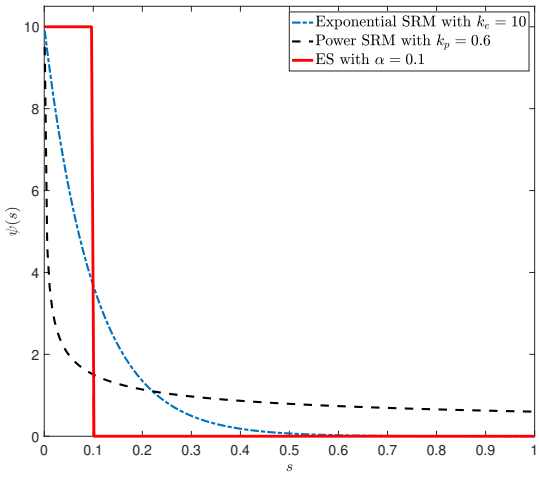


Figure 1. The spectrum function $\phi_t(s)$ in the spectral risk measure.

This framework encompasses several well-studied portfolio optimization problems as special cases. When all risk measure weights are concentrated at the terminal stage (i.e., $\mu_t = \kappa_t = 0$ for $t \leq T-1$ and $\mu_T \geq 0, \kappa_T \geq 0$), the problem reduces to either the dynamic mean-variance optimization model introduced by [22] when $\mu_T = 0$ and $\kappa_T > 0$, or the dynamic mean-CVaR problem studied by [49,50] when $\mu_T > 0$, $\kappa_T = 0$, and CVaR is chosen as a specific instance of SRM. Furthermore, when considering only variance-based risk measures distributed across multiple periods (i.e., $\mu_t = 0$ and $\kappa_t > 0$ for all $t \leq T$), our framework corresponds to the dynamic mean-variance problem with intertemporal variance controls developed by [24].

While the formulation $(\hat{\mathcal{P}})$ captures the desired economic objectives, solving it directly poses significant challenges. The multi-period stochastic setting necessitates dealing with nonanticipativity constraints and evolving information structures. The primary computational hurdle stems from the spectral risk measure term, $\mathrm{SRM}_{\phi_t}(x_t)$, defined via the quantile function $G_t(s)$ of the wealth x_t. Obtaining an analytical expression for $G_t(s)$ is generally intractable, as the distribution of x_t is a complex result of the sequence of portfolio decisions $(\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{t-1})$ and the underlying random returns. Additionally, the presence of variance terms and expectation constraints introduces non-linearities and couplings across time and scenarios. These complexities make direct analytical or numerical solutions difficult. To develop a computationally tractable approach, we adopt a scenario-based representation of the underlying uncertainty, which is common in multi-stage stochastic programming (see [51,52]).

We consider a finite set of N discrete scenarios for the random excess return vector $\boldsymbol{r}_t$, where $\boldsymbol{r}_t^i$ represents the i-th scenario (realization) of the excess return. These discrete scenarios can be generated either through random sampling from the empirical distribution or by directly utilizing historical return data. (Effective scenario generation is crucial for the model's real-world applicability. The N discrete scenarios (ξ_i) for asset returns can

be developed using various established methods, such as historical simulation, Monte Carlo simulation (e.g., based on statistical distributions like multivariate normal/Student's t-test, or time–series models like GARCH), or bootstrapping from historical data.) Given an investment policy $\boldsymbol{u}_t$, the corresponding discrete scenarios of the terminal wealth $x_{t+1}^i = r_{t+1}^f x_t + (\boldsymbol{r}_{t+1}^i)^\top \boldsymbol{u}_t$, for $i = 1, \cdots, N$.

We then focus on reformulating the SRM in a linear functional form. Similar to the one in [51,52], we discretize the interval $[0, 1]$ by taking $N+1$ points evenly as $s_i = \frac{i-1}{N}$ for $i = 1, \ldots, N+1$. For given spectrum function $\phi_t(s)$, we define

$$\phi_t^i := \int_{s_i}^{s_{i+1}} \phi_t(s) ds \approx \frac{1}{N} \phi_t(s_i)$$

for $i = 1, \ldots, N$ as the discretization of $\phi_t(\cdot)$, which also satisfies $\sum_{i=1}^N \phi_t^i = 1$. Following the similar method by [51], for given $\boldsymbol{u}_t$, the discretized SRM can be expressed as

$$\mathrm{SRM}_{\phi_t}^{(N)}(x_t) = -\sum_{i=1}^{N} \phi_{t,i} x_t^{i:N}, \tag{3}$$

Here, $x_t^{i:N}$ denotes the i-th smallest element of x_t^i for $i = 1, \ldots, N$ at time t.

In formulation (3), we need to sort the realizations of x_t^i for $i = 1, \ldots, N$ in ascending order. However, as the terminal wealth is a random variable affected by the decision variable $\boldsymbol{u}_t$, the order of $\{x_t^i\}|_{i=1}^N$ is also affected by the portfolio decision, which prevents us from using the formulation (3) directly [51]. Fortunately, this difficulty can be conquered by formulating the sorting procedure as an optimization problem with auxiliary variables, which is shown in Section 3.

3. Methodology

In this section, we apply the Progressive Hedging Algorithm (PHA) in combination with linkage objective and constraint reformulation techniques to solve the problem $(\mathcal{P})$, which incorporates the intertemporal SRMs.

3.1. Scenario-Based Decomposition and Progressive Hedging Algorithm

Let Ξ be the set of N empirical realizations, which is $\Xi = (\xi_1, \xi_2, \cdots, \xi_N)$, where ξ_i is the i-th realization. Let the objective function be $f(\boldsymbol{U}, \Xi)$, which is defined as

$$f(\boldsymbol{U}, \Xi) = \sum_{t=1}^{T} \mu_t \mathrm{SRM}_{\phi_t}(x_t) + \sum_{t=1}^{T} \kappa_t \mathrm{Var}(x_t),$$

where $\boldsymbol{U} \in \boldsymbol{R}^{T \times M} = (\boldsymbol{u}_0, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_{T-1})$. Let the feasible policy set be $\mathcal{A}$, which is defined as

$$\mathcal{A} = \{\boldsymbol{u_t} \in \mathcal{A}_t, \boldsymbol{E}[x_t] = d_t, \{x_t, \boldsymbol{u}_t\} \text{ satisfy } (1), \forall\, t \in \{0, \cdots, T-1\}\}.$$

Moreover, the multi-stage decisions should be *nonanticipative*; that is, decisions at each stage should be based only on information available up to that point in time, without knowledge of future random outcomes or events. Let $\mathcal{N}$ be the nonanticipativity action space. Hence, the problem $(\mathcal{P})$ can be rewritten as

$$\min_{\boldsymbol{U} \in \mathcal{A} \cap \mathcal{N}} f(\boldsymbol{U}, \Xi). \tag{4}$$

PHA operates by decomposing the original stochastic optimization problem into a set of deterministic subproblems, each corresponding to a specific scenario ξ. These subproblems are then solved iteratively, and their solutions are aggregated to converge

towards a globally optimal solution. Formally, the original problem is decomposed into N scenario-based subproblems, each of which can be expressed as:

$$\min_{\boldsymbol{U}\in\mathcal{A}} f(\boldsymbol{U},\xi). \tag{5}$$

To ensure convergence to a solution that is both feasible and optimal for the original stochastic problem, PHA incorporates penalty terms that progressively enforce consensus among scenario-specific decisions. Let $\boldsymbol{V}$ represent the dual variables (also referred to as the price system), where $\boldsymbol{V} \in \mathcal{M}$, and $\mathcal{M}$ is the orthogonal complement of $\mathcal{N}$. Consequently, the N scenario-based subproblems that PHA solves in each iteration can be formulated as:

$$\hat{\boldsymbol{U}}(\xi) = \arg\min_{\boldsymbol{U}\in\mathcal{A}} f(\boldsymbol{U},\xi) - \langle \boldsymbol{U}, \boldsymbol{V}(\xi)\rangle. \tag{6}$$

Following the resolution of these N scenario-based subproblems, PHA aggregates the obtained solutions $\hat{\boldsymbol{U}}(\xi)$ and updates both the primal and dual variables. The complete procedure for PHA is summarized in Algorithm 1. For further details and the corresponding proof, refer to [36].

Algorithm 1 Original progressive hedging.

Require: Initial primal–dual pair $(\boldsymbol{U}^0 \in \mathcal{N}, \boldsymbol{V}^0 \in \mathcal{M})$, penalty parameter $r > 0$
Ensure: Optimal solution $(\boldsymbol{U}^*, \boldsymbol{V}^*)$
Set iteration $k \leftarrow 0$
repeat
  // Scenario decomposition
  for all $\xi \in \Xi$ **do**
    $\hat{\boldsymbol{U}}^k(\xi) \leftarrow \arg\min_{\boldsymbol{U}\in\mathcal{A}}\{f(\boldsymbol{U},\xi) - \langle \boldsymbol{U}^k, \boldsymbol{V}^k(\xi)\rangle + \frac{r}{2}\|\boldsymbol{U} - \boldsymbol{U}^k(\xi)\|_F^2\}$
  end for
  // Update primal and dual variables
  $\boldsymbol{U}^{k+1}(\cdot) \leftarrow P_{\mathcal{N}}(\hat{\boldsymbol{U}}^k(\cdot))$ ▷ Project onto subspace $\mathcal{N}$
  $\boldsymbol{V}^{k+1}(\cdot) \leftarrow \boldsymbol{V}^k(\cdot) - rP_{\mathcal{M}}(\hat{\boldsymbol{U}}^k(\cdot))$ ▷ Project onto subspace $\mathcal{M}$
  $k \leftarrow k + 1$
until convergence
return solution $(\boldsymbol{U}^k, \boldsymbol{V}^k)$

The standard PHA framework outlined in Algorithm 1 is effective for problems where the objective function and constraints are scenario-separable. However, our problem $(\mathcal{P})$ includes terms like $\mathrm{SRM}_{\phi_t}(x_t)$ and $\mathrm{Var}(x_t)$, and the constraint $\boldsymbol{E}[x_t] = d_t$, which inherently link decisions across different scenarios. These "linkage" terms violate the separability assumption crucial for the direct application of PHA. Therefore, modifications are necessary to handle these interdependencies.

3.2. Reformulation of Linkage Objectives

We begin by reformulating intertemporal SRMs. Let the summation of intertemporal SRM be ISRM:

$$\mathrm{ISRM} = \sum_{t=1}^{T} \mu_t \mathrm{SRM}_{\phi_t}(x_t). \tag{7}$$

Let us introduce the function $F(\boldsymbol{U},\boldsymbol{B})$, which serves as a key component in our reformulation:

$$F(\boldsymbol{U},\boldsymbol{B}) = \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta_{t,i} + \frac{1}{1-\alpha_i} \int_{\xi} [-x_t(\xi) - \beta_{t,i}]^+ p(\xi) d\xi \right\} \right), \tag{8}$$

where ξ represents the scenario of random variables capturing the uncertainty in our model, $\alpha_i = \frac{i}{N}$ for $i \in \{1, \ldots, N\}$, and the coefficients $\lambda_{t,i}$ are defined as $\lambda_{t,i} = \phi_{t,i} - \phi_{t,i+1}$ for $i \in \{1, \ldots, N-1\}, \lambda_{t,N} = \phi_{t,N}$. $\boldsymbol{B} = (\boldsymbol{\beta}_1^N, \boldsymbol{\beta}_2^N, \ldots, \boldsymbol{\beta}_T^N) \in \boldsymbol{R}^{T \times N}$ is a matrix of auxiliary variables, with each $\boldsymbol{\beta}_t^N = (\beta_{t,1}, \beta_{t,2}, \ldots, \beta_{t,N}) \in \boldsymbol{R}^N$ representing a vector of these variables for a specific time period t. $[\cdot]^+$ represents the positive part function.

This formulation leverages the known representation of CVaR as the solution to an optimization problem [26] and the relationship between SRM and CVaR (as shown in Appendix A, Lemma A1). It allows us to express the intertemporal SRM in terms of a minimization problem over the auxiliary variables $\boldsymbol{B}$, as we will demonstrate in the subsequent theorem.

Theorem 1. *The ISRM can be expressed as the minimum of the function $F(\boldsymbol{U}, \boldsymbol{B})$ over the auxiliary variables $\boldsymbol{B}$:*

$$ISRM = \min_{\boldsymbol{B} \in \boldsymbol{R}^{T \times N}} F(\boldsymbol{U}, \boldsymbol{B}) \tag{9}$$

Minimize the ISRM associated with $\boldsymbol{U}$ over all $\boldsymbol{U} \in \mathcal{A}$ is equivalent to minimize $F(\boldsymbol{U}, \boldsymbol{B})$ over all $(\boldsymbol{U}, \boldsymbol{B}) \in \mathcal{A} \times \boldsymbol{R}^{T \times N}$:

$$\arg\min_{\boldsymbol{U} \in \mathcal{A}} ISRM = \arg\min_{(\boldsymbol{U}, \boldsymbol{B}) \in \mathcal{A} \times \boldsymbol{R}^{T \times N}} F(\boldsymbol{U}, \boldsymbol{B}). \tag{10}$$

For the optimal solution $\boldsymbol{B}^*$, each element $\beta_{t,i}^*$ is equal to the VaR at level α_i for the loss at time t:

$$\beta_{t,i}^* = VaR_{\alpha_i}(x_t), \quad \forall t \in \{1, \ldots, T\}, i \in \{1, \ldots, N\} \tag{11}$$

The detailed proof is provided in Appendix A. Theorem 1 establishes the equivalence between minimizing ISRM and minimizing the function $F(\boldsymbol{U}, \boldsymbol{B})$ in a continuous probability space. For practical implementation within our scenario-based framework, we need to work with the discrete empirical distribution derived from the scenarios. This leads to a discrete approximation of the function $F(\boldsymbol{U}, \boldsymbol{B})$.

Let us assume we have collected N such realizations. We can then formulate a discrete approximation of the ISRM by sampling from the probability distribution of ξ:

$$\text{ISRM} = \min_{\boldsymbol{B} \in \boldsymbol{R}^{T \times N}} \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta_{t,i} + \frac{N}{N-i} \boldsymbol{E}_{\xi} [-x_t(\xi) - \beta_{t,i}]^+ \right\} \right), \tag{12}$$

These discretized formulations offer computationally tractable approximations of the ISRM. Furthermore, these discretizations facilitate the decomposition of the objective function into scenario-based sub-objectives, rendering them amenable to the PHA framework.

Utilizing the function $F(\boldsymbol{U}, \boldsymbol{B})$, we can reformulate the original problem as:

$$\begin{aligned} \min_{\boldsymbol{U} \in \mathcal{A} \cap \mathcal{N}, \boldsymbol{B} \in \boldsymbol{R}^{T \times N}} f(\boldsymbol{U}, \boldsymbol{B}, \Xi) = & \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta_{t,i} + \frac{N}{N-i} \boldsymbol{E}_{\xi} [-x_t(\xi) - \beta_{t,i}]^+ \right\} \right) \\ & + \sum_{t=1}^{T} \kappa_t \left(\boldsymbol{E}_{\xi} [x_t(\xi) - d_t]^2 \right) \end{aligned} \tag{13}$$

which can be subsequently decomposed into scenario-based sub-objectives:

$$\begin{aligned} \min_{\boldsymbol{U} \in \mathcal{A}, \boldsymbol{B} \in \boldsymbol{R}^{T \times N}} f(\boldsymbol{U}, \boldsymbol{B}, \Xi) = & \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta_{t,i} + \frac{1}{N-i} [-x_t(\xi) - \beta_{t,i}]^+ \right\} \right) \\ & + \sum_{t=1}^{T} \frac{\kappa_t}{N} \left((x_t(\xi) - d_t)^2 \right) \end{aligned} \tag{14}$$

This decomposition property is crucial for the efficient application of PHA to our multi-period portfolio optimization problem $(\mathcal{P})$.

Despite this increased separability, a challenge persists regarding the requirement for $\boldsymbol{B}$ to remain consistent across all scenarios ξ. Specifically, $\boldsymbol{B}$ is subject to the nonanticipativity constraint. Hence, we treat it as an additional first-stage decision variable. Consequently, $\boldsymbol{B}$ must converge to a constant value to satisfy the nonanticipativity condition. The intricacies of projecting $\boldsymbol{B}$ onto the nonanticipativity subspace and updating its corresponding dual variables are delineated in Algorithm 2.

3.3. Reformulation of Linkage Constraints

Having successfully decomposed the linkage objective function into scenario-specific components through the reformulation utilizing $F(\boldsymbol{U}, \boldsymbol{B})$, we now address the linkage constraint $\boldsymbol{E}[x_t] = d_t$. This constraint is classified as a linkage constraint due to its incorporation of the expectation over all scenarios, thereby coupling decisions across different realizations. Such coupling precludes straightforward scenario-based decomposition.

To address the linkage constraint, one approach involves employing augmented Lagrangian relaxation techniques. This method entails introducing a Lagrange multiplier associated with the constraint and incorporating it into the objective function. Subsequently, the linkage constraint can be decomposed using the methodology presented in Section 3.2. However, the direct integration of PHA with Lagrangian methods results in a nested iterative process. The outer iteration involves updating the Lagrangian multiplier, while the inner iteration comprises the PHA update steps with a fixed Lagrangian multiplier. To ensure global convergence, it becomes necessary to iterate the PHA steps until convergence for each outer iteration, which is computationally inefficient. Although some research has attempted to consolidate the nested iterative process into a single iteration [43], the guarantee of convergence to optimality and the associated convergence rate remain challenging to theoretically establish.

To address the linkage constraint without resorting to Lagrangian multipliers, we propose an alternative approach that introduces auxiliary variables and modifies the nonanticipativity subspace. We begin by reformulating the linkage constraint as follows:

$$\boldsymbol{E}[x_t] = d_t \Longleftrightarrow d_t - \sum_{\xi} p(\xi) x_t(\boldsymbol{U}(\xi), \xi) = 0, \forall t \in \{1, \ldots, T\}. \tag{15}$$

This constraint can be equivalently expressed using auxiliary variables:

$$\exists \omega_t(\xi) : d_t - x_t(\boldsymbol{U}(\xi), \xi) = \omega_t(\xi) \forall \xi \in \Xi \text{and} \sum_{\xi} p(\xi) \omega_t(\xi) = 0, \tag{16}$$

where $\boldsymbol{\Omega} \in \boldsymbol{R}^{T \times N} = (\boldsymbol{\omega}(\xi_1), \ldots, \boldsymbol{\omega}(\xi_N))$ and $\boldsymbol{\omega}(\xi_i) \in \boldsymbol{R}^T = (\omega_1(\xi_i), \ldots, \omega_T(\xi_i))$ is a vector of auxiliary variables.

Therefore, the feasible set $\mathcal{A} \cap \mathcal{N}$ can be modified to $\tilde{\mathcal{A}} \cap \tilde{\mathcal{N}}$, which are defined as:

$$\begin{aligned} \tilde{\mathcal{A}} = \{ & \boldsymbol{u_t} \in \mathcal{A}_t, d_t - x_t(\boldsymbol{U}(\xi), \xi) = \omega_t(\xi) \\ & \forall \xi \in \Xi, \{x_t, \boldsymbol{u}_t\} \text{ satisfy } (1), \forall\, t \in \{0, \cdots, T-1\}\}, \end{aligned} \tag{17}$$

$$\tilde{\mathcal{N}} = \left\{ \boldsymbol{U} \in \mathcal{N}, \sum_{\xi} p(\xi) \omega_t(\xi) = 0, \forall\, t \in \{0, \cdots, T-1\} \right\}. \tag{18}$$

The auxiliary variables $\omega_t(\xi)$ represent the deviation of the scenario-specific wealth realization $x_t(\xi)$ from the target mean d_t. The constraint $\sum_{\xi} p(\xi)\omega_t(\xi) = 0$ ensures that the expected deviation is zero, thereby satisfying the original linkage constraint $\boldsymbol{E}[x_t] = d_t$.

This constraint on $\boldsymbol{\Omega}$ is handled within the PHA projection step. The projection onto the augmented nonanticipativity space is modified due to the introduction of the additional constraint $\sum_{\xi} p(\xi)\omega_t(\xi) = 0$. Let $\hat{\omega}_t^k(\xi)$ denote the optimal results obtained from all scenarios Ξ, and let $\tau_t^k(\xi)$ represent the associated dual variables. The projection of $\hat{\omega}_t^k(\xi)$ onto the augmented nonanticipativity space $\tilde{\mathcal{N}}$ is given by:

$$\omega_t^{k+1}(\xi) = \hat{\omega}_t^k(\xi) - \sum_{\xi} p(\xi)\hat{\omega}_t(\xi), \forall \xi \in \Xi, \forall t \in \{0, \cdots, T-1\}, \tag{19}$$

while the update of the dual variables is expressed as:

$$\tau_t^{k+1}(\xi) = \tau_t^k(\xi) - r\sum_{\xi} p(\xi)\hat{\omega}_t(\xi), \forall \xi \in \Xi, \forall t \in \{0, \cdots, T-1\}. \tag{20}$$

The integration of these modifications with PHA is delineated in Algorithm 2.

This reformulation allows us to decompose the linkage constraint into scenario-specific components while maintaining the overall expectation constraint through the auxiliary variables. By incorporating these auxiliary variables into the PHA framework, we can effectively handle the linkage constraint without introducing nested iterations, thereby improving computational efficiency.

3.4. Algorithm for Problem $(\mathcal{P})$

Having developed methods to reformulate both the linkage objectives (Section 3.2) and the linkage constraints (Section 3.3) into forms compatible with scenario decomposition, we can now integrate these techniques into the Progressive Hedging Algorithm. This leads to a modified PHA specifically designed to solve problem $(\mathcal{P})$ in this section. Specifically, we integrate the scenario-decomposition approach of the original PHA outlined in Section 3.1 with the techniques developed for addressing linkage objectives and constraints, as detailed in Sections 3.2 and 3.3, respectively. This integration results in a unified algorithmic framework tailored to efficiently solve our multi-period portfolio optimization problem with intertemporal SRMs.

To implement the modified PHA for problem $(\mathcal{P})$, we consider several sets of variables that require iterative calculation. The algorithm operates on three primal–dual pairs: $(\boldsymbol{U}, \boldsymbol{V})$, $(\boldsymbol{B}, \boldsymbol{P})$, and $(\boldsymbol{\Omega}, \boldsymbol{T})$. Here, $\boldsymbol{U}$ represents the original decision variables of the portfolio optimization problem, $\boldsymbol{B}$ denotes the auxiliary variables introduced to decompose the linkage objective, and $\boldsymbol{\Omega}$ comprises the auxiliary variables used to decompose the linkage constraint. Their respective dual variables are represented by $\boldsymbol{V}$, $\boldsymbol{P}$, and $\boldsymbol{T}$. For notational convenience, we define enlarged primal and dual variables as $\tilde{\boldsymbol{U}} = (\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\Omega})$ and $\tilde{\boldsymbol{V}} = (\boldsymbol{V}, \boldsymbol{P}, \boldsymbol{T})$, respectively.

Algorithm 2 presents a detailed exposition of the modified PHA tailored for problem $(\mathcal{P})$, which yields an optimal solution that respects both the nonanticipativity constraints and the linkage constraints of the original problem. The initial conditions for the dual variables $\boldsymbol{P}^0$ and $\boldsymbol{T}^0$ ensure they lie within the appropriate subspace orthogonal to the nonanticipativity constraints imposed on $\boldsymbol{B}$ and $\boldsymbol{\Omega}$, respectively.

Having presented the modified PHA (Algorithm 2) designed to solve problem $(\mathcal{P})$, we now turn to its theoretical validation. A crucial aspect of any iterative optimization algorithm is ensuring that it reliably converges to a correct solution. The following theorem formally establishes this convergence guarantee for our proposed method and characterizes the speed at which it approaches the optimum. The detailed proof is provided in Appendix B.

Algorithm 2 Modified Progressive Hedging Algorithm for problem $(\mathcal{P})$.

Require: Initial primal–dual pairs:
1: $(\boldsymbol{U}^0, \boldsymbol{V}^0)$, where $\boldsymbol{U}^0 \in \mathcal{N}$ and $\boldsymbol{V}^0 \in \mathcal{M}$;
2: $(\boldsymbol{B}^0, \boldsymbol{P}^0)$, where $\boldsymbol{B}^0 \in \boldsymbol{R}^{T\times Q} : \beta_{i,t}(\xi) = \beta_{i,t} \forall i,t$, where $\beta_{i,t}$ are constants and $\boldsymbol{P}^0 = \{\rho^0_{t,i}(\cdot) : \Xi \to \boldsymbol{R} \mid \boldsymbol{E}_\xi[\rho^0_{t,i}(\xi)] = 0, \forall t, i\}$;
3: $(\boldsymbol{\Omega}^0, \boldsymbol{T}^0)$, where $\boldsymbol{\Omega}^0(\cdot) : \Xi \to \boldsymbol{R}^T$ with $\boldsymbol{E}_\xi[\omega^0_t(\xi)] = 0, \forall t$ and $\boldsymbol{T}^0(\cdot) : \Xi \to \boldsymbol{R}$ with $\tau^0_t(\xi) = \tau_t \forall t, \xi$, where τ_t are constants;
4: penalty parameter $r > 0$.
Ensure: Optimal solution $(\boldsymbol{U}^*, \boldsymbol{V}^*)$
5: Set iteration $k \leftarrow 0$
6: **repeat**
7: // Scenario decomposition
8: **for all** $\xi \in \Xi$ **do**

$$(\hat{\boldsymbol{U}}^k(\xi), \hat{\boldsymbol{B}}^k(\xi), \hat{\boldsymbol{\omega}}^k(\xi)) \leftarrow \underset{(\boldsymbol{U},\boldsymbol{B},\boldsymbol{\omega})\in\tilde{\mathcal{A}}}{\arg\min} \Big\{ f(\boldsymbol{U}, \boldsymbol{B}, \xi) - \langle \boldsymbol{U}, \boldsymbol{V}^k(\xi)\rangle - \langle \boldsymbol{B}, \boldsymbol{P}^k(\xi)\rangle - \langle \boldsymbol{\omega}, \boldsymbol{\tau}^k(\xi)\rangle + \frac{r}{2}\Big(\|\boldsymbol{U} - \boldsymbol{U}^k(\xi)\|_F^2 + \|\boldsymbol{B} - \boldsymbol{B}^k(\xi)\|_F^2 + \|\boldsymbol{\omega} - \boldsymbol{\omega}^k(\xi))\|^2\Big)\Big\}$$

9: **end for**
10: // Update primal variables
11: $\boldsymbol{U}^{k+1}(\cdot) \leftarrow P_{\mathcal{N}}(\hat{\boldsymbol{U}}^k(\cdot))$
12: $\boldsymbol{B}^{k+1} \leftarrow \frac{1}{N}\sum_{i=1}^N \hat{\boldsymbol{B}}^k(\xi)$
13: $\omega^{k+1}_t(\xi) \leftarrow \hat{\omega}^k_t(\xi) - \sum_{\tilde{\xi}} p(\xi)\hat{\omega}^k_t(\xi), \forall \xi \in \Xi, t \in \{0, \ldots, T-1\}$
14: // Update dual variables
15: $\boldsymbol{V}^{k+1}(\cdot) \leftarrow \boldsymbol{V}^k(\cdot) - rP_{\mathcal{M}}(\hat{\boldsymbol{U}}^k(\cdot))$
16: $\boldsymbol{P}^{k+1}(\xi) \leftarrow \boldsymbol{P}^k(\xi) - r(\hat{\boldsymbol{B}}^k(\xi) - \boldsymbol{B}^{k+1}), \forall \xi \in \Xi$
17: $\tau^{k+1}_t(\xi) \leftarrow \tau^k_t(\xi) - r\sum_{\tilde{\xi}} p(\xi)\hat{\omega}^k_t(\xi), \forall \xi \in \Xi, t \in \{0, \ldots, T-1\}$
18: $k \leftarrow k + 1$
19: **until** convergence
20: **return** solution $(\tilde{\boldsymbol{U}}^k, \tilde{\boldsymbol{V}}^k)$

Theorem 2. *(Convergence guarantee and convergence rate) Let $\{(\tilde{\boldsymbol{U}}^k, \tilde{\boldsymbol{V}}^k)\}$ be the sequence generated by Algorithm 2 from any initial point $(\tilde{\boldsymbol{U}}^0, \tilde{\boldsymbol{V}}^0)$. The sequence converges to an optimal solution $(\tilde{\boldsymbol{U}}^*, \tilde{\boldsymbol{V}}^*)$ of problem $(\mathcal{P})$. The convergence occurs at a q-linear rate with respect to the norm $\|\cdot\|_r$ defined as:*

$$\|(\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}})\|_r = \sqrt{\|\tilde{\boldsymbol{U}}\|_F^2 + \frac{1}{r^2}\|\tilde{\boldsymbol{V}}\|_F^2}, \tag{21}$$

where $r > 0$ is the penalty parameter used in the algorithm. Specifically, there exists a constant $q \in [0, 1)$ such that for all $k \geq 0$:

$$\|(\tilde{\boldsymbol{U}}^{k+1} - \tilde{\boldsymbol{U}}^*, \tilde{\boldsymbol{V}}^{k+1} - \tilde{\boldsymbol{V}}^*)\|_r \leq q\|(\tilde{\boldsymbol{U}}^k - \tilde{\boldsymbol{U}}^*, \tilde{\boldsymbol{V}}^k - \tilde{\boldsymbol{V}}^*)\|_r, \tag{22}$$

with strict inequality holding unless $(\tilde{\boldsymbol{U}}^k, \tilde{\boldsymbol{V}}^k) = (\tilde{\boldsymbol{U}}^*, \tilde{\boldsymbol{V}}^*)$.

Remark 1. *The q-linear convergence rate implies that the distance to the optimal solution decreases by at least a fixed fraction in each iteration, which is a strong form of convergence. The strict inequality in the convergence rate, unless at the optimal point, ensures that the algorithm makes progress in each iteration until the optimal solution is reached.*

Theorem 2 provides a theoretical guarantee for the overall convergence of Algorithm 2. It is worth noting, however, that the practical speed of convergence, quantified by the q-linear rate q, is influenced by several parameters related to the problem formulation and the

algorithm itself. A crucial factor is the PHA penalty parameter r; selecting an appropriate value is key, as it must balance enhancing the strong convexity of the subproblems (favoring faster convergence) against potentially slowing the propagation of information through the dual variable updates.

Furthermore, the objective function weights, particularly the variance weight κ_t, impact the problem's conditioning. Larger values of κ_t strengthen the quadratic term in the objective, which can improve the strong convexity with respect to the allocation variables $\boldsymbol{U}$ and thus potentially decrease q (i.e., accelerate convergence). The SRM weights μ_t also contribute to the objective landscape. Additionally, the inherent structure, including dimensionality (T, M, N) and asset return characteristics, along with the nature of the feasible sets $\mathcal{A}_t$, affects the overall conditioning and convergence behavior. While convergence is guaranteed, practical implementation often involves tuning r to achieve efficient performance for a specific problem instance.

An important consideration related to the problem formulation is the choice of the spectral function $\phi_t(s)$. While Theorem 2 guarantees q-linear convergence for any valid spectral function (including step functions), the specific form of $\phi_t(s)$ can affect the practical convergence speed by altering the complexity of the scenario subproblems solved in Step 4 of Algorithm 2. Specifically:

- If $\phi_t(s)$ is a single-step function (corresponding to CVaR at a specific quantile level, e.g., [49,50]), the SRM reformulation $F(\boldsymbol{U}, \boldsymbol{B})$ requires only one auxiliary variable $\beta_{t,i}$ per time period t. This significantly simplifies the scenario subproblem, potentially leading to faster computation within each PHA iteration.
- If $\phi_t(s)$ is a step function with K steps($1 < K < N$), the reformulation involves K auxiliary variables $\beta_{t,i}$ per time t. The subproblem complexity is intermediate between the CVaR case and the general case analyzed.
- If $\phi_t(s)$ is a more general function approximated by an N-step function(as implicitly achieved in our discretization and reformulation), the subproblem involves N auxiliary variables $\beta_{t,i}$ per time t, representing the most complex case considered here.

Therefore, using simpler spectral functions (like step functions with fewer steps) can reduce the computational burden per iteration, potentially leading to faster overall convergence in practice. This presents a possible trade-off between the flexibility of the spectral risk measure used and the computational efficiency of the algorithm. While convergence is guaranteed, practical implementation often involves tuning r and considering the structure of $\phi_t(s)$ to achieve efficient performance for a specific problem instance.

Moreover, the practical performance of the algorithm heavily depends on the efficiency with which the scenario-specific subproblems (Step 4 in Algorithm 2) can be solved in each iteration. We now turn our attention to the structure of these subproblems and investigate possibilities for obtaining analytical or semi-analytical solutions under certain conditions.

While these convex subproblems can generally be solved numerically, deriving an analytical solution, even under simplifying assumptions, provides valuable insight into the structure of the problem and can potentially lead to faster implementations. We now consider the specific case where there are no explicit constraints (like non-negativity or budget limits) imposed on the portfolio allocation decisions $\boldsymbol{u}_t$ within the set $\mathcal{A}_t$. The following theorem presents the resulting analytical expression for the optimal solution $(\hat{\boldsymbol{U}}^k(\xi), \hat{\boldsymbol{B}}^k(\xi), \hat{\boldsymbol{\omega}}^k(\xi))$ of the scenario subproblem under this condition. The detailed proof is provided in Appendix C.

Theorem 3. *Suppose there are no restrictions on how assets can be allocated in the portfolio* $\hat{\boldsymbol{U}}^k(\xi)$*. In that case, the optimal solution* $(\hat{\boldsymbol{U}}^k(\xi), \hat{\boldsymbol{B}}^k(\xi), \hat{\boldsymbol{\omega}}^k(\xi))$ *for the scenario-based subproblem can be expressed as:*

$$\begin{cases} \hat{\boldsymbol{U}}^k(\xi) = \left(r\boldsymbol{I} + \sum_{t=1}^{T} \frac{2\kappa_t}{N} \sum_{\xi} D_{0:t}^{\top} D_{0:t}\right)^{-1} \Bigg(\boldsymbol{V}^k(\xi) + r\boldsymbol{U}^k(\xi) - \sum_{t=1}^{T} \frac{2\kappa_t}{N} \left(\sum_{\xi} (\gamma_0 x_0 - d_t) D_{0:t}\right) \\ \qquad\qquad - \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \frac{1}{N-i} \sum_{\xi} D_{0:t} \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} \right\} \right) \Bigg) \\ \hat{\boldsymbol{B}}^k(\xi) = \frac{1}{r}\left(\boldsymbol{P}^k(\xi) + r\boldsymbol{B}^k(\xi) - \Delta\boldsymbol{B}\right) \\ \hat{\boldsymbol{\omega}}^k(\xi) = \boldsymbol{\omega}^k(\xi) + \frac{1}{r}\boldsymbol{\tau}^k(\xi) \end{cases}$$

where $D_{0:t} = [D_0, D_1, \cdots, D_{T-1}]$*, with the i-th element defined as* $D_i = \prod_{j=i+1}^{t} r_j^f \cdot \boldsymbol{r}_{i+1}^{\top}$ *if* $i < t$*, and* $D_i = \mathbf{0}$ *otherwise.* $\Delta\boldsymbol{B}$ *is a matrix defined as:*

$$\Delta\boldsymbol{B} = \begin{bmatrix} \mu_1\lambda_{1,1}\left(1 - \frac{1}{N}\sum_{\xi} \mathbf{1}_{\{-x_1(\xi) > \beta_{1,0}\}}\right) & \cdots & \mu_T\lambda_{T,1}\left(1 - \frac{1}{N}\sum_{\xi} \mathbf{1}_{\{-x_T(\xi) > \beta_{T,0}\}}\right) \\ \vdots & \ddots & \vdots \\ \mu_1\lambda_{1,N}\left(1 - \frac{1}{1}\sum_{\xi} \mathbf{1}_{\{-x_1(\xi) > \beta_{1,N-1}\}}\right) & \cdots & \mu_T\lambda_{T,N}\left(1 - \frac{1}{1}\sum_{\xi} \mathbf{1}_{\{-x_T(\xi) > \beta_{T,N-1}\}}\right) \end{bmatrix}.$$

However, it is important to note that this solution is not explicitly solvable in a single step. The expressions for $\hat{\boldsymbol{U}}^k(\xi)$ and $\hat{\boldsymbol{B}}^k(\xi)$ exhibit a mutual dependence through the indicator terms $\mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}}$. This interdependence arises because the optimal portfolio allocation $\hat{\boldsymbol{U}}^k(\xi)$ affects the wealth $x_t(\xi)$, which in turn influences the optimal risk thresholds $\hat{\boldsymbol{B}}^k(\xi)$. Conversely, the risk thresholds $\beta_{t,i}$ determine which scenarios are considered in the tail risk calculations, thereby impacting the optimal portfolio allocation $\hat{\boldsymbol{U}}^k(\xi)$. This circular relationship between $\hat{\boldsymbol{U}}^k(\xi)$ and $\hat{\boldsymbol{B}}^k(\xi)$ creates a complex optimization landscape. It precludes the possibility of obtaining a closed-form solution through direct algebraic manipulation. Instead, this interdependence necessitates a more sophisticated approach to find the optimal solution.

Fortunately, the sequence $\{\beta_{t,i}\}_{i=0}^{N-1}$ is monotonically increasing for each t, a property that stems from $\beta_{t,i}$ representing the VaR at progressively higher confidence levels. This monotonicity allows us to employ an efficient binary search algorithm to determine the optimal values. Leveraging this property, we can characterize $\{\mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}}\}_{i=0}^{N-1}$ as a binary sequence: a contiguous series of ones followed by a contiguous series of zeros. Consequently, our task reduces to finding the transition point between these two series. This can be achieved by iteratively updating $\hat{\boldsymbol{U}}^k(\xi)$ and $\hat{\boldsymbol{B}}^k(\xi)$ while employing binary search to efficiently locate the transition point in the indicator sequence for each t. This approach significantly reduces the computational complexity compared to exhaustively evaluating all possible combinations of $\hat{\boldsymbol{U}}^k(\xi)$ and $\hat{\boldsymbol{B}}^k(\xi)$, thereby providing an efficient method to solve the scenario-based subproblem.

4. Numerical Results

To assess the effectiveness of our proposed approach, we conducted some numerical experiments using a multi-period portfolio optimization problem incorporating intertemporal spectral risk measures. The portfolio optimization problem is defined over a three-period time horizon ($T = 3$), with a risk-free rate of 1.05. The portfolio consists of two risky assets, characterized by risk–return pairs $[[1.3, 0.95], [1.2, 1]]$. This setup yields 4 scenarios per time period, resulting in a total of $4^3 = 64$ scenarios across the full horizon. We set targeted mean returns $d_t = 1.1^t$ for each time period $t \in \{1, 2, 3\}$.

For the spectral risk measure, we employed an exponential spectrum defined as $\phi_t(i) = (k_e \cdot e^{-k_e \cdot i})/(1 - e^{-k_e})$, where $k_e = 5$. This spectrum, illustrated in Figure 2, provides a nuanced representation of risk preferences across different probability levels, assigning greater weight to extreme losses. The function exhibits a high initial value for extreme losses and decreases exponentially as the loss magnitude diminishes, reflecting a risk-averse attitude. This configuration emphasizes severe, low-probability losses over more frequent, moderate ones. For comparison, we also plotted the CVaR spectrum with an 80% quantile, represented by a step function in Figure 2. While CVaR assigns uniform weight to the worst 20% of outcomes and zero to others, our exponential spectrum offers a more gradual transition. This approach facilitates a comprehensive risk assessment across all potential outcomes while maintaining an emphasis on tail risk.

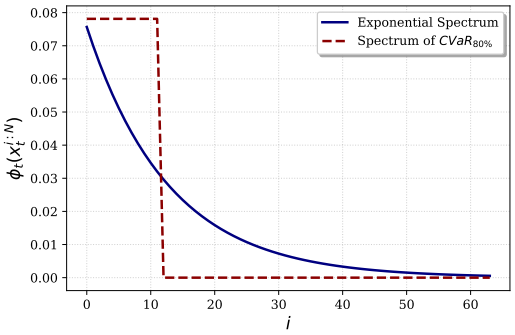


Figure 2. Comparison of exponential spectrum ($k_e = 5$) and spectrum of CVaR (80% quantile) used in the SRM. The exponential spectrum provides a more gradual weighting of risk across probability levels compared to the step function of CVaR.

We solved this optimization problem using the Progressive Hedging Algorithm, which was implemented with a step size of $r = 1$ for updating dual variables, a maximum of 200 iterations, and a convergence tolerance of 10^{-2}. Two distinct weighting schemes for the spectral risk measures and variance terms were explored to evaluate their impact on portfolio performance.

We compare the performance of two distinct weighting schemes: an intertemporal scheme that considers risk (SRM and variance) across all time periods and a terminal scheme that focuses solely on the risk at the final period. The first scheme, referred to as the "intertemporal weights" scheme, assigns equal importance to all time periods, setting $\mu_t = 1$ and $\kappa_t = 50$ for each $t \in \{1, 2, 3\}$. The second scheme, referred to as the "terminal weights" scheme, focuses exclusively on the final period by setting $\mu_T = 1$ and $\kappa_T = 50$, while nullifying the weights for intermediate periods ($\mu_t = \kappa_t = 0$ for $t \in \{1, 2\}$). The relatively large value of κ_t compared to μ_t was chosen to balance the scales between the variance and SRM terms in the objective function, as the variance term tends to be smaller in magnitude than the SRM term.

Figures 3 and 4 demonstrate the convergence behavior of the PHA for intertemporal weighting schemes. Figure 3 tracks several performance metrics—average wealth, variance, SRM, and objective function value—which all stabilize after approximately 100 iterations. More importantly, Figure 4 shows the monotonic decrease in the r-norm, which measures the combined distance of primal and dual variables from their optimal values. The strictly decreasing pattern of the r-norm confirms the q-linear convergence rate established in Theorem 2. The empirical evidence of both the metric stabilization and the r-norm's monotonic decrease validates the theoretical convergence properties of our modified PHA in solving the multi-period portfolio optimization problem with intertemporal risk measures.

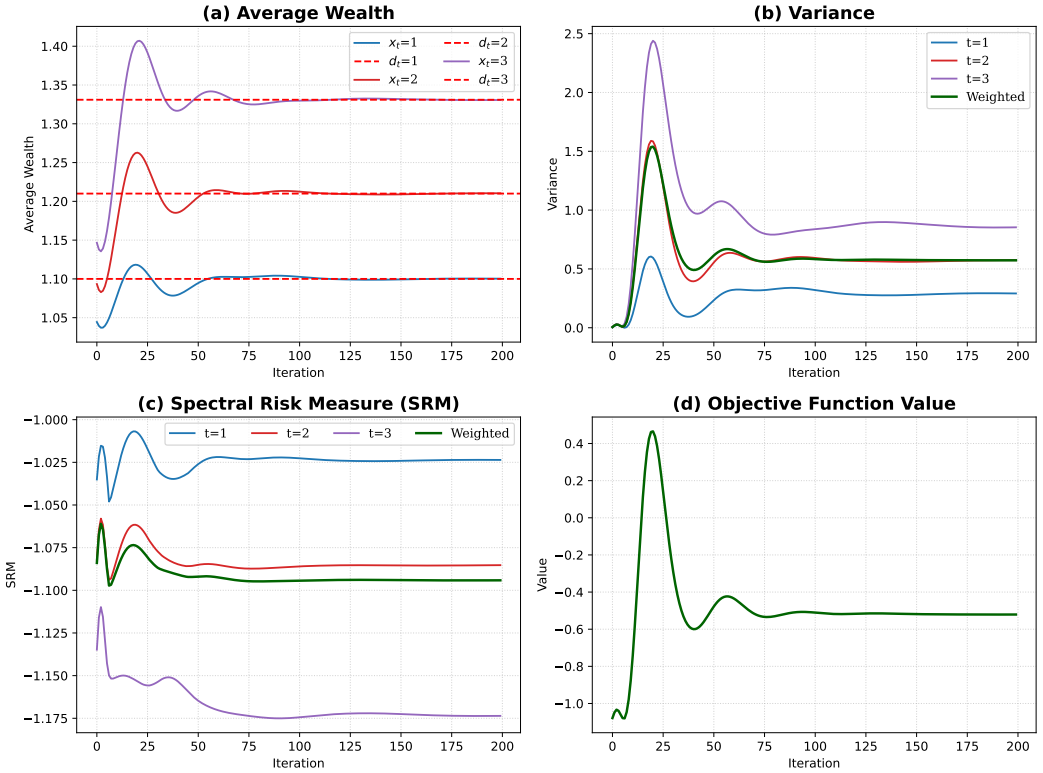


Figure 3. Convergence of average wealth, variance, SRM, and objective value over iterations for the terminal weighting scheme.

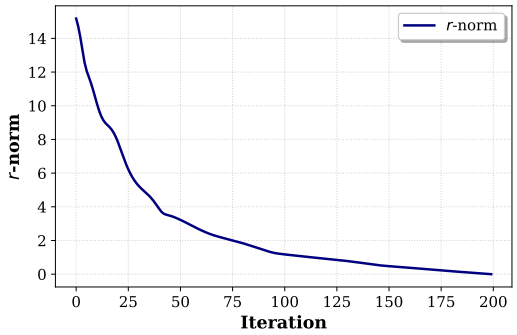


Figure 4. Convergence of the r-norm over iterations, demonstrating the q-linear convergence rate of the modified PHA. The r-norm measures the combined distance of primal and dual variables from their optimal values.

Figure 5 illustrates the wealth trajectories across all scenarios for both the intertemporal and terminal weighting schemes over the three-period investment horizon. The intertemporal scheme exhibits a tighter clustering of wealth trajectories, especially in the intermediate period ($t = 1, 2$). This tighter clustering suggests that the intertemporal approach provides more consistent wealth outcomes across different scenarios, indicating better risk management throughout the investment horizon. In contrast, the terminal weighting scheme shows a wider dispersion of wealth trajectories, particularly in the intermediate periods. This increased variability likely stems from the scheme's focus on optimizing only the final period's risk, allowing for greater fluctuations in earlier periods.

The observed tighter clustering of wealth trajectories under the intertemporal scheme (Figure 5), suggesting more consistent wealth outcomes, aligns with findings in other dynamic asset allocation studies, where strategies considering intermediate performance often lead to reduced path diversity and smoother wealth accumulation. For instance, [24]

noted that incorporating intermediate variance controls in a multi-period setting, albeit with different risk measures, helped stabilize portfolio performance across scenarios.

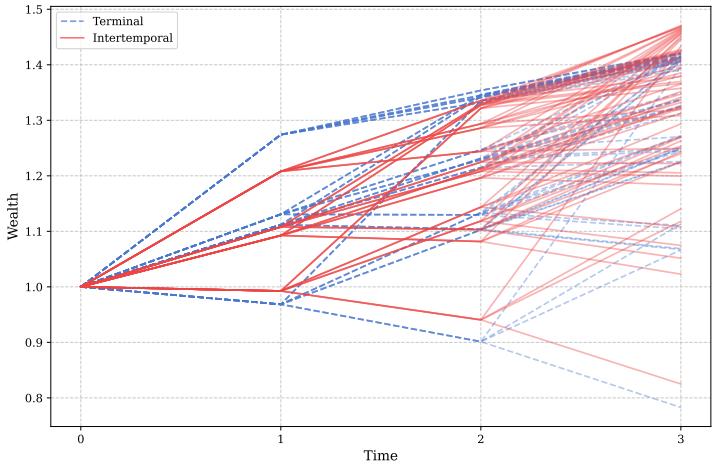


Figure 5. Wealth trajectories across all scenarios for intertemporal and terminal weighting schemes over the investment horizon. The intertemporal scheme exhibits tighter clustering of wealth outcomes, particularly in intermediate periods, indicating more consistent risk management throughout the investment horizon.

While the trajectory plots offer a useful visual comparison, a more rigorous evaluation requires examining quantitative performance and risk metrics. More measure metrics are shown in Table 1. The intertemporal weighting scheme outperforms the terminal scheme in terms of the overall objective function value. This suggests that considering risk across all time periods leads to a more optimal solution according to our defined criteria. The intertemporal approach also achieves better time-weighted SRM and time-weighted variance scores.

Looking at the time-specific metrics, we observe that the intertemporal scheme generally provides better risk management in the earlier periods ($t = 1, 2$) across various measures, including variance, SRM, skewness, kurtosis, and VaR. This is consistent with its design to consider risk throughout the investment horizon. The terminal scheme, however, shows slightly better performance in some metrics for the final period ($t = 3$), such as variance and SRM, aligning with its focus on terminal risk.

In terms of risk-adjusted return measures, we consider three widely used ratios: the Sharpe ratio, the Sortino ratio, and the Omega ratio. The Sharpe ratio measures excess return per unit of total risk, providing a general measure of risk-adjusted performance. The Sortino ratio is similar but focuses on downside risk, penalizing only returns below a target threshold. The Omega ratio offers a comprehensive view by considering the entire return distribution, weighing potential gains against losses. In our analysis, the intertemporal scheme demonstrates superior performance in Sortino and Omega ratios across all time periods, indicating better downside risk management. The Sharpe ratios are comparable, with the intertemporal scheme showing an advantage in earlier periods and the terminal scheme slightly outperforming in the final period.

Moreover, we extend the time horizon to 5 periods, resulting in $4^5 = 1024$ scenarios. Figure 6 presents the distribution of terminal wealth under both the intertemporal and terminal schemes. The terminal scheme exhibits a leptokurtic (high-kurtosis) distribution with heavy tails, indicating a higher probability of extreme outcomes that may be undesirable for risk-averse investors. In contrast, the intertemporal scheme yields a more balanced wealth distribution with moderate tails and lower kurtosis, which better aligns with typical investor preferences for stable and predictable investment outcomes.

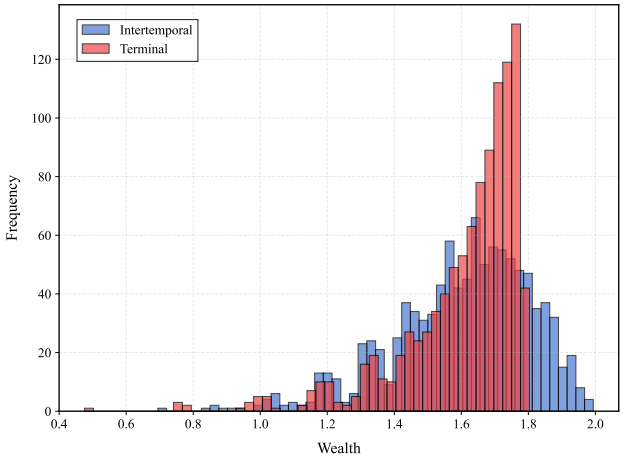


Figure 6. Distribution of terminal wealth under intertemporal and terminal schemes over a 5-period horizon. The terminal scheme exhibits a leptokurtic distribution with heavy tails, while the intertemporal scheme shows a more moderate distribution pattern.

Table 1. Comparison of intertemporal and terminal metrics. OBJ stands for objective function value, WSRM stands for weighted SRM, and WVar stands for weighted variance.

Metric	**OBJ**	**WSRM**	**WVar**	**Variance**			**SRM**		
				$t = 1$	$t = 2$	$t = 3$	$t = 1$	$t = 2$	$t = 3$
intertemporal	$-$**0.5209**	$-$**1.0941**	**0.5733**	**0.0058**	**0.0115**	0.0171	$-$**1.0236**	$-$**1.0853**	$-$1.1736
terminal	$-$0.3815	$-$1.0931	0.7116	0.0117	0.0155	**0.0154**	$-$1.0127	$-$1.0850	$-$**1.1817**
Metric	Skewness			Kurtosis			VaR		
	$t = 1$	$t = 2$	$t = 3$	$t = 1$	$t = 2$	$t = 3$	$t = 1$	$t = 2$	$t = 3$
intertemporal	**0.0000**	$-$**0.8335**	$-$**1.3775**	$-$**1.0205**	**0.1556**	**2.1985**	$-$**0.9926**	$-$**0.9618**	$-$**1.0796**
terminal	**0.0000**	$-$1.0297	$-$1.9513	$-$1.0151	0.5171	4.4911	$-$0.9685	$-$0.9315	$-$1.0742
Metric	Sharpe Ratio			Sortino Ratio			Omega Ratio		
	$t = 1$	$t = 2$	$t = 3$	$t = 1$	$t = 2$	$t = 3$	$t = 1$	$t = 2$	$t = 3$
intertemporal	**0.6585**	**1.4966**	2.1477	**1.7513**	**5.8666**	**9.9094**	**4.5027**	**24.4664**	**72.2525**
terminal	**0.6585**	1.4485	**2.2610**	1.7488	4.8562	8.4150	4.4975	20.4249	68.3198

Note: Bold values indicate better metrics.

Overall, these results suggest that the intertemporal weighting scheme provides more consistent risk management across the entire investment horizon, leading to generally better performance metrics, especially in earlier periods. The terminal scheme, while focusing on and slightly outperforming in some final period metrics, appears to sacrifice some risk management in earlier periods.

5. Conclusions

This paper presents a comprehensive framework for multi-period portfolio optimization incorporating intertemporal spectral risk measures. Our approach makes several significant contributions to the literature on dynamic portfolio management and risk measurement.

First, we introduced a novel multi-period portfolio optimization model that integrates SRMs with variance measures across all time periods. This formulation fills an important gap by explicitly considering both types of risk in an intertemporal fashion. Second, we propose an efficient solution methodology based on a modified Progressive Hedging Algorithm, enhanced with specialized reformulations for handling linkage objectives and constraints.Third, our research provides strong theoretical underpinnings and empirical

validation. We have established the q-linear convergence rate of our modified PHA under mild conditions, offering a solid theoretical foundation for its computational efficiency. Furthermore, comprehensive numerical experiments have validated the model's effectiveness, demonstrating that an intertemporal focus on risk leads to superior downside risk protection—as evidenced by improved Sortino and Omega ratios—and more balanced terminal wealth distributions compared to traditional terminal-focused strategies.

The framework and findings of this research carry several practical implications for portfolio managers, institutional investors, financial advisors, and, potentially, financial regulators. The proposed model serves as a practical tool for investors to implement more sophisticated dynamic risk-management strategies. By considering both tail risk and volatility at each period, managers can better navigate uncertain market conditions and tailor strategies to achieve more consistent performance throughout the investment horizon, rather than focusing solely on terminal outcomes.

The ability to specify time-varying risk preferences via the spectral function allows investment strategies to be precisely aligned with evolving institutional needs or individual investor life-cycle considerations, such as a pension fund gradually de-risking its portfolio as it approaches its liability horizon.

Critically, our numerical results strongly suggest that incorporating ISRMs leads to better protection against significant losses, which is crucial for risk-averse investors, endowments, and any entity with strict drawdown limits or capital preservation mandates. Furthermore, many investors are concerned with performance and risk exposure during intermediate periods; our intertemporal approach directly addresses this by optimizing risk–return characteristics throughout the investment journey, leading to smoother wealth paths and increased confidence in meeting ongoing obligations. This allows for more informed decision-making in the design and management of complex financial products like structured products or target-date funds that require a nuanced approach to risk evolution.

In essence, our approach empowers investment professionals with a flexible and robust methodology to proactively manage multifaceted risks across time, leading to potentially more stable and desirable investment outcomes.

Moreover, the framework developed in this paper opens several promising directions for future research. These include the incorporation of transaction costs and market frictions, and the extension to robust optimization frameworks, for instance, to handle uncertainty in the spectral function itself. A crucial area for future investigation involves a comprehensive analysis of the model's sensitivity to key assumptions—such as those related to market behavior and investor preferences—and its robustness under a wider range of market conditions. This could include evaluating alternative scenario generation methods that capture diverse market dynamics, and exploring the impact of varying investor preference parameters, like different spectral functions or risk aversion weights, or even extending to models where these preferences are endogenously determined or strategically designed [53]. Further research could also focus on adaptation to other complex market dynamics. Additionally, the methodological advances presented here may find applications beyond portfolio optimization, potentially benefiting other areas of financial risk management and dynamic decision-making under uncertainty.

In conclusion, our work provides both theoretical insights and practical tools for implementing dynamic risk-management strategies that effectively balance multiple objectives across time. The framework's flexibility and computational tractability make it particularly valuable for institutional investors and asset managers seeking to implement sophisticated risk-management strategies in practice.

Author Contributions: Conceptualization, C.J. and J.G.; Formal analysis, J.G.; Funding acquisition, J.G.; Investigation, C.J.; Methodology, C.J. and J.G.; Project administration, J.G.; Software, C.J.; Supervision, J.G.; Validation, C.J. and J.G.; Visualization, C.J.; Writing—original draft, C.J.; Writing—review and editing, C.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No. 71971132, 72192832) and by Shanghai Research Center for Data Science and Decision Technology. Additionally, it was partially funded by the InnoHK initiative of the Government of the HKSAR, and by the Laboratory for AI-Powered Financial Technologies and supported by the Key Laboratory of Interdisciplinary Research of Computation and Economics, Ministry of Education, Shanghai University of Finance and Economics.

Data Availability Statement: The original contributions presented in this study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author(s).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Proof of Theorem 1

This appendix section establishes the fundamental equivalence between intertemporal spectral risk measures (ISRMs) and their reformulation as a convex optimization problem through three key components. First, we demonstrate how discrete spectral risk measures can be represented as linear combinations of Conditional Value-at-Risk (CVaR) measures (Lemma A1). Second, we analyze the convexity and differentiability properties of the composite objective function $F(\boldsymbol{U}, \boldsymbol{B})$ (Lemma A2). Finally, we synthesize these results to prove the main theorem, showing that minimizing ISRM is equivalent to solving an optimized convex combination of CVaR constraints. This reformulation enables the decomposition of linkage objectives required for our Progressive Hedging Algorithm implementation.

Lemma A1. *Given N empirical realizations, the spectral risk measure (SRM) admits a representation as a linear combination of Conditional Value-at-Risk (CVaR) measures:*

$$SRM_{\phi_t}(x_t) = \sum_{i=1}^{N} \lambda_{t,i} CVaR_{\alpha_i}(x_t), \tag{A1}$$

where $\alpha_i = \frac{i}{N}$ for $i \in \{1, \ldots, N\}$, and the coefficients $\lambda_{t,i}$ are defined as: $\lambda_{t,i} = \phi_{t,i} - \phi_{t,i+1}$ for $i \in \{1, \ldots, N-1\}, \lambda_{t,N} = \phi_{t,N}$.

Proof of Lemma A1. Recall that given N empirical realizations, the discretized approximation of SRM is given by

$$\mathrm{SRM}_{\phi_t}^{(N)}(x_t) = -\sum_{i=1}^{N} \phi_i x_t^{i:N}. \tag{A2}$$

The ordered statistics $x_t^{i:N}$ can be related to the VaR by the identity $\mathrm{VaR}_{\frac{i}{N}}(x_t) = -x_t^{i:N}$. Consequently, the discretized SRM can be expressed as

$$\mathrm{SRM}_{\phi_t}(x_t) = \sum_{i=1}^{N} \phi_t \mathrm{VaR}_{\frac{i}{N}}(x_t). \tag{A3}$$

Moreover, it holds that $\mathrm{VaR}_{\frac{i}{N}}(x_t) = \mathrm{CVaR}_{\frac{i}{N}}(x_t) - \mathrm{CVaR}_{\frac{i-1}{N}}(x_t)$, for $i = 1, \ldots, N$ if we let $\mathrm{CVaR}_{\frac{0}{N}}(x_t) = 0$. Therefore, we can rewrite the SRM as

$$\mathrm{SRM}_{\phi_t}(x_t) = \sum_{i=1}^{N-1} (\phi_{t,i} - \phi_{t,i+1}) \mathrm{CVaR}_{\frac{i}{N}}(x_t) + \phi_{t,N} \mathrm{CVaR}_{\frac{N}{N}}(x_t). \tag{A4}$$

This expression demonstrates that the SRM can indeed be represented as a linear combination of CVaR measures, with coefficients $\lambda_{t,i} = \phi_{t,i} - \phi_{t,i+1}$ for $i \in \{1, \ldots, N-1\}$ and $\lambda_{t,N} = \phi_{t,N}$, thus completing the proof. □

Remark A1. *The formulation in equation (A4) aligns with the reformulation presented [51] when* $CVaR_{\frac{N}{N}}(x_t)$ *is replaced with* $-\frac{1}{N}\sum_{i=1}^{N} x_T^{i:N}$ *and employ the CVaR optimization framework from [26] for* $CVaR_{\frac{i}{N}}(x_T)$.

Lemma A2. *The function* $F(\boldsymbol{U}, \boldsymbol{B})$ *is convex and continuously differentiable. Its partial derivative with respect to* $\beta_{t,i}$ *is given by:*

$$\frac{\partial F(\boldsymbol{U}, \boldsymbol{B})}{\partial \beta_{t,i}} = \mu_t \lambda_{t,i} \frac{\Psi(-x_t, \beta_{t,i}) - \alpha_i}{1 - \alpha_i}, \tag{A5}$$

where $\Psi(-x_t, \beta_{t,i})$ *is the cumulative distribution function for the loss:*

$$\Psi(-x_t, \beta_{t,i}) = \int_{-x_t(\xi) \leq \beta_{t,i}} p(\xi) d\xi. \tag{A6}$$

Proof of Lemma A2. We first establish the convexity of $F(\boldsymbol{U}, \boldsymbol{B})$. This follows from the convexity of the positive part function $[\cdot]^+$ and the preservation of convexity under linear operations such as integration and summation. To prove continuous differentiability and derive the partial derivative, we focus on a single term in the sum corresponding to $\beta_{t,i}$:

$$f_{t,i}(\beta_{t,i}) = \beta_{t,i} + \frac{1}{1 - \alpha_i} \int_{\xi} [-x_t(\xi) - \beta_{t,i}]^+ p(\xi) d\xi.$$

Applying the Leibniz integral rule, we obtain:

$$\frac{\partial f_{t,i}}{\partial \beta_{t,i}} = 1 - \frac{1}{1 - \alpha_i} \int_{\xi} \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} p(\xi) d\xi,$$

where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function. Recognizing that the integral term represents the probability of the loss exceeding $\beta_{t,i}$, we can express this using the cumulative distribution function $\Psi(-x_t, \beta_{t,i})$ defined in (A6):

$$\frac{\partial f_{t,i}}{\partial \beta_{t,i}} = 1 - \frac{1 - \Psi(-x_t, \beta_{t,i})}{1 - \alpha_i} = \frac{\Psi(-x_t, \beta_{t,i}) - \alpha_i}{1 - \alpha_i}.$$

The derivative exists and is continuous for all $\beta_{t,i}$, as $\Psi(-x_t, \beta_{t,i})$ is a cumulative distribution function and thus continuous. The partial derivative of the full function $F(\boldsymbol{U}, \boldsymbol{B})$ with respect to $\beta_{t,i}$ is then obtained by multiplying by the corresponding weights:

$$\frac{\partial F(\boldsymbol{U}, \boldsymbol{B})}{\partial \beta_{t,i}} = \mu_t \lambda_{t,i} \frac{\Psi(-x_t, \beta_{t,i}) - \alpha_i}{1 - \alpha_i}.$$

This partial derivative is continuous in $\beta_{t,i}$ for all t and i, which proves that $F(\boldsymbol{U}, \boldsymbol{B})$ is continuously differentiable. □

Having established the necessary preliminary results in Lemmas A1 and A2, we now synthesize them to construct the proof for Theorem 1. The proof will proceed by minimizing the function $F(\boldsymbol{U}, \boldsymbol{B})$ with respect to the auxiliary variables $\boldsymbol{B}$, using the derivative calculated in Lemma A2, and demonstrating that the resulting minimum value precisely recovers the ISRM expression linked to CVaR in Lemma A1.

Proof. The proof leverages the properties of $F(\boldsymbol{U}, \boldsymbol{B})$ established in Lemma A2. We know that $F(\boldsymbol{U}, \boldsymbol{B})$ is convex and continuously differentiable with respect to $\boldsymbol{B}$. To find the minimum of $F(\boldsymbol{U}, \boldsymbol{B})$, we set the partial derivative of F for each t and q to zero:

$$\mu_t \lambda_{t,i} \frac{\Psi(-x_t, \beta_{t,i}) - \alpha_i}{1 - \alpha_i} = 0$$

Since $\mu_t > 0$, $\lambda_{t,i} > 0$, and $1 - \alpha_i > 0$, this equation simplifies to:

$$\Psi(-x_t, \beta_{t,i}) = \alpha_i$$

This condition is precisely the definition of the α_i-quantile of the loss distribution at time t. In other words, $\beta_{t,i}$ that satisfies this condition is the Value-at-Risk (VaR) at the α_i level for the loss at time t:

$$\beta^*_{t,i} = \mathrm{VaR}_{\alpha_i}(x_t).$$

Let $\boldsymbol{B}^*$ be the matrix of optimal $\beta^*_{t,i}$ values. We can then express:

$$\begin{aligned}
\min_{\boldsymbol{B}} F(\boldsymbol{U}, \boldsymbol{B}) &= F(\boldsymbol{U}, \boldsymbol{B}^*) \\
&= \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta^*_{t,i} + \frac{1}{1-\alpha_i} \int_{\xi} \left[-x_t(\xi) - \beta^*_{t,i} \right]^+ p(\xi) d\xi \right\} \right) \\
&= \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta^*_{t,i} + \frac{1}{1-\alpha_i} \int_{-x_t(\xi) \le \beta_{t,i}} \left[-x_t(\xi) - \beta^*_{t,i} \right] p(\xi) d\xi \right\} \right) \\
&= \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta^*_{t,i} + \frac{1}{1-\alpha_i} \left[\int_{-x_t(\xi) \le \beta_{t,i}} -x_t(\xi) p(\xi) d\xi - \beta^*_{t,i} \int_{-x_t(\xi) \le \beta_{t,i}} p(\xi) d\xi \right] \right\} \right) \\
&= \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta^*_{t,i} + \frac{1}{1-\alpha_i} \left[(1-\alpha_i) \mathrm{CVaR}_{\alpha_i}(x_t) - \beta^*_{t,i}(1-\alpha_i) \right] \right\} \right) \\
&= \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \mathrm{CVaR}_{\alpha_i}(x_t) \right) = \mathrm{ISRM}.
\end{aligned}$$

This proves that the minimum value of $F(\boldsymbol{U}, \boldsymbol{B})$ is indeed equal to ISRM, and that the optimal $\boldsymbol{B}$ corresponds to the VaR values at each time and quantile level. Furthermore, it demonstrates that:

$$\arg\min_{\boldsymbol{U} \in \mathcal{A}} \mathrm{ISRM} = \arg\min_{(\boldsymbol{U}, \boldsymbol{B}) \in \mathcal{A} \times \boldsymbol{R}^{T \times N}} F(\boldsymbol{U}, \boldsymbol{B}).$$

as the minimizing $\boldsymbol{B}$ in both cases is $\boldsymbol{B}^*$, where each $\beta^*_{t,i} = \mathrm{VaR}_{\alpha_i}(x_t)$. □

Appendix B. Proof of Theorem 2

The proof leverages the standard convergence theory for the PHA, as detailed by [36]. We demonstrate that our modified algorithm (Algorithm 2) fits within this framework. Specifically, we verify that the reformulations using auxiliary variables $(\boldsymbol{B}, \boldsymbol{\Omega})$ correctly handle the linkage terms, preserving the necessary convexity and decomposability properties. We also confirm that the primal–dual updates and projections onto the augmented spaces are consistent with PHA principles. Satisfying these conditions allows us to directly apply the established PHA convergence results, guaranteeing q-linear convergence for the sequence $\{(\tilde{\boldsymbol{U}}^k, \tilde{\boldsymbol{V}}^k)\}$.

Proof. The proof relies on the classical convergence theorem for PHA as presented in Section 5 of [36]. We demonstrate that our problem, after the reformulations detailed in Sections 3.1–3.3, satisfies the conditions required for this theorem to apply.

First of all, the problem is defined over a finite set of scenarios Ξ, which is a key assumption in the PHA convergence theorem. The reformulation of the problem with linkage objective and constraints, as shown in Sections 3.2 and 3.3, transforms $(\mathcal{P})$ into a separable minimization problem. This allows for the application of the PHA convergence theorem. The objective function $f(\boldsymbol{U}, \boldsymbol{B}, \xi)$ is convex and piecewise linear for all ξ. This follows from the properties of the CVaR function and the linear nature of the portfolio dynamics. The problem $(\mathcal{P})$ is assumed to be feasible, and the feasibility set $\tilde{\mathcal{A}} \cap \tilde{\mathcal{N}}$ is convex due to the linear nature of the constraints and the convexity-preserving operations used in the reformulation.

Given that these conditions are satisfied, the convergence theorem from [36] directly applies to our modified PHA (Algorithm 2). This ensures convergence to an optimal solution with a q-linear rate relative to the $\|\cdot\|_r$ norm. For a complete proof of these convergence properties, we refer the reader to [36]. □

Appendix C. Proof of Theorem 3

This proof derives the analytical expressions for the optimal solution $(\hat{\boldsymbol{U}}^k(\xi), \hat{\boldsymbol{B}}^k(\xi), \hat{\boldsymbol{\omega}}^k(\xi))$ to the scenario-specific subproblem solved in each iteration of Algorithm 2. The derivation focuses on the case where there are no explicit constraints on the portfolio allocation vector $\boldsymbol{u}_t$ (i.e., $\mathcal{A}_t = \mathbb{R}^M$). We apply first-order optimality conditions by calculating the partial derivatives of the subproblem's augmented Lagrangian objective function with respect to each block of variables ($\boldsymbol{U}$, $\boldsymbol{B}$, and $\boldsymbol{\omega}$), setting these derivatives to zero, and solving the resulting system of equations.

Proof. The Lagrangian objective function $L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega})$ for the scenario-based subproblem is defined as:

$$\begin{aligned} L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega}) = f(\boldsymbol{U}, \boldsymbol{B}, \xi) - \langle \boldsymbol{U}^k, \boldsymbol{V}^k(\xi) \rangle - \langle \boldsymbol{B}, \boldsymbol{P}^k(\xi) \rangle - \langle \boldsymbol{\omega}^k(\xi), \boldsymbol{\tau}^k(\xi) \rangle \\ + \frac{r}{2} \Big(\|\boldsymbol{U} - \boldsymbol{U}^k(\xi)\|_F^2 + \|\boldsymbol{B} - \boldsymbol{B}^k(\xi)\|_F^2 + \|\boldsymbol{\omega} - \boldsymbol{\omega}^k(\xi))\|^2 \Big) \end{aligned}$$

Substituting the definition of $f(\boldsymbol{U}, \boldsymbol{B}, \xi)$ from equation (14), we obtain:

$$\begin{aligned} L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega}) = \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \beta_{t,i} + \frac{1}{N-i} \sum_{\xi} [-x_t(\xi) - \beta_{t,i}]^+ \right\} \right) + \sum_{t=1}^{T} \frac{\kappa_t}{N} \left(\sum_{\xi} (x_t(\xi) - d_t)^2 \right) \\ - \langle \boldsymbol{U}^k, \boldsymbol{V}^k(\xi) \rangle - \langle \boldsymbol{B}^k, \boldsymbol{P}^k(\xi) \rangle - \langle \boldsymbol{\omega}^k(\xi), \boldsymbol{\tau}^k(\xi) \rangle \\ + \frac{r}{2} \Big(\|\boldsymbol{U} - \boldsymbol{U}^k(\xi)\|_F^2 + \|\boldsymbol{B} - \boldsymbol{B}^k(\xi)\|_F^2 + \|\boldsymbol{\omega} - \boldsymbol{\omega}^k(\xi))\|^2 \Big) \end{aligned}$$

To determine the optimal solution $(\hat{\boldsymbol{U}}^k(\xi), \hat{\boldsymbol{B}}^k(\xi), \hat{\boldsymbol{\omega}}^k(\xi))$, we compute the partial derivatives of $L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega})$ with respect to $\boldsymbol{U}$, $\boldsymbol{B}$, and $\boldsymbol{\omega}$, and set them to zero. The partial derivative with respect to $\boldsymbol{U}$ is given by:

$$\begin{aligned} \frac{\partial L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega})}{\partial \boldsymbol{U}} = \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \frac{1}{N-i} \sum_{\xi} \left[-\frac{\partial x_t(\xi)}{\partial \boldsymbol{U}} \right] \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} \right\} \right) \\ + \sum_{t=1}^{T} \frac{2\kappa_t}{N} \left(\sum_{\xi} (x_t(\xi) - d_t) \frac{\partial x_t(\xi)}{\partial \boldsymbol{U}} \right) - \boldsymbol{V}^k(\xi) + r(\boldsymbol{U} - \boldsymbol{U}^k(\xi)) \end{aligned}$$

The derivative $\frac{\partial x_t(\xi)}{\partial \boldsymbol{U}}$ can be expressed as:

$$\begin{aligned}\frac{\partial x_t(\xi)}{\partial \boldsymbol{U}} &= \left[\frac{\partial x_t(\xi)}{\partial \boldsymbol{u}_0}, \quad \frac{\partial x_t(\xi)}{\partial \boldsymbol{u}_1}, \quad \cdots, \quad \frac{\partial x_t(\xi)}{\partial \boldsymbol{u}_{t-1}}, \quad \mathbf{0}, \quad \cdots, \quad \mathbf{0}\right] \\ &= \left[\prod_{j=1}^{t} r_j^f \cdot \boldsymbol{r}_1^\top, \quad \prod_{j=2}^{t} r_j^f \cdot \boldsymbol{r}_2^\top, \quad \cdots, \quad r_t^f \cdot \boldsymbol{r}_t^\top, \quad \mathbf{0}, \quad \cdots, \quad \mathbf{0}\right]\end{aligned}$$

Let $D_{0:t} = [D_0, D_1, \cdots, D_{T-1}]$, where the i-th element is defined as $D_i = \prod_{j=i+1}^{t} r_j^f \cdot \boldsymbol{r}_{i+1}^\top$ if $i < t$, and $D_i = \mathbf{0}$ otherwise. Consequently, we have $\frac{\partial x_t(\xi)}{\partial \boldsymbol{U}} = D_{0:t}$. Substituting this into the partial derivative expression yields:

$$\begin{aligned}\frac{\partial L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega})}{\partial \boldsymbol{U}} =& \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \frac{1}{N-i} \sum_{\xi} [-D_{0:t}] \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} \right\} \right) \\ &+ \sum_{t=1}^{T} \frac{2\kappa_t}{N} \left(\sum_{\xi} (x_t(\xi) - d_t) D_{0:t} \right) - \boldsymbol{V}^k(\xi) + r(\boldsymbol{U} - \boldsymbol{U}^k(\xi))\end{aligned}$$

Given that $x_t(\xi) = \langle D_{0:t}, \boldsymbol{U} \rangle + \gamma_0 x_0$, we can further simplify the expression:

$$\begin{aligned}\frac{\partial L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega})}{\partial \boldsymbol{U}} =& \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \frac{1}{N-i} \sum_{\xi} [-D_{0:t}] \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} \right\} \right) \\ &+ \sum_{t=1}^{T} \frac{2\kappa_t}{N} \left(\sum_{\xi} (\langle D_{0:t}, \boldsymbol{U} \rangle + \gamma_0 x_0 - d_t) D_{0:t} \right) - \boldsymbol{V}^k(\xi) + r(\boldsymbol{U} - \boldsymbol{U}^k(\xi))\end{aligned}$$

Setting $\frac{\partial L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega})}{\partial \boldsymbol{U}} = 0$ and solving for $\boldsymbol{U}$, we obtain:

$$\begin{aligned}\left(r\boldsymbol{I} + \sum_{t=1}^{T} \frac{2\kappa_t}{N} \sum_{\xi} D_{0:t}^\top D_{0:t} \right) \boldsymbol{U} =& \boldsymbol{V}^k(\xi) + r\boldsymbol{U}^k(\xi) - \sum_{t=1}^{T} \frac{2\kappa_t}{N} \left(\sum_{\xi} (\gamma_0 x_0 - d_t) D_{0:t} \right) \\ &- \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \frac{1}{N-i} \sum_{\xi} D_{0:t} \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} \right\} \right)\end{aligned}$$

The matrix $\left(r\boldsymbol{I} + \sum_{t=1}^{T} \frac{2\kappa_t}{N} \sum_{\xi} D_{0:t}^\top D_{0:t} \right)$ is invertible, as it is a positive definite matrix. This follows from the fact that $r\boldsymbol{I}$ is positive definite for $r > 0$, and the sum of positive semidefinite matrices $D_{0:t}^\top D_{0:t}$ is positive semidefinite. The sum of a positive definite matrix and a positive semidefinite matrix is positive definite, thus invertible. Consequently, we can express the optimal solution $\hat{\boldsymbol{U}}^k(\xi)$ as:

$$\begin{aligned}\hat{\boldsymbol{U}}^k(\xi) =& \left(r\boldsymbol{I} + \sum_{t=1}^{T} \frac{2\kappa_t}{N} \sum_{\xi} D_{0:t}^\top D_{0:t} \right)^{-1} \left(\boldsymbol{V}^k(\xi) + r\boldsymbol{U}^k(\xi) - \sum_{t=1}^{T} \frac{2\kappa_t}{N} \left(\sum_{\xi} (\gamma_0 x_0 - d_t) D_{0:t} \right) \right. \\ &\left. - \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \frac{1}{N-i} \sum_{\xi} D_{0:t} \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} \right\} \right) \right)\end{aligned}$$

The partial derivative of the Lagrangian with respect to $\beta_{t,i}$ is given by:

$$\frac{\partial L(\boldsymbol{U}, \boldsymbol{B}, \boldsymbol{\omega})}{\partial \beta_{t,i}} = \mu_t \lambda_{t,i} \left(1 - \frac{1}{N-i} \sum_{\xi} \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} \right) - P_{t,i}^k(\xi) + r(\beta_{t,i} - B_{t,i}^k(\xi))$$

Consequently, the partial derivative with respect to $\boldsymbol{B}$ can be expressed as:

$$\frac{\partial L(\boldsymbol{U},\boldsymbol{B},\boldsymbol{\omega})}{\partial \boldsymbol{B}} = \Delta \boldsymbol{B} - \boldsymbol{P}^k(\xi) + r(\boldsymbol{B} - \boldsymbol{B}^k(\xi))$$

where $\Delta \boldsymbol{B}$ is a matrix defined as:

$$\Delta \boldsymbol{B} = \begin{bmatrix} \mu_1 \lambda_{1,1}\left(1 - \frac{1}{N}\sum_{\xi} \mathbf{1}_{\{-x_1(\xi) > \beta_{1,0}\}}\right) & \cdots & \mu_T \lambda_{T,1}\left(1 - \frac{1}{N}\sum_{\xi} \mathbf{1}_{\{-x_T(\xi) > \beta_{T,0}\}}\right) \\ \vdots & \ddots & \vdots \\ \mu_1 \lambda_{1,N}\left(1 - \frac{1}{1}\sum_{\xi} \mathbf{1}_{\{-x_1(\xi) > \beta_{1,N-1}\}}\right) & \cdots & \mu_T \lambda_{T,N}\left(1 - \frac{1}{1}\sum_{\xi} \mathbf{1}_{\{-x_T(\xi) > \beta_{T,N-1}\}}\right) \end{bmatrix}$$

Setting $\frac{\partial L(\boldsymbol{U},\boldsymbol{B},\boldsymbol{\omega})}{\partial \boldsymbol{B}} = 0$ and solving for $\hat{\boldsymbol{B}}^k(\xi)$ yields:

$$\hat{\boldsymbol{B}}^k(\xi) = \frac{1}{r}\left(\boldsymbol{P}^k(\xi) + r\boldsymbol{B}^k(\xi) - \Delta \boldsymbol{B}\right)$$

Analogously, the partial derivative with respect to $\boldsymbol{\omega}$ is:

$$\frac{\partial L(\boldsymbol{U},\boldsymbol{B},\boldsymbol{\omega})}{\partial \boldsymbol{\omega}} = -\boldsymbol{\tau}^k(\xi) + r(\boldsymbol{\omega} - \boldsymbol{\omega}^k(\xi))$$

Setting this equal to zero and solving for $\hat{\boldsymbol{\omega}}^k(\xi)$ results in:

$$\hat{\boldsymbol{\omega}}^k(\xi) = \boldsymbol{\omega}^k(\xi) + \frac{1}{r}\boldsymbol{\tau}^k(\xi)$$

Consequently, the optimal solution $(\hat{\boldsymbol{U}}^k(\xi), \hat{\boldsymbol{B}}^k(\xi), \hat{\boldsymbol{\omega}}^k(\xi))$ for the scenario-based subproblem can be expressed as:

$$\begin{cases} \hat{\boldsymbol{U}}^k(\xi) = \left(r\boldsymbol{I} + \sum_{t=1}^{T} \frac{2\kappa_t}{N} \sum_{\xi} D_{0:t}^{\top} D_{0:t}\right)^{-1} \Bigg(\boldsymbol{V}^k(\xi) + r\boldsymbol{U}^k(\xi) - \sum_{t=1}^{T} \frac{2\kappa_t}{N} \left(\sum_{\xi} (\gamma_0 x_0 - d_t) D_{0:t}\right) \\ \qquad\qquad - \sum_{t=1}^{T} \mu_t \left(\sum_{i=1}^{N} \lambda_{t,i} \left\{ \frac{1}{N-i} \sum_{\xi} D_{0:t} \mathbf{1}_{\{-x_t(\xi) > \beta_{t,i}\}} \right\}\right) \Bigg) \\ \hat{\boldsymbol{B}}^k(\xi) = \frac{1}{r}\left(\boldsymbol{P}^k(\xi) + r\boldsymbol{B}^k(\xi) - \Delta \boldsymbol{B}\right) \\ \hat{\boldsymbol{\omega}}^k(\xi) = \boldsymbol{\omega}^k(\xi) + \frac{1}{r}\boldsymbol{\tau}^k(\xi) \end{cases}$$

□

References

1. Markowitz, H.M. Portfolio Selection. *J. Financ.* **1952**, *7*, 1063–1070.
2. Artzner, P.; Delbaen, F.; Eber, J.M.; Heath, D. Coherent measure of risk. *Math. Financ.* **1999**, *9*, 203–228. [CrossRef]
3. Föllmer, H.; Schied, A. Convex measures od risk and trading contraints. *Financ. Stoch.* **2002**, *6*, 429–447. [CrossRef]
4. Kolm, P.N.; Tütüncü, R.; Fabozzi, F.J. 60 years of portfolio optimization: Practical challenges and current trends. *Eur. J. Oper. Res.* **2014**, *234*, 356–371. [CrossRef]
5. Gao, J.J.; Zhou, K.; Li, D.; Cao, X.R. Dynamic mean-LPM and mean-CVaR portfolio optimization in continuous-time. *SIAM J. Control Optim.* **2017**, *55*, 1377–1397. [CrossRef]
6. Zhou, K.; Gao, J.J.; Li, D.; Cui, X.Y. Dynamic mean-VaR portfolio selection in continuous time. *Quant. Financ.* **2017**, *17*, 1631–1643. [CrossRef]
7. He, X.D.; Zhou, X.Y. Dynamic portfolio choice when risk is measured by weighted VaR. *Math. Oper. Res.* **2015**, *40*, 773–796. [CrossRef]
8. Ortobelli, S.; Rachev, S.; Stoyanov, S.; Fabozzi, F.; Biglova, A. The proper use of risk measures in portfolio theory. *Int. J. Theor. Appl. Financ.* **2008**, *8*, 1107–1133. [CrossRef]
9. Adam, A.; Houkari, M.; Laurent, J.P. Spectral risk measures and portfolio selection. *J. Bank. Financ.* **2008**, *32*, 1870–1882. [CrossRef]

10. Steuer, R.E.; Na, P. Multiple criteria decison making combined with finance: A categorized bibliographic study. *Eur. J. Oper. Res.* **2003**, *150*, 496–515. [CrossRef]
11. Roman, D.; Mitra, G. Portfolio selection models: A review and new directions. *Wilmott J.* **2009**, *1*, 69–85. [CrossRef]
12. Wong, K.C.; Yam, S.C.P.; Zeng, J. Mean-risk portfolio management with bankruptcy prohibition. *Insur. Math. Econ.* **2019**, *85*, 153–172. [CrossRef]
13. Liu, H.Y.; Mao, T.T. Distributionally robust reinsurance with Value-at-Risk and Conditional Value-at-Risk. *Insur. Math. Econ.* **2022**, *107*, 393–417. [CrossRef]
14. Mi, H.; Xu, Z.Q. Optimal portfolio selection with VaR and portfolio insurance constraints under rank-dependent expected utility theory. *Insur. Math. Econ.* **2023**, *110*, 82–105. [CrossRef]
15. Gao, J.J.; Xiong, Y.; Li, D. Dynamic mean-risk portfolio selection with multiple risk measures in continuous-Time. *Eur. J. Oper. Res.* **2016**, *249*, 647–656. [CrossRef]
16. Roman, D.; Darby-Dowman, K.; Mitra, G. Mean-risk models using two risk measures: A multi-objective approach. *Quant. Financ.* **2007**, *7*, 443–458. [CrossRef]
17. Cesarone, F.; Martino, M.L.; Tardella, F. Mean-Variance-VaR portfolios: MIQP formulation and performance analysis. *Spectr.* **2023**, *45*, 1043–1069. [CrossRef]
18. Acerbi, C. Spectral measures of risk: A coherent representation of subjective risk aversion. *J. Bank. Financ.* **2002**, *26*, 1505–1518. [CrossRef]
19. Cai, J.; Li, J.Y.M.; Mao, T. Distributionally robust optimization under distorted expectations. *Oper. Res.* **2025**, *73*, 969–985. [CrossRef]
20. van Staden, P.M.; Dang, D.M.; Forsyth, P.A. On the distribution of terminal wealth under dynamic mean-variance optimal investment strategies. *SIAM J. Financ. Math.* **2021**, *12*. [CrossRef]
21. Costa, O.L.V.; Nabholz, R.B. Multiperiod Mean-Variance Optimization with Intertemporal Restrictions. *J. Optim. Theory Appl.* **2007**, *134*, 257–274. [CrossRef]
22. Li, D.; Ng, W.L. Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Math. Financ.* **2000**, *10*, 387–406. [CrossRef]
23. Cui, X.Y.; Li, X.; Li, D. Unified Framework of Mean-Field Formulations for Optimal Multi-Period Mean-Variance Portfolio Selection. *IEEE Trans. Autom. Control* **2014**, *59*, 1833–1844. [CrossRef]
24. Wu, W.; Yu, D.; Wang, T.; Gao, J. Multiperiod mean-variance portfolio selection with intertemporal restrictions and correlated returns. In Proceedings of the 2018 Chinese Control and Decision Conference (CCDC), Shenyang, China, 9–11 June 2018 ; IEEE: Piscataway, NJ, USA, 2018; pp. 2347–2352.
25. Pan, Q.; Lin, Y.; Wu, J.; Zhong, J. Intertemporal constrained portfolio optimization problem with tracking-error. *Appl. Math. Sci. Eng.* **2025**, *33*, 2463456. [CrossRef]
26. Rockafellar, R.T.; Uryasev, S. Optimization of conditional value-at-risk. *J. Risk* **2000**, *2*, 21–41. [CrossRef]
27. Utz, S.; Wimmer, M.; Hirschberger, M.; Steuer, R.E. Tri-criterion inverse portfolio optimization with application to socially responsible mutual funds. *Eur. J. Oper. Res.* **2014**, *234*, 491–498. [CrossRef]
28. Basak, S.; Shapiro, A. Value-at-risk-based risk management: Optimal policies and asset prices. *Rev. Financ. Stud.* **2001**, *14*, 371–405. [CrossRef]
29. Bensoussan, A.; Hoe, S.R.; Kim, J.; Yan, Z.F. A risk extended version of Merton's optimal consumption and portfolio selection. *Eur. J. Oper. Res.* **2022**, *70*, 815–829. [CrossRef]
30. Chiu, M.C.; Wong, H.Y.; Zhao, J. Dynamic safety first expected utility model. *Eur. J. Oper. Res.* **2018**, *271*, 141–154. [CrossRef]
31. Abad, C.; Iyengar, G. Portfolio Selection with Multiple Spectral Risk Constraints. *SIAM J. Financ. Math.* **2015**, *6*, 467–486. [CrossRef]
32. Guo, S.Y.; Xu, H.F. Robust spectral risk optimization when the subjective riskaversion is ambiguous: A moment-type approach. *Math. Program.* **2022**, *194*, 305–340. [CrossRef]
33. Moghimi, M.; Ku, H. Beyond CVaR: Leveraging Static Spectral Risk Measures for Enhanced Decision-Making in Distributional Reinforcement Learning. *arXiv* **2025**, arXiv:2501.02087.
34. Chiu, M.C.; Wong, H.Y.; Li, D. Roy's safety-first portfolio principle in financial risk management of disastrous events. *Risk Anal.* **2012**, *32*, 1856–1872. [CrossRef] [PubMed]
35. Madan, D.; Pistorius, M.; Stadje, M. On dynamic spectral risk measures, a limit theorem and optimal portfolio allocation. *Financ. Stoch.* **2017**, *21*, 1073–1102. [CrossRef]
36. Rockafellar, R.T.; Wets, R.J.B. Scenarios and policy aggregation in optimization under uncertainty. *Math. Oper. Res.* **1991**, *16*, 119–147. [CrossRef]
37. Watson, J.P.; Woodruff, D.L. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Comput. Manag. Sci.* **2011**, *8*, 355–370. [CrossRef]

38. Carpentier, P.L.; Gendreau, M.; Bastin, F. Long-term management of a hydroelectric multireservoir system under uncertainty using the progressive hedging algorithm. *Water Resour. Res.* **2013**, *49*, 2812–2827. [CrossRef]
39. Gul, S.; Denton, B.T.; Fowler, J.W. A progressive hedging approach for surgery planning under uncertainty. *INFORMS J. Comput.* **2015**, *27*, 755–772. [CrossRef]
40. Hvattum, L.M.; Løkketangen, A. Using scenario trees and progressive hedging for stochastic inventory routing problems. *J. Heuristics* **2009**, *15*, 527–557. [CrossRef]
41. Crainic, T.G.; Fu, X.; Gendreau, M.; Rei, W.; Wallace, S.W. Progressive hedging-based metaheuristics for stochastic network design. *Networks* **2011**, *58*, 114–124. [CrossRef]
42. Rockafellar, R.T. Solving stochastic programming problems with risk measures by progressive hedging. *Set-Valued Var. Anal.* **2018**, *26*, 759–768. [CrossRef]
43. Jin, C.; Wu, W.; Xie, J. Responsible Investing: A Multi-Period Portfolio Selection Model. 2024. Available online: https://ssrn.com/abstract=4814019 (accessed on 21 May 2025).
44. Kusuoka, S. *On Law Invariant Coherent Risk Measures*; Advances in mathematical economics; Springer: Tokyo, Japan, 2001; pp. 83–95.
45. Brandtner, M. "Spectral Risk Measures: Properties and Limitation": Comment on Dowd, Cotter, and Sorwar. *J. Financ. Serv. Res.* **2016**, *49*, 121–131. [CrossRef]
46. Bertsimas, D.; Lasuprete, G.J.; Samarov, A. Shortfall as a risk measure: Properties, optimization and applications. *J. Econ. Dyn. Control* **2004**, *28*, 1353–1381. [CrossRef]
47. Dowd, K.; Cotter, J.; Sorwar, G. Spectral Risk Measures: Properties and Limitations. *J. Financ. Serv. Res.* **2008**, *34*, 61–75. [CrossRef]
48. Brandtner, M. Conditional Value-at-Risk, spectral risk measures and (non-) diversification in portfolio selection problems-A comparison with mean-variance analysis. *J. Bank. Financ.* **2013**, *37*, 5526–5537. [CrossRef]
49. Strub, M.S.; Li, D.; Cui, X.; Gao, J. Discrete-time mean-CVaR portfolio selection and time-consistency induced term structure of the CVaR. *J. Econ. Dyn. Control* **2019**, *108*, 103751. [CrossRef]
50. Cui, X.; Gao, J.; Shi, Y.; Zhu, S. Time-consistent and self-coordination strategies for multi-period mean-conditional value-at-risk portfolio selection. *Eur. J. Oper. Res.* **2019**, *276*, 781–789. [CrossRef]
51. Acerbi, C.; Simonetti, P. Portfolio optimization with spectral measure of risk. *arXiv* **2002**, arXiv:cond-mat/0203607.
52. Benati, S.; Rizzi, R. A mixed integer linear programming formulation of the optimal mean/value-at-risk portfolio problem. *Eur. J. Oper. Res.* **2007**, *176*, 423–434. [CrossRef]
53. Liu, S.; Zhu, Q. Stackelberg risk preference design. *Math. Program.* **2025**, *209*, 785–823. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.