CONVEX DUALITY IN MEAN-VARIANCE HEDGING UNDER CONVEX TRADING CONSTRAINTS

CHRISTOPH CZICHOWSKY,* *University of Vienna*

MARTIN SCHWEIZER,** *ETH Zürich and Swiss Finance Institute*

Abstract

We study mean-variance hedging under portfolio constraints in a general semimartingale model. The constraints are formulated via predictable correspondences, meaning that the trading strategy is restricted to lie in a closed convex set which may depend on the state and time in a predictable way. To obtain the existence of a solution, we first establish the closedness in L^2 of the space of all gains from trade (i.e. the terminal values of stochastic integrals with respect to the price process of the underlying assets). This is a first main contribution which enables us to tackle the problem in a systematic and unified way. In addition, using the closedness allows us to explain and generalise in a systematic way the convex duality results obtained previously by other authors via ad-hoc methods in specific frameworks.

Keywords: Mean-variance hedging; constraints; stochastic integral; convex duality

2010 Mathematics Subject Classification: Primary 60G48; 91G10; 93E20; 49N10;
60H05

1. Introduction

Mean-variance hedging and mean-variance portfolio selection are two classical problems in finance. The latter is also called the Markowitz problem and involves finding a trading strategy whose resulting final wealth has an optimal risk–reward profile, where reward and risk are measured via the mean and variance. Understanding and solving this problem is vastly simplified by a good knowledge about the general mean-variance or quadratic hedging problem. We study this in a general semimartingale financial market with general convex constraints on strategies.

In more mathematical terms, let $S = (S_t)_{0 \le t \le T}$ be an $\mathbb{R}^d$-valued semimartingale modelling the discounted prices of d risky assets. A self-financing trading strategy is described by its initial wealth $x \in \mathbb{R}$ and an $\mathbb{R}^d$-valued predictable process $\vartheta = (\vartheta_t)_{0 \le t \le T}$ describing the number of shares held dynamically over time. Its resulting final wealth is

$$V_T(x, \vartheta) := x + \int_0^T \vartheta_s \, \mathrm{d}S_s =: x + G_T(\vartheta),$$

and if the $\mathcal{F}_T$-measurable random variable H gives the time-T payoff of a financial product, *mean-variance hedging* for H requires solving the (linear-quadratic control) problem

$$\text{minimise } \mathrm{E}[|H - x - G_T(\vartheta)|^2]$$

either over $\vartheta \in \Theta_S(C)$ for fixed x or over $(x, \vartheta) \in \mathbb{R} \times \Theta_S(C)$. The space $\Theta_S(C)$ of 'allowed' integrands of course imposes a square-integrability condition on the stochastic integral process $\int \vartheta \, dS$, and the argument C in brackets indicates that we have *trading constraints* in the sense that $\vartheta_t(\omega)$ must lie in a convex closed subset $C(\omega, t)$ of $\mathbb{R}^d$. This can depend on ω and t in a predictable way, as made precise later, and it is worth pointing out that the $C(\omega, t)$ need not be cones in general. One strength of our contribution is that the above setup is essentially the most general formulation for mean-variance hedging under constraints. Under very weak local square-integrability and no-arbitrage-type assumptions on S, we give in Theorem 3.1 a necessary and sufficient condition (jointly on S and C) for the space $G_T(\Theta_S(C))$ to be closed in $L^2(\mathrm{P})$. This allows us to easily prove in Theorem 4.1 the *existence* of a solution to the mean-variance hedging problem for any $H \in L^2(\mathrm{P})$. To obtain more information on the *structure* of this solution, we then use *convex duality* tools. We introduce a dual problem whose variables and objective function both involve the constraints C through their support function. We then prove in Theorems 5.1 and 5.2 the existence of a solution to the dual problem, show how it is related to the solution of the primal problem, and give properties of the corresponding (primal and dual) value functions. There are two sets of results because we give two formulations—one in terms of static variables and the other in terms of dynamic variables.

Conceptually and resultwise, our duality approach is analogous to the classical convex duality techniques familiar from utility maximisation problems; see, for example, the work by Cvitanić and Karatzas [5], Kramkov and Schachermayer [18], and Karatzas and Žitković [17]. However, the mathematics are somewhat different since our 'quadratic random utility' $U(x, \omega) = -\frac{1}{2}|x - H(\omega)|^2$ is not increasing in x and the duality is taken in a different space. A fairly close precursor of our work is due to Labbé and Heunis [19], who studied the same problem when S is given by a complete Itô process model and the constraints do not depend on ω and t. Their duality is very similar to parts of our Theorem 5.2, but their formulations and in particular their proofs strongly depend on the availability and use of Itô's representation theorem. We do not need that at all, since S and the underlying filtration $\mathbb{F}$ are general in our setting.

The paper is structured as follows. Section 2 contains a precise problem formulation, including basic results on correspondences that we use for the modelling of constraints. Section 3 contains the central closedness result for $G_T(\Theta_S(C))$, and in Section 4 we use this to prove existence of a solution to the mean-variance hedging problem under constraints. Finally, in Section 5 we present the duality results. We first give a careful motivation for the way the dual problem is set up, both for static and dynamic variables. Then we prove the main duality theorems in those two settings, and we close the section with more detailed comments on and comparison to the literature.

2. Formulation of the problem

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ satisfying the usual conditions of completeness and right continuity, where $T > 0$ is a fixed and finite time horizon. Hence we can and do choose, for every local P-martingale, a version whose trajectories are right continuous with left limits (RCLL). For all unexplained notions concerning stochastic integration, we refer the reader to [24].

We consider a *financial market* consisting of one riskless asset, whose (discounted) price is 1, and d risky assets described by an $\mathbb{R}^d$-valued semimartingale S. We denote by $\mathcal{H}^2(\mathrm{P})$ the Banach space of all square-integrable semimartingales, i.e. special semimartingales Y with canonical decomposition $Y = Y_0 + M^Y + A^Y$, where M^Y is a square-integrable martingale Y

null at 0, $M^Y \in \mathcal{M}_0^2(\mathrm{P})$, and A^Y is a predictable finite variation RCLL process null at 0, such that

$$\|Y\|_{\mathcal{H}^2(\mathrm{P})} := \|Y_0\|_{L^2(\mathrm{P})} + \|([M^Y]_T)^{1/2}\|_{L^2(\mathrm{P})} + \left\| \int_0^T |\mathrm{d}A_s^Y| \right\|_{L^2(\mathrm{P})} < \infty.$$

Note that $\mathcal{H}_{\mathrm{loc}}^2(\mathrm{P})$ coincides with the semimartingale space $\mathcal{S}_{\mathrm{loc}}^2(\mathrm{P})$. We suppose that S is a locally square-integrable semimartingale, denoted by $S \in \mathcal{H}_{\mathrm{loc}}^2(\mathrm{P})$, with canonical decomposition $S = S_0 + M + A$. Then there exists a predictable increasing RCLL process B, e.g. $B = \sum_{i=1}^d (\langle M^i \rangle + \int |\mathrm{d}A^i|)$, with $\langle M^i, M^j \rangle \ll B$ and $A^i \ll B$ for $i, j = 1, \ldots, d$. We define an $\mathbb{R}^{d \times d}$-valued predictable process c^M and an $\mathbb{R}^d$-valued predictable process a by $(c^M)^{ij} = \mathrm{d}\langle M^i, M^j \rangle / \mathrm{d}B$ and $a^i = \mathrm{d}A^i / \mathrm{d}B$. We set $\overline{\Omega} := \Omega \times [0, T]$ and $\mathrm{P}_B := \mathrm{P} \otimes B$, and view $\mathbb{R}^d$-valued predictable processes as $\mathcal{P}$-measurable random variables, i.e. elements of $\mathcal{L}^0(\overline{\Omega}, \mathcal{P}; \mathbb{R}^d)$. For *trading strategies*, we take

$$\Theta := \Theta_S := \left\{ \vartheta \in \mathcal{L}(S) \,\middle|\, \int \vartheta \, \mathrm{d}S \in \mathcal{H}^2(\mathrm{P}) \right\},$$

where $\mathcal{L}(Y)$ denotes the space of all $\mathbb{R}^d$-valued, Y-integrable, predictable processes for a semimartingale Y. Note that we work with processes without identifying ϑ and ϑ' when $\int \vartheta \, \mathrm{d}S = \int \vartheta' \, \mathrm{d}S$; hence we write $\mathcal{L}(S)$, not $L(S)$. By the uniqueness of the canonical decomposition we have $\Theta_S = \mathcal{L}^2(M) \cap \mathcal{L}^2(A)$ with

$$\mathcal{L}^2(M) := \left\{ \vartheta \in \mathcal{L}^0(\overline{\Omega}, \mathcal{P}; \mathbb{R}^d) \,\middle|\, \|\vartheta\|_{\mathcal{L}^2(M)} := \left(\mathrm{E}\left[\int_0^T \vartheta_s^\top c_s^M \vartheta_s \, \mathrm{d}B_s \right] \right)^{1/2} < \infty \right\},$$

$$\mathcal{L}^2(A) := \left\{ \vartheta \in \mathcal{L}^0(\overline{\Omega}, \mathcal{P}; \mathbb{R}^d) \,\middle|\, \|\vartheta\|_{\mathcal{L}^2(A)} := \left(\mathrm{E}\left[\left(\int_0^T |\vartheta_s^\top a_s| \, \mathrm{d}B_s \right)^2 \right] \right)^{1/2} < \infty \right\}.$$

The wealth process generated up to time $t \in [0, T]$ by a self-financing trading strategy ϑ with initial capital $x \in \mathbb{R}$ is

$$V_t(x, \vartheta) := x + \int_0^t \vartheta_s \, \mathrm{d}S_s =: x + G_t(\vartheta),$$

where the process $G(\vartheta)$ denotes the cumulative gains from trading. The set of all outcomes of self-financing trading strategies with zero initial wealth is

$$G_T(\Theta_S) = \{G_T(\vartheta) \mid \vartheta \in \Theta_S\},$$

and the set of attainable payoffs is

$$\mathcal{A}(\Theta_S) = \mathbb{R} + G_T(\Theta_S).$$

In contrast to strategies, we identify here final wealths that are equal P-almost surely (P-a.s.). Owing to the definition of Θ_S, the sets $G_T(\Theta_S)$ and $\mathcal{A}(\Theta_S)$ are thus linear subspaces of $L^2(\mathrm{P})$ by the following result.

Proposition 2.1. *Let Y be a semimartingale in $\mathcal{H}^2(\mathrm{P})$, and let $Y_t^* := \sup_{0 \leq s \leq t} |Y_s|$. Then*

$$\mathrm{E}[(Y_T^*)^2] \leq 8\|Y\|_{\mathcal{H}^2(\mathrm{P})}^2.$$

Proof. See Theorem IV.5 of [24]. $\square$

A square-integrable $\mathcal{F}_T$-measurable random variable H is called a *contingent claim*. We assume that an investor wants to hedge a contingent claim by means of a self-financing trading strategy. However, since the market is usually incomplete, perfect replication of the contingent claim, in the sense that $H = V_T(x, \vartheta)$, P-a.s. for some x and ϑ, is in general impossible. So the investor wants to optimise the hedging performance of his trading strategy according to some criterion. One possible choice, especially when the investor simultaneously considers buying or selling H, is the minimisation of the mean-squared hedging error, which leads to the approximation problem

$$\mathrm{E}\left[\left|H - x - \int_0^T \vartheta_s \, \mathrm{d}S_s\right|^2\right] = \min_{(x, \vartheta) \in \mathbb{R} \times \Theta_S} !$$

For a fixed initial capital $x \in \mathbb{R}$, we obtain the problem of *mean-variance hedging*, i.e.

$$\mathrm{E}\left[\left|H - x - \int_0^T \vartheta_s \, \mathrm{d}S_s\right|^2\right] = \min_{\vartheta \in \Theta_S} !$$

Mathematically, this amounts to projecting $H - x$ onto $G_T(\Theta_S)$ or H onto $\mathcal{A}(\Theta_S)$. Therefore, a solution for every $H \in L^2(\mathrm{P})$ exists if and only if $G_T(\Theta_S)$ and $\mathcal{A}(\Theta_S)$ are closed in $L^2(\mathrm{P})$. Note that both problems are naturally studied in $L^2(\mathrm{P})$ rather than $\mathcal{L}^2(\mathrm{P})$.

Before we introduce the mean-variance hedging problem under *trading constraints*, we make the following simple observation. In the unconstrained case, when $G_T(\Theta_S)$ and $\mathcal{A}(\Theta_S)$ are closed linear subspaces, the problem admits a unique solution by elementary Hilbert space arguments. Under trading constraints, this is still true if the subsets in which we want to find the best approximation are closed and convex subsets of $L^2(\mathrm{P})$. Despite its simplicity, this observation is very useful. We shall see that mean-variance hedging problems can be embedded into this framework even under the additional constraint that the trading strategy only takes values in a closed convex set, which is allowed to depend on the state ω and time t in a predictable way. This allows us to treat these problems in a systematic and unified way.

To model *'predictable trading constraints'*, we formulate them via predictable correspondences. This idea is analogous to [15], where it is used to study the existence of the numéraire portfolio under predictable convex constraints.

Definition 2.1. A mapping $C \colon \overline{\Omega} \to 2^{\mathbb{R}^d}$ is called a *correspondence*. We say that a correspondence C is *predictable* if $C^{-1}(F) = \{(\omega, t) \mid C(\omega, t) \cap F \neq \varnothing\}$ is a predictable set (i.e. in $\mathcal{P}$) for all closed $F \subseteq \mathbb{R}^d$. The *domain* dom(C) of a correspondence is given by dom$(C) = \{(\omega, t) \mid C(\omega, t) \neq \varnothing\}$. A *(predictable) selector* of a (predictable) correspondence C is a (predictable) process ψ with $\psi(\omega, t) \in C(\omega, t)$ for all $(\omega, t) \in \text{dom}(C)$.

For convenience, we recall some results on predictable correspondences which ensure the existence of predictable selectors in all situations relevant for us.

Proposition 2.2. (Castaing.) *For a correspondence $C \colon \overline{\Omega} \to 2^{\mathbb{R}^d}$ with closed values, the following statements are equivalent:*

(a) *C is predictable,*

(b) dom(C) *is predictable and there exists a* Castaing representation *of C, i.e. a sequence (ψ^n) of predictable selectors of C such that*

$$C(\omega, t) = \overline{\{\psi^1(\omega, t), \psi^2(\omega, t), \ldots\}} \quad \text{for each } (\omega, t) \in \text{dom}(C).$$

In particular, every predictable closed-valued C admits a predictable selector ψ.

Proof. See Corollary 18.14 of [1] or Theorem 1B of [26].

Proposition 2.3. *Let* $C\colon \overline{\Omega} \to 2^{\mathbb{R}^d}$ *be a predictable correspondence with closed values, and let* $f\colon \overline{\Omega} \times \mathbb{R}^m \to \mathbb{R}^d$ *and* $g\colon \overline{\Omega} \times \mathbb{R}^d \to \mathbb{R}^m$ *be* Carathéodory *functions, which means that* $f(\omega, t, y)$ *and* $g(\omega, t, x)$ *are predictable with respect to* (ω, t) *and continuous in* y *and* x. *Then the mappings* C' *and* C'' *given by* $C'(\omega, t) = \{y \in \mathbb{R}^m \mid f(\omega, t, y) \in C(\omega, t)\}$ *and* $C''(\omega, t) = \overline{\{g(\omega, t, x) \mid x \in C(\omega, t)\}}$ *are predictable correspondences (from* $\overline{\Omega}$ *to* $2^{\mathbb{R}^m}$*) with closed values.*

Proof. See Corollaries 1P and 1Q of [26].

Proposition 2.4. *Let* $C^n\colon \overline{\Omega} \to 2^{\mathbb{R}^d}$ *for each* $n \in \mathbb{N}$ *be a predictable correspondence with closed values, and define the correspondences* C' *and* C'' *by* $C'(\omega, t) = \bigcap_{n \in \mathbb{N}} C^n(\omega, t)$ *and* $C''(\omega, t) = \bigcup_{n \in \mathbb{N}} C^n(\omega, t)$. *Then* C' *and* C'' *are predictable and* C' *is closed valued.*

Proof. See Theorem 1M of [26] and Lemma 18.4 of [1].

For a predictable correspondence $C\colon \overline{\Omega} \to 2^{\mathbb{R}^d} \setminus \{\varnothing\}$, we denote by

$$\mathcal{C} := \mathcal{C}^S := \{\psi \in \mathcal{L}(S) \mid \psi(\omega, t) \in C(\omega, t) \text{ for all } (\omega, t) \in \overline{\Omega}\}$$

the set of *C-valued* or *C-constrained integrands* for S and by

$$\Theta_S(C) = \Theta_S \cap \mathcal{C} = \{\vartheta \in \Theta_S \mid \vartheta(\omega, t) \in C(\omega, t) \text{ for all } (\omega, t) \in \overline{\Omega}\}$$

the set of all *C-constrained trading strategies*. With this formulation, the set of all outcomes of C-constrained self-financing trading strategies with zero initial wealth is

$$G_T(\Theta_S(C)) = \{G_T(\vartheta) \mid \vartheta \in \Theta_S(C)\},$$

and the set of payoffs that are attainable with C-constrained trading strategies is

$$\mathcal{A}(\Theta_S(C)) = \mathbb{R} + G_T(\Theta_S(C)).$$

Note that one can, for example, model prohibition of short-selling or rectangular constraints in this formulation; see Examples 4.1 in Section 5.4 of [16]. The *mean-variance hedging problem under trading constraints* is now formulated as

$$\mathrm{E}\!\left[\left|H - x - \int_0^T \vartheta_s \, \mathrm{d}S_s\right|^2\right] = \min_{\vartheta \in \Theta_S(C)}!$$

for a fixed initial capital $x \in \mathbb{R}$, and we also study

$$\mathrm{E}\!\left[\left|H - x - \int_0^T \vartheta_s \, \mathrm{d}S_s\right|^2\right] = \min_{(x, \vartheta) \in \mathbb{R} \times \Theta_S(C)}!$$

when including the initial capital into the approximation problem. As already explained above, these problems admit solutions if $G_T(\Theta_S(C))$ and $\mathcal{A}(\Theta_S(C))$ are closed and convex subsets of $L^2(\mathrm{P})$. By the convexity of C, the convexity of $G_T(\Theta_S(C))$ and $\mathcal{A}(\Theta_S(C))$ immediately follows. The closedness will be established in the next section.

3. The closedness of $G_T(\Theta_S(C))$ and $\mathcal{A}(\Theta_S(C))$

In this section we show that the set of all outcomes of C-constrained self-financing trading strategies with zero initial wealth and the set of all payoffs that are attainable with C-constrained trading strategies are both closed in $L^2(\mathrm{P})$. To that end, we use the concept of $\mathcal{E}$-martingales, which was introduced and developed by Choulli, Krawczyk and Stricker in [4] to deduce the closedness of the analogous subspaces in the unconstrained case. For easy reference, we start by briefly recalling some definitions and results.

For a semimartingale Y, we denote its stochastic exponential by $\mathcal{E}(Y)$. Throughout this paper, let N be a fixed local P-martingale starting at 0. For any stopping time τ, we denote the process Y stopped at τ by Y^τ and the process Y started at τ by ${}^\tau Y = Y - Y^\tau$, but we set ${}^\tau\mathcal{E} = {}^\tau\mathcal{E}(N) = \mathcal{E}(N - N^\tau)$. So, for the stochastic exponential, ${}^\tau\mathcal{E}(N)$ denotes a multiplicative rather than an additive restarting. Since N is P-a.s. RCLL, it has at most a finite number of jumps with $\Delta N = -1$, and, moreover, each ${}^\tau\mathcal{E}(N)$ can jump to 0 at most once. Therefore, we can define an increasing sequence of stopping times by $\hat{T}_0 = 0$ and $\hat{T}_{n+1} = \inf\{t > \hat{T}_n \mid {}^{\hat{T}_n}\mathcal{E}(N)_t = 0\} \wedge T = \inf\{t > \hat{T}_n \mid \Delta N_t = -1\} \wedge T$, and $(\hat{T}_n)$ tends to T stationarily; this follows from the representation of the stochastic exponential in Theorem II.37 of [24].

Definition 3.1. An adapted RCLL process Y is an *$\mathcal{E}$-local martingale* if the product of ${}^{\hat{T}_n}Y$ and ${}^{\hat{T}_n}\mathcal{E}$ is a local P-martingale for any $n \in \mathbb{N}$. It is an *$\mathcal{E}$-martingale* if, for any $n \in \mathbb{N}$, we have $\mathrm{E}[|Y_{\hat{T}_n}{}^{\hat{T}_n}\mathcal{E}_{\hat{T}_{n+1}}|] < \infty$ and the above product is a (true) P-martingale.

The next two propositions, which are Corollaries 3.16 and 3.17 of [4], give some information about the structure of $\mathcal{E}$-martingales.

Proposition 3.1. *Let Y be a special semimartingale, and let $Y = Y_0 + M^Y + A^Y$ be its canonical decomposition. Then Y is an $\mathcal{E}$-local martingale if and only if $[M^Y, N]$ is locally P-integrable and $A^Y = -\langle M^Y, N \rangle$.*

Proposition 3.2. *A semimartingale $Y = Y_0 + M^Y - \langle M^Y, N \rangle$ with $\mathrm{E}[Y_T^*({}^{\hat{T}_n}\mathcal{E})_T^*] < \infty$ for any $n \in \mathbb{N}$ is an $\mathcal{E}$-martingale.*

We also need the following definitions.

Definition 3.2. We say that $\mathcal{E}$ is *regular* if ${}^{\hat{T}_n}\mathcal{E}$ is a P-martingale for any n.

Definition 3.3. We say that $\mathcal{E}$ satisfies *the reverse Hölder inequality $R_2(\mathrm{P})$* if there exists a constant $c \geq 1$ such that $\mathrm{E}[|{}^t\mathcal{E}_T|^2 \mid \mathcal{F}_t] \leq c$ for any t.

The next proposition is a partial statement of Proposition 3.9 of [4].

Proposition 3.3. *Assume that $\mathcal{E}$ satisfies $R_2(\mathrm{P})$. Then $\mathcal{E}$ is regular if and only if ${}^\tau\mathcal{E}$ is a P-martingale for any stopping time τ, and in that case, ${}^\tau\mathcal{E}$ is a P-square-integrable P-martingale.*

Finally, a combination of Theorem 4.9 of [4] and Proposition 2.1 gives the following equivalence of norms.

Proposition 3.4. *Assume that $\mathcal{E}$ is regular and satisfies $R_2(\mathrm{P})$. Then there exists a constant c such that*

$$\frac{1}{c}\|Y\|_{\mathcal{H}^2(\mathrm{P})} \leq \|Y_T\|_{L^2(\mathrm{P})} \leq c\|Y\|_{\mathcal{H}^2(\mathrm{P})}$$

for every $\mathcal{E}$-martingale Y. We write this for short as $\|Y\|_{\mathcal{H}^2(\mathrm{P})} \sim \|Y_T\|_{L^2(\mathrm{P})}$.

Note that, when $^0\mathcal{E}(N)$ is a strictly positive P-martingale, the definition of an $\mathcal{E}$-local martingale coincides with the notion of a local martingale under the measure Q defined by $\mathrm{dQ} = {}^0\mathcal{E}(N)_T\,\mathrm{dP}$. This will be called the *classical case*.

As explained in the previous section, we consider a possibly incomplete financial market composed of one riskless asset, whose price is 1, and d risky assets described by an $\mathbb{R}^d$-valued semimartingale $S \in \mathcal{H}^2_{\mathrm{loc}}(\mathrm{P})$ with canonical decomposition $S = S_0 + M + A$. We suppose that there exists $N \in \mathcal{M}_{0,\mathrm{loc}}(\mathrm{P})$ such that S is an $\mathcal{E}$-local martingale. By Proposition 3.1, this implies that $\langle M, N \rangle$ exists and $A = -\langle M, N \rangle$. Moreover, we assume that $\mathcal{E}(N)$ satisfies $R_2(\mathrm{P})$, so N is locally P-square-integrable and in bmo_2, i.e. there exists a constant $c > 0$ such that $\mathrm{E}[\langle N \rangle_T - \langle N \rangle_t \mid \mathcal{F}_t] \leq c$ for all $t \in [0, T]$; see Proposition 3.10 of [4]. An application of the Kunita–Watanabe decomposition yields $N = -\int \lambda\,\mathrm{d}M + L$ with $\lambda \in \mathcal{L}^2(M)$ and $L \in \mathcal{M}^2_0(\mathrm{P})$ strongly P-orthogonal to M, and, hence, S satisfies the *structure condition*, i.e.

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

Since N is in bmo_2, $\int \lambda\,\mathrm{d}M$ is also in bmo_2, which implies by Theorem 3.3 of [9] the inequality $D_2(\mathrm{P})$, i.e. there exists a constant $c > 0$ such that $\|\vartheta\|_{\mathcal{L}^2(A)} \leq c\|\vartheta\|_{\mathcal{L}^2(M)}$ for all $\vartheta \in \mathcal{L}^2(M)$. As a consequence, we have $\Theta_S = \mathcal{L}^2(M)$. To motivate the closedness proof under trading constraints, we give below the argument for the unconstrained case, which is due to Choulli *et al.* [4, Theorem 5.2].

Proposition 3.5. *Assume that $\mathcal{E} = \mathcal{E}(N)$ is regular and satisfies $R_2(\mathrm{P})$, and that $S \in \mathcal{H}^2_{\mathrm{loc}}(\mathrm{P})$ is an $\mathcal{E}$-local martingale. Then the following statements hold.*

(a) *For each σ-field $\mathcal{B}_0 \subseteq \mathcal{F}_0$ and each $Y_0 \in L^2(\mathcal{B}_0)$, the process $Y_0 + \int \vartheta\,\mathrm{d}S \in \mathcal{H}^2(\mathrm{P})$ is an $\mathcal{E}$-martingale.*

(b) *The spaces $G_T(\Theta_S)$, $\mathcal{A}(\Theta_S)$, and $L^2(\mathcal{B}_0) + G_T(\Theta_S)$ for any σ-field $\mathcal{B}_0 \subseteq \mathcal{F}_0$ are closed in $L^2(\mathrm{P})$.*

Proof. (a) The stochastic integral $\int \vartheta\,\mathrm{d}S$ is for each $\vartheta \in \Theta_S$ a special semimartingale in $\mathcal{H}^2(\mathrm{P})$ with canonical decomposition $\int \vartheta\,\mathrm{d}S = \int \vartheta\,\mathrm{d}M + \int \vartheta\,\mathrm{d}A$. Since S is an $\mathcal{E}$-local martingale, we have $A = -\langle M, N \rangle$ by Proposition 3.1; so $\int \vartheta\,\mathrm{d}A = -\langle \int \vartheta\,\mathrm{d}M, N \rangle$, and, therefore, $\int \vartheta\,\mathrm{d}S$ is an $\mathcal{E}$-local martingale again by Proposition 3.1. Since $\mathcal{E}$ is regular and satisfies $R_2(\mathrm{P})$, Proposition 3.3 states that $^\tau\mathcal{E}$ is a square-integrable martingale for each stopping time τ. By Doob's inequality and Proposition 2.1, $\{^\tau\mathcal{E}\}^*_T$ and $\{G(\vartheta)\}^*_T$ are in $L^2(\mathrm{P})$ so that $\{G(\vartheta)\}^*_T\{^\tau\mathcal{E}\}^*_T$ is in $L^1(\mathrm{P})$ for every stopping time τ. Proposition 3.2 now implies that $G(\vartheta)$ is an $\mathcal{E}$-martingale. Replacing $G(\vartheta)$ by Y_0 shows in the same way that the constant process Y_0 is an $\mathcal{E}$-martingale for any $Y_0 \in L^2(\mathcal{B}_0)$, and, hence, so is $Y_0 + G(\vartheta)$.

(b) Let $(Y_0^n + G_T(\vartheta^n))$ be a sequence in $L^2(\mathcal{B}_0) + G_T(\Theta_S)$ converging to H in $L^2(\mathrm{P})$. By part (a), each $Y_0^n + G(\vartheta^n)$ is an $\mathcal{E}$-martingale and, therefore, the sequence $(Y_0^n + G(\vartheta^n))$ is a Cauchy sequence in the Banach space $\mathcal{H}^2(\mathrm{P})$ by Proposition 3.4, and, hence, convergent to some $Y \in \mathcal{H}^2(\mathrm{P})$ which satisfies $Y_T = H$. Since the space (of processes) $Y_0 + G(\Theta_S)$ is closed in $\mathcal{H}^2(\mathrm{P})$ (either by the construction of the stochastic integral as in Section IV.2 of [24] or by Theorem V.4 of [20]), there exists some $\vartheta \in \Theta_S$ with $Y = Y_0 + G(\vartheta)$, and, therefore, $L^2(\mathcal{B}_0) + G_T(\Theta_S)$ is closed in $L^2(\mathrm{P})$. Choosing first $\mathcal{B}_0 = \{\varnothing, \Omega\}$ and then $Y_0^n = 0$ for all $n \in \mathbb{N}$ in the above argument, we obtain the closedness of $\mathcal{A}(\Theta_S)$ and $G_T(\Theta_S)$ in $L^2(\mathrm{P})$, which completes the proof. $\qquad\square$

Remark 3.1. Assuming that there exists $N \in \mathcal{M}_{0,\mathrm{loc}}(\mathrm{P})$ such that $\mathcal{E}(N)$ is regular and P-square-integrable and such that S is an $\mathcal{E}$-local martingale implies the weak no-arbitrage condition that

$G_T(\Theta_S)$ admits *no approximate profits in L^2*; this means that $1 \notin \overline{G_T(\Theta_S)}$, where the overline denotes the closure in $L^2(P)$. See Section 4 of [27].

To obtain the closedness under constraints, we observe the following. If $(Y_0^n + G_T(\vartheta^n))$ is a sequence in $L^2(\mathcal{B}_0) + G_T(\Theta_S(C))$ converging to some H in $L^2(P)$, then there exist, under the assumptions of Proposition 3.5, some $Y_0 \in L^2(\mathcal{B}_0)$ and some $\vartheta \in \Theta_S$ such that $Y_0 + G_T(\vartheta) = H$ and $(Y_0^n + G(\vartheta^n))$ converges to $Y_0 + G(\vartheta)$ even in $\mathcal{H}^2(P)$. The question is then whether ϑ can be chosen to be C-valued. In general (even if S is a martingale), the answer is negative, as a simple counterexample in Section 3 of [6] illustrates. The reason behind this is that the linear dependence of the different components of S can make some of the risky assets *redundant* in the sense that one of them can be replicated on some predictable set by trading in the other ones. As a consequence, there may exist different strategies with the same gains process, and so a trading strategy is not uniquely determined (even up to indistinguishability) by its gains process. Indeed, by the construction of the stochastic integral, two strategies ψ and φ in Θ_S have the same gains process up to indistinguishability if and only if $c^M(\psi - \varphi) = 0$ and $a^\top(\psi - \varphi) = 0$, P_B-almost everywhere (P_B-a.e.); see Lemma 5.1 of [6]. Therefore, the convergence of the gains processes $G(\vartheta^n)$ need not imply the pointwise convergence of the strategies ϑ^n, and so the pointwise closedness of $C(\omega, t)$ is not sufficient to deduce that ϑ can be chosen to be C-valued. However, the convergence of the gains processes is sufficient to prove that we do get a convergent sequence (ψ^n) of modified strategies which have the same gains processes as the strategies ϑ^n. This sequence is given by the projection of (ϑ^n) on the predictable range of S.

Proposition 3.6. *For each $\mathbb{R}^d$-valued semimartingale Y, there exists an $\mathbb{R}^{d \times d}$-valued predictable process Π^Y, called the* projection on the predictable range of Y, *which takes values in the orthogonal projections in $\mathbb{R}^d$ and has the following property. If $\vartheta \in \mathcal{L}(Y)$ and φ is predictable, then φ is in $\mathcal{L}(Y)$ with $\int \varphi \, dY = \int \vartheta \, dY$ (up to indistinguishability) if and only if $\Pi^Y \vartheta = \Pi^Y \varphi$, P_B-a.e. We choose and fix one version of Π^Y.*

Proof. See Lemma 5.3 of [6].

Remark 3.2. Suppose that S satisfies the structure condition. Then the construction of the stochastic integral gives $\int \psi \, dS = 0$ if and only if $\int \psi \, dM = 0$, and, therefore, that Π^S and Π^M coincide (P_B-a.e.). Since $\int \psi \, dM = 0$ if and only if $c^M \psi = 0$, P_B-a.e., we see that $\psi = \vartheta - \Pi^M \vartheta$ is P_B-a.e. valued in $\text{Ker}(c^M)$, the kernel of the matrix c^M. Moreover, since Π^M is an orthogonal projection, so is $\text{Id}_{d \times d} - \Pi^M$, and, therefore, $\Pi^M = \Pi^S$ can be chosen as the pointwise orthogonal projection on $\text{range}(c^M(\omega, t))$, the range of the matrix $c^M(\omega, t)$, which is equal to $\text{Ker}(c^M(\omega, t))^\perp$.

Using the projection on the predictable range, we can now prove the closedness in the constrained case.

Theorem 3.1. *Assume that $\mathcal{E} = \mathcal{E}(N)$ is regular and satisfies $R_2(P)$, and that $S \in \mathcal{H}^2_{\text{loc}}(P)$ is an $\mathcal{E}$-local martingale. Let $C: \overline{\Omega} \to 2^{\mathbb{R}^d}$ be a predictable correspondence with closed values and such that $\Theta_S(C)$ is nonempty. Then the spaces $G_T(\Theta_S(C))$, $\mathcal{A}(\Theta_S(C))$, and $L^2(\mathcal{B}_0) + G_T(\Theta_S(C))$ for any σ-field $\mathcal{B}_0 \subseteq \mathcal{F}_0$ are closed in $L^2(P)$ if and only if the projection of C on the predictable range of S is closed, i.e. $\Pi^S(\omega, t) C(\omega, t)$ is closed P_B-a.e.*

Proof. For the '$\Leftarrow$' part, let $(Y_0^n + G_T(\vartheta^n))$ be a sequence in $L^2(\mathcal{B}_0) + G_T(\Theta_S(C))$ converging to H in $L^2(P)$. Then there exist some $Y_0 \in L^2(\mathcal{B}_0)$ and some $\vartheta \in \Theta_S$ such that $Y_0 + G_T(\vartheta) = H$, P-a.s., and $(Y_0^n + G(\vartheta^n))$ converges to $Y_0 + G(\vartheta)$ in $\mathcal{H}^2(P)$, by the proof

of Proposition 3.5. The convergence in $\mathcal{H}^2(\mathrm{P})$ implies the convergence in the semimartingale topology by Theorem V.14 and the lemma preceding Theorem IV.12 of [24]. By Theorem 4.5 of [6], the space of stochastic integrals of C-valued integrands is closed in the semimartingale topology if the projection of C on the predictable range of S is closed. Thus, there exists $\widetilde{\vartheta} \in \Theta_S(C)$ such that $G(\widetilde{\vartheta}) = G(\vartheta)$, and, therefore, $L^2(\mathcal{B}_0) + G_T(\Theta_S(C))$ is closed in $L^2(\mathrm{P})$. As in the proof of Proposition 3.5, choosing $\mathcal{B}_0 = \{\varnothing, \Omega\}$ and $Y_0^n = 0$ for all $n \in \mathbb{N}$ gives the closedness of $\mathcal{A}(\Theta_S)$ and $G_T(\Theta_S)$.

For the '$\Rightarrow$' part, first note that, for any stopping time τ, the projection Π^{S^τ} on the predictable range of S^τ simply $\Pi^S \mathbb{I}_{[\![0,\tau]\!]}$, where $[\![0,\tau]\!]$ denotes, as in [24], the stochastic interval from 0 to τ and $\mathbb{I}.$ denotes the indicator funtion. Recall that $S \in \mathcal{H}^2_{\mathrm{loc}}(\mathrm{P})$. Arguing by contradiction, we choose a stopping time τ such that $S^\tau \in \mathcal{H}^2(\mathrm{P})$ and Π^{S^τ} is not closed. Applying Lemma 4.4 of [6] with S^τ and using the fact that $\int \varphi \, \mathrm{d}S^\tau = \int \varphi \mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S$ for any $\varphi \in \mathcal{L}(S)$ implies that there exist $\vartheta \in \mathcal{L}(S)$ and a sequence (ψ^n) of C-valued integrands such that $(\int \psi^n \mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S)$ converges to $\int \vartheta \mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S$ in the semimartingale topology, but there is no C-valued integrand ψ such that $\int \psi \mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S = \int \vartheta \mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S$. An inspection of the proof of Lemma 4.4 of [6] shows that we can choose ϑ and (ψ^n) such that $(\Pi^S \vartheta)\mathbb{I}_{[\![0,\tau]\!]}$ and $(\Pi^S \psi^n)\mathbb{I}_{[\![0,\tau]\!]}$ are uniformly bounded and $(\Pi^S \psi^n)\mathbb{I}_{[\![0,\tau]\!]} \to (\Pi^S \vartheta)\mathbb{I}_{[\![0,\tau]\!]}$ uniformly in (ω, t). Since

$$\int \psi^n \mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S = \int (\Pi^S \psi^n)\mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S \quad \text{and} \quad \int \vartheta \mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S = \int (\Pi^S \vartheta)\mathbb{I}_{[\![0,\tau]\!]} \, \mathrm{d}S,$$

we have, by dominated convergence, $\int (\psi^n \mathbb{I}_{[\![0,\tau]\!]} + \varphi \mathbb{I}_{]\!]\tau,T]\!]}) \, \mathrm{d}S \to \int \widetilde{\vartheta} \, \mathrm{d}S$ in $\mathcal{H}^2(\mathrm{P})$ for any $\varphi \in \Theta_S(C)$, with $\widetilde{\vartheta} = \vartheta \mathbb{I}_{[\![0,\tau]\!]} + \varphi \mathbb{I}_{]\!]\tau,T]\!]}$, and, hence, also $G_T(\psi^n \mathbb{I}_{[\![0,\tau]\!]} + \varphi \mathbb{I}_{]\!]\tau,T]\!]}) \to G_T(\widetilde{\vartheta})$ in $L^2(\mathrm{P})$ by Proposition 2.1. But, because there exists by construction of $\widetilde{\vartheta}$ no C-valued integrand ψ with $G(\psi) = G(\widetilde{\vartheta})$ and since $G(\widetilde{\vartheta})$ is uniquely determined in $\mathcal{H}^2(\mathrm{P})$ by its terminal value $G_T(\widetilde{\vartheta})$ by Proposition 3.4, there cannot be any $\psi \in \Theta_S(C)$ with $G_T(\psi) = G_T(\widetilde{\vartheta})$. This contradicts the closedness of $G_T(\Theta_S(C))$ in $L^2(\mathrm{P})$ and, therefore, completes the proof.

For a better understanding of our assumptions, we now spell them out in a multidimensional Itô process model. This is one standard example of a financial market, and it illustrates that our assumptions are weaker than those in [13], [14], and [19].

Example 3.1. Let W be an $\mathbb{R}^n$-valued Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathrm{P})$ with a filtration satisfying the usual conditions. Note that, for our results, $\mathbb{F}$ need not be the P-augmentation of the filtration generated by W; we do not use Itô's representation theorem. Let $\bar{S} = (\bar{S}^i)_{i=1,\ldots,d}$ be the undiscounted price processes of the d risky assets, and let $\bar{S}^0$ be the undiscounted price of the bank account. These processes are given as the solutions to the stochastic differential equations (SDEs)

$$\mathrm{d}\bar{S}^i_t = \bar{S}^i_t \left(\mu^i_t \, \mathrm{d}t + \sum_{j=1}^n \sigma^{ij}_t \, \mathrm{d}W^j_t \right), \qquad \bar{S}^i_0 = s^i_0,$$

for $i = 1, \ldots, d$ with constants $s^i_0 > 0$ and

$$\mathrm{d}\bar{S}^0_t = \bar{S}^0_t r_t \, \mathrm{d}t, \qquad \bar{S}^0_0 = 1,$$

with predictable $\mathbb{R}^d$-, $\mathbb{R}$-, and $\mathbb{R}^{d \times n}$-valued processes μ, r, and σ that are P-a.s. on $[0, T]$ Lebesgue-integrable and Lebesgue-square-integrable, respectively. In our abstract setup, we

consider the discounted prices $S^i = \bar{S}^i/\bar{S}^0$. The SDEs for the S^i then take the form

$$\mathrm{d}S_t^i = S_t^i \left((\mu_t^i - r_t)\, \mathrm{d}t + \sum_{j=1}^n \sigma_t^{ij}\, \mathrm{d}W_t^j \right), \qquad S_0^i = s_0^i,$$

and we explicitly have

$$\mathrm{d}\langle M\rangle_t = \mathrm{diag}(S_t)\sigma_t\sigma_t^\top \mathrm{diag}(S_t)\, \mathrm{d}t =: c_t^M\, \mathrm{d}t,$$

and

$$\mathrm{d}A_t = \mathrm{diag}(S_t)(\mu_t - r_t\mathbf{1})\, \mathrm{d}t =: a_t\, \mathrm{d}t$$

with $\mathbf{1} = (1,\dots,1)^\top \in \mathbb{R}^d$. Up to integrability conditions on μ, r, and σ, the process S satisfies the structure condition if and only if $(\mu - r\mathbf{1}) \in \mathrm{range}(\sigma\sigma^\top)$ $\mathrm{dP}\otimes\mathrm{d}t$-a.e., since $S_t^i > 0$ for $i = 1,\dots,d$. For a sufficient condition for this, we assume that $\sigma\sigma^\top$ is $\mathrm{dP}\otimes\mathrm{d}t$-a.e. invertible, which means that $n \geq d$ and that σ has $\mathrm{dP}\otimes\mathrm{d}t$-a.e. full rank d. This condition also implies that $\Pi^S = \mathrm{Id}_{d\times d}$ and, therefore, that the projection of *any* closed-valued predictable correspondence C on the predictable range of S is closed. A natural candidate to obtain a local martingale N such that S is an $\mathcal{E}(N)$-martingale is $N = -\int \lambda\, \mathrm{d}M = -\int \varphi\, \mathrm{d}W$, where $\varphi = \sigma^\top\lambda = \sigma^\top(\sigma\sigma^\top)^{-1}(\mu - r\mathbf{1})$ is the *instantaneous market price of risk*. Here we make the frequently used assumption that the *mean-variance tradeoff (MVT) process*

$$K_t := \int_0^t \lambda_s^\top\, \mathrm{d}\langle M\rangle_s \lambda_s = \left\langle \int \lambda\, \mathrm{d}M \right\rangle_t = \left\langle \int \varphi\, \mathrm{d}W \right\rangle_t = \int_0^t |\varphi_s|^2\, \mathrm{d}s,$$

which coincides in this setup with the integrated squared market price of risk, is uniformly bounded in t and ω. This is sufficient to guarantee that $\mathcal{E}(-\int \lambda\, \mathrm{d}M)$ is a true martingale and satisfies $R_2(\mathrm{P})$ by Proposition 3.7 of [4]. As M is continuous, $\mathcal{E}(-\int \lambda\, \mathrm{d}M)$ is strictly positive and $\mathrm{d}\hat{\mathrm{P}} = \mathcal{E}(-\int \lambda\, \mathrm{d}M)_T\, \mathrm{dP}$ defines an equivalent local martingale measure (ELMM) for the process S, the so-called *minimal martingale measure*; see [27]. Thus, we can conclude that if the MVT process K is uniformly bounded in t and ω, and $\sigma\sigma^\top$ is $\mathrm{dP}\otimes\mathrm{d}t$-a.e. invertible, the assumptions of Theorem 3.1 are satisfied and $G_T(\Theta_S(C))$ and $\mathcal{A}(\Theta_S(C))$ are closed in $L^2(\mathrm{P})$ for *all* closed-valued predictable correspondences C. If we suppose in addition that $\bar{S}_T^0$ and $1/\bar{S}_T^0$ are in $L^\infty(\mathrm{P})$, which holds for instance if r is uniformly bounded in t and ω, then also the corresponding sets $\bar{S}_T^0 G_T(\Theta_S(C))$, $\bar{S}_T^0(x + G_T(\Theta_S(C)))$, and $\bar{S}_T^0 \mathcal{A}(\Theta_S(C))$ of undiscounted payoffs attainable with constrained trading strategies considered in [13], [14], and [19] are closed in $L^2(\mathrm{P})$.

Our assumptions are clearly far less restrictive than completeness of the (unconstrained) financial market. The latter is imposed in [13] and [19] by the conditions that μ, r, and σ are uniformly bounded in t and ω, that σ^{-1} exists and is uniformly bounded in t and ω as well, and that $\mathbb{F}$ is the P-augmentation of the filtration generated by W. The last two assumptions allow us to use Itô's representation theorem and then rewrite integrals of W as integrals of S.

In [13] and [19], the constraints are not formulated in terms of the *number of shares* ϑ^i, but in terms of the *cash amounts* $\pi^i := \vartheta^i \bar{S}^i$ invested *in the different assets*. To see that this can also be handled in our formulation, let C^π be a closed-valued predictable correspondence which describes constraints on the cash amounts. Extending [13] and [19], this need not be deterministic. Since $\bar{S}^i > 0$, we can define the correspondence C^ϑ by

$$C^\vartheta(\omega, t) := \mathrm{diag}(\bar{S}^i(\omega, t))^{-1} C^\pi(\omega, t),$$

which is, by Proposition 2.3 again, a closed-valued predictable correspondence and describes by definition the same constraints as C^π, but in the number of shares. Alternatively, we can consider the dynamics of the gains process parametrised in cash amounts, i.e.

$$\begin{aligned}
\mathrm{d}G_t(\vartheta) &= \vartheta_t^\top \operatorname{diag}(S_t)(\mu_t - r_t \mathbf{1})\,\mathrm{d}t + \vartheta_t^\top \operatorname{diag}(S_t)\sigma_t\,\mathrm{d}W_t \\
&= \pi_t^\top \frac{1}{\bar{S}_t^0}(\mu_t - r_t\mathbf{1})\,\mathrm{d}t + \pi_t^\top \frac{1}{\bar{S}_t^0}\sigma_t\,\mathrm{d}W_t \\
&= \pi_t^\top\,\mathrm{d}X_t, \\
G_0(\vartheta) &= G_0(\pi) = 0,
\end{aligned}$$

with the discounted returns process

$$\mathrm{d}X_t = \frac{1}{\bar{S}_t^0}(\mu_t - r_t\mathbf{1})\,\mathrm{d}t + \frac{1}{\bar{S}_t^0}\sigma_t\,\mathrm{d}W_t, \qquad X_0 = \mathbf{1},$$

as the integrator. Then we can apply our results to the stochastic integrals $\int \pi\,\mathrm{d}X$ with

$$\pi \in \Theta_X(C^\pi) := \left\{ \pi \in \mathcal{L}(X) \;\middle|\; \int \pi\,\mathrm{d}X \in \mathcal{H}^2(\mathrm{P}) \text{ and} \right.$$
$$\left. \pi(\omega, t) \in C^\pi(\omega, t) \text{ for all } (\omega, t) \in \overline{\Omega} \right\}$$

rather than $\int \vartheta\,\mathrm{d}S$ with $\vartheta \in \Theta_S(C^\vartheta)$ to show that $G_T(\Theta_X(C^\pi)) = G_T(\Theta_S(C^\vartheta))$ is closed in $L^2(\mathrm{P})$. In this parametrisation, each (undiscounted) payoff in $\bar{S}_T^0(x + G_T(\Theta_X(C^\pi)))$ is the final value $\bar{V}_T(x, \pi)$ of the wealth process of a self-financing trading strategy, where $\bar{V}(x, \pi)$ is given by the solution of the SDE

$$\mathrm{d}\bar{V}_t(x, \pi) = (\bar{V}_t(x, \pi)r_t + \pi_t^\top(\mu_t - r_t\mathbf{1}))\,\mathrm{d}t + \pi_t^\top\sigma_t\,\mathrm{d}W_t, \qquad \bar{V}_0(x, \pi) = x.$$

For no short-selling constraints, i.e. $C^\pi = [0, +\infty)^d$, Jin and Zhou [14] did not (and did not need to) assume invertibility of $\sigma\sigma^\top$ to obtain a solution to the constrained Markowitz problem. The reason behind this is that as $[0, +\infty)^d$ is a polyhedral set, all its projections are closed. Of course, our results cover this case as well.

To obtain (the existence of) a solution to the mean-variance hedging problem under constraints, we assume in addition to the conditions of Theorem 3.1 that $C \colon \overline{\Omega} \to 2^{\mathbb{R}^d}$ also takes convex values. This gives the following relations to *predictable convexity* and *stability* which come up naturally in dynamic portfolio optimisation problems. The notion of predictable convexity was introduced in [11] to obtain an optional decomposition theorem under constraints, and stability of a set of strategies is usually assumed to establish a dynamic programming principle. The next result and its proof are inspired by Theorems 3 and 4 of [8].

Proposition 3.7. *Assume that $\mathcal{E} = \mathcal{E}(N)$ is regular and satisfies $R_2(\mathrm{P})$, and that $S \in \mathcal{H}_{\mathrm{loc}}^2(\mathrm{P})$ is an $\mathcal{E}$-local martingale. Let $\mathfrak{C} \subseteq \Theta_S$ be nonempty and such that $G_T(\mathfrak{C})$ is closed in $L^2(\mathrm{P})$. Then the following statements are equivalent.*

(a) *The set $\mathfrak{C}$ is predictably convex, i.e. for all ϑ and φ in $\mathfrak{C}$ and all $[0, 1]$-valued predictable processes k, the strategy $k\vartheta + (1 - k)\varphi$ is in $\mathfrak{C}$.*

(b) *The set $\mathfrak{C}$ is convex and stable, i.e. for all ϑ and φ in $\mathfrak{C}$, all $t \in [0, T]$, and all $F \in \mathcal{F}_t$, the strategy $\vartheta\mathbb{I}_{F^c} + (\vartheta\mathbb{I}_{[\![0,t]\!]} + \varphi\mathbb{I}_{]\!]t,T]\!]})\mathbb{I}_F$ is in $\mathfrak{C}$.*

(c) *There exists a predictable correspondence $C: \overline{\Omega} \to 2^{\mathbb{R}^d}$ with nonempty, closed, and convex values such that the projection of C on the predictable range of S is closed, i.e. $\Pi^S(\omega, t)C(\omega, t)$ is closed P_B-a.e., and*

$$G(\mathfrak{C}) = \{G(\varphi) \mid \varphi \in \mathfrak{C}\} = \{G(\vartheta) \mid \vartheta \in \Theta_S(C)\} = G(\Theta_S(C)).$$

Proof. The implication (a)$\Rightarrow$(b) is obvious. For the remaining implications, we observe that, by Proposition 3.4, the closedness of $G_T(\mathfrak{C})$ in $L^2(P)$ is equivalent to that of $G(\mathfrak{C})$ in $\mathcal{H}^2(P)$. The equivalence (c)$\Leftrightarrow$(a) then follows from part 2 of Remark 4.12 of [6], and (b)$\Rightarrow$(a) from (the arguments in the proof of) Lemma 11 of [8].

4. Existence of a solution

Having established the closedness of $G_T(\Theta_S(C))$ and $\mathcal{A}(\Theta_S(C))$, we are able to prove the existence of a solution to the mean-variance hedging problem under trading constraints by using the best approximation theorem in Hilbert spaces; see Theorem 1.4.1 of [2]. Although this looks very easy, it is worth pointing out that our result is given for a very general framework. It covers for instance the existence of a solution in the Itô process setting of Labbé and Heunis [19], Hu and Zhou [13], and Jin and Zhou [14]. We also emphasise that our approach provides a unified treatment for the above papers, which use different and more situation-based arguments like convex duality for Itô processes, Itô's representation theorem, linear-quadratic optimal control, and backward SDE techniques.

Theorem 4.1. *Assume that $\mathcal{E} = \mathcal{E}(N)$ is regular and satisfies $R_2(P)$, and that $S \in \mathcal{H}^2_{\mathrm{loc}}(P)$ is an $\mathcal{E}$-local martingale. Let $C: \overline{\Omega} \to 2^{\mathbb{R}^d}$ be a predictable correspondence with closed convex values such that $\Theta_S(C)$ is nonempty. Then the following statements hold for every $H \in L^2(P)$.*

(a) *There exists a solution $\hat{\vartheta}(x) \in \Theta_S(C)$ to the problem*

$$\mathrm{E}[|x + G_T(\vartheta) - H|^2] = \min_{\vartheta \in \Theta_S(C)} !$$

(b) *There exists a solution $(\hat{x}, \hat{\vartheta}(\hat{x})) \in \mathbb{R} \times \Theta_S(C)$ to the problem*

$$\mathrm{E}[|x + G_T(\vartheta) - H|^2] = \min_{(x,\vartheta) \in \mathbb{R} \times \Theta_S(C)} !$$

Proof. By Theorem 3.1, $G_T(\Theta_S(C))$ and $\mathcal{A}(\Theta_S(C))$ are closed and convex subsets of $L^2(P)$. Therefore, Theorem 1.4.1 of [2] implies the existence of a unique best approximation of $H - x$ by an element in $G_T(\Theta_S(C))$. This can be identified uniquely with an element $G(\hat{\vartheta}(x))$ in $G(\Theta_S(C))$ which gives some $\hat{\vartheta}(x) \in \Theta_S(C)$ and proves (a). In the same way, we obtain a unique element $\hat{v}$ in $\mathcal{A}(\Theta_S(C))$ which is the best approximation to H in $L^2(P)$, and $\hat{v}$ can again be identified with an element $(\hat{x}, \hat{\vartheta}(\hat{x}))$ in $\mathbb{R} \times \Theta_S(C)$.

Remark 4.1. (a) As explained in Example 3.1, the assumptions of Theorem 4.1 are satisfied in the Itô process framework of [13] and [14]. By the argument in the proof of Theorem 11 of [14], obtaining a solution to the *constrained Markowitz problem*, i.e.

$$\begin{aligned} \text{minimise } \mathrm{var}[\bar{V}_T(x, \pi)] &= \mathrm{E}[|\bar{V}_T(x, \pi)|^2] - z^2 \\ \text{subject to } \pi &\in \Theta_X(K^\pi) \text{ and } \mathrm{E}[\bar{V}_T(x, \pi)] = z \end{aligned} \tag{4.1}$$

for $z \geq x\,\mathrm{E}[\bar{S}^0_T]$ and a predictable correspondence K^π with closed and convex *cones* as values,

is equivalent to finding a solution to

$$\mathrm{E}\!\left[\left(\bar{V}_T(x,\pi) - \left(m_1 - m_2\mathcal{E}\!\left(-\int \lambda\,\mathrm{d}M - \int r\,\mathrm{d}t\right)_T\right)\right)^2\right] = \min_{\pi\in\Theta_X(K^\pi)}\,!$$

for a suitable pair $(m_1, m_2) \in \mathbb{R}^2$ of Lagrange multipliers. Therefore, the existence of a solution to (4.1) follows from Theorem 4.1. The dynamic structure of this solution in a general semimartingale framework is established in [7], which generalises the results obtained for a complete Itô process model in Theorem 6.3 of [13].

(b) The problem studied in [19] in an Itô process framework is

$$\mathrm{E}\!\left[\tfrac{1}{2}(\bar{a}|\bar{V}_T(x,\pi)|^2 + \bar{c}V_T(x,\pi))\right] + q = \min_{\pi\in\Theta_X(C^\pi)}\,!,$$

where $\bar{a} > 0$ and $1/\bar{a}$ are in $L^\infty(\mathrm{P})$, $\bar{c} \in L^2(\mathrm{P})$, $q \in \mathbb{R}$, and $C^\pi \equiv K \subseteq \mathbb{R}^d$ is a fixed closed and convex set; see Equation (5.2) of [19]. To obtain a solution to this problem, we observe that $\bar{a}\bar{S}_T^0(x + G_T(\Theta_X(C^\pi)))$ is convex and closed in $L^2(\mathrm{P})$, since $\bar{a}$ and $1/\bar{a}$ are in $L^\infty(\mathrm{P})$, and that we can write

$$\mathrm{E}\!\left[\frac{1}{2}(\bar{a}|\bar{V}_T(x,\pi)|^2 + \bar{c}V_T(x,\pi))\right] + q = \frac{1}{2}\,\mathrm{E}\!\left[\left|\bar{a}\bar{V}_T(x,\pi) + \frac{\bar{c}}{2\bar{a}}\right|^2\right] - \mathrm{E}\!\left[\left|\frac{\bar{c}}{2\bar{a}}\right|^2\right] + q.$$

Then the existence of a solution follows as in the proof of Theorem 4.1 by the best approximation theorem.

In order to also handle constraints on the trajectory of the wealth process, we use a simple martingale argument which already appears in [3], in a complete Itô process model. For that, we define the set of ELMMs for S with P-square-integrable density by

$$\mathbb{P}_e^2(S) = \left\{\mathrm{Q} \sim \mathrm{P} \;\middle|\; S \text{ is a local Q-martingale and } \frac{\mathrm{dQ}}{\mathrm{dP}} \in L^2(\mathrm{P})\right\}.$$

Proposition 4.1. *Suppose that $S \in \mathcal{H}_{\mathrm{loc}}^2(\mathrm{P})$ and that $\mathbb{P}_e^2(S) \neq \varnothing$. Let J be a closed interval in $\mathbb{R}$. Then the following statements hold for any $\vartheta \in \Theta_S$ and any $x \in \mathbb{R}$.*

(a) *$G(\vartheta)$ takes values in J, P-a.s. if and only if its final value $G_T(\vartheta)$ does.*

(b) *The wealth process $V(x,\vartheta)$ takes values in J, P-a.s. if and only if its final value $V_T(x,\vartheta)$ does.*

Proof. Since $V(x,\vartheta) = x + G(\vartheta)$, the proofs for (a) and (b) are completely analogous, and the 'only if' part is obvious. For the 'if' part, choose $\mathrm{Q} \in \mathbb{P}_e^2(S)$ and write $J = [b_1, b_2]$ with $b_1, b_2 \in \mathbb{R}$. Because Q and P are equivalent, we can write a.s. without specifying which measure is meant. Moreover, the density process Z^{Q} is strictly positive and can be represented as a stochastic exponential $Z^Q = \mathcal{E}(L^Q)$ with $L^Q = \int(1/Z_-^Q)\,\mathrm{d}Z^Q$. In the proof of part (a) of Proposition 3.5, the only point where we use the assumption that $R_2(\mathrm{P})$ is satisfied is when ensuring that ${}^\tau\mathcal{E}(N)$ is a P-square-integrable P-martingale for all stopping times τ by Proposition 3.3. However, as this is already known for $Z^{\mathrm{Q}} = \mathcal{E}(L^{\mathrm{Q}})$, we can apply the same arguments here to show that $G(\vartheta)$ is a Q-martingale for all $\vartheta \in \Theta_S$. Hence, $b_1 \leq G_T(\vartheta) \leq b_2$ a.s. implies that $b_1 \leq G_t(\vartheta) \leq b_2$ for all $t \in [0, T]$ a.s. For an infinite interval, the argument is analogous.

Combining the previous result with the L^2-closedness of the space of terminal gains allows us to solve the mean-variance hedging problem also under constraints on the trajectory of the wealth process, again via the best approximation theorem.

Proposition 4.2. *Suppose that $S \in \mathcal{H}^2_{\mathrm{loc}}(\mathrm{P})$ and that there exists $\mathrm{Q} \in \mathbb{P}^2_e(S)$ such that its density process Z^{Q} satisfies $R_2(\mathrm{P})$. Then the following statements hold for every $H \in L^2(\mathrm{P})$ and every closed interval J in $\mathbb{R}$.*

(a) *With $\mathcal{G}_c(\Theta_S) = \{G_T(\vartheta) \in G_T(\Theta_S) \mid G_t(\vartheta) \in J \text{ for all } t \in [0, T], \text{ P-a.s.}\}$, there exists a unique solution $\hat{g}(x) \in \mathcal{G}_c(\Theta_S)$ to the problem*

$$\mathrm{E}[|x + g - H|^2] = \min_{g \in \mathcal{G}_c(\Theta_S)} !$$

(b) *With $\mathcal{A}_c(\Theta_S) = \{V_T(x, \vartheta) \in \mathcal{A}(\Theta_S) \mid V_t(x, \vartheta) \in J \text{ for all } t \in [0, T], \text{ P-a.s.}\}$, there exists a unique solution $\hat{v} \in \mathcal{A}_c(\Theta_S)$ to the problem*

$$\mathrm{E}[|v - H|^2] = \min_{v \in \mathcal{A}_c(\Theta_S)} !$$

Proof. Thanks to Proposition 3.5, $G_T(\Theta_S)$ and $\mathcal{A}(\Theta_S)$ are closed in $L^2(\mathrm{P})$. By Proposition 4.1, we have

$$\{G_T(\vartheta) \in G_T(\Theta_S) \mid G_t(\vartheta) \in J \text{ for all } t \in [0, T], \text{ P-a.s.}\} = \{g \in G_T(\Theta_S) \mid g \in J, \text{ P-a.s.}\},$$

$$\{V_T(x, \vartheta) \in \mathcal{A}(\Theta_S) \mid V_t(x, \vartheta) \in J \text{ for all } t \in [0, T], \text{ P-a.s.}\} = \{a \in \mathcal{A}(\Theta_S) \mid a \in J, \text{ P-a.s.}\}.$$

Moreover, we have $\{g \in G_T(\Theta_S) \mid g \in J, \text{ P-a.s.}\} = G_T(\Theta_S) \cap \{f \in L^2(\mathrm{P}) \mid f \in J, \text{ P-a.s.}\}$ and $\{v \in \mathcal{A}(\Theta_S) \mid v \in J, \text{ P-a.s.}\} = \mathcal{A}(\Theta_S) \cap \{f \in L^2(\mathrm{P}) \mid f \in J, \text{ P-a.s.}\}$. Since $J \subseteq \mathbb{R}$ is closed, the set $\{f \in L^2(\mathrm{P}) \mid f \in J, \text{ P-a.s.}\}$ is closed in $L^2(\mathrm{P})$ and, therefore, so are $\{g \in G_T(\Theta_S) \mid g \in J, \text{ P-a.s.}\}$ and $\{a \in \mathcal{A}(\Theta_S) \mid a \in J, \text{ P-a.s.}\}$. An application of the best approximation theorem completes the proof.

5. Convex duality

While the existence result in Theorem 4.1 is valid in a general framework, its easy proof has the drawback that it only gives the existence of a solution without any further properties. This is one motivation to study mean-variance hedging problems under trading constraints by means of convex duality. Typically, this provides additional insights into the structure of the solution, e.g. that the value functions of the primal and dual problems are continuously differentiable, strictly concave or convex, respectively, and conjugate to each other. Moreover, the solution of the primal problem is linked via the inverse of the 'marginal utility' to the solution of the dual problem.

The general outline of these arguments follows the classical approach of Kramkov and Schachermayer [18] of maximising the expected utility from terminal wealth. The main idea we adopt from there is to treat the problem first as a static optimisation problem. This can be handled easily since we can apply duality theory in a Hilbert space. The obtained duality and existence results are then transferred back to the level of stochastic processes. As for the existence of a solution, the required condition to establish this dual formulation is the closedness of the set $G_T(\Theta_S(C))$. Of course, there is a lot of related work in the literature; we discuss this in more detail in Section 5.3.

To emphasise the analogy with utility maximisation, we rewrite the mean-variance hedging problem as a maximisation problem. This is then our primal problem and consists of finding the optimal trading strategy over time.

Primal Problem. (*Stochastic processes.*)

$$E\left[-\tfrac{1}{2}|x + G_T(\vartheta) - H|^2\right] = \max_{\vartheta \in \Theta_S(C)} ! \tag{5.1}$$

The objective function in (5.1) is $U(x, \omega) = -\tfrac{1}{2}|x - H(\omega)|^2$, which depends on the state ω and is strictly convex and continuously differentiable in x. Its derivative and the inverse of that are $U'(x, \omega) = -x + H(\omega)$ and $I(y, \omega) = (U')^{-1}(y, \omega) = -y + H(\omega)$. Since U fails to be monotonic in x, it is not a utility function in the proper sense. But, as it satisfies the other properties of a utility function and represents our preferences, we call it a *'quadratic utility function'*.

Now observe that (5.1) involves only the terminal wealth $x + G_T(\vartheta)$. Hence, we do not change the optimal value if we regard (5.1) as an optimisation problem over the set of square-integrable random variables defined by

$$\mathcal{C}(x) = \{f \in L^2(P) \mid f = x + G_T(\vartheta) \text{ for some } \vartheta \in \Theta_S(C)\} = x + G_T(\Theta_S(C)).$$

This leads to the corresponding static optimisation problem, which runs only over a set of random variables.

Primal Problem. (*Random variables.*)

$$E[U(f)] = E\left[-\tfrac{1}{2}|f - H|^2\right] = \max_{f \in \mathcal{C}(x)} ! \tag{5.2}$$

By construction, both problems have the same value function

$$u(x) = \sup_{\vartheta \in \Theta_S(C)} E[U(x + G_T(\vartheta))] = \sup_{f \in \mathcal{C}(x)} E[U(f)].$$

5.1. Duality for static variables

If $C \colon \overline{\Omega} \to 2^{\mathbb{R}^d}$ is a predictable correspondence with closed convex values and the assumptions of Theorem 3.1 are satisfied, it follows from there that $G_T(\Theta_S(C))$ is a closed convex subset of $L^2(P)$. Thus, the set of primal variables has the general structure

$$\mathcal{C}(x) = x + \mathcal{G}_c,$$

where

$$\mathcal{G}_c \text{ is a nonempty, closed, and convex subset of } L^2(P). \tag{5.3}$$

Moreover, the assumptions of Theorem 3.1 also imply that with $\mathcal{G} = \text{span}\{\mathcal{G}_c\}$, the set

$$\mathbb{P}_s^2(\mathcal{G}) = \left\{ Q \ll P \,\middle|\, Q \text{ is a signed measure with } Q[\Omega] = 1, \frac{dQ}{dP} \in L^2(P), \text{ and } \frac{dQ}{dP} \in \mathcal{G}^\perp \right\}$$

of signed $\mathcal{G}$-martingale measures is nonempty, where $\mathcal{G}^\perp$ denotes the orthogonal complement of $\mathcal{G}$ in $L^2(P)$. Hence, we also suppose in our abstract static setting that

$$\mathbb{P}_s^2(\mathcal{G}) \neq \varnothing. \tag{5.4}$$

We emphasise that these simple structural properties will be enough to establish the desired duality results in the static setting. To obtain a dual characterisation of the primal variables, we use the following characterisation of closed convex sets $\mathfrak{K}$ in a Hilbert space $\mathfrak{H}$; see Theorem 2.5.1 of [2]. For any $k \in \mathfrak{H}$,

$$k \in \mathfrak{K} \quad \Longleftrightarrow \quad (h, k)_{\mathfrak{H}} \leq \sup_{k' \in \mathfrak{K}} (h, k')_{\mathfrak{H}} =: \delta(h|\mathfrak{K}) \quad \text{for all } h \in \mathfrak{H},$$

where $(\cdot, \cdot)_{\mathfrak{H}}$ denotes the scalar product in $\mathfrak{H}$ and

$$\delta(h|\mathfrak{K}) := \sup_{k' \in \mathfrak{K}} (h, k')_{\mathfrak{H}}$$

is the *support function* of $\mathfrak{K}$. It is easy to see that the support function of a general nonempty set is positively homogeneous, convex, lower semicontinuous, and bounded from below by $-\min_{k \in \mathfrak{K}} \|k\|_{\mathfrak{H}} \|h\|_{\mathfrak{H}}$, which is finite if $\mathfrak{K}$ is nonempty; see Proposition 2.5.1 of [2]. Applying this characterisation to $\mathcal{G}_c$ and $L^2(P)$, we obtain, for any $g \in L^2(P)$,

$$g \in \mathcal{G}_c \quad \Longleftrightarrow \quad E[hg] \leq \sup_{g' \in \mathcal{G}_c} E[hg'] =: \delta(h|\mathcal{G}_c) \quad \text{for all } h \in L^2(P), \qquad (5.5)$$

where $\delta(\cdot|\mathcal{G}_c)$ is the support function of the set $\mathcal{G}_c$.

To deduce dual variables from characterisation (5.5), we observe from [18], [21], and [23] that, for the general outline of the arguments for the dual approach to hold, the dual variables should have the following properties. First, they should be defined in such a way that the dual problem, which is an optimisation problem over the set of dual variables, attains its solution. Since the primal problem is a maximisation for a concave function, the dual problem is a minimisation for a convex function. Thus, it should be enough for the existence of a solution that the set of dual variables is convex and closed. Second, one should be able to establish a duality relation between the sets of primal and dual variables that allows one to show that a natural candidate derived from the dual solution lies in the set of primal variables. This candidate is given by the inverse I of the quadratic 'marginal utility' applied to the dual solution, as follows typically from the first-order condition for optimality in the dual problem. Third, to show that the value functions of the primal and dual problems are conjugate, the product of the parameters x and y of the primal and dual problem should appear in the upper bound for the expectation of a primal and a dual variable for the corresponding parameters.

Let us start with the last point. For a primal variable $f \in \mathcal{C}(x)$ and an element h of $L^2(P)$, the general structure of the primal variables gives

$$E[fh] = E[(x + g)h] \leq x E[h] + \delta(h|\mathcal{G}_c).$$

This motivates us to define the static dual variables by

$$\mathcal{D}(y) = \{h \in L^2(P) \mid E[h] = y, \ \delta(h|\mathcal{G}_c) < \infty\} \quad \text{for } y \in \mathbb{R},$$

because this gives the third of the above properties, i.e.

$$f \in \mathcal{C}(x), \ h \in \mathcal{D}(y) \quad \Longrightarrow \quad E[fh] \leq xy + \delta(h|\mathcal{G}_c). \qquad (5.6)$$

By continuity and linearity of the expectation and lower semicontinuity and convexity of the support function, the set $\mathcal{D}(y)$ is closed and convex in $L^2(P)$ and, thus, also likely to satisfy the first property listed above. Note that $\mathcal{D}(y)$ contains $y\, dQ/dP$ for any $Q \in \mathbb{P}_s^2(\mathcal{G})$ so that it is nonempty due to (5.4). The second property from above will follow via the dual characterisation (5.5) of convex closed sets in $L^2(P)$; see the proof of Theorem 5.1 below.

Remark 5.1. (a) For a linear subspace $\mathcal{G}_c = \mathcal{G}$, characterisation (5.5) simplifies to orthogonality, and the dual domain becomes $\mathcal{D}(y) = \{h \in L^2(\mathrm{P}) \mid \mathrm{E}[h] = y, \ h \in \mathcal{G}_c^\perp\}$. Moreover, we have $\mathcal{D}(y) = y\mathbb{P}_s^2(\mathcal{G})$ for any $y \neq 0$. This is exploited in [12] for the dual formulation in the unconstrained case.

(b) If $\mathcal{G}_c$ is a *cone*, the support function $\delta(\cdot|\mathcal{G}_c)$ only takes the values 0 and ∞ and (5.5) therefore reduces to the *bipolar relation*

$$g \in \mathcal{G}_c \quad \Longleftrightarrow \quad \mathrm{E}[hg] \leq 0 \quad \text{for all } h \in \mathcal{G}_c^\circ,$$

where $\mathcal{G}_c^\circ = \{h \in L^2(\mathrm{P}) \mid \mathrm{E}[hg] \leq 0 \text{ for all } g \in \mathcal{G}_c\}$ is the polar of $\mathcal{G}_c$. Since $\mathcal{G}_c^\circ$ is again a cone, we have $\mathcal{D}(y) = y\mathcal{D}(1)$ for $y > 0$ and $\mathcal{D}(y) = |y|\mathcal{D}(-1)$ for $y < 0$. The sets $\mathcal{D}(1)$ and $\mathcal{D}(-1)$ can then be interpreted as the sets of all Radon–Nikodým derivatives of signed $\mathcal{G}_c$-supermartingale and $\mathcal{G}_c$-submartingale measures, respectively. The above simplification explains why the majority of papers concentrate on constraints given by closed convex *cones*.

Returning to the general case, we work as usual with the Legendre transform V in x of $-U(-\cdot, \omega)$ to derive the formulation of the dual problem. The function V is given by

$$V(y, \omega) = \sup_{x \in \mathbb{R}}\{U(x, \omega) - xy\} = U(I(y, \omega)) - I(y, \omega)y = \tfrac{1}{2}y^2 - yH(\omega);$$

it depends on the state ω and is continuously differentiable and strictly convex in y. The motivation for using the Legendre transform comes from looking for the sharpest inequality such that

$$U(x, \omega) \leq V(y, \omega) + xy \quad \text{for all } x, y \in \mathbb{R} \text{ and all } \omega \in \Omega.$$

Plugging in $f \in \mathcal{C}(x)$ and $h \in \mathcal{D}(y)$ for x and y in the above inequality, taking expectations, and optimising on both sides gives, via (5.6),

$$
\begin{aligned}
u(x) = \ & \sup_{f \in \mathcal{C}(x)} \mathrm{E}[U(f)] \\
& \leq \inf_{y \in \mathbb{R}}\left\{ \inf_{h \in \mathcal{D}(y)} \mathrm{E}[V(h) + gh] \right\} \\
& \leq \inf_{y \in \mathbb{R}}\left\{ \inf_{h \in \mathcal{D}(y)} \{\mathrm{E}[V(h)] + \delta(h|\mathcal{G}_c)\} + xy \right\} \\
& = \inf_{y \in \mathbb{R}}\{v(y) + xy\},
\end{aligned}
\tag{5.7}
$$

where the value function of the dual problem on the level of random variables is

$$v(y) = \inf_{h \in \mathcal{D}(y)}\{\mathrm{E}[V(h)] + \delta(h|\mathcal{G}_c)\}.$$

Note that the objective function of the dual problem explicitly involves the constraints via the support function δ.

Dual Problem. (*Random variables.*)

$$\Psi(h) := \mathrm{E}[V(h)] + \delta(h|\mathcal{G}_c) = \min_{h \in \mathcal{D}(y)} ! \tag{5.8}$$

From the inequalities in (5.7), we see that if we can find for a given x a pair $(\hat{f}(x), \hat{h}(y))$ of primal and dual variables such that the equality holds in (5.7), we have also found a solution to the primal problem (5.2), as $\hat{f}(x)$ attains the supremum. Of course, y will then also depend

on x. So an abstract recipe for solving the primal problem is as follows.

1. Find the solution $\hat{h}(y)$ to the dual problem (5.8) for any $y \in \mathbb{R}$.

2. Find the minimiser $\hat{y}(x)$ for the indirect dual problem $v(y) + xy = \min_{y \in \mathbb{R}}!$ for any $x \in \mathbb{R}$.

3. Define $\hat{h} := \hat{h}(\hat{y}(x))$ and $\hat{f}(x) := I(\hat{h})$, and show that $\mathrm{E}[I(\hat{h})\hat{h}] = x\hat{y}(x) + \delta(\hat{h}|\mathcal{G}_c)$.

4. If we can show that $\hat{f}(x) \in \mathcal{C}(x)$ then $\hat{f}(x)$ solves (5.2), since we have, by combining (5.7) with steps 1–3,

$$
\begin{aligned}
u(x) &\leq \inf_{y \in \mathbb{R}} \{v(y) + xy\} \\
&= v(\hat{y}(x)) + x\hat{y}(x) \\
&= \mathrm{E}[V(\hat{h})] + \delta(\hat{h}|\mathcal{G}_c) + x\hat{y}(x) \\
&= \mathrm{E}[U(I(\hat{h})) - I(\hat{h})\hat{h}] + \delta(\hat{h}|\mathcal{G}_c) + x\hat{y}(x) \\
&= \mathrm{E}[U(\hat{f}(x))] \\
&\leq u(x).
\end{aligned}
$$

To solve the primal problem (5.2), it now remains to implement the above recipe. We start by solving the dual problem, making use of the following result from convex analysis; see Proposition 1.2 of [10].

Proposition 5.1. *Let B be a reflexive Banach space, let K be a nonempty closed convex subset of B, and let F be a strictly convex, coercive, and lower-semicontinuous function from B into $\mathbb{R} \cup \{+\infty\}$ that is proper on K. Then there exists a unique solution $\hat{b} \in K$ to*

$$
F(b) = \min_{b \in K}!
$$

Taking $B = L^2(\mathrm{P})$ and $K = \mathcal{D}(y)$, we only need to check that the dual objective function Ψ satisfies the properties of F to apply the proposition.

Lemma 5.1. *For every $H \in L^2(\mathrm{P})$, the mapping*

$$
h \mapsto \Psi(h) = \mathrm{E}[V(h)] + \delta(h|\mathcal{G}_c) = \mathrm{E}\left[\tfrac{1}{2}h^2 - hH\right] + \delta(h|\mathcal{G}_c)
$$

from $L^2(\mathrm{P})$ into $\mathbb{R} \cup \{+\infty\}$ is strictly convex, lower semicontinuous, coercive, and uniformly bounded from below by $-\frac{1}{2}(\|H\|_{L^2(\mathrm{P})} + \min_{g \in \mathcal{G}_c} \|g\|_{L^2(\mathrm{P})})^2$.

Proof. We begin by proving that the mapping $h \mapsto \mathrm{E}[V(h)]$ from $L^2(\mathrm{P})$ into $\mathbb{R}$ is strictly convex and continuous. The first property follows immediately from the strict convexity of the function $V(\cdot, \omega)$ for all $\omega \in \Omega$. If (h_n) converges to h in $L^2(\mathrm{P})$ then (h_n) is bounded in $L^2(\mathrm{P})$ and the Cauchy–Schwarz inequality gives

$$
\begin{aligned}
&\left| \mathrm{E}\left[\tfrac{1}{2}h_n^2 - h_n H\right] - \mathrm{E}\left[\tfrac{1}{2}h^2 - hH\right] \right| \\
&\quad = \left| \mathrm{E}\left[(h_n - h)\left(\tfrac{1}{2}(h_n + h) - H\right)\right] \right| \\
&\quad \leq \|h_n - h\|_{L^2(\mathrm{P})} \left(\tfrac{1}{2}\left(\sup_{n \in \mathbb{N}} \|h_n\|_{L^2(\mathrm{P})} + \|h\|_{L^2(\mathrm{P})} \right) + \|H\|_{L^2(\mathrm{P})} \right) \\
&\quad \to 0 \quad \text{as } n \to \infty,
\end{aligned}
$$

which proves the claimed continuity. Since $\delta(\cdot|\mathcal{G}_c)$ is convex and lower semicontinuous, the sum $\Psi(\cdot) = \mathrm{E}[V(\cdot)] + \delta(\cdot|\mathcal{G}_c)$ is strictly convex and lower semicontinuous. Moreover, the Cauchy–Schwarz inequality implies that

$$\Psi(h) = \mathrm{E}\left[\tfrac{1}{2}h^2 - hH\right] + \delta(h|\mathcal{G}_c)$$
$$\geq \tfrac{1}{2}\|h\|_{L^2(\mathrm{P})}^2 - \|h\|_{L^2(\mathrm{P})}\left(\|H\|_{L^2(\mathrm{P})} + \min_{g \in \mathcal{G}_c}\|g\|_{L^2(\mathrm{P})}\right), \tag{5.9}$$

which gives coercivity, since the right-hand side tends to ∞ as $\|h\|_{L^2(\mathrm{P})} \to \infty$ because $\min_{g \in \mathcal{G}_c}\|g\|_{L^2(\mathrm{P})}$ is finite. Minimising the right-hand side over $\|h\|_{L^2(\mathrm{P})}$ also gives the asserted lower bound, which completes the proof.

From the definition of $\mathcal{D}(y)$, we know that $y\,\mathrm{d}Q/\mathrm{d}P$ is in $\mathcal{D}(y)$ for every $y \in \mathbb{R}$ and $Q \in \mathbb{P}_s^2(\mathcal{G})$. Since $\delta(y\,\mathrm{d}Q/\mathrm{d}P|\mathcal{G}_c) = 0$, assumption (5.4) implies that Ψ is proper on $\mathcal{D}(y)$ for each $y \in \mathbb{R}$. Therefore, all the conditions of Proposition 5.1 are satisfied in the setting of the dual problem, and the existence of a solution to the dual problem follows by combining Lemma 5.1 with Proposition 5.1. This gives the following result.

Proposition 5.2. *Under assumption (5.3) and assumption (5.4), there exists a unique solution $\hat{h}(y) \in \mathcal{D}(y)$ to the dual problem (5.8) for every $H \in L^2(\mathrm{P})$ and each $y \in \mathbb{R}$, i.e.*

$$\Psi(\hat{h}(y)) = \mathrm{E}[V(\hat{h}(y))] + \delta(\hat{h}(y)|\mathcal{G}_c) = \inf_{h \in \mathcal{D}(y)} \mathrm{E}[V(h)] + \delta(h|\mathcal{G}_c) = v(y).$$

By Proposition 5.2, the function v inherits all nice properties of Ψ, which enables us to solve the indirect dual problem by again using Proposition 5.1.

Lemma 5.2. *Under assumptions (5.3) and (5.4), the function v is strictly convex, continuous, and coercive.*

Proof. If $y_1, y_2 \in \mathbb{R}$ and $\mu \in (0, 1)$, then $\mu\hat{h}(y_1) + (1 - \mu)\hat{h}(y_2) \in \mathcal{D}(\mu y_1 + (1 - \mu)y_2)$; so

$$\mu v(y_1) + (1 - \mu)v(y_2) = \mu\Psi(\hat{h}(y_1)) + (1 - \mu)\Psi(\hat{h}(y_2))$$
$$> \Psi(\mu\hat{h}(y_1) + (1 - \mu)\hat{h}(y_2))$$
$$\geq v(\mu y_1 + (1 - \mu)y_2)$$

by Proposition 5.2 and the strict convexity of Ψ. Hence, v is strictly convex, and continuous like any convex function on $\mathbb{R}$ with finite values; see Corollary II.10.1.1 of [25]. By Jensen's inequality, $\|\hat{h}(y)\|_{L^2(\mathrm{P})} \geq \mathrm{E}[\hat{h}(y)] = y$ tends to ∞ as $y \to \infty$. Thus, coercivity of Ψ implies coercivity of v, again by Proposition 5.2. Note that in view of (5.9), $v(y)$ even grows quadratically as $|y| \to \infty$.

Since a continuous function is obviously proper, applying Proposition 5.1 to the strictly convex, continuous, and coercive mapping $y \mapsto v(y) + xy$ on $\mathbb{R}$ immediately gives the following result.

Corollary 5.1. *Assume that (5.3) and (5.4) hold. For every $x \in \mathbb{R}$, there exists a unique $\hat{y}(x) \in \mathbb{R}$ that solves*

$$v(y) + xy = \min_{y \in \mathbb{R}}!$$

Now we have everything in place to formulate and prove the abstract static version of the main result of this section.

Theorem 5.1. *Suppose as in (5.3) that $\mathcal{G}_c$ is a nonempty, convex, and closed subset of $L^2(\mathrm{P})$, and impose assumption (5.4) that $\mathbb{P}^2_s(\mathcal{G}) \neq \varnothing$. Then the following statements hold.*

(a) *For every $x \in \mathbb{R}$, there exists a unique solution $\hat{f}(x) \in \mathcal{C}(x)$ to*

$$\mathrm{E}\left[-\tfrac{1}{2}|f - H|^2\right] = \max_{f \in \mathcal{C}(x)} !$$

It is given by

$$\hat{f}(x) = I(\hat{h}(\hat{y}(x))) = -\hat{h}(\hat{y}(x)) + H,$$

where $\hat{h}(\hat{y}(x)) \in \mathcal{D}(\hat{y}(x))$ and $\hat{y}(x) \in \mathbb{R}$ are respectively the unique solutions to

$$\Psi(h) = \mathrm{E}\left[\tfrac{1}{2}h^2 - hH\right] + \delta(h|\mathcal{G}_c) = \min_{h \in \mathcal{D}(\hat{y}(x))} !$$

and

$$v(y) + xy = \min_{y \in \mathbb{R}}! \tag{5.10}$$

(b) *The value functions u and v are conjugate, i.e.*

$$u(x) = \inf_{y \in \mathbb{R}}\{v(y) + xy\}, \qquad v(y) = \sup_{x \in \mathbb{R}}\{u(x) - xy\},$$

and continuously differentiable. Furthermore, u is strictly concave and v is strictly convex.

(c) *We have the equivalent relations*

$$\mathrm{E}[\hat{f}(x)\hat{h}(\hat{y}(x))] = x\hat{y}(x) + \delta(\hat{h}(\hat{y}(x))|\mathcal{G}_c), \tag{5.11a}$$

$$\mathrm{E}[\hat{f}(x)U'(\hat{f}(x))] = xu'(x) + \delta(U'(\hat{f}(x))|\mathcal{G}_c), \tag{5.11b}$$

$$\mathrm{E}[\hat{h}(\hat{y}(x))V'(\hat{h}(\hat{y}(x)))] = \hat{y}(x)v'(\hat{y}(x)) - \delta(\hat{h}(\hat{y}(x))|\mathcal{G}_c). \tag{5.11c}$$

Proof. (a) Since $\hat{y}(x)$ and $\hat{h} := \hat{h}(\hat{y}(x))$ solve problems (5.10) and (5.8), the definition of $\mathcal{D}(y)$ implies that $\hat{h}(\hat{y}(x))$ is also the solution to

$$\Psi(h) + x\,\mathrm{E}[h] = \min_{h \in L^2(\mathrm{P})} ! = \min_{h \in \mathcal{D}(\hat{y}(x))} ! \tag{5.12}$$

For $\varepsilon \in (0, 1)$ and $h \in L^2(\mathrm{P})$, set $h_\varepsilon = \hat{h} + \varepsilon h$. Then optimality of $\hat{h}$ for (5.12) gives

$$0 \leq \liminf_{\varepsilon \searrow 0} \frac{\Psi(h_\varepsilon) + x\,\mathrm{E}[h_\varepsilon] - (\Psi(\hat{h}) + x\,\mathrm{E}[\hat{h}])}{\varepsilon}$$

$$= \liminf_{\varepsilon \searrow 0}\left\{\mathrm{E}[(\hat{h} - H + x)h] + \frac{1}{2}\varepsilon\,\mathrm{E}[h^2] + \frac{\delta(\hat{h} + \varepsilon h|\mathcal{G}_c) - \delta(\hat{h}|\mathcal{G}_c)}{\varepsilon}\right\}, \tag{5.13}$$

where the last expression is well defined as $\delta(\hat{h}|\mathcal{G}_c)$ is finite. Hence, we obtain, by using $I(\hat{h}) = -\hat{h} + H$ and the sublinearity of $\delta(\cdot|\mathcal{G}_c)$,

$$\mathrm{E}[(I(\hat{h}) - x)h] \leq \delta(h|\mathcal{G}_c) \quad \text{for all } h \in L^2(\mathrm{P}), \tag{5.14}$$

and, thus, $I(\hat{h}) - x \in \mathcal{G}_c$, i.e. $I(\hat{h}) \in \mathcal{C}(x)$, by the characterisation of closed convex sets in (5.5). Substituting $h = -\hat{h}$ into (5.13) and using the positive homogeneity of $\delta(\cdot|\mathcal{G}_c)$ gives

$$\mathrm{E}[(I(\hat{h}) - x)(-\hat{h})] \leq -\delta(\hat{h}|\mathcal{G}_c).$$

Combining this with (5.14) for $h = \hat{h}$ gives

$$\delta(\hat{h}|\mathcal{G}_c) = \mathrm{E}[(I(\hat{h}) - x)\hat{h}] = \mathrm{E}[I(\hat{h})\hat{h}] - x\hat{y}(x). \tag{5.15}$$

Hence, we obtain from (5.15), as in step 4 of the abstract recipe,

$$
\begin{aligned}
u(x) &\geq \mathrm{E}[U(I(\hat{h}))] \\
&= \mathrm{E}[V(\hat{h}) + I(\hat{h})\hat{h}] \\
&= \mathrm{E}[V(\hat{h})] + \delta(\hat{h}|\mathcal{G}_c) + x\hat{y}(x) \\
&= v(\hat{y}(x)) + x\hat{y}(x) \\
&\geq u(x),
\end{aligned}
\tag{5.16}
$$

which shows that $\hat{f}(x) := I(\hat{h}(\hat{y}(x)))$ indeed maximises $\mathrm{E}[U(f)]$ over $\mathcal{C}(x)$.

(b) Since we have equality in (5.16) and $\hat{y}(x)$ attains $\inf_{y \in \mathbb{R}}\{v(y) + xy\}$, we also have $u(x) = \inf_{y \in \mathbb{R}}\{v(y) + xy\}$ for all $x \in \mathbb{R}$ and then $v(y) = \sup_{x \in \mathbb{R}}\{u(x) - xy\}$ by the biconjugate property of the Legendre transform; see Theorem III.12.2 of [25]. To show the strict concavity of u, we fix $x_1, x_2 \in \mathbb{R}$ and $\mu \in (0, 1)$. Then $\mu\hat{f}(x_1) + (1 - \mu)\hat{f}(x_2)$ is in $\mathcal{C}(\mu x_1 + (1 - \mu)x_2)$ and so part (a) yields, by the strict concavity of $U(\cdot, \omega)$,

$$
\begin{aligned}
\mu u(x_1) + (1 - \mu)u(x_2) &= \mathrm{E}[\mu U(\hat{f}(x_1)) + (1 - \mu)U(\hat{f}(x_2))] \\
&< \mathrm{E}[U(\mu\hat{f}(x_1) + (1 - \mu)\hat{f}(x_2))] \\
&\leq u(\mu x_1 + (1 - \mu)x_2).
\end{aligned}
$$

Continuous differentiability of u and v follows since the Legendre transform of a strictly convex function is differentiable; see Theorems V.24.1 and V.26.3 of [25]. Since v is continuously differentiable, we obtain for the minimiser $\hat{y}(x)$ of $v(y) + xy$ over $y \in \mathbb{R}$ the relation $v'(\hat{y}(x)) = -x$. Again by general results on the Legendre transform, we have

$$V'(\cdot, \omega) = -(U')^{-1}(\cdot, \omega) = -I(\cdot, \omega)$$

and $v' = -(u')^{-1}$; see Theorem V.23.5 of [25]. Combining this with $v'(\hat{y}(x)) = -x$, $\hat{f}(x) = I(\hat{h})$, and (5.15) gives the relations in (5.11). This completes the proof. ∎

5.2. Duality for dynamic variables

Under the assumptions of Theorem 3.1, Theorem 5.1 already implies the existence of a unique solution to the primal problem (5.1) by choosing $\mathcal{G}_c = G_T(\Theta_S(C))$, i.e. there exists an optimal trading strategy $\hat{\vartheta}(x) \in \Theta_S(C)$ such that $\hat{f}(x) = x + G_T(\hat{\vartheta}(x))$. In particular, we recover Theorem 4.1(b).

To establish an analogous duality result on the level of stochastic processes, we need a dynamic version for the dual variables. If we assume for simplicity that

$$\mathcal{F} = \mathcal{F}_T,$$

we can identify every $h \in L^2(\mathrm{P})$ with a square-integrable RCLL martingale $Z = Z(h)$ given by $Z_t = \mathrm{E}[h \mid \mathcal{F}_t]$ for $t \in [0, T]$. The Kunita–Watanabe decomposition then yields

$$Z_t = \mathrm{E}[h \mid \mathcal{F}_0] + \int_0^t \eta_s\, \mathrm{d}M_s + R_t, \qquad t \in [0, T],$$

with $\eta \in L^2(M)$ and $R \in \mathcal{M}_0^2(\mathrm{P})$ strongly P-orthogonal to M. We choose this parametrisation because it makes it easy to calculate the dynamics of the product of a gains process and a dual variable. Moreover, it is similar to [5], where dual variables are supermartingale measures for the gains processes of constrained trading strategies. The parametrisation in [5] can be obtained by applying the Kunita–Watanabe decomposition to the stochastic logarithm of the density process, and in the Brownian filtration of [5], this decomposition can of course be replaced by Itô's representation theorem; see [5] and Example 3.2 of [21].

Lemma 5.3. *Suppose that S is in $\mathcal{H}_{\mathrm{loc}}^2(\mathrm{P})$ and satisfies the structure condition. For every $\vartheta \in \Theta_S$ and every $Z \in \mathcal{M}^2(\mathrm{P})$, the process*

$$\left(G_t(\vartheta) Z_t - \int_0^t (\eta_s + Z_{s-}\lambda_s)^\top c_s^M \vartheta_s \, dB_s \right)_{0 \le t \le T} \tag{5.17}$$

is then a P-martingale with P-integrable supremum, i.e. a martingale in $\mathcal{H}^1(\mathrm{P})$.

Proof. Applying the product rule and using the fact that S satisfies the structure condition gives

$$d(G(\vartheta)Z) = Z_- \vartheta \, dM + Z_- \vartheta^\top d\langle M \rangle \lambda + G_-(\vartheta) \, dZ + d\left[Z, \int \vartheta \, dA \right] + d\left[Z, \int \vartheta \, dM \right].$$

Clearly, $\int Z_- \vartheta \, dM$ and $\int G_-(\vartheta) \, dZ$ are local P-martingales. Moreover, $\langle Z, \int \vartheta \, dM \rangle$ exists because $Z \in \mathcal{M}^2(\mathrm{P})$ and $\vartheta \in \mathcal{L}^2(M)$, and $[Z, \int \vartheta \, dA]$ is a local martingale by Yoeurp's lemma. Writing '$\overset{\mathrm{mart}}{=}$' for equality up to a local P-martingale and using the fact that R is strongly P-orthogonal to M, we thus obtain

$$d(G(\vartheta)Z) \overset{\mathrm{mart}}{=} Z_- \vartheta^\top d\langle M \rangle \lambda + \vartheta^\top d\langle Z, M \rangle \overset{\mathrm{mart}}{=} (\eta + Z_-\lambda)^\top d\langle M \rangle \vartheta.$$

This shows that the process in (5.17) is a local P-martingale. To check integrability, we first observe that, by Doob's inequality and Proposition 2.1, $(ZG(\vartheta))_T^*$ is in $L^1(\mathrm{P})$ since $Z \in \mathcal{M}^2(\mathrm{P})$ and $G(\vartheta) \in \mathcal{H}^2(\mathrm{P})$, and that the Kunita–Watanabe inequality gives

$$\mathrm{E}\left[\left(\int \eta^\top d\langle M \rangle \vartheta \right)_T^* \right] \le \mathrm{E}\left[\int_0^T |\eta_s^\top c_s^M \vartheta_s| \, dB_s \right] \le \|\eta\|_{\mathcal{L}^2(M)} \|\vartheta\|_{\mathcal{L}^2(M)} < \infty.$$

Moreover, using $\int |\vartheta \, dA| = \int |\vartheta^\top c^M \lambda| \, dB$ and the Cauchy–Schwarz and Doob inequalities allows us to estimate the remaining term by

$$\mathrm{E}\left[\left(\int Z_- \vartheta^\top d\langle M \rangle \lambda \right)_T^* \right] \le \mathrm{E}\left[Z_T^* \int_0^T |\vartheta_s^\top c_s^M \lambda_s| \, dB_s \right] \le 2\|Z_T\|_{L^2(\mathrm{P})} \|\vartheta\|_{\mathcal{L}^2(A)}.$$

Replacing $c_t^M \, dB_t$ by $d\langle M \rangle_t$ and using the estimate

$$\left(G(\vartheta)Z - \int (\eta + Z_-\lambda)^\top c^M \vartheta \, dB \right)_T^*$$
$$\le (ZG(\vartheta))_T^* + \left(\int \eta^\top d\langle M \rangle \vartheta \right)_T^* + \left(\int Z_- \vartheta^\top d\langle M \rangle \lambda \right)_T^*$$

then shows that the local P-martingale in (5.17) has a P-integrable supremum.

Using Lemma 5.3 and optimising over ϑ immediately gives, for every $Z \in \mathcal{M}^2(P)$,

$$
\sup_{\vartheta \in \Theta_S(C)} \mathrm{E}[G_T(\vartheta) Z_T] = \sup_{\vartheta \in \Theta_S(C)} \mathrm{E}\left[\int_0^T (\eta_s + Z_{s-}\lambda_s)^\top c_s^M \vartheta_s \, dB_s \right]
$$

$$
\leq \mathrm{E}\left[\int_0^T \delta(c_s^M(\eta_s + Z_s\lambda_s)|C) \, dB_s \right] \tag{5.18}
$$

by the definition of the support function $\delta(\cdot|C)$, because each ϑ_s has values in C. The next result shows that we even have equality in (5.18). Note that we use the same symbol δ for support functions in two different Hilbert spaces—$L^2(P)$ on the left-hand side and $\mathbb{R}^d$ on the right-hand side of (5.19).

Lemma 5.4. *Suppose that S is in $\mathcal{H}^2_{\mathrm{loc}}(P)$ and satisfies the structure condition. For every $Z \in \mathcal{M}^2(P)$,*

$$
\delta(Z_T|G_T(\Theta_S(C))) = \sup_{\vartheta \in \Theta_S(C)} \mathrm{E}[G_T(\vartheta) Z_T] = \mathrm{E}\left[\int_0^T \delta(c_s^M(\eta_s + Z_{s-}\lambda_s)|C) \, dB_s \right]. \tag{5.19}
$$

Proof. Without loss of generality, we can assume that $0 \in C(\omega, t)$. Indeed, let φ be in $\Theta_S(C)$ and set $C' = C - \varphi$, which is, by Proposition 2.2, a predictable correspondence with $0 \in C'(\omega, t)$. Then $\Theta_S(C) = \Theta_S(C') + \varphi$ and, therefore,

$$
\delta(Z_T|G_T(\Theta_S(C))) = \delta(Z_T|G_T(\Theta_S(C'))) + \mathrm{E}[Z_T G_T(\varphi)]
$$

and

$$
\mathrm{E}\left[\int_0^T \delta(c_s^M(\eta_s + Z_{s-}\lambda_s)|C) \, dB_s \right]
$$

$$
= \mathrm{E}\left[\int_0^T \delta(c_s^M(\eta_s + Z_{s-}\lambda_s)|C') \, dB_s + \int_0^T \varphi_s^\top c_s^M(\eta_s + Z_{s-}\lambda_s) \, dB_s \right].
$$

Since $\mathrm{E}[Z_T G_T(\varphi)] = \mathrm{E}[\int_0^T \varphi_s^\top c_s^M(\eta_s + Z_{s-}\lambda_s) \, dB_s]$ by Lemma 5.3, (5.19) holds for C if and only if it holds for C'.

In view of (5.18), it remains to show that

$$
\sup_{\vartheta \in \Theta_S(C)} \mathrm{E}[G_T(\vartheta) Z_T] \geq \mathrm{E}\left[\int_0^T \delta(c_s^M(\eta_s + Z_{s-}\lambda_s)|C) \, dB_s \right].
$$

To this end, we construct a sequence (ϑ^n) of C-constrained trading strategies such that

$$
\lim_{n \to \infty} \mathrm{E}[G_T(\vartheta^n) Z_T] = \mathrm{E}\left[\int_0^T \delta(c_s^M(\eta_s + Z_{s-}\lambda_s)|C) \, dB_s \right].
$$

Define a function $f: \overline{\Omega} \times \mathbb{R}^d \to \mathbb{R}$ by

$$
f((\omega, t), z) := (\eta(\omega, t) + Z(\omega, t-)\lambda(\omega, t))^\top c^M(\omega, t) y
$$

and, for each $n \in \mathbb{N}$, a predictable correspondence C^n by

$$
C^n(\omega, t) := C(\omega, t) \cap \overline{B_n(0)} \subseteq \mathbb{R}^d
$$

for $(\omega, t) \in \overline{\Omega}$, where $\overline{B_n(0)}$ denotes the closure of the ball of radius n in $\mathbb{R}^d$. Note that the C^n have convex and compact values and that $0 \in C^n(\omega, t)$ for $(\omega, t) \in \overline{\Omega}$ and each $n \in \mathbb{N}$.

Moreover, $f(\cdot, z)$ is predictable for $z \in \mathbb{R}^d$ and $f((\omega, t), \cdot)$ is continuous for $(\omega, t) \in \overline{\Omega}$, i.e. f is a Carathéodory function. Let $\{x^{n,m} \mid m \in \mathbb{N}\}$ be a Castaing representation for C^n as in Proposition 2.2, and define

$$g^n(\omega, t) := \delta(c^M(\omega, t)(\eta(\omega, t) + Z(\omega, t-)\lambda(\omega, t))|C^n(\omega, t)) = \sup_{x \in C^n(\omega, t)} f((\omega, t), x).$$

As $g^n(\omega, t) = \sup_{m \in \mathbb{N}} f((\omega, t), x^{n,m}(\omega, t))$ by Proposition 2.2, g^n is predictable and finite-valued by the compactness of $C^n(\omega, t)$. Combining Propositions 2.3 and 2.4 tells us that

$$D^n(\omega, t) = \{z \in C^n(\omega, t) \mid f((\omega, t), z) = g^n(\omega, t)\}$$

is a predictable correspondence with nonempty, convex, and compact values. Let y^n be a predictable selector of D^n. Then $g^n(\omega, t) = f((\omega, t), y^n(\omega, t))$ and

$$\delta(c^M(\omega, t)(\eta(\omega, t) + Z(\omega, t-)\lambda(\omega, t))|C(\omega, t)) = \lim_{n \to \infty} f((\omega, t), y^n(\omega, t)),$$

where the limit is increasing since $C(\omega, t) = \bigcup_{n \in \mathbb{N}} C^n(\omega, t)$. Let $(\tau_m)_{m \in \mathbb{N}}$ be a localising sequence such that S^{τ_m} is in $\mathcal{H}^2(\mathrm{P})$. Since $|y^n(\omega, t)| \leq n$ and each $C(\omega, t)$ contains 0, the process $\vartheta^n := y^n \mathbb{I}_{[\![0, \tau_n]\!]}$ is in $\Theta_S(C)$ for each $n \in \mathbb{N}$. Hence, Lemma 5.3 and monotone integration yield

$$\lim_{n \to \infty} \mathrm{E}[G_T(\vartheta^n) Z_T] = \lim_{n \to \infty} \mathrm{E}\left[\int_0^{\tau_n} (y_s^n)^\top c_s^M (\eta_s + Z_{s-}\lambda_s) \, \mathrm{d}B_s\right]$$
$$= \mathrm{E}\left[\int_0^T \delta(c_s^M (\eta_s + Z_{s-}\lambda_s)|C) \, \mathrm{d}B_s\right],$$

which completes the proof.

As Lemma 5.4 relates the support function $\delta(\cdot|G_T(\Theta_S(C)))$ to the expectation of the terminal value of a stochastic process, we are led to reformulate the dual problem (5.8) on the level of stochastic processes in the following way.

Dual Problem. (*Stochastic processes.*)

$$\Psi(Z_T) = \mathrm{E}\left[\frac{1}{2} Z_T^2 - Z_T H + \int_0^T \delta(c_s^M (\eta_s + Z_{s-}\lambda_s)|C) \, \mathrm{d}B_s\right] = \min_{Z \in \mathcal{Y}(y)} !, \qquad (5.20)$$

where

$$\mathcal{Y}(y) = \left\{ Z \in \mathcal{M}^2(\mathrm{P}) \,\middle|\, Z = Z_0 + \int \eta \, \mathrm{d}M + R \text{ with } Z_0 \in L^2(\mathrm{P}, \mathcal{F}_0), \, \eta \in L^2(M), \right.$$
$$\left. R \in \mathcal{M}_0^2(\mathrm{P}) \text{ strongly P-orthogonal to } M, \text{ and } \mathrm{E}[Z_T] = \mathrm{E}[Z_0] = y \right\}.$$

Remark 5.2. In the Itô process framework of Example 3.1 and if $\mathbb{F}$ is the P-augmentation of the filtration generated by W, the above dual problem (5.20) for stochastic processes specialises to the dual problem considered in [19, Equation 5.37]; this only requires some adjustments to the notation, along the lines of Remark 4.1(b).

Now set $\mathcal{G}_c := G_T(\Theta_S(C))$ so that Lemma 5.4 gives an explicit representation of the support function $\delta(\cdot|\mathcal{G}_c)$. Then the functions Ψ in (5.8) and (5.20) coincide, and by identifying each $h \in L^2(\mathrm{P})$ with the corresponding square-integrable martingale, it also follows that (5.8) and (5.20) have the same optimal value and, therefore, the same value function

$$
\begin{aligned}
v(y) &= \inf_{Z \in \mathcal{Y}(y)} \mathrm{E}\left[\frac{1}{2}Z_T^2 - Z_T H + \int_0^T \delta(c_s^M(\eta_s + Z_{s-}\lambda_s)|C)\, \mathrm{d}B_s\right] \\
&= \inf_{h \in \mathcal{D}(y)} \{\mathrm{E}[V(h)] + \delta(h|\mathcal{G}_c)\}.
\end{aligned}
\tag{5.21}
$$

Moreover, we have the following relation between the primal and dual variables on the level of stochastic processes.

Lemma 5.5. *Suppose that S is in $\mathcal{H}^2_{\mathrm{loc}}(\mathrm{P})$ and satisfies the structure condition. For every $x \in \mathbb{R}$, $\vartheta \in \Theta_S(C)$, and $Z \in \mathcal{M}^2(\mathrm{P})$ with $\delta(Z_T|\mathcal{G}_c) < \infty$, the process*

$$
\left((x + G_t(\vartheta))Z_t - \int_0^t \delta(c_s^M(\eta_s + Z_{s-}\lambda_s)|C)\, \mathrm{d}B_s\right)_{0 \le t \le T}
$$

is a P-supermartingale.

Proof. The process $(x + G(\vartheta))Z - \int_0^t (\eta + Z_-\lambda)^\top c^M \vartheta\, \mathrm{d}B$ is a P-martingale by Lemma 5.3 and because Z is a P-martingale. Moreover,

$$
\int \delta(c^M(\eta + Z_-\lambda)|C)\, \mathrm{d}B - \int (\eta + Z_-\lambda)^\top c^M \vartheta\, \mathrm{d}B
$$

is adapted and increasing by the definition of the support function $\delta(\cdot|\mathcal{G}_c)$, and integrable due to Lemma 5.4 since $\delta(Z_T|\mathcal{G}_c) < \infty$. Taking the difference gives the result. ∎

Remark 5.3. In our formulation, the process $(\int \delta(c^M(\eta + Z_-\lambda)|C)\, \mathrm{d}B)$ plays a similar role as the upper variation process $A(\mathrm{Q})$ in the optional decomposition of Föllmer and Kramkov of [11]; see also Example 3.2 of [21].

Combining Lemma 5.4 with Lemma 5.5 gives a result for stochastic processes which is analogous to Theorem 5.1.

Theorem 5.2. *Assume that $\mathcal{E} = \mathcal{E}(N)$ is regular and satisfies $R_2(\mathrm{P})$, and that $S \in \mathcal{H}^2_{\mathrm{loc}}(\mathrm{P})$ is an $\mathcal{E}$-local martingale. Let $C : \overline{\Omega} \to 2^{\mathbb{R}^d}$ be a predictable correspondence with closed convex values such that $\Theta_S(C)$ is nonempty and the projection of C on the predictable range of S is closed, i.e. $\Pi^S(\omega, t)C(\omega, t)$ is closed P_B-a.e. Then the following statements hold.*

(a) *For every $x \in \mathbb{R}$ and $H \in L^2(\mathrm{P})$, there exists a solution $\hat{\vartheta}(x) \in \Theta_S(C)$ to*

$$
\mathrm{E}\left[-\frac{1}{2}|x + G_T(\vartheta) - H|^2\right] = \max_{\vartheta \in \Theta_S(C)} !
$$

All solutions $\hat{\vartheta}(x)$ have the same gains process $G(\hat{\vartheta}(x))$ and satisfy

$$
x + G_T(\hat{\vartheta}(x)) = I(\hat{Z}_T) = -\hat{Z}_T + H,
$$

where $\hat{Z} \in \mathcal{Y}(\hat{y}(x))$ and $\hat{y}(x) \in \mathbb{R}$ are respectively the unique solutions to

$$
\Psi(Z_T) = \mathrm{E}\left[\frac{1}{2}Z_T^2 - Z_T H + \int_0^T \delta(c_s^M(\eta_s + Z_{s-}\lambda_s)|C)\, \mathrm{d}B_s\right] = \min_{Z \in \mathcal{Y}(\hat{y}(x))} !
$$

and

$$
v(y) + xy = \min_{y \in \mathbb{R}} !
$$

(b) *The value functions u and v are conjugate, i.e.*

$$u(x) = \inf_{y \in \mathbb{R}} \{v(y) + xy\}, \qquad v(y) = \sup_{x \in \mathbb{R}} \{u(x) - xy\},$$

and continuously differentiable. Furthermore, u is strictly concave and v is strictly convex.

(c) *The process* $(x + G(\hat{\vartheta}(x)))\hat{Z} - \int \delta(c^M(\hat{\eta} + \hat{Z}_-\lambda)|C)\,\mathrm{d}B$ *is a* P*-martingale for all solutions* $\hat{\vartheta}(x)$*, and* $(\hat{\eta}_s + \hat{Z}_{s-}\lambda_s)^\top c_s^M \hat{\vartheta}_s(x) = \delta(c_s^M(\hat{\eta}_s + \hat{Z}_{s-}\lambda_s)|C)$*,* P_B*-a.e.*

Proof. (a) By of Theorem 3.1(c), $\mathcal{G}_c = G_T(\Theta_S(C))$ is a nonempty, closed, and convex subset of $L^2(\mathrm{P})$. Hence, we can apply Theorem 5.1 to obtain unique solutions $\hat{f}(x) \in \mathcal{C}(x)$ to (5.2) and $\hat{h}(\hat{y}(x)) \in \mathcal{D}(\hat{y}(x))$ to (5.8). Since $\hat{f}(x) - x$ is in $\mathcal{G}_c$, there exists some $\hat{\vartheta}(x) \in \Theta_S(C)$ with $x + G_T(\hat{\vartheta}(x)) = \hat{f}(x)$ which is a solution to (5.1), and, since $\hat{f}(x)$ is unique, this equality must hold for all solutions. As $G(\hat{\vartheta}(x))$ is an $\mathcal{E}$-martingale, it is uniquely determined by its terminal value and so all solutions $\hat{\vartheta}(x)$ have this as their gains process. Identifying $\hat{h}(\hat{y}(x))$ with $\hat{Z}$ shows that $\hat{Z}$ solves (5.20); this uses the observation before (5.21) that the functions Ψ in (5.8) and (5.20) coincide due to Lemma 5.4.

(b) Since the value functions of (5.1) and (5.2), and (5.20) and (5.8), respectively, coincide, the assertion follows immediately from Theorem 5.1(b).

(c) By Lemma 5.5, the process $(x + G(\hat{\vartheta}(x)))\hat{Z} - \int \delta(c^M(\hat{\eta} + \hat{Z}_-\lambda)|C)\,\mathrm{d}B$ is a P-super-martingale with initial value $x\hat{y}(x)$ and final value

$$(x + G_T(\hat{\vartheta}(x)))\hat{Z}_T - \int_0^T \delta(c_s^M(\hat{\eta}_s + \hat{Z}_{s-}\lambda_s)|C)\,\mathrm{d}B_s$$
$$= \hat{f}(x)\hat{h}(\hat{y}(x)) - \int_0^T \delta(c_s^M(\hat{\eta}_s + \hat{Z}_{s-}\lambda_s)|C)\,\mathrm{d}B_s.$$

Moreover, Lemma 5.4 shows that

$$\mathrm{E}\left[\int_0^T \delta(c_s^M(\hat{\eta}_s + \hat{Z}_{s-}\lambda_s)|C)\,\mathrm{d}B_s\right] = \delta(\hat{Z}_T|G_T(\Theta_S(C))) = \delta(\hat{h}(\hat{y}(x))|\mathcal{G}_c).$$

Hence, relation (5.11a) implies that the above process has constant expectation and is therefore a P-martingale. Combining this with Lemma 5.3 tells us that the increasing process $\int \delta(c^M(\hat{\eta} + \hat{Z}_-\lambda)|C)\,\mathrm{d}B - \int (\hat{\eta} + \hat{Z}_-\lambda)^\top c^M \hat{\vartheta}(x)\,\mathrm{d}B$ is a martingale null at 0 and, hence, indistinguishable from the zero process. Since the definition of the support function yields $(\hat{\eta}_s + \hat{Z}_{s-}\lambda_s)^\top c_s^M \hat{\vartheta}_s(x) \leq \delta(c_s^M(\hat{\eta}_s + \hat{Z}_{s-}\lambda_s)|C)$, we must have equality P_B-a.e., and this completes the proof. ∎

5.3. Related work

Our approach combines duality techniques and constraints with quadratic optimisation problems and so has connections to several areas, in particular utility maximisation under constraints. Very informally, our results can be viewed as the special case of a state-dependent quadratic utility $U(x, \omega) = -\frac{1}{2}|x - H(\omega)|^2$. But, they cannot be deduced directly because this 'utility function' is not increasing in x and since the duality must be taken in a different setting (L^2 instead of L_+^0). Let us explain the relations in more detail.

The oldest neoclassical work on utility maximisation under constraints is probably by Cvitanić and Karatzas [5]. In an Itô process setting, they introduced the basic ideas of using convex duality and working with the support function of the constraint set to describe the dual variables and also the dual criterion. The seminal work of Kramkov and Schachermayer [18]

extended the duality idea to general semimartingale models without trading constraints. One key idea there was to separate the duality arguments into a static level of random variables and a dynamic level of stochastic processes, as in Sections 5.1 and 5.2. For the static level, this also needed a bipolar theorem in L^0_+. In Karatzas and Žitković [17], general semimartingale models were combined with *cone* constraints on trading strategies, and the optional decomposition theorem under constraints from [11] was used to obtain the basic duality characterisation of superreplicable consumption-investment pairs. In contrast to [5], the support function δ of the constraint set did not show up explicitly since the latter was a cone; see Remark 5.1. However, Karatzas and Žitković [17] obtained a full duality result in the sense that, as in [18], they could prove the existence of an optimiser for the dual problem and then use that to construct an optimiser for the primal problem. The paper by Mnif and Pham [21] is more general in that it allows American-type as well as convex (not necessarily conic) constraints and does not impose nonnegativity for (intermediate values of) the wealth process. The last fact makes it impossible to parametrise strategies by fractions of wealth, and this in turn forces one to use the additive form of the optional decomposition under constraints. Together with the general convex constraints, this leads to an additional term in the objective function for the dual problem. Owing to these complications, Mnif and Pham [21] only obtained a partial (verification) duality result; they showed how to construct a primal from a dual optimiser, but did not prove existence of a dual optimiser.

The utility paper closest to our results is probably that of Pham [22], who worked in finite discrete time with cone constraints (so that, as explained in Remark 5.1, the dual objective function has no explicit extra term), with the key (superreplication) duality resting on the monotonicity of the utility function. But, as in our approach, Pham did not impose nonnegativity constraints on the wealth, and the underlying duality was formulated in an (L^p, L^q)-setting.

The second area of related work is mean-variance hedging and mean-variance portfolio selection. Like utility maximisation, this is huge, and we focus only on a small sample of papers. (An attempt at a broader overview can be found in [28].) Duality for mean-variance hedging without constraints is discussed by Hou and Karatzas [12]. An abstract and static formulation of Markowitz-type problems under cone constraints is given by Sun and Wang [29]; this is similar to Section 5.1, but gives no duality and is considerably simpler since constraints are conic and there is no contingent claim. Labbé and Heunis [19] studied quadratic utility maximisation problems in an Itô process model whose completeness is destroyed by having convex constraints on trading strategies. They introduced (in a fairly complicated way, to our mind) a dual problem for certain processes, showed that this has a solution, and constructed from that a solution to the original problem. Via Itô's representation theorem, the last step crucially exploits the completeness of the unconstrained market. The existence proof for the dual optimiser is analogous to our Proposition 5.2, and, as in Lemma 5.1, the objective function involves an extra term from the support function of the constraint set. It is a matter of taste whether our results are simpler or more natural than those in [19], but they are definitely much more general.

Markowitz problems in complete and incomplete Itô process models are also studied by Hu and Zhou [13] and Jin and Zhou [14]. The former has cone constraints on strategies, the latter imposes no short sale constraints (which are also described by cones), and both use (quadratic or linear) BSDEs to obtain a solution. This setup has a lot of extra structure, and the continuity of asset prices simplifies matters considerably. For an extension to general semimartingale models with cone constraints and a more detailed discussion, we refer the reader to forthcoming work in [7].

Acknowledgements

References

[1] ALIPRANTIS, C. D. AND BORDER, K. C. (2006). *Infinite Dimensional Analysis*, 3rd edn. Springer, Berlin.

[2] AUBIN, J.-P. (2000). *Applied Functional Analysis*, 2nd edn. Wiley-Interscience, New York.

[3] BIELECKI, T. R., JIN, H., PLISKA, S. R. AND ZHOU, X. Y. (2005). Continuous-time mean-variance portfolio selection with bankruptcy prohibition. *Math. Finance* **15,** 213–244.

[4] CHOULLI, T., KRAWCZYK, L. AND STRICKER, C. (1998). $\mathcal{E}$-martingales and their applications in mathematical finance. *Ann. Prob.* **26,** 853–876.

[5] CVITANIĆ, J. AND KARATZAS, I. (1992). Convex duality in constrained portfolio optimization. *Ann. Appl. Prob.* **2,** 767–818.

[6] CZICHOWSKY, C. AND SCHWEIZER, M. (2011). Closedness in the semimartingale topology for spaces of stochastic integrals with constrained integrands. In *Séminaire de Probabilités XLIII* (Lecture Notes Math. **2006**), Springer, Berlin, pp. 413–436

[7] CZICHOWSKY, C. AND SCHWEIZER, M. (2012). Cone-constrained continuous-time Markowitz problems. To appear in *Ann. Appl. Prob.*

[8] DELBAEN, F. (2006). The structure of m-stable sets and in particular of the set of risk neutral measures. In *In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX* (Lecture Notes Math. **1874**), Springer, Berlin, pp. 215–258.

[9] DELBAEN, F. *et al.* (1997). Weighted norm inequalities and hedging in incomplete markets. *Finance Stoch.* **1,** 181–227.

[10] EKELAND, I. AND TEMAM, R. (1976). *Convex Analysis and Variational Problems*. North-Holland, Amsterdam.

[11] FÖLLMER, H. AND KRAMKOV, D. (1997). Optional decompositions under constraints. *Prob. Theory Relat. Fields* **109,** 1–25.

[12] HOU, C. AND KARATZAS, I. (2004). Least-squares approximation of random variables by stochastic integrals. In *Stochastic Analysis and Related Topics in Kyoto* (Adv. Stud. Pure Math. **41**), Mathematical Society, Japan, Tokyo, pp. 141–166.

[13] HU, Y. AND ZHOU, X. Y. (2005). Constrained stochastic LQ control with random coefficients, and application to portfolio selection. *SIAM J. Control Optimization* **44,** 444–466.

[14] JIN, H. AND ZHOU, X. Y. (2007). Continuous-time Markowitz's problems in an incomplete market, with no-shorting portfolios. In *Stochastic Analysis and Applications* (Abel Symp. **2**), Springer, Berlin, pp. 435–459.

[15] KARATZAS, I. AND KARDARAS, C. (2007). The numéraire portfolio in semimartingale financial models. *Finance Stoch.* **11,** 447–493.

[16] KARATZAS, I. AND SHREVE, S. E. (1998). *Methods of Mathematical Finance* (Appl. Math. **39**). Springer, New York.

[17] KARATZAS, I. AND ŽITKOVIĆ, G. (2003). Optimal consumption from investment and random endowment in incomplete semimartingale markets. *Ann. Prob.* **31,** 1821–1858.

[18] KRAMKOV, D. AND SCHACHERMAYER, W. (1999). The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Prob.* **9,** 904–950.

[19] LABBÉ, C. AND HEUNIS, A. J. (2007). Convex duality in constrained mean-variance portfolio optimization. *Adv. Appl. Prob.* **39,** 77–104.

[20] MÉMIN, J. (1980). Espaces de semi martingales et changement de probabilité. *Z. Wahrscheinlichkeitsth.* **52,** 9–39.

[21] MNIF, M. AND PHAM, H. (2001). Stochastic optimization under constraints. *Stoch. Process. Appl.* **93,** 149–180.

[22] PHAM, H. (2000). Dynamic L^p-hedging in discrete time under cone constraints. *SIAM J. Control Optimization* **38,** 665–682.

[23] PHAM, H. (2002). Minimizing shortfall risk and applications to finance and insurance problems. *Ann. Appl. Prob.* **12,** 143–172.

[24] PROTTER, P. E. (2005). *Stochastic Integration and Differential Equations* (Stoch. Modelling Appl. Prob. **21**). Springer, Berlin.

[25] ROCKAFELLAR, R. T. (1970). *Convex Analysis*. (Princeton Math. Ser. **28**). Princeton University Press, Princeton, NJ.

[26] ROCKAFELLAR, R. T. (1976). Integral functionals, normal integrands and measurable selections. In *Nonlinear Operators and the Calculus of Variations* (Lecture Notes Math. **543**), Springer, Berlin, pp. 157–207

[27] SCHWEIZER, M. (2001). A guided tour through quadratic hedging approaches. In *Option Pricing, Interest Rates and Risk Management.* Cambridge University Press, pp. 538–574.

[28] SCHWEIZER, M. (2010). Mean-variance hedging. In *Encyclopedia of Quantitative Finance*, ed. R. Cont, John Wiley, pp. 1177–1181.

[29] SUN, W. G. AND WANG, C. F. (2006). The mean-variance investment problem in a constrained financial market. *J. Math. Econom.* **42,** 885–895.