A Comparison of Two Quadratic Approaches to Hedging in Incomplete Markets

David Heath, Eckhard Platen
School of Finance and Economics and
School of Mathematical Sciences
University of Technology Sydney
City Campus Broadway
GPO Box 123
Broadway, NSW 2007
Australia

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

Abstract: This paper provides comparative theoretical and numerical results on risks, values and hedging strategies for local risk-minimization versus mean-variance hedging in a class of stochastic volatility models. We explain the theory for both hedging approaches in a general framework, specialize to a Markovian situation and analyze in detail variants of the well-known Heston (1993) and Stein/Stein (1991) stochastic volatility models. Numerical results are obtained mainly by PDE and simulation methods. In addition, we take special care to check that all our examples do satisfy the conditions required by the general theory.

This version: February 27, 2001

* corresponding author

0. Introduction

The valuation and hedging of derivatives in incomplete financial markets is a frequently studied problem in mathematical finance. Several different approaches have been developed in the literature, but no agreement on one uniformly superior method has emerged so far. The present paper contributes to this discussion by comparing in some case studies the two main competing *quadratic* hedging approaches: local risk-minimization and mean-variance hedging.

In a nutshell, the main difference between these two approaches is the following: one has either simple solutions for hedging strategies (local risk-minimization) or a control over total costs and risks (mean-variance hedging), but not both. We therefore compare the total risks for the two methods in order to understand better how much is lost by using the simpler solution. More precisely, we provide comparative theoretical and numerical results on risks, option values and hedging strategies in a class of stochastic volatility models. We offer improvements over the few existing studies in this area in several directions. The papers by Hipp (1996, 1998) compare in a diffusion framework the mean-variance optimal strategy for hedging a constant to an alternative (called "locally optimal") strategy. However, the latter is not the locally risk-minimizing strategy and risks and values can be computed explicitly in Hipp's examples. Hence these papers need no numerical computations except for the simulation of a few trajectories. Grünewald/Trautmann (1997) study the same basic questions as we do. They work in a jump-diffusion model and their numerical results are based on Monte Carlo simulations. In contrast, we derive and numerically solve partial differential equations for values and risks in diffusion models, thus obtaining higher accuracy than with Monte Carlo methods. In a very recent paper, Bertsimas/Kogan/Lo (1999) use PDE techniques and dynamic programming arguments to obtain mean-variance optimal strategies in a number of situations. While their derivation of the Hamilton-Jacobi-Bellman equation seems to be based on purely algebraic manipulations, we make a deliberate effort to prove that our examples do satisfy the conditions required by the theory.

Both approaches under comparison here are based on a quadratic criterion and as such of course subject to possible criticism. The main objections raised in this connection are inconsistency with increasing utility of wealth and the fact that profits and losses are both equally punished. Thus anyone who prefers more to less or thinks of an asymmetric positively skewed profit and loss distribution is likely to be dissatisfied with our approaches. In other words, economic intuition does not fit together well with quadratic methods.

Having said that, we can also point out some advantages. One is tractability: quadratic criteria typically lead to fairly explicit solutions even in very general models. Moreover, our quadratic approaches yield arbitrage-free valuations and in the case of a complete market reproduce the usual unique arbitrage-free prices and riskless hedging strategies. Hence they

can be viewed as one possible consistent extension from the complete to the incomplete market case.

The paper is structured as follows. We first explain in section 1 the general theoretical background for the two approaches in a uniform framework. Section 2 specializes these results to a Markovian situation and more specifically to a general Markovian model for an asset with stochastic volatility. As examples, we study variants of the well-known Heston (1993) and Stein/Stein (1991) stochastic volatility models which are presented in detail in section 3. In particular, we provide there the justifications for applying the general theory to these specific examples. Section 4 then contains our numerical results. Their main conclusion is that mean-variance hedging can be applied in practice with a tolerable amount of extra computational work (in our examples, about 20%) in comparison to local risk-minimization. However, this only holds in situations where the theory for mean-variance hedging is already sufficiently developed, and these are not yet numerous. All proofs are collected in the appendix.

1. General theory

This section lays out the general framework, explains the basic problem under consideration and presents the two hedging methods we want to compare. We start with a probability space $(\Omega, \mathcal{F}, P)$, a fixed time horizon $T \in (0, \infty)$ and a filtration $I\!\!F = (\mathcal{F}_t)_{0 \le t \le T}$ satisfying the usual conditions of right-continuity and completeness. For all unexplained terminology from martingale theory, we refer to Dellacherie/Meyer (1982). We consider a frictionless market in continuous time with $d+1$ primary assets available for trade; their price processes are $S^i = (S^i_t)_{0 \le t \le T}$ for $i = 0, 1, \ldots, d$. We assume that S^0 has a strictly positive price, use S^0 as numeraire and for ease of notation immediately pass to quantities discounted with S^0. Thus asset 0 will have (discounted) price 1 at all times and the other assets' (discounted) prices are $X^i = \frac{S^i}{S^0}$ for $i = 1, \ldots, d$. Without further mention, all subsequently appearing prices and values will be expressed in discounted units.

Mathematically, the basic asset prices are described by the constant process 1 and an $I\!\!R^d$-valued stochastic process $X = (X_t)_{0 \le t \le T}$ adapted to the filtration $I\!\!F$. To exclude arbitrage opportunities, we assume that X admits an *equivalent local martingale measure (ELMM) Q*, i.e., there exists a probability $Q \approx P$ such that X is a local Q-martingale. If $I\!\!P$ denotes the convex set of all ELMMs Q, we thus assume $I\!\!P \ne \emptyset$. Incompleteness of the market given by $(1, X)$ and $I\!\!F$ is taken to mean that $I\!\!P$ contains more than one element (and therefore infinitely many). The basic problem is then to deal with a general European-type *contingent claim*; this is a random payoff at time T, formally described by an $\mathcal{F}_T$-measurable random variable H. Before elaborating on possible ways of "dealing with" H, we illustrate the preceding concepts by an example.

Example. All our subsequent numerical work is done in the context of a stochastic volatility model with a single risky asset ($d = 1$). If Y denotes the volatility of this asset X, we consider a stochastic differential equation (SDE) of the form

(1.1)
$$\frac{dX_t}{X_t} = \mu(t, X_t, Y_t)\, dt + Y_t\, dW_t,$$

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

with suitable functions μ, a, b and P-Brownian motions W, W' with instantaneous correlation $\varrho(t, X_t, Y_t)$, i.e., $d\langle W, W'\rangle_t = \varrho(t, X_t, Y_t)\, dt$. These are constructed from independent P-Brownian motions W^1, W^2 on $(\Omega, \mathcal{F}, P)$ by setting $W := W^1$ and $W' := \int \varrho(u, X_u, Y_u)\, dW_u^1 + \int \sqrt{1 - \varrho^2(u, X_u, Y_u)}\, dW_u^2$. The filtration $\mathbb{F}$ is generated by W^1, W^2, made complete and right-continuous. We thus have $d = 1$ asset X; the additional factor Y models the volatility of X and induces some additional randomness, but is not available as a traded asset. A typical example of a contingent claim is a European call option on X with strike K and maturity T; its (net) payoff at time T is $H = (X_T - K)^+$. But $\mathcal{F}_T$-measurability also allows very general path-dependent payoffs and H could even depend explicitly on the evolution of the volatility process Y.

Mild assumptions on μ, a, b guarantee that X admits an ELMM Q; it is enough to remove the drift μ by a Girsanov transformation. The stipulation that X should be a local Q-martingale uniquely determines that transformation's effect on W, but imposes no restrictions on how W^2 is affected. Hence there is clearly no unique ELMM so that we have an incomplete market. We remark that this rests of course on our assumption that the only dynamically traded asset is the underlying X; standard options may be available, but not yet traded liquidly. This ends the present discussion of this example.

In "dealing with" a contingent claim H, one may want to do two things: *valuation* (i.e., assign a value to H at times $t < T$) and *hedging* (i.e., cover oneself against potential losses arising from a sale of H) which brings up the notion of a strategy via trading in X. Since the assumption $\mathbb{P} \neq \emptyset$ implies that X is a semimartingale under P, one can use stochastic integrals with respect to X and we denote by $L(X)$ the space of $\mathbb{R}^d$-valued predictable X-integrable processes ϑ; see Dellacherie/Meyer (1982) for details. For $\vartheta \in L(X)$, the stochastic integral $\int \vartheta\, dX$ is well-defined, but elements of $L(X)$ are too general to yield economically reasonable strategies; we shall presently impose some additional integrability conditions.

Definition. A *pre-strategy* is any pair $\varphi = (\vartheta, \eta)$, where $\vartheta \in L(X)$ and $\eta = (\eta_t)_{0 \leq t \leq T}$ is a real-valued adapted process such that the *value process* $V_t(\varphi) := \vartheta_t^{\mathrm{tr}} X_t + \eta_t$, $0 \leq t \leq T$, is right-continuous. The (cumulative) *cost process* $C(\varphi)$ is then defined by

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

Intuitively, a pre-strategy should be viewed as a portfolio varying dynamically over time. The controllable quantities ϑ_t^i and η_t denote the respective numbers of shares of assets i and 0 held at time t so that $V_t(\varphi)$ is obviously the value of the portfolio one owns at time t. $C_t(\varphi)$ describes the total costs incurred by φ over the interval $[0, t]$; note that these costs arise from trading because of the fluctuations in the price process X and are not due to transaction costs. By allowing ϑ and η to vary independently, we also admit portfolios that are not self-financing and thus may generate profits or losses over time.

Definition. A pre-strategy φ is called *self-financing* if its cost process is P-a.s. constant and *mean-self-financing* if $C(\varphi)$ is a martingale (under P). If $C(\varphi)$ is square-integrable, the *risk process* of φ is defined by

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

Since H is $\mathcal{F}_T$-measurable and η is adapted, there always exist pre-strategies with $V_T(\varphi) = H$; a simple example is "pay at the end", i.e., $\vartheta \equiv 0$ and $\eta_t = H I_{\{t=T\}}$. The goal in both approaches here is to find a suitably integrable pre-strategy φ with $V_T(\varphi) = H$ which "minimizes risk" in a sense to be made precise. We emphasize that we are therefore looking at *hedging approaches* designed to control the riskiness of a pre-strategy as measured by its cost fluctuations. If there is an optimal strategy φ^*, we can use $V_t(\varphi^*)$ as a value of H at time t, but this valuation is a by-product and not our primary objective.

Clearly, the absolute minimum of $R(\varphi)$ is the zero process. This can be attained if and only if there is a self-financing pre-strategy φ with $V_T(\varphi) = H$. Hence contingent claims with zero risk are exactly those which are *attainable* in the sense that they can be replicated by a self-financing pre-strategy. The interesting questions thus arise for non-attainable claims, i.e., in an incomplete market.

1.1. Local risk-minimization

In the first approach we consider, one insists on the requirement $V_T(\varphi) = H$ P-a.s. If H is not attainable, this forces one to go away from self-financing pre-strategies and so the objective becomes to minimize the risk process $R(\varphi)$ in a suitable way. For the case where X is a local P-martingale, this criterion has been defined and developed by Föllmer/Sondermann (1986) under the name of *risk-minimization*. Existence and uniqueness of risk-minimizing hedging strategies are proved there as an application of the well-known Galtchouk-Kunita-Watanabe decomposition theorem. The generalization to the semimartingale case is due to Schweizer (1988, 1991) who called the resulting concept *local risk-minimization*. Its basic idea is to control hedging errors at each instant by minimizing the conditional variances of

instantaneous cost increments sequentially over time. This involves (local) variances and thus requires more specific assumptions on X.

More precisely, we assume that X can be decomposed as $X = X_0 + M + A$ where $M \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ is an $\mathbb{R}^d$-valued locally square-integrable local P-martingale null at 0 and A is an $\mathbb{R}^d$-valued adapted *continuous* process of finite variation also null at 0. We denote by $\langle M \rangle = \left(\langle M \rangle^{ij} \right)_{i,j=1,\ldots,d} = \left(\langle M^i, M^j \rangle \right)_{i,j=1,\ldots,d}$ the matrix-valued covariance process of M and we suppose that A is absolutely continuous with respect to $\langle M \rangle$ in the sense that

$$A^i_t = \left(\int\limits_0^t d\langle M \rangle_s \, \widehat{\lambda}_s \right)^i := \sum_{j=1}^d \int\limits_0^t \widehat{\lambda}^j_s \, d\langle M^i, M^j \rangle_s \qquad , \qquad 0 \le t \le T, i = 1, \ldots, d$$

for some $\mathbb{R}^d$-valued predictable process $\widehat{\lambda}$ such that

$$\widehat{K}_t := \int\limits_0^t \widehat{\lambda}^{\mathrm{tr}}_s \, d\langle M \rangle_s \, \widehat{\lambda}_s = \sum_{i,j=1}^d \int\limits_0^t \widehat{\lambda}^i_s \widehat{\lambda}^j_s \, d\langle M^i, M^j \rangle_s < \infty \qquad P\text{-a.s. for each } t \in [0,T].$$

This complex of conditions on X is sometimes called the structure condition (SC). Since $\mathbb{P} \ne \emptyset$, it is automatically satisfied if X is continuous; see Theorem 1 of Schweizer (1995). Note also that the stochastic integral $\int \widehat{\lambda} \, dM$ is then well-defined and that its variance process is $\left\langle \int \widehat{\lambda} \, dM \right\rangle = \widehat{K}$; this will be used later on.

Definition. Θ 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, ϑ has to be predictable with

$$E\left[\int\limits_0^T \vartheta^{\mathrm{tr}}_s \, d[M]_s \, \vartheta_s + \left(\int\limits_0^T |\vartheta^{\mathrm{tr}}_s \, dA_s| \right)^2 \right] < \infty.$$

(This does not use (SC); the last formulation only requires X to be a special semimartingale.) An L^2-*strategy* is any pre-strategy $\varphi = (\vartheta, \eta)$ with $\vartheta \in \Theta$ and such that $V(\varphi)$ is square-integrable, i.e., $V_t(\varphi) \in L^2(P)$ for each $t \in [0,T]$. An L^2-strategy φ is called a *pseudo-locally risk-minimizing strategy* if its cost process $C(\varphi)$ is a square-integrable P-martingale and strongly P-orthogonal to M, the P-martingale part of X.

We have avoided the original definition of locally risk-minimizing strategies from Schweizer (1991) because this is rather delicate in continuous time. Under mild technical assumptions on X, the notions "locally risk-minimizing" and "pseudo-locally risk-minimizing" actually coincide; this is proved (for $d = 1$) in Proposition 2.3 of Schweizer (1991). Moreover, Proposition (2.24) of Föllmer/Schweizer (1991) shows that finding a pseudo-locally

risk-minimizing strategy for a given contingent claim $H \in L^2(P)$ is equivalent to finding a decomposition of H as

$$
(1.2) \qquad H = H_0 + \int_0^T \xi_u^H \, dX_u + L_T^H
$$

with $H_0 \in L^2(\mathcal{F}_0, P)$, $\xi^H \in \Theta$ and a square-integrable P-martingale L^H null at 0 and strongly P-orthogonal to M. Once we have (1.2), the desired strategy φ^{lr} is given by

$$
(1.3) \qquad \vartheta_t^{\mathrm{lr}} = \xi_t^H \qquad , \qquad 0 \le t \le T
$$

and

$$
(1.4) \qquad C_t(\varphi^{\mathrm{lr}}) = H_0 + L_t^H \qquad , \qquad 0 \le t \le T;
$$

η^{lr} is then determined via the value process

$$
V_t(\varphi^{\mathrm{lr}}) = C_t(\varphi^{\mathrm{lr}}) + \int_0^t \vartheta_u^{\mathrm{lr}} \, dX_u = H_0 + \int_0^t \xi_u^H \, dX_u + L_t^H \qquad , \qquad 0 \le t \le T.
$$

In the literature, (1.2) has been called the *Föllmer-Schweizer decomposition* of H and has been studied by several authors; see for instance Monat/Stricker (1995).

In the case where X is *continuous*, there is a well-known approach to constructing a candidate for the decomposition (1.2) with the help of the so-called minimal ELMM $\widehat{P}$. We first define the strictly positive continuous local P-martingale

$$
(1.5) \qquad \widehat{Z}_t := \mathcal{E}\left(-\int \widehat{\lambda} \, dM\right)_t = \exp\left(-\int_0^t \widehat{\lambda}_u \, dM_u - \frac{1}{2}\widehat{K}_t\right) \quad , \quad 0 \le t \le T.
$$

If $\widehat{Z}$ is a square-integrable P-martingale, then

$$
\frac{d\widehat{P}}{dP} := \widehat{Z}_T \in L^2(P)
$$

defines a probability measure $\widehat{P} \approx P$ which is in $I\!\!P$ since one easily verifies that $\widehat{Z}X$ is a local P-martingale. For reasons explained in Föllmer/Schweizer (1991) and Schweizer (1995), $\widehat{P}$ is called the *minimal ELMM*. Note that $H \in L^1(\widehat{P})$ because H and $\widehat{Z}_T$ are both in $L^2(P)$ and so the $\widehat{P}$-martingale

$$
V_t^{H,\widehat{P}} := \widehat{E}[H|\mathcal{F}_t] \qquad , \qquad 0 \le t \le T
$$

is well-defined. Because the local $\widehat{P}$-martingale X is continuous, $V^{H,\widehat{P}}$ admits a *Galtchouk-Kunita-Watanabe decomposition* under $\widehat{P}$ with respect to X as

$$(1.6) \qquad V_t^{H,\widehat{P}} = V_0^{H,\widehat{P}} + \int_0^t \xi_u^{H,\widehat{P}} \, dX_u + L_t^{H,\widehat{P}} \qquad , \qquad 0 \leq t \leq T$$

where $\xi^{H,\widehat{P}}$ is in $L(X)$ and $L^{H,\widehat{P}}$ is a local $\widehat{P}$-martingale null at 0 and strongly $\widehat{P}$-orthogonal to X. Thanks to the continuity of X, $L^{H,\widehat{P}}$ is then also a local P-martingale and strongly P-orthogonal to M; see Theorem 9 of Schweizer (1995).

Since $V_T^{H,\widehat{P}} = H$, (1.6) yields in particular a decomposition of H which looks very similar to (1.2); the only difference is in fact that we have better integrability properties in (1.2). If we already know that H admits a decomposition (1.2), we can argue as in Theorem (3.14) of Föllmer/Schweizer (1991) to conclude that $H_0 = V_0^{H,\widehat{P}}$, $\xi^H = \xi^{H,\widehat{P}}$ and $L^H = L^{H,\widehat{P}}$. Thus (1.6) provides with (1.3) and (1.4) in this case a convenient way to *identify* the strategy φ^{lr}. If we have no existence result guaranteeing (1.2), we can still use (1.6) as a starting point and then try to prove that $V_0^{H,\widehat{P}} \in L^2(P)$, that $\xi^{H,\widehat{P}} \in \Theta$ and that $L^{H,\widehat{P}}$ is a square-integrable P-martingale. This would then constructively imply that H has a decomposition (1.2) and therefore admits a pseudo-locally risk-minimizing strategy given by (1.3) and (1.4). Note that whenever we can get (1.2) from (1.6), the value process $V(\varphi^{\mathrm{lr}})$ coincides with $V^{H,\widehat{P}}$. Finding the decomposition (1.6) in a Markovian situation will be discussed in section 2.

1.2. Mean-variance hedging

This subsection explains the second of the two approaches compared here. While local risk-minimization is concerned with very short-term optimality properties, mean-variance hedging minimizes the global risk over a long term. Returning to our basic model, we now assume in addition to $I\!P \neq \emptyset$ that X is *continuous*. By Theorem 1 of Schweizer (1995), X then satisfies (SC). We denote by

$$I\!P_e^2 := \left\{ Q \in I\!P \; \middle| \; \frac{dQ}{dP} \in L^2(P) \right\} \subseteq I\!P$$

the set of all ELMMs with square-integrable density and assume that $I\!P_e^2 \neq \emptyset$ and also that $\mathcal{F}_0$ is trivial. Our formulation of mean-variance hedging follows Gouriéroux/Laurent/Pham (1998) rather than Rheinländer/Schweizer (1997) because the former framework turns out to be more flexible for applications.

Definition. Θ' denotes the space of all processes $\vartheta \in L(X)$ for which $\int_0^T \vartheta_u \, dX_u$ is in $L^2(P)$

7

and the stochastic integral process $\int \vartheta \, dX$ is a Q-martingale for each $Q \in I\!\!P_e^2$. An *MV-strategy* is any pair (V_0, ϑ) such that $\vartheta \in \Theta'$ and $V_0 \in I\!\!R$. If $H \in L^2(P)$ is a contingent claim, an MV-strategy $(V_0^{\mathrm{mvo}}, \vartheta^{\mathrm{mvo}})$ is called a *mean-variance optimal strategy* for H if it minimizes

$$
E\left[\left(H - V_T(V_0, \vartheta)\right)^2\right] = \left\| H - V_0 - \int_0^T \vartheta_u \, dX_u \right\|_{L^2(P)}^2
$$

over all MV-strategies (V_0, ϑ).

To give an interpretation for this criterion, we first note that each MV-strategy (V_0, ϑ) can be identified with a self-financing pre-strategy $\varphi = (\vartheta, \eta)$ by setting $\eta := V_0 + \int \vartheta \, dX - \vartheta^{\mathrm{tr}} X$. The difference $H - V_T(V_0, \vartheta)$ is then the net loss at time T from paying out the claim H after having traded according to (V_0, ϑ) and mean-variance hedging simply minimizes the expected net squared loss. Alternatively, we can recognize $H - V_0 - \int_0^T \vartheta_u \, dX_u$ as the cost on $(0, T]$ of a pre-strategy with $V_T(\varphi) = H$, initial capital V_0 and stock component ϑ. With this interpretation, we minimize the risk at time 0 only instead of the entire risk process as in the previous subsection. Since R_0 depends only on V_0 and ϑ, it is unnecessary to minimize over the entire pair $\varphi = (\vartheta, \eta)$.

Remark. In purely mathematical terms, mean-variance hedging amounts to projecting the random variable H on the linear space spanned by constants and stochastic integrals of X. In the special case where X is a local P-martingale so that the original measure P is in $I\!\!P$, this problem is solved by the well-known Galtchouk-Kunita-Watanabe decomposition theorem. Moreover, ϑ^{mvo} turns out to coincide with ϑ^{lr} in the martingale case, but this is not necessarily true for a general semimartingale X. Very briefly, the intuition for this result is that local and global projections behave differently if X has a global drift; see Schweizer (1999) for more details on this issue.

To describe the mean-variance optimal strategy, we need some more notation.

Definition. The *variance-optimal ELMM* $\widetilde{P}$ is the unique element of $I\!\!P_e^2$ that minimizes $\left\| \frac{dQ}{dP} \right\|_{L^2(P)} = \sqrt{1 + \mathrm{Var}_P\left[\frac{dQ}{dP}\right]}$ over all $Q \in I\!\!P_e^2$.

Actually, the existence of $\widetilde{P}$ for continuous processes is a nontrivial result due to Delbaen/Schachermayer (1996) and slightly generalized by Gouriéroux/Laurent/Pham (1998).

Moreover, both these papers also show that

$$\widetilde{Z}_t := \widetilde{E}\left[\frac{d\widetilde{P}}{dP}\bigg|\mathcal{F}_t\right] = \widetilde{Z}_0 + \int_0^t \widetilde{\zeta}_u \, dX_u \qquad , \qquad 0 \le t \le T$$

for some $\widetilde{\zeta} \in \Theta'$; in particular, $\widetilde{Z}$ is like X continuous. As the next result shows, $\widetilde{P}$, $\widetilde{Z}$ and $\widetilde{\zeta}$ all turn up in the solution of the mean-variance hedging problem.

Theorem 1. *Let* $H \in L^2(P)$ *be a contingent claim and write the Galtchouk-Kunita-Watanabe decomposition of* H *under* $\widetilde{P}$ *with respect to* X *as*

$$(1.7) \qquad H = \widetilde{E}[H] + \int_0^T \xi_u^{H,\widetilde{P}} \, dX_u + L_T^{H,\widetilde{P}} = V_T^{H,\widetilde{P}}$$

with

$$V_t^{H,\widetilde{P}} := \widetilde{E}[H|\mathcal{F}_t] = \widetilde{E}[H] + \int_0^t \xi_u^{H,\widetilde{P}} \, dX_u + L_t^{H,\widetilde{P}} \qquad , \qquad 0 \le t \le T.$$

Then the mean-variance optimal strategy for H *is given by*

$$(1.8) \qquad V_0^{\mathrm{mvo}} = \widetilde{E}[H]$$

and

$$(1.9) \qquad \vartheta_t^{\mathrm{mvo}} = \xi_t^{H,\widetilde{P}} - \frac{\widetilde{\zeta}_t}{\widetilde{Z}_t}\left(V_{t-}^{H,\widetilde{P}} - \widetilde{E}[H] - \int_0^t \vartheta_u^{\mathrm{mvo}} \, dX_u\right)$$

$$= \xi_t^{H,\widetilde{P}} - \widetilde{\zeta}_t \int_0^{t-} \frac{1}{\widetilde{Z}_u} \, dL_u^{H,\widetilde{P}} \qquad , \qquad 0 \le t \le T.$$

Proof. See appendix.

Proposition 2. *Let* $H \in L^2(P)$ *be a contingent claim and* $(V_0^{\mathrm{mvo}}, \vartheta^{\mathrm{mvo}})$ *the associated mean-variance-optimal strategy. The process* $U := V^{H,\widetilde{P}} - V(V_0^{\mathrm{mvo}}, \vartheta^{\mathrm{mvo}})$ *then satisfies the stochastic differential equation*

$$(1.10) \qquad dU_t = U_{t-}\frac{\widetilde{\zeta}_t}{\widetilde{Z}_t} \, dX_t + dL_t^{H,\widetilde{P}} \qquad , \qquad U_0 = 0$$

9

and is explicitly given by

$$(1.11) \qquad U_t = \widetilde{Z}_t \int_0^t \frac{1}{\widetilde{Z}_u} \, dL_u^{H,\widetilde{P}} \qquad , \qquad 0 \le t \le T.$$

Moreover, the minimal total risk of H is given by

$$(1.12) \qquad R_0^* := E\left[\left(H - V_T(V_0^{\mathrm{mvo}}, \vartheta^{\mathrm{mvo}}) \right)^2 \right] = E\left[\int_0^T \frac{Z_u^{\widetilde{P}}}{\widetilde{Z}_u} \, d\big[L^{H,\widetilde{P}}\big]_u \right]$$

where $Z_t^{\widetilde{P}} := E\left[\frac{d\widetilde{P}}{dP} \Big| \mathcal{F}_t \right], 0 \le t \le T$, is the density process of $\widetilde{P}$ with respect to P. (Note the difference between $\widetilde{Z}$ and $Z^{\widetilde{P}}$.)

Proof. See appendix.

To make the preceding results practically applicable, we still need two more ingredients. Since the Galtchouk-Kunita-Watanabe decomposition (1.7) of H under $\widetilde{P}$ will be dealt with in section 2, we focus here on finding the processes $\widetilde{\zeta}$, $\widetilde{Z}$ and $Z^{\widetilde{P}}$. This involves a more detailed study of the structure of $\widetilde{P}$ and we first recall a rather special case solved by Pham/ Rheinländer/Schweizer (1998).

Lemma 3. *If $\widehat{K}_T$ is deterministic, then $\widetilde{P} = \widehat{P}$,*

$$Z_t^{\widetilde{P}} = Z_t^{\widehat{P}} = \widehat{Z}_t = \mathcal{E}\left(-\int \widehat{\lambda} \, dM \right)_t \qquad , \qquad 0 \le t \le T,$$

$$\widetilde{Z}_t = \widehat{E}\left[\frac{d\widehat{P}}{dP} \Big| \mathcal{F}_t \right] = e^{\widehat{K}_T} \mathcal{E}\left(-\int \widehat{\lambda} \, dX \right)_t \qquad , \qquad 0 \le t \le T,$$

$$\widetilde{\zeta}_t = -e^{\widehat{K}_T} \mathcal{E}\left(-\int \widehat{\lambda} \, dX \right)_t \widehat{\lambda}_t = -\widetilde{Z}_t \widehat{\lambda}_t \qquad , \qquad 0 \le t \le T$$

and

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

Proof. See appendix.

While Lemma 3 is a pleasingly simple result, its assumption is usually too restrictive. In a multidimensional diffusion model, Laurent/Pham (1999) have shown how to obtain $Z^{\widetilde{P}}$, hence

$\widetilde{P}$, and the ratio $\frac{\widetilde{\zeta}}{\widetilde{Z}}$ in terms of the canonical decomposition of a certain semimartingale j. This j is the solution of a stochastic control problem and in general hard to obtain explicitly. But following Laurent/Pham (1999), we shall see in the next section how j, hence also $\widetilde{\zeta}$, $Z^{\widetilde{P}}$ and $\widetilde{Z}$, can be determined in a Markovian framework under some additional assumptions.

1.3. Comparing the two approaches

It should be clear from the preceding definitions that both mean-variance hedging and local risk-minimization aim essentially at minimizing expected squared hedging costs. The only difference is that mean-variance hedging does this over one global time step whereas local risk-minimization applies the quadratic criterion "on each infinitesimal interval". While this leads to a more involved definition, the resulting solution is actually simpler than for mean-variance hedging because it only uses the local structure of X. In particular, finding and implementing a locally risk-minimizing strategy is almost straightforward. On the other hand, such a strategy is only locally optimal and gives no control over the total errors in a given time interval. Such a control can be achieved by using mean-variance hedging, but at the expense of more complicated strategies and (at least at present) of more restrictive settings for examples; the theory is not yet far enough.

Given this situation, we ask here the question: How much does one gain from using the more complicated approach of mean-variance hedging? We therefore compare the expected squared total costs

$$E\left[\left(H - V_0^{\mathrm{mvo}} - \int\limits_0^T \vartheta_u^{\mathrm{mvo}}\, dX_u\right)^2\right] = R_0^*$$

and

$$E\left[\left(H - V_0(\varphi^{\mathrm{lr}}) - \int\limits_0^T \vartheta_u^{\mathrm{lr}}\, dX_u\right)^2\right] = E\left[\left(C_T(\varphi^{\mathrm{lr}}) - C_0(\varphi^{\mathrm{lr}})\right)^2\right] = \mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right],$$

using that $C(\varphi^{\mathrm{lr}})$ is a martingale. By definition, it is clear that mean-variance hedging then dominates local risk-minimization in the sense that $R_0^* \le \mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right]$. On the other hand, we can give a reverse estimate under the assumptions of Lemma 3 because $\widetilde{P} = \widehat{P}$ then implies that $L^{H,\widetilde{P}} = L^{H,\widehat{P}}$. Since we also have $\frac{Z_t^{\widetilde{P}}}{\widetilde{Z}_t} = e^{-(\widehat{K}_T - \widehat{K}_t)}$, (1.12), (1.4) and (1.6) yield

$$R_0^* = E\left[\int\limits_0^T e^{-(\widehat{K}_T - \widehat{K}_u)}\, d\left[L^{H,\widetilde{P}}\right]_u\right] \ge e^{-\widehat{K}_T} E\left[\left(L_T^{H,\widehat{P}}\right)^2\right] = e^{-\widehat{K}_T}\, \mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right].$$

Because this estimate is rather crude, we prefer to compute R_0^* and $\text{Var}\left[C_T(\varphi^{\text{lr}})\right]$ more explicitly and compare their values numerically in examples.

2. A Markovian framework

In this section, we specialize the preceding results to a Markovian situation. For concreteness and with a view to the numerical examples below, we do this in a stochastic volatility model, but most of the results extend to more general cases. Let X and Y be given by (1.1). A look at section 1 shows that both hedging approaches outlined there involve two steps:

a) Determine an appropriate ELMM Q and find the dynamics of (X,Y) under Q. More precisely, we need $Q = \widehat{P}$ for local risk-minimization and $Q = \widetilde{P}$ for mean-variance hedging. In addition, we may need some processes associated to Q.

b) Find the Galtchouk-Kunita-Watanabe decomposition of H with respect to X under Q.

The optimal strategies are then given by (1.3), (1.4) and (1.8), (1.9) respectively. Once a) and b) are completed, we thus know all we need for solving our two hedging problems.

Because b) is conceptually the same in both cases, we first explain how to do this. Suppose that step a) has already been taken and that (X,Y) is a Markov process under the ELMM Q; this will be the case in all our subsequent examples. Thus we can write

$$\frac{dX_t}{X_t} = Y_t\, dB_t,$$

$$dY_t = a^Q(t, X_t, Y_t)\, dt + b(t, X_t, Y_t)\, dB'_t$$

for Q-Brownian motions B, B' with instantaneous correlation $\varrho(t, X_t, Y_t)$; these are driven by independent Q-Brownian motions B^1, B^2. In comparison to (1.1), the drift μ has disappeared from X (because Q is an ELMM) and the drift in Y has changed from a to a^Q. Neither the correlation ϱ nor the volatility b of Y is affected by the change of measure from P to Q.

For a contingent claim of the form $H = h(X_T, Y_T)$ with a function $h(x,y)$ on $[0,\infty) \times \mathbb{R}$, finding the Galtchouk-Kunita-Watanabe decomposition of H under Q reduces to a PDE problem if one exploits the Markovian structure. This is already explained in Pham/Rheinländer/ Schweizer (1998) and so we just outline the argument here. Under regularity assumptions on a^Q, b, ϱ, h, the Markov property implies that

$$V_t^{H,Q} := E_Q[H|\mathcal{F}_t] = E_Q[h(X_T, Y_T)|\mathcal{F}_t] = v^Q(t, X_t, Y_t)$$

for a function $v^Q(t,x,y)$ on $[0,T] \times [0,\infty) \times \mathbb{R}$ satisfying the partial differential equation

$$(2.1) \qquad \frac{\partial v^Q}{\partial t} + a^Q \frac{\partial v^Q}{\partial y} + \frac{1}{2}\left(x^2 y^2 \frac{\partial^2 v^Q}{\partial x^2} + b^2 \frac{\partial^2 v^Q}{\partial y^2} + 2xyb\varrho \frac{\partial^2 v^Q}{\partial x \partial y}\right) = 0$$

$$\text{on } (0,T) \times (0,\infty) \times \mathbb{R}$$

with boundary condition

$$(2.2) \qquad v^Q(T, x, y) = h(x, y) \qquad \text{on } [0, \infty) \times I\!\!R.$$

Applying Itô's formula to $V^{H,Q}$ and comparing to the Galtchouk-Kunita-Watanabe decomposition

$$V_t^{H,Q} = V_0^{H,Q} + \int\limits_0^t \xi_u^{H,Q} \, dX_u + L_t^{H,Q} \qquad , \qquad 0 \le t \le T$$

under Q then shows that $\xi^{H,Q}$ and $L^{H,Q}$ are given by

$$(2.3) \qquad \xi_t^{H,Q} = \frac{\partial v^Q}{\partial x}(t, X_t, Y_t) + \frac{1}{X_t Y_t}\left(b\varrho \frac{\partial v^Q}{\partial y}\right)(t, X_t, Y_t) \qquad , \qquad 0 \le t \le T$$

and

$$(2.4) \qquad L_t^{H,Q} = \int\limits_0^t \left(b\sqrt{1 - \varrho^2}\frac{\partial v^Q}{\partial y}\right)(u, X_u, Y_u) \, dB_u^2 \qquad , \qquad 0 \le t \le T.$$

Apart from integrability questions, this completes step b). We now turn to step a) from above which has to be done separately for the two approaches.

2.1. Local risk-minimization: Finding $\widehat{P}$ and the corresponding dynamics

According to section 1, the density process of the minimal ELMM $\widehat{P}$ with respect to P is given by $\widehat{Z} = \mathcal{E}\left(-\int \widehat{\lambda} \, dM\right)$ from (1.5) and so we first have to determine the canonical decomposition $X = X_0 + M + \int d\langle M\rangle \, \widehat{\lambda}$ of X under P. But (1.1) immediately yields

$$M_t = \int\limits_0^t X_u Y_u \, dW_u \qquad , \qquad A_t := \int\limits_0^t d\langle M\rangle_u \, \widehat{\lambda}_u = \int\limits_0^t X_u \, \mu(u, X_u, Y_u) \, du$$

so that $\langle M\rangle_t = \int\limits_0^t X_u^2 Y_u^2 \, du$ and $\widehat{\lambda}_t = \frac{dA_t}{d\langle M\rangle_t} = \frac{\mu(t, X_t, Y_t)}{X_t Y_t^2}$. This implies that

$$(2.5) \qquad \widehat{K}_t = \int\limits_0^t \left(\frac{\mu(u, X_u, Y_u)}{Y_u}\right)^2 du$$

and therefore

$$\widehat{Z}_t = \mathcal{E}\left(-\int \widehat{\lambda} \, dM\right)_t = \exp\left(-\int\limits_0^t \frac{\mu(u, X_u, Y_u)}{Y_u} \, dW_u - \frac{1}{2}\int\limits_0^t \left(\frac{\mu(u, X_u, Y_u)}{Y_u}\right)^2 du\right).$$

If $\widehat{Z}$ is a true P-martingale, Girsanov's theorem implies that $\widehat{W}^1 := W^1 + \int \frac{\mu(u, X_u, Y_u)}{Y_u}\, du$ and $\widehat{W}^2 := W^2$ are independent $\widehat{P}$-Brownian motions. Under the minimal ELMM $\widehat{P}$, the SDE for X and Y therefore takes the form

$$\frac{dX_t}{X_t} = Y_t\, d\widehat{W}_t,$$

$$dY_t = a^{\widehat{P}}(t, X_t, Y_t)\, dt + b(t, X_t, Y_t)\, d\widehat{W}'_t$$

for $\widehat{P}$-Brownian motions $\widehat{W}, \widehat{W}'$ with instantaneous correlation $\varrho(t, X_t, Y_t)$ and

(2.6) $$a^{\widehat{P}}(t, x, y) = a(t, x, y) - \frac{1}{y}(b\varrho\mu)(t, x, y).$$

To apply the results from section 1, it remains to check whether $\widehat{Z}$ is a true P-martingale and square-integrable under P. A well-known sufficient condition for both is boundedness of $\widehat{K}$ (uniformly in t and ω) and we shall see other sufficient conditions in the examples below. This completes step a) for local risk-minimization; observe how easily one obtains $\widehat{P}$ by simply looking at the canonical decomposition of X under P.

2.2. Mean-variance hedging: Finding $\widetilde{P}$ and related quantities

For mean-variance hedging, step a) is in general a bit more involved. If X has a deterministic mean-variance tradeoff, Lemma 3 gives $\widetilde{P} = \widehat{P}$ and so there is no problem. This will cover two of our models below. For the other two, we use recent results obtained by Laurent/Pham (1999) in a multidimensional diffusion model and translate these directly into our present context. First of all, we need to assume that $\varrho \equiv 0$ so that the Brownian motions W and W' driving X and Y in (1.1) are *independent*. We also have to assume that the coefficient functions $\mu(t, x, y)$, $a(t, x, y)$ and $b(t, x, y)$ in (1.1) all do not depend on x so that (1.1) reduces to

(2.7) $$\frac{dX_t}{X_t} = \mu(t, Y_t)\, dt + Y_t\, dW_t,$$

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

under P. This is a "Black-Scholes model in a random environment": the diffusion Y determines the drift and volatility of X, but is independent of X and in that sense completely exogenous. It would be highly desirable to remove these restrictions and allow in particular nonzero correlation between X and Y, but despite some recent advances in Biagini/Guasoni/Pratelli (2000), this problem still awaits further research.

So consider the specialized model (2.7). Under suitable assumptions on μ, a, b, (X, Y) is a Markov process under P. From (6.4) and (6.1) of Laurent/Pham (1999), we then obtain

$$(2.8) \qquad j_t := E\left[\left(\frac{Z_T^{\widetilde{P}}}{Z_t^{\widetilde{P}}}\right)^2 \Bigg| \mathcal{F}_t\right] = \frac{\widetilde{Z}_t}{Z_t^{\widetilde{P}}} = \exp\left(J(t, Y_t)\right) \qquad , \qquad 0 \leq t \leq T$$

where the function $J(t, y)$ is explicitly given by

$$(2.9) \qquad J(t, y) := -\log E\left[\exp\left(-\int_t^T \left(\frac{\mu(u, Y_u^{t,y})}{Y_u^{t,y}}\right)^2 du\right)\right]$$

with

$$dY_u^{t,y} = a(u, Y_u^{t,y})\, du + b(u, Y_u^{t,y})\, dW_u' \qquad , \qquad Y_t^{t,y} = y.$$

If J is sufficiently smooth, Itô's formula yields the canonical decomposition of j as

$$dj_t = j_t \frac{\partial J}{\partial y}(t, Y_t) b(t, Y_t)\, dW_t' + \ldots dt$$

and Theorem 4.2 of Laurent/Pham (1999) then yields

$$\frac{\widetilde{\zeta}_t}{\widetilde{Z}_t} = -\widetilde{a}_t^{\mathrm{GLP}} = -\frac{\mu(t, Y_t)}{X_t Y_t^2} \qquad , \qquad 0 \leq t \leq T$$

and

$$Z^{\widetilde{P}} = \mathcal{E}\left(-\int \widehat{\lambda}\, dM - \int \widetilde{\nu}\, dW'\right) = \mathcal{E}\left(-\int \frac{\mu(u, Y_u)}{Y_u}\, dW_u - \int \widetilde{\nu}_u\, dW_u'\right)$$

with

$$(2.10) \qquad \widetilde{\nu}_t = \frac{\partial J}{\partial y}(t, Y_t) b(t, Y_t) \qquad , \qquad 0 \leq t \leq T.$$

By Girsanov's theorem, $\widetilde{W} := W + \int \frac{\mu(u, Y_u)}{Y_u}\, du$ and $\widetilde{W}' := W' + \int \widetilde{\nu}_u\, du$ are *independent* Brownian motions under $\widetilde{P}$. Hence the SDE for X and Y under $\widetilde{P}$ takes the form

$$\frac{dX_t}{X_t} = Y_t\, d\widetilde{W}_t,$$

$$dY_t = a^{\widetilde{P}}(t, Y_t)\, dt + b(t, Y_t)\, d\widetilde{W}_t'$$

with

$$(2.11) \qquad a^{\widetilde{P}}(t, y) = a(t, y) - b^2(t, y) \frac{\partial J}{\partial y}(t, y)$$

15

thanks to (2.10). This completes step a) for mean-variance hedging.

2.3. On the comparison of the two approaches

Due to the Markovian structure of our models, we can give more explicit expressions for the quantities R_0^* and $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right]$ required for our comparison. In fact, (1.4) and (2.4) imply

$$(2.12) \quad \mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] = E\left[\left(L_T^{H,\widehat{P}}\right)^2\right] = E\left[\int_0^T \left(b^2(1-\varrho^2)\left(\frac{\partial v^{\widehat{P}}}{\partial y}\right)^2\right)(u, X_u, Y_u)\, du\right]$$

in general and for $\varrho \equiv 0$

$$(2.13) \quad R_0^* = E\left[\int_0^T e^{-J(u, Y_u)} b^2(u, Y_u)\left(\frac{\partial v^{\widetilde{P}}}{\partial y}(u, X_u, Y_u)\right)^2 du\right]$$

by (1.12), (2.8) and (2.4) for $Q = \widetilde{P}$, because $\varrho \equiv 0$.

Remark. We emphasize that neither local risk-minimization nor mean-variance hedging introduce memory effects into valuation in this situation. This is obvious from the fact that values are given by conditional expectations and (X, Y) is Markovian under both $\widetilde{P}$ and $\widehat{P}$. For hedging, things are a bit more subtle. The recursive structure of (1.9) shows that the mean-variance optimal strategy depends in a possibly complicated way on the past evolution of gains from trade. However, (1.3) and (2.3) also show that the locally risk-minimizing strategy is in general not a simple delta hedge; this was already pointed out by Frey (1997).

3. Specific models

To compare performances, we now explicitly compute in some examples the optimal strategies, the corresponding value processes and in particular the expected squared total costs for local risk-minimization versus mean-variance hedging. We do this for *stochastic volatility models* and the work in this section goes into verifying that we can apply the preceding theory.

We consider four models. Two are of the *Stein/Stein (1991) type*, two of the *Heston (1993) type* and this terminology in itself calls for explanation. We first point out that most of the literature on option pricing under stochastic volatility does not completely specify the underlying model. Authors usually describe the volatility dynamics under a specific ELMM that they use for pricing and they often do this by stating assumptions on the market price of volatility risk. But hardly any paper gives explicitly the drift of the stock price under

the original measure P. The reason is that almost all papers consider only pricing, but not hedging, and their arguments for choosing a particular ELMM for pricing do not involve the original measure P. In both approaches studied here, the situation is completely different. We are mainly concerned with hedging and so we start from a model for X and its volatility Y under the *original* ("real-world" or "objective") measure P. This is essential because we want to *quantify the riskiness* of our hedging strategies and this should of course be done under the original measure P, not under an ELMM. In that sense, the description under P is really taken seriously here. Each of the two criteria then selects its corresponding ELMM — $\widehat{P}$ for local risk-minimization, $\widetilde{P}$ for mean-variance hedging — and so the resulting martingale dynamics depend on the specification of the P-drift μ.

Having made this important point, we now turn to our examples. We always start from an SDE of the form

$$(2.7) \qquad \frac{dX_t}{X_t} = \mu(t, Y_t)\, dt + Y_t\, dW_t,$$

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

with P-Brownian motions W, W' with constant instantaneous correlation ϱ, i.e., $d\langle W, W'\rangle_t = \varrho\, dt$. The four models are then summarized by the following table:

Model	Type	Coefficients of Y	Drift of X	Correlation ϱ
S1	Stein/Stein	$a_1(t,y) = -\delta(y - \beta)$ $b_1(t,y) = k$	$\mu_1(t,y) = \Delta y$	0
S2	Stein/Stein	same as S1	$\mu_2(t,y) = \gamma y^2$	0
H1	Heston	$a_2(t,y) = \frac{(4\kappa\theta - \Sigma^2)}{8y} - \frac{\kappa}{2}y$ $b_2(t,y) = \frac{\Sigma}{2}$	$\mu_1(t,y) = \Delta y$	ϱ_0
H2	Heston	same as H1	$\mu_2(t,y) = \gamma y^2$	0

Table 3.1 *Summary of the four models*

The constants $\delta, \beta, k, \kappa, \theta, \Sigma$ are all nonnegative, Δ, γ are real constants and ϱ_0 is in $(-1, 1)$.

In the Stein/Stein type models, Y is independent of W and an Ornstein-Uhlenbeck process around some level β. In the Heston type models, Y is the square root of a Cox-Ingersoll-Ross type process; this means that $F := Y^2$ follows the SDE

$$(3.1) \qquad dF_t = \kappa(\theta - F_t)\, dt + \Sigma\sqrt{F_t}\, dW'_t$$

as one easily sees from Itô's formula. For each of these two types, we consider two specially chosen drift functions μ_1 and μ_2. The first choice makes $\widehat{K}$ deterministic so that Lemma 3 applies; the second drift is chosen to allow an explicit computation of the function J from section 2. To avoid integrability problems, we take a contingent claim of the form $H = h(X_T)$ with a *bounded* function h. The computations in the next section use a put option where $h(x) = (K - x)^+$.

Remarks. **1)** We have not yet said anything about the boundary conditions for Y at the origin. In the Heston type models, we assume $\kappa\theta \geq \frac{1}{2}\Sigma^2$ since this guarantees by Feller's test for explosions a strictly positive solution for F in (3.1). For the Stein/Stein type models, it was already pointed out in Stein/Stein (1991) that only Y^2 enters the calculations for option values and strategies; hence our (and their) formulation is equivalent to putting a reflecting barrier at 0 for Y.

2) For fairly general stochastic volatility models, Sin (1996) has studied the question whether X is a true or only a local martingale under a given ELMM Q; see also Sin (1998). For our context, this gives results about the properties of X under the minimal ELMM $\widehat{P}$, but not in general under the variance-optimal ELMM $\widetilde{P}$ because Y has under $\widetilde{P}$ time-dependent coefficients. For the models S1, S2, H2 with correlation 0, we obtain that X is always a true martingale under $\widehat{P}$. For the H1 model, Theorem 4.5.2 of Sin (1996) shows that X is under $\widehat{P}$ a true martingale if and only if the correlation ϱ_0 is nonpositive. For $\varrho_0 > 0$, X is for H1 a strict local $\widehat{P}$-martingale. But in any case, all these results are not good enough for our purposes because we need additional integrability properties.

3.1. The simple drift models S1 and H1

In the much easier case of the drift function μ_1, (2.5) implies that

$$(3.2) \qquad \widehat{K}_t = \int\limits_0^t \left(\frac{\mu_1(u, Y_u)}{Y_u} \right)^2 du = \Delta^2 t$$

is deterministic, hence also bounded uniformly in $t \in [0, T]$ and ω. This implies that $\widehat{P} \in \mathbb{P}_e^2 \neq \emptyset$ and that $\widetilde{P} = \widehat{P}$ by Lemma 3. Step a) of the scheme in section 2 is therefore immediate and needs in particular no assumption on the correlation; note that Lemma 3 provides all required quantities related to $\widetilde{P}$. Step b) is the same for both approaches and so subsection 2.2 gives all we need for mean-variance hedging.

For local risk-minimization, some work remains to be done because the Galtchouk-Kunita-Watanabe decomposition (1.6) does not immediately yield the Föllmer-Schweizer decomposition (1.2); see the discussion at the end of subsection 1.1. But since X is continuous and $\widehat{K}$ is bounded, Corollary 5 of Pham/Rheinländer/Schweizer (1998) implies that

each contingent claim $H \in L^2(P)$ admits a decomposition (1.2). In particular, this ensures the existence of a pseudo-locally risk-minimizing strategy φ^{lr} and we can identify this from (1.6) without having to worry about integrability problems. The resulting strategy φ^{lr} is by construction pseudo-locally risk-minimizing. To show that it is also locally risk-minimizing in the original sense, one has to check whether the assumptions of Proposition 2.3 of Schweizer (1991) are satisfied. This is straighforward for model S1, but less so for model H1. In fact, one can verify these hypotheses by the methods from subsection 3.2 if $\varrho_0 = 0$ and if the time horizon is sufficiently short, but it is in theory not clear if they hold in general.

Remark. The above discussion has glossed over the fact that the PDE arguments in section 2 require some regularity assumptions. This is a tricky point because the standard sufficient conditions (e.g., Lipschitz-continuity) are not satisfied in our models. However, this problem can be overcome in all examples considered here; see Heath/Schweizer (2000) for more details.

3.2. The models S2 and H2

Now we consider the drift function μ_2. This involves additional work because now

$$\widehat{K}_t = \int\limits_0^t \left(\frac{\mu_2(u, Y_u)}{Y_u} \right)^2 du = \int\limits_0^t \gamma^2 Y_u^2 \, du$$

is random and unbounded. We first want to obtain moment estimates for X and $\widehat{Z}$ (to show that $\widehat{P} \in I\!\!P_e^2$) and the following result provides the key tool for this. For a stochastic process $U = (U_t)_{0 \leq t \leq T}$, $U_t^* := \sup\limits_{0 \leq s \leq t} |U_t|$ for $0 \leq t \leq T$ denotes the associated maximal process.

Lemma 4. On a filtered probability space $(\Omega, \mathcal{F}, I\!\!F, P)$, let W be a Brownian motion and B an $I\!\!F$-predictable process with $\int\limits_0^T B_u^2 \, du < \infty$ P-a.s. For $\alpha, \beta \in I\!\!R$, define the processes

$$L_t(\alpha) := \mathcal{E} \left(\alpha \int B \, dW \right)_t = \exp \left(\alpha \int\limits_0^t B_u \, dW_u - \tfrac{1}{2}\alpha^2 \int\limits_0^t B_u^2 \, du \right) \qquad , \qquad 0 \leq t \leq T$$

and

$$D_t(\beta) := \mathcal{E} \left(\beta \int B_u^2 \, du \right)_t = \exp \left(\beta \int\limits_0^t B_u^2 \, du \right) \qquad , \qquad 0 \leq t \leq T.$$

For every $p \in [2, \infty)$, we can then find a constant c depending only on α, β, p such that

$$(3.3) \qquad\qquad E\left[|D_T^*(\beta)|^p \right] \leq E\left[\exp\left(cT(B_T^*)^2 \right) \right]$$

19

and if B is independent of W,

$$(3.4) \qquad E\left[|L_T^*(\alpha)|^p\right] \le c\left(1 + T^{\frac{p}{2}-1} E\left[\exp\left(c(T + T^{\frac{2}{p}})(B_T^*)^2\right)\right]\right).$$

Proof. See appendix.

Corollary 5. *Consider either model S2 or model H2. For each $p \in [2, \infty)$, there exists a time horizon $T_0(p) > 0$ such that*

$$(3.5) \qquad \sup_{0 \le t \le T} X_t \in L^p(P)$$

and

$$(3.6) \qquad \sup_{0 \le t \le T} \widehat{Z}_t \in L^p(P)$$

for $0 < T \le T_0(p)$. In particular, we obtain for a sufficiently short time horizon T that

(3.7) $I\!P_e^2 \ne \emptyset,$

(3.8) X *is a $\widehat{P}$-martingale*

and

(3.9) Θ *contains all uniformly bounded predictable processes.*

Proof. See appendix.

Remark. It may seem strange that one has to restrict the time horizon T to obtain sufficient integrability properties. However, this is an intrinsic feature of the models used here and not a technical issue raised by our approach. Consider for instance the Stein/Stein type models where Y is an Ornstein-Uhlenbeck process. Conditionally on Y, $\widehat{Z}_T$ is lognormal with parameters involving $I_T(\gamma) := \int_0^T \gamma^2 Y_u^2 \, du$ and so any unconditional moment of $\widehat{Z}_T$ of order $p > 1$ is an exponential moment of $I_T(\gamma)$. But the latter can be computed explicitly and becomes infinite unless γ and T are small enough. The same applies to X instead of $\widehat{Z}$. For all our examples, however, there were no problems: $T = 100$ (years) was small enough in the sense that numerical computation of $E\left[\exp\left(I_T(\gamma)\right)\right]$ was stable and gave a finite value.

For the rest of this section, we choose T sufficiently small so that the conclusions of Corollary 5 hold.

3.2.1. Local risk-minimization in models S2 and H2

To obtain a locally risk-minimizing strategy, we first note that similar arguments as in the proof of Corollary 5 show that the hypotheses of Proposition 2.3 of Schweizer (1991) are satisfied. Thus pseudo-locally risk-minimizing strategies are also locally risk-minimizing and it remains to obtain the Föllmer-Schweizer decomposition (1.2) of H. To that end, we determine as in step b) in section 2 by PDE methods the decomposition (1.6) under $\widehat{P}$ and refer again to Heath/Schweizer (2000) for a justification. Because we no longer have a general existence result for a Föllmer-Schweizer decomposition, we then have to show (as explained in section 1) that $V_0^{H,\widehat{P}}$, $\xi^{H,\widehat{P}}$ and $L^{H,\widehat{P}}$ possess the required integrability properties. Since $\mathcal{F}_0$ is trivial, $V_0^{H,\widehat{P}}$ as a constant is of course in $L^2(P)$ and since the contingent claim H is bounded, so is the process $V^{H,\widehat{P}}$; hence

$$(3.10) \qquad \left(V^{H,\widehat{P}}\right)_T^* \in L^2(P).$$

Provided that $\xi^{H,\widehat{P}}$ is bounded, it is in Θ by (3.9) so that $\left(\int \xi^{H,\widehat{P}}\, dX\right)_T^*$ is in $L^2(P)$. Taking differences in (1.6) and using (3.10) then implies that $\left(L^{H,\widehat{P}}\right)_T^* \in L^2(P)$ so that the local P-martingale $L^{H,\widehat{P}}$ is a square-integrable true P-martingale. Putting all this together, we conclude that $\xi^{H,\widehat{P}}$ and $L^{H,\widehat{P}}$ do indeed provide the locally risk-minimizing strategy for H.

It remains to show that $\xi^{H,\widehat{P}}$ is bounded and for our European put option, this can be done with a standard convexity argument. In fact, conditioning on Y yields

$$V_t^{H,\widehat{P}} = \widehat{E}\left[(K-X_T)^+\big|\mathcal{F}_t\right] = \widehat{E}\left[\widehat{E}\left[(K-X_T)^+\big|\sigma(Y)\vee\mathcal{F}_t\right]\Big|\mathcal{F}_t\right]$$

and by the independence of X and Y under $\widehat{P}$, the conditional distribution of X_T given Y and $\mathcal{F}_t$ is lognormal. Hence the inner conditional expectation is simply the Black-Scholes put price at time t in a model with known instantaneous volatility $\left(Y_s(\omega)\right)_{0\le s\le T}$ and in particular a convex differentiable function of X_t with values in $[0,(K-X_t)^+]$. Since these properties are preserved under averaging over the conditional distribution of Y given $\mathcal{F}_t$, we obtain

$$0 \le V_t^{H,\widehat{P}} = v^{\widehat{P}}(t,X_t,Y_t) \le (K-X_t)^+.$$

Because $v^{\widehat{P}}$ is convex and differentiable in x, taking partial derivatives and using $\varrho = 0$ yields

$$0 \ge \frac{\partial v^{\widehat{P}}}{\partial x}(t,X_t,Y_t) = \xi_t^{H,\widehat{P}} \ge -1;$$

hence $\xi^{H,\widehat{P}}$ is uniformly bounded. In summary, then, we have shown that we can obtain the locally risk-minimizing strategy for a European put in models S2 and H2 as in section 2.

3.2.2. Mean-variance hedging in models S2 and H2

Now we turn to mean-variance hedging. Theorem 1 and (3.7) guarantee the existence of a mean-variance optimal strategy for both models S2 and H2. Moreover, arguments as in Heath/Schweizer (2000) show that the Galtchouk-Kunita-Watanabe decomposition (1.7) of H under $\widetilde{P}$ can be obtained by PDE techniques as in section 2. Thus only step a) remains and subsection 2.2 shows that this boils down to computing the function $J(t,y)$ in (2.9). But

$$J(t,y) = -\log E\left[\exp\left(-\gamma^2 \int_t^T (Y_u^{t,y})^2\,du\right)\right]$$

by the specific choice of μ_2 and so $J(t,Y_t)$ is just the logarithm of the conditional Laplace transform of $\int_t^T Y_u^2\,du$ given $\mathcal{F}_t$. For both our models, this can be determined explicitly:

In model S2, Y is an Ornstein-Uhlenbeck process and so $J(t,y)$ can be computed along the lines of Leblanc (1996). We cannot use Leblanc's actual formula since his computations contain some errors, but following his steps leads after some lengthy calculations to

$$J_{\text{S2}}(t,y) = f_0(T-t) + f_1(T-t)\frac{y}{k} + f_2(T-t)\frac{y^2}{k^2}$$

with

$$f_2(\tau) := \frac{\lambda\gamma_1 e^{-2\gamma_1\tau}}{\lambda + \gamma_1 - \lambda e^{-2\gamma_1\tau}} - \lambda,$$

$$f_1(\tau) := \frac{1}{1 + 2\lambda\psi(\tau)}\left((2D - D')e^{-2\gamma_1\tau} - 2De^{-\gamma_1\tau}\right) + D',$$

$$f_0(\tau) := \frac{1}{2}\log\left(1 + 2\lambda\psi(\tau)\right) - \left(\lambda + \frac{\delta^2\beta^2}{2k^2}(\frac{\delta^2}{\gamma_1^2} - 1)\right)\tau - \frac{2D^2\psi(\tau)}{1 + 2\lambda\psi(\tau)}$$

$$+ \frac{\delta^2\beta}{k\gamma_1^2}\left(\frac{1}{1 + 2\lambda\psi(\tau)}\left(2De^{-\gamma_1\tau} - (D - \frac{1}{2}D')e^{-2\gamma_1\tau}\right) - (D + \frac{1}{2}D')\right)$$

with constants $\gamma_1 := \sqrt{2k^2\gamma^2 + \delta^2}$, $\lambda := \frac{\delta - \gamma_1}{2}$, $D := \frac{\delta\beta}{2k}\left(1 - \frac{\delta^2}{\gamma_1^2}\right)$, $D' := \frac{\delta\beta}{k}\left(1 - \frac{\delta}{\gamma_1}\right)$ and the function $\psi(\tau) := \frac{1 - e^{-2\gamma_1\tau}}{2\gamma_1}$. According to (2.11), the $\widetilde{P}$-dynamics of Y in model S2 are then

$$dY_t = \left(\delta\beta - kf_1(T-t) - \left(\delta + 2f_2(T-t)\right)Y_t\right)dt + k\,d\widetilde{W}_t'.$$

22

This completes the required specifications for mean-variance hedging in model S2.

In model H2, Y^2 is a Cox-Ingersoll-Ross type process; finding the above Laplace transform is thus equivalent to computing zero coupon bond prices in the CIR model. This gives

$$J_{\text{H2}}(t, y) = - \log g(T - t) + \alpha(T - t)\gamma^2 y^2$$

with $\Gamma := \sqrt{2\gamma^2 \Sigma^2 + \kappa^2}$ and the functions

$$\alpha(\tau) := \frac{2\left(e^{\Gamma\tau} - 1\right)}{(\Gamma + \kappa)\left(e^{\Gamma\tau} - 1\right) + 2\Gamma},$$

$$g(\tau) := \left(\frac{2\Gamma e^{\frac{\Gamma+\kappa}{2}\tau}}{(\Gamma + \kappa)(e^{\Gamma\tau} - 1) + 2\Gamma}\right)^{\frac{2\kappa\theta}{\Sigma^2}}.$$

Using again (2.11) gives the $\widetilde{P}$-dynamics of Y in model H2 as

(3.11)
$$dY_t = \left(a_2(t, Y_t) - \frac{1}{2}\Sigma^2\gamma^2\alpha(T - t)Y_t\right) dt + \frac{\Sigma}{2} d\widetilde{W}'_t,$$

specifying all we need for mean-variance hedging in model H2.

4. Numerical results

This section presents numerical results on the effects of using local risk-minimization or mean-variance hedging. We show values, expected squared total costs and strategies for the Heston models H1 and H2 from section 3. Similar results obtained for the Stein/Stein models S1 and S2 are omitted for the sake of brevity. We comment below on some of these additional results and refer to Heath/Platen/Schweizer (1999) for more details.

4.1. Methods and parameters

To compute values under mean-variance hedging and local risk-minimization, we solve the PDE (2.1), (2.2) for v^Q with $Q = \widetilde{P}$ and $Q = \widehat{P}$ respectively. The drift functions $a^{\widetilde{P}}$ and $a^{\widehat{P}}$ for the H2 model are given by (2.11) with $J = J_{\text{H2}}$ from subsection 3.2.2 and by (2.6) with $\varrho = 0$. To obtain approximate numerical solutions to the PDEs for $v^{\widetilde{P}}$ and $v^{\widehat{P}}$, we used finite difference methods based mainly on the Crank-Nicholson scheme. The basic method was embedded in a fractional step procedure to accommodate the two spatial dimensions corresponding to the components X and Y. Additional information on the use of these

and other numerical PDE techniques can be found in sections 8.2 – 8.5 of Fletcher (1988) and chapters 11 and 14 of Hoffman (1993). We also experimented with different boundary conditions to obtain the most accurate and stable results. For example, option values for high and low volatilities were computed by setting the second partial derivative of the value with respect to volatility equal to zero.

The expected squared total cost for local risk-minimization is given by (2.12) with $\varrho = 0$. For mean-variance hedging, it is given by (2.13) in model H2 and by

$$
(4.1) \qquad R_0^* = E\left[\int_0^T e^{-\Delta^2(T-u)}(1-\varrho_0^2)\frac{\Sigma^2}{4}\left(\frac{\partial v^{\widehat{P}}}{\partial y}(u, X_u, Y_u)\right)^2 du\right]
$$

for model H1; this follows from (1.12), Lemma 3, (3.2) and (2.4). To compute all these quantities, we again employed PDE techniques; see Heath/Platen/Schweizer (1999) for details on the methods used for orthogonalization and interpolation.

The hedging strategies ϑ^{lr} and ϑ^{mvo} for local risk-minimization and mean-variance hedging are obtained from (1.3) and (1.9) respectively. This requires the calculation of the integrands $\xi^{H,\widehat{P}}$ and $\xi^{H,\widetilde{P}}$ as explained in section 2. To compute these numerically for a given sample path, we used an order 1.0 weak predictor-corrector scheme as in section 15.5 of Kloeden/Platen (1992) and applied this to the system in (2.7) to first obtain a simulated sample path for the vector process (X, Y). This was then used to obtain sample paths for the integrands $\xi^{H,\widehat{P}}$ and $\xi^{H,\widetilde{P}}$ by using pre-computed and interpolated values for the valuation functions $v^{\widehat{P}}$ and $v^{\widetilde{P}}$. More details on this procedure are again provided in Heath/Platen/Schweizer (1999).

All the numerical experiments described here were done for a European put option with payoff $h(X_T) = (K - X_T)^+$. The default parameter values for the Heston models H1 and H2 were $\kappa = 5.0$, $\theta = 0.04$, $\Sigma = 0.6$, $\Delta = 0.5$, $\gamma = 2.5$ and $\varrho_0 = 0$. We also considered nonzero correlation cases for the H1 model; see the results in Figure 4.2. Default starting values for X and Y were $X_0 = 100$, $Y_0 = 0.2$ and default option parameters were strike $K = 100$ and time to maturity $T = 1.0$. We do not show the results for the Stein/Stein models S1 and S2, but we give for completeness the parameters we have used. These are $\delta = \kappa = 5.0$, $\beta = \sqrt{\theta} = 0.2$ and $k = \Sigma/2 = 0.3$.

Our choice of the above parameter values can be motivated as follows. According to (3.1), the chosen θ ensures a mean reversion level of $\sqrt{\theta} = 0.2$ for the volatility process Y in the Heston models. The parameter $\kappa = 5.0$ then produces a half-life of approximately 0.14 years for the deterministic part of (3.1). Our positivity condition $\kappa\theta \geq \frac{1}{2}\Sigma^2$ for F entails $\Sigma \leq \sqrt{2\kappa\theta} \approx 0.63$ for $\theta = 0.04$ and $\kappa = 5.0$. Hence the value $\Sigma = 0.6$ produces a strong stochastic volatility effect while just staying within the allowable range for this parameter. To obtain the same mean reversion level and rate for the Stein/Stein models, the values for δ

and β were set at $\delta = \kappa = 5.0$ and $\beta = \sqrt{\theta} = 0.2$ respectively. The value $k = \frac{\Sigma}{2} = 0.3$ ensures that the diffusion terms b_1 and b_2 in Table 3.1 are equal for both models. Finally, $\Delta = 0.5$ implies in the S1 and H1 models an average drift rate for X of 0.1 because of the previously specified mean reversion level of 0.2 for Y. The same average drift rate is obtained for the S2 and H2 models with $\gamma = 2.5$.

4.2. Values

Let us first examine the time 0 values $\widehat{E}[H]$ and $\widetilde{E}[H]$ under local risk-minimization and mean-variance hedging respectively. Since the H1 model has $\widetilde{P} = \widehat{P}$ by Lemma 3, there is nothing to compare there so that we focus on the H2 model. Figure 4.1 shows the resulting value differences as functions of time to maturity T and moneyness $\log(X_0/K)$. Typical value differences for at-the-money options were of the order of $2 - 3\%$. With $T = 1.0$ and $\log(X_0/K) = 0$, for example, the computed values were $v^{\widehat{P}}(0, X_0, Y_0) = 7.6945$ and $v^{\widetilde{P}}(0, X_0, Y_0) = 7.8927$. For out-of-the-money options, relative value differences increased substantially; for instance, $T = 1.0$ and $\log(X_0/K) = 0.3$ gave computed values of $v^{\widehat{P}}(0, X_0, Y_0) = 0.7640$ and $v^{\widetilde{P}}(0, X_0, Y_0) = 0.8486$, hence a difference of about 10%.
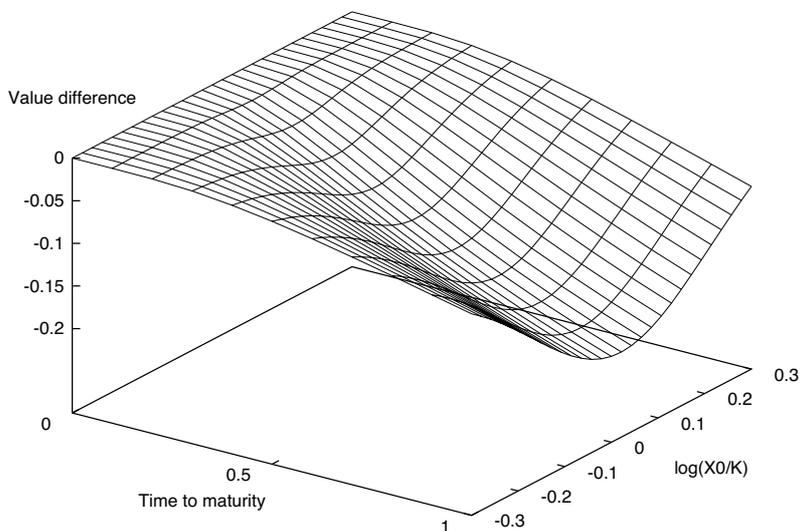
Figure 4.1 *Value differences* $v^{\widehat{P}}(0, X_0, Y_0) - v^{\widetilde{P}}(0, X_0, Y_0)$ *for the H2 model*

Note that values under local risk-minimization seem to lie systematically below values for mean-variance hedging. Since these results are for puts, the inequalities would be reversed for call values.

4.3. Expected squared total costs

Our main interest is in comparing R_0^* and $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right]$, the expected squared total costs for the two competing approaches. The (nonnegative) differences $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] - R_0^*$ are shown in Figure 4.2 for the H1 model for different values of the correlation parameter ϱ_0 and the moneyness $\log(X_0/K)$. As for values, differences tend to be maximal at the money and also for correlation 0. For the other extreme case $|\varrho_0| = 1$, the expected squared total cost for both hedging approaches is zero; this can be seen from (2.12) and (4.1) and is also clear since $|\varrho_0| = 1$ yields a complete model.
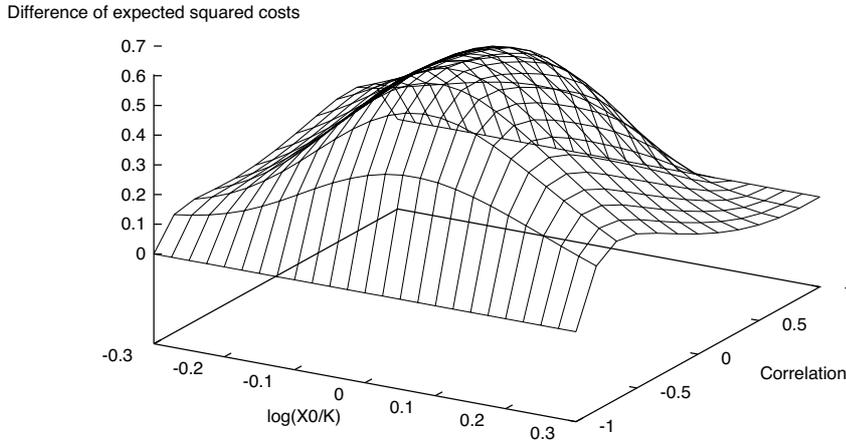
Figure 4.2 *Expected squared cost differences* $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] - R_0^*$ *for the H1 model*

To give some idea of the relative sizes of values and expected squared total costs, we mention that $\log(X_0/K) = 0$ and $\varrho_0 = 0$ yielded the values $v^{\widehat{P}}(0, X_0, Y_0) = v^{\widetilde{P}}(0, X_0, Y_0) = 7.6910$, $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] = 4.2573$ and $R_0^* = 3.6855$. For $\log(X_0/K) = 0$ and $\varrho_0 = -0.5$ the corresponding values were $v^{\widehat{P}}(0, X_0, Y_0) = 10.6616$, $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] = 4.4290$ and $R_0^* = 3.8361$ and for an out-of-the-money put option with $\log(X_0/K) = 0.3$ and $\varrho_0 = 0.5$, the resulting values were $v^{\widehat{P}}(0, X_0, Y_0) = 2.0382$, $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] = 1.5510$ and $R_0^* = 1.3351$.

Figure 4.3 provides a different view on the expected squared total cost. Here we plot this with the default parameters for the H2 model and for both hedging approaches over the time interval $[0, T]$ with $T = 1.0$. Similar results were obtained for the H1 model. As noted above, the hedging cost for both approaches becomes smaller in general as one moves out of the money. In the H2 model with $\log(X_0/K) = 0$ and $T = 1.0$, for example, the computed costs were $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] = 4.2364$ and $R_0^* = 3.8322$. With $\log(X_0/K) = 0.3$ and $T = 1.0$, the corresponding values were $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] = 1.1304$ and $R_0^* = 1.1287$. This means that Figure 4.3 with $\log(X_0/K) = 0$ represents the most informative case.
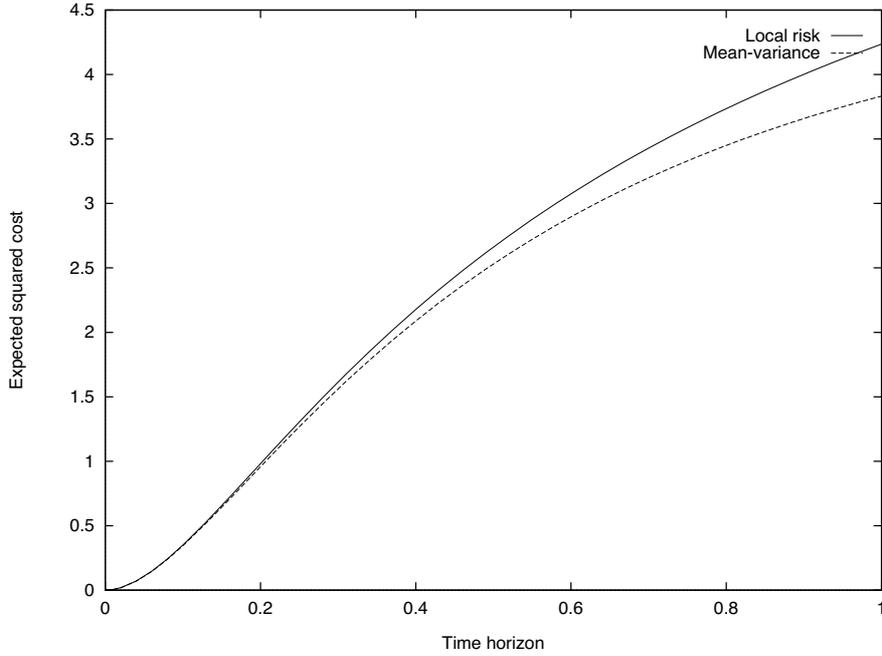
Figure 4.3 *Expected squared costs* $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right]$ *and* R_0^* *for the H2 model*

If one looks at Figure 4.3, it seems as if the expected squared total costs were increasing with T, but this is not the case. For large time horizons ($T \approx 100$), our numerical experiments indicated that R_0^* approaches zero whereas $\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right]$ seems to remain above a certain fixed positive value. These results were obtained for all four models and seem to indicate an interesting difference between the asymptotic properties of the two hedging methods. This is pointed out in Heath/Platen/Schweizer (1999) and deserves further investigation.

Remark. In subsection 1.3, we have derived the quantity $e^{-\widehat{K}_T}$ as a lower bound for the ratio $R_0^*/\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right]$ in the H1 and S1 models where Lemma 3 applies. Not surprisingly, this bound is rather crude for larger T. For example, the H1 model with $T = 0.01$ gave $R_0^*/\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] = 0.9982$ and $e^{-\widehat{K}_T} = 0.9975$ whereas the values obtained for $T = 1.0$ were $R_0^*/\mathrm{Var}\left[C_T(\varphi^{\mathrm{lr}})\right] = 0.8657$ and $e^{-\widehat{K}_T} = 0.7788$.

4.4. Strategies

Since we have no explicit formulas for the optimal strategies, a reasonable comparison is rather difficult. Just to illustrate the qualitative behaviour to some extent, we show in Figure 4.4 the strategies ϑ^{lr} and ϑ^{mvo} in the H2 model for one asset price sample path ending in the money and in Figure 4.5 for another path ending out of the money. The corresponding two trajectories of $X/100$ and Y are displayed in Figure 4.6.

27
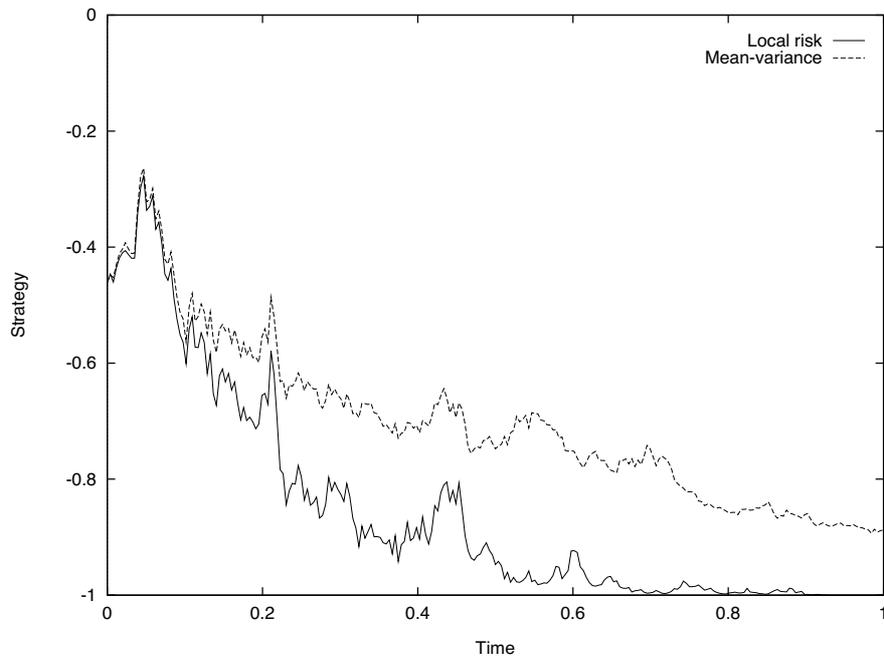
Figure 4.4 *Strategies ϑ^{lr} and ϑ^{mvo} for an in-the-money sample path: H2 model*
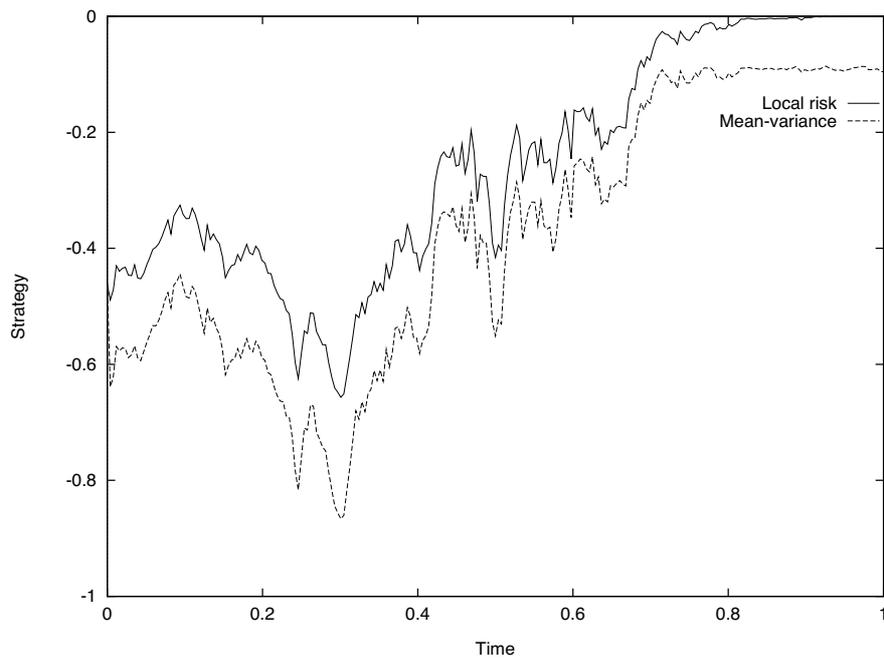
Figure 4.5 *Strategies ϑ^{lr} and ϑ^{mvo} for an out-of-the-money sample path: H2 model*
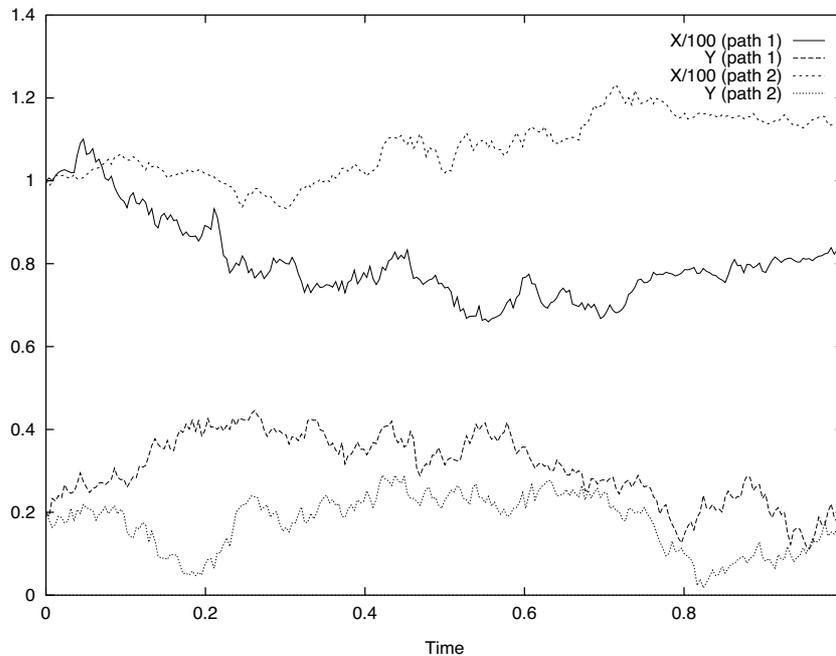
Figure 4.6 *Trajectories for $X/100$ and Y for the two sample paths used: H2 model*

4.5. Computational efforts

To conclude this section, we give some indications of the computational efforts involved in producing the above results. As previously mentioned, computing option values requires for both approaches the numerical solution of a (spatially) two-dimensional PDE. Because the function J is available in analytic form, the extra computational load for the mean-variance optimal strategy is small when compared to local risk-minimization. However, valuing under the variance-optimal measure requires re-initialization of the implicit solvers used at each time step; this is not necessary for local risk-minimization. All in all, we found for mean-variance hedging a typical increase of computation time of the order of 20%.

Expected squared total costs for both approaches were also computed by numerically solving a two-dimensional PDE; see Heath/Platen/Schweizer (1999) for a more detailed derivation. By using the same grid and similar levels of precision, we could therefore obtain these costs with approximately the same amount of computational work needed to generate values. Hedging strategies for both valuation measures are quickly and efficiently obtained by using finite differences as estimates for the required partial derivatives; for mean-variance hedging, (1.9) gives an additional term involving the cumulative trading gains from the past.

For all the plots shown in Figures 4.1 – 4.6, we used 160 time steps on a spatial grid of 80 points for the X component times 20 points for the Y component. A combined value/cost estimate for local risk-minimization then took about 1.0 seconds on a Pentium 233 MMX notebook computer. The corresponding time for mean-variance hedging was about 1.2 sec-

29

onds. In summary, one could say that for those situations where mean-variance hedging is sufficiently developed in theory, it is also applicable in practice with a tolerable amount of extra work.

5. Appendix: Proofs

This appendix collects the proofs of all the results in the paper.

Proof of Theorem 1. (1.8) follows from Proposition 2 of Schweizer (1996). According to Corollary 16 of Schweizer (1996), ϑ^{mvo} is then obtained by projecting the random variable $H - \widetilde{E}[H]$ in $L^2(P)$ on the space of all stochastic integrals $\int\limits_0^T \vartheta_u \, dX_u$ with $\vartheta \in \Theta'$ and this is in principle dealt with in Rheinländer/Schweizer (1997) (RS, for short). The representation (1.9) is very similar to their Theorem 6, but we cannot directly use their results since our examples may not satisfy their closedness assumption (1.2). Thus we use results from Gouriéroux/ Laurent/Pham (1998) (GLP, for short) and this involves a second change of measure. Because $\widetilde{Z}$ is a strictly positive P-martingale and $\widetilde{Z}_0$ is deterministic, we can define a probability measure $\widetilde{R} \approx \widetilde{P} \approx P$ by $\frac{d\widetilde{R}}{d\widetilde{P}} := \frac{\widetilde{Z}_T}{\widetilde{Z}_0}$. Clearly, the $I\!\!R^{d+1}$-valued process $Y = \begin{pmatrix} 1/\widetilde{Z} \\ X/\widetilde{Z} \end{pmatrix}$ is then a continuous local $\widetilde{R}$-martingale since $\widetilde{P} \in I\!\!P$. The density of $\widetilde{R}$ with respect to P is $\frac{d\widetilde{R}}{dP} = \frac{\widetilde{Z}_T^2}{\widetilde{Z}_0}$ and because $\mathcal{F}_0$ is trivial, $H \in L^2(P)$ if and only if $\frac{H}{\widetilde{Z}_T} \in L^2(\widetilde{R})$. The idea of GLP is now to use $\frac{\widetilde{Z}}{\widetilde{Z}_0}$ as a new numeraire, rewrite the original problem in terms of the corresponding quantities and apply the Galtchouk-Kunita-Watanabe decomposition theorem to $\frac{H}{\widetilde{Z}_T}$ under $\widetilde{R}$ with respect to Y. This yields

$$(5.1) \qquad \frac{H}{\widetilde{Z}_T} = E_{\widetilde{R}}\left[\frac{H}{\widetilde{Z}_T}\right] + \int\limits_0^T \psi_u \, dY_u + L_T$$

for some $I\!\!R^{d+1}$-valued $\psi \in L(Y)$ with $\int \psi \, dY \in \mathcal{M}_0^2(\widetilde{R})$ and some $L \in \mathcal{M}_0^2(\widetilde{R})$ strongly $\widetilde{R}$-orthogonal to Y. By Theorem 5.1 and the subsequent remark in GLP, ϑ^{mvo} is given by

$$(5.2) \qquad (\vartheta_t^{\mathrm{mvo}})^i = \psi_t^i + \widetilde{\zeta}_t^i \left(\frac{V_0^{\mathrm{mvo}}}{\widetilde{Z}_0} + \int\limits_0^t \psi_u \, dY_u - \psi_t^{\mathrm{tr}} Y_t \right) \qquad , \qquad 0 \le t \le T, i = 1, \dots, d$$

if we note that the relation between their terminology and ours is given by $V(\widetilde{a}^{\mathrm{GLP}}) = \frac{\widetilde{Z}}{\widetilde{Z}_0}$,

$X^i(\widetilde{a}^{\mathrm{GLP}}) = \widetilde{Z}_0 Y^i$ and $\widetilde{a}^{\mathrm{GLP}} = -\dfrac{\widetilde{\zeta}}{\widetilde{Z}}$. By using Proposition 8 of RS, (5.2) can be rewritten as

$$(5.3) \qquad\qquad \vartheta^{\mathrm{mvo}} = \frac{V_0^{\mathrm{mvo}}}{\widetilde{Z}_0}\widetilde{\zeta} + \vartheta$$

with ϑ corresponding to ψ from (5.1) via equation (4.6) in RS. Hence it only remains to obtain ϑ or ψ in terms of the decomposition (5.1) and this is basically already contained in RS if one looks carefully enough. More precisely, we start from (5.1) and argue as in Proposition 10 of RS to express the quantities in the decomposition (1.7) in terms of ψ and L. Note that as long as we make no integrability assertions, that argument only uses Proposition 8 of RS which holds as soon as $I\!\!P_e^2(X) \neq \emptyset$; see Remark 2) following that Proposition 8. The uniqueness of the Galtchouk-Kunita-Watanabe decomposition then implies that $L^{H,\widetilde{P}} = \int \widetilde{Z}_u\, dL_u$ and

$$\xi_t^{H,\widetilde{P}} = \frac{\widetilde{E}[H]}{\widetilde{Z}_0}\widetilde{\zeta}_t + \vartheta_t + L_{t-}\widetilde{\zeta}_t \qquad , \qquad 0 \leq t \leq T.$$

Solving this for ϑ and plugging the result into (5.3) yields by (1.8) the second expression in (1.9). The first then follows similarly as in the proof of Theorem 6 of RS.

<div align="right">**q.e.d.**</div>

Proof of Proposition 2. By (1.9), we have $\vartheta^{\mathrm{mvo}} = \xi^{H,\widetilde{P}} - \dfrac{\widetilde{\zeta}}{\widetilde{Z}}U_-$ and therefore

$$dU = dV^{H,\widetilde{P}} - \vartheta^{\mathrm{mvo}}\, dX = (\xi^{H,\widetilde{P}} - \vartheta^{\mathrm{mvo}})\, dX + dL^{H,\widetilde{P}} = U_-\frac{\widetilde{\zeta}}{\widetilde{Z}}\, dX + dL^{H,\widetilde{P}}$$

which is (1.10). Since $[L^{H,\widetilde{P}}, X] = 0$ by the strong $\widetilde{P}$-orthogonality of $L^{H,\widetilde{P}}$ and X and the continuity of X, (1.11) follows by Itô's formula. The proof of Theorem 1 shows that $\int \frac{1}{\widetilde{Z}}\, dL^{H,\widetilde{P}} = L$ is in $\mathcal{M}_0^2(\widetilde{R})$. By (1.11) and the definition of $\widetilde{R}$, we therefore have

$$R_0^* = \widetilde{Z}_0 E_{\widetilde{R}}[L_T^2] = \widetilde{Z}_0 E_{\widetilde{R}}\left[\int_0^T \frac{1}{\widetilde{Z}_u^2}\, d[L^{H,\widetilde{P}}]_u\right] = E\left[\widetilde{Z}_T Z_T^{\widetilde{P}}\int_0^T \frac{1}{\widetilde{Z}_u^2}\, d[L^{H,\widetilde{P}}]_u\right],$$

where the last equality uses $\widetilde{Z}_T = Z_T^{\widetilde{P}}$ and the definition of $\widetilde{R}$. Since $\widetilde{Z}$ is a $\widetilde{P}$-martingale, $\widetilde{Z}Z^{\widetilde{P}}$ is a P-martingale and so (1.12) follows from Theorem VI.57 of Dellacherie/Meyer (1982).

<div align="right">**q.e.d.**</div>

Proof of Lemma 3. Since X is continuous and $I\!\!P \neq \emptyset$, X satisfies (SC); hence 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 final assertion follows because

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

q.e.d.

Proof of Lemma 4. Because $\left(\int B_u^2\,du\right)_T^* \le T(B_T^*)^2$, the estimate (3.3) is straightforward. As a stochastic exponential, $L(\alpha)$ is a continuous local martingale starting at 1 with

(5.4) $$\langle L(\alpha)\rangle_t = \alpha^2 \int_0^t L_u^2(\alpha)B_u^2\,du.$$

In the following estimates, c denotes a generic constant that may vary from line to line. From the Burkholder-Davis-Gundy inequality, (5.4) and Hölder's inequality, we obtain a constant depending only on p, α such that

$$E\left[|L_T^*(\alpha)|^p\right] \le c\left(1 + E\left[\left|(L(\alpha)-1)_T^*\right|^p\right]\right)$$

$$\le c\left(1 + E\left[\left(\langle L(\alpha)\rangle_T\right)^{\frac{p}{2}}\right]\right)$$

$$\le c\left(1 + T^{\frac{p}{2}-1}E\left[\int_0^T L_u^p(\alpha)|B_u|^p\,du\right]\right).$$

Since B is independent of W, the conditional distribution of $L_u(\alpha)$ given B is lognormal and this implies that

$$E\left[L_u^p(\alpha)\big|B\right] = E\left[\exp\left(p\alpha\int_0^u B_s\,dW_s - \tfrac{1}{2}p\alpha^2\int_0^u B_s^2\,ds\right)\bigg|B\right]$$

$$= \exp\left(\left(-\frac{1}{2}p\alpha^2 + \frac{1}{2}(p\alpha)^2\right)\int_0^u B_s^2\,ds\right)$$

$$\le \exp\left(cT(B_T^*)^2\right)$$

holds uniformly in $u \in [0,T]$. Hence Fubini's theorem and conditioning on B yield

$$E\left[|L_T^*(\alpha)|^p\right] \le c\left(1 + T^{\frac{p}{2}-1}\int_0^T E\left[E\left[L_u^p(\alpha)|B\right]|B_u|^p\right]du\right)$$

$$\le c\left(1 + T^{\frac{p}{2}-1}E\left[\exp\left(cT(B_T^*)^2\right)\int_0^T |B_u|^p\,du\right]\right).$$

Because

$$\int_0^T |B_u|^p\,du \le T|B_T^*|^p = \left(T^{\frac{2}{p}}(B_T^*)^2\right)^{\frac{p}{2}} \le \exp\left(\frac{p}{2}T^{\frac{2}{p}}(B_T^*)^2\right)$$

32

by the trivial estimate $x \le e^x$, we obtain (3.4) by putting everything together.

q.e.d.

Proof of Corollary 5. From the SDEs for X and $\widehat{Z}$ and the choice of μ_2, we obtain

$$X_t = X_0 \mathcal{E}\left(\int Y \, dW\right)_t \mathcal{E}\left(\int \gamma Y_u^2 \, du\right)_t \qquad , \qquad 0 \le t \le T$$

and

$$\widehat{Z}_t = \mathcal{E}\left(-\int \gamma Y \, dW\right)_t \qquad , \qquad 0 \le t \le T.$$

Since $\varrho = 0$, W and W' are independent and so Y is also independent of W. Moreover, Y is adapted and continuous, hence predictable and P-a.s. bounded and therefore (3.5) and (3.6) follow from Lemma 4 and Hölder's inequality if we can show that

(5.5) $$E\left[\exp\left(\alpha(Y_T^*)^2\right)\right] < \infty \qquad \text{for } \alpha > 0 \text{ sufficiently small.}$$

In model S2, this follows immediately from Theorem 3.2 of Adler (1990) because Y as an Ornstein-Uhlenbeck process is a deterministic translate of a centered continuous Gaussian process. In model H2, the process $F = Y^2$ satisfies the SDE

$$dF_t = \kappa(\theta - F_t) \, dt + \Sigma \sqrt{F_t} \, dW_t'$$

with $\kappa > 0$. Choose now $m \in I\!\!N$ with $\frac{1}{4}\Sigma^2 m > \kappa\theta$ and define U as the solution of the SDE

$$dU_t = \frac{1}{4}\Sigma^2 m \, dt + \Sigma\sqrt{U_t} \, dW_t' \qquad , \qquad U_0 = F_0.$$

Since $\frac{1}{4}\Sigma^2 m > \kappa\theta \ge \kappa(\theta - x)$ for all $x \ge 0$, the comparison theorem for SDEs implies that $P\left[U_t \ge F_t, 0 \le t \le T\right] = 1$; see Theorem VI.1.1 of Ikeda/Watanabe (1989). But $\bar{U} := \frac{4U}{\Sigma^2}$ satisfies the SDE

$$d\bar{U}_t = 2\sqrt{\bar{U}_t} \, dW_t' + m \, dt$$

and is therefore the square of a Bessel diffusion with index m; see Example IV.8.3 of Ikeda/Watanabe (1989). Thus $\bar{U}$ has the same law as the squared norm of an m-dimensional Brownian motion $\bar{W} = (\bar{W}^1 \dots \bar{W}^m)^{\text{tr}}$ starting from $\frac{4F_0}{\Sigma^2}$ and so we obtain

$$(Y_T^*)^2 \le U_T^* \stackrel{\text{(d)}}{=} \frac{1}{4}\Sigma^2 \left(\bar{W}_T^*\right)^2$$

where $\stackrel{\text{(d)}}{=}$ denotes equality in distribution. Since $\bar{W}^1, \dots, \bar{W}^m$ are independent, we get

$$E\left[\exp\left(\alpha(Y_T^*)^2\right)\right] \le E\left[\exp\left(\frac{\alpha}{4}\Sigma^2 m \left((\bar{W}^1)_T^*\right)^2\right)\right]$$

33

and the right-hand side is again finite for α sufficiently small by Theorem 3.2 of Adler (1990). This proves (5.5), hence (3.5) and (3.6).

Taking $p = 2$ in (3.6) implies that $\widehat{Z}$ is a square-integrable P-martingale and thus yields $\widehat{P} \in I\!\!P_e^2$, hence (3.7). Combining this with (3.5) for $p = 2$ immediately gives (3.8) because X is then a local $\widehat{P}$-martingale whose supremum is in $L^1(\widehat{P})$ by Hölder's inequality. Finally, (3.9) follows if we show that $[M]_T$ and $|A|_T^2 = \left(\int\limits_0^T |dA_u| \right)^2$ are both in $L^1(P)$. But since

$$|A|_T^2 = \left(\int\limits_0^T |X_u\, \mu_2(u, Y_u)|\, du \right)^2 = \left(\int\limits_0^T X_u \gamma Y_u^2\, du \right)^2 \leq \gamma^2 T \int\limits_0^T X_u^2 Y_u^4\, du,$$

the same techniques as in the proof of Lemma 4 allow us to bound $E\left[|A|_T^2 \right]$ by an exponential moment of $(Y_T^*)^2$ which is finite for small T by (5.5). Then A_T^* is in $L^2(P)$, hence by (3.5) also M_T^* and so $[M]_T \in L^1(P)$ by the Burkholder-Davis-Gundy inequality.

<div align="right">

q.e.d.

</div>

References

R. J. Adler (1990), "An Introduction to Continuity, Extrema and Related Topics for General Gaussian Processes", *IMS Lecture Notes Monograph Series, Vol. 12, Hayward, California*

D. Bertsimas, L. Kogan and A. Lo (1999), "Hedging Derivative Securities and Incomplete Markets: An ε-Arbitrage Approach", *LFE working paper No. 1027-99R, Sloan School of Management, MIT, Cambridge MA; to appear in Operations Research*

F. Biagini, P. Guasoni and M. Pratelli (2000), "Mean-Variance Hedging for Stochastic Volatility Models", *Mathematical Finance 10*, 109–123

F. Delbaen and W. Schachermayer (1996), "The Variance-Optimal Martingale Measure for Continuous Processes", *BERNOULLI 2*, 81–105; Amendments and corrections (1996), *BERNOULLI 2*, 379–380

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

C. A. J. Fletcher (1988), "Computational Techniques for Fluid Dynamics", *Springer*

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

R. Frey (1997), "Derivative Asset Analysis in Models with Level-Dependent and Stochastic Volatility", *CWI Quarterly 10/1*, 1–34

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

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

D. Heath, E. Platen and M. Schweizer (1999), "Numerical Comparison of Local Risk-Minimisation and Mean-Variance Hedging", *preprint FMRR No. 99.001, Australian National University, Canberra; to appear in: E. Jouini, J. Cvitanić, M. Musiela (eds.), Advances in Mathematical Finance, Cambridge University Press*

D. Heath and M. Schweizer (2000), "Martingales versus PDEs in Finance: An Equivalence Result with Examples", *to appear in Journal of Applied Probability*

S. L. Heston (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", *Review of Financial Studies 6*, 327–343

C. Hipp (1996), "Hedging and Insurance Risk", *preprint 1/96, University of Karlsruhe*

C. Hipp (1998), "Hedging General Claims in Diffusion Models", *preprint, University of Karlsruhe*

J. D. Hoffman (1993), "Numerical Methods for Engineers and Scientists", *McGraw-Hill*

N. Ikeda and S. Watanabe (1989), "Stochastic Differential Equations and Diffusion Processes", *North-Holland*

P. E. Kloeden and E. Platen (1992), "Numerical Solution of Stochastic Differential Equations", *Springer*

J. P. Laurent and H. Pham (1999), "Dynamic Programming and Mean-Variance Hedging", *Finance and Stochastics 3*, 83–110

B. Leblanc (1996), "Une Approche Unifiée pour une Forme Exacte du Prix d'une Option dans les Différents Modèles à Volatilité Stochastique", *Stochastics and Stochastics Reports 57*, 1–35

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

H. Pham, T. Rheinländer and M. Schweizer (1998), "Mean-Variance Hedging for Continuous Processes: New Proofs and Examples", *Finance and Stochastics 2*, 173–198

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 (1991), "Option Hedging for Semimartingales", *Stochastic Processes and their Applications 37*, 339–363

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

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

M. Schweizer (1999), "A Guided Tour through Quadratic Hedging Approaches", *SFB 373 Discussion Paper 96, Humboldt University, Berlin; to appear in: E. Jouini, J. Cvitanić, M. Musiela (eds.), Advances in Mathematical Finance, Cambridge University Press*

C. A. Sin (1996), "Strictly Local Martingales and Hedge Ratios on Stochastic Volatility Models", *PhD thesis, Cornell University*

C. A. Sin (1998), "Complications with Stochastic Volatility Models", *Advances in Applied Probability 30*, 256–268

E. M. Stein and J. C. Stein (1991), "Stock Price Distributions with Stochastic Volatility: An Analytic Approach", *Review of Financial Studies 4*, 727–752