DISS. ETH NO. 28838

DYNAMIC PROGRAMMING APPROACHES FOR THE MEAN–VARIANCE PORTFOLIO SELECTION PROBLEM

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

ZHOUYI TAN

M.S. in Mathematics
New York University, United States
born on April 22, 1992
citizen of China

accepted on the recommendation of

Prof. Dr. Martin Schweizer	examiner
Prof. Dr. Beatrice Acciaio	co-examiner
Prof. Dr. Aleš Černý	co-examiner

2022

DOI: 10.3929/ethz-b-000590444
ISBN: 978-3-907363-09-6

To my wife, Cong Xie.
致我的妻子，谢璁。

Abstract

In this thesis, we develop new dynamic programming approaches for solving the mean–variance portfolio selection (MVPS) problem in both discrete and continuous time. Let $G_T(\theta)$ be the final wealth of a self-financing strategy θ investing in the underlying assets S. The MVPS problem consists of maximising $E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)]$ over a suitable set Θ of stochastic processes θ for a risk aversion parameter $\xi > 0$.

Chapters I and II develop results for the market cloning technique in discrete and continuous time, respectively. This approach consists of constructing independent copies of the market, studying the auxiliary problem to maximise the expectation of empirical mean minus ξ times empirical variance of L individual copies of the final wealth, and passing to the limit as L goes to infinity. To tackle these auxiliary problems, we use dynamic programming. In Chapter I, a systematic backward recursive computation leads to both the value process and an optimal strategy for each auxiliary problem in finite discrete time. In Chapter II, with a continuum number of time steps, such a recursion is no longer available. We use a guess-and-verify procedure for solving the auxiliary problem for continuous processes. In both chapters, our general framework allows us to go beyond the i.i.d. innovations or Brownian-driven SDE models typically assumed in the current literature for this kind of approach.

In Chapter III, we develop a deterministic dynamic programming principle (DPP) for a general class of open-loop McKean–Vlasov control problems in finite discrete time. We embed the original problem into a sequence of deterministic tail problems whose criterion and optimisation at time t involve only an expectation, not a conditional expectation, of variables from $t + 1$ onward. This works for controlled processes without specifying any dynamics for the underlying process. The resulting DPP gives in full generality a systematic backward recursion for both the value and an optimal strategy for the MVPS problem in discrete time.

Kurzfassung

In dieser Arbeit entwickeln wir neue Ansätze via dynamische Programmierung zur Lösung des μ-σ-Portfoliooptimierungsproblems sowohl in diskreter als auch in stetige Zeit. Sei $G_T(\theta)$ das Endvermögen einer selbstfinanzierenden Strategie θ, die in die zugrunde liegenden Anlagen S investiert. Das μ-σ-Problem besteht dann aus der Maximierung von $E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)]$ über eine geeignete Menge Θ stochastischer Prozesse θ für einen Risikoaversionsparameter $\xi > 0$.

Kapitel I und II entwickeln Ergebnisse für die sogenannte Marktklontechnik in diskreter bzw. stetiger Zeit. Dieser Ansatz besteht darin, unabhängige Kopien des Marktes zu konstruieren, als Hilfsproblem dann Erwartung des empirischen Mittelwerts minus ξ mal empirische Varianz von L einzelnen Kopien des Endvermögens zu maximieren, und den Limes für $L \to \infty$ zu betrachten. Um die Hilfsprobleme anzugehen, verwenden wir dynamische Programmierung. In Kapitel I führt eine systematische rückwärts rekursive Berechnung sowohl zum Wertprozess als auch zu einer optimalen Strategie für jedes Hilfsproblem in endlicher diskreter Zeit. In Kapitel II, mit einer kontinuierlichen Anzahl von Zeitschritten, ist eine solche Rekursion nicht mehr verfügbar. Wir verwenden stattdessen ein "guess-and-verify"-Verfahren zur Lösung des Hilfsproblems für stetige Prozesse. In beiden Kapiteln erlaubt uns unser allgemeiner Rahmen, über die Modelle hinauszugehen, die typischerweise in der aktuellen Literatur für diese Art von Ansatz angenommen werden, nämlich i.i.d. Innovationen oder stochastische Differentialgleichungen, die von einer Brownschen Bewegung getrieben werden.

In Kapitel III entwickeln wir ein deterministisches Prinzip der dynamischen Programmierung (DPP) für eine allgemeine Klasse von open-loop-McKean–Vlasov-Kontrollproblemen in endlicher diskreter Zeit. Wir betten das ursprüngliche Problem in eine Folge von deterministischen Endstück-Problemen ein, deren Kriterium und Optimierung zum Zeitpunkt t nur eine Erwartung, keine bedingte Erwartung, von Variablen ab $t + 1$ beinhaltet. Dies funktioniert für kontrollierte Prozesse, ohne eine Dynamik für den zugrunde liegenden Prozess vorzugeben.

Das resultierende DPP liefert in voller Allgemeinheit eine systematische Rückwärtsrekursion sowohl für den Wert als auch für eine optimale Strategie für das μ-σ-Problem in diskreter Zeit.

Acknowledgements

First, I want to express my sincerest gratitude to my advisor Martin Schweizer for his trust, guidance and constant support throughout my journey at ETH Zurich. His unique blend of mathematical taste, rigour and elegance has set a perfect example for me and profoundly shaped my view of mathematics. I will never forget his patience and encouragement, especially when things do not go as expected. I also wish to extend my thanks to Beatrice Acciaio and Aleš Černý for agreeing to act as co-examiners and for their valuable feedback on this thesis.

Second, my appreciation must go to my wonderful colleagues and friends in Zurich. I want to thank my officemates and colleagues Bálint, Chen, Daniel, Martin, Philippe, Philipp and Wahid, for sharing the past years and having many diverse and stimulating conversations in our double-storey office. I also want to thank Dàniel, David, Robert, Marco, Hanna, Chiara, Daria, Moritz, Laurin, Florian, Jakob and all members of Group 3 for creating such a cordial atmosphere. I am very fortunate to spend my leisure time with Chen, Chong, Bian, Jinzhou and Songyan. My flatmates Hanlin, Yilin and Yinglu gave me a warm and smooth first year in Zurich.

Third, I must thank all the teachers and friends in New York, Berkeley and Dalian while switching to mathematics from a distant subject. Constant support and encouragement from my master's thesis advisor Raghu Varadhan and professors Yuri Bakhtin, Percy Deift and Georg Stadler strongly motivated my pursuit of a PhD. Along the way, I had a wonderful time with my friends Zhe, Licong, Yukuan, Zhuolun, Hongxi, Zhilei, Jize, Huize, Ning, Dai, Chenyu, Wei, Kunchun, Ke, Zheng, Yihan and Guolong.

Last but not least, I want to thank my whole family, particularly my parents, grandparents, parents-in-law and my wife, for their love, patience, and unlimited confidence in me. Without them, I could have never reached where I am.

Contents

Chapter 0

Overview

1 Setup and motivation

Let $T \in \mathbb{N}$ be a time horizon and fix a time index set $\mathbb{T} \subseteq [0, T]$ which contains the terminal time T. Consider a financial market with $d + 1$ assets with $d \in \mathbb{N}$, among which there are d risky assets and one riskless asset. For simplicity, all prices are discounted by the riskless asset and expressed by units of 1. Trading in these assets with respect to a dynamic strategy yields cumulative profits and losses at each time $t \in \mathbb{T}$.

To convert the above concepts into mathematical terms, we consider an $\mathbb{R}^d$-valued stochastic process $S = (S_t)_{t\in\mathbb{T}}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$. The components of S_t represent the discounted prices at time t of the d risky assets in the market. We denote by Θ a suitable set of dynamic trading strategies $\theta = (\theta_t)_{t\in\mathbb{T}}$. For each strategy $\theta \in \Theta$, there is a corresponding gains process $G(\theta) = (G_t(\theta))_{t\in\mathbb{T}}$ recording the cumulative profits and losses of the strategy θ at every time $t \in \mathbb{T}$. For a risk aversion parameter $\xi > 0$, the *mean–variance portfolio selection (MVPS)* problem is to

$$\text{maximise } E[G_T(\theta)] - \xi \text{Var}[G_T(\theta)] \text{ over all } \theta \in \Theta. \tag{1.1}$$

Under mild assumptions, it is not difficult to obtain the existence and uniqueness (in the sense that $G_T(\widehat{\theta})$ is unique) of an optimal strategy $\widehat{\theta}$ for problem (1.1). However, we are more interested in a dynamic description of $\widehat{\theta}$. From this perspective, the problem (1.1) is outside the scope of classic tools from standard stochastic control theory, like dynamic programming or a stochastic maximum principle, due to the variance term in (1.1). However, it turns out that one can construct a solution for (1.1) from a solution for the so-called *pure hedging problem* to

$$\text{minimise } E\big[\big(1 - G_T(\theta)\big)^2\big] \text{ over all } \theta \in \Theta, \tag{1.2}$$

which is a special case of the *mean–variance hedging (MVH)* problem to

$$\text{minimise } E\big[\big(H - G_T(\theta)\big)^2\big] \text{ over all } \theta \in \Theta \tag{1.3}$$

for a given contingent claim H (i.e., random variable $H \in L^2(\mathcal{F}_T, P)$).

The MVH problem as formulated in (1.3) or similarly has been studied extensively since the 1990s; see e.g. Duffie and Richardson [26], Schweizer [58, 59, 62], Rheinländer and Schweizer [56], Gouriéroux et al. [31] for early developments of the general theory, or Schweizer [63] for a survey of the early works. Since then, there has been a lot of literature making progress from different aspects of the general theory; see e.g. Kohlmann and Tang [42], Lim and Zhou [48], Lim [47] and Arai [6], which culminates in a general description of the structure of the solution to the MVH problem given by Černý and Kallsen [17]. Although this storyline is complete to some extent, there are much fewer results which manage to compute the optimal strategy explicitly. For some papers in this direction, we refer to Schweizer [61], Bertsimas et al. [9], Gugushvili [32] and Černý [15] in finite discrete time, and to Laurent and Pham [43], Biagini et al. [10], Hobson [35] and Černý and Kallsen [18] in continuous time.

Our initial motivation is driven by the question whether one can do explicit computations for the MVPS problem (1.1) systematically like for the MVH problem (1.2). However, problems (1.1) and (1.2) are quite distinct from this perspective. One can easily state, at least formally, a dynamic programming principle for (1.2), whereas it is not obvious how to do so for the MVPS problem (1.1). For a more detailed discussion of the problem, see for instance Björk and Murgoci [13], Björk et al. [11] and Björk et al. [12, Chapters 8 and 18]. In this thesis, we attempt to develop tools from a control perspective to tackle the MVPS problem (1.1) in a general probabilistic framework. Since there is a deep correspondence between dynamic programming and the stochastic maximum principle, we focus on giving results based on the former and leave the latter for future research.

The rest of this chapter is organised as follows. We first elaborate on the connection between the MVPS problem and the pure hedging problem. Then we give a quick review of the main existing results for the pure hedging problem

(1.2). The chapter ends with a brief introduction to the approaches proposed in this thesis.

2 An abstract mean–variance portfolio selection (MVPS) problem

The idea of connecting the MVPS problem to an auxiliary linear–quadratic stochastic control (LQSC) problem dates back at least to Li and Ng [45]. This technique is later improved by Sun and Wong [65], Xia and Yan [66] and Fontana and Schweizer [30] to write the solution to the MVPS problem in terms of a solution to the pure hedging problem (1.2). Following [30], we give a simple presentation of this connection. The idea is to forget the temporal structure and look at problem (1.1) statically.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Denote by L^2 the space of all (equivalence classes of) real-valued square-integrable random variables. Consider a non-empty subset $\Gamma \subseteq L^2$. For a fixed risk aversion parameter $\xi > 0$, we consider an abstract/static version of the MVPS problem, namely to

$$\text{maximise } E[g] - \xi \text{Var}[g] \text{ over } g \in \Gamma. \tag{2.1}$$

Note that we have already abstracted away the temporal structure in the original MVPS problem (1.1).

Assumption 2.1. 1) Γ is a closed linear subspace of L^2.

2) Γ does not contain the constant (payoff) 1.

Under Assumption 2.1, we can write uniquely $L^2 = \Gamma \oplus \Gamma^\perp$, where $\Gamma^\perp$ is the orthogonal complement of Γ in L^2. Thus we can decompose any random variable $Y \in L^2$ uniquely into

$$Y = g^Y + \pi(Y) \quad \text{with } g^Y \in \Gamma \text{ and } \pi(Y) \in \Gamma^\perp. \tag{2.2}$$

Now we argue that the solution to the abstract MVPS problem (2.1) can be read off from the solution to the abstract pure hedging problem to

$$\text{minimise } E[(1-g)^2] \text{ over } g \in \Gamma. \tag{2.3}$$

Note that using the notation (2.2), we have $1 = g^1 + \pi(1)$, and thus the solution to (2.3) can be simply written as g^1.

Theorem 2.2. *Suppose that Assumption* 2.1 *is satisfied. Then problem* (2.1) *has a unique solution* $g^{\mathrm{mv}} \in \Gamma$, *explicitly given by*

$$g^{\mathrm{mv}} = \frac{1}{2\xi} \frac{1}{E[1-g^1]}\, g^1. \tag{2.4}$$

Proof. See Fontana and Schweizer [30, Proposition 3.4]. □

To study the abstract pure hedging problem (2.1), it turns out that an important quantity is the *variance-optimal signed* Γ*-martingale measure* defined as follows.

Definition 2.3. A signed measure Q on $(\Omega, \mathcal{F})$ is called a *signed* Γ*-martingale measure* if $Q[\Omega] = 1$, $Q \ll P$ with $\frac{\mathrm{d}Q}{\mathrm{d}P} \in L^2 = L^2(P)$ and

$$E\left[\frac{\mathrm{d}Q}{\mathrm{d}P}\, g\right] = 0 \qquad \text{for all } g \in \Gamma.$$

The set of all signed Γ-martingale measures is denoted by $\mathbb{P}^2_s(\Gamma)$. We call a signed Γ-martingale measure $\widetilde{P}_\Gamma$ *variance-optimal for* Γ if $\widetilde{P}_\Gamma$ satisfies

$$\widetilde{P}_\Gamma = \operatorname*{arg\,min}_{Q \in \mathbb{P}^2_s(\Gamma)} \operatorname{Var}\left[\frac{\mathrm{d}Q}{\mathrm{d}P}\right].$$

Whenever $\mathbb{P}^2_s(\Gamma)$ is nonempty, a variance-optimal $\widetilde{P}_\Gamma$ exists and is unique because it can be obtained by minimising the strictly convex functional $Q \mapsto \|\frac{\mathrm{d}Q}{\mathrm{d}P}\|^2_{L^2}$ over the closed convex set $\mathbb{P}^2_s(\Gamma)$ (where we identify elements of Q of $\mathbb{P}^2_s(\Gamma)$ with their density $\frac{\mathrm{d}Q}{\mathrm{d}P} \in L^2$).

Lemma 2.4. *Suppose that Assumption* 2.1 *is satisfied. Then*

$$\frac{\mathrm{d}\widetilde{P}_\Gamma}{\mathrm{d}P} = \frac{1-g^1}{E[1-g^1]}, \tag{2.5}$$

where g^1 *is a solution to the abstract pure hedging problem* (2.3).

In view of Lemma 2.4, it suffices to study the variance-optimal signed Γ-martingale measure. We end this subsection by translating the correspondence (2.4) between the abstract random variables into a correspondence between strategies. To proceed, we first need as a property for the gains process that

$$G(\theta) \text{ is linear in } \theta, \tag{2.6}$$

without giving a precise definition of $G(\theta)$. If we want to apply the abstract result in Theorem 2.2 to the case $\Gamma = G_T(\Theta)$, Assumption 2.1 becomes a corresponding (implicit) assumption on the set Θ of strategies.

Assumption 2.5. 1) Θ is a linear space, and $G_T(\Theta)$ is closed in L^2.
2) $G_T(\Theta)$ does not contain the constant payoff 1.

Note that due to (2.6), Assumption 2.5, 1) implies that $G_T(\Theta)$ is a closed linear subspace of L^2.

Corollary 2.6. *Suppose that Assumption* 2.5 *is satisfied. Then problem* (1.1) *has a solution* $\widehat{\theta}^{\mathrm{mv}} \in \Theta$ *which can be written explicitly as*

$$\widehat{\theta}_t^{\mathrm{mv}} = \frac{1}{2\xi} \frac{1}{E[1 - G_T(\widehat{\theta}^1)]} \, \widehat{\theta}_t^1, \quad t \in \mathbb{T}, \tag{2.7}$$

where $\widehat{\theta}^1$ *is a solution to the pure hedging problem* (1.2)*. In consequence, the gains process* $G(\widehat{\theta}^{\mathrm{mv}})$ *satisfies*

$$G_t(\widehat{\theta}^{\mathrm{mv}}) = \frac{1}{2\xi} \frac{1}{E[1 - G_T(\widehat{\theta}^1)]} \, G_t(\widehat{\theta}^1), \quad t \in \mathbb{T}. \tag{2.8}$$

Proof. Because $G_T(\Theta)$ is closed in L^2, we can write

$$L^2 = G_T(\Theta) \oplus G_T(\Theta)^\perp.$$

This yields $1 = g^1 + \pi(1)$, where g^1 and $\pi(1)$ are the orthogonal projections of 1 onto $G_T(\Theta)$ and $G_T(\Theta)^\perp$, respectively. Because $g^1 \in G_T(\Theta)$, there exists a strategy $\widehat{\theta}^1 \in \Theta$ such that $G_T(\widehat{\theta}^1) = g^1$, and thus $\widehat{\theta}^1$ is a solution to the pure hedging problem (1.2). Inserting the above identity into (2.4) yields

$$g^{\mathrm{mv}} = \frac{1}{2\xi} \frac{1}{E[1 - G_T(\widehat{\theta}^1)]} \, G_T(\widehat{\theta}^1).$$

Defining $\widehat{\theta}^{\mathrm{mv}}$ as in (2.7), using the linearity of the gains process from (2.6) and invoking the last display, we have

$$G_T(\widehat{\theta}^{\mathrm{mv}}) = \frac{1}{2\xi} \frac{1}{E[1 - G_T(\widehat{\theta}^1)]} \, G_T(\widehat{\theta}^1) = g^{\mathrm{mv}}.$$

Because $\widehat{\theta}^{\mathrm{mv}}$ is in Θ due to the linearity of Θ from Assumption 2.5, 1), this shows that $\widehat{\theta}^{\mathrm{mv}}$ is a solution to the MVPS problem (1.1). Finally, the identity (2.8) is a direct consequence of (2.7) and the linearity of the gains assumed in (2.6). □

3 Results for the pure hedging problem

As mentioned earlier, there is a lot of literature on the study of the MVH problem (1.3). We refer to Schweizer [64] for a general overview. From a methodological perspective, we can categorise these works as follows:

(a) Viewing problem (1.3) as an L^2-projection problem, one obtains an optimality criterion of projection-type for (1.3), which leads to the usual *martingale techniques*. For works along this line, see e.g. Schweizer [58, 59, 61, 62], Rheinländer and Schweizer [56] and Černý and Kallsen [17].

(b) Viewing problem (1.3) as a stochastic control problem, one uses tools from standard stochastic control theory to tackle problem (1.3). For works using general tools like dynamic programming, see e.g. Bertsimas et al. [9], Gugushvili [32], Černý [15], Laurent and Pham [43] and Jeanblanc et al. [38]. For works using more specialised results from LQSC and backward stochastic differential equations (BSDEs), see e.g. Kohlmann and Tang [42], Lim and Zhou [48] and Lim [47].

Although it is difficult to draw a clear boundary between (a) and (b), we hope this crude classification sheds some light on the ideas used in the existing literature to tackle the MVH problem.

Let us now discuss methods from (a) in some detail because they give a general structure of the solution to the pure hedging problem (1.2) more efficiently. Recall from Section 1 the basic setup: we have a time horizon $T \in \mathbb{N}$, a time index set $\mathbb{T} \subseteq [0, T]$ and a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. On that space, there is an $\mathbb{R}^d$-valued process S whose components model the discounted prices of d risky assets in a financial market. In Section 1, we only verbally introduced trading strategies and the associated gains processes. We now frame these in mathematical terms. As a first assumption, the price process S is required to be adapted to the filtration $\mathbb{F}$. A *trading strategy* is a pair (v_0, θ), where $v_0 \in \mathbb{R}$ is the initial capital and θ is an $\mathbb{R}^d$-valued predictable process. We also want our strategies (v_0, θ) to be *self-financing* so that the *wealth process* of (v_0, θ) is given by

$$V_t(v_0, \theta) = v_0 + \int_0^t \theta_s \, \mathrm{d}S_s =: v_0 + G_t(\theta), \qquad t \in \mathbb{T}. \tag{3.1}$$

The process $G(\theta)$ is called the *gains process* of θ. Because (3.1) involves a stochastic integral with respect to S, the presentation is different for discrete-time and continuous-time processes. We only discuss the latter case here and refer to Section I.2 for the much simpler case of discrete time. For continuous time $\mathbb{T} = [0, T]$, we impose that S is a semimartingale with respect to the filtration $\mathbb{F}$ and that θ is S-integrable. This ensures that the (real-valued) stochastic integral $\int \theta \,\mathrm{d}S$ is well defined. For a precise definition, we refer to Jacod and Shiryaev [37, Sections I.4 and III.6]. Thanks to (3.1), there is for fixed v_0 a correspondence between a trading strategy (v_0, θ) and a process θ. Since v_0 is always fixed from now on, we use the term "trading strategy" to refer to an $\mathbb{R}^d$-valued, $\mathbb{F}$-predictable, S-integrable process θ. We denote the set of all these θ by $L(S)$. To ease notation, we **assume** $d = 1$ **throughout this section**.

3.1 General structure

Černý and Kallsen [17] provide a complete description of the solution to the pure hedging problem. To briefly present their results, we need some terminology. By $\mathcal{S}^2 = \mathcal{S}^2(P)$, we mean the space of semimartingales admitting a decomposition $X = X_0 + M^X + A^X$ with $M^X \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ and A^X of square-integrable variation (and therefore predictable, so that X is special).

Assumption 3.1. 1) The price process S is in $\mathcal{S}^2_{\mathrm{loc}}$.

2) There is some equivalent σ-martingale measure with square-integrable density, i.e., some probability measure $Q \approx P$ with $\frac{\mathrm{d}Q}{\mathrm{d}P} \in L^2(P)$ and such that S is a Q-σ-martingale.

Next we specify the set of trading strategies. A process $\theta \in L(S)$ is said to be *admissible* if there exists a sequence $(\theta^n)_{n\in\mathbb{N}}$ of simple (i.e., piecewise constant on finitely many stochastic intervals) strategies such that

$$\begin{aligned} &G_t(\theta^n) \longrightarrow G_t(\theta) \text{ in probability for any } t \in [0, T], \text{ and} \\ &G_T(\theta^n) \longrightarrow G_T(\theta) \text{ in } L^2. \end{aligned}$$

(Note that this admissibility is different from the one used by Delbaen and Schachermayer [24].) We consider

$$\Theta_{\mathrm{CK}} = \{\theta \in L(S) : \theta \text{ is admissible}\}. \tag{3.2}$$

The pure hedging problem in this setup is to

$$\text{minimise } E\big[\big(1 - G_T(\theta)\big)^2\big] \text{ over } \theta \in \Theta_{\mathrm{CK}}. \tag{3.3}$$

Lemma 3.2. *Suppose Assumption* 3.1 *is satisfied. Then* $G_T(\Theta_{\mathrm{CK}})$ *is equal to the* L^2*-closure of* $G_T(\Theta_{\mathrm{S}})$*, where*

$$\Theta_{\mathrm{S}} := \{\theta \in L(S) : G(\theta) \in \mathcal{S}^2\}. \tag{3.4}$$

Proof. See Černý and Kallsen [17, Corollary 2.9]. □

Lemma 3.2 implies that $G_T(\Theta_{\mathrm{CK}})$ is closed in L^2. Because $G_T(\Theta_{\mathrm{CK}})$ is already linear by the definition of admissible strategies, we obtain that Assumption 2.1 holds with $\Gamma = G_T(\Theta_{\mathrm{CK}})$. In view of Lemma 2.4 with $\Gamma = G_T(\Theta_{\mathrm{CK}})$, it is therefore enough to construct the variance-optimal signed $G_T(\Theta_{\mathrm{CK}})$-martingale measure. To this end, a central object is the so-called *opportunity process* which basically records a family of optimal values for certain conditional problems, as follows. **In the rest of this subsection,** Assumption 3.1 is imposed throughout.

For a stopping time τ, we define

$$\Theta_{\mathrm{CK}}(\tau) := \{\theta \in \Theta_{\mathrm{CK}} : \theta = 0 \text{ on } [\![0,\tau]\!]\}$$

and call a strategy $\lambda^{(\tau)} \in \Theta_{\mathrm{CK}}$ τ*-efficient* if $\lambda^{(\tau)}$ minimises $\theta \mapsto E[(1-G_T(\theta))^2]$ over $\theta \in \Theta_{\mathrm{CK}}(\tau)$. For any stopping time τ, such a τ-efficient strategy $\lambda^{(\tau)}$ exists as shown in Černý and Kallsen [17, Lemma 3.1]. For $t \in [0,T]$ and a t-efficient strategy $\lambda^{(t)}$, we then define

$$q_t = E\big[\big(1-G_T(\lambda^{(t)})\big)^2\big|\mathcal{F}_t\big], \quad t \in [0,T], \tag{3.5}$$

and call $q = (q_t)_{t\in[0,T]}$ the *opportunity process*. Lemma 3.2 and Corollary 3.4 in [17] show that we can (and do) choose an RCLL version of q and that q is a semimartingale (even a submartingale). It is shown in [17, Lemma 3.7] that there exists a process $\widetilde{a} \in L(S)$ such that

$$1-G(\lambda^{(\tau)}) = \mathcal{E}\big(G(-\widetilde{a}\mathbb{1}_{]\!]\tau,T]\!]})\big) = 1 - G\Big(\widetilde{a}\mathbb{1}_{]\!]\tau,T]\!]}\mathcal{E}\big(G(-\widetilde{a}\mathbb{1}_{]\!]\tau,T]\!]})\big)_-\Big)$$

for any stopping time τ. We call this (possibly non-unique) process $\widetilde{a}$ an *adjustment process.* Then we can define a signed measure Q^* by

$$\frac{\mathrm{d}Q^*}{\mathrm{d}P} := \frac{\mathcal{E}(G(-\widetilde{a}))_T}{E[\mathcal{E}(G(-\widetilde{a}))_T]}. \tag{3.6}$$

Theorem 3.3. *Suppose Assumption* 3.1 *is satisfied. Then the signed measure* Q^* *is a well-defined signed* $G_T(\Theta_{\mathrm{CK}})$*-martingale measure and variance-optimal for*

$G_T(\Theta_{\mathrm{CK}})$ *in the sense of Definition* 2.3.

Proof. See Černý and Kallsen [17, Proposition 3.13]. □

To convert this result into a solution for the pure hedging problem (3.3), we define a predictable process $\overline{a}$ as the solution of

$$a_t = -\widetilde{a}_t\big(G_{t-}(a) - 1\big), \quad t \in [0,T]. \tag{3.7}$$

Theorem 3.4. *Suppose Assumption* 3.1 *is satisfied. Then the equation* (3.7) *has a unique solution* $\overline{a}$ *in* $L(S)$*. Moreover, the process* $\overline{a}$ *is in* Θ_{CK} *and is an optimal strategy for the pure hedging problem* (3.3)*.*

Proof. See Černý and Kallsen [17, Lemma 4.9 and Theorem 4.10]. □

Finally, we convert the solution $\overline{a}$ for the pure hedging problem (3.3) to a solution $\widehat{\theta}^{\mathrm{mv}}$ for the MVPS problem (1.1) with $\Theta = \Theta_{\mathrm{CK}}$. Because Assumption 3.1 implies that Assumption 2.5 is satisfied with $\Theta = \Theta_{\mathrm{CK}}$, we use Corollary 2.6, (2.7) and (2.8) and the linearity of the gains process to obtain

$$\widehat{\theta}^{\mathrm{mv}}_t = \frac{1}{2\xi}\frac{1}{E[\mathcal{E}(G(-\widetilde{a}))_T]}\,\overline{a}_t, \quad G_t(\widehat{\theta}^{\mathrm{mv}}) = \frac{1}{2\xi}\frac{1}{E[\mathcal{E}(G(-\widetilde{a}))_T]}\,G_t(\overline{a}), \quad t \in [0,T],$$

and hence

$$\begin{aligned}\widehat{\theta}^{\mathrm{mv}}_t &= -\frac{1}{2\xi}\frac{1}{E[\mathcal{E}(G(-\widetilde{a}))_T]}\,\widetilde{a}_t\big(G_{t-}(\overline{a}) - 1\big)\\ &= -\widetilde{a}_t\left(G_{t-}(\widehat{\theta}^{\mathrm{mv}}) - \frac{1}{2\xi E[\mathcal{E}(G(-\widetilde{a}))_T]}\right), \quad t \in [0,T].\end{aligned} \tag{3.8}$$

Example 3.5. We discuss a special case where the adjustment process $\widetilde{a}$ can be determined explicitly. Suppose Assumption 3.1 is satisfied. Schweizer [60, Theorem 1] says that S satisfies the structure condition, meaning that

$$S = S_0 + M + A = S_0 + M + \int \lambda\,\mathrm{d}\langle M\rangle$$

with $M \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ and $\lambda \in L^2_{\mathrm{loc}}(M)$. If the entire mean–variance tradeoff (MVT) process $K := \int \lambda\,\mathrm{d}\langle M\rangle$ is deterministic, we can define a martingale measure $\widehat{P}$ via its density $\frac{\mathrm{d}\widehat{P}}{\mathrm{d}P} = \mathcal{E}(\int -\lambda\,\mathrm{d}M)_T$. This is called the *minimal signed martingale measure* for S by Schweizer [60]. Moreover, [60, Theorem 8] there shows that the variance-optimal local martingale measure for S coincides with $\widehat{P}$. It is not difficult to argue that a variance-optimal local martingale measure for

S is the same as a variance-optimal $G_T(\Theta_{\mathrm{CK}})$-martingale measure in this setup. Therefore $\widetilde{a}$ can be chosen as λ. From (3.8), we then obtain an optimal strategy for the MVPS problem explicitly given by

$$\widehat{\theta}_t^{\mathrm{mv}} = -\lambda_t \left(G_{t-}(\widehat{\theta}^{\mathrm{mv}}) - \frac{1}{2\xi E[\mathcal{E}(G(-\lambda))_T]} \right), \quad t \in [0,T]. \tag{3.9}$$

If in addition S is continuous, then we only need to assume that the final value K_T is deterministic to obtain $\widetilde{a} = \lambda$ and hence the identity (3.9).

3.2 Results in discrete time

In finite discrete time, we have more explicit results about the solutions for the pure hedging problem (3.3) and the MVH problem (1.3). Schweizer [61] first fully worked out the explicit structure of the solutions for both problems with $\Theta = \Theta_{\mathrm{S}}$ given in (3.4). In discrete time, the requirement in (3.4) that $G(\theta) \in \mathcal{S}^2$ translates into the condition that $G_t(\theta) \in L^2$ for $t = 1, \dots, T$. Later, various papers including Gugushvili [32], Černý [15] and Melnikov and Nechaev [50] extended the results to more general spaces of strategies. But the explicit structure of the solutions for both problems remains the same as in Schweizer [61]. We first follow the recent presentation given in Černý and Kallsen [17] and then discuss the connections to other papers.

The opportunity process q and adjustment process $\widetilde{a}$ are given more explicitly. Indeed, we have from [17, Example 3.32] that

$$\begin{aligned} q_{t-1} &= E[q_t|\mathcal{F}_{t-1}] - \frac{(E[q_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[q_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}, \qquad q_T = 1, \\ \widetilde{a}_t &= \frac{E[q_t \triangle S_t|\mathcal{F}_{t-1}])}{E[q_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}, \qquad t = 1, \dots, T. \end{aligned} \tag{3.10}$$

Thanks to the discrete-time structure, the stochastic exponential reads

$$\mathcal{E}\big(G(-\widetilde{a})\big)_t = \prod_{s=1}^{t} (1 - \widetilde{a}_s \triangle S_s), \quad t = 1, \dots, T.$$

From (3.6), we also have an explicit expression for the variance-optimal signed $G_T(\Theta_{\mathrm{CK}})$-martingale measure.

The signed measure Q^* from (3.6) in finite discrete time has the density

$$\frac{\mathrm{d}Q^*}{\mathrm{d}P} = \frac{\mathcal{E}(G(-\widetilde{a}))_T}{E[\mathcal{E}(G(-\widetilde{a}))_T]} = \frac{\prod_{t=1}^T (1 - \widetilde{a}_t \triangle S_t)}{E[\prod_{t=1}^T (1 - \widetilde{a}_t \triangle S_t)]}.$$

We translate (3.7) into a solution $\overline{a}$ for the pure hedging problem (3.3) in discrete time via

$$\overline{a}_t = -\widetilde{a}_t \big(G_{t-1}(\overline{a}) - 1 \big), \quad t = 1, \dots, T,$$

where the adjustment process $\widetilde{a}$ is given explicitly by (3.10). Similarly, we translate (3.8) into a solution $\widehat{\theta}^{\mathrm{mv}}$ for the MVPS problem (1.1) with $\Theta = \Theta_{\mathrm{CK}}$ in finite discrete time via

$$\widehat{\theta}_t^{\mathrm{mv}} = -\widetilde{a}_t \left(G_{t-1}(\widehat{\theta}^{\mathrm{mv}}) - \frac{1}{2\xi E[\mathcal{E}(G(-\widetilde{a}))_T]} \right), \quad t = 1, \dots, T.$$

We now discuss the connections to earlier literature. First, Schweizer [61] solved the pure hedging problem (3.3) with $\Theta = \Theta_{\mathrm{S}}$ under the assumption that the mean–variance tradeoff process is uniformly bounded. Melnikov and Nechaev [50] considered the space

$$\Theta_{\mathrm{MN}} = \{\theta = (\theta_t)_{t=1,\dots,T} : \theta \text{ real-valued}, \mathbb{F}\text{-predictable and } G_T(\theta) \in L^2\} \quad (3.11)$$

and solved the problem (3.3) with $\Theta = \Theta_{\mathrm{MN}}$. Note that we immediately have $\Theta_{\mathrm{S}} \subseteq \Theta_{\mathrm{MN}}$. To compare Θ_{MN} and Θ_{CK} in finite discrete time, we need the result below.

Lemma 3.6. *Suppose that Assumption* 3.1 *is satisfied. Then in finite discrete time, we have*

$$\Theta_{\mathrm{CK}} = \Theta_{\mathrm{MN}}.$$

Proof. See Černý and Kallsen [19, Proposition 8.4]. □

3.3 A backward stochastic differential equation (BSDE) for the opportunity process

Although (3.7) gives a general structure for the solution of the pure hedging problem (3.3), the optimal strategy requires to calculate either the opportunity process q or the adjustment process $\widetilde{a}$, both of which are notoriously difficult to find explicitly. In this subsection, we present a BSDE approach proposed by

Jeanblanc et al. [38], which gives a different characterisation of the opportunity process q given in (3.5).

We denote by $\mathbb{P}^2_{\mathrm{e},\sigma}(S)$ the set of all probability measures $Q \approx P$ on $\mathcal{F}_T$ such that S is a Q-σ-martingale and $\frac{\mathrm{d}Q}{\mathrm{d}P} \in L^2(P)$, and we consider the spaces

$$\begin{aligned} \Theta_{\mathrm{S}} &:= \{\theta \in L(S) : G(\theta) \in \mathcal{S}^2(P)\}, \\ \Theta_{\mathrm{S}}(\tau) &:= \{\theta \in \Theta_{\mathrm{S}} : \theta = 0 \text{ on } [\![0,\tau]\!]\} \end{aligned}$$

of strategies, where τ is an $\mathbb{F}$-stopping time. A basic relation between Θ_{S} and Θ_{CK} in (3.2) is that $G_T(\Theta_{\mathrm{CK}})$ is the L^2-closure of $G_T(\Theta_{\mathrm{S}})$, i.e.

$$\begin{aligned} \overline{G_T(\Theta_{\mathrm{S}})} &= G_T(\Theta_{\mathrm{CK}}), \\ \overline{G_T\big(\Theta_{\mathrm{S}}(\tau)\big)} &= G_T\big(\Theta_{\mathrm{CK}}(\tau)\big), \end{aligned}$$

for any stopping time τ; see for instance Černý and Kallsen [17, Corollary 2.9] for a proof. In view of these identities, we can equivalently rewrite the definition (3.5) of q as

$$q_t = \operatorname*{ess\,inf}_{\theta\in\Theta_{\mathrm{CK}}(t)} E\big[\big(1-G_T(\theta)\big)^2\big|\mathcal{F}_t\big] = \operatorname*{ess\,inf}_{\theta\in\Theta_{\mathrm{S}}(t)} E\big[\big(1-G_T(\theta)\big)^2\big|\mathcal{F}_t\big]. \tag{3.12}$$

A natural way to compute q is then to solve the family of conditional problems given by the right-most expression in (3.12).

To proceed, we assume that Assumption 3.1 is satisfied, i.e., the process S is in $\mathcal{S}^2_{\mathrm{loc}}$ and $\mathbb{P}^2_{\mathrm{e},\sigma}(S) \neq \emptyset$. The latter condition is one way of imposing absence of arbitrage for the market. Moreover, this implies that S satisfies the so-called *structure condition*, namely that S has the form

$$S = S_0 + M^S + A^S = S_0 + M + \int \lambda \,\mathrm{d}\langle M\rangle,$$

with $M = M^S \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ and λ being predictable and in $L^2_{\mathrm{loc}}(M)$. Before stating the BSDE, we recall some results and notations from the general theory of processes. For any locally bounded process X, we denote by ${}^{\mathrm{p}}X$ its predictable projection. If X is any process of locally integrable variation, then its compensator X^{p} exists. Let $Y = Y_0 + N^Y + B^Y$ be a bounded (hence special) semimartingale. The property (2.9) in Jeanblanc et al. [38] shows that if the process $[N^Y,[S]]$ is of locally integrable variation, then its compensator $[N^Y,[S]]^{\mathrm{p}}$ is absolutely continuous with respect to $\langle M \rangle$ and has a predictable density

$$g_t(Y) := \frac{\mathrm{d}[N^Y, [S]]^{\mathrm{p}}_t}{\mathrm{d}\langle M \rangle_t}, \quad t \in [0,T].$$

Finally, we introduce the notation $\mathcal{N}(Y)$ given by

$$\mathcal{N}_t(Y) := {}^{\mathrm{p}}Y_t(1 + \lambda_t^2 \triangle \langle M \rangle_t) + g_t(Y), \quad t \in [0,T].$$

Now we consider the backward equation

$$\mathrm{d}Y_t = \frac{\psi_t + \lambda_t({}^{\mathrm{p}}Y_t)}{\mathcal{N}_t(Y)} \,\mathrm{d}\langle M \rangle_t + \psi \,\mathrm{d}M_t + \mathrm{d}L_t, \quad Y_T = 1. \tag{3.13}$$

A solution to (3.13) is a triplet (Y, ψ, L), where L is a local P-martingale strongly P-orthogonal to M, ψ is in $L^1_{\mathrm{loc}}(M)$ and Y is a P-special semimartingale with $[N^Y, [S]]$ of locally integrable variation. For fixed $t \in [0,T]$, define the stochastic exponential starting from time t of a semimartingale X by

$${}^t\mathcal{E}(X)_u = 1 + \int_t^u {}^t\mathcal{E}(X)_{r-} \,\mathrm{d}X_r = \mathcal{E}(X - X^t)_u, \quad u \in [t,T].$$

Theorem 3.7. *Suppose that the process S is in $\mathcal{S}^2_{\mathrm{loc}}$ and $\mathbb{P}^2_{\mathrm{e},\sigma}(S) \neq \emptyset$. Then the following two statements are equivalent:*

1) *For every $t \in [0,T]$, there exists an optimal strategy $\theta^{*,t} \in \Theta_{\mathrm{S}}(t)$ for the second conditional problem in* (3.12).

2) *There exists a solution (Y, ψ, L) to the BSDE* (3.13) *having $L \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ strongly P-orthogonal to M, $\psi \in L^2_{\mathrm{loc}}(M)$, Y bounded and strictly positive and such that for every $t \in [0,T]$, the process ${}^t\mathcal{E}(G(-\frac{\psi + \lambda({}^{\mathrm{p}}Y)}{\mathcal{N}(Y)}))$ is in $\mathcal{S}^2(P)$.*

If either 1) *or* 2) *holds, then for each $t \in [0,T]$ the optimal strategy $\theta^{*,t}$ for the second conditional problem in* (3.12) *is given by*

$$\theta^{*,t}_u = -\frac{\psi_u + \lambda_u({}^{\mathrm{p}}Y_u)}{\mathcal{N}_u(Y)} \, {}^t\mathcal{E}\left(G\Big(-\frac{\psi + \lambda({}^{\mathrm{p}}Y)}{\mathcal{N}(Y)}\Big)\right)_u, \quad u \in [t,T], \tag{3.14}$$

and the opportunity process q is the unique bounded strictly positive solution Y of (3.13).

If in addition there is some $Q \in \mathbb{P}^2_{\mathrm{e},\sigma}(S)$ satisfying the reverse Hölder inequality $R_2(P)$, then q is the unique solution to (3.13) *in the class of processes satisfying $c \leq Y \leq C$ for some $c, C > 0$.*

4 A brief introduction to our approach

In this thesis, we develop two types of dynamic programming techniques for the MVPS problem.

4.1 Chapters I and II

The first method is inspired by the market cloning technique originally proposed by Ankirchner and Dermoune [5] and extended by Fischer and Livieri [29]. For convenience, by a financial market, we mean here a tuple $(\Omega, \mathcal{F}, \mathbb{F}, P, \Theta, S)$, where $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is a filtered probability space, Θ is a set of strategies and S is the discounted price process. We recall from (1.1) that the MVPS problem in this market is to

$$\text{maximise } E[G_T(\theta)] - \xi \text{Var}[G_T(\theta)] \text{ over all } \theta \in \Theta. \tag{4.1}$$

Due to the variance term in (4.1), one cannot easily state a dynamic programming principle for this problem. The idea to get around this issue relies on a law of large numbers which formally says that the empirical mean $\frac{1}{L}\sum_{\ell=1}^{L} X_\ell$ of i.i.d. random variables $(X_\ell)_{\ell\in\mathbb{N}}$ converges to the expectation of the random variable X_1 as $L \to \infty$. Inspired by this observation, one then blows up the problem by first constructing an extended market $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbb{G}^{(L)}, \mathbf{P}^{(L)}, \mathbf{\Theta}^{(L)}, \mathbf{S}^{(L)})$ which supports L i.i.d. copies of the original financial market $(\Omega, \mathcal{F}, P, \Theta, S)$, then defining a vector gains process $\mathbf{G}(\boldsymbol{\vartheta})$ whose ℓ-th coordinate records the profits and losses of investment in $\mathbf{S}^{\ell,(L)}$ via $\boldsymbol{\vartheta}^{\ell,(L)}$ according to a vector $\boldsymbol{\vartheta}^{(L)} = (\boldsymbol{\vartheta}^{\ell,(L)})_{\ell=1,\dots,L}$ of strategies, and finally replacing the expectation and variance in (4.1) by empirical means and empirical variances. This leads us to consider the extended problem to

$$\text{maximise } \mathbf{E}^{(L)}\big[\text{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big) - \xi \text{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big)\big] \text{ over all } \boldsymbol{\vartheta}^{(L)} \in \mathbf{\Theta}^{(L)}, \tag{4.2}$$

where em$(\cdot)$ and evar$(\cdot)$ denote the empirical mean and variance, respectively. The key observation is now that problem (4.2) is a standard stochastic control problem with state variable $\mathbf{G}(\boldsymbol{\vartheta}^{(L)})$ and control $\boldsymbol{\vartheta}^{(L)}$. So we can state and prove a dynamic programming principle for (4.2), which then gives a systematic way of computing an optimal strategy $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for the extended problem.

Having obtained an optimal $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for each auxiliary problem (4.2), we then pass to the limit as $L \to \infty$ in order to solve the original problem (4.1). Let us introduce handy notations J^{mv} and $J^{(L)}$ by

$$J^{\mathrm{mv}}(\theta) := G_T(\theta) - \xi\big(G_T(\theta) - E[G_T(\theta)]\big)^2, \quad \theta \in \Theta, \tag{4.3}$$

$$J^{(L)}(\boldsymbol{\vartheta}^{(L)}) := \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big), \quad \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}, \tag{4.4}$$

respectively, so that $E[J^{\mathrm{mv}}(\theta)] = E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)]$. Using (4.3) and (4.4), we can equivalently write problems (4.1) and (4.2) as to

$$\text{maximise } E[J^{\mathrm{mv}}(\theta)] \text{ over all } \theta \in \Theta,$$
$$\text{maximise } \mathbf{E}^{(L)}[J^{(L)}(\boldsymbol{\vartheta}^{(L)})] \text{ over all } \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)},$$

respectively. For any $\theta \in \Theta$, we now construct a vector strategy $\theta^{\otimes L} \in \boldsymbol{\Theta}^{(L)}$ consisting of L i.i.d. copies of θ, which in turn yields that the vector final gain $\mathbf{G}_T(\theta^{\otimes L})$ consists of L i.i.d. copies of $G_T(\theta)$. Because $J^{(L)}$ in (4.4) involves only the empirical mean and variance of the i.i.d. random variables $(\mathbf{G}_T^{\ell,(L)}(\theta^{\otimes L}))_{\ell=1,\ldots,L}$, we can use a law of large numbers argument to show that

$$\mathbf{E}^{(L)}[J^{(L)}(\theta^{\otimes L})] \longrightarrow E[J^{\mathrm{mv}}(\theta)] \quad \text{as } L \to \infty.$$

In view of $\theta^{\otimes L} \in \boldsymbol{\Theta}^{(L)}$ and the optimality of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for the auxiliary problem, we always have

$$\mathbf{E}^{(L)}[J^{(L)}(\theta^{\otimes L})] \le \mathbf{E}^{(L)}[J^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})].$$

Combining the above two inequalities yields

$$E[J^{\mathrm{mv}}(\theta)] \le \limsup_{L\to\infty} \mathbf{E}^{(L)}[J^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] \quad \text{for all } \theta \in \Theta.$$

Suppose that θ^{mv} is a solution to the original MVPS problem. We then have

$$E[J^{\mathrm{mv}}(\theta^{\mathrm{mv}})] \le \limsup_{L\to\infty} \mathbf{E}^{(L)}[J^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})]. \tag{4.5}$$

This can be used to construct a candidate $\widehat{\theta}$ for the solution to the original problem in the following manner. For the sake of argument, suppose that an explicit formula for $\widehat{\boldsymbol{\vartheta}}^{(L)}$ can be obtained. By passing to the limit formally or likewise, we can then construct a candidate $\widehat{\theta} \in \Theta$ for the original problem based on that formula for $\widehat{\boldsymbol{\vartheta}}^{(L)}$. If we can show that $E[J^{\mathrm{mv}}(\widehat{\theta})]$ is equal to the right-hand side of (4.5), we obtain $E[J^{\mathrm{mv}}(\theta^{\mathrm{mv}})] \le E[J^{\mathrm{mv}}(\widehat{\theta})]$, which yields the optimality of $\widehat{\theta}$ for the original problem (4.1).

The market cloning technique as explained in the above paragraphs does not seem to rely on the specific dynamics of the price process S. From a methodological perspective, a unified mathematical framework is thus worth to be worked out in detail and constitutes our first contribution to the existing literature.

In Chapter I, we construct the extended market and formulate the auxiliary problem in finite discrete time. The DPP embeds the auxiliary problem for fixed L into a temporal sequence of (conditional) stochastic control problems and asserts that the (optimal) value of the conditional problem at time $t-1$ can be computed in terms of the solution to the problem at time t. This gives us a systematic approach for constructing an optimal strategy for the auxiliary problem, which consists of solving the problem at time t by finding an optimiser, then plugging that into the value of the time-t problem, and iterating this procedure backward in time starting from $t = T$. Finally, we construct a candidate for the MVPS problem by exploiting the just obtained strategy and passing to the limit.

In Chapter II, we develop this approach in continuous time with the same basic idea as in Chapter I. But because we have a continuum of time steps, we attack the auxiliary problem by guessing an affine–quadratic structure for the value process and then constructing candidates for both the value process and an optimal strategy based on a martingale optimality principle. In both chapters, we manage to solve the auxiliary problem under the assumption that the MVT process is deterministic. Although in these cases, an optimal strategy for the MVPS problem can be alternatively obtained as in (3.9), the market cloning technique exploits a completely different angle and has more the flavour of McKean–Vlasov control problems which typically consider i.i.d. innovations and Brownian-driven SDE models. Therefore, our results nicely extend this technique and give our second contribution to the existing literature.

4.2 Chapter III

The second type of dynamic programming techniques for the MVPS problem is inspired by recent progress in McKean–Vlasov control theory. Because the MVPS problem belongs to a category of such problems, it is natural to use results from that theory to tackle it. Andersson and Djehiche [4] propose a solution technique based on a stochastic maximum principle for McKean–Vlasov control problems. Pham and Wei [52] develop a dynamic programming principle (DPP) for processes driven by i.i.d. innovations and apply that DPP to solve the MVPS problem. The market cloning technique proposed by Ankirchner and Dermoune [5] and extended by Fischer and Livieri [29] has some flavour of this type. But at present, it seems that the current literature mostly considers processes driven by i.i.d. innovations or Brownian-driven SDE models. This is an aspect worth some improvement.

We first describe the problem. Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t=0,1,\dots,T}$. Consider a family $(X_t^\theta)_{t=0,1,\dots,T}$ of stochastic processes controlled by an $\mathbb{F}$-predictable process $\theta = (\theta_t)_{t=1,\dots,T}$, meaning that X_t^θ depends on θ only via $\theta_1, \dots, \theta_t$. For a generic random quantity Y, we denote by P_Y the distribution of Y under P. For the sake of argument, we suppose that all random variables are real-valued and denote by $\mathcal{P}(\mathbb{R})$ the set of all probability distributions on $\mathbb{R}$. For a set Θ of controls θ, we are interested in a criterion j of McKean–Vlasov type given by

$$j(\theta) = E\Bigg[\sum_{t=0}^{T-1} f(t, X_t^\theta, \theta_{t+1}, P_{X_t^\theta, \theta_{t+1}}) + g(X_T^\theta, P_{X_T^\theta})\Bigg],$$

where f and g are real-valued (measurable) functions with appropriate domains. The problem is to

$$\text{maximise } j(\theta) \text{ over all } \theta \in \Theta. \tag{4.6}$$

Although the criterion j looks quite formidable, we can embed this problem into a sequence of *deterministic* tail problems where both the value $v(t, \theta)$ and the optimisation of $v(t, \theta)$ involve only an expectation, not a conditional expectation, of variables from $t + 1$ onward. This then yields a DPP which asserts that the (optimal) value at $t-1$ can be obtained in terms of the solution to the tail problem at time t. Since the tail problem for $t = 0$ corresponds to the original problem (4.6), we thus obtain a systematic approach via the above backward recursion to compute both the value and an optimal strategy for (4.6). This idea already appears in Pham and Wei [52, Lemma 3.1], but the authors work there with closed-loop controls in an i.i.d. framework. We first extend their results to open-loop controlled processes without specifying any dynamics for the underlying process. Because the MVPS problem has a linear–quadratic (LQ) structure, we study and obtain some structural results for a general class of single-period LQ problems. We finally piece these one-step results together to solve the MVPS problem (4.1) in full generality in finite discrete time.

Chapter I

Mean field approach for MVPS – discrete time

1 Introduction

Mean–variance portfolio selection is a classic problem in finance. In financial terms, the goal is to maximise the expectation and minimise the variance of the final wealth $G_T(\theta) = \int_0^T \theta_s \, \mathrm{d}S_s$ of a self-financing strategy θ investing in the underlying assets S. Mathematically, this is formulated as

$$\text{maximise } E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)] \text{ over } \theta \in \Theta,$$

for a suitable set Θ of stochastic processes and a risk aversion parameter $\xi > 0$. Compared to its single-period version studied by Markowitz [49], the multi-period mean–variance portfolio selection (MVPS) problem seems to have gained less popularity over the decades. One reason is the fact that the variance term in the mean–variance criterion makes the problem not amenable to a (Bellman-type) dynamic programming principle or other standard tools from stochastic control theory. The contribution of this chapter is to extend a novel solution technique proposed by Ankirchner and Dermoune [5] and apply that to solve the MVPS problem in finite discrete time.

While the toolbox for tackling the multi-period MVPS problem in generality seems to be almost empty, there still has been some significant progress on this problem. The connection to other literature is discussed in detail in Section 6. Roughly speaking, there are three approaches for the multi-period MVPS problem:

(a) One considers a time-consistent or equilibrium solution for the MVPS problem. This approach goes away from the original formulation and studies a different type of optimality for which a solution enjoys a dynamic programming principle.

(b) One exploits the connection of the MVPS problem to a class of auxiliary pure hedging problems which turn out to be standard linear–quadratic stochastic control problems. Then one solves the latter and translates the resulting solution into a solution for the original MVPS problem.

(c) One views the MVPS problem as a special case of a McKean–Vlasov control problem and uses tools from there to tackle the MVPS it.

The market cloning technique of Ankirchner and Dermoune [5] seems to lie at the intersection of (b) and (c) as explained in Section 0.2 in Chapter 0. The authors also consider a class of auxiliary standard control problems. However, these look more analogous to the original MVPS problem than the class of auxiliary problems typically considered in (b). We extend the approach of [5] in two aspects. First, from a methodological perspective, there is a general mathematical structure behind their technique which is worth developing systematically and rigorously. Second, in this more general framework, we can apply their technique to study the MVPS problem with an underlying price process S having non-independent increments. Of course, some results for the MVPS problem with a general underlying process S can be obtained by the approaches in (b). But since the market cloning technique has more the flavour of (c) where one typically studies processes driven by i.i.d. innovations, our results nicely complement the current literature.

This chapter is structured as follows. In Section 2, we first introduce the market and formulate the MVPS problem in precise terms. Then we introduce the market cloning technique from Ankirchner and Dermoune [5] by constructing an extended market which supports independent copies of the original market and formulating in that market an auxiliary problem, with more care, rigour and in unified form. The section ends with the classic martingale optimality principle (MOP) tailored for the auxiliary problem and some results for martingales under shrinkage of filtrations. The latter results allow us later to go beyond the case of i.i.d. innovations. We keep the development in this section as general as possible and impose only some mild assumptions.

In Section 3, we study the auxiliary problem in more detail. Using the abstract MOP in Section 2, we establish a dynamic programming result for the extended problem and reformulate that into a sequence of single-period (conditional) problems. Then we go away from the general presentation and provide the necessary details for a concrete setup where we perform the actual computation of an optimal strategy for the auxiliary problem backward in time in Section 4.

Section 4 proposes a recipe for the systematic computation of an optimal strategy for the auxiliary problem based on the dynamic programming principle from Section 3. Following that recipe, we first maximise a conditional problem at time T and look for a structure that propagates backward in time. But a desired simple affine–quadratic structure comes at the cost of the extra assumption that the mean–variance tradeoff process for the price process S is deterministic; this appears naturally along the computations. Assuming that condition, we present in the main result of this section (Theorem 4.12) an affine–quadratic structure for the entire value process and an optimal strategy for the auxiliary problem.

Finally, we explore in Section 5 the connection between the auxiliary and the original MVPS problems. The connections are two-fold. First, with the help of the optimal strategy from Theorem 4.12 for the auxiliary problem, we construct a candidate for an optimal strategy for the original MVPS problem and present a simple verification procedure. The verification step is new and can be easily extended to a more general class of problems whose criterion involves a nonlinear function of an expectation. Second, we show that the gains of the optimal strategy for the auxiliary problem converge to the gains of the optimal strategy for the original MVPS problem, as the number of copies in the extended market goes to infinity, with a precise rate of convergence. This improves the corresponding results in Ankirchner and Dermoune [5].

In Section 6, we discuss connections to the literature in detail.

2 Problem formulation and general preliminaries

2.1 Mean–variance portfolio selection (MVPS) in a financial market

Let us begin with some necessary preparation. We first introduce a financial market and describe investment in this market, both in mathematical terms. Then we formulate the MVPS problem in this financial market. Finally, we study the existence and uniqueness of a solution to the MVPS problem under some natural assumptions.

Let $T \in \mathbb{N}$ be a time horizon and fix a *time index set* $\mathbb{T} \subseteq [0,T]$. The two main examples are $\mathbb{T} = \{0,1,\dots,T\}$ and $\mathbb{T} = [0,T]$, which stand for the case of discrete and continuous time, respectively.

Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$. We always assume that $\mathcal{F}_0$ is P-trivial, meaning that $\mathcal{F}_0$ only contains events with probability 0 or 1. In continuous time, i.e. $\mathbb{T} = [0,T]$, we additionally assume that $\mathbb{F}$ satisfies the usual conditions of right-continuity and P-completeness.

Consider a financial market with $d+1$ assets with $d \in \mathbb{N}$. There are d risky assets and one riskless asset. For simplicity, all prices are discounted by the riskless asset and expressed by units of 1. In mathematical terms, the price of the riskless asset is modelled by $S^0 = (S^0_t)_{t\in\mathbb{T}} \equiv 1$, and the (discounted) prices of the risky assets are modelled by an $\mathbb{R}^d$-valued stochastic process $S = (S_t)_{t\in\mathbb{T}}$. As a first assumption, we impose that S is adapted to $\mathbb{F}$.

Next, we mathematically describe trading behaviours in this market. A *trading strategy* is a pair (v_0, θ), where $v_0 \in \mathbb{R}$ is the initial capital and θ is an $\mathbb{R}^d$-valued predictable process. Because we want our strategies to be self-financing, the *wealth process* of (v_0,θ) is given by

$$V_t(v_0,\theta) = V_t = v_0 + \int_0^t \theta_s \,\mathrm{d}S_s =: v_0 + G_t(\theta), \quad t \in \mathbb{T}. \tag{2.1}$$

The process $G(\theta)$ is called the *gains process* of θ. We also use $\theta \bullet S$ to denote the stochastic integral process of θ with respect to S. Hence, $\theta \bullet S$ and $\int \theta \,\mathrm{d}S$ are used interchangeably.

Note that the self-financing condition involves a stochastic integral with respect to S, which requires different assumptions for discrete-time and continuous-time processes. Therefore, we divide the discussion into two cases.

Trading in discrete time: $\mathbb{T} = \{0,1,\dots,T\}$. This requires no additional assumptions on either θ or S. We introduce the notation

$$\triangle X_t := X_t - X_{t-1}, \quad t \in \mathbb{N},$$

for increments of any discrete-time process $X = (X_t)_{t\in\mathbb{N}_0}$. Note that to make sense of

$$\int_0^t \theta_s \,\mathrm{d}S_s := \sum_{s=1}^t \theta_s^\top \triangle S_s, \quad t = 0,1,\dots,T,$$

we indeed do not need any further assumptions. Here we adopt the standard convention that any sum over an empty set is equal to 0, which yields in particular $\int_0^0 \theta_s \, \mathrm{d}S_s = 0$.

Trading in continuous time: $\mathbb{T} = [0, T]$. We make a particular choice of well-defined stochastic integration theories. That is, we impose that S is a semimartingale with respect to the filtration $\mathbb{F}$. Moreover, we require that θ is S-integrable. This ensures that the (real-valued) stochastic integral $\int \theta \, \mathrm{d}S$ is well defined. For the precise definition, we refer to Jacod and Shiryaev [37, Sections I.4 and III.6].

To summarise, we consider two classes of assumptions:

- In discrete time, S needs no further assumptions. A (self-financing) trading strategy (v_0, θ) satisfies that $v_0 \in \mathbb{R}$ and θ is an $\mathbb{R}^d$-valued, predictable process.

- In continuous time, we assume further that S is a semimartingale with respect to $\mathbb{F}$. A (self-financing) trading strategy (v_0, θ) satisfies that $v_0 \in \mathbb{R}$ and θ is an $\mathbb{R}^d$-valued, predictable, S-integrable process.

Remark 2.1. In discrete time, any adapted $\mathbb{R}^d$-valued process S is automatically a semimartingale, and any $\mathbb{R}^d$-valued process is automatically S-integrable.

Now we turn to formulating the MVPS problem. Let Θ be a set of processes such that for any $v_0 \in \mathbb{R}$ and any $\theta \in \Theta$, the pair (v_0, θ) is a (self-financing) trading strategy. Fix a risk tolerance parameter $\gamma > 0$ and an initial capital $v_0 \in \mathbb{R}$. The MVPS problem is to

$$\text{maximise } E[V_T(v_0, \theta)] - \frac{1}{2\gamma} \mathrm{Var}[V_T(v_0, \theta)] \text{ over all } \theta \in \Theta, \tag{2.2}$$

which by (2.1) is structurally equivalent to

$$\text{maximise } E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)] \text{ over all } \theta \in \Theta, \tag{2.3}$$

where ξ is a positive real number standing for a generic risk aversion parameter, and we can choose $\xi = \frac{1}{2\gamma}$ for consistency with (2.2) when necessary.

The remaining task is to make additional assumptions on S and Θ so that the MVPS problem (2.3) is mathematically well defined.

Assumption 2.2. Θ satisfies the following:

1) $G_T(\Theta) := \{G_T(\theta) : \theta \in \Theta\}$ is a closed subspace of $L^2 := L^2(P)$, i.e., $G_T(\Theta) \subseteq L^2$ is a linear space and $G_T(\Theta)$ is closed in L^2.

2) The constant (payoff) 1 is not in $\overline{G_T(\Theta)}$, the L^2-closure of $G_T(\Theta)$.

Remark 2.3. 1) Note that Assumption 2.2, 1) is not a specific choice of Θ. Instead, it is a generic condition that Θ should satisfy. We shall make concrete choices of Θ only in Section 3.

2) Assumption 2.2, 2) is a weak kind of no-arbitrage condition. If we also have 1), it is of course equivalent to $1 \notin G_T(\Theta)$.

We now show that under Assumption 2.2, the MVPS problem is indeed well posed as follows.

Theorem 2.4. *Under Assumption* 2.2*, the MVPS problem* (2.3) *has a maximiser* $\widehat{\theta} \in \Theta$*, and the corresponding* $G_T(\widehat{\theta})$ *is unique.*

Proof. Consider the functional $F : L^2 \to \mathbb{R}$ given by

$$F(g) := E[g] - \xi \mathrm{Var}[g]. \tag{2.4}$$

1) We first show that F is L^2-coercive on $G_T(\Theta)$, i.e., if $(g_n)_{n\in\mathbb{N}}$ is a sequence in $G_T(\Theta)$ with $E[g_n^2] \to \infty$, then $F(g_n) \to -\infty$. To this end, we denote by $\pi(1)$ the L^2-orthogonal projection of 1 onto $G_T(\Theta)^\perp$, the orthogonal complement in L^2 of $G_T(\Theta)$. Note that $\pi(1)$ exists and is nonzero. Indeed, the linearity of $G_T(\Theta)$ from Assumption 2.2, 1) gives existence and uniqueness of $\pi(1)$ by the projection result in Hilbert spaces, and Assumption 2.2, 2) implies that $\pi(1) \not\equiv 0$. In particular, we have $E[\pi(1)^2] = \inf_{g\in G_T(\Theta)} E[(1-g)^2] > 0$.

Let $g \in G_T(\Theta)$. If $E[g] \neq 0$, then $g/E[g] \in G_T(\Theta)$ by the linearity of $G_T(\Theta)$, and hence factoring out $(E[g])^2$ and using $E[\pi(1)^2] = \inf_{g\in G_T(\Theta)} E[(1-g)^2]$ yields

$$\mathrm{Var}[g] = E\big[(g - E[g])^2\big] = (E[g])^2 E\Big[\Big(\frac{g}{E[g]} - 1\Big)^2\Big] \geq (E[g])^2 E[\pi(1)^2]. \tag{2.5}$$

Of course, if $E[g] = 0$, (2.5) still holds simply because $\mathrm{Var}[g] \geq 0$.

Now take a sequence $(g_n)_{n\in\mathbb{N}}$ in $G_T(\Theta)$ with $E[g_n^2] \to \infty$ as $n \to \infty$ and let $(n_k)_{k\in\mathbb{N}}$ be a subsequence. If $(E[g_{n_k}])_{k\in\mathbb{N}}$ is bounded (say by C), then

$$F(g_{n_k}) = E[g_{n_k}] - \xi \mathrm{Var}[g_{n_k}] \leq C + \xi C^2 - \xi E[g_{n_k}^2] \longrightarrow -\infty \quad \text{as } k \to \infty.$$

Suppose that $(E[g_{n_k}])_{k\in\mathbb{N}}$ is not bounded. By taking a further subsequence, we may assume that $E[g_{n_k}] \to \infty$ as $k \to \infty$. Then (2.5) with $E[\pi(1)^2] > 0$ gives

$$F(g_{n_k}) = E[g_{n_k}] - \xi \mathrm{Var}[g_{n_k}] \leq E[g_{n_k}] - \xi (E[g_{n_k}])^2 E[\pi(1)^2] \longrightarrow -\infty \quad \text{as } k \to \infty.$$

This establishes that every subsequence $(F(g_{n_k}))_{k\in\mathbb{N}}$ of $(F(g_n))_{n\in\mathbb{N}}$ has a further subsequence converging to $-\infty$ and thus verifies $F(g_n) \to -\infty$ as $n \to \infty$. Therefore the functional F is L^2-coercive.

2) L^2-coercivity of F implies that there exists $C > 0$ such that

$$\sup_{g\in G_T(\Theta)} F(g) = \sup_{g\in G_T(\Theta):\|g\|_{L^2(P)}\le C} F(g) =: c.$$

Note that due to $F(g) = E[g] - \xi \mathrm{Var}[g] \le E[g] + \xi E[g^2]$, we have the inequality $c \le C + \xi C^2 < \infty$. Set

$$D := \{g \in G_T(\Theta) : \|g\|_{L^2(P)} \le C\} \tag{2.6}$$

and take a sequence $(g_n)_{n\in\mathbb{N}}$ in D such that $F(g_n) \uparrow c$ as $n \to \infty$. From its definition in (2.6), D is closed and bounded in L^2. So by the Eberlein–Šmulian theorem, see e.g. Bühler and Salamon [14, Theorem 3.4.1], there exists $\widehat{g} \in D$ such that $(g_n)_{n\in\mathbb{N}}$ converges to $\widehat{g}$ in the weak topology of L^2 as $n \to \infty$. This implies $E[g_n] = E[g_n 1] \to E[\widehat{g}]$ and hence $(E[g_n])^2 \to (E[\widehat{g}])^2$ as $n \to \infty$. Because the norm is lower semicontinuous in the weak topology, we have in addition $E[\widehat{g}^2] \le \liminf_{n\to\infty} E[g_n^2]$. Combining this with $F(g_n) \uparrow c$ and the definition of F in (2.4), we obtain

$$\begin{aligned} c = \limsup_{n\to\infty} F(g_n) &= \limsup_{n\to\infty} \big(E[g_n] + \xi(E[g_n])^2 - \xi E[g_n^2]\big) \\ &= E[\widehat{g}] + \xi(E[\widehat{g}])^2 - \xi \liminf_{n\to\infty} E[g_n^2] \\ &\le E[\widehat{g}] + \xi(E[\widehat{g}])^2 - \xi E[\widehat{g}^2] = F(\widehat{g}). \end{aligned}$$

This shows $F(\widehat{g}) \ge c \ge F(g)$ for all $g \in G_T(\Theta)$. Because $\widehat{g} \in D \subseteq G_T(\Theta)$ and hence there exists a corresponding $\widehat{\theta} \in \Theta$ such that $\widehat{g} = G_T(\widehat{\theta})$, we conclude that $\widehat{\theta}$ is a maximiser to the MVPS problem.

3) We turn to the uniqueness of $\widehat{g} = G_T(\widehat{\theta})$. By a perturbation argument, we obtain that any optimiser $\widehat{g}$ satisfies

$$2\xi \mathrm{Cov}(\widehat{g}, g) - E[g] = 0, \quad \forall g \in G_T(\Theta). \tag{2.7}$$

Indeed, fix $g \in G_T(\Theta)$ and $\varepsilon > 0$. By the linearity of $G_T(\Theta)$, we have that $\widehat{g} \pm \varepsilon g$ is in $G_T(\Theta)$. Then using that $\widehat{g}$ is an optimiser and the definition (2.4) of F, we get

$$\begin{aligned} 0 &\leq F(\widehat{g}) - F(\widehat{g} \pm \varepsilon g) \\ &= E[\widehat{g}] - \xi \mathrm{Var}[\widehat{g}] - E[\widehat{g} \pm \varepsilon g] + \xi \mathrm{Var}[\widehat{g} \pm \varepsilon g] \\ &= \mp \varepsilon E[g] + 2\xi \mathrm{Cov}\big(\widehat{g}, \pm \varepsilon g\big) + \xi \mathrm{Var}[\pm \varepsilon g] \\ &= \mp \varepsilon E[g] \pm 2\varepsilon \xi \mathrm{Cov}(\widehat{g}, g) + \varepsilon^2 \xi \mathrm{Var}[g]. \end{aligned} \tag{2.8}$$

Dividing (2.8) by ε and sending $\varepsilon \to 0$ yields $0 \leq \mp E[g] \pm 2\xi \mathrm{Cov}(\widehat{g}, g)$ and hence (2.7).

Now let g_1, g_2 be two maximisers. Then (2.7) with $\widehat{g} \in \{g_1, g_2\}$ and $g = g_1 - g_2$ gives

$$2\xi \mathrm{Cov}(g_1, g_1 - g_2) - E[g_1 - g_2] = 2\xi \mathrm{Cov}(g_2, g_1 - g_2) - E[g_1 - g_2] = 0,$$

which yields

$$0 = 2\xi \mathrm{Cov}(g_1 - g_2, g_1 - g_2) = 2\xi \mathrm{Var}[g_1 - g_2]. \tag{2.9}$$

Moreover, we have also by (2.7) that $E[g_2] = 2\xi \mathrm{Cov}(g_1, g_2) = E[g_1]$. Combining this with (2.9), we conclude that $g_1 = g_2$ P-a.s., which proves the uniqueness of $\widehat{g}$. □

Having the abstract result in Theorem 2.4 is nice, but our main goal is to obtain a dynamic description of the maximiser $\widehat{\theta}$. In this regard, Problem (2.3) is a well-known non-standard stochastic control problem. The non-standard part of its cost criterion is a nonlinear (quadratic) function of an expected value. This makes the problem not directly amenable to a dynamic programming principle, and hence the solution is difficult to construct explicitly.

2.2 An auxiliary problem

To tackle the MVPS problem (2.3) beyond mere existence, we follow the idea from Ankirchner and Dermoune [5] as explained in Chapter 0. This first requires constructing a new probability space supporting independent copies of the original filtration $\mathbb{F}$ and process S. Then, similarly to Section 2.1, we introduce a financial market and discuss investments in that new market. Finally, we formulate in that market a *standard* stochastic control problem which is closely related to the original MVPS problem (2.3).

Fix $L \in \mathbb{N} \cup \{\infty\}$. We construct a probability space $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$ that supports L independent copies of $(\mathbb{F}, S)$. The construction is fairly standard. Nevertheless, we give details below for any L. To this end, consider the product space $(\Omega^{(L)}, \mathcal{F}^{(L)})$ defined by $\Omega^{(L)} = \prod_{\ell=1}^{L} \Omega$, the L-fold Cartesian product of Ω, and the σ-algebra $\mathcal{F}^{(L)}$ generated by all finite rectangles with $\mathcal{F}$-measurable sides, i.e., by the family $\mathcal{Z}$ of all cylinders of the form

$$Z = \{\omega^{(L)} = (\omega_1, \dots, \omega_L) : \omega_{\ell_j} \in E_{\ell_j} \text{ for } j = 1, \dots, k\}$$

for some $k \in \mathbb{N}$ and $E_{\ell_j} \in \mathcal{F}$, $j = 1, \dots, k$. For $L = \infty$, the notation $(\omega_1, \dots, \omega_L)$ stands for $(\omega_1, \omega_2, \dots)$. Also for $L = \infty$, we construct a product measure on $\mathcal{F}^{(\infty)}$ by

$$\mathbf{P}^{(\infty)}[E_{\ell_1} \times E_{\ell_2} \times \cdots \times E_{\ell_k} \times \Omega \times \cdots] = \prod_{j=1}^{k} P[E_{\ell_j}], \quad E_{\ell_j} \in \mathcal{F}, k \in \mathbb{N}. \tag{2.10}$$

By standard results from measure theory, (2.10) defines a unique probability measure on $\mathcal{F}^{(\infty)}$. Having defined $\mathbf{P}^{(\infty)}$, we construct $\mathbf{P}^{(L)}$ for $L \in \mathbb{N}$ as follows. Consider the projection $\pi_{(L)} : \Omega^{(\infty)} \to \Omega^{(L)}$ onto the first L coordinates given by $\pi_{(L)}(\omega_1, \omega_2, \dots) = (\omega_1, \dots, \omega_L)$. Then we have or set

$$\begin{aligned} \Omega^{(L)} &= \pi_{(L)}(\Omega^{(\infty)}), \\ \mathcal{F}^{(L)} &= \{E \subseteq \Omega : \pi_{(L)}^{-1}(E) \in \mathcal{F}^{(\infty)}\}, \\ \mathbf{P}^{(L)} &:= \mathbf{P}^{(\infty)} \circ \pi_{(L)}^{-1}, \end{aligned}$$

where $\circ$ denotes the standard operation of functional composition. Indeed, the first identity is straightforward by definition and the second involves a standard measure-theoretic argument. From the third, by (2.10) and the definition of $\pi_{(L)}$, we have explicitly

$$\begin{aligned} \mathbf{P}^{(L)}[E_1 \times E_2 \times \cdots \times E_L] &= \mathbf{P}^{(L)}[\pi_{(L)}^{-1}(E_1 \times E_2 \times \cdots \times E_L)] \\ &= \prod_{\ell=1}^{L} P[E_\ell], \quad E_\ell \in \mathcal{F}, L \in \mathbb{N}. \end{aligned} \tag{2.11}$$

The previous paragraph gives a new probability space $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$ for $L \in \mathbb{N} \cup \{\infty\}$. Let us make connections between the original space $(\Omega, \mathcal{F}, P)$ and the new space. Consider the canonical projection $\pi_{\ell,L} : \Omega^{(L)} \to \Omega$ from $\Omega^{(L)}$ onto its ℓ-th coordinate for $\ell = 1, \dots, L$. In particular, by $\ell = 1, \dots, \infty$, we mean that

$\ell \in \mathbb{N}$. Via (2.11), we can recover P from $\mathbf{P}^{(L)}$ by

$$P[E] = \mathbf{P}^{(L)}[\pi_{\ell,L}^{-1}(E)], \quad E \in \mathcal{F}, \ell = 1, \dots, L. \tag{2.12}$$

Thanks to (2.12), we can probabilistically identify any event $E \in \mathcal{F}$ with the event $\pi_{\ell,L}^{-1}(E) \in \mathcal{F}^{(L)}$ for any L and $\ell = 1, \dots, L$. More explicitly, we consider

$$\mathcal{F}^{\ell,(L)} := \pi_{\ell,L}^{-1}(\mathcal{F}) := \{\pi_{\ell,L}^{-1}(E) : E \in \mathcal{F}\} \subseteq \mathcal{F}^{(L)}, \quad \ell = 1, \dots, L, \tag{2.13}$$

and identify $\mathcal{F}$ with $\mathcal{F}^{\ell,(L)}$.

Similarly to (2.13), we also consider for the filtration $\mathbb{F}$ and any $L \in \mathbb{N} \cup \{\infty\}$ the collection $(\mathbb{F}^{\ell,(L)})_{\ell=1,\dots,L}$ of filtrations defined by

$$\mathcal{F}_t^{\ell,(L)} := \pi_{\ell,L}^{-1}(\mathcal{F}_t) = \{\pi_{\ell,L}^{-1}(E) : E \in \mathcal{F}_t\}, \quad \ell = 1, \dots, L, t \in \mathbb{T}. \tag{2.14}$$

From (2.12)–(2.14), it is easy to see that

$$\begin{aligned}&\text{the family } (\mathcal{F}^{\ell,(L)})_{\ell=1,\dots,L} \text{ is } \mathbf{P}^{(L)}\text{-independent and hence}\\ &(\mathcal{F}_t^{\ell,(L)})_{\ell=1,\dots,L} \text{ are } \mathbf{P}^{(L)}\text{-independent for any } t \in \mathbb{T}.\end{aligned} \tag{2.15}$$

Via (2.14) and (2.15), we can thus interpret each $\mathbb{F}^{\ell,(L)}$ as a copy of $\mathbb{F}$ in the product space $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$, and the individual copies are independent. Finally, we construct $\mathbf{S}^{(L)}$ on $\Omega^{(L)}$ from S as $\mathbf{S}_t^{\ell,(L)} = S_t \circ \pi_{\ell,L}$ for $t \in \mathbb{T}$ and $\ell = 1, \dots, L$. More explicitly, we have

$$\mathbf{S}_t^{\ell,(L)}(\omega_1, \dots, \omega_L) := S_t\big(\pi_{\ell,L}(\omega_1, \dots, \omega_L)\big) = S_t(\omega_\ell), \quad \ell = 1, \dots, L, t \in \mathbb{T}. \tag{2.16}$$

Notice that the above identity for $L = \infty$ means $\mathbf{S}_t^{\ell,(\infty)}(\omega_1, \omega_2, \dots) = S_t(\omega_\ell)$ for $\ell \in \mathbb{N}$ and $t \in \mathbb{T}$. From (2.16), (2.14), and (2.15), it is clear that each process $\mathbf{S}^{\ell,(L)}$ is $\mathbb{F}^{\ell,(L)}$-adapted and that the processes $(\mathbf{S}^{\ell,(L)})_{\ell=1,\dots,L}$ are $\mathbf{P}^{(L)}$-independent. Moreover, by (2.12) and (2.16), we see that each process $\mathbf{S}^{\ell,(L)}$ has under $\mathbf{P}^{(L)}$ the same finite-dimensional distributions as S under P. Therefore, we can view each $(\mathbb{F}^{\ell,(L)}, \mathbf{S}^{\ell,(L)})$ as a copy of $(\mathbb{F}, S)$ on the space $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$, and the individual copies are independent. However, it is unnatural to use a collection $(\mathbb{F}^{\ell,(L)})_{\ell=1,\dots,L}$ of filtrations for recording the evolution of information. Instead, we want to have a single filtration to which every $\mathbf{S}^{\ell,(L)}$ is adapted for $\ell = 1, \dots, L$.

To this end, we introduce a new filtration $\mathbb{G}^{(L)} := (\mathcal{G}_t^{(L)})_{t\in\mathbb{T}}$ via

$$\mathcal{G}_t^{(L)} := \sigma\Bigg(\bigcup_{\ell=1}^{L} \mathcal{F}_t^{\ell,(L)}\Bigg), \quad t \in \mathbb{T}. \tag{2.17}$$

In analogy to $(\Omega, \mathcal{F}, P, \mathbb{F})$, the resulting filtered space $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)}, \mathbb{G}^{(L)})$ is our desired space where we next discuss investments.

We pause here to introduce a generic way of "lifting" an object living on the original space $(\Omega, \mathcal{F}, P)$ to the product space $(\Omega^{(L)}, \mathbf{P}^{(L)}, \mathcal{F}^{(L)})$. Given a process X, we define $X^{\otimes L}$ by

$$X^{\ell,\otimes L} := X \circ \pi_{\ell,L}, \quad \ell = 1, \ldots, L, \tag{2.18}$$

where $\pi_{\ell,L}$ is as above the canonical projection from $\Omega^{(L)}$ onto its ℓ-th coordinate. Roughly speaking, each $X^{\ell,\otimes L}$ is the same process X applied to coordinates on $\Omega^{(L)}$ which are independent because $\mathbf{P}^{(L)}$ is the product measure. In view of the above notation, (2.16) can be written as

$$\mathbf{S}^{(L)} = S^{\otimes L}.$$

In the remainder of this subsection, we assume $L \in \mathbb{N}$. Recall that in the original market, we have a set Θ standing for an abstraction of trading strategies. Analogously, we define

$$\Theta^{\ell,\otimes L} = \{\theta \circ \pi_{\ell,L} : \theta \in \Theta\}. \tag{2.19}$$

Like $\mathcal{F}^{\ell,(L)}$ in (2.14), each $\Theta^{\ell,\otimes L}$ can be viewed as a direct analogue of Θ associated with $(\mathbf{S}^{\ell,(L)}, \mathbb{F}^{\ell,(L)})$ for each $\ell = 1, \ldots, L$, and in particular by (2.14) and (2.19), every process in $\Theta^{\ell,\otimes L}$ is $\mathbb{F}^{\ell,(L)}$-predictable. However, because of that coordinatewise predictability restriction, we do *not* use processes from $(\Theta^{\ell,\otimes L})_{\ell=1,\ldots,L}$ as (part of) our trading strategies for a new problem that will be presented below shortly. Instead, we consider

$$\begin{aligned}\boldsymbol{\Theta}^{(L)} := \{(\boldsymbol{\vartheta}_t^{(L)})_{t\in\mathbb{T}} : \;& \boldsymbol{\vartheta}^{(L)} \text{ is } \mathbb{R}^{d\times L}\text{-valued, } \mathbb{G}^{(L)}\text{-predictable, and} \\ & \boldsymbol{\vartheta}^{\ell,(L)} \text{ satisfies the integrability condition} \\ & \text{specified for } \Theta^{\ell,\otimes L}, \text{ for } \ell = 1, \ldots, L\}.\end{aligned} \tag{2.20}$$

The key point here is that while the coordinates of $\mathbf{S}^{(L)} = S^{\otimes L}$ as well as of a generic element of $(\theta^{\ell,\otimes L})_{\ell=1,\dots,L}$ are independent, the coordinates of $\boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}$ are not independent because each depends on $\mathbb{G}^{(L)}$. Because we have not yet specified any particular integrability condition on Θ or equivalently on $\Theta^{\ell,\otimes L}$, the third condition in (2.20) is only a formal description of the trading strategies, which will turn into a precise definition after we give more details about S and Θ. At the moment, let us work at this level of generality and proceed to introduce a notation for the gains process on the new space. For each $\boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}$, we define the vector gains process $\mathbf{G}(\boldsymbol{\vartheta}^{(L)}) = (\mathbf{G}^{\ell}(\boldsymbol{\vartheta}^{(L)}))_{\ell=1,\dots,L}$ by

$$\mathbf{G}^{\ell}_t(\boldsymbol{\vartheta}^{(L)}) := \int_0^t \boldsymbol{\vartheta}^{\ell,(L)}_s \, \mathrm{d}\mathbf{S}^{\ell,(L)}_s, \quad \ell = 1, \dots, L, t \in \mathbb{T}. \tag{2.21}$$

As a general (and vague) summary, we shall use strategies from $\boldsymbol{\Theta}^{(L)}$ to invest in $\mathbf{S}^{(L)}$ on the filtered space $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)}, \mathbb{G}^{(L)})$. This motivates the following definition.

Definition 2.5. For each $L \in \mathbb{N}$, we call the tuple $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)}, \mathbb{G}^{(L)}, \mathbf{S}^{(L)})$ an *L-extended market*. For $L = 1$, the extended market coincides with the original market $(\Omega, \mathcal{F}, P, \mathbb{F}, S)$. When there is no need to mention the underlying probability space, we also refer to the triple $(\mathbf{P}^{(L)}, \mathbb{G}^{(L)}, \mathbf{S}^{(L)})$ as the L-extended market.

Note that in this L-extended market, all quantities are L-tuples of things we know from the original market. To reduce from L-tuples to the original size, we form averages. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^L$, we set

$$\mathbf{x} \odot \mathbf{y} := (\mathbf{x}^{\ell}\mathbf{y}^{\ell})_{\ell=1,\dots,L}, \quad (\mathbf{x})^2 := \mathbf{x} \odot \mathbf{x}, \tag{2.22}$$

$$\mathrm{em}(\mathbf{x}) := \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{x}^{\ell}, \tag{2.23}$$

$$\mathrm{evar}(\mathbf{x}) := \mathrm{em}\big((\mathbf{x})^2\big) - \big(\mathrm{em}(\mathbf{x})\big)^2 = \mathrm{em}(\mathbf{x} \odot \mathbf{x}) - \big(\mathrm{em}(\mathbf{x})\big)^2. \tag{2.24}$$

This means that

$$\mathrm{evar}(\mathbf{x}) = \frac{1}{L}\sum_{\ell=1}^{L} (\mathbf{x}^{\ell})^2 - \bigg(\frac{1}{L}\sum_{\ell=1}^{L} \mathbf{x}^{\ell}\bigg)^2.$$

If $b \in \mathbb{R}$, we mean by $\mathrm{em}(b)$ the average $\mathrm{em}(\mathbf{b})$ with $\mathbf{b} = (b, b, \dots, b) \in \mathbb{R}^L$, so that of course $\mathrm{em}(b) = b$.

Now we are ready to present a standard stochastic control problem in the L-extended market. This problem is closely related to (2.3) and forms the main subject of the next few sections. The idea is to replace (2.3) by

$$\text{maximise } \mathbf{E}^{(L)}\big[\mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big)\big] \text{ over all } \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}, \quad (2.25)$$

where $\mathbf{E}^{(L)}$ denotes the expectation under the measure $\mathbf{P}^{(L)}$ and $\xi > 0$ as before is a positive real number. Here, we also assume $L \geq 2$ to avoid the triviality that $\mathrm{evar}(\mathbf{x}) = 0$ by (2.24) for any $\mathbf{x} \in \mathbb{R}^1$.

The key point (as we argue in detail in Section 2.3 below) is that problem (2.25) becomes a standard stochastic control problem with the state variable $\mathbf{G}(\boldsymbol{\vartheta}^{(L)})$ and control $\boldsymbol{\vartheta}^{(L)}$. Then we can proceed via a dynamic programming principle which leads to linear systems of equations derived from the first order condition for optimality. Under appropriate conditions, a solution $\widehat{\boldsymbol{\vartheta}}^{(L)}$ to the first order condition can be computed explicitly for every $L \in \mathbb{N}$ with $L \geq 2$.

We end this section by outlining the remaining steps for solving the original problem (2.3); their details are elaborated in Section 5. The idea is to make comparisons between problems (2.3) and (2.25). Let us introduce handy notations J_T^{mv} and $J_T^{(L)}$ by

$$J_T^{\mathrm{mv}}(\theta) := G_T(\theta) - \xi\big(G_T(\theta) - E[G_T(\theta)]\big)^2, \quad \theta \in \Theta, \quad (2.26)$$

$$J_T^{(L)}(\boldsymbol{\vartheta}^{(L)}) := \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big), \quad \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}, \quad (2.27)$$

respectively, so that $E[J_T^{\mathrm{mv}}(\theta)] = E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)]$. Using (2.26) and (2.27), we can equivalently write problems (2.3) and (2.25) as to

$$\text{maximise } E[J_T^{\mathrm{mv}}(\theta)] \text{ over all } \theta \in \Theta, \quad (2.28)$$

$$\text{maximise } \mathbf{E}^{(L)}[J_T^{(L)}(\boldsymbol{\vartheta}^{(L)})] \text{ over all } \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}, \quad (2.29)$$

respectively. Viewing the original market as a coordinate of the extended market, we consider analogously to (2.26) for any $\boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}$ the quantity

$$J_T^{\mathrm{mv},\ell}(\boldsymbol{\vartheta}^{\ell,(L)}) = \mathbf{G}_T^{\ell}(\boldsymbol{\vartheta}^{\ell,(L)}) - \xi\big(\mathbf{G}_T^{\ell}(\boldsymbol{\vartheta}^{\ell,(L)}) - E^{\mathbf{P}^{(L)}}[\mathbf{G}_T^{\ell}(\boldsymbol{\vartheta}^{\ell,(L)})]\big)^2, \quad (2.30)$$

for $\ell = 1, \ldots, L$. More specifically, for each $L \in \mathbb{N}$ and $\theta \in \Theta$, we can use the lifting technique (2.18) to find a product-type strategy $\theta^{\otimes L} \in \boldsymbol{\Theta}^{(L)}$ consisting of (independent) copies of θ in the extended market. By construction, the final gains $\mathbf{G}_T^{\ell}(\theta^{\otimes L})$ are then $\mathbf{P}^{(L)}$-independent and have the same distribution as $G_T(\theta)$. Inserting $G_T(\theta)$, $\mathbf{G}_T^{\ell}(\theta^{\otimes L})$ into (2.26), (2.30) respectively and using the i.i.d. property of $\mathbf{G}_T(\theta^{\otimes L})$, then using some form of a law of large numbers, and finally using the optimality of a solution $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for (2.29), we should get

$$\begin{aligned} E[J_T^{\mathrm{mv}}(\theta)] &= \mathbf{E}^{(L)}[J_T^{\mathrm{mv},\ell}(\theta^{\otimes L})] \\ &= \lim_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\theta^{\otimes L})] \\ &\leq \limsup_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})]. \end{aligned} \tag{2.31}$$

Finally, we go back to the original market and try to construct a strategy $\widehat{\theta} \in \Theta$ such that $E[J_T^{\mathrm{mv}}(\widehat{\theta})] = \limsup_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})]$, by exploiting the formal limit of the explicit expression for $\widehat{\boldsymbol{\vartheta}}^{(L)}$ as $L \to \infty$. Then (2.31) readily implies that $\widehat{\theta}$ is an optimal strategy for (2.28).

As we can see from the presentation so far, the superscript $^{(L)}$ makes the notations quite heavy. **For ease of notation, we drop the superscript** $^{(L)}$ whenever we are working only with the extended market and L is fixed. In particular, we still write $\Omega, \mathcal{F}$, but use $\mathbf{P}$, $(\mathbb{F}^\ell)_{\ell=1,\dots,L}$, $\mathbf{S} = (\mathbf{S}^\ell)_{\ell=1,\dots,L}$ instead of $\mathbf{P}^{(L)}$, $(\mathbb{F}^{\ell,(L)})_{\ell=1,\dots,L}$ and $\mathbf{S}^{(L)} = (\mathbf{S}^{\ell,(L)})_{\ell=1,\dots,L}$, respectively. Note the difference between $P, \mathbf{P}$ and $S, \mathbf{S}$, respectively.

Convention 2.6. When L is clear from the context, we write the L-extended market in Definition 2.5 as $(\Omega, \mathcal{F}, \mathbf{P}, \mathbb{G}, \mathbf{S})$ or simply $(\mathbf{P}, \mathbb{G}, \mathbf{S})$. Moreover, the expectation $\mathbf{E}^{(L)}$ is written as $\mathbf{E}$. We also use L^p to denote the equivalence classes of p-integrable random variables when the reference probability measure is clear. For consistency, we use letters in boldface only to refer to quantities in the extended market.

2.3 Martingale optimality principle

We have claimed above that (2.25) is a standard stochastic control problem. Let us argue this in this subsection and therefore collect some general results on the martingale optimality principle (MOP) for the optimisation problem (2.25), whose assumptions will be verified later. In particular, these results hold both in discrete and continuous time and do not need a specific choice of Θ or $\boldsymbol{\Theta}$.

First, we fix $L \in \mathbb{N}$ with $L \geq 2$. Recall the extended market $(\mathbf{P}, \mathbb{G}, \mathbf{S})$ from Definition 2.5 and Convention 2.6, as well as the filtration $\mathbb{G}$ from (2.17). We introduce the notation $\boldsymbol{\Theta}$ standing for an abstraction of a set of trading strategies in the extended market. For any $\mathbb{T}$-valued $\mathbb{G}$-stopping time τ and $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$, define

$$\boldsymbol{\Theta}(\tau, \boldsymbol{\vartheta}) := \{\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta} : \widetilde{\boldsymbol{\vartheta}} = \boldsymbol{\vartheta} \text{ on } \rrbracket 0, \tau \rrbracket \cap \mathbb{T}\}, \tag{2.32}$$

$$J_\tau(\widetilde{\boldsymbol{\vartheta}}) := \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widetilde{\boldsymbol{\vartheta}})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\widetilde{\boldsymbol{\vartheta}})\big)\big|\mathcal{G}_\tau\big]. \tag{2.33}$$

The latter is consistent with (2.27) (without the superscript $^{(L)}$) when $\tau = T$. Also, we define the *value family* $\mathcal{V}(\boldsymbol{\vartheta})$ to problem (2.25) for $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and any $\mathbb{T}$-valued $\mathbb{G}$-stopping time τ by

$$V_\tau(\boldsymbol{\vartheta}) := \operatorname{ess\,sup}\{J_\tau(\widetilde{\boldsymbol{\vartheta}}) : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(\tau, \boldsymbol{\vartheta})\}. \tag{2.34}$$

For $J_\tau(\widetilde{\boldsymbol{\vartheta}})$ to be well defined, given that

$$J_\tau(\widetilde{\boldsymbol{\vartheta}}) = \mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_\tau] \tag{2.35}$$

in view of (2.33), we need

$$J_T(\widetilde{\boldsymbol{\vartheta}}) \in L^1, \quad \forall \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta},$$

which gives a first condition on $\boldsymbol{\Theta}$. Here we adopt Convention 2.6 to write $L^1(\mathbf{P})$ as L^1. To compute $V_T(\boldsymbol{\vartheta})$, we use (2.34), then $\boldsymbol{\Theta}(T, \boldsymbol{\vartheta}) = \{\boldsymbol{\vartheta}\}$ by (2.32) and (2.33) to obtain

$$\begin{aligned} V_T(\boldsymbol{\vartheta}) &= \operatorname{ess\,sup}\{J_T(\widetilde{\boldsymbol{\vartheta}}) : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(T, \boldsymbol{\vartheta})\} \\ &= J_T(\boldsymbol{\vartheta}) \\ &= \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big). \end{aligned} \tag{2.36}$$

For the other boundary case $V_0(\boldsymbol{\vartheta})$, we similarly use (2.34), then $\boldsymbol{\Theta}(0, \boldsymbol{\Theta}) = \boldsymbol{\Theta}$ and (2.35) plus the $\mathbf{P}$-triviality of $\mathcal{G}_0$ to deduce that

$$V_0(\boldsymbol{\vartheta}) = \operatorname*{ess\,sup}_{\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(0, \boldsymbol{\vartheta})} J_0(\widetilde{\boldsymbol{\vartheta}}) = \sup_{\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}} \mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}})] =: V_0 \tag{2.37}$$

is the value of the optimisation problem (2.25) and that this expression is independent of $\boldsymbol{\vartheta}$.

Before we state the MOP, we add more conditions and list all of them below.

Condition 2.7. 1) $J_T(\widetilde{\boldsymbol{\vartheta}}) \in L^1$ for all $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}$.

2) The family $\{J_\tau(\widetilde{\boldsymbol{\vartheta}}) : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(\tau, \boldsymbol{\vartheta})\}$ is upward directed for any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and

any $\mathbb{T}$-valued $\mathbb{G}$-stopping time τ.

3) $\boldsymbol{\Theta}(\tau,\boldsymbol{\vartheta}) \subseteq \boldsymbol{\Theta}(\sigma,\boldsymbol{\vartheta})$ for any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and any $\mathbb{T}$-valued $\mathbb{G}$-stopping times σ,τ with $\sigma \le \tau$ $\mathbf{P}$-a.s.

4) $V_0 = \sup_{\widetilde{\boldsymbol{\vartheta}}\in\boldsymbol{\Theta}} \mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}})] < \infty$.

Now fix $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and consider the value family

$$\mathcal{V}(\boldsymbol{\vartheta}) = \{V_\tau(\boldsymbol{\vartheta}) : \tau \text{ is a } \mathbb{T}\text{-valued } \mathbb{G}\text{-stopping time}\}$$

of random variables. Recall that $\mathcal{V}(\boldsymbol{\vartheta})$ is called a *supermartingale system* (resp., a *martingale system*) if $\mathcal{V}(\boldsymbol{\vartheta}) \subseteq L^1$ and

$$\mathbf{E}[V_\tau(\boldsymbol{\vartheta})|\mathcal{G}_\sigma] \le V_\sigma(\boldsymbol{\vartheta}) \qquad \big(\text{resp. } \mathbf{E}[V_\tau(\boldsymbol{\vartheta})|\mathcal{G}_\sigma] = V_\sigma(\boldsymbol{\vartheta})\big) \tag{2.38}$$

for any $\mathbb{T}$-valued $\mathbb{G}$-stopping times σ,τ with $\sigma \le \tau$ $\mathbf{P}$-a.s.

Lemma 2.8. *Suppose that Condition* 2.7 *is satisfied. Then the following statements hold:*

1) *For any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$, the family $\mathcal{V}(\boldsymbol{\vartheta})$ is a supermartingale system in the filtration $\mathbb{G}$.*

2) *Suppose that $\boldsymbol{\vartheta}^* \in \boldsymbol{\Theta}$. Then $\boldsymbol{\vartheta}^*$ is optimal for* (2.25) *if and only if $\mathcal{V}(\boldsymbol{\vartheta}^*)$ is a martingale system in the filtration $\mathbb{G}$.*

Proof. Fix $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and a stopping time τ. Because $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}(\tau,\boldsymbol{\vartheta})$ and $J_T(\boldsymbol{\vartheta})$ is in L^1 by Condition 2.7, 1), we obtain by (2.35) that $J_\tau(\boldsymbol{\vartheta}) = \mathbf{E}[J_T(\boldsymbol{\vartheta})|\mathcal{G}_\tau] \in L^1$ and hence

$$V_\tau(\boldsymbol{\vartheta}) \ge J_\tau(\boldsymbol{\vartheta}) > -\infty \quad \mathbf{P}\text{-a.s.} \tag{2.39}$$

This implies that the conditional expectation in (2.38) is well defined with values in $(-\infty,+\infty]$. Now we argue the two statements separately.

1) Let σ be a stopping time with $\sigma \le \tau$ $\mathbf{P}$-a.s. By Condition 2.7, 2) and due to (2.39), there exists a sequence $(\overline{\boldsymbol{\vartheta}}^n)_{n\in\mathbb{N}}$ in $\boldsymbol{\Theta}(\tau,\boldsymbol{\vartheta})$ such that $J_\tau(\overline{\boldsymbol{\vartheta}}^n) \uparrow V_\tau(\boldsymbol{\vartheta})$ $\mathbf{P}$-a.s. as $n \to \infty$. Again using Condition 2.7, 2), we can find $\boldsymbol{\vartheta}^n \in \boldsymbol{\Theta}(\tau,\boldsymbol{\vartheta})$ such that $\max(J_\tau(\boldsymbol{\vartheta}), J_\tau(\overline{\boldsymbol{\vartheta}}^n)) \le J_\tau(\boldsymbol{\vartheta}^n)$ for each $n \in \mathbb{N}$. This clearly yields a sequence $(\boldsymbol{\vartheta}^n)_{n\in\mathbb{N}}$ in $\boldsymbol{\Theta}(\tau,\boldsymbol{\vartheta})$ such that $J_\tau(\boldsymbol{\vartheta}^n) \uparrow V_\tau(\boldsymbol{\vartheta})$ and $J_\tau(\boldsymbol{\vartheta}^n) \ge J_\tau(\boldsymbol{\vartheta}) \in L^1$. Thus we can use the monotone convergence theorem, then $(\boldsymbol{\vartheta}^n)_{n\in\mathbb{N}} \subseteq \boldsymbol{\Theta}(\tau,\boldsymbol{\vartheta})$, next Condition 2.7, 3), and finally (2.34) and $\mathbf{E}[J_\tau(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_\sigma] = J_\sigma(\widetilde{\boldsymbol{\vartheta}})$ from the definition

(2.35) to obtain the supermartingale property

$$\begin{aligned}\mathbf{E}[V_\tau(\boldsymbol{\vartheta})|\mathcal{G}_\sigma] &= \mathbf{E}\Big[\lim_{n\to\infty} J_\tau(\boldsymbol{\vartheta}^n)\Big|\mathcal{G}_\sigma\Big] \\ &= \lim_{n\to\infty} \mathbf{E}[J_\tau(\boldsymbol{\vartheta}^n)|\mathcal{G}_\sigma] \\ &\le \operatorname{ess\,sup}\{\mathbf{E}[J_\tau(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_\sigma] : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(\tau,\boldsymbol{\vartheta})\} \\ &\le \operatorname{ess\,sup}\{\mathbf{E}[J_\tau(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_\sigma] : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(\sigma,\boldsymbol{\vartheta})\} \\ &= V_\sigma(\boldsymbol{\vartheta}). \end{aligned} \tag{2.40}$$

We now show $V_\tau(\boldsymbol{\vartheta}) \in L^1$ with the help of (2.40) and in particular that both $V_\tau^+(\boldsymbol{\vartheta}), V_\tau^-(\boldsymbol{\vartheta})$ are in L^1. For the negative part, the inequality (2.39) readily implies that $V_\tau^-(\boldsymbol{\vartheta}) \le J_\tau^-(\boldsymbol{\vartheta}) \in L^1$. For the positive part, using the identity $V_\tau^+(\boldsymbol{\vartheta}) = V_\tau(\boldsymbol{\vartheta}) + V_\tau^-(\boldsymbol{\vartheta})$, then the supermartingale property (2.40), and finally Condition 2.7, 4) with $V_\tau^-(\boldsymbol{\vartheta}) \in L^1$, we get

$$\mathbf{E}[V_\tau^+(\boldsymbol{\vartheta})] = \mathbf{E}[V_\tau(\boldsymbol{\vartheta})] + \mathbf{E}[V_\tau^-(\boldsymbol{\vartheta})] \le V_0 + \mathbf{E}[V_\tau^-(\boldsymbol{\vartheta})] < \infty.$$

This completes the proof of 1).

2) Because $\mathcal{V}(\boldsymbol{\vartheta}^*)$ is a supermartingale system due to 1), it is a martingale system if and only if it has constant expectation, which is in turn equivalent to $\mathbf{E}[V_T(\boldsymbol{\vartheta}^*)] = \mathbf{E}[V_0(\boldsymbol{\vartheta}^*)] = V_0$ thanks to $\mathbf{P}$-triviality of $\mathcal{G}_0$. We then write this equivalent equality $\mathbf{E}[V_T(\boldsymbol{\vartheta}^*)] = V_0$, using $V_T(\boldsymbol{\vartheta}^*) = J_T(\boldsymbol{\vartheta}^*)$ by (2.36) and $V_0 = \sup\{\mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}})] : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}\}$ by (2.37), as

$$\mathbf{E}[J_T(\boldsymbol{\vartheta}^*)] = \mathbf{E}[V_T(\boldsymbol{\vartheta}^*)] = V_0 = \sup_{\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}} \mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}})],$$

which is equivalent to the optimality of $\boldsymbol{\vartheta}^*$ for problem (2.25) by (2.33). □

Lemma 2.8 may look slightly different from well-known forms of the martingale optimality principle, due to the presence of the family $\mathcal{V}(\boldsymbol{\vartheta})$ and the concept of supermartingale systems. However, this formulation is entirely classical; see for instance the masterful presentation in El Karoui [27, Chapter I]. Moreover, the formulation in Lemma 2.8 gives a unified presentation that holds both for discrete and continuous time.

In discrete time, Lemma 2.8 directly reduces to the following statement.

Lemma 2.9. *Suppose that Condition* 2.7 *is satisfied and* $\mathbb{T} = \{0, 1, \dots, T\}$. *Then the following statements hold:*

1) *For any* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$*, the process* $(V_t(\boldsymbol{\vartheta}))_{t=0,1,\dots,T}$ *is a supermartingale in the filtration* $\mathbb{G}$.

2) *Suppose that* $\boldsymbol{\vartheta}^* \in \boldsymbol{\Theta}$*. Then* $\boldsymbol{\vartheta}^*$ *is optimal for* (2.25) *if and only if* $(V_t(\boldsymbol{\vartheta}^*))_{t=0,1,\dots,T}$ *is a martingale in the filtration* $\mathbb{G}$.

In continuous time, when working with a supermartingale or martingale X, one usually wants to have (at least a version such) that the path $t \mapsto X_t$ is RCLL $\mathbf{P}$-a.s. But the process given by $V(\boldsymbol{\vartheta}) = (V_t(\boldsymbol{\vartheta}))_{t\in[0,T]}$ is just a collection of random variables contained in $\mathcal{V}(\boldsymbol{\vartheta})$ and a priori has no path regularity properties for $t \mapsto V_t(\boldsymbol{\vartheta})$. We devote the rest of this subsection to this delicate issue occurring only in continuous time. The reader may jump directly to the next subsection if he/she is only interested in results in discrete time.

Condition 2.10. For each $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$, there exists an adapted RCLL process $\widetilde{V}(\boldsymbol{\vartheta})$ such that for each $\mathbb{T}$-valued $\mathbb{G}$-stopping time τ, we have $\widetilde{V}_\tau(\boldsymbol{\vartheta}) = V_\tau(\boldsymbol{\vartheta})$ $\mathbf{P}$-a.s.

The process $\widetilde{V}(\boldsymbol{\vartheta})$ in Condition 2.10 *aggregates* the family $\mathcal{V}(\boldsymbol{\vartheta})$ into an RCLL process, which is also a version of $V(\boldsymbol{\vartheta})$. If Condition 2.10 is satisfied, we fix such a $\widetilde{V}(\boldsymbol{\vartheta})$ and work with it. For ease of notation, we then also still write V instead of $\widetilde{V}$. Combining Conditions 2.7 and 2.10 with Lemma 2.8, we finally present the following version of a martingale optimality principle.

Lemma 2.11. *Suppose that Conditions* 2.7 *and* 2.10 *are satisfied and choose* $V(\boldsymbol{\vartheta})$ *to be an RCLL aggregation of* $\mathcal{V}(\boldsymbol{\vartheta})$ *if necessary. Then the following statements hold:*

1) *For any* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$*, the process* $(V_t(\boldsymbol{\vartheta}))_{t\in\mathbb{T}}$ *is a supermartingale in the filtration* $\mathbb{G}$.

2) *Suppose* $\boldsymbol{\vartheta}^* \in \boldsymbol{\Theta}$*. Then* $\boldsymbol{\vartheta}^*$ *is optimal for* (2.25) *if and only if* $(V_t(\boldsymbol{\vartheta}^*))_{t\in\mathbb{T}}$ *is a martingale in the filtration* $\mathbb{G}$.

We end this subsection with a sanity check — Condition 2.10 automatically holds when $\mathbb{T} = \{0, 1, \dots, T\}$ because we can then choose $\widetilde{V}_t(\boldsymbol{\vartheta}) = V_t(\boldsymbol{\vartheta})$ for $t = 0, 1, \dots, T$ and use that any path $t \mapsto X_t$ is continuous in discrete time. Hence Lemma 2.11 reduces to Lemma 2.9 when $\mathbb{T}$ is discrete.

2.4 Some technical results on shrinkage of filtration

To use Lemma 2.11, we need to verify in later sections that some process is a supermartingale/martingale in the filtration $\mathbb{G}$. In this subsection, we prepare some tools that can be used to reduce conditional expectations with respect to the filtration $\mathbb{G}$ given in (2.17) to simpler filtrations like $\mathbb{F}^\ell$ for $\ell = 1, \dots, L$. All results below refer to an abstract probability space $(\Omega, \mathcal{A}, P)$.

Lemma 2.12. *Suppose $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are two independent σ-algebras and X, Y are $\widetilde{\mathcal{A}}$- resp. $\widetilde{\mathcal{B}}$-measurable integrable random variables with $XY \in L^1$. Then*

$$E[XY|\sigma(\mathcal{A},\mathcal{B})] = E[X|\mathcal{A}]\, E[Y|\mathcal{B}] \tag{2.41}$$

for all σ-algebras $\mathcal{A} \subseteq \widetilde{\mathcal{A}}$ and $\mathcal{B} \subseteq \widetilde{\mathcal{B}}$. In particular, for $Y \equiv 1$, we have

$$E[X|\sigma(\mathcal{A},\mathcal{B})] = E[X|\mathcal{A}]. \tag{2.42}$$

Proof. The RHS of (2.41) is clearly $\sigma(\mathcal{A},\mathcal{B})$-measurable. To check the averaging property, fix $A \in \mathcal{A}$ and $B \in \mathcal{B}$. We use the independence of $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$, the definition of the conditional expectations $E[X|\mathcal{A}]$ and $E[Y|\mathcal{A}]$ and again the independence of $\mathcal{A}$ and $\mathcal{B}$ to obtain

$$\begin{aligned} E[XY\mathbb{1}_A\mathbb{1}_B] &= E[X\mathbb{1}_A]E[Y\mathbb{1}_B] \\ &= E\big[E[X|\mathcal{A}]\mathbb{1}_A\big]E\big[E[Y|\mathcal{B}]\mathbb{1}_B\big] \\ &= E\big[E[X|\mathcal{A}]E[Y|\mathcal{B}]\mathbb{1}_A\mathbb{1}_B\big]. \end{aligned} \tag{2.43}$$

Because $\mathcal{D} := \{D \in \sigma(\mathcal{A},\mathcal{B}) : E[XY\mathbb{1}_D] = E[E[X|\mathcal{A}]E[Y|\mathcal{B}]\mathbb{1}_D]\}$ is a λ-system and contains by (2.43) the π-system $\{A \cap B : A \in \mathcal{A}, B \in \mathcal{B}\}$, we see from Dynkin's π-λ theorem that $\mathcal{D} = \sigma(\mathcal{A},\mathcal{B})$. The definition of $\mathcal{D}$ then yields that (2.41) is true. □

Lemma 2.13. *Let $\mathbb{A}$ and $\mathbb{B}$ be two filtrations with $\mathcal{B}_T$ independent of $\mathcal{A}_T$. If $(X_t)_{t\in\mathbb{T}}$ is a martingale with respect to $\mathbb{A}$, then it is also a martingale in the filtration $\mathbb{A}\vee\mathbb{B}$ given by $(\mathbb{A}\vee\mathbb{B})_t = \sigma(\mathcal{A}_t,\mathcal{B}_t)$. Conversely, if $(X_t)_{t\in\mathbb{T}}$ is a martingale in the filtration $\mathbb{A}\vee\mathbb{B}$ and adapted to $\mathbb{A}$, it is also a martingale in the filtration $\mathbb{A}$.*

Proof. Since the adaptedness, integrability and (if needed) path regularity of $(X_t)_{t\in\mathbb{T}}$ are all clear, we only check the martingale properties. Let $s, t \in \mathbb{T}$ with $s < t$. Because $X_t \in L^1$ for $t \in \mathbb{T}$, and $\mathcal{A}_T$ and $\mathcal{B}_T$ are independent and hence also $\mathcal{A}_s$, $\mathcal{B}_s$ are independent, we can use (2.42) with $(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}, \mathcal{A}, \mathcal{B}) = (\mathcal{A}_T, \mathcal{B}_T, \mathcal{A}_s, \mathcal{B}_s)$ and the martingale property of $(X_t)_{t\in\mathbb{T}}$ with respect to the filtration $\mathbb{A}$ to obtain $E[X_t|\sigma(\mathcal{A}_s,\mathcal{B}_s)] = E[X_t|\mathcal{A}_s] = X_s$. This verifies the martingale property for the first statement. For the second statement, we only check the martingale property of $(X_t)_{t\in\mathbb{T}}$ with respect to the filtration $\mathbb{A}$. Using the tower property, the martingale property of $(X_t)_{t\in\mathbb{T}}$ with respect to $\mathbb{A}\vee\mathbb{B}$, and that $(X_t)_{t\in\mathbb{T}}$ is adapted to $\mathbb{A}$, we obtain $E[X_t|\mathcal{A}_s] = E[E[X_t|\sigma(\mathcal{A}_s,\mathcal{B}_s)]|\mathcal{A}_s] = E[X_s|\mathcal{A}_s] = X_s$. □

Lemma 2.14. *Suppose that $(X_t)_{t\in\mathbb{T}}$ and $(Y_t)_{t\in\mathbb{T}}$ are martingales in the filtrations $\mathbb{A}$ and $\mathbb{B}$, respectively, with $\mathcal{A}_T$ independent of $\mathcal{B}_T$. Then the process $(X_tY_t)_{t\in\mathbb{T}}$ is a martingale in the filtration $\mathbb{A}\vee\mathbb{B}$.*

Proof. The adaptedness and path regularity of $(X_tY_t)_{t\in\mathbb{T}}$ are again clear. For each t, X_t is $\mathcal{A}_t$-measurable and Y_t is $\mathcal{B}_t$-measurable. Because $\mathcal{A}_T\supseteq\mathcal{A}_t$ and $\mathcal{B}_T\supseteq\mathcal{B}_t$ are independent, this implies that X_t and Y_t are independent, and so $X_tY_t\in L^1$ as $E[|X_tY_t|] = E[|X_t|]E[|Y_t|] < \infty$. Hence it remains to check the martingale property. Because X_t, Y_t, X_tY_t are all in L^1 for $t\in\mathbb{T}$ and $\mathcal{A}_T,\mathcal{B}_T$ are independent and hence $\mathcal{A}_s,\mathcal{B}_s$ are independent, we can use (2.41) with the choice $(\widetilde{\mathcal{A}},\widetilde{\mathcal{B}},\mathcal{A},\mathcal{B}) = (\mathcal{A}_T,\mathcal{B}_T,\mathcal{A}_s,\mathcal{B}_s)$ and the martingale properties of $(X_t)_{t\in\mathbb{T}}$ and $(Y_t)_{t\in\mathbb{T}}$ with respect to the filtrations $\mathbb{A}$ and $\mathbb{B}$, respectively, to obtain

$$E[X_tY_t|\sigma(\mathcal{A}_s,\mathcal{B}_s)] = E[X_t|\mathcal{A}_s]E[Y_t|\mathcal{B}_s] = X_sY_s.$$

This completes the proof. □

We need the following result to apply Lemmas 2.13 and 2.14 to multiple independent σ-algebras.

Lemma 2.15. *Let $(\mathcal{F}^i)_{i\in\mathbb{I}}$ be a family of arbitrarily many independent σ-algebras and let $\mathcal{I}$ be an arbitrary disjoint partition of $\mathbb{I}$. Then the family $(\mathcal{F}^I)_{I\in\mathcal{I}}$ of σ-algebras given by $\mathcal{F}^I = \sigma(\mathcal{F}^i, i\in I)$ is also independent.*

Proof. See Kallenberg [40, Corollary 4.7]. □

3 The auxiliary problem in finite discrete time

In this section, we elaborate on the auxiliary problem (2.25) to

$$\text{maximise } \mathbf{E}\big[\text{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) - \xi\text{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big)\big] \text{ over all } \boldsymbol{\vartheta}\in\boldsymbol{\Theta},$$

specifically in finite discrete time. **Throughout this section**, $L\in\mathbb{N}$ with $L\geq 2$ is fixed and $\mathbb{T} = \{0,1,\dots,T\}$. From (2.36), we recall that

$$J_T(\boldsymbol{\vartheta}) = \text{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) - \xi\text{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big).$$

Using this equality, the corresponding dynamic value process $V(\boldsymbol{\vartheta})$ in (2.34) is

$$V_t(\boldsymbol{\vartheta}) = \operatorname{ess\,sup}\{\mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_t] : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t,\boldsymbol{\vartheta})\}, \quad t = 0,1,\dots,T, \tag{3.1}$$

and the value of the auxiliary problem (2.25) is $V_0 = V_0(\boldsymbol{\vartheta})$ by (2.37). Because we want to obtain both V_0 and an optimal strategy $\widehat{\boldsymbol{\vartheta}}$, we need to compute the entire process $(V_t(\boldsymbol{\vartheta}))_{t=0,1,\dots,T}$ for any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$. As a first step, we present an abstract tool for reducing the computation of the family of complicated essential suprema in (3.1) to a sequence of one-step problems by exploiting the martingale optimality principle in Lemma 2.9. Next, we go away from the abstract presentation and provide a concrete setup where we can ultimately solve the auxiliary problem (2.25) in the next section.

3.1 Dynamic programming in discrete time

We start by rewriting the global result in Lemma 2.9 as a sequence of local results.

Lemma 3.1. *Suppose that Condition* 2.7, 1)–3) *are satisfied. For any* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ *and any* $t = 1,\dots,T$, *we then have*

$$V_{t-1}(\boldsymbol{\vartheta}) = \operatorname{ess\,sup}\big\{\mathbf{E}[V_t(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_{t-1}] : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})\big\}. \tag{3.2}$$

Proof. We argue analogously to Lemma 2.8. Fix $\boldsymbol{\vartheta}$ and $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})$. For "$\le$" in (3.2), we observe that (2.35) and (2.34) yield $\mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_t] = J_t(\widetilde{\boldsymbol{\vartheta}}) \le V_t(\widetilde{\boldsymbol{\vartheta}})$, and hence again by (2.35)

$$J_{t-1}(\widetilde{\boldsymbol{\vartheta}}) = \mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_{t-1}] \le \mathbf{E}[V_t(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_{t-1}]. \tag{3.3}$$

Note that we use here Condition 2.7, 1) to ensure that all the conditional expectations are well defined. Taking essential suprema over $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})$ on both sides of (3.3) and using the definition (2.34) of $V_{t-1}(\boldsymbol{\vartheta})$ yields "$\le$" in (3.2).

For "$\ge$" in (3.2), we fix $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})$. By its definition in (3.1), we have

$$V_t(\widetilde{\boldsymbol{\vartheta}}) = \operatorname{ess\,sup}\{\mathbf{E}[J_T(\overline{\boldsymbol{\vartheta}})|\mathcal{G}_t] : \overline{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t,\widetilde{\boldsymbol{\vartheta}})\},$$

and thus we can use Condition 2.7, 2) as in the proof of Lemma 2.8, 1) to find a sequence $(\widetilde{\boldsymbol{\vartheta}}^n)_{n\in\mathbb{N}}$ in $\boldsymbol{\Theta}(t,\widetilde{\boldsymbol{\vartheta}})$ such that $J_t(\widetilde{\boldsymbol{\vartheta}}) \le J_t(\widetilde{\boldsymbol{\vartheta}}^n) = \mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}}^n)|\mathcal{G}_t] \uparrow V_t(\widetilde{\boldsymbol{\vartheta}})$ with $J_t(\widetilde{\boldsymbol{\vartheta}}) \in L^1$. Next, we observe $\widetilde{\boldsymbol{\vartheta}}^n \in \boldsymbol{\Theta}(t,\widetilde{\boldsymbol{\vartheta}}) \subseteq \boldsymbol{\Theta}(t-1,\widetilde{\boldsymbol{\vartheta}}) = \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})$, where the last equality uses that $\widetilde{\boldsymbol{\vartheta}}$ is in $\boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})$ and the definition (2.32) of

$\Theta(t-1,\boldsymbol{\vartheta})$ so that the restrictions on a process imposed up to $t-1$ by $\widetilde{\boldsymbol{\vartheta}}$ and by $\boldsymbol{\vartheta}$ are the same. Then we use monotone convergence and $\widetilde{\boldsymbol{\vartheta}}^n \in \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})$ for all $n \in \mathbb{N}$ to obtain

$$\mathbf{E}[V_t(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_{t-1}] = \lim_{n\to\infty} \mathbf{E}\big[\mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}}^n)|\mathcal{G}_t]\big|\mathcal{G}_{t-1}\big] = \lim_{n\to\infty} \mathbf{E}[J_T(\widetilde{\boldsymbol{\vartheta}}^n)|\mathcal{G}_{t-1}] \le V_{t-1}(\boldsymbol{\vartheta}).$$

Since $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})$ is arbitrary, we take essential suprema over $\widetilde{\boldsymbol{\vartheta}}$ on the left-hand side of the above inequality to obtain "$\ge$" in (3.2). □

Neither stating nor proving Lemma 3.1 requires Lemma 2.9 or Lemma 2.11. In finite discrete time, we can therefore just rely on Lemma 3.1 to proceed. This is, however, not possible in continuous time, which is why we still presented the two lemmas in Section 2. Note also that Lemma 3.1 does not need Condition 2.7, 4).

One important consequence of (3.2) is that $V_{t-1}(\boldsymbol{\vartheta})$ depends only on the restriction of $\boldsymbol{\vartheta}$ to $\rrbracket 0, t-1\rrbracket \cap \mathbb{T}$, or, more explicitly, on $\boldsymbol{\vartheta}_1,\dots,\boldsymbol{\vartheta}_{t-1}$, for any $t=1,\dots,T$ and any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$. Now fix $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$. To get $V_{t-1}(\boldsymbol{\vartheta})$ from (3.2), we need to consider $V_t(\widetilde{\boldsymbol{\vartheta}})$, where $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})$. Since $\widetilde{\boldsymbol{\vartheta}}_s = \boldsymbol{\vartheta}_s$ for all $s=1,\dots,t-1$ and the latter variables are fixed, the value of $V_t(\widetilde{\boldsymbol{\vartheta}})$ depends on $\widetilde{\boldsymbol{\vartheta}}$ only through $\widetilde{\boldsymbol{\vartheta}}_t$. As a result, it is sufficient to optimise over random variables $\widetilde{\boldsymbol{\vartheta}}_t$ rather than over stochastic processes $\widetilde{\boldsymbol{\vartheta}}$. This observation allows us to simplify (3.2). For $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and for any $\mathcal{G}_{t-1}$-measurable $\mathbb{R}^d$-valued $\boldsymbol{\delta}_t$, define

$$\boldsymbol{\Theta}^{[t]}(\boldsymbol{\vartheta}) := \{\widetilde{\boldsymbol{\vartheta}}_t : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}(t-1,\boldsymbol{\vartheta})\}, \quad \boldsymbol{\vartheta}(t,\boldsymbol{\delta}_t) := (\boldsymbol{\vartheta}_1,\dots,\boldsymbol{\vartheta}_{t-1},\boldsymbol{\delta}_t). \tag{3.4}$$

Then (3.2) can be rewritten as

$$V_{t-1}(\boldsymbol{\vartheta}) = \operatorname{ess\,sup}\big\{\mathbf{E}\big[V_t\big(\boldsymbol{\vartheta}(t,\boldsymbol{\delta}_t)\big)\big|\mathcal{G}_{t-1}\big] : \boldsymbol{\delta}_t \in \boldsymbol{\Theta}^{[t]}(\boldsymbol{\vartheta})\big\} \tag{3.5}$$

for any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}$ and $t=1,\dots,T$. (Note the slight abuse of notation — V_t should be a function of some $\boldsymbol{\vartheta} = (\boldsymbol{\vartheta}_1,\dots,\boldsymbol{\vartheta}_T)$, whereas $\boldsymbol{\vartheta}(t,\boldsymbol{\delta}_t)$ has only length t. But this is no problem because V_t only depends on the first t coordinates of $\boldsymbol{\vartheta}$ anyway.) This is now a recursive sequence of one-step (conditional) problems that we can tackle backward in time, starting with $V_T(\boldsymbol{\vartheta}) = J_T(\boldsymbol{\vartheta})$, to find $(V_t(\boldsymbol{\vartheta}))_{t=0,1,\dots,T}$ as well as the optimal strategy $\widehat{\boldsymbol{\vartheta}}$.

3.2 A concrete setup in dimension 1

The presentation so far is still very abstract. To give more concrete results for the MVPS problem (2.3), we now turn to a specific setup. The goal of this subsection is to give conditions on Θ and S such that Assumption 2.2 is satisfied. **We assume from now on** that $d = 1$ so that S is one-dimensional.

To make a specific choice of Θ, we consider

$$\Theta_{\mathrm{S}} := \{\theta := (\theta_t)_{t=1,\dots,T} : \theta \text{ is real-valued}, \mathbb{F}\text{-predictable and } \theta_t \triangle S_t \in L^2 \text{ for all } t = 1,\dots,T\}. \tag{3.6}$$

Note that Θ_{S} is non-empty because it contains the constant process 0. But beyond this, it is not clear whether Θ_{S} contains other nonzero processes without knowing anything about the integrability of S. To proceed, we introduce the following assumption.

Assumption 3.2. S is square-integrable, meaning that for $t = 0, 1, \dots, T$ we have $S_t \in L^2$.

Remark 3.3. Because $\mathbb{T}$ is finite, Assumption 3.2 is equivalent to saying that $\sup_{t\in\mathbb{T}} |S_t| \in L^2$. In continuous time, these two assumptions become different.

We use Assumption 3.2 and apply Doob's decomposition in the filtration $\mathbb{F}$ to obtain a square-integrable martingale M and a square-integrable predictable process A, both null at 0 and both with respect to $\mathbb{F}$, such that $S = S_0 + M + A$. Explicitly, for $t = 1, \dots, T$, we have

$$M_t = \sum_{s=1}^{t} (S_s - E[S_s|\mathcal{F}_{s-1}]), \quad A_t = \sum_{s=1}^{t} (E[S_s|\mathcal{F}_{s-1}] - S_{s-1}). \tag{3.7}$$

To give more structure on the process S and ensure that Θ_{S} defined in (3.6) satisfies Assumption 2.2, we need to introduce some definitions and notations. For a square-integrable process $(X_t)_{t=0,1,\dots,T}$ adapted to the filtration $\mathbb{F}$, we introduce the two processes $[X]$ and $\langle X\rangle$ via

$$[X]_0 := 0, \quad \triangle[X]_t := (\triangle X_t)^2, \quad t = 1, \dots, T, \tag{3.8}$$

$$\langle X\rangle_0 := 0, \quad \triangle\langle X\rangle_t := E\big[\triangle[X]_t \big| \mathcal{F}_{t-1}\big] = E[(\triangle X_t)^2|\mathcal{F}_{t-1}], \quad t = 1, \dots, T. \tag{3.9}$$

From (3.9), the predictability of A and (3.8), we have $\triangle\langle A\rangle_t = (\triangle A_t)^2 = \triangle[A]_t$, and using $E[\triangle M_t \triangle A_t|\mathcal{F}_{t-1}] = 0$ due to the predictability of A, the martingale property of M and the integrability property from Assumption 3.2 additionally yields

$$\triangle\langle S\rangle_t = E[(\triangle S_t)^2|\mathcal{F}_{t-1}] = E[(\triangle M_t)^2|\mathcal{F}_{t-1}]+(\triangle A_t)^2 = \triangle\langle M\rangle_t+(\triangle A_t)^2. \quad (3.10)$$

Note also that by (3.9) and (3.7),

$$\triangle\langle M\rangle_t = E[(\triangle M_t)^2|\mathcal{F}_{t-1}] = \mathrm{Var}[\triangle S_t|\mathcal{F}_{t-1}]. \quad (3.11)$$

Assumption 3.4. The process S satisfies the *structure condition*, meaning that the process A is absolutely continuous with respect to the process $\langle M\rangle$. We write $A \ll \langle M\rangle$.

Because (3.10) always gives $\langle M\rangle \ll \langle S\rangle$, Assumption 3.4 implies that we have $\langle S\rangle \approx \langle M\rangle$ and can thus define the predictable processes

$$\lambda = \frac{\mathrm{d}A}{\mathrm{d}\langle M\rangle}\,, \qquad \widetilde{\lambda} = \frac{\mathrm{d}A}{\mathrm{d}\langle S\rangle}\,. \quad (3.12)$$

Assumption 3.4 also says that $A = \int \lambda\,\mathrm{d}\langle M\rangle = \sum \lambda\triangle\langle M\rangle$ and therefore implies both that

$$\triangle A_t = \triangle A_t \mathbb{1}_{\{\triangle\langle M\rangle_t\neq 0\}}, \quad t = 1,\dots,T, \quad (3.13)$$

and, using (3.7) and (3.11),

$$\triangle A_t = E[\triangle S_t|\mathcal{F}_{t-1}] = 0 \quad \text{on } \{\mathrm{Var}[\triangle S_t|\mathcal{F}_{t-1}] = 0\}. \quad (3.14)$$

As a consequence of (3.13), we get

$$A_t = \sum_{s=1}^{t} \triangle A_s = \sum_{s=1}^{t} \triangle A_s \mathbb{1}_{\{\triangle\langle M\rangle_s\neq 0\}} = \sum_{s=1}^{t} \frac{\triangle A_s}{\triangle\langle M\rangle_s} \mathbb{1}_{\{\triangle\langle M\rangle_s\neq 0\}}\triangle\langle M\rangle_s.$$

The uniqueness of the Radon–Nikodým derivatives in (3.12) therefore implies that

$$\lambda_t = \frac{\triangle A_t}{\triangle\langle M\rangle_t}\mathbb{1}_{\{\triangle\langle M\rangle_t\neq 0\}}, \quad t = 1,\dots,T. \quad (3.15)$$

Similarly, we use additionally that $\{\triangle\langle S\rangle_t \neq 0\} = \{\triangle\langle M\rangle_t \neq 0\}$ to obtain

$$\widetilde{\lambda}_t = \frac{\triangle A_t}{\triangle\langle S\rangle_t}\mathbb{1}_{\{\triangle\langle S\rangle_t\neq 0\}} = \frac{\triangle A_t}{\triangle\langle M\rangle_t + (\triangle A_t)^2}\mathbb{1}_{\{\triangle\langle M\rangle_t\neq 0\}}, \quad t = 1,\dots,T. \quad (3.16)$$

Note that $\lambda_t = 0$ on $\{\triangle\langle M\rangle_t = 0\}$. With the convention $\frac{0}{0} := 0$, we can write

(3.15) and (3.16) more compactly as

$$\lambda_t = \frac{\triangle A_t}{\triangle\langle M\rangle_t}, \quad \widetilde{\lambda}_t = \frac{\triangle A_t}{\triangle\langle M\rangle_t + (\triangle A_t)^2}, \quad t = 1, \dots, T. \tag{3.17}$$

Definition 3.5. We define the *mean–variance tradeoff (MVT) process* of S to be

$$K_t := \int_0^t \lambda_s \, \mathrm{d}A_s, \quad t = 1, \dots, T, \tag{3.18}$$

and the *extended mean–variance tradeoff (EMVT) process* of S to be

$$\widetilde{K}_t := \int_0^t \widetilde{\lambda}_s \, \mathrm{d}A_s, \quad t = 1, \dots, T. \tag{3.19}$$

From (3.17)–(3.19), we get explicit expressions for the increments of the MVT and EMVT processes for $t = 1, \dots, T$ as

$$\triangle K_t = \lambda_t \triangle A_t = \frac{(\triangle A_t)^2}{\triangle\langle M\rangle_t}, \qquad \triangle\widetilde{K}_t = \widetilde{\lambda}_t \triangle A_t = \frac{(\triangle A_t)^2}{\triangle\langle M\rangle_t + (\triangle A_t)^2}. \tag{3.20}$$

Moreover, (3.20) also yields $1+\triangle K_t = \frac{\triangle\langle M\rangle_t+(\triangle A_t)^2}{\triangle\langle M\rangle_t}$ and $1-\triangle\widetilde{K}_t = \frac{\triangle\langle M\rangle_t}{\triangle\langle M\rangle_t+(\triangle A_t)^2}$, which implies that

$$\triangle K_t = \frac{\triangle\widetilde{K}_t}{1-\triangle\widetilde{K}_t}, \quad \triangle\widetilde{K}_t = \frac{\triangle K_t}{1+\triangle K_t}, \quad t = 1, \dots, T. \tag{3.21}$$

Note that writing (3.21) assumes implicitly that $\triangle\widetilde{K}_t < 1$ P-a.s. This is evidently true on $\{\triangle\langle M\rangle_t \neq 0\}$ by (3.20). But using Assumption 3.4, we also have $\triangle\widetilde{K}_t = 0$ on $\{\triangle\langle M\rangle_t = 0\}$, which is consistent with (3.20) and our convention that $\frac{0}{0} = 0$.

The above notations provide handy tools which can be used to impose conditions on S so that Θ_{S} fulfils Assumption 2.2. In view of Theorem 2.4, we then know that Θ_{S} is at least not a bad place to look for a maximiser to the MVPS problem (2.3).

Lemma 3.6. *Suppose Assumptions* 3.2 *and* 3.4 *are satisfied. If the MVT process* K *in* (3.18) *is bounded, then* Θ_{S} *in* (3.6) *satisfies Assumption* 2.2, *i.e.,* $G_T(\Theta_{\mathrm{S}})$ *is closed in* L^2 *and* $1 \notin \overline{G_T(\Theta_{\mathrm{S}})}$.

Proof. For Assumption 2.2, 1), we refer to Schweizer [61, Theorem 2.1]. Note that $\widehat{K}$ there is the same as K here and that (ND) there is equivalent to boundedness of $\widehat{K}$ there.

For 2), we want to show that $\overline{G_T(\Theta_S)}$, or equivalently here $G_T(\Theta_S)$, does not contain the constant payoff 1. To this end, we recall from Schweizer [63] that a *signed $G_T(\Theta_S)$-martingale measure* is a signed measure Q such that $Q[\Omega] = 1$, $Q \ll P$ with $\frac{dQ}{dP} \in L^2$ and

$$E\left[\frac{dQ}{dP}g\right] = 0 \quad \text{for all } g \in G_T(\Theta_S), \tag{3.22}$$

and we denote by $\mathcal{Q}$ the set of all signed $G_T(\Theta_S)$-martingale measures. Also from Schweizer [63, Lemma 4.1], we recall that $G_T(\Theta_S)$ does not contain the constant payoff 1 if and only if $\mathcal{Q} \neq \emptyset$. We now show that the latter is true by constructing an element in $\mathcal{Q}$ directly, using that the MVT process K is bounded. Define

$$Z_0 = 1, \quad Z_t := Z_{t-1}(1 - \lambda_t \triangle M_t) = \prod_{s=1}^{t}(1 - \lambda_s \triangle M_s), \qquad t = 1, \dots, T. \tag{3.23}$$

We claim that the measure Q given by $\frac{dQ}{dP} = Z_T$ is an element of $\mathcal{Q}$. The proof of this claim is divided into four parts as follows.

a) We first argue that $Z_t \in L^2$ for $t = 1, \dots, T$. Because λ_t^2 and $(\triangle M_t)^2$ are nonnegative, we use $\mathcal{F}_{t-1}$-measurability of λ_t^2 and the explicit expressions of $\triangle\langle M\rangle_t, \lambda_t, \triangle K_t$ in (3.9), (3.17) and (3.20), respectively, to obtain for $t = 1, \dots, T$ that

$$E[(\lambda_t \triangle M_t)^2 | \mathcal{F}_{t-1}] = \lambda_t^2 \triangle\langle M\rangle_t = \frac{(\triangle A_t)^2}{(\triangle\langle M\rangle_t)^2}\triangle\langle M\rangle_t = \triangle K_t. \tag{3.24}$$

Taking expectations in (3.24) and using boundedness of the MVT process K by assumption yields

$$\lambda_t \triangle M_t \in L^2, \quad t = 1, \dots, T. \tag{3.25}$$

We now use the integrability in (3.25), the martingale property of M and (3.24) to obtain for $t = 1, \dots, T$ that

$$\begin{aligned} E[(1 - \lambda_t \triangle M_t)^2 | \mathcal{F}_{t-1}] &= E\big[\big(1 - 2\lambda_t \triangle M_t + (\lambda_t \triangle M_t)^2\big)\big|\mathcal{F}_{t-1}\big] \\ &= 1 + \triangle K_t. \end{aligned} \tag{3.26}$$

Because Z_{t-1}^2 and $(1 - \lambda_t \triangle M_t)^2$ are nonnegative, we next use $\mathcal{F}_{t-1}$-measurability of Z_{t-1} from its definition in (3.23) and then (3.26) to get for $t = 1, \dots, T$ that

$$E[Z_t^2 | \mathcal{F}_{t-1}] = E\big[Z_{t-1}^2 E[(1 - \lambda_t \triangle M_t)^2 | \mathcal{F}_{t-1}]\big] = Z_{t-1}^2(1 + \triangle K_t). \tag{3.27}$$

Taking expectations in (3.27) and using induction and boundedness of the MVT process K, we obtain $E[Z_t^2] = E[\prod_{s=1}^{t}(1 + \triangle K_s)] < \infty$ and hence $Z_t \in L^2$ for $t = 1, \ldots, T$.

b) We next show that $(Z_t)_{t=1,\ldots,T}$ is a martingale and hence $E[Z_T] = Z_0 = 1$ by its definition in (3.23). The adaptedness is clear and the integrability is given by part a). For the martingale property, fix $t \in \{1, \ldots, T\}$. Note that both Z_t and $1 - \lambda_t \triangle M_t$ are in L^2 by part a) and (3.25), respectively. We combine this observation with $\mathcal{F}_{t-1}$-measurability of Z_{t-1} and the martingale property of M to obtain

$$E[Z_t|\mathcal{F}_{t-1}] = Z_{t-1} E[1 - \lambda_t \triangle M_t|\mathcal{F}_{t-1}] = Z_{t-1}, \quad t = 1, \ldots, T.$$

This shows that $(Z_t)_{t=1,\ldots,T}$ is a martingale. Hence by (3.23), we also get that $E[Z_T] = E[Z_0] = 1$.

c) We now argue that $(Z_t S_t)_{t=0,1,\ldots,T}$ is also a martingale. The adaptedness is clear. Fix $t \in \{1, \ldots, T\}$. Because Z_t, $1 - \lambda_t \triangle M_t$ and S_t are all in L^2 by part a), (3.25) and Assumption 3.2, respectively, we also get that $Z_t S_t$ and the product $(1 - \lambda_t \triangle M_t)\triangle S_t$ are both in L^1. We use this observation and $\mathcal{F}_{t-1}$-measurability of Z_{t-1}, then $\triangle S_t = \triangle M_t + \triangle A_t$, next $\mathcal{F}_{t-1}$-measurability of λ_t and $\triangle A_t$, that $\lambda_t \triangle M_t \triangle A_t, \lambda_t(\triangle M_t)^2$ are both in L^1 due to (3.25) and Assumption 3.2, and finally the explicit expressions for λ_t and $\triangle\langle M\rangle_t$ in (3.17) and (3.9) to obtain that

$$\begin{aligned}
E[Z_t \triangle S_t|\mathcal{F}_{t-1}] &= Z_{t-1} E[(1 - \lambda_t \triangle M_t)\triangle S_t|\mathcal{F}_{t-1}] \\
&= Z_{t-1}\big(\triangle A_t - E[\lambda_t \triangle M_t \triangle A_t|\mathcal{F}_{t-1}] \\
&\qquad + E[\triangle M_t|\mathcal{F}_{t-1}] - E[\lambda_t(\triangle M_t)^2|\mathcal{F}_{t-1}]\big) \\
&= Z_{t-1}\big(\triangle A_t - \lambda_t \triangle A_t E[\triangle M_t|\mathcal{F}_{t-1}] \\
&\qquad + E[\triangle M_t|\mathcal{F}_{t-1}] - \lambda_t E[(\triangle M_t)^2|\mathcal{F}_{t-1}]\big) \\
&= Z_{t-1}\left(\triangle A_t - \frac{\triangle A_t}{\triangle\langle M\rangle_t}\triangle\langle M\rangle_t\right) = 0, \qquad t = 1, \ldots, T. \qquad (3.28)
\end{aligned}$$

This yields $E[Z_t(S_t - S_{t-1})|\mathcal{F}_{t-1}] = 0$ and therefore

$$E[Z_t S_t|\mathcal{F}_{t-1}] = E[Z_t S_{t-1}|\mathcal{F}_{t-1}] = Z_{t-1} S_{t-1}$$

because Z is a martingale by part b).

d) Lastly, we turn to verifying (3.22) based on parts a)–c). For $g \in G_T(\Theta_{\mathrm{S}})$, we know by definition that $g = \sum_{t=1}^{T} \theta_t \triangle S_t$ for some $\theta = (\theta_t)_{t=1,\dots,T} \in \Theta_{\mathrm{S}}$. Using $\frac{\mathrm{d}Q}{\mathrm{d}P} = Z_T$ with this equality, then $\mathcal{F}_t$-measurability of $\theta_t \triangle S_t$ with part b) and $Z_T \theta_t \triangle S_t \in L^1$ by part a) and $\theta_t \triangle S_t \in L^2$ from the definition (3.6) of Θ_{S}, next $\mathcal{F}_{t-1}$-measurability of θ_t with $Z_t \theta_t \triangle S_t \in L^1$ similarly as $Z_T \theta_t \triangle S_t \in L^1$ and $Z_t \triangle S_t \in L^1$ by part c) and finally (3.28), we write (3.22) as

$$\begin{aligned} E\bigg[\frac{\mathrm{d}Q}{\mathrm{d}P} g\bigg] &= \sum_{t=1}^{T} E\Big[E\big[E[Z_T \theta_t \triangle S_t | \mathcal{F}_t]\big|\mathcal{F}_{t-1}\big]\Big] \\ &= \sum_{t=1}^{T} E\big[E[Z_t \theta_t \triangle S_t | \mathcal{F}_{t-1}]\big] \\ &= \sum_{t=1}^{T} E\big[\theta_t E[Z_t \triangle S_t | \mathcal{F}_{t-1}]\big] \\ &= 0. \end{aligned}$$

This verifies (3.22). Together with $Z_T \in L^2$ from part b) and $Q \ll P$ by construction, we conclude that $Q \in \mathcal{Q}$. This completes the proof. □

3.3 A concrete setup for the auxiliary problem

The previous subsection provides a setup to study the MVPS problem (2.3). Analogously, we give in this subsection a concrete setup for studying the auxiliary problem (2.25). Recall the extended market $(\mathbf{P}, \mathbb{G}, \mathbf{S})$ from Definition 2.5 and Convention 2.6. Although Lemma 3.6 gives sufficient conditions in terms of the model (Θ, S) such that the MVPS problem (2.3) is well-posed by Theorem 2.4, we do *not* seek a similar result in the extended market $(\mathbf{P}, \mathbb{G}, \mathbf{S})$. Instead, our strategy is to use the dynamic programming result from Lemma 3.1 to directly construct an optimiser for the auxiliary problem (2.25). To do this, we still need to choose a good space $\boldsymbol{\Theta}$ which satisfies the premises of Lemma 3.1, and this is of course motivated by (2.20).

In analogy to Θ_{S} given in (3.6), we set

$$\begin{aligned} \boldsymbol{\Theta}_{\mathrm{S}} := \{\boldsymbol{\vartheta} := (\boldsymbol{\vartheta}_t)_{t=1,\dots,T} : \boldsymbol{\vartheta}^\ell &\text{ is real-valued, } \mathbb{G}\text{-predictable and} \\ &\boldsymbol{\vartheta}_t^\ell \triangle \mathbf{S}_t^\ell \in L^2 \text{ for } t = 1,\dots,T, \ell = 1,\dots,L\}. \end{aligned} \tag{3.29}$$

Next, note that Assumption 3.2 and (2.16) imply that $\mathbf{S}^\ell$ is square-integrable under the measure $\mathbf{P}$ for $\ell = 1, \dots, L$. It immediately follows that in this extended market, we can write $\mathbf{S}^\ell = \mathbf{S}_0 + \mathbf{M}^\ell + \mathbf{A}^\ell$, where $\mathbf{M}^\ell, \mathbf{A}^\ell$ are from the Doob decomposition of $\mathbf{S}^\ell$ in the filtration $\mathbb{G}$ given in (2.17), and explicitly like Doob's decomposition (3.7) for S,

$$\mathbf{M}_t^\ell = \sum_{s=1}^t (\mathbf{S}_s^\ell - \mathbf{E}[\mathbf{S}_s^\ell | \mathcal{G}_{s-1}]), \quad \mathbf{A}_t^\ell = \sum_{s=1}^t (\mathbf{E}[\mathbf{S}_s^\ell | \mathcal{G}_{s-1}] - \mathbf{S}_{s-1}^\ell), \tag{3.30}$$

for $t = 1, \dots, T$ and $\ell = 1, \dots, L$. Recall for $\ell = 1, \dots, L$ the filtration $\mathbb{F}^\ell$ from (2.14). The two decompositions (3.7) (in $\mathbb{F}$) and (3.30) (in $\mathbb{G}$) are related in the following way.

Lemma 3.7. *Suppose that Assumption* 3.2 *is satisfied. Then the following statements are true:*

1) *For each* $\ell = 1, \dots, L$, *the Doob decomposition of* $\mathbf{S}^\ell$ *in the filtration* $\mathbb{G}$ *is the same as in the filtration* $\mathbb{F}^\ell$ *given in* (2.14).

2) *For* $\ell, m = 1, \dots, L$ *and* $\ell \neq m$, *the process* $(\mathbf{M}_t^\ell \mathbf{M}_t^m)_{t=0,1,\dots,T}$ *is a martingale in the filtration* $\mathbb{G}$. *Equivalently, for* $t = 1, \dots, T$, *we have*

$$\mathbf{E}[\triangle \mathbf{M}_t^\ell \triangle \mathbf{M}_t^m | \mathcal{G}_{t-1}] = 0 \quad \text{for } \ell \neq m, \tag{3.31}$$

which also means that the martingales $\mathbf{M}^\ell$ *and* $\mathbf{M}^m$ *for* $\ell \neq m$ *are strongly orthogonal (under* $\mathbf{P}$ *with respect to* $\mathbb{G}$*).*

Proof. 1) We first argue that $(\mathbf{M}_t^\ell)_{t=0,1,\dots,T}$ is a martingale in the filtration $\mathbb{F}^\ell$. From (2.17), we can write $\mathcal{G}_t = \sigma(\mathcal{F}_t^\ell, \mathcal{B}_t^\ell)$ with $\mathcal{B}_t^\ell := \sigma(\cup_{j \neq \ell} \mathcal{F}_t^j)$ for $t = 1, \dots, T$. Because $(\mathcal{F}_t^\ell)_{\ell=1,\dots,L}$ are independent by (2.15), we get from Lemma 2.15 that $\mathcal{F}_t^\ell$ and $\mathcal{B}_t^\ell$ are also independent for $t = 0, 1, \dots, T$. Then applying Lemma 2.13 with $\mathbb{A} = \mathbb{F}^\ell$, $\mathbb{B} = (\mathcal{B}_t^\ell)_{t=0,1,\dots,T}$ gives that $(\mathbf{M}_t^\ell)_{t=0,1,\dots,T}$ is a martingale in $\mathbb{F}^\ell$. Inserting (2.42) with $(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}, \mathcal{A}, \mathcal{B}) = (\mathcal{F}_T^\ell, \mathcal{B}_T^\ell, \mathcal{F}_{s-1}^\ell, \mathcal{B}_{s-1}^\ell)$ for $s = 1, \dots, t$ into (3.30), we obtain

$$\mathbf{A}_t^\ell = \sum_{s=1}^t \left(\mathbf{E}[\mathbf{S}_s^\ell | \mathcal{F}_{s-1}^\ell \vee \mathcal{B}_{s-1}^\ell] - \mathbf{S}_{s-1}^\ell \right) = \sum_{s=1}^t (\mathbf{E}[\mathbf{S}_s^\ell | \mathcal{F}_{s-1}^\ell] - \mathbf{S}_{s-1}^\ell).$$

Together with the fact that $(\mathbf{S}_t^\ell)_{t=0,1,\dots,T}$ is $\mathbb{F}^\ell$-adapted by the constructions (2.14) and (2.16), we see that $(\mathbf{A}_t^\ell)_{t=0,1,\dots,T}$ is $\mathbb{F}^\ell$-predictable. By the uniqueness of the Doob decomposition, we finally get that $\mathbf{S}^\ell = \mathbf{S}_0 + \mathbf{M}^\ell + \mathbf{A}^\ell$ is the Doob decomposition of $\mathbf{S}^\ell$ in the filtration $\mathbb{F}^\ell$.

2) The adaptedness is clear. For integrability, we use Assumption 3.2 and the Cauchy–Schwarz inequality to obtain that $\mathbf{M}_t^\ell \mathbf{M}_t^m \in L^1$ for $t = 0, 1, \dots, T$. By part 1), $(\mathbf{M}_t^\ell)_{t=0,1,\dots,T}$ and $(\mathbf{M}_t^m)_{t=0,1,\dots,T}$ are martingales in the filtrations $\mathbb{F}^\ell$

and $\mathbb{F}^m$, respectively. Because $\mathcal{F}_T^\ell$ and $\mathcal{F}_T^m$ are independent by (2.15), we obtain first by Lemma 2.14 with $\mathbb{A} = \mathbb{F}^\ell$, $\mathbb{B} = \mathbb{F}^m$ that $(\mathbf{M}_t^\ell \mathbf{M}_t^m)_{t=0,1,\dots,T}$ is a martingale in the filtration $\mathbb{F}^\ell \vee \mathbb{F}^m$ given by $\mathcal{F}_t^\ell \vee \mathcal{F}_t^m = \sigma(\mathcal{F}_t^\ell, \mathcal{F}_t^m)$. Because $\sigma(\mathcal{F}_T^\ell, \mathcal{F}_T^m)$ and $\sigma(\mathcal{F}_T^j, j \neq \ell, j \neq m)$ are independent, we can then apply Lemma 2.13 with $\mathbb{A} = \mathbb{F}^\ell \vee \mathbb{F}^m$ and $\mathbb{B}$ given by $\mathcal{B}_t = \sigma(\cup_{j \neq \ell, j \neq m} \mathcal{F}_t^j)$ to conclude that $(\mathbf{M}_t^\ell \mathbf{M}_t^m)_{t=0,1,\dots,T}$ is also a martingale in the filtration $\mathbb{G}$. The rest is then obvious from the abstract result that for any two square-integrable martingales M, N with respect to a filtration $\mathbb{H}$, we have $E[\triangle M_t \triangle N_t | \mathcal{H}_s] = E[\triangle [M,N]_t | \mathcal{H}_s] = E[\triangle (MN)_t | \mathcal{H}_s]$ because $MN - [M,N]$ is an $\mathbb{H}$-martingale. This completes the proof. □

Remark 3.8. Fix $\ell \in \{1, \dots, L\}$. We remark that by Lemma 3.7, 1), the decomposition in (3.30) agrees with $\mathbf{M}^\ell = M \circ \pi_{\ell,L}$ and $\mathbf{A}^\ell = A \circ \pi_{\ell,L}$, where M, A are from (3.7) and $\pi_{\ell,L}$ is the canonical projection onto the ℓ-th coordinate. Therefore $\mathbf{M}^\ell$ and $\mathbf{A}^\ell$ have the same distributions as M and A, respectively. Indeed, using (2.42) with $(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}, \mathcal{A}, \mathcal{B}) = (\mathcal{F}_T^\ell, \sigma(\cup_{j \neq \ell} \mathcal{F}_T^j), \mathcal{F}_{t-1}^\ell, \sigma(\cup_{j \neq \ell} \mathcal{F}_{t-1}^j))$, we obtain $\mathbf{E}[\mathbf{S}_t^\ell | \mathcal{G}_{t-1}] = \mathbf{E}[\mathbf{S}_t^\ell | \mathcal{F}_{t-1}^\ell]$ for $t = 1, \dots, T$. By (2.12), (2.14) and (2.16), the last expression is equal to $E[S_t | \mathcal{F}_{t-1}] \circ \pi_{\ell,L}$. Combining this identity with (2.16), we get $\mathbf{S}_t^\ell - \mathbf{E}[\mathbf{S}_t^\ell | \mathcal{G}_{t-1}] = (S_t - E[S_t | \mathcal{F}_{t-1}]) \circ \pi_{\ell,L}$. We use this identity and the explicit formula for M in (3.7) to write (3.30) for $t = 1, \dots, T$ and $\ell = 1, \dots, L$ as

$$\mathbf{M}_t^\ell = \sum_{s=1}^{t} (\mathbf{S}_s^\ell - \mathbf{E}[\mathbf{S}_s^\ell | \mathcal{G}_{s-1}]) = \left(\sum_{s=1}^{t} (S_s - E[S_s | \mathcal{F}_{s-1}]) \right) \circ \pi_{\ell,L} = M_t \circ \pi_{\ell,L}.$$

The identity $\mathbf{A}^\ell = A \circ \pi_{\ell,L}$ holds because $\mathbf{A}^\ell = \mathbf{S}^\ell - \mathbf{S}_0^\ell - \mathbf{M}^\ell$ and $A = S - S_0 - M$. In view of the notation $X^{\ell, \otimes L}(\omega^{(L)}) = X(\omega_\ell)$ for $\ell = 1, \dots, L$ from (2.18), we can simply write $\mathbf{M} = M^{\otimes L}$ and $\mathbf{A} = A^{\otimes L}$. In other words, $\mathbf{M}$ and $\mathbf{A}$ simply consist of independent copies of M and A, respectively.

Now we translate the main notations from (3.8) to (3.21) into corresponding quantities in the extended market $(\mathbf{P}, \mathbb{G}, \mathbf{S})$. While (3.8) remains the same, the filtration used in (3.9) is changed accordingly in the extended market. For a square-integrable process $\mathbf{X} = (\mathbf{X}^\ell)_{\ell=1,\dots,L}$, meaning that $\mathbf{X}_t^\ell \in L^2$ for $t = 1, \dots, T$ and $\ell = 1, \dots, L$, we set

$$\triangle \langle \mathbf{X}^\ell \rangle_0 = 0, \quad \triangle \langle \mathbf{X}^\ell \rangle_t = \mathbf{E}\big[\triangle [\mathbf{X}^\ell]_t \big| \mathcal{G}_{t-1}\big] = \mathbf{E}[(\triangle \mathbf{X}_t^\ell)^2 | \mathcal{G}_{t-1}]. \tag{3.32}$$

By Remark 3.8, $\mathbf{M}^\ell$ and $\mathbf{A}^\ell$ have the same distributions as M and A, respectively. Moreover, using (3.32) and (2.42) in the same way as in Remark 3.8, we obtain $\triangle \langle \mathbf{M}^\ell \rangle_t = \mathbf{E}[(\triangle \mathbf{M}_t^\ell)^2 | \mathcal{G}_{t-1}] = \mathbf{E}[(\triangle \mathbf{M}_t^\ell)^2 | \mathcal{F}_{t-1}^\ell]$. Again in view of the argument in

Remark 3.8, we get $\langle \mathbf{M}^\ell \rangle = \langle M \rangle \circ \pi_{\ell,L}$, and hence $\langle \mathbf{M}^\ell \rangle$ has the same distribution as $\langle M \rangle$. Therefore, Assumption 3.4 carries over to $\mathbf{S}^\ell$, i.e. $\mathbf{A}^\ell \ll \langle \mathbf{M}^\ell \rangle$ and hence $\langle \mathbf{M}^\ell \rangle \approx \langle \mathbf{S}^\ell \rangle$ for $\ell = 1, \dots, L$. In view of (3.32), (3.10) corresponds to

$$\triangle \langle \mathbf{S}^\ell \rangle_t = \triangle \langle \mathbf{M}^\ell \rangle_t + (\triangle \mathbf{A}^\ell_t)^2, \quad t = 1, \dots, T, \ \ell = 1, \dots, L. \tag{3.33}$$

For $\ell = 1, \dots, L$, define

$$\boldsymbol{\lambda}^\ell = \frac{\mathrm{d}\mathbf{A}^\ell}{\mathrm{d}\langle \mathbf{M}^\ell \rangle}, \quad \widetilde{\boldsymbol{\lambda}}^\ell = \frac{\mathrm{d}\mathbf{A}^\ell}{\mathrm{d}\langle \mathbf{S}^\ell \rangle}, \quad \mathbf{K}^\ell = \int \boldsymbol{\lambda}^\ell \, \mathrm{d}\mathbf{A}^\ell, \quad \widetilde{\mathbf{K}}^\ell = \int \widetilde{\boldsymbol{\lambda}}^\ell \, \mathrm{d}\mathbf{A}^\ell. \tag{3.34}$$

The proofs of (3.15) and (3.16) and the definition (3.34) translate (3.17) into

$$\boldsymbol{\lambda}^\ell_t = \frac{\triangle \mathbf{A}^\ell_t}{\triangle \langle \mathbf{M}^\ell \rangle_t}, \quad \widetilde{\boldsymbol{\lambda}}^\ell_t = \frac{\triangle \mathbf{A}^\ell_t}{\triangle \langle \mathbf{M}^\ell \rangle_t + (\triangle \mathbf{A}^\ell_t)^2}, \quad t = 1, \dots, T, \ell = 1, \dots, L. \tag{3.35}$$

In view of (3.34) and (3.35), (3.20) reads for $t = 1, \dots, T$ and $\ell = 1, \dots, L$ as

$$\triangle \mathbf{K}^\ell_t = \boldsymbol{\lambda}^\ell_t \triangle \mathbf{A}^\ell_t = \frac{(\triangle \mathbf{A}^\ell)^2}{\triangle \langle \mathbf{M}^\ell \rangle}, \quad \triangle \widetilde{\mathbf{K}}^\ell_t = \widetilde{\boldsymbol{\lambda}}^\ell_t \triangle \mathbf{A}^\ell_t = \frac{(\triangle \mathbf{A}^\ell)^2}{\triangle \langle \mathbf{M}^\ell \rangle + (\triangle \mathbf{A}^\ell)^2}. \tag{3.36}$$

Remark 3.9. 1) We see from Remark 3.8 and the subsequent discussion that we have $\mathbf{M}^\ell = M \circ \pi_{\ell,L}$, $\mathbf{A}^\ell = A \circ \pi_{\ell,L}$ and $\langle \mathbf{M}^\ell \rangle = \langle M \rangle \circ \pi_{\ell,L}$ for $\ell = 1, \dots, L$. This implies that the processes $\boldsymbol{\lambda}$, $\widetilde{\boldsymbol{\lambda}}$, $\mathbf{K}$ and $\widetilde{\mathbf{K}}$ in (3.34) agree with the vector processes $\lambda^{\otimes L}$, $\widetilde{\lambda}^{\otimes L}$, $K^{\otimes L}$ and $\widetilde{K}^{\otimes L}$, respectively. Thus each coordinate of these processes has the same distributions as λ, $\widetilde{\lambda}$, K and $\widetilde{K}$, respectively. Moreover, due to (2.15), the processes $\mathbf{Y}^\ell$ and $\mathbf{Y}^m$ are independent for $\ell \neq m$ and for any $\mathbf{Y} \in \{\mathbf{M}, \mathbf{A}, \boldsymbol{\lambda}, \widetilde{\boldsymbol{\lambda}}, \mathbf{K}, \widetilde{\mathbf{K}}\}$.

2) Because $\mathbf{M}^\ell$ and $\mathbf{M}^m$ are independent for $\ell \neq m$, we have $\langle \mathbf{M}^\ell, \mathbf{M}^m \rangle \equiv 0$ for $\ell \neq m$; see Lemma 3.7, 2). Therefore the matrix-valued process $\langle \mathbf{M} \rangle$ has a diagonal form, and so it is enough to look only at $\langle \mathbf{M}^\ell \rangle$ for $\ell = 1, \dots, L$.

Finally, we show that $\boldsymbol{\Theta}_{\mathrm{S}}$ in (3.29) is indeed a good choice for dynamic programming in the sense that $\boldsymbol{\Theta} = \boldsymbol{\Theta}_{\mathrm{S}}$ satisfies the assumptions of Lemma 3.1.

Lemma 3.10. *Suppose Assumptions* 3.2 *and* 3.4 *are satisfied. If the MVT process K is bounded, then Condition* 2.7, 1), 2) and 3) *hold with the choice* $\boldsymbol{\Theta} = \boldsymbol{\Theta}_{\mathrm{S}}$ *given in* (3.29).

Proof. Fix $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$. For Condition 2.7, 1), we show $J_T(\boldsymbol{\vartheta}) \in L^1$. By the definition of $\boldsymbol{\Theta}_{\mathrm{S}}$ in (3.29), we have $\mathbf{G}^\ell_T(\boldsymbol{\vartheta}) \in L^2$ for $\ell = 1, \dots, L$. Using this integrability and $0 \leq \mathrm{evar}(\mathbf{x}) = \mathrm{em}((\mathbf{x})^2) - (\mathrm{em}(\mathbf{x}))^2 \leq \mathrm{em}((\mathbf{x})^2)$ from (2.24), we obtain

$$0 \leq \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) \leq \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})^2\big) \in L^1$$

and hence from the explicit expression (2.36) for $J_T(\boldsymbol{\vartheta})$ that

$$J_T(\boldsymbol{\vartheta}) \leq \big|\mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big)\big| + \xi\big|\mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big)\big| \in L^1.$$

Next, we prove Condition 2.7, 3). For $\mathbb{T}$-valued stopping times σ, τ with $\sigma \leq \tau$, we observe from the definition (2.32) of $\boldsymbol{\Theta}_{\mathrm{S}}(\tau, \boldsymbol{\vartheta})$ that any element in $\boldsymbol{\Theta}_{\mathrm{S}}(\tau, \boldsymbol{\vartheta})$ agrees with $\boldsymbol{\vartheta}$ on $]\!]0,\tau]\!] \cap \mathbb{T}$ and thus agrees with $\boldsymbol{\vartheta}$ on $]\!]0,\sigma]\!] \cap \mathbb{T}$ since $]\!]0,\sigma]\!] \subseteq]\!]0,\tau]\!]$. Because elements of $\boldsymbol{\Theta}_{\mathrm{S}}(\sigma, \boldsymbol{\vartheta})$ and of $\boldsymbol{\Theta}_{\mathrm{S}}(\tau, \boldsymbol{\vartheta})$ have the same measurability and integrability conditions, we obtain $\boldsymbol{\Theta}_{\mathrm{S}}(\tau, \boldsymbol{\vartheta}) \subseteq \boldsymbol{\Theta}_{\mathrm{S}}(\sigma, \boldsymbol{\vartheta})$.

To argue Condition 2.7, 2), we first show that $\boldsymbol{\Theta}_{\mathrm{S}}$ is stable under bifurcation, i.e., that for any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$, $\mathbb{T}$-valued $\mathbb{G}$-stopping time τ, $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}_{\mathrm{S}}(\tau, \boldsymbol{\vartheta})$ and event $F \in \mathcal{G}_\tau$, the process $\widetilde{\boldsymbol{\vartheta}}^F := \mathbb{1}_F \boldsymbol{\vartheta} + \mathbb{1}_{F^c} \widetilde{\boldsymbol{\vartheta}}$ is again in $\boldsymbol{\Theta}_{\mathrm{S}}$. In view of the definition (3.29) of $\boldsymbol{\Theta}_{\mathrm{S}}$, we first show that $\widetilde{\boldsymbol{\vartheta}}^F$ is $\mathbb{G}$-predictable. Indeed, because τ is a $\mathbb{G}$-stopping time, both $\mathbb{1}_{]\!]0,\tau]\!]}$ and $\mathbb{1}_{]\!]\tau,T]\!]}$ are $\mathbb{G}$-predictable processes. Moreover, $F \in \mathcal{G}_\tau$ implies that $\mathbb{1}_F \mathbb{1}_{]\!]\tau,T]\!]}$ is $\mathbb{G}$-predictable. Therefore, using $\widetilde{\boldsymbol{\vartheta}} \mathbb{1}_{]\!]0,\tau]\!]} = \boldsymbol{\vartheta} \mathbb{1}_{]\!]0,\tau]\!]}$ from $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}_{\mathrm{S}}(\tau, \boldsymbol{\vartheta})$ yields $\widetilde{\boldsymbol{\vartheta}}^F = \mathbb{1}_F \boldsymbol{\vartheta} + \mathbb{1}_{F^c} \widetilde{\boldsymbol{\vartheta}} = \mathbb{1}_{]\!]0,\tau]\!]} \boldsymbol{\vartheta} + \mathbb{1}_F \mathbb{1}_{]\!]\tau,T]\!]} \boldsymbol{\vartheta} + \mathbb{1}_{F^c} \mathbb{1}_{]\!]\tau,T]\!]} \widetilde{\boldsymbol{\vartheta}}$, which readily shows the $\mathbb{G}$-predictability of $\widetilde{\boldsymbol{\vartheta}}^F$. For the integrability condition in (3.29), we use $(\widetilde{\boldsymbol{\vartheta}}^{F,\ell}_t \triangle \mathbf{S}^\ell_t)^2 \leq (\boldsymbol{\vartheta}^\ell_t \triangle \mathbf{S}^\ell_t)^2 + (\widetilde{\boldsymbol{\vartheta}}^\ell_t \triangle \mathbf{S}^\ell_t)^2 \in L^1$ for $t = 1, \ldots, T$ and $\ell = 1, \ldots, L$. This proves that $\widetilde{\boldsymbol{\vartheta}}^F$ is in $\boldsymbol{\Theta}_{\mathrm{S}}$.

We now turn to verifying Condition 2.7, 2), which says that for any $\mathbb{T}$-valued $\mathbb{G}$-stopping time τ and $\widetilde{\boldsymbol{\vartheta}}^{(1)}, \widetilde{\boldsymbol{\vartheta}}^{(2)} \in \boldsymbol{\Theta}_{\mathrm{S}}(\tau, \boldsymbol{\vartheta})$, there exists $\widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}_{\mathrm{S}}(\tau, \boldsymbol{\vartheta})$ such that $\max\{J_\tau(\widetilde{\boldsymbol{\vartheta}}^{(1)}), J_\tau(\widetilde{\boldsymbol{\vartheta}}^{(2)})\} \leq J_\tau(\widetilde{\boldsymbol{\vartheta}})$. We take such $\widetilde{\boldsymbol{\vartheta}}^{(1)}, \widetilde{\boldsymbol{\vartheta}}^{(2)}$, set

$$F := \{J_\tau(\widetilde{\boldsymbol{\vartheta}}^{(1)}) \geq J_\tau(\widetilde{\boldsymbol{\vartheta}}^{(2)})\} \in \mathcal{G}_\tau$$

and define $\widetilde{\boldsymbol{\vartheta}} := \mathbb{1}_F \widetilde{\boldsymbol{\vartheta}}^{(1)} + \mathbb{1}_{F^c} \widetilde{\boldsymbol{\vartheta}}^{(2)} \in \boldsymbol{\Theta}_{\mathrm{S}}$ by the stability under bifurcation proved above. Using the definition of F, then the definition (2.33) of J_τ with $F \in \mathcal{G}_\tau$, next the definition of $\widetilde{\boldsymbol{\vartheta}}$ and finally (2.33) again, we write

$$
\begin{aligned}
\max\{J_\tau(\widetilde{\boldsymbol{\vartheta}}^{(1)}), J_\tau(\widetilde{\boldsymbol{\vartheta}}^{(2)})\} &= \mathbb{1}_F J_\tau(\widetilde{\boldsymbol{\vartheta}}^{(1)}) + \mathbb{1}_{F^c} J_\tau(\widetilde{\boldsymbol{\vartheta}}^{(2)}) \\
&= E[\mathbb{1}_F J_T(\widetilde{\boldsymbol{\vartheta}}^{(1)}) + \mathbb{1}_{F^c} J_T(\widetilde{\boldsymbol{\vartheta}}^{(2)})|\mathcal{G}_\tau] \\
&= E[\mathbb{1}_F J_T(\widetilde{\boldsymbol{\vartheta}}) + \mathbb{1}_{F^c} J_T(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_\tau] \\
&= E[J_T(\widetilde{\boldsymbol{\vartheta}})|\mathcal{G}_\tau] \\
&= J_\tau(\widetilde{\boldsymbol{\vartheta}}),
\end{aligned}
$$

which shows Condition 2.7, 2). This completes the proof. □

4 Recursive computation of the value process for dimension 1

In this section, we solve the auxiliary problem (3.5) with the choice of $\boldsymbol{\Theta} = \boldsymbol{\Theta}_{\mathrm{S}}$. This will need an extra assumption in addition to Assumptions 3.2 and 3.4. In view of the discussion at the end of Section 3.1, the idea is to compute the value process $(V_t(\boldsymbol{\vartheta}))_{t=0,1,\dots,T}$ in (3.5) backward in time. Let us begin with rewriting (3.5) for $t = 1, \dots, T$ as

$$
V_{t-1}(\boldsymbol{\vartheta}) = \operatorname{ess\,sup}\big\{\mathbf{E}\big[V_t\big(\boldsymbol{\vartheta}(t,\boldsymbol{\delta}_t)\big)\big|\mathcal{G}_{t-1}\big] : \boldsymbol{\delta}_t \in \boldsymbol{\Theta}_{\mathrm{S}}^{[t]}(\boldsymbol{\vartheta})\big\}, \quad t = 1, \dots, T, \tag{4.1}
$$

with $\boldsymbol{\vartheta}(t,\boldsymbol{\delta}_t) = (\boldsymbol{\vartheta}_1, \dots, \boldsymbol{\vartheta}_{t-1}, \boldsymbol{\delta}_t)$ from (3.4) and $V_T(\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T))$ of the general form

$$
\begin{aligned}
&V_T\big(\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T)\big) = a_T \mathrm{em}\big(\mathbf{G}_T\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T)\big) - b_T \mathrm{evar}\big(\mathbf{G}_T\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T)\big) + c_T, \\
&\text{where } a_T, b_T, c_T \text{ are nonrandom constants with } b_T > 0.
\end{aligned} \tag{4.2}
$$

While (4.2) may look like spurious generality because we actually have $a_T = 1$, $b_T = \xi, c_T = 0$, it turns out to be useful to write things in this generality. Moreover, using the definitions of $\boldsymbol{\Theta}(\tau,\boldsymbol{\vartheta})$ and $\boldsymbol{\Theta}_{\mathrm{S}}$ from (2.32) and (3.29), respectively, we write $\boldsymbol{\Theta}_{\mathrm{S}}^{[t]}(\boldsymbol{\vartheta})$ in (3.4) more explicitly as

$$
\begin{aligned}
\boldsymbol{\Theta}_{\mathrm{S}}^{[t]}(\boldsymbol{\vartheta}) &= \{\widetilde{\boldsymbol{\vartheta}}_t : \widetilde{\boldsymbol{\vartheta}} \in \boldsymbol{\Theta}_{\mathrm{S}}(t-1,\boldsymbol{\vartheta})\} \\
&= \{\boldsymbol{\delta}_t : \boldsymbol{\delta}_t^\ell \text{ is real-valued, } \mathcal{G}_{t-1}\text{-measurable} \\
&\qquad \text{and } \boldsymbol{\delta}_t^\ell \triangle \mathbf{S}_t^\ell \in L^2,\ \ell = 1, \dots, L\}.
\end{aligned} \tag{4.3}
$$

In the rest of this section, we present a solution technique for (4.1) in two parts. First, we consider (4.1) for $t = T$ for a fixed $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ and note that (4.1) says that $V_{T-1}(\boldsymbol{\vartheta})$ is obtained by maximising the conditional expectation $\mathbf{E}[V_T(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T))|\mathcal{G}_{T-1}] =: F(\boldsymbol{\delta}_T)$ over $\boldsymbol{\delta}_T$. Because of (4.2), $F(\boldsymbol{\delta}_T)$ is an affine–quadratic function of $\boldsymbol{\delta}_T$, and hence the first-order condition (FOC) for the optimisation of $F(\boldsymbol{\delta}_T)$ over $\boldsymbol{\delta}_T$ is affine. Plugging its solution back in should yield that $V_{T-1}(\boldsymbol{\vartheta})$ is an affine–quadratic function of $\mathbf{G}_{T-1}(\boldsymbol{\vartheta})$, like $V_T(\boldsymbol{\vartheta})$ of $\mathbf{G}_T(\boldsymbol{\vartheta})$. Moreover, it seems plausible that this reasoning can be iterated backwards until we obtain V_0. In the first part, we analyse the above programme rigorously and argue that this solution technique is indeed iteratable under an extra assumption. In the second part, we state the main result of this section — a recursive description of the entire value process and the optimal strategy for (4.1).

We first give a step-by-step recipe that will be implemented below.

Recipe 4.1. 1) Compute $\mathbf{E}[V_T(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T))|\mathcal{G}_{T-1}] =: F(\boldsymbol{\delta}_T)$ as a function of $\boldsymbol{\delta}_T$ explicitly.

2) Maximise $\boldsymbol{\delta}_T \mapsto F(\boldsymbol{\delta}_T)$ via solving an FOC for the optimality.

3) Verify that the candidate maximiser obtained in 2) is in $\boldsymbol{\Theta}_{\mathrm{S}}^{[T]}(\boldsymbol{\vartheta})$ and is indeed a maximiser for (4.1). Plug it back into F to obtain an explicit formula for $V_{T-1}(\boldsymbol{\vartheta})$.

4) Argue carefully how and why steps 1)–3) can be extended to general $t < T$ under an extra assumption.

4.1 Step 1: Computing $F(\boldsymbol{\delta}_T) := \mathbf{E}[V_T(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T))|\mathcal{G}_{T-1}]$

In this subsection, we embark on the programme described in Recipe 4.1 and implement its step 1), from which we recall $F(\boldsymbol{\delta}_T) = \mathbf{E}[V_T(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T))|\mathcal{G}_{T-1}]$. This relies on the Doob decomposition of $\mathbf{S}$ in the filtration $\mathbb{G}$. As we can see from (4.4) and (4.5) below, the conditional expectation $F(\boldsymbol{\delta}_T)$ is indeed an affine–quadratic function in $\boldsymbol{\delta}_T$. Recall from (2.22) the notation $\odot$ for the coordinatewise multiplication.

Lemma 4.2. *Suppose that Assumptions* 3.2 *and* 3.4 *are satisfied. If* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ *and* $V_T(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T))$ *is given by* (4.2) *for* $\boldsymbol{\delta}_T \in \boldsymbol{\Theta}_{\mathrm{S}}^{[T]}(\boldsymbol{\vartheta})$, *then we have*

$$\mathbf{E}\big[V_T\big(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)\big)|\mathcal{G}_{T-1}\big] = a_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) - b_T \mathrm{evar}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + c_T + R_T(\boldsymbol{\delta}_T), \tag{4.4}$$

where

$$
\begin{aligned}
R_T(\boldsymbol{\delta}_T) := {}& a_T \mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T) - 2b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}) \odot \boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T\big) \\
&+ 2b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T) \\
&- b_T \mathrm{em}\Big((\boldsymbol{\delta}_T)^2 \odot \big((1 - L^{-1}) \triangle \langle \mathbf{M} \rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\
&+ b_T \big(\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T)\big)^2. \qquad (4.5)
\end{aligned}
$$

Remark 4.3. In analogy to the empirical variance evar, we define for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^L$ the *empirical covariance* between $\mathbf{x}$ and $\mathbf{y}$ by

$$
\mathrm{ecov}(\mathbf{x}, \mathbf{y}) = \mathrm{em}(\mathbf{x} \odot \mathbf{y}) - \mathrm{em}(\mathbf{x})\, \mathrm{em}(\mathbf{y}). \qquad (4.6)
$$

Using (4.6), we can write the identity (4.5) more compactly and suggestively as

$$
\begin{aligned}
R_T(\boldsymbol{\delta}_T) = {}& a_T \mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T) - 2b_T \mathrm{ecov}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}), \boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T\big) \\
&- b_T \mathrm{em}\Big((\boldsymbol{\delta}_T)^2 \odot \big((1 - L^{-1}) \triangle \langle \mathbf{M} \rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\
&+ b_T \big(\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T)\big)^2. \qquad (4.7)
\end{aligned}
$$

Proof of Lemma 4.2. Let us recall from (2.24) that

$$
\mathrm{evar}(\mathbf{x}) = \mathrm{em}(\mathbf{x}^2) - \big(\mathrm{em}(\mathbf{x})\big)^2. \qquad (4.8)
$$

Taking conditional expectations in (4.2) and using (4.8) and the non-randomness of a_T, b_T, c_T gives

$$
\begin{aligned}
\mathbf{E}\big[V_T\big(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)\big) \big| \mathcal{G}_{T-1}\big] &= a_T \mathbf{E}\Big[\mathrm{em}\Big(\mathbf{G}_T\big(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)\big)\Big) \Big| \mathcal{G}_{T-1}\Big] \\
&\quad - b_T \mathbf{E}\Big[\mathrm{evar}\Big(\mathbf{G}_T\big(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)\big)\Big) \Big| \mathcal{G}_{T-1}\Big] + c_T \\
&= a_T \mathbf{E}\Big[\mathrm{em}\Big(\mathbf{G}_T\big(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)\big)\Big) \Big| \mathcal{G}_{T-1}\Big] \\
&\quad - b_T \mathbf{E}\Big[\mathrm{em}\Big(\mathbf{G}_T\big(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)\big)^2\Big) \Big| \mathcal{G}_{T-1}\Big] \\
&\quad + b_T \mathbf{E}\bigg[\bigg(\mathrm{em}\Big(\mathbf{G}_T\big(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)\big)\Big)\bigg)^2 \bigg| \mathcal{G}_{T-1}\bigg] + c_T. \qquad (4.9)
\end{aligned}
$$

To expand the terms in (4.9), we recall from (2.21), (3.4) and (2.23) the equalities

$$\mathbf{G}_T\big(\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T)\big) = \mathbf{G}_{T-1}(\boldsymbol{\vartheta}) + \boldsymbol{\delta}_T \odot \triangle\mathbf{S}_T, \tag{4.10}$$

$$\mathrm{em}(\mathbf{x}+\mathbf{y}) = \mathrm{em}(\mathbf{x}) + \mathrm{em}(\mathbf{y}), \quad \mathrm{em}(\mathbf{x}) = \frac{1}{L}\sum_{\ell=1}^{L}\mathbf{x}^\ell. \tag{4.11}$$

Next, we recall from Assumption 3.2 and (3.30) for $\ell = 1,\dots,L$ that

$$\triangle\mathbf{S}_T^\ell = \triangle\mathbf{M}_T^\ell + \triangle\mathbf{A}_T^\ell, \quad \mathbf{E}[\triangle\mathbf{M}_T^\ell|\mathcal{G}_{T-1}] = 0, \quad \mathbf{E}[\triangle\mathbf{S}_T^\ell|\mathcal{G}_{T-1}] = \triangle\mathbf{A}_T^\ell. \tag{4.12}$$

We also recall the angle bracket notation from (3.32) and (3.33) as

$$\triangle\langle\mathbf{X}^\ell\rangle_T = \mathbf{E}[(\triangle\mathbf{X}_T^\ell)^2|\mathcal{G}_{T-1}], \quad \triangle\langle\mathbf{S}^\ell\rangle_T = \triangle\langle\mathbf{M}^\ell\rangle_T + (\triangle\mathbf{A}_T^\ell)^2. \tag{4.13}$$

Finally, we recall from Assumption 3.2, (3.29) and (4.3) that for $\ell = 1,\dots,L$,

$$\triangle\mathbf{S}_T^\ell, \triangle\mathbf{M}_T^\ell, \triangle\mathbf{A}_T^\ell, \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta}), \boldsymbol{\delta}_T^\ell\triangle\mathbf{S}_T^\ell, \boldsymbol{\delta}_T^\ell\triangle\mathbf{M}_T^\ell, \boldsymbol{\delta}_T^\ell\triangle\mathbf{A}_T^\ell \text{ are all in } L^2. \tag{4.14}$$

For the first term in (4.9), we now use (4.10), (4.11) and $\triangle\mathbf{S}_T = \triangle\mathbf{M}_T + \triangle\mathbf{A}_T$, then $\mathcal{G}_{T-1}$-measurability of $\mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}))$ and $\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle\mathbf{A}_T)$ and the identity $\mathbf{E}[\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle\mathbf{M}_T)|\mathcal{G}_{T-1}] = 0$, which follows from $\mathcal{G}_{T-1}$-measurability of $\boldsymbol{\delta}_T^\ell$, (4.11), (4.12) and (4.14), to obtain

$$\begin{aligned} a_T\mathbf{E}\Big[\mathrm{em}\Big(\mathbf{G}_T\big(\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T)\big)\Big)\Big|\mathcal{G}_{T-1}\Big] &= a_T\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\big|\mathcal{G}_{T-1}\big] \\ &\quad + a_T\mathbf{E}[\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle\mathbf{S}_T)|\mathcal{G}_{T-1}] \\ &= a_T\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \\ &\quad + a_T\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle\mathbf{A}_T). \end{aligned} \tag{4.15}$$

The last equality also uses that $\mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}))$ and $\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle\mathbf{S}_T)$ are in $L^2 \subseteq L^1$ thanks to (4.11) and the integrability property in (4.14).

For the second term in (4.9), we use (4.10) and (4.11) successively, then the

$\mathcal{G}_{T-1}$-measurability of $\mathbf{G}^\ell_{T-1}(\boldsymbol{\vartheta})$, (4.12) and (4.13) to obtain

$$\begin{aligned}
&b_T\mathbf{E}\Big[\mathrm{em}\Big(\mathbf{G}_T\big(\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T)\big)^2\Big)\Big|\mathcal{G}_{T-1}\Big]\\
&= b_T\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})^2\big)\big|\mathcal{G}_{T-1}\big] + 2b_T\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\odot\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T\big)\big|\mathcal{G}_{T-1}\big]\\
&\quad + b_T\mathbf{E}\big[\mathrm{em}\big((\boldsymbol{\delta}_T)^2\odot(\triangle\mathbf{S}_T)^2\big)\big|\mathcal{G}_{T-1}\big]\\
&= b_T\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})^2\big)\big|\mathcal{G}_{T-1}\big] + 2b_T\frac{1}{L}\sum_{\ell=1}^{L}\mathbf{E}[\mathbf{G}^\ell_{T-1}(\boldsymbol{\vartheta})\boldsymbol{\delta}^\ell_T\triangle\mathbf{S}^\ell_T|\mathcal{G}_{T-1}]\\
&\quad + b_T\frac{1}{L}\sum_{\ell=1}^{L}\mathbf{E}[(\boldsymbol{\delta}^\ell_T\triangle\mathbf{S}^\ell_T)^2|\mathcal{G}_{T-1}]\\
&= b_T\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})^2\big) + 2b_T\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\odot\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T\big)\\
&\quad + b_T\mathrm{em}\Big((\boldsymbol{\delta}_T)^2\odot\big(\triangle\langle\mathbf{M}\rangle_T + (\triangle\mathbf{A}_T)^2\big)\Big). \qquad (4.16)
\end{aligned}$$

The last equality also uses (4.11), the integrability property in (4.14) directly and that $(\mathbf{G}^\ell_{T-1}(\boldsymbol{\vartheta}))^2$, $\mathbf{G}^\ell_{T-1}(\boldsymbol{\vartheta})\boldsymbol{\delta}^\ell_T\triangle\mathbf{S}^\ell_T$ and $(\boldsymbol{\delta}^\ell_T\triangle\mathbf{S}^\ell_T)^2$ are in L^1 for $\ell = 1,\dots,L$ due to (4.14).

For the third term in (4.9), we need several steps. First, we again use (4.10) and (4.11), then $\mathcal{G}_{T-1}$-measurability of $\mathbf{G}^\ell_{T-1}(\boldsymbol{\vartheta})$ and that, as argued in (4.15), $\mathbf{E}[\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T)|\mathcal{G}_{T-1}] = \mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T)$, to write the third term as

$$\begin{aligned}
&b_T\mathbf{E}\bigg[\Big(\mathrm{em}\Big(\mathbf{G}_T\big(\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T)\big)\Big)\Big)^2\bigg|\mathcal{G}_{T-1}\bigg]\\
&= b_T\mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\Big)^2\Big|\mathcal{G}_{T-1}\Big] + 2b_T\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T)\big|\mathcal{G}_{T-1}\big]\\
&\quad + b_T\mathbf{E}\big[\big(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T)\big)^2\big|\mathcal{G}_{T-1}\big]\\
&= b_T\Big(\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\Big)^2 + 2b_T\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T)\\
&\quad + b_T\mathbf{E}\big[\big(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T)\big)^2\big|\mathcal{G}_{T-1}\big]. \qquad (4.17)
\end{aligned}$$

The second equality also uses that both the terms $\mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}))\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T)$ and $(\mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})))^2$ are in L^1 due to (4.11) and (4.14). Then we expand the last term in (4.17). Using $\triangle\mathbf{S}_T = \triangle\mathbf{M}_T + \triangle\mathbf{A}_T$, (4.11) and $\mathcal{G}_{T-1}$-measurability of $(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T))^2$, we write that last term in (4.17) as

$$\begin{aligned}
&b_T\mathbf{E}\big[\big(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T)\big)^2\big|\mathcal{G}_{T-1}\big]\\
&= b_T\mathbf{E}\big[\big(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{M}_T)\big)^2\big|\mathcal{G}_{T-1}\big] + 2b_T\mathbf{E}[\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{M}_T)\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T)|\mathcal{G}_{T-1}]\\
&\quad + b_T\big(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T)\big)^2. \qquad (4.18)
\end{aligned}$$

For the first term in (4.18), we use (4.11), then $\mathcal{G}_{T-1}$-measurability of $\boldsymbol{\delta}_T^\ell$, (4.13) and $\mathbf{E}[\triangle\mathbf{M}_T^\ell\triangle\mathbf{M}_T^m|\mathcal{G}_{T-1}] = 0$ for $\ell \neq m$ by (3.31), and finally (4.11) again to obtain

$$\begin{aligned}\mathbf{E}\big[\big(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{M}_T)\big)^2\big|\mathcal{G}_{t-1}\big] &= \sum_{\ell=1}^{L}\frac{1}{L^2}\mathbf{E}[(\boldsymbol{\delta}_T^\ell\triangle\mathbf{M}_T^\ell)^2|\mathcal{G}_{t-1}]\\ &\quad+\sum_{\ell\neq m}^{L}\frac{1}{L^2}\mathbf{E}[\boldsymbol{\delta}_T^\ell\boldsymbol{\delta}_T^m\triangle\mathbf{M}_T^\ell\triangle\mathbf{M}_T^m|\mathcal{G}_{t-1}]\\ &=\sum_{\ell=1}^{L}\frac{1}{L^2}(\boldsymbol{\delta}_T^\ell)^2\mathbf{E}[(\triangle\mathbf{M}_T^\ell)^2|\mathcal{G}_{t-1}]\\ &=\frac{1}{L}\mathrm{em}\big((\boldsymbol{\delta}_T)^2\odot\triangle\langle\mathbf{M}\rangle_T\big).\end{aligned}\tag{4.19}$$

The second equality in (4.19) also uses (4.14) and its consequence that the terms $\triangle\mathbf{M}_T^\ell\triangle\mathbf{M}_T^m$ and $\boldsymbol{\delta}_t^\ell\boldsymbol{\delta}_t^m\triangle\mathbf{M}_t^\ell\triangle\mathbf{M}_t^m$ are in L^1 for $\ell, m = 1,\dots,L$. For the second term in (4.18), we argue similarly as in the computation of the conditional expectation $\mathbf{E}[\mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}))\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T)|\mathcal{G}_{T-1}]$ in (4.17) and use $\mathbf{E}[\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{M}_T)|\mathcal{G}_{T-1}] = 0$ as argued in (4.15) to obtain

$$\begin{aligned}&\mathbf{E}[\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{M}_T)\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T)|\mathcal{G}_{T-1}]\\ &\quad=\mathbf{E}[\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{M}_T)|\mathcal{G}_{T-1}]\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T)=0.\end{aligned}\tag{4.20}$$

Inserting (4.18) with (4.19) and (4.20) into (4.17) yields

$$\begin{aligned}&b_T\mathbf{E}\big[\big(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{S}_T)\big)^2\big|\mathcal{G}_{T-1}\big]\\ &\quad=b_T\frac{1}{L}\mathrm{em}\big((\boldsymbol{\delta}_T)^2\odot\triangle\langle\mathbf{M}\rangle_T\big)+\big(\mathrm{em}(\boldsymbol{\delta}_T\odot\triangle\mathbf{A}_T)\big)^2.\end{aligned}\tag{4.21}$$

Finally, plugging (4.15), (4.16) and (4.21) back into (4.9) and reordering the terms yields (4.4) and (4.5). □

4.2 Step 2: Maximising $\boldsymbol{\delta}_T \mapsto F(\boldsymbol{\delta}_T)$

In this subsection, we implement Recipe 4.1, step 2). Let us consider the map $\boldsymbol{\delta}_T \mapsto F(\boldsymbol{\delta}_T) = E[V_T(\boldsymbol{\vartheta}(T,\boldsymbol{\delta}_T))|\mathcal{G}_{T-1}]$. We maximise F over $\boldsymbol{\delta}_T$ *without* any constraint. Formal differentiation of F with respect to $\boldsymbol{\delta}_T$ yields a (formal) first order condition (FOC) for optimality. Then solving that FOC gives a candidate $\widehat{\boldsymbol{\delta}}_T$ for the maximiser. Finally, we verify that the candidate $\widehat{\boldsymbol{\delta}}_T$ is a true maximiser for F by completing the square.

Proposition 4.4. *Suppose that Assumptions* 3.2 *and* 3.4 *are satisfied. If* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ *and* $V_T(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T))$ *is given by* (4.2)*, then a maximiser* $\widehat{\boldsymbol{\delta}}_T$ *for* $\boldsymbol{\delta}_T \mapsto F(\boldsymbol{\delta}_T)$ *is a solution to the system of linear equations*

$$\boldsymbol{\delta}_T^\ell = \frac{\boldsymbol{\lambda}_T^\ell}{(1 - L^{-1}) + \triangle \mathbf{K}_T^\ell} \bigg(\frac{a_T}{2b_T} - \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + \mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T) \bigg), \quad \ell = 1, \dots, L. \tag{4.22}$$

Moreover, a solution $\widehat{\boldsymbol{\delta}}_T$ *to* (4.22) *exists and satisfies*

$$\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) = \frac{\mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)} \odot (\frac{a_T}{2b_T} - \mathbf{G}_{T-1}(\boldsymbol{\vartheta}) + \mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}))))}{1 - \mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)})}, \tag{4.23}$$

where

$$\triangle \widetilde{\mathbf{K}}_T^{\ell,(L)} := \frac{\triangle \mathbf{K}_T^\ell}{(1 - L^{-1}) + \triangle \mathbf{K}_T^\ell}, \quad \ell = 1, \dots, L. \tag{4.24}$$

Explicitly, we have for $\ell = 1, \dots, L$ *that*

$$\widehat{\boldsymbol{\delta}}_T^\ell = \frac{\boldsymbol{\lambda}_T^\ell}{(1 - L^{-1}) + \triangle \mathbf{K}_T^\ell} \bigg(\frac{a_T}{2b_T} - \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + e \bigg), \tag{4.25}$$

where e *is the right-hand side of* (4.23)*. The solution* $\widehat{\boldsymbol{\delta}}_T$ *given in* (4.25) *is indeed a maximiser for* $\boldsymbol{\delta}_T \mapsto F(\boldsymbol{\delta}_T)$*, and the resulting* $R_T(\widehat{\boldsymbol{\delta}}_T)$ *from* (4.5) *satisfies*

$$R_T(\widehat{\boldsymbol{\delta}}_T) = \frac{a_T}{2} \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) - b_T \mathrm{ecov}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}), \widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T\big). \tag{4.26}$$

Here we use the notation ecov *defined in* (4.6).

Proof. 1) Lemma 4.2 shows that $\boldsymbol{\delta}_T \mapsto \mathbf{E}[V_T(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)) | \mathcal{G}_{T-1}]$ depends on $\boldsymbol{\delta}_T$ only through $R_T(\boldsymbol{\delta}_T)$ given in (4.5). So formally differentiating $R_T(\boldsymbol{\delta}_T)$ in (4.5) with respect to $\boldsymbol{\delta}_T = (\boldsymbol{\delta}_T^\ell)_{\ell=1,\dots,L}$ yields a formal FOC for optimality as

$$\begin{aligned}
0 &= \frac{\partial}{\partial \boldsymbol{\delta}_T^\ell} R_T(\boldsymbol{\delta}_T) \\
&= \frac{\partial}{\partial \boldsymbol{\delta}_T^\ell} \Big(a_T \mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T) - 2b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}) \odot \boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T\big) \\
&\qquad + 2b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T) \\
&\qquad - b_T \mathrm{em}\Big((\boldsymbol{\delta}_T)^2 \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\
&\qquad + b_T\big(\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T)\big)^2\Big), \quad \ell = 1,\dots,L. \qquad (4.27)
\end{aligned}$$

Computing the RHS of (4.27) explicitly and multiplying on both sides by L gives

$$\begin{aligned}
0 = {} & a_T \triangle \mathbf{A}_T^\ell - 2b_T \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta})\triangle \mathbf{A}_T^\ell + 2b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\triangle \mathbf{A}_T^\ell \\
& - 2b_T \boldsymbol{\delta}_T^\ell\big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T^\ell + (\triangle \mathbf{A}_T^\ell)^2\big) + 2b_T \mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T)\triangle \mathbf{A}_T^\ell. \qquad (4.28)
\end{aligned}$$

Moving $-2b_T \boldsymbol{\delta}_T^\ell\big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T^\ell + (\triangle \mathbf{A}_T^\ell)^2\big)$ to the other side and dividing on both sides of (4.28) by $2b_T\big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T^\ell + (\triangle \mathbf{A}_T^\ell)^2\big)$ yields (4.22).

2) We now construct a solution to (4.22). For this, we first recall from (3.35) and (3.36) that for $\ell = 1,\dots,L$,

$$\boldsymbol{\lambda}_T^\ell = \frac{\triangle \mathbf{A}_T^\ell}{\triangle\langle \mathbf{M}^\ell\rangle_T}, \quad \triangle \mathbf{K}_T^\ell = \frac{(\triangle \mathbf{A}_T^\ell)^2}{\triangle\langle \mathbf{M}^\ell\rangle_T}, \quad \boldsymbol{\lambda}_T^\ell \triangle \mathbf{A}_T^\ell = \triangle \mathbf{K}_T^\ell. \qquad (4.29)$$

We multiply (4.22) by $\triangle \mathbf{A}_T^\ell$ and use (4.29) and then average over ℓ to obtain the equation

$$\begin{aligned}
\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T) = \mathrm{em}\bigg(\triangle \widetilde{\mathbf{K}}_T^{(L)} \odot \Big(\frac{a_T}{2b_T} - \mathbf{G}_{T-1}(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \\
+ \mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T)\Big)\bigg), \qquad (4.30)
\end{aligned}$$

where $\triangle \widetilde{\mathbf{K}}_T^{(L)}$ is given in (4.24). For (4.23), moving $\mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)})\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T)$ to the left side of (4.30), we see that (4.30) has a solution given by (4.23) if and only if $1 - \mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)}) \neq 0$. But the latter is readily verified because we have $\triangle \widetilde{\mathbf{K}}_t^{\ell,(L)} < 1$ $\mathbf{P}$-a.s. for $\ell = 1,\dots,L$ due to (4.24) and our convention $\frac{0}{0} = 0$. Now we construct $\widehat{\boldsymbol{\delta}}_T$ by inserting the right side of (4.23) into (4.22) to replace the term $\mathrm{em}(\boldsymbol{\delta}_T \odot \triangle \mathbf{A}_T)$ (with $\boldsymbol{\delta}_T$ replaced by $\widehat{\boldsymbol{\delta}}_T$). This yields the explicit expression for $\widehat{\boldsymbol{\delta}}_T$ given in (4.25). To show that the constructed $\widehat{\boldsymbol{\delta}}_T$ solves (4.22), we multiply both sides of (4.25) by $\triangle \mathbf{A}_T^\ell$ and average over ℓ to obtain again (4.23). This implies that we can replace the term e in (4.25) by $\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)$, which says exactly that $\widehat{\boldsymbol{\delta}}_T$ solves (4.22).

3) Let us argue that a solution $\widehat{\boldsymbol{\delta}}_T$ to (4.22) is indeed a global maximiser for (4.1) at time $t = T$. It is sufficient to prove that $\widehat{\boldsymbol{\delta}}_T$ is a global maximiser for $\boldsymbol{\delta}_T \mapsto R_T(\boldsymbol{\delta}_T)$ with R_T given in (4.5). We show this by completing the square. Let $\boldsymbol{\delta}_T \in \boldsymbol{\Theta}_{\mathrm{S}}^{[T]}$. We write $\boldsymbol{\delta}_T = \boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T + \widehat{\boldsymbol{\delta}}_T$, insert the latter into (4.7) and expand and reorder the terms as expressions involving $\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T$ and $\widehat{\boldsymbol{\delta}}_T$ separately to obtain

$$
\begin{aligned}
&R_T(\boldsymbol{\delta}_T)\\
&= R_T(\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T + \widehat{\boldsymbol{\delta}}_T)\\
&= a_T \mathrm{em}((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T) - 2 b_T \mathrm{ecov}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}), (\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big) \qquad (4.31)\\
&\quad - b_T \mathrm{em}\Big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1 - L^{-1}) \triangle \langle \mathbf{M} \rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big)\\
&\quad + b_T \Big(\mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big)\Big)^2\\
&\quad + a_T \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) - 2 b_T \mathrm{ecov}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}), \widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T\big)\\
&\quad - b_T \mathrm{em}\Big((\widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1 - L^{-1}) \triangle \langle \mathbf{M} \rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big)\\
&\quad + b_T \Big(\mathrm{em}\big(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T\big)\Big)^2\\
&\quad - 2 b_T \mathrm{em}\Big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \widehat{\boldsymbol{\delta}}_T \odot \big((1 - L^{-1}) \triangle \langle \mathbf{M} \rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \qquad (4.32)\\
&\quad + 2 b_T \mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big) \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T). \qquad (4.33)
\end{aligned}
$$

Using (4.11) and that $\widehat{\boldsymbol{\delta}}_T$ satisfies (4.22) and

$$
\frac{\boldsymbol{\lambda}_T^\ell}{(1 - L^{-1}) + \triangle \mathbf{K}_T^\ell} \big((1 - L^{-1}) \triangle \langle \mathbf{M}^\ell \rangle_T + (\triangle \mathbf{A}_T^\ell)^2\big) = \triangle \mathbf{A}_T^\ell
$$

by (4.29), we unravel the definition of the term in (4.32) as

$$
\begin{aligned}
&2b_T \mathrm{em}\Big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \widehat{\boldsymbol{\delta}}_T \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\
&= 2b_T \frac{1}{L}\sum_{\ell=1}^{L} (\boldsymbol{\delta}_T^\ell - \widehat{\boldsymbol{\delta}}_T^\ell) \frac{\boldsymbol{\lambda}_T^\ell}{(1-L^{-1}) + \triangle \mathbf{K}_T^\ell} \\
&\qquad\qquad \times \left(\frac{a_T}{2b_T} - \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)\right) \\
&\qquad\qquad \times \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big) \\
&= 2b_T \frac{1}{L}\sum_{\ell=1}^{L} (\boldsymbol{\delta}_T^\ell - \widehat{\boldsymbol{\delta}}_T^\ell) \\
&\qquad\qquad \times \left(\frac{a_T}{2b_T} - \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)\right) \triangle \mathbf{A}_T^\ell \\
&= a_T \mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big) - 2b_T \mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta}) \odot \triangle \mathbf{A}_T\big) \\
&\quad + 2b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big) \\
&\quad + 2b_T \mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big) \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) \\
&= a_T \mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big) - 2b_T \mathrm{ecov}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}), (\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big) \\
&\quad + 2b_T \mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big) \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T).
\end{aligned}
$$

The last line uses the definition (4.6) of ecov. The above equation brings a significant cancellation in (4.33); indeed, it shows that the three lines (4.31), (4.32) and (4.33) sum up to 0. We exploit this cancellation, then reorder terms according to the involvement of $\widehat{\boldsymbol{\delta}}_T$ and $\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T$ and use the alternative expression (4.7) for $R_T(\boldsymbol{\delta}_T)$ to further simplify (4.33) to

$$
\begin{aligned}
R_T(\boldsymbol{\delta}_T) &= a_T \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) - 2b_T \mathrm{ecov}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}), \widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T\big) \\
&\quad - b_T \mathrm{em}\Big((\widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\
&\quad + b_T \Big(\mathrm{em}\big(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T\big)\Big)^2 \\
&\quad - b_T \mathrm{em}\Big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\
&\quad + b_T \Big(\mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big)\Big)^2 \\
&= R_T(\widehat{\boldsymbol{\delta}}_T) - b_T \Bigg(\mathrm{em}\Big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\
&\qquad\qquad - \Big(\mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big)\Big)^2\Bigg).
\end{aligned} \tag{4.34}
$$

By part 1), $\widehat{\boldsymbol{\delta}}_T$ also satisfies (4.28). Hence we multiply both sides of (4.28) by $\widehat{\boldsymbol{\delta}}_T^{\ell}$, average over ℓ and use the definition (4.6) of ecov to write (4.28), with $\boldsymbol{\delta}_T$ replaced by $\widehat{\boldsymbol{\delta}}_T$, as

$$\begin{aligned} 0 &= a_T \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) - 2b_T \mathrm{ecov}(\mathbf{G}_{T-1}, \widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) \\ &\quad - 2b_T \mathrm{em}\Big((\widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\ &\quad + 2b_T \Big(\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)\Big)^2. \end{aligned} \tag{4.35}$$

In view of (4.35), (4.7) becomes

$$\begin{aligned} R_T(\widehat{\boldsymbol{\delta}}_T) &= b_T \mathrm{em}\Big((\widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\ &\quad - b_T\Big(\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)\Big)^2 \qquad (4.36) \\ &= \frac{a_T}{2}\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) - b_T \mathrm{ecov}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}), \widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T\big) \\ &\quad + b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T). \qquad (4.37) \end{aligned}$$

Inserting (4.36) into (4.34) and using (4.8) gives

$$\begin{aligned} R_T(\boldsymbol{\delta}_T) &= b_T\bigg(\mathrm{em}\Big((\widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\ &\qquad - \Big(\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)\Big)^2\bigg) \\ &\quad - b_T\bigg(\mathrm{em}\Big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T)^2 \odot \big((1-L^{-1})\triangle\langle \mathbf{M}\rangle_T + (\triangle \mathbf{A}_T)^2\big)\Big) \\ &\qquad - \Big(\mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big)\Big)^2\bigg) \\ &= b_T\bigg(\mathrm{em}\big((\widehat{\boldsymbol{\delta}}_T)^2 \odot (1-L^{-1})\triangle\langle \mathbf{M}\rangle_T\big) + \mathrm{evar}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)\bigg) \\ &\quad - b_T\bigg(\mathrm{em}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T)^2 \odot (1-L^{-1})\triangle\langle \mathbf{M}\rangle_T\big) \\ &\qquad + \mathrm{evar}\big((\boldsymbol{\delta}_T - \widehat{\boldsymbol{\delta}}_T) \odot \triangle \mathbf{A}_T\big)\bigg). \end{aligned} \tag{4.38}$$

This is clearly a quadratic form in $\boldsymbol{\delta}_T$, and the sum of the last two lines of (4.38) is always nonpositive because $b_T > 0$ from (4.2), $\triangle\langle \mathbf{M}\rangle_T \geq 0$ by (4.13) and $\mathrm{evar}(\cdot) \geq 0$. So we obtain a maximum over $\boldsymbol{\delta}_T$ in (4.38) if and only if that term is 0, which happens if and only if $\boldsymbol{\delta}_T = \widehat{\boldsymbol{\delta}}_T$. Finally, (4.26) is given by (4.37). □

We end this subsection by summarising the key aspects of Recipe 4.1, steps

1) and 2) or, more precisely, Lemma 4.2 and Proposition 4.4, because similar structures will appear later recursively. First, writing (4.1) allows us to consider a single-step optimisation with one single variable (here, $\boldsymbol{\delta}_T$). Differentiating the one-step objective function with respect to that variable yields a first-order condition (4.22) which can be solved explicitly as in (4.25). Moreover, the optimiser $\widehat{\boldsymbol{\delta}}_T$ has the (feedback) form $f(T, \mathbf{G}_{T-1}(\boldsymbol{\vartheta}))$ for a function $f(T, \mathbf{x})$ (see the expression from (4.25)) which does not explicitly depend on $\boldsymbol{\vartheta}$ itself. From an abstract perspective, Lemma 4.2 and Proposition 4.4 can be viewed as a (partial) solution technique for (4.1) at the level of (4.2). The partialness is due to that we have not yet shown that the maximiser $\widehat{\boldsymbol{\delta}}_T$ for the map F lies in $\boldsymbol{\Theta}_{\mathrm{S}}^{[T]}(\boldsymbol{\vartheta})$. This is the content of Recipe 4.1, step 3) and is done in the next subsection.

4.3 Step 3: The candidate maximiser $\widehat{\boldsymbol{\delta}}_T$ is in $\boldsymbol{\Theta}_{\mathrm{S}}^{[T]}$

In this subsection, we implement Recipe 4.1, step 3) by showing that $\widehat{\boldsymbol{\delta}}_T$ is in $\boldsymbol{\Theta}_{\mathrm{S}}^{[T]}(\boldsymbol{\vartheta})$ for $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ and deriving an explicit formula for $V_{T-1}(\boldsymbol{\vartheta})$.

Proposition 4.5. *Suppose that Assumptions* 3.2 *and* 3.4 *are satisfied and that* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ *and* $V_T(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T))$ *is given by* (4.2)*. If the MVT process* K *is bounded, then* $\widehat{\boldsymbol{\delta}}_T$ *given in* (4.22) *is in* $\boldsymbol{\Theta}_{\mathrm{S}}^{[T]}(\boldsymbol{\vartheta})$ *and hence is a maximiser for* (4.1) *at time* $t = T$*. Moreover,* $V_{T-1}(\boldsymbol{\vartheta})$ *from* (4.1) *has the form*

$$V_{T-1}(\boldsymbol{\vartheta}) = a_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) - b_T \mathrm{evar}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + c_T + R_T(\widehat{\boldsymbol{\delta}}_T), \tag{4.39}$$

where $R_T(\widehat{\boldsymbol{\delta}}_T)$ *is given by* (4.23)*. Explicitly, we have*

$$\begin{aligned} R_T(\widehat{\boldsymbol{\delta}}_T) = {} & a_T \mathrm{em}\left(\frac{\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)}) - \triangle\widetilde{\mathbf{K}}_T^{(L)}}{1 - \mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})} \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})\right) \\ & + b_T \mathrm{em}\big(\triangle\widetilde{\mathbf{K}}_T^{(L)} \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})^2\big) \\ & + b_T \frac{1}{1 - \mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})} \Big(\mathrm{em}\big((1 - \triangle\widetilde{\mathbf{K}}_T^{(L)} \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\Big)^2 \\ & - b_T \Big(\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\Big)^2 + \frac{a_T^2 \mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})}{4 b_T \big(1 - \mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})\big)}, \end{aligned} \tag{4.40}$$

and hence

$$\begin{aligned} V_{T-1}(\boldsymbol{\vartheta}) = a_T \mathrm{em}\Big(&\frac{1-\triangle\widetilde{\mathbf{K}}_T^{(L)}}{1-\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})} \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})\Big) \\ &- b_T\big(1-\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})\big)\Bigg(\mathrm{em}\Big(\frac{1-\triangle\widetilde{\mathbf{K}}_T^{(L)}}{1-\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})} \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})^2\Big) \\ &\qquad\qquad - \Bigg(\mathrm{em}\Big(\frac{1-\triangle\widetilde{\mathbf{K}}_T^{(L)}}{1-\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})} \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})\Big)\Bigg)^2\Bigg) \\ &+ \frac{a_T^2 \mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})}{4b_T\big(1-\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})\big)}. \end{aligned} \tag{4.41}$$

Proof. 1) We first prove that $\widehat{\boldsymbol{\delta}}_T$ given by (4.22) is in $\boldsymbol{\Theta}_{\mathrm{S}}^{[T]}(\boldsymbol{\vartheta})$, which shows the optimality of $\widehat{\boldsymbol{\delta}}_T$ for (4.1) at time $t = T$ by Proposition 4.4. The explicit definition in (4.25) readily shows that $\widehat{\boldsymbol{\delta}}_T$ is $\mathcal{G}_{T-1}$-measurable; note that $\boldsymbol{\lambda}, \mathbf{K}$ and $\widetilde{\mathbf{K}}$ are all $\mathbb{G}$-predictable. In view of the definition (4.3) of $\boldsymbol{\Theta}_{\mathrm{S}}^{[T]}(\boldsymbol{\vartheta})$, we thus only need to show that $\widehat{\boldsymbol{\delta}}_T^\ell \triangle \mathbf{S}_T^\ell \in L^2$ for $\ell = 1, \dots, L$. To this end, we use $\mathcal{G}_{T-1}$-measurability of $\widehat{\boldsymbol{\delta}}_T^\ell$ with $(\widehat{\boldsymbol{\delta}}_T^\ell \triangle \mathbf{S}_T^\ell)^2, (\triangle \mathbf{S}_T^\ell)^2 \geq 0$, then the explicit expression for $\widehat{\boldsymbol{\delta}}_T^\ell$ from (4.22) and the angle bracket identity in (4.13) and its consequence

$$(\boldsymbol{\lambda}_T^\ell)^2\big(\triangle\langle \mathbf{M}^\ell\rangle_T + (\triangle \mathbf{A}_T^\ell)^2\big) = \triangle \mathbf{K}_T^\ell + (\triangle \mathbf{K}_T^\ell)^2$$

to obtain

$$\begin{aligned} \mathbf{E}[(\widehat{\boldsymbol{\delta}}_T^\ell \triangle \mathbf{S}_T^\ell)^2] &= \mathbf{E}\big[(\widehat{\boldsymbol{\delta}}_T^\ell)^2 \mathbf{E}[(\triangle \mathbf{S}_T^\ell)^2 | \mathcal{G}_{T-1}]\big] \\ &= \mathbf{E}\Bigg[\frac{\triangle \mathbf{K}_T^\ell + (\triangle \mathbf{K}_T^\ell)^2}{((1-L^{-1}) + \triangle \mathbf{K}_T^\ell)^2} \\ &\qquad \times \Bigg(\frac{a_T}{2b_T} - \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta}) \\ &\qquad\qquad + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)\Bigg)^2\Bigg]. \end{aligned} \tag{4.42}$$

We note in (4.42) that because $L \geq 2$ and $\triangle \mathbf{K}_T^\ell \geq 0$ by (4.29), we obtain for $\ell = 1, \dots, L$ the inequality $((1-L^{-1}) + \triangle \mathbf{K}_T^\ell)^2 \geq (1-L^{-1})^2 \geq \frac{1}{4}$. Using this lower bound and that $\triangle K_T$ and hence $\triangle \mathbf{K}_T^\ell$ is bounded (say by C), we obtain

$$\frac{\triangle \mathbf{K}_T^\ell + (\triangle \mathbf{K}_T^\ell)^2}{((1-L^{-1}) + \triangle \mathbf{K}_T^\ell)^2} \leq 4(C + C^2). \tag{4.43}$$

This bound implies that we only need to show that each term in the large round parenthesis in (4.42) is in L^2. The first three terms are evidently in L^2 by the nonrandomness of a_T and b_T, due to $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$, and by the definition of $\boldsymbol{\Theta}_{\mathrm{S}}$ in (3.29). For the last term $\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)$, by (4.11), it is sufficient to show that $\widehat{\boldsymbol{\delta}}_T^\ell \triangle \mathbf{A}_T^\ell \in L^2$ for $\ell = 1, \ldots, L$. By (4.23), we see that

$$\widehat{\boldsymbol{\delta}}_T^\ell \triangle \mathbf{A}_T^\ell = \frac{\triangle \widetilde{\mathbf{K}}_T^{\ell,(L)}}{1 - \mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)})} \left(\frac{a_T}{2 b_T} - \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \right). \tag{4.44}$$

We claim that the factor $\frac{\triangle \widetilde{\mathbf{K}}_T^{\ell,(L)}}{1 - \mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)})}$ is bounded uniformly over ℓ and $\omega \in \Omega$. Then from the explicit expression in (4.44) and the fact that each term in the parenthesis in (4.44) is in L^2, it is evident that $\widehat{\boldsymbol{\delta}}_T^\ell \triangle \mathbf{A}_T^\ell \in L^2$ for $\ell = 1, \ldots, L$. For the boundedness of $\frac{\triangle \widetilde{\mathbf{K}}_T^{\ell,(L)}}{1 - \mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)})}$, we observe by (4.24) and $\triangle \mathbf{K}_T^\ell \leq C$ that

$$1 - \triangle \widetilde{\mathbf{K}}_T^{\ell,(L)} = \frac{1 - L^{-1}}{(1 - L^{-1}) + \triangle \mathbf{K}_T^\ell} \geq \frac{1 - L^{-1}}{1 - L^{-1} + C}, \quad \ell = 1, \ldots, L. \tag{4.45}$$

Using (4.11) with $\mathrm{em}(1) = 1$ and (4.45), we obtain the two inequalities

$$\begin{aligned} \frac{1}{1 - \mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)})} &= \frac{1}{\mathrm{em}(1 - \triangle \widetilde{\mathbf{K}}_T^{(L)})} \\ &= \left(\frac{1}{L} \sum_{\ell=1}^{L} (1 - \triangle \widetilde{\mathbf{K}}_T^{\ell,(L)}) \right)^{-1} \\ &\leq \frac{1 - L^{-1} + C}{1 - L^{-1}}, \end{aligned} \tag{4.46}$$

$$\triangle \widetilde{\mathbf{K}}_T^{\ell,(L)} \leq 1 - \frac{1 - L^{-1}}{1 - L^{-1} + C} = \frac{C}{(1 - L^{-1}) + C}, \quad \ell = 1, \ldots, L. \tag{4.47}$$

Hence by combining (4.46) and (4.47), we find that

$$\frac{\triangle \widetilde{\mathbf{K}}_T^{\ell,(L)}}{1 - \mathrm{em}(\triangle \widetilde{\mathbf{K}}_T^{(L)})} \leq \frac{C}{1 - L^{-1}} \leq 2C, \quad \ell = 1, \ldots, L, L \geq 2.$$

This completes the proof that $\widehat{\boldsymbol{\delta}}_T \in \boldsymbol{\Theta}_{\mathrm{S}}^{[T]}(\boldsymbol{\vartheta})$.

2) Next, we turn to verifying (4.40). Using (4.26) with the definition (4.6) of ecov, then (4.11), (4.22) and $\frac{\boldsymbol{\lambda}_T^\ell}{(1 - L^{-1}) + \triangle \mathbf{K}_T^\ell} \triangle \mathbf{A}_T^\ell = \triangle \mathbf{K}_T^{\ell,(L)}$ by (4.29) and (4.24),

we obtain

$$
\begin{aligned}
&R_T(\widehat{\boldsymbol{\delta}}_T) \\
&= \frac{a_T}{2}\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) - b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}) \odot \widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T\big) \\
&\quad + b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) \\
&= \frac{a_T}{2}\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) \\
&\quad - \frac{b_T}{L}\sum_{\ell=1}^{L} \mathbf{G}^{\ell}_{T-1}(\boldsymbol{\vartheta}) \triangle \widetilde{\mathbf{K}}^{\ell,(L)}_T \bigg(\frac{a_T}{2b_T} - \mathbf{G}^{\ell}_{T-1}(\boldsymbol{\vartheta}) \\
&\qquad\qquad\qquad + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + \mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)\bigg) \\
&\quad + b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) \\
&= \frac{a_T}{2}\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) - \frac{a_T}{2}\mathrm{em}\big(\triangle \widetilde{\mathbf{K}}^{(L)}_T \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + b_T \mathrm{em}\big(\triangle \widetilde{\mathbf{K}}^{(L)}_T \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})^2\big) \\
&\quad - b_T \mathrm{em}\big(\triangle \widetilde{\mathbf{K}}^{(L)}_T \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \\
&\quad - b_T \mathrm{em}\big(\triangle \widetilde{\mathbf{K}}^{(L)}_T \odot \mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T) \\
&\quad + b_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T).
\end{aligned}
\tag{4.48}
$$

Then using the explicit expression for $\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle \mathbf{A}_T)$ from (4.23) in (4.48) yields (4.40) after a tedious but straightforward computation. Finally, (4.41) follows from the optimality of $\widehat{\boldsymbol{\delta}}_T$ proved in part 1) and by inserting (4.40) into (4.4). □

4.4 Step 4: Adding an assumption to allow iteration

The first three steps give a complete solution for (4.1) at time $t = T$. Naturally, we hope to extend this technique to all earlier times $t < T$. In this subsection, we first give an overview of Recipe 4.1, steps 1)–3) and argue why this solution technique cannot be directly applied. Next, we introduce an extra assumption that makes the problem (4.1) more tractable at times $t < T$, and simplify the results from previous steps accordingly.

If we look carefully at Lemma 4.2 and Propositions 4.4 and 4.5, we can see a clear structure. We start with an affine–quadratic objective

$$
\begin{aligned}
V_T(\boldsymbol{\vartheta}) &= \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) \\
&= a_T \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) - b_T \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) + c_T
\end{aligned}
\tag{4.49}
$$

with $a_T = 1$, $b_T = \xi$ and $c_T = 0$. Then we try to maximise

$$\boldsymbol{\delta}_T \mapsto \mathbf{E}\big[V_T\big(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)\big)\big|\mathcal{G}_{T-1}\big] \tag{4.50}$$

over $\boldsymbol{\delta}_T \in \boldsymbol{\Theta}_{\mathrm{S}}^{[T]}$, where $\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_T)$ is given by (3.4). It turns out, see (4.25), that the optimiser has the form

$$\widehat{\boldsymbol{\delta}}_T = f\big(T, \mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big), \tag{4.51}$$

where the function $f(T, \mathbf{x})$ is affine in $\mathbf{x} \in \mathbb{R}^L$ and $\mathrm{em}(\mathbf{x})$, and thus $\widehat{\boldsymbol{\delta}}_T$ is affine in $\mathbf{G}_{T-1}(\boldsymbol{\vartheta})$ and $\mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}))$. Using the recursion (4.1) for $(V_t(\boldsymbol{\vartheta}))_{t=0,1,\dots,T}$ and plugging that $\widehat{\boldsymbol{\delta}}_T$ into (4.50) then yields, see (4.39), that

$$V_{T-1}(\boldsymbol{\vartheta}) = \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) - \xi \mathrm{evar}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + R_T(\widehat{\boldsymbol{\delta}}_T),$$

where the term $R_T(\widehat{\boldsymbol{\delta}}_T)$, see (4.26), is an explicit affine–quadratic expression involving $\widehat{\boldsymbol{\delta}}_T$. In view of the affine structure in (4.51), we therefore obtain for $V_{T-1}(\boldsymbol{\vartheta})$ an affine–quadratic function of the variables $\mathbf{G}_{T-1}(\boldsymbol{\vartheta})$ and $\mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}))$. Comparing this to (4.49) suggests that we should be able to iterate Lemma 4.2 and Propositions 4.4 and 4.5 by simply replacing T everywhere by t.

Unfortunately, this does not work as easily as one would hope. When looking, in analogy to (4.49), at the conditional expectation $\boldsymbol{\delta}_t \mapsto \mathbf{E}[V_t(\boldsymbol{\vartheta}(T, \boldsymbol{\delta}_t))|\mathcal{G}_{t-1}]$, one can still pull out the $\mathcal{G}_{t-1}$-measurable quantity $\boldsymbol{\delta}_t$. But what remains inside the conditional expectation, see (4.41), is no longer – in contrast to the case $t = T$ in (4.49) – a simple affine–quadratic expression in $\mathbf{G}_{t-1}(\boldsymbol{\vartheta})$ and $\mathrm{em}(\mathbf{G}_{t-1}(\boldsymbol{\vartheta}))$. It also involves conditional expectations of random quantities measurable with respect to $\mathcal{G}_t, \mathcal{G}_{t+1}, \dots, \mathcal{G}_T$ in a way that stacks up recursively and hence cannot be managed in a transparent manner. At present, we can only make further progress if each $V_t(\boldsymbol{\vartheta})$ has the form

$$V_t(\boldsymbol{\vartheta}) = a_t \mathrm{em}\big(\mathbf{G}_t(\boldsymbol{\vartheta})\big) - b_t \mathrm{evar}\big(\mathbf{G}_t(\boldsymbol{\vartheta})\big) + c_t, \tag{4.52}$$

with *nonrandom* real-valued coefficients a_t, b_t, c_t and $b_t > 0$. (4.53)

If we have (4.52) and (4.53), the coefficients can then be taken out of conditional expectations which makes the computation from t to $t-1$ completely analogous to the one from T to $T-1$. Thus Recipe 4.1, steps 1)–3) or, more precisely, Lemma 4.2 and Propositions 4.4 and 4.5, can be applied to all earlier times $t < T$ with only purely mechanical modifications.

We now take a closer look at $V_{T-1}(\boldsymbol{\vartheta})$ in (4.41). A careful inspection of that expression reveals that the random vector $\frac{1-\triangle\widetilde{\mathbf{K}}_T^{(L)}}{1-\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})}$ appears in all coefficients of the terms involving $\mathbf{G}_{T-1}(\boldsymbol{\vartheta})$ and hence makes $V_{T-1}(\boldsymbol{\vartheta})$ fail to satisfy (4.52) and (4.53). We thus get a first natural condition to eliminate this unpleasant consequence.

Condition 4.6. The vector $\triangle\widetilde{\mathbf{K}}_T^{(L)}$ in (4.24) has coordinates which do not depend on ℓ, so that $\triangle\widetilde{\mathbf{K}}_T^{\ell,(L)} = \frac{\triangle K_T}{(1-L^{-1})+\triangle K_T} =: \triangle\widetilde{K}_T^{(L)}$ for $\ell = 1,\dots,L$.

If Condition 4.6 holds, we can simplify in (4.41) the terms

$$\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)}) = \triangle\widetilde{K}_T^{(L)}, \qquad \frac{1-\triangle\widetilde{\mathbf{K}}_T^{\ell,(L)}}{1-\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})} = \frac{1-\triangle\widetilde{K}_T^{(L)}}{1-\triangle\widetilde{K}_T^{(L)}} = 1, \quad \ell = 1,\dots,L.$$

Plugging these back into (4.41) and using the definition of $\triangle\widetilde{K}_T^{(L)}$ from Condition 4.6 yields

$$\begin{aligned} V_{T-1}(\boldsymbol{\vartheta}) &= a_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \\ &\quad - b_T(1-\triangle\widetilde{K}_T^{(L)})\Big(\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})^2\big) - \Big(\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\Big)^2\Big) \\ &\quad + \frac{a_T^2\triangle\widetilde{K}_T^{(L)}}{4b_T(1-\triangle\widetilde{K}_T^{(L)})} \\ &= a_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \\ &\quad - b_T\frac{1-L^{-1}}{(1-L^{-1})+\triangle K_T}\Big(\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})^2\big) - \Big(\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\Big)^2\Big) \\ &\quad + \frac{a_T^2\triangle K_T}{4b_T(1-L^{-1})}. \end{aligned} \tag{4.54}$$

This readily gives $V_{T-1}(\boldsymbol{\vartheta})$ the form (4.52) for $t = T-1$. However, like $\triangle K_T$, the coefficients can still be random so that (4.53) need not hold.

Let us look at Condition 4.6 from a different perspective. Note that the random variables $(\triangle\widetilde{\mathbf{K}}_T^{\ell,(L)})_{\ell=1,\dots,L}$ are independent because each random variable $\triangle\widetilde{\mathbf{K}}_T^{\ell,(L)}$ is $\mathcal{F}_T^\ell$-measurable by the explicit formula (3.34) for $\triangle\mathbf{K}_T$ and $(\mathcal{F}_T^\ell)_{\ell=1,\dots,L}$ are independent σ-algebras by their constructions (2.15). So Condition 4.6 implies that the random variable $\triangle\widetilde{K}_T^{(L)}$ is independent of itself and hence $\triangle\widetilde{K}_T^{(L)}$ must be deterministic. Because $\triangle\widetilde{K}_T^{(L)} = \frac{\triangle K_T}{(1-L^{-1})+\triangle K_T}$, we see that $\triangle\widetilde{K}_T^{(L)}$ is deterministic if and only if $\triangle K_T$ is deterministic and this motivates the following.

Condition 4.7. The increment $\triangle K_T$ of the MVT process of S is deterministic.

Remark 4.8. The paragraph above Condition 4.7 shows that Condition 4.6 implies Condition 4.7. By the definition $\triangle\widetilde{\mathbf{K}}_T^{\ell,(L)} = \frac{\triangle\mathbf{K}_T^\ell}{(1-L^{-1})+\triangle\mathbf{K}_T^\ell}$ in (4.24) and because $\triangle\mathbf{K}_T^\ell$ has the same distribution as $\triangle K_T$ (see Remark 3.9), we immediately see that Condition 4.7 implies Condition 4.6, and hence Conditions 4.6 and 4.7 are equivalent. We use in the sequel Condition 4.7 because this is formulated directly in terms of the original model.

We summarise below the decisive simplification of Proposition 4.5 brought by Condition 4.7.

Proposition 4.9. *Suppose that Assumptions* 3.2, 3.4 *and Condition* 4.7 *are satisfied. If* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$, *then the remainder term* $R_T(\widehat{\boldsymbol{\delta}}_T)$ *from* (4.26) *simplifies to*

$$R_T(\widehat{\boldsymbol{\delta}}_T) = \frac{a_T^2 \triangle K_T}{4b_T(1-L^{-1})} + b_T \frac{\triangle K_T}{(1-L^{-1}) + \triangle K_T} \mathrm{evar}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big), \tag{4.55}$$

and the value $V_{T-1}(\boldsymbol{\vartheta})$ *for* (4.1) *at time* $t = T$ *simplifies to*

$$\begin{aligned} V_{T-1}(\boldsymbol{\vartheta}) = {} & a_T \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) - b_T(1 - \triangle\widetilde{K}_T^{(L)}) \mathrm{evar}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \\ & + \frac{a_T^2 \triangle\widetilde{K}_T^{(L)}}{4b_T(1-\triangle\widetilde{K}_T^{(L)})}. \end{aligned} \tag{4.56}$$

In particular, $V_{T-1}(\boldsymbol{\vartheta})$ *is also of the form*

$$V_{T-1}(\boldsymbol{\vartheta}) = a_{T-1}\mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) - b_{T-1}\mathrm{evar}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + c_{T-1} \tag{4.57}$$

as in (4.52) *and* (4.53), *with the coefficients satisfying the recursive relations*

$$a_{T-1} = a_T, \tag{4.58}$$

$$b_{T-1} = b_T(1 - \triangle\widetilde{K}_T^{(L)}) = b_T \frac{1-L^{-1}}{(1-L^{-1}) + \triangle K_T}, \tag{4.59}$$

$$c_{T-1} = c_T + \frac{a_T^2 \triangle\widetilde{K}_T^{(L)}}{4b_T(1-\triangle\widetilde{K}_T^{(L)})} = c_T + \frac{a_T^2 \triangle K_T}{4b_T(1-L^{-1})}. \tag{4.60}$$

With a_{T-1}, b_{T-1} *defined in* (4.58), (4.59) *respectively, the maximiser* $\widehat{\boldsymbol{\delta}}_T$ *is given explicitly, for* $\ell = 1, \dots, L$, *by*

$$\widehat{\boldsymbol{\delta}}_T^\ell = \frac{\boldsymbol{\lambda}_T^\ell}{(1-L^{-1}) + \triangle\mathbf{K}_T^\ell} \left(\frac{a_{T-1}}{2b_{T-1}} - \mathbf{G}_{T-1}^\ell(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) \right). \tag{4.61}$$

Proof. We have already seen (4.56) in (4.54) which readily gives (4.57). Now

(4.55) follows from (4.56) and (4.4). The relations (4.58)–(4.60) can be read off directly by comparing (4.2) to (4.57). For (4.61), we first note the identity $\triangle\widetilde{\mathbf{K}}_T^{\ell,(L)} = \triangle K_T^{(L)}$ for $\ell = 1,\dots,L$, which follows from Condition 4.7 and Remark 4.8. In consequence, $\triangle\widetilde{\mathbf{K}}_T^{(L)}$ can be taken out of empirical averages, then $\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)}) = \triangle K_T^{(L)}$, and finally the additivity of the empirical average in (4.11) with $\mathrm{em}(b) = b$ for $b \in \mathbb{R}$ allows us to simplify (4.23) as

$$\begin{aligned}\mathrm{em}(\widehat{\boldsymbol{\delta}}_T \odot \triangle\mathbf{A}_T) &= \frac{\mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)} \odot (\frac{a_T}{2b_T} - \mathbf{G}_{T-1}(\boldsymbol{\vartheta}) + \mathrm{em}(\mathbf{G}_{T-1}(\boldsymbol{\vartheta}))))}{1 - \mathrm{em}(\triangle\widetilde{\mathbf{K}}_T^{(L)})}\\ &= \frac{\triangle\widetilde{K}_T^{(L)}}{1-\triangle\widetilde{K}_T^{(L)}}\mathrm{em}\Big(\frac{a_T}{2b_T} - \mathbf{G}_{T-1}(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\Big)\\ &= \frac{a_T}{2b_T}\frac{\triangle\widetilde{K}_T^{(L)}}{1-\triangle\widetilde{K}_T^{(L)}}.\end{aligned} \tag{4.62}$$

Inserting (4.62) into (4.25) and using (4.58) and (4.59) yields

$$\begin{aligned}\widehat{\boldsymbol{\delta}}_T^{\ell} &= \frac{\boldsymbol{\lambda}_T^{\ell}}{(1-L^{-1}) + \triangle\mathbf{K}_T^{\ell}}\left(\frac{a_T}{2b_T} - \mathbf{G}_{T-1}^{\ell}(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big) + \frac{a_T}{2b_T}\frac{\triangle\widetilde{K}_T^{(L)}}{1-\triangle\widetilde{K}_T^{(L)}}\right)\\ &= \frac{\boldsymbol{\lambda}_T^{\ell}}{(1-L^{-1}) + \triangle\mathbf{K}_T^{\ell}}\left(\frac{a_T}{2b_T(1-\triangle\widetilde{K}_T^{(L)})} - \mathbf{G}_{T-1}^{\ell}(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\right)\\ &= \frac{\boldsymbol{\lambda}_T^{\ell}}{(1-L^{-1}) + \triangle\mathbf{K}_T^{\ell}}\left(\frac{a_{T-1}}{2b_{T-1}} - \mathbf{G}_{T-1}^{\ell}(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{T-1}(\boldsymbol{\vartheta})\big)\right), \quad \ell = 1,\dots,L.\end{aligned}$$

□

Under the extra Condition 4.7, the results in Lemma 4.2 and Propositions 4.4, 4.5 and 4.9 can be used as a generic solution technique to be applied iteratively backward in time; the simple affine–quadratic structure (4.52) and (4.53) is passed from any t to $t-1$ provided that $\triangle K_t$ is deterministic. Therefore, we only need to state formally the entire recursive structure. This is done in the next subsection.

4.5 Complete recursion for the computation of the value process $V(\boldsymbol{\vartheta})$

In this subsection, we state our main results about the solution to (4.1). As argued in the previous subsection, we do not need to give explicit proofs to any of the results. It is enough to replace T by t in Condition 4.7, and then state the analogues of the results in Lemma 4.2 and Propositions 4.4, 4.5 and 4.9.

Lemma 4.10. *Suppose that Assumptions* 3.2 *and* 3.4 *are satisfied. Let* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ *and* $t \in \{1, \dots, T\}$. *If* $\triangle K_t$ *is deterministic and* $V_t(\boldsymbol{\vartheta})$ *is of the form* (4.52) *and* (4.53), *then the following statements hold:*

1) *We have*

$$\mathbf{E}\big[V_t\big(\boldsymbol{\vartheta}(t, \boldsymbol{\delta}_t)\big)\big|\mathcal{G}_{t-1}\big] = a_t \mathrm{em}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big) - b_t \mathrm{evar}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big) + c_t + R_t(\boldsymbol{\delta}_t), \quad (4.63)$$

where

$$\begin{aligned} R_t(\boldsymbol{\delta}_t) = {} & a_t \mathrm{em}(\boldsymbol{\delta}_t \odot \triangle \mathbf{A}_t) - 2b_t \mathrm{em}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta}) \odot \boldsymbol{\delta}_t \odot \triangle \mathbf{A}_t\big) \\ & + 2b_t \mathrm{em}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big)\mathrm{em}(\boldsymbol{\delta}_t \odot \triangle \mathbf{A}_t) \\ & - b_t \mathrm{em}\Big((\boldsymbol{\delta}_t)^2 \odot \big((1 - L^{-1})\triangle\langle \mathbf{M}\rangle_t + (\triangle \mathbf{A}_t)^2\big)\Big) \\ & + b_t\big(\mathrm{em}(\boldsymbol{\delta}_t \odot \triangle \mathbf{A}_t)\big)^2. \end{aligned} \quad (4.64)$$

2) *A maximiser* $\widehat{\boldsymbol{\delta}}_t$ *for* $\boldsymbol{\delta}_t \mapsto \mathbf{E}[V_t(\boldsymbol{\vartheta}(t, \boldsymbol{\delta}_t))|\mathcal{G}_{t-1}]$ *exists. It is given explicitly, for* $\ell = 1, \dots, L$, *by*

$$\widehat{\boldsymbol{\delta}}_t^{\ell} = \frac{\boldsymbol{\lambda}_t^{\ell}}{(1 - L^{-1}) + \triangle K_t}\bigg(\frac{1}{2b_{t-1}^{(L)}} - \mathbf{G}_{t-1}^{\ell}(\boldsymbol{\vartheta}) + \mathrm{em}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big)\bigg). \quad (4.65)$$

The remainder $R_t(\widehat{\boldsymbol{\delta}}_t)$ *satisfies*

$$R_t(\widehat{\boldsymbol{\delta}}_t) = \frac{a_t^2 \triangle \widetilde{K}_t^{(L)}}{4b_t(1 - \triangle \widetilde{K}_t^{(L)})} + b_t \triangle \widetilde{K}_t^{(L)} \mathrm{evar}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big), \quad (4.66)$$

and hence by inserting (4.66) *into* (4.63),

$$\begin{aligned} V_{t-1}(\boldsymbol{\vartheta}) = {} & a_t \mathrm{em}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big) - b_t(1 - \triangle \widetilde{K}_t^{(L)})\mathrm{evar}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big) \\ & + \frac{a_t^2 \triangle \widetilde{K}_t^{(L)}}{4b_t(1 - \triangle \widetilde{K}_t^{(L)})}. \end{aligned} \quad (4.67)$$

In particular, $V_{t-1}(\boldsymbol{\vartheta})$ *is also of the form* (4.52) *and* (4.53), *namely*

$$V_{t-1}(\boldsymbol{\vartheta}) = a_{t-1} \mathrm{em}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big) - b_{t-1} \mathrm{evar}\big(\mathbf{G}_{t-1}(\boldsymbol{\vartheta})\big) + c_{t-1} \quad (4.68)$$

with the coefficients satisfying the recursive relations

$$a_{t-1} = a_t, \tag{4.69}$$

$$b_{t-1} = b_t \frac{1 - L^{-1}}{(1 - L^{-1}) + \triangle K_t}, \tag{4.70}$$

$$c_{t-1} = c_t + \frac{a_t^2 \triangle K_t}{4b_t(1 - L^{-1})}. \tag{4.71}$$

Proof. 1) The identity (4.63) with (4.64) is obtained like Lemma 4.2, simply with T replaced by t.

2) The identities (4.66)–(4.71) are formal restatements of (4.55)–(4.60), and (4.65) corresponds to (4.61), all with T replaced by t. □

In view of Lemma 4.10, the entire structure of (4.52) is maintained if all the $\triangle K_t$ are deterministic. This motivates the following assumption.

Assumption 4.11. The mean–variance tradeoff process K is deterministic.

Combining Assumption 4.11 with Lemma 4.10, we can state the main result of this section, which is effectively just a formality. For later reference, we also reinstate the superscript $^{(L)}$ to stress the dependence of the solution on L.

Theorem 4.12. *Suppose that Assumptions* 3.2, 3.4 *and* 4.11 *are satisfied, meaning that S is square-integrable and satisfies the structure condition (SC), and the MVT process K is deterministic. Then:*

1) *For any $\boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$, the entire value process $(V_t^{(L)}(\boldsymbol{\vartheta}^{(L)}))_{t=0,1,\ldots,T}$ of* (3.5) *is of the form* (4.52) *and* (4.53), *i.e., for $t = 0, 1, \ldots, T$,*

$$V_t^{(L)}(\boldsymbol{\vartheta}^{(L)}) = a_t^{(L)} \mathrm{em}\big(\mathbf{G}_t(\boldsymbol{\vartheta}^{(L)})\big) - b_t^{(L)} \mathrm{evar}\big(\mathbf{G}_t(\boldsymbol{\vartheta}^{(L)})\big) + c_t^{(L)}, \tag{4.72}$$

with deterministic coefficients satisfying the recursions, for $t = 1, \ldots, T$,

$$a_{t-1}^{(L)} = a_t^{(L)}, \quad a_T^{(L)} = 1, \tag{4.73}$$

$$b_{t-1}^{(L)} = b_t^{(L)} \frac{1 - L^{-1}}{(1 - L^{-1}) + \triangle K_t}, \quad b_T^{(L)} = \xi, \tag{4.74}$$

$$c_{t-1}^{(L)} = c_t^{(L)} + \frac{(a_t^{(L)})^2 \triangle K_t}{4b_t^{(L)}(1 - L^{-1})}, \quad c_T^{(L)} = 0. \tag{4.75}$$

Explicitly, (4.73)–(4.75) *can also be written as*

$$a_t^{(L)} = 1, \quad t = 0, 1, \ldots, T, \tag{4.76}$$

$$b_t^{(L)} = \xi \prod_{u=t+1}^{T} \frac{1 - L^{-1}}{(1 - L^{-1}) + \triangle K_u}, \quad t = 0, 1, \ldots, T, \tag{4.77}$$

$$c_t^{(L)} = \sum_{u=t+1}^{T} \frac{\triangle K_u}{4 b_u^{(L)} (1 - L^{-1})}, \quad t = 0, 1, \ldots, T, \tag{4.78}$$

and (4.77) *can be written with the help of the stochastic exponential* $\mathcal{E}$ *as*

$$b_t^{(L)} = \xi \frac{\mathcal{E}(\frac{K}{1-L^{-1}})_t}{\mathcal{E}(\frac{K}{1-L^{-1}})_T}, \quad t = 0, 1, \ldots, T. \tag{4.79}$$

2) *For every* $\boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$, *the solution to the conditional problem* (3.1) *at time* t *is given, for* $\ell = 1, \ldots, L$, *by*

$$\begin{aligned}
\widehat{\boldsymbol{\vartheta}}_u^{\ell,(L)} &= \boldsymbol{\vartheta}_u^{(L)}, \quad u = 1, \ldots, t, \\
\widehat{\boldsymbol{\vartheta}}_u^{\ell,(L)} &= \frac{\boldsymbol{\lambda}_u^{\ell,(L)}}{(1 - L^{-1}) + \triangle K_u} \\
&\quad \times \left(\frac{1}{2 b_{u-1}^{(L)}} - \mathbf{G}_{u-1}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}) + \mathrm{em}\big(\mathbf{G}_{u-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) \right), \quad u = t+1, \ \ldots, T.
\end{aligned}$$

In particular, the solution to (3.1) *at time 0 and hence to the global problem* (2.25) *is given, for* $\ell = 1, \ldots, L$, *by*

$$\begin{aligned}
\widehat{\boldsymbol{\vartheta}}_u^{\ell,(L)} &= \frac{\boldsymbol{\lambda}_u^{\ell,(L)}}{(1 - L^{-1}) + \triangle K_u} \\
&\quad \times \left(\frac{1}{2 b_{u-1}^{(L)}} - \mathbf{G}_{u-1}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}) + \mathrm{em}\big(\mathbf{G}_{u-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) \right), \quad u = 1, \ldots, T.
\end{aligned} \tag{4.80}$$

Proof. We use in both parts the equivalence between (3.1) and (4.1) which follows from Lemma 3.1, the abstract rewriting in the discussion after Lemma 3.1 and the concrete specification in the beginning of Section 4.

1) The recursions (4.73)–(4.75) are (4.69)–(4.71) with the terminal conditions $a_T^{(L)} = 1$, $b_T^{(L)} = \xi$ and $c_T^{(L)} = 0$ from (4.1). Also (4.73)–(4.75) immediately yield the explicit expressions in (4.76)–(4.78). The expression (4.79) is a re-writing of (4.77).

2) We only need to show that $\widehat{\boldsymbol{\vartheta}}^{(L)}$ given by (4.80) is in $\boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$. The $\mathbb{G}$-pre-

dictability of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ follows from its definition (4.80) and the $\mathbb{G}$-predictability of $\boldsymbol{\lambda}$ by (3.34). To establish the integrability requirement in (3.29), we need to show that $\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^{\ell,(L)}$ is in L^2 for $t = 1, \dots, T$ and $\ell = 1, \dots, L$. Let us argue this claim inductively. In view of the definition (4.3) of $\boldsymbol{\Theta}_{\mathrm{S}}^{[t]}$, part 1) of the proof of Proposition 4.5 with T replaced by t indeed shows that if $\mathbf{G}_{t-1}^{\ell}(\boldsymbol{\vartheta}^{(L)}) \in L^2$ for $\ell = 1, \dots, L$, then $\widehat{\boldsymbol{\delta}}_t^{(L)} =: f(t, \mathbf{G}_{t-1}(\boldsymbol{\vartheta}^{(L)}))$ given in (4.65) (with the superscript $^{(L)}$ added) satisfies $\widehat{\boldsymbol{\delta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^{\ell,(L)} \in L^2$ for $\ell = 1, \dots, L$. Comparing the definitions of $\widehat{\boldsymbol{\vartheta}}_t^{(L)}$ with $\widehat{\boldsymbol{\delta}}_t^{(L)}$ in (4.80) and (4.65), respectively, gives $\widehat{\boldsymbol{\vartheta}}_t^{(L)} = f(t, \mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}))$ and hence yields the induction step for the claim. Because $\mathbf{G}_0^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}) = 0$ is obviously in L^2, the base case is also true by the previous reasoning. □

5 Taking limits and verification

Having obtained from Theorem 4.12 a complete description of a solution $\widehat{\boldsymbol{\vartheta}}^{(L)}$ to the auxiliary problem (2.25), we aim to construct an optimal strategy to the original MVPS problem (2.3) with the help of $\widehat{\boldsymbol{\vartheta}}^{(L)}$. To this end, we elaborate on the steps sketched earlier in (2.26)–(2.31) to construct an optimal strategy $\widehat{\theta}$ for the MVPS problem (2.3), or equivalently (2.28), with $\Theta = \Theta_{\mathrm{S}}$ given in (3.6).

5.1 Embedding the finite-L results

Before delving into the details in (2.26)–(2.31), we devote this subsection to some formalities for ease of notation. Readers only interested in the construction and verification of an optimal strategy to the MVPS problem (2.3) may skip this subsection and jump directly to the next subsection.

Recall that for each $L \in \mathbb{N}$ with $L \geq 2$, we solve the auxiliary problem (2.25) in the L-extended market $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)}, \mathbb{G}^{(L)}, \mathbf{S}^{(L)})$ and obtain an optimal strategy $\widehat{\boldsymbol{\vartheta}}^{(L)}$ as in (4.80). This in particular implies that the strategies $(\widehat{\boldsymbol{\vartheta}}^{(L)})_{L \in \mathbb{N}, L \geq 2}$ live on different probability spaces. However, for studying their convergence behaviour, it is more natural to lift them to one common probability space. Let us recall from the beginning of Section 2.2 the (sequence) probability space $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, \mathbf{P}^{(\infty)})$, where $\Omega^{(\infty)}$ is the infinite Cartesian product of Ω, $\mathcal{F}^{(\infty)}$ is the σ-algebra generated by all finite rectangles with $\mathcal{F}$-measurable sides, and $\mathbf{P}^{(\infty)}$ is the infinite product measure of P on $\mathcal{F}^{(\infty)}$, defined by

$$\mathbf{P}^{(\infty)}[E_{\ell_1} \times E_{\ell_2} \times \cdots \times E_{\ell_k} \times \Omega \times \cdots] = \prod_{j=1}^{k} P[E_{\ell_j}], \quad E_{\ell_j} \in \mathcal{F}, k \in \mathbb{N},$$

and $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$ for $L \in \mathbb{N}$ is constructed so that it depends only on the first L coordinates. Given a (possibly vector-valued) function $f^{(L)}$ defined on $\Omega^{(L)}$ for $L \in \mathbb{N}$, we always identify it with $F^{f^{(L)}}$ on $\Omega^{(\infty)}$ defined by

$$F^{f^{(L)}}(\omega_1, \omega_2, \dots) := f^{(L)}(\omega_1, \dots, \omega_L) \tag{5.1}$$

and write $F^{f^{(L)}}$ as $f^{(L)}$ for ease of notation. In the infinite product space context, the superscript $^{(L)}$ then means that the identification (5.1) is used, unless a different meaning is pointed out. We remark that (5.1) can also be written more formally as $F^{f^{(L)}} = f^{(L)} \circ \pi_{(L)}$ by pre-composing $f^{(L)}$ with the canonical projection $\pi_{(L)}$ of $\Omega^{(\infty)}$ onto $\Omega^{(L)}$. However, using projections adds only formality rather than clarity and hence (5.1) is preferred.

Given an $\mathbb{R}^L$-valued process $\mathbf{X}^{(L)}$ initially defined on $\Omega^{(L)}$ for $L \in \mathbb{N}$, we adopt (5.1) to lift it to $\Omega^{(\infty)}$ and then extend it to be $\mathbb{R}^\infty$-valued by setting

$$\mathbf{X}^{\ell,(L)} \equiv 0 \quad \text{for } \ell > L. \tag{5.2}$$

Finally, let us translate Section 3.3 to this infinite product space setting. Recall that the construction of $(\mathbb{F}^{\ell,(L)})_{\ell=1,\dots,L}$ and $(\mathbf{S}^{\ell,(L)})_{\ell=1,\dots,L}$ in (2.14)–(2.16) for $L = \infty$ gives now infinite sequences $(\mathbb{F}^\ell)_{\ell\in\mathbb{N}}$ and $(\mathbf{S}^\ell)_{\ell\in\mathbb{N}}$ of independent filtrations and processes, respectively. Moreover, all the $\mathbf{S}^\ell$ have the same distribution as S. For each $L \in \mathbb{N} \cup \{\infty\}$, the construction of $\mathbb{G}^{(L)}$ in (2.17) formally yields in this context a filtration, still written as $\mathbb{G}^{(L)}$, given by $\mathcal{G}_t^{(L)} = \sigma(\cup_{\ell=1}^L \mathcal{F}_t^\ell)$ for $t = 1, \dots, T$. We can repeat the construction in Section 3.3 to first obtain the Doob decomposition of $\mathbf{S}^\ell$ with respect to $\mathbb{G}^{(L)}$. A priori, this depends on L; but Lemma 3.7, 1) implies that it agrees with the Doob decomposition with respect to $\mathbb{F}^\ell$ as long as $\ell \le L$. So by $\mathbf{S}^\ell = \mathbf{S}_0 + \mathbf{M}^\ell + \mathbf{A}^\ell$, we always refer to the Doob decomposition of $\mathbf{S}^\ell$ with respect to some $\mathbb{G}^{(L)}$ with $L \ge \ell$, or, equivalently, with respect to $\mathbb{F}^\ell$. This then yields for $\ell = 1, \dots, L$ and $L \ge \ell$ that

$$\triangle \mathbf{S}_t^\ell = \triangle \mathbf{M}_t^\ell + \triangle \mathbf{A}_t^\ell, \quad \mathbf{E}^{(\infty)}[\triangle \mathbf{M}_t^\ell | \mathcal{G}_{t-1}^{(L)}] = 0, \quad \mathbf{E}^{(\infty)}[\triangle \mathbf{S}_t^\ell | \mathcal{G}_{t-1}^{(L)}] = \triangle \mathbf{A}_t^\ell. \tag{5.3}$$

Translating the strong orthogonality between $\mathbf{M}^\ell$ and $\mathbf{M}^m$ for $\ell, m = 1, \dots, L$ and $\ell \neq m$ in (3.31) into the current setup, we also have

$$\mathbf{E}^{(\infty)}[\triangle \mathbf{M}_t^\ell \triangle \mathbf{M}_t^m | \mathcal{G}_{t-1}^{(L)}] = 0, \quad \ell, m = 1, \dots, L, \ell \neq m. \tag{5.4}$$

For a square-integrable process $\mathbf{X} = (\mathbf{X}^\ell)_{\ell\in\mathbb{N}}$ such that each $\mathbf{X}^\ell$ is $\mathbb{F}^\ell$-adapted, we define and we have for $\ell \in \mathbb{N}$ and $\ell \leq L$ that

$$\triangle\langle \mathbf{X}^\ell\rangle_0 = 0, \quad \triangle\langle \mathbf{X}^\ell\rangle_t = \mathbf{E}^{(\infty)}[(\triangle \mathbf{X}_t^\ell)^2 | \mathcal{F}_{t-1}^\ell] = \mathbf{E}^{(\infty)}[(\triangle \mathbf{X}_t^\ell)^2 | \mathcal{G}_{t-1}^{(L)}]. \tag{5.5}$$

The last equality uses the $\mathbb{F}^\ell$-adaptedness of $\mathbf{X}^\ell$ and the proof of Lemma 3.7, 1) or more precisely (2.42) with $(\widetilde{\mathcal{A}}, \widetilde{\mathcal{B}}, \mathcal{A}, \mathcal{B}) = (\mathcal{F}_T^\ell, \sigma(\cup_{j\neq\ell}\mathcal{F}_T^j), \mathcal{F}_{t-1}^\ell, \sigma(\cup_{j\neq\ell}\mathcal{F}_{t-1}^j))$. This then gives the process $\langle \mathbf{M}^\ell\rangle$ and allows us to repeat (3.34) to define

$$\boldsymbol{\lambda}^\ell = \frac{\mathrm{d}\mathbf{A}^\ell}{\mathrm{d}\langle \mathbf{M}^\ell\rangle}, \quad \mathbf{K}^\ell = \int \boldsymbol{\lambda}^\ell \,\mathrm{d}\mathbf{A}^\ell, \quad \ell \in \mathbb{N}. \tag{5.6}$$

Moreover, Remark 3.9, saying that functions in the L-extended market like $\boldsymbol{\lambda}^{\ell,(L)}$ and $\mathbf{K}^{\ell,(L)}$ agree with the corresponding functions applied to the ℓ-th coordinate, more precisely translates into

$$\mathbf{Y}^\ell(\omega_1, \omega_2, \dots) = Y(\omega_\ell), \quad \ell \in \mathbb{N}, \tag{5.7}$$

for $\mathbf{Y}^\ell \in \{\mathbf{M}^\ell, \mathbf{A}^\ell, \langle \mathbf{M}^\ell\rangle, \boldsymbol{\lambda}^\ell, \mathbf{K}^\ell\}$ and $Y \in \{M, A, \langle M\rangle, \lambda, K\}$. Let $\boldsymbol{\vartheta}^{(L)}$ be any strategy in $\boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$. Applying the identification and extension from (5.1) and (5.2) to both $\boldsymbol{\vartheta}^{(L)}$ and $\mathbf{G}(\boldsymbol{\vartheta}^{(L)})$, we can write $\mathbf{G}(\boldsymbol{\vartheta}^{(L)})$ from (2.21) in the current setting as

$$\mathbf{G}^\ell(\boldsymbol{\vartheta}^{(L)}) = \int \boldsymbol{\vartheta}^{\ell,(L)} \,\mathrm{d}\mathbf{S}^\ell, \quad \ell \leq L, \qquad \mathbf{G}^\ell(\boldsymbol{\vartheta}^{(L)}) \equiv 0, \quad \ell > L. \tag{5.8}$$

In particular, because $\boldsymbol{\vartheta}^{\ell,(L)} \equiv 0$ for $\ell > L$ by (5.2), the definition (5.8) can be consistently written as

$$\mathbf{G}^\ell(\boldsymbol{\vartheta}^{(L)}) = \int \boldsymbol{\vartheta}^{\ell,(L)} \,\mathrm{d}\mathbf{S}^\ell, \quad \ell \in \mathbb{N}. \tag{5.9}$$

For $\mathbf{x}^{(L)} \in \mathbb{R}^\infty$ with $\mathbf{x}^{\ell,(L)} = 0$ for $\ell > L$, we still write

$$\mathrm{em}(\mathbf{x}^{(L)}) = \frac{1}{L}\sum_{\ell=1}^{L} \frac{\mathbf{x}^{\ell,(L)}}{L}. \tag{5.10}$$

With the above preparation, we can write $\widehat{\boldsymbol{\vartheta}}^{(L)}$ from (4.80) in this setup as

$$\begin{aligned} \widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} &= \frac{\boldsymbol{\lambda}_t^\ell}{(1 - L^{-1}) + \triangle K_t}\left(\frac{1}{2b_{t-1}^{(L)}} - \mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) + \mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\right) \\ &\qquad \text{for } \ell = 1, \dots, L, t = 1, \dots, T, \\ \widehat{\boldsymbol{\vartheta}}^{\ell,(L)} &\equiv 0 \quad \text{for } \ell > L. \end{aligned} \tag{5.11}$$

The first formula of (5.11) uses (4.80), (5.1), (5.9) and (5.10). It also uses (5.7) and (5.1) to obtain $\boldsymbol{\lambda}^{\ell,(L)}(\omega^{(L)}) = \lambda(\omega_\ell) = \boldsymbol{\lambda}^\ell$ for $\ell = 1, \dots, L$. The second formula of (5.11) uses (5.2).

5.2 Verification – preparation

In the next two subsections, we construct an optimal strategy $\widehat{\theta}$ for the MVPS problem (2.28). Recall from (2.26) and (2.27) with $\Theta = \Theta_{\mathrm{S}}$ and $\boldsymbol{\Theta}^{(L)} = \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$ that

$$J_T^{\mathrm{mv}}(\theta) = G_T(\theta) - \xi\big(G_T(\theta) - E[G_T(\theta)]\big)^2, \quad \theta \in \Theta_{\mathrm{S}}, \tag{5.12}$$

$$J_T^{(L)}(\boldsymbol{\vartheta}^{(L)}) = \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big), \quad \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}, \tag{5.13}$$

respectively. Note that (5.12) and (5.13) are defined on the spaces $(\Omega, \mathcal{F}, P)$ and $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$, respectively. However, all the quantities involving L are lifted to $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, \mathbf{P}^{(\infty)})$ as discussed in Section 5.1 and can equivalently be considered there. Also recall from (2.28) and (2.29) that we can equivalently write the MVPS problem (2.3) and the auxiliary problem (2.25) as

$$\text{maximise } E[J_T^{\mathrm{mv}}(\theta)] \text{ over all } \theta \in \Theta_{\mathrm{S}}, \tag{5.14}$$

$$\text{maximise } \mathbf{E}^{(L)}[J_T^{(L)}(\boldsymbol{\vartheta}^{(L)})] \text{ over all } \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}, \tag{5.15}$$

respectively. We summarise below the programme sketched in (2.26)–(2.31).

Recipe 5.1. 1) Let $L \in \mathbb{N}$. For X on $(\Omega, \mathcal{F}, P)$, we recall from (2.18) that $X^{\ell,\otimes L}(\omega^{(L)}) = X \circ \pi_{\ell,L}(\omega^{(L)}) = X(\omega_\ell)$ for $\ell = 1, \dots, L$, where $\pi_{\ell,L}$ is the canonical projection of $\Omega^{(L)}$ onto its ℓ-th coordinate. Note that $(X^{\ell,\otimes L})_{\ell=1,\dots,L}$ are always (under $\mathbf{P}^{(L)}$) independent and have the same distribution as X because of the identity $P = \mathbf{P}^{(L)} \circ \pi_{\ell,L}^{-1}$ by (2.12). Applying this lifting technique to θ and $G(\theta)$ gives two processes $\theta^{\otimes L}$ and $\mathbf{G}(\theta^{\otimes L})$ both with i.i.d. coordinates. Moreover, this technique maps Θ_{S} into a subset of $\boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$.

2) Given $\theta \in \Theta_{\mathrm{S}}$, Step 1) or, more precisely, the i.i.d. property of the coordinates of $\mathbf{G}_T(\theta^{\otimes L}) \in \mathbf{G}_T(\boldsymbol{\Theta}_{\mathrm{S}}^{(L)})$, allows us to prove a form of law of large numbers (LLN). Because $J_T^{(L)}$ from (5.13) involves empirical averages and variances, the LLN and the optimality of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for the auxiliary problem (5.15) obtained in Theorem 4.12 then yield

$$E[J_T^{\mathrm{mv}}(\theta)] = \lim_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\theta^{\otimes L})] \le \limsup_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})]. \tag{5.16}$$

Based on (5.16), we get an abstract verification result stating that $\widehat{\theta} \in \Theta_{\mathrm{S}}$ is optimal for the original MVPS problem (5.14) *if* $\mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] \to E[J_T^{\mathrm{mv}}(\widehat{\theta})]$ as $L \to \infty$.

3) We construct $\widehat{\theta} \in \Theta_{\mathrm{S}}$ explicitly such that the last condition in step 2) is satisfied.

We close this subsection with two preparatory results.

Lemma 5.2. *Let $L \in \mathbb{N}$ and X be a random quantity defined on $(\Omega, \mathcal{F}, P)$. Then $(X^{\ell,\otimes L})_{\ell=1,\dots,L}$ defined on $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$ is a family of independent random variables each of which has the same distribution as X. Moreover, if $\theta \in \Theta_{\mathrm{S}}$ (resp. $g \in G_T(\Theta_{\mathrm{S}})$), then the corresponding $\theta^{\otimes L}$ is in $\boldsymbol{\Theta}_{\mathrm{S}}$ (resp. $g^{\otimes L}$ is in $\mathbf{G}_T(\boldsymbol{\Theta}_{\mathrm{S}}^{(L)})$), and $g^{\otimes L} = \mathbf{G}_T(\theta^{\otimes L})$.*

Proof. Let $L \in \mathbb{N}$. For each $\ell = 1, \dots, L$, we recall from (2.18) that

$$X^{\ell,\otimes L} = X \circ \pi_{\ell,L}, \quad \ell = 1, \dots, L, \tag{5.17}$$

where $\pi_{\ell,L}$ is the canonical projection of $\Omega^{(L)}$ onto its ℓ-th coordinate. In view of $P = \mathbf{P}^{(L)} \circ \pi_{\ell,L}^{-1}$ from (2.12), we immediately obtain that $(X^{\ell,\otimes L})_{\ell=1,\dots,L}$ are under $\mathbf{P}^{(L)}$ independent and have the same distribution as X under P.

Next assume that θ is in Θ_{S} given in (3.6). This means that θ is $\mathbb{F}$-predictable and $\theta_t \triangle S_t \in L^2$ for $t = 1, \dots, T$. To show $\theta^{\otimes L}$ is in $\boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$ given in (3.29), we claim that $\theta^{\ell,\otimes L}$ is $\mathbb{F}^{\ell,(L)}$-predictable (and hence $\mathbb{G}^{(L)}$-predictable) and $\theta_t^{\ell,\otimes L} \triangle \mathbf{S}_t^{\ell,(L)}$ is in L^2 for $\ell = 1, \dots, L$ and $t = 1, \dots, T$. Indeed, both the measurability and integrability properties are carried over from θ to $\theta^{\otimes L}$ via (5.17) because of the equalities $\mathcal{F}_t^{\ell,(L)} = \{\pi_{\ell,L}^{-1}(E) : E \in \mathcal{F}_t\}$ and $P = \mathbf{P}^{(L)} \circ \pi_{\ell,L}^{-1}$ from (2.14) and (2.12), respectively.

Finally, let $g = G_T(\theta)$. To show that $g^{\otimes L}$ is in $\mathbf{G}_T(\boldsymbol{\Theta}_{\mathrm{S}}^{(L)})$, we claim that $g^{\ell,\otimes L}$ is equal to $\mathbf{G}_T^{\ell}(\theta^{\otimes L})$ for $\ell = 1, \dots, L$. This then yields the assertion because we know that $\theta^{\otimes L} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$ from the previous paragraph. Writing (5.17) as $X^{\ell,\otimes L}(\omega_1, \dots, \omega_L) = X(\omega_\ell)$ for $\ell = 1, \dots, L$ and using the definitions of $G(\theta)$ and $\mathbf{G}^{\ell}(\theta^{\otimes L})$ from (2.1) and (2.21) respectively, we obtain

$$\begin{aligned} g^{\ell,\otimes L}(\omega_1, \dots, \omega_L) &= G_T(\theta)(\omega_\ell) \\ &= \left(\int_0^T \theta \, \mathrm{d}S \right)(\omega_\ell) \\ &= \left(\int_0^T \theta^{\ell,\otimes L} \, \mathrm{d}\mathbf{S}^{\ell,(L)} \right)(\omega_1, \dots, \omega_L) = \mathbf{G}_T^{\ell}(\theta^{\otimes L})(\omega_1, \dots, \omega_L) \end{aligned}$$

for $\ell = 1, \ldots, L$. This establishes the claim and hence completes the proof. $\square$

Proposition 5.3. *If there exists* $\widehat{\theta} \in \Theta_{\mathrm{S}}$ *such that* $\mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] \to E[J_T^{\mathrm{mv}}(\widehat{\theta})]$ *as* $L \to \infty$*, then* $G_T(\widehat{\theta})$ *is the optimal final gain for* (5.14) *and* $\widehat{\theta}$ *is an optimal strategy for* (5.14)*.*

Proof. Let $\theta \in \Theta_{\mathrm{S}}$, $g = G_T(\theta)$ and $L \in \mathbb{N}$. By Lemma 5.2, the random variable $g^{\otimes L}$ is in $\mathbf{G}_T(\boldsymbol{\Theta}_{\mathrm{S}}^{(L)})$ and $(g^{\ell,\otimes L})_{\ell=1,\ldots,L}$ is a family of independent random variables each having the same distribution as g. Next, we establish a version of the law of large numbers. Let us recall from (2.23) that

$$\mathrm{em}(\mathbf{x}^{(L)}) = \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{x}^{\ell,(L)}. \tag{5.18}$$

Because $(g^{\ell,\otimes L})_{\ell=1,\ldots,L}$ are identically distributed according to g (we write this as $g^{\ell,\otimes L} \overset{d}{=} g$), we use (5.18) in the next two lines to obtain

$$\mathbf{E}^{(L)}[\mathrm{em}(g^{\otimes L})] = \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{E}^{(L)}[g^{\ell,\otimes L}] = \frac{1}{L}\sum_{\ell=1}^{L} E[g] = E[g] \longrightarrow E[g], \tag{5.19}$$

$$\begin{aligned}\mathbf{E}^{(L)}\big[\mathrm{em}\big((g^{\otimes L})^2\big)\big] &= \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{E}^{(L)}[(g^{\ell,\otimes L})^2] \\ &= \frac{1}{L}\sum_{\ell=1}^{L} E[g^2] = E[g^2] \longrightarrow E[g^2]\end{aligned} \tag{5.20}$$

as $L \to \infty$. Similarly, we use (5.18), then that $(g^{\ell,\otimes L})_{\ell=1,\ldots,L}$ are $\mathbf{P}^{(L)}$-independent and finally $g^{\ell,\otimes L} \overset{d}{=} g$ for $\ell = 1, \ldots, L$ to obtain

$$\begin{aligned}\mathbf{E}^{(L)}\big[\big(\mathrm{em}(g^{\otimes L})\big)^2\big] &= \frac{1}{L^2}\sum_{\ell=1}^{L} \mathbf{E}^{(L)}[(g^{\ell,\otimes L})^2] + \frac{1}{L^2}\sum_{\ell\neq m} \mathbf{E}^{(L)}[g^{\ell,\otimes L} g^{m,\otimes L}] \\ &= \frac{1}{L^2}\sum_{\ell=1}^{L} \mathbf{E}^{(L)}[(g^{\ell,\otimes L})^2] + \frac{1}{L^2}\sum_{\ell\neq m} \mathbf{E}^{(L)}[g^{\ell,\otimes L}]\mathbf{E}^{(L)}[g^{m,\otimes L}] \\ &= \frac{1}{L}E[g^2] + \frac{L^2 - L}{L^2}(E[g])^2 \longrightarrow (E[g])^2 \quad \text{as } L \to \infty.\end{aligned} \tag{5.21}$$

Using (5.12) and (5.13), we can write (5.19)–(5.21) more compactly as

$$E[J_T^{\mathrm{mv}}(\theta)] = \lim_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\theta^{\otimes L})]. \tag{5.22}$$

But for every $L \geq 2$, we get by the optimality of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ from Theorem 4.12 that

$$\mathbf{E}^{(L)}[J_T^{(L)}(\theta^{\otimes L})] \leq \mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})]. \tag{5.23}$$

Now by assumption, we have $\mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] \to E[J_T^{\mathrm{mv}}(\widehat{\theta})]$ for some $\widehat{\theta} \in \Theta_{\mathrm{S}}$. Using (5.22), (5.23) and the last property, we get for all $\theta \in \Theta_{\mathrm{S}}$ that

$$E[J_T^{\mathrm{mv}}(\theta)] = \lim_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\theta^{\otimes L})] \leq \lim_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] = E[J_T^{\mathrm{mv}}(\widehat{\theta})].$$

This proves the proposition. □

5.3 Verification – construction

In view of Proposition 5.3, the main task is to implement Recipe 5.1, 3) by showing that there exists a strategy $\widehat{\theta} \in \Theta_{\mathrm{S}}$ satisfying the assumptions of Proposition 5.3. This can be done in two steps:

Recipe 5.4. 1) Using the explicit formula for $\widehat{\boldsymbol{\vartheta}}^{(L)}$ from (4.80) and taking limits formally as $L \to \infty$, we construct a candidate $\widehat{\theta} \in \Theta_{\mathrm{S}}$ according to the limiting formula of $\widehat{\boldsymbol{\vartheta}}^{(L)}$.

2) For $L \geq 2$, we combine $J_T^{(L)} = V_T^{(L)}$ from (2.36), that $V^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})$ is martingale by the martingale optimality principle (see Lemmas 2.9 and 3.1) and $V_0(\widehat{\boldsymbol{\vartheta}}^{(L)}) = c_0^{(L)}$ by the explicit expression for the process $V^{(L)}(\boldsymbol{\vartheta}^{(L)})$ in (4.72) to obtain $\mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \mathbf{E}^{(L)}[V_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] = V_0^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)}) = c_0^{(L)}$. In view of this equality and Proposition 5.3, we then only need to show that we have $\lim_{L\to\infty} c_0^{(L)} = E[J_T^{\mathrm{mv}}(\widehat{\theta})]$, where $\widehat{\theta}$ is from step 1).

Let us implement Recipe 5.4. First, recall from (4.73)–(4.75) that

$$\begin{aligned}
a_{t-1}^{(L)} &= a_t^{(L)}, \quad a_T^{(L)} = 1,\\
b_{t-1}^{(L)} &= b_t^{(L)} \frac{1-L^{-1}}{(1-L^{-1}) + \triangle K_t}, \quad b_T^{(L)} = \xi,\\
c_{t-1}^{(L)} &= c_t^{(L)} + \frac{(a_t^{(L)})^2 \triangle K_t}{4 b_t^{(L)} (1-L^{-1})}, \quad c_T^{(L)} = 0.
\end{aligned}$$

Sending $L \to \infty$ formally in the above three recursions gives the recursions,

$$a_{t-1}^{(\infty)} = a_t^{(\infty)}, \quad a_T^{(\infty)} = 1, \tag{5.24}$$

$$b_{t-1}^{(\infty)} = b_t^{(\infty)} \frac{1}{1+\triangle K_t}, \quad b_T^{(\infty)} = \xi, \tag{5.25}$$

$$c_{t-1}^{(\infty)} = c_t^{(\infty)} + \frac{(a_t^{(\infty)})^2 \triangle K_t}{4b_t^{(\infty)}} = c_t^{(\infty)} + \frac{\triangle K_t}{4b_t^{(\infty)}}, \quad c_T^{(\infty)} = 0, \tag{5.26}$$

or explicitly, either formally sending $L \to \infty$ in the explicit formulas for $a^{(L)}$, $b^{(L)}$, $c^{(L)}$ in (4.76)–(4.78) or directly solving (5.24)–(5.26),

$$a_t^{(\infty)} = 1, \qquad t = 0, 1, \ldots, T, \tag{5.27}$$

$$b_t^{(\infty)} = \xi \prod_{u=t+1}^{T} \frac{1}{1+\triangle K_u}, \qquad t = 0, 1, \ldots, T, \tag{5.28}$$

$$c_t^{(\infty)} = \sum_{u=t+1}^{T} \frac{\triangle K_u}{4b_u^{(\infty)}}, \qquad t = 0, 1, \ldots, T. \tag{5.29}$$

From the explicit expressions in (4.76)–(4.78) and (5.27)–(5.29), it is easy to see that

$$x_t^{(L)} \longrightarrow x_t^{(\infty)} \quad \text{as } L \to \infty \text{ for } x \in \{a, b, c\}. \tag{5.30}$$

Indeed, $a_t^{(L)} = 1$ and $c_t^{(L)} = \sum_{u=t+1}^{T} \frac{\triangle K_u}{4b_u^{(L)}(1-L^{-1})}$ for $t = 0, 1, \ldots, T$ by (4.76) and (4.78). By (5.27), we have $a_t^{(L)} = a_t^{(\infty)} = 1$ for $t = 0, 1, \ldots, T$. By (5.29), we have $c_t^{(L)} \to c_t^{(\infty)}$ provided that $b_t^{(L)} \to b_t^{(\infty)}$ as $L \to \infty$ for $t = 0, 1, \ldots, T$. The latter convergence can also be directly read off from the formulas in (4.77) and (5.28) because

$$\begin{aligned} b_t^{(L)} &= \xi \prod_{u=t+1}^{T} \frac{1-L^{-1}}{(1-L^{-1})+\triangle K_u} \\ &\longrightarrow \xi \prod_{u=t+1}^{T} \frac{1}{1+\triangle K_u} = b_t^{(\infty)} \quad \text{for } t = 0, 1, \ldots, T, \text{ as } L \to \infty. \end{aligned} \tag{5.31}$$

This completes the justification of the claim (5.30). In particular, we note that $\lim_{L\to\infty} c_0^{(L)} = c_0^{(\infty)}$.

Note at this point that the only formal aspect up to here is the "derivation" of (5.24)–(5.26) from (4.73)–(4.75). Once we have guessed how the "limit quantities" $a^{(\infty)}, b^{(\infty)}, c^{(\infty)}$ should behave, we can (and did) rigorously solve for them and deduce the convergence in (5.30).

Next we try to construct a candidate for the MVPS problem (5.14). Recall from (4.80) that the optimal strategy for the L-extended problem is given by

$$\begin{aligned}\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} &= \frac{\boldsymbol{\lambda}_t^{\ell,(L)}}{(1-L^{-1})+\triangle K_t}\\ &\quad\times\left(\frac{1}{2b_{t-1}^{(L)}} - \mathbf{G}_{t-1}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}) + \mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\right), \qquad \ell=1,\dots,L, t=1,\dots,T.\end{aligned}$$

Fix $\ell \le L$. Let us formally analyse the limiting behaviour of the above expression as $L\to\infty$. For the factor in the first line, we recall from (3.35) the explicit expression $\boldsymbol{\lambda}_t^{\ell,(L)} = \frac{\triangle \mathbf{A}_t^{\ell,(L)}}{\triangle\langle \mathbf{M}^{\ell,(L)}\rangle_t}$. But by Remark 3.9, we have $\mathbf{A}_t^{\ell,(L)}(\omega^{(L)}) = A_t(\omega_\ell)$ and $\triangle\langle \mathbf{M}^{\ell,(L)}\rangle_t(\omega^{(L)}) = \langle M\rangle_t(\omega_\ell)$; hence their dependence on L is artificial and the limit of the first factor formally reads

$$\frac{\boldsymbol{\lambda}_t^{\ell,(L)}(\omega_1,\dots,\omega_L)}{(1-L^{-1})+\triangle K_t} \longrightarrow \frac{\lambda_t(\omega_\ell)}{1+\triangle K_t} \quad \text{as } L\to\infty.$$

For the factor in the second line, the convergence $\frac{1}{2b_{t-1}^{(L)}} \to \frac{1}{2b_{t-1}^{(\infty)}}$ is evident from (5.31). We then expect that the empirical mean converges to the expectation, say e_{t-1}, as $L\to\infty$ by a law of large numbers effect. Suppose this is true and denote the formal limit of $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ by $\widehat{\theta}^{\ell,\otimes\infty}$. Because the dependence on other coordinates via the empirical average disappears in the limit, we expect that $\widehat{\theta}^{\ell,\otimes\infty}$ and hence $\mathbf{G}^{\ell}(\widehat{\theta}^{\otimes\infty})$ depend only on ω_ℓ. The symmetry of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ among $\ell=1,\dots,L$ also suggests that the expectation e_{t-1} does not depend on ℓ and is equal to the expectation of $\mathbf{G}_{t-1}^{\ell}(\widehat{\theta}^{\otimes\infty})$ for any $\ell=1,\dots,L$. Summarising the above analysis and using that $\widehat{\theta}^{\ell,\otimes\infty}$ can be written as a function of ω_ℓ so that we replace ω_ℓ by ω motivates us to consider in the original space $(\Omega,\mathcal{F},P)$ the candidate

$$\widehat{\theta}_t := \frac{\lambda_t}{1+\triangle K_t}\left(\frac{1}{2b_{t-1}^{(\infty)}} - G_{t-1}(\widehat{\theta}) + E[G_{t-1}(\widehat{\theta})]\right), \quad t=1,\dots,T. \tag{5.32}$$

Note that because $G_{t-1}(\widehat{\theta})$ depends on $\widehat{\theta}_1,\dots,\widehat{\theta}_{t-1}$ but not on $\widehat{\theta}_t$, (5.32) gives a well-defined predictable process $\widehat{\theta}$. To finish the implementation of Recipe 5.4, 1), we now argue that $\widehat{\theta}$ from above belongs to Θ_{S}. We frequently use below the identities, resulting from the explicit expressions $\lambda_t = \frac{\triangle A_t}{\triangle\langle M\rangle_t}$, $\triangle K_t = \frac{(\triangle A_t)^2}{\triangle\langle M\rangle_t}$ and $\triangle\langle S\rangle_t = \triangle\langle M\rangle_t + (\triangle A_t)^2$ in (3.20) and (3.10), that

$$\lambda_t\triangle A_t = \triangle K_t, \quad (\lambda_t)^2\triangle\langle S\rangle_t = \triangle K_t + (\triangle K_t)^2, \qquad t=1,\dots,T. \tag{5.33}$$

Lemma 5.5. *Suppose that Assumptions* 3.2 *and* 3.4 *are satisfied. If the MVT process K is bounded, then $\widehat{\theta}$ from* (5.32) *is in* Θ_{S}.

Proof. From its recursive definition, the process $\widehat{\theta}$ is clearly $\mathbb{F}$-predictable. In view of the definition of Θ_{S} from (3.6), we only need to show that $\widehat{\theta}_t \triangle S_t \in L^2$ for $t = 1, \dots, T$. To this end, we use first (5.32) and $\triangle\langle S\rangle_t = \triangle\langle M\rangle_t + (\triangle A_t)^2$ from (3.10), then the Cauchy–Schwarz inequality and (5.33), and finally that $\triangle K_t$ is bounded (say by C) and hence $\frac{\triangle K_t + (\triangle K_t)^2}{(1+\triangle K_t)^2} = \frac{\triangle K_t}{1+\triangle K_t} \le C$ as well as Jensen's inequality to get for $t = 1, \dots, T$ that

$$\begin{aligned} E[(\widehat{\theta}_t \triangle S_t)^2] &= E\bigg[\Big(\frac{\lambda_t}{1+\triangle K_t}\Big)^2\Big(\frac{1}{2b^{(\infty)}_{t-1}} - G_{t-1}(\widehat{\theta}) + E[G_{t-1}(\widehat{\theta})]\Big)^2 \triangle\langle S\rangle_t\bigg] \\ &\le 3E\bigg[\frac{\triangle K_t + (\triangle K_t)^2}{(1+\triangle K_t)^2}\bigg(\Big(\frac{1}{2b^{(\infty)}_{t-1}}\Big)^2 + \big(G_{t-1}(\widehat{\theta})\big)^2 + \big(E[G_{t-1}(\widehat{\theta})]\big)^2\bigg)\bigg] \\ &\le 3CE\bigg[\Big(\frac{1}{2b^{(\infty)}_{t-1}}\Big)^2 + 2\big(G_{t-1}(\widehat{\theta})\big)^2\bigg]. \end{aligned} \tag{5.34}$$

Using the explicit formula (5.28) for $b^{(\infty)}_t$ and that K is increasing and bounded, we also obtain that

$$0 \le \frac{1}{2b^{(\infty)}_{t-1}} = \frac{1}{2\xi}\prod_{u=t}^{T}(1+\triangle K_u) \le \frac{1}{2\xi}\prod_{u=1}^{T}(1+\triangle K_u) \le \frac{1}{2\xi}(1+K_T)^T$$

is bounded. This implies that with a new constant, we can rewrite (5.34) as

$$E[(\widehat{\theta}_t \triangle S_t)^2] \le C + CE\big[\big(G_{t-1}(\widehat{\theta})\big)^2\big], \quad t = 1, \dots, T. \tag{5.35}$$

The claim $\widehat{\theta}_t \triangle S_t \in L^2$ for $t = 1, \dots, T$ now follows easily by induction. Indeed, setting $t = 1$ in (5.35) and using $G_0(\widehat{\theta}) = 0$ gives the base case. The induction step follows directly from (5.35). □

Lemma 5.5 finishes the implementation of Recipe 5.4, 1). The next step is to show that

$$c^{(\infty)}_0 = E[J^{\mathrm{mv}}_T(\widehat{\theta})] = E[G_T(\widehat{\theta})] - \xi \mathrm{Var}[G_T(\widehat{\theta})]. \tag{5.36}$$

To this end, we define a sequence $(\widetilde{V}_t)_{t=0,1,\dots,T}$ by

$$\widetilde{V}_t := a^{(\infty)}_t E[G_t(\widehat{\theta})] - b^{(\infty)}_t \mathrm{Var}[G_t(\widehat{\theta})] + c^{(\infty)}_t, \quad t = 0, 1, \dots, T. \tag{5.37}$$

Lemma 5.6. *Suppose that Assumptions* 3.2, 3.4 *and* 4.11 *are satisfied. Then the* *sequence* $(\widetilde{V}_t)_{t=0,1,\dots,T}$ *given in* (5.37) *is constant.*

Proof. We first find recursive formulas for the quantities $(E[G_t(\widehat{\theta})])_{t=0,1,\dots,T}$ and $(\mathrm{Var}[G_t(\widehat{\theta})])_{t=0,1,\dots,T}$. Let us recall that Assumption 3.2 says that $\triangle S_t \in L^2$. Lemma 5.5 shows that $\widehat{\theta} \in \Theta_{\mathrm{S}}$, and then the definition of Θ_{S} in (3.6) implies that $\widehat{\theta}_t \triangle S_t$ and $G_{t-1}(\widehat{\theta})$ are in L^2. We summarise these properties as

$$\triangle S_t, \widehat{\theta}_t \triangle S_t, G_{t-1}(\widehat{\theta}) \text{ are all in } L^2. \tag{5.38}$$

Conditioning on $\mathcal{F}_{t-1}$ with (5.38), then using the explicit expression (5.32) for $\widehat{\theta}_t$ and $\lambda_t \triangle A_t = \triangle K_t$ from (5.33) and finally observing that $\triangle K_t$ is deterministic and so is thus $\frac{1}{2b^{(\infty)}_{t-1}}$, we get

$$\begin{aligned} E[\widehat{\theta}_t \triangle S_t] &= E[\widehat{\theta}_t \triangle A_t] \\ &= E\left[\frac{\triangle K_t}{1+\triangle K_t}\left(\frac{1}{2b^{(\infty)}_{t-1}} - G_{t-1}(\widehat{\theta}) + E[G_{t-1}(\widehat{\theta})]\right)\right] \\ &= \frac{\triangle K_t}{1+\triangle K_t}\frac{1}{2b^{(\infty)}_{t-1}}, \quad t = 1, \dots, T, \end{aligned} \tag{5.39}$$

and hence

$$E[G_t(\widehat{\theta})] = E[G_{t-1}(\widehat{\theta})] + \frac{\triangle K_t}{1+\triangle K_t}\frac{1}{2b^{(\infty)}_{t-1}}, \quad t = 1, \dots, T. \tag{5.40}$$

The recursion for the variance term $(\mathrm{Var}[G_t(\widehat{\theta})])_{t=0,1,\dots,T}$ is found to satisfy, for $t = 1, \dots, T$, that

$$\mathrm{Var}[G_t(\widehat{\theta})] = \left(1 - \frac{\triangle K_t}{1+\triangle K_t}\right)\mathrm{Var}[G_{t-1}(\widehat{\theta})] + \frac{\triangle K_t}{(1+\triangle K_t)^2}\left(\frac{1}{2b^{(\infty)}_{t-1}}\right)^2, \tag{5.41}$$

which will be verified later. To show the main assertion, we now use (5.41) and the recursion (5.25) for $b^{(\infty)}$ to obtain

$$\begin{aligned} b^{(\infty)}_t \mathrm{Var}[G_t(\widehat{\theta})] &= b^{(\infty)}_t \frac{1}{1+\triangle K_t}\mathrm{Var}[G_{t-1}(\widehat{\theta})] + b^{(\infty)}_t \frac{\triangle K_t}{(1+\triangle K_t)^2}\left(\frac{1}{2b^{(\infty)}_{t-1}}\right)^2 \\ &= b^{(\infty)}_{t-1}\mathrm{Var}[G_{t-1}(\widehat{\theta})] + \frac{\triangle K_t}{4b^{(\infty)}_{t-1}(1+\triangle K_t)}, \quad t = 1, \dots, T. \end{aligned} \tag{5.42}$$

Inserting (5.40) and (5.42) into (5.37), then using the recursions (5.25), (5.26) for

$b^{(\infty)}, c^{(\infty)}$ and (5.37) gives

$$\begin{aligned}
\widetilde{V}_t &= E[G_{t-1}(\widehat{\theta})] + \frac{\triangle K_t}{1+\triangle K_t}\frac{1}{2b_{t-1}^{(\infty)}} - b_{t-1}^{(\infty)}\mathrm{Var}[G_{t-1}(\widehat{\theta})] - \frac{\triangle K_t}{4b_{t-1}^{(\infty)}(1+\triangle K_t)} + c_t^{(\infty)} \\
&= E[G_{t-1}(\widehat{\theta})] - b_{t-1}^{(\infty)}\mathrm{Var}[G_{t-1}(\widehat{\theta})] + c_t^{(\infty)} + \frac{\triangle K_t}{4b_{t-1}^{(\infty)}(1+\triangle K_t)} \\
&= E[G_{t-1}(\widehat{\theta})] - b_{t-1}^{(\infty)}\mathrm{Var}[G_{t-1}(\widehat{\theta})] + c_t^{(\infty)} + \frac{\triangle K_t}{4b_t^{(\infty)}} \\
&= E[G_{t-1}(\widehat{\theta})] - b_{t-1}^{(\infty)}\mathrm{Var}[G_{t-1}(\widehat{\theta})] + c_{t-1}^{(\infty)} \\
&= \widetilde{V}_{t-1}, \qquad t = 1, \dots, T.
\end{aligned}$$

This proves the main assertion.

It remains to verify (5.41). Expanding the variance term as

$$\mathrm{Var}[G_t(\widehat{\theta})] = \mathrm{Var}[G_{t-1}(\widehat{\theta})] + 2\mathrm{Cov}\big(G_{t-1}(\widehat{\theta}), \widehat{\theta}_t\triangle S_t\big) + \mathrm{Var}[\widehat{\theta}_t\triangle S_t], \tag{5.43}$$

we need to compute the second and third terms in (5.43) more explicitly. We condition on $\mathcal{F}_{t-1}$, then use (5.32) and that $\triangle K_t$ is deterministic from Assumption 4.11, and finally (5.39) to compute

$$\begin{aligned}
\mathrm{Cov}\big(G_{t-1}(\widehat{\theta}), \widehat{\theta}_t\triangle S_t\big) &= E[G_{t-1}(\widehat{\theta})\widehat{\theta}_t\triangle A_t] - E[G_{t-1}(\widehat{\theta})]E[\widehat{\theta}_t\triangle A_t] \\
&= E\left[G_{t-1}(\widehat{\theta})\frac{\triangle K_t}{1+\triangle K_t}\left(\frac{1}{2b_{t-1}^{(\infty)}} - G_{t-1}(\widehat{\theta}) + E[G_{t-1}(\widehat{\theta})]\right)\right] \\
&\quad - E[G_{t-1}(\widehat{\theta})]E[\widehat{\theta}_t\triangle A_t] \\
&= \frac{\triangle K_t}{1+\triangle K_t}\Bigg(E[G_{t-1}(\widehat{\theta})]\frac{1}{2b_{t-1}^{(\infty)}} - E[G_{t-1}(\widehat{\theta})^2] \\
&\qquad + \big(E[G_{t-1}(\widehat{\theta})]\big)^2\Bigg) - E[G_{t-1}(\widehat{\theta})]E[\widehat{\theta}_t\triangle A_t] \\
&= \frac{\triangle K_t}{1+\triangle K_t}\left(E[G_{t-1}(\widehat{\theta})]\frac{1}{2b_{t-1}^{(\infty)}} - \mathrm{Var}[G_{t-1}(\widehat{\theta})]\right) \\
&\quad - E[G_{t-1}(\widehat{\theta})]\frac{\triangle K_t}{1+\triangle K_t}\frac{1}{2b_{t-1}^{(\infty)}} \\
&= -\frac{\triangle K_t}{1+\triangle K_t}\mathrm{Var}[G_{t-1}(\widehat{\theta})].
\end{aligned} \tag{5.44}$$

The conditioning step in the first equality of (5.44) is ensured thanks to (5.38) and its consequence $G_{t-1}(\widehat{\theta})\widehat{\theta}_t\triangle S_t \in L^1$. For the third term in (5.43), we condition on $\mathcal{F}_{t-1}$, then use the first line in (5.34) with (5.33) and finally combine the fact that as $\triangle K_t$ is deterministic, so is $b^{(\infty)}_{t-1}$, with (5.39) to get

$$\begin{aligned}
\mathrm{Var}[\widehat{\theta}_t \triangle S_t] &= E[(\widehat{\theta}_t \triangle S_t)^2] - (E[\widehat{\theta}_t \triangle A_t])^2 \\
&= E\left[\frac{\triangle K_t + (\triangle K_t)^2}{(1+\triangle K_t)^2}\left(\frac{1}{2b^{(\infty)}_{t-1}} - G_{t-1}(\widehat{\theta}) + E[G_{t-1}(\widehat{\theta})]\right)^2\right] \\
&\quad - (E[\widehat{\theta}_t \triangle A_t])^2 \\
&= \frac{\triangle K_t + (\triangle K_t)^2}{(1+\triangle K_t)^2}\left(\left(\frac{1}{2b^{(\infty)}_{t-1}}\right)^2 + \mathrm{Var}[G_{t-1}(\widehat{\theta})]\right) - \left(\frac{\triangle K_t}{1+\triangle K_t}\frac{1}{2b^{(\infty)}_{t-1}}\right)^2 \\
&= \frac{\triangle K_t}{1+\triangle K_t}\mathrm{Var}[G_{t-1}(\widehat{\theta})] + \frac{\triangle K_t}{(1+\triangle K_t)^2}\left(\frac{1}{2b^{(\infty)}_{t-1}}\right)^2. \qquad (5.45)
\end{aligned}$$

The conditioning step in (5.45) uses directly (5.38). Inserting (5.44) and (5.45) into (5.43) yields (5.41). □

Lemma 5.6 implies that $c^{(\infty)}_0 = \widetilde{V}_0 = \widetilde{V}_T = E[J^{\mathrm{mv}}_T(\widehat{\theta})]$ by (5.24)–(5.26). This proves (5.36) and finishes Recipe 5.4, 2), and hence we can state our main verification result below.

Theorem 5.7. *Suppose that Assumptions* 3.2*,* 3.4 *and* 4.11 *are satisfied, meaning that S is square-integrable and satisfies the structure condition (SC) with a deterministic MVT process K. Then $\widehat{\theta}$ defined in* (5.32) *and given explicitly by*

$$\widehat{\theta}_t := \frac{\lambda_t}{1+\triangle K_t}\left(\frac{1}{2b^{(\infty)}_{t-1}} - G_{t-1}(\widehat{\theta}) + E[G_{t-1}(\widehat{\theta})]\right), \quad t = 1, \dots, T, \qquad (5.46)$$

is an optimal strategy for the MVPS problem (5.14).

5.4 Convergence of strategies – preparation

Starting from this subsection, we turn to study the convergence behaviour of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ which is given explicitly in (5.11). We follow the convention described in Section 5.1 and try to argue that $\widehat{\boldsymbol{\vartheta}}^{(L)}$ converges in some sense to the optimal strategy $\widehat{\theta}$ for the MVPS problem given in (5.46). In this subsection, we do some preliminary work by collecting some identities and inequalities that will be used for proving the main assertion.

Let us briefly recap some consequences of Section 5.1. In the next two sections, we only work with the infinite product space. For a quantity $\mathbf{X}^{(L)}$, the superscript $^{(L)}$ indicates that it is originally defined on $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathcal{F}^{(L)})$ for $L \in \mathbb{N}$ and is lifted to the infinite product space via

$$\mathbf{X}^{(L)}(\omega_1, \omega_2, \dots) = \mathbf{X}^{(L)}(\omega_1, \dots, \omega_L) \tag{5.47}$$

as in (5.1). We also make $\mathbf{X}^{(L)}$ $\mathbb{R}^\infty$-valued by setting $\mathbf{X}^{\ell,(L)} \equiv 0$ for $\ell > L$ as in (5.2), with the first L components unchanged. Now note that $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ lives on $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, \mathbf{P}^{(\infty)})$ while $\widehat{\theta}$ lives on $(\Omega, \mathcal{F}, P)$. To make a comparison possible, we first recall from (2.18) that we have defined $X^{\otimes L}$ for any $L \in \mathbb{N} \cup \{\infty\}$ by

$$X^{\otimes L} := X \circ \pi_{\ell,L}, \qquad \ell = 1, \dots, L.$$

So we can use the process $\widehat{\theta}^{\otimes\infty}$ given by

$$\widehat{\theta}^{\ell,\otimes\infty}(\omega^{(\infty)}) = \widehat{\theta} \circ \pi_{\ell,\infty}(\omega^{(\infty)}) = \widehat{\theta}(\omega_\ell), \quad \ell \in \mathbb{N}, \tag{5.48}$$

to lift $\widehat{\theta}$ to $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, \mathbf{P}^{(\infty)})$ and obtain by the explicit formula for $\widehat{\theta}$ in (5.46) and (5.48) that

$$\widehat{\theta}_t^{\ell,\otimes\infty} = \frac{\boldsymbol{\lambda}_t^\ell}{1 + \triangle K_t}\left(\frac{1}{2b_{t-1}^{(\infty)}} - \mathbf{G}_{t-1}^\ell(\widehat{\theta}^{\otimes\infty}) + \mathbf{E}^{(\infty)}[\mathbf{G}_{t-1}^\ell(\widehat{\theta}^{\otimes\infty})]\right) \tag{5.49}$$

for $t = 1, \dots, T$ and $\ell \in \mathbb{N}$. Thanks to (5.48) and $P = \mathbf{P}^{(\infty)} \circ \pi_{\ell,\infty}$ in (2.12), we easily see that $(\widehat{\theta}^{\ell,\otimes\infty})_{\ell\in\mathbb{N}}$ are independent and each has the same distribution as $\widehat{\theta}$. This link between $\widehat{\theta}^{\ell,\otimes\infty}$ and $\widehat{\theta}$ is further exploited below. From now on, we write $\mathbf{E}$ instead of $\mathbf{E}^{(\infty)}$ for ease of notation. **In all what follows below, we have $L < \infty$.**

Lemma 5.8. *Suppose Assumptions* 3.2, 3.4 *and* 4.11 *are satisfied. Then for* $L \geq 2$ *and each* $t = 0, 1, \dots, T$*, the random variables* $(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}))_{\ell=1,\dots,L}$ *have the same first two moments.*

Proof. We first use (5.7), the explicit expressions $\lambda_t = \frac{\triangle A_t}{\triangle\langle M\rangle_t}, \triangle K_t = \frac{(\triangle A_t)^2}{\triangle\langle M\rangle_t}$ and $\triangle\langle S\rangle_t = \triangle\langle M\rangle_t + (\triangle A_t)^2$ in (3.20) and (3.10) and that $\triangle K_t$ is deterministic so that $\triangle\mathbf{K}_t^\ell = \triangle K_t$ for $\ell \in \mathbb{N}$ to obtain for $\ell \in \mathbb{N}$ that

$$\boldsymbol{\lambda}_t^\ell \triangle\mathbf{A}_t^\ell = \triangle K_t, \;\; (\boldsymbol{\lambda}_t^\ell)^2 \triangle\langle\mathbf{M}^\ell\rangle_t = \triangle K_t, \;\; (\boldsymbol{\lambda}_t^\ell)^2 \triangle\langle\mathbf{S}^\ell\rangle_t = \triangle K_t + (\triangle K_t)^2. \tag{5.50}$$

This is used repeatedly in the proof below. We also use the two properties that

$$\text{the random variables } \triangle \mathbf{S}_t^\ell,\ \widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell \text{ and } \mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) \text{ are all in } L^2, \tag{5.51}$$

$$\text{the random variables } \triangle K_t \text{ and hence } b_t^{(L)} \text{ and } b_t^{(\infty)} \text{ are deterministic.} \tag{5.52}$$

Via $\mathbf{S}^\ell(\omega^{(\infty)}) = S(\omega_\ell) = \mathbf{S}^{\ell,(L)}(\omega^{(L)})$ and (5.47) with $\boldsymbol{X}^{(L)} = \widehat{\boldsymbol{\vartheta}}^{(L)}$, property (5.51) can be translated from a statement in the L-extended market that $\mathbf{S}_t^{\ell,(L)}$, $\widehat{\boldsymbol{\vartheta}}_t^{(L)} \triangle \mathbf{S}_t^{\ell,(L)}$ and $\mathbf{G}_t^\ell(\boldsymbol{\vartheta}^{(L)})$ are in L^2. Note that the latter statement refers to the L-extended market, while (5.51) is a property in the infinite product space. Recall that the first property in (5.51) is a direct consequence of $S_t \in L^2$ by Assumption 3.2, and the other two follow from $\widehat{\boldsymbol{\vartheta}}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$ by Theorem 4.12 and the integrability requirement in the definition of $\boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$ in (3.29). Property (5.51) is frequently used below for taking out $\mathcal{G}_{t-1}^{(L)}$-measurable quantities from conditional expectations. The statement (5.52) uses that $\triangle K_t$ is deterministic by Assumption 4.11 and the explicit formulas of $b_t^{(L)}$ and $b_t^{(\infty)}$ in (4.77) and (5.28). We now separately show that for $t = 1, \dots, T$, all coordinates of the random vector $(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}))_{\ell=1,\dots,L}$ have the same first two moments.

1) For the first moment, we condition on $\mathcal{G}_{t-1}^{(L)}$, then use the explicit formula (5.11) for $\widehat{\boldsymbol{\vartheta}}^{(L)}$, (5.50) and (5.52) to compute, for $\ell = 1, \dots, L$,

$$\begin{aligned} \mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell] &= \mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{A}_t^\ell] \\ &= \frac{\triangle K_t}{(1 - L^{-1}) + \triangle K_t} \\ &\quad \times \left(\frac{1}{2b_{t-1}^{(L)}} - \mathbf{E}[\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})] + \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] \right). \end{aligned} \tag{5.53}$$

The conditioning step in (5.53) also uses (5.51). Then (5.53) yields by induction (forward in time), used for the empirical mean, that $\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell]$ does not depend on ℓ and is given by

$$\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell] = \frac{\triangle K_t}{(1 - L^{-1}) + \triangle K_t} \frac{1}{2b_{t-1}^{(L)}}, \quad \ell = 1, \dots, L,\ t = 1, \dots, T. \tag{5.54}$$

Because $\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) = \sum_{s=1}^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)} \triangle \mathbf{S}_s^\ell$ for $\ell = 1, \dots, L$ by (2.21), we obtain that $\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \mathbf{E}[\mathbf{G}_t^m(\widehat{\boldsymbol{\vartheta}}^{(L)})]$ for $\ell, m = 1, \dots, L$. Note that this also gives

$$\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big], \quad \ell = 1, \dots, L. \tag{5.55}$$

2) For the second moment, we claim for $\ell, m = 1, \dots, L$ and $t = 0, 1, \dots, T$ that

$$\mathbf{E}\big[\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] = \mathbf{E}\big[\big(\mathbf{G}_t^m(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big], \tag{5.56}$$

$$\mathbf{E}\big[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] = \mathbf{E}\big[\mathbf{G}_t^m(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big]. \tag{5.57}$$

We prove these equalities by induction. The base case $t = 0$ is evident. Suppose the claim is true for $t - 1$. Below we use from (2.21) and (2.23) the identities

$$\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) = \mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) + \widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell, \quad t = 1, \dots, T,\ \ell \in \mathbb{N}, \tag{5.58}$$

$$\mathrm{em}(\mathbf{x} + \mathbf{y}) = \mathrm{em}(\mathbf{x}) + \mathrm{em}(\mathbf{y}). \tag{5.59}$$

We use (5.58), then condition on $\mathcal{G}_{t-1}^{(L)}$ and use the $\mathcal{G}_{t-1}^{(L)}$-measurability of both $\widehat{\boldsymbol{\vartheta}}_t^{(L)}$ and $\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})$, the identity $\mathbf{E}[\triangle \mathbf{S}_t^\ell | \mathcal{G}_{t-1}^{(L)}] = \triangle \mathbf{A}_t^\ell$ from (5.3) and the definition of $\triangle\langle \mathbf{S}^\ell \rangle$ in (5.5) to obtain

$$\begin{aligned}
\mathbf{E}\big[\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] &= \mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] + 2\mathbf{E}[\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell] \\
&\quad + \mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell)^2] \\
&= \mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] + 2\mathbf{E}[\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{A}_t^\ell] \\
&\quad + \mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)})^2 \triangle \langle \mathbf{S}^\ell \rangle_t].
\end{aligned} \tag{5.60}$$

The conditioning step uses (5.51) as well. For the second term on the right-hand side of (5.60), we insert the explicit formula for $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ in (5.11) into that second term and use (5.50) with (5.52) to obtain

$$\begin{aligned}
\mathbf{E}[\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{A}_t^\ell] &= \frac{\triangle K_t}{(1 - L^{-1}) + \triangle K_t} \\
&\quad \times \Bigg(\frac{\mathbf{E}[\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{2b_{t-1}^{(L)}} - \mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] \\
&\qquad + \mathbf{E}\big[\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] \Bigg).
\end{aligned} \tag{5.61}$$

Similarly plugging the explicit formula for $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ in (5.11) into the third term in (5.60) and using again (5.50) with (5.52) gives together with the equality

$\mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \mathbf{E}[\mathrm{em}(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$ from (5.55) that

$$\begin{aligned}\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t)^2\triangle\langle\mathbf{S}^\ell\rangle_t] &= \frac{\triangle K_t + (\triangle K_t)^2}{((1-L^{-1})+\triangle K_t)^2}\mathbf{E}\bigg[\bigg(\frac{1}{2b^{(L)}_{t-1}} - \mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) \\ &\qquad + \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big]\bigg)^2\bigg] \\ &= \frac{\triangle K_t + (\triangle K_t)^2}{((1-L^{-1})+\triangle K_t)^2}\bigg(\Big(\frac{1}{2b^{(L)}_{t-1}}\Big)^2 + \mathbf{E}\big[\big(\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] \\ &\qquad - 2\mathbf{E}\big[\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] \\ &\qquad + \Big(\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big]\Big)^2\bigg). \end{aligned} \tag{5.62}$$

Inserting (5.61) and (5.62) into (5.60) and using the induction hypotheses for (5.56) and (5.57), we get $\mathbf{E}[(\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2] = \mathbf{E}[(\mathbf{G}^m_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2]$ for $\ell, m = 1, \ldots, L$. This completes the induction step for (5.56).

It remains to show $\mathbf{E}[\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))] = \mathbf{E}[\mathbf{G}^m_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$. To this end, we use (5.58) and (5.59) to obtain

$$\begin{aligned}&\mathbf{E}\big[\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] \\ &= \mathbf{E}\big[\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] + \mathbf{E}\big[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t\triangle\mathbf{S}^\ell_t\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] \\ &\quad + \mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle\mathbf{S}_t)] + \mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t\triangle\mathbf{S}^\ell_t\mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle\mathbf{S}_t)]. \end{aligned} \tag{5.63}$$

We now argue that each term on the RHS of (5.63) does not depend on ℓ. For the first term, this is true by the induction hypothesis for (5.57). For the second term in (5.63), we condition on $\mathcal{G}^{(L)}_{t-1}$, then invoke the $\mathcal{G}^{(L)}_{t-1}$-measurability of $\widehat{\boldsymbol{\vartheta}}^{(L)}_t$ and $\mathrm{em}(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}))$ and finally use the explicit formula for $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ in (5.11), (5.50) and (5.52) to compute

$$\begin{aligned}\mathbf{E}\big[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t\triangle\mathbf{S}^\ell_t\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] &= \frac{\triangle K_t}{(1-L^{-1})+\triangle K_t} \\ &\quad \times \mathbf{E}\bigg[\bigg(\frac{1}{2b^{(L)}_{t-1}} - \mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) + \mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\bigg) \\ &\qquad \times \mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\bigg]. \end{aligned} \tag{5.64}$$

The conditioning step also uses (5.51). By the induction hypothesis for (5.57), the last quantity is independent of ℓ for $\ell = 1, \ldots, L$. For the third term in (5.63), we argue similarly as in (5.64) and use (5.59) to get

$$\begin{aligned}
&\mathbf{E}[\mathbf{G}^{\ell}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{S}_t)] \\
&= \mathbf{E}\bigg[\mathbf{G}^{\ell}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) \frac{\triangle K_t}{(1-L^{-1})+\triangle K_t} \\
&\qquad \times \mathrm{em}\bigg(\frac{1}{2b^{(L)}_{t-1}} - \mathbf{G}^{\ell}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) + \mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\bigg)\bigg] \\
&= \mathbf{E}\bigg[\mathbf{G}^{\ell}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) \frac{\triangle K_t}{(1-L^{-1})+\triangle K_t} \frac{1}{2b^{(L)}_{t-1}}\bigg].
\end{aligned}$$

Because all coordinates of $(\mathbf{G}^{\ell}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))_{\ell=1,\dots,L}$ have by part 1) the same first moment for $t=0,1,\dots,T$ (see (5.55)), the last expression in the above display is independent of ℓ for $\ell=1,\dots,L$. For the fourth term in (5.63), we first use $\triangle \mathbf{S}^{\ell}_t = \triangle \mathbf{M}^{\ell}_t + \triangle \mathbf{A}^{\ell}_t$ and (5.59) to expand

$$\begin{aligned}
&\mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{S}^{\ell}_t \mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{S}_t)] \\
&= \mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{M}^{\ell}_t \mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{M}_t)] + \mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{M}^{\ell}_t \mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{A}_t)] \\
&\quad + \mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{A}^{\ell}_t \mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{M}_t)] + \mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{A}^{\ell}_t \mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{A}_t)]. \qquad (5.65)
\end{aligned}$$

We then condition on $\mathcal{G}^{(L)}_{t-1}$ with the integrability property from (5.51) and use the $\mathcal{G}^{(L)}_{t-1}$-measurability of $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t$ and $\mathbf{E}[\triangle \mathbf{M}^{\ell}_t \triangle \mathbf{M}^{m}_t | \mathcal{G}^{(L)}_{t-1}] = 0$ for $\ell, m = 1,\dots,L$ and $\ell \neq m$ from (5.4) to obtain $\mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{M}^{\ell}_t \widehat{\boldsymbol{\vartheta}}^{m,(L)}_t \triangle \mathbf{M}^{m}_t] = 0$ for $\ell, m = 1,\dots,L$ and $\ell \neq m$. Using $\mathrm{em}(\mathbf{x}^{(L)}) = \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{x}^{\ell,(L)}$, the last identity and the definition of $\triangle\langle \mathbf{M}^{\ell}\rangle_t$ in (5.5) simplifies the first term in (5.65) to

$$\mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{M}^{\ell}_t \mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{M}_t)] = \frac{1}{L}\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t)^2 \triangle \langle \mathbf{M}^{\ell}\rangle_t].$$

By a similar conditioning step and the identity $\mathbf{E}[\triangle \mathbf{M}^{\ell}_t | \mathcal{G}^{(L)}_{t-1}] = 0$ from (5.3), we get $\mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{A}^{\ell}_t \widehat{\boldsymbol{\vartheta}}^{m,(L)}_t \triangle \mathbf{M}^{m}_t] = 0$ for all $\ell, m = 1,\dots,L$. Averaging this equality over ℓ and m respectively, we obtain that

$$\mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{M}^{\ell}_t \mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{A}_t)] = \mathbf{E}[\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{A}^{\ell}_t \mathrm{em}(\widehat{\boldsymbol{\vartheta}}^{(L)}_t \odot \triangle \mathbf{M}_t)] = 0.$$

We insert the last two displays into (5.65) and then use (5.11), (5.50) and (5.52) to compute

$$\begin{aligned}\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell \mathrm{em}(\widehat{\boldsymbol{\vartheta}}_t^{(L)}\odot\triangle\mathbf{S}_t)] &= \frac{1}{L}\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)})^2\triangle\langle\mathbf{M}^\ell\rangle_t] \\ &\quad+\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{A}_t^\ell \mathrm{em}(\widehat{\boldsymbol{\vartheta}}_t^{(L)}\odot\triangle\mathbf{A}_t)] \\ &= \frac{1}{L}\frac{(\triangle K_t)^2}{((1-L^{-1})+\triangle K_t)^2} \\ &\quad\times\mathbf{E}\bigg[\bigg(\frac{1}{2b_{t-1}^{(L)}}-\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})+\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\bigg)^2\bigg] \\ &\quad+\frac{\triangle K_t}{(1-L^{-1})+\triangle K_t}\frac{1}{2b_{t-1}^{(L)}}\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{A}_t^\ell]. \end{aligned} \tag{5.66}$$

Getting the last quantity in the second equality in (5.66) also uses

$$\begin{aligned}\mathrm{em}(\widehat{\boldsymbol{\vartheta}}_t^{(L)}\odot\triangle\mathbf{A}_t) &= \frac{\triangle K_t}{(1-L^{-1})+\triangle K_t} \\ &\quad\times\mathrm{em}\bigg(\frac{1}{2b_{t-1}^{(L)}}-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)+\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\bigg) \\ &= \frac{\triangle K_t}{(1-L^{-1})+\triangle K_t}\frac{1}{2b_{t-1}^{(L)}},\end{aligned}$$

obtained by multiplying $\triangle\mathbf{A}_t^\ell$ in the explicit formula (5.11) for $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$, averaging over ℓ and using (5.50). Expanding the square in the second-to-last line in (5.66), we can see that it depends on ℓ through the expectations of $\mathbf{G}_{t-1}^\ell(\boldsymbol{\vartheta})$, $(\mathbf{G}_{t-1}^\ell(\boldsymbol{\vartheta}))^2$ and $\mathbf{G}_{t-1}^\ell(\boldsymbol{\vartheta})\mathrm{em}(\mathbf{G}_{t-1}(\boldsymbol{\vartheta}))$. All of these terms are independent of ℓ for $\ell=1,\dots,L$ by part 1), see (5.55), and the induction hypotheses for (5.56) and (5.57). The last term in (5.66) is also independent of ℓ by $\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{A}_t^\ell]=\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]$ and (5.54). We can now conclude that all terms in (5.63) are independent of ℓ for $\ell=1,\dots,L$. This finishes the induction step for (5.57). □

Lemma 5.9. *Suppose Assumptions* 3.2*,* 3.4 *and* 4.11 *are satisfied. Then the following statements hold:*

1) *For each* $t=0,1,\dots,T$*, we have*

$$\sup_{L\in\mathbb{N}}\bigg(\frac{1}{b_t^{(L)}}\bigg)\le\frac{1}{\xi}\prod_{u=t+1}^{T}(1+2K_T)<\infty. \tag{5.67}$$

2) *There exists a constant* $C>0$ *such that for* $t=1,\dots,T$ *and* $\ell=1,\dots,L$*,*

we have

$$\sup_{L\in\mathbb{N},L\geq 2}\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell)^2]\leq 3C\sup_{L\in\mathbb{N},L\geq 2}\left(\frac{1}{2b_{t-1}^{(L)}}\right)^2 + 6C\sup_{L\in\mathbb{N},L\geq 2}\mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big], \tag{5.68}$$

and hence, for each $\ell\in\mathbb{N}$

$$\sup_{L\in\mathbb{N}}\mathbf{E}\big[\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big]<\infty,\quad t=0,1,\dots,T. \tag{5.69}$$

Here we use the convention (5.8) *so that* $\mathbf{G}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\equiv 0$ *for* $\ell>L$.

3) *For* $t=1,\dots,T$,

$$\sum_{\ell\neq m}^{L}\frac{\mathrm{Cov}^{\mathbf{P}}(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell,\widehat{\boldsymbol{\vartheta}}_t^{m,(L)}\triangle\mathbf{S}_t^m)}{L^2} = -\frac{1}{L}\left(\frac{\triangle K_t}{(1-L^{-1})+\triangle K_t}\right)^2\mathbf{E}\big[\mathrm{evar}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big]. \tag{5.70}$$

Proof. 1) Because all the $\triangle K_t$ are deterministic by Assumption 4.11, we use the explicit formula (4.77) for $b^{(L)}$, $1-L^{-1}\geq\frac{1}{2}$ for $L\geq 2$ and $\triangle K_u\leq K_T$ for $u=1,\dots,T$ from (3.20) to get

$$\sup_{L\in\mathbb{N}}\left(\frac{1}{b_t^{(L)}}\right)=\sup_{L\in\mathbb{N}}\left(\frac{1}{\xi}\prod_{u=t+1}^{T}\frac{(1-L^{-1})+\triangle K_u}{1-L^{-1}}\right)\leq\frac{1}{\xi}(1+2K_T)^T<\infty.$$

2) Let us show (5.68). First, we obviously get (as in (4.43)) that

$$\frac{\triangle K_t+(\triangle K_t)^2}{((1-L^{-1})+\triangle K_t)^2}\leq\frac{\triangle K_t+(\triangle K_t)^2}{(1-L^{-1})^2}\leq 4(K_T+K_T^2)=:C. \tag{5.71}$$

Recall that we have computed $\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell)^2]$ in (5.62). We insert (5.71) into the first line of (5.62), then apply Cauchy–Schwarz successively and finally use

(5.56) to get

$$
\begin{aligned}
\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell)^2] &\le C\mathbf{E}\Bigg[\bigg(\frac{1}{2b_{t-1}^{(L)}} - \mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) + \mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\bigg)^2\Bigg] \\
&\le 3C\bigg(\Big(\frac{1}{2b_{t-1}^{(L)}}\Big)^2 + \mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] \\
&\qquad + \mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big]\bigg) \\
&\le 3C\bigg(\Big(\frac{1}{2b_{t-1}^{(L)}}\Big)^2 + \mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] + \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})^2\big)\big]\bigg) \\
&= 3C\bigg(\Big(\frac{1}{2b_{t-1}^{(L)}}\Big)^2 + 2\mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big]\bigg).
\end{aligned}
$$

This gives (5.68). We now prove (5.69) by induction. The claim for $t = 0$ is trivial. Suppose (5.69) holds for $t-1$. Then we use (5.68) to estimate

$$
\begin{aligned}
\sup_{L\in\mathbb{N},L\ge 2} \mathbf{E}\big[\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] &\le 2 \sup_{L\in\mathbb{N},L\ge 2} \mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] + 2 \sup_{L\in\mathbb{N},L\ge 2} \mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell)^2] \\
&\le 6C \sup_{L\in\mathbb{N},L\ge 2} \bigg(\frac{1}{2b_{t-1}^{(L)}}\bigg)^2 \\
&\quad + (2+12C) \sup_{L\in\mathbb{N},L\ge 2} \mathbf{E}\big[\big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big].
\end{aligned}
$$

The last quantity is finite thanks to (5.67) and the induction hypothesis. This proves the induction step.

3) Let $\ell, m \in \{1, \dots, L\}$ and $\ell \ne m$. Conditioning on $\mathcal{G}_{t-1}^{(L)}$, then pulling out $\mathcal{G}_{t-1}^{(L)}$-measurable quantities and using the martingale property in (5.3) and the strong orthogonality in (5.4), we get

$$
\begin{aligned}
\mathrm{Cov}^{\mathbf{P}}(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell, \widehat{\boldsymbol{\vartheta}}_t^{m,(L)} \triangle \mathbf{S}_t^m) &= \mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \widehat{\boldsymbol{\vartheta}}_t^{m,(L)} (\triangle \mathbf{M}_t^\ell + \triangle \mathbf{A}_t^\ell)(\triangle \mathbf{M}_t^m + \triangle \mathbf{A}_t^m)] \\
&\quad - \mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} (\triangle \mathbf{M}_t^\ell + \triangle \mathbf{A}_t^\ell)] \\
&\quad \times \mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{m,(L)} (\triangle \mathbf{M}_t^m + \triangle \mathbf{A}_t^m)] \\
&= \mathrm{Cov}^{\mathbf{P}}(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{A}_t^\ell, \widehat{\boldsymbol{\vartheta}}_t^{m,(L)} \triangle \mathbf{A}_t^m). \qquad (5.72)
\end{aligned}
$$

The conditioning step also uses the integrability properties in (5.51). Recall $\triangle \widetilde{K}_t^{(L)} = \frac{\triangle K_t}{(1-L^{-1})+\triangle K_t}$ from Condition 4.6, which is satisfied due to the equivalence between Conditions 4.6 and 4.7 as argued in Remark 4.8, and that the MVT process K is deterministic by Assumption 4.11. Using the formula (5.11) for $\widehat{\boldsymbol{\vartheta}}^{(L)}$ with (5.50), then (5.52) and finally $E[\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}))] = 0$ from

(5.55), we get

$$
\begin{aligned}
&\mathrm{Cov}^{\mathbf{P}}(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{A}_t^{\ell},\widehat{\boldsymbol{\vartheta}}_t^{m,(L)}\triangle\mathbf{A}_t^{m})\\
&=\mathrm{Cov}^{\mathbf{P}}\bigg(\frac{\triangle\widetilde{K}_t^{(L)}}{2b_{t-1}^{(L)}},\frac{\triangle\widetilde{K}_t^{(L)}}{2b_{t-1}^{(L)}}\bigg)\\
&\quad+\mathrm{Cov}^{\mathbf{P}}\bigg(\frac{\triangle\widetilde{K}_t^{(L)}}{2b_{t-1}^{(L)}},\triangle\widetilde{K}_t^{(L)}\Big(\mathbf{G}_{t-1}^{m}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\bigg)\\
&\quad+\mathrm{Cov}^{\mathbf{P}}\bigg(\frac{\triangle\widetilde{K}_t^{(L)}}{2b_{t-1}^{(L)}},\triangle\widetilde{K}_t^{(L)}\Big(\mathbf{G}_{t-1}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\bigg)\\
&\quad+\mathrm{Cov}^{\mathbf{P}}\bigg(\triangle\widetilde{K}_t^{(L)}\Big(\mathbf{G}_{t-1}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big),\\
&\qquad\qquad\triangle\widetilde{K}_t^{(L)}\Big(\mathbf{G}_{t-1}^{m}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\bigg)\\
&=(\triangle\widetilde{K}_t^{(L)})^2\\
&\quad\times\mathrm{Cov}^{\mathbf{P}}\Big(\mathbf{G}_{t-1}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big),\mathbf{G}_{t-1}^{m}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\\
&=(\triangle\widetilde{K}_t^{(L)})^2\mathbf{E}\Big[\Big(\mathbf{G}_{t-1}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\\
&\qquad\qquad\times\Big(\mathbf{G}_{t-1}^{m}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\Big].
\end{aligned}
\tag{5.73}
$$

Inserting (5.73) into (5.72), adding and subtracting the (diagonal, $\ell=m$) term $\frac{1}{L^2}\sum_{\ell=1}^{L}(\triangle\widetilde{K}_t^{(L)})^2\mathbf{E}[(\mathbf{G}_{t-1}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]$, and finally using the expansion $(\sum_{\ell=1}^{L}\mathbf{x}^\ell)^2 = \sum_{\ell,m=1}^{L}\mathbf{x}^\ell\mathbf{x}^m$, we get

$$\begin{aligned}
&\frac{1}{L^2}\sum_{\ell\neq m}^{L}\mathrm{Cov}^{\mathbf{P}}\big(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell,\widehat{\boldsymbol{\vartheta}}_t^{m,(L)}\triangle\mathbf{S}_t^m\big)\\
&=\frac{1}{L^2}\sum_{\ell\neq m}^{L}(\triangle\widetilde{K}_t^{(L)})^2\mathbf{E}\Big[\Big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\\
&\qquad\qquad\times\Big(\mathbf{G}_{t-1}^m(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\Big]\\
&=\frac{1}{L^2}\sum_{\ell\neq m}^{L}(\triangle\widetilde{K}_t^{(L)})^2\mathbf{E}\Big[\Big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\\
&\qquad\qquad\times\Big(\mathbf{G}_{t-1}^m(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\Big]\\
&\quad+\frac{1}{L^2}\sum_{\ell=1}^{L}(\triangle\widetilde{K}_t^{(L)})^2\mathbf{E}\Big[\Big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big]\\
&\quad-\frac{1}{L^2}\sum_{\ell=1}^{L}(\triangle\widetilde{K}_t^{(L)})^2\mathbf{E}\Big[\Big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big]\\
&=(\triangle\widetilde{K}_t^{(L)})^2\mathbf{E}\Big[\Big(\frac{1}{L}\sum_{\ell=1}^{L}\Big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\Big)^2\Big]\\
&\quad-\frac{1}{L^2}\sum_{\ell=1}^{L}(\triangle\widetilde{K}_t^{(L)})^2\mathbf{E}\Big[\Big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big]\\
&=-\frac{(\triangle\widetilde{K}_t^{(L)})^2}{L}\mathbf{E}\big[\mathrm{evar}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big].
\end{aligned}$$

The last equality uses (5.18) and the definition of the empirical variance in (2.24). This completes the proof. □

5.5 Convergence of strategies – main results

We are ready to present the main result of these two subsections:

$$\max_{\ell=1,\dots,L}\mathbf{E}\Big[\max_{t=1,\dots,T}\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})-\mathbf{G}_t^\ell(\widehat{\theta}^{\otimes\infty})\big)^2\Big]=O(L^{-1})\longrightarrow 0\quad\text{as } L\to\infty. \tag{5.74}$$

As a byproduct, we obtain a second proof of Theorem 5.7, i.e., the optimality of $\widehat{\theta}$ for the MVPS problem (2.3).

First, we adapt from the convergence of $b^{(L)}$ to $b^{(\infty)}$ given in (5.31) to obtain

$$\max_{t=0,1,\dots,T}\left|\frac{1}{b_t^{(L)}}-\frac{1}{b_t^{(\infty)}}\right|\leq\frac{1}{\xi}\max_{t=0,1,\dots,T-1}\prod_{u=t+1}^{T}\left|\frac{\triangle K_u}{1-L^{-1}}-\triangle K_u\right|=O(L^{-1})\quad(5.75)$$

because $b_T^{(L)}=\xi=b_T^{(\infty)}$ and each factor satisfies for $u=1,\dots,T$ that

$$\left|\frac{\triangle K_u}{1-L^{-1}}-\triangle K_u\right|=\frac{\triangle K_u}{L-1}=O(L^{-1})$$

due to the fact that $\triangle K_u$ is deterministic from Assumption 4.11. Next, we establish an L^2-weak law of large numbers for the optimal final gains $(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}))_{\ell=1,\dots,L}$ as $L\to\infty$. Note that the empirical average em implicitly depends on L.

Proposition 5.10. *Suppose Assumptions* 3.2, 3.4 *and* 4.11 *are satisfied. Then for any $\ell\in\mathbb{N}$, we have for $t=0,1,\dots,T$ that*

$$\mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)-\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})]\Big)^2\Big]=O(L^{-1})\quad\textit{for } L\in\mathbb{N}\setminus\{1\}.\qquad(5.76)$$

Proof. We prove this result by induction over t. The statement is trivial for $t=0$. Suppose (5.76) holds for $t-1$. For the induction step, thanks to Cauchy–Schwarz, we only need to prove that

$$\max_{\ell=1,\dots,L}\mathbf{E}\big[\big(\mathrm{em}(\widehat{\boldsymbol{\vartheta}}_t^{(L)}\odot\triangle\mathbf{S}_t)-\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]\big)^2\big]=O(L^{-1})\quad\text{for } L\in\mathbb{N}\setminus\{1\}.$$

First we recall from (5.55) that

$$\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]=\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell',(L)}\triangle\mathbf{S}_t^{\ell'}]\quad\text{for any } \ell,\ell'\in\{1,\dots,L\}\qquad(5.77)$$

and use this to obtain

$$\begin{aligned}&\max_{\ell'=1,\dots,L}\mathbf{E}\big[\big(\mathrm{em}(\widehat{\boldsymbol{\vartheta}}_t^{(L)}\odot\triangle\mathbf{S}_t)-\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell',(L)}\triangle\mathbf{S}_t^\ell]\big)^2\big]\\&=\mathbf{E}\big[\big(\mathrm{em}(\widehat{\boldsymbol{\vartheta}}_t^{(L)}\odot\triangle\mathbf{S}_t)-\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]\big)^2\big].\end{aligned}\qquad(5.78)$$

Then we use (5.78) and the definition of the empirical average recalled in (5.18), expand $(\sum_{\ell=1}^{L}\mathbf{x}^\ell)^2$ and finally invoke (5.77) to replace ℓ' by ℓ to write

$$\begin{aligned}
&\max_{\ell'=1,\dots,L}\mathbf{E}\Big[\big(\mathrm{em}(\widehat{\boldsymbol{\vartheta}}_t^{(L)}\odot\triangle\mathbf{S}_t)-\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell',(L)}\triangle\mathbf{S}_t^\ell]\big)^2\Big]\\
&=\mathbf{E}\Big[\big(\mathrm{em}(\widehat{\boldsymbol{\vartheta}}_t^{(L)}\odot\triangle\mathbf{S}_t)-\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]\big)^2\Big]\\
&=\mathbf{E}\Bigg[\sum_{\ell'=1}^{L}\bigg(\frac{\widehat{\boldsymbol{\vartheta}}_t^{\ell',(L)}\triangle\mathbf{S}_t^{\ell'}-\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]}{L}\bigg)^2\Bigg]\\
&=\mathbf{E}\Bigg[\sum_{\ell=1}^{L}\bigg(\frac{\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^{\ell}-\mathbf{E}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]}{L}\bigg)^2\Bigg]\\
&=\sum_{\ell=1}^{L}\frac{\mathrm{Var}^{\mathbf{P}}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]}{L^2}+\sum_{\ell\neq m}^{L}\frac{\mathrm{Cov}^{\mathbf{P}}(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell,\widehat{\boldsymbol{\vartheta}}_t^{m,(L)}\triangle\mathbf{S}_t^m)}{L^2}.
\end{aligned}\tag{5.79}$$

For the variance term in (5.79), we use (5.68) and (5.69) to obtain

$$\sup_{L\in\mathbb{N},L\geq 2}\mathrm{Var}^{\mathbf{P}}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]\leq\sup_{L\in\mathbb{N},L\geq 2}\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell)^2]<\infty$$

and hence

$$\sum_{\ell=1}^{L}\frac{\mathrm{Var}^{\mathbf{P}}[\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell]}{L^2}\leq\frac{\sup_{L\in\mathbb{N}}\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell)^2]}{L}=O(L^{-1})\quad\text{for } L\in\mathbb{N}\setminus\{1\}.$$

For the covariance term in (5.79), we use (5.70) and $\frac{\triangle K_t}{(1-L^{-1})+\triangle K_t}\leq 1$ with $\mathrm{evar}(\mathbf{x})\leq\mathrm{em}(\mathbf{x}^2)$ from the definition of evar in (2.24) to obtain

$$\begin{aligned}
\Bigg|\sum_{\ell\neq m}^{L}\frac{\mathrm{Cov}^{\mathbf{P}}(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell,\widehat{\boldsymbol{\vartheta}}_t^{m,(L)}\triangle\mathbf{S}_t^m)}{L^2}\Bigg|&=\frac{1}{L}\bigg(\frac{\triangle K_t}{(1-L^{-1})+\triangle K_t}\bigg)^2\\
&\quad\times\mathbf{E}\Big[\mathrm{evar}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big]\\
&\leq\frac{1}{L}\sup_{L\in\mathbb{N},L\geq 2}\mathbf{E}\Big[\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})^2\big)\Big]\\
&=O(L^{-1})
\end{aligned}$$

The last step uses that $\sup_{L\in\mathbb{N},L\geq 2}\mathbf{E}[\mathrm{em}(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})^2)]<\infty$, which follows from (5.69) and the identity $\mathbf{E}[\mathrm{em}(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})^2)]=\mathbf{E}[(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2]$ for $\ell=1,\dots,L$ by (5.56). Since both terms in (5.79) are $O(L^{-1})$ for $L\in\mathbb{N}\setminus\{1\}$, the proof is complete. □

Now we control the L^2-distance between $\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}\triangle\mathbf{S}_t^\ell$ and $\widehat{\theta}_t^{\ell,\otimes\infty}\triangle\mathbf{S}_t^\ell$ in terms of differences between $\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)}$ and $\widehat{\theta}_t^{\ell,\otimes\infty}$. Note that both $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ and $\widehat{\theta}^{\ell,\otimes\infty}$ are defined on $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, \mathbf{P}^{(\infty)})$. In particular, each ℓ-th coordinate of the process $\widehat{\theta}^{\otimes\infty}(\omega^{(\infty)})$ is simply equal to $\widehat{\theta}(\omega_\ell)$, where $\widehat{\theta}$ is the optimal strategy from (5.46) for the MVPS problem (5.14).

Lemma 5.11. *Suppose that Assumptions* 3.2, 3.4 *and* 4.11 *are satisfied. Then there exists a constant $C > 0$ such that for each $t = 1, \dots, T$ and $\ell = 1, \dots, L$, we have*

$$\begin{aligned}
&\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell - \widehat{\theta}_t^{\ell,\otimes\infty} \triangle \mathbf{S}_t^\ell)^2] \\
&\le 6C\mathbf{E}\bigg[\bigg(\frac{1}{b_{t-1}^{(L)}} - \frac{1}{b_{t-1}^{(\infty)}}\bigg)^2 + \big(\mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_{t-1}^\ell(\widehat{\theta}^{\otimes\infty})\big)^2 \\
&\qquad + \Big(\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - \mathbf{E}[\mathbf{G}_{t-1}^\ell(\widehat{\theta}^{\otimes\infty})]\Big)^2\bigg] \\
&\quad + \frac{2C}{L^2}\mathbf{E}\bigg[\bigg(\frac{1}{2b_{t-1}^{(\infty)}} - \mathbf{G}_{t-1}^\ell(\widehat{\theta}^{\otimes\infty}) + \mathbf{E}[\mathbf{G}_{t-1}^\ell(\widehat{\theta}^{\otimes\infty})]\bigg)^2\bigg].
\end{aligned} \tag{5.80}$$

Proof. First we introduce a few notations for some quantities in the definitions of $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ and $\widehat{\theta}^{\ell,\otimes\infty}$, namely

$$\begin{aligned}
U_1 &= \frac{1}{(1 - L^{-1}) + \triangle K_t}, \qquad W_1 = \frac{1}{2b_{t-1}^{(L)}}, \\
X_1 &= \mathbf{G}_{t-1}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}), \qquad Y_1 = \mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big), \\
U_2 &= \frac{1}{1 + \triangle K_t}, \qquad W_2 = \frac{1}{2b_{t-1}^{(\infty)}}, \\
X_2 &= \mathbf{G}_{t-1}^\ell(\widehat{\theta}^{\otimes\infty}), \qquad Y_2 = \mathbf{E}[\mathbf{G}_{t-1}^\ell(\widehat{\theta}^{\otimes\infty})].
\end{aligned} \tag{5.81}$$

To obtain the bound (5.80), we use the affine structures of $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ and $\widehat{\theta}^{\ell,\otimes\infty}$ (see (5.11) and (5.49)) and operate at the level of (5.81). Conditioning on $\mathcal{G}_{t-1}^{(L)}$ and using (5.81), then invoking $(\boldsymbol{\lambda}_t^\ell)^2 \triangle \langle \mathbf{S}^\ell \rangle_t = \triangle K_t + (\triangle K_t)^2$ in (5.50), we can write the LHS of (5.80) as

$$\begin{aligned}
\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell - \widehat{\theta}_t^{\ell,\otimes\infty} \triangle \mathbf{S}_t^\ell)^2] &= \mathbf{E}\big[\big(U_1(W_1 - X_1 + Y_1) - U_2(W_2 - X_2 + Y_2)\big)^2 \\
&\qquad \times (\boldsymbol{\lambda}_t^\ell)^2 \triangle \langle \mathbf{S}^\ell \rangle_t\big] \\
&= \mathbf{E}\big[\big(U_1(W_1 - X_1 + Y_1) - U_2(W_2 - X_2 + Y_2)\big)^2 \\
&\qquad \times \big(\triangle K_t + (\triangle K_t)^2\big)\big].
\end{aligned} \tag{5.82}$$

The conditioning step uses the $\mathcal{G}_{t-1}^{(L)}$-measurability of U_i, W_i, X_i, Y_i for $i \in \{1, 2\}$ and the integrability property in (5.51). Thanks to the deterministic property in

(5.52), we can write $\triangle K_t + (\triangle K_t)^2 \le C := K_T + K_T^2$. Then we add and subtract $-U_1(W_2 - X_2 + Y_2)$ on the RHS of (5.82) and use the Cauchy–Schwarz inequality successively to get

$$\begin{aligned}\mathbf{E}[(\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} \triangle \mathbf{S}_t^\ell - \widehat{\theta}_t^{\ell,\otimes\infty} \triangle \mathbf{S}_t^\ell)^2] &\le C\mathbf{E}\big[\big((U_1(W_1 - X_1 + Y_1) - U_1(W_2 - X_2 + Y_2) \\ &\qquad + U_1(W_2 - X_2 + Y_2) - U_2(W_2 - X_2 + Y_2)\big)^2\big] \\ &\le 2C\big(\mathbf{E}[U_1^2(W_1 - W_2 - X_1 + X_2 + Y_1 - Y_2)^2] \\ &\qquad + \mathbf{E}[(U_1 - U_2)^2(W_2 - X_2 + Y_2)^2]\big) \\ &\le 6C\mathbf{E}\big[U_1^2\big((W_1 - W_2)^2 + (X_1 - X_2)^2 \\ &\qquad + (Y_1 - Y_2)^2\big)\big] \\ &\quad + 2C\mathbf{E}[(U_1 - U_2)^2(W_2 - X_2 + Y_2)^2]\big). \end{aligned} \tag{5.83}$$

Because $L \ge 2$ and $\triangle K_t \ge 0$ by (3.20), we have

$$U_1^2 = \frac{1}{((1 - L^{-1}) + \triangle K_t)^2} \le \frac{1}{(\frac{1}{2} + \triangle K_t)^2} \le 4, \tag{5.84}$$

$$\begin{aligned}(U_1 - U_2)^2 &= \left(\frac{1}{(1 - L^{-1}) + \triangle K_t} - \frac{1}{1 + \triangle K_t}\right)^2 \\ &= \frac{L^{-2}}{((1 - L^{-1}) + \triangle K_t)^2(1 + \triangle K_t)^2} \le 4L^{-2}. \end{aligned} \tag{5.85}$$

Inserting (5.84) and (5.85) into (5.83) yields the desired bound (5.80). □

Now we are ready to prove the main convergence result (5.74).

Theorem 5.12. *Suppose that Assumptions* 3.2, 3.4 *and* 4.11 *are satisfied, meaning that S is square-integrable and satisfies the structure condition (SC), and the MVT process K is deterministic. Then for every $t = 0, 1, \dots, T$, we have*

$$\max_{\ell=1,\dots,L} \mathbf{E}\big[\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_t^\ell(\widehat{\theta}^{\otimes\infty})\big)^2\big] = O(L^{-1}) \quad \textit{for } L \in \mathbb{N} \setminus \{1\}. \tag{5.86}$$

In consequence, (5.74) *holds, i.e.*

$$\max_{\ell=1,\dots,L} \mathbf{E}\Big[\max_{t=0,1,\dots,T} \Big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_t^\ell(\widehat{\theta}^{\otimes\infty})\Big)^2\Big] = O(L^{-1}) \longrightarrow 0 \quad \textit{as } L \to \infty.$$

Remark 5.13. Because the MVPS problem has an origin in finance, we provide a financial interpretation of the above result. From an investment perspective, it is natural to compare two strategies $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ and $\widehat{\theta}^{\ell,\otimes\infty}$ via their respective gains processes. The above result then says that the difference of the gains processes between $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ and $\widehat{\theta}^{\ell,\otimes\infty}$ converges to 0 in the strong sense above as $L \to \infty$.

Proof of Theorem 5.12. We argue (5.86) by induction forward in time. The induction basis for $t = 0$ is trivial. Suppose the statement is true for $t-1$, i.e., $\max_{\ell=1,\ldots,L} \mathbf{E}\big[\big(\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})\big)^2\big] = O(L^{-1})$ for $L \in \mathbb{N} \setminus \{1\}$. For the induction step, we only need to show that

$$\max_{\ell=1,\ldots,L} \mathbf{E}[(\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{S}^\ell_t - \widehat{\theta}^{\ell,\otimes\infty}_t \triangle \mathbf{S}^\ell_t)^2] = O(L^{-1}) \quad \text{for } L \in \mathbb{N} \setminus \{1\}.$$

Note that the constant C in the bound (5.80) is independent of ℓ, L and t. So we take the maximum over $\ell = 1, \ldots, L$ on both sides of (5.80) to obtain

$$\begin{aligned}
&\max_{\ell=1,\ldots,L} \mathbf{E}[(\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_t \triangle \mathbf{S}^\ell_t - \widehat{\theta}^{\ell,\otimes\infty}_t \triangle \mathbf{S}^\ell_t)^2] \\
&\le 6C\mathbf{E}\bigg[\bigg(\frac{1}{b^{(L)}_{t-1}} - \frac{1}{b^{(\infty)}_{t-1}}\bigg)^2\bigg] + 6C \max_{\ell=1,\ldots,L} \mathbf{E}\big[\big(\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})\big)^2\big] \\
&\quad + 6C \max_{\ell=1,\ldots,L} \mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - \mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})]\Big)^2\Big] \\
&\quad + \frac{2C}{L^2} \max_{\ell=1,\ldots,L} \mathbf{E}\bigg[\bigg(\frac{1}{2b^{(\infty)}_{t-1}} - \mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty}) + \mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})]\bigg)^2\bigg] \\
&=: D^{(L)}_1 + D^{(L)}_2 + D^{(L)}_3 + D^{(L)}_4.
\end{aligned} \tag{5.87}$$

For $D^{(L)}_1$, we use (5.75) to get

$$\max_{t=0,1,\ldots,T} \left| \frac{1}{b^{(L)}_t} - \frac{1}{b^{(\infty)}_t} \right| = O(L^{-1})$$

and hence $D^{(L)}_1 = O(L^{-1})$ for $L \in \mathbb{N}\setminus\{1\}$. For $D^{(L)}_4$, note that $\widehat{\theta}^{\ell,\otimes\infty}(\omega^{(\infty)}) = \widehat{\theta}(\omega_\ell)$ and hence $\mathbf{G}^\ell(\widehat{\theta}^{\otimes\infty})(\omega^{(\infty)}) = G(\widehat{\theta})(\omega_\ell)$ by (5.48). Then using $P = \mathbf{P} \circ \pi^{-1}_{\ell,\infty}$ and $G_t(\widehat{\theta}) \in L^2(P)$ by Lemma 5.5, we have

$$\begin{aligned}
D^{(L)}_4 &= \frac{2C}{L^2} \max_{\ell=1,\ldots,L} \mathbf{E}\bigg[\bigg(\frac{1}{2b^{(\infty)}_{t-1}} - \mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty}) + \mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})]\bigg)^2\bigg] \\
&\le \frac{2C}{L^2} \max_{\ell=1,\ldots,L} \bigg(2\Big(\frac{1}{2b^{(\infty)}_{t-1}}\Big)^2 + 2\mathrm{Var}^{\mathbf{P}^{(\infty)}}[\mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})]\bigg) \\
&= \frac{4C}{L^2}\bigg(\Big(\frac{1}{2b^{(\infty)}_{t-1}}\Big)^2 + \mathrm{Var}[G_{t-1}(\widehat{\theta})]\bigg) \\
&= O(L^{-2}).
\end{aligned}$$

For $D_2^{(L)}$, we use the induction hypothesis to obtain

$$\max_{\ell=1,\dots,L} \mathbf{E}\big[\big(\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})\big)^2\big] = O(L^{-1}) \tag{5.88}$$

for $L \in \mathbb{N} \setminus \{1\}$. Now, for $D_3^{(L)}$, we obtain from (5.88) and Jensen's inequality that

$$\begin{aligned}
&\max_{\ell=1,\dots,L} \Big(\big(\mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})] - \mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})]\big)^2\Big) \\
&\le \max_{\ell=1,\dots,L} \mathbf{E}\big[\big(\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})\big)^2\big] \\
&= O(L^{-1})
\end{aligned} \tag{5.89}$$

for $L \in \mathbb{N} \setminus \{1\}$. Next, we use $(x+y)^2 \le 2x^2 + 2y^2$, then (5.76) and (5.89) to get for $L \in \mathbb{N} \setminus \{1\}$ that

$$\begin{aligned}
&\max_{\ell=1,\dots,L} \mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - \mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})]\Big)^2\Big] \\
&\le 2 \max_{\ell=1,\dots,L} \mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - \mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})]\Big)^2\Big] \\
&\quad + 2 \max_{\ell=1,\dots,L} \Big(\mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\boldsymbol{\vartheta}}^{(L)})] - \mathbf{E}[\mathbf{G}^\ell_{t-1}(\widehat{\theta}^{\otimes\infty})]\Big)^2 \\
&= O(L^{-1}).
\end{aligned} \tag{5.90}$$

This shows that in (5.87), we have $D_1^{(L)} + D_2^{(L)} + D_3^{(L)} + D_4^{(L)} = O(L^{-1})$ for $L \in \mathbb{N} \setminus \{1\}$ and thus completes the induction step.

For (5.74), we observe that

$$\begin{aligned}
&\max_{\ell=1,\dots,L} \mathbf{E}\Big[\max_{t=1,\dots,T} \big(\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}^\ell_t(\widehat{\theta}^{\otimes\infty})\big)^2\Big] \\
&\le \max_{\ell=1,\dots,L} \mathbf{E}\Big[\sum_{t=1}^{T} \big(\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}^\ell_t(\widehat{\theta}^{\otimes\infty})\big)^2\Big].
\end{aligned}$$

The latter term is still $O(L^{-1})$ as $L \to \infty$ because each summand is $O(L^{-1})$. This proves (5.74). □

We end this chapter with an alternative proof of Theorem 5.7. In view of Proposition 5.3, we again show that

$$\lim_{L\to\infty} \mathbf{E}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] = E[J_T^{\mathrm{mv}}(\widehat{\theta})]. \tag{5.91}$$

Instead of using $\mathbf{E}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] = V_0^{(L)}$ and showing $V_0^{(L)} \to E[J_T^{\mathrm{mv}}(\widehat{\theta})]$ as $L \to \infty$, we directly argue that

$$\begin{aligned}\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] &\longrightarrow E[G_T(\widehat{\theta})],\\ \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})^2\big)\big] &\longrightarrow E\big[\big(G_T(\widehat{\theta})\big)^2\big],\\ \mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big] &\longrightarrow \big(E[G_T(\widehat{\theta})]\big)^2\end{aligned}$$

as $L \to \infty$. This also shows (5.91) and hence provides a different proof of Theorem 5.7.

Corollary 5.14. *Suppose Assumptions* 3.2, 3.4 *and* 4.11 *are satisfied. Then*

$$\mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - \mathbf{E}[\mathbf{G}_T^{\ell}(\widehat{\theta}^{\otimes\infty})]\Big)^2\Big] \longrightarrow 0 \quad \textit{as } L \to \infty. \tag{5.92}$$

Proof. Estimate similarly as in (5.90) and use (5.76) and Theorem 5.12. □

Corollary 5.15. *Suppose Assumptions* 3.2, 3.4 *and* 4.11 *are satisfied. Then*

$$\begin{aligned}\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] &\longrightarrow E[G_T(\widehat{\theta})] \quad \textit{as } L \to \infty,\\ \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})^2\big)\big] &\longrightarrow E\big[\big(G_T(\widehat{\theta})\big)^2\big] \quad \textit{as } L \to \infty,\\ \mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big] &\longrightarrow \big(E[G_T(\widehat{\theta})]\big)^2 \quad \textit{as } L \to \infty.\end{aligned}$$

In particular, (5.91) *holds and thus $\widehat{\theta}$ is an optimal strategy for the MVPS problem* (5.14).

Proof. Using $\mathbf{G}^{\ell}(\widehat{\theta}^{\otimes\infty})(\omega^{(\infty)}) = G(\widehat{\theta})(\omega_\ell)$ and $P = \mathbf{P} \circ \pi_{\ell,\infty}^{-1}$ by (5.48) and (2.12) respectively, we only need to show

$$\begin{aligned}\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] &\longrightarrow \mathbf{E}[\mathbf{G}_T^{\ell}(\widehat{\theta}^{\otimes\infty})] \quad \text{as } L \to \infty,\\ \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})^2\big)\big] &\longrightarrow \mathbf{E}\big[\big(\mathbf{G}_T^{\ell}(\widehat{\theta}^{\otimes\infty})\big)^2\big] \quad \text{as } L \to \infty,\\ \mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big] &\longrightarrow \big(\mathbf{E}[\mathbf{G}_T^{\ell}(\widehat{\theta}^{\otimes\infty})]\big)^2 \quad \text{as } L \to \infty.\end{aligned}$$

The first two convergence results follow from

$$\mathbf{E}\big[\big(\mathbf{G}_T^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^i\big] \longrightarrow \mathbf{E}\big[\big(\mathbf{G}_T^{\ell}(\widehat{\theta}^{\otimes\infty})\big)^i\big] \quad \text{for } \ell \in \mathbb{N}, i \in \{1,2\}, \text{ as } L \to \infty$$

by Theorem 5.12 and a Cesàro-type statement that if $a_n \to a$ as $n \to \infty$, then $\frac{1}{N}\sum_{n=1}^N a_n \to a$ as $N \to \infty$. For the last convergence, we use (5.92) and the generic bound

$$|\mathbf{E}[x^2 - y^2]| = |\mathbf{E}[(x-y)(x+y)]| \leq \sqrt{\mathbf{E}[(x-y)^2]}\sqrt{\mathbf{E}[(x+y)^2]}$$

with $x = \mathrm{em}(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)}))$ and $y = \mathbf{E}[\mathbf{G}_T^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})]$. □

6 Connections to the literature

In this final section, we comment on related work in the literature. We give an overview of papers that attack the MVPS problem in finite discrete time. The first work on portfolio selection under a mean–variance criterion can be found in the classic paper of Markowitz [49] in a single-period setting. Due to its variance term, the multi-period formulation of the MVPS problem as in (2.3) is much more difficult and a first breakthrough appears nearly 50 years later in Li and Ng [45]. We recall from Section 1 the three approaches for the multi-period MVPS problem:

(a) The equilibrium approach goes away from the original formulation and studies a different type of optimality for which a solution satisfies a dynamic programming principle.

(b) The embedding approach connects the MVPS problem to a class of auxiliary hedging problems which turn out to be standard linear–quadratic stochastic control (LQSC) problems. One solves these LQSC problems and translates the resulting solutions into a solution for the original MVPS problem.

(c) The mean-field approach views the MVPS problem as a special case of a McKean–Vlasov control problem and uses tools from there to tackle it.

We do not discuss (a) here. For related work in (b), the pioneering paper is Li and Ng [45] which embeds the original MVPS problem into a family of auxiliary LQSC problems. This embedding technique is later used widely to study generalised MVPS problems (e.g. including the running mean and variance) whose underlying dynamics is driven by i.i.d. innovations. These i.i.d. innovations sometimes have parameters described by Markov chains. For various levels of generalisation along this line, we refer to Costa and de Oliveira [21] and He et al. [34] and references therein. Sun and Wang [65] and Fontana and Schweizer [30], with slightly different formulations, later show that one can write the solution to the MVPS problem in terms of the solution to the pure hedging problem to

$$\text{minimise } E\big[\big(1 - G_T(\theta)\big)^2\big] \text{ over } \theta \in \Theta, \tag{6.1}$$

which has a lot of results to invoke. The hedging problem (6.1) in finite discrete time is first fully worked out by Schweizer [61] with $\Theta = \Theta_{\mathrm{S}}$ given as in (3.6) even before Li and Ng [45]. Various papers later, e.g. Gugushvili [32], Černý [15], Melnikov and Nechaev [50], Černý and Kallsen [19] extended the results to more general spaces of strategies than Θ_{S}. But the explicit structure of the solution for (6.1) with various spaces Θ remains the same as in [61]. Note that this represents one end of the spectrum, i.e. a complicated recursive expression of an optimal strategy for the pure hedging problem (6.1). The other end of the spectrum is that we also know the explicit results for the pure hedging problem (6.1) from e.g. Schweizer [61] when the mean–variance tradeoff process is deterministic (This covers the case considered in Theorem 5.7.) However, there seem to be much fewer results in the middle. For an attempt in this direction in finite discrete time, we refer to Černý and Kallsen [16] and Hubelak et al. [36] for results that refine the general theory by expressing all exogenous coefficients of the solution in closed form in a Lévy-type setup, and to Kallsen et al. [41] where the authors obtain semi-explicit results for general affine stochastic volatility models.

This chapter is inspired by the market cloning approach proposed by Ankirchner and Dermoune [5]. A simple Google Scholar search reveals that this approach seems to remain almost unknown to the community, although it has a continuous-time extension by Fischer and Livieri [29]. Our results seem to be among the first attempts to further develop this approach in finite discrete time. Because this approach has more the flavour of (c), we round off this section with a brief discussion of some related work there. Andersson and Djehiche [4] first obtain and use a stochastic maximum principle from McKean–Vlasov control theory to solve the MVPS problem (in continuous time). In finite discrete time, Pham and Wei [52] develop a dynamic programming principle (DPP) for McKean–Vlasov control problems whose controlled process is driven by i.i.d. innovations, and apply the resulting DPP to solve the MVPS problem in this restricted setup. A similar result to our work is Basei and Pham [8] whose verification result in Lemma 3.1 is analogous to our Lemma 5.6.

Chapter II

Mean field approach for MVPS – continuous time

1 Introduction

This chapter studies the mean–variance portfolio selection (MVPS) problem in continuous time. A mathematical formulation of this problem is as follows. Let $G_T(\theta) = \int_0^T \theta_s \, \mathrm{d}S_s$ be the final wealth of a self-financing strategy trading in underlying price processes S from initial capital 0. We are interested in the problem to

$$\text{maximise } E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)] \text{ over } \theta \in \Theta$$

for a suitable set Θ of stochastic processes and a risk aversion parameter $\xi > 0$. We have studied this problem in finite discrete time in Chapter I and now replace the temporal structure $\mathbb{T} = \{0, \dots, T\}$ by the interval $\mathbb{T} = [0, T]$; so the present chapter is a natural sequel to Chapter I. Our main contribution here is to further develop the market cloning technique originally proposed by Ankirchner and Dermoune [5] and extended by Fischer and Livieri [29], and to apply that to solve the MVPS problem in continuous time.

As in finite discrete time, the main idea is to attack the problem by constructing an extended market which supports i.i.d. copies of the original financial market, then solve in that extended market an auxiliary but now standard stochastic control problem, and finally pass to the limit as the number of copies goes to infinity, to obtain an optimal strategy for the original MVPS problem. This approach, due to Ankirchner and Dermoune [5] and Fischer and Livieri [29], is further extended to continuous time in two aspects. First, we build a basic framework within which the market cloning technique can be performed for general semimartingales in $\mathcal{S}^2_{\mathrm{loc}}$. Second, in this more general framework, we apply the market cloning technique to study the MVPS problem with a *continuous* price process S, hence going beyond the Brownian-driven stochastic differential equation (SDE) models. In fact, we adopt a top-down approach and do not specify any dynamics for S. Because the market cloning technique has more the flavour of McKean–Vlasov-type control problems which often study Brownian-driven SDE models, this chapter is a nice addition to the current literature.

This chapter is structured as follows. Section 2 consists of results recalled and adapted from Chapter I for the market, the MVPS problem and the market cloning technique. We first recall the market and the MVPS problem from Chapter I. Then we recall the construction of an extended market which supports i.i.d. copies of the original market and the auxiliary problem in that extended market. We modify the filtration from Chapter I to obtain a filtration $\mathbb{G}$ for the extended market in order to invoke standard results from stochastic calculus in continuous time. The section ends with a version of the martingale optimality principle (MOP) which is used in Section 4 as a verification tool for solving the auxiliary problem.

In Section 3, we present a concrete setup where we solve the auxiliary problem later in Section 4. First we recall from Schweizer [59] a classic framework for studying the original MVPS problem. Based on the results for shrinkage of filtrations from Chapter I, we translate this framework into the extended market with the modified filtration $\mathbb{G}$. This allows us later to study the auxiliary problem with more general processes than Brownian-driven SDE models.

Section 4 proposes and implements a recipe for the construction of a candidate for the value process family for the auxiliary problem. This programme is different from Chapter I because solving a sequence of one-step (conditional) problems is no longer possible in continuous time. We first make the educated guess that the value process has the same affine–quadratic structure as in finite discrete time. Assuming that the underlying price process S is continuous, we then compute the canonical decomposition for the value process. Using the MOP from Section 2 as a tool, we heuristically derive from the martingale/supermartingale properties some differential equations for the coefficients in the affine–quadratic expression for the value process. While these differential equations can always be solved under mild assumptions, proving that the guessed affine–quadratic structure satisfies the martingale/supermartingale properties can at present only be done under the extra assumption that the mean–variance tradeoff process for the price process S is deterministic. Assuming this extra condition, we give in the main result of this section (Theorem 4.14) an explicit optimal strategy for the auxiliary problem.

In Section 5, we construct an optimal strategy $\widehat{\theta}$ for the MVPS problem with the help of the optimal strategies from Theorem 4.14 for the auxiliary problems, and we study the convergence behaviour of the latter strategies as the number of copies go to infinity. We formally take the limit of the expressions for the optimal strategies from Theorem 4.14 to obtain a formula for $\widehat{\theta}$ and show that the values of these auxiliary problems converge to the time-0 value of the MVPS criterion of $\widehat{\theta}$. Thanks to the verification result from Chapter I, this already shows the optimality of $\widehat{\theta}$ for the MVPS problem. Then we show that the gains of the optimal strategy for the auxiliary problems converge to the gains of $\widehat{\theta}$ as the number of copies goes to infinity, with a precise rate of convergence. This extends and improves the corresponding results in Fischer and Livieri [29].

Finally in Section 6, we discuss the connection to other literature in detail.

2 Problem formulation and preliminaries

2.1 MVPS problem in continuous time

In this section, we recall and adapt (in continuous time) the financial market, the mean–variance portfolio selection (MVPS) problem and the result for existence and uniqueness of its solution, as developed in Section I.2.1.

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space with a finite time horizon $T > 0$ and such that $\mathcal{F}_0$ is P-trivial. As a standard assumption in continuous time, we assume that the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions of right-continuity and P-completeness.

Consider a financial market with d risky assets and 1 riskless asset. All prices are discounted by the riskless asset and expressed by units of 1. The price of the riskless asset is given by $S^0 = (S^0_t)_{t\in[0,T]} \equiv 1$, and the discounted prices of the risky assets are given by an $\mathbb{R}^d$-valued stochastic process $S = (S_t)_{t\in[0,T]}$ adapted to the filtration $\mathbb{F}$.

To mathematically discuss trading activities in this financial market, we assume that S is a semimartingale. A (self-financing) *trading strategy* is a pair (v_0, θ), where $v_0 \in \mathbb{R}$ is the initial capital and θ is an $\mathbb{R}^d$-valued predictable process with respect to $\mathbb{F}$ such that the wealth process of (v_0, θ) is given by

$$V_t(v_0, \theta) = v_0 + G_t(\theta) := v_0 + \int_0^t \theta_s \,\mathrm{d}S_s, \quad t \in [0, T].$$

Note that this includes the requirement that θ is S-integrable so that the (real-valued) stochastic integral process $\int \theta \, \mathrm{d}S$ is well defined.

Let Θ be a set of processes such that for any $v_0 \in \mathbb{R}$ and $\theta \in \Theta$, the pair (v_0, θ) is a self-financing strategy. Fix a generic risk aversion constant $\xi > 0$. We recall from (I.2.3) an equivalent form of the MVPS problem *in continuous time* as to

$$\text{maximise } E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)] \text{ over all } \theta \in \Theta. \tag{2.1}$$

Assumption 2.1. Θ satisfies the following properties:

1) $G_T(\Theta) := \{G_T(\theta) : \theta \in \Theta\}$ is a closed subspace of L^2, i.e., $G_T(\Theta)$ is a linear space and $G_T(\Theta)$ is closed in L^2.

2) The constant payoff 1 is not in the L^2-closure of $G_T(\Theta)$. In view of 1), this is equivalent to $1 \notin G_T(\Theta)$.

We recall from Theorem I.2.4 the result about the existence and uniqueness of a solution to the MVPS problem. Note that it holds irrespective of the underlying temporal structure and hence requires no additional proof.

Theorem 2.2. *Suppose that Assumption* 2.1 *is satisfied. Then the MVPS problem* (2.1) *has a maximiser* $\widehat{\theta} \in \Theta$, *and the resulting* $G_T(\widehat{\theta})$ *is unique.*

2.2 An auxiliary problem in continuous time

As in the previous chapter, we are interested in a dynamic description of a maximiser $\widehat{\theta}$ to the MVPS problem (2.1). To this end, we recall the extended market and the auxiliary problem discussed in Section I.2.2 and adapt those constructions accordingly in continuous time.

Fix $L \in \mathbb{N} \cup \{\infty\}$. Following the construction in Section I.2.2, we can construct a probability space $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$ supporting L independent filtrations $(\mathbb{F}^{\ell,(L)})_{\ell=1,\dots,L}$ and processes $(\mathbf{S}^{\ell,(L)})_{\ell=1,\dots,L}$ such that each $\mathbf{S}^{\ell,(L)} = (\mathbf{S}_t^{\ell,(L)})_{t \in [0,T]}$ is adapted to $\mathbb{F}^{\ell,(L)}$ and has the same law as S. For $L = \infty$, we mean by $\ell = 1, \dots, L$ that we consider a sequence indexed by $\ell \in \mathbb{N}$.

So far, all the ingredients require only formal adjustments of notations from Chapter I. A first technical issue is brought by the filtration $\mathbb{G}^{(L)}$ used in Chapter I; see (I.2.17). We relabel $\mathbb{G}^{(L)}$ there by $\widetilde{\mathbb{G}}^{(L)}$ here, which is given explicitly by

$$\widetilde{\mathcal{G}}_t^{(L)} := \sigma\bigg(\bigcup_{\ell=1}^{L} \mathcal{F}_t^{\ell,(L)} \bigg), \quad t \in [0, T]. \tag{2.2}$$

To invoke standard results from stochastic calculus in continuous time, we wish to work with a filtration that satisfies the usual conditions. Although the filtration $\mathbb{F}$ is assumed to have this property, we argue below that the usual conditions are not automatically carried over to the filtration $\widetilde{\mathbb{G}}^{(L)}$ as long as the underlying probability space is rich enough.

Lemma 2.3. *Suppose that $L \in \mathbb{N}$ is finite. If there exists $F \in \mathcal{F}_0$ such that $P[F] > 0$ (and because $\mathcal{F}_0$ is P-trivial, this means $P[F] = 1$) and F contains a nonmeasurable set $\widetilde{F}$, then the filtration $\widetilde{\mathbb{G}}^{(L)}$ is not complete.*

Proof. We first show that $\widetilde{\mathcal{G}}_t^{(L)}$ is equal to the L-fold product σ-algebra $\otimes_{\ell=1}^L \mathcal{F}_t$ formed by $\mathcal{F}_t$. Indeed, let $\mathcal{C}$ be the class of finite intersections of sets in $\cup_{\ell=1}^L \mathcal{F}_t^{\ell,(L)}$; then clearly $\widetilde{\mathcal{G}}_t^{(L)} = \sigma(\mathcal{C})$ by (2.2). Recall $\mathcal{F}_t^{\ell,(L)} = \{\pi_{\ell,L}^{-1}(E) : E \in \mathcal{F}_t\}$ from (I.2.14), where $\pi_{\ell,L} : \Omega^{(L)} \to \Omega$ is the canonical projection onto the ℓ-th coordinate. We then use

$$\bigcap_{\ell=1}^{L} \pi_{\ell,L}^{-1}(E_\ell) = E_1 \times E_2 \times \cdots \times E_L, \quad E_\ell \in \mathcal{F}_t,$$

to obtain $\mathcal{C} = \{E_1 \times \cdots \times E_L : E_\ell \in \mathcal{F}_t, \ell = 1, \ldots, L\}$. This then yields by definition $\sigma(\mathcal{C}) = \otimes_{\ell=1}^L \mathcal{F}_t$ and hence the claim.

Now we show that $\otimes_{\ell=1}^L \mathcal{F}_0$ is not complete. Without loss of generality, we assume $L = 2$. For $\Lambda \subseteq \Omega \times \Omega$, we consider $\Lambda_x := \{y \in \Omega : (x, y) \in \Lambda\}$. A standard result from measure theory states that

$$\Lambda_x \in \mathcal{F}_0 \quad \text{for all } \Lambda \in \otimes_{\ell=1}^2 \mathcal{F}_0 = \mathcal{F}_0 \otimes \mathcal{F}_0 \text{ and } x \in \Omega; \tag{2.3}$$

see e.g. Salamon [57, Lemma 7.2]. Suppose for a contradiction that $\mathcal{F}_0 \otimes \mathcal{F}_0$ is complete. Let $N \subseteq \Omega$ be a nonempty P-null set and $\widetilde{F}$ a nonmeasurable set contained in F. Then from $N \times \widetilde{F} \subseteq N \times F$ and $\mathbf{P}^{(2)}[N \times F] = P[N]P[F] = 0$ because $\mathbf{P}^{(2)}$ is the product measure, we get that $N \times \widetilde{F} \in \mathcal{F}_0 \otimes \mathcal{F}_0$ due to the completeness of $\mathcal{F}_0 \otimes \mathcal{F}_0$. By (2.3) and because $\widetilde{F} = (N \times \widetilde{F})_x$ for any $x \in N$, we obtain $\widetilde{F} \in \mathcal{F}_0$. But $\widetilde{F}$ is nonmeasurable by assumption, and this is a contradiction. Note that we need $P[F] > 0$ because if $P[F] = 0$, then $\widetilde{F} \subseteq F$ must be in $\mathcal{F}_0$ because $\mathcal{F}_0$ if P-complete. □

Because of the above result, we work with $\mathbb{G}^{(L)}$ — the standard augmented filtration of $\widetilde{\mathbb{G}}^{(L)}$. To give a precise definition of that, we introduce some notations. For two classes $\mathcal{A}, \mathcal{B}$ of sets, we set $\mathcal{A} \vee \mathcal{B} := \sigma(\mathcal{A}, \mathcal{B})$. For a filtration $(\mathcal{H}_t)_{t\in[0,T]}$, we denote its right-continuous version by $\mathcal{H}_t^+ := \bigcap_{\varepsilon>0} \mathcal{H}_{t+\varepsilon}$. The augmented filtration $\mathbb{G}^{(L)}$ of $\widetilde{\mathbb{G}}^{(L)}$ is given by

$$\mathcal{G}_t^{(L)} = (\widetilde{\mathcal{G}}_t^{(L)} \vee \mathcal{N}^{(L)})^+, \tag{2.4}$$

where $\mathcal{N}^{(L)}$ is the class of $\mathbf{P}^{(L)}$-null sets in the underlying σ-algebra $\widetilde{\mathcal{G}}_T^{(L)}$. Once the filtration $\mathbb{G}^{(L)}$ is fixed, the rest of this subsection goes completely in parallel to Chapter I thanks to the general presentation given in Section I.2.2. For convenience, we recall the necessary notations in order to present the auxiliary problem (I.2.25). Defined as in (I.2.19), $\mathbf{\Theta}^{(L)}$ consists of all $\mathbb{R}^{d\times L}$-valued, $\mathbb{G}^{(L)}$-predictable processes $\boldsymbol{\vartheta}^{(L)}$ such that each coordinate $\boldsymbol{\vartheta}^{\ell,(L)}$ satisfies the integrability condition of Θ. Strategies from $\mathbf{\Theta}^{(L)}$ are used to make investments in $\mathbf{S}^{(L)} = (\mathbf{S}^{\ell,(L)})_{\ell=1,\dots,L}$. Given $\boldsymbol{\vartheta}^{(L)} \in \mathbf{\Theta}^{(L)}$, its (vector) gains process is given by

$$\mathbf{G}_t^{\ell}(\boldsymbol{\vartheta}^{(L)}) = \int_0^t \boldsymbol{\vartheta}_s^{\ell,(L)} \,\mathrm{d}\mathbf{S}_s^{\ell,(L)}, \quad \ell = 1,\dots,L,\ t \in [0,T]. \tag{2.5}$$

In view of Definition I.2.5, by an L-extended market, we still mean the tuple $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)}, \mathbb{G}^{(L)}, \mathbf{S}^{(L)})$ and write $(\mathbf{P}^{(L)}, \mathbb{G}^{(L)}, \mathbf{S}^{(L)})$ whenever the underlying probability space is clear from the context. Recall from (I.2.23) and (I.2.24) for $\mathbf{x}^{(L)}, \mathbf{y}^{(L)}$ in $\mathbb{R}^L$ the operations

$$\mathbf{x}^{(L)} \odot \mathbf{y}^{(L)} = (\mathbf{x}^{\ell,(L)}\mathbf{y}^{\ell,(L)})_{\ell=1,\dots,L}, \quad (\mathbf{x}^{(L)})^2 = \mathbf{x}^{(L)} \odot \mathbf{x}^{(L)}, \tag{2.6}$$

$$\mathrm{em}(\mathbf{x}^{(L)}) = \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{x}^{\ell,(L)}, \tag{2.7}$$

$$\mathrm{evar}(\mathbf{x}^{(L)}) = \mathrm{em}\big((\mathbf{x}^{(L)})^2\big) - \big(\mathrm{em}(\mathbf{x}^{(L)})\big)^2. \tag{2.8}$$

The auxiliary problem — a standard stochastic control problem — of interest in the L-extended market is still to

$$\text{maximise } \mathbf{E}^{(L)}\big[\mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big)\big] \quad \text{over } \boldsymbol{\vartheta}^{(L)} \in \mathbf{\Theta}^{(L)}, \tag{2.9}$$

where $\mathbf{E}^{(L)}$ denotes the expectation under the measure $\mathbf{P}^{(L)}$ and $\xi > 0$ as before is a risk aversion parameter. As we argued in Section I.2.3, the auxiliary problem (2.9) is a standard stochastic control problem and enjoys a martingale optimality principle as in Lemma I.2.11 *provided* that (I.2.10) for aggregation is satisfied. In contrast to Chapter I, solving (2.9) via a sequence of one-step (conditional) problems is no longer possible because we have a continuum of time steps in continuous time. Instead, we proceed by guessing and deriving conditions on the value process from the martingale optimality principle, and then constructing candidates for both the value process of the auxiliary problem (2.9) and its optimal strategy directly based on our guess and derived conditions. Finally, we use the martingale optimality principle as a tool to verify that our candidates are the true value process and optimal strategy.

After an optimal strategy $\widehat{\boldsymbol{\vartheta}}^{(L)}$ to the auxiliary problem (2.9) is obtained, the programme becomes again parallel to Chapter I. We prefer to give only a high-level overview here and refer the interested reader to Section I.5.2 for results in detail. Let us recall from (I.2.26)–(I.2.29) the handy notations

$$J_T^{\mathrm{mv}}(\theta) := G_T(\theta) - \xi\big(G_T(\theta) - E[G_T(\theta)]\big)^2, \quad \theta \in \Theta, \tag{2.10}$$
$$J_T^{(L)}(\boldsymbol{\vartheta}^{(L)}) := \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big), \quad \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}, \tag{2.11}$$

respectively. Using (2.10) and (2.11), we can equivalently write the MVPS and auxiliary problems as to

$$\text{maximise } E[J_T^{\mathrm{mv}}(\theta)] \text{ over all } \theta \in \Theta, \tag{2.12}$$
$$\text{maximise } \mathbf{E}^{(L)}[J_T^{(L)}(\boldsymbol{\vartheta}^{(L)})] \text{ over all } \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}^{(L)}, \tag{2.13}$$

respectively. The idea is to build a link between the MVPS and the auxiliary problem. As a first step, we construct quantities so that the MVPS criterion J^{mv} makes sense in the L-extended market. For this, we can simply define, for $\ell = 1, \dots, L$,

$$J_T^{\mathrm{mv},\ell}(\boldsymbol{\vartheta}^{(L)}) := \mathbf{G}_T^{\ell}(\boldsymbol{\vartheta}^{(L)}) - \xi\big(\mathbf{G}_T^{\ell}(\boldsymbol{\vartheta}^{(L)}) - \mathbf{E}^{(L)}[\mathbf{G}_T^{\ell}(\boldsymbol{\vartheta}^{(L)})]\big)^2.$$

Now we recall from (I.2.18) the generic lifting operation for a process X given by $X^{\otimes L} = (X^{\ell,\otimes L})_{\ell=1,\dots,L}$ with

$$X^{\ell,\otimes L} := X \circ \pi_{\ell,L}, \quad \ell = 1, \dots, L. \tag{2.14}$$

Below we use the notation $\omega^{(L)}$ for an sample element of $\Omega^{(L)}$ and ω_ℓ for the ℓ-th coordinate of $\omega^{(L)}$. For any strategy $\theta \in \Theta$ and $L \in \mathbb{N}$, we can always use (2.14) to "lift" θ to the L-extended market via $\theta^{\otimes L}$ given by

$$\theta^{\ell,\otimes L}(\omega^{(L)}) = (\theta \circ \pi_{\ell,L})(\omega^{(L)}) = \theta(\omega_\ell)$$

for $\ell = 1, \ldots, L$. It turns out that the resulting gains also satisfy

$$\mathbf{G}_T^\ell(\theta^{\otimes L})(\omega^{(L)}) = \big(G_T(\theta) \circ \pi_{\ell,L}\big)(\omega^{(L)}) = G_T(\theta)(\omega_\ell). \tag{2.15}$$

Using this and $P = \mathbf{P}^{(L)} \circ \pi_{\ell,L}^{-1}$, we then have $E[J_T^{\mathrm{mv}}(\theta)] = \mathbf{E}^{(L)}[J_T^{\mathrm{mv},\ell}(\boldsymbol{\vartheta}^{(L),\theta})]$ for $\ell = 1, \ldots, L$.

Since $\theta^{\otimes L}$ lives in the L-extended market, we can of course also evaluate it against the auxiliary criterion by considering $\mathbf{E}^{(L)}[J_T^{(L)}(\theta^{\otimes L})]$. The key point here is to observe that the above lifting technique, or more precisely the identity $\mathbf{G}_T^\ell(\theta^{\otimes L})(\omega^{(L)}) = G_T(\theta)(\omega_\ell)$, produces i.i.d. random variables $(\mathbf{G}_T^\ell(\theta^{\otimes L}))_{\ell=1,\ldots,L}$, where each has the same distribution as $G_T(\theta)$, so that $\mathbf{E}^{(L)}[\mathbf{G}_T^\ell(\theta^{\otimes L})] = E[G_T(\theta)]$ and $\mathrm{Var}^{\mathbf{P}^{(L)}}[\mathbf{G}_T^\ell(\theta^{\otimes L})] = \mathrm{Var}[G_T(\theta)]$ for $\ell = 1, \ldots, L$. Thus using some form of a law of large numbers and the fact that the auxiliary criterion $J_T^{(L)}$ only involves the empirical averages and variances, we should obtain

$$\mathbf{E}^{(L)}[J_T^{(L)}(\theta^{\otimes L})] \longrightarrow E[G_T(\theta)] - \xi \mathrm{Var}[G_T(\theta)] = E[J_T^{\mathrm{mv}}(\theta)] \quad \text{as } L \to \infty.$$

The final piece for the link between the MVPS and the auxiliary problem comes from the observation that $\theta^{\otimes L}$ lives in the L-extended market and hence we have the bound $\mathbf{E}^{(L)}[J_T^{(L)}(\theta^{\otimes L})] \le \mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})]$ for any $L \in \mathbb{N}$ with $L \ge 2$ by the optimality of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for the auxiliary problem. Sending $L \to \infty$ and using the convergence result in the above display, we obtain

$$E[J_T^{\mathrm{mv}}(\theta)] \le \limsup_{L\to\infty} \mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] \quad \text{for all } \theta \in \Theta.$$

This bound gives a clue to both the construction and verification for an optimal strategy $\widehat{\theta}$ to the MVPS problem. Again from the perspective of the law of large numbers and the symmetry of the auxiliary problem (2.11), the quantity $\mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})]$ should converge to a mean–variance expression as $L \to \infty$. If we can construct $\widehat{\theta} \in \Theta$ such that $E[J_T^{\mathrm{mv}}(\widehat{\theta})]$ equals that limit, we can already claim that $\widehat{\theta}$ is an optimal strategy to the MVPS problem. To construct such a process $\widehat{\theta}$, we rely on the explicit formula for $\widehat{\boldsymbol{\vartheta}}^{(L)}$ and formally take limits as $L \to \infty$.

2.3 Martingale optimality principle: a verification tool

In this subsection, we present a version of the martingale optimality principle for the auxiliary problem (2.11) so that it can be used as an abstract tool to verify that a candidate for the value process of the auxiliary problem is the true value process. As we work only with the L-extended market below, we **drop the superscript** $^{(L)}$ for ease of notation.

We first recall from (I.2.32)–(I.2.34) some necessary notations to define the dynamic value process associated to the auxiliary problem (2.11). Fix $\mathbf{\Theta}$ which stands for an abstraction of a set of trading strategies in the extended market. For any $[0,T]$-valued $\mathbb{G}$-stopping time τ and $\boldsymbol{\vartheta} \in \mathbf{\Theta}$, recall

$$\mathbf{\Theta}(\tau,\boldsymbol{\vartheta}) = \{\widetilde{\boldsymbol{\vartheta}} \in \mathbf{\Theta} : \widetilde{\boldsymbol{\vartheta}} = \boldsymbol{\vartheta} \text{ on } [\![0,\tau]\!]\}, \tag{2.16}$$

$$J_\tau(\widetilde{\boldsymbol{\vartheta}}) = \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widetilde{\boldsymbol{\vartheta}})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\widetilde{\boldsymbol{\vartheta}})\big)\big|\mathcal{G}_\tau\big], \tag{2.17}$$

$$V_\tau(\boldsymbol{\vartheta}) = \operatorname*{ess\,sup}\big\{J_\tau(\widetilde{\boldsymbol{\vartheta}}) : \widetilde{\boldsymbol{\vartheta}} \in \mathbf{\Theta}(\tau,\boldsymbol{\vartheta})\big\}. \tag{2.18}$$

Note that $G_T(\boldsymbol{\vartheta})$ does not depend on $\boldsymbol{\vartheta}_0$, and hence neither do $J_\tau(\boldsymbol{\vartheta})$ nor $V_\tau(\boldsymbol{\vartheta})$. The verification result that we use to solve the auxiliary problem (2.9) is the version below of a martingale optimality principle.

Lemma 2.4. *Suppose that $(\widetilde{V}(\boldsymbol{\vartheta}))_{\boldsymbol{\vartheta}\in\mathbf{\Theta}}$ is a family of processes with the following properties:*

1) *For any $\boldsymbol{\vartheta} \in \mathbf{\Theta}$, we have $\widetilde{V}_T(\boldsymbol{\vartheta}) = J_T(\boldsymbol{\vartheta})$ and $\widetilde{V}_t(\boldsymbol{\vartheta}) = \widetilde{V}_t(\widetilde{\boldsymbol{\vartheta}})$ whenever $\widetilde{\boldsymbol{\vartheta}} \in \mathbf{\Theta}(t,\boldsymbol{\vartheta})$; so $\widetilde{V}_0(\boldsymbol{\vartheta})$ is independent of $\boldsymbol{\vartheta}$. This common value is denoted by $\widetilde{V}_0$.*
2) *For any $\boldsymbol{\vartheta} \in \mathbf{\Theta}$, the process $(\widetilde{V}_t(\boldsymbol{\vartheta}))_{t\in[0,T]}$ is a $\mathbb{G}$-supermartingale.*
3) *There is $\boldsymbol{\vartheta}^* \in \mathbf{\Theta}$ such that the process $(\widetilde{V}_t(\boldsymbol{\vartheta}^*))_{t\in[0,T]}$ is a $\mathbb{G}$-martingale.*

Then $\boldsymbol{\vartheta}^*$ is optimal for the auxiliary problem (2.9)*. In particular, $\widetilde{V}_t(\boldsymbol{\vartheta}^*) = V_t(\boldsymbol{\vartheta}^*)$ for each $t \in [0,T]$, and $\widetilde{V}_0 = V_0$.*

Proof. First, let us note from (2.18) or recall from (I.2.36) that

$$V_T(\boldsymbol{\vartheta}) = J_T(\boldsymbol{\vartheta}) \quad \text{for all } \boldsymbol{\vartheta} \in \mathbf{\Theta}. \tag{2.19}$$

Moreover, 1) gives

$$\widetilde{V}_T(\boldsymbol{\vartheta}) = J_T(\boldsymbol{\vartheta}) = V_T(\boldsymbol{\vartheta}) \quad \text{for all } \boldsymbol{\vartheta} \in \mathbf{\Theta}. \tag{2.20}$$

Let $\boldsymbol{\vartheta} \in \mathbf{\Theta}$. By (2.20) and the supermartingale property of $\widetilde{V}(\boldsymbol{\vartheta})$ from 2), we get

$$\mathbf{E}[J_T(\boldsymbol{\vartheta})] = \mathbf{E}[\widetilde{V}_T(\boldsymbol{\vartheta})] \le \mathbf{E}[\widetilde{V}_0(\boldsymbol{\vartheta})] = \widetilde{V}_0.$$

This inequality holds for all $\boldsymbol{\vartheta} \in \mathbf{\Theta}$, with equality attained at $\boldsymbol{\vartheta}^* \in \mathbf{\Theta}$ because of 3). So we take the supremum above over $\boldsymbol{\vartheta} \in \mathbf{\Theta}$ and use the martingale property of $\widetilde{V}(\boldsymbol{\vartheta}^*)$ in 3) and $\widetilde{V}_T(\boldsymbol{\vartheta}^*) = J_T(\boldsymbol{\vartheta}^*)$ in 1) to obtain

$$\sup_{\boldsymbol{\vartheta}\in\boldsymbol{\Theta}} \mathbf{E}[J_T(\boldsymbol{\vartheta})] = \sup_{\boldsymbol{\vartheta}\in\boldsymbol{\Theta}} \mathbf{E}[\widetilde{V}_T(\boldsymbol{\vartheta})] = \widetilde{V}_0 = \mathbf{E}[\widetilde{V}_T(\boldsymbol{\vartheta}^*)] = \mathbf{E}[J_T(\boldsymbol{\vartheta}^*)].$$

This shows that $\boldsymbol{\vartheta}^*$ is optimal. For the last statement, observe that $\widetilde{V}(\boldsymbol{\vartheta}^*)$ and $V(\boldsymbol{\vartheta}^*)$ are $\mathbb{G}$-martingales due to 3) and Lemma I.2.11, respectively. They share the same terminal value $J_T(\boldsymbol{\vartheta}^*)$ by (2.20). Taking conditional expectations with respect to $\mathcal{G}_t$ for $t \in [0,T]$ yields the claim. □

3 The auxiliary problem in continuous time

In this section, we elaborate on the auxiliary problem (2.9) to

$$\text{maximise } \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big)\big] \text{ over all } \boldsymbol{\vartheta} \in \boldsymbol{\Theta}$$

in continuous time. **Throughout this section**, $L \in \mathbb{N}$ with $L \geq 2$ is fixed. We present below a concrete setup where we can ultimately solve (2.9) in the next section.

3.1 A concrete setup in continuous time

To give concrete results for the MVPS problem (2.1) in continuous time, we present a continuous-time framework analogous to the one in Chapter I. This setup was first introduced in Schweizer [59]. We first present this in detail and choose a space Θ of strategies. Then we collect some basic properties of Θ and show that it is indeed a good choice for studying (2.1) in the sense that Θ satisfies Assumption 2.1 under which the existence and uniqueness result in Theorem 2.2 for the MVPS problem holds. All this is on $(\Omega, \mathcal{F}, \mathbb{F}, P)$.

Let $\mathcal{S}^2$ be the space of $\mathbb{R}^d$-valued semimartingales admitting a special semimartingale decomposition $X = X_0 + M^X + A^X$, where $M^X \in \mathcal{M}^2_{0,\mathrm{loc}}$ is a square-integrable martingale and A^X is a predictable process of square-integrable variation, both null at 0. Denote by $\mathcal{S}^2_{\mathrm{loc}}$ the localised class of $\mathcal{S}^2$.

To make a specific choice of Θ, we consider

$$\Theta_{\mathrm{S}} := \bigg\{\theta = (\theta_t)_{t\in[0,T]} : \theta \text{ is } \mathbb{R}^d\text{-valued, } \mathbb{F}\text{-predictable,} \\ S\text{-integrable and } G(\theta) = \int \theta \,\mathrm{d}S \in \mathcal{S}^2\bigg\}. \tag{3.1}$$

Assumption 3.1. *S is an $\mathbb{R}^d$-valued semimartingale in $\mathcal{S}^2_{\mathrm{loc}}$.*

The continuous-time analogue of the structure condition in Assumption I.3.4 needs a bit more preparation. Using Assumption 3.1, we write

$$S^i - S^i_0 = (M^S)^i + (A^S)^i =: M^i + A^i$$

for $i = 1, \dots, d$ for the canonical decomposition of S^i with respect to the filtration $\mathbb{F}$, and we use the abbreviation to M^i, A^i for ease of notation whenever the reference process is clear from the context. Thanks to Assumption 3.1 and in particular because M^i are locally square-integrable local martingales, the *angle brackets* $\langle M^i \rangle$ of M^i exist for $i = 1, \dots, d$; see e.g. Jacod and Shiryaev [37, Theorem I.4.2].

Assumption 3.2. *For $i = 1, \dots, d$, we have $A^i \ll \langle M^i \rangle$ with predictable density $\alpha^i = (\alpha^i_t)_{0 \le t \le T}$.*

Fix an increasing predictable RCLL process $B = (B_t)_{t \in [0,T]}$ with $\langle M^i \rangle \ll B$ for $i = 1, \dots, d$, e.g. $B = \sum_{i=1}^d \langle M^i \rangle$. By the Kunita-Watanabe inequality, this implies $\langle M^i, M^j \rangle \ll B$ for $i, j = 1, \dots, d$. Define predictable processes σ and γ by

$$\sigma^{ij}_t = \frac{\mathrm{d}\langle M^i, M^j \rangle_t}{\mathrm{d}B_t}, \quad i, j = 1, \dots, d, t \in [0,T], \tag{3.2}$$

$$\gamma^i_t = \alpha^i_t \sigma^{ii}_t = \frac{\mathrm{d}A^i_t}{\mathrm{d}B_t}, \quad i = 1, \dots, d, t \in [0,T]. \tag{3.3}$$

Definition 3.3. We say that S *satisfies the structure condition (SC)* if Assumptions 3.1 and 3.2 are satisfied and there exists a predictable $\mathbb{R}^d$-valued process $\lambda = (\lambda_t)_{t \in [0,T]}$ such that

$$\sigma_t \lambda_t = \gamma_t \qquad P\text{-a.s. for all } t \in [0,T] \tag{3.4}$$

and

$$K_T := \int_0^T \lambda_s^\top \gamma_s \, \mathrm{d}B_s = \int_0^T (\lambda_s)^\top \sigma_s \lambda_s \, \mathrm{d}B_s < \infty, \tag{3.5}$$

where σ and γ are defined in (3.2) and (3.3), respectively. We then call the increasing (and finite-valued) process $K := \int \lambda^\top \gamma \, \mathrm{d}B = \int \lambda^\top \sigma \lambda \, \mathrm{d}B$ the *mean–variance tradeoff (MVT) process* of S.

We can equivalently describe Θ_{S} given in (3.1) via some integrability conditions, which are easier to verify.

Definition 3.4. The space $L^2_{(\mathrm{loc})}(M)$ consists of all $\mathbb{R}^d$-valued $\mathbb{F}$-predictable processes θ such that the integral process $\int \theta^\top \sigma \theta \,\mathrm{d}B$ is (locally) integrable. The space $L^2_{(\mathrm{loc})}(A)$ consists of all $\mathbb{R}^d$-valued $\mathbb{F}$-predictable processes θ such that the integral process $\int |\theta^\top \gamma| \,\mathrm{d}B$ is (locally) square-integrable.

We use the following observations made in Schweizer [59].

Lemma 3.5. *Suppose that Assumptions* 3.1 *and* 3.2 *are satisfied. Then the following statements hold:*

1) *If $\theta \in L^2_{(\mathrm{loc})}(M)$, then the stochastic integral $\int \theta \,\mathrm{d}M$ is well defined, in $\mathcal{M}^2_{0(,\mathrm{loc})}$, and for $\psi \in L^2_{(\mathrm{loc})}(M)$,*

$$\left\langle \int \theta \,\mathrm{d}M, \int \psi \,\mathrm{d}M \right\rangle_t = \int_0^t \theta_s^\top \sigma_s \psi_s \,\mathrm{d}B_s, \qquad t \in [0,T]. \tag{3.6}$$

In particular, if $\theta \in L^2(M)$, then

$$E\left[\sup_{t\in[0,T]} \left| \int_0^t \theta_s \,\mathrm{d}M_s \right|^2\right] \leq 4E\left[\left(\int_0^T \theta_s \,\mathrm{d}M_s\right)^2\right] = 4E\left[\int_0^t \theta_s^\top \sigma_s \theta_s \,\mathrm{d}B_s\right]. \tag{3.7}$$

2) *If $\theta \in L^2_{(\mathrm{loc})}(A)$, then the process $\int \theta^\top \,\mathrm{d}A := \sum_{i=1}^d \int \theta^i \,\mathrm{d}A^i$ is well defined as a Lebesgue–Stieltjes integral, (locally) square-integrable, predictable and satisfies*

$$\int_0^t \theta_s^\top \,\mathrm{d}A_s = \int_0^t \theta_s^\top \gamma_s \,\mathrm{d}B_s, \qquad t \in [0,T]. \tag{3.8}$$

Proof. 1) The statement that the stochastic integral $\int \theta \,\mathrm{d}M$ is well defined and in $\mathcal{M}^2_{0(,\mathrm{loc})}$ for $\theta \in L^2_{(\mathrm{loc})}(M)$ is standard and can be found in Jacod and Shiryaev [37, Section I.4]. The equality (3.6) follows from Itô's isometry, see e.g. [37, (I.4.6)], and (3.2). Because $\int \theta \,\mathrm{d}M \in \mathcal{M}^2_0$ for $\theta \in L^2(M)$, we can apply Doob's inequality (see [37, I.1.43]) to obtain the first inequality in (3.7). Using Itô's isometry and inserting (3.6) into (3.7) yields the second equality in (3.7).

2) The statement that the process $\int \theta^\top \,\mathrm{d}A$ is a well-defined Lebesgue-Stieltjes integral and (locally) square-integrable and predictable is again standard. The equality (3.8) follows from the definition that $\int \theta^\top \,\mathrm{d}A = \sum_{i=1}^d \int \theta^i \,\mathrm{d}A^i$ and (3.3). □

Lemma 3.6. *If Assumptions* 3.1 *and* 3.2 *are satisfied, then*

$$\Theta_{\mathrm{S}} = L^2(M) \cap L^2(A). \tag{3.9}$$

If in addition S satisfies (SC) and the MVT process is bounded, then $\Theta_{\mathrm{S}} = L^2(M)$.

Proof. See Schweizer [59, Lemma 2]. Note that because K is increasing, it is bounded if and only if K_T is. □

Assumption 3.7. S satisfies the structure condition (SC) given in Definition 3.3.

We end this section with the result that the space Θ_{S} satisfies the premises of the existence and uniqueness result in Theorem 2.2, which is analogous to Lemma I.3.6 in the discrete time case.

Lemma 3.8. *Suppose that Assumption* 3.7 *is satisfied. If the MVT process K is bounded, then Θ_{S} satisfies Assumption* 2.1.

Proof. For Assumption 2.1, 1) that the space $G_T(\Theta_{\mathrm{S}})$ is closed in L^2, we refer to Monat and Stricker [51, Theorem 4.1].

For Assumption 2.1, 2) that the constant payoff 1 is not in $G_T(\Theta_{\mathrm{S}})$, the idea is similar to the proof of Lemma I.3.6 in discrete time. Again, we recall that a signed $G_T(\Theta_{\mathrm{S}})$-martingale measure is a signed measure Q such that $Q[\Omega] = 1$, $Q \ll P$ with $\frac{\mathrm{d}Q}{\mathrm{d}P} \in L^2$ and

$$E\left[\frac{\mathrm{d}Q}{\mathrm{d}P}\, g\right] = 0 \quad \text{for all } g \in G_T(\Theta_{\mathrm{S}}), \tag{3.10}$$

and we call $\mathcal{Q}$ the set of all signed $G_T(\Theta_{\mathrm{S}})$-martingale measures. Looking at the proof of Lemma I.3.6, we see that we only need to show $\mathcal{Q} \neq \emptyset$. For that, we consider $Z = \mathcal{E}(-\int \lambda \,\mathrm{d}M)$, where λ is given in (3.4), and claim that Q defined by $\frac{\mathrm{d}Q}{\mathrm{d}P} = Z_T$ is in $\mathcal{Q}$ under the assumption that the MVT process K is bounded. The proof of the last claim can be found below Corollary 16 in Schweizer [59]. We nevertheless give a proof here for completeness. Because K_T is bounded, Théorème II.2 of Lepingle and Mémin [44] shows that Z is a square-integrable martingale. This gives $Q[\Omega] = E[Z_T] = 1$ and $\frac{\mathrm{d}Q}{\mathrm{d}P} = Z_T \in L^2$ and hence establishes the first two properties in the definition of a $G_T(\Theta_{\mathrm{S}})$-martingale measure. It is left to check (3.10). To this end, we use the product formula, then $\mathrm{d}Z = -Z_{-}\lambda \,\mathrm{d}M$ from the definition of Z and finally (3.6), (3.8) and (3.4) to get

$\langle Z, \int \theta \, \mathrm{d}M \rangle = -\langle \int Z_- \lambda \, \mathrm{d}M, \int \theta \, \mathrm{d}M \rangle = -\int Z_- \theta^\top \, \mathrm{d}A$ and

$$\begin{aligned}
ZG(\theta) &= \int Z_- \theta \, \mathrm{d}S + \int G_-(\theta) \, \mathrm{d}Z + \Big[Z, \int \theta \, \mathrm{d}M\Big] + \Big[Z, \int \theta^\top \, \mathrm{d}A\Big] \\
&= \int Z_- \theta \, \mathrm{d}S + \Big\langle Z, \int \theta \, \mathrm{d}M \Big\rangle + \int G_-(\theta) \, \mathrm{d}Z \\
&\quad + \Big[Z, \int \theta \, \mathrm{d}M\Big] - \Big\langle Z, \int \theta \, \mathrm{d}M \Big\rangle + \Big[Z, \int \theta^\top \, \mathrm{d}A\Big] \\
&= \int Z_- \theta \, \mathrm{d}M + \int G_-(\theta) \, \mathrm{d}Z \\
&\quad + \Big[Z, \int \theta \, \mathrm{d}M\Big] - \Big\langle Z, \int \theta \, \mathrm{d}M \Big\rangle + \Big[Z, \int \theta^\top \, \mathrm{d}A\Big].
\end{aligned}$$

Because the first term can be written as $\int Z_- \, \mathrm{d}(\int \theta \, \mathrm{d}M)$ and $\int \theta \, \mathrm{d}M$ is a local martingale due to $\theta \in \Theta_{\mathrm{S}} \subseteq L^2(M)$ by Lemma 3.6, the first two terms are both stochastic integrals of a locally bounded integrand with respect to a local martingale and thus are local martingales themselves. The difference of the third and fourth terms is a local martingale by the definition of the angle bracket. Yoeurp's lemma (see Dellacherie and Meyer [25, Theorem VII.36]) implies that the last term is a local martingale. Therefore $ZG(\theta)$ is in $\mathcal{M}_{0,\mathrm{loc}}$ for any $\theta \in \Theta_{\mathrm{S}}$. Next we claim that $E[\sup_{t\in[0,T]} |Z_t G_t(\theta)|^2] < \infty$, which readily shows that $ZG(\theta) \in \mathcal{M}_0^2$ and hence yields (3.10) as $Z_T = \frac{\mathrm{d}Q}{\mathrm{d}P}$. For proving that claim, we apply the Cauchy–Schwarz inequality to get

$$\Big(E\Big[\sup_{t\in[0,T]} |Z_t G_t(\theta)|\Big]\Big)^2 \le E\Big[\sup_{t\in[0,T]} |Z_t|^2\Big] E\Big[\sup_{t\in[0,T]} |G_t(\theta)|^2\Big].$$

Due to $Z \in \mathcal{M}^2$ and Doob's inequality, we get $E[\sup_{t\in[0,T]} |Z_t|^2] \le 4E[Z_T^2] < \infty$ and thus only need to show $E[\sup_{t\in[0,T]} |G_t(\theta)|^2] < \infty$. Using Cauchy–Schwarz, then (3.7) and (3.8) and that B is increasing, and finally (3.9) from Lemma 3.6 and Definition 3.4, we obtain

$$\begin{aligned}
E\Big[\sup_{t\in[0,T]} |G_t(\theta)|^2\Big] &\le 2E\Big[\sup_{t\in[0,T]} \Big|\int_0^t \theta_s \, \mathrm{d}M_s\Big|^2\Big] + 2E\Big[\sup_{t\in[0,T]} \Big|\int_0^t \theta_s^\top \, \mathrm{d}A_s\Big|^2\Big] \\
&\le 8E\Big[\int_0^T \theta_s^\top \sigma_s \theta_s \, \mathrm{d}B_s\Big] + 2E\Big[\Big|\int_0^T |\theta_s^\top \gamma_s| \, \mathrm{d}B_s\Big|^2\Big] \\
&< \infty.
\end{aligned}$$

This completes the proof. □

3.2 A concrete setup for the auxiliary problem

The previous subsection sets up a good framework for studying the MVPS problem (2.1) in continuous time. In this subsection, we provide a suitable setup for studying the auxiliary problem (2.9). First, we specify a space of strategies in the extended market similarly to Θ_S. Then we study how the structure condition (SC) for the process S from Assumption 3.7 and Definition 3.3 as well as various other quantities translate into the extended market with respect to the filtration $\mathbb{G}^{(L)}$.

We start by recalling from Section I.2.2 some basic components of the extended market. Fix $L \in \mathbb{N}$. The underlying probability space $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$ is obtained by taking the L-fold product of the original probability space $(\Omega, \mathcal{F}, P)$. Indeed, $\Omega^{(L)} = \prod_{\ell=1}^{L} \Omega$ is the Cartesian product of Ω, the σ-algebra $\mathcal{F}^{(L)}$ is generated by all finite rectangles with $\mathcal{F}$-measurable sides, and the measure $\mathbf{P}^{(L)}$ is the standard L-fold product measure of P. More importantly, we construct from the filtration $\mathbb{F}$ a family $(\mathbb{F}^{\ell,(L)})_{\ell=1,\dots,L}$ of filtrations via

$$\mathcal{F}_t^{\ell,(L)} := \{\pi_{\ell,L}^{-1}(E) : E \in \mathcal{F}_t\}, \quad \ell = 1, \dots, L, t \in [0, T], \tag{3.11}$$

where $\pi_{\ell,L} : \Omega^{(L)} \to \Omega$ is the canonical projection of $\Omega^{(L)}$ onto its ℓ-th coordinate. The process $\mathbf{S}^{(L)}$ is defined by

$$\mathbf{S}_t^{\ell,(L)} := S_t \circ \pi_{\ell,L} = S^{\ell,\otimes L}, \quad \ell = 1, \dots, L, t \in [0, T]. \tag{3.12}$$

The last equality uses the lifting operation introduced in (2.14). Because $\mathbf{P}^{(L)}$ is the product measure (or see (I.2.15)), we get that

$$\text{the } \sigma\text{-algebras } (\mathcal{F}_t^{\ell,(L)})_{\ell=1,\dots,L} \text{ are } \mathbf{P}^{(L)}\text{-independent for any } t \in [0, T]. \tag{3.13}$$

From (3.11) and (3.12), we immediately see that each $\mathbf{S}^{\ell,(L)}$ is $\mathbb{F}^{\ell,(L)}$-adapted, and thus $\mathbf{S}^{(L)}$ satisfies that

$$(\mathbf{S}^{\ell,(L)})_{\ell=1,\dots,L} \text{ are } \mathbf{P}^{(L)}\text{-independent and each has the same law as } S. \tag{3.14}$$

Finally, we recall from (2.2) that the filtration $\widetilde{\mathbb{G}}^{(L)}$ is given by

$$\widetilde{\mathcal{G}}_t^{(L)} := \sigma\bigg(\bigcup_{\ell=1}^{L} \mathcal{F}_t^{\ell,(L)}\bigg), \quad t \in [0, T], \tag{3.15}$$

and $\mathbb{G}^{(L)}$ is the augmented filtration of $\widetilde{\mathbb{G}}^{(L)}$ defined in (2.4) and given explicitly by

$$\mathcal{G}_t^{(L)} = (\widetilde{\mathcal{G}}_t^{(L)} \vee \mathcal{N}^{(L)})^+, \quad t \in [0,T]. \tag{3.16}$$

In the rest of this subsection, we drop the superscript $^{(L)}$ as usual.

We turn to give a suitable setup for studying the auxiliary problem. Let us first specify a space of strategies in the extended market. In analogy to Θ_{S} given in (3.1), we define

$$\begin{aligned}\mathbf{\Theta}_{\mathrm{S}} := \{\boldsymbol{\vartheta} = (\boldsymbol{\vartheta}^\ell)_{\ell=1,\dots,L} : \boldsymbol{\vartheta}^\ell &\text{ is } \mathbb{R}^d\text{-valued, } \mathbb{G}\text{-predictable, } S\text{-integrable and} \\ &\mathbf{G}^\ell(\boldsymbol{\vartheta}) \in \mathcal{S}^2 \text{ for } \ell = 1,\dots,L\}.\end{aligned} \tag{3.17}$$

Because each $\mathbf{S}^\ell$ is an independent copy of S under the measure $\mathbf{P}$ from (3.14), we expect that Assumption 3.7 implies that $\mathbf{S}^\ell$ satisfies (SC) with respect to the filtration $\mathbb{G}$ for $\ell = 1,\dots,L$. However, this is not immediate because $\mathbb{G}$ is larger than $\mathbb{F}^\ell$. To argue in detail, we need to collect some results on how semimartingale and martingale properties change from $\mathbb{F}^\ell$ to $\mathbb{G}$ and vice versa. We cite the following result from Aksamit and Jeanblanc [3, Proposition 1.12].

Lemma 3.9. *If $\mathbb{H}$ and $\mathbb{I}$ are two right-continuous filtrations such that $\mathcal{H}_T$ and $\mathcal{I}_T$ are independent, then the filtration $(\sigma(\mathcal{H}_t,\mathcal{I}_t))_{t\in[0,T]}$ is also right-continuous.*

Recall from (3.16) that $\mathbb{G}$ is the augmented filtration of the filtration $\widetilde{\mathbb{G}}$ and is given by $\mathcal{G}_t = (\widetilde{\mathcal{G}}_t \vee \mathcal{N})^+$ for $t \in [0,T]$. A standard result from measure theory (see e.g. Kallenberg [40, Lemma 9.8]) further gives $\mathcal{G}_t = \widetilde{\mathcal{G}}_t^+ \vee \mathcal{N}$. Because $(\mathbb{F}^\ell)_{\ell=1,\dots,L}$ are right-continuous like $\mathbb{F}$ and $(\mathcal{F}_T^\ell)_{\ell=1,\dots,L}$ are independent from (3.13), we apply Lemma 3.9 to obtain

$$\mathcal{G}_t = \widetilde{\mathcal{G}}_t \vee \mathcal{N}, \quad t \in [0,T]. \tag{3.18}$$

Lemma 3.10. *The sigma-algebras $\mathcal{F}_t^1, \mathcal{F}_t^2, \dots, \mathcal{F}_t^L, \sigma(\mathcal{N})$ are $\mathbf{P}$-independent for $t \in [0,T]$.*

Proof. Let $F^\ell \in \mathcal{F}_t^\ell$ for $\ell = 1,\dots,L$ and $N \in \mathcal{N}$. Consider the class $\mathcal{D}$ of sets D such that we have the identity

$$\mathbf{P}[F^1 \cap F^2 \cap \dots \cap F^L \cap D] = \mathbf{P}[F^1]\mathbf{P}[F^2]\cdots\mathbf{P}[F^L]\mathbf{P}[D].$$

We claim $\sigma(\mathcal{N}) \subseteq \mathcal{D}$. Note that this already implies that the sigma-algebras

$\mathcal{F}_t^1, \mathcal{F}_t^2, \dots, \mathcal{F}_t^L, \sigma(\mathcal{N})$ are independent. Because $(\cap_{\ell=1}^L F^\ell) \cap N \subseteq N$, we have

$$\mathbf{P}\left[\left(\bigcap_{\ell=1}^{L} F^\ell\right) \cap N\right] = 0 = \mathbf{P}[F^1]\mathbf{P}[F^2]\cdots\mathbf{P}[F^L]\mathbf{P}[N]$$

and hence $\mathcal{N} \subseteq \mathcal{D}$. Obviously $\mathcal{N}$ is a π-system and $\mathcal{D}$ is a Dynkin system. From Dynkin's lemma, we conclude that $\sigma(\mathcal{N}) \subseteq \mathcal{D}$. This completes the proof. □

Recall from (I.2.42) the identity

$$E[X|\sigma(\mathcal{A}, \mathcal{B})] = E[X|\mathcal{A}] \tag{3.19}$$

whenever X is independent of $\mathcal{B}$. We first summarise and translate the results in Lemmas I.2.13–I.2.15 using the current notation.

Lemma 3.11. 1) *Any $\mathbb{F}^\ell$-martingale is a $\widetilde{\mathbb{G}}$-martingale.*

2) *If X^ℓ and X^m are $\mathbb{F}^\ell$ and $\mathbb{F}^m$-martingales, respectively, then the product process $X^\ell X^m$ is a $\widetilde{\mathbb{G}}$-martingale.*

Next, we show that the above results can be extended to local martingales.

Lemma 3.12. *Any local $\widetilde{\mathbb{G}}$-martingale is a local $\mathbb{G}$-martingale, and any local $\mathbb{F}^\ell$-martingale is a local $\mathbb{G}$-martingale for $\ell = 1, \dots, L$.*

Proof. Let X be a local $\widetilde{\mathbb{G}}$-martingale. The path property of X is not affected by a change of filtration. We take a $\widetilde{\mathbb{G}}$-localising sequence $(\tau_n)_{n\in\mathbb{N}}$ for X, meaning that $(\tau_n)_{n\in\mathbb{N}}$ is a sequence of $\widetilde{\mathbb{G}}$-stopping times such that X^{τ_n} is a $\widetilde{\mathbb{G}}$-martingale for $n \in \mathbb{N}$ (i.e., both with respect to the filtration $\widetilde{\mathbb{G}}$). Because $\widetilde{\mathcal{G}}_t \subseteq \mathcal{G}_t$ for $t \in \mathbb{N}$, we get that each τ_n is still a stopping time in the larger filtration $\mathbb{G}$, and the stopped process X^{τ_n} is still adapted to $\mathbb{G}$. Now we check the martingale property for X^{τ_n} with respect to the filtration $\mathbb{G}$. Let $s \in [0,t]$. Because each X^{τ_n} is adapted to $\widetilde{\mathbb{G}}$, we get by Lemma 3.10 that $X_t^{\tau_n}$ is independent of $\mathcal{N}$ for all $t \in [0,T]$. So we can use (3.18) and (3.19) with $(\mathcal{A}, \mathcal{B}) = (\widetilde{\mathcal{G}}_s, \mathcal{N})$ and finally the martingale property of X^{τ_n} with respect to $\widetilde{\mathbb{G}}$ to obtain

$$\mathbf{E}[X_t^{\tau_n}|\mathcal{G}_s] = \mathbf{E}[X_t^{\tau_n}|\sigma(\widetilde{\mathcal{G}}_s, \mathcal{N})] = \mathbf{E}[X_t^{\tau_n}|\widetilde{\mathcal{G}}_s] = X_s^{\tau_n}.$$

Combining the same localisation argument with Lemma 3.11, 1) yields that any local $\mathbb{F}^\ell$-martingale is a local $\widetilde{\mathbb{G}}$-martingale. So the second statement follows from the first. □

Lemma 3.13. *Suppose that Assumption* 3.1 *is satisfied. Then the following statements hold:*

1) *For each* $\ell = 1, \dots, L$*, the special semimartingale decomposition of* $\mathbf{S}^\ell$ *in the filtration* $\mathbb{G}$ *is the same as in the filtration* $\mathbb{F}^\ell$ *given in* (3.11).

2) *For* $\ell, m = 1, \dots, L$ *with* $\ell \neq m$ *and* $i, j = 1, \dots, d$*, the processes* $\mathbf{M}^{\ell,i}$ *and* $\mathbf{M}^{m,j}$ *are strongly orthogonal, meaning that* $\langle \mathbf{M}^{\ell,i}, \mathbf{M}^{m,j} \rangle = 0$*, in the filtration* $\mathbb{G}$.

Proof. 1) In view of Lemma 3.12, the proof of part 1) is the same as in discrete time. We start with the canonical decomposition of $\mathbf{S}^\ell = \mathbf{S}_0 + \mathbf{M}^{\mathbb{F}^\ell} + \mathbf{A}^{\mathbb{F}^\ell}$ in the filtration $\mathbb{F}^\ell$. Because $\mathbf{M}^{\mathbb{F}^\ell}$ is still a local martingale in the filtration $\mathbb{G}$ and $\mathbf{A}^{\mathbb{F}^\ell}$ is still predictable in $\mathbb{G}$, we get by the uniqueness of the canonical decomposition that $\mathbf{S}^\ell = \mathbf{S}_0 + \mathbf{M}^{\mathbb{F}^\ell} + \mathbf{A}^{\mathbb{F}^\ell}$ is the canonical decomposition of $\mathbf{S}^\ell$ in the filtration $\mathbb{G}$.

2) We claim that the product $\mathbf{M}^{\ell,i}\mathbf{M}^{m,j}$ is a local martingale in $\mathbb{G}$. This then yields the assertion by the definition of the angle bracket. We first use part 1) and Assumption 3.1 to view $\mathbf{M}^{\ell,i}$ and $\mathbf{M}^{m,j}$ as elements in $\mathcal{M}^2_{0,\mathrm{loc}}(\mathbb{F}^\ell)$ and $\mathcal{M}^2_{0,\mathrm{loc}}(\mathbb{F}^m)$, respectively. Thus we may choose localising sequences $(\tau_n)_{n\in\mathbb{N}}$ and $(\sigma_n)_{n\in\mathbb{N}}$ (in the filtrations $\mathbb{F}^\ell$ and $\mathbb{F}^m$, respectively) such that for $n \in \mathbb{N}$, the stopped processes $(\mathbf{M}^{\ell,i})^{\tau_n}$ and $(\mathbf{M}^{m,j})^{\sigma_n}$ are in $\mathcal{M}^2_0(\mathbb{F}^\ell)$ and $\mathcal{M}^2_0(\mathbb{F}^m)$, respectively, and hence are still independent. So we can apply Lemma 3.11 to obtain that the product $(\mathbf{M}^{\ell,i})^{\tau_n}(\mathbf{M}^{m,j})^{\sigma_n}$ is a martingale in the filtration $\widetilde{\mathbb{G}}$, and then so is the process $(\mathbf{M}^{\ell,i}\mathbf{M}^{m,j})^{\tau_n\wedge\sigma_n} = ((\mathbf{M}^{\ell,i})^{\tau_n}(\mathbf{M}^{m,j})^{\sigma_n})^{\tau_n\wedge\sigma_n}$. Here we use that $(\tau_n \wedge \sigma_n)_{n\in\mathbb{N}}$ is a sequence of stopping times in the filtration $\widetilde{\mathbb{G}}$. Indeed, using the definitions of τ_n and σ_n and (3.15), we get

$$\{\tau_n \wedge \sigma_n \le t\} = \{\tau_n \le t\} \cup \{\sigma_n \le t\} \in \sigma(\mathcal{F}^\ell_t, \mathcal{F}^m_t) \subseteq \widetilde{\mathcal{G}}_t, \quad t \in [0,T].$$

This shows that $\tau_n \wedge \sigma_n$ is indeed a stopping time with respect to $\widetilde{\mathbb{G}}$ and hence the process $\mathbf{M}^{\ell,i}\mathbf{M}^{m,j}$ is a local martingale in the filtration $\widetilde{\mathbb{G}}$. Finally by Lemma 3.12, the process $\mathbf{M}^{\ell,i}\mathbf{M}^{m,j}$ is a local martingale in the filtration $\mathbb{G}$, as desired. □

Corollary 3.14. *Suppose that Assumption* 3.7 *is satisfied. Then* $\mathbf{S}^\ell$ *satisfies the structure condition (SC) with respect to the filtration* $\mathbb{G}$ *for* $\ell = 1, \dots, L$*, meaning that:*

1) $\mathbf{S}^\ell$ *is an* $\mathbb{R}^d$*-valued* $\mathbb{G}$*-semimartingale in* $\mathcal{S}^2_{\mathrm{loc}}$ *for* $\ell = 1, \dots, L$.

2) *For* $i = 1, \dots, d$ *and* $\ell = 1, \dots, L$*, we have* $\mathbf{A}^{\ell,i} \ll \langle \mathbf{M}^{\ell,i} \rangle$ *with a* $\mathbb{G}$*-predictable density* $\alpha^{\ell,i}$.

3) *For* $\mathbf{B}^\ell = \sum_{i=1}^d \langle \mathbf{M}^{\ell,i} \rangle$ *and*

$$\boldsymbol{\sigma}^{\ell,ij} = \frac{\mathrm{d}\langle \mathbf{M}^{\ell,i}, \mathbf{M}^{\ell,j} \rangle}{\mathrm{d}\mathbf{B}^\ell}, \qquad i,j = 1,\dots,d,\ \ \ell = 1,\dots,L, \tag{3.20}$$

$$\boldsymbol{\gamma}^{\ell,i} = \boldsymbol{\alpha}^{\ell,i}\boldsymbol{\sigma}^{\ell,ii} = \frac{\mathrm{d}\mathbf{A}^{\ell,i}}{\mathrm{d}\mathbf{B}^\ell}, \qquad i = 1,\dots,d,\ \ \ell = 1,\dots,L, \tag{3.21}$$

there exist $\mathbb{G}$*-predictable* $\mathbb{R}^d$*-valued processes* $\boldsymbol{\lambda}^\ell$ *such that*

$$\boldsymbol{\sigma}^\ell_t \boldsymbol{\lambda}^\ell_t = \boldsymbol{\gamma}^\ell_t \qquad \mathbf{P}\text{-a.s. for all } t \in [0,T], \tag{3.22}$$

$$\mathbf{K}^\ell_T := \int_0^t (\boldsymbol{\lambda}^\ell_s)^\top \boldsymbol{\gamma}^\ell_s \,\mathrm{d}\mathbf{B}^\ell_s = \int_0^t (\boldsymbol{\lambda}^\ell_s)^\top \boldsymbol{\sigma}^\ell_s \boldsymbol{\lambda}^\ell_s \,\mathrm{d}\mathbf{B}^\ell_s < \infty \qquad \mathbf{P}\text{-a.s.} \tag{3.23}$$

Moreover, the key quantities in the structure condition can be constructed explicitly by

$$\mathbf{X}^\ell(\omega^{(L)}) = X(\omega_\ell), \quad \ell = 1,\dots,L, \tag{3.24}$$

for $\mathbf{X} \in \{\mathbf{M}, \mathbf{A}, \langle \mathbf{M} \rangle, \boldsymbol{\alpha}, \mathbf{B}, \boldsymbol{\sigma}, \boldsymbol{\gamma}, \boldsymbol{\lambda}, \mathbf{K}\}$ *and* $X \in \{M, A, \langle M \rangle, \alpha, B, \sigma, \gamma, \lambda, K\}$.

Proof. Due to $\mathbf{S}^\ell(\omega^{(L)}) = S(\omega_\ell)$ from (3.12) and the relation (3.11) between $\mathbb{F}$ and $\mathbb{F}^\ell$, we immediately see that $\mathbf{S}^\ell$ satisfies the structure condition (SC) in the filtration $\mathbb{F}^\ell$ for $\ell = 1,\dots,L$. For $\mathbf{X} \in \{\mathbf{M}^{\mathbb{F}^\ell}, \mathbf{A}^{\mathbb{F}^\ell}, \langle \mathbf{M}^{\mathbb{F}^\ell} \rangle, \boldsymbol{\alpha}^{\mathbb{F}^\ell}, \mathbf{B}^{\mathbb{F}^\ell}, \boldsymbol{\sigma}^{\mathbb{F}^\ell}, \boldsymbol{\gamma}^{\mathbb{F}^\ell}, \boldsymbol{\lambda}^{\mathbb{F}^\ell}, \mathbf{K}^{\mathbb{F}^\ell}\}$, which denote the key quantities describing (SC) for $\mathbf{S}^\ell$ with respect to the filtration $\mathbb{F}^\ell$, and for $X \in \{M, A, \langle M \rangle, \alpha, B, \sigma, \gamma, \lambda, K\}$, we have

$$\mathbf{X}^\ell(\omega^{(L)}) = X(\omega_\ell), \quad \ell = 1,\dots,L \tag{3.25}$$

This then *yields* that $\mathbf{S}^\ell$ satisfies the structure condition (SC) in the filtration $\mathbb{G}$. Indeed, because of the fact from Lemma 3.13, 1) that the special semimartingale decomposition of $\mathbf{S}^\ell$ in the filtration $\mathbb{G}$ is the same as in the filtration $\mathbb{F}^\ell$, we obtain $(\mathbf{M}^\ell, \mathbf{A}^\ell) = (\mathbf{M}^{\mathbb{F}^\ell}, \mathbf{A}^{\mathbb{F}^\ell})$. This also establishes 1).

For 2), due to the uniqueness characterisation of the angle bracket, we obtain $\langle \mathbf{M}^\ell \rangle = \langle \mathbf{M}^{\mathbb{F}^\ell} \rangle$. For $i = 1,\dots,d$, we can choose $\boldsymbol{\alpha}^{\ell,i} = \boldsymbol{\alpha}^{\mathbb{F}^\ell,i}$ because $\boldsymbol{\alpha}^{\mathbb{F}^\ell,i}$ is also $\mathbb{G}$-predictable and satisfies

$$\mathbf{A}^{\ell,i} = \mathbf{A}^{\mathbb{F}^\ell,i} = \int \boldsymbol{\alpha}^{\mathbb{F}^\ell,i}\,\mathrm{d}\langle \mathbf{M}^{\mathbb{F}^\ell,i} \rangle = \int \boldsymbol{\alpha}^{\mathbb{F}^\ell,i}\,\mathrm{d}\langle \mathbf{M}^{\ell,i} \rangle.$$

So far we have proved that

$$(\mathbf{M}^\ell, \mathbf{A}^\ell, \langle \mathbf{M}^\ell \rangle, \boldsymbol{\alpha}^\ell) = (\mathbf{M}^{\mathbb{F}^\ell}, \mathbf{A}^{\mathbb{F}^\ell}, \langle \mathbf{M}^{\mathbb{F}^\ell} \rangle, \boldsymbol{\alpha}^{\mathbb{F}^\ell}).$$

Using this identity and the definitions of $\mathbf{B}, \boldsymbol{\sigma}, \boldsymbol{\gamma}$ in 3) yields

$$(\mathbf{B}^\ell, \boldsymbol{\sigma}^\ell, \boldsymbol{\gamma}^\ell) = (\mathbf{B}^{\mathbb{F}^\ell}, \boldsymbol{\sigma}^{\mathbb{F}^\ell}, \boldsymbol{\gamma}^{\mathbb{F}^\ell}).$$

Because $\boldsymbol{\lambda}^{\mathbb{F}^\ell}$ is also $\mathbb{G}$-predictable, we can choose $\boldsymbol{\lambda}^\ell = \boldsymbol{\lambda}^{\mathbb{F}^\ell}$ so that the requirements (3.22) and (3.23) are satisfied. In summary, we obtain that

$$\begin{aligned}&(\mathbf{M}^\ell, \mathbf{A}^\ell, \langle \mathbf{M}^\ell \rangle, \boldsymbol{\alpha}^\ell, \mathbf{B}^\ell, \boldsymbol{\sigma}^\ell, \boldsymbol{\gamma}^\ell, \boldsymbol{\lambda}^\ell, \mathbf{K}^\ell)\\ &= (\mathbf{M}^{\mathbb{F}^\ell}, \mathbf{A}^{\mathbb{F}^\ell}, \langle \mathbf{M}^{\mathbb{F}^\ell} \rangle, \boldsymbol{\alpha}^{\mathbb{F}^\ell}, \mathbf{B}^{\mathbb{F}^\ell}, \boldsymbol{\sigma}^{\mathbb{F}^\ell}, \boldsymbol{\gamma}^{\mathbb{F}^\ell}, \boldsymbol{\lambda}^{\mathbb{F}^\ell}, \mathbf{K}^{\mathbb{F}^\ell}).\end{aligned}$$

Combining the above display with (3.25) yields (3.24). □

Definition 3.4 about $L^2_{(\mathrm{loc})}(M)$ and $L^2_{(\mathrm{loc})}(A)$ corresponds to the following definition.

Definition 3.15. The space $L^2_{(\mathrm{loc})}(\mathbf{M})$ consists of all $\mathbb{R}^{d\times L}$-valued $\mathbb{G}$-predictable processes $(\boldsymbol{\vartheta}^\ell)_{\ell=1,\dots,L}$ such that the process $\int (\boldsymbol{\vartheta}^\ell)^\top \boldsymbol{\sigma}^\ell \boldsymbol{\vartheta}^\ell \,\mathrm{d}\mathbf{B}^\ell$ is (locally) integrable for $\ell = 1,\dots,L$. The space $L^2_{(\mathrm{loc})}(\mathbf{A})$ consists of all $\mathbb{R}^{d\times L}$-valued $\mathbb{G}$-predictable processes $(\boldsymbol{\vartheta}^\ell)_{\ell=1,\dots,L}$ such that the process $\int |(\boldsymbol{\vartheta}^\ell)^\top \boldsymbol{\gamma}^\ell| \,\mathrm{d}\mathbf{B}^\ell$ is (locally) square-integrable for $\ell = 1,\dots,L$.

Comparing (3.17) with (3.1), we obtain the following results corresponding to Lemmas 3.5 and 3.6, respectively.

Lemma 3.16. *Suppose that Assumption* 3.7 *is satisfied. Then the following statements hold:*

1) *If* $\boldsymbol{\vartheta} \in L^2_{(\mathrm{loc})}(\mathbf{M})$, *then the stochastic integral* $\int \boldsymbol{\vartheta}^\ell \,\mathrm{d}\mathbf{M}^\ell$ *is well defined, in* $\mathcal{M}^2_{0(,\mathrm{loc})}$ *for* $\ell = 1,\dots,L$, *and satisfies for* $\boldsymbol{\psi} \in L^2_{(\mathrm{loc})}(\mathbf{M})$ *that*

$$\left\langle \int \boldsymbol{\vartheta}^\ell \,\mathrm{d}\mathbf{M}^\ell, \int \boldsymbol{\psi}^\ell \,\mathrm{d}\mathbf{M}^\ell \right\rangle_t = \int_0^t (\boldsymbol{\vartheta}^\ell_s)^\top \boldsymbol{\sigma}^\ell_s \boldsymbol{\psi}^\ell_s \,\mathrm{d}\mathbf{B}^\ell_s, \quad t \in [0,T], \ell = 1,\dots,L. \tag{3.26}$$

In particular, if $\boldsymbol{\vartheta} \in L^2(\mathbf{M})$, *then*

$$\mathbf{E}\left[\sup_{t\in[0,T]} \left| \int_0^t \boldsymbol{\vartheta}^\ell_s \,\mathrm{d}\mathbf{M}^\ell_s \right|^2\right] \le 4\mathbf{E}\left[\int_0^t (\boldsymbol{\vartheta}^\ell_s)^\top \boldsymbol{\sigma}^\ell_s \boldsymbol{\vartheta}^\ell_s \,\mathrm{d}\mathbf{B}^\ell_s\right], \quad \ell = 1,\dots,L. \tag{3.27}$$

2) *If* $\boldsymbol{\vartheta} \in L^2_{(\mathrm{loc})}(\mathbf{A})$, *then the process* $\int (\boldsymbol{\vartheta}^\ell)^\top \,\mathrm{d}\mathbf{A}^\ell := \sum_{i=1}^d \int_0^t \boldsymbol{\vartheta}^{\ell,i}_s \,\mathrm{d}\mathbf{A}^{\ell,i}_s$ *is well defined as a Lebesgue–Stieltjes integral for* $\ell = 1,\dots,L$, *(locally) square-*

integrable, predictable and satisfies

$$\int_0^t (\boldsymbol{\vartheta}_s^\ell)^\top \, \mathrm{d}\mathbf{A}_s^\ell = \int_0^t (\boldsymbol{\vartheta}_s^\ell)^\top \boldsymbol{\gamma}_s^\ell \, \mathrm{d}\mathbf{B}_s^\ell, \qquad t \in [0,T], \ell = 1, \dots, L. \tag{3.28}$$

Lemma 3.17. *Suppose that Assumption* 3.7 *is satisfied. Then*

$$\boldsymbol{\Theta}_{\mathrm{S}} = L^2(\mathbf{M}) \cap L^2(\mathbf{A}), \tag{3.29}$$

and $\mathbf{E}[\sup_{s\in[0,t]} |\mathbf{G}_s^\ell(\boldsymbol{\vartheta})|^2] < \infty$ *for* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ *and* $\ell = 1, \dots, L$. *If in addition the MVT process* K *is bounded, then*

$$\boldsymbol{\Theta}_{\mathrm{S}} = L^2(\mathbf{M}).$$

Proof. The property $\mathbf{E}[\sup_{s\in[0,t]} |\mathbf{G}_s^\ell(\boldsymbol{\vartheta})|^2] < \infty$ follows from the same argument as for $E[\sup_{s\in[0,t]} |G_t(\theta)|^2] < \infty$ for $\theta \in \Theta_{\mathrm{S}}$ in the last part of the proof of Lemma 3.8. □

4 Construction of a value process in dimension 1

In this section, we construct a candidate $(\widetilde{V}(\boldsymbol{\vartheta}))_{\boldsymbol{\vartheta}\in\boldsymbol{\Theta}_{\mathrm{S}}}$ for the value process family for the auxiliary problem (2.9). The main idea is to use Lemma 2.4 as follows.

Recipe 4.1. 1) From the discrete-time result (I.4.72), we first make the educated guess that $\widetilde{V}$ has the same affine–quadratic form in continuous time. We then apply Itô's lemma to find the canonical decomposition of $\widetilde{V}(\boldsymbol{\vartheta})$ for $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$.

2) If we want to apply Lemma 2.4, then each $\widetilde{V}(\boldsymbol{\vartheta})$ should satisfy certain supermartingale or martingale properties. Using heuristic arguments, we derive from these desired supermartingale/martingale properties some differential equations for the coefficients in the affine–quadratic expression of $\widetilde{V}$.

3) Finally, we solve those differential equations explicitly, possibly with additional assumptions, and we verify that the resulting candidate family $(\widetilde{V}(\boldsymbol{\vartheta}))_{\boldsymbol{\vartheta}\in\boldsymbol{\Theta}_{\mathrm{S}}}$ indeed satisfies the supermartingale/martingale conditions in Lemma 2.4.

Before we embark on the above programme, let us **restrict ourselves to** $d = 1$ so that the notations are simpler. To avoid the technical difficulties brought by processes having jumps, we only consider the case where the process S is continuous. For convenience, let us state this additional assumption together with Assumption 3.7. We can omit Assumption 3.1 because S is locally bounded by continuity and hence in $\mathcal{S}^2_{\mathrm{loc}}$.

Assumption 4.2. The process S is a real-valued continuous semimartingale satisfying the structure condition (SC).

4.1 Implementing Recipe 4.1, 1) – canonical decomposition of the value process

In this subsection, we implement Recipe 4.1, 1). We first recall several notations in the extended market from (2.5)–(2.8), namely

$$\mathbf{G}_t^\ell(\boldsymbol{\vartheta}) = \int_0^t \boldsymbol{\vartheta}_s^\ell \,\mathrm{d}\mathbf{S}_s^\ell, \quad \ell = 1, \dots, L, \tag{4.1}$$

$$\mathbf{x} \odot \mathbf{y} = \sum_{\ell=1}^L \mathbf{x}^\ell \mathbf{y}^\ell, \quad \mathbf{x}^2 = \mathbf{x} \odot \mathbf{x},$$

$$\mathrm{em}(\mathbf{x}) = \frac{1}{L} \sum_{\ell=1}^L \mathbf{x}^\ell, \tag{4.2}$$

$$\mathrm{evar}(\mathbf{x}) = \mathrm{em}(\mathbf{x}^2) - \big(\mathrm{em}(\mathbf{x})\big)^2. \tag{4.3}$$

Inspired by the discrete-time formula (I.4.72), we guess and here *assume* that $\widetilde{V}$ takes the form

$$\widetilde{V}_t(\boldsymbol{\vartheta}) = a_t \mathrm{em}\big(\mathbf{G}_t(\boldsymbol{\vartheta})\big) - b_t \mathrm{evar}\big(\mathbf{G}_t(\boldsymbol{\vartheta})\big) + c_t, \quad t \in [0, T],$$
$$\text{where the processes } a = (a_t)_{t\in[0,T]}, b = (b_t)_{t\in[0,T]}, c = (c_t)_{t\in[0,T]}$$
$$\text{are adapted, continuous and of finite variation.} \tag{4.4}$$

As in Lemma 2.4, we also *assume* that $\widetilde{V}_T(\boldsymbol{\vartheta}) = J_T(\boldsymbol{\vartheta})$ for all $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$. So from (2.20) and (2.17), we have

$$\widetilde{V}_T(\boldsymbol{\vartheta}) = J_T(\boldsymbol{\vartheta}) = \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta})\big), \quad \boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}.$$

Comparing the last expression with (4.4), we obtain for the processes a, b, c the terminal conditions

$$a_T = 1, \quad b_T = \xi, \quad c_T = 0. \tag{4.5}$$

With $d = 1$, the processes $\mathbf{S}^\ell$ and $\mathbf{M}^\ell$, $\mathbf{A}^\ell$ from the canonical decomposition of $\mathbf{S}^\ell$ in the filtration $\mathbb{G}$,

$$\mathbf{S}^\ell = \mathbf{S}_0 + \mathbf{M}^\ell + \mathbf{A}^\ell, \tag{4.6}$$

are all 1-dimensional for $\ell = 1, \dots, L$. The quantities describing (SC) are simplified as follows. Setting $\mathbf{B}^\ell = \langle \mathbf{M}^\ell \rangle$, we get from the expressions (3.20) for $\boldsymbol{\sigma}^\ell$ and (3.21) for $\boldsymbol{\gamma}^\ell$ that for $\ell = 1, \dots, L$,

$$\boldsymbol{\sigma}^\ell_t = \frac{\mathrm{d}\langle \mathbf{M}^\ell \rangle_t}{\mathrm{d}\langle \mathbf{M}^\ell \rangle_t} = 1, \quad \boldsymbol{\gamma}^\ell_t = \boldsymbol{\alpha}^\ell_t \boldsymbol{\sigma}^\ell_t = \boldsymbol{\alpha}^\ell_t \qquad \mathbf{P}\text{-a.s. for all } t \in [0,T]. \tag{4.7}$$

Hence (3.22) for $\boldsymbol{\lambda}^\ell$ yields $\boldsymbol{\lambda}^\ell = \boldsymbol{\alpha}^\ell$ for $\ell = 1, \dots, L$. We obtain from this and (3.23) for $\mathbf{K}^\ell$ that

$$\mathbf{A}^\ell_t = \int_0^t \boldsymbol{\lambda}^\ell_s \,\mathrm{d}\langle \mathbf{M}^\ell \rangle_s \qquad \mathbf{P}\text{-a.s. for all } t \in [0,T],\ \ell = 1, \dots, L, \tag{4.8}$$

$$\mathbf{K}^\ell_t = \int_0^t (\boldsymbol{\lambda}^\ell_s)^2 \,\mathrm{d}\langle \mathbf{M}^\ell \rangle_s \qquad \mathbf{P}\text{-a.s. for all } t \in [0,T],\ \ell = 1, \dots, L. \tag{4.9}$$

Lemma 4.3. *Suppose that Assumption* 4.2 *is satisfied and assume* (4.4)*. Then for $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$, the canonical decomposition of $\widetilde{V}(\boldsymbol{\vartheta})$ in the filtration $\mathbb{G}$ is*

$$\begin{aligned}
&\widetilde{V}_t(\boldsymbol{\vartheta}) \\
&= \widetilde{V}_0 + \widetilde{M}_t(\boldsymbol{\vartheta}) + \frac{1}{L}\int_0^t a_s \sum_{\ell=1}^L \boldsymbol{\vartheta}^\ell_s \boldsymbol{\lambda}^\ell_s \,\mathrm{d}\langle \mathbf{M}^\ell \rangle_s \\
&\quad - \frac{1}{L}\sum_{\ell=1}^L \int_0^t b_s \Big(2\Big(\mathbf{G}^\ell_s(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big)\boldsymbol{\vartheta}^\ell_s \boldsymbol{\lambda}^\ell_s + (1 - L^{-1})(\boldsymbol{\vartheta}^\ell_s)^2 \Big) \,\mathrm{d}\langle \mathbf{M}^\ell \rangle_s \\
&\quad + \int_0^t \mathrm{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\,\mathrm{d}a_s - \int_0^t \mathrm{evar}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\,\mathrm{d}b_s + c_t, \qquad t \in [0,T],
\end{aligned} \tag{4.10}$$

where $(\widetilde{M}_t(\boldsymbol{\vartheta}))_{t\in[0,T]}$ is a continuous local martingale in $\mathbb{G}$.

Proof. The proof is mainly to apply Itô's lemma to (4.4) in the filtration $\mathbb{G}$. Let $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ and note that (4.4) implies that $\widetilde{V}(\boldsymbol{\vartheta})$ is continuous. Using (4.1) and (4.3), we get

$$\mathrm{d}\Big(\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big) = \frac{1}{L}\sum_{\ell=1}^L \boldsymbol{\vartheta}^\ell \,\mathrm{d}\mathbf{S}^\ell. \tag{4.11}$$

We then use the product formula with (4.1) and (4.6) with the continuity of $\mathbf{S}^\ell$

for $\ell = 1, \dots, L$ from Assumption 4.2 to obtain

$$\begin{aligned} \mathrm{d}\Big(\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})^2\big)\Big) &= \frac{1}{L}\sum_{\ell=1}^{L}\big(2\mathbf{G}^\ell_-(\boldsymbol{\vartheta})\,\mathrm{d}\mathbf{G}^\ell(\boldsymbol{\vartheta}) + \mathrm{d}[\mathbf{G}^\ell(\boldsymbol{\vartheta})]\big) \\ &= \frac{1}{L}\sum_{\ell=1}^{L}\big(2\mathbf{G}^\ell(\boldsymbol{\vartheta})\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{S}^\ell + (\boldsymbol{\vartheta}^\ell)^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle\big). \end{aligned} \tag{4.12}$$

Applying the product formula again with the continuity of $\mathbf{S}^\ell$ and then using (4.11), we also obtain

$$\begin{aligned} \mathrm{d}\Big(\Big(\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big)^2\Big) &= 2\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\,\mathrm{d}\Big(\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big) + \mathrm{d}\big[\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\big] \\ &= \frac{1}{L}\sum_{\ell=1}^{L} 2\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{S}^\ell + \frac{1}{L^2}\sum_{\ell,m=1}^{L}\boldsymbol{\vartheta}^\ell\boldsymbol{\vartheta}^m\,\mathrm{d}[\mathbf{S}^\ell,\mathbf{S}^m] \\ &= \frac{1}{L}\sum_{\ell=1}^{L} 2\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{S}^\ell + \frac{1}{L^2}\sum_{\ell=1}^{L}(\boldsymbol{\vartheta}^\ell)^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle. \end{aligned} \tag{4.13}$$

The last line of (4.13) uses (4.6) and $\langle\mathbf{M}^\ell,\mathbf{M}^m\rangle = 0$ for $\ell \neq m$ by Lemma 3.13, 2). Now combining (4.12) and (4.13) with the definition (4.3) of $\mathrm{evar}(\mathbf{x})$ yields

$$\begin{aligned} \mathrm{d}\Big(\mathrm{evar}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big) &= \mathrm{d}\Big(\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})^2\big)\Big) - \mathrm{d}\Big(\Big(\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big)^2\Big) \\ &= \frac{1}{L}\sum_{\ell=1}^{L}\bigg(2\Big(\mathbf{G}^\ell(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big)\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{S}^\ell \\ &\qquad\qquad + (1 - L^{-1})(\boldsymbol{\vartheta}^\ell)^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle\bigg). \end{aligned} \tag{4.14}$$

Then we apply the product formula to (4.4) and use the continuity of $\mathbf{S}^\ell$ and that the processes a, b and c are of finite variation to obtain

$$\big\langle a, \mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\big\rangle = \big\langle b, \mathrm{evar}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\big\rangle = 0.$$

Using the above identity and $\mathrm{d}(\mathrm{em}(\mathbf{G}(\boldsymbol{\vartheta})))$, $\mathrm{d}(\mathrm{evar}(\mathbf{G}(\boldsymbol{\vartheta})))$ from (4.11) and (4.14), respectively and finally $\mathbf{A}^\ell = \int \boldsymbol{\lambda}^\ell \,\mathrm{d}\langle \mathbf{M}^\ell\rangle$ from (4.8) yields

$$
\begin{aligned}
\mathrm{d}\widetilde{V}(\boldsymbol{\vartheta}) &= a\,\mathrm{d}\Big(\mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big) + \mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\,\mathrm{d}a - b\,\mathrm{d}\Big(\mathrm{evar}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big) \\
&\quad - \mathrm{evar}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\,\mathrm{d}b + \mathrm{d}c \\
&= \frac{1}{L}\sum_{\ell=1}^{L} a\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{S}^\ell + \mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\,\mathrm{d}a \\
&\quad - \frac{1}{L}\sum_{\ell=1}^{L}\bigg(2b\Big(\mathbf{G}^\ell(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big)\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{S}^\ell + b(1-L^{-1})(\boldsymbol{\vartheta}^\ell)^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle\bigg) \\
&\quad - \mathrm{evar}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\,\mathrm{d}b + \mathrm{d}c \\
&= \frac{1}{L}\sum_{\ell=1}^{L} a\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{M}^\ell - \frac{1}{L}\sum_{\ell=1}^{L} 2b\Big(\mathbf{G}^\ell(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big)\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{M}^\ell \\
&\quad + \frac{1}{L}\sum_{\ell=1}^{L} a\boldsymbol{\vartheta}^\ell\boldsymbol{\lambda}^\ell\,\mathrm{d}\langle\mathbf{M}^\ell\rangle + \mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\,\mathrm{d}a \\
&\quad - \frac{1}{L}\sum_{\ell=1}^{L} b\bigg(2\Big(\mathbf{G}^\ell(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\Big)\boldsymbol{\vartheta}^\ell\boldsymbol{\lambda}^\ell + (1-L^{-1})(\boldsymbol{\vartheta}^\ell)^2\bigg)\,\mathrm{d}\langle\mathbf{M}^\ell\rangle \\
&\quad - \mathrm{evar}\big(\mathbf{G}(\boldsymbol{\vartheta})\big)\,\mathrm{d}b + \mathrm{d}c.
\end{aligned}
\tag{4.15}
$$

Writing (4.15) in integral form, we obtain (4.10). It remains to show that

$$
\begin{aligned}
\widetilde{M}_t(\boldsymbol{\vartheta}) &= \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t a_s\boldsymbol{\vartheta}_s^\ell\,\mathrm{d}\mathbf{M}_s^\ell \\
&\quad - \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t 2b_s\Big(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big)\boldsymbol{\vartheta}_s^\ell\,\mathrm{d}\mathbf{M}_s^\ell, \quad t\in[0,T],
\end{aligned}
$$

is a local martingale. Because $(a_t)_{t\in[0,T]}$ is continuous and $\boldsymbol{\vartheta}\in L^2(\mathbf{M})$ by (3.29), we get $a\odot\boldsymbol{\vartheta} = (a\boldsymbol{\vartheta}^\ell)_{\ell=1,\ldots,L}\in L^2_{\mathrm{loc}}(\mathbf{M})$, and hence $\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t a_s\boldsymbol{\vartheta}_s^\ell\,\mathrm{d}\mathbf{M}_s^\ell$ is a local martingale by Lemma 3.16, 1). By the same reasoning, we only need to show that the process $\int(\mathbf{G}^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}(\boldsymbol{\vartheta})))\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{M}^\ell$ is a local martingale for $\ell = 1,\ldots,L$. In view of (4.2), it suffices to show that $\int\mathbf{G}^m(\boldsymbol{\vartheta})\boldsymbol{\vartheta}^\ell\,\mathrm{d}\mathbf{M}^\ell$ is a local martingale for $\ell, m = 1,\ldots,L$. This simply follows from $\boldsymbol{\vartheta}\in L^2(\mathbf{M})$ and Lemma 3.16, 1) and because $\mathbf{G}^m(\boldsymbol{\vartheta})$ is continuous by Assumption 4.2 and hence locally bounded. □

4.2 Implementing Recipe 4.1, 2) – a heuristic derivation of equations

In this subsection, we implement Recipe 4.1, 2) – to derive from the desired (super)martingale conditions in Lemma 2.4 equations for the processes a, b, and c in the assumed representation (4.4) of $\widetilde{V}$.

As a first simplification, we *assume*

$$a_t = 1, \qquad t \in [0,T].$$

This is also suggested by the discrete-time formula (I.4.72). Then (4.10) becomes

$$\begin{aligned}\widetilde{V}_t(\boldsymbol{\vartheta}) &= \widetilde{V}_0 + \widetilde{M}_t(\boldsymbol{\vartheta}) \\ &\quad - \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \Bigg(\bigg(2b_s\Big(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big) - 1\bigg)\boldsymbol{\vartheta}_s^\ell\boldsymbol{\lambda}_s^\ell \\ &\qquad\qquad + b_s(1-L^{-1})(\boldsymbol{\vartheta}_s^\ell)^2\Bigg)\,\mathrm{d}\langle\mathbf{M}^\ell\rangle_s \\ &\quad - \int_0^t \mathrm{evar}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\,\mathrm{d}b_s + c_t \\ &=: \widetilde{V}_0 + \widetilde{M}_t(\boldsymbol{\vartheta}) + \widetilde{A}_t(\boldsymbol{\vartheta})\end{aligned} \tag{4.16}$$

where $(\widetilde{M}_t(\boldsymbol{\vartheta}))_{t\in[0,T]}$ is a local martingale. The (super)martingale conditions on $\widetilde{V}(\boldsymbol{\vartheta})$ then translate into the statement that

$$\widetilde{A}(\boldsymbol{\vartheta}) \text{ is decreasing for all } \boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}, \text{ and } \widetilde{A}(\boldsymbol{\vartheta}^*) \equiv 0 \text{ for some } \boldsymbol{\vartheta}^* \in \boldsymbol{\Theta}_{\mathrm{S}}. \tag{4.17}$$

We apply the identity $\alpha x^2 + \beta x = \alpha(x + \frac{\beta}{2\alpha})^2 - \frac{\beta^2}{4\alpha}$ to the integrand in the second and third line of (4.16) to complete the square and reorder the terms to obtain, with $\alpha = b_s(1-L^{-1})$, $\beta = (2b_s\Big(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big) - 1)\boldsymbol{\lambda}_s^\ell$, $x = \boldsymbol{\vartheta}_s^\ell$,

$$\begin{aligned}&\bigg(2b_s\Big(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big) - 1\bigg)\boldsymbol{\vartheta}_s^\ell\boldsymbol{\lambda}_s^\ell + b_s(1-L^{-1})(\boldsymbol{\vartheta}_s^\ell)^2 \\ &= b_s(1-L^{-1})\bigg(\boldsymbol{\vartheta}_s^\ell + \frac{2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s(1-L^{-1})}\boldsymbol{\lambda}_s^\ell\bigg)^2 \\ &\quad - \frac{(2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1)^2}{4b_s(1-L^{-1})}(\boldsymbol{\lambda}_s^\ell)^2, \quad s \in [0,t].\end{aligned}$$

Inserting the last equality into (4.16) and using $\mathbf{K}^\ell = \int(\boldsymbol{\lambda}^\ell)^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle$ from (4.9) yields

$$\begin{aligned}\widetilde{A}_t(\boldsymbol{\vartheta}) = &-\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t b_s(1-L^{-1}) \\ &\qquad\times\left(\boldsymbol{\vartheta}_s^\ell + \frac{2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s(1-L^{-1})}\boldsymbol{\lambda}_s^\ell\right)^2 \mathrm{d}\langle \mathbf{M}^\ell\rangle_s \\ &+ \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \frac{(2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1)^2}{4b_s(1-L^{-1})}\,\mathrm{d}\mathbf{K}_s^\ell \\ &- \int_0^t \mathrm{evar}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\,\mathrm{d}b_s + c_t. \end{aligned} \tag{4.18}$$

The monotonicity of the first term in (4.18) completely depends on the sign of the process b. To analyse the other terms in (4.18), we introduce the following condition.

Condition 4.4. The process $\mathbf{K}^\ell$ is independent of ℓ, meaning that $\mathbf{K}^\ell, \mathbf{K}^m$ are indistinguishable for $\ell \neq m$.

In view of Lemma 3.13 and under Condition 4.4, each $\mathbf{K}^\ell$ is indistinguishable from the MVT process K. We use that fact, then square out and finally use the definitions of $\mathrm{em}(\mathbf{x})$ and $\mathrm{evar}(\mathbf{x})$ from (4.2) and (4.3), respectively, to write the second line in (4.18) as

$$\begin{aligned}&\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \frac{(2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1)^2}{4b_s(1-L^{-1})}\,\mathrm{d}\mathbf{K}_s^\ell \\ &= \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \frac{(2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1)^2}{4b_s(1-L^{-1})}\,\mathrm{d}K_s \\ &= \frac{1}{L}\int_0^t \sum_{\ell=1}^{L} \frac{4b_s^2(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta})))^2 + 1}{4b_s(1-L^{-1})}\,\mathrm{d}K_s \\ &\quad - \int_0^t \frac{1}{4b_s(1-L^{-1})}\sum_{\ell=1}^{L}\frac{4b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta})))}{L}\,\mathrm{d}K_s \\ &= \int_0^t \frac{4b_s^2\mathrm{evar}(\mathbf{G}_s(\boldsymbol{\vartheta})) + 1}{4b_s(1-L^{-1})}\,\mathrm{d}K_s \\ &= \int_0^t \left(\frac{b_s\mathrm{evar}(\mathbf{G}_s(\boldsymbol{\vartheta}))}{1-L^{-1}} + \frac{1}{4b_s(1-L^{-1})}\right)\mathrm{d}K_s. \end{aligned} \tag{4.19}$$

The last expression in (4.19) significantly simplifies the middle line in (4.18). Since this is an integral with respect to the MVT process K, it is convenient to introduce the following condition so that we can write the last two lines in (4.18) as a single expression to conduct further analysis. Recall from the definition (3.3) of (SC) that the process B is chosen such that $\langle M\rangle \ll B$ when $d=1$. For simplicity, we take $B = \langle M\rangle$.

Condition 4.5. The processes b, c in (4.4) are absolutely continuous with respect to the process $\langle M\rangle$.

Plugging (4.19) back into (4.18) and using Condition 4.5, we get

$$
\begin{aligned}
\widetilde{A}_t(\boldsymbol{\vartheta}) &= -\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t b_s(1-L^{-1}) \\
&\qquad\qquad \times \left(\boldsymbol{\vartheta}_s^\ell + \frac{2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s(1-L^{-1})}\boldsymbol{\lambda}_s^\ell\right)^2 \mathrm{d}\langle \mathbf{M}^\ell\rangle_s \\
&\quad + \int_0^t \bigg(\Big(\frac{b_s \mathrm{evar}(\mathbf{G}_s(\boldsymbol{\vartheta}))}{1-L^{-1}} + \frac{1}{4b_s(1-L^{-1})}\Big)\frac{\mathrm{d}K_s}{\mathrm{d}\langle M\rangle_s} - \mathrm{evar}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\frac{\mathrm{d}b_s}{\mathrm{d}\langle M\rangle_s} \\
&\qquad\qquad + \frac{\mathrm{d}c_s}{\mathrm{d}\langle M\rangle_s}\bigg)\mathrm{d}\langle M\rangle_s \\
&= -\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t b_s(1-L^{-1}) \\
&\qquad\qquad \times \left(\boldsymbol{\vartheta}_s^\ell + \frac{2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s(1-L^{-1})}\boldsymbol{\lambda}_s^\ell\right)^2 \mathrm{d}\langle \mathbf{M}^\ell\rangle_s \\
&\quad + \int_0^t \mathrm{evar}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big(\frac{b_s}{1-L^{-1}}\frac{\mathrm{d}K_s}{\mathrm{d}\langle M\rangle_s} - \frac{\mathrm{d}b_s}{\mathrm{d}\langle M\rangle_s}\Big)\mathrm{d}\langle M\rangle_s \\
&\quad + \int_0^t \frac{1}{4b_s(1-L^{-1})}\frac{\mathrm{d}K_s}{\mathrm{d}\langle M\rangle_s} + \frac{\mathrm{d}c_s}{\mathrm{d}\langle M\rangle_s}\,\mathrm{d}\langle M\rangle_s.
\end{aligned}
\tag{4.20}
$$

Now the (super)martingale property in (4.17), which should hold for any $\boldsymbol{\vartheta}$, suggests that we make the last two integrals vanish. Using

$$
\frac{\mathrm{d}K_t}{\mathrm{d}\langle M\rangle_t} = \lambda_t^2, \quad t\in[0,T], \tag{4.21}
$$

from (3.5) when $d=1$ and $B=\langle M\rangle$ together with the terminal conditions from (4.5), we obtain the system of backward random differential equations

$$
\begin{aligned}
&\frac{\mathrm{d}b}{\mathrm{d}\langle M\rangle} - \lambda^2\frac{b}{(1-L^{-1})} = 0, \quad b_T = \xi, \\
&\frac{\mathrm{d}c}{\mathrm{d}\langle M\rangle} + \frac{1}{4b(1-L^{-1})} = 0, \quad c_T = 0.
\end{aligned}
\tag{4.22}
$$

If the processes b, c are solutions to (4.22), then the last two lines in (4.20) vanish, and inserting the resulting simplified (4.20) back into (4.16) yields

$$\begin{aligned}\widetilde{V}_t(\boldsymbol{\vartheta}) &= \widetilde{V}_0 + \widetilde{M}_t(\boldsymbol{\vartheta}) \\ &\quad - \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t b_s(1-L^{-1}) \\ &\qquad \times \left(\boldsymbol{\vartheta}_s^\ell + \frac{2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s(1-L^{-1})}\boldsymbol{\lambda}_s^\ell\right)^2 \mathrm{d}\langle \mathbf{M}^\ell\rangle_s. \quad (4.23)\end{aligned}$$

By (4.22), the process b is given by a continuous (stochastic) exponential and hence positive; so we readily see that the process $\widetilde{V}(\boldsymbol{\vartheta})$ is a *local* supermartingale in the filtration $\mathbb{G}$ for any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}$ because it is a local martingale plus a decreasing process by (4.23). Moreover, (4.23) suggests that $\widetilde{V}(\boldsymbol{\vartheta})$ is a local martingale if and only if $\boldsymbol{\vartheta}$ makes the integrand in (4.23) vanish and hence formally satisfies

$$\boldsymbol{\vartheta}_s^\ell = -\frac{2b_s(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s(1-L^{-1})}\boldsymbol{\lambda}_s^\ell, \quad s \in [0,T], \ell = 1, \ldots, L. \qquad (4.24)$$

(4.24) also gives a candidate for a solution of the auxiliary problem (2.9).

Let us briefly recap this subsection. We assume (4.4) and start with a (simplified) canonical decomposition of $\widetilde{V}(\boldsymbol{\vartheta})$ as in (4.16). Then we manipulate the terms (mainly by completing a square) and introduce some additional conditions so that we arrive at a system of integral equations (4.22) as a (promising) sufficient condition for the (super)martingale property of $\widetilde{V}(\boldsymbol{\vartheta})$. The simplified expression (4.23) also suggests a candidate (4.24) for the optimal strategy to the auxiliary problem (2.9). We now end this heuristic subsection with more detailed instructions to implement Recipe 4.1, 3).

Recipe 4.6. 1) Construct processes b, c by solving the system (4.22) of random ODEs. Note that this also a posteriori verifies Condition 4.5.

2) Verify that Condition 4.4 is true, possibly under additional assumptions.

3) Prove that the resulting $\widetilde{V}(\boldsymbol{\vartheta})$ has the properties desired in Lemma 2.4.

Note that up to here, we make guesses and assumptions and used heuristic arguments. In the sequel, we now provide rigorous results.

4.3 Implementing Recipe 4.6 – construction and verification

In this subsection, we fix $L \in \mathbb{N}$ with $L \geq 2$ and follow Recipe 4.6 to construct, based on the explicit formula (4.24), a candidate $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for the optimal strategy to the auxiliary problem (2.9). First, we solve the system (4.22) explicitly, derive from (4.24) a system of affine SDEs and study the existence and uniqueness of a solution to that system, which rigorously yields a candidate $\widehat{\boldsymbol{\vartheta}}^{(L)}$. Next, we argue that Condition 4.4 is equivalent to the assumption that the MVT process K is deterministic, as in discrete time. Finally, we verify that $\widehat{\boldsymbol{\vartheta}}^{(L)}$ is indeed optimal for the auxiliary problem (2.9) by proving that $\widehat{\boldsymbol{\vartheta}}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}$ and that the conditions in Lemma 2.4 are satisfied. **Throughout this subsection, we keep the superscript $^{(L)}$.**

Recall that Assumption 4.2 says that S is a continuous semimartingale satisfying the structure (SC). We now implement Recipe 4.6, 1) and solve the system (4.22) of ODES

$$\begin{aligned} \frac{\mathrm{d}b}{\mathrm{d}\langle M\rangle} - \lambda^2 \frac{b}{(1-L^{-1})} &= 0, \quad b_T = \xi, \\ \frac{\mathrm{d}c}{\mathrm{d}\langle M\rangle} + \frac{1}{4b(1-L^{-1})} &= 0, \quad c_T = 0. \end{aligned} \tag{4.25}$$

Lemma 4.7. *Suppose that Assumption* 4.2 *is satisfied. Then the system* (4.25) *has a unique strong solution* $(b^{(L)}, c^{(L)})$ *in the space of* $\mathbb{F}$*-adapted continuous processes on* $(\Omega, \mathcal{F}, P)$*. Explicitly, the processes* $b^{(L)}, c^{(L)}$ *are continuous and given by*

$$b_t^{(L)} = \xi \frac{\mathcal{E}(\frac{K}{1-L^{-1}})_t}{\mathcal{E}(\frac{K}{1-L^{-1}})_T} = \xi \exp\left(\frac{K_t - K_T}{1-L^{-1}}\right), \quad t \in [0,T], \tag{4.26}$$

$$c_t^{(L)} = -\frac{1}{4\xi}\left(1 - \frac{\xi}{b_t^{(L)}}\right) = -\frac{1}{4\xi}\left(1 - \exp\left(-\frac{K_t - K_T}{1-L^{-1}}\right)\right), \quad t \in [0,T]. \tag{4.27}$$

Consequently, both $b^{(L)}$ *and* $\frac{1}{b^{(L)}}$ *are positive and bounded on* $[0,T]$ *(uniformly in* ω*) whenever* K *is bounded.*

Proof. It is straightforward to differentiate (4.26) and (4.27) and verify that they indeed solve the system (4.25) with $\mathrm{d}K = \lambda^2\,\mathrm{d}\langle M\rangle$ as in (4.21). For uniqueness of $b^{(L)}$, it is enough to observe that its equation in (4.25) is a P-a.s. linear (random) ODE and thus has a unique solution. The uniqueness of $c^{(L)}$ is obtained from the uniqueness of $b^{(L)}$. The last assertion follows from (4.26). □

We translate this result into the L-extended market. Recall from (4.8) and (4.9) with the superscript $^{(L)}$ reinstated that

$$\mathbf{A}_t^{\ell,(L)} = \int_0^t \boldsymbol{\lambda}_s^{\ell,(L)} \,\mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle_s \quad \mathbf{P}\text{-a.s. for all } t \in [0,T],\ \ell = 1,\dots,L, \tag{4.28}$$

$$\mathbf{K}_t^{\ell,(L)} = \int_0^t (\boldsymbol{\lambda}_s^{\ell,(L)})^2 \,\mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle_s \quad \mathbf{P}\text{-a.s. for all } t \in [0,T],\ \ell = 1,\dots,L. \tag{4.29}$$

The system (4.25) translates into a corresponding system of ODEs

$$\begin{aligned}
\frac{\mathrm{d}\boldsymbol{b}^{\ell,(L)}}{\mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle} - (\boldsymbol{\lambda}^{\ell,(L)})^2 \frac{\boldsymbol{b}^{\ell,(L)}}{(1-L^{-1})} &= 0, \quad \boldsymbol{b}_T^{\ell,(L)} = \xi,\\
\frac{\mathrm{d}\boldsymbol{c}^{\ell,(L)}}{\mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle} + \frac{1}{4\boldsymbol{b}^{\ell,(L)}(1-L^{-1})} &= 0, \quad \boldsymbol{c}_T^{\ell,(L)} = 0,
\end{aligned} \tag{4.30}$$

for $\ell = 1,\dots,L$.

Lemma 4.8. *Suppose that Assumption* 4.2 *is satisfied. Then for* $\ell = 1,\dots,L$, *the system* (4.30) *of ODEs has a unique strong solution* $(\boldsymbol{b}^{\ell,(L)}, \boldsymbol{c}^{\ell(L)})$ *in the space of* $\mathbb{G}^{(L)}$*-adapted continuous processes defined on* $(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$. *Explicitly, the processes* $\boldsymbol{b}^{\ell,(L)}, \boldsymbol{c}^{\ell,(L)}$ *are continuous and given, for* $t \in [0,T]$, *by*

$$\boldsymbol{b}_t^{\ell,(L)} = \xi \frac{\mathcal{E}\big(\frac{\mathbf{K}^{\ell,(L)}}{1-L^{-1}}\big)_t}{\mathcal{E}\big(\frac{\mathbf{K}^{\ell,(L)}}{1-L^{-1}}\big)_T} = \xi \exp\left(\frac{\mathbf{K}_t^{\ell,(L)} - \mathbf{K}_T^{\ell,(L)}}{1-L^{-1}}\right), \tag{4.31}$$

$$\boldsymbol{c}_t^{\ell,(L)} = -\frac{1}{4\xi}\left(1 - \frac{\xi}{\boldsymbol{b}_t^{\ell,(L)}}\right) = -\frac{1}{4\xi}\left(1 - \exp\Big(-\frac{\mathbf{K}_t^{\ell,(L)} - \mathbf{K}_T^{\ell,(L)}}{1-L^{-1}}\Big)\right). \tag{4.32}$$

Consequently, both $\boldsymbol{b}^{\ell,(L)}$ *and* $\frac{1}{\boldsymbol{b}^{\ell,(L)}}$ *are positive and bounded on* $[0,T]$ *(uniformly in* $\omega^{(L)}$*) whenever* K *is bounded.*

Proof. This is directly translated from Lemma 4.7. The last statement follows from the fact that the boundedness of K implies the boundedness of $\mathbf{K}^{\ell,(L)}$ for $\ell = 1,\dots,L$ due to the relation $\mathbf{K}^{\ell,(L)} = K^{\ell,\otimes L}$ from (3.24), using the lifting notation (2.14). □

Given (4.26) and (4.27), we now study (4.24) rigorously. Consider the system of affine SDEs

$$\mathrm{d}\mathbf{X}_t^{\ell,(L)} = -\frac{2\boldsymbol{b}_t^{\ell,(L)}(\mathbf{X}_t^{\ell,(L)} - \mathrm{em}(\mathbf{X}_t^{(L)})) - 1}{2\boldsymbol{b}_t^{\ell,(L)}(1-L^{-1})} \boldsymbol{\lambda}_t^{\ell,(L)} \,\mathrm{d}\mathbf{S}_t^{\ell,(L)}, \quad \mathbf{X}_0^{\ell,(L)} = 0, \tag{4.33}$$

for $\ell = 1,\dots,L$.

Lemma 4.9. *Suppose that Assumption* 4.2 *is satisfied. Then the system* (4.33) *has a unique strong solution in the class of* $\mathbb{G}^{(L)}$*-adapted continuous processes. If in addition the MVT process* K *is absolutely continuous with respect to a deterministic and increasing process* D*, with density* $\kappa = \frac{\mathrm{d}K}{\mathrm{d}D}$ *bounded by a constant* C_0*, then there exist* $C_1, C_2 > 0$ *such that*

$$\mathbf{E}\Big[\sup_{s\in[0,t]} |\mathbf{X}_s^{(L)}|^2\Big] \le C_1 \exp\big(C_2(C_0^2 + C_0)\big).$$

Proof. The function

$$f_t^{\ell,(L)}(\mathbf{x}) := \frac{1 - 2\boldsymbol{b}_t^{\ell,(L)}(\mathbf{x}^\ell - \mathrm{em}(\mathbf{x}))}{2\boldsymbol{b}_t^{\ell,(L)}(1 - L^{-1})}\boldsymbol{\lambda}_t^{\ell,(L)} \tag{4.34}$$

is well defined thanks to $\boldsymbol{b}_t^{\ell,(L)} > 0$ from the last assertion in Lemma 4.8, and affine in $\mathbf{x} \in \mathbb{R}^L$ and hence Lipschitz (with a time-dependent $\mathbb{G}^{(L)}$-adapted coefficient). The existence and uniqueness of a strong solution to (4.33) therefore follows from standard theory; see e.g. Protter [55, Theorem V.6]. Note that (4.33) can be written more compactly using (4.34) as

$$\mathrm{d}\mathbf{X}_t^{\ell,(L)} = f_t^{\ell,(L)}(\mathbf{X}_t^{(L)})\,\mathrm{d}\mathbf{S}_t^{\ell,(L)}, \quad \mathbf{X}_0^{\ell,(L)} = 0. \tag{4.35}$$

Now suppose that Assumption 4.2 is satisfied and $K \ll D$ for a deterministic and increasing process D such that $\kappa = \frac{\mathrm{d}K}{\mathrm{d}D}$ is bounded by a constant C_0. The identity $\mathbf{K}^{\ell,(L)}(\omega^{(L)}) = K^{\ell,\otimes L}(\omega^{(L)}) = K(\omega_\ell)$ from (3.24) then implies that $\mathbf{K}^{\ell,(L)} \ll D$ with $\kappa^{\ell,(L)} := \frac{\mathrm{d}\mathbf{K}^{\ell,(L)}}{\mathrm{d}D} = \kappa^{\ell,\otimes L}$ bounded again by C_0. We summarise this as

$$\sup_{s\in[0,T]} |\kappa_s^{\ell,L}| \le C_0, \text{ uniformly on } \Omega^{(L)}, \tag{4.36}$$

for future reference. From the explicit expressions (4.28) for $\mathbf{A}^{\ell,(L)}$ and (4.29) for $\mathbf{K}^{\ell,(L)}$, we obtain

$$\mathrm{d}\mathbf{K}^{\ell,(L)} = (\boldsymbol{\lambda}^{\ell,(L)})^2\,\mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle = \boldsymbol{\lambda}^{\ell,(L)}\,\mathrm{d}\mathbf{A}^{\ell,(L)} = \kappa^{\ell,(L)}\,\mathrm{d}D. \tag{4.37}$$

Define $g_t^{(L)}(\mathbf{x})$ by

$$g_t^{\ell,(L)}(\mathbf{x}) = \frac{1 - 2\boldsymbol{b}_t^{\ell,(L)}(\mathbf{x}^\ell - \mathrm{em}(\mathbf{x}))}{2\boldsymbol{b}_t^{\ell,(L)}(1 - L^{-1})}, \quad \ell = 1, \dots, L. \tag{4.38}$$

Comparing (4.38) with (4.34) for $f_t^{\ell,(L)}$ yields $f_t^{\ell,(L)} = g_t^{\ell,(L)}\boldsymbol{\lambda}_t^{\ell,(L)}$. We use the elementary inequality $(x+y)^2 \le 2x^2 + 2y^2$, then the above identity and finally the Cauchy–Schwarz inequality and (4.37), (4.36) to get from (4.35) that

$$\begin{aligned}(\mathbf{X}_s^{\ell,(L)})^2 &\le 2\Big(\int_0^s f_r^{\ell,(L)}(\mathbf{X}_r^{(L)})\,\mathrm{d}\mathbf{A}_r^{\ell,(L)}\Big)^2 + 2\Big(\int_0^s f_r^{\ell,(L)}(\mathbf{X}_r^{(L)})\,\mathrm{d}\mathbf{M}_r^{\ell,(L)}\Big)^2\\ &= 2\Big(\int_0^s g_r^{\ell,(L)}(\mathbf{X}_r^{(L)})\,\mathrm{d}\mathbf{K}_r^{\ell,(L)}\Big)^2 + 2\Big(\int_0^s f_r^{\ell,(L)}(\mathbf{X}_r^{(L)})\,\mathrm{d}\mathbf{M}_r^{\ell,(L)}\Big)^2\\ &\le 2C_0^2\int_0^s \big(g_r^{\ell,(L)}(\mathbf{X}_r^{(L)})\big)^2\,\mathrm{d}D_r + 2\Big(\int_0^s f_r^{\ell,(L)}(\mathbf{X}_r^{(L)})\,\mathrm{d}\mathbf{M}_r^{\ell,(L)}\Big)^2. \qquad (4.39)\end{aligned}$$

Using (4.39), the identity $\sup_{s\in[0,t]}|\int_0^s (g_r^{\ell,(L)}(\mathbf{X}_r^{(L)}))^2\,\mathrm{d}D_r| = \int_0^t (g_r^{\ell,(L)}(\mathbf{X}_r^{(L)}))^2\,\mathrm{d}D_r$ and the BDG inequality yields

$$\begin{aligned}\mathbf{E}\Big[\sup_{s\in[0,t]}(\mathbf{X}_s^{\ell,(L)})^2\Big] &\le \mathbf{E}\Big[2C_0^2\int_0^t \big(g_s^{\ell,(L)}(\mathbf{X}_s^{(L)})\big)^2\,\mathrm{d}D_s\\ &\qquad + 8\int_0^t \big(f_s^{\ell,(L)}(\mathbf{X}_s^{(L)})\big)^2\,\mathrm{d}\langle\mathbf{M}^{\ell,(L)}\rangle_s\Big]\\ &= \mathbf{E}\Big[2C_0^2\int_0^t \big(g_s^{\ell,(L)}(\mathbf{X}_s^{(L)})\big)^2\,\mathrm{d}D_s\\ &\qquad + 8\int_0^t \big(g_s^{\ell,(L)}(\mathbf{X}_s^{(L)})\big)^2\kappa_s^{\ell,(L)}\,\mathrm{d}D_s\Big]\\ &= (2C_0^2 + 8C_0)\int_0^t \mathbf{E}\big[\big(g_s^{\ell,(L)}(\mathbf{X}_s^{(L)})\big)^2\big]\,\mathrm{d}D_s. \qquad (4.40)\end{aligned}$$

The first equality uses $f_t^{\ell,(L)}(\mathbf{x}) = g_t^{\ell,(L)}(\mathbf{x})\boldsymbol{\lambda}_t^{\ell,(L)}$ and $\mathrm{d}\mathbf{K}^{\ell,(L)} = \kappa^{\ell,(L)}\,\mathrm{d}D$ from (4.37), and the second uses the bound on $\kappa^{\ell,(L)}$ from (4.36) and Fubini's theorem due to the non-randomness of the increasing process D. Denote

$$g_t^0 := g_t^{\ell,(L)}(0) = \frac{1}{2\boldsymbol{b}_t^{\ell,(L)}(1-L^{-1})} \ge 0.$$

Because $g_t^{\ell,(L)}$ is Lipschitz-continuous, we use the triangle inequality to get

$$\begin{aligned}|g_t^{\ell,(L)}(\mathbf{x})| &\le g_t^{\ell,(L)}(0) + |g_t^{\ell,(L)}(\mathbf{x}) - g_t^{\ell,(L)}(0)|\\ &= g_t^0 + \frac{|\mathbf{x}^\ell - \mathrm{em}(\mathbf{x})|}{1-L^{-1}}\\ &\le g_t^0 + C_{\ell,(L)}|\mathbf{x}| \qquad (4.41)\end{aligned}$$

for some constant $C_{\ell,(L)} > 0$. By the symmetry from

$$g_t^{\ell,(L)}(\mathbf{x}) - g_t^{\ell,(L)}(0) = -\frac{\mathbf{x}^\ell - \mathrm{em}(\mathbf{x})}{1 - L^{-1}}, \quad \ell = 1, \dots, L,$$

the constant $C_{\ell,(L)}$ can be chosen independent of ℓ; so we denote this common value by $C_{(L)}$. Inserting (4.41) with $C_{\ell,(L)} = C_{(L)}$ and $(x+y)^2 \le 2x^2 + 2y^2$ into (4.40), we get

$$\begin{aligned}
\mathbf{E}\Big[\sup_{s\in[0,t]} |\mathbf{X}_s^{(L)}|^2\Big] &= \mathbf{E}\Big[\sup_{s\in[0,t]} \sum_{\ell=1}^{L} (\mathbf{X}_s^{\ell,(L)})^2\Big] \\
&\le (2C_0^2 + 8C_0) \int_0^t \sum_{\ell=1}^{L} \mathbf{E}\big[\big(g_s^{\ell,(L)}(\mathbf{X}_s^{(L)})\big)^2\big] \, \mathrm{d}D_s \\
&\le L(2C_0^2 + 8C_0) \int_0^t \big(2\mathbf{E}[(g_s^0)^2] + 2C_{(L)}^2 \mathbf{E}[|\mathbf{X}_s^{(L)}|^2]\big) \, \mathrm{d}D_s \\
&\le L(2C_0^2 + 8C_0) \int_0^t \Big(2\mathbf{E}[(g_s^0)^2] + 2C_{(L)}^2 \mathbf{E}\Big[\sup_{r\le s} |\mathbf{X}_r^{(L)}|^2\Big]\Big) \, \mathrm{d}D_s.
\end{aligned}$$

It follows from Gronwall's lemma that

$$\mathbf{E}\Big[\sup_{s\in[0,t]} |\mathbf{X}_s^{(L)}|^2\Big] \le C \exp\big(LC_{(L)}^2 (4C_0 + 16) C_0\big) < \infty$$

for $C = L(2C_0^2 + 8C_0) D_T \sup_{t\in[0,T]} \mathbf{E}[2(g_t^0)^2]$. This completes the proof. □

Lemma 4.9 rigorously justifies the system (4.24) of formal equations because it allows us to define $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for $t \in [0,T]$ by

$$\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} = -\frac{2\boldsymbol{b}_t^{\ell,(L)}(\mathbf{X}_t^{\ell,(L)} - \mathrm{em}(\mathbf{X}_t^{(L)})) - 1}{2\boldsymbol{b}_t^{\ell,(L)}(1 - L^{-1})} \boldsymbol{\lambda}_t^\ell = f_t^{\ell,(L)}(\mathbf{X}_t^{\ell,(L)}), \ \ell = 1, \dots, L. \tag{4.42}$$

Then (4.42) and (4.35) yield the identity

$$\mathbf{G}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) = \int \widehat{\boldsymbol{\vartheta}}^{\ell,(L)} \, \mathrm{d}\mathbf{S}^{\ell,(L)} = \int f^{\ell,(L)}(\mathbf{X}^{(L)}) \, \mathrm{d}\mathbf{S}^{\ell,(L)} = \mathbf{X}^{\ell,(L)} \tag{4.43}$$

because $\mathbf{X}_0^{\ell,(L)} = 0 = \mathbf{G}_0^{\ell,(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})$. Replacing $\mathbf{X}^{\ell,(L)}$ by $\mathbf{G}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})$ in (4.42) thus gives for $\ell = 1, \dots, L$ that

$$\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} = -\frac{2\boldsymbol{b}_t^{\ell,(L)}(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})) - 1}{2\boldsymbol{b}_t^{\ell,(L)}(1 - L^{-1})} \boldsymbol{\lambda}_t^\ell, \quad t \in [0,T]. \tag{4.44}$$

This gives a rigorous formulation for (4.24).

Corollary 4.10. *Suppose that Assumption 4.2 is satisfied. If the MVT process K is absolutely continuous with respect to D for a deterministic and increasing process D with a density $\kappa = \frac{\mathrm{d}K}{\mathrm{d}D}$ bounded by a constant C_0, then the process $\mathbf{G}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})$ is in $\mathcal{S}^2$ for each $\ell = 1, \dots, L$. In particular, $\widehat{\boldsymbol{\vartheta}}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$.*

Proof. Because $K_T = \int_0^T \kappa_t \,\mathrm{d}D_t \le C_0 D_T$, the MVT process K is bounded. So we can apply Lemma 3.17 to get $\boldsymbol{\Theta}_{\mathrm{S}}^{(L)} = L^2(\mathbf{M}^{(L)})$. Now we only need to show that $\widehat{\boldsymbol{\vartheta}}^{(L)} \in L^2(\mathbf{M})$. Using Definition 3.15 with $\boldsymbol{\sigma}^\ell \equiv 1$ and $\mathbf{B}^{\ell,(L)} = \langle \mathbf{M}^{\ell,(L)} \rangle$ for $\ell = 1, \dots, L$ in dimension 1 by (4.7) and the line above it, we compute

$$
\begin{aligned}
\mathbf{E}\bigg[\int_0^T \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)} \boldsymbol{\sigma}_s^{\ell,(L)} \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)} \,\mathrm{d}\mathbf{B}_s^{\ell,(L)}\bigg] &= \mathbf{E}\bigg[\int_0^T (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)})^2 \,\mathrm{d}\langle \mathbf{M}^{\ell,(L)} \rangle_s\bigg] \\
&\le 2C_0 \int_0^T \mathbf{E}\bigg[\bigg(\frac{(\mathbf{X}_s^{\ell,(L)} - \mathrm{em}(\mathbf{X}_s^{(L)}))}{1 - L^{-1}}\bigg)^2\bigg] \,\mathrm{d}D_s \\
&\quad + 2C_0 \int_0^T \mathbf{E}\bigg[\bigg(\frac{1}{2\boldsymbol{b}_s^{\ell,(L)}(1 - L^{-1})}\bigg)^2\bigg] \,\mathrm{d}D_s \\
&\le \frac{2C_0}{(1 - L^{-1})^2} D_T \mathbf{E}\Big[\sup_{s \in [0,T]} \big(\mathbf{X}_s^{\ell,(L)} - \mathrm{em}(\mathbf{X}_s^{(L)})\big)^2\Big] \\
&\quad + 2C_0 \int_0^T \mathbf{E}\bigg[\bigg(\frac{1}{2\boldsymbol{b}_s^{\ell,(L)}(1 - L^{-1})}\bigg)^2\bigg] \,\mathrm{d}D_s \\
&< \infty.
\end{aligned}
$$

The first inequality uses the explicit formula (4.44) for the candidate $\widehat{\boldsymbol{\vartheta}}^{(L)}$, the identity $(\boldsymbol{\lambda}^{\ell,(L)})^2 \,\mathrm{d}\langle \mathbf{M}^{\ell,(L)} \rangle = \kappa^{\ell,(L)} \,\mathrm{d}D$ from (4.37) and the bound on $\kappa^{\ell,(L)}$ from (4.36) as well as Fubini's theorem because D is non-random. The last inequality uses $\mathbf{E}[\sup_{s\in[0,T]} |\mathbf{X}_s^{\ell,(L)}|^2] < \infty$ for $\ell = 1, \dots, L$ from Lemma 4.9 and that $\frac{1}{\boldsymbol{b}^{\ell,(L)}}$ is bounded due to the boundedness of K; see Lemma 4.8. □

We now discuss Condition 4.4. Due to Lemma 3.13, 1), each process $\mathbf{K}^{\ell,(L)}$ is also $\mathbb{F}^{\ell,(L)}$-adapted, and hence $(\mathbf{K}^{\ell,(L)})_{\ell=1,\dots,L}$ are independent. Then Condition 4.4 implies that each $\mathbf{K}^{\ell,(L)}$ is indistinguishable from $\mathbf{K}^{j,(L)}$ for $\ell \ne j$ and is also independent of $\mathbf{K}^{j,(L)}$. Thus it is independent of itself and therefore must be deterministic. We remarked below Condition 4.4 that $\mathbf{K}^{\ell,(L)}$ and the MVT process K are indistinguishable, and so the previous reasoning implies that K is deterministic. Conversely, if the MVT process K is deterministic, then Condition 4.4 is automatically satisfied. Therefore, the following assumption is an equivalent restatement of Condition 4.4, which is also needed in discrete time (see Assumption I.4.11).

Assumption 4.11. The mean–variance tradeoff process K is deterministic.

An immediate consequence of Assumption 4.11 is that $\mathbf{K}^{\ell,(L)} = K$ and thus the processes $b^{(L)}$ and $c^{(L)}$ are deterministic. Moreover, in Lemma 4.9 and Corollary 4.10, we can take $D \equiv K$ and $\kappa \equiv 1$.

Corollary 4.12. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied. Then* $b^{(L)}, c^{(L)}$ *given by* (4.26) *and* (4.27) *are deterministic and thus* $\boldsymbol{b}^{\ell,(L)} = b^{(L)}$ *and* $\boldsymbol{c}^{\ell,(L)} = c^{(L)}$ *for* $\boldsymbol{b}^{\ell,(L)}, \boldsymbol{c}^{\ell,(L)}$ *from Lemma* 4.8.

Proof. The first claim directly follows from the explicit formulas

$$b_t^{(L)} = \xi \exp\Big(\frac{K_t - K_T}{1 - L^{-1}}\Big), \quad c_t^{(L)} = -\frac{1}{4\xi}\bigg(1 - \exp\Big(\frac{K_T - K_t}{1 - L^{-1}}\Big)\bigg)$$

given in (4.26) and (4.27). The second follows from comparing formula (4.31) for $\boldsymbol{b}^{\ell,(L)}$ and (4.32) for $\boldsymbol{c}^{\ell,(L)}$ with the above display and using that $\mathbf{K}^{\ell,(L)} = K$ for $\ell = 1, \dots, L$ when K is deterministic. □

Inspired by the guess (4.4) with the process a there chosen equal to constant 1, we consider $\widetilde{\mathbf{V}}^{(L)}(\boldsymbol{\vartheta})$ with coordinates

$$\widetilde{\mathbf{V}}^{\ell,(L)}(\boldsymbol{\vartheta}) := \mathrm{em}\big(\mathbf{G}_t(\boldsymbol{\vartheta})\big) - \boldsymbol{b}_t^{\ell,(L)}\mathrm{evar}\big(\mathbf{G}_t(\boldsymbol{\vartheta})\big) + \boldsymbol{c}_t^{\ell,(L)}, \quad \boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}, t \in [0,T],$$
$$\text{where the coefficient processes } (\boldsymbol{b}_t^{\ell,(L)})_{t\in[0,T]}, (\boldsymbol{c}_t^{\ell,(L)})_{t\in[0,T]} \text{ satisfy (4.30).}$$

When the MVT process K is deterministic, we have $\boldsymbol{b}^{\ell,(L)} = b^{(L)}$ and $\boldsymbol{c}^{\ell,(L)} = c^{(L)}$ due to Corollary 4.12 and thus $\widetilde{\mathbf{V}}^{\ell,(L)}(\boldsymbol{\vartheta}) = \widetilde{V}^{(L)}(\boldsymbol{\vartheta})$ for $\ell = 1, \dots, L$, with $\widetilde{V}^{(L)}(\boldsymbol{\vartheta})$ given by

$$\widetilde{V}_t^{(L)}(\boldsymbol{\vartheta}) := \mathrm{em}\big(\mathbf{G}_t(\boldsymbol{\vartheta})\big) - b_t^{(L)}\mathrm{evar}\big(\mathbf{G}_t(\boldsymbol{\vartheta})\big) + c_t^{(L)}, \quad \boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}, t \in [0,T],$$
$$\text{where the coefficient processes } (b_t^{(L)})_{t\in[0,T]}, (c_t^{(L)})_{t\in[0,T]} \text{ satisfy (4.22).} \tag{4.45}$$

Let us now work with $\widetilde{V}^{(L)}$ given above. The canonical decomposition obtained in Lemma 4.3 under the assumption (4.4) can be further simplified if we look at $\widetilde{V}^{(L)}(\boldsymbol{\vartheta})$.

Lemma 4.13. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied. Then for any* $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$*, the canonical decomposition of* $\widetilde{V}^{(L)}(\boldsymbol{\vartheta})$ *defined in* (4.45) *is given* *by*

$$\begin{aligned}
&\widetilde{V}_t^{(L)}(\boldsymbol{\vartheta}) \\
&= \widetilde{V}_0^{(L)} + \widetilde{M}_t^{(L)}(\boldsymbol{\vartheta}) + \widetilde{A}_t^{(L)}(\boldsymbol{\vartheta}) \\
&= \widetilde{V}_0^{(L)} + \widetilde{M}_t^{(L)}(\boldsymbol{\vartheta}) \\
&\quad - \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t b_s^{(L)}(1-L^{-1}) \\
&\qquad\qquad \times \left(\boldsymbol{\vartheta}_s^{\ell,(L)} + \frac{2b_s^{(L)}(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s^{(L)}(1-L^{-1})}\boldsymbol{\lambda}_s^{\ell,(L)}\right)^2 \mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle_s,
\end{aligned} \tag{4.46}$$

where $\widetilde{M}^{(L)}(\boldsymbol{\vartheta})$ *is a local martingale in the filtration* $\mathbb{G}^{(L)}$*.*

Proof. For this argument, we repeat the computation from (4.16) to (4.23). First, we insert $a_t^{(L)} = 1$ for $t \in [0,T]$ into the canonical decomposition (4.10) for $\widetilde{V}^{(L)}(\boldsymbol{\vartheta})$ to obtain $\widetilde{V}_t^{(L)}(\boldsymbol{\vartheta}) = \widetilde{V}_0 + \widetilde{M}_t^{(L)}(\boldsymbol{\vartheta}) + \widetilde{A}_t^{(L)}(\boldsymbol{\vartheta})$, where $\widetilde{M}^{(L)}(\boldsymbol{\vartheta})$ is a local martingale in the filtration $\mathbb{G}^{(L)}$ and the process $\widetilde{A}^{(L)}(\boldsymbol{\vartheta})$ is given as in (4.16) and (4.18) by

$$\begin{aligned}
&\widetilde{A}_t^{(L)}(\boldsymbol{\vartheta}) \\
&= -\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \Bigg(\Big(2b_s^{(L)}\Big(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big) - 1\Big)\boldsymbol{\vartheta}_s^{\ell,(L)}\boldsymbol{\lambda}_s^{\ell,(L)} \\
&\qquad\qquad + b_s^{(L)}(1-L^{-1})(\boldsymbol{\vartheta}_s^{\ell,(L)})^2\Bigg)\,\mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle_s \\
&\quad - \int_0^t \mathrm{evar}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\,\mathrm{d}b_s^{(L)} + c_t^{(L)} \\
&= -\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t b_s^{(L)}(1-L^{-1}) \\
&\qquad\qquad \left(\boldsymbol{\vartheta}_s^{\ell,(L)} + \frac{2b_s^{(L)}(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s^{(L)}(1-L^{-1})}\boldsymbol{\lambda}_s^{\ell,(L)}\right)^2 \mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle_s \\
&\quad + \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \frac{(2b_s^{(L)}(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1)^2}{4b_s^{(L)}(1-L^{-1})}\,\mathrm{d}\mathbf{K}_s^{\ell,(L)} \\
&\quad - \int_0^t \mathrm{evar}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\,\mathrm{d}b_s^{(L)} + c_t^{(L)}.
\end{aligned} \tag{4.47}$$

The second equality is derived exactly as in (4.18) by using the elementary identity $\alpha x^2 + \beta x = \alpha(x + \frac{\beta}{2\alpha})^2 - \frac{\beta^2}{4\alpha}$ and $\mathrm{d}\mathbf{K}^{\ell,(L)} = (\boldsymbol{\lambda}^{\ell,(L)})^2\,\mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle$ from (4.29). Because

K is deterministic by Assumption 4.11, we have $\mathbf{K}^{\ell,(L)} = K$ for $\ell = 1, \dots, L$. As calculated in (4.19), we then obtain the crucial simplification

$$\begin{aligned}
&\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \frac{(2b_s^{(L)}(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1)^2}{4b_s^{(L)}(1-L^{-1})}\,\mathrm{d}\mathbf{K}_s^{\ell,(L)} \\
&= \int_0^t \left(\frac{b_s^{(L)}\mathrm{evar}(\mathbf{G}_s(\boldsymbol{\vartheta}))}{1-L^{-1}} + \frac{1}{4b_s^{(L)}(1-L^{-1})}\right)\mathrm{d}K_s. \qquad (4.48)
\end{aligned}$$

Now because $b^{(L)}, c^{(L)}$ by construction satisfy the system (4.22) of ODEs, we can insert (4.48) into (4.47) and use the equation for $b^{(L)}$ in (4.22) to write the last two terms in (4.47) as integrals with respect to $\langle M\rangle$ to obtain as in (4.23) that the last two lines in (4.47) cancel out. This yields

$$\begin{aligned}
&\widetilde{A}_t^{(L)}(\boldsymbol{\vartheta}) \\
&= -\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t b_s^{(L)}(1-L^{-1}) \\
&\qquad \times \left(\boldsymbol{\vartheta}_s^{\ell,(L)} + \frac{2b_s^{(L)}(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}))) - 1}{2b_s^{(L)}(1-L^{-1})}\boldsymbol{\lambda}_s^{\ell,(L)}\right)^2 \mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle_s
\end{aligned}$$

and hence (4.46). □

We are now ready to implement Recipe 4.6, 3) and present the main result of this section.

Theorem 4.14. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied, meaning that the price process S is a real-valued continuous semimartingale satisfying the structure condition (SC) and the MVT process K is deterministic. Then $\widetilde{V}^{(L)}(\boldsymbol{\vartheta})$ given by* (4.45) *satisfies $\widetilde{V}_T^{(L)}(\boldsymbol{\vartheta}) = J_T(\boldsymbol{\vartheta}) = \mathrm{em}(\mathbf{G}_T(\boldsymbol{\vartheta})) - \xi\mathrm{evar}(\mathbf{G}_T(\boldsymbol{\vartheta}))$ and is a supermartingale for any $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$. It is a martingale for $\boldsymbol{\vartheta}^* = \widehat{\boldsymbol{\vartheta}}^{(L)}$ defined by* (4.42)*, and we have explicitly for $\ell = 1, \dots, L$ that*

$$\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} = -\frac{2b_t^{(L)}(\mathbf{G}_t^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))) - 1}{2b_t^{(L)}(1-L^{-1})}\boldsymbol{\lambda}_t^{\ell}, \quad t \in [0,T]. \qquad (4.49)$$

In particular, $\widehat{\boldsymbol{\vartheta}}^{(L)}$ is an optimal strategy for the auxiliary problem (2.9).

Proof. Let $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$. We recall from Lemma 3.17 the property

$$\sup_{t\le T}|\mathbf{G}_t(\boldsymbol{\vartheta})| \in L^2. \qquad (4.50)$$

Thanks to Assumptions 4.2 and 4.11, we can use Lemma 4.13, (4.46) to obtain the local martingale $\widetilde{M}^{(L)}(\boldsymbol{\vartheta}) = \widetilde{V}^{(L)}(\boldsymbol{\vartheta}) - \widetilde{V}_0^{(L)} - \widetilde{A}^{(L)}(\boldsymbol{\vartheta})$ for

$$\begin{aligned}
&\widetilde{A}_t^{(L)}(\boldsymbol{\vartheta}) \\
&= -\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t b_s^{(L)}(1-L^{-1}) \\
&\qquad \times \left(\boldsymbol{\vartheta}_s^{\ell,(L)} + \frac{2b_s^{(L)}(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}^{(L)}))) - 1}{2b_s^{(L)}(1-L^{-1})}\boldsymbol{\lambda}_s^{\ell,(L)}\right)^2 \mathrm{d}\langle \mathbf{M}^{\ell,(L)}\rangle_s \\
&=: -\frac{1}{L}\sum_{\ell=1}^{L} R_t^{\ell,(L)}(\boldsymbol{\vartheta}), \qquad t\in[0,T].
\end{aligned} \tag{4.51}$$

From the first equality in (4.51) and $b^{(L)} > 0$ by Lemma 4.7, we observe that $\widetilde{A}^{(L)}(\boldsymbol{\vartheta})$ is a decreasing process for all $\boldsymbol{\vartheta} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}$. Because $\boldsymbol{\vartheta} = \widehat{\boldsymbol{\vartheta}}^{(L)}$ satisfies (4.49), we get that the integrand in the first equality of (4.51) vanishes and thus $\widetilde{A}^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)}) \equiv 0$. So $\widetilde{V}^{(L)}(\boldsymbol{\vartheta})$ is always equal to a local martingale $\widetilde{M}^{(L)}(\boldsymbol{\vartheta})$ plus a decreasing process $\widetilde{A}^{(L)}(\boldsymbol{\vartheta})$, and $\widetilde{V}^{(L)}(\boldsymbol{\vartheta})$ is equal to a local martingale when $\boldsymbol{\vartheta} = \widehat{\boldsymbol{\vartheta}}^{(L)}$. To deduce the first assertion of the theorem from this, we now prove that $\widetilde{M}^{(L)}(\boldsymbol{\vartheta})$ is a true martingale. To this end, it is sufficient to show that $\sup_{t\in[0,T]}|\widetilde{M}_t^{(L)}(\boldsymbol{\vartheta})| \in L^1$. But $\widetilde{V}_t^{(L)}(\boldsymbol{\vartheta})$ is an affine–quadratic transformation of $\mathbf{G}_t(\boldsymbol{\vartheta})$ by (4.45), and so we can use (4.50) with $\mathrm{evar}(\mathbf{x}) \le \mathrm{em}(\mathbf{x}^2)$ by its definition (4.3) and that $b^{(L)}$ and $c^{(L)}$ are bounded due to (4.26) and (4.27) to obtain $\sup_{t\in[0,T]}|\widetilde{V}_t^{(L)}(\boldsymbol{\vartheta})| \in L^1$. Note that together with the integrability of $\widetilde{M}^{(L)}(\boldsymbol{\vartheta})$, this implies that also $\sup_{t\in[0,T]}|\widetilde{A}_t^{(L)}(\boldsymbol{\vartheta})| \in L^1$, and so the local supermartingale $\widetilde{V}^{(L)}(\boldsymbol{\vartheta})$ is a true supermartingale. For the term in the last equality of (4.51), we use the elementary inequalities $(x+y)^2 \le 2x^2 + 2y^2$, $1 \ge 1 - L^{-1} \ge \frac{1}{2}$ for $L \ge 2$ and $\mathbf{K}^{\ell,(L)} = \int(\boldsymbol{\lambda}^{\ell,(L)})^2\,\mathrm{d}\langle\mathbf{M}^{\ell,(L)}\rangle$ from (4.29) with $\mathbf{K}^{\ell,(L)} = K$ by Assumption 4.11 to get

$$\begin{aligned}
&\frac{1}{L}\sum_{\ell=1}^{L} R_t^{\ell,(L)}(\boldsymbol{\vartheta}) \\
&\le \frac{1}{L}\sum_{\ell=1}^{L} 2\int_0^t b_s^{(L)}(\boldsymbol{\vartheta}_s^{\ell})^2\,\mathrm{d}\langle\mathbf{M}^{\ell,(L)}\rangle_s \\
&\quad + \frac{1}{L}\sum_{\ell=1}^{L} 2\int_0^t \frac{1}{b_s^{(L)}}\Big(2b_s^{(L)}\Big(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \mathrm{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big) - 1\Big)^2\,\mathrm{d}K_s.
\end{aligned} \tag{4.52}$$

For the second term in (4.52), we use the explicit expressions $\mathrm{em}(\mathbf{x}) = \frac{1}{L}\sum_{\ell=1}^{L}\mathbf{x}^{\ell}$,

$\operatorname{evar}(\mathbf{x}) = \frac{1}{L}\sum_{\ell=1}^{L}(\mathbf{x}^\ell - \operatorname{em}(\mathbf{x}))^2$ and the property $\operatorname{evar}(\mathbf{x}) \leq \operatorname{em}(\mathbf{x}^2)$ all from (4.2) and (4.3) to compute

$$\begin{aligned}
&\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \frac{1}{b_s^{(L)}}\Big(2b_s^{(L)}\Big(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \operatorname{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big) - 1\Big)^2 \,\mathrm{d}K_s \\
&= \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t 4b_s^{(L)}\Big(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \operatorname{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big)^2 \,\mathrm{d}K_s \\
&\quad - \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t 4\Big(\mathbf{G}_s^{\ell,(L)}(\boldsymbol{\vartheta}) - \operatorname{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big)\Big) \,\mathrm{d}K_s + \int_0^t \frac{1}{b_s^{(L)}} \,\mathrm{d}K_s \\
&= 4\int_0^t b_s^{(L)} \operatorname{evar}\big(\mathbf{G}_s(\boldsymbol{\vartheta})\big) \,\mathrm{d}K_s + \int_0^t \frac{1}{b_s^{(L)}} \,\mathrm{d}K_s \\
&\leq 4\int_0^t b_s^{(L)} \operatorname{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})^2\big) \,\mathrm{d}K_s + \int_0^t \frac{1}{b_s^{(L)}} \,\mathrm{d}K_s. \qquad (4.53)
\end{aligned}$$

Inserting (4.53) into (4.52), we get

$$\begin{aligned}
\frac{1}{L}\sum_{\ell=1}^{L} R_t^{\ell,(L)}(\boldsymbol{\vartheta}) &\leq \frac{1}{L}\sum_{\ell=1}^{L} 2\int_0^t b_s^{(L)}(\boldsymbol{\vartheta}_s^\ell)^2 \,\mathrm{d}\langle \mathbf{M}^\ell\rangle_s + 4\int_0^t b_s^{(L)} \operatorname{em}\big(\mathbf{G}_s(\boldsymbol{\vartheta})^2\big) \,\mathrm{d}K_s \\
&\quad + \int_0^t \frac{1}{b_s^{(L)}} \,\mathrm{d}K_s \\
&=: F_t^1 + F_t^2 + F_t^3. \qquad (4.54)
\end{aligned}$$

Due to $\sup_{t\in[0,T]}|\mathbf{G}_t(\boldsymbol{\vartheta})| \in L^2$ by (4.50) and because K and $b^{(L)}$ are bounded, it is obvious that $\sup_{t\in[0,T]}|F_t^2| \in L^1$. The terms $\sup_{t\in[0,T]}|F_t^1|$ and $\sup_{t\in[0,T]}|F_t^3|$ are in L^1 because both terms $\sup_{t\in[0,T]}|b_t^{(L)}|$ and $\sup_{t\in[0,T]}|\frac{1}{b_t^{(L)}}|$ are bounded as pointed out in Lemma 4.7, K_T is bounded (even deterministic) by Assumption 4.11 and finally $\boldsymbol{\vartheta} \in L^2(\mathbf{M}^{(L)})$ by Lemma 3.17. Therefore combining (4.54) and (4.51) gives

$$\sup_{t\in[0,T]}|\widetilde{M}_t^{(L)}(\boldsymbol{\vartheta})| \leq 2\sup_{t\in[0,T]}|\widetilde{V}_t^{(L)}(\boldsymbol{\vartheta})| + \sup_{t\in[0,T]}|F_t^1| + \sup_{t\in[0,T]}|F_t^2| + \sup_{t\in[0,T]}|F_t^3| \in L^1.$$

This shows that $\widetilde{M}^{(L)}(\boldsymbol{\vartheta})$ is a martingale as desired. As we pointed out below (4.51) that $\widetilde{A}^{(L)}(\boldsymbol{\vartheta})$ is decreasing, $\widetilde{V}^{(L)}(\boldsymbol{\vartheta}) = \widetilde{V}_0^{(L)}(\boldsymbol{\vartheta}) + \widetilde{M}^{(L)}(\boldsymbol{\vartheta}) + \widetilde{A}^{(L)}(\boldsymbol{\vartheta})$ is a supermartingale. Moreover, (4.44) implies $\widetilde{A}^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)}) = -\frac{1}{L}\sum_{\ell=1}^{L} R_t^{\ell,(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)}) = 0$ and hence $\widetilde{V}^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)}) = \widetilde{V}_0^{(L)} + \widetilde{M}(\widehat{\boldsymbol{\vartheta}}^{(L)})$ is a martingale. In view of Lemma 2.4, this completes the proof. □

5 Taking limits and verification

Having obtained from Theorem 4.14 a complete description of a solution $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for the auxiliary problem (2.9), we aim to construct an optimal strategy for the original MVPS problem with the help of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ and study the convergence behaviour of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ as $L \to \infty$. Note that both the verification and convergence require an analysis of various limits as $L \to \infty$.

Let $L \in \mathbb{N} \cup \{\infty\}$ with $L \geq 2$. A basic relation is that for any $\ell = 1, \dots, L$,

$$\Omega = \pi_{\ell,L}(\Omega^{(L)}), \quad \mathcal{F} = \{E : \pi_{\ell,L}^{-1}(E) \in \mathcal{F}^{(L)}\}, \quad P = \mathbf{P}^{(L)} \circ \pi_{\ell,L}^{-1}, \tag{5.1}$$

where $\pi_{\ell,L} : \Omega^{(L)} \to \Omega$ is the canonical projection of $\Omega^{(L)}$ onto its ℓ-th coordinate. Thanks to Lemma 3.13, 1), the process $\mathbf{S}^{\ell,(L)}$ has the same special semimartingale decomposition with respect to the filtrations $\mathbb{F}^{\ell,(L)}$ and $\mathbb{G}^{(L)}$. Recall from (3.11) and (3.12) that we have $\mathcal{F}_t^{\ell,(L)} = \{E : \pi_{\ell,L}^{-1}(E) \in \mathcal{F}_t\}$ and $\mathbf{S}^{\ell,(L)}(\omega^{(L)}) = S^{\ell,\otimes L}(\omega^{(L)}) = S(\omega_\ell)$, respectively. By the uniqueness of $\mathbf{X}^\ell$ for $\mathbf{X} \in \{\mathbf{M}^{(L)}, \mathbf{A}^{(L)}, \langle \mathbf{M}^{(L)} \rangle\}$, we then also obtain

$$\mathbf{Y}^\ell(\omega^{(L)}) = Y^{\ell,\otimes L}(\omega^{(L)}) = Y(\omega_\ell)$$

for $\mathbf{Y} \in \{\mathbf{M}^{(L)}, \mathbf{A}^{(L)}, \langle \mathbf{M}^{(L)} \rangle\}$ and $Y \in \{M, A, \langle M \rangle\}$. For the quantities describing (SC) in (3.20)–(3.23) in dimension 1, we recall from the definition (3.3) and (4.28), (4.29) that we only require the existence of a predictable RCLL process $\boldsymbol{\lambda}^\ell$ such that $\mathbf{A}^{\ell,(L)} = \int \boldsymbol{\lambda}^{\ell,(L)} \,\mathrm{d}\langle \mathbf{M}^{\ell,(L)} \rangle$ and $\mathbf{K}_T^{\ell,(L)} = \int_0^T (\boldsymbol{\lambda}_s^{\ell,(L)})^2 \,\mathrm{d}\langle \mathbf{M}^{\ell,(L)} \rangle_s$ is $\mathbf{P}^{(L)}$-a.s. finite. Due to the identity in the above display, we can and do choose $\boldsymbol{\lambda}^\ell = \lambda^{\ell,\otimes L}$ for $\ell = 1, \dots, L$ to obtain

$$\begin{aligned}
\mathbf{A}_t^{\ell,(L)} &= A_t^{\ell,\otimes L} = \int_0^t \lambda_s^{\ell,\otimes L} \,\mathrm{d}\langle M^{\ell,\otimes L} \rangle_s = \int_0^t \boldsymbol{\lambda}^{\ell,(L)} \,\mathrm{d}\langle \mathbf{M}^{\ell,(L)} \rangle, \\
\mathbf{K}_t^{\ell,(L)} &= \int_0^t (\boldsymbol{\lambda}_s^{\ell,(L)})^2 \,\mathrm{d}\langle \mathbf{M}^{\ell,(L)} \rangle_s = \int_0^t (\lambda_s^{\ell,\otimes L})^2 \,\mathrm{d}\langle M^{\ell,\otimes L} \rangle_s < \infty,
\end{aligned}$$

both $\mathbf{P}^{(L)}$-a.s. for all $t \in [0, T]$. This yields ultimately

$$\mathbf{Y}^\ell(\omega^{(L)}) = Y^{\ell,\otimes L}(\omega^{(L)}) = Y(\omega_\ell) \tag{5.2}$$

for $\mathbf{Y} \in \{\mathbf{S}^{(L)}, \mathbf{M}^{(L)}, \mathbf{A}^{(L)}, \langle \mathbf{M}^{(L)} \rangle, \boldsymbol{\lambda}^{(L)}, \mathbf{K}^{(L)}\}$ and $Y \in \{S, M, A, \langle M \rangle, \lambda, K\}$. We emphasize that both (5.1) and (5.2) hold with $L = \infty$ as well. Hence we can view $(\Omega, \mathcal{F}, P)$ as a coordinate of both finite-L and infinite product spaces

$(\Omega^{(L)}, \mathcal{F}^{(L)}, \mathbf{P}^{(L)})$.

5.1 Verification

In this subsection, we aim to construct an optimal strategy $\widehat{\theta}$ for the MVPS problem (2.1) in continuous time. We begin by recalling from Section I.5.1 some formal connections between the original and extended markets.

Let us write as in (2.10) and (2.11)

$$J_T^{\mathrm{mv}}(\theta) = G_T(\theta) - \xi\big(G_T(\theta) - E[G_T(\theta)]\big)^2, \quad \theta \in \Theta_{\mathrm{S}}, \tag{5.3}$$

$$J_T^{(L)}(\boldsymbol{\vartheta}^{(L)}) = \mathrm{em}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big) - \xi \mathrm{evar}\big(\mathbf{G}_T(\boldsymbol{\vartheta}^{(L)})\big), \quad \boldsymbol{\vartheta}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)}, \tag{5.4}$$

and recall the result from Proposition I.5.3. As that result is abstract, it holds equally well in discrete and continuous time.

Proposition 5.1. *If there exists $\widehat{\theta} \in \Theta_{\mathrm{S}}$ such that $\mathbf{E}^{(L)}[J_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] \to E[J_T^{\mathrm{mv}}(\widehat{\theta})]$ as $L \to \infty$, then $G_T(\widehat{\theta})$ is the optimal final gain for the MVPS problem* (2.12) *and $\widehat{\theta}$ is an optimal strategy for* (2.12).

In view of Proposition 5.1, we want to construct a strategy $\widehat{\theta} \in \Theta_{\mathrm{S}}$ satisfying a certain limit property. As in discrete time, this candidate $\widehat{\theta}$ can be found by formally sending $L \to \infty$ in the solution of the auxiliary problem in the L-extended market. We recall from (4.49) the explicit formula giving $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for $\ell = 1, \dots, L$ as

$$\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} = -\frac{2b_t^{(L)}(\mathbf{G}_t^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))) - 1}{2b_t^{(L)}(1 - L^{-1})} \boldsymbol{\lambda}_t^{\ell,(L)}, \quad t \in [0,T], \tag{5.5}$$

which is well defined due to the existence of a unique solution to the system (4.33) of SDEs explicitly given by

$$\mathrm{d}\mathbf{X}_t^{\ell,(L)} = -\frac{2b_t^{(L)}(\mathbf{X}_t^{\ell,(L)} - \mathrm{em}(\mathbf{X}_t^{(L)})) - 1}{2b_t^{(L)}(1 - L^{-1})} \boldsymbol{\lambda}_t^{\ell,(L)} \,\mathrm{d}\mathbf{S}_t^{\ell,(L)}, \quad \mathbf{X}_0^{\ell,(L)} = 0, \tag{5.6}$$

for $\ell = 1, \dots, L$. In order to define a candidate $\widehat{\boldsymbol{\vartheta}}^{(\infty)}$, we need to identify a limit of the above system as $L \to \infty$. In view of the explicit expression (4.26) for $b_t^{(L)}$, the convergence $b_t^{(L)} = \xi \exp(\frac{K_t - K_T}{1 - L^{-1}}) \to \xi \exp(K_t - K_T) =: b_t^{(\infty)}$ is clear. Because $\boldsymbol{\lambda}^{\ell,(L)}(\omega^{(L)}) = \lambda^{\ell,\otimes L}(\omega^{(L)}) = \lambda(\omega_\ell)$ and $\mathbf{S}^{\ell,(L)}(\omega^{(L)}) = S^{\ell,\otimes L}(\omega^{(L)}) = S(\omega_\ell)$, their dependence on L is artificial and their limits should still be $\lambda(\omega_\ell)$ and $S(\omega_\ell)$, respectively. Now we expect by a law of large numbers effect that the empirical mean $\mathrm{em}(\mathbf{X}_t^{(L)})$ converges to an expectation, say e_t, as $L \to \infty$. Suppose this is true and denote the formal limit of $\mathbf{X}^{\ell,(L)}$ by $\mathbf{X}^{\ell,(\infty)}$ for $\ell \in \mathbb{N}$. Because the dependence on other coordinates $\mathbf{X}^{m,(L)}$ for $m \neq \ell$ via the empirical average disappears in the limit, we expect that $\mathbf{X}^{\ell,(\infty)}$ depends only on ω_ℓ. The symmetry of $\mathbf{X}^{(L)}$ among $\ell = 1, \dots, L$ from (5.6) also suggests that the expectation e_t does not depend on ℓ and is equal to the expectation of $\mathbf{X}_t^{\ell,(\infty)}$ for any ℓ. Summarising the above analysis and using that $\mathbf{X}^{\ell,(\infty)}$ depends only on ω_ℓ motivates us to replace ω_ℓ by ω and consider instead of (5.6) on the original space $(\Omega, \mathcal{F}, P)$ the affine McKean–Vlasov SDE

$$\mathrm{d}\widehat{X}_t = -\frac{2b_t^{(\infty)}(\widehat{X}_t - E[\widehat{X}_t]) - 1}{2b_t^{(\infty)}}\lambda_t \,\mathrm{d}S_t, \quad \widehat{X}_0 = 0. \tag{5.7}$$

Suppose that (5.7) can be solved uniquely. We then define, in analogy to (5.5),

$$\widehat{\theta}_t := -\frac{2b_t^{(\infty)}(\widehat{X}_t - E[\widehat{X}_t]) - 1}{2b_t^{(\infty)}}\lambda_t = -\lambda_t\left(\widehat{X}_t - E[\widehat{X}_t] - \frac{1}{2b_t^{(\infty)}}\right), \quad t \in [0,T], \tag{5.8}$$

so that we get

$$G(\widehat{\theta}) = \int \widehat{\theta}\,\mathrm{d}S = \int -\frac{2b^{(\infty)}(\widehat{X} - E[\widehat{X}]) - 1}{2b^{(\infty)}}\lambda \,\mathrm{d}S = \widehat{X} \tag{5.9}$$

due to (5.7) and $G_0(\widehat{\theta}) = 0 = \widehat{X}_0$. Plugging this back into (5.10) shows that $\widehat{\theta}$ satisfies

$$\widehat{\theta}_t = -\frac{2b_t^{(\infty)}(G_t(\widehat{\theta}) - E[G_t(\widehat{\theta})]) - 1}{2b_t^{(\infty)}}\lambda_t, \quad t \in [0,T], \tag{5.10}$$

and gives a well-defined process which is $\mathbb{F}$-predictable due to the continuity of $G(\widehat{\theta})$. So to argue that $\widehat{\theta} \in \Theta_{\mathrm{S}}$, it is enough to prove $\widehat{\theta} \in L^2(M)$ by the identity $\Theta_{\mathrm{S}} = L^2(M)$ from (3.9). We now summarise what precisely needs to be done.

Recipe 5.2. 1) Establish the existence and uniqueness of a solution to the SDE (5.7), as well as the integrability property of the solution. Moreover, show that $\widehat{\theta} \in L^2(M)$.

2) Show that $\mathbf{E}^{(L)}[J^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] \to E[J_T^{\mathrm{mv}}(\widehat{\theta})]$ as $L \to \infty$.

Let us implement this programme now. Suppose Assumption 4.2 is satisfied. This means that S is a real-valued continuous semimartingale satisfying the structure condition (SC). We begin by recalling from (3.2)–(3.5) that

$$A_t = \int_0^t \lambda_s \,\mathrm{d}\langle M\rangle_s, \quad K_t = \int_0^t \lambda_s \,\mathrm{d}A_s = \int_0^t \lambda_s^2 \,\mathrm{d}\langle M\rangle_s, \quad t \in [0,T], \tag{5.11}$$

when $d = 1$, and from (4.26) and (4.27) that for $L \in \mathbb{N} \setminus \{1\}$

$$\begin{aligned} b_t^{(L)} &= \xi \frac{\mathcal{E}(\frac{K}{1-L^{-1}})_t}{\mathcal{E}(\frac{K}{1-L^{-1}})_T} = \xi \exp\left(\frac{K_t - K_T}{1 - L^{-1}}\right), \quad t \in [0,T], \\ c_t^{(L)} &= -\frac{1}{4\xi}\left(1 - \frac{\xi}{b_t^{(L)}}\right) = -\frac{1}{4\xi}\left(1 - \exp\left(-\frac{K_t - K_T}{1 - L^{-1}}\right)\right), \quad t \in [0,T]. \end{aligned}$$

Sending $L \to \infty$ in the above two identities yields the candidate limits

$$b_t^{(\infty)} := \xi \exp(K_t - K_T), \quad t \in [0,T], \tag{5.12}$$

$$c_t^{(\infty)} := -\frac{1}{4\xi}\left(1 - \frac{\xi}{b_t^{(\infty)}}\right), \quad t \in [0,T]. \tag{5.13}$$

It is easy to verify that $(b^{(\infty)}, c^{(\infty)})$ is a solution to the system

$$\frac{\mathrm{d}b_t^{(\infty)}}{\mathrm{d}K_t} - b_t^{(\infty)} = 0, \quad b_T^{(\infty)} = \xi, \tag{5.14}$$

$$\frac{\mathrm{d}c_t^{(\infty)}}{\mathrm{d}K_t} + \frac{1}{4b_t^{(\infty)}} = 0, \quad c_T^{(\infty)} = 0, \tag{5.15}$$

and satisfies

$$\sup_{t\in[0,T]} |b_t^{(L)} - b_t^{(\infty)}| + \sup_{t\in[0,T]} |c_t^{(L)} - c_t^{(\infty)}| + \sup_{t\in[0,T]} \left|\frac{1}{b_t^{(L)}} - \frac{1}{b_t^{(\infty)}}\right| \longrightarrow 0 \tag{5.16}$$

as $L \to \infty$ under the assumption that the MVT process K is bounded.

Lemma 5.3. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied. Then there exists a unique strong solution* $\widehat{X}$ *to the SDE* (5.7) *in the class of* $\mathbb{F}$*-adapted continuous processes. It satisfies* $E[\sup_{t\in[0,T]} \widehat{X}_s^2] < \infty$ *and has the expectation*

$$E[\widehat{X}_t] = \int_0^t \frac{1}{2b_s^{(\infty)}} \,\mathrm{d}K_s, \quad t \in [0,T]. \tag{5.17}$$

Moreover, the process $\widehat{\theta}$ *defined in* (5.8) *is in* Θ_{S} *and satisfies* $G(\widehat{\theta}) = \widehat{X}$ *and* (5.10).

Proof. 1) If we have a solution $\widehat{X}$ to (5.7), we can *formally* take expectations in (5.7) and use (5.11) and the non-randomness of K from Assumption 4.11 to obtain

$$E[\widehat{X}_t] = \int_0^t -E\left[\frac{2b_s^{(\infty)}(\widehat{X}_s - E[\widehat{X}_s]) - 1}{2b_s^{(\infty)}}\right] \mathrm{d}K_s = \int_0^t \frac{1}{2b_s^{(\infty)}} \mathrm{d}K_s.$$

To make this rigorous, let us define the process $m = (m_t)_{t\in[0,T]}$ by

$$m_t = \int_0^t \frac{1}{2b_s^{(\infty)}} \mathrm{d}K_s \geq 0, \quad t \in [0,T], \tag{5.18}$$

and consider the standard SDE

$$dX_t = -\frac{2b_t^{(\infty)}(X_t - m_t) - 1}{2b_t^{(\infty)}}\lambda_t \,\mathrm{d}S_t, \quad X_0 = 0. \tag{5.19}$$

Because the map

$$f_t(x) := -\frac{2b_t^{(\infty)}(x - m_t) - 1}{2b_t^{(\infty)}}\lambda_t \tag{5.20}$$

is Lipschitz in x with a time-dependent $\mathcal{F}_t$-measurable coefficient λ_t, the SDE

$$dX_t = f_t(X_t)\,\mathrm{d}S_t, \quad X_0 = 0, \tag{5.21}$$

has a unique strong solution X (see Protter [55, Theorem V.6]) in the space of $\mathbb{F}$-adapted continuous processes. Because (5.21) is equivalent to (5.19), X is of course a solution to (5.19). We now argue that $E[\sup_{t\in[0,T]}|X_t|^2] < \infty$ by collecting some similar estimates as in the proof of Lemma 4.9. Writing $g_t(x) = -\frac{2b_t^{(\infty)}(x-m_t)-1}{2b_t^{(\infty)}}$ and $g_t^0 := g_t(0) = m_t + \frac{1}{2b_t^{(\infty)}} \geq 0$, we see that g_t is affine in x with Lipschitz constant 1, which yields

$$|g_t(x)| \leq g_t(0) + |g_t(x) - g_t(0)| \leq g_t^0 + |x|. \tag{5.22}$$

Using $(x+y)^2 \leq 2x^2 + 2y^2$ in (5.21), then $f_t(x) = g_t(x)\lambda_t$ and $\mathrm{d}K_t = \lambda_t\,\mathrm{d}A_t$ by

(5.11) and finally the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}X_s^2 &\le 2\Big(\int_0^s f_r(X_r)\,\mathrm{d}A_r\Big)^2 + 2\Big(\int_0^s f_r(X_r)\,\mathrm{d}M_r\Big)^2\\ &= 2\Big(\int_0^s g_r(X_r)\,\mathrm{d}K_r\Big)^2 + 2\Big(\int_0^s f_r(X_r)\,\mathrm{d}M_r\Big)^2\\ &\le 2K_s\int_0^s \big(g_r(X_r)\big)^2\,\mathrm{d}K_r + 2\Big(\int_0^s f_r(X_r)\,\mathrm{d}M_r\Big)^2. \qquad (5.23)\end{aligned}$$

We use $\sup_{s\in[0,t]} |K_s \int_0^s (g_r(X_r))^2\,\mathrm{d}K_r| \le K_t \int_0^t (g_s(X_s))^2\,\mathrm{d}K_s$ with $\lambda^2\,\mathrm{d}\langle M\rangle = \mathrm{d}K$ and the BDG inequality in (5.23) and then (5.22) to get

$$\begin{aligned}E\Big[\sup_{s\in[0,t]} X_s^2\Big] &\le E\Big[2K_t\int_0^t \big(g_s(X_s)\big)^2\,\mathrm{d}K_s + 8\int_0^t \big(g_s(X_s)\big)^2\,\mathrm{d}K_s\Big]\\ &\le (2K_t+8)E\Big[\int_0^t 2\big((g_s^0)^2 + X_s^2\big)\,\mathrm{d}K_s\Big]\\ &\le (4K_t+16)\int_0^t \Big((g_s^0)^2 + E\Big[\sup_{r\in[0,s]} X_r^2\Big]\Big)\,\mathrm{d}K_s.\end{aligned}$$

Note that the last step uses that K is deterministic so that we can apply Fubini's theorem. In view of Gronwall's lemma, the above bound yields that

$$E\Big[\sup_{t\in[0,T]} X_t^2\Big] \le (4K_T+16)\Big(\int_0^T (g_s^0)^2\,\mathrm{d}K_s\Big)\exp\big((4K_T+16)K_T\big). \qquad (5.24)$$

Because $\sup_{t\in[0,T]} \frac{1}{2b_t^{(\infty)}} < \infty$ by the explicit expression in (5.12), we use the definitions of $g_t^0 = g_t(0) = m_t + \frac{1}{2b_t^{(\infty)}} \ge 0$ and $m_t \ge 0$ and Assumption 4.11 to obtain

$$\sup_{t\in[0,T]} g_t^0 = \sup_{t\in[0,T]}\Big(m_t + \frac{1}{2b_t^{(\infty)}}\Big) = \sup_{t\in[0,T]}\Big(\int_0^t \frac{1}{2b_s^{(\infty)}}\,\mathrm{d}K_s + \frac{1}{2b_t^{(\infty)}}\Big) < \infty. \quad (5.25)$$

Inserting (5.25) into (5.24) gives $E[\sup_{t\in[0,T]} X_t^2] < \infty$ as desired.

2) Next we show that the solution X of (5.19) indeed solves (5.7). For this, in view of the uniqueness of the solution to (5.7), it suffices to show that $m_t = E[X_t]$ for $t\in[0,T]$. Consider $D_t := X_t - m_t$; then we argue that $E[D_t] = 0$ for all $t\in[0,T]$. Applying Itô's lemma to D, then using formulas (5.19) and (5.18) for $\mathrm{d}X$ and $\mathrm{d}m$, respectively, and finally using $\lambda\,\mathrm{d}A = \mathrm{d}K$ and the definition (5.20)

of f_t yields

$$\begin{aligned}
\mathrm{d}D_t &= -\frac{2b_t^{(\infty)}(X_t - m_t) - 1}{2b_t^{(\infty)}}\lambda_t \,\mathrm{d}S_t - \frac{1}{2b_t^{(\infty)}}\,\mathrm{d}K_t \\
&= -\frac{2b_t^{(\infty)}(X_t - m_t) - 1}{2b_t^{(\infty)}}\lambda_t \,\mathrm{d}M_t - \frac{2b_t^{(\infty)}(X_t - m_t) - 1}{2b_t^{(\infty)}}\,\mathrm{d}K_t - \frac{1}{2b_t^{(\infty)}}\,\mathrm{d}K_t \\
&= -f_t(X_t)\,\mathrm{d}M_t - D_t\,\mathrm{d}K_t. \qquad (5.26)
\end{aligned}$$

We again use $f_t(x) = g_t(x)\lambda_t$, $\mathrm{d}K = \lambda^2\,\mathrm{d}\langle M\rangle$ by (5.11) and that K is deterministic by Assumption 4.11, then $(x+y)^2 \le 2x^2 + 2y^2$ with (5.22), and finally (5.25) and $E[\sup_{t\in[0,T]} X_t^2] < \infty$ to obtain

$$\begin{aligned}
E\Big[\int_0^T \big(f_t(X_t)\big)^2\,\mathrm{d}\langle M\rangle_t\Big] &= \int_0^T E\big[\big(g_t(X_t)\big)^2\big]\,\mathrm{d}K_t \\
&\le 2K_T\Big(\sup_{t\in[0,T]}(g_t^0)^2 + E\Big[\sup_{t\in[0,T]} X_t^2\Big]\Big) \\
&< \infty.
\end{aligned}$$

This shows that $(f_t(X_t))_{t\in[0,T]} \in L^2(M)$ and hence the process $\int f(X)\,\mathrm{d}M$ is a martingale by Lemma 3.5, 1). So taking expectations in (5.26) and using Fubini's theorem with the non-randomness of K from Assumption 4.11 yields

$$E[D_t] = E\Big[\int_0^t (-D_s)\,\mathrm{d}K_s\Big] = -\int_0^t E[D_s]\,\mathrm{d}K_s, \quad t \in [0,T]. \qquad (5.27)$$

Note that using $D_t = X_t - m_t$ and $\sup_{t\in[0,T]}(m_t)^2 < \infty$ and $\sup_{t\in[0,T]} \frac{1}{2b_t^{(\infty)}} < \infty$ from (5.18) and (5.12), respectively, we obtain

$$\sup_{t\in[0,T]} |D_t| \le \sup_{t\in[0,T]} |X_t| + \sup_{t\in[0,T]} m_t.$$

Then the fact that $\sup_{t\in[0,T]} |X_t| \in L^2$ obtained in part 1) and $L^2 \subseteq L^1$ implies that $\sup_{t\in[0,T]} |D_t|$ is in L^1 and

$$E\Big[\int_0^T |D_t|\,\mathrm{d}K_t\Big] \le E\Big[\sup_{t\in[0,T]} |D_t|\Big] K_T < \infty.$$

In view of $\mathrm{d}P \otimes \mathrm{d}K$ as a measure on $\Omega \times [0,T]$, this ensures that Fubini's theorem can be applied. Because $E[D_0] = X_0 - m_0 = 0$, the only solution to the integral equation (5.27) is the constant process 0 and hence $E[D_t] = 0$ for all $t \in [0,T]$.

3) By parts 1) and 2), the equation (5.7) has a unique strong solution $\widehat{X}$ which is given by the solution X to (5.19). Thus we can define $\widehat{\theta}$ by (5.8). Because $\widehat{\theta} = f(X)$, the fact that $\widehat{\theta} \in L^2(M)$ is just a restatement of $f(X) \in L^2(M)$ which has been obtained in part 2). In view of $\Theta_{\mathrm{S}} = L^2(M)$ by Lemma 3.6 and Assumption 4.11, we obtain $\widehat{\theta} \in \Theta_{\mathrm{S}}$. As pointed out in (5.9), we also have $G(\widehat{\theta}) = \widehat{X}$. Using this in (5.8) yields finally (5.10). □

We shall see later that $\widehat{\theta}$ is a solution to the MVPS problem (2.1). We also have from Lemma 5.3 that $G(\widehat{\theta}) = \widehat{X}$ and therefore

$$E[G_t(\widehat{\theta})] = E[\widehat{X}_t] = \int_0^t \frac{1}{2b_s^{(\infty)}} \,\mathrm{d}K_s, \quad t \in [0,T].$$

It is interesting to note that this is always nonnegative and actually like K null at 0 and increasing, because $b^{(\infty)} \geq 0$. In other words, the expected gains from the optimal strategy $\widehat{\theta}$ increase over time.

Lemma 5.3 summarises the implementation of Recipe 5.2, 1). We proceed to the next step. Define a process $(\widehat{V}_t^{(\infty)})_{t\in[0,T]}$ by

$$\widehat{V}_t^{(\infty)} = E[G_t(\widehat{\theta})] - b_t^{(\infty)} \mathrm{Var}[G_t(\widehat{\theta})] + c_t^{(\infty)}, \quad t \in [0,T]. \tag{5.28}$$

Note the analogy of (5.28) and Lemma 5.5 below to (I.5.37) and Lemma I.5.6 in finite discrete time. We prove below that the process $\widehat{V}^{(\infty)}$ is constant. Before doing that, let us recall a sufficient condition under which a stochastic integral with respect to a local martingale is a martingale.

Lemma 5.4. *If θ, ψ are predictable processes such that $\int \theta\psi \,\mathrm{d}M$ is a continuous local martingale and we have $\sup_{t\in[0,T]} |\theta_t| \in L^2$ and $\int_0^T (\psi_t)^2 \,\mathrm{d}\langle M\rangle_t \in L^1$, then the process $\int \theta\psi \,\mathrm{d}M$ is a continuous martingale.*

Proof. Since $\int \theta\psi \,\mathrm{d}M$ is a continuous local martingale, we apply the BDG inequality, Itô's isometry and Cauchy–Schwarz to obtain

$$\begin{aligned}
E\Big[\sup_{t\in[0,T]}\Big|\int_0^t \theta_s\psi_s\,\mathrm{d}M_s\Big|\Big] &\le CE\Big[\Big\langle\int\theta\psi\,\mathrm{d}M\Big\rangle_T^{\frac{1}{2}}\Big]\\
&= CE\Big[\Big(\int_0^T(\theta_t\psi_t)^2\,\mathrm{d}\langle M\rangle_t\Big)^{\frac{1}{2}}\Big]\\
&\le CE\Big[\Big(\sup_{t\in[0,T]}|\theta_t|\Big)\Big(\int_0^T(\psi_t)^2\,\mathrm{d}\langle M\rangle_t\Big)^{\frac{1}{2}}\Big]\\
&\le CE\Big[\sup_{t\in[0,T]}|\theta_t|^2\Big]^{\frac{1}{2}}E\Big[\int_0^T(\psi_t)^2\,\mathrm{d}\langle M\rangle_t\Big]^{\frac{1}{2}}\\
&< \infty.
\end{aligned}$$

This yields the assertion. □

Lemma 5.5. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied. Then* $\widehat{V}^{(\infty)}$ *is constant.*

Proof. First, we know by the last assertion in Lemma 5.3 that $G(\widehat{\theta}) = \widehat{X}$, where $\widehat{X}$ satisfies (5.7). To ease notation, we work with $\widehat{X}$. Observe from (5.17) that the expectation process $E[\widehat{X}]$ satisfies

$$E[\widehat{X}_t] = \int_0^t \frac{1}{2b_s^{(\infty)}}\,\mathrm{d}K_s, \quad t\in[0,T]. \tag{5.29}$$

Also by Itô's lemma, we get from $\widehat{X} = G(\widehat{\theta})$ and $\mathrm{d}A = \lambda\,\mathrm{d}\langle M\rangle$ in (5.11) that

$$\begin{aligned}
\widehat{X}_t^2 &= \int_0^t 2\widehat{X}_s\,\mathrm{d}\widehat{X}_s + [\widehat{X}]_t\\
&= \int_0^t 2\widehat{X}_s\widehat{\theta}_s\,\mathrm{d}A_s + \int_0^t 2\widehat{X}_s\widehat{\theta}_s\,\mathrm{d}M_s + \int_0^t \widehat{\theta}_s^2\,\mathrm{d}\langle M\rangle_s\\
&= \int_0^t 2\widehat{X}_s\widehat{\theta}_s\lambda_s\,\mathrm{d}\langle M\rangle_s + \int_0^t 2\widehat{X}_s\widehat{\theta}_s\,\mathrm{d}M_s + \int_0^t \widehat{\theta}_s^2\,\mathrm{d}\langle M\rangle_s.
\end{aligned}\tag{5.30}$$

We want to take expectations in (5.30) to eliminate the (local) martingale term. So we need to show that the process $(\int_0^t 2\widehat{X}_s\widehat{\theta}_s\,\mathrm{d}M_s)_{t\in[0,T]}$ is a true martingale. This immediately follows from $\sup_{t\in[0,T]}|\widehat{X}_t| \in L^2$ and $\widehat{\theta}\in L^2(M)$ together with Lemma 5.4. Therefore, using (5.30), the explicit formula (5.8) for $\widehat{\theta}$ and Fubini's theorem with $\mathrm{d}K = \lambda^2 \,\mathrm{d}\langle M\rangle$ from (5.11), we get that

$$\begin{aligned}
E[\widehat{X}_t^2] &= E\bigg[\int_0^t \big(2\widehat{X}_s\widehat{\theta}_s\lambda_s + \widehat{\theta}_s^2\big)\,\mathrm{d}\langle M\rangle_s\bigg] \\
&= E\bigg[\int_0^t (\widehat{\theta}_s + \widehat{X}_s\lambda_s)^2\,\mathrm{d}\langle M\rangle_s\bigg] - E\bigg[\int_0^t (\widehat{X}_s\lambda_s)^2\,\mathrm{d}\langle M\rangle_s\bigg] \\
&= E\bigg[\int_0^t \bigg(E[\widehat{X}_s] + \frac{1}{2b_s^{(\infty)}}\bigg)^2 \lambda_s^2\,\mathrm{d}\langle M\rangle_s\bigg] - E\bigg[\int_0^t (\widehat{X}_s\lambda_s)^2\,\mathrm{d}\langle M\rangle_s\bigg] \\
&= \int_0^t \bigg(\frac{1}{2b_s^{(\infty)}} + E[\widehat{X}_s]\bigg)^2\,\mathrm{d}K_s - \int_0^t E[\widehat{X}_s^2]\,\mathrm{d}K_s. \qquad (5.31)
\end{aligned}$$

Here Fubini's theorem is applicable because the integrands are positive. Of course, from the explicit formula (5.29) for $E[\widehat{X}_t]$ and Itô's lemma, we have

$$(E[\widehat{X}_t])^2 = \int_0^t \frac{E[\widehat{X}_s]}{b_s^{(\infty)}}\,\mathrm{d}K_s. \qquad (5.32)$$

Combining (5.31) and (5.32), we get

$$\begin{aligned}
\mathrm{Var}[\widehat{X}_t] &= E[\widehat{X}_t^2] - (E[\widehat{X}_t])^2 \\
&= \int_0^t \bigg(\frac{1}{2b_s^{(\infty)}} + E[\widehat{X}_s]\bigg)^2\,\mathrm{d}K_s - \int_0^t E[\widehat{X}_s^2]\,\mathrm{d}K_s - \int_0^t \frac{E[\widehat{X}_s]}{b_s^{(\infty)}}\,\mathrm{d}K_s \\
&= \int_0^t \bigg(\frac{1}{(2b_s^{(\infty)})^2} - \mathrm{Var}[\widehat{X}_s]\bigg)\,\mathrm{d}K_s. \qquad (5.33)
\end{aligned}$$

Applying Itô's product rule to $b_t^{(\infty)}\mathrm{Var}[\widehat{X}_t]$ and using (5.33) and the differential equation (5.14) for $b^{(\infty)}$, we get

$$\begin{aligned}
b_t^{(\infty)}\mathrm{Var}[\widehat{X}_t] &= \int_0^t \mathrm{Var}[\widehat{X}_s]\,\mathrm{d}b_s^{(\infty)} + \int_0^t b_s^{(\infty)}\,\mathrm{d}\mathrm{Var}[\widehat{X}_s] \\
&= \int_0^t \mathrm{Var}[\widehat{X}_s]b_s^{(\infty)}\,\mathrm{d}K_s + \int_0^t b_s^{(\infty)}\bigg(\frac{1}{(2b_s^{(\infty)})^2} - \mathrm{Var}[\widehat{X}_s]\bigg)\,\mathrm{d}K_s \\
&= \int_0^t \frac{1}{4b_s^{(\infty)}}\,\mathrm{d}K_s. \qquad (5.34)
\end{aligned}$$

Therefore, we can use the explicit form (5.28) for $\widehat{V}^{(\infty)}$ and then combine the integral representations (5.29), (5.34) and (5.15) for $E[\widehat{X}]$, $b^{(\infty)}\mathrm{Var}[\widehat{X}]$ and $c^{(\infty)}$, respectively, to obtain

$$\begin{aligned}\widehat{V}_t^{(\infty)} &= E[\widehat{X}_t] - b_t^{(\infty)}\mathrm{Var}[\widehat{X}_t] + c_t^{(\infty)} \\ &= \int_0^t \frac{1}{2b_s^{(\infty)}}\,\mathrm{d}K_s - \int_0^t \frac{1}{4b_s^{(\infty)}}\,\mathrm{d}K_s + c_0^{(\infty)} - \int_0^t \frac{1}{4b_s^{(\infty)}}\,\mathrm{d}K_s \\ &= c_0^{(\infty)},\end{aligned}$$

which is constant as desired. □

Our first main result now gives an optimal strategy for the MVPS problem in explicit form.

Theorem 5.6. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied, meaning that the price process S is a real-valued continuous semimartingale satisfying the structure condition (SC) and the MVT process K is deterministic. Then the process $\widehat{\theta}$ defined in* (5.10) *is an optimal strategy for the MVPS problem* (2.1).

Proof. By the optimality of $\widehat{\boldsymbol{\vartheta}}^{(L)}$ shown in Theorem 4.14, the process $\widetilde{V}^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})$ is a martingale. Thus using the explicit form (4.45) for $\widetilde{V}^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})$, $c_0^{(L)} \to c_0^{(\infty)}$ from (5.16), the constancy of $\widehat{V}^{(\infty)}$ from Lemma 5.5 and finally the explicit form (5.28) for $\widehat{V}_T^{(\infty)}$, we have

$$\mathbf{E}^{(L)}[\widetilde{V}_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \widetilde{V}_0^{(L)} = c_0^{(L)} \longrightarrow c_0^{(\infty)} = \widehat{V}_0^{(\infty)} = \widehat{V}_T^{(\infty)} = E[J_T^{\mathrm{mv}}(\widehat{\theta})].$$

We then conclude the proof by Proposition 5.1. □

5.2 Convergence of strategies – preparation

From now on, we turn to the study of the convergence behaviour of the sequence $(\widehat{\boldsymbol{\vartheta}}^{(L)})_{L\in\mathbb{N}}$ of processes as $L \to \infty$. In this subsection, we do some preparation work by recalling from Section I.5.1 some formalities of lifting/embedding quantities from L-extended markets to a single infinite product space and collecting some basic properties of the lifted optimal strategy (still written as) $\widehat{\boldsymbol{\vartheta}}^{(L)}$ for the auxiliary problem (5.4).

Below we work only with the probability space $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, \mathbf{P}^{(\infty)})$ which supports infinite sequences $(\mathbb{F}^{\ell,(\infty)})_{\ell\in\mathbb{N}}$ and $(\mathbf{S}^{\ell,(\infty)})_{\ell\in\mathbb{N}}$ of filtrations and processes. The processes $\mathbf{S}^{\ell,(\infty)}$ are independent and all distributed according to S. For $L \in \mathbb{N}$, the filtration $\widetilde{\mathbb{G}}^{(L)}$ given in (2.2) changes to $\widetilde{\mathcal{G}}_t = \sigma(\cup_{\ell=1}^{L} \mathcal{F}_t^{\ell,(\infty)})$, and the filtration $\mathbb{G}^{(L)}$ as in (3.18) is now given by $\mathcal{G}_t^{(L)} = \widetilde{\mathcal{G}}_t^{(L)} \vee \mathcal{N}^{(\infty)}$, where $\mathcal{N}^{(\infty)}$ is the class of $\mathbf{P}^{(\infty)}$-null sets in $\widetilde{\mathcal{G}}_T^{(\infty)}$. The properties (5.1) and (5.2) for $L = \infty$ are

$$\Omega = \pi_{\ell,\infty}(\Omega^{(\infty)}), \quad \mathcal{F} = \{E : \pi_{\ell,\infty}^{-1}(E) \in \mathcal{F}^{(\infty)}\}, \quad P = \mathbf{P}^{(\infty)} \circ \pi_{\ell,(\infty)}^{-1}, \tag{5.35}$$

$$\mathbf{Y}^{\ell}(\omega^{(L)}) = Y^{\ell,\otimes\infty}(\omega^{(L)}) = Y(\omega_\ell), \tag{5.36}$$

for $\mathbf{Y} \in \{\mathbf{S}^{(\infty)}, \mathbf{M}^{(\infty)}, \mathbf{A}^{(\infty)}, \langle \mathbf{M}^{(\infty)} \rangle, \boldsymbol{\lambda}^{(\infty)}, \mathbf{K}^{(\infty)}\}$ and $Y = \{S, M, A, \langle M \rangle, \lambda, K\}$. Recall from (2.14) that we have defined $X^{\otimes L}$ for both $L \in \mathbb{N}$ and $L = \infty$ by

$$X^{\ell,\otimes L} = X \circ \pi_{\ell,L}, \quad \ell = 1, \dots, L. \tag{5.37}$$

In the rest of this chapter, we drop the superscript $^{(\infty)}$ and write

$$\mathbb{F}^{\ell,(\infty)} = \mathbb{F}^{\ell}, \quad \ell \in \mathbb{N}, \tag{5.38}$$

$$(\mathbf{S}^{(\infty)}, \mathbf{M}^{(\infty)}, \mathbf{A}^{(\infty)}, \langle \mathbf{M}^{(\infty)} \rangle, \boldsymbol{\lambda}^{(\infty)}, \mathbf{K}^{(\infty)}) = (\mathbf{S}, \mathbf{M}, \mathbf{A}, \langle \mathbf{M} \rangle, \boldsymbol{\lambda}, \mathbf{K}). \tag{5.39}$$

Following the above convention, we use a slight modification of the arguments in Lemma 3.13, 1) and the $\mathbb{F}^{\ell}$-adaptedness of $\mathbf{S}^{\ell}$ to obtain that the canonical decompositions of $\mathbf{S}^{\ell}$ with respect to $\mathbb{G}^{(L)}$ and $\mathbb{F}^{\ell}$ are the same as long as $\ell \leq L$. We always refer to $\mathbf{S}^{\ell} = S_0 + \mathbf{M}^{\ell} + \mathbf{A}^{\ell}$ as the canonical decomposition of $\mathbf{S}^{\ell}$ with respect to any $\mathbb{G}^{(L)}$ such that $\ell \leq L$ or equivalently with respect to $\mathbb{F}^{\ell}$.

For a quantity $\mathbf{X}^{(L)}$ and $L \in \mathbb{N}$, the superscript $^{(L)}$ indicates that it is originally defined on $(\Omega^{(L)}, \mathcal{F}^{(L)}), \mathbf{P}^{(L)})$ and is identified with a quantity still written as $\mathbf{X}^{(L)}$ on $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, \mathbf{P}^{(\infty)})$ via

$$\mathbf{X}^{(L)}(\omega_1, \omega_2, \dots) = \mathbf{X}^{(L)}(\omega_1, \omega_2, \dots, \omega_L). \tag{5.40}$$

We also make $\mathbf{X}^{(L)}$ $\mathbb{R}^{\infty}$-valued by setting $\mathbf{X}^{\ell,(L)} \equiv 0$ for $\ell > L$. These two practices are **always assumed** in the rest of the chapter so that we can use results from Section 4 in the current setting and thus safely work with the probability space $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, \mathbf{P}^{(\infty)})$ alone. For ease of notation, we write $\mathbf{P}$ and $\mathbf{E}$ for $\mathbf{P}^{(\infty)}$ and $\mathbf{E}^{(\infty)}$, respectively.

Now we lift $\widehat{\theta}$ which lives on $(\Omega, \mathcal{F}, P)$ to the infinite product space via (5.37) as in discrete time to obtain $\widehat{\theta}^{\otimes\infty}$. Using the explicit expression (5.10) for $\widehat{\theta}$ and $P = \mathbf{P} \circ \pi_{\ell,\infty}^{-1}$ from (5.35), we get explicitly

$$\widehat{\theta}_t^{\ell,\otimes\infty} = -\frac{2b_t^{(\infty)}(\mathbf{G}_t^{\ell}(\widehat{\theta}^{\otimes\infty}) - \mathbf{E}[\mathbf{G}_t^{\ell}(\widehat{\theta}^{\otimes\infty})]) - 1}{2b_t^{(\infty)}} \boldsymbol{\lambda}_t^{\ell}, \quad \ell \in \mathbb{N}, t \in [0, T]. \tag{5.41}$$

As in (5.9), we also get

$$\mathbf{G}^{\ell}(\widehat{\theta}^{\otimes\infty})(\omega^{(\infty)}) = \big(G(\widehat{\theta})^{\ell,\otimes\infty}\big)(\omega^{(\infty)}) = G(\widehat{\theta})(\omega_{\ell}), \quad \ell \in \mathbb{N}. \tag{5.42}$$

For convenience, we also recall from (4.49) the explicit expression for $\widehat{\boldsymbol{\vartheta}}^{(L)}$ that for $\ell = 1, \dots, L$,

$$\widehat{\boldsymbol{\vartheta}}_t^{\ell,(L)} = -\frac{2b_t^{(L)}(\mathbf{G}_t^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))) - 1}{2b_t^{(L)}(1 - L^{-1})}\boldsymbol{\lambda}_t^{\ell}, \quad t \in [0,T], \tag{5.43}$$

and $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)} \equiv 0$ for $\ell > L$. The explicit definitions

$$\mathrm{em}(\mathbf{x}) = \frac{1}{L}\sum_{\ell=1}^{L}\mathbf{x}^{\ell}, \quad \mathrm{evar}(\mathbf{x}) = \mathrm{em}(\mathbf{x}^2) - \big(\mathrm{em}(\mathbf{x})\big)^2 \tag{5.44}$$

are frequently used, and so is the property in Lemma 4.7 that

$$\sup_{t\in[0,T]} b_t^{(L)} \text{ and } \sup_{t\in[0,T]} \frac{1}{b_t^{(L)}} \text{ are bounded uniformly in } \omega \in \Omega, \tag{5.45}$$

whenever Assumption 4.2 is satisfied and the MVT process K is bounded. Finally, due to $\mathrm{d}K = \lambda\,\mathrm{d}A = (\lambda)^2\,\mathrm{d}\langle M\rangle$ from (5.11), the relation (5.36) between $(\mathbf{A}^{\ell}, \langle\mathbf{M}^{\ell}\rangle, \boldsymbol{\lambda}^{\ell}, \mathbf{K}^{\ell})$ and $(A, \langle M\rangle, \lambda, K)$, and the fact that K is deterministic by Assumption 4.11, we have

$$\mathrm{d}K = \boldsymbol{\lambda}^{\ell}\,\mathrm{d}\mathbf{A}^{\ell} = (\boldsymbol{\lambda}^{\ell})^2\,\mathrm{d}\langle\mathbf{M}^{\ell}\rangle, \quad \ell \in \mathbb{N}. \tag{5.46}$$

Lemma 5.7. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied. Then we have for* $L \in \mathbb{N}$ *and* $\ell = 1, \dots, L$ *that*

$$\mathbf{E}[\mathbf{G}_t^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big], \tag{5.47}$$

$$\mathbf{E}[\mathbf{G}_t^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})^2] = \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})^2\big)\big]. \tag{5.48}$$

In particular, the random variables $(\mathbf{G}_t^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}))_{\ell=1,\dots,L}$ *all have the same first two moments for* $t \in [0,T]$.

Proof. Let us prove (5.47) and (5.48) separately.

1) For (5.47), we directly use the explicit expression (5.43) for $\widehat{\boldsymbol{\vartheta}}^{(L)}$, then $\mathrm{d}K = \boldsymbol{\lambda}^{\ell}\,\mathrm{d}\mathbf{A}^{\ell}$ from (5.46), and finally Fubini's theorem with deterministic K to

compute

$$\begin{aligned}
&\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})] \\
&= \mathbf{E}\bigg[\int_0^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{A}_s^\ell + \int_0^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{M}_s^\ell\bigg] \\
&= \mathbf{E}\bigg[\int_0^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{A}_s^\ell\bigg] \\
&= -\int_0^t \mathbf{E}\bigg[\frac{\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{1-L^{-1}} - \frac{1}{2b_s^{(L)}(1-L^{-1})}\bigg]\,\mathrm{d}K_s. \qquad (5.49)
\end{aligned}$$

The second equality uses that $\int \widehat{\boldsymbol{\vartheta}}^{\ell,(L)}\,\mathrm{d}\mathbf{M}^\ell$ are martingales for $\ell = 1,\dots,L$ in the infinite product space due to $\widehat{\boldsymbol{\vartheta}}^{(L)} \in \boldsymbol{\Theta}_{\mathrm{S}}^{(L)} = L^2(\mathbf{M})$ which follows from Corollary 4.10, Lemma 3.17 and the embedding (5.40) of the L-extended market into the infinite product space. The application of Fubini's theorem uses

$$\int_0^t \mathbf{E}\bigg[\bigg|\frac{\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{1-L^{-1}}\bigg| + \frac{1}{2b_s^{(L)}(1-L^{-1})}\bigg]\,\mathrm{d}K_s < \infty \qquad (5.50)$$

due to $\sup_{t\in[0,T]}|\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})| \in L^2$ for $\ell = 1,\dots,L$, hence by Cauchy–Schwarz also $\sup_{t\in[0,T]}|\mathrm{em}(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}))| \in L^2$, and finally the boundedness of $\sup_{t\in[0,T]}\frac{1}{b_t^{(L)}}$ from (5.45) and the boundedness of K implied by its non-randomness from Assumption 4.11. Then inserting the definition (5.44) of $\mathrm{em}(\mathbf{x})$ into (5.49) and using that definition again yields

$$\begin{aligned}
\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] &= -\frac{1}{L}\sum_{\ell=1}^L \int_0^t \mathbf{E}\bigg[\frac{\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{1-L^{-1}} \\
&\qquad\qquad - \frac{1}{2b_s^{(L)}(1-L^{-1})}\bigg]\,\mathrm{d}K_s \\
&= \int_0^t \frac{1}{2b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s. \qquad (5.51)
\end{aligned}$$

Set $D_t = \mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$ for $t \in [0,T]$. We then subtract (5.51) from (5.49) to obtain

$$D_t = -\int_0^t \frac{D_s}{1-L^{-1}}\,\mathrm{d}K_s.$$

Because $D_0 = 0$, we obtain by the uniqueness of the solution (in the class of continuous (or even RCLL) processes) to this integral equation that $D_t = 0$ for all $t \in [0,T]$. This implies the desired identity $\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \mathbf{E}[\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$

for all $\ell = 1, \dots, L$.

2) To prove (5.48), we first establish the identities

$$\mathbf{E}\Big[\Big(\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big] = \mathbf{E}\big[\mathrm{evar}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big], \tag{5.52}$$

$$\mathbf{E}\big[\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] = \mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big], \tag{5.53}$$

for $\ell = 1, \dots, L$ and $t \in [0, T]$ in order. The desired identity (5.48) is a direct consequence of (5.52) and (5.53).

Like in part 1), we consider the difference between the left- and right-hand sides in (5.52). We show that it satisfies a linear integral equation starting from 0 and thus must be identically 0. This gives (5.52). For this, we need to find integral representations for $\mathbf{E}[(\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2], \mathbf{E}[\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))], \mathbf{E}[(\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]$ and $\mathbf{E}[\mathrm{evar}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$.

First, we apply Itô's lemma and use the explicit formula (5.43) for $\widehat{\boldsymbol{\vartheta}}^{(L)}$, the identity $\mathrm{d}K = \boldsymbol{\lambda}^\ell \,\mathrm{d}\mathbf{A}^\ell$ from (5.46) and Fubini's theorem to compute

$$\begin{aligned}
\mathbf{E}\big[\big(\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] &= 2\mathbf{E}\bigg[\int_0^t \mathbf{G}^\ell_s(\widehat{\boldsymbol{\vartheta}}^{(L)})\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_s \,\mathrm{d}\mathbf{S}^\ell_s\bigg] + \mathbf{E}\bigg[\int_0^t (\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_s)^2 \,\mathrm{d}\langle \mathbf{M}^\ell\rangle_s\bigg] \\
&=: 2I^\ell_a + Q^\ell \\
&= 2\int_0^t \bigg(- \frac{\mathbf{E}[(\mathbf{G}^\ell_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2 - \mathbf{G}^\ell_s(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{1 - L^{-1}} \\
&\qquad\qquad + \frac{\mathbf{E}[\mathbf{G}^\ell_s(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{2b^{(L)}_s(1 - L^{-1})}\bigg) \,\mathrm{d}K_s + Q^\ell,
\end{aligned} \tag{5.54}$$

where I^ℓ_a and Q^ℓ are defined by

$$I^\ell_a := \mathbf{E}\bigg[\int_0^t \mathbf{G}^\ell_s(\widehat{\boldsymbol{\vartheta}}^{(L)})\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_s \,\mathrm{d}\mathbf{S}^\ell_s\bigg], \tag{5.55}$$

$$Q^\ell := \mathbf{E}\bigg[\int_0^t (\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}_s)^2 \,\mathrm{d}\langle \mathbf{M}^\ell\rangle_s\bigg]. \tag{5.56}$$

The third equality in (5.54) also uses that $\int \mathbf{G}^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\widehat{\boldsymbol{\vartheta}}^{\ell(L)} \,\mathrm{d}\mathbf{M}^\ell$ are martingales for $\ell = 1, \dots, L$ by Lemma 5.4 because $\sup_{t\in[0,T]} |\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})| \in L^2$ and $\widehat{\boldsymbol{\vartheta}}^{(L)}$ is in $L^2(\mathbf{M})$. These integrability properties are also used to verify that the integrand of the first integral in the last equality of (5.54) is in $L^1(\mathbf{P} \otimes K)$ so that Fubini's theorem can be applied. The term Q^ℓ defined in (5.56) can be computed similarly as

$$
\begin{aligned}
Q^\ell &= \mathbf{E}\bigg[\int_0^t (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)})^2 \,\mathrm{d}\langle \mathbf{M}^\ell\rangle_s\bigg] \\
&= \int_0^t \mathbf{E}\bigg[\bigg(\frac{\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{1-L^{-1}}\bigg)^2 - \frac{\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{b_s^{(L)}(1-L^{-1})^2} \\
&\qquad\qquad + \bigg(\frac{1}{2b_s^{(L)}(1-L^{-1})}\bigg)^2\bigg] \,\mathrm{d}K_s \\
&= \int_0^t \mathbf{E}\bigg[\bigg(\frac{\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{1-L^{-1}}\bigg)^2 + \bigg(\frac{1}{2b_s^{(L)}(1-L^{-1})}\bigg)^2\bigg] \,\mathrm{d}K_s. \qquad (5.57)
\end{aligned}
$$

The last line also uses that the expectations of $\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})$ and $\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))$ are the same by (5.47). Averaging over the formulas (5.54) and (5.57) for I_a^ℓ and Q^ℓ, respectively, and using $\mathrm{evar}(\mathbf{x}) = \mathrm{em}(\mathbf{x}^2) - (\mathrm{em}(\mathbf{x}))^2$ from (5.44), we also obtain

$$
\frac{1}{L}\sum_{\ell=1}^L I_a^\ell = \int_0^t -\frac{\mathbf{E}[\mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{1-L^{-1}} + \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{2b_s^{(L)}(1-L^{-1})} \,\mathrm{d}K_s, \qquad (5.58)
$$

$$
\frac{1}{L}\sum_{\ell=1}^L Q^\ell = \int_0^t \mathbf{E}\bigg[\frac{\mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{(1-L^{-1})^2} + \bigg(\frac{1}{2b_s^{(L)}(1-L^{-1})}\bigg)^2\bigg] \,\mathrm{d}K_s, \qquad (5.59)
$$

respectively. These identities are used multiple times later.

Second, we use the product rule with the continuity of the processes from Assumption 4.2, and the definition (5.44) of the empirical average, and finally $\langle \mathbf{M}^\ell, \mathbf{M}^m\rangle \equiv 0$ for $\ell \neq m$ from Lemma 3.13, 2) to get

$$
\begin{aligned}
\mathbf{E}\big[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] &= \mathbf{E}\bigg[\frac{1}{L}\sum_{m=1}^L \int_0^t \mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\widehat{\boldsymbol{\vartheta}}_s^{m,(L)} \,\mathrm{d}\mathbf{A}_s^m\bigg] \\
&\quad + \mathbf{E}\bigg[\int_0^t \mathrm{em}\big(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)} \,\mathrm{d}\mathbf{A}_s^\ell\bigg] \\
&\quad + \mathbf{E}\bigg[\frac{1}{L}\int_0^t (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)})^2 \,\mathrm{d}\langle \mathbf{M}^\ell\rangle_s\bigg] \\
&=: I_b^\ell + I_c^\ell + L^{-1}Q^\ell, \qquad (5.60)
\end{aligned}
$$

where

$$I_b^\ell = \mathbf{E}\bigg[\frac{1}{L}\sum_{m=1}^{L}\int_0^t \mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\widehat{\boldsymbol{\vartheta}}_s^{m,(L)}\,\mathrm{d}\mathbf{A}_s^m\bigg], \tag{5.61}$$

$$I_c^\ell = \mathbf{E}\bigg[\int_0^t \mathrm{em}\big(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{A}_s^\ell\bigg] \tag{5.62}$$

and $Q^\ell = \mathbf{E}[\int_0^t (\widehat{\boldsymbol{\vartheta}}^{\ell,(L)})_s^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle_s]$ is from (5.56). The first equality also uses that all the integral terms with respect to $\mathbf{M}^m$ are martingales for $m = 1,\dots,L$ due to a similar reasoning as for (5.54). For the first term I_b^ℓ, we use the explicit formula (5.43) for $\widehat{\boldsymbol{\vartheta}}^{(L)}$, then $\mathrm{d}K = (\boldsymbol{\lambda}^m)^2\,\mathrm{d}\langle\mathbf{M}^m\rangle$ for $m = 1,\dots,L$ from (5.46), next the definition (5.44) of the empirical average and finally Fubini's theorem to obtain

$$\begin{aligned}
I_b^\ell &= \mathbf{E}\bigg[\frac{1}{L}\sum_{m=1}^{L}\int_0^t \bigg(-\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) \\
&\qquad\qquad \times \frac{2b_s^{(L)}(\mathbf{G}_s^m(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})) - 1}{2b_s^{(L)}(1-L^{-1})}(\boldsymbol{\lambda}_s^m)^2\bigg)\,\mathrm{d}\langle\mathbf{M}^m\rangle_s\bigg] \\
&= \mathbf{E}\bigg[\frac{1}{L}\sum_{m=1}^{L}\int_0^t \mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\frac{1}{2b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s\bigg] \\
&= \int_0^t \frac{\mathbf{E}[\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{2b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s.
\end{aligned}$$

The applicability of Fubini's theorem is ensured because $\sup_{t\in[0,T]}|\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})|$ is in L^2, similarly as in (5.50). For the second term I_c^ℓ, we analogously obtain

$$\begin{aligned}
I_c^\ell = \int_0^t \bigg(&-\frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - (\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]}{1-L^{-1}} \\
&+ \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{2b_s^{(L)}(1-L^{-1})}\bigg)\,\mathrm{d}K_s.
\end{aligned}$$

Averaging in the above display yields

$$\frac{1}{L}\sum_{\ell=1}^{L} I_c^\ell = \int_0^t \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{2b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s. \tag{5.63}$$

We insert I_b^ℓ and I_c^ℓ from above into $\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))] = I_b^\ell + I_c^\ell + L^{-1}Q^\ell$ due to (5.60) and use that the expectations of $\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})$ and $\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))$ are the same from (5.47) to obtain

$$\begin{aligned}
&\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)] \\
&= \int_0^t \Bigg(\frac{\mathbf{E}[\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{b_s^{(L)}(1-L^{-1})} \\
&\qquad - \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - (\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]}{1-L^{-1}} \Bigg) \,\mathrm{d}K_s + L^{-1}Q^\ell. \quad (5.64)
\end{aligned}$$

Third, we seek an integral representation for $\mathbf{E}[(\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]$. Recall from (4.13) that

$$\mathrm{d}\Big(\big(\mathrm{em}\big(\mathbf{G}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big)^2\Big) = \frac{1}{L}\sum_{\ell=1}^{L} 2\mathrm{em}\big(\mathbf{G}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}\,\mathrm{d}\mathbf{S}^\ell + \frac{1}{L^2}\sum_{\ell=1}^{L}(\widehat{\boldsymbol{\vartheta}}^{\ell,(L)})^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle.$$

Using the definitions (5.62) and (5.56) for I_2^ℓ and Q^ℓ, respectively, we obtain

$$\mathbf{E}\Big[\big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big)^2\Big] = \frac{1}{L}\sum_{\ell=1}^{L} 2I_c^\ell + \frac{1}{L^2}\sum_{\ell=1}^{L} Q^\ell. \quad (5.65)$$

Now we can get an integral formula for $\mathbf{E}[(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]$ in (5.52). Using (5.54) for $\mathbf{E}[(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2]$, (5.64) for $\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$ and (5.65) for $\mathbf{E}[(\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]$ yields

$$\begin{aligned}
&\mathbf{E}\Big[\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big)^2\Big] \\
&= \mathbf{E}\big[\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)^2\big] - 2\mathbf{E}[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)] + \mathbf{E}\Big[\big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big)^2\Big] \\
&= 2\int_0^t \Bigg(- \frac{\mathbf{E}[(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2 - \mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{1-L^{-1}} + \frac{\mathbf{E}[\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{2b_s^{(L)}(1-L^{-1})} \Bigg) \,\mathrm{d}K_s \\
&\quad + Q^\ell \\
&\quad - 2\int_0^t \Bigg(\frac{\mathbf{E}[\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{b_s^{(L)}(1-L^{-1})} - \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - (\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]}{1-L^{-1}} \Bigg) \,\mathrm{d}K_s \\
&\quad - 2L^{-1}Q^\ell + \frac{1}{L}\sum_{m=1}^{L} 2I_c^m + \frac{1}{L^2}\sum_{m=1}^{L} Q^m \\
&= -\int_0^t \frac{\mathbf{E}[\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s - 2\int_0^t \frac{\mathbf{E}[(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2]}{1-L^{-1}}\,\mathrm{d}K_s \\
&\quad + (1-2L^{-1})Q^\ell + \frac{1}{L}\sum_{m=1}^{L} 2I_c^m + \frac{1}{L^2}\sum_{m=1}^{L} Q^m. \quad (5.66)
\end{aligned}$$

Fourth, we look for an expression for $\mathbf{E}[\operatorname{evar}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$. Recall from (4.14) that the dynamics of $\operatorname{evar}(\mathbf{G}(\widehat{\boldsymbol{\vartheta}}^{(L)}))$ is given by

$$\mathrm{d}\Big(\operatorname{evar}\big(\mathbf{G}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big) = \frac{1}{L}\sum_{m=1}^{L}\bigg(2\Big(\mathbf{G}^m(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \operatorname{em}\big(\mathbf{G}(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)\widehat{\boldsymbol{\vartheta}}^{m,(L)}\,\mathrm{d}\mathbf{S}^m + (1-L^{-1})(\widehat{\boldsymbol{\vartheta}}^{m,(L)})^2\,\mathrm{d}\langle\mathbf{M}^m\rangle\bigg).$$

Using the definitions (5.55) and (5.56) of I_a^ℓ and Q^ℓ, respectively, yields

$$\mathbf{E}\big[\operatorname{evar}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] = \frac{1}{L}\sum_{m=1}^{L}2(I_a^m - I_c^m) + \frac{1-L^{-1}}{L}\sum_{m=1}^{L}Q^m. \tag{5.67}$$

This ends the preparation; we now prove the desired formula (5.52). We set $F_t = \mathbf{E}[(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \operatorname{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2 - \operatorname{evar}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$. Subtracting the expression (5.67) for $\mathbf{E}[\operatorname{evar}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$ from (5.66) for $\mathbf{E}[(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \operatorname{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]$, we obtain

$$\begin{aligned}F_t &= -\int_0^t \frac{\mathbf{E}[\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s - 2\int_0^t \frac{\mathbf{E}[(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \operatorname{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2]}{1-L^{-1}}\,\mathrm{d}K_s\\ &\quad + (1-2L^{-1})Q^\ell + \frac{1}{L}\sum_{m=1}^{L}2I_c^m + \frac{1}{L^2}\sum_{m=1}^{L}Q^m - \frac{1}{L}\sum_{m=1}^{L}2(I_a^m - I_c^m)\\ &\quad - \frac{1-L^{-1}}{L}\sum_{m=1}^{L}Q^m\\ &= -2\int_0^t \frac{\mathbf{E}[(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \operatorname{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))^2 - \operatorname{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{1-L^{-1}}\,\mathrm{d}K_s\\ &\quad + (1-2L^{-1})Q^\ell + \frac{1}{L^2}\sum_{m=1}^{L}Q^m - \frac{1-L^{-1}}{L}\sum_{m=1}^{L}Q^m.\end{aligned} \tag{5.68}$$

In the second equality, we use the formulas (5.58) and (5.63) for $\frac{1}{L}\sum_{m=1}^{L} I_a^m$ and

$\frac{1}{L}\sum_{m=1}^{L} I_c^m$, respectively, to derive the identity

$$\begin{aligned}\frac{1}{L}\sum_{m=1}^{L}(4I_c^m - 2I_a^m) &= 2\int_0^t \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s \\ &\quad + 2\int_0^t \frac{\mathbf{E}[\mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{1-L^{-1}} - \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{2b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s \\ &= \int_0^t \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s + 2\int_0^t \frac{\mathbf{E}[\mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{1-L^{-1}}\end{aligned}$$

and then invoke the identity $\mathbf{E}[\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$ for $s \in [0,t]$ asserted in part 1). Using the expressions (5.57) for Q^ℓ and (5.59) for $\frac{1}{L}\sum_{m=1}^{L} Q^m$ simplifies the last line to

$$\begin{aligned}&(1-2L^{-1})Q^\ell + \frac{1}{L^2}\sum_{m=1}^{L} Q^m - \frac{1-L^{-1}}{L}\sum_{m=1}^{L} Q^m \\ &= (1-2L^{-1})\Big(Q^\ell - \frac{1}{L}\sum_{m=1}^{L} Q^m\Big) \\ &= (1-2L^{-1})\int_0^t \frac{\mathbf{E}[(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2 - \mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{(1-L^{-1})^2}\,\mathrm{d}K_s. \quad (5.69)\end{aligned}$$

Inserting (5.69) into (5.68) yields

$$F_t = -\frac{1}{(1-L^{-1})^2}\int_0^t F_s\,\mathrm{d}K_s, \quad t \in [0,T].$$

Because $F_0 = 0$, we get that $F_t = 0$ for all $t \in [0,T]$ is the unique solution to the above integral equation. This proves (5.52).

We use a similar technique to prove (5.53). Set

$$H_t = \mathbf{E}\Big[\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - \Big(\mathrm{em}(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}))\Big)^2\Big], \quad t \in [0,T].$$

Note that due to (5.52) proved just above, we deduce from the third equality of (5.69) that

$$\begin{aligned}Q^\ell - \frac{1}{L}\sum_{m=1}^{L} Q^m &= \int_0^t \frac{\mathbf{E}[(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2 - \mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{(1-L^{-1})^2}\,\mathrm{d}K_s \\ &= 0.\end{aligned}$$

Subtracting $\mathbf{E}[(\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]$ given in (5.65) from $\mathbf{E}[\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$ given in (5.64), and then using $\frac{1}{L}\sum_{m=1}^L 2I^m_c = \int_0^t \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{b^{(L)}_s(1-L^{-1})}\,\mathrm{d}K_s$ from the explicit formula (5.63) and finally the above display, we obtain

$$\begin{aligned} H_t &= \int_0^t \left(\frac{\mathbf{E}[\mathbf{G}^\ell_s(\widehat{\boldsymbol{\vartheta}}^{(L)})]}{b^{(L)}_s(1-L^{-1})} - \frac{\mathbf{E}[\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))\mathbf{G}^\ell_s(\widehat{\boldsymbol{\vartheta}}^{(L)}) - (\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})))^2]}{1-L^{-1}} \right) \mathrm{d}K_s \\ &\quad + L^{-1}Q^\ell - \frac{1}{L}\sum_{m=1}^L 2I^m_c - \frac{1}{L^2}\sum_{m=1}^L Q^m \\ &= -\int_0^t \frac{H_s}{1-L^{-1}}\,\mathrm{d}K_s + L^{-1}\Big(Q^\ell - \frac{1}{L}\sum_{m=1}^L Q^m \Big) \\ &= -\int_0^t \frac{H_s}{1-L^{-1}}\,\mathrm{d}K_s. \end{aligned}$$

Again because $H_0 = 0$, the unique solution to the above integral equation is $H_t = 0$ for $t \in [0,T]$. This proves (5.53). □

5.3 Convergence of strategies – main results

We are ready to prove the main result of these two subsections: for $\ell \in \mathbb{N}$,

$$\max_{\ell=1,\dots,L} \mathbf{E}\Big[\sup_{t\in[0,T]} \big(\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}^\ell_t(\widehat{\theta}^{\otimes\infty}) \big)^2 \Big] \longrightarrow 0 \quad \text{as } L \to \infty.$$

Suppose that Assumptions 4.2 and 4.11 are satisfied so that S is a continuous semimartingale satisfying the structure condition (SC) and the MVT process K is deterministic. Define the process $m^{(L)}$ by $m^{(L)}_t := \mathbf{E}[\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})]$ for $t \in [0,T]$. By Lemma 5.7, (5.47), this quantity is equal to $\mathbf{E}[\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$ and thus does not depend on ℓ. Using the explicit formula (5.51) for $\mathbf{E}[\mathrm{em}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$, we obtain for $t \in [0,T]$ that

$$m^{(L)}_t = \mathbf{E}[\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})] = \mathbf{E}\big[\mathrm{em}\big(\mathbf{G}^\ell_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] = \int_0^t \frac{1}{2b^{(L)}_s(1-L^{-1})}\,\mathrm{d}K_s. \tag{5.70}$$

Because $\widehat{\theta}^{\ell,\otimes\infty}(\omega^{(\infty)}) = \widehat{\theta}(\omega_\ell)$, $(\mathbf{G}^\ell(\widehat{\theta}^{\otimes\infty}))(\omega^{(\infty)}) = G(\widehat{\theta})(\omega_\ell)$ and $P = \mathbf{P} \circ \pi^{-1}_{\ell,\infty}$ from (5.37), (5.42) and (5.35), respectively, the expectation $\mathbf{E}[\mathbf{G}^\ell_t(\widehat{\theta}^{\otimes\infty})]$ is also independent of ℓ. Hence we can define $m^{(\infty)}$ as below and use the identity $\mathbf{G}^\ell(\widehat{\theta}^{\otimes\infty})(\omega^{(\infty)}) = \widehat{X}(\omega_\ell)$ from Lemma 5.3 and the explicit expression (5.17) for

$E[\widehat{X}_t]$ to obtain

$$m_t^{(\infty)} := \mathbf{E}[\mathbf{G}_t^{\ell}(\widehat{\theta}^{\otimes\infty})] = E[G_t(\widehat{\theta})] = E[\widehat{X}_t] = \int_0^t \frac{1}{2b_s^{(\infty)}}\,\mathrm{d}K_s, \quad t \in [0,T]. \tag{5.71}$$

The elementary inequality $|\exp(x)-1| = O(x)$ as $x \to 0$ yields with $x = \pm\frac{K_t-K_T}{L-1}$ that

$$\begin{aligned}
&\sup_{t\in[0,T]}\left|\exp\Big(\pm\frac{K_t-K_T}{1-L^{-1}}\Big) - \exp\big(\pm(K_t-K_T)\big)\right| \\
&= \sup_{t\in[0,T]}\left\{\exp\big(\pm(K_t-K_T)\big)\Big(\exp\Big(\pm\frac{K_t-K_T}{L-1}\Big)-1\Big)\right\} \\
&= O(L^{-1}).
\end{aligned}$$

Then using $b_t^{(L)} = \xi\exp(\frac{K_t-K_T}{1-L^{-1}})$ from (4.26), $b_t^{(\infty)} = \xi\exp(K_t-K_T)$ from (5.12) and the above display, we recall and refine the convergence of $b^{(L)}$ to $b^{(\infty)}$ given in (5.16) as

$$\sup_{t\in[0,T]}|b_t^{(L)} - b_t^{(\infty)}| + \sup_{t\in[0,T]}\left|\frac{1}{b_t^{(L)}} - \frac{1}{b_t^{(\infty)}}\right| = O(L^{-1}) \quad \text{as } L\to\infty. \tag{5.72}$$

This in turn yields

$$\begin{aligned}
\sup_{t\in[0,T]}|m_t^{(L)} - m_t^{(\infty)}| &\le K_T\bigg(\frac{1}{1-L^{-1}}\sup_{s\in[0,T]}\left|\frac{1}{2b_s^{(L)}} - \frac{1}{2b_s^{(\infty)}}\right| \\
&\qquad + \left|\frac{1}{1-L^{-1}} - 1\right|\sup_{s\in[0,T]}\frac{1}{2b_s^{(\infty)}}\bigg) \\
&= O(L^{-1}) \quad \text{as } L\to\infty.
\end{aligned} \tag{5.73}$$

Proposition 5.8. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied. Then*

$$\mathbf{E}\Big[\sup_{t\in[0,T]}\Big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - m_t^{(L)}\Big)^2\Big] = O(L^{-1}), \tag{5.74}$$

$$\mathbf{E}\Big[\sup_{t\in[0,T]}\Big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - m_t^{(\infty)}\Big)^2\Big] = O(L^{-1}). \tag{5.75}$$

as $L\to\infty$.

Proof. The second property (5.75) immediately follows from (5.74) and the estimate $\sup_{t\in[0,T]}|m_t^{(L)} - m_t^{(\infty)}| = O(L^{-1})$ given in (5.73); so we only need to prove (5.74). We use the product rule and that $m^{(L)}$ has finite variation by its explicit formula (5.70) to obtain

$$(m_t^{(L)})^2 = 2\int_0^t m_s^{(L)}\,\mathrm{d}m_s^{(L)} = \int_0^t \frac{m_s^{(L)}}{b_s^{(L)}(1-L^{-1})}\,\mathrm{d}K_s. \tag{5.76}$$

On the other hand, using the definition (5.44) of $\mathrm{em}(\mathbf{x})$, then $(x+y)^2 \le 2x^2+2y^2$, the formulas for $\widehat{\boldsymbol{\vartheta}}^{(L)}$ from (5.43) and $m^{(L)}$ from (5.70), respectively, and finally $\mathrm{d}K = \boldsymbol{\lambda}^\ell\,\mathrm{d}\mathbf{A}^\ell$ from (5.46) yields

$$\begin{aligned}
&\Big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - m_t^{(L)}\Big)^2 \\
&= \Big(\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{M}_s^\ell + \frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{A}_s^\ell - m_t^{(L)}\Big)^2 \\
&\le 2\Big(\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{M}_s^\ell\Big)^2 \\
&\quad + 2\Big(\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \Big(-\frac{2b_s^{(L)}(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))) - 1}{2b_s^{(L)}(1-L^{-1})} \\
&\qquad\qquad - \frac{1}{2b_s^{(L)}(1-L^{-1})}\Big)\,\mathrm{d}K_s\Big)^2 \\
&= 2\Big(\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{M}_s^\ell\Big)^2.
\end{aligned}$$

The last equality again uses the definition (5.44) of the empirical average. Hence due to the above inequality, the BDG inequality and $\langle \mathbf{M}^\ell, \mathbf{M}^m\rangle \equiv 0$ for $\ell \neq m$ from Lemma 3.13, 2), we obtain

$$\begin{aligned}
\mathbf{E}\Big[\sup_{t\in[0,T]}\Big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - m_t^{(L)}\Big)^2\Big] &\le 2\mathbf{E}\Big[\sup_{t\in[0,T]}\Big(\frac{1}{L}\sum_{\ell=1}^{L}\int_0^t \widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)}\,\mathrm{d}\mathbf{M}_s^\ell\Big)^2\Big] \\
&\le 8\mathbf{E}\Big[\frac{1}{L^2}\sum_{\ell=1}^{L}\int_0^T (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)})^2\,\mathrm{d}\langle \mathbf{M}^\ell\rangle_s\Big].
\end{aligned}$$

We now claim that

$$\sup_{L\in\mathbb{N}, L\ge 2}\Big(\mathbf{E}\Big[\frac{1}{L}\sum_{\ell=1}^{L}\int_0^T (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)})^2\,\mathrm{d}\langle \mathbf{M}^\ell\rangle_s\Big]\Big) < \infty. \tag{5.77}$$

Suppose (5.77) is true. Then inserting this bound into the inequality just before gives

$$\mathbf{E}\Big[\sup_{t\in[0,T]}\Big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)-m_t^{(L)}\Big)^2\Big]$$
$$\leq \frac{8}{L}\sup_{L\in\mathbb{N},L\geq 2}\left(\mathbf{E}\Big[\frac{1}{L}\sum_{\ell=1}^{L}\int_0^T(\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)})^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle_s\Big]\right)=O(L^{-1})$$

as $L\to\infty$ as desired. Now let us prove (5.77). By the definition (5.56) of Q^ℓ and the formula (5.59) for $\frac{1}{L}\sum_{\ell=1}^{L}Q^\ell$, we get

$$\mathbf{E}\Big[\frac{1}{L}\sum_{\ell=1}^{L}\int_0^T(\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)})^2\,\mathrm{d}\langle\mathbf{M}^\ell\rangle_s\Big]$$
$$=\frac{1}{L}\sum_{\ell=1}^{L}Q^\ell$$
$$=\int_0^T\mathbf{E}\left[\frac{\mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{(1-L^{-1})^2}+\left(\frac{1}{2b_s^{(L)}(1-L^{-1})}\right)^2\right]\mathrm{d}K_s.$$

Because $\sup_{t\in[0,T]}|\frac{1}{b_t^{(L)}}-\frac{1}{b_t^{(\infty)}}|=O(L^{-1})$ by (5.72), we have obviously

$$\sup_{L\in\mathbb{N},L\geq 2}\sup_{t\in[0,T]}\frac{1}{2b_t^{(L)}}<\infty. \tag{5.78}$$

Using this observation in the equation above and also that K_T is bounded (even deterministic), we obtain that a sufficient condition for (5.77) is

$$\sup_{L\in\mathbb{N},L\geq 2}\sup_{t\in[0,T]}\mathbf{E}\big[\mathrm{evar}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big]<\infty. \tag{5.79}$$

To prove the latter, we insert the explicit formulas (5.58) for $\frac{1}{L}\sum_{m=1}^{L}I_a^m$, (5.63) for $\frac{1}{L}\sum_{m=1}^{L}I_c^m$ and (5.59) for $\frac{1}{L}\sum_{m=1}^{L}Q^m$ into the identity

$$\mathbf{E}\big[\mathrm{evar}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big]=\frac{1}{L}\sum_{m=1}^{L}\big(2I_a^m-2I_c^m-(1-L^{-1})Q^m\big)$$

from (5.67) and due to the nonnegativity of $\mathbf{E}[\mathrm{evar}(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)}))]$ obtain

$$\begin{aligned}
&\mathbf{E}\big[\mathrm{evar}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] \\
&= -2\int_0^t \frac{\mathbf{E}[\mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))]}{1-L^{-1}}\,\mathrm{d}K_s \\
&\quad + (1-L^{-1})\int_0^t \mathbf{E}\bigg[\frac{\mathrm{evar}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{(1-L^{-1})^2} + \bigg(\frac{1}{2b_s^{(L)}(1-L^{-1})}\bigg)^2\bigg]\,\mathrm{d}K_s \\
&\le \int_0^t \bigg(\frac{1}{2b_s^{(L)}(1-L^{-1})}\bigg)^2\,\mathrm{d}K_s.
\end{aligned}$$

The last line also uses that $b^{(L)}$ is non-random by Corollary 4.12 because the MVT process K is deterministic due to Assumption 4.11. Thus using the elementary fact that $(1-L^{-1})^2 \ge \frac{1}{4}$ for $L \ge 2$ and (5.78) in the above inequality yields

$$\sup_{L\in\mathbb{N},L\ge 2}\ \sup_{t\in[0,T]} \mathbf{E}\big[\mathrm{evar}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] \le K_T \sup_{L\in\mathbb{N},L\ge 2}\ \sup_{t\in[0,T]} \bigg(\frac{1}{b_t^{(L)}}\bigg)^2 < \infty.$$

This proves (5.79) and completes the proof. □

Theorem 5.9. *Suppose that Assumptions* 4.2 *and* 4.11 *are satisfied, meaning that S is square-integrable and satisfies the structure condition (SC) with a deterministic MVT process K. Then $\widehat{\boldsymbol{\vartheta}}^{(L)} \to \widehat{\theta}^{\otimes\infty}$ uniformly in ℓ in the sense that*

$$\max_{\ell=1,\dots,L} \mathbf{E}\Big[\sup_{t\in[0,T]} \Big(\mathbf{G}_t^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_t^{\ell}(\widehat{\theta}^{\otimes\infty})\Big)^2\Big] = O(L^{-1}) \longrightarrow 0 \qquad \text{as } L\to\infty. \tag{5.80}$$

Remark 5.10. As in discrete time, we again take a look at (5.80) from a financial perspective. Because the processes $\mathbf{G}^{\ell}(\widehat{\boldsymbol{\vartheta}}^{(L)})$ and $\mathbf{G}^{\ell}(\widehat{\theta}^{\otimes\infty})$ represent the gains of the two strategies, the convergence (5.80) simply says that the maximum difference between the profits and losses of the two strategies vanishes in the limit (in the above sense).

Proof. The idea of the proof is analogous to that of Proposition 5.8. We start by writing the gains process as a stochastic integral and use the Cauchy–Schwarz inequality to obtain

$$
\begin{aligned}
\big(\mathbf{G}_t^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_t^\ell(\widehat{\theta}^{\otimes\infty})\big)^2 &= \Big(\int_0^t (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)} - \widehat{\theta}_s^{\ell,\otimes\infty})\,\mathrm{d}\mathbf{S}_s^\ell\Big)^2 \\
&\le 2\Big(\int_0^t (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)} - \widehat{\theta}_s^{\ell,\otimes\infty})\,\mathrm{d}\mathbf{M}_s^\ell\Big)^2 \\
&\quad + 2\Big(\int_0^t (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)} - \widehat{\theta}_s^{\ell,\otimes\infty})\,\mathrm{d}\mathbf{A}_s^\ell\Big)^2 \\
&= 2I_t^a + 2I_t^b.
\end{aligned}
\tag{5.81}
$$

We claim for $i \in \{a, b\}$ that for some constant $C > 0$ not depending on ℓ,

$$
\mathbf{E}\Big[\sup_{s\in[0,t]} I_s^i\Big] \le O(L^{-1}) + \int_0^t C\mathbf{E}\Big[\sup_{r\in[0,s]} \big(\mathbf{G}_r^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_r^\ell(\widehat{\theta}^{\otimes\infty})\big)^2\Big]\,\mathrm{d}K_s \tag{5.82}
$$

as $L \to \infty$. Then the main assertion (5.80) follows from combining (5.81) and (5.82), taking the supremum over ℓ and using Gronwall's inequality.

Let us discuss the cases $i \in \{a, b\}$ separately. For $i = 1$, we apply the BDG inequality, then use the formulas (5.43) for $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ and (5.41) for $\widehat{\theta}^{\ell,\otimes\infty}$, and finally invoke $K = \int (\boldsymbol{\lambda}^\ell)^2\,\mathrm{d}\langle \mathbf{M}^\ell\rangle$ from (5.46) and Fubini's theorem to estimate

$$
\begin{aligned}
\mathbf{E}\Big[\sup_{s\in[0,t]} I_s^a\Big] &\le 4\mathbf{E}\Big[\int_0^t (\widehat{\boldsymbol{\vartheta}}_s^{\ell,(L)} - \widehat{\theta}_s^{\ell,\otimes\infty})^2\,\mathrm{d}\langle \mathbf{M}^\ell\rangle_s\Big] \\
&= 4\int_0^t \mathbf{E}\Big[\Big(-\frac{2b_s^{(L)}(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}^{(L)}))) - 1}{2b_s^{(L)}(1-L^{-1})} \\
&\qquad + \frac{2b_s^{(\infty)}(\mathbf{G}_s^\ell(\widehat{\theta}^{\otimes\infty}) - m_s^{(\infty)}) - 1}{2b_s^{(\infty)}}\Big)^2\Big]\,\mathrm{d}K_s \\
&\le 12\Big(\int_0^t \mathbf{E}\Big[\Big(-\frac{\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})}{1-L^{-1}} + \mathbf{G}_s^\ell(\widehat{\theta}^{\otimes\infty})\Big)^2\Big] \\
&\qquad + \mathbf{E}\Big[\Big(\frac{\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{1-L^{-1}} - m_s^{(\infty)}\Big)^2\Big] \\
&\qquad + \Big(\frac{1}{2b_s^{(L)}(1-L^{-1})} - \frac{1}{2b_s^{(\infty)}}\Big)^2\,\mathrm{d}K_s\Big).
\end{aligned}
\tag{5.83}
$$

Above, we have grouped the right-hand side in the equality as a sum of three terms and then used the Cauchy–Schwarz inequality to obtain the last inequality. The convergence results from Proposition 5.8, (5.75) and from (5.72) imply that the sum of the last two terms in the last inequality of (5.83) is $O(L^{-1})$. But due to the appearance of the factor $\frac{1}{1-L^{-1}}$, we nevertheless show this claim. Using first $m_s^{(\infty)} = \frac{1}{1-L^{-1}} m_s^{(\infty)} - \frac{L^{-1}}{1-L^{-1}} m_s^{(\infty)}$ and $(x+y)^2 \leq 2x^2 + 2y^2$, then the fact that $\frac{1}{1-L^{-1}} \leq 2$ for $L \geq 2$, and finally the convergence result in Proposition 5.8, (5.75) and $\sup_{t\in[0,T]} m_t^{(\infty)} < \infty$ from (5.71) gives

$$\begin{aligned}
&\int_0^t \mathbf{E}\Big[\Big(\frac{\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)}))}{1-L^{-1}} - m_s^{(\infty)}\Big)^2\Big] \,\mathrm{d}K_s \\
&\leq 2\int_0^t \mathbf{E}\Big[\Big(\frac{\mathrm{em}(\mathbf{G}_s(\widehat{\boldsymbol{\vartheta}}^{(L)})) - m_s^{(\infty)}}{1-L^{-1}}\Big)^2\Big] \,\mathrm{d}K_s + 2\int_0^t \Big(\frac{L^{-1}}{1-L^{-1}} m_s^{(\infty)}\Big)^2 \,\mathrm{d}K_s \\
&\leq 8K_T\Big(\mathbf{E}\Big[\sup_{t\in[0,T]} \Big(\mathrm{em}\big(\mathbf{G}_t(\widehat{\boldsymbol{\vartheta}}^{(L)})\big) - m_t^{(\infty)}\Big)^2\Big] + L^{-2} \sup_{t\in[0,T]} (m_t^{(\infty)})^2\Big) \\
&= O(L^{-1}) \qquad \text{as } L \to \infty.
\end{aligned} \tag{5.84}$$

Similarly, we use $\sup_{t\in[0,T]} |\frac{1}{b_t^{(L)}} - \frac{1}{b_t^{(\infty)}}| = O(L^{-1})$ as $L \to \infty$ from (5.72) to obtain

$$\begin{aligned}
\int_0^t \Big(\frac{1}{2b_s^{(L)}(1-L^{-1})} - \frac{1}{2b_s^{(\infty)}}\Big)^2 \,\mathrm{d}K_s &\leq 2\int_0^t \Big(\frac{1}{2b_s^{(L)}} - \frac{1}{2b_s^{(\infty)}}\Big)^2 \frac{1}{(1-L^{-1})^2} \,\mathrm{d}K_s \\
&\quad + 2\int_0^t \Big(\frac{L^{-1}}{2b_s^{(\infty)}(1-L^{-1})}\Big)^2 \,\mathrm{d}K_s \\
&\leq 8K_T\Bigg(\sup_{t\in[0,T]} \Big(\frac{1}{2b_s^{(L)}} - \frac{1}{2b_s^{(\infty)}}\Big)^2 \\
&\qquad + L^{-2} \sup_{t\in[0,T]} \Big(\frac{1}{b_s^{(\infty)}}\Big)^2\Bigg) \\
&= O(L^{-2}) \quad \text{as } L \to \infty.
\end{aligned} \tag{5.85}$$

With the same trick applied to the first term in (5.83), we obtain

$$\begin{aligned}
&\int_0^t \mathbf{E}\Big[\Big(-\frac{\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)})}{1-L^{-1}} + \mathbf{G}_s^\ell(\widehat{\theta}^{\otimes\infty})\Big)^2\Big] \,\mathrm{d}K_s \\
&\leq 2\int_0^t \mathbf{E}\Big[\Big(\frac{\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_s^\ell(\widehat{\theta}^{\otimes\infty})}{1-L^{-1}}\Big)^2\Big] \,\mathrm{d}K_s + 2\int_0^t \mathbf{E}\Big[\Big(\frac{L^{-1}\mathbf{G}_s^\ell(\widehat{\theta}^{\otimes\infty})}{1-L^{-1}}\Big)^2\Big] \,\mathrm{d}K_s \\
&\leq 8\int_0^t \mathbf{E}\Big[\sup_{r\in[0,s]} \big(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_s^\ell(\widehat{\theta}^{\otimes\infty})\big)^2\Big] \,\mathrm{d}K_s + 8K_T L^{-2} \sup_{t\in[0,T]} \mathbf{E}\big[\big(\mathbf{G}_t^\ell(\widehat{\theta}^{\otimes\infty})\big)^2\big] \\
&= 8\int_0^t \mathbf{E}\Big[\sup_{r\in[0,s]} \big(\mathbf{G}_s^\ell(\widehat{\boldsymbol{\vartheta}}^{(L)}) - \mathbf{G}_s^\ell(\widehat{\theta}^{\otimes\infty})\big)^2\Big] \,\mathrm{d}K_s + O(L^{-2}) \qquad \text{as } L \to \infty.
\end{aligned} \tag{5.86}$$

The difference here is that we bound the first term in the first inequality by the running supremum process rather than the supremum over $[0,T]$. The last equality uses that $\max_{\ell=1,\dots,L} \mathbf{E}[\sup_{t\in[0,T]}(\mathbf{G}_t^\ell(\widehat{\theta}^{\otimes\infty}))^2] = E[\sup_{t\in[0,T]}(G_t(\widehat{\theta}))^2] < \infty$ due to Lemma 5.3 and the identity $\mathbf{G}_t^\ell(\widehat{\theta}^{\otimes\infty})(\omega^{(\infty)}) = G_t(\widehat{\theta})(\omega_\ell)$ from (5.42). Inserting (5.84)–(5.86) into (5.83) yields (5.82) for $i = a$.

It is completely analogous to prove (5.82) for $i = b$. We use the formulas for $\widehat{\boldsymbol{\vartheta}}^{\ell,(L)}$ from (5.43) and $\widehat{\theta}^{\ell,\otimes\infty}$ from (5.41), then $K = \int \boldsymbol{\lambda}^\ell \, \mathrm{d}\mathbf{A}^\ell$ from (5.46), and finally the Cauchy–Schwarz inequality to get

$$\begin{aligned}
\sup_{s\in[0,t]} I_s^b &= \sup_{s\in[0,t]} \bigg(\int_0^s (\widehat{\boldsymbol{\vartheta}}_r^{\ell,(L)} - \widehat{\boldsymbol{\vartheta}}_r^{\ell,(L)}) \, \mathrm{d}\mathbf{A}_r^\ell \bigg)^2 \\
&= \sup_{s\in[0,t]} \bigg(\int_0^s \Big(- \frac{2b_r^{(L)}(\mathbf{G}_r^\ell(\boldsymbol{\vartheta}^{(L)}) - \mathrm{em}(\mathbf{G}_r(\boldsymbol{\vartheta}^{(L)}))) - 1}{2b_r^{(L)}(1 - L^{-1})} \\
&\qquad + \frac{2b_r^{(\infty)}(\mathbf{G}_r^\ell(\widehat{\theta}^{\otimes\infty}) - m_r^{(\infty)}) - 1}{2b_r^{(\infty)}} \Big) \, \mathrm{d}K_r \bigg)^2 \\
&\leq K_t \int_0^t \Big(- \frac{2b_s^{(L)}(\mathbf{G}_s^\ell(\boldsymbol{\vartheta}^{(L)}) - \mathrm{em}(\mathbf{G}_s(\boldsymbol{\vartheta}^{(L)}))) - 1}{2b_s^{(L)}(1 - L^{-1})} \\
&\qquad + \frac{2b_s^{(\infty)}(\mathbf{G}_s^\ell(\widehat{\theta}^{\otimes\infty}) - m_s^{(\infty)}) - 1}{2b_s^{(\infty)}} \Big)^2 \, \mathrm{d}K_s.
\end{aligned}$$

Taking expectations, using that K is deterministic by Assumption 4.11, and comparing the resulting expression with (5.83) for $\mathbf{E}[\sup_{s\in[0,t]} I_s^a]$ reduces the proof to that for $i = a$. □

As a byproduct, we also give an alternative proof for Theorem 5.6. The argument is completely identical to the discrete-time case in Corollary I.5.15 and therefore omitted.

Corollary 5.11. *Suppose Assumptions* 4.2 *and* 4.11 *are satisfied meaning that S is square-integrable and satisfies the structure condition (SC) with a deterministic MVT process K. Then*

$$\begin{aligned}
\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\big] &\longrightarrow E[G_T(\widehat{\theta})] \quad \textit{as } L \to \infty, \\
\mathbf{E}\big[\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})^2\big)\big] &\longrightarrow E\big[\big(G_T(\widehat{\theta})\big)^2\big] \quad \textit{as } L \to \infty, \\
\mathbf{E}\Big[\Big(\mathrm{em}\big(\mathbf{G}_T(\widehat{\boldsymbol{\vartheta}}^{(L)})\big)\Big)^2\Big] &\longrightarrow \big(E[G_T(\widehat{\theta})]\big)^2 \quad \textit{as } L \to \infty.
\end{aligned}$$

In particular, $\mathbf{E}[\widetilde{V}_T^{(L)}(\widehat{\boldsymbol{\vartheta}}^{(L)})] \to E[G_T(\widehat{\theta})] - \xi \mathrm{Var}[G_T(\widehat{\theta})]$ *as* $L \to \infty$, *and* $\widehat{\theta}$ *is an optimal strategy for the MVPS problem* (2.1).

6 Connections to the literature

In this final section, we discuss the related work in the literature. The same approach as in Li and Ng [45] for finite discrete time is used in Zhou and Li [67] to embed the MVPS problem into a family of auxiliary linear–quadratic control (LQSC) problems for diffusion models with deterministic coefficients and later extended to random coefficients in complete and incomplete markets by Lim and Zhou [48] and Lim [46], respectively. The authors use a completion of squares to obtain a stochastic Riccati equation (SRE). The optimal strategy for the auxiliary problem can then be written in terms of a solution to the SRE. It turns out that the SRE can be associated with the so-called variance-optimal martingale measure. This kind of connection is furthermore explored by Sun and Wang [65], Xia and Yan [66] and later by Fontana and Schweizer [30] to link the MVPS problem and the variance-optimal martingale measure (or the mean–variance hedging problem). These results allow one to write the solution to the MVPS problem in terms of the solution to the mean–variance hedging (MVH) problem to

$$\text{minimise } E\big[\big(1 - G_T(\theta)\big)^2\big] \quad \text{over } \theta \in \Theta.$$

The MVH problem in continuous time has been studied since the early 1990s. We refer to Duffie and Richardson [26], Schweizer [58, 59, 62], Rheinländer and Schweizer [56], Gouriéroux et al. [31] for early developments of the general theory. Černý and Kallsen [17] provide a characterisation of the optimal hedging strategy in terms of the so-called opportunity-neutral measure for general semimartingales. Stochastic control methods for the MVH problem are thoroughly analysed in Jeanblanc et al. [38]. Among these works, [56] and [31] focus more concretely on continuous processes. Studying particular classes of models often leads to more explicit expressions of the optimal strategy for the MVH problem. First, when the entire mean–variance tradeoff (MVT) process K is deterministic, or when the final value K_T is deterministic and the underlying price process is continuous, the optimal strategy for the MVH problem is given explicitly in Schweizer [59]. As explained in Example 0.3.5, we can then exploit the connection between the MVPS and MVH problems to recover our optimal strategy $\widehat{\theta}$ in (5.8) for the MVPS problem. Beyond this, Biagini et al. [10], Hobson [35], Černý and Kallsen [18] and Chiu and Wong [20] give more explicit results on the MVH problem for stochastic volatility model with or without correlation. More recently, MVPS and MVH problems have been considered for affine and quadratic rough volatility models in Han and Wong [33] and Abi Jaber et al. [1].

This chapter is inspired by the market cloning approach proposed by Ankirchner and Dermoune [5] for finite discrete time and its continuous-time extension by Fischer and Livieri [29]. Our results seem still to be among the first attempts to further develop this approach in continuous time. Because this approach has more the flavour of McKean–Vlasov control theory, we end this section with a brief discussion of some related work there. Andersson and Djehiche [4] first obtain and use a stochastic maximum principle from McKean–Vlasov control theory to solve the MVPS problem for the Black-Scholes model. Pham and Wei [53, 54] develop a dynamic programming principle (DPP) for McKean–Vlasov control problems for diffusion models with deterministic coefficients (in some cases with common noise) and apply the resulting DPP to solve the MVPS problem in this setup. An analogue of the martingale optimality principle for McKean–Vlasov control problem for diffusion models appears in Basei and Pham [8].

Chapter III

A deterministic dynamic programming principle for McKean–Vlasov control problems in finite discrete time

1 Introduction

This chapter studies the so-called McKean–Vlasov control problems in finite discrete time with a view towards the mean–variance portfolio selection (MVPS) problem in full generality. Roughly speaking, given a class of control processes θ and controlled processes X^θ, a McKean–Vlasov-type criterion is an expectation of functions that involve a direct dependence on the probability distribution of X^θ. Mathematically, we study the problem to

$$\text{maximise } E\Big[\sum_{u=1}^{T} f(u, X_u^\theta, \theta_u, P_{X_u^\theta,\theta_u}) + g(X_T^\theta, P_{X_T^\theta})\Big] \quad \text{over } \theta \in \Theta$$

for a suitable set Θ of control processes, where f and g are appropriate deterministic functions and the notation P_Y denotes the law of a random object Y. Our main results include a deterministic dynamic programming principle (DPP) for this type of problems and an application of this to solve the MVPS problem in full generality in finite discrete time.

The basic idea to obtain a DPP is inspired by Pham and Wei [52]. Namely, we embed the McKean–Vlasov control problem into a family of *deterministic* instead of stochastic control problems. Our contribution to the current literature has two aspects. First, we notice that the DPP in [52, Lemma 3.1] uses neither the i.i.d. structure for the innovations driving the controlled processes nor the restriction to the class of closed-loop controls imposed throughout the paper, and thus can be extended to open-loop controlled processes without specifying any dynamics. Second, we apply the corresponding DPP to solve the general MVPS problem in finite discrete time. Along the way, we also obtain a structural result for general linear–quadratic McKean–Vlasov problems.

This chapter is structured as follows. In Section 2, we first introduce the setup and formulate a McKean–Vlasov control problem in mathematical terms. Then we state and prove a deterministic DPP which embeds the original problem into a sequence of deterministic tail problems for $t = 0, 1, \dots, T$, where both the criterion and the optimisation are restricted onto variables from $t + 1$ onward. Finally, we rewrite the sequence of tail problems into a sequence of single-period problems which we can attack backward in time starting at $t = T$. In this section, we keep the presentation as general as possible; this allows us later to study general controlled processes in finite discrete time.

In Section 3, we study a general class of McKean–Vlasov problems whose criterion has a linear–quadratic (LQ) structure. More precisely, we present a solution technique for the single-period problems obtained by the rewriting in Section 2 for any time t by assuming that the time-t criterion has a particular LQ structure and that the controlled process depends linearly on the control. First, we express the LQ criterion explicitly in terms of the control variable at time t and derive a first order condition for optimality in that control variable. Second, due to the LQ dependence of the criterion on the control variable, the resulting FOC is linear and can be explicitly solved under some extra conditions, which yields an optimal strategy for the one-step problem at time t and shows that the optimal value of the time-t problem preserves the LQ structure with some yet-to-be verified properties of the appearing coefficients. This section ends with a review of the main steps and a discussion of the still missing ingredients.

We turn to solving the MVPS problem in Section 4. First, we formulate the problem in precise terms, show that the MVPS problem fits into the framework of the LQ problems discussed in Section 3, and present a recipe to provide the missing ingredients mentioned in Section 3. All missing conditions depend on the properties of a crucial process $\widetilde{Z}$. Following the recipe, we first show that this process $\widetilde{Z}$ is well defined. Then we translate the structural results for LQ problems in Section 3 into corresponding properties for the MVPS problem. For each one-step problem induced by the MVPS problem, this yields a simplified expression for the criterion, a first order condition (FOC) for optimality in the single-period control, and a solution for that FOC which provides an optimiser for the one-step problem. Based on the missing ingredients for general LQ problems, we verify that most of the missing conditions are satisfied for the MVPS problem and leave one last condition as the extra assumption that the solution to the FOC lies in the proposed space of strategies. Under that extra assumption, we piece together the one-step optimisers over time to obtain in the main result (Theorem 4.11) an affine–quadratic structure for the value function and a recursion for the optimal strategy for the MVPS problem. The next subsection consists of the verification of the just mentioned extra assumption in concrete cases. Finally, we work with the most general space Θ_{MN} of strategies, show that the strategy obtained in the previous subsections is again well defined and optimal in the space Θ_{MN}, and finish the storyline for the MVPS problem in finite discrete time.

In Section 5, we discuss connections to the literature in detail.

2 A deterministic DPP in finite discrete time

2.1 Formulation of a stochastic control problem

We first introduce a general stochastic control problem of mean-field type in finite discrete time and its setup, motivated by the so-called *mean–variance portfolio selection (MVPS)* problem studied in the earlier chapters. This class of problems refers to maximising or minimising the expectation of (deterministic) mean-field type functions over a class of controlled random objects X^θ. The mean-field type functions can depend directly on the law of X^θ, e.g. via a nonlinear function of $E[X^\theta]$.

Let us start with some necessary mathematical basics. Fix a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t=0,1,\dots,T}$. Suppose that there is a family $X^\theta = (X^\theta_t)_{t=0,1,\dots,T}$ of stochastic processes controlled by a family $\theta = (\theta_t)_{t=1,\dots,T}$ of $\mathbb{F}$-predictable processes, meaning that X^θ_t depends on θ only through $\theta_1, \dots, \theta_t$ for $t = 1, \dots, T$. The two processes X^θ and θ take values in two measurable spaces $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ and $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$, where $\mathcal{B}(\cdot)$ denotes the Borel σ-algebra on the underlying space.

For a generic $\mathcal{Y}$-valued random variable Y, we denote by $P_Y = P \circ Y^{-1}$ the distribution on $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ induced by Y. For simplicity, we slightly abuse the notation to denote the space of probability distributions on $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ by $\mathcal{P}(\mathcal{Y})$.

Let f and g be (measurable) functions on $\{0,\dots,T-1\}\times\mathcal{X}\times\mathcal{U}\times\mathcal{P}(\mathcal{X}\times\mathcal{U})$ and $\mathcal{X}\times\mathcal{P}(\mathcal{X})$, respectively, taking values in $\mathbb{R}$. The performance criterion associated to a control θ is of mean-field type given by

$$j(\theta)=E\bigg[\sum_{u=1}^{T} f(u,X_u^\theta,\theta_u,P_{X_u^\theta,\theta_u})+g(X_T^\theta,P_{X_T^\theta})\bigg]. \tag{2.1}$$

Here $P_{X_u^\theta,\theta_u}$ means the joint distribution of (X_u^θ,θ_u) for $u=1,\dots,T$. The criterion (2.1) has a continuous–time analogue studied in Acciaio et al. [2], which has the form

$$j^{\mathrm{c}}(\theta)=E\bigg[\int_0^T f(t,X_t^\theta,\theta_t,P_{X_t^\theta,\theta_t})\,\mathrm{d}t+g(X_T^\theta,P_{X_T^\theta})\bigg]$$

for a continuous-time process $X^\theta=(X_t^\theta)_{t\in[0,T]}$. In finite discrete time, a similar criterion has been considered in Pham and Wei [52]. Although the authors also develop a dynamic programming approach for solving that kind of problem, they assume that the controlled process X^θ is driven by i.i.d. innovations and work at the more abstract level of probability distributions.

Now our task is to

$$\text{maximise } j(\theta) \text{ over all } \theta\in\Theta, \tag{2.2}$$

where Θ is a suitable set of $\mathbb{F}$-predictable processes. Note that the functions f and g in the criterion (2.1) depend explicitly on the laws $P_{X_u^\theta,\theta_{u+1}}$ and $P_{X_T^\theta}$, respectively, and can be nonlinear. This makes the control problem (2.2) time-inconsistent in the sense that the standard dynamic programming principle from stochastic control does not apply. Our goal is to develop a method so that one can still calculate an optimal strategy for problem (2.2) recursively. We end this subsection with two examples.

Example 2.1. 1) Suppose that $f(u,x,y,z)\equiv f(u,x,y)$ and $g(y,z)\equiv g(y)$. Then (2.1) reads

$$E\bigg[\sum_{u=1}^{T} f(u,X_u^\theta,\theta_u)+g(X_T^\theta)\bigg].$$

This recovers the standard stochastic control problem.

2) A specific example of an explicit dependence of f via $P_{X_u^\theta,\theta_u}$ is given by

$$f(u,X_u^\theta,\theta_u,P_{X_u^\theta,\theta_u})=h(\theta_u X_u^\theta-E[\theta_u X_u^\theta])$$

for a function h, e.g. $h(x) = -|x|^p$ for $p \geq 1$ or $h(x) = r(x^+)$ for a "reward function" r.

2.2 A deterministic dynamic programming principle

We state and prove here a deterministic dynamic programming principle (DPP) for problem (2.2). The idea is to introduce and study a tail problem and relevant quantities for each time $t = 0, \dots, T-1$, where both the criterion and the optimisation involve only variables from $t+1$ onward for $t = 0, \dots, T-1$. We then establish the promised deterministic DPP which asserts that the (optimal) value of the tail problem at $t-1$ can be computed in terms of the solution to the tail problem at t. If we can find an optimiser for the tail problem at t, then plugging that into the value of the tail problem at t yields the value of the tail problem at $t-1$. Because the tail problem for $t = 0$ corresponds to our original problem (2.2), the deterministic DPP yields in particular a systematic approach to compute both the value, via the above backward recursion, and the optimal strategy, by piecing together the above one-step optimisers, for the original problem (2.2).

To proceed, we follow Pham and Wei [52] and extend Lemma 3.1 from there by allowing open-loop controls and working without an i.i.d. structure. For $t = 0, 1, \dots, T$, define

$$\Theta(t,\theta) := \{\widetilde{\theta} \in \Theta : \widetilde{\theta} = \theta \text{ on } \rrbracket 0, t \rrbracket \cap \mathbb{N}\}, \tag{2.3}$$

$$j(t,\widetilde{\theta}) := E\Bigg[\sum_{u=t+1}^{T} f(u, X_u^{\widetilde{\theta}}, \widetilde{\theta}_u, P_{X_u^{\widetilde{\theta}}, \widetilde{\theta}_u}) + g(X_T^{\widetilde{\theta}}, P_{X_T^{\widetilde{\theta}}})\Bigg], \tag{2.4}$$

$$v(t,\theta) := \sup_{\widetilde{\theta} \in \Theta(t,\theta)} j(t,\widetilde{\theta}). \tag{2.5}$$

Note that (2.4) is an expectation, not a conditional expectation. Fix $\theta \in \Theta$ and $t \in \{1, \dots, T\}$. The quantity $j(t,\widetilde{\theta})$ on the right-hand side in (2.5) is affected directly by $\widetilde{\theta}$ only via $\widetilde{\theta}_{t+1}, \dots, \widetilde{\theta}_T$; this explains the name "*a tail problem*". For $t = T$, we use (2.5), $\Theta(T,\theta) = \{\theta\}$ by (2.3) and the convention that any sum over an empty set is 0 to get

$$v(T,\theta) = \sup_{\widetilde{\theta} \in \{\theta\}} j(T,\widetilde{\theta}) = j(T,\theta) = E[g(X_T^{\theta}, P_{X_T^{\theta}})]. \tag{2.6}$$

For $t = 0$, we use (2.5), (2.4) and $\Theta(0,\theta) = \Theta$ to deduce that

$$v(0,\theta) = \sup_{\widetilde{\theta}\in\Theta} E\Bigg[\sum_{u=1}^{T} f(u, X_u^{\widetilde{\theta}}, \widetilde{\theta}_u, P_{X_u^{\widetilde{\theta}},\widetilde{\theta}_u}) + g(X_T^{\widetilde{\theta}}, P_{X_T^{\widetilde{\theta}}})\Bigg] =: v_0 \tag{2.7}$$

is the value of the (original) optimisation problem (2.2) and that v_0 is independent of θ.

To state and prove a dynamic programming principle, we collect a few abstract conditions.

Condition 2.2. 1) The random variables $f(t, X_t^\theta, \theta_t, P_{X_t^\theta,\theta_t})$ and $g(X_T^\theta, P_{X_T^\theta})$ are in L^1, and consequently $j(t,\theta) < \infty$, for all $\theta \in \Theta$ and $t = 1,\dots,T$.

2) $\Theta(s,\theta) \supseteq \Theta(t,\theta)$ for any $\theta \in \Theta$ and $s,t \in \{0,\dots,T\}$ with $s \le t$.

Proposition 2.3. *Suppose that Condition* 2.2 *is satisfied. Then for any* $\theta \in \Theta$, *we have*

$$v(t-1,\theta) = \sup_{\widetilde{\theta}\in\Theta(t-1,\theta)} \big(v(t,\widetilde{\theta}) + E[f(t, X_t^{\widetilde{\theta}}, \widetilde{\theta}_t, P_{X_t^{\widetilde{\theta}},\widetilde{\theta}_t})]\big) \tag{2.8}$$

for $t = 1,\dots,T$.

Proof. We argue analogously to Lemma I.3.1. Fix θ and $\widetilde{\theta} \in \Theta(t-1,\theta)$. For the inequality "$\le$" in (2.8), we first use Condition 2.2, 1), (2.4) and then (2.5) with $\widetilde{\theta} \in \Theta(t,\widetilde{\theta})$ to obtain

$$\begin{aligned} j(t-1,\widetilde{\theta}) &= E\Bigg[f(t, X_t^{\widetilde{\theta}}, \widetilde{\theta}_t, P_{X_t^{\widetilde{\theta}},\widetilde{\theta}_t}) \\ &\qquad + \sum_{u=t+1}^{T} f(u, X_u^{\widetilde{\theta}}, \widetilde{\theta}_u, P_{X_u^{\widetilde{\theta}},\widetilde{\theta}_u}) + g(X_T^{\widetilde{\theta}}, P_{X_T^{\widetilde{\theta}}})\Bigg] \\ &\le E[f(t, X_t^{\widetilde{\theta}}, \widetilde{\theta}_t, P_{X_t^{\widetilde{\theta}},\widetilde{\theta}_t})] + v(t,\widetilde{\theta}). \end{aligned}$$

Taking suprema on both sides over $\widetilde{\theta} \in \Theta(t-1,\theta)$ and using the definition (2.5) of $v(t-1,\theta)$ yields "$\le$" in (2.8).

For the inequality "$\ge$", let us denote by $h(t,\widetilde{\theta})$ the quantity inside the supremum on the right-hand side of (2.8). We use $v(t,\widetilde{\theta}) = \sup\{j(t,\overline{\theta}) : \overline{\theta} \in \Theta(t,\widetilde{\theta})\}$ from (2.5) to find a sequence $(\overline{\theta}^n)_{n\in\mathbb{N}}$ in $\Theta(t,\widetilde{\theta})$ such that $j(t,\overline{\theta}^n) \uparrow v(t,\widetilde{\theta})$ as n goes to infinity. Because $\widetilde{\theta} \in \Theta(t-1,\theta)$ and $\overline{\theta}^n \in \Theta(t,\widetilde{\theta}) \subseteq \Theta(t-1,\widetilde{\theta}) = \Theta(t-1,\theta)$ for $n \in \mathbb{N}$ by Condition 2.2, 2), we see that $\overline{\theta}^n$ is in $\Theta(t-1,\theta)$ for $n \in \mathbb{N}$. Moreover, because $X_t^{\overline{\theta}^n}$ depends on $\overline{\theta}^n$ only via $\overline{\theta}_1^n,\dots,\overline{\theta}_t^n$ and $\overline{\theta}^n \in \Theta(t,\widetilde{\theta})$ so that $\overline{\theta}^n$ and $\widetilde{\theta}$

agree up to time t, we also obtain

$$f(t, X_t^{\widetilde{\theta}}, \widetilde{\theta}_t, P_{X_t^{\widetilde{\theta}}, \widetilde{\theta}_t}) = f(t, X_t^{\overline{\theta}^n}, \overline{\theta}_t^n, P_{X_t^{\overline{\theta}^n}, \overline{\theta}_t^n}).$$

Using these facts, (2.4) and (2.5), we get

$$\begin{aligned} h(t, \widetilde{\theta}) &= \lim_{n \to \infty} \big(j(t, \overline{\theta}^n) + E[f(t, X_t^{\widetilde{\theta}}, \widetilde{\theta}_t, P_{X_t^{\widetilde{\theta}}, \widetilde{\theta}_t})] \big) \\ &= \lim_{n \to \infty} \big(j(t, \overline{\theta}^n) + E[f(t, X_t^{\overline{\theta}^n}, \overline{\theta}_t^n, P_{X_t^{\overline{\theta}^n}, \overline{\theta}_t^n})] \big) \\ &= \lim_{n \to \infty} j(t-1, \overline{\theta}^n) \\ &\le v(t-1, \theta). \end{aligned}$$

Since $\widetilde{\theta} \in \Theta(t-1, \theta)$ is arbitrary, we can take the supremum over $\widetilde{\theta}$ to deduce "$\ge$" in (2.8). $\square$

Proposition 2.3 shows that recursively solving the single-period problems given on the right-hand side in (2.8) yields a solution to the original problem (2.2). Because of the one-step nature of the problem (2.8), this represents an important simplification which mirrors an analogue in (I.3.5) in Chapter I. By the definition (2.3) of $\Theta(t-1, \theta)$, each $\widetilde{\theta} \in \Theta(t-1, \theta)$ agrees with θ on $⟧0, t-1⟧ \cap \mathbb{N}$. Then the term inside the supremum in (2.8), denoted by $h(t, \widetilde{\theta})$, only depends on the restriction of θ to $\{1, \dots, t-1\}$, and hence so does the left-hand side $v(t-1, \theta)$. Now fix $\theta \in \Theta$. To compute $v(t-1, \theta)$, we still need to optimise the right-hand side of (2.8) which depends on $\widetilde{\theta} \in \Theta(t-1, \theta)$. Because $\widetilde{\theta}_s = \theta_s$ for $s = 1, \dots, t-1$ and θ is fixed, the value of $h(t, \widetilde{\theta})$ depends on $\widetilde{\theta}$ only through $\widetilde{\theta}_t$. Consequently, it is enough to optimise over random variables $\widetilde{\theta}_t$ rather than over stochastic processes $\widetilde{\theta}$. We then are able to simplify (2.8), as follows.

For $t = 1, \dots, T$, $\theta \in \Theta$ and any $\mathcal{F}_{t-1}$-measurable random variable δ_t, we define

$$\Theta^{[t]}(\theta) := \{\widetilde{\theta}_t : \widetilde{\theta} \in \Theta(t-1, \theta)\}, \quad \theta(t, \delta_t) := (\theta_1, \dots, \theta_{t-1}, \delta_t). \tag{2.9}$$

The last paragraph points out that $v(t, \theta)$ depends on θ only through its first t elements. We can then extend the definition of v by identifying $v(t, \theta(t, \delta_t))$ with any $v(t, \psi)$, where $\psi \in \Theta$ agrees with $\theta(t, \delta_t)$ on $⟧0, t⟧ \cap \mathbb{N}$, i.e., $\psi \in \Theta(t, \theta(t, \delta_t))$. **This extended definition is always assumed** in the rest of this chapter. Since the notation $v(t, \theta(t, \delta_t))$ is clear enough, we do not make a separate definition. Next, because X^θ is controlled by θ, X_t^θ depends on θ only via $\theta_1, \dots, \theta_t$ for

$t = 1, \dots, T$. As a result, we can define

$$X_s^{\theta(t,\delta_t)} := X_s^\theta, \quad s = 0, \dots, t-1 \tag{2.10}$$

and rewrite the deterministic DPP (2.8) as

$$v(t-1,\theta) = \sup_{\delta_t \in \Theta^{[t]}(\theta)} \Big(v\big(t, \theta(t,\delta_t)\big) + E[f(t, X_t^{\theta(t,\delta_t)}, \delta_t, P_{X_t^{\theta(t,\delta_t)},\delta_t})] \Big). \tag{2.11}$$

Here we use the extended definition of v introduced before (2.10). The above is now a sequence of single-period problems that we can solve backward in time, starting with $v(T,\theta) = j(T,\theta) = E[g(X_T^\theta, P_{X_T^\theta})]$. More precisely, we need to compute an optimiser for the one-step problem at t and then plug that in to obtain the value function at $t-1$. Repeating this step yields $v(t,\theta)$ for $t = 1, \dots, T$ as well as an optimal strategy θ^*, by piecing together the one-step optimisers over time.

Remark 2.4. For $\rho \in \mathcal{P}(\mathcal{X} \times \mathcal{U})$ and $\mu \in \mathcal{P}(\mathcal{X})$, we set

$$\begin{aligned} \widehat{f}(t,\rho) &:= \int_{\mathcal{X}\times\mathcal{U}} f(t,x,u,\rho)\,\mathrm{d}\rho(x,u), \quad t = 0, \dots, T-1, \\ \widehat{g}(\mu) &:= \int_{\mathcal{X}} g(x,\mu)\,\mathrm{d}\mu(x). \end{aligned}$$

Using (2.1) and the notations introduced above, we can rewrite problem (2.2) as

$$\text{maximise } \sum_{t=1}^{T} \widehat{f}(t, P_{X_t^\theta,\theta_t}) + \widehat{g}(P_{X_T^\theta}) \text{ over } \theta \in \Theta,$$

and the dynamic programming principle (2.8) reads

$$\begin{aligned} v(T,\theta) &= \widehat{g}(P_{X_T^\theta}), \\ v(t-1,\theta) &= \sup_{\widetilde{\theta} \in \Theta(t-1,\theta)} \big(v(t,\widetilde{\theta}) + \widehat{f}(t, P_{X_t^{\widetilde{\theta}},\widetilde{\theta}_t}) \big) \\ &= \sup_{\delta_t \in \Theta^{[t]}(\theta)} \Big(v\big(t,\theta(t,\delta_t)\big) + \widehat{f}(t, P_{X_t^{\theta(t,\delta_t)},\delta_t}) \Big) \quad \text{for } t = 1, \dots, T. \end{aligned}$$

Note that both $\widehat{f}$ and $\widehat{g}$ are deterministic. If we define $p^\theta = (p_t^\theta)_{t=1,\dots,T}$ by $p_t^\theta = P_{X_t^\theta,\theta_t}$ for $t = 1, \dots, T$, then viewing p^θ as a (measure-valued) controlled process turns (2.2) into a deterministic (measure-valued) control problem admitting *deterministic* state variables (but not control variables). This is the perspective from Pham and Wei [52].

We end this section by presenting a deterministic version of the martingale optimality principle. This is not used in the rest of the chapter; so the reader may safely jump to the next section.

Corollary 2.5. *Suppose that* Condition 2.2 *is satisfied. Then the following statements hold:*

1) *For any $\theta \in \Theta$, the function $t \mapsto v(t,\theta) + E[\sum_{s=1}^{t} f(s, X_s^\theta, \theta_s, P_{X_s^\theta,\theta_s})]$ is decreasing.*

2) *Suppose that $\theta^* \in \Theta$. Then θ^* is optimal for problem* (2.2) *if and only if the function $t \mapsto v(t,\theta^*) + E[\sum_{s=1}^{t} f(s, X_s^{\theta^*}, \theta_s^*, P_{X_s^{\theta^*},\theta_s^*})]$ is constant.*

Proof. 1) Due to the dynamic programming relation (2.8) and because $\theta \in \Theta(t,\theta)$ from the definition (2.3), we obtain

$$\begin{aligned} v(t,\theta) &= \sup_{\widetilde{\theta}\in\Theta(t,\theta)} \Big(v(t+1,\widetilde{\theta}) + E[f(t+1, X_{t+1}^{\widetilde{\theta}}, \widetilde{\theta}_{t+1}, P_{X_{t+1}^{\widetilde{\theta}},\widetilde{\theta}_{t+1}})] \Big) \\ &\geq v(t+1,\theta) + E[f(t+1, X_t^\theta, \theta_{t+1}, P_{X_{t+1}^\theta,\theta_{t+1}})]. \end{aligned}$$

Adding $E[\sum_{s=1}^{t} f(s, X_s^\theta, \theta_s, P_{X_s^\theta,\theta_s})]$ on both sides justifies the claim.

2) For notational convenience, we set

$$h(t,\theta) = v(t,\theta) + E\left[\sum_{s=1}^{t} f(s, X_s^\theta, \theta_s, P_{X_s^\theta,\theta_s})\right], \quad t = 1, \dots, T.$$

Recall from (2.7) that $v_0 := v(0,\theta)$ is the value of the original optimisation problem (2.2) and is independent of θ. This yields $v_0 = h(0,\theta)$ for any θ. So the optimality of θ^* is equivalent to

$$h(0,\theta^*) = v_0 = E\left[\sum_{u=1}^{T} f(u, X_u^{\theta^*}, \theta_u^*, P_{X_u^{\theta^*},\theta_u^*}) + g(X_T^{\theta^*}, P_{X_T^{\theta^*}})\right] = h(T,\theta^*).$$

By part 1), we already know that $h(t,\theta^*)$ is decreasing in t. The above is then equivalent to the constancy of $h(t,\theta^*)$ in t. This completes the proof. $\square$

3 Some results for linear–quadratic (LQ) problems

In this section, we study a specific class of problems like (2.2). Let us start by providing a specific setup. Suppose that $S = (S_t)_{t=0,1,\dots,T}$ is an $\mathbb{R}^d$-valued and $\mathbb{F}$-adapted stochastic process representing the discounted prices of d risky assets in a financial market. This market contains in addition a traded riskless asset whose discounted price at all times is 1. Let Θ be a suitable subspace of the set of all $\mathbb{R}^d$-valued, $\mathbb{F}$-predictable processes. The notation

$$\triangle X_t := X_t - X_{t-1}, \quad t \in \mathbb{N},$$

is used to denote increments of any discrete-time process $X = (X_t)_{t\in\mathbb{N}_0}$. For $\theta \in \Theta$, we define the *gains process* to be

$$G_t(\theta) = \int_0^t \theta_s \, \mathrm{d}S_s = \sum_{s=1}^{t} \theta_s^\top \triangle S_s, \quad t = 0, \dots, T.$$

We use the standard convention that the sum over an empty set is always 0 so that $G_0(\theta) = 0$ for any $\theta \in \Theta$.

In the control problem (2.2) to

$$\text{maximise } E\bigg[\sum_{u=1}^{T} f(u, X_u^\theta, \theta_u, P_{X_u^\theta,\theta_u}) + g(X_T^\theta, P_{X_T^\theta})\bigg]$$

over a set Θ of predictable processes, we let the controlled process X^θ be the gains process $G(\theta)$ and impose that f is identically 0 and g has a special linear–quadratic structure

$$g(x, \mu) = a_T x + b_T x^2 + c_T \bigg(\int_{\mathcal{X}} z \, \mathrm{d}\mu(z)\bigg)^2 + d_T$$

for deterministic quantities a_T, b_T, c_T, d_T. This leads more simply to

$$\text{maximise } E\big[a_T G_T(\theta) + b_T \big(G_T(\theta)\big)^2\big] + c_T \big(E[G_T(\theta)]\big)^2 + d_T \tag{3.1}$$

over a suitable class Θ of predictable processes. **Throughout this section, we take the dimension $d = 1$ for simplicity.**

3.1 Overview of ideas

In this section, we give an overview of the ideas we use here for solving the LQ problem (3.1). We start by introducing a specific set Θ of strategies. Define

$$\Theta_{\mathrm{S}} := \{\theta := (\theta_t)_{t=1,\dots,T} : \theta \text{ is real-valued}, \mathbb{F}\text{-predictable and } \theta_t \triangle S_t \in L^2 \text{ for all } t = 1, \dots, T\}. \tag{3.2}$$

Note that $\theta \in \Theta_{\mathrm{S}}$ is equivalent to saying that $G_t(\theta) \in L^2$ for $t = 0, 1, \dots, T$ or also to $\sup_{0 \le s \le t} |G_s(\theta)| \in L^2$ for $t = 0, 1, \dots, T$. Let us consider the problem (3.1) with $\Theta = \Theta_{\mathrm{S}}$. To avoid abuse of notation, we reserve the notation $v(t, \theta)$ for the value function of the general problem (2.2), whereas we denote by $w(t, \theta)$ the value function of the LQ problem (3.1). The basic idea is to employ the deterministic DPP to compute both w and an optimal strategy recursively.

Lemma 3.1. *Condition* 2.2 *holds for problem* (3.1) *with the choice* $\Theta = \Theta_{\mathrm{S}}$.

Proof. Fix $\theta \in \Theta_{\mathrm{S}}$. For Condition 2.2, 1), because of the explicit LQ expression in (3.1), it is enough to show that $a_T G_T(\theta) + b_T (G_T(\theta))^2 + c_T (E[G_T(\theta)])^2$ is in L^1. This evidently follows from $G_T(\theta) \in L^2$ thanks to the definition (3.2) of Θ_{S}. Next we show Condition 2.2, 2). For $s, t \in \{1, \dots, T\}$ with $s \le t$, we observe from the definition (2.3) that any element in $\Theta_{\mathrm{S}}(t, \theta)$ agrees with θ up to t and thus agrees with θ up to s as well. Because $\Theta_{\mathrm{S}}(s, \theta)$ and $\Theta_{\mathrm{S}}(t, \theta)$ have the same measurability and integrability conditions, we obtain $\Theta_{\mathrm{S}}(s, \theta) \supseteq \Theta_{\mathrm{S}}(t, \theta)$. □

Lemma 3.1 justifies that the deterministic DPP in Proposition 2.3 indeed applies to the problem (3.1) with $\Theta = \Theta_{\mathrm{S}}$. Moreover, the rewriting (2.11) (with v replaced by w everywhere) yields a sequence of one-step problems

$$w(t-1, \theta) = \sup_{\delta_t \in \Theta_{\mathrm{S}}^{[t]}(\theta)} w\big(t, \theta(t, \delta_t)\big), \quad t = 1, \dots, T, \tag{3.3}$$

with $\theta(t, \delta_t) = (\theta_1, \dots, \theta_{t-1}, \delta_t)$ from (2.9), $\Theta_{\mathrm{S}}^{[t]}(\theta)$ given explicitly by

$$\Theta_{\mathrm{S}}^{[t]}(\theta) = \{\delta_t : \delta_t \text{ is real-valued, } \mathcal{F}_{t-1}\text{-measurable and } \delta_t \triangle S_t \in L^2\} \tag{3.4}$$

for $t = 1, \dots, T$, and $w(T, \theta(T, \delta_T))$, thanks to (2.6) and (3.1), given by

$$w(T, \theta(T, \delta_T)) = a_T E\big[G_T\big(\theta(T, \delta_T)\big)\big] + b_T E\Big[\Big(G_T\big(\theta(T, \delta_T)\big)\Big)^2\Big] + c_T \Big(E\big[G_T\big(\theta(T, \delta_T)\big)\big]\Big)^2 + d_T. \tag{3.5}$$

Lemma 3.2. *The set* $\Theta_{\mathrm{S}}^{[t]}(\theta)$ *given in* (3.4) *is a linear space for all* $t = 0, \dots, T$ *and* θ.

Proof. This is immediate from its expression (3.4). □

We first sketch the idea for solving problem (3.3) for $t = T$. Given $\theta \in \Theta_{\mathrm{S}}$, the value function $w(T-1, \theta)$ is obtained by maximising $w(T, \theta(T, \delta_T))$ over δ_T. Note that in the expression (3.5) for $w(T, \theta(T, \delta_T))$, the term $G_T(\theta(T, \delta_T))$ is linear in δ_T. Thus the quantity $w(T, \theta(T, \delta_T))$ is an affine–quadratic expression of $G_T(\theta(T, \delta_T)) = G_{T-1}(\theta) + \delta_T \triangle S_T$, and hence the first order condition (FOC) for the optimisation of $w(T, \theta(T, \delta_T))$ over δ_T is affine. Plugging its solution back in should yield that $w(T-1, \theta)$ is again an affine–quadratic functional of $G_{T-1}(\theta)$, like $w(T, \theta)$ of $G_T(\theta)$, possibly with more complicated coefficients.

To gain more flexibility and yet focus on the one-step nature in (3.3), we next present a solution technique instead of giving a general theory for (3.3). From this perspective, we *assume* that $w(t, \theta(t, \delta_t))$ has the general form

$$w\big(t, \theta(t, \delta_t)\big) = a_t E\big[Z_t G_t\big(\theta(t, \delta_t)\big)\big] + b_t E\Big[Z_t \Big(G_t\big(\theta(t, \delta_t)\big)\Big)^2\Big] + c_t \Big(E\big[Z_t G_t\big(\theta(t, \delta_t)\big)\big]\Big)^2 + d_t, \tag{3.6}$$

$$\text{where } a_t, b_t, c_t, d_t \text{ are deterministic with } b_t \neq 0, \tag{3.7}$$

$$\text{and } Z_t \text{ is bounded, nonnegative and } \mathcal{F}_t\text{-measurable.} \tag{3.8}$$

For $t = T$, this looks like a spurious rewriting due to the terminal condition (3.5). But it turns out that under extra assumptions, the affine–quadratic functional form in (3.6)–(3.8) propagates back from $w(t, \theta)$ to $w(t-1, \theta)$. Therefore, the programme described above for $t = T$ should in principle be applicable to all $t \in \{1, \dots, T\}$ and yield an iteration backward in t. Now we end this subsection with a concrete programme for solving problem (3.3) at a fixed $t = 1, \dots, T$, assuming that $w(t, \theta(t, \delta_t))$ has the structure in (3.6)–(3.8).

Recipe 3.3. $1)_t$ Write $w(t, \theta(t, \delta_t)) =: F_t(\delta_t)$ as a functional of δ_t explicitly and derive a first order condition for optimality.

$2)_t$ Possibly under extra conditions, solve the first order condition to obtain a candidate maximiser $\widehat{\delta}_t$. Then verify its optimality and plug it back into F_t to obtain an explicit formula for $w(t-1, \theta)$.

3.2 Step $1)_t$: Computing $w(t,\theta(t,\delta_t))$ and deriving a first order condition

Let $t \in \{1,\dots,T\}$ be a generic fixed time index. We implement Recipe 3.3, Step $1)_t$. This requires a patient organisation of the terms in $w(t,\theta(t,\delta_t))$, based on which we derive a first order condition (FOC) for optimality.

We start with a natural assumption on the process S.

Assumption 3.4. The increment $\triangle S_t$ is in L^2.

Because of the condition (3.7) and (3.8) on the quantities Z_t, a_t, b_t, c_t, Assumption 3.4 and the definitions (3.2) and (3.4) of Θ_{S} and $\Theta_{\mathrm{S}}^{[t]}(\theta)$, respectively, we have that

$$Y \text{ is bounded for } Y \in \{a_t Z_t, b_t Z_t, c_t Z_t\}, \tag{3.9}$$

$$G_{t-1}(\theta),\ \triangle S_t \text{ and } \delta_t \triangle S_t \text{ are in } L^2. \tag{3.10}$$

These two points are used frequently. In a first step, we express $w(t,\theta(t,\delta_t))$ in terms of $G_{t-1}(\theta)$ (instead of $G_t(\theta)$) and δ_t.

Lemma 3.5. *Suppose that Assumption* 3.4 *is satisfied. For* $\theta \in \Theta_{\mathrm{S}}$, $\delta_t \in \Theta_{\mathrm{S}}^{[t]}(\theta)$ *and* w *given by* (3.6)–(3.8)*, we then have*

$$\begin{aligned} w\big(t,\theta(t,\delta_t)\big) = {} & E\big[a_t Z_t G_{t-1}(\theta) + b_t Z_t \big(G_{t-1}(\theta)\big)^2\big] + c_t\big(E[Z_t G_{t-1}(\theta)]\big)^2 + d_t \\ & + r_t(\delta_t), \end{aligned} \tag{3.11}$$

where

$$\begin{aligned} r_t(\delta_t) = {} & E[a_t Z_t \delta_t \triangle S_t + 2 b_t Z_t G_{t-1}(\theta)\delta_t \triangle S_t + b_t Z_t (\delta_t \triangle S_t)^2] \\ & + 2 c_t E[Z_t G_{t-1}(\theta)] E[Z_t \delta_t \triangle S_t] + c_t (E[Z_t \delta_t \triangle S_t])^2. \end{aligned} \tag{3.12}$$

Proof. We first use the definition (3.4) of $\theta(t,\delta_t)$ and the expression of the gains process $G(\theta)$ to obtain

$$G_t\big(\theta(t,\delta_t)\big) = G_{t-1}(\theta) + \delta_t \triangle S_t.$$

Inserting this identity into the formula (3.6) for $w(t,\theta(t,\delta_t))$ and squaring out yields

$$\begin{aligned}
w\big(t,\theta(t,\delta_t)\big) = E\big[&a_tZ_tG_{t-1}(\theta) + a_tZ_t\delta_t\triangle S_t \\
&+ b_tZ_t\big(G_{t-1}(\theta)\big)^2 + 2b_tZ_tG_{t-1}(\theta)\delta_t\triangle S_t + b_tZ_t(\delta_t\triangle S_t)^2\big] \\
&+ c_t\big(E[Z_tG_{t-1}(\theta)]\big)^2 + 2c_tE[Z_tG_{t-1}(\theta)]E[Z_t\delta_t\triangle S_t] \\
&+ c_t(E[Z_t\delta_t\triangle S_t])^2 + d_t. \qquad (3.13)
\end{aligned}$$

Thanks to (3.9), (3.10) and the Cauchy–Schwarz inequality, we obtain that all random variables inside the expectations in (3.13) are in L^1. Therefore, we can define $r_t(\delta_t)$ as in (3.12) and group the terms in (3.13) according to the dependence on δ_t to get (3.11). □

In view of the decomposition in (3.11), maximisation of $\delta_t \mapsto F_t(\delta_t)$ reduces to that of $\delta_t \mapsto r_t(\delta_t)$. Based on this observation, we derive an FOC for maximisation of the latter.

Lemma 3.6. *Suppose that Assumption* 3.4 *is satisfied and* (3.6)–(3.8) *hold. If* $\theta \in \Theta_{\mathrm{S}}$ *and* r_t *is given by* (3.12)*, then any maximiser* $\widehat{\delta}_t$ *for* $\delta_t \mapsto r_t(\delta_t)$ *is a solution to the linear equation*

$$\begin{aligned}
\delta_t = -&\frac{E[Z_t\triangle S_t|\mathcal{F}_{t-1}]}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]} \\
&\times\left(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t}E\big[Z_t\big(G_{t-1}(\theta) + \delta_t\triangle S_t\big)\big]\right). \qquad (3.14)
\end{aligned}$$

Proof. Let $\eta_t \in \Theta_{\mathrm{S}}^{[t]}(\theta)$ be arbitrary. Using the optimality of $\widehat{\delta}_t$ and (3.12) and expanding the terms, we compute

$$\begin{aligned}
0 &\ge r_t(\widehat{\delta}_t + \eta_t) - r_t(\widehat{\delta}_t) \\
&= E[a_tZ_t\eta_t\triangle S_t + 2b_tZ_tG_{t-1}(\theta)\eta_t\triangle S_t + 2b_tZ_t\widehat{\delta}_t\eta_t(\triangle S_t)^2 + b_tZ_t(\eta_t\triangle S_t)^2] \\
&\quad + 2c_tE[Z_tG_{t-1}(\theta)]E[Z_t\eta_t\triangle S_t] + 2c_tE[Z_t\widehat{\delta}_t\triangle S_t]E[Z_t\eta_t\triangle S_t] + c_t(E[Z_t\eta_t\triangle S_t])^2.
\end{aligned}$$

Because $\Theta_{\mathrm{S}}^{[t]}(\theta)$ is a linear space by Lemma 3.2, the random variable $\pm\frac{1}{n}\eta_t$ is in $\Theta_{\mathrm{S}}^{[t]}(\theta)$ for every $n \in \mathbb{N}$. So replacing η_t by $\pm\frac{1}{n}\eta_t$ in the above display, multiplying by n and using the dominated convergence theorem (which applies because of

(3.9), (3.10) and the Cauchy–Schwarz inequality) gives

$$\begin{aligned}0 &= E[a_t Z_t \eta_t \triangle S_t + 2b_t Z_t G_{t-1}(\theta)\eta_t \triangle S_t + 2b_t Z_t \widehat{\delta}_t \eta_t (\triangle S_t)^2] \\ &\quad + 2c_t E[Z_t G_{t-1}(\theta)] E[Z_t \eta_t \triangle S_t] + 2c_t E[Z_t \widehat{\delta}_t \triangle S_t] E[Z_t \eta_t \triangle S_t] \\ &= E\Big[Z_t \eta_t \triangle S_t \Big(a_t + 2b_t G_{t-1}(\theta) + 2b_t \widehat{\delta}_t \triangle S_t \\ &\qquad + 2c_t E\big[Z_t\big(G_{t-1}(\theta) + \widehat{\delta}_t \triangle S_t\big)\big]\Big)\Big]. \end{aligned} \tag{3.15}$$

The second equality in the above display uses that c_t is deterministic by (3.7). Because (3.15) holds for all $\eta_t \in \Theta_S^{[t]}(\theta)$, we can take $\eta_t = \mathbb{1}_H$ for $H \in \mathcal{F}_{t-1}$ (this uses that $\triangle S_t \in L^2$ by Assumption 3.4) to obtain

$$0 = E\Big[Z_t \triangle S_t \Big(a_t + 2b_t G_{t-1}(\theta) + 2b_t \widehat{\delta}_t \triangle S_t + 2c_t E\big[Z_t\big(G_{t-1}(\theta) + \widehat{\delta}_t \triangle S_t\big)\big]\Big)\Big| \mathcal{F}_{t-1}\Big].$$

Using (3.9), (3.10) and the Cauchy–Schwarz inequality, we obtain that each individual summand inside the above conditional expectation is in L^1 and has a well-defined conditional expectation. So moving $E[2b_t \widehat{\delta}_t Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]$ to the left, taking out the $\mathcal{F}_{t-1}$-measurable $\widehat{\delta}_t$ and dividing by $2b_t$ thanks to $b_t \neq 0$ by (3.7) yields

$$\begin{aligned}-E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]\widehat{\delta}_t &= E[Z_t \triangle S_t | \mathcal{F}_{t-1}] \\ &\quad \times \left(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t} E\big[Z_t\big(G_{t-1}(\theta) + \widehat{\delta}_t \triangle S_t\big)\big]\right).\end{aligned}$$

By (3.8), Z_t is nonnegative and bounded, and the Cauchy–Schwarz inequality yields $(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2 \leq E[Z_t|\mathcal{F}_{t-1}] E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]$. These two points in particular imply $\{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}] = 0\} \subseteq \{E[Z_t \triangle S_t|\mathcal{F}_{t-1}] = 0\}$. Hence we can use the convention $\frac{0}{0} = 0$ and divide by the term $-E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]$ on both sides above to obtain that $\widehat{\delta}_t$ is a solution to (3.14). □

3.3 Step $2)_t$: Maximising $\delta_t \mapsto w(t, \theta(t, \delta_t)) =: F_t(\delta_t)$ by solving a linear equation

In this subsection, we implement Recipe 3.3, $2)_t$. In view of Lemma 3.6, maximising $\delta_t \mapsto r_t(\delta_t)$ amounts to solving the linear equation (3.14). Because the right-hand side of (3.14) involves an expectation of (a function of) the unknown δ_t, we multiply on both sides by $Z_t \triangle S_t$ and take expectations to get an auxiliary equation now for the unknown $E[Z_t \delta_t \triangle S_t]$. Solving that equation and plugging its solution back in (3.14) gives a candidate for a solution to the original equation (3.14). However, carrying out the computation rigorously needs extra conditions.

Lemma 3.7. *Suppose that Assumption* 3.4 *is satisfied and* (3.6)–(3.8) *hold.*

1) *If*

$$\frac{E[Z_t \triangle S_t | \mathcal{F}_{t-1}]}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} Z_t \triangle S_t \in L^2, \tag{3.16}$$

$$1 + \frac{c_t}{b_t} E\left[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right] \neq 0, \tag{3.17}$$

then a solution $\widehat{\delta}_t$ *to* (3.14) *exists and satisfies*

$$\begin{aligned} E[Z_t \widehat{\delta}_t \triangle S_t] &= -\left(1 + \frac{c_t}{b_t} E\left[\frac{E[Z_t \triangle S_t | \mathcal{F}_{t-1}]}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} Z_t \triangle S_t\right]\right)^{-1} \\ &\quad \times E\left[\frac{E[Z_t \triangle S_t | \mathcal{F}_{t-1}]}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} Z_t \triangle S_t \right. \\ &\quad\quad \left. \times \left(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t} E[Z_t G_{t-1}(\theta)]\right)\right] \\ &= -\left(1 + \frac{c_t}{b_t} E\left[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right]\right)^{-1} \\ &\quad \times E\left[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \left(\frac{a_t}{2b_t} + G_{t-1}(\theta) \right.\right. \\ &\quad\quad \left.\left. + \frac{c_t}{b_t} E[Z_t G_{t-1}(\theta)]\right)\right]. \end{aligned} \tag{3.18}$$

Explicitly, we have

$$\widehat{\delta}_t = -\frac{E[Z_t \triangle S_t | \mathcal{F}_{t-1}]}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \left(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t}\big(E[Z_t G_{t-1}(\theta)] + e_t\big)\right), \tag{3.19}$$

where e_t *is given by the right-hand side of the last equality in* (3.18).

2) *If, in addition,*

$$b_t + c_t E[Z_t] \leq 0 \text{ and the solution } \widehat{\delta}_t \text{ to (3.14) is in } \Theta_{\mathrm{S}}^{[t]}(\theta), \tag{3.20}$$

then $\widehat{\delta}_t$ *maximises* $\delta_t \mapsto r_t(\delta_t)$.

3) *If the conditions* (3.16), (3.17) *and* (3.20) *are satisfied, then the resulting value function* $w(t-1, \theta) = w(t, \theta(t, \widehat{\delta}_t))$ *from* (3.3) *has a similar structure as*

$w(t,\theta(t,\delta_t))$ *given in* (3.6) *in the sense that*

$$\begin{aligned} w(t-1,\theta) = E\big[a_{t-1}Z_{t-1}G_{t-1}(\theta) + b_{t-1}Z_{t-1}\big(G_{t-1}(\theta)\big)^2\big] \\ + c_{t-1}\big(E[Z_{t-1}G_{t-1}(\theta)]\big)^2 + d_{t-1}, \end{aligned} \tag{3.21}$$

where

$$Z_{t-1} = E[Z_t|\mathcal{F}_{t-1}] - \frac{(E[Z_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}, \tag{3.22}$$

$$a_{t-1} = a_t\left(1 + \frac{c_t}{b_t}E\left[\frac{(E[Z_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1}, \tag{3.23}$$

$$b_{t-1} = b_t, \tag{3.24}$$

$$c_{t-1} = c_t\left(1 + \frac{c_t}{b_t}E\left[\frac{(E[Z_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1}, \tag{3.25}$$

$$\begin{aligned} d_{t-1} = d_t - \left(1 + \frac{c_t}{b_t}E\left[\frac{(E[Z_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1}, \\ \times \frac{a_t^2}{4b_t}E\left[\frac{(E[Z_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]. \end{aligned} \tag{3.26}$$

Proof. 1) First we show that the right-hand side e_t of (3.18) is well defined. Indeed, the denominator in (3.18) is nonzero due to the condition (3.17). Also, the right-hand side in the first equality of (3.18) is finite thanks to (3.16) and because $G_{t-1}(\theta)$ is in L^2 by (3.10). This also allows us to take conditional expectations with respect to $\mathcal{F}_{t-1}$ to obtain the second equality of (3.18). So we can construct $\widehat{\delta}_t$ by inserting the right-hand side e_t of (3.18) into (3.14) to replace the term $E[Z_t\delta_t\triangle S_t]$. This gives the explicit expression (3.19) for $\widehat{\delta}_t$. To verify that the constructed $\widehat{\delta}_t$ indeed solves (3.14), we multiply both sides of (3.19) by $Z_t\triangle S_t$. Then (3.16) again allows us to take expectations on both sides and organise the terms to obtain again (3.18). Therefore we can replace the term e_t in (3.19) by $E[Z_t\widehat{\delta}_t\triangle S_t]$, which shows that $\widehat{\delta}_t$ solves (3.14).

2) Let us now argue that $r_t(\widehat{\delta}_t) \geq r_t(\delta_t)$ for all $\delta_t \in \Theta_{\mathrm{S}}^{[t]}(\theta)$. To that end, we write $\delta_t = \delta_t - \widehat{\delta}_t + \widehat{\delta}_t$ and note by (3.20) that $\delta_t - \widehat{\delta}_t$ is in $\Theta_{\mathrm{S}}^{[t]}(\theta)$ like δ_t and $\widehat{\delta}_t$ due to the linearity of $\Theta_{\mathrm{S}}^{[t]}(\theta)$ from Lemma 3.2. So we insert this rewriting into the expression (3.12) for $r_t(\delta_t)$, square out and reorder the terms as expressions involving $\delta_t - \widehat{\delta}_t$ and $\widehat{\delta}_t$ only and use the first equality in the FOC (3.15) (with

$\eta_t = \delta_t - \widehat{\delta}_t$) to obtain

$$\begin{aligned}
r_t(\delta_t) &= r_t(\delta_t - \widehat{\delta}_t + \widehat{\delta}_t) \\
&= E\big[a_t Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t + a_t Z_t \widehat{\delta}_t \triangle S_t + 2b_t Z_t G_{t-1}(\theta)(\delta_t - \widehat{\delta}_t)\triangle S_t \\
&\qquad + 2b_t Z_t G_{t-1}(\theta)\widehat{\delta}_t \triangle S_t + b_t Z_t\big((\delta_t - \widehat{\delta}_t)\triangle S_t\big)^2 + 2b_t Z_t(\delta_t - \widehat{\delta}_t)\widehat{\delta}_t(\triangle S_t)^2 \\
&\qquad + b_t Z_t(\widehat{\delta}_t \triangle S_t)^2\big] + 2c_t E[Z_t G_{t-1}(\theta)]E[Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t] \\
&\quad + 2c_t E[Z_t G_{t-1}(\theta)]E[Z_t \widehat{\delta}_t \triangle S_t] + c_t\big(E[Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t]\big)^2 \\
&\quad + 2c_t E[Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t]E[Z_t \widehat{\delta}_t \triangle S_t] + c_t(E[Z_t \widehat{\delta}_t \triangle S_t])^2 \\
&= E\big[a_t Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t + 2b_t Z_t G_{t-1}(\theta)(\delta_t - \widehat{\delta}_t)\triangle S_t \\
&\qquad + 2b_t Z_t(\delta_t - \widehat{\delta}_t)\widehat{\delta}_t(\triangle S_t)^2\big] + 2c_t E[Z_t G_{t-1}(\theta)]E[Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t] \\
&\quad + 2c_t E[Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t]E[Z_t \widehat{\delta}_t \triangle S_t] + E\big[b_t Z_t\big((\delta_t - \widehat{\delta}_t)\triangle S_t\big)^2\big] \\
&\quad + c_t\big(E[Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t]\big)^2 + r_t(\widehat{\delta}_t) \\
&= E\big[b_t Z_t\big((\delta_t - \widehat{\delta}_t)\triangle S_t\big)^2\big] + c_t\big(E[Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t]\big)^2 + r_t(\widehat{\delta}_t) \\
&\le (b_t + c_t E[Z_t])E\big[Z_t\big((\delta_t - \widehat{\delta}_t)\triangle S_t\big)^2\big] + r_t(\widehat{\delta}_t) \\
&\le r_t(\widehat{\delta}_t).
\end{aligned}$$

The third equality also uses that the terms in the second equality involving only $\widehat{\delta}_t$ indeed sum to $r_t(\widehat{\delta}_t)$ by the expression (3.12) for r_t. The second-to-last line uses the Cauchy–Schwarz inequality applied to the term

$$\big(E[Z_t(\delta_t - \widehat{\delta}_t)\triangle S_t]\big)^2 = \big(E[\sqrt{Z}_t\sqrt{Z}_t(\delta_t - \widehat{\delta}_t)\triangle S_t]\big)^2$$

and the non-randomness of b_t, c_t from (3.7). The last line uses $b_t + c_t E[Z_t] \le 0$ from (3.20). Because $\widehat{\delta}_t \in \Theta_{\mathrm{S}}^{[t]}(\theta)$ by (3.20) again, this verifies the optimality of $\widehat{\delta}_t$ for $\delta_t \mapsto r_t(\delta_t)$.

3) By (3.3) and the optimality of $\widehat{\delta}_t$ from 2), we have $w(t-1,\theta) = w(t, \theta(t, \widehat{\delta}_t))$. So we prove that the latter quantity has the functional form in (3.21), where the quantities $Z_{t-1}, a_{t-1}, b_{t-1}, c_{t-1}, d_{t-1}$ are given in (3.22)–(3.26). First we use the FOC (3.15) to simplify the expression for $r_t(\widehat{\delta}_t)$. Indeed, we set $\eta_t = \widehat{\delta}_t$ in the FOC (3.15) to obtain

$$\begin{aligned}
&E\big[\big(a_t + 2b_t G_{t-1}(\theta) + 2c_t E[Z_t G_{t-1}(\theta)]\big) Z_t \widehat{\delta}_t \triangle S_t\big] \\
&= -E[2b_t Z_t(\widehat{\delta}_t \triangle S_t)^2] - 2c_t(E[Z_t \widehat{\delta}_t \triangle S_t])^2.
\end{aligned}$$

Then we organise the terms and factor out $Z_t\widehat{\delta}_t\triangle S_t$ in the expression (3.12) for

$r_t(\widehat{\delta}_t)$ and use the above identity to replace the quadratic terms to get

$$\begin{aligned} r_t(\widehat{\delta}_t) &= E\big[\big(a_t + 2b_t G_{t-1}(\theta) + 2c_t E[Z_t G_{t-1}(\theta)]\big) Z_t \widehat{\delta}_t \triangle S_t\big] \\ &\quad + E[b_t Z_t (\widehat{\delta}_t \triangle S_t)^2] + c_t (E[Z_t \widehat{\delta}_t \triangle S_t])^2 \\ &= E\bigg[\bigg(\frac{1}{2}a_t + b_t G_{t-1}(\theta) + c_t E[Z_t G_{t-1}(\theta)]\bigg) Z_t \widehat{\delta}_t \triangle S_t\bigg] \\ &= E\bigg[b_t\bigg(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t} E[Z_t G_{t-1}(\theta)]\bigg)\widehat{\delta}_t E[Z_t \triangle S_t | \mathcal{F}_{t-1}]\bigg]. \end{aligned}$$

The last step uses $b_t \neq 0$ by (3.7) and iterative conditioning on $\mathcal{F}_{t-1}$ thanks to $\widehat{\delta}_t \triangle S_t \in L^2$, $G_{t-1}(\theta) \in L^2$, and the boundedness of Z_t from (3.20), (3.10) and (3.8), respectively. Using the explicit expressions (3.19) for $\widehat{\delta}_t$ and then (3.18) for the deterministic quantity e_t yields

$$\begin{aligned} r_t(\widehat{\delta}_t) &= E\bigg[-b_t \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\bigg(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t} E[Z_t G_{t-1}(\theta)]\bigg)^2\bigg] \\ &\quad - E\bigg[c_t \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\bigg(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t} E[Z_t G_{t-1}(\theta)]\bigg) e_t\bigg] \\ &= E\bigg[-b_t \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\bigg(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t} E[Z_t G_{t-1}(\theta)]\bigg)^2\bigg] \\ &\quad + c_t \bigg(1 + \frac{c_t}{b_t} E\bigg[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\bigg]\bigg)^{-1} \\ &\quad \times \bigg(E\bigg[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\bigg(\frac{a_t}{2b_t} + G_{t-1}(\theta) + \frac{c_t}{b_t} E[Z_t G_{t-1}(\theta)]\bigg)\bigg]\bigg)^2. \end{aligned}$$

Note from the above display that $r_t(\widehat{\delta}_t)$ is a sum of two squares. To make the expressions lighter, we introduce the shorthand notations

$$x = E[Z_t G_{t-1}(\theta)], \tag{3.27}$$

$$y = E\bigg[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} G_{t-1}(\theta)\bigg], \tag{3.28}$$

$$z = E\bigg[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\bigg]. \tag{3.29}$$

Using the notations (3.27)–(3.29) to square out the terms in $r_t(\widehat{\delta}_t)$ yields

$$\begin{aligned}
r_t(\widehat{\delta}_t) = &-\frac{a_t^2}{4b_t}z - E\left[b_t\frac{(E[Z_t\triangle S_t|\mathcal{F}_{t-1}])^2}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\big(G_{t-1}(\theta)\big)^2\right] - \frac{c_t^2}{b_t}zx^2 - a_t y - \frac{a_t c_t}{b_t}zx \\
&- 2c_t yx \\
&+ c_t\left(1+\frac{c_t}{b_t}z\right)^{-1}\left(\frac{a_t^2}{4b_t^2}z^2 + y^2 + \frac{c_t^2}{b_t^2}z^2x^2 + \frac{a_t}{b_t}zy + \frac{a_t c_t}{b_t^2}z^2 x + \frac{2c_t}{b_t}yzx\right).
\end{aligned}$$

We now use the explicit expression (3.11) for $w(t,\theta(t,\widehat{\delta}_t))$, then the shorthand notations (3.27)–(3.29) again and finally the above display to obtain

$$\begin{aligned}
&w\big(t,\theta(t,\widehat{\delta}_t)\big) \\
&= E\big[a_t Z_t G_{t-1}(\theta) + b_t Z_t\big(G_{t-1}(\theta)\big)^2\big] + c_t\big(E[Z_t G_{t-1}(\theta)]\big)^2 + d_t \\
&\quad + r_t(\widehat{\delta}_t) \\
&= a_t x + E\big[b_t Z_t\big(G_{t-1}(\theta)\big)^2\big] + c_t x^2 + d_t - \frac{a_t^2}{4b_t}z \\
&\quad - E\left[b_t\frac{(E[Z_t\triangle S_t|\mathcal{F}_{t-1}])^2}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\big(G_{t-1}(\theta)\big)^2\right] - \frac{c_t^2}{b_t}zx^2 - a_t y - \frac{a_t c_t}{b_t}zx - 2c_t yx \\
&\quad + c_t\left(1+\frac{c_t}{b_t}z\right)^{-1}\left(\frac{a_t^2}{4b_t^2}z^2 + y^2 + \frac{c_t^2}{b_t^2}z^2x^2 + \frac{a_t}{b_t}zy + \frac{a_t c_t}{b_t^2}z^2 x + \frac{2c_t}{b_t}yzx\right) \\
&= a_t x - a_t y - \frac{a_t c_t}{b_t}zx + c_t\left(1+\frac{c_t}{b_t}z\right)^{-1}\left(\frac{a_t}{b_t}zy + \frac{a_t c_t}{b_t^2}z^2 x\right) \\
&\quad + E\big[b_t Z_t\big(G_{t-1}(\theta)\big)^2\big] - E\left[b_t\frac{(E[Z_t\triangle S_t|\mathcal{F}_{t-1}])^2}{E[Z_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\big(G_{t-1}(\theta)\big)^2\right] \\
&\quad + c_t x^2 - \frac{c_t^2}{b_t}zx^2 - 2c_t yx + c_t\left(1+\frac{c_t}{b_t}z\right)^{-1}\left(y^2 + \frac{c_t^2}{b_t^2}z^2x^2 + \frac{2c_t}{b_t}yzx\right) \\
&\quad + d_t - \frac{a_t^2}{4b_t}z + c_t\left(1+\frac{c_t}{b_t}z\right)^{-1}\frac{a_t^2}{4b_t^2}z^2 \\
&= w_a + w_b + w_c + w_d
\end{aligned} \tag{3.30}$$

where w_a, w_b, w_c, w_d each denotes one of the four lines in the third equality, respectively. In the third equality, we use the fact from (3.27)–(3.29) that both x and y are linear in $G_{t-1}(\theta)$ and z does not depend on $G_{t-1}(\theta)$ to reorder the terms. More precisely, w_a contains terms that are linear in $G_{t-1}(\theta)$, w_b contains terms that are linear in $(G_{t-1}(\theta))^2$, w_c contains terms that involve products of expectations of linear functions of $G_{t-1}(\theta)$, and w_d does not depend on $G_{t-1}(\theta)$.

Now we simplify the terms in (3.30) for w_a, factor out the term $(1+\frac{c_t}{b_t}z)^{-1}$ and reinstate the explicit expressions (3.27)–(3.29) for x, y, z to obtain explicitly that

$$\begin{aligned}
w_a &= a_t x - a_t y - \frac{a_t c_t^2}{b_t} zx + c_t \left(1 + \frac{c_t}{b_t} z\right)^{-1} \left(\frac{a_t}{b_t} zy + \frac{a_t c_t}{b_t^2} z^2 x\right) \\
&= a_t x \left(1 - \frac{c_t}{b_t} z + \left(1 + \frac{c_t}{b_t} z\right)^{-1} \left(\frac{c_t}{b_t} z\right)^2\right) - a_t y \left(1 - \left(1 + \frac{c_t}{b_t} z\right)^{-1} \frac{c_t}{b_t} z\right) \\
&= \left(1 + \frac{c_t}{b_t} z\right)^{-1} \left(a_t x \left(1 - \frac{c_t^2}{b_t^2} z^2 + \frac{c_t^2}{b_t^2} z^2\right) - a_t y \left(1 + \frac{c_t}{b_t} z - \frac{c_t}{b_t} z\right)\right) \\
&= a_t \left(1 + \frac{c_t}{b_t} z\right)^{-1} (x - y) \\
&= E\Bigg[a_t \left(1 + \frac{c_t}{b_t} E\Big[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\Big]\right)^{-1} \\
&\qquad \times \left(Z_t - \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right) G_{t-1}(\theta)\Bigg] \\
&= E\Bigg[a_t \left(1 + \frac{c_t}{b_t} E\Big[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\Big]\right)^{-1} \\
&\qquad \times \left(E[Z_t | \mathcal{F}_{t-1}] - \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right) G_{t-1}(\theta)\Bigg]. \qquad (3.31)
\end{aligned}$$

The iterative conditioning in (3.31) is allowed thanks to the boundedness of Z_t from (3.8) and $G_{t-1}(\theta) \in L^2$ from (3.10). Similarly, we obtain

$$\begin{aligned}
w_b &= E\big[b_t Z_t \big(G_{t-1}(\theta)\big)^2\big] - E\left[b_t \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \big(G_{t-1}(\theta)\big)^2\right] \\
&= E\left[b_t \left(E[Z_t | \mathcal{F}_{t-1}] - \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right) \big(G_{t-1}(\theta)\big)^2\right]. \qquad (3.32)
\end{aligned}$$

In w_c, we factor out the term $(1 + \frac{c_t}{b_t} z)^{-1} c_t$, reinstate the shorthand notations x, y, z from (3.27)–(3.29) and use the tower property thanks to the boundedness of Z_t and $G_{t-1}(\theta) \in L^2$ by (3.8) and (3.10) to obtain

$$
\begin{aligned}
w_c &= c_t x^2 - \frac{c_t^2}{b_t} z x^2 - 2c_t y x + c_t \left(1 + \frac{c_t}{b_t} z\right)^{-1} \left(y^2 + \frac{c_t^2}{b_t^2} z^2 x^2 + \frac{2c_t}{b_t} yzx\right) \\
&= \left(1 + \frac{c_t}{b_t} z\right)^{-1} c_t \\
&\quad \times \left(x^2 + \frac{c_t}{b_t} x^2 z - \frac{c_t}{b_t} z x^2 - \frac{c_t^2}{b_t^2} z^2 x^2 - 2yx - 2\frac{c_t}{b_t} yxz \right. \\
&\qquad \left. + y^2 + \frac{c_t^2}{b_t^2} z^2 x^2 + 2\frac{c_t}{b_t} yzx\right) \\
&= \left(1 + \frac{c_t}{b_t} z\right)^{-1} c_t (x-y)^2 \\
&= \left(1 + \frac{c_t}{b_t} E\Big[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\Big]\right)^{-1} c_t \\
&\quad \times \left(E\Big[\Big(Z_t - \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\Big) G_{t-1}(\theta)\Big]\right)^2 \\
&= \left(1 + \frac{c_t}{b_t} E\Big[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\Big]\right)^{-1} c_t \\
&\quad \times \left(E\Big[\Big(E[Z_t | \mathcal{F}_{t-1}] - \frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\Big) G_{t-1}(\theta)\Big]\right)^2. \qquad (3.33)
\end{aligned}
$$

Finally, we simplify the terms in w_d and insert the explicit expression (3.29) for z to obtain

$$
\begin{aligned}
w_d &= d_t - \frac{a_t^2}{4b_t} z + c_t \left(1 + \frac{c_t}{b_t} z\right)^{-1} \frac{a_t^2}{4b_t^2} z^2 \\
&= d_t + \left(1 + \frac{c_t}{b_t} z\right)^{-1} \left(-\frac{a_t^2}{4b_t} z - \frac{a_t^2}{4b_t} \frac{c_t}{b_t} z^2 + c_t \frac{a_t^2}{4b_t^2} z^2\right) \\
&= d_t - \left(1 + \frac{c_t}{b_t} z\right)^{-1} \frac{a_t^2}{4b_t} z \\
&= d_t - \left(1 + \frac{c_t}{b_t} E\left[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right]\right)^{-1} \\
&\qquad \times \frac{a_t^2}{4b_t} E\left[\frac{(E[Z_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[Z_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right]. \qquad (3.34)
\end{aligned}
$$

Inserting the expressions (3.31)–(3.34) for w_a, w_b, w_c, w_d into (3.30) yields the expressions (3.21)–(3.26) for $w(t-1, \theta), Z_{t-1}, a_{t-1}, b_{t-1}, c_{t-1}, d_{t-1}$, respectively. □

3.4 Summary and the missing ingredients

Lemmas 3.5–3.7 in the previous subsection finish the implementation of Recipe 3.3. We now summarise the key aspects.

The problem of interest is to

$$\text{maximise } E\big[a_T G_T(\theta) + b_T\big(G_T(\theta)\big)^2\big] + c_T\big(E[G_T(\theta)]\big)^2 + d_T \text{ over } \theta \in \Theta_{\mathrm{S}},$$

where a_T, b_T, c_T, d_T are deterministic and Θ_{S} is the space of all $\mathbb{F}$-predictable processes such that the associated gains processes have square-integrable increments. The deterministic DPP in Proposition 2.3 and its rewriting (2.11) yield that it is sufficient to solve the sequence of single-step problems (3.3)–(3.5). We then develop a technique in Recipe 3.3 to optimise a general affine–quadratic objective

$$\begin{aligned}\delta_t \mapsto w\big(t, \theta(t,\delta_t)\big) = E\Big[a_t Z_t G_t\big(\theta(t,\delta_t)\big) + b_t Z_t \Big(G_t\big(\theta(t,\delta_t)\big)\Big)^2\Big]\\ + c_t\Big(E\big[Z_t G_t\big(\theta(t,\delta_t)\big)\big]\Big)^2 + d_t,\end{aligned}$$

where a_t, b_t, c_t, d_t are deterministic with $b_t \neq 0$,

and Z_t is bounded, nonnegative and $\mathcal{F}_t$-measurable,

with respect to a single (random) variable δ_t. Indeed, formal calculus of variations yields a first order condition (3.14) which can be solved explicitly as in (3.19) and gives an affine–quadratic objective $w(t-1,\theta)$ with coefficients $Z_{t-1}, a_{t-1}, b_{t-1}, c_{t-1}, d_{t-1}$ completely analogous as in the above display. Moreover, these coefficients satisfy recursive relations given in (3.22)–(3.26). Therefore, Recipe 3.3 or more precisely Lemmas 3.5–3.7 can be viewed provisionally as a solution technique for our problem (3.3) by iterating Recipe 3.3 backward in time.

However, there are several missing ingredients. The implementation of Recipe 3.3, Step $2)_t$ requires the conditions (3.16), (3.17) and (3.20) for $t = 1, \dots, T$. In addition, the coefficients $Z_{t-1}, a_{t-1}, b_{t-1}, c_{t-1}, d_{t-1}$ in $w(t-1,\theta)$ need to satisfy analogous properties as in (3.7) and (3.8), namely that

$$a_{t-1}, b_{t-1}, c_{t-1}, d_{t-1} \text{ are deterministic with } b_t \neq 0, \tag{3.35}$$

$$\text{and } Z_{t-1} \text{ is bounded, nonnegative and } \mathcal{F}_{t-1}\text{-measurable.} \tag{3.36}$$

These conditions and properties should be verified in principle by backward induction and so depend specifically on the terminal conditions Z_T, a_T, b_T, c_T, d_T.

Since the present section focuses on demonstrating the methodology, we study the verification of these conditions later in a concrete application.

4 Application to the MVPS problem

4.1 Problem formulation

A classic example of problem (2.2) and (3.1) is the *mean–variance portfolio selection (MVPS)* problem. We briefly present this problem below and refer to Section I.2.1 for a detailed exposition.

Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t=0,1,\dots,T}$. Suppose that $S = (S_t)_{t=0,1,\dots,T}$ is an $\mathbb{F}$-adapted and $\mathbb{R}^d$-valued stochastic process representing the discounted prices of d risky assets in a financial market. **We take $d = 1$ for notational simplicity.** Results in higher dimensions only induce more notations, but do not need substantially new ideas. Let Θ be a suitable set of $\mathbb{F}$-predictable processes standing for investment strategies. For a generic risk tolerance parameter $\xi > 0$, the MVPS criterion is

$$j^{\mathrm{mv}}(\theta) = E\big[G_T(\theta) - \xi\big(G_T(\theta) - E[G_T(\theta)]\big)^2\big], \tag{4.1}$$

and the MVPS problem is to

$$\text{maximise } j^{\mathrm{mv}}(\theta) \text{ over all } \theta \in \Theta. \tag{4.2}$$

Note that problem (4.2) is not a standard stochastic control problem due to the appearance of a quadratic term in the expected final gains, as pointed out in the earlier chapters. Instead, the criterion j^{mv} is of the form (2.1). Indeed, define the functional $g^{\mathrm{mv}} : \mathbb{R} \times \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ by

$$g^{\mathrm{mv}}(x, \mu) = x - \xi\left(x - \int_{\mathbb{R}} z \,\mathrm{d}\mu(z)\right)^2.$$

Then comparing with (4.1) leads to $j^{\mathrm{mv}}(\theta) = E[g^{\mathrm{mv}}(G_T(\theta), P_{G_T(\theta)})]$.

Let us now give more details for problem (4.2). Recall from (3.2) the space of strategies

$$\begin{aligned}\Theta_{\mathrm{S}} := \{\theta := (\theta_t)_{t=1,\dots,T} : \theta &\text{ is real-valued}, \mathbb{F}\text{-predictable and} \\ &\theta_t \triangle S_t \in L^2 \text{ for all } t = 1, \dots, T\}.\end{aligned} \tag{4.3}$$

Our goal is to obtain a dynamic description of the globally optimal strategy $\widehat{\theta}^{\mathrm{mv}}$ for problem (4.2) with $\Theta = \Theta_{\mathrm{S}}$.

4.2 Applying the results for the LQ problem

In this subsection, we discuss the problem (4.2) with the choice of $\Theta = \Theta_{\mathrm{S}}$ and connect it to the general results obtained for LQ problems in Section 3.

Comparing the MVPS problem (4.2) with the general LQ problem in (3.1) both with $\Theta = \Theta_{\mathrm{S}}$ leads to the observation that the MVPS problem is a special case of the latter with $a_T = 1, b_T = -\xi, c_T = \xi, d_T = 0$. We recall the structural properties (3.3)–(3.5) for the value function w there and relabel w here as v^{mv} to deduce a sequence of one-step problems

$$v^{\mathrm{mv}}(t-1,\theta) = \sup_{\delta_t \in \Theta_{\mathrm{S}}^{[t]}(\theta)} v^{\mathrm{mv}}\big(t,\theta(t,\delta_t)\big), \quad t = 1,\dots,T, \tag{4.4}$$

with $v^{\mathrm{mv}}(T,\theta(T,\delta_T))$ having the form (3.6) as

$$\begin{aligned} v^{\mathrm{mv}}\big(T,\theta(T,\delta_T)\big) = E\Big[&\widetilde{a}_t \widetilde{Z}_T G_T\big(\theta(T,\delta_T)\big) + \widetilde{b}_T \widetilde{Z}_T\Big(G_T\big(\theta(T,\delta_T)\big)\Big)^2\Big] \\ &+ \widetilde{c}_T\Big(E\big[\widetilde{Z}_T G_T\big(\theta(T,\delta_T)\big)\big]\Big)^2 + \widetilde{d}_T, \end{aligned} \tag{4.5}$$

where

$$\widetilde{Z}_T = 1, \quad \widetilde{a}_T = 1, \quad \widetilde{b}_T = -\xi, \quad \widetilde{c}_T = \xi, \quad \widetilde{d}_T = 0. \tag{4.6}$$

To use the results from Section 3, we introduce some relevant quantities. In view of the recursive formulas (3.22)–(3.26) and the terminal condition (4.6), we define processes $\widetilde{Z}, \widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d}$ by

$$\widetilde{Z}_{t-1} = E[\widetilde{Z}_t|\mathcal{F}_{t-1}] - \frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}, \quad \widetilde{Z}_T = 1, \tag{4.7}$$

$$\widetilde{a}_{t-1} = \widetilde{a}_t \left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1}, \quad \widetilde{a}_T = 1, \tag{4.8}$$

$$\widetilde{b}_{t-1} = \widetilde{b}_t, \quad \widetilde{b}_T = -\xi, \tag{4.9}$$

$$\widetilde{c}_{t-1} = \widetilde{c}_t \left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1}, \quad \widetilde{c}_T = \xi, \tag{4.10}$$

$$\begin{aligned}\widetilde{d}_{t-1} = \widetilde{d}_t - &\left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1} \\ &\times \frac{\widetilde{a}_t^2}{4\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right], \quad \widetilde{d}_T = 0.\end{aligned} \tag{4.11}$$

Now we need to implement Recipe 3.3 or Lemmas 3.5–3.7 in our special case and provide the missing ingredients pointed out in Section 3.4. Specifically, translating Lemmas 3.5 and 3.6 needs that every $\triangle S_t$ is in L^2 and the quantities $\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t$ are well defined. If we can prove this, then we can immediately state the corresponding results for Lemmas 3.5 and 3.6 by replacing $(Z_t, a_t, b_t, c_t, d_t)$ there with $(\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t)$ given in (4.7)–(4.11). Finally, we need to verify the conditions (3.16), (3.17), (3.20), (3.35) and (3.36) all with $(Z_t, a_t, b_t, c_t, d_t)$ replaced by $(\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t)$. Observe from (4.7)–(4.11) that the processes $\widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d}$ depend crucially on the process $\widetilde{Z}$. So we analyse this process first. These tasks are summarised in the following recipe.

Recipe 4.1. 1) Show that the process $\widetilde{Z}$ is well defined.

2) Show that the processes $\widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d}$ are well defined and translate Lemmas 3.5 and 3.6 accordingly.

3) For each $t = 1, \dots, T$, verify conditions (3.16), (3.17), (3.20), (3.35) and (3.36) with $(Z_t, a_t, b_t, c_t, d_t) = (\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t)$.

4.3 Step 1): Well-definedness of the process $\widetilde{Z}$

In this subsection, we implement Recipe 4.1, Step 1). Namely, we prove that the process $\widetilde{Z}$ recursively defined in (4.7) is well defined. Assumption 3.4 requires that $\triangle S_t$ for a fixed t is in L^2. It is natural to impose that the entire process S is square-integrable.

Assumption 4.2. The process S is square-integrable, meaning that $S_t \in L^2$ for

$t = 0, 1, \dots, T$.

We first give some properties of $\widetilde{Z}$ from the recursion (4.7). To simplify the exposition, we define

$$\widetilde{U}_{t,T} := \prod_{u=t}^{T} \left(1 - \frac{E[\widetilde{Z}_u \triangle S_u | \mathcal{F}_{u-1}]}{E[\widetilde{Z}_u (\triangle S_u)^2 | \mathcal{F}_{u-1}]} \triangle S_u \right), \tag{4.12}$$

with the usual convention that a product over an empty set is 1. In particular, $\widetilde{U}_{T+1,T} = 1$.

Lemma 4.3. *Suppose Assumption* 4.2 *is satisfied. Then we have for* $t = 1, \dots, T$ *that*

$$\widetilde{Z}_{t-1} \textit{ is well defined and has values in } [0,1], \tag{4.13}$$

$$\widetilde{U}_{t,T} \textit{ is well defined and in } L^2, \tag{4.14}$$

$$\widetilde{U}_{t+1,T} \frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \triangle S_t \in L^2, \tag{4.15}$$

$$\widetilde{Z}_{t-1} = E[\widetilde{U}_{t,T} | \mathcal{F}_{t-1}] = E[(\widetilde{U}_{t,T})^2 | \mathcal{F}_{t-1}]. \tag{4.16}$$

In consequence, Condition (3.16) *is satisfied with* $Z_t = \widetilde{Z}_t$ *for* $t = 1, \dots, T$. *Explicitly, this means that we have*

$$\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \widetilde{Z}_t \triangle S_t \in L^2, \quad t = 1, \dots, T. \tag{4.17}$$

Proof. This result is formulated differently in Schweizer [62, Lemma 3]. For completeness, we give a proof here. Let us argue (4.13)–(4.16) by backward induction. Recall the convention that $\frac{0}{0} = 0$. For $t = T$, using $\widetilde{Z}_T = 1$ and the Cauchy–Schwarz inequality in the recursive definition (4.7) for $\widetilde{Z}$ yields that

$$\widetilde{Z}_{T-1} = 1 - \frac{(E[\triangle S_T | \mathcal{F}_{T-1}])^2}{E[(\triangle S_T)^2 | \mathcal{F}_{T-1}]} \text{ has values in } [0,1].$$

This establishes (4.13) for the base case $t = T$ and also shows (4.14) for $t = T$, i.e. $\widetilde{U}_{T,T}$ is well defined and in L^2. Now we consider

$$Y_T := \frac{E[\widetilde{Z}_T \triangle S_T | \mathcal{F}_{T-1}]}{E[\widetilde{Z}_T (\triangle S_T)^2 | \mathcal{F}_{T-1}]} \triangle S_T = \frac{E[\triangle S_T | \mathcal{F}_{T-1}]}{E[(\triangle S_T)^2 | \mathcal{F}_{T-1}]} \triangle S_T = 1 - \widetilde{U}_{T,T} \tag{4.18}$$

and claim that $E[Y_T^2 | \mathcal{F}_{T-1}] \le 1$. This then yields (4.15) for $t = T$; note that

$\widetilde{U}_{T+1,T} = 1$. To justify the claim, we use the fact that the conditional expectation of a nonnegative random variable always exists and take out the $\mathcal{F}_{T-1}$-measurable quantity to get

$$E[Y_T^2|\mathcal{F}_{T-1}] = \frac{(E[\triangle S_T|\mathcal{F}_{T-1}])^2}{(E[(\triangle S_T)^2|\mathcal{F}_{T-1}])^2} E[(\triangle S_T)^2|\mathcal{F}_{T-1}] = \frac{(E[\triangle S_T|\mathcal{F}_{T-1}])^2}{E[(\triangle S_T)^2|\mathcal{F}_{T-1}]}. \quad (4.19)$$

The Cauchy–Schwarz inequality then yields the desired claim. Next, because both Y_T and $\triangle S_T$ are in L^2, a similar conditioning argument gives that $E[Y_T|\mathcal{F}_{T-1}]$ is equal to the right-hand side in (4.19) and hence is equal to $E[Y_T^2|\mathcal{F}_{T-1}]$. In view of (4.18), the fact that $E[Y_T|\mathcal{F}_{T-1}]$ is equal to the right-hand side in (4.19), and the recursion (4.7) for $t = T$ with $\widetilde{Z}_T = 1$, we get

$$E[\widetilde{U}_{T,T}|\mathcal{F}_{T-1}] = E[1 - Y_T|\mathcal{F}_{T-1}] = 1 - \frac{(E[\triangle S_T|\mathcal{F}_{T-1}])^2}{E[(\triangle S_T)^2|\mathcal{F}_{T-1}]} = \widetilde{Z}_{T-1}.$$

This yields the first equality in (4.16) for $t = T$. Using (4.18) and the identity $E[Y_T|\mathcal{F}_{T-1}] = E[Y_T^2|\mathcal{F}_{T-1}]$ from (4.18) and (4.19), we get

$$E[(1 - \widetilde{U}_{T,T})^2|\mathcal{F}_{T-1}] = E[Y_T^2|\mathcal{F}_{T-1}] = E[Y_T|\mathcal{F}_{T-1}] = E[1 - \widetilde{U}_{T,T}|\mathcal{F}_{T-1}].$$

Squaring out the terms yields the second equality in (4.16) for $t = T$, namely $E[\widetilde{U}_{T,T}|\mathcal{F}_{T-1}] = E[(\widetilde{U}_{T,T})^2|\mathcal{F}_{T-1}]$.

Suppose now that (4.13)–(4.16) are satisfied for $t+1$. We justify the induction step for (4.15), (4.14), (4.16) and (4.13) all for t in order. Consider

$$Y_t := \widetilde{U}_{t+1,T} \frac{E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]} \triangle S_t, \quad (4.20)$$

and note that Y_t is not $\mathcal{F}_t$-measurable. Similarly to the base case, we use the tower property, (4.16) for $t + 1$ from the induction hypothesis and the tower property again to get

$$\begin{aligned}
E[Y_t^2|\mathcal{F}_{t-1}] &= E\big[E[Y_t^2|\mathcal{F}_t]\big|\mathcal{F}_{t-1}\big] \\
&= E\left[E[\widetilde{U}_{t+1,T}^2|\mathcal{F}_t]\left(\frac{E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t\right)^2\middle|\mathcal{F}_{t-1}\right] \\
&= E\left[\widetilde{Z}_t\left(\frac{E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t\right)^2\middle|\mathcal{F}_{t-1}\right] \\
&= \frac{(E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}.
\end{aligned} \tag{4.21}$$

Applying the Cauchy–Schwarz inequality and (4.13) for $t+1$ to (4.21) gives

$$E[Y_t^2|\mathcal{F}_{t-1}] = \frac{(E[\widetilde{Z}_t^{\frac{1}{2}}\widetilde{Z}_t^{\frac{1}{2}}\triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]} \le E[\widetilde{Z}_t|\mathcal{F}_{t-1}] \le 1. \tag{4.22}$$

In particular, this implies $E[Y_t^2] \le 1$ and hence gives (4.15) for t. The definition (4.12) for $\widetilde{U}_{t,T}$ yields the identity

$$\widetilde{U}_{t,T} = \widetilde{U}_{t+1,T} - \widetilde{U}_{t+1,T}\frac{E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t = \widetilde{U}_{t+1,T} - Y_t. \tag{4.23}$$

In particular, this shows that $\widetilde{U}_{t,T}$ is well defined because $\widetilde{U}_{t+1,T}$ and Y_t are. The additional property $\widetilde{U}_{t,T} \in L^2$ in (4.14) for t then follows from the middle equality in (4.23), the just proved property (4.15) for t, and the induction hypothesis (4.14) for $t+1$.

The recursion (4.7) for $\widetilde{Z}_{t-1}$ and (4.21) yield $\widetilde{Z}_{t-1} = E[\widetilde{Z}_t|\mathcal{F}_{t-1}] - E[Y_t^2|\mathcal{F}_{t-1}]$, and thus $\widetilde{Z}_{t-1}$ is well defined. We also obtain that $\widetilde{Z}_{t-1} \le E[\widetilde{Z}_t|\mathcal{F}_{t-1}] \le 1$ due to the induction hypothesis (4.13) for $t+1$. The remaining assertion in (4.13) for t is that $\widetilde{Z}_{t-1} \ge 0$, which is a consequence of (4.16) for t.

We turn to proving (4.16) for t. Due to (4.22) and Assumption 4.2, both Y_t and $\triangle S_t$ are in L^2. So we can repeat the steps in (4.21) to compute (with now

Y_t instead of Y_t^2)

$$
\begin{aligned}
E[Y_t|\mathcal{F}_{t-1}] &= E\big[E[Y_t|\mathcal{F}_t]\big|\mathcal{F}_{t-1}\big] \\
&= E\left[E[\widetilde{U}_{t+1,T}|\mathcal{F}_t]\frac{E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t\middle|\mathcal{F}_{t-1}\right] \\
&= E\left[\widetilde{Z}_t\frac{E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t\middle|\mathcal{F}_{t-1}\right] \\
&= \frac{(E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]} \\
&= E[Y_t^2|\mathcal{F}_{t-1}].
\end{aligned}
\tag{4.24}
$$

The conditioning in the second equality uses that $\widetilde{U}_{t+1,T} \in L^2$ from (4.14) for $t+1$, and the conditioning in the second-to-last equality uses that $\widetilde{Z}_t\triangle S_t \in L^1$ by $\widetilde{Z}_t \in [0,1]$ and $\triangle S_t \in L^2$. The first equality in (4.16) for t is now a direct consequence of (4.23), the second-to-last equality in (4.24), (4.16) for $t+1$ and finally the recursion (4.7) for $\widetilde{Z}_{t-1}$ via

$$
\begin{aligned}
E[\widetilde{U}_{t,T}|\mathcal{F}_{t-1}] &= E[\widetilde{U}_{t+1,T}|\mathcal{F}_{t-1}] - E[Y_t|\mathcal{F}_{t-1}] \\
&= E[\widetilde{Z}_t|\mathcal{F}_{t-1}] - \frac{(E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]} \\
&= \widetilde{Z}_{t-1}.
\end{aligned}
$$

Using (4.23) repeatedly, then iterative conditioning and finally invoking (4.16) for $t+1$ and (4.24) yields

$$
\begin{aligned}
E[\widetilde{U}_{t,T}^2|\mathcal{F}_{t-1}] &= E[(\widetilde{U}_{t+1,T} - Y_t)^2|\mathcal{F}_{t-1}] \\
&= E\big[E[\widetilde{U}_{t+1,T}^2|\mathcal{F}_t]\big|\mathcal{F}_{t-1}\big] - 2E\big[E[\widetilde{U}_{t+1,T}Y_t|\mathcal{F}_t]\big|\mathcal{F}_{t-1}\big] + E[Y_t^2|\mathcal{F}_{t-1}] \\
&= E[\widetilde{U}_{t+1,T}|\mathcal{F}_{t-1}] - 2E[Y_t|\mathcal{F}_{t-1}] + E[Y_t|\mathcal{F}_{t-1}] \\
&= E[\widetilde{U}_{t,T}|\mathcal{F}_{t-1}].
\end{aligned}
$$

This gives the second equality in (4.16) for t. The second-to-last line uses the tower property and the identity

$$
E[\widetilde{U}_{t+1,T}Y_t|\mathcal{F}_t] = E[\widetilde{U}_{t+1,T}^2|\mathcal{F}_t]\frac{E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t = E[Y_t|\mathcal{F}_t]
$$

due to the definition (4.20) for Y_t, the second equality in (4.16) for $t+1$ and

(4.20) again. This completes the proof of (4.13)–(4.16).

Finally, Condition (3.16) for $Z_t = \widetilde{Z}_t$ reads explicitly in (4.17) as

$$\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \widetilde{Z}_t \triangle S_t \in L^2.$$

But because $0 \leq \widetilde{Z}_t \leq 1$ by (4.13) for $t+1$, we obtain as in (4.22) that

$$\begin{aligned} E\left[\left(\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \widetilde{Z}_t \triangle S_t\right)^2 \middle| \mathcal{F}_{t-1}\right] &\leq E\left[\left(\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right)^2 \widetilde{Z}_t \triangle S_t^2 \middle| \mathcal{F}_{t-1}\right] \\ &= \frac{(E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \\ &\leq 1. \end{aligned}$$

□

The next result immediately follows from the construction of $\widetilde{Z}$. It has apparently not been noticed in Schweizer [62], but analogous results appear for instance in Černý and Kallsen [17, Corollary 3.4] or Jeanblanc et al. [38, Lemma 1.5].

Corollary 4.4. *Suppose that Assumption* 4.2 *is satisfied. Then the process* $\widetilde{Z}$ *defined by* (4.7) *is a submartingale.*

Proof. Due to (4.13), we have $\widetilde{Z}_t \in L^1$. Moreover, the recursive formula (4.7) yields $\widetilde{Z}_{t-1} \leq E[\widetilde{Z}_t | \mathcal{F}_{t-1}]$ for $t = 1, \dots, T$. □

4.4 Step 2): Translating the results for the LQ problem

In this subsection, we implement Recipe 4.1, Step 2). Namely, we show that the deterministic processes $\widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d}$ are well defined and then translate the results in Lemmas 3.5 and 3.6 to the present setting.

To make the expressions in (4.14)–(4.16) lighter, we introduce a process β via

$$\beta_t := \frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}, \quad t = 1, \dots, T. \tag{4.25}$$

Note that if $\widetilde{Z}$ is deterministic, this would coincide with the process $\widetilde{\lambda}$ from Chapter I; see (I.3.17), (I.3.10) and (I.3.9). Under Assumption 4.2, the process β is also well-defined with the convention $\frac{0}{0} = 0$ because $\widetilde{Z}$ is well defined by

Lemma 4.3 and as in (4.22),

$$(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2 \leq E[\widetilde{Z}_t|\mathcal{F}_{t-1}]E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]$$

due to the Cauchy–Schwarz inequality. Using (4.25) and (4.17), we get

$$\widetilde{Z}_t\beta_t\triangle S_t \in L^1 \quad \text{and} \quad E[\widetilde{Z}_t\beta_t\triangle S_t|\mathcal{F}_{t-1}] = \frac{(E[\widetilde{Z}_t\triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}. \tag{4.26}$$

The property (4.26) allows us to write the recursive definitions (4.7)–(4.11) more lightly as

$$\widetilde{Z}_{t-1} = E[\widetilde{Z}_t(1-\beta_t\triangle S_t)|\mathcal{F}_{t-1}], \qquad \widetilde{Z}_T = 1, \tag{4.27}$$

$$\widetilde{a}_{t-1} = \widetilde{a}_t\left(1+\frac{\widetilde{c}_t}{\widetilde{b}_t}E[\widetilde{Z}_t\beta_t\triangle S_t]\right)^{-1}, \qquad \widetilde{a}_T = 1, \tag{4.28}$$

$$\widetilde{b}_{t-1} = \widetilde{b}_t, \qquad \widetilde{b}_T = -\xi, \tag{4.29}$$

$$\widetilde{c}_{t-1} = \widetilde{c}_t\left(1+\frac{\widetilde{c}_t}{\widetilde{b}_t}E[\widetilde{Z}_t\beta_t\triangle S_t]\right)^{-1}, \qquad \widetilde{c}_T = \xi, \tag{4.30}$$

$$\widetilde{d}_{t-1} = \widetilde{d}_t - \left(1+\frac{\widetilde{c}_t}{\widetilde{b}_t}E[\widetilde{Z}_t\beta_t\triangle S_t]\right)^{-1}\frac{\widetilde{a}_t^2}{4\widetilde{b}_t}E[\widetilde{Z}_t\beta_t\triangle S_t], \qquad \widetilde{d}_T = 0. \tag{4.31}$$

Moreover, we also write (4.12) as

$$\widetilde{U}_{t,T} = \prod_{u=t}^{T}(1-\beta_u\triangle S_u) = \frac{\mathcal{E}(-\int\beta\,\mathrm{d}S)_T}{\mathcal{E}(-\int\beta\,\mathrm{d}S)_{t-1}} =: \mathcal{E}_{t,T}\Big(-\int\beta\,\mathrm{d}S\Big).$$

Now we show that $\widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t$ and $\widetilde{d}_t$ are well defined for $t = 0, 1, \dots, T$. By the definitions in (4.8)–(4.11) of these quantities, it suffices to show for $t = 1, \dots, T$ that

$$1+\frac{\widetilde{c}_t}{\widetilde{b}_t}E[\widetilde{Z}_t\beta_t\triangle S_t] \neq 0. \tag{4.32}$$

Note that this is simply the condition (3.17) in our special case which needs to be shown anyway in view of the discussion in Section 3.4. To establish (4.32), we need a further assumption.

Assumption 4.5. 1) The space Θ_{S} satisfies Assumption I.2.2, 2) which says that the L^2-closure of $G_T(\Theta_{\mathrm{S}})$ does not contain the constant payoff 1.

2) The process S satisfies the structure condition (SC), meaning that the process $(E[\triangle S_t|\mathcal{F}_{t-1}])_{t=1,\dots,T}$ is absolutely continuous with respect to the process

$(\mathrm{Var}[\triangle S_t|\mathcal{F}_{t-1}])_{t=1,\dots,T}$.

We recall from (I.3.14) that Assumption 4.5, 2) implies for $t = 1, \dots, T$ that $E[\triangle S_t|\mathcal{F}_{t-1}] \ll \mathrm{Var}[\triangle S_t|\mathcal{F}_{t-1}]$ in the sense that

$$E[\triangle S_t|\mathcal{F}_{t-1}] = 0 \quad \text{on } \{\mathrm{Var}[\triangle S_t|\mathcal{F}_{t-1}] = 0\}. \tag{4.33}$$

Lemma 4.6. *Suppose Assumptions* 4.2 *and* 4.5 *are satisfied. Then the following statements hold:*

1) $E[\widetilde{Z}_t] > 0$ *for* $t = 0, 1, \dots, T$.
2) $\frac{\widetilde{c}_t}{\widetilde{b}_t} = -\frac{1}{E[\widetilde{Z}_t]}$ *for* $t = 0, 1, \dots, T$, *and* (4.32) *holds for* $t = 1, \dots, T$.
3) *The quantities* $\widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t$ *and* $\widetilde{d}_t$ *are well defined for* $t = 0, 1, \dots, T$.

Proof. 1) Delbaen and Schachermayer [23, Lemma 2.1] asserts that Assumption 4.5, 1) is equivalent to the statement that the set of signed $G_T(\Theta_{\mathrm{S}})$-martingale measures (see (I.3.22) for the definition) is nonempty, under which the result $E[\widetilde{Z}_t] > 0$ for $t = 0$ is deduced from Corollary 4 and Theorem 5 in Schweizer [62]. For $t = 1, \dots, T$, we use the submartingale property of $\widetilde{Z}$ from Lemma 4.3 to get $E[\widetilde{Z}_t] \geq E[\widetilde{Z}_0] > 0$.

2) For convenience, we introduce a process h by $h_t = \frac{\widetilde{c}_t}{\widetilde{b}_t}$ for $t = 0, 1, \dots, T$, which by the recursions (4.29) for $\widetilde{b}_{t-1}$ and (4.30) for $\widetilde{c}_{t-1}$ satisfies

$$h_{t-1} = \frac{\widetilde{c}_{t-1}}{\widetilde{b}_{t-1}} = \frac{\widetilde{c}_t}{\widetilde{b}_t}\left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E[\widetilde{Z}_t \beta_t \triangle S_t]\right)^{-1} = h_t(1 + h_t E[\widetilde{Z}_t \beta_t \triangle S_t])^{-1} \tag{4.34}$$

whenever $1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E[\widetilde{Z}_t \beta_t \triangle S_t] \neq 0$. We prove that $h_t = -\frac{1}{E[\widetilde{Z}_t]}$ for $t = 0, 1, \dots, T$ and (4.32) for $t = 1, \dots, T$ by backward induction. For $t = T$, the identity $h_T = -\frac{1}{E[\widetilde{Z}_T]}$ immediately follows from

$$h_T = \frac{\widetilde{c}_T}{\widetilde{b}_T} = \frac{\xi}{-\xi} = -1 = -\frac{1}{E[\widetilde{Z}_T]}$$

thanks to the terminal conditions $\widetilde{Z}_T = 1, \widetilde{b}_T = -\xi, \widetilde{c}_T = \xi$ in (4.27), (4.29) and (4.30). Note that $E[\triangle S_T|\mathcal{F}_{T-1}] = 0$ on $\{\mathrm{Var}[\triangle S_T|\mathcal{F}_{T-1}] = 0\}$ from (4.33) and the convention $\frac{0}{0} = 0$ yield

$$\frac{(E[\triangle S_T|\mathcal{F}_{T-1}])^2}{E[(\triangle S_T)^2|\mathcal{F}_{T-1}]} = \frac{(E[\triangle S_T|\mathcal{F}_{T-1}])^2}{\mathrm{Var}[\triangle S_T|\mathcal{F}_{T-1}] + (E[\triangle S_T|\mathcal{F}_{T-1}])^2} < 1 \quad P\text{-a.s.}$$

Combining the above with $\widetilde{Z}_T = 1, \widetilde{b}_T = -\xi, \widetilde{c}_T = \xi$, we get

$$1 + \frac{\widetilde{c}_T}{\widetilde{b}_T} E\left[\frac{(E[\widetilde{Z}_T \triangle S_T | \mathcal{F}_{T-1}])^2}{E[\widetilde{Z}_T (\triangle S_T)^2 | \mathcal{F}_{T-1}]}\right] = 1 - E\left[\frac{(E[\triangle S_T | \mathcal{F}_{T-1}])^2}{E[(\triangle S_T)^2 | \mathcal{F}_{T-1}]}\right] > 0.$$

This proves condition (4.32) for $t = T$.

Suppose now that $h_t = -\frac{1}{E[\widetilde{Z}_t]}$ and (4.32) holds for t. Then the identity (4.34) for h_{t-1} is true. Using (4.34), the identity $h_t = -\frac{1}{E[\widetilde{Z}_t]}$ from the induction hypothesis and the new recursion (4.27) for $\widetilde{Z}_{t-1}$, we get

$$\begin{aligned} h_{t-1} &= h_t (1 + h_t E[\widetilde{Z}_t \beta_t \triangle S_t])^{-1} \\ &= -\frac{1}{E[\widetilde{Z}_t]} \frac{1}{1 - \frac{1}{E[\widetilde{Z}_t]} E[\widetilde{Z}_t \beta_t \triangle S_t]} \\ &= -\frac{1}{E[\widetilde{Z}_t (1 - \beta_t \triangle S_t)]} \\ &= -\frac{1}{E[\widetilde{Z}_{t-1}]}. \end{aligned}$$

Next, we use the definition $h_t = \frac{\widetilde{c}_t}{\widetilde{b}_t}$, plug the above equality back into the left-hand side of (4.32) for t and use the new recurrence relation (4.27) for $\widetilde{Z}_{t-1}$ and $E[\widetilde{Z}_t] > 0$ from part 1) to obtain

$$\begin{aligned} 1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E[\widetilde{Z}_t \beta_t \triangle S_t] &= 1 + h_t E[\widetilde{Z}_t \beta_t \triangle S_t] \\ &= 1 - \frac{1}{E[\widetilde{Z}_t]} E[\widetilde{Z}_t \beta_t \triangle S_t] \\ &= \frac{E[\widetilde{Z}_t (1 - \beta_t \triangle S_t)]}{E[\widetilde{Z}_t]} \\ &= \frac{E[\widetilde{Z}_{t-1}]}{E[\widetilde{Z}_t]} > 0. \end{aligned}$$

This proves (4.32) for $t - 1$.

3) This is an immediate consequence of part 2) and the definitions (4.28)–(4.31) for $\widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d}$. □

Corollary 4.7. *Suppose Assumptions* 4.2 *and* 4.5 *are satisfied. Then the process* $\widetilde{a}$ *given in* (4.28) *satisfies for* $t = 0, 1, \dots, T$ *that*

$$\widetilde{a}_t = -\frac{\widetilde{c}_t}{\widetilde{b}_t} = \frac{1}{E[\widetilde{Z}_t]}. \tag{4.35}$$

Proof. The second equality in (4.35) is obtained in Lemma 4.6, 2). We prove the first equality in (4.35) by backward induction. The base case for (4.35) reads $\widetilde{a}_T = -\frac{\widetilde{c}_T}{\widetilde{b}_T}$, which is clear from $\widetilde{a}_T = 1, \widetilde{b}_T = -\xi, \widetilde{c}_T = \xi$ given in (4.28), (4.29) and (4.30). Now suppose that $\widetilde{a}_t = -\frac{\widetilde{c}_t}{\widetilde{b}_t}$ is true for t. Then the well-definedness of $\widetilde{a}, \widetilde{b}, \widetilde{c}$ from Lemma 4.6, 3) allows us to use the recursions (4.29) for $\widetilde{b}_{t-1}$ and (4.30) for $\widetilde{c}_{t-1}$, the induction hypothesis (4.35) for t and finally the recursion (4.28) for $\widetilde{a}_{t-1}$ to obtain

$$\frac{\widetilde{c}_{t-1}}{\widetilde{b}_{t-1}} = \frac{\widetilde{c}_t}{\widetilde{b}_t}\left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E[\widetilde{Z}_t \beta_t \triangle S_t]\right)^{-1} = -\widetilde{a}_t \left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E[\widetilde{Z}_t \beta_t \triangle S_t]\right)^{-1} = -\widetilde{a}_{t-1}.$$

This completes the induction step and hence proves (4.35). □

In view of the discussion before Recipe 4.1 and the well-definedness of the quantities $\widetilde{Z}, \widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d}$ from Lemmas 4.3 and 4.6, we can now mechanically implement Recipe 3.3, Step $1)_t$ for $t = 0, 1, \dots, T$.

Proposition 4.8. *Suppose Assumptions* 4.2 *and* 4.5 *are satisfied. Fix a time* $t \in \{1, \dots, T\}$. *If* $\theta \in \Theta_{\mathrm{S}}$, $\delta_t \in \Theta_{\mathrm{S}}^{[t]}(\theta)$ *and* v^{mv} *satisfies*

$$v^{\mathrm{mv}}\big(t, \theta(t, \delta_t)\big) = E\Big[\widetilde{a}_t \widetilde{Z}_t G_t\big(\theta(t, \delta_t)\big) + \widetilde{b}_t \widetilde{Z}_t \Big(G_t\big(\theta(T, \delta_t)\big)\Big)^2\Big] + \widetilde{c}_t \Big(E\big[\widetilde{Z}_t G_t\big(\theta(T, \delta_t)\big)\big]\Big)^2 + \widetilde{d}_t, \tag{4.36}$$

$$\textit{where } \widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t \textit{ are given by } (4.7)\text{–}(4.11), \tag{4.37}$$

then the following statements hold:

1) *We have*

$$v^{\mathrm{mv}}\big(t, \theta(t, \delta_t)\big) = E\big[\widetilde{a}_t \widetilde{Z}_t G_{t-1}(\theta) + \widetilde{b}_t \widetilde{Z}_t \big(G_{t-1}(\theta)\big)^2\big] + \widetilde{c}_t \big(E[\widetilde{Z}_t G_{t-1}(\theta)]\big)^2 + \widetilde{d}_t + r_t^{\mathrm{mv}}(\delta_t), \tag{4.38}$$

where

$$r_t^{\mathrm{mv}}(\delta_t) = E\big[\widetilde{a}_t \widetilde{Z}_t \delta_t \triangle S_t + 2\widetilde{b}_t \widetilde{Z}_t G_{t-1}(\theta) \delta_t \triangle S_t + \widetilde{b}_t \widetilde{Z}_t (\delta_t \triangle S_t)^2 + 2\widetilde{c}_t E[\widetilde{Z}_t G_{t-1}(\theta)] E[\widetilde{Z}_t \delta_t \triangle S_t] + \widetilde{c}_t (E[\widetilde{Z}_t \delta_t \triangle S_t])^2\big]. \tag{4.39}$$

2) *Any maximiser for* $\delta_t \mapsto r_t^{\mathrm{mv}}(\delta_t)$ *is a solution to the linear equation*

$$\delta_t = -\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \times \bigg(\frac{\widetilde{a}_t}{2\widetilde{b}_t} + G_{t-1}(\theta) + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\big[\widetilde{Z}_t\big(G_{t-1}(\theta) + \delta_t \triangle S_t\big)\big] \bigg). \tag{4.40}$$

A solution $\widetilde{\delta}_t$ *to the linear equation* (4.40) *is explicitly given by*

$$\widetilde{\delta}_t = -\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \bigg(\frac{\widetilde{a}_t}{2\widetilde{b}_t} + G_{t-1}(\theta) + \frac{\widetilde{c}_t}{\widetilde{b}_t} \big(E[\widetilde{Z}_t G_{t-1}(\theta)] + \widetilde{e}_t \big) \bigg), \tag{4.41}$$

where $\widetilde{e}_t$ *is given by*

$$\widetilde{e}_t = -\bigg(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\bigg[\frac{(E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \bigg] \bigg)^{-1} \times E\bigg[\frac{(E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \bigg(\frac{\widetilde{a}_t}{2\widetilde{b}_t} + G_{t-1}(\theta) + \frac{\widetilde{c}_t}{\widetilde{b}_t} E[\widetilde{Z}_t G_{t-1}(\theta)] \bigg) \bigg] \tag{4.42}$$

and satisfies $\widetilde{e}_t = E[\widetilde{Z}_t \widetilde{\delta}_t \triangle S_t]$.

Proof. Parts 1) and 2) are translated from Lemmas 3.5–3.7. Indeed, (4.38)–(4.40) are translated from (3.11), (3.12) and (3.14), respectively. To do the translation properly, we need to verify that the proposed form (4.36), (4.37) of v^{mv} satisfies (3.6)–(3.8). Obviously, the affine–quadratic structure in (4.36) is the same as (3.6). To verify (3.7) and (3.8), we observe from the recursive relations (4.7)–(4.11) for $\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t$ that $\widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t$ are deterministic, $\widetilde{b}_t = \widetilde{b}_T = -\xi \neq 0$ and $\widetilde{Z}_t$ is $\mathcal{F}_t$-measurable. By (4.16), $\widetilde{Z}_t$ also has values in $[0, 1]$ so that it is nonnegative and bounded.

Finally, we apply Lemma 3.7, 1) with $(Z_t, a_t, b_t, c_t, d_t) = (\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t)$ to solve the linear equation (4.40) and obtain (4.41) and (4.42). Thanks to (4.17), we have (3.16) with $Z_t = \widetilde{Z}_t$. By (4.32) argued in Lemma 4.6, we also have (3.17). Therefore, Lemma 3.7, 1) can be applied as desired; (4.41) and (4.42) is then translated from (3.19) and (3.18). □

4.5 Step 3): Fulfilling the missing requirement

Having Proposition 4.8, we can complete the implementation of Recipe 4.1, Steps 1) and 2). To implement Recipe 4.1, Step 3), what needs to be done is to verify the conditions (3.16), (3.17), (3.20), (3.35) and (3.36) with $(Z_t, a_t, b_t, c_t, d_t)$ there equal to $(\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t)$. These conditions are explicitly given by

$$\frac{E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\widetilde{Z}_t \triangle S_t \in L^2, \tag{4.43}$$

$$1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right] \neq 0, \tag{4.44}$$

$$\widetilde{b}_t + \widetilde{c}_t E[\widetilde{Z}_t] \leq 0 \text{ and the solution } \widetilde{\delta}_t \text{ to (4.40) is in } \Theta_{\mathrm{S}}^{[t]}(\theta), \tag{4.45}$$

$$\widetilde{a}_{t-1}, \widetilde{b}_{t-1}, \widetilde{c}_{t-1}, \widetilde{d}_{t-1} \text{ are deterministic with } \widetilde{b}_t \neq 0, \tag{4.46}$$

$$\widetilde{Z}_{t-1} \text{ is bounded, nonnegative and } \mathcal{F}_{t-1}\text{-measurable,} \tag{4.47}$$

for $t = 1, \ldots, T$. Note that (4.43) is already given by (4.17) and (4.44) is proved by Lemma 4.6, 2). In Corollary 4.7, we also obtained $\widetilde{a}_t = -\frac{\widetilde{c}_t}{\widetilde{b}_t} = \frac{1}{E[\widetilde{Z}_t]}$ which yields the first half of the condition (4.45) because

$$\widetilde{b}_t + \widetilde{c}_t E[\widetilde{Z}_t] = \widetilde{b}_t(1 - \widetilde{a}_t E[\widetilde{Z}_t]) = 0.$$

Moreover, the conditions (4.46) and (4.47) have been established in the proof of Proposition 4.8. So among the conditions (4.43)–(4.47), it only remains to prove the second half of (4.45).

Let us take a quick look at this condition. We recall from (4.41) that

$$\widetilde{\delta}_t = -\frac{E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\left(\frac{\widetilde{a}_t}{2\widetilde{b}_t} + G_{t-1}(\theta) + \frac{\widetilde{c}_t}{\widetilde{b}_t}\big(E[\widetilde{Z}_t G_{t-1}(\theta)] + \widetilde{e}_t\big)\right)$$

for a constant $\widetilde{e}_t$ given in (4.42). So the second half of (4.45) for $t = T$ reads equivalently

$$\left(\frac{E[\widetilde{Z}_T \triangle S_T|\mathcal{F}_{T-1}]}{E[\widetilde{Z}_T(\triangle S_T)^2|\mathcal{F}_{T-1}]}\triangle S_T G_{T-1}(\theta)\right)^2 \in L^1.$$

This is satisfied because $E[(\frac{E[\widetilde{Z}_T \triangle S_T|\mathcal{F}_{T-1}]}{E[\widetilde{Z}_T(\triangle S_T)^2|\mathcal{F}_{T-1}]}\triangle S_T)^2|\mathcal{F}_{T-1}] \leq 1$, as argued in (4.18) and (4.19), and $G_{T-1}(\theta) \in L^2$ due to $\theta \in \Theta_{\mathrm{S}}$. For a general t, we have no control over the factor $\frac{E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t$, while the other factor $G_{t-1}(\theta)$ is only known to be in L^2 due to the arbitrariness of θ. At present, we can only proceed in implementing Recipe 4.1, Step 2) by leaving the above condition (i.e., the second half of (4.45)) as an extra assumption.

Proposition 4.9. *Suppose Assumptions* 4.2 *and* 4.5 *are satisfied. Fix* $\theta \in \Theta_{\mathrm{S}}$ *and* $t \in \{1, \ldots, T\}$. *If* $v^{\mathrm{mv}}(t, \cdot)$ *is of the form* (4.36) *and* (4.37), *then the following statements hold:*

1) *Suppose that*

$$\text{the solution } \widetilde{\delta}_t \text{ given in (4.41) and (4.42) satisfies } \widetilde{\delta}_t \triangle S_t \in L^2. \tag{4.48}$$

Then a maximiser for $\delta_t \mapsto r_t^{\mathrm{mv}}(\delta_t)$ *exists and is in* $\Theta_{\mathrm{S}}^{[t]}(\theta)$. *It is a solution to* (4.40) *and given explicitly by* $\widetilde{\delta}_t$ *in* (4.41) *and* (4.42).

2) *If* (4.48) *is satisfied, then the resulting*

$$v^{\mathrm{mv}}(t-1, \theta) = \sup_{\delta_t \in \Theta_{\mathrm{S}}^{[t]}(\theta)} v\big(t, \theta(t, \delta_t)\big) = v^{\mathrm{mv}}\big(t, \theta(t, \widetilde{\delta}_t)\big)$$

from (4.4) *is also of the form* (4.36) *and* (4.37). *Precisely, we have*

$$\begin{aligned} v^{\mathrm{mv}}(t-1, \theta) = E\big[&\widetilde{a}_{t-1} \widetilde{Z}_{t-1} G_{t-1}(\theta) + \widetilde{b}_{t-1} \widetilde{Z}_{t-1} \big(G_{t-1}(\theta)\big)^2\big] \\ &+ \widetilde{c}_{t-1} \big(E[\widetilde{Z}_{t-1} G_{t-1}(\theta)]\big)^2 + \widetilde{d}_t, \end{aligned} \tag{4.49}$$

where $\widetilde{Z}_{t-1}, \widetilde{a}_{t-1}, \widetilde{b}_{t-1}, \widetilde{c}_{t-1}, \widetilde{d}_{t-1}$ *are given by* (4.7)–(4.11), *namely*

$$\widetilde{Z}_{t-1} = E[\widetilde{Z}_t | \mathcal{F}_{t-1}] - \frac{(E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}, \tag{4.50}$$

$$\widetilde{a}_{t-1} = \widetilde{a}_t \left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right]\right)^{-1}, \tag{4.51}$$

$$\widetilde{b}_{t-1} = \widetilde{b}_t, \tag{4.52}$$

$$\widetilde{c}_{t-1} = \widetilde{c}_t \left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right]\right)^{-\frac{1}{2}}, \tag{4.53}$$

$$\begin{aligned} \widetilde{d}_{t-1} = \widetilde{d}_t - &\left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right]\right)^{-1} \\ &\times \frac{\widetilde{a}_t^2}{4\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}\right]. \end{aligned} \tag{4.54}$$

Proof. Thanks to the extra assumption (4.48), the conditions (3.16), (3.17) and (3.20) with $(Z_t, a_t, b_t, c_t, d_t) = (\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t)$ are satisfied. So Lemma 3.7, 2) and 3) can be applied and translated into the desired results. The identity (4.49) is from (3.21). The recursions (4.50)–(4.54) are (3.22)–(3.26) with

$(Z_t, a_t, b_t, c_t, d_t) = (\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t)$, which obviously agrees with the definitions (4.7)–(4.11) of $\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t$. □

4.6 Complete recursion and solution for the MVPS problem

In this subsection, we apply the results from the previous two subsections to solve the MVPS problem (4.2) with $\Theta = \Theta_{\mathrm{S}}$.

We first recall from (2.5) with $v = v^{\mathrm{mv}}, j = j^{\mathrm{mv}}, \Theta(t,\theta) = \Theta_{\mathrm{S}}(t,\theta)$ and the expression (4.1) for j^{mv} that

$$v^{\mathrm{mv}}(t,\theta) := \sup_{\widetilde{\theta} \in \Theta_{\mathrm{S}}(t,\theta)} E\big[G_T(\widetilde{\theta}) - \xi\big(G_T(\widetilde{\theta}) - E[G_T(\widetilde{\theta})]\big)^2\big] \tag{4.55}$$

for $t = 0, 1, \dots, T$. Combining this with (4.4), we need to solve the sequence of problems

$$v^{\mathrm{mv}}(t,\theta) = \sup_{\delta_{t+1} \in \Theta_{\mathrm{S}}^{[t+1]}(\theta)} v^{\mathrm{mv}}\big(t+1, \theta(t+1, \delta_{t+1})\big), \tag{4.56}$$

for $t = 0, \dots, T-1$ with $v^{\mathrm{mv}}(T,\theta) = j^{\mathrm{mv}}(\theta)$. In view of Proposition 4.9, the entire linear–quadratic structure (4.36) and (4.37) for v^{mv} is maintained if (4.48) is satisfied for all $t = 1, \dots, T$. This motivates a final assumption, which is studied later in special cases.

Assumption 4.10. For all $t = 1, \dots, T$ and $\theta \in \Theta_{\mathrm{S}}$, the solution $\widetilde{\delta}_t$ to (4.40) satisfies $\widetilde{\delta}_t \triangle S_t \in L^2$.

We piece everything together to state the main result of this entire section, which is effectively a formality.

Theorem 4.11. *Suppose that Assumptions* 4.2, 4.5 *and* 4.10 *are satisfied. Then:*

1) *For any* $\theta \in \Theta_{\mathrm{S}}$ *and any* $t = 0, 1, \dots, T$, *the value function* $v^{\mathrm{mv}}(t,\theta)$ *is of the form* (4.36) *and* (4.37), *i.e.*

$$\begin{aligned} v^{\mathrm{mv}}\big(t, \theta(t,\delta_t)\big) = E\Big[\widetilde{a}_t \widetilde{Z}_t G_t\big(\theta(t,\delta_t)\big) + \widetilde{b}_t \widetilde{Z}_t \Big(G_t\big(\theta(t,\delta_t)\big)\Big)^2\Big] \\ + \widetilde{c}_t \Big(E\big[\widetilde{Z}_t G_t\big(\theta(t,\delta_t)\big)\big]\Big)^2 + \widetilde{d}_t, \end{aligned} \tag{4.57}$$

where $\widetilde{Z}_t, \widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t$ *are given by*

$$\widetilde{Z}_{t-1} = E[\widetilde{Z}_t|\mathcal{F}_{t-1}] - \frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}, \quad \widetilde{Z}_T = 1, \tag{4.58}$$

$$\widetilde{a}_{t-1} = \widetilde{a}_t \left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1}, \quad \widetilde{a}_T = 1, \tag{4.59}$$

$$\widetilde{b}_{t-1} = \widetilde{b}_t, \quad \widetilde{b}_T = -\xi, \tag{4.60}$$

$$\widetilde{c}_{t-1} = \widetilde{c}_t \left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1}, \quad \widetilde{c}_T = \xi, \tag{4.61}$$

$$\begin{aligned}\widetilde{d}_{t-1} = \widetilde{d}_t - &\left(1 + \frac{\widetilde{c}_t}{\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right]\right)^{-1} \\ &\times \frac{\widetilde{a}_t^2}{4\widetilde{b}_t} E\left[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right], \quad \widetilde{d}_T = 0.\end{aligned} \tag{4.62}$$

Explicitly, with the help of the process β *defined by*

$$\beta_t = \frac{E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}, \quad t = 1, \dots, T, \tag{4.63}$$

and the stochastic exponential $\mathcal{E}$, *the identities* (4.58)–(4.62) *can be written as*

$$\begin{aligned}\widetilde{Z}_t &= E[\widetilde{Z}_{t+1}(1 - \beta_{t+1} \triangle S_{t+1})|\mathcal{F}_t] \\ &= E\left[\frac{\mathcal{E}(-\int \beta \, \mathrm{d}X)_T}{\mathcal{E}(-\int \beta \, \mathrm{d}X)_t}\middle|\mathcal{F}_t\right] \\ &= E\left[\prod_{u=t+1}^{T} (1 - \beta_u \triangle S_u)\middle|\mathcal{F}_t\right], \quad t = 0, 1, \dots, T,\end{aligned} \tag{4.64}$$

$$\widetilde{a}_t = \frac{1}{E[\widetilde{Z}_t]}, \quad t = 0, 1, \dots, T, \tag{4.65}$$

$$\widetilde{b}_t = -\xi, \quad t = 0, 1, \dots, T, \tag{4.66}$$

$$\widetilde{c}_t = \frac{\xi}{E[\widetilde{Z}_t]} = \xi \widetilde{a}_t, \quad t = 0, 1, \dots, T, \tag{4.67}$$

$$\widetilde{d}_t = \frac{1}{4\xi}\left(\frac{1}{E[\widetilde{Z}_t]} - 1\right) = \frac{1}{4\xi}(\widetilde{a}_t - 1), \quad t = 0, 1, \dots, T. \tag{4.68}$$

2) *For every* $\theta \in \Theta_{\mathrm{S}}$, *the solution to the problem* (4.55) *at time* t *is given by*

$\widetilde{\theta}_u = \theta_u$ *for* $u = 1, \dots, t$ *and*

$$\widetilde{\theta}_u = -\frac{E[\widetilde{Z}_u \triangle S_u | \mathcal{F}_{u-1}]}{E[\widetilde{Z}_u (\triangle S_u)^2 | \mathcal{F}_{u-1}]} \left(\frac{\widetilde{a}_u}{2\widetilde{b}_u} + G_{u-1}(\widetilde{\theta}) + \frac{\widetilde{c}_u}{\widetilde{b}_u} \big(E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})] + \widetilde{e}_u \big) \right) \tag{4.69}$$

$$= -\beta_u \left(G_{u-1}(\widetilde{\theta}) - \frac{(2\xi)^{-1} + E[\widetilde{Z}_{u-1} G_{u-1}(\widetilde{\theta})]}{E[\widetilde{Z}_{u-1}]} \right), \quad u = t+1, \dots, T, \tag{4.70}$$

where $\widetilde{e}_u$ *is given by*

$$\widetilde{e}_u = -(1 - \widetilde{a}_u E[\widetilde{Z}_u \beta_u \triangle S_u])^{-1} \times E\left[\widetilde{Z}_u \beta_u \triangle S_u \left(\frac{\widetilde{a}_u}{2\widetilde{b}_u} + G_{u-1}(\widetilde{\theta}) - \widetilde{a}_u E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})] \right) \right]. \tag{4.71}$$

In particular, the solution to the problem (4.55) *at time* 0 *and hence to the MVPS problem* (4.2) *with* $\Theta = \Theta_{\mathrm{S}}$ *is given by*

$$\widetilde{\theta}^{\mathrm{mv}}_u = -\frac{E[\widetilde{Z}_u \triangle S_u | \mathcal{F}_{u-1}]}{E[\widetilde{Z}_u (\triangle S_u)^2 | \mathcal{F}_{u-1}]} \times \left(G_{u-1}(\widetilde{\theta}^{\mathrm{mv}}) - \frac{(2\xi)^{-1} + E[\widetilde{Z}_{u-1} G_{u-1}(\widetilde{\theta}^{\mathrm{mv}})]}{E[\widetilde{Z}_{u-1}]} \right) \tag{4.72}$$

$$= -\beta_u \left(G_{u-1}(\widetilde{\theta}^{\mathrm{mv}}) - \frac{1}{2\xi E[\widetilde{Z}_0]} \right), \quad u = 1, \dots, T. \tag{4.73}$$

Proof. We use in both parts the equivalence between (4.55) and (4.56) which follows from the dynamic programming result in Proposition 2.3, the abstract rewriting in (2.11) after that result and the concrete specification in the beginning of Section 4.2.

1) Assumptions 4.2, 4.5 and 4.10 allow us to apply Propositions 4.8 and 4.9 repeatedly backward in t, starting from $t = T$, at which the structure (4.57)–(4.62) of the value function $v^{\mathrm{mv}}(T, \theta) = j^{\mathrm{mv}}(\theta)$ is exogenously given by the MVPS problem as in (4.5) and (4.6). This yields the structure (4.57)–(4.62) of v^{mv} for all $t = 0, 1, \dots, T$. The explicit expressions (4.64)–(4.68) for $\widetilde{Z}, \widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d}$ are derived from their recursive counterparts (4.58)–(4.62). Indeed, the expressions (4.64) and (4.65) for $\widetilde{Z}$ and $\widetilde{a}$ have been obtained in (4.16) and Lemma 4.6, 2), respectively. The explicit expression $\widetilde{b} \equiv -\xi$ for $\widetilde{b}$ is immediate from its recursive definition (4.60). For $\widetilde{c}$, we recall from (4.35) that $\frac{\widetilde{c}_t}{\widetilde{b}_t} = -\widetilde{a}_t$ for all $t = 1, \dots, T$. Then inserting the explicit formulas for $\widetilde{a}$ and $\widetilde{b}$ into that identity yields (4.67). Finally, we use $\frac{\widetilde{c}_t}{\widetilde{b}_t} = -\widetilde{a}_t$, $\frac{\widetilde{a}_t^2}{4\widetilde{b}_t} = \frac{\widetilde{a}_t^2}{4\xi}$ by (4.66) and finally the recursion (4.58) for

$\widetilde{Z}_{t-1}$ to compute for $t = 1, \dots, T$ that

$$\begin{aligned}\widetilde{d}_{t-1} - \widetilde{d}_t &= \frac{1}{4\xi}\widetilde{a}_t \frac{\widetilde{a}_t E[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}] - 1 + 1}{1 - \widetilde{a}_t E[\frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}]} \\ &= \frac{1}{4\xi}\frac{1}{E[\widetilde{Z}_t]}\left(\frac{1}{1 - \frac{1}{E[\widetilde{Z}_t]}\frac{E[(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]]}} - 1\right) \\ &= \frac{1}{4\xi}\left(\frac{1}{E[\widetilde{Z}_t] - \frac{E[(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]]}} - \frac{1}{E[\widetilde{Z}_t]}\right) \\ &= \frac{1}{4\xi}\left(\frac{1}{E[\widetilde{Z}_{t-1}]} - \frac{1}{E[\widetilde{Z}_t]}\right).\end{aligned}$$

Summing both sides of this display from $t+1$ to T and using $\widetilde{d}_T = 0$ as well as $E[\widetilde{Z}_T] = \widetilde{Z}_T = 1$ yields the explicit expression (4.68) for $\widetilde{d}$.

2) For $t = 1, \dots, T$ and $\theta \in \Theta_{\mathrm{S}}$, we use Assumptions 4.2 and 4.5 to apply Proposition 4.8, 2) and Proposition 4.9, 1) repeatedly from t to T. This yields a strategy $\widetilde{\theta}$ whose expression in (4.69) and (4.71) is obtained by repeatedly translating (4.41) and (4.42) again starting from $t+1$ to T. More precisely, the expression (4.41) for $\widetilde{\delta}_t$ allows us to write $\widetilde{\delta}_t =: f(t, G_{t-1}(\theta))$ and $\widetilde{\theta}$ is obtained by setting $\widetilde{\theta}_s = \theta_s$ for $s = 1, \dots, t$ and $\widetilde{\theta}_u = f(u, G_{u-1}(\widetilde{\theta}))$ for $u = t+1, \dots, T$. Assumption 4.10 says that $f(u, G_{u-1}(\theta))\triangle S_u \in L^2$ for any $u = 1, \dots, T$ and $\theta \in \Theta_{\mathrm{S}}$. We now argue $\widetilde{\theta} \in \Theta_{\mathrm{S}}(t, \theta)$ by induction. First, we have $\widetilde{\theta}_s = \theta_s$ for $s = 1, \dots, t$ by construction. Comparing (4.41) for $\widetilde{\delta}_{t+1}$ with (4.69) for $\widetilde{\theta}_{t+1}$ yields $\widetilde{\theta}_{t+1} = \widetilde{\delta}_{t+1} = f(t+1, G_t(\theta))$ and thus $\widetilde{\theta}_{t+1}\triangle S_{t+1} \in L^2$ due to Assumption 4.10. Suppose next that $\widetilde{\theta}_s \triangle S_s \in L^2$ for $s = t+1, \dots, u-1$. Consider the strategy $\varphi^{u-1} := (\theta_1, \dots, \theta_t, \widetilde{\theta}_{t+1}, \dots, \widetilde{\theta}_{u-1}, 0, \dots, 0) \in \Theta_{\mathrm{S}}$. Because $\widetilde{\theta}_s = \varphi^{u-1}_s$ for $s = 1, \dots, u-1$ by construction, we have $G_{u-1}(\widetilde{\theta}) = G_{u-1}(\varphi^{u-1})$, and therefore $\widetilde{\theta}_u = f(u, G_{u-1}(\widetilde{\theta})) = f(u, G_{u-1}(\varphi^{u-1}))$ satisfies $\widetilde{\theta}_u \triangle S_u \in L^2$ due to $\varphi^{u-1} \in \Theta_{\mathrm{S}}$ and Assumption 4.10. This completes the induction step and shows that $\widetilde{\theta}$ is in $\Theta_{\mathrm{S}}(t, \theta)$.

Due to the optimality of $\widetilde{\delta}_u = f(u, G_{u-1}(\theta))$ for the one-step problem (4.56) for $u = t+1, \dots, T$, we get by the equivalence between (4.55) and (4.56) that $\widetilde{\theta}$ is optimal for the problem (4.55) at time t.

To get (4.70), we first use $\frac{\widetilde{c}_u}{\widetilde{b}_u} = -\widetilde{a}_u$ from (4.35), then factor out $\widetilde{a}_u$ and

$1 - \widetilde{a}_u E[\widetilde{Z}_u \beta_u \triangle S_u]$ and finally use the expression (4.71) for $\widetilde{e}_u$ to obtain

$$\begin{aligned}
&\frac{\widetilde{a}_u}{2\widetilde{b}_u} + \frac{\widetilde{c}_u}{\widetilde{b}_u}\big(E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})] + \widetilde{e}_u\big)\\
&= \widetilde{a}_u \left(\frac{1}{2\widetilde{b}_u} - E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})] - \widetilde{e}_u \right)\\
&= \frac{\widetilde{a}_u}{1 - \widetilde{a}_u E[\widetilde{Z}_u \beta_u \triangle S_u]} \left(\frac{1 - \widetilde{a}_u E[\widetilde{Z}_u \beta_u \triangle S_u]}{2\widetilde{b}_u} - E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})] \right.\\
&\qquad + \widetilde{a}_u E[\widetilde{Z}_u \beta_u \triangle S_u] E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})]\\
&\qquad \left. + E\left[\widetilde{Z}_u \beta_u \triangle S_u \left(\frac{\widetilde{a}_u}{2\widetilde{b}_u} + G_{u-1}(\widetilde{\theta}) - \widetilde{a}_u E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})] \right) \right] \right)\\
&= \frac{\widetilde{a}_u}{1 - \widetilde{a}_u E[\widetilde{Z}_u \beta_u \triangle S_u]} \left(\frac{1}{2\widetilde{b}_u} - E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})] + E[\widetilde{Z}_u \beta_u \triangle S_u G_{u-1}(\widetilde{\theta})] \right)\\
&= \frac{1}{E[\widetilde{Z}_{u-1}]} \left(-\frac{1}{2\xi} - E[\widetilde{Z}_{u-1} G_{u-1}(\widetilde{\theta})] \right).
\end{aligned}$$

The second-to-last equality cancels out the terms involving $\frac{\widetilde{a}_u E[\widetilde{Z}_u \beta_u \triangle S_u]}{2\widetilde{b}_u}$ and $\widetilde{a}_u E[\widetilde{Z}_u \beta_u \triangle S_u] E[\widetilde{Z}_u G_{u-1}(\widetilde{\theta})]$. The last equality uses $\widetilde{a}_u = \frac{1}{E[\widetilde{Z}_u]}$ from (4.65) and $E[\widetilde{Z}_{t-1}] = E[\widetilde{Z}_t(1 - \beta_t \triangle S_t)]$ from the first equality in (4.64) twice as well as $\widetilde{b}_u = -\xi$ from (4.66). Inserting the above equality into (4.69) yields (4.70).

The expression (4.72) follows from (4.70). To show (4.73), we next argue for $t = 0, 1, \ldots, T$ that

$$\frac{(2\xi)^{-1} + E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]}{E[\widetilde{Z}_t]} = \frac{1}{2\xi E[\widetilde{Z}_0]}. \tag{4.74}$$

Denote by z the right-hand side. For $t = 0$, (4.74) is clear because $G_0(\widetilde{\theta}^{\mathrm{mv}}) = 0$. Suppose the above is true for $t - 1$. We then use the explicit formula (4.72) for $\widetilde{\theta}^{\mathrm{mv}}_t$ with $\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]}$ replaced by β_t due to (4.63) and the induction hypothesis

(4.74) for $t-1$ to get

$$
\begin{aligned}
E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})] &= E\big[\widetilde{Z}_t\big(G_{t-1}(\widetilde{\theta}^{\mathrm{mv}}) + \widetilde{\theta}^{\mathrm{mv}}_t \triangle S_t\big)\big] \\
&= E\Bigg[\widetilde{Z}_t\Bigg(G_{t-1}(\widetilde{\theta}^{\mathrm{mv}}) \\
&\qquad\qquad - \beta_t \triangle S_t\Bigg(G_{t-1}(\widetilde{\theta}^{\mathrm{mv}}) - \frac{(2\xi)^{-1} + E[\widetilde{Z}_{t-1} G_{t-1}(\widetilde{\theta}^{\mathrm{mv}})]}{E[\widetilde{Z}_{t-1}]}\Bigg)\Bigg)\Bigg] \\
&= E\Big[\widetilde{Z}_t\Big(G_{t-1}(\widetilde{\theta}^{\mathrm{mv}}) - \beta_t \triangle S_t\big(G_{t-1}(\widetilde{\theta}^{\mathrm{mv}}) - z\big)\Big)\Big] \\
&= E[\widetilde{Z}_t(1-\beta_t \triangle S_t) G_{t-1}(\widetilde{\theta}^{\mathrm{mv}})] + z E[\widetilde{Z}_t \beta_t \triangle S_t] \\
&= E[\widetilde{Z}_{t-1} G_{t-1}(\widetilde{\theta}^{\mathrm{mv}})] + z E[\widetilde{Z}_t \beta_t \triangle S_t].
\end{aligned}
$$

The last line uses the recursion in the first equality of (4.64) for $\widetilde{Z}_{t-1}$ and the tower property thanks to the boundedness of $\widetilde{Z}$ and $G_{t-1}(\widetilde{\theta}^{\mathrm{mv}}) \in L^2$. This integrability property plus $\widetilde{Z}_t \beta_t \triangle S_t \in L^1$ from (4.26) also ensures that the decomposition in the third line is allowed. Now inserting the above display into (4.74) for t, then using the identity $(2\xi)^{-1} + E[\widetilde{Z}_{t-1} G_{t-1}(\widetilde{\theta}^{\mathrm{mv}})] = z E[\widetilde{Z}_{t-1}]$ deduced from the induction hypothesis (4.74) for $t-1$ and finally invoking from (4.64) the recursion $E[\widetilde{Z}_{t-1}] = E[\widetilde{Z}_t(1-\beta_t \triangle S_t)]$ yields

$$
\begin{aligned}
\frac{(2\xi)^{-1} + E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]}{E[\widetilde{Z}_t]} &= \frac{(2\xi)^{-1} + E[\widetilde{Z}_{t-1} G_{t-1}(\widetilde{\theta}^{\mathrm{mv}})] + z E[\widetilde{Z}_t \beta_t \triangle S_t]}{E[\widetilde{Z}_t]} \\
&= \frac{z E[\widetilde{Z}_{t-1}] + z E[\widetilde{Z}_t \beta_t \triangle S_t]}{E[\widetilde{Z}_t]} \\
&= z.
\end{aligned}
$$

This completes the induction step and justifies the claim (4.74) for $t = 0, 1, \dots, T$.

Finally, the MVPS problem (4.2) is the same as the problem (4.55) at time 0 thanks to $\Theta_{\mathrm{S}}(0,\theta) = \Theta_{\mathrm{S}}$. The strategy $\widetilde{\theta}^{\mathrm{mv}}$ is simply $\widetilde{\theta}$ in (4.69) starting from $u = 1$. □

4.7 Discussion and special cases

In this subsection, we provide some concrete examples for the MVPS problem with $\Theta = \Theta_{\mathrm{S}}$. To this end, we need to verify (in these examples) the assumptions for Theorem 4.11, among which Assumption 4.10 is the main focus. The presentation below is divided into three parts. The first is a bottom-up and technical discussion of Assumption 4.10, while the second is more top-down and abstract. In the last part, we give examples where some explicit computations are possible.

For the sake of the argument, let us suppose that Assumptions 4.2 and 4.5 are satisfied, meaning that S is square-integrable and satisfies the structure condition (SC), and the L^2-closure of $G_T(\Theta_S)$ does not contain the constant (payoff) 1. Recall that Assumption 4.10 says that the solution $\widetilde{\delta}_t$ to the linear equation (4.40) satisfies $\widetilde{\delta}_t \triangle S_t \in L^2$ for all $t = 1, \ldots, T$ and $\theta \in \Theta_S$. For convenience, we recall from (4.41) that $\widetilde{\delta}_t$ is given explicitly by

$$\widetilde{\delta}_t = -\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \left(\frac{\widetilde{a}_t}{2\widetilde{b}_t} + G_{t-1}(\theta) + \frac{\widetilde{c}_t}{\widetilde{b}_t} \big(E[\widetilde{Z}_t G_{t-1}(\theta)] + \widetilde{e}_t \big) \right), \tag{4.75}$$

where $\widetilde{e}_t$ is a real number given explicitly in (4.42). Because all quantities inside the large parenthesis except for $G_{t-1}(\theta)$ in (4.75) are deterministic, we need $\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \triangle S_t G_{t-1}(\theta) \in L^2$. Due to $G_{t-1}(\theta) \in L^2$ by $\theta \in \Theta_S$, we might expect that we need $\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \triangle S_t$ or $E[(\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \triangle S_t)^2 | \mathcal{F}_{t-1}]$ to be in L^∞. We can give some partial results in this direction.

Lemma 4.12. *Suppose that Assumption* 4.2 *is satisfied and U is a bounded random variable. If $U \geq \ell$ for some $\ell > 0$, then we have*

$$E\left[\left(\frac{E[U \triangle S_t | \mathcal{F}_{t-1}]}{E[U (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \triangle S_t \right)^2 \middle| \mathcal{F}_{t-1} \right] \leq \left(\frac{\|U\|_\infty}{\ell} \right)^2, \quad t = 1, \ldots, T. \tag{4.76}$$

Proof. First we note that $\frac{E[U \triangle S_t | \mathcal{F}_{t-1}]}{E[U (\triangle S_t)^2 | \mathcal{F}_{t-1}]}$ is well defined by the positivity of U, the Cauchy–Schwarz inequality and the convention $\frac{0}{0} = 0$. Because U is bounded and satisfies $U \geq \ell$, we have

$$\left(\frac{E[U \triangle S_t | \mathcal{F}_{t-1}]}{E[U (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \triangle S_t \right)^2 \leq \left(\frac{\|U\|_\infty}{\ell} \right)^2 \left(\frac{E[\triangle S_t | \mathcal{F}_{t-1}]}{E[(\triangle S_t)^2 | \mathcal{F}_{t-1}]} \triangle S_t \right)^2.$$

Taking conditional expectations with respect to $\mathcal{F}_{t-1}$ on both sides and using the tower property and the Cauchy–Schwarz inequality yields the desired bound. □

Lemma 4.13. *Suppose that Assumptions* 4.2 *and* 4.5 *are satisfied. If the process $\widetilde{Z}$ given by* (4.7) *is uniformly bounded below by a positive real number, then Assumption* 4.10 *is satisfied.*

Proof. Thanks to the explicit formula (4.75) for $\widetilde{\delta}_t$, we only need to show that

$$\left(\frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t (\triangle S_t)^2 | \mathcal{F}_{t-1}]} \triangle S_t G_{t-1}(\theta) \right)^2 \in L^1.$$

We write $\widetilde{Z}_t \geq \widetilde{z} > 0$ by the assumption that $\widetilde{Z}$ is uniformly lower bounded by a positive real number. Then using (4.76) with $(U, \ell) = (\widetilde{Z}_t, \widetilde{z})$ and taking expectations yields

$$\begin{aligned} E\bigg[\Big(\frac{E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t G_{t-1}(\theta)\Big)^2\bigg] &= E\bigg[E\Big[\Big(\frac{E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\triangle S_t\Big)^2\Big|\mathcal{F}_{t-1}\Big] \\ &\qquad \times \big(G_{t-1}(\theta)\big)^2\bigg] \\ &\leq \bigg(\frac{\|\widetilde{Z}_t\|_\infty}{\widetilde{z}}\bigg)^2 E\big[\big(G_{t-1}(\theta)\big)^2\big] \\ &< \infty. \end{aligned}$$

The last inequality uses that $G_{t-1}(\theta) \in L^2$ by $\theta \in \Theta_{\mathrm{S}}$, the definition (4.3) of Θ_{S} and $0 \leq \widetilde{Z}_t \leq 1$ from Lemma 4.3. □

Next, we turn to giving a concrete example such that $\widetilde{Z}$ is uniformly bounded from below. We first introduce some terminology. Let $\mathbb{P}^2_{\mathrm{e}}(S)$ be the set of all probability measures Q equivalent to P on $\mathcal{F}_T$ such that the Radon–Nikodým derivative $\frac{\mathrm{d}Q}{\mathrm{d}P}$ is in $L^2(P)$ and S is a Q-martingale. The density process of Q with respect to P is denoted by $Z^Q := (Z^Q_t)_{t=0,1,\dots,T}$.

Lemma 4.14. *If there exists $Q \in \mathbb{P}^2_{\mathrm{e}}(S)$ which satisfies the reverse Hölder inequality $R_2(P)$, meaning that there exists a constant $C > 0$ such that*

$$E[(Z^Q_T)^2|\mathcal{F}_t] \leq C(Z^Q_t)^2, \quad t = 0, 1, \dots, T, \tag{4.77}$$

then the process $\widetilde{Z}$ is uniformly bounded from below by a positive real number, and Assumption 4.10 *is satisfied.*

Proof. This is shown in Lemma 2.1 in Jeanblanc et al. [38]. Our process $\widetilde{Z}$ corresponds to the process q there. Indeed, that process q is shown in [38, Theorem 2.4] to be equal to a solution Y to the BSDE (2.18) in [38]. Later in the last displayed equation in [38, Section 5.2], the authors work out the recursion for Y in finite discrete time. That recursion is the same as the recursion (4.58) for our process $\widetilde{Z}$. □

The above result is not completely satisfactory because its assumption is not described explicitly in terms of the price process S. To improve this, we use Assumption 4.2 and apply Doob's decomposition for S in the filtration $\mathbb{F}$ to obtain a square-integrable martingale M and a square-integrable predictable process A, both null at 0 and both with respect to $\mathbb{F}$, such that $S = S_0 + M + A$. For an explicit expression, we refer to (I.3.7). Also recall from (I.3.8) and (I.3.9) the two bracket notations for any square-integrable process X, namely

$$[X]_0 := 0, \quad \triangle[X]_t := (\triangle X_t)^2, \quad t = 1, \dots, T,$$
$$\langle X \rangle_0 := 0, \quad \triangle\langle X \rangle_t := E\big[\triangle[X]_t \big| \mathcal{F}_{t-1}\big] = E[(\triangle X_t)^2 | \mathcal{F}_{t-1}], \quad t = 1, \dots, T.$$

The mean–variance tradeoff (MVT) process $K \geq 0$ of S defined in Definition I.3.5 has its increments explicitly given by

$$\lambda_t = \frac{\triangle A_t}{\triangle\langle M \rangle_t}, \quad \triangle K_t = \lambda_t \triangle A_t = \frac{(\triangle A_t)^2}{\triangle\langle M \rangle_t}, \quad t = 1, \dots, T. \tag{4.78}$$

Using this definition and the identity $E[(\triangle S_t)^2 | \mathcal{F}_{t-1}] = \triangle\langle M \rangle_t + (\triangle A_t)^2$ for $t = 1, \dots, T$, we can also write for $t = 0, 1, \dots, T$ that

$$\frac{E[\triangle S_t | \mathcal{F}_{t-1}]}{E[(\triangle S_t)^2 | \mathcal{F}_{t-1}]} = \frac{\lambda_t}{1 + \triangle K_t}, \quad \frac{(E[\triangle S_t | \mathcal{F}_{t-1}])^2}{E[(\triangle S_t)^2 | \mathcal{F}_{t-1}]} = \frac{\triangle K_t}{1 + \triangle K_t}. \tag{4.79}$$

Lemma 4.15. *Suppose Assumptions* 4.2 *and* 4.5, 2) *are satisfied. If the MVT process K is uniformly bounded and $\lambda_t \triangle M_t < 1$ for $t = 1, \dots, T$, then Assumption* 4.10 *is satisfied.*

Proof. In view of Lemma 4.14, we construct a measure $Q \in \mathbb{P}^2_{\mathrm{e}}(S)$ that satisfies the reverse Hölder inequality $R_2(P)$. Define Q via $\frac{\mathrm{d}Q}{\mathrm{d}P} := Z^Q_T$ with

$$Z^Q_0 = 1, \quad Z^Q_t = Z^Q_{t-1}(1 - \lambda_t \triangle M_t) = \prod_{s=1}^{t} (1 - \lambda_s \triangle M_s), \qquad t = 1, \dots, T.$$

The proof of Lemma I.3.6 shows that Q is a signed measure such that $Q[\Omega] = 1$, $Q \ll P$ with $\frac{\mathrm{d}Q}{\mathrm{d}P} \in L^2$ and $E[Z^Q_T G_T(\theta)] = 0$ for all $\theta \in \Theta_{\mathrm{S}}$. The equality (I.3.27) proved there also shows for $t = 1, \dots, T$ that

$$E[(Z^Q_t)^2 | \mathcal{F}_{t-1}] = E\big[(Z^Q_{t-1})^2 E[(1 - \lambda_t \triangle M_t)^2 | \mathcal{F}_{t-1}]\big] = (Z^Q_{t-1})^2 (1 + \triangle K_t).$$

Iterative conditioning and using the above identity together with $0 \leq \triangle K_u \leq K_u$ repeatedly yields

$$E[(Z^Q_T)^2 | \mathcal{F}_t] \leq (Z^Q_t)^2 \prod_{u=t+1}^{T} (1 + \|K_u\|_\infty)$$

and hence the desired inequality (4.77). □

Now we begin the second part of this subsection. Let us look at the role of the L^2-integrability of $\widetilde{\delta}_t \triangle S_t$ from a different angle. Fix $\theta \in \Theta_{\mathrm{S}}$. In Proposition 4.9 or Lemma 3.7, that integrability is used to ensure that the candidate maximiser $\widetilde{\delta}_t$ for the map $\delta_t \mapsto r_t^{\mathrm{mv}}(\delta_t)$, or equivalently for the map $\delta_t \mapsto v_t^{\mathrm{mv}}(t, \theta(t, \delta_t))$, is still in $\Theta_{\mathrm{S}}^{[t]}(\theta)$, which is in turn used to argue that the map $\delta_t \mapsto v_t^{\mathrm{mv}}(t, \theta(t, \delta_t))$ has a maximiser in $\Theta_{\mathrm{S}}^{[t]}(\theta)$. But if the existence of an optimiser $\widetilde{\delta}_t$ (in $\Theta_{\mathrm{S}}^{[t]}(\theta)$) for the map $\delta_t \mapsto v_t^{\mathrm{mv}}(t, \theta(t, \delta_t))$ is known in advance, then it must solve the linear equation (4.40) and yields $\widetilde{\delta}_t$ given by (4.75) again plus $\widetilde{\delta}_t \triangle S_t \in L^2$.

Lemma 4.16. *Suppose that Assumptions* 4.2 *and* 4.5 *are satisfied. Then the following statements hold:*

1) *If the map* $\delta_t \mapsto v^{\mathrm{mv}}(t, \theta(t, \delta_t))$ *has a maximiser in* $\Theta_{\mathrm{S}}^{[t]}(\theta)$ *for* $t = 1, \dots, T$, *then Assumption* 4.10 *in Theorem* 4.11 *can be dropped.*

2) *If the MVT tradeoff process* K *is bounded, then the map* $\delta_t \mapsto v^{\mathrm{mv}}(t, \theta(t, \delta_t))$ *has a maximiser in* $\Theta_{\mathrm{S}}^{[t]}(\theta)$ *for* $t = 1, \dots, T$.

Proof. a) Part 1) is clear from the discussion preceding Lemma 4.16. We only argue part 2) here. Recall from the proof of Theorem I.2.4 that the functional $F(g) = E[g] - \xi \mathrm{Var}[g]$ has a maximiser in $G_T(\Theta_{\mathrm{S}})$ as soon as $G_T(\Theta_{\mathrm{S}})$ is closed in L^2 and does not contain the constant payoff 1. Fix $\theta \in \Theta_{\mathrm{S}}$. Recall by the definitions (2.5) and (2.6) of v^{mv} (which is there called v only; see the beginning of Section 4.2) that we have

$$v^{\mathrm{mv}}(t, \theta) = \sup_{\widetilde{\theta} \in \Theta(t,\theta)} j^{\mathrm{mv}}(\widetilde{\theta}) = \sup_{\widetilde{\theta} \in \Theta(t,\theta)} v^{\mathrm{mv}}(T, \widetilde{\theta}) = \sup_{\widetilde{\theta} \in \Theta(t,\theta)} F\big(G_T(\widetilde{\theta})\big). \tag{4.80}$$

For convenience, we also recall from (2.3) that

$$\Theta_{\mathrm{S}}(t, \theta) = \{\widetilde{\theta} \in \Theta_{\mathrm{S}} : \widetilde{\theta} = \theta \text{ on } [\![0, t]\!] \cap \mathbb{N}\}. \tag{4.81}$$

The same proof as for Theorem I.2.4 implies that the functional F also has a maximiser in the space

$$\Psi_t := G_t(\theta) + G_{t,T}\big(\Theta_{\mathrm{S}}(t, \theta)\big) := \left\{ G_t(\theta) + \sum_{u=t+1}^{T} \widetilde{\theta}_u \triangle S_u : \widetilde{\theta} \in \Theta_{\mathrm{S}}(t, \theta) \right\},$$

because the space Ψ_t is also closed in L^2 and does not contain the constant payoff 1. This is argued in step b) below. Using this observation and continuing from

(4.80), we can write

$$\begin{aligned} v^{\mathrm{mv}}\big(t, \theta(t, \delta_t)\big) &= \sup_{\widetilde{\theta} \in \Theta(t, \theta(t, \delta_t))} F\big(G_T(\widetilde{\theta})\big) \\ &= F\Big(G_t\big(\theta(t, \delta_t)\big) + \sum_{u=t+1}^{T} \theta^*_u \triangle S_u\Big) \\ &= F\Big(G_T\big(\theta^*(\delta_t)\big)\Big) \end{aligned}$$

for a maximiser $\theta^*(\delta_t) = (\theta_1, \dots, \theta_{t-1}, \delta_t, \theta^*_{t+1}, \dots, \theta^*_T)$. So maximising the map $\delta_t \mapsto v^{\mathrm{mv}}(t, \theta(t, \delta_t))$ is the same as maximising $\delta_t \mapsto F(G_T(\theta^*(\delta_t)))$. By a completely analogous argument as in step b) for fixed $\theta \in \Theta_{\mathrm{S}}$, $t \in \{1, \dots, T\}$ and the space $\Psi_t = G_t(\theta) + G_{t,T}(\Theta_{\mathrm{S}}(t, \theta))$, we obtain that the space

$$\big\{G_T\big(\theta^*(\delta_t)\big) : \delta_t \in \Theta^{[t]}_{\mathrm{S}}(\theta)\big\}$$

is closed in L^2 and does not contain the payoff 1. So the same proof as for Theorem I.2.4 yields the desired conclusion.

b) To argue that Ψ_t is closed in L^2, take a sequence $(g^n)_{n \in \mathbb{N}}$ in Ψ_t which converges in L^2 to some g^∞. We need to show that $g^\infty \in \Psi_t$. Because the MVT process K is bounded, the proof of Schweizer [61, Theorem 2.1] shows that $G_{t,T}(\Theta_{\mathrm{S}}) = \{\sum_{u=t+1} \widetilde{\theta}_u \triangle S_u : \widetilde{\theta} \in \Theta_{\mathrm{S}}\}$ is closed in L^2 for $t = 0, 1, \dots, T$. Because $g^n \in \Psi_t \subseteq G_T(\Theta_{\mathrm{S}})$ for $n \in \mathbb{N}$ and $G_T(\Theta_{\mathrm{S}}) = G_{0,T}(\Theta_{\mathrm{S}})$ is closed in L^2, we get $g^\infty \in G_T(\Theta_{\mathrm{S}})$. We now claim that $g^\infty = g^\infty_t + g^\infty_{t,T}$ for $g^\infty_t = G_t(\theta)$ and some $g^\infty_{t,T} \in G_{t,T}(\Theta_S(t, \Theta))$. To see this, let us write $g^n = g^n_t + g^n_{t,T}$ for $g^n_t = G_t(\theta)$, $g^n_{t,T} \in G_{t,T}(\Theta_{\mathrm{S}}(t, \theta))$, $n \in \mathbb{N}$. Because $g^n_t = G_t(\theta)$ for all $n \in \mathbb{N}$, we clearly have $g^n_t \to g^\infty_t = G_t(\theta)$ in L^2. Combining this with $g^n \to g^\infty$ in L^2, we get that $g^n_{t,T} = g^n - g^n_t \to g^\infty - G_t(\theta)$ in L^2. But $g^n_{t,T} \in G_{t,T}(\Theta_{\mathrm{S}}(t, \theta)) = G_{t,T}(\Theta_{\mathrm{S}})$ and $G_{t,T}(\Theta_{\mathrm{S}})$ is closed in L^2. Hence we get $g^n_{t,T} \to h_{t,T} = \sum_{u=t+1}^{T} \varphi_u \triangle S_u$ for some $\mathcal{F}_{u-1}$-measurable φ_u with $\varphi_u \triangle S_u \in L^2$ for $u = t+1, \dots, T$. Now

$$g^\infty = \lim_{n \to \infty} (g^n_t + g^n_{t,T}) = G_t(\theta) + h_{t,T},$$

and the strategy $(\theta_1, \dots, \theta_t, \varphi_{t+1}, \dots, \varphi_T)$ is in $\Theta_{\mathrm{S}}(t, \theta)$; see (4.81). This proves that Ψ_t is closed in L^2. Finally, the inclusion $\Psi_t \subseteq G_T(\Theta_{\mathrm{S}})$ shows that $\overline{\Psi_t}$ does not contain the constant 1 due to Assumption 4.5, 1).

□

Corollary 4.17. *Suppose that Assumptions* 4.2 *and* 4.5 *are satisfied. If the MVT* *tradeoff process K is uniformly bounded, then the conclusion of Theorem* 4.11 *holds without Assumption* 4.10.

Proof. The result follows directly from Lemma 4.16. □

Finally, we turn to some explicit examples. If the entire MVT process K is deterministic, we obtain an explicit expression for $\widetilde{Z}$ and the optimal strategy $\widetilde{\theta}^{\mathrm{mv}}$. This also recovers the result in Theorem I.5.7 in Chapter I.

Theorem 4.18. *Suppose Assumptions* 4.2 *and* 4.5 *are satisfied, meaning that S is square-integrable and satisfies the structure condition (SC). If the MVT process K is deterministic, then the following statements hold:*

1) *The process $\widetilde{Z}$ is also deterministic and explicitly given, for $t = 0, 1, \dots, T$, by*

$$\widetilde{Z}_t = \prod_{u=t+1}^{T} \frac{1}{1 + \triangle K_u} = \frac{\mathcal{E}(K)_t}{\mathcal{E}(K)_T}. \tag{4.82}$$

2) *The optimal strategy $\widetilde{\theta}^{\mathrm{mv}}$ for the MVPS problem* (4.2) *with $\Theta = \Theta_{\mathrm{S}}$ is given by*

$$\begin{aligned}
\widetilde{\theta}^{\mathrm{mv}}_u &= -\frac{E[\triangle S_u | \mathcal{F}_{u-1}]}{E[(\triangle S_u)^2 | \mathcal{F}_{u-1}]} \left(G_{u-1}(\widetilde{\theta}^{\mathrm{mv}}) - E[G_{u-1}(\widetilde{\theta}^{\mathrm{mv}})] - \frac{1}{2\xi \widetilde{Z}_{u-1}} \right) &(4.83)\\
&= \frac{E[\triangle S_u | \mathcal{F}_{u-1}]}{E[(\triangle S_u)^2 | \mathcal{F}_{u-1}]} \left(G_{u-1}(\widetilde{\theta}^{\mathrm{mv}}) - \frac{1}{2\xi \widetilde{Z}_0} \right) &(4.84)\\
&= \frac{\lambda_u}{1 + \triangle K_u} \left(G_{u-1}(\widetilde{\theta}^{\mathrm{mv}}) - \frac{1}{2\xi \widetilde{Z}_0} \right), \qquad u = 1, \dots, T. &(4.85)
\end{aligned}$$

3) *The value function $v^{\mathrm{mv}}(t, \widetilde{\theta}^{\mathrm{mv}})$ is explicitly given, for $t = 0, 1, \dots, T$, by*

$$v^{\mathrm{mv}}(t, \widetilde{\theta}^{\mathrm{mv}}) = E[G_t(\widetilde{\theta}^{\mathrm{mv}})] - \xi \widetilde{Z}_t \mathrm{Var}[G_t(\widetilde{\theta}^{\mathrm{mv}})] + \frac{1}{4\xi} \left(\frac{1}{\widetilde{Z}_t} - 1 \right). \tag{4.86}$$

Consequently, the optimal strategy $\widetilde{\theta}^{\mathrm{mv}}$ and the value function v^{mv} coincide with the strategy $\widehat{\theta}$ and the process $\widetilde{V}$ given in (I.5.32) *and* (I.5.37), *respectively.*

Proof. 1) We prove this assertion by backward induction. For $t = T$, (4.82) is clear because the product over an empty set is 1. Suppose that $\widetilde{Z}_t$ is deterministic. Then by the recursive formula (4.58) for $\widetilde{Z}$, non-randomness of $\widetilde{Z}_t$ and the second equality in (4.79), we obtain

$$\begin{aligned}\widetilde{Z}_{t-1} &= E[\widetilde{Z}_t|\mathcal{F}_{t-1}] - \frac{(E[\widetilde{Z}_t \triangle S_t|\mathcal{F}_{t-1}])^2}{E[\widetilde{Z}_t(\triangle S_t)^2|\mathcal{F}_{t-1}]}\\ &= \widetilde{Z}_t\left(1 - \frac{(E[\triangle S_t|\mathcal{F}_{t-1}])^2}{E[(\triangle S_t)^2|\mathcal{F}_{t-1}]}\right)\\ &= \widetilde{Z}_t \frac{1}{1+\triangle K_t}.\end{aligned}$$

Because $\triangle K_t$ is deterministic by assumption, we get that $\widetilde{Z}_{t-1}$ is deterministic as well. This finishes the induction step and also justifies the first equality in (4.82). The second equality directly follows from the definition of the stochastic exponential $\mathcal{E}$ that $\mathcal{E}(K)_t = \prod_{u=1}^t (1+\triangle K_u)$.

2) Because $\widetilde{Z}$ is deterministic by part 1), the submartingale property of $\widetilde{Z}$ from Corollary 4.4 implies that $\widetilde{Z}$ is increasing. This in particular yields that $\widetilde{Z}_t \geq \widetilde{Z}_0 > 0$ and Assumption 4.10 is satisfied. We recall from Lemma I.3.6 that $G_T(\Theta_{\mathrm{S}})$ is closed in L^2 and does not contain the constant payoff 1 whenever the MVT process K is bounded. So the non-randomness of K implies that the L^2-closure of $G_T(\Theta_{\mathrm{S}})$ (equals $G_T(\Theta_{\mathrm{S}})$ and) does not contain the constant payoff 1, which means that Assumption 4.5, 1) is satisfied. Therefore, we can apply Theorem 4.11 and use that $\widetilde{Z}$ is deterministic to simplify the optimal strategy $\widetilde{\theta}$ given in (4.72) and (4.73) to obtain (4.83) and (4.84). The equality (4.85) is then due to the first equality in (4.79).

3) Recall from Theorem 4.11 that the value function $v^{\mathrm{mv}}(t, \widetilde{\theta}^{\mathrm{mv}})$ is given by

$$\begin{aligned}v^{\mathrm{mv}}(t,\widetilde{\theta}^{\mathrm{mv}}) = E\big[&\widetilde{a}_t\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}}) + \widetilde{b}_t\widetilde{Z}_t\big(G_t(\widetilde{\theta}^{\mathrm{mv}})\big)^2\big]\\ &+ \widetilde{c}_t\big(E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]\big)^2 + \widetilde{d}_t.\end{aligned} \tag{4.87}$$

We then use the explicit formulas for $\widetilde{a}_t, \widetilde{b}_t, \widetilde{c}_t, \widetilde{d}_t$ from (4.65)–(4.68) and the non-randomness of $\widetilde{Z}_t$ to obtain for $t = 0, 1, \dots, T$ that

$$\begin{aligned}\widetilde{a}_t\widetilde{Z}_t &= \frac{\widetilde{Z}_t}{E[\widetilde{Z}_t]} = 1,\\ \widetilde{b}_t\widetilde{Z}_t &= -\xi\widetilde{Z}_t,\\ \widetilde{c}_t\widetilde{Z}_t &= \frac{\xi}{E[\widetilde{Z}_t]}\,\widetilde{Z}_t = \xi\widetilde{Z}_t\\ \widetilde{d}_t &= \frac{1}{4\xi}\left(\frac{1}{\widetilde{Z}_t} - 1\right).\end{aligned}$$

Inserting these expressions into (4.87) yields (4.86).

Finally, let us recall from (I.5.32) and (I.5.37) in Chapter I that

$$\widehat{\theta}_t = \frac{\lambda_t}{1+\triangle K_t}\left(\frac{1}{2b_{t-1}^{(\infty)}} - G_{t-1}(\widehat{\theta}) + E[G_{t-1}(\widehat{\theta})]\right), \quad t = 1, \dots, T, \tag{4.88}$$

and

$$\widetilde{V}_t = a_t^{(\infty)} E[G_t(\widehat{\theta})] - b_t^{(\infty)} \mathrm{Var}[G_t(\widehat{\theta})] + c_t^{(\infty)}, \quad t = 0, 1, \dots, T, \tag{4.89}$$

where the coefficients $a^{(\infty)}, b^{(\infty)}, c^{(\infty)}$ are given by

$$a_t^{(\infty)} = 1, \qquad t = 0, 1, \dots, T, \tag{4.90}$$

$$b_t^{(\infty)} = \xi \prod_{u=t+1}^{T} \frac{1}{1+\triangle K_u}, \qquad t = 0, 1, \dots, T, \tag{4.91}$$

$$c_t^{(\infty)} = \sum_{u=t+1}^{T} \frac{\triangle K_u}{4b_u^{(\infty)}}, \qquad t = 0, 1, \dots, T. \tag{4.92}$$

To prove that $\widetilde{\theta}^{\mathrm{mv}}$ agrees with $\widehat{\theta}$ given in (4.88), we use that $\frac{E[\triangle S_t|\mathcal{F}_{t-1}]}{E[(\triangle S_t)^2|\mathcal{F}_{t-1}]}$ agrees with $\frac{\lambda_t}{1+\triangle K_t}$ by (4.79) and from (4.82) and (4.91) that $\xi\widetilde{Z}_t$ agrees with $b_t^{(\infty)}$. Then a comparison of (4.83) with (4.88) yields that $\widetilde{\theta}^{\mathrm{mv}}$ and $\widehat{\theta}$ satisfy the same recursion and start from $G_0(\widetilde{\theta}^{\mathrm{mv}}) = 0 = G_0(\widehat{\theta})$. This yields the claim. To prove that v^{mv} agrees with $\widetilde{V}$, we first note that their first two terms are the same. By using $b_t^{(\infty)} = \xi\widetilde{Z}_t$ and the explicit formula (4.82) for $\widetilde{Z}$ repeatedly, we obtain in (4.92) that

$$\begin{aligned} c_t^{(\infty)} &= \sum_{u=t+1}^{T} \frac{\triangle K_u}{4b_u^{(\infty)}} \\ &= \frac{1}{4\xi} \sum_{u=t+1}^{T} \frac{\triangle K_u}{\widetilde{Z}_u} \\ &= \frac{1}{4\xi} \sum_{u=t+1}^{T} \triangle K_u \prod_{v=u+1}^{T} (1+\triangle K_v) \\ &= \frac{1}{4\xi}\left(\frac{1}{\widetilde{Z}_t} - 1\right). \end{aligned}$$

The last equality uses the elementary identity

$$1+\sum_{u=t+1}^{T} x_u \prod_{v=u+1}^{T}(1+x_v)=\prod_{u=t+1}^{T}(1+x_u)$$

and again (4.82). □

We end this subsection with a concrete example.

Example 4.19. Suppose that the price process S has independent returns. More precisely, this means that it satisfies the dynamics

$$S_t = S_0\,\mathcal{E}(R)_t = S_0\prod_{s=1}^{t}(1+\triangle R_s), \quad t=0,1,\dots,T,$$

with all the $\triangle R_t > -1$ and the random variables $\triangle R_1,\dots,\triangle R_T$ are independent. Let us compute the MVT increment $\triangle K$ for S. By independence, we get

$$E[\triangle S_t|\mathcal{F}_{t-1}] = E[S_{t-1}(1+\triangle R_t)-S_{t-1}|\mathcal{F}_{t-1}] = S_{t-1}E[\triangle R_t] \tag{4.93}$$

and

$$\begin{aligned} E[(\triangle S_t)^2|\mathcal{F}_{t-1}] &= E[S_{t-1}^2(1+\triangle R_t)^2 - 2S_{t-1}^2(1+\triangle R_t)+S_{t-1}^2|\mathcal{F}_{t-1}] \\ &= S_{t-1}^2E[(\triangle R_t)^2]. \end{aligned} \tag{4.94}$$

From the identity $\triangle\langle M\rangle_t = E[(\triangle S_t)^2|\mathcal{F}_{t-1}]-(E[\triangle S_t|\mathcal{F}_{t-1}])^2$ and the definition (4.78) for $\triangle K_t$, we therefore obtain

$$\triangle K_t = \frac{(E[\triangle S_t|\mathcal{F}_{t-1}])^2}{E[(\triangle S_t)^2|\mathcal{F}_{t-1}]-(E[\triangle S_t|\mathcal{F}_{t-1}])^2} = \frac{S_{t-1}^2(E[\triangle R_t])^2}{S_{t-1}^2\mathrm{Var}[\triangle R_t]} = \frac{(E[\triangle R_t])^2}{\mathrm{Var}[\triangle R_t]},$$

which is deterministic. Thus we can apply Theorem 4.18 to get an explicit formula for $\widetilde{\theta}^{\mathrm{mv}}$. Let us set $\mu_t = E[\triangle R_t]$ and $\sigma_t^2 = \mathrm{Var}[\triangle R_t]$. Using the explicit formula (4.82), we get

$$\widetilde{Z}_t = \prod_{u=t+1}^{T}\frac{1}{1+\frac{\mu_u^2}{\sigma_u^2}} = \prod_{u=t+1}^{T}\frac{\sigma_u^2}{\sigma_u^2+\mu_u^2}.$$

Inserting (4.93), (4.94) and the above formula into (4.84), we get that the optimal strategy $\widetilde{\theta}^{\text{mv}}$ is given by

$$\widetilde{\theta}_t^{\text{mv}} = -\frac{\mu_t}{S_{t-1}(\sigma_t^2 + \mu_t^2)} \left(G_{t-1}(\widetilde{\theta}^{\text{mv}}) - \frac{1}{2\xi} \prod_{u=1}^{T} \frac{\sigma_u^2 + \mu_u^2}{\sigma_u^2} \right), \quad t = 1, \dots, T.$$

This agrees with (4.19) in Pham and Wei [52]. Indeed, Pham and Wei work with i.i.d. returns so that we can set $\mu_t \equiv \mu$ and $\sigma_t \equiv \sigma$ and identify the relevant parameters to recover their formula with initial wealth equal to 0.

4.8 More general strategies

In this subsection, we argue that the optimal strategy $\widetilde{\theta}^{\text{mv}}$ for the MVPS problem (4.2) with $\Theta = \Theta_{\text{S}}$ obtained in Theorem 4.11 is also optimal in the larger set Θ_{MN} of strategies, which only imposes that the final gains are square-integrable. This space has been considered in Melnikov and Nechaev [50]. The MVPS problem (4.2) with $\Theta = \Theta_{\text{MN}}$ is to

$$\text{maximise } j^{\text{mv}}(\theta) = E[G_T(\theta)] - \xi \text{Var}[G_T(\theta)] \text{ over all } \theta \in \Theta_{\text{MN}}, \tag{4.95}$$

where

$$\begin{aligned} \Theta_{\text{MN}} = \{\theta := (\theta_t)_{t=1,\dots,T} : \theta \text{ is real-valued, } \mathbb{F}\text{-predictable} \\ \text{and } G_T(\theta) \in L^2\}. \end{aligned} \tag{4.96}$$

Comparing (4.3) with (4.96), we immediately see that

$$\Theta_{\text{S}} \subseteq \Theta_{\text{MN}}. \tag{4.97}$$

Remark 4.20. 1) Note that Θ_{MN} is the most general set Θ of strategies such that the MVPS criterion $j^{\text{mv}}(\theta)$ in (4.95) is finite for $\theta \in \Theta$. Moreover, the space $G_T(\Theta_{\text{MN}})$ is closed in L^2. Indeed, take a sequence $(\theta^n)_{n \in \mathbb{N}}$ in Θ_{MN} such that $g_n = G_T(\theta^n)$ converges to g in L^2. We now argue that $g = G_T(\theta^\infty)$ for some $\theta^\infty \in \Theta_{\text{MN}}$. Consider the space

$$\Theta_{\text{A}} := \{\theta = (\theta_t)_{t=1,\dots,T} : \theta \text{ is real-valued and } \mathbb{F}\text{-predictable}\}.$$

We know from Remark 1 in Chapter 2 of Kabanov and Safarian [39] that $G_T(\Theta_{\text{A}})$ is always closed in the space L^0 of all random variables equipped with the topology of convergence in probability. This result (specific to finite discrete time) is known as Stricker's lemma. Because $G_T(\Theta_{\mathrm{MN}}) \subseteq G_T(\Theta_{\mathrm{A}})$ and $g_n \to g$ in L^2 implies $g_n \to g$ in L^0, we get from the L^0-closedness of $G_T(\Theta_{\mathrm{A}})$ that $g = G_T(\theta^\infty)$ for some $\theta^\infty \in \Theta_{\mathrm{A}}$. To show $\theta^\infty \in \Theta_{\mathrm{MN}}$, we use $g_n \in L^2$ for all $n \in \mathbb{N}$ by (4.96) and $g_n \to g$ in L^2 to obtain $G_T(\theta^\infty) = g \in L^2$.

2) With the extra assumption that the L^2-closure of $G_T(\theta_{\mathrm{MN}})$, which of course equals $G_T(\Theta_{\mathrm{MN}})$ by 1), does not contain the constant payoff 1, Θ_{MN} satisfies Assumption I.2.2 under which the MVPS problem always has an optimiser by Theorem I.2.4. More precisely, that result tells us that the MVPS problem (4.95) has a solution in Θ_{MN}.

Recall that our optimal strategy $\widetilde{\theta}^{\mathrm{mv}}$ is obtained by solving a linear equation derived from the first order condition (FOC) with respect to a one-step variable δ_t at each $t = 1, \dots, T$. The idea to prove the optimality of $\widetilde{\theta}^{\mathrm{mv}}$ in Θ_{MN} is to obtain an FOC now with respect to the final gains $G_T(\theta)$. By the linear–quadratic structure of the MVPS criterion $j^{\mathrm{mv}}(\theta)$, satisfying that FOC is equivalent to being an optimiser of the MVPS problem (4.95). So it is enough to verify that $G_T(\widetilde{\theta}^{\mathrm{mv}})$ satisfies that FOC. Moreover, we remove Assumption 4.10 because we only need $\widetilde{\theta}^{\mathrm{mv}}$ to lie in Θ_{MN} rather than Θ_{S}. This leads to a programme we implement in the rest of this subsection.

Recipe 4.21. 1) Derive a first order equation for optimality of some $G_T(\widehat{\theta})$ in $G_T(\Theta_{\mathrm{MN}})$ with respect to the final gains.

2) Show that the final gain $G_T(\widetilde{\theta}^{\mathrm{mv}})$ of $\widetilde{\theta}^{\mathrm{mv}}$ lies in L^2 without Assumption 4.10 and satisfies the FOC in 1).

Lemma 4.22. *Suppose that Assumption* 4.2 *is satisfied. If* $\widehat{\theta} \in \Theta_{\mathrm{MN}}$ *satisfies*

$$E\big[\big(1 - 2\xi G_T(\widehat{\theta}) + 2\xi E[G_T(\widehat{\theta})]\big) G_T(\eta)\big] = 0, \quad \forall \eta \in \Theta_{\mathrm{MN}}, \tag{4.98}$$

or equivalently

$$E\big[\big(1 - 2\xi G_T(\widehat{\theta}) + 2\xi E[G_T(\widehat{\theta})]\big) \triangle S_t \big| \mathcal{F}_{t-1}\big] = 0, \quad t = 1, \dots, T, \tag{4.99}$$

then $\widehat{\theta}$ *is an optimal strategy for the MVPS problem with* $\Theta = \Theta_{\mathrm{MN}}$.

Proof. We first show that (4.98) yields the optimality for the MVPS problem (4.95). Let $\widehat{\theta} \in \Theta_{\mathrm{MN}}$ satisfy (4.98). For any $\theta \in \Theta_{\mathrm{MN}}$, we use the definition (4.95) of j^{mv}, the FOC (4.98) with $\eta = \theta - \widehat{\theta}$ and the Cauchy–Schwarz inequality to obtain

$$\begin{aligned}
j^{\mathrm{mv}}(\theta) &= j^{\mathrm{mv}}(\theta - \widehat{\theta} + \widehat{\theta}) \\
&= E\big[G_T(\theta - \widehat{\theta}) + G_T(\widehat{\theta}) \\
&\qquad - \xi\big(G_T(\theta - \widehat{\theta})\big)^2 - 2\xi\big(G_T(\theta - \widehat{\theta})\big)G_T(\widehat{\theta}) - \xi\big(G_T(\widehat{\theta})\big)^2 \\
&\qquad + \xi\big(E[G_T(\theta - \widehat{\theta})]\big)^2 + 2\xi E[G_T(\theta - \widehat{\theta})]E[G_T(\widehat{\theta})] + \xi\big(E[G_T(\widehat{\theta})]\big)^2\big] \\
&= -\xi E\big[\big(G_T(\theta - \widehat{\theta})\big)^2 - \big(E[G_T(\theta - \widehat{\theta})]\big)^2\big] + j^{\mathrm{mv}}(\widehat{\theta}) \\
&= -\xi \mathrm{Var}[G_T(\theta - \widehat{\theta})] + j^{\mathrm{mv}}(\widehat{\theta}) \\
&\leq j^{\mathrm{mv}}(\widehat{\theta}).
\end{aligned}$$

This shows that $\widehat{\theta}$ is optimal for the MVPS problem (4.95).

Now we show the equivalence between (4.98) and (4.99). For "$\Rightarrow$", it is enough to take $\eta = \mathbb{1}_{H\times\{t\}}$ with $H \in \mathcal{F}_{t-1}$ for $t = 1, \dots, T$ and use that $G_T(\widehat{\theta})$ and $\triangle S_t$, hence also $G_T(\eta) = \mathbb{1}_H \triangle S_t$, are in L^2 due to $\widehat{\theta} \in \Theta_{\mathrm{MN}}$ and Assumption 4.2, respectively. To prove "$\Leftarrow$", let $\eta \in \Theta_{\mathrm{MN}}$ and consider

$$H_t^n := \{|\eta_s| \leq n \text{ for } s = 1, \dots, t\} \in \mathcal{F}_{t-1}$$

for $n \in \mathbb{N}$ and $t = 1, \dots, T$. Note that $H_t^n \in \mathcal{F}_{t-1}$ uses that η is $\mathbb{F}$-predictable. To ease notation, we set $\widehat{F} := 1 - 2\xi G_T(\widehat{\theta}) + 2\xi E[G_T(\widehat{\theta})]$. Note that $G_T(\widehat{\theta}) \in L^2$ implies

$$\widehat{F} \in L^2. \tag{4.100}$$

By the definition of H_t^n, we get

$$Y_t^n := E[\widehat{F}|\mathcal{F}_t]G_t(\eta)\mathbb{1}_{H_t^n} \longrightarrow E[\widehat{F}|\mathcal{F}_t]G_t(\eta) \quad P\text{-a.s.} \tag{4.101}$$

In view of (4.98), it is enough to show $E[\widehat{F}G_T(\eta)] = 0$. To this end, we show for $t = 0, 1, \dots, T$ that

$$E[\widehat{F}|\mathcal{F}_t]G_t(\eta) = E[\widehat{F}G_T(\eta)|\mathcal{F}_t]. \tag{4.102}$$

For $t = T$, (4.102) is trivial. Suppose (4.102) is true for t. For the induction step, it is enough to show $E[E[\widehat{F}|\mathcal{F}_t]G_t(\eta)|\mathcal{F}_{t-1}] = E[\widehat{F}|\mathcal{F}_{t-1}]G_{t-1}(\eta)$. Now for $n \in \mathbb{N}$ and $s = 1, \dots, t$, $\eta_s \mathbb{1}_{H_t^n}$ is bounded. Because S is square-integrable, this implies that $G_t(\theta)\mathbb{1}_{H_t^n} \in L^2$, and $\widehat{F} \in L^2$ by (4.100). Hence we get for $n \in \mathbb{N}$ that $\widehat{F}G_t(\eta)\mathbb{1}_{H_t^n}$ and $\widehat{F}\triangle S_t$ are in L^1. This allows us to condition on $\mathcal{F}_{t-1}$ in the definition (4.101) of Y_t^n, take out the $\mathcal{F}_{t-1}$-measurable quantities $G_{t-1}(\eta)\mathbb{1}_{H_t^n}$ and

$\eta_t \mathbb{1}_{H_t^n}$ and use the FOC (4.99), which reads $E[\widehat{F}\triangle S_t|\mathcal{F}_{t-1}] = 0$, to obtain

$$\begin{aligned} E[Y_t^n|\mathcal{F}_{t-1}] &= E\big[E[\widehat{F}|\mathcal{F}_t]\big(G_{t-1}(\eta)\mathbb{1}_{H_t^n} + \eta_t\mathbb{1}_{H_t^n}\triangle S_t\big)\big|\mathcal{F}_{t-1}\big] \\ &= E[\widehat{F}|\mathcal{F}_{t-1}]G_{t-1}(\eta)\mathbb{1}_{H_t^n} + E[\widehat{F}\triangle S_t|\mathcal{F}_{t-1}]\mathbb{1}_{H_t^n} \\ &= E[\widehat{F}|\mathcal{F}_{t-1}]G_{t-1}(\eta)\mathbb{1}_{H_t^n} \\ &\longrightarrow E[\widehat{F}|\mathcal{F}_{t-1}]G_{t-1}(\eta) \quad P\text{-a.s. as } n \to \infty, \end{aligned} \tag{4.103}$$

for $t = 1, \dots, T$. Note that the induction hypothesis (4.102) also implies that $E[\widehat{F}|\mathcal{F}_t]G_t(\eta) = E[\widehat{F}G_T(\eta)|\mathcal{F}_t] \in L^1$. The definition (4.101) for Y_t^n obviously yields $|Y_t^n| \le |E[\widehat{F}|\mathcal{F}_t]G_t(\eta)|$ and so the dominated convergence theorem implies $Y_t^n \to E[\widehat{F}|\mathcal{F}_t]G_t(\eta)$ in L^1 and that $E[Y_t^n|\mathcal{F}_{t-1}] \to E[\widehat{F}|\mathcal{F}_{t-1}]G_{t-1}(\eta)$ in L^1 due to (4.103). We therefore obtain

$$E\big[E[\widehat{F}|\mathcal{F}_t]G_t(\eta)\big|\mathcal{F}_{t-1}\big] = \lim_{n\to\infty} E[Y_t^n|\mathcal{F}_{t-1}] = E[\widehat{F}|\mathcal{F}_{t-1}]G_{t-1}(\eta)$$

as desired. □

To implement Recipe 4.21, Step 2), we now add the extra assumption mentioned in Remark 4.20, 2) that $1 \notin \overline{G_T(\Theta_{\mathrm{MN}})} = G_T(\Theta_{\mathrm{MN}})$. For convenience, we modify Assumption 4.5, 1) by replacing Θ_{S} with Θ_{MN} and keeping Assumption 4.5, 2) unchanged and refer it to Assumption 4.23 below. Although Assumption 4.23, 1) is stronger than Assumption 4.5, 1) due to $G_T(\Theta_{\mathrm{S}}) \subseteq G_T(\Theta_{\mathrm{MN}})$ from (4.97), we believe this is more natural because we work in this subsection with Θ_{MN} only.

Assumption 4.23. 1) The space Θ_{MN} satisfies $1 \notin G_T(\Theta_{\mathrm{MN}})$.

2) The process S satisfies the structure condition (SC), meaning that the process $(E[\triangle S_t|\mathcal{F}_{t-1}])_{t=1,\dots,T}$ is absolutely continuous with respect to the process $(\mathrm{Var}[\triangle S_t|\mathcal{F}_{t-1}])_{t=1,\dots,T}$.

Recall the strategy $\widetilde{\theta}^{\mathrm{mv}}$ and some related quantities from Theorem 4.11. The strategy $\widetilde{\theta}^{\mathrm{mv}}$ is given by

$$\begin{aligned} \widetilde{\theta}_u^{\mathrm{mv}} &= -\frac{E[\widetilde{Z}_u\triangle S_u|\mathcal{F}_{u-1}]}{E[\widetilde{Z}_u(\triangle S_u)^2|\mathcal{F}_{u-1}]} \\ &\qquad \times \left(G_{u-1}(\widetilde{\theta}^{\mathrm{mv}}) - \frac{(2\xi)^{-1} + E[\widetilde{Z}_{u-1}G_{u-1}(\widetilde{\theta}^{\mathrm{mv}})]}{E[\widetilde{Z}_{u-1}]}\right) \end{aligned} \tag{4.104}$$

$$= -\beta_u\left(G_{u-1}(\widetilde{\theta}^{\mathrm{mv}}) - \frac{1}{2\xi E[\widetilde{Z}_0]}\right), \quad u = 1, \dots, T. \tag{4.105}$$

Note that $\widetilde{\theta}^{\mathrm{mv}}$ is well defined whenever β is well defined and $E[\widetilde{Z}_t] > 0$ for $t = 0, 1, \ldots, T$. Now if we look at Lemma 4.3 and the definition (4.25) of β, then we see that β is well defined under Assumption 4.2. It remains to check that $\widetilde{Z}$ has strictly positive expectations. In view of the proof of Lemma 4.6, 1), this holds under the assumption that the L^2-closure of $G_T(\Theta_{\mathrm{S}})$ does not contain the constant payoff 1. But we have seen in the discussion above Assumption 4.23 that this follows from Assumption 4.23, 1). To sum up, we can safely proceed by replacing Assumption 4.5 with Assumption 4.23. The quantities $\widetilde{Z}, \widetilde{a}, \widetilde{b}, \widetilde{c}$ are given by

$$\widetilde{Z}_{t-1} = E\left[\widetilde{Z}_t\left(1 - \frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)|\mathcal{F}_{t-1}]}\triangle S_t\right)\middle|\mathcal{F}_{t-1}\right], \quad \tilde{Z}_T = 1,$$

$$\widetilde{a}_t = \frac{1}{E[\widetilde{Z}_t]}, \quad t = 0, 1, \ldots, T, \tag{4.106}$$

$$\widetilde{b}_t = -\xi, \quad t = 0, 1, \ldots, T, \tag{4.107}$$

$$\widetilde{c}_t = \frac{\xi}{E[\widetilde{Z}_t]}, \quad t = 0, 1, \ldots, T. \tag{4.108}$$

With the process β defined as in (4.25) by

$$\beta_t = \frac{E[\widetilde{Z}_t \triangle S_t | \mathcal{F}_{t-1}]}{E[\widetilde{Z}_t(\triangle S_t)|\mathcal{F}_{t-1}]}, \quad t = 1, \ldots, T,$$

we can also write

$$\widetilde{Z}_{t-1} = E[\widetilde{Z}_t(1 - \beta_t \triangle S_t)|\mathcal{F}_{t-1}] = E\left[\prod_{u=t}^{T}(1 - \beta_u \triangle S_u)\middle|\mathcal{F}_{t-1}\right] \tag{4.109}$$

for $t = 1, \ldots, T$. After this preparation, we start to implement the second step in Recipe 4.21.

Lemma 4.24. *Suppose that Assumptions* 4.2 *and* 4.23 *are satisfied. Then we have for* $t = 1, \ldots, T$ *that*

$$G_t(\widetilde{\theta}^{\mathrm{mv}}) - \frac{1}{2\xi E[\widetilde{Z}_0]} = -\frac{1}{2\xi E[\widetilde{Z}_0]}\prod_{s=1}^{t}(1 - \beta_s \triangle S_s), \tag{4.110}$$

$$\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}}) \in L^2, \tag{4.111}$$

$$\begin{aligned} &E\big[\big(1 - 2\xi G_T(\widetilde{\theta}^{\mathrm{mv}}) + 2\xi E[G_T(\widetilde{\theta}^{\mathrm{mv}})]\big)\big|\mathcal{F}_t\big] \\ &= \widetilde{a}_t \widetilde{Z}_t + 2\widetilde{b}_t \widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}}) + 2\widetilde{c}_t \widetilde{Z}_t E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]. \end{aligned} \tag{4.112}$$

Proof. 1) We first prove (4.110). Using $G_t(\widetilde{\theta}^{\rm mv}) = G_{t-1}(\widetilde{\theta}^{\rm mv}) + \widetilde{\theta}^{\rm mv}_t \triangle S_t$ and the explicit formula (4.105) for $\widetilde{\theta}^{\rm mv}$, we obtain

$$\begin{aligned} G_t(\widetilde{\theta}^{\rm mv}) - \frac{1}{2\xi E[\widetilde{Z}_0]} &= G_{t-1}(\widetilde{\theta}^{\rm mv}) - \beta_t \triangle S_t \bigg(G_{t-1}(\widetilde{\theta}^{\rm mv}) - \frac{1}{2\xi E[\widetilde{Z}_0]}\bigg) - \frac{1}{2\xi E[\widetilde{Z}_0]} \\ &= (1-\beta_t\triangle S_t)\bigg(G_{t-1}(\widetilde{\theta}^{\rm mv}) - \frac{1}{2\xi E[\widetilde{Z}_0]}\bigg) \end{aligned}$$

for $t = 1, \dots, T$. Iterating the above identity in t and using $G_0(\widetilde{\theta}^{\rm mv}) = 0$ yields (4.110).

2) To argue $\widetilde{Z}_t G_t(\widetilde{\theta}^{\rm mv}) \in L^2$, we use the second formula in (4.109) for $\widetilde{Z}_t$, then $(x-y)^2 \le 2x^2 + 2y^2$ and (4.110) to obtain

$$\begin{aligned} &\big(\widetilde{Z}_t G_t(\widetilde{\theta}^{\rm mv})\big)^2 \\ &\le \bigg(\frac{1}{2\xi E[\widetilde{Z}_0]}\bigg)^2 E\bigg[\bigg(\prod_{u=t+1}^{T}(1-\beta_u\triangle S_u)^2\bigg)\bigg(1-\prod_{s=1}^{t}(1-\beta_s\triangle S_s)\bigg)^2\bigg|\mathcal{F}_t\bigg] \\ &\le \bigg(\frac{1}{2\xi E[\widetilde{Z}_0]}\bigg)^2 E\bigg[\bigg(\prod_{u=t+1}^{T}(1-\beta_u\triangle S_u)^2\bigg)2\bigg(1+\prod_{s=1}^{t}(1-\beta_s\triangle S_s)^2\bigg)\bigg|\mathcal{F}_t\bigg] \\ &= \bigg(\frac{\sqrt{2}}{2\xi E[\widetilde{Z}_0]}\bigg)^2 E\bigg[\prod_{u=t+1}^{T}(1-\beta_u\triangle S_u)^2 + \prod_{u=1}^{T}(1-\beta_u\triangle S_u)^2\bigg|\mathcal{F}_t\bigg]. \end{aligned}$$

Taking expectations, using the explicit expression (4.109) for $\widetilde{Z}$, Jensen's inequality and the bound $|\widetilde{Z}_t| \le 1$ for $t = 1, \dots, T$ from (4.16), we get

$$E\big[\big(\widetilde{Z}_t G_t(\widetilde{\theta}^{\rm mv})\big)^2\big] \le \bigg(\frac{\sqrt{2}}{2\xi E[\widetilde{Z}_0]}\bigg)^2 (E[\widetilde{Z}_t^2] + E[\widetilde{Z}_0^2]) \le \bigg(\frac{1}{\xi E[\widetilde{Z}_0]}\bigg)^2,$$

as desired. Note that directly using $|\widetilde{Z}_t| \le 1$ to estimate

$$\big(\widetilde{Z}_t G_t(\widetilde{\theta}^{\rm mv})\big)^2 \le \big(G_t(\widetilde{\theta}^{\rm mv})\big)^2$$

does not help because without extra assumptions, we do not know whether $G_t(\widetilde{\theta}^{\rm mv})$ is in L^2.

3) We establish (4.112) by induction. First by (4.111) for $t = T$ and due to $\widetilde{Z}_T = 1$, the quantity inside the conditional expectation is in L^2 and so that conditional expectation is well defined. Suppose the identity is true for $t+1$. To

prove the induction step, we show that

$$
\begin{aligned}
&E\big[\widetilde{a}_{t+1}\widetilde{Z}_{t+1} + 2\widetilde{b}_{t+1}\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}}) + 2\widetilde{c}_{t+1}\widetilde{Z}_{t+1}E[\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}})]\big|\mathcal{F}_t\big] \\
&= \widetilde{a}_t\widetilde{Z}_t + 2\widetilde{b}_t\widetilde{Z}_tG_t(\widetilde{\theta}^{\mathrm{mv}}) + 2\widetilde{c}_t\widetilde{Z}_tE[\widetilde{Z}_tG_t(\widetilde{\theta}^{\mathrm{mv}})].
\end{aligned}
$$

Using the explicit expressions (4.109) and (4.110) for $\widetilde{Z}$ and $G(\widetilde{\theta}^{\mathrm{mv}})$, respectively, and $\prod_{u=t}^{T}(1-\beta_u\triangle S_u) = \widetilde{U}_{t,T} \in L^2$ for $t = 1, \dots, T$ from (4.25), (4.14) and (4.12), we obtain

$$
\begin{aligned}
\widetilde{Z}_uG_u(\widetilde{\theta}^{\mathrm{mv}}) &= \frac{1}{2\xi E[\widetilde{Z}_0]}\left(\widetilde{Z}_u - \widetilde{Z}_u\prod_{s=1}^{u}(1-\beta_s\triangle S_s)\right) \\
&= \frac{1}{2\xi E[\widetilde{Z}_0]}\left(\widetilde{Z}_u - E\left[\prod_{s=u+1}^{T}(1-\beta_s\triangle S_s)\middle|\mathcal{F}_u\right]\prod_{s=1}^{u}(1-\beta_s\triangle S_s)\right) \\
&= \frac{1}{2\xi E[\widetilde{Z}_0]}E\left[\widetilde{Z}_u - \prod_{s=1}^{T}(1-\beta_s\triangle S_s)\middle|\mathcal{F}_u\right] \qquad (4.113)
\end{aligned}
$$

for $u = 0, \dots, T$. Subtracting the right-hand side of (4.113) for $u = t$ from $u = t+1$ and conditioning on $\mathcal{F}_u$ yields

$$
E[\widetilde{Z}_{u+1}G_{u+1}(\widetilde{\theta}^{\mathrm{mv}}) - \widetilde{Z}_uG_u(\widetilde{\theta}^{\mathrm{mv}})|\mathcal{F}_u] = \frac{1}{2\xi E[\widetilde{Z}_0]}E[\widetilde{Z}_{u+1} - \widetilde{Z}_u|\mathcal{F}_u] \qquad (4.114)
$$

for $u = 0, \dots, T-1$. Taking expectations in (4.113) and using the explicit expression (4.109) for $\widetilde{Z}$ leads to

$$
E[\widetilde{Z}_uG_u(\widetilde{\theta}^{\mathrm{mv}})] = \frac{1}{2\xi E[\widetilde{Z}_0]}(E[\widetilde{Z}_u] - E[\widetilde{Z}_0]) \qquad (4.115)
$$

for $u = 0, \dots, T$. Then we use the recursive definition (4.109) for $\widetilde{Z}_t$ to obtain

$$
\begin{aligned}
E[\triangle(\widetilde{a}_{t+1}\widetilde{Z}_{t+1})|\mathcal{F}_t] &= E[\widetilde{a}_{t+1}\widetilde{Z}_{t+1} - \widetilde{a}_t\widetilde{Z}_{t+1}(1-\beta_{t+1}\triangle S_{t+1})|\mathcal{F}_t] \\
&= E[(\widetilde{a}_{t+1} - \widetilde{a}_t)\widetilde{Z}_{t+1} + \widetilde{a}_t\widetilde{Z}_{t+1}\beta_{t+1}\triangle S_{t+1}|\mathcal{F}_t]. \qquad (4.116)
\end{aligned}
$$

Using (4.114) and $\widetilde{b} \equiv -\xi$ from (4.107) as well as (4.109), we get

$$\begin{aligned} E\big[\triangle\big(2\widetilde{b}_{t+1}\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}})\big)\big|\mathcal{F}_t\big] &= -2\xi E[\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}}) - \widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})|\mathcal{F}_t] \\ &= -\frac{1}{E[\widetilde{Z}_0]}E[\widetilde{Z}_{t+1} - \widetilde{Z}_t|\mathcal{F}_t] \\ &= -\frac{1}{E[\widetilde{Z}_0]}E[\widetilde{Z}_{t+1}\beta_{t+1}\triangle S_{t+1}|\mathcal{F}_t]. \end{aligned} \tag{4.117}$$

Now we use the recursive definition (4.109) for $\widetilde{Z}_t$, then the expression (4.115) for $E[\widetilde{Z}_u G_u(\widetilde{\theta}^{\mathrm{mv}})]$ with $u \in \{t+1, t\}$ and $\widetilde{c}_t = \xi\widetilde{a}_t$ from (4.106) and (4.108) and finally $\widetilde{a}_t = \frac{1}{E[\widetilde{Z}_t]}$ from (4.106) to obtain

$$\begin{aligned} &E\big[\triangle(2\widetilde{c}_{t+1}\widetilde{Z}_{t+1}E[\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}})])\big|\mathcal{F}_t\big] \\ &= E\big[2\widetilde{c}_{t+1}\widetilde{Z}_{t+1}E[\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}})] - 2\widetilde{c}_t\widetilde{Z}_{t+1}(1-\beta_t\triangle S_t)E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]\big|\mathcal{F}_t\big] \\ &= E\big[\big(2\widetilde{c}_{t+1}E[\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}})] - 2\widetilde{c}_t E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]\big)\widetilde{Z}_{t+1} \\ &\qquad + 2\widetilde{c}_t\widetilde{Z}_{t+1}\beta_t\triangle S_t E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]\big|\mathcal{F}_t\big] \\ &= 2\xi E\Bigg[\bigg(\widetilde{a}_{t+1}\frac{E[\widetilde{Z}_{t+1}] - E[\widetilde{Z}_0]}{2\xi E[\widetilde{Z}_0]} - \widetilde{a}_t\frac{E[\widetilde{Z}_t] - E[\widetilde{Z}_0]}{2\xi E[\widetilde{Z}_0]}\bigg)\widetilde{Z}_{t+1} \\ &\qquad + \widetilde{a}_t\widetilde{Z}_{t+1}\beta_t\triangle S_t\frac{E[\widetilde{Z}_t] - E[\widetilde{Z}_0]}{2\xi E[\widetilde{Z}_0]}\bigg|\mathcal{F}_t\Bigg] \\ &= \frac{1}{E[\widetilde{Z}_0]}E\big[\big(\widetilde{a}_{t+1}(E[\widetilde{Z}_{t+1}] - E[\widetilde{Z}_0]) - \widetilde{a}_t(E[\widetilde{Z}_t] - E[\widetilde{Z}_0])\big)\widetilde{Z}_{t+1} \\ &\qquad + \widetilde{a}_t\widetilde{Z}_{t+1}\beta_t\triangle S_t(E[\widetilde{Z}_t] - E[\widetilde{Z}_0])\big|\mathcal{F}_t\big] \\ &= E\Bigg[(\widetilde{a}_t - \widetilde{a}_{t+1})\widetilde{Z}_{t+1} + \bigg(\frac{1}{E[\widetilde{Z}_0]} - \widetilde{a}_t\bigg)\widetilde{Z}_{t+1}\beta_{t+1}\triangle S_{t+1}\Bigg|\mathcal{F}_t\Bigg]. \end{aligned} \tag{4.118}$$

Adding (4.116)–(4.118), we obtain

$$E\big[\triangle\big(\widetilde{a}_{t+1}\widetilde{Z}_{t+1} + 2\widetilde{b}_{t+1}\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}}) + 2c^2_{t+1}\widetilde{Z}_{t+1}E[\widetilde{Z}_{t+1}G_{t+1}(\widetilde{\theta}^{\mathrm{mv}})]\big)\big|\mathcal{F}_t\big] = 0$$

as desired. This completes the proof. □

We now state the main result of this subsection.

Theorem 4.25. *Suppose that Assumptions* 4.2 *and* 4.23 *are satisfied, meaning that the price process S is square-integrable and satisfies the structure condition (SC) and the space $G_T(\Theta_{\mathrm{MN}})$ does not contain the constant payoff 1. Then the* *final payoff* $G_T(\widetilde{\theta}^{\mathrm{mv}})$ *satisfies* (4.99), *namely*

$$E\big[\big(1-2\xi G_T(\widetilde{\theta}^{\mathrm{mv}})+2\xi E[G_T(\widetilde{\theta}^{\mathrm{mv}})]\big)\triangle S_t\big|\mathcal{F}_{t-1}\big]=0,\quad t=1,\ldots,T. \qquad (4.119)$$

Consequently, $\widetilde{\theta}^{\mathrm{mv}}$ *given in* (4.104)–(4.108) *is an optimal strategy for the problem* (4.95), *i.e. the MVPS problem with* $\Theta=\Theta_{\mathrm{MN}}$.

Proof. The second assertion follows from (4.119) and Lemma 4.22. So we only need to show (4.119). To this end, we first use (4.112) and that $G_T(\widetilde{\theta}^{\mathrm{mv}})\triangle S_t$ and $\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})\triangle S_t$ are in L^1 due to (4.111) and Assumption 4.2 to obtain

$$\begin{aligned}
&E\big[\big(1-2\xi G_T(\widetilde{\theta}^{\mathrm{mv}})+2\xi E[G_T(\widetilde{\theta}^{\mathrm{mv}})]\big)\triangle S_t\big|\mathcal{F}_{t-1}\big]\\
&=E\Big[E\big[\big(1-2\xi G_T(\widetilde{\theta}^{\mathrm{mv}})+2\xi E[G_T(\widetilde{\theta}^{\mathrm{mv}})]\big)\big|\mathcal{F}_t\big]\triangle S_t\Big|\mathcal{F}_{t-1}\Big]\\
&=E\big[\big(\widetilde{a}_t\widetilde{Z}_t+2\widetilde{b}_t\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})+2\widetilde{c}_t\widetilde{Z}_t E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]\big)\triangle S_t\big|\mathcal{F}_{t-1}\big]\\
&=E\bigg[\bigg(\frac{\widetilde{Z}_t}{E[\widetilde{Z}_t]}-2\xi\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})+2\xi\frac{1}{E[\widetilde{Z}_t]}\widetilde{Z}_t E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]\bigg)\triangle S_t\bigg|\mathcal{F}_{t-1}\bigg]. \qquad (4.120)
\end{aligned}$$

This last equality uses the explicit expressions (4.106)–(4.108) for $\widetilde{a}_t,\widetilde{b}_c,\widetilde{c}_t$. Inserting the expression (4.115) for $E[\widetilde{Z}_t\widetilde{G}_t(\widetilde{\theta}^{\mathrm{mv}})]$ into (4.120), then using (4.113) for $\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})$ and (4.109) for $\widetilde{Z}_t$ and finally cancelling out terms yields

$$\begin{aligned}
&E\bigg[\bigg(\frac{\widetilde{Z}_t}{E[\widetilde{Z}_t]}-2\xi\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})+2\xi\frac{1}{E[\widetilde{Z}_t]}\widetilde{Z}_t E[\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})]\bigg)\triangle S_t\bigg|\mathcal{F}_{t-1}\bigg]\\
&=E\bigg[\bigg(\frac{\widetilde{Z}_t}{E[\widetilde{Z}_t]}-2\xi\widetilde{Z}_t G_t(\widetilde{\theta}^{\mathrm{mv}})+\frac{1}{E[\widetilde{Z}_t]E[\widetilde{Z}_0]}\widetilde{Z}_t(E[\widetilde{Z}_t]-E[\widetilde{Z}_0])\bigg)\triangle S_t\bigg|\mathcal{F}_{t-1}\bigg]\\
&=E\bigg[\bigg(\frac{\widetilde{Z}_t}{E[\widetilde{Z}_t]}-\frac{1}{E[\widetilde{Z}_0]}\Big(\widetilde{Z}_t-\prod_{s=1}^{T}(1-\beta_s\triangle S_s)\Big)\\
&\qquad+\frac{1}{E[\widetilde{Z}_t]E[\widetilde{Z}_0]}\widetilde{Z}_t(E[\widetilde{Z}_t]-E[\widetilde{Z}_0])\bigg)\triangle S_t\bigg|\mathcal{F}_{t-1}\bigg]\\
&=E\bigg[\frac{\triangle S_t}{E[\widetilde{Z}_0]}\prod_{u=1}^{T}(1-\beta_u\triangle S_u)\bigg|\mathcal{F}_{t-1}\bigg]\\
&=0.
\end{aligned}$$

The last step uses Schweizer [62, Corollary 4]. ∎

5 Connections to the literature

In this final section, we discuss related work in the literature by giving an overview of papers that attack the MVPS problem via McKean–Vlasov control theory. A first successful attempt is made by Andersson and Djehiche [4] who obtain a stochastic maximum principle for McKean–Vlasov control problems for diffusion models and use that to solve the MVPS problem for a (time-dependent) Black–Scholes model. In finite discrete time, Elliott et al. [28] present a similar result for a general linear–quadratic problem whose controlled system is linear and driven by a martingale difference sequence. The present chapter is mainly inspired by Pham and Wei [52] where the authors also adopt a dynamic programming (DP) approach for solving McKean–Vlasov control problems. However, they work at the level of probability distributions and assume that the controlled process X^θ is driven by i.i.d. innovations. Cui et al. [22] and its extension by Barbieri and Costa [7] propose a similar DP approach for a system X^θ driven by i.i.d. innovations, and they work more explicitly with the state variables $E[X^\theta]$ and $X^\theta - E[X^\theta]$. A similar result to our deterministic version of the martingale optimality principle in Corollary 2.5 can be found in Basei and Pham [8, Lemma 3.1] for SDE models in continuous time.

The process $\widetilde{Z}$ in the expressions (4.72) and (4.73) for the optimal strategy $\widetilde{\theta}^{\mathrm{mv}}$ turns out to be the opportunity process L in the sense of Černý and Kallsen [17] or the process called q in Jeanblanc et al. [38]. This is not surprising because the MVPS problem is closely connected to the pure hedging problem to

$$\text{minimise } E\Big[\big(1 - G_T(\theta)\big)^2\Big] \quad \text{over } \theta \in \Theta$$

as explained in Fontana and Schweizer [30]. For a general theory of the pure hedging problem, we refer to Schweizer [61, 62] or more comprehensively to Černý and Kallsen [17].

Index

Bibliography

[1] E. Abi Jaber, E. Miller, and H. Pham. Markowitz portfolio selection for multivariate affine and quadratic Volterra models. *SIAM Journal on Financial Mathematics*, 12:369–409, 2021.

[2] B. Acciaio, J. Backhoff-Veraguas, and R. Carmona. Extended mean field control problems: Stochastic maximum principle and transport perspective. *SIAM Journal on Control and Optimization*, 57:3666–3693, 2019.

[3] A. Aksamit and M. Jeanblanc. *Enlargement of Filtration with Finance in View*. Springer, Cham, 2017.

[4] D. Andersson and B. Djehiche. A maximum principle for SDEs of mean-field type. *Applied Mathematics and Optimization*, 63:341–356, 2011.

[5] S. Ankirchner and A. Dermoune. Multiperiod mean-variance portfolio optimization via market cloning. *Applied Mathematics and Optimization*, 64:135–154, 2011.

[6] T. Arai. An extension of mean-variance hedging to the discontinuous case. *Finance and Stochastics*, 9:129–139, 2005.

[7] F. Barbieri and O. Costa. A mean-field formulation for the mean-variance control of discrete-time linear systems with multiplicative noises. *International Journal of Systems Science*, 51:1825–1846, 2020.

[8] M. Basei and H. Pham. A weak martingale approach to linear-quadratic McKean–Vlasov stochastic control problems. *Journal of Optimization Theory and Applications*, 181:347–382, 2019.

[9] D. Bertsimas, L. Kogan, and A. W. Lo. Hedging derivative securities and incomplete markets: An ε-arbitrage approach. *Operations Research*, 49:372–397, 2001.

[10] F. Biagini, P. Guasoni, and M. Pratelli. Mean-variance hedging for stochastic volatility models. *Mathematical Finance*, 10:109–123, 2000.

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

[12] T. Björk, M. Khapko, and A. Murgoci. *Time-Inconsistent Control Theory with Finance Applications*. Springer, Cham, 2021.

[13] T. Björk and A. Murgoci. A theory of Markovian time-inconsistent stochastic control in discrete time. *Finance and Stochastics*, 18:545–592, 2014.

[14] T. Bühler and D. A. Salamon. *Functional Analysis*. American Mathematical Society, Providence, RI, 2018.

[15] A. Černý. Dynamic programming and mean-variance hedging in discrete time. *Applied Mathematical Finance*, 11:1–25, 2004.

[16] A. Černý. Optimal continuous-time hedging with leptokurtic returns. *Mathematical Finance*, 17:175–203, 2007.

[17] A. Černý and J. Kallsen. On the structure of general mean-variance hedging strategies. *Annals of Probability*, 35:1479–1531, 2007.

[18] A. Černý and J. Kallsen. Mean–variance hedging and optimal investment in Heston's model with correlation. *Mathematical Finance*, 18:473–492, 2008.

[19] A. Černý and J. Kallsen. Hedging by sequential regressions revisited. *Mathematical Finance*, 19:591–617, 2009.

[20] M. C. Chiu and H. Y. Wong. Mean–variance portfolio selection with correlation risk. *Journal of Computational and Applied Mathematics*, 263:432–444, 2014.

[21] O. Costa and A. de Oliveira. Optimal mean–variance control for discrete-time linear systems with Markovian jumps and multiplicative noises. *Automatica*, 48:304–315, 2012.

[22] X. Cui, X. Li, and D. Li. Unified framework of mean-field formulations for optimal multi-period mean-variance portfolio selection. *IEEE Transactions on Automatic Control*, 59:1833–1844, 2014.

[23] F. Delbaen and W. Schachermayer. The variance-optimal martingale measure for continuous processes. *Bernoulli*, 2:81–105, 1996.

[24] F. Delbaen and W. Schachermayer. *The Mathematics of Arbitrage*. Springer Berlin, Heidelberg, 2006.

[25] C. Dellacherie and P.-A. Meyer. *Probabilities and Potential B. Theory of Martingales*. North-Holland, Amsterdam, 1982.

[26] D. Duffie and H. Richardson. Mean-variance hedging in continuous time. *Annals of Applied Probability*, 1:1 – 15, 1991.

[27] N. El Karoui. Les aspects probabilistes du contrôle stochastique. In P. L. Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour IX-1979*, Lecture Notes in Mathematics 876, pages 73–238. Springer Berlin Heidelberg, 1981.

[28] R. Elliott, X. Li, and Y.-H. Ni. Discrete time mean-field stochastic linear-quadratic optimal control problems. *Automatica*, 49:3222–3233, 2013.

[29] M. Fischer and G. Livieri. Continuous time mean-variance portfolio optimization through the mean field approach. *ESAIM: Probability and Statistics*, 20:30–44, 2016.

[30] C. Fontana and M. Schweizer. Simplified mean-variance portfolio optimisation. *Mathematics and Financial Economics*, 6:125–152, 2012.

[31] C. Gouriéroux, J.-P. Laurent, and H. Pham. Mean-variance hedging and numéraire. *Mathematical Finance*, 8:179–200, 1998.

[32] S. Gugushvili. Dynamic programming and mean-variance hedging in discrete time. *Georgian Mathematical Journal*, 10:237–246, 2003.

[33] B. Han and H. Y. Wong. Mean–variance portfolio selection under Volterra Heston model. *Applied Mathematics and Optimization*, 84:683–710, 2021.

[34] J. He, Q. Wang, P. Cheng, J. Chen, and Y. Sun. Multi-period mean-variance portfolio optimization with high-order coupled asset dynamics. *IEEE Transactions on Automatic Control*, 60:1320–1335, 2015.

[35] D. Hobson. Stochastic volatility models, correlation, and the q-optimal measure. *Mathematical Finance*, 14:537–556, 2004.

[36] F. Hubalek, J. Kallsen, and L. Krawczyk. Variance-optimal hedging for processes with stationary independent increments. *Annals of Applied Probability*, 16:853–885, 2006.

[37] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer Berlin, Heidelberg, 2003. Second Edition.

[38] M. Jeanblanc, M. Mania, M. Santacroce, and M. Schweizer. Mean-variance hedging via stochastic control and BSDEs for general semimartingales. *Annals of Applied Probability*, 22:2388–2428, 2012.

[39] Y. Kabanov and M. Safarian. *Markets with Transaction Costs: Mathematical Theory*. Springer Berlin, Heidelberg, 2009.

[40] O. Kallenberg. *Foundations of Modern Probability*. Springer, Cham, 2021. Third Edition.

[41] J. Kallsen, J. Muhle-Karbe, N. Shenkman, and R. Vierthauer. Discrete-time variance-optimal hedging in affine stochastic volatility models. In R. Kiesel, Scherer M., and Zagst R., editors, *Alternative Investments and Strategies*. World Scientific, 2009.

[42] M. Kohlmann and S. Tang. Global adapted solution of one-dimensional backward stochastic Riccati equations, with application to the mean–variance hedging. *Stochastic Processes and their Applications*, 97:255–288, 2002.

[43] J.-P. Laurent and H. Pham. Dynamic programming and mean-variance hedging. *Finance and Stochastics*, 3:83–110, 1999.

[44] D. Lépingle and J. Mémin. Sur l'intégrabilité uniforme des martingales exponentielles. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 42:175–203, 1978.

[45] D. Li and W.-L. Ng. Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical Finance*, 10:387–406, 2000.

[46] A. E. B. Lim. Quadratic hedging and mean-variance portfolio selection with random parameters in an incomplete market. *Mathematics of Operations Research*, 29:132–161, 2004.

[47] A. E. B. Lim. Mean-variance hedging when there are jumps. *SIAM Journal on Control and Optimization*, 44:1893–1922, 2005.

[48] A. E. B. Lim and X. Y. Zhou. Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research*, 27:101–120, 2002.

[49] H. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

[50] A. V. Melnikov and M. L. Nechaev. On the mean-variance hedging problem. *Theory of Probability and Its Applications*, 43:588–603, 1999.

[51] P. Monat and C. Stricker. Föllmer-Schweizer decomposition and mean-variance hedging for general claims. *Annals of Probability*, 23:605–628, 1995.

[52] H. Pham and X. Wei. Discrete time McKean-Vlasov control problem: a dynamic programming approach. *Applied Mathematics and Optimization*, 74:487–506, 2016.

[53] H. Pham and X. Wei. Dynamic programming for optimal control of stochastic McKean–Vlasov dynamics. *SIAM Journal on Control and Optimization*, 55:1069–1101, 2017.

[54] H. Pham and X. Wei. Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM: Control, Optimisation and Calculus of Variations*, 24:437–461, 2018.

[55] P. E. Protter. *Stochastic Integration and Differential Equations.* Springer Berlin, Heidelberg, 2005. Second edition. Version 2.1, Corrected third printing.

[56] T. Rheinländer and M. Schweizer. On L^2-projections on a space of stochastic integrals. *Annals of Probability*, 25:1810–1831, 1997.

[57] D. A. Salamon. *Measure and Integration.* European Mathematical Society (EMS), Zürich, 2016.

[58] M. Schweizer. Mean-variance hedging for general claims. *Annals of Applied Probability*, 2:171–179, 1992.

[59] M. Schweizer. Approximating random variables by stochastic integrals. *Annals of Probability*, 22:1536–1575, 1994.

[60] M. Schweizer. On the minimal martingale measure and the Föllmer-Schweizer decomposition. *Stochastic Analysis and Applications*, 13:573–599, 1995.

[61] M. Schweizer. Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, 20:1–32, 1995.

[62] M. Schweizer. Approximation pricing and the variance-optimal martingale measure. *Annals of Probability*, 24:206–236, 1996.

[63] M. Schweizer. A guided tour through quadratic hedging approaches. In É. Jouini, J. Cvitanić, and M. Musiela, editors, *Option pricing, interest rates and risk management*, pages 538–574. Cambridge University Press, Cambridge, 2001.

[64] M. Schweizer. Mean–variance hedging. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. Wiley, 2010.

[65] W. G. Sun and C. F. Wang. The mean-variance investment problem in a constrained financial market. *Journal of Mathematical Economics*, 42:885–895, 2006.

[66] J. Xia and J.-A. Yan. Markowitz's portfolio optimization in an incomplete market. *Mathematical Finance*, 16:203–216, 2006.

[67] X. Y. Zhou and D. Li. Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Applied Mathematics and Optimization*, 42:19–33, 2000.