DYNAMIC MEAN-VARIANCE PROBLEM: RECOVERING TIME-CONSISTENCY

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

INDUSTRIAL ENGINEERING

By

Seyit Emre Düzoylum

August 2021

Dynamic mean-variance problem: recovering time-consistency

By Seyit Emre Düzoylum

August 2021

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Çağın Ararat(Advisor)

Savaş Dayanık

Diclehan Tezcaner Öztürk

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

DYNAMIC MEAN-VARIANCE PROBLEM: RECOVERING TIME-CONSISTENCY

Seyit Emre Düzoylum
M.S. in Industrial Engineering
Advisor: Çağın Ararat
August 2021

As the foundation of modern portfolio theory, Markowitz's mean-variance portfolio optimization problem is one of the fundamental problems of financial mathematics. The dynamic version of this problem in which a positive linear combination of the mean and variance objectives is minimized is known to be time-inconsistent, hence the classical dynamic programming approach is not applicable. Following the dynamic utility approach in the literature, we consider a less restrictive notion of time-consistency, where the weights of the mean and variance are allowed to change over time. Precisely speaking, rather than considering a fixed weight vector throughout the investment period, we consider an adapted weight process. Initially, we start by extending the well-known equivalence between the dynamic mean-variance and the dynamic mean-second moment problems in a general setting. Thereby, we utilize this equivalence to give a complete characterization of a time-consistent weight process, that is, a weight process which recovers the time-consistency of the mean-variance problem according to our definition. We formulate the mean-second moment problem as a biobjective optimization problem and develop a set-valued dynamic programming principle for the biobjective setup. Finally, we retrieve back to the dynamic mean-variance problem under the equivalence results that we establish and propose a backward-forward dynamic programming scheme by the methods of vector optimization. Consequently, we compute both the associated time-consistent weight process and the optimal solutions of the dynamic mean-variance problem.

Keywords: mean-variance problem, time-consistency, portfolio optimization, set-valued dynamic programming, vector optimization.

ÖZET

BEKLENTİ-DEĞİŞİNTİ PROBLEMİ: ZAMANDA TUTARLILIĞIN GERİ KAZANIMI

Seyit Emre Düzoylum
Endüstri Mühendisliği, Yüksek Lisans
Tez Danışmanı: Çağın Ararat
Ağustos 2021

Markowitz'in beklenti-değişinti portföy eniyileme problemi, modern portföy teorisinin başlangıç noktası olarak finans matematiğinin en temel problemlerinden biridir. Bu problemin dinamik uyarlaması, beklenti ve değişintinin pozitif bir doğrusal kombinasyonunun enküçüklenmesi ile gerçekleşir ve bu dinamik problemin zamanda tutarsız olduğu bilinmektedir. Bu sebeple dinamik programlamanın standart yöntemlerini bu problem üzerinde uygulamak mümkün değildir. Bu tez içerisinde, bilimsel yazında dinamik fayda yöntemi olarak bilinen bir yaklaşım takip edilerek zamanda tutarlılığın daha az kısıtlayıcı bir tanımı kullanılmıştır. Bu tanım altında, tüm yatırım süreci boyunca sabit bir ağırlık vektörü almak yerine, bu vektörlerin zaman içerisinde değişmesine izin verilerek, uyarlı bir ağırlık süreci değerlendirilmektedir. İlk aşama olarak bilimsel yazında bilinen bir sonuç olan beklenti-değişinti ve beklenti-ikinci moment problemlerinin arasındaki denklik ilişkisinin doğruluğu daha genel bir kurgu altında gösterilmiştir. Bu denklik altında, beklenti-değişinti probleminin zamanda tutarlılığını, kullanılan tanıma göre geri kazandıran zamanda tutarlı ağırlık sürecinin nitelendirmesi yapılmıştır. Devamında, beklenti-ikinci moment problemi iki amaçlı bir vektör eniyileme problemi olarak kurgulanmış, bu problem için küme değerli bir dinamik programlama ilkesi elde edilmiştir. Önceden elde edilen denklik sonuçları kullanılarak beklenti-değişinti problemi için, öncelikle zamanda geriye doğru, sonrasında zamanda ileriye doğru çalışan bir dinamik programlama yöntemi önerilmiştir. Böylelikle beklenti-değişinti probleminin eniyi çözümleri ve zamanda tutarlı ağırlık süreci dinamik bir biçimde hesaplanmıştır.

Anahtar sözcükler: beklenti değişinti problemi, zamanda tutarlılık, portföy eniyileme, küme değerli dinamik programlama, vektör eniyileme.

Acknowledgement

First and foremost, I would like to express my sincerest gratitude to my advisor Asst. Prof. Çağın Ararat. I can assure that every single little detail in this thesis owes its existence to Çağın Hocam. I will be in debt to him for the rest of my days for his never-ending support, understanding, patience, and guidance during these last three years. Each and every weak, I felt more and more privileged for getting the chance to have Çağın Hocam as my advisor and mentor. My admiration for the passion and interest that he puts into his work and students is immense, and it is one of my most genuine wishes to be able to show even a glance of it in my own future career.

I would like to thank Prof. Savaş Dayanık and Asst. Prof. Diclehan Tezcaner Öztürk, for allocating their valuable time to read and review this thesis and for their valuable feedback.

I would like to acknowledge the financial support provided by TÜBİTAK, The Scientific and Technological Research Council of Turkey, within project 117F438.

I am thankful to my dear friends Efe Sertkaya, İsmail Burak Taş, Barış Bilir, Kerem Ayöz, Kağan Kan, Kerem Avcı, Efe Orhun Şahin, Fırat Uçar, Samet İriş, Ece Önen, Çağrı Utku Sokat, Deniz Karazeybek, Cem Mirzaoğlu, Başak Ulutaş, Büşra Dişli, Şifa Çelik, Deniz Şimşek, Alperen Turan, Vakuralp Mor, Abdullah Buğra Kaya, Metin Ozantürk and everyone else that I could not list. Some of them I could not, unfortunately, get a break for the entirety of the last decade, whereas some of them I only got the chance to meet in the overtime. Their friendship has been a major pillar in my life that I could always lean on and find some support, and I am grateful to have each of them as a friend. My special thanks go to İrem Nur Keskin for always being there for me during these challenging times, for being the best friend there is, and for making my every single day a little bit brighter with her lovely smile.

Last but not least, I would like to thank my dear family, my sister Fatma, and my parents Ahmet and Hacer. Without their unconditional love and endless support, I definitely could not be here in this position. They have always, and always been there for me, which is worth more than anything. They are the best family that one can wish for, and I would like to devote this thesis to them.

Contents

List of Figures

Chapter 1

Introduction

As one of the fundamental problems of financial mathematics, after almost 70 years, the seminal work of Harry Markowitz on the mean-variance portfolio optimization problem [1] continues to attract extensive research. Though being the cornerstone of the modern financial analysis of the tradeoff between return and risk, to this day, there is not a unanimously agreed-upon approach for its extension to the multi-period setting. The fundamental issue is that the classical dynamic mean-variance problem turns out to be time-inconsistent. That is, an optimal portfolio for an investor at initial time may fail to be optimal at a later investment period, under the revelation of new information. Hence, Bellman's principle does not hold for the dynamic mean-variance problem in general; therefore, the classical methods of dynamic programming are not applicable. In the last twenty years, there has been a renewed interest in the time-consistency of the dynamic mean-variance problem. There are three main approaches in the literature for handling the issues incidental to time-inconsistency, which we review briefly.

The first one is the so-called *precommitment* approach, where the investor's notion of optimality only respects initial time. That is, it assumes that the investors cannot, or prefer not to, deviate from the original portfolio that they choose at the initial time. Hence, the investor precommits themselves into the utilization of the initial *optimal* portfolio, even though it may fail to be optimal at a later stage. To list some prominent studies that utilize the precommitment approach in discrete- and continuous-time settings, we recommend [2], [3], [4], [5], [6], and the references therein.

The *game-theoretic* approach, or sometimes called the *consistent planning* approach, is initially introduced by [7] for general time-inconsistent utility maximization problems and more recently revitalized for the dynamic mean-variance setting by [8]. The main idea is to consider dynamic portfolio optimization as an infinite sequential game, where the opponents are the investor's future reincarnations. These reincarnations are assumed to establish the local, in the sense of time, optimality of their portfolios for themselves. Hence, optimal strategies for the investor form a subgame perfect Nash equilibria. Some examples in the literature that utilize the game-theoretic approach are [9], [10], and [11].

The last approach, and in particular the one we utilize in this thesis, is what we prefer to call the *moving weight*, or the *dynamic utility* approach, where the main argument is that time-inconsistency occurs due to the incomplete formulation of the dynamic mean-variance problem. Indeed, the biobjective nature of the problem is typically encorporated by taking a linear combination of the two objectives for some appropriate weight vector $\lambda_0 = (\lambda_{0,1}, \lambda_{0,2})^\mathsf{T} \in \mathbb{R}^2_+$ and considering a scalar problem (see ($\mathrm{M}_0(v_0, \lambda_0)$) in Chapter 3) of the form

$$\text{minimize} \quad -\lambda_{0,1}\mathbb{E}(v_T^\pi) + \lambda_{0,2}\operatorname{Var}(v_T^\pi) \quad \text{subject to} \quad \pi \in \Phi_0(v_0), \tag{1.1}$$

where $\Phi_0(v_0)$ is the set of admissible portfolios with initial wealth $v_0 \in \mathbb{R}_+$ and v_T^π denotes the terminal wealth under an admissible portfolio π. At each intermediate time, a similar problem can be formulated using conditional expectation and variance; however, the same weight vector λ_0 is used throughout the investment period. This particular modeling feature cannot incorporate the investor's *dynamic attitude* towards risk as time progresses. [10] indicate that the assumption of constant risk aversion, that is, the constant relative weight assigned to the variance, is not realistic as the investor's attitude towards risk should change based on the amount of wealth they possess, while the investment period progresses.

Therefore, they consider a state-dependent weight vector based on the current wealth of the investor at the beginning of each investment period. We note that they still apply the game-theoretic approach for their problem formulation, hence the approach in [10] can be seen as a combination of both of these approaches. Moreover, [12] demonstrate that when the weights of the mean-variance problems are selected appropriately, indeed, time-consistency can be recovered. On the other hand, [13] investigate the dynamic behavior of the mean-variance problem when the weight process is constrained to take only nonnegative values (see Remark 3.3.3 below). They indicate that an investor can sustain a free cash flow out of the market while still having a time-consistent solution with respect to their definition.

Recently, [14] introduce a new methodology for the dynamic mean-risk problem, where the risk is measured by a time-consistent dynamic coherent risk measure, and develop a set-valued dynamic programming principle for the analogous problem. They start by formulating the mean-risk problem as a biobjective vector optimization framework. Moreover, rather than working towards obtaining a scalar Bellman's principle; they utilize the upper images of the vector optimization problems as the set-valued value functions for their analysis. Under this construction, they prove that a set-valued Bellman's equation holds for the mean-risk problem.

In this thesis, we formulate the dynamic-mean variance problem in discrete time, under a financial market where each feasible portfolio satisfies the self-financing property, and the resulting portfolio value is square-integrable. Moreover, in line with [12], we utilize a broader notion of time-consistency that generalizes the classical one, where the family of mean-variance problems are allowed to use different weights at different investment periods. We note that the scalar problem in (1.1) can be written equivalently as

$$\text{minimize} \quad -\lambda_{0,1}\mathbb{E}(v_T^\pi) - \lambda_{0,2}\left(\mathbb{E}(v_T^\pi)\right)^2 + \lambda_{0,2}\mathbb{E}\left((v_T^\pi)^2\right) \quad \text{subject to} \quad \pi \in \Phi_0(v_0).$$

[4] and [3] point out that the nonlinear function of the expected terminal wealth

in the above formulation is the fundamental cause of time-inconsistency. However, we observe that, indeed this nonlinear formulation can be manipulated, and utilized for recovering time-consistency itself. To that end, we start by extending the observations of [4] on the equivalence between the mean-variance and mean-second moment problems (see $\mathrm{A}_t(v_t, \rho_t)$ in Chapter 3), which they utilize to formulate their linear-quadratic embedding. We give a full characterization for the equivalent weights of the two problems, under which they share their optimal solutions. Moreover, we observe that the dynamic mean-second moment problem is time-consistent in the classical sense. Hence, we utilize the equivalence to give a complete characterization of a *time-consistent weight process*, that is, a weight process that would make the family of mean-variance problems time-consistent with respect to our definition (see Definition 3.0.3), which is given by the formula (see Theorem 3.3.2 below)

$$\lambda_t = \begin{pmatrix} \lambda_{0,1} + 2\lambda_{0,2}\left(\mathbb{E}\left(v_T^{\pi^\star}\right) - \mathbb{E}_t\left(v_T^{\pi^\star}\right)\right) \\ \lambda_{0,2} \end{pmatrix},$$

where $\pi^\star$ is the optimal solution of the initial mean-variance problem, $\lambda_{0,1}$ and $\lambda_{0,2}$ are the initial weights assigned by the investor to mean and variance, respectively. In order to obtain these conclusions, we heavily benefit from random set theory, and the decomposability of these problems.

We observe that, as the dynamic mean-second moment problem is time-consistent in the classical sense, that is, with a constant weight, it satisfies an associated scalar Bellman's equation. Therefore, in theory, our go-to method for formulating a dynamic programming scheme should incorporate this scalar Bellman's equation. However, we notice that the equivalent weights between these two problems are functions of the optimal terminal wealth, that is, for a given initial weight $\lambda_0 \in \mathbb{R}^2_+$ for the mean-variance problem, the corresponding weight for the mean-second moment problem can be calculated *provided that* the optimal terminal wealth is already known. Hence, to make the transition from the mean-variance problem into this dynamic programming scheme for an arbitrary initial weight λ_0, one has to know the optimal terminal wealth from the very beginning. This is usually not the case unless, for instance, the optimal solutions have a closed form expression as in [4]. For this reason, using the scalar Bellman's principle associated to the mean-second moment problem is generally an invalid approach when finding a time-consistent weight process for the mean-variance problem. Therefore, we move on to developing a recursive dynamic programming scheme that has a set-valued basis to obtain both optimal solutions and a time-consistent weight process for the dynamic-mean variance problem.

To that end, we follow a parallel approach to [14] and reformulate the mean-second moment problem within a vector optimization setting. The main advantage of this setting is the disembodiment from weighted sum scalarizations. With this advancement, we acquire the ability to study the dynamic behavior of the mean-second moment problem, free from the choice for a particular weight vector. Moreover, we still preserve the connection to the scalarized mean-second moment problem, hence the time-consistent weight process for the mean-variance problem, by the means of weakly minimal solutions. In alignment with [14], we let the value function of this dynamic vector optimization problem to be set-valued, and we choose it to be the so-called upper images of the vector optimization problems. Within this formulation, we obtain the complete set-valued analogue of the scalar Bellman's equation for the family of mean-second moment problems, with all the additional benefits mentioned above. We note that, as the scalar problem itself is time-consistent with a constant weight, under a relatively relaxed notion of time-consistency within the vector optimization framework, this conclusion is somewhat expected.

Under this new set-valued Bellman's equation, we introduce a recursive backward-forward dynamic programming method for obtaining the optimal solutions and the associated time-consistent weight processes of the mean-variance problem dynamically. First, we solve a series of backward one-step vector optimization problems to obtain the upper image of the mean-second moment problem at each step in the investment horizon, recursively. Then, as an intermediate step, we apply some transformations on the upper image of the mean-second problem at initial time in order to revert back to the mean-variance problem. Thanks to this advancement, we bypass the issues of scalar Bellman's equation, and we are able to find the associated initial weight of the mean-second moment problem for every weight vector $\lambda_0 \in \mathbb{R}^2_+$ that the investors assigns for the mean-variance problem at initial time. Finally, once we obtain the associated weight vector, we solve a series of one-step scalar optimization problems, which are indeed scalarized with the respect to the associated weights obtained in the previous step. As a result, we are able to obtain both the optimal solutions and the corresponding time-consistent weight process for every initial weight for the dynamic mean-variance problem systematically, and dynamically.

The organization for remainder of this thesis is as follows. In Chapter 2, we establish the notation and the underlying financial market structure. In Chapter 3, we give a full characterization of the time-consistent weight process for the mean-variance problem, and investigate the equivalence between the mean-variance and mean-second moment problems. In Chapter 4, we introduce a scalar Bellman's equation and develop a scalar dynamic programming principle for the mean-second moment problem. In Chapter 5, we reformulate the mean-second moment problem under a vector optimization framework, and obtain a set-valued analogue of the scalar Bellman's equation and dynamic programming principle. Finally, in Chapter 6, we propose a new dynamic programming methodology, which utilizes our set-valued dynamic programming principle, to find both the optimal solutions and the time-consistent weight processes of the dynamic mean-variance problem via vector optimization. We apply our methodology on two different financial markets and announce our findings.

Chapter 2

Preliminaries and notation

In this chapter, we introduce the structure of the financial market on which we will study the mean-variance problem. We also fix some notation for the rest of the thesis.

We work in a discrete-time setting with index set $\mathbb{T} = \{0, \ldots, T\}$ for some $T \in \mathbb{N} := \{1, 2, \ldots\}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\mathcal{F}_t)_{t\in\mathbb{T}}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathcal{F}_0$ being trivial and $\mathcal{F}_T = \mathcal{F}$. Let $n \in \mathbb{N}$. For $p \in [1, \infty)$, we denote the space of all equivalence classes of p-integrable and $\mathcal{F}_t$-measurable random variables taking values in $\mathbb{R}^n$ by $L^p_t(\mathbb{R}^n) := L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^n)$, whereas $L^\infty_t(\mathbb{R}^n) := L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^n)$ denotes the space of all equivalence classes of essentially bounded and $\mathcal{F}_t$-measurable random variables taking values in $\mathbb{R}^n$. The space $L^p_t(\mathbb{R}^n)$ is a Banach space with respect to the norm $x \mapsto \|x\|_p := \mathbb{E}(\|x\|^p)^{1/p}$ for $p \in [1, \infty)$ and $x \mapsto \|x\|_\infty$ for $p = \infty$, where

$$\|x\|_\infty := \inf\{c \in \mathbb{R}_+ \mid \|x\| \leq c\}.$$

In particular, $L^2_t(\mathbb{R}^n)$ is a Hilbert space with the inner product $(x, y) \mapsto \langle x, y\rangle := \mathbb{E}(x^\mathsf{T} y)$. For a subset $A \subseteq \mathbb{R}^n$, we denote the set of all random variables in $L^p_t(\mathbb{R}^n)$ that take values in A by $L^p_t(A)$. Given $D, E \subseteq L^p_t(\mathbb{R}^n)$, $D + E := \{x + y \mid x \in D,\ y \in E\}$ denotes the Minkowski sum of D, E. When $D = \{x\}$ for some

$x \in L_t^p(\mathbb{R}^n)$, we write $x + E \coloneqq \{x\} + E$. We denote the closure, interior, convex hull, linear hull, and conic hull of $D \subseteq L_t^p(\mathbb{R}^n)$ by $\mathrm{cl}(D)$, $\mathrm{int}(D)$, $\mathrm{conv}(D)$, $\mathrm{lin}(D)$, and $\mathrm{cone}(D)$, respectively. In general, many cones, and in particular $L_t^1(\mathbb{R}_+)$ and $L_t^2(\mathbb{R}_+)$, which we utilize often throughout the thesis, have empty interior when $L^p(\mathbb{R}^n) \coloneqq L_T^p(\mathbb{R}^n)$ is infinite dimensional. Hence, we make use of a weaker notion of interior for such sets. To that end, we denote the *quasi-interior* of D by $\mathrm{qi}(D)$, which is introduced by [15], and defined as

$$\mathrm{qi}(D) \coloneqq \{x \in D \mid \mathrm{cl}(\mathrm{cone}(D - x)) = L_t^p(\mathbb{R}^n)\}.$$

We denote the collection of all nonempty closed subsets of $\mathbb{R}^n$ by $\mathcal{C}(\mathbb{R}^n)$. We call a set-valued function $F\colon \Omega \to \mathcal{C}(\mathbb{R}^n)$ a *random closed set* if

$$\{\omega \in \Omega \mid F(\omega) \cap A \neq \emptyset\} \in \mathcal{F}$$

for every open set $A \subseteq \mathbb{R}^n$, see [16, Definition 1.1.1]. Moreover, we call a random variable $x\colon \Omega \to \mathbb{R}^n$ a *measurable selection* of F if $x(\omega) \in F(\omega)$ for $\mathbb{P}$-almost every $\omega \in \Omega$. We call a given set $D \subseteq L_t^p(\mathbb{R}^n)$ *decomposable* (with respect to $\mathcal{F}_t$) if $\mathbf{1}_B x^1 + \mathbf{1}_{B^c} x^2 \in D$ for every $x^1, x^2 \in D$, and $B \in \mathcal{F}_t$, where $\mathbf{1}_B\colon \Omega \to \mathbb{R}$ is the indicator function of B defined by

$$\mathbf{1}_B(\omega) = \begin{cases} 1 & \omega \in B, \\ 0 & \omega \in B^c. \end{cases}$$

Throughout the thesis, equalities and inequalities between random variables should be understood in the $\mathbb{P}$-almost sure ($\mathbb{P}$-a.s.) sense. Furthermore, addition, multiplication or composition of random variables should be understood pointwise. For the sake of readability, for each $t \in \mathbb{T}$, we denote the conditional expectation and conditional variance given $\mathcal{F}_t$ by $\mathbb{E}_t(\cdot) = \mathbb{E}(\,\cdot \mid \mathcal{F}_t)$ and $\mathrm{Var}_t(\cdot) = \mathrm{Var}(\,\cdot \mid \mathcal{F}_t)$, respectively.

We consider a financial market with $d \in \mathbb{N}$ assets which follow a d-dimensional square-integrable and $(\mathcal{F}_t)_{t\in\mathbb{T}}$-adapted discounted price process $(S_t)_{t\in\mathbb{T}}$. An investor in this market utilizes an essentially bounded and $(\mathcal{F}_t)_{t\in\mathbb{T}}$-predictable portfolio process $(\pi_t)_{t\in\mathbb{T}\setminus\{0\}}$, where $\pi_t = (\pi_t^1, \ldots, \pi_t^n)^{\mathsf{T}}$. In particular, for each $t \in \mathbb{T}\setminus\{0\}$ and $k \in \{1, \ldots, d\}$, π_t^k denotes the number of physical units of asset k to be held in period t, that is, the duration between times $t-1$ and t. The investor enters the market with an initial wealth $v_0 \in \mathbb{R}_+$ and they do not withdraw or deposit any wealth at an intermediate time step, thus each admissible portfolio should be self-financing. Furthermore, we suppose that the market can impose some additional constraints on the portfolio positions, for which we incorporate into our model as the convex constraints $\varphi_t(\pi_{t+1}) \geq 0$ for every $t \in \mathbb{T}\setminus\{T\}$, where $\varphi_t \colon L_t^\infty(\mathbb{R}^d) \to L_t^\infty(\mathbb{R}^m)$ is an upper semicontinuous and concave function for some image space dimension $m \in \mathbb{N}$. Note that the sequence of market constraints is identical for every initial wealth of the investor, that is, the structure of $(\varphi_t)_{t\in\mathbb{T}\setminus\{T\}}$ is independent of the initial wealth $v_0 \in \mathbb{R}_+$. Under these assumptions, we denote the set of all admissible portfolios for an investor with a starting wealth $v_t \in L_t^2(\mathbb{R})$ at time $t \in \mathbb{T}\setminus\{T\}$ by $\Phi_t(v_t)$, which is given by

$$\begin{aligned}\Phi_t(v_t) \coloneqq \big\{(\pi_s)_{s\geq t+1} \mid \pi_{t+1}^{\mathsf{T}} S_t = v_t,\ &\forall s \in \{t+1, \ldots, T-1\} \colon \pi_s^{\mathsf{T}} S_s = \pi_{s+1}^{\mathsf{T}} S_s,\\ &\forall s \in \{t+1, \ldots, T\} \colon \varphi_{s-1}(\pi_s) \geq 0,\\ &\forall s \in \{t+1, \ldots, T\} \colon \pi_s \in L_{s-1}^\infty(\mathbb{R}^d)\big\}.\end{aligned}$$

We note that, for each $t \in \mathbb{T}\setminus\{T\}$ and $v_t \in L_t^2(\mathbb{R})$, the feasible set $\Phi_t(v_t)$ is a convex subset of $L_t^\infty(\mathbb{R}^d) \times \ldots \times L_{T-1}^\infty(\mathbb{R}^d)$ by construction. For every portfolio $\pi = (\pi_s)_{s\geq t+1} \in \Phi_t(v_t)$, we define its value at time $s \in \{t+1, \ldots, T\}$ as

$$v_{t,s}^\pi \coloneqq \pi_s^{\mathsf{T}} S_s. \tag{2.1}$$

Furthermore, whenever $t = 0$, for the sake of readability, we drop t from the notation and write $v_s^\pi \coloneqq v_{0,s}^\pi$.

Chapter 3

Time-consistent weight process

Due to its multi-objective nature, the dynamic mean-variance problem is usually treated by considering a linear combination of the two objectives for some weight vector $\lambda_0 = (\lambda_{0,1}, \lambda_{0,2}) \in \mathbb{R} \times \mathbb{R}_+$, which yields a scalar objective function. Following this classical methodology, we define the mean-variance problem at time zero with given wealth $v_0 \in \mathbb{R}_+$ as

$$\inf_{\pi \in \Phi_0(v_0)} -\lambda_{0,1}\mathbb{E}(v_T^\pi) + \lambda_{0,2}\mathrm{Var}(v_T^\pi). \tag{$\mathrm{M}_0(v_0, \lambda_0)$}$$

Similarly, as we are interested in the dynamic behavior, we formulate the corresponding problem at an intermediate time step $t \in \mathbb{T}\backslash\{T\}$ with some initial wealth $v_t \in L_t^2(\mathbb{R})$ as

$$\operatorname*{ess\,inf}_{\pi \in \Phi_t(v_t)} -\lambda_{0,1}\mathbb{E}_t(v_{t,T}^\pi) + \lambda_{0,2}\mathrm{Var}_t(v_{t,T}^\pi). \tag{$\mathrm{M}_t(v_t, \lambda_0)$}$$

We start by announcing the standard definition of time-consistency for the family $(\mathrm{M}_t(\cdot, \lambda_0))_{t\in\mathbb{T}\backslash\{T\}}$.

Definition 3.0.1. *The family* $(\mathrm{M}_t(\cdot, \lambda_0))_{t\in\mathbb{T}\backslash\{T\}}$ *of optimization problems with initial wealth* $v_0 \in \mathbb{R}_+$ *is called* time-consistent *if every optimal solution* $\pi^\star$ *of* $(\mathrm{M}_0(v_0, \lambda_0))$ *continues to be optimal at any future time, that is,* $(\pi_s^\star)_{s\geq t+1}$ *is an optimal solution of* $(\mathrm{M}_t(v_t^{\pi^\star}, \lambda_0))$ *for all* $t \in \mathbb{T}\backslash\{T\}$.

It is well-known in the literature that the family $(\mathrm{M}_t(\cdot, \lambda_0))_{t\in\mathbb{T}\setminus\{T\}}$ fails to be time-consistent in the sense of Definition 3.0.1 (see [8]). In line with the works [10], [12] and [13] that work under different settings, we will argue that time-consistency can be recovered if one wishes to replace the static weight λ_0 that is used at all times in $(\mathrm{M}_t(\cdot, \lambda_0))_{t\in\mathbb{T}\setminus\{T\}}$ by an adapted stochastic weight process $(\lambda_t)_{t\in\mathbb{T}\setminus\{T\}}$. This, in return, will incorporate the potential change of the relative weight assigned to each objective by the investor during the investment horizon under different financial outcomes.

Therefore, we define the mean-variance problem at time $t \in \mathbb{T}\setminus\{T\}$ with some initial wealth $v_t \in L_t^2(\mathbb{R})$ and random weight $\lambda_t = (\lambda_{t,1}, \lambda_{t,2})^\mathsf{T} \in L_t^2(\mathbb{R}) \times L_t^\infty(\mathbb{R}_+)$ as

$$\operatorname*{ess\,inf}_{\pi\in\Phi_t(v_t)} -\lambda_{t,1}\mathbb{E}_t(v_{t,T}^\pi) + \lambda_{t,2}\mathrm{Var}_t(v_{t,T}^\pi). \tag{$\mathrm{M}_t(v_t, \lambda_t)$}$$

Remark 3.0.2. We note that, apart from the premise of yielding time-consistency for the mean-variance problem, this particular methodology has a financial rationale as well. In the literature, the weight λ_0 is interpreted as the risk aversion of an investor at initial time, as the components of λ_0 are the relative weights that the investor assigns to risk and expected return. Therefore, it can be argued that considering a fixed weight throughout the entire investment period is equivalent to determining the risk aversion of the investor at the future apriori, before the revelation of additional information. In fact, [10] notes that if the investor accumulates some additional wealth during the investment process, as the perception of risk should change significantly depending on the amount of wealth present to the investor, their risk aversion should be increasing as well.

Under the new structure of the dynamic mean-variance problem with random weights, we extend our definition of time-consistency accordingly.

Definition 3.0.3. *The family $(\mathrm{M}_t(\cdot, \lambda_t))_{t\in\mathbb{T}\setminus\{T\}}$ of optimization problems with initial wealth $v_0 \in \mathbb{R}_+$ is called* time-consistent *under the weight process $(\lambda_t)_{t\in\mathbb{T}\setminus\{T\}}$ if every optimal solution $\pi^\star$ of $(\mathrm{M}_0(v_0, \lambda_0))$ continues to be optimal at any future time, that is, $(\pi_s^\star)_{s\geq t+1}$ is an optimal solution of $\mathrm{M}_t(v_t^{\pi^\star}, \lambda_t)$ for all $t \in \mathbb{T}\setminus\{T\}$.*

In this chapter, our main goal is to characterize the associated weight process $(\lambda_t)_{t\in\mathbb{T}\backslash\{T\}}$ for a given initial weight $\lambda_0 \in \mathbb{R}^2_+$ and an initial wealth $v_0 \in \mathbb{R}_+$, for which the family $(\mathrm{M}_t(v_t,\lambda_t))_{t\in\mathbb{T}\backslash\{T\}}$ becomes time-consistent according to Definition 3.0.3. To that end, we utilize the auxiliary *mean-second moment problem*, which in fact possesses an inherent relationship with the mean-variance problem. Precisely, we announce the mean-second moment problem at time $t \in \mathbb{T}\backslash\{T\}$ for some initial wealth $v_t \in L^2_t(\mathbb{R})$ and random weight $\rho_t = (\rho_{t,1}, \rho_{t,2})^\mathsf{T} \in L^2_t(\mathbb{R}_+) \times L^\infty_t(\mathbb{R}_+)$ as

$$\operatorname*{ess\,inf}_{\pi\in\Phi_t(v_t)} -\rho_{t,1}\mathbb{E}_t(v^\pi_{t,T}) + \rho_{t,2}\mathbb{E}_t((v^\pi_{t,T})^2). \qquad (\mathrm{A}_t(v_t,\rho_t))$$

Notice that the notion of time-consistency is directly tied with the optimal strategies of the family $(\mathrm{M}_t(\cdot,\lambda_t))_{t\in\mathbb{T}\backslash\{T\}}$ of optimization problems. Therefore, for completeness, throughout the rest of the thesis, we shall assume the following assumption on the existence of an optimal strategy for both $(\mathrm{M}_t(v_t,\lambda_t))$ and $(\mathrm{A}_t(v_t,\rho_t))$.

Assumption 3.0.4. *For each $t \in \mathbb{T}\backslash\{T\}$, $v_t \in L^2_t(\mathbb{R})$, $\lambda_t \in L^2_t(\mathbb{R}) \times L^\infty_t(\mathbb{R}_+)$, and $\rho_t \in L^2_t(\mathbb{R}_+) \times L^\infty_t(\mathbb{R}_+)$, the optimal values of $(\mathrm{M}_t(v_t,\lambda_t))$ and $(\mathrm{A}_t(v_t,\rho_t))$ are attained. That is, there exist an optimal strategies $\pi^\star \in \Phi_t(v_t)$ and $\bar{\pi} \in \Phi_t(v_t)$ for the problems $(\mathrm{M}_t(v_t,\lambda_t))$ and $(\mathrm{A}_t(v_t,\rho_t))$, respectively.*

As we observe later on with Theorem 3.2.1, there is a fundamental equivalence between these two problems in terms of their optimal solutions, which is based on the appropriate choice of their respective weights λ_t and ρ_t. Moreover, with Proposition 3.3.1, we observe that the mean-second moment problem turns out to be time-consistent in the classical sense, that is, in the spirit of Definition 3.0.1. We note that these observations are well-known in the literature, and the initial observations are recognized by [4], be it in a simpler finite probabilistic setup and under special financial market settings. Therefore, our main results in this chapter extend these observations to a possibly infinite-dimensional probabilistic setup with as much generality as possible, which turns out to be non-trivial and benefits heavily from the random set theory. To that end, we start our analysis with the so-called decomposability of the two problems of interest.

3.1 Decomposability of $(\mathrm{M}_t(v_t, \lambda_t))$ and $(\mathrm{A}_t(v_t, \rho_t))$

We start by defining the following two sets which are the images of the mean-variance and mean-second moment pairs, respectively:

$$\mathcal{M}_t(v_t) := \left\{ \begin{pmatrix} -\mathbb{E}_t(v^\pi_{t,T}) \\ \mathrm{Var}_t(v^\pi_{t,T}) \end{pmatrix} \mid \pi \in \Phi_t(v_t) \right\}, \ \mathcal{A}_t(v_t) := \left\{ \begin{pmatrix} -\mathbb{E}_t(v^\pi_{t,T}) \\ \mathbb{E}_t((v^\pi_{t,T})^2) \end{pmatrix} \mid \pi \in \Phi_t(v_t) \right\},$$

where $v_t \in L^2_t(\mathbb{R})$ is given. Moreover, for the sake of completeness, for every $v_T \in L^2_T(\mathbb{R})$, we define

$$\mathcal{A}_T(v_T) := \left(v_T, \, (v_T)^2\right)^\mathsf{T}.$$

Since $\mathrm{Var}_t(v^\pi_{t,T}) = \mathbb{E}_t((v^\pi_{t,T})^2) - (\mathbb{E}_t(v^\pi_{t,T}))^2$ for every $t \in \mathbb{T}\setminus\{T\}$ and $\pi \in \Phi_t(v_t)$, we immediately have

$$\mathcal{M}_t(v_t) = \left\{ (x_1, x_2 - (x_1)^2)^\mathsf{T} \mid x \in \mathcal{A}_t(v_t) \right\}. \tag{3.1}$$

For future use and readability, we introduce the following notation and re-express (3.1) compactly. Let us define a function $\mathcal{T} \colon \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\mathcal{T}(z) = (z_1, z_2 - (z_1)^2)^\mathsf{T}, \tag{3.2}$$

whose inverse function exists and is given by $\mathcal{T}^{-1}(z) = (z_1, z_2 + (z_1)^2)^\mathsf{T}$; both $\mathcal{T}$ and $\mathcal{T}^{-1}$ are continuous on $\mathbb{R}^2$. Moreover, in order to extend the pointwise transformation to random variables as well, for each $t \in \mathbb{T}$, we define $\widehat{\mathcal{T}}_t : L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R}) \to L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})$ by

$$\widehat{\mathcal{T}}_t(x) = \mathcal{T} \circ x.$$

Thus, we may rewrite (3.1) as

$$\mathcal{M}_t(v_t) = \left\{ \widehat{\mathcal{T}}_t(x) \mid x \in \mathcal{A}_t(v_t) \right\}. \tag{3.3}$$

Accordingly, the problems $(\mathrm{M}_t(v_t, \lambda_t))$ and $(\mathrm{A}_t(v_t, \rho_t))$ can be rewritten as

$$\operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x) \quad (\mathrm{M}_t(v_t, \lambda_t)), \qquad \operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \rho_t^\mathsf{T} x \quad (\mathrm{A}_t(v_t, \rho_t)). \tag{3.4}$$

Lemma 3.1.1. *Let $t \in \mathbb{T}\setminus\{T\}$ and $v_t \in L_t^2(\mathbb{R})$. Then, the image set $\mathcal{A}_t(v_t)$ is a decomposable subset of $L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})$.*

Proof. Let us consider $x^1, x^2 \in \mathcal{A}_t(v_t)$ and $B \in \mathcal{F}_t$. Then, there exist $\pi^1, \pi^2 \in \Phi_t(v_t)$ such that $x^1 = (-\mathbb{E}_t(v_{t,T}^{\pi^1}), \mathbb{E}_t((v_{t,T}^{\pi^1})^2))^\mathsf{T}$ and $x^2 = (-\mathbb{E}_t(v_{t,t}^{\pi^2}), \mathbb{E}_t((v_{t,T}^{\pi^2})^2))^\mathsf{T}$. Now let $\pi \coloneqq \mathbf{1}_B\pi^1 + \mathbf{1}_{B^c}\pi^2$. Then, we observe that $\pi_{t+1}^\mathsf{T} S_t = v_t$,

$$\pi_s^\mathsf{T} S_s = \mathbf{1}_B(\pi_s^1)^\mathsf{T} S_s + \mathbf{1}_{B^c}(\pi_s^2)^\mathsf{T} S_s = \mathbf{1}_B(\pi_{s+1}^1)^\mathsf{T} S_s + \mathbf{1}_{B^c}(\pi_{s+1}^2)^\mathsf{T} S_s = \pi_{s+1}^\mathsf{T} S_s$$

for all $s \in \{t+1, \dots, T-1\}$, and

$$\varphi_{s-1}(\pi_s) = \mathbf{1}_B\varphi_{s-1}(\pi_s^1) + \mathbf{1}_{B^c}\varphi_{s-1}(\pi_s^2) \geq 0,\ \pi_s \in L_{s-1}^\infty(\mathbb{R}^n)$$

for all $s \in \{t+1, \dots, T\}$. Hence, $\pi \in \Phi_t(v_t)$. By the definition of value process in (2.1), we have

$$v_{t,T}^\pi = v_{t,T}^{\mathbf{1}_B\pi^1 + \mathbf{1}_{B^c}\pi^2} = \mathbf{1}_B v_{t,T}^{\pi^1} + \mathbf{1}_{B^c} v_{t,T}^{\pi^2}. \tag{3.5}$$

Furthermore, due to the properties of indicator function, we have the following useful identity:

$$(v_{t,t}^\pi)^2 = (\mathbf{1}_B v_{t,T}^{\pi^1})^2 + (\mathbf{1}_{B^c} v_{t,T}^{\pi^2})^2 + 2(\mathbf{1}_B\mathbf{1}_{B^c} v_{t,T}^{\pi^1} v_T^{\pi^2}) = \mathbf{1}_B(v_{t,T}^{\pi^1})^2 + \mathbf{1}_{B^c}(v_{t,T}^{\pi^2})^2. \tag{3.6}$$

Then, by combining (3.5) and (3.6) with the linearity of conditional expectation and the $\mathcal{F}_t$-measurablity of $\mathbf{1}_B$, we obtain

$$\begin{pmatrix} -\mathbb{E}_t(v_{t,T}^\pi) \\ \mathbb{E}_t((v_{t,T}^\pi)^2) \end{pmatrix} = \begin{pmatrix} -\mathbf{1}_B\mathbb{E}_t(v_{t,T}^{\pi^1}) - \mathbf{1}_{B^c}\mathbb{E}_t(v_{t,T}^{\pi^2}) \\ \mathbf{1}_B\mathbb{E}_t((v_{t,T}^{\pi^1})^2) + \mathbf{1}_{B^c}\mathbb{E}_t((v_{t,T}^{\pi^2})^2) \end{pmatrix} = \mathbf{1}_B x^1 + \mathbf{1}_{B^c} x^2.$$

Hence, we conclude that $\mathbf{1}_B x^1 + \mathbf{1}_{B^c} x^2 \in \mathcal{A}_t(v_t)$. □

For a given $v_t \in L_t^2(\mathbb{R})$, we observe that the image sets $\mathcal{A}_t(v_t)$ and $\mathcal{M}_t(v_t)$ fail to be convex in general, which can be very inconvenient, especially under the current optimization framework. However, by the next two lemmas, we recover the convexity of both images sets by adding the cone $\{0\} \times L_t^1(\mathbb{R}_+)$ to each of them. Furthermore, as we observe later in Lemma 3.1.5, this addition has no significant effect for our purposes.

Lemma 3.1.2. *Let $t \in \mathbb{T}\setminus\{T\}$, $v_t \in L^2_t(\mathbb{R})$, and define*

$$\overline{\mathcal{A}_t(v_t)} := \mathrm{cl}\left(\mathcal{A}_t(v_t) + \{0\} \times L^1_t(\mathbb{R}_+)\right),$$

where the closure is taken with respect to the product topology on $L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})$. Then, $\overline{\mathcal{A}_t(v_t)}$ is a closed, convex and decomposable subset of $L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})$.

Proof. Note that $\overline{\mathcal{A}_t(v_t)}$ is closed by definition. Moreover, observe that $\mathcal{A}_t(v_t)$ is decomposable by Lemma 3.1.1 and $\{0\} \times L^1_t(\mathbb{R}_+)$ is decomposable by definition. Hence, as the sum and closure of decomposable sets is again decomposable, we conclude that $\overline{\mathcal{A}_t(v_t)}$ is decomposable. The convexity of $\mathcal{A}_t(v_t) + \{0\} \times L^1_t(\times\mathbb{R}_+)$ follows from the convexity of $\Phi_t(v_t)$, the linearity of conditional expectation and the convexity of second moment. Therefore, as the closure of a convex set, we conclude that $\overline{\mathcal{A}_t(v_t)}$ is a convex set as well. □

In order to obtain an analogue of Lemma 3.1.2 for $\mathcal{M}_t(v_t)$, we need the following lemma that establishes the conditions on the continuity of $\widehat{\mathcal{T}}_t$.

Lemma 3.1.3. *Let $t \in \mathbb{T}\setminus\{T\}$ and $v_t \in L^2_t(\mathbb{R})$. Then, the function $\widehat{\mathcal{T}}_t$ is continuous on $\overline{\mathcal{A}_t(v_t)}$. Furthermore, we have*

$$\widehat{\mathcal{T}}_t\left[\overline{\mathcal{A}_t(v_t)}\right] = \mathrm{cl}\,\widehat{\mathcal{T}}_t\left[\mathcal{A}_t(v_t) + \{0\} \times L^1_t(\mathbb{R}_+)\right].$$

Proof. To show the continuity claim, let $(x^n)_{n\in\mathbb{N}} \subseteq \overline{\mathcal{A}_t(v_t)}$ be a sequence that converges to some $x \in \overline{\mathcal{A}_t(v_t)}$ in $L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})$. Note that

$$\widehat{\mathcal{T}}_t(x^n) = \begin{pmatrix} x_1^n \\ x_2^n - (x_1^n)^2 \end{pmatrix}, \ n \in \mathbb{N}, \qquad \widehat{\mathcal{T}}_t(x^n) = \begin{pmatrix} x_1 \\ x_2 - (x_1)^2 \end{pmatrix}.$$

By the construction of the sequence, $(x_1^n)_{n\in\mathbb{N}}$ converges to x_1 in $L^2_t(\mathbb{R})$ and $(x_2^n)_{n\in\mathbb{N}}$ converges to x_2 in $L^1_t(\mathbb{R})$. In particular, $((x_1^n)^2)_{n\in\mathbb{N}}$ converges to $(x_1)^2$ in $L^1_t(\mathbb{R})$ as well. Hence, $(x_2^n - (x_1^n)^2)_{n\in\mathbb{N}}$ converges to $(x_2 - (x_1)^2)$ in $L^1_t(\mathbb{R})$. It follows that

$(\widehat{\mathcal{T}}_t(x^n))_{n\in\mathbb{N}}$ converges to $\widehat{\mathcal{T}}_t(x)$ in $L^2_t(\mathbb{R})\times L^1_t(\mathbb{R})$, hence the continuity of $\widehat{\mathcal{T}}_t$ follows. Therefore, the forward inclusion $\widehat{\mathcal{T}}_t[\overline{\mathcal{A}_t(v_t)}]\subseteq \operatorname{cl}\widehat{\mathcal{T}}_t\left[\mathcal{A}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+)\right]$ is straightforward.

To prove the reverse inclusion, we observe that the inverse function $\widehat{\mathcal{T}}_t^{-1}\colon L^2_t(\mathbb{R})\times L^1_t(\mathbb{R})\to L^2_t(\mathbb{R})\times L^1_t(\mathbb{R})$ exists by construction, and it is given by $\widehat{\mathcal{T}}_t^{-1}(x)=\mathcal{T}^{-1}\circ x$. Moreover, similar to the continuity proof above, it can be checked that $\widehat{\mathcal{T}}_t^{-1}$ is continuous on $\operatorname{cl}\widehat{\mathcal{T}}_t\left[\mathcal{A}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+)\right]$. Since the arguments are similar, we omit the proof for brevity. Then, we have

$$\begin{aligned}\widehat{\mathcal{T}}_t^{-1}\left[\operatorname{cl}\widehat{\mathcal{T}}_t\left[\mathcal{A}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+)\right]\right]&\subseteq \operatorname{cl}\left(\widehat{\mathcal{T}}_t^{-1}\left[\widehat{\mathcal{T}}_t\left[\mathcal{A}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+)\right]\right]\right)\\&\subseteq \operatorname{cl}\left(\mathcal{A}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+)\right)=\overline{\mathcal{A}_t(v_t)},\end{aligned}$$

and when we apply the transformation $\widehat{\mathcal{T}}_t$ once more to both sides of the set inclusion, it implies that

$$\operatorname{cl}\widehat{\mathcal{T}}_t\left[\mathcal{A}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+)\right]\subseteq\widehat{\mathcal{T}}_t\left[\overline{\mathcal{A}_t(v_t)}\right],$$

which completes the proof. □

Lemma 3.1.4. *Let $t\in\mathbb{T}\setminus\{T\}$ and $v_t\in L^2_t(\mathbb{R})$. Then, we have*

$$\overline{\mathcal{M}_t(v_t)}\coloneqq \operatorname{cl}(\mathcal{M}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+))=\{\widehat{\mathcal{T}}_t(x)\mid x\in\overline{\mathcal{A}_t(v_t)}\}. \tag{3.7}$$

where the closure is taken with respect to the product topology on $L^2_t(\mathbb{R})\times L^1_t(\mathbb{R})$. Further, $\overline{\mathcal{M}_t(v_t)}$ is a closed, convex and decomposable subset of $L^2_t(\mathbb{R})\times L^1_t(\mathbb{R})$.

Proof. First, we note that by Lemma 3.1.3 we have

$$\left\{\widehat{\mathcal{T}}_t(x)\mid x\in\overline{\mathcal{A}_t(v_t)}\right\}=\widehat{\mathcal{T}}_t\left[\overline{\mathcal{A}_t(v_t)}\right]=\operatorname{cl}\widehat{\mathcal{T}}_t\left[\mathcal{A}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+)\right]. \tag{3.8}$$

Therefore, under (3.3), it suffices to show the following equality holds:

$$\widehat{\mathcal{T}}_t\big[\mathcal{A}_t(v_t)+\{0\}\times L^1_t(\mathbb{R}_+)\big]=\widehat{\mathcal{T}}_t[\mathcal{A}_t(v_t)]+\{0\}\times L^1_t(\mathbb{R}_+).$$

To that end, let $x \in \mathcal{A}_t(v_t)$ and $r \in \{0\} \times L^1_t(\mathbb{R}_+)$, then as $r_1 = 0$, we obtain that

$$\widehat{\mathcal{T}}_t(x+r) = (x_1, x_2+r_2-(x_1)^2)^\mathsf{T} = (x_1, x_2-(x_1)^2)^\mathsf{T} + (0, r_2)^\mathsf{T} = \widehat{\mathcal{T}}_t(x)+r. \quad (3.9)$$

Hence, under (3.8) and (3.9), we conclude that

$$\begin{aligned}\{\widehat{\mathcal{T}}_t(x) \mid x \in \overline{\mathcal{A}_t(v_t)}\} &= \operatorname{cl}(\widehat{\mathcal{T}}_t[\mathcal{A}_t(v_t)] + \{0\} \times L^1_t(\mathbb{R}_+)) \\ &= \operatorname{cl}(\mathcal{M}_t(v_t) + \{0\} \times L^1_t(\mathbb{R}_+)) = \overline{\mathcal{M}_t(v_t)}.\end{aligned}$$

We note that, $\overline{\mathcal{M}_t(v_t)}$ is closed in $L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})$ by definition. Moreover, the convexity of $\overline{\mathcal{M}_t(v_t)}$ follows from the convexity of $\Phi_t(v_t)$, the linearity of the conditional expectation and the convexity of the conditional variance. In terms of decomposability, similar to the proof of Lemma 3.1.2, it suffices to show that $\mathcal{M}_t(v_t)$ is decomposable. To that end, let $m^1, m^2 \in \mathcal{M}_t(v_t)$ and $B \in \mathcal{F}_t$. Then, there exists $x^1, x^2 \in \mathcal{A}_t(v_t)$ such that $\widehat{\mathcal{T}}_t(x^1) = m^1$ and $\widehat{\mathcal{T}}_t(x^2) = m^2$. Further, we let $x = \mathbf{1}_B x^1 + \mathbf{1}_{B^c} x^2$ then as $\mathcal{A}_t(v_t)$ is decomposable by Lemma 3.1.1, we have $x \in \mathcal{A}_t(v_t)$. Therefore, following a similar argument to (3.6), we obtain

$$\mathbf{1}_B m^1 + \mathbf{1}_{B^c} m^2 = \mathbf{1}_B \widehat{\mathcal{T}}_t(x^1) + \mathbf{1}_{B^c} \widehat{\mathcal{T}}_t(x^2) = \widehat{\mathcal{T}}_t(\mathbf{1}_B x^1 + \mathbf{1}_{B^c} x^2) = \widehat{\mathcal{T}}_t(x).$$

Thus, as it is the image of some $x \in \mathcal{A}_t(v_t)$ under $\widehat{\mathcal{T}}_t$, we conclude that $\mathbf{1}_B m^1 + \mathbf{1}_{B^c} m^2 \in \mathcal{M}_t(v_t)$ by (3.7). □

Our next result indicates that when we replace the image set $\mathcal{A}_t(v_t)$ in (3.4) with the newly defined $\overline{\mathcal{A}_t(v_t)}$, although the feasible region is larger, both problems have the same optimal value.

Lemma 3.1.5. *Let $t \in \mathbb{T} \setminus \{T\}$, $v_t \in L^2_t(\mathbb{R})$. Then, it holds that*

$$\operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x) = \operatorname*{ess\,inf}_{x \in \overline{\mathcal{A}_t(v_t)}} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x)$$

for every $\lambda_t \in L^2_t(\mathbb{R}) \times L^\infty_t(\mathbb{R}_+)$ in the case of the mean-variance problem $(\mathrm{M}_t(v_t, \lambda_t))$, and

$$\operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \rho_t^\mathsf{T} x = \operatorname*{ess\,inf}_{x \in \overline{\mathcal{A}_t(v_t)}} \rho_t^\mathsf{T} x$$

for every $\rho_t \in L^2_t(\mathbb{R}_+) \times L^\infty_t(\mathbb{R}_+)$ in the case of the mean-second moment problem $(\mathrm{A}_t(v_t, \rho_t))$.

Proof. We only provide the proof of the case for $(\mathrm{M}_t(v_t, \lambda_t))$ to prevent repetition. Let us fix some $\lambda_t \in L^2_t(\mathbb{R}) \times L^\infty_t(\mathbb{R}_+)$. As $\mathcal{A}_t(v_t) \subseteq \overline{\mathcal{A}_t(v_t)}$ by construction, it is clear that

$$\operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x) \geq \operatorname*{ess\,inf}_{x \in \overline{\mathcal{A}_t(v_t)}} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x).$$

To obtain the reverse inequality, we start with the observation that, for every $x \in \mathcal{A}_t(v_t)$ and $r \in \{0\} \times L^1_t(\mathbb{R}_+)$, it holds

$$\lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x + r) = \lambda_t^\mathsf{T} (\widehat{\mathcal{T}}_t(x) + r) = \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x) + \lambda_t^\mathsf{T} r \geq \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x),$$

as we have $\lambda_t^\mathsf{T} r \geq 0$. Thus, by the definition essential infimum,

$$\operatorname*{ess\,inf}_{x' \in \mathcal{A}_t(v_t)} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x') \leq \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x) \leq \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x + r) \tag{3.10}$$

for all $x \in \mathcal{A}_t(v_t)$ and $r \in \{0\} \times L^1_t(\mathbb{R}_+)$. As the next step, we consider $\overline{x} \in \overline{\mathcal{A}_t(v_t)}$ such that

$$\lim_{n \to \infty} (x^n + r^n) = \overline{x} \text{ in } L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R}),$$

where $x^n \in \mathcal{A}_t(v_t) \in$ and $r^n \in L^2_t(\{0\} \times \mathbb{R}_+)$ for each $n \in \mathbb{N}$. Then, as $\widehat{\mathcal{T}}_t$ is continuous on $\overline{\mathcal{A}_t(v_t)}$ by Lemma 3.1.3 and $\lambda \in L^2_t(\mathbb{R}) \times L^\infty_t(\mathbb{R})$, we observe that the map $x \mapsto \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x)$ is continuous on $\overline{\mathcal{A}_t(v_t)}$ into $L^1_t(\mathbb{R})$. Therefore, we obtain that

$$\lim_{n \to \infty} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x^n + r^n) = \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(\overline{x}) \text{ in } L^1_t(\mathbb{R}).$$

Hence, there exists a subsequence $(\lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x^{n_k} + r^{n_k}))_{k \in N}$ that converges to $\lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(\overline{x})$ almost surely. Moreover, under (3.10), we observe that

$$\operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x) \leq \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x^{n_k} + r^{n_k}) \tag{3.11}$$

for all $k \in \mathbb{N}$. Therefore, after taking the limit of both sides on (3.11), we may conclude that

$$\operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x) \leq \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(\overline{x}).$$

Since $\overline{x} \in \overline{\mathcal{A}_t(v_t)}$ is arbitrary, by the definition of essential infimum, we obtain

$$\operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(x) \leq \operatorname*{ess\,inf}_{x \in \overline{\mathcal{A}_t(v_t)}} \lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(x),$$

as desired. $\square$

Under the conclusions of Lemma 3.1.2, $\overline{\mathcal{A}_t(v_t)}$ is a closed and decomposable subset of $L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})$. In the theory of random sets, it is well-known [16, Theorem 2.1.10] that a closed and decomposable subset of $L_t^p(\mathbb{R}^n)$, for some $n \in \mathbb{N}$ and $p \in [1, +\infty]$, can be characterized by the set of all p-integrable selectors of a random closed set. Moreover, this classical result can be generalized for the product space setting. The reader is referred to [16, Section 2] for more details on the relationship between decomposable sets and random sets. As a result of [16, Theorem 2.1.10], we guarantee the existence of a closed random set $F_t^{v_t} \colon \Omega \to 2^{\mathbb{R}^2}$ such that

$$S_t^{2,1}(F_t^{v_t}) \coloneqq \{x \in L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R}) \,|\, \omega \in \Omega,\, x(\omega) \in F_t^{v_t}(\omega) \text{ } \mathbb{P}\text{-a.s.}\} = \overline{\mathcal{A}_t(v_t)}, \quad (3.12)$$

where $S_t^{2,1}(F_t^{v_t})$ is the set of all $\mathcal{F}_t$-measurable selectors of $F_t^{v_t}$ in $L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})$. Moreover, we introduce the decomposed analogue of (3.3) for the multivalued random variable $F_t^{v_t}$, that is, we define $G_t^{v_t} \colon \Omega \to 2^{\mathbb{R}^2}$ such that

$$G_t^{v_t}(\omega) \coloneqq \{\mathcal{T}(z) \mid z \in F_t^{v_t}(\omega)\}, \quad \omega \in \Omega. \quad (3.13)$$

Then, as $\overline{\mathcal{M}_t(v_t)}$ is closed and decomposable by Lemma 3.1.4, by their constructions in (3.7) and (3.13), we observe that a similar analogue of (3.12) holds true for $G_t^{v_t}$ as well, that is, we have

$$S_t^{2,1}(G_t^{v_t}) \coloneqq \{x \in L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R}) \mid \mathbb{P}\{\omega \in \Omega \mid x(\omega) \in G_t^{v_t}(\omega)\} = 1\} = \overline{\mathcal{M}_t(v_t)}.$$

Furthermore, as $\overline{\mathcal{A}_t(v_t)}$ and $\overline{\mathcal{M}_t(v_t)}$ are convex subsets of $L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})$ under Lemma 3.1.2 and Lemma 3.1.4, by [16, Proposition 2.1.7], we conclude that both $F_t^{v_t}(\omega)$ and $G_t^{v_t}(\omega)$ are convex subsets of $\mathbb{R}^2$ for almost every $\omega \in \Omega$.

As the final step, we introduce the following decomposed analogues of the problems $(\mathrm{M}_t(v_t, \lambda_t))$ and $(\mathrm{A}_t(v_t, \rho_t))$:

$$\inf_{z \in F_t^{v_t}(\omega)} \lambda_t(\omega)^\mathsf{T} \begin{pmatrix} z_1 \\ z_2 - (z_1)^2 \end{pmatrix} \qquad (\mathrm{m}_t(v_t, \lambda_t, \omega)), \qquad\qquad \inf_{z \in F_t^{v_t}(\omega)} \rho_t(\omega)^\mathsf{T} z \qquad (\mathrm{a}_t(v_t, \rho_t, \omega))$$

for each $\omega \in \Omega$. Furthermore, by (3.13), we can rewrite the problem $\mathrm{m}_t(v_t, \lambda_t, \omega)$ equivalently in a linear form for each $\omega \in \Omega$, as

$$\inf_{z \in F_t^{v_t}(\omega)} \lambda_t(\omega)^\mathsf{T} \mathcal{T}(z) = \inf_{z \in \mathcal{T}[F_t^{v_t}(\omega)]} \lambda_t(\omega)^\mathsf{T} z = \inf_{z \in G_t^{v_t}(\omega)} \lambda_t(\omega)^\mathsf{T} z. \tag{3.14}$$

Our next lemma ensures that instead of working on the stochastic problems $(\mathrm{M}_t(v_t, \lambda_t))$ and $(\mathrm{A}_t(v_t, \rho_t))$, we may pursue our analysis on the more manageable decomposed problems $(\mathrm{m}_t(v_t, \lambda_t, \omega))$ and $(\mathrm{a}_t(v_t, \rho_t, \omega))$.

Lemma 3.1.6. *Let $t \in \mathbb{T} \setminus \{T\}$ and $v_t \in L_t^2(\mathbb{R})$. Let $\pi^\star \in \Phi_t(v_t)$ and define*

$$x^\star := (-\mathbb{E}_t(v_{t,T}^{\pi^\star}), \mathbb{E}_t((v_{t,T}^{\pi^\star})^2))^\mathsf{T}.$$

Then, the following equivalences hold.

(i) For every $\lambda_t \in L_t^2(\mathbb{R}) \times L_t^\infty(\mathbb{R}_+)$, $\pi^\star$ is the optimal solution of $(\mathrm{M}_t(v_t, \lambda_t))$ if and only if $x^\star(\omega)$ is the optimal solution of $(\mathrm{m}_t(v_t, \lambda_t, \omega))$ for almost every $\omega \in \Omega$.

(ii) For every $\rho_t \in L_t^2(\mathbb{R}_+) \times L_t^\infty(\mathbb{R}_+)$, $\pi^\star$ is the optimal solution of $(\mathrm{A}_t(v_t, \rho_t))$ if and only if $x^\star(\omega)$ is the optimal solution of $(\mathrm{a}_t(v_t, \rho_t, \omega))$ for almost every $\omega \in \Omega$.

Proof. As both results are similar in nature, to prevent repetition we only provide the proof of the first part. To that end, let us fix some $\lambda_t \in L_t^2(\mathbb{R}) \times L_t^\infty(\mathbb{R}_+)$ and suppose that $\pi^\star$ is an optimal solution of $(\mathrm{M}_t(v_t, \lambda_t))$. Then, by Lemma 3.1.5, we have $\lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(x^\star) \leq \lambda_t^\mathsf{T} \widehat{\mathcal{T}}_t(\overline{x})$ for all $\overline{x} \in \overline{\mathcal{A}_t(v_t)}$. Hence, as $x^\star \in \mathcal{A}_t(v_t) \subseteq \overline{\mathcal{A}_t(v_t)}$, by the monotonicity of expectation we obtain that

$$\mathbb{E}(\lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(x^\star)) = \inf_{\overline{x} \in \mathcal{A}_t(v_t)} \mathbb{E}(\lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(\overline{x})).$$

Let

$$\xi(\omega) := \inf_{z \in F_t^{v_t}(\omega)} \lambda_t(\omega)^{\mathsf{T}}\, \mathcal{T}(z), \quad \omega \in \Omega.$$

Then, by [16, Theorem 2.1.20], ξ is an $\mathcal{F}_t$-measurable random variable, and further we have

$$\inf_{\overline{x} \in \mathcal{A}_t(v_t)} \mathbb{E}(\lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(\overline{x})) = \mathbb{E}(\xi),$$

since we satisfy necessary assumptions, that is, we have $\mathbb{E}(\lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(x^\star)) < \infty$, and $z \mapsto \lambda_t(\omega)^{\mathsf{T}}\, \mathcal{T}(z)$ is continuous on $\mathbb{R}^2$ for every $\omega \in \Omega$. Therefore, we conclude that

$$\mathbb{E}(\xi) = \mathbb{E}(\lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(x^\star)). \tag{3.15}$$

Now, assume to the contrary that $x^\star(\omega)$ is not optimal solution of $(\mathrm{m}_t(v_t, \lambda_t, \omega))$ on an event with strictly positive probability, that is, we have $\mathbb{P}\{\xi < \lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(x^\star)\} > 0$. Note that, by definition we have $\mathbb{P}\{\xi \leq \lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(x^\star)\} = 1$. Then, under the monotonicity of expectation we arrive at a contradiction to (3.15) as

$$\mathbb{E}(\xi) < \mathbb{E}(\lambda_t^{\mathsf{T}} \widehat{\mathcal{T}}_t(x^\star)).$$

For the converse, let us suppose that $x^\star(\omega)$ is an optimal solution of $(\mathrm{m}_t(v_t, \lambda_t, \omega))$ for almost every $\omega \in \Omega$ and assume that $\pi^\star$ is not an optimal solution for $(\mathrm{M}_t(v_t, \lambda_t))$. Then, by Assumption 3.0.4, there exist an optimal solution $\pi \in \Phi_t(v_t)$ with $x = (-\mathbb{E}_t(v_T^\pi)\ , \mathbb{E}_t((v_T^\pi)^2))^{\mathsf{T}}$ and $B \in \mathcal{F}_t$ with $\mathbb{P}(B) > 0$, such that

$$\lambda_t^{\mathsf{T}}(\bar{\omega})\, \mathcal{T}(x(\omega)) < \lambda_t^{\mathsf{T}}(\bar{\omega})\, \mathcal{T}(x^\star(\omega))$$

for all $\omega \in B$. However, we arrive at a contradiction as $x^\star(\omega)$ is not optimal for the problem $(\mathrm{m}_t(v_t, \lambda_t, \omega))$ on a set with strictly positive probability. $\square$

3.2 The equivalence of $(\mathrm{M}_t(v_t, \lambda_t))$ and $(\mathrm{A}_t(v_t, \rho_t))$

With the results of Section 3.1, we are ready to announce the following theorem which indicates an essential equivalence, in terms of their optimal solutions, between the problems $(\mathrm{M}_t(v_t, \lambda_t))$ and $(\mathrm{A}_t(v_t, \rho_t))$ based on the choice of their weights λ_t and ρ_t.

Theorem 3.2.1. *Let $t \in \mathbb{T}\setminus\{T\}$, $v_t \in L^2_t(\mathbb{R})$, and $\pi^\star \in \Phi_t(v_t)$.*

(i) If $\pi^\star$ is an optimal solution of $(\mathrm{A}_t(v_t, \rho_t))$ for some $\rho_t \in L^2_t(\mathbb{R}_+) \times L^\infty_t(\mathbb{R}_+)$, then it is also an optimal solution of $(\mathrm{M}_t(v_t, \bar{\lambda}_t))$ with corresponding weight

$$\bar{\lambda}_t = \begin{pmatrix} \rho_{t,1} - 2\rho_{t,2}\mathbb{E}_t(v^{\pi^\star}_{t,T}) \\ \rho_{t,2} \end{pmatrix}. \tag{3.16}$$

(ii) If $\pi^\star$ is an optimal solution of $(\mathrm{M}_t(v_t, \lambda_t))$ for some $\lambda_t \in L^2_t(\mathbb{R}) \times L^\infty_t(\mathbb{R}_+)$, then it is also an optimal solution of $(\mathrm{A}_t(v_t, \bar{\rho}_t))$ with corresponding weight

$$\bar{\rho}_t = \begin{pmatrix} \lambda_{t,1} + 2\lambda_{t,2}\mathbb{E}_t(v^{\pi^\star}_{t,T}) \\ \lambda_{t,2} \end{pmatrix}. \tag{3.17}$$

Proof. For convenience, given $\gamma \in \mathbb{R}^2$, we define a nonlinear function $f(\cdot\,;\gamma)\colon \mathbb{R}^2 \to \mathbb{R}$ and a linear function $h(\cdot\,;\gamma)\colon \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x;\gamma) = \gamma^\mathsf{T}\,\mathcal{T}(x), \qquad h(x;\gamma) = \gamma^\mathsf{T} x.$$

We first prove (i). Let us fix some $\rho_t \in L^2_t(\mathbb{R}_+) \times L^\infty_t(\mathbb{R}_+)$ and suppose that $\pi^\star$ is an optimal solution of $(\mathrm{A}_t(v_t, \rho_t))$. Let us define $x^\star_t := (-\mathbb{E}_t(v^{\pi^\star}_{t,T}), \mathbb{E}_t((v^{\pi^\star}_{t,T})^2))^\mathsf{T}$ so that the definition of $\bar{\lambda}_t$ in (3.16) reads as

$$\bar{\lambda}_t = \begin{pmatrix} \rho_{t,1} + 2\rho_{t,2}x^\star_{t,1} \\ \rho_{t,2} \end{pmatrix}.$$

Then by Lemma 3.1.6(ii), $x^\star_t(\omega)$ is optimal for $(\mathrm{a}_t(v_t, \rho_t, \omega))$ for almost every $\omega \in \Omega$. Moreover, by Lemma 3.1.6(i), it suffices to show that $x^\star_t(\omega)$ is optimal for

$(\mathrm{m}_t(v_t, \bar{\lambda}_t, \omega))$ for almost every $\omega \in \Omega$. To that end, let us fix $\omega \in \Omega \setminus N$, where

$N := \{\omega \in \Omega \mid x_t^\star(\omega)$ is not optimal for $\mathrm{a}_t(v_t, \rho_t, \omega)$ or $G_t^{v_t}(\omega)$ is not a convex set$\}$.

Note that, both events that define N are $\mathcal{F}_t$-measurable null events, hence we have $N \in \mathcal{F}_t$ and $\mathbb{P}(N) = 0$. For readability, let us define

$$x^\star := x_t^\star(\omega), \qquad \bar{\lambda} := \bar{\lambda}_t(\omega), \qquad \rho := \rho_t(\omega).$$

We start by analyzing the gradient of the function $f(\cdot\,; \bar{\lambda})$. We observe that it is given at a point $x \in \mathbb{R}^2$ as

$$\nabla f(x; \bar{\lambda}) = \begin{pmatrix} \bar{\lambda}_1 - 2\bar{\lambda}_2 x_1 \\ \bar{\lambda}_2 \end{pmatrix} = \begin{pmatrix} \rho_1 + 2\rho_2 x_1^\star - 2\rho_2 x_1 \\ \rho_2 \end{pmatrix}.$$

Thus, at the point $x^\star$, the gradients of the functions $f(\cdot\,; \bar{\lambda})$ and $h(\cdot\,; \rho)$ are equal:

$$\nabla f(x^\star; \bar{\lambda}) = \rho = \nabla h(x^\star; \rho). \tag{3.18}$$

Under (3.18), in a small neighborhood around $x^\star$, both functions should behave similarly. We claim that $x^\star$ is a local minimum of the problem $(\mathrm{m}_t(v_t, \bar{\lambda}_t, \omega))$ in the sense that, there exists $\epsilon > 0$ such that $f(x^\star; \bar{\lambda}) \leq f(x; \bar{\lambda})$ for every $x \in B(x^\star, \epsilon) \cap F_t^{v_t}(\omega)$. To see that, suppose to the contrary for every $\epsilon > 0$, there exists $e \in B(0, \epsilon)$ such that $x^\star + e \in F_t^{v_t}(\omega)$ and $f(x^\star; \bar{\lambda}) > f(x^\star + e; \bar{\lambda})$. Then, by Taylor's theorem, we may write

$$f(x^\star + e; \bar{\lambda}) = f(x^\star; \bar{\lambda}) + \nabla f(x^\star; \bar{\lambda})^\mathsf{T} e + o(e) \tag{3.19}$$

where $o(e) := -2\bar{\lambda}_2 (e_1)^2$ is the nonlinear error term. Further, since $o(e)$ is of higher order, as $e \to 0$, it converges faster than the linear term $\nabla f(x^\star; \bar{\lambda})^\mathsf{T} e$, that is, there exist $\epsilon^\star > 0$ and $e^\star \in B(0, \epsilon^\star)$ such that

$$\left|\nabla f(x^\star; \bar{\lambda})^\mathsf{T} (e^\star)\right| > |o(e^\star)|\,. \tag{3.20}$$

Then, by combining our assumption that $f(x^\star; \bar{\lambda}) > f(x^\star + e^\star; \bar{\lambda})$ with (3.19), we obtain that $\nabla f(x^\star;\bar{\lambda})^\mathsf{T} e^\star + o(e^\star) < 0$. Hence, under (3.20), we have $\nabla f(x^\star;\bar{\lambda})^\mathsf{T}(e^\star) < 0$. Therefore, by (3.18),

$$\nabla h(x^\star;\rho)^\mathsf{T} e^\star = \nabla f(x^\star;\bar{\lambda})^\mathsf{T} e^\star < 0. \tag{3.21}$$

Finally, as $h(\cdot;\rho)$ is linear, similarly by Taylor's theorem, we can express $h(x^\star + e^\star;\rho)$ as

$$h(x^\star + e^\star;\rho) = h(x^\star;\rho) + \nabla h(x^\star;\rho)^\mathsf{T} e^\star.$$

Therefore, by (3.21), we obtain that $h(x^\star;\rho) > h(x^\star + e^\star;\rho)$, where $x^\star + e^\star \in F_t^{v_t}(\omega)$. Clearly, this contradicts the optimality of $x^\star$ for $(\mathrm{a}_t(v_t,\rho_t,\omega))$ due to $\omega \in \Omega \setminus N$. Thus, the claim holds, that is, $x^\star$ is locally optimal for $(\mathrm{m}_t(v_t,\bar{\lambda}_t,\omega))$.

As the final step, we generalize this local optimality of $x^\star$ to a global one. To that end, let $y^\star = \mathcal{T}(x^\star)$. By the feasibility of $x^\star$ for $(\mathrm{m}_t(v_t,\bar{\lambda}_t,\omega))$, we have $x^\star \in F_t^{v_t}(\omega)$. Hence, $y^\star \in G_t^{v_t}(\omega)$ by (3.13). Moreover, by the continuity of $\mathcal{T}^{-1}$, there exists $\delta > 0$ such that

$$\mathcal{T}^{-1}[B(y^\star,\delta)] \subseteq B(x^\star,\epsilon).$$

Hence, we indeed preserve the local optimality of $x^\star$ when we transform the problem into a linear setting, that is, for every $y \in B(y^\star,\delta) \cap G_t^{v_t}(\omega)$, we have $\mathcal{T}^{-1}(y) \in B(x^\star,\varepsilon) \cap F_t^{v_t}(\omega)$ so that

$$h(y^\star;\bar{\lambda}) = f(x^\star;\bar{\lambda}) \leq f(\mathcal{T}^{-1}(y);\bar{\lambda}) = h(y;\bar{\lambda}), \tag{3.22}$$

thanks to the local optimality of $x^\star$ and the definitions of f and h. Finally, to get a contradiction, let us assume that there exits $y \in G_t^{v_t}(\omega)$ such that $h(y;\bar{\lambda}) < h(y^\star;\bar{\lambda})$. Then, due to the linearity of $h(\cdot\,;\bar{\lambda})$ we have

$$\nabla h(y^\star;\bar{\lambda})^\mathsf{T}(y - y^\star) = h(y;\bar{\lambda}) - h(y^\star;\bar{\lambda}) < 0. \tag{3.23}$$

Moreover, we note that there exists small enough $\delta' \in (0,\delta)$ such that

$$\alpha := \frac{\delta'}{|y - y^\star|} \in (0,1).$$

In particular $e' := \alpha(y - y^\star) \in B(0, \delta)$. Furthermore, as $G_t^{v_t}(\omega)$ is convex and $y, y^\star \in G_t^{v_t}(\omega)$, we observe that $y^\star + e' \in B(y^\star, \delta) \cap G_t^{v_t}(\omega)$ since

$$y^\star + e' = y^\star + \alpha(y - y^\star) = \alpha y + (1 - \alpha) y^\star.$$

Finally, we write the image of $y^\star + e'$ under the linear objective $h(\cdot\,; \bar{\lambda})$ to obtain

$$h(y^\star + e'; \bar{\lambda}) = h(y^\star; \bar{\lambda}) + \nabla h(y^\star; \bar{\lambda})^\mathsf{T} e' = h(y^\star; \bar{\lambda}) + \alpha \nabla h(y^\star; \bar{\lambda})^\mathsf{T} (y - y^\star) \quad (3.24)$$

Thus, by (3.23) and (3.24), we conclude that $h(y^\star + e'; \bar{\lambda}) < h(y^\star; \bar{\lambda})$, which results in a contradiction with the local optimality of $y^\star$ in (3.22). Therefore, we conclude that $y^\star$ is an optimal solution of the problem (3.14), or equivalently, $x^\star$ is an optimal solution of the problem $(\mathrm{m}_t(v_t, \bar{\lambda}_t, \omega))$, which finishes the proof of (i).

To prove (ii), let us fix some $\lambda_t \in L_t^2(\mathbb{R}) \times L_t^\infty(\mathbb{R}_+)$ and suppose that $\pi^\star$ is an optimal solution of $(\mathrm{M}_t(v_t, \lambda_t))$. Let $x_t^\star := (-\mathbb{E}_t(v_{t,T}^{\pi^\star}), \mathbb{E}_t((v_{t,T}^{\pi^\star})^2))^\mathsf{T}$. Then, the definition of $\bar{\rho}_t$ in (3.17) reads as

$$\bar{\rho}_t = \begin{pmatrix} \lambda_{t,1} - 2\lambda_{t,2} x_{t,1}^\star \\ \lambda_{t,2} \end{pmatrix}. \quad (3.25)$$

Similar to the proof of (i), by Lemma 3.1.6, it is enough to show that $x_t^\star(\omega)$ is an optimal solution of $(\mathrm{a}_t(v_t, \bar{\rho}_t, \omega))$ for almost every $\omega \in \Omega$. To that end, let us fix $\omega \in \Omega \setminus N$, where

$$N := \{\omega \in \Omega \mid x_t^\star(\omega) \text{ is not optimal for } \mathrm{m}_t(v_t, \lambda_t, \omega)\} \in \mathcal{F}_t,$$

and $\mathbb{P}(N) = 0$. Furthermore, analogous to the first part, let us define

$$x^\star := x_t^\star(\omega), \qquad \lambda := \lambda_t(\omega), \qquad \bar{\rho} := \bar{\rho}_t(\omega).$$

To get a contradiction, let us assume the existence of $x \in F_t^{v_t}(\omega)$ such that

$h(x;\bar{\rho}) < h(x^\star;\bar{\rho})$. Then,

$$\bar{\rho}_1 x_1 + \bar{\rho}_2 x_2 < \bar{\rho}_1 x_1^\star + \bar{\rho}_2 x_2^\star.$$

Then, we substitute $\bar{\rho}$ by its definition in (3.25), and we observe

$$\lambda_1 x_1 - 2\lambda_2 x_1^\star x_1 + \lambda_2 x_2 < \lambda_1 x_1^\star - 2\lambda_2 (x_1^\star)^2 + \lambda_2 x_2^\star. \tag{3.26}$$

Moreover, we reorder the terms of inequality (3.26) as follows:

$$(\lambda_1 x_1 + \lambda_2 x_2) - (\lambda_1 x_1^\star - \lambda_2 (x_1^\star)^2 + \lambda_2 x_2^\star) < 2\lambda_2 x_1^\star x_1 - \lambda_2 (x_1^\star)^2$$

As the final step, we add the term $-\lambda_2(x_1)^2$ to both sides of the inequality and obtain

$$f(x;\lambda) - f(x^\star;\lambda) = (\lambda_1 x_1 - \lambda_2 (x_1)^2 + \lambda_2 x_2) - (\lambda_1 x_1^\star - \lambda_2 (x_1^\star)^2 + \lambda_2 x_2^\star) < -\lambda_2 (x_1^\star - x_1)^2.$$

Then, as $\lambda_2 \geq 0$ by definition, we observe that $f(x;\lambda) < f(x^\star;\lambda)$, where $x \in F_t^{v_t}(\omega)$. This indicates that $x^\star$ is not an optimal solution ($\mathrm{m}_t(v_t, \lambda_t, \omega)$), which results in a contradiction. □

The reader can notice that the equivalent weights in Theorem 3.2.1 depend on the terminal values of the optimal portfolios. Therefore, the uniqueness of the equivalent weights, and their potential effects on time-consistency fully depend on the uniqueness of the optimal terminal values. Therefore, our next series of results establish necessary conditions for obtaining the desired uniqueness.

Lemma 3.2.2. *Let $x \in L_T^2(\mathbb{R}^2)$ with $\mathbb{P}\{x_1 \neq x_2\} > 0$, and let $\alpha \in (0,1)$. Then,*

$$\mathbb{E}((\alpha x_1 + (1-\alpha)x_2)^2) < \alpha \mathbb{E}((x_1)^2) + (1-\alpha)\mathbb{E}((x_2)^2). \tag{3.27}$$

In particular, if $\mathbb{E}(x_1) = \mathbb{E}(x_2)$, then

$$\operatorname{Var}(\alpha x_1 + (1-\alpha)x_2) < \alpha \operatorname{Var}(x_1) + (1-\alpha)\operatorname{Var}(x_2).$$

Proof. By the strict convexity of the square function, we have

$$(\alpha x_1(\omega) + (1-\alpha)x_2(\omega))^2 < \alpha x_1(\omega)^2 + (1-\alpha)x_2(\omega)^2$$

for every $\omega \in \{x_1 \neq x_2\}$. Moreover, we have nonstrict inequality almost surely, that is,

$$(\alpha x_1 + (1-\alpha)x_2)^2 \leq \alpha(x_1)^2 + (1-\alpha)(x_2)^2.$$

Then, as $\mathbb{P}\{x_1 \neq x_2\} > 0$, under the monotonicity and linearity of expectation, (3.27) follows. Next, suppose that $\mathbb{E}(x_1) = \mathbb{E}(x_2)$ and let $\tilde{x}_1 := x_1 - \mathbb{E}(x_1)$, $\tilde{x}_2 := x_2 - \mathbb{E}(x_2)$. Clearly, $\mathbb{P}\{\tilde{x}_1 \neq \tilde{x}_2\} > 0$ holds. Thus, applying (3.27) to $\tilde{x}_1$ and $\tilde{x}_2$ gives the inequality for variance. □

Proposition 3.2.3. *Let $v_0 \in \mathbb{R}_+$ and $\lambda_0 \in (\mathbb{R} \setminus \{0\}) \times \mathbb{R}_{++}$. Moreover, let $\pi^1, \pi^2 \in \Phi_0(v_0)$ be two optimal solutions of $(\mathrm{M}_0(v_0, \lambda_0))$. Then, they admit the same terminal value, that is, $v_T^{\pi^1} = v_T^{\pi^2}$.*

Proof. As an initial step, we claim that $\mathrm{Var}(v_T^{\pi^1}) = \mathrm{Var}(v_T^{\pi^2})$. Let us assume to the contrary that $\mathrm{Var}(v_T^{\pi_1}) \neq \mathrm{Var}(v_T^{\pi_2})$. In particular, we must have

$$\mathbb{P}\{v_T^{\pi^1} - \mathbb{E}(v_T^{\pi^1}) \neq v_T^{\pi^2} - \mathbb{E}(v_T^{\pi^2})\} > 0. \tag{3.28}$$

Let $\alpha \in (0,1)$ and consider $\pi := \alpha\pi^1 + (1-\alpha)\pi^2$. We have $\pi \in \Phi_0(v_0)$ as the feasible region is convex. Then, by the definition of value process in (2.1), we have

$$v_T^{\pi} = v_T^{\alpha\pi^1 + (1-\alpha)\pi^2} = \alpha v_T^{\pi^1} + (1-\alpha)v_T^{\pi^2}.$$

so that

$$\mathrm{Var}(v_T^{\pi}) = \mathrm{Var}(\alpha v_T^{\pi^1} + (1-\alpha)v_T^{\pi^2}) = \mathrm{Var}\left(\alpha(v_T^{\pi^1} - \mathbb{E}(v_T^{\pi^1})) + (1-\alpha)(v_T^{\pi^2} - \mathbb{E}(v_T^{\pi^2}))\right).$$

Then, by (3.28) and Lemma 3.2.2, we conclude that

$$\mathrm{Var}(v_t^{\pi}) < \alpha\,\mathrm{Var}(v_T^{\pi^1}) + (1-\alpha)\,\mathrm{Var}(v_T^{\pi^2}). \tag{3.29}$$

Furthermore, by the linearity of expectation,

$$\mathbb{E}(v_T^{\pi}) = \mathbb{E}(\alpha v_T^{\pi^1} + (1-\alpha)v_T^{\pi^2}) = \alpha\mathbb{E}(v_T^{\pi^1}) + (1-\alpha)\mathbb{E}(v_T^{\pi^2}). \tag{3.30}$$

Therefore, as λ is strictly positive in the second component, combining (3.29) and (3.30) gives

$$\lambda_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi}) \\ \mathrm{Var}(v_T^{\pi}) \end{pmatrix} < \alpha\lambda_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi^1}) \\ \mathrm{Var}(v_T^{\pi^1}) \end{pmatrix} + (1-\alpha)\lambda_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi^2}) \\ \mathrm{Var}(v_T^{\pi^2}) \end{pmatrix}. \tag{3.31}$$

Then, as π^1 and π^2 are optimal solutions of $(\mathrm{M}_0(v_0, \lambda_0))$, hence sharing the same objective value, we rearrange (3.31) to get

$$\lambda_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi}) \\ \mathrm{Var}(v_T^{\pi}) \end{pmatrix} < \lambda_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi^1}) \\ \mathrm{Var}(v_T^{\pi^1}) \end{pmatrix},$$

which is a contradiction to the optimality of π^1. Thus, we conclude that $\mathrm{Var}(v_T^{\pi^1}) = \mathrm{Var}(v_T^{\pi^2})$. Moreover, as λ is non-zero in both components, we conclude that $\mathbb{E}(v_T^{\pi^1}) = \mathbb{E}(v_T^{\pi^2})$ as well. Therefore, by Theorem 3.2.1, both π^1 and π^2 are optimal solutions of the problem $\mathrm{A}_0(v_0, \rho_0)$, where

$$\rho_0 := \begin{pmatrix} \lambda_{0,1} + 2\lambda_{0,2}\mathbb{E}_t(v_T^{\pi^1}) \\ \lambda_{0,2} \end{pmatrix}.$$

Finally, to get a contradiction, let us assume that $\mathbb{P}\{v_T^{\pi^1} \neq v_T^{\pi^2}\} > 0$. Then, by Lemma 3.2.2, we observe that

$$\mathbb{E}((v_T^{\pi})^2) = \mathbb{E}((\alpha v_T^{\pi^1} + (1-\alpha)v_T^{\pi^2})^2) < \alpha\mathbb{E}((v_T^{\pi^1})^2) + (1-\alpha)\mathbb{E}((v_T^{\pi^2})^2). \tag{3.32}$$

Then, following a similar argument as for the mean-variance problem, by combining (3.30) and (3.32), we get

$$\rho_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi}) \\ \mathbb{E}((v_T^{\pi})^2) \end{pmatrix} < \alpha\rho_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi^1}) \\ \mathbb{E}((v_T^{\pi^1})^2) \end{pmatrix} + (1-\alpha)\rho_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi^2}) \\ \mathbb{E}((v_T^{\pi^2})^2) \end{pmatrix}$$

since $\rho_{0,2} = \lambda_{0,2} > 0$ by assumption. Moreover, as π^1 and π^2 are both optimal solutions of $(\mathrm{A}_0(v_0,\rho))$, they share the same objective value. Therefore, since $\mathbb{E}(v_T^{\pi^1}) = \mathbb{E}(v_T^{\pi^2})$, we observe that $\mathbb{E}((v_T^{\pi^1})^2) = \mathbb{E}((v_T^{\pi^2})^2)$. Finally, we obtain

$$\rho_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi}) \\ \mathbb{E}((v_T^{\pi})^2) \end{pmatrix} < \rho_0^{\mathsf{T}} \begin{pmatrix} -\mathbb{E}(v_T^{\pi^1}) \\ \mathbb{E}((v_T^{\pi^1})^2) \end{pmatrix},$$

which contradicts the optimality of π^1 for $(\mathrm{A}_0(v_0,\rho_0))$. Hence, $v_T^{\pi^1} = v_T^{\pi^2}$. □

3.3 Time-consistency of the mean-second moment problem

After establishing the equivalence between the mean-variance and mean-second moment problems, we turn our interest to the time-consistency of the mean-second moment problem, as provided by the next result. Note that a fixed weight is already sufficient to establish time-consistency in this case.

Proposition 3.3.1. *Let $v_0 \in \mathbb{R}_+$ and $\rho_0 \in \mathbb{R}^2_+$. Then, the family $(\mathrm{A}_t(\cdot,\rho_0))_{t\in\mathbb{T}\setminus\{T\}}$ of optimization problems with initial wealth v_0 is time-consistent, that is, for every optimal solution $\pi^\star$ of $(\mathrm{A}_0(v_0,\rho_0))$ and $t \in \mathbb{T}\setminus\{T\}$, $(\pi^\star_s)_{s\geq t+1}$ is optimal for $(\mathrm{A}_t(v_t^{\pi^\star},\rho_0))$.*

Proof. Let $\pi^\star$ be an optimal solution of $(\mathrm{A}_0(v_0,\rho_0))$. To get a contradiction, assume that there exists $t \in \mathbb{T}\setminus\{0,T\}$ such that $(\pi^\star_s)_{s\geq t+1}$ is not optimal for $(\mathrm{A}_t(v_t^{\pi^\star},\rho_0))$. Then, by Assumption 3.0.4, there exist an optimal solution $\pi \in \Phi_t(v_t^{\pi^\star})$ of $(A_t(v_t^{\pi^\star},\rho_0))$ and an event $B \in \mathcal{F}_t$ with $\mathbb{P}(B) > 0$ such that

$$-\rho_{0,1}\mathbb{E}_t(v_{t,T}^{\pi})(\omega)+\rho_{0,2}\mathbb{E}_t((v_{t,T}^{\pi})^2)(\omega) < -\rho_{0,1}\mathbb{E}_t(v_{t,T}^{\pi^\star})(\omega)+\rho_{0,2}\mathbb{E}_t((v_{t,T}^{\pi^\star})^2)(\omega) \quad (3.33)$$

for every $\omega \in B$. On the other hand, as π is optimal for $(\mathrm{A}_t(v_t^{\pi^\star},\rho_0))$, we have

$$-\rho_{0,1}\mathbb{E}_t(v_{t,T}^{\pi}) + \rho_{0,2}\mathbb{E}_t((v_{t,T}^{\pi})^2) \leq -\rho_{0,1}\mathbb{E}_t(v_{t,T}^{\pi^\star}) + \rho_{0,2}\mathbb{E}_t((v_{t,T}^{\pi^\star})^2). \quad (3.34)$$

Now, we consider $\bar{\pi} := ((\pi^\star_s)_{s\leq t},\pi)$. By construction, we have $\bar{\pi} \in \Phi_0(v_0)$ and

$v_T^{\bar{\pi}} = v_{t,T}^{\pi}$. Finally, we take the expectation of both sides of (3.34) and, by the strict monotonicity of expectation due to (3.33), we obtain

$$\mathbb{E}(-\rho_{0,1}\mathbb{E}_t(v_{t,T}^{\bar{\pi}}) + \rho_{0,2}\mathbb{E}_t((v_{t,T}^{\bar{\pi}})^2)) < \mathbb{E}(-\rho_{0,1}\mathbb{E}_t(v_{t,T}^{\pi^\star}) + \rho_{0,2}\mathbb{E}_t((v_{t,T}^{\pi^\star})^2)).$$

After making the substitutions $v_T^{\bar{\pi}} = v_{t,T}^{\pi}$ and $v_T^{\pi^\star} = v_{t,T}^{\pi^\star}$, by linearity and tower property, we arrive at

$$-\rho_{0,1}\mathbb{E}(v_T^{\bar{\pi}}) + \rho_{0,2}\mathbb{E}((v_T^{\bar{\pi}})^2) < -\rho_{0,1}\mathbb{E}(v_T^{\pi^\star}) + \rho_{0,2}\mathbb{E}((v_T^{\pi^\star})^2),$$

which results in a contradiction with $\pi^\star$ being optimal for $(\mathrm{A}_0(v_0, \rho_0))$. $\square$

The next theorem is the main result about time-consistency achieved by using a stochastic weight process in the mean-variance problem.

Theorem 3.3.2. *Let $v_0 \in \mathbb{R}_+$ and $\lambda_0 \in \mathbb{R} \times \mathbb{R}_+$. Suppose that $\pi^\star$ is an optimal solution of $(\mathrm{M}_0(v_0, \lambda_0))$ and let $t \in \mathbb{T}\backslash\{T\}$. Then, the process $(\pi_s^\star)_{s \geq t+1} \in \Phi_t(v_t^{\pi^\star})$ is an optimal solution of $(\mathrm{M}_t(v_t^{\pi^\star}, \lambda_t))$, where*

$$\lambda_t := \begin{pmatrix} \lambda_{0,1} + 2\lambda_{0,2}(\mathbb{E}(v_T^{\pi^\star}) - \mathbb{E}_t(v_T^{\pi^\star})) \\ \lambda_{0,2} \end{pmatrix}.$$

Furthermore, if $\lambda_0 \in (\mathbb{R}\setminus\{0\}) \times \mathbb{R}_{++}$, then the family $(\mathrm{M}_t(\cdot, \lambda_t))_{t\in\mathbb{T}\backslash\{T\}}$ of mean-variance problems with initial wealth v_0 is time-consistent in the sense of Definition 3.0.3.

Proof. By Theorem 3.2.1(ii), $\pi^\star$ is also an optimal solution of $(\mathrm{A}_0(v_0, \rho_0))$, where

$$\rho_0 := \begin{pmatrix} \lambda_{0,1} + 2\lambda_{0,2}\mathbb{E}(v_T^{\pi^\star}) \\ \lambda_{0,2} \end{pmatrix}.$$

Moreover, by Proposition 3.3.1, for every $t \in \mathbb{T}\backslash\{T\}$, the feasible portfolio $(\pi_s^\star)_{s\geq t+1}$ is an optimal solution of $\mathrm{A}_t(v_t^{\pi^\star}, \rho_0)$. Then, by Theorem 3.2.1(i), for every $t \in \mathbb{T}\backslash\{T\}$, the feasible portfolio $(\pi_s^\star)_{s\geq t+1}$ is an optimal solution of

$(\mathrm{M}_t(v_t^{\pi^\star}, \lambda_t))$, where

$$\lambda_t := \begin{pmatrix} \rho_{0,1} - 2\rho_{0,2}\mathbb{E}_t(v_T^{\pi^\star}) \\ \rho_{0,2} \end{pmatrix} = \begin{pmatrix} \lambda_{0,1} + 2\lambda_{0,2}(\mathbb{E}(v_T^{\pi^\star}) - \mathbb{E}_t(v_T^{\pi^\star})) \\ \lambda_{0,2} \end{pmatrix}.$$

Now, let us suppose further that $\lambda_0 \in (\mathbb{R}\backslash\{0\}) \times \mathbb{R}_{++}$. Then, by Proposition 3.2.3, given two arbitrary optimal solutions π^1, $\pi^2 \in \Phi_0(v_0)$ of $(\mathrm{M}_0(v_0, \lambda_0))$, we have $v_T^{\pi^1} = v_T^{\pi^2}$. Hence, we conclude that the conditional expectation processes of their respective terminal values are equal, that is,

$$\mathbb{E}_t(v_T^{\pi^1}) = \mathbb{E}_t(v_T^{\pi^2}), \quad t \in \mathbb{T},$$

and as a result,the time-consistent weight processes corresponding to π^1 and π^2 are equal, which is $(\lambda_t)_{t\in\mathbb{T}\backslash\{T\}}$. Therefore, for every optimal solution π' of $(\mathrm{M}_0(v_0, \lambda_0))$, $(\pi'_s)_{s\geq t+1}$ is an optimal solution of $(\mathrm{M}_t(v_t^{\pi'}, \lambda_t))$. Hence, by the previous step, we conclude that $(\mathrm{M}_t(\cdot, \lambda_t))_{t\in\mathbb{T}\backslash\{T\}}$ becomes time-consistent. $\square$

Remark 3.3.3. The reader can notice that throughout this chapter, we avoid restricting the weight for the mean-variance problem to have nonnegative first component, even though doing so is the standard to recover the classical case where one aims to maximize expected return while minimizing variance. The main reason for us to seek such generality is mainly due to the results of Theorem 3.3.2. Indeed, the time-consistent weight process, depending on the initial risk aversion and the behavior of the market, can possibly have a negative first component. The reader is referred to [13] for a more detailed analysis of this phenomena, where the authors investigate the time-consistency of the mean-variance problem when the weight process is constrained to take only nonnegative values.

Chapter 4

A scalar dynamic programming principle

After establishing the time-consistency of the mean-second moment problem with a fixed weight process in Proposition 3.3.1, the next natural step is to investigate its associated Bellman's equation. For this purpose, for each $t \in \mathbb{T}\setminus\{T\}$, we define the value function $V_t\colon L_t^2(\mathbb{R}) \to L_t^1(\mathbb{R})$ of $(\mathrm{A}_t(\cdot,\rho_0))$ with a fixed weight $\rho_0 \in \mathbb{R}_+^2$ directly as the optimal objective value, that is,

$$V_t(v_t) := \operatorname*{ess\,inf}_{\pi\in\Phi_t(v_t)} \left(-\rho_{0,1}\mathbb{E}_t(v_{t,T}^{\pi}) + \rho_{0,2}\mathbb{E}_t((v_{t,T}^{\pi})^2)\right), \quad v_t \in L_t^2(\mathbb{R}). \tag{4.1}$$

Moreover, as the terminal condition, we define the value function at time T by

$$V_T(v_T) := -\rho_{0,1}v_T + \rho_{0,2}(v_T)^2, \quad v_T \in L_T^2(\mathbb{R}). \tag{4.2}$$

Under this construction, we have the following dynamic programming principle.

Proposition 4.0.1. *Let $t \in \mathbb{T}\setminus\{T\}$ and $v_t \in L_t^2(\mathbb{R})$. Then, it holds*

$$V_t(v_t) = \operatorname*{ess\,inf}_{\substack{\pi\in L_t^{\infty}(\mathbb{R}^d):\\ \pi^{\mathsf{T}}S_t=v_t\\ \varphi_t(\pi)\geq 0}} \mathbb{E}_t\left(V_{t+1}(v_{t,t+1}^{\pi})\right). \tag{4.3}$$

Proof. First, let us consider the last period from $T-1$ to T. Then, by the linearity of conditional expectation, and the definition of the value function in (4.1) and (4.2), it is clear that

$$\begin{aligned}
&\operatorname*{ess\,inf}_{\substack{\pi\in L^\infty_{T-1}(\mathbb{R}^d):\\ \pi^\intercal S_{T-1}=v_{T-1}\\ \varphi_{T-1}(\pi)\geq 0}} \mathbb{E}_{T-1}\left(V_T(v^\pi_{T-1,T})\right)\\
&\quad= \operatorname*{ess\,inf}_{\substack{\pi\in L^\infty_{T-1}(\mathbb{R}^d):\\ \pi^\intercal S_{T-1}=v_{T-1}\\ \varphi_{T-1}(\pi)\geq 0}} \left(-\rho_{0,1}\mathbb{E}_{T-1}\left(v^\pi_{T-1,T}\right)+\rho_{0,2}\mathbb{E}_{T-1}\left((v^\pi_{T-1,T})^2\right)\right) = V_{T-1}(v_{T-1}).
\end{aligned}$$

Now let us suppose that $t\in\mathbb{T}\backslash\{T-1,T\}$, and consider an arbitrary $v_{t+1}\in L^2_{t+1}(\mathbb{R})$. By the monotonicity of conditional expectation, we have

$$\mathbb{E}_t\left(V_{t+1}(v_{t+1})\right)\leq \mathbb{E}_t\left(-\rho_{0,1}\mathbb{E}_{t+1}(v^\pi_{t+1,T})+\rho_{0,2}\mathbb{E}_{t+1}((v^\pi_{t+1,T})^2)\right)$$

for every $\pi\in\Phi_{t+1}(v_{t+1})$. Therefore, by linearity and tower property,

$$\mathbb{E}_t\left(V_{t+1}(v_{t+1})\right)\leq \operatorname*{ess\,inf}_{\pi\in\Phi_{t+1}(v_{t+1})}\left(-\rho_{0,1}\mathbb{E}_t(v^\pi_{t+1,T})+\rho_{0,2}\mathbb{E}_t((v^\pi_{t+1,T})^2)\right).$$

However, by Assumption 3.0.4, we observe that the essential infimum in the definition of $V_{t+1}(v_{t+1})$ is attained at some $\pi^\star\in\Phi_{t+1}(v_{t+1})$. In particular, we have

$$\mathbb{E}_t(V_{t+1}(v_{t+1}))=\mathbb{E}_t\left((-\rho_{0,1}\mathbb{E}_{t+1}(v^{\pi^\star}_{t+1,T})+\rho_{0,2}\mathbb{E}_{t+1}((v^{\pi^\star}_{t+1,T})^2))\right).$$

Hence, we conclude that

$$\mathbb{E}_t(V_{t+1}(v_{t+1}))=\operatorname*{ess\,inf}_{\pi\in\Phi_{t+1}(v_{t+1})}\left(-\rho_{0,1}\mathbb{E}_t(v^\pi_{t+1,T})+\rho_{0,2}\mathbb{E}_t((v^\pi_{t+1,T})^2)\right). \tag{4.4}$$

Next, we separate the feasible region for $V_t(v_t)$ into two parts and apply (4.4) to obtain

$$
\begin{aligned}
V_t(v_t) &= \operatorname*{ess\,inf}_{\substack{\pi\in L_t^\infty(\mathbb{R}^d):\\ \pi^\intercal S_t=v_t\\ \varphi_t(\pi)\geq 0}} \operatorname*{ess\,inf}_{\tilde{\pi}\in\Phi_{t+1}(v^\pi_{t,t+1})} \left(-\rho_{0,1}\mathbb{E}_t(v^{\tilde{\pi}}_{t+1,T})+\rho_{0,2}\mathbb{E}_t((v^{\tilde{\pi}}_{t+1,T})^2)\right) \\
&= \operatorname*{ess\,inf}_{\substack{\pi\in L_t^\infty(\mathbb{R}^d):\\ \pi^\intercal S_t=v_t\\ \varphi_t(\pi)\geq 0}} \mathbb{E}_t(V_{t+1}(v^\pi_{t,t+1})).
\end{aligned}
$$

Hence, (4.3) follows. □

Although the dynamic programming principle in Proposition 4.0.1 can be helpful when studying the dynamic mean-second moment problem for its own sake, its usefulness is questionable for the mean-variance problem. Indeed, under Theorem 3.2.1, for a given initial weight $\lambda_0 \in \mathbb{R}\times\mathbb{R}_+$ for the mean-variance problem, one first neeeds to calculate the corresponding weight $\bar{\rho}_0 \in \mathbb{R}^2$ for the mean second-moment problem given by (3.17). But this would require knowing the optimal wealth for the initial mean-variance problem $(\mathrm{M}_0(v_0,\lambda_0))$ beforehand, in which case the dynamic programming principle in Proposition 4.0.1 becomes irrelevant.

Fortunately, we can cope up with this issue by considering all possible pairs $(\rho_0, \mathbb{E}(v_T^{\pi^\star}))$ simultaneously, where $\pi^\star$ is an optimal solution of $(\mathrm{A}_0(v_0,\rho_0))$. That is, if we can come up with a methodology that yields all such pairs dynamically, then we may find $\bar{\rho}_0$ corresponding to the initial weight $\lambda_0 \in \mathbb{R}\times\mathbb{R}_+$ of the mean-variance problem. To achieve this goal, in Chapter 5, we will reformulate the mean-second moment problem as a biobjective optimization problem for which a dynamic programming principle is available in an extended sense.

Chapter 5

Set-valued dynamic programming

In this chapter, we prove a set-valued analogue of the scalar dynamic programming principle in Proposition 4.0.1 for the biobjective mean-second moment problem.

Recently, [14] have introduced a new methodology for the dynamic mean-risk problem, where the risk is measured by a time-consistent dynamic coherent risk measure. They pass to a biobjective formulation of this problem and define a set-valued value function for this formulation. Accordingly, they obtain a set-valued dynamic programming principle.

Here, we follow a similar approach for the mean-second moment problem. To that end, let us first introduce the biobjective formulation of the mean-second moment problem at time $t \in \mathbb{T} \setminus \{T\}$ with starting wealth $v_t \in L_t^2(\mathbb{R})$ as

$$\begin{aligned} \text{minimize} \quad & \begin{pmatrix} -\mathbb{E}_t(v_{t,T}^\pi) \\ \mathbb{E}_t((v_{t,T}^\pi)^2) \end{pmatrix} \quad \text{with respect to} \quad \leq = \leq_{L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)} \\ \text{subject to} \quad & \pi \in \Phi_t(v_t). \end{aligned} \tag{$\mathrm{C}_t(v_t)$}$$

Note that this is a *vector optimization* problem where the vector-valued objective is to be minimized with respect to the partial order $\leq = \leq_{L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)}$, which corresponds to the componentwise almost sure ordering on $L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})$. We use the concepts of minimal element and efficient solution from vector optimization theory, which we recall in Section 5.1.

5.1 Solution concepts and equivalence with scalar counterparts

We consider the partially ordered vector space $(L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R}), \leq_{L^2_t(\mathbb{R}_+)\times L^1_t(\mathbb{R}_+)})$ and recall the notion of minimal element for a subset of this space.

Definition 5.1.1. *Let $D \subseteq L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})$ be a nonempty set and let $\bar{x} \in D$.*

(i) $\bar{x}$ is called a minimal *element of D if $(\bar{x} - L^2_t(\mathbb{R}_+) \times L^1_t(\mathbb{R}_+)) \cap D = \{\bar{x}\}$, or equivalently, if there exists no $x \in D \setminus \{\bar{x}\}$ such that $x \leq \bar{x}$.*

(ii) $\bar{x}$ is called a weakly minimal *element of D if $(\bar{x} - \mathrm{qi}(L^2_t(\mathbb{R}_+) \times L^1_t(\mathbb{R}_+))) \cap D = \emptyset$, or equivalently, if there exists no $x \in D$ such that $x \in \bar{x} - \mathrm{qi}(L^2_t(\mathbb{R}_+) \times L^1_t(\mathbb{R}_+))$.*

We note that, the above definition of weakly minimal element is slightly different from the classical one in the literature which uses usual interior instead of quasi-interior (see [17, Section 4]), and it was introduced by [18]. In our case, we prefer working with quasi-interior since both $\mathrm{int}(L^2_t(\mathbb{R}_+))$ and $\mathrm{int}(L^1_t(\mathbb{R}_+))$ are empty unless they are finite-dimensional. Hence, the classical definition becomes trivial in infinite-dimensional settings. Moreover, we follow the convention in [18, Remark 4] and keep using the terminology *weakly minimal*, as this definition is in fact a generalization of the classical one.

Note that the image set, that is, the set of all objective vectors, of the biobjective problem $(\mathrm{C}_t(v_t))$ is already established in Chapter 3 as $\mathcal{A}_t(v_t)$. Next, we recall the definition of efficient and weakly efficient solutions for $(\mathrm{C}_t(v_t))$, which are standard in vector optimization.

Definition 5.1.2. *[17, Definition 11.5] Let $t \in \mathbb{T}\backslash\{T\}$ and $v_t \in L_t^2(\mathbb{R})$. A feasible solution $\pi \in \Phi_t(v_t)$ is called an* (weakly) efficient solution *of* $(\mathrm{C}_t(v_t))$ *if its image $(-\mathbb{E}_t(v_{t,T}^{\pi}), \mathbb{E}_t((v_{t,T}^{\pi})^2))^{\mathsf{T}}$ is a (weakly) minimal element of $\mathcal{A}_t(v_t)$.*

We note that, in order to make use of [18, Corollary 9], which we utilize for the characterization of the weakly efficient solutions of $(\mathrm{C}_t(v_t))$, we need the following assumption, which is necessary for the separation arguments within [18, Corollary 9].

Assumption 5.1.3. *For every $t \in \mathbb{T}$ and $v_t \in L_t^2(\mathbb{R})$ it holds that $\mathcal{A}_t(v_t) + \mathrm{qi}\,(L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)) = \mathrm{qi}\,(\mathcal{A}_t(v_t) + L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+))$.*

Remark 5.1.4. We emphasize that, [18] follow a similar path in terms of the assumption we made, and further they conjecture that the condition on Assumption 5.1.3 holds as long as both the image set and the ordering cone of the vector optimization problem are convex sets. However, more recently [19] came up with a counter example to the conjecture of [18] thereby, we note that the assumption made here is non-trivial.

Remark 5.1.5. The quasi-interior of the positive cone $L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)$ is $L_t^2(\mathbb{R}_{++}) \times L_t^1(\mathbb{R}_{++})$ [20], and this is completely analogous to the finite dimensional case, where we have $\mathrm{int}(\mathbb{R}_+^n) = \mathbb{R}_{++}^n$. Therefore, this particular observation justifies the use of quasi-interior in our setting. Moreover, it indicates that Definition 5.1.1 and Definition 5.1.2 provide adequate extensions of the classical theory for the infinite-dimensional vector space $L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})$.

Then, under [18, Corollary 9], we conclude that $x^\star$ is a weakly minimal element of $\mathcal{A}_t(v_t)$ if and only if there exists some nonzero $\rho_t \in L_t^2(\mathbb{R}_+) \times L_t^\infty(\mathbb{R}_+)$ such that

$$\mathbb{E}(\rho_t^{\mathsf{T}} x^\star) \leq \mathbb{E}(\rho_t^{\mathsf{T}} x) \tag{5.1}$$

for every $x \in \mathcal{A}_t(v_t)$. Moreover, by Lemma 3.1.5, we can extend (5.1) for every $x \in \overline{\mathcal{A}_t(v_t)}$ as

$$\inf_{x \in \mathcal{A}_t(v_t)} \mathbb{E}(\rho_t^{\mathsf{T}} x) = \mathbb{E}\left(\operatorname*{ess\,inf}_{x \in \mathcal{A}_t(v_t)} \rho_t^{\mathsf{T}} x\right) = \mathbb{E}\left(\operatorname*{ess\,inf}_{x \in \overline{\mathcal{A}_t(v_t)}} \rho_t^{\mathsf{T}} x\right) = \inf_{x \in \overline{\mathcal{A}_t(v_t)}} \mathbb{E}(\rho_t^{\mathsf{T}} x).$$

Therefore, under [16, Theorem 2.1.20] we have

$$\inf_{x\in\overline{\mathcal{A}_t(v_t)}} \mathbb{E}(\rho_t^\mathsf{T} x) = \int_\Omega \inf_{x\in F_t^{v_t}(\omega)} \left(\rho_t(\omega)^\mathsf{T} x\right)\mathbb{P}(d\omega).$$

Hence, we conclude that $x^\star$ is a weakly minimal element of $\mathcal{A}_t(v_t)$ if and only if we can find some nonzero vector $\rho_t \in L_t^2(\mathbb{R}_+)\times L_t^\infty(\mathbb{R}_+)$ for which $x^\star(\omega)$ is an optimal solution of $(\mathrm{a}_t(v_t,\rho_t,\omega))$ for almost every $\omega\in\Omega$. Equivalently, by Lemma 3.1.6, we observe that $x^\star$ is a weakly minimal element of $\mathcal{A}_t(v_t)$ if and only if it is an optimal solution of $(\mathrm{A}_t(v_t,\rho_t))$ for some nonzero $\rho_t \in L_t^2(\mathbb{R}_+)\times L_t^\infty(\mathbb{R}_+)$.

The above observations also yield an equivalence between the weakly efficient solutions of $(\mathrm{C}_t(v_t))$, and the optimal solutions of the scalar mean-second moment problems of the form $(\mathrm{A}_t(v_t,\rho_t))$, where $\rho_t \in L_t^2(\mathbb{R}_+)\times L_t^\infty(\mathbb{R}_+)$. In fact, for every weakly efficient solution $\pi^\star$ of $(\mathrm{C}_t(v_t))$, there is a corresponding weight $\rho_t \in L_t^2(\mathbb{R}_+)\times L_t^1(\mathbb{R}_+)$ for which $\pi^\star$ is an optimal solution of the scalar problem $(\mathrm{A}_t(v_t,\rho_t))$, and vice versa. The main benefit of passing to the vector optimization setup is that we are fully able to study the dynamic behavior of mean-second moment problem, free from the choice of a particular weight. This is explored in detail in the next section.

5.2 Set-valued Bellman's equation

In order to establish a set-valued Bellman principle for the problem $(\mathrm{C}_t(v_t))$, we utilize the so-called *upper images* as our candidate value function. To that end, for $t\in\mathbb{T}$ and $v_t\in L_t^2(\mathbb{R})$, we define the *upper image* of $(\mathrm{C}_t(v_t))$ by

$$\mathcal{P}_t(v_t) \coloneqq \mathrm{cl}\left(\mathcal{A}_t(v_t) + L_t^2(\mathbb{R}_+)\times L_t^1(\mathbb{R}_+)\right).$$

Remark 5.2.1. Note that, in general the upper image $\mathcal{P}_t(v_t)$ is a strict superset of $\overline{\mathcal{A}_t(v_t)}$ defined in Lemma 3.1.2. The difference is due to the choice of the ordering cone which is $L_t^2(\mathbb{R}_+)\times L_t^1(\mathbb{R}_+)$ for $(\mathrm{C}_t(v_t))$ as opposed to $\{0\}\times L_t^1(\mathbb{R}_+)$ for $\overline{\mathcal{A}_t(v_t)}$. It should be noted that $\{0\}\times L_t^1(\mathbb{R}_+)$ is generally not a suitable ordering cone for vector optimization since its positive dual cone $L^2_t(\mathbb{R}) \times L^\infty_t(\mathbb{R}_+)$ contains a line. This issue becomes more apparent for the computations in Chapter 6; see Remark 6.2.1 there.

For each $t \in \mathbb{T}$, we define the set-valued extension of conditional expectation as a function $\mathcal{E}_t \colon L^2_T(\mathbb{R}) \times L^1_T(\mathbb{R}) \to 2^{L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})}$ given by

$$\mathcal{E}_t(x) := \mathbb{E}_t(x) + L^2_t(\mathbb{R}_+) \times L^1_t(\mathbb{R}_+)$$

for each $x \in L^2_T(\mathbb{R}) \times L^1_T(\mathbb{R})$. Next, we establish some basic properties of $\mathcal{E}_t$.

Lemma 5.2.2. *For every $t \in \mathbb{T}$, the set-valued function $\mathcal{E}_t$ satisfies the following properties.*

- ***Monotonicity:*** *For every $x, y \in L^2_T(\mathbb{R}) \times L^1_T(\mathbb{R})$ such that $x \leq y$, we have $\mathcal{E}_t(x) \subseteq \mathcal{E}_t(y)$.*
- ***Recursiveness:*** *For every $x \in L^2_T(\mathbb{R}) \times L^1_T(\mathbb{R})$, we have $\mathcal{E}_t(x) = \mathcal{E}_t[\mathcal{E}_s(x)]$ for all $s \geq t$.*

Proof. The monotonicity of $\mathcal{E}_t$ directly follows from the monotonicty of (vector-valued) conditional expectation. For recursiveness, let us fix some $t, s \in \mathbb{T}$ such that $t \leq s$, and $x \in L^2_T(\mathbb{R}) \times L^1_T(\mathbb{R})$. Further, let $e \in \mathcal{E}_t(x)$. Then, by the tower property of conditional expectation we obtain

$$\mathcal{E}_t(\mathbb{E}_s(x)) = \mathbb{E}_t(\mathbb{E}_s(x)) + L^2_t(\mathbb{R}_+) \times L^1_t(\mathbb{R}_+) = \mathcal{E}_t(x).$$

Hence, as $\mathbb{E}_s(x) \in \mathcal{E}_s(x)$ by definition, we conclude $e \in \mathcal{E}_t(\mathbb{E}_s(x)) \subseteq \mathcal{E}_t[\mathcal{E}_s(x)]$. For the reverse inclusion, let $e \in \mathcal{E}_t[\mathcal{E}_s(x)]$. Then, there exists $r \in L^2_s(\mathbb{R}_+) \times L^1_s(\mathbb{R}_+)$ such that $e \in \mathcal{E}_t(\mathbb{E}_s(x) + r)$. Thus, by the linearity and the tower property of conditional expectation we observe

$$\mathcal{E}_t(\mathbb{E}_s(x)+r) = \mathbb{E}_t(\mathbb{E}_s(x)+r)+(L^2_t(\mathbb{R}_+)\times L^1_t(\mathbb{R}_+)) = (\mathbb{E}_t(x)+\mathbb{E}_t(r))+(L^2_t(\mathbb{R}_+)\times L^1_t(\mathbb{R}_+)).$$

Finally, as $\mathbb{E}_t(r) \in L^2_t(\mathbb{R}_+) \times L^1_t(\mathbb{R}_+)$ we obtain that $\mathcal{E}_t(\mathbb{E}_s(x) + r) \subseteq \mathcal{E}_t(x)$. Therefore, $e \in \mathcal{E}_t(x)$. $\square$

We announce our main theorem on dynamic programming, which establishes a backward recursion for the upper images of the biobjective problem ($\mathrm{C}_t(v_t)$) for consecutive values of t.

Theorem 5.2.3. *For every $t \in \mathbb{T}\setminus\{T\}$ and $v_t \in L^2_t(\mathbb{R})$, it holds*

$$\mathcal{P}_t(v_t) = \operatorname{cl} \bigcup_{\substack{\pi\in\Phi_t(v_t)\\ x\in\mathcal{P}_{t+1}(v^{\pi}_{t,t+1})}} \mathcal{E}_t(x), \tag{5.2}$$

where the closure is evaluated in the space $L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})$.

Proof. Let us fix some t $\in \mathbb{T}\setminus\{T\}$, $v_t \in L^2_t(\mathbb{R})$, and define the sets

$$A = \bigcup_{\pi\in\Phi_t(v_t)} \mathcal{E}_t\big(-v^{\pi}_{s,T}, (v^{\pi}_{s,T})^2\big), \qquad B = \bigcup_{\substack{\pi\in\Phi_t(v_t)\\ x\in\mathcal{P}_{t+1}(v^{\pi}_{t,t+1})}} \mathcal{E}_t(x).$$

Note that (5.2) can be rewritten as $\operatorname{cl} A = \operatorname{cl} B$. Then, it suffices to show that $A \subseteq \operatorname{cl} B$ and $B \subseteq \operatorname{cl} A$.

To prove $A \subseteq \operatorname{cl} B$, let $a \in A$. Then, there exists $\pi \in \Phi_t(v_t)$ such that $a \in \mathcal{E}_t(-v^{\pi}_{s,T}, (v^{\pi}_{s,T})^2)$. Now let us consider $\tilde{\pi} = (\pi_s)_{s\geq t+2}$. By construction, it holds $\tilde{\pi} \in \Phi_{t+1}(v^{\pi}_{t,t+1})$. Hence,

$$\mathcal{E}_{t+1}(-v^{\tilde{\pi}}_{t+1,T}, (v^{\tilde{\pi}}_{t+1,T})^2) \subseteq \mathcal{P}_{t+1}(v^{\pi}_{t,t+1}).$$

Moreover, by Lemma 5.2.2,

$$\mathcal{E}_t[\mathcal{E}_{t+1}(-v^{\tilde{\pi}}_{t+1,T}, (v^{\tilde{\pi}}_{t+1,T})^2)] = \mathcal{E}_t(-v^{\tilde{\pi}}_{t+1,T}, (v^{\tilde{\pi}}_{t+1,T})^2).$$

Furthermore, we have $v^{\pi}_{t,T} = v^{\tilde{\pi}}_{t+1,T}$ by the definition of $\tilde{\pi}$. Hence,

$$\begin{aligned}\mathcal{E}_t(-v^{\pi}_{t,T}, (v^{\pi}_{t,T})^2) = \mathcal{E}_t(-v^{\tilde{\pi}}_{t+1,T}, (v^{\tilde{\pi}}_{t+1,T})^2) &= \mathcal{E}_t[\mathcal{E}_{t+1}(-v^{\tilde{\pi}}_{t+1,T}, (v^{\tilde{\pi}}_{t+1,T})^2)] \\ &\subseteq \mathcal{E}_t[\mathcal{P}_{t+1}(v^{\pi}_{t,t+1})] \subseteq B\end{aligned}$$

and $a \in B \subseteq \operatorname{cl} B$.

To prove $B \subseteq \operatorname{cl} A$, let $b \in B$. Then, there exists a tuple (π, x) such that $\pi \in \Phi_t(v_t)$, $x \in \mathcal{P}_{t+1}(v^{\pi}_{t,t+1})$ and $b \in \mathcal{E}_t(x)$. By the definition of $\mathcal{P}_{t+1}(v^{\pi}_{t,t+1})$, there exists a sequence $(x^n = a^n + r^n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} x^n = x \text{ in } L^2_{t+1}(\mathbb{R}) \times L^1_{t+1}(\mathbb{R}),$$

where $a^n \in \mathcal{A}_{t+1}(v^{\pi}_{t,t+1})$ and $r^n \in L^2_{t+1}(\mathbb{R}_+) \times L^1_{t+1}(\mathbb{R}_+)$ for each $n \in \mathbb{N}$. Let $n \in \mathbb{N}$. There exists $\pi^n \in \Phi_{t+1}(v^{\pi}_{t,t+1})$ such that $a^n = \mathbb{E}_{t+1}(-v^{\pi^n}_{t+1,T}, (v^{\pi^n}_{t+1,T})^2)$. Moreover, letting $\tilde{\pi}^n := (\pi_{t+1}, \pi^n)$, we have $v^{\tilde{\pi}^n}_{t,T} = v^{\pi^n}_{t+1,T}$; by Lemma 5.2.2, we also have

$$\mathbb{E}_t(a^n) = \mathbb{E}_t\Big(\mathbb{E}_{t+1}\big(-v^{\pi^n}_{t+1,T}, (v^{\pi^n}_{t+1,T})^2\big)\Big) = \mathbb{E}_t\big(-v^{\tilde{\pi}^n}_{t,T}, (v^{\tilde{\pi}^n}_{t,T})^2\big) \in \mathcal{P}_t(v_t)$$

since $\tilde{\pi}^n \in \Phi_t(v_t)$ by construction. Note that we have $a^n \leq x^n$, hence $\mathbb{E}_t(a^n) \leq \mathbb{E}_t(x^n)$ as well. Therefore, as the upper image $\mathcal{P}_t(v_t)$ is a monotone set by definition, we obtain $\mathbb{E}_t(x^n) \in \mathcal{P}_t(v_t)$. Finally, by the contraction property of conditional expectation conclude that

$$\lim_{n\to\infty} \mathbb{E}_t(x^n) = \mathbb{E}_t\Big(\lim_{n\to\infty} x^n\Big) = \mathbb{E}_t(x) \text{ in } L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R}).$$

Then, as $\mathcal{P}_t(v_t)$ is closed, we obtain $\mathbb{E}_t(x) \in \mathcal{P}_t(v_t)$. Therefore, $b \in \mathcal{E}_t(x) \subseteq \mathcal{P}_t(v_t) = \operatorname{cl} A$. □

Remark 5.2.4. With minor adjustments in the proof of Theorem 5.2.3, the following set-valued Bellman's principle can be shown by changing the ordering cone of $(\mathrm{C}_t(v_t))$ as $\{0\} \times L^1_t(\mathbb{R}_+)$:

$$\overline{\mathcal{A}_t(v_t)} = \operatorname{cl} \bigcup_{\substack{\pi\in\Phi_t(v_t) \\ x\in\overline{\mathcal{A}_{t+1}(v^{\pi}_{t,t+1})}}} \left(\mathbb{E}_t(x) + \{0\} \times L^1_t(\mathbb{R}_+)\right).$$

Hence, the choice of our ordering cone $L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)$ is not necessarily for theoretical convenience, but it is rather for computational purposes as noted in Remark 5.2.1.

Remark 5.2.5. Let us define the hyperspace

$$\mathscr{F} = \left\{ D \subseteq L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R}) \mid \mathrm{cl}(D + L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)) = D \right\},$$

which is the collection of all closed subsets of $L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})$ that are monotone with respect to the cone $L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)$. Then, the superset relation $\supseteq$ defines a partial order on $\mathscr{F}$, and $(\mathscr{F}, \supseteq)$ is indeed a complete lattice. In this order-theoretic framework, the infimum of an arbitrary subset $\mathscr{D} \subseteq \mathscr{F}$ is given as $\inf \mathscr{D} = \mathrm{cl} \bigcup_{D \in \mathscr{D}} D$, which makes it possible to study rigorously the minimization of a function taking values in $\mathscr{D}$ as a *set optimization* problem. The reader is referred to [21] for a survey on the abstract foundations of this approach, and to [14] for an application of this approach on the mean-risk problem. In our case, it is easy to see that

$$\left\{ \mathcal{E}_t \left[\mathcal{P}_{t+1}(v^\pi_{t,t+1}) \right] \mid \pi \in L_t^\infty(\mathbb{R}^d),\ \pi^\mathsf{T} S_t = v_t,\ \varphi_t(\pi) \geq 0 \right\} \subseteq \mathscr{F}.$$

Hence, by [21, Example 2.11], we may rewrite (5.2) as a set optimization problem as follows:

$$\mathcal{P}_t(v_t) = \inf_{\pi \in \Phi_t(v_t)} \mathcal{E}_t \left[\mathcal{P}_{t+1}(v^\pi_{t,t+1}) \right]. \tag{5.3}$$

Note that (5.3) is a set-valued analogue of the scalar Bellman's principle in Proposition 4.0.1.

Similar to Proposition 4.0.1, we may rewrite the dynamic programming principle in Theorem 5.2.3 as

$$\mathcal{P}_t(v_t) = \mathrm{cl} \bigcup_{\substack{\pi \in L_t^\infty(\mathbb{R}^d) \\ \pi^\mathsf{T} S_t = v_t \\ \varphi_t(\pi) \geq 0}} \mathcal{E}_t \left[\mathcal{P}_{t+1}(v^\pi_{t,t+1}) \right] \tag{5.4}$$

as we are only interested in the behavior of a single period portfolio position. Then, we may rearrange the terms in (5.4) as follows in order to obtain the

Minkowski sum of an image set and the ordering cone $L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)$:

$$\mathcal{P}_t(v_t) = \mathrm{cl}\left(\left\{\mathbb{E}_t(x) \;\middle|\; \begin{array}{c} \pi \in L_t^\infty(\mathbb{R}^d),\ \pi^\mathsf{T} S_t = v_t \\ \varphi_t(\pi) \geq 0,\ x \in \mathcal{P}_{t+1}(v_{t,t+1}^\pi) \end{array}\right\} + L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)\right). \quad (5.5)$$

Hence, we may see $\mathcal{P}_t(v_t)$, the upper image of $(\mathrm{C}_t(v_t))$, also as the upper image of the following *one-step* vector optimization problem in which the set-valued function $\mathcal{P}_{t+1}$ appears in the constraints:

$$\begin{array}{lll} \text{minimize} & \mathbb{E}_t(x) \quad \text{with respect to} & \leq_{L_t^2(\mathbb{R}_+)\times L_t^1(\mathbb{R}_+)} \\ \text{subject to} & \pi \in L_t^\infty(\mathbb{R}^d) & \\ & \pi^\mathsf{T} S_t = v_t & \\ & \varphi_t(\pi) \geq 0 & \\ & x \in \mathcal{P}_{t+1}(v_{t,t+1}^\pi). & \end{array} \qquad (\tilde{\mathrm{C}}_t(v_t))$$

Finally, similar to the relationship between the scalar problem $(\mathrm{A}_t(v_t, \rho_t))$ and the vector optimization counterpart $(\mathrm{C}_t(v_t))$, we introduce the scalar analogue of $(\tilde{\mathrm{C}}_t(v_t))$ as a weighted sum scalarization. That is, for every $\rho_t \in L_t^2(\mathbb{R}_+)\times L_t^\infty(\mathbb{R}_+)$, we introduce the scalar problem

$$\operatorname*{ess\,inf}_{\substack{\pi\in L_t^\infty(\mathbb{R}^d) \\ \pi^\mathsf{T} S_t = v_t \\ \varphi_t(\pi)\geq 0 \\ x\in\mathcal{P}_{t+1}(v_{t,t+1}^\pi)}} \rho_t^\mathsf{T}\mathbb{E}_t(x). \qquad (\mathrm{R}_t(v_t, \rho_t))$$

The next two results reinforce our previous observations on the analogy between the scalar and set-valued Bellman's equations (4.3) and (5.3). We first note that if the weight ρ_t in $(\mathrm{R}_t(v_t, \rho_t))$ is chosen to be a constant, then $(\mathrm{R}_t(v_t, \rho_t))$ reduces to the scalar value function V_t defined by (4.1).

Proposition 5.2.6. *Let $t \in \mathbb{T}\setminus\{T\}$, $v_t \in L_t^2(\mathbb{R})$ and $\rho_0 \in \mathbb{R}_+^2 \setminus \{0\}$. Let $(\pi^\star, x^\star)$ be an optimal solution of $(\mathrm{R}_t(v_t, \rho_0))$. Then, it holds*

$$\mathbb{E}_t(\rho_0^\mathsf{T} x^\star) = V_t(v_t).$$

In addition, let $\tilde{\pi} \in \Phi_t(v_{t,t+1}^{\pi^\star})$ be an optimal solution of $(\mathrm{A}_{t+1}(v_{t,t+1}^{\pi^\star}, \rho_0))$. Then,

$\pi = (\pi^\star, \tilde{\pi})$ *is an optimal solution of* $(\mathrm{A}_t(v_t, \rho_0))$. *In particular, if we further assume that* $\rho_0 \in \mathbb{R}^2_{++}$, *then the following identity holds:*

$$x^\star = \left(-\mathbb{E}_{t+1}(v^{\tilde{\pi}}_{t+1,T}), \mathbb{E}_{t+1}((v^{\tilde{\pi}}_{t+1,T})^2)\right)^{\mathsf{T}}. \tag{5.6}$$

Proof. Following similar arguments as in the proof of Lemma 3.1.5, it can be checked that

$$\operatorname*{ess\,inf}_{x \in \mathcal{P}_{t+1}(v_{t+1})} \rho_0^{\mathsf{T}} \mathbb{E}_t(x) = \operatorname*{ess\,inf}_{x \in \mathcal{A}_{t+1}(v_{t+1})} \rho_0^{\mathsf{T}} \mathbb{E}_t(x) \tag{5.7}$$

for every $v_{t+1} \in L^2_{t+1}(\mathbb{R})$. The proof is omitted for brevity. Then, as $\rho_0 \in \mathbb{R}^2_+$, by (4.4) in the proof of Proposition 4.0.1, we have

$$\operatorname*{ess\,inf}_{x \in \mathcal{A}_{t+1}(v_{t+1})} \rho_0^{\mathsf{T}} \mathbb{E}_t(x) = \operatorname*{ess\,inf}_{x \in \mathcal{A}_{t+1}(v_{t+1})} \mathbb{E}_t(\rho_0^{\mathsf{T}} x) = \mathbb{E}_t\left(V_{t+1}(v_{t+1})\right) \tag{5.8}$$

for every $v_{t+1} \in L^2_{t+1}(\mathbb{R})$. Furthermore, by decomposing the feasible region into two parts and using (5.7) and (5.8), we rewrite $(\mathrm{R}_t(v_t, \rho))$ as

$$\operatorname*{ess\,inf}_{\substack{\pi \in L^\infty_t(\mathbb{R}^d) \\ \pi^{\mathsf{T}} S_t = v_t \\ \varphi_t(\pi) \geq 0}} \left(\operatorname*{ess\,inf}_{x \in \mathcal{P}_{t+1}(v^\pi_{t,t+1})} \mathbb{E}_t(\rho_0^{\mathsf{T}} x) \right) = \operatorname*{ess\,inf}_{\substack{\pi \in L^\infty_t(\mathbb{R}^d) \\ \pi^{\mathsf{T}} S_t = v_t \\ \varphi_t(\pi) \geq 0}} \mathbb{E}_t\left(V_{t+1}(v^\pi_{t,t+1})\right) = V_t(v_t), \tag{5.9}$$

where the last equality is due to Proposition 4.0.1. As a result, we conclude that

$$\mathbb{E}_t(\rho_0^{\mathsf{T}} x^\star) = V_t(v_t).$$

As the next step, let us suppose that $\tilde{\pi} \in \Phi_t(v^{\pi^\star}_{t,t+1})$ is an optimal solution of $(\mathrm{A}_{t+1}(v^{\pi^\star}_{t,t+1}, \rho_0))$. Then, by the definition of the scalar value function in (4.1), we have

$$V_{t+1}(v^{\pi^\star}_{t,t+1}) = -\rho_{0,1}\mathbb{E}_{t+1}\left(v^{\tilde{\pi}}_{t+1,T}\right) + \rho_{0,2}\mathbb{E}_{t+1}\left((v^{\tilde{\pi}}_{t+1,T})^2\right).$$

Note that $\pi = (\pi^\star, \tilde{\pi}) \in \Phi_t(v_t)$ by construction. Then, under (5.9), by linearity and tower property, we have

$$\begin{aligned}
\mathbb{E}_t\left(V_{t+1}(v^{\pi^\star}_{t,t+1})\right) &= -\rho_{0,1}\mathbb{E}_t\left(v^{\tilde{\pi}}_{t+1,T}\right) + \rho_{0,2}\mathbb{E}_t\left((v^{\tilde{\pi}}_{t+1,T})^2\right) \\
&= -\rho_{0,1}\mathbb{E}_t\left(v^{\pi}_{t,T}\right) + \rho_{0,2}\mathbb{E}_t\left((v^{\pi}_{t,T})^2\right) = V_t(v_t).
\end{aligned}$$

Hence, we conclude that π is an optimal solution of $(\mathrm{A}_t(v_t,\rho_0))$. Finally, let us further assume that $\rho_0 \in \mathbb{R}^2_{++}$. Then, by Proposition 3.2.3, we guarantee the uniqueness of the optimal terminal values of the problem $(\mathrm{A}_{t+1}(v^{\pi^\star}_{t,t+1},\rho_0))$. Moreover, by (5.7), we have $x^\star \in \mathcal{A}_{t+1}(v^{\pi^\star}_{t,t+1})$. Hence, (5.6) follows by (5.8) and the uniqueness of optimal terminal values. $\square$

As a direct consequence of Proposition 5.2.6, our next result indicates that when we solve the family $(\mathrm{R}_t(\cdot,\rho_0))_{t\in\mathbb{T}\setminus\{T\}}$ of one-step scalar optimization problems for a given initial wealth $v_0 \in \mathbb{R}_+$ and weight vector ρ_0, the resulting optimal portfolio process π is in fact an optimal solution for $(\mathrm{A}_t(v^\pi_t,\rho_0))$ for all $t \in \mathbb{T}\setminus\{T\}$. Moreover, if $\rho_0 \in \mathbb{R}_{++}$, that is, if both of the objectives have a nonzero weight, than π is an optimal solution for $(\mathrm{M}_t(v^\pi_t,\lambda_t))$ for every $t \in \mathbb{T}\setminus\{T\}$, where $(\lambda_t)_{t\in\mathbb{T}\setminus\{T\}}$ is the associated time-consistent weight process and it can be calculated by using the optimal solutions of $(\mathrm{R}_t(\cdot,\rho_0))_{t\in\mathbb{T}\setminus\{T\}}$.

Corollary 5.2.7. *Let $v_0 \in \mathbb{R}_+$ and $\rho_0 \in \mathbb{R}^2_+$. For every $t \in \mathbb{T}\setminus\{0\}$, let*

$$(\pi_t, x^t) \in L^\infty_t(\mathbb{R}^d) \times \left(L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})\right),$$

and define the portfolio $\pi = (\pi_s)_{s\in\mathbb{T}\setminus\{0\}}$. Suppose that (π_1, x^1) is an optimal solution of $(\mathrm{R}_0(v_0,\rho_0))$ and (π_t, x^t) is an optimal solution of $(\mathrm{R}_t(v^\pi_t,\rho_0))$ for every $t \in \mathbb{T}\setminus\{0,T\}$. Then, $(\pi_s)_{s\geq t+1}$ is an optimal solution of $(\mathrm{A}_t(v^\pi_t,\rho_0))$ for every $t \in \mathbb{T}\setminus\{0,T\}$, and in particular, $\pi \in \Phi_0(v_0)$ is an optimal solution of $(\mathrm{A}_t(v_0,\rho_0))$. If we further assume that $\rho_0 \in \mathbb{R}^2_{++}$, then $(\pi_s)_{s\geq t+1}$ is an optimal solution of $(\mathrm{M}_t(v^\pi_t,\lambda_t))$ for every $t \in \mathbb{T}\setminus\{T\}$, where

$$\lambda_t = \begin{pmatrix} \rho_{0,1} - 2\rho_{0,2}\mathbb{E}_t\left(x^{t+1}_1\right) \\ \rho_{0,2} \end{pmatrix}$$

for every $t \in \mathbb{T}\setminus\{T\}$.

Proof. We note that, by definition the upper image at the terminal time T is given by

$$\mathcal{P}_T(v_T) := \left\{x \in L^2_T(\mathbb{R}) \times L^1_T(\mathbb{R}) \mid x \geq \left(v_T, v_T^2\right)^\mathsf{T}\right\} = (v_T,(v_T)^2)^\mathsf{T} + L^2_t(\mathbb{R}_+)\times L^1_t(\mathbb{R}_+).$$

Therefore, the equivalence of the feasible regions, hence the optimal solutions of the problems $(\mathrm{A}_{T-1}(v^{\pi}_{T-1},\rho_0))$ and $(\mathrm{R}_{T-1}(v^{\pi}_{T-1},\rho_0))$ is straightforward. Thus, we conclude that π_T is an optimal solution for $(\mathrm{A}_{T-1}(v^{\pi}_{T-1},\rho_0))$ as well. Moreover, under Proposition 5.2.6, by induction, the results for $t \in \{0,\ldots,T-2\}$ follow. Now, let us further assume that $\rho_0 \in \mathbb{R}^2_{++}$. Then, the assertion for $(\mathrm{M}_t(v^{\pi}_t,\lambda_t))$ follows by (5.6) in Proposition 5.2.6 and Theorem 3.2.1. □

5.3 Graph of $\mathcal{P}_t$ and convex projection

Note that, in the one-step vector optimization problem $(\tilde{\mathrm{C}}_t(v_t))$, the constraint $x \in \mathcal{P}_{t+1}(v^{\pi}_{t,t+1})$ requires access to arbitrarily many values of the set-valued function $\mathcal{P}_{t+1}$ as we search over different choices of the feasible portfolio π. Consequently, in view of (5.4), to calculate $\mathcal{P}_t(v_t)$ for a given $v_t \in L^2_t(\mathbb{R})$, we generally need to know the entire function $\mathcal{P}_{t+1}$ rather than a specific value of it. Therefore, it is more convenient to consider the graphs of these set-valued functions. For each $t \in \mathbb{T}$, let us define the graph of $\mathcal{P}_t$ by

$$\operatorname{graph}\mathcal{P}_t \coloneqq \{(v_t,x) \in L^2_t(\mathbb{R}) \times (L^2_t(\mathbb{R}) \times L^1_t(\mathbb{R})) \mid x \in \mathcal{P}_t(v_t)\}.$$

Clearly, we have

$$x \in \mathcal{P}_t(v_t) \iff (v_t,x) \in \operatorname{graph}\mathcal{P}_{t+1}. \tag{5.10}$$

By this reformulation, the functional constraint turns into a single set constraint for the variable (v_t,x) and $(\tilde{\mathrm{C}}_t(v_t))$ becomes

$$\begin{aligned} \text{minimize} \quad & \mathbb{E}_t(x) \quad \text{with respect to} \quad \leq_{L^2_t(\mathbb{R}^2_+)} \\ \text{subject to} \quad & \pi \in L^\infty_t(\mathbb{R}^d) \\ & \pi^{\mathsf{T}} S_t = v_t \\ & \varphi_t(\pi) \geq 0 \\ & (v^{\pi}_{t,t+1},x) \in \operatorname{graph}\mathcal{P}_{t+1}. \end{aligned} \tag{$\tilde{\mathrm{C}}_t(v_t)$}$$

To operate with this reformulation conveniently, we will prove an analogue of Theorem 5.2.3 for the collection $(\operatorname{graph}\mathcal{P}_t)_{t\in\mathbb{T}}$ of graphs. To that end, for the remainder of this thesis we shall assume that the $\operatorname{graph}\mathcal{P}_t$ is a closed subset of $L^2_t(\mathbb{R}^2)\times L^1_t(\mathbb{R})$.

Assumption 5.3.1. *For every $t\in\mathbb{T}$, $\operatorname{graph}\mathcal{P}_t$ is a closed subset of $L^2_t(\mathbb{R}^2)\times L^1_t(\mathbb{R})$.*

The following result prepares us for the study of the dynamic behavior of the collection $(\operatorname{graph}\mathcal{P}_t)_{t\in\mathbb{T}}$.

Lemma 5.3.2. *Let $t\in\mathbb{T}\setminus\{T\}$ and $v_t\in L^2_t(\mathbb{R})$. Then, the ordering cone in* (5.5) *can be dropped, that is,*

$$\mathcal{P}_t(v_t)=\operatorname{cl}\left\{\mathbb{E}_t(x)\mid \pi\in L^\infty_t(\mathbb{R}^d),\ \pi^\mathsf{T}S_t=v_t,\ \varphi_t(\pi)\geq 0,\ x\in\mathcal{P}_{t+1}(v^\pi_{t,t+1})\right\},$$

where the closure is evaluated in the space $L^2_t(\mathbb{R})\times L^1_t(\mathbb{R})$.

Proof. Let A be the set inside the closure above. It suffices to have $A+L^2_t(\mathbb{R}_+)\times L^1_t(\mathbb{R}_+)\subseteq A$. To that end, let $a\in A$ and $r\in L^2_t(\mathbb{R}_+)\times L^1_t(\mathbb{R}_+)$. Then, there exist $\pi\in L^\infty_t(\mathbb{R}^d)$ and $x\in L^2_{t+1}(\mathbb{R})\times L^1_{t+1}(\mathbb{R})$ such that $\pi^\mathsf{T}S_t=v_t$, $\varphi_t(\pi)\geq 0$, $x\in\mathcal{P}_{t+1}(v^\pi_{t,t+1})$, and $a=\mathbb{E}_t(x)$. Moreover, as $L^p_t(\mathbb{R}_+)\subseteq L^p_{t+1}(\mathbb{R}_+)$ for $p\in\{1,2\}$, we observe that $r\in L^2_{t+1}(\mathbb{R}_+)\times L^1_{t+1}(\mathbb{R}_+)$, and $x+r\in\mathcal{P}_{t+1}(v^\pi_{t,t+1})$. Thus, we conclude that $a+r\in A$, as $\mathbb{E}_t(x+r)=\mathbb{E}_t(x)+\mathbb{E}_t(r)=a+r$ due the linearity of conditional expectation and the $\mathcal{F}_t$-measurablity of r. $\square$

The next proposition indicates a convex projection property for the $\operatorname{graph}\mathcal{P}_t$ in the spirit of [22, Section 4].

Proposition 5.3.3. *For every $t\in\mathbb{T}\setminus\{T\}$, it holds that*

$$\operatorname{graph}\mathcal{P}_t=\operatorname{cl}\left(\left\{(v_t,\mathbb{E}_t(x))\;\middle|\;\exists(\pi,x)\colon \begin{array}{c} v_t\in L^2_t(\mathbb{R}),\ \pi\in L^\infty_t(\mathbb{R}^d),\ \pi^\mathsf{T}S_t=v_t\\ \varphi_t(\pi)\geq 0,\ (v^\pi_{t,t+1},x)\in\operatorname{graph}\mathcal{P}_{t+1}\end{array}\right\}\right),$$

where the closure is evaluated in the space $L^2_t(\mathbb{R})\times(L^2_t(\mathbb{R})\times L^1_t(\mathbb{R}))$.

Proof. Let $\mathcal{N}$ denote the set inside the closure above. To show that $\operatorname{graph} \mathcal{P}_t \subseteq \operatorname{cl} \mathcal{N}$, let $(v_t, y) \in \operatorname{graph} \mathcal{P}_t$, that is, $y \in \mathcal{P}_t(v_t)$. By Lemma 5.3.2, there exists a sequence $(\pi^n, x^n)_{n\in\mathbb{N}}$ such that $\pi^n \in L_t^\infty(\mathbb{R}^d)$, $(\pi^n)^\mathsf{T} S_t = v_t$, $\varphi_t(\pi^n) \geq 0$, and $x^n \in \mathcal{P}_{t+1}(v_{t,t+1}^{\pi^n})$ for every $n \in \mathbb{N}$, and

$$\lim_{n\to\infty} \mathbb{E}_t(x^n) = y \text{ in } L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R}).$$

We observe that $(v_t, \mathbb{E}_t(x^n)) \in \mathcal{N}$ for every $n \in \mathbb{N}$. Moreover, as v_t is uniform for the sequence $(\mathbb{E}_t(x^n))_{n\in\mathbb{N}}$ we obtain that

$$\lim_{n\to\infty} (v_t, \mathbb{E}_t(x^n)) = (v_t, y) \text{ in } L_t^2(\mathbb{R}) \times (L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})).$$

Thus, we conclude that $(v_t, y) \in \operatorname{cl} \mathcal{N}$.

For the reverse inclusion, let $(v_t, y) \in \operatorname{cl} \mathcal{N}$. Then, there exists a sequence $(v_t^n, \pi^n, x^n)_{n\in\mathbb{N}}$ such that $v_t^n \in L_t^2(\mathbb{R})$, $(\pi^n)^\mathsf{T} S_t = v_t^n$, $\pi^n \in L_t^\infty(\mathbb{R}^d)$, and $x^n \in \mathcal{P}_{t+1}(v_{t,t+1}^{\pi^n})$ for every $n \in \mathbb{N}$, and

$$\lim_{n\to\infty} (v_t^n, \mathbb{E}_t(x^n)) = (v_t, y) \quad \text{in } L_t^2(\mathbb{R}) \times (L_t^2(\mathbb{R}) \times L_t^1(\mathbb{R})).$$

By Lemma 5.3.2 again, we have $\mathbb{E}_t(x^n) \in \mathcal{P}_t(v_t^n)$ for every $n \in \mathbb{N}$. Therefore, as $\operatorname{graph} \mathcal{P}_t$ is closed by Assumption 5.3.1, we conclude that $(v_t, y) \in \operatorname{graph} \mathcal{P}_t$. Hence, $\operatorname{cl} \mathcal{N} \subseteq \operatorname{graph} \mathcal{P}_t$. □

Note that Proposition 5.3.3 provides a backward recursion for the sequence $(\operatorname{graph} \mathcal{P}_t)_{t\in\mathbb{T}}$. Moreover, similar to (5.5), using Lemma 5.3.2, we may extend Proposition 5.3.3 such that $\operatorname{graph} \mathcal{P}_t$ is equal to

$$\operatorname{cl} \left(\left\{ (v_t, \mathbb{E}_t(x)) \;\middle|\; \exists (\pi, x) \colon \begin{array}{c} v_t \in L_t^2(\mathbb{R}),\, \pi \in L_t^\infty(\mathbb{R}^d),\, \pi^\mathsf{T} S_t = v_t \\ \varphi_t(\pi) \geq 0,\, (v_{t,t+1}^\pi, x) \in \operatorname{graph} \mathcal{P}_{t+1} \end{array} \right\} \right.$$
$$\left. + \{0\} \times (L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)) \right),$$

which has the structure of an upper image of a convex optimization problem whose ordering cone is $\{0\} \times (L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+))$ (see [22, Section 4]). However, in accordance with Remark 5.2.1, this is not a convenient ordering cone for computational purposes. For this reason, in order to utilize the dynamic structure in Proposition 5.3.3, we need to switch to a different convex vector optimization problem with a convenient ordering cone.

Recently, the connection between projection problems and vector optimization has been studied in the literature. [23] established an equivalence between polyhedral projection and linear vector optimization. The more delicate connection between convex projection and convex vector optimization is analyzed by [22]. These articles deal with finite-dimensional projection/optimization problems. Analogous to their approach, for each $t \in \mathbb{T}\backslash\{T\}$, we introduce the following convex vector optimization problem which has an additional objective dimension compared to graph $\mathcal{P}_t$:

$$\begin{aligned} \text{minimize} \quad & \begin{pmatrix} v_t \\ \mathbb{E}_t(x) \\ -\mathbf{1}^\mathsf{T}(v_t, \mathbb{E}_t(x)) \end{pmatrix} \quad \text{with respect to} \quad \leq_{L^2_t(\mathbb{R}^2_+)\times L^1_t(\mathbb{R}^2_+)} \\ \text{subject to} \quad & v_t \in L^2_t(\mathbb{R}) \\ & \pi \in L^\infty_t(\mathbb{R}^n) \\ & \pi^\mathsf{T} S_t = v_t \\ & \varphi_t(\pi) \geq 0 \\ & (v^\pi_{t,t+1}, x) \in \operatorname{graph} \mathcal{P}_{t+1}. \end{aligned} \tag{G$_t$}$$

We denote the image set and the upper image of the problem (G$_t$) as $\mathcal{I}_t$ and $\mathcal{U}_t$, respectively, which are given by

$$\mathcal{I}_t := \left\{ \left(v_t, \mathbb{E}_t(x), -v_t - \mathbf{1}^\mathsf{T}\mathbb{E}_t(x)\right)^\mathsf{T} \;\middle|\; \begin{array}{c} v_t \in L^2_t(\mathbb{R}),\, \pi \in L^\infty_t(\mathbb{R}^d),\, \pi^\mathsf{T} S_t = v_t \\ \varphi_t(\pi) \geq 0,\, (v^\pi_{t,t+1}, x) \in \operatorname{graph} \mathcal{P}_{t+1} \end{array} \right\},$$

$$\mathcal{U}_t := \operatorname{cl}\left(\mathcal{I}_t + L^2_t(\mathbb{R}^2_+) \times L^1_t(\mathbb{R}^2_+)\right),$$

where the closure is evaluated in $L^2_t(\mathbb{R}^2) \times L^1_t(\mathbb{R}^2)$. Note that $\mathcal{I}_t$ is a subset of the linear subspace $\{x \in L^2_t(\mathbb{R}^2) \times L^1_t(\mathbb{R}^2) \mid \mathbf{1}^\mathsf{T} x = 0\}$.

Furthermore, compared to $\mathcal{I}_t$, the upper image $\mathcal{U}_t$ does not contain any additional points on the subspace $\{x \in L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R}^2) \mid \mathbf{1}^\mathsf{T} x = 0\}$, except its limit points on the space $L_t^1(\mathbb{R}^4)$. The following two results justify our observations on $\mathcal{I}_t$ and $\mathcal{U}_t$.

Proposition 5.3.4. *Let $t \in \mathbb{T} \setminus \{T\}$. Then,*

$$\operatorname{cl} \mathcal{I}_t \subseteq \{x \in \mathcal{U}_t \mid \mathbf{1}^\mathsf{T} x = 0\} \subseteq (\operatorname{cl}_1 \mathcal{I}_t) \cap (L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R}^2)),$$

where cl *denotes the closure operator in $L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R}^2)$ and* cl_1 *denotes the closure operator in $L_t^1(\mathbb{R}^4)$.*

Proof. To prove the first inclusion, note that $\mathcal{I}_t \subseteq \{x \in \mathcal{U}_t \mid \mathbf{1}^\mathsf{T} x = 0\}$ by the definitions of $\mathcal{I}_t$ and $\mathcal{U}_t$. We claim that $\{x \in L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R}^2) \mid \mathbf{1}^\mathsf{T} x = 0\}$ is a closed subspace of $L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R}^2)$. Indeed, let $(x^n)_{n\in\mathbb{N}}$ be a convergent sequence in this subspace with some limit x. Then, $(x^n)_{n\in\mathbb{N}}$ converges to x also in $L_t^1(\mathbb{R}^4)$. By the continuity of addition, $(\mathbf{1}^\mathsf{T} x^n)_{n\in\mathbb{N}}$ converges to $\mathbf{1}^\mathsf{T} x$. However, we have $\mathbf{1}^\mathsf{T} x^n = 0$ for each $n \in \mathbb{N}$. Hence, we must have $\mathbf{1}^\mathsf{T} x = 0$. Therefore, x belongs to the same subspace and the claim follows. On the other hand, $\mathcal{U}_t$ is closed in $L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R}^2)$ by definition. Hence, $\{x \in \mathcal{U}_t \mid \mathbf{1}^\mathsf{T} x = 0\}$ is closed in $L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R}^2)$ so that $\operatorname{cl} \mathcal{I}_t \subseteq \{x \in \mathcal{U}_t \mid \mathbf{1}^\mathsf{T} x = 0\}$.

To prove the second inclusion, let $x \in \mathcal{U}_t$ with $\mathbf{1}^\mathsf{T} x = 0$. By the definition of $\mathcal{U}_t$, there exists a sequence $(x^n, r^n)_{n\in\mathbb{N}}$ such that $x^n \in \mathcal{I}_t$ and $r^n \in L_t^2(\mathbb{R}_+^2) \times L_t^1(\mathbb{R}_+^2)$ for each $n \in \mathbb{N}$, and

$$\lim_{n\to\infty} x^n + r^n = x \text{ in } L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R}^2). \tag{5.11}$$

Note that the above limit is also valid in $L_t^1(\mathbb{R}^4)$. In particular, by the continuity of addition, we have

$$\lim_{n\to\infty} \mathbf{1}^\mathsf{T}(x^n + r^n) = \mathbf{1}^\mathsf{T} x \text{ in } L_t^1(\mathbb{R}^4).$$

On the other hand, we have $\mathbf{1}^\mathsf{T} x = 0$ by supposition, and, for every $n \in \mathbb{N}$, we have $\mathbf{1}^\mathsf{T} x^n = 0$ since $x^n \in \mathcal{I}_t$. Hence,

$$\lim_{n\to\infty} \mathbf{1}^\mathsf{T} r^n = 0 \text{ in } L^1_t(\mathbb{R}). \tag{5.12}$$

Moreover, since $r_n \geq 0$ for every $n \in \mathbb{N}$, we observe that $\left\|\mathbf{1}^\mathsf{T} r^n\right\|_1 = \|r^n\|_1$ for every $n \in \mathbb{N}$. Thus, combining this observation with (5.12), we conclude that

$$\lim_{n\to\infty} r^n = 0 \text{ in } L^1_t(\mathbb{R}^4).$$

Therefore, we conclude that $(x^n)_{n\in\mathbb{N}} \subseteq \mathcal{I}_t$ converges to x in $L^1_t(\mathbb{R}^4)$. Thus, we have $x \in \operatorname{cl}_1 \mathcal{I}_t$. □

Corollary 5.3.5. *Let $t \in \mathbb{T}\backslash\{T\}$, and let us consider the projection $\mathcal{U}^P_t$ of $\mathcal{U}_t$ onto its first three components in the subspace $\{x \in L^2_t(\mathbb{R}^2) \times L^1_t(\mathbb{R}^2) \mid \mathbf{1}^\mathsf{T} x = 0\}$, which is given by*

$$\mathcal{U}^P_t \coloneqq \left\{x \in L^2_t(\mathbb{R}^2) \times L^1_t(\mathbb{R}) \mid \exists y \in L^1_t(\mathbb{R})\colon (x, -y)^\mathsf{T} \in \mathcal{U}_t \ , \ \mathbf{1}^\mathsf{T} x = y\right\}. \tag{5.13}$$

Then, it holds that $\operatorname{cl}_1 \operatorname{graph} \mathcal{P}_t = \operatorname{cl}_1 \mathcal{U}^P_t$, where cl_1 denotes the closure operator in $L^1_t(\mathbb{R}^4)$.

Proof. The inclusion $\operatorname{graph} \mathcal{P}_t \subseteq \operatorname{cl}_1 \mathcal{U}^P_t$ is straightforward since for every $x \in \operatorname{graph} \mathcal{P}_t$, we may set $y = -\mathbf{1}^\mathsf{T} x \in L^1_t(\mathbb{R})$ so that $(x, -y)^\mathsf{T} \in \mathcal{U}_t$. Hence, $\operatorname{cl}_1 \operatorname{graph} \mathcal{P}_t \subseteq \operatorname{cl}_1 \mathcal{U}^P_t$ as well. To prove the reverse inclusion, let $x \in \mathcal{U}^P_t$. Then, there exists $y \in L^1_t(\mathbb{R})$ such that $(x, -y)^\mathsf{T} \in \mathcal{U}_t$ and $\mathbf{1}^\mathsf{T}(x, -y)^\mathsf{T} = 0$. Thus, by Proposition 5.3.4, we obtain that $(x, -y)^\mathsf{T} \in \operatorname{cl}_1 \mathcal{I}_t$. Note that we have

$$\left\{(z, -\mathbf{1}^\mathsf{T} z)^\mathsf{T} \mid z \in \operatorname{graph} \mathcal{P}_t\right\} = \mathcal{I}_t$$

by construction. Then, observe that for each $z \in \operatorname{cl}_1 \operatorname{graph} \mathcal{P}_t$, there exists a convergent sequence $(z^n)_{n\in\mathbb{N}}$ whose limit in $L^1_t(\mathbb{R}^3)$ is z. Then, by the continuity of addition, $(-\mathbf{1}^\mathsf{T} z^n)_{n\in\mathbb{N}}$ converges to $-\mathbf{1}^\mathsf{T} z$. Hence, we conclude that $(z^n, -\mathbf{1}^\mathsf{T} z^n)_{n\in\mathbb{N}} \subseteq \mathcal{I}_t$ converges to $(z, -\mathbf{1}^\mathsf{T} z)$ in $L^1_t(\mathbb{R}^4)$. Therefore, we have $(z, -\mathbf{1}^\mathsf{T} z) \in \operatorname{cl}_1 \mathcal{I}_t$. Conversely, let $(z^n)_{n\in\mathbb{N}} \subseteq \mathcal{I}_t$ be a convergent sequence whose limit is z. Letting $\tilde{z}^n = (z^n_1, z^n_2, z^n_3)^\mathsf{T}$ for each $n \in \mathbb{N}$, we obtain that

$(\tilde{z}^n)_{n\in\mathbb{N}} \subseteq \operatorname{graph}\mathcal{P}_t$ converges to $\tilde{z} = (z_1, z_2, z_3)^\mathsf{T}$ in $L_t^1(\mathbb{R}^3)$. Hence, we have $\tilde{z} \in \operatorname{graph}\mathcal{P}_t$. As a result, we conclude that

$$\left\{(x, -\mathbf{1}^\mathsf{T}x)^\mathsf{T} \mid x \in \operatorname{cl}_1 \operatorname{graph}\mathcal{P}_t\right\} = \operatorname{cl}_1 \mathcal{I}_t.$$

Therefore, we arrive at the desired inclusion, as $(x, -y)^\mathsf{T} \in \operatorname{cl}_1 \mathcal{I}_t$ implies $x \in \operatorname{cl}_1 \operatorname{graph}\mathcal{P}_t$. □

Remark 5.3.6. The reader can notice that both $\mathcal{U}_t^P$ and $\operatorname{graph}\mathcal{P}_t$ are closed subsets of $L_t^2(\mathbb{R}^2) \times L_t^1(\mathbb{R})$, by definition and Assumption 5.3.1, respectively. However, we shall not establish their equality without accounting for their respective limit points in the space $L_t^1(\mathbb{R}^4)$. The main reason is the loss of convergence information when the projections of $\mathcal{U}_t$ and $\mathcal{I}_t$ are evaluated. That is, as we observe in Proposition 5.3.4, there can exist an element of $\mathcal{U}_t^P$ which can only be approached by the elements of $\operatorname{graph}\mathcal{P}_t$ with the $\|\cdot\|_1$-norm. That being covered, if under any circumstance we have $L_t^1(\mathbb{R}) \simeq L_t^2(\mathbb{R})$, that is, if the spaces of all equivalence classes of integrable and square-integrable $\mathcal{F}_t$-measurable random variables are isomorphic to each other, then the necessity for the closure in $L_t^1(\mathbb{R}^4)$ vanishes. This is indeed the case when we consider a finite probability space as in Chapter 6 below.

Chapter 6

Implementation of the recursion

This chapter is dedicated to the recursive implementation of the dynamic set-valued structures obtained in Sections 5.2 and 5.3. For computational reasons, throughout this chapter, we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a finite and atomized probability space where each atom $\omega \in \Omega$ is non-trivial, that is, we have $\mathbb{P}(\{\omega\}) > 0$. We note that, as the underlying probability space is finite, for every $v_0 \in \mathbb{R}_+$ and $t \in \mathbb{T}$, the maximum and minimum wealths over all outcomes, that is,

$$v_t^- := \min_{\substack{\pi \in \Phi_0(v_0) \\ \omega \in \Omega}} v_t^\pi(\omega) \quad \text{and} \quad v_t^+ := \max_{\substack{\pi \in \Phi_0(v_0) \\ \omega \in \Omega}} v_t^\pi(\omega)$$

are attained. Moreover, for every $t \in \mathbb{T}$, and $p \in [1, \infty)$ we have $L_t^p(\mathbb{R}^n) = L_t^\infty(\mathbb{R}^n)$, and $L_t^\infty(\mathbb{R}^n)$ is isomorphic to a finite dimensional Euclidean space, that is, $L_t^\infty \simeq \mathbb{R}^m$ for some $m \in \mathbb{N}$.

6.1 Backward and forward algorithms

After revealing the dynamic behavior of graph $\mathcal{P}_t$, the next natural step is to develop a methodology for calculating graph $\mathcal{P}_t$ for every initial $v_0 \in \mathbb{R}_+$ recursively. Algorithm 1 is developed for this purpose and it is a direct implementation of the backward recursion provided by the one-step problem (G_t) combined with the geometric procedure in Corollary 5.3.5. We note that as we work on a finite probability space, by Remark 5.3.6, the projection property $\operatorname{graph}\mathcal{P}_t = \mathcal{U}_t^P$ holds without the necessity for the closure.

Algorithm 1: Backward recursive computation of $\operatorname{graph}\mathcal{P}_t$

Input: Initial wealth $v_0 \in \mathbb{R}_+$
Set $\operatorname{graph}\mathcal{P}_T = \{(v_T, x) \in L_T^2(\mathbb{R}^2) \times L_T^1(\mathbb{R}) \mid x_1 \geq v_T, x_2 \geq v_T^2\}$.
Set $t = T - 1$.
while $t \in \mathbb{T}$ **do**
  Solve (G_t) using $\operatorname{graph}\mathcal{P}_{t+1}$ as input, and obtain the upper image $\mathcal{U}_t$.
  Compute $\mathcal{U}_t^P$ using Corollary 5.3.5, that is, compute the projection of $\mathcal{U}_t$ on the subspace $\mathbf{1}^\mathsf{T} x = 0$.
  Set $\operatorname{graph}\mathcal{P}_t = \mathcal{U}_t^P$.
  Set $t = t - 1$.
end
Set $\mathcal{P}_0(v_0) = \{x \in \mathbb{R}^2 \mid (v_0, x) \in \operatorname{graph}\mathcal{P}_t\}$.
Output: $\mathcal{P}_0(v_0)$, $\operatorname{graph}\mathcal{P}_0, \ldots, \operatorname{graph}\mathcal{P}_{T-1}$

After obtaining $\mathcal{P}_0(v_0)$ as a consequence of Algorithm 1, the next step for our dynamic programming methodology is to establish the bridge between the mean-variance and mean-second moment weights at initial time. This will bypass the issues of the scalar dynamic programming principle that we explore in Chapter 4. However, as opposed to Lemma 3.1.4, we do not have a parallel result when the ordering cone is chosen to be $L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)$, that is, for every $t \in \mathbb{T} \setminus \{T\}$ we have

$$\widehat{\mathcal{T}}_t [\mathcal{P}_t(v_t)] \supsetneq \operatorname{cl}\left(\mathcal{M}_t(v_t) + L_t^2(\mathbb{R}_+) \times L_t^1(\mathbb{R}_+)\right)$$

in general. In particular, at the initial time, we have $\mathcal{T}[\mathcal{P}_0(v_0)] \neq \operatorname{cl}\left(\mathcal{M}_0(v_0) + \mathbb{R}_+^2\right)$, and consequently,

$$\inf_{x \in \mathcal{P}_0(v_0)} \lambda_0^\mathsf{T} \mathcal{T}(x) \neq \inf_{x \in \mathcal{A}_0(v_0)} \lambda_0^\mathsf{T} \mathcal{T}(x) \tag{6.1}$$

in general. Therefore, the lack of equality here poses an issue when trying to calculate the corresponding weight $\bar{\rho}_0$ of the mean-second moment problem for a given initial weight $\lambda_0 \in \mathbb{R}^2_+$.

To overcome these issues, we introduce an additional step to find a sufficiently small subset of $\mathcal{P}_0(v_0)$ for which we can recover the equality of optimal values in (6.1) when we replace $\mathcal{P}_0(v_0)$ with it. Let us define

$$\tilde{\mathcal{B}} \coloneqq \operatorname*{arg\,min}_{x \in \mathcal{P}_0(v_0)} x_2 \tag{6.2}$$

and fix some

$$\tilde{x} \in \operatorname*{arg\,min}_{x \in \tilde{\mathcal{B}}} x_1.$$

Let us also define

$$\tilde{\mathcal{P}}_0(v_0) \coloneqq \{x \in \mathcal{P}_0(v_0) \mid x_1 \leq \tilde{x}_1\}.$$

The next result shows that $\tilde{\mathcal{P}}_0(v_0)$ is a sufficiently small subset of $\mathcal{P}_0(v_0)$ for our purposes.

Proposition 6.1.1. *For every $v_0 \in \mathbb{R}_+$, and $\lambda_0 \in \mathbb{R}^2_+$ we have*

$$\inf_{x \in \tilde{\mathcal{P}}_0(v_0)} \lambda_0^{\mathsf{T}} \mathcal{T}(x) = \inf_{x \in \mathcal{A}_0(v_0)} \lambda_0^{\mathsf{T}} \mathcal{T}(x).$$

Proof. Notice that under Assumption 3.0.4, for $\lambda_0 = (0,1)^{\mathsf{T}} \in \mathbb{R}^2_+$ the optimal value of $\mathrm{A}_0(v_0, \lambda_0)$ is attained. Therefore, by (6.2), we observe that $\tilde{x} \in \mathcal{A}_0(v_0)$, since for every $x \in \tilde{\mathcal{B}}$ we have $x - \bar{x} \in \mathbb{R}^2_+$. Moreover, again by Assumption 3.0.4, as $v_0 \in \mathbb{R}_+$, we have $\mathbb{E}\left(v_T^{\pi^\star}\right) \geq 0$ for every optimal solution $\pi^\star$ of $\mathrm{M}_0(v_0, \lambda_0)$. Thus, as $\lambda_0 \in \mathbb{R}^2_+$, by Theorem 3.2.1(ii), we also have $\bar{\rho}_0 \in \mathbb{R}^2_+$ for the corresponding weight of the mean-second moment problem. As a result of Lemma 3.1.5 and Theorem 3.2.1,

$$\inf_{x \in \overline{\mathcal{A}_0(v_0)}} \lambda_0^{\mathsf{T}} \mathcal{T}(x) = \inf_{x \in \mathcal{A}_0(v_0)} \lambda_0^{\mathsf{T}} \mathcal{T}(x) = \inf_{x \in \mathcal{A}_0(v_0)} \bar{\rho}_0^{\mathsf{T}} x = \inf_{x \in \overline{\mathcal{A}_0(v_0)}} \bar{\rho}_0^{\mathsf{T}} x.$$

Now, let us consider some $x \in \overline{\mathcal{A}_0(v_0)}$ such that $x_1 > \tilde{x}_1$. Note that, by construction, we have $r = x_2 - \tilde{x}_2 \geq 0$. Hence, $\tilde{x} + (0,r)^{\mathsf{T}} \in \overline{\mathcal{A}_0(v_0)}$. Then,

$$\bar{\rho}_0^{\mathsf{T}}(\tilde{x} + (0,r))^{\mathsf{T}} = \bar{\rho}_{0,1}\tilde{x}_1 + \bar{\rho}_{0,2}(\tilde{x}_2 + r) = \bar{\rho}_{0,1}\tilde{x}_1 + \bar{\rho}_{0,2}x_2 \leq \bar{\rho}_{0,1}x_1 + \bar{\rho}_{0,2}x_2 = \bar{\rho}_0^{\mathsf{T}} x$$

as $\bar{\rho}_0 \in \mathbb{R}_+$. Therefore, it suffices to only consider $x \in \overline{\mathcal{A}_0(v_0)}$ such that $x_1 \leq \tilde{x}_1$, that is, if we define $\tilde{\mathcal{A}}_0(v_0) = \{x \in \overline{\mathcal{A}_0(v_0)} \mid x_1 \leq \tilde{x}_1\}$, then we have

$$\inf_{x \in \tilde{\mathcal{A}}_0(v_0)} \lambda_0^{\mathsf{T}} \mathcal{T}(x) = \inf_{x \in \mathcal{A}_0(v_0)} \lambda_0^{\mathsf{T}} \mathcal{T}(x).$$

Thus, showing that $\tilde{\mathcal{A}}_0(v_0) = \tilde{\mathcal{P}}_0(v_0)$ will finish the proof. To that end, observe that the inclusion $\tilde{\mathcal{A}}_0(v_0) \subseteq \tilde{\mathcal{P}}_0(v_0)$ is straightforward as $\mathcal{A}_0(v_0) \subseteq \mathcal{P}_0(v_0)$ under Remark 5.2.1. For the reverse inclusion, let $x \in \tilde{\mathcal{P}}_0(v_0)$ and initially assume that there exist $a \in \mathcal{A}_0(v_0)$ and $r \in \mathbb{R}^2_+$ such that $x = a + r$. Notice that, if $r_1 = 0$ we already have $x \in \tilde{\mathcal{A}}_0(v_0)$, hence suppose that $r_1 > 0$. Then, let

$$\lambda = \frac{r_1}{\tilde{x}_1 - a_1}, \text{ and } \bar{x} = \lambda x + (1 - \lambda)\tilde{x}.$$

Note that, we have $\lambda \in (0, 1)$ and as $\overline{\mathcal{A}_0(v_0)}$ is convex by Lemma 3.1.2, we have $\bar{x} \in \overline{\mathcal{A}_0(v_0)}$. Moreover, by construction we observe that $\bar{x} + (0, r) = x$ for some $r \in \mathbb{R}_+$, hence we conclude $x \in \tilde{\mathcal{A}}_0(v_0)$. Finally, let $x \in \tilde{\mathcal{P}}_0(v_0)$, then there exists $(a^n + r^n)_{n \in \mathbb{N}}$ such that

$$\lim_{n \to \infty} a^n + r^n = x \text{ in } \mathbb{R}^2_+. \tag{6.3}$$

Then, as $\overline{\mathcal{A}_0(v_0)}$ is closed by Lemma 3.1.2, by our previous arguments, we have $x \in \tilde{\mathcal{P}}_0(v_0)$. $\square$

Thanks to Proposition 6.1.1, the set $\tilde{\mathcal{M}}_0(v_0) \coloneqq \mathcal{T}[\tilde{\mathcal{P}}_0(v_0)]$ is an appropriate subset of $\mathcal{M}_0(v_0) + \mathbb{R}^2_+$ to calculate the corresponding weight vector $\bar{\rho}_0$ of the mean-second moment problem. This procedure is summarized as Algorithm 2.

We stress that, finding $\tilde{\mathcal{B}}$ first can be impractical in general. However, if $\mathcal{P}_0(v_0)$, or at least an approximation of it, is polyhedral, then the computation of $\tilde{\mathcal{B}}$ turns out to be trivial. In fact, we satisfy this condition as a result of our solution methodology for graph $\mathcal{P}_t$, which is explored in detail in Section 6.2.

As the forward part of our proposed dynamic programming methodology, we develop Algorithm 3, which utilizes the outputs of both Algorithms 1 and 2.

Algorithm 2: Computation of the mean-second moment weight $\bar{\rho}_0$

Input: Initial wealth $v_0 \in \mathbb{R}_+$, mean-variance weight $\lambda_0 \in \mathbb{R}^2_+$, upper image $\mathcal{P}_0(v_0)$

Solve $\min\limits_{x \in \mathcal{P}_0(v_0)} x_2$ and find the corresponding set $\tilde{\mathcal{B}}$ of all optimal solutions.

Find $x \in \tilde{\mathcal{B}}$ with minimal first component, that is, solve $\min\limits_{x \in \tilde{\mathcal{B}}} x_1$ and find its optimal solution $\tilde{x}$.

Set $\tilde{\mathcal{P}}_0(v_0) \coloneqq \{x \in \mathcal{P}_0(v_0) \mid x_1 \leq \tilde{x}_1\}$.

Compute $\tilde{\mathcal{M}}_0(v_0) = \{\mathcal{T}(x) \mid x \in \tilde{\mathcal{P}}_0(v_0)\}$.

Solve $\min\limits_{x \in \tilde{\mathcal{M}}_0(v_0)} \lambda_0^\mathsf{T} x$ and obtain its optimal solution $x^\star$.

Compute $\bar{\rho}_0 = (\lambda_{0,1} + 2\lambda_{0,2}\tilde{x}_1,\, \lambda_{0,2})^\mathsf{T}$.

Output: Corresponding equivalent weight $\bar{\rho}_0$ for the mean-second moment problem

For any given initial wealth $v_0 \in \mathbb{R}_+$, and weight $\lambda_0 \in \mathbb{R}^2_+$, Algorithm 3 returns an optimal solution for the mean-variance problem $(\mathrm{M}_0(v_0, \lambda_0))$, the corresponding time-consistent weight process $(\lambda_t)_{t \in \mathbb{T} \setminus \{T\}}$ and the optimal value process $(\mathbb{E}_t(v_T), \mathbb{E}_t\left((v_T)^2\right))_{t \in \mathbb{T}}$. We note that, its results are justified by Corollary 5.2.7.

Algorithm 3: Forward computation of the time-consistent weight process

Input: $v_0 \in \mathbb{R}_+$, $\lambda_0 \in \mathbb{R}^2_+$, $\bar{\rho}_0 \in \mathbb{R}^2_+$, graph $\mathcal{P}_1, \dots,$ graph $\mathcal{P}_T$

Set $t = 0$.

while $t \in \mathbb{T} \setminus \{T\}$ **do**

  Solve $(\mathrm{R}_t(v_t, \bar{\rho}_0))$ using graph $\mathcal{P}_{t+1}$ and obtain an optimal solution $(\pi^\star, x^\star)$.

  Set $v_{t+1} = v^{\pi^\star}_{t,t+1}$.

  Set $\pi_{t+1} = \pi^\star$.

  Set $x_t = x^\star$.

  Compute $\lambda_t = (\bar{\rho}_{0,1} - 2\bar{\rho}_{0,2}\mathbb{E}_t(x_{t,1}),\, \lambda_{0,2})^\mathsf{T}$

  Set $t = t + 1$.

end

Output: A time-consistent weight process $(\lambda_t)_{t \in \mathbb{T} \setminus \{T\}}$, an optimal solution $(\pi_t)_{t \in \mathbb{T} \setminus \{0\}}$ and optimal value process $(x_t)_{t \in \mathbb{T}}$

6.2 Solution methodology for (G_t) and $\operatorname{graph}\mathcal{P}_t$

In this section, we discuss the computational implementation of the algorithms in Section 6.1. We note that it is possible to use a variety of different algorithms for *solving* a convex vector optimization problem in the literature. Hence, line 4 of Algorithm 1 should be considered as calling one such algorithm as a subroutine to solve (G_t). For our purposes, we use the primal Benson-type algorithm proposed in [24], which returns a finite weak ϵ-solution (see Definition 6.2.2 below) as well as inner and outer polyhedral approximations for the upper image, using a given approximation error $\epsilon > 0$.

Remark 6.2.1. The reader can notice that, throughout the thesis, we avoid using the ordering cones $\{0\} \times L^1_t(\mathbb{R}_+)$ and $\{0\} \times L^2_t(\mathbb{R}_+) \times L^1_t(\mathbb{R}_+)$. The main reason is the fact that, even under a finite probability space, both cones have an empty interior. However, the primal algorithm in [24], and generally other finite-dimensional convex vector optimization algorithms in the literature, requires nonempty interior for the ordering cone. Hence, for computational reasons, it becomes impractical to formulate vector optimization problems with ordering cones $\{0\} \times L^1_t(\mathbb{R}_+)$ and $\{0\} \times L^2_t(\mathbb{R}_+) \times L^1_t(\mathbb{R}_+)$.

The primal algorithm in [24] assumes that the problem has a compact feasible region, this is to ensure that certain scalarization problems that are solved within the algorithm have optimal solutions [24, Assumption 4.1]. In our case, the vector optimization problem (G_t) does not satisfy this assumption due to the unbounded set constraint $(v^\pi_{t,t+1}, x) \in \operatorname{graph}\mathcal{P}_{t+1}$. For this reason, we will introduce the following box constraints to (G_t):

$$\kappa_t^{v^-} \leq v_t \leq \kappa_t^{v^+}, \qquad \kappa^{x^-} \leq x \leq \kappa^{x^+},$$

where the corresponding bounds are set as $\kappa_t^{v^-} := v_t^-$, $\kappa_t^{v^+} := v_t^+$, $\kappa^{x^-} := (v_T^-, (v_T^-)^2)^\mathsf{T}$ and $\kappa^{x^+} := (v_T^+, (v_T^+)^2)^\mathsf{T}$ for each $t \in \mathbb{T}\setminus\{T\}$. With these constraints, we introduce the bounded analogue of $\operatorname{graph}\mathcal{P}_t$ as

$$\overline{\operatorname{graph}}\,\mathcal{P}_t := \left\{(v_t, x) \in \operatorname{graph}\mathcal{P}_t \mid \kappa_t^{v^-} \leq v_t \leq \kappa_t^{v^+},\, \kappa^{x^-} \leq x \leq \kappa^{x^+}\right\}. \tag{6.4}$$

Furthermore, by Proposition 5.3.3, we rewrite $\overline{\operatorname{graph}}\,\mathcal{P}_t$ in an equivalent form in terms of $\operatorname{graph}\mathcal{P}_{t+1}$ as

$$\overline{\operatorname{graph}}\,\mathcal{P}_t = \operatorname{cl}\left\{ (v_t, \mathbb{E}_t(x)) \;\middle|\; \begin{array}{c} \kappa_t^{v^-} \le v_t \le \kappa_t^{v^+},\ \kappa^{x^-} \le x \le \kappa^{x^+} \\ \pi^{\mathsf{T}} S_t = v_t,\ \varphi_t(\pi) \ge 0,\ (v^{\pi}_{t,t+1}, x) \in \operatorname{graph}\mathcal{P}_{t+1} \end{array} \right\}.$$

Finally, for each $v_t \in L_t^{\infty}(\mathbb{R})$, and $\pi \in L_t^{\infty}(\mathbb{R}^d)$ such that $\kappa_t^{v^-} \le v_t \le \kappa_t^{v^+}$, $\pi^{\mathsf{T}} S_t = v_t$, and $\varphi_t(\pi) \ge 0$, thanks to (6.4), we have $(v^{\pi}_{t,t+1}, x) \in \overline{\operatorname{graph}}\,\mathcal{P}_{t+1}$ for every $x \in L_t^{\infty}(\mathbb{R}^2)$ such that $\kappa^{x^-} \le x \le \kappa^{x^+}$ and $(v^{\pi}_{t,t+1}, x) \in \operatorname{graph}\mathcal{P}_{t+1}$. Thus, we indeed preserve the dynamic relation in Proposition 5.3.3 when we switch to the bounded setting, that is, we have

$$\overline{\operatorname{graph}}\,\mathcal{P}_t = \operatorname{cl}\left\{ (v_t, \mathbb{E}_t(x)) \;\middle|\; \begin{array}{c} \kappa_t^{v^-} \le v_t \le \kappa_t^{v^+},\ \kappa^{x^-} \le x \le \kappa^{x^+} \\ \pi^{\mathsf{T}} S_t = v_t,\ \varphi_t(\pi) \ge 0,\ (v^{\pi}_{t,t+1}, x) \in \overline{\operatorname{graph}}\,\mathcal{P}_{t+1} \end{array} \right\}. \tag{6.5}$$

Therefore, we introduce the following convex vector optimization problem associated to $\overline{\operatorname{graph}}\,\mathcal{P}_t$, which is indeed the bounded analogue of (G_t), and has a compact feasible region:

$$\begin{array}{lll} \text{minimize} & \begin{pmatrix} v_t \\ \mathbb{E}_t(x) \\ -v_t - \mathbf{1}^{\mathsf{T}}\mathbb{E}_t(x) \end{pmatrix} & \text{with respect to} \quad \le_{L_t^{\infty}(\mathbb{R}^4_+)} \\ \text{subject to} & \pi^{\mathsf{T}} S_t = v_t & \\ & (v^{\pi}_{t,t+1}, x) \in \overline{\operatorname{graph}}\,\mathcal{P}_{t+1} & \\ & \kappa_t^{v^-} \le v_t \le \kappa_t^{v^+} & \\ & \kappa^{x^-} \le x \le \kappa^{x^+} & \\ & \varphi_t(\pi) \ge 0. & \end{array} \tag{$\overline{\mathrm{G}}_t$}$$

The corresponding image set and the upper image of the updated problem $(\overline{\mathrm{G}}_t)$, similarly, are the bounded analogues of the original problem (G_t), and they are given by

$$\overline{\mathcal{I}}_t = \left\{ \left(v_t, \mathbb{E}_t(x), -v_t - \mathbf{1}^{\mathsf{T}}\mathbb{E}_t(x)\right)^{\mathsf{T}} \in \mathcal{I}_t \mid \kappa_t^{v^-} \le v_t \le \kappa^{v^+},\ \kappa^{x^-} \le x \le \kappa^{x^+} \right\},$$

$$\overline{\mathcal{U}}_t = \mathrm{cl}(\overline{\mathcal{I}}_t + L_t^\infty(\mathbb{R}^4_+)),$$

respectively. We also use the approximate solution concept in [24, Definition 3.3] for $(\overline{\mathrm{G}}_t)$ as we recall next.

Definition 6.2.2. *A set $\mathcal{X}$ of feasible solutions is called a* finite weak ϵ-solution *of $(\overline{\mathrm{G}}_t)$ if, $\mathcal{X}$ consists of only weakly efficient solutions (see Definition 5.1.2 or [24, Section 2]), and*

$$\mathrm{conv} \bigcup_{(v_t,x)\in\mathcal{X}} \begin{pmatrix} v_t \\ \mathbb{E}_t(x) \\ -v_t - \mathbf{1}^\mathsf{T}\mathbb{E}_t(x) \end{pmatrix} + L_t^\infty(\mathbb{R}^4) - \epsilon\mathbf{1} \supseteq \overline{\mathcal{U}}_t.$$

For conciseness, we denote the outer polyhedral approximations of a finite weak ϵ-solution $\mathcal{X}$ of (G_t) by $\mathscr{U}_t(\mathcal{X})$, where

$$\mathscr{U}_t(\mathcal{X}) = \mathrm{conv}\left\{\left(v_t, \mathbb{E}_t(x), -v_t - \mathbf{1}^\mathsf{T}\mathbb{E}_t(x)\right)^\mathsf{T} \mid (v_t, x) \in \mathcal{X}\right\} + L_t^\infty(\mathbb{R}^4).$$

We note that, the conclusions of Proposition 5.3.4 and Corollary 5.3.5 continue to hold for the analogue bounded image $\overline{\mathcal{I}}_t$ and upper image $\overline{\mathcal{U}}_t$, as we are interested in only a particular portion of the domain where the projection occurs. Our observation can be justified by the fact that both results do not depend on the structure of the feasible region of (G_t). Thus, we have the following analogous relation between $\overline{\mathrm{graph}}\,\mathcal{P}_t$ and $\overline{\mathcal{U}}_t$ by Corollary 5.3.5 and Remark 5.3.6:

$$\overline{\mathrm{graph}}\,\mathcal{P}_t = \overline{\mathcal{U}}_t^P, \tag{6.6}$$

where $\overline{\mathcal{U}}_t^P := \{x \in L_t^\infty(\mathbb{R}^3) \mid \exists y \in L_t^\infty(\mathbb{R})\colon (x, -y)^\mathsf{T} \in \overline{\mathcal{U}}_t,\ \mathbf{1}^\mathsf{T}x = y\}$. In addition to satisfying the compactness assumption of the primal algorithm in [24] for computational purposes, passing to a bounded setting allows us to consider $\overline{\mathrm{graph}}\,\mathcal{P}_t$ as a canonical bounded projection of $\overline{\mathcal{I}}_t$, in the spirit of [22], onto its first three components. However, as we work on a probabilistic setup, the main difference is the inclusion of the additional dimension for the objective occurs for each $\omega \in \Omega$ as opposed a single dimension. Hence, we are in a trade-off between the precision of the approximations, and the problem complexity with respect to the inclusion of additional objective dimensions. Therefore, in accordance with [22, Definition 3.2], we define a finite ϵ-solution of $\overline{\operatorname{graph}}\,\mathcal{P}_t$ next.

Definition 6.2.3. *A set $\mathcal{X}$ of feasible solutions of $(\overline{\mathrm{G}}_t)$ is called a* finite ϵ-solution *of $\overline{\operatorname{graph}}\,\mathcal{P}_t$ if*

$$\operatorname{conv} \bigcup_{(v_t,x)\in\mathcal{X}} \begin{pmatrix} v_t \\ \mathbb{E}_t(x) \end{pmatrix} + B_t^{\infty}(0,\epsilon) \supseteq \overline{\operatorname{graph}}\,\mathcal{P}_t,$$

where $B_t^{\infty}(0,\epsilon) \coloneqq \{x \in L_t^{\infty}(\mathbb{R}^3) \mid \|x\|_{\infty} \leq \epsilon\}$.

Our next result is motivated by [22, Theorem 3.12], and it is aimed to establish the bridge between the finite weak ϵ-solutions of the vector optimization problem $(\overline{\mathrm{G}}_t)$, and the finite ϵ-solutions of $\overline{\operatorname{graph}}\,\mathcal{P}_t$.

Proposition 6.2.4. *Let $\epsilon > 0$. If $\mathcal{X}$ is a finite weak ϵ-solution of $(\overline{\mathrm{G}}_t)$, then it is a finite 6ϵ-solution of $\overline{\operatorname{graph}}\,\mathcal{P}_t$.*

Proof. Note that the outer polyhedral approximation of $\overline{\mathcal{U}}_t$ under the finite weak ϵ-solution $\mathcal{X}$ is given by $\mathscr{U}_t(\mathcal{X}) = \mathcal{C}_t(\mathcal{X}) + L_t^{\infty}(\mathbb{R}^4_+)$, where

$$\mathcal{C}_t(\mathcal{X}) \coloneqq \operatorname{conv}\left(\bigcup_{(v_t,x)\in\mathcal{X}} \begin{pmatrix} v_t \\ \mathbb{E}_t(x) \\ -v_t - \mathbf{1}^{\mathsf{T}}\mathbb{E}_t(x) \end{pmatrix}\right).$$

Initially, we claim that, for every $(c,\bar{c}) \in \mathcal{C}_t(\mathcal{X})$, where $c \in L_t^{\infty}(\mathbb{R}^3)$ and $\bar{c} \in L_t^{\infty}(\mathbb{R})$, we have $\mathbf{1}^{\mathsf{T}}c = \bar{c}$. Let $(c,\bar{c}) \in \mathcal{C}_t(\mathcal{X})$, then without any loss of generality we have

$$c = \sum_{k=1}^{m} \lambda_k c_k, \text{ and } \bar{c} = \sum_{k=1}^{m} \lambda_k \bar{c}_k$$

for some $\lambda_k \in [0,1]$ and $(c_k,\bar{c}_k) \in \{\big(v_t, \mathbb{E}_t(x), -v_t - \mathbf{1}^{\mathsf{T}}\mathbb{E}_t(x)\big)^{\mathsf{T}} \mid (v_t,x) \in \mathcal{X}\}$ such that $c \in L_t^{\infty}(\mathbb{R}^3)$, $\bar{c} \in L_t^{\infty}(\mathbb{R})$ and $\mathbf{1}^{\mathsf{T}}c_k = \bar{c}_k$ for every $k \in \{1,\ldots,m\}$. Hence, the claim holds as

$$\mathbf{1}^\mathsf{T} c = \mathbf{1}^\mathsf{T} \sum_{k=1}^{m} \lambda_k c_k = \sum_{k=1}^{m} \lambda_k (\mathbf{1}^\mathsf{T} c_k) = \sum_{k=1}^{m} \lambda_k \bar{c}_k = \bar{c}.$$

As $\mathcal{X}$ is a finite weak ϵ-solution of $(\overline{\mathrm{G}}_t)$, we have $\mathscr{U}_t(\mathcal{X}) - \mathbf{1}\epsilon \supseteq \overline{\mathcal{U}}_t$. Furthermore when we apply the transformation (5.13), that is, to take the projection on the subspace $\mathbf{1}^\mathsf{T} x = 0$, by Corollary 5.3.5 and (6.6), we obtain that

$$\left\{x \in L_t^\infty(\mathbb{R}^3) \mid \exists y \in L_t^\infty(\mathbb{R})\colon (x, -y)^\mathsf{T} \in \mathscr{U}_t(\mathcal{X}) - \epsilon \mathbf{1} \ , \ \mathbf{1}^\mathsf{T} x = y\right\} \supseteq \overline{\operatorname{graph}} \, \mathcal{P}_t. \tag{6.7}$$

As the next step, let $(x, -y^\mathsf{T}) \in \mathscr{U}_t(\mathcal{X}) - \epsilon\mathbf{1}$ such that $\mathbf{1}^\mathsf{T} x = y$. Then, by our previous observation, there exist $(c, -\mathbf{1}^\mathsf{T} c) \in \mathcal{C}_t(\mathcal{X})$ and $r = (\bar{r}, r')^\mathsf{T} \in L_t^\infty(\mathbb{R}_+^3) \times L_t^\infty(\mathbb{R}_+)$ such that $x = c + \bar{r} - \epsilon\mathbf{1}$ and $y = \mathbf{1}^\mathsf{T} c - r' + \epsilon$. Thus,

$$\mathbf{1}^\mathsf{T} x = y \iff \mathbf{1}^\mathsf{T} c + \mathbf{1}^\mathsf{T} \bar{r} - 3\epsilon = \mathbf{1}^\mathsf{T} c - r' + \epsilon \iff \mathbf{1}^\mathsf{T} r = 4\epsilon.$$

Therefore, as $r \in L_t^\infty(\mathbb{R}_+^4)$, we have $\|r\|_\infty \leq 4\epsilon$. Hence, $\bar{r} - \epsilon\mathbf{1} \in B_t^\infty(0, 6\epsilon)$ since

$$\|\bar{r} - \epsilon\mathbf{1}\|_\infty \leq \|\bar{r}\|_\infty + \epsilon\,\|\mathbf{1}\|_\infty \leq \|r\|_\infty + \epsilon\,\|\mathbf{1}\|_\infty \leq 4\epsilon + \sqrt{3}\epsilon \leq 6\epsilon. \tag{6.8}$$

Moreover, as a result of (6.8), we conclude that the superset relation (6.7) can be extended as

$$\left\{x \in L_t^\infty(\mathbb{R}^3) \mid \exists y \in L_t^\infty(\mathbb{R})\colon (x, -y)^\mathsf{T} \in \mathcal{C}_t(\mathcal{X}) \ , \ \mathbf{1}^\mathsf{T} x = y\right\} + B_t^\infty(0, 6\epsilon) \supseteq \overline{\operatorname{graph}} \, \mathcal{P}_t.$$

Finally, by the preceding arguments, we observe that the projection of $\mathcal{C}_t(\mathcal{X})$ on the subspace $\mathbf{1}^\mathsf{T} x = 0$ is canonical, that is,

$$\operatorname{conv}\left(\bigcup_{(v_t, x) \in \mathcal{X}} \begin{pmatrix} v_t \\ \mathbb{E}_t(x) \end{pmatrix}\right) + B_t^\infty(0, 6\epsilon) \supseteq \overline{\operatorname{graph}} \, \mathcal{P}_t,$$

as desired. □

After establishing the solution methodology for both the vector optimization problem ($\overline{\mathrm{G}}_t$), and the convex projection of $\overline{\operatorname{graph}}\, \mathcal{P}_t$, we move on to incorporating this methodology for the backward recursive computations, namely Algorithm 1. To that end, we update Algorithm 1 as Algorithm 4 such that, at each iteration, we utilize a polyhedral approximation that we obtain in the previous step to replace $\overline{\operatorname{graph}}\, \mathcal{P}_t$.

Algorithm 4: Polyhedral approximation of $\operatorname{graph} \mathcal{P}_t$

Input: Initial wealth $v_0 \in \mathbb{R}_+$, tolerance levels $\epsilon_0, \dots, \epsilon_{T-1}$

Set $\overline{\operatorname{graph}}\, \mathcal{P}_T = \{(v_T, x) \in L_T^\infty(\mathbb{R}^3) \mid \kappa_T^{v^-} \leq v_T \leq \kappa_T^{v^+},\ \kappa^{x^-} \leq x \leq \kappa^{x^+}\}$.

Solve ($\overline{\mathrm{G}}_t$) with tolerance level ϵ_{T-1} by the primal algorithm in [24].

Obtain a finite weak ϵ-solution $\mathcal{X}_{T-1}$ and its corresponding outer polyhedral approximation $\mathscr{U}_{T-1}(\mathcal{X}_{T-1})$.

Compute the projection of $\mathscr{U}_{T-1}(\mathcal{X}_{T-1})$ on the subspace $\mathbf{1}^\mathsf{T} x = 0$ and obtain $\mathscr{U}_{T-1}^P(\mathcal{X}_{T-1})$.

Set $\mathscr{G}_{T-1} = \mathscr{U}_{T-1}^P(\mathcal{X}_{T-1})$.

Set $t = T - 2$.

while $t \in \mathbb{T}$ **do**

- Replace $\overline{\operatorname{graph}}\, \mathcal{P}_{t+1}$ with the polyhedral approximation $\mathscr{G}_{t+1}$, and solve ($\overline{\mathrm{G}}_t$) with tolerance level ϵ_t by the primal algorithm in [24].
- Obtain a finite weak ϵ-solution $\mathcal{X}_t$ and its corresponding outer polyhedral approximation $\mathscr{U}_t(\mathcal{X}_t)$.
- Compute the projection of $\mathscr{U}_t(\mathcal{X}_t)$ on the subspace $\mathbf{1}^\mathsf{T} x = 0$ and obtain $\mathscr{U}_t^P(\mathcal{X}_t)$.
- Set $\mathscr{G}_t = \mathscr{U}_t^P(\mathcal{X})$.
- Set $t = t - 1$.

end

Set $\mathscr{P}_0(v_0) = \{x \in \mathbb{R}^2 \mid (v_0, x) \in \mathscr{G}_0\}$.

Output: Polyhedral approximations of $\mathcal{P}_0(v_0)$, $\overline{\operatorname{graph}}\, \mathcal{P}_0, \dots, \overline{\operatorname{graph}}\, \mathcal{P}_{T-1}$ as $\mathscr{P}_0$, $\mathscr{G}_0, \dots, \mathscr{G}_{T-1}$, respectively.

Remark 6.2.5. For a finite weak ϵ-solution $\mathcal{X}$ of ($\overline{\mathrm{G}}_t$), its corresponding outer polyhedral approximation $\mathscr{U}_t(\mathcal{X})$ approximates $\overline{\mathcal{U}}_t$ by at most 2ϵ, that is,

$d_H(\mathscr{U}_t(\mathcal{X}),\overline{\mathcal{U}}_t) < 2\epsilon$, where d_H is the usual Hausdorff distance given by

$$d_H(\mathscr{U}_t(\mathcal{X}),\overline{\mathcal{U}}_t) = \sup_{x\in\mathscr{U}_t(\mathcal{X})} \inf_{y\in\overline{\mathcal{U}}_t} \|x-y\|_\infty . \tag{6.9}$$

Moreover, as we implement in Algorithm 4, if we take the projection of $\mathscr{U}_t(\mathcal{X})$ on the subspace $\mathbf{1}^\mathsf{T}x = 0$, the resulting polyhedron $\mathscr{G}_t$ approximates $\overline{\operatorname{graph}}\,\mathcal{P}_t$ by at most 6ϵ, that is, $d_H(\mathscr{G}_t,\overline{\operatorname{graph}}\,\mathcal{P}_t) < 6\epsilon$. Further, under Proposition 6.2.4 it holds

$$\mathscr{G}_t = \operatorname{conv}\left\{(v_t,\mathbb{E}_t(x))^\mathsf{T} \mid (v_t,x)\in\mathcal{X}\right\}$$

Thereby, in accordance with Definition 6.2.3, $d_H(\mathscr{G}_t,\overline{\operatorname{graph}}\,\mathcal{P}_t) < 6\epsilon$ is an equivalent condition for $\mathcal{X}$ to be a finite 6ϵ-solution of $\overline{\operatorname{graph}}\,\mathcal{P}_t$.

Our following result indicates that, when we replace $\overline{\operatorname{graph}}\,\mathcal{P}_t$ with its polyhedral approximation $\mathscr{G}_t$ for each $t\in\mathbb{T}\backslash\{T\}$, the error of the previous approximations propagate to the following iterations in a linear form. That is, at each $t\in\mathbb{T}\backslash\{T\}$, the polyhedral approximation $\mathscr{G}_t$ approximates $\overline{\operatorname{graph}}\,\mathcal{P}_t$ by at most the accumulative sum of the previous errors.

Proposition 6.2.6. *Let $v_0\in\mathbb{R}_+$, and $(\epsilon_t)_{t\in T\backslash\{T\}}$ be sequence of strictly positive tolerance levels. Then, for every $t\in\mathbb{T}\backslash\{T\}$ the polyhedral approximation $\mathscr{G}_t$ which is returned by Algorithm 4 approximates $\overline{\operatorname{graph}}\,\mathcal{P}_t$ by at most $\bar{\epsilon}_t$, that is, $\mathcal{X}_t$ is a finite $\bar{\epsilon}_t$-solution of $\overline{\operatorname{graph}}\,\mathcal{P}_t$, where*

$$\bar{\epsilon}_t \coloneqq 6\sum_{k=t}^{T-1}\epsilon_k.$$

Proof. Note that, by Proposition 6.2.4 and Remark 6.2.5, $\mathcal{X}_{T-1}$ is a $\bar{\epsilon}_{T-1}$-solution of $\overline{\operatorname{graph}}\,\mathcal{P}_{T-1}$, where $\bar{\epsilon}_{T-1} = 6\epsilon_{T-1}$. Now, let us assume that the assertion holds for some $t+1\in\mathbb{T}$. As the first step, we define the following set, where we replace $\overline{\operatorname{graph}}\,\mathcal{P}_{t+1}$ with its polyhedral approximation $\mathscr{G}_{t+1}$ on (6.5):

$$\mathscr{T}_t \coloneqq \operatorname{cl}\left\{(v_t,\mathbb{E}_t(x)) \;\middle|\; \exists(\pi,x)\colon \begin{array}{c} \kappa_t^{v^-}\le v_t\le\kappa_t^{v^+},\ \kappa^{x^-}\le x\le\kappa^{x^+} \\ \pi^\mathsf{T}S_t = v_t,\ \varphi_t(\pi)\ge 0,\ (v_{t,t+1}^\pi,x)\in\mathscr{G}_{t+1}\end{array}\right\}.$$

Note that $\mathscr{T}_t$ can be seen as the best, in the sense of approximation error, theoretical approximation of $\overline{\operatorname{graph}}\,\mathcal{P}_t$, when we solve ($\overline{\mathrm{G}}_t$) while replacing $\overline{\operatorname{graph}}\,\mathcal{P}_{t+1}$ with its polyhedral approximation $\mathscr{G}_{t+1}$. Moreover, for each $v_t \in L_t^\infty(\mathbb{R})$, and $\pi \in L_t^\infty(\mathbb{R}^d)$ such that $\kappa_t^{v^-} \leq v_t \leq \kappa_t^{v^+}$, $\pi^\mathsf{T} S_t = v_t$, and $\varphi_t(\pi) \geq 0$, we observe that for every $x \in L_t^\infty(\mathbb{R}^2)$ such that $\kappa^{x^-} \leq x \leq \kappa^{x^+}$ with $(v^\pi_{t,t+1}, x) \in \mathscr{G}_{t+1}$, we have

$$\sup_{y \in \overline{\operatorname{graph}}\,\mathcal{P}_{t+1}} \left\| (v^\pi_{t,t+1}, x)^\mathsf{T} - y \right\|_\infty \leq 6 \sum_{k=t+1}^{T-1} \epsilon_k = \bar{\epsilon}_{t+1},$$

as $\mathcal{X}_{t+1}$ is a finite $\bar{\epsilon}_{t+1}$-solution of $\overline{\operatorname{graph}}\,\mathcal{P}_{t+1}$. Therefore, under (6.5) we observe that

$$\mathscr{T}_t + \{0\} \times \mathbb{E}_t\left[B^\infty_{t+1}(0, \bar{\epsilon}_{t+1})\right] \supseteq \mathscr{T}_t + \{0\} \times B^\infty_t(0, \bar{\epsilon}_{t+1}) \supseteq \overline{\operatorname{graph}}\,\mathcal{P}_t,$$

where the first inclusion is due to the contraction property of conditional expectation. Thus, in particular we have

$$\mathscr{T}_t + B^\infty_t(0, \bar{\epsilon}_{t+1}) \supseteq \overline{\operatorname{graph}}\,\mathcal{P}_t. \tag{6.10}$$

Moreover, under the construction of Algorithm 4, in a similar spirit to Proposition 6.2.4, we observe that $\mathcal{X}_t$ is a finite $6\epsilon_t$-solution of $\mathscr{T}_t$, that is, under Remark 6.2.5 we have

$$\mathscr{G}_t + B^\infty_t(0, 6\epsilon_t) \supseteq \mathscr{T}_t. \tag{6.11}$$

Therefore, by combining (6.10) and (6.11), we arrive at the desired inclusion as

$$\mathscr{G}_t + B^\infty_t(0, 6\epsilon_t) + B^\infty_t(0, \bar{\epsilon}_{t+1}) = \mathscr{G}_t + B^\infty_t(0, \bar{\epsilon}_t) \supseteq \overline{\operatorname{graph}}\,\mathcal{P}_t$$

Whence, the polyhedral approximation $\mathscr{G}_t$ approximates $\overline{\operatorname{graph}}\,\mathcal{P}_t$ by at most $\bar{\epsilon}_t$, or equivalently, $\mathcal{X}_t$ is a finite $\bar{\epsilon}_t$-solution of $\overline{\operatorname{graph}}\,\mathcal{P}_t$. Thus, the result for each $t \in \mathbb{T} \setminus \{T\}$ follows by induction. $\square$

6.3 Computational results

For all the examples that are covered in this section, we only consider financial markets with shortselling constraints, that is, we take the additional market constraints $(\varphi_t)_{t\in\mathbb{T}\setminus\{T\}}$ to be the identity mapping for each $t \in \{0,\ldots,T-1\}$. Moreover, for every financial market we consider, we assume that the vector-valued price process $(S_t)_{t\in\mathbb{T}}$ has stationary and independent increments. As a result, the graph of the set-valued function $\mathcal{P}_t$ is identical under every realization. That is, for every $\omega_1, \omega_2 \in \Omega$, we have graph $\mathcal{P}_t(\omega_1) =$ graph $\mathcal{P}_t(\omega_2)$. Hence, for each $t \in \mathbb{T}\setminus\{T\}$, there exists $H_t \subseteq \mathbb{R}^3$ such that

$$\overline{\operatorname{graph}}\,\mathcal{P}_t = H_t \times \ldots \times H_t, \tag{6.12}$$

where $\overline{\operatorname{graph}}\,\mathcal{P}_t(\omega) = H_t$ for each $\omega \in \Omega$. Therefore, it suffices to calculate H_t in isolation, and utilize the relation (6.12) to recover $\overline{\operatorname{graph}}\,\mathcal{P}_t$. That is, rather than solving the stochastic vector optimization problem $\overline{\mathrm{G}}_t$, we may only consider a realization of it for an arbitrary $\omega \in \Omega$, that is, we solve the following four-dimensional vector optimization problem:

$$\begin{aligned}
&\text{minimize} \quad \begin{pmatrix} v_t(\omega) \\ \mathbb{E}_t(x)(\omega) \\ -v_t(\omega) - \mathbf{1}^{\mathsf{T}}\mathbb{E}_t(x)(\omega) \end{pmatrix} \quad \text{with respect to} \quad \leq_{\mathbb{R}^4_+} \\
&\text{subject to} \quad \pi^{\mathsf{T}} S_t(\omega) = v_t(\omega) \\
&\qquad\qquad\quad (\pi^{\mathsf{T}} S_{t+1}, x) \in H_{t+1} \times \ldots \times H_{t+1} \\
&\qquad\qquad\quad \kappa_t^{v^-} \leq v_t(\omega) \leq \kappa_t^{v^+} \\
&\qquad\qquad\quad \kappa_t^{x^-} \leq x \leq \kappa_t^{x^+} \\
&\qquad\qquad\quad \pi \geq 0.
\end{aligned} \tag{6.13}$$

Therefore, the resulting upper image of (6.13) constitutes H_t, and we may recover $\overline{\operatorname{graph}}\,\mathcal{P}_t$ under (6.12). We note that, this advancement has no significant effect on the results of Section 6.2, and the implementation of Algorithm 4 in particular, other than decreasing the problem complexity significantly.

Remark 6.3.1. We acknowledge that our observations are unique for financial markets where the underlying price process is Lévy, and they mail fail to hold in general. In such cases, the computation of $\overline{\operatorname{graph}}\,\mathcal{P}_t$ can suffer from the curse of dimensionality greatly as the dimension of the underlying finite probability space increases. Although it is interesting study such markets, and to develop efficient methods for the computation purposes, we prefer to leave it for a future study.

For the conciseness of the examples, we introduce the following notation for each $t \in \mathbb{T}\backslash\{T\}$ and $i \in \{1, \ldots, d\}$ as

$$\tilde{\lambda}_t = \frac{\lambda_{t,1}}{\lambda_{t,1} + \lambda_{t,2}} \text{ and } \tilde{\pi}^i_{t+1} = \frac{\pi^i_{t+1}}{v_t},$$

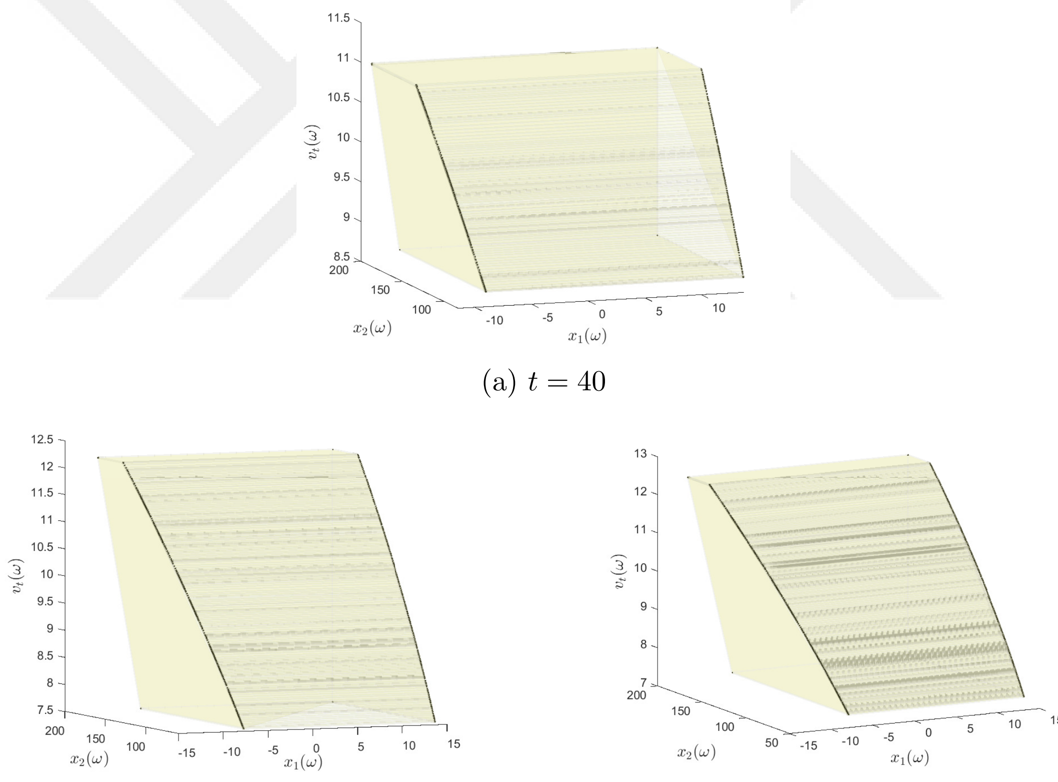
where $\tilde{\lambda}_t$ denotes percentage weight assigned to the expected terminal wealth, whereas $\tilde{\pi}^i_t$ denotes portfolio values as the percentage of current wealth of the investor rather than the physical units. We present our computational results under two different market structures in Example 6.3.2 and Example 6.3.3 below.

Example 6.3.2. For our first example, we consider a trinomial market model with two stocks. Daily trading over half a year is considered, where the market parameters are selected such that, the negatively correlated stock prices $(S^1_t)_{t\in\mathbb{T}}$ and $(S^2_t)_{t\in\mathbb{T}}$ admit 6.4%, and 6.2% semi-annual mean return, respectively. The investor enters the market with initial wealth $v_0 = 10$.

For the illustration of the outputs of Algorithm 4, we present the polyhedral approximations of $\overline{\operatorname{graph}}\,\mathcal{P}_t$ for three different periods on Figure 6.1

We compute the optimal portfolios $(\tilde{\pi}^\star_t)_{t\in\mathbb{T}\backslash\{0\}}$ of the dynamic-mean variance problem by our dynamic programming methodology for three different values of initial weight $\tilde{\lambda}_0$. For each $\tilde{\lambda}_0 \in \{0.3, 0.375, 0.45\}$, we report the time-consistent weight process $(\tilde{\lambda}_t)_{t\in\mathbb{T}\backslash\{T\}}$, the value process $(v^\pi_t)_{t\in\mathbb{T}}$, the conditional expectation process $\left(\mathbb{E}_t\left(v^{\pi^\star}_T\right)\right)_{t\in\mathbb{T}}$ and the conditional variance process $\left(\operatorname{Var}_t\left(v^{\pi^\star}_T\right)\right)_{t\in\mathbb{T}}$ along one path on Figure 6.2.

We note that, for the plots of the portfolio positions $\tilde{\pi}^\star$, at each period the area dedicated each color represent the percentage of the current wealth assigned to their respective assets. Thereby, we can observe that as the investor becomes less risk averse as the time progresses, which we can observe by the increase in the time consistent weight process $(\tilde{\lambda}_t)_{t\in\mathbb{T}\setminus\{T\}}$, the investor tends to assign a larger portion of their current wealth to the first asset, which is the one that accepts a higher return with higher risk. Moreover, we can observe that when we alter the initial weight vector $\tilde{\lambda}_0$, the resulting pictures shows a significant resemblance to each other, which suggests a smoothness property on the initial condition. Though it is interesting to investigate such a behavior, we leave it for a future study.



(a) $t = 40$

(b) $t = 75$

(c) $t = 90$

Figure 6.1: Polyhedral approximations of $\overline{\operatorname{graph}}\,\mathcal{P}_t$ for $t \in \{40, 75, 90\}$.

Example 6.3.3. Initially, let us consider a five-dimensional Black-Scholes model, where $S_0 = \mathbf{1} \in \mathbb{R}^5_+$ and the continuous price dynamics are given by

$$\mathrm{d}S^i_t = S^i_t(r\ \mathrm{d}t + \sigma_i\ \mathrm{d}W^i_t),$$

for each $i \in \{1, \dots, 5\}$. Further, let us allow the Brownian motions $(W^i_t)_{t\in\mathbb{R}_+}$ and $(W^j_t)_{t\in\mathbb{R}_+}$ to be correlated with some constant correlation coefficient $\rho_{ij} \neq 0$ for every $i, j \in \{1, \ldots, 5\}$ such that $i \neq j$. Moreover, we take $r = 0.1$, and the other market parameters are taken as follows:

$$\sigma = (0.2,\, 0.5,\, 0.15,\, 0.3,\, 0.25)^{\mathsf{T}}\,,\ \rho = \begin{bmatrix} 1 & 0.5 & -0.2 & -0.1 & -0.3 \\ 0.5 & 1 & -0.4 & 0.4 & -0.1 \\ -0.2 & -0.4 & 1 & 0.25 & -0.3 \\ -0.1 & 0.4 & 0.25 & 1 & -0.1 \\ -0.3 & -0.1 & -0.3 & -0.1 & 1 \end{bmatrix}.$$

In order to transition into a discrete time setting and apply our dynamic solution methodology, we utilize the decoupling approach proposed by [25] for discretization of the continuous time price process $(S_t)_{t\in\mathbb{R}_+}$. Accordingly, we introduce the log-price dynamics as $X_t^i = \ln\left(S_t^i\right)$ for every $i \in \{1, \ldots, 5\}$ and $t \in \mathbb{R}_+$ where

$$\mathrm{d}X_t^i = \left(r - \frac{1}{2}\sigma_i^2\right)\,\mathrm{d}t + \sigma_i +\,\mathrm{d}W_t^i. \tag{6.14}$$

Note that under (6.14) the state dependence in the diffusion coefficients are no longer present. The main idea behind this method is to apply an appropriate transformation, that converts the original five correlated Brownian motions into independent ones. To that end, [25] exploit the fact that the covariance matrix is positive semi-definite and accepts an Chloesky factorization C. Then under [25, Proposition 1] the column vectors of C, applying on the process $(X_t)_{t\in\mathbb{R}_+}$, is an adequate transformation and the resulting processes are indeed independent. After achieving the independence, the discretization methodology is fairly straightforward for a given discretization interval Δt and inline with the standard methods in the literature. For the resulting discrete time price process, at each node of the finite probability tree there are 2^5 child nodes, where each of them share the same transition probability due do the discretization methodology. Finally, under [25, Proposition 6] the approximated discrete price process converges weakly, in the sense of distribution, to the original price process $(S_t)_{t\in\mathbb{R}_+}$. For more detail on the subject the reader is referred to [25, Section 3].

Accordingly, we consider a weakly trading over a year, hence we take $\Delta t = 1/52$. Moreover, we denote the discretized price process by $(\tilde{S}_t)_{t\in\mathbb{T}}$. Investor enters this market with initial wealth $v_0 = 1$. Similar to our first example, we compute the optimal portfolios $(\tilde{\pi}_t^\star)_{t\in\mathbb{T}\setminus\{0\}}$ of the dynamic-mean variance problem by our dynamic programming methodology for three different values of initial weight $\tilde{\lambda}_0$. For each $\tilde{\lambda}_0 \in \{0.24,\ 0.29,\ 0.37\}$, we report the time consistent weight process $(\tilde{\lambda}_t)_{t\in\mathbb{T}\setminus\{T\}}$, the value process $(v_t^\pi)_{t\in\mathbb{T}}$, the conditional expectation process $\left(\mathbb{E}_t\left(v_T^{\pi^\star}\right)\right)_{t\in\mathbb{T}}$ and the conditional variance process $\left(\mathrm{Var}_t\left(v_T^{\pi^\star}\right)\right)_{t\in\mathbb{T}}$ along one path on Figure 6.3.

Initially, we observe that the smoothness property that we deduced in the previous example continues to hold in a more complex market structure, that is, both the time consistent weight process and the behavior of the investor in terms of the portfolio positions shows smoothness when the initial weight $\tilde{\lambda}_0$ varies. In particular, we can observe the investor's behavior towards the risk more clearly between the periods fifteen and thirty, where the risk aversion of the investor increases, which propagates itself as a diversification in the portfolio positions. We note that, this observation is relevant for each initial weight $\tilde{\lambda}_0$, however, the magnitude of the response changes significantly based on the initial risk aversion of the investor. Furthermore, we observe that after period thirty, as the market moves against the investor, which in return increases the weight $\tilde{\lambda}_t$, the diversification slowly decreases and the investor allocates more and more wealth to the second asset, which accepts the highest return among all five assets.

Let us consider the optimal portfolio $\tilde{\pi}^\star$ for $\tilde{\lambda} = 0.24$ on Figure 6.3, and for each $t \in \{0,\ldots,52\}$ let $x_t^\star = \left(\mathbb{E}_t\left(v_T^{\pi^\star}\right), \mathrm{Var}_t\left(v_T^{\pi^\star}\right)\right)$, where $(x_t^\star)_{t\in\mathbb{T}}$ denotes the process of the optimal values. Then we present the dynamic movement of the process $(x_t^\star)_{t\in\mathbb{T}}$ on the upper images of the dynamic mean-variance problem along one path on Figure 6.4. Indeed, we observe that the optimal values of the portfolio $\tilde{\pi}^\star$ moves on the so called efficient frontier of the mean-variance pairs for every $t \in \mathbb{T}\setminus\{T\}$ under the time consistent weight process $(\tilde{\lambda}_t)_{t\in\mathbb{T}\setminus\{T\}}$.

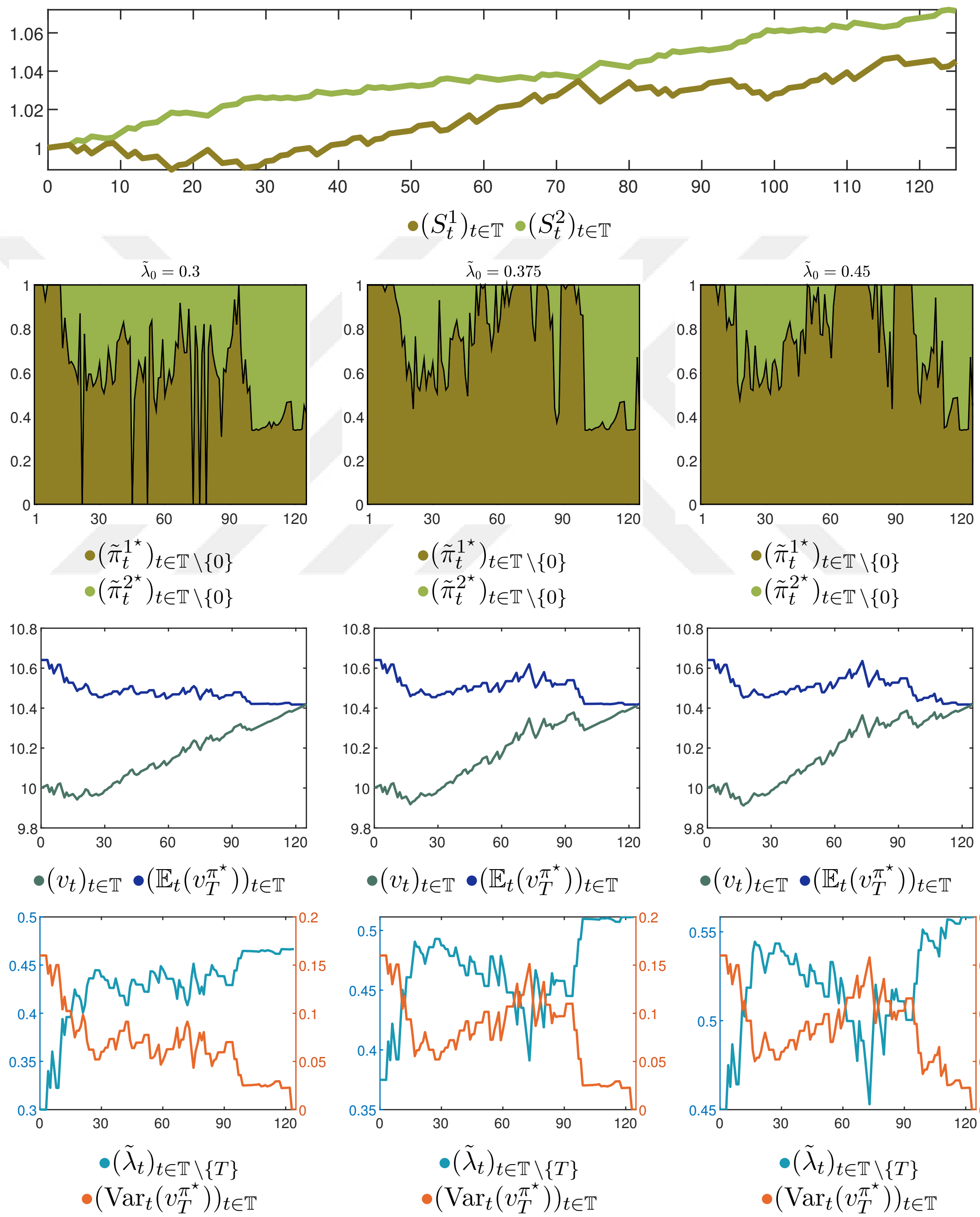


Figure 6.2: The processes $(S_t)_{t\in\mathbb{T}}$, $(\pi_t^\star)_{t\in\mathbb{T}\setminus\{0\}}$, $(\tilde{\lambda}_t)_{t\in\mathbb{T}}$, $(v_t)_{t\in\mathbb{T}}$, $(\mathbb{E}_t(v_T^{\pi^\star}))_{t\in\mathbb{T}}$ and $(\mathrm{Var}_t(v_T^{\pi^\star}))_{t\in\mathbb{T}}$ for $\tilde{\lambda}_0 \in \{0.3, 0.375, 0.45\}$ along one path.

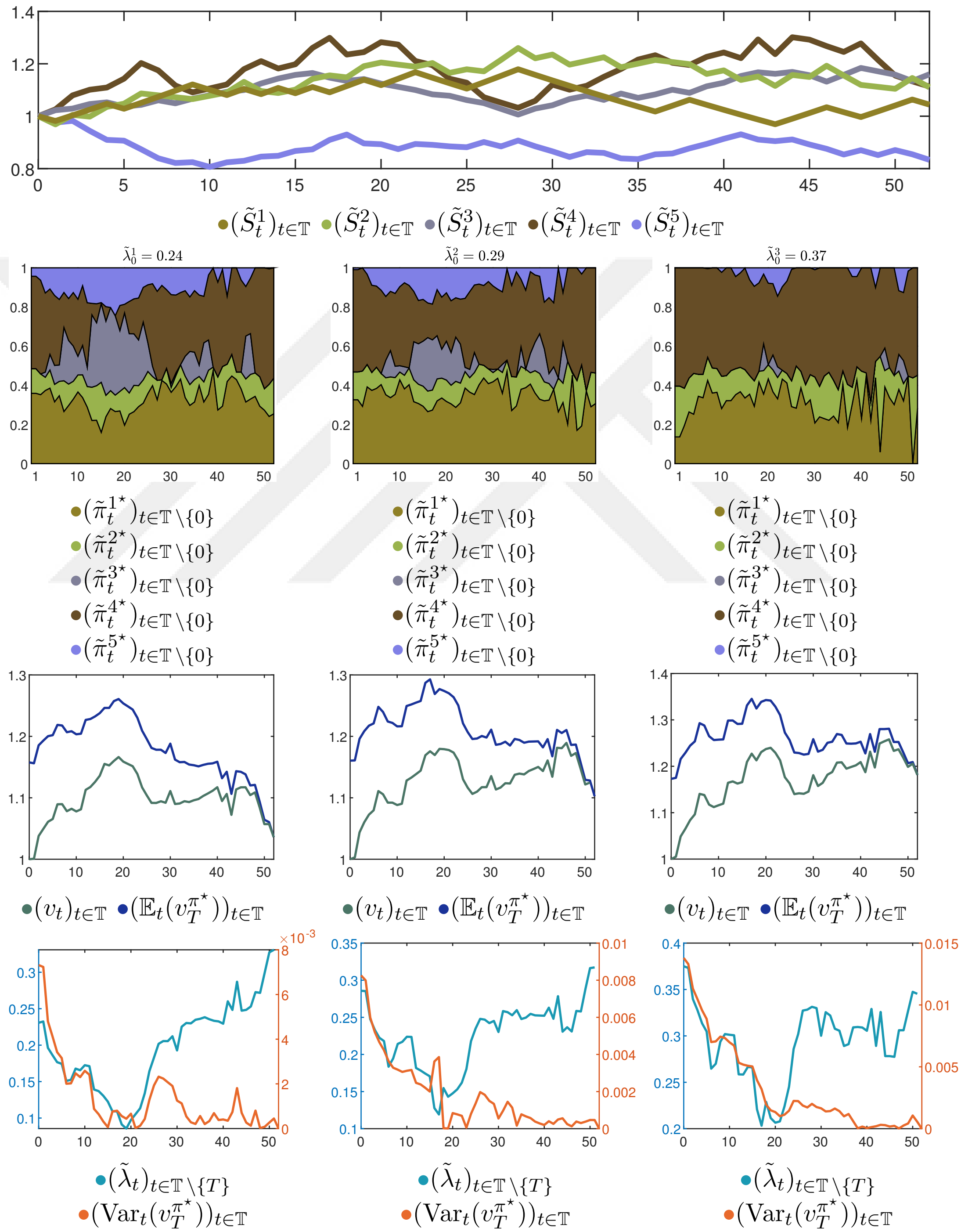


Figure 6.3: The processes $(\tilde{S}_t)_{t\in\mathbb{T}}$, $(\pi^\star_t)_{t\in\mathbb{T}\setminus\{0\}}$, $(\tilde{\lambda}_t)_{t\in\mathbb{T}\setminus\{T\}}$, $(v_t)_{t\in\mathbb{T}}$, $(\mathbb{E}_t(v_T^{\pi^\star}))_{t\in\mathbb{T}}$ and $(\mathrm{Var}_t(v_T^{\pi^\star}))_{t\in\mathbb{T}}$ for $\tilde{\lambda}_0 \in \{0.24, 0.29, 0.37\}$ along one path.

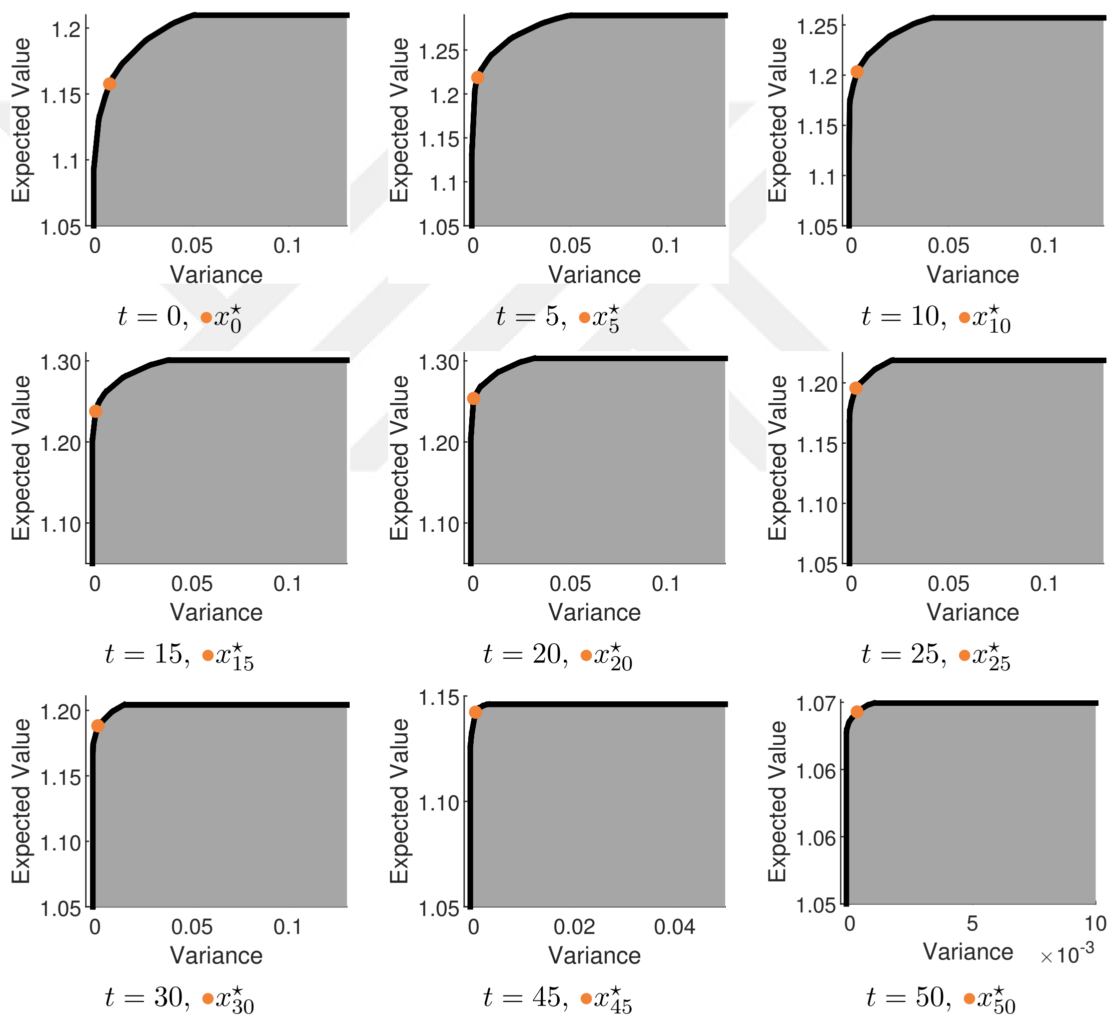


Figure 6.4: Dynamic movement of the optimal objective process $(x_t^\star)_{t\in\mathbb{T}}$ on the upper images along one path.

Chapter 7

Conclusion

In this thesis, we investigate the time-consistency of the dynamic mean-variance problem, where we utilize a less strict notion of time-consistency with Definition 3.0.3. As the initial step, we investigate the equivalence between the mean-variance and mean-second moment problems in Section 3.2. By Theorem 3.2.1 we establish that, when the weights of the mean-variance and the mean-second moment problems are selected approppirately, they become inherently connected as they share the same set of optimal solutions. We move on to study the time-consistency of the mean-second moment problem in Section 3.3, and by Proposition 3.3.1 we conclude that the dynamic-mean second moment problem is time-consistent in the classical sense, i.e., with a constant weight process. Finally, we combine the results of the previous two sections, and provide a full characterization of the time-consistent weight process for every initial mean-variance weight $\lambda_0 \in \mathbb{R}_{++}$ and initial wealth $v_0 \in \mathbb{R}_+$ by Theorem 3.3.2.

In light of Section 3.3, we observe that the mean-second moment problem is time-consistent, hence satisfies a scalar Bellman's equation. Therefore, in Chapter 4, by Proposition 4.0.1 we develop a scalar dynamic programming principle for the dynamic mean-second moment problem. However, we observe that the scalar Bellman's equation is fundamentally inadequate in terms of its utilization for the dynamic study of the mean-variance problem as long as the optimal solutions do not have a closed form expression. Therefore, in order to bypass the issues of the scalar formulation we extend to vector optimization in Chapter 5. We start by formulating the mean-second moment problem as a biobjective optimization problem and observe that it is a natural generalization of its scalar counterpart. We move on to establishing a set-valued Bellman's equation in Section 5.2, where we utilize the upper image $\mathcal{P}_t$ as our set-valued value function, and by Theorem 5.2.3 we obtain the set-valued analogue the results of Chapter 4. Furthermore, we extend the results of the set-valued dynamic programming for graphs of the set-valued function $\mathcal{P}_t$ for computational purposes.

We propose our forward-backward dynamic programming methodology in Section 6.1. With Algorithm 1 we dynamically compute graph $\mathcal{P}_t$ in a backward recursive manner and obtain the upper image $\mathcal{P}_0(v_0)$ at the initial time. Moreover, we observe that we need an additional step for the computation of the equivalent mean-second moment weight $\bar{\rho}_0$ for a given mean-variance weight λ_0. Therefore we introduce Algorithm 2 which establishes the procedure for this purpose. As the last step, with Algorithm 3 we compute the time consistent weight process and an optimal solution for the dynamic mean-variance problem for every initial weight λ_0 and initial wealth v_0. Finally, we demonstrate our dynamic solution methodology on two financial market settings in Chapter 6.

Bibliography

[1] H. Markowitz, "Portfolio selection," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[2] I. Bajeux-Besnainou and R. Portait, "Dynamic asset allocation in a mean-variance framework," *Management Science*, vol. 44, no. 11-part-2, pp. S79–S95, 1998.

[3] X. Y. Zhou and D. Li, "Continuous-time mean-variance portfolio selection: A stochastic lq framework," *Applied Mathematics and Optimization*, vol. 42, pp. 19–33, Jan 2000.

[4] D. Li and W.-L. Ng, "Optimal dynamic portfolio selection: Multiperiod mean-variance formulation," *Mathematical Finance*, vol. 10, no. 3, pp. 387–406, 2000.

[5] A. E. B. Lim and X. Y. Zhou, "Mean-variance portfolio selection with random parameters in a complete market," *Mathematics of Operations Research*, vol. 27, no. 1, pp. 101–120, 2002.

[6] X. Li, X. Y. Zhou, and A. E. B. Lim, "Dynamic mean-variance portfolio selection with no-shorting constraints," *SIAM Journal on Control and Optimization*, vol. 40, no. 5, pp. 1540–1555, 2002.

[7] R. H. Strotz, "Myopia and inconsistency in dynamic utility maximization," *The Review of Economic Studies*, vol. 23, no. 3, pp. 165–180, 1955.

[8] S. Basak and G. Chabakauri, "Dynamic Mean-Variance Asset Allocation," *Review of Financial Studies*, vol. 23, pp. 2970–3016, August 2010.

[9] C. Czichowsky, “Time-consistent mean-variance portfolio selection in discrete and continuous time,” *Finance and Stochastics*, vol. 17, pp. 227–271, Apr 2013.

[10] T. Björk, A. Murgoci, and X. Y. Zhou, “Mean–variance portfolio optimization with state-dependent risk aversion,” *Mathematical Finance*, vol. 24, no. 1, pp. 1–24, 2014.

[11] T. Björk, M. Khapko, and A. Murgoci, “On time-inconsistent stochastic control in continuous time,” *Finance and Stochastics*, vol. 21, pp. 331–360, Apr 2017.

[12] C. Karnam, J. Ma, and J. Zhang, “Dynamic approaches for some time-inconsistent optimization problems,” *The Annals of Applied Probability*, vol. 27, no. 6, pp. 3435 – 3477, 2017.

[13] X. Cui, D. Li, S. Wang, and S. Zhu, “Better than dynamic mean-variance: Time inconsistency and free cash flow stream,” *Mathematical Finance*, vol. 22, no. 2, pp. 346–378, 2012.

[14] G. Kováčová and B. Rudloff, “Time consistency of the mean-risk problem,” *Operations Research*, Published online in Articles in Advance 2021.

[15] H. H. Schaefer, *Banach Lattices and Positive Operators*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1974.

[16] I. Molchanov, *Theory of Random Sets*. London: Springer London, 2017.

[17] J. Jahn, *Vector Optimization: Theory, Applications, and Extensions*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.

[18] S.-M. Grad and E.-L. Pop, “Vector duality for convex vector optimization problems by means of the quasi-interior of the ordering cone,” *Optimization*, vol. 63, no. 1, pp. 21–37, 2014.

[19] C. Zalinescu, “On three open problems related to quasi relative interior,” 2014.

[20] L. Narici and E. Beckenstein, *Topological vector spaces, second edition.* 01 2010.

[21] A. H. Hamel, F. Heyde, A. Löhne, B. Rudloff, and C. Schrage, "Set optimization—a rather short introduction," in *Set Optimization and Applications - The State of the Art* (A. H. Hamel, F. Heyde, A. Löhne, B. Rudloff, and C. Schrage, eds.), (Berlin, Heidelberg), pp. 65–141, Springer Berlin Heidelberg, 2015.

[22] G. Kováčová and B. Rudloff, "Convex projection and convex vector optimization," 2021.

[23] A. Löhne and B. Weißing, "Equivalence between polyhedral projection, multiple objective linear programming and vector linear programming," *Mathematical Methods of Operations Research*, vol. 84, pp. 411–426, Oct 2016.

[24] A. Löhne, B. Rudloff, and F. Ulus, "Primal and dual approximation algorithms for convex vector optimization problems," *Journal of Global Optimization*, vol. 60, pp. 713–736, Dec 2014.

[25] R. Korn and S. Müller, "The decoupling approach to binomial pricing of multi-asset options," *Journal of Computational Finance*, vol. 12, pp. 1–30, 2009.