Filtering, Approximation and Portfolio Optimization for Shot-Noise Models and the Heston Model

Oleksandra Putyatina

Department of Mathematics
University of Kaiserslautern

Vom Fachbereich Mathematik der TU Kaiserslautern zur Verleihung des akademisches Grades *Doktor der Naturwissenschaften (Doctor rerum naturalium, Dr. rer. nat.)* genehmigte Dissertation

Referees:
Prof. Dr. J. Sass, University of Kaiserslautern
Prof. Dr. W. J. Runggaldier, Universit degli Studi di Padova, Italy

Date of defence: September 24, 2012

D 386

Acknowledgements

I would like to express gratitude, great appreciation and respect to my supervisor Prof. Dr. Jörn Sass for excellent supervision of my PhD studies, for giving valuable recommendations and useful pieces of advice.

Special thanks to my parents Dr. Galyna Putyatina and Prof. Dr. Yevgeniy Putyatin for continuous support, encouragement and advice.

I am very gratefull to my husbend Andrey Botsman for his patience.

I am grateful to my brother Dr. Vladyslav Putyatin and his family for encouraging me.

I also acknowledge DAAD for giving me the opportunity to do PhD research in financial mathematics. I would like to thank Dr. Falk Triebsch for the help.

This thesis was printed with the financial support of DAAD.

Abstract

We consider a continuous time market model in which stock returns satisfy a stochastic differential equation with stochastic drift, e.g. following an Ornstein-Uhlenbeck process. The driving noise of the stock returns consists not only of Brownian motion but also of a jump part (shot noise or compound Poisson process). The investor's objective is to maximize expected utility of terminal wealth under partial information which means that the investor only observes stock prices but does not observe the drift process. Since the drift of the stock prices is unobservable, it has to be estimated using filtering techniques. E.g., if the drift follows an Ornstein-Uhlenbeck process and without jump part, Kalman filtering can be applied and optimal strategies can be computed explicitly. Also in other cases, like for an underlying Markov chain, finite-dimensional filters exist.

But for certain jump processes (e.g. shot noise) or certain non-linear drift dynamics explicit computations, based on discrete observations, are no longer possible or existence of finite dimensional filters is no longer valid. The same computational difficulties apply to the optimal strategy since it depends on the filter. In this case the model may be approximated by a model where the filter is known and can be computed. E.g., we use statistical linearization for non-linear drift processes, finite-state-Markov chain approximations for the drift process and/or diffusion approximations for small jumps in the noise term.

In the approximating models, filters and optimal strategies can often be computed explicitly. We analyze and compare different approximation methods, in particular in view of performance of the corresponding utility maximizing strategies.

Abstract

Wir betrachten ein zeitstetiges Marktmodell, in dem Renditen der Aktien einer stochastischen Differentialgleichung mit stochastischer Drift genügen, die einem Ornstein-Uhlenbeck Prozess folgt. Die Störungen der Renditen der Aktien ergeben sich nicht nur aus einer Brownschen Bewegung, sondern auch aus Sprüngen (Shot-Noise- oder zusammengesetzter Poisson-Prozess). Das Ziel eines Investors ist, die Maximierung des erwarteten Nutzens des Endvermögens unter partieller Information. Das bedeuted, dass ein Investor nur Aktienpreise und nicht die Werte der Drift beobachtet. Da die Drift der Aktienpreise nicht beobachtet wird, muss die Drift gefiltert werden. Beispielsweise, wenn die Drift einem Ornstein-Uhlenbeck Prozess ohne Sprünge folgt, kann der Kalman-Filter angewendet werden und die optimalen Strategien können explizit berechnet werden. Endlichdimensionale Filter existieren auch in anderen Fällen, wie, zum Bespiel für Markov Ketten.

Aber für manche Sprungprozesse (z.B. Shot-Noise) oder für nichtlineare Driftdynamik sind explizite Berechnungen, die auf diskreten Beobachtungen basieren, nicht mehr möglich oder die Existenz der endlichdimensionalen Filter ist nicht mehr gegeben. Dieselben Rechenschwierigkeiten kommen auch bei den optimalen Strategien vor, weil diese von den Filtern abhängen. In diesem Fall kann das Modell von einem anderen Modell mit bekanntem Filter approximiert werden. Beispielsweise benutzen wir statistische Linearisierung für nichtlineare Driftprozesse, Approximationen mit der endlichen Markov Kette für Driftprozesse und/oder Approximationen mit der Diffusion für kleine Sprünge.

In den Approximationsmodellen können Filter und optimale Strategien oft explizit berechnet werden. Wir analysieren und vergleichen unterschiedliche Approximationsmethoden, vor allem im Hinblick auf die Güte der entsprechenden Strategien für die Maximierung des erwarteten Nutzens.

Contents

List of Figures

List of Tables

Introduction

The filtering theory was initially developed for engineering problems. Imagine that there is a physical system that performs some functions. In order to know that the system is performing correctly, an engineer has to know the state of the system at any time instant. Random noises may interfere with the system. The engineer takes some measurements (which may be disturbed by noise) of the system, and tries to determine the state of the system from noisy observations. This problem is called filtering. The filtering problem for linear physical systems can have nice explicit solutions and is often used in practice, for example, in orbit mechanics.

The problem of non-linear infinite-dimensional filtering has been studied for a long time. Stochastic differential equations with non-linear coefficients and jumps occur in many practical applications and quite often the filtering of an unobservable variable has to be carried out. The term 'filtering' roughly means finding the 'best estimate' of the unobservable variable having got the observations of another variable (observable), which depends on the unobservable variable. Assume now that both unobservable and observable variables follow stochastic differential equations. If the coefficients of these SDEs are linear, then there exists an explicit solution to the filtering problem, called Kalman filter. There are very few cases, when the filtering problem can be solved explicity. In all other cases the solution to the filtering problem has to be approximated.

In this thesis a financial market is considered for possible applications of non-linear infinite-dimensional filtering. We consider different types of stochastic processes for the risky asset: jump-diffusion processes, diffusion processes with non-linear coefficients etc. The drift of the asset price process follows a diffusion process (with linear or non-linear coefficients). The drift process cannot be observed directly, only asset prices are observable in the market. The filtering problem would be to estimate the drift values having the asset price observations. The estimated values of the drift process can further be applied for solving a portfolio optimization problem.

Therefore, portfolio optimization is considered to be the application of the filtering results. The portfolio consists of a risky asset and a bank account, where money investments are made according to a portfolio strategy. The main task is to find the optimal portfolio strategy so as to maximize the expected utility of the portfolio wealth at terminal time. The portfolio optimization framework under partial information is in detail given in [22]. The term 'partial information' means that some variables of the market model are unobservable (i.e. drift of the asset price process). If the drift of the asset price process follows linear Gaussian dynamics, the solution is given in [23], and if it is a continuous-time Markov chain, then the solution to the portfolio optimization problem is given in [33].

In Part II of this thesis we consider asset price dynamics, which is influenced by Brownian motion and a shot-noise process. The drift of the asset prices is not constant, but follows an Ornstein-Uhlenbeck process. The investor's objective is to estimate the values of the unobservable drift in order to optimize the portfolio value. The filtering of the shot-noise driven asset price process, for the process observed in discrete time, would be infinite dimensional. Therefore, one has to try to approximate the shot-noise process by Brownian motion, according to [7] and apply Kalman filtering. Portfolio optimization can then be carried out according to the framework of [6]. In order to assess the quality of this approximate portfolio optimization, the theoretical solution to portfolio optimization problem for shot-noise driven processes has to be derived.

It would also be nice to just consider a compound Poisson process instead of the shot-noise process. Like the shot-noise process, the compound Poisson process can be approximated by a Brownian motion, according to [7]. So, the aim is to perform the filtering of the unobservable drift of the asset price driven by a Brownian motion and a compound Poisson process, and to perform the portfolio optimization (under full and partial information). These tasks will also be solved in Part II of the thesis.

In Part III of the thesis the Heston's stochastic volatility model is considered. This model consists of the asset price diffusion equation, where the diffusion coefficient is a non-linear function of a stochastic volatility process. The drift coefficient also depends on the stochastic volatilty process. In order to perform filtering and portfolio optimization one has to linearize all nonlinear coefficients of the asset price model and then apply Kalman filtering to the discretized model. In order to optimize the portfolio and filter the unobservable stochastic volatility process one can also apply the extended Kalman filter.

Part I of the thesis is an introductery part, that gives an overview over the main books ([2], [4], [5]) and articles on infinite-dimensional filtering, Lévy processes and portfolio optimization ([14], [15], [28], [32], [38]). Further, we present in subsection 1.6 a Markov chain approximation for the filtering problem based on [8]. For each part, a summary is presented (Sections 5, 9, 13), and a conclusion is given in Section 14.

Part I

Some results and definitions from filtering, Lévy processes and portfolio optimization

In the introductory part of the thesis the basic framework for filtering, Lévy processes and portfolio optimization is discussed.

Assume that there is a stochastic (diffusion or jump-diffusion (Lévy)) process that cannot be observed directly, i.e. the unobservable signal process, but can be observed through another stochastic (diffusion or jump-diffusion (Lévy)) process, i.e. the observation process. The filtering problem is to evaluate the conditional expectation of the unobservable process having the observations. The filter satisfies Zakai or Kushner-Stratonovich equations. Generally, these equations are infinite-dimensional and cannot be solved explicitly. There are a few cases of the finite-dimensional filters: Kalman-Bucy filter, Beneš filter, Wonham filter. Infinite-dimensional filters have to be solved numerically.

Filtering (finite- or infinite-dimensional) can have many practical applications, but in this thesis we consider applications in the financial market. Asset prices are regarded as an observation, and the drift of the asset prices is an unobservable signal. The filter or the conditional expectation of the drift process can further be applied to solve the portfolio optimization problem.

1 Filtering theory overview

1.1 The filtering problem

Let us consider several definitions and basic results concerning general filtering theory, following the book [4].

Let (Ω, F, P) be a probability space together with a filtration $(F_t)_{t\geq 0}$ which satisfies the usual conditions:

- F is complete i.e. $A \subset B$, $B \in F$ and $P(B) = 0$ implies that $A \in F$ and $P(A) = 0$.
- The filtration F_t is right continuous i.e. $F_t = F_{t+}$.
- F_0 (and all F_t for $t \geq 0$) contains all P-null sets.

On (Ω, F, P) consider a stochastic process $X = \{X_t, t \geq 0\}$ which takes values in a complete separable metric space S (the corresponding Borel σ-algebra is denoted by $B(S)$):

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dV_s,$$

where $a(x)$ is vector in the Euclidiean r-space, $b(x)$ is an $r \times r$ matrix, V_t is the standard r-dimensional Brownian motion. The process X_t is unobservable and is called a signal.

Let W_t be a standard Brownian motion on (Ω, F, P) independent of V_t and let the observation process Y_t satisfy the following evolution equation

$$Y_t = Y_0 + \int_0^t h(X_s)ds + W_t,$$

where $h(x) : S \to \mathrm{R}^r$. Let $\{F_t^Y, t \geq 0\}$ be the usual augmentation of the filtration associated with the process Y_t:

$$F_t^Y = \sigma(Y_s, s \in [0, t]),$$

$$F^Y = \vee_{t \in R_+} F_t^Y$$

Definition 1.1. (Bain, Crisan) The filtering problem consists of determining the conditional expectation π_t of the unobservable signal X_t at time t given information from observing Y_t in the interval $[0, t]$ i.e.

$$\pi_t(\phi) = E_P[\phi(X_t)|F_t^Y].$$

One of the approaches to obtain the evolution equation for π_t is to change the measure. According to this approach the new measure is constructed, which transforms Y_t into Brownian motion and then π_t can be considered in terms of its unnormalized version ρ_t which satisfies the linear equation that is more easy to solve in some cases.

Further, some basic notions are introduced, that will be used for the derivation of equations. Let $Z = \{Z_t, t \geq 0\}$ be the process of the following form:

$$Z_t = \exp\left(-\int_0^t h(X_s)dW_s - \frac{1}{2}\int_0^t (h(X_s))^2 ds\right).$$

The process $Z_t, t \geq 0$ is a martingale if Novikov's condition is satisfied:

$$E_P\left[\exp\left(\frac{1}{2}\int_0^t (h(X_s))^2\right)\right] < \infty.$$

Since $Z_t > 0$, a new probability measure Q on F_t is introduced by specifying its Radon-Nikodym derivative with respect to P:

$$\frac{dQ}{dP} = Z_t.$$

Then, under the new measure Q the observation process Y_t is a Brownian motion independent of signal X_t. The law of signal process X_t under Q is the same as its law under P.

Let $\{Z_t^{-1}, t \geq 0\}$ be the process defined by

$$Z_t^{-1} = \exp\left(\int_0^t h(X_s)dW_s + \frac{1}{2}\int_0^t (h(X_s))^2 ds\right).$$

Then $E_Q[Z_t^{-1}] = E_P[Z_t^{-1}Z_t] = 1$, so Z_t^{-1} is an F_t-adapted martingale under Q and also

$$\frac{dP}{dQ} = Z_t^{-1}.$$

1.2 Unnormalized conditional distribution

The Kallianpur-Striebel formula is one of the central tools of the filtering theory: for every $\phi \in B(S)$ it holds

$$\pi_t(\phi) = \frac{E_Q[Z_t^{-1}\phi(X_t)|F^Y]}{E_Q[Z_t^{-1}|F^Y]}.$$

Let $\xi = \{\xi_t, t \geq 0\}$ be the process defined as follows:

$$\xi_t = E_Q[Z_t^{-1}|F_t^Y].$$

Definition 1.2. (Bain, Crisan) The unnormalized conditional distribution of X_t is the measure-valued process $\rho = \{\rho_t, t \geq 0\}$ which is determined by $\rho_t(\phi) = \pi_t(\phi)\xi_t$ for $\phi \in B(S)$. Furthermore, $\rho_t(\phi) = E_Q[Z_t^{-1}\phi(X_t)|F_t^Y]$, $Q(P)$-a.s.

In other terms,

$$\pi_t(\phi) = \frac{\rho_t(\phi)}{\rho_t(1)}.$$

Here one sees that $\rho_t(1)$ is the normalization factor.

1.3 The Zakai and Kushner-Stratonovich equations

The Zakai equation is a stochastic differential equation, whose solution is the unnormalized conditional distribution ρ_t. Unlike the Kushner-Stratonovich equation, the Zakai equation is linear and sometimes easier to solve. The Zakai equation can be presented in two different forms, both of which will be considered below.

If the following conditions

- $E_P\left[\int_0^t \|h(X_s)\|^2 ds\right] < \infty$, $E_P\left[\int_0^t Z_s\|h(X_s)\|^2 ds\right] < \infty$, for all $t > 0$
- $Q\left[\int_0^t [\rho_s(\|h\|)]^2 ds < \infty\right] = 1$, for all $t > 0$

are satisfied, then the process ρ_t satisfies the following evolution equation, called Zakai equation [4]:

$$\rho_t(\phi) = \rho_0(\phi) + \int_0^t \rho_s(A\phi)ds + \int_0^t \rho_s(\phi h)dY_s, \tag{1}$$

Q-a.s., for all $t > 0$, A is the generator assosiated with unobservable signal process X_t

$$Af(x) = a(x)\frac{\partial f(x)}{\partial x} + \frac{1}{2}(b(x))^2 \frac{\partial^2 f(x)}{\partial x^2}.$$

This is the classical form of the Zakai equation.

For the original form let us refer to the paper [37]. The normalized conditional density $\pi_t = E_P(\phi(X_t)|F_t^Y)$ can be rewritten using Kallianpur-Striebel formula:

$$E_P(\phi(X_t)|F_t^Y) = \frac{E_Q(\phi(X_t)Z_t^{-1}|F_t^Y)}{E_Q(Z_t^{-1}|F_t^Y)},$$

or rewritten using the tower law

$$E_P(\phi(X_t)|F_t^Y) = \frac{E_Q\left\{E_Q(Z_t^{-1}|F_t^{Y,X})\phi(X_t)|F_t^Y\right\}}{E_Q\left\{E_Q(Z_t^{-1}|F_t^{Y,X})|F_t^Y\right\}}.$$

Let $\tilde{Z}$ be the value of some version of $E_Q\left(Z_t^{-1}|F_t^{Y,X}\right)$ at $X_t = u$. Let $P(u,t) = Prob\{X_t \leq u\}$. Since, under Q, F_t^X and F_t^Y are independent, by Fubini theorem it follows that

$$E_Q\left\{E_Q(Z_t^{-1}|F_t^{Y,X})\phi(X_t)|F_t^Y\right\} = \int_E \phi(u)\tilde{Z}P(du,t) \tag{2}$$

and

$$E_P(\phi(X_t)|F_t^Y) = \frac{\int_E \phi(u)\tilde{Z}P(du,t)}{\int_E \tilde{Z}P(du,t)}.$$

Since $\phi(X_t)$ is arbitrary, a version of the conditional probability of X_t conditioned on Y_t with respect to P is given by:

$$Prob\{X_t \in \Gamma | F_t^Y\} = \frac{\int_\Gamma \tilde{Z}P(du,t)}{\int_E \tilde{Z}P(du,t)},$$

where Γ is any Borel set in E. Furthermore, if $P(u,t)$ is absolutely continuous and $P(du,t) = p(u,t)du$, then a version of $Prob\{X_t \in \Gamma | F_t^Y\}$ is also absolutely continuous with respect to the Lebesgue measure and its conditional density exists and satisfies:

$$p(u,t|F_t^Y) = \frac{\tilde{Z}p(u,t)}{\int_E \tilde{Z}p(u,t)du}.$$

Setting $\Phi(u,t) = \tilde{Z}p(u,t)$, define the density in the following way:

$$p(u,t|F_t^Y) = \frac{\Phi(u,t)}{\int_E \Phi(u,t)du}.$$

The connection between the unnormalized conditional distribution ρ_t and $\Phi(u,t)$, taking into account (2) and $P(du,t) = p(u,t)du$, is as in [4]:

$$\rho_t(\phi) = \int_R \phi(u)\Phi(t,u)du.$$

In this setting $\Phi(t,u)$ is the density of ρ_t.

The idea is to find a stochastic differential equation for $\Phi(u,t)$. According to the corollary in the paper [37] $\Phi(u,t)$ satisfies:

$$\Phi(u,t) = p(u,t) + \int_0^t \int_E h(z)\Phi(z,s)p_z(u,t-s)dzdY_s, \tag{3}$$

where $p(u,t)$ is the density of $Prob[X_t \in \Gamma]$ and $p_z(u,t-s)$ is the transition density of the transition probability $Prob[X_t \in \Gamma | X_s = z]$.

In order to obtain this equation in a more explicit form, we assume that the unobservable process X_t posesses a transition density $p_z(u,t)$ (the corresponding transition probability $Prob[X_t \in \Gamma | X_0 = z]$ is absolutely continuous with respect to the Lebesgue measure) for $t > 0$. Let A^* denote the forward Kolmogorov differential operator

$$A^* f(x) = -\frac{\partial}{\partial x}(a(x)f(x)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(b^2(x)f(x)).$$

Operator A^* is the Hermitian adjoint of the operator A i.e.

$$< A^*\psi, \phi >=< \psi, A\phi >$$

for suitably chosen functions ϕ, ψ.

A real valued function $f(x)$, $x \in E$ belongs to $C^{(2,\lambda)}$ if $f(x)$ and its first and second partial derivatives are bounded, continuous and satisfy on E Hölder condition with exponent $\lambda > 0$

$$f(x) - f(y) \leq C|x - y|^\lambda, \text{ for all } x, y \in E.$$

A transition density is of class A if it satisfies the following conditions as in [37]:

- If $f(x)$ is real valued, continuous and bounded on E and $0 \leq s \leq t$ then
$$U_t f(x) = U_s f(x) + \int_s^t A^* U_\theta f(u) d\theta,$$
where U_t denotes the operator
$$U_t f(x) = \int_E f(z) p_z(u, t) dz$$
for $t > 0$ and $U_0 f(u) = f(u)$.

- If $f(u, \theta)$, $u \in E$, $\theta \in [\theta_1, \theta_2]$ and its first and second partial derivatives with respect to the u are bounded and continuous on $E \times [\theta_1, \theta_2]$, and $f(u, \theta)$ is $C^{(2,\lambda)}$ in u, uniformly in $[\theta_1, \theta_2]$, then $U_t f(u, \theta)$ and its first and second partial derivatives with respect to the u are continuous on $(0, T) \times E \times [\theta_1, \theta_2]$ and bounded on $[0, T] \times G \times [\theta_1, \theta_2]$, where G is any bounded subset of E.

The following theorem is the basic theorem of [37].

Theorem 1.3. *(Zakai) Assume that unobservable process X_t posesses a transition density which is of class A. Assume that $\Phi(u, t)$ satisfies equation (3) and a.s. $\Phi(u, t)$ and $h(u)\Phi(u, t)$ together with their first and second derivatives with respect to u are bounded and continuous in $E \times [t_1, t_2]$ and $\Phi(u, t)$, $h(u)\Phi(u, t)$ are $C^{(2,\lambda)}$ in u, uniformly in $[t_1, t_2]$, then $\Phi(u, t)$ also satisfies the evolution-type equation*

$$\Phi(u, t) = \Phi(u, s) + \int_s^t A^* \Phi(u, \theta) d\theta + \int_s^t h(u) \Phi(u, \theta) dY_\theta, \tag{4}$$

$t_1 \leq s \leq t \leq t_2$.

Equation (4) is also called Zakai equation.

If the same conditions as for the Zakai equation (1) are satisfied, then the conditional distribution π_t of the signal X_t satisfies the following Kushner-Stratonovich evolution equation [4]:

$$\pi_t(\phi) = \pi_0(\phi) + \int_0^t \pi_s(A\phi)ds + \int_0^t (\pi_s(\phi h) - \pi_s(h)\pi_s(\phi))(dY_s - \pi_s(h)ds),$$

for any $\phi \in D(A)$.

1.4 Finite-Dimensional Filters

Generally filters are infinite-dimensional. The heuristical definition of the infinite-dimensional filter states that the distribution of such a filter cannot be described by a finite number of parameters. Therefore, the corresponding Zakai equation would be impossible to solve explicitly. Quite few finite-dimensional filters are known, basically: Kalman-Bucy filter (coefficients of the signal and observation equations are linear functions), Beneš filter (coefficients of the signal and observation equations satisfy specific conditions) and Wonham filter (signal process is a Markov chain).

1.4.1 Kalman-Bucy filter

One of the most often used finite-dimensional filters is the Kalman-Bucy filter [4].

Let X_t be the solution of the linear SDE driven by Brownian motion process V_t:

$$X_t = X_0 + \int_0^t (A_s X_s + a_s)ds + \int_0^t b_s dV_s,$$

where A_s, b_s and a_s are measurable and locally bounded functions. Assume that $X_0 \sim N(x_0, r_0)$ is independent of V_t. Assume that W_t is a standard F_t-adapted m-dimensional Brownian motion on (Ω, F, P) independent of X_t and let Y_t be the process satisfying the following evolution equation

$$Y_t = Y_0 + \int_0^t (H_s X_s + h_s)ds + W_t.$$

In the case when the coefficients of the signal and observation processes are linear functions, then the Kushner-Stratonovich equation takes a simpler form as given in the proposition below.

Proposition 1.4. *(Bain, Crisan) Let $\hat{X} = \{\hat{X}, t \geq 0\}$ be the conditional mean of the signal, i.e. $\hat{X} = E[X_t|F_t^Y]$. Define also $R = \{R_t, t \geq 0\}$ to be the conditional covariance matrix of the signal, i.e.*

$$R_t = E[X_t^2|F_t^Y] - E[X_t|F_t^Y]E[X_t|F_t^Y].$$

Then $\hat{X}$ satisfies the stochastic differential equation

$$d\hat{X}_t = (A_t\hat{X}_t + a_t)dt + R_tH_t(dY_t - (H_t\hat{X}_t + h_t)dt),$$

and R satisfies the deterministic Riccati equation

$$\frac{dR_t}{dt} = b_t^2 + 2F_tR_t - R_t^2H_t^2.$$

If the filter is linear, then the normalized conditional distribution π_t is normally distributed with mean value $\hat{X}_t$ and covariance R_t.

The solution to the Kalman-Bucy filtering problem has an explicit form.

Proposition 1.5. *(Bain, Crisan) The conditional expectation of X_t given the observation σ-algebra is given by the following formula:*

$$\pi_t(\phi) = \frac{1}{(2\pi)^{n/2}} \int_R \phi(\hat{X} + R_t^{1/2}z) \exp(-\frac{1}{2}\|z\|^2)dz,$$

for any $\phi \in B(R)$.

The Kalman-Bucy filter can be solved numerically in the discrete form.

Discrete Kalman filter. This framework is in detail described in [5]. The discrete version of the filtering problem is as follows:

$$X_{k+1} = A_kX_k + B_kW_k^{(1)} = A_kX_k + w_k^{(1)}, \tag{5}$$

$$Y_k = C_kX_k + D_kW_k^{(2)} = C_kX_k + w_k^{(2)}, \tag{6}$$

where the parameters A_k, B_k, C_k, D_k of the system are known. The noise processes $W_k^{(1)}$ and $W_k^{(2)}$ are normally distributed with the following properties:

$$E[W_k^1] = 0; E[(W_k^{(1)}])^2] = Q_k,$$

$$E[W_k^{(2)}] = 0; E[(W_k^{(2)}])^2] = R_k,$$

$$E[W_k^{(1)}W_k^{(2)}] = 0.$$

The processes $w_k^{(1)}$, $w_k^{(2)}$ have the following parameters:

$$E[w_k^{(1)}] = 0; E[(w_k^{(1)}])^2] = B_k^2 Q_k,$$

$$E[w_k^{(2)}] = 0; E[(w_k^{(2)}])^2] = D_k^2 R_k,$$

$$E[w_k^{(1)} w_k^{(2)}] = 0.$$

It is necessary to construct a linear unbiased estimate $\hat{X}_k$ of X_k having the observations $Y_1, Y_2, ..., Y_2$. The estimation error is denoted by $\tilde{X}_k = X_k - \hat{X}_k$. The variance of the error is denoted by $\tilde{P}_k$ and defined as follows:

$$\tilde{P}_k = E[(\tilde{X}_k)^2].$$

The discrete Kalman filter can be evaluated by iterating the following equations:

- filter equations;

$$X_k^* = A_{k-1}\hat{X}_{k-1}, \tag{7}$$

$$\hat{X}_k = X_k^* + K_k\{Y_k - C_k X_k^*\}; \tag{8}$$

- variance of the estimation error and the coefficient K;

$$P_k^* = A_{k-1}^2 \tilde{P}_{k-1} + B_{k-1}^2 Q_{k-1}, \tag{9}$$

$$K_k = P_k^* C_k \{C_k^2 P_k^* + D_k^2 R_k\}^{-1}, \tag{10}$$

$$\tilde{P}_k = P_k^* - K_k C_k P_k^*. \tag{11}$$

Note that $P_{k+1}^* = E[(A_k \tilde{X}_k + w_k^{(1)})^2]$.

Discrete Kalman filter in the case, when the driving noise of the processes is given by the sum of Wiener processes. The model to be considered looks as follows:

$$X_{k+1} = A_k X_k + \sum_{i=1}^{m} B_k^{(i)} W_k^{(i)} = A_k X_k + \sum_{i=1}^{m} w_k^{(i)}, \tag{12}$$

$$Y_k = C_k X_k + \sum_{i=1}^{m} D_k^{(i)} V_k^{(i)} = C_k X_k + \sum_{i=1}^{m} v_k^{(i)}, \tag{13}$$

where the parameters $A_k, B_k^{(i)}, C_k, D_k^{(i)}, i = 1, ..., m$ of the system are known. The noise processes $W_k^{(i)}$ and $V_k^{(i)}$ are normally distributed with the following parameters:

$$E[W_k^{(i)}] = 0; E[(W_k^{(i)}])^2] = Q_k^{(i)},$$

$$E[V_k^{(i)}] = 0; E[(V_k^{(i)}])^2] = R_k^{(i)},$$
$$E[W_k^{(i)} V_k^{(j)}] = 0.$$

The processes $w_k^{(i)}$, $v_k^{(i)}$ have the following parameters:

$$E[w_k^{(i)}] = 0; E[(w_k^{(i)}])^2] = (B_k^{(i)})^2 Q_k^{(i)},$$

$$E[v_k^{(i)}] = 0; E[(v_k^{(i)}])^2] = (D_k^{(i)})^2 R_k^{(i)},$$

$$E[w_k^{(i)} v_k^{(i)}] = 0, \ for \ i, j = 1, ..., m.$$

The discrete Kalman filter in this case consists of the following equations:

- filter equations;

$$X_k^* = A_{k-1} \hat{X}_{k-1}, \tag{14}$$

$$\hat{X}_k = X_k^* + K_k \{Y_k - C_k X_k^*\}; \tag{15}$$

- variance of the estimation error and the coefficient K;

$$P_k^* = A_{k-1}^2 \tilde{P}_{k-1} + \sum_{i=1}^{m} (B_{k-1}^{(2)})^2 Q_{k-1}^{(i)}, \tag{16}$$

$$K_k = P_k^* C_k \{C_k^2 P_k^* + \sum_{i=1}^{m} D_i^2(k) R_i(k)\}^{-1}, \tag{17}$$

$$\tilde{P}_k = P_k^* - K_k C_k P_k^*. \tag{18}$$

Note that $P_{k+1}^* = E[(A_k \tilde{X}_k + \sum_{i=1}^m w_k^{(i)})^2]$.

1.4.2 Beneš filter

The second well known finite-dimensional filter is Beneš filter [4].

Assume that unobservable signal process X_t satisfies the following stochastic differential equation:

$$X_t = X_0 + \int_0^t a(X_s) ds + \sigma V_t,$$

where V_t is Brownian motion and function a:R→ R satisfies

$$|a(x) - a(y)| \leq K|x - y|,$$

for all x, y and some constant K.

This condition ensures that the SDE for signal process has a unique solution. It is also assumed that W_t is a standard Brownian motion independent of V_t and the observation process Y_t satisfies the following equation:

$$Y_t = \int_0^t h(X_s)ds + W_t.$$

Also assume that function h: R→ R is linear, i.e. $h(x) = h_1 x + h_2$, for $x \in$R and $h_1, h_2 \in$R. Function $a(x)$ is not linear. In order to apply Beneš filter, the Beneš condition should be satisfied, i.e.

$$a'(x) + a^2(x)\sigma^{-2} + h^2(x) = P(x),$$

where $x \in$R and a' is the derivative of a and $P(x)$ is a second order polynomial with positive leading order coefficient.

Proposition 1.6. *(Bain, Crisan) If the Beneš condition is satisfied, then for arbitrary bounded Borel-measurable ϕ it follows that the conditional expectation $\pi_t(\phi)$ satisfies the following explicit formula:*

$$\pi_t = \frac{1}{c_t}\int_{-\infty}^{\infty} \phi(z)\exp(F(z)\sigma^{-2} + Q_t(z))dz,$$

where $Q_t(z)$ is the second order polynomial

$$Q_t(z) = z\Big(h_1\sigma \int_0^t \frac{\sinh(sp\sigma)}{\sinh(tp\sigma)}dY_s + \frac{q + p^2 x_0}{p\sigma\sinh(tp\sigma)} - \frac{q}{p\sigma}\coth(tp\sigma)\Big) - \frac{p\coth(tp\sigma)}{2\sigma}z^2,$$

and c_t is the corresponding normalizing constant

$$c_t = \int_{\infty}^{\infty} \exp(A(z)\sigma^2 + Q_t(z))dz,$$

where A is antiderivative of a and p, q are the coefficients of another form for polynomial $P(x) = p^2x^2 + 2qx + r$, where p, q, $r \in R$ are arbitrary.

1.4.3 Wonham filter

Another well-known finite-dimensional filter is Wonham filter [36].

Let X_t be discrete, real-valued random variable with range of values $a_1, ..., a_k$ and a priori probability distribution $\{p_0^{(j)}, j = 1, ..., K\}$ at $t = 0$. Suppose that one observes process Y_t, which is the solution of the following stochastic differential equation:

$$dY_t = X_t dt + b_t dW_t,\ Y_0 = 0,$$

where function b_t is continuously differentiable and bounded away from 0 for $t \geq 0$; W_t is a Brownian motion which is independent of X_t with $P[W_0 = 0] = 1$.

Posterior distribution is introduced:

$$p_t^{(j)} = P\{X_t = a_j | Y_s, 0 \leq s \leq t\},\ j = 1, ..., K.$$

The explicit form of $p_t^{(j)}$ is given by the following formula [36]:

$$p_t^{(j)} = \frac{p_0^{(j)} \exp\left[a_j \int_0^t b_s^{-2} dY_s - \frac{1}{2} a_j \int_0^t b_s^{-2} ds\right]}{\sum_{k=1}^K p_0^{(k)} \exp\left[a_k \int_0^t b_s^{-2} dY_s - \frac{1}{2} a_k \int_0^t b_s^{-2} ds\right]}$$

$$= \frac{p_0^{(j)} \exp\left[a_j \left(\int_0^\tau b_s^{-2} dY_s + \int_\tau^t b_s^{-2} dY_s\right) - \frac{1}{2} a_j \left(\int_0^\tau b_s^{-2} ds + \int_\tau^t b_s^{-2} ds\right)\right]}{\sum_{k=1}^K p_0^{(k)} \exp\left[a_k \left(\int_0^\tau b_s^{-2} dY_s + \int_\tau^t b_s^{-2} dY_s\right) - \frac{1}{2} a_k \left(\int_0^\tau b_s^{-2} ds + \int_\tau^t b_s^{-2} ds\right)\right]}$$

$$= \frac{p_0^{(j)} \exp\left[a_j \int_0^\tau b_s^{-2} dY_s - \frac{1}{2} a_j \int_0^\tau b_s^{-2} ds\right] \exp\left[a_j \int_\tau^t b_s^{-2} dY_s - \frac{1}{2} a_j \int_\tau^t b_s^{-2} ds\right]}{\sum_{k=1}^K p_0^{(k)} \exp\left[a_k \int_0^\tau b_s^{-2} dY_s - \frac{1}{2} a_k \int_0^\tau b_s^{-2} ds\right] \exp\left[a_k \int_\tau^t b_s^{-2} dY_s - \frac{1}{2} a_k \int_\tau^t b_s^{-2} ds\right]}$$

$$= \frac{p_\tau^{(j)} \exp\left[a_j \int_\tau^t b_s^{-2} dY_s - \frac{1}{2} a_j \int_\tau^t b_s^{-2} ds\right]}{\sum_{k=1}^K p_\tau^{(k)} \exp\left[a_k \int_\tau^t b_s^{-2} dY_s - \frac{1}{2} a_k \int_\tau^t b_s^{-2} ds\right]}. \tag{19}$$

Consider joint process $\{X_t, p_t, t \geq 0\}$, where $p_t = [p_t^{(1)}, ..., p_t^{(K)}]$ and X_t is regarded as fixed random variable with a priori probability distribution $\{p_0^{(j)}\}$. From (19) one sees that p_t depends only on X_τ, p_τ and the increments of W_s for $\tau < s < t$. These increments are independent of p_τ, X_τ and the W_s increments for $0 < s < \tau$, on which p_τ depends. It follows that the conditional distribution p_t of X_t, given X_s, p_s, $0 < s < \tau$, is a function of X_τ, p_τ alone. Therefore, the process $\{X_t, p_t, t \geq 0\}$ is Markov.

In the paper [36] the evolution in time of the $p^{(j)}$'s is described by means of a system of SDE's. Applying the result of Dynkin the author states the following evolution equation:

$$dp_t^{(j)} = m_j(t, X, p_t) + \sigma_j(t, X, p_t) dW_t,\ j = 1, ..., K.$$

The functions m_j and σ_j have probabilistic meaning

$$m_j(t, \xi, p) = \lim_{h \to 0} E\Big[\frac{p_{t+h}^{(j)} - p_t^{(j)}}{h} \Big| X_t = \xi,\ p_t = p\Big] = b_t^{-2}(X_t - \bar{X})(a_j - \bar{X}) p^{(j)},$$

$$\sigma_i(t,\xi,p)\sigma_j(t,\xi,p) = \lim_{h\to 0} E\Big[\frac{[p^{(i)}_{t+h} - p^{(i)}_t][p^{(j)}_{t+h} - p^{(j)}_t]}{h}\Big| X_t = \xi,\ p_t = p\Big],$$
$$i,j = 1,...,K,$$

where $\bar{X} = \sum_{k=1}^K a_k p^k$ and

$$\sigma_j(t,x,p) = b_t^{-1}(a_j - \bar{X})p^{(j)}.$$

Therefore, the evolution equation for $p^{(j)}$'s takes the following form:

$$\begin{aligned} dp_t^{(j)} &= b_t^{-2}(X_t - \bar{X}_t)(a_j - \bar{X}_t)p_t^{(j)}dt + b_t^{-1}(a_j - \bar{X}_t)p^{(j)}dW_t \\ &= -b_t^{-2}(x_t - \bar{x}_t)(a_j - \bar{x}_t)p_t^{(j)}dt + b_t^{-2}(a_j - \bar{x}_t)p^{(j)}dY_t. \end{aligned}$$

This system of stochastic differential equations specifies the dynamic structure of the filter.

For a derivation based on the Zakai equation see [4].

The Wonham filter or Hidden Markov model filter can be used to compute the approximating filter in case when, for example, a diffusion process is approximated by a Markov chain as in [8], which will be discussed later.

1.5 Numerical methods for solving the non-linear filtering problem

Very often a non-linear filtering problem can be quite successfully solved numerically. In this subsection three numerical methods, see Section 8 in [4], are considered, the first of them will be applied further in the thesis.

1.5.1 Extended Kalman filter

This method allows to apply Kalman-Bucy filter to a non-linear filtering problem, see Section 8 in [4]. For this, the non-linear coefficient functions of the signal and observation equations have to be linearized using Taylor expansion.

Let (X_t, Y_t) be the solution of the following 1-dimensional system of non-linear stochastic differential equations:

$$dX_t = a(X_t)dt + b(X_t)dV_t,$$
$$dY_t = h(X_t)dt + dW_t,$$

and assume that $(X_0, Y_0) = (x_0, 0)$. Define $\bar{x}_t$ to be the solution of the ordinary differential equation

$$\frac{d\bar{x}_t}{dt} = a(\bar{x}_t),$$

$\bar{x}_0 = x_0$.

It is assumed that the contribution of the stochastic term in the above system of SDEs remains small, at least within a small window of time $[0;\epsilon]$. The following Taylor-like expansion is applied in order to linearize the filtering problem:

$$dX_t \approx (a'(\bar{x}_t)(X_t - \bar{x}_t) + a(\bar{x}_t))dt + b(\bar{x}_t)dV_t,$$

$$dY_t \approx (h'(\bar{x}_t)(X_t - \bar{x}_t) + h(\bar{x}_t))dt + dW_t,$$

where a' and h' are the derivatives of a and h respectively.

One can assume that for a small time window, the equation satisfied by the pair (X_t, Y_t) is nearly linear. Therefore, it can be assumed that the conditional expectation π_t is normal with mean $\hat{x}_t$ and with covariance R_t, which satisfy the following system of equations:

$$d\hat{x}_t = (a'(\bar{x}_t)\hat{x}_t + a(\bar{x}_t) - a'(\bar{x}_t)\bar{x}_t)dt + R_t h'(\bar{x}_t)(dY_t - (h'(\bar{x}_t)\hat{x}_t + h(\bar{x}_t) - h'(\bar{x}_t)\bar{x}_t)dt),$$

$$\frac{dR_t}{dt} = \sigma_t^2(\bar{x}_t) + 2a'(\bar{x}_t)R_t - R_t^2 h'(\bar{x}_t)^2.$$

with $\hat{x}_0 = x_0$ and $R_0 = r_0$.

The extended Kalman filter is not mathematically proved, but is used in practice. The extended Kalman filter will give a good estimate if the initial position of the signal is well approximated (r_0 is 'small'), the coefficients a and b are nearly linear, h is injective and the system is stable. In the last chapter of the thesis this method will be applied to the case, when the non-linear coefficient functions are square root functions. The square root function can be heurictically considered to be a 'slightly' non-linear function. The filtering results turn out to be quite good for square root functions.

1.5.2 The spectral approach to solving the filtering problem

The density of the unnormalised conditional distribution of the signal is the solution of the stochastic partial differential equation (4), called the Zakai equation:

$$\Phi(u,t) = \Phi(u,s) + \int_s^t A^*\Phi(u,\theta)d\theta + \int_s^t h(u)\Phi(u,\theta)dY_\theta.$$

The spectral approach, see Section 8 in [4] is based on decomposing $\Phi(u,t)$ into a sum of the form:

$$\Phi(u,t) = \sum_\alpha \frac{1}{\sqrt{\alpha!}}\phi_\alpha(t,z)\xi_\alpha(Y), \tag{20}$$

where $\xi_\alpha(Y)$ are Wick polynomials

$$\xi_\alpha = \prod_{k,l}\Big(\frac{H_{\alpha_k^l}(\xi_{k,l})}{\sqrt{\alpha_k^l!}}\Big),$$

where $H_n(x)$ are Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2/2}\frac{d^n}{dx^n}e^{-x^2/2}$$

and $\xi_{k,l} = \int_0^t m_k(s)dY^l(s)$ are random variables, and $\{m_k\} = \{m_k(s)\}_{k\geq 1}$ is an orthonormal system in the space $L^2([0,t])$. Further, a collection $\alpha = (\alpha_k^l)_{1\leq l\leq d, k\geq 1}$ of nonnegative integers is called a d-dimensional multi-index if only finitely many of α_k^l are different from zero. Let J be the set of all d-dimensional multi-indices. For $\alpha \in J$ define

- length of α: $\alpha = \sum_{l,k}\alpha_k^l$;
- the order of α: $d(\alpha) = \max\{k \geq 1,\ \alpha_k^l\ for\ some\ 1 \leq l \leq d\}$;
- $\alpha! = \prod_{k,l}\alpha_k^l!$.

Returning back to (20), consider $\phi_\alpha(t,z)$, which are deterministic Hermite-Fourier coeffcients in the Cameron-Martin orthogonal decomposition of $\Phi(t,z)$ and which satisfy the following system of equations:

$$\frac{d\phi_\alpha(t,z)}{dt} = A^*\phi_\alpha(t,z) + \sum_{k,l}\alpha_k^l m_k(t)h^l(z)\phi_{\alpha(k,l)}(t,z)$$

$$\phi_\alpha(0,z) = \pi_0(z)1_{\alpha=0},$$

where $\alpha(i,j)$ stands for the multi-index $(\tilde{\alpha}_k^l)_{1\leq l\leq d,k\geq 1}$ is defined as $\tilde{\alpha}_t{}^l = \alpha_k^l$, if $k \neq i$ or $l \neq j$ or both, and $\tilde{\alpha}_t{}^l = \max\{0, \alpha_i^j - 1\}$ if $k = i$ and $l = j$.

Theorem 1.7. *(Bain, Crisan) Under certain technical assumptions, the series*

$$\sum_\alpha \frac{1}{\sqrt{\alpha!}}\phi_\alpha(t,z)\xi_\alpha(Y)$$

converges in $L^2(\Omega,\tilde{P})$ and in $L^1(\Omega,P)$ and

$$\Phi(t,z) = \sum_\alpha \frac{1}{\sqrt{\alpha!}}\phi_\alpha(t,z)\xi_\alpha(Y),$$

P-a.s.

This expansion separates the parameters from the observations: the Hermite-Fourier coeffcients are determined only by the coeffcients of the signal process, its initial distribution and the observation function h, whereas the polynomials $\xi_\alpha(Y)$ are completely determined by the observation process.

For computational purposes one has to truncate the sum in the expansion of $\Phi(t,z)$.

1.5.3 The partial differential equation method for solving the Zakai equation

The density $\Phi(t,z)$ of the of the unnormalized conditional distibution of the signal is the solution of the stochastic PDE, i.e. Zakai equation

$$\Phi(u,t) = \Phi(u,s) + \int_s^t A^*\Phi(u,\theta)d\theta + \int_s^t h(u)\Phi(u,\theta)dY_\theta,$$

to which the splitting-up algorithm is applied, following Section 8 in [4].

Let $0 = t_0 < t_1 < ... < t_n < ...$ be a uniform partition of the interval $[0,\infty)$ with the time step $\Delta = t_n - t_{n-1}$. The density $\Phi(t_n, z)$ is approximated by $\phi_n^\Delta(z)$. The transition from $\phi_{n-1}^\Delta(z)$ to $\phi_n^\Delta(z)$ is made in two steps:

- The prediction step consists in solving the following Fokker-Planck equation for the time interval $[t_{n-1,t_n}]$:

$$\frac{\partial \phi_t^n(z)}{\partial t} = A^*\phi_t^n(z)$$

$$\phi_{t_{n-1}}^n = \phi_{n-1}^\Delta.$$

  Denote the prior estimate $\bar{\phi}_n^\Delta = \phi_{t_n}^n$. The Fokker-Planck equation is solved by using the implicit Euler scheme

$$\bar{\phi}_n^\Delta - \Delta A^*\bar{\phi}_n^\Delta = \bar{\phi}_{n-1}^\Delta.$$

  One can approximate the solution to this equation by using a finite difference scheme in order to approximate the differential operator A^*.

- The second step is called the correction step and uses the new observation Y_{t_n} to update $\bar{\phi}_n^\Delta$ for $z \in R$

$$\phi_n^\Delta(z) = c_n \psi_n^\Delta \bar{\phi}_n^\Delta(z),$$

  where $\psi_n^\Delta(z) = \exp\Big(-\frac{1}{2}\Delta||z_n^\Delta - h(z)||^2\Big)$, c_n is a normalization constant chosen such that

$$\int_R \phi_n^\Delta(z)dz = 1$$

and

$$z_n^\Delta = \frac{1}{\Delta}(Y_{t_n} - Y_{t_{n-1}}).$$

1.6 Markov Chain approximations for solving the non-linear filtering problem

Here we consider a non-linear filtering problem, which consists of an unobservable signal, driven by a diffusion process with non-linear drift and diffusion coefficients, and an observation that also follows a diffusion process with a non-linear drift coefficient. Such a filtering problem cannot be solved explicitly, that is why a Markov chain approximation for the diffusion signal process is sought. The way how to construct such a Markov chain is proposed in paper [8].

1.6.1 Continuous-time Markov chain approximation for nonlinear filtering problem without jumps

Here one of the basic papers on approximate filter construction is considered. In the work [8] an approximate solution for the non-linear filtering problem is derived.

A diffusion process is approximated by a Markov chain and then it is shown, that an approximate filter, derived for an approximating Markov Chain, converges to the unknown filter of the original diffusion process. The framework of the paper [8] is considered below.

Consider an 1-dimensional unobservable signal process X_t, which is the solution to the following SDE:

$$dX_t = a(X_t)dt + b(X_t)dV_t,\ 0 \le t \le T, \tag{21}$$

where a and b are R-valued functions respectively on R$\times[0;T]$ and V_t is an 1-dimensional standard Wiener process. The process X_t is partially observed through the 1-dimensional observation process Y_t, that solves another SDE:

$$dY_t = h(X_t)dt + dW_t,$$

where h is an R-valued function on R$\times[0,T]$ and W_t is a 1-dimensional standard Wiener process independent of V_t.

Let F_t^Y be the σ-algebra generated by Y_s, $0 \le s \le t$ and let ϕ be a real-valued meaurable function on R. The nonlinear filtering problem is to evaluate for each t the conditional expectation

$$\pi(\phi(X_t)) = E\{\phi(X_t)|F_t^Y\}.$$

Assumption 1.8. Functions a, b, h and ϕ are continuous and bounded.

The authors of paper [8] consider the Kallianpur-Striebel formula as the way to evaluate the filter. Let $B[0,T]$ be the Borel σ-algebra induced on $C[0,T]$ by the sup norm topology. Let $\bar{X}_t$ be a version of X_t independent of Y_t in the sense that $\bar{X}_t$ induces on $B[0,T]$ the same probability measure as X_t and, $\bar{X}_s, 0 \leq s \leq T$ and $Y_s, 0 \leq s \leq T$ are independent. Define

$$V_\phi(t) = E\Big[\phi(\bar{X}_t)\exp\Big(\int_0^t h(\bar{X}_s)dY_s - \frac{1}{2}\int_0^t h^2(\bar{X}_s)ds\Big)|F_t^Y\Big]. \qquad (22)$$

Then,

$$E\{\phi(X_t)|F_t^Y\} = \frac{V_\phi(t)}{V_1(t)}\ a.s,$$

where $V_1(t)$ is given by (22) in case if $\phi(x) = 1$.

The conditional expectation solves the Kushner-Stratanovich equation, which is in most cases not easy to solve. That is why an approximate solution is searched in [8].

In the paper the original diffusion process X_t is approximated by a continuous-time Markov Chain which is obtained by spatial discretization of the unobservable process X_t. As the discretization step goes to zero, the chain converges weakly to the diffusion and the corresponding function ϕ of the chain converges to the function ϕ of the original diffusion [8]. That is how the filter is approximated.

For the actual computation of the approximating functionals the process X_t is stopped at its first exit from a certain bounded region. In this way an approximating chain with finite number of states is obtained. The diffusion process X_t is approximated by a continuous-time Markov chain $\{X^k, 0 \leq t \leq T\}$ with state space $\mathrm{R}^k = \{z \in \mathrm{R}\text{: } z = X_0 + nk,\ n \text{ integer}\}$.

Proposition 1.9. *(Di Masi, Runggaldier) An approximating Markov Chain in the above sense is given by the following transition rates:*

$$Q_+^k(X_t) = \frac{a^+(X_t)}{k} + \frac{b^2(X_t)}{2k^2};$$

$$Q_-^k(X_t) = \frac{a^-(X_t)}{k} + \frac{b^2(X_t)}{2k^2};$$

$$Q^k(X_t) = -\frac{|a(X_t)|}{k} - \frac{b^2(X_t)}{k^2};$$

where $a^+ = \max(a, 0)$ *and* $a^- = [-a]^+$.

The transition probabilities in [8] are chosen in the following form:

$$P[X^k_{t+\Delta} = x + k | X^k_t = x] = \int_t^{t+\Delta} Q^k_+(X_s)ds;$$

$$P[X^k_{t+\Delta} = x - k | X^k_t = x] = \int_t^{t+\Delta} Q^k_-(X_s)ds;$$

$$\sum_{n \in Z} P[X^k_{t+\Delta} = x + nk | X^k_t = x] = o(\Delta),\ n \neq -1, 0, +1;$$

$$P[X^k_{t+\Delta} = x | X^k_t = x] = 1 - \int_t^{t+\Delta} Q^k(X_t) - o(\Delta).$$

The problem is considered on a bounded region. Let $G \subset$R be open and bounded and let $G^k = G \cap$ R^k and define the stopping times

$$\tau = \begin{cases} \text{T if X}_t \in G \text{ for all } t \in [0, T] \\ \inf\{t : 0 \leq t \leq T, X_t \notin G\} \text{ otherwise} \end{cases}$$

$$\tau^k = \begin{cases} \text{T if X}^k_t \in G^k \text{ for all } t \in [0, T] \\ \inf\{t : 0 \leq t \leq T, X^k_t \notin G^k\} \text{ otherwise} \end{cases}$$

and the corresponding stopped processes

$$\tilde{X}_t = X_{t \wedge \tau},\ \tilde{X}^k_t = X^k_{t \wedge \tau^k}.$$

The number of states of $\tilde{X}^k$ is denoted by N^k and $Q^k_+ + Q^k_- + Q^k = 0$ for $X \in \partial G^k$, where $\partial G^k = \{x \in$R$^k - G^k : \min |x - y| < k,\ y \in G^k$ is the boundary of the state space of $\tilde{X}^k$.

Let $\tilde{Y}_t$ be the observation corresponding to the stopped process $\tilde{X}_t$:

$$\tilde{Y}_t = Y_0 + \int_0^t h(\tilde{X}_s)ds + W_t$$

and $\tilde{F}^Y_t$ be the σ-algebra generated by $\tilde{Y}_s, 0 \leq s \leq t$. Then the problem of filtering the stopped process $\tilde{X}$ looks as follows:

$$E\{\phi(\tilde{X}_t) | \tilde{F}^Y_t\} = \frac{\tilde{V}_\phi(t)}{\tilde{V}_1(t)}\ a.s,$$

where

$$\tilde{V}_\phi(t) = E\Big[\phi(\hat{X}_t)\exp\Big(\int_0^t h(\hat{X}_s)d\tilde{Y}_s - \frac{1}{2}\int_0^t h^2(\hat{X}_s)ds\Big)\Big|\tilde{F}_t^Y\Big], \qquad (23)$$

where $\hat{X}_t$ is a version of $\tilde{X}_t$ independent of $\tilde{Y}_t$. Let $\hat{X}^k$ be a version of X^k independent of $\tilde{Y}_t$, then the approximating functional for evaluating the conditional expectation is as follows

$$\tilde{V}_\phi(t) = E\Big[\phi(\hat{X}_t^k)\exp\Big(\int_0^t h(\hat{X}_s^k)dY_s - \frac{1}{2}\int_0^t h^2(\hat{X}_s^k)ds\Big)\Big|\tilde{F}_t^Y\Big], \qquad (24)$$

In the paper [8] it is shown, that X^k converges weakly to X_t as $k \to \infty$. The proof of convergence is done in the following way:

- First, it is shown that a sequence of approximating chains X^k, $k \to 0$ is sequentially compact, namely there exists a weakly converging subsequence. Denote $D^n[0,T]$ to be the space of R^n-valued functions on $[0,T]$ that are right-continuous, left continuous at T and have left-hand limits, endowed with metric with respect to which $D^n[0,T]$ is separable and complete [8]. On $D^n[0,T]$ weak sequential compactness is equivalent to tightness. Define the following:

$$A_t^k = \int_0^t a(X_s^k)ds, \qquad (25)$$

$$B_t^k = X_t^k - X_0 - \int_0^t a(X_s^k)ds, \qquad (26)$$

so that $X_t^k = X_0 + A_t^k + B_t^k$.

Proposition 1.10. *(Di Masi, Runggaldier) Every sequence* $\{(X^k, A^k, B^k),$ $k \to 0\}$ *is tight in* $D^3[0,T]$ *and the limit of every converging subsequence has trajectories in* $C^3[0,T]$ *a.s.*

For the proof it is sufficient to show, that there exist positive constants K, α and β such that for each scalar component η^k of the sequence $\{(X^k, A^k, B^k)\}$ and for each $t, t+\Delta \in [0,T]$ holds

$$\lim E[|\eta_{t+\Delta}^k - \eta_t^k|^\alpha] \leq K\Delta^{1+\beta},\ k \to 0. \qquad (27)$$

The proof is given in [8] in detail.

- Second, for the proof of convergence one has to show, that every weakly converging sequence $\{X^k,\ k \to 0\}$ converges to the same limit, which is precisely the original diffusion X_t. It is assumed that for each initial condition X_0 equation (21) has on $[0,T]$ a unique solution in the weak sense.

  Theorem 1.11. *(Di Masi, Runggaldier) For each weakly converging sequence* $\{X^k, A^k, B^k,\ k \to 0\}$ *there exists a standard Wiener process* $\tilde{V}_t$ *such that the limiting process is given by* (ξ, A, B) *with*

$$\xi_t = X_0 + A_t + B_t, \tag{28}$$

$$A_t = \int_0^t a(\xi_s)ds, \tag{29}$$

$$B_t = \int_0^t b(\xi_s)d\tilde{V}_s. \tag{30}$$

The convergence of the stopped processes is considered further. For this define $\bar{G}$ to be the closure of G and define the stopping time

$$\tau' = \begin{cases} \text{T if X}_t \in \bar{G} \text{ for all } t \in [0,T] \\ \inf\{t : 0 \le t \le T, X_t \notin \bar{G}\} \text{ otherwise.} \end{cases}$$

Then the theorem in [8] states, that if $P\{\tau' = \tau\} = 1$ then $\tilde{X}_t^k$ converges weakly to $\tilde{X}_t$ as $k \to 0$.

The next theorem shows the convergence of the filters for the original diffusion process and its Markov chain approximation.

Theorem 1.12. *(Di Masi, Runggaldier) If* $P\{\tau' = \tau\} = 1$*, then* $\tilde{V}_\phi^k(t)$ *converges to* $\tilde{V}_\phi(t)$ *a.s. for each* $t \in [0,T]$*, as* $k \to 0$.

Proof. (Di Masi, Runggaldier)
Consider

$$\tilde{V}_\phi(t) = E\Big[\phi(\hat{X}_t^k)\exp\Big(\int_0^t h(\hat{X}_s^k)h(\tilde{X}_s)ds + \int_0^t h(\hat{X}_s^k)dW_s - \frac{1}{2}\int_0^t h^2(\hat{X}_s^h)ds\Big)\,\Big|\tilde{F}_t^Y\Big]. \tag{31}$$

Using the Skorohod imbedding, every sequence $\{\hat{X}^k, h \to 0\}$ converges with probability 1 and uniformly on $[0, T]$ to $\hat{X}$. Due to the boundedness of ϕ and h

$$\phi(\hat{X}_t^k) \to \phi(\hat{X}_t),$$

$$\int_0^t h(\hat{X}_s^k) h(\tilde{X}_s) ds \to \int_0^t h(\hat{X}_s) h(\tilde{X}_s) ds,$$

$$\int_0^t h^2(\hat{X}_s^k) ds \to \int_0^t k^2(\hat{X}_s) ds,$$

$$\int_0^t |h(\hat{X}_s^k) - h(\hat{X}_s)|^2 ds \to 0$$

converge with probability 1. □

1.6.2 Continuous-time Markov chain approximations for the nonlinear filtering problem with discontinuous observations

This subsuction is based on the paper [9]. Let a partially observable process X_t, Y_t, $t \in [0, T]$ be given on a probability space (Ω, F, P). The unobservable component X_t, called the signal process follows the diffusion process and is given by the following SDE as in [8]:

$$dX_t = a(X_t)dt + b(X_t)dV_t, \ 0 \le t \le T,$$

where a and b are R-valued functions respestively on R$\times[0; T]$, V_t is a 1-dimensional standard Wiener process. The signal process X_t is partially observed through the observation process Y_t which follows

$$dY_t = h(X_t)dt + dW_t + dN_t,$$

where h is an R-valued function on R$\times[0, T]$, W_t is a 1-dimensional standard Wiener process and N_t is doubly stochastic Poisson process with rate $\lambda(X_t)$. Also W_t is independent of V_t and N_t.

Let F_t^Y be the σ-algebra generated by Y_s, $0 \le s \le t$ and let ϕ be a real-valued meaurable Borel function on R. As it was already mentioned earlier, the nonlinear filtering problem is to evaluate the conditional expectation

$$\pi(\phi(X_t)) = E_P\{\phi(X_t)|F_t^Y\},$$

for each time moment t.

Further, consider in detail how such a filter can be approximated. Assume, that $\lambda(x)$ and $\phi(x)$ are bounded and continuous functions; $h(x)$ is bounded and of class C^2; also functions a and b are bounded and continuous.

The new measure Q on (Ω, P) is defined in [9]:

$$\frac{dQ}{dP} = \exp\Big[-\int_0^T h(X_s)dW_s - \frac{1}{2}\int_0^T h^2(X_s)ds$$

$$-\int_0^T \log\lambda(X_s)dN_s - \int_0^T [1-\lambda(X_s)]ds\Big],$$

such that $Y_t^c = Y_t - N_t$ is an (F,Q)-standard Wiener process, N_t is (F,Q)-standard Poisson process and under Q the processes Y_t and X_t are independent. Define

$$L_t = E_Q\left\{\frac{dP}{dQ}\Big|F_t^X \vee F_t^Y\right\},$$

where $F_t^X = \sigma(X_s, 0 \leq s \leq t)$ and $F_t^Y = \sigma(Y_s, 0 \leq s \leq t)$.

The process L_t is given by

$$L_t = \exp\Big[\int_0^T h(X_s)dY_s^c - \frac{1}{2}\int_0^T h^2(X_s)ds$$

$$+\int_0^T \log\lambda(X_s)dN_s + \int_0^T [1-\lambda(X_s)]ds\Big],$$

where Y_t^c is the continuous part of the process Y_t, and the following filter representation holds:

$$\pi(\phi(X_t)) = E_P[\phi(X_t)|F_t^Y] = \frac{E_Q[\phi(X_t)L_t|F_t^Y]}{E_Q[L_t|F_t^Y]}$$

$$= \frac{V_t(y,\phi)}{V_t(y,1)} = G_t(y),\ P-a.s.$$

The idea in the paper [9] is to approximate $V_t(y,\phi)$. First the signal process X_t is approximated by a sequence of weakly-converging, continuous-time and finite-state Markov chains X_t^k. The exact way how it can be done was considered in the previous section. Then the sequence of approximating functionals is defined:

$$V_t^k(y,\phi) = E_Q[\phi(X_t^k)L_t^k|F_t^Y],$$

where L_t^k differs from L_t only by changing X_t into X_t^k.

One can prove that under certain assumptions

$$\lim_{h\to 0} V_t^h(y,\phi) = V_t(y,\phi),$$

$t \in [0, T]$ P-a.s. and therefore

$$G_t^k(y) = \frac{V_t^k(y, \phi)}{V_t(y, 1)} \approx G_t(y).$$

In the paper [9] it is proved that the optimal filter $G_t(y)$ is robust, i.e. for each $t \in [0, T]$ the filter $G_t(y)$ is P-a.s. continuous. The sequence of approximating filters is also robust.

2 Introduction to Lévy processes

This section gives a short overview over the Lévy or jump-diffusion processes, which will be widely used further in the thesis. This theoretical backgroung is from the book [2].

Let $X = (X_t, t \geq 0)$ be a stochastic process defined on a probability space (Ω, F, P), then X is a Lévy process if

- $X(0) = 0$ a.s.
- X has independent and stationary increments
- X is stochastically continuous, i.e. for all $a > 0$ and for all $s \geq 0$

$$\lim_{t \to s} P(|X_t - X_s| > a) = 0.$$

The Lévy-Khintchine formula for a Lévy process $X = (X_t, t \geq 0)$ is

$$E\left(e^{i(u,X_t)}\right) = \exp\Big(t\Big[i(b,u) - \frac{1}{2}(u, Au) + \int_{R-\{0\}} [e^{i(u,y)} - 1 - i(u,y) I_{\hat{B}}(y)] v(dy)\Big]\Big),$$

for each $t \geq 0$, $u \in R$, where (b, A, v) are the characteristics of X and v is Lévy measure; I is the characteristic function, $\hat{B} = B_1(0)$ is the unit ball.

Lévy measure v is a Borel measure defined on $R - \{0\} = \{x \in R, x \neq 0\}$ such that

$$\int_{R-\{0\}} (|y|^2 \wedge 1) v(dy) < \infty \; or \; \int_{R-\{0\}} \frac{|y|^2}{1 + |y|^2} v(dy) < \infty.$$

Any Lévy process is a Markov process i.e.

$$E(f(X_t)|F_s^X) = E(f(X_t)|X_s)$$

a.s. for $s > t$.

Proposition 2.1. ***The Lévy-Ito decomposition***: *If X is a Lévy process, then there exists $b \in R$, a Brownian motion B^A with covariance matrix A and an independent Poisson random measure N on $R^+ \times (R - \{0\})$ such that for each $t \geq 0$*

$$X_t = bt + B_t^A + \int_{|x|<1} x\tilde{N}(t, dx) + \int_{|x|\geq 1} xN(t, dx).$$

Poisson random measure $N(t, A)(\omega)$ is for each $t > 0$, $\omega \in \Omega$ a counting measure on $B(R - \{0\})$:

$$N(t, A)(\omega) = \{0 \leq s \leq t; \Delta X_s(\omega) \in A\}.$$

For each bounded A, $(N(t, A), t \geq 0)$ is a Poisson process with intensity $v(A) = E(N(1, A))$ and $\tilde{N}$ is a compensated Poisson random measure (martingale measure) defined by $\tilde{N}(t, A) = N(t, A) - tv(A)$ for bounded A.

Every Lévy process is a semimartingale.

Next, consider the generators of Lévy processes. Let X be a Lévy process with characteristics (b, a, v) and let A be the infinitesimal generator of X:

$$(Af)(x) = b\frac{\partial f(x)}{\partial x} + \frac{1}{2}a\frac{\partial^2 f(x)}{\partial x^2} + \int_{R-\{0\}} [f(x+y) - f(x) - y\frac{\partial f(x)}{\partial x} I_{\hat{B}}(y)]v(dy).$$

Ito formula for Lévy process. Let X be a Lévy process with stochastic differential

$$dX_t = G_t dt + F_t dB_t + \int_{|x|<1} H(t, x)\tilde{N}(dt, dx) + \int_{|x|>1} K(t, x)N(dt, dx). \quad (32)$$

Continuous part is $X_t^c = G_t dt + F_t dB_t$ and discontinuous part is $X_t^d = \int_{|x|<1} H(t, x)\tilde{N}(dt, dx) + \int_{|x|>1} K(t, x)N(dt, dx)$, therefore $X_t = X_0 + X_t^c + X_t^d$.

Theorem 2.2. ***Ito formula***: *If X is a Lévy-type stochastic process of the form (32), then for each $f \in C^2(R)$, $t \geq 0$, with probability 1*

$$f(X_t) - f(X_0) = \int_0^t \frac{\partial f(X_{s-})}{\partial x} dX_s^c + \frac{1}{2}\int_0^t \frac{\partial^2 f(X_{s-})}{\partial x^2} d[X^c, X^c](s)$$
$$+ \int_0^t \int_{|x|\geq 1} [f(X_s + K(x, s)) - f(X_{s-})]dN(ds, dx)$$
$$+ \int_0^t \int_{|x|<1} [f(X_s + H(x, s)) - f(X_{s-})]d\tilde{N}(ds, dx)$$
$$+ \int_0^t \int_{|x|<1} [f(X_s + H(x, s)) - f(X_{s-}) - H(x, s)\frac{\partial f(X_{s-})}{\partial x}]v(dx)ds.$$

Ito's product formula. Consider the Lévy-type stochastic processes $X^{(1)}$ and $X^{(2)}$ of the form (32), then for all $t \geq 0$ with probability 1

$$d(X_t^{(1)}, X_t^{(2)}) = X_{t-}^{(1)} dX_t^{(2)} + X_{t-}^{(2)} dX_t^{(1)} + d[X^{(1)}, X^{(2)}]_t,$$

where Ito correlation is as follows:

$$d[X^{(1)}, X^{(2)}]_t = F^{(1)}(s)F^{(2)}(s)ds + \int_{|x|<1} H^{(1)}(s,x)H^{(2)}(s,x)\tilde{N}(ds,dx)$$

$$+ \int_{|x|\geq 1} K^{(1)}(s,x)K^{(2)}(s,x)N(ds,dx).$$

Exponential martingales. Consider the Lévy-type process

$$dX_t = G_t dt + F_t dB_t + \int_{|x|<1} H(t,x)\tilde{N}(dt,dx) + \int_{|x|>1} K(t,x)N(dt,dx).$$

Also consider the process $e^X = (e^{X_t}, t \geq 0)$ and a condition, under which this process is a martingale.

Corollary 2.3. *(Appelbaum) e^Y is a local martingale if and only if*

$$G(s) + \frac{1}{2}F^2(s) + \int_{|x|<1} \Big(e^{H(s,x)} - 1 - H(s,x)\Big) v(dx)$$

$$+ \int_{|x|\geq 1} \Big(e^{K(s,x)} - 1\Big) v(dx) = 0$$

almost surely and for almost all $s \geq 0$.

The heuristic explanation to the above result can be obtained in the following way. By Ito formula

$$e^{X_t} = 1 + \int_0^t e^{X_{s-}} F(s) dB_s + \int_0^t \int_{|x|<1} \left(e^{H(s,x)} - 1\right) \tilde{N}(ds,dx)$$

$$+ \int_0^t \int_{|x|\geq 1} \left(e^{K(s,x)} - 1\right) \tilde{N}(ds,dx)$$

$$+ \int_0^t e^{X_{s-}} [G(s) + \frac{1}{2}F^2(s) + \int_{|x|<1} \Big(e^{H(s,x)} - 1 - H(s,x)\Big) v(dx)$$

$$+ \int_{|x|\geq 1} \Big(e^{K(s,x)} - 1\Big) v(dx)\Big] ds.$$

In order e^{X_t} to be a martingale, the drift term (the term at ds) should be zero. This is stated in the corollary above.

Let e^X be an exponential martingale, then an equivalent probability measure Q on (Ω, F) can be defined as follows:

$$\frac{dQ_t}{dP_t} = e^{X_t}.$$

These were the most inportant facts, definitions and theorems concerning Lévy processes.

3 Non-linear filtering problems for Lévy processes

The recent publications consider filtering problems for the jump-diffusion processes. There are several works, where an attempt is made to solve explicitly the non-linear filtering problem, in which the signal process or the observation process or both are jump-diffusions. In two of the following works a corresponding Zakai equation for jump-diffusion processes is derived and in one work this Zakai equation is solved theoretically.

3.1 The case when the unobservable process follows the diffusion process and the observation process follows the jump-diffusion process

This theoretical approach is based on the paper [28]. Consider the following non-linear filtering problem where X_t is the signal process driven by standard Wiener process V_t

$$dX_t = a(X_t)dt + b(X_t)dV_t, \tag{33}$$

and Y_t is observation process

$$dY_t = h(X_t)dt + dW_t + \int_{R_0} \zeta N_\lambda(dt, d\zeta) \tag{34}$$

driven by standard Wiener process W_t and an integer-valued random measure N_λ. This integer-valued random measure has a predictable compensator $\hat{\mu}(dt, d\zeta, \omega) = \lambda(t, X_t, \zeta)dt\nu(d\zeta)$, where ν is a Lévy measure. It is assumed that (V_t, W_t) is a Wiener process independent of N_λ.

The aim is to determine the conditional expectation

$$E[f(X_t)|F_t^Y],$$

where f is a Borel function and F_t^Y is the σ-algebra, generated by $\{Y_s, 0 \leq s \leq t\}$. Consider conditional distribution $P[X_t|F_t^Y]$ such that $P[X_t \in dx|F_t^Y](\omega) = p(t,x,\omega)dx$, and also consider the process Φ which is related to $p(t,x,\omega)$ in the following way:

$$p(t,x,\omega) = \frac{\Phi(t,x,\omega)}{\int_R \Phi(t,x,\omega)dx}.$$

This function Φ fullfills Zakai equation (4). The equivalent measure μ on (Ω, F) is defined via $d\pi = \Lambda_t d\mu$ with Radon-Nikodym density

$$\Lambda_t = \exp\Big\{ \int_0^t h(X_s)dW_t - \frac{1}{2}\int_0^t h^2(X_s)ds$$
$$+ \int_0^t \int_{R_0} \log \lambda(s, X_s, \zeta) N_\lambda(ds, d\zeta) + \int_0^t \int_{R_0} (1 - \lambda(s, X_s, \zeta))ds\nu(d\zeta) \Big\}.$$

Processes (33)-(34) get transformed under measure μ in the following way:

$$dX_t = a(X_t)dt + b(X_t)dV_t,$$
$$dY_t = B_t + L_t,$$

where Y_t is a Lévy process independent of X_t under μ with

$$B_t = W_t - \int_0^t h(X_s)ds$$

and

$$L_t = \int_0^t \int_{R_0} \zeta N(ds, d\zeta),$$

where under transformed measure μ we have $\lambda = 1$ and we write $N(ds, d\zeta) = N_1(ds, d\zeta)$.

Thererfore, the observation is decomposed into a Brownian motion part and jump part.

Under this setting the function Φ satisfies the following Zakai equation [28]:

$$\Phi(t,x) = \int_0^t L^*\Phi(s,x)ds + \int_0^t h(x)\Phi(s,x)dB_s$$
$$+ \int_0^t \int_{R_0} (\lambda(s,x,\zeta) - 1)\Phi(s,x)\tilde{N}(ds, d\zeta),$$

where L^* is the adjoint operator of the generator L of X_t, $\tilde{N}(ds, d\zeta) = N(ds, d\zeta) - ds\nu(d\zeta)$ and also $\Phi(0,x) = p_0(x)$, where $p_0(x)$ is the density function of the initial condition X_0.

Operator L^* can be written as $L^* = A - c$ according to [28], where A is the generator of a diffusion and the function c is in $C_b^{1,2}(R_+ \times R) \cap C^{2+\beta}(R_+ \times R)$, where $C_b^{n,m}(R_+ \times R)$ is the space of continuously differentiable functions (n times in t, m times in x) with all partial derivatives bounded and $C^{n+\beta}(R_+ \times R)$ denotes the space of functions whose partial derivatives up to order n are Hölder-continuous of order $0 < \beta \leq 1$.

The authors of [28] make several assumptions, the most important of which are the following:
- The Lévy measure v is finite ;
- The generator A is uniformly elliptic;
- The coefficients a and b are Hölder continuous and belong to $C_b^{1,3}(R_+ \times R)$; the observation function h is in $C_b^{1,2}(R_+ \times R) \cap C^{2+\beta}(R_+ \times R)$ and also the intensity rate λ is in the same space and is strictly positive;
- The initial distribution $p_0(x)$ is positive and is an element of $C_b^{2\beta}(R)$.

Using results from white-noise analysis and the Feynman-Kač representation theorem, the authors of [28] get:

Theorem 3.1. *(Meyer-Brandis, Proske) The solution $\Phi(t, x)$ is a unique strong solution in $L^p(\mu)$, $p \geq 1$ and twice continuously differentiable in x. It takes the following form:*

$$\Phi(t, x) = E^x \Big[p_0(X_t(\theta)) \exp\Big(\int_0^t c(X_{t-s}(\theta)) ds \Big) \times$$

$$\times \exp\Big(\int_0^t h(X_s(\theta)) dW_t(\omega) - \frac{1}{2} \int_0^t h^2(X_s(\theta)) ds$$

$$+ \int_0^t \int_{R_0} \log \lambda(s, X_s(\theta), \zeta) \tilde{N}_\lambda(ds, d\zeta, \omega)$$

$$+ \int_0^t \int_{R_0} (\log(\lambda(s, X_{t-s}(\theta), \zeta)) - \lambda(s, X_s(\theta), \zeta) - 1)) ds \nu(d\zeta) \Big) \Big) \Big],$$

where $X_s(\theta) = X_s^x(\theta)$ is a diffusion process associated with A, which starts at time zero in x and which is defined on the auxiliary probability space (Θ, G, ϑ), and E^x is the expectation with respect to the measure ϑ of $X_s = X_s^x$.

3.2 The case when the unobservable process follows the jump-diffusion process and the observation process follows the diffusion process

This subsection os based on the paper [32].

Consider the nonliner filtering problem of the form

$$dX_t = a(X_t)dt + b(X_t)dV_t + dJ_t,$$

where a and b are bounded continuous functions on R, V_t is standard Brownian motion and J_t the jump process $J_t = \int_0^t \int_\Gamma g(X_{s-}, \rho)N(ds, d\rho)$, where $N(ds, d\rho)$ is Poisson random measure on the Borel sets $[0, \infty) \times \Gamma$, independent of the Brownian motion V_t with mean rate equal to $EN(t + \Delta, A) - EN(t, A) = \lambda \Delta F(A)$. Here λ is the jump rate and $F(.)$ is the probability measure on the space of jump hights Γ.

The process $\{X_t, t \geq 0\}$ is called the signal or unobservable process and is defined on the probability space (Ω, F, P). On the same probability space there is an obserbable process $\{Y_t, t \geq 0\}$ defined as follows:

$$dY_t = h(X_t)dt + dW_t,$$

where function h is continuous and bounded and the standard Brownian motion W is independent of X_t. Denote by $F_t^Y = \sigma\{Y_s, s \leq t\}$ the sigma-algebra generated by the observations. The filtering problem consists of computing the least squares estimate of $f(X_t)$, where $f \in C_0^2(R)$, given the observations F_t^Y:

$$E_P[f(X_t)|F_t^Y].$$

In the paper [32] the optimal filter $E_P[f(X_t)|F_t^Y]$ is defined through the Kallianpur-Striebel formula:

$$E_P[f(X_t)|F_t^Y] = \frac{E_Q[f(X_t)Z_t|F_t^Y]}{E_Q[Z_t|F_t^Y]},$$

where Z_t is a Radon-Nikodym derivative of the form

$$\frac{dQ}{dP} = Z_t = \exp\left[-\int_0^t h(X_s)dW_s - \frac{1}{2}\int_0^t h^2(X_s)ds\right].$$

this density according to Girsanov theorem allows to define a new measure Q such that the observations Y_t become independent from the unobservable signal process X_t. So, under Q, Y_t is a standard Brownian motion and, therefore, is independent of X_t. In the paper [32] the equation for the unnormalized conditional density $\rho_t(f) = E_Q[f(X_t)Z_t|F_t^Y]$ is derived

$$\rho_t(f) = \rho_0(f) + \int_0^t \rho_{s-}(fh)dY_s + \int_0^t \rho_{s-}(Lf)ds$$

$$-\int_0^t \rho_{s-}\left(\lambda \int_\Gamma (f(\cdot + g(\cdot, \rho)) - f(\cdot))F(d\rho)\right) ds,$$

where

$$Lf(x) = f'(x)a(x) + \frac{1}{2}f''(x)b^2(x) + \lambda \int_\Gamma (f(x+g(x,\rho)) - f(x))F(d\rho)$$

is the differential operator of the diffusion process X_t.

The normalized filtering equation is the equation that is satisfied by $\pi_t(f) = E_P[f(X_t)|F_t^Y]$ or $\pi_t(f) = \rho_t(f)/\rho_t(1)$. The optimal filter π_f satisfies

$$\pi_t(f) = \pi_0(f) + \int_0^t [\pi_{s-}(fh) - \pi_{s-}(f)\pi_{s-}(h)]dY_s + \int_0^t \pi_{s-}(Lf)ds + \int_0^t \pi_{s-}^2(h)ds$$
$$- \int_0^t \pi_{s-}\left(\lambda \int_\Gamma (f(\cdot + g(\cdot,\rho)) - f(\cdot))F(d\rho)\right) ds.$$

Remark 3.2. In [15] the authors follow the arguments for diffusions in [4] as close as possible to obtain a Zakai equation for observation and signal following jump-diffusions. For details we refer to [15].

Remark 3.3. All the approaches discussed above, are purely theoretical and are difficult in most cases to implement in practice, since they require the observations of the jumps as it is given in continuous time and not for discrete observations. That is why in the remaining parts of the thesis an approximate solution to the non-linear jump-diffusion filtering problem is considered.

4 Portfolio optimization

In this section a general approach to solving the portfolio optimization problem is discussed. Portfolio optimization is one of many industrial applications, where the approximation to the non-linear jump-diffusion filter can be used.

Let (Ω, F, P) be a probability space. On this space there are two correlated Brownian motions V_t, W_t with correlation ρ, which is often considered to be of zero value.

The portfolio optimization problem consists of maximizing the expected utility of the terminal wealth at time T with respect to utility function. Consider the definition from the paper [14].

Definition 4.1. A function $U : [0,\infty) \to R \cup \{-\infty\}$ is called a utility function, if U is stricktly increasing, strictly concave, twice continuously differentiable on $(0,\infty)$ and satisfies the Inada conditions

$$U'(\infty) = \lim_{x\to\infty} U'(x) = 0,\ U'(0+) = \lim_{x\to 0} U'(x) = \infty.$$

Let us consider power utility functions $U(x) = \frac{1}{\gamma}x^{\gamma}$, $x \geq 0$ and $\gamma < 1$, $\gamma \neq 0$. Assume that the investor has the following investment opportunities:

- bank account with the dynamics $dP_t = P_t r dt$, where risk-free rate r is constant;
- risky asset whose price dynamics follows SDE

$$dS_t = S_t[\mu(z_t, t)dt + \sigma(z_t, t)dW_t],$$

where $dz_t = a(z_t, t)dt + b(z_t, t)dV_t$, a, b are real valued measurable functions. Let F be the filtration generated by W and V and F^S the filtration generated by S, both augmented by null sets. Let us construct the portfolio that consists of the bank account and risky assets.

Definition 4.2. For initial capital $x_0 > 0$ the wealth process $X^{\pi} = (X_t^{\pi})_{t\in[0,T]}$ is defined by

$$dX_t^{\pi} = \pi_t[\mu(z_t, t)dt + \sigma(z_t, t)dW_t] + (X_t^{\pi} - \pi_t)rdt,\ X_0^{\pi} = x_0. \tag{35}$$

The investor rebalances his portfolio dynamically by choosing at any time $s \in [t, T]$ the amount π_s to be invested into the stock account.

Definition 4.3. A trading strategy $\pi = (\pi_t)_{t\in[0,T]}$ is a 1-dimensional F-adapted or F^S-adapted, measurable process satisfying $\int_0^T \sigma_t \pi_t dt < \infty$ and $\int_0^T |\mu_t \pi_t| dt < \infty$.

In the first case (F-adapted) we have *full information*, in the second case (F^S-adapted) we have *partial information*.

It is assumed that all coefficients of the above SDEs are progressively measurable with respect to the Brownian filtration $\{F_t\}_{t\geq 0}$ i.e.

$$\int_0^T (|\mu(z_s, s)| + (\sigma(z_s, s))^2)ds < \infty \ a.s.,$$

$$\int_0^T (|\mu(z_s, s)\pi_s| + \pi_t^2\sigma^2(z_s, s))ds < \infty \ a.s.,$$

and that the above SDEs have unique solutions. The global Lipschitz conditions and linear growth conditions on the coefficients $\mu(z, t)$, $\sigma(z, t)$, $a(z, t)$, $b(z, t)$:$R \times [0, T] \to R$ should be satisfied.

The portfolio optimization problem is to maximize the expected utility of the terminal wealth

$$\max_{\pi} E\Big[U(X_T^{\pi})\Big] = \max_{\pi} E\Big[\frac{1}{\gamma}(X_T^{\pi})^{\gamma}\Big].$$

The value function is defined in the following way:

$$J(t,x,z) = \max_{\pi} E^{t,x,z}\Big(\frac{1}{\gamma}(X_T^{\pi})^{\gamma}\Big).$$

In order to solve the optimization problem, let us first consider the optimization method from the control theory.

Hamilton-Jacobi-Bellman equation for stochastic control. Consider the following stochastic system (see [20]):

$$dX_t = a(t,X_t,u_t)dt + b(t,X_t,u_t)dW_t, \tag{36}$$
$$X_0 = x,$$

where W_t is Brownian motion, $t \in [0,T]$, u_t is control function. Let the cost functional be defined as follows:

$$V(t,X,u(\cdot)) = E\left\{\int_0^T f(t,X_t,u_t)dt + h(X_T)\right\}.$$

One has to maximize $V(t,X_t,u(\cdot))$ over all admissible $u(\cdot)$.

The value function is defined in the following way:

$$J(t,X_t) = \sup_{u(\cdot)} V(t,X_t,u(\cdot)),$$
$$J(T,x) = h(x).$$

The HJB equation is as follows, assuming that $J \in C^{1,2}$:

$$J_t + \sup_u \left\{\frac{1}{2}b^2(t,X_t,u)J_{xx} + a(t,X_t,u)J_x + f(t,X_t,u)\right\} = 0, \tag{37}$$
$$J|_{t=T} = h(x).$$

If one has a stochastic process that itself depends on another stochastic process, for example:

$$dX_t = a(t,X_t,c_t,u_t)dt + b(t,X_t,u_t)dW_t, \tag{38}$$
$$dc_t = l(t,c_t)dt + g(t,c_t)dV_t,$$

where V_t is a Brownian motion and if the value function J itself depends on t, X_t and c_t, i.e. is of the form $J(t,x,c)$, then the HJB equation looks as follows:

$$J_t + \sup_u\Big\{\frac{1}{2}b^2(t,x,u)J_{xx} + a(t,x,c,u)J_x$$
$$+\frac{1}{2}g^2(t,c)J_{cc} + l(t,c)J_c + f(t,x,u)\Big\} = 0, \tag{39}$$

if V and W are independent. If they are correlated there will be an additional term J_{xc} due to the covariation.

4.1 Solving the optimal portfolio problem

This framework is based on the papers [21] and [38]. In order to find the optimal startegy π_t^* under full information, so as to maximize the expected utility of the terminal wealth, given be equation (35), the corresponding Hamilton-Bellman-Jacobi equation has to be solved, following [38]:

$$\sup_{\pi}\Big(J_t + x(\mu(z,t)-r)\pi J_x + a(z,t)J_z + \frac{1}{2}x^2(\sigma(z,t))^2\pi^2 J_{xx}$$
$$+b(z,t)x\sigma(z,t)\rho J_{xz} + \frac{1}{2}(b(z,t))^2 J_{zz}\Big) = 0, \tag{40}$$

where J is the candidate for the value function and the terminal condition for the HJB equation $J(T,x,z) = \frac{1}{\gamma}x^\gamma$, i.e. we use the utility function $U(x) = \frac{x^\gamma}{\gamma}$ with $\gamma < 1$, $\gamma \neq 0$ as in [38]. The value function is assumed to be of the form $J(t,x,z) = \frac{1}{\gamma}(f(t,z))^c$. In order to find the coefficient c and the function $f(z,t)$ the value function is substituted into the HJB equation, the derivative with respect to π_t is calculated and set to zero. This is how the optimal strategy is computed. Then the optimal strategy π_t^* is substituted in the HJB equation (40) and coefficient c together with the function $f(t,z)$ is determined.

According to the paper [38] the value function J is as follows:

$$J(t,x,z) = \frac{x^\gamma}{\gamma}(f(t,z))^{\frac{1-\gamma}{1-\gamma+\rho^2\gamma}}.$$

Function f: $\mathrm{R}\times[0,T] \to \mathrm{R}^+$ solves the linear parabolic equation

$$f_t + \frac{1}{2}b^2(z,t)f_{zz} + \left[a(z,t) + \rho\frac{\gamma\mu(z,t)b(z,t)}{(1-\gamma)\sigma(z,t)}\right]f_z$$
$$+\frac{\gamma(1-\gamma+\rho^2\gamma)}{1-\gamma}\left[r + \frac{(\mu(z,t)-r)^2}{2(\sigma(z,t))^2(1-\gamma)}\right]f = 0,$$

with the terminal condition $f(z,T) = 1$. The optimal portfolio strategy π^* is as follows:

$$\pi_t^* = \frac{1}{1-\gamma}\frac{\mu(z_t,t)}{\sigma^2(z_t,t)} + \frac{1}{1-\gamma}c\cdot\rho\frac{b(z_t,t)f_z(z_t,t)}{\sigma(z_t,t)f(z_t,t)}.$$

Next consider the change of measure from P to Q which is defined by Girsanov density

$$Z_t = \frac{dQ}{dP} = \exp\left(-\frac{1}{2}\left(\frac{\gamma}{1-\gamma}\right)^2\rho^2\int_0^t\frac{\mu^2(z_t,t)}{\sigma^2(z_t,t)}ds + \frac{\gamma}{1-\gamma}\rho\int_0^t\frac{\mu(z_t,t)}{\sigma(z_t,t)}dW_s\right).$$

The above results are only meaningful if Z is a density and f is well-defined.

The verification theorem for this portfolio optimization result was formulated and proved in the paper [21]. In this paper it was assumed that the coefficient γ of the power utility function is in $(0, 1)$.

Definition 4.4. (Kraft) A portfolio strategy is said to be admissible if the following conditions are satisfied:

- π is progressively measurable;
- for all initial conditions $(t_0, x_0, z_0) \in [0, T] \times (0, \infty)^2$ the wealth process X^π with $X^\pi_{t_0} = x_0$ has a pathwise unique solution $\{X^\pi_t\}_{t\in[t_0,T]}$;
- $E^{t_0,x_0,z_0}\left(\left[\frac{1}{\gamma}(X^\pi_T)^\gamma\right]^-\right) < +\infty$;
- $X^\pi \geq 0$. The set of admissible strategies is denoted by A. Besides, A_2 denotes the subset of all admissible strategies π that belong to $L^2[0, T]$, i.e. $E\left(\int_0^T \pi_s^2 ds\right) < \infty$.

Definition 4.5. (Kraft)
Property U: assume that Z is well-defined and that we have $f \in C^{1,2}[0, T]$. Let $\pi \in A$. If for all sequences of stopping times $\{\theta_p\}_{p\in N}$ with $0 \leq \theta_p \leq T$ the sequence $\{J(\theta_p, X^\pi_{\theta_p}, z_{\theta_p})\}$ is uniformly integrable, then π has property U.

Assumption 4.6. (Kraft) For all bounded sets $I \subset [0, T] \times [0, \infty)$ there exists some constant K such that $|a(z, t)| + |b(z, t)| \leq K$ for all $(z, t) \in I$.

Theorem 4.7. *(Kraft)* ***Verification Result****: assume that Z is well-defined, that $\gamma \in (0, 1)$ and that $f \in C^{1,2}[0, T]$. Then*

$$E^{t_0,x_0,z_0}\left(\frac{1}{\gamma}(X^\pi_T)^\gamma\right) < J(t_0, x_0, z_0)$$

for all $\pi \in A$. Let $\pi^* \in A_2$. If Assumption 4.6 holds and π^* has property U, then

$$E^{t_0,x_0,z_0}\left(\frac{1}{\gamma}(X^\pi_T)^\gamma\right) = J(t_0, x_0, z_0).$$

5 Summary of Part I

In this part of the thesis a general portfolio optimization problem for diffusion processes was discussed. It was assumed that both asset prices and drift values of the asset price model were observable (the case of full information). Such situation cannot occur in practice, because in the real market only asset price can be observed, but not the drift values (the case of partial information). In order to evaluate the drift of the asset price model one has to apply filtering techniques. The resulting filter will be applied for solving the portfolio optimization problem under partial information. Such a portfolio optimization problem has to be solved in a slightly different way compared to the portfolio optimization problem under full information.

Part II

Filtering, approximation, portfolio optimization for jump-diffusions

Jump-diffusion processes are used in many applications and quite often they come up in modelling the financial markets. Abrupt movements of asset prices, default situations and etc. can be modelled by jump-diffusion processes. The filtering problem for jump-diffusion processes cannot be easily solved explicitly. In the introductery part of the thesis two different papers, concerning jump-diffusion filtering problem were considered in 3.1 and 3.2. The explicit solution to the jump-diffusion filtering problem is presented only in 3.1, although this solution can hardly be implemented.

Therefore, in this part of the thesis an approximate solution to the jump-diffusion filtering problem is sought.

We consider here the following asset price model: the observable asset price follows a linear jump-diffusion process and the unobservable drift of the asset price follows a linear diffusion process (Ornstein-Uhlenbeck process). The jump part of the asset price process is either driven by a shot-noise process or by a compound Poisson prosess. The aim is to filter the unobservable drift of the asset price, which will be further used for solving the portfolio optimization problem. The portfolio consists of the bank account and the risky asset that follows the jump-diffusion process. The portfolio optimization problem is to find the optimal trading strategy so as to maximize the expected utility of the terminal wealth of the portfolio.

6 Filtering shot noise process

6.1 Motivation

The shot-noise process has already been studied for quite a long time. Very often the shot-noise process is used to model the stochastic intensity of a Poisson process. Such a process is called a doubly stochastic Poisson process or the Cox process. In the paper [7] it is shown that both shot-noise process and doubly stochastic Poisson process have normal approximations, which is applied for pricing reinsurance contracts.

In the paper [19] the asymptotic properties of the explosive Poisson shot-noise process are described and the shot-noise process is applied to modelling the delay in claim settlement.

In the work [17] the Cox process with shot-noise intensity is used to model the default time and to derive the survival probability to price the defaultable zero-coupon bonds with zero recovery. Also the shot-noise process is used for pricing extreme insurance claims [16].

Some of the recent publications propose to model extreme movements of asset prices using shot-noise process, for example paper [34]. Usually any abrupt movements of the asset price are modelled by the jump-diffusion process. This model implies that stock prices follow a geometric Brownian motion, but at random times can jump to a new level (upwards or downwards) and then again follow the geometric Brownian motion process. In the real market the jump effect can fade away as time passes by. The most appropriate model for this type of jumps is the shot noise process.

Remark 6.1. The general shot-noise process is

$$\lambda_t = \sum_{i=1}^{N_t} J^i_{t-s_i},$$

where J^i_t, $i = 1, 2, ...$ satisfies the SDE

$$dJ^i_t = a(t, J^i_t, Y_i)dt + b(t, J^i_t, Y_i)dB^i_t, \tag{41}$$

where a, b: $\mathrm{R}^+ \times \mathrm{R} \times \mathrm{R} \to \mathrm{R}$ are smooth functions, Y_i, $i = 1, 2, ...$ are iid random variables and B^i_t, $i = 1, 2, ...$ are independent Brownian motions. Also, N_t is a Cox process with intensity ρ and with jump times s_i.

The general shot-noise process is not a Markov process, only in one specific case this process will become a Markov process. If in (41) we set $b = 0$ and $a(t, J^i_t, Y_i) = Y_i h'(t)$, then $J^i_t = Y_i h(t)$. The shot-noise process λ_t is Markovian only if $h(t) = e^{-ct}$ as it was shown in the paper [13]. Exactly this type of process will be considered further in this thesis.

6.2 Definition of the shot-noise process and asset price model with shot noise

Consider the basic Black-Scholes stock price model:

$$\tilde{S}_t = \tilde{S}_0 e^{\int_0^t \mu_s ds - \frac{\nu^2}{2}t + \nu W_t},$$

where $\tilde{S}_t$ is stock price, μ_t is the drift process, ν is the volatility and W_t is Wiener process.

The shot noise effects can be added to the standard model in the following way [34]:

$$S_t = \tilde{S}_t \exp(\lambda_t) = \tilde{S}_0 e^{\int_0^t \mu_s ds - \frac{\nu^2}{2} t + \nu W_t} e^{\lambda_t}, \tag{42}$$

where S_t is stock price, μ_t is the drift that follows the Ornstein-Uhlenbeck process

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t,$$

where $\bar{\mu}$ is the mean of the drift, κ is the speed of mean reversion, ω is volatility of the drift and V_t is Wiener process independent of W_t. The process λ_t is the shot noise process, defined as follows:

$$\lambda_t = \lambda_0 e^{-\delta t} + \sum_{i=1}^{M_t} Y_i e^{-\delta(t-s_i)},$$

where

- λ_0 is the initial value of λ_t;
- $\{Y_i\}_{i=1,2,\ldots}$ is the sequence of iid random variables with distribution function $F(y)$ and $E(Y_i) = \mu_1$;
- $\{s_i\}_{i=1,2,\ldots}$ is the sequence representing the event times of a Poisson process M_t with constant intensity ρ;
- δ is the rate of exponential decay.

In the stock price model (42) the distribution of the random variables $\{Y_i\}_{i=1,2,\ldots}$ can be arbitrary, for example normal distribution, i.e. $Y_i \sim N(\mu_1, \sigma^2)$, $\mu_1, \ \sigma \in R$. Even if Y_i is negative for some i and even if λ_t is negative for some t, stock prices will not become negative.

If $\delta = 0$, then one obtains a pure jump-diffusion process. This case will be considered separately below.

6.3 Differential form of the shot noise process

Let us first derive the differential form of the shot noise process. Shot noise process can be rewritten in the following form:

$$\lambda_t = \lambda_0 e^{-\delta t} + \sum_{i=1}^{M_t} Y_i e^{-\delta(t-s_i)}$$

$$= \lambda_0 e^{-\delta t} + e^{-\delta t} \sum_{i=1}^{M_t} Y_i e^{\delta s_i},$$

setting $H(s_i, y_i) := Y_i e^{\delta s_i}$ one obtains the following:

$$\lambda_t = e^{-\delta t}(\lambda_0 + \sum_{i, s_i \leq t} H(s_i, y_i))$$

$$= e^{-\delta t}(\lambda_0 + \int_0^t \int_E H(s, y) N(ds, dy)),$$

where $N(ds, dy)$ is a Poisson random measure with $E[N(ds, dy)] = \rho dt F(dy)$, where ρ is the intensity of the Poisson process M_t and F is the distribution function of the jumps Y_i. The differential form of λ_t is obtained by differentiating λ_t as a product of two functions:

$$d\lambda_t = -e^{-\delta t}\delta\lambda_0 dt - \delta e^{-\delta t}\int_0^t \int_E H(s, y)N(ds, dy)dt + e^{-\delta t}\int_E H(t, y)N(dt, dy)$$

$$= -\delta\lambda_t dt + e^{-\delta t}\int_E y e^{\delta t} N(dt, dy)$$

$$= -\delta\lambda_t dt + \int_E y N(dt, dy).$$

Thus, we obtain the following result:

Lemma 6.2. *The differential form of the shot noise process λ_t is as follows:*

$$d\lambda_t = -\delta\lambda_t dt + \int_E y N(dt, dy).$$

Instead of the shot noise process, one can add compound Poisson process as a noise term, i.e.

$$J_t = \sum_{i=1}^{M_t} Y_i = \sum_{i, s_i \leq t} Y_i = \int_0^t \int_E H(s, y) N(ds, dy),$$

where $H(s_i, Y_i) = Y_i$. The differential form of the compound Poisson process is as follows:

$$dJ_t = \int_E H(s, y) N(ds, dy) = \int_E y N(ds, dy).$$

The case of the Compound Poisson process as a noise term will be considered separately below.

Next we apply the Ito Theorem 2.2 to the asset price process (function $f(x) = \exp(x)$). The asset price process is as follows:

$$S_t = S_0 e^{L_t} = S_0 e^{\int_0^t \mu_s ds - \frac{\nu^2}{2}t + \nu W_t + \lambda_t},$$

where
$$dL_t = \left(\mu_t - \frac{\nu^2}{2} - \delta\lambda_t\right) dt + \nu dW_t + \int_E yN(dt, dy).$$

Thus,

$$\begin{aligned} dS_t = S_0 e^{L_{t-}} dL_t^c + \frac{1}{2} S_0 e^{L_{t-}} \nu^2 dt + \int_E (S_0 e^{L_{t-}+y} - S_0 e^{L_{t-}}) N(dt, dy) \\ = S_{t-} \left[(\mu_t - \delta\lambda_t)dt + \nu dW_t + \int_E (e^y - 1)N(dt, dy)\right] \\ = S_{t-} \left[\mu_t dt + \nu dW_t + d\lambda_t + \int_E (e^y - 1 - y)N(dt, dy)\right] \end{aligned}$$

If on average the values of jumps Y_i are less than some small ε, then the last term is quite small and can be neglected. Then the differential form of the equation (42) is given by the following SDE:

$$dS_t \approx \mu_t S_{t-} dt + \nu S_{t-} dW_t + S_{t-} d\lambda_t. \tag{43}$$

Another way to add shot noise jumps to the standard stock price equation is considered in the paper [1]:

$$\begin{aligned} S_t = S_0 \exp\Big(\Big(\mu - \frac{\nu^2}{2}\Big)t + \nu W_t\Big) \prod_{i=1}^{M_t} \Big(1 + Y_i e^{-\delta(t-S_i)}\Big) \\ = S_0 \exp\Big(\Big(\mu - \frac{\nu^2}{2}\Big)t + \nu W_t\Big) \exp\Big(\sum_{i=1}^{M_t} \ln(1 + Y_i e^{-\delta(t-S_i)})\Big). \end{aligned} \tag{44}$$

The shot noise process in this case is as follows:

$$\lambda_t^1 = \sum_{i=1}^{M_t} \ln(1 + Y_i e^{-\delta(t-S_i)}).$$

In this case one needs $Y_i > -1$; for the purposes of simulations one can take log-normal distribution for Y_i.

To sum up this subsection, we consider the asset price model (42) driven by shot-noise process. Assuming, that the jumps of the shot-noise process are small, we will further consider the approximation (43) of the asset price model (42). The approximated model is given by the following pair of equations:

$$dS_t \approx \mu_t S_t dt + \nu S_t dW_t + S_{t-} d\lambda_t, \tag{45}$$

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t, \tag{46}$$

where S_t is stock price, μ_t is the drift process of the stock price, ν is the volatility of the stock price, $d\lambda_t$ is the shot noise process and W_t is Wiener process. In the second equation $\bar{\mu}$ is the mean of the drift, ω is volatility of the drift and V_t is Wiener process independent of W_t.

The filtering problem consists of approximating the value of μ_t for each t taking into account the values of S_s, $s \leq t$, i.e.

$$\hat{\mu_t} = E[\mu_t | F_t^S],$$

where F_t^S is the filtration based on the stock prices up to time t.

The filteing problem with shot-noise process is infinite dimensional and cannot be solved explicitly. One way to solve this problem is to approximate shot noise λ_t with the Brownian motion and then to apply Kalman filtering.

6.4 Approximating shot noise process

According to the paper [7] the shot noise process can be approximated by Brownian motion.

Let the shot noise look as follows:

$$\lambda_t = \lambda_0 e^{-\delta t} + \sum_{i=1}^{M_t} Y_i e^{-\delta(t-s_i)}, \tag{47}$$

where

- λ_0 is the initial value of λ_t;
- $\{Y_i\}_{i=1,2,\dots}$ is the sequence of iid random variables with distribution function $F(y)$, $E(Y_i) = \mu_1$ and $E(Y_i^2) = \mu_2$;
- $\{s_i\}_{i=1,2,\dots}$ is the sequence representing the event times of a Poisson process M_t with constant intensity ρ;
- δ is the rate of exponential decay.

The expectation of the shot noise process λ_t, assuming that λ_0 is known, is as follows from [7]:

$$E(\lambda_t) = \frac{\mu_1 \rho}{\delta} + \left(\lambda_0 - \frac{\mu_1 \rho}{\delta}\right) e^{-\delta t} \to \frac{\mu_1 \rho}{\delta} \text{ as } t \to \infty, \tag{48}$$

and, moreover, if the initial value λ_0 equals $\mu_1\rho/\delta$, then we have a stationary case and the mean value $E(\lambda_t)$ will be equal to $\mu_1\rho/\delta$ and will not depend on time t.

The variance of the shot noise process λ_t is as in [7]:

$$\mathrm{Var}(\lambda_t) = \frac{\mu_2 \rho}{2\delta}(1 - e^{-2\delta t}) \to \frac{\mu_2 \rho}{2\delta} \text{ as } t \to \infty. \tag{49}$$

Consider the following linear transformation:

$$Z_t^{(\rho)} = \frac{\lambda_t - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}. \tag{50}$$

The central result of the paper [7] is that $Z_t^{(\rho)}$ converges in law to some Z_t that is normally distributed.

For the proof of this main result, the authors in [7] additionally define the following:

$$V_t^{(\rho)} = \frac{J_t - \mu_1\rho t}{\sqrt{\mu_2\rho/2\delta}},$$

where $J_t = \sum_{i=1}^{M_t} Y_i$, and prove the convergence of $V_t^{(\rho)}$ to the Brownian motion. The following lemma states this convergence result:

Lemma 6.3. *(Lemma 2, Dassios, Jang) For $\rho \to \infty$,*

$$V_t^{(p)} \to \sqrt{2\delta}B_t,$$

where B_t is Brownian motion.

This lemma is an auxilarry result which is used in the proof of the main theorem that shows the convergence of

$$Z_t^{(\rho)} = \frac{\lambda_t - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}$$

to a Guassian random variable. The proof of the main result we provide in details.

Theorem 6.4. *(Theorem 2, Dassios, Jang) Assume that $\rho \to \infty$ and that λ_0 is a random variable independent of everything else, such that $(\lambda_0 - (\mu_1\rho/\delta))(\mu_1\rho/2\delta)^{-1/2}$ converges in distribution to Z_0. Then $Z_t^{(\rho)}$ converges in law to Z_t, where*

$$dZ_t = -\delta Z_t dt + \sqrt{2\delta}dB_t, \tag{51}$$

where B_t is standard Brownian motion.

Proof. (Dassios, Jang) Rewrite Z_t^{ρ} as follows:

$$Z_t^{(\rho} = \frac{\lambda_t - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}$$

$$= \frac{\lambda_0 e^{-\delta t} + \sum_{i=1}^{M_t} Y_i e^{-\delta(t-S_i)} - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}$$

$$= \frac{\lambda_0 - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}} e^{-\delta t} - \frac{\mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}(1 - e^{-\delta t}) + e^{-\delta t} \int_0^t \int_E \frac{y e^{\delta s} N(ds, dy)}{\sqrt{\mu_2\rho/2\delta}}$$

$$= \frac{\lambda_0 - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}} e^{-\delta t} - \frac{\mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}(1 - e^{-\delta t}) + e^{-\delta t} \int_0^t e^{\delta s} \int_E \frac{y N(ds, dy)}{\sqrt{\mu_2\rho/2\delta}}$$

$$= \frac{\lambda_0 - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}} e^{-\delta t} - \frac{\mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}(1 - e^{-\delta t}) + e^{-\delta t} \int_0^t e^{\delta s} \frac{dJ_s}{\sqrt{\mu_2\rho/2\delta}}$$

$$= \frac{\lambda_0 - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}} e^{-\delta t} - \frac{\mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}}(1 - e^{-\delta t}) + \frac{J_t}{\sqrt{\mu_2\rho/2\delta}} - \delta \int_0^t e^{-\delta(t-u)} \frac{dJ_u}{\sqrt{\mu_2\rho/2\delta}}.$$

The last expression was obtained using integration by parts formula. Since $\delta \int_0^t u e^{-\delta(t-u)} du = t - (1 - e^{-\delta t})/\delta$, one can substitute in the expression above the following formula $-(1 - e^{-\delta t})/\delta = t - \delta \int_0^t u e^{-\delta(t-u)} du$ and obtain the expression below:

$$Z_t^{(\rho)} = \frac{\lambda_0 - \mu_1\rho/\delta}{\sqrt{\mu_2\rho/2\delta}} e^{-\delta t} + \frac{J_t - \mu_1\rho t}{\sqrt{\mu_2\rho/2\delta}} - \delta \int_0^t e^{-\delta(t-u)} \frac{J_u - \mu_1\rho u}{\sqrt{\mu_2\rho/2\delta}} du,$$

Therefore, Z_t^ρ converge to

$$Z_t = Z_0 e^{-\delta t} + \sqrt{2\delta}(B_t - \delta \int_0^t e^{-\delta(t-s)} B_s ds)$$

$$= Z_0 e^{-\delta t} + \sqrt{2\delta} \int_0^t e^{-\delta(t-s)} dB_s,$$

where B_t is standard Brownian motion. This finishes the proof. □

This theorem implies that Z_t follows an Ornstein-Uhlenbeck process and is normally distributed with mean $E(Z_t) = Z_0 e^{-\delta t} \to 0$ as $t \to \infty$ and variance $Var(Z_t) = 1 - e^{-2\delta t} \to 1$ as $t \to \infty$. If $\lambda_0 = \mu_1\rho/\delta$, then $Z_0 = 0$ and, therefore, $E(Z_t) = 0$

Following the linear transformation (50) the shot-noise process λ_t has the following form:

$$\lambda_t = \frac{\mu_1\rho}{\delta} + Z_t^\rho \sqrt{\frac{\mu_2\rho}{2\delta}}.$$

Define $\hat{\lambda}_t$ as Gaussian approximation of λ_t as follows:

$$\hat{\lambda}_t = \frac{\mu_1\rho}{\delta} + Z_t \sqrt{\frac{\mu_2\rho}{2\delta}}.$$

Consider the mean of the approximation $\hat{\lambda}_t$:

$$E(\hat{\lambda}_t) = \frac{\mu_1\rho}{\delta} + Z_0 e^{-\delta t}\sqrt{\frac{\mu_2\rho}{2\delta}},$$

and, taking into account that for the case when $\lambda_0 = \mu_1\rho/\delta$ (therefore $Z_0 = 0$),

$$E(\hat{\lambda}_t) = \frac{\mu_1\rho}{\delta}$$

one sees that the mean value of the approximation $\hat{\lambda}_t$ is consistent with the mean value (48) of the shot noise λ_t.

The variance of the approximation $\hat{\lambda}_t$ is

$$\mathrm{Var}(\hat{\lambda}_t) = \frac{\mu_2\rho}{2\delta}\mathrm{Var}(Z_t) = \frac{\mu_2\rho}{2\delta}(1 - e^{-2\delta t}),$$

which corresponds to the variance (49) of the shot noise process λ_t.

6.5 Numerical experiments

In this subsection we simulate the model (45)-(46). Choose the following parameters for this model:

$$S_t = \tilde{S}_0 e^{(\int_0^t \mu_s ds - \frac{0.3^2}{2})t + 0.3W_t} e^{\lambda_t}, \tag{52}$$

$$d\mu_t = 0.5(0.5 - \mu_t)dt + 0.3dV_t, \tag{53}$$

The rate of the Poisson process M_t used to generate the shot noise λ_t is equal to $\rho = 1.6$. In order to apply Kalman filtering one has to approximate the shot noise process λ_t by $\hat{\lambda}_t$. Applying the results from the previous section,

$$\frac{dS_t}{S_t} \approx \mu_t dt + \nu dW_t + d\hat{\lambda}_t = \mu_t dt + \nu dW_t + dZ_t\sqrt{\frac{\mu_2\rho}{2\delta}}$$

$$= \left(\mu_t - \delta\sqrt{\frac{\mu_2\rho}{2\delta}}Z_t\right)dt + \nu dW_t + \sqrt{\mu_2\rho}dB_t,$$

where B_t is a standard Brownian motion and

$$dZ_t = -\delta Z_t dt + \sqrt{2\delta}dB_t.$$

Consider the term $c := \delta\sqrt{\mu_2\rho/2\delta}$. As we have assumed before, the jumps y of the shot noise process λ_t are 'small', therefore, the value of c will also be small, because μ_2 is 'small'. Also, cZ_t is 'small' comparing to μ_t and, therefore, the term cZ_t can be neglected. The model to be considered for the numerical experiments is as follows:

$$\frac{dS_t}{S_t} \approx \mu_t dt + \nu dW_t + \sqrt{\mu_2 \rho} dB_t,$$
$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t.$$

The discrete-time version of the model (45)-(46) looks as follows:

- signal (unobservable)

$$\mu_{k+1} = (1 - \kappa \Delta t)\mu_k + \kappa \bar{\mu} \Delta t + \omega \sqrt{\Delta t} V_k, \tag{54}$$

- observation

$$\frac{\Delta S_k}{S_k} = y_k = \mu_k \Delta t + \nu \sqrt{\Delta t} W_k + \sqrt{\Delta t} B_k \sqrt{\frac{\mu_2 \rho}{2\delta}}. \tag{55}$$

For the numerical experiments we simulate the origianl asset price model (the one driven by Brownian motion and the shot-noise process), but for estimating the drift process of the asset price we use the approximated model (the one where the shot-noise process is approximated by Brownian motion). The results of discrete Kalman filtering of the asset price model with approximated shot noise process are shown below:

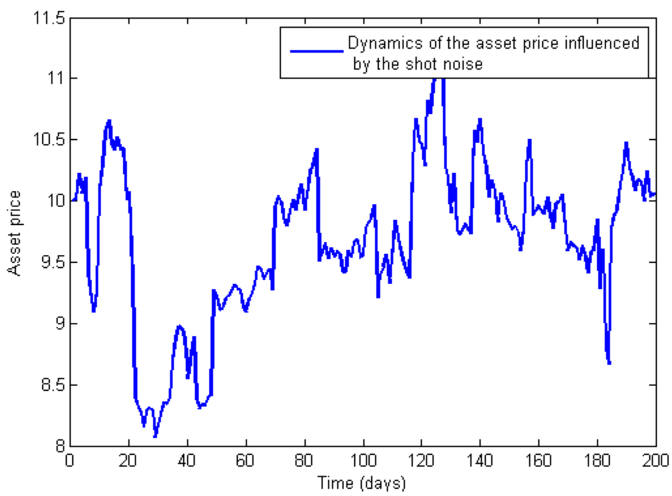

Figure 1: Stock price behavior with the effects of the shot noise (first example)

For the filtering example, shown in Figures 1 and 2, the mean error between the true and the filtered values of the drift process is equal to 0.0330.

In the second example, shown in Figures 3 and 4, the mean error between the true and the filtered values of the drift process is equal to 0.0166.

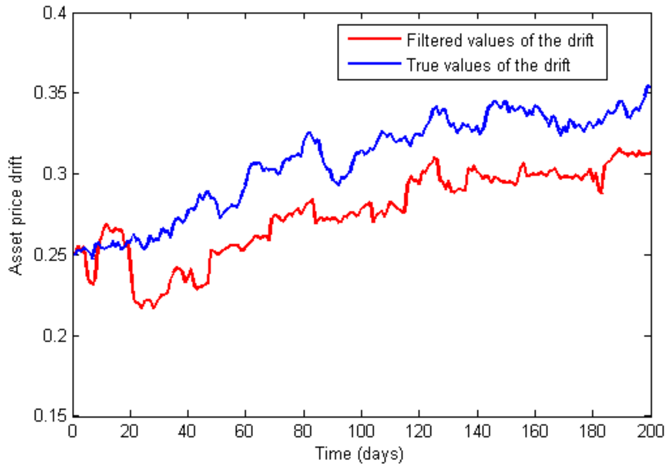

Figure 2: True and filtered behavior of the drift process μ_t (first example)

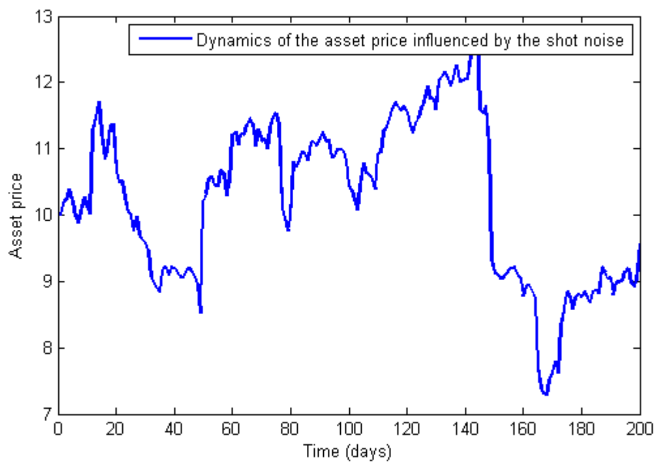

Figure 3: Stock price behavior with the effects of the shot noise (second example)

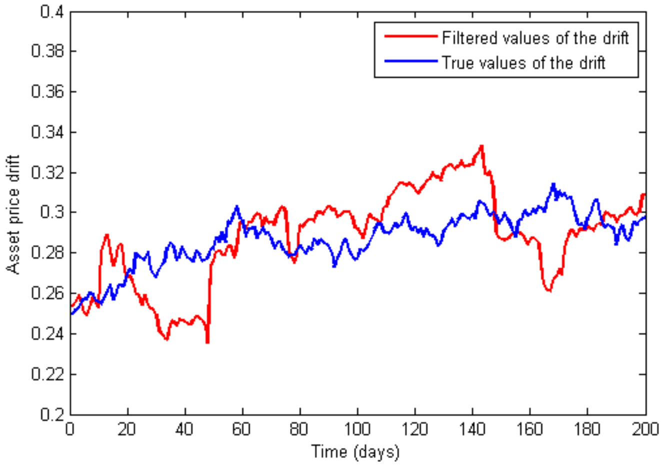

Figure 4: True and filtered behavior of the drift process μ_t (second example)

6.6 Filtering of multi-dimensional asset price model with shot noise process as jump part (first version) and numerical experiments

A model for n assets, each driven by the sum of the same Wiener processes, is considered. Therefore, the assets are correlated. The shot-noise process is different for each asset price.

Assume that there are n assets $S_i,\ i = 1, ..., n$ whose prices satisfy the following model:

$$S_i(t) = S_i(0) \exp\left(\int_0^t \mu_i(s)ds - \frac{1}{2} \sum_{j=1}^{n} \nu_{ij}^2 t + \sum_{j=1}^{n} \nu_{ij} W_j(t) \right) \exp\left(\lambda_i(t)\right),$$

or the approximated model (under assumption that the jumps of each shot-noise process $\lambda_i(t)$ are 'small')

$$dS_i(t) \approx \mu_i(t)S_i(t)dt + S_i(t) \sum_{j=1}^{n} \nu_{ij} dW_j(t) + S_i(t)d\lambda_i(t),$$

where $W_i(t), i = 1, ..., n$ are independent Wiener processes, ν_{ij} are diffusion coefficients, $\lambda_i(t)$ is the shot noise process for the corresponding asset price $S_i(t)$. The drift processes $\mu_i(t)$ follow the SDEs below:

$$d\mu_i(t) = \kappa_i(\bar{\mu}_i - \mu_i(t))dt + \omega_i dV_i(t),$$

where ω_i is volatility, $\bar{\mu}_i$ is the mean, κ_i is the mean-reversion coefficient and $V_i(t)$ is Wiener process independent of $W_i(t), i = 1, ..., n$.

In order to solve this filtering problem, one has to approximate shot noise processes with Brownian motion:

$$dS_i(t) \approx \mu_i(t)S_i(t)dt + S_i(t)\sum_{j=1}^{n}\nu_{ij}dW_j(t) + S_i(t)d\lambda_i(t) =$$

$$= S_i(t)\Big(\Big(\mu_i(t) - \delta_i\sqrt{\frac{\mu_2^{(i)}\rho_i}{2\delta_i}}Z_i(t)\Big)dt + \sum_{j=1}^{n}\nu_{ij}dW_j(t) + \sqrt{\mu_2^{(i)}\rho_i}dB_i(t)\Big),$$

where

$$dZ_i(t) = -\delta_i Z_i(t)dt + \sqrt{2\delta_i}dB_i(t),$$

where $B_i(t)$ is a standard Brownian motion for $i = 1...n$ and δ_i is the rate of exponential decay for the shot-noise process $\lambda_i(t)$ and $\mu_2^{(i)}$ is the second-order raw moment of the distribution of the jumps of the shot-noise process $\lambda_i(t)$. The jumps of shot noise-processes $\lambda_i(t)$ for $i = 1...n$ are assumed to be 'small', therefore, $c_i := \delta_i\sqrt{\mu_2^{(i)}\rho_i/2\delta_i}$ is 'small' and $c_iZ_i(t)$ can be neglected as it was done in Subsection 6.5.

This multi-dimensional asset price model has to be discretized in order to apply Kalman filtering as described in section 1.4.1.

Let $n = 2$, this means we have two assets which satisfy the following SDEs:

$$S_1(t) = S_1(0)\exp\Big(\int_0^t \mu_1(s)ds - \frac{1}{2}(0.3)^2t - \frac{1}{2}(0.2)^2t$$
$$+0.3W_1(t) + 0.2W_2(t) + \lambda_1(t)\Big),$$
$$d\mu_1(t) = 0.5(0.5 - \mu_1(t))dt + 0.3dV_1(t)$$

and

$$S_2(t) = S_2(0)\exp\Big(\int_0^t \mu_2(s)ds - \frac{1}{2}(0.2)^2t - \frac{1}{2}(0.4)^2t$$
$$+0.2W_1(t) + 0.4W_2(t) + \lambda_2(t)\Big),$$
$$d\mu_2(t) = 0.7(0.3 - \mu_2(t))dt + 0.3dV_2(t),$$

where shot noise process $\lambda_i, i = 1, 2$ is given by (47). For $i = 1$: $\delta_1 = 1.5$, intensity of the Poisson process $M_t^{(1)}$ is $\rho_1 = 1.6$ and $Y \sim N(0, 0.02)$. For $i = 2$: $\delta_2 = 2$, intensity of the Poisson process $M_t^{(2)}$ is $\rho_2 = 1.6$ and $Y \sim N(0, 0.03)$.

The results of numerical simulations are shown in the figures below. The simulations of the price of the first asset and its drift process (true and filtered) are shown in Figures 5 and 6. Figures 7 and 8 show the the simulations of the price of the second asset and its drift process (true and filtered). Mean error for the first filter equals 0.0091 and for the second 0.0090.

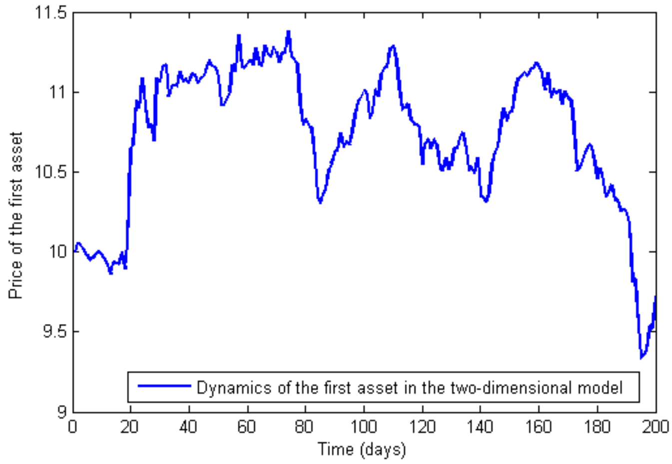

Figure 5: Price behavior of the first stock with the effects of the shot noise in the two dimensional model

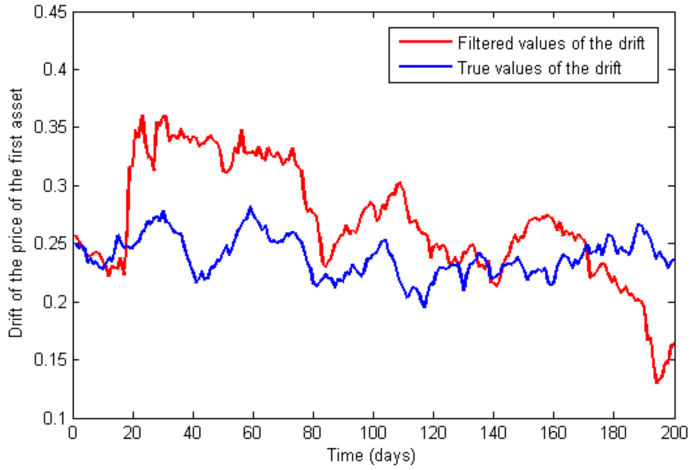

Figure 6: True and filtered behavior of the drift process μ_t of the first stock in the two dimensional model

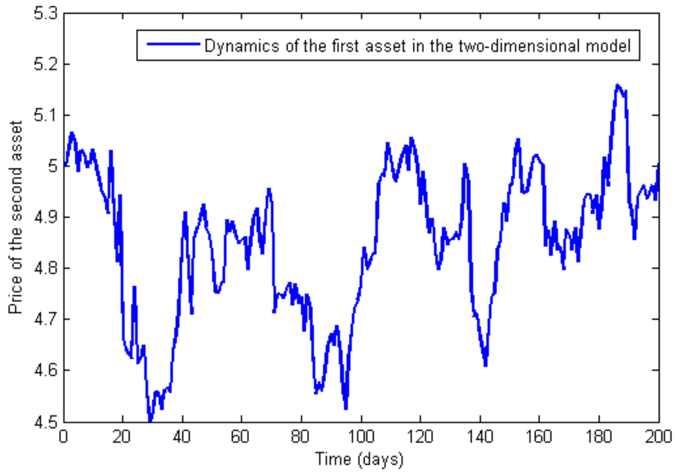

Figure 7: Price behavior of the second stock with the effects of the shot noise in the two dimensional model

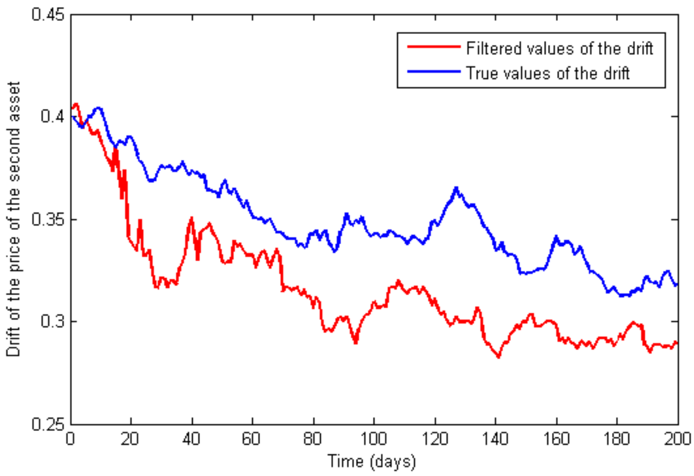

Figure 8: True and filtered behavior of the drift process μ_t of the second stock in the two dimensional model

6.7 Filtering of multi-dimensional asset price model with shot noise process as jump part (second version) and numerical experiments

Unlike the asset price model of the previous subsection, here we assume that there are n shot-noise process $\lambda_j(t)$, $j = 1, ..., n$, all of which are present in each asset price equation. In other words, each asset price is influenced by the sum of the same shot-noise processes.

Assume that there are n assets S_i, $i = 1, ..., n$, whose prices satisfy the following model:

$$S_i(t) = S_i(0) \exp\left(\int_0^t \mu_i(s)ds - \frac{1}{2}\nu_i^2 t + \nu_i W_i(t)\right) \prod_{i=1}^{n} \exp\left(\lambda_i(t)\right),$$

or the approximated model (under assumption that the jumps of each shot-noise process $\lambda_i(t)$ are 'small')

$$dS_i(t) \approx \mu_i(t)S_i(t)dt + S_i(t)\nu_i dW_i(t) + S_i(t)\sum_{j=1}^{n} d\lambda_j(t),$$

where $W_i(t), i = 1, ..., n$ are independent Wiener processes each for the corresponding stock price $S_i(t)$, ν_{ij} are diffusion coefficients, $\lambda_i(t), i = 1, ...,$ are the shot noise processes. The drift processes $\mu_i(t)$ follow the SDEs below:

$$d\mu_i(t) = \kappa_i(\bar{\mu}_i - \mu_i(t))dt + \omega_i dV_i(t),$$

where ω_i is volatility, κ_i is mean reversion coefficient and $V_i(t)$ is Wiener process independent of $W_i(t), i = 1, ..., n$.

As before, it is proposed to approximate the shot-noise process with Brownian motion:

$$dS_i(t) \approx \mu_i(t)S_i(t)dt + S_i(t)\nu_i dW_i(t) + S_i(t)\sum_{j=1}^{n} d\lambda_j(t)) =$$

$$= S_i(t)\Big(\Big(\mu_i(t) - \sum_{j=1}^{n} \delta_i \sqrt{\frac{\mu_2^{(i)}\rho_i}{2\delta_i}} Z_i(t)\Big)dt + \nu_i dW_i(t) + \sum_{j=1}^{n} \sqrt{\mu_2^{(i)}\rho_i} dB_i(t)\Big),$$

where

$$dZ_i(t) = -\delta_i Z_i(t)dt + \sqrt{2\delta_i} dB_i(t),$$

where $B_i(t)$ is a standard Brownian motion for $i = 1...n$ and δ_i is the rate of exponential decay for the shot-noise process $\lambda_i(t)$ and $\mu_2^{(i)}$ is the second-order raw moment of the distribution of the jumps of the shot-noise process $\lambda_i(t)$. The jumps of shot noise-processes $\lambda_i(t)$ for $i = 1...n$ are assumed to be 'small', therefore, $c_i := \delta_i \sqrt{\mu_2^{(i)}\rho_i/2\delta_i}$ is 'small' and $c_i Z_i(t)$ can be neglected as it was done in 6.5.

One discretizes this multi-dimensional asset price model and solves separately n filtering problems (for each of the n assets) using Kalman filtering.

Let $n = 2$, and let the SDEs look as follows:

$$S_1(t) = S_1(0) \exp\left(\int_0^t \mu_1(s)ds - \frac{1}{2}(0.3)^2 t + 0.3W_1(t) + \lambda_1(t) + \lambda_2(t)\right),$$

$$d\mu_1(t) = 0.5(0.5 - \mu_1(t))dt + 0.3dV_1(t)$$

and

$$S_2(t) = S_2(0) \exp\left(\int_0^t \mu_2(s)ds - \frac{1}{2}(0.2)^2 t + 0.2W_2(t) + \lambda_1(t) + \lambda_2(t)\right),$$

$$d\mu_2(t) = 0.7(0.5 - \mu_2(t))dt + 0.3dV_2(t),$$

where shot noise process $\lambda_i, 1 = 1, 2$ is given by (47). For $i = 1$: $\delta_1 = 1.5$, intensity of the Poisson process $M_t^{(1)}$ is $\rho_1 = 1.2$ and $Y \sim N(0, 0.008)$. For $i = 2$: $\delta_2 = 2$, intensity of the Poisson process $M_t^{(2)}$ is $\rho_2 = 1.6$ and $Y \sim N(0, 0.007)$.

The first simulation example consists of the following figures: Figure 9 shows the behaviour of both asset prices $S_1(t)$ and $S_2(t)$; Figure 10 shows the

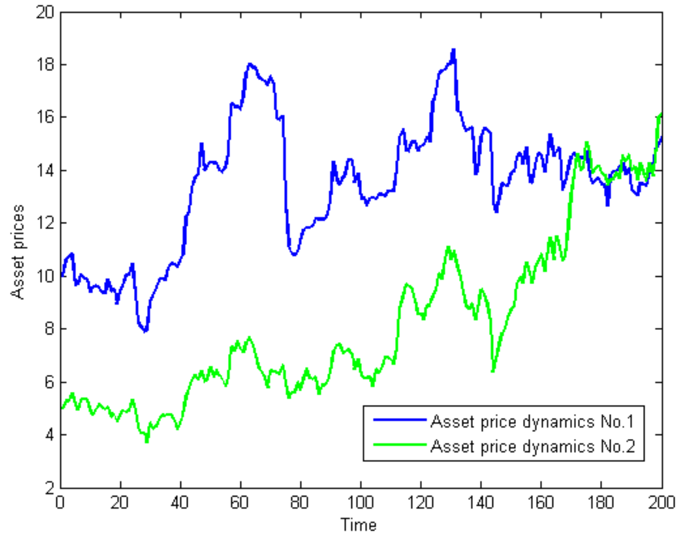

Figure 9: Two dynamics of asset prices influenced by the mixture of shot noise processes (first example)

true and filtered drift processes of the first asset $S_1(t)$ and Figure 11 shows the true and filtered drift processes of the second asset $S_2(t)$.

Observing two such asset price dynamics one can visually determine the times of jumps of shot-noise by comparing the similar jump parts of the asset price processes; jumps can be eliminated if needed. In Figure 9 we see that one of the shot-noise processes jumped at time $T = 60$ and the second jump is at time $T = 130$ and a small peak is present at time $T = 100$. All other jumps in both processes, probably, come from Brownian motions and not from shot-noise processes.

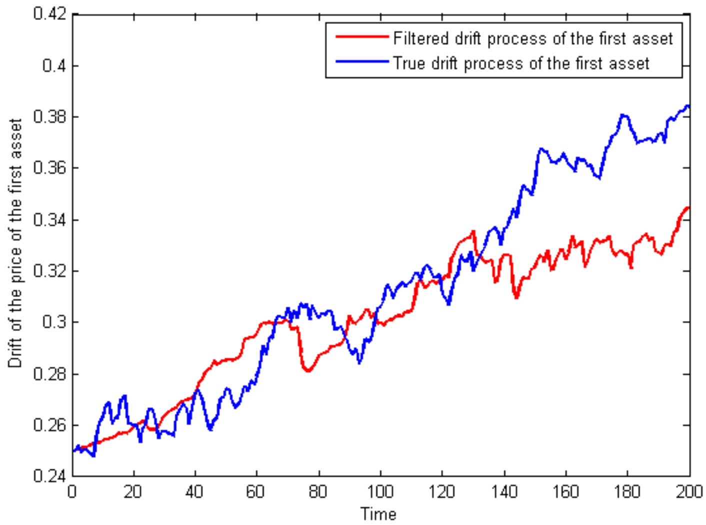

Figure 10: True and filtered behavior of the drift process μ_t of the first asset (first example)

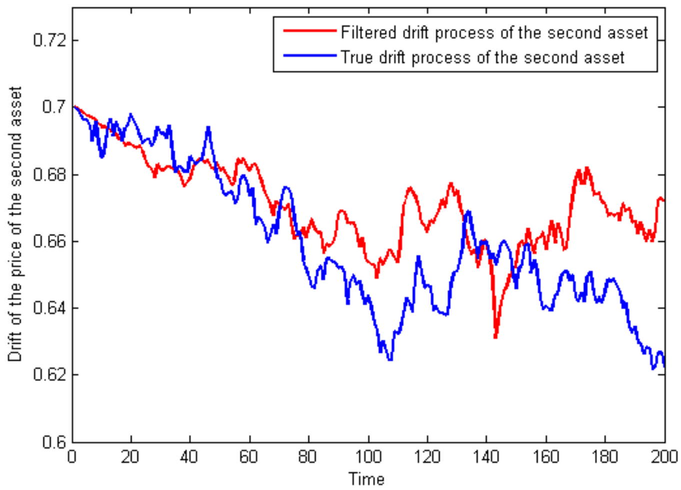

Figure 11: True and filtered behavior of the drift process μ_t of the second asset (first example)

The second numerical example for the asset price model with the same parameters is shown below. Figures 12, 13 and 14 show the behaviour of both asset prices, and the corresponding drift processes (true and filtered ones).

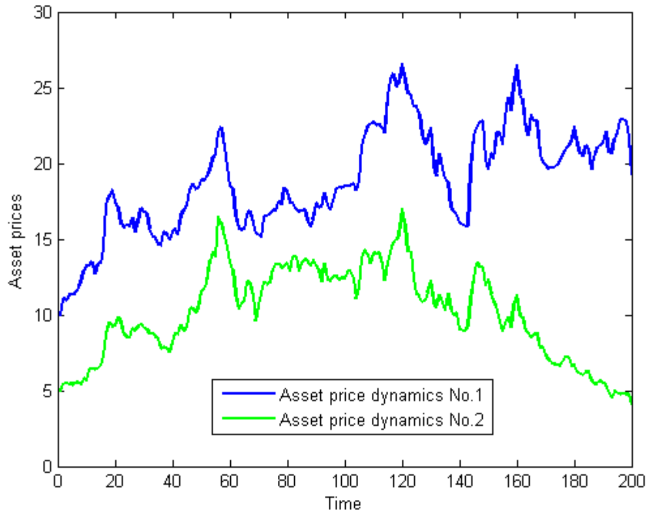

Figure 12: Two dynamics of asset prices influenced by the mixture of shot noise processes (second example)

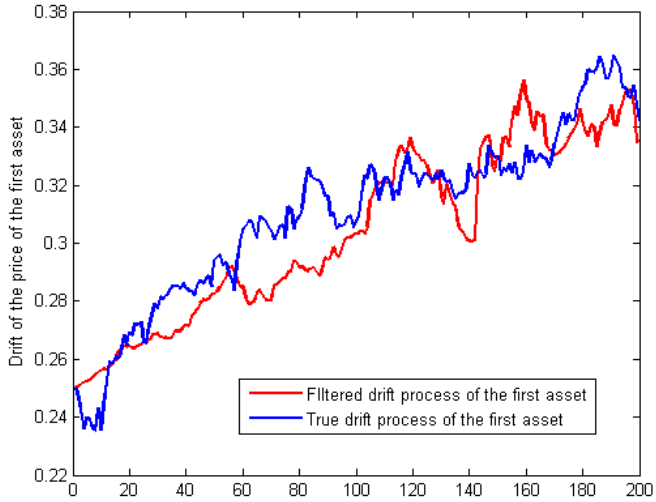

Figure 13: True and filtered behavior of the drift process μ_t of the first asset (second example)

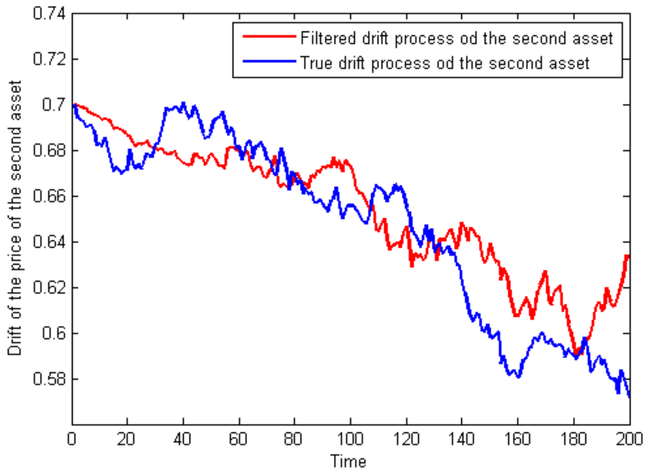

Figure 14: True and filtered behavior of the drift process μ_t of the second asset (second example)

7 Portfolio optimization for shot-noise driven processes

This section starts with the basic framework described in the paper [6]. In this work a portfolio optimization problem is considered for the case when both unobservable signal (drift process) and observation (asset price) processes follow the SDEs with linear and constant coefficients and are driven by the sum of Brownian motions.

7.1 Portfolio optimization for an asset price model with approximation of shot noise process

Consider a financial market with one riskfree asset (bank account) and one risky asset. The riskfree asset pays a constant interest rate. The risky asset is modelled as follows:

$$S_t = \tilde{S}_0 e^{\int_0^t \mu_s ds - \frac{\nu^2}{2}t + \nu W_t} e^{\lambda_t}, \tag{56}$$

where S_t is stock price, λ_t is a shot-noise process, ν is the volatility of the stock price, W_t is Wiener process, μ_t is the drift that follows the Ornstein-Uhlenbeck process

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t,$$

where $\bar{\mu}$ is the mean of the drift, κ is the speed of mean reversion, ω is volatility of the drift and V_t is Wiener process independent of W_t.

We assume that the jumps of the shot-noise process are 'small', therefore, the approximation idea from subsection 6.3 can applied, and, thus, the risky asset (56) can be approximated in the following way:

$$dS_t \approx \mu_t S_t dt + \nu S_t dW_t + S_t d\lambda_t, \tag{57}$$

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t. \tag{58}$$

According to [6], we assume that the riskfree rate is equal to zero.

Recall the definition of the shot-noise λ_t process:

$$\lambda_t = \lambda_0 e^{-\delta t} + \sum_{i=1}^{M_t} Y_i e^{-\delta(t-s_i)},$$

where

- λ_0 is the initial value of λ_t;

- $\{Y_i\}_{i=1,2,\ldots}$ is the sequence of iid random variables with distribution function $F(y)$ and $E(Y_i) = \mu_1$, $E(Y_i^2) = \mu_2$;
- $\{s_i\}_{i=1,2,\ldots}$ is the sequence representing the event times of a Poisson process M_t with constant intensity ρ;
- δ is the rate of exponential decay.

As shown above, shot-noise processes can be approximated by Brownian motion as follows:

$$\frac{dS_t}{S_t} \approx \mu_t dt + \nu dW_t + d\hat{\lambda}_t = \mu_t dt + \nu dW_t + dZ_t\sqrt{\frac{\mu_2\rho}{2\delta}}$$

$$= \Big(\mu_t - \delta\sqrt{\frac{\mu_2\rho}{2\delta}}Z_t\Big)dt + \nu dW_t + \sqrt{\mu_2\rho}dB_t,$$

where $dZ_t = -\delta Z_t dt + \sqrt{2\delta}dB_t$, B_t is a standard Brownian motion.

Therefore, the asset price model for the case, when the shot-noise process is approximated by Brownian motion (taking into account, that the jumps of the shot-noise process λ_t are 'small'), looks as follows:

$$\frac{dS_t}{S_t} = \Big(\mu_t - \delta\sqrt{\frac{\mu_2\rho}{2\delta}}Z_t\Big)dt + \nu dW_t + \sqrt{\mu_2\rho}dB_t, \tag{59}$$

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t, \tag{60}$$

$$dZ_t = -\delta Z_t dt + \sqrt{2\delta}dB_t. \tag{61}$$

This asset price model is used to construct a portfolio that consists of a bank account and the asset. The interest rate of the bank account is constant and can be set to 0 without loss of generality. Then the portfolio wealth X_t follows the process:

$$dX_t = X_t\pi_t\frac{dS_t}{S_t},$$

where π_t denotes the fraction of wealth invested in the risky asset. Initial capital is x. The aim is to select the trading strategy π_t in order to maximize the expected utility $E[U(X_T)]$ of the portfolio wealth at terminal time T, where $U(x)$ is the utility function. At each time moment t the investor chooses how much of the wealth will be invested in the risky asset.

Define the value function

$$J(t,\mu,z,x) = \sup_{\pi} E[U(X_T^{\pi})|X_t = x, Z_t = z, \mu_t = \mu].$$

In order to solve this portfolio optimization problem one uses the HJB equation, considered in section 4. The HJB equation for the stated portfolio optimization problem is as follows (μ stands for μ_t, z stands for Z_t and x stands for X_t):

$$\max_{\pi}\Big\{\frac{\partial}{\partial t}J(t,\mu,z,x)+\frac{\partial}{\partial x}J(t,\mu,z,x)\pi(\mu-cz)x+\frac{1}{2}\frac{\partial^2}{\partial x^2}J(t,\mu,z,x)\pi^2\Sigma_S x^2$$
$$+\frac{1}{2}\frac{\partial^2}{\partial \mu^2}J(t,\mu,z,x)\Sigma_\mu+\frac{\partial}{\partial x\partial z}J(t,\mu,z,x)\pi Rx+\frac{\partial}{\partial \mu}J(t,\mu,z,x)\kappa\bar{\mu}$$
$$-\frac{\partial}{\partial \mu}J(t,\mu,z,x)\kappa\mu-\frac{\partial}{\partial z}J(t,\mu,z,x)\delta z+\frac{1}{2}\frac{\partial^2}{\partial z^2}J(t,\mu,z,x)\Sigma_z\Big\}=0,$$

where $c=\delta\sqrt{\mu_2\rho/2\delta}$, $\Sigma_S=\nu^2+\mu_2\rho$, $\Sigma_\mu=\omega^2$, $\Sigma_z=2\delta$, $R=\sqrt{2\delta\mu_2\rho}$.

Differentiating with respect to π one obtains the following formula for the optimal portfolio strategy $\pi^*=\pi^*(\mu,z)$:

$$\pi^*(t,\mu,z,x)=\frac{(\mu-cz)\frac{\partial}{\partial x}J(t,\mu,z,x)-R\frac{\partial}{\partial x\partial z}J(t,\mu,z,x)}{x\Sigma_S\frac{\partial^2}{\partial x^2}J(t,\mu,z,x)}.$$

If one sets the specific form of the value function $J(t,\mu,z,x)$, then the optimal trading strategy can be evaluated assuming that the investor possesses the full information i.e. observes the drift μ_t and knows the values of Z_t. If one does not observe these values, then portfolio optimization problem for such an asset price model is difficult to solve, since no explicit solutions are known for the filtering problem given discrete observations.

One can rewrite the model (59)-(60)-(61) in a simpler form. The coefficients κ and δ of this model represent the speed with which the loss of information occures in both drift and shot-noise processes. That is why it would be reasonable to set $\kappa=\delta$. Therefore, the idea is to transform the drift process μ_t by substracting cZ_t, the differential form of μ_t-cZ_t is as follows:

$$d(\mu_t-cZ_t)=\kappa(\bar{\mu}-(\mu_t-cZ_t))dt+\omega dV_t-c\sqrt{2k}dB_t.$$

Denoting $\eta_t=\mu_t-cZ_t$, the model (59)-(60)-(61) can be rewritten as follows:

$$\frac{dS_t}{S_t}=\eta_t dt+\nu dW_t+\sqrt{\mu_2\rho}dB_t, \tag{62}$$

$$d\eta_t=\kappa(\bar{\mu}-\eta_t)dt+\omega dV_t-c\sqrt{2\kappa}dB_t. \tag{63}$$

The portfolio optimization problem for the portfolio, that consists of a bank account and an asset, whoose price follows the model (62)-(63), can be solved using the framework in [6].

Following [6], let us for convenience rewrite the model (62)-(63) in the following form:

$$dS_t = S_t\Big(\eta_t dt + \sum_{i=1}^{3} \sigma_{S,i} dW_t^{(i)}\Big), \tag{64}$$

$$d\eta_t = \kappa(\bar{\mu} - \eta_t)dt + \sum_{i=1}^{3} \sigma_{\eta,i} dW_t^{(i)}, \tag{65}$$

where $W_t^{(1)}$, $W_t^{(2)}$, $W_t^{(3)}$ are independent Wiener processes, $W_t^{(1)}$ stands for W_t, $W_t^{(2)}$ stands for B_t, $W_t^{(3)}$ stands for V_t and $\sigma_S = (\sigma_{S,1}, \sigma_{S,2}, \sigma_{S,3}) = (\nu, \sqrt{\mu_2\rho}, 0)$ and $\sigma_\eta = (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3}) = (0, c\sqrt{2\kappa}, \omega)$.

The portfolio consists of an asset and of a bank account. Then the portfolio value X_t follows the following process:

$$dX_t = X_t \pi_t \frac{dS_t}{S_t},$$

where π_t denotes the fraction of wealth invested in the risky asset. Initial capital is x. Investor has to select the trading strategy π_t in order to maximize $E[U(X_T)]$, where $U(x) = \frac{x^\gamma}{\gamma}, \gamma < 1$, $\gamma \neq 0$ is the form of the utility function. At each time moment t the investor chooses how much of the wealth will be invested in the risky asset.

The quadratic variation per unit time of the risky asset returns dS_t/S_t is given by:

$$\Sigma_S = \sigma_S \cdot \sigma_S^T = \sigma_{S,1}^2 + \sigma_{S,2}^2 = \nu^2 + \mu_2\rho,$$

the quadratic variation of the drift of the risky asset is

$$\Sigma_\eta = \sigma_\eta \cdot \sigma_\eta^T = \omega^2 + 2c^2\kappa,$$

The covariation between the price of the risky asset and the value of the drift is denoted by R:

$$R = \sigma_S \cdot \sigma_\eta = c\sqrt{2\delta\mu_2\rho}.$$

Theoretically, there are two possible scenarios:

1. Assume that both asset price and drift values are observable. This is the case of full information.

2. Assume that only asset prices are observable but not the drift values. This is the case of incomplete information.

Define the value function

$$J(t, \eta, x) = \sup_{\pi} E[U(X_T^\pi)|X_t = x, \eta_t = \eta].$$

In order to solve this portfolio optimization problem consider the HJB equation considered in Section 4.

7.1.1 Solving portfolio optimization problem under full information

Following [6], the value function under full information is of the form $J(t,\eta,x)$ (η stands for η_t and x stands for X_t) and satisfies the HJB equation (39):

$$\max_{\pi}\Big\{\frac{\partial}{\partial t}J(t,\eta,x)+\frac{\partial}{\partial x}J(t,\eta,x)\pi\eta x+\frac{1}{2}\frac{\partial^2}{\partial x^2}J(t,\eta,x)\pi^2\Sigma_S x^2$$

$$+\frac{1}{2}\frac{\partial^2}{\partial \eta^2}J(t,\eta,x)\Sigma_\eta+\frac{\partial}{\partial x\partial \eta}J(t,\eta,x)\pi Rx+\frac{\partial}{\partial \eta}J(t,\eta,x)\kappa\bar{\mu}-\frac{\partial}{\partial \eta}J(t,\eta,x)\kappa\eta\Big\}=0, \tag{66}$$

with boundary condition $J(T,\eta,x)=\frac{x^\gamma}{\gamma}$, $\gamma<1$, $\gamma\neq 0$. The solution is assumed to be of the following form:

$$\tilde{J}(t,\eta,x)=\frac{x^\gamma}{\gamma}\exp\left(\eta^2A_t+\eta B_t+C_t\right),$$

where functions A_t, B_t, C_t are to be determined later.

The optimal portfolio strategy π^* can be obtained by substituting the value function $\tilde{J}(t,x,\eta)$ into (66). The HJB equation gets the following form:

$$\max_{\pi}\Big\{\Big[x^\gamma\left\{\eta^2\frac{dA_t}{dt}+\eta\frac{dB_t}{dt}+\frac{dC_t}{dt}\right\}+\frac{1}{2}\Sigma_S x^2\pi^2(\gamma(\gamma-1))x^{\gamma-2}$$

$$+\frac{1}{2}\Sigma_\eta x^\gamma(4\eta^2A_t^2)+4\eta A_tB_t+2A_t+B_t^2)$$

$$+\pi Rx\gamma x^{\gamma-1}(2\eta A_t+B_t)+\pi\eta x\gamma x^{\gamma-1}+\kappa\bar{\mu}(2\eta A_t+B_t)$$

$$-\kappa\eta(2\eta A_t+B_t))\Big]\exp\left(\eta^2A_t+\eta B_t+C_t\right)\Big\}=0. \tag{67}$$

Differentiating the expression in the brackets with respect to π and setting it to zero one obtains the following:

$$\Sigma_S x^2\pi(\gamma(\gamma-1))x^{\gamma-2}+Rx\gamma x^{\gamma-1}(2\eta A_t+B_t)+\eta x\gamma x^{\gamma-1}=0,$$

then

Proposition 7.1. *(Brendle) The maximizer in (67) (optimal portfolio strategy after Theorem 7.5 below) is as follows:*

$$\pi^*(\eta,t)=\frac{2R\eta A_t+\eta+RB_t}{\Sigma_S(1-\gamma)}. \tag{68}$$

Coefficients A_t, B_t and C_t are obtained by setting π^* into (67):

$$\eta^2 \frac{dA_t}{dt} + \eta \frac{dB_t}{dt} + \frac{dC_t}{dt} + \frac{1}{2}\Sigma_S \left(\frac{2R\eta A_t + \eta + RB_t}{\Sigma_S(1-\gamma)} \right)^2 (\gamma(1-\gamma))$$

$$+\frac{1}{2}\Sigma_\eta x^\gamma (4\eta^2 A_t^2) + 4\eta A_t B_t + 2A_t + B_t^2)$$

$$+\frac{2R\eta A_t + \eta + RB_t}{\Sigma_S(1-\gamma)} R\gamma(2\eta A_t + B_t) + \frac{2R\eta A_t + \eta + RB_t}{\Sigma_S(1-\gamma)}\eta\gamma$$

$$+\kappa\bar{\mu}(2\eta A_t + B_t) - \kappa\eta(2\eta A_t + B_t) = 0 \tag{69}$$

Next, combine together the terms with η^2, η and η^0. For example, for η^2 one obtains:

$$\eta^2 \frac{dA_t}{dt} + \frac{1}{2}\gamma\eta^2 \frac{(2RA_t + 1)^2}{\Sigma_S(1-\gamma)} + \frac{1}{2}\Sigma_\eta 4A_t^2\eta^2$$

$$+\eta^2 \frac{\gamma}{\Sigma_S(1-\gamma)}(2RA_t + 1) - 2\kappa A_t \eta^2 + \eta^2 \frac{\gamma R}{\Sigma_S(1-\gamma)}(2RA_t^2 + 2A_t) = 0$$

Removing η^2, one obtains an ODE for A_t. The same thing is done with terms containing η and η^0.

Proposition 7.2. *(Brendle) Thus, coefficients A_t, B_t and C_t satisfy the following ordinary differential equations*

$$\frac{dA_t}{dt} = -2A_t \left\{ \Sigma_\eta + \frac{\gamma}{1-\gamma} R^2 \Sigma_S^{-1} \right\} A_t + 2\kappa A_t - \frac{2\gamma}{1-\gamma}\Sigma_S^{-1} R A_t - \frac{1}{2}\frac{\gamma}{1-\gamma}\Sigma_S^{-1}$$

$$\frac{dB_t}{dt} = -2A_t \left\{ \Sigma_\eta + \frac{\gamma}{1-\gamma} R^2 \Sigma_S^{-1} \right\} B_t + 2\kappa A_t \bar{\mu} + \kappa B_t - \frac{\gamma}{1-\gamma}\Sigma_S^{-1} R B_t$$

$$\frac{dC_t}{dt} = -\Sigma_\eta A_t - \frac{1}{2}\left\{ \Sigma_\eta + \frac{\gamma}{1-\gamma} R^2 \Sigma_S^{-1} \right\} B_t - B_t \kappa \bar{\mu}$$

with boundary condition $A_T = B_T = C_T = 0$.

The same result for the optimal strategy π_t^* can be obtained if one considers the utility function of the form $U(x) = -x^\gamma$ for $\gamma < 0$.

Verification theorem. Such a theorem can be formulated similarly as it was done in the paper [3].

For the verification theorem let us assume that the utility function is $U(x) = \frac{x^\gamma}{\gamma}$ for $\gamma < 1$, $\gamma \neq 0$.

Let us introduce the new notation for the functions and operators.

The function $h(t,\eta)$ is the part of the value function $\tilde{J}(t,\eta,x)$

$$h(t,\eta) = \exp(\eta^2 A_t + \eta B_t + C_t).$$

The operator H^π is the operator of the HJB equation:

$$H^\pi h(t,\eta) = L^\pi h(t,\eta) + Gh(t,\eta),$$

where

$$L^\pi h(t,\eta) = h(t,\eta)\Big[\pi\eta + \frac{1}{2}\Sigma_S \pi^2(\gamma-1) + \pi R(2\eta A_t + B_t)\Big]$$

and

$$Gh(t,\eta) = \frac{h(t,\eta)}{\gamma}\Big[\eta^2 \frac{dA_t}{dt} + \eta\frac{dB_t}{dt} + \frac{dC_t}{dt}$$

$$+\frac{1}{2}\Sigma_\eta(4\eta^2 A_t^2 + 4\eta A_t B_t + 2A_t + B_t^2) + \kappa\bar{\mu}(2\eta A_t + B_t)$$

$$-\kappa\eta(2\eta A_t + B_t)\Big].$$

Operator L^π depends on the portfolio strategy π and G does not depend on π.

The HJB equation is as follows:

$$\sup_\pi\{H^\pi h(t,\eta)\} = 0. \tag{70}$$

Before the verification theorem is stated, the auxilary lemma has to be formulated, based on the paper [3].

For $\tilde{J}(t,\eta,x) = \frac{x^\gamma}{\gamma}h(t,\eta)$ and any admissible portfolio strategy π_t with wealth $X_t = X_t^\pi$ we have

$$d\tilde{J}(t,\eta_t,X_t) = X_t^\gamma \cdot H^{\pi_t}h(t,\eta_t)dt + d\psi_t^\pi, \tag{71}$$

The following lemma gives the explicit form of ψ_t^π.

Lemma 7.3.

$$\psi_t^\pi = \int_0^t \Big[X_s^\gamma h(s,\eta_s)\Big[(B_s + 2A_s\eta_s)\sum_{i=1}^{3}\sigma_{\eta,i}dW_s^{(i)} + \gamma\pi_s\sum_{i=1}^{3}\sigma_{S,i}dW_s^{(i)}\Big]\Big] \tag{72}$$

Proof. Let us denote $F_t = X_t^\gamma$ which satisfies the following SDE (applying Ito lemma)

$$dF_t = \gamma F_t\Big(\eta_t\pi_t + \frac{1}{2}(\gamma-1)\Sigma_S\pi_t^2\Big)dt + \gamma F_t\pi_t\sum_{i=1}^{3}\sigma_{S,i}dW_t^{(i)}.$$

Taking into account $J(t,\eta_t,X_t) = \frac{1}{\gamma}X_t^\gamma h(t,\eta_t) = \frac{1}{\gamma}F_t h(t,\eta_t)$, the product rule yields

$$dJ(t,\eta_t,X_t) = \frac{1}{\gamma}F_t dh(t,\eta_t) + \frac{1}{\gamma}h(t,\eta_t)dF_t.$$

The differential form of $h(t,\eta_t)$ can be calculated using Ito lemma:

$$\begin{aligned} dh(t,\eta_t) &= d(e^{\eta_t^2 A_t+\eta_t B_t+C_t}) \\ &= e^{\eta_t^2 A_t+\eta_t B_t+C_t}\Big[\eta_t^2\frac{dA_t}{dt} + \eta_t\frac{dB_t}{dt} + \frac{dC_t}{dt} \\ &+\kappa(\bar{\mu}-\eta_t)(B_t+2A_t\eta_t)dt + (B_t+2A_t\eta_t)\sum_{i=1}^{3}\sigma_{\eta,i}dW_t^{(i)} \\ &+\frac{1}{2}\Sigma_\mu(4\eta_t^2A_t^2 + 4\eta_t A_t B_t + B_t^2 + 2A_t)dt\Big]. \end{aligned}$$

Therefore,

$$\begin{aligned} dJ(t,\eta_t,X_t) &= \frac{1}{\gamma}X_t^\gamma(\underbrace{L^\pi h(t,\eta_t) + Gh(t,\eta_t)}_{H^\pi h(t,\eta_t)})dt \\ &+\frac{1}{\gamma}X_t^\gamma h(t,\eta_t)\Big[(B_t+2A_t\eta_t)\sum_{i=1}^{3}\sigma_{\eta,i}dW_t^{(i)} + \gamma\pi_t\sum_{i=1}^{3}\sigma_{S,i}dW_t^{(i)}\Big]. \end{aligned}$$

Therefore,

$$\psi_t^\pi = \int_0^t\Big[X_s^\gamma h(s,\eta_s)\Big[(B_s+2A_s\eta_s)\sum_{i=1}^{3}\sigma_{\eta,i}dW_s^{(i)} + \gamma\pi_s\sum_{i=1}^{3}\sigma_{S,i}dW_s^{(i)}\Big]\Big].$$

□

For the following verification theorem we need that ψ^π is a martingale. We simply require this for the strategies π we consider and refer to Remark 7.6 for a short discussion.

So we consider

$$\mathrm{A}(t,\eta,x) = \{\pi = (\pi_s)_{s\in[t,T]} | \pi \text{ admissible, } \psi^\pi \text{ is martingale}\}$$

and require

Assumption 7.4. For $\pi_t^* = \pi^*(\eta_t, t)$ in (68), $\psi_t^{\pi^*}$ is a martingale.

Further, the value function $J(t, \eta, x)$ will be considered only for π that are in $\mathrm{A}(t, \eta, x)$:

$$J(t, \eta, x) = \sup_{\pi \in \mathrm{A}(t,\eta,x)} E[U(X_T^\pi) | X_t = x, \eta_t = \eta], \tag{73}$$

and $\tilde{J}(t, \eta, x)$ is as above.

Further, a verification theorem similar to the one in paper [3] is introduced.

Theorem 7.5. *(Verification Theorem) Under Assumption 7.4, we have that $\tilde{J}(t, \eta, x) = \frac{x^\gamma}{\gamma} h(t, \eta)$ is the value function when considering strategies in $\mathrm{A}(t, \eta, x)$ and π_t^*, given by (68), is the optimal strategy for the given portfolio optimization problem.*

Proof. Under Assumption 7.4, the optimal portfolio strategy π^* is in $\mathrm{A}(t, \eta, x)$ and is given by the maximizer for the HJB equation (70). Further, h satisfiees the HJB equation (70) with $h(T, \eta_T) = 1$.

For any portfolio strategy $\pi \in \mathrm{A}(t, \eta, x)$ the function $\tilde{J}(t, \mu_t, X_t^\pi)$ satisfies decomposition (71):

$$\tilde{J}(T, \eta_T, X_T^\pi) = \tilde{J}(t, \eta_t, X_t^\pi) + \int_t^T F_s H^{\pi_s} h(s, \eta_s) ds + \psi_T^\pi - \psi_t^\pi,$$

where ψ_t^π is a martingale with zero expectation, $\tilde{J}(t, \eta_t, X_t^\pi) = \frac{1}{\gamma}(X_t^\pi)^\gamma h(t, \eta_t)$, $F_t = (X_t^\pi)^\gamma$. Since $h(t, \eta_t)$ satisfies the HJB equation (70), then $H^\pi h(t, \eta_t) \leq 0$. Therefore,

$$\int_t^T F_s H^{\pi_s} h(s, \eta_s) ds \leq 0.$$

Moreover,

$$\tilde{J}(T, \eta_T, X_T^\pi) \leq \tilde{J}(t, \eta_t, X_t^\pi) + \psi_T^\pi - \psi_t^\pi.$$

Taking into account that $\tilde{J}(T, \eta_T, X_T^\pi) = U(X_T^\pi)$, the conditional expectation is taken on both sides:

$$E^{t,\eta,x}[U(X_T^\pi)] \leq \tilde{J}(t, \eta, x).$$

Further, taking the supremum over all admissible trading strategies on both sides, one obtaines the following result:

$$J(t, \eta, x) := \sup_{\pi \in \mathrm{A}(t,\eta,x)} E^{t,\eta,x}[U(X_T^\pi)] \leq \tilde{J}(t, \eta, x). \tag{74}$$

Since π^* exist, we have an equality in (74). For π_t^* we have the following:

$$\int_t^T F_s \cdot H^{\pi_s^*} h(s, \eta_s) ds = 0.$$

Therefore,

$$J(t, \eta, x) = \tilde{J}(t, \eta, x).$$

This finishes the proof.

□

Remark 7.6.
- For $\gamma \in (0, 1)$, ψ_t^π is a positive local martingale and, thus, a supermartingale. Then we would consider in Theorem 7.5 all admissible π and would need the martingale property only for π^*.
- Conditions for Assumption 7.4 could presumably be obtained following similar arguments as e.g. in [23] or [31] based on repeated application of the Hölder inequality and Burkholder-Davis-Gundy inequality.

7.1.2 Solving portfolio optimization problem under partial information

Following [6], the value function under partial information is denoted by $\hat{J}(t, \hat{\eta}, \Omega, x)$, where $\hat{\eta}$ and Ω correspond to the filter: conditional mean

$$\hat{\eta}_t = E[\eta_t | F_t^S]$$

and conditional variance

$$\Omega_t = E[(\eta_t - \hat{\eta}_t)^2 | F_t^S],$$

where F_t^S is augmented filtration generated by the asset prices. Both are the solution of the Kalman-Bucy filtering problem. In continuous time $\hat{\eta}_t$ and Ω_t satisfy the pair of differential equations which can also be solved numerically in discrete time:

$$d\hat{\eta}_t = -\kappa(\hat{\eta}_t - \bar{\mu}_t)dt + (\Omega_t + R)\Sigma_S^{-1} \left(\frac{dS_t}{S_t} - \hat{\eta}_t dt \right)$$

$$\frac{d\Omega_t}{dt} = \Sigma_\eta - 2\kappa\Omega_t - (\Omega_t + R)^2 \Sigma_S^{-1}.$$

Since Ω_t is deterministic, we have $\hat{J}(t, \hat{\eta}, \Omega, x) = \hat{J}(t, \hat{\eta}, x)$.

The value function $\hat{J}(t, \hat{\eta}, x)$ ($\hat{\eta}$ stands for $\hat{\eta}_t$ and x stands for X_t) satisfies the Bellman equation:

$$\max_\pi \Big\{ \frac{\partial}{\partial t} \hat{J}(t, \hat{\eta}, x) + \frac{\partial}{\partial x} \hat{J}(t, \hat{\eta}, x) \pi \hat{\eta} x + \frac{1}{2} \frac{\partial^2}{\partial x^2} \hat{J}(t, \hat{\eta}, x) \pi^2 \Sigma_S x^2$$

$$+\frac{1}{2}\frac{\partial^2}{\partial\hat{\eta}^2}\hat{J}(t,\hat{\eta},x)\Sigma_S^{-1}(\Omega_t+R)^2+\frac{\partial}{\partial x\partial\hat{\eta}}\hat{J}(t,\hat{\eta},x)\pi(R+\Omega_t)x$$

$$+\frac{\partial}{\partial\hat{\eta}}\hat{J}(t,\hat{\eta},x)\kappa\bar{\mu}-\frac{\partial}{\partial\hat{\eta}}\hat{J}(t,\hat{\eta},x)\kappa\hat{\eta}\Big\}=0$$

with boundary condition $\hat{J}(T,\hat{\eta},x)=\frac{x^\gamma}{\gamma}$. The solution can be written in the following form:

$$\hat{J}(t,\hat{\eta},x)=\frac{x^\gamma}{\gamma}\exp{(\hat{\eta}^2\hat{A}_t+\hat{\eta}\hat{B}_t+\hat{C}_t)}.$$

Proposition 7.7. *(Brendle) Functions $\hat{A}_t$, $\hat{B}_t$, $\hat{C}_t$ satisfy the ordinary differential equations:*

$$\frac{d\hat{A}_t}{dt}=-\frac{2\hat{A}_t^2}{1-\gamma}(\Omega_t+R)^2\Sigma_S^{-1}+2\kappa\hat{A}_t-\frac{2\gamma}{1-\gamma}\Sigma_S^{-1}(\Omega_t+R)\hat{A}_t$$

$$\frac{d\hat{B}_t}{dt}=-\frac{2\hat{A}_t}{1-\gamma}(\Omega_t+R)^2\Sigma_S^{-1}\hat{B}_t+2\kappa\hat{A}_t\bar{\mu}+\kappa\hat{B}_t-\frac{\gamma}{1-\gamma}\Sigma_S^{-1}(\Omega_t+R)\hat{B}_t$$

$$\frac{d\hat{C}_t}{dt}=-\Sigma_S^{-1}(\Omega_t+R)^2\hat{A}_t-\frac{1}{2}\frac{1}{1-\gamma}\hat{B}_t^2(\Omega_t+R)^2-\hat{B}_t\kappa\bar{\mu}$$

with boundary condition $\hat{A}_T=\hat{B}_T=\hat{C}_T=0$. This system is solved backwards in time.

Proposition 7.8. *(Brendle) The optimal portfolio strategy $\hat{\pi}^*$ under partial information is obtained by solving the HJB in the same way as above:*

$$\hat{\pi}^*(t,\hat{\eta})=\frac{(1+2(\Omega_t+R)\hat{A}_t)\hat{\eta}+(\Omega_t+R)\hat{B}_t)}{\Sigma_S(1-\gamma)}$$

$$=\frac{(1+2RA_t)(1-2\Omega_tA_t)^{-1}(\hat{\eta}+\Omega_tB_t)+RB_t}{\Sigma_S(1-\gamma)}.$$

The same result for the optimal strategy $\hat{\pi}^*$ can be obtained if one considers the utility function of the form $U(x)=-x^\gamma$ for $\gamma<1$, $\gamma\neq 0$.

7.1.3 Solving portfolio optimization problem for the case $\kappa \neq \delta$

Consider the model (59)-(60)-(61):

$$\frac{dS_t}{S_t}=\Big(\mu_t-cZ_t\Big)dt+\nu dW_t+\sqrt{\mu_2\rho}dB_t,$$

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t,$$
$$dZ_t = -\delta Z_t dt + \sqrt{2\delta} dB_t,$$

where $c = \delta\sqrt{\mu_2\rho/2\delta}$.

If $\kappa \neq \delta$, then one cannot transform this model into the model (62)-(63), since in the case $\kappa \neq \delta$ we cannot switch to η as we did above in (63).

Consider the value of c. In order to apply the framework of [6] for solving the portfolio optimization problem, it was assumed that the jumps of the shot-noise process were 'small'. The second-order raw moment μ_2 of the distribution of the jump hight is also 'small', and, therefore, c is 'small' and cZ_t is also 'small' compared to the values of the drift μ_t. Thus, it would be reasonable to neglect the term cZ_t. The new approximated model looks as follows:

$$\frac{dS_t}{S_t} = \mu_t dt + \nu dW_t + \sqrt{\mu_2\rho} dB_t, \tag{75}$$
$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t, \tag{76}$$

where S_t is an asset price, μ_t is the drift process of the stock price, ν is the volatility of the stock price, W_t is Wiener process, ρ is the jump intensity of the shot-noise process λ_t, B_t is the standard Brownian motion. In the second equation $\bar{\mu}$ is the mean of the drift, κ is the rate of mean reversion, ω is volatility of the drift and V_t is Wiener process independent of W_t and B_t.

The model (75)-(76) in terms of the framework in [6] can be rewritten as follows:

$$dS_t = S_t\Big(\mu_t dt + \sum_{i=1}^{3} \sigma_{S,i} dW_t^{(i)}\Big), \tag{77}$$

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \sum_{i=1}^{3} \sigma_{\mu,i} dW_t^{(i)}, \tag{78}$$

where $W_t^{(1)}$, $W_t^{(2)}$, $W_t^{(3)}$ are independent Wiener processes, $W_t^{(1)}$ stands for W_t, $W_t^{(2)}$ stands for B_t, $W_t^{(3)}$ stands for V_t and $\sigma_S = (\sigma_{S,1}, \sigma_{S,2}, \sigma_{S,3}) = (\nu, \sqrt{\mu_2\rho}, 0)$ and $\sigma_\mu = (\sigma_{\mu,1}, \sigma_{\mu,2}, \sigma_{\mu,3}) = (0, 0, \omega)$.

The quadratic variation per unit time of the risky asset returns dS_t/S_t is given by

$$\Sigma_S = \sigma_S \cdot \sigma_S^T = \sigma_{S,1}^2 + \sigma_{S,2}^2 = \nu^2 + \mu_2\rho,$$

the quadratic variation of the drift of the risky asset is

$$\Sigma_\mu = \sigma_\mu \cdot \sigma_\mu^T = \omega^2,$$

The covariation between the price of the risky asset and the value of the drift is denoted by R:

$$R = \sigma_S \cdot \sigma_\mu = 0.$$

In order to solve this portfolio optimization problem one applies the framework of [6], that was discussed in 7.1.1 and 7.1.2.

The model approximation (75)-(76) will be used in order to obtain the simulated numerical results for solving the portfolio optimization problem for the model (56),(58).

7.1.4 Logarithmic utility function

This is an extension of the framework in [6] for the case of the logarithmic utility. The method for solving the portfolio optimization problem is the same as discussed in 7.1.1 and 7.1.2, only the form of the value function is different.

The portfolio optimization problem is solved for the model (62)-(63).

Define the logarithmic utility function $U(x) = \ln x$. As before define the value function

$$J(t, \eta, x) = \sup_\pi E[U(X_T)|X_t = x, \eta_t = \eta].$$

Assume that the value function has the following form

$$J(t, \eta, x) = \ln x + (\eta^2 A_t + \eta B_t + C_t),$$

where the functions A_t, B_t and C_t to be determined later as in the framework above. The value function $J(t, \eta, x)$ (where x stands for X_t and η stands for η_t) satisfies the HJB equation (66):

$$\begin{aligned}\max_\pi\Big\{ &\frac{\partial}{\partial t}J(t,\eta,x) + \frac{\partial}{\partial x}J(t,\eta,x)\pi\eta x + \frac{1}{2}\frac{\partial^2}{\partial x^2}J(t,\eta,x)\pi^2\Sigma_S x^2 \\ &+\frac{1}{2}\frac{\partial^2}{\partial \eta^2}J(t,\eta,x)\Sigma_\eta + \frac{\partial}{\partial x\partial\eta}J(t,\mu,x)\pi R x \\ &+\frac{\partial}{\partial \eta}J(t,\eta,x)\kappa\bar{\mu} - \frac{\partial}{\partial \eta}J(t,\eta,x)\kappa\eta\Big\} = 0, \qquad (79)\end{aligned}$$

with boundary condition $J(T, \eta, x) = \ln x$. The optimal portfolio strategy π^* can be defined by substituting the value function $J(t, \eta, x)$ into the upper equation. The HJB equation takes the following form:

$$\max_\pi\Big\{\eta^2\frac{dA_t}{dt} + \eta\frac{dB_t}{dt} + \frac{dC_t}{dt} - \frac{1}{2}\Sigma_S x^2\pi^2\left(\frac{1}{x^2}\right) + \frac{1}{2}\Sigma_\eta 2A_t$$

$$+\pi Rx \cdot 0 + \pi_t \eta x \frac{1}{x} + \kappa\bar{\mu}(2\eta A_t + B_t) - \kappa\eta(2\eta A_t + B_t)\Big\} = 0. \quad (80)$$

Differentiating the expression in the brackets with respect to π_t and setting it to zero one obtaines the following expression:

$$\eta - \pi\Sigma_S = 0.$$

Proposition 7.9. *The optimal portfolio strategy is*

$$\pi^*(\eta, t) = \frac{\eta}{\Sigma_S}.$$

Coefficients A_t, B_t and C_t satisfy the following ordinary differential equations

$$\frac{dA_t}{dt} = 2\kappa A_t - \frac{1}{2\Sigma_S},$$

$$\frac{dB_t}{dt} = -2\kappa\bar{\mu}A_t + \kappa B_t,$$

$$\frac{dC_t}{dt} = -\kappa\bar{\mu}B_t - \Sigma_\eta A_t,$$

with boundary condition $A_T = B_T = C_T = 0$.

For the case of partial information denote the value function as $\hat{J}(t, \hat{\eta}, x)$, which also depends on deterministic Ω_t, where $\hat{\eta}$ stands for $\hat{\eta}_t$ and, $\hat{\eta}_t$ and Ω_t as in the previous framework denote the conditional mean and covariance matrix of η_t. Both are the solution of the Kalman-Bucy filtering problem.

The function $\hat{J}(t, \hat{\eta}, x)$ (where x stands for X_t) satisfies the Bellman equation:

$$\max_\pi \Big\{ \frac{\partial}{\partial t}\hat{J}(t, \hat{\eta}, x) + \frac{\partial}{\partial x}\hat{J}(t, \hat{\eta}, x)\pi\hat{\eta}x + \frac{1}{2}\frac{\partial^2}{\partial x^2}\hat{J}(t, \hat{\eta}, x)\pi^2\Sigma_S x^2$$

$$+\frac{1}{2}\frac{\partial^2}{\partial\hat{\eta}^2}\hat{J}(t, \hat{\eta}, x)\Sigma_S^{-1}(\Omega_t + R)^2 + \frac{\partial}{\partial x\partial\hat{\eta}}\hat{J}(t, \hat{\eta}, x)\pi_t(R + \Omega_t)x$$

$$+\frac{\partial}{\partial\hat{\eta}}\hat{J}(t, \hat{\eta}, x)\kappa\bar{\mu} - \frac{\partial}{\partial\hat{\eta}}\hat{J}(t, \hat{\mu}, x)\kappa\hat{\eta}\Big\} = 0$$

with boundary condition $\hat{J}(T, \hat{\eta}, x) = \ln x$. The solution can be written in the following form:

$$\hat{J}(t, \hat{\eta}, x) = \ln x + (\hat{\eta}^2\hat{A}_t + \hat{\eta}\hat{B}_t + \hat{C}_t),$$

The optimal portfolio strategy π_t^* can be defined by substituting the value function $J(t, \hat{\eta}, x)$ into the upper equation. The HJB equation takes the following form:

$$\max_{\pi}\Big\{\hat{\eta}^2\frac{d\hat{A}_t}{dt} + \hat{\eta}\frac{d\hat{B}_t}{dt} + \frac{d\hat{C}_t}{dt} - \frac{1}{2}\Sigma_S x^2\pi^2\left(\frac{1}{x^2}\right)$$
$$+2\frac{1}{2}\Sigma_S^{-1}(R+\Omega_t)^2 A_t + \pi(R+\Omega_t)x\cdot 0 + \pi\hat{\eta}x\frac{1}{x} + \kappa\bar{\mu}(2\hat{\eta}\hat{A}_t + \hat{B}_t)$$
$$-\kappa\hat{\eta}(2\hat{\eta}\hat{A}_t + \hat{B}_t))\Big\} = 0. \tag{81}$$

Differentiating the expression in the brackets with respect to π and setting it to zero one obtaines the following ezpression:

$$\hat{\eta} - \pi\Sigma_S = 0.$$

Proposition 7.10. *The optimal portfolio strategy*

$$\hat{\pi}^*(\hat{\eta}, t) = \frac{\hat{\eta}}{\Sigma_S}.$$

Thus, the optimal strategy under partial information $\hat{\pi}^*$ converges to the optimal strategy under full information π^*, if the filter $\hat{\eta}_t$ converges to the true drift process η_t.

Coefficients $\hat{A}_t$, $\hat{B}_t$ and $\hat{C}_t$ satisfy the following ordinary differential equations

$$\frac{d\hat{A}_t}{dt} = 2\kappa\hat{A}_t - \frac{1}{\Sigma_S},$$
$$\frac{d\hat{B}_t}{dt} = \kappa\hat{B}_t - 2\kappa\hat{A}_t\bar{\mu},$$
$$\frac{d\hat{C}_t}{dt} = -\kappa\bar{\mu}\hat{B}_t - \frac{\hat{A}_t(R+\Omega_t)^2}{\Sigma_S},$$

with boundary condition $\hat{A}_T = \hat{B}_T = \hat{C}_T = 0$.

7.2 Portfolio optimization for the case when asset price model is influenced by shot noise process

7.2.1 Portfolio optimization for shot noise without approximation

In this subsection we discuss the theoretical solution to the optimization problem for the portfolio, consisting of an asset, whose dynamics follows jump diffusion process, driven by Brownian motion and shot noise process, and a bank account with the constant interest rate.

Suppose in the financial market there are two possible investments:

- bank account with price dynamics:

$$dP_t = rP_t dt$$

- risky asset

$$S_t = S_0 e^{\int_0^t \mu_s ds - \frac{\nu^2}{2}t + \nu W_t + \lambda_t}$$

or

$$dS_t = S_{t-}\left[(\mu_t - \delta\lambda_t)dt + \nu dW_t + \int_E (e^y - 1)N(dt, dy)\right],$$

drift process μ_t is described by the following SDE:

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t,$$

where V_t and W_t are independent Wiener processes, λ_t is the shot noise process whose differential form is as follows:

$$d\lambda_t = -\delta\lambda_t dt + \int_E yN(dt, dy),$$

where $N(dt, dy)$ is the Poisson random measure with $E[N(dt, dy)] = \rho F(dy)dt$, where ρ is the jump intensity and $F(dy)$ is the distribution of the jumps hight.

In this framework we do not intend to approximate the shot-noise process and obtain approximate solution to the portfolio optimization problem, but instead we intend to obtain a purely theoretical solution.

Define the wealth process X_t as follows:

$$dX_t = X_t(1 - \pi_t)\frac{dP_t}{P_t} + X_t\pi_t\frac{dS_t}{S_t}$$

$$= (r(1 - \pi_t) + (\mu_t - \delta\lambda_t)\pi_t)X_t dt + \nu\pi_t X_t dW_t + \pi_t X_{t-}\int_E g(y)N(dt, dy),$$

where $g(y) = e^y - 1$, π is the self-financing portfolio strategy, which denotes the fraction of wealth invested in risky assets. For simplicity one can assume $r = 0$. Denote the process π_t to be the control process and the process $Z_t = Z_t^{(\pi)} = (t, X_t, \mu_t, \lambda_t)$ to be the controled jump diffusion process. Consider the following performance criterion

$$\Phi(z) = E^z[U(X_t)],$$

where $U(x) = \frac{x^\gamma}{\gamma}$, for $\gamma < 1$, $\gamma \neq 0$. The same result for the optimal trading strategy π^* can be obtained if one considers the power utility function of the form $U(x) = -x^\gamma$, $\gamma < 0$ as in [6].

The problem is to find the optimal trading strategy π^* so as to maximize $E^z[U(X_T)]$ and the value function is

$$J(z) = J(t, x, \mu, \lambda) = \sup_{\pi} E^z[U(X_T)|X_t = x, \mu_t = \mu, \lambda_t = \lambda].$$

The generator of $Z_t^{(\pi)}$ is as follows:

$$\begin{aligned} AJ(z) = \frac{\partial J}{\partial s} + \pi x(\mu - \delta\lambda)\frac{\partial J}{\partial x} + \frac{1}{2}\nu^2\pi^2x^2\frac{\partial^2 J}{\partial x^2} \\ + \int_E \{J(x + x\pi g(y)) - J(x)\}v(dy) + \frac{1}{2}\omega^2\frac{\partial^2 J}{\partial \mu^2} + \kappa\bar{\mu}\frac{\partial J}{\partial \mu} - \kappa\mu\frac{\partial J}{\partial \mu} \\ -\delta\lambda\frac{\partial J}{\partial \lambda} + \int_E \{J(\lambda + y) - J(\lambda)\}v(dy), \end{aligned}$$

where $v(dy) = \rho F(dy)$, x stands for X_t, μ stands for μ_t, π stands for π_t and λ stands for λ_t.

According to the HJB equation for optimal control of jump diffusions [29], the value function $J(Z)$ is the solution of the following equation:

$$\max_{\pi} \{AJ\} = 0. \tag{82}$$

Assume now that the value function has the following form:

$$\tilde{J}(z) = \tilde{J}(t, x, \mu, \lambda) = \frac{x^\gamma}{\gamma}h(t, \mu, \lambda),$$

then

$$\begin{aligned} A\tilde{J}(t, x, \mu, \lambda) = \frac{x^\gamma}{\gamma}\frac{\partial h(t, \mu, \lambda)}{\partial t} \\ +\pi x(\mu - \delta\lambda)x^{\gamma-1}h(t, \mu, \lambda) \\ +\frac{1}{2}\nu^2\pi^2x^2(\gamma - 1)x^{\gamma-2}h(t, \mu, \lambda) \\ +h(t, \mu, \lambda)\frac{1}{\gamma}\int_E \{(x + x\pi g(y))^\gamma - x^\gamma\}v(dy) \\ +\frac{1}{2\gamma}\omega^2x^\gamma\frac{\partial^2 h(t, \mu, \lambda)}{\partial \mu^2} + \frac{1}{\gamma}\kappa\bar{\mu}x^\gamma\frac{\partial h(t, \mu, \lambda)}{\partial \mu} \\ -\frac{1}{\gamma}\kappa\mu x^\gamma\frac{\partial h(t, \mu, \lambda)}{\partial \mu} - \frac{1}{\gamma}\delta\lambda x^\gamma\frac{\partial h(t, \mu, \lambda)}{\partial \lambda} \\ +\frac{1}{\gamma}x^\gamma\int_E \{h(t, \mu, \lambda + y) - h(t, \mu, \lambda)\}v(dy) = p(\pi). \end{aligned} \tag{83}$$

Denote the right hand side of the expression above by $p(\pi)$ and differentiate $p(\pi)$ with respect to π, then set the derivative to zero:

$$\frac{dp(\pi)}{d\pi} = \mu - \delta\lambda + \nu^2\pi(\gamma-1) + \int_E \{(1+\pi g(y))^{\gamma-1} g(y)\} v(dy) = 0.$$

Proposition 7.11. *The maximizer* $\pi^* = \pi(t, \mu, \lambda)$ *in (83) (the optimal trading strategy after Theorem 7.16 below) solves the integral equation*

$$\mu - \delta\lambda + \nu^2\pi(\gamma-1) + \int_E \{(1+\pi g(y))^{\gamma-1} g(y)\} v(dy) = 0. \tag{84}$$

if

$$\frac{d^2p}{d\pi^2}(\pi(t, \mu, \lambda)) < 0$$

for all possible values of t, μ, λ.

The same result for the optimal strategy π^* can be obtained if one considers the utility function of the form $U(x) = -x^\gamma$ for $\gamma < 0$.

Further, we comment on the condition in Proposition 7.11. The trading strategy π is a maximizer of $h(\pi)$ if

$$\frac{d^2p(\pi)}{d\pi^2} < 0,$$

therefore,

$$\frac{d^2p(\pi)}{d\pi^2} = \nu^2(\gamma-1) + (\gamma-1)\int_E (1+\pi g(y))^{\gamma-2} g^2(y) v(dy) \le 0$$

or

$$(1+\pi g(y)) > 0$$

as $\gamma - 1 < 0$.

It follows that

$$\pi < -\frac{1}{g(y)} \text{ for } g(y) < 0$$

and

$$\pi > -\frac{1}{g(y)} \text{ for } g(y) > 0.$$

Therefore,

$$\underline{\pi} = \sup\Big\{-\frac{1}{g(y)}\Big| y \in supp(v),\ g(y) > 0\Big\} < \pi$$
$$< \inf\Big\{-\frac{1}{g(y)}\Big| y \in supp(v),\ g(y) < 0\Big\} = \overline{\pi}.$$

If the jumps are normally distributed, then y can take values from $-\infty$ to ∞, and $g(y) = e^y - 1 \in (-1, \infty)$, therefore, $\underline{\pi} = 0$ and $\overline{\pi} = 1$. For other types of distributions $\underline{\pi}$ can be less then zero and $\overline{\pi}$ can be bigger then one.

Thus, the trading strategy π, that maximizes $p(\pi)$, belongs to the interval $(\underline{\pi}, \overline{\pi})$ and, therefore, one can set a condition for the existence of the solution of (84) on $(\underline{\pi}, \overline{\pi})$:

$$\int_E \{(1 + \pi g(y))^{\gamma-1} g(y)\} v(dy) > 0 \text{ for } \pi \to \underline{\pi},$$

and

$$\int_E \{(1 + \pi g(y))^{\gamma-1} g(y)\} v(dy) < 0 \text{ for } \pi \to \overline{\pi}.$$

7.2.2 Approximation of the optimal portfolio strategy and the value function

Equation (84) can be solved numerically. As the equation (84) cannot be solved explicitly, it would be reasonable to try to find an appropriate approximation $\tilde{\pi}^*$ of the optimal trading strategy π^*. Denote

$$H(\eta, \pi) = H(\mu - \delta\lambda, \pi) = \mu - \delta\lambda + \nu^2 \pi(\gamma - 1) + \int_E \{(1 + \pi g(y))^{\gamma-1} g(y)\} v(dy) = 0,$$

where $v(dy) = \rho F(dy)$, $F(dy)$ is the probability density of y.

The solution π to the equation (84) is a function of $\eta = \mu - \delta\lambda$, i.e.

$$\pi = f(\mu - \delta\lambda) = f(\eta),$$

and, therefore, $H(\eta, f(\eta)) = 0$. The derivative of H with respect to η is

$$\frac{\partial H}{\partial \eta} = H_\eta + H_\pi f'(\eta) = 0,$$

therefore,

$$f'(\eta) = -\frac{H_\eta(\eta, f(\eta))}{H_\pi(\eta, f(\eta))},$$

or

$$f'(\eta) = -\frac{1}{(\gamma - 1)(\nu^2 + \rho \int_E (1 + \pi g(y))^{\gamma-2} g^2(y) F(dy))}.$$

Applying the first order Taylor expansion to $\pi = f(\eta)$ one obtains

$$\pi = f(\eta) \approx f(c) + f'(c)(\eta - c).$$

As point c for the expansion one can use

$$c := \eta = \mu - \delta\lambda = -(\gamma - 1)\nu^2\pi(c) - \rho\int_E (1 + \pi(c)g(y))^{\gamma-1}g(y)F(dy)$$

Setting $\pi(c) = 0$, obtain the following:

$$c = -\rho\int_E g(y)F(dy).$$

If y is normally distributed with mean m and standard deviation σ, then $c = -\rho(e^{m+\sigma^2} - 1)$.

Therefore, the Taylor expansion of function π is

$$\pi = f(\eta) \approx \pi(c) + f'(c)(\eta - c) = 0 - \frac{1}{(\gamma - 1)(\nu^2 + \rho\int_E g^2(y)F(dy))}(\eta - c),$$

therefore,

$$\tilde{\pi}^* = -\frac{1}{(\gamma - 1)(\nu^2 + \rho\int_E g^2(y)F(dy))}(\mu - \delta\lambda - c).$$

If $y \sim N(m, \sigma)$, then $\int_E g^2(y)v(dy) = \left(e^{m+\sigma^2} - 1\right)^2$.

Proposition 7.12. *The first order approximation $\tilde{\pi}^*$ of the optimal trading strategy π^* is*

$$\tilde{\pi}^* = c_1(\mu - \delta\lambda - c), \tag{85}$$

where

$$c_1 = \frac{1}{(1 - \gamma)(\nu^2 + \rho\int_E g^2(y)F(dy))}.$$

Remark 7.13. The same result for the approximation $\tilde{\pi}^*$ of the optimal trading strategy π^* can be obtained, if one approximates the integral in the equation (84) in the following way:

$$\int_E \{(1 + \pi g(y))^{\gamma-1}g(y)\}v(dy) \approx \int_E \{(1 + (\gamma - 1)\pi g(y))g(y)\}v(dy)$$

$$= \rho\int_E g(y)F(dy) + (\gamma - 1)\pi\int_E g^2(y)F(dy)$$

for 'small' values of y.

Having the explicit form of the approximation $\tilde{\pi}^*$ of the optimal trading strategy π^*, one can set the explicit form of the approximating value function

$$\tilde{J}_{approx}(t,x,\mu,\lambda) = \frac{x^\gamma}{\gamma} e^{(\mu-\delta\lambda)^2\tilde{A}_t + (\mu-\delta\lambda)\tilde{B}_t + \tilde{C}_t},$$

where coefficients $\tilde{A}_t$, $\tilde{B}_t$, $\tilde{C}_t$ depend on the approximation $\tilde{\pi}^*$ of the optimal trading strategy.

Setting $\tilde{\pi}^*$ into the HJB equation, one obtains the following:

$$\frac{1}{\gamma}\Big\{(\mu-\delta\lambda)^2\frac{d\tilde{A}_t}{dt} + (\mu-\delta\lambda)\frac{d\tilde{B}_t}{dt} + \frac{d\tilde{C}_t}{dt}\Big\}$$

$$+c_1(\mu-\delta\lambda-c)(\mu-\delta\lambda) + \frac{1}{2}\nu^2 c_1^2(\mu-\delta\lambda-c)^2(\gamma-1)$$

$$+c_1(\mu-\delta\lambda-c)\int_E\{g(y)\}v(dy) + \frac{1}{2}c_1^2(\mu-\delta\lambda-c)^2\int_E\{(\gamma-1)g^2(y)\}v(dy)$$

$$+\frac{1}{2\gamma}\omega^2((2\tilde{A}_t(\mu-\delta\lambda)+\tilde{B}_t)^2 + 2\tilde{A}_t) + \frac{1}{\gamma}\kappa\bar{\mu}(2\tilde{A}_t(\mu-\delta\lambda)+\tilde{B}_t)$$

$$-\frac{1}{\gamma}\kappa\mu(2\tilde{A}_t(\mu-\delta\lambda)+\tilde{B}_t) - \frac{1}{\gamma}\delta\lambda(-2\delta\tilde{A}_t(\mu-\delta\lambda) - \delta\tilde{B})$$

$$+\frac{1}{\gamma}\int_E\{-2\delta y(\mu-\delta\lambda)A_t + \delta^2y^2A_t - \delta yB_t\}v(dy) = 0.$$

The following approximations were used:

$$\int_E\{(1+\pi g(y))^\gamma - 1\}v(dy) \approx \int_E\Big\{\gamma\pi g(y) + \frac{1}{2}\gamma(\gamma-1)\pi^2g^2(y)\Big\}v(dy)$$

and

$$e^{-2\delta y(\mu-\delta\lambda)A_t + \delta^2y^2A_t - \delta yB_t} - 1 \approx -2\delta y(\mu-\delta\lambda)A_t + \delta^2y^2A_t - \delta yB_t$$

for 'small' values of y.

The expression

$$-\frac{1}{\gamma}\kappa\mu(2\tilde{A}_t(\mu-\delta\lambda)+\tilde{B}_t) - \frac{1}{\gamma}\delta\lambda(-2\delta\tilde{A}_t(\mu-\delta\lambda) - \delta\tilde{B})$$

can be rewritten as

$$-\frac{1}{\gamma}2\delta\tilde{A}_t(\mu-\delta\lambda)(\mu-\delta\lambda) - \frac{1}{\gamma}\delta\tilde{B}_t(\mu-\delta\lambda)$$

for $\kappa = \delta$.

Therefore, the system of ODEs for the coefficients $\tilde{A}_t$, $\tilde{B}_t$, $\tilde{C}_t$ is as follows:

$$\frac{d\tilde{A}_t}{dt} = -2\omega^2 \tilde{A}_t^2 + 2\delta \tilde{A}_t - \gamma c_1 - \frac{1}{2}\gamma(\gamma-1)\nu^2 c_1^2 + \frac{\gamma}{2}(\gamma-1)c_1^2 \int_E g^2(y)v(dy),$$

$$\frac{d\tilde{B}_t}{dt} = \gamma c_1 c + \gamma(\gamma-1)\nu^2 c_1^2 c - c_1\gamma \int_E g(y)v(dy) + \gamma(\gamma-1)c_1^2 c \int_E g^2(y)v(dy) -$$
$$-2\omega^2 \tilde{A}_t \tilde{B}_t - 2\kappa\bar{\mu}\tilde{A}_t + \delta\tilde{B}_t + 2\delta\tilde{A}_t \int_E yv(dy),$$

$$\frac{d\tilde{C}_t}{dt} = -\frac{1}{2}\nu^2\gamma(\gamma-1)c_1^2c^2 + \gamma c_1 c \int_E g(y)v(dy) + \gamma(\gamma-1)c_1^2c^2 \int_E g^2(y)v(dy) - \omega^2\tilde{A}_t -$$
$$-\frac{1}{2}\omega^2\tilde{B}_t^2 - \kappa\bar{\mu}\tilde{B}_t - \int_E (\delta^2 y^2 \tilde{A}_t - \delta y \tilde{B}_t)v(dy),$$

where $A_T = B_T = C_T = 0$, for $\kappa = \delta$.

7.2.3 Verification result for the theoretical solution to the portfolio optimization problem (without approximation)

The theoretical solution to the portfolio optimization problem with asset price driven by shot noise process and Brownian motion was obtained in the previous section. The aim of the verification result is to show that the solution to the HJB equation is also the solution to the initial portfolio optimization problem. Before the verification theorem is stated, the auxilary lemma has to be formulated, based on the paper [3].

In order to formulate this auxilary lemma let us introduce the new notation for the functions and operators.

The function $h(t, \mu, \lambda)$ was earlier defined as the part of the value function $\tilde{J}(t, x, \mu, \lambda)$. The operator H^π corresponds to the generator A of Z_t and consists of two parts:

$$H^\pi h(t, \mu, \lambda) = L^\pi h(t, \mu, \lambda) + Gh(t, \mu, \lambda),$$

where

$$L^\pi h(t, \mu, \lambda) = h(t, \mu, \lambda)\Big[\pi(\mu - \delta\lambda) + \frac{1}{2}\nu^2\pi^2(\gamma-1) + \frac{1}{\gamma}\int_E ((1+\pi g(y))^\gamma - 1)v(dy)\Big]$$

and

$$Gh(t, \mu, \lambda) = \frac{1}{\gamma}\frac{\partial h(t, \mu, \lambda)}{\partial t}$$

$$+\frac{1}{2\gamma}\omega^2\frac{\partial^2 h(t,\mu,\lambda)}{\partial\mu^2}+\frac{1}{\gamma}\kappa\bar{\mu}\frac{\partial h(t,\mu,\lambda)}{\partial\mu}$$
$$-\frac{1}{\gamma}\kappa\mu_t\frac{\partial h(t,\mu,\lambda)}{\partial\mu}-\frac{1}{\gamma}\delta\lambda\frac{\partial h(t,\mu,\lambda)}{\partial\lambda}$$
$$+\frac{1}{\gamma}\int_E\{h(t,\mu,\lambda+y)-h(t,\mu,\lambda)\}v(dy).$$

The HJB equation can be written as

$$\sup_{\pi}\{H^{\pi}h(t,\mu,\lambda,\pi)\}=0. \tag{86}$$

For $\tilde{J}(Z)=\tilde{J}(t,x,\mu,\lambda)=\frac{1}{\gamma}x^{\gamma}h(t,\mu,\lambda)$ and any admissible portfolio strategy π_t with wealth $X_t=X_t^{\pi}$ we have

$$d\tilde{J}(t,X_t,\mu_t,\lambda_t)=\frac{1}{\gamma}X_t^{\gamma}\cdot H^{\pi_t}h(t,\mu_t,\lambda_t)dt+d\psi_t^{\pi}, \tag{87}$$

Further, we formulate the lemma, that gives the explicit form of ψ_t^{π}.

Lemma 7.14.

$$\psi_t^{\pi}=\int_0^t\Big[X_s^{\gamma}\Big[\int_E\{h(t,\mu_t,\lambda_t+y)-h(t,\mu_t,\lambda_t)\}\tilde{N}(dt,dy)$$
$$+\omega\frac{\partial h(t,\mu_t,\lambda_t)}{\partial\mu_t}dV_t+h(t,\mu_t,\lambda_t)\nu\pi_t dW_t+h(t,\mu_t,\lambda_t)\int_E((1+\pi_t g(y))^{\gamma}-1)\tilde{N}(dt,dy)\Big]\Big] \tag{88}$$

Proof. Let us denote $F_t=X_t^{\gamma}$ which satisfies the following SDE (applying Ito lemma)

$$dF_t=\gamma F_t\Big((\mu_t-\delta\lambda_t)\pi_t+\frac{1}{2}(\gamma-1)\nu^2\pi_t^2\Big)dt+\gamma F_t\nu\pi_t dW_t$$
$$+\gamma F_t\int_E((1+\pi_t g(y))^{\gamma}-1)N(dt,dy).$$

Taking into account $J(t,X_t,\mu_t,\lambda_t)=\frac{1}{\gamma}X_t^{\gamma}h(t,\mu_t,\lambda_t)=\frac{1}{\gamma}F_t h(t,\mu_t,\lambda_t)$, the product rule yields

$$dJ(t,X_t,\mu_t,\lambda_t)=\frac{1}{\gamma}F_t dh(t,\mu_t,\lambda_t)+\frac{1}{\gamma}h(t,\mu_t,\lambda_t)dF_t.$$

The differential form of $h(t,\mu,\lambda)$ can be calculated using Ito lemma:

$$dh(t,\mu,\lambda) = \frac{\partial h(t,\mu,\lambda)}{\partial t}dt$$

$$+\frac{1}{2}\omega^2\frac{\partial^2 h(t,\mu,\lambda)}{\partial \mu^2}dt + \kappa\bar{\mu}\frac{\partial h(t,\mu,\lambda)}{\partial \mu}dt$$

$$-\kappa\mu_t\frac{\partial h(t,\mu,\lambda)}{\partial \mu}dt + \omega\frac{\partial h(t,\mu,\lambda)}{\partial \mu}dV_t - \delta\lambda\frac{\partial h(t,\mu,\lambda)}{\partial \lambda}dt$$

$$+\int_E \{h(t,\mu,\lambda+y) - h(t,\mu,\lambda)\}N(dt,dy).$$

Notice that $\tilde{N}(dt,dy) = N(dt,dy) - v(dy)dt$, therefore,

$$dJ(t,X_t,\mu_t,\lambda_t) = \frac{1}{\gamma}X_t^\gamma(\underbrace{L^\pi h(t,\mu_t,\lambda_t) + Gh(t,\mu_t,\lambda_t)}_{H^\pi h(t,\mu_t,\lambda_t)})dt$$

$$+\frac{1}{\gamma}X_t^\gamma\Big[\int_E \{h(t,\mu_t,\lambda_t+y) - h(t,\mu_t,\lambda_t)\}\tilde{N}(dt,dy)$$

$$+\omega\frac{\partial h(t,\mu_t,\lambda_t)}{\partial \mu_t}dV_t + h(t,\mu_t,\lambda_t)\nu\pi_t dW_t + h(t,\mu_t,\lambda_t)\int_E((1+\pi_t g(y))^\gamma - 1)\tilde{N}(dt,dy)\Big].$$

Therefore,

$$\psi_t^\pi = \int_0^t\Big[X_s^\gamma\Big[\int_E \{h(t,\mu_t,\lambda_t+y) - h(t,\mu_t,\lambda_t)\}\tilde{N}(dt,dy)$$

$$+\omega\frac{\partial h(t,\mu_t,\lambda_t)}{\partial \mu_t}dV_t + h(t,\mu_t,\lambda_t)\nu\pi_t dW_t + h(t,\mu_t,\lambda_t)\int_E((1+\pi_t g(y))^\gamma - 1)\tilde{N}(dt,dy)\Big]\Big]$$

□

For the following verification theorem we need that ψ_t^π is a martingale. We simply require this for the portfolio strategies π we consider.

Consider

$$\mathrm{A}(t,x,\mu,\lambda) = \{\pi = (\pi_s)_{s\in[t,T]} | \pi \text{ admissible, } \psi^\pi \text{ is martingale}\}$$

and require

Assumption 7.15. For $\pi_t^* = \pi^*(t,\mu_t,\lambda_t)$ in (84), $\psi_t^{\pi^*}$ is a martingale.

Further, the value function $J(t,x,\mu,\lambda)$ will be considered only for π that are in $\mathrm{A}(t,x,\mu,\lambda)$:

$$J(t,x,\mu,\lambda) = \sup_{\pi\in \mathrm{A}(t,x,\mu,\lambda)} E[U(X_T^\pi)|X_t = x, \mu_t = \mu, \lambda_t = \lambda], \tag{89}$$

and $\tilde{J}(t,x,\mu,\lambda)$ is as above.

Further, a verification theorem similar to the one in paper [3] is introduced.

Theorem 7.16. *(Verification Theorem) Under Assumption 7.15, we have that* $\tilde{J}(t,x,\mu,\lambda) = \frac{1}{\gamma}x^\gamma h(t,\mu,\lambda)$ *is the value function when considering strategies in* $\mathrm{A}(t,x,\mu,\lambda)$ *and* π_t^**, given by* (84)*, is the optimal strategy for the given portfolio optimization problem.*

Proof. Under Assumption 7.15, π_t^* is in $\mathrm{A}(t,x,\mu,\lambda)$ and is given by the maximizer for the HJB equation (86). Further, h satisfies HJB equation (86).

For any $\pi \in \mathrm{A}$ we apply Lemma 7.14 to $\tilde{J}(t, X_t^\pi, \mu_t, \lambda_t)$:

$$\tilde{J}(T, X_T^\pi, \mu_T, \lambda_T) = \tilde{J}(t, X_t^\pi, \mu_t, \lambda_t) + \int_t^T F_s H^{\pi_s} h(s,\mu_s,\lambda_s) ds + \psi_T^\pi - \psi_t^\pi,$$

where ψ_t^π is a martingale with zero expectation, $\tilde{J}(t, X_t^\pi, \mu_t, \lambda_t) = \frac{1}{\gamma}(X_t^\pi)^\gamma h(t,\mu_t,\lambda_t)$, $F_t = (X_t^\pi)^\gamma$. Since $h(t,\mu,\lambda)$ satisfies the HJB equation (86), then $Hh(t,\mu_t,\lambda_t) \leq 0$. Therefore,

$$\int_t^T F_s H^{\pi_s} h(s,\mu_s,\lambda_s) \leq 0.$$

Moreover,

$$\tilde{J}(T, X_T^\pi, \mu_T, \lambda_T) \leq \tilde{J}(t, X_t^\pi, \mu_t, \lambda_t) + \psi_T^\pi - \psi_t^\pi.$$

Taking into account that $\tilde{J}(T, X_T^\pi, \mu_T, \lambda_T) = U(X_T^\pi)$, the conditional expectation is taken on both sides:

$$E^{t,x,\mu,\lambda}[U(X_t^\pi)] \leq \tilde{J}(t,x,\mu,\lambda).$$

Further, taking the supremum over all admissible trading strategies on both sides, one obtaines the following result:

$$J(t,x,\mu,\lambda) := \sup_{\pi\in \mathrm{A}(t,x,\mu,\lambda)} E^{t,x,\mu,\lambda}[U(X_T^\pi)] \leq \tilde{J}(t,x,\mu,\lambda). \tag{90}$$

Since π^* exists, we have an equality in (90). For π_t^* we have the following:

$$\int_t^T F_s \cdot H^{\pi_s^*} h(s,\mu_s,\lambda_s) ds = 0.$$

Therfore,

$$J(t,x,\mu,\lambda) = \tilde{J}(t,x,\mu,\lambda).$$

This finishes the proof.

□

The Remark 7.6 is also valid for this case.

7.2.4 Logarithmic utility function

Under the portfolio optimization problem of the previous section assume that the utility function is

$$U(x) = \ln x.$$

The value function, that solves the HJB equation (82), is assumed, with respect to the logarithmic utility, as follows

$$\tilde{J}(z) = \tilde{J}(t,x,\mu,\lambda) = \ln x + h(t,\mu,\lambda).$$

Substituting this value function into the HJB equation (82) and then differentiating with respect to π one obtaines the following result:

Proposition 7.17. *The optimal trading strategy π_t solves the integral equation*

$$\mu_t - \delta\lambda_t - \nu^2\pi_t + \int_E \frac{g(y)}{1+\pi_t g(y)} v(dy) = 0,$$

which can be solved numerically.

7.3 Numerical experiments

Consider the following asset price equation:

$$S_t = S_0 e^{\int_0^t (\mu_s - \frac{0.5^2}{2})ds + 0.5W_t + \lambda_t}, \tag{91}$$

where μ_t is the drift process, which follows the equation:

$$d\mu_t = 0.5(0.1 - \mu_t)dt + 0.2dV_t, \tag{92}$$

where shot noise λ_t has following parameters: intensity of ρ the Poisson process M_t equals 1.04 and the jump size y is normally distributed $N(0, 0.007)$.

The portfolio X_t, consisting of an asset and a bank account is constructed, and we want to find the optimal trading strategy π_t^* (the proportion of wealth invested into stocks at each time moment) so that to maximize the expected utility $E[U(X_T)]$ at terminal time T, where U is the utility function of the form $U(x) = -x^\gamma$ with $\gamma = -0.2$ or $U(x) = x^\gamma/\gamma$.

This portfolio optimization problem is solved using the theoretical result (and the first-order approximation of the theoretical result), stated in 7.2, and the result for the case when the shot-noise process is approximated by Brownian motion, stated in 7.1.3.

In the simulated figures that are shown below, several possible solutions to the portfolio optimization problem (91)-(92) are compared. First, this problem is solved using the method proposed in the previous subsection (theoretical (unapproximated and approximated) solutions to portfolio optimization problem for the asset price driven by shot noise process), second, the above problem is solved under full and partial information using the method proposed in the paper [6] and described in Subsection 7.1.

Figure 15 shows the stock price dynamics, in Figure 16 true and filtered drift process are shown. Figure 17 compares the optimal trading strategies: two strategies according to framework of subsection 7.1 (under full and partial information), and two according to the theoretical unapproximated (84) and approximated (85) solutions to the portfolio optimization problem from the previous subsection. The dynamics of the wealth growth under four optimal trading strategies is shown in Figure 18.

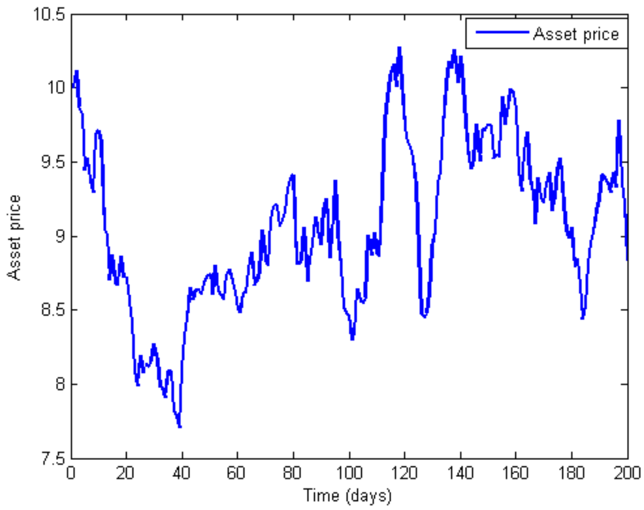

Figure 15: Asset price behaviour (first example)

The average error between real and filtered drift processes is 0.0077.

The average error between the strategy, obtained under theoretical solution (unapproximated), and the strategy, obtained under approximation of

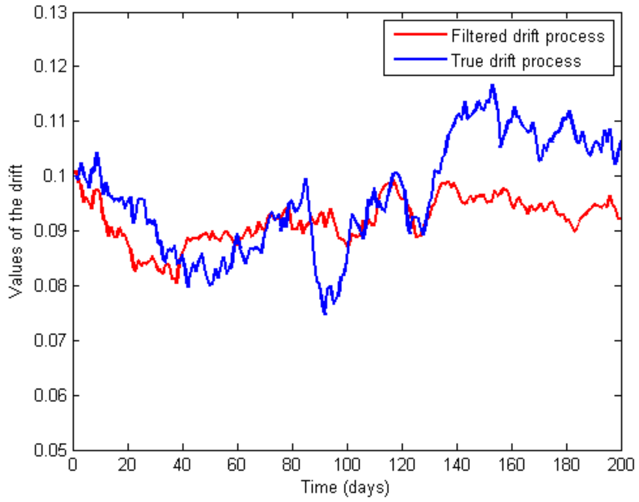

Figure 16: True and filtered drift processes (first example)

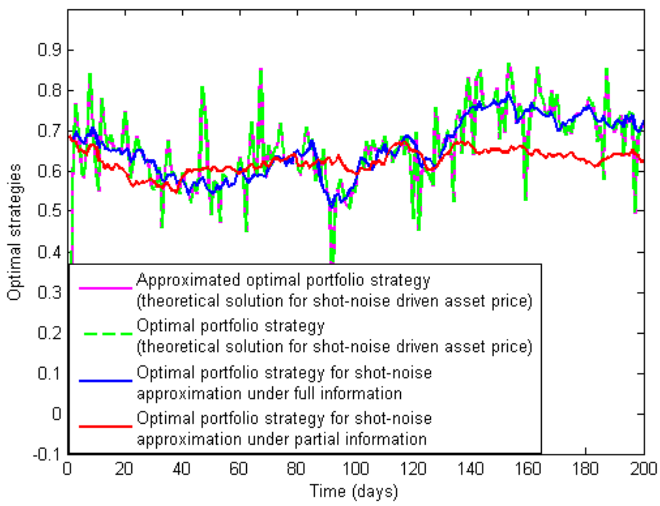

Figure 17: Optimal startegies (first example)

shot-noise and filtering of the drift, is 0.0747. The optimal trading strategies (theoretical unapproximated (84) and approximated (85) solutions) are very close to one another with the average error 0.0043.

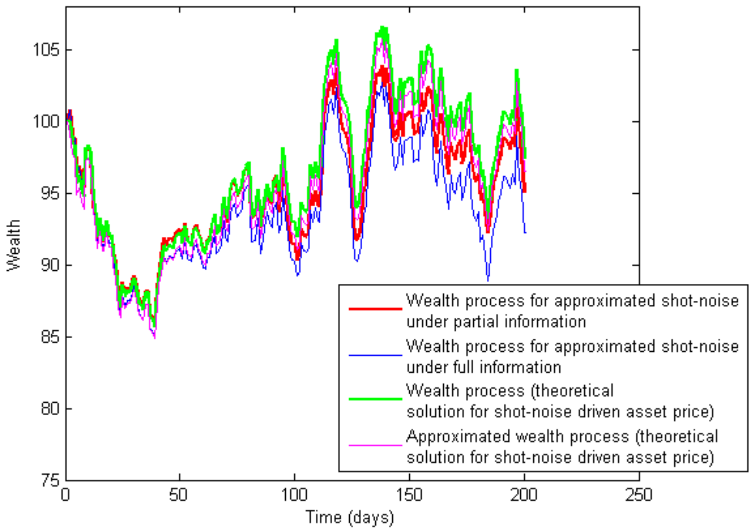

Figure 18: Wealth processes (first example)

The average error between theoretical wealth process and the wealth process, obtained under approximation of the shot noise process and filtering of the drift, is 1.2732.

The second simulation example of the same model is shown in Figures 19, 20, 21 and 22.

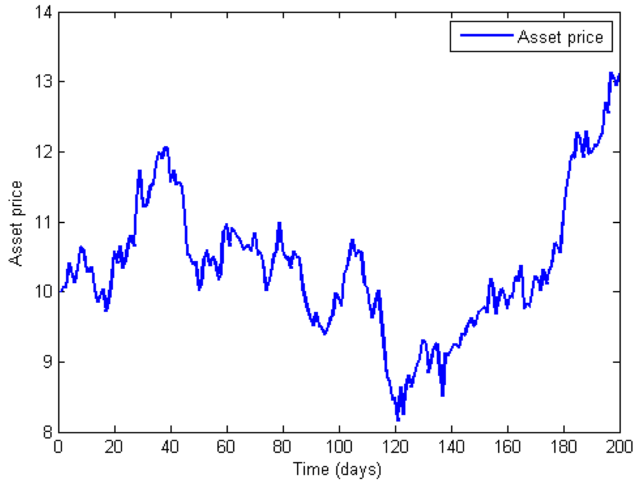

Figure 19: Asset price behaviour (second example)

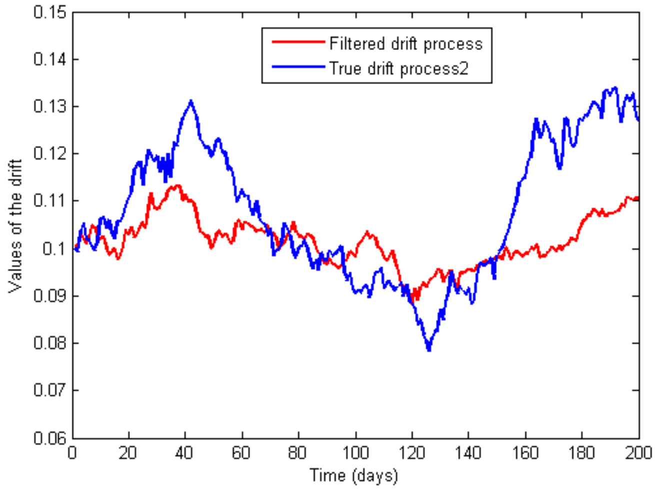

Figure 20: True and filtered drift processes (second example)

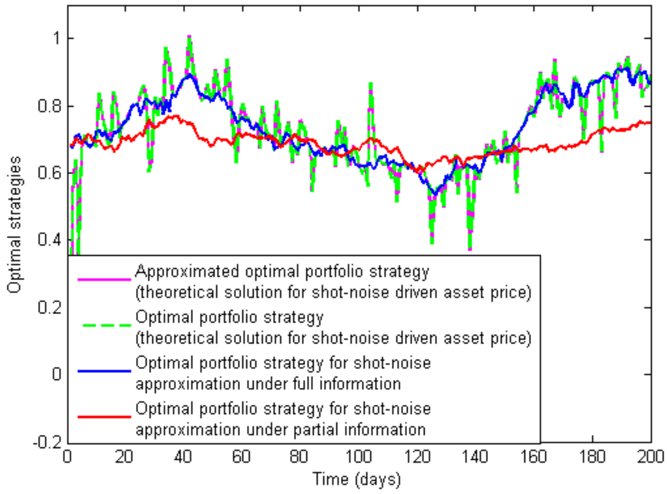

Figure 21: Optimal startegies (second example)

Figure 22: Wealth processes (second example)

Error estimation. The error between the theoretical and approximated portfolio wealth can be estimated by the loss in expected utility, i.e.

$$E[U(X_T^\pi)] - E[U(X_T^{approx})],$$

where X_T^π is the optimal portfolio wealth and X_T^{approx} is the approximated portfolio wealth at terminal time T and the utility function is $U(x) = x^\gamma/\gamma$, where $\gamma = -0.2$.

The loss in the utility between the portfolio wealth, obtained by applying the theoretical (upapproximated) optimal portfolio strategy, and the wealth, obtained by applying the opimal trading strategy under shot-noise approximation (for full and partial information), is 0.0033 and 0.0601 respectively. The loss in the utility between the portfolio wealth, obtained by applying the theoretical (unapproximated) optimal strategy and theoretical approximated optimal strategy is 0.0275. The expected value is estimated as an average over 1000 simulations.

The table below shows the expected utility of the terminal wealth for two different utility functions. The terminal wealth is calculated using the optimal portfolio strategies, obtained in different ways (theoretical solution (unapproximated and approximated) and the shot-noise approximated solution (under full and partial information)). In brackets one can see the standard deviation of the expected utility.

Table 1: Expected utility $E[U(X_T)]$ and standard deviation of expected utility

	$U(x) = \ln x$	$U(x) = x^\gamma/\gamma$
Theoretical solution (unapproximated)	4.6142	−2.0272(0.0573)
Theoretical solution (approximated)		−2.0547(0.0668)
Approximate solution under full information	4.4888	−2.0305(0.0590)
Approximate solution under partial information	4.3222	−2.0873(0.1184)

In this table 'Theoretical solution (unapproximated)' denotes the solution, obtained without approximating the shot-noise process and without approximating the optimal portfolio strategy, as stated in 7.2.1. 'Theoretical solution (approximated)' means, that the solution is obtained without approximating the shot-noise process but with approximation of the optimal portfolio strategy, given in 7.2.2. Both 'Approximate solution under full information' and 'Approximate solution under partial information' refer to the framework by [6] under full and partial information respectively, as considered in 7.1.

7.4 Theoretical extension: Portfolio optimization for the case when asset price is influenced by shot noise process: observing big jumps

The shot noise process can consist of the jumps that are of big size and of small size. There is a criterion in the literature that determines whether the jump is big or not, and it will be considered later.

If the jumps are big, they can be visually observed and eliminated. The asset price model with jumps is

$$dS_t = S_{t-}\left[(\mu_t - \delta\lambda_t)dt + \nu dW_t + \int_E (e^y - 1)N(dt, dy)\right],$$

and the asset price model without big jumps is

$$dS_t = S_{t-}\left[(\mu_t - \delta\lambda_t)dt + \nu dW_t\right],$$

drift process μ_t is described by the following SDE:

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t,$$

where V_t and W_t are independent Wiener processes.

One can estimate the drift μ_t using Kalman filtering. After estimating the drift process, one can optimize the portfolio as proposed below.

Rewrite the stock price process using the innovation process:

$$dS_t = S_{t-}\left[(\hat{\mu}_t - \delta\lambda_t)dt + \nu d\tilde{W}_t + \int_E (e^y - 1)N(dt, dy)\right], \tag{93}$$

where $\tilde{W}_t = W_t + \frac{1}{2}\int_0^t(\mu_s - \hat{\mu_s})ds$ is the innovation process. Substituting $\tilde{W}_t$ into (93) we obtain the original stock price equation. According to the proposition in [4] $\tilde{W}_t$ is a F_t^S-adapted Brownian motion under measure P. Further, one can solve the optimization problem using the theoretical method stated in subsection 7.2.

The next question is how to distinguish and observe the jumps and particularly big jumps. According to the paper [24] there is a statistical test that determines whether there is a jump at the current moment. Suppose that the discussed Lévy process is of the form:

$$d\log S_t = \mu_t dt + \nu_t dW_t + dL_t, \tag{94}$$

where L_t is an F_t-adapted Lévy process independent of W_t. The processes μ_t and ν_t satisfy

Assumption 7.18. (Lee, Hannig) μ_t and ν_t are smooth, can be stochastic and depend on the price process; these processes do not have big jumps over a short time period; ν_t is always positive.

Definition 7.19. (Lee, Hannig) The piecewise constant process $T(t)$ which is used to test whether there is a jump in the time period $(t_{i-1}, t_i]$ is defined in the following way:

$$T(t) = \frac{\log (S_{t_i}/S_{t_{i-1}})}{\hat{\nu}_{t_i} \Delta t^{1/2}},$$

where for any $g > 0$, $0 < \omega < 1/2$ and $t \in K$

$$\hat{\nu_t}^2 = \frac{\Delta t^{-1}}{K} \sum_{j=i-K}^{i-1} (\log S_{t_{j-m+1}}/S_{t_{j-m}})^2 I_{\{\log S_{t_{j-m+1}}/S_{t_{j-m}} \leq g\Delta t^{\omega}\}},$$

where $\Delta t = T/n$, n is the number of observations over $[0, T]$.

Theorem 7.20. *(Lee, Hannig) Let* $T(t)$ *be as in the definition* 7.19 *and* $K \to \infty$ *and* $\Delta t K \to 0$*. Suppose Lévy process follows the dynamics (*94*) and Assumption 7.18 is satisfied:*
A. For any stopping time τ *such that* $\Delta S(\tau) = 0$ *almost surely (there is no jump at time* τ *almost surely), as* $\Delta t \to 0$

$$T(\tau) \to N(0, 1),$$

where $N(0, 1)$ *denotes a standard normal distribution.*
B. Define the time of k^{th} *jump bigger then* h

$$\tau_{k,h} = \inf \{t > \tau_{k-1,h}, \Delta L_t > h\}.$$

Then, as $\Delta t \to \infty$

$$P\left(\min_k \frac{T(\tau_{k,h})}{h/(\nu_{\tau_{k,h}}\sqrt{\Delta t})} \geq 1\right) \to 1.$$

Therefore, $T(\tau_{k,h}) \to \infty$ *as* $\Delta t \to 0$*.*

In the paper [24] it is proposed to test for the presence and arrivals of big jumps using the following rule. It is hypothesized that our realized returns come from model without jump term $d \log S_t = \mu_t dt + \nu_t dW_t$, and it is considered how large the magnitude of $T(t)$ can be. For this, the asymptotic null distribution of its maximums is studied in the following proposition in [24]. This offers the big-jump detection region.

Proposition 7.21. *(Lee, Hannig: big Lévy jump detection rule): Let $T(t)$ be as in Definition and $K \to \infty$ and $\Delta t K \to 0$. Suppose the process is without jumps and Assumption 1 is satisfied. Then, as $\Delta t \to 0$,*

$$\frac{\max_{t\in(t_{i-1},t_i]\ for\ 0\le i\le n}|T(t)| - C_n}{S_n} \to \xi,$$

where ξ has a cumulative distribution function $P(\xi \le x) = \exp(-e^{-x})$,

$$C_n = (2\log n)^{1/2} - \frac{\log\pi + \log\log n}{2(2\log n)^{1/2}},$$

$$S_n = \frac{1}{(2\log n)^{1/2}}.$$

Notice that the test $T(t)$ is defined as a piecewise constant and the maximum in the Proposition 7.21 is the same as if it were taken at all observation times t_i.

The main use of Proposition 7.21 is to set up the big-jump rejection regions of the test: *namely, detect a big jump arrival at testing time t_i if the absolute value of the test statistic is bigger than $q_\alpha S_n + C_n$, where q_α is the α quantile of the limiting distribution of maximum ξ.*

In the real market situation the shot noise process λ_t may consist of a component $\lambda_t^{(1)}$ that is generated by small jumps and a component $\lambda_t^{(2)}$ that is generated by big jumps, i.e.

$$\lambda_t = \lambda_t^{(1)} + \lambda_t^{(2)},$$

where $\lambda_t^{(k)} = \lambda_0^{(k)} e^{-\delta t} + \sum_{i=1}^{M_t^{(k)}} Y_i^{(k)} e^{-\delta(t-S_i)}$ for $k = 1, 2$. Here we have two independent Poisson processes $M_t^{(1)}$ and $M_t^{(2)}$, and random varibles $Y_i^{(1)}$, $Y_i^{(2)}$ have different distributions, corresponging to big and small jumps.

In this case small jumps can be approximated by Brownian motion and the optimal filter can be computed according to [6] as described in Subsection 7.1. Big jumps can be tested and eliminated and then small jumps can again be approximated by Brownian motion. Therefore, a portfolio optimization problem with additional observable jumps would have to be solved.

8 Compound Poisson process as a noise term of the asset price: filtering and portfolio optimization

Asset price model that includes discontinuities has already been studied for quite a long time. Compound Poisson process as a noise term was first proposed in the papers [30] and [27]. The general Lévy process as a noise term was proposed in [26] as early as in 1963. From that time there have been many studies (concerning Lévy processes as noise terms) in option pricing, portfolio optimization, filtering and many more. Lévy process in the asset price models discribe certain unexpected abrupt movements due to bankrupcy, natural catastrophes, wars and etc.

Theoretical solutions to the filtering problem with jump-diffusion processes were considered in the introductory part of the thesis (papers [15], [28], [32]). But these theoretical solutions are most often not applicable in practice. In this section of the thesis we propose an approximate solution to the filtering and portfolio potimization problems, where asset price model is driven by Brownian motion and compound Poisson process.

8.1 Filtering problem with compound Poisson process as a noise term

Compound Poisson process is defined as follows:

$$J_t = \sum_{i=1}^{M_t} Y_i = \sum_{i,s_i \le t} Y_i = \int_0^t \int_E H(s,y) N(ds,dy),$$

where $H(s_i, Y_i) = Y_i$. This definition is a special case of the shot-noise process with $\delta = 0$.

Compound Poisson process J_t has the following differential form:

$$dJ_t = \int_E H(s,y) N(ds,dy) = \int_E y N(ds,dy).$$

Asset price process S_t is described by the similar model as it was assumed in the shot-noise framework:

$$S_t = S_0 e^{L_t} = S_0 e^{\int_0^t \mu_s ds - \frac{\nu^2}{2} t + \nu W_t + J_t},$$

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t,$$

where V_t is the Brownian motion independent of Brownian motion W_t, ω is the diffusion coefficient, $\bar{\mu}$ is mean value and κ is the rate of mean-reversion; ν is volatility of the asset price equation, and J_t is the compound Poisson prosess. This model structure of the asset price is mentioned in [10].

Let

$$dL_t = \left(\mu_t - \frac{\nu^2}{2}\right) dt + \nu dW_t + \int_E y N(dt,dy).$$

Thus,

$$dS_t = S_0 e^{L_{t-}} dL_t^c + \frac{1}{2} S_0 e^{L_{t-}} \nu^2 dt + \int_E (S_0 e^{L_{t-}+y} - S_0 e^{L_{t-}}) N(dt, dx) =$$

$$= S_{t-} \left[\mu_t dt + \nu dW_t + \int_E (e^y - 1) N(dt, dy) \right],$$

where dL_t^c is the continuous part of dL_t.

The same assumption as in the case of shot-noise process is made: if on average the values of jumps Y_i are less than some 'small' ε, then the last term can be approximated by linear function i.e. $e^y - 1 \approx y$. Then the SDE, that corresponds to the stock price process, can be written in the following approximate form:

$$dS_t = \mu_t S_{t-} dt + \nu S_{t-} dW_t + S_{t-} \int_E (e^y - 1) N(dt, dy) \approx$$

$$\approx \mu_t S_{t-} dt + \nu S_{t-} dW_t + S_{t-} dJ_t.$$

Returning back to the approximation of shot-noise process, we recall the auxilary result from [7]

$$V_t^{(\rho)} = \frac{J_t - \mu_1 \rho t}{\sqrt{\mu_2 \rho / 2\delta}} \to \sqrt{2\delta} B_t, \ as \ \rho \to \infty,$$

where $J_t = \sum_{i=1}^{M_t} Y_i$, M_t is Poisson process with intensity ρ and B_t is Brownian motion.

This means that compound Poisson process $J_t = \sqrt{\mu_2 \rho / 2\delta} V_t^\rho + \mu_1 \rho t$ can be approximated by $B_t \sqrt{\mu_2 \rho} + \mu_1 \rho t$, thus

$$dJ_t \approx \sqrt{\mu_2 \rho} dB_t + \mu_1 \rho dt. \tag{95}$$

and, therefore,

$$dS_t \approx (\mu_t + \mu_1 \rho) S_{t-} dt + \nu S_{t-} dW_t + S_{t-} dB_t \sqrt{\mu_2 \rho},$$

where B_t and W_t are independent Brownian motions.

Therefore, the following asset price model is considered:

- unobservable drift of the stock price

$$d\mu_t = \kappa(\bar{\mu} - \mu_t) dt + \omega dV_t, \tag{96}$$

where V_t is Brownian motion independent of W_t and B_t, ω is the diffusion coefficient, $\bar{\mu}$ is mean value and κ is the rate of mean-reversion;

- observable stock price

$$\frac{dS_t}{S_t} = \mu_t dt + \nu dW_t + dJ_t, \tag{97}$$

or

$$dS_t \approx (\mu_t + \mu_1 \rho) S_{t-} dt + \nu S_{t-} dW_t + S_{t-} dB_t \sqrt{\mu_2 \rho}.$$

If the drift process is unobservable, then it can be filtered by applying Kalman filtering to the approximated asset price model (asset price equation driven by two independent Brownian motions).

For numerical simulations the following parameters of the model are chosen: $\nu = 0.3$, $\kappa = 0.5$, $\bar{\mu} = 0.5$, $\omega = 0.3$, $\rho = 0.008$ and $Y \sim N(0, 0.0002)$.

Compound Poisson process is simulated so that jumps occur very rarely (not more than 4 jumps during the whole time period). The convergence of the normalized compound Poisson process to the Brownian motion is valid only when $\rho \to \infty$. Therefore, approximation (95) should be valid for 'big' ρ (ρ is the jump intensity of the Poisson process M_t). When ρ is large, then it is difficult to simulate the asset price process, since very large values occure and the asset price process explodes. In the examples below we see that even for a 'small' ρ one obtaines a quite reasonable filtering results.

Two examples are presented below. Figures 23 and 26 show two compound Poisson processes, Figures 24 and 27 show the asset price dynamics influenced by Brownian motion and compound Poisson process. And, finally, Figures 25 and 28 show the corresponding true drift processes that follow Ornstein-Uhlenbeck and the filtered drift processes, that were obtained using Kalman filtering.

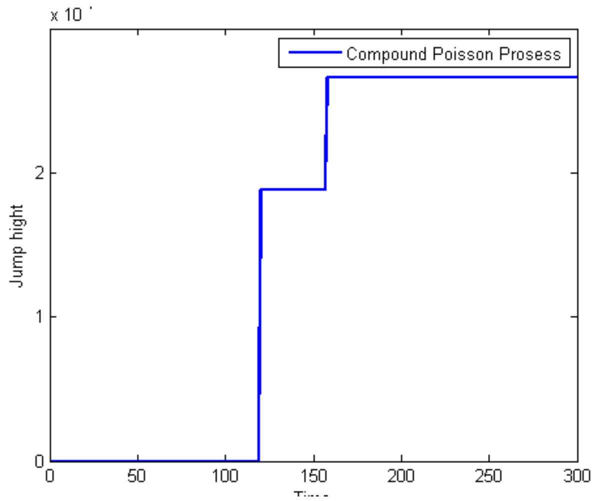

Figure 23: Compound Poisson process (first example)

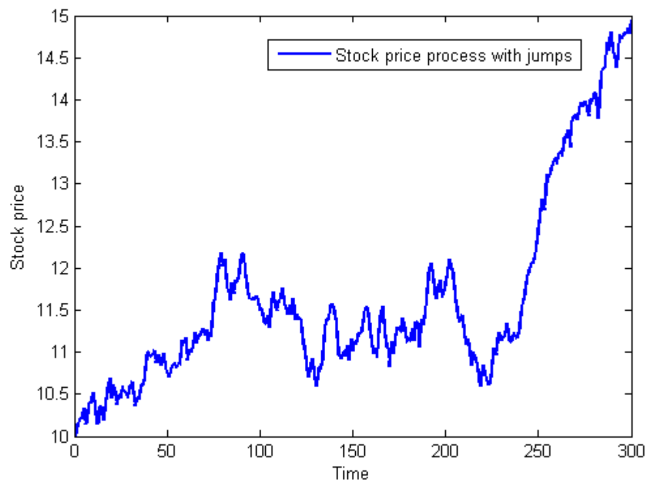

Figure 24: Dynamics of asset price influenced by compound Poisson process (first example)

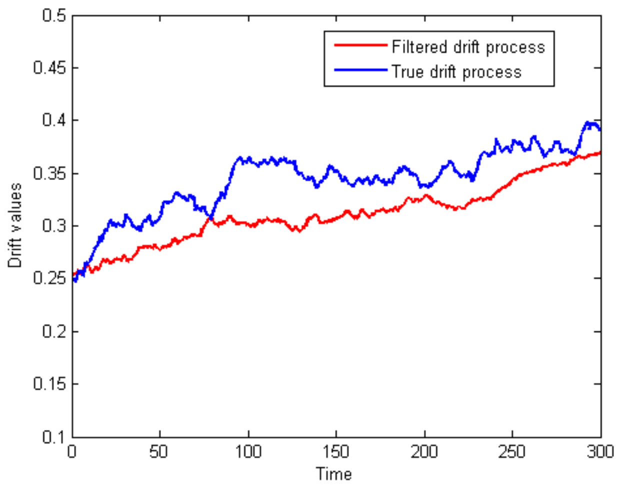

Figure 25: True and filtered behavior of the drift process μ_t (first example)

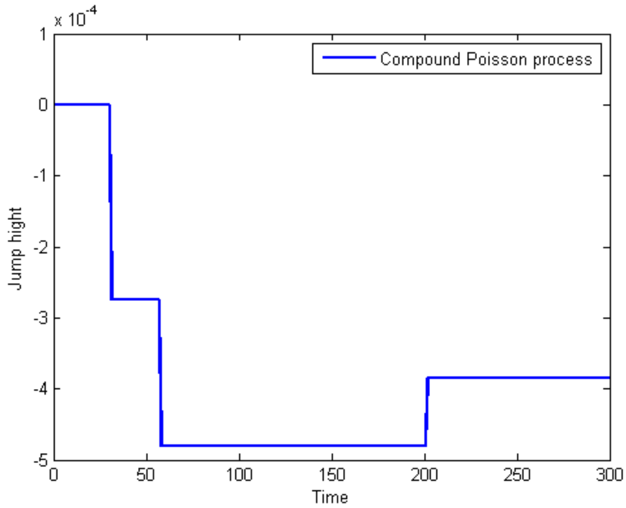

Figure 26: Compound Poisson process (second example)

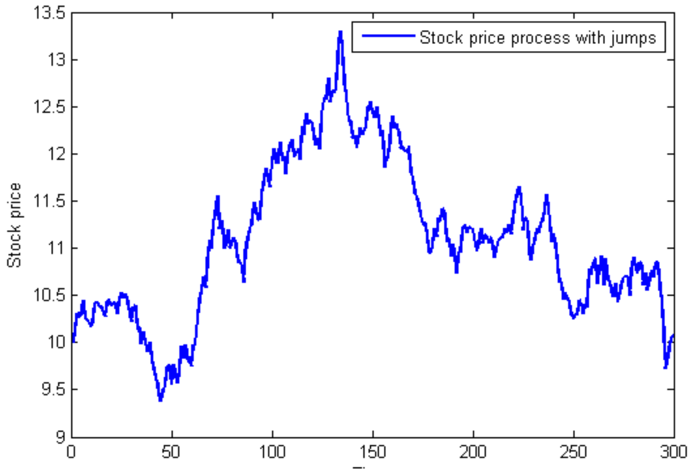

Figure 27: Dynamics of asset price influenced by compound poisson process (second example)

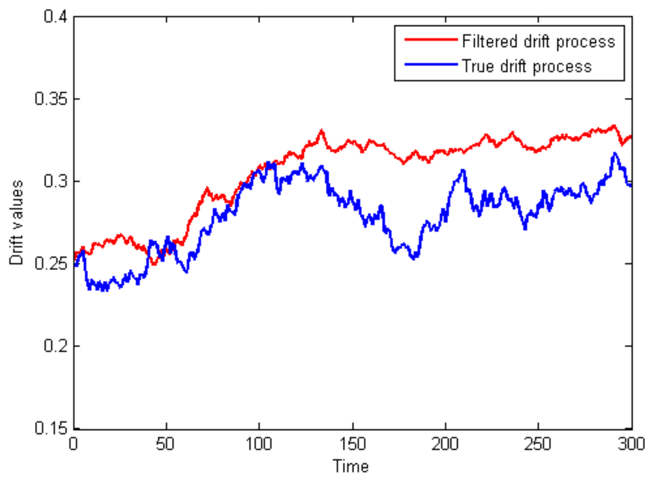

Figure 28: True and filtered behavior of the drift process μ_t (second example)

8.2 Portfolio optimization for the case when asset price model is driven by compound Poisson process: theoretical solution

This framework is similar to the framework of Subsection 7.2, where asset price with shot-noise process was discussed. Here we consider the portfolio optimization problem and its theoretical solution for the case when asset price model is driven by Brownian motion and compound Poisson process. Compound Poisson process is not approximated.

If one constructs a portfolio consisting of a risky asset, where compound Poisson process in approximated by Brownian motion

$$dS_t \approx (\mu_t + \mu_1 \rho) S_{t-} dt + \nu S_{t-} dW_t + S_{t-} dB_t \sqrt{\mu_2 \rho},$$

and a bank account

$$dP_t = rP_t dt,$$

then one can use the framework of the paper [6](or subsection 7.1 of this thesis) in order to obtain an approximate solution to the optimal portfolio problem.

Further we present the theoretical solution to the optimal portfolio problem.

Suppose in the market there are two possible investments:

- bank account with price dynamics:

$$dP_t = rP_t dt$$

- risky asset

$$S_t = S_0 e^{\int_0^t \mu_s ds - \frac{\nu^2}{2} t + \nu W_t + J_t}$$

  or

$$dS_t = S_{t-} \left[\mu_t dt + \nu dW_t + \int_E (e^y - 1) N(dt, dy) \right],$$

drift process μ_t is described by the following SDE:

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \omega dV_t,$$

where V_t and W_t are independent Wiener processes, J_t is the compound Poisson process, $N(dt, dy)$ is the Poisson random measure with $E[N(dt, dy)] = \rho F(dy)dt$, where ρ is the jump intensity and $F(dy)$ is the distribution of the jumps hight.

Define the wealth process X_t as follows:

$$dX_t = X_t(1-\pi_t)\frac{dP_t}{P_t} + X_t\pi_t\frac{dS_t}{S_t} =$$

$$= [r(1-\pi_t) + \mu_t\pi_t]X_t dt + \nu\pi_t X_t dW_t + \pi_t X_{t-}\int_E g(y)N(dt,dy),$$

where $g(y) = e^y - 1$, π is the self-financing portfolio strategy, which denotes the fraction of wealth invested in risky assets. For simplicity one can assume $r = 0$. The process π_t is the control process and process $Z_t = Z_t^{(\pi)} = (t, X_t, \mu_t)$ is a controlled jump diffusion process. Consider performance criterion

$$\Phi(z) = E^z[U(X_t)],$$

where $U(x) = x^\gamma/\gamma$, $\gamma < 1$, $\gamma \neq 1$. The problem is to find the optimal trading strategy π^* so as to maximize $E^z[U(X_T)]$.

The generator of $Z_t^{(\pi)}$ is as follows:

$$AJ(z) = \frac{\partial J}{\partial t} + \pi x\mu\frac{\partial J}{\partial x} + \frac{1}{2}\nu^2\pi^2x^2\frac{\partial^2 J}{\partial x^2} +$$

$$+\int_E \{J(x + x\pi g(y)) - J(x)\}v(dy) + \frac{1}{2}\omega^2\frac{\partial^2 J}{\partial\mu^2} + \kappa\bar{\mu}\frac{\partial J}{\partial\mu} - \kappa\mu\frac{\partial J}{\partial\mu},$$

where $v(dy) = \rho F(dy)$ and x stands for X_t, μ stands for μ_t and π stands for π_t.

According to HJB for optimal control of jump diffusions [29], the value function $J(z)$ is the solution of the following equation:

$$\max_\pi\{AJ\} = 0.$$

Set $\tilde{J}(z) = \tilde{J}(t,x,\mu) = \frac{x^\gamma}{\gamma}h(t,\mu)$, then

$$A\tilde{J}(t,x,\mu) = \frac{x^\gamma}{\gamma}\frac{\partial h(t,\mu)}{\partial t}$$

$$+\pi x\mu x^{\gamma-1}h(t,\mu,\lambda)$$

$$+\frac{1}{2}\nu^2\pi^2x^2(\gamma-1)x^{\gamma-2}h(t,\mu,\lambda)$$

$$+h(t,\mu,\lambda)\frac{1}{\gamma}\int_E \{(x + x\pi g(y))^\gamma - x^\gamma\}v(dy)$$

$$+\frac{1}{2\gamma}\omega^2x^\gamma\frac{\partial^2 h(t,\mu,\lambda)}{\partial\mu^2} + \frac{1}{\gamma}\kappa\bar{\mu}x^\gamma\frac{\partial h(t,\mu,\lambda)}{\partial\mu}$$

$$-\frac{1}{\gamma}\kappa\mu x^{\gamma}\frac{\partial h(t,\mu,\lambda)}{\partial \mu}=p(\pi).$$

Denote the right hand side of the expression above by $h(\pi)$ and differentiate $h(\pi)$ with respect to π, then set the derivative to zero.

Proposition 8.1. *The optimal trading strategy satisfies the following integral equation:*

$$\mu+\nu^2\pi(\gamma-1)+\int_E\{(1+\pi g(y))^{\gamma-1}g(y)\}v(dy)=0.$$

if

$$\frac{d^2p}{d\pi^2}(\pi(t,\mu))<0$$

for all possible values of t, μ.

This equation has to be solved numerically.

Remark 8.2. The integral equation in the above proposition can be obtained by setting $\delta=0$ in the integral equation (84) in the Proposition 7.11. By setting $\delta=0$, the shot-noise process λ_t is transformed into the compound Poisson process J_t.

For numerical simulations consider the following asset price model:

$$S_t=S_0e^{L_t}=S_0e^{\int_0^t\mu_s ds-\frac{0.5^2}{2}t+0.5W_t+J_t}, \tag{98}$$

with drift model:

$$d\mu_t=0.5(0.05-\mu_t)dt+0.1dV_t,$$

where compound Poisson process J_t has following parameters: intensity of ρ the Poisson process M_t equals 0.008 and the jump size y is normally distributed with $N(0,0.0002)$.

The portfolio X_t, consisting of an asset and a bank account is constructed, and we want to find the optimal trading strategy π_t^* (the proportion of wealth invested into stocks at each time moment) so that to maximize the expected utility $E[U(X_T)]$ at terminal time T, where U is the utility function of the form $U(x)=-x^{\gamma}$ with $\gamma=-0.2$ or $U(x)=x^{\gamma}/\gamma$.

Figure 29 shows the asset price process according to (98). True and filtered drift processes of the model (98) are shown in Figure 30. Figure 31 compares the optimal portfolio strategies: theoretical solution (framework of this subsection) and approximate solutions (under full and partial information) based on [6]. And, finally, Figure 32 shows three wealth processes

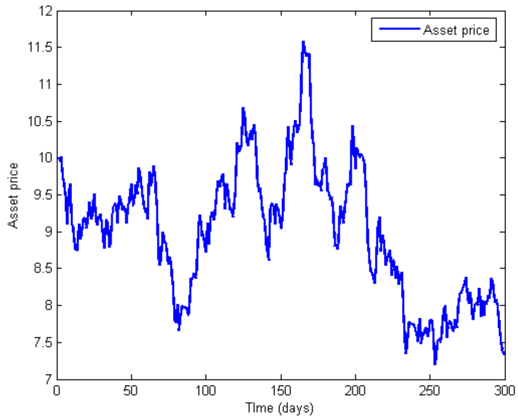

Figure 29: Asset price behaviour influenced by compound Poisson process (third example)

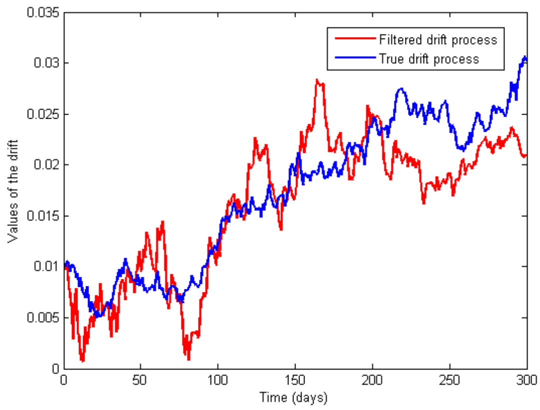

Figure 30: True and filtered drift processes (third example)

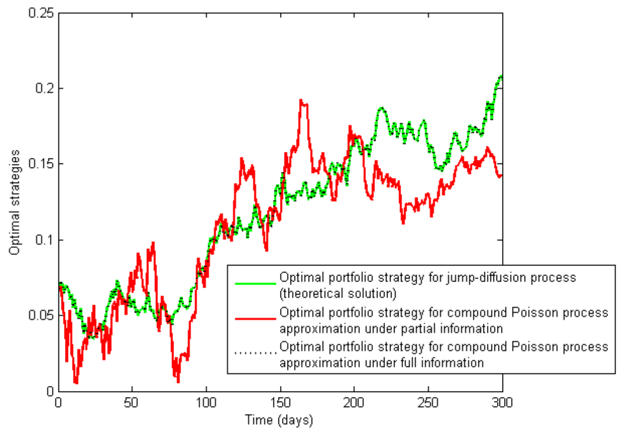

Figure 31: Optimal startegies (third example)

calculated under three different optimal portfolio strategies, that are shown in Figure 31.

Average error between real and filtered drift processes is 0.0033.

Average error between jump-diffusion theoretical strategy and the strategy for approximated and filtered case is 0.0225.

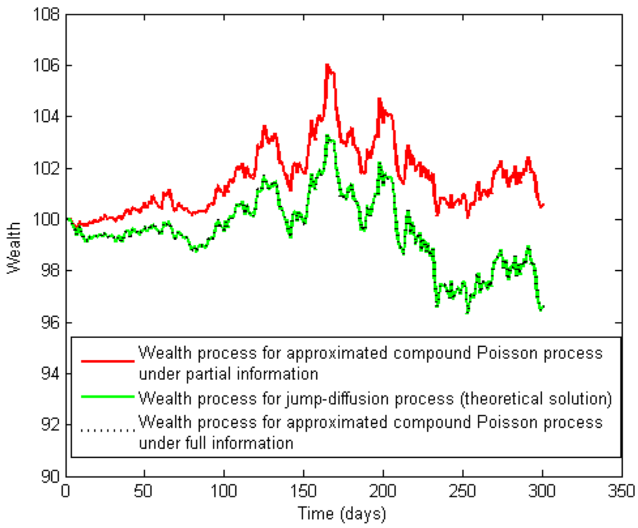

Figure 32: Wealth processes (third example)

Average error between theoretical wealth process and appoximate wealth process is 2.0363.

9 Summary of Part II

In this part of the thesis filtering problem together with portfolio optimization for jump-duffusion processes was considered. Two particular cases of jump processes were discussed: shot-noise process and compound Poisson process. A filtering problem, that contains these jump processes as noise terms, is infinite-dimensional and and cannot be solved explicitly. Therefore, an approximate way of solving this filtering problem was suggested. The idea was to approximate the shot-noise and compound Poisson processes by the Brownian motion in order to obtain a diffusion process instead of a jump-duffusion process and to apply the Kalman filtering.

The theoretical solution to the portfolio optimization problem of the jump-diffusion asset price model was derived and the verification theorem (for the case of the shot-noise driven asset price) was formulated and proved. The portfolio optimization problem for the approximated diffusion asset price model was solved by applying the results of the paper [6]. The numerical experiments for filtering and portfolio optimization were carried out.

In this part of the thesis the asset price model under consideration consists of a jump-diffusion asset price process and a diffusion process that discribes the drift of the asset price process. In a more general case the drift process can also follow a jump diffusion process (with shot-noise or compound Poisson process as a jump part). Then shot-noise and compound Poisson processes can be approximated by Brownian motion and Kalman filtering can be applied. Portfolio optimization can be carried out using the framework in [6]. The theoretical solution to the portfolio optimization problem for the case, when both unobservable and observation processes follow jump-diffusion equations (with shot-noise or compound Poisson process as a jump part) is an idea for the future research.

Part III

Filtering and Portfolio Optimization for Heston's Stochastic Volatility Model

Heston's stochastic volatility model is an asset price model, driven by Brownian motion, where the volatility of the asset price is not constant or deterministic, but itself follows a stochastic process. The drift of the asset price is also driven by the same stochastic process as the volatility. The aim of the filtering problem is to evaluate the stochastic volatility process using the observations of the asset price. The coefficients of the Heston model are not constant or linear, therefore, we have to deal with non-linear infinite-dimensional filtering, which has to be solved in an approximate way. The main idea of an approximate solution is to linearize the non-linear coefficients of both asset price and volatility equations, and then to apply Kalman filtering to the linearized model.

Portfolio optimization problem for Heston's stochastic volatility model has to be solved under full and under partial information. The case of partial information implies that the volatility process is unobservable and has to be filtered having asset prices as observations.

10 Filtering of Heston model

10.1 Heston model

Consider Heston model in [21] that describes the dynamics of the stock price process $S(t)$, $t > 0$:

$$dS_t = S_t(r + \lambda z_t)dt + S_t\sqrt{z_t}dW_t, \tag{99}$$

$$dz_t = \kappa(\theta - z_t)dt + \sigma\sqrt{z_t}dV_t, \tag{100}$$

where z_t, $t > 0$ is the stochastic volatility, r is short rate, $\lambda \in R$, $\theta > 0$ is the average volatility to which the expected value of z_t tends, $\kappa > 0$ is the rate at which z_t reverts to θ, σ is the variance of the volatility z_t. Both W_t and V_t are independent Wiener processes.

In the financial market one gets the information about the stock prices but does not know the actual values of the volatility, i.e. stock prices are observed in discrete time only. Let $S_1, S_2, ..., S_k,\ k \in N$ be the discrete observations of the stock price.

The aim of this part of the thesis is the following: having the discrete observations S_k of the stock price under the model (99)-(100) one should estimate the values of the stochastic volatility z_k.

This is the problem of non-linear filtering. In order to solve this problem, statistically linearized filter (SLF), mentioned in the introduction of [35], is applied.

Statistically linearized filter is obtained by linearizing all non-linear stochastic functions in the equations (99) and (100) and then by applying Kalman filtering.

10.2 Method of statistical linearization

This method is in detail presented in the book [18].

Let z_t and y_t be random functions:

$$y_t = \varphi(z_t), \tag{101}$$

where φ is non-linear function.

The non-linear transform (101) is to be approximated by a linearized dependence U_t between random functions z_t and y_t. Decompose the random functions z_t and y_t as follows:

$$z_t = m_z(t) + z_t^0, \tag{102}$$

$$y_t = m_y(t) + y_t^0, \tag{103}$$

where $m_z(t)$, $m_y(t)$ are expectations of z_t and y_t respectively, z_t^0, y_t^0 are centered random variables.

The function U_t, that would approximate the non-linear function φ, is assumed to be as follows:

$$U_t = \varphi_0 + k_1 z_t^0, \tag{104}$$

where φ_0 and k_1 are the coefficients which have to be determined.

Random functions are considered to be statistically equivalent if their expectations and variances are the same, given a distribution law. So, coefficients φ_0 and k_1 have to be determined in such a way that y_t and U_t are statistically equivalent.

There are two ways to determine the coefficients. The first one is to set the expectations and variances of y_t and U_t to be equal:

$$E[y_t] = E[U_t], \tag{105}$$

$$E[y_t - m_y(t)]^2 = E[U_t - m_u(t)]^2, \tag{106}$$

From (105) and (104) one obtains the value of the coefficient φ_0:

$$\varphi_0 = m_y(t). \tag{107}$$

From (106), (104) and the expression for $y(t)$ one obtains the formulae for k_1:

$$k_1 = \pm\frac{\sigma_y(t)}{\sigma_z(t)}, \tag{108}$$

where $\sigma_y(t) = \sqrt{\mathrm{Var}(y_t)}$ and $\sigma_z(t) = \sqrt{\mathrm{Var}(z_t)}$. The sign is chosen in such a way that the signs of the true and approximating functions coincide.

The second way to determine the coefficients φ_0 and k_1 is as follows:

$$E[y_t - U_t]^2 \to \min. \tag{109}$$

In this case the coefficients φ_0 and k_1 are the following:

$$\varphi_0 = m_y(t), \tag{110}$$

$$k_1 = \frac{\mathrm{Corr}(z_t, y_t)}{\mathrm{Var}(z_t)}, \tag{111}$$

One possibility to calculate the coefficients φ_0 and k_1 is to estimate the expectations and variances of the stochastic processes using empirical mean and empirical variance. Note that the distribution is not known in case of non-linear models. The only observation data that one gets in the financial market is the value of the stock prices $S_1, S_2, ..., S_k$, $k \in N$ or the stock price increments $\frac{S_2 - S_1}{S_2}, \frac{S_3 - S_2}{S_3}, ...$ The values of the stochastic volatility z_t are not known, but is is possible to simulate these values using the model (100). Based on that, one can calculate the estimates for m_z, m_y, σ_z and σ_y.

Therefore, Heston model can be approximated and linearized in the following way:

$$dS_t \approx S_t(r + \lambda z_t)dt + S_t(\varphi_0 + k_1(z_t - m_z))dW_t, \tag{112}$$

$$dz_t \approx \kappa(\theta - z_t)dt + \sigma(\varphi_0 + k_1(z_t - m_z))dV_t, \tag{113}$$

where m_z is the expected value of the random variable z_t.

10.3 Numerical results

The discrete-time version of Heston model (99)-(100) is as follows:

$$\frac{\Delta S_{k+1}}{S_k} = (r + \lambda z_k)\Delta t + \sqrt{z_k}\sqrt{\Delta t}W_k \tag{114}$$

$$\triangle z_{k+1} = \kappa(\theta - z_k)\Delta t + \sigma\sqrt{z_k}\sqrt{\Delta t}V_k. \tag{115}$$

In order to apply Kalman filtering one should first linearize all non-linear functions in the model. The terms in the drift part are linear according to the model, so one has to linearize only the square root function $\sqrt{z_k}$. In terms of the method of statistical linearization the function to be linearized is $y_k = \sqrt{z_k}$ and the approximating function is $U_t = \varphi_0 + k_1 z_t^0 = \varphi_0 + k_1(z_t - m_z)$. Denoting $\frac{\triangle S_{k+1}}{S_k}$ by s_k the discrete version of the linearized model looks as follows:

- signal (unobservable)

$$z_{k+1} = (1 - \kappa\Delta t)z_k + \kappa\theta\Delta t + (\sigma\varphi_0\sqrt{\Delta t} - \sigma k_1\sqrt{\Delta t}m_z + \sigma k_1 z_k\sqrt{\Delta t})V_k, \tag{116}$$

- observation

$$s_k = \lambda\Delta t z_k + r\Delta t + (z_k k_1\sqrt{\Delta t} + \varphi_0\sqrt{\Delta t} - m_z k_1\sqrt{\Delta t})W_k, \tag{117}$$

where $s_k = \Delta S_{k+1}/S_k$ are the returns from the stock price.

The model to be simulated is as follows:

$$\frac{\Delta S_{k+1}}{S_k} = (0.2 + 1 \cdot z_k)\Delta t + \sqrt{z_k}\sqrt{\Delta t}W_k, \tag{118}$$

$$\triangle z_{k+1} = 0.5(0.5 - z_k)\Delta t + 0.3\sqrt{z_k}\sqrt{\Delta t}V_k, \tag{119}$$

where $z_0 = 0.25$, $S_0 = 10$ and $V \sim N(0, 0.3)$, $W \sim N(0, 0.3)$. The first example follows below.

Figure 33 shows the simulation of the stochastic volatility process of Heston model and its version with linearized coefficients. One notices that the method of statistical linearization performs quite well, as both processes are very similar. Figure 34 shows the simulation of the stock price returns. The first part of the figure shows the returns of the stock prices according to the Heston model, the second one shows the linearized version. The method of statistical linearization shows good performance.

Figures 35 and 36 show the stock price process and the drift processes (true and filtered) respectively.

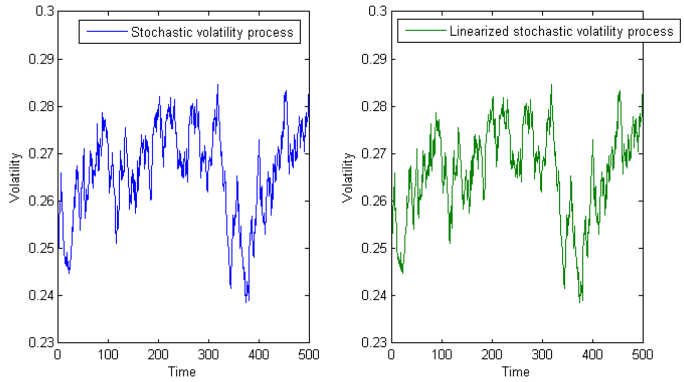

Figure 33: Stochastic volatility of the asset price: original and linearized processes (first example)

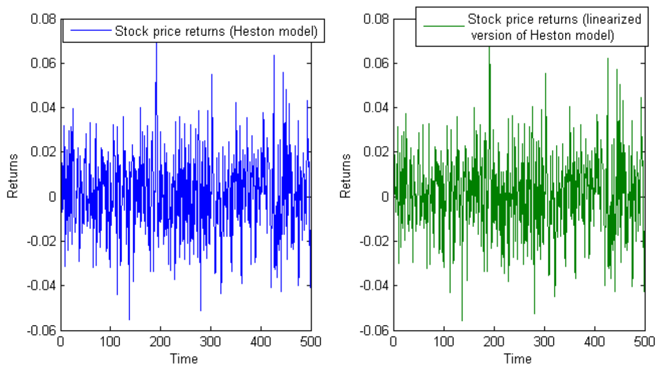

Figure 34: Simulated returns of the stock price process: Heston model and linearized version of Heston model (first example)

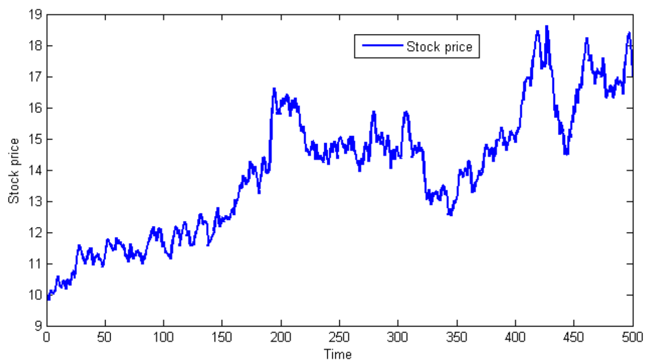

Figure 35: Stock price process (first example)

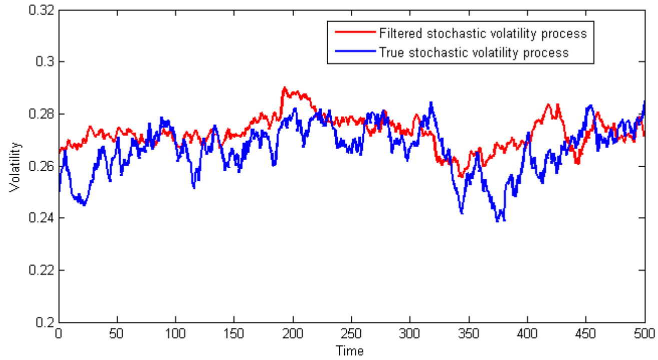

Figure 36: True and filtered stochastic volatility processes (first example)

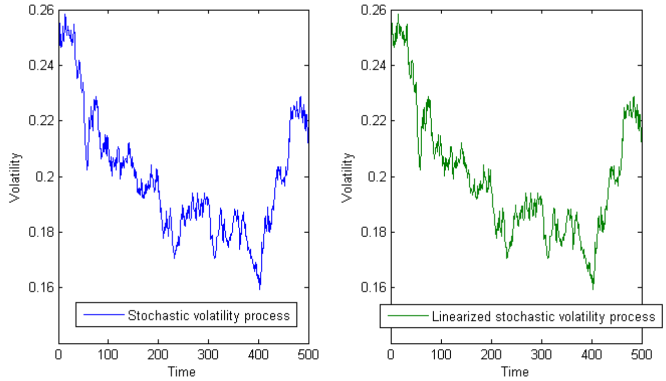

Figure 37: Stochastic volatility of the asset price: original and linearized processes (second example)

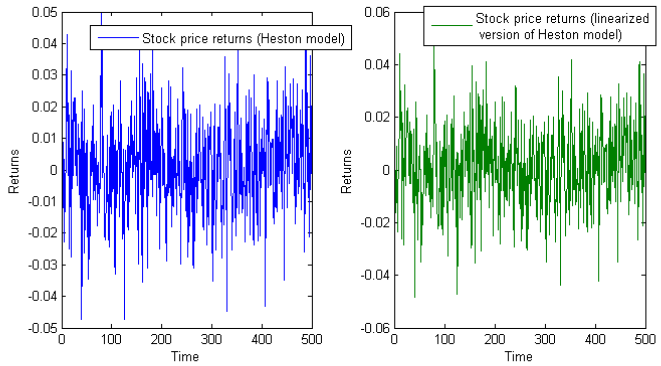

Figure 38: Simulated returns of the stock price process: Heston model and linearized version of Heston model (second example)

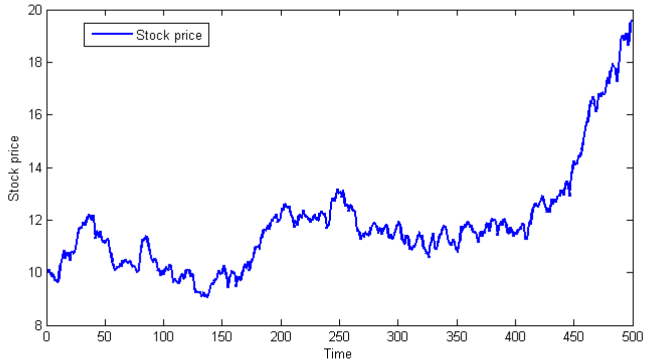

Figure 39: Stock price process (second example)

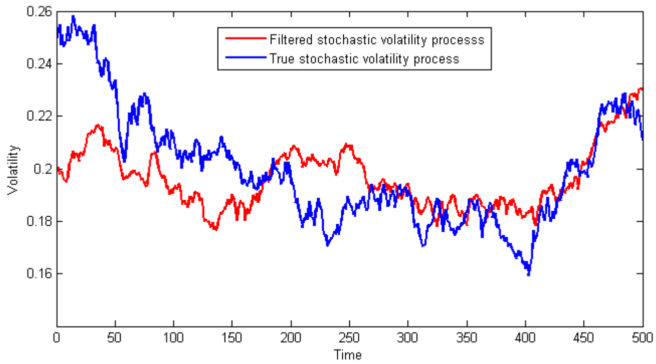

Figure 40: True and filtered stochastic volatility processes (second example)

One sees that the dynamics of the true and filtered processes are quite similar.

The second example simulates the same model and is shown in Figures 37, 38, 39 and 40.

11 Portfolio optimization for Heston model

The aim of this section is to solve the portfolio optimization problem for Heston's stochastic volatility model under full and under partial information. Under full information (both asset prices and volatility values are observable) one can apply the results of the paper [21]. The way how to solve the portfolio optimization problem under partial information (only asset prices are observable) is developed in this section.

11.1 Portfolio optimization for Heston model under full information

Consider the general portfolio optimization problem that was discussed in Section 4:

$$dS_t = S_t[\mu(z_t, t)dt + \sigma(z_t, t)dW_t],$$
$$dz_t = a(z_t, t)dt + b(z_t, t)dV_t,$$

where a, b are real valued functions, ρ is the correlation between Brownian motions W_t and V_t.

If we set $a(z_t, t) = \kappa(\theta - z_t)$, $b(z_t, t) = \sigma\sqrt{z_t}$, $\mu(z_t, t)) = r + \lambda z_t$ and $\sigma(z_t, t) = \sqrt{z_t}$, then we obtain the Heston model:

$$dS_t = S_t(r + \lambda z_t)dt + S(t)\sqrt{z_t}dW_t,$$

$$dz_t = \kappa(\theta - z_t)dt + \sigma\sqrt{z_t}dV_t.$$

Following the paper [21], we define the portfolio wealth X_t^π such that the portfolio startegy π_t denotes the fraction of wealth invested into the stock:

$$\begin{aligned} dX_t^\pi &= X_t^\pi \pi_t[\mu(z_t, t)dt + \sigma(z_t, t)dW_t] + X_t^\pi(1 - \pi_t)rdt \\ &= X_t^\pi \pi_t[(\mu(z_t, t) - r)dt + \sigma(z_t, t)dW_t] + X_t^\pi rdt \\ &= X_t^\pi[(r + (\mu(z_t, t) - r)\pi_t)dt + \pi_t\sqrt{z_t}dW_t] \\ &= X_t^\pi[(r + \lambda z_t\pi_t)dt + \pi_t\sqrt{z_t}dW_t], \end{aligned}$$

and the portfolio optimization problem is find the optimal portfolio strategy so as to maximize the expected utility of the terminal wealth X_T^π i.e.

$$\max_\pi E\Big[\frac{1}{\gamma}(X_T^\pi)^\gamma\Big].$$

Without loss of generality we set $r = 0$ as in [6].

Theorem 11.1. *(Kraft) The function f in the value function $J(t,x,z) = \frac{(x)^\gamma}{\gamma}(f(t,z))^{\frac{1-\gamma}{1-\gamma+\rho^2\gamma}}$ is real-valued and has the representation*

$$f(t,z_t) = \exp\left(\frac{\gamma}{c}r(T-t) - A_f(t,T) - B_f(t,T)z_t\right) \tag{120}$$

if

$$\frac{\gamma}{1-\gamma}\lambda\left(\frac{\kappa\rho}{\sigma} + \frac{\lambda}{2}\right) < \frac{\kappa^2}{2\sigma^2}, \tag{121}$$

where $A_f(t,T)$ is a real-valued C^1-function and $B_f(t,T) = 2\tilde{\beta}\frac{e^{\tilde{\alpha}(T-t)}-1}{e^{\tilde{\alpha}(T-t)}(\tilde{k}+\alpha)-\tilde{k}+\tilde{\alpha}}$, where $\tilde{\beta} = -\frac{1}{2c}\frac{\gamma}{1-\gamma}\lambda^2$, $\tilde{\alpha} = \sqrt{\tilde{k}^2 + 2\tilde{\beta}\sigma^2}$, $c = \frac{1-\gamma}{1-\gamma+\rho^2\gamma}$, and $\tilde{k} = \kappa - \frac{\gamma}{1-\gamma}\rho\lambda\sigma$ with $\tilde{k} > 0$.

To recall, $J(t,x,z)$ is the value function $\max_\pi E^{t,x,z}\left(\frac{1}{\gamma}(X_t^\pi)^\gamma\right)$, which is the solution to the HJB equation (40).

The candidates for the value function and for the optimal portfolio strategy are well-defined under assumption (121).

Proposition 11.2. *(Kraft) The candidate for the optimal strategy is*

$$\pi_t^* = \frac{1}{1-\gamma}\lambda - \frac{1}{1-\gamma}c\rho\sigma B_f(t,T).$$

This result can be obtained in the following way. The following HJB equation for the optimization problem has to be considered (x stands for X_t^π, z stands for z_t and π stands for π_t):

$$\max_\pi\Big(\frac{\partial}{\partial t}J(t,x,z) + x(r + (\mu(z,t) - r)\pi_t)\frac{\partial}{\partial x}J(t,x,z)$$
$$+a(z,t)\frac{\partial}{\partial z}J(t,x,z) + \frac{1}{2}x^2(\sigma(z,t))^2\pi^2\frac{\partial^2}{\partial x^2}J(t,x,z)$$
$$+\pi b(z,t)x\sigma(z,t)\rho\frac{\partial^2}{\partial xz}J(t,x,z) + \frac{1}{2}(b(z,t))^2\frac{\partial^2}{\partial z^2}J(t,x,z)\Big) = 0$$

and substitute the value function $J(t,x,z)$ (f given by (120)) into the HJB equation:

$$\max_\pi\Big\{\exp(\gamma(T-t) - cA_f(t,T) - cB_f(t,T)z)\times$$
$$\times\Big[\frac{1}{\gamma}(x^\gamma\Big[-\gamma - c\frac{dA_f(t,T)}{dt} - cz\frac{dB_f(t,T)}{dt}\Big] + x(r + (\mu(z,t) - r)\pi)\frac{\gamma}{\gamma}x^{\gamma-1}$$
$$+a(z,t)\frac{1}{\gamma}x^\gamma[-cB_f(t,T)] + \frac{1}{2}x^2(\sigma(z,t))^2\pi^2\frac{\gamma(\gamma-1)}{\gamma}x^{\gamma-2}$$

$$+b(z,t)x\sigma(z,t)\rho\frac{\gamma}{\gamma}x^{\gamma-1}(-cB_f(t,T))+\frac{1}{2}(b(z,t))^2\frac{1}{\gamma}x^{\gamma}c^2B_f^2(t,T)\Big]\Big\}=0.$$

Next, the expression in the brackets is differentiated with respect to π and the derivative is set to zero:

$$\mu(z,t)-r-b(z,t)\rho c\sigma(z,t)B_f(t,T)+\pi(\gamma-1)(\sigma(z,t))^2=0,$$

therefore

$$\pi^*(z,t)=\frac{\mu(z,t)-r-b(z,t)\rho c\sigma(z,t)B_f(t,T)}{(1-\gamma)(\sigma(z,t))^2}.$$

Substituting the coefficients of the Heston model we obtain the following optimal portfolio strategy:

$$\pi^*(z,t)=\frac{\lambda}{1-\gamma}-\frac{\rho c\sigma B_f(t,T)}{(1-\gamma)}$$

and one sees that if $\rho=0$ then

$$\pi_t^*=\lambda/(1-\gamma) \tag{122}$$

is constant.

Consider the linearized Heston model (method of statistical linearization applied):

$$dS_t\approx S_t(r+\lambda z_t)dt+S_t(\varphi_0+k_1(z_t-m_z))dW_t, \tag{123}$$

$$dz_t\approx\kappa(\theta-z_t)dt+\sigma(\varphi_0+k_1(z_t-m_z))dV_t, \tag{124}$$

then in terms of the general model we have the following: $a(z_t,t)=\kappa(\theta-z_t)$, $b(z_t,t)=\sigma(\varphi_0+k_1(z_t-m_z))$, $\mu(z_t,t))=r+\lambda z_t$ and $\sigma(z_t,t)=(\varphi_0+k_1(z_t-m_z))$, where m_z is the mean value of z_t.

The market price of risk in this case $\zeta_t=\mu(z_t,t)/\sigma(z_t,t)=\lambda z_t/(\varphi_0+k_1(z_t-m_z))$ is bounded, that means that Z is density and is well-defined.

Therefore, assuming that $\rho=0$, the optimal portfolio strategy is

$$\pi^*(t,z)=\frac{\lambda z}{(1-\gamma)(\varphi_0+k_1(z-m_z))^2}. \tag{125}$$

11.2 Portfolio optimization for Heston model under partial information

The term 'partial information' implies that only the asset prices S_t are observable but not the stochastic volatility process z_t. For the portfolio optimization problem one has to filter the unobservable variable z_t. Recall that Heston model looks as follows:

$$ds_t=\frac{dS_t}{S_t}=(r+\lambda z_t)dt+\sqrt{z_t}dW_t, \tag{126}$$

$$dz_t = \kappa(\theta - z_t)dt + \sigma\sqrt{z_t}dV_t. \tag{127}$$

Notice, that because the diffusion coefficients in both SDEs are not constant, we cannot use Kalman filtering. Also for the case of the statistically linearized version (123)-(124) of the Heston model Kalman filtering in continuous time cannot be applied, because diffusion coefficients are linear functions but not constant. In discrete time one could apply statistical linearization to the original Heston model and then apply Kalman filtering in order to estimate the unobservable variable z_t.

But the extended Kalman filter (subsubsection 1.5.1) can be applied in order to estimate the unobservable variable z_t in continuous time. Consider the following general form of signal and observation equations (126)-(127):

$$dz_t = f(z_t)dt + \sigma(z_t)dV_t,$$

$$ds_t = h(z_t)dt + g(z_t)dW_t,$$

where V_t and W_t are independent Brownian motions, and f, σ, h, g are linear or non-linear functions. In the case of Heston model (126)-(127) these non-linear functions are the following:

$$f(z_t) = \kappa(\theta - z_t),\ \sigma(z_t) = \sigma\sqrt{z_t},\ h(z_t) = r + \lambda z_t,\ g(z_t) = \sqrt{z_t}.$$

The SDEs with non-linear coefficients can be approximated by means of Taylor expansion:

$$dz_t \approx (f'(\bar{z}_t)(z_t - \bar{z}_t) + f(\bar{z}_t))dt + \sigma(\bar{z}_t)dV_t,$$

$$ds_t \approx (h'(\bar{z}_t)(z_t - \bar{z}_t) + h(\bar{z}_t))dt + g(\bar{z}_t)dW_t,$$

where f' and h' are the derivatives of f and h respectively and $\bar{z}$ is the solution of the following ODE:

$$\frac{d\bar{z}_t}{dt} = f(\bar{z}_t),\ \bar{z}_0 = z_0. \tag{128}$$

The standard Kalman filter can be applied to the Taylor approximated signal and observation SDEs.

Taking into account that $f(z_t) = \kappa(\theta - z_t)$, the solution to (128) is as follows:

$$\bar{z}_t = (z_0 - \theta)e^{-\kappa t} + \theta.$$

Further, $f'(z) = -\kappa$ and $h'(z) = \lambda$, therefore:

$$dz_t \approx (-\kappa(z_t - \bar{z}_t) + \kappa(\theta - \bar{z}_t))dt + \sigma(\bar{z}_t)dV_t,$$

$$ds_t \approx (\lambda(z_t - \bar{z}_t) + r + \lambda\bar{z}_t)dt + g(\bar{z}_t)dW_t$$

or

$$\begin{aligned} dz_t &\approx \kappa(\theta - z_t)dt + \sigma(\bar{z}_t)dV_t = \kappa(\theta - z_t)dt + \sigma\sqrt{\bar{z}_t}dV_t \\ &= f(z_t)dt + \sigma\sqrt{\bar{z}_t}dV_t, \qquad (129) \\ ds_t &\approx (r + \lambda z_t)dt + g(\bar{z}_t)dW_t = (r + \lambda z_t)dt + \sqrt{\bar{z}_t}dW_t \\ &= h(z_t)dt + \sqrt{\bar{z}_t}dW_t. \qquad (130) \end{aligned}$$

Now one can refer to the paper [6] and find the optimal portfolio strategy under partial information. The conditional expectation $\hat{z}_t = E[z_t|F_t^S]$ (F_t^S is the filtration based on the observations of asset price process) of the drift process z_t satisfies

$$d\hat{z}_t = \kappa(\theta - \hat{z}_t)dt + (\Omega_t + R)(g(\bar{z}_t))^{-2}(ds_t - (\lambda\hat{z}_t + r)dt),$$

where Ω_t is the conditional covariance martix which satisfies the Ricatti equation

$$\frac{d\Omega_t}{dt} = \sigma(\bar{z}_t) - 2\kappa\Omega_t - (\Omega_t + \rho)(g(\bar{z}_t))^{-2}(\Omega_t + \rho).$$

In the above filtering equations $R = \rho g(\bar{z}_t)\sigma(\bar{z}_t)$.

The wealth process X_t^π is as follows (taking into account the linearized model (129)-(130)):

$$\begin{aligned} dX_t^\pi &= X_t^\pi \pi_t[h(z_t)dt + \sqrt{\bar{z}_t}dW_t] + X_t^\pi(1 - \pi_t)rdt \\ &= X_t^\pi \pi_t[(h(z_t) - r)dt + \sqrt{\bar{z}_t}dW_t] + X_t^\pi rdt \\ &= X_t^\pi[(r + (h(z_t) - r)\pi_t)dt + \sqrt{\bar{z}_t}dW_t] \\ &= X_t^\pi[(r + \lambda z_t\pi_t)dt + \pi_t\sqrt{\bar{z}_t}dW_t], \end{aligned}$$

and the portfolio optimization problem is as before:

$$\max_\pi E\Big[\frac{1}{\gamma}(X_T^\pi)^\gamma\Big].$$

Without loss of generality we can set $r = 0$ as in [6].

Consider the HJB equation for our optimization problem ((x stands for X_t^π, z stands for z_t and π stands for π_t)):

$$\max_\pi\Big(\frac{\partial}{\partial t}J(t,z,x) + x(r + \lambda\hat{z}\pi)\frac{\partial}{\partial x}J(t,z,x) + \kappa\theta\frac{\partial}{\partial z}J(t,z,x)$$

$$-\kappa\hat{z}\frac{\partial}{\partial z}J(t,z,x) + \frac{1}{2}x^2(g(\bar{z}_t))^2\pi^2\frac{\partial^2}{\partial x^2}J(t,z,x)$$

$$+\pi x(R+\Omega_t)\frac{\partial^2}{\partial xz}J(t,z,x)+\frac{1}{2}(\Omega_t+R)^2(g(\bar{z}_t))^{-2}\frac{\partial^2}{\partial z^2}J(t,z,x)\Big)=0$$

and substitute the value function $J(t,x,\hat{z})=\frac{(x)^\gamma}{\gamma}\exp\left(\hat{z}^2A_t+\hat{z}B_t+C_t\right)$ into the HJB equation:

$$\max_\pi\Big\{\frac{1}{\gamma}\Big\{\hat{z}^2\frac{dA_t}{dt}+\hat{z}\frac{dB_t}{dt}+\frac{dC_t}{dt}\Big\}+\frac{1}{2}\pi^2(g(\bar{z}_t))^2(\gamma-1)$$

$$+\frac{1}{2\gamma}(\Omega_t+R)(g(\bar{z}_t))^{-2}(4\hat{z}^2A_t^2+2\hat{z}A_tB_t+2A_t+B_t^2)+\pi(\Omega_t+R)(2\hat{z}A_t+B_t)$$

$$+\pi\lambda\hat{z}+r+\frac{\kappa}{\gamma}\theta(2\hat{z}A_t+B_t)-\frac{\kappa}{\gamma}\hat{z}(2\hat{z}A_t+B_t)\Big\}=0.$$

Proposition 11.3. *Differentiating the above expression with respect to π, one obtains the following optimal portfolio strategy:*

$$\pi^*(t,\hat{z})=\frac{\lambda\hat{z}+(\Omega_t+R)(2\hat{z}A_t+B_t)}{(1-\gamma)(g(\bar{z}_t))^2},\tag{131}$$

where

$$\frac{dA_t}{dt}=-\frac{\gamma(\lambda+2(\Omega_t+R)A_t)^2}{2(1-\gamma)(g(\bar{z}_t))^2}-\frac{2(\Omega_t+R)^2A_t}{(g(\bar{z}_t))^2}+2\kappa A_t,$$

$$\frac{dB_t}{dt}=-\frac{\gamma}{2(1-\gamma)(g(\bar{z}_t))^2}(3\lambda r+4r(\Omega_t+R)A_t+2\lambda(\Omega+R)B_t-4(\Omega_t+R)^2A_tB_t)-$$

$$-\frac{2(\Omega_t+R)^2A_tB_t}{(g(\bar{z}_t))^2}-2\kappa\theta A_t+\kappa B_t,$$

$$\frac{dC_t}{dt}=-\frac{\gamma}{2(1-\gamma)(g(\bar{z}_t))^2}(r^2+3(\Omega_t+R)rB_t+(\Omega_t+R)^2B_t^2)-$$

$$-\kappa\theta B_t-\frac{(\Omega_t+R)^2(2A_t+B_t^2)}{2(g(\bar{z}_t))^2}.$$

11.3 Numerical experiments

The model to be simulated looks as follows:

$$\frac{dS_t}{S_t}=(0+1\cdot z_t)dt+\sqrt{z_t}dW_t,$$

$$dz_t=0.5(0.5-z_t)dt+0.3\sqrt{z_t}dV_t,$$

where $dV_t\sim N(0,0.3)$ and $dW_t\sim N(0,0.5)$.

We present several figures below. On figure 41 the asset price dynamics is shown. Figure 42 shows the true stochastic volatility process z_t, the linearized stochastic volatility process and filter approximation, obtained by using the extended Kalman filter. One cannotice that both approimations give very similar results.

Figure 43 represents the optimal trading strategies for Heston model under full information (122), under partial information (131), and for statistically linearized Heston model (125).

Figure 44 shows the wealth dynamics under three different trading strategies presented in Figure 43.

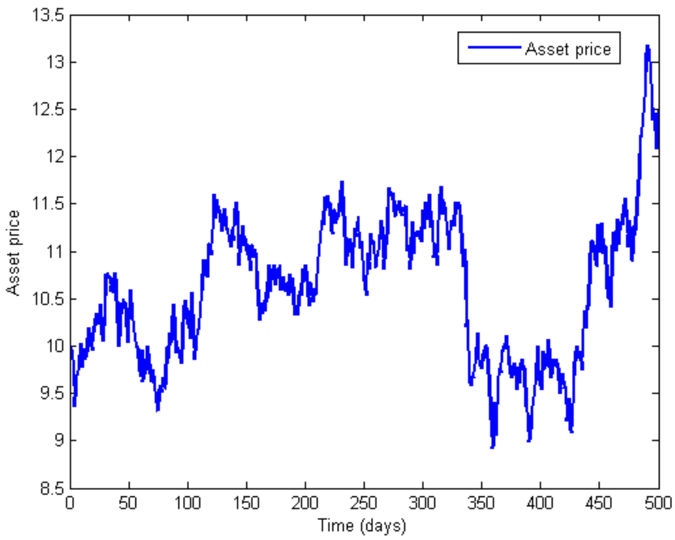

Figure 41: Stock price process driven by Heston model

As a short summary let us discuss the average errors in the presented example. The average error between the true process z_t and the extended Kalman filter is 0.0198.

Average errors between the optimal portfolio strategy under full information and the strategies under partial information and statistical linearization are 0.7620 and 0.0558 respectively.

And at last, the average errors between the wealth, obtained by applying the optimal strategy under full information, and the wealth obtained under partial information and statistical linearization are 4.9344 and 0.3496 respectively.

The error between the optimal and approximated portfolio wealth can be estimated by the loss in expected utility, i.e. $E[U(X_T^{\pi})] - E[U(X_T^{approx})]$, where X_T^{π} is the optimal portfolio wealth and X_T^{approx} is the approximated portfolio wealth at terminal time T.

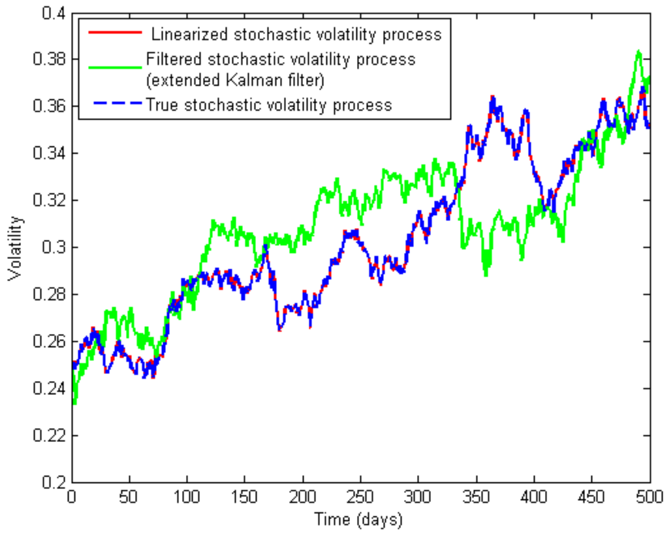

Figure 42: True and filtered stochastic volatility process of the stock price driven by Heston model

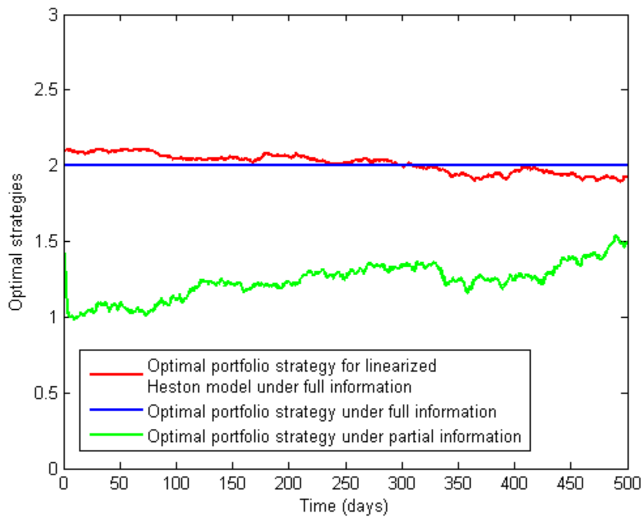

Figure 43: Optimal portfolio strategies for Heston model (under full and partial information and statistically linearized strategy)

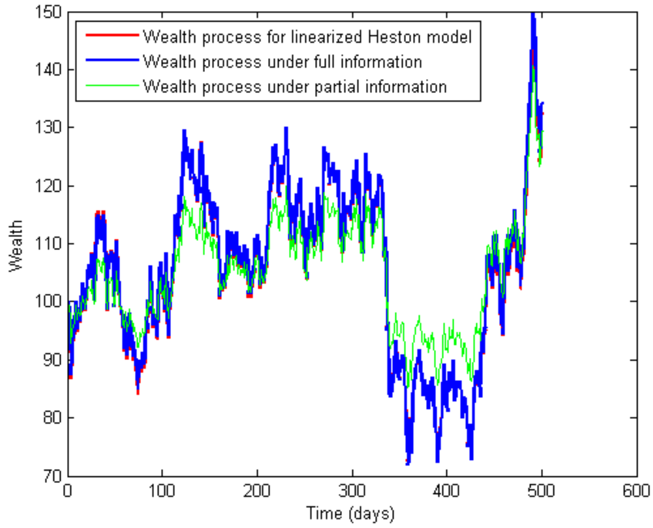

Figure 44: Wealth processes for Heston model under three different optimal portfolio strategies

The loss in the utility beween the portfolio wealth, obtained by applying the optimal strategy under full information, and the wealth obtained under partial information and statistical linearization are 0.0062 and $7,56 \cdot 10^{-4}$ respectively. The utility function is $U(x) = x^{\gamma}/\gamma$, where $\gamma = 0.5$. The expected value is estimated as an average over 1000 simulations.

12 Filtering of Heston model with switching of states

In this section the Heston model will be extended. The extension consists of adding the switching of states of a certain parameter of the model. First, the switching in the volatility of the stochastic volatility z_t will be considered, and then, the switching in the parameter λ in the drift of the asset price will be considered. The switching of states of a certain parameter corresponds to good and bad market conditions. For example, when the volatility of the variable z_t is big (bad market condition), then the estimate of z_t (filter) may deviate stronger from the original process z_t, than in the case when the volatility of z_t is small, i.e. good market condition.

The portfolio optimization problem for the Heston model with switching of states will not be considered in this section and is left for the future research.

12.1 Switching in the stochastic volatility process

Consider the following modification of the Heston model:

$$ds_t = \frac{dS_t}{S_t} = (r + \lambda z_t)dt + \sqrt{z_t}dW_t, \tag{132}$$

$$dz_t = \kappa(\theta - z_t)dt + (\sigma_1 1_{M_t=1} + \sigma_2 1_{M_t=2})\sqrt{z_t}dV_t, \tag{133}$$

In this case it is assumed that the volatility coefficient σ of the stochastic volatility process z_t is not always the same but can switch from one state to another, respectively σ_1 and σ_2. The switching happens according to the variable M_t which is a Markov chain with $K = 2$ states. So we have a Markov Switching model. The filtering of the state of M_t is based on the observations of the variance z_t. But z_t is not observed directly, one observes only the stock prices S_t.

In this case it would be reasonable to combine Kalman filtering and filtering of the Markov Switching model.

12.1.1 Statistical inference for Markov Switching model in the stochastic volatility process

Let us assume that the parameters in (133) are known and also the transition probabilities $A = (a_{ij})$, $i, j = 1, ..., K$ of the Markov chain M_t are known.

Inference about the state of M_t at time t given information $F_t^z = (z_1, ..., z_t)$ is defined by the probability distribution $P(M_t = l|F_t^z)$, where $l = 1, ..., K$.

The filtering algorithm is considered in detail in the book [12]. This algorithm consists of recursive performance of two steps: prediction step and filtering step. In order to obtain the inference about the state of the Markov chain M_t at time t one needs the filtered probabilities $P(M_{t-1} = n|F_{t-1}^z)$ of the previous step (time $t-1$) and the current value z_t:

- First step $t = 1$

  $$P(M_1 = l|F_0^z) = \sum_{n=1}^{K} a_{nl} P(M_0 = n),$$

  where $P(M_0 = n)$ is the initial distribution of the filtering problem. This initial distribution can be chosen in many ways. If the transition probability matrix is known, then one can determine the stationary distribution of the Markov chain and take it as the initial distribution $P(M_0 = n)$.

- Prediction of the state of the Markov chain:

  $$P(M_t = l|F_{t-1}^z) = \sum_{n=1}^{K} a_{nl} P(M_{t-1} = n|F_{t-1}^z). \tag{134}$$

  This formula gives the 'prior' distribution of S_t given the information up to time $t-1$;

- Filtering of the state of the Markov chain:

  $$P(M_t = l|F_t^z) = \frac{p(z_t|M_t = l, F_{t-1}^z) P(M_t = l|F_{t-1}^z)}{p(z_t|F_{t-1}^z)}, \tag{135}$$

  where

  $$p(z_t|F_{t-1}^z) = \sum_{n=1}^{K} p(z_t|M_t = n, F_{t-1}^z) P(M_t = n|F_{t-1}^z). \tag{136}$$

  Formula 134 allows to improve the prediction step having the information from the current observation z_t.

The probability density function $p(z_t|M_t = l)$ can be determined only if the functions in (133) are linearized and (133) is discretized. In this case the density will be Gaussian.

But in order to estimate the state of the Markov chain M_t the volatility z_t should be first estimated through the stock prices S_t. This is done by first linearizing and discretizing the processes (133)-(132) and then by applying discrete Kalman filtering. In order to linearize the processes (133)-(132) one applies the method of statistical linearization, discussed in subsection 13.2. The linearized continuous-time version of the model (133)-(132) is

$$ds_t = \frac{dS_t}{S_t} \approx (r + \lambda z_t)dt + (\varphi_0 + k_1(z_t - m_z))dW_t,$$

$$dz_t \approx \kappa(\theta - z_t)dt + (\sigma_1 1_{M_t=1} + \sigma_2 1_{M_t=2})(\varphi_0 + k_1(z_t - m_z))dV_t,$$

Before the algorithm starts one has to set the transition matrix A of the Markov chain M_t and its stationary distribution π. Also the initial state of the Markov chain $l = 1$ has to be specified. So, summing up, one gets the following algorithm on the n^{th}, $n = 1, ..., N$ iteration:

- Kalman filtering
  - filter equations;

$$z_k^* = (1 - \kappa\Delta t)\hat{z}_{k-1} + \kappa\theta\Delta t,$$

$$\hat{z}_k = z_k^* + K_k\Big\{\Delta s_k - r\Delta t - \lambda\Delta t z_k^*\Big\};$$

  - variance of the estimation error and the coefficient K;

$$P_k^* = (1 - \kappa\Delta t)^2\tilde{P}_{k-1} + \sigma_l^2\Delta t(\varphi_0 + k_1(z_{k-1}^* - m_z))^2 Q,$$

$$K_k = P_k^*\lambda\Delta t\Big\{(\lambda\Delta t)^2 P_k^* + (\varphi_0 + k_1(z_k^* - m_z))^2 R\Big\}^{-1},$$

$$\tilde{P}_k = P_k^* - K_k\lambda\Delta t P_k^*,$$

  where φ_0 and k_1 are the coefficients of the statistical linearization method, $E[V_t^2] = Q$, $E[W_t^2] = R$, the term Δt is due to the discretization of the linearized version of the model (132)-(133) and l is the current state of the Markov Chain M_t;

- filtering of switching Markov model
  - prediction step

$$P(M_t = l|F_{t-1}^z) = \sum_{n=1}^{2} a_{nl}P(M_{t-1} = n|F_{t-1}^z),$$

  - fitering step

$$P(M_t = l|F_t^z) = \frac{p(z_t|M_t = l, F_{t-1}^z)P(M_t = l|F_{t-1}^z)}{p(z_t|F_{t-1}^z)},$$

where

$$p(z_t|F^z_{t-1}) = \sum_{n=1}^{2} p(z_t|M_t = n, F^z_{t-1})P(M_t = n|F^z_{t-1}),$$

$$l = \arg\max_{i=1,2} P(M_t = i|F^z_t).$$

One has to repeat the algorithm as many iterations as necessary, each time with the updated value for the state l of the Markov chain M_t.

12.1.2 Numerical simulations of switching in stochastic volatility process

For this example the Markov chain with the following transition probabilities matrix A was chosen:

$$A = \begin{pmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{pmatrix}$$

The states of the Markov chain M_t are the following: in the state $M_t = 1$ the volatility term is $\sigma_1 = 0.3$, in the state $M_t = 2$ the volatility term is $\sigma_2 = 1$. The other parameters of the model are the same as in numerical examples in Subsection 10.3. Further, the simulated figures are shown.

The first example consists of Figures 45, 46, 47, 48.

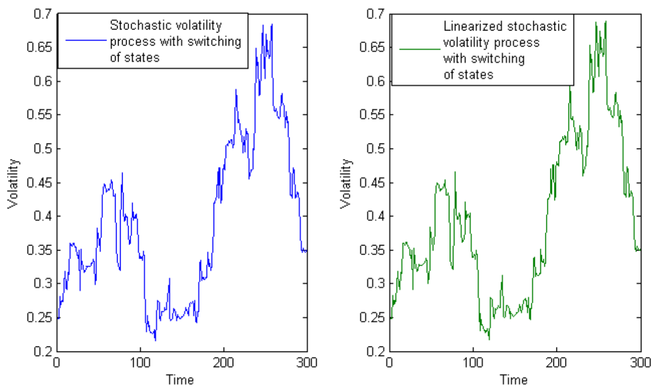

Figure 45: Stochastic volatility (with switching of states) of the asset price: original and linearized processes (first example)

One can notice, that due to the switching of states of the diffusion coefficient in the stochastic volatility process, there are abrupt jumps in the volatility process, as shown in Figures 45 and 49.

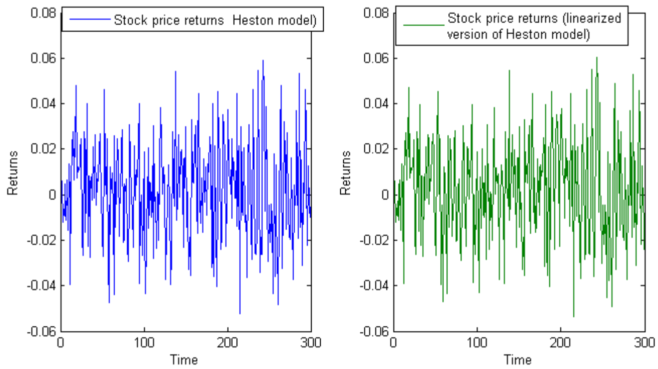

Figure 46: Simulated returns of the stock price process: Heston model and linearized version of Heston model (first example)

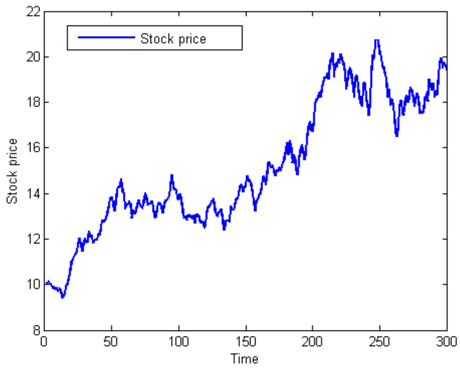

Figure 47: Stock price process (first example)

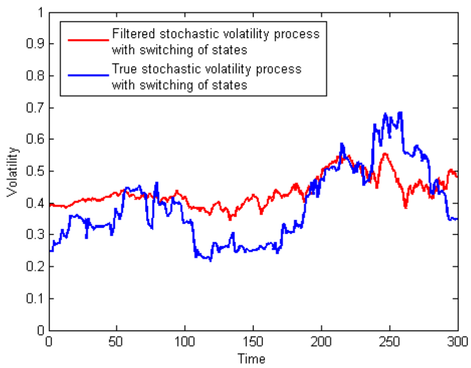

Figure 48: True and filtered stochastic volatility processes with switching of states (first example)

In the next example the transition matrix A of the Markov chain M_t is as in the previous example and the states of the Markov chain M_t are the following: in the state $M_t = 1$ the the volatility term is the following: $\sigma_1 = 0.3$, in the state $M_t = 2$ the volatility term is given by $\sigma_2 = 2$. The second example consists of Figures 49, 50, 51, 52.

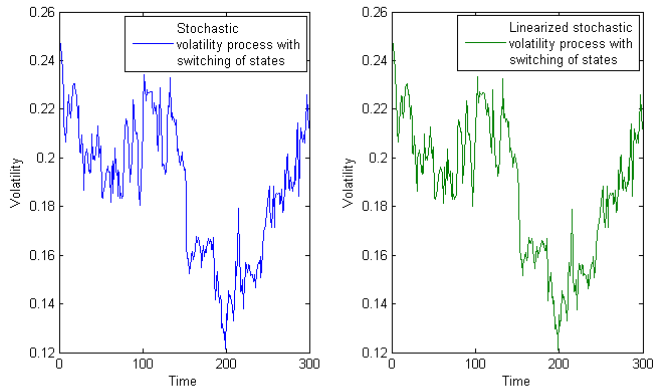

Figure 49: Stochastic volatility (with switching of states) of the asset price: original and linearized processes (second example)

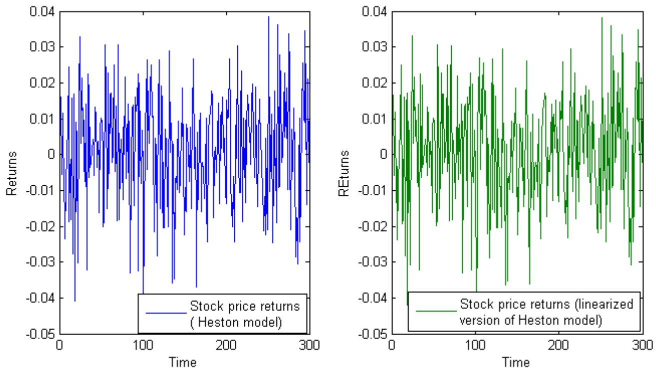

Figure 50: Simulated returns of the stock price process: Heston model and linearized version of Heston model (second example)

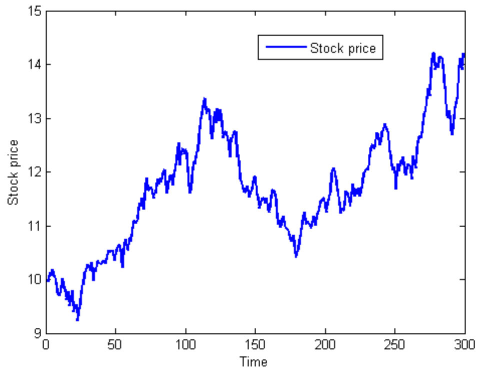

Figure 51: Stock price process (second example)

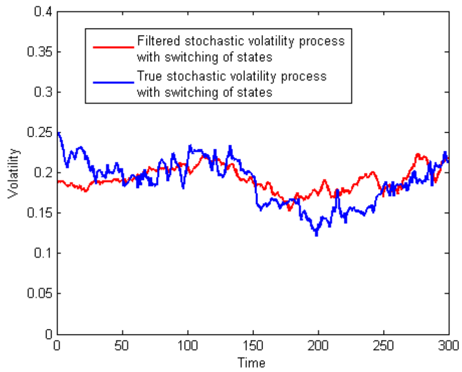

Figure 52: True and filtered stochastic volatility processes with switching of states (second example)

12.2 Switching of states in the asset price process

Consider another modification of the Heston model:

$$ds_t = \frac{dS_t}{S_t} = (r + (\lambda_1 \cdot 1_{M_t=1} + \lambda_2 \cdot 1_{M_t=2})z_t)dt + S_t\sqrt{z_t}dW_t, \tag{137}$$

$$dz_t = \kappa(\theta - z_t)dt + \sigma\sqrt{z_t}dV_t, \tag{138}$$

In this case it is assumed that the parameter λ is not always the same but can switch from one state to another. The switching happens according the the variable M_t which is a Markov chain with $K = 2$ states. So we have a Markov Switching model. Unlike the previous case, the filtering of the state of M_t is based on the observations of the variance S_t. But S_t itself depends on the unobservable variable z_t.

As in the case above it would also be reasonable to combine Kalman filtering and filtering of the Markov Switching model, but in a slightly different way.

12.2.1 Statistical inference for Markov Switching model in the asset price process

Assume that the parameters in (137) are known and also the transition probabilities $A = (a_{ij})$, $i, j = 1, ..., K$ of the Markov chain M_t are known.

Inference about the state of M_t at time t given information about asset prices $F_t^S = (S_1, ..., S_t)$ is defined by the probability distribution $P(M_t = l|F_t^S)$, where $l = 1, ..., K$.

Filtering Algorithm is very similar to the filtering algorithm from the previous subsection:

- First step $t = 1$

  $$P(M_1 = l|F_0^S) = \sum_{n=1}^{K} a_{nl}P(M_0 = n),$$

  where $P(M_0 = k)$ is the initial distribution of the filtering problem. This initial distribution can be chosen in many ways. If the matrix of transition probabilities is known, then one can determine the stationary distribution of the Markov chain and take it as the initial distribution $P(M_0 = k)$.

- Prediction of the state of the Markov chain:

  $$P(M_t = l|F_{t-1}^S) = \sum_{n=1}^{K} a_{nl}P(M_{t-1} = n|F_{t-1}^S), \tag{139}$$

- Filtering for the state of the Markov chain:

$$P(M_t = l|F_t^S) = \frac{p(s_t|M_t = l, F_{t-1}^S)P(M_t = l|F_{t-1}^S)}{p(s_t|F_{t-1}^S)}, \tag{140}$$

where

$$p(s_t|F_{t-1}^S) = \sum_{n=1}^{K} p(s_t|M_t = k, F_{t-1}^S)P(M_t = n|F_{t-1}^S). \tag{141}$$

Instead of stock prices S_t, it would be reasonable to consider increments s_t. The probability density function $p(s_t|M_t = l)$ can be determined only if the functions in (138) are linearized and discretized. In this case the density will be Gaussian.

In order to estimate the state of the Markov chain M_t the volatility z_t should be first estimated through the increments of the stock prices s_t. This is done by first linearizing and discrertizing the processes (138)-(137) and then by applying Kalman filtering.

In order to linearize the processes (138)-(137) one applies the method of statistical linearization, discussed in subsection 13.2. The linearized continuous-time version of the model (138)-(137) is

$$ds_t = \frac{dS_t}{S_t} \approx (r + (\lambda_1 \cdot 1_{M_t=1} + \lambda_2 \cdot 1_{M_t=2})z_t)dt + (\varphi_0 + k_1(z_t - m_z))dW_t,$$

$$dz_t \approx \kappa(\theta - z_t)dt + \sigma(\varphi_0 + k_1(z_t - m_z))dV_t,$$

Before the algorithm starts one has to set the transition matrix A of the Markov chain M_t and its stationary distribution π. Also the initial state of the Markov chain $l = 1$ has to be specified. So, summing up, one gets the following algorithm on the n^{th}, $n = 1, ..., N$ iteration:

- Kalman filter
  - filter equations;

$$z_k^* = (1 - \kappa\Delta t)\hat{z}_{k-1} + \kappa\theta\Delta t,$$

$$\hat{z}_k = z_k^* + K_k\Big\{\Delta s_k - r\Delta t - \lambda_l \Delta t z_k^*\Big\};$$

  - variance of the estimation error and the coefficient K;

$$P_k^* = (1 - \kappa\Delta t)^2\tilde{P}_{k-1} + \sigma^2\Delta t(\varphi_0 + k_1(z_{k-1}^* - m_z))^2 Q,$$

$$K_k = P_k^*\lambda_l\Delta t\Big\{(\lambda_l\Delta t)^2 P_k^* + (\varphi_0 + k_1(z_k^* - m_z))^2 R\Big\}^{-1},$$

$$\tilde{P}_k = P_k^* - K_k \lambda_l \Delta t P_k^*,$$

where φ_0 and k_1 are the coefficients of the statistical linearization method, $E[V_t^2] = Q$, $E[W_t^2] = R$, the term Δt is due to the discretization of the linearized version of the model (137)-(138) and l is the current state of the Markov Chain M_t;

- Switching Markov model filtering
  - prediction step

$$P(M_t = l|F_{t-1}^S) = \sum_{n=1}^{2} a_{nl} P(M_{t-1} = n|F_{t-1}^S),$$

  - filtering step

$$P(M_t = l|F_t^S) = \frac{p(s_t|M_t = l, F_{t-1}^S)P(M_t = l|F_{t-1}^S)}{p(s_t|F_{t-1}^S)},$$

where

$$p(s_t|F_{t-1}^S) = \sum_{n=1}^{2} p(s_t|M_t = n, F_{t-1}^S)P(M_t = n|F_{t-1}^S),$$

$$l = \arg\max_{i=1,2} P(M_t = i|F_t^S).$$

One has to repeat the algorithm as many iterations as necessary, each time with the updated value for the state l of the Markov chain M_t.

Therefore, one sees that the filtering algorithm is very similar to the filtering algorithm described in the previous subsection.

12.2.2 Numerical simulations of switching in the asset price process

For this example the Markov chain with the following transition probabilities matrix A was chosen:

$$A = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}$$

The states of the Markov chain M_t are the following: in the state $M_t = 1$ the parameter λ in (137) equals $\lambda_1 = 1$, in state $M_t = 2$ the parameter λ equals $\lambda_2 = 3$. The other parameters of the model are the same as in numerical examples in Subsection 10.3. The simulated numerical experiment is shown in Figures 53, 54, 55, 56.

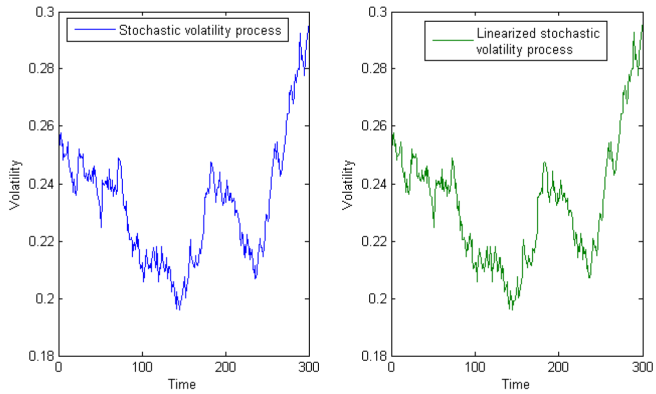

Figure 53: Stochastic volatility (with switching of states) of the asset price: original and linearized processes (third example)

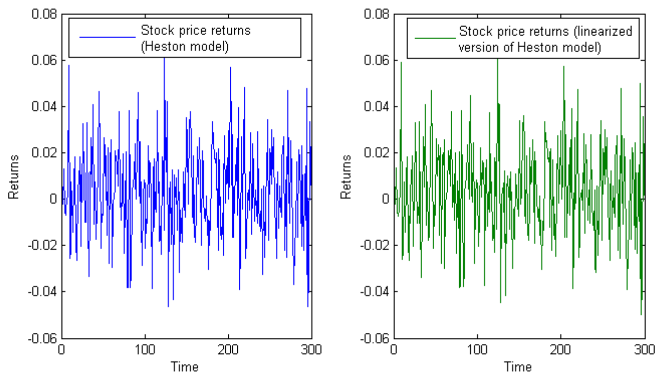

Figure 54: Simulated returns of the stock price process: Heston model and linearized version of Heston model (third example)

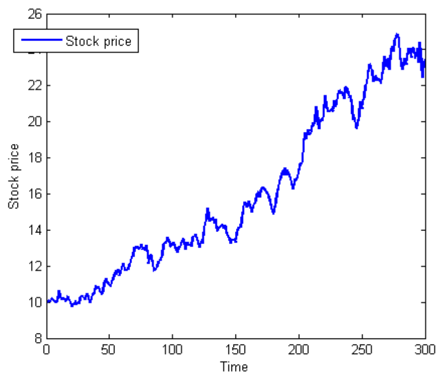

Figure 55: Stock price process (third example)

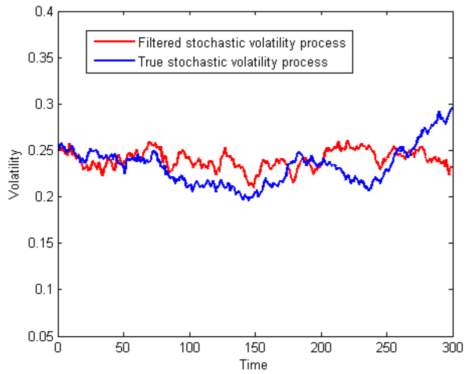

Figure 56: True and filtered stochastic volatility processes with switching of states (third example)

13 Summary of Part III

In the last part of the thesis filtering and portfolio optimization for Heston's stochastic volatility model was considered. For solving the non-linear filtering problem two approximation methods were applied: the method of statistical linearization together with Kalman filtering and the extended Kalman filter. Both methods give quite similar results. The extended Kalman filter can be further applied in order to solve the portfolio optimization problem under partial information.

14 Conclusion

In this thesis the problem of infinite-dimensional non-linear filtering with applications to finance was considered. Two different models were discussed in details:

- asset price model, driven by Brownian motion and shot-noise process and the drift of the asset price equation modelled by Ornstein-Uhlenbeck process;
- Heston's stochastic volatility model.

The natural application of the filtering results for market models is portfolio optimization. The problem of optimizing the portfolio was solved in the thesis under different assumptions (full and partial information).

The existing results and papers ([15], [28], [32]) show that it is quite difficult to solve the non-linear infinite-dimensional filtering problem explicitly. Therefore, approximate solutions were derived.

If the asset price is not only influenced by Brownian motion, but also by shot-noise process, then the stochastic drift of the asset price cannot be filtered explicitly, because the filter would be infinite-dimensional. Instead, it was proposed to approximate the shot-noise process by Brownian motion as in [7] and apply Kalman filtering to the approximated model. The portfolio optimization problem was solved for the approximated model (under full and partial information) according to [6]. A verification theorem for the framework of [6] was formulated and proved under suitable assumptions. An important result of this thesis is a theoretical solution to the portfolio optimization problem for the unapproximated (shot-noise driven) model. A verification theorem for this theoretical result was formulated. In the real market one cannot obtain the theoretical solution to the portfolio optimization problem (because drift is unobservable), but only the solution for the approximated model under partial information (the drift process filtered). The simulations show that the approximated optimal solution is still quite good.

Also, the case of compound Poisson process as the noise term was considered in the thesis. Instead of the shot-noise process one can consider a compound Poisson process, which can also be approximated by a Brownian motion, following [7]. Portfolio optimization was carried out under full and partial information, and also a theoretical solution to this portfolio optimization problem was derived.

Heston's stochastic volatilty model is quite often used in practice. Therefore, for practical needs one has to know how to filter the unobservable stochastic volatility process. This can be done in two ways: linearize all non linear coefficients in Heston model and apply Kalman filtering or apply the extended Kalman filter. For the purposes of portfolio optimization under partial information one uses the filter results, obtained by applying extended Kalman filter. Both was done after discretizing the model.

References

[1] T. Altmann, T. Schmidt, W. Stute, "A shot noise model for financial assets", International Journal of Theoretical and Applied Finance, 2008

[2] D. Applebaum, "Lévy Processes and Stochastic Calculus", Cambridge University Press, New York, 2009

[3] N. Bäuerle, U. Rieder, "Portfolio optimization with jumps and unobservable intensity process", Mathematical Finance, Vol. 17, No. 2, pp. 205-224, 2007

[4] A. Bain, D. Crisan, "Fundamentals of Stochastic Filtering", Springer Science+Business Media, LLC 2009

[5] K. Brammer, G. Siffling, "Kalman-Bucy Filters", Artech House, Inc, Norwood, MA 1989

[6] S. Brendle, "Portfolio selection under incomplete information", Stochastic Processes and their Applications 116, pp. 701-723, 2006

[7] A. Dassios, J.-W. Jang, "Kalman-Bucy filtering for linear systems driven by the Cox process with shot noise intensity and its application to the pricing of reinsurance contracts", Journal of Applied Probability 42, 93-107, 2005

[8] G.B. Di Masi, W.J. Runggaldier, "Continuous-time approximations for the nonlinear filtering problem", Applied Mathematics and Optimization 7, pp. 233-245, 1981

[9] G.B. Di Masi, W.J. Runggaldier, "On robust approximations in nonlinear filtering", Stochastic Differential Systems, Lect. Notes in Control and Info, Sci. 43, Springer-Verlag, 1982

[10] J.E. Figueroa-Lopez, "Jump-diffusion models driven by lévy processes", Chapter in Handbook of Computational Finance, Springer-Verlag, pp. 61-88, 2012

[11] A. Friedman, "Stochastic Differential Equations and Applications", Vols. 1 and 2, Academic Press, New York, 1976

[12] Sylvia Frühwirth-Schnatter, "Finite Mixture and Markov Switching Models", Springer, 2008

[13] R.M. Gaspar, T. Schmidt, "Quadratic models for portfolio credit risk with shot-noise effects", SSE/EFI Working Paper Series in Economics and Finance No. 616, November 2005

[14] M.Hahn, W. Putschögl, J. Sass, "Portfolio optimization with nonconstant volatility and partial information", Brazilian Journal of Probability and Statistics 21, pp. 27-61, 2007

[15] M. Hess, "Nonlinear double-jump stochastic filtering using generalized Lévy-type processes", Manuscript, 2011

[16] J.-W. Jang, Y. Krvavych, "Shot noise process and pricing of extreme insurance claims in an economic environment", Working Paper, Actuarial Studies, UNSW, Sydrey, Australia, 2003

[17] J.-W. Jang, "Credit derivatives pricing using the cox process with shot noise intensity", the 15th International AFIR Colloquium, Zuerich, 6-9 Sep. 2005

[18] I. Kazakov, B. Dostupov, "Statisticheskaya Dinamica Nelineynih Avtomaticheskih Sistem"(in russian), Moscow, 1962

[19] C. Klüppelberg, T. Mikosch, "Explosive poisson shot noise processes with applications to risk reserves", Bernoulli 1(1/2), pp. 125-147, 1995

[20] R. Korn, E. Korn, "Option Pricing and Portfolio Optimization: Modern Methods of Financial Mathematics", American Mathematical Society, 2001

[21] H. Kraft, "Optimal portfolio and Heston's stochastic volatility model: an explicit solution for power utility", Quantitative Finance, Volume 5, Issue 3, pp. 303-313, 2005

[22] P. Lakner, "Utility maximization with partial information", Stochastic Processes and their Applications 56, pp. 247-273, 1995

[23] P. Lakner, "Optimal trading strategy for an investor: the case of partial information", Stochastic Processes and their Applications 76, pp. 77-97, 1998

[24] S. S. Lee and J. Hannig, "Detecting jumps from Lévy jump diffusion processes", Journal of Financial Economics, 96, pp. 271-290, 2010

[25] A. Løkka, B. Øksendal, F. Proske, "Stochastic partial differential equations driven by Lévy space time white noise", The Annals of Applied Probability, Vol.14, No.3, pp. 1506-1528, 2004

[26] B. Mandelbrot, "The variation of certain speclulative prices", The Journal of Business, 35:394-419, 1963

[27] R.C. Merton, "Option pricing when underlying stock returns are discontinuous", Journal of Financial Economics, 3:125-144, 1976

[28] T. Meyer-Brandis, F. Proske, "Explicit solution of a non-linear filtering problem for Lévy processes with application to finance", Applied Mathemetics and Optimization 50, pp. 119-134, 2004

[29] B. Øksendal, A. Sulem, "Applied Stochastic Control of Jump Diffusions", Springer-Verlag Berlin Heidelberg, 2005

[30] S.J. Press, "A compound event model for security prices", The Journal of Business, 40:317-335, 1967

[31] W. Putschögl, J. Sass, "Optimal consumption and investment under partial information", Decisions in Economics and Finance 31, pp. 131-170, 2008

[32] D. Rupnik-Poklukar, "Nonlinear filtering for jump-diffusions", Journal of Computational and Applied Mathematics 197, pp. 558-567, 2006

[33] J. Sass, U.G. Haussmann, "Optimizing the terminal wealth under partial information: the drift process as a continuous-time Markov chain", Finance and Stochastics 8, pp. 553-577, 2004.

[34] T. Schmidt, W. Stute, "Shot-noise process and the minimal martingale measure", Statistics and Probability Letters, Vol. 77, Nr. 12, pp. 1332-1338, 2007

[35] E.-H. Shin, "Sigma point-based estimation", Journal of Global Positioning Systems, Vol.8, No.2, pp. 188-196, 2009

[36] W.M. Wonham, "Some applications of stochastic differential equationa to optimal nonlinear filtering", J.SIAM Control, Ser. A, Vol. 2, No.3, 1965

[37] M. Zakai, "On the optimal filtering of diffusion processes", Z. Wahrscheinlichkeitstheorie verw. Geb. 11, pp. 230-243, 1969

[38] T. Zariphopoulou, "A solution approach to valuation with unhedgeable risks", Finance and Stochastics 5, pp. 61-82, 2001

Publications

This thesis uses the following conference talks:

1. O. Putyatina, J. Sass: Approximation of Filters and Portfolio Optimization in Market Models with Patrial Information and Jumps. *16th INFORMS Applied Probability Conference "The Applied Probability Society Conference"*, Stockholm. 2011

2. O. Putyatina, J. Sass: Approximations and Filtering for Portfolio Optimization with Patrial Information and Jumps. *10th German Probability and Statistics Days "Stochastik-Tage Mainz"*, Mainz, 2012

Scientific Career

- 1991-2001: Secondary School No.132, Kharkiv, Ukraine.
- 2001-2006: Kharkiv National University of Radioelectronics, Faculty of Applied Mathematics and Management
  - 2005: Bachelors Degree Honours in Applied Mathematics
  - 2006: Masters Degree Honours in Applied Mathematics
  - 2007: Diploma of specialist in Economical Cybernetics
- 2007-2012: Technical University of Kaiserslautern, Department of mathematics
  - 2007-2009: Master study of Management Mathematics (programme Mathematics International) with specialization in Financial Mathematics. Degree: Master of Science (M.Sc.)
  - 2009-2012: PhD studies at the Department of Mathematics with specialization in Financial Mathematics

Wissenschaftlicher Werdegang

- 1991-2001: Oberschule 132, Kharkov, Ukraine.
- 2001-2006: Charkower Nationaluniversität für Radioelektronik, Fakultät für angewandte Mathematik und Management
  - 2005: Bachelor in angewandter Mathematik, (Diplom mit Auszeichnung)
  - 2006: Magister Artium in angewandter Mathematik, (Diplom mit Auszeichnung)
  - 2007: Diplom in Ökonomische Kybernrtik
- 2007-2012: Technische Universität Kaiserslautern, Fachbereich Mathematik
  - 2007-2009: Master-Studium im Studiengang Wirtschaftsmathematik (Studienprogramm Mathematics International) mit dem Studienschwerpunkt Finanzmathematik. Abschluss: Master of Science (M.Sc.)
  - 2009-2012: Promotion im Fach Mathematik mit dem Studienschwerpunkt Finanzmathematik