GAU
შპს ქართულ-ამერიკული უნივერსიტეტი
GEORGIAN AMERICAN UNIVERSITY, LLC

Stochastic Analysis: Applications to Statistics and Finance

Edited by T. Toronjadze

Stochastic Analysis: Applications to Statistics and Finance

Edited by T. Toronjadze

Georgian-American University (GAU) Business School
Business Research Center

Stochastic Analysis: Applications to Statistics and Finance

Edited by T. Toronjadze

Technical Editor: M. Kvinikadze

Tbilisi, Georgia, 2018

Contents

Preface

Stochastic analysis is important and powerful tool to study uncertainty in numerous problems of statistics and modern finance theory. This book presents a selected collection of some of the research of GAU Business Research Center during the last 10 years in this important area. All these researches were published in international impact-factor journals.

First chapter by Lazrieva and Toronjadze presents the optimal B-robust estimate for multidimensional parameter in drift coefficient of diffusion type process with small noise. Chapter 2 by Mania et al. presents the mean-variance hedging problem under partial information. Chapter 3 by Tevzadze is devoted to the proof of the existence of unique solution of general backward stochastic differential equation with quadratic grows driven by martingales. In chapter 4, Mania et al. discuss the mean-variance hedging problem in case where flow of observable events do not contain the full information on the underlying asset price process. In chapter 5, Lazrieva and Toronjadze present general results concerning the asymptotic behavior of the Polyak average of the solution of the Robbins-Monro type stochastic differential equation. In chapter 6, Lazrieva and Toronjadze present the recourse estimation problem of a one-dimensional parameter in the trend coefficient of diffusion process. In chapter 7, Tevzadze et al. study the robust maximization of terminal wealth utility in diffusion financial market models. In chapter 8, Chikvinidze and Mania give new proofs of some well known results of BMO martingales and improves some estimate of BMO norms. In chapter 9, Lazrieva and Toronjadze study the recursive estimation problem of one-dimensional parameter of statistical models associated with semimartingales. Chapter 10 by Mania and Tevzadze studies regularity properties of the dynamic value function of the primal and dual problems of optimal investing for utility functions defined only whole real line. Chapter 11 by Mania and Tevzadze establishes connections between system of Forward and Backward SDEs and backward stochastic PDEs related to the utility maximization problem.

Editor T. Toronjadze

GAU Business School

OPTIMAL ROBUST MEAN-VARIANCE HEDGING IN INCOMPLETE FINANCIAL MARKETS

N. LAZRIEVA AND T. TORONJADZE

Abstract. Optimal B-robust estimate is constructed for multidimensional parameter in drift coefficient of diffusion type process with small noise. Optimal mean-variance robust (optimal V-robust) trading strategy is find to hedge in mean-variance sense the contingent claim in incomplete financial market with arbitrary information structure and misspecified volatility of asset price, which is modelled by multidimensional continuous semimartingale. Obtained results are applied to stochastic volatility model, where the model of latent volatility process contains unknown multidimensional parameter in drift coefficient and small parameter in diffusion term.

Key words and phrases: Stochastic volatility, small diffusion, robust parameter estimate, optimal mean-variance robust hedging

MSC 2010: 60G22, 62F35, 91B28, 62F35, 62M05, 62M09

1. Introduction, Motivation and Results

The hedging and pricing of contingent claims in incomplete financial markets, and dynamic portfolio selection problems are important issues in modern theory of finance. These problems are associated due to the so-called mean-variance approach.

For determining a "good" hedging strategy in incomplete market with arbitrary information structure $F = (\mathcal{F})_{0\le t\le T}$, one riskless asset and d, $d \ge 1$, risky assets, whose price process is a semimartingale X, the mean-variance approach suggests to use the quadratic criterion to measure the hedging error, i.e. to solve the mean-variance hedging problem introduced by Föllmer and Sondermann [10]:

$$\text{minimize} \quad E\left(H - x - \int_0^T \theta_t dX_t\right)^2 \quad \text{over all} \quad \theta \in \Theta, \tag{1.1}$$

where contingent claim H is a $\mathcal{F}_T$-measurable square-integrable random variable (r.v.), x is an initial investment, Θ is a class of admissible trading strategies, T is an investment horizon.

The mean-variance formulation by Markowitz [26], provides a foundation for a single period portfolio selection (see, also Merton [27]). In recent paper of Li and Ng [22] the concept of Markowitz's mean-variance formulation for finding the optimal portfolio policy and determining the efficient frontier in analytical form has been extended to multiperiod portfolio selection.

Published in *J. Math. Sci. (N.Y.)* **153** (2008), no. 3, 262–290.

As it pointed out in Li and Ng [22] the results on multiperiod mean-variance formulation with one riskless asset can be derived using the results of the mean-variance hedging formulation.

Therefore, the mean-variance hedging is s powerful approach for both above mentioned major problems.

The problem (1.1) was intensively investigated in last decade (see, e.g., Dufiie and Richardson [9], Schwezer [36], [37], [38], Delbaen et al. [8], Monat and Striker [28], Rheinländer and Schweizer [33], (RSch hereafter), Pham et al. [31], Gourieroux et al. [11] (GLP hereafter), Laurent and Pham [18]).

A stochastic volatility model, proposed by Hull and White [13] and Scott [39], where the stock price volatility is an random process, is a popular model of incomplete market, where the mean-variance hedging approach can be used (see, e.g., Laurent and Pham [18], Biagini et al. [13], Mania and Tevzadze [24], Pham et al. [31]).

Consider the stochastic volatility model described by the following system of SDE

$$\begin{gathered} dX_t = X_t\, dR_t, \quad X_0 > 0, \\ dR_t = \mu_t(R_t, Y_t)\, dt + \sigma_. dw_t^R, \quad R_0 = 0, \\ \sigma_t^2 = f(Y_t), \\ dY_t = a(t, Y_t; \alpha)\, dt + \varepsilon\, dw_t^\sigma, \quad Y_0 = 0, \end{gathered} \tag{1.2}$$

where $w = (w^R, w^\sigma)$ is a standard two-dimensional Wiener process, defined on complete probability space $(\Omega, \mathcal{F}, P)$, $F^w = (\mathcal{F}_t^w)_{0 \le t \le T}$ is the P-augmentation of the natural filtration $\mathcal{F}_t^w = \sigma(w_s, 0 \le s \le t)$, $0 \le t \le T$, generated by w, $f(\cdot)$ is a continuous one-to-one positive locally bounded function (e.g., $f(x) = e^x$), $\alpha = (\alpha_1, \dots, \alpha_m)$, $m \ge 1$, is a vector of unknown parameters, and ε, $0 < \varepsilon \ll 1$, is a small number. Assume that the system (1.2) has an unique strong solution.

This model is analogous to the model proposed by Renault and Touzi [32] (RT hereafter). The principal difference is the presence of small parameter ε in our model, which due to the assumption that the volatility of randomly fluctuated volatility process is small (see, also Sircar and Papanicolau [40]). Thus assumption enables us to use the prices of trading options with short, nearest to the current time value maturities for volatility process filtration and parameter estimation purposes (see below). In contrast, RT [32] needs to assume that there exist trading derivatives with any (up to the infinity) maturities.

Important feature of the stochastic volatility models is that volatility process Y is unobservable (latent) process. To obtain explicit form of optimal trading strategy full knowledge of the model of the process Y is necessary and hence one needs to estimate the unknown parameter $\alpha = (\alpha_1, \dots, \alpha_m)$, $m \ge 1$.

A variety of estimation procedures are used, which involve either direct statistical analysis of the historical data or the use of implied volatilities extracted from prices of existing traded derivatives.

For example, one can use the following method based on historical data.

Fix the time variable t. From observations $X_{t_0^{(n)}}, \dots, X_{t_n^{(n)}}$, $0 = t_0^{(n)} < \dots < t_n^{(n)} = t$, $\max_j [t_{j+1}^{(n)} - t_j^{(n)}] \to 0$, as $n \to 0$, calculate the realization of yield process $R_t = \int_0^t \frac{dX_s}{X_s}$, and then calculate the sum

$$S_n(t) = \sum_{j=0}^{n-1} |R_{t_{j+1}^{(n)}} - R_{t_j^{(n)}}|^2.$$

It is well-known (see, e.g., Lipster and Shiryaev [23]) that

$$S_n(t) \xrightarrow{P} \int_0^t \sigma_s^2\, ds \quad \text{as} \quad n \to \infty.$$

Since $\sigma_t^2(\omega) = f(Y_t)$ is a continuous process we get

$$\sigma_t^2(\omega) = \lim_{\Delta \downarrow 0} \frac{F(t+\Delta,\omega) - F(t,\omega)}{\Delta},$$

where $F(t,\omega) = \int\limits_0^t \sigma_s^2(\omega) ds$.

Hence, the realization $(y_t)_{0\le t\le T}$ of the process Y can be found by the formula $y_t = f^{-1}(\sigma_t^2)$, $0 \le t \le T$.

More sofisticated methods using the same idea can be found, e.g., in Chesney et al. [5], Pastorello [30].

We can use the reconstructed sample path (y_t), $0 \le t \le T.$ to estimate the unknown parameter α in the drift coefficient of diffusion process Y.

The second, market price adjusted procedure of reconstruction the sample path of volatility process Y and parameter estimate was suggested by RT [32], where they used implied volatility data.

We present a quick review of this method, adapted to our model (1.2).

Suppose that the volatility risk premium $\lambda^\sigma \equiv 0$, meaning that the risk from the volatility process is non-compensated (or can be diversified away). Then the price $C_t(\sigma)$ of European call option can be calculated by the Hull and White formula (see, e.g., RT [32]), and Black-Scholes (BS) implied volatility $\sigma^i(\sigma)$ can be found as an unique solution of the equation

$$C_t(\sigma) = C_t^{BS}(\sigma^i(\sigma)),$$

where $C^{BS}(\sigma)$ denotes the standard BS formula written as a function of the volatility parameter σ.

Here (for further estimational purposes) only at-the-money options are used.

Under some technical assumptions (see Proposition 5.1 of RT [32], and Bujeux and Rochet [23] for general diffusion of volatility process)

$$\frac{\partial \sigma_t^i(\sigma,\alpha)}{\partial \sigma_t} > 0 \tag{1.3}$$

(remember that the drift coefficient of process Y depends on unknown parameter α).

Fix current value of time parameter t, $0 \le t \le T$, and let $0 < T_1 < T_2 < \cdots < T_{k-1} < t < T_k$ be the maturity times of some traded at-the-money options.

Let $\sigma_{t_j^\varepsilon}^{i^*}$ be the observations of an implied volatility at the time moments $0 = t_0^\varepsilon < t_1^\varepsilon < \cdots < t_{[\frac{t}{\varepsilon}]} = t$, $\max\limits_j [t_{j+1}^\varepsilon - t_j^\varepsilon] \to 0$, as $\varepsilon \to 0$.

Then, using (1.3), and solving the equation

$$\sigma_{t_j^\varepsilon}^i(\sigma_{t_j^\varepsilon}, \alpha) = \sigma_{t_j^\varepsilon}^{i^*},$$

one can obtained the realization $\{\widetilde{\sigma}_{t_j^\varepsilon}\}$ of the volatility (σ_t), and thus, using the formula $y_{t_j^\varepsilon} = f^{-1}(\widetilde{\sigma}^2_{t_j^\varepsilon})$, the realization $\{y_{t_j^\varepsilon}\}$ of volatility process (Y_t), which can be viewed as the realization of nonlinear AR(1) process:

$$Y_{t_{j+1}^\varepsilon} - Y_{t_j^\varepsilon} = a(t_j^\varepsilon, Y_{t_j^\varepsilon}; \alpha)(t_{j+1}^\varepsilon - t_j^\varepsilon) + \varepsilon(w^\sigma_{t_{j+1}^\varepsilon} - w^\sigma_{t_j^\varepsilon}).$$

Using the data $\{y_{t_j^\varepsilon}\}$ one can construct the MLE $\widehat{\alpha}_t^\varepsilon$ of parameter α, see, e.g., Chitashvili et al. [25], [26], Lazrieva and Toronjadze [19].

Remember the scheme of construction of MLE. Rewrite the previous AR(1) process, using obvious simple notation, in form

$$Y_{j+1} - Y_j = a(t_j, Y_j; \alpha)\Delta + \varepsilon \Delta w_j^\sigma.$$

Then

$$\frac{\partial}{\partial y} P\{Y_{j+1} \le y \mid Y_j\} = \frac{1}{\sqrt{2\pi\Delta\varepsilon}} \exp\left(-\frac{(y - Y_j - a(t_j, Y_j; \alpha)\Delta)^2}{2\varepsilon^2\Delta}\right) =: \varphi_{j+1}(y, Y_j; \alpha),$$

and the log-derivative of the likelihood process $\ell_t = (\ell_t^{(1)}, \dots, \ell_t^{(m)})$ is given by the relation

$$\ell_t^{(i)} = \sum_j \ell_{j+1}^{(i)}, \quad i = \overline{1, m},$$

where

$$\ell_{j+1}^{(i)}(y; \alpha) = \frac{\partial}{\partial \alpha_i} \ln \varphi_{j+1}(y, Y_j; \alpha) = \frac{1}{\varepsilon^2\Delta}(y - Y_j - a(t_j, Y_j; \alpha)\Delta)\dot{a}^{(i)}(t_j, Y_j; \alpha)\Delta.$$

Hence MLE is a solution (under some conditions) of the system of equations

$$\frac{1}{\varepsilon^2\Delta} \sum_j (y_{j+1} - y_j - a(t_j, y_j; \alpha)\Delta)\dot{a}^{(i)}(t_j, y_j; \alpha)\Delta = 0, \quad i = \overline{1, m},$$

where the reconstructed data $\{y_j\} = \{y_{t_j^\varepsilon}\}$ are substituted).

Following RT [32] let us introduce the functionals

$$HW_\varepsilon^{-1} : \widehat{\alpha}_t^\varepsilon(p) \to \left(y_{t_j^\varepsilon}^{(p+1)},\ 0 \le j \le \left[\frac{t}{\varepsilon}\right]\right),$$

$$MLE_\varepsilon : \left(y_{t_j^\varepsilon}^{(p+1)},\ 0 \le j \le \left[\frac{t}{\varepsilon}\right]\right) \to \widehat{\alpha}_t^\varepsilon(p+1)$$

and

$$\phi_\varepsilon = MLE_\varepsilon \circ HW_\varepsilon^{-1}.$$

Starting with some constant initial value (or preliminary estimate obtained, e.g., from historical data) one can compute a sequence of estimates

$$\widehat{\alpha}_t^\varepsilon(p+1) = \phi_\varepsilon(\widetilde{\alpha}_t^\varepsilon(p)), \quad p \ge 1.$$

If the operator ϕ_ε is a strong contraction in the neighborhood of the true value of the parameter α^0, for a small enough ε, then one can define the estimate $\widehat{\alpha}_t^\varepsilon$ as the limits of the sequence $\{\widehat{\alpha}_t^\varepsilon(p)\}_{p \ge 1}$. It was proved in RT [32] that $\widehat{\alpha}_t^\varepsilon$ is a strong consistent estimate of the parameter α.

Return to our consideration.

Interpolating on some way the corresponding (to the estimate $\widehat{\alpha}_t^\varepsilon$) realization $\{y_{t_j^\varepsilon}\}$ we get the reconstructed continuous sample path $(y_s)_{0\le s\le t}$ of the latent process Y, which can be used for further analysis.

Unfortunately, both described statistical procedures are highly sensitive w.r.t errors in all steps of parameter identification process.

Hence, this is a natural place for introducing the robust procedure of parameter estimates.

Suppose that the sample path $(y_s)_{0\le s\le t}$ comes from the observation of process $(\widetilde{Y}_s)_{0\le s\le t}$ with distribution $\widetilde{P}_\alpha^\varepsilon$ from the shrinking contamination neighborhood of the distribution P_α^ε of the basic process $Y=(Y_s)_{0\le s\le t}$. That is

$$\frac{d\widetilde{P}_\alpha^\varepsilon}{dP_\alpha^\varepsilon}\Big|\,\mathcal{F}_t^w=\mathcal{E}_t(\varepsilon N^\varepsilon), \tag{1.4}$$

where $N^\varepsilon=(N_s^\varepsilon)_{0\le s\le t}$ is a P_α^ε-square integrable martingale, $\mathcal{E}_t(M)$ is the Dolean exponential of martingale M.

In the diffusion-type framework (1.4) represents the Huber gross error model (as it explain in Remark 2.2). The model of type (1.4) of contamination of measures for statistical models with filtration was suggested by Lazrieva and Toronjadze [20], [21].

In Section 2 we study the problem of construction of robust estimates for contamination model (1.4).

In subsection 2.1 we give a description of the basic model and definition of consistent uniformly linear asymptotically normal (CULAN) estimates, connected with the basic model (Definition 2.1).

In subsection 2.2 we introduce a notion of shrinking contamination neighborhood, described in terms of contamination of nominal distribution, which naturally leads to the class of alternative measures (see (2.18) and (2.19)).

In subsection 2.3 we study the asymptotic behaviour of CULAN estimates under alternative measures (Proposition 2.2), which is the basis for the formulation of the optimization problem.

In subsection 2.4 the optimization problem is solved which leads to construction of optimal B-robust estimate (Theorem 2.1).

Based on the limit theorem (subsection 2.1), one can construct the asymptotic confidence region of level γ for unknown parameter α

$$\lim_{\varepsilon\to\infty}P_\alpha^\varepsilon\left(\varepsilon^{-2}(\alpha-\alpha_t^{*,\varepsilon})'V^{-1}(\psi^*;\alpha_t^{*,\varepsilon})(\alpha-\alpha_t^{*,\varepsilon})\le\chi_\gamma^2\right)=1-\gamma,$$

where χ_γ^2 is a quantile of order $1-\gamma$ of χ^2-distribution with m degree of freedom, and $V(\psi^*;\alpha)$ is given by (2.17).

This region shrinks to the estimate $\alpha_t^{*,0}$, as $\varepsilon\to 0$.

Now if the coefficient $a(t,y;\alpha)$ in (1.2) is such that the solution $Y_t^\varepsilon(\alpha)$ of SDE (1.2) is continuous w.r.t parameter α (see, e.g., Krylov [16]), then the confidence region of parameter α is mapped to the confidence interval for $Y_t^\varepsilon(\alpha)$, which shrinks to $Y_t^*=Y_t^0(\alpha_t^{*,0})$, Further, by the function f, the latter interval is mapped to the confidence interval for σ_t, which shrinks to $\sigma_t^*=f^{1/2}(Y_t^0(\alpha_t^{*,0}))$. Denote σ_t^0 the center of this interval. Then the interval can be written in the form

$$\sigma_t=\sigma_t^0+\delta(\varepsilon)h_t,$$

where $\delta(\varepsilon) \to 0$, as $\varepsilon \to 0$, and $h \in \mathcal{H}$ (see (3.18)).

Thus, we arrive at the asset price model (1.2) with misspecified volatility, and it is natural to consider the problem of construction of the robust trading strategy to hedge a contingent claim H.

We investigate this problem in the mean-variance setting in Section 3. We consider the general situation, when the asset price is modelled by d-dimensional continuous semimartingale and the information structure is given by some general filtration.

In subsection 3.1 we give a description of the financial market model.

In subsection 3.2 we collect the facts concerning the variance-optimal equivalent local martingale measure, which plays a key role in the mean-variance hedging approach.

In last subsection 3.3 we construct "optimal robust hedging strategy" (Theorem 3.1) by approximating the optimization problem (3.25) by the problem (3.27). As it is mentioned in Remark 3.2, such approach and term are common in robust statistic theory. In contact to optimal B-robustness (see Section 2), here we develop the approach, known in robust statistics as optimal V-robustness, see Hampel et al. [12].

Note that our approach allows incorporating current information on the underlying model, and hence is adaptive. Namely, passing from time value t to $t+\tau$, $\tau > 0$, when more information about market prices are available, the asymptotic variance-covariance of the constructed estimate $\alpha_t^{*,\varepsilon}$ becomes smaller, and hence the estimation procedure becomes more precise.

In the paper of Runggaldier and Zaccaria [35] the adaptive approach to risk management under general uncertainty (restricted information) was developed. As it is mentioned in this paper there exist a series of investigations dealt with various type of adaptive approaches (see list of references in [35]). But in all these papers (except Runggaldier and Zaccaria [35]) the uncertainty is only in the stock appreciation rate in contrast to our consideration, where the model misspecification is due to the volatility parameter.

The consideration of misspecified asset price models was initiated by Avellaneda et al. [1], Avellaneda and Paras [2].

Various authors in different settings attacked the robustness problem. The method used in Section 3 was suggested by Toronjadze [41] for asset price process modelled by the one-dimensional process. As it will be shown in Remark 3.2 below, in simplest case when asset price process is a martingale w.r.t initial measure P, and it is possible to find the solution of "exact" optimization problem (3.25), this solution coincides with the solution of an approximating optimization problem (3.27). In more general situation (when asset price process is not more the P-martingale) investigation of the problem (3.25) by, e.g., control theory methods seems sufficiently difficult. Anyway, we do not know the solution of the problem (3.25).

Return to the stochastic volatility model (1.2) and describe successive steps of our approach:

1) For each current time value t, $0 < t < T$, reconstruct the sample path $(y_s)_{0\le s\le t}$, using the historical data or the tradable derivatives prices;

2) Using the approach developed in Section 2, calculate the value $\alpha_t^{*,\varepsilon}$ of the robust estimate of parameter α (i.e. construct the deterministic function $t \to \alpha_t^{*,\varepsilon} \in R^m$) and then find the confidence region for parameter α;

3) Based on the volatility process model find the confidence interval for $Y_t(\alpha)$;

4) Denoting $a^*(t,y) = a(t,y;\alpha_t^{*,\varepsilon})$, where $a(t,y;\alpha)$ is a drift coefficient of volatility process, consider the stochastic volatility model with misspecified asset price model and fully specified volatility process model

$$dX_t = X_t\,dR_t, \quad X_0 > 0,$$
$$dR_t = (\sigma_t^0 + \delta(\varepsilon)h_t)dM_t^0, \quad R_0 = 0,$$
$$dY_t = a^*(t,Y_t)\,dt + \varepsilon\,dw_t^\sigma, \quad Y_0 = 0, \quad 0 \le t \le T,$$

where

$$dM_t^0 = k_t\,dt + dw_t^R,$$

$h \in \mathcal{H}$ and σ_t^0 is the center of the confidence interval of volatility.

Using Theorem 3.1 construct the optimal robust hedging strategy by the formula (3.44),

$$\theta_t^* = \frac{1}{\sigma_t^0}\left[\psi_t^{1,H} + \zeta_t(V_t^* - (\psi_t^H)'U_t\right],$$

where all objects are defined in Theorem 3.1. □

It should be mentioned that if one constructs a hedging strategy $\widetilde{\theta}_t^*$ by the above-given formula with $\sigma_t^{*,\varepsilon} = f^{1/2}(Y_t^\varepsilon(\alpha_t^{*,\varepsilon}))$ instead of σ_t^0, then the strategies $\widetilde{\theta}_t^*$ and θ_t^* would be different, since $\sigma_t^{*,\varepsilon} \ne \sigma_t^0$, in general. Hence the value $\Delta_t = |\sigma_t^{*,\varepsilon} - \sigma_t^0|$ defines the correction term between the robust, θ_t^* and non-robust, $\widetilde{\theta}^*$ strategies.

In nontrivial case, when $k_t = k(Y_t)$ the variance-optimal martingale measure $\widetilde{P}$ is given by (3.17), $\zeta_t = -k_t\mathcal{E}_t(-k\cdot M^0)$ (see subsection 3.2), and the process $(X_t,Y_t)_{0\le t\le T}$ is the Markov process. If $H = h(X_T,Y_T)$ ($h(x,y)$ is some function), then $\widetilde{V}_t^H = E^{\widetilde{P}}(H|\mathcal{F}_t^w) = E^{\widetilde{P}}(h(X_T,Y_T)|\mathcal{F}_t^w) = v(t,X_T,Y_T)$ and if, e.g., $v(t,x,y) \in C^{1,2,2}$, then v is an unique solution of the following partial differential equation

$$\frac{\partial v}{\partial t} + a^*\frac{\partial v}{\partial y} + \frac{1}{2}\left(\varepsilon^2\frac{\partial^2 v}{\partial y^2} + x^2v^2\frac{\partial^2 v}{\partial x^2}\right) = 0,$$

with the boundary condition $v(T,x,t) = h(x,y)$. More general situation with nonsmooth v is considered in Laurent and Pham [18], Mania and Tevzadze [24].

Further, one can find the Galtchouck–Kunita–Watanabe decomposition of r.v. H (see, e.g., Pham et al. [31]) putting

$$\xi_t^H = \frac{\partial v(t,X_t,Y_t)}{\partial x}, \quad L_T^H = \varepsilon\int_0^T \frac{\partial v}{\partial y}(t,X_t,Y_t)\,dw_t^\sigma,$$

and calculate ψ_t^H, L_T and V_t^* using (4.13) and (4.14) of RSch [33].

Thus one get the explicit solution of the mean-variance hedging problem.

Finally, here is the short summary of approach:

a) Incorporate the robust procedure in statistical analysis of volatility process. That is construct and use in the model optimal B-robust estimate of unknown parameter in drift coefficient of volatility process.

Parameter estimation naturally leads to the asset price model misspecification.

b) Incorporate the second robust procedure in financial analysis of contingent claim hedging. That is construct and use for hedging purposes optimal V-robust trading strategy.

In our opinion this “double robust” strategy should be more attractive to protect the hedger against the possible errors.

The general asymptotic theory of estimation can be found in Ibragimov and Khas’miskii [14]; the theory of robust statistics is developed in Hampel et al. [12] and in Rieder [34]; the theory of the trend parameter estimates for diffusion process with small noise is developed in Kutoyants [17]; the book of Musiela and Rutkowsky [29] is devoted to the mathematical theory of finance and finally, the general theory of martingales can be found in Jacod and Shiryaev [15].

2. Optimal B-Robust Estimates

2.1. **Basic model. CULAN estimates.** The basic model of observations is described by the SDE

$$dY_s = a(s,Y;\alpha)\,ds + \varepsilon\,dw_s, \quad Y_0 = 0, \quad 0 \le s \le t, \tag{2.1}$$

where t is a fixed number, $w = (w_s)_{0\le s\le t}$ is a standard Wiener process defined on the filtered probability space $(\Omega,\mathcal{F},F = (\mathcal{F}_s)_{0\le s\le t},P)$ satisfying the usual conditions, $\alpha = (\alpha_1,\dots,\alpha_m)$, $m \ge 1$, is an unknown parameter to be estimated, $\alpha \in \mathcal{A} \subset R^m$, $\mathcal{A}$ is an open subset of R^m, ε, $0 < \varepsilon \ll 1$, is a small parameter (index of series). In our further considerations all limits correspond to $\varepsilon \to 0$.

Denote $(C_t,\mathcal{B}_t)$ a measurable space of continuous on $[0,t]$ functions $x = (x_s)_{0\le s\le t}$ with σ-algebra $\mathcal{B}_t = \sigma(x : x_s, s \le t)$. Put $\mathcal{B}_s = \sigma(x : x_u, u \le s)$.

Assume that for each $\alpha \in \mathcal{A}$ the drift coefficients $a(s,x;\alpha)$, $0 \le s \le t$, $x \in C_t$ is a known nonanticipative (i.e. $\mathcal{B}_s$-measurable for each s, $0 \le s \le t$) functional satisfying the functional Lipshitz and linear growth conditions **L**:

$$|a(s,x^1;\alpha) - a(s,x^2;\alpha)| \le L_1 \int_0^s |x_u^1 - x_u^2|\,dk_u + L_2|x_s^1 - x_s^2|,$$

$$|a(s,x;\alpha)| \le L_1 \int_0^s (1+|x_u|)\,dk_u + L_2(1+|x_s|),$$

where L_1 and L_2 are constants, which do not depend on α, $k = (k(s))_{0\le s\le t}$ is a non-decreasing right-continuous function, $0 \le k(s) \le k_0$, $0 : k_0 < \infty$, $x^1, x^2 \in C_t$.

Then, as it is well-known (see, e.g., Liptser and Shiryaev [23]), for each $\alpha \in \mathcal{A}$ the equation (2.1) has an unique strong solution $Y^\varepsilon(\alpha) = (Y_s^\varepsilon(\alpha))_{0\le s\le t}$, and in addition (see Kutoyants [17])

$$\sup_{0\le s\le t} |Y_s^\varepsilon(\alpha) - Y_s^0(\alpha)| \le C\varepsilon \sup_{0\le s\le t} |w_s| \quad P\text{-a.s.},$$

with some constant $C = C(L_1,L_2,k_0,t)$, where $Y^0(\alpha) = (Y_s^0(\alpha))_{0\le s\le t}$ is the solution of the following nonperturbated differential equation

$$dY_s = a(s,Y;\alpha)\,ds, \quad Y_0 = 0. \tag{2.2}$$

Change the initial problem of estimation of parameter α by the equivalent one, when the observations are modelled according to the following SDE

$$dX_s = a_\varepsilon(s,X;\alpha)\,ds + dw_s, \quad X_0 = 0, \tag{2.3}$$

where $a_\varepsilon(s,x;\alpha) = \frac{1}{\varepsilon}\,a(s,\varepsilon x;\alpha)$, $0 \le s \le t$, $x \in C_t$, $\alpha \in \mathcal{A}$.

It is clear that if $X^\varepsilon(\alpha) = (X^\varepsilon_s(\alpha))_{0\le s\le t}$ is the solution of SDE (2.3), then for each $s \in [0,t]$ $\varepsilon X^\varepsilon_s(\alpha) = Y^\varepsilon_s(\alpha)$.

Denote by P^ε_α the distribution of process $X^\varepsilon(\alpha)$ on the space $(C_t, \mathcal{B}_t)$, i.e. P^ε_α is the probability measure on $(C_t, \mathcal{B}_t)$ induced by the process $X^\varepsilon(\alpha)$. Let P^w be a Wiener measure on $(C_t, \mathcal{B}_t)$. Denote $X = (X_s)_{0\le s\le t}$ a coordinate process on $(C_t, \mathcal{B}_t)$, that is $X_s(x) = x_s$, $x \in C_t$.

The conditions $\mathbf{L}$ guarantee that for each $\alpha \in \mathcal{A}$ the measures P^ε_α and P^w are equivalent ($P^\varepsilon_\alpha \sim P^w$), and if we denote $z^{\alpha,\varepsilon}_s = \frac{dP^\varepsilon_\alpha}{dP^w}|\mathcal{B}_s$ the density process (likelihood ratio process), then

$$z^{\alpha,\varepsilon}_s(X) = \mathcal{E}_s(a_\varepsilon(\alpha)\cdot X) := \exp\left\{ \int_0^s a_\varepsilon(u, X; \alpha)\, dX_u - \frac{1}{2}\int_0^s a^2_\varepsilon(u, X; \alpha)\, du\right\}.$$

Introduce class Ψ of R^m-valued nonanticipative functionals ψ, $\psi : [0,t]\times C_t \times \mathcal{A} \to R^m$ such that for each $\alpha \in \mathcal{A}$ and $\varepsilon > 0$

$$1)\quad E^\varepsilon_\alpha \int_0^t |\psi(s, X; \alpha)|^2 ds < \infty, \tag{2.4}$$

$$2)\quad \int_0^t |\psi(s, Y^0(\alpha); \alpha)|^2 ds < \infty, \tag{2.5}$$

3) uniformly in α on each compact $K \subset \mathcal{A}$

$$P^\varepsilon_\alpha - \lim_{\varepsilon\to 0} \int_0^t |\psi(s, \varepsilon X; \alpha) - \psi(s, Y^0(\alpha); \alpha)|^2 ds = 0, \tag{2.6}$$

where $|\cdot|$ is an Euclidean norm in R^m, $P^\varepsilon_\alpha - \lim\limits_{\varepsilon\to 0} \zeta_\varepsilon = \zeta$ denotes the convergence $P^\varepsilon_\alpha\{|\zeta_\varepsilon - \zeta| > \rho\} \to 0$, as $\varepsilon \to 0$, for all ρ, $\rho > 0$.

Assume that for each $s \in [0,t]$ and $x \in C_t$ the functional $a(s, x; a)$ is differentiable in α and gradient $\dot{a} = \left(\frac{\partial}{\partial\alpha_1} a, \dots, \frac{\partial}{\partial\alpha_m} a\right)'$ belongs to Ψ ($\dot{a} \in \Psi$), where the sign "$'$" denoted a transposition.

Then the Fisher information process

$$I^\varepsilon_s(X; \alpha) := \int_0^s \dot{a}_\varepsilon(u, X; \alpha)[\dot{a}_\varepsilon(u, X; \alpha)]' du, \quad 0 \le s \le t,$$

is well-defined and, moreover, uniformly in α on each compact

$$P^\varepsilon_\alpha - \lim_{\varepsilon\to 0} \varepsilon^2 I^\varepsilon_t(\alpha) = I^0(\alpha), \tag{2.7}$$

where

$$I^0(\alpha) := \int_0^t \dot{a}(s, Y^0(\alpha); \alpha)[\dot{a}(s, Y^0(\alpha); \alpha]' ds.$$

For each $\psi \in \Psi$, introduce the functional $\psi_\varepsilon(s, x; \alpha) := \frac{1}{\varepsilon}\,\psi(s, \varepsilon x; \alpha)$ and matrices $\Gamma^\psi_\varepsilon(\alpha)$ and $\gamma^\psi_\varepsilon \alpha$:

$$\Gamma^\psi_\varepsilon(X; \alpha) := \int_0^t \psi_\varepsilon(s, X; \alpha)[\psi_\varepsilon(s, X; \alpha)]' ds, \tag{2.8}$$

$$\gamma^\psi_\varepsilon(X; \alpha) := \int_0^t \psi_\varepsilon(s, X; \alpha)[\dot{a}_\varepsilon(s, X; \alpha)]' ds. \tag{2.9}$$

Then from (2.6) it follows that uniformly in α on each compact

$$P_\alpha^\varepsilon - \lim_{\varepsilon\to 0} \varepsilon^2 \Gamma_\varepsilon^\psi(\alpha) = \Gamma_0^\psi(\alpha), \tag{2.10}$$

$$P_\alpha^\varepsilon - \lim_{\varepsilon\to 0} \varepsilon^2 \gamma_\varepsilon^\psi(\alpha) = \gamma_0^\psi(\alpha), \tag{2.11}$$

where the matrices $\Gamma_0^\psi(\alpha)$ and $\gamma_0^\psi(\alpha)$ are defined as follows

$$\Gamma_0^\psi(\alpha) := \int_0^t \psi(s, Y^0(\alpha); \alpha)[\psi(s, Y^0(\alpha); \alpha)]' ds, \tag{2.12}$$

$$\gamma_0^\psi(\alpha) := \int_0^t \psi(s, Y^0(\alpha); \alpha)[\dot{a}(s, Y^0(\alpha); \alpha)]' ds. \tag{2.13}$$

Note that,by virtue of (2.4), (2.5) and $\dot{a} \in \Psi$, matrices given by (2.8), (2.9), (2.12) and (2.13) are well-defined.

Denote Ψ_0 the subset of Ψ such that for each $\psi \in \Psi_0$ and $\alpha \in \mathcal{A}$, $\operatorname{rank} \Gamma_0^\psi(\alpha) = m$ and $\operatorname{rank} \gamma_0^\psi(\alpha) = m$.

Assume that $\dot{a} \in \Psi_0$.

For each $\psi \in \Psi_0$, define a P_α^ε-square integrable martingale $L^{\psi,\varepsilon}(\alpha) = (L_s^{\psi,\varepsilon}(\alpha))_{0\le s\le t}$ as follows

$$L_s^{\psi,\varepsilon}(X; \alpha) = \int_0^s \psi_\varepsilon(u, X; \alpha)(dX_u - \alpha_\varepsilon(u, X; \alpha)\, du). \tag{2.14}$$

Now we give a definition of CULAN M-estimates.

Definition 2.1. An estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0} = (\alpha_{1,t}^{\psi,\varepsilon}, \dots, \alpha_{m,t}^{\psi,\varepsilon})'_{\varepsilon>0}$, $\psi \in \Psi_0$, is called consistent uniformly lineal asymptotically normal (CULAN) if it admits the following expansion

$$\alpha_t^{\psi,\varepsilon} = \alpha + [\gamma_0^\psi(\alpha)]^{-1} \varepsilon^2 L_t^{\psi,\varepsilon}(\alpha) + r_{\psi,\varepsilon}(\alpha), \tag{2.15}$$

where uniformly in α on each compact

$$P_\alpha^\varepsilon - \lim_{\varepsilon\to 0} \varepsilon^{-1} r_{\psi,\varepsilon}(\alpha) = 0. \tag{2.16}$$

It is well-known (see Lazrieva, Toronjadze [19]) that under the above conditions uniformly in α on each compact

$$\mathcal{L}\{\varepsilon^{-1}(\alpha_t^{\psi,\varepsilon} - \alpha) \mid P_\alpha^\varepsilon\} \xrightarrow{w} N(0, V(\psi; \alpha)),$$

with

$$V(\psi; \alpha) := [\gamma_0^\psi(\alpha)]^{-1} \Gamma_0^\psi(\alpha)([\gamma_0^\psi(\alpha)]^{-1})', \tag{2.17}$$

where $\mathcal{L}(\zeta|P)$ denotes the distribution of random vector ζ calculated under measure P, symbol "$\xrightarrow{w}$" denotes the weak convergence of measures, $N(0, V(\psi; \alpha))$ is a distribution of Gaussian vector with zero mean and covariance matrix $V(\psi; \alpha)$.

Remark 2.1. In context of diffusion type processes the M-estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$ is defined as a solution of the following stochastic equation

$$L_t^{\psi,\varepsilon}(X; \alpha) = 0,$$

where $L_t^{\psi,\varepsilon}(X; \alpha)$ is defined by (2.14), $\psi \in \Psi_0$.

The asymptotic theory of M-estimates for general statistical models with filtration is developed in Chitashvili et al. [7]. Namely, the problem of existence and global behaviour of solutions is studied. In particular, the conditions of regularity and ergodicity type are established, under which M-estimates have a CULAN property.

For our model, in case when $\mathcal{A} = R^m$, the sufficient conditions for CULAN property take the form:

(1) for all s, $0 \le s \le t$, and $x \in C_t$ the functionals $\psi(s,x;\alpha)$ and $\dot{a}(s,x;\alpha)$ are twice continuously differentiable in α with bounded derivatives satisfying the functional Lipshitz conditions with constants, which do not depend on α.

(2) the equation (w.r.t y)

$$\Delta(\alpha, y) := \int_0^t \psi(s, Y^0(\alpha); y)(a(s, Y^0(\alpha); \alpha) - a(s, Y^0(\alpha); y))\, ds = 0,$$

has an unique solution $y = \alpha$.

The MLE is a special case of M-estimates when $\psi = \dot{a}$.

Remark 2.2. According to (2.7) the asymptotic covariance matrix of MLE $(\widehat{\alpha}_t^\varepsilon)_{\varepsilon>0}$ is $[I_0(\alpha)]^{-1}$. By the usual technique one can show that for each $\alpha \in \mathcal{A}$ and $\psi \in \Psi_0$, $I_0^{-1}(\alpha) \le V(\psi, \alpha)$ (see (2.17)), where for two symmetric matrices B and C the relation $B \le C$ means that the mattix $C - B$ is nonnegative definite.

Thus, the MLE has a minimal covariance matrix among all M-estimates.

2.2. **Shrinking contamination neighborhoods.** In this subsection we give a notion of a contamination of the basic model (2.3), described in terms of shrinking neighborhoods of basic measures $\{P_\alpha^\varepsilon,\, \alpha \in \mathcal{A},\, \varepsilon > 0\}$, which is an analog of the Huber gross error model (see, e.g., Hampel et.al. [12] and also, Remark 2.3 below).

Let $\mathcal{H}$ be a family of bounded nonanticipative functionals $h : [0,t] \times C_t \times \mathcal{A} \to R^1$ such that for all $s \in [0,t]$ and $\alpha \in \mathcal{A}$ the functional $h(s,x;\alpha)$ is continuous at the point $x_0 = Y^0(\alpha)$.

Let for each $h \in \mathcal{H}$, $\alpha \in \mathcal{A}$ and $\varepsilon > 0$, $P_\alpha^{\varepsilon,h}$ be a measure on $(C_t, \mathcal{B}_t)$ such that

$$\begin{aligned} &1) \qquad P_\alpha^{\varepsilon,h} \sim P_\alpha^\varepsilon, \\ &2) \qquad \frac{dP_\alpha^{\varepsilon,h}}{dP_\alpha^\varepsilon} = \mathcal{E}_t(\varepsilon N_\alpha^{\varepsilon,h}), \end{aligned} \tag{2.18}$$

where

$$3) \qquad N_{\alpha,s}^{\varepsilon,h} := \int_0^s h_s(u, X; \alpha)(dX_u - a_\varepsilon(u, X; \alpha)\, du), \tag{2.19}$$

with $h_\varepsilon(s,x;\alpha) := \frac{1}{\varepsilon}\, h(s, \varepsilon x; \alpha)$, $0 \le s \le t$, $x \in C_t$.

Denote $\mathbf{P}_\alpha^{\varepsilon,\mathcal{H}}$ a class of measures $P_\alpha^{\varepsilon,h}$, $h \in \mathcal{H}$, that is

$$\mathbf{P}_\alpha^{\varepsilon,\mathcal{H}} = \{P_\alpha^{\varepsilon,h};\;\; h \in \mathcal{H}\}.$$

We call $(\mathbf{P}_\alpha^{\varepsilon,\mathcal{H}})_{\varepsilon>0}$ a shrinking contamination neighborhoods of the basic measures $(P_\alpha^\varepsilon)_{\varepsilon>0}$, and the element $(P_\alpha^{\varepsilon,h})_{\varepsilon>0}$ of these neighborhoods is called alternative measure (or simply alternative).

Obviously for each $h \in \mathcal{H}$ and $\alpha \in \mathcal{A}$, the process $N_\alpha^{\varepsilon,h} = (N_{\alpha,s}^{\varepsilon,h})_{0\le s\le t}$ defined by (2.19) is a P_α^ε-square integrable martingale. Since under measure P_α^ε the process $\overline{w} = (\overline{w}_s)_{0\le s\le t}$ defined as

$$\overline{w}_s := X_s - \int_0^s a_\varepsilon(u, X; \alpha)\, du, \quad 0 \le s \le t,$$

is a Wiener process, then by virtue of the Girsanov Theorem the process $\widetilde{w} := \overline{w} + \langle \overline{w}, \varepsilon N_\alpha^{\varepsilon,h}\rangle$ is a Wiener process under changed measure $P_\alpha^{\varepsilon,h}$. But by the definition

$$\widetilde{w}_s = X_s - \int_0^s (a_\varepsilon(u, X; \alpha) + \varepsilon h_\varepsilon(u, X; \alpha))\, du,$$

and hence, one can conclude that $P_\alpha^{\varepsilon,h}$ is a weak solution of SDE

$$dX_s = (a_\varepsilon(s, X; \alpha) + \varepsilon h_\varepsilon(s, X; \alpha))\, ds + dw_s, \quad X_0 = 0.$$

This SDE can be viewed as a “small” perturbation of the basic model (2.3).

Remark 2.3. 1) In the case of i.i.d. observations $X_1, X_2, \dots, X_n$, $n \ge 1$, the Huber gross error model in shrinking setting is defined as follows

$$f^{n,h}(x; \alpha) := (1 - \varepsilon_n) f(x; \alpha) + \varepsilon_n h(x; \alpha),$$

where $f(x; \alpha)$ is a basic (core) density of distribution of r.v. X_i (w.r.t some dominating measure μ), $h(x; \alpha)$ is a contaminating density, $f^{n,h}(x; \alpha)$ is a contaminated density, $\varepsilon_n = O(n^{-1/2})$. If we denote by P_α^n and $P_\alpha^{n,h}$ the measures on $(R^n, \mathcal{B}(R^n))$, generated by $f(x; \alpha)$ and $f^{n,h}(x; \alpha)$, respectively, then

$$\frac{dP_\alpha^{n,h}}{dP_\alpha^n} = \prod_{i=1}^n \frac{f^{n,h}(X_i; \alpha)}{f(X_i; \alpha)} = \prod_{i=1}^n (1 + \varepsilon_n H(X_i; \alpha)) = \mathcal{E}_n(\varepsilon_n \cdot N_\alpha^{n,h}),$$

where $H = \frac{h-f}{f}$, $N_\alpha^{n,h} = (N_{\alpha,m}^{n,h})_{1\le m\le n}$, $N_{\alpha,m}^{n,h} = \sum_{i=1}^m H(X_i; \alpha)$, $N_\alpha^{n,h}$ is a P_α^n-martingale, $\mathcal{E}_n(\varepsilon_n N_\alpha^{n,h}) = \prod_{i=1}^n (1 + \varepsilon_n \Delta N_{\alpha,i}^{n,h})$ is the Dolean exponential in discrete time case.

Thus

$$\frac{dP_\alpha^{n,h}}{dP_\alpha^n} = \mathcal{E}(\varepsilon_n \cdot N_\alpha^{n,n}), \tag{2.20}$$

and the relation (2.18) is a direct analog of (2.20).

2) The concept of shrinking contamination neighborhoods, expressed in the form of (2.18) was proposed in Lazrieva and Toronjadze [20] for more general situation, concerning with the contamination areas for semimartingale statistical models with filtration. □

Note here that the power of the small parameter ε is crucial. One cannot consider the perturbation of measure with different power of ε if he/she wish to get nontrivial result.

In the remainder of this subsection we study the asymptotic properties of CULAN estimates under alternatives.

For this aim we first consider the problem of contiguity of measures $(P_\alpha^{\varepsilon,h})_{\varepsilon>0}$ to $(P_\alpha^\varepsilon)_{\varepsilon>0}$.

Let $(\varepsilon_n)_{n\ge1}$, $\varepsilon_n \downarrow 0$, and $(\alpha_n)_{n\ge1}$, $\alpha_n \in K$, $K \subset \mathcal{A}$ is a compact, be arbitrary sequences.

Proposition 2.1. *For each $h \in \mathcal{H}$ the sequence of measures $(P^{\varepsilon_n,h}_{\alpha_n})$ is contiguous to sequence of measures $(P^{\varepsilon_n}_{\alpha_n})$, i.e.*

$$(P^{\varepsilon_n,h}_{\alpha_n}) \lhd (P^{\varepsilon_n}_{\alpha_n}).$$

Proof. From the predictable criteria of contiguity (see, e.g., Jacod and Shiryaev [15]), follows that we have to verify the relation

$$\lim_{N\to\infty} \limsup_{n\to\infty} P^{\varepsilon_n,h}_{\alpha_n} \left\{ h^n_t \left(\frac{1}{2}\right) > N \right\} = 0, \tag{2.21}$$

where $h^n(\frac{1}{2}) = (h^n_s(\frac{1}{2}))_{0\le s\le t}$ is the Hellinger process of order $\frac{1}{2}$.

By the definition of Hellinger process (see, e.g., Jacod and Shiryaev [15]) we have

$$h^n_t\left(\frac{1}{2}\right) = h^n_t\left(\frac{1}{2}, P^{\varepsilon_n,h}_{\alpha_n}, P^{\varepsilon_n}_{\alpha_n}\right) = \frac{1}{8}\int_0^t [h(s,\varepsilon_n X;\alpha_n)]^2\,ds,$$

and since $h \in \mathcal{H}$, and hence is bounded, $h^n_t(\frac{1}{2})$ is bounded too, which provides (2.21). □

Proposition 2.2. *For each estimate $(\alpha^{\varepsilon,\psi}_t)_{\varepsilon>0}$ with $\psi \in \Psi_0$ and each alternative $(P^{\varepsilon,h}_\alpha)_{\varepsilon>0} \in (\mathbf{P}^{\varepsilon,\mathcal{H}}_\alpha)_{\varepsilon>0}$ the following relation holds true*

$$\mathcal{L}\left\{\varepsilon^{-1}(\alpha^{\psi,\varepsilon}_t - \alpha) \mid P^{\varepsilon,h}_\alpha\right\} \xrightarrow{w} N\left([\gamma^\psi_0(\alpha)]^{-1} b(\psi,h;\alpha), V(\psi,\alpha)\right),$$

where

$$b(\psi,h;\alpha) := \int_0^t \psi(s,Y^0(\alpha);\alpha)h(s,Y^0(\alpha);\alpha)\,ds.$$

Proof. Proposition 2.1 together with (2.16) provides that uniformly in α on each compact

$$P^{\varepsilon,h}_\alpha - \lim_{\varepsilon\to 0} \varepsilon^{-1} r_{\psi,\varepsilon}(\alpha) = 0,$$

and therefore we have to establish the limit distribution of random vector $[\gamma^\psi_0(\alpha)]^{-1}\varepsilon L^{\psi,\varepsilon}_t$ under the measures $(P^{\varepsilon,h}_\alpha)_{\varepsilon>0}$.

By virtue of the Girsanov Theorem the process $L^{\psi,\varepsilon}(\alpha) = (L^{\psi,\varepsilon}_s(\alpha))_{0\le s\le t}$ is a semimartingale with canonical decomposition

$$L^{\psi,\varepsilon}_s(\alpha) = \widetilde{L}^{\psi,\varepsilon}_s(\alpha) + b_{\varepsilon,s}(\psi,h;\alpha), \quad 0 \le s \le t, \tag{2.22}$$

where $\widetilde{L}^{\psi,\varepsilon}(\alpha) = (\widetilde{L}^{\psi,\varepsilon}_s(\alpha))_{0\le s\le t}$ is a $P^{\varepsilon,h}_\alpha$-square integrable martingales defined as follows

$$\widetilde{L}^{\psi,\varepsilon}_s(X;\alpha) := \int_0^s \psi_\varepsilon(u,X;\alpha)\ (dX_u - (a_\varepsilon(u,X;\alpha) + \varepsilon h_\varepsilon(u,X;\alpha))\ du,$$

and

$$b_{\varepsilon,s}(\psi,h;\alpha) := \varepsilon\int_0^s \psi_\varepsilon(u,X;\alpha)h_\varepsilon(u,X;\alpha)\,du.$$

But $\langle \widetilde{L}^{\psi,\varepsilon}(\alpha)\rangle_t = \Gamma^\psi_\varepsilon(\alpha)$, where $\Gamma^\psi_\varepsilon(\alpha)$ is defined by (2.8). On the other hand, from Proposition 2.1 and (2.10) it follows that

$$P^{\varepsilon,h}_\alpha - \lim_{\varepsilon\to 0}\langle \varepsilon\widetilde{L}^{\psi,\varepsilon}(\alpha)\rangle_t = P^{\varepsilon,h}_\alpha - \lim_{\varepsilon\to 0}\varepsilon^2\Gamma^\psi_\varepsilon(\alpha) = P^\varepsilon_\alpha - \lim_{\varepsilon\to 0}\varepsilon^2\Gamma^\psi_\varepsilon(\alpha) = \Gamma^\psi_0(\alpha)$$

uniformly in α on each compact, and hence

$$\mathcal{L}\left\{[\gamma^\psi_0(\alpha)]^{-1}\varepsilon\widetilde{L}^{\psi,\varepsilon}_t(\alpha) \mid P^{\varepsilon,h}_\alpha\right\} \xrightarrow{w} N(0, V(\psi;\alpha)). \tag{2.23}$$

Finally, relation (2.23) together with (2.22) and relation

$$P_\theta^{\varepsilon,h} - \lim_{\varepsilon\to 0} \varepsilon b_{\varepsilon,t}(\psi,h;\alpha) = \int_0^t \psi(s,Y^0(\alpha);\alpha)h(s,Y^0(\alpha);\alpha)\,ds = b(\psi,h;\alpha),$$

provides the desirable results. □

2.3. **Optimization criteria. Construction of optimal B-robust estimates.** In this subsection we state and solve an optimization problem, which results in construction of optimal B-robust estimate.

Initially, it should be stressed that the bias vector $\widetilde{b}(\psi,h;\alpha) := [\gamma_0^\psi(\alpha)]^{-1} \times b(\psi,h;\alpha)$ can be viewed as the influence functional of the estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$ w.r.t. alternative $(P_\alpha^{\psi,h})_{\varepsilon>0}$.

Indeed, the expansion (2.15) together with (2.22) and (2.23) allows to conclude that

$$\mathcal{L}\left\{\varepsilon^{-1}(\alpha_t^{\psi,\varepsilon} - \alpha - \varepsilon^2[\gamma_0^\psi(\alpha)]^{-1}b_\varepsilon(\psi,h;\alpha)) \mid P_\alpha^{\varepsilon,h}\right\} \xrightarrow{w} N(0,V(\psi,\alpha)),$$

and, hence, the expression

$$\alpha + \varepsilon^2[\gamma_0^\psi(\alpha)]^{-1}b_\varepsilon(\psi,h;\alpha) - \alpha = \varepsilon^2[\gamma_0^\psi(\alpha)]^{-1}b_\varepsilon(\psi,h;\alpha),$$

plays the role of bias on the "fixed step ε" and it seems natural to interpret the limit

$$P_\alpha^{\varepsilon,h} - \lim_{\varepsilon\to 0} \frac{\alpha + \varepsilon^2[\gamma_0^\psi(\alpha)]^{-1}b_\varepsilon(\psi,h;\alpha) - \alpha}{\varepsilon} = [\gamma_0^\psi(\alpha)]^{-1}b(\psi,h;\alpha),$$

as the influence functional.

For each estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$, $\psi\in\Psi_0$, define the risk functional w.r.t. alternative $(P_\alpha^{\varepsilon,h})_{\varepsilon>0}$, $h\in\mathcal{H}$, as follows:

$$D(\psi,h;\alpha) = \lim_{a\to\infty}\lim_{\varepsilon\to 0} E_\alpha^{\varepsilon,h}\left((\varepsilon^{-2}|\alpha_t^{\psi\varepsilon} - \alpha|^2)\wedge a\right),$$

where $x\wedge\alpha = \min(x,a)$, $a>0$, $E_\alpha^{\varepsilon,h}$ is an expectation w.r.t. measure $P_\alpha^{\varepsilon,h}$.

Using Proposition 2.2 it is not hard to verify that

$$D(\psi,h;\alpha) = |\widetilde{b}(\psi,h;\alpha)|^2 + \operatorname{tr} V(\psi,\alpha),$$

where $\operatorname{tr} A$ denotes the trace of matrix A.

By Proposition 2.2

$$\varepsilon^{-1}(\alpha_t^{\psi,\varepsilon} - \alpha) \xrightarrow{d} N\big(\widetilde{b}(\psi,h;\alpha), V(\psi;\alpha)\big),$$

where $\xrightarrow{d}$ denotes the convergence by distribution (by distribution $P_\alpha^{\varepsilon,h}$ in our case), $N(\widetilde{b},V)$ is a Gaussian random vector with mean $\widetilde{b}$ and covariation matrix V.

But if $\xi = (\xi_1,\dots,\xi_m)'$ is a Gaussian vector with parameters (μ,σ^2), then

$$E|\xi|^2 = \sum_{i=1}^m E\xi_i^2 = \sum_{i=1}^m (E\xi_i)^2 + \sum_{i=1}^m D\xi_i = |\mu|^2 + \operatorname{tr}\sigma^2,$$

as it was required.

Connect with each $\psi\in\Psi_0$ the function $\widetilde{\psi}$ as follows

$$\widetilde{\psi}(s,x;\alpha) = [\gamma_0^\psi(\alpha)]^{-1}\psi(s,x;\alpha), \quad 0\le s\le t, \quad x\in C_t, \quad \alpha\in\mathcal{A}.$$

Then $\widetilde{\psi} \in \Psi_0$ and

$$\gamma_0^{\widetilde{\psi}}(\alpha) = Id,$$

where Id is an unit matrix,

$$V(\psi;\alpha) = V(\widetilde{\psi};\alpha) = \Gamma_0^{\widetilde{\psi}}(\alpha), \quad \widetilde{b}(\psi, h;\alpha) = \widetilde{b}(\widetilde{\psi}, h;\alpha) = b(\widetilde{\psi}, h;\alpha).$$

Therefore

$$D(\psi, h;\alpha) = D(\widetilde{\psi}, h;\alpha) = |b(\widetilde{\psi}, h;\alpha)|^2 + \operatorname{tr}\Gamma_0^{\widetilde{\psi}}(\alpha). \tag{2.24}$$

Denote $\mathcal{H}_r$, a set of functions $h \in \mathcal{H}$ such that for each $\alpha \in \mathcal{A}$

$$\int_0^t |h(s, Y^0(\alpha);\alpha)|\, ds \leq r,$$

where r, $r > 0$, is a constant.

Since, for each $r > 0$,

$$\sup_{h\in\mathcal{H}_r} |b(\widetilde{\psi}, h;\alpha)| \leq const(r) \sup_{0\leq s\leq t} |\widetilde{\psi}(s, Y^0(\alpha);\alpha)|,$$

where constant depends on r, we call the function $\widetilde{\psi}$ an influence function of estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$ and a quantity

$$\gamma_\psi^*(\alpha) = \sup_{0\leq s\leq t} |\widetilde{\psi}(s, Y^0(\alpha);\alpha)|$$

is named as the (unstandardized) gross error sensitivity at point α of estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$.

Define

$$\Psi_{0,c} = \left\{ \psi \in \Psi_0 \ : \ \int_0^t \psi(s, Y^0(\alpha);\alpha)[\dot{a}(s, Y^0(\alpha);\alpha)]' ds = Id, \right. \tag{2.25}$$

$$\left. \gamma_\psi^*(\alpha) \leq c \right\}, \tag{2.26}$$

where $c \in [0, \infty)$ is a generic constant.

Take into account the expression (2.24) for the risk functional we come to the following optimization problem, known in robust estimation theory as Hampel's optimization problem: minimize the trace of the asymptotic covariance matrix of estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$ over the class $\Psi_{0,c}$, that is

$$\text{minimize} \quad \int_0^t \psi(s, Y^0(\alpha);\alpha)[\psi(s, Y^0(\alpha);\alpha)]' ds \tag{2.27}$$

under the side conditions (2.25) and (2.26).

Define the Huber function $h_c(z)$, $z \in R^m$, $c > 0$, as follows

$$h_c(z) := z \min\left(1, \frac{c}{|z|}\right).$$

For arbitrary nondegenerate matrix A denote $\psi_c^A = h_c(A\dot{a})$.

Theorem 2.1. *Assume that for given constant c there exists a nondegenerate $m \times m$-matrix $A_c^*(\alpha)$, which solves the equation (w.r.t. matrix A)*

$$\int_0^t \psi_c^A(s, Y^0(\alpha);\alpha)[\dot{a}(s, Y^0(\alpha);\alpha)]' ds = Id. \tag{2.28}$$

Then the function $\psi_c^{A_c^*(\alpha)} = h_c(A_c^*(\alpha)\dot{a})$ *solves the optimization problem* (2.27).

Proof. We follow Hampel et al. [12].

Let A be an arbitrary $m \times m$-matrix.

Since for each $\psi \in \Psi_{0,c}$, $\int \psi(\dot{a})' = Id$, $\int \dot{a}[\dot{a}]' = I^0(\alpha)$ (see (2.7)), then

$$\int (\psi - A\dot{a})(\psi - A\dot{a})' = \int \psi\psi' - A - A' + AI^0(\alpha)A'$$

(here and below we use simple evident notation for integrals).

Therefore since the trace is an additive functional instead of minimizing of $\operatorname{tr}\int \psi\psi'$ we can minimize

$$\operatorname{tr}\int (\psi - A\dot{a})(\psi - A\dot{a})' = \int |\psi - A\dot{a}|^2.$$

Note that for each z

$$\arg\min_{|y|\le c} |z-y|^2 = h_c(z).$$

Indeed, it is evident that minimizing y has the form $y = \beta z$, where β, $0 \le \beta \le 1$, is constant. Then

$$\min_{|y|\le c} |z-y|^2 = \min_{\beta \le \frac{c}{|z|}} (1-\beta)^2 |z|^2.$$

Thus we have to find

$$\arg\min_{\beta\le\frac{c}{|z|}} (1-\beta)^2 = \min\Big(1, \frac{c}{|z|}\Big).$$

But last relation is trivially satisfied. Hence the minimizing $y^* = z\min(1, \frac{c}{|z|})$ and

$$\arg\min_{|\psi|\le c} |\psi - A\dot{a}|^2 = h_c(A\dot{a}).$$

From the other side,

$$|h_c(z)|^2 = |z|^2 I_{\{|z|\le c\}} + \frac{|z|^2}{|z|^2}\, c^2\, I_{(|z|\ge c)} \le c^2.$$

Hence

$$|h_c(z)| \le c \quad \text{for all} \quad z$$

and therefore $h_c(A\dot{a})$ satisfies the condition (2.26) for each A.

Now it is evident that a function $h_c(A\dot{a})$ minimizes the expression under integral sign, and hence the integral itself over all functions $\psi \in \Psi_0$ satisfying (2.26).

At the same time the condition (2.25), generally speaking, can be violated. But, since a matrix A is arbitrary, we can choose $A = A_c^*(\alpha)$ from (2.28) which, of course, guarantees the validity of (2.25) for $\psi_c^* = \psi_c^{A_c^*(\alpha)}$. □

As we have seen the resulting optimal influence functions ψ_c^* is defined along the process $Y^0(\alpha) = (Y_s^0(\alpha))_{0\le s\le t}$, which is a solution of equation (2.2).

But for constructing optimal estimate we need a function $\psi_c^*(s, x; \alpha)$ defined on whole space $[0, t] \times C_t \times \mathcal{A}$.

For this purpose define $\psi_c^*(s, x; \alpha)$ as follows;

$$\psi_c^*(s, x; \alpha) = \psi_c^{A_c^*(\alpha)}(s, x; \alpha) = h_\varepsilon(A_c^*(\alpha)\dot{a}(s, x; \alpha)), \tag{2.29}$$

and as usual $\psi^*_{c,\varepsilon}(s,x;\alpha) = \frac{1}{\varepsilon}\psi^*_c(s,\varepsilon x;\alpha)$, $0 \le s \le t$, $x \in C_t$, $\alpha \in \mathcal{A}$.

Definition 2.2. We say that $\psi^*_c(s,x;\alpha)$, $0 \le s \le t$, $x \in C_t$, $\alpha \in \mathcal{A}$, is an influence function of optimal B-robust estimate $(\alpha_t^{*,\varepsilon})_{\varepsilon>0} = (\alpha_t^{\psi^*_c,\varepsilon})_{\varepsilon>0}$ over the class of CULAN estimates $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$, $\psi \in \Psi_{0,c}$, if the matrix $A^*(\alpha)$ is differentiable in α.

From (2.9), (2.11), (2.28) and (2.29) it directly follows that

$$\gamma_0^{\psi^*_c}(\alpha) = P^\varepsilon_\alpha - \lim_{\varepsilon to 0} \varepsilon^2 \gamma_\varepsilon^{\psi^*_c}(\alpha) = \int_0^t \psi^*_c(s, Y^0(\alpha);\alpha)(\dot a(s, Y^0(\alpha);\alpha))' ds = Id.$$

Besides, for each alternative $(P_\alpha^{\varepsilon,h})_{\varepsilon>0}$, $h \in \mathcal{H}$, according to the Proposition 2.2 we have

$$\mathcal{L}\left\{\varepsilon^{-1}(\alpha_t^{*,\varepsilon} - \alpha) \mid P_\alpha^{\varepsilon,h}\right\} \xrightarrow{w} N(b(\psi^*_c, h;\alpha), V(\psi^*_c;\alpha)) \quad \text{as} \quad \varepsilon \to 0,$$

where

$$b(\psi^*_c, h;\alpha) = \int_0^t \psi^*_c(s, Y^0(\alpha);\alpha) h(s, Y^0(\alpha);\alpha)\, ds,$$

and $V(\psi^*_c;\alpha) = \Gamma_0^{\psi^*_c}(\alpha)$.

Hence, the risk functional for estimate $(\alpha_t^{*,\varepsilon})_{\varepsilon>0}$ is

$$D(\psi^*_c, h;\alpha) = |b(\psi^*_c, h;\alpha)|^2 + \operatorname{tr}\Gamma_0^{\psi^*_c}, \quad h \in \mathcal{H},$$

and the (unstandardized) gross error sensitivity of $(\alpha_t^{*,\varepsilon})_{\varepsilon>0}$ is

$$\gamma_{\psi^*_c}(\alpha) = \sup_{0 \le s \le t} |\psi^*_c(s, Y^0(\alpha);\alpha)| \le c.$$

From above reasons, we may conclude that $(\alpha_t^{*,\varepsilon})_{\varepsilon>0}$ is the optimal B-robust estimate over the class of estimates $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$, $\psi \in \Psi_{0,c}$ in the following sense: the trace of asymptotic covariance matrix of $(\alpha_t^{*,\varepsilon})_{\varepsilon>0}$ is minimal among all estimates $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$ with bounded by constant gross error sensitivity, that is

$$\Gamma_0^{\psi^*_c}(\alpha) \le \Gamma_0^{\psi}(\alpha) \quad \text{for all} \quad \psi \in \Psi_{0,c}\,. \qquad \Box$$

Note that for each estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$ and alternatives $(P_\alpha^{\varepsilon,h})_{\varepsilon>0}$, $h \in \mathcal{H}$, the influence functional is bounded by $const(r) \cdot c$. Indeed, we have for $\psi \in \Psi_{0,c}$,

$$\sup_{h \in \mathcal{H}_r} |b(\psi, h;\alpha)| \le const(r) \cdot c := C(r,c),$$

and since from (2.24)

$$\inf_{\psi \in \Psi_{0,c}} \sup_{h \in \mathcal{H}_r} D(\psi, h;\alpha) \le C^2(r,c) + \operatorname{tr}\Gamma_0^{\psi^*_c}(\alpha),$$

we can choose "optimal level" of truncation, minimizing the expression

$$C^2(r,c) + \operatorname{tr}\Gamma_0^{\psi^*_c}(\alpha)$$

over all constants c, for which the equation (2.28) has a solution $A^*_c(\alpha)$. This can be done using the numerical methods.

For the problem of existence and uniqueness of solution of equation (2.28) we address to Rieder [34].

In the case of one-dimensional parameter α (i.e. $m = 1$) the optimal level c^* of truncation is given as an unique solution of the following equation (see Lazrieva and Toronjadze [20], [21])

$$r^2c^2 = \int_0^t [\dot{a}(s, Y^0(\alpha); \alpha)]_{-c}^{c} \dot{a}(s, Y^0(\alpha); \alpha)\, ds - \int_0^t ([\dot{a}(s, Y^0(\alpha); \alpha)]_{-c}^{c})^2\, ds,$$

where $[x]_a^b = (x \wedge b) \vee a$ and the resulting function

$$\psi^*(s, x; \alpha) = [\dot{a}(s, x; \alpha)]_{-c^*}^{c^*}, \quad 0 \le s \le t, \quad x \in C_t,$$

is $(\Psi_0, \mathcal{H}_r)$ optimal in the following minimax sense:

$$\sup_{h \in \mathcal{H}_r} D(\psi^*, h; \alpha) = \inf_{\psi \in \Psi} \sup_{h \in \mathcal{H}_r} D(\psi, h; \alpha).$$

3. Optimal Mean-Variance Robust Hedging

3.1. **A financial market model.** Let $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{0 \le t \le T}, P)$ be a filtered probability space with filtration F satisfying the usual conditions, where $T \in (0, \infty]$ is a fixed time horizon. Assume that $\mathcal{F}_0$ is a trivial and $\mathcal{F}_T = \mathcal{F}$.

There exist $d + 1$, $d \ge 1$ primitive assets: one bound, whose price process is assumed to be 1 at all times and d risky assets (stocks), whose R^d-valued price process $X = (X_t)_{0 \le t \le T}$ is a continuous semimartingale given by the relation:

$$dX_t = \operatorname{diag}(X_t)\, dR_t, \quad X_0 > 0, \tag{3.1}$$

where $\operatorname{diag}(X)$ denotes the diagonal $d \times d$-matrix with diagonal elements $X^1, \dots, X^d$, and the yield process $R = (R_t)_{0 \le t \le T}$ is a R^d-valued continuous semimartingale satisfying the stricture condition (SC). That is (see Schweizer [37])

$$dR_t = d\langle \widetilde{M} \rangle_t \lambda_t + d\widetilde{M}_t, \quad R_0 = 0, \tag{3.2}$$

where $\widetilde{M} = (\widetilde{M}_t)_{0 \le t \le T}$ is a R^d-valued continuous martingale, $\widetilde{M} \in \mathcal{M}_{0,\mathrm{loc}}^2(P)$, $\lambda = (\lambda_t)_{0 \le t \le T}$ is a F-predictable R^d-valued process, and the mean-variance tradeoff (MVT) process $\widetilde{\mathcal{K}} = (\widetilde{\mathcal{K}}_t)_{0 \le t \le T}$ of process R

$$\widetilde{\mathcal{K}}_t := \int_0^t \lambda_s' d\langle \widetilde{M} \rangle_s \lambda_s = \langle \lambda' \cdot \widetilde{M} \rangle_t < \infty \;\; P\text{-a.s.}, \;\; t \in [0, T]. \tag{3.3}$$

Remark 3.1. Remember that all vectors are assumed to be column vectors.

Suppose that the martingale $\widetilde{M}$ has the form

$$\widetilde{M} = \sigma \cdot M, \tag{3.4}$$

where $M = (M_t)_{0 \le t \le T}$ is a R^d-valued continuous martingale, $M \in \mathcal{M}_{0,\mathrm{loc}}^2(P)$ with $d\langle M^i, M^j \rangle_t = I_{i_j}^{d \times d} dC_t$, $I^{d \times d}$ is the identity matrix, $C = (C_t)_{0 \le t \le T}$ is a continuous increasing bounded process with $C_0 = 0$.

Further, let $\sigma = (\sigma_t)_{0 \le t \le T}$ is a $d \times d$-matrix valued, F-predictable process with $\operatorname{rank}(\sigma_t) = d$ for any t, P-a.s., the process $(\sigma_t^{-1})_{0 \le t \le T}$ is locally bounded, and

$$\int_0^T \sigma_t\, d\langle M \rangle_t \sigma_t' < \infty \;\; P\text{-a.s.} \tag{3.5}$$

Assume now that the following condition be satisfied:
There exist fixed R^d-valued, F-predictable process $k = (k_t)_{0 \le t \le T}$ such that

$$\lambda = \lambda(\sigma) = (\sigma')^{-1} k. \tag{3.6}$$

In the case from (3.2) we get

$$\begin{aligned} dR_t &= d\langle \widetilde{M} \rangle_t \lambda_t + d\widetilde{M}_t = \sigma_t d\langle M \rangle_t \sigma_t' (\sigma_t')^{-1} k_t + \sigma_t dM_t \\ &= \sigma_t (d\langle M \rangle_t k_t + dM_t), \end{aligned} \tag{3.7}$$

and

$$\begin{aligned} \widetilde{\mathcal{K}}_t &= \int_0^t \lambda_s' d\langle \widetilde{M} \rangle_s \lambda_s = \int_0^t k_t' ((\sigma_t')^{-1})' \sigma_t d\langle \widetilde{M} \rangle_t \sigma_t' (\sigma_t')^{-1} k_t \\ &= \int_0^t k_t' d\langle M \rangle_t k_t = \langle k \cdot M \rangle_t := \mathcal{K}_t. \end{aligned}$$

From (3.3) we have

$$\mathcal{K}_t < \infty \quad P\text{-}a.s. \ \text{ for all } \ t \in [0, T]. \tag{3.8}$$

Thus, of we introduce the process $M^0 = (M_t^0)_{0 \le t \le T}$ by the relarion

$$dM_t^0 = d\langle M \rangle_t k_t + dM_t, \quad M_0^0 = 0, \tag{3.9}$$

then the MVT process $\mathcal{K} = (\mathcal{K}_t)_{0 \le t \le T}$ of R^d-valued semimartingale M^0 is finite, and hence M^0 satisfies SC.

Finally, the scheme (3.1), (3.2), (3.4), (3.6) and (3.9) can be rewritten in the following form

$$\begin{aligned} dX_t &= \operatorname{diag}(X_t)\, dR_t, \quad X_0 > 0, \\ dR_t &= \sigma_t\, dM_t^0, \quad R_0 = 0, \\ dM_t^0 &= d\langle M \rangle_t k_t + dM_t, \quad M_0 = 0, \end{aligned} \tag{3.10}$$

where σ and k satisfy (3.5) and (3.8), respectively.

This is our financial market model.

3.2. **Characterization of variance-optimal ELMM (equivalent local martingale measure).** A key role in mean-variance hedging plays variance-optimal ELMM (see, e.g., RSch [33], GLP [11]). Here we collect some facts characterizing this measure.

We start with remark that the sets ELMMs for processes X, R and M^0 form (3.10) coincide. Hence we can and will consider the simplest process M^0.

Introduce the notation

$$\mathcal{M}_2^e := \left\{ Q \sim P \,:\, \frac{dQ}{dP} \in L^2(P), \ \ M^0 \ \text{ is a } Q\text{-local martingale} \right\},$$

and suppose that

$$(c.1) \qquad\qquad \mathcal{M}_2^e \neq \varnothing.$$

The solution $\widetilde{P}$ of the optimization problem

$$E\mathcal{E}_T^2(\mathcal{M}^Q) \to \inf_{Q \in \mathcal{M}_2^e}$$

is called variance-optimal ELMM.

Here

$$\frac{dQ}{dP}\Big|_{\mathcal{F}_T} = \mathcal{E}_T(M^Q),$$

and $(\mathcal{E}_t(M^Q))_{0\le t\le T}$ is the Dolean exponential of martingale M^Q.

It is well-known (see, e.g., Schweizer [37], [38]) that under condition (c.1) variance-optimal ELMM $\widetilde{P}$ exist.

Denote

$$\widetilde{z}_T := \frac{d\widetilde{P}}{dP}\Big|_{\mathcal{F}_T},$$

and introduce RCLL process $\widetilde{z} = (\widetilde{z}_t)_{0\le t\le T}$ by the relation

$$\widetilde{z}_t = E^{\widetilde{P}}(\widetilde{z}_T/\mathcal{F}_T), \quad 0 \le t \le T.$$

Then, by Schweizer [37], [38]

$$\widetilde{z}_T = \widetilde{z}_0 + \int_0^T \zeta_t'\, dM_t^0, \tag{3.11}$$

where $\zeta = (\zeta_t)_{0\le t\le T}$ is the R^d-valued F-predictable process with

$$\int_0^T \zeta_t'\, d\langle M\rangle_t \zeta_t < \infty,$$

and the process $\big(\int\limits_0^t \zeta_s' dM_s^0\big)_{0\le t\le T}$ is a $\widetilde{P}$-martingale.

Relation (3.11) easily implies that the process $\widetilde{z}$ is actually continuous.

Suppose, in addition to (c.1), that the following condition is satisfied:

(c.*) all P-local martingales are continuous.

This technical assumption is satisfied in stochastic volatility models, where $F = F^w$ is the natural filtration generated by the Wiener process.

It shown in Mania and Tevzadze [34], Mania et al. [25] that under conditions (c.1) and (c*) density $\widetilde{z}_T$ of variance optimal ELMM is uniquely characterized by the relation

$$\widetilde{z}_T = \frac{\mathcal{E}_T((\varphi-k)'\cdot M^0)}{E\mathcal{E}_T((\varphi-k)'\cdot M^0)}, \tag{3.12}$$

where φ together with the pair (L, c) is the unique solution of the following equation

$$\frac{\mathcal{E}_T((\varphi-2k)'\cdot M)}{\mathcal{E}_T(L)} = c\mathcal{E}_T^2(-k'\cdot M), \tag{3.13}$$

where $L \in M^2_{0,\mathrm{loc}}(P)$, $\langle L, M\rangle = 0$, c is a constant.

Moreover, the process $\zeta = (\zeta_t)_{0\le t\le T}$ from (3.11) has the form

$$\zeta_t = (\varphi_t - k_t)\mathcal{E}_t((\varphi-k)'\cdot M^0). \tag{3.14}$$

Here $\varphi = (\varphi_t)_{0\le t\le T}$ is a R^d-valued, F-predictable process with

$$\int_0^T \varphi_t'\, d\langle M\rangle_t \varphi_t < \infty.$$

Let τ be F-stopping time.

Denote $\langle k'\cdot M\rangle_{T\tau} = \langle k'\cdot M\rangle_T - \langle k'\cdot M\rangle_\tau$.

Proposition 3.1 (see also Biagini et al. [3], LLaurent and Pham [18]).

1. *Equation* (3.13) *is equivalent to equation*

$$\frac{\mathcal{E}_T(\varphi' \cdot M^*)}{\mathcal{E}_T(L)} = ce^{\langle k' \cdot M \rangle_T}, \tag{3.15}$$

where the R^d-valued process $M^* = (M_t^*)_{0 \le t \le T}$ is given by the relation

$$dM_t^* = 2d\langle M \rangle_t k_t + dM_t, \quad M_0^* = 0.$$

2. a) *If there exists the martingale $m = (m_t)_{0 \le t \le T}$, $m \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ such that*

$$e^{-\langle k' \cdot M \rangle_T} = c + m_T, \quad \langle m, M \rangle = 0, \tag{3.16}$$

then $\varphi \equiv 0$ and $L_T = \int\limits_0^T \frac{1}{c+m}\, dm_s$ solve the equation (3.15).

In this case

$$\widetilde{z}_T = \frac{\mathcal{E}_T(-k' \cdot M^0)}{E\mathcal{E}_T(-k' \cdot M^0)}, \tag{3.17}$$

process $\zeta = (\zeta_t)_{0 \le t \le T}$ from (3.11) *is equal to*

$$\zeta_t = -k_t \mathcal{E}_t(-k' \cdot M^0),$$

and

$$E\left[\left(\frac{\widetilde{z}_T}{\widetilde{z}_\tau}\right)^2 \Big/ \mathcal{F}_\tau\right] = \frac{1}{E(e^{-\langle k' \cdot M \rangle_{T\tau}} / \mathcal{F}_\tau)}.$$

b) *If there exist R^d-valued F-predictable process $\ell = (\ell_t)_{0 \le t \le T}$, $\int\limits_0^T \ell_t' d\langle M \rangle \ell_t < \infty$ and*

$$e^{\langle k' \cdot M \rangle_T} = c + \int_0^T \ell_t'\, dM_t^*,$$

then $L \equiv 0$ and $\varphi_t = \frac{\ell_t}{c + \int_0^t \ell_s' dM_s^*}$ solve the equation (3.15).

In this case

$$\widetilde{z}_T = \mathcal{E}_T(-k' \cdot M) \ \ (:= \widehat{z}_T, \ \ \textit{the density of minimal martingale measure } \ \widehat{P}),$$

and

$$E\left(\left(\frac{\widetilde{z}_T}{\widetilde{z}_\tau}\right)^2 \Big/ \mathcal{F}_\tau\right) = E^{P^*}(e^{\langle k' \cdot M \rangle_{T\tau}} / \mathcal{F}_\tau),$$

where $dP^* = \mathcal{E}_T(-2k' \cdot M)dP$.

Proof. 1. By the Yor formula

$$\mathcal{E}_T\left((\varphi - 2k)' \cdot M\right) = \mathcal{E}_T(\psi' \cdot M - 2k' \cdot M)$$

$$= \mathcal{E}_T\left(\varphi' \cdot \left(M + 2\int_0^\cdot d\langle M \rangle_t k_t\right) - 2\int_0^\cdot \psi_t' d\langle M \rangle_t k_t - 2k' \cdot M\right)$$

$$= \mathcal{E}_T(\varphi' \cdot M^*)\mathcal{E}_T(-2k' \cdot M),$$

and

$$\mathcal{E}_T^2(-k' \cdot M) = \mathcal{E}_T(-2k' \cdot M)e^{\langle k' \cdot M \rangle_T}.$$

Assertion follows.

2. a) Note at first that $\langle L, M\rangle = 0$. Further, by the formula we can write

$$\ln(c+m_t) - \ln c = \int_0^t \frac{1}{c+m_s}\, dm_s - \frac{1}{2}\int_0^t \frac{1}{(c+m_s)^2}\, d\langle m\rangle_s.$$

Hence

$$e^{\ln(c+m_T)-\ln c} = \mathcal{E}_T(L),$$

and thus

$$\mathcal{E}_T(L) = \frac{c+m_T}{c} = \frac{e^{-\langle k'\cdot M\rangle_T}}{c}.$$

Finally, by the Bayes rule and the Girasnov Theorem

$$E\left(\left(\frac{\widetilde{z}_T}{\widetilde{z}_\tau}\right)^2 \Big/ \mathcal{F}_\tau\right) = \frac{E(\mathcal{E}_T(-2k'\cdot M)e^{-\langle k'\cdot M\rangle_T}/\mathcal{F}_\tau)}{E^2(\mathcal{E}_T(-k'\cdot M)e^{-\langle k'\cdot M\rangle_T}/\mathcal{F}_\tau)}$$
$$= \frac{E^*(c+m_T/\mathcal{F}_\tau)\mathcal{E}_T^2(-k'\cdot M)}{(\widehat{E}(c+m_\tau/\mathcal{F}_\tau))^2\mathcal{E}_T^2(-2k'\cdot M)} = \frac{c+m_\tau}{(c+m_\tau)^2}\cdot e^{\langle k'\cdot M\rangle_\tau}$$
$$= \frac{1}{E(e^{\langle k'\cdot M\rangle_{T\tau}}/\mathcal{F}_\tau)}.$$

The proof of case 2 b) is quite analogous. □

3.3. **Misspecified asset price model and robust hedging.** Denote by $\mathrm{Ball}_L(0,r)$, $r\in[0,\infty)$ the closed r-radius ball in the space $L = L_\infty(dt\times dP)$, with the center at the origin, and let

$$\begin{aligned}\mathcal{H} := \big\{h = \{h_{ij}\},\ i,j=\widehat{1,d} : h \text{ is } F\text{-predictable } d\times d\text{-matrix}\\ \text{valued process},\ \ \mathrm{rank}(h)=d,\ \ h_{ij}\in \mathrm{Ball}_L(0,r),\ \ r\in[0,\infty)\big\}.\end{aligned} \tag{3.18}$$

Class $\mathcal{H}$ is called the class of alternatives.

Fix the value of small parameter $\delta > 0$, as well as $d\times d$-matrix valued, F-predictable process $\sigma^0 = (\sigma^0_t)_{0\le t\le T} = (\{\sigma^0_{ij,t}\},\ 1\le i,j\le d)_t$ such that $|\sigma^0_{ij,t}| \le const$, $\forall i,j,t$, the matrix $(\sigma^0)^2 = \sigma^0(\sigma^0)'$ is uniformly elliptic, i.e. for each vector $v_t = (v_t', \dots, v_t^d)$ with probability 1

$$\sum_{i,j=1}^d (\sigma^0)^2_{ij,t} v_t^i v_t^j \ge c\sum_{i=1}^d |v_t^i|^2, \quad c>0, \quad 0\le t\le T, \tag{3.19}$$

and denote

$$A_\delta = \{\sigma : \sigma = \sigma^0 + \delta h,\ \ h\in\mathcal{H}\}. \tag{3.20}$$

Proposition 3.2. *Every σ from the class A_δ for sufficiently small δ is F-predictable $d\times d$-valued process with bounded elements and the matrix $\sigma^2 = \sigma\sigma'$ is uniformly elliptic.*

Proof. The process σ is F-predictable as linear combination of F-predictable processes. Further,

$$|\sigma_{ij,t}| = |\sigma^0_{ij,t} + \delta h_{ij,t}| \le const + \delta r, \quad 0<\delta\ll 1.$$

From (3.19) and (3.20) for each vector $\nu_t = (\nu_t^1, \ldots, \nu_t^d)$ we have

$$\begin{aligned}\sum_{i,j=1}^{d} (\sigma^2)_{ij,t}\nu_t^i\nu_t^j &= \sum_{i,j=1}^{d} (\sigma^0+\delta h)(\sigma^0+\delta h)'_{ij,t}\nu_t^i\nu_t^j \\ &= \sum_{i,j=1}^{d} (\sigma^0(\sigma^0)')_{ij,t}\nu_t^i\nu_t^j + \delta \sum_{i,j=1}^{d} (\sigma^0 h')_{ij,t}\nu_t^i\nu_t^j \\ &\quad + \delta \sum_{i,j=1}^{d} (h(\sigma^0)')_{ij,t}\nu_t^i\nu_t^j + \delta^2 \sum_{i,j=1}^{d} (hh')_{ij,t}\nu_t^i\nu_t^j. \end{aligned} \tag{3.21}$$

Note now that the elements of matrices σ^0 and h are bounded. Hence choosing δ sufficiently small we get

$$\max\left(\delta|(\sigma^0 h')_{ij,t}|, \delta|(h(\sigma^0))_{ij,t}|, \delta^2|(hh')_{ij,t}|\right) \le \frac{\varepsilon}{3}\,.$$

Therefore from (3.19) and (3.21) we get

$$\sum_{i,j=1}^{d} \sigma^2_{ij,t}\nu_t^i\nu_t^j \ge (c - const\cdot\varepsilon)\sum_{i,j=1}^{d} |\nu_t^i|^2 \quad \text{for each } \varepsilon > 0.$$

Proposition is proved. □

Consider the set of processes $\{R^\sigma(\text{or } X^\sigma),\ \sigma \in A_\delta\}$, which represents the misspecified of asset price model.

Define the class of admissible trading strategies $\Theta = \Theta(\sigma^0)$.

Proposition 3.3. *For each R^d-valued F-predictable process $\theta = (\theta_t)_{0\le t\le T}$ and for each $\sigma \in A_\delta$, $\delta > 0$,*

$$aE\int_0^T |\theta_t|^2 dC_t \le E\int_0^T \theta_t'\sigma_t d\langle M\rangle_t\sigma_t'\theta_t = E\int_0^T \theta_t'\sigma_t\sigma_t'\theta_t dC_t \le AE\int_0^T |\theta_t|^2 dC_t,$$

where the constants a, A are such that $0 < a \le A < \infty$, and the parameter $\delta > 0$ is sufficiently small.

Proof. Remember that $d\langle M\rangle_t = d\langle M^i, M^j\rangle_t = I_{ij}^{d\times d}dC_t$. Hence

$$E\int_0^T \theta_t'\sigma_t d\langle M\rangle_t\sigma_t'\theta_t = E\int_0^T \theta_t'\sigma_t\sigma_t'\theta_t\, dC_t.$$

Further, since $\sigma = \sigma^0 + \delta h$ and elements of matrices σ^0 and h are bounded, then the same is true for the elements of matrix σ with $0 \le \delta \le const$. Thus using the inequality $ab \le 2(a^2 + b^2)$ we get

$$E\int_0^T \theta_t'\sigma_t\sigma_t'\theta_t\, dC_t \le AE\int_0^T |\theta_t|^2 dC_t.$$

On the other hand, by Proposition 3.2 the matrix $\sigma^2 = \sigma\sigma'$ is uniformly elliptic for sufficiently small δ, which yields the first inequality. □

Definition 3.1. The class $\Theta = \Theta(\sigma^0)$ is a class of R^d-valued F-predictable processes $\theta = (\theta_t)_{0\le t\le T}$ such that

$$E\int_0^T |\theta_t|^2 dC_t < \infty. \tag{3.22}$$

Let $\theta \in \Theta$ be the dollar amount (rather than the number of shares) invested in the stock X^σ, $\sigma \in A_\delta$. Then for each $\sigma \in A_\delta$ the trading gains induced by the self-financing portfolio strategy associated to θ has the form

$$G_t(\sigma,\theta) = \int_0^t \theta'_s\, dR^\sigma_s, \quad 0 \le t \le T, \tag{3.23}$$

where $R^d = (R^d_t)_{0\le t\le T}$ is the yield process given by (3.10).

Introduce the condition:

(c.2) There exists ELMM $\overline{Q}$ such that the density process $z = z^{\overline{Q}}$ satisfies the reverse Hölder inequality $R_2(P)$, see definition in RSch [33].

It is well-known that under the conditions (c.1) and (c.2) the density process $\widetilde{z} = (\widetilde{z}_t)_{\le t\le T}$ of the variance-optimal ELMM satisfies $R_2(P)$ as well, see Dolean et al. [8].

Now under the conditions (c.1) and (c.2) the r.v. $G_T(\sigma,\theta) \in L^2(P)$, $\forall \sigma \in A_\delta$, and the space $G_T(\sigma,\Theta)$ is closed in $L^2(P)$, $\forall\sigma \in A_\delta$ (see, e.g., Theorem 2 of RSch [33]).

A contingent claim is an $\mathcal{F}_T$-measurable square-integrable r.v. H, which models the pay-off from a financial product at the maturity date T.

The problem we are interested in is to find the robust hedging strategy for a contingent claim H in the above described incomplete financial market model with misspecified asset price process X^σ, $\sigma \in A_\delta$, using mean-variance approach.

For each $\sigma \in A_\delta$, the total loss of a hedger, who starts with the initial capital x, uses the strategy θ, believes that the stock price process follows X^σ, and has to pay a random amount H at the date T, is H-x-$G_T(\sigma,\theta)$.

Denote

$$\mathcal{J}(\sigma,\theta) := E(H - x - G_T(\sigma,\theta))^2. \tag{3.24}$$

One setting of the robust mean-variance hedging problem consist in solving the optimization problem

$$\text{minimize} \sup_{\sigma\in A_\delta} \mathcal{J}(\sigma,\theta) \;\; \text{over all strategies} \;\; \theta \in \Theta. \tag{3.25}$$

We "slightly" change this problem using the approach developed in Toronjadze [41] which based on the following approximation

$$\begin{aligned}
\sup_{\sigma\in A_\delta} \mathcal{J}(\sigma,\theta) &= \exp\big\{\sup_{h\in\mathcal{H}} \ln \mathcal{J}(\sigma^0 + \delta h,\theta)\big\} \\
&\simeq \exp\left\{\sup_{h\in\mathcal{H}}\left[\ln \mathcal{J}(\sigma^0,\theta) + \delta\,\frac{D\mathcal{J}(\sigma^0,h,\theta)}{\mathcal{J}(\sigma^0,\theta)}\right]\right\} \\
&= \mathcal{J}(\sigma^0,\theta)\exp\left\{\delta\sup_{h\in\mathcal{H}} \frac{D\mathcal{J}(\sigma^0,h,\theta)}{\mathcal{J}(\sigma^0,\theta)}\right\},
\end{aligned}$$

where

$$D\mathcal{J}(\sigma^0,h,\theta) := \frac{d}{d\delta}\,\mathcal{J}(\sigma^0+\delta h,\theta)|_{\delta=0} = \lim_{\delta\to 0}\frac{\mathcal{J}(\sigma^0+\delta h,\theta) - \mathcal{J}(\sigma^0,\theta)}{\delta}\,,$$

is the Gateaux differential of the functional $\mathcal{J}$ at the point σ^0 in the direction h.

Approximate (in leading order δ) the optimization problem (3.25) by the problem

$$\text{minimize}\ \ \mathcal{J}(\sigma^0,\theta)\exp\left\{\delta \sup_{h\in\mathcal{H}} \frac{D\mathcal{J}(\sigma^0,h,\theta)}{\mathcal{J}(\sigma^0,\theta)}\right\}$$
$$\text{over all strategies}\ \ \theta\in\Theta. \tag{3.26}$$

Note that each solution θ^* of the problem (3.26) minimizes $\mathcal{J}(\sigma^0,\theta)$ under the constraint

$$\sup_{h\in\mathcal{H}} \frac{D\mathcal{J}(\sigma^0,h,\theta)}{\mathcal{J}(\sigma^0,\theta)} \le c := \sup_{h\in\mathcal{H}} \frac{D\mathcal{J}(\sigma^0,h,\theta^*)}{\mathcal{J}(\sigma^0,\theta^*)}\,.$$

This characterization of an optimal strategy θ^* of the problem (3.26) leads to the

Definition 3.2. The trading strategy $\theta^*\in\Theta$ is called optimal mean-variance robust trading strategy against the class of alternatives $\mathcal{H}$ if it is a solution of the optimization problem

$$\text{minimize}\ \ \mathcal{J}(\sigma^0,\theta)\ \ \text{over all strategies}\ \ \theta\in\Theta,\ \ \text{subject to constraint}$$
$$\sup_{h\in\mathcal{H}} \frac{D\mathcal{J}(\sigma^0,h,\theta)}{\mathcal{J}(\sigma^0,\theta)} \le c, \tag{3.27}$$

where c is some generic constant.

Remark 3.2. In contrast to "mean-variance robust" trading strategy which associates with optimization problem (3.25) and control theory, we find the "optimal mean-variance robust" strategy in the sense of Definition 3.2. Such approach and term are common in robust statistics theory (see, e.g., Hampel et al. [12], Rieder [34]).

Does the suggested approach provide "good" approximation? Consider the case.

Diffusion model with zero drift. Let a standard Wiener process $w=(w_t)_{0\le t\le T}$ be given on the complete probability space $(\Omega,\mathcal{F},P)$. Denote by $F^w=(\mathcal{F}_t^w,\,0\le t\le T)$ the P-augmentation of the natural filtration $\mathcal{F}_t^w=\sigma(w_s,\,0\le s\le t)$, $0\le t\le T$, generated by w.

Let the stock price process be modeled by the equation

$$dX_t^\sigma = X_t^\sigma\cdot\sigma_t\,dw_t,\ \ X_0^\sigma>0,\ \ 0\le t\le T,$$

where $\sigma\in A_\delta$ with

$$\int_0^T (\sigma_t^0)^2\,dt<\infty$$

and $h\in \mathrm{Ball}_{L_\infty(dt\times dP)}(0,r)$, $0<r<\infty$. All considered processes are real-valued.

Denote by R^σ the yield process, i.e.,

$$dR_t^\sigma=\sigma_t\,dw_t,\ \ R_0^\sigma=0,\ \ 0\le t\le T.$$

The wealth at maturity T, with the initial endowment x, is equal to

$$V_T^{x,\theta}(\sigma)=x+\int_0^T\theta_t\,dR_t^\sigma.$$

Let, further, the contingent claim H be $\mathcal{F}_T^w$-measurable P-square-integrable r.v.

Consider the optimization problem (3.25). It is easy to see that if $\sigma \in A_\delta$; then

$$\sigma_t^0 - \delta r \le \sigma_t \le \sigma_t^0 + \delta r, \quad 0 \le t \le T, \quad P\text{-}a.s.,$$

By the martingale representation theorem

$$H = EH + \int_0^T \varphi_t^H \, dw_t, \quad P\text{-}a.s.,$$

where φ^H is the F^w-predictable process with

$$E \int_0^T (\varphi_t^H)^2 \, dt < \infty. \tag{3.28}$$

Hence

$$E\big(H - V_T^{x,\theta}(\sigma)\big)^2 = (EH - x)^2 + E \int_0^T (\varphi_t^H - \sigma_t \theta_t)^2 \, dt.$$

From this it directly follows that the process

$$\begin{aligned} \sigma_t^*(\theta) &= (\sigma_t^0 - \delta r) I_{\{\frac{\varphi_t^H}{\theta_t} \ge \sigma_t^0\}} I_{\{\theta_t \ne 0\}} \\ &\quad + (\sigma_t^0 + \delta r) I_{\{\frac{\varphi_t^H}{\theta_t} < \sigma_t^0\}} I_{\{\theta_t \ne 0\}}, \quad 0 \le t \le T, \end{aligned} \tag{3.29}$$

is a solution of the optimization problem

$$\text{maximize} \ \ E\big(H - V_T^{x,\theta}(\sigma)\big)^2 \ \ \text{over all} \ \ \sigma \in A_\delta, \ \ \text{with a given} \ \ \theta \in \Theta.$$

It remains to minimize (w.r.t. θ) the expression

$$E \int_0^T \big(\varphi_t^H - \sigma_t^*(\theta)\theta_t\big)^2 \, dt.$$

From (3.29) it easily follows that the equation (w.r.t. θ)

$$\varphi_t^H - \sigma_t^*(\theta)\theta_t = 0,$$

has no solution, but

$$\theta_t^* = \frac{\varphi_t^H}{\sigma_t^0} I_{\{\sigma_t^0 \ne 0\}}, \quad 0 \le t \le T, \tag{3.30}$$

solves problem. We assume that $0/0 := 0$.

Consider now the optimization problem (3.27).

For each fixed h

$$\begin{aligned} J(\sigma, \theta) &= E\Big(H - x - \int_0^T \theta_t \, dR_t^\sigma\Big)^2 \\ &= E\bigg(H - x - \int_0^T \theta_t \sigma_t^0 \, dw_t - \delta \int_0^T \theta_t h_t \, dw_t\bigg)^2 \\ &= J(\sigma^0, \theta) - 2\delta E\bigg[\Big(EH - x + \int_0^T \big(\varphi_t^H - \theta_t \sigma_t^0\big) \, dw_t\Big) \int_0^T \theta_t h_t \, dw_t\bigg] \\ &\quad + \delta^2 E \int_0^T \theta_t^2 h_t^2 \, dt, \end{aligned}$$

and hence

$$DJ(\sigma^0, h; \theta) = 2E \int_0^T \big(\theta_t \sigma_t^0 - \varphi_t^H\big)\theta_t h_t \, dt, \tag{3.31}$$

as follows from (3.28), the definition of the class $\mathcal{H}$ and the estimation

$$\Big(E \int_0^T \big(\theta_t \sigma_t^0 - \varphi_t^H\big)\theta_t h_t \, dt\Big)^2 \le E \int_0^T \big(\theta_t \sigma_t^0 - \theta_t^H\big)^2 \, dt \; E \int_0^T \theta_t^2 h_t^2 \, dt$$

$$\le \text{const} \cdot r^2 \bigg(E \int_0^T \theta_t^2 (\sigma_t^0)^2 \, dt + E \int_0^T (\varphi_t^H)^2 \, dt\bigg) E \int_0^T \theta_t^2 \, dt < \infty. \tag{3.32}$$

Since, further, $DJ(\sigma^0, h; \theta) = 0$ for $h \equiv 0$, using (3.32) we get

$$0 \le \sup_{h \in \mathcal{H}} DJ(\sigma^0, h; \theta) < \infty.$$

Hence we can take $0 \le c < \infty$ in problem (6). Now if we substitute θ^* from (3.30) into (3.31), we get $DJ(\sigma^0, h; \theta^*) = 0$ for each h, and thus

$$\frac{\sup\limits_{h \in \mathcal{H}} DJ(\sigma^0, h; \theta^*)}{J(\sigma^0, \theta^*)} = 0.$$

If we recall that $\theta^* = \arg \min\limits_{\theta \in \Theta_{A_\delta}} J(\sigma^0, \theta)$, we get that θ^* defined by (3.30) is a solution of this optimization problem as well.

Thus we prove that

(a) *the mean-variance robust trading strategy* $\theta^* = (\theta_t^*)_{0 \le t \le T}$ *for the optimization problem* (3.25) *is given by the formula*

$$\theta_t^* = \frac{\varphi_t^H}{\sigma_t^0} \, I_{\{\sigma_t^0 \ne 0\}};$$

(b) *at the same time this strategy is an optimal mean-variance robust trading strategy for the optimization problem* (3.27).

Hence in this case the suggested approach leads to the perfect solution of initial problem (3.25).

To solve the problem (3.27) in general case we need to calculate $D\mathcal{J}(\sigma^0, h, \theta)$. Suppose that $k = (k_t)_{0 \le t \le T} = (k_{i,t}, \, 1 \le i \le d)_{0 \le t \le T}$ from (3.10) is such that $|k_{i,t}| \le \mathit{const} \; \forall i, t$.

Following RSch [33] and GLP [11] introduce the probability measure $\widetilde{Q} \sim P$ on $\mathcal{F}_T$ by the relation

$$d\widetilde{Q} = \frac{\widetilde{z}_T}{\widetilde{z}_0} \, d\widetilde{P} \quad \Big(\text{and hence } \; d\widetilde{Q} = \frac{\widetilde{z}_T^2}{\widetilde{z}_0} \, dP\Big). \tag{3.33}$$

Using Proposition 5.1 of GLP [11] we can write

$$\mathcal{J}(\sigma, \theta) = E \, \frac{\widetilde{z}_T^2}{\widetilde{z}_0^2} \, \frac{\widetilde{z}_0^2}{\widetilde{z}_T^2} \bigg(H - x - \int_0^T \theta_t' \, dR_t^\sigma\bigg)^2 = \widetilde{z}_0^{-1} E^{\widetilde{Q}} \, \frac{\widetilde{z}_0^2}{\widetilde{z}_T^2} \bigg(H - x - \int_0^T \theta_t' \sigma_t \, dM_t^0\bigg)^2$$

$$= \widetilde{z}_0^{-1} E^{\widetilde{Q}} \bigg(\frac{H \widetilde{z}_0}{\widetilde{z}_T} - x - \int_0^T \psi_t^0(\sigma) \, d \, \frac{\widetilde{z}_0^2}{\widetilde{z}_t^2} - \int_0^T \big(\psi_t^1(\sigma)\big)' \, d \, \frac{M_t^0}{\widetilde{z}_t} \, \widetilde{z}_0\bigg)^2$$

$$:= \overline{\mathcal{J}}(\sigma, \psi^0, \psi^1) \quad (\text{or } \; \overline{\mathcal{J}}(\sigma, \psi) \; \text{ with } \; \psi = (\psi^0, \psi^1)'), \tag{3.34}$$

where

$$\psi_t^1 = \psi_t^1(\sigma) = \sigma_t'\theta_t,$$
$$\psi_t^0 = \psi_t^0(\sigma) = \int_0^t \theta_s'\sigma_s dM_s^0 - \theta_t'\sigma_t M_t^0, \quad 0 \le t \le T. \tag{3.35}$$

Thus

$$\psi_t^1(\sigma) = \psi_t^1(\sigma^0) + \delta\psi_t^1(h), \quad \psi_t^0(\sigma) = \psi_t^0(\sigma^0) + \delta\psi_t^0(h).$$

Let (following RSch [33])

$$\frac{H}{\widetilde{z}_T}\,\widetilde{z}_0 = E\left(\frac{H}{\widetilde{z}_T}\,\widetilde{z}_0\right) + \int_0^T (\psi_t^H)' dU_t + L_T, \tag{3.36}$$

be the Galtchouk–Kunita–Watanabe decomposition of r.v. $\frac{H}{\widetilde{z}_T}\,\widetilde{z}_0$ w.r.t. $R^{(d+1)}$-valued $\widetilde{Q}$-local martingale $U = \left(\frac{\widetilde{z}_0}{\widetilde{z}}, \frac{M^0}{\widetilde{z}}\,\widetilde{z}_0\right)'$, where $\psi^H = (\psi^{0,H}, \psi^{1,H})' \in L^2(U, \widetilde{Q})$, the space of F-predictable processes ψ such that $\int \psi' dU \in \mathcal{M}^2(\widetilde{Q})$ of martingale, and $L \in \mathcal{M}^2_{0,\mathrm{loc}}(\widetilde{Q})$, L is $\widetilde{Q}$-strongly orthogonal to U.

Remember that

$$\psi = (\psi^0, \psi^1)'. \tag{3.37}$$

Then, using (3.34), (3.35) and (3.36) we can write for each h

$$\begin{aligned}
&\mathcal{J}(\sigma^0 + \delta h, \psi) = \mathcal{J}(\sigma^0, \psi) + \delta \cdot 2\widetilde{z}_0^{-1} \\
&\times E^{\widetilde{Q}}\left\{\left[\left(x - E^{\widetilde{Q}}\,\frac{H}{\widetilde{z}_T}\,\widetilde{z}_0\right) - L_T + \int_0^T (\psi_t(\sigma^0) - \psi_t^H)' dU_t\right] \int_0^T (\overline{\psi}_t(h))' dU_t\right\} \\
&\quad + \delta^2 \widetilde{z}_0^{-1} E^{\widetilde{Q}}\left[\int_0^T (\overline{\psi}_t(h))' dU_t\right]^2 \\
&= \mathcal{J}(\sigma^0, \psi) + \delta \cdot 2\widetilde{z}_0^{-1} E^{\widetilde{Q}}\left[\int_0^T (\psi_t(\sigma^0) - \psi_t^H)' dU_t \int_0^T (\overline{\psi}_t(h))' dU_t\right] \\
&\quad + \delta^2 \widetilde{z}_0^{-1} E^{\widetilde{Q}}\left[\int_0^T (\overline{\psi}_t(h))' dU_t\right]^2.
\end{aligned} \tag{3.38}$$

Using Proposition 8 of RSch [33] we have for each h

$$\frac{\widetilde{z}_0}{\widetilde{z}_T}\,G_r(h, \Theta) = \left\{\int_0^T (\psi(h))' dU_t : \psi(h) \in L^2(U, \widetilde{Q})\right\},$$

and hence by (3.23)

$$\begin{aligned}
E^{\widetilde{Q}}\left(\int_0^T (\psi_t(h))' dU_t\right)^2 &= E^{\widetilde{Q}}\,\frac{\widetilde{z}_0^2}{\widetilde{z}_T^2}\,G_T^2(h, \theta) = \widetilde{z}_0 E G_T^2(h, \theta) = \widetilde{z}_0 E\left(\int_0^T \theta_t'\, dR_t^h\right)^2 \\
&= \widetilde{z}_0 E\left(\int_0^T \theta_t' h_t dM_t^0\right)^2 = \widetilde{z}_0 E\left(\int_0^T \theta_t' h_t d\langle M\rangle_t k_t + \int_0^T \theta_t' h_t dM_t\right)^2
\end{aligned}$$

$$\leq const\Bigg[E\bigg(\int_0^T |\theta_t' h_t d\langle M\rangle_t k_t|\bigg)^2 + E\bigg(\int_0^T \theta_t' h_t dM_t\bigg)^2\Bigg]$$

$$\leq const\ r^2 E\int_0^T |\theta_t|^2 dC_t < \infty. \tag{3.39}$$

Further,

$$\Bigg(E^{\widetilde{Q}}\bigg[\int_0^T (\psi_t(\sigma^0)-\psi_t^H)' dU_t \int_0^T (\psi_t(h))' dU_t\bigg]\Bigg)^2$$

$$\leq E^{\widetilde{Q}}\bigg(\int_0^T (\psi_t(\sigma^0)-\psi_t^H)' dU_t\bigg)^2 E^{\widetilde{Q}}\bigg(\int_0^T (\psi_t(h))' dU_t\bigg)^2 < \infty. \tag{3.40}$$

From these estimates we conclude that:

$$1)\quad D\overline{\mathcal{J}}(\sigma^0,h,\psi) = 2\widetilde{z}_0^{-1}E^{\widetilde{Q}}\int_0^T (\psi_t(\sigma^0)-\psi_t^H)' d\langle U\rangle_t \psi_t(h) < \infty, \tag{3.41}$$

thanks to (3.39).

2) $D\overline{\mathcal{J}}(\sigma^0,h,\psi)|_{h\equiv 0} = 0$, since $\psi(0)=0$ by (3.37) and (3.35).

Thus

$$\sup_{h\in\mathcal{H}} D\overline{\mathcal{J}}(\sigma^0,h,\psi) \geq 0. \tag{3.42}$$

3) From (3.40) and (3.39) we get

$$(D\overline{\mathcal{J}}(\sigma^0,h,\psi))^2 \leq const\ \widetilde{z}_0^{-2} r^2$$

$$\times E^{\widetilde{Q}}\int_0^T (\psi_t(\sigma^0)-\psi_t^H)' d\langle U\rangle_t (\psi_t(\sigma^0)-\psi_t^H) E\int_0^T |\theta_t|^2 dC_t < \infty.$$

Thus $|D\overline{\mathcal{J}}(\sigma^0,h,\psi)|$ is estimated by the expression which does not depend on h, and is equal to zero if we substitute $\psi_t(\sigma^0)\equiv\psi_t^H$, $0\leq t\leq T$.

Hence, by (3.42)

$$0 \leq \sup_{h\in\mathcal{H}} D\overline{\mathcal{J}}(\sigma^0,h,\psi)|_{\psi\equiv\psi^H} \leq \sup_{h\in\mathcal{H}} |D\overline{\mathcal{J}}(\sigma^0,h,\psi)|\big|_{\psi\equiv\psi^H} = 0 \tag{3.43}$$

Further, from (3.42) follows that we can take $c\in[0,\infty)$ in (3.27).

Now substituting $\psi\equiv\psi^H$ into $\overline{\mathcal{J}}(\sigma^0,\psi)$ and $D\overline{\mathcal{J}}(\sigma^0,h,\psi)$ we get

$$\overline{\mathcal{J}}(\sigma^0,\psi^H) = \min_{\psi}\overline{\mathcal{J}}(\sigma^0,\psi) = \widetilde{z}_0^{-1}(E^{\widetilde{P}}H - x)^2 + \widetilde{z}_0^{-1}E^{\widetilde{Q}}L_T^2$$

(see Lemma 5.1 of GLP [11]) and

$$\sup_{h\in\mathcal{H}}\frac{D\overline{\mathcal{J}}(\sigma^0,h,\psi^H)}{\overline{\mathcal{J}}(\sigma^0,\psi^H)} = 0.$$

Hence the constraint of problem (3.27) is satisfied.

Remark 3.3. If $x = E^{\widetilde{P}}H$ and $L_T\equiv 0$, then we get

$$\frac{D\overline{\mathcal{J}}(\sigma^0,h,\psi^H)}{\overline{\mathcal{J}}(\sigma^0,\psi^H)} = \frac{0}{0}$$

which is assumed to be zero, since if we consider the shifted risk functional $\widetilde{\mathcal{J}} = \overline{\mathcal{J}} + 1$, the optimization problem and the optimal trading strategy will not change, but $D\widetilde{\mathcal{J}}(\sigma^0, h, \psi^H) = D\overline{\mathcal{J}}(\sigma^0, h, \psi^H) = 0$ and $\widetilde{\mathcal{J}}(\sigma^0, \psi^H) = 1$.

Finally, using Proposition 8 of RSch [33] we arrive at the following

Theorem 3.1. *In Model* (3.10) *under conditions* $(\mathrm{c}.1)$ *and* $(\mathrm{c}.2)$ *the optimal mean-variance robust trading strategy* $($*in the sense of Definition* $3.1)$ *is given by the formula*

$$\theta_t^* = ((\sigma_t^0)')^{-1}[\psi_t^{1,H} + \zeta_t(V_t^* - (\psi_t^H)'U_t)], \quad 0 \le t \le T, \tag{3.44}$$

where

$$\psi_t^H = (\psi_t^{0,H}, \psi_t^{1,H})', \quad U_t = \left(\frac{\widetilde{z}_0}{\widetilde{z}_t}, \frac{M_t^0}{\widetilde{z}_t}\,\widetilde{z}_0\right)',$$

$$V_t^* = \frac{\widetilde{z}_0}{\widetilde{z}_t}\left(x + \int_0^t (\psi_t^H)' dU_t\right),$$

ψ_t^H *and* ζ_t *are given by the relations* (3.36) *and* (3.11), *respectively,* $\widetilde{z}_t$ *is defined in* (3.11).

REFERENCES

[1] M. Avellaneda, A. Levy and A. Paras, Pricing and hedging derivative securities in markets with uncertain volatilities. *Appl. Math. Finance* **2** (1995), 73–88.

[2] M. Avellaneda and A. Paras, Managing the volatility risk of portfolios of derivative securities: The Lagrangian uncertain volatility model. *Appl. Math. Finance* **3** (1996), 21–52.

[3] F. Biagini, P. Guasoni and M. Pratelli, Mean-variance hedging for stochastic volatility models. *INFORMS Applied Probability Conference* (*Ulm,* 1999). *Math. Finance* **10** (2000), No. 2, 109–123.

[4] I. Bajeux-Besnainou and J. Rochet, Dynamic spanning: Are options an appropriate instrument? *Math. Finance* **6** (1996), No. 1, 1–16.

[5] M. Chesney, R. J. Elliott, D. Madan and H. Yang, Diffusion coefficient estimation and asset pricing when risk premia and sensitivities are time varying. *Math. Finance* **3** (1993), No. 2, 85–99.

[6] R. Chitashvili, N. Lazrieva and T. Toronjadze, Asymptotic theory of M-estimators in general statistical models. Part I. On asymptotic behaviour of estimators under model perturbance. *Centrum voor Woskunde en Informatica Amsterdam,* Report BS-R9019, 1990.

[7] R. Chitashvili, N. Lazrieva and T. Toronjadze, Asymptotic theory of M-estimators in general statistical models. Part II. On asymptotic behaviour of estimators in the presence of nuisance parameter. *Centrum voor Woskunde en Informatica Amsterdam,* Report BS-R9020, 1990.

[8] F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer and C. Stricker, Weighted norm inequalities and hedging in incomplete markets. *Finance Stoch.* **1** (1997), No. 3, 181–227.

[9] D. Duffie and H. R. Richardson, Mean-variance hedging in continuous time. *Ann. Appl. Probab.* **1** (1991), No. 1, 1–15.

[10] H. Föllmer and D. Sondermann, Hedging of nonredundant contingent claims. *Contributions to mathematical economics,* 205–223, *North-Holland, Amsterdam,* 1986.

[11] C. Gourieroux, J. P. Laurent and H. Pham, Mean-variance hedging and numéraire. *Math. Finance* **8** (1998), No. 3, 179–200.

[12] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw and W. A. Stahel, Robust statistics. The approach based on influence functions. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York,* 1986.

[13] J. Hull and A. White, The pricing of options on assets with stochastic voltilities. *J. Finance* **3** (1987), 281–300.

[14] I. A. Ibragimov, R. Z. Has'minskiĭ, Statistical estimation. Asymptotic theory. (Translated from the Russian) *Applications of Mathematics,* 16. *Springer-Verlag, New York-Berlin,* 1981.

[15] J. Jacod and A. N. Shiryaev, Limit theorems for stochastic processes. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*], 288. *Springer-Verlag, Berlin,* 1987.

[16] N. V. Krylov, Controlled diffusion processes. (Translated from the Russian) *Applications of Mathematics,* 14. *Springer-Verlag, New York-Berlin,* 1980.

[17] Yu. Kutoyants, Identification of dynamical systems with small noise. *Mathematics and its Applications,* 300. *Kluwer Academic Publishers Group, Dordrecht,* 1994.

[18] J. P. Laurent and H. Pham, Dynamic programming and mean-variance hedging. *Finance Stoch.* **3** (1999), No. 1, 83–110.

[19] N. L. Lazrieva and T. A. Torondzhadze, Asymptotic properties of an M-estimate in a general statistical experiment scheme. (Russian) *Statistics and control of random processes* (Russian) (*Preila,* 1987), 105–112, *Nauka, Moscow,* 1989.

[20] N. Lazrieva and T. Toronjadze, Robust estimators in statistical models with filtration. Shrinking neighbourhoods. *Seminarberichte, Fachbereich Mathematik, Fernuniversität, Hagen, Germany* **48** (1994), 50-68.

[21] N. Lazrieva and T. Toronjadze, Robust estimators in statistical models associated with semimartingales. *Proc. A. Razmadze Math. Inst.* **118** (1998), 73–100.

[22] D. Li and W.-L. Ng, Optimal dynamic portfolio selection: multiperiod mean-variance formulation. *Math. Finance* **10** (2000), No. 3, 387–406.

[23] R. S. Liptser and A. N. Shiryayev, Statistics of random processes. I. General theory. (Translated from the Russian) *Applications of Mathematics,* Vol. 5. *Springer-Verlag, New York-Heidelberg,* 1977.

[24] M. Mania and R. Tevzadze, A semimartingale Bellman equation and the variance-optimal martingale measure. *Georgian Math. J.* **7** (2000), No. 4, 765–792.

[25] M. Mania, M. Santacroce and R. Tevzadze, A semimartingale backward equation related to the p-optimal martingale measure and the lower price of a contingent claim. *Stochastic processes and related topics* (*Siegmundsburg,* 2000), 189–212, *Stochastics Monogr.,* 12, *Taylor* & *Francis, London,* 2002.

[26] H. M. Markowitz, Mean-variance analysis in portfolio choice and capital markets. *Basil Blackwell, Oxford,* 1987.

[27] R. C. Merton, An analytic derivation of the efficient portfolio frontier. *J. Financ. Qualitat. Anal.* **7** (1972), 1851–1872.

[28] P. Monat and C. Stricker, Föllmer-Schweizer decomposition and mean-variance hedging for general claims. *Ann. Probab.* **23** (1995), No. 2, 605–628.

[29] M. Musiela and M. Rutkowski, Martingale methods in financial modelling. *Applications of Mathematics (New York),* 36. *Springer-Verlag, Berlin,* 1997.

[30] S. Pastorello, Diffusion coefficient estimation and asset pricing when risk premia and sensitivities are time varying: A comment. *Math. Finance* **6** (1996), No. 1, 111–117.

[31] H. Pham, T. Rheinländer and M. Schweizer, Mean-variance hedging for continuous processes: new proofs and examples. *Finance Stoch.* **2** (1998), No. 2, 173–198.

[32] E. Renault and N. Touzi, Option hedging and implied volatilities in a stochastic volatility model. *Math. Finance* **6** (1996), No. 3, 279–302.

[33] T. Rheinländer and M. Schweizer, On L^2-projections on a space of stochastic integrals. *Ann. Probab.* **25** (1997), No. 4, 1810–1831.

[34] H. Rieder, Robust asymptotic statistics. *Springer Series in Statistics. Springer-Verlag, New York,* 1994.

[35] W. J. Runggaldier and A. Zaccaria, A stochastic control approach to risk management under restricted information. *INFORMS Applied Probability Conference* (*Ulm,* 1999). *Math. Finance* **10** (2000), No. 2, 277–288.

[36] M. Schweizer, Mean-variance hedging for general claims. *Ann. Appl. Probab.* **2** (1992), No. 1, 171–179.

[37] M. Schweizer, Approximating random variables by stochastic integrals. *Ann. Probab.* **22** (1994), No. 3, 1536–1575.

[38] M. Schweizer, Approximation pricing and the variance-optimal martingale measure. *Ann. Probab.* **24** (1996), No. 1, 206–236.

[39] L. Scott, Option pricing when the variance changes randomly: theory, estimation and application. *J. Financ. Qualitat. Anal.* **22** (1987), 419–438

[40] K. R. Sircar and G. G. Papanicolaou, Stochastic volatility, smile and asymptotics. *Appl. Math. Finance* **6** (1999), No. 2, 107–145.

[41] T. Toronjadze, Optimal mean-variance robust hedging under asset price model misspecification. *Georgian Math. J.* **8** (2001), No. 1, 189–199.

MEAN-VARIANCE HEDGING UNDER PARTIAL INFORMATION

M. MANIA, R. TEVZADZE AND T. TORONJADZE

Abstract. We consider the mean-variance hedging problem under partial information. The underlying asset price process follows a continuous semimartingale, and strategies have to be constructed when only part of the information in the market is available. We show that the initial mean-variance hedging problem is equivalent to a new mean-variance hedging problem with an additional correction term, which is formulated in terms of observable processes. We prove that the value process of the reduced problem is a square trinomial with coefficients satisfying a triangle system of backward stochastic differential equations and the filtered wealth process of the optimal hedging strategy is characterized as a solution of a linear forward equation.

Key words and phrases: Backward stochastic differential equation, semimartingale market model, incomplete markets, mean-variance hedging, partial information

MSC 2010: 90A09, 60H30, 90C39

1. INTRODUCTION

In the problem of derivative pricing and hedging it is usually assumed that the hedging strategies have to be constructed by using all market information. However, in reality, investors acting in a market have limited access to the information flow. For example, an investor may observe just stock prices, but stock appreciation rates depend on some unobservable factors; one may think that stock prices can be observed only at some time intervals or up to some random moment before an expiration date, or an investor would like to price and hedge a contingent claim whose payoff depends on an unobservable asset, and he observes the prices of an asset correlated with the underlying asset. Besides, investors may not be able to use all available information even if they have access to the full market flow. In all such cases, investors are forced to make decisions based on only a part of the market information.

We study a mean-variance hedging problem under partial information when the asset price process is a continuous semimartingale and the flow of observable events do not necessarily contain all information on prices of the underlying asset.

We assume that the dynamics of the price process of the asset traded on the market is described by a continuous semimartingale $S = (S_t, t \in [0,T])$ defined on a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{A}_t, t \in [0,T]), P)$, satisfying the usual conditions, where $\mathcal{A} = \mathcal{A}_T$ and $T < \infty$ is the fixed time horizon. Suppose that the interest rate is equal to zero and the asset price process satisfies the structure condition; i.e., the process S admits the decomposition

$$S_t = S_0 + N_t + \int_0^t \lambda_u d\langle N \rangle_u, \quad \langle \lambda \cdot N \rangle_T < \infty \quad \text{a.s.,} \tag{1.1}$$

Published in *SIAM J. Control Optim.* **47** (2008), no. 5, 2381–2409.

where N is a continuous $\mathcal{A}$-local martingale and λ is an $\mathcal{A}$-predictable process.

Let G be a filtration smaller than $\mathcal{A}$:

$$G_t \subseteq \mathcal{A}_t \quad \text{for every} \quad t \in [0, T].$$

The filtration G represents the information that the hedger has at his disposal; i.e., hedging strategies have to be constructed using only information available in G.

Let H be a P-square integrable $\mathcal{A}_T$-measurable random variable, representing the payoff of a contingent claim at time T.

We consider the mean-variance hedging problem

$$\text{to minimize} \quad E[(X_T^{x,\pi} - H)^2] \quad \text{over all} \quad \pi \in \Pi(G), \tag{1.2}$$

where $\Pi(G)$ is a class of G-predictable S-integrable processes. Here $X_t^{x,\pi} = x + \int_0^t \pi_u dS_u$ is the wealth process starting from initial capital x, determined by the self-financing trading strategy $\pi \in \Pi(G)$.

In the case $G = \mathcal{A}$ of complete information, the mean-variance hedging problem was introduced by Föllmer and Sondermann [8] in the case when S is a martingale and then developed by several authors for a price process admitting a trend (see, e.g., [6], [12], [25], [26], [24], [10], [11]).

Asset pricing with partial information under various setups has been considered. The mean-variance hedging problem under partial information was first studied by Di Masi, Platen, and Runggaldier [3] when the stock price process is a martingale and the prices are observed only at discrete time moments. For general filtrations and when the asset price process is a martingale, this problem was solved by Schweizer [27] in terms of G-predictable projections. Pham [22] considered the mean-variance hedging problem for a general semimartingale model, assuming that the observable filtration contains the augmented filtration F^S generated by the asset price process S

$$F_t^S \subseteq G_t \quad \text{for every} \quad t \in [0, T]. \tag{1.3}$$

In this paper, using the variance-optimal martingale measure with respect to the filtration G and suitable Kunita–Watanabe decomposition, the theory developed by Gourieroux, Laurent, and Pham [10] and Rheinländer and Schweizer [23] to the case of partial information was extended.

If G is not containing F^S, then S is not a G-semimartingale and the problem is more involved. Let us introduce an additional filtration $F = (F_t, t \in [0, T])$, which is an augmented filtration generated by F^S and G.

Then the price process S is a continuous F-semimartingale, and the canonical decomposition of S with respect to the filtration F is of the form

$$S_t = S_0 + \int_0^t \widehat{\lambda}_u^F d\langle M\rangle_u + M_t, \tag{1.4}$$

where $\widehat{\lambda}^F$ is the F-predictable projection of λ and

$$M_t = N_t + \int_0^t [\lambda_u - \widehat{\lambda}_u^F] d\langle N\rangle_u$$

is a continuous F-local martingale. Besides $\langle M\rangle = \langle N\rangle$, and these brackets are F^S-predictable.

Throughout the paper we shall make the following assumptions:
(A) $\langle M\rangle$ is G-predictable and $d\langle M\rangle_t dP$ a.e. $\widehat{\lambda}^F = \widehat{\lambda}^G$; hence P-a.s. for each t

$$E(\lambda_t|F^S_{t-}\vee G_t) = E(\lambda_t|G_t);$$

(B) any G-martingale is an F-local martingale;
(C) the filtration G is continuous; i.e., all G-local martingales are continuous;
(D) there exists a martingale measure for S (on F_T) that satisfies the reverse Hölder condition.

Remark. It is evident that if $F^S \subseteq G$, then $\langle M\rangle$ is G-predictable. Besides, in this case $G = F$, and conditions (A) and (B) are satisfied.

We shall use the notation $\widehat{Y}_t$ for the process of the G-projection of Y. Condition (A) implies that

$$\widehat{S}_t = E(S_t|G_t) = S_0 + \int_0^t \widehat{\lambda}_u d\langle M\rangle_u + \widehat{M}_t.$$

Let

$$H_t = E(H|F_t) = EH + \int_0^t h_u dM_u + L_t$$

and

$$H_t = EH + \int_0^t h^G_u d\widehat{M}_u + L^G_t$$

be the Galtchouk–Kunita–Watanabe (GKW) decompositions of $H_t = E(H|F_t)$ with respect to local martingales M and $\widehat{M}$, where h and h^G are F-predictable processes and L and L^G are local martingales strongly orthogonal to M and $\widehat{M}$, respectively.

We show (Theorem 3.1) that the initial mean-variance hedging problem (1.2) is equivalent to the problem to minimize the expression

$$E\bigg[\bigg(x + \int_0^T \pi_u d\widehat{S}_u - \widehat{H}_T\bigg)^2 + \int_0^T \Big(\pi_u^2\left(1-\rho_u^2\right) + 2\pi_u\widetilde{h}_u\Big)\, d\langle M\rangle_u\bigg] \tag{1.5}$$

over all $\pi \in \Pi(G)$, where

$$\widetilde{h}_t = \widehat{h^G_t}\rho_t^2 - \widehat{h}_t \quad \text{and} \quad \rho_t^2 = \frac{d\langle \widehat{M}\rangle_t}{d\langle M\rangle_t}.$$

Thus, the problem (1.5), equivalent to (1.2), is formulated in terms of G-adapted processes. One can say that (1.5) is the mean-variance hedging problem under complete information with an additional correction term.

Let us introduce the value process of the problem (1.5):

$$\begin{aligned} V^H(t,x) = \operatorname*{ess\,inf}_{\pi\in\Pi(G)} E\bigg[&\bigg(x + \int_t^T \pi_u d\widehat{S}_u - \widehat{H}_T\bigg)^2 \\ &+ \int_t^T \Big[\pi_u^2\left(1-\rho_u^2\right) + 2\pi_u\widetilde{h}_u\Big]\, d\langle M\rangle_u \Big| G_t\bigg]. \end{aligned} \tag{1.6}$$

We show in Theorem 4.1 that the value function of the problem (1.5) admits a representation

$$V^H(t,x) = V_t(0) - 2V_t(1)x + V_t(2)x^2,$$

where the coefficients $V_t(0), V_t(1)$, and $V_t(2)$ satisfy a triangle system of backward stochastic differential equations (BSDEs). Besides, the filtered wealth process of the optimal hedging strategy is characterized as a solution of the linear forward equation

$$\widehat{X}^*_t = x - \int_0^t \frac{\rho_u^2\varphi_u(2) + \widehat{\lambda}_u V_u(2)}{1-\rho_u^2+\rho_u^2 V_u(2)} \widehat{X}^*_u d\widehat{S}_u + \int_0^t \frac{\rho_u^2\varphi_u(1) + \widehat{\lambda}_u V_u(1) + \widetilde{h}_u}{1-\rho_u^2+\rho_u^2 V_u(2)} d\widehat{S}_u. \quad (1.7)$$

Note that if $F^S \subseteq G$, then

$$\rho = 1, \quad \tilde{h} = 0, \quad \widehat{M} = M, \quad \text{and} \quad \widehat{S} = S. \quad (1.8)$$

In the case of complete information ($G = \mathcal{A}$), in addition to (1.8) we have $\widehat{\lambda} = \lambda$ and $\widehat{M} = N$, and (1.7) gives equations for the optimal wealth process from [19].

In section 5 we consider a diffusion market model, which consists of two assets S and η, where S_t is a state of a process being controlled and η_t is the observation process. Suppose that S_t and η_t are governed by

$$dS_t = \mu_t dt + \sigma_t dw_t^0,$$

$$d\eta_t = a_t dt + b_t dw_t,$$

where w^0 and w are Brownian motions with correlation ρ and the coefficients μ, σ, a, and b are $\mathcal{F}^\eta$-adapted. In this case $\mathcal{A}_t = \mathcal{F}_t = \mathcal{F}_t^{S,\eta}$, and the flow of observable events is $\mathcal{G}_t = \mathcal{F}_t^\eta$. As an application of Theorem 4.1 we also consider a diffusion market model with constant coefficients and assume that an investor observes the price process S only up to a random moment τ before the expiration date T. In this case we give an explicit solution of (1.2).

2. Main definitions and auxiliary facts

Denote by $\mathcal{M}^e(F)$ the set of equivalent martingale measures for S, i.e., the set of probability measures Q equivalent to P such that S is a F-local martingale under Q.

Let

$$\mathcal{M}_2^e(F) = \{Q \in \mathcal{M}^e(F) : EZ_T^2(Q) < \infty\},$$

where $Z_t(Q)$ is the density process (with respect to the filtration F) of Q relative to P. We assume that $\mathcal{M}_2^e(F) \neq \emptyset$.

***Remark* 2.1.** Note that $\mathcal{M}_2^e(\mathcal{A}) \neq \emptyset$ implies that $\mathcal{M}_2^e(F) \neq \emptyset$ (see Remark 2.1 from Pham [22].

It follows from (1.4) and condition (A), that the density process $Z_t(Q)$ of any element Q of $\mathcal{M}^e(F)$ is expressed as an exponential martingale of the form

$$\mathcal{E}_t(-\widehat{\lambda} \cdot M + L),$$

where L is a F-local martingale strongly orthogonal to M and $\mathcal{E}_t(X)$ is the Doleans–Dade exponential of X.

If the local martingale $Z_t^{min} = \mathcal{E}_t(-\widehat{\lambda} \cdot M)$ is a true martingale, $dQ^{min}/dP = Z_T^{min}$ defines the minimal martingale measure for S.

Recall that a measure Q satisfies the reverse Hölder inequality $R_2(P)$ if there exists a constant C such that

$$E\left(\frac{Z_T^2(Q)}{Z_\tau^2(Q)} | F_\tau\right) \leq C, \qquad P\text{-a.s.}$$

for every F-stopping time τ.

***Remark* 2.2.** If there exists a measure $Q \in \mathcal{M}^e(F)$ that satisfies the reverse Hölder inequality $R_2(P)$, then according to Theorem 3.4 of Kazamaki [14] the martingale $M^Q = -\widehat{\lambda} \cdot M + L$ belongs to the class BMO and hence $-\widehat{\lambda} \cdot M$ also belongs to BMO, i.e.,

$$E\left(\int_\tau^T \widehat{\lambda}_u^2 d\langle M\rangle_u | F_\tau\right) \leq \text{const} \tag{2.1}$$

for every stopping time τ. Therefore, it follows from Theorem 2.3 of [14] that $\mathcal{E}_t(-\widehat{\lambda} \cdot M)$ is a true martingale. So, condition (D) implies that the minimal martingale measure exists (but Z^{min} is not necessarily square integrable).

Let us make some remarks on conditions (B) and (C).

***Remark* 2.3.** Condition (B) is satisfied if and only if the σ-algebras $F_t^S \vee G_t$ and G_T are conditionally independent given G_t for all $t \in [0, T]$ (see Theorem 9.29 from Jacod [13]).

***Remark* 2.4.** Condition (C) is weaker than the assumption that the filtration F is continuous. The continuity of the filtration F and condition (B) imply the continuity of the filtration G, but the converse is not true in general. Note that filtrations F and F^S can be discontinuous. Recall that the continuity of a filtration means that all local martingales with respect to this filtration are continuous.

By μ^K we denote the Dolean measure of an increasing process K. For all unexplained notations concerning the martingale theory used below, we refer the reader to [5], [18], [13].

Let $\Pi(F)$ be the space of all F-predictable S-integrable processes π such that the stochastic integral

$$(\pi \cdot S)_t = \int_0^t \pi_u dS_u, \quad t \in [0, T],$$

is in the $\mathcal{S}^2$ space of semimartingales, i.e.,

$$E\left(\int_0^T \pi_s^2 d\langle M\rangle_s\right) + E\left(\int_0^T |\pi_s \widehat{\lambda}_s| d\langle M\rangle_s\right)^2 < \infty.$$

Denote by $\Pi(G)$ the subspace of $\Pi(F)$ of G-predictable strategies.

***Remark* 2.5.** Since $\widehat{\lambda} \cdot M \in BMO$ (see Remark 2.2), it follows from the proof of Theorem 2.5 of Kazamaki [14] that

$$E\left(\int_0^T |\pi_u \widehat{\lambda}_u| d\langle M\rangle_u\right)^2 = E\langle |\pi| \cdot M, |\widehat{\lambda}| \cdot M\rangle_T^2 \leq 2||\widehat{\lambda} \cdot M||_{\text{BMO}} E\int_0^T \pi^2 d\langle M\rangle_u < \infty.$$

Therefore, under condition (D) the G-predictable (resp., F-predictable) strategy π belongs to the class $\Pi(G)$ (resp., $\Pi(F)$) if and only if $E\int_0^T \pi_s^2 d\langle M\rangle_s < \infty$.

Define $J_T^2(F)$ and $J_T^2(G)$ as spaces of terminal values of stochastic integrals, i.e.,

$$J_T^2(F) = \{(\pi \cdot S)_T : \pi \in \Pi(F)\}.$$

$$J_T^2(G) = \{(\pi \cdot S)_T : \pi \in \Pi(G)\}.$$

For convenience we give some assertions from [4], which establishes necessary and sufficient conditions for the closedness of the space $J^2_T(F)$ in L^2.

Proposition 2.1. *Let S be a continuous semimartingale. Then the following assertions are equivalent:*

(1) *There is a martingale measure $Q \in \mathcal{M}^e(F)$, and $J^2_T(F)$ is closed in L^2.*
(2) *There is a martingale measure $Q \in \mathcal{M}^e(F)$ that satisfies the reverse Hölder condition $R_2(P)$.*
(3) *There is a constant C such that for all $\pi \in \Pi(F)$ we have*

$$||\sup_{t\le T}(\pi\cdot S)_t||_{L^2(P)} \le C||(\pi\cdot S)_T||_{L^2(P)}.$$

(4) *There is a constant c such that for every stopping time τ, every $A \in \mathcal{F}_\tau$, and every $\pi \in \Pi(F)$, with $\pi = \pi I_{]\tau,T]}$, we have*

$$||I_A - (\pi\cdot S)_T||_{L^2(P)} \ge cP(A)^{1/2}.$$

Note that assertion (4) implies that for every stopping time τ and for every $\pi \in \Pi(G)$ we have

$$E\bigg(\bigg(1+\int_\tau^T \pi_u dS_u\bigg)^2 \Big/ F_\tau\bigg) \ge c. \tag{2.2}$$

Now we recall some known assertions from the filtering theory. The following proposition can be proved similarly to [18].

Proposition 2.2. *If conditions* (A), (B)*, and* (C) *are satisfied, then for any continuous F-local martingale M, with $M_0 = 0$, and any G-local martingale m^G*

$$\widehat{M}_t = E(M_t|G_t) = \int_0^t \frac{d\widehat{\langle M, m^G\rangle}_u}{d\langle m^G\rangle_u} dm^G_u + L^G_t, \tag{2.3}$$

where L^G is a local martingale orthogonal to m^G.

It follows from this proposition that for any G-predictable, M-integrable process π and any G-martingale m^G

$$\begin{aligned}\langle\widehat{(\pi\cdot M)}, m^G\rangle_t &= \int_0^t \pi_u \frac{d\widehat{\langle M, m^G\rangle}_u}{d\langle m^G\rangle_u} d\langle m^G\rangle_u \\ &= \int_0^t \pi_u d\langle \widehat{M}, m^G\rangle_u = \langle \pi\cdot\widehat{M}, m^G\rangle_t.\end{aligned}$$

Hence, for any G-predictable, M-integrable process π

$$\widehat{(\pi\cdot M)}_t = E\left(\int_0^t \pi_s dM_s|G_t\right) = \int_0^t \pi_s d\widehat{M}_s. \tag{2.4}$$

Since π, λ, and $\langle M\rangle$ are G-predictable, from (2.4) we have

$$\widehat{(\pi\cdot S)}_t = E\left(\int_0^t \pi_u dS_u|G_t\right) = \int_0^t \pi_u d\widehat{S}_u, \tag{2.5}$$

where

$$\widehat{S}_t = S_0 + \int_0^t \widehat{\lambda}_u d\langle M\rangle_u + \widehat{M}_t.$$

3. Separation principle: The optimality principle

Let us introduce the value function of the problem (1.2) defined as

$$U^H(t,x) = \operatorname*{ess\,inf}_{\pi\in\Pi(G)} E\Big(\Big(x + \int_t^T \pi_u dS_u - H\Big)^2 |G_t\Big). \tag{3.1}$$

By the GKW decomposition

$$H_t = E(H|F_t) = EH + \int_0^t h_u dM_u + L_t \tag{3.2}$$

for a F-predictable, M-integrable process h and a local martingale L strongly orthogonal to M. We shall use also the GKW decompositions of $H_t = E(H|F_t)$ with respect to the local martingale $\widehat{M}$

$$H_t = EH + \int_0^t h_u^G d\widehat{M}_u + L_t^G, \tag{3.3}$$

where h^G is a F-predictable process and L^G is a F-local martingale strongly orthogonal to $\widehat{M}$.

It follows from Proposition 2.2 (applied for $m^G = \widehat{M}$) and Lemma A.1 that

$$\langle E(H|G_\cdot), \widehat{M}\rangle_t = \int_0^t \widehat{h_u^G} \rho_u^2 d\langle M\rangle_u. \tag{3.4}$$

We shall use the notation

$$\widetilde{h}_t = \widehat{h_t^G}\rho_t^2 - \widehat{h}_t. \tag{3.5}$$

Note that $\widetilde{h}$ belongs to the class $\Pi(G)$ by Lemma A.2.

Let us introduce now a new optimization problem, equivalent to the initial mean-variance hedging problem (1.2), to minimize the expression

$$E\Bigg[\Big(x + \int_0^T \pi_u d\widehat{S}_u - \widehat{H}_T\Big)^2 + \int_0^T \Big(\pi_u^2\left(1-\rho_u^2\right) + 2\pi_u\tilde{h}_u\Big)d\langle M\rangle_u\Bigg] \tag{3.6}$$

over all $\pi \in \Pi(G)$. Recall that $\widehat{S}_t = E(S_t|G_t) = S_0 + \int_0^t \widehat{\lambda}_u d\langle M\rangle_u + \widehat{M}_t$.

Theorem 3.1. *Let conditions* (A), (B), *and* (C) *be satisfied. Then the initial mean-variance hedging problem* (1.2) *is equivalent to the problem* (3.6). *In particular, for any* $\pi \in \Pi(G)$ *and* $t \in [0,T]$

$$E\Bigg[\Big(x + \int_t^T \pi_u dS_u - H\Big)^2 |G_t\Bigg] = E\Bigg[\Big(H - \widehat{H}_T\Big)^2 |G_t\Bigg]$$
$$+E\Bigg[\Big(x + \int_t^T \pi_u d\widehat{S}_u - \widehat{H}_T\Big)^2 + \int_t^T \Big(\pi_u^2\left(1-\rho_u^2\right) + 2\pi_u\tilde{h}_u\Big)d\langle M\rangle_u |G_t\Bigg]. \tag{3.7}$$

Proof. We have

$$\begin{aligned}
&E\left[\left(x+\int_t^T \pi_u dS_u - H\right)^2 |G_t\right] \\
&\quad = E\left[\left(x+\int_t^T \pi_u d\widehat{S}_u - H + \int_t^T \pi_u d\left(M_u - \widehat{M}_u\right)\right)^2 |G_t\right] \\
&\quad = E\left[\left(x+\int_t^T \pi_u d\widehat{S}_u - H\right)^2 |G_t\right] \\
&\qquad + \ 2E\left[\left(x+\int_t^T \pi_u d\widehat{S}_u - H\right)\left(\int_t^T \pi_u d\left(M_u - \widehat{M}_u\right)\right)|G_t\right] \\
&\qquad + \ E\left[\left(\int_t^T \pi_u d\left(M_u - \widehat{M}_u\right)\right)^2 |G_t\right] = I_1 + 2I_2 + I_3.
\end{aligned} \tag{3.8}$$

It is evident that

$$I_1 = E\left[\left(x+\int_t^T \pi_u d\widehat{S}_u - \widehat{H}_T\right)^2 |G_t\right] + E\left[\left(H - \widehat{H}_T\right)^2 |G_t\right]. \tag{3.9}$$

Since π, $\widehat{\lambda}$, and $\langle\widehat{M}\rangle$ are G_T-measurable and the σ-algebras $F_t^S \vee G_t$ and G_T are conditionally independent given G_t (see Remark 2.3), it follows from (2.4) that

$$\begin{aligned}
&E\left[\int_t^T \pi_u \widehat{\lambda}_u d\langle M\rangle_u \int_t^T \pi_u d\left(M_u - \widehat{M}_u\right)|G_t\right] \\
&\quad = E\left[\int_t^T \pi_u \widehat{\lambda}_u d\langle M\rangle_u \int_0^T \pi_u d\left(M_u - \widehat{M}_u\right)|G_t\right] \\
&\qquad - E\left[\int_t^T \pi_u \widehat{\lambda}_u d\langle M\rangle_u \int_0^t \pi_u d\left(M_u - \widehat{M}_u\right)|G_t\right] \\
&\quad = E\left[\int_t^T \pi_u \widehat{\lambda}_u d\langle M\rangle_u E\left(\int_0^T \pi_u d\left(M_u - \widehat{M}_u\right)|G_T\right)|G_t\right] \\
&\qquad - E\left[\int_t^T \pi_u \widehat{\lambda}_u d\langle M\rangle_u |G_t\right] E\left[\int_0^t \pi_u d\left(M_u - \widehat{M}_u\right)|G_t\right] \\
&\quad = 0.
\end{aligned} \tag{3.10}$$

On the other hand, by using decomposition (3.2), equality (3.4), properties of square characteristics of martingales, and the projection theorem, we obtain

$$\begin{aligned}
&E\left[H\int_t^T \pi_u d\left(M_u - \widehat{M}_u\right)|G_t\right] \\
&\quad = E\left[H\int_t^T \pi_u dM_u |G_t\right] - E\left[\widehat{H}_T \int_t^T \pi_u d\widehat{M}_u |G_t\right] \\
&\quad = E\left[\int_t^T \pi_u d\langle M, E(H|F_\cdot)\rangle_u |G_t\right] - E\left[\int_t^T \pi_u d\langle \widehat{H}, \widehat{M}\rangle_u |G_t\right]
\end{aligned}$$

$$= E\Big[\int_t^T \pi_u h_u d\langle M\rangle_u|G_t\Big] - E\Big[\int_t^T \pi_u \widehat{h_u^G}\rho_u^2 d\langle M\rangle_u|G_t\Big]$$

$$= E\Big[\int_t^T \pi_u \left(\widehat{h}_u - \widehat{h_u^G}\rho_u^2\right) d\langle M\rangle_u|G_t\Big] = -E\Big[\int_t^T \pi_u \widetilde{h}_u d\langle M\rangle_u|G_t\Big]. \quad (3.11)$$

Finally, it is easy to verify that

$$2E\Big[\int_t^T \pi_u \widehat{M}_u \int_t^T \pi_u d\left(M_u - \widehat{M}_u\right)|G_t\Big] + E\Big[\Big(\int_t^T \pi_u d\left(M_u - \widehat{M}_u\right)\Big)^2|G_t\Big]$$

$$= E\Big[\Big(\int_t^T \pi_u^2 d\langle M\rangle_u - \int_t^T \pi_u^2 d\langle\widehat{M}\rangle_u\Big)|G_t\Big] = E\Big[\int_t^T \pi_u^2\left(1-\rho_u^2\right) d\langle M\rangle_u|G_t\Big]. \quad (3.12)$$

Therefore (3.8), (3.9), (3.10), (3.11), and (3.12) imply the validity of equality (3.7). □

Thus, it follows from Theorem 3.1 that the optimization problems (1.2) and (3.6) are equivalent. Therefore it is sufficient to solve the problem (3.6), which is formulated in terms of G-adapted processes. One can say that (3.6) is a mean-variance hedging problem under complete information with a correction term and can be solved by using methods for complete information.

Let us introduce the value process of the problem (3.6)

$$V^H(t,x) = \operatorname*{ess\,inf}_{\pi\in\Pi(G)} E\Bigg[\Big(x + \int_t^T \pi_u d\widehat{S}_u - \widehat{H}_T\Big)^2 + \int_t^T \Big[\pi_u^2\left(1-\rho_u^2\right) + 2\pi_u\tilde{h}_u\Big] d\langle M\rangle_u|G_t\Bigg]. \quad (3.13)$$

It follows from Theorem 3.1 that

$$U^H(t,x) = V^H(t,x) + E\big[(H-\widehat{H}_T)^2|G_t\big]. \quad (3.14)$$

The optimality principle takes in this case the following form.

Proposition 3.1 (optimality principle). *Let conditions* (A), (B) *and* (C) *be satisfied. Then*

(a) *for all $x \in R$, $\pi \in \Pi(G)$, and $s \in [0,T]$ the process*

$$V^H\Big(t, x + \int_s^t \pi_u d\widehat{S}_u\Big) + \int_s^t \Big[\pi_u^2\left(1-\rho_u^2\right) + 2\pi_u\tilde{h}_u)\Big] d\langle M\rangle_u$$

is a submartingale on $[s,T]$, admitting an right continuous with left limits (RCLL) modification.

(b) *π^* is optimal if and only if the process*

$$V^H\Big(t, x + \int_s^t \pi_u^* d\widehat{S}_u\Big) + \int_s^t \Big[(\pi_u^*)^2\left(1-\rho_u^2\right) + 2\pi_u^*\tilde{h}_u\Big] d\langle M\rangle_u$$

is a martingale.

This assertion can be proved in a standard manner (see, e.g., [7], [15]). The proof more adapted to this case one can see in [19].

Let

$$V(t,x) = \operatorname*{ess\,inf}_{\pi\in\Pi(G)} E\left[\left(x + \int_t^T \pi_u d\widehat{S}_u\right)^2 + \int_t^T \pi_u^2 \left(1-\rho_u^2\right) d\langle M\rangle_u | G_t\right]$$

and

$$V_t(2) = \operatorname*{ess\,inf}_{\pi\in\Pi(G)} E\left[\left(1 + \int_t^T \pi_u d\widehat{S}_u\right)^2 + \int_t^T \pi_u^2 \left(1-\rho_u^2\right) d\langle M\rangle_u | G_t\right].$$

It is evident that $V(t,x)$ (resp., $V_t(2)$) is the value process of the optimization problem (3.6) in the case $H = 0$ (resp., $H = 0$ and $x = 1$), i.e.,

$$V(t,x) = V^0(t,x) \quad \text{and} \quad V_t(2) = V^0(t,1).$$

Since $\Pi(G)$ is a cone, we have

$$\begin{aligned} V(t,x) &= x^2 \operatorname*{ess\,inf}_{\pi\in\Pi(G)} E\left[\left(1 + \int_t^T \frac{\pi_u}{x} d\widehat{S}_u\right)^2 + \int_t^T \left(\frac{\pi_u}{x}\right)^2 \left(1-\rho_u^2\right) d\langle M\rangle_u | G_t\right] \\ &= x^2 V_t(2). \end{aligned} \tag{3.15}$$

Therefore from Proposition 3.1 and equality (3.15) we have the following.

Corollary 3.1. (a) *The process*

$$V_t(2)\left(1 + \int_s^t \pi_u d\widehat{S}_u\right)^2 + \int_s^t (\pi_u)^2 (1-\rho_u^2) d\langle M\rangle_u,$$

$t \geq s$, *is a submartingale for all* $\pi \in \Pi(G)$ *and* $s \in [0,T]$.

(b) π^* *is optimal if and only if*

$$V_t(2)\left(1 + \int_s^t \pi_u^* d\widehat{S}_u\right)^2 + \int_s^t (\pi_u^*)^2 (1-\rho_u^2) d\langle M\rangle_u,$$

$t \geq s$, *is a martingale.*

Note that in the case $H = 0$ from Theorem 3.1 we have

$$\begin{aligned} &E\left[\left(1 + \int_t^T \pi_u dS_u\right)^2 | G_t\right] \\ &\qquad = E\left[\left(1 + \int_t^T \pi_u d\widehat{S}_u\right)^2 + \int_t^T \pi_u^2 \left(1-\rho_u^2\right) d\langle M\rangle_u | G_t\right] \end{aligned} \tag{3.16}$$

and, hence,

$$V_t(2) = U^0(t,1). \tag{3.17}$$

Lemma 3.1. *Let conditions* (A)–(D) *be satisfied. Then there is a constant* $1 \geq c > 0$ *such that* $V_t(2) \geq c$ *for all* $t \in [0,T]$ *a.s. and*

$$1 - \rho_t^2 + \rho_t^2 V_t(2) \geq c \quad \mu^{\langle M\rangle} a.e. \tag{3.18}$$

Proof. Let

$$V_t^F(2) = \operatorname*{ess\,inf}_{\pi \in \Pi(F)} E\left[\left(1 + \int_t^T \pi_u dS_u\right)^2 |F_t\right].$$

It follows from assertion (4) of Proposition 2.1 that there is a constant $c > 0$ such that $V_t^F(2) \geq c$ for all $t \in [0, T]$ a.s. Note that $c \leq 1$ since $V^F \leq 1$. Then by (3.17)

$$\begin{aligned} V_t(2) = U^0(t,1) &= \operatorname*{ess\,inf}_{\pi \in \Pi(G)} E\left[\left(1 + \int_t^T \pi_u dS_u\right)^2 |G_t\right] \\ &= \operatorname*{ess\,inf}_{\pi \in \Pi(G)} E\left[E\left(\left(1 + \int_t^T \pi_u dS_u\right)^2 |F_t\right)|G_t\right] \\ &\geq E(V_t^F(2)|G_t) \geq c. \end{aligned}$$

Therefore, since $\rho_t^2 \leq 1$ by Lemma A.1,

$$1 - \rho_t^2 + \rho_t^2 V_t(2) \geq 1 - \rho_t^2 + \rho_t^2 c \geq \inf_{r \in [0,1]} (1 - r + rc) = c. \qquad \square$$

4. BSDEs FOR THE VALUE PROCESS

Let us consider the semimartingale backward equation

$$Y_t = Y_0 + \int_0^t f(u, Y_u, \psi_u) d\langle m \rangle_u + \int_0^t \psi_u dm_u + L_t \tag{4.1}$$

with the boundary condition

$$Y_T = \eta, \tag{4.2}$$

where η is an integrable G_T-measurable random variable, $f : \Omega \times [0,T] \times R^2 \to R$ is $\mathcal{P} \times \mathcal{B}(R^2)$ measurable, and m is a local martingale. A solution of (4.1)–(4.2) is a triple (Y, ψ, L), where Y is a special semimartingale, ψ is a predictable m-integrable process, and L a local martingale strongly orthogonal to m. Sometimes we call Y alone the solution of (4.1)–(4.2), keeping in mind that $\psi \cdot m + L$ is the martingale part of Y.

Backward stochastic differential equations have been introduced in [1] for the linear case as the equations for the adjoint process in the stochastic maximum principle. The semimartingale backward equation, as a stochastic version of the Bellman equation in an optimal control problem, was first derived in [2]. The BSDE with more general nonlinear generators was introduced in [21] for the case of Brownian filtration, where the existence and uniqueness of a solution of BSDEs with generators satisfying the global Lifschitz condition was established. These results were generalized for generators with quadratic growth in [16], [17] for BSDEs driven by a Brownian motion and in [20], [28] for BSDEs driven by martingales. But conditions imposed in these papers are too restrictive for our needs. We prove here the existence and uniqueness of a solution by directly showing that the unique solution of the BSDE that we consider is the value of the problem.

In this section we characterize optimal strategies in terms of solutions of suitable semimartingale backward equations.

Theorem 4.1. *Let H be a square integrable F_T-measurable random variable, and let conditions* (A), (B), (C), *and* (D) *be satisfied. Then the value function of the problem* (3.6) *admits a representation*

$$V^H(t,x) = V_t(0) - 2V_t(1)x + V_t(2)x^2, \tag{4.3}$$

where the processes $V_t(0), V_t(1)$, and $V_t(2)$ satisfy the following system of backward equations:

$$\begin{aligned} Y_t(2) = Y_0(2) &+ \int_0^t \frac{\big(\psi_s(2)\rho_s^2 + \widehat{\lambda}_s Y_s(2)\big)^2}{1-\rho_s^2+\rho_s^2 Y_s(2)} d\langle M\rangle_s \\ &+ \int_0^t \psi_s(2) d\widehat{M}_s + L_t(2), \quad Y_T(2) = 1, \end{aligned} \tag{4.4}$$

$$\begin{aligned} Y_t(1) = Y_0(1) &+ \int_0^t \frac{\big(\psi_s(2)\rho_s^2 + \widehat{\lambda}_s Y_s(2)\big)\big(\psi_s(1)\rho_s^2 + \widehat{\lambda}_s Y_s(1) - \tilde{h}_s\big)}{1-\rho_s^2+\rho_s^2 Y_s(2)} d\langle M\rangle_s \\ &+ \int_0^t \psi_s(1) d\widehat{M}_s + L_t(1), \quad Y_T(1) = E(H|G_T), \end{aligned} \tag{4.5}$$

$$\begin{aligned} Y_t(0) = Y_0(0) &+ \int_0^t \frac{\big(\psi_s(1)\rho_s^2 + \widehat{\lambda}_s Y_s(1) - \tilde{h}_s\big)^2}{1-\rho_s^2+\rho_s^2 Y_s(2)} d\langle M\rangle_s \\ &+ \int_0^t \psi_s(0) d\widehat{M}_s + L_t(0), \quad Y_T(0) = E^2(H|G_T), \end{aligned} \tag{4.6}$$

where $L(2), L(1)$, and $L(0)$ are G-local martingales orthogonal to $\widehat{M}$.

Besides, the optimal filtered wealth process $\widehat{X}_t^{x,\pi^*} = x + \int_0^t \pi_u^* d\widehat{S}_u$ is a solution of the linear equation

$$\begin{aligned} \widehat{X}_t^* = x &- \int_0^t \frac{\rho_u^2\psi_u(2) + \widehat{\lambda}_u Y_u(2)}{1-\rho_u^2+\rho_u^2 Y_u(2)} \widehat{X}_u^* d\widehat{S}_u \\ &+ \int_0^t \frac{\psi_u(1)\rho_u^2 + \widehat{\lambda}_u Y_u(1) - \tilde{h}_u}{1-\rho_u^2+\rho_u^2 Y_u(2)} d\widehat{S}_u. \end{aligned} \tag{4.7}$$

Proof. Similarly to the case of complete information one can show that the optimal strategy exists and that $V^H(t,x)$ is a square trinomial of the form (4.3) (see, e.g., [19]). More precisely the space of stochastic integrals

$$J_{t,T}^2(G) = \left\{ \int_t^T \pi_u dS_u : \pi \in \Pi(G) \right\}$$

is closed by Proposition 2.1, since $\langle M\rangle$ is G-predictable.

Hence there exists optimal strategy $\pi^*(t,x) \in \Pi(G)$ and $U^H(t,x) = E[|H - x - \int_t^T \pi_u^*(t,x) dS_u|^2 | G_t]$.

Since $\int_t^T \pi_u^*(t,x) dS_u$ coincides with the orthogonal projection of $H - x \in L^2$ on the closed subspace of stochastic integrals, then the optimal strategy is linear with respect to x, i.e., $\pi_u^*(t,x) = \pi_u^0(t) + x\pi_u^1(t)$. This implies that the value function $U^H(t,x)$ is a square trinomial. It follows from the equality (3.14) that $V^H(t,x)$ is also a square trinomial, and it admits the representation (4.3).

Let us show that $V_t(0), V_t(1)$, and $V_t(2)$ satisfy the system (4.4)–(4.6). It is evident that

$$V_t(0) = V^H(t,0) = \operatorname*{ess\,inf}_{\pi\in\Pi(G)} E\Bigg[\left(\int_t^T \pi_u d\widehat{S}_u - \widehat{H}_T\right)^2 + \int_t^T [\pi_u^2\left(1-\rho_u^2\right) + 2\pi_u\tilde{h}_u] d\langle M\rangle_u | G_t\Bigg] \tag{4.8}$$

and

$$V_t(2) = V^0(t,1) = \operatorname*{ess\,inf}_{\pi\in\Pi(G)} E\left[\left(1+\int_t^T \pi_u d\widehat{S}_u\right)^2 + \int_t^T \pi_u^2\left(1-\rho_u^2\right) d\langle M\rangle_u | G_t\right]. \tag{4.9}$$

Therefore, it follows from the optimality principle (taking $\pi = 0$) that $V_t(0)$ and $V_t(2)$ are RCLL G-submartingales and

$$\begin{aligned} &V_t(2) \le E(V_T(2)|G_t) \le 1,\\ &V_t(0) \le E(E^2(H|G_T)|G_t) \le E(H^2|G_t). \end{aligned}$$

Since

$$V_t(1) = \frac{1}{2}(V_t(0) + V_t(2) - V^H(t,1)), \tag{4.10}$$

the process $V_t(1)$ is also a special semimartingale, and since $V_t(0) - 2V_t(1)x + V_t(2)x^2 = V^H(t,x) \ge 0$ for all $x \in R$, we have $V_t^2(1) \le V_t(0)V_t(2)$; hence

$$V_t^2(1) \le E\left(H^2|G_t\right).$$

Expressions (4.8), (4.9), and (3.13) imply that $V_T(0) = E^2(H|G_T)$, $V_T(2) = 1$, and $V^H(T,x) = (x - E(H|G_T))^2$. Therefore from (4.10) we have $V_T(1) = E(H|G_T)$, and $V(0), V(1)$, and $V(2)$ satisfy the boundary conditions.

Thus, the coefficients $V_t(i), i = 0,1,2$, are special semimartingales, and they admit the decomposition

$$V_t(i) = V_0(i) + A_t(i) + \int_0^t \varphi_s(i) d\widehat{M}_s + m_t(i), \quad i = 0,1,2, \tag{4.11}$$

where $m(0), m(1)$, and $m(2)$ are G-local martingales strongly orthogonal to $\widehat{M}$ and $A(0)$, $A(1)$, and $A(2)$ are G-predictable processes of finite variation.

There exists an increasing continuous G-predictable process K such that

$$\langle M\rangle_t = \int_0^t \nu_u dK_u, \quad A_t(i) = \int_0^t a_u(i) dK_u, \quad i = 0,1,2,$$

where ν and $a(i), i = 0,1,2$, are G-predictable processes.

Let $\widehat{X}_{s,t}^{x,\pi} \equiv x + \int_s^t \pi_u d\widehat{S}_u$ and

$$Y_{s,t}^{x,\pi} \equiv V^H\left(t, \widehat{X}_{s,t}^{x,\pi}\right) + \int_s^t \left[\pi_u^2\left(1-\rho_u^2\right) + 2\pi_u\tilde{h}_u\right] d\langle M\rangle_u.$$

Then by using (4.3), (4.11), and the Itô formula for any $t \geq s$ we have

$$\left(\widehat{X}^{x,\pi}_{s,t}\right)^2 = x + \int_s^t \left[2\pi_u \widehat{\lambda}_u \widehat{X}^{x,\pi}_{s,u} + \pi_u^2 \rho_u^2\right] d\langle M\rangle_u + 2\int_s^t \pi_u \widehat{X}^{x,\pi}_{s,u} d\widehat{M}_u \tag{4.12}$$

and

$$\begin{aligned} Y^{x,\pi}_{s,t} - V^H(s,x) = &\int_s^t \left[\left(\widehat{X}^{x,\pi}_{s,u}\right)^2 a_u(2) - 2\widehat{X}^{x,\pi}_{s,u} a_u(1) + a_u(0)\right] dK_u \\ &+ \int_s^t \left[\pi_u^2 \left(1-\rho_u^2+\rho_u^2 V_{u-}(2)\right) + 2\pi_u \widehat{X}^{x,\pi}_{s,u}\left(\widehat{\lambda}_u V_{u-}(2) + \varphi_u(2)\rho_u^2\right)\right. \\ &\left. - 2\pi_u\left(V_{u-}(1)\widehat{\lambda}_u + \varphi_u(1)\rho_u^2 - \tilde{h}_u\right)\right]\nu_u dK_u + m_t - m_s, \end{aligned} \tag{4.13}$$

where m is a local martingale.

Let

$$\begin{aligned} G(\pi,x) = G(\omega,u,\pi,x) = &\pi^2\left(1-\rho_u^2+\rho_u^2 V_{u-}(2)\right) + 2\pi x\left(\widehat{\lambda}_u V_{u-}(2) + \varphi_u(2)\rho_u^2\right) \\ &- 2\pi(V_{u-}(1)\widehat{\lambda}_u + \varphi_u(1)\rho_u^2 - \tilde{h}_u). \end{aligned}$$

It follows from the optimality principle that for each $\pi \in \Pi(G)$ the process

$$\int_s^t \left[\left(\widehat{X}^{x,\pi}_{s,u}\right)^2 a_u(2) - 2\widehat{X}^{x,\pi}_{s,u} a_u(1) + a_u(0)\right] dK_u + \int_s^t G\left(\pi_u, \widehat{X}^{x,\pi}_{s,u}\right)\nu_u dK_u \tag{4.14}$$

is increasing for any s on $s \leq t \leq T$, and for the optimal strategy π^* we have the equality

$$\int_s^t \left[\left(\widehat{X}^{x,\pi^*}_{s,u}\right)^2 a_u(2) - 2\widehat{X}^{x,\pi^*}_{s,u} a_u(1) + a_u(0)\right] dK_u = -\int_s^t G\left(\pi^*_u, \widehat{X}^{x,\pi^*}_{s,u}\right)\nu_u dK_u. \tag{4.15}$$

Since $\nu_u dK_u = d\langle M\rangle_u$ is continuous, without loss of generality one can assume that the process K is continuous (see [19] for details). Therefore, by taking in (4.14) $\tau_s(\varepsilon) = \inf\{t \geq s : K_t - K_s \geq \varepsilon\}$ instead of t, we have that for any $\varepsilon > 0$ and $s \geq 0$

$$\begin{aligned} &\frac{1}{\varepsilon}\int_s^{\tau_s(\varepsilon)} \left[\left(\widehat{X}^{x,\pi}_{s,u}\right)^2 a_u(2) - 2\widehat{X}^{x,\pi}_{s,u} a_u(1) + a_u(0)\right] dK_u \\ &\qquad \geq -\frac{1}{\varepsilon}\int_s^{\tau_s(\varepsilon)} G\left(\pi_u, \widehat{X}^{x,\pi}_{s,u}\right)\nu_u dK_u. \end{aligned} \tag{4.16}$$

By passing to the limit in (4.16) as $\varepsilon \to 0$, from Proposition B of [19] we obtain

$$x^2 a_u(2) - 2x a_u(1) + a_u(0) \geq -G(\pi_u, x)\nu_u, \qquad \mu^K\text{-a.e.},$$

for all $\pi \in \Pi(G)$. Similarly from (4.15) we have that μ^K-a.e.

$$x^2 a_u(2) - 2x a_u(1) + a_u(0) = -G(\pi^*_u, x)\nu_u$$

and hence

$$x^2 a_u(2) - 2x a_u(1) + a_u(0) = -\nu_u \operatorname*{ess\,inf}_{\pi\in\Pi(G)} G(\pi_u, x). \tag{4.17}$$

The infimum in (4.17) is attained for the strategy

$$\hat{\pi}_t = \frac{V_t(1)\widehat{\lambda}_t + \varphi_t(1)\rho_t^2 - \tilde{h}_t - x(V_t(2)\widehat{\lambda}_t + \varphi_t(2)\rho_t^2)}{1-\rho_t^2+\rho_t^2 V_t(2)}. \tag{4.18}$$

From here we can conclude that

$$\operatorname*{ess\,inf}_{\pi\in\Pi(G)} G(\pi_t, x) \ge G(\hat{\pi}_t, x)$$
$$= -\frac{\left(V_t(1)\widehat{\lambda}_t + \varphi_t(1)\rho_t^2 - \tilde{h}_t - x\left(V_t(2)\widehat{\lambda}_t + \varphi_t(2)\rho_t^2\right)\right)^2}{1-\rho_t^2+\rho_t^2 V_t(2)}. \tag{4.19}$$

Let $\pi_t^n = I_{[0,\tau_n[}(t)\hat{\pi}_t$, where $\tau_n = \inf\{t : |V_t(1)| \ge n\}$.

It follows from Lemmas A.2, 3.1, and A.3 that $\pi^n \in \Pi(G)$ for every $n \ge 1$ and hence

$$\operatorname*{ess\,inf}_{\pi\in\Pi(G)} G(\pi_t, x) \le G(\pi_t^n, x)$$

for all $n \ge 1$. Therefore

$$\operatorname*{ess\,inf}_{\pi\in\Pi(G)} G(\pi_t, x) \le \lim_{n\to\infty} G(\pi_t^n, x) = G(\hat{\pi}_t, x). \tag{4.20}$$

Thus (4.17), (4.19), and (4.20) imply that

$$x^2 a_t(2) - 2x a_t(1) + a_t(0)$$
$$= \nu_t \frac{(V_t(1)\widehat{\lambda}_t + \varphi_t(1)\rho_t^2 - \tilde{h}_t - x(V_t(2)\widehat{\lambda}_t + \varphi_t(2)\rho_t^2))^2}{1-\rho_t^2+\rho_t^2 V_t(2)}, \quad \mu^K\text{-a.e.}, \tag{4.21}$$

and by equalizing the coefficients of square trinomials in (4.21) (and integrating with respect to dK) we obtain

$$A_t(2) = \int_0^t \frac{(\varphi_s(2)\rho_s^2 + \widehat{\lambda}_s V_s(2))^2}{1-\rho_s^2+\rho_s^2 V_s(2)} d\langle M\rangle_s, \tag{4.22}$$

$$A_t(1) = \int_0^t \frac{(\varphi_s(2)\rho_s^2 + \widehat{\lambda}_s V_s(2))(\varphi_s(1)\rho_s^2 + \widehat{\lambda}_s V_s(1) - \tilde{h}_s)}{1-\rho_s^2+\rho_s^2 V_s(2)} d\langle M\rangle_s, \tag{4.23}$$

$$A_t(0) = \int_0^t \frac{(\varphi_s(1)\rho_s^2 + \widehat{\lambda}_s V_s(1) - \tilde{h}_s)^2}{1-\rho_s^2+\rho_s^2 V_s(2)} d\langle M\rangle_s, \tag{4.24}$$

which, together with (4.11), implies that the triples $(V(i), \varphi(i), m(i))$, $i = 0, 1, 2$, satisfy the system (4.4)–(4.6).

Note that $A(0)$ and $A(2)$ are integrable increasing processes and relations (4.22) and (4.24) imply that the strategy $\hat{\pi}$ defined by (4.18) belongs to the class $\Pi(G)$.

Let us show now that if the strategy $\pi^* \in \Pi(G)$ is optimal, then the corresponding filtered wealth process $\widehat{X}_t^{\pi^*} = x + \int_0^t \pi_u^* d\widehat{S}_u$ is a solution of (4.7).

By the optimality principle the process

$$Y_t^{\pi^*} = V^H\left(t, \widehat{X}_t^{\pi^*}\right) + \int_0^t \left[(\pi_u^*)^2\left(1-\rho_u^2\right) + 2\pi_u^*\tilde{h}_u\right] d\langle M\rangle_u$$

is a martingale. By using the Itô formula we have

$$Y_t^{\pi^*} = \int_0^t \left(\widehat{X}_u^{\pi^*}\right)^2 dA_u(2) - 2\int_0^t \widehat{X}_u^{\pi^*} dA_u(1) + A_t(0) + \int_0^t G\left(\pi_u^*, \widehat{X}_u^{\pi^*}\right) d\langle M\rangle_u + N_t,$$

where N is a martingale. Therefore by applying equalities (4.22), (4.23), and (4.24) we obtain

$$Y_t^{\pi^*} = \int_0^t \Bigg(\pi_u^* - \frac{V_u(1)\widehat{\lambda}_u + \varphi_u(1)\rho_u^2 - \tilde{h}_u}{1-\rho_u^2+\rho_u^2 V_u(2)} + \widehat{X}_u^{\pi^*} \frac{V_u(2)\widehat{\lambda}_u + \varphi_u(2)\rho_u^2}{1-\rho_u^2+\rho_u^2 V_u(2)} \Bigg)^2 \big(1-\rho_u^2+\rho_u^2 V_u(2)\big)\, d\langle M\rangle_u + N_t,$$

which implies that $\mu^{\langle M\rangle}$-a.e.

$$\pi_u^* = \frac{V_u(1)\widehat{\lambda}_u + \varphi_u(1)\rho_u^2 - \tilde{h}_u}{1-\rho_u^2+\rho_u^2 V_u(2)} - \widehat{X}_u^{\pi^*} \frac{\Big(V_u(2)\widehat{\lambda}_u + \varphi_u(2)\rho_u^2\Big)}{1-\rho_u^2+\rho_u^2 V_u(2)}.$$

By integrating both parts of this equality with respect to $d\widehat{S}$ (and adding then x to the both parts), we obtain that $\widehat{X}^{\pi^*}$ satisfies (4.7). $\square$

The uniqueness of the system (4.4)–(4.6) we shall prove under following condition (D*), stronger than condition (D).

Assume that

(D*)

$$\int_0^T \frac{\widehat{\lambda}_u^2}{\rho_u^2} d\langle M\rangle_u \le C.$$

Since $\rho^2 \le 1$ (Lemma A.1), it follows from (D*) that the mean-variance tradeoff of S is bounded, i.e.,

$$\int_0^T \widehat{\lambda}_u^2 d\langle M\rangle_u \le C,$$

which implies (see, e.g., Kazamaki [14]) that the minimal martingale measure for S exists and satisfies the reverse Hölder condition $R_2(P)$. So, condition (D*) implies condition (D). Besides, it follows from condition (D*) that the minimal martingale measure $\widehat{Q}^{min}$ for $\widehat{S}$

$$d\widehat{Q}^{min} = \mathcal{E}_T\bigg(-\frac{\widehat{\lambda}}{\rho^2}\cdot \widehat{M}\bigg)$$

also exists and satisfies the reverse Hölder condition. Indeed, condition (D*) implies that $\mathcal{E}_t(-2\frac{\widehat{\lambda}}{\rho^2}\cdot\widehat{M})$ is a G-martingale and hence

$$E\bigg(\mathcal{E}_{tT}^2\bigg(-\frac{\widehat{\lambda}}{\rho^2}\cdot\widehat{M}\bigg)|G_t\bigg) = E\bigg(\mathcal{E}_{tT}\bigg(-2\frac{\widehat{\lambda}}{\rho^2}\cdot\widehat{M}\bigg)e^{\int_t^T \frac{\widehat{\lambda}_u^2}{\rho_u^2}d\langle M\rangle_u} G_t\bigg) \le e^C.$$

Recall that the process Z belongs to the class D if the family of random variables $Z_\tau I_{(\tau\le T)}$ for all stopping times τ is uniformly integrable.

Theorem 4.2. *Let conditions* (A), (B), (C), *and* (D*) *be satisfied. If a triple* $(Y(0), Y(1), Y(2))$, *where* $Y(0)\in D$, $Y^2(1)\in D$, *and* $c\le Y(2)\le C$ *for some constants* $0<c<C$, *is a solution of the system* (4.4)–(4.6), *then such a solution is unique and coincides with the triple* $(V(0), V(1), V(2))$.

Proof. Let $Y(2)$ be a bounded strictly positive solution of (4.4), and let

$$\int_0^t \psi_u(2)d\widehat{M}_u + L_t(2)$$

be the martingale part of $Y(2)$.

Since $Y(2)$ solves (4.4), it follows from the Itô formula that for any $\pi \in \Pi(G)$ the process

$$Y_t^\pi = Y_t(2)\left(1 + \int_s^t \pi_u d\widehat{S}_u\right)^2 + \int_s^t \pi_u^2\left(1-\rho_u^2\right)d\langle M\rangle_u, \tag{4.25}$$

$t \geq s$, is a local submartingale.

Since $\pi \in \Pi(G)$, from Lemma A.1 and the Doob inequality we have

$$\begin{aligned} &E\sup_{t\leq T}\left(1+\int_0^t \pi_u d\widehat{S}\right)^2 \\ &\qquad \leq \text{const}\left(1 + E\int_0^T \pi_u^2\rho_u^2 d\langle M\rangle_u\right) + E\left(\int_0^T |\pi_u\widehat{\lambda}_u| d\langle M\rangle_u\right)^2 < \infty. \end{aligned} \tag{4.26}$$

Therefore, by taking in mind that $Y(2)$ is bounded and $\pi \in \Pi(G)$ we obtain

$$E\big(\sup_{s\leq u\leq T} Y_u^\pi\big)^2 < \infty,$$

which implies that $Y^\pi \in D$. Thus Y^π is a submartingale (as a local submartingale from the class D), and by the boundary condition $Y_T(2) = 1$ we obtain

$$Y_s(2) \leq E\left(\left(1+\int_s^T \pi_u d\widehat{S}_u\right)^2 + \int_s^T \pi_u^2\left(1-\rho_u^2\right)d\langle M\rangle_u | G_s\right)$$

for all $\pi \in \Pi(G)$ and hence

$$Y_t(2) \leq V_t(2). \tag{4.27}$$

Let

$$\tilde{\pi}_t = -\frac{\widehat{\lambda}_t Y_t(2) + \psi_t(2)\rho_t^2}{1-\rho_t^2+\rho_t^2 Y_t(2)}\mathcal{E}_t\left(-\frac{\widehat{\lambda}Y(2)+\psi(2)\rho^2}{1-\rho^2+\rho^2 Y(2)}\cdot\widehat{S}\right).$$

Since $1 + \int_0^t \tilde{\pi}_u d\widehat{S}_u = \mathcal{E}_t(-\frac{\widehat{\lambda}Y(2)+\psi(2)\rho^2}{1-\rho^2+\rho^2 Y(2)}\cdot\widehat{S})$, it follows from (4.4) and the Itô formula that the process $Y^{\tilde{\pi}}$ defined by (4.25) is a positive local martingale and hence a supermartingale. Therefore

$$Y_s(2) \geq E\left(\left(1+\int_s^T \tilde{\pi}_u d\widehat{S}_u\right)^2 + \int_s^T \tilde{\pi}_u^2\left(1-\rho_u^2\right)d\langle M\rangle_u | G_s\right). \tag{4.28}$$

Let us show that $\tilde{\pi}$ belongs to the class $\Pi(G)$.

From (4.28) and (4.27) we have for every $s \in [0,T]$

$$E\left(\left(1+\int_s^T \tilde{\pi}_u d\widehat{S}_u\right)^2 + \int_s^T \tilde{\pi}_u^2\left(1-\rho_u^2\right)d\langle M\rangle_u | G_s\right) \leq Y_s(2) \leq V_s(2) \leq 1 \tag{4.29}$$

and hence

$$E\Big(1+\int_0^T \tilde{\pi}_u d\widehat{S}_u\Big)^2 \le 1, \tag{4.30}$$

$$E\int_0^T \tilde{\pi}_u^2\left(1-\rho_u^2\right)d\langle M\rangle_u \le 1. \tag{4.31}$$

By (D^*) the minimal martingale measure $\widehat{Q}^{min}$ for $\widehat{S}$ satisfies the reverse Hölder condition, and hence all conditions of Proposition 2.1 are satisfied. Therefore the norm

$$E\Big(\int_0^T \tilde{\pi}_s^2\rho_s^2 d\langle M\rangle_s\Big)+E\Big(\int_0^T |\tilde{\pi}_s\widehat{\lambda}_s| d\langle M\rangle_s\Big)^2$$

is estimated by $E\big(1+\int_0^T \tilde{\pi}_u d\widehat{S}_u\big)^2$ and hence

$$E\int_0^T \tilde{\pi}_u^2\rho_u^2 d\langle M\rangle_u<\infty,\quad E\Big(\int_0^T |\tilde{\pi}_s\widehat{\lambda}_s| d\langle M\rangle_s\Big)^2<\infty.$$

It follows from (4.31) and the latter inequality that $\tilde{\pi}\in\Pi(G)$, and from (4.28) we obtain

$$Y_t(2)\ge V_t(2),$$

which together with (4.27) gives the equality $Y_t(2)=V_t(2)$.

Thus $V(2)$ is a unique bounded strictly positive solution of (4.4). Besides,

$$\int_0^t \psi_u(2)d\widehat{M}_u=\int_0^t \varphi_u(2)d\widehat{M}_u,\quad L_t(2)=m_t(2) \tag{4.32}$$

for all t, P-a.s.

Let $Y(1)$ be a solution of (4.5) such that $Y^2(1)\in D$. By the Itô formula the process

$$\begin{aligned}R_t=&\,Y_t(1)\mathcal{E}_t\Big(-\frac{\varphi(2)\rho^2+\widehat{\lambda}V(2)}{1-\rho^2+\rho^2V(2)}\cdot\widehat{S}\Big)\\&+\int_0^t \mathcal{E}_u\Big(-\frac{\varphi(2)\rho^2+\widehat{\lambda}V(2)}{1-\rho^2+\rho^2V(2)}\cdot\widehat{S}\Big)\frac{(\varphi_u(2)\rho_u^2+\widehat{\lambda}_uV_u(2))\tilde{h}_u}{1-\rho_u^2+\rho_u^2V_u(2)}d\langle M\rangle_u\end{aligned} \tag{4.33}$$

is a local martingale. Let us show that R_t is a martingale.

As was already shown, the strategy

$$\tilde{\pi}_u=\frac{\psi_u(2)\rho_u^2+\widehat{\lambda}_uY_u(2)}{1-\rho^2+\rho^2Y_u(2)}\mathcal{E}_u\Big(-\frac{\psi(2)\rho^2+\widehat{\lambda}Y(2)}{1-\rho^2+\rho^2Y(2)}\cdot\widehat{S}\Big)$$

belongs to the class $\Pi(G)$.

Therefore (see (4.26)),

$$E\sup_{t\le T}\mathcal{E}_t^2\Big(-\frac{\psi(2)\rho^2+\widehat{\lambda}Y(2)}{1-\rho^2+\rho^2Y(2)}\cdot\widehat{S}\Big)=E\sup_{t\le T}\Big(1+\int_0^t\tilde{\pi}_u d\widehat{S}\Big)^2<\infty, \tag{4.34}$$

and hence

$$Y_t(1)\mathcal{E}_t\Big(-\frac{\varphi(2)\rho^2+\widehat{\lambda}V(2)}{1-\rho^2+\rho^2V(2)}\cdot\widehat{S}\Big)\in D.$$

On the other hand, the second term of (4.33) is the process of integrable variation, since $\tilde{\pi} \in \Pi(G)$ and $\tilde{h} \in \Pi(G)$ (see Lemma A.2) imply that

$$E\int_0^T \left|\mathcal{E}_u\Big(-\frac{\varphi(2)\rho^2+\widehat{\lambda}V(2)}{1-\rho^2+\rho^2V(2)}\cdot\widehat{S}\Big)\frac{(\varphi_u(2)\rho_u^2+\widehat{\lambda}_uV_u(2))\tilde{h}_u}{1-\rho_u^2+\rho_u^2V_u(2)}\right|d\langle M\rangle_u$$
$$=E\int_0^T|\tilde{\pi}_u\tilde{h}_u|d\langle M\rangle_u \le E^{1/2}\int_0^T\tilde{\pi}_u^2d\langle M\rangle_uE^{1/2}\int_0^T\tilde{h}_u^2d\langle M\rangle_u<\infty.$$

Therefore, the process R_t belongs to the class D, and hence it is a true martingale. By using the martingale property and the boundary condition we obtain

$$Y_t(1)=E\bigg(\widehat{H}_T\mathcal{E}_{tT}\Big(-\frac{\varphi(2)\rho^2+\widehat{\lambda}V(2)}{1-\rho^2+\rho^2V(2)}\cdot\widehat{S}\Big)$$
$$+\int_t^T\mathcal{E}_{tu}\Big(-\frac{\varphi(2)\rho^2+\widehat{\lambda}V(2)}{1-\rho^2+\rho^2V(2)}\cdot\widehat{S}\Big)\frac{(\varphi_u(2)\rho_u^2+\widehat{\lambda}_uV_u(2))\tilde{h}_u}{1-\rho_u^2+\rho_u^2V_u(2)}d\langle M\rangle_u|G_t\bigg). \quad (4.35)$$

Thus, any solution of (4.5) is expressed explicitly in terms of $(V(2),\varphi(2))$ in the form (4.35). Hence the solution of (4.5) is unique, and it coincides with $V_t(1)$.

It is evident that the solution of (4.6) is also unique. □

***Remark* 4.1.** In the case $F^S \subseteq G$ we have $\rho_t=1, \tilde{h}_t=0$, and $\widehat{S}_t=S_t$, and (4.7) takes the form

$$\widehat{X}_t^*=x-\int_0^t\frac{\psi_u(2)+\widehat{\lambda}_uY_u(2)}{Y_u(2)}\widehat{X}_u^*dS_u+\int_0^t\frac{\psi_u(1)+\widehat{\lambda}_uY_u(1)}{Y_u(2)}dS_u.$$

Corollary 4.1. *In addition to conditions* (A)–(C) *assume that ρ is a constant and the mean-variance tradeoff $\langle\widehat{\lambda}\cdot M\rangle_T$ is deterministic. Then the solution of* (4.4) *is the triple* $(Y(2),\psi(2),L(2))$, *with* $\psi(2)=0, L(2)=0$, *and*

$$Y_t(2)=V_t(2)=\nu\left(\rho,1-\rho^2+\langle\widehat{\lambda}\cdot M\rangle_T-\langle\widehat{\lambda}\cdot M\rangle_t\right), \quad (4.36)$$

where $\nu(\rho,\alpha)$ is the root of the equation

$$\frac{1-\rho^2}{x}-\rho^2\ln x=\alpha. \quad (4.37)$$

Besides,

$$Y_t(1)=E\bigg(H\mathcal{E}_{tT}\Big(-\frac{\widehat{\lambda}V(2)}{1-\rho^2+\rho^2V(2)}\cdot\widehat{S}\Big)$$
$$+\int_t^T\mathcal{E}_{tu}\Big(-\frac{\widehat{\lambda}V(2)}{1-\rho^2+\rho^2V(2)}\cdot\widehat{S}\Big)\frac{\lambda_uV_u(2)\tilde{h}_u}{1-\rho^2+\rho^2V_u(2)}d\langle M\rangle_u|G_t\bigg) \quad (4.38)$$

uniquely solves (4.5), *and the optimal filtered wealth process satisfies the linear equation*

$$\widehat{X}_t^*=x-\int_0^t\frac{\widehat{\lambda}_uV_u(2)}{1-\rho^2+\rho^2V_u(2)}\widehat{X}_u^*d\widehat{S}_u+\int_0^t\frac{\varphi_u(1)\rho^2+\widehat{\lambda}_uV_u(1)-\tilde{h}_u}{1-\rho^2+\rho^2V_u(2)}d\widehat{S}_u. \quad (4.39)$$

Proof. The function $f(x)=\frac{1-\rho^2}{x}-\rho^2\ln x$ is differentiable and strictly decreasing on $]0,\infty[$ and takes all values from $]-\infty,+\infty[$. So (4.37) admits a unique solution for all α. Besides, the inverse function $\alpha(x)$ is differentiable. Therefore $Y_t(2)$ is a process of finite variation, and it is adapted since $\langle\widehat{\lambda}\cdot M\rangle_T$ is deterministic.

By definition of $Y_t(2)$ we have that for all $t\in[0,T]$

$$\frac{1-\rho^2}{Y_t(2)}-\rho^2\ln Y_t(2)=1-\rho^2+\langle\widehat{\lambda}\cdot M\rangle_T-\langle\widehat{\lambda}\cdot M\rangle_t.$$

It is evident that for $\alpha=1-\rho^2$ the solution of (4.37) is equal to 1, and it follows from (4.36) that $Y(2)$ satisfies the boundary condition $Y_T(2)=1$. Therefore

$$\begin{aligned}\frac{1-\rho^2}{Y_t(2)}-\rho^2\ln Y_t(2)-\left(1-\rho^2\right)&=-\left(1-\rho^2\right)\int_t^T d\frac{1}{Y_u(2)}+\rho^2\int_t^T d\ln Y_u(2)\\&=\int_t^T\left(\frac{1-\rho^2}{Y_u^2(2)}+\frac{\rho^2}{Y_u(2)}\right)dY_u(2)\end{aligned}$$

and

$$\int_t^T\frac{1-\rho^2+\rho^2Y_u(2)}{Y_u^2(2)}dY_u(2)=\langle\widehat{\lambda}\cdot M\rangle_T-\langle\widehat{\lambda}\cdot M\rangle_t$$

for all $t\in[0,T]$. Hence

$$\int_0^t\frac{1-\rho^2+\rho^2Y_u(2)}{Y_u^2(2)}dY_u(2)=\langle\widehat{\lambda}\cdot M\rangle_t,$$

and, by integrating both parts of this equality with respect to $Y(2)/(1-\rho^2+\rho^2Y(2))$, we obtain that $Y(2)$ satisfies

$$Y_t(2)=Y_0(2)+\int_0^t\frac{Y_u^2(2)\widehat{\lambda}_u^2}{1-\rho^2+\rho^2Y_u(2)}d\langle M\rangle_u,\tag{4.40}$$

which implies that the triple $(Y(2),\psi(2)=0,L(2)=0)$ satisfies (4.4) and $Y(2)=V(2)$ by Theorem 4.2. Equations (4.38) and (4.39) follow from (4.35) and (4.7), respectively, by taking $\varphi(2)=0$. $\square$

***Remark* 4.2.** In case $F^S\subseteq G$ we have $\widehat{M}=M$ and $\rho=1$. Therefore (4.40) is linear and $Y_t(2)=e^{\langle\widehat{\lambda}\cdot M\rangle_t-\langle\widehat{\lambda}\cdot M\rangle_T}$. In the case $\mathcal{A}=G$ of complete information, $Y_t(2)=e^{\langle\lambda\cdot N\rangle_t-\langle\lambda\cdot N\rangle_T}$.

5. Diffusion market model

***Example* 1.** Let us consider the financial market model

$$\begin{aligned}d\tilde{S}_t&=\tilde{S}_t\mu_t(\eta)dt+\tilde{S}_t\sigma_t(\eta)dw_t^0,\\d\eta_t&=a_t(\eta)dt+b_t(\eta)dw_t,\end{aligned}$$

subjected to initial conditions. Here w^0 and w are correlated Brownian motions with $Edw_t^0dw_t=\rho dt,\rho\in(-1,1)$.

Let us write

$$w_t=\rho w_t^0+\sqrt{1-\rho^2}w_t^1,$$

where w^0 and w^1 are independent Brownian motions. It is evident that $w^\perp = -\sqrt{1-\rho^2}w^0 + \rho w^1$ is a Brownian motion independent of w, and one can express Brownian motions w^0 and w^1 in terms of w and $w^\perp$ as

$$w_t^0 = \rho w_t - \sqrt{1-\rho^2}w_t^\perp, \quad w_t^1 = \sqrt{1-\rho^2}w_t + \rho w_t^\perp. \tag{5.1}$$

Suppose that $b^2 > 0,\ \sigma^2 > 0$, and coefficients μ, σ, a, and b are such that $F_t^{S,\eta} = F_t^{w^0,w}$ and $F_t^\eta = F_t^w$.

We assume that an agent would like to hedge a contingent claim H (which can be a function of S_T and η_T) using only observations based on the process η. So the stochastic basis will be $(\Omega, \mathcal{F}, F_t, P)$, where F_t is the natural filtration of (w^0, w) and the flow of observable events is $G_t = F_t^w$.

Also denote $dS_t = \mu_t dt + \sigma_t dw_t^0$, so that $d\tilde{S}_t = \tilde{S}_t dS_t$ and S is the return of the stock.

Let $\tilde{\pi}_t$ be the number of shares of the stock at time t. Then $\pi_t = \tilde{\pi}_t \tilde{S}_t$ represents an amount of money invested in the stock at the time $t \in [0,T]$. We consider the mean-variance hedging problem

$$\text{to minimize} \quad E\left[\left(x + \int_0^T \tilde{\pi}_t d\tilde{S}_t - H\right)^2\right] \quad \text{over all} \quad \tilde{\pi} \quad \text{for which} \quad \tilde{\pi}\tilde{S} \in \Pi(G), \tag{5.2}$$

which is equivalent to studying the mean-variance hedging problem

$$\text{to minimize} \qquad E\left[\left(x + \int_0^T \pi_t dS_t - H\right)^2\right] \quad \text{over all} \quad \pi \in \Pi(G).$$

***Remark* 5.1.** Since S is not G-adapted, $\widetilde{\pi}_t$ and $\widetilde{\pi}_t\widetilde{S}_t$ cannot be simultaneously G-predictable and the problem

$$\text{to minimize} \qquad E\left[\left(x + \int_0^T \tilde{\pi}_t d\tilde{S}_t - H\right)^2\right] \quad \text{over all} \quad \tilde{\pi} \in \Pi(G)$$

is not equivalent to the problem (5.2). In this setting, condition (A) is not satisfied, and it needs separate consideration.

By comparing with (1.1) we get that in this case

$$M_t = \int_0^t \sigma_s dw_s^0, \quad \langle M \rangle_t = \int_0^t \sigma_s^2 ds, \quad \lambda_t = \frac{\mu_t}{\sigma_t^2}.$$

It is evident that w is a Brownian motion also with respect to the filtration F^{w^0,w^1} and condition (B) is satisfied. Therefore by Proposition 2.2

$$\widehat{M}_t = \rho \int_0^t \sigma_s dw_s.$$

By the integral representation theorem the GKW decompositions (3.2) and (3.3) take the following forms:

$$c_H = EH, \ \ H_t = c_H + \int_0^t h_s \sigma_s dw_s^0 + \int_0^t h_s^1 dw_s^1, \tag{5.3}$$

$$H_t = c_H + \rho \int_0^t h_s^G \sigma_s dw_s + \int_0^t h_s^\perp dw_s^\perp. \tag{5.4}$$

By putting expressions (5.1) for w^0 and w^1 in (5.3) and equalizing integrands of (5.3) and (5.4), we obtain

$$h_t = \rho^2 h_t^G - \sqrt{1-\rho^2}\frac{h_t^\perp}{\sigma_t}$$

and hence

$$\widehat{h}_t = \rho^2 \widehat{h_t^G} - \sqrt{1-\rho^2}\,\frac{\widehat{h}_t^\perp}{\sigma_t}.$$

Therefore by the definition of $\widetilde{h}$

$$\widetilde{h}_t = \rho^2 \widehat{h_t^G} - \widehat{h}_t = \sqrt{1-\rho^2}\,\frac{\widehat{h}_t^\perp}{\sigma_t}. \tag{5.5}$$

By using notations

$$Z_s(0) = \rho\sigma_s\varphi_s(0),\ Z_s(1) = \rho\sigma_s\varphi_s(1),\ Z_s(2) = \rho\sigma_s\varphi_s(2),\ \ \theta_s = \frac{\mu_s}{\sigma_s},$$

we obtain the following corollary of Theorem 4.1.

Corollary 5.1. *Let H be a square integrable F_T-measurable random variable. Then the processes $V_t(0), V_t(1)$, and $V_t(2)$ from* (4.3) *satisfy the following system of backward equations:*

$$V_t(2) = V_0(2) + \int_0^t \frac{(\rho Z_s(2) + \theta_s V_s(2))^2}{1-\rho^2+\rho^2 V_s(2)}ds + \int_0^t Z_s(2)dw_s, \quad V_T(2) = 1, \tag{5.6}$$

$$V_t(1) = V_0(1) + \int_0^t \frac{(\rho Z_s(2) + \theta_s V_s(2))(\rho Z_s(1) + \theta_s V_s(1) - \sqrt{1-\rho^2}\,\widehat{h}_s^\perp)}{1-\rho^2+\rho^2 V_s(2)}ds$$

$$+ \int_0^t Z_s(1)dw_s, \quad V_T(1) = E(H|G_T), \tag{5.7}$$

$$V_t(0) = V_0(0) + \int_0^t \frac{(\rho Z_s(1) + \theta_s V_s(1) - \sqrt{1-\rho^2}\,\widehat{h}_s^\perp)^2}{1-\rho^2+\rho^2 V_s(2)}ds$$

$$+ \int_0^t Z_s(0)dw_s, \quad V_T(0) = E^2(H|G_T). \tag{5.8}$$

Besides, the optimal wealth process $\widehat{X}^*$ satisfies the linear equation

$$\widehat{X}_t^* = x - \int_0^t \frac{\rho Z_s(2) + \theta_s V_s(2)}{1-\rho^2+\rho^2 V_s(2)}\widehat{X}_s^*(\theta_s ds + \rho dw_s)$$

$$+ \int_0^t \frac{\rho Z_s(1) + \theta_s V_s(1) - \sqrt{1-\rho^2}\,\widehat{h}_s^\perp}{1-\rho^2+\rho^2 V_s(2)}(\theta_s ds + \rho dw_s). \tag{5.9}$$

Suppose now that θ_t and σ_t are deterministic. Then the solution of (5.6) is the pair $(V_t(2), Z_t(2))$, where $Z(2) = 0$ and $V(2)$ satisfies the ordinary differential equation

$$\frac{dV_t(2)}{dt} = \frac{\theta_t^2 V_t^2(2)}{1-\rho^2+\rho^2 V_t(2)}, \quad V_T(2) = 1. \tag{5.10}$$

By solving this equation we obtain

$$V_t(2) = \nu\left(\rho, 1-\rho^2 + \int_t^T \theta_s^2 ds\right) \equiv \nu_t^{\theta,\rho}, \tag{5.11}$$

where $\nu(\rho, \alpha)$ is the solution of (4.37). From (5.10) it follows that

$$\left(\ln \nu_t^{\theta,\rho}\right)' = \frac{\theta_t^2 \nu_t^{\theta,\rho}}{1-\rho^2+\rho^2\nu_t^{\theta,\rho}} \quad \text{and} \quad \ln \frac{\nu_s^{\theta,\rho}}{\nu_t^{\theta,\rho}} = \int_t^s \frac{\theta_r^2 \nu_r^{\theta,\rho} dr}{1-\rho^2+\rho^2\nu_r^{\theta,\rho}}. \tag{5.12}$$

If we solve the linear BSDE (5.7) and use (5.12), we obtain

$$V_t(1) = E\left[\widehat{H}_T(w)\mathcal{E}_{tT}\left(-\int_0^\cdot \frac{\theta_r \nu_r^{\theta,\rho}}{1-\rho^2+\rho^2\nu_r^{\theta,\rho}}(\theta_r dr + \rho dw_r)\right)|G_t\right],$$

$$\int_t^T \frac{\theta_s \nu_s^{\theta,\rho}\sigma_s}{1-\rho^2+\rho^2\nu_s^{\theta,\rho}} E\left[\tilde{h}_s(w)\mathcal{E}_{ts}\left(-\int_0^\cdot \frac{\theta_r \nu_r^{\theta,\rho}}{1-\rho^2+\rho^2\nu_r^{\theta,\rho}}(\theta_r dr + \rho dw_r)\right)|G_t\right] ds$$

$$= \nu_t^{\theta,\rho} E\left[\widehat{H}_T(w)\mathcal{E}_{tT}\left(-\int_0^\cdot \frac{\theta_r \nu_r^{\theta,\rho}}{1-\rho^2+\rho^2\nu_r^{\theta,\rho}}\rho dw_r\right)|G_t\right]$$

$$+ \nu_t^{\theta,\rho}\int_t^T \frac{\mu_s}{1-\rho^2+\rho^2\nu_s^{\theta,\rho}} E\left[\tilde{h}_s(w)\mathcal{E}_{ts}\left(-\int_0^\cdot \frac{\theta_r \nu_r^{\theta,\rho}}{1-\rho^2+\rho^2\nu_r^{\theta,\rho}}\rho dw_r\right)|G_t\right] ds.$$

By using the Girsanov theorem we finally get

$$V_t(1) = \nu_t^{\theta,\rho} E\left[\widehat{H}_T\left(\rho\int_0^\cdot \frac{\theta_r \nu_r^{\theta,\rho}}{1-\rho^2+\rho^2\nu_r^{\theta,\rho}}dr + w\right)|G_t\right]$$

$$+ \nu_t^{\theta,\rho}\int_t^T \frac{\mu_s}{1-\rho^2+\rho^2\nu_s^{\theta,\rho}} E\left[\tilde{h}_s\left(\rho\int_0^\cdot \frac{\theta_r \nu_r^{\theta,\rho}}{1-\rho^2+\rho^2\nu_r^{\theta,\rho}}dr + w\right)|G_t\right] ds. \tag{5.13}$$

Besides, the optimal strategy is of the form

$$\pi_t^* = -\frac{\theta_t V_t(2)}{(1-\rho^2+\rho^2 V_t(2))\sigma_t}\widehat{X}_t^* + \frac{\rho Z_t(1) + \theta_t V_t(1) - \sqrt{1-\rho^2}\,\widehat{h}_t^\perp}{(1-\rho^2+\rho^2 V_t(2))\sigma_t}. \qquad \square$$

If, in addition, μ and σ are constants and the contingent claim is of the form $H = \mathcal{H}(S_T, \eta_T)$, then one can give an explicit expressions also for $\tilde{h}$, $\widehat{h}^\perp$, $\widehat{H}$, and $Z(1)$.

***Example* 2.** In Frey and Runggaldier [9] the incomplete-information situation arises, assuming that the hedger is unable to monitor the asset continuously but is confined to observations at discrete random points in time $\tau_1, \tau_2, \dots, \tau_n$. Perhaps it is more natural to assume that the hedger has access to price information on full intervals $[\sigma_1, \tau_1], [\sigma_2, \tau_2], \dots, [\sigma_n, \tau_n]$. For the models with nonzero drifts, even the case $n = 1$ is nontrivial. Here we consider this case in detail.

Let us consider the financial market model

$$d\widetilde{S}_t = \mu \widetilde{S}_t dt + \sigma \widetilde{S}_t dW_t, \quad S_0 = S,$$

where W is a standard Brownian motion and the coefficients μ and σ are constants. Assume that an investor observes only the returns $S_t - S_0 = \int_0^t \frac{1}{\widetilde{S}_u} d\widetilde{S}_u$ of the stock prices up to a random moment τ before the expiration date T. Let $\mathcal{A}_t = F_t^S$, and let τ be a stopping time with respect to F^S. Then the filtration G_t of observable events is equal to the filtration $F_{t\wedge\tau}^S$.

Consider the mean-variance hedging problem

$$\text{to minimize} \quad E\left[\left(x + \int_0^T \pi_t dS_t - H\right)^2\right] \quad \text{over all} \quad \pi \in \Pi(G),$$

where π_t is a dollar amount invested in the stock at time t.

By comparing with (1.1) we get that in this case

$$N_t = M_t = \sigma W_t, \quad \langle M\rangle_t = \sigma^2 t, \quad \lambda_t = \frac{\mu}{\sigma^2}.$$

Let $\theta = \frac{\mu}{\sigma}$. The measure Q defined by $dQ = \mathcal{E}_T(\theta W)dP$ is a unique martingale measure for S, and it is evident that Q satisfies the reverse Hölder condition. It is also evident that any G-martingale is F^S-martingale and that conditions (A)–(C) are satisfied. Besides,

$$E(W_t|G_t) = W_{t\wedge\tau}, \quad \widehat{S}_t = \mu t + \sigma W_{t\wedge\tau} \quad \text{and} \quad \rho_t = I_{\{t\le\tau\}}. \tag{5.14}$$

By the integral representation theorem

$$E\left(H|F_t^S\right) = EH + \int_0^t h_u \sigma dW_u \tag{5.15}$$

for F-predictable W-integrable process h. On the other hand, by the GKW decomposition with respect to the martingale $W^\tau = (W_{t\wedge\tau}, t\in[0,T])$,

$$E\left(H|F_t^S\right) = EH + \int_0^t h_u^G \sigma dW_u^\tau + L_t^G \tag{5.16}$$

for F^S-predictable process h^G and F^S martingale L^G strongly orthogonal to W^τ. Therefore, by equalizing the right-hand sides of (5.15) and (5.16) and taking the mutual characteristics of both parts with W^τ, we obtain $\int_0^{t\wedge\tau}(h_u^G\rho_u^2 - h_u)du = 0$ and hence

$$\int_0^t \widetilde{h}_u du = \int_0^t \left(\widehat{h}_u^G I_{(u\le\tau)} - \widehat{h}_u\right) du = -\int_0^t I_{(u>\tau)} E\left(h_u|F_\tau^S\right) du. \tag{5.17}$$

Therefore, by using notations

$$Z_s(0) = \rho\sigma\varphi_s(0),\ Z_s(1) = \rho\sigma\varphi_s(1),\ Z_s(2) = \rho\sigma\varphi_s(2),$$

it follows from Theorem 4.1 that the processes $(V_t(2), Z_t(2))$ and $(V_t(1), Z_t(1))$ satisfy the following system of backward equations:

$$\begin{aligned} V_t(2) &= V_0(2) + \int_0^{t\wedge\tau} \frac{\left(Z_s(2)+\theta V_s(2)\right)^2}{V_s(2)} ds + \int_{t\wedge\tau}^t \theta^2 V_s^2(2) ds \\ &\quad + \int_0^{t\wedge\tau} Z_s(2) dW_s, \quad V_T(2) = 1, \end{aligned} \tag{5.18}$$

$$\begin{aligned} V_t(1) &= V_0(1) + \int_0^{t\wedge\tau} \frac{\left(Z_s(2)+\theta V_s(2)\right)\left(Z_s(1)+\theta V_s(1)\right)}{V_s(2)} ds \\ &\quad + \int_{t\wedge\tau}^t \theta V_s(2)\left(\theta V_s(1) + E\left(h_s|F_\tau^S\right)\right) ds \\ &\quad + \int_0^{t\wedge\tau} Z_s(1) dW_s, \quad V_T(1) = E(H|G_T). \end{aligned} \tag{5.19}$$

Equation (5.18) admits in this case an explicit solution. To obtain the solution one should solve first the equation

$$U_t = U_0 + \int_0^t \theta^2 U_s^2 ds, \quad U_T = 1, \tag{5.20}$$

in the time interval $[\tau, T]$ and then the BSDE

$$V_t(2) = V_0(2) + \int_0^t \frac{\big(Z_s(2) + \theta V_s(2)\big)^2}{V_s(2)} ds + \int_0^t Z_s(2) dW_s \tag{5.21}$$

in the interval $[0, \tau]$, with the boundary condition $V_\tau(2) = U_\tau$. The solution of (5.20) is

$$U_t = \frac{1}{1 + \theta^2(T - t)},$$

and the solution of (5.21) is expressed as

$$V_t(2) = \frac{1}{E\big((1 + \theta^2(T - \tau))\mathcal{E}^2_{t,\tau}(-\theta W)|F^S_t\big)}$$

(this can be verified by applying the Itô formula for the process $V_t^{-1}(2)\mathcal{E}_t^2(-\theta W)$ and by using the fact that this process is a martingale). Therefore

$$V_t(2) = \begin{cases} \frac{1}{1+\theta^2(T-t)} & \text{if} \quad t \geq \tau, \\ \frac{1}{E\big((1+\theta^2(T-\tau))\mathcal{E}^2_{t,\tau}(-\theta W)|F^S_t\big)} & \text{if} \quad t \leq \tau. \end{cases} \tag{5.22}$$

According to (4.37), taking in mind (5.14), (5.17), and the fact that $e^{-\int_t^T \theta^2 V_u(2) du} = \frac{1}{1+\theta^2(T-t)}$ on the set $t \geq \tau$, the solution of (5.19) is equal to

$$\begin{aligned} V_t(1) = & \, E\bigg(\frac{H}{1 + \theta^2(T - t)} + \int_t^T \frac{\theta V_u(2) h_u du}{1 + \theta^2(T - u)} |F^S_\tau\bigg) I_{(t>\tau)} \\ & + E\bigg(\mathcal{E}_{t,\tau}\bigg(-\frac{\varphi(2) + \lambda V(2)}{V(2)} \cdot S\bigg)\bigg(\frac{H}{1 + \theta^2(T - \tau)} \\ & + \int_\tau^T \frac{\theta V_u(2) h_u du}{1 + \theta^2(T - u)}\bigg)|F^S_t\bigg) I_{(t\leq\tau)}. \end{aligned} \tag{5.23}$$

By Theorem 4.1 the optimal filtered wealth process is a solution of a linear SDE, which takes in this case the following form:

$$\begin{aligned} \widehat{X}^*_t = & \, x - \int_0^{t\wedge\tau} \frac{\varphi_u(2) + \theta V_u(2)}{V_u(2)} \widehat{X}^*_u(\theta du + dW_u) - \int_{t\wedge\tau}^t \theta^2 V_u(2) \widehat{X}^*_u du \\ & + \int_0^{t\wedge\tau} \frac{\varphi_u(1) + \theta V_u(1)}{V_u(2)} (\theta du + dW_u) + \int_{t\wedge\tau}^t \left(\theta^2 V_u(1) + \mu E\left(h_u|F^S_\tau\right)\right) du. \end{aligned} \tag{5.24}$$

The optimal strategy is equal to

$$\begin{aligned} \pi^*_t = & \left[-\frac{\varphi_t(2) + \theta V_t(2)}{V_t(2)} I_{(t\leq\tau)} - \theta^2 V_t(2) I_{(t>\tau)}\right] \widehat{X}^*_t \\ & + \frac{\varphi_t(1) + \theta V_t(1)}{V_t(2)} I_{(t\leq\tau)} + \left(\theta^2 V_t(1) + \mu E\left(h_t|F^S_\tau\right)\right) I_{(t>\tau)}, \end{aligned} \tag{5.25}$$

where $\widehat{X}^*_t$ is a solution of the linear equation (5.24), $V(2)$ and $V(1)$ are given by (5.22) and (5.23), and $\varphi(2)$ and $\varphi(1)$ are integrands of their martingale parts, respectively. In particular the optimal strategy in time interval $[\tau, T]$ (i.e., after interrupting observations) is of the form

$$\pi^*_t = -\theta^2 V_t(2) \widehat{X}^*_t + \theta^2 V_t(1) + \mu E\left(h_t|F^S_\tau\right), \tag{5.26}$$

where

$$\widehat{X}_t^* = \frac{\widehat{X}_\tau^*}{1+\theta^2(t-\tau)} - \int_\tau^t \left(\theta^2 V_u(1) - \mu E\left(h_u|F_\tau^S\right)\right)\frac{1}{1+\theta^2(t-u)}du.$$

For instance, if τ is deterministic, then $V_t(2)$ is also deterministic:

$$V_t(2) = \begin{cases} \frac{1}{1+\theta^2(T-t)} & \text{if} \quad t \ge \tau, \\ \frac{1}{1+\theta^2(T-t)} e^{-\theta^2(\tau-t)} & \text{if} \quad t \le \tau, \end{cases}$$

and $\varphi(2) = 0$.

Note that it is not optimal to do nothing after interrupting observations, and in order to act optimally one should change the strategy deterministically as it is given by (5.26).

Appendix A

For convenience we give the proofs of the following assertions used in the paper.

Lemma A.1. *Let conditions* (A)–(C) *be satisfied and* $\widehat{M}_t = E(M_t|G_t)$. *Then* $\langle\widehat{M}\rangle$ *is absolutely continuous w.r.t.* $\langle M\rangle$ *and* $\mu^{\langle M\rangle}$ *a.e.*

$$\rho_t^2 = \frac{d\langle\widehat{M}\rangle_t}{d\langle M\rangle_t} \le 1.$$

Proof. By (2.4) for any bounded G-predictable process h

$$\begin{aligned} E\int_0^t h_s^2 d\langle\widehat{M}\rangle_s &= E\left(\int_0^t h_s d\widehat{M}_s\right)^2 = E\left(E\left(\int_0^t h_s dM_s|G_t\right)\right)^2 \\ &\le E\left(\int_0^t h_s dM_s\right)^2 = E\int_0^t h_s^2 d\langle M\rangle_s, \end{aligned} \tag{A.1}$$

which implies that $\langle\widehat{M}\rangle$ is absolutely continuous w.r.t. $\langle M\rangle$, i.e.,

$$\langle\widehat{M}\rangle_t = \int_0^t \rho_s^2 d\langle M\rangle_s$$

for a G-predictable process ρ. □

Moreover (A.1) implies that the process $\langle M\rangle$–$\langle\widehat{M}\rangle$ is increasing and hence $\rho^2 \le 1$ $\mu^{\langle M\rangle}$ a.e.

Lemma A.2. *Let* $H \in L^2(P, F_T)$, *and let conditions* (A)–(C) *be satisfied. Then*

$$E\int_0^T \tilde{h}_u^2 d\langle M\rangle_u < \infty.$$

Proof. It is evident that

$$E\int_0^T (h_u^G)^2 d\langle\widehat{M}\rangle_u < \infty, \quad E\int_0^T h_u^2 d\langle M\rangle_u < \infty.$$

Therefore, by the definition of $\tilde{h}$ and Lemma A.1,

$$E\int_0^T \tilde{h}_u^2 d\langle M\rangle_u \le 2E\int_0^T \widehat{h}_u^2 d\langle M\rangle_u + 2E\int_0^T \left(\widehat{h_u^G}\right)^2 \rho_u^4 d\langle M\rangle_u$$
$$\le 2E\int_0^T h_u^2 d\langle M\rangle_u + 2E\int_0^T \left(h_u^G\right)^2 \rho_u^2 d\langle \widehat{M}\rangle_u < \infty.$$

Thus $\tilde{h} \in \Pi(G)$ by Remark 2.5. □

Lemma A.3. (a) *Let $Y = (Y_t, t \in [0,T])$ be a bounded positive submartingale with the canonical decomposition*

$$Y_t = Y_0 + B_t + m_t,$$

where B is a predictable increasing process and m is a martingale. Then $m \in BMO$.

(b) *In particular the martingale part of $V(2)$ belongs to BMO. If H is bounded, then martingale parts of $V(0)$ and $V(1)$ also belong to the class BMO, i.e., for $i = 0, 1, 2$,*

$$E\left(\int_\tau^T \varphi_u^2(i)\rho_u^2 d\langle M\rangle_u | G_\tau\right) + E\left(\langle m(i)\rangle_T - \langle m(i)\rangle_\tau | G_\tau\right) \le C \tag{A.2}$$

for every stopping time τ.

Proof. By applying the Itô formula for $Y_T^2 - Y_\tau^2$ we have

$$\langle m\rangle_T - \langle m\rangle_\tau + 2\int_\tau^T Y_u dB_u + 2\int_\tau^T Y_u dm_u = Y_T^2 - Y_\tau^2 \le \text{const} \tag{A.3}$$

Since Y is positive and B is an increasing process, by taking conditional expectations in (A.3) we obtain

$$E(\langle m\rangle_T - \langle m\rangle_\tau | F_\tau) \le \text{const}$$

for any stopping time τ, and hence $m \in BMO$.

(A.2) follows from assertion (a) applied for positive submartingales $V(0), V(2)$, and $V(0) + V(2) - 2V(1)$. For the case $i = 1$ one should take into account also the inequality

$$\langle m(1)\rangle_t \le \text{const}(\langle m(0) + m(2) - 2m(1)\rangle_t + \langle m(0)\rangle_t + \langle m(2)\rangle_t). \qquad \square$$

ACKNOWLEDGMENTS

We would like to thank N. Lazrieva and anonymous referees for useful remarks and comments.

This work was supported by Georgian National Science Foundation grant STO07/3-172.

REFERENCES

[1] J. M. BISMUT, *Conjugate convex functions in optimal stochastic control*, J. Math. Anal. Appl., 44 (1973), pp. 384–404.

[2] R. CHITASHVILI, *Martingale Ideology in the Theory of Controlled Stochastic Processes*, Lecture Notes in Math. 1021, Springer-Verlag, New York, 1983, pp. 73–92.

[3] G. B. DI MASI, E. PLATEN, AND W. J. RUNGGALDIER, *Hedging of options under discrete observation on assets with stochastic volatility*, in Seminar on Stochastics Analysis and Random Fields Applications, Progr. Probab. 36, Birkhäuser-Verlag, Basil, Switzerland, 1995, pp. 359–364.

[4] F. DELBAEN, P. MONAT, W. SCHACHERMAYER, W. SCHWEIZER, AND C. STRICKER, *Weighted norm inequalities and hedging in incomplete markets*, Finance Stoch., 1 (1997), pp. 181–227.

[5] C. Dellacherie and P. A. Meyer, *Probabilités et Potentiel*, II. Hermann, Paris, 1980.

[6] D. Duffie and H. R. Richardson, *Mean-variance hedging in continuous time*, Ann. Appl. Probab., 1 (1991), pp. 1–15.

[7] N. El Karoui and M. C. Quenez, *Dynamic programming and pricing of contingent claims in an incomplete market*, SIAM J. Control Optim., 33 (1995), pp. 29–66.

[8] H. Föllmer and D. Sondermann, *Hedging non-redundant contingent claims*, in Contrib. Math. Econ., Hon. G. Debreu, W. Hildenbrand, and A. Mas-Collel, eds., North–Holland, Amsterdam, 1986, pp. 205–223.

[9] R. Frey and W. J. Runggaldier, *Risk-minimizing hedging strategies under restricted information: The case of stochastic volatility models observable only at discrete random times. Financial optimization*, Math. Methods Oper. Res., 50 (1999), pp. 339–350.

[10] C. Gourieroux, J. P. Laurent, and H. Pham, *Mean-variance hedging and numeraire*, Math. Finance, 8 (1998), pp. 179–200.

[11] D. Heath, E. Platen, and M. Schweizer, *A comparision of two quadratic approaches to hedging in incomplete markets*, Math. Finance, 11 (2001), pp. 385–413.

[12] C. Hipp, *Hedging general claims*, in Proceedings of the 3rd AFIR Colloquium, 2, Rome, 1993, pp. 603–613.

[13] J. Jacod, *Calcule Stochastique et Problèmes des Martingales*, Lecture Notes in Math. 714, Springer-Verlag, Berlin, 1979.

[14] N. Kazamaki, *Continuous Exponential Martingales and BMO*, Lecture Notes in Math. 1579, Springer-Verlag, New York, 1994.

[15] D. O. Kramkov, *Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets*, Probab. Theory Related Fields, 105 (1996), pp. 459–479.

[16] M. Kobylanski, *Backward stochastic differential equation and partial differential equations with quadratic growth*, Ann. Prob., 28 (2000), pp. 558–602.

[17] J. P. Lepeltier and J. San Martin, *Existence for BSDE with superlinear-quadratic coefficient*, Stoch. Stoch. Rep., 63 (1998), pp. 227–240.

[18] R. Sh. Liptzer and A. N. Shiryayev, *Martingale Theory*, Nauka, Moscow, 1986.

[19] M. Mania and R. Tevzadze, *Backward stochastic PDE and imperfect hedging,* Int. J. Theor. Appl. Finance, 6 (2003), pp. 663–692.

[20] M. A. Morlais, *Quadratic Backward Stochastic Differential Equations (BSDEs) Driven by a Continuous Martingale and Application to the Utility Maximization Problem*, http://hal.ccsd.cnrs.fr/ccsd-00020254/

[21] E. Pardoux and S. G. Peng, *Adapted solution of a backward stochastic differential equation,* Systems Control Lett., 14 (1990), pp. 55–61.

[22] H. Pham, *Mean-variance hedging for partially observed drift processes*, Int. J. Theor. Appl. Finance, 4 (2001), pp. 263–284.

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

[24] M. Schäl, *On quadratic cost criteria for option hedging*, Math. Oper. Res., 19 (1994), pp. 121–131.

[25] M. Schweizer, *Mean-variance hedging for general claims*, Ann. Appl. Probab., 2 (1992), pp. 171–179.

[26] M. Schweizer, *Approximating random variables by stochastic integrals*, Ann. Probab., 22 (1994), pp. 1536–1575.

[27] M. Schweizer, *Risk-minimizing hedging strategies under restricted information*, Math. Finance, 4 (1994), pp. 327–342.

[28] R. Tevzadze, *Solvability of backward stochastic differential equations with quadratic growth*, Stochastic Proc. Appl., 118 (2008), pp. 503–515.

SOLVABILITY OF BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS WITH QUADRATIC GROWTH

R. TEVZADZE

Abstract. We prove the existence of the unique solution of a general Backward Stochastic Differential Equation with quadratic growth driven by martingales. Some kind of comparison theorem is also proved.

Key words and phrases: Backward Stochastic Differential Equation, Contraction principle, BMO-martingale

MSC 2010: 90A09, 60H30, 90C39

1. Introduction

In this paper we show a general result of existence and uniqueness of Backward Stochastic Differential Equation (BSDE) with quadratic growth driven by continuous martingale. Backward stochastic differential equations have been introduced by Bismut [1] for the linear case as equations of the adjoint process in the stochastic maximum principle. A nonlinear BSDE (with Bellman generator) was first considered by Chitashvili [4]. He derived the semimartingale BSDE (or SBE), which can be considered as a stochastic version of the Bellman equation for a stochastic control problem, and proved the existence and uniqueness of a solution. The theory of BSDEs driven by the Brownian motion was developed by Pardoux and Peng [22] for more general generators. The results of Pardoux and Peng were generalized by Kobylansky [11], Lepeltier and San Martin [12] for generators with quadratic growth. In the work of Hu at all [8] BMO-martingales were used for BSDE with quadratic generators in Brownian setting and in [15], [16], [17], [18], [19], [21] for BSDEs driven by martingales. By Chitashvili [4], Buckdahn [3],and El Karoui and Huang [7] the well posedness of BSDE with generators satisfying Lipschitz type conditions was established. Here we suggest new approach including an existence and uniqueness of the solution of general BSDE with quadratic growth. In the earlier papers [15], [16], [17], [18], [19] we studied, as well as Bobrovnytska and Schweizer [2], the particular cases of BSDE with quadratic nonlinearities related to the primal and dual problems of Mathematical Finance. In these works the solutions were represented as a value function of the corresponding optimization problems.

The paper is organized as follows. In Section 2 we give some basic definitions and facts used in what follows. In Section 3 we show the solvability of the system of BSDEs for sufficiently small initial condition and further prove the solvability of one dimensional BSDE for arbitrary bounded initial data. At the end of Section 4 we prove the comparison theorem, which generalizes the results of Mania and Schweizer [14], and apply this results to the uniqueness of the solution.

Published in *Stochastic Process. Appl.* **118** (2008), no. 3, 503–515.

2. Some Basic Definitions and Assumptions

Let $(\Omega, \mathcal{F}, \mathbf{F} = (F_t)_{t\geq 0}, P)$ be filtered probability space satisfying the usual conditions. We assume that all local martingales with respect to $\mathbf{F}$ are continuous. Here the time horizon $T < \infty$ is a stopping time and $\mathcal{F} = F_T$. Let us consider Backward Stochastic Differential Equation (BSDE) of the form

$$dY_t = -f(t, Y_t, \sigma_t^* Z_t)dK_t - d\langle N\rangle_t g_t + Z_t^* dM_t + dN_t, \tag{2.1}$$

$$Y_T = \xi, \tag{2.2}$$

We suppose that

- $(M_t, t \geq 0)$ is an R^n-valued continuous martingale with cross-variations matrix $\langle M\rangle_t = (\langle M^i, M^j\rangle_t)_{1\leq i,j\leq n}$,
- $(K_t, t\geq 0)$ is a continuous, adapted, increasing process, such that $\langle M\rangle_t = \int_0^t \sigma_s \sigma_s^* dK_s$ for some predictable, non degenerate $n \times n$ matrix σ,
- ξ is $\mathcal{F}$-measurable an R^d-valued random variable,
- $f : \Omega \times R^+ \times R^d \times R^{n\times d} \to R^d$ is a stochastic process, such that for any $(y, z) \in R^d \times R^{n\times d}$ the process $f(\cdot, \cdot, y, z)$ is predictable,
- $g : \Omega \times R^+ \to R^{d\times d}$ is a predictable process.

The notation $R^{n\times d}$ here denotes the space of $n \times d$-matrix C with Euclidian norm $|C| = \sqrt{\mathrm{tr}(CC^*)}$. For some stochastic process X_t and sopping times $\tau,\ \nu$, such that $\tau \geq \nu$ we denote $X_{\nu,\tau} = X_\tau - X_\nu$. For all unexplained notations concerning the martingale theory used below we refer [9], [5] and [13]. About BMO-martingales see [6] or [10].

A solution of the BSDE is a triple (Y, Z, N) of stochastic processes, such that (2.1), (2.2) is satisfied and

- Y is an adapted R^d-valued continuous process,
- Z is an $R^{n\times d}$-valued predictable process,
- N is an R^d-valued continuous martingale, orthogonal to the basic martingale M.

One says that (f, g, ξ) is a generator of BSDE (2.1),(2.2).

We introduce the following spaces

- $L^\infty(R^d) = \{X : \Omega \to R^d, \mathcal{F}_T - \text{measurable}, \|X\|_\infty = \operatorname*{ess\,sup}_{\Omega} |X(\omega)| < \infty\}$,
- $S^\infty(R^d) = \{\varphi : \Omega\times R^+ \to R^d,\ \text{continuous, adapted}, \|\varphi\|_\infty = \operatorname*{ess\,sup}_{[[0,T]]} |\varphi(t,\omega)| < \infty\}$,
- $$H^2(R^{n\times d}, \sigma) = \Big\{\varphi : \Omega \times R^+ \to R^{n\times d},\ \text{predictable},$$

$$\|\varphi\|_H^2 = \operatorname*{ess\,sup}_{[[0,T]]} E\Big(\int_t^T |\sigma_s^* \varphi_s|^2 dK_s|\mathcal{F}_t\Big) \equiv \operatorname*{ess\,sup}_{[[0,T]]} E(\mathrm{tr}\langle \varphi\cdot M\rangle_{tT}|\mathcal{F}_t) < \infty\Big\}, \tag{2.3}$$

- $\mathrm{BMO}(Q) = \{N,\ R^d\text{-valued } Q\text{-martingale } \|N\|_Q^2 = \operatorname*{ess\,sup}_{[[0,T]]} E^Q(tr\langle N\rangle_{tT}|\mathcal{F}_t) < \infty\}$.

We also use the notation $|r|_{2,\infty}$ for the norm $\| \int_0^T r_s^2 dK_s\|_\infty$.
The norm of the triple is defined as

$$\|(Y, Z, N)\|^2 = \|Y\|^2 + \|Z\|_H^2 + \|N\|_P^2.$$

Throughout the paper we use the condition

A) There exist a constant θ and predictable processes

$$\alpha : \Omega \times R^+ \to R^d,\ \Gamma : \Omega \times R^+ \to Lin(R^{n\times d}, R^d),\quad r : \Omega \times R^+ \to R,$$

such that the following conditions $\int_0^T r_s dK_s,\ \int_0^T r_s^2 dK_s \in L^\infty$, $\Gamma(\sigma^{-1}) \in H_T^2$, $|\alpha_t| \le r_t$, $|g_t| \le \theta^2$ and

$$\begin{aligned} &|f(t,y_1,z_1) - f(t,y_2,z_2) - \alpha_t(y_1 - y_2) - \Gamma_t(z_1 - z_2)| \\ &\le (r_t|y_1 - y_2| + \theta|z_1 - z_2|)(r_t(|y_1| + |y_2|) + \theta(|z_1| + |z_2|)) \end{aligned} \tag{2.4}$$

are satisfied.

Sometimes we use the more restrictive conditions

B1) $\int_0^T |f(t,0,0)|dK_t + |g_t| \le \theta^2$ for all $t \in [0,T]$,
B2) $|f_y(t,y,z)| \le r_t,\ |f_z(t,y,z)| \le r_t + \theta|z|$ for all (t,y,z),
B3) $|f_{yy}(t,y,z)| \le r_t^2,\ |f_{yz}(t,y,z)| \le \theta r_t,\ |f_{zz}(t,y,z)| \le \theta^2$ for all (t,y,z).

Remark 1. Condition A) follow from conditions B1)-B3), since using notations $\delta y = y_1 - y_2,\ \delta z = z_1 - z_2$ for $\alpha_t = f_y(t,0,0),\ \Gamma_t = f_z(t,0,0)$ by the mean value theorem we have

$$\begin{aligned} &|f(t,y_1,z_1) - f(t,y_2,z_2) - \alpha_t\delta y - \Gamma_t(\delta z)| \\ &= |f_y(t,\nu y_1 + (1-\nu)y_2, \nu z_1 + (1-\nu)z_2)\delta y - f_y(t,0,0)\delta y| \\ &+ f_z(t,\nu y_1 + (1-\nu)y_2, \nu z_1 + (1-\nu)z_2)(\delta z) - f_z(t,0,0)(\delta z)|, \end{aligned}$$

for some $\nu \in [0,1]$. Using again mean value theorem we obtain that

$$\begin{aligned} &|f(t,y_1,z_1) - f(t,y_2,z_2) - \alpha_t\delta y - \Gamma_t(\delta z)| \\ &\le (|\nu y_1 + (1-\nu)y_2| \max_{y,z}|f_{yy}(t,y,z)| + |\nu z_1 + (1-\nu)z_2| \max_{y,z}|f_{yz}(t,y,z)|)|\delta y| \\ &+ (|\nu y_1 + (1-\nu)y_2| \max_{y,z}|f_{yz}(t,y,z)| + |\nu z_1 + (1-\nu)z_2| \max_{y,z}|f_{zz}(t,y,z)|)|\delta z| \\ &\le [r_t^2(|y_1| + |y_2|) + r_t\theta(|z_1| + |z_2|)]|\delta y| + [r_t\theta(|y_1| + |y_2|) + \theta^2(|z_1| + |z_2|)]|\delta z| \\ &= (r_t|\delta y| + \theta|\delta z|)(r_t(|y_1| + |y_2|) + \theta(|z_1| + |z_2|). \end{aligned}$$

Remark 2. If $d = 1$ the operator Γ_t is given by an n-dimensional vector γ_t such that $\Gamma_t(z) = \gamma_t^* z$. Thus inequality in A) can be rewritten as

$$\begin{aligned} &|f(t,y_1,z_1) - f(t,y_2,z_2) - \alpha_t\delta y - \gamma_t^*\delta z| \\ &\le (r_t|\delta y| + \theta|\delta z|)(r_t(|y_1| + |y_2|) + \theta(|z_1| + |z_2|)). \end{aligned}$$

The main statement of the paper is the following

Theorem 1. *Let $\xi \in L^\infty$, $d = 1$ and conditions* B1)–B3) *are satisfied. Then there exists a unique triple (Y, Z, N), where $Y \in S^\infty$, $Z \in H^2$, $N \in BMO$, that satisfies equation* (2.1), (2.2).

3. Existence of the Solution

First we prove the existence and uniqueness of the solution for a sufficiently small initial data.

Proposition 1. *Let f and g satisfy condition* A) *with $\alpha = 0$ and $\gamma_t = 0$. Then for ξ with the norm $\|\xi\|_\infty < \frac{1}{32\beta}$, $\beta = 8\max(|r|^2_{2,\infty}, \theta^2)$ there exists a unique solution (Y, Z, N) of BSDE*

$$dY_t = (f(t,0,0) - f(t, Y_t, \sigma_t^* Z_t))dK_t + d\langle N\rangle_t g_t + Z_t^* dM_t + dN_t, \qquad Y_T = \xi, \tag{3.1}$$

with the norm $\|(Y,Z,N)\| \leq R$, where R is a constant satisfying the inequality $4\|\xi\|^2_\infty + \beta^2 R^4 \leq R^2$, namely $R = 2\sqrt{2}\|\xi\|_\infty$.

Moreover, if $\|\xi\|_\infty + \| \int_0^\infty |f(s,0,0)|dK_s\|_\infty$ is small enough then BSDE (2.1) *admits a unique solution.*

Proof. We define the mapping $(Y, Z, N) = F(y, z, n)$, n is orthogonal to M, $(y, z\cdot M + n) \in S_T^\infty \times BMO(P)$ by the relation

$$dY_t = (f(t,0,0) - f(t, y_t, \sigma_t^* z_t))dK_t + d\langle n\rangle_t g_t + Z_t^* dM_t + dN_t, \qquad Y_T = \xi. \tag{3.2}$$

Using the Itô formula for $|Y_t|^2$ we obtain that

$$|Y_t|^2 = |\xi|^2 + 2\int_t^T Y_s^*(f(s, y_s, \sigma_s^* z_s) - f(s,0,0))dK_t$$
$$+2\int_t^T Y_s^* d\langle n\rangle_s g_s - \int_t^T \mathrm{tr} Z_s^* d\langle M\rangle_s Z_s - \mathrm{tr}\langle N\rangle_{tT} - \int_t^T Y_s^* Z_s^* dM_s - \int_t^T Y_s^* dN_s.$$

If we take the conditional expectation and use (2.3) and the elementary inequality $2ab \leq \frac{1}{4}a^2 + 4b^2$, we get

$$|Y_t|^2 + E\bigg(\int_t^T |\sigma_s^* Z_s|^2 dK_s + tr\langle N\rangle_{tT}|\mathcal{F}_t\bigg) \leq \|\xi\|^2 + \frac{1}{4}\|Y\|^2_\infty$$
$$+4E^2\bigg(\int_t^T |f(s, y_s, \sigma_s^* z_s) - f(s,0,0)|dK_s + \int_t^T |g_s| d\mathrm{tr}\langle n\rangle_s|\mathcal{F}_t\bigg). \tag{3.3}$$

Thus using condition A), identities

$$\mathrm{tr}\langle z\cdot M\rangle_t = \mathrm{tr}\int_0^t z_s^* d\langle M\rangle_s z_s = \int_0^t \mathrm{tr}(z_s^* \sigma_s \sigma_s^* z_s)dK_s = \int_0^t |\sigma_s^* z_s|^2 dK_s \tag{3.4}$$

and explicit inequalities

$$\frac{1}{2}(\|Y\|^2_\infty + \|Z\cdot M + N\|^2_{\text{BMO}}) \leq \max(\|Y\|^2_\infty, \|Z\cdot M + N\|^2_{\text{BMO}})$$
$$\leq \operatorname*{ess\,sup}_{[[0,T]]}\bigg[|Y_t|^2 + E\bigg(\int_t^T |\sigma_s^* Z_s|^2 dK_s + tr\langle N\rangle_{tT}|\mathcal{F}_t\bigg)\bigg]$$

we obtain from (3.3)

$$\begin{aligned}
&\frac{1}{4}\|Y\|_\infty^2+\frac{1}{2}\|Z\cdot M+N\|_{\text{BMO}}^2\\
&\le \|\xi\|^2+4\operatorname*{ess\,sup}_{[[0,T]]}E^2\bigg(\int_t^T |f(s,y_s,\sigma_s^* z_s)-f(s,0,0)|dK_t+\theta^2\mathrm{tr}\langle n\rangle_{tT}|\mathcal{F}_t\bigg) \qquad (3.5)\\
&\le \|\xi\|^2+16\operatorname*{ess\,sup}_{[[0,T]]}E^2\bigg(\int_t^T r_s^2y_s^2dK_s+\theta^2\mathrm{tr}\langle z\cdot M+n\rangle_{tT}|\mathcal{F}_t\bigg)\\
&\le \|\xi\|^2+16|r|_{2,\infty}^4\|y\|_\infty^4+16\theta^4\|z\cdot M+n\|_{\text{BMO}}^4.
\end{aligned}$$

Therefore

$$\begin{aligned}
\|Y\|_\infty^2+\|Z\cdot M+N\|_{\text{BMO}}^2 &\le 4\|\xi\|^2+64|r|_{2,\infty}^4\|y\|_\infty^4+64\theta^4\|z\cdot M+n\|_{\text{BMO}}^4\\
&\le 4\|\xi\|^2+\beta^2(\|y\|_\infty^2+\|z\cdot M+n\|_{\text{BMO}}^2)^2,
\end{aligned}$$

where $\beta=8\max(|r|_{2,\infty}^2,\theta^2)$. We can pick R such that

$$4\|\xi\|^2+\beta^2R^4\le R^2$$

if and only if $\|\xi\|_\infty\le\frac{1}{4\beta}$. For instance $R=2\sqrt{2}\|\xi\|_\infty$ satisfies this quadratic inequality. Therefore the ball

$$\mathcal{B}_R=\{(Y,Z\cdot M+N)\in S^\infty\times\text{BMO},\ N\perp M,\ \ \|Y\|_\infty^2+\|Z\cdot M+N\|_{\text{BMO}}^2\le R^2\}$$

is such that $F(\mathcal{B}_R)\subset\mathcal{B}_R$.

Similarly, for $(y^j,z^j\cdot M+n^j)\in\mathcal{B}_R$, $j=1,2$, using the notations $\delta y=y^1-y^2$, $\delta z=z^1-z^2$, $\delta n=n^1-n^2$, we can show that

$$\begin{aligned}
\|\delta Y\|_\infty^2+\|\delta Z\cdot M+\delta N\|_{\text{BMO}}^2 &\le 4\operatorname*{ess\,sup}_{[[0,T]]}E^2\bigg(\int_t^T |f(s,y_s^1,\sigma_s^*z_s^1)-f(s,y_s^2,\sigma_s^*z_s^2)|dK_s\\
&\qquad+\int_t^T|g_s|d\mathrm{var}(\mathrm{tr}\langle\delta n,n^1+n^2\rangle)_s|\mathcal{F}_t\bigg)\\
&\le 8\operatorname*{ess\,sup}_{[[0,T]]}E\bigg(\int_t^T(r_s^2|\delta y_s|^2+\theta^2|\sigma_s^*\delta z_s|^2dK_s)|\mathcal{F}_t\bigg)\\
&\qquad\times E\bigg(\int_t^T(r_s(|y_s^1|+|y_s^2|)+\theta(|\sigma_s^*z_s^1|+|\sigma_s^*z_s^2|)^2dK_s)|\mathcal{F}_t\bigg)\\
&\qquad+\theta^2E(\mathrm{tr}\langle\delta n\rangle_{tT}|\mathcal{F}_s)E(\mathrm{tr}\langle n^1+n^2\rangle_{tT}|\mathcal{F}_t)
\end{aligned}$$

Again using the equalities (3.4) we can pass to the norm. Thus

$$\begin{aligned}
\|\delta Y\|_\infty^2+\|\delta Z\cdot M+\delta N\|_{\text{BMO}}^2 &\le 8(|r|_{2,\infty}^2\|\delta y\|_\infty^2+\theta^2\|\delta z\cdot M\|_{\text{BMO}}^2)\\
&\times(|r|_{2,\infty}^2(\|y^1\|_\infty^2+\|y^2\|_\infty^2)+\theta^2(\|z^1\cdot M\|_P^2+\|z^2\cdot M\|_P^2)\\
&+2\theta^2\|\delta n\|_{\text{BMO}}^2(\|n^1\|_{\text{BMO}}^2+\|n^2\|_{\text{BMO}}^2)^2).
\end{aligned}$$

Since $\|z^1\cdot M\|,\|z^2\cdot M\|\le R$, $\|n^1\|,\|n^2\|\le R$, we get

$$\begin{aligned}
\|\delta Y\|_\infty^2+\|\delta Z\cdot M+\delta N\|_{\text{BMO}}^2 &\le 128\beta^2R^2(\|\delta y\|_\infty^2+\|\delta z\cdot M\|_{\text{BMO}}^2)+4\beta^2R^2\|\delta n\|_{\text{BMO}}^2\\
&\le 128\beta^2R^2(\|\delta y\|_\infty^2+\|\delta z\cdot M+\delta n\|_{\text{BMO}}^2). \qquad (3.6)
\end{aligned}$$

Now we can take $R = 2\sqrt{2}\|\xi\|_\infty < \frac{1}{8\sqrt{2}\beta}$. This means that $\|\xi\|_\infty < \frac{1}{32\beta}$ and F is contraction on $\mathcal{B}_R$. By contraction principle the mapping F admits a unique fixed point, which is the solution of (3.1). □

From now we suppose that $d = 1$.

Lemma 1. *Let condition* A) *is satisfied. Then the generator* $(\bar f, \bar g, \bar\xi)$, *where*

$$\bar f(t,\bar y,\bar z) = e^{\int_0^t \alpha_s dK_s}(f(t, e^{-\int_0^t e^{\alpha_s dK_s}}\bar y, e^{-\int_0^t e^{\alpha_s dK_s}}\bar z) - f(t,0,0)) - \alpha_t\bar y - \gamma_t^*\bar z,$$

$$\bar g_t = e^{-\int_0^t \alpha_s dK_s} g_t \quad \text{and} \quad \bar\xi = e^{\int_0^T \alpha_s dK_s}\xi,$$

satisfies condition A) *with* $\alpha = 0$, $\gamma = 0$, $\bar r_t = r_t e^{\|\int_0^\infty r_s dK_s\|_\infty}$ *and* $\bar\theta = \theta e^{\|\int_0^T r_s dK_s\|_\infty}$.

Moreover, (Y, Z, N) *is a solution of BSDE* (3.1) *if and only if*

$$(\bar Y_t, \bar Z_t, \bar N_t) = \left(e^{\int_0^t \alpha_s dK_s} Y_t, e^{\int_0^t \alpha_s dK_s} Z_t, \int_0^t e^{\int_0^s \alpha_u dK_u} dN_s\right)$$

is a solution w.r.t. measure $d\bar P = \mathcal{E}_T((\gamma\sigma^{-1})\cdot M)dP$ *of BSDE*

$$\begin{aligned} d\bar Y_t = -\bar f(t,\bar Y_t, \sigma_t^*\bar Z_t)dK_t - d\langle\bar N\rangle_t \bar g_t + \bar Z_t^* d\bar M_t + d\bar N_t,\\ \bar Y_T = \bar\xi, \end{aligned} \tag{3.7}$$

where $\bar M_t = M_t - \langle(\gamma\sigma^{-1})\cdot M, M\rangle_t$.

Proof. Condition A) for $(\bar f, \bar g, \bar\xi)$ is satisfied since by (2.4)

$$\begin{aligned} |\bar f(t,\bar y_1,\bar z_1) - \bar f(t,\bar y_2,\bar z_2)| &\le e^{\int_0^t \alpha_s dK_s}(r_t|\delta\bar y| + \theta|\delta\bar z|)(r_t(|\bar y_1| + |\bar y_2|) + \theta(|\bar z_1| + |\bar z_2|))\\ &\le (\bar r_t|\delta\bar y| + \bar\theta|\delta\bar z|)(\bar r_t(|\bar y_1| + |\bar y_2|) + \bar\theta(|\bar z_1| + |\bar z_2|)). \end{aligned}$$

On the other hand, using the Itô formula we have

$$\begin{aligned} d\bar Y_t &= e^{\int_0^t \alpha_s dK_s} dY_t + \alpha_t e^{\int_0^t \alpha_s dK_s} Y_t dK_t\\ &= e^{\int_0^t \alpha_s dK_s}(f(t,0,0) - f(t,Y_t,\sigma_t^* Z_t))dK_t + e^{\int_0^t \alpha_s dK_s} d\langle N\rangle_t g_t\\ &\quad + e^{\int_0^t \alpha_s dK_s} Z_t^* dM_t + e^{\int_0^t \alpha_s dK_s} dN_t + \alpha_t\bar Y_t dK_t \end{aligned}$$

Taking into account that

$$e^{\int_0^t \alpha_s dK_s}(f(t,0,0) - f(t,Y_t,\sigma_t^* Z_t)) + \alpha_t\bar Y_t = -\bar f(t,\bar Y_t,\sigma_t^*\bar Z_t) - \gamma_t\sigma_t^*\bar Z_t,$$

$$e^{\int_0^t \alpha_s dK_s} d\langle N\rangle_t g_t = d\langle\bar N\rangle_t e^{-\int_0^t \alpha_s dK_s} g_t = d\langle\bar N\rangle_t\bar g_t$$

and

$$\bar Z\cdot M - \int_0^\cdot \gamma_t\sigma_t^*\bar Z_t dK_t = \bar Z\cdot M - \int_0^\cdot \gamma_t\sigma_t^{-1}\sigma_t\sigma_t^*\bar Z_t dK_t$$

$$= \bar Z\cdot M - \int_0^\cdot \gamma_t\sigma_t^{-1} d\langle M\rangle_t\bar Z_t = \bar Z\cdot M - \langle(\gamma\cdot\sigma^{-1})\cdot M, \bar Z\cdot M\rangle = \bar Z\cdot\bar M$$

we obtain

$$d\bar Y_t = -\bar f(t,\bar Y_t,\sigma_t^*\bar Z_t)dK_t - d\langle\bar N\rangle_t\bar g_t + \bar Z_t d\bar M_t + d\bar N_t.$$

Here $\bar M$ is a local martingale w.r.t. $\bar P$ by Girsanov theorem. □

Corollary 1. *Let f and g satisfy condition A) and $\|\xi\|_\infty \le \frac{1}{32\beta}\exp(-2\|\int_0^T r_s dK_s\|_\infty)$. Then there exist the solution of* (3.1) *with the norm $\|Y\|_\infty^2 + \|Z\cdot\bar{M}+N\|^2_{\mathrm{BMO}(\bar{\mathrm{P}})} \le \frac{1}{128\beta^2}$.*

Proof. It is obvious that

$$\|Y\|_\infty^2 + \|Z\cdot\bar{M}+N\|^2_{\mathrm{BMO}(\bar{\mathrm{P}})} \le \left(\|\bar{Y}\|_\infty^2 + \|\bar{Z}\cdot\bar{M}+\bar{N}\|^2_{\mathrm{BMO}(\bar{\mathrm{P}})}\right)\exp\left(2\|\int_0^T r_s dK_s\|_\infty\right)$$
$$\le 8\|\bar{\xi}\|_\infty^2 \exp\left(2\|\int_0^T r_s dK_s\|_\infty\right) \le 8\|\xi\|_\infty^2 \exp\left(4\|\int_0^T r_s dK_s\|_\infty\right).$$

From $\|\xi\|_\infty \le \frac{1}{32\beta}\exp(-2\|\int_0^T r_s dK_s\|_\infty)$ follows that $8\|\xi\|_\infty \exp(4\|\int_0^T r_s dK_s\|_\infty) \le \frac{1}{128\beta^2}$. Hence we get $\|Y\|_\infty^2 + \|Z\cdot\bar{M}+N\|^2_{\mathrm{BMO}(\bar{\mathrm{P}})} \le \frac{1}{128\beta^2}$. □

Corollary 2. *Let generator (f,g,ξ) satisfies conditions* B1)–B3) *and $(\tilde{Y}_t, \tilde{Z}_t, \tilde{N}_t)$ be a solution of* (3.1)*. Then BSDE*

$$\begin{aligned} d\hat{Y}_t = (f(t,\tilde{Y}_t,\sigma_t^*\tilde{Z}_t) - f(t,\hat{Y}_t+\tilde{Y}_t,\sigma_t^*\hat{Z}_t+\sigma_t^*\tilde{Z}_t))dK_t \\ -d(\langle\hat{N}\rangle_t + 2\langle\tilde{N},\hat{N}\rangle_t)g_t + \hat{Z}_t^* dM_t + d\hat{N}_t, \\ \hat{Y}_T = \hat{\xi} \end{aligned} \tag{3.8}$$

satisfy condition A) *with $-\hat{f}(t,y,z) = f(t,\tilde{Y}_t,\sigma_t^*\tilde{Z}_t) - f(t,y+\tilde{Y}_t,z+\sigma_t^*\tilde{Z}_t)$, $\alpha_t = f_y(t,\tilde{Y}_t,\sigma_t^*\tilde{Z}_t)$, $\gamma_t = f_z(t,\tilde{Y}_t,\sigma_t^*\tilde{Z}_t)$ and the new probability measure $\mathcal{E}_T(2g\cdot\tilde{N})dP$. Moreover (3.8) admits a unique solution $(\hat{Y}_t,\hat{Z}_t,\hat{N}_t)$ if $\|\hat{\xi}\|_\infty \le \frac{1}{32\beta}\exp(-2\|\int_0^{\cdot} r_s dK_s\|_\infty)$.*

Proof. Using a change of measure equation (3.8) reduces to equation of type (3.1). By previous corollary we obtain the existence and uniqueness of the BSDE. □

Lemma 2. *Let conditions* B1)–B3) *be satisfied and random variables $\tilde{\xi}$ and $\hat{\xi}$ be such that $max(\|\tilde{\xi}\|_\infty, \|\hat{\xi}\|_\infty) \le \frac{1}{32\beta}e^{-2\|\int_0^T r_s^2 dK_s\|_\infty}$. Then there exist solutions of BSDEs (3.8) and*

$$\begin{aligned} d\tilde{Y}_t = (f(t,0,0) - f(t,\tilde{Y}_t,\sigma_t^*\tilde{Z}_t))dK_t - d\langle\tilde{N}\rangle_t g_t + \tilde{Z}_t^* dM_t + d\tilde{N}_t, \\ \tilde{Y}_T = \tilde{\xi} \end{aligned} \tag{3.9}$$

and the triple $(Y,Z,N) = (\tilde{Y}+\hat{Y}, \tilde{Z}+\hat{Z}, \tilde{N}+\hat{N})$ satisfies BSDE

$$\begin{aligned} dY_t = (f(t,0,0) - f(t,Y_t,\sigma_t^* Z_t))dK_t - d\langle N\rangle_t g_t + Z_t^* dM_t + dN_t, \\ Y_T = \tilde{\xi} + \hat{\xi}. \end{aligned}$$

Proof. Similarly to the Remark from Section 1 we can show that for

$$\hat{f}(t,y,z) = f(t,\tilde{Y}_t,\sigma_t^*\tilde{Z}_t) - f(t,y+\tilde{Y}_t,\sigma_t^* z+\sigma_t^*\tilde{Z}_t),$$

$\alpha_t = f_y(t,\tilde{Y}_t,\sigma_t^*\tilde{Z}_t)$, $\gamma_t = f_z(t,\hat{Y}_t,\sigma_t^*\hat{Z}_t)$, the estimate

$$\begin{aligned} |\hat{f}(t,y_1,z_1) - \hat{f}(t,y_2,z_2) - \alpha_t\delta y - \gamma_t^*\delta z| \\ \le (r_t|\delta y| + \theta|\delta z|)(r_t(|y_1|+|y_2|) + \theta(|z_1|+|z_2|)) \end{aligned}$$

holds.

Now by Lemma 1 and Corollary 2 of Lemma 1 we obtain the solvability of both equations (3.9), (3.8). □

Proposition 2. *Let f and g satisfy condition* B1)–B3) *and $\xi \in L^\infty$. Then BSDE* (2.1) *admits a solution $(Y, Z \cdot M + N) \in S^\infty \times \mathrm{BMO}$.*

Proof. An arbitrary $\xi \in L^\infty(R)$ can be represented as sum $\xi = \sum_{i=1}^m \xi_i$ with $\|\xi_i\|_\infty \le \frac{1}{32\beta}\exp(-2\|\int_0^\cdot r_s dK_s\|_\infty)$. Denote by $(Y^j, Z^j, N^j),\ j = 1, \dots, m$, the solution of

$$
\begin{aligned}
dY_t^j &= (f(t, Y_t^0 + \dots + Y_t^{j-1}, \sigma_t^*(Z_t^0 + \dots + Z_t^{j-1})) \\
&\quad - f(t, Y_t^0 + \dots + Y_t^j, \sigma_t^*(Z_t^0 + \dots + Z_t^j))dK_t \\
&\quad - d(\langle N^j\rangle_t + 2\langle N^j, N^0 + \dots + N^{j-1}\rangle_t)g_t + Z_t^{j*}dM_t + dN_t^j, \qquad (3.10)\\
Y_T^j &= \xi^j, \\
Y^0 &= 0, \quad Z^0 = 0 \quad N^0 = 0.
\end{aligned}
$$

By Corollary 1 we get

$$\|Y^j\|_\infty^2 + \|Z^j \cdot M^j + N^j\|^2_{\mathrm{BMO(P^j)}} \le \frac{1}{128\beta^2},$$

where $dP^j = \mathcal{E}_T(\int_0^\cdot f_z(s, Y_s^0 + \dots + Y_s^{j-1}, \sigma_s^*(Z_s^0 + \dots + Z_s^{j-1}))\sigma_s^{-1}dM_s)dP$, and $M^j = M - \langle f_z(\cdot, Y^0 + \dots + Y^{j-1}, \sigma^*(Z^0 + \dots + Z^{j-1}))\sigma^{-1} \cdot M, M\rangle$.

Using Lemma 2 we get the existence of a solution for BSDE

$$
\begin{aligned}
d\bar{Y}_t = (f(t,0,0) - f(t, \bar{Y}_t, \sigma_t^* Z_t))dK_t - d\langle N\rangle_t g_t + Z_t^* dM_t + dN_t, \\
\bar{Y}_T = \xi.
\end{aligned}
$$

Since $\int_0^T f(t,0,0)dK_t$ is bounded we can apply the above argument with f replaced by $\bar{f}(t,y,z) = f(t, y - \int_0^t f(s,0,0)dK_s, z)$ to get the existence of solution

$$
\begin{aligned}
d\bar{Y}_t = (f(t,0,0) - f(t, \bar{Y}_t - \int_0^t f(s,0,0)dK_s, \sigma_t^* Z_t))dK_t - d\langle N\rangle_t g_t + Z_t^* dM_t + dN_t, \\
\bar{Y}_T = \xi + \int_0^T f(s,0,0)dK_s.
\end{aligned}
$$

Obviously, $Y_t = \bar{Y}_t - \int_0^t f(s,0,0)dK_s$ is a solution of BSDE (2.1), (2.2). □

4. A Comparison Theorem for BSDEs

Let us consider BSDE (2.1),(2.2) in the case $d = 1$.

Lemma 3. *Let $\xi \in L_\infty$ and assume that there are positive constants $C(f), C(g)$, increasing function $\lambda : R^+ \to R^+$, bounded on all bounded subsets and a predictable process $k \in H^2(R, 1)$ such that*

$$|f(t,y,z)| \le k_t^2 \lambda(|y|) + C(f)z^2, \tag{4.1}$$

$$|g(t)| \le C(g). \tag{4.2}$$

Then the martingale part of any bounded solution of (2.1), (2.2) *belongs to the space* $\mathrm{BMO}(P)$.

Proof. Let Y be a solution of (2.1), (2.2) and there is a constant $C > 0$ such that

$$|Y_t| \leq C \quad \text{a.s.} \quad \text{for all} \quad t.$$

Applying the Itô formula for $\exp\{\beta Y_T\} - \exp\{\beta Y_\tau\}$ and using the boundary condition $Y_T = \xi$ we have

$$\begin{aligned}\frac{\beta^2}{2}\int_\tau^T e^{\beta Y_s} Z_s^* d\langle M\rangle_s Z_s + \frac{\beta^2}{2}\int_\tau^T e^{\beta Y_s} d\langle N\rangle_s - \beta\int_\tau^T e^{\beta Y_s} f(s, Y_s, Z_s) dK_s \\ - \beta\int_\tau^T e^{\beta Y_s} g(s) d\langle N\rangle_s + \beta\int_\tau^T e^{\beta Y_s} Z_s^* dM_s + \beta\int_\tau^T e^{\beta Y_s} dN_s \\ = e^{\beta\xi} - e^{\beta Y_\tau} \leq e^{\beta C}, \quad (4.3)\end{aligned}$$

where β is a constant yet to be determined.

If $Z \cdot M$ and N are square integrable martingales taking conditional expectations in (4.3) we obtain

$$\begin{aligned}\frac{\beta^2}{2} E\Big(\int_\tau^T e^{\beta Y_s} Z_s^* d\langle M\rangle_s Z_s | F_\tau\Big) + \frac{\beta^2}{2} E\Big(\int_\tau^T e^{\beta Y_s} d\langle N\rangle_s | F_\tau\Big) \\ \leq e^{\beta C} + \beta E\Big(\int_\tau^T e^{\beta Y_s} |f(s, Y_s, Z_s)| dK_s | F_\tau\Big) + \beta E\Big(\int_\tau^T e^{\beta Y_s} |g(s)| d\langle N\rangle_s | F_\tau\Big).\end{aligned}$$

Now if we use the estimates (4.1), (4.2), we get

$$\begin{aligned}\frac{\beta^2}{2} E\Big(\int_\tau^T e^{\beta Y_s} Z_s^* d\langle M\rangle_s Z_s | F_\tau\Big) + \frac{\beta^2}{2} E\Big(\int_\tau^T e^{\beta Y_s} d\langle N\rangle_s | F_\tau\Big) \\ \leq e^{\beta C} + \beta\lambda(C) E\Big(\int_\tau^T e^{\beta Y_s} k_s^2 dK_s | F_\tau\Big) \\ + \beta C(f) E\Big(\int_\tau^T e^{\beta Y_s} |\sigma_s^* Z_s|^2 dK_s | F_\tau\Big) + \beta E\Big(\int_\tau^T e^{\beta Y_s} |g(s)| d\langle N\rangle_s | F_\tau\Big) \\ \leq e^{\beta C} + \beta\lambda(C) E\Big(\int_\tau^T e^{\beta Y_s} k_s^2 dK_s | F_\tau\Big) \\ + \beta C(f) E\Big(\int_\tau^T e^{\beta Y_s} |Z_s^* d\langle M\rangle_s Z_s|^2 | F_\tau\Big) + C(g)\beta E\Big(\int_\tau^T e^{\beta Y_s} d\langle N\rangle_s | F_\tau\Big).\end{aligned}$$

Conditions (4.1) and (4.2) imply that

$$\begin{aligned}\Big(\frac{\beta^2}{2} - \beta C(f)\Big) E\Big(\int_\tau^T e^{\beta Y_s} Z_s^* d\langle M\rangle_s Z_s | F_\tau\Big) \\ + \Big(\frac{\beta^2}{2} - \beta C(g)\Big) E\Big(\int_\tau^T e^{\beta Y_s} d\langle N\rangle_s | F_\tau\Big) \\ \leq e^{\beta C} + \beta\lambda(C) E\Big(\int_\tau^T e^{\beta Y_s} k_s^2 dK_s | F_\tau\Big). \quad (4.4)\end{aligned}$$

Taking $\beta = 4\overline{C}$, where $\overline{C} = \max(C(f), C(g))$, from (4.4) we have

$$4\overline{C}^2\Big[E\Big(\int_\tau^T e^{\beta Y_s} Z_s^* d\langle M\rangle_s Z_s|F_\tau\Big) + E\Big(\int_\tau^T e^{\beta Y_s} d\langle N\rangle_s|F_\tau\Big)\Big]$$
$$\le e^{4C\overline{C}}\big(4\overline{C}\lambda(C)\|k\|_H + 1\big).$$

Since $Y \ge -C$, from the latter inequality we finally obtain the estimate

$$E\big(\langle Z\cdot M\rangle_{\tau T}|F_\tau\big) + E\big(\langle N\rangle_{\tau T}|F_\tau\big) \le \frac{e^{8C\overline{C}}[4\overline{C}\lambda(C)\|k\|_H + 1]}{4\overline{C}^2} \tag{4.5}$$

for any stopping time τ, hence $Z\cdot M, N \in BMO$.

For general $Z\cdot M$ and N we stop at τ_n and derive (4.5) with T replaced τ_n. Letting $n\to\infty$, the proof is completed. □

Further we use some notations. Let (Y, Z), $(\widetilde{Y}, \widetilde{Z})$ be two pairs of processes and (f, g, ξ), $(\tilde{f}, \tilde{g}, \tilde{\xi})$ be two triples of generators. We denote:

$$\delta f = f - \tilde{f},\ \ \delta g = g - \tilde{g},\ \ \delta\xi = \xi - \tilde{\xi},$$

$$\partial_y f(t, Y_t, \widetilde{Y}_t, Z_t) \equiv \partial f_y(t) = \frac{f(t, Y_t, Z_t) - f(t, \widetilde{Y}_t, Z_t)}{Y_t - \widetilde{Y}_t}$$

$$\text{for all } \ j = 1, \dots, n,\ \ \partial_j f(t, \widetilde{Y}_t, Z_t, \widetilde{Z}_t) \equiv \partial_j f(t)$$

$$= \frac{f(t, \widetilde{Y}_t, Z_t^1, \dots, Z_t^{j-1}, Z_t^j, \widetilde{Z}_t^{j+1}, \dots, \widetilde{Z}_t^n) - f(t, \widetilde{Y}_t, Z_t^1, \dots, Z_t^{j-1}, \widetilde{Z}_t^j, \widetilde{Z}_t^{j+1}, \dots, \widetilde{Z}_t^n)}{Z_t^j - \widetilde{Z}_t^j},$$

$$\nabla f(t) = (\partial_1 f(t), \dots, \partial_n f(t))^*.$$

Thus we have

$$f(t, Y_t, Z_t) - f(t, \widetilde{Y}_t, \widetilde{Z}_t) = \partial_y f(t)\delta Y_t + \nabla f(t)^*\delta Z_t. \tag{4.6}$$

Theorem 2. *Let Y and $\widetilde{Y}$ be the bounded solutions of SBE* (2.1) *with generators* (f, g, ξ) *and* $(\tilde{f}, \tilde{g}, \tilde{\xi})$ *respectively, satisfying the conditions of Lemma* 3.

If $\xi \ge \tilde{\xi}$ (a.s), $f(t, y, z) \ge \tilde{f}(t, y, z)$ (μ^K-a.e.), $g(t) \ge \tilde{g}(t)$ ($\mu^{\langle N\rangle}$-a.e.) and f (or $\tilde{f}$) satisfies the following Lipschitz condition:

L1) *for any $Y, \widetilde{Y}, Z$*

$$\frac{f(t, Y_t, Z_t) - f(t, \widetilde{Y}_t, Z_t)}{Y_t - \widetilde{Y}_t} \in S^\infty,$$

L2) *for any $Z, \widetilde{Z} \in H^2$ and any bounded process Y*

$$(\sigma_t\sigma_t^*)^{-1}\nabla f(t, Y_t, Z_t, \widetilde{Z}_t) \in H^2(R^n, \sigma),$$

then $Y_t \ge \widetilde{Y}_t$ a.s. for all $t \in [0, T]$.

Proof. Taking the difference of the equations (2.1), (2.2) with generators (f, g, ξ) and $(\widetilde{f}, \widetilde{g}, \widetilde{\xi})$ respectively, we have

$$Y_t - \widetilde{Y}_t = Y_0 - \widetilde{Y}_0 - \int_0^t [f(s, Y_s, Z_s) - f(s, \widetilde{Y}_s, \widetilde{Z}_s)]dK_s$$

$$-\int_0^t [f(s,\widetilde{Y}_s,\widetilde{Z}_s)-\tilde{f}(s,\widetilde{Y}_s,\widetilde{Z}_s)]dK_s-\int_0^t [g(s)-\tilde{g}(s)]d\langle N\rangle_s$$
$$-\int_0^t \widetilde{g}(s)d(\langle N\rangle_s-\langle\widetilde{N}\rangle_s)+\int_0^t (Z_s-\widetilde{Z}_s)dM_s+N_t-\widetilde{N}_t. \tag{4.7}$$

Let us define the measure Q by $dQ=\mathcal{E}_T(\Lambda)dP$, where

$$\Lambda_t=\int_0^t \nabla f(s)^*(\sigma_s\sigma_s^*)^{-1}dM_s+\int_0^t \widetilde{g}(s)d(N_s+\widetilde{N}_s).$$

By Lemma 3 $Z,\widetilde{Z}\in H^2$ and N, $\widetilde{N}$ are BMO- martingales. Therefore Condition $L1),L2)$ and (4.2) imply that $\Lambda\in BMO$ and hence Q is a probability measure equivalent to P.

Denote by $\bar{\Lambda}$ the martingale part of $\delta Y=Y-\widetilde{Y}$, i.e.,

$$\bar{\Lambda}=(Z-\widetilde{Z})\cdot M+N-\widetilde{N}.$$

Therefore, by Girsanov's Theorem and by (4.6) the process

$$\delta Y_t+\int_0^t (\partial_y f(s)\delta Y_s+\nabla f(s)^*\delta Z_s)dK_s+\int_0^t \delta f(s,\widetilde{Y}_s,\widetilde{Z}_s)dK_s+\int_0^t \delta g(s)d\langle N\rangle_s$$
$$=\delta Y_t+\int_0^t (\partial_y f(s)\delta Y_s+\delta f(s,\widetilde{Y}_s,\widetilde{Z}_s))dK_s$$
$$+\int_0^t \nabla f(s)^*(\sigma_s\sigma_s^*)^{-1}d\langle M\rangle_s\delta Z_s+\int_0^t \delta g(s)d\langle N\rangle_s$$
$$=-\int_0^t \widetilde{g}(s)d(\langle N\rangle_s-\langle\widetilde{N}\rangle_s)+\int_0^t (Z_s-\widetilde{Z}_s)dM_s+N_t-\widetilde{N}_t=\bar{\Lambda}_t-\langle\Lambda,\bar{\Lambda}\rangle_t$$

is a local martingale under Q. Moreover, since by Lemma 3 $\bar{N}\in BMO$, Proposition 11 of [6] implies that

$$\bar{\Lambda}_t-\langle\Lambda,\bar{\Lambda}\rangle_t\in BMO(Q).$$

Thus, using the martingale property and the boundary conditions $Y_T=\xi,\widetilde{Y}_T=\widetilde{\xi}$ we have

$$Y_t-\widetilde{Y}_t$$
$$=E^Q\left(e^{\int_t^T\partial_y f_s dK_s}(\xi-\tilde{\xi})+\int_t^T e^{\int_t^s\partial_y f_u dK_u}(f(s,\widetilde{Y}_s,\widetilde{Z}_s)-\tilde{f}(s,\widetilde{Y}_s,\widetilde{Z}_s))dK_s|F_t\right)$$
$$+E^Q\left(\int_t^T e^{\int_t^s\partial_y f_u dK_u}(g(s)-\widetilde{g}(s))d\langle N\rangle_s|F_t\right),$$

which implies that $Y_t\geq\widetilde{Y}_t$ a.s. for all $t\in[0,T]$. □

Corollary 3. *Let condition* A) *be satisfied. Then if the solution of* (2.1), (2.2) *exists, it is unique.*

The proof of **Theorem 1** follows now from the last corollary and Proposition 2.

Remark 3. Conditions L1), L2) are satisfied if there is constant $C>0$ such that

$$|f(t,y,z)-f(t,\tilde{y},\tilde{z})|\leq C|y-\tilde{y}|+C|z-\tilde{z}|(|z|+|\tilde{z}|)$$

and $tr(\sigma_t\sigma_t^*)^{-1} \leq C$ for all $y, \tilde{y} \in R,\ z, \tilde{z} \in R^n\ t \in [0,T]$. Conditions L1),L2) are also fulfilled if $f(t,y,z)$ satisfies the global Lipschitz condition and $M \in BMO$.

ACKNOWLEDGEMENTS

I am thankful to M. Mania for useful discussions and remarks. I also would like to thank the referee for valuable remarks and suggestions which lead to many improvements in the original version of the paper .

REFERENCES

[1] Bismut J. M. (1973) Conjugate convex functions in optimal stochastic control, *J. Math. Anal. Appl.* **44**, 384–404.

[2] Bobrovnytska O. and Schweizer M. (2004) Mean-Variance Hedging and Stochastic Control: Beyond the Brownian Setting, IEEE Transactions on Automatic Control 49, 396-408.

[3] Buckdahn R., (1993) Backward stochastic differential equation driven by Martingale, Preprint.

[4] Chitashvili R., (1983) Martingale ideology in the theory of controlled stochastic processes,*Lect. Notes in Math.*, Springer, Berlin, N. 1021, pp. 73-92.

[5] Dellacherie C. and Meyer P. A., (1980) *Probabilités et potentiel. Chapitres V a VIII. Théorie des martingales.* Actualités Scientifiques et Industrielles Hermann, Paris.

[6] Doleans-Dade K. and Meyer P. A., (1979) Inégalités de normes avec poinds, Séminaire de Probabilités XIII, Lect. Notes in Math., Springer, Berlin, N. 721, pp. 204-215.

[7] El Karoui N., Huang S.J., (1997) A general result of existence and uniqueness of backward stochastic differential equations. Pitman Res. Notes Math. Ser.,364, Longman, Harlow, 27-36.

[8] Hu Y., Imkeller P. and Müller M., (2005), Utility maximization in incomplete markets. Ann. Appl. Probab. 15 , no. 3, 1691-1712.

[9] Jacod J., (1979) *Calcule Stochastique et problèmes des martingales. Lecture Notes in Math.*, Springer, Berlin, N.714.

[10] Kazamaki N., (1994) *Continuous exponential martingales and BMO, Lecture Notes in Math.*, Springer, Berlin, N. 1579.

[11] Kobylanski M., (2000) Backward stochastic differential equation and partial differential equations with quadratic growth, The Annals of Probability, vol. 28, N2, 558-602.

[12] Lepeltier J.P. and San Martin J., (1998) Existence for BSDE with superlinear-quadratic coefficient, Stoch. Stoch. Rep. 63, 227-240.

[13] Liptser R.Sh. and Shiryaev A.N., (1986) *Martingale theory*, Nauka, Moscow.

[14] Mania M. and Schweizer M., (2005) Dynamic exponential utility indifference valuation, Ann. Appl. Probab. 15, no. 3, 2113–2143.

[15] Mania M. and Tevzadze R., (2000) A Semimartingale Bellman equation and the variance-optimal martingale measure, *Georgian Math. J.* **7** 4 , 765–792.

[16] Mania M. and R. Tevzadze R. (2003) A Unified Characterization of q-optimal and minimal entropy martingale measures by Semimartingale Backward Equations,*Georgian Math. J.* vol.10, No 2, 289-310.

[17] Mania M. and Tevzadze R. (2003) Backward Stochastic PDE and Imperfect Hedging, International Journal of Theoretical and Applied Finance, vol.6, 7,663-692.

[18] Mania M. Santacroce M. and Tevzadze R. (2003) A Semimartingale BSDE related to the minimal entropy martingale measure, Finance and Stochastics, vol.7, No 3, 385-402.

[19] Mania M. and Tevzadze R. (2003) A Semimartingale Bellman equation and the variance-optimal martingale measure under general information flow, SIAM Journal on Control and Optimization, Vol. 42, N5, 1703-1726.

[20] Mania M. and Tevzadze R. (2005) An exponential martingale equation, From Stochastic Calculus to Mathematical Finance, The Shiryaev Festschrift, Springer, 507-516.

[21] Morlais M.A., Quadratic Backward Stochastic Differential Equations (BSDEs) Driven by a Continuous Martingale and Application to the Utility Maximization Problem, http://hal.ccsd.cnrs.fr/ccsd-00020254/

[22] Pardoux E. and Peng S.G. (1990) Adapted solution of a backward stochastic differential equation, *Systems Control Lett.* **14** ,55–61.

L^2-APPROXIMATING PRICING UNDER RESTRICTED INFORMATION

M. MANIA, R. TEVZADZE AND T. TORONJADZE

ABSTRACT. We consider the mean-variance hedging problem under partial information in the case where the flow of observable events does not contain the full information on the underlying asset price process. We introduce a certain type martingale equation and characterize the optimal strategy in terms of the solution of this equation. We give relations between this equation and backward stochastic differential equations for the value process of the problem.

Key words and phrases: Semimartingale, incomplete markets, mean-variance hedging, partial information, backward stochastic differential equation.

MSC 2010: 90A09, 60H30, 90C39.

1. INTRODUCTION

We assume that the dynamics of the price process of the asset traded on a market is described by a continuous semimartingale $S = (S_t, t \in [0,T])$ defined on a filtered probability space $(\Omega, F, \mathcal{F} = (\mathcal{F}_t, t \in [0,T], P)$, satisfying the usual conditions, where $F = \mathcal{F}_T$ and $T < \infty$ is the fixed time horizon. Suppose that the interest rate is equal to zero and the asset price process satisfies the structure condition, i.e., the process S admits the decomposition

$$S_t = S_0 + M_t + \int_0^t \lambda_u d\langle M \rangle_u, \quad \langle \lambda \cdot M \rangle_T < \infty \quad \text{a.s.}, \tag{1.1}$$

where M is a continuous $\mathcal{F}-$local martingale and λ is a $\mathcal{F}$-predictable process.

Let us introduce an additional filtration smaller than $\mathcal{F}$

$$\mathcal{G}_t \subseteq \mathcal{F}_t, \quad \text{for every} \quad t \in [0,T].$$

The filtration $\mathcal{G}$ represents the information that the hedger has at his disposal, i.e., hedging strategies have to be constructed using only information available in $\mathcal{G}$. We assume that the filtration $\mathcal{G}$ also satisfies the usual conditions.

Let H be a P-square integrable $\mathcal{F}_T$-measurable random variable, representing the payoff of a contingent claim at time T.

We consider the mean-variance hedging problem

$$\text{to minimize} \quad E[(X_T^{x,\pi} - H)^2] \quad \text{over all} \quad \pi \in \Pi(\mathcal{G}), \tag{1.2}$$

where $\Pi(\mathcal{G})$ is a class of $\mathcal{G}$-predictable S-integrable processes to be specified later. Here $X_t^{x,\pi} = x + \int_0^t \pi_u dS_u$ is the wealth process starting from initial capital x, determined by the self-financing trading strategy $\pi \in \Pi(\mathcal{G})$.

In the case $\mathcal{G} = \mathcal{F}$ of complete information the mean-variance hedging problem was introduced by Föllmer and Sondermann [9] in the case when S is a martingale and then

Published in *Appl. Math. Optim.* **60** (2009), no. 1, 39–70.

developed by several authors for price process admitting a trend (see, e.g., [7], [13], [26],[27], [25], [11], [12], [1]).

Asset pricing with partial information under various setups has been considered. The mean-variance hedging problem under partial information was first studied by Di Masi, Platen and Runggaldier (1995) when the stock price process is a martingale and the prices are observed only at discrete time moments. For a general filtrations and when the asset price process is a martingale this problem was solved by Schweizer (1994) in terms of $\mathcal{G}$-predictable projections. Pham (2001) considered the mean-variance hedging problem for a general semimartingale model, assuming that the observable filtration contains the augmented filtration $\mathcal{F}^S$ generated by the asset price process S

$$\mathcal{F}_t^S \subseteq \mathcal{G}_t, \quad \text{for every} \quad t \in [0,T]. \tag{1.3}$$

In this paper, using the variance-optimal martingale measure with respect to the filtration $\mathcal{G}$ and suitable Kunita-Watanabe decomposition, the theory developed by Gourieroux, Laurent and Pham (1998) and Rheinländer and Schweizer (1997) was extended to the case of partial information .

If $\mathcal{F}^S \subseteq \mathcal{G}$, then the price process is a $\mathcal{G}$-semimartingale and the sharp bracket $\langle M\rangle$ is $\mathcal{G}$-adapted. If $\mathcal{G}$ is not containing $\mathcal{F}^S$, then S is not a $\mathcal{G}$-semimartingale and the problem is more involved. At the beginning of Section 3 under mild conditions (see Proposition 3.1) we derive a 'forward-backward' equation which gives a necessary condition of optimality. In the case when S is a martingale this equation admits an explicit solution and gives the optimal strategy constructed by Schweizer (1994). We focus our attention to the case when the filtration $\mathcal{G}$ of observable events does not contain the full information about the asset price process S. In this case this equation is hard to solve and we require the following additional assumptions:

A) $\langle M\rangle$ and λ are $\mathcal{G}$-predictable,

B) any $\mathcal{G}$- martingale is an $\mathcal{F}$-local martingale,

D) there exists a martingale measure for S that satisfies the Reverse Hölder condition (see definition in Section 2).

Denote by $\widehat{Y}$ and ${}^pY$ - the $\mathcal{G}$-optional and $\mathcal{G}$ predictable projections of a process Y. For the processes of finite variation, by Y^p we denote the dual $\mathcal{G}$-predictable projections. Condition A) implies that

$$\widehat{S}_t = E(S_t|\mathcal{G}_t) = S_0 + \int_0^t \lambda_u d\langle M\rangle_u + \widehat{M}_t.$$

Let

$$H_t = E[H] + \int_0^t h_u dM_u + L_t^H$$

and

$$H_t = E[H] + \int_0^t h_u^{\mathcal{G}} d\widehat{M}_u + L_t^{H,\mathcal{G}}$$

be the Galtchouk-Kunita-Watanabe (GKW) decompositions of $H_t = E(H|\mathcal{F}_t)$ with respect to local martingales M and $\widehat{M}$, where $h, h^{\mathcal{G}}$ are $\mathcal{F}$-predictable process and $L^H, L^{H,\mathcal{G}}$ are $\mathcal{F}$-local martingales strongly orthogonal to M and $\widehat{M}$ respectively. We shall use also notations:

$$\rho_t^2 = \frac{d\langle\, \widehat{M}\, \rangle_t}{d\langle M\rangle_t}, \quad \widetilde{h}_t = {}^p(h^{\mathcal{G}})_t \rho_t^2 - {}^p h_t \quad \text{and} \quad \widetilde{H} = \widehat{H}_T - \int_0^T \frac{\widetilde{h}_t}{1-\rho_t^2} d\widehat{S}_t - x.$$

We introduce the following martingale equation

$$\widetilde{Y}_T = \widetilde{H} - \int_0^T \frac{1}{1-\rho_t^2}\left[\lambda_t \widetilde{Y}_{t-} + \rho_t^2 \widetilde{\psi}_t\right]\left(\lambda_t d\langle M\rangle_t + d\widehat{M}_t\right). \tag{1.4}$$

The solution of this equation is a pair $(\widetilde{Y}, \widetilde{\psi})$, where $\widetilde{Y}$ is a square integrable martingale and $\widetilde{\psi}$ is defined by the GKW decomposition of $\widetilde{Y}$

$$\widetilde{Y}_t = \widetilde{Y}_0 + \int_0^t \widetilde{\psi}_u d\widehat{M}_u + L_t, \quad \langle\, \widehat{M}, L\rangle = 0. \tag{1.5}$$

Now we formulate the main result of the paper which is proved in Section 3.

Theorem. *Let conditions A), B) and D) be satisfied. Assume also that $E\widetilde{H}^2 < \infty$ and $\rho_t^2 < 1$ for all $t \in [0,T]$. Then there exists a unique solution $(\tilde{Y}, \tilde{\psi})$ of equation (1.4) and the strategy π^* is optimal if and only if it admits the representation*

$$\pi_t^* = \frac{1}{1-\rho_t^2}\left(\widetilde{h}_t + \lambda_t \widetilde{Y}_{t-} + \rho_t^2 \widetilde{\psi}_t\right). \tag{1.6}$$

In Section 4 (see Propositions 4.2 and 4.3), we establish connections between equation (1.4) and BSDEs for the value process of the problem (1.2) derived in [21], additionally assuming that

C) the filtration $\mathcal{G}$ is continuous, i.e., all $\mathcal{G}$-local martingales are continuous.

It was shown in [21] that the optimal strategy is determined by

$$\pi_t^* = \frac{\lambda_t V_t(1) + \rho_t^2 \varphi_t(1) - \widehat{X}_t^{\pi^*}(\lambda_t V_t(2) + \rho_t^2 \varphi_t(2))}{1-\rho_t^2 + \rho_t^2 V_t(2)}, \quad \widehat{X}_0^{\pi^*} = x, \tag{1.7}$$

where the triples $(V(1), \varphi(1), L(1))$ and $(V(2), \varphi(2), L(2))$ satisfy the following system of BSDEs

$$dV_t(1) = \frac{(\lambda_t V_t(2) + \rho_t^2 \varphi_t(2))(\lambda_t V_t(1) + \rho_t^2 \varphi_t(1))}{1-\rho_t^2 + \rho_t^2 V_t(2)} d\langle M\rangle_t + \varphi_t(1) d\widehat{M}_t + dL_t(1), \tag{1.8}$$

$$V_T(1) = \widetilde{H}.$$

$$dV_t(2) = \frac{(\lambda_t V_t(2) + \rho_t^2 \varphi_t(2))^2}{1-\rho_t^2 + \rho_t^2 V_t(2)} d\langle M\rangle_t + \varphi_t(2)\, d\widehat{M}_t + dL_t(2), \quad V_T(2) = 1, \tag{1.9}$$

Here $L(1)$ and $L(2)$ are $\mathcal{G}$-local martingales strongly orthogonal to $\widehat{M}$.

Note that, to construct the optimal strategy (1.6) we need to solve only equation (1.4), which is easier to solve than equation (1.9), whereas for the construction of the optimal strategy by (1.7) one should solve equation (1.9) and two linear equations (1.7) and (1.8). Besides proving the main theorem we don't need the continuity of the filtration $\mathcal{G}$ imposed in [21]. On the other hand the construction by (1.4), (1.6) does not contain the case of the full information, since in this case $\rho^2 = 1$ and the integral in (1.4) is not defined (this case can be included only by using certain limiting procedures), but the construction (1.7)- (1.9) includes this case directly.

The relations between these equations are as follows (here we assume that condition C) is satisfied):
If $(\tilde{Y}, \tilde{\psi})$ is a solution of (1.4) for H equal to strictly positive constant c, then the processes $\tilde{Y}_t$, $c - \int_0^t \pi_u^* d\widehat{S}_u$ are strictly positive and the process

$$U_t = \frac{\widetilde{Y}_t}{c - \int_0^t \pi_u^* d\widehat{S}_u},$$

where π^* is defined by (1.6), satisfies the BSDE (1.9).

On the other hand, if the triples $(V(1), \varphi(1), L(1))$ and $(V(2), \varphi(2), L(2))$ satisfy (1.8)-(1.9), then the pair $(\widetilde{Y}, \widetilde{\psi})$, where $\widetilde{Y}_t = V_t(1) - \widehat{X}_t^{\pi^*} V_t(2)$ and $\widetilde{\psi}_t = \varphi_t(1) - V_t(1)\pi_t^* - \varphi_t(2)\widehat{X}_t^{\pi^*}$ (π^* and $\widehat{X}_t^{\pi^*}$ are defined by (1.7)), is a solution of equation (1.4).

In Section 5, we consider a diffusion market model which consists of two assets S and η, where S_t is a state of a process being controlled and η_t is the observation process. Suppose that S_t and η_t are governed by

$$dS_t = \mu_t(\eta)dt + \sigma_t(\eta)dw_t^0,$$
$$d\eta_t = a_t(\eta)dt + b_t(\eta)dw_t,$$

where w^0 and w are Brownian motions with correlation ρ. In this case $\mathcal{F}_t = \mathcal{F}_t^{S,\eta}$ and the flow of observable events is $\mathcal{G}_t = \mathcal{F}_t^{\eta}$. We give in the case of markovian coefficients solution of the problem (1.2) in terms of parabolic differential equations (PDE) and an explicit solution when coefficients and the contingent claim are deterministic.

2. Main Definitions and Auxiliary Facts

Denote by $\mathcal{M}^e(\mathcal{F})$ the set of equivalent martingale measures for S, i.e., set of probability measures Q equivalent to P such that S is a $\mathcal{F}$-local martingale under Q.

Let

$$\mathcal{M}_2^e(\mathcal{F}) = \{Q \in \mathcal{M}^e(\mathcal{F}) : EZ_T^2(Q) < \infty\},$$

where $Z_t(Q)$ is the density process (with respect to the filtration $\mathcal{F}$) of Q relative to P.

Remark 2.1. If S is continuous, then the existence of an equivalent martingale measure and the Girsanov theorem imply that the structure condition (1.1) is satisfied.

Note that the density process $Z_t(Q)$ of any element Q of $\mathcal{M}^e(\mathcal{F})$ is expressed as an exponential martingale of the form

$$\mathcal{E}_t(-\lambda \cdot M + N),$$

where N is a $\mathcal{F}$-local martingale strongly orthogonal to M and $\mathcal{E}_t(X)$ is the Doleans-Dade exponential of X.

If the local martingale $Z_t^{min} = \mathcal{E}_t(-\lambda \cdot M)$ is a true martingale, $dQ^{min}/dP = Z_T^{min} dP$ defines an equivalent probability measure called the minimal martingale measure for S.

Recall that a measure Q satisfies the Reverse Hölder inequality $R_2(P)$ if there exists a constant C such that

$$E\big(\frac{Z_T^2(Q)}{Z_\tau^2(Q)}|\mathcal{F}_\tau\big) \le C, \quad P-a.s.$$

for every $\mathcal{F}$-stopping time τ.

Remark 2.2. If there exists a measure $Q \in \mathcal{M}^e(\mathcal{F})$ that satisfies the Reverse Hölder inequality $R_2(P)$, then according to Theorem 3.4 of Kazamaki [16] the martingale $M^Q = -\lambda \cdot M + N$ belongs to the class BMO and hence $-\lambda \cdot M$ also belongs to BMO, i.e.,

$$E\big(\int_\tau^T \lambda_u^2 d\langle M\rangle_u|\mathcal{F}_\tau\big) \le const \tag{2.1}$$

for every stopping time τ. Therefore, it follows from Theorem 2.3 of Kazamaki [16] that $\mathcal{E}_t(-\lambda\cdot M)$ is a true martingale. So, condition D) implies that the minimal martingale measure exists (but Z^{min} is not necessarily square integrable).

For all unexplained notations concerning the martingale theory used below we refer the reader to [6], [19], [15].

Let $\Pi(\mathcal{F})$ be the space of all $\mathcal{F}$-predictable S-integrable processes π such that the stochastic integral

$$(\pi \cdot S)_t = \int_0^t \pi_u dS_u, \;\; t \in [0,T],$$

is in the $\mathcal{S}^2$ space of semimartingales , i.e.,

$$E\bigg(\int_0^T \pi_s^2 d\langle M\rangle_s\bigg) + E\bigg(\int_0^T |\pi_s\lambda_s| d\langle M\rangle_s\bigg)^2 < \infty.$$

Denote by $\Pi(\mathcal{G})$ the subspace of $\Pi(\mathcal{F})$ of $\mathcal{G}$-predictable strategies.

Remark 2.3. Since $\lambda \cdot M \in BMO$ (see Remark 2.2), it follows from the proof of Theorem 2.5 of Kazamaki [16]

$$E\bigg(\int_0^T |\pi_u\lambda_u| d\langle M\rangle_u\bigg)^2 = E\langle|\pi|\cdot M, |\lambda|\cdot M\rangle_T^2 \le 2||\lambda\cdot M||_{\text{BMO}} E\int_0^T \pi_u^2 d\langle M\rangle_u < \infty.$$

Therefore, under condition D) the strategy π belongs to the class $\Pi(\mathcal{G})$ if and only if $E\int_0^T \pi_s^2 d\langle M\rangle_s < \infty$.

Define $J_T^2(\mathcal{F})$ and $J_T^2(\mathcal{G})$ as spaces of terminal values of stochastic integrals, i.e.,

$$J_T^2(\mathcal{F}) = \{(\pi\cdot S)_T : \pi \in \Pi(\mathcal{F})\}, \;\; J_T^2(\mathcal{G}) = \{(\pi\cdot S)_T : \pi \in \Pi(\mathcal{G})\}.$$

Let us make a comment on condition B).

Remark 2.4. Condition B) is satisfied if and only if the σ-algebras $\mathcal{F}_t$ and $\mathcal{G}_T$ are conditionally independent given $\mathcal{G}_t$ for all $t \in [0,T]$ (see Theorem 9.29 from Jacod 1978). Note that one can weaken this condition imposing that any $\mathcal{G}$-martingale is a $\mathcal{A}$- local martingale, where $\mathcal{A}$ is the augmented filtration generated by $\mathcal{F}^S$ and $\mathcal{G}$. This condition is automatically satisfied if $\mathcal{F}_t^S \subseteq \mathcal{G}_t$. In this case instead of (1.1) one should use the decomposition

$$S_t = S_0 + \int_0^t {}^p\lambda_u d\langle M\rangle_u + N_t, \tag{2.2}$$

where

$$N_t = M_t + \int_0^t [\lambda_u - {}^p\lambda_u] d\langle N\rangle_u \tag{2.3}$$

is a $\mathcal{A}$-local martingale and ${}^p\lambda$ is an $\mathcal{A}-$predictable projection of λ.

Now we recall some known assertions from the filtering theory.

Let $A = (A_t, t \in [0,T])$ be a RCLL process and there is a sequence $(\tau_n, n \geq 1)$ of $\mathcal{G}$-stopping times such that $E\int_0^{\tau_n} |dA_u| < \infty$ for all $n \geq 1$. Then there exists a unique $\mathcal{G}$-predictable process A^p of finite variation (see Jacod 1978), called a $\mathcal{G}$-dual predictable projection of A such that

$$E(A_t|\mathcal{G}_t) - A_t^p \quad \text{is a } \ \mathcal{G} - \text{local martingale.}$$

Denote by $\mathcal{M}^2_{\text{loc}}(\mathcal{G})$ the class of locally square integrable $\mathcal{G}$-martingales.

For reader's convenience, we give the proof of the following assertion, which is proved similarly to [19].

Proposition 2.1. *If conditions $A)$ and $B)$ are satisfied, then for any $m^{\mathcal{G}} \in \mathcal{M}^2_{loc}(\mathcal{G})$*

$$\widehat{M_t} = E(M_t|\mathcal{G}_t) = \int_0^t {}^p\left(\frac{d\langle M, m^{\mathcal{G}}\rangle}{d\langle m^{\mathcal{G}}\rangle}\right)_u dm_u^{\mathcal{G}} + N_t^{\mathcal{G}}, \tag{2.4}$$

where $N^{\mathcal{G}} \in \mathcal{M}^2_{loc}(\mathcal{G})$ and is strongly orthogonal to $m^{\mathcal{G}}$.

Proof. Condition A) and the continuity of M imply that $\widehat{M} \in \mathcal{M}^2_{\text{loc}}(\mathcal{G})$. Therefore $\widehat{M}$ admits the GKW decomposition

$$\widehat{M_t} = E(M_t|\mathcal{G}_t) = \int_0^t f_u dm_u^{\mathcal{G}} + N_t^{\mathcal{G}}, \tag{2.5}$$

where $f_u = \frac{d\langle \widehat{M}, m^{\mathcal{G}}\rangle_u}{d\langle m^{\mathcal{G}}\rangle_u}$. Thus, it is sufficient to show that $d\langle m^{\mathcal{G}}\rangle_t dP$ -a.e.

$$\frac{d\langle\, \widehat{M}, m^{\mathcal{G}}\rangle_u}{d\langle m^{\mathcal{G}}\rangle_u} = {}^p\left(\frac{d\langle M, m^{\mathcal{G}}\rangle}{d\langle m^{\mathcal{G}}\rangle}\right)_u. \tag{2.6}$$

By condition B) $m^{\mathcal{G}} \in \mathcal{M}^2_{\text{loc}}(\mathcal{G})$ implies $m^{\mathcal{G}} \in \mathcal{M}^2_{\text{loc}}(\mathcal{F})$ and the process $M_t m_t^{\mathcal{G}} - \langle M, m^{\mathcal{G}}\rangle_t$ is a $\mathcal{F}$-local martingale. It follows from condition A) that $M_t m_t^{\mathcal{G}}$ and $\langle M, m^{\mathcal{G}}\rangle_t$ are $\mathcal{G}$-locally integrable. Therefore the processes $E(M_t m_t^{\mathcal{G}} - \langle M, m^{\mathcal{G}}\rangle_t|\mathcal{G}_t)$ and $E(\langle M, m^{\mathcal{G}}\rangle_t|\mathcal{G}_t) - \langle M, m^{\mathcal{G}}\rangle_t^p$ are $\mathcal{G}$-local martingales and hence the process

$$E(M_t m_t^{\mathcal{G}}|\mathcal{G}_t) - \langle M, m^{\mathcal{G}}\rangle_t^p \tag{2.7}$$

is also a $\mathcal{G}$- local martingale.

On the other hand $E(M_t m_t^{\mathcal{G}}|\mathcal{G}_t) = \widehat{M}_t m_t^{\mathcal{G}}$ and the process $\widehat{M}_t m_t^{\mathcal{G}} - \langle\, \widehat{M}, m^{\mathcal{G}}\rangle_t$ is a $\mathcal{G}$-local martingale. Therefore the process

$$E(M_t m_t^{\mathcal{G}}|\mathcal{G}_t) - \langle \widehat{M}, m^{\mathcal{G}}\rangle_t$$

is also a $\mathcal{G}$- local martingale. This, together with (2.7), implies that

$$\langle\, \widehat{M}, m^{\mathcal{G}}\rangle_t = \langle M, m^{\mathcal{G}}\rangle_t^p. \tag{2.8}$$

But

$$\langle M, m^{\mathcal{G}}\rangle^p = \left(\int_0^t \frac{d\langle M, m^{\mathcal{G}}\rangle_u}{d\langle m^{\mathcal{G}}\rangle_u} d\langle m^{\mathcal{G}}\rangle_u\right)^p = \int_0^t {}^p\left(\frac{d\langle M, m^{\mathcal{G}}\rangle}{d\langle m^{\mathcal{G}}\rangle}\right)_u d\langle m^{\mathcal{G}}\rangle_u,$$

which proves equality (2.6) and (2.4) holds. □

Corollary 2.1. *For any* $\pi \in \Pi(\mathcal{G})$

$$\widehat{(\pi \cdot S)}_t = E\left(\int_0^t \pi_u dS_u | \mathcal{G}_t\right) = \int_0^t \pi_u d\widehat{S}_u. \tag{2.9}$$

Proof. It follows from Proposition 2.1 that for any $\mathcal{G}$-predictable, M-integrable process π and any $m^{\mathcal{G}} \in \mathcal{M}^2_{\text{loc}}(\mathcal{G})$ that

$$\langle \widehat{(\pi \cdot M)}, m^{\mathcal{G}}\rangle = \int_0^t {}^p\left(\frac{d\langle M, m^{\mathcal{G}}\rangle}{d\langle m^{\mathcal{G}}\rangle}\right)_u \pi_u d\langle m^{\mathcal{G}}\rangle_u = \int_0^t \pi_u d\langle \widehat{M}, m^{\mathcal{G}}\rangle_u = \langle \pi \cdot \widehat{M}, m^{\mathcal{G}}\rangle_t.$$

Hence, for any $\mathcal{G}$-predictable, M-integrable process π

$$\widehat{(\pi \cdot M)}_t = E\left(\int_0^t \pi_s dM_s | \mathcal{G}_t\right) = \int_0^t \pi_s d\widehat{M}_s. \tag{2.10}$$

Since π, λ and $\langle M\rangle$ are $\mathcal{G}$-predictable, from (2.10) we obtain (2.9).

Remark 2.5. In particular, equality (2.8) implies that

$$\langle M, \widehat{M}\ \rangle^p = \langle\ \widehat{M}\ \rangle \tag{2.11}$$

and

$$\langle M, L\rangle^p = 0 \tag{2.12}$$

if L is a $\mathcal{G}$-local martingale orthogonal to $\widehat{M}$.

Lemma 2.1. *Let conditions* A), B) *be satisfied and* $\widehat{M}_t = E(M_t|\mathcal{G}_t)$. *Then* $\langle\ \widehat{M}\ \rangle$ *is absolutely continuous w.r.t* $\langle M\rangle$ *and*

$$\rho_t^2 = \frac{d\langle\ \widehat{M}\ \rangle_t}{d\langle M\rangle_t} \le 1.$$

Moreover, if $A = \{(\omega, t) : \rho_t^2 = 1\}$ *then a.s. for all t*

$$\int_0^t I_A(u) dM_u = \int_0^t I_A(u) d\widehat{M}_u. \tag{2.13}$$

Proof. By (2.10) for any bounded $\mathcal{G}$-predictable process f

$$\begin{aligned}
E\int_0^t f_s^2 d\langle\ \widehat{M}\ \rangle_s &= E\left(\int_0^t f_s d\widehat{M}_s\right)^2 \\
&= E\left(E\left(\int_0^t f_s dM_s|\mathcal{G}_t\right)\right)^2 \le EE\left(\left(\int_0^t f_s dM_s\right)^2 |\mathcal{G}_t\right) \\
&= E\left(\int_0^t f_s dM_s\right)^2 = E\int_0^t f_s^2 d\langle M\rangle_s
\end{aligned} \tag{2.14}$$

which implies that $\langle\widehat{M}\rangle$ is absolutely continuous w.r.t $\langle M\rangle$, i.e.,

$$\langle \widehat{M} \rangle_t = \int_0^t \rho_s^2 d\langle M\rangle_s$$

for a $\mathcal{G}$-predictable process ρ. Moreover (2.14) implies that the process $\langle M\rangle - \langle \widehat{M} \rangle$ is increasing and hence $\rho^2 \le 1\ \mu^{\langle M\rangle}$ a.e.

Let us show now the equality (2.13). By definition of the set A, $\int_0^t I_A(u)d\langle M\rangle_u = \int_0^t I_A(u)d\langle\widehat{M}\rangle_u$. Since the set A is $\mathcal{G}$-predictable and $\langle M,\widehat{M}\rangle^p = \langle\widehat{M}\rangle$, by Proposition 2.2

$$E\Big(\int_0^t I_A(u)dM_u - \int_0^t I_A(u)d\widehat{M}_u\Big)^2 = E\int_0^t I_A(u)d\langle M-\widehat{M}\,\rangle_u$$

$$= E\int_0^t I_A(u)d\langle M-\widehat{M}\,\rangle_u^p = E\int_0^t I_A(u)d\langle M\rangle_u - E\int_0^t I_A(u)d\langle\,\widehat{M}\,\rangle_u = 0.$$

□

Corollary 2.2. *If $\rho_t^2 = 1$ for all t, then $M = \widehat{M}$ and therefore M is a $\mathcal{G}$-local martingale.*

We shall use the following Lemma proved in [5].

Lemma 2.2. *Let $N = (N_t, t \in [0,T])$ be a square integrable martingale such that $N_0 > 0$. Let $\tau = \inf\{t : N_t \le 0\} \wedge T$ [1]) be a predictable stopping time announced by a sequence of stopping times $(\tau_n; n \ge 1)$. Then*

$$E\left(\frac{N_T^2}{N_{\tau_n}^2}|\mathcal{G}_{\tau_n}\right) \to \infty \text{ on the set } (N_\tau = 0)$$

Proof.

$$1 = E\left(\frac{N_T}{N_{\tau_n}}|\mathcal{G}_{\tau_n}\right) = E\left(\frac{N_T}{N_{\tau_n}}I_{(N_\tau=0)}|\mathcal{G}_{\tau_n}\right)$$

$$\le E^{\frac12}\left(\frac{N_T^2}{N_{\tau_n}^2}|\mathcal{G}_{\tau_n}\right)E^{\frac12}\left(I_{(N_\tau=0)}|\mathcal{G}_{\tau_n}\right). \tag{2.15}$$

By the Levy theorem $\lim_{n\to\infty} E\left(I_{(N_\tau=0)}|\mathcal{G}_{\tau_n}\right) = I_{(N_\tau=0)}$ is equal to 0 on the set $(N_\tau = 0)$. Therefore it follows from (2.15) that $E\left(\frac{N_T^2}{N_{\tau_n}^2}|\mathcal{G}_{\tau_n}\right) \to \infty$ on $(N_\tau = 0)$. □

3. MEAN-VARIANCE HEDGING AND FORWARD-BACKWARD EQUATION

Let $X_t^{x,\pi^*} = x + \int_0^t \pi_s dS_s$ be the wealth process corresponding to the optimal strategy π^* and initial capital x. Without loss of generality we assume that $x = 0$ and denote by $X(\pi^*) \equiv X^* = X^{0,\pi^*}$. Let $H_t = E[H|\mathcal{F}_t]$, $c_H = E[H]$ and let

$$H_t = E(H|\mathcal{F}_t) = c_H + \int_0^t h_u dM_u + L_t^H \tag{3.1}$$

be the Galtchouk-Kunita-Watanabe (GKW) decomposition of H_t, where L^H is an $\mathcal{F}$-martingale orthogonal to M and h is $\mathcal{F}$-predictable M-integrable process.

[1] It is assumed that $\inf\emptyset = \infty$ and $a \wedge b$ denotes $\min\{a, b\}$

In the following proposition we don't need the continuity of the process S which we assumed throughout the paper.

Proposition 3.1. *Let S be a special semimartingale satisfying the structure condition* (1.1) *and M is a locally square integrable $\mathcal{F}$-martingale. If $\pi^* \in \Pi(\mathcal{G})$ is the optimal strategy of the problem* (1.2)*, then $d\langle M\rangle_t dP$-a.e.*

$$\pi_t^* = \frac{d\left(\int_0^t [h_u + \psi_u + \lambda_u H_u + \lambda_u Y_u - \lambda_u X_u^*] d\langle M\rangle_u\right)^p}{d\langle M\rangle_t^p}, \tag{3.2}$$

where the triple (Y, ψ, N), $\langle N, M\rangle = 0$ is a solution of BSDE

$$dY_t = \pi_t^* \lambda_t d\langle M\rangle_t + \psi_t dM_t + dN_t, \;\; Y_T = 0. \tag{3.3}$$

In particular, if $\langle M\rangle$ is $\mathcal{G}$-predictable then $d\langle M\rangle_t dP$-a.e.

$$\pi_t^* = {}^p h_t + {}^p\psi_t + {}^p\big(\lambda(H_\cdot + Y - X^*)\big)_t. \tag{3.4}$$

Proof. The variational principle gives that

$$E(H - X_T(\pi^*))X_T(\pi) = 0, \;\; \forall \pi \in \Pi(\mathcal{G}).$$

Since $\pi^* \in \Pi(\mathcal{G})$ we have that $E\big(\int_0^T \pi_u^* \lambda_u d\langle M\rangle_u\big)^2 < \infty$ and by the GKW decomposition

$$-\int_0^T \pi_u^* \lambda_u d\langle M\rangle_u = c + \int_0^T \psi_u dM_u + N_T, \quad \langle M, N\rangle = 0, \tag{3.5}$$

where $\psi \cdot M$ and N are square integrable martingales. Using the martingale property, it follows from (3.5) that the triple (Y, ψ, N), where

$$Y_t = E\bigg(\int_t^T \pi_u^* \lambda_u d\langle M\rangle_u | \mathcal{F}_t\bigg)$$

and ψ, N are defined by (3.5), satisfies the BSDE

$$Y_t = Y_0 + \int_0^t \pi_u^* \lambda_u d\langle M\rangle_u + \int_0^t \psi_u dM_u + N_t, \;\; Y_T = 0. \tag{3.6}$$

Note that $Y_0 = c = E\int_0^T \pi_u^* \lambda_u d\langle M\rangle_u$.

Therefore by (3.1) and (3.5) we have

$$\begin{aligned}
&E(H - X_T(\pi^*))X_T(\pi) \\
&\quad = E\bigg(-\int_0^T \pi_t^* \lambda_t d\langle M\rangle_t - \int_0^T \pi_t^* dM_t + H\bigg)\bigg(\int_0^T \pi_t dS_t\bigg) \\
&\quad = E\bigg(Y_0 + \int_0^T \psi_t dM_t + N_T - \int_0^T \pi_t^* dM_t + H\bigg)\bigg(\int_0^T \pi_t dS_t\bigg) \\
&\quad = E\bigg(Y_0 + N_T + \int_0^T (\psi_t - \pi_t^*)\, dM_t + H\bigg)\bigg(\int_0^T \pi_t \lambda_t d\langle M\rangle_t\bigg) \qquad (3.7) \\
&\qquad + E\bigg(Y_0 + N_T + \int_0^T (\psi_t - \pi_t^*)\, dM_t + c_H
\end{aligned}$$

$$+\int_0^T h_t dM_t + L_T^H\Big)\Big(\int_0^T \pi_t dM_t\Big) = 0. \tag{3.8}$$

Using the formula of integration by parts in (3.7) and properties of mutual characteristics of martingales in (3.8) we obtain the equality

$$E\int_0^T \left(Y_0 + N_t + \int_0^t (\psi_u - \pi_u^*)\, dM_t + H_t\right) \pi_t\lambda_t d\langle M\rangle_t$$
$$+ E\int_0^T (\psi_t + h_t - \pi_t^*)\, \pi_t d\langle M\rangle_t = 0.$$

Inserting the solution Y of BSDE (3.6) in the latter equality gives

$$E\int_0^T \left(Y_0 + H_t + Y_t - \int_0^t \lambda_u\pi_u^* d\langle M\rangle_u - \int_0^t \pi_u^* dM_t\right) \pi_t\lambda_t d\langle M\rangle_t$$
$$+E\int_0^T (\psi_t + h_t - \pi_t^*)\, \pi_t d\langle M\rangle_t$$
$$= E\int_0^T \big((H_t + Y_t - X_t^*)\lambda_t + \psi_t + h_t - \pi_t^*\big)\pi_t d\langle M\rangle_t = 0.$$

It follows from the latter equality that

$$E\int_0^T \pi_t d\Big(\int_0^t \pi_u^* d\langle M\rangle_u\Big)$$
$$= E\int_0^T \pi_t d\Big(\int_0^t [h_u + \psi_u + \lambda_u H_u + \lambda_u Y_u - \lambda_u X_u^*] d\langle M\rangle_u\Big)$$

and using the properties of $\mathcal{G}$-dual projections

$$E\int_0^T \pi_t\pi_u^* d\langle M\rangle_u^p = E\int_0^T \pi_t d\big(\int_0^t [h_u + \psi_u + \lambda_u H_u + \lambda_u Y_u - \lambda_u X_u^*] d\langle M\rangle_u\big)^p.$$

By arbitrariness of $\pi \in \Pi(\mathcal{G})$ we get

$$\int_0^t \pi_u^* d\langle M\rangle_u^p = \Big(\int_0^t [h_u + \psi_u + \lambda_u H_u + \lambda_u Y_u - \lambda_u X_u^*] d\langle M\rangle_u\Big)^p. \tag{3.9}$$

It is evident that if $A \ll B$, then $A^p \ll B^p$. Therefore, taking the Radon-Nicodym derivatives in (3.9) equality (3.2) follows. It is also evident that if $\langle M\rangle$ is $\mathcal{G}$-predictable, then

$$\frac{d(\int_0^t [h_u + \psi_u + \lambda_u H_u + \lambda_u Y_u - \lambda_u X_u^*] d\langle M\rangle_u)^p}{d\langle M\rangle_t^p} = {}^p h_t + {}^p\psi_t + {}^p\big(\lambda(H_\cdot + Y - X^*)\big)_t,$$

which gives (3.4). □

Corollary 3.1 (Schweizer [28]). *If the price process S is a locally square integrable martingale, then the optimal strategy is of the form*

$$\pi_t^* = \frac{d\big(\int_0^t h_u d\langle M\rangle_u\big)^p}{d\langle M\rangle_t^p}.$$

Proof. It follows from (3.2) and (3.3), since in this case $\lambda = 0$ and the triply ($Y = 0$, $\psi = 0$, $L = 0$) satisfies the BSDE (3.3). □

We shall use also the GKW decomposition of $H_t = E(H|\mathcal{F}_t)$ with respect to the local martingale $\widehat{M}$

$$H_t = c_H + \int_0^t h_u^{\mathcal{G}} d\widehat{M}_u + L_t^{H,\mathcal{G}}. \tag{3.10}$$

Here $h^{\mathcal{G}}$ is a $\mathcal{F}$-predictable process and $L^{H,\mathcal{G}}$ is a $\mathcal{F}$- local martingale strongly orthogonal to $\widehat{M}$.

It follows from Proposition 2.1 (applied for $m^{\mathcal{G}} = \widehat{M}$) and Lemma 2.1 that

$$\langle E(H|\mathcal{G}_\cdot), \widehat{M}\rangle_t = \int_0^t {}^p(h^{\mathcal{G}})_u d\langle\widehat{M}\rangle_u = \int_0^t {}^p(h^{\mathcal{G}})_u \rho_u^2 d\langle M\rangle_u. \tag{3.11}$$

Corollary 3.2. *Let conditions* A) *and* B) *be satisfied. Then* (3.2), (3.3) *is equivalent to the system of Forward-Backward equations*

$$d\widehat{X}_t^* = \left({}^p h_t + \frac{d\langle\widehat{M}, m\rangle_t}{d\langle M\rangle_t} + \lambda_t(\widehat{H}_{t-} + \widehat{Y}_{t-} - \widehat{X}_{t-}^*)\right) d\widehat{S}_t, \quad \widehat{X}_0^* = 0, \tag{3.12}$$

$$d\widehat{Y}_t = \lambda_t \left({}^p h_t + \frac{d\langle\widehat{M}, m\rangle_t}{d\langle M\rangle_t} + \lambda_t(\widehat{H}_{t-} + \widehat{Y}_{t-} - \widehat{X}_{t-}^*)\right) d\langle M\rangle_t + dm_t, \;\; \widehat{Y}_T = 0. \tag{3.13}$$

Proof. Since $\lambda, \pi^*, \langle M\rangle$ are $\mathcal{G}$-adapted,

$${}^p(\lambda H_\cdot + \lambda Y - \lambda X^*)_t = \lambda_t(\widehat{H}_{t-} + \widehat{Y}_{t-} - \widehat{X}_{t-}^*)$$

and

$$m_t \equiv \int_0^t \psi_s dM_s + N_t$$

is a $\mathcal{G}$-martingale (this follows from (3.6) and condition B), since under this condition Y_t is also equal to $E(\int_t^T \pi_u^* \lambda_u d\langle M\rangle_u|\mathcal{G}_t)$). Therefore, by (2.6)

$$\widehat{\psi}_t = {}^p\left(\frac{d\langle M, m\rangle}{d\langle M\rangle}\right)_t = \frac{d\langle\widehat{M}, m\rangle_t}{d\langle M\rangle_t}$$

and it follows from (3.4) and (3.3) that the optimal strategy π^* satisfies the system

$$\pi_t^* = {}^p h_t + \frac{d\langle\widehat{M}, m\rangle_t}{d\langle M\rangle_t} + \lambda_t(\widehat{H}_{t-} + \widehat{Y}_{t-} - \widehat{X}_{t-}^*), \tag{3.14}$$

$$d\widehat{Y}_t = \pi_t^* \lambda_t d\langle M\rangle_t + dm_t, \;\; \widehat{Y}_T = 0. \tag{3.15}$$

If we insert the expression (3.14) for π^* in (3.15) and then integrate both parts of equation (3.14) with respect to $\widehat{S}$, we obtain the system of Forward–Backward equations (3.12), (3.13). $\square$

Remark 3.1. If we use the GKW decomposition of m with respect to $\widehat{M}$ and filtration $\mathcal{G}$

$$m_t = \int_0^t \widetilde{\psi}_u d\widehat{M}_u + \widetilde{L}_t, \quad \langle\widehat{M}, \widetilde{L}\rangle = 0,$$

then by (2.6) $\widehat{\psi}_t = \rho_t^2 \widetilde{\psi}_t$ and one can write the Forward-Backward equations (3.12), (3.13) in the form

$$d\widehat{X}_t^* = \left({}^p h_t + \rho_t^2 \widetilde{\psi}_t + \lambda_t(\widehat{H}_{t-} + \widehat{Y}_{t-} - \widehat{X}_{t-}^*)\right) d\widehat{S}_t, \quad \widehat{X}_0^* = 0$$

$$d\widehat{Y}_t = \lambda_t \left({}^p h_t + \rho_t^2 \widetilde{\psi}_t + \lambda_t(\widehat{H}_{t-} + \widehat{Y}_{t-} - \widehat{X}_{t-}^*)\right) d\langle M\rangle_t + \widetilde{\psi}_t d\widehat{M}_t + d\widetilde{L}_t, \quad \widehat{Y}_T = 0.$$

From now on we assume **E**) $\rho_t^2 < 1$ for all $t \in [0, T]$.

Let us introduce the operator AY defined for any $Y \in \mathcal{M}^2(\mathcal{G}, P)$ by

$$(AY)_t = E\left(\int_0^T \frac{1}{1-\rho_u^2}\left[\lambda_u Y_{u-} + \rho_u^2 \psi_u\right]\left(\lambda_u d\langle M\rangle_u + d\widehat{M}_u\right)\Big|\mathcal{G}_t\right).$$

We shall use the following notations;

$$\widetilde{h}_t = {}^p h_t - {}^p h_t^{\mathcal{G}} \rho_t^2, \quad \widetilde{H} = \widehat{H}_T - \int_0^T \frac{\widetilde{h}_t}{1-\rho_t^2} d\widehat{S}_t. \tag{3.16}$$

Let us consider equation

$$\widetilde{Y}_T = \widetilde{H} - \int_0^T \frac{1}{1-\rho_t^2}\left[\lambda_t \widetilde{Y}_{t-} + \rho_t^2 \widetilde{\psi}_t\right]\left(\lambda_t d\langle M\rangle_t + d\widehat{M}_t\right), \tag{3.17}$$

which can be written in the form $\widetilde{Y}_T = \widetilde{H} - (A\widetilde{Y})_T$.

Theorem 3.1. *Let conditions* A), B) *and* E) *be satisfied and let* $E\tilde{H}^2 < \infty$. *Then the equation* (3.17) *admits a unique solution* $\widetilde{Y} \in \mathcal{M}^2(\mathcal{G}, P)$ *satisfying* $E|\widetilde{Y}_T|^2 \le E|\widetilde{H}|^2$.

Proof. We need only to show that A is a non-negative operator. Indeed, for $Y_t = c + \int_0^t \varphi_s d\widehat{M}_s + L_t,\ \langle \widehat{M}, L\rangle = 0$ we have

$$\begin{aligned}(Y, AY) = E\Big(&Y_T \int_0^T \frac{1}{1-\rho_t^2} Y_{t-}\lambda_t^2 d\langle M\rangle_t + Y_T \int_0^T \frac{1}{1-\rho_t^2} Y_{t-}\lambda_t d\widehat{M}_t \\ &+ Y_T \int_0^T \frac{\rho_t^2}{1-\rho_t^2}\varphi_t \lambda_t d\langle M\rangle_t + Y_T \int_0^T \frac{\rho_t^2}{1-\rho_t^2}\varphi_t d\widehat{M}_t\Big).\end{aligned}$$

Since $\langle Y, \widehat{M}\rangle_t = \int_0^t \varphi_u \rho_u^2 d\langle M\rangle_u$ and $EY_T \int_0^T g_u d\langle M\rangle_u = E\int_0^T Y_u g_u d\langle M\rangle_u$ for any $\mathcal{G}$-predictable process g, we obtain that

$$\begin{aligned}(Y, AY) &= E\Big(\int_0^T \frac{1}{1-\rho_t^2} Y_{t-}^2 \lambda_t^2 d\langle M\rangle_t + \int_0^T \frac{1}{1-\rho_t^2} Y_{t-}\lambda_t \varphi_t d\langle \widehat{M}\rangle_t \\ &\quad + \int_0^T \frac{\rho_t^2}{1-\rho_t^2} Y_{t-}\varphi_t \lambda_t d\langle M\rangle_t + \int_0^T \frac{\rho_t^2}{1-\rho_t^2}\varphi_t^2 d\langle \widehat{M}\rangle_t\Big) \\ &= E\Big(\int_0^T \frac{1}{1-\rho_t^2} Y_{t-}^2 \lambda_t^2 d\langle M\rangle_t + \int_0^T \frac{\rho_t^2}{1-\rho_t^2} Y_{t-}\lambda_t \varphi_t d\langle M\rangle_t \\ &\quad + \int_0^T \frac{\rho_t^2}{1-\rho_t^2} Y_{t-}\varphi_t \lambda_t d\langle M\rangle_t + \int_0^T \frac{\rho_t^4}{1-\rho_t^2}\varphi_t^2 d\langle M\rangle_t\Big) \\ &= E\int_0^T \frac{1}{1-\rho_t^2}\left(Y_{t-}\lambda_t + \rho_t^2 \varphi_t\right)^2 d\langle M\rangle_t \ge 0.\end{aligned}$$

Thus $Y + AY$ is a strictly positive operator, $(Id + A)^{-1}$ is bounded with the norm less than one and $Y = (Id + A)^{-1}\widetilde{H}$ is a unique solution of (3.17). □

Remark 3.2. Condition $E\tilde{H}^2 < \infty$ is satisfied if $EH^2 < \infty$ and $\rho_t^2 \le 1-\varepsilon$ for all $t \in [0, T]$, for some $\varepsilon > 0$.

Remark 3.3. If $(\tilde{Y}, \tilde{\psi})$ is a solution of equation (3.17), then it follows from the proof of Theorem 3.1 that

$$E\int_0^T \frac{1}{1-\rho_t^2}\left(\widetilde{Y}_{t-}\lambda_t + \rho_t^2\widetilde{\psi}_t\right)^2 d\langle M\rangle_t = (\widetilde{Y}, A\widetilde{Y}) = E\widetilde{Y}_T\widetilde{H} - E\widetilde{Y}_T^2 < \infty. \tag{3.18}$$

Theorem 3.2. *Let conditions* A), B), D) *and* E) *be satisfied,* $E\tilde{H}^2 < \infty$ *and* $E\int_0^T \tilde{h}_u^2(1-\rho_u^2)^{-1}d\langle M\rangle_u < \infty$. *Then the strategy* π^* *is optimal if and only if it admits the representation*

$$\pi_t^* = \frac{1}{1-\rho_t^2}\left(\widetilde{h}_t + \lambda_t\widetilde{Y}_{t-} + \rho_t^2\widetilde{\psi}_t\right), \tag{3.19}$$

where the pair $(\tilde{Y}, \tilde{\psi})$ *satisfies equation* (3.17).

Proof. Let us show that if the strategy π^* is optimal, then it is of the form (3.19). By Proposition 2.1 $\widehat{X}_t(\pi^*) = \int_0^t \pi_s^*\lambda_s d\langle M\rangle_s + \int_0^t \pi_s^* d\widehat{M}_s$. Introducing notations

$$\widetilde{Y}_t = \widehat{Y}_t + \widehat{H}_t - \widehat{X}_t(\pi^*), \quad \widetilde{m}_t = m_t + \widehat{H}_t - \int_0^t \pi_s^* d\widehat{M}_s \tag{3.20}$$

(note that $\widetilde{Y} = \widetilde{m}$ by (3.15)) we have

$$\pi_t^* = {}^p h_t + \frac{d\langle\widehat{M}, \widetilde{m}\rangle_t}{d\langle M\rangle_t} + \pi_t^*\frac{d\langle\widehat{M}\rangle_t}{d\langle M\rangle_t} - \frac{d\langle\widehat{M}, \widehat{H}\rangle_t}{d\langle M\rangle_t} + \lambda_t\widetilde{Y}_{t-},$$
$$d\widetilde{Y}_t = d\widetilde{m}_t, \quad \widetilde{Y}_T = \widehat{H}_T - \widehat{X}_T(\pi^*),$$

which gives (since $\rho_t^2 < 1$ for all t)

$$\pi_t^* = \frac{1}{1-\rho_t^2}\left[{}^p h_t - \frac{d\langle\widehat{M}, \widehat{H}\rangle_t}{d\langle M\rangle_t} + \frac{d\langle\widehat{M}, \widetilde{Y}\rangle_t}{d\langle M\rangle_t} + \lambda_t\widetilde{Y}_{t-}\right], \tag{3.21}$$

$$\widetilde{Y}_T = \widehat{H}_T - \widehat{X}_T(\pi^*). \tag{3.22}$$

Integrating (3.21) with respect to $\widehat{S}$

$$\widehat{X}_T(\pi^*) = \int_0^T \frac{1}{1-\rho_t^2}\left[{}^p h_t - \frac{d\langle\widehat{M}, \widehat{H}\rangle_t}{d\langle M\rangle_t} + \frac{d\langle\widehat{M}, \widetilde{Y}\rangle_t}{d\langle M\rangle_t} + \lambda_t\widetilde{Y}_{t-}\right] d\widehat{S}_t.$$

Since $\frac{d\langle\widehat{M}, \widehat{H}\rangle_t}{d\langle M\rangle_t} = {}^p h_t^{\mathcal{G}}\rho_t^2$, inserting the latter equality into (3.22) and taking in mind (3.16), we obtain the equation for the martingale $\widetilde{Y}$

$$\widetilde{Y}_T = \widetilde{H} - \int_0^T \frac{1}{1-\rho_t^2}\left[\frac{d\langle\widehat{M}, \widetilde{Y}\rangle_t}{d\langle M\rangle_t} + \lambda_t\widetilde{Y}_{t-}\right]\left(\lambda_t d\langle M\rangle_t + d\widehat{M}_t\right). \tag{3.23}$$

We remark that if $\widetilde{Y}_t = \widetilde{Y}_0 + \int_0^t \widetilde{\psi}_s d\widehat{M}_s + \widetilde{L}_t$ is the GKW decomposition of $\widetilde{Y}$ then (3.23) can be rewritten as (3.17). Thus, if the optimal strategy π^* exists, then the pair $(\widetilde{Y}, \widetilde{\psi})$, where $\widetilde{Y}$ is defined by (3.20), satisfies equation (3.17) and by (3.21) π^* is of the form (3.19).

Let us show now that if the strategy π^* is of the form (3.19), then it is optimal. Let first verify that $\pi^* \in \Pi(\mathcal{G})$. Since by (3.17) and (3.16)

$$\int_0^T \pi_u^* d\widehat{S}_u = \int_0^T \frac{1}{1-\rho_t^2}\Big[\lambda_t \widetilde{Y}_{t-} + \rho_t^2 \widetilde{\psi}_t + \tilde{h}_t\Big]\Big(\lambda_t d\langle M\rangle_t + d\widehat{M}_t\Big) = \widehat{H}_T - \widetilde{Y}_T,$$

it follows from Theorem 3.1 that $E\big(\int_0^T \pi_u^* d\widehat{S}_u\big)^2 < \infty$. Therefore,

$$E\bigg(\int_0^T \pi_u^* dS_u\bigg)^2 = E\bigg(\int_0^T \pi_u^* d\widehat{S}_u + \int_0^T \pi_u^* d(\widehat{M}_u - M_u)\bigg)^2$$

$$\le 2E\bigg(\int_0^T \pi_u^* d\widehat{S}_u\bigg)^2 + 2E\int_0^T (\pi_u^*)^2 d\langle \widehat{M} - M\rangle_u < \infty,$$

since it follows from (2.11) and (3.18) that

$$E\int_0^T (\pi_u^*)^2 d\langle \widehat{M} - M\rangle_u = E\int_0^T (\pi_u^*)^2 d\langle \widehat{M} - M\rangle_u^p$$

$$= E\int_0^T (\pi_u^*)^2 (1-\rho_u^2) d\langle M\rangle_u \le 2E\int_0^T \frac{\tilde{h}_u^2}{1-\rho_u^2} d\langle M\rangle_u$$

$$+2E\int_0^T \frac{1}{1-\rho_u^2}(\lambda_u \widetilde{Y}_{u-} + \rho_u^2 \widetilde{\psi}_u)^2 d\langle M\rangle_u < \infty.$$

Thus $E\big(\int_0^T \pi_u^* dS_u\big)^2 < \infty$ and by condition D) and Theorem 4.9 from [2] (see also [4])

$$E\int_0^T (\pi_u^*)^2 \langle M\rangle_u \le const\; E\bigg(\int_0^T \pi_u^* dS_u\bigg)^2 < \infty \tag{3.24}$$

and $\pi^* \in \Pi(\mathcal{G})$ by Remark 2.3.

By the variational principle it is sufficient to show that

$$E\bigg(H - \int_0^T \pi_u^* dS_u\bigg)\bigg(\int_0^T \pi_u dS_u\bigg) = 0, \quad \forall \pi \in \Pi(\mathcal{G}). \tag{3.25}$$

From equation (3.22) we have that

$$-\int_0^T \pi_u^* \lambda_u d\langle M\rangle_u = \widetilde{Y}_T - \widehat{H}_T + \int_0^T \pi_u^* d\widehat{M}_u.$$

Therefore

$$E\bigg(H - \int_0^T \pi_u^* dS_u\bigg)\bigg(\int_0^T \pi_u dS_u\bigg)$$

$$= E\bigg(\widetilde{Y}_T + H - \widehat{H}_T + \int_0^T \pi_u^* d(\widehat{M}_u - M_u)\bigg)\bigg(\int_0^T \pi_u \lambda_u d\langle M\rangle_u + \int_0^T \pi_u dM_u\bigg).$$

Since λ and $\langle M\rangle$ are $\mathcal{G}$-adapted

$$E\left(H-\widehat{H}_T\right)\left(\int_0^T \pi_u\lambda_u d\langle M\rangle_u\right)=0$$

and by Proposition 2.1

$$E\int_0^T \pi_u^* d(\widehat{M}_u-M_u)\bigg(\int_0^T \pi_u\lambda_u d\langle M\rangle_u\bigg)$$

$$=E\int_0^T \pi_u\lambda_u d\langle M\rangle_u E\bigg(\int_0^T \pi_u^* d(\widehat{M}_u-M_u)|\mathcal{G}_T\bigg)=0.$$

Since $\widetilde{Y}$ is a martingale

$$E\widetilde{Y}_T\bigg(\int_0^T \pi_u\lambda_u d\langle M\rangle_u\bigg)=E\int_0^T \pi_u\lambda_u\widetilde{Y}_u d\langle M\rangle_u. \tag{3.26}$$

Using the GKW decomposition for $\widetilde{Y}_t$ and relations (2.11), (2.12)

$$E\widetilde{Y}_T\int_0^T \pi_u dM_u=E\int_0^T \pi_u\widetilde{\psi}_u d\langle M,\widehat{M}\rangle_u=E\int_0^T \pi_u\widetilde{\psi}_u\rho_u^2 d\langle M\rangle_u. \tag{3.27}$$

Using decompositions (3.1), (3.10) for H, projection theorem and again relations (2.11), (2.12)

$$\begin{aligned}E(H-\widehat{H}_T)\bigg(\int_0^T \pi_u dM_u\bigg)&=E\int_0^T \pi_u h_u d\langle M\rangle_u-E\int_0^T \pi_u h_u^{\mathcal{G}} d\langle M,\widehat{M}\rangle_u\\&=E\int_0^T \pi_u\widetilde{h}_u d\langle M\rangle_u.\end{aligned} \tag{3.28}$$

Taking the sum of right-hand sides of (3.26), (3.27) and (3.28) we obtain

$$\begin{aligned}&E\int_0^T \pi_u\lambda_u\widetilde{Y}_u d\langle M\rangle_u+E\int_0^T \pi_u\widetilde{\psi}_u\rho_u^2 d\langle M\rangle_u+E\int_0^T \pi_u\widetilde{h}_u d\langle M\rangle_u\\&=E\int_0^T \pi_u(\widetilde{h}_u+\lambda_u\tilde{Y}_u+\widetilde{\psi}_u\rho_u^2)d\langle M\rangle_u\\&=E\int_0^T \pi_u\pi_u^*(1-\rho_u^2)d\langle M\rangle_u,\end{aligned} \tag{3.29}$$

since π^* is of the form (3.19). Finally,

$$\begin{aligned}E\int_0^T \pi_u^* d(\widehat{M}_u-M_u)\int_0^T \pi_u dM_u&=E\int_0^T \pi_u^*\pi_u d\langle\widehat{M},M\rangle_u-E\int_0^T \pi_u^*\pi_u d\langle M\rangle_u\\&=-E\int_0^T \pi_u\pi_u^*(1-\rho_u^2)d\langle M\rangle_u,\end{aligned}$$

which, together with (3.29), implies that (3.25) is fulfilled and hence π^* is optimal. □

Remark 3.4. Theorem 3.2 remains true if instead of condition D) we assume that: $\lambda \cdot M \in$ BMO and $\rho_t^2 \le 1-\varepsilon$ for all $t \in [0,T]$, for some $\varepsilon > 0$. Indeed, in the proof of Theorem 3.2 condition D) is used only to show that $E\int_0^T (\pi_u^*)^2 \langle M\rangle_u < \infty$. But if $\rho_t^2 \le 1-\varepsilon$, then

$$\begin{aligned}
E\int_0^T (\pi_u^*)^2 \langle M\rangle_u &= E\int_0^T \frac{1}{(1-\rho_t^2)^2}\left(\widetilde{h}_t + \widetilde{Y}_t\lambda_t + \rho_t^2\widetilde{\varphi}_t\right)^2 d\langle M\rangle_t \\
&\le \frac{1}{\varepsilon} E\int_0^T \frac{1}{1-\rho_t^2}\left(\widetilde{h}_t + \widetilde{Y}_t\lambda_t + \rho_t^2\widetilde{\varphi}_t\right)^2 d\langle M\rangle_t < \infty
\end{aligned}$$

according to Remarks 3.2 and 3.3.

4. Relations to BSDEs for the Value Process

In this section we express the solution of equation (3.23) in terms of the value process of the problem (1.2) and show that equation (3.23) is equivalent to the BSDE derived in [21].

To this end we consider equation

$$\widetilde{Y}_T = \zeta - \int_\tau^T \frac{1}{1-\rho_t^2}\left(\lambda_t\widetilde{Y}_t + \rho_t^2\widetilde{\psi}_t\right) d\widehat{S}_t \tag{4.1}$$

for any stopping time $\tau \le T$. Similarly to Theorem 3.1 one can show that if $E\zeta^2 < \infty$, then there exists a unique solution $(\widetilde{Y}, \widetilde{\psi})$ of (4.1), where $\widetilde{Y}$ is a square integrable martingale. Throughout this section we assume that conditions A)-E) are fulfilled. Thus, unlike to previous sections we also assume the continuity of the filtration G.

Lemma 4.1. *Let $(\widetilde{Y}^\tau, \widetilde{\psi}^\tau)$ and $(\widetilde{Y}, \widetilde{\psi})$ be solutions of equations*

$$\widetilde{Y}_T = c - \int_0^T \frac{1}{1-\rho_t^2}\left(\lambda_t\widetilde{Y}_t + \rho_t^2\widetilde{\psi}_t\right) d\widehat{S}_t \tag{4.2}$$

and

$$\widetilde{Y}_T^\tau = 1 - \int_\tau^T \frac{1}{1-\rho_u^2}(\lambda_u\widetilde{Y}_u^\tau + \rho_u^2\widetilde{\psi}_u^\tau) d\widehat{S}_u, \tag{4.3}$$

respectively. Let

$$\widetilde{\pi}_u = \frac{1}{1-\rho_u^2}(\lambda_u\widetilde{Y}_u + \rho_u^2\widetilde{\psi}_u), \quad \widetilde{\pi}_u^\tau = \frac{1}{1-\rho_u^2}(\lambda_u\widetilde{Y}_u^\tau + \rho_u^2\widetilde{\psi}_u^\tau).$$

Then

$$\begin{aligned}
&\widetilde{Y}_t = \widetilde{Y}_t^\tau(c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u),\ \widetilde{\psi}_t = \widetilde{\psi}_t^\tau(c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u), \\
&\widetilde{\pi}_t = \widetilde{\pi}_t^\tau(c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u), \quad t \ge \tau.
\end{aligned} \tag{4.4}$$

Proof. Multiplying both parts of equation (4.3) by $c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u$ we get

$$\begin{aligned}
\widetilde{Y}_T^\tau(c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u) &= c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u \\
&\quad - \left(c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u\right)\int_\tau^T \frac{1}{1-\rho_u^2}(\lambda_u\widetilde{Y}_u^\tau + \rho_u^2\widetilde{\psi}_u^\tau) d\widehat{S}_u.
\end{aligned}$$

Since $c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u$ is $\mathcal{G}_\tau$-measurable, using properties of stochastic integrals we have

$$\widetilde{Y}_T^\tau(c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u) = c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u$$
$$- \int_\tau^T \frac{1}{1-\rho_u^2} \left(\lambda_u \widetilde{Y}_u^\tau (c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u) + \rho_u^2 \widetilde{\psi}_u^\tau (c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u) \right) d\widehat{S}_u.$$

On the other hand,

$$\widetilde{Y}_T = c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u - \int_\tau^T \frac{1}{1-\rho_t^2} \left(\lambda_t \widetilde{Y}_t + \rho_t^2 \widetilde{\psi}_t \right) d\widehat{S}_t$$

and relations (4.4) follow from the uniqueness of a solution of equation (4.1) with $\zeta = c - \int_0^\tau \widetilde{\pi}_u d\widehat{S}_u$. □

Let us define the process

$$\widetilde{V}_t = E\Big[\Big(1 - \int_t^T \frac{1}{1-\rho_u^2} \left(\lambda_u \widetilde{Y}_u^t + \rho_u^2 \widetilde{\psi}_u^t\right) d\widehat{S}_u\Big)^2$$
$$+ \int_t^T \frac{1}{1-\rho_u^2} (\lambda_u \widetilde{Y}_u^t + \rho_u^2 \widetilde{\psi}_u^t)^2 \, d\langle M\rangle_u | \mathcal{G}_t\Big].$$

Lemma 4.2. *$\widetilde{V}_t > 0, a.s.$ for all $t \in [0,T]$ and the process*

$$\widetilde{V}_t(c - \int_0^t \widetilde{\pi}_u dS_u)^2 + \int_0^t \widetilde{\pi}_u^2 (1-\rho_u^2) d\langle M\rangle_u$$

is a martingale.

Proof. It is evident that $\widetilde{V}_t$ is non-negative. Let us show that it is strictly positive. Assume that there exist $t \in [0,T], B \in \mathcal{G}_t$ such that $P(B) > 0$ and

$$E\Big[\Big(1 - \int_t^T \frac{1}{1-\rho_u^2} \left(\lambda_u \widetilde{Y}_u^t + \rho_u^2 \widetilde{\psi}_u^t\right) d\widehat{S}_u\Big)^2$$
$$+ \int_t^T \frac{1}{1-\rho_u^2} (\lambda_u \widetilde{Y}_u^t + \rho_u^2 \widetilde{\psi}_u^t)^2 \, d\langle M\rangle_u | \mathcal{G}_t\Big] I_B = 0.$$

This implies that

$$I_B - \int_t^T I_B \widetilde{\pi}_u^t d\widehat{S}_u = 0, \tag{4.5}$$

$$\int_t^T I_B \widetilde{\pi}_u^{t2} (1-\rho_u^2) d\langle M\rangle_u = 0. \tag{4.6}$$

Since $\rho_u < 1$, it follows from (4.6) that $\int_t^T I_B \widetilde{\pi}_u^t d\widehat{S}_u = 0$. Therefore, from (4.5) we obtain $I_B = 0$ a.s., which gives a contradiction. Thus $P(B) = 0$ and $\widetilde{V}$ is strictly positive. Let us check now the martingale property. Using elementary properties of conditional expectations and stochastic integrals it follows from Lemma 4.1 that

$$\widetilde{V}_t(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u)^2 = E\Big[\Big(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u - (c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u)\int_t^T \widetilde{\pi}^t_u\, d\widehat{S}_u\Big)^2$$
$$+ (c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u)^2 \int_t^T (1-\rho_u^2)|\widetilde{\pi}^t_u|^2\, d\langle M\rangle_u|\mathcal{G}_t\Big]$$
$$= E\Big[\Big(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u - \int_t^T (c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u)\widetilde{\pi}^t_u\, d\widehat{S}_u\Big)^2$$
$$+ \int_t^T (1-\rho_u^2)|(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u)\widetilde{\pi}^t_u|^2\, d\langle M\rangle_u|\mathcal{G}_t\Big]$$
$$= E\Big[\Big(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u - \int_t^T \widetilde{\pi}_u d\widehat{S}_u\Big)^2 + \int_t^T (1-\rho_u^2)\widetilde{\pi}^2_u d\langle M\rangle_u|\mathcal{G}_t\Big].$$

Therefore, for any $t\in[0,T]$

$$\widetilde{V}_t(c-\int_0^t \widetilde{\pi}_u dS_u)^2 + \int_0^t \widetilde{\pi}^2_u(1-\rho_u^2)d\langle M\rangle_u$$
$$= E\Big[\Big(c-\int_0^T \widetilde{\pi}_u d\widehat{S}_u\Big)^2 + \int_0^T (1-\rho_u^2)\widetilde{\pi}^2_u d\langle M\rangle_u|\mathcal{G}_t\Big],$$

which proves that this process is a martingale. □

Proposition 4.1. *The solution of* (4.2) *is strictly positive, i.e.,* $\widetilde{Y}_t>0$ *a.s. for all* $t\in[0,T]$.

Proof. Let first show that $E\widetilde{Y}_T>0$. Multiplying both parts of equation (4.2) by $\widetilde{Y}_T$ and taking expectations (as in the proof of Theorem 3.1) we obtain that

$$E\widetilde{Y}^2_T = cE\widetilde{Y}_T - \int_0^T \frac{1}{1-\rho_u^2}(\widetilde{Y}_u\lambda_u+\rho_u^2\widetilde{\psi}_u)^2 d\langle M\rangle_u.$$

Therefore $cE\widetilde{Y}_T \ge E\widetilde{Y}^2_T>0$, hence $E\widetilde{Y}_T>0$.

Let us consider the process

$$Z_t = \widetilde{Y}_t(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u) + \int_0^t \widetilde{\pi}^2_u(1-\rho_u^2)d\langle M\rangle_u. \tag{4.7}$$

It follows from the Ito formula that Z is a martingale and using the martingale property from (4.2) we have

$$\widetilde{Y}_t\Big(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u\Big) = E\Big(\widetilde{Y}^2_T + \int_t^T \widetilde{\pi}^2_u(1-\rho_u^2)d\langle M\rangle_u|\mathcal{G}_t\Big). \tag{4.8}$$

Besides the process $\widetilde{Z}_t = \widetilde{Y}_t(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u)$ is a supermartingale and

$$\widetilde{Y}_t\Big(c-\int_0^t \widetilde{\pi}_u d\widehat{S}_u\Big) \ge E(\widetilde{Y}^2_T|\mathcal{G}_t). \tag{4.9}$$

Let us define $\tau=\inf\{t:\widetilde{Y}_t=0\}\wedge T$. Then τ is a predictable stopping time and there exists a sequence of stopping times $(\tau_n; n\ge 1)$ such that $\lim\tau_n=\tau$ and $\tau_n<\tau$ for every n on $\tau>0$. Note that $\widetilde{Y}_{\tau_n}>0$ by definition of τ_n, since $\widetilde{Y}_0=E\widetilde{Y}_T>0$.

Taking τ_n instead of t in (4.9) and dividing both parts of this inequality by $\widetilde{Y}_{\tau_n}$, we obtain

$$E\bigg(\frac{\widetilde{Y}_T^2}{\widetilde{Y}_{\tau_n}^2}|\mathcal{G}_{\tau_n}\bigg) \leq \frac{c - \int_0^{\tau_n} \widetilde{\pi}_u d\widehat{S}_u}{\widetilde{Y}_{\tau_n}}. \tag{4.10}$$

It follows from the Lemma 2.2 (applied for the martingale $\widetilde{Y}_t = E(\widetilde{Y}_T|\mathcal{G}_t)$) that

$$E\bigg(\frac{\widetilde{Y}_T^2}{\widetilde{Y}_{\tau_n}^2}|\mathcal{G}_{\tau_n}\bigg) \to \infty \text{ on the set } \{\widetilde{Y}_\tau = 0\}. \tag{4.11}$$

By Lemma 4.2 and (4.8) the processes $\widetilde{V}_t(c - \int_0^t \widetilde{\pi}_u d\widehat{S}_u)^2 + \int_0^t \widetilde{\pi}_u^2(1-\rho_u^2)d\langle M\rangle_u$ and $\widetilde{Y}_t(c - \int_0^t \widetilde{\pi}_u d\widehat{S}_u) + \int_0^t \widetilde{\pi}_u^2(1-\rho_u^2)d\langle M\rangle_u$ are martingales and their values at time T coincide, hence they are indistinguishable. Thus

$$\widetilde{V}_t\bigg(c - \int_0^t \widetilde{\pi}_u d\widehat{S}_u\bigg)^2 = \widetilde{Y}_t\bigg(c - \int_0^t \widetilde{\pi}_u d\widehat{S}_u\bigg) \tag{4.12}$$

which, together with (4.10), implies that

$$E\bigg(\frac{\widetilde{Y}_T^2}{\widetilde{Y}_{\tau_n}^2}|\mathcal{G}_{\tau_n}\bigg) \leq \frac{c - \int_0^{\tau_n} \widetilde{\pi}_u d\widehat{S}_u}{\widetilde{Y}_{\tau_n}} = \frac{1}{\widetilde{V}_{\tau_n}}.$$

Since $\widetilde{V}_t > 0$, it follows from the latter inequality that

$$\lim_{n\to\infty} E\bigg(\frac{\widetilde{Y}_T^2}{\widetilde{Y}_{\tau_n}^2}|\mathcal{G}_{\tau_n}\bigg) < \infty \text{ on the set } \{\widetilde{Y}_\tau = 0\},$$

which contradicts to (4.11). Therefore $P(\widetilde{Y}_\tau = 0) = 0$ and hence $\widetilde{Y}_t > 0$ for all $t \in [0, T]$. □

Corollary 4.1. *For all* $t \in [0, T]$

$$c - \int_0^t \widetilde{\pi}_u d\widehat{S}_u \geq \widetilde{Y}_t \tag{4.13}$$

and

$$\widetilde{V}_t = \frac{\widetilde{Y}_t}{c - \int_0^t \widetilde{\pi}_u d\widehat{S}_u}. \tag{4.14}$$

Proof. By (4.9) and the Jensen inequality

$$\widetilde{Y}_t(c - \int_0^t \widetilde{\pi}_u d\widehat{S}_u) \geq E(\widetilde{Y}_T^2|\mathcal{G}_t) \geq \widetilde{Y}_t^2 \tag{4.15}$$

and since $\widetilde{Y}_t > 0$, we obtain inequality (4.13). Therefore the process $c - \int_0^t \widetilde{\pi}_u d\widehat{S}_u$ is also strictly positive and equality (4.14) follows from (4.12). □

Remark 4.1. $\widetilde{V}_t$ coincides with the value process V_t of optimization problem

$$\min_{\pi\in\Pi(\mathcal{G})} E\bigg(1 - \int_0^T \pi_u dS_u\bigg)^2$$

defined by

$$V_t = \operatorname*{ess\,inf}_{\pi \in \Pi(\mathcal{G})} E\Big(\Big(1 - \int_t^T \pi_u dS_u\Big)^2 |\mathcal{G}_t\Big).$$

This follows from Theorem 3.2 and from Theorem 3.1 of [21]. But we shall show this equality, proving that $\widetilde{V}$ satisfies the BSDE for the value process V, derived in [21].

Proposition 4.2. *Let* $(\widetilde{Y}_t, \widetilde{\psi}_t)$ *satisfies the equation*

$$\widetilde{Y}_T = c - \int_0^T \frac{1}{1-\rho_t^2}\,(\lambda_t \widetilde{Y}_t + \rho_t^2 \widetilde{\psi}_t)\, d\widehat{S}_t \tag{4.16}$$

and $\pi_t^* = \dfrac{1}{1-\rho_t^2}\,(\lambda_t \widetilde{Y}_t + \rho_t^2 \widetilde{\psi}_t)$. *Then* $c - (\pi^* \cdot \widehat{S})_t \equiv c - \widehat{X}_t^{\pi^*}$ *is strictly positive and*

$$U_t = \frac{\widetilde{Y}_t}{c - \widehat{X}_t^{\pi^*}} \tag{4.17}$$

is a solution of BSDE

$$dU_t = \frac{(\lambda_t U_t + \rho_t^2 \psi_t^U)^2}{1 - \rho_t^2 + \rho_t^2 U_t}\, d\langle M\rangle_t + \psi_t^U\, d\widehat{M}_t + dL_t^U, \quad U_T = 1. \tag{4.18}$$

Proof. By Corollary 4.1 and Lemma 4.1 $c - \widehat{X}_t^{\pi^*} > 0$ P-a.s. for all t. Therefore U_t is a $\mathcal{G}$-semimartingale. This semimartingale admits the decomposition

$$U_t = A_t + \int_0^t \psi_s^U d\widehat{M}_s + L_t^U,$$

where A_t is $\mathcal{G}$-predictable process of finite variation and L^U is a $\mathcal{G}$-local martingale strongly orthogonal to $\widehat{M}$.

By the Itô formula

$$\begin{aligned} d\widetilde{Y}_t &= d((c - \widehat{X}_t^{\pi^*})U_t) \\ &= (c - \widehat{X}_t^{\pi^*})(dA_t + \psi_t^U d\widehat{M}_t + dL_t^U) - U_t \pi_t^* d\widehat{S}_t - \pi_t^* \psi_t^U \rho_t^2 d\langle M\rangle_t \\ &= ((c - \widehat{X}_t^{\pi^*})\psi_t^U - \pi_t^* U_t) d\widehat{M}_t + (c - \widehat{X}_t^{\pi^*}) dL_t^U \\ &\quad + (c - \widehat{X}_t^{\pi^*}) dA_t - (\lambda_t U_t \pi_t^* + \rho_t^2 \psi_t^U \pi_t^*) d\langle M\rangle_t. \end{aligned} \tag{4.19}$$

Since $\widetilde{Y}$ is a martingale with the decomposition

$$\widetilde{Y}_t = \widetilde{Y}_0 + \int_0^t \widetilde{\psi}_u d\widehat{M}_u + \widetilde{L}_t \tag{4.20}$$

comparing the decomposition terms of (4.19) and (4.20) we have

$$\widetilde{\psi}_t = (c - \widehat{X}_t^{\pi^*})\psi_t^U - \pi_t^* U_t, \tag{4.21}$$

$$A_t = \int_0^t \frac{\lambda_s U_s + \rho_s^2 \psi_s^U}{c - \widehat{X}_s^{\pi^*}}\, \pi_s^* d\langle M\rangle_s. \tag{4.22}$$

From (4.17) and (4.21)

$$\pi_t^* = \frac{1}{1-\rho_t^2}\,(\lambda_t \widetilde{Y}_t + \rho_t^2 \widetilde{\psi}_t) = \frac{1}{1-\rho_t^2}\,\Big(\lambda_t U_t (c - \widehat{X}_t^{\pi^*}) - \rho_t^2 U_t \pi_t^* - \rho_t^2 (c - \widehat{X}_t^{\pi^*})\psi_t^U\Big)$$

which gives

$$\pi_t^* = \frac{\lambda_t U_t + \rho_t^2 \psi_t^U}{1-\rho_t^2+\rho_t^2 U_t}\,(c - \widehat{X}_t^{\pi^*}). \tag{4.23}$$

Finally from (4.23) and (4.22) we obtain the equality

$$A_t = \int_0^t \frac{(\lambda_s U_s + \rho_s^2 \psi_s^U)^2}{1-\rho_s^2+\rho_s^2 U_s} d\langle M\rangle_s$$

which means that U_t satisfies (4.18). □

Proposition 4.3. *Let the triple* $(V_t(1), V_t(2), \widehat{X}_t^{\pi^*})$ *satisfies the Forward-Backward stochastic differential equation*

$$dV_t(1) = \frac{(\lambda_t V_t(2) + \rho_t^2\varphi_t(2))(\lambda_t V_t(1) + \rho_t^2\varphi_t(1))}{1-\rho_t^2+\rho_t^2 V_t(2)}\, d\langle M\rangle_t + \varphi_t(1) d\widehat{M}_t + dL_t(1), \quad V_T(1) = \widetilde{H}, \tag{4.24}$$

$$dV_t(2) = \frac{(\lambda_t V_t(2) + \rho_t^2\varphi_t(2))^2}{1-\rho_t^2+\rho_t^2 V_t(2)}\, d\langle M\rangle_t + \varphi_t(2)\, d\widehat{M}_t + dL_t(2), \quad V_T(2) = 1, \tag{4.25}$$

$$\pi_t^* = \frac{\lambda_t V_t(1) + \rho_t^2\varphi_t(1) - \widehat{X}_t^{\pi^*}(\lambda_t V_t(2) + \rho_t^2\varphi_t(2))}{1-\rho_t^2+\rho_t^2 V_t(2)}, \quad \widehat{X}_0^{\pi^*} = 0, \tag{4.26}$$

$$\langle L(1), \widehat{M}\rangle = \langle L(2), \widehat{M}\rangle = 0. \tag{4.27}$$

Then the pair $(\widetilde{Y}, \widetilde{\psi})$*, where*

$$\widetilde{Y}_t = V_t(1) - \widehat{X}_t^{\pi^*} V_t(2) \quad \textit{and} \quad \widetilde{\psi}_t = \varphi_t(1) - V_t(2)\pi_t^* - \varphi_t(2)\widehat{X}_t^{\pi^*}, \tag{4.28}$$

is a solution of equation

$$\widetilde{Y}_T = \widetilde{H} - \int_0^T \frac{1}{1-\rho_t^2}\,(\lambda_t \widetilde{Y}_t + \rho_t^2 \widetilde{\psi}_t)\, d\widehat{S}_t. \tag{4.29}$$

Proof. By the Itô formula

$$\begin{aligned} d\widetilde{Y}_t = &-\big(V_t(2)\pi_t^* + \widehat{X}_t^{\pi^*}\varphi_t(2) - \varphi_t(1)\big)\, d\widehat{M}_t - \widehat{X}_t^{\pi^*} dL_t(2) + dL_t(1) \\ &- \Bigg(V_t(2)\lambda_t\pi_t^* + \widehat{X}_t^{\pi^*}\frac{(\lambda_t V_t(2) + \rho_t^2\varphi_t(2))^2}{1-\rho_t^2+\rho_t^2 V_t(2)} + \rho_t^2\varphi_t(2)\pi_t^* \\ &- \frac{(\lambda_t V_t(1) + \rho_t^2\varphi_t(1))(\lambda_t V_t(2) + \rho_t^2\varphi_t(2))}{1-\rho_t^2+\rho_t^2 V_t(2)}\Bigg) d\langle M\rangle_t. \end{aligned}$$

It follows from (4.26) that the expression in the latter bracket is equal to zero. Thus $\widetilde{Y}_t$ is martingale and $\widetilde{\psi}_t = \varphi_t(1) - V_t(2)\pi_t^* - \varphi_t(2)\widehat{X}_t^{\pi^*}$. By (4.28)

$$\widetilde{Y}_T = \widetilde{H} - \widehat{X}_T^{\pi^*}$$

and inserting $(\widetilde{Y}, \widetilde{\psi})$ in (4.29) we claim

$$\widehat{X}_T^{\pi^*} = \int_0^T \frac{1}{1-\rho_t^2}\,(\lambda_t V_t(1) - \lambda_t V_t(2)\widehat{X}_t^{\pi^*} + \rho_t^2\varphi_t(1) - \rho_t^2 V_t(2)\pi_t^* - \rho_t^2 \widehat{X}_t^{\pi^*})\, d\widehat{S}_t.$$

This means that

$$\pi_t^* = \frac{1}{1-\rho_t^2}\left(\lambda_t V_t(1) + \rho_t^2\varphi_t(1) - \widehat{X}_t^{\pi^*}(\lambda_t V_t(2) + \rho_t^2\varphi_t(2)) - \rho_t^2 V_t(2)\pi_t^*\right)$$

and

$$(1-\rho_t^2+\rho_t^2 V_t(2))\pi_t^* = \lambda_t V_t(1) + \rho_t^2\varphi_t(1) - \widehat{X}_t^{\pi^*}(\lambda_t V_t(2) + \rho_t^2\varphi_t(2)).$$

Obviously this equality coincides with (4.26). Therefore $(\widetilde{Y},\widetilde{\psi})$ satisfies (4.29). □

5. Diffusion Market Model

Let us consider the financial market model

$$dS_t = \mu_t(\eta)dt + \sigma_t(\eta)dw_t^0,$$
$$d\eta_t = a_t(\eta)dt + b_t(\eta)dw_t,$$

subjected to initial conditions, where only the second component η is observed. Here w^0 and w are correlated Brownian motions with $Edw_t^0 dw_t = \rho dt, \rho\in(-1,1)$.

Let us write

$$w_t = \rho w_t^0 + \sqrt{1-\rho^2}w_t^1,$$

where w^0 and w^1 are independent Brownian motions. It is evident that $w^\perp = -\sqrt{1-\rho^2}w^0 + \rho w^1$ is a Brownian motion independent of w and one can express Brownian motions w^0, w^1 in terms of w and $w^\perp$ as

$$w_t^0 = \rho w_t - \sqrt{1-\rho^2}w_t^\perp, \ \ w_t^1 = \sqrt{1-\rho^2}w_t + \rho w_t^\perp. \tag{5.1}$$

We assume that $b^2>0,\ \sigma^2>0$ and coefficients μ,σ,a and b are such that

$$\mathcal{F}^{S,\eta} = \mathcal{F}^{w^0,w}, \mathcal{F}^\eta = \mathcal{F}^w.$$

So the stochastic basis will be $(\Omega, F, \mathcal{F}, P)$, where $\mathcal{F}$ is the natural filtration of (w^0,w) and the flow of observable events is $\mathcal{G}=\mathcal{F}^w$.

We consider the mean variance hedging problem

$$\text{to minimize} \quad E[(x+\int_0^T \pi_t dS_t - H)^2] \quad \text{over all} \quad \pi\in\Pi(\mathcal{G}),$$

where $H\in L^2(\mathcal{F}_T)$ and π_t is a dollar amount invested in the stock at time t.

Comparing with (1.1) we get that in this case

$$M_t = \int_0^t \sigma_s dw_s^0, \quad \langle M\rangle_t = \int_0^t \sigma_s^2 ds, \quad \lambda_t = \frac{\mu_t}{\sigma_t^2}.$$

It is evident that w is a Brownian motion also with respect to the filtration $\mathcal{F}^{w^0,w^1}$ and condition B) is satisfied. Therefore by Proposition 2.1

$$\widehat{M}_t = \rho\int_0^t \sigma_s dw_s.$$

By the integral representation theorem the GKW decompositions (3.1), (3.10) take the following forms

$$H_t = c_H + \int_0^t h_s\sigma_s dw_s^0 + \int_0^t h_s^1 dw_s^1, \tag{5.2}$$

$$H_t = c_H + \rho \int_0^t h_s^{\mathcal{G}} \sigma_s dw_s + \int_0^t h_s^{\perp} dw_s^{\perp}. \tag{5.3}$$

Putting expressions for $w, w^{\perp}$ in (5.3) and equalizing integrands of (5.2) and (5.3) we obtain that

$$h_t = h_t^{\mathcal{G}} \rho^2 - \sqrt{1-\rho^2} \frac{h_t^{\perp}}{\sigma_t}$$

and hence

$${}^p h_t = {}^p(h^{\mathcal{G}})_t \rho^2 - \sqrt{1-\rho^2} \, \frac{{}^p h_t^{\perp}}{\sigma_t}.$$

Therefore by definition of $\widetilde{h}$ (equation (3.16))

$$\widetilde{h}_t = {}^p h_t - {}^p(h^{\mathcal{G}})_t \rho^2 = -\sqrt{1-\rho^2} \, \frac{{}^p h_t^{\perp}}{\sigma_t}. \tag{5.4}$$

It is evident that $\frac{d\langle \widehat{M} \rangle_t}{d\langle M \rangle_t} = \rho^2$ and (3.23) takes the form

$$\widetilde{Y}_T = \widetilde{H} - \frac{1}{1-\rho^2} \int_0^T \widetilde{Y}_t \theta_t \left(\theta_t dt + \rho dw_t\right) - \frac{\rho}{1-\rho^2} \int_0^T \widetilde{\varphi}_t \left(\theta_t dt + \rho dw_t\right) \tag{5.5}$$

for $\widetilde{Y}_t = c + \int_0^t \widetilde{\psi}_s \rho \sigma_s dw_s \equiv c + \int_0^t \widetilde{\varphi}_s dw_s$, where $\theta_t = \frac{\mu_t}{\sigma_t}$.

Since $H \in L^2(\mathcal{F}_T)$, it follows from (5.4) that

$$E \int_0^T \frac{\widetilde{h}_u^2}{1-\rho^2} d\langle M \rangle_u < \infty, \quad E\widetilde{H}^2 < \infty$$

and all conditions of Theorems 3.1 and 3.2 are satisfied. Therefore, there exists a unique solution $(\widetilde{Y}, \widetilde{\varphi})$ of equation (5.5) and the optimal strategy in this case is

$$\pi_t^* = \frac{1}{1-\rho^2} \left(\theta_t \widetilde{Y}_t + \rho \widetilde{\varphi}_t - \sqrt{1-\rho^2} \, {}^p h_t^{\perp} \right) \sigma_t^{-1}. \tag{5.6}$$

Note that one can write the equation (5.5) also in terms of a random variable ξ

$$\begin{aligned} \xi = \widetilde{H} &- \frac{1}{1-\rho^2} \int_0^T E[\xi | \mathcal{F}_t^w] \theta_t \left(\theta_t dt + \rho dw_t\right) \\ &- \frac{\rho}{1-\rho^2} \int_0^T E[D_t \xi | \mathcal{F}_t^w] \left(\theta_t dt + \rho dw_t\right), \end{aligned} \tag{5.7}$$

where D is the stochastic derivative.

Remark 5.1. Let $\rho = 0$ and let θ be deterministic. In this case $w = w^1, \; w^{\perp} = -w^0, \; \mathcal{G} = \mathcal{F}^{w^1}$ and

$$\widehat{M}_t = E\left(\int_0^t \sigma_s dw_s^0 \big| \mathcal{F}_t^{w^1} \right) = 0$$

Therefore equation (3.17) takes the form

$$\widetilde{Y}_T = \widetilde{H} - \int_0^T \widetilde{Y}_t \theta_t^2 dt. \tag{5.8}$$

Since $\widetilde{Y}$ is a $\mathcal{G}$-martingale, by the integral representation theorem

$$\widetilde{Y}_t = \widetilde{Y}_0 + \int_0^t l_s dw_s^1.$$

Note that in the decomposition (1.5) for $\widetilde{Y}$

$$\int_0^t \widetilde{\psi}_u d\widehat{M}_u = 0 \quad \text{and} \quad \widetilde{L}_t = \int_0^t l_s dw_s^1.$$

Besides, in the decomposition (3.10) for $H_t = E(H|\mathcal{F}_t)$

$$\int_0^t h_u^{\mathcal{G}} d\widehat{M}_u = 0, \quad L_t^{H,\mathcal{G}} = \int_0^t h_s \sigma_s dw_s^0 + \int_0^t h_s^1 dw_s^1.$$

Using the integral representation of $\widetilde{H} = \tilde{c}_H + \int_0^T \widetilde{h}_s^1 dw_s^1$ and the formula of the integration by parts we have

$$\widetilde{Y}_T = \widetilde{c}_H + \int_0^T \widetilde{h}_t^1 dw_t^1 - \widetilde{Y}_T \int_0^T \theta_s^2 ds + \int_0^T \int_0^t \theta_u^2 du d\widetilde{Y}_t \quad \text{and}$$

$$\widetilde{Y}_T \left(1 + \int_0^T \theta_s^2 ds \right) = \widetilde{c}_H + \int_0^T \widetilde{h}_t^1 dw_t^1 + \int_0^T l_t \int_0^t \theta_u^2 du dw_t^1.$$

On the other hand,

$$\widetilde{Y}_T \left(1 + \int_0^T \theta_s^2 ds \right) = \widetilde{Y}_0 \left(1 + \int_0^T \theta_s^2 ds \right) + \int_0^T \left(1 + \int_0^T \theta_s^2 ds \right) l_t dw_t^1.$$

Comparing the last two equalities we obtain

$$\widetilde{Y}_0 = \frac{\widetilde{c}_H}{1 + \int_0^T \theta_s^2 ds}, \quad l_t = \frac{\widetilde{h}_t^1}{1 + \int_t^T \theta_s^2 ds}.$$

Therefore the solution of (5.8) is expressed as

$$\widetilde{Y}_t = \frac{\widetilde{c}_H}{1 + \int_0^T \theta_s^2 ds} + \int_0^t \frac{\widetilde{h}_s^1}{1 + \int_s^T \theta_u^2 du} dw_s^1.$$

Since $\widetilde{h} = {}^p h$, the optimal strategy is

$$\pi_t^* = {}^p h_t + \frac{\widetilde{c}_H \lambda_t}{1 + \int_0^T \lambda_s^2 \sigma_s^2 ds} + \lambda_t \int_0^t \frac{\widetilde{h}_s^1}{1 + \int_s^T \lambda_u^2 \sigma_u^2 du} dw_s^1.$$

Proposition 5.1. *Suppose that $H = c_H$, $\eta_t = w_t$ and $\frac{\mu_t}{\sigma_t} = \theta(t, w_t)$ for some continuous function θ, such that the nonlinear PDE*

$$u_t + \frac{1}{2} u_{xx} = \frac{(\theta(t,x)u + \rho u_x)^2}{1 - \rho^2 + \rho^2 u}, \quad u(T, x) = 1 \tag{5.9}$$

admits the sufficiently smooth solution u. Then the solution of (3.23) *can be represented as*

$$\widetilde{Y}_t = c_H u(t, w_t) \mathcal{E}_t \left(- \int_0^\cdot \frac{\theta(s, w_s)u(s, w_s) + \rho u_x(s, w_s)}{1 - \rho^2 + \rho^2 u(s, w_s)} (\theta(s, w_s) ds + \rho dw_s) \right)$$

and the optimal strategy is

$$\pi_t^* = c_H \sigma^{-1}(t,w_t) \frac{\theta(t,w_t)u(t,w_t) + \rho u_x(t,w_t)}{1-\rho^2+\rho^2 u(t,w_t)}$$
$$\times \mathcal{E}_t\left(- \int_0^{\cdot} \frac{\theta(s,w_s)u(s,w_s) + \rho u_x(s,w_s)}{1-\rho^2+\rho^2 u(s,w_s)} (\theta(s,w_s)ds + \rho dw_s) \right). \tag{5.10}$$

Sketch of the proof. It is well known that if $u(t,x)$ is the solution of (5.9), then $V_t(2) = u(t,w_t)$ will be the solution of (4.25). On the other hand $V_t(1) = cV_t(2)$ and $c - \widehat{X}_t^{\pi^*} = c\mathcal{E}_t\big(- \int_0^{\cdot} \frac{\lambda_s V_s + \phi_s \rho_s^2}{1-\rho_s^2+\rho_s^2 V_s} \, d\widehat{S}_s\big)$. Moreover, similarly to Proposition 4.3 it can be verified that $\widetilde{Y}_t = (c - \widehat{X}_t^{\pi^*})V_t(2)$ satisfies equation (3.23) and it follows from (4.23) that π^* is of the form (5.10).

The detailed proof we shall give in Appendix A.

Example. If $\theta(t,x) = \theta(t)$ then the solution of (5.9) is of the form $u(t,x) = u(t)$, where u satisfies

$$\frac{du(t)}{dt} = \frac{\theta^2(t)u^2(t)}{1-\rho^2+\rho^2 u(t)}, \quad u(T) = 1.$$

Thus

$$-\frac{1-\rho^2}{u(s)} + \rho^2 \ln u(s)\Big|_t^T = \int_t^T \theta^2(s)ds.$$

If we denote by $\nu(\rho,\alpha)$ the unique solution of

$$\frac{1-\rho^2}{u} - \rho^2 \ln u = \alpha,$$

then $u(t) = \nu(\rho, 1-\rho^2 + \int_t^T \theta^2(s)ds)$ and the solution of (5.7) is explicitly given by

$$\xi = c_H \mathcal{E}_T\left(- \int_0^{\cdot} \frac{\theta(s)\nu(\rho, 1-\rho^2+\int_s^T \theta^2(u)du)}{1-\rho^2+\rho^2\nu(\rho,1-\rho^2+\int_s^T \theta^2(u)du)} (\theta(s)ds + \rho dw_s) \right).$$

Appendix A. Appendix

Proof of Proposition 5.1. It easy to see that (5.9) is equivalent to

$$u_t - \rho \frac{\theta(t,x)uu_x + \rho u_x^2}{1-\rho^2+\rho^2 u} + \frac{1}{2}u_{xx} = \frac{\theta^2(t,x)u^2 + \rho\theta(t,x)uu_x}{1-\rho^2+\rho^2 u}, \quad u(T,x) = 1.$$

If u is a solution of (5.9) then using the notation $g = -\frac{\theta(t,x)u + \rho u_x}{1-\rho^2+\rho^2 u}$, the Feynmann-Kac formula and Girsanov's theorem we can write

$$u(t,x) = E\left(\mathcal{E}_{tT}\left(\int_0^{\cdot} g(s,w_s)(\theta(s,w_s)ds + \rho dw_s) \right) \big| w_t = x \right).$$

Hence the integrand $\tilde{\varphi}$ of the integral representation of the martingale $\widetilde{Y}_t = c_H E\left[\mathcal{E}_T\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)|\mathcal{F}_t^w\right]$ can be calculated as follows

$$
\begin{aligned}
&\widetilde{\varphi}_t dw_t = d\widetilde{Y}_t\\
&= c_H d\left(\mathcal{E}_t\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)u(t,w_t)\right)\\
&= c_H\mathcal{E}_t\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds+\rho dw_s)\right)\\
&\quad\times\left(u_x(t,w_t)dw_t - g(t,w_t)\left(\theta(t,w_t)u(t,w_t)+\rho u_x(t,w_t)\right)dt\right)\\
&\quad + c_H u(t,w_t)\mathcal{E}_t\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)g(t,w_t)\left(\theta(t,w_t)dt + \rho dw_t\right)\\
&\quad + c_H\mathcal{E}_t\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)\rho g(t,w_t)u_x(t,w_t)dt\\
&= c_H\mathcal{E}_t\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)(u_x(t,w_t) + \rho u(t,w_t)g(t,w_t))dw_t.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\frac{1}{1-\rho^2}\int_0^T \widetilde{Y}_t\theta(t,w_t)(\theta(t,w_t)dt + \rho dw_t) + \frac{\rho}{1-\rho^2}\int_0^T \widetilde{\varphi}_t(\theta(t,w_t)dt + \rho dw_t)\\
&\quad = c_H\frac{1}{1-\rho^2}\int_0^T \left(u(t,w_t)\theta(t,w_t) + \rho u_x(t,w_t) + \rho^2 g(t,w_t)u(t,w_t)\right)\\
&\qquad\quad\times\mathcal{E}_t\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)(\theta(t,w_t)dt + \rho dw_t).
\end{aligned}
$$

Since $u\theta + \rho u_x + \rho^2 gu = (\rho^2-1)g$, then

$$
\begin{aligned}
&\frac{1}{1-\rho^2}\int_0^T \widetilde{Y}_t\theta(t,w_t)(\theta(t,w_t)dt + \rho dw_t) + \frac{\rho}{1-\rho^2}\int_0^T \widetilde{\varphi}_t(\theta(t,w_t)dt + \rho dw_t)\\
&= -c_H\int_0^T g(t,w_t)\mathcal{E}_t\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)\left(\theta(t,w_t)dt + \rho dw_t\right).
\end{aligned}
$$

On the other hand,

$$
\widetilde{Y}_T = c_H\mathcal{E}_T\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)
$$

$$
= c_H + \int_0^T c_H g(t,w_t)\mathcal{E}_t\left(\int_0^\cdot g(s,w_s)(\theta(s,w_s)ds + \rho dw_s)\right)(\theta(t,w_t)dt + \rho dw_t).
$$

Hence (5.5) is satisfied. The expression for π^* is obtained from the representation

$$
\begin{aligned}
\widetilde{Y}_T =& c_H\mathcal{E}_T\left(-\int_0^\cdot \frac{\theta(s,w_s)u(s,w_s) + \rho u_x(s,w_s)}{1-\rho^2+\rho^2 u(s,w_s)}(\theta(s,w_s)ds + \rho dw_s)\right)\\
=& - c_H\int_0^T \frac{\theta(s,w_s)u(s,w_s) + \rho u_x(s,w_s)}{1-\rho^2+\rho^2 u(s,w_s)}
\end{aligned}
$$

$$\times\, \mathcal{E}_s\left(-\int_0^{\cdot} \frac{\theta u + \rho u_x}{1-\rho^2+\rho^2 u}(\theta du + \rho dw_u)\right)(\theta(s,w_s)ds + \rho dw_s)$$

and equations (5.5) and (5.6). □

ACKNOWLEDGMENTS

We would like to thank N. Lazrieva and the anonymous referee for valuable remarks and suggestions.

This work was supported by Georgian National Science Foundation grant STO07/3-172.

REFERENCES

[1] A. Cerny and J. Kallsen , On The Structure of General Mean-Variance Hedging Strategies, *Ann. Probab.*, **35**, 4, (2007), 1479-1531 .

[2] T. Choulli, L. Krawczyk and C. Stricker, E-martingales and their applications in mathematical finance, *Ann. Probab.* **26**, 2, (1998), 853-876.

[3] G. B. Di Masi, E. Platen and W. J. Runggaldier , Hedging of Options under Discrete Observation on Assets with Stochastic Volatility, *Seminar on Stoch. Anal. Rand. Fields Appl.*, (1995) 359-364.

[4] F. Delbaen, P. Monat, W. Schachermayer, W. Schweizer and C. Stricker, Weighted norm inequalities and hedging in incomplete markets, *Finance Stoch.* **1, No. 2**, (1997), 181-227.

[5] F. Delbaen and W. Schachermayer, Variance-optimal martingale measure for continuous processes, *Bernoulli* **2**, 1 (1996), 81-105.

[6] C. Dellacherie and P. A. Meyer, Probabilités et potentiel, II. *Hermann, Paris,* 1980.

[7] D. Duffie and H. R. Richardson, Mean-Variance hedging in continuous time. *Ann. Appl. Probab.* **1**(1991), 1–15.

[8] N.El Karoui and M.C. Quenez, Dynamic programming and pricing of contingent claims in an incomplete market, *SIAM J. Control Optim.* **33** 1 (1995), 29-66.

[9] H. Föllmer and D. Sondermann, Hedging non-redundant contingent claims. *Contributions to mathematical economics hon. G. Debreu* (*W. Hildenbrand and A. Mas-Collel, eds.*),. *North Holland, Amsterdam,* (1986), 205–223.

[10] R. Frey and W. J. Runggaldier, Risk-minimizing hedging strategies under restricted information: the case of stochastic volatility models observable only at discrete random times. Financial optimization. *Math. Methods Oper. Res.* **50**, no.2, (1999), 339–350.

[11] C. Gourieroux, J. P. Laurent, and H. Pham, Mean-variance hedging and numeraire. *Math. Finance* **8**, No. 3(1998), 179–200.

[12] D. Heath, E. Platen, and M. Schweizer, A comparision of two quadratic approaches to hedging in incomplete markets. *Math. Finance* **11** No. 4 (2001) 385–413.

[13] C. Hipp, Hedging general claims. *Proc. 3rd AFIR Colloquium* 2, 603–613, *Rome,* 1993.

[14] N. Hoffman, E. Platen and M. Schweizer, Option pricing under incompleteness and stochastic volatility, *Math. Finance*, **2**, N. 3, 1992, pp. 153-187.

[15] J. Jacod, Calcule Stochastique et problèmes des martingales. *Lecture Notes in Math.* **714**, *Springer, Berlin etc.,* (1979).

[16] N. Kazamaki, Continuous exponential martingales and BMO, *Lecture Notes in Math.* **1579**, Springer, New York, 1994.

[17] D.O.Kramkov, Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets, *Probab. Theory Related Fields* **105** (1996), 459–479.

[18] R.Sh. Liptser and A. N. Shiryaev, Statistics of random processes, *Springer, Berlin etc.*, (1977).

[19] R.Sh. Liptzer and A.N. Shiryayev, Martingale theory, Nauka, Moscow, 1986.

[20] M. Mania and R. Tevzadze, Backward Stochastic PDE and Imperfect Hedging, *Int. J. Theor. Appl. Finance*, **6**, 7,(2003),663-692.

[21] M. Mania, R. Tevzadze and T. Toronjadze, Mean-variance Hedging Under Partial Information, *SIAM J. Control Optim.*, **47**, 5, (2008) , 2381-2409.

[22] E. Pardoux and S.G.Peng, Adapted solution of a backward stochastic differential equation, *Systems Control Lett.* **14** (1990),55–61.

[23] H. Pham, Mean-variance hedging for partially observed drift processes *Int. J. Theor. Appl. Finance* **4**, No. 2, (2001), 263–284.

[24] T. Rheinländer and M. Schweizer, On L^2- Projections on a space of Stochastic Integrals, *Ann. Probab.* **25** (1997), 1810–1831.

[25] M. Schäl, On quadratic cost criteria for option hedging, *Math. Oper. Res.* **19** (1994), 121–131.

[26] M. Schweizer, Mean-variance hedging for general claims. *Ann. Appl. Probab.* **2**(1992), 171–179.

[27] M. Schweizer, Approximating random variables by stochastic integrals. *Ann. Probab.* **22**,(1994), No. 3, 1536–1575.

[28] M. Schweizer, Risk-minimizing hedging strategies under restricted information. *Math. Finance* **4**, No.4, (1994), 327-342.

THE ROBBINS–MONRO TYPE STOCHASTIC DIFFERENTIAL EQUATIONS. III. POLYAK'S AVERAGING

N. LAZRIEVA AND T. TORONJADZE

Abstract. General results concerning the asymptotic behaviour of the Polyak averaging $\overline{z} = (\overline{z}_t)_{t\geq 0}$ of the solution of the Robbins–Monro type stochastic differential equation are presented. It is shown that the suitable normed process $\overline{z}$ admits an asymptotic expansion which enables one to obtain its asymptotic distribution from a Central Limit Theorem for martingales.

Key words and phrases: Stochastic approximation; Robbins–Monro type SDE; Recursive estimation; Polyak's averaging

MSC 2010: 62L20; 60H10

INTRODUCTION

In recent years the interest to stochastic approximation and its applications substantially increased including such fields of applications as statistics [5], medicine and engineering [17], adaptive control, signal processing, weirless communication [4], [1] and even mathematical finance [19].

An important approach to stochastic approximation problems has been proposed by Polyak [13] and Ruppert [15]. The main idea of this approach is the use of averaging iterates obtained from primary schemes. In [14], [18], [3], [8] and [9] it was proved that such algorithms can provide strongly consistent estimates which are asymptotically efficient. In several cases the averaging can provide the rate of convergence higher than the rate of the primary process. Le Breton and Novikov [10], [11] concentrate their attention on a general multidimensional linear regression model with Gaussian errors. They demonstrate that averaging can also provide consistent estimates for asymptotic covariances of estimates. Melnikov and Valkeila [16] consider the averaging procedure both for the so-called standard RM procedure and for procedures with slowly varying gains. They have proved the convergence of averaging procedures and studies the asymptotic properties of these procedures.

In the present paper we study the asymptotic behaviour of the Polyak averaging procedure for the Robbins–Monro type (RM type) SDE introduced in [6]

$$z_t = z_0 + \int_0^t H(s, z_{s-})\, dK_s + \int_0^t M(ds, z_{s-}), \tag{0.1}$$

where $K = \{K_t,\ t \geq 0\}$ is an increasing predictable process, $H(t,u)$ and $M(t,u)$, $t \geq 0$, $u \in R^1$, are random fields given on some stochastic basis. We assume that for each $t \geq 0$

$$H(t,0) = 0, \quad H(t,u)u < 0 \quad \text{for} \quad u \neq 0 \quad P\text{-a.s.},$$

Published in *Stochastics* **82** (2010), no. 1-3, 165–188.

for each $u \in R^1$, $M(u) = \{M(t,u),\ t \geq 0\} \in \mathcal{M}^2_{loc}$, the symbol $\int_0^t M(ds, z_{s-})$ is used for the stochastic line integral (see [6] for more details).

Equation (0.1) naturally includes both generalized RM stochastic approximation algorithm with martingale noises [16] and recursive estimation procedures for parametric semimartingale statistical models, and enables one to study them by a common approach.

The present work is the final part of series of papers [6], [7] concerning the asymptotic behaviour of solution $z = (z_t)_{t\geq 0}$ of equation (0.1).

We define the Polyak averaging procedure for the process $z = (z_t)_{t\geq 0}$ by the formula

$$\overline{z}_t = \frac{1}{\mathcal{E}_t^{-1}(-g \circ K)} \int_0^t z_s d\mathcal{E}_s^{-1}(-g \circ K), \tag{0.2}$$

where $g_t \geq 0$, $g_t \Delta K_t < 1$, $g \circ K_t < \infty$ for all $t \geq 0$, $g \circ K_\infty = \infty$ P-a.s. Here $g \circ K_t = \int_0^t g_s\, dK_s$, $\mathcal{E}(X)$ is the Dolean exponential of X. Denote $\mathcal{E}_t^{-1} := \mathcal{E}_t^{-1}(-g \circ K)$.

The aim of the present paper is to study the asymptotic properties of process $\overline{z} = (\overline{z}_t)_{t\geq 0}$ defined by (0.2).

First note that if

$$z_t \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.},$$

then since $(\mathcal{E}_t^{-1})_{t\geq 0}$ is an increasing process, $\mathcal{E}_\infty^{-1} = \infty$ the Toeplitz lemma (see Appendix A) yields

$$\overline{z}_t \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.}$$

We show that under sufficiently mild conditions the normed process $\overline{z} = (\overline{z}_t)_{t\geq 0}$ admits the following representation

$$\mathcal{E}_t^{-1} \widetilde{B}_t^{1/2} \overline{z}_t = \frac{\int_0^t (B_t - B_{s-}) dL_s}{\widetilde{B}_t^{1/2}} + \overline{R}_t, \quad \overline{R}_t \xrightarrow{P} 0, \quad \text{as} \quad t \to \infty, \tag{0.3}$$

where

$$B_t := \int_0^t \Gamma_s^{-1} d\mathcal{E}_s^{-1}, \quad \widetilde{B}_t := \int_0^t (B_t - B_{s-})^2 d\langle L \rangle_s \tag{0.4}$$

and objects $(\Gamma_t)_{t\geq 0}$, $(L_t)_{t\geq 0}$ and $(\langle L \rangle_t)_{t\geq 0}$ are defined by Eq. (1.1) below.

For instance, if we choose $\mathcal{E}_t^{-1} := 1 + \int_0^t \Gamma_s^2 \langle L \rangle_s^{-1} \beta_s dK_s$, then $\mathcal{E}^{-1}\widetilde{B} \sim \mathcal{E}^{-1/2}$ (see Appendix A, Definition A.1) and the asymptotic distribution of process $\mathcal{E}^{-1/2}\overline{z}$ coincides with the asymptotic distribution of $\widetilde{B}_t^{-1/2} \int_0^t (B_t - B_{s-}) dL_s$ (up to the constant factor) as $t \to \infty$. As a special cases we obtain the results of [16] concerning the asymptotics of averaging procedure for standard RM stochastic approximation algorithm, as well as for RM algorithms with slowly varying gains.

The paper is organized as follows: In Section 1 the main objects and assumptions are introduced. In Section 2 we study the asymptotic properties of averaging procedure $\overline{z}$ in the linear case. The general case we consider in Section 3. In Section 4 the special cases are considered. Section 5 is devoted to proof of results. In Appendix A some definitions and technical results are given and in Appendix B we collect necessary results from [6] and [7] for convenience of readers.

All notation and facts concerning martingale theory can be found in [12].

1. Preliminaries

Let on a filtered probability space $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geq 0}, P)$ satisfying the usual conditions the following objects be defined:

(1) A predictable increasing process $K = (K_t)_{t\geq 0}$.
(2) A random field $\{H(t,u),\ t \geq 0,\ u \in R^1\}$ such that $(H(t,u))_{t\geq 0}$ is a predictable process for each $u \in R^1$ and

$$\text{(A)}\qquad \begin{aligned} &H(t,0) = 0,\\ &H(t,u)u < 0, \quad u \neq 0, \end{aligned}$$

for all $t \geq 0$ P-a.s.
(3) A random field $\{M(t,u),\ t \geq 0,\ u \in R^1\}$ such that for each $u \in R^1$

$$M(u) = (M(t,u))_{t\geq 0} \in \mathcal{M}^2_{loc}(P), \quad M(t,0) \neq 0, \quad t \geq 0, \ \ P\text{-a.s.},$$
$$\langle M(u), M(v)\rangle_t = h(u,v) \circ K_t, \quad h(0,0) \circ K_\infty < \infty.$$

Denote $\ell^2_t := h_t(0,0)$, $M_t := M(t,0)$. Evidently, $\langle M\rangle_\infty < \infty$.

Assume that there exists an unique strong solution $z = (z_t)_{t\geq 0}$ of Eq. (0.1) on the whole time interval $[0,\infty)$ such that (see [6])

$$(\widetilde{M}_t)_{t\geq 0} := \left(\int_0^t M(ds, z_{s-})\right)_{t\geq 0} \in \mathcal{M}^2_{loc}(P).$$

Let us denote

$$\beta_t := -\lim_{u\to 0} \frac{H(t,u)}{u}$$

assuming that this limit exists for each $t \geq 0$ P-a.s. and define

$$\beta_t(u) = \begin{cases} -\frac{H(t,u)}{u} & \text{if } u \neq 0,\\ \beta_t & \text{if } u = 0. \end{cases}$$

It follows from (A) that for all $t \geq 0$ and $u \in R^1$

$$\beta_t \geq 0 \quad \text{and} \quad \beta_t(u) \geq 0 \quad P\text{-a.s.}$$

Throughout of this paper we are working under the following assumptions:

Assumption 1: For all $t \geq 0$ $\beta_t \Delta K_t \neq 1$, $\beta_t\Delta K_t < 1$ eventually, $\lim_{t\to\infty} \beta_t \Delta K_t = m$, $0 \leq m \leq 1$, $\beta \circ K_t < \infty$, $\beta \circ K_\infty = \infty$ P-a.s.

Define further the objects: for all $t \geq 0$

$$\Gamma_t := \mathcal{E}_t^{-1}(-\beta \circ K), \quad L_t := \int_0^t \Gamma_s\, dM_s, \quad \langle L\rangle_t := 1 + \int_0^t \Gamma_s^2 \ell_s^2\, dK_s. \tag{1.1}$$

Assumption 2: $\langle L\rangle_\infty = \infty$ P-a.s.

Assumption 3: There exists a predictable increasing process $\gamma = (\gamma_t)_{t\geq 0}$ equivalent to $\left(\Gamma_t^2 \langle L\rangle_t^{-1}\right)_{t\geq 0}$ with

$$\lim_{t\to\infty} \frac{\Gamma_t^2 \langle L\rangle_t^{-1}}{\gamma_t} = \widetilde{\gamma}^{-1}, \quad 0 < \widetilde{\gamma} < \infty \quad P\text{-a.s.}$$

Besides, we assume that $\gamma_0 = 1$, $\gamma_\infty = \infty$, $\frac{\Delta\gamma_t}{\gamma_t} < 1$ for all $t \geq 0$ P-a.s., also $\gamma_t = 1 + \int_0^t \widetilde{g}_s \, dK_s$,for some appropriate $\widetilde{g}$.

Remark 1. From Assumption 1 it directly follows that $(\Gamma_t)_{t\geq 0}$ is an increasing process eventually, $\Gamma_0 = 1$, $\Gamma_\infty = \infty$ with $\lim\limits_{t\to\infty} \Gamma_t^2 \langle L\rangle_t^{-1} = \infty$ P-a.s.

It is not hard to show (see [7]) that the process $z = (z_t)_{t\geq 0}$ can be written as

$$z_t = \Gamma_t^{-1} \left(z_0 + L_t + \int_0^t \Gamma_s \, d\overline{R}_s \right), \tag{1.2}$$

where

$$\begin{aligned} \overline{R}_t &:= \overline{R}_t^1 + \overline{R}_t^2, \\ \overline{R}_t^1 &:= \int_0^t \left(\beta_s - \beta_s(z_{s-})\right) z_{s-} \, dK_s, && (1.3) \\ \overline{R}_t^2 &:= \int_0^t \left(M(ds, z_{s-}) - M(ds, 0)\right). && (1.4) \end{aligned}$$

Moreover, the normed process $(z_t)_{t\geq 0}$ admits the following asymptotic expansion:

$$\Gamma_t \langle L\rangle_t^{-1/2} z_t = \frac{L_t}{\langle L\rangle_t^{1/2}} + R_t, \tag{1.5}$$

where

$$R_t := \frac{z_0}{\langle L\rangle_t^{1/2}} + \frac{1}{\langle L\rangle_t^{1/2}} \int_0^t \Gamma_s \, d\overline{R}_s \,.$$

The conditions sufficient for the convergence

$$R_t \xrightarrow{P} 0 \quad \text{as } \ t \to \infty, \tag{1.6}$$

were studied in [7].

From Assumptions 1, 2 and 3, Eqs. (1.5) and (1.6) we obtain

$$\lim_{t\to\infty} \mathcal{L}(\gamma_t^{1/2} z_t) = \lim_{t\to\infty} \mathcal{L}\left(\widetilde{\gamma} \, \frac{L_t}{\langle L\rangle_t^{1/2}} \right)$$

in the sense of weak convergence.

2. Asymptotic Properties of Process $\overline{z}$ in the Linear Case

In this section we consider the following linear equation

$$dz_t = -\beta_t z_{t-} \, dK_t + dM_t, \quad z_0. \tag{2.1}$$

Solving this equation we have

$$z_t = \Gamma_t^{-1} \left(z_0 + \int_0^t \Gamma_s \, dM_s \right). \tag{2.2}$$

Since $(\Gamma_t)_{t\geq 0}$ is an increasing process, $\Gamma_\infty = \infty$ (see Remark 1 of Section 1), $\langle M\rangle_\infty < \infty$ it follows from the stochastic version of Kronecker lemma (see [12, Ch. 2, § 6, Lemma 3]) that

$$z_t \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.}$$

Multiplying both sides of Eq. (2.2) by $\Gamma_t \langle L\rangle_t^{-1/2}$ yields

$$\Gamma_t \langle L\rangle_t^{-1/2} z_t = z_0 \langle L\rangle_t^{-1/2} + L_t/\langle L\rangle_t^{1/2} \tag{2.3}$$

and hence (by Assumption 2)

$$\lim_{t\to\infty} \mathcal{L}\left(\Gamma_t \langle L\rangle_t^{-1/2} z_t\right) = \lim_{t\to\infty} \mathcal{L}\left(L_t/\langle L\rangle_t^{1/2}\right) \tag{2.4}$$

in the sense of a weak convergence.

Further, substituting (2.2) in (0.2) we get

$$\overline{z}_t = z_0\, \frac{B_t}{\mathcal{E}_t^{-1}} + \frac{L \circ B_t}{\mathcal{E}_t^{-1}}\,. \tag{2.5}$$

To continue we need additional

Assumption 4 : The processes $(\beta_t)_{t\geq 0}$, $(\ell_t^2)_{t\geq 0}$, $(g_t)_{t\geq 0}$, $(K_t)_{t\geq 0}$ are deterministic.

Since $(B_t)_{t\geq 0}$ is deterministic, then applying the Itô formula we get $L \circ B_t = \int_0^t (B_t - B_s)\, dL_s$. Multiplying Eq. (2.5) by $\mathcal{E}_t^{-1}(\widetilde{B}_t)^{-1/2}$ we obtain

$$\mathcal{E}_t^{-1}(\widetilde{B}_t)^{-1/2}\overline{z}_t = z_0\, \frac{B_t}{\widetilde{B}_t^{1/2}} + \frac{\int_0^t (B_t - B_{s-})\, dL_s}{\widetilde{B}_t^{1/2}}\,, \tag{2.6}$$

where $(B_t)_{t\geq 0}$, $(\widetilde{B}_t)_{t\geq 0}$ are defined by (0.4).

The main result of this section is

Proposition 2.1. *Let the following conditions be satisfied:*

(1) $\langle L\rangle \circ B_\infty = \infty$,

(2) $(\langle L\rangle \circ B) \circ B_\infty = \infty$,

(3) $\displaystyle\lim_{t\to\infty} \frac{\langle L\rangle_t \Delta B_t}{\langle L\rangle \circ B_t} = c, \quad 0 \leq c < 2.$

Then

$$\frac{B_t}{\widetilde{B}_t^{1/2}} \to 0 \quad \textit{as} \quad t \to \infty,$$

and

$$\lim_{t\to\infty} \mathcal{L}\left(\mathcal{E}_t^{-1}(\widetilde{B}_t)^{-1/2}\overline{z}_t\right) = \lim_{t\to\infty} \mathcal{L}\biggl(\frac{\int_0^t (B_t - B_{s-})\, dL_s}{\widetilde{B}_t^{1/2}}\biggr). \tag{2.7}$$

Proposition 2.2. *Let*

$$\mathcal{E}_t^{-1} = 1 + \int_0^t \Gamma_s^2 \langle L\rangle_s^{-1} \beta_s\, dK_s. \tag{2.8}$$

Then conditions (1)–(3) *of Proposition* 2.1 *are satisfied. Moreover,*

$$\lim_{t\to\infty} \frac{\widetilde{B}_t}{\mathcal{E}_t^{-1}} = 2 - m \tag{2.9}$$

and

$$\lim_{t\to\infty} \mathcal{L}(\mathcal{E}_t^{-1/2}\overline{z}_t) = \lim_{t\to\infty} \mathcal{L}\left(\sqrt{2-m}\, \frac{\int_0^t (B_t - B_{s-})\, dL_s}{\widetilde{B}_t^{1/2}}\right). \tag{2.10}$$

Remark 2. Eq. (2.8) gives the natural choice of $\mathcal{E}_t^{-1}$ which becomes clear from the proof of this proposition, see Eqs. (5.3) and (5.4).

From now we define averaging procedure (0.2) with $\mathcal{E}_t^{-1}$ given by (2.8).

3. The Asymptotic Properties of Process $\overline{z}$ in General Case

In this section we study asymptotic properties of $\overline{z} = (\overline{z})_{t\geq 0}$ defined by Eq. (0.2) with $\mathcal{E}_t^{-1}$ given by (2.8).

Substituting (1.2) in (0.2) and multiplying both sides of resulting equation by $\mathcal{E}_t^{-1}\widetilde{B}_t^{-1/2}$ we obtain

$$\mathcal{E}_t^{-1}\widetilde{B}_t^{-1/2}\overline{z}_t = z_0\, \frac{B_t}{\widetilde{B}_t^{1/2}} + \frac{\int_0^t (B_t - B_{s-})\, dL_s}{\widetilde{B}_t^{1/2}} + r_t\,, \tag{3.1}$$

where

$$r_t = \frac{1}{\widetilde{B}_t^{1/2}}\left(\int_0^t R_s^1\, dB_s + \int_0^t R_s^2\, dB_s\right)$$

and $R_t^i := \int_0^t \Gamma_s\, d\overline{R}\,_s^i$, $\overline{R}\,_s^i$, $i=1,2$, are defined by (1.3) and (1.4), respectively.

Denote

$$r_t^i := \frac{1}{\widetilde{B}_t^{1/2}} \int_0^t R_s^i\, dB_s, \quad i=1,2.$$

As is seen, the first two terms in the right-hand side of Eq. (3.1) coincides with those in Eq. (2.6). Hence, our problem is to prove that

$$r_t \xrightarrow{P} 0 \quad \text{as} \quad t\to\infty.$$

Since $\widetilde{B} \sim \mathcal{E}^{-1}$, it is sufficient to establish conditions under which

$$\widetilde{r}_t^i := \frac{1}{\mathcal{E}_t^{-1/2}} \int_0^t R_s^i\, dB_s \xrightarrow{P} 0 \quad \text{as} \quad t\to\infty, \quad i=1,2.$$

First we consider the case $i=1$.

Lemma 3.1. *Suppose that the following conditions are satisfied:*

(i) $$\frac{\Gamma_t^2\langle L\rangle_t^{-1}}{1+\int_0^t \Gamma_s^2\langle L\rangle_s^{-1}\beta_s\, dK_s} < 1 \quad \textit{eventually},$$

(ii) $$\int_0^\infty \mathcal{E}_s^{-1/2}|\beta_s - \beta_s(z_{s-})|\,|z_{s-}|\, dK_s < \infty \quad P\text{-a.s.} \quad 0\leq c<1,$$

Then

$$\widetilde{r}_t^1 \to 0 \quad \textit{as} \quad t\to\infty, \quad P\text{-a.s.} \tag{3.2}$$

Suppose that the following assumption holds.

Assumption 5 : (1) $z_t \to 0$ as $t \to \infty$ P-a.s.,
(2) For each δ, $0 < \delta < \delta_0$, $0 < \delta_0 \leq 1$,

$$\gamma_t^\delta z_t^2 \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.},$$

where $\gamma = (\gamma_t)_{t \geq 0}$ is an increasing predictable process presented in Assumption 3 with $\gamma_0 = 0$, $\gamma_\infty = \infty$ P-a.s.

Remark 3. The conditions sufficient for (1) and (2) were studied in [6] and [7], respectively.

Corollary 3.2. (1) *Let the following condition be satisfied:*
(ii)$'$ *There exists δ, $0 < \delta < \frac{\delta_0}{2}$, such that*

$$\int_0^\infty \mathcal{E}_t^{-1/2} |\beta_t - \beta_t(z_{t-})| \gamma_{t-}^{-\delta} \, dK_t < \infty. \tag{3.3}$$

Then (3.2) *holds.*
(2) *Assume in addition that*

$$H(t,u) = -\beta_t u + v(t,u)$$

$$\lim_{u \to 0} \frac{v(t,u)}{u^2} = v_t,$$

where $(v_t)_{t \geq 0}$ some predictable process.
Suppose that the following condition (ii)$''$ *is satisfied.*
(ii)$''$ *There exists δ, $0 < \delta < \frac{\delta_0}{2}$, such that*

$$\int_0^\infty \mathcal{E}_t^{-1/2} |v_t| \gamma_{t-}^{-2\delta} \, dK_t < \infty \quad P\text{-}a.s.$$

Then (ii)$''$$\Rightarrow$(ii)$'$$\Rightarrow$(ii).

Let us consider the case $i = 2$. Denote $N_t = \int_0^t \Gamma_s [M(ds, z_{s-}) - M(ds, 0)]$. Then

$$\widetilde{r}_t^2 = \frac{1}{\mathcal{E}_t^{-1/2}} \int_0^t N_s \, dB_s.$$

Lemma 3.3. *Let the condition*

$$\frac{\langle N \rangle_t}{\langle L \rangle_t} \xrightarrow{P} 0 \quad as \quad t \to \infty, \tag{3.4}$$

be satisfied. Then

$$\widetilde{r}_t^2 \xrightarrow{P} 0 \quad as \quad t \to \infty.$$

All above-given results we summarize in the following

Theorem 3.4. *Let conditions of Lemmas* 3.1 *and* 3.3 *be satisfied. Then* (2.10) *holds true.*

4. Special Cases

Case 1. This case illustrate that the rate of convergence of $\overline{z}$ is higher than of z.

In Eq. (2.1) let $K_t = t$, $\beta_t = \beta(1+t)^{-(\frac{1}{2}+\alpha)}$, $\ell_t^2 = \sigma_t^2(1+t)^{-(\frac{3}{2}+\alpha)}$, where α, β are some constants, $\beta > 0$, $0 < \alpha < \frac{1}{2}$, $0 < \sigma_t^2 < c$, $\lim\limits_{t\to\infty} \sigma_t^2 = \sigma^2$, $\sigma^2 > 0$.

It is not hard to observe that $\langle L\rangle_\infty = \infty$.

Put $\gamma_t = 1+t$. Then since $\Delta K_t = 0$, $\gamma_t^{-1}\,\frac{\beta_t}{\ell_t^2} = \frac{\beta}{\sigma_t^2} \to \frac{\beta}{\sigma^2}$, and $\gamma_t^{-1}\,\frac{\lambda_t}{\ell_t^2} = \frac{1}{\sigma_t^2}\,(1+t)^{\alpha-\frac{1}{2}} \to 0$ as $t\to\infty$, conditions (a), (b) and (c) of Proposition A.3 are satisfied with $c_1 = c_3 = 0$, $c_2 = \frac{\beta}{\sigma^2}$. Thus

$$\lim_{t\to\infty} \frac{\Gamma_t^2 \langle L\rangle_t^{-1}}{1+t} = 2\,\frac{\beta}{\sigma^2}\,.$$

Then from Eqs. (2.3) and (2.4) follows

$$(1+t)^{1/2} z_t \xrightarrow{d} N\left(0, \frac{\sigma^2}{2\beta}\right),$$

where $N(a,\sigma^2)$ stands for normal distribution with parameters a and σ^2.

Obviously in (2.9) $m = 0$ and hence

$$\lim_{t\to\infty} \mathcal{L}(\mathcal{E}_t^{-1/2}\overline{z}_t) = \lim_{t\to\infty} \mathcal{L}\Bigg(\sqrt{2}\,\frac{\int_0^t (B_t - B_s)\,dL_s}{\widetilde{B}^{1/2}}\Bigg). \tag{4.1}$$

On the other hand, it is not hard to check that

$$\frac{\mathcal{E}_t^{-1}}{(1+t)^{(\frac{3}{2}-\alpha)}} \to \frac{4\beta^2}{\sigma^2(3-2\alpha)} \quad \text{as} \quad t\to\infty. \tag{4.2}$$

Hence, from (4.1) and (4.2) using the usual technique developed in the proof of Lemma 3.3 we obtain

$$(1+t)^{\frac{1}{2}(\frac{3}{2}-\alpha)}\overline{z}_t \xrightarrow{d} N\left(0, \left(\frac{3}{2}-\alpha\right)\frac{\sigma^2}{\beta^2}\right).$$

Note that $\frac{1}{2}(\frac{3}{2}-\alpha) > \frac{1}{2}$.

Case 2. Standard Linear Procedure. Let $\gamma = (\gamma_t)_{t\ge0}$ be an increasing predictable process, $\gamma_0 = 1$, $\gamma_t = 1 + \int_0^t \widetilde{g}\,dK_s$, $\widetilde{g}_t \ge 0$, $\widetilde{g}\circ K_\infty = \infty$ P-a.s. Obviously, γ_t can be written as $\gamma_t = \mathcal{E}_t(\lambda\circ K)$, where $\lambda_t = \frac{\widetilde{g}_t}{\gamma_{t-}}$.

Put in (2.1) $\beta_t = \frac{\beta}{\gamma_{t-}}$, $M_t = \int_0^t \frac{\sigma_s}{\gamma_{s-}}\,dm_s$, $\langle m\rangle_t = K_t$, where $(\sigma_t)_{t\ge0}$ has the same properties as in the previous case 1.

Thus, we consider the following SDE

$$dz_t = -\frac{\beta}{\gamma_{t-}}\,z_{t-}\,dK_t + \frac{\sigma_t}{\gamma_{t-}}\,dm_t, \quad z_0.$$

Proposition 4.1. *Assume that the following conditions are satisfied: P-a.s.*

$$\text{(i)}\ \int_0^\infty \frac{dK_t}{\gamma_{t-}} = \infty, \quad \text{(ii)}\ \int_0^\infty \frac{dK_t}{\gamma_{t-}^2} < \infty, \quad \text{(iii)}\ \sum_{t\ge0}\left(\frac{\Delta K_t}{\gamma_{t-}}\right)^2 < \infty,$$

$$\text{(iv)}\ \lim_{t\to\infty}\widetilde{g}_t = \widetilde{g}, \quad \widetilde{g}\ge 0, \quad \widetilde{g}\ \textit{is a constant}, \qquad \text{(v)}\ 2\beta > \widetilde{g}.$$

Then the following assertions are true:

(1) $\langle L\rangle_\infty = \infty \quad P$-*a.s.*;

(2) $\lim\limits_{t\to\infty} \dfrac{\Gamma_t^2 \langle L\rangle_t^{-1}}{\gamma_t} = \dfrac{2\beta - \widetilde{g}}{\sigma^2} \quad P$-*a.s.*;

(3) $\lim\limits_{t\to\infty} \mathcal{L}(\gamma_t^{1/2} z_t) = \lim\limits_{t\to\infty} \mathcal{L}\biggl(\dfrac{\sigma}{\sqrt{2\beta - \widetilde{g}}}\, \dfrac{L_t}{\langle L\rangle_t^{1/2}}\biggr)$;

(4) *if* $(K_t)_{t\geq 0}$ *and* $(\gamma_t)_{t\geq 0}$ *are deterministic, then*

$$\lim_{t\to\infty} \mathcal{L}((1+K_t)^{1/2}\overline{z}_t) = \lim_{t\to\infty} \mathcal{L}\biggl(\frac{\sqrt{2}\sigma}{\sqrt{\beta(2\beta-\widetilde{g})}}\, \frac{\int_0^t (B_t - B_{s-})\, dL_s}{\widetilde{B}_t^{1/2}}\biggr).$$

Note that the primary process z has the rate of convergence $\gamma_t^{1/2}$, while averaging process $\overline{z}$ has the rate $(1+K_t)^{1/2}$ in all cases. For instance, if $\gamma_t = (1+K_t)$, then since $\widetilde{g}_t = 1$, we obtain

$$\lim_{t\to\infty} \mathcal{L}((1+K_t)^{1/2} z_t) = \lim_{t\to\infty} \mathcal{L}\biggl(\frac{\sigma}{\sqrt{2\beta-1}}\, \frac{L_t}{\langle L\rangle_t^{1/2}}\biggr)$$

and

$$\lim_{t\to\infty} \mathcal{L}((1+K_t)^{1/2}\overline{z}_t) = \lim_{t\to\infty} \mathcal{L}\biggl(\frac{\sigma}{\sqrt{\beta(2\beta-1)}}\, \frac{\int_0^t (B_t - B_{s-})\, dL_s}{\widetilde{B}_t^{1/2}}\biggr).$$

If $\gamma_t = (1+K_t)^r$, $\frac{1}{2} < r < 1$, then, since $\widetilde{g}_t \to 0$ as $t\to\infty$ (see (A.3)) we get

$$\lim_{t\to\infty} \mathcal{L}((1+K_t)^{r/2} z_t) = \lim_{t\to\infty} \mathcal{L}\biggl(\frac{\sigma}{\sqrt{2\beta}}\, \frac{L_t}{\langle L\rangle_t^{1/2}}\biggr)$$

and

$$\lim_{t\to\infty} \mathcal{L}((1+K_t)^{1/2}\overline{z}_t) = \lim_{t\to\infty} \mathcal{L}\biggl(\frac{\sigma}{\beta}\, \frac{\int_0^t (B_t - B_{s-})\, dL_s}{\widetilde{B}_t^{1/2}}\biggr).$$

Case 3. RM stochastic approximation algorithm with slowly varying gains. This case may be considered as a summurized example, where we demonstrate our methodology developed in [6], [7] and in the present paper in full capacity and details.

Consider the following SDE

$$dz_t = \frac{R(z_{t-})}{(1+K_{t-})^r}\, dK_t + \frac{\sigma_t}{(1+K_{t-})^r}\, dm_t, \quad z_0, \tag{4.3}$$

where $K = (K_t)_{t\geq 0}$ is an increasing predictable process, $K_\infty = \infty$ P-a.s., such that

$$\sum_t \biggl(\frac{\Delta K_t}{(1+K_{t-})^r}\biggr)^2 < \infty \quad P\text{-a.s.}, \tag{4.4}$$

the process $(\sigma_t)_{t\geq 0}$ is predictable with $0 < \sigma_t^2 \leq c$, $\lim\limits_{t\to\infty} \sigma_t^2 = \sigma^2$, $m = (m_t)_{t\geq 0} \in \mathcal{M}_{loc}^2$, $d\langle m\rangle_t = dK_t$, $R(u)$, $u\in R^1$, is some deterministic function satisfying the following condition:

(A) $\quad R(0) = 0, \quad R(u)u < 0, \quad u \neq 0,$

$$R(u) = -\beta u + v(u) \quad \text{with} \quad v(u) = O(u^2) \quad \text{when} \quad u\to 0, \tag{4.5}$$

β is a positive constant, $\frac{1}{2} < r < 1$.

In our notation

$$H(t,u) = \frac{R(u)}{(1+K_{t-})^r}, \quad M(t,u) \equiv M_t := \int_0^t \frac{\sigma_s}{(1+K_{s-})^r}\, dm_s.$$

We demonstrate our approach step by step.

Suppose, as usual, that equation (4.3) admits unique strong solution $z = (z_t)_{t\geq 0}$ on the whole time interval $[0,\infty)$.

Step 1. Convergence $z_t \to 0$ as $t \to \infty$ P-a.s. We refer to Proposition B.1 of Appendix B.

Introduce the following objects

$$a_t(u) := 2H(t,u)u = 2\,\frac{R(u)u}{(1+K_{t-})^r} < 0, \quad u \neq 0, \quad a_t(0) = 0,$$

$$b_t(u) := H^2(t,u)\,\Delta K_t = \frac{R^2(u)}{(1+K_{t-})^{2r}}\,\Delta K_t.$$

Proposition 4.2. *Let the following conditions be satisfied:*

(i) $$\frac{|R(u)|\,|u|}{(1+K_{t-})^r}\left[-2 + \frac{|R(u)|}{|u|}\,\frac{\Delta K_t}{(1+K_{t-})^r}\right]^+ \leq D_t(1+u^2), \quad D_t \geq 0, \quad D \circ K_\infty < \infty,$$

(ii) $$\inf_{\varepsilon \leq |u| < \frac{1}{\varepsilon}} \frac{|R(u)|\,|u|}{(1+K_{t-})^r}\Bigg\{2I_{\{\Delta K_t = 0\}} + \Bigg[-2 + \frac{|R(u)|}{|u|}\,\frac{\Delta K_t}{(1+K_{t-})^r}\Bigg]^- I_{\{\Delta K_t \neq 0\}}\Bigg\} \circ K_\infty = \infty, \quad P\text{-a.s.}$$

Then

$$z_t \to 0 \quad as \quad t \to \infty \quad P\text{-a.s.}$$

Remark 4. Suppose that the function $R(u)$ satisfies the following condition: there exist some positive constants G, $\widetilde{G}$, $G < \widetilde{G}$, such that

$$G|u| \leq |R(u)| < \widetilde{G}|u|.$$

Then conditions (i) and (ii) are satisfied. Indeed, if we put

$$D_t = \frac{\widetilde{G}}{(1+K_{t-})^r}\left[-2 + \widetilde{G}\,\frac{\Delta K_t}{(1+K_{t-})^r}\right]^+,$$

then since according to (4.4) $\frac{\Delta K_t}{(1+K_{t-})^r} \to 0$ as $t \to \infty$, we may conclude that $D = (D_t)_{t\geq 0}$ is equal to zero eventually. Thus (i) follows.

As for condition (ii) we have

$$\frac{|R(u)|\,|u|}{(1+K_{t-})^r}\left\{2I_{\{\Delta K_t=0\}} + \left[-2I_{\{\Delta K_t \neq 0\}} + \frac{|R(u)|}{|u|}\,\frac{\Delta K_t}{(1+K_{t-})^r}\right]^-\right\} \geq \frac{2G|u|^2}{(1+K_{t-})^r} \quad \text{eventually}.$$

Now (ii) follows from (see Proposition A.4 (1))

$$\int_0^\infty \frac{dK_t}{(1+K_{t-})^r} = \infty.$$

Step 2. Rate of convergence of $z = (z_t)_{t\geq 0}$**.** In this subsection we assume that $z_t \to 0$ as $t \to \infty$ P-a.s.

Since the process $\Gamma^2\langle L\rangle^{-1} = (\Gamma_t^2\langle L\rangle_t^{-1})_{t\geq 0}$ is equivalent to the process $((1+K_t)^r)_{t\geq 0}$ (see Proposition A.4, (4)) the natural choice of normalizing process γ_t is $\gamma_t = (1+K_t)^r$, $t \geq 0$.

Thus we have to prove that conditions imposed on function $R(u)$, $u \in R^1$ and process $K = (K_t)_{t\geq 0}$ together with $z_t \to 0$ as $t \to \infty$, P-a.s. ensure the following asymptotic property of $(z_t)_{t\geq 0}$: for all δ, $0 < \delta < \frac{\delta_0}{2}$, $0 < \delta_0 < 1$,

$$(1+K_t)^{r\delta} z_t \to 0 \quad \text{as} \quad t \to \infty \;\; P\text{-a.s.} \tag{4.6}$$

For this aim we refer to Proposition B.2 of Appendix B. First we focus our attention on condition (6), which allows us to obtain the value of δ_0 such that for all $\delta < \delta_0$ the following condition is satisfied:

$$\int_0^\infty (1+K_t)^{r\delta} \frac{1}{(1+K_{t-})^{2r}} < \infty, \quad P\text{-a.s.} \tag{4.7}$$

But (4.7) holds true if $2r - r\delta > 1$ (see (A.1) and (A.2)). Hence, $\delta < 2 - \frac{1}{r} := \delta_0$, $0 < \delta_0 < 1$. Condition (1) is trivially satisfied. Indeed, since $\Delta K_t/(1+K_{t-}) \to 0$ as $t \to \infty$,

$$\frac{\gamma_t}{\gamma_{t-}} = \left(1 + \frac{\Delta K_t}{(1+K_{t-})}\right)^r \to 1 \quad \text{as} \quad t \to \infty, \quad P\text{-a.s.}$$

As for condition (2), we have from (4.5)

$$\beta_t(z_{t-})\Delta K_t = \beta \frac{\Delta K_t}{(1+K_{t-})^r} + \frac{v(z_{t-})}{z_{t-}^2} z_{t-} I(z_{t-} \neq 0) \frac{\Delta K_t}{(1+K_{t-})^r} \to 0 \quad \text{as} \quad t \to \infty, \quad P\text{-a.s.}$$

Let us check condition (3). Note that

$$(1+K_t)^r = 1 + \int_0^t \widetilde{g}_s \, dK_s,$$

where

$$\widetilde{g}_t = r(1+K_t)^{r-1} I_{\{\Delta K_t = 0\}} + \frac{(1-K_t)^r - (1+K_{t-})^r}{\Delta K_t} I_{\{\Delta K_t \neq 0\}} \, .$$

Therefore

$$\int_0^\infty \left[\delta \frac{\widetilde{g}_t}{\gamma_t} - \beta_t(z_{t-})\right]^+ dK_t^c$$
$$= \int_0^\infty \left[r\delta \frac{(1+K_t)^{r-1}}{(1+K_t)^r} - \frac{\beta}{(1+K_t)^r} + \frac{v(z_{t-})}{z_{t-}^2} z_{t-} \frac{1}{(1+K_t)^r} I_{\{z_{t-}\neq 0\}}\right]^+ dK_t^c < \infty,$$

since the integrand in the last expression equals zero eventually. Indeed, the integrand is equal to

$$\frac{1}{(1+K_t)^r}\left[r\delta\,\frac{1}{(1+K_t)^{1-r}}-\beta+\frac{v(z_{t-})}{z_{t-}^2}\,z_{t-}I_{\{z_{t-}\neq 0\}}\right]^+=0 \quad \text{eventually}$$

because $(1+K_t)^{r-1}\to 0$, also $\frac{v(z_{t-})}{z_{t-}^2}\,z_{t-}\to 0$ as $t\to\infty$ (see (4.5)).

Further, since $\beta_t(z_{t-})\Delta K_t\to 0$ as $t\to\infty$, condition (5) is trivially satisfied.

Thus to finish the proof of (4.6) we have to check condition (4).

We have, for any δ, $0<\delta<1$,

$$1-\left(1-\frac{\Delta\gamma_t}{\gamma_t}\right)^{\delta}=1-\left(1-\frac{\Delta K_t}{1+K_t}\right)^{r\delta}\le r\delta\,\frac{\Delta K_t}{1+K_t}+(1-r\delta)\left(\frac{\Delta K_t}{1+K_t}\right)^2.$$

Therefore

$$\begin{aligned}\left[1-\beta_t(z_{t-})-\left(1-\frac{\Delta\gamma_t}{\gamma_t}\right)^{\delta}\right]^+ &\le \frac{\Delta K_t}{(1+K_{t-})^r}\Bigg[r\delta(1+K_{t-})^{r-1}\\ &\quad+(1-r\delta)\frac{\Delta K_t}{(1+K_{t-})^{2-r}}-\beta+\frac{v(z_{t-})}{z_{t-}^2}\,z_{t-}I_{\{z_{t-}\neq 0\}}\Bigg]^+=0 \quad \text{eventually.}\end{aligned}$$

Step 3. Asymptotic expansion for $z=(z_t)_{t\ge 0}$. In this subsection we assume that $\gamma_t^{\delta}z_t\to 0$ as $t\to\infty$ P-a.s., for all $0<\delta<\frac{\delta_0}{2}$, $0<\delta_0<1$. Recall that $\gamma_t=(1+K_t)^r$ and $\delta_0=2-\frac{1}{r}$. Assume that $r>\frac{4}{5}$.

According to Remark 6 to Proposition B.3 if we prove that

$$\int_0^\infty |\beta_t-\beta_t(z_{t-})|\gamma_{t-}^{\varepsilon}\,dK_t<\infty \quad P\text{-a.s.} \tag{4.8}$$

for some ε, $\frac{1}{2}-\frac{\delta_0}{2}<\varepsilon<\frac{1}{2}$, then the normed process $(\Gamma_t\langle L\rangle_t^{-1/2}z_t)_{t\ge 0}$ admits the following asymptotic expansion

$$\Gamma_t\langle L\rangle_t^{-1/2}z_t=\frac{z_0}{\langle L\rangle_t}+\frac{L_t}{\langle L\rangle_t^{1/2}}+R_t \tag{4.9}$$

with $R_t\to 0$ as $t\to\infty$ P-a.s.

Let is check condition (4.8). For each δ, $0<\delta<\frac{\delta_0}{2}$, we have

$$\begin{aligned}\int_0^\infty |\beta_t-\beta_t(z_{t-})|\gamma_{t-}^{\varepsilon}dK_t &= \int_0^\infty \frac{|v(z_{t-})|}{z_{t-}^2}\,|z_{t-}|(1+K_{t-})^{-r-r\varepsilon}dK_t\\ &\le const(\omega)\int_0^\infty (1+K_{t-})^{-r(1+\delta-\varepsilon)}\,dK_t.\end{aligned}$$

Therefore for given r, $\frac{1}{2}<r<1$, if there exists the pair (ε,δ) such that

$$\frac{1}{2r}-\frac{1}{2}<\varepsilon<\frac{1}{2}, \quad 0<\delta<1-\frac{1}{2r}, \quad r(1+\delta-\varepsilon)>1,$$

then condition (4.8) will be satisfied. But such a pair (ε,δ) exists only if $r>\frac{4}{5}$.

Note that from Eq. (4.9) follows that

$$\lim_{t\to\infty}\mathcal{L}(\Gamma_t\langle L\rangle_t^{-1/2}z_t)=\lim_{t\to\infty}\mathcal{L}\bigg(\frac{L_t}{\langle L\rangle_t^{1/2}}\bigg).$$

Thus, one can obtain the asymptotic distribution of the process $((1+K_t)^{1/2}z_t)_{t>0}$ using the appropriate form of Central Limit Theorem for locally square integrable martingales.

Namely, since $\lim\limits_{t\to\infty}\frac{\Gamma_t^2\langle L\rangle_t^{-1}}{(1+K_t)^r}=2\,\frac{\beta}{\sigma^2}$ (see Proposition A.4 (4)), then

$$\lim_{t\to\infty}\mathcal{L}\left((1+K_t)^{r/2}z_t\right)=\lim_{t\to\infty}\mathcal{L}\bigg(\sqrt{\frac{\sigma^2}{2\beta}}\,\frac{L_t}{\langle L\rangle_t^{1/2}}\bigg). \tag{4.10}$$

For instance, consider the case when all processes under consideration are continuous. Define for any sequence $(t_n)_{n\geq 1}$ of positive numbers with $\lim\limits_{n\to\infty}t_n=\infty$, the sequence $Y^n=(Y_u^n,\mathcal{F}_u^n)_{u\in[0,1]}\in\mathcal{M}_{loc}^2(P)$, where for $u\in[0,1]$

$$Y_u^n=\frac{L_{t_nu}}{\langle L\rangle_{t_n}^{1/2}}\,,\quad \mathcal{F}_u^n=\mathcal{F}_{t_nu}.$$

Then $\langle Y^n\rangle_1=1$ and hence $Y_1^n\xrightarrow{d}N(0,1)$ as $n\to\infty$. But, $Y_1^n=\frac{L_{t_n}}{\langle L\rangle_{t_n}^{1/2}}$ and therefore

$$\frac{L_t}{\langle L\rangle_t^{1/2}}\xrightarrow{d}N(0,1)\quad\text{as}\quad t\to\infty. \tag{4.11}$$

Finally, from (4.10) we obtain

$$(1+K_t)^{1/2}z_t\xrightarrow{d}N\left(0,\frac{\sigma^2}{2\beta}\right).$$

In the case when $K=(K_t)_{t\geq 0}$ and $(m_t)_{t\geq 0}$ are discontinuous the following Lindeberg condition (L_2) (see [12]) ensures (4.11):

$(L_2)\quad x^2I_{\{|x|>\delta\}}*\nu_1^n\xrightarrow{P}0,\quad \delta\in(0,1],$

where ν^n is the compensator of jump measure of $(Y_u^n)_{u\in[0,1]}$.

Note that (L_2) can be expressed in terms of the jump measure of $m=(m_t)_{\geq 0}$.

Step 4. Asymptotic properties of $\overline{z}=(\overline{z}_t)_{t\geq 0}$. First we study the asymptotic properties of $\overline{z}=(\overline{z}_t)_{t\geq 0}$ in linear case, when in (2.1) $\beta_t=\frac{\beta}{(1+K_{t-})^r}$. Then if we put in Proposition 4.1 $\gamma_t=(1+K_t)^r$, all conditions of this proposition are satisfies with $\widetilde{g}=0$ (see Eq. (4.4) and Proposition A.4). Further, by virtue of the Toeplitz lemma,

$$\lim_{t\to\infty}\frac{\mathcal{E}_t^{-1}}{1+K_t}=\lim_{t\to\infty}\frac{1+\int_0^t\Gamma_s^2\langle L\rangle_s^{-1}\beta(1+K_{s-})^{-r}dK_s}{1+K_t}=\lim_{t\to\infty}\frac{\beta\Gamma_t^2\langle L\rangle_t^{-1}}{(1+K_{t-})^r}=\frac{2\beta^2}{\sigma^2}\,,$$

since $\frac{1+K_t}{1+K_{t-}}\to 1$ as $t\to\infty$. Finally, from assertion (4) of Proposition 4.1 we obtain

$$\lim_{t\to\infty}\mathcal{L}\left((1+K_t)^{1/2}\overline{z}_t\right)=\lim_{t\to\infty}\mathcal{L}\bigg(\frac{\sigma}{\beta}\,\frac{\int_0^t(B_t-B_{s-})dL_s}{\widetilde{B}_t^{1/2}}\bigg).$$

Let us return to the general case. Assume now $r > \frac{5}{6}$ (in Step 3 $r > \frac{4}{5}$). First note that in (3.1) $r_t^2 \equiv 0$.

Now because the difference between the representations of (2.5) and (3.3) for linear and general cases, respectively consists only in the remainder term r_t, to obtain the asymptotic distribution of $\overline{z}$ in general case it is enough to prove

$$r_t^1 \to 0 \quad \text{as} \quad t \to \infty \quad P\text{-a.s.} \tag{4.12}$$

For this aim let us refer to Corollary 3.2, (2), and note that in the considered case

$$v(t,u) = \frac{v(u)}{(1+K_{t-})^r},$$
$$|v_t| = \lim_{u\to 0} \frac{|v(t,u)|}{u^2} = \lim_{u\to 0} \frac{|v(u)|}{u^2} \frac{1}{(1+K_{t-})^r} \le const \, \frac{1}{(1+K_{t-})^r}\,.$$

Thus to prove (4.12) we have to check that condition (ii)$''$ of Corollary 3.2 is satisfied.

But condition (ii)$''$ will be satisfied if there exists δ, $0 < \delta < \frac{\delta_0}{2}$, $\frac{\delta_0}{2} = 1 - \frac{1}{2r}$, such that

$$\int_0^\infty (1+K_{t-})^{1/2}(1+K_{t-})^{-r}(1+K_{t-})^{-2r\delta}\, dK_t < \infty.$$

Such a δ exists only if $r > \frac{5}{6}$.

Thus we obtain

$$\lim_{t\to\infty} \mathcal{L}((1+K_t)^{-1/2}\overline{z}_t) = \lim_{t\to\infty} \mathcal{L}\biggl(\frac{\sigma^2}{\beta^2} \frac{\int_0^t (B_t - B_{s-})dL_s}{\widetilde{B}_t^{1/2}}\biggr).$$

5. Proof of Results

Proof of Proposition 2.1: Applying the Itô formula to the process $\widetilde{B}_t$ after simple calculations it is easy to check that

$$\widetilde{B}_t = 2\,(\langle L\rangle \circ B)\circ B_t - (\langle L\rangle \Delta B)\circ B_t. \tag{5.1}$$

From conditions (2) and (3) applying the Toeplitz lemma we obtain

$$\lim_{t\to\infty} \frac{\widetilde{B}_t}{(\langle L\rangle \circ B)\circ B_t} = 2 - \lim_{t\to\infty} \frac{(\langle L\rangle \Delta B)\circ B_t}{(\langle L\rangle \circ B)\circ B_t} = 2 - \lim_{t\to\infty} \frac{\frac{\langle L\rangle \Delta B}{\langle L\rangle \circ B}(\langle L\rangle \circ B)\circ B_t}{(\langle L\rangle \circ B)\circ B_t}$$
$$= 2 - \lim_{t\to\infty} \frac{\langle L\rangle_t \Delta B_t}{\langle L\rangle \circ B_t} = 2 - c.$$

Further,

$$\lim_{t\to\infty} \frac{B_t^2}{\widetilde{B}_t} = \lim_{t\to\infty} \frac{B_t^2}{(\langle L\rangle \circ B)\circ B_t} \frac{(\langle L\rangle \circ B)\circ B_t}{\widetilde{B}_t} = \frac{1}{2-c} \lim_{t\to\infty} \frac{B_t^2}{(\langle L\rangle \circ B)\circ B_t}\,.$$

Therefore, it remains to show that

$$\lim_{t\to\infty} \frac{B_t^2}{(\langle L\rangle \circ B)\circ B_t} = 0. \tag{5.2}$$

We have

$$\frac{B_t^2}{(\langle L\rangle\circ B)\circ B_t}=\frac{\int_0^t(B_{s-}+B_s)\,dB_s}{(\langle L\rangle\circ B)\circ B_t}\le\frac{2\int_0^t B_s\,dB_s}{(\langle L\rangle\circ B)\circ B_t}$$
$$=\frac{2\int_0^t B_s(\langle L\rangle\circ B)_s^{-1}(\langle L\rangle\circ B)_s\,dB_s}{(\langle L\rangle\circ B)\circ B_t}\,.$$

Now, applying once again the Toeplitz lemma to the last term of this inequality and using condition (2) we obtain

$$\lim_{t\to\infty}\frac{2\int_0^t B_s\,dB_s}{(\langle L\rangle\circ B)\circ B_t}=2\,\overline{\lim_{t\to\infty}}\,\frac{B_t}{\langle L\rangle\circ B_t}=2\lim_{t\to\infty}\frac{\int_0^t\langle L\rangle_s^{-1}\langle L\rangle_s\,dB_s}{\langle L\rangle\circ B_t}=2\lim_{t\to\infty}\frac{1}{\langle L\rangle_t}=0\,.$$

Therefore

$$\overline{\lim_{t\to\infty}}\,\frac{B_t^2}{(\langle L\rangle\circ B)\circ B_t}\le 2\lim_{t\to\infty}\frac{2\int_0^t B_s\,dB_s}{(\langle L\rangle\circ B)\circ B_t}=0$$

from which (5.2) follows. Eq. (2.7) directly follows from Eq. (2.6). □

Proof of Proposition 2.2: First, show that conditions (1) and (2) of Proposition 2.1 are satisfied. We have

$$\langle L\rangle\circ B_t=\langle L\rangle\Gamma^{-1}\circ\mathcal{E}_t^{-1}=\int_0^t\langle L_s\rangle\Gamma_s^{-1}\Gamma_s^2\langle L_s\rangle^{-1}\beta_s\,dK_s$$
$$=\int_0^t d\Gamma_s=\Gamma_t-1\to\infty\ \text{ as }\ t\to\infty.$$

Thus, condition (1) is satisfied.

Further, since $\mathcal{E}_\infty^{-1}=\infty$ and

$$\frac{(\langle L\rangle\circ B)\circ B_t}{\mathcal{E}_t^{-1}}=\frac{(\langle L\rangle\circ B)\Gamma^{-1}\circ\mathcal{E}_t^{-1}}{\mathcal{E}_t^{-1}}\,,$$

one can apply the Toeplitz lemma to obtain

$$\lim_{t\to\infty}\frac{(\langle L\rangle\circ B)\circ B_t}{\mathcal{E}_t^{-1}}=\lim_{t\to\infty}\frac{\int_0^t d\Gamma_s}{\Gamma_t}=1,$$

from which condition (2) follows.

Condition (3) immediately follows from Assumption 1. Indeed,

$$\frac{L_t\Delta B_t}{\langle L\rangle\circ B_t}=\frac{\Delta\Gamma_t}{\Gamma_t-1}=\beta_t\Delta K_t\,\frac{1}{(1-\frac{1}{\Gamma_t})}\to m\quad\text{as}\quad t\to\infty.$$

To prove (2.9) rewrite $\widetilde{B}_t$ in the following form

$$\widetilde{B}_t=(2\langle L\rangle\circ B-\langle L\rangle\Delta B)\circ B_t=(2\langle L\rangle\circ B-\langle L\rangle\Delta B)\,\Gamma^{-1}\circ\mathcal{E}_t^{-1}.$$

Applying the Toeplitz lemma we have

$$\lim_{t\to\infty}\frac{\widetilde{B}_t}{\mathcal{E}_t^{-1}}=\lim_{t\to\infty}\frac{2\langle L\rangle\circ B_t-\langle L\rangle_t\Delta B_t}{\Gamma_t}=\lim_{t\to\infty}\frac{2(\langle L\rangle\Gamma^{-1})\circ\mathcal{E}^{-1}-\langle L\rangle_t\Gamma_t^{-1}\Delta\mathcal{E}_t^{-1}}{\Gamma_t}$$
$$=\lim_{t\to\infty}\frac{2\langle L\rangle\Gamma^{-2}\beta^{-1}g\mathcal{E}^{-1}\circ\Gamma_t-\langle L\rangle_t\Gamma_t^{-2}\beta_t^{-1}g_t\mathcal{E}_t^{-1}\Delta\Gamma_t}{\Gamma_t}\,.$$

Then, since $\Gamma_\infty = \infty$, applying the Toeplitz lemma again we obtain

$$\lim_{t\to\infty} \frac{\widetilde{B}_t}{\mathcal{E}_t^{-1}} = 2 \lim_{t\to\infty} \langle L\rangle_t \Gamma_t^{-2}\beta_t^{-1} g_t \mathcal{E}_t^{-1} - \lim_{t\to\infty} \langle L\rangle_t \Gamma_t^{-2}\beta_t^{-1} g_t \mathcal{E}_t^{-1} \frac{\Delta\Gamma_t}{\Gamma_t} \tag{5.3}$$

if these limits exist. Now note that

$$g_t \mathcal{E}_t^{-1} = \Gamma_t^2 \langle L\rangle_t^{-1} \beta_t, \tag{5.4}$$

therefore from (5.3) we get (2.9). Now Eq. (2.10) is a direct consequence of Eq. (2.7). □

Proof of Lemma 3.1: Since

$$d\mathcal{E}_t^{-1} = (\mathcal{E}_t^{-1/2} + \mathcal{E}_{t-}^{-1/2})\, d\mathcal{E}_t^{-1/2},$$

we can write $\widetilde{r}_t^i$ as follows:

$$\widetilde{r}_t^1 = \frac{1}{\mathcal{E}_t^{-1/2}} \int_0^t R_s^1 \Gamma_s^{-1} (\mathcal{E}_s^{-1/2} + \mathcal{E}_{s-}^{-1/2})\, d\mathcal{E}_s^{-1/2}.$$

Let us show first that

$$\lim_{t\to\infty} \frac{|R_t^1|(\mathcal{E}_t^{-1/2} + \mathcal{E}_{t-}^{-1/2})}{\Gamma_t} = 0.$$

Indeed, using (1.3) we get

$$\begin{aligned} \frac{|R_t^1|(\mathcal{E}_t^{-1/2} + \mathcal{E}_{t-}^{-1/2})}{\Gamma_t} &\le 2 \frac{|R_t^1|}{\Gamma_t \mathcal{E}_t^{1/2}} \\ &= \frac{\int_0^t (\Gamma_s \mathcal{E}_s^{-1/2}) \mathcal{E}_s^{-1/2} |\beta_s - \beta_s(z_{s-})|\, |z_{s-}|\, dK_s}{\Gamma_t \mathcal{E}_t^{1/2}}. \end{aligned} \tag{5.5}$$

Show that the limit of the last expression equals zero. For this aim establish first that

$$\lim_{t\to\infty} \Gamma_t \mathcal{E}_t^{1/2} = \infty. \tag{5.6}$$

We have

$$\Gamma_t^{-2}\mathcal{E}_t^{-1} = \frac{1 + \int_0^t \Gamma_s^2 \langle L\rangle_s^{-1} \beta_s\, dK_s}{\Gamma_t^2} = \frac{1 + \int_0^t \Gamma_s^2 \langle L\rangle_s^{-1} \Gamma_s^{-2} \left(2 - \frac{\Delta\Gamma_s}{\Gamma_s}\right)^{-1} d\Gamma_s^2}{\Gamma_t^2}. \tag{5.7}$$

Applying the Toeplitz lemma to the last term of Eq. (5.7), we get

$$\lim_{t\to\infty} \Gamma_t^{-2}\mathcal{E}_t^{-1} = \lim_{t\to\infty} \langle L\rangle_t^{-1} \left(2 - \frac{\Delta\Gamma_t}{\Gamma_t}\right)^{-1} = 0.$$

Further, show that condition (i) implies that $(\Gamma_t \mathcal{E}_t^{1/2})_{t\ge0}$ is an increasing process eventually.

Using the Itô formula to the process $(\Gamma_t^2 \mathcal{E}_t)_{t\ge0}$ we have

$$\begin{aligned} d(\Gamma_t^2 \mathcal{E}_t) &= d\left(\frac{\Gamma_t^2}{\mathcal{E}_t^{-1}}\right) = \frac{1}{\mathcal{E}_{t-}^{-1}} (\Gamma_t + \Gamma_{t-})\, d\Gamma_t - \Gamma_t^2 \frac{1}{\mathcal{E}_t^{-1}\mathcal{E}_{t-}^{-1}}\, d(\mathcal{E}_t^{-1}) \\ &= \frac{1}{\mathcal{E}_{t-}^{-1}} \Gamma_t^2 \left[\left(2 - \frac{\Delta\Gamma_t}{\Gamma_t}\right) - \frac{1}{\mathcal{E}_t^{-1}} \Gamma_t^2 \langle L\rangle_t^{-1}\right] \beta_t\, dK_t > 0 \quad \text{eventually.} \end{aligned}$$

Hence $(\Gamma_t^2 \mathcal{E}_t^{1/2})_{t\ge 0}$ is an increasing process eventually. This fact together with (5.6) allows one to apply the Kronecker lemma to observe that condition (ii) ensures that the limit of the last term of Eq. (5.5) equals zero. Now assertion of the lemma follows from the Toeplitz lemma. □

Proof of Corollary 3.2: (1) As $\gamma_t^\delta z_t \to 0$ as $t \to \infty$, P-a.s. for all δ, $0 < \delta < \frac{\delta_0}{2}$, then $|\gamma_t^\delta z_t(\omega)|$ is bounded by some constant (depending on ω) for all δ, $0 < \delta < \frac{\delta_0}{2}$, P-a.s. Therefore, for all δ, $0 < \delta < \frac{\delta_0}{2}$,

$$\int_0^\infty \mathcal{E}_t^{-1/2} |\beta_t - \beta_t(z_{t-})|\, |z_{t-}|\, dK_t = \int_0^\infty \mathcal{E}_t^{-1/2} |\beta_t - \beta_t(z_{t-})| \gamma_{t-}^{-\delta} |\gamma_{t-}^\delta z_{t-}|\, dK_t$$
$$\le const(\omega) \int_0^\infty \mathcal{E}_t^{-1/2} |\beta_t - \beta_t(z_{t-})| \gamma_{t-}^{-\delta}\, dK_t \quad P\text{-a.s.}$$

Thus, (ii)$'\Rightarrow$(ii).

(2)The arguments are clear. □

Proof of Lemma 3.3: Show that

$$\widetilde{r}_t^2 = \frac{1}{\mathcal{E}_t^{-1/2}} \int_0^t R_s^2\, dB_s = \frac{1}{\mathcal{E}_t^{-1/2}} \int_0^t N_s\, dB_s \xrightarrow{P} 0 \quad \text{as} \quad t \to \infty.$$

Since $B = (B_t)_{t\ge 0}$ is deterministic, we have

$$\widetilde{r}_t^2 = \frac{1}{\mathcal{E}_t^{-1/2}} \int_0^t N_s\, dB_s = \frac{1}{\mathcal{E}_t^{-1/2}} \int_0^t (B_t - B_{s-})\, dN_s.$$

Further, for any sequence $(t_n)_{n\ge 1}$, $t_n \to \infty$ as $n \to \infty$, let us consider a sequence of martingales $Y^n = (Y_u^n, \mathcal{F}_u^n)_{u\in[0,1]} \in \mathcal{M}_{loc}^2(P)$, $n \ge 1$, where $\mathcal{F}_u^n = \mathcal{F}_{t_n u}$,

$$Y_u^n = \frac{1}{\mathcal{E}_{t_n}^{-1/2}} \int_0^{t_n u} (B_{t_n} - B_{s-})\, dN_s.$$

Now, if we show that $\langle Y^n \rangle_1 \xrightarrow{P} 0$ as $n \to \infty$, then from the well-known fact that $\langle Y^n \rangle_1 \xrightarrow{P} 0 \Rightarrow Y_1^n \xrightarrow{P} 0$ (see, e.g., [12]) we will get $\widetilde{r}_{t_n}^2 \to 0$ as $n \to \infty$, and hence $\widetilde{r}_t^2 \to 0$ as $t \to \infty$.

Thus we have to show that $\langle Y^n \rangle_1 \xrightarrow{P} 0$ as $n \to \infty$. Since

$$\langle Y^n \rangle_1 = \frac{1}{\mathcal{E}_{t_n}^{-1}} \int_0^{t_n} (B_{t_n} - B_{s-})^2 d\langle N \rangle_s,$$

we have to show

$$\frac{1}{\mathcal{E}_t^{-1}} \int_0^t (B_t - B_{s-})^2 d\langle N \rangle_s \xrightarrow{P} 0 \quad \text{as} \quad t \to \infty.$$

We use (5.1) with $\langle N \rangle$ instead of $\langle L \rangle$ to obtain

$$\int_0^t (B_t - B_{s-})^2 d\langle N \rangle_s \le 2(\langle N \rangle \circ B) \circ B_t.$$

Hence, it is enough to prove that (3.4) implies

$$\frac{(\langle N \rangle \circ B) \circ B_t}{\mathcal{E}_t^{-1}} \to 0 \quad \text{as} \quad t \to \infty.$$

Since $\mathcal{E}_\infty^{-1} = \infty$ and

$$\frac{(\langle N\rangle \circ B) \circ B_t}{\mathcal{E}_t^{-1}} = \frac{\int_0^t \int_0^s \langle N\rangle_u dB_u \, \Gamma_s^{-1} d\mathcal{E}_s^{-1}}{\mathcal{E}_t^{-1}}\,,$$

applying the Toeplitz lemma yields

$$\lim_{t\to\infty} \frac{(\langle N\rangle \circ B) \circ B_t}{\mathcal{E}_t^{-1}} = \lim_{t\to\infty} \frac{1}{\Gamma_t} \int_0^t \langle N\rangle_s \, dB_s$$
$$= \lim_{t\to\infty} \frac{1}{\Gamma_t} \int_0^t \langle N\rangle_s \Gamma_s^{-1} \, d\mathcal{E}_s^{-1} = \lim_{t\to\infty} \frac{1}{\Gamma_t} \int_0^t \langle N\rangle_s \langle L\rangle_s^{-1} \, d\Gamma_s = \lim_{t\to\infty} \frac{\langle N\rangle_t}{\langle L\rangle_t} = 0,$$

which completes the proof. □

Proof of Proposition 4.1 : (1) By the definition

$$\langle L\rangle_t = 1 + \int_0^t \Gamma_s^2 \ell_s^2 \, dK_s = 1 + \int_0^t \Gamma_s^2 \frac{\sigma_s^2}{\gamma_{s-}^2} \, dK_s$$
$$= 1 + \int_0^t \Gamma_s^2 \gamma_{s-}^{-1} \sigma_s^2 \frac{dK_s}{\gamma_{s-}} \geq 1 + \int_0^t \Gamma_s^2 \gamma_s^{-1} \sigma_s^2 \frac{dK_s}{\gamma_{s-}}\,.$$

Now consider the process $(\Gamma_t^{-2}\gamma_t)_{t\geq 0}$. Using the Yor formula $\mathcal{E}_t(X)\mathcal{E}_t(Y) = \mathcal{E}_t(X + Y + [X,Y])$, we have

$$\Gamma_t^{-2}\gamma_t = \mathcal{E}_t^2\left(-\frac{\beta}{\gamma_-} \circ K\right) \mathcal{E}_t(\lambda \circ K)$$
$$= \mathcal{E}_t\Bigg(\int_0^{\cdot} [2\beta - \widetilde{g}_s] \frac{dK_s}{\gamma_{s-}} + \sum_{s\leq \cdot} \left[\beta_s^2 - 2\beta\widetilde{g}_s + \beta^2 \widetilde{g}_s \frac{\Delta K_s}{\gamma_{s-}}\right] \left(\frac{\Delta K_s}{\gamma_{s-}}\right)^2\Bigg). \tag{5.8}$$

Further, note that by virtue of conditions (i), (iii), (iv) and (v)

$$\int_0^\infty [2\beta - \widetilde{g}_s] \frac{dK_s}{\gamma_{s-}} = \infty \quad P\text{-a.s.}$$

and

$$\sum_{t\geq 0} \left[\beta^2 - 2\beta\widetilde{g}_t + \beta^2 \widetilde{g}_t \frac{\Delta K_t}{\gamma_{t-}}\right] \left(\frac{\Delta K_t}{\gamma_{t-}}\right)^2$$
$$\leq \sum_{t\geq 0} \beta^2 \left[1 + \widetilde{g}_t \frac{\Delta K_t}{\gamma_{t-}}\right] \left(\frac{\Delta K_t}{\gamma_{t-}}\right)^2 < \infty \quad P\text{-a.s.}$$

These relations together with (5.8) imply

$$\lim_{t\to\infty} \Gamma_t^2 \gamma_t^{-1} = \infty$$

which in turn provides $\langle L\rangle_\infty = \infty$ P-a.s.

(2) Let us note that conditions (a), (b) and (c) of Proposition A.3 are satisfied with $c_1 = 0$ (see (iii)), $c_2 = \frac{\beta}{\sigma^2}$, $c_3 = \frac{\widetilde{g}}{\sigma^2}$ (see (iv)).

Assertion (3) follows from (2.4).

On the other hand, applying the Toeplitz lemma and assertion (2) we obtain

$$\lim_{t\to\infty} \frac{\mathcal{E}_t^{-1}}{1+K_t} = \lim_{t\to\infty} \Gamma_t^2 \langle L\rangle_t^{-1} \frac{\beta}{\gamma_{t-}} = \frac{\beta(2\beta - \widetilde{g})}{\sigma^2}\,.$$

Now the last assertion (4) follows from (2.9) with $m = 0$ since $\frac{\Delta\Gamma_t}{\Gamma_t} = \beta_t \Delta K_t = \beta\,\frac{\Delta K_t}{\gamma_{t-}} \to 0$. □

Proof of Proposition 4.2: Conditions (i) and (ii) of this proposition are the same as the corresponding conditions in Proposition B.1 written for the considered case. Condition (B) of Proposition B.1 follow from Eq. (4.4). □

REFERENCES

[1] Han-Fu Chen, *Stochastic Approximation and its Applications,* Nonconvex Optimization and its Applications, 64. Kluwer Academic Publishers, Dordrecht, 2002.

[2] C. Dellacherie, *Capacités et Processus Stochastiques,* (French) Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 67. Springer-Verlag, Berlin-New York, 1972.

[3] B. Delyon and A. B. Juditsky, *Stochastic optimization with averaging of trajectories,* Stochastics Stochastics Rep. 39 (1992), pp. 107–118.

[4] H. J. Kushner and G. G. Yin, *Stochastic approximation and recursive algorithms and applications,* Second edition. Applications of Mathematics (New York), 35. Stochastic Modelling and Applied Probability. Springer-Verlag, New York, 2003.

[5] Tze Leung Lai, *Stochastic approximation: invited paper,* Dedicated to the memory of Herbert E. Robbins. Ann. Statist. 31 (2003), no. 2, pp. 391–406.

[6] N. Lazrieva, T. Sharia and T. Toronjadze, *The Robbins–Monro type SDEs. I. Convergence of solutions,* Stochastic Stochastics Rep. 61 (1997), pp. 67–87.

[7] N. Lazrieva, T. Sharia and T. Toronjadze, *The Robbins–Monro type SDEs. II. Asymptotic behaviour of solutions,* Stochastic Stochastics Rep. 75 (2003), no. 5, pp. 153–180.

[8] A. Le Breton, *About Gaussian schemes in stochastic approximation,* Stochastic Process. Appl. 50 (1992), pp. 101–116.

[9] A. Le Breton, *About the averaging approach in Gaussian schemes for stochastic approximation,* Math. Methods Statist. 2 (1993), pp. 295–315.

[10] A. Le Breton and A. A. Novikov, *Averaging for estimating covariances in stochastic approximation,* Math. Methods Statist. 3 (1994), pp. 244–266.

[11] A. Le Breton and A. Novikov, *S*ome results about averaging in stochastic approximation, Second International Conference on Mathematical Statistics (Smolenice Castle, 1994). Metrika 42 (1995), no. 3-4, pp. 153–171.

[12] R. S. Lipster and A. N. Shiryayev, *Theory of Martingales,* Kluwer, Dordrecht, 1986 (original work published in Nauka, Moscow, 1986).

[13] B. T. Polyak, *New method of stochastic approximation type,* Automate Remote Control. 51 (1990), pp. 937–946.

[14] B. T. Polyak and A. B. Juditsky, *Acceleration of stochastic approximation by averaging,* SIAM J. Cont. Optim. 30 (1992), pp. 838–855.

[15] D. Ruppert, *Efficient estimations from a slowly convergent Robbins-Monro process,* Tech. Rep. 781, School of Oper. Res. and Indust. Eng., Cornell Univ., 1988.

[16] E. Valkiela and A.V. Melnikov, *Martingale models of stochastic approximation and their convergence,* (Russian) Probab. Theory Appl. 44 (1999), no. 2, pp. 278–312.

[17] M. T. Wasan, *Stochastic Approximation,* Series: Cambridge Tracts in Mathematics, No. 58, June 2004.

[18] G. Yin, *On extensions of Polyak's averaging approach to stochastic approximation,* Stochastics Stochastics Rep. 36 (1991), pp. 245–264.

[19] G. Yin, Q. Zhang, F. Liu, R. H. Liu and Y. Cheng, *Stock liquidation via stochastic approximation using NASDAQ daily and intra-day data,* Math. Finance 16 (2006), no. 1, pp. 217–236.

APPENDIX A

Here we collect some necessary definitions and technical results concerning main objects of this paper.

Definition A.1. We say that two processes $\xi = (\xi_t)_{t\geq 0}$ and $\eta = (\eta_t)_{t\geq 0}$ are equivalent and write $\xi \sim \eta$ if there exists some constant c, $0 < c < \infty$, such that

$$\lim_{t\to\infty} \frac{\xi_t}{\eta_t} = c \quad P\text{-a.s.}$$

Definition A.2. We say that the process $\xi = (\xi_t)_{t\geq 0}$ has some property eventually if for every ω in a set Ω_0 of P probability 1, the trajectory $(\xi_t(\omega))_{t\geq 0}$ of the process has this property on the set $[t_0(\omega), \infty)$ for some $t_0(\omega) < \infty$.

Toeplitz's lemma: *Let $L = (L_t)_{t\geq 0}$ be a predictable increasing process with $L_\infty = \infty$ P-a.s., $Y = (Y_t)_{t\geq 0} \in D$ be some process such that $Y \circ L_t < \infty$ P-a.s. for all $t \geq 0$ and $\lim\limits_{t\to\infty} Y_\infty < \infty$ P-a.s. Then*

$$\lim_{t\to\infty} \frac{\int_0^t Y_s\, dL_s}{1 + L_t} = \lim_{t\to\infty} Y_t = Y_\infty \quad P\text{-a.s.}$$

Assume that the following objects are given: $K = (K_t)_{t\geq 0}$ – predictable increasing process with $K_\infty = \infty$ P-a.s. $\Gamma = (\Gamma_t)_{t\geq 0}$, $\Gamma_t = \mathcal{E}_t^{-1}(-\beta \circ K)$, where $(\beta_t)_{t\geq 0}$ is a positive predictable process such that $\beta_t \Delta K_t < 1$, $\beta \circ K_t < \infty$ for all $t \geq 0$ and $\beta \circ K_\infty = \infty$ P-a.s. $L = (L_t)_{t\geq 0}$, $L_t = \int_0^t \Gamma_s\, dm_s$, where $(m_t)_{t\geq 0} \in \mathcal{M}^2_{loc}(P)$, $d\langle m\rangle_t = \ell_t^2\, dK_t$, $\ell^2 \circ K_\infty < \infty$ P-a.s.

Let $\gamma = (\gamma_t)_{t\geq 0}$ be some increasing predictable process such that $\gamma_t = 1 + \widetilde{g} \circ K_t$, $\widetilde{g}_t \geq 0$, $\widetilde{g} \circ K_t < \infty$ for all $t \geq 0$ and $\widetilde{g} \circ K_\infty = \infty$ P-a.s. Obviously process $(\gamma_t)_{t\geq 0}$ can be written as follows: $\gamma_t = \mathcal{E}_t(\lambda \circ K)$, where $\lambda_t = \widetilde{g}_t / \gamma_t$.

Everywhere below we assume that $\langle L\rangle_\infty = \infty$ P-a.s.

Proposition A.3. *Let the following conditions be satisfied: P-a.s.*

$$\text{(a)} \quad \beta_t \Delta K_t \to c_1 \quad \textit{as} \quad t \to \infty, \quad 0 \leq c_1 < 1,$$

$$\text{(b)} \quad \gamma_{t-}^{-1} \frac{\beta_t}{\ell_t^2} \to c_2 \quad \textit{as} \quad t \to \infty, \quad 0 < c_2 < \infty,$$

$$\text{(c)} \quad \gamma_t^{-1} \frac{\lambda_t}{\ell_t^2} \to c_3 \quad \textit{as} \quad t \to \infty, \quad 0 \leq c_3 < c_2(2 - c_1).$$

Then

$$\lim_{t\to\infty} \frac{\Gamma_t^2 \langle L\rangle_t^{-1}}{\gamma_t} = c_2(2 - c_1) - c_3 \quad P\text{-a.s.}$$

Proof. Applying the Itô formula to the process $(\Gamma_t^2 \gamma_t^{-1})_{t\geq 0}$ and having in mind that $d\Gamma_t^2 = (\Gamma_t + \Gamma_{t-})\, d\Gamma_t$, $d\Gamma_t = \Gamma_t \beta_t\, dK_t$, $d\langle L\rangle_t = \Gamma_t^2 \ell_t^2\, dK_t$ and $d\gamma_t^{-1} = -\gamma_t^{-1} \lambda_t\, dK_t$ we obtain

$$\frac{\Gamma_t^2 \langle L\rangle_t^{-1}}{\gamma_t} = \frac{\Gamma_t^2 \gamma_t^{-1}}{\langle L\rangle_t} = \frac{\int_0^t \gamma_{s-}^{-1} \left(2 - \frac{\Delta\Gamma_s}{\Gamma_s}\right) \frac{\beta_s}{\ell_s^2}\, d\langle L\rangle_s - \int_0^t \gamma_s^{-1} \ell_s^{-2} \lambda_s d\langle L\rangle_s}{\langle L\rangle_t}.$$

Since $\langle L\rangle_\infty = \infty$ using the Toeplitz lemma and conditions (a), (b) and (c) we have

$$\begin{aligned}\lim_{t\to\infty} \frac{\Gamma_t^2 \langle L\rangle_t^{-1}}{\gamma_t} &= \lim_{t\to\infty} \frac{\Gamma_t^2 \gamma_t^{-1}}{\langle L\rangle_t} \\ &= \lim_{t\to\infty} \gamma_{t-}^{-1}\left(2 - \frac{\Delta\Gamma_t}{\Gamma_t}\right)\frac{\beta_t}{\ell_t^2} - \lim_{t\to\infty} \gamma_t^{-1}\ell_t^{-2}\lambda_t = c_2(2-c_1) - c_3. \quad \square\end{aligned}$$

Remark 5. Under condition (b) the condition (c) is equivalent to

(c)$'$ $\quad \gamma_t^{-1}\widetilde{g}_t/\beta_t \to \frac{c_3}{c_2}$ as $t\to\infty$, P-a.s.

This fact immediately follows from the relation

$$\gamma_t^{-1}\lambda_t/\ell_t^2 = (\gamma_t^{-1}\,\widetilde{g}_t/\beta_t)(\gamma_{t-}^{-1}\,\beta_t/\ell_t^2).$$

Proposition A.4. *Let $K = (K_t)$ be an increasing predictable process, $K_\infty = \infty$ P-a.s., satisfying* (4.4) *with $\frac{1}{2} < r < 1$.*

Put $\beta_t = \frac{\beta}{(1+K_-)^r}$, $\ell_t^2 = \frac{\sigma_t^2}{(1+K_{t-})^{2r}}$, $\lim_{t\to\infty} \sigma_t^2 = \sigma^2$.

In this case the following assertions hold true: P-a.s.

(1) $(1+K_-)^{-r}\circ K_\infty = \infty$*; moreover,* $\lim_{t\to\infty} \frac{(1+K_-)^{-r}\circ K_t}{(1+K_t)^{1-r}} = \frac{1}{1-r}$*;*

(2) $(1+K_-)^{2r}\circ K_\infty < \infty$*;*

(3) $\langle L\rangle_\infty = \infty$*;*

(4) $\lim_{t\to\infty} \frac{\Gamma_t^2\langle L\rangle_t^{-1}}{(1+K_t)^r} = 2\,\frac{\beta}{\sigma^2}$.

Proof. Assertion (1) immediately follows from the following time change formula (see, e.g., [2])

$$\int_0^\infty \frac{dK_t}{(1+K_{t-})^r} = \int_0^{K_\infty} \frac{dt}{(1+K_{c(t)-})^r} \ge \int_0^{K_\infty} \frac{dt}{(1+t)^r} = \infty,$$

where $c(t) = \inf\{s>0 : K_s > t\}$. Note that $K_{c(t)-} \le t$.

(2) We will prove more general result. Let $\alpha > 0$. Then the condition

$$\sum_t \frac{(\Delta K_t)^2}{(1+K_{t-})^{2+\alpha}} < \infty \qquad \text{(A.1)}$$

is sufficient for

$$(1+K_-)^{-(1+\alpha)}\circ K_\infty < \infty. \qquad \text{(A.2)}$$

Using the Itô formula to the process $(1+K_t)^{-\alpha}$ one easily obtain

$$\begin{aligned}(1+K_-)^{-(1+\alpha)}\circ K_\infty &= \frac{1}{\alpha}\left(1-(1+K_t)^{-\alpha}\right) \\ &+ \frac{1}{\alpha}\sum_t \Big[(1+K_t)^{-\alpha} - (1+K_{t-})^{-\alpha} + \alpha(1+K_{t-})^{-(1+\alpha)}\Delta K_t\Big].\end{aligned}$$

But

$$\sum_t \left|(1+K_t)^{-\alpha} - (1+K_{t-})^{-\alpha} + \alpha(1+K_{t-})^{-(1+\alpha)}\Delta K_t\right|$$
$$\leq \alpha(1+\alpha)\sum_t \frac{(\Delta K_t)^2}{(1+K_{t-})^{2+\alpha}} < \infty.$$

In the case when $1+\alpha = 2r$, for the convergence $(1+K_-)^{-2r} \circ K_\infty < \infty$ it is sufficient the following condition

$$\sum_t \frac{(\Delta K_t)^2}{(1+K_{t-})^{2r+1}} < \infty$$

which in turn follows from (4.4).

Note that (3) follows from Proposition 4.2 with $\gamma_t = (1+K_t)^r$.

(4) Put $\gamma(t) = (1+K_t)^r$ and let us check conditions (a), (b) and (c) of Proposition A.3. From the condition (4.4) directly follows that condition (a) is satisfied with $c_1 = 0$. Condition (b) is trivially satisfied with $c_2 = \frac{\beta}{\sigma^2}$. As for condition (c), it is not hard to check that

$$\gamma_t = 1 + \int_0^t \widetilde{g}_s \, dK_s, \tag{A.3}$$

where

$$\widetilde{g}_t = r(1+K_{t-})^{r-1} I_{\{\Delta K_t = 0\}} + \frac{(1+K_t)^r - (1+K_{t-})^r}{\Delta K_t} I_{\{\Delta K_t \neq 0\}}$$
$$\leq r(1+K_t)^{r-1} \to 0 \text{ as } t \to \infty. \quad \square$$

Appendix B

In this section for convenience of readers we formulate some results from [6], [7].

The following proposition is the second part of Theorem 3.1 from [7].

Proposition B.1. *Let the following conditions be satisfied:*

(i) $[a_t(u) I_{\{\Delta K_t \neq 0\}} + b_t(u)]^+ \leq D_t(1+u^2), \quad D_t \geq 0, \quad D \circ K_\infty < \infty,$ (B.1)

(ii) *For each* $\varepsilon > 0$

$$\inf_{\varepsilon \leq |u| \leq \frac{1}{\varepsilon}} \left\{ |a(u) I_{\{\Delta K_t = 0\}} + [a(u) I_{\{\Delta K_t \neq 0\}} + b(u)]^- \right\} \circ K_\infty = \infty \quad P\text{-a.s.}$$

(B) $\langle M \rangle_\infty < \infty \quad P$-*a.s.*

Then $z_t \to 0$ *as* $t \to \infty$ P-*a.s.*

Combining the results of Theorem 2.1 and Corollary 2.1 from [7] we obtain

Proposition B.2. *Let the following conditions be satisfied:*

(1) $\dfrac{\gamma}{\gamma_-}$ *is eventually bounded,*

(2) $(\beta_t(z_{t-})\Delta K_t)_{t \geq 0}$ *is eventually bounded,*

for all δ, $0 < \delta < \frac{\delta_0}{2}$, P-*a.s.*

(3) $\left[\delta\,\dfrac{\widetilde{g}}{\gamma} - \beta(z_-)\right]^+ \circ K^c_\infty < \infty$, *where* $\widetilde{g}_t = \frac{d\gamma_t}{dK_t}$,

(4) $\displaystyle\sum_{t\geq 0}\left[1 - \beta_t(z_{t-})\Delta K_t - \left(1 - \frac{\Delta\gamma_t}{\gamma_t}\right)^\delta\right]^+ I_{\{\beta_t(z_{t-})\Delta K_t \leq 1\}} < \infty$,

(5) $\displaystyle\sum_{t\geq 0}\left[\beta_t(z_{t-})\Delta K_t - 1 - \left(1 - \frac{\Delta\gamma_t}{\gamma_t}\right)^\delta\right]^+ I_{\{\beta_t(z_{t-})\Delta K_t \geq 1\}} < \infty$,

(6) $\displaystyle\int_0^\infty \gamma_t^{2\delta}\ell_t^2\, dK_t < \infty$.

Then for all δ, $0 < \delta < \frac{\delta_0}{2}$

$$\gamma_t^\delta z_t \to 0 \quad as \quad t \to \infty \quad P\text{-}a.s.$$

Proposition B.3 below is the repetition of Theorem 3.1 from [7] for the case when $\langle L\rangle_t$, $t \geq 0$, is deterministic, $\beta_t\Delta K_t < 1$ eventually, $M(t,u) \equiv M(t,0) := M_t$.

Proposition B.3. *Let the following condition be satisfied: there exists* ε, $\frac{1}{2} - \delta_0 < \varepsilon < \frac{1}{2}$ *such that*

$$\frac{1}{\langle L\rangle_t}\int_0^t |\beta_s - \beta_s(z_{s-})|\gamma_{s-}^\varepsilon \langle L\rangle_s\, dK_s \to 0 \quad as \quad t\to\infty \quad P\text{-}a.s. \tag{B.2}$$

Then

$$\Gamma_t\langle L\rangle_t^{-1/2} z_t = \frac{z_0}{\langle L\rangle_t^{1/2}} + \frac{L_t}{\langle L\rangle_t^{1/2}} + R_t$$

with $R_t \xrightarrow{P} 0$ *as* $t \to \infty$.

Remark 6. The condition: there exists ε, $\frac{1}{2} - \delta_0 < \varepsilon < \frac{1}{2}$ such that

$$\int_0^\infty |\beta_t - \beta_t(z_{t-})|\gamma_{t-}^\varepsilon\, dK_t < \infty \quad P\text{-a.s.}$$

is sufficient for (B.2).

RECURSIVE PARAMETER ESTIMATION IN THE TREND COEFFICIENT OF A DIFFUSION PROCESS

N. LAZRIEVA AND T. TORONJADZE

Abstract. The recursive estimation problem of a one-dimensional parameter in the trend coefficient of a diffusion process is considered. The asymptotic properties of recursive estimators are derived, based on the results on the asymptotic behaviour of a Robbins–Monro type SDE. Various special cases are considered.

Key words and phrases: Diffusion, recursive parameter estimation, stochastic approximation, Robbins–Monro type SDE

MSC 2010: 62F12, 60J60

Dedicated to Professor Yu. V. Prokhorov on the occasion of his 80th birthday

0. Introduction

The asymptotic theory of maximum likelihood estimation of an unknown parameter in the trend coefficient of diffusion processes was developed (in some generality) in a series of works ([2], [3], [12] and others) in the context of Hajek–Le Cam theory. In particular, it was shown that under some regularity conditions the maximum likelihood estimator (MLE) is asymptotically normal and efficient.

As is well known, the maximum likelihood estimator can be constructed solving the MLE equation

$$\int_0^t \frac{\dot{a}(X_s,\theta)}{(\sigma(X_s))^2}\,(dX_s - a(X_s,\theta)ds) = 0$$

which is nonlinear (in general) w.r.t. the parameter θ, and, in addition, requires repeated calculations of a stochastic integral.

On the other hand, to avoid these difficulties Nevelson and Khas'minskiĭ (1972) [11], Albert and Gardner (1967) [1] and others ([9], [10], [14], [15]) introduced recursive procedures of constructing estimators that are asymptotically equivalent to a MLE for special cases of statistical models (i.i.d., diffusion processes, etc.). They also suggested to study the asymptotic properties of recursive procedures by the methods of stochastic approximation.

In [6], [7], the authors proposed the method of constructing recursive estimation procedures for semimartingale statistical models. Later, in [4], [5] they introduced the Robbins–Monro type stochastic differential equation (RM type SDE) and studied the asymptotic properties of solutions (convergence, rate of convergence, asymptotic expansion) based on a general theory of martingales and stochastic calculus.

Published in *Georgian Math. J.* **17** (2010), no. 4, 683–704.

It should be noticed that an RM type SDE covers both general stochastic approximation algorithms with martingale noises and recursive estimation procedures for semimartingale statistical models.

In the present paper we study the problem of recursive estimation of a one-dimensional parameter in the trend coefficient of a homogeneous diffusion process, embedding the recursive estimation procedures (SDEs) in the RM type SDE, and derive the asymptotic properties of recursive estimators based on the results of [4], [5].

In Section 1, we study the general case.

In Section 2, we consider the special cases of ergodic diffusion with the trend coefficient of a separated parameter and phase variables, Ornstein–Uhlenbeck process and ergodic diffusion.

1. General Case

We consider the problem of recursive estimation of the one-dimensional parameter in the trend coefficient of a diffusion process $\xi = \{\xi_t,\ t \geq 0\}$ with

$$d\xi_t = a(\xi_t, \theta)\,dt + \sigma(\xi_t)\,dw_t,\ \ \xi_0, \tag{1.1}$$

where $w = \{w_t,\ t \geq 0\}$ is a standard Wiener process, $a(\cdot, \theta)$ is the known function, $\theta \in \Theta \subseteq R$ is a parameter to be estimated, Θ is some open subset of R, $\sigma^2(\cdot)$ is the known diffusion coefficient.

We assume that there exists a unique strong solution of equation (1.1).

For each $\theta \in \Theta$ denote by P^θ the distribution of the process ξ on $(C_{[0,\infty)}, \mathcal{B})$.

Let $X = \{X_t,\ t \geq 0\}$ be the coordinate process, that is, for each $x = \{x_t, t \geq 0\} \in C_{[0,\infty)},\ X_t(x) = x_t, t \geq 0$.

Fix some $\theta' \in \Theta$ and assume that for each $\theta \in \Theta$, $P^\theta \overset{(loc)}{\sim} P^{\theta'}$. Then the density process $\rho_t(X, \theta)$ can be written as

$$\rho_t(X,\theta) := \frac{dP_t^\theta}{dP_t^{\theta'}}(X) = \exp\left\{ \int_0^t \frac{a(X_s,\theta) - a(X_s,\theta')}{\sigma(X_s)}\, \frac{(dX_s - a(X_s,\theta')ds)}{\sigma(X_s)} \right. \\ \left. - \frac{1}{2}\int_0^t \left(\frac{a(X_s,\theta) - a(X_s,\theta')}{\sigma(X_s)}\right)^2 ds. \right.$$

Recall that if for all $t \geq 0$ P^θ-a.s.

$$\int_0^1 \sigma^2(X_s)\,ds < \infty, \tag{1.2}$$

then the process $X^c := \left\{X_t - \int\limits_0^t a(X_s,\theta)\,ds,\ t \geq 0\right\} \in M_{loc}^2(P^\theta)$ with $\langle X^c\rangle_t = \int\limits_0^t \sigma^2(X_s)\,ds$.

Under suitable regularity conditions if we assume that for all $t \geq 0$ P^θ-a.s.

$$\int_0^t \left(\frac{\dot a(X_s,\theta)}{\sigma(X_s)}\right)^2 ds < \infty, \tag{1.3}$$

we will have

$$\left\{\frac{\partial}{\partial\theta}\ln\rho_t(X,\theta)=\int_0^t\left(\frac{\dot a(X_s,\theta)}{\sigma(X_s)}\right)^2 dX_s^c,\ \ t\geq 0\right\}\in M_{loc}^2(P^\theta),$$

where $\dot a(\cdot,\theta)$ denotes the derivative of $a(\cdot,\theta)$ w.r.t θ.

Below we assume that conditions (1.2) and (1.3) are satisfied.

Introduce the Fisher information process

$$\widehat{I}_t(\theta)=\int_0^t\left(\frac{\dot a(X_s,\theta)}{\sigma(X_s)}\right)^2 ds.$$

Suppose that for each θ, $\widehat{I}_t(\theta)\to\infty$ as $t\to\infty$ P^θ-a.s. and there exists some positive predictable non-increasing process $\{I_t(\theta),\ t\geq 0\}$ such that

$$\widehat{I}_t(\theta)I_t(\theta)\to 1\quad\text{as}\ \ t\to\infty\ \ P^\theta\text{-a.s.}$$

Then, according to equation (1.4.11) from [6], the SDE for constructing the recursive estimator $(\theta_t, t\geq 0)$ has the form

$$d\theta_t=I_t(\theta_t)\left[\frac{\dot a(X_t,\theta_t)}{\sigma^2(X_s)}\,dX_t^c+\frac{\dot a(X_t,\theta_t)}{\sigma^2(X_t)}\ (a(X_t,\theta)-a(X_t,\theta_t))\,dt\right]. \tag{1.4}$$

Fix some $\theta\in\Theta$. To study the asymptotic properties of the recursive estimator $\{\theta_t, t\geq 0\}$ as $t\to\infty$ under measure P^θ let us denote $z_t=\theta_t-\theta$ and rewrite (1.4) in the following form:

$$\begin{aligned}dz_t=I_t(\theta+z_t)\Bigg[&\frac{\dot a(X_t,\theta+z_t)}{\sigma^2(X_s)}\,dX_t^c\\&+\frac{\dot a(X_t,\theta+z_t)}{\sigma^2(X_t)}\ (a(X_t,\theta)-a(X_t,\theta+z_t))\,dt\Bigg].\end{aligned} \tag{1.5}$$

In the sequel we assume that there exists a unique strong solution of equation (1.5) such that

$$\left\{\int_0^t I_s(\theta+z_s)\,\frac{\dot a(X_s,\theta+z_s)}{\sigma^2(X_s)}\,dX_s^c,\ \ t\geq 0\right\}\in M_{loc}^2(P_\theta),$$

that is, for each $t\geq 0$ P^θ-a.s.

$$\int_0^t I_s^2(\theta+z_s)\left(\frac{\dot a(X_s,\theta+z_s)}{\sigma(X_s)}\right)^2 ds<\infty.$$

We will study the asymptotic properties of the process $z=\{z_t,\ t\geq 0\}$ as $t\to\infty$ (under the measure P^θ) using the results of [4], [5] concerning the asymptotic behaviour of solutions of the Robbins–Monro type SDE

$$z_t=z_0+\int_0^t H_s(z_{s-})\,dK_s+\int_0^t M(ds,z_{s-}). \tag{1.6}$$

Note that equation (1.6) covers equation (1.5) with $K_t = t$,

$$H_t(u) := H_t^\theta(u) = I_t(\theta + u)\,\frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)} \times \left(a(X_t,\theta) - a(X_t,\theta+u)\right), \quad H_t^\theta(0) = 0, \tag{1.7}$$

$$M(u) := M^\theta(u) = \left\{ M^\theta(t,u) = \int_0^t I_s(\theta+u)\,\frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)}\,dX_s^c,\ \ t \ge 0 \right\}. \tag{1.8}$$

Let for each $u \in R$ the process $M^\theta(u) \in M^2_{loc}(P^\theta)$. Then

$$\langle M^\theta(u), M^\theta(v)\rangle_t = \int_0^t h_s(u,v)\,ds,$$

where

$$h_t(u,v) = h_t^\theta(u,v) = I_t(\theta+u)I_t(\theta+v)\,\frac{\dot a(X_t,\theta+u)\dot a(X_t,\theta+v)}{\sigma^2(X_t)}\,. \tag{1.9}$$

Let us introduce the following objects:

$$\beta_t^\theta = -\lim_{u\to 0}\frac{H_t^\theta(u)}{u} = I_t(\theta)\left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2, \tag{1.10}$$

$$\beta_t^\theta(u) = \begin{cases} -\dfrac{H_t^\theta(u)}{u}\,, & u \ne 0, \\ \beta_t^\theta, & u = 0, \end{cases} \tag{1.11}$$

$$\Gamma_t^\theta = \exp\left\{ \int_0^t \beta_s^\theta\ ds \right\}, \tag{1.12}$$

$$L_t^\theta = \int_0^t \Gamma_s^\theta\ dM_s^\theta(0). \tag{1.13}$$

Suppose $\Gamma_t^\theta \to \infty$ as $t \to \infty$ P^θ-a.s. and

$$L^\theta = \{L_t^\theta,\ t \ge 0\} \in M^2_{loc}(P^\theta).$$

Below to simplify the notation we omit θ in some formulas.

Theorem 1.1. (I) *Convergence. Let the following conditions be satisfied:*

(A) *for each $t \ge 0$ P^θ-a.s.*

$$H_t^\theta(u)u < 0 \ \textit{ for all } \ u \ne 0, \ \ \theta + u \in \Theta;$$

(i) *for all u, $\theta + u \in \Theta$ and $t \ge 0$ P^θ-a.s.*

$$I_t^2(\theta+u)\left(\frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)}\right)^2 \le B_t(1+|u|^2),$$

where $B = \{B_t,\ t \ge 0\}$ is a predictable positive process (maybe depending on θ) such that

$$\int_0^\infty B_s ds < \infty, \quad P^\theta\textit{-a.s.;}$$

(ii) *for each* ε, $0 < \varepsilon < 1$,

$$\int_0^\infty \inf_{\varepsilon \le |u| \le \frac{1}{\varepsilon}} \left| I_t(\theta+u) \frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)} \left(a(X_t,\theta) - a(X_t,\theta+u)\right) u \right| dt = \infty \ P^\theta\text{-a.s.}$$

Then for any initial value z_0

$$z_t \to 0 \quad as \ \ t \to \infty \ \ P^\theta\text{-a.s.} \tag{1.14}$$

(II) Rate of convergence. *Assume that* (1.14) *holds true.*

Let $\gamma = \{\gamma_t,\ t \ge 0\}$ *be an increasing process such that* $\gamma_t = 1 + \int_0^t g_s ds$, $\gamma_\infty = \infty$, P^θ*-a.s.*

Suppose the following conditions are satisfied:

(i) *for all* δ, $0 < \delta < 1$, P^θ*-a.s.*

$$\int_0^\infty \gamma_t^\delta I_t^2(\theta+z_t) \left(\frac{\dot a(X_t,\theta+z_t)}{\sigma(X_t)} \right)^2 dt < \infty;$$

(ii) *for all* δ, $0 < \delta < \frac{1}{2}$, P^θ*-a.s.*

$$\int_0^\infty \Bigg[\delta \frac{g_t}{\gamma_t} - I_t(\theta) \left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)} \right)^2 \Bigg(I_{\{z_t=0\}} + \frac{I_t(\theta+z_t)}{I_t(\theta)} \frac{\dot a(X_t,\theta+z_t)}{\dot a(X_t,\theta)} \frac{a(X_t,\theta+z_t) - a(X_t,\theta)}{\dot a(X_t,\theta) z_t} I_{\{z_t \ne 0\}} \Bigg) \Bigg]^+ dt < \infty,$$

where $a^+ = \max[0,a]$.

Then for all δ, $0 < \delta < \frac{1}{2}$,

$$\gamma_t^\delta z_t \to 0 \quad as \ \ t \to \infty \ \ P^\theta\text{-a.s.}, \tag{1.15}$$

for any initial value z_0.

(III) Asymptotic expansion. *Assume that* (1.15) *holds true and the following conditions are fulfilled:*

(i) $\langle L \rangle_\infty = \int_0^\infty \Gamma_t^2 \, d\langle M^\theta(0) \rangle_t = \infty \ \ P^\theta$*-a.s.;*

(ii) $\Gamma^2 \langle L \rangle^{-1} \simeq \gamma^*$ (*i.e. these two processes are asymptotically equivalent*);

(iii) $\langle L \rangle = \{\langle L \rangle_t,\ t \ge 0\}$ *is a deterministic process or* $\langle L \rangle \simeq \langle \widetilde{L} \rangle$, *where* $\langle \widetilde{L} \rangle$ *is a deterministic process;*

(iv) *there exists* ε, $\varepsilon > 0$, *such that* P^θ*-a.s.*

$$\frac{1}{\langle L \rangle_t} \int_0^t I_s(\theta) \left(\frac{\dot a(X_s,\theta)}{\sigma(X_s)} \right)^2 \Bigg| 1 - \frac{I_s(\theta+z_s)}{I_s(\theta)} \frac{\dot a(X_s,\theta+z_s)}{\dot a(X_s,\theta)} \times \frac{a(X_s,\theta+z_s) - a(X_s,\theta)}{\dot a(X_s,\theta) z_s} \Bigg| |z_s|^{-\varepsilon} I_{\{z_s \ne 0\}} \langle L \rangle_s ds \to 0 \ \ as \ \ t \to \infty;$$

---

*For the exact definition of "$\simeq$" see Definition 2.2 from [5],

(v) $$\frac{1}{\langle L\rangle_t}\int_0^t \Gamma_s^2 I_s(\theta)\left(\frac{\dot a(X_s,\theta)}{\sigma(X_s)}\right)^2\left[1-\frac{I_s(\theta+z_s)}{I_s(\theta)}\times\frac{\dot a(X_s,\theta+z_s)}{\dot a(X_s,\theta)}\right]^2 ds \xrightarrow{P^\theta} 0 \;\; as \;\; t\to\infty.$$

Then process $z=\{z_t,\ t\ge 0\}$ admits the following representation

$$\Gamma_t\langle L\rangle_t^{-1/2} z_t = \frac{L_t}{\langle L\rangle_t^{1/2}} + R_t$$

with $R_t \xrightarrow{P^\theta} 0$ as $t\to\infty$.

Proof. (I) It is enough to note that conditions (A), (I)(i) and (I)(ii) of the theorem ensure that the conditions (A), (B) and (I) of Theorem 3.1 from [4] are satisfied with

$$h_t(u) = I_t^2(\theta+u)\left(\frac{\dot a(X_t,\theta+u)}{\sigma(X_t)}\right)^2$$

and

$$\alpha_t(u) = I_t(\theta+u)\,\frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)}\,\left(a(X_t,\theta)-a(X_t,\theta+u)\right)u.$$

Thus (1.14) holds true.

(II) Condition (II)(i) of the theorem is the same as condition (2.4) of Theorem 2.1 from [5] with $h_t(u)$ defined as above.

After a simple calculation one can check that condition (II)(ii) is the same as condition (2.6) from [5], taking into account that

$$\beta_t(u) := \beta_t^\theta(u) = -\,I_t(\theta+u)\frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)}\,\frac{a(X_t,\theta)-a(X_t,\theta+u)}{u}\,I_{\{u\ne 0\}} + I_t(\theta)\left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2 I_{\{u=0\}}\,.$$

Then (1.15) directly follows from Corollary 2.1 and Theorem 2.1 from [5], noting that all conditions of Corollary 2.1 are satisfied because $\{K_t,\ t\ge 0\}$ and $\{\gamma_t,\ t\ge 0\}$ are continuous.

(III) One can easily check that conditions (III) (i)–(v) imply all conditions of Theorem 3.1 from [5]. □

2. Special Cases

Case 2.1. Ergodic process with separated phase and parameter variables in the trend coefficient. Let in equation (1.1) the trend coefficient $a(x,\theta) = a(\theta)\varphi(x)$. Suppose that conditions (1.2) and (1.3) are satisfied. Assume also that for any $\theta\in\Theta\subseteq R$ the process ξ has the ergodic property, that is, the functions $a(\cdot,\theta)$ and $\sigma^2(\cdot)$ satisfy conditions (1.44) and (1.45) from [3] with the density of invariant distribution

$$f(x,\theta) = \frac{1}{G(\theta)\sigma^2(x)}\,\exp\left\{2\int_0^x \frac{a(v,\theta)}{\sigma^2(v)}\,dv\right\},$$

where

$$G(\theta) = \int_{-\infty}^{\infty} \sigma^{-2}(y) \exp\left\{2\int_0^y \frac{a(v,\theta)}{\sigma^2(v)}\, dv\right\} dy < \infty.$$

Assume that $a(\theta)$ is continuously differentiable with $\dot{a}(\theta) \neq 0$ for all $\theta \in \Theta$.

In this case

$$\widehat{I}_t(\theta) = (\dot{a}(\theta))^2 \Phi_t(X),$$

where

$$\Phi_t(X) = \int_0^t \left(\frac{\varphi(X_s)}{\sigma(X_s)}\right)^2 ds$$

with the assumption $\Phi_t(X) \to \infty$ as $t \to \infty$ P^θ-a.s.

Put in (1.5)

$$I_t(\theta) = \left[(\dot{a}(\theta))^2 (1 + \Phi_t(X))\right]^{-1}.$$

Proposition 2.1. (I) *Convergence. Suppose that for each $u \neq 0$, $u + \theta \in \Theta$, the following conditions are satisfied:*

(a) $\dot{a}(\theta + u)\, (a(\theta) - a(\theta + u))\, u < 0$*;*

(b) $[\dot{a}(\theta + u)]^{-2} \leq c(\theta)(1 + |u|^2)$*;*

(c) *for each ε, $0 < \varepsilon < 1$,*

$$\inf_{\varepsilon \leq |u| \leq \frac{1}{\varepsilon}} \left| \frac{\dot{a}(\theta + u)}{\dot{a}(\theta)} \, \frac{a(\theta) - a(\theta + u)}{\dot{a}(\theta) u} \right| > 0.$$

Then

$$z_t \to 0 \quad as \quad t \to +\infty \;\; P^\theta\text{-a.s.} \tag{2.1.1}$$

(II) *Rate of convergence. Put $\gamma_t = 1 + (\dot{a}(\theta))^2 \Phi_t(X)$. Then for all δ, $0 < \delta < 1$,*

$$\gamma_t^\delta z_t \to 0 \quad as \;\; t \to \infty, \;\; P^\theta\text{-a.s.} \tag{2.1.2}$$

(III) *Asymptotic expansion. Let the following condition be fulfilled: there exist some $\varepsilon > 0$, $\gamma > 0$ and $c(\theta)$ such that*

$$|\dot{a}(\theta + u) - \dot{a}(\theta + v)| \leq c|u - v|^\gamma$$

for all $u, v \in O_\varepsilon(\theta)$ (here $O_\varepsilon(\theta)$ is some ε-neighbourhood of θ). Then the process $z = \{z_t,\ t \geq 0\}$ admits the following asymptotic expansion

$$\gamma_t^{1/2} z_t = \frac{L_t(\theta)}{\langle L \rangle_t^{1/2}} + R_t \quad with \quad R_t \xrightarrow{P^\theta} 0 \quad as \quad t \to \infty. \tag{2.1.3}$$

Proof. (I) Condition (I) (A), of Theorem 1.1 directly follows from condition (I) (a) of Proposition 2.1.

Condition (I) (i) of Theorem 1.1 also directly follows from condition (I) (b) of Proposition 2.1 with

$$B_t = \left(\frac{\varphi(X_t)}{\sigma(X_t)}\right)^2 (1 + \Phi_t(X))^{-2},$$

Indeed,

$$h_t(u) := h_t(u,u) = I_t^2(\theta+u)\left(\frac{\dot a(X_t,\theta+u)}{\sigma(X_t)}\right)^2$$
$$= [\dot a(\theta+u)]^{-2} B_t \le c(\theta) B_t (1+|u|^2)$$

with

$$\int_0^\infty B_t\,dt = \int_0^\infty (1+\Phi_t(X))^{-2}\,d\Phi_t(X) < \infty \quad P^\theta\text{-a.s.}$$

(II) We will check conditions (II)(i) and (ii) of Theorem 1.1.

Condition (II)(i) for the considered case is of the following form: for all δ, $0<\delta<1$

$$\int_0^\infty \big(1+(\dot a(\theta))^2\Phi_t(X)\big)^\delta \Big[(\dot a(\theta+z_t))^2\,(1+\Phi_t(X))\Big]^{-2} (\dot a(\theta+z_t))^2\,d\Phi_t(X)$$
$$= \int_0^\infty \left(\frac{1+(\dot a(\theta))^2\Phi_t(X)}{1+\Phi_t(X)}\right)^\delta (\dot a(\theta+z_t))^{-2}\,(1+\Phi_t(X))^{\delta-2}\,d\Phi_t(X) < \infty \quad P^\theta\text{-a.s.}$$

But the last integral is finite since

$$\int_0^\infty \frac{d\Phi_t(X)}{(1+\Phi_t(X))^{2-\delta}} < \infty \quad P^\theta\text{-a.s.}$$

and

$$\left(\frac{1+(\dot a(\theta))^2\Phi_t(X)}{1+\Phi_t(X)}\right)^\delta [\dot a(\theta+z_t)]^{-2} \to (\dot a(\theta))^{2(\delta-1)} \quad \text{as } t\to\infty \ \ P^\theta\text{-a.s.}$$

Further condition (II)(ii) of Theorem 1.1 for the considered case means that for all δ, $0<\delta<1$, P^θ-a.s.

$$(\dot a(\theta))^2 \int_0^\infty \Bigg[\delta - \frac{1+(\dot a(\theta))^2\Phi_t(X)}{(\dot a(\theta))^2(1+\Phi_t(X))}\Bigg(I_{\{z_t=0\}}$$
$$+\frac{\dot a(\theta)}{\dot a(\theta+z_t)}\,\frac{a(\theta+z_t)-a(\theta)}{\dot a(\theta)z_t}\,I_{\{z_t\neq 0\}}\Bigg)\Bigg]^+ \frac{d\Phi_t(X)}{1+(\dot a(\theta))^2\Phi_t(X)} < \infty$$

and is fulfilled since

$$\Bigg[\delta - \frac{1+(\dot a(\theta))^2\Phi_t(X)}{(\dot a(\theta))^2(1+\Phi_t(X))}\Bigg(I_{\{z_t=0\}}$$
$$+\frac{\dot a(\theta)}{\dot a(\theta+z_t)}\,\frac{a(\theta+z_t)-a(\theta)}{\dot a(\theta)z_t}\,I_{\{z_t\neq 0\}}\Bigg)\Bigg]^+ = 0 \ \text{ eventually.}$$

Indeed,

$$I_{\{z_t=0\}} + \frac{\dot a(\theta)}{\dot a(\theta+z_t)}\,\frac{a(\theta+z_t)-a(\theta)}{\dot a(\theta)z_t}\,I_{\{z_t\neq 0\}} \to 1 \quad \text{as } t\to\infty \ \ P^\theta\text{-a.s.}$$

and also

$$\frac{1+(\dot a(\theta))^2\Phi_t(X)}{\dot a(\theta)^2(1+\Phi_t(X))} \to 1 \quad \text{as } t\to\infty \ \ P^\theta\text{-a.s.}$$

(III) Using (1.11), (1.12) and (1.13) we have

$$\Gamma_t := \Gamma_t^\theta = 1 + \Phi_t(X) \to \infty \quad \text{as} \quad t \to \infty \quad P^\theta\text{-a.s.},$$
$$\langle L^\theta\rangle_t \, (\dot a(\theta))^2 \, \Phi_t(X) \to \infty \quad \text{as} \quad t \to \infty \quad P^\theta\text{-a.s.},$$
$$\lim_{t\to\infty} \frac{\Gamma_t^2 \langle L^\theta\rangle_t^{-1}}{\gamma_t} = 1$$

and, moreover, if we denote

$$i(\theta) = \int_R \left(\frac{\varphi(x)}{\sigma(x)}\right)^2 f(x,\theta)\, dx, \tag{2.1.4}$$

then

$$\frac{\langle L^\theta\rangle_t}{(1+t)i(\theta)} \to 1 \quad \text{as} \quad t \to \infty \quad P^\theta\text{-a.s.}$$

Hence conditions (III) (i)–(iii) of Theorem 1.1 are satisfied.

Further, according to condition (III) (iv) of Theorem 1.1 we have to show that there exists ε, $\varepsilon > 0$, such that P^θ-a.s.

$$\frac{1}{\langle L\rangle_t} \int_0^t \left| 1 - \frac{a(\theta+z_s) - a(\theta)}{\dot a(\theta+z_s) z_s} \right| |z_s|^{-\varepsilon}\, I_{\{z_s\neq 0\}} \langle L^\theta\rangle_s \, \frac{d\Phi_s(X)}{1+\Phi_s(X)} \to 0 \text{ as } t \to \infty.$$

But according the the stochastic version of the Kronecker lemma it is enough to show that P^θ-a.s.

$$\int_0^\infty \left| 1 - \frac{a(\theta+z_s) - a(\theta)}{\dot a(\theta+z_s) z_s} \right| |z_s|^{-\varepsilon}\, I_{\{z_s\neq 0\}} \, \frac{d\Phi_s(X)}{1+\Phi_s(X)} < \infty.$$

On the other hand,

$$\left| 1 - \frac{a(\theta+z_s) - a(\theta)}{\dot a(\theta+z_s) z_s} \right| = \frac{|\dot a(\theta+z_s) - \dot a(\theta+\widetilde z_s)|}{|\dot a(\theta+z_s)|} \leq const(\theta) |z_s - \widetilde z_s|^\gamma,$$

where $0 < |\widetilde z_s| < |z_s|$. Therefore for all δ, $0<\delta<1$, $\varepsilon < \gamma$

$$\begin{aligned}
&\int_0^\infty \left| 1 - \frac{a(\theta+z_s) - a(\theta)}{\dot a(\theta+z_s) z_s} \right| |z_s|^{-\varepsilon}\, I_{\{z_s\neq 0\}} \, \frac{d\Phi_s(X)}{1+\Phi_s(X)} \\
&\quad \leq const(\theta) \int_0^\infty |\dot a(\theta+z_s)|^{-1}\, |z_s - \widetilde z_s|^\gamma |z_s|^{-\varepsilon} I_{\{z_s\neq 0\}} \, \frac{d\Phi_s(X)}{1+\Phi_s(X)} \\
&\quad \leq const(\theta) \int_0^\infty |z_s|^{\gamma-\varepsilon} I_{\{z_s\neq 0\}}\, d\Phi_s(X) \\
&\quad \leq const(\theta) \int_0^\infty |\gamma_s^\delta z_s|^{\gamma-\varepsilon} \gamma_s^{-\delta(\gamma-\varepsilon)} \, \frac{d\Phi_s(X)}{1+\Phi_s(X)} \\
&\quad \leq const(\theta) \int_0^\infty \frac{d\Phi_s(X)}{(1+\Phi_s(X))^{1+\delta(\gamma-\varepsilon)}} < \infty \quad P^\theta\text{-a.s.}
\end{aligned}$$

It remains to check condition (III)(v) which for the considered case becomes

$$\frac{1}{1+(\dot a(\theta))^2 \Phi_t(X)} \int_0^t (\dot a(\theta))^{-2} \left[1 - \frac{\dot a(\theta)}{\dot a(\theta+z_s)}\right]^2 d\Phi_s(X) \xrightarrow{P^\theta} 0 \quad \text{as} \quad t \to \infty,$$

and is satisfied by virtue of the Toeplitz lemma and since $\Phi_\infty(X)=\infty$, $1-\frac{\dot a(\theta)}{\dot a(\theta+z_s)}\to 0$ as $t\to\infty$ P^θ-a.s. $\square$

Remark 2.1. 1. Conditions (I) (a) and (c) of Proposition 2.1 are satisfied if $\dot a(\cdot)$ is a strongly monotone function.

2. If $a(\theta)=\theta$, all conditions of Proposition 2.1 will be fulfilled if the process is ergodic and

$$\int_R \left(\frac{\varphi(x)}{\sigma(x)}\right)^2 f(x,\theta)\,dx<\infty,$$

where $f(x,\theta)$ is the density of an invariant distribution.

3. From equation (2.1.3) it directly follows that

$$\mathcal{L}_\theta\left\{\sqrt{t}\,z_t\right\}\Rightarrow N\left(0,\left((\dot a(\theta))^2 i(\theta)\right)^{-1}\right),$$

where $i(\theta)$ is defined by (2.1.4), $\mathcal{L}_\theta$ denotes the probability distribution of process $z=(z_t)_{t\geq 0}$ under the measure P^θ, and the symbol "$\Rightarrow$" is used to denote weak convergence of distributions. Indeed, from (2.1.3) it is evident that the weak limits of $\mathcal{L}_\theta(\gamma_t^{1/2}z_t)$ and $\mathcal{L}_\theta\left\{\frac{L_t(\theta)}{\langle L_t\rangle^{1/2}}\right\}$ coincide and according to the Central Limit Theorem for martingales one obtains

$$\mathcal{L}_\theta\left(\gamma_t^{1/2}z_t\right)\Rightarrow N(0,1).$$

It remains to note that

$$\lim_{t\to\infty}\frac{\gamma_t}{t}=(\dot a(\theta))^2\,i(\theta).$$

Example 2.1. The Ornstein–Uhlenbeck process: $a(x,\theta)=-\theta x$, $\theta>0$, that is, $\Theta=\{\theta,\ \theta>0\}$, $\sigma(x)\equiv 1$ (for simplicity).

As is well known, the MLE of θ is

$$\widehat\theta_t=\theta+\frac{\int_0^t X_s dw_s}{\int_0^t X_s^2\,ds},$$

where $dw_t=dX_t+\theta X_t\,dt$ is a Wiener process. The asymptotic behaviour of the normed process $\sqrt{t}(\widehat\theta_t-\theta)=\sqrt{t}\,z_t$ directly can be obtain using martingale limit theorems, namely,

$$\mathcal{L}_\theta(\sqrt{t}\,z_t)\Rightarrow N(0,2\theta).$$

The same result can be obtained if we rewrite the process $z=(z_t)_{t\geq 0}$ in the recurrent form

$$dz_t=-z_t\,\frac{X_t^2}{\int_0^t X_s^2\,ds}\,dt+\frac{X_t}{\int_0^t X_s^2 ds}\,dw_t$$

and study its asymptotic behaviour using the results of Proposition 2.1. Since the Ornstein–Uhlenbeck process has the ergodic property as it has been mentioned in Remark 2.1, all conditions of this proposition are satisfied and we obtain the same results since $\lim\limits_{t\to\infty}\frac{\gamma_t}{t}=2\theta$.

Case 2.2. Ergodic diffusion. We consider the diffusion process $\xi = \{\xi_t,\ t \geq 0\}$ defined by (1.1) and assume that for any $\theta \in \Theta \subset R$ the process ξ has the ergodic property. Besides, the law of large numbers is fulfilled: for any measurable function $\varphi(\cdot, \theta)$ with

$$\int_R |\varphi(x,\theta)|\, f(x,\theta)\, dx < \infty$$

we have the convergence

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \varphi(S_s, \theta) ds = \int_R \varphi(x,\theta) f(x,\theta)\, dx$$

(see, e.g., [3], [13]).

According to this law,

$$\lim_{t\to\infty} \frac{1}{1+t} \widehat{I}_t(\theta) = \int_R \left(\frac{\dot{a}(x,\theta)}{\sigma(x)} \right)^2 f(x,\theta)\, dx := I(\theta).$$

Below we assume that for each $\theta \in \Theta$ $0 < I(\theta) < \infty$. Hence one can put

$$I_t(\theta) = [(1+t)I(\theta)]^{-1}. \tag{2.2.1}$$

We will study the asymptotic properties of the solution $z = \{z_t,\ t \geq 0\}$ of the recursive SDE (1.5) with $I_t(\theta)$ defined by (2.2.1) based on the results of Theorem 1.1.

For this purpose define the following objects:

$$\begin{aligned}
\gamma_t &= 1 + \int_0^t \left(\frac{\dot{a}(X_s,\theta)}{\sigma(X_s)} \right)^2 ds, \\
\Gamma_t &= \exp\left\{ \int_0^t I_s(\theta) \left(\frac{\dot{a}(X_s,\theta)}{\sigma(X_s)} \right)^2 ds \right\} = \exp\left\{ \int_0^t I_s(\theta)\, d\gamma_s \right\}, \\
L_t &= \int_0^t \Gamma_s \, dM^\theta(s,0),
\end{aligned}$$

where

$$M^\theta(t,0) = \int_0^t I_s(\theta) \frac{\dot{a}(X_s,\theta)}{\sigma^2(X_s)}\, d(X_s - a(X_s,\theta) ds).$$

Lemma 2.1. (1) *$\Gamma_t \to \infty$ as $t \to \infty$ P^θ-a.s.*

(2) *If*

$$\int_0^\infty |I_t(\theta)\gamma_t - 1|\, d\ln\gamma_t < \infty \quad P^\theta\text{-a.s.}, \tag{2.2.2}$$

then $\Gamma_t \simeq \gamma_t$.

(3) *$\langle L \rangle_\infty = \infty$ P^θ-a.s.*

(4) *$\lim_{t\to\infty} \dfrac{\Gamma_t^2 \langle L \rangle_t^{-1}}{\gamma_t} = 1$ P^θ-a.s.*

Proof. (1) Since $\gamma_t \to \infty$ as $t \to \infty$ P^θ-a.s. and

$$\ln \Gamma_t = \int_0^t I_s(\theta)\gamma_s\, d\ln\gamma_s,$$

one can apply the Toeplitz lemma to obtain

$$\lim_{t\to\infty}\frac{\ln\Gamma_t}{\ln\gamma_t}=\lim_{t\to\infty}I_t(\theta)\gamma_t=1\quad P^\theta\text{-a.s.}\tag{2.2.3}$$

Moreover, equation (2.2.3) allows one to conclude that

$$\lim_{t\to\infty}\Gamma_t^2\gamma_t^{-1}=\infty\quad P^\theta\text{-a.s.}\tag{2.2.4}$$

Indeed,

$$\lim_{t\to\infty}\ln\left(\Gamma_t^2\gamma_s^{-1}\right)=\lim_{t\to\infty}\ln\gamma_t\left(2\,\frac{\ln\Gamma_t}{\ln\gamma_t}-1\right)=\infty\quad P^\theta\text{-a.s.}$$

Assertion (2) directly follows from the equality

$$\frac{\Gamma_t}{\gamma_t}=\exp^{\int_0^t(I_s(\theta)\gamma_s-1)d\ln\gamma_s}\ .$$

(3) We have

$$\langle L\rangle_t=\int_0^t\Gamma_s^2I_s^2(\theta)\,d\gamma_s=\int_0^t\Gamma_s^2\gamma_s^{-1}(\gamma_sI_s(\theta))^2\,d\ln\gamma_s\to\infty\ \text{ as }\ t\to\infty$$

since $\Gamma_t^2\gamma_t^{-1}\to\infty$, $\gamma_sI_s(\theta)\to1$ as $t\to\infty$ P^θ-a.s.

(4) We will check conditions (a), (b) and (c) of Proposition A3 from [8].

Condition (a) is trivially satisfied because $K_t=t$, so $c_1=0$.

As for condition (b), note that

$$\langle M^\theta(0)\rangle_t=\int_0^t\ell_s^2ds,$$

where

$$\ell_t^2=I_t^2(\theta)\left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2.$$

Hence

$$\frac{1}{\gamma_t}\,\frac{\beta_t}{\ell_t^2}=\frac{1}{\gamma_tI_t(\theta)}\to1\quad\text{as }\ t\to\infty\ \ P^\theta\text{-a.s.}$$

Thus $c_2=1$.

It is not difficult to show (in the same way as above) that condition (c) is also satisfied with $c_3=1$. Therefore according to Proposition A3 from [8]

$$\lim_{t\to\infty}\frac{\Gamma_t^2\langle L\rangle_t^{-1}}{\gamma_t}=2(c_2-c_1)-c_3=1.\qquad\square$$

Remark 2.2. (I) Suppose that the process $\xi=\{\xi_t,\ t\ge0\}$ defined by (1.1) is stationary and the following conditions are satisfied:

a) there exist some constants c_1 and p_1 (not depending on θ) such that for all $\theta\in\Theta$

$$\left(\frac{\dot a(x,\theta)}{\sigma(x)}\right)^2\le c_1\left(1+|x|^{p_1}\right);$$

b) there exist the constants c_2 and p_2 such that

$$\left|\sigma^{-1}(x)\right|\le c_2\left(1+|x|^{p_2}\right);$$

c) for each θ

$$\overline{\lim_{|x|\to\infty}} \operatorname{sgn}(x)\, \frac{a(x,\theta)}{\sigma^2(x)} < 0.$$

If there exists some $\varepsilon > 0$ such that

$$\int \left(\frac{\dot a(x,\theta)}{\sigma(x)}\right)^{2+\varepsilon} f(x,\theta)\, dx < \infty, \tag{2.2.5}$$

then condition (2.2.2) of Lemma 2.1 is satisfied.

Indeed, we have

$$\begin{aligned}\int_0^\infty |I_t(\theta)\gamma_t - 1|\; d\ln\gamma_t &= \int_0^\infty |I_t(\theta)\gamma_t - 1|\; \left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2 \gamma_t^{-1}\, dt \\ &= \int_0^\infty |I_t(\theta)\gamma_t - 1|\; I_t(\theta)\; \left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2 \alpha_t\, dt,\end{aligned}$$

where $\alpha_t = (\gamma_t I_t(\theta))^{-1} \to 1$ as $t\to\infty$ P^θ-a.s. Therefore it is enough to show that

$$\int_0^\infty |I_t(\theta)\gamma_t - 1|\; I_t(\theta)\; \left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2 dt < \infty \quad P^\theta\text{-a.s.} \tag{2.2.6}$$

Further, we have (using the Hölder inequality with $p = 1+\frac{2}{\varepsilon}$ and $q = 1+\frac{\varepsilon}{2}$, $\varepsilon>0$)

$$\begin{aligned}&E_\theta \int_0^\infty |I_t(\theta)\gamma_t - 1|\; I_t(\theta)\; \left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2 dt \\ &\le \int_0^\infty \frac{1}{(1+t)I(\theta)}\; \left(E_\theta\, |I_t(\theta)\gamma_t - 1|^{1+\frac{2}{\varepsilon}}\right)^{\frac{\varepsilon}{2+\varepsilon}} \left(E_\theta\left[\left(\frac{\dot a(\xi,\theta)}{\sigma(\xi)}\right)^2\right]^{1+\frac{\varepsilon}{2}}\right)^{\frac{\varepsilon}{2+\varepsilon}} dt \\ &\le \left(E_\theta\left(\frac{\dot a(\xi,\theta)}{\sigma(\xi)}\right)^{2+\varepsilon}\right)^{\frac{\varepsilon}{2+\varepsilon}} \\ &\qquad \times \int_0^\infty \frac{1}{(1+t)I(\theta)}\; \left(E_\theta\, |I_t(\theta)\gamma_t - 1|^{1+\frac{2}{\varepsilon}}\right)^{\frac{\varepsilon}{2+\varepsilon}} dt,\end{aligned} \tag{2.2.7}$$

where ξ is a random variable with a distribution density $f(x,\theta)$.

On the other hand, under conditions a), b) and c) we have for all $p>0$ (see Proposition 1.18 from [3])

$$E_\theta\left|\frac{1}{1+t}\int_0^{1+t}\left(\frac{\dot a(X_s,\theta)}{\sigma(X_s)}\right)^2 ds - I(\theta)\right|^p \le \frac{c_p}{(1+t)^{\frac{p}{2}}}\,.$$

Using the last inequality with $p = \frac{2}{\varepsilon}+1$ we obtain

$$\left(E_\theta\, |I_t(\theta)\gamma_t - 1|^{1+\frac{2}{\varepsilon}}\right)^{\frac{\varepsilon}{2+\varepsilon}} \le const \left(\frac{1}{(1+t)^{\frac12+\frac1\varepsilon}}\right)^{\frac{\varepsilon}{2+\varepsilon}} = const\, \frac{1}{(1+t)^{\frac12}}\,.$$

Substituting this inequality in (2.2.7) we obtain

$$E_\theta \int_0^\infty |I_t(\theta)\gamma_t - 1|\; I_t(\theta)\; \left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2 dt \le const \int_0^\infty (1+t)^{-\frac32} dt < \infty.$$

Theorem 2.1. (I) *Convergence. Let the following conditions be fulfilled:*

(A) *for each* $x \in R$

$$\dot{a}(x, \theta + u)(a(x, \theta) - a(x, \theta + u))u < 0 \quad \textit{for all} \quad u \neq 0$$

(for instance, this condition is satisfied if for each $x \in R$ *the function* $a(x, \theta)$ *is strongly monotone in* θ*).*

(i) *for each* x

$$\left[\frac{\dot{a}(x, \theta + u)I(\theta)}{\dot{a}(x, \theta)I(\theta + u)}\right]^2 \leq c(1 + |u|^2),$$

where c *is a constant;*

(ii) *for each* ε, $0 < \varepsilon < 1$,

$$\inf_{\varepsilon \leq |u| \leq \frac{1}{\varepsilon}} \inf_x \left| \frac{I(\theta)}{I(\theta + u)} \frac{\dot{a}(x, \theta + u)}{\dot{a}(x, \theta)} \frac{a(x, \theta) - a(x, \theta + u)}{\dot{a}(x, \theta)u} \right| > 0.$$

Then

$$z_t \to 0 \quad \textit{as} \quad t \to \infty \quad P^\theta\textit{-a.s.}$$

(II) *Rate of convergence. Suppose that the following conditions are satisfied:*

$$\text{(i)} \quad \sup_x \left| \frac{I(\theta)}{I(\theta + u)} \frac{\dot{a}(x, \theta + u)}{\dot{a}(x, \theta)} - 1 \right| \to 0 \quad \textit{as} \quad u \to 0,$$

$$\text{(ii)} \quad \sup_x \left| \frac{a(x, \theta + u) - a(x, \theta)}{\dot{a}(x, \theta)u} - 1 \right| \to 0 \quad \textit{as} \quad u \to 0.$$

Then for each δ, $0 < \delta < \frac{1}{2}$,

$$\gamma_t^\delta z_t \to 0 \quad \textit{as} \quad t \to \infty \quad P^\theta\textit{-a.s.}$$

(III) *Asymptotic expansion. Let the following conditions be satisfied:*

(i) *there exists some* $\gamma > 0$ *such that*

$$\sup_x \left| \frac{I(\theta)}{I(\theta + u)} \frac{\dot{a}(x, \theta + u)}{\dot{a}(x, \theta)} \frac{a(x, \theta + u) - a(x, \theta)}{\dot{a}(x, \theta)u} - 1 \right| = O(|u|^\gamma) \quad \textit{as } u \to 0;$$

$$\text{(ii)} \quad \sup_x \left| 1 - \frac{I(\theta)}{I(\theta + u)} \frac{\dot{a}(x, \theta + u)}{\dot{a}(x, \theta)} \right| \to 0 \quad \textit{as} \quad u \to 0.$$

Then if condition (2) *of Lemma* 2.1 *is fulfilled, the process* $z = \{z_t, \ t \geq 0\}$ *admits the following asymptotic expansion*

$$\Gamma_t \langle L \rangle_t^{-1/2} z_t = \frac{L_t}{\langle L \rangle_t^{1/2}} + R_t$$

with $R_t \xrightarrow{P^\theta} 0$ *as* $t \to 0$.

Proof. (I) We have to check conditions (I) (A), (i) and (ii) of Theorem 1.1. Condition (I) (A) is satisfied.

As for condition (i) we have

$$I_t^2(\theta+u)\left(\frac{\dot a(X_t,\theta+u)}{\sigma(X_t)}\right)^2 = \frac{1}{[(1+t)I(\theta)]^2}\left[\frac{\dot a(X_t,\theta+u)I(\theta)}{\dot a(X_t,\theta)I(\theta+u)}\right]\left[\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right]^2$$
$$\le c\left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2(1+|u|^2).$$

Put

$$B_t = \left(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\right)^2 I_t^2(\theta).$$

Then

$$\int_0^\infty B_t\,dt < \infty \quad P^\theta\text{-a.s.}$$

Indeed, we have, recalling the definition of $I_t(\theta)$ and γ_t,

$$\int_0^\infty B_t\,dt = \int_0^\infty \gamma_t^2 I_t^2(\theta)\,\frac{d\gamma_t}{\gamma_t^2} < \infty$$

since $\gamma_t^2 I_t(\theta) \to 1$ as $t\to\infty$, and $\int_0^\infty \frac{d\gamma_t}{\gamma_t^2} < \infty$ P^θ-a.s.

According to (I) (iii) of Theorem 1.1 we have to show

$$\int_0^\infty \frac{1}{1+t}\inf_{\varepsilon\le|u|\le\frac{1}{\varepsilon}}\left|\frac{u}{I(\theta+u)}\,\frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)}\,(a(X_t,\theta)-a(X_t,\theta+u))\right|dt$$
$$< \infty \quad P^\theta\text{-a.s.} \tag{2.2.8}$$

But the integral of the last expression can be written as

$$\int_0^\infty \gamma_t I_t(\theta)\inf_{\varepsilon\le|u|\le\frac{1}{\varepsilon}}|u|^2\left|\frac{I(\theta)}{I(\theta+u)}\,\frac{\dot a(X_t,\theta+u)}{\dot a(X_t,\theta)}\,\frac{a(X_t,\theta)-a(X_t,\theta+u)}{\dot a(X_t,\theta)u}\right|\frac{d\gamma_t}{\gamma_t}$$
$$\ge \varepsilon^2\inf_{\varepsilon\le|u|\le\frac{1}{\varepsilon}}\inf_x\left|\frac{I(\theta)}{I(\theta+u)}\,\frac{\dot a(X_t,\theta+u)}{\dot a(X_t,\theta)}\,\frac{a(X_t,\theta)-a(X_t,\theta+u)}{\dot a(X_t,\theta)u}\right|$$
$$\times\int_0^\infty \gamma_t I_t(\theta)\,\frac{d\gamma_t}{\gamma_t} = \infty,$$

this follows from condition (I) (ii) of Theorem 2.1, and from facts that $\gamma_t I_t(\theta)\to 1$ as $t\to\infty$ and $\int_0^\infty \frac{d\gamma_t}{\gamma_t} = \infty$ P^θ-a.s.

(II) (i) We will check condition (II)(i) of Theorem 1.1, which in the case under consideration takes the form: for each δ, $0<\delta<1$,

$$\int_0^\infty \gamma_t^\delta I_t^2(\theta+z_t)\left(\frac{\dot a(X_t,\theta+z_t)}{\sigma(X_t)}\right)^2 dt < \infty \quad P^\theta\text{-a.s.} \tag{2.2.9}$$

But

$$\begin{aligned}\int_0^\infty \gamma_t^\delta I_t^2(\theta+z_t)\Big(\frac{\dot a(X_t,\theta+z_t)}{\sigma(X_t)}\Big)^2 dt\\ &= \int_0^\infty \gamma_t^\delta I_t^2(\theta)\Big(\frac{I(\theta)}{I(\theta+z_t)}\frac{\dot a(X_t,\theta+z_t)}{\dot a(X_t,\theta)}\Big)^2 \gamma_t^2 d\gamma_t\\ &\le \int_0^\infty I_t^2(\theta)\gamma_t^2 \sup_x\Big(\frac{I(\theta)}{I(\theta+z_t)}\frac{\dot a(X_t,\theta+z_t)}{\dot a(X_t,\theta)}\Big)^2\frac{d\gamma_t}{\gamma_t^{2-\delta}}<\infty,\end{aligned}$$

which directly follows from the relation $I_t(\theta)\gamma_t \to 1$ as $t\to\infty$, condition (II)(i) of Theorem 2.1 and the inequality

$$\int_0^\infty \frac{d\gamma_t}{\gamma_t^{2-\delta}}<\infty \quad P^\theta\text{-a.s.}$$

(ii) We have, after simple calculations (recall that $g_t=\big(\frac{\dot a(X_t,\theta)}{\sigma(X_t)}\big)^2$):

$$\begin{aligned}\Big[\delta\frac{g_t}{\gamma_t}-\beta_t(z_t)\Big]^+ &= \frac{g_t}{\gamma_t}\Big[\delta-\gamma_t I_t(\theta)+\gamma_t I_t(\theta)\Big(1-\frac{I(\theta)}{I(\theta+z_t)}\frac{\dot a(X_t,\theta+z_t)}{\dot a(X_t,\theta)}\\ &\quad\times\frac{\dot a(X_t,\theta+z_t)-a(X_t,\theta)}{a(X_t,\theta)z_t}\Big)I_{\{z_t\neq 0\}}\Big]^+ = 0 \ \text{ eventually.}\end{aligned}\tag{2.2.10}$$

Indeed, from conditions (II) (i), (ii) of Theorem 2.1 it directly follows that

$$\sup_x\Big|\frac{I(\theta)}{I(\theta+z_t)}\frac{\dot a(x,\theta+z_t)}{\dot a(x,\theta)}\frac{a(x,\theta+z_t)-a(x,\theta)}{\dot a(x,\theta)z_t}-1\Big|\to 0 \ \text{ as } \ t\to\infty \ \ P^\theta\text{-a.s.}$$

and (2.2.10) can be derived using the same arguments as in the proof of (II)(ii) of Case 2.1.

(III) By virtue of Lemma 2.1, conditions (III) (i)–(iii) of Theorem 1.1 are satisfied.

Let us check condition (iv) which for the considered case is formulated as follows: there exists ε, $\varepsilon>0$, such that

$$\begin{aligned}\frac{1}{\langle L\rangle_t}\int_0^t [(1+s)I(\theta)]^{-1}\Big|1-\frac{I(\theta+z_t)}{I(\theta)}\frac{\dot a(X_s,\theta+z_s)}{\dot a(X_s,\theta)}\\ \times\frac{a(X_s,\theta+z_s)-a(X_s,\theta)}{\dot a(X_s,\theta)z_s}\Big||z_s|^{-\varepsilon}\langle L\rangle_s\,ds\to 0 \ \text{ as } \ t\to\infty \ \ P^\theta\text{-a.s.}\end{aligned}$$

Since $\langle L\rangle_\infty=\infty$, P^θ-a.s. according to the Kronecker lemma it is enough to show that there exists $\varepsilon>0$ such that

$$\begin{aligned}\int_0^\infty [(1+s)I(\theta)]^{-1}\Big(\frac{\dot a(X_s,\theta)}{\sigma(X_s)}\Big)^2\Big|1-\frac{I(\theta+z_t)}{I(\theta)}\frac{\dot a(X_s,\theta+z_s)}{\dot a(X_s,\theta)}\\ \times\frac{a(X_s,\theta+z_s)-a(X_s,\theta)}{\dot a(X_s,\theta)z_s}I_{\{z_t\neq 0\}}\Big||z_s|^{-\varepsilon}\,ds<\infty \ \ P^\theta\text{-a.s.}\end{aligned}$$

Further, using condition (III)(i) the last integral is less than

$$const \int_0^\infty \gamma_t^{-1} |z_t|^{\gamma-\varepsilon} d\gamma_t = const \int_0^\infty \gamma_t^{-1} |\gamma_t^\delta z_t|^{\gamma-\varepsilon} \gamma_t^{-\delta(\gamma-\varepsilon)} d\gamma_t$$
$$\leq const \int_0^\infty \gamma_t^{-(1+\delta(\gamma-\varepsilon))} d\gamma_t < \infty$$

for $\varepsilon < \gamma$.

It remains to check condition (v) of Theorem 1.1. For this purpose rewrite it as

$$\frac{1}{\langle L\rangle_t} \int_0^t \left[1 - \frac{I(\theta)}{I(\theta+z_s)} \frac{\dot{a}(X_s, \theta+z_s)}{\dot{a}(X_s,\theta)}\right]^2 d\langle L\rangle_s \xrightarrow{P_\theta} 0. \tag{2.2.11}$$

Now, applying the Toeplitz lemma and taking into account (III)(ii) of Theorem 2.1 we obtain the desired result (2.2.11). □

Acknowledgement

This work was supported by grant GNSF/ST007/3-172.

References

[1] A. E. Albert and L. A. Gardner, Jr., *Stochastic Approximations and Nonlinear Regression*, M.I.T. Press Research Monograph No. 42. The M.I.T. Press, Cambridge, Mass. 1967.

[2] I. A. Ibragimov and R. Z. Khas'minskiĭ, *Statistical Estimation. Asymptotic Theory*, Springer-verlag, New-York, 1982.

[3] Yu. A. Kutoyants, *Statistical Inference for Ergodic Diffusion Processes*, Springer-Verlag, 2004.

[4] N. Lazrieva, T. Sharia and T. Toronjadze, The Robbins-Monro type stochastic differential equations. I. Convergence of solutions, *Stochastics Stochastics Rep.* **61**:1-2 (1997), 67–87.

[5] N. Lazrieva, T. Sharia and T. Toronjadze, The Robbins–Monro type stochastic differential equations. II. Asymptotic behaviour of solutions, *Stochastics Stochastics Rep.* **75**:3 (2003), 153–180.

[6] N. Lazrieva, T. Sharia and T. Toronjadze, Semimartingale stochastic approximation procedures, *J. Math. Sci.* **153** (2008), 211–261.

[7] N. Lazrieva and T. Toronjadze, Itô–Ventzel's formula for semimartingales, asymptotic properties of MLE and recursive estimation, in: *Stochastic differential systems, Proc. IFIP-WG* 7/1 *Work. Conf., Eisenach/GDR* 1986, Lect. Notes Control Inf. Sci. 96, 346–355 (1987).

[8] N. Lazrieva and T. Toronjadze, The Robbins–Monro type stochastic differential equation. III. Polyak's averaging, *Stochastics: An International Journal of Probabilities and Stochastic Processes*, accepted for publication, 2009.

[9] D. Levanony, A. Shwartz, and O. Zeitouni, Continuous-time recursive identification, in: E. Arikan (ed.) *Communication, Control and Signal Processing,* 1725–1732; Elsevier, Amsterdam, 1990.

[10] D. Levanony, A. Shwartz, and O. Zeitouni, Recursive identification in continuous-time stochastic processes, *Stochastic Process. Appl.* **49**:2 (1994), 245–275.

[11] M. B. Nevelson and R. Z. Khas'minskiĭ, *Stochastic Approximation and Recurrent Estimation*, (Russian) Monographs in Probability Theory and Mathematical Statistics. Nauka, Moscow, 1972.

[12] B. L. S. Prakasa Rao, *Statistical Inference for Diffusion Type Processes*, Arnold, London, 1999.

[13] A. V. Skorokhod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, (Translated from the Russian) Translations of Mathematical Monographs, 78. American Mathematical Society, Providence, RI, 1989; Russian original: Naukova Dumka, Kiev, 1987.

[14] T. Sharia, On the recursive parameter estimation in the general discrete time statistical model, *Stochastic Process. Appl.* **73**:2 (1998), 151–172.

[15] P. Spreij, Recursive approximate maximum likelihood estimation for a class of counting process models, *J. Multivariate Anal.* **39**:2 (1991), 236–245.

ROBUST UTILITY MAXIMIZATION FOR A DIFFUSION MARKET MODEL WITH MISSPECIFIED COEFFICIENTS

R. TEVZADZE, T. TORONJADZE AND T. UZUNASHVILI

Abstract. The paper studies the robust maximization of terminal wealth utility in a diffusion financial market model. The underlying model consists of a risky tradable asset whose price is described by the diffusion process with misspecified trend and volatility coefficients, and a non-tradable asset with the known parameter. The robust functional is defined in terms of a utility function. An explicit characterization of the problem solution is given using the solution of the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation.

Key words and phrases: Maximin problem, saddle point, Hamilton-Jacobi-Bellman-Isaacs equation, robust utility maximization, generalized control.

MSC 2010: 60H10, 60H30, 90C47.

1. INTRODUCTION

The purpose of the present paper is to study the robust maximization of terminal wealth utility in a diffusion financial market model where the trend and volatility of an asset price are uncertain.

The concept of robustness was introduced by P. Huber (see [19]) in the context of statistical estimation of an unknown distribution parameter. The essence of our approach is as follows. Suppose we need to estimate the mean of some symmetric distribution. If the estimation is based on "pure" observations, then the effective estimate is the sample mean. But if observations are contaminated by outliers, then the situation completely changes. Huber introduced the so-called gross error model (the contaminated neighborhood of a true distribution) and showed that an optimal estimate is a maximum likelihood estimate constructed for the so-called least favorable distribution. Analytically, this means that we need to solve a minimax problem analogous to the problem given by formula (2.4) below with the asymptotic mean square error as a risk function. In some limiting cases, an optimal estimate is a median,but not a sample mean. In mathematical finance, for most approaches and settings it is implicitly supposed that the underlying asset model is fully specified: the parameters (trend and volatility) of the model are known.Actually, we have all the same to estimate these parameters and construct, say, confidence intervals for them. Hence we only know that a pair (μ, σ) belongs with high probability to the rectangle $[\mu_-, \mu_+] \times [\sigma_-, \sigma_+]$. In that case there arises a problem of construction of robust trading strategies where an optimal strategy is the best strategy against the worst state of Nature. If the risk function of the problem is the expected terminal wealth utility, then our definition of the optimization problem (2.4) is an exact one.

Published in *Finance Stoch.* **17** (2013), no. 3, 535–563.

In 1999, Chen and Epstein introduced a continuous time intertemporal version of a multiple-priors utility function for Brownian filtration. In that case, beliefs are represented by a set $\mathcal{P}$ of probability measures and the utility is defined as a minimum of the expected utilities over the set $\mathcal{P}$. Independently, Cvitanic and Karatzas [7] studied, for a given option, the hedging strategies which minimize the expected "shortfall", i.e. the difference between the payoff and the terminal wealth. They considered the problem of determining the "worst-case" model $\tilde{Q}$, i.e. the model which maximizes a minimal shortfall risk over all possible priors $Q \in \mathcal{P}$. It was shown that under certain assumptions their maximin problem could be written as a minimax problem. In 2004, Quenez [30] studied the problem of utility maximization in an incomplete multiple-priors model, where asset prices are semimartingales. This problem corresponds to a maximin problem where the maximum is taken over the set of feasible wealth X (or portfolios) and where the minimum is taken over the set of priors $\mathcal{P}$. The author showed that, under suitable conditions, there exists a saddle point for this problem. Moreover, Quenez developed the dual approach which consists in solving a dual minimization problem over the set of priors and supermartingale measures and showed how the solution of the dual problem leads to a solution of the primal problem.

The above maximin problems can also be called robust optimization problems since optimization involves an entire class $\mathcal{P}$ of possible probabilistic models and thus takes into account the model risk. Optimal investment problems for such robust utility functionals were considered in particular by Talay and Zheng [33], Quenez [30], Schied [31], Korn and Menkens [23], Gundel [15], Bordigoni [5], Föllmer and Gundel [13], Dokuchaev [12], Hernández-Hernández and Schied [16, 17].

The majority of the relevant published works are concerned with the case where one of the parameters is known exactly. For the unknown drift coefficient, the existence of a saddle point of the corresponding minimax problem was established and the characterization of an optimal strategy obtained in [7, 16, 15]. For the unknown volatility coefficients, the hedging strategy was constructed in [2, 4, 3, 6, 25, 10, 35].

The most difficult case is to characterize the optimal strategy of the maximin problem under the uncertainty of both drift and volatility terms.

Talay and Zheng [33] applied the PDE-based approach to the minimax problem and characterized the value as a viscosity solution of the corresponding Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation. In general, such a problem does not contain a saddle point. Moreover, in robust maximization problems, the maximin should be taken instead of the minimax used by Talay and Zheng. Recently in the work of Denis and Kervarec [11] the general problem of the utility maximization encompassing the case of the uncertain volatility was studied and a duality theory for robust utility maximization in this framework was established.

During the referring process we have found the preprint of Matoussi, Possamai and Zhou [28] which is also devoted to the robust utility maximization problem. To study the exponential, power, and logarithmic utility maximization, the authors use the 2BSDE theory (this theory was thoroughly developed by Cheridito, Soner, Touzi, Victoir and Zhang in works [8, 32]). They obtained explicit solutions in some particular cases, which is one of the tasks of our paper too. Despite some advantages of their approach (non-Markovian models, the existence of a saddle point, a general contingent claim), we should say that that approach is not sufficiently general for our model. Namely

a) only the volatility matrix is misspecified in their model. In our case both coefficients (drift and volatility) are misspecified,

b) the volatility matrix $\sqrt{a_t}$ satisfies the condition $\underline{a} \le a_t \le \overline{a}$, where $\underline{a}$ and $\overline{a}$ are given matrices, which does not cover our "partially misspecified volatility" case since in our paper matrices $a_t = \begin{pmatrix} \sigma_t^2 & \rho\sigma_t \\ \rho\sigma_t & 1 \end{pmatrix}$, $\underline{a} = \begin{pmatrix} \sigma_-^2 & \rho\sigma_- \\ \rho\sigma_- & 1 \end{pmatrix}$ and $\overline{a} = \begin{pmatrix} \sigma_+^2 & \rho\sigma_+ \\ \rho\sigma_+ & 1 \end{pmatrix}$ are non-comparable to each other.

Moreover in the non-Markovian case the BSDE corresponding to our problem won't be 2BSDE (see Remark 3.2). And, besides, we cannot even get our BSDE as a particular case of the 2BSDE given in [28]. So we can conclude that [28] has little in common with our paper.

In this paper, we consider the incomplete diffusion financial market model which resembles the model considered by Schied [31], Hernández-Hernández and Schied [16, 17]. We suppose that the market consists of a risk-free asset, a risky tradable asset with misspecified trend and volatility and a non-tradable asset with known parameters. As different from the approach of Quenez [30] and Schied [31], we solve the maximin problem using the HJBI equation which corresponds to the primal problem. When the trend and volatility coefficients are uncertain, such a maximin problem has no saddle point in general. We extend the set of model coefficients, i.e. carry out some "randomization" and obtain as a result a minimax problem with a saddle point. This makes it possible to replace the maximin problem by a minimax problem which is easier to study using the HJBI equation properties. In particular, we have found such a form of this equation that coincides with the equation derived by Hernández-Hernández and Schied [16] when the volatility is assumed to be known. We establish the solvability of the obtained equation in the classical sense and solve the HJBI equation explicitly for the specific drift coefficient. The saddle point (an optimal portfolio and optimal coefficients) of the considered maximin problem has been found as well. An explicit characterization of the optimal strategies of the maximin problem for the case of power and exponential utilities in terms of the solution of the HJBI equation is the main result of the paper.

To illustrate our approach, we present a simple quadratic hedging problem. Let $(B, B^\perp)$ be the 2-dimensional Brownian motion and $F^B = (\mathcal{F}_t^B)_{t\in[0,T]}$, $F^{B,B^\perp} = (\mathcal{F}_t^{B,B^\perp})_{t\in[0,T]}$ denote the augmented filtrations generated by B and $(B, B^\perp)$, respectively. We consider the filtration $F = (\mathcal{F}_t)_{t\in[0,T]}$ satisfying the usual conditions and $F^B \subset F \subset F^{B,B^\perp}$. Let H be a square integrable $\mathcal{F}_T^B$-measurable random variable. Denote by Π^2 the set of square-integrable predictable processes with respect to the filtration F. Let $\mathcal{P}([\sigma_-, \sigma_+])$ be the set of probability measures on $[\sigma_-, \sigma_+]$ and $\mathcal{U}$, $\tilde{\mathcal{U}}$ denote the set of predictable processes with respect to the filtration F with values in $[\sigma_-, \sigma_+]$ and $\mathcal{P}$,respectively. We use the notation $f \cdot \nu$ for $\int_{\sigma_-}^{\sigma_+} f(\sigma) d\nu(\sigma)$, $f \in C[\sigma_-, \sigma_+]$, $\nu \in \mathcal{P}([\sigma_-, \sigma_+])$. The wealth process corresponding to a portfolio process $\pi \in \Pi^2$ and volatility $\sigma \in \mathcal{U}$ is defined as

$$X_t(\pi, \sigma) = c + \int_0^t \pi_s \sigma_s dB_s. \tag{1.1}$$

The problem is to find $\pi^* \in \Pi^2$ minimizing the worst case mean-variance hedging error

$$\max_{\sigma\in\mathcal{U}} E|H - X_T(\pi^*, \sigma)|^2 = \min_{\pi\in\Pi^2} \max_{\sigma\in\mathcal{U}} E|H - X_T(\pi, \sigma)|^2, \tag{1.2}$$

Such π^* is called a robust hedging strategy.

Let us extend problem (1.2) as follows. For each $\nu \in \tilde{\mathcal{U}}$ we define the processes

$$W_t^\nu = \int_0^t \frac{p \cdot \nu_s}{\sqrt{p^2 \cdot \nu_s}} dB_s + \int_0^t \sqrt{1 - \frac{(p \cdot \nu_s)^2}{p^2 \cdot \nu_s}} dB_s^\perp,$$
$$W_t^{\nu,\perp} = \int_0^t \sqrt{1 - \frac{(p \cdot \nu_s)^2}{p^2 \cdot \nu_s}} dB_s - \int_0^t \frac{p \cdot \nu_s}{\sqrt{p^2 \cdot \nu_s}} dB_s^\perp,$$

where $p,\ p^2$ are the functions $p(\sigma) = \sigma, p^2(\sigma) = \sigma^2$ respectively. One can easily check that $(W^\nu, W^{\nu,\perp})$ is also 2-dimensional Brownian motion and the equation

$$B_t = \int_0^t \frac{p \cdot \nu_s}{\sqrt{p^2 \cdot \nu_s}} dW_s^\nu + \int_0^t \sqrt{1 - \frac{(p \cdot \nu_s)^2}{p^2 \cdot \nu_s}} dW_s^{\nu,\perp} \tag{1.3}$$

is satisfied.

For each $\pi \in \Pi^2,\ \nu \in \tilde{\mathcal{U}}$ we define

$$X_t(\pi, \nu) = c + \int_0^t \pi_s \sqrt{p^2 \cdot \nu_s} dW_s^\nu. \tag{1.4}$$

It is clear that $\mathcal{U} \subset \tilde{\mathcal{U}}$ and for $\nu \in \mathcal{U}$, $W^\nu = B$ and (1.1) coincides with (1.4). Hence we can consider the minimax problem

$$\min_{\pi \in \Pi^2} \max_{\nu \in \tilde{\mathcal{U}}} E|H - X_T(\pi, \nu)|^2, \tag{1.5}$$

which is the extension of problem (1.2).

For the sake of simplicity, it is assumed that $c = EH$ and, using the stochastic integral representation

$$\begin{aligned} H &= EH + \int_0^T h_t dB_t \\ &= EH + \int_0^T h_t \frac{p \cdot \nu_t}{\sqrt{p^2 \cdot \nu_t}} dW_t^\nu + \int_0^T h_t \sqrt{1 - \frac{(p \cdot \nu_t)^2}{p^2 \cdot \nu_t}} dW_t^{\nu,\perp}, \end{aligned}$$

(1.5) is rewritten as

$$\begin{aligned} &\min_{\pi \in \Pi^2} \max_{\nu \in \tilde{\mathcal{U}}} \left[E \int_0^T |h_t \frac{p \cdot \nu_t}{\sqrt{p^2 \cdot \nu_t}} - \pi_t \sqrt{p^2 \cdot \nu_t}|^2 dt + E \int_0^T h_t^2 \left(1 - \frac{(p \cdot \nu_t)^2}{p^2 \cdot \nu_t} \right) dt \right] \\ &= \min_{\pi \in \Pi^2} \max_{\nu \in \tilde{\mathcal{U}}} E \int_0^T [\pi_t^2 (p^2 \cdot \nu_t) - 2 h_t \pi_t (p \cdot \nu_t) + h_t^2] dt. \end{aligned}$$

Since for each $\pi \in \Pi^2$

$$\max_{\nu \in \tilde{\mathcal{U}}} E \int_0^T [\pi_t^2 (p^2 \cdot \nu_t) - 2 h_t \pi_t (p \cdot \nu_t) + h_t^2] dt = \max_{\sigma \in \mathcal{U}} E \int_0^T [\pi_t^2 \sigma_t^2 - 2 h_t \pi_t \sigma_t + h_t^2] dt,$$

we have

$$\min_{\pi \in \Pi^2} \max_{\sigma \in \mathcal{U}} E|H - X_T(\pi, \sigma)|^2 = \min_{\pi \in \Pi^2} \max_{\nu \in \tilde{\mathcal{U}}} E|H - X_T(\pi, \nu)|^2.$$

We will see below that this expression is positive. Moreover,

$$\max_{\sigma\in\mathcal{U}}\min_{\pi\in\Pi^2} E|H - X_T(\pi,\sigma)|^2 = \max_{\sigma\in\mathcal{U}}\min_{\pi\in\Pi^2} E\int_0^T |h_t - \pi_t\sigma_t|^2 dt = 0.$$

This means that the saddle point does not exist for the problem (1.2).

On the other hand, the function G defined on $\Pi^2 \times \tilde{\mathcal{U}}$ by

$$G(\pi,\nu) = E\int_0^T [\pi_t^2(p^2\cdot\nu_t) - 2h_t\pi_t(p\cdot\nu_t) + h_t^2]dt.$$

is convex in π and linear in ν. Then by the Neumann theorem (see Theorem 8 of [1], Chapt. 6) there exists a saddle point $(\pi^*,\sigma^*) \in \Pi^2 \times \tilde{\mathcal{U}}$. Therefore we have

$$\begin{aligned}
0 &= \max_{\sigma\in\mathcal{U}}\min_{\pi\in\Pi^2} E|H - X_T(\pi,\sigma)|^2 \\
&< \min_{\pi\in\Pi^2}\max_{\sigma\in\mathcal{U}} E|H - X_T(\pi,\sigma)|^2 = \min_{\pi\in\Pi^2}\max_{\nu\in\tilde{\mathcal{U}}} E|H - X_T(\pi,\nu)|^2 \\
&= G(\pi^*,\nu^*) = \max_{\nu\in\tilde{\mathcal{U}}}\min_{\pi\in\Pi^2} E|H - X_T(\pi,\nu)|^2 \\
&= \max_{\nu\in\tilde{\mathcal{U}}}\min_{\pi\in\Pi^2}\left[E\int_0^T \left|h_t\frac{p\cdot\nu_t}{\sqrt{p^2\cdot\nu_t}} - \pi_t\sqrt{p^2\cdot\nu_t}\right|^2 dt + E\int_0^T h_t^2\left(1 - \frac{(p\cdot\nu_t)^2}{p^2\cdot\nu_t}\right)dt\right] \\
&= \max_{\nu\in\tilde{\mathcal{U}}} E\int_0^T h_t^2\left(1 - \frac{(p\cdot\nu_t)^2}{p^2\cdot\nu_t}\right)dt.
\end{aligned}$$

It is easy to see that the saddle point is [1)]

$$\nu_t^* = \frac{\sigma_-}{\sigma_+ + \sigma_-}\delta_{\sigma_+} + \frac{\sigma_+}{\sigma_+ + \sigma_-}\delta_{\sigma_-}, \quad \pi_t^* = h_t\frac{p\cdot\nu_t^*}{p^2\cdot\nu_t^*} = \frac{2h_t}{\sigma_- + \sigma_+}.$$

Thus

$$\min_{\pi\in\Pi^2}\max_{\sigma\in\mathcal{U}} E|H - X_T(\pi,\sigma)|^2 = F(\pi^*,\nu^*) = \left(\frac{\sigma_- - \sigma_+}{\sigma_- + \sigma_+}\right)^2 E\int_0^T h_t^2 dt.$$

As we see, the extension of the problem allows us to find the robust strategy and the worst case mean-variance hedging error for the original problem (1.2). In Section 2, we will obtain this result by means of the HJBI equation in the case of a terminal contingent claim $H(B_T)$.

Notice, that the problem (1.2) can be solved also directly, but in more general cases (e.g. for the models with nonzero drift) such “explicit computations” are complicated and in our knowledge does not exist in the literature. The aim of this work is to show that the existence of a saddle point in the extended problem simplifies solving the original problem and enables us to find “explicit solutions”.

The paper is organized as follows. In Section 2, we describe the model and consider the misspecified coefficients as generalized controls. Furthermore, we show the existence of a saddle point of the generalized maximin problem and derive the HJBI equation for the value function. Some examples are also discussed. In Section 3, we prove the solvability in the classical sense of obtained PDE in the case of power and exponential utility and give an explicit PDE-characterization of the robust maximization problem.

[1)] δ_a denotes the measure with support at a point a

2. Generalized Coefficients and the Existence of a Saddle Point

Suppose that the financial market consists of a risk-free asset

$$dS_t^0 = r(Y_t)S_t^0 dt \tag{2.1}$$

with $r(y) \geq 0$ and a risky financial assets whose prices are defined through the stochastic differential equation (SDE)

$$\frac{dS_t}{S_t} = (\tilde{b}(Y_t) + \mu_t)dt + \sigma_t dW_t. \tag{2.2}$$

Here W_t is a standard Brownian motion and Y_t denotes an economical factor process modeled by the SDE

$$dY_t = \beta(Y_t)dt + \left(\rho dW_t + \sqrt{1-\rho^2}dW_t^{\perp}\right), \tag{2.3}$$

for some correlation factor $\rho \in [0,1]$ and standard Brownian motion $W^{\perp}$ which is independent of W. Let $(\mathcal{F}_t)_{t\in[0,T]}$ denote the augmented filtration generated by $W, W^{\perp}$. Denote $b = \tilde{b} - r$ and assume that

A1) $b(y),\ \beta(y),\ r(y)$ belong to $C_b^1(\mathbb{R})$,

A2) $b'(y),\ r'(y)$ belong to $C_0(\mathbb{R})$,

where $C_b^1(\mathbb{R})$ is the class of bounded continuous functions with bounded derivatives and $C_0(\mathbb{R})$ denotes the class of continuous functions with compact support.

Introduce the set $\mathcal{P}(K)$ of probability distributions with support on $K = \times[\sigma_-,\sigma_+]$ ($\mathcal{P}(K)$ is a compact metric space in a weak topology), where $0 \leq \mu_- \leq \mu_+,\ 0 < \sigma_- \leq \sigma_+$. Let $\tilde{\mathcal{U}}_K$ be the set of predictable $\mathcal{P}(K)$-valued processes with respect to filtration $(\mathcal{F}_t)_{t\in[0,T]}$. Such type process usually called the generalized control in control theory [36]. We identify the set of predictable K-valued processes $\mathcal{U}_K$ to the subset of $\tilde{\mathcal{U}}_K$ assigning to each (μ_t,σ_t) from $\mathcal{U}_K$ the $\mathcal{P}(K)$-valued process $\delta_{(\mu_t,\sigma_t)}$.

By Π^2 we denote the set of predictable processes with finite $L^2([0,T]\times\Omega)$-norm. The objective of economic agent is to find the optimal robust strategy of the problem

$$\max_{\pi\in\Pi^2} \min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{\mu,\sigma}(\pi), Y_T), \tag{2.4}$$

with

$$\begin{aligned} dX_t &= r(Y_t)X_t dt + \pi_t(b(Y_t)+\mu_t)dt + \pi_t\sigma_t dW_t, \quad X_0 = x, \\ dY_t &= \beta(Y_t)dt + \rho dW_t + \sqrt{1-\rho^2}dW_t^{\perp}, \quad Y_0 = y, \end{aligned} \tag{2.5}$$

where $U(x,y)$ is a continuous function defined on $\mathbb{R}^2$ satisfying the quadratic growth condition.

If we denote by $f\cdot\nu_t$ the integral $\int_K f(\mu,\sigma)\nu_t(d\mu d\sigma)$, where $f(\mu,\sigma)$ is an arbitrary continuous function, and by $p_\mu,\ p_\sigma$ the functions $p_\mu(\mu,\sigma) = \mu,\ p_\sigma(\mu,\sigma) = \sigma$, respectively, we can consider the following extended maximin problem

$$\max_{\pi\in\Pi^2}\min_{\nu\in\tilde{\mathcal{U}}_K} EU(X_T^\nu(\pi), Y_T^\nu), \tag{2.6}$$

$$\begin{aligned} dX_t &= r(Y_t)X_t dt + \pi_t(b(Y_t) + p_\mu\cdot\nu_t)dt + \pi_t\sqrt{p_\sigma^2\cdot\nu_t}dW_t, \quad X_0 = x,\\ dY_t &= \beta(Y_t)dt + \rho\frac{p_\sigma\cdot\nu_t}{\sqrt{p_\sigma^2\cdot\nu_t}}dW_t + \sqrt{1-\rho^2\frac{(p_\sigma\cdot\nu_t)^2}{p_\sigma^2\cdot\nu_t}}dW_t^\perp, \quad Y_0 = y. \end{aligned} \tag{2.7}$$

As follows from results of [14] there exists the strong solution of (2.7) with $E(\sup_{t\le T}|X_t|^2 + \sup_{t\le T}|Y_t|^2) < \infty$ for each $(\pi,\nu)\in\Pi^2\times\tilde{\mathcal{U}}_K$. Notice that for $(\mu,\sigma)\in\mathcal{U}_K$ the equation (2.7) coincides with (2.5). Our aim is to show that

$$\max_{\pi\in\Pi^2}\min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{\mu,\sigma}(\pi), Y_T) = \max_{\pi\in\Pi^2}\min_{\nu\in\tilde{\mathcal{U}}_K} EU(X_T^\nu(\pi), Y_T^\nu) \tag{2.8}$$

and the latter problem admits a saddle point (π^*,ν^*). It is clear that then π^* will be an optimal robust strategy of the initial problem (2.4),(2.5).

The link between problems (2.4),(2.5) and (2.6),(2.7) will be discussed in Theorem 1 below.

Remark 2.1. Let $\mathcal{B}[0,T]$ be the Borel $\sigma-$algebra on $[0,T]$ and $\tilde{\mathcal{F}}$ be some $\sigma-$algebra with $\mathcal{F}_T\subset\tilde{\mathcal{F}}$. Then the $\mathcal{B}[0,T]\bigotimes\tilde{\mathcal{F}}$-measurable process (μ_t,σ_t) (not necessarily adapted to $(\mathcal{F}_t)_{t\in[0,T]}$) with values in the set K, defines the element $\nu\in\tilde{\mathcal{U}}_K$ by the formula $P((\mu_t,\sigma_t)\in B|\mathcal{F}_t) = \nu_t(B)$. More precisely, denoting ${}^pY$ the predictable projection of a process Y (see [26]), we have the equalities ${}^p\mu_t = \int_K \mu\nu_t(d\mu d\sigma)$, ${}^p\sigma_t = \int_K \sigma\nu_t(d\mu d\sigma)$. Hence instead of (2.7) we can write

$$\begin{aligned} dX_t &= r(Y_t)X_t dt + \pi_t(b(Y_t) + {}^p\mu_t)dt + \pi_t\sqrt{{}^p\sigma_t^2}dW_t, \quad X_0 = x,\\ dY_t &= \beta(Y_t)dt + \rho\frac{{}^p\sigma_t}{\sqrt{{}^p\sigma_t^2}}dW_t + \sqrt{1-\rho^2\frac{({}^p\sigma_t)^2}{{}^p\sigma_t^2}}dW_t^\perp, \quad Y_0 = y. \end{aligned} \tag{2.9}$$

Remark 2.2. The main Theorems of the paper is valid if instead of $\Pi^2\times\tilde{\mathcal{U}}_K$ will be considered the set of Markovian strategies and coefficients, i.e. the set of Borel measurable $\mathbb{R}\times\mathcal{P}(K)$-valued function $(\pi(t,x,y),\nu(t,x,y))$ such that there exist weak solution (X,Y) of (2.9) satisfying condition $E\int_0^T\pi_t^2(X_t,Y_t)dt<\infty$.

Since

$$\begin{aligned} &\begin{pmatrix} \pi_t\sqrt{p_\sigma^2\cdot\nu_t} & 0\\ \rho\frac{p_\sigma\cdot\nu_t}{\sqrt{p_\sigma^2\cdot\nu_t}} & \sqrt{1-\rho^2\frac{(p_\sigma\cdot\nu_t)^2}{p_\sigma^2\cdot\nu_t}} \end{pmatrix} \begin{pmatrix} \pi_t\sqrt{p_\sigma^2\cdot\nu_t} & \rho\frac{p_\sigma\cdot\nu_t}{\sqrt{p_\sigma^2\cdot\nu_t}}\\ 0 & \sqrt{1-\rho^2\frac{(p_\sigma\cdot\nu_t)^2}{p_\sigma^2\cdot\nu_t}} \end{pmatrix}\\ &\qquad = \begin{pmatrix} (p_\sigma^2\cdot\nu_t)\pi_t^2 & \rho(p_\sigma\cdot\nu_t)\pi_t\\ \rho(p_\sigma\cdot\nu_t)\pi_t & 1 \end{pmatrix} \end{aligned} \tag{2.10}$$

the generator of the process (X_t,Y_t) can be given by the function

$$\frac{1}{2}\pi^2(p_\sigma^2\cdot\nu)q_{11} + \rho\pi(p_\sigma\cdot\nu)q_{12} + \frac{1}{2}q_{22} + xr(y)p_1 + \pi b(y)p_1 + \pi(p_\mu\cdot\nu)p_1 + \beta(y)p_2.$$

For all $\nu \in \mathcal{P}(K), \pi \in \mathbb{R}, (\mu, \sigma) \in K$ and $(x, y, p, q) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^3$ we set

$$\begin{aligned}\mathcal{H}^{\pi,\mu,\sigma}(x,y,p,q) &= \frac{1}{2}\,\pi^2\sigma^2 q_{11} + \rho\pi\sigma q_{12}\\ &\quad + \frac{1}{2}\,q_{22} + xr(y)p_1 + \pi b(y)p_1 + \pi\mu p_1 + \beta(y)p_2, \qquad (2.11)\\ \mathcal{H}^{\pi,\nu}(x,y,p,q) &= \mathcal{H}^{\pi,\cdot,\cdot}(x,y,p,q)\cdot\nu \qquad (2.12)\end{aligned}$$

and

$$\mathcal{H}(x,y,p,q) = \max_{\pi\in\mathbb{R}}\ \min_{\nu\in\mathcal{P}(K)} \mathcal{H}^{\pi,\nu}(x,y,p,q). \qquad (2.13)$$

Proposition 2.1. *For each fixed $(x, y, p, q) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^3$, with $q_{11} < 0$ the function $(\pi, \nu) \to \mathcal{H}^{\pi,\nu}(x, y, p, q)$ admits a saddle point (π^*, ν^*), i.e.*

$$\mathcal{H}^{\pi^*,\nu^*}(x,y,p,q) = \max_{\pi\in\mathbb{R}}\ \min_{\nu\in\mathcal{P}(K)} \mathcal{H}^{\pi,\nu}(x,y,p,q) = \min_{\nu\in\mathcal{P}(K)}\ \max_{\pi\in\mathbb{R}} \mathcal{H}^{\pi,\nu}(x,y,p,q). \qquad (2.14)$$

Moreover,

$$\max_{\pi\in\mathbb{R}}\ \min_{\nu\in\mathcal{P}(K)} \mathcal{H}^{\pi,\nu}(x,y,p,q) = \max_{\pi\in\mathbb{R}}\ \min_{(\mu,\sigma)\in K} \mathcal{H}^{\pi,\mu,\sigma}(x,y,p,q). \qquad (2.15)$$

Proof. By the Neumann theorem (see Theorem 8 of [1], Chapt.6) for each fixed point (x, y, p, q) the function of $\pi \in \mathbb{R}$ and $\nu \in \mathcal{P}(K)$

$$(\pi,\nu) \to \mathcal{H}^{\pi,\nu}(x,y,p,q)$$

admits a saddle point (π^*, ν^*), i.e.

$$\max_{\pi\in\mathbb{R}}\ \min_{\nu\in\mathcal{P}(K)} \mathcal{H}^{\pi,\nu}(x,y,p,q) = \min_{\nu\in\mathcal{P}(K)}\ \max_{\pi\in\mathbb{R}} \mathcal{H}^{\pi,\nu}(x,y,p,q) = \mathcal{H}^{\pi^*,\nu^*}(x,y,p,q). \qquad (2.16)$$

It is obvious that

$$\pi^* = -\frac{b(y)p_1 + (p_\mu\cdot\nu^*)p_1 + (p_\sigma\cdot\nu^*)\rho q_{12}}{(p_\sigma^2\cdot\nu)q_{11}}.$$

Moreover, for each continuous function f on K

$$\min_{\nu\in\mathcal{P}(K)} f\cdot\nu = \min_{(\mu,\sigma)\in K} f(\mu,\sigma),$$

since for $\nu^* = \arg\min_\nu f\cdot\nu$ we have $\operatorname{supp}\nu^* \subseteq \{(\mu^*,\sigma^*) | f(\mu^*,\sigma^*) = \min f(\mu,\sigma)\}$. Hence

$$\min_{\nu\in\mathcal{P}(K)} \mathcal{H}^{\pi,\nu}(x,y,p,q) = \min_{(\mu,\sigma)\in K} \mathcal{H}^{\pi,\mu,\sigma}(x,y,p,q)$$

and equality (2.15) is satisfied. □

Now we define the value functions

$$\begin{aligned} v^-(t,x,y) &= \max_{\pi\in\Pi^2}\ \min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{t,x,y}, Y_T^{t,x,y}),\\ v^+(t,x,y) &= \min_{(\mu,\sigma)\in\tilde{\mathcal{U}}_K}\ \max_{\pi\in\Pi^2} EU(X_T^{t,x,y}, Y_T^{t,x,y}). \end{aligned} \qquad (2.17)$$

Since the Isaacs condition is satisfied (by virtue of Proposition 2.1), there exists, as we will see below, a value of the differential game $v \equiv v^+ = v^-$, which will be a solution of the HJBI equation

$$\begin{aligned} &\frac{\partial}{\partial t} v(t,x,y) \\ &+\mathcal{H}(x,y,v_x(t,x,y),v_y(t,x,y),v_{xx}(t,x,y),v_{xy}(t,x,y),v_{yy}(t,x,y)) = 0, \quad (2.18)\\ &v(T,x,y) = U(x,y). \quad (2.19)\end{aligned}$$

The latter equation can be rewritten as

$$\begin{aligned} &\frac{\partial}{\partial t} v(t,x,y) + \frac{1}{2} v_{yy}(t,x,y) + \beta(y) v_y(t,x,y) + x r(y) v_x(t,x,y) \\ &\quad + \min_{\nu\in\mathcal{P}(K)} \max_{\pi\in\mathbb{R}} \Big[\frac{1}{2}(p_\sigma^2\cdot\nu) v_{xx}(t,x,y)\pi^2 + (p_\sigma\cdot\nu)\rho v_{xy}(t,x,y)\pi \\ &\quad + (b(y) + p_\mu\cdot\nu) v_x(t,x,y)\pi\Big] = 0, \quad (2.20)\\ &v(T,x,y) = U(x,y). \quad (2.21)\end{aligned}$$

Simplifying (2.15) we get

$$\begin{aligned} &\min_{\nu\in\mathcal{P}(K)} \max_{\pi\in\mathbb{R}} \Big[\frac{1}{2}(p_\sigma^2\cdot\nu) q_{11}\pi^2 + (p_\sigma\cdot\nu)\rho q_{12}\pi + b(y)p_1\pi + (p_\mu\cdot\nu)p_1\pi\Big] \\ &= \min_{\nu\in\mathcal{P}(K)} \left[\frac{((p_\sigma\cdot\nu)\rho q_{12} + (b(y)+p_\mu\cdot\nu)p_1)^2}{-2(p_\sigma^2\cdot\nu)q_{11}}\right] \\ &= \begin{cases} -\dfrac{p_1^2}{2q_{11}} \min\limits_{\nu\in\mathcal{P}(K)} \left[\dfrac{((p_\sigma\cdot\nu)\kappa + b(y) + p_\mu\cdot\nu)^2}{p_\sigma^2\cdot\nu}\right] & \text{if } p_1 \neq 0, \\ -\dfrac{\rho^2 q_{12}^2}{2\sigma_M q_{11}} & \text{if } p_1 = 0, \end{cases} \quad (2.22)\end{aligned}$$

where we suppose that $q_{11} < 0$ and use the notation $\kappa = \frac{\rho q_{12}}{p_1}$, $\sigma_M = \frac{\sigma_- + \sigma_+}{2}$.

For the sake of simplicity we assume in addition that

A3) $b(y) + \mu_- \geq 0$, for all $y \in \mathbb{R}$.

By $\varphi(z)$ we denote the linear function of $z \in \left[-\frac{\mu_+}{\sigma_-}, -\frac{\mu_-}{\sigma_+}\right]$ with $\varphi(-\frac{\mu_+}{\sigma_-}) = \sigma_-$, $\varphi(-\frac{\mu_-}{\sigma_+}) = \sigma_+$. Then the pair

$$(l(z),m(z)) = \begin{cases} (\mu_+, \frac{\mu_+}{z} + \frac{\sigma_-\sigma_+}{\sigma_M}) & \text{if } z \in \left(-\infty, \dfrac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}\right], \\ (\mu_+, \sigma_-) & \text{if } z \in \left(\dfrac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}, -\dfrac{\mu_+}{\sigma_-}\right], \\ (-z\varphi(z), \varphi(z)) & \text{if } z \in \left(-\dfrac{\mu_+}{\sigma_-}, -\dfrac{\mu_-}{\sigma_+}\right], \\ (\mu_-, \sigma_+) & \text{if } z \in \left(-\dfrac{\mu_-}{\sigma_+}, \dfrac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}\right], \\ (\mu_-, \frac{\mu_-}{z} + \frac{\sigma_-\sigma_+}{\sigma_M}) & \text{if } z \in \left(\dfrac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}, \infty\right) \end{cases} \quad (2.23)$$

is a continuous, piecewise smooth function of $z \in (-\infty, \infty)$.

Proposition 2.2. *There exists $\nu^* \in \mathcal{P}(K)$ of the form $\nu^* = \alpha\delta_{\mu_a,\sigma_-} + (1-\alpha)\delta_{\mu_a,\sigma_+}$, with some $(\alpha, a) \in [0,1] \times \{-,+\}$, such that*

$$\min_{\nu\in\mathcal{P}(K)} \left[\frac{(b(y) + p_\mu \cdot \nu + \kappa p_\sigma \cdot \nu)^2}{p_\sigma^2 \cdot \nu}\right] = \frac{(b(y) + p_\mu \cdot \nu^* + \kappa p_\sigma \cdot \nu^*)^2}{p_\sigma^2 \cdot \nu^*}$$

$$= \begin{cases} \dfrac{\kappa(2(b(y)+\mu_+)\sigma_M + \kappa\sigma_-\sigma_+)}{\sigma_M^2} & \text{if } \kappa \in \left(-\infty, \dfrac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}\right], \\ \dfrac{(b(y) + \mu_+ + \kappa\sigma_-)^2}{\sigma_-^2} & \text{if } \kappa \in \left(\dfrac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}, -\dfrac{\mu_+}{\sigma_-}\right], \\ 0 & \text{if } \kappa \in \left(-\dfrac{\mu_+}{\sigma_-}, -\dfrac{\mu_-}{\sigma_+}\right], \\ \dfrac{(b(y) + \mu_- + \kappa\sigma_+)^2}{\sigma_+^2} & \text{if } \kappa \in \left(-\dfrac{\mu_-}{\sigma_+}, \dfrac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}\right], \\ \dfrac{\kappa(2(b(y)+\mu_-)\sigma_M + \kappa\sigma_-\sigma_+)}{\sigma_M^2} & \text{if } \kappa \in \left(\dfrac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}, \infty\right) \end{cases} \tag{2.24}$$

and $(p_\mu \cdot \nu^*, p_\sigma \cdot \nu^*) = (l(\kappa), m(\kappa))$, where (l, m) is defined by (2.23).

The proof is given in Appendix.

Corollary 1.

$$\min_{\nu\in\mathcal{P}(K)} \left[\frac{(b(y) + p_\mu \cdot \nu)p_1 + (p_\sigma \cdot \nu)\rho q_{12})^2}{-2p_\sigma^2 \cdot \nu q_{11}}\right] = \min_{(\mu,\sigma)\in K} \left[\frac{(b(y)p_1 + \mu p_1 + \sigma\rho q_{12})^2}{-2(2\sigma_M\sigma - \sigma_-\sigma_+)q_{11}}\right]$$

$$\begin{aligned} &= -\frac{\rho q_{12}(2p_1(b(y) + \mu_+)\sigma_M + \rho q_{12}\sigma_-\sigma_+)}{2q_{11}\sigma_M^2}\chi\left(\frac{\rho q_{12}}{p_1} \in \left(-\infty, \frac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}\right]\right) \\ &\quad - \frac{(p_1(b(y) + \mu_+) + \rho q_{12}\sigma_-)^2}{2q_{11}\sigma_-^2}\chi\left(\frac{\rho q_{12}}{p_1} \in \left(\frac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}, -\frac{\mu_+}{\sigma_-}\right]\right) \\ &\quad - \frac{(p_1(b(y) + \mu_-) + \rho q_{12}\sigma_+)^2}{2q_{11}\sigma_+^2}\chi\left(\frac{\rho q_{12}}{p_1} \in \left(-\frac{\mu_-}{\sigma_+}, \frac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}\right]\right) \\ &\quad - \frac{\rho q_{12}(2p_1(b(y) + \mu_-)\sigma_M + \rho q_{12}\sigma_-\sigma_+)}{2q_{11}\sigma_M^2}\chi\left(\frac{\rho q_{12}}{p_1} \in \left(\frac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}, \infty\right)\right) \\ &\quad - \frac{\rho^2 q_{12}^2}{2\sigma_M}\chi(p_1 = 0), \end{aligned} \tag{2.25}$$

where $\chi(A)$ denotes the indicator of a set A.

Proof. It is sufficient to verify that for $\nu_\pm^* = \alpha\delta_{\mu_\pm,\sigma_-} + (1-\alpha)\delta_{\mu_\pm,\sigma_+}$, $0 \le \alpha \le 1$, we get $p_\sigma^2 \cdot \nu_\pm^* = 2\sigma_M(p_\sigma \cdot \nu_\pm^*) - \sigma_-\sigma_+$. □

From this Corollary we obtain that the HJBI equation has the form

$$\frac{\partial}{\partial t}v(t,x,y) + \frac{1}{2}v_{yy}(t,x,y) + \beta(y)v_y(t,x,y) + xr(y)v_x(t,x,y)$$

$$+\min_{(\mu,\sigma)\in K}\frac{(b(y)v_x(t,x,y)+\mu v_x(t,x,y)+\rho\sigma v_{xy}(t,x,y))^2}{-2(2\sigma_M\sigma-\sigma_-\sigma_+)v_{xx}(t,x,y)}=0, \tag{2.26}$$

$$v(T,x,y)=U(x,y). \tag{2.27}$$

A classical solution $v(t,x,y)$ of this equation defines the pair of continuous, piecewise smooth functions of (t,x,y)

$$(\bar{l}(t,x,y),\bar{m}(t,x,y))=\left(l\left(\frac{\rho v_{xy}(t,x,y)}{v_x(t,x,y)}\right),m\left(\frac{\rho v_{xy}(t,x,y)}{v_x(t,x,y)}\right)\right) \tag{2.28}$$

by the formula (2.23).

The following Theorem is the Verification Theorem of [29] adapted to our setting.

Theorem 1 (Verification Theorem)**.** *Let $v(t,x,y)$ be a classical solution of* (2.20), (2.21) *such that $v_{xx}(t,x,y)<0$ and*

$$\begin{aligned}|v(t,x,y)|\le& L(1+|x|+|y|)^p,\quad \left|\frac{v_x(t,x,y)}{v_{xx}(t,x,y)}\right|\le L(1+|x|+|y|),\\ \left|\frac{v_{xy}(t,x,y)}{v_{xx}(t,x,y)}\right|&\le L(1+|x|+|y|),\end{aligned} \tag{2.29}$$

holds for some constants $L>0,\ p\ge 1$. Suppose also that the triplet $(\pi^*(t,x,y),p_\mu\cdot\nu^*(t,x,y),p_\sigma\cdot\nu^*(t,x,y))$ satisfies the Lipschitz condition on each compact subsets of $[0,T]\times\mathbb{R}\times\mathbb{R}$, where

$$\pi^*(t,x,y)=-\frac{(b(y)+p_\mu\cdot\nu^*(t,x,y))v_x(t,x,y)+p_\sigma\cdot\nu^*(t,x,y)\rho v_{xy}(t,x,y)}{(2\sigma_M p_\sigma\cdot\nu^*(t,x,y)-\sigma_-\sigma_+)v_{xx}(t,x,y)}, \tag{2.30}$$

and $(p_\mu\cdot\nu^*(t,x,y),p_\sigma\cdot\nu^*(t,x,y))$ coincides with $(\bar{l}(t,x,y),\bar{m}(t,x,y))$ defined by (2.28)*. Then (π^*,ν^*) is saddle point of the problem* (2.6), (2.7) *and*

$$\begin{aligned}&\max_{\pi\in\Pi^2}\min_{(\mu,\sigma)\in\mathcal{U}_K}EU(X_T^{\mu,\sigma}(\pi),Y_T)\\&\quad=\max_{\pi\in\Pi^2}\min_{\nu\in\tilde{\mathcal{U}}_K}EU(X_T^\nu(\pi),Y_T^\nu)=\min_{\nu\in\tilde{\mathcal{U}}_K}\max_{\pi\in\Pi^2}EU(X_T^\nu(\pi),Y_T^\nu).\end{aligned}$$

Proof. By the definition of (2.30),(2.28) the pair $(\pi^*(t,x,y),\nu^*(t,x,y))$ is a saddle point of the function

$$\begin{aligned}f(t,x,y,\pi,\nu)=&\frac{1}{2}(p_\sigma^2\cdot\nu)v_{xx}(t,x,y)\pi^2+(p_\sigma\cdot\nu)\rho v_{xy}(t,x,y)\pi\\&+(b(y)+p_\mu\cdot\nu)v_x(t,x,y)\pi\end{aligned}$$

for each (t,x,y). It is easy to see that this pair is a continuous, piecewise-smooth function of variables $(t,x,y)\in[0,T]\times\mathbb{R}\times\mathbb{R}$. By the definition, the triplet of functions $(\pi^*(t,x,y),p_\mu\cdot\nu^*(t,x,y),p_\sigma\cdot\nu^*(t,x,y))$ consists of Lipschitz functions on the each compact subset. Since $p_\sigma^2\cdot\nu^*(t,x,y)=2\sigma_M(p_\sigma\cdot\nu^*(t,x,y))-\sigma_-\sigma_+\ge\sigma_-^2$ is satisfied, $\frac{1}{p_\sigma^2\cdot\nu^*(t,x,y)}$ is also the Lipschitz function on the each compact subset. The linear growth condition for the triplet is also satisfied since $|p_\mu\cdot\nu^*(t,x,y)|\le\mu_+$, $\sigma_-\le|p_\sigma\cdot\nu^*(t,x,y)|\le\sigma_+$ and inequalities

$$|\pi^*(t,x,y)|=\left|\frac{(b(y)+p_\mu\cdot\nu^*(t,x,y))v_x(t,x,y)+p_\sigma\cdot\nu^*(t,x,y)\rho v_{xy}(t,x,y)}{(2\sigma_M p_\sigma\cdot\nu^*(t,x,y)-\sigma_-\sigma_+)v_{xx}(t,x,y)}\right|$$

$$\le \frac{\max_y(b(y)+\mu_+)}{\sigma_-^2}\left|\frac{v_x(t,x,y)}{v_{xx}(t,x,y)}\right| + \frac{\rho\sigma_+}{\sigma_-^2}\left|\frac{v_{xy}(t,x,y)}{v_{xx}(t,x,y)}\right|$$
$$\le \bar{L}(1+|x|+|y|)$$

hold for some constant $\bar{L}$ thanks to condition (2.29).

Thus SDE

$$\begin{aligned}
dX_t^* &= r(Y_t^*)X_t^*dt + \pi^*(t,X_t^*,Y_t^*)(b(Y_t^*)+p_\mu\cdot\nu^*(t,X^*,Y^*))dt \\
&\quad + \pi^*(t,X_t^*,Y_t^*)\sqrt{p_\sigma^2\cdot\nu^*(t,X_t^*,Y_t^*)}dW_t, \\
X_0 &= x, \\
dY_t^* &= \beta(Y_t^*)dt + \rho\frac{p_\sigma\cdot\nu^*(t,X_t^*,Y_t^*)}{\sqrt{p_\sigma^2\cdot\nu^*(t,X_t^*,Y_t^*)}}dW_t + \sqrt{1-\rho^2\frac{p_\sigma\cdot\nu^*(t,X_t^*,Y_t^*)^2}{p_\sigma^2\cdot\nu^*(t,X_t^*,Y_t^*)}}dW_t^\perp, \\
Y_0 &= y,
\end{aligned}$$

defining an optimal wealth process has the coefficients which are Lipschitz functions on each $\{(t,x,y):\ |x|\le R,|y|\le R\}$ and satisfy the linear growth condition. Hence there exists unique strong solution of SDE with $E\sup_{t\le T}|X_t^*|^k<\infty,\ E\sup_{t\le T}|Y_t^*|^k<\infty$, for each $k\ge 1$ (see Theorem 2.3, Chapter V of [14]) and $E\int_0^T\pi^{*2}(t,X^*,Y^*)dt<\infty$. For each control pair $(\pi_t,\nu_t)\in\Pi^2\times\tilde{\mathcal{U}}_K$ we denote by $(X_t(\pi^*,\nu),Y_t(\pi^*,\nu)),\ (X_t(\pi,\nu^*),Y_t(\pi,\nu^*))$ the solution of the system (2.7) corresponding to π_t^*,ν_t and π_t,ν_t^* respectively.

Let $\tau_R = T\wedge\inf\{t:|X_t^*|\ge R,|Y_t^*|\ge R\}$. Since

$$\begin{aligned}
&\frac{\partial}{\partial t}v + \mathcal{H}^{\pi^*,\nu^*}(x,y,v_x,v_y,v_{xx},v_{xy},v_{yy}) \\
&\qquad \equiv \frac{\partial}{\partial t}v + \frac{1}{2}v_{yy} + \beta(y)v_y + xr(y)v_x + f(t,x,y,\pi^*,\cdot)\cdot\nu^* = 0
\end{aligned}$$

and $v_x\pi^*,\ v_y$ are the continuous bounded functions on each ball, we can apply Ito's formula to $v(t,X_t^*,Y_t^*)$ and get $v(t,x,y)=Ev(X_{\tau_R}^{*t,x,y},Y_{\tau_R}^{*t,x,y})$. Passing to the limit as $R\to\infty$ we obtain

$$v(t,x,y)=EU(X_T^{*t,x,y},Y_T^{*t,x,y}),$$

since by the integrability of $\sup_{t\le T}|X_t^*|^p+\sup_{t\le T}|Y_t^*|^p$ we have

$$P(\tau_R<T)\le P(\sup_{t\le T}|X_t^*|\ge R,\ \sup_{t\le T}|Y_t^*|\ge R)\to 0 \quad\text{as}\quad R\to\infty$$

and

$$\begin{aligned}
&|Ev(X_T^{*t,x,y},Y_T^{*t,x,y}) - Ev(\tau_R,X_{\tau_R}^{*t,x,y},Y_{\tau_R}^{*t,x,y})| \\
&\quad\le |Ev(X_T^{*t,x,y},Y_T^{*t,x,y})\chi(\tau_R<T)| + |Ev(X_{\tau_R}^{*t,x,y},Y_{\tau_R}^{*t,x,y})\chi(\tau_R<T)| \\
&\quad\le 2LE(1+\sup_{t\le T}|X_t^*|^p+\sup_{t\le T}|Y_t^*|^p)\chi(\tau_R<T)\to 0 \ \text{ as }\ R\to\infty.
\end{aligned}$$

Similarly, using Ito's formula for the processes $v(t,X_t(\pi^*,\nu),Y_t(\pi^*,\nu))$, $v(t,X_t(\pi,\nu^*),Y_t(\pi,\nu^*))$ and taking into account the inequalities

$$\begin{aligned}
f(t,x,y,\pi,\cdot)\cdot\nu^*(t,x,y) &\le f(t,x,y,\pi^*(t,x,y),\cdot)\cdot\nu^*(t,x,y) \\
&\le f(t,x,y,\pi^*(t,x,y),\cdot)\cdot\nu
\end{aligned}$$

we get

$$EU(X_T^{t,x,y}(\pi,\nu^*),Y_T^{t,x,y}(\pi,\nu^*)) \le v(t,x,y) \le EU(X_T^{t,x,y}(\pi^*,\nu),Y_T^{t,x,y}(\pi^*,\nu)).$$

Finally, we obtain

$$\begin{aligned} EU(X_T^{t,x,y}(\pi,\nu^*),Y_T^{t,x,y}(\pi,\nu^*)) &\le EU(X_T^{*t,x,y},Y_T^{*t,x,y})) \\ &\le EU(X_T^{t,x,y}(\pi^*,\nu),Y_T^{t,x,y}(\pi^*,\nu)). \end{aligned}$$

This means that the pair (π^*,ν^*) is a saddle point of problem (2.6).

Since $v(t,x,y) = \inf_{\nu\in\tilde{\mathcal{U}}_K} EU(X_T^{t,x,y}(\pi^*,\nu),Y_T^{t,x,y}(\pi^*,\nu))$ satisfies the HJB equation of the stochastic control problem and

$$\min_{\nu\in\mathcal{P}(K)} f(t,x,y,\pi^*(t,x,y),\cdot)\cdot\nu = \min_{(\mu,\sigma)\in K} f(t,x,y,\pi^*(t,x,y),\mu,\sigma),$$

we conclude that

$$\begin{aligned} v(t,x,y) &= \min_{\nu\in\tilde{\mathcal{U}}_K} EU(X_T^{t,x,y}(\pi^*,\nu),Y_T^{t,x,y}(\pi^*,\nu)) \\ &= \min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{t,x,y}(\pi^*,\mu,\sigma),Y_T^{t,x,y}). \end{aligned}$$

Thus

$$\begin{aligned} &\min_{\nu\in\tilde{\mathcal{U}}_K}\max_{\pi\in\Pi^2} EU(X_T^{t,x,y}(\pi,\nu),Y_T^{t,x,y}(\pi,\nu))) \\ &\quad\le \max_{\pi\in\Pi^2} EU(X_T^{t,x,y}(\pi,\nu^*),Y_T^{t,x,y}(\pi,\nu^*)) \le v(t,x,y) \\ &\quad= \min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{t,x,y}(\pi^*,\mu,\sigma),Y_T^{t,x,y}) \\ &\quad\le \max_{\pi\in\Pi^2}\min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{t,x,y}(\pi,\mu,\sigma),Y_T^{t,x,y}). \end{aligned}$$

On the other hand,

$$\begin{aligned} &\max_{\pi\in\Pi^2}\min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{t,x,y}(\pi,\mu,\sigma),Y_T^{t,x,y}) \\ &\quad\le \min_{\nu\in\tilde{\mathcal{U}}_K}\max_{\pi\in\Pi^2} EU(X_T^{t,x,y}(\pi,\nu),Y_T^{t,x,y}(\pi,\nu)). \end{aligned}$$

Therefore we get that the values of problems (2.4),(2.5) and (2.6), (2.7) are equal to

$$\begin{aligned} &\min_{\nu\in\tilde{\mathcal{U}}_K}\max_{\pi\in\Pi^2} EU(X_T^{t,x,y}(\pi,\nu),Y_T^{t,x,y}(\pi,\nu)) \\ &\quad= \max_{\pi\in\Pi^2}\min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{t,x,y}(\pi,\mu,\sigma),Y_T^{t,x,y}). \qquad\square \end{aligned}$$

Corollary 2. *The optimal strategy of the robust utility maximization problem* (2.4), (2.5) *is given by*

$$\pi^*(t,x,y) = -\frac{(b(y)+\bar{l}(t,x,y))v_x(t,x,y)+\bar{m}(t,x,y)\rho v_{xy}(t,x,y)}{(2\sigma_M\bar{l}(t,x,y)-\sigma_-\sigma_+)v_{xx}(t,x,y)}, \tag{2.31}$$

where the pair $(\bar{l}(t,x,y),\bar{m}(t,x,y))$ *is defined by* (2.28) *and* $v(t,x,y)$ *is a solution of* (2.26), (2.27).

Example 2. Let us consider the robust mean-variance hedging problem with zero drift and unknown volatility

$$\min_{\pi\in\Pi^2}\max_{\sigma_t\in[\sigma_-,\sigma_+]} E(H(Y_T)-X_T(\pi,\sigma))^2, \tag{2.32}$$

$$\begin{aligned} &dX_t = rX_t dt + \pi_t\sigma_t dW_t, \quad X_0 = x,\\ &dY_t = \beta(Y_t)dt + \rho dW_t + \sqrt{1-\rho^2}dW_t^\perp, \quad Y_0 = y. \end{aligned} \tag{2.33}$$

Therefore we have $U(x,y) = -(x-H(y))^2$, $(x,y)\in\mathbb{R}^2$, $\mu_- = \mu_+ = 0$, $r'(y) = 0$. We assume that H is a continuous bounded function. By equation (2.28) we get

$$(p_\mu\cdot\nu^*(t,x,y), p_\sigma\cdot\nu^*(t,x,y)) = \left(0,\ \frac{\sigma_-\sigma_+}{\sigma_M}\right)$$

since $p_\sigma\cdot\nu(t,x,y) = \frac{2\sigma_-\sigma_+}{\sigma_++\sigma_-} = \frac{\sigma_-\sigma_+}{\sigma_M}$ (this means $\nu^*(t,x,y) = \frac{\sigma_-}{\sigma_++\sigma_-}\delta_{(0,\sigma_+)} + \frac{\sigma_+}{\sigma_++\sigma_-}\delta_{(0,\sigma_-)}$). Thus

$$\arg\min_{\sigma\in[\sigma_-,\sigma_+]}\frac{\rho^2\sigma^2}{-2(2\sigma_M\sigma-\sigma_-\sigma_+)} = \frac{\sigma_-\sigma_+}{\sigma_M}$$

and from (2.26) follows

$$\begin{aligned} &\frac{\partial}{\partial t}v(t,x,y) + \frac{1}{2}v_{yy}(t,x,y) + \beta(y)v_y(t,x,y) + xr(y)v_x(t,x,y)\\ &\qquad + \min_{\sigma\in[\sigma_-,\sigma_+]}\frac{\rho^2\sigma^2 v_{xy}^2(t,x,y)}{-2(2\sigma_M\sigma-\sigma_-\sigma_+)v_{xx}(t,x,y)}\\ &\quad\equiv \frac{\partial}{\partial t}v(t,x,y) + \frac{1}{2}v_{yy}(t,x,y) + \beta(y)v_y(t,x,y) + xr(y)v_x(t,x,y)\\ &\qquad - \rho^2\frac{\sigma_-\sigma_+}{2\sigma_M^2}\frac{v_{xy}^2(t,x,y)}{v_{xx}(t,x,y)} = 0, \end{aligned} \tag{2.34}$$

$$v(T,x,y) = -(x-H(y))^2. \tag{2.35}$$

The solution of (2.34), (2.35) can be given as a quadratic polynomial in x

$$v(t,x,y) = -A(t,y)x^2 + 2B(t,y)x - C(t,y),$$

where the triplet (A,B,C) satisfies the system of PDEs

$$\begin{aligned} &\frac{\partial}{\partial t}A(t,y) + \frac{1}{2}A_{yy}(t,y) + \beta(y)A_y(t,y) + 2rA(t,y) + \rho^2\frac{\sigma_-\sigma_+}{2\sigma_M^2}\frac{A_y^2(t,y)}{A(t,y)} = 0,\\ &A(T,y) = 1,\\ &\frac{\partial}{\partial t}B(t,y) + \frac{1}{2}B_{yy}(t,y) + \beta(y)B_y(t,y) + 2rB(t,y) + \rho^2\frac{\sigma_-\sigma_+}{2\sigma_M^2}\frac{A_y(t,y)B_y(t,y)}{A(t,y)} = 0,\\ &B(T,y) = H(y),\\ &\frac{\partial}{\partial t}C(t,y) + \frac{1}{2}C_{yy}(t,y) + \beta(y)C_y(t,y) + \rho^2\frac{\sigma_-\sigma_+}{2\sigma_M^2}\frac{B_y^2(t,y)}{A(t,y)} = 0,\\ &C(T,y) = H^2(y). \end{aligned}$$

The system admits an explicit solution

$$A(t,y) = e^{2r(T-t)}, \quad B(t,y) = e^{2r(T-t)} EH(Y_T^{t,y}),$$

$$C(t,y) = \rho^2 \frac{\sigma_-\sigma_+}{2\sigma_M^2} e^{2r(T-t)} \int_t^T EB_y^2(s, Y_s^{t,y})ds + EH^2(Y_T^{t,y})$$

(notice that $B_y(t,y) = e^{2r(T-t)} EH_y(Y_T^{t,y}) e^{\int_t^T \beta_y(Y_s^{t,y})ds}$, when H is differentiable). The optimal strategy then takes the form

$$\begin{aligned}
\pi^*(t,x,y) &= -\frac{\rho\frac{\sigma_-\sigma_+}{\sigma_M} v_{xy}(t,x,y)}{(2\sigma_M\frac{\sigma_-\sigma_+}{\sigma_M} - \sigma_-\sigma_+)v_{xx}(t,x,y)} \\
&= -\frac{\rho}{\sigma_M}\frac{v_{xy}(t,x,y)}{v_{xx}(t,x,y)} = -\frac{\rho}{\sigma_M}\frac{B_y(t,y) - xA_y(t,y)}{-A(t,y)} \\
&= \frac{\rho}{\sigma_M}\frac{B_y(t,y)}{A(t,y)} = \frac{\rho}{\sigma_M} e^{-2r(T-t)} B_y(t,y).
\end{aligned}$$

$B(t,y) = e^{2r(T-t)} EH(Y_T^{t,y})$ is a classical bounded solution of the corresponding linear parabolic equation with bounded continuous $B_y(t,y)$ and continuous $B_{yy}(t,y)$ (see [14] formulae (5.20)–(5.22) of Chapter VI). It is clear that

$$|v(t,x,y)| \le L(1+|x|^2), \quad \left|\frac{v_x(t,x,y)}{v_{xx}(t,x,y)}\right| \le L(1+|x|), \quad \left|\frac{v_{xy}(t,x,y)}{v_{xx}(t,x,y)}\right| \le L$$

for some $L > 0$ and $B_y(t,y)$ is the locally Lipschitz function. Hence the pair $(\pi^*(t,x,y), \nu^*(t,y))$ satisfies all conditions of Theorem 1.

The case $\rho = 1$, $r = 0$, $\beta \equiv 0$ is discussed in the introduction. In this case the second equation in (2.7) defines the Brownian motion $Y_t = B_t$ for all non-anticipating strategies $\nu_t(Y) \equiv \nu_t(B)$ and (2.7) coincides with (1.3).

In the case of objective function $U(x,y)$ defined on $\mathbb{R}_+ \times \mathbb{R}$ it is convenient to determine the wealth process as a solution of SDE

$$\begin{aligned}
dX_t &= r(Y_t)X_t dt + \pi_t X_t(b(Y_t) + \mu_t)dt + \pi_t X_t \sigma_t dW_t, \quad X_0 = x, \\
dY_t &= \beta(Y_t)dt + \rho dW_t + \sqrt{1-\rho^2} dW_t^\perp, \quad Y_0 = y
\end{aligned} \tag{2.36}$$

The set of admissible strategies Π we define as the set of all predictable processes π such that $\int \pi_s dW_s$ is BMO-martingale (as regards BMO-martingales see [21]). It is clear that for each $(\pi, \mu, \sigma) \in \Pi \times \mathcal{U}_K$, $\int \pi_s \sigma_s dW_s$ is also BMO-martingale, a solution of (2.36) is positive and maximin problem

$$\max_{\pi\in\Pi} \min_{(\mu,\sigma)\in\mathcal{U}_K} EU(X_T^{\mu,\sigma}(\pi), Y_T), \tag{2.37}$$

make sense.

As in previous case of problem (2.4),(2.5) we consider the following extended maximin problem

$$\max_{\pi\in\Pi} \min_{\nu\in\tilde{\mathcal{U}}_K} EU(X_T^\nu(\pi), Y_T^\nu), \tag{2.38}$$

$$
\begin{aligned}
dX_t &= r(Y_t)X_t dt + \pi_t X_t(b(Y_t) + p_\mu \cdot \nu_t)dt + \pi_t X_t \sqrt{p_\sigma^2 \cdot \nu_t} dW_t, \quad X_0 = x, \\
dY_t &= \beta(Y_t)dt + \rho \frac{p_\sigma \cdot \nu_t}{\sqrt{p_\sigma^2 \cdot \nu_t}} dW_t + \sqrt{1 - \rho^2 \frac{(p_\sigma \cdot \nu_t)^2}{p_\sigma^2 \cdot \nu_t}} dW_t^\perp, \quad Y_0 = y.
\end{aligned} \tag{2.39}
$$

It is easy to see that HJBI equation for the value $v(t,x,y)$ of this problem is the solution of the same equation (2.26), but π^* is defined now by

$$
\pi^*(t,x,y) = -\frac{(b(y) + \bar{l}(t,x,y))v_x(t,x,y) + \bar{m}(t,x,y)\rho v_{xy}(t,x,y)}{x(2\sigma_M \bar{l}(t,x,y) - \sigma_- \sigma_+)v_{xx}(t,x,y)}, \tag{2.40}
$$

where the pair $(\bar{l}(t,x,y), \bar{m}(t,x,y))$ is defined by (2.28).

Theorem 1′. *Let* $v(t,x,y)$ *be a classical solution of* (2.20), (2.21) *such that* $v_{xx}(t,x,y) < 0$ *and*

$$
\begin{aligned}
|v(t,x,y)| \le & L(1 + |x| + |y|)^p, \quad \left| \frac{v_x(t,x,y)}{v_{xx}(t,x,y)} \right| \le Lx, \\
\left| \frac{v_{xy}(t,x,y)}{v_{xx}(t,x,y)} \right| & \le Lx, \quad (t,x,y) \in [0,T] \times \mathbb{R}_+ \times \mathbb{R},
\end{aligned} \tag{2.41}
$$

holds for some constants $L > 0, \ p \ge 1$. *Suppose also that the triplet* $(\pi^*(t,x,y), p_\mu \cdot \nu^*(t,x,y), p_\sigma \cdot \nu^*(t,x,y))$ *satisfies the Lipschitz condition on each compact subsets of* $[0,T] \times \mathbb{R}_+ \times \mathbb{R}$, *where*

$$
\pi^*(t,x,y) = -\frac{(b(y) + p_\mu \cdot \nu^*(t,x,y))v_x(t,x,y) + p_\sigma \cdot \nu^*(t,x,y)\rho v_{xy}(t,x,y)}{x(2\sigma_M p_\sigma \cdot \nu^*(t,x,y) - \sigma_- \sigma_+)v_{xx}(t,x,y)}, \tag{2.42}
$$

and $(p_\mu \cdot \nu^*(t,x,y), p_\sigma \cdot \nu^*(t,x,y))$ *coincides with* $(\bar{l}(t,x,y), \bar{m}(t,x,y))$ *defined by* (2.28). *Then* (π^*, ν^*) *is saddle point of the problem* (2.38), (2.39) *and*

$$
\begin{aligned}
&\max_{\pi \in \Pi} \min_{(\mu,\sigma) \in \mathcal{U}_K} EU(X_T^{\mu,\sigma}(\pi), Y_T) \\
&\quad = \max_{\pi \in \Pi} \min_{\nu \in \tilde{\mathcal{U}}_K} EU(X_T^\nu(\pi), Y_T^\nu) = \min_{\nu \in \tilde{\mathcal{U}}_K} \max_{\pi \in \Pi} EU(X_T^\nu(\pi), Y_T^\nu).
\end{aligned}
$$

Proof. The strategy defined by (2.42) is bounded since for all $(t,x,y) \in [0,T] \times \mathbb{R}_+ \times \mathbb{R}$

$$
|\pi^*(t,x,y)| \le \frac{\max_y(b(y) + \mu_+)}{\sigma_-^2} \left| \frac{v_x(t,x,y)}{xv_{xx}(t,x,y)} \right| + \frac{\rho\sigma_+}{\sigma_-^2} \left| \frac{v_{xy}(t,x,y)}{xv_{xx}(t,x,y)} \right| \le \bar{L}
$$

for some constant $\bar{L}$. Hence $(\pi^*(t, X_t^*, Y_t^*))_{t \in [0,T]} \in \Pi$, where (X^*, Y^*) is the corresponding solution of (2.39). The rest of the proof follows the proof of Theorem 1. □

3. Power and Exponential Utility Cases

Now let us consider the robust utility maximization problem with power utility $U(x) = \frac{1}{q}x^q, \ x > 0$ with $0 < q < 1$

$$
\max_{\pi \in \Pi_x^\delta} \min_{(\mu,\sigma) \in \mathcal{U}_K} \frac{1}{q} E(X_T^{\mu,\sigma}(\pi))^q, \tag{3.1}
$$

subject to

$$
\begin{aligned}
&dX_t = r(Y_t)X_t dt + \pi_t X_t(b(Y_t) + \mu_t)dt + \pi_t X_t \sigma_t dW_t, \quad X_0 = x, \\
&dY_t = \beta(Y_t)dt + \rho dW_t + \sqrt{1-\rho^2}dW_t^{\perp}, \quad Y_0 = y.
\end{aligned} \tag{3.2}
$$

In this case, the HJBI equation (2.26),(2.27) gets

$$
\begin{aligned}
&\frac{\partial}{\partial t}v(t,x,y) + \frac{1}{2}v_{yy}(t,x,y) + \beta(y)v_y(t,x,y) + xr(y)v_x(t,x,y) \\
&\qquad + \min_{(\mu,\sigma)\in K} \frac{((b(y)+\mu)v_x(t,x,y) + \rho\sigma v_{xy}(t,x,y))^2}{-2(2\sigma_M\sigma - \sigma_-\sigma_+)v_{xx}(t,x,y)} = 0, \qquad (3.3)\\
&v(T,x,y) = \frac{1}{q}x^q. \qquad (3.4)
\end{aligned}
$$

A solution of this equation is of the form $v(t,x,y) = \frac{1}{q}x^q e^{u(t,y)}$, where u satisfies the equation

$$
\begin{aligned}
&\frac{\partial}{\partial t}u(t,y) + \frac{1}{2}u_{yy}(t,y) + \beta(y)u_y(t,y) + \frac{1}{2}u_y^2(t,y) + qr(y) \\
&\qquad - \frac{q}{2(q-1)} \min_{(\mu,\sigma)\in K} \frac{(b(y)+\mu+\rho\sigma u_y(t,y))^2}{2\sigma_M\sigma - \sigma_-\sigma_+} = 0, \qquad (3.5)\\
&u(T,y) = 0. \qquad (3.6)
\end{aligned}
$$

The pair $(p_\mu \cdot \nu^*(t,x,y), p_\sigma \cdot \nu^*(t,x,y))$ from Theorem $1'$ takes the form

$$
(p_\mu \cdot \nu^*(t,y), p_\sigma \cdot \nu^*(t,y)) = (l(\rho u_y(t,y)), m(\rho u_y(t,y))), \tag{3.7}
$$

where (l,m) is defined by (2.23).

Remark 3.1. By Corollary 1 and (2.28), equation (3.5) can be written as

$$
\begin{aligned}
&\frac{\partial}{\partial t}u(t,y) + \frac{1}{2}u_{yy}(t,y) + \beta(y)u_y(t,y) + \frac{1}{2}u_y^2(t,y) + qr(y) \\
&\quad - \frac{q\rho u_y(t,y)}{2(q-1)\sigma_M^2}(2(b(y)+\mu_+)\sigma_M + \sigma_-\sigma_+\rho u_y(t,y))\chi\left(\rho u_y(t,y) \le \frac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}\right) \\
&\quad - \frac{q}{2(q-1)\sigma_-^2}(b(y)+\mu_+ + \rho\sigma_- u_y(t,y))^2\chi\left(\frac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-} < \rho u_y(t,y) \le -\frac{\mu_+}{\sigma_-}\right) \\
&\quad - \frac{q}{2(q-1)\sigma_+^2}(b(y)+\mu_- + \rho\sigma_+ u_y(t,y))^2\chi\left(-\frac{\mu_-}{\sigma_+} < \rho u_y(t,y) \le \frac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}\right) \\
&- \frac{q\rho u_y(t,y)}{2(q-1)\sigma_M^2}(2(b(y)+\mu_-)\sigma_M + \sigma_-\sigma_+\rho u_y(t,y)) \\
&\qquad\qquad \times \chi\left(\rho u_y(t,y) > \frac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}\right) = 0, \qquad (3.8)\\
&u(T,y) = 0. \qquad (3.9)
\end{aligned}
$$

Theorem 3. *Under conditions* A1)–A3) *the Cauchy problem* (3.5), (3.6) *admits a classical solution with bounded* $u_y(t,y)$ *and a saddle point* $(\nu^*(t,y), \pi^*(t,y))$ *of the problem*

(2.38), (2.39) *is defined by equation* (3.7) *and by the formula*

$$\pi^*(t,y) = \frac{1}{1-q}\left(\frac{b(y)+p_\mu\cdot\nu^*(t,y)}{p_\sigma^2\cdot\nu^*(t,y)} + \rho\frac{p_\sigma\cdot\nu^*(t,y)}{p_\sigma^2\cdot\nu^*(t,y)}u_y(t,y)\right). \tag{3.10}$$

Moreover, $\pi^*(t,y)$ *is the optimal strategy of robust utility maximization problem* (3.1), (3.2).

Proof. By the Proposition of Appendix B there exists a classical solution of (3.5),(3.6) with bounded $u_y(t,y)$. From the continuity of $u_{yy}(t,y)$ follows that $u_y(t,y)$ is the locally Lipschitz function. By the Lemma A.1 the pair $(p_\mu\cdot\nu^*(t,y), p_\sigma\cdot\nu^*(t,y))$, where $\nu^*(t,y)$ defined by (3.7), is the locally Lipschitz function. Since $p_\sigma^2\cdot\nu^*(t,y) = 2\sigma_M(p_\sigma\cdot\nu^*(t,y)) - \sigma_-\sigma_+ \geq \sigma_-^2$ is satisfied, $\frac{1}{p_\sigma^2\cdot\nu^*(t,y)}$ is also the locally Lipschitz function. Hence

$$\begin{aligned}
\pi^*(t,y) &= -\frac{1}{q-1}\frac{b(y)+l(\rho u_y(t,y)) + m(\rho u_y(t,y))\rho u_y(t,y)}{2m(\rho u_y(t,y))\sigma_M - \sigma_-\sigma_+}\\
&= \frac{1}{1-q}\frac{b(y)+p_\mu\cdot\nu^*(t,y)+p_\sigma\cdot\nu^*(t,y)\rho u_y(t,y)}{2p_\sigma\cdot\nu^*(t,y)\sigma_M-\sigma_-\sigma_+}\\
&= \frac{1}{1-q}\left(\frac{b(y)+p_\mu\cdot\nu^*(t,y)}{p_\sigma^2\cdot\nu^*(t,y)} + \rho\frac{p_\sigma\cdot\nu^*(t,y)}{p_\sigma^2\cdot\nu^*(t,y)}u_y(t,y)\right)
\end{aligned} \tag{3.11}$$

following from (2.30) is also the Lipschitz function. It is obvious that $\pi^*\in\Pi$ for each $\nu\in\mathcal{U}_K$ (since $X(\pi^*,\nu)$ is a solution of the linear SDE), $v_{xx}(t,x,y) = (q-1)x^{q-2}e^{u(t,y)} < 0$ and all the conditions of Theorem $1'$ are satisfied. Therefore we can conclude that $(\pi^*(t,y),\nu^*(t,y))$ is the saddle point of the problem (3.1), (2.9). □

Corollary 3. *If* $b=0$, $r=0$, *then*

$$u(t,y) = -\frac{q}{2(q-1)}(T-t)\min_{(\mu,\sigma)\in K}\frac{\mu^2}{2\sigma_M\sigma-\sigma_-\sigma_+} = -\frac{q}{2(q-1)}(T-t)\frac{\mu_-^2}{\sigma_+^2}$$

is a solution of (3.5) *and a saddle point of the maximin problem can be given explicitly*

$$(\mu_t^*,\sigma_t^*) = (\mu_-,\sigma_+),\quad \pi^*(t,x,y) = -\frac{\mu_-}{2(q-1)\sigma_+^2}x.$$

Example 4. When $\sigma_- = \sigma_+ = \sigma_M$ we obtain

$$\begin{aligned}
&\frac{\partial}{\partial t}u(t,y) + \frac{1}{2}u_{yy}(t,y) + \beta(y)u_y(t,y) + \frac{1}{2}u_y^2(t,y)\\
&\quad - \frac{q}{2(q-1)\sigma_M^2}\min_{\mu_-\leq\mu\leq\mu_+}(b(y)+\mu+\rho\sigma_M u_y(t,y))^2\\
&\equiv \frac{\partial}{\partial t}u(t,y) + \frac{1}{2}u_{yy}(t,y) + (2\rho\sigma_M b(y)+\beta(y))u_y(t,y) + \frac{1}{2}\left(1-\frac{q\rho^2\sigma_M}{q-1}\right)u_y^2(t,y)\\
&\quad - \frac{q}{2(q-1)\sigma_M^2}\min_{\mu_-\leq\mu\leq\mu_+}((b(y)+\mu)^2 + 2\mu\rho\sigma_M u_y(t,y)) = 0,\\
&u(T,y) = 0.
\end{aligned}$$

Applications of such type equations in finance and the existence of a classical solution are discussed in [16].

Remark 3.2. Instead of PDE (3.5) we can use the BSDE with quadratic growth

$$\begin{aligned} dV_t &= -\bigg(\frac{1}{2}Z_t^2 + qr(Y_t) \\ &\quad - \frac{q}{2(q-1)} \min_{(\mu,\sigma)\in K} \frac{(b(Y_t)+\mu+\rho\sigma Z_t)^2}{2\sigma_M\sigma - \sigma_-\sigma_+}\bigg)dt + Z_t dW_t + Z_t^{\perp} dW_t^{\perp}, \\ V_T &= 0. \end{aligned}$$

solvability of which follows from the results of [22, 34]. The solution of the BSDE can be constructed using the solution of PDE (3.5) by the formulas

$$V_t = u(t,Y_t),\ Z_t = \rho u_y(t,Y_t),\ Z_t^{\perp} = \sqrt{1-\rho^2}u_y(t,Y_t).$$

The optimal strategy $\pi_t^* = \pi^*(t,Y_t)$ is defined by the linear equation

$$\pi_t^* = \frac{1}{1-q}\left(\frac{b(Y_t)+p_\mu\cdot\hat{\nu}_t^*(Z_t)}{p_\sigma^2\cdot\hat{\nu}_t^*(Z_t)} + \frac{p_\sigma\cdot\hat{\nu}_t^*(Z_t)}{p_\sigma^2\cdot\hat{\nu}_t^*(Z_t)}Z_t\right),$$

following from (3.10). As follows from (3.7), the pair $(p_\mu\cdot\hat{\nu}_t^*(z), p_\sigma\cdot\hat{\nu}_t^*(z))$ coincides with $(l(z), m(z))$ defined by (2.23). ☐

Suppose now $U(x,y) = -e^{-\gamma(x-H(y))}, (x,y)\in\mathbb{R}^2,\ \gamma>0$ and $r=0$. This case corresponds to the exponential hedging problem of the contingent claim $H(y)$, depending only on the non-tradable asset. We assume that $H\in C_b(\mathbb{R})$. Now following [27] we consider the restricted class of strategies $\Pi = \{\pi\in\Pi^2 : \int_0^t \pi_s dW_s$ is BMO-martingale$\}$ and minimax problem

$$\min_{\pi\in\Pi}\max_{(\mu,\sigma)\in\mathcal{U}_K} Ee^{\gamma(H(Y_T)-X_T^{\mu,\sigma}(\pi))} \tag{3.12}$$

subject to (2.5). It is easy to verify that a solution of (2.26),(2.27) is of the form $v(t,x,y) = -e^{\gamma u(t,y)-\gamma x}$, where $u(t,y)$ is a bounded solution of

$$\begin{aligned} &\frac{\partial}{\partial t}u(t,y) + \frac{1}{2}u_{yy}(t,y) + \beta(y)u_y(t,y) + \frac{1}{2}\gamma u_y^2(t,y) \\ &\qquad + \frac{1}{2\gamma}\min_{(\mu,\sigma)\in K}\frac{(b(y)+\mu+\rho\gamma\sigma u_y(t,y))^2}{2\sigma_M\sigma-\sigma_-\sigma_+} = 0, \end{aligned} \tag{3.13}$$

$$u(T,y) = H(y). \tag{3.14}$$

The existence of a classical bounded solution of (3.13),(3.14) with bounded u_y for the case

$$H' \in C_0(\mathbb{R}) \tag{3.15}$$

follows from Proposition B.1. Thus $\frac{v_x(t,x,y)}{v_{xx}(t,x,y)} = -\frac{1}{\gamma}$, $\frac{v_{xy}(t,x,y)}{v_{xx}(t,x,y)} = -u_y(t,y)$ are bounded. One can checks that all conditions of Theorem 1 except of the polynomial growth condition of $v(t,x,y)$ are satisfied. The robust optimal portfolio is

$$\pi^*(t,y) = -\frac{1}{\gamma}\frac{b(y)+p_\mu\cdot\nu^*(t,y)-\gamma\rho p_\sigma\cdot\nu^*(t,y)u_y(t,y)}{2\sigma_M p_\sigma\cdot\nu^*(t,y)-\sigma_-\sigma_+},$$

where $(p_\mu\cdot\nu^*(t,y), p_\sigma\cdot\nu^*(t,y))$ is defined by (2.28). Thus $(\pi^*(t,y), p_\mu\cdot\nu^*(t,y), p_\sigma\cdot\nu^*(t,y))$ is the bounded, locally Lipschitz function of (t,y) and $X_s^{t,x,y}(\pi^*,\nu^*)$, $s\geq t$, is BMO-martingale. Hence $\{e^{\gamma(X_\tau^{t,x,y}(\pi^*,\nu^*)-u(\tau,Y_\tau^{t,x,y}))}$, τ is stopping times, $t\leq\tau\leq T\}$ is uniformly integrable family of random variables. This enable us to pass in the limit in the Theorem 1

$$Ev(X_T^{*t,x,y}, Y_T^{*t,x,y}) = Ee^{\gamma(X_{\tau_R}^{*t,x,y} - u(\tau_R, Y_{\tau_R}^{*t,x,y}))} \to Ee^{\gamma(X_T^{*t,x,y} - u(T, Y_T^{*t,x,y}))}$$
$$= Ev(T, X_T^{*t,x,y}, Y_T^{*t,x,y}) \quad \text{as} \quad R \to \infty$$

without the polynomial growth condition of $v(t,x,y)$. Hence we have proved

Theorem 5. *Under conditions* A1)–A3) *and* (3.15) *the Cauchy problem* (3.13), (3.14) *admits a classical solution with bounded* $u_y(t,y)$ *and a saddle point* $(\pi^*(t,y), \nu^*(t,y))$ *of the problem* (3.12) *is defined by the equation*

$$(p_\mu \cdot \nu^*(t,y), p_\sigma \cdot \nu^*(t,y)) = (l(\rho\gamma u_y(t,y)), m(\rho\gamma u_y(t,y)))$$

and by the formula

$$\pi^*(t,y) = -\frac{1}{\gamma} \frac{b(y) + p_\mu \cdot \nu^*(t,y) - \gamma\rho(p_\sigma \cdot \nu^*(t,y))u_y(t,y)}{2\sigma_M p_\sigma \cdot \nu^*(t,y) - \sigma_- \sigma_+}.$$

Moreover $\pi^*(t,y)$ *is the optimal strategy of the robust exponential hedging problem* (3.12), (2.5).

Appendix A

Each measure ν may be realized as a distribution of a pair of random variables (ξ, η) with the value in K. Simplifying the notation we denote $b(y) + \mu$ by μ again. Our aim is to characterize the dependence of the minimizer of the problem

$$\min_{\nu \in \mathcal{P}(K)} \left[\frac{(p_\mu \cdot \nu + \kappa p_\sigma \cdot \nu)^2}{p_\sigma^2 \cdot \nu} \right] = \min_{(\xi,\eta) \in K} \left[\frac{(E\xi + \kappa E\eta)^2}{E\eta^2} \right]$$

on a parameter $\kappa \in \mathbb{R}$.

Proposition A.1. *Let*

$$(\xi^*, \eta^*) = \arg \min_{(\xi,\eta) \in K} \left[\frac{(E\xi + \kappa E\eta)^2}{E\eta^2} \right].$$

Then ξ^* *is a number,* η^* *is the Bernoulli random variable with value in the set* $\{\sigma_-, \sigma_+\}$ *and the expectation of the pair* (ξ^*, η^*) *is given by the formula*

$$(\xi^*, E\eta^*) = \begin{cases} \left(\mu_+, \dfrac{\mu_+}{\kappa} + \dfrac{\sigma_-\sigma_+}{\sigma_M}\right) & \text{if } \kappa \in \left(-\infty, \dfrac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}\right], \\ (\mu_+, \sigma_-) & \text{if } \kappa \in \left(\dfrac{\mu_+\sigma_M}{\sigma_M\sigma_- - \sigma_+\sigma_-}, -\dfrac{\mu_+}{\sigma_-}\right], \\ (\kappa, -1)\text{constant} & \text{if } \kappa \in \left(-\dfrac{\mu_+}{\sigma_-}, -\dfrac{\mu_-}{\sigma_+}\right], \\ (\mu_-, \sigma_+) & \text{if } \kappa \in \left(-\dfrac{\mu_-}{\sigma_+}, \dfrac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}\right], \\ \left(\mu_-, \dfrac{\mu_-}{\kappa} + \dfrac{\sigma_-\sigma_+}{\sigma_M}\right) & \text{if } \kappa \in \left(\dfrac{\mu_-\sigma_M}{\sigma_M\sigma_+ - \sigma_+\sigma_-}, \infty\right). \end{cases}$$

Moreover,

$$\frac{(\xi^*+\kappa E\eta^*)^2}{E\eta^{*2}}=\begin{cases}\dfrac{\kappa(2\mu_+\sigma_M+\kappa\sigma_-\sigma_+)}{\sigma_M^2} & \text{if } \kappa\in\left(-\infty,\dfrac{\mu_+\sigma_M}{\sigma_M\sigma_--\sigma_+\sigma_-}\right],\\ \dfrac{(\mu_++\kappa\sigma_-)^2}{\sigma_-^2} & \text{if } \kappa\in\left(\dfrac{\mu_+\sigma_M}{\sigma_M\sigma_--\sigma_+\sigma_-},-\dfrac{\mu_+}{\sigma_-}\right],\\ 0 & \text{if } \kappa\in\left(-\dfrac{\mu_+}{\sigma_-},-\dfrac{\mu_-}{\sigma_+}\right],\\ \dfrac{(\mu_-+\kappa\sigma_+)^2}{\sigma_+^2} & \text{if } \kappa\in\left(-\dfrac{\mu_-}{\sigma_+},\dfrac{\mu_-\sigma_M}{\sigma_M\sigma_+-\sigma_+\sigma_-}\right],\\ \dfrac{\kappa(2\mu_-\sigma_M+\kappa\sigma_-\sigma_+)}{\sigma_M^2} & \text{if } \kappa\in\left(\dfrac{\mu_-\sigma_M}{\sigma_M\sigma_+-\sigma_+\sigma_-},\infty\right).\end{cases}$$

Proof. Let $(\mu_++\kappa\sigma_-)(\mu_-+\kappa\sigma_+)\le 0$. Then by the continuity of a function $\mu+\kappa\sigma$, $(\mu,\sigma)\in K$, there exists $(\hat\mu,\hat\sigma)$ such that $\hat\mu+\kappa\hat\sigma=0$. Thus $(\hat\mu,\hat\sigma)$ is proportional to $(\kappa,-1)$ and $\left[\frac{(E\xi^*+\kappa E\eta^*)^2}{E\eta^{*2}}\right]=0$. If $(\mu_++\kappa\sigma_-)(\mu_-+\kappa\sigma_+)>0$, then either $\kappa>\frac{\mu_-}{\sigma_+}$ and $\xi^*=\mu_-$ or $\kappa<-\frac{\mu_+}{\sigma_-}$ and $\xi^*=\mu_+$. Thus it is sufficient to study the minimization problem

$$\min_{\eta\in[\sigma_-,\sigma_+]}\left[\frac{(\mu_a+\kappa E\eta)^2}{E\eta^2}\right]\quad\text{for}\quad a=+,-.$$

Now we show that η^* is of the form $\eta^*=\sigma_-\chi_B+\sigma_+\chi_{B^c}$ for some event B. Indeed, if $E\eta^*=y$, then $E\eta^{*2}=2\sigma_My-\sigma_-\sigma_+$ and η^* is the maximizer of the problem $\max_{\eta,E\eta=y}E\eta^2$, since for any η, with $E\eta=y$ we have

$$\begin{aligned}E\eta^2&=E(\eta-\sigma_M)^2+2\sigma_My-\sigma_M^2\\&\le\left(\frac{\sigma_+-\sigma_-}{2}\right)^2+2\sigma_My-\sigma_M^2=2\sigma_My-\sigma_-\sigma_+=E\eta^{*2}.\end{aligned}$$

Hence

$$\min_{\eta\in[\sigma_-,\sigma_+]}\left[\frac{(\mu_a+\kappa E\eta)^2}{E\eta^2}\right]=\min_{\sigma_-\le y\le\sigma_+}\psi_a(y),$$

where $\psi_a(y)=\frac{(\mu_a+\kappa y)^2}{2\sigma_My-\sigma_-\sigma_+}$. Since

$$\psi_a'(y)=\frac{\kappa^2}{2\sigma_M}-\frac{\kappa^2}{2\sigma_M}\frac{(2\sigma_M\frac{\mu_a}{\kappa}+\sigma_-\sigma_+)^2}{(2\sigma_My-\sigma_-\sigma_+)^2}$$

the equation $\psi_a'(y)=0$ has two roots:

$$y_1^a=-\frac{\mu_a}{\kappa},\quad y_2^a=\frac{\mu_a}{\kappa}+\frac{\sigma_-\sigma_+}{\sigma_M}.$$

If $y_1^a=-\frac{\mu_a}{\kappa}\in[\sigma_-,\sigma_+]$, then $y_2^a=\frac{\mu_a}{\kappa}+\frac{\sigma_-\sigma_+}{\sigma_M}\in[-\sigma_++\frac{\sigma_-\sigma_+}{\sigma_M},-\sigma_-+\frac{\sigma_-\sigma_+}{\sigma_M}]$ and vice versa. Moreover, $[\sigma_-,\sigma_+]\cap[-\sigma_++\frac{\sigma_-\sigma_+}{\sigma_M},-\sigma_-+\frac{\sigma_-\sigma_+}{\sigma_M}]=\emptyset$. Since $\lim_{y\to\pm\infty}\psi_a(y)=\pm\infty$, the smallest root is the maximizer and the biggest one is the minimizer. The case of $y_1^a\in[\sigma_-,\sigma_+]$ is equivalent to the case $\kappa\in\left[-\frac{\sigma_+}{\mu_a},-\frac{\sigma_-}{\mu_a}\right]$, which yields $\min\psi_a(y)=\psi_a(y_1^a)=0$. From the relation $y_2^a\in[\sigma_-,\sigma_+]$ follows that $-\sigma_++\frac{\sigma_-\sigma_+}{\sigma_M}\le-\frac{\mu_a}{\kappa}\le-\sigma_+-\frac{\sigma_-\sigma_+}{\sigma_M}$, which is equivalent to the relation $\kappa \in \big(-\infty, \frac{\mu_a}{\sigma_- - \frac{\sigma_-\sigma_+}{\sigma_M}}\big] \cup \big[\frac{\mu_a}{\sigma_+ - \frac{\sigma_-\sigma_+}{\sigma_M}}, \infty\big)$. In that case, $\min_{\sigma_- \le y \le \sigma_+} \psi_a(y) = \psi_a(y_2^a) = \kappa \frac{2\mu_a + \kappa\sigma_-\sigma_+}{\sigma_M^2}$.

Now we will consider step by step all the possibilities of displacement of κ in the intervals formulated in the proposition.

1) $\kappa \in (-\infty, \frac{\mu_a}{\sigma_- - \frac{\sigma_-\sigma_+}{\sigma_M}}]$. Since $\frac{\mu_a}{\sigma_- - \frac{\sigma_-\sigma_+}{\sigma_M}} \le -\frac{\mu_+}{\sigma_-}$, we have $\kappa \in (-\infty, -\frac{\mu_+}{\sigma_-}]$ and $\xi^* = \mu_+$. Moreover, $\min \psi_+(y) = \psi_+(y_2^+) = \kappa \frac{2\mu_+ + \kappa\sigma_-\sigma_+}{\sigma_M^2}$.

2) $\kappa \in (\frac{\mu_+}{\sigma_- - \frac{\sigma_-\sigma_+}{\sigma_M}}, -\frac{\mu_+}{\sigma_-}]$. From $\kappa \le -\frac{\mu_+}{\sigma_-}$ it follows that $y_1^+ = -\frac{\mu_+}{\kappa} < \sigma_-$ and from $\kappa > \frac{\mu_+}{\sigma_- - \frac{\sigma_-\sigma_+}{\sigma_M}}$ it follows that $y_2^+ = \frac{\mu_+}{\kappa} + \frac{\sigma_-\sigma_+}{\sigma_M} < \sigma_-$. Hence $\psi_+(y)$ is increasing on $[\sigma_-, \sigma_+]$ and $\arg\min_{\sigma_- \le y \le \sigma_+} \psi_+(y) = \sigma_-$.

3) $\kappa \in (-\frac{\mu_+}{\sigma_-}, -\frac{\mu_-}{\sigma_+}]$. Then $y_1^+ = -\frac{\mu_+}{\kappa} \in [\sigma_-, \sigma_+]$ and $\min \psi_+(y) = 0$.

4) $\kappa \in (-\frac{\mu_-}{\sigma_+}, \frac{\mu_-}{\sigma_+ - \frac{\sigma_-\sigma_+}{\sigma_M}}]$. Then $\frac{\mu_-}{\kappa} > \sigma_+ - \frac{\sigma_-\sigma_+}{\sigma_M}$ and $y_1^- = -\frac{\mu_-}{\kappa} < -\sigma_+ + \frac{\sigma_-\sigma_+}{\sigma_M} < \sigma_-$, $y_2^- = \frac{\mu_-}{\kappa} + \frac{\sigma_-\sigma_+}{\sigma_M} > \sigma_+$. Hence $\psi_-(y)$ is decreasing on $[\sigma_-, \sigma_+]$ and $\arg\min \psi_+(y) = \sigma_+$.

5) $\kappa \in (\frac{\mu_-}{\sigma_+ - \frac{\sigma_-\sigma_+}{\sigma_M}}, \infty]$. Then $\kappa > \frac{\mu_-}{\sigma_+}$ and $\xi^* = \mu_-$. On the other hand, from $\frac{\mu_-}{\kappa} < \sigma_+ - \frac{\sigma_-\sigma_+}{\sigma_M}$ it follows that $y_2^- \in [\sigma_-, \sigma_+]$. Hence $\min_{\sigma_- \le y \le \sigma_+} \psi_-(y) = \psi_-(y_2^-)$. □

Lemma A.1. *Let $u(y)$, $f_1(z), f_2(z), \ldots, f_N(z)$ be Lipschitz functions and $-\infty = a_0 < a_1 < \cdots < a_N = \infty$ are such points that $f_k(a_k) = f_{k+1}(a_k)$, $k = 1, \ldots, N-1$. Then the function*

$$\nu(y) = f_k(u(y)), \ \text{if} \ a_{k-1} < u(y) \le a_k, \ k = 2, \ldots, N,$$

is also a Lipschitz function.

Proof. For the sake of simplicity we consider the case $N = 2$. It is clear that $f_k(u(y))$, $k = 1, 2, 3$, are Lipschitz functions, i.e. there exists a constant $C > 0$ such that $|f_k(u(y_1) - f_k(u(y_2))| \le C|y_1 - y_2|$. Suppose $A_1 = \{y : u(y) \le a_1\}$, $A_2 = \{y : u(y) > a_1\}$ and set $y_1 \in A_1$ $y_2 \in A_2$. Since $u(y_1) \le a_1 \le u(y_2)$, by the continuity of u there exists $\bar{y}$ such that $u(\bar{y}) = a_1$, $y_1 \le \bar{y} \le y_2$. Hence we have

$$\begin{aligned}|\nu(y_1) - \nu(y_2)| &= |f_1(u(y_1)) - f_2(u(y_2))| = |f_1(u(y_1)) - f_1(a_1) + f_2(a_2) - f_2(u(y_2))| \\ &\le |f_1(u(y_1)) - f_1(u(\bar{y}))| + |f_2(u(\bar{y})) - f_2(u(y_2))| \\ &\le C|y_1 - \bar{y}| + C|y_2 - \bar{y}| = C(y_2 - y_1).\end{aligned}$$ □

APPENDIX B

Let $\beta, a, b, H \in C_b(\mathbb{R})$ and γ, c, g be some constants. We consider the Cauchy problem

$$\frac{\partial}{\partial t}u(t,y) + \frac{1}{2}u_{yy}(t,y) + \beta(y)u_y(t,y) + \gamma u_y^2(t,y) + a(y) + c \min_{(\mu,\sigma)\in K} \frac{(b(y) + \mu + g\sigma u_y(t,y))^2}{2\sigma_M\sigma - \sigma_-\sigma_+} = 0, \tag{B.1}$$

$$u(T,y) = H(y). \tag{B.2}$$

Proposition B.1. *Let β, a, b, H be such that $a', b', H' \in C_0(\mathbb{R})$. Then the Cauchy problem $(B.1)$, $(B.2)$ admits a classical solution with bounded $u(t,y)$, $u_y(t,y)$.*

Proof. By condition of the proposition there exists $N \geq 0$ such that $a'(y), b'(y) = 0,\ H'(y) = 0$, if $|y| > N$. Thus $a(y) = a^+,\ b(y) = b^+,\ H(y) = H^+$, if $y \geq N$ and $a(y) = a^-,\ b(y) = b^-,\ H(y) = H^-$, if $y \leq -N$ for some constants $a^+, a^-, b^+, b^-, H^+, H^-$. The solutions of (3.5) on the intervals $(-\infty, -N]$ and $[N, \infty)$ are $u^-(t) = a^-(T-t) + c\frac{(b^- + \mu_-)^2}{\sigma_+^2}(T-t) + H^-$ and $u^+(t) = a^+(T-t) + c\frac{(b^+ + \mu_-)^2}{\sigma_+^2}(T-t) + H^+$, respectively. Now let us consider the Cauchy-Dirichlet problem on the bounded domain $(0,T) \times (-N, N)$

$$\begin{aligned}
&\frac{\partial}{\partial t}u(t,y) + \frac{1}{2}u_{yy}(t,y) + \beta(y)u_y(t,y) + \gamma u_y^2(t,y) + d(y) \\
&\qquad + c \min_{(\mu,\sigma)\in K} \frac{(b(y) + \mu + g\sigma u_y(t,y))^2}{2\sigma_M\sigma - \sigma_-\sigma_+} = 0, \\
&u(T,y) = H(y), \quad u(t, \pm N) = u^{\pm}(t).
\end{aligned}$$

Suppose

$$a_1(t,y,u,p) = \frac{1}{2}p, \quad a(t,y,u,p) = -\beta(y)p - \gamma p^2 - d(y) - c \min_{(\mu,\sigma)\in K} \frac{(b(y) + \mu + g\sigma p)^2}{2\sigma_M\sigma - \sigma_-\sigma_+}.$$

Hence we get the Cauchy-Dirichlet problem for $\tilde{u}(t,y) = u(T-t, y)$ in the form of [24]

$$\begin{aligned}
&\frac{\partial}{\partial t}\tilde{u}(t,y) - \frac{\partial}{\partial y}a_1(t,y,\tilde{u}(t,y),\tilde{u}_y(t,y)) + a(t,y,\tilde{u}(t,y),\tilde{u}_y(t,y)) = 0, \\
&\qquad \tilde{u}(0,y) = H(y),\ \tilde{u}(t,\pm N) = u^{\pm}(T-t).
\end{aligned}$$

It is easy to see that $a(t,y,u,p)$ is the Lipschitz function on the each ball of its domain, $a(t,y,u,0)u$ is a lower-bounded by a quadratic function of the type $-b_1u^2 - b_2,\ b_1, b_2 > 0$ and all the rest conditions of Theorem 6.2 (Chapter V, p. 457) of [24] are satisfied. Therefore there exists a classical solution of (3.5), (3.6) with bounded $u_y(t,y)$. $\square$

Remark B.1. The existence of classical solution of equation (B.3) with boundary conditions $u(T,y) = 0,\ u_y(t,\pm N) + u(t,\pm N) = u^{\pm}(t)$ follows also from Example 3.6 of [18].

ACKNOWLEDGEMENT

This study was supported by Georgian National Science Foundation grant GNSF/ST09-471-3-104.

REFERENCES

[1] Aubin J.-P. and Ekeland I., Applied nonlinear analysis, John Wiley & Sons, Inc., New York (1984)

[2] Avelaneda M., Levy A. and Paras A., Pricing and hedging derivative securities in market with uncertain volatilities, Applied Mathematical Finance, 2, 73–88 (1995).

[3] Ahn H., Muni A. and Swindle G., Misspecified asset price models and robust hedging strategies, Applied Mathematical Finance, 4, 21–36 (1997).

[4] Ahn H., Muni A. and Swindle G., Optimal hedging strategies for misspecified asset price models, Applied Mathematical Finance, 6, 197–208 (1999).

[5] Bordigoni G., Matoussi A. and Schweizer M., A Stochastic Control Approach to a Robust Utility Maximization Problem, in: F. E. Benth et al. (eds.), Stochastic Analysis and Applications. Proceedings of the Second Abel Symposium, Oslo, Springer, 125–151 (2007)

[6] Buff R., Uncertain volatility models – theory and applications, Springer-Verlag, Berlin (2002)

[7] Cvitanic J. and Karatzas I., On dynamic measures of risk, Finance Stoch., 3(4), 451–482 (1999)

[8] Cheridito P., Soner H.M., Touzi N. and Victoir N., Second Order Backward Stochastic Differential Equations and Fully Non-Linear Parabolic PDEs, Comm. Pure Appl. Math., 60(7), 1081–1110 (2007)
[9] Chen Z. and Epstein L., Ambiguity, Risk, and Asset Returns in Continuous Time, Econometrica, Vol. 70, 4, pp. 1403-1443, (2002)
[10] Denis L. and Martini C., A theoretical framework for the pricing of contingent claims in the presence of model uncertainty, Ann. Appl. Probab., 16(2), 827–852 (2006)
[11] Denis L. and Kervarec M., Utility functions and optimal investment in non-dominated models, http://hal.archives-ouvertes.fr/hal-00371215/en/, (2009)
[12] Dokuchaev N., Maximin investment problems for discounted and total wealth, IMA J. Manag. Math., 19(1), 63–74 (2008)
[13] Föllmer H. and Gundel A., Robust projections in the class of martingale measures, Illinois J. Math., 50(1-4), 439–472 (2006)
[14] Friedman A., Stochastic differential equations and applications, vol. I-II, Academic Press, New York-London (1975)
[15] Gundel A., Robust utility maximization for complete and incomplete market models, Finance Stoch., 9, 151–176 (2005)
[16] Hernandez-Hernandez D. and Schied A., Robust utility maximization in a stochastic factor model, Statist. Decisions 24(1), 109–125 (2006)
[17] Hernández-Hernández D. and Schied A., A control approach to robust utility maximization with logarithmic utility and time-consistent penalties, Stochastic Process. Appl., 117(8), 980–1000 (2007)
[18] Henry D. , Geometric theory of semilinear parabolic equations, Springer-Verlag, Berlin-New York (1981)
[19] Huber P. J., Robust Statistics. N.-Y., John Wiley & Sons, Inc., New York (1981)
[20] Kramkov D. and Schachermayer W., The asymptotic elasticity of utility functions and optimal investment in incomplete markets. Ann. Appl. Probab., 9(3), 904–950 (1999)
[21] Kazamaki N., Continuous exponential martingales and BMO, Lecture Notes in Mathematics, 1579. Springer-Verlag, Berlin (1994)
[22] Kobylanski M., Backward stochastic differential equations and partial differential equations with quadratic growth, Ann. Appl. Probab., 28(2), 558–602 (2000)
[23] Korn R. and Menkens O., Worst-case scenario portfolio optimization: a new stochastic control approach, Math. Methods Oper. Res., 62(1), 123–140 (2005)
[24] Ladyzhenskaya O. A., Solonnikov V. A., and Ural'tseva N. N., Linear and quasi-linear equations of parabolic type, American Mathematical Society (1988).
[25] Lazrieva N. and Toronjadze T., Optimal robust mean-variance hedging in incomplete financial markets, J. Math. Sci. (N. Y.), 153(3), 262–290 (2008)
[26] Liptzer R. Sh. and Shiryayev A. N., Martingale theory, Nauka, Moscow (1986)
[27] Mania M. and Santacroce M., Exponential hedging under partial information, Finance Stoch., 14(3), 419–448 (2010)
[28] Matoussi A., Possamai D. and Zhou C., Robust Utility Maximization In Non-Dominated Models With 2BSDEs, Preprint (2012), http://arxiv.org/abs/1201.0769
[29] Nisio M., Stochastic differential games and viscosity solutions of Isaacs equations, Nagoya Math. J., 110, 163–184 (1988)
[30] Quenez M-C., Optimal portfolio in a multiple-priors model, Seminar on Stochastic Analysis, Random Fields and Applications IV, 291–321, Progr. Probab., 58, Birkhäuser, Dalang R.C., Dozzi M. Russo F. Editors , Basel (2004)
[31] Schied A., Robust optimal control for a consumption-investment problem, Math. Methods Oper. Res., 67(1), 1–20 (2008)
[32] Soner H. M., Touzi N. and Zhang J., Wellposedness of second order BSDEs, Probability Theory and Related Fields, Vol. 153, N. 1-2 (2012), 149-190, (2012).
[33] Talay D. and Zheng Z., Worst case model risk management, Finance Stoch., 6(4), 517–537 (2002)
[34] Tevzadze R., Solvability of backward stochastic differential equations with quadratic growth, Stochastic Process. Appl., 118(3), 503–515 (2008)
[35] Touzi N., Direct Characterization of the value of super-replication under stochastic volatility and portfolio constraints, Stochastic Process. Appl., 88(2), 305–328 (2000)
[36] Warga J., Optimal control of differential and functional equations, Academic press, New York-London (1976)

NEW PROOFS OF SOME RESULTS ON BMO MARTINGALES USING BSDES

B. CHIKVINIDZE AND M. MANIA

Abstract. Using properties of backward stochastic differential equations we give new proofs of some well known results on BMO martingales and improve some estimates of BMO norms.

Key words and phrases: BMO martingales, Girsanov's transformation, Backward stochastic differential equation

MSC 2010: 60G44.

1. INTRODUCTION

The BMO martingale theory is extensively used to study backward stochastic differential equations (BSDEs). Some properties of BMO martingales was already used by Bismut [3] when he discussed the existence and uniqueness of a solution of some particular backward stochastic Riccati equations, choosing the BMO space for the martingale part of the solution process. In the work of Delbaen et al [6] conditions for the closedness of stochastic integrals with respect to semimartingales in L^2 were established in relation to the problem of hedging contingent claims and linear BSDEs. Most of these conditions deal with BMO martingales and reverse Hölder inequalities. BMO martingales naturally arise in BSDEs with quadratic generators. When the generator of a BSDE has quadratic growth then the martingale part of any bounded solution of the BSDE is a BMO martingale. This fact was proved in [10, 13, 14, 15, 17, 20] under various degrees of generality. Note that in [4] the existence of a solution was proved to BSDE with quadratic growth and unbounded terminal condition, where the terminal value satisfies certain exponential moment condition. In this case the martingale part of a solution of such equation is not a BMO martingale in general, but the stochastic exponential of the martingale part (as for BMO martingales) is a uniformly integrable martingale (see [18] for details). Later, the BMO norms were used to prove an existence, uniqueness and stability results for BSDEs, among others in [1, 2, 5, 7, 9, 16, 19, 20].

The aim of this paper is to do the converse: to prove some results on BMO martingales using the BSDE technique.

It is well known that if M is a BMO martingale, then the mapping $\phi\ :\ \mathcal{L}(P)\ni X\longrightarrow \tilde{X}=\langle X,M\rangle - X\in\mathcal{L}(\tilde{P})$ is an isomorphism of $BMO(P)$ onto $BMO(\tilde{P})$, where $d\tilde{P}=\mathcal{E}_T(M)dP$. E. g., it was proved by Kazamaki [11, 12] that the inequality

$$\|\tilde{X}\|_{BMO(\tilde{P})}\leq C_{Kaz}(\tilde{M})\cdot\|X\|_{BMO(P)}$$

is valid for all $X\in BMO(P)$, where the constant $C_{Kaz}(\tilde{M})>0$ is independent of X but depends on the martingale M. Using the properties of a suitable BSDE we prove this

Published in *J. Theoret. Probab.* **27** (2014), no. 4, 1213–1228.

inequality with a constant $C(\tilde{M})$ which we express as a linear function of the $BMO(\tilde{P})$ norm of $\tilde{M} = \langle M \rangle - M$ and which is less than $C_{Kaz}(\tilde{M})$ for all values of this norm.

Using properties of BSDEs we also prove the well known equivalence between BMO property, Muckenhoupt and reverse Hölder conditions (Doleans-Dade and Meyer [8], Kazamaki [12]) and obtain BMO norm estimates in terms of reverse Hölder and Muckenhaupt constants.

2. Reverse Hölder and Muckenhoupt Conditions and Relations with BSDEs

We start with a probability space $(\Omega, \mathcal{F}, P)$, a finite time horizon $0 < T < \infty$ and a filtration $F = (\mathcal{F}_t)_{0 \leq t \leq T}$ satisfying the usual conditions of right-continuity and completeness.

We recall definitions of BMO martingales, Reverse Hölder and Muckenhaupt conditions (see, e.g., Doleans-Dade and Meyer [8], or Kazamaki [12]).

Definition 1. A continuous, uniformly integrable martingale $(M_t, \mathcal{F}_t)$ with $M_0 = 0$ is said to be from the class BMO if

$$\|M\|_{BMO} = \sup_{\tau} \Big\| E\big[\langle M \rangle_T - \langle M \rangle_\tau | \mathcal{F}_\tau\big]^{1/2} \Big\|_\infty < \infty,$$

where the supremum is taken over all stopping times $\tau \in [0, T]$ and $\langle M \rangle$ is the sharp bracket of M.

Denote by $\mathcal{E}(M)$ the stochastic exponential of a continuous local martingale M:

$$\mathcal{E}_t(M) = \exp\Big\{M_t - \frac{1}{2}\langle M \rangle_t\Big\}.$$

Throughout the paper we assume that M is a continuous local martingale with $\langle M \rangle_T < \infty$ P- a.s. This implies that $\mathcal{E}_t(M) > 0$ P-a.s. for all $t \in [0, T]$, which allows to define $\mathcal{E}_{\tau,T}(M)$ as $\mathcal{E}_{\tau,T}(M) = \mathcal{E}_T(M)/\mathcal{E}_\tau(M)$.

Definition 2. Let $1 < p < \infty$. $\mathcal{E}(M)$ is said to satisfy (R_p) condition if the reverse Hölder inequality

$$E\Big[\big\{\mathcal{E}_{\tau,T}(M)\big\}^p \Big| \mathcal{F}_\tau\Big] \leq C_p$$

is valid for every stopping time τ, with a constant $C_p > 0$ depending only on p.

If $\mathcal{E}(M)$ is a uniformly integrable martingale then by the Jensen inequality we also have that $E\Big[\big\{\mathcal{E}_{\tau,T}(M)\big\}^p \Big| \mathcal{F}_\tau\Big] \geq 1$.

A condition dual to (R_p) is the Muckenhoupt condition (A_p).

Definition 3. $\mathcal{E}(M)$ is said to satisfy (A_p) condition for $1 < p < \infty$ if there is a constant $D_p > 0$ such that for every stopping time $\tau \in [0, T]$

$$E\Big[\big\{\mathcal{E}_{\tau,T}(M)\big\}^{-\frac{1}{p-1}} \Big| \mathcal{F}_\tau\Big] \leq D_p.$$

Note that, since $\mathcal{E}(M)$ is a supermartingale, the Jensen inequality implies the converse inequality

$$E\Big[\big\{\mathcal{E}_{\tau,T}(M)\big\}^{-\frac{1}{p-1}} \Big| \mathcal{F}_\tau\Big] \geq \Big\{E\big[\mathcal{E}_{\tau,T}(M) \big| \mathcal{F}_\tau\big]\Big\}^{-\frac{1}{p-1}} \geq 1.$$

In this paper we shall consider only linear BSDEs of the type

$$Y_t = Y_0 - \int_0^t [\alpha Y_s + \beta \psi_s] d\langle M \rangle_s + \int_0^t \psi_s dM_s + N_t, \quad Y_T = 1,$$

where α and β are constants.

A solution of such a BSDE is a triple (Y, ψ, N), where Y is a special semimartingale, ψ is a predictable M-integrable process and N is a locally square integrable martingale with $\langle N, M \rangle = 0$.

Let us define the space $S^\infty \times BMO(P) \times H^2(P)$ equipped with the following norms

$$\|Y\|_\infty = \|Y_T^*\|_{L^\infty}, \quad \text{where} \quad Y_T^* = \sup_{t \in [0,T]} |Y_t|,$$

$$\|\psi \cdot M\|_{BMO(P)} = \sup_\tau \Big\| E\Big[\int_\tau^T \psi_s^2 d\langle M \rangle_s | \mathcal{F}_\tau \Big]^{1/2} \Big\|_\infty,$$

$$\|N\|_{H^2} = E^{\frac{1}{2}}[N]_T,$$

where $[N]$ is the square bracket of N.

Note that, since the martingale M is assumed to be continuous, only the latter term of this equation may have the jumps, i.e., $\Delta Y = \Delta N$. In order to avoid the definition of BMO norms for right-continuous martingales, we are using the H^2 norms for orthogonal martingale parts. This is sufficient for our goals, since the generators of equations under consideration do not depend on orthogonal martingale parts.

Sometimes we call Y alone the solution of BSDE, keeping in mind that $\psi \cdot M + N$ is the martingale part of Y.

Lemma 1. *Let M be a continuous local martingale.*
a) *$\mathcal{E}(M)$ satisfies (R_p) if and only if there exists a bounded, positive solution of BSDE*

$$\begin{cases} Y_t = Y_0 - \int_0^t [\frac{p(p-1)}{2} Y_s + p\psi_s] d\langle M \rangle_s + \int_0^t \psi_s dM_s + N_t, \\ Y_T = 1. \end{cases} \tag{1}$$

b) *$\mathcal{E}(M)$ satisfies (A_p) if and only if there exists a bounded, positive solution of equation*

$$\begin{cases} X_t = X_0 - \int_0^t [\frac{p}{2(p-1)^2} X_s - \frac{1}{p-1}\varphi_s] d\langle M \rangle_s + \int_0^t \varphi_s dM_s + L_t, \\ X_T = 1. \end{cases} \tag{2}$$

Proof. **a)** Let first show that if $\mathcal{E}(M)$ satisfies (R_p), then the process $Y_t = E\Big[\big\{\mathcal{E}_{t,T}(M)\big\}^p \Big| \mathcal{F}_t\Big]$ is a solution of BSDE (1). It is evident that Y is a bounded positive process and that $Y_t\big\{\mathcal{E}_t(M)\big\}^p$ is a uniformly integrable martingale. Therefore, since $\mathcal{E}_t(M) > 0$, the process Y will be a special semimartingale. Let $Y_t = Y_0 + A_t + m_t$ be the canonical decomposition of Y, where m is a locally square integrable martingale and A a predictable process of bounded variation. Using the Galtchouk-Kunita-Watanabe decomposition for m, we get

$$Y_t = Y_0 + A_t + \int_0^t \psi_s dM_s + N_t, \tag{3}$$

where N is a locally square integrable martingale strongly orthogonal to M.

Now using the Itô formula we have

$$Y_t\big\{\mathcal{E}_t(M)\big\}^p = Y_0 + \int_0^t \big[\frac{p(p-1)}{2}Y_s + p\psi_s\big]\big\{\mathcal{E}_s(M)\big\}^p d\langle M\rangle_s + \int_0^t \big\{\mathcal{E}_s(M)\big\}^p dA_s + \tilde{m}_t, \tag{4}$$

where $\tilde{m}$ is a local martingale.

Because $Y_t\big\{\mathcal{E}_t(M)\big\}^p$ is a martingale, equalizing the part of bounded variation to zero, we obtain that

$$A_t = -\int_0^t \Big[\frac{p(p-1)}{2}Y_s + p\psi_s\Big]d\langle M\rangle_s,$$

which implies that $Y_t = E\Big[\big\{\mathcal{E}_{t,T}(M)\big\}^p\Big|\mathcal{F}_t\Big]$ is a solution of equation (1).

Now let equation (1) admits a bounded positive solution Y_t. Using the Itô formula for the process $Y_t\big\{\mathcal{E}_t(M)\big\}^p$ we get that $Y_t\big\{\mathcal{E}_t(M)\big\}^p$ is a local martingale. Hence it is a supermartingale, as a positive local martingale. Therefore, from the supermartingale inequality and the boundary condition $Y_T = 1$ we obtain that $E\Big[\big\{\mathcal{E}_{t,T}(M)\big\}^p\Big|\mathcal{F}_t\Big] \le Y_t$. Because Y is bounded, this implies that $\mathcal{E}(M)$ satisfies (R_p) condition.

b) The proof is similar to the proof of the part a), we only need to replace p by $-\frac{1}{p-1}$. □

Let $\mathcal{E}(M)$ be a uniformly integrable martingale. Denote by $\tilde{P}$ a new probability measure defined by $d\tilde{P} = \mathcal{E}_T(M)dP$ and let $\tilde{M} = \langle M\rangle - M$.

Now we shall give a new proof of the well known equivalence (Doleans-Dade and Meyer [8], Kazamaki [12]) between BMO property, Muckenhoupt and reverse Hölder conditions.

Theorem 1. *Let $\mathcal{E}(M)$ be a uniformly integrable martingale. Then the following conditions are equivalent:*
i) $\tilde{M} \in BMO(\tilde{P})$.
ii) $\mathcal{E}(M)$ *satisfies the* (R_p) *condition for some* $p > 1$.
iii) $M \in BMO(P)$.
iv) $\mathcal{E}(M)$ *satisfies the* (A_p) *condition for some* $p > 1$.

Proof. For the sake of simplicity, in all proofs given here, we shall assume without loss of generality that all stochastic integrals are martingales, otherwise one can use the localization arguments.

i) $\Longrightarrow$ ii) Let $\tilde{M} \in BMO(\tilde{P})$. According to Lemma 1 it is sufficient to show that equation (1) admits a bounded positive solution for some $p > 1$. Let us rewrite equation (1) in terms of the $\tilde{P}$-martingale $\tilde{M}$:

$$\begin{cases} Y_t = Y_0 - \int_0^t [\frac{p(p-1)}{2}Y_s + (p-1)\psi_s]d\langle M\rangle_s - \int_0^t \psi_s d\tilde{M}_s + N_t, \\ Y_T = 1. \end{cases}$$

Since $\langle N, M\rangle = 0$, N is a local $\tilde{P}$- martingale orthogonal to $\tilde{M}$.

Define the mapping $H\ :\ S^\infty \times BMO(\tilde{P}) \times H^2(\tilde{P})$ into itself, which maps $(y, \psi, n) \in S^\infty \times BMO(\tilde{P}) \times H^2(\tilde{P})$ onto the solution (Y, Ψ, N) of the BSDE (1), i.e.,

$$Y_t = E^{\tilde{P}}\left[1 + \int_t^T \Big[\frac{p(p-1)}{2} y_s + (p-1)\psi_s\Big] d\langle M\rangle_s \Big| \mathcal{F}_t\right]$$

and

$$-\int_0^t \Psi_s d\tilde{M}_s + N_t = E^{\tilde{P}}\left[\int_0^T \Big[\frac{p(p-1)}{2} y_s + (p-1)\psi_s\Big] d\langle M\rangle_s \Big| \mathcal{F}_t\right]$$
$$-E^{\tilde{P}} \int_0^T \Big[\frac{p(p-1)}{2} y_s + (p-1)\psi_s\Big] d\langle M\rangle_s.$$

We shall show that there exists $p > 1$ such that this mapping is a contraction.

Let

$$\delta Y = Y^1 - Y^2,\ \delta y = y^1 - y^2,\ \delta\Psi = \Psi^1 - \Psi^2,\ \delta\psi = \psi^1 - \psi^2,\ \delta N = N^1 - N^2.$$

It is evident that $\delta Y_T = 0$ and

$$\delta Y_t = \delta Y_0 - \int_0^t \Big[\frac{p(p-1)}{2}\delta y_s + (p-1)\delta\psi_s\Big] d\langle M\rangle_s - \int_0^t \delta\Psi_s d\tilde{M}_s + \delta N_t.$$

Applying the Itô formula to $(\delta Y_\tau)^2 - (\delta Y_T)^2$ and taking conditional expectations we have

$$(\delta Y_\tau)^2 + E^{\tilde{P}}\left[\int_\tau^T (\delta\Psi_s)^2 d\langle M\rangle_s \Big| \mathcal{F}_\tau\right] + E^{\tilde{P}}\Big[[\delta N]_T - [\delta N]_\tau \Big| \mathcal{F}_\tau\Big]$$
$$= E^{\tilde{P}}\left[\int_\tau^T p(p-1)\delta Y_s \delta y_s d\langle M\rangle_s \Big| \mathcal{F}_\tau\right] + E^{\tilde{P}}\left[\int_\tau^T 2(p-1)\delta Y_s \delta\psi_s d\langle M\rangle_s \Big| \mathcal{F}_\tau\right]$$

and using elementary inequalities we obtain

$$(\delta Y_\tau)^2 + E^{\tilde{P}}\left[\int_\tau^T (\delta\Psi_s)^2 d\langle M\rangle_s \Big| \mathcal{F}_\tau\right] + E^{\tilde{P}}\Big[[\delta N]_T - [\delta N]_\tau \Big| \mathcal{F}_\tau\Big]$$
$$\leq \frac{p(p-1)}{2}\|\tilde{M}\|^2_{BMO(\tilde{P})} \cdot \|\delta Y\|^2_\infty + \frac{p(p-1)}{2}\|\tilde{M}\|^2_{BMO(\tilde{P})} \cdot \|\delta y\|^2_\infty$$
$$+ (p-1)\|\tilde{M}\|^2_{BMO(\tilde{P})} \cdot \|\delta Y\|^2_\infty + (p-1)\Big\|\int \delta\psi d\tilde{M}\Big\|^2_{BMO(\tilde{P})}.$$

Because the right-hand side of the inequality does not depend on τ, we will have

$$\Big(1 - \frac{3p(p-1)}{2}\|\tilde{M}\|^2_{BMO(\tilde{P})} - 3(p-1)\|\tilde{M}\|^2_{BMO(\tilde{P})}\Big)\|\delta Y\|^2_\infty$$
$$+ \Big\|\int \delta\Psi d\tilde{M}\Big\|^2_{BMO(\tilde{P})} + \|\delta N\|^2_{H^2(\tilde{P})}$$
$$\leq \frac{3p(p-1)}{2}\|\tilde{M}\|^2_{BMO(\tilde{P})}\|\delta y\|^2_\infty + 3(p-1)\Big\|\int \delta\psi d\tilde{M}\Big\|^2_{BMO(\tilde{P})}. \tag{5}$$

Since

$$1 - \frac{3}{2}(p-1)(p+2)\|\tilde{M}\|^2_{BMO(\tilde{P})} > 0$$

for p sufficiently close to 1, one can make the constant of $\|\delta Y\|_\infty^2$ in the left-hand side of (5) positive and we finally obtain the inequality

$$\|\delta Y\|_\infty^2 + \Big\| \int \delta\Psi d\tilde{M} \Big\|_{BMO(\tilde{P})}^2 + \|\delta N\|_{H^2(\tilde{P})}^2$$

$$\leq \alpha(p)\cdot \|\delta y\|_\infty^2 + \beta(p)\cdot \Big\| \int \delta\psi d\tilde{M} \Big\|_{BMO(\tilde{P})}^2, \tag{6}$$

where

$$\alpha(p) = \frac{3p(p-1)\|\tilde{M}\|_{BMO(\tilde{P})}^2}{2-3(p-1)(p+2)\|\tilde{M}\|_{BMO(\tilde{P})}^2},$$

$$\beta(p) = \frac{6(p-1)}{2-3(p-1)(p+2)\|\tilde{M}\|_{BMO(\tilde{P})}^2}.$$

It is easy to see that $\lim_{p\downarrow 1}\alpha(p) = \lim_{p\downarrow 1}\beta(p) = 0$. So, if we take p^* such that $\alpha(p^*) < 1$ and $\beta(p^*) < 1$ we obtain that there exists $0 < C < 1$ such that

$$\|\delta Y\|_\infty^2 + \Big\| \int \delta\Psi d\tilde{M} \Big\|_{BMO(\tilde{P})}^2 + \|\delta N\|_{H^2(\tilde{P})}^2$$

$$\leq C\big(\|\delta y\|_\infty^2 + \Big\| \int \delta\psi d\tilde{M} \Big\|_{BMO(\tilde{P})}^2 + \|\delta n\|_{H^2(\tilde{P})}^2\big), \tag{7}$$

for any $(y,\psi,n)\in S^\infty \times BMO(\tilde{P})\times H^2(\tilde{P})$.

Thus, the mapping H is a contraction and there exists a fixed-point of H, which is the unique solution (Y,Ψ,N) of (1) in $S^\infty \times BMO(\tilde{P})\times H^2(\tilde{P})$.

Since $\alpha(p)$ and $\beta(p)$ are decreasing functions of $p\in(1,\infty)$, the norms $\|Y\|_\infty$ and $\|\Psi\cdot\tilde{M}\|_{BMO(\tilde{P})}$ are uniformly bounded, as functions of p for $p\in[1,p^*]$. Therefore, for any $p\in[1,p^*]$ we have

$$Y_t = E^{\tilde{P}}\Big[1 + \int_t^T \Big[\frac{p(p-1)}{2}Y_s + (p-1)\Psi_s\Big] d\langle M\rangle_s \Big| \mathcal{F}_t\Big] \tag{8}$$

and

$$Y_t \geq 1 - \frac{p(p-1)}{2}\|Y\|_\infty \|\tilde{M}\|_{BMO(\tilde{P})}^2 - \frac{p-1}{2}\|\tilde{M}\|_{BMO(\tilde{P})}^2$$

$$-\frac{p-1}{2}\|\Psi\cdot\tilde{M}\|_{BMO(\tilde{P})}^2 \geq 0$$

for some p sufficiently close to 1. Hence, there exists a bounded positive solution of equation (1) for some $p > 1$, which implies that $\mathcal{E}(M)$ satisfies the R_p condition, according to Lemma 1.

ii) $\Longrightarrow$ iii) Let $\mathcal{E}(M)$ be a uniformly integrable martingale and satisfies the (R_p) condition for some $p > 1$. Then the process $Y_t = E\Big[\big\{\mathcal{E}_{t,T}(M)\big\}^p\Big|\mathcal{F}_t\Big]$ is a solution of equation (1) and satisfies the two-sided inequality

$$1 \leq Y_t \leq C_p. \tag{9}$$

Applying the Itô formula to $e^{-\beta Y_t}$, integrating from τ to T and taking conditional expectations we have

$$e^{-\beta} - e^{-\beta Y_\tau} = \beta\frac{p(p-1)}{2}E\Big[\int_\tau^T Y_s e^{-\beta Y_s} d\langle M\rangle_s \Big| \mathcal{F}_\tau\Big]$$

$$+E\Big[\int_\tau^T e^{-\beta Y_s}\Big(\frac{\beta^2}{2}\psi_s^2 + \beta p\psi_s\Big)d\langle M\rangle_s\Big|\mathcal{F}_\tau\Big] + \frac{\beta^2}{2}E\Big[\int_\tau^T e^{-\beta Y_s} d\langle N^c\rangle_s\Big|\mathcal{F}_\tau\Big]$$

$$+E\Big[\Sigma_{\tau<s\le T}\big(e^{-\beta Y_s} - e^{-\beta Y_{s-}} + \beta e^{-\beta Y_{s-}}\Delta Y_s\big)\Big|\mathcal{F}_\tau\Big].$$

Since $\frac{\beta^2}{2}\psi_s^2 + \beta p\psi_s \ge -\frac{p^2}{2}$ and $e^{-\beta Y_s} - e^{-\beta Y_{s-}} + \beta e^{-\beta Y_{s-}}\Delta Y_s \ge 0$ we obtain the inequality

$$\frac{p}{2}E\big[\int_\tau^T \big(\beta(p-1)Y_s - p\big)e^{-\beta Y_s} d\langle M\rangle_s|\mathcal{F}_\tau\big] \le e^{-\beta} - e^{-\beta Y_\tau}.$$

Then from the two-sided inequality (9) it follows that for any $\beta > \frac{p}{p-1}$

$$\frac{p}{2}(\beta(p-1)-p)e^{-\beta C_p}E\Big[\langle M\rangle_T - \langle M\rangle_\tau\Big|\mathcal{F}_\tau\Big] \le e^{-\beta} - e^{-\beta C_p}, \tag{10}$$

which implies that

$$\|M\|^2_{BMO(P)} \le \frac{2(e^{\beta(C_p-1)}-1)}{p(\beta(p-1)-p)},$$

since the right-hand side of (10) does not depends on τ.

iii) $\Longrightarrow$ iv) If M is a $BMO(P)$ martingale, then according to Lemma 1 it is sufficient to show that equation (2) admits bounded positive solution for some $p > 1$, which can be proved similarly to the implication i) $\Longrightarrow$ ii). By the same way one can show that for the mapping H

$$X_t = E\Big[1 + \int_t^T \Big[\frac{p}{2(p-1)^2}x_s - \frac{1}{p-1}\varphi_s\Big]d\langle M\rangle_s\Big|\mathcal{F}_t\Big],$$

where $-\int_0^t \Phi_s dM_s + L_t$ is the martingale part of X, the inequality (6) holds with

$$\alpha(p) = \frac{3p\|M\|^2_{BMO(P)}}{2(p-1)^2 - (9p-6)\|M\|^2_{BMO(P)}},$$

$$\beta(p) = \frac{6(p-1)}{2(p-1)^2 - (9p-6)\|M\|^2_{BMO(P)}},$$

where $\lim_{p\to\infty}\alpha(p) = \lim_{p\to\infty}\beta(p) = 0$. So if we take p large enough we obtain that the mapping H is a contraction.

iv) $\Longrightarrow$ i) The proof is similar to the proof of the implication ii) $\Longrightarrow$ iii) and we only give a brief sketch of the proof.

Since $\mathcal{E}(M)$ satisfies the (A_p) condition for some $p > 1$, according to Lemma 1 the process $X_t = E\Big[\big\{\mathcal{E}_{t,T}(M)\big\}^{-\frac{1}{p-1}}\Big|\mathcal{F}_t\Big]$ is a bounded positive solution of equation (2), which can be written in the following equivalent form

$$X_t = X_0 - \int_0^t [\frac{p}{2(p-1)^2}X_s - \frac{p}{p-1}\varphi_s]d\langle M\rangle_s - \int_0^t \varphi_s d\tilde{M}_s + L_t$$

in terms of $\tilde{P}$ martingale $\tilde{M} = \langle M \rangle - M$. Note that $\langle \tilde{M} \rangle = \langle M \rangle$ and L is also a local $\tilde{P}$ martingale orthogonal to $\tilde{M}$.

Applying the Itô formula for $e^{-\beta X_T} - e^{-\beta X_\tau}$, using successively the elementary inequality $\frac{\beta^2}{2}\varphi_s^2 - \frac{\beta p}{p-1}\varphi_s \geq -\frac{p^2}{2(p-1)^2}$, the convexity of the function $e^{-\beta x}$ and the two-sided inequality $1 \leq X_t \leq D_p$, similarly to the implication ii) $\Longrightarrow$ iii) we obtain the following estimate for the BMO norm of $\tilde{M}$

$$\|\tilde{M}\|^2_{BMO(\tilde{P})} \leq \frac{2(p-1)^2}{p(\beta - p)}\left(e^{\beta(D_p - 1)} - 1\right)$$

valid for any $\beta > p$, where D_p is a constant from Definition 3. □

3. Girsanov's Transformation of BMO Martingales and BSDEs

Let M be a continuous local P-martingale such that $\mathcal{E}(M)$ is a uniformly integrable martingale and let $d\tilde{P} = \mathcal{E}_T(M)dP$. To each continuous local martingale X we associate the process $\tilde{X} = \langle X, M \rangle - X$, which is a local $\tilde{P}$-martingale according to Girsanov's theorem. We denote this map by $\varphi : \mathcal{L}(P) \to \mathcal{L}(\tilde{P})$, where $\mathcal{L}(P)$ and $\mathcal{L}(\tilde{P})$ are classes of P and $\tilde{P}$ local martingales.

Let consider the process

$$Y_t = E^{\tilde{P}}\left[\langle X \rangle_T - \langle X \rangle_t \middle| \mathcal{F}_t\right] = E\left[\mathcal{E}_{t,T}(M)(\langle X \rangle_T - \langle X \rangle_t)\middle|\mathcal{F}_t\right]. \tag{11}$$

Since $\langle \tilde{X} \rangle = \langle X \rangle$ under either probability measure, it is evident that

$$\|Y\|_\infty = \|\tilde{X}\|^2_{BMO(\tilde{P})}. \tag{12}$$

Let $M \in BMO(P)$. According to Theorem 1 condition (R_p) is satisfied for some $p > 1$. The (R_p) condition and conditional energy inequality (Kazamaki [12], page 29) imply that for any $X \in BMO(P)$ the process Y is bounded, i.e., φ maps $BMO(P)$ into $BMO(\tilde{P})$. Moreover, as proved by Kazamaki [11, 12], $BMO(P)$ and $BMO(\tilde{P})$ are isomorphic under the mapping ϕ and for all $X \in BMO(P)$ the inequality

$$\|\tilde{X}\|^2_{BMO(\tilde{P})} \leq C^2_{Kaz}(\tilde{M}) \cdot \|X\|^2_{BMO(P)} \tag{13}$$

is valid, where

$$C^2_{Kaz}(\tilde{M}) = 2p \cdot 2^{1/p} \sup_\tau \left\| E^{\tilde{P}}\left[\left\{\mathcal{E}_{\tau,T}(\tilde{M})\right\}^{-\frac{1}{p-1}} \middle| \mathcal{F}_\tau\right]\right\|_\infty^{(p-1)/p}, \tag{14}$$

and $p > 1$ is such that

$$\|\tilde{M}\|_{BMO(\tilde{P})} < \sqrt{2}(\sqrt{p} - 1). \tag{15}$$

The conditional expectation in (14) is bounded, if p satisfies inequality (15), according to Theorem 2.4 from [12]. Note also that the similar inequality holds for the inverse mapping ϕ^{-1} by the closed graph theorem.

Similarly to Lemma 1 one can show that for any $X \in BMO(P)$ the process Y(defined by (11)) is a positive bounded solution of the $BSDE$

$$Y_t = Y_0 - \langle X \rangle_t - \int_0^t \varphi_s d\langle M \rangle_s + \int_0^t \varphi_s dM_s + L_t, \quad Y_T = 0. \tag{16}$$

Indeed, it is evident that $(Y_t + \langle X\rangle_t)\mathcal{E}_t(M)$ is a local martingale. Since $\mathcal{E}_t(M) > 0$ P-a.s. for all $t \in [0,T]$, the process Y will be a special semimartingale with the decomposition

$$Y_t = Y_0 + A_t + \int_0^t \varphi_s dM_s + N_t, \tag{17}$$

where A is a predictable process of bounded variation and N is a local martingale orthogonal to M.

By the Itô formula

$$(Y_t + \langle X\rangle_t)\mathcal{E}_t(M) = \int_0^t \mathcal{E}_s(M)\big[dA_s + d\langle X\rangle_s + \varphi_s d\langle M\rangle_s\big] + local \ \ martingale,$$

which implies that $A_t = -\langle X\rangle_t - \int_0^t \varphi_s d\langle M\rangle_s$. Therefore, it follows from (17) that Y satisfies equation (16).

Now we give an alternative proof of the inequality (13) with a constant expressed as a linear function of the BMO norm of the martingale $\tilde{M}$.

Theorem 2. *If $M \in BMO(P)$, then $\phi \ : \ X \to \tilde{X}$ is an isomorphism of $BMO(P)$ onto $BMO(\tilde{P})$. In particular, the inequality*

$$\frac{1}{\Big(1 + \frac{\sqrt{2}}{2}\|M\|_{BMO(P)}\Big)}\|X\|_{BMO(P)} \le \|\tilde{X}\|_{BMO(\tilde{P})}$$

$$\le \Big(1 + \frac{\sqrt{2}}{2}\|\tilde{M}\|_{BMO(\tilde{P})}\Big)\|X\|_{BMO(P)}. \tag{18}$$

is valid for any $X \in BMO(P)$.

Proof. Applying the Itô formula to $(Y_\tau + \varepsilon)^p - (Y_T + \varepsilon)^p$ (for $0 < p < 1, \ \varepsilon > 0$) and taking conditional expectations we obtain

$$\big(Y_\tau + \varepsilon\big)^p - \varepsilon^p = E\Big[\int_\tau^T p(Y_s+\varepsilon)^{p-1} d\langle X\rangle_s \Big| \mathcal{F}_\tau\Big] + \frac{p(1-p)}{2} E\Big[\int_\tau^T (Y_s+\varepsilon)^{p-2} d\langle L^c\rangle_s \Big| \mathcal{F}_\tau\Big]$$

$$+E\Big[\int_\tau^T \Big(\frac{p(1-p)}{2}(Y_s+\varepsilon)^{p-2}\varphi_s^2 + p(Y_s+\varepsilon)^{p-1}\varphi_s\Big) d\langle M\rangle_s \Big| \mathcal{F}_\tau\Big]$$

$$-E\Big[\Sigma_{\tau < s \le T}\big((Y_s+\varepsilon)^p - (Y_{s-}+\varepsilon)^p - p(Y_{s-}+\varepsilon)^{p-1}\Delta Y_s\big)\Big|\mathcal{F}_\tau\Big]. \tag{19}$$

Because $f(x) = x^p$ is concave for $p \in (0,1)$, the last term in (19) is positive. Therefore, using the inequality

$$\frac{p(1-p)}{2}(Y_s+\varepsilon)^{p-2}\varphi_s^2 + p(Y_s+\varepsilon)^{p-1}\varphi_s + \frac{p}{2(1-p)}(Y_s+\varepsilon)^p \ge 0$$

from (19) we obtain

$$(Y_\tau + \varepsilon)^p - \varepsilon^p \ge E\Big[\int_\tau^T p(Y_s+\varepsilon)^{p-1} d\langle X\rangle_s \Big| \mathcal{F}_\tau\Big]$$

$$-\frac{p}{2(1-p)} E\Big[\int_\tau^T (Y_s+\varepsilon)^p d\langle M\rangle_s \Big| \mathcal{F}_\tau\Big]. \tag{20}$$

Since $0 < p < 1$

$$p\big(\|Y\|_\infty + \varepsilon\big)^{p-1} E\Big[\langle X\rangle_T - \langle X\rangle_\tau \Big| \mathcal{F}_\tau\Big] \le E\Big[\int_\tau^T p(Y_s+\varepsilon)^{p-1} d\langle X\rangle_s \Big| \mathcal{F}_\tau\Big],$$

from (20) we have

$$p\big(\|Y\|_\infty + \varepsilon\big)^{p-1} E\Big[\langle X\rangle_T - \langle X\rangle_\tau \Big| \mathcal{F}_\tau\Big]$$

$$\le \big(Y_\tau + \varepsilon\big)^p - \varepsilon^p + \frac{p}{2(1-p)} E\Big[\int_\tau^T (Y_s+\varepsilon)^p d\langle M\rangle_s \Big| \mathcal{F}_\tau\Big]$$

and taking norms in the both sides of the latter inequality we obtain

$$p\big(\|Y\|_\infty+\varepsilon\big)^{p-1}\cdot\|X\|^2_{BMO(P)} \le \big(\|Y\|_\infty+\varepsilon\big)^p - \varepsilon^p + \frac{p}{2(1-p)}\big(\|Y\|_\infty+\varepsilon\big)^p\cdot\|M\|^2_{BMO(P)}.$$

Taking the limit when $\varepsilon \to 0$ we will have that for all $p \in (0,1)$

$$\|X\|^2_{BMO(P)} \le \Big(\frac{1}{p} + \frac{1}{2(1-p)}\|M\|^2_{BMO(P)}\Big)\cdot\|Y\|_\infty.$$

Therefore,

$$\|X\|^2_{BMO(P)} \le \min_{p\in(0,1)} \Big(\frac{1}{p} + \frac{1}{2(1-p)}\|M\|^2_{BMO(P)}\Big)\cdot\|Y\|_\infty$$

$$= \Big(1 + \frac{\sqrt{2}}{2}\|M\|_{BMO(P)}\Big)^2\cdot\|Y\|_\infty, \tag{21}$$

since the minimum of the function $f(p) = \frac{1}{p} + \frac{1}{2(1-p)}\|M\|^2_{BMO(P)}$ is attained for $p^* = \sqrt{2}/(\sqrt{2} + \|M\|_{BMO(P)})$ and $f(p^*) = \Big(1 + \frac{\sqrt{2}}{2}\|M\|_{BMO(P)}\Big)^2$.

Thus, from (21) and (12) we obtain

$$\frac{1}{\Big(1 + \frac{\sqrt{2}}{2}\|M\|_{BMO(P)}\Big)}\|X\|_{BMO(P)} \le \|\tilde{X}\|_{BMO(\tilde{P})}.$$

Now we can use inequality (21) for the Girsanov transform of $\tilde{X}$.

Since $dP/d\tilde{P} = \mathcal{E}_T^{-1}(M) = \mathcal{E}_T(\tilde{M}),\ \ \tilde{M}, \tilde{X} \in BMO(\tilde{P})$ and

$$\varphi(\tilde{X}) = \langle \tilde{X}, \tilde{M}\rangle - \tilde{X} = X,$$

from (21) we get the inverse inequality:

$$\|\tilde{X}\|_{BMO(\tilde{P})} \le \Big(1 + \frac{\sqrt{2}}{2}\|\tilde{M}\|_{BMO(\tilde{P})}\Big)\|X\|_{BMO(P)}. \tag{22}$$

The theorem is proved. □

Comparison of constants $C(\tilde{M})$ and $C_{Kaz}(\tilde{M})$. Let us compare the constant

$$C(\tilde{M}) = 1 + \frac{\sqrt{2}}{2}\|\tilde{M}\|_{BMO(\tilde{P})}$$

from (18) with the corresponding constant $C_{Kaz}(\tilde{M})$ from (13) (Kazamaki [12]).

Since by the Jensen inequality

$$E^{\tilde{P}}\Big[\big\{\mathcal{E}_{\tau,T}(\tilde{M})\big\}^{-\frac{1}{p-1}}\Big|\mathcal{F}_\tau\Big] \geq 1,$$

it follows from (14) that the constant $C_{Kaz}(\tilde{M})$ is more than $\sqrt{2p}$, where p is such that $\|\tilde{M}\|_{BMO(\tilde{P})} < \sqrt{2}(\sqrt{p} - 1)$. Since the last inequality is equivalent to the inequality

$$p > \big(1 + \frac{\sqrt{2}}{2}\|\tilde{M}\|_{BMO(\tilde{P})}\big)^2,$$

we obtain from (14) that at least

$$C^2(\tilde{M}) < \frac{1}{2}C^2_{Kaz}(\tilde{M}).$$

It is evident that in the trivial case $M = 0$ we have that $\tilde{P} = P$ and $\tilde{X} = X$. Note that, if $M = 0$ then (18) gives the two-sided inequality

$$\|X\|_{BMO(P)} \leq \|\tilde{X}\|_{BMO(\tilde{P})} \leq \|X\|_{BMO(P)},$$

implying the equality $\tilde{X} = X$, whereas from (13) we only have

$$\frac{1}{2}\|X\|_{BMO(P)} \leq \|\tilde{X}\|_{BMO(\tilde{P})} \leq 2\|X\|_{BMO(P)}.$$

This shows that the following simple corollary can not be deduced from inequality (13).

Corollary. *Let* $(M^n, n \geq 1)$ *be a sequence of* $BMO(P)$ *martingales such that* $\lim_{n\to\infty}\|M^n\|_{BMO(P)} = 0$. *Let* $dP^n = \mathcal{E}_T(M^n)dP$ *and* $\tilde{X}^n = \langle X, M^n\rangle - X$. *Then for any* $X \in BMO(P)$

$$\lim_{n\to\infty}\|\tilde{X}^n\|_{BMO(P^n)} = \|X\|_{BMO(P)}.$$

Proof. The second inequality of (18), applied for $X = M^n$ and $M = M^n$ gives

$$\|\tilde{M}^n\|_{BMO(P^n)} \leq \Big(1 + \frac{\sqrt{2}}{2}\|\tilde{M}^n\|_{BMO(P^n)}\Big)\|M^n\|_{BMO(P)}.$$

Therefore,

$$\frac{1}{\frac{\sqrt{2}}{2} + 1/\|\tilde{M}^n\|_{BMO(P^n)}} \leq \|M^n\|_{BMO(P)},$$

which implies that $\lim_{n\to\infty}\|\tilde{M}^n\|_{BMO(P^n)} = 0$. Now, passing to the limit in the two-sided inequality (18) we obtai

$$\|X\|_{BMO(P)} \leq \lim_{n\to\infty}\|\tilde{X}^n\|_{BMO(P^n)} \leq \|X\|_{BMO(P)}. \qquad \square$$

Remark. Note that the converse of Theorem 2 is also true. I.e., if M is a continuous local martingale and $\mathcal{E}(M)$ is a uniformly integrable martingale, Schachermayer [21] proved that if $M \notin BMO(P)$ then the map φ is not an isomorphism from $BMO(P)$ into $BMO(\tilde{P})$.

ACKNOWLEDGMENTS

The work was supported by the Rusthaveli National Scientific Foundation Grant N.FR/69/5-104/12. We would like to thank the referees for useful remarks and comments.

REFERENCES

[1] S. Ankirchner, P. Imkeller, and G. Reis, " Classical and variational differentiability of bsdes with quadratic growth" *Electronic Journal of Probability*, Vol. 12 , pp. 1418-1453, 2007.

[2] P. Barrieu, N. Cazanave and N. El Karoui, "Closedness results for BMO semi-martingales and application to quadratic BSDEs," *Comptes Rendus Mathematique*, Vol. 346, pp. 881-886, 2008.

[3] J. M. Bismut, "Controle des systemes lineaires quadratiques: applications de l'integrale stochastique," *Seminaire de Probabilites XII (eds.: C. Dellacherie, P. A. Meyer, and M. Weil), Lecture Notes in Mathematics 649*, Springer-Verlag, Berlin/Heidelberg, pp. 180-264, 1978.

[4] Ph. Briand and Y. Hu, " BSDE with quadratic growth and unbounded terminal value", Probability Theory and Related Fields 136, N. 4, pp. 604-618, 2006.

[5] B. Chikvinidze, "Backward stochastic differential equations with a convex generator," *Georgian Mathematical Journal*, Vol. 19, pp. 63-92, 2012.

[6] F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer and C. Stricker, "Weighted norm inequalities and hedging in incomplete markets," *Finance and Stochastics* , Vol. 1, pp. 181-227, 1997.

[7] F. Delbaen and S. Tang, "Harmonic analysis of stochastic equations and backward stochastic differential equations," *Probability Theory and Related Fields* , Vol. 146, pp. 291-336, 2010.

[8] C. Doleans-Dade and P. A. Meyer, "Inegalites de normes avec poids," *Universite de Strasbourg Seminaire de Probabilites* , XIII, pp. 313-331, 1979.

[9] C. Frei, M Mocha and N. Westray, "BSDEs in Utility Maximization with BMO Market Price of Risk", *Working Paper*, arXiv:1107.0183v1, 2011.

[10] Y. Hu, P. Imkeller, and M. Müller, "Utility maximization in incomplete markets," *Annals of Applied Probability* , Vol. 15, pp. 1691-1712, 2005.

[11] N. Kazamaki, "On transforming the class of BMO-martingales by a change of law," *Tohoku Mathematical Journal* , Vol. 31, pp. 117-125, 1979.

[12] N. Kazamaki, *Continuous Exponential Martingales and BMO* , Vol. 1579 of *Lecture Notes in Mathematics* , Springer, Berlin-Heidelberg, 1994.

[13] M. Kohlmann and S. Tang, "Minimization of risk and linear quadratic optimal control theory," *SIAM Journal on Control and Optimization* , Vol. 42, pp. 1118-1142, 2003.

[14] M. Mania and R. Tevzadze, "A semimartingale Bellman equation and the variance-optimal martingale measure," *Georgian Mathematical Journal* , Vol. 7, pp. 765-792, 2000.

[15] M. Mania and M. Schweizer, "Dynamic exponential indifference valuation," *Annals of Applied Probability* , Vol. 15, pp. 2113-2143, 2005.

[16] M. Mania and R. Tevzadze, "Martingale equation of exponential type," *Electronic communication in probability* , Vol. 11, pp. 206-216, 2006.

[17] M. Mania, M. Santacroce and R. Tevzadze, "A semimartingale BSDE related to the minimal entropy martingale measure",*Finance and Stochastics*, Vol. 7, No. 3, pp. 385-402, 2003.

[18] M. Mocha and N. Westray, "Quadratic Semimartingale BSDEs Under an Exponential Moments Condition", Seminaire de Probabilites XLIV, Lecture Notes in Mathematics, pp 105-139, 2012.

[19] M.A. Morlais, "Quadratic BSDEs driven by a continuous martingale and application to utility maximization problem", *Finance and Stochastics*, Vol. 13, No. 1, pp. 121-150, 2009.

[20] R. Tevzadze, "Solvability of backward stochastic differential equations with quadratic growth," *Stochastic Processes and their Applications* , Vol. 118, pp. 503-515, 2008.

[21] W. Schachermayer, "A characterization of the closure of H_∞ in BMO," *Seminaire de Probabilites XXX, Lecture Notes in Mathematics 1626*, Springer, Berlin, pp. 344-356, 1996.

RECURSIVE ESTIMATION PROCEDURES FOR ONE-DIMENSIONAL PARAMETER OF STATISTICAL MODELS ASSOCIATED WITH SEMIMARTINGALES

N. LAZRIEVA AND T. TORONJADZE

Abstract. The recursive estimation problem of a one-dimensional parameter for statistical models associated with semimartingales is considered. The asymptotic properties of recursive estimators are derived, based on the results on the asymptotic behaviour of a Robbins–Monro type SDE. Various special cases are considered.

Key words and phrases: Stochastic approximation, Robbins–Monro type SDE, semimartingale statistical models, recursive estimation, asymptotic properties

MSC 2010: 62L20, 60H10, 60H30

INTRODUCTION

Beginning from the paper [1] of A. Albert and L. Gardner a link between Robbins–Monro (RM) stochastic approximation algorithm (introduced in [19]) and recursive parameter estimation procedures was intensively exploited. Later on recursive parameter estimation procedures for various special models (e.g., i.i.d models, non i.i.d. models in discrete time, etc.) have been studied by a number of authors using methods of stochastic approximation (see, e.g., [2, 3, 4, 7, 8, 14, 15, 20, 21, 22]). It would be mentioned the fundamental book [18] by M. B. Nevelson and R. Z. Khas'minski (1972) between them.

In 1987 by N. Lazrieva and T. Toronjadze a heuristic algorithm of a construction of the recursive parameter estimation procedures for statistical models associated with semimartingales (including both discrete and continuous time semimartingale statistical models) was proposed [9]. These procedures could not be covered by the generalized stochastic approximation algorithm with martingale noises (see, e.g., [17]), while in discrete time case the classical RM algorithm contains recursive estimation procedures.

To recover the link between the stochastic approximation and recursive parameter estimation in [10, 11, 12] by Lazrieva, Sharia and Toronjadze the semimartingale stochastic differential equation was introduced, which naturally includes both generalized RM stochastic approximation algorithms with martingale noises and recursive parameter estimation procedures for semimartingale statistical models.

In the present work we are concerning with the construction of recursive estimation procedures for semimartingale statistical models asymptotically equivalent to the MLE and M-estimators, embedding these procedures in the Robbins–Monro type equation. For this reason in Section 1 we shortly describe the Robbins–Monro type SDE and give necessary objects to state results concerning the asymptotic behavior of recursive estimator procedures.

Published in *Trans. A. Razmadze Math. Inst.* **171** (2017), no. 1, 57–75.

In Section 2 we give a heuristic algorithm of constructing recursive estimation procedures for one-dimensional parameter of semimartingale statistical models. These procedures provide estimators asymptotically equivalent to MLE. To study the asymptotic behavior of these procedures we rewrite them in the form of the Robbins–Monro type SDE. Besides, we give a detailed description of all objects presented in this SDE, allowing us separately study special cases (e.g. discrete time case, diffusion processes, point processes, etc.).

In Section 4 we formulate main results concerning the asymptotic behaviour of recursive procedures, asymptotically equivalent to the MLE.

In Section 5, we develop recursive procedures, asymptotically equivalent to M-estimators.

Finally, in Section 6, we give various examples demonstrating the usefulness of our approach.

1. The Robbins–Monro Type SDE

Let on the stochastic basis $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geq 0}, P)$ satisfying the usual conditions the following objects be given:

a) the random field $H = \{H_t(u),\ t \geq 0,\ u \in R^1\} = \{H_t(\omega, u),\ t \geq 0,\ \omega \in \Omega,\ u \in R^1\}$ such that for each $u \in R^1$ the process $H(u) = (H_t(u))_{t\geq 0} \in \mathcal{P}$ (i.e. is predictable);

b) the random field $M = \{M(t,u),\ t \geq 0,\ u \in R^1\} = \{M(\omega, t, u),\ \omega \in \Omega,\ t \geq 0,\ u \in R^1\}$ such that for each $u \in R^1$ the process $M(u) = (M(t,u))_{t\geq 0} \in \mathcal{M}^2_{\text{loc}}(P)$;

c) the predictable increasing process $K = (K_t)_{t\geq 0}$ (i.e. $K \in \mathcal{V}^+ \cap \mathcal{P}$).

In the sequel we restrict ourselves to the consideration of the following particular case: for each $u \in R^1$ $M(u) = \varphi(u) \cdot m + W(u) * (\mu - \nu)$, where $m \in \mathcal{M}^c_{\text{loc}}(P)$, μ is an integer-valued random measure on $(R \times E, \mathcal{B}(R_+) \times \mathcal{E})$, ν is its P-compensator, $(E, \mathcal{E})$ is the Blackwell space, $W(u) = (W(t,x,u),\ t \geq 0,\ x \in E) \in \mathcal{P} \otimes \mathcal{E}$. Here we also mean that all stochastic integrals are well-defined.[1]

Later on by the symbol $\int\limits_0^t M(ds, u_s)$, where $u = (u_t)_{t\geq 0}$ is some predictable process, we denote the following stochastic line integrals:

$$\int_0^t \varphi(s, u_s)\, dm_s + \int_0^t \int_E W(s, x, u_s)(\mu - \nu)(ds, dx)$$

provided the latters are well-defined.

Consider the following semimartingale stochastic differential equation

$$z_t = z_0 + \int_0^t H_s(z_{s-})\, dK_s + \int_0^t M(ds, z_{s-}), \quad z_0 \in \mathcal{F}_0. \tag{1.1}$$

We call SDE (1.1) the Robbins–Monro (RM) type SDE if the drift coefficient $H_t(u)$, $t \geq 0$, $u \in R^1$ satisfies the following conditions: for all $t \in [0, \infty)$ P-*a.s.*

$$\text{(A)} \qquad \begin{aligned} &H_t(0) = 0, \\ &H_t(u)u < 0 \quad \text{for all} \quad u \neq 0. \end{aligned}$$

The question of strong solvability of SDE (1.1) is well-investigated (see, e.g., [5]).

[1] See [16] for basic concepts and notations.

We assume that there exists a unique strong solution $z = (z_t)_{t\geq 0}$ of equation (1.1) on the whole time interval $[0, \infty)$ and such that $\widetilde{M} \in \mathcal{M}^2_{\mathrm{loc}}(P)$, where

$$\widetilde{M}_t = \int_0^t M(ds, z_{s-}).$$

Sufficient conditions for the latter can be found in [5].

The unique solution $z = (z_t)_{t\geq 0}$ of RM type SDE (1.1) can be viewed as a semimartingale stochastic approximation procedure.

In [10] and [11], the asymptotic properties of the process $z = (z_t)_{t\geq 0}$ as $t \to \infty$ are investigated, namely, convergence ($z_t \to 0$ as $t \to \infty$ P-a.s.), rate of convergence (that means that for all $\delta < \frac{1}{2}$, $\gamma_t^\delta z_t \to 0$ as $t \to \infty$ P-a.s., with the specially chosen normalizing sequence $(\gamma_t)_{t\geq 0}$) and asymptotic expansion

$$\chi_t^2 z_t^2 = \frac{L_t}{\langle L\rangle_t^{1/2}} + R_t$$

with the specially chosen normalizing sequence χ_t^2 and martingale $L = (L_t)_{t\geq 0}$, where $R_t \to 0$ as $t \to \infty$ (see [10] and [11] for definition of objects χ_t^2, L_t and R_t).

2. Basic Model and Regularity

Our object of consideration is a parametric filtered statistical model

$$\mathcal{E} = (\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \{P_\theta; \theta \in R\})$$

associated with one-dimensional $\mathbb{F}$-adapted RCLL process $X = (X_t)_{t\geq 0}$ in the following way: for each $\theta \in R^1$ P_θ is assumed to be the unique measure on $(\Omega, \mathcal{F})$ such that under this measure X is a semimartingale with predictable characteristics $(B(\theta), C(\theta), \nu_\theta)$ (w.r.t. standard truncation function $h(x) = xI_{\{|x|\leq 1\}}$). For simplicity assume that all P_θ coincide on $\mathcal{F}_0$.

Suppose that for each pair (θ, θ') $P_\theta \overset{\mathrm{loc}}{\sim} P_{\theta'}$. Fix some $\theta_0 \in R$ and denote $P = P_{\theta_0}$, $B = B(\theta_0)$, $C = C(\theta_0)$, $\nu = \nu_{\theta_0}$.

Let $\rho(\theta) = (\rho_t(\theta))_{t\geq 0}$ be a local density process (likelihood ratio process)

$$\rho_t(\theta) = \frac{dP_{\theta,t}}{dP_t},$$

where for each θ $P_{\theta,t} := P_\theta|\mathcal{F}_t$, $P_t := P|\mathcal{F}_t$ are restrictions of measures P_θ and P on $\mathcal{F}_t$, respectively.

As it is well-known (see, e.g., [6, Ch. III, §3d, Th. 3.24]) for each θ there exists a $\widetilde{\mathcal{P}}$-measurable positive function

$$Y(\theta) = \{Y(\omega, t, x; \theta),\ \ (\omega, t, x) \in \Omega \times R_+ \times R\},$$

and a predicable process $\beta(\theta) = (\beta_t(\theta))_{t\geq 0}$ with

$$|h(Y(\theta) - 1)| * \nu \in \mathcal{A}^+_{\mathrm{loc}}(P), \quad \beta^2(\theta) \circ C \in \mathcal{A}^+_{\mathrm{loc}}(P),$$

and such that

$$\begin{aligned} &(1)\quad B(\theta) = B + \beta(\theta) \circ C + h(Y(\theta) - 1) * \nu, \\ &(2)\quad C(\theta) = C, \qquad (3)\quad \nu_\theta = Y(\theta) \cdot \nu. \end{aligned} \tag{2.1}$$

In addition, the function $Y(\theta)$ can be chosen in such a way that

$$a_t := \nu(\{t\}, R) = 1 \Longleftrightarrow a_t(\theta) := \nu_\theta(\{t\}, R) = \int Y(t, x; \theta)\nu(\{t\})dx = \widehat{Y}_t(\theta) = 1.$$

We give a definition of the regularity of the model based on the following representation of the density process as exponential matringale:

$$\rho(\theta) = \mathcal{E}(M(\theta)),$$

where

$$M(\theta) = \beta(\theta) \cdot X^c + \left(Y(\theta) - 1 + \frac{\widehat{Y}(\theta) - a}{1 - a} I_{\{0<a<1\}} \right) * (\mu - \nu) \in \mathcal{M}_{\mathrm{loc}}(P), \quad (2.2)$$

$\mathcal{E}_t(M)$ is the Dolean exponential of the martingale M (see, e.g., [16]). Here X^c is a continuous martingale part of X under measure P.

We say that the model is regular if for almost all (ω, t, x) the functions $\beta : \theta \to \beta_t(\omega; \theta)$ and $Y : \theta \to Y(\omega, t, x; \theta)$ are differentiable (notation $\dot{\beta}(\theta) := \frac{\partial}{\partial\theta}\beta(\theta)$, $\dot{Y}(\theta) := \frac{\partial}{\partial\theta}Y(\theta)$) and differentiability under integral sign is possible. Then

$$\frac{\partial}{\partial\theta} \ln \rho(\theta) = L(\dot{M}(\theta), M(\theta)) := L(\theta) \in \mathcal{M}_{\mathrm{loc}}(P_\theta),$$

where $L(m, M)$ is the Girsanov transformation defined as follows: if $m, M \in \mathcal{M}_{\mathrm{loc}}(P)$ and $Q \ll P$ with $\frac{dQ}{dP} = \mathcal{E}(M)$, then

$$L(m, M) := m - (1 + \Delta M)^{-1} \circ [m, M] \in \mathcal{M}_{\mathrm{loc}}(Q).$$

It is not hard to verify that

$$L(\theta) = \dot{\beta}(\theta) \cdot (X^c - \beta(\theta) \circ C) + \Phi(\theta) * (\mu - \nu_\theta), \quad (2.3)$$

where

$$\Phi(\theta) = \frac{\dot{Y}(\theta)}{Y(\theta)} + \frac{\dot{a}(\theta)}{1 - a(\theta)}$$

with $I_{\{a(\theta)=1\}}\dot{a}(\theta) = 0$, and $0/0 = 0$ (recall that $\frac{\partial}{\partial\theta}\widehat{Y}(\theta) = \dot{a}(\theta)$).

Indeed, due to the regularity of the model, we have

$$\dot{M}(\theta) = \dot{\beta}(\theta) \cdot X^c + \left(\dot{Y}(\theta) - \frac{\dot{a}(\theta)}{1 - a} I_{(0<a<1)} \right) * (\mu - \nu)$$

and (2.3) simply follows from (1.16)–(1.18) of [13, Part I] with

$$g(\theta) = Y(\theta) - 1 + \frac{a(\theta) - a}{1 - a} I_{(0<a<1)} ,$$

$$\psi(\theta) = \dot{Y}(\theta) - \frac{\dot{a}(\theta)}{1 - a} I_{(0<a<1)} .$$

The empirical Fisher information process is $\widehat{I}_t(\theta) = [L(\theta), L(\theta)]_t$ and if we assume that for each $\theta \in R^1$ $L(\theta) \in \mathcal{M}^2_{\mathrm{loc}}(P_\theta)$, then the Fisher information process is

$$I_t(\theta) = \langle L(\theta), L(\theta) \rangle_t.$$

3. Recursive Estimation Procedure for MLE

In [9], a heuristic algorithm was proposed for the construction of recursive estimators of unknown parameter θ asymptotically equivalent to the maximum likelihood estimator (MLE).

This algorithm was derived using the following reasons:

Consider the MLE $\widehat{\theta} = (\widehat{\theta}_t)_{t\geq 0}$, where $\widehat{\theta}_t$ is a solution of estimational equation

$$L_t(\theta) = 0.$$

The question of solvability of this equation is considered in [13, Part II].

Assume that

1) for each $\theta \in R^1$, $I_t(\theta) \to \infty$ as $t \to \infty$, P_θ-a.s., the process $(\widehat{I}_t(\theta))^{1/2}(\widehat{\theta}_t - \theta)$ is P_θ-stochastically bounded and, in addition, the process $(\widehat{\theta}_t)_{t\geq 0}$ is a P_θ-semimartingale;
2) for each pair (θ', θ) the process $L(\theta') \in \mathcal{M}^2_{\mathrm{loc}}(P_{\theta'})$ and is a P_θ-special semimartingale;
3) the family $(L(\theta), \theta \in R^1)$ is such that the Itô–Ventzel formula is applicable to the process $(L(t, \widehat{\theta}_t))_{t\geq 0}$ w.r.t. P_θ for each $\theta \in R^1$;
4) for each $\theta \in R^1$ there exists a positive increasing predictable process $(\gamma_t(\theta))_{t\geq 0}$, $\gamma_0 > 0$, asymptotically equivalent to $\widehat{I}_t^{-1}(\theta)$, i.e.

$$\gamma_t(\theta)\widehat{I}_t(\theta) \xrightarrow{P_\theta} 1 \quad \text{as} \quad t \to \infty.$$

Under these assumptions using the Ito–Ventzel formula for the process $(L(t, \widehat{\theta}_t))_{t\geq 0}$ we get an "implicit" stochastic equation for $\widehat{\theta} = (\widehat{\theta}_t)_{t\geq 0}$. Analyzing the orders of infinitesimality of terms of this equation and rejecting the high order terms we get the following SDE (recursive procedure)

$$d\theta_t = \gamma_t(\theta_{t-})L(dt, \theta_{t-}), \tag{3.1}$$

where $L(dt, u_t)$ is a stochastic line integral w.r.t. the family $\{L(t, u),\, u \in R^1,\, t \in R_+\}$ of P_θ-special semimartingales along the predictable curve $u = (u_t)_{t\geq 0}$.

Note that in many cases under consideration one can choose $\gamma_t(\theta) = (I_t^{-1}(\theta) + 1)^{-1}$, or in ergodic situations such as i.i.d. case, ergodic diffusion one can replace $I_t(\theta)$ by another process equivalent to them (see examples).

To give an explicit form to the SDE (3.1) for the statistical model associated with the semimartingale X assume for a moment that for each (u, θ) (including the case $u = \theta$)

$$|\Phi(u)| * \mu \in \mathcal{A}^+_{\mathrm{loc}}(P_\theta). \tag{3.2}$$

Then for each pair (u, θ) we have

$$\Phi(u) * (\mu - \nu_u) = \Phi(u) * (\mu - \nu_\theta) + \Phi(u)\left(1 - \frac{Y(u)}{Y(\theta)}\right) * \nu_\theta.$$

Based on this equality one can obtain the canonical decomposition of P_θ-special semimartingale $L(u)$ (w.r.t. measure P_θ):

$$\begin{aligned} L(u) &= \dot{\beta}(u) \circ (X^c - \beta(\theta) \circ C) + \Phi(u) * (\mu - \nu_\theta) \\ &\quad + \dot{\beta}(u)(\beta(\theta) - \beta(u)) \circ C + \Phi(u)\left(1 - \frac{Y(u)}{Y(\theta)}\right) * \nu_\theta. \end{aligned} \tag{3.3}$$

Now, using (3.3) the meaning of $L(dt, u_t)$ is

$$\begin{aligned}\int_0^t L(ds, u_{s-}) &= \int_0^t \dot\beta_s(u_{s-}) d(X^c - \beta(\theta)\circ C)_s + \int_0^t \int \Phi(s,x,u_{s-})(\mu - \nu_\theta)(ds,dx)\\ &\quad + \int_0^t \dot\beta_s(u_s)(\beta_s(\theta) - \beta_s(u_s)) dC_s\\ &\quad + \int_0^t \int \Phi(s,x,u_{s-})\left(1 - \frac{Y(s,x,u_{s-})}{Y(s,x,\theta)}\right)\nu_\theta(ds,dx).\end{aligned}$$

Finally, the recursive SDE (3.1) takes the form

$$\begin{aligned}\theta_t = \theta_0 &+ \int_0^t \gamma_s(\theta_{s-})\dot\beta_s(\theta_{s-}) d(X^c - \beta(\theta)\circ C)_s\\ &+ \int_0^t \int \gamma_s(\theta_{s-})\Phi(s,x,\theta_{s-})(\mu - \nu_\theta)(ds,dx)\\ &+ \int_0^t \gamma_s(\theta)\dot\beta_s(\theta_s)(\beta_s(\theta) - \beta_s(\theta_s)) dC_s\\ &+ \int_0^t \int \gamma_s(\theta_{s-})\Phi(s,x,\theta_{s-})\left(1 - \frac{Y(s,x,\theta_{s-})}{Y(s,x,\theta)}\right)\nu_\theta(ds,dx). \qquad (3.4)\end{aligned}$$

Remark 3.1. One can give more accurate than (3.2) sufficient conditions (see, e.g., [6, 16]) to ensure the validity of decomposition (3.3).

Assume that there exists an unique strong solution $(\theta_t)_{t\geq 0}$ of the SDE (3.4).

Fox arbitrary $\theta \in R^1$. To investigate the asymptotic properties, under measure P_θ, of recursive estimators $(\theta_t)_{t\geq 0}$ as $t \to \infty$, namely, a strong consistency, rate of convergence and asymptotic expansion we reduce the SDE (3.4) to the Robbins–Monro type SDE.

For this aim denote $z_t = \theta_t - \theta$. Then (3.4) can be rewritten as

$$\begin{aligned}z_t = z_0 &+ \int_0^t \gamma_s(\theta + z_{s-})\dot\beta(\theta + z_{s-})(\beta_s(\theta) - \beta_s(\theta + z_{s-}) dC_s\\ &+ \int_0^t \int \gamma_s(\theta + z_{s-})\Phi(s,x,\theta + z_{s-})\left(1 - \frac{Y(s,x,\theta + z_{s-})}{Y(s,x,\theta)}\right)\nu_\theta(ds,dx)\\ &+ \int_0^t \gamma_s(\theta + z_s)\dot\beta_s(\theta + z_s) d(X^c - \beta(\theta)\circ C)_s\\ &+ \int_0^t \int \gamma_s(\theta + z_{s-})\Phi(s,x,\theta + z_{s-})(\mu - \nu_\theta)(ds,dx). \qquad (3.5)\end{aligned}$$

For the definition of the objects K^θ, $\{H^\theta(u), u \in R^1\}$ and $\{M^\theta(u), u \in R^1\}$ we consider such a version of characteristics (C, ν_θ) that

$$\begin{aligned}C_t &= c^\theta \circ A_t^\theta,\\ \nu_\theta(\omega, dt, dx) &= dA_t^\theta B_{\omega,t}^\theta(dx),\end{aligned}$$

where $A^\theta = (A^\theta_t)_{t\geq 0} \in \mathcal{A}^+_{\mathrm{loc}}(P_\theta)$, $c^\theta = (c^\theta_t)_{t\geq 0}$ is a nonnegative predictable process, and $B^\theta_{\omega,t}(dx)$ is a transition kernel from $(\Omega \times R_+, \mathcal{P})$ in $(R, \mathcal{B}(R))$ with $B^\theta_{\omega,t}(\{0\}) = 0$ and

$$\Delta A^\theta_t B^\theta_{\omega,t}(R) \leq 1$$

(see [6, Ch. 2, §2, Prop. 2.9]).

Put $K^\theta_t = A^\theta_t$,

$$\begin{aligned} H^\theta_t(u) &= \gamma_t(\theta+u)\bigg\{\dot\beta_t(\theta+u)(\beta_t(\theta) - \beta_t(\theta+u))c^\theta_t \\ &\quad + \int \Phi(t,x,\theta+u)\bigg(1 - \frac{Y(t,x,\theta+u)}{Y(t,x,\theta)}\bigg)B^\theta_{\omega,t}(dx)\bigg\}, \qquad (3.6) \\ M^\theta(t,u) &= \int_0^t \gamma_s(\theta+u)\dot\beta_s(\theta+u)d(X^c - \beta(\theta)\circ C)_s \\ &\quad + \int_0^t \int \gamma_s(\theta+u)\Phi(s,x,\theta+u)(\mu - \nu_\theta)(ds,dx). \qquad (3.7) \end{aligned}$$

Assume that for each u, $u \in R$, $M^\theta(u) = (M^\theta(t,u))_{t\geq 0} \in \mathcal{M}^2_{\mathrm{loc}}(P_\theta)$. Then

$$\begin{aligned} \langle M^\theta(u)\rangle_t &= \int_0^t (\gamma_s(\theta+u)\dot\beta_s(\theta+u))^2 c^\theta_s dA^\theta_s \\ &\quad + \int_0^t \gamma^2_s(\theta+u)\bigg(\int \Phi^2(s,x,\theta+u)B^\theta_{\omega,s}(dx)\bigg)dA^{\theta,c}_s \\ &\quad + \int_0^t \gamma^2_s(\theta+u)B^\theta_{\omega,t}(R)\bigg\{\int \Phi^2(s,x,\theta+u)q^\theta_{\omega,s}(dx) \\ &\quad - a_s(\theta)\bigg(\int \Phi(s,x,\theta+u)q^\theta_{\omega,s}(dx)\bigg)^2\bigg\}dA^{\theta,d}_s, \end{aligned}$$

where $a_s(\theta) = \Delta A^\theta_s B^\theta_{\omega,s}(R)$, $q^\theta_{\omega,s}(dx)I_{\{a_s(\theta)>0\}} = \frac{B^\theta_{\omega,s}(dx)}{B^\theta_{\omega,s}(R)}\, I_{\{a_s(\theta)>0\}}$.

Now we give a more detailed description of $\Phi(\theta)$, $I(\theta)$, $H^\theta(u)$ and $\langle M^\theta(u)\rangle$. This allows us to study the special cases separately (see Remark 3.2 below). Denote

$$\frac{d\nu^c_\theta}{d\nu^c} := F(\theta), \quad \frac{q^\theta_{\omega,t}(dx)}{q_{\omega,t}(dx)} := f_{\omega,t}(x,\theta) \quad (:= f_t(\theta)).$$

Then

$$Y(\theta) = F(\theta)I_{\{a=0\}} + \frac{a(\theta)}{a}\, f(\theta)I_{\{a>0\}}$$

and

$$\dot Y(\theta) = \dot F(\theta)I_{\{a=0\}} + \bigg(\frac{\dot a(\theta)}{a}\, f(\theta) + \frac{a(\theta)}{a}\, \dot f(\theta)\bigg)\, I_{\{a>0\}}.$$

Therefore

$$\Phi(\theta) = \frac{\dot F(\theta)}{F(\theta)}\, I_{\{a=0\}} + \bigg\{\frac{\dot f(\theta)}{f(\theta)} + \frac{\dot a(\theta)}{a(\theta)(1-a(\theta))}\bigg\}I_{\{a>0\}} \qquad (3.8)$$

with $I_{\{a(\theta)>0\}} \int \frac{\dot f(\theta)}{f(\theta)}\, q^\theta(dx) = 0$.

Remark 3.2. Denote $\dot{\beta}(\theta) = \ell^c(\theta)$, $\frac{\dot{F}(\theta)}{F(\theta)} := \ell^\pi(\theta)$, $\frac{\dot{f}(\theta)}{f(\theta)} := \ell^\delta(\theta)$, $\frac{\dot{a}(\theta)}{a(\theta)(1-a(\theta))} := \ell^b(\theta)$.

Indices $i = c, \pi, \delta, b$ carry the following loads: "c" corresponds to the continuous part, "π" to the Poisson type part, "δ" to the predictable moments of jumps (including a main special case – the discrete time case), "b" to the binomial type part of the likelihood score $\ell(\theta) = (\ell^c(\theta), \ell^\pi(\theta), \ell^\delta(\theta), \ell^b(\theta))$.

In these notations we have for the Fisher information process:

$$I_t(\theta) = \int_0^t (\ell_s^c(\theta))^2 dC_s + \int_0^t \int (\ell_s^\pi(x;\theta))^2 B_{\omega,s}^\theta(dx) dA_s^{\theta,c} + \int_0^t B_{\omega,s}^\theta(R)\Big[\int (\ell_s^\delta(x;\theta))^2 q_{\omega,s}^\theta(dx)\Big] dA_s^{\theta,d} + \int_0^t (\ell_s^b(\theta))^2 (1-a_s(\theta)) dA_s^{\theta,d}. \tag{3.9}$$

For the random field $H^\theta(u)$ we have

$$H_t^\theta(u) = \gamma_t(\theta+u)\Big\{\ell_t^c(\theta+u)(\beta_t(\theta) - \beta_t(\theta+u))c_t^\theta \tag{3.10}$$

$$+ \int \ell_t^\pi(x;\theta+u)\Big(1 - \frac{F_t(x;\theta+u)}{F_t(x;\theta)}\Big) B_{\omega,t}^\theta(dx) I_{\{\Delta A_t^\theta = 0\}}$$

$$+ \Big\{ \int \ell_t^\delta(x;\theta+u) q_{\omega,t}^\theta(dx) \ell_t^b(\theta+u)\, \frac{a_t(\theta) - a_t(\theta+u)}{a_t(\theta)} \Big\} B_{\omega,t}^\theta(R) I_{\{\Delta A_t^\theta > 0\}}. \tag{3.11}$$

Finally, we have for $\langle M^\theta(u)\rangle$:

$$\langle M^\theta(u)\rangle_t = (\gamma(\theta+u)\ell^c(\theta+u))^2\, c^\theta \circ A_t^\theta + \int_0^t \gamma_s^2(\theta+u) \int (\ell_s^\pi(x;\theta+u))^2 B_{\omega,s}^\theta(dx) dA_s^{\theta,c} + \int_0^t \gamma_s^2(\theta+u) B_{\omega,s}^\theta(R)\Big\{ \int (\ell_s^\delta(x;\theta+u) + \ell_s^b(\theta+u))^2 q_{\omega,s}^\theta(dx) - a_s(\theta)\Big(\int (\ell_s^\delta(x;\theta+u) + \ell_s^b(\theta+u)) q_{\omega,s}^\theta(dx)\Big)^2 \Big\} dA_s^{\theta,d}. \tag{3.12}$$

Thus, we reduced SDE (3.5) to the Robbins–Monro type SDE with $K_t^\theta = A_t^\theta$, and $H^\theta(u)$ and $M^\theta(u)$ defined by (3.6) and (3.7), respectively.

As it follows from (3.6), (3.11)

$$H_t^\theta(0) = 0 \quad \text{for all} \quad t \geq 0, \quad P_\theta\text{-}a.s.$$

As for condition (A) to be satisfied it ie enough to require that for all $t \geq 0$, $u \neq 0$ P_θ-$a.s.$

$$\dot{\beta}_t(\theta+u)(\beta_t(\theta) - \beta_t(\theta+u)) < 0,$$

$$\Big(\int \frac{\dot{F}(t,x,\theta+u)}{F(t,x,\theta+u)}\Big(1 - \frac{F(t,x;\theta+u)}{F(t,x;\theta)}\Big) B_{\omega,t}^\theta(dx)\Big) I_{\{\Delta A_t^\theta = 0\}} u < 0,$$

$$\Big(\int \frac{\dot{f}(t,x;\theta+u)}{f(t,x;\theta+u)}\, q_t^\theta(dx)\Big) I_{\{\Delta A_t^\theta > 0\}} u < 0,$$

$$\dot{a}_t(\theta+u)(a_t(\theta) - a_t(\theta+u))u < 0,$$

and the simplest sufficient conditions for the latter ones is the strong monotonicity (P-a.s.) of functions $\beta(\theta)$, $F(\theta)$, $f(\theta)$ and $a(\theta)$ w.r.t. θ.

4. Main Results

We are ready to formulate main results about asymptotic properties of recursive estimators $\{\theta_t,\ t \geq 0\}$ as $t \to \infty$, (P_θ-a.s.), which is the same of solution z_t, $t \geq 0$, of equation (3.5).

For simplicity we restrict ourselves by the case when semimartingale $X = (X_t)_{t \geq 0}$ is left quasi-continuous, so $\nu(\omega; \{t\}, R) = 0$ for all $t \geq 0$, P-a.s., and $A^\theta = (A_t^\theta)_{t \geq 0}$ is a continuous process. In this case

$$H_t^\theta(u) = \gamma_t(\theta+u)\biggl\{\dot\beta_t(\theta+u)(\beta_t(\theta) - \beta_t(\theta+u))c_t^\theta + \int \frac{\dot F_t(x;\theta+u)}{F_t(x;\theta+u)}\left(1 - \frac{\dot F_t(x;\theta+u)}{F_t(x;\theta)}\right)B_{\omega,t}^\theta(dx)\biggr\}, \tag{4.1}$$

$$\langle M^\theta(u)\rangle_t = \int_0^t (\gamma_s(\theta+u)\dot\beta_s(\theta+u))^2 dA_s^\theta + \int_0^t \gamma_s^2(\theta+u)\biggl(\int \left(\frac{\dot F_s(x;\theta+u)}{F_s(x;\theta+u)}\right)^2 B_{\omega,s}^\theta(dx)\biggr)dA_s^\theta, \tag{4.2}$$

$$I_t(\theta) = \int_0^t (\dot\beta_s(\theta))^2 c_s^\theta dA_s^\theta + \int_0^t \int \left(\frac{\dot F_s(x;\theta)}{F_s(x;\theta)}\right)^2 B_{\omega,s}(dx) dA_s^\theta. \tag{4.3}$$

Theorem 4.1 (Strong consistency)**.** *Let for all $t \geq 0$, P_θ-a.s. the following conditions be satisfied:*

(A) $H_t^\theta(0) = 0,\quad H_t^\theta(u)u < 0,\ \ u \neq 0,$

(B) $h_t^\theta(u) \leq B_t^\theta(1+u^2)$, *where $B^\theta = (B_t^\theta)_{t \geq 0}$ is a predictable process, $B_t^\theta \geq 0$, $B^\theta \circ A_\infty^\theta < \infty$,*

$$h_t^\theta(u) = \frac{d\langle M^\theta(u)\rangle_t}{dA_t^\theta}\,, \tag{4.4}$$

(C) *for each ε, $\varepsilon > 0$,*

$$\inf_{\varepsilon \leq |u| \leq \frac{1}{\varepsilon}} |H^\theta(u)u| \circ A_\infty^\theta = \infty.$$

Then for each $\theta \in R^1$

$$\widehat\theta_t \to 0 \quad (\mathit{or}\ \ z_t \to 0), \quad \mathit{as}\ \ t \to \infty,\ \ P_\theta\text{-}a.s.$$

Proof. Immediately follows from conditions of Theorem 3.1 of [10] applied to prespecified by (4.1)–(4.3) objects. □

In the sequel we assume that for each $\theta \in R^1$

$$P_\theta\left(\lim_{t\to\infty} \frac{\widehat I_t(\theta)}{I_t(\theta)} = 1\right) = 1,$$

from which it follows that $\gamma_t(\theta) = I_t^{-1}(\theta)$. Denote

$$g_t^\theta = \frac{dI_t(\theta)}{dA_t^\theta} = (\dot\beta_t(\theta))^2 c_t^\theta + \int \left(\frac{\dot F_t(x;\theta)}{F_t(x;\theta)}\right)^2 B_{\omega,t}(dx). \tag{4.5}$$

We assume also that $z_t \to 0$ as $t \to \infty$, P_θ-a.s.

Theorem 4.2 (Rate of convergence)**.** *Suppose that for each δ, $0 < \delta < 1$, the following conditions are satisfied:*

$$(i) \qquad \int_0^\infty \left[\delta \, \frac{g_t^\theta}{I_t^\theta} - 2\beta_t^\theta(z_t)\right]^+ dA_t^\theta < \infty, \quad P_\theta\text{-a.s.},$$

$$\text{where} \quad \beta_t^\theta(u) = \begin{cases} -\frac{H_t^\theta(u)}{u}, & u \neq 0, \\ -\lim\limits_{u\to 0} \frac{H_t^\theta(u)}{u}, & u = 0, \end{cases} \tag{4.6}$$

$$(ii) \qquad \int_0^\infty (I_t(\theta))^\delta h_t^\theta(z_t) dA_t^\theta < \infty, \quad P_\theta\text{-a.s.}$$

Then for each $\theta \in R^1$, δ, $0 < \delta < 1$,

$$I_t^\delta(\theta) z_t^2 \to 0 \quad \text{as} \;\; t \to \infty, \quad P_\theta\text{-a.s.}$$

Proof. It is enough to note that conditions (2.3) and (2.4) of Theorem 2.1 from [11] are satisfied with $I_t(\theta)$ instead of γ_t, $\delta g_t^\theta / I_t(\theta)$ instead of r_t^δ and $\beta_t^\theta(u)$ instead of $\beta_t(u)$. □

In the sequel we assume that for all δ, $0 < \delta < \frac{1}{2}$,

$$I_t^\delta(\theta) z_t \to 0 \quad \text{as} \;\; t \to \infty, \quad P_\theta\text{-a.s.}$$

It is not hard to verify that the following expansion holds true

$$I_t^{1/2}(\theta) z_t = \frac{L_t^\theta}{\langle L^\theta \rangle_t^{1/2}} + R_t^\theta, \tag{4.7}$$

where L_t^θ, R_t^θ will be specified below.

Indeed, according to "Preliminary and Notation" section of [11]

$$\overline{\beta}_t^\theta = -\lim_{u \to 0} \frac{H_t^\theta(u)}{u} = -I_t^{-1}(\theta) g_t^\theta.$$

Further,

$$-\overline{\beta}^\theta \circ A_t^\theta = \int_0^t I_s^{-1}(\theta) \, \frac{dI_s(\theta)}{dA_s(\theta)} \, dA_s^\theta = \ln I_t(\theta).$$

Therefore

$$\Gamma_t^\theta = \varepsilon_t^{-1}(-\overline{\beta}^\theta \circ A_t^\theta) = I_t(\theta) \tag{4.8}$$

and

$$L_t^\theta = \int_0^t \Gamma_s^\theta dM^\theta(s, 0)$$

with

$$\langle L^\theta \rangle_t = \int_0^t (\Gamma_s^\theta)^2 d\langle M^\theta(0)\rangle_s = \int_0^t I_s^2(\theta) I_s^{-2}(\theta) dI_s(\theta) = I_t(\theta). \tag{4.9}$$

Finally, we obtain

$$\chi_t^\theta = \Gamma_t^\theta \langle L^\theta \rangle_t^{-1/2} = I_t^{1/2}(\theta). \tag{4.10}$$

As for R_t^θ, one can use the definition of R_t from the same section by replacing of objects by the corresponding objects with upperscipts "θ", e.g. $\overline{\beta}_t$ by $\overline{\beta}_t^\theta$, L_t by L_t^θ, etc.

Theorem 4.3 (Asymptotic expansion)**.** *Let the following conditions be satisfied:*

(i) $\langle L^\theta\rangle_t$ *is a deterministic process,* $\langle L^\theta\rangle_\infty = \infty$,

(ii) *there exists* ε, $0 < \varepsilon < \frac{1}{2}$, *such that*

$$\frac{1}{\langle L^\theta\rangle_t}\int_0^t |\beta_s^\theta - \beta_s^\theta(z_s)| I_s^{-\varepsilon}(\theta)\langle L^\theta\rangle_s dA_s^\theta \to 0 \quad as \ t\to\infty, \quad P_\theta\text{-}a.s.,$$

(iii)

$$\frac{1}{\langle L^\theta\rangle_t}\int_0^t I_t^2(\theta)(h_s^\theta(z_s,z_s) - 2h_s^\theta(z_s,0) + h_s(0,0))dA_s^\theta \overset{P_\theta}{\to} 0 \quad as \ t\to\infty,$$

where

$$h_t^\theta(u,v) = \frac{d\langle M^\theta(u), M^\theta(v)\rangle}{dA_t^\theta}\,. \tag{4.11}$$

Then in equation (4.7) *for each* $\theta\in R$

$$R_t^\theta \overset{P_\theta}{\to} 0 \quad as \ t\to\infty.$$

Proof. It is not hard to verify that all conditions of Theorem 3.1 from [11] are satisfied with $\langle L^\theta\rangle_t$ instead of $\langle L\rangle_t$, $\beta_s^\theta(u)$ instead of $\beta_s(u)$, $I_\theta^{-1}(\theta)$ instead of γ_t, A_t^θ instead of χ_t, Γ_s^θ instead Γ_s, and $I_t^{1/2}(\theta)$ instead of χ_t, $h_t^\theta(u,v)$ instead of $h_t(u,v)$, and, finally, P^θ instead of P. □

Remark. It follows from equation (4.7) and Theorem 4.3 that, using the Central Limit Theorem for martingales

$$I_t^{1/2}(\theta)(\theta_t - \theta) \overset{d}{\to} N(0,1).$$

5. Recursive Procedure for M-Estimators

As stated in previous section the maximum likelihood equation has the form

$$L_t(\theta) = L_t(\dot M_\theta, M_\theta) = 0.$$

This equation is the special member of the following family of estimational equations

$$L_t(m_\theta, M_\theta) = 0 \tag{5.1}$$

with certain P-martingales m_θ, $\theta\in R_1$. These equations are of the following sense: their solutions are viewed as estimators of unknown parameter θ, so-called M-estimators. To preserve the classical terminology we shall say that the martingale m_θ defines the M-estimator, and P_θ-martingale $L(m_\theta, M_\theta)$ is the influence martingale.

As it is well known M-estimators play the important role in robust statistics, besides they are sources to obtain asymptotically normal estimators.

Since for each $\theta\in R_1$ P_θ is an unique measure such that under this measure $X = (X_t)_{t\ge 0}$ is a semimartingale with characteristics $(B(\theta), c(\theta), \nu_\theta)$ all P_θ-martingales admit an integral representation property w.r.t. continuous martingale part and martingale measure $(\mu - \nu_\theta)$ of X. In particular, the P-martingale M_θ has the form (see Eq. (2.2))

$$M_\theta = \beta(\theta)\circ X^s + \psi * (\mu - \nu), \tag{5.2}$$

where

$$\psi(s,x,\theta) = Y(t,x,\theta) - 1 + \frac{\widehat{Y}(t,\theta) - a}{1-a}\, I_{(0<a<1)}$$

and $m_\theta \in \mathcal{M}_{\mathrm{loc}}(P)$ can be represented as

$$m(\theta) = g(\theta) \circ X^c + G(\theta) * (\mu - \nu) \tag{5.3}$$

with certain functions $g(\theta)$ and $G(\theta)$.

It can be easily shown that P_θ-martingale $L(m_\theta, M_\theta)$ can be represented as

$$L(m_\theta, M_\theta) = \varphi_m(\theta) \cdot (X^c - \beta(\theta) \circ C) + \Phi_m(\theta) * (\mu - \nu_\theta), \tag{5.4}$$

where the functions φ_m and Φ_m are expressed in terms of functions $\beta(\theta)$, $\psi(\theta)$, $g(\theta)$ and $G(\theta)$.

On the other hand, it can be easily shown that each P_θ-martingale $\widetilde{M}_\theta$ can be expressed as $L(\widetilde{m}_\theta, M_\theta)$ with P-martingale $\widetilde{m}_\theta$ defined as

$$\widetilde{m}_\theta = L(\widetilde{M}_\theta, L(-M_\theta, M_\theta)) \in \mathcal{M}_{\mathrm{loc}}(P)$$

(since $\frac{dP}{dP_\theta} = \mathcal{E}(L(-M_\theta, M_\theta))$, according to the generalized Girsanov theorem $L(\widetilde{M}_\theta, L(-M_\theta, M_\theta)) \in \mathcal{M}_{\mathrm{loc}}(P)$).

Therefore without loss of generality one can consider the M-estimator associated with the parametric family $(\widetilde{M}_\theta,\ \theta \in R)$ of P_θ-martingale as the solution of the estimational equation

$$\widetilde{M}_t(\theta) = 0. \tag{5.5}$$

In the sequel we assume that for each $\theta \in R_1$, $\widetilde{M}_\theta \in \mathcal{M}^2_{\mathrm{loc}}(P_\theta)$. Assume also that there exists a positive decreasing predictable process $\widetilde{\gamma}_t(\theta)$ with $\widetilde{\gamma}_0(\theta) = 1$ such that $\widetilde{\gamma}_t(\theta)\langle \widetilde{M}_\theta \rangle_t \overset{P_\theta}{\to} 1$ as $t \to \infty$.

Now using the same arguments as in Section 3 we introduce the following recursive procedure for constructing estimator $(\widetilde{\theta}_t,\ t \geq 0)$ asymptotically equivalent to the M-estimator defined by relation (5.5) as the solution of the following SDE

$$d\widetilde{\theta}_t = \widetilde{\gamma}_t(\theta) \widetilde{M}(dt, \widetilde{\theta}_{t-}). \tag{5.6}$$

To obtain the explicit form of the last SDE, recall that $\widetilde{M}_\theta$ has an integral representation property

$$\widetilde{M}_t(\theta) = \widetilde{\varphi}(\theta) \circ (X^c - \beta(\theta) \circ \langle X^c \rangle) + \widetilde{\Phi}(\theta) * (\mu - \nu_\theta).$$

We can obtain the canonical decomposition of P_θ-semimartingale $\widetilde{M}_t(u)$, $u \in R^1$ (w.r.t. measure P_θ)

$$\begin{aligned} \widetilde{M}(u) &= \widetilde{\varphi}(u) \circ (X^c - \beta(\theta) \circ C) + \widetilde{\Phi}(u) * (\mu - \nu_\theta) \\ &\quad + [\widetilde{\varphi}(u)(\beta(\theta) - \beta(u))] \circ C + \widetilde{\Phi}(u) \left(1 - \frac{y(u)}{y(\theta)}\right) * (\mu - \nu_\theta). \end{aligned}$$

Based on the last expression we can derive the explicit form of SDE (5.5)

$$\theta_t = \theta_0 + \int_0^t \widetilde{\gamma}_s(\widetilde{\theta}_{s-}) \widetilde{\varphi}(s, \theta_{s-}) d(X^c - \beta(\theta) \circ C)$$

$$
\begin{aligned}
&+\int_0^t\int\widetilde{\gamma}_s(\theta_{s-})\widetilde{\Phi}(s,x,\widetilde{\theta}_{s-})(\mu-\nu_\theta)(ds,dx)\\
&+\int_0^t\widetilde{\gamma}_s(\theta_{s-})\widetilde{\varphi}(s,\widetilde{\theta}_{s-})(\beta_s(\theta)-\beta_s(\theta_{s-}))dC_s\\
&+\int_0^t\int\gamma_s(\theta_{s-})\widetilde{\Phi}(s,x,\widetilde{\theta}_{s-})\bigg(1-\frac{Y(s,x,\widetilde{\theta}_{s-})}{Y(s,x,\theta)}\bigg)\nu_\theta(ds,dx). \qquad (5.7)
\end{aligned}
$$

To study the asymptotic properties of the solution of this equation $(\widetilde{\theta}_t,\ t\geq 0)$ (e.g. consistency, rate of convergence, asymptotic normality) is more convenient to rewrite this equation as $(z_t=\widetilde{\theta}_t-\theta)$

$$
\begin{aligned}
z_t=z_0&+\int_0^t\widetilde{\gamma}_s(\theta+z_{s-})\widetilde{\varphi}(s,\theta+z_{s-})d(X^c-\beta(\theta)\circ C)\\
&+\int_0^t\int\widetilde{\gamma}_s(\theta+z_{s-})\widetilde{\Phi}(s,x,\theta+z_{s-})(\mu-\nu_\theta)(ds,dx)\\
&+\int_0^t\widetilde{\gamma}_s(\theta+z_{s-})\widetilde{\varphi}(s,\theta+z_{s-})(\beta_s(\theta)-\beta_s(\theta_s+z_{s-}))dC_s\\
&+\int_0^t\int\widetilde{\gamma}_s(\theta+z_{s-})\widetilde{\Phi}(s,x,\theta+z_{s-})\bigg(1-\frac{Y(s,x,\theta+z_{s-})}{Y(s,x,\theta)}\bigg)\nu_\theta(ds,dx). \qquad (5.8)
\end{aligned}
$$

6. Examples

To make the things more clear let us begin with the simplest case of i.i.d. observations.

Example 1. Let $\{p_\theta,\ \theta\in R_1\}$ be the family of probability measures defined on some measurable space $(X,\mathcal{B})$ such that for each pair θ,θ', $p_\theta\sim p_{\theta'}$.

Put $\Omega=X^\infty$, $\mathcal{F}_n=\mathcal{B}(X^n)$, $\mathcal{F}=\mathcal{B}(X^\infty)$, $P_\theta=p_\theta\times p_\theta\times\cdots$. Then for θ,θ', $P_\theta\overset{\text{loc}}{\sim}P_{\theta'}$. Fix some $\theta_0\in R_1$ and denote $p=p_{\theta_0}$. Let $dp_\theta/dp=f(x,\theta)$. Then the local density process

$$
\rho_n(\theta)=\frac{dP_{n,\theta}}{dP_n}=\prod_{i=1}^n f(X_i,\theta)=\mathcal{E}_n(M_\theta), \qquad (6.1)
$$

where

$$
M(\theta)=\sum_{i=1}^n(f(X_i,\theta)-1)
$$

is a P-martingale. Here $(X_n)_{n\geq 1}$ is a coordinate process, $X_n(\omega)=x_n$.

Assume that for all x, $f(x,\theta)$ is continuous differentiable in θ and denote $\frac{\partial}{\partial\theta}f(X,\theta)=\dot{f}(X,\theta)$. Assume also that $\frac{\partial}{\partial\theta}\int f(x,\theta)p(dx)=\int\dot{f}(x,\theta)p(dx)$. Then $\dot{M}_n(\theta)=\sum\limits_{i=1}^n\dot{f}(X_i,\theta)$ is a P-martingale.

In these notation the MLE takes the form

$$
L_n(\dot{M}(\theta),M_\theta)=\sum_{i=1}^n\frac{\dot{f}(X_i,\theta)}{f(X_i,\theta)}=0.
$$

The Fischer information process

$$I_n(\theta) = \langle L(\dot{M}_\theta, M_\theta)\rangle = nI(\theta), \tag{6.2}$$

where $I(\theta) = E_\theta\big(\frac{\dot f(\cdot,\theta)}{f(\cdot,\theta)}\big)^2$, assuming that the last integral is finite.

The recursive estimation procedure to obtain the estimator θ_n, asymptotically equivalent to MLE is well known:

$$\theta_n = \theta_{n-1} + \frac{1}{nI(\theta_{n-1})}\,\frac{\dot f(X_n,\theta_{n-1})}{f(X_n,\theta_{n-1})}\,. \tag{6.3}$$

Let us derive this equation from the general recursive SDE.

For this aim consider the process $S_n = \sum\limits_{i=1}^{n} X_i$, $n \ge 1$. This process is a semimartingale with the jump measure

$$\mu(\omega,[0,n]\times B) = \sum_{i\le n} I_{\{X_i\in B\}}$$

and its P_θ-compensator is

$$\nu_\theta(\omega,[0,n]\times B) = \sum_{i\le n} P_\theta(X_i\in B) = n\int_B f(x,\theta)p(dx).$$

Note that $a_n(\theta) = \nu(\omega,\{n\};X) = 1$ for all $n\ge 1$ and $\theta\in R_1$.

It is obvious that $\nu_\theta = Y\cdot\nu$, where $Y_\theta(\omega,n,x)\equiv f(x,\theta)$. Besides,

$$\Phi(\theta) = \frac{\dot Y(\theta)}{Y(\theta)} + \frac{\dot a(\theta)}{1-a(\theta)} = \frac{\dot f(\cdot,\theta)}{f(\cdot,\theta)}\,.$$

At the same time the general recursive SDE for this special case can be written as

$$\theta_n = \theta_{n-1} + \frac{1}{nI(\theta_{n-1})}\,\frac{\dot f(x_n,\theta_{n-1})}{f(x_n,\theta_{n-1})} - \frac{1}{nI(\theta_{n-1})}\int \frac{\dot f(x,u)}{f(x,u)}\,\frac{f(x,u)}{f(x,\theta)}\,f(x,\theta)\,d\mu|_{u=\theta_{n-1}}\,.$$

But $\int \dot f(x,u)\,d\mu = 0$ and thus the last term equals zero and we come to equation (6.3).

In terms of $z_n = \theta_n - \theta$ equation (6.3) takes the form

$$z_n = z_{n-1} + \frac{1}{nI(\theta+z_{n-1})}\,b(\theta,z_{n-1}) + \frac{1}{nI(\theta+z_{n-1})}\,\Delta m_n,$$

where

$$b(\theta,u) = \int \frac{\dot f(x,u)}{f(x,u)}\,f(x,\theta)\,d\mu,\quad \Delta m_n = \Delta m_n(u),\quad \Delta m_n = \frac{\dot f(x,u)}{f(x,u)} - b(\theta,u).$$

Concerning to M-estimators recall that by the definition the estimational equation is

$$L_n(m(\theta),M(\theta)) = 0, \tag{6.4}$$

where $m(\theta)$ is some P-martingale, $m_n(\theta) = \sum\limits_{i\le n} g(X_i,\theta)$ with $\int g(x,\theta)\,dp = 0$.

Equation (6.4) can be written as

$$\sum_{i\le n}\frac{g(X_i,\theta)}{f(X_i,\theta)} = 0.$$

Thus, without loss of generality, we can define M-estimator as the solution of the equation

$$\widetilde{M}_n(\theta) = \sum_{i \le n} \psi(X_i, \theta) = 0, \tag{6.5}$$

where

$$\int \psi(x_i, \theta) f(x_i, \theta)\, \mu(dx) = 0, \quad \langle \widetilde{M}(\theta) \rangle_n = n \int \psi^2(x, \theta) f(x, \theta)\, \mu(dx) = n I_\psi(\theta).$$

Now using the same arguments as in the case of MLE we obtain the following recursive procedure for construction the estimator asymptotically equivalent to the M-estimator defined by (6.5)

$$\theta_n = \theta_{n-1} + \frac{1}{n I_\psi(\theta_{n-1})}\, \psi(X_n, \theta_{n-1}).$$

Example 2. Discrete time case.

Let $X_0, X_1, \dots, X_n, \dots$ be observations taking values in some measurable space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ such that the regular conditional densities of distributions (w.r.t. some measure p) $f_i(x_i, \theta | x_{i-1}, \dots, x_0)$, $i \le n$, $n \ge 1$ exist, $f_0(x_0, \theta) \equiv f_0(x_0)$, $\theta \in R^1$ is the parameter to be estimated. Denote P_θ corresponding distribution on $(\Omega, \mathcal{F}) := (\mathcal{X}^\infty, \mathcal{B}(\mathcal{X}^\infty))$. Identify the process $X = (X_i)_{i \ge 0}$ with coordinate process and denote $\mathcal{F}_0 = \sigma(X_0)$, $\mathcal{F}_n = \sigma\ (X_i, i \le n)$. If $\psi = \psi(X_i, X_{i-1}, \dots, X_0)$ is a r.v., then under $E_\theta(\psi | \mathcal{F}_{i-1})$ we mean the following version of conditional expectation

$$E_\theta(\psi \mid \mathcal{F}_{i-1}) := \int \psi(z, X_{i-1}, \dots, X_0) f_i(z, \theta \mid X_{i-1}, \dots, X_0) \mu(dz),$$

if the last integral exists.

Assume that the usual regularity conditions are satisfied and denote

$$\frac{\partial}{\partial \theta} f_i(x_i, \theta \mid x_{i-1}, \dots, x_0) := \dot{f}_i(x_i, \theta \mid x_{i-1}, \dots, x_0),$$

the maximum likelihood scores

$$l_i(\theta) := \frac{\dot{f}_i}{f_i}\,(X_i, \theta \mid X_{i-1}, \dots, X_0)$$

and the empirical Fisher information

$$I_n(\theta) := \sum_{i=1}^{n} E_\theta(l_i^2(\theta) \mid \mathcal{F}_{i-1}).$$

Denote also

$$b_n(\theta, u) := E_\theta(l_n(\theta + u) \mid \mathcal{F}_{n-1})$$

and indicate that for each $\theta \in R^1$, $n \ge 1$

$$b_n(\theta, 0) = 0 \ \ (P_\theta\text{-}a.s.). \tag{6.6}$$

Using the same arguments as in the case of i.i.d. observations we come to the following recursive procedure

$$\theta_n = \theta_{n-1} + I_n^{-1}(\theta_{n-1}) l_n(\theta_{n-1}), \quad \theta_0 \in \mathcal{F}_0.$$

Fix θ, denote $z_n = \theta_n - \theta$ and rewrite the last equation in the form

$$\begin{aligned} z_n &= z_{n-1} + I_n^{-1}(\theta + z_{n-1}) b_n(\theta, z_{n-1}) + I_n^{-1}(\theta + z_{n-1}) \Delta m_n, \\ z_0 &= \theta - \theta, \end{aligned} \tag{6.7}$$

where $\Delta m_n = \Delta m(n, z_{n-1})$ with $\Delta m(n,u) = l_n(\theta+u) - E_\theta(l_n(\theta+u)|\mathcal{F}_{n-1})$.

Note that the algorithm (6.7) is embedded in SDE (1.1) with

$$\begin{aligned} H_n(u) &= I_n^{-1}(\theta+u) b_n(\theta, u) \in \mathcal{F}_{n-1}, \quad \Delta K_n = 1, \\ \Delta M(n,u) &= I_n^{-1}(\theta+u) \Delta m(n,u). \end{aligned}$$

This example clearly shows the necessity of consideration of random fields $H_n(u)$ and $M(n,u)$.

The discrete time case was considered by T. Sharia in [20, 21].

Example 3. Recursive parameter estimation in the trend coefficient of a diffusion process.

Here we consider the problem of recursive estimation of the one-dimensional parameter in the trend coefficient of a diffusion process $\xi = \{\xi_t,\ t \geq 0\}$ with

$$d\xi_t = a(\xi_t, \theta)\, dt + \sigma(\xi_t)\, dw_t, \ \ \xi_0, \tag{6.8}$$

where $w = \{w_t,\ t \geq 0\}$ is a standard Wiener process, $a(\cdot, \theta)$ is the known function, $\theta \in \Theta \subseteq R$ is a parameter to be estimated, Θ is some open subset of R, $\sigma^2(\cdot)$ is the known diffusion coefficient.

We assume that there exists a unique weak solution of equation (6.8).

For each $\theta \in \Theta$ denote by P^θ the distribution of the process ξ on $(C_{[0,\infty)}, \mathcal{B})$.

Let $X = \{X_t,\ t \geq 0\}$ be the coordinate process, that is, for each $x = \{x_t, t \geq 0\} \in C_{[0,\infty)},\ X_t(x) = x_t, t \geq 0$.

Fix some $\theta \in \Theta$ and assume that for each $\theta' \in \Theta$, $P^\theta \overset{(loc)}{\sim} P^{\theta'}$. Then the density process $\rho_t(X, \theta)$ can be written as

$$\rho_t(X,\theta) := \frac{dP_t^\theta}{dP_t^{\theta'}}(X) = \exp\left\{ \int_0^t \frac{a(X_s,\theta) - a(X_s,\theta')}{\sigma(X_s)} \frac{(dX_s - a(X_s,\theta')ds)}{\sigma(X_s)} \right. \\ \left. - \frac{1}{2} \int_0^t \left(\frac{a(X_s,\theta) - a(X_s,\theta')}{\sigma(X_s)} \right)^2 ds. \right.$$

Recall that if for all $t \geq 0$ P^θ-a.s.

$$\int_0^1 \sigma^2(X_s)\, ds < \infty, \tag{6.9}$$

then the process $\left\{ X_t - \int\limits_0^t a(X_s, \theta)\, ds,\ t \geq 0 \right\} \in M_{loc}^2(P^\theta)$ with the square characteristic $\int\limits_0^t \sigma^2(X_s)\, ds$.

Under suitable regularity conditions if we assume that for all $t \geq 0$ P^θ-a.s.

$$\int_0^t \left(\frac{\dot{a}(X_s,\theta)}{\sigma(X_s)} \right)^2 ds < \infty, \tag{6.10}$$

we will have

$$\left\{\frac{\partial}{\partial\theta}\ln\rho_t(X,\theta)=\int_0^t\left(\frac{\dot a(X_s,\theta)}{\sigma(X_s)}\right)d(X_s-a(X_s,\theta)ds),\ \ t\ge 0\right\}\in M^2_{loc}(P^\theta),$$

where $\dot a(\cdot,\theta)$ denotes the derivative of $a(\cdot,\theta)$ w.r.t. θ.

Below we assume that conditions (6.9) and (6.10) are satisfied.

Introduce the Fisher information process

$$I_t(\theta)=\int_0^t\left(\frac{\dot a(X_s,\theta)}{\sigma(X_s)}\right)^2ds.$$

Then, according to equation (3.4), the SDE for constructing the recursive estimator $(\theta_t, t\ge 0)$ has the form

$$d\theta_t=I_t(\theta_t)\left[\frac{\dot a(X_t,\theta_t)}{\sigma^2(X_s)}\,dX_t^c+\frac{\dot a(X_t,\theta_t)}{\sigma^2(X_t)}\,(a(X_t,\theta)-a(X_t,\theta_t))\,dt\right]. \tag{6.11}$$

Fix some $\theta\in\Theta$. To study the asymptotic properties of the recursive estimator $\{\theta_t, t\ge 0\}$ as $t\to\infty$ under measure P^θ let us denote $z_t=\theta_t-\theta$ and rewrite (6.11) in the following form:

$$dz_t=I_t(\theta+z_t)\left[\frac{\dot a(X_t,\theta+z_t)}{\sigma^2(X_s)}\,dX_t^c+\frac{\dot a(X_t,\theta+z_t)}{\sigma^2(X_t)}\,(a(X_t,\theta)-a(X_t,\theta+z_t))\,dt\right]. \tag{6.12}$$

In the sequel we assume that there exists a unique strong solution of equation (6.12) such that

$$\left\{\int_0^t I_s(\theta+z_s)\,\frac{\dot a(X_s,\theta+z_s)}{\sigma^2(X_s)}\,dX_s^c,\ \ t\ge 0\right\}\in M^2_{loc}(P_\theta),$$

that is, for each $t\ge 0$ P^θ-a.s.

$$\int_0^t I_s^2(\theta+z_s)\left(\frac{\dot a(X_s,\theta+z_s)}{\sigma(X_s)}\right)^2ds<\infty.$$

To study the asymptotic properties of the process $z=\{z_t,\ t\ge 0\}$ as $t\to\infty$ (under the measure P^θ) one can use the results of Theorems 4.1–4.3 concerning the asymptotic behaviour of solutions of the Robbins–Monro type SDE

$$z_t=z_0+\int_0^t H_s(z_{s-})\,dK_s+\int_0^t M(ds,z_{s-}). \tag{6.13}$$

Note that equation (6.13) covers equation (6.12) with $K_t=t$,

$$H_t(u):=H_t^\theta(u)=I_t(\theta+u)\,\frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)}\,(a(X_t,\theta)-a(X_t,\theta+u)),\ \ H_t^\theta(0)=0, \tag{6.14}$$

$$M(u):=M^\theta(u)=\left\{M^\theta(t,u)=\int_0^t I_s(\theta+u)\,\frac{\dot a(X_t,\theta+u)}{\sigma^2(X_t)}\,dX_s^c,\ \ t\ge 0\right\}. \tag{6.15}$$

Let for each $u\in R$ the process $M^\theta(u)\in M^2_{loc}(P^\theta)$. Then

$$\langle M^\theta(u),M^\theta(v)\rangle_t=\int_0^t h_s(u,v)\,ds,$$

where

$$h_t(u,v) = h_t^\theta(u,v) = I_t(\theta+u)I_t(\theta+v)\,\frac{\dot a(X_t,\theta+u)\dot a(X_t,\theta+v)}{\sigma^2(X_t)}\,. \tag{6.16}$$

This problem is fully studied by Lazrieva and Toronjadze in [9].

Example 4. Let $(\Omega, \mathcal{F} = (\mathcal{F}_t)_{t\ge 0}, P, P_\theta, \theta \in R_1)$ be filtered probability space and $M = (M_t)_{t\ge 0}$ be a P-martingale with the deterministic characteristic $\langle M\rangle_t$, $\langle M\rangle_\infty = \infty$. Let for each $\theta \in R_1$ P_θ be unique measure on $(\Omega,\mathcal{F})$ such that the process $X(t)$ follows the equation

$$X_t = X_0 + a(\theta)\langle M\rangle_t + M_t,$$

where $a(\theta)$ is known function depending on the unknown parameter θ. Then for each pair (θ,θ'), $P_\theta \overset{\text{loc}}{\sim} P_{\theta'}$. Fix some $\theta_0 \in R_1$. Then the local density process

$$\rho_t(\theta) = \frac{dP_{\theta,t}}{dP_{\theta_0,t}} = \mathcal{E}_t(M(\theta)),$$

where

$$M_t(\theta) = (a(\theta)-a(\theta_0))(X_t - a(\theta_0)\langle M\rangle_t). \tag{6.17}$$

Assume that $a(\theta)$ is strongly monotone function continuously differentiable in θ. Then

$$L_t(\theta) = \frac{\partial}{\partial\theta}\ln\rho_t(\theta) = L_t(\dot M(\theta), M(\theta)) = \dot a(\theta)(X_t - a(\theta)\langle M\rangle_t)$$

and the Fischer information process os

$$I_t(\theta) = \langle L(\theta), L(\theta)\rangle_t = [\dot a(\theta)]^2\langle M\rangle_t\,.$$

Put $\gamma_t(\theta) = [\dot a(\theta)]^{-2}\frac{1}{\langle M\rangle_t+1} = [\dot a(\theta)]^{-2}\gamma_t^{-1}$ (with the obvious notation $\gamma_t = \langle M\rangle_t + 1$). Therefore the recursive estimation procedure to obtain estimator asymptotically equivalent to the MLE θ_t is

$$\begin{aligned}\theta_t = \theta_0 &+ \int_0^t \frac{1}{\langle M\rangle_s+1}\,\frac{a(\theta)-a(\theta_s)}{\dot a(\theta_s)}\,d\langle M\rangle_s \\ &+ \int_0^t \frac{1}{1+\langle M\rangle_s}\,\frac{1}{\dot a(\theta_s)}\,d(X_s - a(\theta)\langle M\rangle_s).\end{aligned} \tag{6.18}$$

Denote $z_t = \theta_t - \theta$ and rewrite the last equation

$$\begin{aligned}dz_t = &\frac{1}{\langle M\rangle_t+1}\,\frac{a(\theta)-a(\theta+z_t)}{\dot a(\theta+z_t)}\,d\langle M\rangle_t \\ &+ \frac{1}{\langle M\rangle_t+1}\,\frac{1}{\dot a(\theta+z_t)}\,d(X_t - a(\theta)\langle M\rangle_t).\end{aligned} \tag{6.19}$$

Further, denote

$$\begin{aligned}H_t(\theta,u) &= \frac{1}{\langle M\rangle_t+1}\,\frac{a(\theta)-a(\theta+z_t)}{\dot a(\theta+z_t)}\,, \\ M_t(\theta,u) &= \int_0^t \frac{1}{\langle M\rangle_s+1}\,\frac{1}{\dot a(\theta+u)}\,d(X_s - a(\theta)\langle M\rangle_t).\end{aligned}$$

In these notation equation (6.19) is the Robbins–Monro type equation

$$dz_t = H_t(\theta, z_t)d\langle M\rangle_t + dM_t(\theta, z_t). \tag{6.20}$$

Indeed, condition (A) of Theorem 4.1 is satisfied since

$$H_t(\theta, 0) = 0 \quad \text{and} \quad H_t(\theta, u)u < 0 \quad \text{for all} \;\; u \neq 0.$$

We study the asymptotic behavior of z_t as $t \to \infty$ under measure P_θ.

1) Convergence: $z_t \to 0$ as $t \to \infty$ P_θ-a.s. or $\theta_t \to \theta$ as $t \to \infty$ P_θ-a.s. (strong consistency).

Proposition 6.1. *Let the following condition be satisfied*

$$[\dot a(\theta + u)]^2(1 + u^2) \geq c, \tag{6.21}$$

where c os some constant depending on θ. Then

$$z_t \to 0 \quad \textit{as} \;\; t \to \infty \;\; P_\theta\textit{-a.s.}$$

Proof. Let us check conditions (A), (B), (C) of Theorem 4.1. (A) is evident. Concerning condition (B) note that

$$\langle M(\theta, u)\rangle_t = \frac{1}{(\dot a(\theta + u))^2} \int_0^t \frac{1}{(\langle M\rangle_s + 1)^2}\, d\langle M\rangle_s$$

and

$$h_t(\theta, u) = \frac{1}{(\dot a(\theta + u))^2}\frac{1}{(\langle M\rangle_t + 1)^2}\,.$$

Then if we denote $B_t = \frac{1}{(\langle M\rangle_t + 1)^2}$, taking into account equation (6.21) we simply obtain

$$h_t(\theta, u) \leq B_t(1 + u^2) \quad \text{with} \quad B \circ \langle M\rangle_\infty < \infty\,.$$

As for condition (C), we have to verify that for each $\varepsilon > 0$

$$\inf_{\varepsilon \leq u \leq \frac{1}{\varepsilon}} \left|\frac{a(\theta) - a(\theta + u)}{\dot a(\theta + u)}\right| \int_0^\infty \frac{d\langle M\rangle_t}{\langle M\rangle_t + 1} = \infty.$$

The last condition is satisfied if for each $\varepsilon > 0$

$$\inf_{\varepsilon \leq |u| \leq \frac{1}{\varepsilon}} \left|\frac{a(\theta) - a(\theta + u)}{\dot a(\theta + u)}\right| > 0,$$

which holds since $\dot a(\theta)$ is continuous. □

2) Rate of convergence. Here we assume that $z_t \to 0$ as $t \to \infty$ P_θ-a.s.

Proposition 6.2. *For all δ, $0 < \delta < \frac{1}{2}$, we have*

$$\gamma_t^\delta z_t = (\langle M\rangle_t + 1)^\delta z_t \to 0 \quad \textit{as} \quad t \to \infty, \;\; P_\theta\textit{-a.s.}$$

Proof. We have to check conditions (i) and (ii) of Theorem 4.2.

Condition (ii) is satisfied. Indeed, for all $0 < \delta < 1$

$$\int_0^\infty (\langle M\rangle_t + 1)^\delta[\dot a(\theta + u)]^{-2}\frac{1}{(\langle M\rangle_t + 1)^2}\, d\langle M\rangle_t < \infty.$$

As for condition (i), it is enough to verify that for all δ, $0 < \delta < \frac{1}{2}$,

$$\int_0^\infty \frac{1}{\langle M\rangle_t + 1}\left[\delta - I_{(z_t=0)} - \frac{a(\theta) - a(\theta + z_t)}{z_t\,\dot a(\theta + z_t)}\right]^+ d\langle M\rangle_t < \infty\,.$$

But $\left[\delta - I_{(z_t=0)} - \frac{a(\theta)-a(\theta+z_t)}{z_t\,\dot a(\theta+z_t)}\, I_{\{z_t\neq 0\}}\right]^+ = 0$ eventually since $z_t \to 0$. □

3) Asymptotic expansion. Here we assume that for all δ, $0 < \delta < \frac{1}{2}$, $\gamma_t^\delta z_t \to 0$ as $t \to \infty$ P_θ-a.s.

Proposition 6.3. *Let there exist some $\varepsilon > 0$, $\gamma > 0$ and $c(\theta)$ such that*

$$|\dot a(\theta + u) - \dot a(\theta + v)| \le c|u - v|^\gamma \tag{6.22}$$

for all $(u, v) \in O_\varepsilon(0)$, then all conditions of Theorem 4.3 are satisfied and the following asymptotic expansion holds true

$$(1 + \langle M\rangle_t)^{1/2}\dot a(\theta) z_t = \frac{L_t}{\langle L\rangle_t^{1/2}} + R_t\,,$$

where $R_t \to 0$ as $t \to \infty$ P-a.s., $L_t = [\dot a(\theta)]^{-1}(X_t - a(\theta)\langle M\rangle_t)$.

Example 5 (Point process with continuous compensator). Let Ω be a space of piecewise constant functions $x = (x_t)_{t\ge 0}$ such that $x_0 = 0$, $x_t = x_{t-} + (0 \text{ or } 1)$, $\mathcal{F} = \sigma\{x : x_s, s \ge 0\}$ and $\mathcal{F}_t = \sigma\{x : x_s, 0 < s \le t\}$. Let for $x \in \Omega$

$$\tau_n(x) = \inf\{s : s > 0,\ \ x_s = n\}$$

setting $\tau_n(\infty) = \infty$ if $\lim\limits_{t\to\infty} x_t < n$. Let $\tau_\infty(x) = \lim\limits_{n\to\infty} \tau_n(x)$.

Note that $x = (x_t)_{t\ge 0}$ can be written as

$$x_t = \sum_{n\ge 1} I_{\{\tau_n(x)\le t\}},$$

and so $(x_t)_{t\ge 0}$ and the family of σ-algebras $(\mathcal{F}_t)_{t\ge 0}$ are right-continuous.

Let for each $\theta \in R_1$ P_θ be a probability measure on $(\Omega, \mathcal{F})$ such that under this measure the coordinate process $X_t(\omega) = x_t$ if $\omega = (x_t)_{t\ge 0}$ is a point process with compensator $A_t(\theta) = A(\theta)A(t)$, where $A(t) = A(t, \omega)$ is an increasing process with continuous trajectories (P_θ-a.s.), $A(0) = 0$, $P_\theta\{A_\infty = \infty\} = 1$, and for each $t > 0$ $P_{\dot\theta}(A_t < \infty) = 1$, $A(\theta)$ is a strongly monotone deterministic function, $A(\theta) > 0$, and $A(\theta)$ is continuously differentiable (denote $\dot A(\theta) = \frac{d}{d\theta}\, A(\theta)$).

Assume that for each pair (θ, θ'), $P_\theta \overset{loc}{\sim} P_{\theta'}$. Fix as usual some $\theta_0 \in R_1$. Then the local density process $\rho_t(\theta) = \frac{dP_{\theta,t}}{dP_{\theta_0,t}}$ can be represented as

$$\rho_t(\theta) = \mathcal{E}_t(M(\theta)),$$

where

$$M_t(\theta) = \left(\frac{A(\theta)}{A(\theta_0)} - 1\right)(X_t - A(\theta_0)A_t).$$

Therefore $L_t(\theta) = \frac{\partial}{\partial\theta}\, \ln \rho_t(\theta)$ has the form

$$L_t(\theta) = L_t(\dot M(\theta), M(\theta)) = \frac{\dot A(\theta)}{A(\theta)}\,(X_t - A(\theta)A(t)).$$

The Fisher information process is

$$I_t(\theta) = \langle L(\dot{M}(\theta), M(\theta)) \rangle_t = \left[\frac{\dot{A}(\theta)}{A(\theta)} \right]^2 A(\theta) A(t).$$

Put $\gamma_t(\theta) = \frac{A(\theta)}{[\dot{A}(\theta)]^2} \frac{1}{A(t)+1}$. It is evident that

$$\lim_{t \to \infty} \gamma_t(\theta) I_t(\theta) = 1.$$

Note that the process $(X_t)_{t \geq 0}$ is a P_θ-semimartingale with the triplet of characteristics $(A(\theta)A(t), 0, A(\theta)A(t))$. Therefore, according to Section 3,

$$F(\theta) = F(\omega, t, x, \theta) = \frac{A(\theta)}{A(\theta_0)}, \quad \Phi(\theta) = \frac{\dot{A}(\theta)}{A(\theta)},$$

$$\ell^c(\theta) = \ell^\delta(\theta) = \ell^b(\theta) = 0, \quad \ell^\pi(\theta) = \frac{\dot{A}(\theta)}{A(\theta)}.$$

Thus from (3.11) we obtain

$$H_t^\theta(u) = \frac{1}{A(t)+1} \frac{A(\theta) - A(\theta+u)}{\dot{A}(\theta+u)},$$

$$M^\theta(t,u) = \frac{1}{\dot{A}(\theta+u)} \int_0^t \frac{1}{A(s)+1} \, d(X_s - A(\theta)A(s)),$$

and the equation for $z_t = \theta_t - \theta$ is

$$dz_t = \frac{1}{A(t)+1} \frac{A(\theta) - A(\theta+z_t)}{\dot{A}(\theta+z_t)} \, dA(t) + \frac{1}{A(t)+1} \frac{1}{\dot{A}(\theta+z_t)} \, d(X_t - A(\theta)A(t)), \quad (6.23)$$

where $(\theta_t)_{t \geq 0}$ is recursive estimation satisfying the equation

$$d\theta_t = \frac{1}{A(t)+1} \frac{A(\theta) - A(\theta_t)}{\dot{A}(\theta_t)} \, dA(t) + \frac{1}{A(t)+1} \frac{1}{\dot{A}(\theta_t)} \, d(X_t - A(\theta)A(t)).$$

As one can see the equation (6.23) is quite similar to (6.19) with $A(\theta)$ instead of $a(\theta)$ and $A(t)$ instead of $\langle M \rangle_t$.

Now if conditions (6.21) and (6.22) with $A(\theta)$ instead of $a(\theta)$ and $A(t)$ instead of $\langle M \rangle_t$ are satisfied, then the asymptotic expansion holds true

$$(A(t)+1)^{1/2} \dot{A}(\theta) z_t = \frac{L_t}{\langle L \rangle_t^{1/2}} + R_t,$$

where $R_t \to 0$ as $t \to \infty$ P_θ-a.s., $L_t = [\dot{A}(\theta)]^{-1}(X_t - A(\theta)A(t))$.

References

[1] A. E. Albert and L. A. Gardner, Jr., Stochastic approximations and nonlinear regression, M.I.T. Press Research Monograph No. 42. The M.I.T. Press, Cambridge, Mass. 1967.

[2] A. Le Breton, About Gaussian schemes in stochastic approximation, Stochastic Process. Appl. 50 (1994), no. 1, 101–115.

[3] H. F. Chen, Asymptotically efficient stochastic approximation, Stochastics Stochastics Rep. 45 (1993), no. 1-2, 1–16.

[4] H. F. Chen and W. Zhao, Recursive Identification and Parameter Estimation, CRC Press, Boca Raton, FL, 2014.

[5] L. I. Gal'čuk, On the existence and uniqueness of solutions of stochastic equations with respect to semimartingales, Teor. Veroyatnost. i Primenen. 23 (1978), no. 4, 782–795 (in Russian).
[6] J. Jacod and A. N. Shiryaev, Limit theorems for stochastic processes, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 288. Springer-Verlag, Berlin, 1987.
[7] H. J. Kushner and G. G. Yin, Stochastic approximation algorithms and applications, Applications of Mathematics (New York), 35. Springer-Verlag, New York, 1997.
[8] H. J. Kushner and G. G. Yin, Stochastic Approximation and Recursive Algorithms and Applications, Second edition. Applications of Mathematics (New York), 35. Stochastic Modelling and Applied Probability. Springer-Verlag, New York, 2003.
[9] N. L. Lazrieva and T. A. Toronjadze, Ito–Ventzel's formula for semimartingales, asymptotic properties of MLE and recursive estimation, Stochastic differential systems, Proc. IFIP-WG 7/1 Work. Conf., Eisenach/GDR 1986, Lect. Notes Control Inf. Sci. 96 (1987), 346–355.
[10] N. Lazrieva, T. Sharia and T. Toronjadze, The Robbins-Monro type stochastic differential equations. I. Convergence of solutions, Stochastics Stochastics Rep. 61 (1997), no. 1-2, 67–87.
[11] N. Lazrieva, T. Sharia and T. Toronjadze, T. The Robbins-Monro type stochastic differential equations. II. Asymptotic behaviour of solutions, Stochastics Stochastics Rep. 75 (2003), no. 3, 153–180.
[12] N. Lazrieva, T. Sharia and T. Toronjadze, Semimartingale stochastic approximation procedure and recursive estimation. Martingale theory and its application, J. Math. Sci. (N. Y.) 153 (2008), no. 3, 211–261.
[13] N. Lazrieva and T. Toronjadze, Asymptotic theory of M-estimators in general statistical models. Parts I and II, Centrum Voor Wiskunde en Informatica, Reports BS-R9010-20 (1990).
[14] D. Levanony, A. Shwartz and O. Zeitouni, Recursive identification in continuous-time stochastic processes, Stochastic Process. Appl. 49 (1994), no. 2, 245–275.
[15] Y. Liang, A. Tovaneswaran and B. Abraham, Recent developments in recursive estimation for time-series models, Inter. J. Statistics Probab. 5 (2016), No. 2.
[16] R. Sh. Liptser and A. N. Shiryayev, Martingale Theory, Probability Theory and Mathematical Statistics. Nauka, Moscow, 1986 (in Russian).
[17] A. V. Melnikov and È. Valkeĭla, Martingale models of stochastic approximation and their convergence, Teor. Veroyatnost. i Primenen. 44 (1999), no. 2, 278–311 (in Russian); translation in Theory Probab. Appl. 44 (2000), no. 2, 330–360.
[18] M. B. Nevelson and R. Z. Khas'minskiĭ, Stochastic Approximation and Recurrent Estimation, Monographs in Probability Theory and Mathematical Statistics. Nauka, Moscow, 1972 (in Russian).
[19] H. Robbins and S. Monro, A stochastic approximation method, Ann. Math. Statistics 22 (1951), 400–407.
[20] T. Sharia, On the recursive parameter estimation in the general discrete time statistical model, Stochastic Process. Appl. 73 (1998), no. 2, 151–172.
[21] T. Sharia, Recursive parameter estimation: convergence, Stat. Inference Stoch. Process. 11 (2008), no. 2, 157–175.
[22] P. Spreij, Recursive approximate maximum likelihood estimation for a class of counting process models, J. Multivariate Anal. 39 (1991), no. 2, 236–245.

ON REGULARITY OF PRIMAL AND DUAL DYNAMIC VALUE FUNCTIONS RELATED TO INVESTMENT PROBLEM AND THEIR REPRESENTATIONS AS BACKWARD STOCHASTIC PDE SOLUTIONS

M. MANIA AND R. TEVZADZE

Abstract. We study regularity properties of the dynamic value functions of primal and dual problems of optimal investing for utility functions defined on the whole real line. Relations between decomposition terms of value processes of primal and dual problems and between optimal solutions of basic and conditional utility maximization problems are established. These properties are used to show that the value function satisfies a corresponding backward stochastic partial differential equation. In the case of complete markets we give conditions on the utility function when this equation admits a solution.

Key words and phrases: Utility maximization, Complete and incomplete markets, Duality, Backward stochastic partial differential equation, Value function

MSC 2010: 90A09, 60H30, 90C39

1. Introduction

We consider a financial market model, where the dynamics of asset prices is described by the continuous semimartingale S defined on the complete probability space $(\Omega, \mathcal{F}, P)$ with continuous filtration $F = (F_t, t \in [0,T])$, where $\mathcal{F} = F_T$ and $T < \infty$. We work with discounted terms, i.e. the bond is assumed to be a constant.

Denote by $\mathcal{M}^e$ (resp. $\mathcal{M}^a$) the set of probability measures Q equivalent (resp. absolutely continuous with respect) to P such that S is a local martingale under Q.

Throughout the paper we assume that the filtration F is continuous (i.e. all F-local martingales are continuous) and

$$\mathcal{M}^e \neq \emptyset. \tag{1}$$

The continuity of F and the existence of an equivalent martingale measure imply that the structure condition is satisfied, i.e. S admits the decomposition

$$S_t = M_t + \int_0^t \lambda_s \, d\langle M\rangle_s, \quad \int_0^t \lambda_s^2 \, d\langle M\rangle_s < \infty$$

for all t P-a.s., where M is a continuous local martingale and λ is a predictable process.

Let $U = U(x) : \mathbb{R} \to \mathbb{R}$ be a utility function taking finite values at all points of real line $\mathbb{R}$ such that U is continuously differentiable, increasing, strictly concave and satisfies the Inada conditions

$$U'(\infty) = \lim_{x\to\infty} U'(x) = 0, \quad U'(-\infty) = \lim_{x\to-\infty} U'(x) = \infty. \tag{2}$$

Published in *SIAM J. Financial Math.* **8** (2017), no. 1, 483–503.

We also assume that U satisfies the condition of reasonable asymptotic elasticity (see [6] and [13] for details), i.e.

$$\limsup_{x\to\infty}\frac{xU'(x)}{U(x)}<1,\quad \liminf_{x\to-\infty}\frac{xU'(x)}{U(x)}>1. \tag{3}$$

We consider the utility maximization problem, i.e. the problem of finding a trading strategy $(\pi_t, t\in[0,T])$ such that the expected utility of terminal wealth $X_T^{x,\pi}$ becomes maximal. The wealth process, determined by a self-financing trading strategy π and initial capital x, is defined as a stochastic integral

$$X_t^{x,\pi}=x+\int_0^t \pi_u dS_u,\quad 0\le t\le T.$$

The predictable, $S-$integrable process π we call admissible if the stochastic integral $(\int_0^t \pi_u dS_u, t\in[0,T])$ is uniformly bounded from below.

The value function V associated to the problem is given by

$$V(x)=\sup_{\pi\in\Pi}E\Big[U\Big(x+\int_0^T \pi_u\,dS_u\Big)\Big], \tag{4}$$

where Π is the class of admissible strategies.

For the utility function U we denote by $\widetilde{U}$ its convex conjugate

$$\widetilde{U}(y)=\sup_x(U(x)-xy),\quad y>0. \tag{5}$$

The dual problem to (4) is

$$\widetilde{V}(y)=\inf_{Q\in\mathcal{M}^e}E[\widetilde{U}(y\rho_T^Q)],\quad y>0, \tag{6}$$

where $\rho_t^Q=dQ_t/dP_t$ is the density process of the measure $Q\in\mathcal{M}^e$ relative to the basic measure P.

Let τ be a stopping time valued in $[0,T]$. Denote by Π_τ the class of admissible processes, such that $\pi=\pi 1_{[\tau,T]}$. Define

$$\mathcal{Z}_{\tau,y}=\{Y: Y=y\frac{\rho_T}{\rho_\tau},\ \rho_T=\frac{dQ}{dP},\ Q\in\mathcal{M}^e(S)\}.$$

The dynamic value functions of primal and dual problems are defined as

$$V(\tau,x)=\operatorname*{ess\,sup}_{\pi\in\Pi_\tau}E\Big[U\Big(x+\int_\tau^T \pi_u\,dS_u\Big)\Big|F_\tau\Big], \tag{7}$$

$$\widetilde{V}(\tau,y)=\operatorname*{ess\,inf}_{Y\in\mathcal{Z}_{\tau,y}}E\Big[\widetilde{U}(Y)\mid F_\tau\Big],\quad y>0. \tag{8}$$

For $V(0,x)$ and $\widetilde{V}(0,y)$ we use the notation $V(x)$ and $\widetilde{V}(y)$ respectively. Following [13] we make the following assymption.

Assumption 1. For each $y>0$ the dual value function $\widetilde{V}(y)$ is finite and the minimizer $Q^*(y)\in\mathcal{M}^e$ (called the minimax martingale measure) exists.

We shall also need two complementary assumptions:

Assumption 2. For the process $Z_T(y) = y\frac{dQ^*(y)}{dP} = y\rho_T^*(y)$ let

$$\liminf_{y\to\infty} Z_T(y)/y > 0.$$

This assumption we need to ensure an existence of the inverse flow of the optimal wealth $X_t(x)$ (more exactly, to ensure the relation $\lim_{x\to-\infty} X_t(x) = -\infty$), see Theorem 1.4 from [11].

Assumption 3. The utility function U is two times differentiable and there are constants $c_1 > 0$ and $c_2 > 0$ such that

$$c_1 < -\frac{U''(x)}{U'(x)} < c_2, \quad x \in \mathbb{R}. \tag{9}$$

The last condition is similar to the condition on relative risk-aversion introduced in [5]. Note that for exponential utility function the risk-aversion coefficient $-\frac{U''(x)}{U'(x)} = \gamma$ is a constant and condition (9) is also satisfied for linear combinations of exponential utility functions with different risk-aversion parameters.

Let Π_x be the class of predictable S integrable processes π such that $U(x + (\pi \cdot S)_T) \in L^1(P)$ and $\pi\cdot S$ is a supermartingale under each $Q \in \mathcal{M}^a$ with finite $\widetilde{U}$-expectation $E\widetilde{U}(\frac{dQ}{dP})$, where the notation $\pi \cdot S$ stands for the stochastic integral.

Denote $Q(x) = Q^*(y) = Q^*(V'(x))$.

It was proved in [12] that under Assumption 1 the optimal strategy $\pi(x) \in \Pi_x$ of problem (4) exists, is unique and $V(x) = EU(X_T(x))$, where the optimal wealth $X_T(x) = x + \int_0^T \pi_u(x)\,dS_u$ is a uniformly integrable $Q(x)$-martingale.

In addition, the following duality relations hold true almost surely:

$$U'(X_T(x)) = Z_T(y), \quad y = V'(x), \tag{10}$$

$$V'\left(t, x + \int_0^t \pi_u(x)\,dS_u\right) = Z_t(y), \quad t \in [0,T], \tag{11}$$

where $y = V'(x)$ (see [13] and Proposition A3 from [11] for the dynamic version). Hereafter we shall use these results without further comments.

It is well known (see, e.g., [10]) that for any $x \in \mathbb{R}$ the process $(V(t,x), t \in [0,T])$ is a supermartingale admitting an RCLL (right-continuous with left limits) modification.

Therefore, using the Galchouk–Kunita–Watanabe (GKW) decomposition, the value function is represented as

$$V(t,x) = V(0,x) - A(t,x) + \int_0^t \psi(s,x)\,dM_s + L(t,x), \tag{12}$$

where for any $x \in \mathbb{R}$ the process $A(t,x)$ is increasing and $L(t,x)$ is a local martingale orthogonal to M.

Let us consider the following assumptions:

a) $V(t,x)$ is two-times continuously differentiable at x P- a.s. for any $t \in [0,T]$,

b) for any $x \in \mathbb{R}$ the process $V(t,x)$ is a special semimartingale with bounded variation part absolutely continuous with respect to $\langle M\rangle$, i.e.

$$A(t,x) = \int_0^t a(s,x)\,d\langle M\rangle_s,$$

for some real-valued function $a(s,x)$ which is predictable and $\langle M\rangle$-integrable for any $x\in\mathbb{R}$,

c) for any $x\in\mathbb{R}$ the process $V'(t,x)$ is a special semimartingale with the decomposition

$$V'(t,x)=V'(0,x)-\int_0^t a'(s,x)\,d\langle M\rangle_s+\int_0^t \psi'(s,x)\,dM_s+L'(t,x),$$

where V', a', ψ' and L' are partial derivatives at x of V, a, ψ and L, respectively.

We shall say that $(V(t,x), t\in[0,T])$ is a regular family of semimartingales if for V conditions a), b) and c) are satisfied.

We shall consider also the following conditions:

d) the conditional optimization problem (7) admits a solution, i.e., for any $t\in[0,T]$ and $x\in\mathbb{R}$ there exists a strategy $\pi(t,x)$ such that

$$V(t,x)=E\big(U(x+\int_t^T \pi_u(t,x)dS_u)|F_t\big), \tag{13}$$

e) for each $s\in[t,T]$ the function $(X_s(t,x)=x+\int_t^s \pi_u(t,x)dS_u,\ s\geq t)$ is continuous at (t,x) $P-$a.s. .

The aim of the paper is to study the properties of the dynamic value functions and the optimal solutions corresponding to primal and dual problems, their representations and existence of a regular solution of backward stochastic partial differential equation (BSPDE) . Although such results are interesting to derive BSPDEs, to study conditions a)-e) separately is also important as they bring information on the structure of such objects.

It was shown in [8, 9, 10] (see, e.g., Theorem 3.1 from [10]) that if the value function satisfies conditions a)-e), then it solves the BSPDE

$$\begin{aligned}V(t,x)=V(0,x)&+\frac{1}{2}\int_0^t \frac{(\varphi'(s,x)+\lambda(s)V'(s,x))^2}{V''(s,x)}\,d\langle M\rangle_s\\&+\int_0^t \varphi(s,x)\,dM_s+L(t,x),\quad V(T,x)=U(x),\end{aligned} \tag{14}$$

and optimal wealth satisfies the following SDE

$$X_t(x)=x-\int_0^t \frac{\varphi'(s,X_s(x))+\lambda(s)V'(s,X_s(x))}{V''(s,X_s(x))}dS_s.$$

One of our main goal is to study conditions on the basic objects (on the asset price model and on the objective function U) which will guarantee that the value function $V(t,x)$ is a regular family of semimartingales and conditions d) and e) are also satisfied, in order to show that the solution of equation (14) exists. This goal, for general objective functions, is reached only in case of complete markets. In Theorem 5.1 sufficient conditions on utility functions are given to ensure properties a)-e) and thus existence of a solution to the BSPDE (14) is established.

The typical example, where all conditions a)-e) are satisfied in the incomplete market, is the case of exponential utility function $U(x)=-e^{-\gamma x}$ with risk aversion parameter $\gamma\in(0,\infty)$. In this case $\widetilde{U}(y)=\frac{y}{\gamma}\big(\ln\frac{y}{\gamma}-1\big)$ and Assumption 1 is equivalent to the existence of $Q \in \mathcal{M}^e$ with finite relative entropy $EZ_T^Q \ln Z_T^Q$ (see e.g. [1]). The corresponding value function is of the form $V(t,x) = -e^{-\gamma x}V_t$, where

$$V_t = \operatorname*{ess\,inf}_{\pi\in\Pi} E(e^{-\gamma(\int_t^T \pi_u dS_u)}|\mathcal{F}_t) \tag{15}$$

is a special semimartingale and the BSPDE (14) for $V(t,x)$ is transformed into a usual backward stochastic differential equation (BSDE) for V_t

$$V_t = V_0 + \frac{1}{2}\int_0^t \frac{(\varphi_s + \lambda_s V_s)^2}{V_s} d\langle M\rangle_s + \int_0^t \varphi_s dM_s + L_t, \quad V_T = 1, \tag{16}$$

where L is a local martingale strongly orthogonal to M. It is evident that for $V(t,x) = -e^{-\gamma x}V_t$ conditions a)- c) are satisfied and Theorem 3.1 from [10] implies that solution of (16) exists. On the other hand an existence of a solution of equation (16) follows also from general theory of quadratic BSDEs, but in the theory of BSPDEs there are no results implying an existence of a solution of equations of type (14) (to our knowledge, the theory about existence of these equations covers only the quasi and semi-linear case). In Theorem 5.1 of Section 5, as mentioned above, we provide conditions in the case of complete markets when a solution of this equation exists.

The first main result of the paper is given as Theorem 2.1 and proves that under Assumptions 1-3, if the dual problem is well posed so is the dynamic primal problem. We also relate the optimal strategy of the static problem (that is, $V(0,x)$) with the one of the dynamic problem, associated to $V(t,x)$.

It was shown in [13] that if we start at time τ with the optimal wealth $X_\tau(x)$, then the optimal value in (7) is attained by $\pi(\tau,x) = \pi(0,x)I_{]\tau,T]}$, i.e.,

$$E[U(X_T(x))|F_\tau] \geq E[U(X_\tau(x) + \int_\tau^T \pi_u dS_u) \mid F_\tau], \ \pi \in \Pi_\tau,$$

which is well understood from the Bellman principle.

Under Assumptions 1-3 we show (see Theorem 2.1) that if we start at time τ with the wealth equal to arbitrary amount x, then the optimal strategy $\pi(\tau,x)$ of (7) is expressed in terms of the optimal strategy $\pi(x) = \pi(0,x)$ and the optimal wealth $X_\tau(x) = X_\tau(0,x)$ of (4) at time τ by the equality

$$\pi_t(\tau,x) = \pi_t(X_\tau^{-1}(x)), \quad t \geq \tau \quad \mu^{\langle S\rangle} - a.e.,$$

where $X_t^{-1}(x)$ is the inverse of the optimal wealth $X_t(x)$ and $\mu^{\langle S\rangle}$ is the Doleans measure of $\langle S\rangle$.

In Section 3, we establish the relation between Doob-Meyer decomposition terms of the value process $V(t,x)$ (12) with corresponding terms of the dual value process $\widetilde{V}(t,y)$. The second main result is stated as Theorem 3.1, where conditions are given when regularity of the primal value function implies the regularity of the dual value function and we derive BSPDE for $\widetilde{V}(t,y)$ from BSPDE (14). To obtain this result in addition to continuity of filtration we require an integral representation property with respect to M and an orthogonal local martingale $M^\perp$, in order to avoid stochastic line integrals in BSPDE representation of $\widetilde{V}(t,y)$.

The problem related with condition a) was studied in [5] for utility functions defined on the positive real line for value functions at time 0 and in [11] for dynamic value function

$V(t,x)$ corresponding to utility functions defined on the whole real line. In particular, it was shown in [11] that for any $t \in [0,T]$ the value function is continuously differentiable at x and the second derivative exists in probability, which is weaker than condition a). In addition, in [11] was proved that under Assumptions 1-3 for any t the optimal wealth is an increasing function of x P-a.s. and an adapted inverse of $X_t(x)$ exists. The problems related with conditions a), b) and c) we connect with an existence of the inverse flow $X_t^{-1}(x)$ of the optimal wealth. In Proposition 4.1 of Section 4, under stronger conditions we derive a stochastic differential equation for the inverse of the optimal wealth $\psi_t(x) = X_t^{-1}(x)$ and deduce from it that the finite variation part of the value process is absolutely continuous with respect to the square characteristic$\langle S\rangle$ of the asset price process. This result is the main step for obtaining properties a)-c) in Proposition 4.2 of Section 4.

Finally we formulate the main result of Section 5. In this section we assume that the market is complete.

Let

$$R_1(x) = -\frac{U''(x)}{U'(x)}, \quad R_2(x) = -\frac{U'''(x)}{U''(x)}, \quad x \in \mathbb{R}. \tag{17}$$

We shall use one of the following conditions:

r1) U is three-times differentiable, $R_1(x)$ is bounded away from zero and infinity and $R_2(x)$ is bounded and Lipschitz continuous,

r2) U is four-times differentiable and the density Z_T of the unique martingale measure is bounded.

Theorem 5.1. *Assume that the market is complete and that one of the condition* $r1)$ *or* $r2)$ *be satisfied. Then conditions* a)-e) *are fulfilled and the value function* $V(t,x)$ *satisfies* $BSPDE$ (14).

In the paper [3] a new approach was developed, where the solution of the problem (4) was reduced to the solvability of a system of Forward-Backward equations which is also a heavy task. Note that they showed that in case of complete markets this system admits a solution under conditions similar to condition r1) given above.

In the work [4] the wealth inverse process and duality relations are used to derive some type $SPDE$ and SDE for the forward dynamic utility (defined on the half real line), its derivative and Fenchel conjugate. In forward utility framework, in contrast to the classical utility theory, there is no prespecified terminal time at the end of which the utility datum is assigned. Thanks to this freedom at terminal time, it was shown in [4] that there exists a whole class of dynamic value functions satisfying regularity conditions of the present paper, which is hard to do for traditional utilities (since the value function is obliged to satisfy the boundary condition $V(T,x) = U(x)$) and needs stronger conditions on the basic objects.

2. The Relation between the Basic and Conditional Utility Maximization Problems

In this section we study basic and conditional utility maximization problems in incomplete markets for utility functions defined on the whole real line and establish relations between optimal strategies of these problems.

To this end we first give some definitions and auxiliary assertions.

We shall say that an adapted stochastic process $(X_t, t \in [\tau, T])$ is a generalized martingale (resp. supermartingale) if

1) $E(|X_t|/F_\tau) < \infty$, for any $t \in [\tau, T]$

2) $E(X_t/F_{t'}) = X_{t'}$ (resp. $\leq X_{t'}$) for any $t' \leq t$, where $t', t \in [\tau, T]$

(see the definition of generalized conditional expectations and of generalized supermartingales for discrete time in [14])

Definition 2.1. A predictable S integrable process π is in $\Pi_{x,\tau}$, if $E(U(x + \int_\tau^T \pi_u dS_u)/F_\tau)$ is finite and $((\pi \cdot S)_t, t \geq \tau)$ is a generalized supermartingale under each $Q \in \mathcal{M}^a$ with finite $\widetilde{U}$-expectation $E\widetilde{U}(\frac{dQ}{dP})$.

The proof of the following assertion follows from Theorem 4.1 and Proposition 3.1 of [11].

Proposition 2.1. *Let Assumptions* 1–3 *be satisfied. Then for any* $t \in [0, T]$ *there exists a modification of the optimal wealth process* $(X_t(x), x \in \mathbb{R})$ *(resp. of* $Z_t(y)$*) almost all paths of which are strictly increasing and absolutely continuous with respect to* dx *(resp.* dy*). Besides*

$$X'_t(x) > 0, \quad E^{Q(x)}(X'_T(x))^2 \leq C, \tag{18}$$

$$\lim_{x \to \infty} X_t(x) = \infty, \quad \lim_{x \to -\infty} X_t(x) = -\infty \tag{19}$$

P*-a.s. for any* $t \in [0, T]$ *and the adapted inverse* $X_t^{-1}(x)$ *(resp.* $Z_t^{-1}(y)$*) of the optimal wealth process exists.*

We shall need also the continuity properties of the square characteristics $\langle X(x) - X(y) \rangle$ which can be deduced from Proposition 2.1.

Lemma 2.1. *Let conditions of Proposition* 2.1 *be satisfied. Then, for any* $t \in [0, T]$ *the random field* $(\langle X(x) - X(y) \rangle_t, x, y \in \mathbb{R})$ *admits a continuous modification.*

Proof. It follows from Proposition 2.1 that $X_t(b) - X_t(a) = \int_a^b X'_t(x) dx$ and

$$\int_a^b E^{Q(x)} \langle X'(x) \rangle_T dx = \int_a^b E^{Q(x)} \big(X'_T(x)\big)^2 dx < \infty$$

and by the Fubini theorem $\int_a^b \frac{U'(X_T(x))}{V'(x)} \langle X'(x) \rangle_T dx < \infty,\ P - a.s.$ Thus by continuity of $\frac{V'(x)}{U'(X_T(x))}$ we obtain

$$\int_a^b \langle X'(x) \rangle_T dx \leq \max_{x \in [a,b]} \frac{V'(x)}{U'(X_T(x))} \int_a^b \frac{U'(X_T(x))}{V'(x)} \langle X'(x) \rangle_T dx < \infty,\ P - a.s.$$

Therefore, using the Kunita-Watanabe and Hölder's inequalities we have

$$\begin{aligned}
\langle X(b) - X(a) \rangle_t &= \int_a^b \int_a^b \langle X'(x), X'(y) \rangle_t dx dy \\
&\leq \int_a^b \int_a^b \langle X'(x) \rangle_t^{1/2} \langle X'(y) \rangle_t^{1/2} dx dy = \left(\int_a^b \langle X'(x) \rangle_t^{1/2} dx \right)^2 \\
&\leq (b - a) \int_a^b \langle X'(x) \rangle_t dx < \infty,\ P - a.s.
\end{aligned}$$

and it follows from inequality

$$\begin{aligned}&\langle X(b')-X(a')\rangle_t-\langle X(b)-X(a)\rangle_t\\&\quad\le\langle X(b')-X(b)\rangle_t^{1/2}\langle X(b')-X(a')+X(b)-X(a)\rangle_t^{1/2}\\&\quad\quad+\langle X(a')-X(a)\rangle_t^{1/2}\langle X(b')-X(a')+X(b)-X(a)\rangle_t^{1/2}\end{aligned}$$

that $\langle X(b_n)-X(a_n)\rangle_t\to\langle X(b)-X(a)\rangle_t,\ P-a.s.$ when $b_n\to b,\ a_n\to a$. Thus the stochastic field defined by

$$\langle X(x)-X(y)\rangle_t^*=\begin{cases}\lim_{r\to a,r'\to b}\langle X(r)-X(r')\rangle_t, & r,r' \text{ are rational,}\\ 0, & \text{if the limit does not exist}\end{cases}$$

is continuous and stochastically equivalent to $\langle X(x)-X(y)\rangle_t$. □

Theorem 2.1. *Let Assumptions* 1–3 *be satisfied. Then there exist the maximizer of* (7) *and the minimizer of* (8) *in the classes* $\Pi_{\tau,x}$ *and* $\mathcal{Z}_{\tau,y}$ *respectively and equalities*

$$X_T(\tau,x)=X_T(X_\tau^{-1}(x)),\ \pi_t(\tau,x)=\pi_t(X_\tau^{-1}(x)),t\ge\tau,\tag{20}$$

$$Y(\tau,y)=Z_T(Z_\tau^{-1}(y)),\ \rho_T^{Q^*}(\tau,y)=\rho_\tau^{Q^*}(y)\frac{Z_T(Z_\tau^{-1}(y))}{y}\tag{21}$$

are satisfied.

Moreover P*-a.s.*

$$\begin{aligned}V(\tau,x)&=E\Big[U\Big(x+\int_\tau^T\pi_u(X_\tau^{-1}(x))dS_u\Big)\mid F_\tau\Big],\\ \widetilde V(\tau,y)&=E\Big[\widetilde U(Z_T(Z_\tau^{-1}(y)))\mid F_\tau\Big],\end{aligned}\tag{22}$$

the duality relation

$$U'\left(x+\int_\tau^T\pi_u(X_\tau^{-1}(x))dS_u\right)=Z_T(Z_\tau^{-1}(y)),\ y=V'(\tau,x)\tag{23}$$

holds and the process

$$Z_t(Z_\tau^{-1}(y))X_t(X_\tau^{-1}(x)),\quad t\in[\tau,T],\ \textit{ where }\ y=V'(\tau,x),\tag{24}$$

is a generalized martingale.

Proof. By the optimality principle (see, e.g. [10]) $V(t,X_t(x))$ is a martingale and since $V(T,x)=U(x)$ we have that for any $x\in\mathbb{R}$

$$V(\tau,X_\tau(x))=E\big(U(X_T(x))/\mathcal{F}_\tau\big)\quad P-a.s.\tag{25}$$

Since for any τ the functions $V(\tau,x)$ and $X_\tau(x)$ are continuous for almost all $\omega\in\Omega$, the equality (25) holds P-a.s. for all $x\in\mathbb{R}$ and substituting $X_\tau^{-1}(x)$ in this equality we obtain that

$$V(\tau,x)=E\big(U(X_T(X_\tau^{-1}(x)))/F_\tau\big)\quad P-a.s.,$$

which means the maximality of $X_T(X_\tau^{-1}(x))$. Let us show that $X_T(X_\tau^{-1}(x))$ is equal to the stochastic integral

$$X_T(X_\tau^{-1}(x))=x+\int_\tau^T\pi_u(X_\tau^{-1}(x))dS_u\tag{26}$$

and that $\pi(X_\tau^{-1}(x))$ belongs to the class $\Pi_{\tau,x}$

In order to show equality (26) it is enough to show that $\int_\tau^T \pi_u(x)dS_u\big|_{x=\xi} = \int_\tau^T \pi_u(\xi)dS_u$, for $\xi = X_\tau^{-1}(x)$.

Let us consider the sequence of simple random variables $\xi_n = \sum_{k=-\infty}^{\infty} c_k 1_{A_k}$, where $A_k = (\frac{k}{n} \le \xi < \frac{k+1}{n}),\ c_k = \frac{k}{n}$. We have $\xi_n \to \xi$ uniformly and

$$\begin{aligned}\int_\tau^T \pi_u(\xi_n)dS_u &= \sum_{k=-\infty}^{\infty} \int_\tau^T \pi_u(c_k)1_{A_k}dS_u \\ &= \sum_{k=-\infty}^{\infty} 1_{A_k}\int_\tau^T \pi_u(c_k)dS_u = \int_\tau^T \pi_u(x)dS_u\big|_{x=\xi_n}.\end{aligned}$$

On the other hand

$$\int_\tau^T \pi_u(x)dS_u\big|_{x=\xi_n} - \int_\tau^T \pi_u(x)dS_u\big|_{x=\xi} = X_T(\xi_n) - X_\tau(\xi_n) - (X_T(\xi) - X_\tau(\xi)) \to 0,$$

as $n \to \infty$, since $X_t(x)$ is continuous and

$$\begin{aligned}&\int_\tau^T (\pi_u(\xi_n) - \pi_u(\xi))^2 d\langle S\rangle_u \\ &= \langle X(x) - X(y)\rangle_T - \langle X(x) - X(y)\rangle_\tau|_{x=\xi_n, y=\xi} \to 0, \quad P-a.s.\end{aligned}$$

as $n \to \infty$, by continuity of $\langle X(x) - X(y)\rangle_t$. Hence $\int_\tau^T \pi_u(\xi_n)dS_u \to \int_\tau^T \pi_u(\xi)dS_u$ in probability and $\int_\tau^T \pi_u(x)dS_u\big|_{x=\xi} = \int_\tau^T \pi_u(\xi)dS_u$ – P-a.s.

Since $E|U(X_T(x))| < \infty$ and $E^Q|X_t(x)| < \infty, t \in [0,T]$ for any $Q \in \mathcal{M}^a$ and $X_\tau^{-1}(x)$ is F_τ-measurable we have that

$$E[|U(X_T(X_\tau^{-1}(x)))| \mid F_\tau] < \infty, \quad E^Q(|X_t(X_\tau^{-1}(x))|/F_\tau) < \infty \quad P\text{-a.s.}, t \ge \tau.$$

On the other hand, since for any $t \in [0,T]$ the function $(X_t(x), x \in \mathbb{R})$ is continuous and increasing, the supermartingale inequality $E^Q(X_t(x)/F_{t'}) \le X_{t'}(x)$, $t' \le t \le T$, implies that

$$E^Q(X_t(X_\tau^{-1}(x))/F_{t'}) \le X_{t'}(X_\tau^{-1}(x)), \quad \tau \le t' \le t \le T,$$

for any $Q \in \mathcal{M}^a$, hence $\pi(\tau,x) = \pi(X_\tau^{-1}(x))$ belongs to the class $\Pi_{\tau,x}$ and the equality (22) holds. Similarly one can show the minimality of $Z_T(Z_\tau^{-1}(y))$, so conditional density of the minimax martingale measure to the problem (8) is $\frac{Z_T(Z_\tau^{-1}(y))}{y}$.

Since for any $t \in [0,T]$ the functions $V'(t,x), x \in \mathbb{R}$ and $Z_t(y), y > 0$ are continuous and the inverse of $Z_t(y)$ exists, from (11) we have that P-a.s.

$$Z_\tau^{-1}(V'(\tau,x)) = V'(X_\tau^{-1}(x)) \tag{27}$$

which together with (10) implies the conditional duality relation (23).

Note also that since $Z_t(y)X_t(x)$ is a martingale (see Theorem 1 from [13]), by continuity of $X(x)$ and $Z(y)$ the process $(Z_t(V'(X_\tau^{-1}))X_t(X_\tau^{-1}(x)), t \ge \tau)$ will be a generalized martingale and by equality (27) this is equivalent to (24). □

3. Relations between Decomposition Terms of the Value Processes of Primal and Dual Problems

In this section in addition to the continuity of the filtration F we assume that any orthogonal to M local martingale L is represented as a stochastic integral with respect to the given continuous local martingale $M^{\perp}$. Therefore, the value process $V(t,x)$ admits the decomposition

$$V(t,x) = V(0,x) - A(t,x) + \int_0^t \varphi(s,x)dM_s + \int_0^t \varphi_{\perp}(s,x)dM_s^{\perp},$$

where $A(t,x)$ is an increasing process for any $x \in \mathbb{R}$, φ and $\varphi_{\perp}$ are M and $M^{\perp}$ integrable predictable processes respectively. Since the value process $\widetilde{V}(t,y)$ of the dual problem is a submartingale for each $y>0$ it is decomposable as

$$\widetilde{V}(t,y) = \widetilde{V}(0,y) + \widetilde{A}(t,y) + \int_0^t \widetilde{\varphi}(s,y)dM_s + \int_0^t \widetilde{\varphi}_{\perp}(s,y)dM_s^{\perp}, \tag{28}$$

with M and $M^{\perp}$ integrable predictable processes $\widetilde{\varphi}$ and $\widetilde{\varphi}_{\perp}$ and an increasing process $\widetilde{A}(t,y)$.

It is known that the value processes of the primal and dual problems are related by the equality

$$V(t, -\widetilde{V}'(t,y)) = \widetilde{V}(t,y) - y\widetilde{V}'(t,y). \tag{29}$$

We are interested in how the decomposition terms A, φ and $\varphi_{\perp}$ are related to $\widetilde{A}, \widetilde{\varphi}$ and $\widetilde{\varphi}_{\perp}$, respectively.

Theorem 3.1. *Assume that the filtration F is continuous and any orthogonal to M local martingale L is represented as a stochastic integral with respect to a local martingale $M^{\perp}$. Assume that $V(t,x)$ is a regular family of semimartingales (i.e., satisfies conditions a)-c) of the introduction) and that $\widetilde{V}'(t,y)$ is a semimartingale with the decomposition*

$$\widetilde{V}'(t,y) = \widetilde{V}'(0,y) + \widetilde{B}(t,y) + \int_0^t \widetilde{\varphi}'(s,y)dM_s + \int_0^t \widetilde{\varphi}'_{\perp}(s,y)dM_s^{\perp}, \tag{30}$$

where $\widetilde{B}(t,y)$ is the process of finite variation for any y.

Then $(\widetilde{V}(t,y), y>0)$ is a regular family of semimartingales and

$$\widetilde{\varphi}(s,y) = \varphi(s, -\widetilde{V}'(s,y)), \quad \mu^{\langle M\rangle} \ \text{a.e.}, \tag{31}$$

$$\widetilde{\varphi}_{\perp}(s,y) = \varphi_{\perp}(s, -\widetilde{V}'(s,y)), \quad \mu^{\langle M\rangle} \ \text{a.e.}, \tag{32}$$

$$\widetilde{A}(t,y) = \int_0^t a(s, -\widetilde{V}'(s,y))d\langle M\rangle_s - \frac{1}{2}\int_0^t \frac{(\varphi'(s,-\widetilde{V}'(s,y)))^2}{V''(s,-\widetilde{V}'(s,y))}d\langle M\rangle_s - \frac{1}{2}\int_0^t \frac{(\varphi'_{\perp}(s,-\widetilde{V}'(s,y)))^2}{V''(s,-\widetilde{V}'(s,y))}d\langle M^{\perp}\rangle_s. \tag{33}$$

In addition, $\widetilde{V}(t,y)$ *satisfies the BSPDE*

$$\widetilde{V}(t,y) = \widetilde{V}(0,y) + \int_0^t \big(y\lambda_s\widetilde{\varphi}'(s,y) - \frac{1}{2}y^2\lambda_s^2\widetilde{V}''(s,y)\big)d\langle M\rangle_s$$
$$+\frac{1}{2}\int_0^t \frac{(\widetilde{\varphi}'_\perp(s,y))^2}{\widetilde{V}''(s,y)}d\langle M^\perp\rangle_s + \int_0^t \widetilde{\varphi}(s,y)dM_s$$
$$+\int_0^t \widetilde{\varphi}_\perp(s,y)dM_s^\perp, \quad \widetilde{V}(T,y) = \widetilde{U}(y). \tag{34}$$

Proof. Using the duality relation (29) and the Itô-Ventzel formula (see, e.g., [7] or [15]) we have

$$V(t,-\widetilde{V}'(t,y)) = V(0,-\widetilde{V}'(0,y)) + \int_0^t \varphi(s,-\widetilde{V}'(s,y))dM_s + \int_0^t \varphi_\perp(s,-\widetilde{V}'(s,y))dM_s^\perp$$
$$-\int_0^t V'(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'(s,y)dM_s - \int_0^t V'(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'_\perp(s,y)dM_s^\perp$$
$$+\int_0^t a(s,-\widetilde{V}'(s,y))d\langle M\rangle_s - \int_0^t V'(s,-\widetilde{V}'(s,y))d\widetilde{B}(s,y)$$
$$+\frac{1}{2}\int_0^t V''(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'(s,y)^2 d\langle M\rangle_s$$
$$+\frac{1}{2}\int_0^t V''(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'_\perp(s,y)^2 d\langle M^\perp\rangle_s$$
$$-\int_0^t \varphi'(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'(s,y)d\langle M\rangle_s - \int_0^t \varphi'_\perp(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'_\perp(s,y)d\langle M^\perp\rangle_s$$
$$= \widetilde{A}(t,y) + \int_0^t \widetilde{\varphi}(s,y)dM_s + \int_0^t \widetilde{\varphi}_\perp(s,y)dM_s^\perp$$
$$-y\widetilde{B}(t,y) - y\int_0^t \widetilde{\varphi}'(s,y)dM_s - y\int_0^t \widetilde{\varphi}'_\perp(s,y)dM_s^\perp. \tag{35}$$

Since $V'(s,-\widetilde{V}'(s,y)) = y$, from (35) we obtain that

$$\int_0^t \varphi(s,-\widetilde{V}'(s,y))dM_s + \int_0^t \varphi_\perp(s,-\widetilde{V}'(s,y))dM_s^\perp$$
$$+\int_0^t a(s,-\widetilde{V}'(s,y))d\langle M\rangle_s + \frac{1}{2}\int_0^t V''(s,-\widetilde{V}'(s,y))(\widetilde{\varphi}'(s,y))^2 d\langle M\rangle_s$$
$$+\frac{1}{2}\int_0^t V''(s,-\widetilde{V}'(s,y))(\widetilde{\varphi}'_\perp(s,y))^2 d\langle M^\perp\rangle_s$$
$$-\int_0^t \varphi'(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'(s,y))d\langle M\rangle_s - \int_0^t \varphi'_\perp(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'_\perp(s,y))d\langle M^\perp\rangle_s$$
$$= \widetilde{A}(t,y) + \int_0^t \widetilde{\varphi}(s,y)dM_s + \int_0^t \widetilde{\varphi}_\perp(s,y)dM_s^\perp. \tag{36}$$

Equalizing the martingale parts in (36) we obtain equalities (31) and (32). Since $\widetilde{V}(t,y)$ is two-times differentiable and

$$\widetilde{V}''(t,y) = -\frac{1}{V''(t,-\widetilde{V}'(t,y))}, \tag{37}$$

we have that $\widetilde{\varphi}(s,y)$ and $\widetilde{\varphi}_\perp(s,y)$ are also differentiable and

$$\widetilde{\varphi}'(s,y) = -\varphi'(s,-\widetilde{V}'(s,y))\widetilde{V}''(s,y) = \frac{\varphi'(s,-\widetilde{V}'(s,y))}{V''(s,-\widetilde{V}'(s,y))}, \quad \mu^{\langle M\rangle} \ \ a.e., \tag{38}$$

$$\widetilde{\varphi}'_\perp(s,y) = -\varphi'_\perp(s,-\widetilde{V}'(s,y))\widetilde{V}''(s,y) = \frac{\varphi'_\perp(s,-\widetilde{V}'(s,y))}{V''(s,-\widetilde{V}'(s,y))}, \quad \mu^{\langle M^\perp\rangle} \ \ a.e. \tag{39}$$

Therefore,

$$\varphi'(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'(s,y)) = V''(s,-\widetilde{V}'(s,y))(\widetilde{\varphi}'(s,y))^2, \quad \mu^{\langle M\rangle} \ \ a.e.,$$
$$\varphi'_\perp(s,-\widetilde{V}'(s,y))\widetilde{\varphi}'_\perp(s,y) = V''(s,-\widetilde{V}'(s,y))(\widetilde{\varphi}'_\perp(s,y))^2, \quad \mu^{\langle M^\perp\rangle} \ \ a.e.$$

and equalizing the finite variation parts in (36) we deduce that equality (33) holds.

Let us show now that $\widetilde{V}(t,y)$ satisfies the BSPDE (34). It follows from (14) that

$$a(s,x) = \frac{1}{2}\frac{(\lambda_s V'(s,x) + \varphi'(s,x))^2}{V''(s,x)},$$

$$\int_0^t a(s,-\widetilde{V}'(s,y))d\langle M\rangle_s = \frac{1}{2}\int_0^t \frac{(y\lambda_s + \varphi'(s,-\widetilde{V}'(s,y)))^2}{V''(s-,\widetilde{V}'(s,y))}d\langle M\rangle_s$$
$$= \int_0^t \big(y\lambda_s\widetilde{\varphi}'(s,y) - \frac{1}{2}y^2\lambda_s^2\widetilde{V}''(s,y)\big)d\langle M\rangle_s + \frac{1}{2}\int_0^t \frac{(\varphi'(s,-\widetilde{V}'(s,y)))^2}{V''(s,-\widetilde{V}'(s,y))}d\langle M\rangle_s,$$

which (together with (33)) implies that

$$\widetilde{A}(t,y) = \int_0^t \big(y\lambda_s\widetilde{\varphi}'(s,y) - \frac{1}{2}y^2\lambda_s^2\widetilde{V}''(s,y)\big)d\langle M\rangle_s + \frac{1}{2}\int_0^t \frac{(\widetilde{\varphi}'_\perp(s,y))^2}{\widetilde{V}''(s,y)}d\langle M^\perp\rangle_s. \tag{40}$$

Now, (28) and (40) imply that $\widetilde{V}(t,y)$ satisfies (34). □

Remark 1. It follows from (28), (33) and (37) that $\widetilde{V}(t,y)$ satisfies also the forward SPDE derived in [4], which takes in this case the following form

$$\widetilde{V}(t,y) = \widetilde{V}(0,y) + \int_0^t a(s,-\widetilde{V}'(s,y))d\langle M\rangle_s + \frac{1}{2}\int_0^t (\varphi'(s,-\widetilde{V}'(s,y))^2\widetilde{V}''(s,y)d\langle M\rangle_s$$
$$+ \frac{1}{2}\int_0^t (\varphi'_\perp(s,-\widetilde{V}'(s,y)))^2\widetilde{V}''(s,y)d\langle M^\perp\rangle_s$$
$$+ \int_0^t \varphi(s,-\widetilde{V}'(s,y))dM_s + \int_0^t \varphi_\perp(s,-\widetilde{V}'(s,y))dM^\perp_s.$$

4. Differential Equation for the Inverse Flow of the Optimal Wealth

By Proposition 2.1, if the filtration F is continuous and Assumptions 1-3 are satisfied then the adapted inverse $X_t^{-1}(x)$ of the optimal wealth process exists. Under stronger conditions we shall derive for the inverse process $X_t^{-1}(x)$ a Stochastic Differential Equation (SDE) which will be used to show the absolute continuity of bounded variation parts of $V(t,x)$ and $V'(t,x)$ with respect to square characteristic $\langle S\rangle$.

For stochastic process $\xi_t(x)$ by $\xi_t'(x)$ we denote the derivative with respect to x, $\mu^{\langle S\rangle}$ denotes Dolean's measure for $\langle S\rangle$, i.e. the measure $d\langle S\rangle dP$ on $[0,T]\times\Omega$. If $F(t,x)$ is a family of semimartigales, then $\int_0^T F(ds,\xi_s)$ denotes a generalized stochastic integral (see [7], Chapter 3), or stochastic line integral by terminology from [2]. If $F(t,x)=xG_t$, where G_t is a semimartingale, then the generalized stochastic integral coincides with the usual one denoted by $\int_0^T \xi_s dG_s$ or $(\xi\cdot G)_T$.

Now we shall derive an SDE for the inverse of the optimal wealth $\psi_t(x)=X_t^{-1}(x)$ of the form

$$d\psi_t=\sigma_t(\psi_t)dS_t+\mu_t(\psi_t)d\langle S\rangle_t,\ \psi_0=x, \tag{41}$$

where $\sigma_t(z)=-\frac{\pi_t(z)}{X_t'(z)},\ \mu_t(z)=\frac{1}{2X_t'(z)}\left(\frac{\pi_t^2(z)}{X_t'(z)}\right)'$.

Proposition 4.1. *Let $X_t''(x),\ \pi_t'(x)$ exist and be locally Lipschitz functions with respect to x $\mu^{\langle S\rangle}-$a.e.. Then SDE (41) or equivalently*

$$d\psi_t=-\frac{\pi_t(\psi_t)}{X_t'(\psi_t)}dS_t+\frac{\pi_t'(\psi_t)\pi_t(\psi_t)}{X_t'(\psi_t)^2}d\langle S\rangle_t-\frac{1}{2}\frac{X_t''(\psi_t)\pi_t^2(\psi_t)}{X_t'(\psi_t)^3}d\langle S\rangle_t, \tag{42}$$

$$\psi_0=x \tag{43}$$

admits a unique maximal solution and it coincides with $X_t^{-1}(x)$.

Proof. The SDE (41) admits unique maximal solution up to time $\tau(x)\le T$, where $|\psi_{\tau(x)-}|=\infty$ if $\tau(x)<T$ (see [7], Theorem 3.4.5). Applying the Itô–Ventzel formula for $X_t(\psi_t)\equiv X(t,\psi_t)$ (see [7], Chapter 3 or [15]) and using that ψ_t satisfies (42) we get

$$\begin{aligned}
dX(t,\psi_t)&=X(dt,\psi_t)+X'(t,\psi_t)d\psi_t+\frac{1}{2}X''(t,\psi_t)d\langle\psi\rangle_t\\
&\quad+d\left\langle\int_0^\cdot X'(dr,\psi_r(x)),\psi(x)\right\rangle_t\\
&=\pi_t(\psi_t)dS_t+X_t'(\psi_t)\Big[-\frac{\pi_t(\psi_t)}{X_t'(\psi_t)}dS_t+\frac{\pi_t'(\psi_t)\pi_t(\psi_t)}{X_t'(\psi_t)^2}d\langle S\rangle_t\\
&\quad-\frac{1}{2}\frac{X_t''(x)\pi_t^2(\psi_t)}{X_t'(\psi_t)^3}d\langle S\rangle_t\Big]+\frac{1}{2}\frac{X_t''(x)\pi_t^2(\psi_t)}{X_t'(\psi_t)^2}d\langle S\rangle_t-\frac{\pi_t'(\psi_t)\pi_t(\psi_t)}{X_t'(\psi_t)}d\langle S\rangle_t=0,\\
\psi_0(x)&=x.
\end{aligned}$$

Hence $X(t,\psi_t(x))=x$ on $[0,\tau(x))$. Since $|X_{\tau(x)}^{-1}(x)|<\infty$, we have $\tau(x)=T\ P-$a.s. and $\psi_t(x)=X_t^{-1}(x)$. □

Remark 2. Let $\pi_t(x)=H_t(X_t(x))$. Then

$$d\psi_t=-\frac{H_t(X_t(\psi_t))}{X_t'(\psi_t)}dS_t+\frac{H_t'(X_t(\psi_t))H_t(X_t(\psi_t))}{X_t'(\psi_t)^2}d\langle S\rangle_t-\frac{1}{2}\frac{X_t''(\psi_t)H_t^2(X_t(\psi_t))}{X_t'(\psi_t)^3}d\langle S\rangle_t.$$

Using equalities $X_t(\psi_t(x))=x,\ \frac{1}{X_t'(\psi_t(x))}=\psi_t'(x),\ and\ -\frac{X_t''(\psi_t(x))}{X_t'(\psi_t(x))}=\frac{\psi_t''(x)}{\psi_t'(x)^2}$ we obtain the linear partial SDE

$$d\psi_t(x)=-H_t(x)\psi_t'(x)dS_t+H_t'(x)H_t(x)\psi_t'(x)d\langle S\rangle_t+\frac{1}{2}H_t^2(x)\psi_t''(x)d\langle S\rangle_t$$

or an SPDE in the divergence form

$$d\psi_t(x)=-H_t(x)\psi_t'(x)dS_t+\frac{1}{2}(H_t^2(x)\psi_t'(x))'d\langle S\rangle_t.$$

Let us define martingale random fields

$$\mathcal{M}(t,x)=E[U(X_T(x)|F_t],$$
$$\overline{\mathcal{M}}(t,x)=E[U'(X_T(x)|F_t].$$

Proposition 4.2. *Let conditions of Proposition 4.1 be satisfied.*

i) *If $\mathcal{M}(t,x)$ is two times continuously differentiable with respect to x, then the finite variation part of $V(t,x)=\mathcal{M}(t,\psi_t(x))$ is absolutely continuous with respect to $\langle S\rangle$.*

ii) *If $\overline{\mathcal{M}}(t,x)$ is two times continuously differentiable with respect to x, then $V'(t,x)$ is a special semimartingale and the finite variation part of $V'(t,x)=\overline{\mathcal{M}}(t,\psi_t(x))$ is absolutely continuous with respect to $\langle S\rangle$. Besides $V'(t,x)$ admits the decomposition*

$$V'(t,x)=V'(0,x)-\int_0^t a'(s,x)\,d\langle M\rangle_s+\int_0^t\psi'(s,x)\,dM_s+L'(t,x). \tag{44}$$

Proof. i) By the optimality principle $V(t,X_t(x))$ is a martingale and since $V(T,x)=U(x)$ we have that $V(t,X_t(x))=E[U(X_T(x))|F_t]=\mathcal{M}(t,x)$. Therefore by duality relation (11)

$$\mathcal{M}'(t,x)=V'(t,X_t(x))X_t'(x)=Z_t(y)X_t'(x) \tag{45}$$

is a martingale and let

$$\mathcal{M}'(t,x)=V'(x)+\int_0^t h_r(x)dM_r+L_t(x),\quad L(x)\perp M$$

be the GKW decomposition of $\mathcal{M}'(t,x)$. From (42) we have

$$\left\langle\int_0^\cdot\mathcal{M}'(dr,\psi_r(x)),\psi(x)\right\rangle_t=-\int_0^t h_r(\psi_r(x))\frac{\pi_r(\psi_r(x))}{X_r'(\psi_r(x))}d\langle S\rangle_r. \tag{46}$$

Since $V(t,x)=\mathcal{M}(t,X_t^{-1}(x))$, by the Itô–Ventzel formula we get

$$V(t,x)=V(0,x)+\int_0^t\mathcal{M}(ds,\psi_s)+\int_0^t\mathcal{M}'(s,\psi_s)d\psi_s$$
$$+\frac{1}{2}\int_0^t\mathcal{M}''(s,\psi_s)d\langle\psi\rangle_s+\left\langle\int_0^\cdot\mathcal{M}'(dr,\psi_r(x)),\psi(x)\right\rangle_t. \tag{47}$$

In view of (42) and (46) one can verify that all finite variation members of (47) are integrals with respect to $\langle S\rangle$. Namely,

$$-A(t,x) = \int_0^t \mathcal{M}'(r,\psi_r(x)) \left(\frac{\pi_r'(\psi_r(x))\pi_r(\psi_r(x))}{X_r'(\psi_r(x))^2} - \frac{1}{2}\frac{X_r''(\psi_r(x))\pi_r^2(\psi_r(x))}{X_r'(\psi_r(x))^3} \right) d\langle S\rangle_r$$
$$+ \int_0^t \left(\frac{1}{2}\mathcal{M}''(r,\psi_r(x))\frac{\pi_r^2(\psi_r(x))}{X_r'(\psi_r(x))^2} - h_r(\psi_r(x))\frac{\pi_r(\psi_r(x))}{X_r'(\psi_r(x))} \right) d\langle S\rangle_r.$$

ii) It follows from (10) and (11) that

$$\overline{\mathcal{M}}(t,x) = E[U'(X_T(x))|F_t] = E[Z_T(y)|F_t] = Z_t(y) = V'(t,X_t(x)), \tag{48}$$

which (together with (45)) implies that $\mathcal{M}$ and $\overline{\mathcal{M}}$ are related as

$$\mathcal{M}'(t,x) = \overline{\mathcal{M}}(t,x)X_t'(x) \tag{49}$$

and $V'(t,x) = \overline{\mathcal{M}}(t,X_t^{-1}(x))$. It follows from (48) that $\overline{\mathcal{M}}'(t,x) = Z_t'(y)V''(x)$ is a martingale and

$$\left\langle \int_0^\cdot \overline{\mathcal{M}}'(dr,\psi_r(x)),\psi(x)\right\rangle_t = -\int_0^t \bar{h}_r(\psi_r(x))\frac{\pi_r(\psi_r(x))}{X_r'(\psi_r(x))}d\langle S\rangle_r, \tag{50}$$

where $\overline{\mathcal{M}}'(t,x) = \bar{V}''(x) + \int_0^t \bar{h}_r(x)dM_r + \bar{L}_t(x),\ \bar{L}(x)\perp M$ is the GKW decomposition of $\overline{\mathcal{M}}'(t,x)$. Therefore the Ito-Ventzel formula implies that $V'(t,x) = \overline{\mathcal{M}}(t,X_t^{-1}(x))$ is a special semimartingale and similarly to i) one can show that the finite variation part of $V'(t,x)$ is absolutely continuous with respect to $\langle S\rangle$. Therefore, $V'(t,x)$ is decomposable as

$$V'(t,x) = V'(0,x) + \int_0^t b(r,x)d\langle M\rangle_r + \int_0^t g(r,x)dM_r + N(t,x), \tag{51}$$

for some local martingale $N(t,x)$ orthogonal to M for any $x\in\mathbb{R}$ and M and $\langle M\rangle$ integrable processes g and b respectively. The Itô-Ventzel formula and conditions of this proposition also imply that $b(r,x)$ and $g(r,x)$ are continuous at x. Therefore, integrating the equation (51) with respect to dx (over a finite interval) and using the stochastic Fubini theorem (taking decomposition (12) in mind), we obtain (44). □

5. The Case of Complete Markets

In this section for the case of complete markets we provide sufficient conditions on the utility function U which guarantee existence of a solution of BSPDE (14).

Hereafter we shall assume that the market is complete, i.e.

$$dQ = Z_T dP, \quad where \quad Z_T = \mathcal{E}_T(-\lambda\cdot M),$$

is the unique martingale measure.

Lemma 5.1. *Let the market be complete and condition r1) be satisfied. Then the optimal wealth $X_T(x)$ is two-times differentiable and the derivatives $X_T'(x), X_T''(x)$ are bounded and Lipschitz continuous.*

Proof. Since $\widetilde{U}(y)$ and $U(x)$ are conjugate, $\widetilde{U}(y)$ is also three-times differentiable and

$$\widetilde{U}''(y) = -\frac{1}{U''(x)}, \quad \widetilde{U}'''(y) = -\frac{U'''(x)}{(U''(x))^3}, \quad y = U'(x). \tag{52}$$

Therefore the functions $B_1(y)$ and $B_2(y)$, where

$$B_1(y) = y\widetilde{U}''(y) = 1/R_1(x), \quad B_2(y) = y^2\widetilde{U}'''(y) = R_2(x)/R_1^2(x) \tag{53}$$

respectively, are also bounded. This implies that the second and third order derivatives of $\widetilde{U}(yZ_T)$ are bounded, hence the function $\widetilde{V}(y) = E\widetilde{U}(yZ_T)$ is three-times differentiable and

$$\widetilde{V}'''(y) = E^Q\widetilde{U}'''(yZ_T)Z_T^2.$$

Since $\widetilde{V}(y)$ and $V(x)$ are conjugate, $V(x)$ is also three-times differentiable.

The duality relation (10) takes in this case the following form

$$U'(X_T(x)) = yZ_T, \quad X_T(x) = -\widetilde{U}'(yZ_T), \quad y = V'(x). \tag{54}$$

This relation implies that the function $X_T(x)$ is two-times differentiable for all $\omega \in \Omega' = (Z_T > 0)$ with $P(\Omega') = 1$ and differentiating the first equality in (54) we have that

$$U''(X_T(x))X_T'(x) = V''(x)Z_T, \tag{55}$$

$$U'''(X_T(x))(X_T'(x))^2 + U''(X_T(x))X_T''(x) = V'''(x)Z_T. \tag{56}$$

From (54) and (55) we obtain that

$$X_T'(x) = \frac{V''(x)}{V'(x)}\,\frac{U'(X_T(x))}{U''(X_T(x))}.$$

By condition $r1)$ and Proposition 1.2 from [11] $c_1 \le -\frac{V''(x)}{V'(x)} \le c_2$. Therefore this implies that $X_T'(x)$ is bounded, in particular

$$\frac{c_1}{c_2} \le X_T'(x) \le \frac{c_2}{c_1}, \tag{57}$$

where c_1 and c_2 are constants from (9).

Comparing equations (55) and (56) we have that

$$X_T''(x) + \frac{U'''(X_T(x))}{U''(X_T(x))}(X_T'(x))^2 = \frac{V'''(x)}{V''(x)}X_T'(x). \tag{58}$$

Since $E^QX_T'(x) = 1$ and $E^QX_T''(x) = 0$, taking expectations with respect to the measure Q in equation (58) we get

$$\frac{V'''(x)}{V''(x)} = E^Q\frac{U'''(X_T(x))}{U''(X_T(x))}(X_T'(x))^2, \tag{59}$$

which together with (57) and condition $r1)$ implies that $\frac{V'''(x)}{V''(x)}$ is bounded.

Therefore, it follows from (58) that $X_T''(x)$ is also bounded, hence $X_T'(x)$ is Lipschitz continuous.

Since the product of bounded Lipschitz continuous functions are Lipschitz continuous, it follows from (59) that $\frac{V'''(x)}{V''(x)}$ is Lipschitz continuous and (58) implies that $X_T''(x)$ is also Lipschitz continuous, since all terms in (58) are bounded and Lipschitz continuous. □

Lemma 5.2. *Let the market be complete and condition $r2)$ be satisfied. Then the optimal wealth $X_T(x)$ is three-times differentiable, $X'_T(x)$ is strictly positive and the derivatives $X'_T(x), X''_T(x)$ and $X'''_T(x)$ are uniformly bounded on every compact $[a,b] \subset \mathbb{R}$.*

Proof. Since $U(x)$ and $\widetilde{U}(y)$ are conjugate, Condition $r2)$ implies that $\widetilde{U}(y)$ is also four times differentiable and the derivatives of $\widetilde{U}(yZ_T)$ are bounded for any $y \in \mathbb{R}$, hence the function $\widetilde{V}(y) = E\widetilde{U}(yZ_T)$ is four-times differentiable.

Then $V(x)$ is also four-times differentiable, since $V'(x)$ is the inverse of $-\widetilde{V}'(y)$. Therefore, the duality relation

$$X_T(x) = -\widetilde{U}'(V'(x)Z_T)$$

implies that the optimal wealth $X_T(x)$ is three-times differentiable and the derivatives $X'_T(x), X''_T(x)$ and $X'''_T(x)$ are bounded on every compact $[a,b] \in \mathbb{R}$. Therefore the derivatives $X'_T(x), X''_T(x)$ satisfy the local Lipschitz condition.

Besides,

$$X'_T(x) = -V''(x)Z_T\widetilde{U}''(V'(x)Z_T)) > 0$$

since $V''(x) < 0$ and $\widetilde{U}''(y) > 0$. □

Corollary 5.1. *The process $(X''_t(x),\ (t,x) \in [0,T] \times \mathbb{R})$ admits a continuous modification.*

Proof. Since $X''_t(x)$ is a $Q-$martingale, by the Doob inequality and the mean value theorem we get

$$E^Q \sup_{t \le T} |X''_t(x_1) - X''_t(x_2)|^2 \le c_1 E^Q |X''_T(x_1) - X''_T(x_2)|^2$$
$$\le c_1|x_1 - x_2| E^Q \sup_{\alpha \in [0,1]} |X'''_T(\alpha x_1 + (1-\alpha)x_2)|^2 \le c_2|x_1 - x_2|^2$$

for some constants c_1, c_2. By the Kolmogorov theorem the map

$$\mathbb{R} \ni x \to X''_\cdot(x) \in C[0,T]$$

admits a continuous modification, which implies the continuity of $X''_t(x)$ with respect to the variables (t,x), $P-$a.s.. □

Proposition 5.1. *Assume that the market is complete and that either condition $r1)$ or $r2)$ is satisfied.*

Then the optimal wealth $X_t(x)$, the optimal strategy $\pi_t(x)$ ($\mu^{\langle S \rangle}$-a.e.), martingale flows $\mathcal{M}(t,x)$ and $\overline{\mathcal{M}}(t,x)$ are two-times continuously differentiable at x for all t, $P-$a.s. and the coefficients of equation (42) satisfy the local Lipschitz condition.

Proof. First assume that condition r1) is satisfied. According to Lemma 5.1 the optimal wealth $X_T(x)$ is two-times differentiable and the derivatives $X'_T(x), X''_T(x)$ are bounded and Lipschitz continuous.

To show an existence of $\pi'(x)$ we use the decomposition $X'_T(x) = 1 + \int_0^T \pi_r^{(x)} dS_r$ with some predictable S-integrable integrand $\pi^{(1)}(x)$ and inequalities

$$E^Q \int_0^T \left(\pi_t^{(1)}(x+\varepsilon) - \pi_t^{(1)}(x)\right)^2 d\langle S\rangle_t = E^Q \left\langle X'(x+\varepsilon) - X'(x)\right\rangle_T$$
$$= E^Q \left(X'_T(x+\varepsilon) - X'_T(x)\right)^2 \le \varepsilon^2 E^Q \max_{0\le s\le 1} |X''_T(x+s\varepsilon)|^2$$
$$\le \varepsilon^2 Const,$$

By the Kolmogorov theorem $\pi^{(1)}(x)$ is continuous with respect to x $\mu^{\langle S\rangle}$-a.e.

Note that, if instead of r1) condition r2) is satisfied, then we shall have that there exists a $\mu^{\langle S\rangle}$-a.e. continuous modification of $\pi^{(1)}(x)$ on each compact of $\mathbb{R}$ which will imply an existence of continuous modification on the whole real line.

Thus by the stochastic Fubini theorem (see [15])

$$x_2 - x_1 + \int_0^T (\pi_r(x_2) - \pi_r(x_1)) dS_r = X_T(x_2) - X_T(x_1)$$
$$= \int_{x_1}^{x_2} X'_T(x)dx = x_2 - x_1 + \int_0^T \int_{x_1}^{x_2} \pi_r^{(1)}(x) dx dS_r$$

and consequently $\pi_r(x_2) - \pi_r(x_1) = \int_{x_1}^{x_2} \pi_r^{(1)}(x)dx$ $\mu^{\langle S\rangle}$-a.e.. Hence $\pi^{(1)}(x) = \pi'(x)$ $\mu^{\langle S\rangle}$-a.e. and

$$X'_T(x) = 1 + \int_0^T \pi'_r(x) dS_r \tag{60}$$

for all x $P-$a.s.

It follows from (60) and from the Fubini theorem that

$$X_t(x_2) - X_t(x_1) = x_2 - x_1 + \int_0^t (\pi_r(x_2) - \pi_r(x_1)) dS_r$$
$$= x_2 - x_1 + \int_0^t \int_{x_1}^{x_2} \pi'_r(x) dx dS_r = \int_{x_1}^{x_2} X'_t(x)dx$$

for any $x_2 \ge x_1$ $P-$a.s. and lemma $A3$ from [11] implies that for each fixed t there exists a modification of $(X_t(x), x \in \mathbb{R})$ which is absolutely continuous with respect to the Lebesgue measure dx. Since $(X'_t(x), t \in [0,T])$ is a Q-martingale

$$|X'_t(x_2) - X'_t(x_1)| \le E^Q(|X'_T(x_2) - X'_T(x_1)|/F_t) \le C|x_2 - x_1| \tag{61}$$

for any $x_2 \ge x_1$ $P-$a.s. and Lemma 5.1 and Corollary 5.1 imply that there exists $\Omega' \subset \Omega,\ P(\Omega') = 1$, such that at each $\omega \in \Omega'$ the inequality (61) is fulfilled for all (t, x).

Since $EX''_T(x) = 0$ and the market is complete we have $X''_T(x) = \int_0^T \pi_r^{(2)}(x) dS_r$ for some predictable S-integrable integrand $\pi^{(2)}$. Similarly as above one can show that $\pi^{(2)}(x)$ is continuous at x $\mu^{\langle S\rangle}$-a.e., $\pi^{(2)}(x) = \pi''(x)$ $\mu^{\langle S\rangle}$-a.e. and, hence $X''_t(x)$ admits the representation

$$X''_t(x) = \int_0^t \pi''_r(x) dS_r.$$

Similarly we can show that one can choose a modification of $X_t(x)$ which is two-times differentiable and such that $X''(x)$ is Lipschitz continuous.

In case when instead of r1) the condition r2) is fulfilled $X''(x)$ will satisfy the local Lipschitz condition. So, in both cases (i.e., if condition r1) or r2) is satisfied) the coefficients of equation (42) will be locally Lipschitz continuous.

Since the market is complete $\overline{\mathcal{M}}(t,x) = V'(x)Z_t$ and it is evident that $\overline{\mathcal{M}}(t,x)$ is two-times continuously differentiable. Besides, equality (49) implies that $\mathcal{M}(t,x)$ is also two-times continuously differentiable at x. □

Theorem 5.1. *Assume that the market is complete and that one of the condition* $r1)$ *or* $r2)$ *be satisfied. Then conditions* a)-e) *are fulfilled and the value function* $V(t,x)$ *satisfies* $BSPDE$ (14).

Proof. It is evident that boundedness of $B_1(y)$ and $B_2(y)$ (defined by (53)) implies that the dual value function $\widetilde{V}(t,y) = E(\widetilde{U}(y\frac{Z_T}{Z_t})/F_t)$ is two-times continuously differentiable. Since

$$V''(t,x) = -\frac{1}{\widetilde{V}''(t,y)}, \quad y = V'(x),$$

the value function $V(t,x)$ is also two-times continuously differentiable, hence condition a) is fulfilled.

It follows from Proposition 5.1 that under the presence assumptions all conditions of Propositions 4.1 and 4.2 are satisfied, therefore these propositions imply that $V(t,x)$ satisfies conditions b) and c), hence $V(t,x)$ is a regular family of semimartingales.

Let us show that the condition e) is also satisfied. By optimality principle (see [10]) for any $t \in [0,T]$ the process $(V(s, X_s(t,x)), s \geq t)$ is a martingale, where $X_s(t,x) = x + \int_t^s \pi_u(t,x)dS_u$ is the solution of the conditional optimization problem (14). This implies that P-a.s.

$$V(t,x) = E(V(s, X_s(t,x))/F_t). \tag{62}$$

On the other hand using again the optimality principle we have

$$V(t, X_t(x)) = E(V(s, X_s(x))/F_t),$$

and substituting in this equality the inverse of the optimal capital $X_t(x)$ we get

$$V(t,x) = E(V(s, X_s(X_t^{-1}(x))/F_t). \tag{63}$$

Since for any t the function $(V(t,x), x \in \mathbb{R})$ is strictly convex, comparing (62) and (63) we obtain that P-a.s $X_s(t,x) = X_s(X_t^{-1}(x))$. By continuity at (t,x) of $X_t^{-1}(x)$ as a solution of SDE (42) we obtain that condition e) is satisfied.

Thus, all conditions of Theorem 3.1 from [10] are satisfied which implies that $V(t,x)$ is a solution of the $BSPDE$ (14). □

Corollary 5.2. *Let conditions of Theorem* 5.1 *be satisfied. Then the process*

$$\widetilde{V}(t,y) = E\Big(\widetilde{U}(y\frac{Z_T}{Z_t})/F_t\Big), \quad t \in [0,T],$$

satisfies the BSPDE (34).

Proof. According to Theorem 3.1 it is sufficient to verify that the process

$$\widetilde{V}'(t,y) = E(\frac{Z_T}{Z_t}\widetilde{U}'(y\frac{Z_T}{Z_t})/F_t), \quad t \in [0,T],$$

is a special semimartingale.

Let $\overline{V}(t.y) = E(Z_T\widetilde{U}'(yZ_T)/F_t)$. It is evident that $\widetilde{V}'(t,y) = \frac{1}{Z_t}\overline{V}(t,\frac{y}{Z_t})$. But by the duality relation (10) $\overline{V}(t.y) = E(Z_T\widetilde{U}'(yZ_T)/F_t) = -Z_tX_t(x)$ and the martingale field $\overline{V}(t.y)$ is two-times differentiable by Proposition 5.1. Therefore the Itô-Ventzel formula implies that $\frac{1}{Z_t}\overline{V}(t,\frac{y}{Z_t})$ is a special semimartingale, hence so is the process $\widetilde{V}'(t,y)$. □

We would like to thank anonymous referees for useful remarks and comments.

REFERENCES

[1] F. Bellini and M. Fritelli, *On the existence of minimax martingale measures*, Math. Finance, 12 (2002), no. 1, pp. 1–21.

[2] R. Chitashvili, *Martingale ideology in the theory of controlled stochastic processes*, Probability theory and mathematical statistics. Proc. 4th USSR-Jap. Symp., Tbilisi, 1982, Lecture Notes in Math. 1021, pp. 73–92, Springer, Berlin etc., 1983.

[3] U. Horst, Ying Hu, P. Imkeller, A. Reveillac and J. Zhang, *Forward-backward systems for expected utility maximizatio*n, Stochastic Processes and their Applications, 124 (2014), no. 5, pp. 1813–1848.

[4] N. El Karoui and M. Mrad, *An Exact Connection between Two Solvable SDEs and a Nonlinear Utility Stochastic PDE*, SIAM Journal on Financial Mathematics, 4 (2013), no. 1, pp. 697–736.

[5] D. Kramkov and M. Sirbu, *On the two-times differentiability of the value functions in the problem of optimal investment in incomplete markets*, The Annals of Applied Probability, 16 (2006), no. 3, pp. 1352–1384.

[6] D. Kramkov and W. Schachermayer, *The asymptotic elasticity of utility functions and optimal investment in incomplete markets*, The Annals of Applied Probability, 9 (1999), no. 9, pp. 904–950.

[7] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge University Press, 1990.

[8] M. Mania and R. Tevzadze, *Backward Stochastic PDE and Imperfect Hedging*, International Journal of Theoretical and Applied Finance, 6 (2003), no. 7, pp. 663–692.

[9] M. Mania and R. Tevzadze, *Backward stochastic partial differential equations related to utility maximization and hedging, Martingale theory and its application*, J. Math. Sci. (N. Y.) 153 (2008), no. 3, pp. 291–380.

[10] M. Mania and R. Tevzadze, *Backward stochastic PDEs related to utility maximization problem*, Georgian Math. J., 17 (2010), no. 4, pp. 705–741.

[11] M. Mania and R. Tevzadze, *On the properties of dynamic value functions in the problem of optimal investment in incomplete markets*, Georgian Math. J., 22 (2015), no. 1.

[12] W. Schachermayer, *Optimal investment in incomplete markets when wealth may be negative*, The Annals of Applied Probability, 11 (2001), no. 3, pp. 694-734.

[13] W. Schachermayer, *A Super-Martingale Property of the Optimal Portfolio Process*, Finance and Stochastics, 7 (2003), no. 4, pp. 433–456.

[14] A. N. Shiryaev, *Essentials of stochastic finance*, World Scintific, 1999.

[15] M. Veraar, *The stochastic Fubini theorem revisited*, Stochastics, An International Journal of Probability and Stochastic Processes, 84, (2012), no. 4, pp. 543–551.

CONNECTIONS BETWEEN A SYSTEM OF FORWARD-BACKWARD SDES AND BACKWARD STOCHASTIC PDES RELATED TO THE UTILITY MAXIMIZATION PROBLEM

M. MANIA AND R. TEVZADZE

Abstract. Connections between a system of Forward-Backward SDEs derived in [4] and Backward Stochastic PDEs (from [9]) related to the utility maximization problem is established. Besides, we derive another version of Forward-Backward SDE of the same problem and prove the existence of solution.

Key words and phrases: Utility maximization problem, backward stochastic partial differential equation, forward backward stochastic differential Equation

MSC 2010: 90A09, 60H30, 90C39

1. Introduction

We consider a financial market model, where the dynamics of asset prices is described by the continuous R^d-valued continuous semimartingale S defined on a complete probability space $(\Omega, \mathcal{F}, P)$ with filtration $F = (F_t, t \in [0,T])$ satisfying the usual conditions, where $\mathcal{F} = F_T$ and $T < \infty$. We work with discounted terms, i.e. the bond is assumed to be constant.

Let $U = U(x) : R \to R$ be a utility function taking finite values at all points of real line R such that U is continuously differentiable, increasing, strictly concave and satisfies the Inada conditions

$$U'(\infty) = \lim_{x\to\infty} U'(x) = 0, \quad U'(-\infty) = \lim_{x\to-\infty} U'(x) = \infty. \tag{1.1}$$

We also assume that U satisfies the condition of reasonable asymptotic elasticity (see [5] and [14] for a detailed discussion of these conditions), i.e.,

$$\limsup_{x\to\infty} \frac{xU'(x)}{U(x)} < 1, \quad \liminf_{x\to-\infty} \frac{xU'(x)}{U(x)} > 1. \tag{1.2}$$

For the utility function U we denote by $\widetilde{U}$ its convex conjugate

$$\widetilde{U}(y) = \sup_x (U(x) - xy), \quad y > 0. \tag{1.3}$$

Denote by $\mathcal{M}^e$ (resp. $\mathcal{M}^a$) the set of probability measures Q equivalent (resp. absolutely continuous) with respect to P such that S is a local martingale under Q.

Published in *Trans. A. Razmadze Math. Inst.* **172** (2018), no. 3.

Let $\mathcal{M}_U^a$ (resp. $\mathcal{M}_U^e$) be the convex set of probability measures $Q \in \mathcal{M}^a$ (resp. $\mathcal{M}^e$) such that

$$E\widetilde{U}\Big(\frac{dQ_T}{dP_T}\Big) < \infty. \tag{1.4}$$

It follows from Proposition 4.1 of [13] that (1.4) implies $E\widetilde{U}\big(y\frac{dQ_T}{dP_T}\big) < \infty$ for any $y > 0$.

Throughout the paper we assume that

$$\mathcal{M}_U^e \neq \emptyset. \tag{1.5}$$

The wealth process, determined by a self-financing trading strategy π and initial capital x, is defined as a stochastic integral

$$X_t^{x,\pi} = x + \int_0^t \pi_u dS_u, \quad 0 \leq t \leq T.$$

We consider the utility maximization problem with random endowment H, where H is a liability that the agent must deliver at the terminal time T. H is an F_T-measurable random variable which for simplicity is assumed to be bounded (one can use also weaker assumption 1.6 from [11]). The value function $V(x)$ associated to the problem is defined by

$$V(x) = \sup_{\pi \in \Pi_x} E\bigg[U\bigg(x + \int_0^T \pi_u \, dS_u + H\bigg)\bigg], \tag{1.6}$$

where Π_x is a class of strategies which (following [14] and [11]) we define as the class of predictable S- integrable processes π such that $U(x + (\pi \cdot S)_T + H) \in L^1(P)$ and $\pi \cdot S$ is a supermartingale under each $Q \in \mathcal{M}_U^a$.

The dual problem to (1.6) is

$$\widetilde{V}(y) = \inf_{Q \in \mathcal{M}_U^e} E[\widetilde{U}(y\rho_T^Q) + y\rho_T^Q H], \quad y > 0, \tag{1.7}$$

where $\rho_t^Q = dQ_t/dP_t$ is the density process of the measure $Q \in \mathcal{M}^e$ relative to the basic measure P.

It was shown in [11] that under assumptions (2) and (5) an optimal strategy $\pi(x)$ in the class Π_x exists. There exists also an optimal martingale measure $Q(y)$ to the problem (1.7), called the minimax martingale measure and by $\rho^* = (\rho_t^*(y), t \in [0,T])$ we denote the density process of this measure relative to the measure P.

It follows also from [11] that under assumptions (2) and (5) optimal solutions $\pi^* i(x) \in \Pi_x i$ and $Q(y) \in \mathcal{M}_U^e$ are related as

$$U^{'}\bigg(x + \int_0^T \pi_u^*(x) dS_u + H\bigg) = y\rho_T^*(y), \quad P\text{-a.s.} \tag{1.8}$$

The continuity of S and the existence of an equivalent martingale measure imply that the structure condition is satisfied, i.e. S admits the decomposition

$$S_t = M_t + \int_0^t d\langle M\rangle_s \lambda_s, \quad \int_0^t \lambda_s^T \, d\langle M\rangle_s \lambda_s < \infty$$

for all t P-a.s., where M is a continuous local martingale and λ is a predictable process. The sign ^{T} here denotes the transposition.

Let us introduce the dynamic value function of problem (1.6) defined as

$$V(t,x) = \operatorname*{ess\,sup}_{\pi \in \Pi_x} E\bigg(U\bigg(x + \int_t^T \pi_u \, dS_u + H\bigg) \,\Big|\, F_t\bigg). \tag{1.9}$$

It is well known that for any $x \in R$ the process $(V(t,x), t \in [0,T])$ is a supermartingale admitting an RCLL (right-continuous with left limits) modification.

Therefore, using the Galchouk–Kunita–Watanabe (GKW) decomposition, the value function is represented as

$$V(t,x) = V(0,x) - A(t,x) + \int_0^t \psi(s,x) \, dM_s + L(t,x),$$

where for any $x \in R$ the process $A(t,x)$ is increasing and $L(t,x)$ is a local martingale orthogonal to M.

Definition 1.1. We say that $(V(t,x), t \in [0,T])$ is a regular family of semimartingales if

a) $V(t,x)$ is two-times continuously differentiable at x P- a.s. for any $t \in [0,T]$,

b) for any $x \in R$ the process $V(t,x)$ is a special semimartingale with bounded variation part absolutely continuous with respect to an increasing predictable process $(K_t, t \in [0,T])$, i.e.

$$A(t,x) = \int_0^t a(s,x) \, dK_s$$

for some real-valued function $a(s,x)$ which is predictable and K-integrable for any $x \in R$,

c) for any $x \in R$ the process $V'(t,x)$ is a special semimartingale with the decomposition

$$V'(t,x) = V'(0,x) - \int_0^t a'(s,x) \, dK_s + \int_0^t \psi'(s,x) \, dM_s + L'(t,x),$$

where a', φ' and L' are partial derivatives of a, φ and L respectively.

If $F(t,x)$ is a family of semimartigales, then $\int_0^T F(ds, \xi_s)$ denotes a generalized stochastic integral, or a stochastic line integral (see [6], or [2]). If $F(t,x) = xG_t$, where G_t is a semimartingale, then the stochastic line integral coincides with the usual stochastic integral denoted by $\int_0^T \xi_s dG_s$ or $(\xi \cdot G)_T$.

It was shown in [7, 8, 9] (see, e.g., Theorem 3.1 from [9]) that if the value function satisfies conditions a)–c), then it solves the following BSPDE

$$\begin{aligned} V(t,x) &= V(0,x) \\ &+ \frac{1}{2}\int_0^t \frac{1}{V''(s,x)} (\varphi'(s,x) + \lambda(s)V'(s,x))^T \, d\langle M\rangle_s (\varphi'(s,x) + \lambda(s)V'(s,x)) \\ &+ \int_0^t \varphi(s,x) \, dM_s + L(t,x), \quad V(T,x) = U(x), \end{aligned} \tag{1.10}$$

and optimal wealth satisfies the SDE

$$X_t(x) = x - \int_0^t \frac{\varphi'(s, X_s(x)) + \lambda(s)V'(s, X_s(x))}{V''(s, X_s(x))} dS_s. \tag{1.11}$$

This assertion is a verification theorem since conditions are required directly on the value function $V(t,x)$ and not on the basic objects (on the asset price model and on the objective function U) only. In the case of complete markets ([10]) conditions on utility functions are given to ensure properties a)-c) and thus existence of a solution to the BSPDE (1.10), (1.11) is established. Note that the BSPDE (1.10), (1.11) is of the same form for random utility functions $U(\omega, x)$, for utility functions defined on half real line and properties a)-c) are also satisfied for standard (exponential, power and logarithmic) utility functions.

In the paper [4] a new approach was developed, where a characterization of optimal strategies to the problem (1.6) in terms of a system of Forward-Backward Stochastic Differential Equations (FBSDE) in the Brownian framework was given. The key observation was an existence of a stochastic process Y with $Y_T = H$ such that $U'(X_t + Y_t)$ is a martingale. The same approach was used in [12], where these results were generalized in semimartingale setting with continuous filtration rejecting also some technical conditions imposed in [4]. The FBSDE for the pair (X, Y) (where X is the optimal wealth and Y the process mentioned above) is of the form (see, [12])

$$Y_t = Y_0 + \int_0^t \left[\lambda_s^T \frac{U'(X_s+Y_s)}{U''(X_s+Y_s)} - \frac{1}{2}\lambda_s^T \frac{U'''(X_s+Y_s)U'(X_s+Y_s)^2}{U''(X_s+Y_s)^3} + Z_s^T\right] d\langle M\rangle_s \lambda_s$$

$$- \frac{1}{2}\int_0^t \frac{U'''(X_s+Y_s)}{U''(X_s+Y_s)} d\langle N\rangle_s + \int_0^t Z_s dM_s + N_t, \quad Y_T = H; \tag{1.12}$$

$$X_t = x - \int_0^t \big(\lambda_s \frac{U'(X_s+Y_s)}{U''(X_s+Y_s)} + Z_s\big) dS_s, \tag{1.13}$$

where N is a local martingale orthogonal to M.

Note that in ([4]) and ([12]) an existence of a solution of FBSDE (1.12), (1.13) is not proved, since not all conditions of corresponding theorems are formulated in terms of basic objects. E.g., in both papers it is imposed that $E(U'(X_T^* + H))^2 < \infty$ and it is not clear if an optimal strategy satisfying this condition exists. Note that in [4] in the case of complete markets an existence of a solution of FBSDE (1.12), (1.13) is proved under certain regularity assumptions on the objective function U.

One of our goal is to derive another version of FBSDE (1.12), (1.13) and to prove the existence of a solution which will imply the existence of a solution of the system (1.12), (1.13) also.

The second goal is to establish relations between equations BSPDE (1.10), (1.11) and FBSDE (1.12), (1.13). Solutions of these equations give constructions of the optimal strategy of the same problem. BSPDE (3.6),(3.7) can be considered as a generalization of Hamiltom-Jacobi-Bellman equation to the non Markovian case and FBSDE (1.12), (1.13) is linked with the stochastic maximum principle (see [4]), although equation (1.12)–(1.13) is not obtained directly from the maximum principle. It is well known that the relation between Bellman's dynamic programming and the Pontriagin's maximum principle in optimal control is of the form $\psi_t = V'(t, X_t)$, where V is the value function, X an optimal solution and ψ is an adjoint process (see, e.g., [1], [15]). Therefore, somewhat similar relation between above mentioned equations should be expected. In particular, it is shown in Theorem 3.1, that the first components of solutions of these equations are related by the equality

$$Y_t = -\widetilde{U}'(V'(t, X_t)) - X_t.$$

In addition, conditions are given when the existence of a solution of BSPDE (3.6),(3.7) imply the existence of a solution of the system (1.12)–(1.13) and vice versa.

2. Another Version of the Forward-Backward System (1.12)–(1.13)

In this section we derive another version of the Forward-Backward system (1.12), (1.13) in which the backward component P_t is a process, such that $P_t + U'(X_t)$ is a martingale.

Theorem 2.1. *Let utility function U be three-times continuously differentiable and let the filtration F be continuous. Assume that conditions* (1.2) *and* (1.5) *are satisfied. Then there exists a quadruple (P, ψ, L, X), where P and X are continuous semimartingales, ψ is a predictable $M-$integrable process and L is a local martingale orhogonal to M, that satisfies the FBSDE*

$$X_t = x - \int_0^t \frac{\lambda_s P_s + \lambda_s U'(X_s) + \psi_s}{U''(X_s)} dS_s, \tag{2.1}$$

$$P_t = P_0 + \int_0^t \left[\lambda_s - \frac{1}{2} U'''(X_s) \frac{(\lambda_s P_s + \lambda_s U'(X_s) + \psi_s)}{U''(X_s)^2}\right]^T d\langle M\rangle_s \big(\lambda_s P_s + \lambda_s U'(X_s) + \psi_s\big)$$

$$+ \int_0^t \psi_s dM_s + L_t, \qquad P_T = U'(X_T + H) - U'(X_T). \tag{2.2}$$

In addition the optimal strategy is expressed as

$$\pi_t^* = -\frac{\lambda_t P_t + \lambda_t U'(X_t) + \psi_t}{U''(X_t)} \tag{2.3}$$

and the optimal wealth X^* coincides with X.

Proof. Define the process

$$P_t = E(U'(X_T^* + H)/F_t) - U'(X_t^*). \tag{2.4}$$

Note that the integrability of $U'(X_T^* + H)$ follows from the duality relation (1.8). It is evident that $P_T = U'(X_T^* + H) - U'(X_T^*)$.

Since U is three-times differentiable, $U'(X_t^*)$ is a continuous semimartingale and P_t admits the decomposition

$$P_t = P_0 + A_t + \int_0^t \psi_u dM_u + L_t, \tag{2.5}$$

where A is a predictable process of finite variations and L is a local martingale orthogonal to M.

Since ρ_t^* is the density of a martingale measure, it is of the form $\rho_t^* = \mathcal{E}_t(-\lambda \cdot M + R)$, $R \perp M$. Therefore, (1.8) and (2.4) imply that

$$E(U'(X_T^* + H)/F_t) = y\rho_t^* = y - \int_0^t \lambda_s y \rho_s^* dM_s + \tilde{R}_t$$

$$= y - \int_0^t \big(P_s + U'(X_s^*)\big)\lambda_s dM_s + \tilde{R}_t, \tag{2.6}$$

where $y = EU'(X_T^* + H)$ and $\tilde{R}$ is a local martingale orthogonal to M.

By definition of the process P_t, using the Itô formula for $U'(X_t^*)$ and taking decompositions (2.5), (2.6) in mind, we obtain

$$\begin{aligned}P_0 + A_t + \int_0^t \psi_s dM_s + L_t = y - \int_0^t \big(P_s + U'(X_s^*)\big)\lambda_s dM_s + \tilde{R}_t \\ -U'(x) - \int_0^t U''(X_s^*)\pi_s^{*T} d\langle M\rangle_s \lambda_s - \frac{1}{2}\int_0^t U'''(X_s^*)\pi_s^{*T} d\langle M\rangle_s \pi_s^* \\ - \int_0^t U''(X_s^*)\pi_s^* dM_s. \qquad (2.7)\end{aligned}$$

Equalizing the integrands of stochastic integrals with respect to dM we have that $\mu^{\langle M\rangle}$-a.e.

$$\pi_t^* = -\frac{\lambda_t P_t + \lambda_t U'(X_t^*) + \psi_t}{U''(X_t^*)} \qquad (2.8)$$

Equalizing the parts of finite variations in (2.7) we get

$$A_t = -\int_0^t \Big(U''(X_s^*)\lambda_s + \frac{1}{2}U'''(X_s^*)\pi_s^*\Big)^T d\langle M\rangle_s \pi_s^* \qquad (2.9)$$

and from (2.8), substituting the expression for π^* in (2.9) we obtain that

$$\begin{aligned}A_t = \int_0^t \left[\lambda_s - \frac{1}{2}U'''(X_s)\frac{\big(\lambda_s P_s + \lambda_s U'(X_s) + \psi_s\big)}{U''(X_s)^2}\right]^T d\langle M\rangle_s \\ \times\big(\lambda_s P_s + \lambda_s U'(X_s) + \psi_s\big). \qquad (2.10)\end{aligned}$$

Therefore, (2.10) and (2.5) imply that P_t satisfies equation (2.2). Integrating both parts of equality (2.8) with respect to dS and adding the initial capital we obtain equation (2.1) for the optimal wealth. □

Corollary. *Let conditions of Theorem* 2.1 *be satisfied. Then there exists a solution of FBSDE* (1.12), (1.13). *In particular, if the pair* (X, P) *is a solution of* (2.1), (2.2)*S, then the pair* (X, Y)*, where*

$$Y_t = -\tilde{U}'(P_t + U'(X_t)) - X_t,$$

satisfies the FBSDE (1.12), (1.13).

Conversely, if the pair (X, Y) *solves the FBSDE* (1.12), (1.13)*, then* $(X_t, P_t = U'(X_t + Y_t) - U'(X_t))$ *satisfies* (2.1), (2.2).

3. Relations between BSPDE (1.10)–(1.11) and FBSDE (1.12)–(1.13)

To establish relations between equations BSPDE (1.10), (1.11) and FBSDE (1.12), (1.13) we need the following

Definition 3.1 ([3])**.** The function $u(t, x)$ is called a decoupling field of the FBSDE (1.12), (1.13) if

$$u(T, x) = H, \quad a.s. \qquad (3.1)$$

and for any $x \in R, s, \tau \in R_+$ such that $0 \le s < \tau \le T$ the FBSDE

$$Y_t = u(s,x)$$
$$+\int_s^t \big(\lambda_r^T \frac{U'(X_r+Y_r)}{U''(X_r+Y_r)} - \frac{1}{2}\lambda_r^T \frac{U'''(X_r+Y_r)U'(X_r+Y_r)^2}{U''(X_r+Y_r)^3} + Z_r^T\big) d\langle M\rangle_r \lambda_r$$
$$-\frac{1}{2}\int_s^t \frac{U'''(X_r+Y_r)}{U''(X_r+Y_r)} d\langle N\rangle_r + \int_s^t Z_r dM_r + N_t - N_s, \quad Y_\tau = u(\tau, X_\tau), \tag{3.2}$$

$$X_t = x - \int_s^t \big(\lambda_r \frac{U'(X_r+Y_r)}{U''(X_r+Y_r)} + Z_r\big) dS_r, \tag{3.3}$$

has a solution (Y, Z, N, X) satisfying

$$Y_t = u(t, X_t), \quad a.s. \tag{3.4}$$

for all $t \in [s,\tau]$. We mean that all integrals are well defined.

We shall say that $u(t,x)$ is a regular decoupling field if it is a regular family of semimartingales (in the sense of Definition 1.1).

If we differentiate equation BSPDE (1.10) at x (assuming that all derivatives involved exist), we obtain the BSPDE

$$V'(t,x) = V'(0,x) + \frac{1}{2}\int_0^t \left(\frac{(\varphi'(s,x)+\lambda_s V'(s,x))^T}{V''(s,x)} d\langle M\rangle_s (\varphi'(s,x)+\lambda_s V'(s,x))\right)'$$
$$+\int_0^t \varphi'(s,x)\, dM_s + L'(t,x), \quad V'(T,x) = U'(x+H). \tag{3.5}$$

Thus, we consider the following BSPDE

$$V'(t,x) = V'(0,x) + \int_0^t \left[\frac{(V''(s,x)\lambda_s + \varphi''(s,x))^T}{V''(s,x)}\right.$$
$$\left. - \frac{1}{2}V'''(s,x)\frac{(V'(s,x)\lambda_s + \varphi'(s,x))^T}{V''(s,x)}\right] d\langle M\rangle_s (V'(s,x)\lambda_s + \varphi'(s,x))$$
$$+\int_0^t \varphi'(s,x)\, dM_s + L'(t,x), \quad V'(T,x) = U'(x+H), \tag{3.6}$$

where the optimal wealth satisfies the same SDE

$$X_t(x) = x - \int_0^t \frac{\varphi'(s, X_s(x)) + \lambda(s)V'(s, X_s(x))}{V''(s, X_s(x))} dS_s. \tag{3.7}$$

The FBSDE (1.12), (1.13) is equivalent, in some sense, to BSPDE (3.6),(3.7) and the following statement establishes a relation between these equations.

Theorem 3.1. *Let the utility function $U(x)$ be three-times continuously differentiable and let the filtration F be continuous.*

a) *If $V'(t,x)$ is a regular family of semimartingales and $(V'(t,x), \varphi'(t,x), L'(t,x), X_t)$ is a solution of BSPDE (3.6),(3.7), then the quadruple (Y_t, Z_t, N_t, X_t), where*

$$Y_t = -\widetilde{U}'(V'(t,X_t)) - X_t, \tag{3.8}$$

$$Z_t = \lambda_t \widetilde{U}'(V'(t,X_t)) + \frac{\varphi'(t,X_t) + \lambda_t V'(t,X_t)}{V''(t,X_t)}, \tag{3.9}$$

$$N_t = -\int_0^t \widetilde{U}''(V'(s,X_s))d\Big(\int_0^s L'(dr,X_r)\Big), \tag{3.10}$$

will satisfy the FBSDE (1.12), (1.13)*. Moreover, the function* $u(t,x) = -\widetilde{U}'(V'(t,x)) - x$ *will be the decoupling field of this FBSDE.*

b) *Let* $u(t,x)$ *be a regular decoupling field of FBSDE* (1.12), (1.13) *and let* $(U'(X_t + Y_t), s \le t \le T)$ *be a true martingale for every* $s \in [0,T]$*. Then* $(V'(t,x), \varphi'(t,x), L'(t,x), X)$ *will be a solution of BSPDE* (3.6), (3.7) *and following relations hold*

$$V'(t,x) = U'(x + u(t,x)), \quad \textit{hence} \quad V'(t,X_t) = U'(X_t + Y_t), \tag{3.11}$$

$$\varphi'(t,X_t) = (Z_t + \lambda_s \frac{U'(X_t+Y_t)}{U''(X_t+Y_t)})V''(t,X_t) - \lambda_t U'(X_t+Y_t), \tag{3.12}$$

$$\int_0^t L'(ds,X_s) = \int_0^t U''(X_s+Y_s)dN_s, \tag{3.13}$$

where $\int_0^t L'(ds,X_s)$ *is a stochastic line integral with respect to the family* $(L'(t,x), x \in R)$ *along the process* X*.*

Proof. a) It follows from BSPDE (3.6), (3.7) and from the Itô-Ventzel formula that $V'(t,X_t)$ is a local martingale with the decomposition

$$V'(t,X_t) = V'(0,x) - \int_0^t \lambda_s V'(s,X_s)dM_s + \int_0^t L'(ds,X_s). \tag{3.14}$$

Let $Y_t = -\widetilde{U}'(V'(t,X_t)) - X_t$. Since U is three-times differentiable (hence so is $\widetilde{U}$), Y_t will be a special semimartingale and by GKW decomposition

$$Y_t = Y_0 + A_t + \int_0^t Z_u dM_u + N_t, \tag{3.15}$$

where A is a predictable process of finite variations and N is a local martingale orthogonal to M.

The definition of the process Y, decompositions (3.14) , (3.15) and the Itô formula for $\widetilde{U}'(V'(t,X_t))$ imply that

$$\begin{aligned} & A_t + \int_0^t Z_s dM_s + N_t \\ &= \int_0^t \widetilde{U}''(V'(s,X_s))V'(s,X_s)\lambda_s dM_s - \int_0^t \widetilde{U}''(V'(s,X_s))d\Big(\int_0^s L'(dr,X_r)\Big) \\ &-\frac{1}{2}\int_0^t \widetilde{U}'''(V'(s,X_s))V'(s,X_s)^2\lambda_s^T d\langle M\rangle_s\lambda_s - \frac{1}{2}\int_0^t \widetilde{U}'''(V'(s,X_s))d\langle\int_0^\cdot L'(dr,X_r)\rangle_s \\ &+\int_0^t \frac{\lambda_s V'(s,X_s) + \varphi'(s,X_s)}{V''(s,X_s)}dM_s + \int_0^t \frac{\lambda_s^T V'(s,X_s) + \varphi'(s,X_s)^T}{V''(s,X_s)}d\langle M\rangle_s\lambda_s. \end{aligned} \tag{3.16}$$

Equalizing the integrands of stochastic integrals with respect to dM in (3.16) we have that $\mu^{\langle M\rangle}$-a.e.

$$Z_s = \frac{\lambda_s V'(s,X_s)+\varphi'(s,X_s)}{V''(s,X_s)} + \widetilde{U}''(V'(s,X_s))V'(s,X_s)\lambda_s. \tag{3.17}$$

Equalizing the orthogonal martingale parts we get P-a.s.

$$N_t = -\int_0^t \widetilde{U}''(V'(s,X_s))d\biggl(\int_0^s L'(dr,X_r)\biggr). \tag{3.18}$$

Equalizing the parts of finite variations in (3.16) we have

$$\begin{aligned}A_t = &\int_0^t \frac{\lambda_s^T V'(s,X_s)+\varphi'(s,X_s)^T}{V''(s,X_s)} d\langle M\rangle_s \lambda_s \\ &-\frac{1}{2}\int_0^t \widetilde{U}'''(V'(s,X_s))V'(s,X_s)^2\lambda_s^T d\langle M\rangle_s\lambda_s \\ &-\frac{1}{2}\int_0^t \widetilde{U}'''(V'(s,X_s))d\Bigl\langle \int_0^\cdot L'(dr,X_r)\Bigr\rangle_s \end{aligned} \tag{3.19}$$

and by equalities (3.17), (3.18) we obtain from (3.19) that

$$\begin{aligned}A_t = &\int_0^t \biggl(Z_s - \widetilde{U}''(V'(s,X_s))V'(s,X_s)\lambda_s - \frac{1}{2}\widetilde{U}'''(V'(s,X_s))V'(s,X_s)^2\lambda_s\biggr)^T d\langle M\rangle_s\lambda_s \\ &-\frac{1}{2}\int_0^t \frac{\widetilde{U}'''(V'(s,X_s))}{\widetilde{U}''(V'(s,X_s))^2} d\langle N\rangle_s. \end{aligned}\tag{3.20}$$

Therefore, using the duality relations

$$V'(t,X_t) = U'(X_t+Y_t),$$

$$\widetilde{U}''(V'(t,X_t)) = -\frac{1}{U''(X_t+Y_t)},$$

$$\widetilde{U}'''(V'(t,X_t)) = -\frac{U'''(X_t+Y_t)}{(U''(X_t+Y_t))^3},$$

we obtain from (3.20) that

$$\begin{aligned}A_t = &\int_0^t \biggl(\lambda_s\frac{U'(X_s+Y_s)}{U''(X_s+Y_s)} - \frac{1}{2}\lambda_s\frac{U'''(X_s+Y_s)U'(X_s+Y_s)^2}{U''(X_s+Y_s)^3} + Z_s\biggr)^T d\langle M\rangle_s\lambda_s \\ &-\frac{1}{2}\int_0^t \frac{U'''(X_s+Y_s)}{U''(X_s+Y_s)} d\langle N\rangle_s \end{aligned}\tag{3.21}$$

Thus, (3.15) and (3.21) imply that Y satisfies equation (1.12).

Since

$$\widetilde{U}''(V'(s,X_s))V'(s,X_s) = -\frac{1}{U''(X_s+Y_s)},$$

from (3.7) and (3.17) we obtain equation (1.13) for the optimal wealth.

The proof that the function $u(t,x) = -\widetilde{U}'(V'(t,x)) - x$ is the decoupling field of the FBSDE (1.12) is similar. One should take integrals from s to t and use the same arguments.

b) Since the quadruple $(Y^{s,x}, Z^{s,x}, N^{s,x}, X^{s,x})$ satisfies the FBSDE (3.2), (3.3), it follows from the Itô formula that for any $t \geq s$

$$U'(X_t^{s,x} + Y_t^{s,x}) = U'(x + u(s,x)) - \int_s^t \lambda_r U'(X_r^{s,x} + Y_r^{s,x}) dM_r + \int_s^t U''(X_r^{s,x} + Y_r^{s,x}) dN_r. \tag{3.22}$$

Thus $U'(X_t^{s,x} + Y_t^{s,x}), t \geq s$, is a local martingale and a true martingale by assumption. Therefore, it follows from (3.1) and (3.4) that

$$U'(X_t^{s,x} + Y_t^{s,x}) = E(U'(X_T^{s,x} + H)/F_t) = V'(t, X_t^{s,x}), \tag{3.23}$$

where the last equality is proved similarly to [13]. For $t = s$ we obtain that

$$U'(x + u(s,x)) = V'(s,x), \tag{3.24}$$

hence

$$u(t,x) = -\tilde{U}'(V'(t,x)) - x. \tag{3.25}$$

Since $U(x)$ is three-times differentiable and $u(t,x)$ is a regular decoupling field, equality (3.24) implies that $V'(t,x)$ will be a regular family of semimartingales. Therefore, using the Itô-Ventzel formula for $V'(t, X_t^{s,x})$ and equalities (3.22) , (3.23) we have

$$\begin{aligned}
&\int_s^t \left[\varphi'(r, X_r^{s,x}) - V''(r, X_r^{s.x})(\lambda_s \frac{U'(X_r^{s,x} + Y_r^{s,x})}{U''(X_r^{s,x} + Y_r^{s,x})} + Z_r^{s,x}) \right] dM_r \\
&+ \int_s^t L'(dr, X_r) + \int_s^t a'(r, X_r^{s,x}) dK_r \\
&- \int_s^t \left(\lambda_r \frac{U'(X_r^{s,x} + Y_r^{s,x})}{U''(X_r^{s,x} + Y_r^{s,x})} + Z_r^{s,x} \right)^T d\langle M \rangle_r (V''(r, X_r^{s,x}) \lambda_r + \varphi''(r, X_r^{s,x})) \\
&- \frac{1}{2} \int_s^t (V'''(r, X_r^{s,x})) \left(\lambda_r \frac{U'(X_r^{s,x} + Y_r^{s,x})}{U''(X_r^{s,x} + Y_r^{s,x})} + Z_r^{s,x} \right)^T d\langle M \rangle_r \\
&\times \left(\lambda_r \frac{U'(X_r^{s,x} + Y_r^{s,x})}{U''(X_r^{s,x} + Y_r^{s,x})} + Z_r^{s,x} \right) \\
&= - \int_s^t \lambda_r U'(X_r^{s,x} + Y_r^{s,x}) dM_r + \int_s^t U''(X_r^{s,x} + Y_r^{s,x}) dN_r.
\end{aligned} \tag{3.26}$$

Equalizing the integrands of stochastic integrals with respect to dM in (3.26) we have that μ^K-a.e.

$$Z_r^{s,x} = \frac{\lambda_r V'(r, X_r^{s,x}) + \varphi'(r, X_r^{s,x})}{V''(r, X_r^{s,x})} - \lambda_r \frac{U'(X_r^{s,x} + Y_r^{s,x})}{U''(X_r^{s,x} + Y_r^{s,x})}. \tag{3.27}$$

Equalizing the parts of finite variations in (3.26), taking (3.27) in mind, we get that for any $t > s$

$$\int_s^t a'(r, X_r^{s,x})dK_r = \int_s^t \Bigg[\frac{(V''(r, X_r^{s,x})\lambda_r + \varphi''(r, X_r^{s,x}))}{V''(r, X_r^{s,x})} \\ -\frac{1}{2}V'''(r, X_r^{s,x})\frac{(V'(r, X_r^{s,x})\lambda_r + \varphi'(r, X_r^{s,x}))}{V''(r, X_r^{s,x})^2}\Bigg]^T \\ \times d\langle M\rangle_r (V'(r, X_r^{s,x})\lambda_r + \varphi'(r, X_r^{s,x})). \tag{3.28}$$

Let $\tau_s(\varepsilon) = \inf\{t \geq s : K_t - K_s \geq \varepsilon\}$. Since $\langle M^i, M^j\rangle << \tilde{K}$ for any $1 \leq i, j \leq d$, where $\tilde{K} = \sum_{i=1}^d \langle M^i\rangle$, taking an increasing process $K + \tilde{K}$ (which we denote again by K), without loss of generality we can assume that $\langle M\rangle << K$ and denote by C_t the matrix of Radon-Nicodym derivatives $C_t = \frac{d\langle M\rangle_t}{dK_t}$. Then from (3.28)

$$\int_s^{\tau_s(\varepsilon)} \Bigg[\frac{(V''(r, X_r^{s,x})\lambda_r + \varphi''(r, X_r^{s,x}))^T C_r (V'(r, X_r^{s,x})\lambda_r + \varphi'(r, X_r^{s,x}))}{V''(r, X_r^{s,x})} \\ -\frac{1}{2}V'''(r, X_r^{s,x})\frac{(V'(r, X_r^{s,x})\lambda_r + \varphi'(r, X_r^{s,x}))^T C_r (V'(r, X_r^{s,x})\lambda_r + \varphi'(r, X_r^{s,x}))}{V''(r, X_r^{s,x})^2} \\ -a'(r, X_r^{s,x})\Bigg] dK_r = 0. \tag{3.29}$$

Since for any $x \in R$ the process $X_r^{s,x}$ is a continuous function on $\{(r, s), r \geq s\}$ with $X_s^{s,x} = x$ (as a solution of equation (3.3)) and $V'(t, x)$ is a regular family of semimartingales, dividing equality (3.29) by ε and passing to the limit as $\varepsilon \to 0$ from [7] (Proposition B1) we obtain that for each x

$$a'(s, x) = \frac{(V''(s, x)\lambda_s + \varphi''(s, x))^T C_s (V'(s, x)\lambda_s + \varphi'(s, x))}{V''(s, x)} \\ -\frac{1}{2}V'''(s, x)\frac{(V'(s, x)\lambda_s + \varphi'(s, x))^T C_s (V'(s, x)\lambda_s + \varphi'(s, x))}{V''(s, x)^2} \\ = \frac{1}{2}\bigg(\frac{(V'(s, x)\lambda_s + \varphi'(s, x))^T C_s (V'(s, x)\lambda_s + \varphi'(s, x))}{V''(s, x)}\bigg)', \ \mu^K\text{-a.e.}, \tag{3.30}$$

which implies that $V'(t, x)$ satisfies the BSPDE

$$V'(t, x) = V'(0, x) + \frac{1}{2}\int_0^t \bigg(\frac{(V'(s, x)\lambda_s + \varphi'(s, x))^T C_s (V'(s, x)\lambda_s + \varphi'(s, x))}{V''(s, x)}\bigg)' dK_s \\ + \int_0^t \varphi'(s, x)\, dM_s + L'(t, x), \quad V'(T, x) = U'(x + H). \tag{3.31}$$

The theorem is proved. □

Remark 3.1. In the proof of the part a) of the theorem we need the condition that $V'(t, x)$ is a regular family of semimartingales only to show equality (3.14) and to obtain representation (3.10). Equality (3.14) can be proved without this assumption (replacing the stochastic line integral by a local martingale orthogonal to M) from the duality relation

$$V'(t, X_t(x)) = \rho_t(y), \quad y = V'(x),$$

where $\rho_t(y)/y$ is the density of the minimax martingale measure (see [14] and [11] for the version with random endowment). Since $\rho_t(y)/y$ is representable in the form $\mathcal{E}(-\lambda\cdot M+D)$, for a local martingale D orthogonal to M, using the Dolean Dade equation we have

$$\begin{aligned} V'(t,X_t) = \rho_t &= y - \int_0^t \lambda_s\rho_s dM_s + \int_0^t \rho_s dD_s \\ &= 1 - \int_0^t \lambda_s V'(s,X_s)dM_s + R_t, \end{aligned}$$

where $R_t \equiv (Z\cdot D)_t$ is a local martingale orthogonal to M. Further the proof will be the same if we always use a local martingale R_t instead of the stochastic line integral $\int_0^t (L'(ds,X_s)$. Hence the representation (3.10) will be of the form

$$N_t = -\int_0^t \widetilde{U}''(V'(s,X_s))dR_t.$$

Remark 3.2. It follows from the proof of Theorem 3.1, that if a regular decoupling field for the FBSDE (1.12), (1.13) exists, then the second component of the solution Z is also of the form $Z_t = g(\omega,t,X_t)$ for some measurable function g and if we assume that any orthogonal to M local martingale L is represented as a stochastic integral with respect to the given continuous local martingale $M^\perp$, then the third component N of the solution will take the same form $N_t = \int_0^t g^\perp(s,X_s)dM_s^\perp$, for some measurable function $g^\perp$.

Remark 3.3. Similarly to Theorem 3.1 b) one can show that $u(t,x) = V'(t,x) - U'(x)$ is the decoupling field of (2.1), (2.2).

REFERENCES

[1] J. M. Bismut, Conjugate convex functions in optimal stochastic control. J. Math. Anal. Appl. 44 (1973), 384–404.

[2] R. Chitashvili, Martingale ideology in the theory of controlled stochastic processes. Probability theory and mathematical statistics (Tbilisi, 1982), 73–92, Lecture Notes in Math., 1021, Springer, Berlin, 1983.

[3] A. Fromm and P. Imkeller, Existence, Uniqueness and Regularity of Decoupling Fields to Multidimensional Fully coupled FBSDEs, arXiv preprint arXiv:1310.0499 (2013).

[4] U. Horst, Y. Hu, P. Imkeller, A. Reveillac and J. Zhang, Forward-backward systems for expected utility maximization. Stochastic Process. Appl. 124 (2014), no. 5, 1813–1848.

[5] D. Kramkov and W. Schachermayer, The asymptotic elasticity of utility functions and optimal investment in incomplete markets. Ann. Appl. Probab. 9 (1999), no. 3, 904–950.

[6] H. Kunita, Stochastic flows and stochastic differential equations. Cambridge Studies in Advanced Mathematics, 24. Cambridge University Press, Cambridge, 1990.

[7] M. Mania and R. Tevzadze, Backward stochastic PDE and imperfect hedging. Int. J. Theor. Appl. Finance 6 (2003), no. 7, 663–692.

[8] M. Mania and R. Tevzadze, Backward stochastic partial differential equations related to utility maximization and hedging. Martingale theory and its application. J. Math. Sci. (N.Y.) 153 (2008), no. 3, 291–380.

[9] M. Mania and R. Tevzadze, Backward stochastic PDEs related to the utility maximization problem. Georgian Math. J. 17 (2010), no. 4, 705–740.

[10] M. Mania and R. Tevzadze, On regularity of primal and dual dynamic value functions related to investment problems and their representations as backward stochastic PDE solutions. SIAM J. Financial Math. 8 (2017), no. 1, 483–503.

[11] M. P. Owen and G. Žitković, Optimal investment with an unbounded random endowment and utility-based pricing. Math. Finance 19 (2009), no. 1, 129–159.

[12] M. Santacroce and B. Trivellato, Forward backward semimartingale systems for utility maximization. SIAM J. Control Optim. 52 (2014), no. 6, 3517–3537.
[13] W. Schachermayer, Optimal investment in incomplete markets when wealth may become negative. Ann. Appl. Probab. 11 (2001), no. 3, 694–734.
[14] W. Schachermayer, A super-martingale property of the optimal portfolio process. Finance Stoch. 7 (2003), no. 4, 433–456.
[15] X. Y. Zhou, The connection between the maximum principle and dynamic programming in stochastic control. Stochastics Stochastics Rep. 31 (1990), no. 1-4, 1–13.