Mean-Variance Hedging and Stochastic Control: Beyond the Brownian Setting

Olga Bobrovnytska and Martin Schweizer

***Abstract*—We show for continuous semimartingales in a general filtration how the mean-variance hedging problem can be treated as a linear-quadratic stochastic control problem. The adjoint equations lead to backward stochastic differential equations for the three coefficients of the quadratic value process, and we give necessary and sufficient conditions for the solvability of these generalized stochastic Riccati equations. Motivated from mathematical finance, this paper takes a first step toward linear-quadratic stochastic control in more general than Brownian settings.**

***Index Terms*—Backward stochastic differential equations, linear-quadratic stochastic control, mean-variance hedging, reverse Hölder inequality, stochastic Riccati equations, variance-optimal martingale measure.**

I. Introduction

STOCHASTIC control methods have a venerable history in the field of financial engineering, and a number of Nobel Prizes bear ample witness to the fruitfulness of this interaction. One can for instance think of Merton's seminal contributions to portfolio optimization and option pricing, among other things. Even earlier, Harry Markowitz was concerned with mean-variance analysis in financial markets, and this topic has retained its popularity even after 50 years; see, for instance, the survey [23] which contains more than 200 references. However, in contrast to portfolio optimization based on utility functions, mean-variance analysis in dynamic intertemporal frameworks has only recently been linked to stochastic control in a more systematic way. Our goal in this paper is to explore this avenue further and to show that it leads to results and insights in stochastic control even beyond the usual settings.

In a given financial market, the *mean-variance hedging* problem is to find for a given payoff a best approximation by means of self-financing trading strategies; the optimality criterion is the expected squared error. In a series of recent papers, this problem has been formulated and treated as a *linear-quadratic* (LQ) *stochastic control* problem at increasing levels of generality; see for instance [11], [13], [8], [7], [10], [24], or [9] for an overview and a historical perspective. In the general case where the market coefficients are random processes, the adjoint equations turn out to lead to a coupled system of *backward stochastic differential equations* (BSDEs) for the coefficients of the (quadratic) value functional. This has led to new interest in and new results on general LQ stochastic control problems, and the mean-variance hedging problem has been treated fairly explicitly by these methods.

From the mathematical finance point of view, one drawback of this approach is that almost all existing papers impose rather restrictive assumptions. To apply general results from LQ stochastic control, authors work with Itô processes and assume that all their coefficients are uniformly bounded, which excludes many practically relevant models. Moreover, the theory of BSDEs is only rarely used beyond the setting of a filtration generated by a Brownian motion and, thus, strongly relies on a martingale representation theorem. On the other hand, the mean-variance hedging problem has been solved in much higher generality by martingale and projection techniques. One can allow continuous semimartingales in general filtrations and only needs an absence-of-arbitrage condition; see [17] and [22] for recent overviews. The present paper is a first step toward a fusion beetween mathematical finance and LQ stochastic control at this more general level. For related recent results, see [16].

The paper is structured as follows. Section II presents the basic model, explains the mean-variance hedging problem and casts it in the form of an LQ stochastic control problem. Combining the martingale optimality principle with the natural guess that the value process of this problem should have a quadratic structure, we then derive a system of BSDEs for the conjectured coefficients a, b, c. Section III gives a necessary and sufficient condition for the first of these BSDEs (for the quadratic coefficient a) to be solvable under the sole assumption that the underlying asset price process S is continuous. One can also show that $1/a$ is the value process of a dual control problem, but we do not dwell on this issue here. Section IV gives sufficient conditions for the other BSDEs (for the linear and constant coefficients b and c, respectively) to be solvable. Apart from continuity of S, we need that the filtration $\mathbb{F}$ is continuous and that the variance-optimal martingale measure $\tilde{P}$ satisfies the reverse Hölder inequality $R_2(P)$. Finally, SectionV shows how one can explicitly construct a solution for the mean-variance hedging problem from the solutions of the BSDEs for a, b, c. This is conceptually well known, but of course technically slightly different than in the usual case of a Brownian filtration.

II. Basic Setup

This section introduces the model and the concepts used throughout the rest of the paper: the mean-variance hedging

Manuscript received October, 2002; revised October, 2003. Recommended by Guest Editor B. Pasik-Duncan. This work was supported by the Deutsche Forschungsgemeinschaft through the Graduiertenkolleg "Stochastische Prozesse und probabilistische Analysis" at the Technical University of Berlin, Berlin, Germany.

O. Bobrovnytska is with the Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin, Germany.

M. Schweizer is with Mathematik, the ETH Zürich, ETH-Zentrum, CH-8092 Zürich, Switzerland (e-mail: mschweiz@math.ethz.ch).

Digital Object Identifier 10.1109/TAC.2004.824468

(MVH) problem, its formulation as a linear-quadratic stochastic control (LQSC) problem, and the associated backward stochastic differential equations (BSDEs). The goal is to study the relations between these objects.

Our starting point is a stochastic model for the evolution of the discounted prices for finitely many assets in a financial market. We begin with a finite time horizon $T > 0$ and a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ satisfying the usual conditions of right-continuity and completeness. $S = (S_t)_{0 \le t \le T}$ is an $\mathbb{R}^d$-valued semimartingale and we think of S_t^i as asset i's discounted price at time t. In addition, there is a riskless asset whose discounted price is 1 at all times.

Throughout this paper, we impose the **standing assumption** that

$$S \text{ is continuous.} \tag{1}$$

Hence, S can be uniquely written as

$$S = S_0 + M + A$$

with an $\mathbb{R}^d$-valued continuous local martingale M and an $\mathbb{R}^d$-valued continuous adapted process A of finite variation (FV), both null at 0. In addition, we assume that A has the form

$$A_t = \int_0^t d\langle M \rangle_s \, \lambda_s, \qquad 0 \le t \le T \tag{2}$$

for an $\mathbb{R}^d$-valued predictable process $\lambda \in L^2_{\text{loc}}(M)$, i.e., satisfying $\int_0^T \lambda_s^\top \, d\langle M \rangle_s \, \lambda_s < \infty$ P-a.s. This is a weak absence-of-arbitrage-type condition on S; see [4] or [20].

Example: The *standard example* to be kept in mind is a multidimensional *Itô process model* of the following type. Let W be an $\mathbb{R}^n$-valued Brownian motion on $(\Omega, \mathcal{F}, P)$ and $\mathbb{F} = \mathbb{F}^W$ the P-augmentation of the filtration generated by W. Let $\bar{S} = (\bar{S}^i)_{i=1,\ldots,d}$ be the solution of the stochastic differential equations

$$d\bar{S}_t^i = \bar{S}_t^i \left(\mu_t^i \, dt + \sum_{j=1}^n \sigma_t^{ij} \, dW_t^j \right), \qquad \bar{S}_0^i > 0$$

with predictable $\mathbb{R}^d$- and $\mathbb{R}^{d \times n}$-valued processes μ and σ that are P-a.s. on $[0, T]$ Lebesgue-integrable and Lebesgue-square-integrable, respectively. $\bar{S}$ models undiscounted prices and $\bar{S}^0$, which is of the same form as the other $\bar{S}^i$, describes an additional asset that we use as a numeraire for discounting; so in our abstract setup, we consider $S^i := \bar{S}^i / \bar{S}^0$ for $i = 1, \ldots, d$. We can for instance think of $\bar{S}^0$ as a reference stock or a zero coupon bond with a suitable maturity. But for concreteness, we focus on the special case where $\sigma^{0j} \equiv 0$ so that $\bar{S}^0$ has finite variation. $\bar{S}^0$ is then called a (classical) *savings account*, $r := \mu^0$ is the instantaneous *short rate*, and S^i takes the form

$$dS_t^i = S_t^i \left((\mu_t^i - r_t) \, dt + \sum_{j=1}^n \sigma_t^{ij} \, dW_t^j \right).$$

The condition that A has the form (2) is then equivalent to assuming that $\mu_t - r_t \underline{1} \in \text{range}(\sigma_t \sigma_t^\top)$ P-a.s. for all t, with $\underline{1} = (1 \ldots 1)^\top \in \mathbb{R}^d$, and then λ is given by $\lambda_t^i = \bar{\lambda}_t^i / S_t^i$ with $\bar{\lambda} = (\sigma\sigma^\top)^{-1}(\mu - r\underline{1})$, provided that σ_t has full rank $d \le n$ P-a.s. for all t. A frequently encountered assumption is that the *market price of risk*, $\varphi := \sigma^\top \bar{\lambda} = \sigma^\top (\sigma\sigma^\top)^{-1}(\mu - r\underline{1})$, is bounded uniformly in t and ω; this implies that

$$\begin{aligned} \int_0^T \lambda_s^\top \, d\langle M \rangle_s \, \lambda_s &= \int_0^T (\mu_s - r_s\underline{1})^\top (\sigma_s \sigma_s^\top)^{-1} (\mu_s - r_s \underline{1}) \, ds \\ &= \int_0^T |\varphi_s|^2 \, ds \end{aligned}$$

is bounded P-a.s. as well. We shall refer to this entire setup as the standard example with the *standard assumptions*.

One very special feature of the standard example is that one has a martingale representation theorem in the filtration $\mathbb{F}^W$. This simplifies many things, and one of our goals here is to avoid this type of assumption. □

We now return to our general setting and introduce the MVH problem. Let Θ be a linear subspace of the set of all $\mathbb{R}^d$-valued predictable S-integrable processes. The stochastic integral process $G(\vartheta) := \int \vartheta \, dS$ is thus well-defined for every $\vartheta \in \Theta$, and we assume that $G_T(\vartheta) \in L^2(P)$ for every $\vartheta \in \Theta$, i.e., $G_T(\Theta) \subseteq L^2(P)$. However, this is not enough to obtain good results; one also needs to impose integrability properties on $G(\vartheta)$ as a process. This issue will be addressed more carefully later on (in Section III) when the need arises.

The *mean-variance hedging problem* (for H, with respect to Θ) is to

$$\text{minimize } E[(H - v_0 - G_T(\vartheta))^2] \quad \text{over all pairs} (v_0, \vartheta) \in \mathbb{R} \times \Theta \tag{3}$$

where $H \in L^2(\mathcal{F}_T, P)$ is a square-integrable $\mathcal{F}_T$-measurable random variable. The financial interpretation of (3) is as follows. Any pair (v_0, ϑ) describes a dynamic trading strategy which starts at time 0 with initial capital v_0, holds ϑ_t^i shares of asset i at time t and is self-financing, thus leading to a wealth of $V_t^\vartheta = v_0 + G_t(\vartheta)$ at time t. The random variable H models the discounted payoff of some financial instrument. This is usually some obligation and so (3) has the goal of minimizing, by the choice of a strategy, the average hedging error measured as the expected quadratic deviation of the final wealth V_T^ϑ from the target H. Mathematically, this amounts to projecting in $L^2(P)$ on the linear subspace $G_T(\Theta)$, and so (3) is solvable for any $H \in L^2(P)$ if and only if $G_T(\Theta)$ is closed in $L^2(P)$. We will return to this in Sections III and V.

The MVH problem is seen to be a special case of a *LQ stochastic control problem* if we rewrite it as

$$dV_t^\vartheta = \vartheta_t \, dS_t = \vartheta_t \, dM_t + \vartheta_t^\top \, d\langle M \rangle_t \, \lambda_t, \qquad V_0^\vartheta = v_0$$

with the objective function

$$F(V, \vartheta) = E\left[\left(V_T^\vartheta - H \right)^2 \right]$$

which is to be minimized by the choice of $\vartheta \in \Theta$. This is actually a slight variation of (3) since v_0 is taken as fixed. However, if one can solve the above problem for any v_0, one can also solve (3) by simply optimizing over v_0.

Example: In the standard example introduced previously, the LQSC problem is first written as

$$dV_t^\vartheta = \vartheta_t^\top \mathrm{diag}(S_t) b_t \, dt + \vartheta_t^\top \mathrm{diag}(S_t)\sigma_t \, dW_t, \qquad V_0^\vartheta = v_0$$

with $b = \mu - r\underline{1}$ and with the same objective function as above. To bring this closer to the usual formulations found in the literature, we introduce

$$\pi_t^i := \vartheta_t^i S_t^i, \qquad 0 \le t \le T$$

which is the amount invested in asset i at time t. Then an obvious change of notation gives

$$dV_t^\pi = \pi_t^\top b_t \, dt + \pi_t^\top \sigma_t \, dW_t, \qquad V_0^\pi = v_0$$

and

$$F(V,\pi) = E\left[(V_T^\pi - H)^2 \right].$$

If we compare this to a typical LQSC problem of the form

$$dx_t = (A(t)x_t + B(t)u_t)\, dt + \left(x_t^\top C(t) + u_t^\top D(t) \right) \, dW_t$$

with random coefficients and objective function

$$F(x,u) = E\left[\int_0^T \left(x_t^\top Q(t) x_t + x_t^\top P(t) u_t + u_t^\top N(t) u_t \right) \, dt \right. \\ \left. + R(T)|x_T - O(T)|^2 \right]$$

we see that we have the special case where $u \mathrel{\widehat{=}} \pi, x \mathrel{\widehat{=}} V$ is one-dimensional and $A \equiv 0, B^\top = b = \mu - r\underline{1}, C \equiv 0, D = \sigma, Q \equiv 0, P \equiv 0, N \equiv 0, R(T) = 1, O(T) = H$. □

In order to study the MVH problem (3), we now fix $H \in L^2(\mathcal{F}_T, P)$ and $v_0 \in I\!R$ and consider the modified MVH problem to

$$\text{minimize } E[(H - v_0 - G_T(\vartheta))^2] \text{ over all } \vartheta \in \Theta. \tag{4}$$

The solution of (4), if it exists, will be denoted by ϑ^* and is called an optimal strategy. As usual in stochastic control, it is helpful to consider a dynamic version of (4). So, we introduce for each $\vartheta \in \Theta$ and $t \in [0,T]$ the set

$${}^t\Theta(\vartheta) := \{ \psi \in \Theta \,|\, \psi = \vartheta \text{ on } \Omega \times [0,t] \}$$

of all strategies that coincide with ϑ up to time t, we define the random variables $\Gamma_t(\psi) := E[(H - v_0 - G_T(\psi))^2 \,|\, \mathcal{F}_t]$, and we set

$$\begin{aligned} J_t(\vartheta) &:= \mathrm{ess\,inf}_{\psi \in {}^t\Theta(\vartheta)} \Gamma_t(\psi) \\ &= \mathrm{ess\,inf} \left\{ E[(H - v_0 - G_T(\psi))^2 \,|\, \mathcal{F}_t] \,|\, \psi \in {}^t\Theta(\vartheta) \right\}. \end{aligned}$$

Since $v_0 + G_T(\psi) = v_0 + G_t(\vartheta) + \int_t^T \psi_u \, dS_u$ for $\psi \in {}^t\Theta(\vartheta), J_t(\vartheta)$ is the minimal expected squared hedging error conditional on $\mathcal{F}_t$ if we start at time t with "initial capital" $V_t^\vartheta = v_0 + G_t(\vartheta)$, the wealth we have achieved by using (v_0, ϑ) up to time t.

Our first result is simply the *martingale optimality principle* for stochastic control in the present setting. Note that this result is valid without the assumption that S is continuous. The proof is a standard dynamic programming argument as for [12, Th. 4.1] and, therefore, omitted. With slight anticipation, we also remark that the imposed assumption on the family of random variables $\Gamma_t(\psi)$ is satisfied for the specific space Θ that we introduce later in Section III.

Proposition 1: Fix $H \in L^2(\mathcal{F}_T, P)$ and $v_0 \in I\!R$. Assume that Θ is such that for fixed t and ϑ, the family $\{\Gamma_t(\psi) \,|\, \psi \in {}^t\Theta(\vartheta)\}$ is stable under taking minima. Then

1) for each $\vartheta \in \Theta$, the process $J(\vartheta)$ is a P-submartingale and has an RCLL version that we again denote by $J(\vartheta)$;
2) a process $\vartheta^* \in \Theta$ solves (4) if and only if $J(\vartheta^*)$ is a P-martingale.

Because our stochastic control problem is quadratic, we guess that its value process has a quadratic structure as well. Since the optimal strategy is obtained by "setting the derivative equal to 0," we also guess that ϑ^* and V^{ϑ^*} should be affinely related. It will turn out that both guesses are correct. Moreover, using the first guess as an a priori assumption will help us to find a systematic way of attacking the MVH problem. So let us see where such an assumption gets us—again even without continuity of S.

Lemma 2: Fix $H \in L^2(\mathcal{F}_T, P)$ and $v_0 \in I\!R$ and suppose that there exists some $\vartheta^* \in \Theta$ such that for each $t \in [0,T], J_t(\vartheta^*)$ is a quadratic function of $V_t^{\vartheta^*}$. Then, there exist adapted stochastic processes a, b, c not depending on ϑ such that we have for every $\vartheta \in \Theta$

$$J_t(\vartheta) = a_t \left(V_t^\vartheta - b_t \right)^2 + c_t \quad P-\text{a.s.}, \qquad 0 \le t \le T. \tag{5}$$

Proof: By assumption, $J_t(\vartheta^*) = a_t(V_t^{\vartheta^*} - b_t)^2 + c_t$ is a quadratic function of $V_t^{\vartheta^*}$, with of course $\mathcal{F}_t$-measurable coefficients. On the other hand, the definition of ${}^t\Theta(\vartheta^*)$ yields

$$\begin{aligned} J_t(\vartheta^*) &= \mathrm{ess\,inf}_{\psi \in {}^t\Theta(\vartheta^*)} \Gamma_t(\psi) \\ &= \mathrm{ess\,inf}_{\psi \in \Theta} E\left[\left(H - V_t^{\vartheta^*} - \int_t^T \psi_u \, dS_u \right)^2 \middle| \mathcal{F}_t \right] \end{aligned}$$

and in the same way for any $\vartheta \in \Theta$

$$J_t(\vartheta) = \mathrm{ess\,inf}_{\psi \in \Theta} E\left[\left(H - V_t^{\vartheta} - \int_t^T \psi_u \, dS_u \right)^2 \middle| \mathcal{F}_t \right].$$

Thus we see that $J_t(\vartheta)$ has the same functional dependence of V_t^ϑ as $J_t(\vartheta^*)$ of $V_t^{\vartheta^*}$. This establishes (5) and also makes it clear that the coefficients a, b, c do not depend on ϑ. q.e.d.

Remarks:

1) That the value process $J(\vartheta^*)$ has a quadratic structure was earlier noticed or conjectured by several authors in particular cases; see for instance [2] for a situation in finite discrete time with finite Ω, [1] for a study of some Markovian models in discrete and continuous time, or [8],

[10], and [24] in the framework of an Itô process model. For the present setting, Mania and Tevzadze have very recently proved in [16] under an additional assumption that $J(\vartheta^*)$ is indeed a quadratic function of V^{ϑ^*}. However, the thrust of Lemma 2 is different because we want to examine $J(\vartheta)$ for arbitrary $\vartheta \in \Theta$.

2) The generality of Lemma 2 comes at a price because it uses an unchecked assumption about ϑ^*, and its proof gives us so far no information about the structure or even path regularity of the processes a, b, c. These issues will be dealt with in future work. ◇

By combining Proposition 1 and Lemma 2, we obtain in a conceptually straightforward way a systematic method to attack and solve the MVH problem. This goes as follows. Suppose we have obtained in some way the quadratic structure (5). *Assuming* that a, b, c from Lemma 2 are nice semimartingales, we exploit the martingale optimality principle in Proposition 1 to derive equations that a, b, c ought to satisfy. Then we study these equations and show how their solutions can be used to construct an optimal strategy for the MVH problem. Hence the equations provide at least *sufficient* conditions for the solvability of the MVH problem. Because the derivation (which is sketched later) of the equations is only heuristic, we do not know yet if they are also necessary conditions. However, we think this is the case and relegate the issues of proving (5) and of a rigorous derivation for the equations to future work. For some very recent progress, see also [24] and [16].

So let us assume that a, b, c are well-behaved semimartingales and that there exists an optimal strategy ϑ^*. Then we know from (5) and Proposition 1 that $J(\vartheta) = a(V_T^\vartheta - b)^2 + c$ is a submartingale for every $\vartheta \in \Theta$ and a martingale for $\vartheta = \vartheta^*$. Hence if we use Itô's formula to compute the canonical decomposition of $J(\vartheta)$, we know that the minimum over ϑ of the finite variation (FV) term must be 0 and attained by ϑ^*. If we do those calculations and formally differentiate with respect to ϑ, we first find that ϑ^* must be an affine function of $G(\vartheta^*)$. (This will actually be confirmed later; see Proposition 10 in Section V.) Plugging this back in leads to an explicit expression for the FV part of $J(\vartheta^*)$, namely a quadratic function of $G(\vartheta^*)$. But $J(\vartheta^*)$ is a martingale, so the FV part must vanish, and setting the coefficients of all powers of $G(\vartheta^*)$ to be 0 produces (under suitable assumptions) the following system of equations for a, b, c:

$$da_t = a_{t-}\left(\lambda_t + \frac{\Lambda_t^a}{a_{t-}}\right)^\top d\langle M\rangle_t \left(\lambda_t + \frac{\Lambda_t^a}{a_{t-}}\right) + \Lambda_t^a\,dM_t + dN_t^a, \qquad a_T = 1 \tag{6}$$

$$db_t = \mu_t^b\,dS_t + \frac{\psi_t^b}{a_{t-}}\,dN_t^a - \frac{\psi_t^b}{a_{t-}^2}\,d\langle N^a\rangle_t + dN_t^b, \qquad b_T = H \tag{7}$$

$$dc_t = -\frac{(\psi_t^b)^2}{a_{t-}} d\langle N^a\rangle_t - a_{t-} d\langle N^b\rangle_t + dN_t^{(c)}, \qquad c_T = 0. \tag{8}$$

Since we study these equations in detail in the rest of the paper, we postpone most explanations for the moment. We just note that (6)–(8) all are BSDEs, that (7) needs ingredients from (6), and (8) even from both (6) and (7). We also mention that solutions of (6)–(8) are, respectively, tuples (a, Λ^a, N^a), (b, μ^b, ψ^b, N^b), and $(c, N^{(c)})$, with properties still to be specified. Our goal is to study the solvability of these equations under minimal assumptions.

Equations (6)–(8) can be viewed as general versions of earlier results derived under more specific assumptions. For the case of an Itô process model, we explain below how we recover some equations from [8]. Prior to that, [1] has studied the MVH problem under Markovian assumptions in both discrete and continuous time; the general discrete-time case has been analyzed (in an unpublished German diploma thesis) by Lehweß–Litzmann who eliminated the assumption in [2] that Ω is finite.

In discrete time, one can rigorously prove by backward induction both the martingale optimality principle and the fact that $J(\vartheta)$ has the quadratic structure (5), without having to assume an *a priori* relation between $J(\vartheta^*)$ and V^{ϑ^*}; this comes out as a consequence. Again by backward induction, one can derive a system of difference equations that a, b, c must satisfy. The BSDEs (6)–(8) are precisely the continuous-time analogues.

In continuous time and in a Markovian model, [1] derives a system of three PDEs from the HJB equation associated to the MVH problem. In fact, the assumption that $J(\vartheta^*)$ is a quadratic function of V^{ϑ^*} implies in that Markovian situation that $J_t(\vartheta^*)$ is a function j^* of $V_t^{\vartheta^*}$ and the underlying Markovian state variables. The martingale optimality principle reduces to the HJB equation for j^*, and the coefficient functions a, b, c of the quadratic function j^* satisfy a coupled system of PDEs which constitute the Markovian analogs of the BSDEs (6)–(8).

Example: Let us look at the structure of (6) for the case of our standard example of an Itô process model. If we simply plug in for M and $\langle M\rangle$, we obtain the equation

$$da_t = a_t\left|\left(\lambda_t + \frac{1}{a_t}\Lambda_t^a\right)^\top \mathrm{diag}(S_t)\sigma_t\right|^2 dt + (\Lambda_t^a)^\top \mathrm{diag}(S_t)\sigma_t dW_t + dN_t^a \tag{9}$$

with a local martingale N^a that must be orthogonal to $M = \int \mathrm{diag}(S)\sigma dW$. Because $I\!\!F = I\!\!F^W$, this means that $N^a = \int \nu^a\,dW$ with $\nu_t^a \in \ker(\sigma_t)$ P-a.s. for all t. So if we set $Y^a := \ell^a + \nu^a$ with $\ell^a := \sigma^\top \mathrm{diag}(S)\Lambda^a$, we see that ℓ_t^a is the projection of Y_t^a on range$(\sigma_t^\top)$ P-a.s. for all t and, thus, $\ell^a = \sigma^\top(\sigma\sigma^\top)^{-1}\sigma Y^a =: \Pi Y^a$. Moreover, $\sigma_t^\top \mathrm{diag}(S_t)\lambda_t = \varphi_t$ is in range$(\sigma_t^\top)$ P-a.s. for all t, so $\varphi^\top \nu^a \equiv 0$ and we obtain by squaring out from (9) the equation

$$\begin{aligned} da_t &= \left(|\varphi_t|^2 a_t + 2\varphi_t^\top Y_t^a + \frac{1}{a_t}(Y_t^a)^\top \sigma_t^\top \times \left(\sigma_t\sigma_t^\top\right)^{-1}\sigma_t Y_t^a\right) dt + Y_t^a\,dW_t \\ &= \left(|\varphi_t|^2 a_t + 2\varphi_t^\top Y_t^a + \frac{1}{a_t}\left(Y_t^a\right)^\top \Pi_t Y_t^a\right) dt + Y_t^a\,dW_t \end{aligned} \tag{10}$$

for the pair (a, Y^a). This coincides (except for the interest rate which is 0 in our discounted situation) with [8, eq. (117)] and, thus, shows that our setting contains parts of that paper as a special case. It is also explained in [8] how (10) reduces to a

PDE under Markovian assumptions; see, in particular, in [8, eq. (40)].

III. Equivalent Martingale Measures and Solvability for a

In this section, we relate the solvability of the BSDE (6) for a to a property of absence of arbitrage. Quite remarkably, it turns out that there is a simple and natural necessary and sufficient condition for a solution of (6) to exist.

Definition: An *equivalent martingale measure (EMM)* for S is a probability measure $Q \approx P$ with $Q = P$ on $\mathcal{F}_0$ and such that S is a local Q-martingale. The *density process* of Q with respect to P is an RCLL version of the strictly positive P-martingale $Z_t^Q = E_P[(dQ/dP) \,|\, \mathcal{F}_t], 0 \le t \le T$.

We denote by $I\!M_e$ the set of all EMM's for S and by $I\!M_e^2$ the subset of those $Q \in I\!M_e$ such that $(dQ/dP) \in L^2(P)$. For readers less familiar with mathematical finance, we point out that equivalent martingale measures for S are the objects which are naturally dual to self-financing trading strategies for S. The assumption that $I\!M_e \neq \emptyset$ is intimately related to a condition of absence of arbitrage for our financial market; see [6] for a precise formulation. We also remark that $I\!M_e \neq \emptyset$ implies the structure (2) for A because S is continuous; see [20, Th. 1]. The more stringent requirement that $I\!M_e^2 \neq \emptyset$ amounts to the condition that the dual of the MVH problem should have nonempty domain.

Provided that $I\!M_e^2 \neq \emptyset$, the *variance-optimal martingale measure* is the unique element $\tilde{P}$ of $I\!M_e^2$ that minimizes $\|(dQ/dP)\|_{L^2(P)}$ over all $Q \in I\!M_e^2$. This uses that S is continuous; see [5] and [21]. Also, because S is continuous, any Z^Q for $Q \in I\!M_e^2$ is of the form $Z^Q = \mathcal{E}(-\int \lambda\, dM + N^Q)$ for some locally P-square-integrable local P-martingale N^Q null at 0 (for brevity, we write $N^Q \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$) with N^Q P-orthogonal to M; see [20, Th. 1]. If $\mathcal{N}^2$ denotes the space of all $N \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ that are P-orthogonal to M and such that $\mathcal{E}(-\int \lambda\, dM + N)$ is strictly positive and in $\mathcal{M}^2(P)$, we can parametrize $I\!M_e^2$ by $\mathcal{N}^2$ via

$$I\!M_e^2 = \left\{ Q \approx P \,\middle|\, \frac{dQ}{dP} = Z_T^Q = \mathcal{E}\left(-\int \lambda\, dM + N^Q\right)_T \right. \\ \left. \text{for some } N^Q \in \mathcal{N}^2 \right\}. \quad (11)$$

Recall that the BSDE for a has the form of (6). A *solution* of (6) is a triplet (a, Λ^a, N^a) satisfying (6) and such that a is a strictly positive RCLL semimartingale with $a_- > 0, \Lambda^a$ is an $I\!R^d$-valued predictable M-integrable process, and N^a is a local P-martingale null at 0 ($N^a \in \mathcal{M}_{0,\mathrm{loc}}(P)$, for short) and P-orthogonal to M ($N^a \perp M$).

Theorem 3: The following statements are equivalent.

1) $I\!M_e^2 \neq \emptyset$.

2) The BSDE (6) has a solution (a, Λ^a, N^a) such that

$$Z(a) := \mathcal{E}\left(-\int \lambda\, dM + \int \frac{1}{a_-}\, dN^a\right) \\ \text{is strictly positive and in } \mathcal{M}^2(P). \quad (12)$$

Moreover, each of these conditions implies that we can find a solution such that

$$a \text{ is bounded from above uniformly in} \\ t, \omega\ (P - \text{ a.s., to be precise}) \quad (13)$$

and

$$\frac{1}{a} \text{ is in the class } \mathcal{D}^2, \text{i.e., } \frac{1}{a}(Z^Q)^2 \text{ is of class}(D) \\ \text{under } P \text{ for every } Q \in I\!M_e^2. \quad (14)$$

Proof: "2) $\Rightarrow$ 1)": If we take a solution (a, Λ^a, N^a) and define $N := \int (1/a_-)\, dN^a$, then N is like N^a in $\mathcal{M}_{0,\mathrm{loc}}(P)$ and P-orthogonal to M. Moreover, $Z := \mathcal{E}(-\int \lambda\, dM + N) = Z(a)$ is by (12) strictly positive and in $\mathcal{M}^2(P)$. This implies $N \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$, hence $N \in \mathcal{N}^2$, and the measure Q with density Z_T with respect to P is in $I\!M_e^2$ which proves 1).

"1) $\Rightarrow$ 2)": If $I\!M_e^2 \neq \emptyset$, the variance-optimal martingale measure $\tilde{P}$ exists in $I\!M_e^2$ since S is continuous, and we know that the process

$$\tilde{Z}_t := E_{\tilde{P}}\left[\frac{d\tilde{P}}{dP}\middle| \mathcal{F}_t\right], \qquad 0 \le t \le T \quad (15)$$

has the form $\tilde{Z} = \tilde{Z}_0 + \int \tilde{\zeta}\, dS$ for a deterministic constant $\tilde{Z}_0$ and an $I\!R^d$-valued predictable S-integrable process $\tilde{\zeta}$; see [5, Lemma 2.2]. Because $\tilde{P} \approx P, \tilde{Z}$ is strictly positive and can thus be written as $\tilde{Z} = \tilde{Z}_0\, \mathcal{E}(\int \gamma\, dS)$ with $\gamma := \tilde{\zeta}/\tilde{Z}$. Note that γ is M-integrable since $1/\tilde{Z}$ is locally bounded, $\tilde{\zeta}$ is S-integrable and S is continuous with martingale part M. On the other hand, the density process $Z^{\tilde{P}}$ has the form $Z^{\tilde{P}} = \mathcal{E}(-\int \lambda\, dM + N^{\tilde{P}})$ for some $N^{\tilde{P}} \in \mathcal{N}^2$. Setting

$$a := \frac{Z^{\tilde{P}}}{\tilde{Z}}$$

and applying Itô's formula yields

$$da = -a_-(\lambda + \gamma)\, dM + a_-\, dN^{\tilde{P}} + a_- \gamma^\top\, d\langle M \rangle\, \gamma.$$

So, if we set

$$\Lambda^a := -a_-(\lambda + \gamma) = -\frac{Z_-^{\tilde{P}}}{\tilde{Z}}\left(\lambda + \frac{\tilde{\zeta}}{\tilde{Z}}\right) \quad (16)$$

and

$$N^a := \int a_-\, dN^{\tilde{P}}$$

we see from $a_T = Z_T^{\tilde{P}}/\tilde{Z}_T = 1$ and by plugging in that the triplet (a, Λ^a, N^a) satisfies (6). Moreover, a and a_- are both strictly positive because $\tilde{Z} > 0$ is continuous and $Z^{\tilde{P}}$ is a strictly positive P-martingale; N^a is like $N^{\tilde{P}}$ a local P-martingale null at 0 and P-orthogonal to M; and Λ^a is predictable and M-integrable because $Z_-^{\tilde{P}}$ and $1/\tilde{Z}$ are locally bounded and λ, γ are both M-integrable. Thus, (a, Λ^a, N^a) is a solution of (6), and $Z(a) = \mathcal{E}(-\int \lambda\, dM + \int (1/a_-)\, dN^a) = Z^{\tilde{P}}$ is in $\mathcal{M}^2(P)$ because $\tilde{P} \in I\!M_e^2$. This proves 2).

Since we know now that 1) and 2) are equivalent, we assume 1) and prove that $a := Z^{\tilde{P}}/\tilde{Z}$ satisfies (13) and (14). By the Bayes formula and Jensen's inequality

$$\tilde{Z}_t = E_{\tilde{P}}\left[\frac{d\tilde{P}}{dP}\middle|\mathcal{F}_t\right] = \frac{1}{Z_t^{\tilde{P}}}E\left[\left(Z_T^{\tilde{P}}\right)^2\middle|\mathcal{F}_t\right] \geq Z_t^{\tilde{P}}$$

so that we get $a_t \leq 1$ $P-$a.s. for all $t \in [0,T]$. This is (13). By the same computation

$$\frac{1}{a_\tau} = \frac{1}{\left(Z_\tau^{\tilde{P}}\right)^2}E\left[\left(Z_T^{\tilde{P}}\right)^2 \mid \mathcal{F}_\tau\right]$$

for any stopping time τ, and since $\tilde{P}$ is variance-optimal, we have

$$E\left[\left(Z_T^{\tilde{P}}\middle/ Z_\tau^{\tilde{P}}\right)^2\middle|\mathcal{F}_\tau\right] \leq E\left[\left(Z_T^{Q}\middle/ Z_\tau^{Q}\right)^2\middle|\mathcal{F}_\tau\right] \quad P - \text{a.s.} \tag{17}$$

for any $Q \in I\!\!M_e^2$. In fact, if (17) fails for some $\bar{Q}$ on a set $B \in \mathcal{F}_\tau$ with $P[B] > 0$, one easily checks that the measure Q' with density process

$$Z' := Z^{\tilde{P}}I_{[[0,\tau]]} + \left(Z^{\tilde{P}}I_{B^c} + \frac{Z_\tau^{\tilde{P}}}{Z_\tau^{\bar{Q}}}Z^{\bar{Q}}I_B\right)I_{]]\tau,T]]}$$

is in $I\!\!M_e^2$ and has $E[(dQ'/dP)^2] < E[(d\tilde{P}/dP)^2]$, contradicting the optimality of $\tilde{P}$. Hence, we get

$$0 \leq \frac{1}{a_\tau}\left(Z_\tau^Q\right)^2 \leq E\left[\left(Z_T^Q\right)^2\middle|\mathcal{F}_\tau\right]$$

for any stopping time τ, and since $Z_T^Q \in L^2(P)$ for $Q \in I\!\!M_e^2$, this shows that $\frac{1}{a}(Z^Q)^2$ is of class (D) under P. Hence, we also have (14). q.e.d.

Remarks:

1) Under the additional assumption that the filtration $I\!\!F$ is continuous (i.e., all local P-martingales are continuous), we could also derive the implication "1) ⇒ 2)" in Theorem 3 from the results of Mania and Tevzadze in [14]. These authors consider the process [note the typo "sup" in their equation (2.7)]

$$V_t := \operatorname{ess\,inf}_{Q\in I\!\!M_e^2}E\left[\left(Z_T^Q\middle/ Z_t^Q\right)^2\middle|\mathcal{F}_t\right], \qquad 0 \leq t \leq T \tag{18}$$

and derive in their Corollary 1 a BSDE (3.30) that V must satisfy. The process V is the value process of the optimization problem dual to the MVH problem, namely that of finding the variance-optimal martingale measure $\tilde{P}$. Hence

$$\begin{aligned} V_t &= E\left[\left(Z_T^{\tilde{P}}\middle/ Z_t^{\tilde{P}}\right)^2\middle|\mathcal{F}_t\right] \\ &= \frac{1}{Z_t^{\tilde{P}}}E_{\tilde{P}}\left[Z_T^{\tilde{P}}\middle|\mathcal{F}_t\right] = \frac{\tilde{Z}_t}{Z_t^{\tilde{P}}} \end{aligned}$$

by the Bayes formula, and we see that a and V are related by $a = 1/V$. If we now start with the BSDE for V and use Itô's formula, we find that $1/V$ solves our BSDE (6). However, this was not the way we originally derived (6).

If $I\!\!F$ is not continuous, matters become more complicated. The recent paper [15] again contains a BSDE for the process V from (18), but both the equation and its derivation become technically much more difficult. Hence, working directly with a, as we do, appears to be the better approach. This has recently also been done in [16].

2) We have already seen at the end of Section II how the BSDE (6) boils down to (10) in the framework of our standard Itô process example. In the terminology of LQ stochastic control, (10) is a particular case of a *stochastic Riccati equation* (SRE). Such equations have recently attracted a lot of attention; see for instance [7], [10], and [24], or [9] for a survey. However, all these papers stay within the framework of a Brownian filtration $I\!\!F = I\!\!F^W$. Our BSDE (6) could be called a *generalized SRE*, and Theorem 3 is then an existence result in a general filtration. To be fair, we should add that the sharpness of our Theorem 3 (necessary and sufficient conditions) is due to the fact that our SRE has more structure than a general SRE—both (6) itself as well as its reduced form (10). ⋄

We have so far not made any uniqueness assertion about the solution of (6). In preparation for that, we now first explain how one can construct the variance-optimal martingale measure $\tilde{P}$ from a nice solution of the BSDE (6).

Proposition 4: Suppose that $I\!\!M_e^2 \neq \emptyset$, and take any solution (a, Λ^a, N^a) of (6) satisfying (12) and (14). Then, the density process of the variance-optimal martingale measure $\tilde{P}$ is given by

$$Z^{\tilde{P}} = \mathcal{E}\left(-\int\lambda\,dM + \int\frac{1}{a_-}\,dN^a\right). \tag{19}$$

Proof: Since $I\!\!M_e^2 \neq \emptyset$, we know from Theorem 3 that at least one solution with the above properties exists, and that $Z(a) = \mathcal{E}(-\int\lambda\,dM + \int(1/a_-)\,dN^a)$ is for any solution with (12) the density process of some $Q(a) \in I\!\!M_e^2$. So, we have to show that $Q(a) = \tilde{P}$, i.e., $Q(a)$ is variance-optimal, if a has the extra property (14). We claim that for any $Q \in I\!\!M_e^2$ with density process $Z^Q = \mathcal{E}(-\int\lambda\,dM + N^Q)$, Itô's formula yields that $(1/a)(Z^Q)^2$ is a local P-submartingale, and a local P-martingale for $Q = Q(a)$. Thanks to the integrability condition (14), we then actually have (true) P-submartingales and a (true) P-martingale and, thus, since $a_T = 1$

$$E\left[\left(\frac{dQ}{dP}\right)^2\right] = E\left[\frac{1}{a_T}\left(Z_T^Q\right)^2\right] \geq E\left[\frac{1}{a_0}\right]$$

for any $Q \in I\!\!M_e^2$, with equality for $Q = Q(a)$. Hence $Q(a)$ is indeed variance-optimal.

It remains to justify the above claim, and this is mainly a matter of diligence in using Itô's formula; the only complication stems from the fact that N^a and N^Q may be discontinuous. Because $Z^Q = \mathcal{E}(-\int\lambda\,dM + N^Q)$ with $N^Q \perp M$, we get

$$\begin{aligned} d((Z^Q)^2) = (Z_-^Q)^2(&-2\lambda\,dM + 2\,dN^Q \\ &+ \lambda^\top\,d\langle M\rangle\,\lambda + d[N^Q]) \end{aligned} \tag{20}$$

and Itô's formula gives

$$d\left(\frac{1}{a}\right) = -\frac{1}{a_-^2}\,da + \frac{1}{a_-^3}\,d\langle a^c\rangle + \Delta\left(\frac{1}{a}\right) + \frac{1}{a_-^2}\Delta a \quad (21)$$

where we write as usual $\Delta U_t = U_t - U_{t-}$ for the jump at time t of an RCLL process U. To compute $[(1/a), (Z^Q)^2]$, we now use (20), (21), and the BSDE (6) together with $d[a, [N^Q]] = \Delta a\,\Delta[N^Q], d[a, M] = d\langle M\rangle\,\Lambda^a$ and

$$\begin{aligned} d[a, N^Q] &= d[N^a, N^Q] \\ &= d\langle (N^a)^c, (N^Q)^c\rangle + \Delta N^a \Delta N^Q \end{aligned}$$

to obtain

$$\begin{aligned} &d\left[\frac{1}{a}, (Z^Q)^2\right] \\ &\quad = (Z_-^Q)^2 \left(\frac{2}{a_-^2}\lambda^\top\, d\langle M\rangle \Lambda^a - \frac{2}{a_-^2} d\langle (N^a)^c, (N^Q)^c\rangle \right. \\ &\qquad \left. + 2\Delta\left(\frac{1}{a}\right)\Delta N^Q + \Delta\left(\frac{1}{a}\right)\Delta[N^Q]\right). \end{aligned} \quad (22)$$

Note that we also have used

$$\Delta N^a = \Delta a \quad (23)$$

which follows immediately from the BSDE (6) since S is continuous. Now, we apply the product rule to $(1/a)(Z^Q)^2$, use (20)–(22) and plug in from the BSDE (6). Then, we use $d[N^Q] = d\langle (N^Q)^c\rangle + (\Delta N^Q)^2$ and $\Delta[N^Q] = (\Delta N^Q)^2$ and, after collecting terms and simplifying, end up with

$$\begin{aligned} \frac{1}{a}(Z^Q)^2 &= \text{local } P - \text{martingale} \\ &\quad + \int (Z_-^Q)^2 \frac{1}{a}\, d\left\langle \left(N^Q - \int \frac{1}{a_-}\, dN^a\right)^c \right\rangle \\ &\quad + \sum \left(\frac{1}{a_-}\left(\Delta N^Q\right)^2 \right. \\ &\qquad \left. + \Delta\left(\frac{1}{a}\right)(1 + \Delta N^Q)^2 + \frac{1}{a_-^2}\Delta a\right). \end{aligned}$$

Finally, an elementary computation shows that the last sum is equal to

$$\sum \frac{(a_-(1 + \Delta N^Q) - a)^2}{a a_-^2}.$$

Because a is strictly positive, we therefore see that $(1/a)(Z^Q)^2$ is indeed always a local P-submartingale. For $Q = Q(a)$, we have $N^Q = \int (1/a_-)\, dN^a$, hence $\Delta N^Q = (1/a_-)\Delta a = (a/a_-) - 1$ by (23), and plugging in shows that $(1/a)(Z^Q)^2$ is then a local P-martingale since both drift terms vanish. This proves our claim and completes the proof. q.e.d.

We now use Proposition 4 to give a uniqueness result for the BSDE (6).

Proposition 5: Suppose that $I\!M_e^2 \neq \emptyset$. Then (6) has a unique solution (a, Λ^a, N^a) satisfying (12) and (14).

Proof: If we have two such solutions (a^i, Λ^i, N^i) for $i = 1, 2$, Proposition 4 tells us that

$$Z(a^i) = \mathcal{E}\left(-\int \lambda\, dM + \int \frac{1}{a_-^i}\, dN^i\right) = Z^{\tilde{P}} \quad \text{for } i = 1, 2. \quad (24)$$

Moreover, the proof of Proposition 4 shows that $(1/a^i)(Z^{\tilde{P}})^2$ is a P-martingale for $i = 1, 2$ and since $a_T^i = 1$, this implies that $a^1 = a^2$. Because (24) yields $\int (1/a_-^1)\, dN^1 = \int (1/a_-^2)\, dN^2$, we conclude that $N^1 = N^2$ as well. If we use the BSDE (6) to write down the dynamics of $a^1 - a^2$ and then use $a^1 = a^2$ and $N^1 = N^2$, we obtain that the zero process has a canonical decomposition with local P-martingale part $\int (\Lambda^1 - \Lambda^2)\, dM$, and this implies that $\Lambda^1 = \Lambda^2$ (in the space $L^2_{\text{loc}}(M)$, to be precise). q.e.d.

Remark: Results similar to Propositions 4 and 5 can also be found in [14] or [15], but only under the additional assumption that $I\!F$ is continuous or that the minimal martingale measure $\hat{P}$ (the EMM associated to $N \equiv 0$ in (11)) exists and satisfies the reverse Hölder inequality $R_2(P)$. Since we do no need this here, explanation of the latter is postponed for the moment. In the framework of the usual Itô process example, (19) can also be obtained by combining [8, eqs. (134), (125), and (117)]. $\diamond$

The preceding results need no specific properties of the strategies ϑ we consider and so the choice of Θ has not been relevant so far. This is so because (as already remarked and as evident from the proofs) a can be expressed directly in terms of the dual of the MVH problem. For b and c, this is no longer the case, and we have to make a precise choice for Θ. We have decided to work here with the same space of strategies as in [3] and [19].

Definition: The space Θ consists of all $I\!R^d$-valued predictable S-integrable processes ϑ such that the stochastic integral $G(\vartheta) = \int \vartheta\, dS$ is in the space $\mathcal{S}^2(P)$ of semimartingales. Equivalently, $\vartheta \in \Theta$ if and only if ϑ is predictable and

$$\begin{aligned} &E\left[\int_0^T \vartheta_u^\top\, d\langle M\rangle_u\, \vartheta_u + \int_0^T \left|\vartheta_u^\top\, d\langle M\rangle_u\, \lambda_u\right|\right] \\ &\qquad = E\left[\left\langle \int \vartheta\, dM\right\rangle_T + \int_0^T \left|\vartheta_u^\top\, dA_u\right|\right] < \infty. \end{aligned}$$

In [3] and [19], this is also expressed by saying that $\vartheta \in L^2(M) \cap L^2(A)$. It is easy to check for this space Θ that for fixed t and ϑ, the family $\{\Gamma_t(\psi) \,|\, \psi \in {}^t\Theta(\vartheta)\}$ is stable under taking minima. Hence, the assumption of Proposition 1 is satisfied.

For the MVH problem to be solvable for arbitrary $H \in L^2(\mathcal{F}_T, P)$, we want $G_T(\Theta)$ to be closed in $L^2(P)$. It has been shown in [3] that this is here (where S is continuous) equivalent to the condition that the variance-optimal martingale measure $\tilde{P}$ satisfies the reverse Hölder inequality $R_2(P)$; see [3, Th. 4.1]. The condition $R_2(P)$ is defined as follows.

Definition: We say that $Q \in I\!M_e^2$ satisfies the *reverse Hölder inequality* $R_2(P)$ if there is a constant $C \in (0, \infty)$ such that the density process Z^Q satisfies

$$E_P\left[\left(Z_T^Q\right)^2 \,\middle|\, \mathcal{F}_t\right] \le C\left(Z_t^Q\right)^2 \quad P-\text{a.s.}, \qquad 0 \le t \le T.$$

Our next result shows that $R_2(P)$ for $\tilde{P}$ can be directly translated into an additional property of a. This again sharpens previous results in [14] and [15].

Theorem 6: The following statements are equivalent.

1) $\mathbb{M}_e^2 \neq \emptyset$ and $\tilde{P}$ satisfies $R_2(P)$.
2) $\mathbb{M}_e^2 \neq \emptyset$ and some $Q \in \mathbb{M}_e^2$ satisfies $R_2(P)$.
3) The BSDE (6) has a solution (a, Λ^a, N^a) satisfying (12) and in addition that

$$a \text{ is bounded from below uniformly in } t, \omega \; (P - \text{ a.s., to be precise}). \quad (25)$$

Proof: The equivalence of 1) and 2) is directly from [3, Th. 4.1]. If 1) holds, the proof of Theorem 3 shows that we can construct a solution to (6) with

$$\begin{aligned}\frac{1}{a_t} &= \frac{\tilde{Z}_t}{Z_t^{\tilde{P}}} = \frac{1}{Z_t^{\tilde{P}}} E_{\tilde{P}}\left[Z_T^{\tilde{P}} \,\middle|\, \mathcal{F}_t \right] \\ &= \frac{1}{\left(Z_t^{\tilde{P}}\right)^2} E\left[\left(Z_T^{\tilde{P}}\right)^2 \,\middle|\, \mathcal{F}_t \right] \leq C \quad P - \text{ a.s.,} \\ & \qquad 0 \leq t \leq T\end{aligned}$$

where we have used the Bayes formula and $R_2(P)$. This yields 3). Conversely, suppose that 3) holds and use Theorem 3 to choose a solution (a, Λ^a, N^a) such that we also have (12). Then $(1/a)(Z^{\tilde{P}})^2$ is a P-martingale by the proof of Proposition 4 and since $a_T = 1$, we get

$$E\left[\left(Z_T^{\tilde{P}}\right)^2 \middle| \mathcal{F}_t \right] = \frac{1}{a_t} \left(Z_t^{\tilde{P}}\right)^2 \leq C \left(Z_t^{\tilde{P}}\right)^2 \quad P - \text{a.s.,} \quad 0 \leq t \leq T$$

due to (25). Thus, 1) holds and the proof is complete. q.e.d.

IV. BSDEs FOR b AND c

In this section, we prove existence and uniqueness for the solutions (b, μ^b, ψ^b, N^b) and $(c, N^{(c)})$ of (7) and (8) for arbitrary $H \in L^2(\mathcal{F}_T, P)$, assuming that we have already solved (6) for a. This will be done under the additional assumptions that the filtration $\mathbb{F}$ is continuous and that $\tilde{P}$ (or, equivalently by Theorem 6, some $Q \in \mathbb{M}_e^2$) satisfies the reverse Hölder inequality $R_2(P)$. If $\mathbb{F}$ is continuous, then a, b, c are actually all continuous. But for ease of comparison with future extensions, we shall still write a_- and b_- where this is appropriate.

Let us first recall the BSDE (7) for b. We fix $H \in L^2(\mathcal{F}_T, P)$ and a solution triple (a, Λ^a, N^a) of the BSDE (6) for a. Then, we consider (7). A *solution* of (7) is a quadruple (b, μ^b, ψ^b, N^b) satisfying (7) and such that b is an RCLL semimartingale, μ^b is an $\mathbb{R}^d$-valued predictable S-integrable process, ψ^b is a real-valued predictable N^a-integrable process and N^b is a local P-martingale null at 0 and P-orthogonal to both M and N^a.

Theorem 7: Suppose that $\mathbb{M}_e^2 \neq \emptyset$. Take the solution (a, Λ^a, N^a) of (6) satisfying (12) and (14). If the filtration $\mathbb{F}$ is continuous, the BSDE (7) for b then has a solution (b, μ^b, ψ^b, N^b).

Proof: As usual, we denote by $\tilde{P}$ the variance-optimal martingale measure.

1) Because H is in $L^2(P)$, the process

$$b_t := E_{\tilde{P}}[H \,|\, \mathcal{F}_t], \qquad 0 \leq t \leq T \qquad (26)$$

is well-defined, can be chosen RCLL, is a semimartingale and a $\tilde{P}$-martingale and has $b_T = H$. Since S is continuous, b has under $\tilde{P}$ a Galtchouk–Kunita–Watanabe decomposition

$$b = b_0 + \int \mu^b \, dS + L^b$$

where μ^b is $\mathbb{R}^d$-valued, predictable and S-integrable, and L^b is in $M_{0,\mathrm{loc}}(\tilde{P})$ and $\tilde{P}$-orthogonal to S. [If $\tilde{P}$ satisfies in addition the reverse Hölder inequality $R_2(P)$, [19, Th. 3] even implies that $\mu^b \in L^2(M, P)$ and $\sup_{0 \leq t \leq T} |L_t^b| \in L^2(P)$, i.e., $L^b \in \mathcal{R}^2(P)$.] However, (7) prescribes the behavior of b under P, not $\tilde{P}$, and so we need to understand how the local $\tilde{P}$-martingale L^b behaves under P.

2) By Theorem 3 and Propostion 5, (6) has a unique solution (a, Λ^a, N^a) with (12) and (14). Then $\int (1/a_-) \, dN^a$ is in $\mathcal{M}_{0,\mathrm{loc}}^2(P)$ and the density process $Z^{\tilde{P}} \in \mathcal{M}^2(P)$ is given by

$$Z^{\tilde{P}} = \mathcal{E}\left(-\int \lambda \, dM + \int \frac{1}{a_-} \, dN^a \right)$$

according to (19) in Proposition 4. Hence, $\langle Z^{\tilde{P}}, \int (1/a_-) \, dN^a \rangle = \int Z_-^{\tilde{P}} (1/a_-^2) \, d\langle N^a \rangle$ exists and Girsanov's theorem tells us that

$$\begin{aligned}R &:= \int \frac{1}{a_-} \, dN^a - \int \frac{1}{Z_-^{\tilde{P}}} \, d\left\langle Z^{\tilde{P}}, \int \frac{1}{a_-} \, dN^a \right\rangle \\ &= \int \frac{1}{a_-} \, dN^a - \int \frac{1}{a_-^2} \, d\langle N^a \rangle \qquad (27)\end{aligned}$$

is a local $\tilde{P}$-martingale. Because $\mathbb{F}$ is continuous, R is continuous and so L^b has under $\tilde{P}$ a Galtchouk–Kunita–Watanabe decomposition as

$$L^b = \int \psi^b \, dR + N^b$$

where ψ^b is predictable and R-integrable and $N^b \in \mathcal{M}_{0,\mathrm{loc}}(\tilde{P})$ is $\tilde{P}$-orthogonal to R. However, due to (27) and since N^a is continuous, ψ^b is also N^a-integrable, and we obtain

$$db_t = \mu_t^b \, dS_t + \psi_t^b \, dR_t + dN_t^b, \qquad b_T = H$$

which is (7) due to the form of R. [If $\tilde{P}$ satisfies in addition $R_2(P)$, we know from 1) above that $L_T^b \in L^2(P)$, and [19, Th. 3] then implies that $(1/a_-)\psi^b \in L^2(N^a, P)$ and $N^b \in \mathcal{R}^2(P)$.]

3) To show that N^b is in $\mathcal{M}_{0,\mathrm{loc}}(P)$ and P-orthogonal to both M and N^a, we use the assumption that $\mathbb{F}$ is continuous. Because all local martingales (under any $Q \approx P$) are then continuous, orthogonality simply means that the quadratic covariation process must vanish, and since this can be computed pathwise, orthogonality is the same under any $Q \approx P$. So, $N^b \perp R$ implies that $\langle N^b, N^a \rangle = 0$, $L^b \perp S$ implies that $\langle L^b, M \rangle = 0$ and $N^a \perp M$ implies that $\langle R, M \rangle = 0$. Hence we also get $\langle N^b, M \rangle = 0$ and it only remains to show that the local $\tilde{P}$-martingale

N^b is also a local P-martingale. However, (19) implies that $\langle Z^{\tilde P}, N^b\rangle = 0$, hence also $\langle (1/Z^{\tilde P}), N^b\rangle = 0$ since all local martingales are continuous, and so $(1/Z^{\tilde P})N^b \in \mathcal{M}_{0,\mathrm{loc}}(\tilde P)$ and N^b is indeed in $\mathcal{M}_{0,\mathrm{loc}}(P)$. This completes the proof. q.e.d.

Proposition 8: Consider the situation of Theorem 7 and assume in addition that $\tilde P$ satisfies the reverse Hölder inequality $R_2(P)$. Then (7) has a unique solution (b, μ^b, ψ^b, N^b) satisfying

$$\mu^b \in L^2(M), \frac{1}{a_-}\psi^b \in L^2(N^a), \qquad N^b \in \mathcal{R}^2(P). \quad (28)$$

Moreover, b is then in $\mathcal{R}^2(P)$.

Proof: That (7) does have a solution satisfying (28) under the given assumptions has already been argued in the proof of Theorem 7. To prove uniqueness, take two solutions of (7) and denote their difference by (δ, μ, ψ, N). Then (7) gives

$$\begin{aligned}\delta &= \delta_0 + \int \mu\, dS + \int \frac{1}{a_-}\psi\, dN^a - \int \frac{1}{a_-^2}\psi\, d\langle N^a\rangle + N\\ &= \delta_0 + \int \mu\, dS + \int \psi\, dR + N\end{aligned}$$

with R as in (27), and of course $\delta_T = 0$. We want to argue that $\int \mu\, dS, \int \psi\, dR$, and N are all $\tilde P$-martingales and pairwise $\tilde P$-orthogonal; then they as well as δ must all vanish and this will prove uniqueness.

It is clear that $\int \mu\, dS$ is a local $\tilde P$-martingale since $\tilde P \in I\!\!M^2_e$. Moreover, using the reverse Hölder inequality $R_2(P)$ (via [19, Prop. 1]) and the fact that $\mu \in L^2(M)$ due to (28) yields

$$\begin{aligned}&E_{\tilde P}\left[\sup_{0\le t\le T}\left|\int_0^t \mu_u\, dS_u\right|\right]\\ &\le \left\|\frac{d\tilde P}{dP}\right\|_{L^2(P)} \left\|\sup_{0\le t\le T}\left|\int_0^t \mu_u\, dS_u\right|\right\|_{L^2(P)}\\ &\le C\left\|\frac{d\tilde P}{dP}\right\|_{L^2(P)} \|\mu\|_{L^2(M)} < \infty \qquad (29)\end{aligned}$$

so that $\int \mu\, dS$ is a true $\tilde P$-martingale. The proof of Theorem 7 shows that $\int \psi\, dR$ is a local $\tilde P$-martingale, and an analogous argument using $R_2(P)$ and (28) as above implies that $\int \psi\, dR$ is also a true $\tilde P$-martingale. Next, $N \perp M$ and $N \perp N^a$ imply due to the representation (19) of $Z^{\tilde P}$ that N is a local $\tilde P$-martingale since $NZ^{\tilde P}$ is in $\mathcal{M}_{0,\mathrm{loc}}(P)$. Because N is in $\mathcal{R}^2(P)$ due to (28), it is even a true $\tilde P$-martingale. Finally, the pairwise orthogonality of $\int \mu\, dS, \int \psi\, dR$ and N follows as in the proof of Theorem 7 from the assumption that $I\!\!F$ is continuous.

Once we know that b is unique, the proof of Theorem 7 yields $b = b_0 + \int \mu^b\, dS + L^b$ with $L^b \in \mathcal{R}^2(P)$ and $\mu^b \in L^2(M)$. Hence, the estimate (29) shows that $\int \mu^b\, dS$ is in $\mathcal{R}^2(P)$, and then so is b. q.e.d.

Remarks:

1) One could argue with some justification that the above results look slightly artificial because they exploit the already known general structure of the solution to the MVH problem—known, that is, from martingale and projection techniques as opposed to stochastic control and BSDEs. But with some thought, the above solution can also be guessed. In fact, the BSDE (7) is linear and it is well known that a change of measure is then often helpful in finding a solution. So, if we introduce Z (as a density process for a change of measure) via $dZ = Z_-(-\lambda\, dM + (1/a_-)\, dN)$ with $N \perp M$, we get

$$d\left\langle \int \frac{1}{Z_-}\, dZ, \int \frac{\psi^b}{a_-}\, dN\right\rangle = \frac{\psi^b}{a_-^2}\, d\langle N\rangle.$$

Because we want to have

$$db = \mu^b\, dS + \frac{\psi^b}{a_-}\, dN^a - \frac{\psi^b}{a_-^2}\, d\langle N^a\rangle + dN^b$$

for (7), we should thus choose Z such that ZS becomes a local P-martingale (which we have already done with the above structure of Z) and such that $N = N^a$. But then we end up with $Z = Z^{\tilde P}$, which tells us that we should try to obtain b as the $\tilde P$-martingale with final value $b_T = H$—and this is how the proof of Theorem 7 starts.

2) The results in [19] are stated under the assumption that $\mathcal{F}_0$ is trivial. But we can still use them here for general $\mathcal{F}_0$ because the triviality of $\mathcal{F}_0$ is not used except for getting an initial value zero for some martingales.

Example: For our standard example of an Itô process model, the standard assumptions guarantee that the minimal martingale measure $\hat P$ associated to $N \equiv 0$ in (11) is in $I\!\!M^2_e$ and satisfies the reverse Hölder inequality $R_2(P)$. We already know that the BSDE (6) for a takes the form

$$\begin{aligned}da_t = \left(|\varphi_t|^2 a_t + 2\varphi_t^\top Y_t^a + \frac{1}{a_t}\,(Y_t^a)^\top \Pi_t Y_t^a\right)\, dt&\\ + Y_t^a\, dW_t, \qquad a_T = 1&\end{aligned}$$

where $\Pi = \sigma^\top(\sigma\sigma^\top)^{-1}\sigma$ denotes the projection on the range of $\sigma^\top$. This is an equation for the pair (a, Y^a); we also know that

$$N^a = \int \nu^a\, dW = \int ((I - \Pi)Y^a)\, dW \qquad (30)$$

and we see that a is continuous.

By plugging in the dynamics of S and using (30), we can first rewrite (7) as

$$\begin{aligned}db_t = &\left((\mu_t - r_t\underline{1})^\top \mathrm{diag}(S_t)\mu_t^b - \frac{\psi_t^b}{a_t^2}\,|(I - \Pi_t)Y_t^a|^2\right)\, dt\\ &+ \left(\sigma_t^\top \mathrm{diag}(S_t)\mu_t^b + \frac{\psi_t^b}{a_t}(I - \Pi_t)Y_t^a + \nu_t^b\right)\, dW_t \quad (31)\end{aligned}$$

for $(b, \mu^b, \psi^b, \nu^b)$, where ν_t^b should be in $\ker(\sigma_t)$ P-a.s. and satisfy $(\nu_t^b)^\top \nu_t^a = 0$ P-a.s. for all t. This uses that $N^b = \int \nu^b\, dW$ by Itô's representation theorem in $I\!\!F = I\!\!F^W$ and expresses the conditions $N^b \perp M$ and $N^b \perp N^a$.

If we now set $Y^b := \ell^b + (\psi^b/a)(I - \Pi)Y^a + \nu^b$ with $\ell^b := \sigma^\top \mathrm{diag}(S)\mu^b$, we obtain that ℓ_t^b is the projection of Y_t^b

on range$(\sigma_t^\top)$ P-a.s. for all t and thus $\ell^b = \Pi Y^b$. Using that ℓ^b and ν^b are both orthogonal to $\nu^a = (I-\Pi)Y^a$ next yields

$$(\nu^a)^\top Y^b = \frac{\psi^b}{a}|\nu^a|^2.$$

For ψ^b, plugging this into the definition of Y^b and solving for ν^b then gives

$$\nu^b = (I-\Pi)Y^b - \frac{(Y^b)^\top (I-\Pi)Y^a}{|(I-\Pi)Y^a|^2}(I-\Pi)Y^a.$$

Finally, using the definition of the market price of risk φ to write

$$(\mu - r\underline{1})^\top \mathrm{diag}(S)\mu^b = \varphi^\top \Pi Y^b$$

leads from (31) to

$$db_t = \left(Y_t^b\right)^\top \left(\Pi_t\varphi_t - \frac{1}{a_t}(I-\Pi_t)Y_t^a\right) dt + Y_t^b\, dW_t, \qquad b_T = H \quad (32)$$

for the pair (b, Y^b). This coincides (except for the interest rate which is 0 in our discounted situation) with [8, eq. (124)], in view of their equations (125), (117), and (114). $\square$

Let us next recall the BSDE (8) for c. We fix a solution (a, Λ^a, N^a) of (6), some $H \in L^2(\mathcal{F}_T, P)$, a solution (b, μ^b, ψ^b, N^b) of the corresponding equation (7) and consider (8). A *solution* of (8) is a pair $(c, N^{(c)})$ satisfying (8) and such that c is an RCLL semimartingale and $N^{(c)}$ is a local P-martingale. (We deliberately add here parentheses to the superscript c to avoid confusion with the standard notation N^c for the continuous local martingale part of a semimartingale N.)

Theorem 9: Suppose that $\mathbb{M}_e^2 \neq \emptyset$. Take the solution (a, Λ^a, N^a) of (6) satisfying (12) and (14). Assume that the filtration $\mathbb{F}$ is continuous and that $\tilde{P}$ satisfies the reverse Hölder inequality $R_2(P)$. For $H \in L^2(\mathcal{F}_T, P)$, take the solution (b, μ^b, ψ^b, N^b) of (7) satisfying (28). Then the BSDE (8) for c has a unique solution $(c, N^{(c)})$ with the property that

$$N^{(c)} \text{ is not only a local, but a true } P - \text{ martingale.} \quad (33)$$

Moreover, c is then of class (D) under P.

Proof: That we can choose (a, Λ^a, N^a) and (b, μ^b, ψ^b, N^b) as desired follows from Propositions 5 and 8. Moreover, the proof of Theorem 3 shows that $a \leq 1$. Hence, we obtain

$$E\left[\int_0^T \frac{(\psi_s^b)^2}{a_{s-}}\, d\langle N^a\rangle_s\right] \leq E\left[\int_0^T \left(\frac{\psi_s^b}{a_{s-}}\right)^2 d\langle N^a\rangle_s\right] < \infty \quad (34)$$

since ψ^b/a_- is in $L^2(N^a)$ by (28), and

$$E\left[\int_0^T a_{s-}\, d\langle N^b\rangle_s\right] \leq E\left[\langle N^b\rangle_T\right] \leq CE\left[\sup_{0\leq t\leq T}\left|N_t^b\right|^2\right] < \infty \quad (35)$$

by the reverse Hölder inequality $R_2(P)$ via [19, Prop. 1] and by the fact that $N^b \in \mathcal{R}^2(P)$ by (28). So the P-martingale

$$N_t^{(c)} := E\left[\left.\int_0^T \frac{(\psi_s^b)^2}{a_{s-}}\, d\langle N^a\rangle_s + \int_0^T a_{s-}\, d\langle N^b\rangle_s \right| \mathcal{F}_t\right], \qquad 0 \leq t \leq T$$

is well-defined, and choosing an RCLL version and setting

$$\begin{aligned} c_t &:= -\int_0^t \frac{\left(\psi_s^b\right)^2}{a_{s-}}\, d\langle N^a\rangle_s - \int_0^t a_{s-}\, d\langle N^b\rangle_s + N_t^{(c)} \\ &= E\left[\left.\int_t^T \frac{(\psi_s^b)^2}{a_{s-}}\, d\langle N^a\rangle_s + \int_t^T a_{s-}\, d\langle N^b\rangle_s \right| \mathcal{F}_t\right], \qquad 0 \leq t \leq T \quad (36) \end{aligned}$$

clearly yields a solution $(c, N^{(c)})$ of (8) satisfying (33). That this c is of class (D) follows immediately from (36) together with (34) and (35).

For any two solutions satisfying (33), (8) implies that $c^1 - c^2 = c_0^1 - c_0^2 + N^{(c^1)} - N^{(c^2)}$ is a P-martingale with final value 0; hence we have $c^1 \equiv c^2$ which implies that $N^{(c^1)} \equiv N^{(c^2)}$ as well, proving uniqueness. q.e.d.

Example: For our standard example of an Itô process model, we obtain by straightforward calculations and by using the intermediate results from the derivation of the BSDE (32) that the general BSDE (8) reduces to the equation

$$dc_t = -a_t\left|(I-\Pi_t)Y_t^b\right|^2 dt + Y_t^{(c)}\, dW_t, \qquad c_T = 0$$

for the pair $(c, Y^{(c)})$. $\square$

V. Solving the MVH Problem With the Help of BSDES

In this section, we show how to construct the solution to the MVH problems (3) and (4) explicitly from the solutions of the BSDEs (6)–(8) for a, b, c. The basic idea for this is both well-known and very simple. The martingale optimality principle in Proposition 1 tells us to look for a strategy ϑ^* such that $J(\vartheta^*)$ becomes a martingale. Since we guess that $J(\vartheta^*)$ is a quadratic function of V^{ϑ^*} and know from the BSDEs (6)–(8) how its coefficients a, b, c should behave, it is natural to start by computing this quadratic functional more explicitly. For that purpose, we first fix $v_0 \in \mathbb{R}$ and define for any $\vartheta \in \Theta$ the process

$$\begin{aligned} j_t(\vartheta) &:= a_t\left(v_0 + G_t(\vartheta) - b_t\right)^2 + c_t \\ &= a_t\left(V_t^\vartheta - b_t\right)^2 + c_t, \qquad 0 \leq t \leq T. \end{aligned}$$

Proposition 10: Suppose that the filtration $\mathbb{F}$ is continuous. Fix $H \in L^2(\mathcal{F}_T, P)$ and $v_0 \in \mathbb{R}$. If $(a, \Lambda^a, N^a), (b, \mu^b, \psi^b, N^b), (c, N^{(c)})$ are solutions of the

BSDE's (6)–(8), then $j(\vartheta)$ is a local P-submartingale for any $\vartheta \in \Theta$, and a local P-martingale if and only if ϑ satisfies

$$\vartheta = \mu^b - \left(\lambda + \frac{\Lambda^a}{a_-}\right)(v_0 + G(\vartheta) - b_-). \tag{37}$$

Proof: A straightforward but lengthy application of Itô's formula gives

$$\begin{aligned} dj_t(\vartheta) = &\cdots dM_t + \cdots dN_t^a + \cdots dN_t^b + dN_t^{(c)} \\ &+ a_{t-}\left((\vartheta_t - \mu_t^b) + \left(\lambda_t + \frac{\Lambda_t^a}{a_{t-}}\right)\right. \\ &\times \left.(v_0 + G_t(\vartheta) - b_{t-})\right)^\top d\langle M\rangle_t \\ &\times \left((\vartheta_t - \mu_t^b) + \left(\lambda_t + \frac{\Lambda_t^a}{a_{t-}}\right)\right. \\ &\times \left.(v_0 + G_t(\vartheta) - b_{t-})\right) \end{aligned}$$

and the assertion follows. We spare the reader the details of the computation and mention only that the assumption of a continuous filtration is used to replace all square brackets by sharp brackets, which in turn allows one to exploit the orthogonality relations between M, N^a, and N^b. q.e.d.

To obtain a local P-martingale for $j(\vartheta^*)$, we have to find a solution ϑ^* to (37). The next result is the first step in that direction.

Lemma 11: Denote by $L(S)$ the set of all $\mathbb{R}^d$-valued predictable S-integrable processes. For any $\mu, \gamma \in L(S)$ and any semimartingale Y, the equation

$$\vartheta = \mu - \gamma(G(\vartheta) - Y_-) \tag{38}$$

then has a unique solution ϑ^b in $L(S)$.

Proof: Due to [18, Th. V.7], the equation

$$dU = -U_-\gamma\, dS + (\mu + \gamma Y_-)\, dS, \qquad U_0 = 0 \tag{39}$$

has a unique solution $\bar{U}$ in the class of continuous semimartingales because S is continuous. The process $\bar{\vartheta} := \mu - \gamma(\bar{U} - Y_-)$ is then in $L(S)$, and

$$G(\bar{\vartheta}) = \int \bar{\vartheta}\, dS = -\int \bar{U}\gamma\, dS + \int (\mu + \gamma Y_-)\, dS = \bar{U}$$

shows that $\bar{\vartheta}$ by its definition does satisfy (38). For any ϑ satisfying (38), $G(\vartheta)$ satisfies (39); hence $G(\vartheta) = \bar{U}$ by uniqueness and so $\vartheta = \bar{\vartheta}$. q.e.d.

If we choose in Lemma 11 $\mu = \mu^b, \gamma = \lambda + (\Lambda^a/a_-)$, and $Y = b - v_0$, we obtain in particular the existence of a process ϑ^* in $L(S)$ satisfying (37). Our next goal is to prove that (v_0, ϑ^*) is a bona fide strategy by showing that ϑ^* is in Θ.

Proposition 12: Suppose that $\mathbb{M}_e^2 \neq \emptyset, \mathbb{F}$ is continuous and $\tilde{P}$ satisfies $R_2(P)$. Fix $H \in L^2(\mathcal{F}_T, P)$ and choose the solutions $(a, \Lambda^a, N^a), (b, \mu^b, \psi^b, N^b), (c, N^{(c)})$ of (6)–(8) as in Theorem 9. For any $v_0 \in \mathbb{R}$, (37) then has a unique solution ϑ^*, and $\vartheta^* \in \Theta$ if b_0 is bounded. (This last condition is automatically satisfied if $\mathcal{F}_0$ is trivial.)

Proof: Lemma 11 implies that (37) has a unique solution ϑ^* in $L(S)$. To show that ϑ^* is in fact in $\Theta \subseteq L(S)$, we prove that

$$G(\vartheta^*) \text{ is in } \mathcal{R}^2(P). \tag{40}$$

Using $R_2(P)$ via [19, Prop. 1] then yields first that ϑ^* is in $L^2(M)$, because $[S] = [M] = \langle M\rangle$ by the continuity of S, and Theorem 2 of [19] tells us that $L^2(M) = \Theta$, again due to $R_2(P)$. To prove (40), we introduce the process

$$L := a(v_0 + G(\vartheta^*) - b) \tag{41}$$

and claim that

$$L \text{ is in } \mathcal{R}^2(P). \tag{42}$$

Then, we solve for $G(\vartheta^*)$ to get

$$G(\vartheta^*) = \frac{L}{a} - v_0 + b$$

and deduce (40) from (42), the fact that $(1/a)$ is bounded due to $R_2(P)$ according to Theorem 6, and the fact that b is in $\mathcal{R}^2(P)$ by Proposition 8.

To prove (42), we apply the product rule to the definition (41) of L, plug in (6)–(8) and (37) and simplify to obtain after straightforward but lengthy calculations the SDE

$$\begin{aligned} dL_t = L_{t-}\left(-\lambda_t\, dM_t + \frac{1}{a_{t-}}\, dN_t^a\right) - \psi_t^b\, dN_t^a& \\ - a_{t-}\, dN_t^b, \qquad L_0 = a_0(v_0 - b_0)& \end{aligned}$$

for L. This is a linear SDE whose solution is given explicitly in [18, Th. V.52]. If we introduce the processes

$$\bar{L} := -\int \psi^b\, dN^a - \int a_-\, dN^b$$

and

$$\mathcal{E}(\bar{H}) := \mathcal{E}\left(-\int \lambda\, dM + \int \frac{1}{a_-}\, dN^a\right) = Z^{\tilde{P}}$$

by (19) in Proposition 4, we obtain

$$\begin{aligned} L &= Z^{\tilde{P}}\left(L_0 + \int \frac{1}{Z_-^{\tilde{P}}}\, d(\bar{L} - [\bar{L}, \bar{H}])\right) \\ &= Z^{\tilde{P}} L_0 - Z^{\tilde{P}} \int \frac{1}{Z_-^{\tilde{P}}}\left(\psi^b\, dN^a + a_-\, dN^b\right. \\ &\quad \left. - \frac{\psi^b}{a_-}\, d\langle N^a\rangle\right) \\ &= Z^{\tilde{P}} L_0 - Z^{\tilde{P}} \int \frac{a_-}{Z_-^{\tilde{P}}}\left(dN^b + \frac{\psi^b}{a_-}\, dN^a\right. \\ &\quad \left. - \frac{\psi^b}{a_-^2}\, d\langle N^a\rangle\right) \\ &= Z^{\tilde{P}} L_0 - Z^{\tilde{P}} \int \frac{a_-}{Z_-^{\tilde{P}}}\, dL^b \end{aligned}$$

with

$$L^b = N^b + \int \frac{\psi^b}{a_-}\, dN^a - \int \frac{\psi^b}{a_-^2}\, d\langle N^a\rangle$$

$$= N^b + \int \psi^b \, dR$$

as in the proof of Theorem 7. Since $\tilde{P} \in I\!M_e^2, Z^{\tilde{P}}$ is in $\mathcal{M}^2(P) \subseteq \mathcal{R}^2(P)$, and $L_0 = a_0(v_0 - b_0)$ is bounded since $a_0 \leq 1, v_0$ is deterministic and b_0 is bounded. So it remains to show that

$$Z^{\tilde{P}} \int \frac{a_-}{Z_-^{\tilde{P}}} \, dL^b \text{ is in } \mathcal{R}^2(P). \tag{43}$$

But this is a consequence of [19, Lemma 7] or more precisely of the proof given there. To see this, note that L^b is a $\tilde{P}$-martingale and in $\mathcal{R}^2(P)$ since $\tilde{P}$ satisfies $R_2(P)$, as argued in step 1) of the proof of Theorem 7. This implies that $[L^b]_T \in L^1(P)$ by [19, Prop. 1], due to $R_2(P)$. Because a is bounded according to Theorem 3, $\tilde{L} := \int a_- \, dL^b$ is also a $\tilde{P}$-martingale and satisfies $[\tilde{L}]_T \in L^1(P)$. Moreover, $L^b \perp S$ under $\tilde{P}$ implies that $\langle \tilde{L}, M \rangle = 0$ since $I\!F$ is continuous and so the proof of [19, Lemma 7] yields

$$Z^{\tilde{P}} \int \frac{1}{Z_-^{\tilde{P}}} \, d\tilde{L} \in \mathcal{R}^2(P)$$

which is precisely (43). This completes the proof. q.e.d.

Now, we have everything we need to solve the MVH problem. We first prove that $J(\vartheta^*) = j(\vartheta^*)$; this shows that our guess in Section II was correct and justifies *a posteriori* the assumption made in Lemma 2.

Theorem 13: Suppose that $I\!M_e^2 \neq \emptyset, I\!F$ is continuous and $\tilde{P}$ satisfies $R_2(P)$. Fix $H \in L^2(\mathcal{F}_T, P)$ and choose the solutions $(a, \Lambda^a, N^a), (b, \mu^b, \psi^b, N^b), (c, N^{(c)})$ of (6)–(8) as in Theorem 9. Suppose that b_0 is bounded. Fix $v_0 \in I\!R$ and take the solution ϑ^* of (37) given by Proposition 12. Then, we have

$$\begin{aligned} \operatorname{ess\,inf}_{\psi \in {}^t\Theta(\vartheta^*)} E[(H - v_0 - G_T(\psi))^2 \,|\, \mathcal{F}_t] &= J_t(\vartheta^*) \\ &= j_t(\vartheta^*) = a_t \left(V_t^{\vartheta^*} - b_t \right)^2 + c_t \end{aligned} \tag{44}$$

P-a.s. for every $t \in [0, T]$. In particular, $J_t(\vartheta^*)$ is a quadratic function of $V_t^{\vartheta^*}$ for each t.

Proof: The outer two equalities in (44) are just the definitions; only the middle one needs proof. We know from Proposition 10 that $j(\vartheta) = a(v_0 + G(\vartheta) - b)^2 + c$ is a local P-submartingale for any $\vartheta \in \Theta$ and a local P-martingale for $\vartheta = \vartheta^*$ which is in Θ by $j(\vartheta)$ is even a true P-submartingale and $j(\vartheta^*)$ a true P-martingale. To see this, note that $j(\vartheta)$ is of class (D) under P because c is so, by Theorem 9, because a is bounded, by Theorem 3, and because $G(\vartheta) \in \mathcal{S}^2(P) \subseteq \mathcal{R}^2(P)$ for $\vartheta \in \Theta$ and $b \in \mathcal{R}^2(P)$, by Proposition 8. So if we take any $\vartheta = \psi \in {}^t\Theta(\vartheta^*)$ and use the final conditions $a_T = 1, b_T = H, c_T = 0$, we obtain from the submartingale property

$$\begin{aligned} &E[(H - v_0 - G_T(\psi))^2 \,|\, \mathcal{F}_t] \\ &\quad = E[j_T(\psi) \,|\, \mathcal{F}_t] \geq j_t(\psi) \\ &\quad = a_t \left(V_t^{\vartheta^*} - b_t \right)^2 + c_t = j_t(\vartheta^*) \end{aligned}$$

because $v_0 + G_t(\psi) = v_0 + G_t(\vartheta^*) = V_t^{\vartheta^*}$ for $\psi \in {}^t\Theta(\vartheta^*)$. For $\psi = \vartheta^*$, we get equality since $j(\vartheta^*)$ is a P-martingale, and so (44) follows. q.e.d.

Corollary 15: Under the assumptions of Theorem 13, we also have for every $\vartheta \in \Theta$

$$\begin{aligned} &\operatorname{ess\,inf}_{\psi \in {}^t\Theta(\vartheta)} E[(H - v_0 - G_T(\psi))^2 \,|\, \mathcal{F}_t] \\ &\qquad = J_t(\vartheta) = j_t(\vartheta) = a_t \left(V_t^{\vartheta} - b_t \right)^2 + c_t \end{aligned}$$

P-a.s. for every $t \in [0, T]$.

Proof: This follows immediately from Theorem 13 and (the proof of) Lemma 2. q.e.d.

Now, we can solve the MVH problem (4) with fixed initial capital.

Corollary 15: Under the assumptions of Theorem 13, ϑ^* solves the MVH problem (4), and the minimal quadratic risk at time 0 is given by

$$\inf_{\vartheta \in \Theta} E[(H - v_0 - G_T(\vartheta))^2] = E[a_0(v_0 - b_0)^2 + c_0]. \tag{45}$$

Proof: The same argument as in the proof of Theorem 13 yields for any $\vartheta \in \Theta$ that

$$\begin{aligned} E[(H - v_0 - G_T(\vartheta))^2] &= E[j_T(\vartheta)] \\ &\geq E[j_0(\vartheta)] = E[a_0(v_0 - b_0)^2 c_0] \end{aligned}$$

with equality for $\vartheta = \vartheta^*$. This proves the assertion. q.e.d.

By minimizing over v_0, we obtain the solution of the MVH problem (3).

Corollary 16: Under the assumptions of Theorem 13, the solution of the MVH problem (3) is given by $(v_0^*, \vartheta^*(v_0^*))$, where $\vartheta^*(v_0^*)$ is the solution of (37) for the choice $v_0 = v_0^* = E_{\tilde{P}}[H]$. The minimal quadratic risk is given by

$$\begin{aligned} &\inf_{v_0 \in I\!R, \vartheta \in \Theta} E\left[(H - v_0 - G_T(\vartheta))^2 \right] \\ &\qquad = E[a_0] \operatorname{Var}[b_0] + E[c_0]. \end{aligned} \tag{46}$$

Proof: Minimizing the right-hand side of (45) yields $v_0^* = E[a_0 b_0]/E[a_0]$. Since $a = Z^{\tilde{P}}/\tilde{Z}$ and $\tilde{Z}_0$ is deterministic, we obtain $E[a_0] = 1/\tilde{Z}_0$ and

$$E[a_0 b_0] = \frac{1}{\tilde{Z}_0} E\left[Z_0^{\tilde{P}} E_{\tilde{P}}[H \,|\, \mathcal{F}_0] \right] = \frac{1}{\tilde{Z}_0} E_{\tilde{P}}[H]$$

by the definition of b in (26). Plugging this into (45) and using again that $a = Z^{\tilde{P}}/\tilde{Z}$ yields (46) if we note that $Z_0^{\tilde{P}} = 1$ since $\tilde{P} = P$ on $\mathcal{F}_0$. q.e.d.

Remark: Recall that S is assumed continuous throughout. We have seen in Theorem 3 that a necessary and sufficient condition for the solvability of the BSDE (6) for a is that $I\!M_e^2 \neq \emptyset$. Combining Proposition 12, Theorem 9, Proposition 8, and Theorem 3 shows that if the filtration $I\!F$ is continuous, $I\!M_e^2 \neq \emptyset$ plus the reverse Hölder inequality $R_2(P)$ for $\tilde{P}$ together form a *sufficient* condition for the solvability of the BSDE system (6)–(8) for a, b, c. Conversely, $I\!M_e^2 \neq \emptyset$ plus $R_2(P)$ is also *necessary* in a certain sense. More precisely, suppose that $I\!F$ is continuous. If the BSDEs (6)–(8) are solvable for any $H \in L^2(\mathcal{F}_T, P)$ with solutions that satisfy (12), (28), and (33), and if the solution ϑ^* to (37) (which exists thanks to Lemma 11) is in Θ for any $v_0 \in I\!R$, then we obtain that $I\!M_e^2 \neq \emptyset$ and $\tilde{P}$ satisfies $R_2(P)$. In fact, solvability of (6) implies $I\!M_e^2 \neq \emptyset$ by Theorem 3, and the argument for Corollary 15 shows that (4) is solvable by ϑ^* for any $H \in L^2(\mathcal{F}_T, P)$. However, this means that $G_T(\vartheta)$ must be closed in $L^2(P)$ which implies $R_2(P)$ by [3, Th. 4.1].

While the above converse result is gratifying because it shows that our conditions are sharp, it is not really useful since the condition that ϑ^* should be in Θ is difficult to check. ⋄

To round off this section, we briefly explain how the above results are related to the solution of the MVH problem via the martingale and projection approach in [22]. As explained there, one first has to find the variance-optimal martingale measure $\tilde{P}$ and represent the process $\tilde{Z}$ from (15) as

$$\tilde{Z} = \tilde{Z}_0 + \int \tilde{\zeta}\, dS. \tag{47}$$

Then one needs the Galtchouk–Kunita–Watanabe decomposition of H under $\tilde{P}$ as

$$H = \tilde{E}[H \mid \mathcal{F}_0] + \int_0^T \xi_u^{H,\tilde{P}}\, dS_u + L_T^{H,\tilde{P}}.$$

One writes

$$\begin{aligned} V_t^{H,\tilde{P}} &:= E_{\tilde{P}}[H \mid \mathcal{F}_t] = \tilde{E}[H \mid \mathcal{F}_0] \\ &\qquad + \int_0^t \xi_u^{H,\tilde{P}}\, dS_u + L_t^{H,\tilde{P}}, \qquad 0 \le t \le T \end{aligned}$$

and obtains the optimal strategy $(\tilde{V}_0, \tilde{\vartheta})$ for (3) as

$$\widetilde{V_0} = E_{\tilde{P}}[H]$$

and

$$\tilde{\vartheta} = \xi^{H,\tilde{P}} - \frac{\tilde{\zeta}}{\tilde{Z}} \left(V_-^{H,\tilde{P}} - \tilde{E}[H] - \int \tilde{\vartheta}\, dS \right). \tag{48}$$

All this is taken from [22, Th. 4.6].

To relate this to our present setting, we first note that the proof of Theorem 3 gives us $a = Z^{\tilde{P}}/\tilde{Z}$ and therefore

$$\frac{\tilde{\zeta}}{\tilde{Z}} = -\left(\lambda + \frac{\Lambda^a}{a_-} \right) \tag{49}$$

by (16). Because we also have from (47) that $\tilde{Z} = \tilde{Z}_0 \mathcal{E}(\int (\tilde{\zeta}/\tilde{Z})\, dS)$, we get

$$\tilde{Z} = \frac{1}{a_0} \mathcal{E}\left(-\int \left(\lambda + \frac{\Lambda^a}{a_-} \right) dS \right) \tag{50}$$

since $Z_0^{\tilde{P}} = 1$. We remark in passing that this means that $\beta := (1/a_0)(\lambda + (\Lambda^a/a_-))$ is an *adjustment process* in the terminology of [21]. Combining (49) and (50) yields

$$\tilde{\zeta} = -\frac{1}{a_0} \left(\lambda + \frac{\Lambda^a}{a_-} \right) \mathcal{E}\left(-\int \left(\lambda + \frac{\Lambda^a}{a_-} \right) dS \right).$$

From the proof of Theorem 7, we obtain $b = V^{H,\tilde{P}}$, $\mu^b = \xi^{H,\tilde{P}}$ and $L^b = L^{H,\tilde{P}}$. So, the solution to (3) obtained from Corollary 16 can be written as

$$v_0^* = E_{\tilde{P}}[H] = \widetilde{V_0}$$

and

$$\begin{aligned} \vartheta^* &= \mu^b - \left(\lambda + \frac{\Lambda^a}{a_-} \right) (v_0^* + G(\vartheta^*) - b_-) \\ &= \xi^{H,\tilde{P}} - \frac{\tilde{\zeta}}{\tilde{Z}} (V_-^{H,\tilde{P}} - E_{\tilde{P}}[H] - \int \vartheta^*\, dS). \end{aligned}$$

This coincides with the recursive equation (48) for $\tilde{\vartheta}$.

Remark: When this paper was almost completed, we received a preprint version of [16] by Mania and Tevzadze, who also study the mean-variance hedging problem for a continuous semimartingale model. These authors actually consider in [16] the more general problem of minimizing the L^p-norm of the final shortfall and derive a backward equation for the corresponding value process. For the case $p = 2$, they prove in particular that the value function is quadratic and derive a system of BSDEs for its coefficients. So their focus is more on deriving the BSDEs, whereas we concentrate on proving existence and uniqueness and on finding sharp conditions for this. □

References

[1] D. Bertsimas, L. Kogan, and A. Lo, “Hedging derivative securities and incomplete markets: An epsilon-arbitrage approach,” *Oper. Res.*, vol. 49, pp. 372–397, 2001.

[2] A. Černý, *Mean-Variance Hedging in Discrete Time*. London, U.K.: Imperial College Management School, 1999.

[3] F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer, and C. Stricker, “Weighted norm inequalities and hedging in incomplete markets,” *Finance Stoch.*, vol. 1, pp. 181–227, 1997.

[4] F. Delbaen and W. Schachermayer, “The existence of absolutely continuous local martingale measures,” *Ann. Appl. Probab.*, vol. 5, pp. 926–945, 1995.

[5] ——, “The variance-optimal martingale measure for continuous processes,” *BERNOULLI*, vol. 2, pp. 81–105, 1996.

[6] F. Delbaen and W. Schachermayer, “The fundamental theorem of asset pricing for unbounded stochastic processes,” *Mathematische Annalen*, vol. 312, pp. 215–250, 1998.

[7] Y. Hu and X. Y. Zhou, “Indefinite Riccati,” *SIAM J. Control Optim.*, vol. 42, pp. 123–137, 2003.

[8] M. Kohlmann and S. Tang, “Global adapted solution of one-dimensional backward Riccati equations, with application to the mean-variance hedging,” *Stoch. Processes Applicat.*, vol. 97, pp. 255–288, 2002.

[9] ——, “Multidimensional backward Riccati equations and applications,” *SIAM J. Control Optim.*, vol. 41, pp. 1696–1721, 2003.

[10] ——, “Minimization of risk and linear quadratic optimal control theory,” *SIAM J. Control Optim.*, vol. 42, pp. 1118–1142, 2003.

[11] M. Kohlmann and X. Y. Zhou, “Relationship between backward stochastic differential equations and stochastic controls: A linear-quadratic approach,” *SIAM J. Control Optim.*, vol. 38, pp. 1392–1407, 2000.

[12] J. P. Laurent and H. Pham, “Dynamic programming and mean-variance hedging,” *Finance Stoch.*, vol. 3, pp. 83–110, 1999.

[13] A. E. B. Lim and X. Y. Zhou, “LQ control and mean-variance portfolio selections: The stochastic parameter case,” in *Mathematical Finance. Workshop of the Mathematical Finance Research Project*, M. Kohlmann and S. Tang, Eds. Boston, MA: Birkhäuser, 2001, pp. 261–270.

[14] M. Mania and R. Tevzadze, “A semimartingale equation and the variance-optimal martingale measure,” *Georgian Math. J.*, vol. 7, pp. 765–792, 2000.

[15] ——, “A semimartingale backward equation and the variance-optimal martingale measure under general information flow,” *SIAM J. Control Optim.*, vol. 42, pp. 1703–1726, 2003.

[16] ——, “Backward stochastic PDE and imperfect hedging,” *Int. J. Theore. Applied Finance*, vol. 6, pp. 663–692, 2003.

[17] H. Pham, “On quadratic hedging in continuous time,” *Math. Meth. Oper. Res.*, vol. 51, pp. 315–339, 2000.

[18] P. Protter, *Stochastic Integration and Differential Equations. A New Approach*. New York: Springer-Verlag, 1990.

[19] T. Rheinländer and M. Schweizer, “On L^2-projections on a space of stochastic integrals,” *Ann. Probab.*, vol. 25, pp. 1810–1831, 1997.

[20] M. Schweizer, "On the minimal martingale measure and the Föllmer-Schweizer decomposition," *Stoch. Anal. Applicat.*, vol. 13, pp. 573–599, 1995.

[21] ——, "Approximation pricing and the variance-optimal martingale measure," *Ann. Probab.*, vol. 24, pp. 206–236, 1996.

[22] ——, "A guided tour through quadratic hedging approaches," in *Option Pricing, Interest Rates and Risk Management*, E. Jouini, J. Cvitanić, and M. Musiela, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2001, pp. 538–574.

[23] M. C. Steinbach, "Markowitz revisited: Mean-variance models in financial portfolio analysis," *SIAM Rev.*, vol. 43, pp. 31–85, 2001.

[24] S. Tang, "General linear quadratic optimal stochastic control problems with random coefficients: Linear stochastic Hamilton systems and backward Riccati equations," *SIAM J. Control Optim.*, vol. 42, pp. 53–75, 2003.

Olga Bobrovnytska is currently working toward the Ph.D. degree in mathematics at the Technical University of Berlin, Berlin, Germany.

Martin Schweizer received the Ph.D. degree in mathematics in 1988.

He is a Professor of mathematics at the ETH Zurich, Zurich, Switzerland He has held Professor positions at the TU Berlin, Berlin, Germany (1994–2001) and the University of Munich, Munich, Germany (2001–2003) before joining ETH in 2003. His research interests are in mathematical finance, probability theory, and mathematics and his focus in particular is on hedging and valuation problems in incomplete financial markets.