

# A Guided Tour through Quadratic Hedging Approaches

*Martin Schweizer* \*  
*Technische Universität Berlin*  
*Fachbereich Mathematik, MA 7-4*  
*Straße des 17. Juni 136*  
*D – 10623 Berlin*  
*Germany*

**Abstract:** This paper gives an overview of results and developments in the area of pricing and hedging contingent claims in an incomplete market by means of a quadratic criterion. We first present the approach of risk-minimization in the case where the underlying discounted price process  $X$  is a local martingale. We then discuss the extension to local risk-minimization when  $X$  is a semimartingale and explain the relations to the Föllmer-Schweizer decomposition and the minimal martingale measure. Finally we study mean-variance hedging, the variance-optimal martingale measure and the connections to closedness properties of spaces of stochastic integrals.

**Key words:** risk-minimization, locally risk-minimizing, mean-variance hedging, minimal martingale measure, variance-optimal martingale measure, Föllmer-Schweizer decomposition, quadratic hedging criteria, incomplete markets

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

**JEL Classification Numbers:** G10, C60

(in: *E. Jouini, J. Cvitanić, M. Musiela (eds.), “Option Pricing, Interest Rates and Risk Management”, Cambridge University Press (2001), 538–574*)

This version: 15.11.1999

---

\* Research for this paper was partially carried out within Sonderforschungsbereich 373. This paper was printed using funds made available by the Deutsche Forschungsgemeinschaft.

## 0. Introduction

The goal of this paper is to give an overview of some results and developments in the area of pricing and hedging options by means of a quadratic criterion. To put this into a broader perspective, we start in this section with some general ideas and financial motivation before turning to more precise mathematical descriptions. We remark that this borrows extensively from the financial introduction of Delbaen/Monat/Schachermayer/Schweizer/Stricker (1997).

To describe a financial market operating in continuous time, we begin with a probability space  $(\Omega, \mathcal{F}, P)$ , a time horizon  $T \in (0, \infty)$  and a filtration  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ . Intuitively,  $\mathcal{F}_t$  describes the information available at time  $t$ . We have  $d + 1$  basic (primary) assets available for trade with price processes  $S^i = (S_t^i)_{0 \leq t \leq T}$  for  $i = 0, 1, \dots, d$ . To simplify the presentation, we assume that one asset, say  $S^0$ , has a strictly positive price. We then use  $S^0$  as numeraire and immediately pass to quantities discounted with  $S^0$ . This means that asset 0 has (discounted) price 1 at all times and the other assets' (discounted) prices are  $X^i = S^i/S^0$  for  $i = 1, \dots, d$ . Without further mention, all subsequently appearing quantities will be expressed in discounted units.

One central problem of financial mathematics in such a framework is the pricing and hedging of contingent claims by means of dynamic trading strategies based on  $X$ . The best-known example of a contingent claim is a European call option on asset  $i$  with expiration date  $T$  and strike price  $K$ , say. The net payoff at  $T$  to its owner is the random amount  $H(\omega) = \max(X_T^i(\omega) - K, 0) = (X_T^i(\omega) - K)^+$ . More generally, a contingent claim here is simply an  $\mathcal{F}_T$ -measurable random variable  $H$  describing the net payoff at  $T$  of some financial instrument. Hence our claims are of European type in the sense that the date of the payoff is fixed; but the amount to be paid may depend on the whole history of  $X$  up to time  $T$ , or even on more if  $\mathbb{F}$  contains additional information. The problems of pricing and hedging  $H$  can then be formulated as follows: What price should the seller of  $H$  charge the buyer at time 0? And having sold  $H$ , how can he insure or cover himself against the random loss at time  $T$ ?

A natural way to approach these questions is to consider dynamic portfolio strategies of the form  $(\vartheta, \eta) = (\vartheta_t, \eta_t)_{0 \leq t \leq T}$  where  $\vartheta$  is a  $d$ -dimensional predictable process and  $\eta$  is adapted. In such a strategy,  $\vartheta_t^i$  describes the number of units of asset  $i$  held at time  $t$  and  $\eta_t$  is the amount invested in asset 0 at time  $t$ . Predictability of  $\vartheta$  is a mathematical formulation of the informational constraint that  $\vartheta$  is not allowed to anticipate the movement of  $X$ . At any time  $t$ , the value of the portfolio  $(\vartheta_t, \eta_t)$  is given by  $V_t = \vartheta_t^{\text{tr}} X_t + \eta_t$  and the cumulative gains from trade up to time  $t$  are  $G_t(\vartheta) = \int_0^t \vartheta_s dX_s$ . To have the last expression well-defined, we assume that  $X$  is a semimartingale and  $G(\vartheta)$  is then the stochastic integral of  $\vartheta$  with respect to  $X$ . The cumulative costs up to time  $t$  incurred by using  $(\vartheta, \eta)$  are given

by  $C_t = V_t - \int_0^t \vartheta_s dX_s = V_t - G_t(\vartheta)$ . A strategy is called self-financing if its cumulative cost process  $C$  is constant over time or equivalently if its value process  $V$  is given by

$$(0.1) \quad V_t = V_0 + \int_0^t \vartheta_s dX_s = V_0 + G_t(\vartheta)$$

where  $V_0 = C_0$  is the initial outlay required to start the strategy. After time 0, such a strategy is self-supporting: any fluctuations in  $X$  can be neutralized by rebalancing  $\vartheta$  and  $\eta$  in such a way that no further gains or losses result. Note that a self-financing strategy is completely described by  $V_0$  and  $\vartheta$  since the self-financing constraint determines  $V$ , hence also  $\eta$ .

Now fix a contingent claim  $H$  and suppose there exists a self-financing strategy  $(V_0, \vartheta)$  whose terminal value  $V_T$  equals  $H$  with probability one. If our financial market model does not allow arbitrage opportunities, it is clear that the price of  $H$  must be given by  $V_0$  and that  $\vartheta$  furnishes a hedging strategy against  $H$ . This was the basic insight leading to the celebrated Black-Scholes formula for option pricing; see Black/Scholes (1973) and Merton (1973) who solved this problem for the case where  $X$  is a one-dimensional geometric Brownian motion and  $H = (X_T - K)^+$  is a European call option. The mathematical structure of the problem and its connections to martingale theory were subsequently worked out and clarified by J. M. Harrison and D. M. Kreps; a detailed account can be found in Harrison/Pliska (1981). Following their terminology, we call a contingent claim  $H$  attainable if there exists a self-financing strategy with  $V_T = H$   $P$ -a.s. By (0.1), this means that  $H$  can be written as

$$(0.2) \quad H = H_0 + \int_0^T \vartheta_s^H dX_s \quad P\text{-a.s.},$$

i.e., as the sum of a constant  $H_0$  and a stochastic integral with respect to  $X$ . We speak of a complete market if every contingent claim is attainable. Recall that we do not give precise definitions here; for a rigorous mathematical formulation, one has to be rather careful about the integrability conditions imposed on  $H$  and  $\vartheta^H$ .

The importance of the concept of a complete market stems from the fact that it allows the pricing and hedging of contingent claims to be done in a preference-independent fashion. However, completeness is a rather delicate property which is typically destroyed as soon as one considers even minor modifications of a basic complete model. For instance, geometric Brownian motion (the classical Black-Scholes model) becomes incomplete if the volatility is influenced by a second stochastic factor or if one adds a jump component to the model. If one insists on a preference-free approach under incompleteness, one can study the range of possible prices for  $H$  which are consistent with absence of arbitrage in a market containing  $X$ , the riskless asset 1 and  $H$  as traded instruments; this is the idea behind the concept of super-replication. An alternative is to introduce subjective criteria according to which

strategies are chosen and option prices are computed. The goal of this paper is to explain two such criteria in more detail. For a very recent similar survey, see also Pham (1999). A numerical comparison study can be found in Heath/Platen/Schweizer (1999).

For a non-attainable contingent claim, it is by definition impossible to find a strategy with final value  $V_T = H$  which is at the same time self-financing. A first possible approach is to insist on the terminal condition  $V_T = H$ ; since  $\eta$  is allowed to be adapted, this can always be achieved by choice of  $\eta_T$ . But because such strategies cannot be self-financing in general, a “good” strategy should now have a “small” cost process  $C$ . Measuring the riskiness of a strategy by a quadratic criterion was first proposed by Föllmer/Sondermann (1986) for the case where  $X$  is a martingale and subsequently extended to the general semimartingale case in Schweizer (1988, 1991). Under some technical assumptions, such a *locally risk-minimizing* strategy can be characterized by two properties: its cost process  $C$  must be a martingale (so that the strategy is no longer self-financing, but still remains mean-self-financing) and this martingale must be orthogonal to the martingale part  $M$  of the price process  $X$ . Translating this into conditions on the contingent claim  $H$  shows that there exists a locally risk-minimizing strategy for  $H$  if and only if  $H$  admits a decomposition of the form

$$(0.3) \quad H = H_0 + \int_0^T \vartheta_s^H dX_s + L_T^H \quad P\text{-a.s.}$$

where  $L^H$  is a martingale orthogonal to  $M$ . The decomposition (0.3) has been called the *Föllmer-Schweizer decomposition* of  $H$ ; it can be viewed as a generalization to the semimartingale case of the classical Galtchouk-Kunita-Watanabe decomposition from martingale theory. Its financial importance lies in the fact that it directly provides the locally risk-minimizing strategy for  $H$ : the stock component  $\vartheta$  is given by the integrand  $\vartheta^H$  and  $\eta$  is determined by the requirement that the cost process  $C$  should coincide with  $H_0 + L^H$ . Note also that the special case (0.2) of an attainable claim simply corresponds to the absence of the orthogonal term  $L_T^H$ . In particular cases, one can give more explicit constructions for the decomposition (0.3). In the case of finite discrete time,  $\vartheta^H$  and  $L^H$  can be computed recursively backward in time. If  $X$  is continuous, the Föllmer-Schweizer decomposition under  $P$  can be obtained as a Galtchouk-Kunita-Watanabe decomposition, computed under the so-called *minimal martingale measure*  $\hat{P}$ .

One drawback of the preceding approach is the fact that one has to work with strategies which are not self-financing. If one prefers to avoid intermediate costs or an unplanned income, a second idea is to insist on the self-financing constraint (0.1). The possible final outcomes of such strategies are of the form  $V_0 + G_T(\vartheta)$  for some initial capital  $V_0 \in \mathbb{R}$  and some  $\vartheta$  in the set  $\Theta$ , say, of all integrands allowed in (0.1). By definition, a non-attainable claim  $H$  is not of this form and so it seems natural to look for a best approximation of  $H$  by the terminal value  $V_0 + G_T(\vartheta)$  of some pair  $(V_0, \vartheta)$ . The use of a quadratic criterion to measure the quality of this

approximation has been proposed by Bouleau/Lamberton (1989) if  $X$  is both a martingale and a function of a Markov process and by Duffie/Richardson (1991) and Schweizer (1994a), among others, in more general cases. To find such a *mean-variance optimal* strategy, one has to project  $H$  in  $L^2(P)$  on the space  $\mathbb{R} + G_T(\Theta)$  of attainable claims. In particular, this raises the questions whether the space  $G_T(\Theta)$  of stochastic integrals of  $X$  is closed in  $L^2(P)$  and what the structure of the corresponding projection is. Both these problems as well as the computation of the optimal initial capital  $V_0$  turn out to be intimately linked to the so-called *variance-optimal martingale measure*  $\tilde{P}$ .

The paper is structured as follows. Section 1 introduces some general notations and recalls a few preliminaries to complement the preceding discussion. Section 2 explains the above two approaches in the case where  $X$  is a local martingale under  $P$ ; this slightly generalizes the classical results due to Föllmer/Sondermann (1986). Section 3 discusses local risk-minimization in detail and the final section 4 is devoted to mean-variance hedging.

## 1. Notations and preliminaries

In this section, we briefly introduce some notation for later use. This complements the introduction by giving precise definitions. For all standard terminology from martingale theory, we refer to Dellacherie/Meyer (1982).

Mathematically, the basic asset prices are defined on a probability space  $(\Omega, \mathcal{F}, P)$  and described by the constant 1 and an  $\mathbb{R}^d$ -valued stochastic process  $X = (X_t)_{0 \leq t \leq T}$  adapted to a filtration  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$  satisfying the usual conditions of right-continuity and completeness. Adaptedness ensures that time  $t$  prices  $X_t$  are  $\mathcal{F}_t$ -measurable, i.e., observable at time  $t$ . To exclude arbitrage opportunities, we assume that  $X$  admits an *equivalent local martingale measure (ELMM)*  $Q$ , i.e., that there exists a probability measure  $Q \approx P$  such that  $X$  is a local  $Q$ -martingale. With  $\mathbb{IP}$  denoting the convex set of all ELMMs  $Q$  for  $X$ , we thus assume that  $\mathbb{IP} \neq \emptyset$ . *Incompleteness* of the market given by  $X$  and  $\mathbb{F}$  is in our context taken to mean that  $\mathbb{IP}$  contains more than one element (and therefore infinitely many). Finally, a European type *contingent claim* is an  $\mathcal{F}_T$ -measurable random variable  $H$ ; it describes a random payoff to be made at time  $T$ . Before we go on with the general theory, it may be useful to illustrate the preceding concepts by a simple example.

**Example.** Consider one risky asset ( $d = 1$ ) with price process  $X$  and stochastic volatility  $Y$ . More precisely, let  $X$  and  $Y$  satisfy the stochastic differential equations

$$\begin{aligned} \frac{dX_t}{X_t} &= \mu(t, X_t, Y_t) dt + Y_t dW_t^1, \\ dY_t &= a(t, X_t, Y_t) dt + b(t, X_t, Y_t) dW_t^2 \end{aligned}$$

with suitable coefficient functions  $\mu, a, b$  and independent Brownian motions  $W^1, W^2$ . The filtration  $\mathbb{F}$  is the one generated by  $W^1$  and  $W^2$ , made complete and right-continuous. A simple example of a contingent claim here is a European call option on  $X$  with strike  $K$  and maturity  $T$ ; its (net) payoff at time  $T$  is  $H = (X_T - K)^+$ . Note, however, that our abstract framework encompasses much more general (e.g., path-dependent) payoffs and unlike the present example usually assumes no Markovian structure.

In this example, weak assumptions on  $\mu, a, b$  readily guarantee the existence of an ELMM  $Q$ . In fact, it is enough to be able to remove the drift  $\mu$  by a Girsanov transformation. This uniquely determines the transformation's effect on  $W^1$ , but imposes no restrictions on the  $Q$ -drift of  $W^2$ . Hence there is no unique ELMM and we have an incomplete market. This is also intuitively clear because there are two sources of uncertainty  $W^1, W^2$ , but (by assumption) only one risky asset  $X$  for trade. If  $Y$  or some other suitable asset were also tradable, the situation would be different. This ends the present discussion of the example.

Given a contingent claim  $H$ , there are at least two things a potential seller of  $H$  may want to do: *pricing* by assigning a value to  $H$  at times  $t < T$  and *hedging* by covering himself against future losses arising from a sale of  $H$ . The notion of hedging brings up the idea of trading in  $X$  and we formalize this by introducing *trading strategies*. Note first that our assumption  $\mathbb{P} \neq \emptyset$  implies that  $X$  is a semimartingale under  $P$ . It thus makes sense to speak of stochastic integrals with respect to  $X$  and we denote by  $L(X)$  the linear space of all  $\mathbb{R}^d$ -valued predictable  $X$ -integrable processes  $\vartheta$ ; see Dellacherie/Meyer (1982) for additional information. For  $\vartheta \in L(X)$ , the stochastic integral  $\int \vartheta dX$  is well-defined, but some elements of  $L(X)$  are too general to yield economically reasonable strategies. We shall have to impose integrability assumptions later and so we use for the moment the term “pre-strategy”.

**Definition.** A *self-financing pre-strategy* is any pair  $(V_0, \vartheta)$  such that  $\vartheta \in L(X)$  and  $V_0$  is an  $\mathcal{F}_0$ -measurable random variable. Intuitively, one starts out with initial capital  $V_0$  and then holds the dynamically varying number  $\vartheta_t^i$  of shares of asset  $i$  at time  $t$ . The self-financing condition implies that the *value process* of  $(V_0, \vartheta)$  is given by

$$(1.1) \quad V_t(V_0, \vartheta) := V_0 + \int_0^t \vartheta_u dX_u \quad , \quad 0 \leq t \leq T.$$

## 2. The martingale case

We first discuss the two basic quadratic hedging approaches in the simple special case where  $X$  is a *local  $P$ -martingale*; this means that the original measure  $P$  itself is in  $\mathbb{P}$ . We denote by  $[X] = ([X^i, X^j])_{i,j=1,\dots,d}$  the matrix-valued optional covariance process of  $X$  and by  $L^2(X)$

the space of all  $\mathbb{R}^d$ -valued predictable processes  $\vartheta$  such that

$$\|\vartheta\|_{L^2(X)} := \left( E \left[ \int_0^T \vartheta_u^{\text{tr}} d[X]_u \vartheta_u \right] \right)^{\frac{1}{2}} < \infty.$$

Our first result shows that the stochastic integral of  $\vartheta$  with respect to  $X$  is well-defined for  $\vartheta \in L^2(X)$  and has nice properties even if  $X$  is not locally square-integrable. This is because the required integrability is already built into the definition of  $L^2(X)$ . I thank C. Stricker for providing the proof given below.

**Lemma 2.1.** *Suppose that  $X$  is a local  $P$ -martingale. For any  $\vartheta \in L^2(X)$ , the process  $\int \vartheta dX$  is well-defined and in the space  $\mathcal{M}_0^2(P)$  of square-integrable  $P$ -martingales null at 0. Moreover, the space  $\mathcal{I}^2(X) := \{\int \vartheta dX \mid \vartheta \in L^2(X)\}$  of stochastic integrals is a stable subspace of  $\mathcal{M}_0^2(P)$ .*

**Proof.** For  $\vartheta \in L^2(X)$ , the process  $\int \vartheta^{\text{tr}} d[X] \vartheta$  is integrable. Hence  $\int \vartheta dX$  is well-defined and a local  $P$ -martingale by Theorem 4.60 of Jacod (1979), and the Burkholder-Davis-Gundy inequality implies that  $\int \vartheta dX$  is even in  $\mathcal{M}_0^2(P)$ . It is clear that  $\mathcal{I}^2(X)$  is a linear subspace of  $\mathcal{M}_0^2(P)$  and stable under stopping. If  $Y^n = \int \vartheta^n dX$  is a sequence in  $\mathcal{I}^2(X)$  converging to some  $Y$  in  $\mathcal{M}_0^2(P)$ , then  $Y^n$  also converges to  $Y$  in  $\mathcal{M}_0^1(P)$  and so Corollary 2.5.2 of Yor (1978) or Corollary 4.23 of Jacod (1979) (plus Remark III.2 in Stricker (1990) to account for the fact that  $X$  is multidimensional) imply that  $Y = \int \psi dX$  for some  $\psi \in L(X)$ . Since

$$\int_0^T (\vartheta_u^n - \psi_u)^{\text{tr}} d[X]_u (\vartheta_u^n - \psi_u) = [Y^n - Y]_T$$

converges to 0 in  $L^1(P)$  by the convergence of  $Y^n$  to  $Y$  in  $\mathcal{M}_0^2(P)$ , we obtain that  $\psi$  is in  $L^2(X)$ . Hence  $Y \in \mathcal{I}^2(X)$ , so  $\mathcal{I}^2(X)$  is closed in  $\mathcal{M}_0^2(P)$  and this completes the proof.

**q.e.d.**

**Definition.** An *RM-strategy* is any pair  $\varphi = (\vartheta, \eta)$  where  $\vartheta \in L^2(X)$  and  $\eta = (\eta_t)_{0 \leq t \leq T}$  is a real-valued adapted process such that the *value process*  $V(\varphi) := \vartheta^{\text{tr}} X + \eta$  is right-continuous and square-integrable (i.e.,  $V_t(\varphi) \in L^2(P)$  for each  $t \in [0, T]$ ).

Intuitively,  $\vartheta_t^i$  and  $\eta_t$  denote as before the respective numbers of shares of assets  $i$  and 0 held at time  $t$ . (The notation RM anticipates that we shall want to focus on risk-minimization.) But in contrast to section 1, we now also admit strategies that are not self-financing and thus may generate profits or losses over time.

**Definition.** For any RM-strategy  $\varphi$ , the (cumulative) *cost process*  $C(\varphi)$  is defined by

$$C_t(\varphi) := V_t(\varphi) - \int_0^t \vartheta_u dX_u \quad , \quad 0 \leq t \leq T.$$

$C_t(\varphi)$  describes the total costs incurred by  $\varphi$  over the interval  $[0, t]$ ; note that these arise from trading because of the fluctuations of the price process  $X$  and are not due to transaction costs. The *risk process* of  $\varphi$  is defined by

$$R_t(\varphi) := E \left[ (C_T(\varphi) - C_t(\varphi))^2 \mid \mathcal{F}_t \right] \quad , \quad 0 \leq t \leq T.$$

Since a contingent claim  $H$  is  $\mathcal{F}_T$ -measurable and  $\eta$  is allowed to be adapted, we can always find RM-strategies with  $V_T(\varphi) = H$  provided that  $H \in L^2(P)$ . The simplest is “wait, then pay” where  $\vartheta \equiv 0$  and  $\eta_t = HI_{\{t=T\}}$ . But in general, these strategies will not be self-financing; in fact, (1.1) tells us that there is a self-financing RM-strategy  $\varphi$  with  $V_T(\varphi) = H$  if and only if  $H$  admits a representation as the sum of an  $\mathcal{F}_0$ -measurable random variable and a stochastic integral with respect to  $X$ . In that case, the cost process  $C(\varphi)$  is constant and the risk process  $R(\varphi)$  is identically 0. For claims where this is not possible, the idea of Föllmer/Sondermann (1986) in defining *risk-minimization* is to look among all RM-strategies with  $V_T(\varphi) = H$  for one which minimizes the risk process in a suitable sense.

**Definition.** An RM-strategy  $\varphi$  is called *risk-minimizing* if for any RM-strategy  $\tilde{\varphi}$  such that  $V_T(\tilde{\varphi}) = V_T(\varphi)$   $P$ -a.s., we have

$$R_t(\varphi) \leq R_t(\tilde{\varphi}) \quad P\text{-a.s. for every } t \in [0, T].$$

This is not the original definition, but it amounts to the same thing:

**Lemma 2.2.** *An RM-strategy  $\varphi$  is risk-minimizing if and only if*

$$R_t(\varphi) \leq R_t(\tilde{\varphi}) \quad P\text{-a.s.}$$

for every  $t \in [0, T]$  and for every RM-strategy  $\tilde{\varphi}$  which is an *admissible continuation* of  $\varphi$  from  $t$  on in the sense that  $V_T(\tilde{\varphi}) = V_T(\varphi)$   $P$ -a.s.,  $\tilde{\vartheta}_s = \vartheta_s$  for  $s \leq t$  and  $\tilde{\eta}_s = \eta_s$  for  $s < t$ .

**Proof.** See Lemma 2.1 of Schweizer (1994b); this does not use that  $X$  is a local  $P$ -martingale.

**Remark.** The definition in Föllmer/Sondermann (1986) of an *admissible continuation* of  $\varphi$  from  $t$  on is more symmetric because they stipulate that  $\tilde{\vartheta}_s = \vartheta_s$  and  $\tilde{\eta}_s = \eta_s$  hold both for

$s < t$ . In the martingale case and for continuous time, this difference does not matter, but a discrete-time setting or the subsequent generalization to local risk-minimization do need the asymmetric formulation in Lemma 2.2. This also reflects the asymmetry between the requirements on  $\vartheta$  and  $\eta$  since  $\vartheta$  must be predictable while  $\eta$  is allowed to be adapted.

Although RM-strategies with  $V_T(\varphi) = H$  will in general not be self-financing, it turns out that good RM-strategies are still “self-financing on average” in the following sense.

**Definition.** An RM-strategy  $\varphi$  is called *mean-self-financing* if its cost process  $C(\varphi)$  is a  $P$ -martingale.

**Lemma 2.3.** Any risk-minimizing RM-strategy  $\varphi$  is also mean-self-financing.

**Proof.** This proof does not use that  $X$  is a local  $P$ -martingale. Fix  $t_0 \in [0, T]$  and define  $\tilde{\varphi}$  by setting  $\tilde{\vartheta} := \vartheta$  and

$$\tilde{\vartheta}_t^{\text{tr}} X_t + \tilde{\eta}_t = V_t(\tilde{\varphi}) := V_t(\varphi)I_{[0, t_0)}(t) + E \left[ V_T(\varphi) - \int_t^T \vartheta_u dX_u \middle| \mathcal{F}_t \right] I_{[t_0, T]}(t),$$

choosing an RCLL version. Then  $\tilde{\varphi}$  is an RM-strategy with  $V_T(\tilde{\varphi}) = V_T(\varphi)$  and because  $C_T(\varphi) = C_T(\tilde{\varphi})$  and  $C_{t_0}(\tilde{\varphi}) = E[C_T(\tilde{\varphi})|\mathcal{F}_{t_0}]$ ,

$$C_T(\varphi) - C_{t_0}(\varphi) = C_T(\tilde{\varphi}) - C_{t_0}(\tilde{\varphi}) + E[C_T(\tilde{\varphi})|\mathcal{F}_{t_0}] - C_{t_0}(\varphi)$$

implies that

$$R_{t_0}(\varphi) = R_{t_0}(\tilde{\varphi}) + (C_{t_0}(\varphi) - E[C_T(\varphi)|\mathcal{F}_{t_0}])^2.$$

Because  $\varphi$  is risk-minimizing, we conclude that

$$C_{t_0}(\varphi) = E[C_T(\varphi)|\mathcal{F}_{t_0}] \quad P\text{-a.s.}$$

and since  $t_0$  is arbitrary, the assertion follows.

**q.e.d.**

The key result for finding risk-minimizing RM-strategies is the well-known *Galtchouk-Kunita-Watanabe decomposition*. Because  $\mathcal{I}^2(X)$  is a stable subspace of  $\mathcal{M}_0^2(P)$ , any  $H \in L^2(\mathcal{F}_T, P)$  can be uniquely written as

$$(2.1) \quad H = E[H|\mathcal{F}_0] + \int_0^T \vartheta_u^H dX_u + L_T^H \quad P\text{-a.s.}$$

for some  $\vartheta^H \in L^2(X)$  and some  $L^H \in \mathcal{M}_0^2(P)$  which is strongly orthogonal to  $\mathcal{I}^2(X)$ ; this means that  $L^H \int \vartheta dX$  is a  $P$ -martingale for every  $\vartheta \in L^2(X)$ . The next result was obtained by Föllmer/Sondermann (1986) for  $d = 1$  under the assumption that  $X$  is in  $\mathcal{M}^2(P)$ . The observation and proof that it holds for a general local  $P$ -martingale  $X$  seem to be new.

**Theorem 2.4.** *Suppose that  $X$  is a local  $P$ -martingale. Then every contingent claim  $H \in L^2(\mathcal{F}_T, P)$  admits a unique risk-minimizing RM-strategy  $\varphi^*$  with  $V_T(\varphi^*) = H$   $P$ -a.s. In terms of the decomposition (2.1),  $\varphi^*$  is explicitly given by*

$$\begin{aligned}\vartheta^* &= \vartheta^H, \\ V_t(\varphi^*) &= E[H|\mathcal{F}_t] =: V_t^* \quad , \quad 0 \leq t \leq T, \\ C(\varphi^*) &= E[H|\mathcal{F}_0] + L^H.\end{aligned}$$

**Proof.** Note first that the above prescription defines an RM-strategy  $\varphi^*$  with  $V_T(\varphi^*) = H$ . Now fix  $t \in [0, T]$  and any RM-strategy  $\tilde{\varphi}$  with  $V_T(\tilde{\varphi}) = H$ . The same argument as in the proof of Lemma 2.3 shows that we may assume  $C_t(\tilde{\varphi}) = E[C_T(\tilde{\varphi})|\mathcal{F}_t]$  and so we get

$$C_T(\tilde{\varphi}) - C_t(\tilde{\varphi}) = H - \int_t^T \tilde{\vartheta}_u dX_u - E[H|\mathcal{F}_t] = L_T^H - L_t^H + \int_t^T (\vartheta_u^H - \tilde{\vartheta}_u) dX_u$$

by using (2.1) and the martingale property of  $\int \tilde{\vartheta} dX$ . Because  $C(\varphi^*) = C_0(\varphi^*) + L^H$ , the orthogonality of  $L^H$  and  $\mathcal{I}^2(X)$  yields

$$R_t(\tilde{\varphi}) = R_t(\varphi^*) + E \left[ \left( \int_t^T (\vartheta_u^H - \tilde{\vartheta}_u) dX_u \right)^2 \middle| \mathcal{F}_t \right] \geq R_t(\varphi^*).$$

Hence  $\varphi^*$  is risk-minimizing. If some other  $\tilde{\varphi}$  is also risk-minimizing, then  $C(\tilde{\varphi})$  must be a martingale by Lemma 2.3 and then the same argument as before gives for  $t = 0$

$$R_0(\tilde{\varphi}) = R_0(\varphi^*) + E \left[ \int_0^T (\vartheta_u^H - \tilde{\vartheta}_u)^{\text{tr}} d[X]_u (\vartheta_u^H - \tilde{\vartheta}_u) \middle| \mathcal{F}_0 \right].$$

Because  $\tilde{\varphi}$  is risk-minimizing, this implies  $\tilde{\vartheta} = \vartheta^H = \vartheta^*$  and since  $C(\tilde{\varphi})$  is a martingale and  $V_T(\tilde{\varphi}) = V_T(\varphi^*)$ , we also obtain  $\tilde{\varphi} = \varphi^*$ .

**q.e.d.**

**Remark.** The preceding approach relies heavily on the fact that the contingent claim  $H$  only makes one payment at the terminal date  $T$ . For applications to insurance derivatives

as in Møller (1998a), this is not sufficient because such products involve possible payments at any time  $t \in [0, T]$ . An extension of the risk-minimization concept to the case of such payment streams has been developed in Møller (1998b).

An alternative quadratic approach in the martingale case has been studied by Bouleau/Lamberton (1989). They imposed the additional condition that  $X$  is a function of some Markov process to get more explicit results, but their basic idea can also be explained in our general framework. Suppose that instead of insisting on  $V_T(\varphi) = H$   $P$ -a.s., we focus on *self-financing* RM-strategies. Such a strategy is described by a pair  $(V_0, \vartheta)$  in  $L^2(\mathcal{F}_0, P) \times L^2(X)$  and its shortfall at the terminal date  $T$  is

$$H - V_T(V_0, \vartheta) = H - V_0 - \int_0^T \vartheta_u dX_u.$$

If  $H$  is attainable by such a strategy in the sense that  $H = V_T(V_0, \vartheta)$  for some pair  $(V_0, \vartheta)$ , the shortfall can be reduced to 0. But in general, one has a residual risk of

$$J_0(V_0, \vartheta) := E \left[ (H - V_T(V_0, \vartheta))^2 \right]$$

if one uses a quadratic loss function and the idea of Bouleau/Lamberton (1989) is to minimize this residual risk by choice of  $(V_0, \vartheta)$ . This clearly amounts to projecting the random variable  $H$  in  $L^2(P)$  on the linear space spanned by  $L^2(\mathcal{F}_0, P)$  and the stochastic integrals  $\int_0^T \vartheta_u dX_u$  with  $\vartheta \in L^2(X)$  and thanks to (2.1), the solution is given by

$$\begin{aligned} \bar{V}_0 &= E[H|\mathcal{F}_0], \\ \bar{\vartheta} &= \vartheta^H \end{aligned}$$

with a minimal residual risk of

$$J_0(\bar{V}_0, \bar{\vartheta}) = E \left[ (L_T^H)^2 \right] = \text{Var}[L_T^H].$$

In the next two sections, we generalize the preceding two approaches to the case where  $X$  is under  $P$  no longer a local martingale, but only a semimartingale. Risk-minimization will be replaced by local risk-minimization and extending the above projection approach leads to mean-variance hedging. We shall also see that extensions of the Galtchouk-Kunita-Watanabe decomposition play an important role and that it is often very helpful to work with a suitably chosen ELMM.

### 3. Local risk-minimization

Let us now consider the general situation where the original measure  $P$  is not in  $\mathcal{IP}$ . Hence  $X$  is no longer a local  $P$ -martingale, but only a semimartingale under  $P$ . Given a contingent claim  $H$ , we could still look for risk-minimizing strategies  $\varphi$  with  $V_T(\varphi) = H$ . But there is bad news:

**Proposition 3.1.** *If  $X$  is not a local  $P$ -martingale, a contingent claim  $H$  admits in general no risk-minimizing strategy  $\varphi$  with  $V_T(\varphi) = H$   $P$ -a.s.*

**Proof.** We show this by presenting an explicit counterexample given in Schweizer (1988). For simplicity, we work in discrete time. Let  $X = (X_k)_{k=0,1,\dots,T}$  (with  $T \in \mathbb{N}$ ) be a real-valued square-integrable process adapted to a filtration  $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$  and fix  $H \in L^2(\mathcal{F}_T, P)$ . The example below is on a finite probability space so that all integrability requirements are satisfied.

If  $\varphi^*$  is a risk-minimizing strategy with  $V_T(\varphi^*) = H$   $P$ -a.s., Lemma 2.3 implies that  $C(\varphi^*)$  is a  $P$ -martingale so that we get

$$R_k(\varphi^*) = \text{Var}[C_T(\varphi^*)|\mathcal{F}_k] = \text{Var} \left[ H - \sum_{j=k+1}^T \vartheta_j^* \Delta X_j \middle| \mathcal{F}_k \right]$$

by using  $V_T(\varphi^*) = H$  and omitting  $\mathcal{F}_k$ -measurable terms from the conditional variance. By  $\Delta X_j := X_j - X_{j-1}$ , we denote the increment of  $X$  from  $j-1$  to  $j$ . Moreover,

$$\vartheta_k^* X_k + \eta_k^* = V_k(\varphi^*) = C_k(\varphi^*) + \sum_{j=1}^k \vartheta_j^* \Delta X_j = E \left[ H - \sum_{j=1}^T \vartheta_j^* \Delta X_j \middle| \mathcal{F}_k \right] + \sum_{j=1}^k \vartheta_j^* \Delta X_j$$

shows that  $\varphi^*$  is uniquely determined by the predictable process  $\vartheta^*$  and vice versa. Because  $\varphi^*$  is risk-minimizing, any mean-self-financing strategy  $\varphi$  with  $V_T(\varphi) = H$  will satisfy

$$\text{Var} \left[ H - \sum_{j=k+1}^T \vartheta_j \Delta X_j \middle| \mathcal{F}_k \right] = R_k(\varphi) \geq R_k(\varphi^*) = \text{Var} \left[ H - \sum_{j=k+1}^T \vartheta_j^* \Delta X_j \middle| \mathcal{F}_k \right].$$

In particular, this implies that the mapping

$$\vartheta_{k+1} \mapsto \text{Var} \left[ H - \vartheta_{k+1} \Delta X_{k+1} - \sum_{j=k+2}^T \vartheta_j^* \Delta X_j \middle| \mathcal{F}_k \right]$$

attains its minimum at  $\vartheta_{k+1}^*$  and so the first order condition for this problem yields

$$(3.1) \quad \vartheta_{k+1}^* = \frac{\text{Cov} \left( H - \sum_{j=k+2}^T \vartheta_j^* \Delta X_j, \Delta X_{k+1} \middle| \mathcal{F}_k \right)}{\text{Var}[\Delta X_{k+1}|\mathcal{F}_k]}.$$

This backward recursive expression determines a unique candidate for a risk-minimizing strategy  $\varphi^*$ .

For the counterexample, we take  $T = 2$  and consider a random walk  $X$  starting at 0 whose (i.i.d.) increments take the values  $+1, 0, -1$  with respective probabilities  $\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$  under  $P$ . The filtration  $\mathbb{F}$  is generated by  $X$  and the contingent claim is  $H = |X_2|^2$ . Any predictable process  $\vartheta$  is determined by the value of  $\vartheta_1$  and the three possible values of  $\vartheta_2$  on the sets  $\{X_1 = +1\}, \{X_1 = 0\}, \{X_1 = -1\}$  generating  $\mathcal{F}_1$  and we denote the latter by  $\vartheta_2(+1), \vartheta_2(0), \vartheta_2(-1)$  respectively. If there is a risk-minimizing strategy  $\varphi^*$  with  $V_T(\varphi^*) = H$ , then  $\vartheta^*$  must be given by (3.1) and an explicit calculation yields the values  $\vartheta_1^* = -\frac{1}{11}, \vartheta_2^*(+1) = \frac{21}{11}, \vartheta_2^*(0) = -\frac{1}{11}, \vartheta_2^*(-1) = -\frac{23}{11}$  which lead to an initial risk of

$$R_0(\varphi^*) = \frac{24}{66}.$$

But for any mean-self-financing strategy  $\varphi$  with  $V_T(\varphi) = H$ , the initial risk  $R_0(\varphi)$  can also be viewed as a function of the four variables  $\vartheta_1, \vartheta_2(+1), \vartheta_2(0), \vartheta_2(-1)$ . The minimum of this function is found to be attained at  $\bar{\vartheta}_1 = -\frac{1}{11}, \bar{\vartheta}_2(+1) = \frac{59}{33}, \bar{\vartheta}_2(0) = \frac{5}{33}, \bar{\vartheta}_2(-1) = -\frac{71}{33}$  and calculated as

$$R_0(\bar{\varphi}) = \frac{23}{66} < R_0(\varphi^*).$$

This shows that the unique candidate  $\varphi^*$  given by (3.1) is not risk-minimizing and hence there cannot exist any risk-minimizing strategy ending at  $H$ . This completes the proof.

**q.e.d.**

**Remark.** Intuitively, the reason for the failure of the risk-minimization approach in the non-martingale case is a compatibility problem. At any time  $t$ , we minimize  $R_t(\varphi)$  over all admissible continuations from  $t$  on and obtain a continuation which is optimal when viewed in  $t$  only. But for  $s < t$ , the  $s$ -optimal continuation from  $s$  on tells us what to do on the entire interval  $(s, T] \supset (t, T]$  and this may be different from what the  $t$ -optimal continuation from  $t$  on prescribes. The above counterexample shows that this creates indeed a problem in general, and the remarkable result in Theorem 2.4 is that the martingale property of  $X$  guarantees the required compatibility.

Before we turn to the somewhat technical concept of local risk-minimization in continuous time, it may be useful to explain the basic ideas and results in a *discrete-time framework*; an elementary introduction can also be found in Föllmer/Schweizer (1989). We consider for this a situation where trading is only done at dates  $k = 0, 1, \dots, T \in \mathbb{N}$ . At time  $k$ , we choose the numbers  $\vartheta_{k+1}$  of shares to be held over the time period  $(k, k + 1]$  and the number  $\eta_k$  of units of asset 0 to be held over  $[k, k + 1)$ . Note that predictability of  $\vartheta$  forces us to determine the date  $k + 1$  holdings  $\vartheta_{k+1}$  already at date  $k$ . The actual time  $k$  portfolio is  $\varphi_k = (\vartheta_k, \eta_k)$

and its value is  $V_k(\varphi) = \vartheta_k^{\text{tr}} X_k + \eta_k$ . Since we want to minimize risk locally, we now consider the incremental cost incurred by adjusting the portfolio from  $\varphi_k$  to  $\varphi_{k+1}$ . Because  $\vartheta_{k+1}$  is chosen already at time  $k$  with prices given by  $X_k$ , this cost increment is

$$\begin{aligned} C_{k+1}(\varphi) - C_k(\varphi) &= (\vartheta_{k+1} - \vartheta_k)^{\text{tr}} X_k + \eta_{k+1} - \eta_k \\ &= V_{k+1}(\varphi) - V_k(\varphi) - \vartheta_{k+1}^{\text{tr}} (X_{k+1} - X_k) \\ &= \Delta V_{k+1}(\varphi) - \vartheta_{k+1}^{\text{tr}} \Delta X_{k+1} \end{aligned}$$

with the difference operator  $\Delta U_{k+1} := U_{k+1} - U_k$  for any discrete-time stochastic process  $U$ .

For local risk-minimization, our goal is to minimize  $E \left[ (C_{k+1}(\varphi) - C_k(\varphi))^2 \mid \mathcal{F}_k \right]$  with respect to the time  $k$  control variables  $\vartheta_{k+1}$  and  $\eta_k$ . To be accurate, this requires integrability conditions on  $\vartheta$  and  $\eta$ , but we leave these aside for the moment. By using the expression for  $\Delta C_{k+1}(\varphi)$  and the fact that the  $\mathcal{F}_k$ -measurable term  $V_k(\varphi)$  does not influence the conditional variance given  $\mathcal{F}_k$ , we can write

$$\begin{aligned} E \left[ (\Delta C_{k+1}(\varphi))^2 \mid \mathcal{F}_k \right] &= \text{Var} \left[ V_{k+1}(\varphi) - \vartheta_{k+1}^{\text{tr}} \Delta X_{k+1} \mid \mathcal{F}_k \right] \\ &\quad + \left( E \left[ V_{k+1}(\varphi) - \vartheta_{k+1}^{\text{tr}} \Delta X_{k+1} \mid \mathcal{F}_k \right] - V_k(\varphi) \right)^2. \end{aligned}$$

Because the first term on the right-hand side does not depend on  $\eta_k$ , it is clearly optimal to choose  $\eta_k$  in such a way that

$$(3.2) \quad V_k(\varphi) = E \left[ V_{k+1}(\varphi) - \vartheta_{k+1}^{\text{tr}} \Delta X_{k+1} \mid \mathcal{F}_k \right].$$

This is equivalent to

$$0 = E \left[ \Delta V_{k+1}(\varphi) - \vartheta_{k+1}^{\text{tr}} \Delta X_{k+1} \mid \mathcal{F}_k \right] = E \left[ \Delta C_{k+1}(\varphi) \mid \mathcal{F}_k \right]$$

so that an optimal strategy should again be mean-self-financing. Because  $V_T(\varphi) = H$  is fixed, (3.2) implies by a backward induction argument that for the purposes of minimizing  $E \left[ (\Delta C_{k+1}(\varphi))^2 \mid \mathcal{F}_k \right]$  at time  $k$ , the value  $V_{k+1}(\varphi)$  may be considered as given. Thus it only remains to minimize  $\text{Var} \left[ V_{k+1}(\varphi) - \vartheta_{k+1}^{\text{tr}} \Delta X_{k+1} \mid \mathcal{F}_k \right]$  with respect to the  $\mathcal{F}_k$ -measurable quantity  $\vartheta_{k+1}$  and this will be achieved if and only if

$$(3.3) \quad \text{Cov} \left( V_{k+1}(\varphi) - \vartheta_{k+1}^{\text{tr}} \Delta X_{k+1}, \Delta X_{k+1} \mid \mathcal{F}_k \right) = 0.$$

To simplify this, we use the Doob decomposition of  $X$  into a martingale  $\bar{M}$  and a predictable process  $\bar{A}$  given by  $\bar{M}_0 := 0 =: \bar{A}_0$ ,  $\Delta \bar{A}_{k+1} := E[\Delta X_{k+1} \mid \mathcal{F}_k]$  and  $\Delta \bar{M}_{k+1} := \Delta X_{k+1} - \Delta \bar{A}_{k+1}$ . Then (3.3) can be rewritten as

$$0 = \text{Cov} \left( \Delta C_{k+1}(\varphi), \Delta \bar{M}_{k+1} \mid \mathcal{F}_k \right) = E \left[ \Delta C_{k+1}(\varphi) \Delta \bar{M}_{k+1} \mid \mathcal{F}_k \right]$$

which says that the product of the two martingales  $C(\varphi)$  and  $\bar{M}$  must be a martingale or (equivalently) that  $C(\varphi)$  and  $\bar{M}$  must be strongly orthogonal under  $P$ . Thus in discrete time

(3.4) a suitably integrable strategy  $\varphi$  is locally risk-minimizing if and only if its cost process  $C(\varphi)$  is a martingale and strongly orthogonal to the martingale part (here  $\bar{M}$ ) of  $X$ .

Before passing to the continuous-time case, let us point out another useful property which will have an analogue later on. Suppose for simplicity that  $d = 1$ . Because  $\vartheta_{k+1}$  is  $\mathcal{F}_k$ -measurable, we can solve (3.3) for  $\vartheta_{k+1}$  to obtain

$$\vartheta_{k+1} = \frac{\text{Cov}(V_{k+1}(\varphi), \Delta X_{k+1} | \mathcal{F}_k)}{\text{Var}[\Delta X_{k+1} | \mathcal{F}_k]} = \frac{E[V_{k+1}(\varphi) \Delta \bar{M}_{k+1} | \mathcal{F}_k]}{E[(\Delta \bar{M}_{k+1})^2 | \mathcal{F}_k]}.$$

Using  $E[\vartheta_{k+1} \Delta X_{k+1} | \mathcal{F}_k] = \vartheta_{k+1} \Delta \bar{A}_{k+1}$  and plugging into (3.2) yields

$$\begin{aligned} V_k(\varphi) &= E[V_{k+1}(\varphi) - \vartheta_{k+1} \Delta \bar{A}_{k+1} | \mathcal{F}_k] \\ &= E\left[V_{k+1}(\varphi) \left(1 - \frac{\Delta \bar{A}_{k+1}}{E[(\Delta \bar{M}_{k+1})^2 | \mathcal{F}_k]} \Delta \bar{M}_{k+1}\right) \middle| \mathcal{F}_k\right] \\ &= E\left[V_{k+1}(\varphi) \frac{\bar{Z}_{k+1}}{\bar{Z}_k} \middle| \mathcal{F}_k\right] \end{aligned}$$

so that

(3.5) for a locally risk-minimizing strategy  $\varphi$ , the product  $\bar{Z}V(\varphi)$  is a  $P$ -martingale

if the process  $\bar{Z}$  is defined by the difference equation

$$(3.6) \quad \bar{Z}_{k+1} - \bar{Z}_k = \bar{Z}_k \left( \frac{\bar{Z}_{k+1}}{\bar{Z}_k} - 1 \right) = -\bar{Z}_k \bar{\lambda}_{k+1} \Delta \bar{M}_{k+1} \quad , \quad \bar{Z}_0 = 1$$

with the predictable process

$$\bar{\lambda}_{k+1} := \frac{\Delta \bar{A}_{k+1}}{E[(\Delta \bar{M}_{k+1})^2 | \mathcal{F}_k]} = \frac{E[\Delta X_{k+1} | \mathcal{F}_k]}{\text{Var}[\Delta X_{k+1} | \mathcal{F}_k]} \quad , \quad k = 0, 1, \dots, T-1.$$

This property will come up again later in a continuous-time version.

**Remark.** The above definition of local risk-minimization in discrete time is different from the original one. The idea there is to consider at time  $k$  instead of  $E\left[(C_{k+1}(\varphi) - C_k(\varphi))^2 \middle| \mathcal{F}_k\right]$  the risk  $R_k(\varphi) = E\left[(C_T(\varphi) - C_k(\varphi))^2 \middle| \mathcal{F}_k\right]$ . But just like before and in contrast to risk-minimization, this is viewed as a function of the time  $k$  control variables  $\eta_k$  and  $\vartheta_{k+1}$  only and

minimized only locally, i.e., with respect to these local variables. A more formal definition can be found in Schweizer (1988) or Lamberton/Pham/Schweizer (1998) who also prove the equivalence between the two definitions; see the remark on p.25 of Schweizer (1988) or Proposition 2 of Lamberton/Pham/Schweizer (1998). The reason for using  $R_k(\varphi)$  is that this formulation can be generalized to continuous time.

Let us now turn to the case of continuous time. Because we want to work again with local variances, we require more specific assumptions on the price process  $X$  and we start by making these precise. Since  $\mathbb{P} \neq \emptyset$ , we know already that  $X$  is a semimartingale under  $P$ . We now assume that  $X$  is in  $\mathcal{S}_{\text{loc}}^2(P)$  so that it can be decomposed as  $X = X_0 + M + A$  where  $M \in \mathcal{M}_{0,\text{loc}}^2(P)$  is an  $\mathbb{R}^d$ -valued locally square-integrable local  $P$ -martingale null at 0 and  $A$  is an  $\mathbb{R}^d$ -valued predictable process of finite variation also null at 0. We denote by  $\langle M \rangle = (\langle M \rangle^{ij})_{i,j=1,\dots,d} = (\langle M^i, M^j \rangle)_{i,j=1,\dots,d}$  the matrix-valued predictable covariance process of  $M$  and we suppose that  $A$  is absolutely continuous with respect to  $\langle M \rangle$  in the sense that

$$A_t^i = \left( \int_0^t d\langle M \rangle_s \widehat{\lambda}_s \right)^i := \sum_{j=1}^d \int_0^t \widehat{\lambda}_s^j d\langle M^i, M^j \rangle_s \quad , \quad 0 \leq t \leq T, i = 1, \dots, d$$

for some  $\mathbb{R}^d$ -valued predictable process  $\widehat{\lambda}$  such that the *mean-variance tradeoff process*

$$\widehat{K}_t := \int_0^t \widehat{\lambda}_s^{\text{tr}} d\langle M \rangle_s \widehat{\lambda}_s = \sum_{i,j=1}^d \int_0^t \widehat{\lambda}_s^i \widehat{\lambda}_s^j d\langle M^i, M^j \rangle_s$$

is finite  $P$ -a.s. for each  $t \in [0, T]$ . This complex of conditions on  $X$  is sometimes called the *structure condition (SC)*. Since  $\mathbb{P} \neq \emptyset$ , it is for instance automatically satisfied if  $X$  is continuous; see Theorem 1 of Schweizer (1995a) for this and Choulli/Stricker (1996) for more general results in this direction. Additional results on the relation between (SC) and properties of absence of arbitrage for the process  $X$  can be found in Delbaen/Schachermayer (1995). Note that the stochastic integral  $\int \widehat{\lambda} dM$  is well-defined under (SC) and that its variance process is  $\langle \int \widehat{\lambda} dM \rangle = \widehat{K}$ ; this will be used later on.

**Definition.**  $\Theta_S$  denotes the space of all processes  $\vartheta \in L(X)$  for which the stochastic integral  $\int \vartheta dX$  is in the space  $\mathcal{S}^2(P)$  of semimartingales. Equivalently,  $\vartheta$  must be predictable with

$$E \left[ \int_0^T \vartheta_s^{\text{tr}} d[M]_s \vartheta_s + \left( \int_0^T |\vartheta_s^{\text{tr}} dA_s| \right)^2 \right] < \infty.$$

(This equivalence does not use (SC); it only requires  $X$  to be a special semimartingale.)

**Definition.** An  $L^2$ -strategy is a pair  $\varphi = (\vartheta, \eta)$  where  $\vartheta \in \Theta_S$  and  $\eta = (\eta_t)_{0 \leq t \leq T}$  is a real-valued adapted process such that the *value process*  $V(\varphi) := \vartheta^{\text{tr}} X + \eta$  is right-continuous and square-integrable (i.e.,  $V_t(\varphi) \in L^2(P)$  for each  $t \in [0, T]$ ). The cost process  $C(\varphi)$ , the risk process  $R(\varphi)$  and the concept of mean-self-financing are defined as in section 2. Note that in the martingale case  $A \equiv 0$ , we have  $\Theta_S = L^2(X)$  so that the notions of RM-strategy and  $L^2$ -strategy then coincide.

For a formal description of local risk-minimization in continuous time, we now restrict our attention to the case  $d = 1$ . One can proceed in a similar way and obtain analogous results for  $d > 1$ ; the details for this have been worked out and will be presented elsewhere. The only reason for choosing  $d = 1$  here is that this permits references to already published work. Let us first fix some terminology. A *partition* of  $[0, T]$  is a finite set  $\tau = \{t_0, t_1, \dots, t_k\}$  of times with  $0 = t_0 < t_1 < \dots < t_k = T$  and the mesh size of  $\tau$  is  $|\tau| := \max_{t_i, t_{i+1} \in \tau} (t_{i+1} - t_i)$ . The number  $k$  of times is not fixed, but can depend on  $\tau$ . A sequence  $(\tau_n)_{n \in \mathbb{N}}$  of partitions is called *increasing* if  $\tau_n \subseteq \tau_{n+1}$  for all  $n$ ; it *tends to the identity* if  $\lim_{n \rightarrow \infty} |\tau_n| = 0$ .

The next definition translates the idea that changing an optimal strategy over a small time interval should lead to an increase of risk, at least asymptotically. The form of the denominator indicates that the appropriate time scale for these asymptotics is determined by the fluctuations of  $X$  as measured by its predictable quadratic variation.

**Definition.** A *small perturbation* is an  $L^2$ -strategy  $\Delta = (\delta, \varepsilon)$  such that  $\delta$  is bounded, the variation of  $\int \delta dA$  is bounded (uniformly in  $t$  and  $\omega$ ) and  $\delta_T = \varepsilon_T = 0$ . For any subinterval  $(s, t]$  of  $[0, T]$ , we then define the small perturbation

$$\Delta|_{(s,t]} := (\delta I_{(s,t]}, \varepsilon I_{[s,t]}).$$

The asymmetry between  $\delta$  and  $\varepsilon$  reflects the fact that  $\delta$  is predictable and  $\varepsilon$  merely adapted.

**Definition.** For an  $L^2$ -strategy  $\varphi$ , a small perturbation  $\Delta$  and a partition  $\tau$  of  $[0, T]$ , we set

$$r^\tau(\varphi, \Delta) := \sum_{t_i, t_{i+1} \in \tau} \frac{R_{t_i}(\varphi + \Delta|_{(t_i, t_{i+1}]}) - R_{t_i}(\varphi)}{E[\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i} | \mathcal{F}_{t_i}]} I_{(t_i, t_{i+1}]}$$

$\varphi$  is called *locally risk-minimizing* if

$$\liminf_{n \rightarrow \infty} r^{\tau_n}(\varphi, \Delta) \geq 0 \quad (P \otimes \langle M \rangle)\text{-a.e. on } \Omega \times [0, T]$$

for every small perturbation  $\Delta$  and every increasing sequence  $(\tau_n)_{n \in \mathbb{N}}$  of partitions tending to the identity.

**Lemma 3.2.** *Let  $d = 1$  and suppose that  $\langle M \rangle$  is  $P$ -a.s. strictly increasing. If an  $L^2$ -strategy is locally risk-minimizing, it is also mean-self-financing.*

**Proof.** This is Lemma 2.1 of Schweizer (1991); note that its assumption (X1) of square-integrability for  $M$  is not required in the proof.

Thanks to Lemma 3.2, we can in searching for locally risk-minimizing strategies restrict ourselves to the class of mean-self-financing strategies. Together with the terminal condition  $V_T(\varphi) = H$ , this class can be parametrized by processes  $\vartheta \in \Theta_S$  so that we effectively have to deal with one dimension less than before. To proceed, we then split  $r^\tau(\varphi, \Delta)$  into a term depending only on  $\vartheta$  and  $\delta$  and a second term involving  $\eta$  and  $\varepsilon$  as well. The subsequent assumptions ensure that the second term vanishes asymptotically, and the first one is dealt with by means of differentiation results for semimartingales presented in Schweizer (1990). In the end, we then obtain the following result; note that it exactly parallels (3.4).

**Theorem 3.3.** *Suppose that  $X$  satisfies (SC),  $d = 1$ ,  $M$  is in  $\mathcal{M}_0^2(P)$ ,  $\langle M \rangle$  is  $P$ -a.s. strictly increasing,  $A$  is  $P$ -a.s. continuous and  $E[\widehat{K}_T] < \infty$ . Let  $H \in L^2(\mathcal{F}_T, P)$  be a contingent claim and  $\varphi$  an  $L^2$ -strategy with  $V_T(\varphi) = H$   $P$ -a.s. Then  $\varphi$  is locally risk-minimizing if and only if  $\varphi$  is mean-self-financing and the martingale  $C(\varphi)$  is strongly orthogonal to  $M$ .*

**Proof.** This follows immediately from Proposition 2.3 of Schweizer (1991) once we note that

$$E[\widehat{K}_T] = E \left[ \int_0^T |\widehat{\lambda}_u|^2 d\langle M \rangle_u \right] < \infty$$

implies that  $\widehat{\lambda} \in L^2(P \otimes \langle M \rangle)$  so that  $|\widehat{\lambda}| \log^+ |\widehat{\lambda}|$  is  $(P \otimes \langle M \rangle)$ -integrable. Assumption (X5) of Schweizer (1991) ( $X$  continuous at  $T$   $P$ -a.s.) is not used in the proof.

**q.e.d.**

Now we return to the general case  $d \geq 1$ . The preceding result motivates the following

**Definition.** Let  $H \in L^2(\mathcal{F}_T, P)$  be a contingent claim. An  $L^2$ -strategy  $\varphi$  with  $V_T(\varphi) = H$   $P$ -a.s. is called *pseudo-locally risk-minimizing* or *pseudo-optimal* for  $H$  if  $\varphi$  is mean-self-financing and the martingale  $C(\varphi)$  is strongly orthogonal to  $M$ .

For  $d = 1$  and  $X$  sufficiently well-behaved, we have just seen that pseudo-optimal and locally risk-minimizing strategies are the same. But in general, pseudo-optimal strategies are both easier to find and to characterize. This is shown in the next result which is due to Föllmer/Schweizer (1991).

**Proposition 3.4.** *A contingent claim  $H \in L^2(\mathcal{F}_T, P)$  admits a pseudo-optimal  $L^2$ -strategy  $\varphi$  with  $V_T(\varphi) = H$   $P$ -a.s. if and only if  $H$  can be written as*

$$(3.7) \quad H = H_0 + \int_0^T \xi_u^H dX_u + L_T^H \quad P\text{-a.s.}$$

with  $H_0 \in L^2(\mathcal{F}_0, P)$ ,  $\xi^H \in \Theta_S$  and  $L^H \in \mathcal{M}_0^2(P)$  strongly  $P$ -orthogonal to  $M$ . The strategy  $\varphi$  is then given by

$$\vartheta_t = \xi_t^H \quad , \quad 0 \leq t \leq T$$

and

$$C_t(\varphi) = H_0 + L_t^H \quad , \quad 0 \leq t \leq T;$$

its value process is

$$(3.8) \quad V_t(\varphi) = C_t(\varphi) + \int_0^t \vartheta_u dX_u = H_0 + \int_0^t \xi_u^H dX_u + L_t^H \quad , \quad 0 \leq t \leq T$$

so that  $\eta$  is also determined by the above description.

**Proof.** This is Proposition (2.24) of Föllmer/Schweizer (1991), but for completeness we repeat here the simple proof: Write

$$H = V_T(\varphi) = C_T(\varphi) + \int_0^T \vartheta_u dX_u = C_0(\varphi) + \int_0^T \vartheta_u dX_u + C_T(\varphi) - C_0(\varphi)$$

and use the definition of pseudo-optimality.

**q.e.d.**

Quite apart from the connection to local risk-minimization, the decomposition (3.7) is in itself interesting. In the martingale case where  $A \equiv 0$  and  $M = X - X_0$ , it is the well-known Galtchouk-Kunita-Watanabe decomposition (2.1). In the general case, it has been called in the literature the *Föllmer-Schweizer decomposition* of  $H$  and has been studied by several authors. Sufficient conditions for its existence have for instance been given by Buckdahn (1993), Schweizer (1994a), Monat/Stricker (1995), Schweizer (1995a), Delbaen/Monat/Schachermayer/Schweizer/Stricker (1997) or Pham/Rheinländer/Schweizer (1998). The simplest sufficient condition is that the mean-variance tradeoff process  $\widehat{K}$  should be bounded uniformly in  $t$  and  $\omega$ ; see Theorem 3.4 of Monat/Stricker (1995). A survey of some results on the Föllmer-Schweizer decomposition has been given by Stricker (1996).

In view of Theorem 3.3 and Proposition 3.4, finding the Föllmer-Schweizer decomposition of a given contingent claim  $H$  is important because it allows one to obtain a locally risk-minimizing strategy under some additional assumptions. In Buckdahn (1993) or Schweizer

(1994a), the existence of this decomposition is proved by means of backward stochastic differential equations whereas Monat/Stricker (1995) and Pham/Rheinländer/Schweizer (1998) use a fixed point argument. But all these results do not provide a constructive way of finding  $\xi^H$  and  $L^H$  more explicitly. Following Föllmer/Schweizer (1991) and Schweizer (1995a), we therefore explain how one can often obtain (3.7) by switching to a suitably chosen martingale measure for  $X$ ; this notably works in the case where  $X$  is continuous and has a bounded mean-variance tradeoff. Moreover, this approach is in perfect analogy to the situation in discrete time.

Inspired by the difference equation (3.6), we consider the stochastic differential equation

$$d\widehat{Z}_t = -\widehat{Z}_t \widehat{\lambda}_t dM_t \quad , \quad \widehat{Z}_0 = 1.$$

Its unique strong solution is the stochastic exponential  $\widehat{Z} = \mathcal{E}\left(-\int \widehat{\lambda} dM\right)$ ; if  $X$  (hence also  $M$ ) is continuous, this is explicitly given by

$$\widehat{Z}_t = \exp\left(-\int_0^t \widehat{\lambda}_u dM_u - \frac{1}{2} \left\langle \int \widehat{\lambda} dM \right\rangle_t\right) = \exp\left(-\int_0^t \widehat{\lambda}_u dM_u - \frac{1}{2} \widehat{K}_t\right) \quad , \quad 0 \leq t \leq T.$$

It is well known and easily checked that  $\widehat{Z}$  is in general a locally square-integrable local  $P$ -martingale such that

$$(3.9) \quad \widehat{Z}X \text{ is a local } P\text{-martingale,}$$

$$\widehat{Z} \int \vartheta dX \text{ is a local } P\text{-martingale for every } \vartheta \in \Theta_S$$

and

$$(3.10) \quad \widehat{Z}L \text{ is a local } P\text{-martingale for every } L \in \mathcal{M}_{0,\text{loc}}^2(P) \text{ strongly } P\text{-orthogonal to } M;$$

see for instance Theorem (3.5) of Föllmer/Schweizer (1991) or Schweizer (1995a). By (3.8), this implies the analogue of (3.5) that

$$(3.11) \quad \text{for a pseudo-optimal } L^2\text{-strategy } \varphi \text{ for } H, \text{ the product } \widehat{Z}V(\varphi) \text{ is a local } P\text{-martingale.}$$

In the situation of (3.11),  $C(\varphi)$  is a martingale and  $\sup_{0 \leq t \leq T} |V_t(\varphi)| \in L^2(P)$ ; hence  $\widehat{Z}V(\varphi)$  is then a true martingale if  $\widehat{Z}$  itself is a square-integrable martingale.

So suppose now that  $\widehat{Z} \in \mathcal{M}^2(P)$ . A restrictive sufficient condition for this is by Theorem II.2 of Lepingle/Mémin (1978) uniform boundedness of  $\widehat{K}$  in  $t$  and  $\omega$ . In concrete applications, one can also try to check square-integrability directly. If  $\widehat{Z}$  is also strictly positive on  $[0, T]$

(which will certainly hold if  $M$ , hence  $\widehat{Z}$ , is continuous), then

$$(3.12) \quad \frac{d\widehat{P}}{dP} := \widehat{Z}_T = \mathcal{E} \left( -\int \widehat{\lambda} dM \right)_T \in L^2(P)$$

defines a probability measure  $\widehat{P} \approx P$  which is in  $\mathbb{P}$  according to (3.9). For reasons explained below, this measure  $\widehat{P}$  is called the *minimal equivalent local martingale measure* for  $X$ . Since the martingale form of (3.11) says that  $V(\varphi)$  is a  $\widehat{P}$ -martingale for a pseudo-optimal  $L^2$ -strategy  $\varphi$  for  $H$ , we get

$$(3.13) \quad V_t(\varphi) = \widehat{E}[H|\mathcal{F}_t] =: V_t^{H,\widehat{P}} \quad , \quad 0 \leq t \leq T$$

for such a strategy. Hence we are led to study the  $\widehat{P}$ -martingale  $V^{H,\widehat{P}}$  and its relation to the local  $\widehat{P}$ -martingale  $X$ . Note that  $H \in L^1(\widehat{P})$  because  $H$  and  $\widehat{Z}_T$  are both in  $L^2(P)$ ; hence  $V^{H,\widehat{P}}$  is indeed well-defined.

In addition to the previous assumptions, suppose now also that  $X$  is *continuous*. By (3.9),  $X$  is a local  $\widehat{P}$ -martingale and so  $V^{H,\widehat{P}}$  admits a Galtchouk-Kunita-Watanabe decomposition under  $\widehat{P}$  with respect to  $X$  as

$$(3.14) \quad V_t^{H,\widehat{P}} = V_0^{H,\widehat{P}} + \int_0^t \xi_u^{H,\widehat{P}} dX_u + L_t^{H,\widehat{P}} \quad , \quad 0 \leq t \leq T$$

where  $\xi^{H,\widehat{P}} \in L(X)$  and  $L^{H,\widehat{P}}$  is a local  $\widehat{P}$ -martingale null at 0 and strongly  $\widehat{P}$ -orthogonal to  $X$ ; see Ansel/Stricker (1993). For  $t = T$ , this gives in particular a decomposition of the random variable  $H$ . Thanks to the continuity of  $X$ ,  $L^{H,\widehat{P}}$  is also a local  $P$ -martingale strongly  $P$ -orthogonal to  $X$ ; see Ansel/Stricker (1992) or Schweizer (1995a). In many cases, this decomposition gives us what we need; this was already observed in Theorem (3.14) of Föllmer/Schweizer (1991).

**Theorem 3.5.** *Suppose that  $X$  is continuous and hence satisfies (SC) (because  $\mathbb{P} \neq \emptyset$ ). Define the strictly positive local  $P$ -martingale  $\widehat{Z} := \mathcal{E} \left( -\int \widehat{\lambda} dM \right)$  and suppose that*

$$(3.15) \quad \widehat{Z} \in \mathcal{M}^2(P).$$

*Define  $\widehat{P}$  and  $V^{H,\widehat{P}}$  as above by (3.12) and (3.13), respectively. If either*

$$(3.16) \quad H \text{ admits a Föllmer-Schweizer decomposition}$$

*or*

$$(3.17) \quad V_0^{H,\widehat{P}} \in L^2(P), \xi^{H,\widehat{P}} \in \Theta_S \text{ and } L^{H,\widehat{P}} \in \mathcal{M}^2(P),$$

then (3.14) for  $t = T$  gives the Föllmer-Schweizer decomposition of  $H$  and  $\xi^{H, \widehat{P}}$  determines a pseudo-optimal  $L^2$ -strategy for  $H$ . A sufficient condition for (3.15), (3.16) and (3.17) is that  $\widehat{K}$  is uniformly bounded.

**Proof.** This is almost a summary of the preceding arguments. If we have (3.16), then (3.10) implies that  $L^H$  is a local  $\widehat{P}$ -martingale and strongly  $\widehat{P}$ -orthogonal to  $X$ , since  $\langle L^H, X \rangle = \langle L^H, M \rangle = 0$  by the continuity of  $X$ . By the uniqueness of the Galtchouk-Kunita-Watanabe decomposition, (3.7) and (3.14) for  $t = T$  must therefore coincide. If we have (3.17), the argument just before Theorem 3.5 shows that (3.14) for  $t = T$  gives a Föllmer-Schweizer decomposition for  $H$  which by uniqueness must again coincide with (3.7). The assertion about  $\xi^{H, \widehat{P}}$  is then immediate from Proposition 3.4 and that boundedness of  $\widehat{K}$  is sufficient follows from Theorem II.2 of Lepingle/Mémin (1978), Theorem 3.4 of Monat/Stricker (1995) and Lemma 6 of Pham/Rheinländer/Schweizer (1998) respectively.

**q.e.d.**

The basic message of Theorem 3.5 is that for  $X$  continuous, finding a locally risk-minimizing strategy essentially boils down to finding the Galtchouk-Kunita-Watanabe decomposition of  $H$  under the minimal ELMM  $\widehat{P}$ . This is very useful because the density process  $\widehat{Z}$  of  $\widehat{P}$  with respect to  $P$  can immediately be written down explicitly and we can directly see the dynamics of  $X$  under  $\widehat{P}$ . In particular, finding (3.14) can often be reduced to solving a partial differential equation if  $H$  can be written as a function of the final value of some (possibly multidimensional) process which has a Markovian structure under  $\widehat{P}$ . This is explained in Pham/Rheinländer/Schweizer (1998) and for the case of a stochastic volatility model in more detail also in Heath/Platen/Schweizer (1998).

**Remark.** We emphasize that by its very nature, local risk-minimization is a *hedging approach* designed to control the riskiness of a strategy as measured by its local cost fluctuations. If there is an optimal strategy  $\varphi$ , we can use  $V_t(\varphi)$  as a *value* or price of  $H$  at time  $t$ , but two things about this should be kept in mind: Such a valuation is a by-product of the method, not its primary objective, and it is only a valuation with respect to the (subjective) criterion of local risk-minimization.

If we can obtain the Föllmer-Schweizer decomposition of  $H$  via the Galtchouk-Kunita-Watanabe decomposition of  $H$  under  $\widehat{P}$ , we know from (3.13) that the value process of the corresponding pseudo-optimal strategy  $\varphi$  is given by the conditional expectations of  $H$  under  $\widehat{P}$ . Together with the preceding remark, this shows that  $V^{H, \widehat{P}}$  can be interpreted as an *intrinsic valuation process* for  $H$  and identifies  $\widehat{P}$  as the valuation operator naturally associated with the criterion of local risk-minimization. It seems therefore appropriate to

comment briefly on the origins and properties of  $\widehat{P}$  and in particular on the terminology “minimal ELMM”.

The first formal definition of a minimal martingale measure appears in Föllmer/Schweizer (1991). They consider a continuous square-integrable real-valued process  $X$  and focus on equivalent martingale measures  $Q$  for  $X$  that satisfy  $\frac{dQ}{dP} \in L^2(P)$ . A martingale measure  $Q$  from this class is called *minimal* if  $Q = P$  on  $\mathcal{F}_0$  and if any  $L \in \mathcal{M}_0^2(P)$  strongly  $P$ -orthogonal to  $M$  is under  $Q$  still a martingale. Theorem (3.5) of Föllmer/Schweizer (1991) then proves that such a measure is unique and must coincide with  $\widehat{P}$  defined above; existence is therefore equivalent to  $\widehat{Z}$  being in  $\mathcal{M}^2(P)$ . These results have precursors in Schweizer (1988, 1991) for the special case where  $\mathcal{M}^2(P)$  is generated by  $M$  and a second orthogonal  $P$ -martingale  $N$ . In that context, the “minimal” martingale measure is introduced as an equivalent probability that turns  $X$  into a martingale and preserves the martingale property of  $N$ . The terminology “minimal” is there motivated by the fact that apart from turning  $X$  into a martingale, this measure disturbs the overall martingale and orthogonality structures as little as possible.

The original motivation in Schweizer (1988) for introducing a minimal martingale measure  $\widehat{P}$  was its use in finding locally risk-minimizing strategies via a variant of Theorem 3.5. It has subsequently turned out that  $\widehat{P}$  appears quite naturally in a number of other situations as well. Apart from local risk-minimization as discussed above, one can mention here logarithmic utility maximization problems (see Cvitanić/Karatzas (1992), Karatzas (1997), Amendinger/Imkeller/Schweizer (1998)), pricing under local utility indifference (see Davis (1994, 1997), Karatzas/Kou (1996)), equilibrium prices for assets (see Pham/Touzi (1996) or Jouini/Napp (1998)) and value preservation (see Korn (1997, 1998)). In view of this apparent ubiquity of  $\widehat{P}$ , it is natural to ask for a more concise and transparent description of  $\widehat{P}$ , preferably as the solution of a suitable optimization problem. This would give a more precise meaning to the sense in which  $\widehat{P}$  is optimal.

**Proposition 3.6.** *Let  $X$  be a continuous adapted process admitting at least one equivalent local martingale measure  $Q$ . If  $\widehat{P}$  defined by (3.12) is a probability measure equivalent to  $P$ , then  $\widehat{P}$  minimizes the reverse relative entropy  $H(P|Q)$  over all ELMMs  $Q$  for  $X$ .*

**Proof.** See Theorem 1 of Schweizer (1999a).

At present, this seems to be the most general known characterization of  $\widehat{P}$ . For the case of a multidimensional diffusion model for  $X$ , this can also be found in section 5.6 of Karatzas (1997), and Schweizer (1999a) contains a discussion of other less general results. A counterexample in Schweizer (1999a) shows that Proposition 3.6 does not carry over to the case where  $X$  is discontinuous. Finding an analogous description of  $\widehat{P}$  in general seems to be an open problem.

## 4. Mean-variance hedging

Let us now return to the general situation where  $X$  is a semimartingale under  $P$  and  $H$  is a given contingent claim. The key difference between (local) risk-minimization and mean-variance hedging is that we no longer impose on our trading strategies the replication requirement  $V_T = H$   $P$ -a.s., but insist instead on the self-financing constraint (1.1). For a self-financing pre-strategy  $(V_0, \vartheta)$ , the shortfall or loss from hedging  $H$  by  $(V_0, \vartheta)$  is then

$$H - V_T(V_0, \vartheta) = H - V_0 - \int_0^T \vartheta_u dX_u$$

and we want to minimize the  $L^2(P)$ -norm of this quantity by choosing  $(V_0, \vartheta)$ . Note that a symmetric criterion is quite natural in the present context of hedging and pricing options because one does not know at the start whether one is dealing with a buyer or a seller; see Bertsimas/Kogan/Lo (1999) for an amplification of this point. Choosing the  $L^2$ -norm is mainly for convenience because it allows fairly explicit results while at the same time leading to interesting mathematical questions. For brevity, we write  $L^2$  for  $L^2(P)$  if there is no risk of confusion.

We first have to be more specific about our strategies. We do not assume that  $\mathcal{F}_0$  is trivial but we insist on a non-random initial capital  $V_0$ .

**Definition.** We denote by  $\Theta_2$  the set of all  $\vartheta \in L(X)$  such that the stochastic integral process  $G(\vartheta) := \int \vartheta dX$  satisfies  $G_T(\vartheta) \in L^2(P)$ . For a fixed linear subspace  $\Theta$  of  $\Theta_2$ , a  $\Theta$ -strategy is a pair  $(V_0, \vartheta) \in \mathbb{R} \times \Theta$  and its value process is  $V_0 + G(\vartheta)$ . A  $\Theta$ -strategy  $(\tilde{V}_0, \tilde{\vartheta})$  is called  $\Theta$ -mean-variance optimal for a given contingent claim  $H \in L^2$  if it minimizes  $\|H - V_0 - G_T(\vartheta)\|_{L^2}$  over all  $\Theta$ -strategies  $(V_0, \vartheta)$  and  $\tilde{V}_0$  is then called the  $\Theta$ -approximation price for  $H$ .

The preceding definition depends on the choice of the space  $\Theta$  of strategies allowed for trading and we shall be more specific about this later on. For the moment, however, we go in the other direction and consider an even more general framework. Suppose we have chosen a linear subspace  $\Theta$  of  $\Theta_2$ . Then the linear subspace

$$\mathcal{G} := G_T(\Theta) = \left\{ \int_0^T \vartheta_u dX_u \mid \vartheta \in \Theta \right\}$$

of  $L^2$  describes all outcomes of self-financing  $\Theta$ -strategies with initial wealth  $V_0 = 0$  and

$$\mathcal{A} := \mathbb{R} + \mathcal{G} = \left\{ V_0 + \int_0^T \vartheta_u dX_u \mid (V_0, \vartheta) \in \mathbb{R} \times \Theta \right\}$$

is the space of contingent claims replicable by self-financing  $\Theta$ -strategies. Our goal in mean-variance hedging is to find the projection in  $L^2$  of  $H$  on  $\mathcal{A}$  and this can be studied for a general

linear subspace  $\mathcal{G}$  of  $L^2$ . In analogy to the above definition, we introduce a  $\mathcal{G}$ -mean-variance optimal pair  $(\tilde{V}_0, \tilde{g}) \in \mathbb{R} \times \mathcal{G}$  for  $H \in L^2$  and call  $\tilde{V}_0$  the  $\mathcal{G}$ -approximation price for  $H$ . In particular, we need no explicit model for  $X$  or  $\Theta$  at this stage and both a discrete-time or a continuous-time choice for  $X$  fit equally well into this setting. This was first pointed out in Schweizer (1998) and exploited in Schweizer (1999b). Our presentation here follows the latter.

**Definition.** We say that  $\mathcal{G}$  admits no approximate profits in  $L^2$  if  $\bar{\mathcal{G}}$  does not contain the constant 1; the bar  $\bar{\cdot}$  denotes the closure in  $L^2$ .

With our preceding interpretations, this notion is very intuitive: It says that one cannot approximate (in the  $L^2$ -sense) the riskless payoff 1 by a self-financing strategy with initial wealth 0. This is a no-arbitrage condition on the financial market underlying  $\mathcal{G}$ ; see also Stricker (1990).

**Definition.** A signed  $\mathcal{G}$ -martingale measure is a signed measure  $Q$  on  $(\Omega, \mathcal{F})$  with  $Q[\Omega] = 1$ ,  $Q \ll P$  with  $\frac{dQ}{dP} \in L^2$  and

$$E_Q[g] = E \left[ \frac{dQ}{dP} g \right] = 0 \quad \text{for all } g \in \mathcal{G}.$$

$\mathbb{P}_s^2(\mathcal{G})$  denotes the convex set of all signed  $\mathcal{G}$ -martingale measures and an element  $\tilde{P}_G$  of  $\mathbb{P}_s^2(\mathcal{G})$  is called *variance-optimal* if it minimizes  $\left\| \frac{dQ}{dP} \right\|_{L^2} = \sqrt{1 + \text{Var} \left[ \frac{dQ}{dP} \right]}$  over all  $Q \in \mathbb{P}_s^2(\mathcal{G})$ .

**Lemma 4.1.** *Let  $\mathcal{G}$  be a linear subspace of  $L^2$ . Then:*

- 1)  $\mathcal{G}$  admits no approximate profits in  $L^2$  if and only if  $\mathbb{P}_s^2(\mathcal{G}) \neq \emptyset$ .
- 2) If  $\mathcal{G}$  admits no approximate profits in  $L^2$ , then  $\bar{\mathcal{A}} = \mathbb{R} + \bar{\mathcal{G}}$ .
- 3) If  $\mathcal{G}$  admits no approximate profits in  $L^2$ , then the variance-optimal signed  $\mathcal{G}$ -martingale measure  $\tilde{P}_G$  exists, is unique and satisfies

$$(4.1) \quad \frac{d\tilde{P}_G}{dP} \in \bar{\mathcal{A}}.$$

**Proof.** This very simple result goes back to Delbaen/Schachermayer (1996a) and Schweizer (1998); for completeness, we reproduce here the detailed proof of Schweizer (1999b). We use  $(\cdot, \cdot)$  for the scalar product in  $L^2$ .

1) An element  $Q$  of  $\mathbb{P}_s^2(\mathcal{G})$  can be identified with a continuous linear functional  $\Psi$  on  $L^2$  satisfying  $\Psi = 0$  on  $\mathcal{G}$  and  $\Psi(1) = 1$  by setting  $\Psi(U) = E \left[ \frac{dQ}{dP} U \right] = \left( \frac{dQ}{dP}, U \right)$ . Hence 1) is clear from the Hahn-Banach theorem.

2) Any  $g \in \bar{\mathcal{G}}$  is the limit in  $L^2$  of a sequence  $(g_n)$  in  $\mathcal{G}$ ; hence  $c + g_n = a_n$  is a Cauchy sequence in  $\mathcal{A}$  and thus converges in  $L^2$  to a limit  $a \in \bar{\mathcal{A}}$  so that  $c + g = a \in \bar{\mathcal{A}}$ . This gives the inclusion “ $\supseteq$ ” in general. For the converse, we use the assumption that  $\mathcal{G}$  admits no approximate profits in  $L^2$  to obtain from part 1) a signed  $\mathcal{G}$ -martingale measure  $Q$ . The random variable  $Z := \frac{dQ}{dP}$  is then in  $\mathcal{G}^\perp$  and satisfies  $(Z, 1) = Q[\Omega] = 1$ . For any  $a \in \bar{\mathcal{A}}$ , there is a sequence  $a_n = c_n + g_n$  in  $\mathcal{A}$  converging to  $a$  in  $L^2$ . Since  $c_n + g_n \in \mathbb{R} + \mathcal{G}$  for all  $n$ , we conclude that  $c_n = (c_n + g_n, Z) = (a_n, Z)$  converges in  $\mathbb{R}$  to  $(a, Z) =: c$ . Therefore  $g_n = a_n - c_n$  converges in  $L^2$  to  $g := a - c$  and since this limit is in  $\bar{\mathcal{G}}$ , we have  $a = c + g \in \mathbb{R} + \bar{\mathcal{G}}$  which proves the inclusion “ $\subseteq$ ”.

3) Existence and uniqueness of  $\tilde{P}_{\mathcal{G}}$  are clear once we observe that we have to minimize  $\|Z\|$  over the closed convex set  $\mathcal{Z} := \left\{ Z = \frac{dQ}{dP} \mid Q \in \mathbb{P}_s^2(\mathcal{G}) \right\}$  which is non-empty thanks to 1). For any fixed  $Z_0 \in \mathcal{Z}$ , the projection  $\tilde{Z}$  of  $Z_0$  in  $L^2$  on  $\bar{\mathcal{A}}$  is again in  $\mathcal{Z}$ ; in fact, one easily verifies that  $\tilde{\Psi}(U) := (\tilde{Z}, U)$  is 0 on  $\mathcal{G}$  and has  $\tilde{\Psi}(1) = 1$ . Since part 2) tells us that  $\tilde{Z} = \tilde{c} + \tilde{g}$  with  $\tilde{g} \in \bar{\mathcal{G}}$ , we obtain  $(Z, \tilde{Z}) = \tilde{c} = (\tilde{Z}, \tilde{Z})$  for all  $Z \in \mathcal{Z}$  and therefore

$$\|Z\|^2 = \|\tilde{Z}\|^2 + \|Z - \tilde{Z}\|^2 \geq \|\tilde{Z}\|^2 \quad \text{for all } Z \in \mathcal{Z}.$$

Hence we conclude that  $\frac{d\tilde{P}_{\mathcal{G}}}{dP} = \tilde{Z}$  is in  $\bar{\mathcal{A}}$ .

**q.e.d.**

For any  $g \in \mathcal{G}$  and any  $Q \in \mathbb{P}_s^2(\mathcal{G})$ , we have

$$1 = E_Q[1 - g] = E \left[ \frac{dQ}{dP} (1 - g) \right] \leq \left\| \frac{dQ}{dP} \right\|_{L^2} \|1 - g\|_{L^2}$$

by the Cauchy-Schwarz inequality and therefore

$$\frac{1}{\inf_{Q \in \mathbb{P}_s^2(\mathcal{G})} \left\| \frac{dQ}{dP} \right\|_{L^2}} = \sup_{Q \in \mathbb{P}_s^2(\mathcal{G})} \frac{1}{\left\| \frac{dQ}{dP} \right\|_{L^2}} \leq \inf_{g \in \mathcal{G}} \|1 - g\|_{L^2}.$$

This indicates that finding the variance-optimal signed  $\mathcal{G}$ -martingale measure is the dual problem to approximating in  $L^2$  the constant 1 by elements of  $\mathcal{G}$ . This duality is reflected in the next result which gives the  $\mathcal{G}$ -approximation price as an expectation under  $\tilde{P}_{\mathcal{G}}$ .

**Proposition 4.2.** *Suppose that  $\mathcal{G}$  is a linear subspace of  $L^2$  which admits no approximate profits in  $L^2$ . If a contingent claim  $H \in L^2$  admits a  $\mathcal{G}$ -mean-variance optimal pair  $(\tilde{V}_0, \tilde{g})$ , the  $\mathcal{G}$ -approximation price of  $H$  is given by*

$$\tilde{V}_0 = \tilde{E}_{\mathcal{G}}[H]$$

where  $\tilde{E}_{\mathcal{G}}$  denotes expectation under the variance-optimal signed  $\mathcal{G}$ -martingale measure  $\tilde{P}_{\mathcal{G}}$ .

**Proof.** If  $H$  admits a  $\mathcal{G}$ -mean-variance optimal pair  $(\tilde{V}_0, \tilde{g})$ , then  $\tilde{V}_0 + \tilde{g}$  is the projection in  $L^2$  of  $H$  on  $\bar{\mathcal{A}} = \mathbb{R} + \bar{\mathcal{G}}$  by Lemma 4.1. Since  $H - \tilde{V}_0 - \tilde{g}$  is then in the orthogonal complement of  $\bar{\mathcal{A}}$ , (4.1) implies that  $E \left[ (H - \tilde{V}_0 - \tilde{g}) \frac{d\tilde{P}_{\mathcal{G}}}{dP} \right] = 0$  and so we obtain

$$\tilde{V}_0 = E \left[ (H - \tilde{g}) \frac{d\tilde{P}_{\mathcal{G}}}{dP} \right] = \tilde{E}_{\mathcal{G}}[H]$$

because  $\tilde{P}_{\mathcal{G}}$  is in  $\mathbb{P}_s^2(\mathcal{G})$ .

**q.e.d.**

The assumption in Proposition 4.2 that  $H$  admits a  $\mathcal{G}$ -mean-variance optimal pair is obviously unpleasant. We can avoid it by either working a priori with elements from the *closed* linear subspace  $\bar{\mathcal{A}} = \mathbb{R} + \bar{\mathcal{G}}$  or by ensuring in some way that  $\mathcal{G}$  (hence also  $\mathcal{A}$ ) is already closed in  $L^2$ . The simpler first solution is preferable if we are not directly interested in the structure of the optimal element  $\tilde{V}_0 + \tilde{g}$ . This is the case in most situations where we only want to *value* contingent claims by using some quadratic criterion; see for instance Mercurio (1996), Aurell/Simdyankin (1998), Schweizer (1998) or Schweizer (1999b). But for *hedging* purposes, we also want to understand  $\tilde{g}$  itself and therefore we follow here the second idea and return to the framework with a semimartingale  $X$  and a space  $\Theta \subseteq \Theta_2$  of integrands to study the closedness of  $G_T(\Theta)$  in  $L^2$ .

So let  $X = (X_t)_{0 \leq t \leq T}$  be an  $\mathbb{R}^d$ -valued semimartingale which is locally in  $L^2(P)$  in the sense that the maximal process  $X_t^* := \sup_{0 \leq s \leq t} |X_s|$ ,  $0 \leq t \leq T$ , is locally  $P$ -square-integrable.

Let  $(\varrho_n)_{n \in \mathbb{N}}$  be a corresponding localizing sequence of stopping times. A process of the form  $\vartheta = \xi I_{\llbracket \sigma, \tau \rrbracket}$  with  $\sigma \leq \tau$  stopping times with  $\tau \leq \varrho_n$  for some  $n$  and with a bounded  $\mathbb{R}^d$ -valued  $\mathcal{F}_\sigma$ -measurable random variable  $\xi$  is called a *simple integrand* and we denote by  $\Theta_{\text{simple}}$  the linear space spanned by all simple integrands. It is evident that  $\Theta_{\text{simple}} \subseteq \Theta_2$  and easy to verify that  $Q$  is an ELMM for  $X$  with  $\frac{dQ}{dP} \in L^2(P)$  if and only if  $Q$  is in  $\mathbb{P}_s^2(\Theta_{\text{simple}})$  and  $Q \approx P$ . We denote the set of all these probability measures  $Q$  by  $\mathbb{P}_e^2(X)$ .

**Definition.** The *variance-optimal signed martingale measure*  $\tilde{P}$  for  $X$  is defined as the variance-optimal signed  $G_T(\Theta_{\text{simple}})$ -martingale measure.

In general,  $\tilde{P}$  is unfortunately a signed measure. But for a continuous process  $X$ , the situation is better.

**Theorem 4.3.** *If  $X$  is a continuous  $\mathbb{R}^d$ -valued semimartingale and  $\mathbb{P}_e^2(X) \neq \emptyset$ , then  $\tilde{P}$  is*

in  $\mathbb{P}_e^2(X)$ . In other words, the variance-optimal signed martingale measure for  $X$  is then automatically equivalent to  $P$  and in particular a probability measure.

**Proof.** See Theorem 1.3 of Delbaen/Schachermayer (1996a).

In order to study the closedness in  $L^2$  of  $\mathcal{G} := G_T(\Theta)$  and also to relate  $\tilde{P}$  to  $\tilde{P}_{\mathcal{G}}$ , we now consider two specific choices of  $\Theta$ .

**Definition.**  $\Theta_{\text{GLP}}$  consists of all  $\vartheta \in L(X)$  such that  $G_T(\vartheta)$  is in  $L^2(P)$  and the process  $G(\vartheta) = \int \vartheta dX$  is a uniformly  $Q$ -integrable  $Q$ -martingale for every  $Q \in \mathbb{P}_e^2(X)$ .  $\Theta_{\text{S}}$  consists (as in section 3) of all  $\vartheta \in L(X)$  such that  $G(\vartheta)$  is in the space  $\mathcal{S}^2(P)$  of semimartingales.

The space  $\Theta_{\text{S}}$  has been introduced by Schweizer (1994a). At first sight, it appears simpler and more natural because it can be defined directly in terms of the original probability measure  $P$ . Moreover, it obviously generalizes the space  $L^2(X)$  used in section 2 for the martingale case to the semimartingale framework. The space  $\Theta_{\text{GLP}}$  was first used by Delbaen/Schachermayer (1996b) and introduced to hedging by Gouriéroux/Laurent/Pham (1998). Its main advantage (as illustrated by the next two results) is that it is better adapted to duality formulations and easier to handle for certain theoretical aspects. On the other hand, proving for an explicitly given strategy  $\vartheta$  that it is in  $\Theta$  is usually much simpler for  $\Theta = \Theta_{\text{S}}$  than for  $\Theta = \Theta_{\text{GLP}}$ . For additional results on the relation between  $\Theta_{\text{S}}$  and  $\Theta_{\text{GLP}}$ , see also Rheinländer (1999).

**Theorem 4.4.** *Let  $X$  be an  $\mathbb{R}^d$ -valued semimartingale which is locally in  $L^2(P)$  and assume that  $\mathbb{P}_e^2(X) \neq \emptyset$ . Then  $G_T(\Theta_{\text{GLP}})$  is closed in  $L^2(P)$ . If  $X$  is continuous, we have in addition that  $G_T(\Theta_{\text{GLP}}) = \overline{G_T(\Theta_{\text{simple}})}$  where the bar  $\bar{\phantom{x}}$  denotes the closure in  $L^2(P)$ ; this implies in particular that  $\tilde{P} = \tilde{P}_{G_T(\Theta_{\text{GLP}})}$ .*

**Proof.** This is due to Delbaen/Schachermayer (1996b). The first assertion follows from the equivalence of (i) and (ii) in their Theorem 1.2 (note that their  $D_2$  is always closed in  $L^2$ ) and the second uses in addition their Theorem 2.2.

**q.e.d.**

For  $\Theta_{\text{S}}$  instead of  $\Theta_{\text{GLP}}$ , analyzing the closedness question is more delicate.

**Definition.** Let  $Z = (Z_t)_{0 \leq t \leq T}$  be a strictly positive  $P$ -martingale with  $E[Z_0] = 1$ . We say that  $Z$  satisfies the *reverse Hölder inequality*  $R_2(P)$  if there is some constant  $C$  such that

$$E [Z_T^2 | \mathcal{F}_t] \leq CZ_t^2 \quad P\text{-a.s.}$$

for each  $t \in [0, T]$ . A probability measure  $Q \approx P$  is said to satisfy  $R_2(P)$  if its density process  $Z_t^Q := E \left[ \frac{dQ}{dP} \middle| \mathcal{F}_t \right]$ ,  $0 \leq t \leq T$ , satisfies  $R_2(P)$ .

**Theorem 4.5.** *Let  $X$  be a continuous  $\mathbb{R}^d$ -valued semimartingale. Then the following statements are equivalent:*

- 1)  $\mathbb{P}_e^2(X) \neq \emptyset$  and  $G_T(\Theta_S)$  is closed in  $L^2(P)$ .
- 2) There exists some  $Q \in \mathbb{P}_e^2(X)$  satisfying  $R_2(P)$ .
- 3) The variance-optimal martingale measure  $\tilde{P}$  is in  $\mathbb{P}_e^2(X)$  and satisfies  $R_2(P)$ .

**Proof.** This is a partial statement of Theorem 4.1 of Delbaen/Monat/Schachermayer/Schweizer/Stricker (1997).

Once we know that  $G_T(\Theta)$  is closed and does not contain 1, we can obtain  $\Theta$ -mean-variance optimal  $\Theta$ -strategies  $(\tilde{V}_0, \tilde{\vartheta})$  by projecting the given contingent claim  $H \in L^2$  on the space  $\mathcal{A}$  of replicable claims and it becomes interesting to study the structure of the optimal integrand  $\tilde{\vartheta}$  in more detail. Before we do this, let us briefly mention some more recent extensions of the preceding results. It is natural to replace the exponent 2 by  $p \in (1, \infty)$  in the definition of  $\Theta_S$  and to ask if  $G_T(\Theta_S)$  is then closed in  $L^p(P)$ . For the case where  $X$  is continuous, this has been treated in Grandits/Krawczyk (1998) who generalized Theorem 4.5 to an arbitrary  $p \in (1, \infty)$ . The next step is then to eliminate the assumption that  $X$  is continuous. This has been done in Choulli/Krawczyk/Stricker (1998, 1999) who first extended the Doob, Burkholder-Davis-Gundy and Fefferman inequalities from (local) martingales to a class of semimartingales (called  $\mathcal{E}$ -martingales) with a particular structure inspired by the financial background of the problem. They then used this to provide sufficient conditions for the closedness of  $G_T(\Theta_S)$  in  $L^p(P)$  when  $X$  is an  $\mathcal{E}$ -martingale. Moreover, they also generalized earlier results by Delbaen/Monat/Schachermayer/Schweizer/Stricker (1997) on the existence and continuity of the Föllmer-Schweizer decomposition. The problem of finding necessary and sufficient conditions for  $G_T(\Theta_S)$  to be closed in this general setting seems at present still open.

Let us now turn to the problem of finding the integrand  $\tilde{\vartheta}$  in the projection of a given  $H \in L^2$  on the space  $\mathcal{A} = \mathbb{R} + G_T(\Theta)$ . For the case where  $X = (X_k)_{k=0,1,\dots,T}$  is a real-valued square-integrable process in discrete time with a bounded mean-variance tradeoff, explicit recursive formulae for  $\tilde{\vartheta}$  have been given in Schweizer (1995b). These results are for the one-dimensional case  $d = 1$ ; the extension to  $d > 1$  has been worked out and will be presented elsewhere. See also Bertsimas/Kogan/Lo (1999) and Černý (1999) for recent results obtained via dynamic programming arguments. If  $X = (X_t)_{0 \leq t \leq T}$  is an  $\mathbb{R}^d$ -valued semimartingale, the above recursive expressions take under some additional assumptions the

form of a backward stochastic differential equation; see Schweizer (1994a, 1996) for more details. Both types of results simplify considerably if  $\log X$  is a Lévy process in either discrete or continuous time and  $H$  has a particular structure; this has been worked out by Hubalek/Krawczyk (1998). Theoretical and numerical results for mean-variance optimal strategies can be found in Biagini/Guasoni/Pratelli (1999), Guasoni/Biagini (1999) and Heath/Platen/Schweizer (1998) for the case of a stochastic volatility model and more numerically oriented studies in diffusion or jump-diffusion models have been done by Bertsimas/Kogan/Lo (1999), Grünewald/Trautmann (1997) and Hipp (1996, 1998). Additional references can also be found after the next theorem.

The most general results on  $\tilde{\vartheta}$  have been obtained for the case where  $X$  is *continuous* and  $\mathbb{P}_e^2(X) \neq \emptyset$ . By Theorem 4.3, the variance-optimal martingale measure  $\tilde{P}$  for  $X$  then exists and is equivalent to  $P$ . Moreover, the arguments in Delbaen/Schachermayer (1996a) also show that the process

$$\tilde{Z}_t := \tilde{E} \left[ \frac{d\tilde{P}}{dP} \middle| \mathcal{F}_t \right] \quad , \quad 0 \leq t \leq T$$

can be written as

$$\tilde{Z}_t = \tilde{Z}_0 + \int_0^t \tilde{\zeta}_u dX_u \quad , \quad 0 \leq t \leq T$$

for some  $\tilde{\zeta} \in \Theta_{\text{GLP}}$ . In particular,  $\tilde{Z}$  is continuous. Note also that (4.1) implies that  $\tilde{Z}_0$  is a non-random constant. As the next result shows,  $\tilde{P}$ ,  $\tilde{Z}$  and  $\tilde{\zeta}$  all turn up in the solution of the mean-variance hedging problem.

**Theorem 4.6.** *Suppose that  $X$  is a continuous process such that  $\mathbb{P}_e^2(X) \neq \emptyset$ . Let  $H \in L^2(P)$  be a contingent claim and write the Galtchouk-Kunita-Watanabe decomposition of  $H$  under  $\tilde{P}$  with respect to  $X$  as*

$$(4.2) \quad H = \tilde{E}[H|\mathcal{F}_0] + \int_0^T \xi_u^{H,\tilde{P}} dX_u + L_T^{H,\tilde{P}} = V_T^{H,\tilde{P}}$$

with

$$V_t^{H,\tilde{P}} := \tilde{E}[H|\mathcal{F}_t] = \tilde{E}[H|\mathcal{F}_0] + \int_0^t \xi_u^{H,\tilde{P}} dX_u + L_t^{H,\tilde{P}} \quad , \quad 0 \leq t \leq T.$$

Then the mean-variance optimal  $\Theta_{\text{GLP}}$ -strategy for  $H$  is given by

$$(4.3) \quad \tilde{V}_0 = \tilde{E}[H]$$

and

$$\begin{aligned}
(4.4) \quad \tilde{\vartheta}_t &= \xi_t^{H, \tilde{P}} - \frac{\tilde{\zeta}_t}{\tilde{Z}_t} \left( V_{t-}^{H, \tilde{P}} - \tilde{E}[H] - \int_0^t \tilde{\vartheta}_u dX_u \right) \\
&= \xi_t^{H, \tilde{P}} - \tilde{\zeta}_t \left( \frac{V_0^{H, \tilde{P}} - \tilde{E}[H]}{\tilde{Z}_0} + \int_0^{t-} \frac{1}{\tilde{Z}_u} dL_u^{H, \tilde{P}} \right) \quad , \quad 0 \leq t \leq T.
\end{aligned}$$

**Proof.** Thanks to Theorem 4.4, (4.3) follows immediately from Proposition 4.2. According to Corollary 16 of Schweizer (1996),  $\tilde{\vartheta}$  is obtained by projecting the random variable  $H - \tilde{E}[H]$  on  $G_T(\Theta)$  and this is in principle dealt with in Rheinländer/Schweizer (1997). The representation (4.4) is very similar to their Theorem 6, but we cannot directly use their results since they work with  $\Theta_S$  instead of  $\Theta_{GLP}$ . Thus we appeal to some results from Gouriéroux/Laurent/Pham (1998) and this involves a second change of measure. Because  $\tilde{Z}$  is a strictly positive  $P$ -martingale and  $\tilde{Z}_0$  is deterministic, we can define a new probability measure  $\tilde{R} \approx \tilde{P} \approx P$  by setting

$$\frac{d\tilde{R}}{d\tilde{P}} := \frac{\tilde{Z}_T}{\tilde{Z}_0}.$$

Clearly, the  $\mathbb{R}^{d+1}$ -valued process  $Y = \begin{pmatrix} 1/\tilde{Z} \\ X/\tilde{Z} \end{pmatrix}$  is then a continuous local  $\tilde{R}$ -martingale since  $\tilde{P} \in \mathcal{I}_e^2(X)$ . The density of  $\tilde{R}$  with respect to  $P$  is  $\tilde{Z}_T^2/\tilde{Z}_0$  and because  $\tilde{Z}_0$  is deterministic,  $H$  is in  $L^2(P)$  if and only if  $H/\tilde{Z}_T$  is in  $L^2(\tilde{R})$ . The basic idea of Gouriéroux/Laurent/Pham (1998) is now to use  $\tilde{Z}/\tilde{Z}_0$  as a new numeraire, rewrite the original problem in terms of the corresponding new quantities and apply the Galtchouk-Kunita-Watanabe decomposition theorem to  $H/\tilde{Z}_T$  under  $\tilde{R}$  with respect to  $Y$ . This yields

$$(4.5) \quad \frac{H}{\tilde{Z}_T} = E_{\tilde{R}} \left[ \frac{H}{\tilde{Z}_T} \middle| \mathcal{F}_0 \right] + \int_0^T \psi_u dY_u + L_T$$

for some  $\mathbb{R}^{d+1}$ -valued  $\psi \in L(Y)$  such that  $\int \psi dY \in \mathcal{M}_0^2(\tilde{R})$  and some  $L \in \mathcal{M}_0^2(\tilde{R})$  strongly  $\tilde{R}$ -orthogonal to  $Y$ . According to Theorem 5.1 and the subsequent remark in Gouriéroux/Laurent/Pham (1998),  $\tilde{\vartheta}$  is then given by

$$(4.6) \quad \tilde{\vartheta}_t^i = \psi_t^i + \tilde{\zeta}_t^i \left( \frac{\tilde{E}[H]}{\tilde{Z}_0} + \int_0^t \psi_u dY_u - \psi_t^{\text{tr}} Y_t \right) \quad , \quad 0 \leq t \leq T, i = 1, \dots, d$$

if we note that the relation between their terminology and ours is given by  $V(\tilde{a}) = \tilde{Z}/\tilde{Z}_0$ ,  $X^i(\tilde{a}) = \tilde{Z}_0 Y^i$  and  $\tilde{a} = -\tilde{\zeta}/\tilde{Z}$ . By using Proposition 8 of Rheinländer/Schweizer (1997),

(4.6) can be rewritten as

$$(4.7) \quad \tilde{\vartheta} = \frac{\tilde{E}[H]}{\tilde{Z}_0} \tilde{\zeta} + \vartheta$$

with  $\vartheta$  corresponding to  $\psi$  from (4.5) via equation (4.6) in Rheinländer/Schweizer (1997). Hence it only remains to obtain  $\vartheta$  or  $\psi$  in terms of the decomposition (4.5) and this is basically already contained in Rheinländer/Schweizer (1997) if one looks carefully enough. More precisely, we start from (4.5) and argue as in Proposition 10 of Rheinländer/Schweizer (1997) to express the quantities in the decomposition (4.2) in terms of  $\psi$  and  $L$ . Note that as long as we make no integrability assertions, that argument only uses Proposition 8 of Rheinländer/Schweizer (1997) which holds as soon as  $\mathcal{P}_e^2(X) \neq \emptyset$ ; see Remark 2) following that Proposition 8. The uniqueness of the Galtchouk-Kunita-Watanabe decomposition then implies that

$$L_t^{H, \tilde{P}} = \int_0^t \tilde{Z}_u dL_u \quad , \quad 0 \leq t \leq T$$

and

$$\xi_t^{H, \tilde{P}} = \frac{V_0^{H, \tilde{P}}}{\tilde{Z}_0} \tilde{\zeta}_t + \vartheta_t + L_{t-} \tilde{\zeta}_t \quad , \quad 0 \leq t \leq T;$$

note that we have to replace  $\tilde{E}[H]$  in equation (4.14) of Rheinländer/Schweizer (1997) by  $V_0^{H, \tilde{P}}$  since  $\mathcal{F}_0$  need not be trivial. Solving this for  $\vartheta$  and plugging the result into (4.7) yields the second expression in (4.4). The first then follows similarly as in the proof of Theorem 6 of Rheinländer/Schweizer (1997); we again have to replace there  $\tilde{E}[H]$  by  $V_0^{H, \tilde{P}}$ .

**q.e.d.**

While Theorem 4.6 does give a reasonably constructive description of the strategy  $\tilde{\vartheta}$ , it is still not completely satisfactory. For continuous-time processes with discontinuous trajectories, hardly anything is known about  $\tilde{\vartheta}$  except under quite restrictive additional assumptions on  $X$ . Fairly explicit expressions have been found by Hubalek/Krawczyk (1998) if  $X$  is an exponential Lévy process. This relies on earlier results in Schweizer (1994a) who obtained an analogue to (4.4) for the case where  $X$  has a deterministic mean-variance tradeoff; see also Grünewald (1998) who used this in a jump-diffusion setting. Somewhat more generally, Hipp (1993, 1996), Wiese (1998) and Pham/Rheinländer/Schweizer (1998) studied the special case where the minimal martingale measure  $\hat{P}$  and the variance-optimal martingale measure  $\tilde{P}$  coincide. But at present, finding  $\tilde{\vartheta}$  in general is an open problem.

At least for continuous processes  $X$ , Theorem 4.6 makes it clear that a key role in determining  $\tilde{\vartheta}$  is played by the variance-optimal martingale measure  $\tilde{P}$ . For one thing, we need the Galtchouk-Kunita-Watanabe decomposition of  $H$  under  $\tilde{P}$  just as we needed the

Galtchouk-Kunita-Watanabe decomposition of  $H$  under  $\widehat{P}$  in section 3 to find locally risk-minimizing strategies. (This partly explains why the case  $\widetilde{P} = \widehat{P}$  is still solvable.) Thus we have to understand the behaviour of  $X$  under  $\widetilde{P}$  and therefore also the structure of  $\widetilde{P}$  itself in more detail. In addition, the latter is also required for finding  $\widetilde{\zeta}$  and  $\widetilde{Z}$  that appear in (4.4). We first recall a rather special case treated by Pham/Rheinländer/Schweizer (1998).

**Lemma 4.7.** *Suppose that  $X$  is a continuous process such that  $\mathbb{P}_e^2(X) \neq \emptyset$ . For  $Q \in \{\widehat{P}, \widetilde{P}\}$ , we denote by  $Z_t^Q := E\left[\frac{dQ}{dP} \middle| \mathcal{F}_t\right]$ ,  $0 \leq t \leq T$ , the density process of  $Q$  with respect to  $P$ . If the final value  $\widehat{K}_T$  of the mean-variance tradeoff is deterministic, then  $\widetilde{P} = \widehat{P}$ ,*

$$\begin{aligned} Z_t^{\widetilde{P}} = Z_t^{\widehat{P}} = \widehat{Z}_t &= \mathcal{E}\left(-\int \widehat{\lambda} dM\right)_t, & 0 \leq t \leq T, \\ \widetilde{Z}_t = \widehat{E}\left[\frac{d\widetilde{P}}{dP} \middle| \mathcal{F}_t\right] &= e^{\widehat{K}_T} \mathcal{E}\left(-\int \widehat{\lambda} dX\right)_t, & 0 \leq t \leq T, \\ \widetilde{\zeta}_t = -e^{\widehat{K}_T} \mathcal{E}\left(-\int \widehat{\lambda} dX\right)_t \widehat{\lambda}_t &= -\widetilde{Z}_t \widehat{\lambda}_t, & 0 \leq t \leq T \end{aligned}$$

and

$$\frac{Z_t^{\widetilde{P}}}{\widetilde{Z}_t} = e^{-(\widehat{K}_T - \widehat{K}_t)}, \quad 0 \leq t \leq T.$$

**Proof.** Because  $X$  satisfies (SC), the three middle results are simply reformulations of subsection 4.2 of Pham/Rheinländer/Schweizer (1998). The equality of  $\widetilde{P}$  and  $\widehat{P}$  is a consequence of the last remark in section 3 of Pham/Rheinländer/Schweizer (1998) and the last result follows because

$$\widetilde{\zeta}_t = e^{\widehat{K}_T} \mathcal{E}\left(-\int \widehat{\lambda} dM - \widehat{K}\right)_t = e^{\widehat{K}_T} Z_t^{\widetilde{P}} e^{-\widehat{K}_t}.$$

**q.e.d.**

Although Lemma 4.7 is a pleasingly simple result, its assumption is usually too restrictive for practical applications. More general results have been obtained by Laurent/Pham (1999) in a multidimensional diffusion model by dynamic programming arguments. They show how one can represent the ratio process  $\widetilde{Z}/Z^{\widetilde{P}}$  as the solution of a dynamic optimization problem and how its canonical decomposition determines the ratio  $\widetilde{\zeta}/\widetilde{Z}$ . Current work in progress is aimed at extending these results to general continuous semimartingales, but there still remains a lot to be done because no really explicit results have been found so far. If we consider for instance a *stochastic volatility* model for  $X$ , the currently available techniques only work in the case where  $X$  and its volatility are uncorrelated. This unfortunately excludes most models of interest for practical applications and illustrates the need for more research in

this area. For additional details and more recent work, we refer to Biagini/Guasoni/Pratelli (1999), Guasoni/Biagini (1999), Heath/Platen/Schweizer (1998) and Laurent/Pham (1999).

**Acknowledgements.** Instead of putting up a very long list of people who would all deserve thanks, I apologize to all those whose work I have forgotten or misrepresented in any way. Thomas Møller pointed out the need to have  $\mathcal{F}_0$  nontrivial in section 4 and Christophe Stricker was as usual extremely helpful with comments and hints on technical issues.

## References

J. Amendinger, P. Imkeller and M. Schweizer (1998), “Additional Logarithmic Utility of an Insider”, *Stochastic Processes and their Applications* 75, 263–286

J.P. Ansel and C. Stricker (1992), “Lois de Martingale, Densités et Décomposition de Föllmer-Schweizer”, *Annales de l’Institut Henri Poincaré* 28, 375–392

J. P. Ansel and C. Stricker (1993), “Décomposition de Kunita-Watanabe”, *Séminaire de Probabilités XXVII, Lecture Notes in Mathematics 1557, Springer*, 30–32

E. Aurell and S. I. Simdyankin (1998), “Pricing Risky Options Simply”, *International Journal of Theoretical and Applied Finance* 1, 1–23

D. Bertsimas, L. Kogan and A. Lo (1999), “Hedging Derivative Securities and Incomplete Markets: An  $\varepsilon$ -Arbitrage Approach”, *LFE working paper No. 1027-99R, Sloan School of Management, MIT, Cambridge MA*

F. Biagini, P. Guasoni and M. Pratelli (1999), “Mean-Variance Hedging for Stochastic Volatility Models”, *preprint, University of Pisa*

F. Black and M. Scholes (1973), “The Pricing of Options and Corporate Liabilities”, *Journal of Political Economy* 81, 637–654

N. Bouleau and D. Lamberton (1989), “Residual Risks and Hedging Strategies in Markovian Markets”, *Stochastic Processes and their Applications* 33, 131–150

R. Buckdahn (1993), “Backward Stochastic Differential Equations Driven by a Martingale”, *preprint, Humboldt University, Berlin [unpublished]*

A. Černý (1999), “Mean-Variance Hedging in Discrete Time”, *preprint, Imperial College Management School, London*

T. Choulli, L. Krawczyk and C. Stricker (1998), “ $\mathcal{E}$ -Martingales and their Applications

in Mathematical Finance”, *Annals of Probability* 26, 853–876

T. Choulli, L. Krawczyk and C. Stricker (1999), “On Fefferman and Burkholder-Davis-Gundy Inequalities for  $\mathcal{E}$ -Martingales”, *Probability Theory and Related Fields* 113, 571–597

T. Choulli and C. Stricker (1996), “Deux Applications de la Décomposition de Galtchouk-Kunita-Watanabe”, *Séminaire de Probabilités XXX, Lecture Notes in Mathematics 1626*, Springer, 12–23

J. Cvitanić and I. Karatzas (1992), “Convex Duality in Constrained Portfolio Optimization”, *Annals of Applied Probability* 2, 767–818

M. H. A. Davis (1994), “A General Option Pricing Formula”, *preprint, Imperial College, London*

M. H. A. Davis (1997), “Option Pricing in Incomplete Markets”, in: *M. A. H. Dempster and S. R. Pliska (eds.), “Mathematics of Derivative Securities”, Cambridge University Press*, 216–226

F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer and C. Stricker (1997), “Weighted Norm Inequalities and Hedging in Incomplete Markets”, *Finance and Stochastics* 1, 181–227

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 (1996a), “The Variance-Optimal Martingale Measure for Continuous Processes”, *BERNOULLI* 2, 81–105; Amendments and corrections (1996), *BERNOULLI* 2, 379–380

F. Delbaen and W. Schachermayer (1996b), “Attainable Claims with  $p$ 'th Moments”, *Annales de l'Institut Henri Poincaré* 32, 743–763

C. Dellacherie and P. A. Meyer (1982), “Probabilities and Potential B”, *North-Holland*

D. Duffie and H. R. Richardson (1991), “Mean-Variance Hedging in Continuous Time”, *Annals of Applied Probability* 1, 1–15

H. Föllmer and M. Schweizer (1989), “Hedging by Sequential Regression: An Introduction to the Mathematics of Option Trading”, *ASTIN Bulletin* 18, 147–160

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*, 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”, North-Holland, 205–223

C. Gouriéroux, J. P. Laurent and H. Pham (1998), “Mean-Variance Hedging and Numéraire”, *Mathematical Finance* 8, 179–200

P. Grandits and L. Krawczyk (1998), “Closedness of Some Spaces of Stochastic Integrals”, *Séminaire de Probabilités XXXII, Lecture Notes in Mathematics 1686*, Springer, 73–85

B. Grünewald (1998), “Absicherungsstrategien für Optionen bei Kurssprüngen”, *Deutscher UniversitätsVerlag, Wiesbaden*

B. Grünewald and S. Trautmann (1997), “Varianzminimierende Hedgingstrategien für Optionen bei möglichen Kurssprüngen”, in: G. Franke (ed.), “Bewertung und Einsatz von Finanzderivaten”, *Zeitschrift für betriebswirtschaftliche Forschung, Sonderheft 38*, 43–87

P. Guasoni and F. Biagini (1999), “Mean-Variance Hedging with Random Volatility Jumps”, *preprint, University of Pisa*

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

D. Heath, E. Platen and M. Schweizer (1998), “A Comparison of Two Quadratic Approaches to Hedging in Incomplete Markets”, *preprint, Technical University of Berlin*

D. Heath, E. Platen and M. Schweizer (1999), “Numerical Comparison of Local Risk-Minimisation and Mean-Variance Hedging”, *preprint, University of Technology Sydney*

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 1/96, University of Karlsruhe*

C. Hipp (1998), “Hedging General Claims in Diffusion Models”, *preprint 1/98, University of Karlsruhe*

F. Hubalek and L. Krawczyk (1998), “Simple Explicit Formulae for Variance-Optimal Hedging for Processes with Stationary Independent Increments”, *preprint, University of Vienna*

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

E. Jouini and C. Napp (1998), “Continuous Time Equilibrium Pricing of Nonredundant Assets”, *CREST preprint No. 9830, Paris*

- I. Karatzas (1997), “Lectures on the Mathematics of Finance”, *CRM Monograph Series, Vol. 8, American Mathematical Society*
- I. Karatzas and S.-G. Kou (1996), “On the Pricing of Contingent Claims under Constraints”, *Annals of Applied Probability* 6, 321–369
- R. Korn (1997), “Value Preserving Portfolio Strategies in Continuous-Time Models”, *Mathematical Methods of Operations Research* 45, 1–43
- R. Korn (1998), “Value Preserving Portfolio Strategies and the Minimal Martingale Measure”, *Mathematical Methods of Operations Research* 47, 169–179
- D. Lamberton, H. Pham and M. Schweizer (1998), “Local Risk-Minimization under Transaction Costs”, *Mathematics of Operations Research* 23, 585–612
- J. P. Laurent and H. Pham (1999), “Dynamic Programming and Mean-Variance Hedging”, *Finance and Stochastics* 3, 83–110
- D. Lepingale and J. Mémin (1978), “Sur l’Intégrabilité Uniforme des Martingales Exponentielles”, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 42, 175–203
- F. Mercurio (1996), “Mean-Variance Pricing and Risk Preferences”, *Tinbergen Institute discussion paper TI 96-44/2, Erasmus University Rotterdam*
- R. C. Merton (1973), “Theory of Rational Option Pricing”, *Bell Journal of Economics and Management Science* 4, 141–183
- T. Møller (1998a), “Risk-Minimizing Hedging Strategies for Unit-Linked Life Insurance Contracts”, *ASTIN Bulletin* 28, 17–47
- T. Møller (1998b), “Risk-Minimizing Hedging Strategies for Insurance Payment Processes”, *working paper No. 154, University of Copenhagen*
- P. Monat and C. Stricker (1995), “Föllmer-Schweizer Decomposition and Mean-Variance Hedging of General Claims”, *Annals of Probability* 23, 605–628
- H. Pham (1999), “On Quadratic Hedging in Continuous Time”, *preprint, Université de Marne-la-Vallée, Paris*
- H. Pham, T. Rheinländer and M. Schweizer (1998), “Mean-Variance Hedging for Continuous Processes: New Results and Examples”, *Finance and Stochastics* 2, 173–198
- H. Pham and N. Touzi (1996), “Equilibrium State Prices in a Stochastic Volatility Model”, *Mathematical Finance* 6, 215–236
- T. Rheinländer (1999), “Optimal Martingale Measures and their Applications in Math-

ematical Finance”, *PhD thesis, Technical University of Berlin*

T. Rheinländer and M. Schweizer (1997), “On  $L^2$ -Projections on a Space of Stochastic Integrals”, *Annals of Probability* 25, 1810–1831

M. Schweizer (1988), “Hedging of Options in a General Semimartingale Model”, *Diss. ETH Zürich* 8615

M. Schweizer (1990), “Risk-Minimality and Orthogonality of Martingales”, *Stochastics and Stochastics Reports* 30, 123–131

M. Schweizer (1991), “Option Hedging for Semimartingales”, *Stochastic Processes and their Applications* 37, 339–363

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

M. Schweizer (1994b), “Risk-Minimizing Hedging Strategies under Restricted Information”, *Mathematical Finance* 4, 327–342

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

M. Schweizer (1995b), “Variance-Optimal Hedging in Discrete Time”, *Mathematics of Operations Research* 20, 1–32

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

M. Schweizer (1998), “From Actuarial to Financial Valuation Principles”, *preprint, Technical University of Berlin*

M. Schweizer (1999a), “A Minimality Property of the Minimal Martingale Measure”, *Statistics and Probability Letters* 42, 27–31

M. Schweizer (1999b), “Risky Options Simplified”, *International Journal of Theoretical and Applied Finance* 2, 59–82

C. Stricker (1990), “Arbitrage et Lois de Martingale”, *Annales de l’Institut Henri Poincaré* 26, 451–460

C. Stricker (1996), “The Föllmer-Schweizer Decomposition”, in: *H.-J. Engelbert, H. Föllmer and J. Zabczyk (eds.), “Stochastic Processes and Related Topics”, Stochastics Monographs, Vol. 10, Gordon and Breach, 77–89*

A. Wiese (1998), “Hedging stochastischer Verpflichtungen in zeitstetigen Modellen”, *Ver-*

*lag Versicherungswissenschaft, Karlsruhe*

M. Yor (1978), “Sous-Espaces Denses dans  $L^1$  ou  $H^1$  et Représentation des Martingales”,  
*Séminaire de Probabilités XII, Lecture Notes in Mathematics 649, Springer, 265–309*