

**Mean-Variance Hedging for Continuous Processes:  
New Proofs and Examples**

*Huyên Pham\**

*Équipe d'Analyse et de Mathématiques Appliquées*

*Université de Marne-la-Vallée*

*2, rue de la Butte Verte*

*F – 93166 Noisy-le-Grand Cedex*

*France*

*Thorsten Rheinländer, Martin Schweizer\*\**

*Technische Universität Berlin*

*Fachbereich Mathematik, MA 7–4*

*Straße des 17. Juni 136*

*D – 10623 Berlin*

*Germany*

---

\* This author is also affiliated to CREST, Laboratoire de Finance-Assurance, Paris.

\*\* Research for this paper was partially carried out within Sonderforschungsbereich 373.

This paper was printed using funds made available by the Deutsche Forschungsgemeinschaft.

**Abstract:** Let  $X$  be a special semimartingale of the form  $X = X_0 + M + \int d\langle M \rangle \widehat{\lambda}$ , denote by  $\widehat{K} = \int \widehat{\lambda}^{\text{tr}} d\langle M \rangle \widehat{\lambda}$  the mean-variance tradeoff process of  $X$  and by  $\Theta$  the space of predictable processes  $\vartheta$  for which the stochastic integral  $G(\vartheta) = \int \vartheta dX$  is a square-integrable semimartingale. For a given constant  $c \in \mathbb{R}$  and a given square-integrable random variable  $H$ , the mean-variance optimal hedging strategy  $\xi^{(c)}$  minimizes the distance in  $\mathcal{L}^2$  between  $H - c$  and the space  $G_T(\Theta)$ . In financial terms,  $\xi^{(c)}$  provides an approximation of the contingent claim  $H$  by means of a self-financing trading strategy with minimal global risk. If  $\widehat{K}$  is bounded and continuous, we first give a simple new proof of the closedness of  $G_T(\Theta)$  in  $\mathcal{L}^2(P)$  and of the existence of the Föllmer-Schweizer decomposition. If moreover  $X$  is continuous and satisfies an additional condition, we can describe the mean-variance optimal strategy in feedback form, and we provide several examples where it can be computed explicitly. The additional condition states that the minimal and the variance-optimal martingale measures for  $X$  should coincide. We provide examples where this assumption is satisfied, but we also show that it will typically fail if  $\widehat{K}_T$  is not deterministic and includes exogenous randomness which is not induced by  $X$ .

**Key words:** mean-variance hedging, stochastic integrals, minimal martingale measure, Föllmer-Schweizer decomposition, variance-optimal martingale measure

**AMS 1991 subject classification:** 90A09, 60H05, 60G48

**JEL Classification Numbers:** G10, C60

*(Finance & Stochastics 2 (1998), 173–198)*

This version: 19.11.1996

## 0. Introduction

The pricing and hedging of contingent claims in an incomplete market is one of the important questions in modern financial mathematics. In the existing literature, one can find at least three different strands of ideas to attack this problem: the super-replication approach introduced by El Karoui/Quenez (1995), utility-based arguments as employed for instance by Davis (1994) or Karatzas/Kou (1996) and a mean-variance approach that originates with the work of Föllmer/Sondermann (1986). In this paper, we present some new results and a lot of examples on mean-variance hedging for continuous processes. One major contribution is to point out rather precisely to which extent one can or cannot use the minimal martingale measure to study this problem successfully.

Let us first clarify what we mean here by mean-variance hedging. Let  $X$  be a stochastic process describing the discounted price of a stock in a frictionless financial market. Under a very mild condition of absence of arbitrage,  $X$  must be a special semimartingale of the form  $X = X_0 + M + \int d\langle M \rangle \hat{\lambda}$  for some predictable process  $\hat{\lambda}$ , and we call  $\hat{K} := \int \hat{\lambda}^{\text{tr}} d\langle M \rangle \hat{\lambda}$  the *mean-variance tradeoff* process of  $X$ . If the local martingale  $\hat{Z} := \mathcal{E}(-\int \hat{\lambda} dM)$  is strictly positive (which will certainly be the case if  $X$  is continuous) and a true martingale, setting  $\frac{d\hat{P}}{dP} := \hat{Z}_T$  defines a probability measure  $\hat{P}$  equivalent to  $P$  which is called the *minimal martingale measure* for  $X$ . This measure will play an important role in the sequel. A contingent claim is an  $\mathcal{F}_T$ -measurable square-integrable random variable  $H$ ; it models the payoff from a certain financial product one is interested in. A strategy  $\vartheta$  is a predictable process such that the stochastic integral  $G(\vartheta) := \int \vartheta dX$  is well-defined and a square-integrable semimartingale. Intuitively,  $G(\vartheta)$  describes the trading gains induced by the self-financing portfolio strategy associated to  $\vartheta$ , and so  $H - c - G_T(\vartheta)$  is the total loss of a hedger who starts with initial capital  $c$ , uses the strategy  $\vartheta$  and has to pay the random amount  $H$  at date  $T$ . In this context, *mean-variance hedging* means solving the optimization problem

$$(0.1) \quad \text{minimize } E \left[ (H - c - G_T(\vartheta))^2 \right] \text{ over all strategies } \vartheta$$

whose solution will be denoted by  $\xi^{(c)}$  if it exists. More detailed explanations can for instance be found in the financial introduction of Delbaen/Monat/Schachermayer/Schweizer/Stricker (1996).

In section 1, we specify our conditions on  $X$ , define the space  $\Theta$  of trading strategies and explain the mean-variance hedging problem studied in this paper. From section 2 on, we assume that the mean-variance tradeoff process  $\hat{K}$  is bounded and continuous; typical examples are situations where  $X$  itself is continuous as well as models of jump-diffusion type. We prove that the space  $G_T(\Theta)$  is then closed in  $\mathcal{L}^2(P)$  and that every  $H \in \mathcal{L}^2(P)$  admits a

Föllmer-Schweizer decomposition as

$$H = H_0 + \int_0^T \xi_s^H dX_s + L_T^H$$

with  $H_0 \in \mathbb{R}$ ,  $\xi^H \in \Theta$  and a square-integrable martingale  $L^H$  strongly orthogonal to  $M$ . These results are actually well known, but by exploiting the continuity of  $\widehat{K}$ , we can provide a much simpler and unified argument. The closedness of  $G_T(\Theta)$  guarantees of course that (0.1) has indeed a solution for all  $H$  and  $c$ . Moreover, the boundedness of  $\widehat{K}$  implies that  $\widehat{Z}_T$  is in  $\mathcal{L}^2(P)$  and therefore has a Föllmer-Schweizer decomposition as

$$\widehat{Z}_T = E[\widehat{Z}_T^2] - E[\widehat{Z}_T \widehat{L}_T] + \int_0^T \widehat{\zeta}_s dX_s + \widehat{L}_T.$$

In section 3, we show that for  $H \in \mathcal{L}^{2+\varepsilon}(P)$ , the solution  $\xi^{(c)}$  of (0.1) is given in feedback form as

$$\xi_t^{(c)} = \xi_t^H - \frac{\widehat{\zeta}_t}{\widehat{E}[\widehat{Z}_T | \mathcal{F}_t]} \left( H_0 + \int_0^t \xi_s^H dX_s + L_{t-}^H - c - \int_0^t \xi_s^{(c)} dX_s \right)$$

if  $X$  is continuous,  $\widehat{K}$  is bounded and if

$$(0.2) \quad \widehat{L}_T = 0 \quad \text{in the Föllmer-Schweizer decomposition of } \widehat{Z}_T.$$

In the framework of a diffusion model, this result was first stated by Hipp (1993), but his argument fails to show that  $\xi^{(c)}$  does belong to  $\Theta$ . In work done independently of ours, Wiese (1995) has proved the same feedback representation result under (0.2) and otherwise slightly different assumptions; see section 3 for more detailed comments. Finally, section 4 is devoted to a number of examples which illustrate both the usefulness and the limitations of the main result. If  $\widehat{K}_T$  is deterministic, it is rather easy to see that the crucial assumption (0.2) is satisfied. Several examples illustrate the fact already pointed out by Hipp (1993) that (0.2) can still hold in situations where  $\widehat{K}_T$  is not deterministic. Moreover, we also show in examples how to compute the optimal strategy  $\xi^{(c)}$  more explicitly. Nevertheless, (0.2) is a rather special assumption: we prove in fact that it will typically fail as soon as  $\widehat{K}_T$  depends on exogenous randomness which is not generated by  $X$ . Since (0.2) is equivalent to assuming that the minimal and the variance-optimal martingale measures coincide, the conclusion from these counterexamples is that the ultimate solution of (0.1) will involve the variance-optimal rather than the minimal martingale measure. For results in this direction, we refer to Gouriéroux/Laurent/Pham (1996) and Rheinländer/Schweizer (1996).

# 1. Preliminaries and motivation

The purpose of this section is to fix the notation and to introduce the basic problem studied in the rest of the paper. Interpretations and motivation will be provided at the end of this section. Let  $(\Omega, \mathcal{F}, P)$  be a probability space with a filtration  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$  satisfying the usual conditions of right-continuity and completeness, where  $T \in (0, \infty]$  is a fixed time horizon. All processes considered will be indexed by  $t \in [0, T]$ . Let  $X$  be an  $\mathbb{R}^d$ -valued RCLL semimartingale in  $\mathcal{S}_{\text{loc}}^2$ ; this means that  $X$  is a special semimartingale with canonical decomposition  $X = X_0 + M + A$ , where  $M \in \mathcal{M}_{0, \text{loc}}^2$  and  $A$  is predictable and of locally square-integrable variation  $|A|$ . We assume that  $A^i \ll \langle M^i \rangle$  for  $i = 1, \dots, d$ , and we fix an increasing predictable RCLL process  $B$  null at 0 such that

$$\begin{aligned} \langle M^i, M^j \rangle_t &= \int_0^t \sigma_s^{ij} dB_s \quad , \quad 0 \leq t \leq T, \\ A_t^i &= \int_0^t \gamma_s^i dB_s \quad , \quad 0 \leq t \leq T \end{aligned}$$

for  $i, j = 1, \dots, d$ . We also assume that  $X$  satisfies the *structure condition (SC)*, i.e., that there is a predictable  $\mathbb{R}^d$ -valued process  $\hat{\lambda}$  such that

$$\sigma_t \hat{\lambda}_t = \gamma_t \quad P\text{-a.s. for } t \in [0, T]$$

and

$$\hat{K}_t := \int_0^t \hat{\lambda}_s^{\text{tr}} \gamma_s dB_s = \int_0^t \hat{\lambda}_s^{\text{tr}} \sigma_s \hat{\lambda}_s dB_s = \int_0^t \hat{\lambda}_s^{\text{tr}} d\langle M \rangle_s \hat{\lambda}_s < \infty \quad P\text{-a.s. for } t \in [0, T],$$

where  $^{\text{tr}}$  denotes transposition. We then fix an RCLL version of  $\hat{K}$  and call this the *mean-variance tradeoff (MVT) process* of  $X$ .

**Remark.** We shall later on assume that  $\hat{K}$  is continuous; this is equivalent to assuming that  $A$  is continuous. In fact, continuity of  $A$  is obviously sufficient, and the necessity follows from the Cauchy-Schwarz inequality.

**Definition.** For any RCLL process  $Y$ , we denote by  $Y^*$  the process

$$Y_t^* := \sup_{0 \leq s \leq t} |Y_s| \quad , \quad 0 \leq t \leq T.$$

We denote by  $\mathcal{R}^2(P)$  the space of all adapted RCLL processes  $Y$  such that

$$\|Y\|_{\mathcal{R}^2(P)} := \|Y_T^*\|_{\mathcal{L}^2(P)} < \infty.$$

**Definition.** For  $p \geq 1$ ,  $L^p(M)$  denotes the space of all predictable  $\mathbb{R}^d$ -valued processes  $\vartheta$  such that

$$\|\vartheta\|_{L^p(M)} := \left\| \left( \int_0^T \vartheta_s^{\text{tr}} \sigma_s \vartheta_s dB_s \right)^{\frac{1}{2}} \right\|_{\mathcal{L}^p(P)} = \left\| \left( \langle \int \vartheta dM \rangle_T \right)^{\frac{1}{2}} \right\|_{\mathcal{L}^p(P)} < \infty.$$

$L^p(A)$  denotes the space of all predictable  $\mathbb{R}^d$ -valued processes  $\vartheta$  such that

$$\|\vartheta\|_{L^p(A)} := \left\| \int_0^T |\vartheta_s^{\text{tr}} \gamma_s| dB_s \right\|_{\mathcal{L}^p(P)} = \left\| \int \vartheta^{\text{tr}} dA \Big|_T \right\|_{\mathcal{L}^p(P)} < \infty.$$

We remark that this definition of  $L^p(M)$  coincides with the one in Jacod (1979) if  $p = 2$  or if  $X$  is continuous, since  $[M] = \langle M \rangle$  in the latter case. Finally, we set

$$\Theta := L^2(M) \cap L^2(A).$$

As shown in Lemma 2 of Schweizer (1994),  $\Theta$  is the space of all  $\mathbb{R}^d$ -valued predictable  $X$ -integrable processes  $\vartheta$  such that the stochastic integral

$$G(\vartheta) := \int \vartheta dX$$

is in the space  $\mathcal{S}^2$  of semimartingales. If  $\widehat{K}$  is bounded, then  $\Theta = L^2(M)$  by the Cauchy-Schwarz inequality.

For a fixed constant  $c \in \mathbb{R}$  and a given random variable  $H \in \mathcal{L}^2(\mathcal{F}_T, P)$ , we now consider the optimization problem

$$(1.1) \quad \text{Minimize } E \left[ (H - c - G_T(\vartheta))^2 \right] \text{ over all } \vartheta \in \Theta$$

and denote its solution by  $\xi^{(c)}$  if it exists.

**Interpretation.** Problem (1.1) arises naturally in financial mathematics when one studies so-called *mean-variance optimal hedging strategies*. Let us think of  $X_t$  as the discounted price at time  $t$  of a (possibly multivariate) risky asset (e.g., a finite number of stocks) and of  $\vartheta$  as a dynamic portfolio strategy in the sense that  $\vartheta_t^i$  describes the number of shares of asset  $i$  to be held at time  $t$ . Let us also assume the existence of some riskless asset (e.g., a bank account or a zero coupon bond) with discounted price 1 at all times. Every  $\vartheta \in \Theta$  then uniquely determines a self-financing trading strategy by the requirement that the value process should be given by  $c + \int \vartheta dX$ , where  $c \in \mathbb{R}$  denotes a given initial capital at time 0; see Harrison/Pliska (1981). In such a framework, the random variable  $H$  can then be interpreted as a contingent

claim, i.e., as the obligation to provide at time  $T$  the random payoff  $H$ . The net loss resulting from the use of some pair  $(c, \vartheta)$  is then obviously  $H - c - \int_0^T \vartheta_s dX_s$ , and a mean-variance optimal strategy has the property that it provides the best approximation of  $H$  in the mean-square sense by the final wealth obtainable by self-financing trading. Condition (SC) is a consequence of a very weak assumption of absence of arbitrage and therefore very natural for the problem under consideration; see for instance Delbaen/Schachermayer (1995). Finally, the mean-variance tradeoff process  $\widehat{K}$  can be viewed as the integrated squared market price of risk associated to  $X$ ; in the familiar Black-Scholes model of geometric Brownian motion with drift  $b$ , volatility  $v$  and riskless interest rate  $r$ , it is for instance given by  $\widehat{K}_t = \left(\frac{b-r}{v}\right)^2 t$ .

## 2. Closedness of $G_T(\Theta)$ , and the Föllmer-Schweizer decomposition

The optimization problem (1.1) immediately raises the question if the space  $G_T(\Theta)$  of stochastic integrals of  $X$  is closed in  $\mathcal{L}^2(P)$ . It is already known from the results of Monat/Stricker (1994, 1995) that the answer is positive if the MVT process  $\widehat{K}$  is bounded. For a necessary and sufficient condition, see Delbaen/Monat/Schachermayer/Schweizer/Stricker (1996). In this section, we give a simple new proof of the closedness of  $G_T(\Theta)$  under the assumption that  $\widehat{K}$  is bounded and *continuous*. As a by-product of our approach, we also obtain a simple new proof of the existence of a Föllmer-Schweizer decomposition and sharper results on its integrability properties.

**Remark.** Continuity of  $\widehat{K}$  is not a very restrictive assumption since the martingale part of  $X$  can still have some jump component. If  $X$  itself is continuous, then of course so is  $\widehat{K}$ , but for instance all jump-diffusion models whose jump component has a continuous compensator also satisfy this condition.

The first result is actually the heart of our argument. It is inspired by the method used in El Karoui/Peng/Quenez (1997) to derive a priori estimates for solutions of backward stochastic differential equations.

**Proposition 1.** Fix  $\vartheta, \psi \in \Theta$ ,  $V_0 \in \mathcal{L}^2(\mathcal{F}_0, P)$  and  $L \in \mathcal{M}^2(P)$  strongly orthogonal to  $M$ , and define the process  $V$  by

$$V_t := V_0 + \int_0^t \vartheta_s^{\text{tr}} dA_s + \int_0^t \psi_s dM_s + L_t \quad , \quad 0 \leq t \leq T.$$

Let  $C$  be any nonnegative increasing predictable RCLL process. If  $C$  is bounded, then

$$E [C_T V_T^2] \geq E \left[ \int_0^T V_{s-}^2 dC_s - \mu^2 \int_0^T V_{s-}^2 C_s d\widehat{K}_s + \int_0^T C_s \psi_s^{\text{tr}} \sigma_s \psi_s dB_s - \frac{1}{\mu^2} \int_0^T C_s \vartheta_s^{\text{tr}} \sigma_s \vartheta_s dB_s \right]$$

for any  $\mu \neq 0$ .

**Proof.** Since  $C$  is increasing and predictable, we obtain from Theorem VIII.19 of Del-  
lacherie/Meyer (1982), Itô's formula and the definition of  $V$

$$\begin{aligned} C_T V_T^2 - C_0 V_0^2 &= \int_0^T V_{s-}^2 dC_s + \int_0^T C_s d(V_s^2) \\ &= \int_0^T V_{s-}^2 dC_s + 2 \int_0^T C_s V_{s-} dV_s + \int_0^T C_s d[V]_s \\ &= \int_0^T V_{s-}^2 dC_s + 2 \int_0^T C_s V_{s-} \vartheta_s^{\text{tr}} dA_s + \int_0^T C_s d[\int \vartheta^{\text{tr}} dA]_s \\ &\quad + \int_0^T C_s d[\int \psi dM]_s + \int_0^T C_s d[L]_s + 2 \int_0^T C_s d[\int \psi dM, L]_s \\ &\quad + 2 \int_0^T C_s V_{s-} \psi_s dM_s + 2 \int_0^T C_s V_{s-} dL_s + 2 \int_0^T C_s d[\int \vartheta^{\text{tr}} dA, \int \psi dM + L]_s \\ &=: \sum_{i=1}^9 \text{term}(i). \end{aligned}$$

Since  $L \in \mathcal{M}^2(P)$  and  $V \in \mathcal{R}^2(P)$ , the process  $\int V_- dL$  is a martingale: it is a local martingale whose supremum is in  $\mathcal{L}^1(P)$  by the Burkholder-Davis-Gundy inequality, because

$$[\int V_- dL]_T^{\frac{1}{2}} \leq V_T^* [L]_T^{\frac{1}{2}} \in \mathcal{L}^1(P).$$

Since  $C$  is predictable and bounded,  $\int CV_- dL$  is also a martingale, and so term(8) is integrable with expectation 0. The same is true for term(7) since  $\int \psi dM \in \mathcal{M}_0^2(P)$  for  $\psi \in \Theta$ , for term(6) by the strong orthogonality of the processes  $\int \psi dM$  and  $L$  in  $\mathcal{M}_0^2(P)$ , and also for term(9). The last assertion follows from the fact that  $[F, N]$  is a martingale whenever  $N \in \mathcal{M}^2(P)$  and  $F$  is predictable and of square-integrable variation; see the proof of Lemma 6 in Schweizer (1994). Because  $\psi$  is in  $\Theta$  and  $C$  is predictable and bounded, term(4) is



integrable and has the same expectation as

$$\int_0^T C_s d\langle \int \psi dM \rangle_s = \int_0^T C_s \psi_s^{\text{tr}} \sigma_s \psi_s dB_s.$$

Term(3) and term(5) are nonnegative, and term(2) can be estimated below by

$$2 \int_0^T C_s V_{s-} \vartheta_s^{\text{tr}} dA_s = 2 \int_0^T C_s V_{s-} \vartheta_s^{\text{tr}} \sigma_s \widehat{\lambda}_s dB_s \geq -2 \left( \int_0^T C_s \vartheta_s^{\text{tr}} \sigma_s \vartheta_s dB_s \right)^{\frac{1}{2}} \left( \int_0^T C_s V_{s-}^2 d\widehat{K}_s \right)^{\frac{1}{2}}$$

due to the Cauchy-Schwarz inequality. Using now the elementary inequality

$$-2\sqrt{ab} \geq -\frac{1}{\mu^2}a - \mu^2b$$

on the lower bound for term(2) and putting everything together yields

$$C_T V_T^2 \geq \int_0^T V_{s-}^2 dC_s - \mu^2 \int_0^T V_{s-}^2 C_s d\widehat{K}_s + \int_0^T C_s \psi_s^{\text{tr}} \sigma_s \psi_s dB_s - \frac{1}{\mu^2} \int_0^T C_s \vartheta_s^{\text{tr}} \sigma_s \vartheta_s dB_s + N_T$$

for a martingale  $N$  null at 0. But  $V \in \mathcal{R}^2(P)$ ,  $\vartheta, \psi \in \Theta$  and boundedness of  $C$  imply that the expectation of the right-hand side is well-defined in  $[-\infty, +\infty)$ , and so the assertion follows.

**q.e.d.**

**Lemma 2.** *Let  $F$  be an increasing predictable RCLL process null at 0 with jumps bounded by a constant  $b$ . For every  $\beta \in (0, \frac{1}{b})$ , the process*

$$C_t^\beta := \frac{1}{\mathcal{E}(-\beta F)_t}$$

*is then the unique increasing predictable RCLL solution of the equation*

$$C_t = 1 + \int_0^t \beta C_s dF_s \quad , \quad 0 \leq t \leq T.$$

*If  $F$  is bounded, then so is  $C^\beta$ .*

**Proof.** The process  $\beta F$  is obviously a special semimartingale, and the jumps of the predictable process in its canonical decomposition are strictly less than 1. By Theorem (6.13) of Jacod (1979),  $C^\beta$  is therefore the unique solution of the equation

$$C_t = 1 + \int_0^t \beta C_s dF_s \quad , \quad 0 \leq t \leq T,$$

where  ${}^pC$  is the  $\mathbb{F}$ -predictable projection of  $C$ . Since  $C^\beta$  is predictable, the first assertion follows. Moreover, the estimate

$$C_t^\beta = \frac{1}{\mathcal{E}(-\beta F)_t} = e^{\beta F_t} \prod_{0 < s \leq t} \frac{e^{-\beta \Delta F_s}}{1 - \beta \Delta F_s} \leq \frac{1}{1 - \beta b} e^{\beta F_t} \prod_{0 < s \leq t} e^{-\beta \Delta F_s}$$

readily implies the second assertion.

**q.e.d.**

If the process  $F$  in Lemma 2 is continuous, the result is of course obvious: we simply have  $C^\beta = e^{\beta F}$  for any  $\beta > 0$ . This observation leads to the following result:

**Proposition 3.** Fix  $\vartheta, \psi \in \Theta$ ,  $V_0 \in \mathcal{L}^2(\mathcal{F}_0, P)$  and  $L \in \mathcal{M}^2(P)$  strongly orthogonal to  $M$ , and define the process  $V$  by

$$V_t := V_0 + \int_0^t \vartheta_s^{\text{tr}} dA_s + \int_0^t \psi_s dM_s + L_t, \quad 0 \leq t \leq T.$$

If the MVT process  $\widehat{K}$  is continuous and bounded, then

$$\begin{aligned} E \left[ e^{\beta \widehat{K}_T} V_T^2 \right] &\geq (\beta - \mu^2) E \left[ \int_0^T e^{\beta \widehat{K}_s} V_{s-}^2 d\widehat{K}_s \right] \\ &\quad + E \left[ \int_0^T e^{\beta \widehat{K}_s} \psi_s^{\text{tr}} \sigma_s \psi_s dB_s \right] - \frac{1}{\mu^2} E \left[ \int_0^T e^{\beta \widehat{K}_s} \vartheta_s^{\text{tr}} \sigma_s \vartheta_s dB_s \right] \end{aligned}$$

for all  $\beta > 0$  and  $\mu \neq 0$ .

**Proof.** Choose  $C = e^{\beta \widehat{K}}$  in Proposition 1.

**q.e.d.**

As a first application, we obtain a very simple proof of the following result:

**Corollary 4.** If the MVT process  $\widehat{K}$  is continuous and bounded, the space

$$G_T(\Theta) = \left\{ \int_0^T \vartheta_s dX_s \mid \vartheta \in \Theta \right\}$$

is closed in  $\mathcal{L}^2(P)$ , and the expressions  $\|\vartheta\|_{L^2(M)}$  and  $|\vartheta|_2 := \left\| \int_0^T \vartheta_s dX_s \right\|_{\mathcal{L}^2(P)}$  define equivalent norms on  $\Theta$ .

**Proof.** Since  $\widehat{K}$  is bounded, we have  $\Theta = L^2(M)$  and

$$|\vartheta|_2 \leq \left(1 + \|\widehat{K}_T\|_{\infty}^{\frac{1}{2}}\right) \|\vartheta\|_{L^2(M)}$$

by the Cauchy-Schwarz inequality. Applying Proposition 3 with  $\psi = \vartheta$ ,  $V_0 = 0$ ,  $L \equiv 0$  and  $\beta > \mu^2 > 1$  yields

$$\begin{aligned} \left(1 - \frac{1}{\mu^2}\right) \|\vartheta\|_{L^2(M)}^2 &\leq \left(1 - \frac{1}{\mu^2}\right) E \left[ \int_0^T e^{\beta \widehat{K}_s} \vartheta_s^{\text{tr}} \sigma_s \vartheta_s dB_s \right] \\ &\leq E \left[ e^{\beta \widehat{K}_T} \left( \int_0^T \vartheta_s dX_s \right)^2 \right] \\ &\leq e^{\beta \|\widehat{K}_T\|_{\infty}} |\vartheta|_2^2. \end{aligned}$$

We thus obtain the equivalence of the two norms, hence also the closedness of  $G_T(\Theta)$ .

**q.e.d.**

The interesting aspect of Corollary 4 is of course not the result itself, but its proof. It is well known that the conclusion of Corollary 4 is even true under the sole assumption that  $\widehat{K}$  is bounded; see Monat/Stricker (1994, 1995). The proofs in these papers use a “salami technique”: since the result is easy to establish if  $\|\widehat{K}_T\|_{\infty} < 1$ , the interval  $[0, T]$  is chopped up by suitable stopping times into random subintervals where  $\widehat{K}$  grows by less than some constant  $b < 1$ . Jumps of  $\widehat{K}$  whose size exceeds  $b$  have to be dealt with separately. On each subinterval, one can argue as if  $\widehat{K}$  were bounded by  $b < 1$ , and a backward induction argument alternating between jumps and subintervals then completes the proof. However, this method is somewhat unsatisfactory since the chopping-up is *always* required — even if  $\widehat{K}$  or  $X$  itself is continuous. To highlight the problem, consider the process

$$X_t = W_t + t \quad , \quad 0 \leq t \leq 1,$$

where  $\widehat{K}_t = t$  is bounded, but  $\widehat{K}_T = \widehat{K}_1 = 1$ . Even in this very simple case, proving the closedness of the space

$$G_1(\Theta) = \left\{ \int_0^1 \vartheta_s dW_s + \int_0^1 \vartheta_s ds \mid \vartheta \text{ is predictable and satisfies } E \left[ \int_0^1 \vartheta_s^2 ds \right] < \infty \right\}$$

up to now required two steps. Our new proof based on Proposition 1 eliminates this difficulty and provides a short and direct argument.

Actually, our technique can also be used to give a slightly shorter proof of the general result that  $G_T(\Theta)$  is closed in  $\mathcal{L}^2(P)$  if  $\widehat{K}$  is only bounded. In that case, however, we still need to resort to one argument from Monat/Stricker (1994) to deal with large jumps of  $\widehat{K}$ . More precisely, fix some  $b \in (0, 1)$  and denote by  $\tau_j$  the time of the  $j$ -th jump of  $\widehat{K}$  of size  $\geq b$ . Since  $\widehat{K}$  is bounded and predictable, there are only finitely many  $\tau_1, \dots, \tau_n$ , and each  $\tau_j$  is a predictable stopping time. By Lemma 2.2 of Monat/Stricker (1994),

$$\|\vartheta I_{\llbracket \tau_j \rrbracket}\|_{L^2(M)} \leq |\vartheta I_{\llbracket 0, \tau_j \rrbracket}|_2 \quad \text{for } j = 1, \dots, n.$$

On the other hand, applying Lemma 2 with  $F := \int I_{\llbracket \tau_{j-1}, \tau_j \rrbracket} d\widehat{K}$  and  $0 < \beta < \frac{1}{b}$  leads to

$$\|\vartheta I_{\llbracket \tau_{j-1}, \tau_j \rrbracket}\|_{L^2(M)} \leq C \left\| \int_0^{\tau_j^-} \vartheta_s dX_s \right\|_{\mathcal{L}^2(P)} = C |\vartheta I_{\llbracket 0, \tau_j \rrbracket}|_2$$

for  $j = 1, \dots, n$  and some constant  $C$ . To see this, argue as in the proofs of Corollary 4 and Proposition 1 with  $1 < \mu^2 < \beta < \frac{1}{b}$ ,  $C = C^\beta$  and  $\tau_{j-1}, \tau_j^-$  instead of 0 and  $T$ , respectively. The preceding two inequalities imply as in Monat/Stricker (1994) that

$$\|\vartheta\|_{L^2(M)} \leq C |\vartheta|_2 \quad \text{for some constant } C$$

which again yields the equivalence of the two norms.

**Remark.** It does not seem possible to avoid the separate treatment of large jumps of  $\widehat{K}$ . This is due to the fact that Lemma 2 may fail without the assumption of bounded jumps; see Remark (6.15) of Jacod (1979).

As a second application, we can also give a short and transparent proof of the following result:

**Corollary 5.** *If the MVT process  $\widehat{K}$  is continuous and bounded, then every  $H \in \mathcal{L}^2(\mathcal{F}_T, P)$  admits a Föllmer-Schweizer decomposition as*

$$H = H_0 + \int_0^T \xi_s^H dX_s + L_T^H \quad P\text{-a.s.}$$

with  $H_0 \in \mathbb{R}$ ,  $\xi^H \in \Theta$  and  $L^H \in \mathcal{M}^2(P)$  strongly orthogonal to  $M$  with  $E[L_0^H] = 0$ .

**Proof.** Since  $\widehat{K}$  is bounded, we have  $\Theta = L^2(M)$ . Consider the mapping  $J : \Theta \rightarrow \Theta$  which maps  $\vartheta$  into the integrand  $\psi$  of  $M$  in the Galtchouk-Kunita-Watanabe decomposition

of  $H - \int_0^T \vartheta_s^{\text{tr}} dA_s$ , i.e.,

$$H - \int_0^T \vartheta_s^{\text{tr}} dA_s = E \left[ H - \int_0^T \vartheta_s^{\text{tr}} dA_s \right] + \int_0^T \psi_s dM_s + L_T(\vartheta) =: H_0(\vartheta) + \int_0^T \psi_s dM_s + L_T(\vartheta).$$

Finding a Föllmer-Schweizer decomposition is then clearly equivalent to finding a fixed point of  $J$ . For any  $\beta > 0$ , setting

$$\|\vartheta\|_\beta := \left\| \left( \int_0^T e^{\beta \widehat{K}_s} \vartheta_s^{\text{tr}} \sigma_s \vartheta_s dB_s \right)^{\frac{1}{2}} \right\|_{\mathcal{L}^2(P)}$$

defines a norm on  $\Theta$  which is equivalent to  $\|\cdot\|_{L^2(M)}$ . But if we now apply Proposition 3 with  $\beta > \mu^2 > 1$ ,  $\vartheta = \vartheta^1 - \vartheta^2$ ,  $\psi = J(\vartheta^1) - J(\vartheta^2)$ ,  $V_0 = H_0(\vartheta^1) - H_0(\vartheta^2)$ ,  $L = L(\vartheta^1) - L(\vartheta^2)$  and notice that this yields  $V_T = 0$ , we immediately obtain

$$\begin{aligned} \|J(\vartheta^1) - J(\vartheta^2)\|_\beta^2 &= E \left[ \int_0^T e^{\beta \widehat{K}_s} \psi_s^{\text{tr}} \sigma_s \psi_s dB_s \right] \\ &\leq \frac{1}{\mu^2} E \left[ \int_0^T e^{\beta \widehat{K}_s} \vartheta_s^{\text{tr}} \sigma_s \vartheta_s dB_s \right] \\ &= \frac{1}{\mu^2} \|\vartheta^1 - \vartheta^2\|_\beta^2 \end{aligned}$$

so that  $J$  is a contraction on  $(\Theta, \|\cdot\|_\beta)$ .

**q.e.d.**

Like Corollary 4, the preceding result is also true without the assumption that  $\widehat{K}$  is continuous; see Monat/Stricker (1995). The new feature of our approach is again that we do not need any chopping technique if  $\widehat{K}$  is continuous or, more generally, has jumps bounded by some constant  $b < 1$ . The general case can also be handled with our method if we use this time Lemma 3.3 of Monat/Stricker (1995) to deal with the jumps of  $\widehat{K}$  whose size exceeds  $b$ . This is quite similar to the argument before Corollary 5; the details are left to the reader.

A third application concerns the additional integrability properties of the various terms in the Föllmer-Schweizer decomposition if  $X$  is continuous and  $H$  is in  $\mathcal{L}^p(P)$  for some  $p \geq 2$ .

**Lemma 6.** *Assume that  $X$  is continuous. If  $\widehat{K}$  is bounded, then any  $H \in \mathcal{L}^p(\mathcal{F}_T, P)$  with  $p \geq 2$  has a Föllmer-Schweizer decomposition with  $\xi^H \in L^p(M)$  and  $L^H \in \mathcal{M}^p(P)$ .*

**Proof.** Since  $\widehat{K}$  is bounded and continuous, the proof of Corollary 5 shows that the mapping  $J$  is a contraction on  $(\Theta, \|\cdot\|_\beta)$  so that  $\xi^H = \lim_{n \rightarrow \infty} J^n(\vartheta)$  for any  $\vartheta \in \Theta = L^2(M)$ . To prove that  $\xi^H$  is in  $L^p(M)$ , it is therefore enough to show that  $J$  maps  $L^p(M)$  into itself. Since  $\widehat{K}$  is bounded,  $L^p(M) \subseteq L^p(A)$  by the Cauchy-Schwarz inequality. Now fix  $\vartheta \in L^p(M)$  and consider the Galtchouk-Kunita-Watanabe decomposition

$$H - \int_0^T \vartheta_s^{\text{tr}} dA_s = H_0(\vartheta) + \int_0^T \psi_s dM_s + L_T(\vartheta).$$

Since  $H \in \mathcal{L}^p(P)$  and  $\vartheta \in L^p(A)$ , we obviously have

$$H_0(\vartheta) + L_0(\vartheta) = E \left[ H - \int_0^T \vartheta_s^{\text{tr}} dA_s \middle| \mathcal{F}_0 \right] \in \mathcal{L}^p(\mathcal{F}_0, P).$$

Moreover, the continuity of  $X$  implies that

$$[\int \psi dM, L(\vartheta)] = \langle \int \psi dM, L(\vartheta) \rangle = 0$$

by the strong orthogonality of  $L(\vartheta)$  and  $M$ . From the Burkholder-Davis-Gundy and Doob inequalities, we therefore obtain

$$\begin{aligned} \left\| \left( \int_0^T \psi_s^{\text{tr}} \sigma_s \psi_s dB_s \right)^{\frac{1}{2}} + [L(\vartheta)]_T^{\frac{1}{2}} \right\|_{\mathcal{L}^p(P)} &= \left\| [\int \psi dM]_T^{\frac{1}{2}} + [L(\vartheta)]_T^{\frac{1}{2}} \right\|_{\mathcal{L}^p(P)} \\ &\leq \sqrt{2} \left\| [\int \psi dM + L(\vartheta)]_T^{\frac{1}{2}} \right\|_{\mathcal{L}^p(P)} \\ &\leq \text{const.} \left\| (\int \psi dM + L(\vartheta))_T^* \right\|_{\mathcal{L}^p(P)} \\ &\leq \text{const.} \left\| \int_0^T \psi_s dM_s + L_T(\vartheta) \right\|_{\mathcal{L}^p(P)} \\ &\leq \text{const.} \left\| H - \int_0^T \vartheta_s^{\text{tr}} dA_s \right\|_{\mathcal{L}^p(P)} < \infty. \end{aligned}$$

Hence we conclude that  $\psi \in L^p(M)$  and  $L(\vartheta) \in \mathcal{M}^p(P)$ , and this implies the assertion.

**q.e.d.**

**Remarks.** 1) Lemma 6 is a slight sharpening of Corollary 10 of Schweizer (1995) where the conclusion was only that  $\xi^H \in L^r(M)$  and  $L^H \in \mathcal{M}^r(P)$  for every  $r < p$ . This loss

of integrability was due to the fact that the proof given there used the Galtchouk-Kunita-Watanabe decomposition of  $H$  under the minimal martingale measure  $\widehat{P}$  instead of working directly under  $P$  as above.

2) We could also obtain a Föllmer-Schweizer decomposition with  $H_0$   $\mathcal{F}_0$ -measurable and  $L_0^H = 0$  by shifting the initial value of  $L^H$  to  $H_0$ . But to facilitate comparison with other papers, we shall use the decomposition with a constant  $H_0$ .

### 3. A description of the optimal strategy

For practical purposes, the mere existence of a mean-variance optimal strategy is of course not very satisfactory. In this section, we therefore provide a description of  $\xi^{(c)}$  in *feedback form* if  $X$  is continuous and satisfies a special assumption. For the case of a diffusion model with a Brownian filtration, this result was first stated by Hipp (1993), but there is a gap in his proof. We present here a complete proof in the more general case where  $X$  is a continuous semimartingale with a bounded mean-variance tradeoff. In independent work, Wiese (1995) has obtained essentially the same result; we shall comment below on the similarities and the differences in the two formulations. Examples will be given in the next section.

Let  $X$  be a *continuous* semimartingale satisfying the structure condition (SC) and denote by  $\widehat{Z} := \mathcal{E}\left(-\int \widehat{\lambda} dM\right)$  the minimal martingale density. If  $E[\widehat{Z}_T] = 1$ , then Girsanov's theorem shows that

$$\frac{d\widehat{P}}{dP} := \widehat{Z}_T$$

defines an equivalent local martingale measure  $\widehat{P}$  for  $X$ , i.e., a probability  $\widehat{P} \approx P$  under which  $X$  is a local martingale.  $\widehat{P}$  is the *minimal local martingale measure* for  $X$ ; see Föllmer/Schweizer (1991) or Schweizer (1995). If  $\widehat{K}_T$  is *bounded*, we have in addition

$$(3.1) \quad \frac{d\widehat{P}}{dP} \in \mathcal{L}^r(P) \quad \text{for every } r < \infty$$

and

$$(3.2) \quad \frac{dP}{d\widehat{P}} \in \mathcal{L}^r(\widehat{P}) \quad \text{for every } r < \infty.$$

Lemma 6 and (3.1) then yield the Föllmer-Schweizer decomposition

$$(3.3) \quad \frac{d\widehat{P}}{dP} = E[\widehat{Z}_T^2] - E[\widehat{Z}_T \widehat{L}_T] + \int_0^T \widehat{\zeta}_s dX_s + \widehat{L}_T$$

with  $\widehat{L} \in \mathcal{M}^r(P)$  for every  $r < \infty$  and

$$\widehat{\zeta} \in L^r(M) \quad \text{for every } r < \infty.$$

To obtain the constant in (3.3), one uses the fact that  $\int \widehat{\zeta} dX$  and  $\widehat{L}$  are both  $\widehat{P}$ -martingales; this is due to (3.1) and the minimality of  $\widehat{P}$ .

The main result of this section is

**Theorem 7.** *Suppose that  $X$  is continuous and  $\widehat{K}$  is bounded. Assume that  $X$  satisfies the*

$$(3.4) \quad \textbf{special assumption:} \quad \widehat{L}_T = 0 \quad \text{in the decomposition (3.3).}$$

For fixed  $H \in \mathcal{L}^{2+\varepsilon}(\mathcal{F}_T, P)$  with  $\varepsilon > 0$ , the solution  $\xi^{(c)}$  of (1.1) is then given by

$$(3.5) \quad \xi_t^{(c)} = \xi_t^H - \frac{\widehat{\zeta}_t}{\widehat{Z}_t^0} \left( \widehat{V}_{t-} - c - \int_0^t \xi_s^{(c)} dX_s \right),$$

where

$$\widehat{Z}_t^0 := \widehat{E}[\widehat{Z}_T | \mathcal{F}_t] = E[\widehat{Z}_T^2] + \int_0^t \widehat{\zeta}_s dX_s \quad , \quad 0 \leq t \leq T$$

and

$$\widehat{V}_t := \widehat{E}[H | \mathcal{F}_t] = H_0 + \int_0^t \xi_s^H dX_s + L_t^H \quad , \quad 0 \leq t \leq T.$$

**Comment.** As mentioned above, essentially the same result as Theorem 7 has independently also been obtained by Wiese (1995). The special assumption (3.4) is equally present in her paper. While we impose boundedness on the mean-variance tradeoff to ensure a Föllmer-Schweizer decomposition for all contingent claims  $H$ , she assumes the existence of Föllmer-Schweizer decompositions for  $H$  and  $\widehat{Z}_T$  and imposes integrability assumptions directly on the various terms in those decompositions. This has the advantage over our approach that it also includes models with an unbounded mean-variance tradeoff, for instance the Ornstein-Uhlenbeck process. Moreover, the process  $X$  in Wiese's work is one-dimensional, but not necessarily continuous. However, this generalization is bought at the expense of the additional assumption that the martingale  $L^H$  in the decomposition of  $H$  must be continuous, in order to obtain the crucial property  $[L^H, X] = 0$  used in the subsequent argument. For a general claim  $H$ , this condition on  $L^H$  seems rather hard to verify.



**Idea of proof of Theorem 7:** By the projection theorem, the optimal strategy  $\xi^{(c)}$  is characterized by the property that

$$E \left[ \left( H - c - G_T(\xi^{(c)}) \right) G_T(\vartheta) \right] = \widehat{E} \left[ \frac{H - c - G_T(\xi^{(c)})}{\widehat{Z}_T} G_T(\vartheta) \right] = 0 \quad \text{for every } \vartheta \in \Theta.$$

Since  $\widehat{P}$  is a martingale measure for  $X$ , we know that

$$\widehat{E}[N_T G_T(\Theta)] = 0 \quad \text{for all bounded } \vartheta \in \Theta$$

for any  $N \in \mathcal{M}^2(\widehat{P})$  strongly  $\widehat{P}$ -orthogonal to  $X$ . This suggests to look for such an  $N$  with the special property that

$$(3.6) \quad H - c - G_T(\xi^{(c)}) = N_T \widehat{Z}_T = N_T \widehat{Z}_T^0.$$

Now apply the product rule and use the Föllmer-Schweizer decompositions of  $H$  and  $\widehat{Z}_T$  together with the special assumption (3.4) to deduce

$$\begin{aligned} & H - c - G_T(\xi^{(c)}) - N_T \widehat{Z}_T^0 \\ &= H_0 - c - N_0 E[\widehat{Z}_T^2] + \int_0^T \left( \xi_s^H - \xi_s^{(c)} - N_{s-} \widehat{\zeta}_s \right) dX_s + L_T^H - \int_0^T \widehat{Z}_s^0 dN_s - [N, \widehat{Z}^0]_T. \end{aligned}$$

But by the special assumption (3.4) and the continuity of  $X$ ,

$$[N, \widehat{Z}^0] = \int \widehat{\zeta}^{\text{tr}} d[N, X] = \int \widehat{\zeta}^{\text{tr}} d\langle N, X \rangle^{\widehat{P}} = 0$$

since  $N$  is strongly  $\widehat{P}$ -orthogonal to  $X$ . Thus we see that (3.6) will be satisfied with the choices

$$(3.7) \quad N_t := \frac{H_0 - c + L_0^H}{E[\widehat{Z}_T^2]} + \int_0^t \frac{1}{\widehat{Z}_s^0} dL_s^H$$

and

$$(3.8) \quad \xi^{(c)} := \xi^H - N_- \widehat{\zeta}.$$

Observe that  $L^H$  is a  $P$ -martingale strongly  $P$ -orthogonal to  $M$ ; the minimality of  $\widehat{P}$  will therefore imply that  $N$  is — as desired — a  $\widehat{P}$ -martingale strongly  $\widehat{P}$ -orthogonal to  $X$ .

**Lemma 8.** *Define the process  $N$  by (3.7). Then*

$$(3.9) \quad N \in \mathcal{M}^{2+\eta}(P) \quad \text{for every } \eta < \varepsilon,$$

$N$  is a  $\widehat{P}$ -martingale strongly  $\widehat{P}$ -orthogonal to  $X$ , and

$$N_- \widehat{\zeta} \in L^2(M).$$

**Proof.** Due to the special assumption (3.4), the process  $\widehat{Z}^0$  is strictly positive and continuous so that  $N$  is well-defined. Moreover,

$$\frac{1}{\widehat{Z}_t^0} = \frac{1}{\widehat{E}[\widehat{Z}_T | \mathcal{F}_t]} \leq \widehat{E} \left[ \frac{1}{\widehat{Z}_T} \middle| \mathcal{F}_t \right]$$

by Jensen's inequality, and therefore

$$\sup_{0 \leq t \leq T} \frac{1}{\widehat{Z}_t^0} \in \mathcal{L}^r(P) \quad \text{for every } r < \infty$$

due to (3.2), Doob's inequality and (3.1). Since  $L^H \in \mathcal{M}^{2+\varepsilon}(P)$  by Lemma 6, we conclude that

$$[N]_T = \int_0^T \frac{1}{(\widehat{Z}_s^0)^2} d[L^H]_s \leq [L^H]_T \sup_{0 \leq t \leq T} \frac{1}{(\widehat{Z}_t^0)^2} \in \mathcal{L}^{1+\delta}(P)$$

for every  $\delta < \frac{\varepsilon}{2}$  by the Burkholder-Davis-Gundy inequality, and this proves (3.9). Since  $L^H$  is strongly  $P$ -orthogonal to  $M$ , so is  $N$ , and (3.9) and the minimality of  $\widehat{P}$  imply that  $N$  is a  $\widehat{P}$ -martingale strongly  $\widehat{P}$ -orthogonal to  $X$ ; see Theorem (3.5) of Föllmer/Schweizer (1991). Using the fact that  $\widehat{\zeta} \in L^r(M)$  for every  $r < \infty$  finally yields

$$\int_0^T N_{s-} \widehat{\zeta}_s^{\text{tr}} \sigma_s \widehat{\zeta}_s N_{s-} dB_s \leq \int_0^T \widehat{\zeta}_s^{\text{tr}} \sigma_s \widehat{\zeta}_s dB_s \left( \sup_{0 \leq s \leq T} |N_s|^2 \right) \in \mathcal{L}^{1+\delta}(P)$$

for every  $\delta < \frac{\varepsilon}{2}$  which shows that  $N_- \widehat{\zeta}$  is indeed in  $L^2(M)$ .

**q.e.d.**

**Proof of Theorem 7.** We first observe that due to Lemma 6,  $L^H$  is in  $\mathcal{M}^{2+\varepsilon}(P)$  and strongly  $P$ -orthogonal to  $M$ . Since  $\widehat{P}$  is the minimal martingale measure and  $X$  is continuous, Theorem (3.5) of Föllmer/Schweizer (1991) implies that  $L^H$  is a  $\widehat{P}$ -martingale strongly  $\widehat{P}$ -orthogonal to  $X$ . This justifies in particular the second expression for  $\widehat{V}$ . Due to (3.2), we even have

$$L^H \in \mathcal{M}^{2+\eta}(\widehat{P}) \quad \text{for every } \eta < \varepsilon.$$

Since  $\widehat{K}$  is bounded,  $\Theta = L^2(M)$  and so the process  $\xi^{(c)} = \xi^H - N_- \widehat{\zeta}$  is in  $\Theta$  by Lemma 8. It therefore remains to show that  $\xi^{(c)}$  is optimal and satisfies (3.5).

Let us first argue the second point. By Lemma 8,  $N$  is a  $\widehat{P}$ -martingale strongly  $\widehat{P}$ -orthogonal to  $X$ . The special assumption (3.4) and the continuity of  $X$  thus yield

$$[N, \widehat{Z}^0] = \int \widehat{\zeta}^{\text{tr}} d[N, X] = \int \widehat{\zeta}^{\text{tr}} d\langle N, X \rangle^{\widehat{P}} = 0$$

and therefore

$$\begin{aligned} (3.10) \quad N\widehat{Z}^0 &= N_0 E[\widehat{Z}_T^2] + \int N_- \widehat{\zeta} dX + \int \widehat{Z}^0 dN \\ &= H_0 - c + \int (\xi^H - \xi^{(c)}) dX + L^H \\ &= \widehat{V} - c - \int \xi^{(c)} dX \end{aligned}$$

by the definitions of  $\xi^{(c)}$  and  $N$ . Since  $\widehat{Z}^0$  is continuous, we deduce that

$$\xi^{(c)} = \xi^H - N_- \widehat{\zeta} = \xi^H - \frac{\widehat{\zeta}}{\widehat{Z}^0} N_- \widehat{Z}^0$$

satisfies (3.5).

Now we turn to the optimality question. From (3.10) and the definitions of  $\widehat{Z}^0$  and  $\widehat{V}$ , we obtain in particular

$$(3.11) \quad H - c - G_T(\xi^{(c)}) = N_T \widehat{Z}_T^0 = N_T \frac{d\widehat{P}}{dP}.$$

For every  $\vartheta \in \Theta$ , the strong  $\widehat{P}$ -orthogonality of  $N$  and  $X$  implies that  $NG(\vartheta)$  is a local  $\widehat{P}$ -martingale null at 0. In addition,

$$\sup_{0 \leq t \leq T} |N_t G_t(\vartheta)| \in \mathcal{L}^{1+\delta}(P)$$

for every  $\delta < \frac{\varepsilon}{2}$  by (3.9) and Hölder's inequality, hence

$$\sup_{0 \leq t \leq T} |N_t G_t(\vartheta)| \in \mathcal{L}^1(\widehat{P})$$

by (3.1), and so  $NG(\vartheta)$  is even a true  $\widehat{P}$ -martingale. This implies that

$$(3.12) \quad E \left[ \left( H - c - G_T(\xi^{(c)}) \right) G_T(\vartheta) \right] = \widehat{E}[N_T G_T(\vartheta)] = 0 \quad \text{for every } \vartheta \in \Theta$$

and thus proves the optimality of  $\xi^{(c)}$ .

**q.e.d.**

As mentioned above, the solution (3.5) under the special assumption (3.4) was already given by Hipp (1993) in the special case where  $X$  is a one-dimensional Itô process and the filtration  $\mathbb{F}$  is generated by a two-dimensional Brownian motion. However, Hipp's result is not complete: he simply defined  $\xi^{(c)}$  by (3.5) and verified algebraically the optimality property (3.12), but he did not show that  $\xi^{(c)}$  is in  $\Theta$ . Moreover, the alternative description of  $\xi^{(c)}$  in (3.8) — which is needed for showing that  $\xi^{(c)}$  is in  $\Theta$  — was also not given in Hipp (1993).

**Remark.** The special assumption (3.4) can also be formulated in a different manner which can be used as the starting point for a more general approach. Recall first that the *variance-optimal martingale measure*  $\tilde{P}$  is defined as that local martingale measure for  $X$  whose density with respect to  $P$  has minimal  $\mathcal{L}^2(P)$ -norm.  $\tilde{P}$  is unique, if it exists, and equivalent to  $P$  if  $X$  is continuous; see Theorem 1.3 of Delbaen/Schachermayer (1996). The important property of  $\tilde{P}$  is now that its density with respect to  $P$  always satisfies (3.4); see Lemma 2.2 of Delbaen/Schachermayer (1996). Thus the special assumption (3.4) is equivalent to saying that the minimal martingale measure  $\hat{P}$  should coincide with the variance-optimal martingale measure  $\tilde{P}$ . One can use this reformulation of (3.4) to extend the results of the present paper by replacing  $\hat{P}$  with  $\tilde{P}$ . This leads to a representation for  $\xi^{(c)}$  rather similar to (3.5), but the techniques for proving this are completely different from those in the present paper. For this generalization and its ramifications, we refer to Rheinländer/Schweizer (1996).

The next result provides an explicit expression for the minimal risk in the optimization problem (1.1). It extends formula (10) of Hipp (1993) to our general situation; see also Corollary 9 of Schweizer (1994).

**Corollary 9.** *Under the assumptions of Theorem 7, the minimal quadratic risk is given by*

$$J_0 = E \left[ \left( H - c - G_T(\xi^{(c)}) \right)^2 \right] = \frac{(H_0 - c)^2 + E[(L_0^H)^2]}{E[\hat{Z}_T^2]} + \hat{E} \left[ \int_0^T \frac{1}{\hat{Z}_s^0} d[L^H]_s \right].$$

**Proof.** (3.11) and (3.12) yield

$$\begin{aligned} J_0 &= \hat{E} \left[ N_T \left( H - c - G_T(\xi^{(c)}) \right) \right] \\ &= \hat{E} \left[ N_T \left( H_0 - c + G_T(\xi^H - \xi^{(c)}) + L_T^H \right) \right] \\ &= \hat{E}[N_T(H_0 - c)] + \hat{E}[N_T L_T^H]. \end{aligned}$$

According to Lemma 8,  $N$  is a  $\hat{P}$ -martingale, and this implies

$$\hat{E}[N_T(H_0 - c)] = (H_0 - c)\hat{E}[N_0] = \frac{1}{E[\hat{Z}_T^2]} \left( (H_0 - c)^2 + (H_0 - c)\hat{E}[L_0^H] \right) = \frac{1}{E[\hat{Z}_T^2]} (H_0 - c)^2,$$

since  $\widehat{E}[L_0^H] = E[L_0^H] = 0$  because  $\widehat{P} = P$  on  $\mathcal{F}_0$ . Finally, we also know from the preceding arguments that  $L^H$  and  $N$  are both in  $\mathcal{M}^2(\widehat{P})$ . Thus

$$\widehat{E}[N_T L_T^H] = \widehat{E}[N_0 L_0^H] + \widehat{E}[[N, L^H]_T] = \frac{\widehat{E}[(L_0^H)^2]}{E[\widehat{Z}_T^2]} + \widehat{E}\left[\int_0^T \frac{1}{\widehat{Z}_s^0} d[L^H]_s\right]$$

by the definition of  $N$ , and putting everything together completes the proof if we note again that  $\widehat{P} = P$  on  $\mathcal{F}_0$ .

**q.e.d.**

## 4. Examples

In this section, we provide several examples of two different basic types. We first exhibit a number of situations where the special assumption (3.4) is satisfied and show how the optimal strategy  $\xi^{(c)}$  in Theorem 7 and the minimal quadratic risk  $J_0$  in Corollary 9 can be computed more explicitly. Then we go on to justify the expression “special” for the assumption (3.4) by showing that for models with sufficiently random coefficients, it will typically not be satisfied.

### 4.1. Approximating a riskless asset

In our first example, we take for  $H$  the constant 1 and  $c = 0$ . In mathematical terms, we thus look for the projection in  $\mathcal{L}^2(P)$  of 1 on the space  $G_T(\Theta)$  of stochastic integrals with respect to  $X$ . Financially speaking, we want to approximate the riskless payoff 1 by the final value of a self-financing strategy with initial capital 0 by investing in the risky assets  $X^1, \dots, X^d$ ; the quality of the approximation is measured by a quadratic loss function. Note that it is the constraint of zero initial capital that makes this problem interesting: although there is a riskless asset, starting from  $c = 0$  forces us to invest in the risky assets as well.

Under the assumptions of Theorem 7, the solution of (1.1) with  $H = 1$  and  $c = 0$  is given by

$$(4.1) \quad \xi_t^{(0)} = -\mathcal{E}\left(\int \frac{\widehat{\zeta}}{\widehat{Z}^0} dX\right)_t \frac{\widehat{\zeta}_t}{\widehat{Z}_t^0}, \quad 0 \leq t \leq T$$

with a minimal quadratic risk of

$$J_0 = \frac{1}{E[\widehat{Z}_T^2]}.$$

To see this, note first that  $\widehat{V} \equiv 1$  and  $\xi^H \equiv 0$  imply by Theorem 7 that

$$1 - \int \xi^{(0)} dX = 1 + \int \frac{\widehat{\zeta}}{\widehat{Z}^0} (1 - \int \xi^{(0)} dX) dX,$$

hence  $1 - \int \xi^{(0)} dX = \mathcal{E} \left( \int \frac{\widehat{\zeta}}{\widehat{Z}^0} dX \right)$ , and plugging this into (3.5) yields (4.1). The expression for  $J_0$  follows from Corollary 9 by noting that  $H_0 = 1$  and  $L^H \equiv 0$ .

We remark that  $\xi^{(0)}$  is the *approximate arbitrage opportunity portfolio* introduced by Gouriéroux/Laurent (1995) in a discrete-time framework; (4.1) thus provides an explicit expression for this quantity under the assumptions of Theorem 7. Moreover, (4.1) and the definition of  $\xi^{(0)}$  also show that  $-\frac{\widehat{\zeta}}{\widehat{Z}^0}$  is an *adjustment process* in the sense of Schweizer (1996).

**Example 1.** Suppose now that  $H$  is *attainable* in the sense that  $L_T^H = 0$  in the Föllmer-Schweizer decomposition of  $H$  so that

$$(4.2) \quad H = H_0 + \int_0^T \xi_s^H dX_s$$

with  $H_0 \in \mathbb{R}$  and  $\xi^H \in \Theta$ . If we were free to choose not only  $\vartheta$ , but also the initial capital  $c$  in the optimization problem (1.1), the solution would trivially be given by  $c = H_0$  and  $\xi^{(H_0)} = \xi^H$ , with a quadratic risk of 0. For an exogenously given  $c$ , however, the solution is different, namely

$$\xi_t^{(c)} = \xi_t^H - (H_0 - c) \mathcal{E} \left( \int \frac{\widehat{\zeta}}{\widehat{Z}^0} dX \right)_t \frac{\widehat{\zeta}_t}{\widehat{Z}_t^0} = \xi_t^H + (H_0 - c) \xi_t^{(0)} \quad , \quad 0 \leq t \leq T.$$

To see this, simply observe that by the projection theorem, the optimal strategy  $\xi^{(c)}$  for (1.1) is linear as a function of  $H$ . Due to (4.2), the attainable part  $G_T(\xi^H)$  can be perfectly hedged by  $\xi^H$ , and the remainder  $H_0 - c$  is a constant which can be approximated as in subsection 4.1.

#### 4.2. The case where $\widehat{K}_T$ is deterministic

We next provide a whole class of examples where the assumptions of Theorem 7 are satisfied. Let us suppose that  $X$  is a continuous semimartingale satisfying the structure condition (SC). Continuity of  $\widehat{K}$  then yields

$$\widehat{Z} = \mathcal{E} \left( - \int \widehat{\lambda} dM \right) = \mathcal{E} \left( - \int \widehat{\lambda} dX + \widehat{K} \right) = \mathcal{E} \left( - \int \widehat{\lambda} dX \right) e^{\widehat{K}}$$

by Yor's formula and in particular

$$\frac{d\widehat{P}}{dP} = \widehat{Z}_T = e^{\widehat{K}_T} \mathcal{E} \left( - \int \widehat{\lambda} dX \right)_T = e^{\widehat{K}_T} \left( 1 - \int_0^T \mathcal{E} \left( - \int \widehat{\lambda} dX \right)_s \widehat{\lambda}_s dX_s \right).$$

If we now assume that the final value  $\widehat{K}_T$  of the mean-variance tradeoff process is *deterministic*, then  $\widehat{K}$  is clearly bounded and  $\widehat{Z}_T$  can be written as the sum of the constant  $e^{\widehat{K}_T}$  and the stochastic integral with respect to  $X$  of

$$(4.3) \quad \widehat{\zeta} := -e^{\widehat{K}_T} \mathcal{E} \left( - \int \widehat{\lambda} dX \right) \widehat{\lambda}.$$

This shows that the special assumption (3.4) is satisfied, since boundedness of  $\widehat{K}$  readily implies that  $\widehat{\zeta}$  in (4.3) is in  $\Theta$ . Moreover,  $\mathcal{E} \left( - \int \widehat{\lambda} dX \right)$  is a  $\widehat{P}$ -martingale, hence

$$\widehat{Z}_t^0 = \widehat{E}[\widehat{Z}_T | \mathcal{F}_t] = e^{\widehat{K}_T} \mathcal{E} \left( - \int \widehat{\lambda} dX \right)_t, \quad 0 \leq t \leq T,$$

and this yields

$$(4.4) \quad -\frac{\widehat{\zeta}_t}{\widehat{Z}_t^0} = \widehat{\lambda}_t, \quad 0 \leq t \leq T.$$

For a fixed random variable  $H$ , the solution of (1.1) is now given by

$$(4.5) \quad \xi_t^{(c)} = \xi_t^H + \widehat{\lambda}_t \left( \widehat{V}_{t-} - c - \int_0^t \xi_s^{(c)} dX_s \right), \quad 0 \leq t \leq T$$

with a minimal quadratic risk of

$$\begin{aligned} J_0 &= e^{-\widehat{K}_T} \left( (H_0 - c)^2 + E[(L_0^H)^2] + E \left[ \int_0^T e^{\widehat{K}_s} d[L^H]_s \right] \right) \\ &= e^{-\widehat{K}_T} \left( (H_0 - c)^2 + E[(L_0^H)^2] + E \left[ \int_0^T e^{\widehat{K}_s} d\langle L^H \rangle_s^P \right] \right); \end{aligned}$$

this extends Theorem 3 and Corollary 9 of Schweizer (1994) to our present situation. To obtain  $J_0$  from Corollary 9, one uses the explicit expressions for  $\widehat{Z}$  and  $\widehat{Z}^0$ , the fact that  $\frac{1}{Z}$  is the density process of  $P$  with respect to  $\widehat{P}$ , and Theorem VI.61 of Dellacherie/Meyer (1982).

**Remarks. 1)** The significance of the special assumption (3.4) and the fact that it holds for a deterministic mean-variance tradeoff were first pointed out by Hipp (1993) in the context of a diffusion model. He also showed how the feedback formula (3.5) then reduces to (4.5). In a general framework, (4.5) was also obtained in Schweizer (1994) where  $X$  was allowed to have jumps, but the *entire* MVT process  $\widehat{K}$  had to be deterministic in return.

2) The following subsections will show that the special assumption (3.4) can also hold in cases where  $\widehat{K}_T$  is not deterministic. Nevertheless, this is rather the exception than the rule; if  $\widehat{K}_T$  contains some randomness in addition to the one generated by  $X$  itself, (3.4) will typically fail. For a more precise statement of this assertion, see subsection 4.5.

### 4.3. An “almost complete” diffusion model

Consider next the case where  $X$  is given by

$$\frac{dX_t^i}{X_t^i} = (b_t^i - r_t) dt + \sum_{j=1}^n v_t^{ij} dW_t^j = m_t^i dt + \sum_{j=1}^n v_t^{ij} dW_t^j$$

for a Brownian motion  $W$  in  $\mathbb{R}^n$ . The processes  $b$  and  $v$  describe the appreciation rates and volatilities of  $d$  stocks  $S^1, \dots, S^d$ , while  $r$  is the riskless interest rate paid continuously by the bond  $S^0$ . The discounted prices are then given by  $X^i = \frac{S^i}{S^0}$ . This generalized version of the Black-Scholes model was studied in great detail by a number of authors; see for instance the recent paper by Cvitanić/Karatzas (1993).

If we assume (as is usually done in this model) that  $d \leq n$  and that the matrix  $v_t$  has full rank for all  $t$   $P$ -a.s., a straightforward computation yields

$$(4.6) \quad \int \widehat{\lambda} dM = \int (b - r\mathbf{1})^{\text{tr}} (v v^{\text{tr}})^{-1} v dW$$

and

$$\widehat{K} = \left\langle \int \widehat{\lambda} dM \right\rangle = \int (b_s - r_s \mathbf{1})^{\text{tr}} (v_s v_s^{\text{tr}})^{-1} (b_s - r_s \mathbf{1}) ds$$

with  $\mathbf{1} := (1 \dots 1)^{\text{tr}} \in \mathbb{R}^d$ . Thus we see that boundedness of the MVT process is guaranteed by the standard assumption that the “relative risk” process or *market price of risk*,

$$v_s^{\text{tr}} (v_s v_s^{\text{tr}})^{-1} (b_s - r_s \mathbf{1}) \quad , \quad 0 \leq s \leq T,$$

is bounded. For the one-dimensional case  $d = 1$ , this reduces to the condition that

$$\frac{b_s - r_s}{v_s} = \frac{m_s}{v_s} \quad , \quad 0 \leq s \leq T$$

should be bounded.

Up to now, we have deliberately not been specific about the filtration under consideration. If we choose the  $P$ -augmentation  $\mathbb{F}^W$  of the filtration generated by  $W$  and if  $d = n$ , then it is well known that the resulting model is *complete* and that every sufficiently integrable random variable can be written as the sum of a constant and a stochastic integral with respect to  $X$ . In fact, Itô’s representation theorem yields a representation as a constant plus a stochastic



integral of  $W$ , and this can be rewritten in terms of  $X$  by using the Bayes rule and (since  $d = n$ ) the invertibility of  $v_t$ . For details, see for instance Proposition 5.8.6 of Karatzas/Shreve (1988). In particular, the special assumption (3.4) is then automatically satisfied. This is of course not surprising since in the complete case, there is only one equivalent martingale measure, and so it must be at the same time minimal and variance-optimal.

However, completeness is rather too restrictive as an assumption, and it is fortunate that we do not really require its full strength to obtain (3.4). Suppose for instance that  $d = n$  as before, but that the filtration  $\mathbb{F} \supseteq \mathbb{F}^W$  is arbitrary. If we assume that

$$(4.7) \quad \text{the market price of risk } v^{\text{tr}}(vv^{\text{tr}})^{-1}(b - r\mathbf{1}) = v^{\text{tr}}(vv^{\text{tr}})^{-1}m \text{ is adapted to } \mathbb{F}^W,$$

then (4.6) implies that  $\widehat{Z}_T = \mathcal{E}\left(-\int \widehat{\lambda} dM\right)_T$  is  $\mathcal{F}_T^W$ -measurable and therefore representable as a constant plus a stochastic integral of  $X$  by the same argument as above. Thus the special assumption (3.4) is satisfied and we can apply Theorem 7 and Corollary 9 to determine the optimal strategy and the minimal quadratic risk for any given random variable  $H$ . Note that the incompleteness in this model is due to the fact that the filtration  $\mathbb{F}$  contains more information than is given by the discounted prices  $X$  or the prices  $S$ . A simple example would be a payoff  $H$  depending on an additional source of randomness which does not influence the coefficients  $b, r, v$  of our model; see subsection 4.4 for a more concrete situation.

**Example 2.** Take  $d = 1$  and let  $X$  be the solution of the stochastic differential equation

$$(4.8) \quad \frac{dX_t}{X_t} = m(t, X_t) dt + v(t, X_t) dW_t.$$

Under suitable assumptions on the functions  $m$  and  $v$ , (4.8) has a unique strong solution which is therefore adapted to the filtration  $\mathbb{F}^W$ . This implies that condition (4.7) is satisfied, hence (3.4) holds for any filtration  $\mathbb{F}$  containing  $\mathbb{F}^W$ , and we can apply the results of section 3. As in Hipp (1993),  $\mathbb{F}$  could for instance be generated by  $W$  and a second Brownian motion  $W'$  independent of  $W$ . We remark in passing that the SDE (4.8) need not be autonomous for this result; see Hipp (1993).

**Example 3.** Somewhat more generally, the special assumption (3.4) is also satisfied if  $\widehat{Z}_T$  is  $\mathbb{F}_T^X$ -measurable and if  $X$  has the predictable representation property for its own filtration. This is for instance the case if  $X$  is given by

$$\frac{dX_t^i}{X_t^i} = m_t^i dt + \sum_{j=1}^d v^{ij}(X_t) dW_t^j$$

with  $m^i$  bounded and adapted to  $\mathbb{F}^X$  and  $v^{ij}$  sufficiently regular; see Example 4.1.(c) of Jacod (1977) for details. That  $\widehat{Z}_T$  is  $\mathbb{F}_T^X$ -measurable follows from the representation

$$\widehat{Z}_T = e^{\widehat{K}_T} \mathcal{E}\left(-\int \widehat{\lambda} dX\right)_T,$$

the explicit expressions for  $\int \widehat{\lambda} dX$  and  $\widehat{K}$  and the measurability assumptions on  $m$  and  $v$ .

**Example 4.** As in Hipp (1993), let us consider the case where  $d = 1$  and where the filtration  $\mathbb{F}$  is generated by  $W$  and an independent Brownian motion  $W'$ . In the general framework of this subsection, any  $P$ -martingale orthogonal to  $M$  is then a stochastic integral of  $W'$ , and so the orthogonal component in the Föllmer-Schweizer decomposition of  $H$  has the form

$$L^H = L_0^H + \int \eta^H dW'$$

for some integrand  $\eta^H$ . Moreover,  $\mathcal{F}_0$  is trivial and so  $L_0^H = E[L_0^H] = 0$ . The minimal quadratic risk is therefore given by

$$J_0 = \frac{(H_0 - c)^2}{E[\widehat{Z}_T^2]} + \widehat{E} \left[ \int_0^T \frac{1}{\widehat{Z}_s^0} (\eta_s^H)^2 ds \right];$$

note that this corrects an error in formulas (10) and (18) of Hipp (1993).

#### 4.4. Markovian stochastic volatility models

Let us now examine Theorem 7 in the context of a fairly general stochastic volatility model. We consider the stochastic differential equation

$$(4.9) \quad \frac{dX_t}{X_t} = m(t, X_t, Y_t) dt + v(t, X_t, Y_t) dW_t,$$

where  $Y$  is an additional random factor (with values in  $\mathcal{Y}$ , say) which influences the evolution of  $X$ . We shall assume that  $m$ ,  $v$  and  $Y$  are such that (4.9) has a unique strong solution and take as  $\mathbb{F}$  the  $P$ -augmentation of the filtration generated by  $X$  and  $Y$ . Since  $Y$  will typically be a non-traded quantity, the model (4.9) is incomplete in general.

Let us further assume that  $(X, Y)$  is a Markov process under  $P$ ; explicit examples will be given below. If we restrict our attention to contingent claims of the form  $H = h(X_T, Y_T)$ , finding the Föllmer-Schweizer decomposition of  $H$  is straightforward. Since

$$(4.10) \quad \widehat{\lambda}_t = \frac{m(t, X_t, Y_t)}{X_t v^2(t, X_t, Y_t)} \quad , \quad 0 \leq t \leq T$$

and

$$(4.11) \quad \frac{d\widehat{P}}{dP} = \widehat{Z}_T = \mathcal{E} \left( - \int \frac{m(s, X_s, Y_s)}{v(s, X_s, Y_s)} dW_s \right)_T,$$

the Bayes rule and the Markov property of  $(X, Y)$  under  $P$  imply that

$$\widehat{V}_t = \widehat{E}[H|\mathcal{F}_t] = \widehat{E}[h(X_T, Y_T)|\mathcal{F}_t] = \widehat{v}(t, X_t, Y_t)$$

for a function  $\widehat{v} : [0, T] \times \mathbb{R}_+ \times \mathcal{Y} \rightarrow \mathbb{R}$ . This can be used to provide explicit expressions for  $\xi^H$  and  $L^H$  as the following two examples will illustrate. For more detailed studies of the function  $\widehat{v}$  in particular models, see for instance Hull/White (1987) or Pham/Touzi (1996).

**Example 5.** Let  $Y$  be given as a strong solution of the stochastic differential equation

$$(4.12) \quad dY_t = a(t, X_t, Y_t) dt + b(t, X_t, Y_t) dW'_t$$

with an independent Brownian motion  $W'$  under  $P$ . As a special case, this includes the well-known stochastic volatility model introduced by Hull/White (1987). Under regularity assumptions on the coefficient functions  $m, v, a, b$ , the function  $\widehat{v}$  is  $C^{1,2,2}$  on  $[0, T] \times (0, \infty) \times \mathbb{R}$  and the unique solution of the partial differential equation

$$\frac{\partial \widehat{v}}{\partial t} + a \frac{\partial \widehat{v}}{\partial y} + \frac{1}{2} \left( b^2 \frac{\partial^2 \widehat{v}}{\partial y^2} + x^2 v^2 \frac{\partial^2 \widehat{v}}{\partial x^2} \right) = 0 \quad \text{on } (0, T) \times (0, \infty) \times \mathbb{R}$$

with the boundary condition

$$\widehat{v}(T, x, y) = h(x, y) \quad \text{for all } x, y \in \mathbb{R}_+ \times \mathbb{R};$$

see for instance Theorem 5.3 of Friedman (1975). Applying Itô's formula to  $\widehat{V}$  then yields

$$\xi_t^H = \frac{\partial \widehat{v}}{\partial x}(t, X_t, Y_t) \quad , \quad 0 \leq t \leq T$$

and

$$L_t^H = \int_0^t \frac{\partial \widehat{v}}{\partial y}(s, X_s, Y_s) b(s, X_s, Y_s) dW'_s \quad , \quad 0 \leq t \leq T.$$

**Example 6.** Let  $Y$  be a Markov jump process with countable state space  $\mathcal{Y} = \{y_1, y_2, \dots\}$  and generator  $\Pi = (\pi_{ij})$  with  $Y$  and  $W$  independent. This model and generalizations of it were also studied by Di Masi/Kabanov/Runggaldier (1994) and Swishchuk (1995). If we denote by  $(T_n)$  the sequence of jump times of  $Y$ , the random measure describing the jumps of  $Y$  is given by

$$\nu([0, t] \times \Gamma) = \sum_{n=1}^{\infty} I_{\{Y_{T_n} \in \Gamma, T_n \leq t\}}$$

and its predictable compensator is

$$\nu^p(dt, y_j) = \sum_{i \neq j} \pi_{ij} I_{\{Y_t = y_i\}} dt.$$

In this case,  $\widehat{v}(t, x, y)$  is the solution of

$$\frac{\partial \widehat{v}}{\partial t} + \Pi \widehat{v} + \frac{1}{2} x^2 v^2 \frac{\partial^2 \widehat{v}}{\partial x^2} = 0 \quad \text{on } (0, T) \times (0, \infty) \times \mathcal{Y}$$

with boundary condition

$$\widehat{v}(T, x, y) = h(x, y) \quad \text{for all } x, y \in \mathbb{R}_+ \times \mathcal{Y}.$$

As before, an application of Itô's formula leads to the explicit expressions

$$\xi_t^H = \frac{\partial \widehat{v}}{\partial x}(t, X_t, Y_{t-}) \quad , \quad 0 \leq t \leq T$$

and

$$L_t^H = \int_0^t \sum_{j=1}^{\infty} (\widehat{v}(s, X_s, y_j) - \widehat{v}(s, X_s, Y_{s-})) (\nu - \nu^p)(ds, y_j) \quad , \quad 0 \leq t \leq T.$$

This ends Example 6.

In connection with the special assumption (3.4), we next have to study the Föllmer-Schweizer decomposition of  $\widehat{Z}_T$ . We do this in two special cases which represent in a way two extreme opposites. The first result is due to C. Hipp (1996).

**Proposition 10.** *Suppose that  $m$  and  $v$  in (4.9) do not depend on  $y$ . Then the special assumption (3.4) is satisfied and (under suitable regularity conditions on  $m, v$ ) the integrand  $\widehat{\zeta}$  in (3.3) is explicitly given by*

$$(4.13) \quad \widehat{\zeta}_t = \widehat{Z}_t \left( \frac{\partial g}{\partial x}(t, X_t) - \widehat{\lambda}_t g(t, X_t) \right) \quad , \quad 0 \leq t \leq T,$$

where  $g : [0, T] \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is the unique solution of the partial differential equation

$$(4.14) \quad \frac{\partial g}{\partial t} + \frac{1}{2} x^2 v^2 \frac{\partial g}{\partial x^2} - x m \frac{\partial g}{\partial x} + \frac{m^2}{v^2} g = 0 \quad \text{on } (0, T) \times (0, \infty)$$

with the boundary condition

$$g(T, x) = 1 \quad \text{for all } x \in \mathbb{R}_+.$$

Moreover, we have

$$(4.15) \quad -\frac{\widehat{\zeta}_t}{\widehat{Z}_t^0} = \widehat{\lambda}_t - \frac{\frac{\partial g}{\partial x}(t, X_t)}{g(t, X_t)} \quad , \quad 0 \leq t \leq T.$$

**Proof.** Sufficient conditions for the existence and uniqueness of  $g$  are for instance that  $m$  and  $v$  are both bounded and uniformly Lipschitz in  $(t, x)$ , together with a uniform lower bound on  $v$ ; see Theorem 5.3 of Friedman (1975). If  $m$  and  $v$  do not depend on  $y$ , the SDE for  $X$  reduces to (4.8), and so (3.4) follows as in Example 2. Now apply Itô's formula to the product  $U_t := \widehat{Z}_t g(t, X_t)$ , write  $\dot{\cdot}$  and  $\prime$  for the partial derivatives with respect to  $t$  and  $x$ , respectively, and use (4.11), (4.10) and (4.14) to obtain

$$dU_t = \widehat{Z}_t \left( \dot{g} + \frac{1}{2} g'' X_t^2 v^2 - \widehat{\lambda} g' X_t^2 v^2 \right) dt + \widehat{Z}_t g' dX_t - \widehat{Z}_t g \widehat{\lambda}_t dM_t = \widehat{Z}_t \left( g' - \widehat{\lambda} g \right) dX_t.$$

Thus  $U$  is a  $\widehat{P}$ -martingale with  $U_T = \widehat{Z}_T = \widehat{Z}_T^0$ , and this implies  $\widehat{Z}^0 \equiv U$ , hence (4.13) and (4.15).

**q.e.d.**

**Example 7.** In the even more special case where  $m$  and  $v$  also do not depend on  $x$ , we can explicitly write down the solution of the PDE (4.14) as

$$g(t, x) = g(t) = \exp \left( \int_t^T \frac{m^2(s)}{v^2(s)} ds \right),$$

and so (4.15) then reduces to (4.4). This is of course obvious since the MVT process  $\widehat{K} = \int \frac{m^2(s)}{v^2(s)} ds$  is deterministic in that case.

If  $m$  and  $v$  do not depend on the additional randomness generated by  $Y$ , we have just seen that (3.4) will hold. Let us now show that in the opposite extreme case, (3.4) will fail.

**Theorem 11.** *Suppose that  $m, v, a, b$  in (4.12) do not depend on  $x$ . If  $\widehat{K}_T$  is bounded and not deterministic, then the special assumption (3.4) is not satisfied.*

**Proof.** Since  $m, v, a, b$  do not depend on  $x$ , (4.9) and (4.12) imply that  $\widehat{K}_T = \int_0^T \frac{m^2(s, Y_s)}{v^2(s, Y_s)} ds$  is  $\mathcal{F}_T^{W'}$ -measurable. Thus we have

$$e^{\widehat{K}_T} = E \left[ e^{\widehat{K}_T} \right] + \int_0^T \nu_s dW'_s \quad P\text{-a.s.}$$

for some integrand  $\nu$  by Itô's representation theorem, and applying the product rule to the processes  $\mathcal{E} \left( - \int \widehat{\lambda} dX \right)$  and  $U := E \left[ e^{\widehat{K}_T} \right] + \int \nu dW'$  yields

$$(4.16) \quad \frac{d\widehat{P}}{dP} = \mathcal{E} \left( - \int \widehat{\lambda} dX \right)_T U_T$$

$$= E \left[ e^{\widehat{K}_T} \right] - \int_0^T U_s \mathcal{E} \left( - \int_s \widehat{\lambda} dX \right)_s \widehat{\lambda}_s dX_s + \int_0^T \mathcal{E} \left( - \int_s \widehat{\lambda} dX \right)_s \nu_s dW'_s,$$

because  $[W', X] = 0$ . Since  $\widehat{K}_T$  is bounded, so is  $U$ , and this together with (3.1) implies that (4.16) is the Föllmer-Schweizer decomposition of  $\frac{d\widehat{P}}{dP}$ : both integrands are sufficiently integrable and the last term is a  $P$ -martingale strongly  $P$ -orthogonal to  $M$ . But because  $\widehat{K}_T$  is not deterministic,  $\nu$  is different from 0, hence so is the orthogonal term in (4.16), and this shows that (3.4) is not satisfied.

**q.e.d.**

**Remark.** It is clear from the preceding argument that Theorem 11 will also hold for other models of  $Y$ ; the only property we use in the proof is that the process generating the additional randomness (in this case,  $W'$ ) has the predictable representation property.

The conditions of Theorem 11 on  $\widehat{K}_T$  are very easily satisfied; the simplest example is the case where the ratio  $\frac{m}{v}$  is a bounded function of  $t$  and  $y$  which is not constant in  $y$ . In conjunction with the preceding results, this example suggests the general rule that the special assumption (3.4) will typically be violated as soon as the crucial quantity  $\widehat{K}_T$  depends on more randomness than the evolution of  $X$  alone. In the Markovian framework of this subsection, a result of this type has been obtained by C. Hipp (1996). Another result in this direction is provided in the next subsection.

#### 4.5. A Black-Scholes model in a random environment

In order to generate now a whole class of examples where the special assumption (3.4) will fail, we consider the stochastic differential equation

$$(4.17) \quad \frac{dX_t}{X_t} = m_t dt + v_t dW_t,$$

where the processes  $m$  and  $v$  are independent of the Brownian motion  $W$  and supposed to be such that (4.17) has a unique strong solution for almost all realizations of  $m$  and  $v$ . This can be thought of as a Black-Scholes model in a random environment, where the environment is described by the stochastic processes  $m$  and  $v$ . The filtration will be the one generated by  $W$  and augmented at time 0 by the complete knowledge of  $m$  and  $v$ . Intuitively, this means that the environment is randomly chosen at time 0, and that  $X$  then evolves as a usual geometric Brownian motion whose random coefficients are determined by the realization of the environment. This is therefore a precise description of a model in which the whole randomness in the coefficients is exogenous, as opposed to coming from  $X$  itself as in (4.8).

Readers who are familiar with the paper of Delbaen/Schachermayer (1996) will of course recognize this as a jazzed-up version of their example 2.5.

**Theorem 12.** *Suppose that the processes  $m$  and  $v$  in (4.17) are independent of the Brownian motion  $W$  and that  $\frac{m}{v}$  is bounded. If  $\widehat{K}_T$  is not deterministic, the special assumption (3.4) is not satisfied.*

**Proof.** It is clear that  $\widehat{K} = \int \frac{m_s^2}{v_s^2} ds$  and  $\widehat{Z} = \mathcal{E} \left( - \int \frac{m}{v} dW \right)$  so that

$$\widehat{Z}_T = \exp \left( - \int_0^T \frac{m_s}{v_s} dW_s - \frac{1}{2} \widehat{K}_T \right).$$

Since  $m$  and  $v$  are independent of  $W$ , the conditional distribution of  $\log \widehat{Z}_T$  given  $m$  and  $v$  is therefore a normal distribution with mean  $-\frac{1}{2} \widehat{K}_T$  and variance  $\widehat{K}_T$ . This allows us to compute

$$E \left[ \widehat{Z}_T^2 \right] = E \left[ e^{\widehat{K}_T} \right]$$

by conditioning on  $m$  and  $v$ . Since  $m$  and  $v$  are both  $\mathcal{F}_0$ -measurable, the process

$$Z_t := \frac{e^{-\widehat{K}_T}}{E \left[ e^{-\widehat{K}_T} \right]} \widehat{Z}_t$$

is a strictly positive  $P$ -martingale with expectation 1. Since  $\widehat{P}$  is a martingale measure for  $X$ , the product  $ZX$  is also a  $P$ -martingale, and so we can define an equivalent martingale measure  $Q$  for  $X$  by setting  $\frac{dQ}{dP} := Z_T$ . By the same argument as above, we can compute

$$E \left[ Z_T^2 \right] = \frac{E \left[ e^{-2\widehat{K}_T} \widehat{Z}_T^2 \right]}{\left( E \left[ e^{-\widehat{K}_T} \right] \right)^2} = \frac{1}{E \left[ e^{-\widehat{K}_T} \right]}.$$

Thus Jensen's inequality implies that

$$\left\| \frac{dQ}{dP} \right\|_{\mathcal{L}^2(P)} < \left\| \frac{d\widehat{P}}{dP} \right\|_{\mathcal{L}^2(P)}$$

because  $\widehat{K}_T$  is not deterministic. Hence  $\widehat{P}$  is not variance-optimal, and so Lemma 1 of Schweizer (1996) implies that (3.4) does not hold.

**q.e.d.**

This result at the same time limits the scope of Theorem 7 and illustrates the general principle that the special assumption (3.4) is very restrictive and should not be expected

to hold if the coefficients involve more randomness than only the one arising from  $X$ . It is therefore an important question to examine the structure of the solution  $\xi^{(e)}$  of the optimization problem (1.1) also in the general case where (3.4) is not satisfied. Recent results on this problem can be found in Rheinländer/Schweizer (1996).

**Remark.** Although both Theorem 11 and Theorem 12 come to the same conclusion that (3.4) fails, the structure of the two examples is rather different. In Theorem 12, all the independent randomness is already present at the initial date 0, and it suffices to modify  $\widehat{P}$  on  $\mathcal{F}_0$  to construct a martingale measure  $Q$  with smaller  $\mathcal{L}^2(P)$ -norm. (Actually, it is not hard to check that  $Q$  is the variance-optimal martingale measure.) In particular,  $Q$  and  $\widehat{P}$ , hence also  $Q$  and  $P$ , differ on  $\mathcal{F}_0$ . In Theorem 11,  $\mathcal{F}_0$  is trivial and the additional randomness induced by  $Y$  only comes into play after some time (depending also on  $m$  and  $v$ ). This explains to some extent why the proof of Theorem 12 uses far less structure than the one of Theorem 11; in particular, there is no need to have a representation theorem available.

## References

- J. Cvitanić and I. Karatzas (1993), “Hedging Contingent Claims with Constrained Portfolios”, *Annals of Applied Probability* 3, 652–681
- M. H. A. Davis (1994), “A General Option Pricing Formula”, *preprint, Imperial College, London*
- F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer and C. Stricker (1996), “Weighted Norm Inequalities and Hedging in Incomplete Markets”, *to appear in Finance and Stochastics*
- F. Delbaen and W. Schachermayer (1995), “The Existence of Absolutely Continuous Local Martingale Measures”, *Annals of Applied Probability* 5, 926–945
- F. Delbaen and W. Schachermayer (1996), “The Variance-Optimal Martingale Measure for Continuous Processes”, *BERNOULLI* 2, 81–105
- C. Dellacherie and P.-A. Meyer (1982), “Probabilities and Potential B”, *North-Holland*
- G. B. Di Masi, Yu. Kabanov and W. J. Runggaldier (1994), “Mean-Variance Hedging of Options on Stocks with Markov Volatilities”, *Theory of Probability and its Applications* 39, 172–182
- N. El Karoui, S. Peng and M.-C. Quenez (1997), “Backward Stochastic Differential Equations in Finance”, *Mathematical Finance* 7, 1–71



N. El Karoui and M.-C. Quenez (1995), “Dynamic Programming and Pricing of Contingent Claims in an Incomplete Market”, *SIAM Journal on Control and Optimization* 33, 29–66

H. Föllmer and M. Schweizer (1991), “Hedging of Contingent Claims under Incomplete Information”, in: *M. H. A. Davis and R. J. Elliott (eds.), “Applied Stochastic Analysis”, Stochastics Monographs, Vol. 5, Gordon and Breach, London/New York*, 389–414

H. Föllmer and D. Sondermann (1986), “Hedging of Non-Redundant Contingent Claims”, in: *W. Hildenbrand and A. Mas-Colell (eds.), Contributions to Mathematical Economics*, 205–223

A. Friedman (1975), “Stochastic Differential Equations and Applications, Volume 1”, *Academic Press*

C. Gouriéroux and J. P. Laurent (1995), “Dynamic Hedge in Discrete Time”, *preprint, CREST, Paris*

C. Gouriéroux, J. P. Laurent and H. Pham (1996), “Mean-Variance Hedging and Numéraire”, *preprint, Université de Marne-la-Vallée*

J. M. Harrison and S. R. Pliska (1981), “Martingales and Stochastic Integrals in the Theory of Continuous Trading”, *Stochastic Processes and their Applications* 11, 215–260

C. Hipp (1993), “Hedging General Claims”, *Proceedings of the 3rd AFIR Colloquium, Rome, Vol. 2*, 603–613

C. Hipp (1996), “Hedging and Insurance Risk”, *preprint No. 1/96, University of Karlsruhe*

J. Hull and A. White (1987), “The Pricing of Options on Assets with Stochastic Volatilities”, *Journal of Finance* XLII, 281–300

J. Jacod (1977), “A General Theorem of Representation for Martingales”, *Proceedings of Symposia in Pure Mathematics* 31, *American Mathematical Society*, 37–53

J. Jacod (1979), “Calcul Stochastique et Problèmes de Martingales”, *Lecture Notes in Mathematics* 714, *Springer*

I. Karatzas and S.-G. Kou (1996), “On the Pricing of Contingent Claims under Constraints”, *Annals of Applied Probability* 6, 321–369

I. Karatzas and S. E. Shreve (1988), “Brownian Motion and Stochastic Calculus”, *Springer*

P. Monat and C. Stricker (1994), “Fermeture de  $G_T(\Theta)$  et de  $\mathcal{L}^2(\mathcal{F}_0)+G_T(\Theta)$ ”, *Séminaire*

*de Probabilités XXVIII, Lecture Notes in Mathematics 1583, Springer, 189–194*

P. Monat and C. Stricker (1995), “Föllmer-Schweizer Decomposition and Mean-Variance Hedging for General Claims”, *Annals of Probability* 23, 605–628

H. Pham and N. Touzi (1996), “Equilibrium State Prices in a Stochastic Volatility Model”, *Mathematical Finance* 6, 215–242

T. Rheinländer and M. Schweizer (1996), “On  $L^2$ -Projections on a Space of Stochastic Integrals”, *to appear in Annals of Probability*

M. Schweizer (1994), “Approximating Random Variables by Stochastic Integrals”, *Annals of Probability* 22, 1536–1575

M. Schweizer (1995), “On the Minimal Martingale Measure and the Föllmer-Schweizer Decomposition”, *Stochastic Analysis and Applications* 13, 573–599

M. Schweizer (1996), “Approximation Pricing and the Variance-Optimal Martingale Measure”, *Annals of Probability* 24, 206–236

A. V. Swishchuk (1995), “Hedging of Options under Mean-Square Criterion and Semi-Markov Volatility”, *Ukrainian Mathematical Journal* 47, 976–983

A. Wiese (1995), “Hedging stochastischer Verpflichtungen”, in: P. Eichelsbacher and M. Löwe (eds.), “Workshop on Discrete Stochastic Models”, University of Bielefeld, SFB 343 *Ergänzungsreihe 95–002*, 29–33