Diogo Alexandre Pereira

Portfolio optimization of stochastic volatility models through the dynamic programming equations

Dissertação para obtenção do Grau de Mestre em
Matemática e Aplicações

Orientador: Fernanda Cipriano, Professor Associado, FCT-UNL
Co-orientador: Nuno Martins, Professor Auxiliar, FCT-UNL

Presidente: Prof. Doutora Marta Faias
Arguente: Prof. Doutor Pedro Mota
Vogal: Prof. Doutor Nuno Martins

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Novembro 2018

Diogo Alexandre Pereira

Portfolio optimization of stochastic volatility models through the dynamic programming equations

Dissertação para obtenção do Grau de Mestre em Matemática e Aplicações

Orientador: Fernanda Cipriano, Professor Associado, FCT-UNL
Co-Orientador: Nuno Martins, Professor Auxiliar, FCT-UNL

Presidente: Prof. Doutora Marta Faias
Arguente: Prof. Doutor Pedro Mota
Vogal: Prof. Doutor Nuno Martins

Portfolio optimization of stochastic volatility models through the dynamic programming equations

Acknowledgements

I would like to express my gratitude for my advisor Prof. Fernanda Cipriano and co-advisor Prof. Nuno Martins, for all their input, observations, and guidance throughout this thesis. Their advice has shaped this thesis in many ways and hopefully made it better.
I would also like to thank my parents for all their support along the years. I couldn't have written this thesis without them.

Resumo

Neste trabalho estudamos o problema de otimização de carteiras em mercados com volatilidade estocástica. O critério de otimização considerado consiste na maximização de utilidade da riqueza final. O método mais usual para este tipo de problema passa pela resolução de uma equação com derivadas parciais, determinística não-linear, denominada equação Hamilton-Jacobi-Bellman (HJB) ou equação de programação dinâmica. Uma das maiores dificuldades consiste em verificar que a solução da equação de HJB coincide com a função *payoff* do portfólio ótimo. Estes resultados são conhecidos como teoremas de verificação. Neste sentido, seguimos a abordagem de Kraft [13], generalizando os teoremas de verificação para funções de utilidade mais gerais. A contribuição mais significativa deste trabalho consiste na resolução do problema de portfólio ótimo para o modelo de volatilidade estocástica 2-hypergeométrico considerando *power utilities*. Mais concretamente obtemos uma fórmula de Feynman-Kac para a solução da equação de HJB. Com base nesta representação estocástica aplicamos o método Monte Carlo para aproximar a solução da equação HJB, que no caso de ser suficientemente regular coincide com a função *payoff* do portfólio ótimo.

Palavras-chave: Carteiras ótimas, volatilidade estocástica, teoremas de verificação, modelo 2-hipergeométrico.

Abstract

In this work we study the problem of portfolio optimization in markets with stochastic volatility. The optimization criteria considered consists in the maximization of the utility of terminal wealth. The most usual method to solve this type of problem passes by the solution of an equation with partial derivatives, deterministic and non-linear, named the Hamilton-Jacobi-Bellman equation (HJB) or the dynamic programming equation. One of the biggest challenges consists in verifying that the solution to the HJB equation coincides with the payoff of the optimal portfolio. These results are known as verification theorems. In this sense, we follow the approach by Kraft [13], generalizing the verification theorems for more general utility functions. The most significant contribution of this work consists in the resolution of the optimal portfolio problem for the 2-hypergeometric stochastic volatility model considering power utilities. Specifically we obtain a Feynman-Kac formula for the solution of the HJB equation. Based on this stochastic representation we apply the Monte Carlo method to approximate the solution to the HJB equation, which if it sufficiently regular it coincides with the payoff function of the optimal portfolio.

Keywords: Optimal portfolios, stochastic volatility, verification theorems, 2-hypergeometric model.

Contents

List of Figures

Acronyms

1. CIR - Cox–Ingersoll–Ross;
2. HJB - Hamilton-Jacobi-Bellman;

Introduction

Ever since the seminal papers by Merton [16, 17], there has been a lot of research on continuous-time portfolio optimization, in the context of stochastic analysis. Merton considers the portfolio problem for the Black-Scholes model, and solves it in several contexts, for example, lifetime consumption and maximizing utility from terminal wealth. In this thesis we focus on the latter. It is well known that the Black-Scholes model does not describe accurately the reality of financial markets. For instance, it doesn't capture the volatility smile and skew feature of implied volatilities (see for instance the book by Zhu [24]).

To solve these problems, one approach has been to allow the volatility of the stocks to be stochastic, by the introduction of so called stochastic volatility models. Heston [8] proposes a model of this type using the Cox–Ingersoll–Ross process and taking the volatility of the stock to be the square root of this process. This model seems to be the most widely used today among the stochastic volatility models, especially for option pricing. For a model of this type, Liu, in [15], proposes a solution to the continuous-time portfolio optimization problem for power utilities, by solving the dynamic programming equations. However, a solution to these equations may not correspond to the payoff function of the optimal portfolio; in addition one needs to prove a verification theorem (see Kraft [13]). Following the approach by Kraft [13], we generalize his result of verification for general stochastic volatility models and general utilities.

Concerning the Heston model, if the so called Feller condition is not satisfied, the CIR process can reach zero in finite time, and in statistical estimation of the parameters we don't always have this condition. To overcome these issues, in the recent paper [4], the authors develop a new model, the α-Hypergeometric Stochastic Volatility Model. This model always has a positive distribution for the volatility, and at the same time remains tractable. The pricing of derivatives under this model has been object of recent research (see [20, 21]). One of our main goals is to analyse the portfolio problem for this model. More precisely we obtain a Feynman-Kac formula for the solution of the Hamilton-Jacobi-Bellman equation (HJB), which is the natural candidate to be the payoff of the optimal portfolio. Taking into account this stochastic representation we develop and implement a Monte Carlo type of numerical scheme for the solution of the HJB equation.

This work is organized as follows. In section 1, we review some results of stochastic analysis that we need throughout the thesis.

In section 2, we introduce our formulation using the Merton problem, i.e., the portfolio problem in the Black-Scholes framework. The purpose of this section is to give a general idea and a motivation to the general approach that we propose for stochastic volatility models.

In section 3, we formulate the general stochastic volatility portfolio problem. Using the condition of uniform integrability (like in Kraft [13]), we prove the verification theorem for general utilities. We also discuss the problem for power utilities, in which the dynamic programming equations can reduced to a linear differential equation, as proposed in [23].

In section 4, we solve the problem for a family of models that include the Heston model proposed by Liu [15]. For this family of models, the dynamic programming equations can be solved explicitly, and the conditions of the verification theorem can be proved, ensuring the optimality of the solution. This section follows closely the paper of Kraft [13].

In section 5, we propose a general method to solving the problem using the Feynman-Kac formula, from a theoretical and numerical point of view. First, we simplify the equation such that the corresponding partial differential operator coincides with the infinitesimal generator of the stochastic volatility model. Then we write the Feynman-Kac formula using the process with this infinitesimal generator. Such process is appropriate to show that the Feynman-Kac representation is finite valued for any given initial condition. Otherwise, in some cases, if we consider the original partial differential equation, the associated stochastic process is hard to deal with. In addition it is difficult to show that the Feynman-Kac representation is finite valued. This strategy can be applied successfully to the 2-hypergeometric model. Finally, still based on the deduced representation, we develop and implement a Monte Carlo type of method to obtain a numerical approximation.

1 Elements of Stochastic Analysis and Monte Carlo methods

1.1 Introductory concepts

Throughout this work we shall fix a constant $T > 0$. In this section, we remind various results of stochastic analysis which will be needed later. We start with various definitions.

Definition 1.1.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is a family, $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$, of σ-algebras such that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$, for any $0 \leq s \leq t \leq T$. We say that a filtration is right continuous if*

$$\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s \,,$$

for any $t \in [0, T)$.

A probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, is complete if the set $\mathcal{N} = \{A \in \mathcal{F} \,|\, \mathbb{P}(A) = 0\}$ satisfies

$$\forall A \in \mathcal{N},\ \forall B \subseteq \Omega,\ B \subseteq A \implies B \in \mathcal{F}.$$

A filtered probability space satisfying the usual conditions, is an ordered vector, $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, such that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$, and

1. *$(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space;*
2. *$\mathcal{F}_0$ contains all null measure sets of $\mathcal{F}$, i.e., $\mathcal{N} \subseteq \mathcal{F}_0$;*
3. *The filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is right continuous.*

For $0 \leq a \leq b \leq T$, a stochastic process in $[a, b]$, is a mapping from $[a, b] \times \Omega$ into $\mathbb{R}$, such that, it's measurable in the product σ-algebra, $\mathcal{B}([a, b]) \times \mathcal{F}$. We say that a stochastic process in $[a, b]$, $X(s, \omega)$, is $\mathcal{F}_t$-adapted if the random variable $X(s, \cdot)$ is $\mathcal{F}_s$-measurable, for any $s \in [a, b]$.

Definition 1.1.2. *Take a filtered probability space satisfying the usual conditions $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. A $\mathcal{F}_t$ Brownian motion is a stochastic process, $(W_t)_{t\in[0,T]}$, such that*

1. *W_t is $\mathcal{F}_t$-adapted;*
2. *For any $0 \leq s < t \leq T$, the random variable $W_t - W_s$, is independent of $\mathcal{F}_s$;*
3. *$W_0 = 0$ almost surely;*
4. *For any $0 \leq s < t \leq T$, $W_t - W_s \sim N(0, t - s)$;*
5. *For any $0 \leq t_1 < t_2 < \cdots < t_n \leq T$, the random variables*

   $$W_{t_1},\, W_{t_2} - W_{t_1},\, \ldots,\, W_{t_n} - W_{t_{n-1}},$$

   are independent.
6. *The process W_t is pathwise continuous, i.e., there is $A \in \mathcal{F}$ with $\mathbb{P}(A) = 1$ such that $W(\cdot, \omega)$ is continuous for all $\omega \in A$.*

Definition 1.1.3. *For a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, n stochastic processes, $f_1(t)$, $f_2(t)$, …, $f_n(t)$, are said to be independent if for any $0 \leq t_1 < t_2 < \cdots < t_m \leq T$, the m-valued random variables $(f_1(t_1), f_1(t_2), \ldots, f_1(t_m))$, …, $(f_n(t_1), f_n(t_2), \ldots, f_n(t_m))$ are independent.*

Given a filtered probability space satisfying the usual conditions $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, we call a q-dimensional Brownian motion, to a mapping from $[0, T] \times \Omega$ to $\mathbb{R}^q$, $W_t = (W_t^1, \ldots, W_t^q)$, such that

1. *Each mapping W_t^i is an $\mathcal{F}_t$ Brownian motion;*

2. *The stochastic processes, $W_t^1, \ldots, W_t^q$, are independent.*

A q-dimensional reference space is an ordered vector $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W)$ such that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space satisfying the usual conditions and W is a q-dimensional Brownian motion.

Given these definitions one can develop the concept of the stochastic integral. We give a brief review of this concept in the following observation. The full development can be found in [14].

Observation 1.1.4. *Given a filtered probability space satiffeynmasfying the usual conditions $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ we may take the following spaces of stochastic processes:*

1. *$L^2_{ad}([a,b] \times \Omega)$ the space of stochastic processes, $f(t,\omega)$, such that*
   (a) *$f(t,\omega)$ is $\mathcal{F}_t$ adapted;*
   (b) *$\mathbb{E} \int_a^b |f(s)|^2 \, ds < \infty$.*
2. *$\mathcal{L}_{ad}(\Omega, L^2[a,b])$ the space of stochastic processes, $f(t,\omega)$, such that*
   (a) *$f(t,\omega)$ is $\mathcal{F}_t$ adapted;*
   (b) *$\int_a^b |f(s)|^2 \, ds < \infty$, $\mathbb{P}$-a.s.*

We note that $L^2_{ad}([a,b] \times \Omega) \subseteq \mathcal{L}_{ad}(\Omega, L^2[a,b])$. We know that for a $\mathcal{F}_t$ Brownian motion, W_t, and stochastic process, $f \in \mathcal{L}_{ad}(\Omega, L^2[a,b])$, the stochastic integral $\int_a^b f(s,\omega) \, dW_s$ is well defined as an element of $L^0(\Omega)$. In the particular case that $f \in L^2_{ad}([a,b] \times \Omega)$, the stochastic integral $\int_a^b f(s,\omega) \, dW_s$ is an element of $L^2(\Omega)$.

We also know that we can take representatives of $\int_a^t f(s) \, dW_s$, $a \leq t \leq b$, i.e., take elements of the a.s. equivalence classes of $\int_a^t f(s) \, dW_s$, such that the mapping

$$(t,\omega) \mapsto \int_0^t f(s) \, dW_s$$

is a continuous stochastic process (a continuous stochastic process is one that satisfies condition 6 of Definition 1.1.2).

We present now the theorem for Itô's formula. First let's define what we mean by an Itô process.

Definition 1.1.5. *Consider a q-dimensional reference space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W)$. A stochastic process, X_t, defined on $[a,b]$ is an Itô process if*

1. *X_t is $\mathcal{F}_t$-adapted and continuous;*
2. *There exist stochastic processes $f(t,\omega), \sigma_1(t,\omega), \ldots, \sigma_q(t,\omega)$, such that*
   (a) *$f(t,\omega)$ is $\mathcal{F}_t$-adapted and $\int_a^b |f(t,\omega)| \, ds < \infty$, $\mathbb{P}$-a.s.;*
   (b) *$\sigma_i(t,\omega) \in \mathcal{L}_{ad}(\Omega, L^2[a,b])$, $1 \leq i \leq q$;*
   (c) *the equality*

$$X_t = X_a + \int_a^t f(s,\omega) \, ds + \sum_{i=1}^{q} \int_a^t \sigma_i(s,\omega) \, dW_s^i \tag{1}$$

   holds a.s., for any $t \in [a,b]$.

We present Itô's formula now, which we generalize for stopping times.

Theorem 1.1.6. *Take an itô process, X_t, with corresponding stochastic processes $f(t,\omega)$, $\sigma_1(t,\omega),\ldots,\sigma_q(t,\omega)$ as in the previous definition. Then for any function $F:[a,b]\times\mathbb{R}\to\mathbb{R}$, such that it's partial derivatives $\frac{\partial F}{\partial t}$, $\frac{\partial F}{\partial x}$, $\frac{\partial^2 F}{\partial x^2}$, exist and are continuous, and for any stopping time, θ, with values in $[a,b]$, the equality*

$$\begin{aligned} F(\theta,X_\theta) = F(a,X_a) &+ \int_a^b 1_{[a,\theta]}(s)\frac{\partial F}{\partial t}(s,X_s)\,ds + \int_a^b 1_{[a,\theta]}(s)\frac{\partial F}{\partial x}(s,X_s)f(s,\omega)\,ds \\ &+ \sum_{i=1}^q \int_a^b 1_{[a,\theta]}(s)\frac{\partial F}{\partial x}(s,X_s)\sigma_i(s,\omega)\,dW_s^i \\ &+ \frac{1}{2}\sum_{i=1}^q \int_a^b 1_{[a,\theta]}(s)\frac{\partial^2 F}{\partial x^2}(s,X_s)\sigma_i(s,\omega)^2\,ds \end{aligned} \tag{2}$$

holds $\mathbb{P}$-a.s.

We can also generalize this theorem for n dimensions.

Theorem 1.1.7. *Take Itô processes, X_t^i, with corresponding stochastic processes $f_i(t,\omega)$, $\sigma_{i,1}(t,\omega),\ldots,\sigma_{i,q}(t,\omega)$, as in the previous definition ($1\le i\le n$). Then for any function $F:[a,b]\times\mathbb{R}^n\to\mathbb{R}$ such that it's partial derivatives $\frac{\partial F}{\partial t}$, $\frac{\partial F}{\partial x_i}$, $\frac{\partial^2 F}{\partial x_i x_j}$, for $i,j=1,\ldots,n$, exist and are continuous, and for any stopping time, θ, with values in $[a,b]$, the equality*

$$\begin{aligned} F(\theta,X_\theta) = F(a,X_a) &+ \int_a^b 1_{[a,\theta]}(s)\frac{\partial F}{\partial t}(s,X_s)\,ds + \sum_{i=1}^n \int_a^b 1_{[a,\theta]}(s)\frac{\partial F}{\partial x_i}(s,X_s)f_i(s,\omega)\,ds \\ &+ \sum_{i=1}^n\sum_{j=1}^q \int_a^b 1_{[a,\theta]}(s)\frac{\partial F}{\partial x_i}(s,X_s)\,\sigma_{ij}(s,\omega)\,dW_s^j \\ &+ \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^q \int_a^b 1_{[a,\theta]}(s)\frac{\partial^2 F}{\partial x_i x_j}(s,X_s)\sigma_{ik}(s,\omega)\sigma_{jk}(s,\omega)\,ds \end{aligned} \tag{3}$$

holds $\mathbb{P}$-a.s. (we're denoting $X_t=(X_t^1,\ldots,X_t^n)$).

The proof of these generalized theorems can be done using the usual Itô's formula and a property of stochastic integrals with stopping times (see Friedman [6]).

1.2 Stochastic Differential Equations

Take for now a 1-dimensional reference space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P},W)$. A 1-dimensional stochastic differential equation on the interval $[a,b]$, $0\le a\le b\le T$, and domain $D\subseteq\mathbb{R}$ (we always take $D=\mathbb{R}$ or $D=\mathbb{R}^+$) is a differential of the form

$$dX_t = \mu(t,X_t,\omega)\,dt + \sigma(t,X_t,\omega)\,dW_t, \quad X_a=\xi. \tag{4}$$

where $\mu:[a,b]\times D\times\Omega\to\mathbb{R}$ and $\sigma:[a,b]\times D\times\Omega\to\mathbb{R}$ are given mappings, and ξ is a certain $\mathcal{F}_a$-measurable random variable. In a way this differential has no exact mathematical form, it's simply something we refer to for ease of communication. Let's define now precisely what we mean by a solution of this equation.

Definition 1.2.1. *Fix an interval $[a,b]$, domain $D\subseteq\mathbb{R}$, mappings $\mu,\sigma:[a,b]\times D\times\Omega\to\mathbb{R}$ and a $\mathcal{F}_a$-measurable random variable ξ. A solution of the stochastic differential equation (4), is a stochastic process, X_t, defined on $[a,b]$, such that*

1. *X_t takes values in D, is $\mathcal{F}_t$-adapted, and is continuous;*

2. $\mu(t, X_t, \omega)$ is an $\mathcal{F}_t$-adapted process such that $\int_a^b |\mu(s, X_s, \omega)| \, ds < \infty$ for $\mathbb{P}$-a.s.;

3. $\sigma(t, X_t, \omega) \in \mathcal{L}_{ad}(\Omega, L^2[a, b])$, so the stochastic integral is well defined;

4. The equality

$$X_t = \xi + \int_a^t \mu(s, X_s, \omega) \, ds + \int_a^t \sigma(s, X_s, \omega) \, dW_s \tag{5}$$

holds a.s., for any $t \in [a, b]$.

This last equality means that for each fixed $t \in [a, b]$, for any version of the stochastic integral, there is as set $A \in \mathcal{F}$ with $\mathbb{P}(A) = 1$, such that for all $\omega \in A$, $\int_a^t |\mu(s, X_s, \omega)| \, ds < \infty$, so this integral is well defined and the equality (5) holds (for the fixed t and ω).

Definition 1.2.2. *We say that the solution to the stochastic differential equation* (4) *is unique if for any two solutions, X_t, Y_t, as in definition 1.2.1, satisfy*

$$X(t, \omega) = Y(t, \omega), \quad \text{for all } t \in [a, b],$$

in a set of probability one.

We note that this is stronger then $X_t = Y_t$ a.s. for all $t \in [a, b]$, i.e. the statement in the definition says there is a set of probability one such that the paths of the stochastic processes coincide. Still because the processes are continuous by definition, if $X_t = Y_t$ a.s. for all $t \in [a, b]$ one will still get the stronger property of pathwise uniqueness. As such, in our case the two views are equivalent.

We present the usual theorem of existence and uniqueness of solution with Lipschitz coefficients, where the functions μ and σ do not depend on ω

Theorem 1.2.3. *Fix an interval $[a, b]$ and take measurable functions $\mu, \sigma : [a, b] \times \mathbb{R} \to \mathbb{R}$ such that for some $K > 0$*

$$|\mu(t, x) - \mu(t, y)| \leq K|x - y|, \quad |\sigma(t, x) - \sigma(t, y)| \leq K|x - y|, \quad \forall\, t \in [a, b],\, x, y \in \mathbb{R}, \tag{6}$$

and for some $C > 0$

$$|\mu(t, x)| \leq C(1 + |x|) \quad |\sigma(t, x)| \leq C(1 + |x|), \quad \forall t \in [a, b],\, x \in \mathbb{R}. \tag{7}$$

Suppose ξ is a $\mathcal{F}_a$ measurable random variable such that $\mathbb{E}(|\xi|^2) < \infty$. Then the stochastic differential equation (4) *as an unique solution.*

Also a solution to this equation, X_t, satisfies the following estimate

$$\mathbb{E}\left(\sup_{a \leq t \leq b} |X_t|^2 \right) \leq C\, \mathbb{E}(|\xi|^2), \tag{8}$$

for some $C > 0$.

We also have the following Proposition for proving the uniqueness of a solution.

Proposition 1.2.4. *Fix an interval $[a, b]$ and domain $D \subseteq \mathbb{R}$ such that $D = \mathbb{R}$ or $D = \mathbb{R}^+$. Take a $\mathcal{F}_a$-measurable random variable, ξ, and functions $\mu, \sigma : [a, b] \times D \to \mathbb{R}$ such that for any compact set $K \subseteq D$ there exists $C_K > 0$ satisfying*

$$|\mu(t, x) - \mu(t, y)| \leq C_K|x - y|, \quad |\sigma(t, x) - \sigma(t, y)| \leq C_K|x - y|, \quad \forall\, t \in [a, b],\, x, y \in K. \tag{9}$$

If X_t is a solution of (4)*, for these mappings, then this solution is unique.*

This Proposition can be applied in many cases, since functions of C^1 type are always locally Lipschitz. The proof of these results can be found in [9]. We now present a theorem for the solution of linear stochastic differential equations.

Theorem 1.2.5. *Consider an interval $[a, b]$ and ξ a $\mathcal{F}_a$-measurable random variable. Take two $\mathcal{F}_t$-adapted stochastic processes $\mu(t, \omega), \sigma(t, \omega)$, defined on $[a, b]$, such that*

$$\int_a^b |\mu(t, \omega)|\, ds < \infty,\ \mathbb{P} - a.s. \quad and \quad \sigma(t, \omega) \in \mathcal{L}_{ad}(\Omega, L^2[a, b]).$$

Then the linear stochastic differential equation

$$dX_t = \mu(t) X_t\, dt + \sigma(t) X_t\, dW_s, \quad X_a = \xi, \tag{10}$$

has an unique solution given by

$$X_t = \xi \exp\left\{ \int_a^t \mu(s)\, ds - \frac{1}{2} \int_a^t \sigma(s)^2\, ds + \int_a^t \sigma(s)\, dW_s \right\}.$$

See [9] for the proof. Finally, we present a theorem that will be useful later. The proof can be found in [11].

Theorem 1.2.6. *(Doob's submartingale inequality) Let X_t be non-negative continuous submartingale in $[a, b]$, $0 \leq a < b \leq T$, and let $t \in [a, b]$. Then for every $K > 0$*

$$\mathbb{P}\left(\sup_{s \in [a,t]} |X_s| \geq K \right) \leq \frac{\mathbb{E}\,|X_t|}{K}, \tag{11}$$

and for every $p > 1$

$$\mathbb{E}\left(\sup_{s \in [a,t]} |X_s|^p \right) \leq \left(\frac{p}{p-1} \right)^p \mathbb{E}\,|X_t|^p. \tag{12}$$

1.3 Feynman-Kac formula

Here we give the theorem for the Feynman-Kac formula. We generalize the conditions that are usually found in the literature.

Theorem 1.3.1. *(Feynman-Kac formula) Fix continuous functions,*

$$\mu, \sigma, V, f : [0, T] \times \mathbb{R} \to \mathbb{R}.$$

Consider now a function $u \in C^{1,2}([0, T] \times \mathbb{R})$ satisfying

$$\frac{\partial u}{\partial t}(t, x) + \mu(t, x) \frac{\partial u}{\partial x}(t, x) + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 u}{\partial x^2}(t, x) + V(t, x) u(t, x) + f(t, x) = 0, \tag{13}$$

on $[0, T] \times \mathbb{R}$, with final condition $u(T, x) = u_0(x)$ on $\mathbb{R}$, for some $u_0 \in C^2(\mathbb{R})$.

Fix $(t, x) \in [0, T] \times \mathbb{R}$ and suppose that the stochastic differential equation (in $D = \mathbb{R}$)

$$dX_s = \mu(s, X_s)\, ds + \sigma(s, X_s)\, dW_s, \quad X_t = x,$$

has a solution X_s. If for any sequence, $(\theta_n)_{n \in \mathbb{N}}$, of stopping times in $[t, T]$ the collection

$$\left\{ u(\theta_n, X_{\theta_n}) e^{\int_t^{\theta_n} V(\tau, X_\tau)\, d\tau} + \int_t^{\theta_n} f(\tau, X_\tau) e^{\int_t^{T} V(\lambda, X_\lambda)\, d\lambda}\, d\tau \right\}_{n \in \mathbb{N}} \tag{14}$$

is uniformly integrable, we have

$$u(t, x) = \mathbb{E}\left[u_0(X_T) e^{\int_t^T V(\tau, X_\tau)\, d\tau} + \int_t^T f(s, X_s) e^{\int_t^s V(\tau, X_\tau)\, d\tau}\, ds \right]. \tag{15}$$

Proof. We start by noting that we can see the process $e^{\int_t^s V(\tau,X_\tau)\,d\tau}$ as an Itô process since

$$e^{\int_t^s V(\tau,X_\tau)\,d\tau} = 1 + \int_t^s V(\tau,X_\tau) e^{\int_t^\tau V(\lambda,X_\lambda)\,d\lambda}\,d\tau,$$

holds $a.s.$ - $\mathbb{P}$, (on the set such that X_s is continuous). Consider a stopping time, θ, with values in $[t,T]$. Using Itô's formula on $u(\theta,X_\theta)e^{\int_t^\theta V(s,X_s)\,ds}$ we obtain

$$\begin{aligned}
u(\theta,X_\theta)e^{\int_t^\theta V(s,X_s)\,ds} = \; & u(t,x) + \int_t^T 1_{[t,\theta]}(s)\frac{\partial u}{\partial t}(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,ds \\
& + \int_t^T 1_{[t,\theta]}(s)\frac{\partial u}{\partial x}(s,X_s)\mu(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,ds \\
& + \frac{1}{2}\int_t^T 1_{[t,\theta]}(s)\frac{\partial^2 u}{\partial x^2}(s,X_s)\sigma(s,X_s)^2 e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,ds \\
& + \int_t^T 1_{[t,\theta]}(s)u(s,X_s)V(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,ds \\
& + \int_t^T 1_{[t,\theta]}(s)\frac{\partial u}{\partial x}(s,X_s)\sigma(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,dW_s.
\end{aligned} \tag{16}$$

Due to equation (13) we can simplify this equality to

$$\begin{aligned}
u(\theta,X_\theta)e^{\int_t^\theta V(s,X_s)\,ds} = \; & u(t,x) - \int_t^T 1_{[t,\theta]}(s)f(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,ds \\
& + \int_t^T 1_{[t,\theta]}(s)\frac{\partial u}{\partial x}(s,X_s)\sigma(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,dW_s.
\end{aligned} \tag{17}$$

Consider now the stopping times

$$\theta_n(\omega) = \begin{cases} \inf\left\{s\in[t,T]\;\middle|\; \int_t^s |\frac{\partial u}{\partial x}(\tau,X_\tau)\sigma(\tau,X_\tau)e^{\int_t^\tau V(\lambda,X_\lambda)\,d\lambda}|^2\,d\tau \geq n\right\}, & if\ \{s\in[t,T]\,|\,\cdots\}\neq\emptyset; \\ T, & if\ \{s\in[t,T]\,|\,\cdots\}=\emptyset. \end{cases}$$

where the set, $\{s\in[t,T]\,|\,\cdots\}$, is the same one we are taking the inf over. Under these stopping times the stochastic integral in (17) has expectation 0. Therefore we may take the expectation in (17), using these stopping times, and obtain

$$\mathbb{E}\left[u(\theta_n,X_{\theta_n})e^{\int_t^{\theta_n} V(s,X_s)\,ds} + \int_t^{\theta_n} f(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,ds\right] = u(t,x). \tag{18}$$

Now note that

$$\lim_{n\to\infty}\theta_n = T, \quad a.s. - \mathbb{P},$$

and by hypothesis the random variables in the above expectation (18) are uniformly integrable. Therefore we have

$$\begin{aligned}
& \mathbb{E}\left[u(T,X_T)e^{\int_t^T V(s,X_s)\,ds} + \int_t^T f(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,ds\right] \\
&= \lim_{n\to\infty}\mathbb{E}\left[u(\theta_n,X_{\theta_n})e^{\int_t^{\theta_n} V(s,X_s)\,ds} + \int_t^{\theta_n} f(s,X_s)e^{\int_t^s V(\tau,X_\tau)\,d\tau}\,ds\right] \\
&= u(t,x).
\end{aligned} \tag{19}$$

□

To prove the verification theorems, we need to prove conditions about uniform integrability, much like in the previous theorem. To help prove these conditions, we present the following theorems on uniform integrability and local martingales, which we need later.

Theorem 1.3.2. *Let X_s be an adapted stochastic process in $[t,T]$ and let $(\theta_n)_{n\in\mathbb{N}}$ be a sequence of stopping times in $[t,T]$. Then if either*

1. $\mathbb{E}\left[\sup_{s\in[t,T]} |X_s|\right] < \infty$;
2. $\mathbb{E}\left|X_{\theta_n}\right|^q \leq C$, *for some $C > 0$, $q > 1$,*

we have that the sequence $(X_{\theta_n})_{n\in\mathbb{N}}$ is uniformly integrable.

Theorem 1.3.3. *Let X_s be an adapted stochastic process in $[t,T]$.*

If X_s is a non-negative local martingale and $\mathbb{E}\,|X_t| < \infty$ then X_s is a supermartingale.

If X_s is a supermartingale, then for any stopping time, θ, in $[t,T]$

$$\mathbb{E}[X_\theta] \leq \mathbb{E}[X_t].$$

The last statement is a special case of the optional sampling theorem. See [11] for these results.

1.4 Monte Carlo methods

Let X be a real valued random variable and $f : \mathbb{R} \to \mathbb{R}$ some function. Suppose that $f(X)$ is integrable, and that we want to approximate its expected value. By the strong law of large numbers, if we have a sequence of independent random variables $X_1, X_2, \ldots$ all with the distribution of X, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(X_i) = \mathbb{E}\, f(X), \quad \mathbb{P} - a.s.$$

See [11] for the proof. The Monte Carlo method is based around this result. To approximate $\mathbb{E}f(X)$ we simulate N independent random variables, $X_1, X_2, \ldots, X_N$, all with the distribution of X, and calculate $\frac{1}{N}\sum_{i=1}^{N} f(X_i)$.

Since this is a stochastic method, we shall measure the error using the standard deviation. For this note

$$\begin{aligned}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}f(X)\right)^2 &= \frac{1}{N^2}\,\mathbb{E}\left(\sum_{i=1}^{N}\Big(f(X_i) - \mathbb{E}f(X)\Big)\right)^2 \\ &= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \mathbb{E}\left[\Big(f(X_i) - \mathbb{E}f(X)\Big)\Big(f(X_j) - \mathbb{E}f(X)\Big)\right].\end{aligned} \tag{20}$$

In the case that $i \neq j$, X_i is independent of X_j, therefore

$$\mathbb{E}\left[\Big(f(X_i) - \mathbb{E}f(X)\Big)\Big(f(X_j) - \mathbb{E}f(X)\Big)\right] = \mathbb{E}\left[f(X_i) - \mathbb{E}f(X)\right]\mathbb{E}\left[f(X_j) - \mathbb{E}f(X)\right] = 0.$$

Continuing the above equality, we have then

$$\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}f(X)\right)^2 = \frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}\left[f(X_i) - \mathbb{E}f(X)\right]^2 = \frac{1}{N}Var(f(X)), \tag{21}$$

where $Var(f(X)) = \mathbb{E}\left(f(X) - \mathbb{E}f(X)\right)^2$. The standard deviation is therefore

$$\sqrt{\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}f(X)\right)^2} = \frac{1}{\sqrt{N}}\sqrt{Var(f(X))}.$$

It follows that this method approximation is of order $\frac{1}{\sqrt{N}}$, assuming that $Var(f(X))$ is finite. We synthesize this in the following theorem.

Theorem 1.4.1. *Let $X, X_1, X_2, \ldots$ be identically distributed and independent random variables, and suppose that $\mathbb{E}(X^2) < \infty$. Then*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} f(X_i) = \mathbb{E}\, f(X), \quad \mathbb{P} - a.s.$$

and

$$\sqrt{\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}f(X)\right)^2} = \frac{1}{\sqrt{N}}\sqrt{Var(f(X))}.$$

The random variables we have to simulate for the Monte Carlo method may sometimes be solutions of stochastic differential equations. In many cases, we don't know the explicit solution to the stochastic differential equation, so we can't simulate these random variables directly, using methods based on their distribution (see for instance [7]).

Take $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W)$ a 1-dimensional reference space. Let $\mu, \sigma : \mathbb{R} \to \mathbb{R}$ be functions, and suppose that the stochastic differential equation

$$dX_s = \mu(X_s)\, ds + \sigma(X_s)\, dW_s, \quad X_t = x$$

has a solution, X_s, on $[t, T]$, for starting conditions $(t, x) \in [0, T] \times \mathbb{R}$. Consider now a discretization of time $t = t_0 < t_1 < \cdots < t_M = T$, such that $t_i - t_{i-1} = h$, for all $i = 1, \ldots, M$, where $h > 0$ is some constant. The Euler method (or Euler–Maruyama method) approximates the simulation of X_s in the points $t_0, t_1, \ldots, t_M$ by the following recursive approach

$$\begin{cases} \hat{X}_0 = x; \\ \hat{X}_{i+1} = \mu(\hat{X}_i)h + \sigma(\hat{X}_i)\sqrt{h}\mathcal{N}_{i+1} \end{cases} \tag{22}$$

where $\mathcal{N}_i$, $i = 1, \ldots .M$ are independent random variables with normal distribution, mean 0 and variance 1, we need to simulate (see [7]).

For $f : \mathbb{R} \to \mathbb{R}$ in some class of functions $\mathcal{C}$ (see [7]), the so called weak error for this method is given by

$$\sup_{i=0,1,\ldots,M} \left|\mathbb{E}f(\hat{X}_i) - \mathbb{E}f(X_{t_i})\right|.$$

Under some conditions on the coefficients μ and σ, this error can be proven to converge to 0 with a order of 1, i.e., for all $f \in \mathcal{C}$

$$\sup_{i=0,1,\ldots,M} \left|\mathbb{E}f(\hat{X}_i) - \mathbb{E}f(X_{t_i})\right| \le C\frac{1}{M},$$

for some $C > 0$ (see [7] and references therein).

2 Merton Problem

2.1 Formulation of the Problem

The Merton Problem is the control problem for the Black Scholes Model. The Black Scholes model is given by the dynamics

$$\begin{cases} dB_t = rB_t\,dt, & B_0 = 1, \\ dS_t = \mu S_t\,dt + \sigma S_t\,dW_t, & S_0 = s_0. \end{cases} \tag{23}$$

where r, μ are non-negative constants with $\mu \geq r$ and σ, s_0 are positive. This system has the solution

$$\begin{cases} B_t = e^{rt} \\ S_t = s_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}, \end{cases} \tag{24}$$

where we are considering a 1-dimensional reference space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W)$.

We want to manage portfolios on this model. By a portfolio we mean a pair of adapted stochastic processes a_t, b_t such that a_t represents the quantity the holder has of the asset B_t at time t and b_t represents the quantity of the asset S_t the holder has at time t. We can then define the value of the portfolio at time t to be

$$X_t = a_t B_t + b_t S_t. \tag{25}$$

We want to consider portfolios that are self-financed, i.e., portfolios where money is not injected or taken out. We formalize this condition by supposing that the wealth process X_t satisfies the dynamics

$$dX_t = a_t dB_t + b_t dS_t. \tag{26}$$

Developing this equality we obtain

$$dX_t = a_t dB_t + b_t dS_t = (a_t r B_t + b_t \mu S_t)\,dt + b_t \sigma S_t\,dW_t. \tag{27}$$

We also only consider portfolios such that $X_t > 0$ for all $(t, \omega) \in [0, T] \times \Omega$. Since in the Black-Scholes model, prices are always positive this is not too restrictive.

So as to simplify the equation we take φ_t to denote the fraction of the wealth invested in the stock S_t, i.e.

$$\varphi_t = \frac{S_t b_t}{X_t}, \qquad \textit{and by } (25), \qquad 1 - \varphi_t = \frac{a_t B_t}{X_t}.$$

So we have simply

$$dX_t = (r(1 - \varphi_t)X_t + \mu\,\varphi_t X_t)\,dt + \sigma\,\varphi_t X_t\,dW_t, \tag{28}$$

which doesn't depend on S_t, so we can simply look to this equation to formulate the control problem.

Definition 2.1.1. *We define the set of admissible controls, $\mathcal{A}(t)$, defined on $t \in [0, T]$, to be the set of adapted stochastic processes, φ, with domain $[t, T] \times \Omega$, such that*

$$\int_t^T |\varphi_s|^2\,ds < \infty, \quad a.s. \tag{29}$$

We note that for a fixed $\varphi_s \in \mathcal{A}(t)$, the stochastic differential equation (28) has an unique solution for any starting position $X_t = x > 0$ (using theorem 1.2.5). Given $t \in [0, T]$, $x \in \mathbb{R}^+$ and $\varphi_s \in \mathcal{A}(t)$, we denote a corresponding solution to equation (28) by $X_s^{t,x,\varphi}$. Also note that, since the equation is a linear equation, for starting time $x > 0$ we always get $X_s^{t,x,\varphi} > 0$, $\forall s \in [t, T]$, *a.s.*

We are now ready to formulate our control problem. For this we need the concept of an utility function.

Definition 2.1.2. *An utility function is a function $U \in C^2(\mathbb{R}^+)$ such that*

$$U(x) \geq 0, \quad U'(x) \geq 0, \quad \textit{and} \quad U''(x) \leq 0, \quad \textit{for all } x \in \mathbb{R}^+$$

The condition $U(x) \geq 0$ is equivalent to supposing U is lower bounded, this is because the control problem is equivalent for utilities such that its difference is constant. The condition $U'(x) \geq 0$ is so the function is increasing, so that the agent prefers more wealth then less, and the condition $U''(x) \leq 0$ is so the agent is risk averse.

The objective is now to find the best control with the maximum payoff $\mathbb{E}\left[U(X_T^{\varphi})\right]$, for each starting condition. Let's define what we mean by the payoff associated to a control.

Definition 2.1.3. *Consider the set*

$$\mathcal{D} = \left\{(t, x, \varphi) \,\middle|\, t \in [0, T],\ x \in \mathbb{R}^+,\ \varphi \in \mathcal{A}(t)\right\}. \tag{30}$$

For each $(t, x, \varphi) \in \mathcal{D}$ we fix a corresponding solution $X_s^{t,x,\varphi}$ of equation (28). We define the payoff function, $P : \mathcal{D} \to \mathbb{R} \cup \{\infty\}$, for the utility function, U, as being

$$P(t, x, \varphi) = \mathbb{E}\left[U(X_T^{t,x,\varphi})\right]. \tag{31}$$

By pathwise uniqueness, $\mathbb{E}\left[U(X_T^{t,x,\varphi})\right]$ is the same for any solution of equation (28) we fix, so $P(t, x, \varphi)$ doesn't depend on the considered solution to the stochastic equation. Now we define the notion of value function.

Definition 2.1.4. *Let P be the payoff function for a certain utility function U. The value function $V : [0, T] \times \mathbb{R}^+ \to \mathbb{R} \cup \{+\infty\}$, is defined by*

$$V(t, x) = \sup_{\varphi \in \mathcal{A}(t)} P(t, x, \varphi). \tag{32}$$

We shall also denote the value function V, by the lower case v.

2.2 Verification theorem

Our objective is to find a control processes, $\varphi^{t,x} \in \mathcal{A}(t)$, for each $(t, x) \in [0, T] \times \mathbb{R}^+$, such that $P(t, x, \varphi^{t,x}) = V(t, x)$.

One approach is to solve a deterministic differential equation, called the dynamic programing equations, or also the Hamilton-Jacobi-Bellmann equations (HJB). Finding a solution to these equations is not always easy, and in some cases the equation might not have a smooth solution, so one has to use weaker notions of solutions to differential equations, like viscosity solutions (we refer to [3, 22] for these concepts). What we shall do now is prove what is called a Verification theorem. What this theorem gives us, is a guarantee that in the case that we find a smooth solution to the HJB equations, and one has an additional uniform integrability condition, then one knows that that solution is the value function.

We start with a Lemma.

Lemma 2.2.1. *Let $U \in C^2(\mathbb{R}^+)$ be an utility function and let $w \in C^{1,2}([0, T] \times \mathbb{R}^+)$ be a non-negative function satisfying*

1. *for fixed $(t, x) \in [0, T] \times \mathbb{R}^+$, the mapping,*

$$a \to \left[\frac{\partial w}{\partial x}(t, x)\, x\, (r + a(\mu - r)) + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(t, x) x^2 \sigma^2 a^2\right]$$

is bounded above for $a \in \mathbb{R}$;

2. $-\dfrac{\partial w}{\partial t}(t,x) - \sup_{a\in\mathbb{R}}\left[\dfrac{\partial w}{\partial x}(t,x)\,x\,(r+a(\mu-r)) + \dfrac{1}{2}\dfrac{\partial^2 w}{\partial x^2}(t,x)x^2\sigma^2a^2\right] \geq 0$, *on* $[0,T]\times\mathbb{R}^+$*;*

3. $w(T,x) \geq U(x), \quad \forall x\in\mathbb{R}^+$.

Then $w \geq v$ *on* $[0,T]\times\mathbb{R}^+$.

Proof. Let $(t,x,\varphi)\in\mathcal{D}$, and take a corresponding solution $X_s^{t,x,\varphi}$. Since $w\in C^{1,2}([0,T]\times\mathbb{R}^+)$ we may apply itô's formula on $w(\theta, X_\theta^{t,x,\varphi})$ for any stopping time $\theta\in[t,T]$. Doing this we obtain

$$
\begin{aligned}
w(\theta, X_\theta^{t,x,\varphi}) = w(t,x) &+ \int_t^T 1_{[t,\theta]}(s)\left[\frac{\partial w}{\partial t}(s,X_s^{t,x,\varphi}) + \frac{\partial w}{\partial x}(s,X_s^{t,x,\varphi})X_s^{t,x,\varphi}(r+\varphi_s(\mu-r))\right]ds \\
&+ \frac{1}{2}\int_t^T 1_{[t,\theta]}(s)\frac{\partial^2 w}{\partial x^2}(s,X_s^{t,x,\varphi})\left(X_s^{t,x,\varphi}\right)^2\sigma^2\varphi_s^2\,ds \\
&+ \int_t^T 1_{[t,\theta]}(s)\frac{\partial w}{\partial x}(s,X_s^{t,x,\varphi})X_s^{t,x,\varphi}\sigma\varphi_s\,dW_s,
\end{aligned}
\tag{33}
$$

a.s.-$\mathbb{P}$. We define the stopping time, θ_n, for $n\in\mathbb{N}$, and valued in $[t,T]$, by

$$
\theta_n(\omega) = \begin{cases} \inf\left\{s\in[t,T]\;\middle|\; \int_t^s |\frac{\partial w}{\partial x}(u,X_u^{t,x,\varphi})X_u^{t,x,\varphi}\sigma\varphi_u|^2\,du \geq n\right\}, & \textit{if } \{s\in[t,T]\,|\cdots\}\neq\emptyset; \\ T, & \textit{if}\{s\in[t,T]\,|\cdots\}=\emptyset, \end{cases}
$$

where $\{s\in[t,T]\,|\cdots\}$ refers to the same set we're taking the inf over.

For this stopping time the last term in (33) has expectation 0. On the other hand we have

$$
\begin{aligned}
&\frac{\partial w}{\partial t}(s,X_s^{t,x,\varphi}) + \frac{\partial w}{\partial x}(s,X_s^{t,x,\varphi})X_s^{t,x,\varphi}(r+\varphi_s(\mu-r))\,ds \\
&\qquad + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(s,X_s^{t,x,\varphi})\left(X_s^{t,x,\varphi}\right)^2\sigma^2\varphi_s^2 \\
\leq\; &\frac{\partial w}{\partial t}(s,X_s^{t,x,\varphi}) + \sup_{a\in\mathbb{R}}\left\{\frac{\partial w}{\partial x}(s,X_s^{t,x,\varphi})X_s^{t,x,\varphi}(r+a(\mu-r))\,ds\right. \\
&\qquad \left. + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(s,X_s^{t,x,\varphi})\left(X_s^{t,x,\varphi}\right)^2\sigma^2a^2\right\} \\
\leq\; &0,
\end{aligned}
\tag{34}
$$

by condition *2*.

Given all these observations, taking the expectation in (33) and using inequality (34), we have

$$
\mathbb{E}\left[w(\theta_n, X_{\theta_n}^{t,x,\varphi})\right] \leq w(t,x). \tag{35}
$$

Now, noting that $\lim_{n\to\infty}\theta_n(\omega)=T$, a.s.-$\mathbb{P}$, and as a result

$$
\lim_{n\to\infty} w(\theta_n, X_{\theta_n}^{t,x,\varphi}) = w(T, X_T^{t,x,\varphi}),
$$

a.s.-$\mathbb{P}$, we can use Fatou's Lemma to obtain

$$
P(t,x,\varphi) = \mathbb{E}\left[U(X_T^{t,x,\varphi})\right] \leq \mathbb{E}\left[w(T,X_T^{t,s,\varphi})\right] \leq \liminf_{n\to\infty}\mathbb{E}\left[w(\theta_n, X_{\theta_n}^{t,x,\varphi})\right] \leq w(t,x). \tag{36}
$$

We note that we can indeed use Fatou's Lemma because w is a non-negative function. Finally, by taking the sup over $\varphi\in\mathcal{A}(t)$ in the last inequality, we obtain

$$
v(t,x) \leq w(t,x).
$$

□

We note that if we find w in the above conditions we also prove that the value function is always finite. We present now the Verification theorem.

Theorem 2.2.2. *(Verification theorem) Let U be an utility function and let $w \in C^{1,2}([0,T] \times \mathbb{R}^+)$ be a non-negative function satisfying*

1. *for fixed $(t,x) \in [0,T] \times \mathbb{R}^+$, the mapping,*

$$a \to \left[\frac{\partial w}{\partial x}(t,x)\, x\,(r + a(\mu - r)) + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(t,x)x^2\sigma^2 a^2\right]$$

is bounded above for $a \in \mathbb{R}$;

2. $-\frac{\partial w}{\partial t}(t,x) - \sup_{a\in\mathbb{R}}\left[\frac{\partial w}{\partial x}(t,x)\, x\,(r + a(\mu - r)) + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(t,x)x^2\sigma^2 a^2\right] = 0,$ *on* $[0,T] \times \mathbb{R}^+$*;*

3. $w(T,x) = U(x), \quad \forall x \in \mathbb{R}^+.$

Let $\alpha : [0,T] \times \mathbb{R}^+ \to \mathbb{R}$ be a function such that

$$\alpha(t,x) \in \arg\max_{a\in\mathbb{R}}\left\{\frac{\partial w}{\partial x}(t,x)\, x\,(r + a(\mu - r)) + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(t,x)x^2\sigma^2 a^2\right\}, \quad on\ [0,T] \times \mathbb{R}^+ \tag{37}$$

(such a function always exists by condition 1). If for $(t,x) \in [0,T] \times \mathbb{R}^+$ the equation

$$d\hat{X}_s^{t,x} = \hat{X}_s^{t,x}(r + \alpha(s,\hat{X}_s^{t,x})(\mu - r))\, ds + \hat{X}_s^{t,x}\sigma\alpha(s,\hat{X}_s^{t,x})\, dW_s, \quad \hat{X}_t^{t,x} = x, \tag{38}$$

has a solution, $\hat{X}_s^{t,x}$, with $\alpha(s,\hat{X}_s^{t,x}) \in \mathcal{A}(t)$ and if for any sequence, $(\tau_n)_{n\in\mathbb{N}}$, of stopping times in $[t,T]$ the collection

$$\left\{w(\tau_n, \hat{X}_{\tau_n}^{t,x})\right\}_{n\in\mathbb{N}}, \tag{39}$$

is uniformly integrable, we have

$$w(t,x) = P(t,x,\alpha(s,\hat{X}_s^{t,x})) = v(t,x).$$

Proof. By the last Lemma it's immediate that $w \geq v$ on $[0,T]\times\mathbb{R}^+$. Fix now $(t,x) \in [0,T]\times\mathbb{R}^+$. If we denote $\varphi_s^* = \alpha(s,\hat{X}_s^{t,x})$ it should be immediate that $\hat{X}_s^{t,x} = X_s^{t,x,\varphi^*}$ for all $s \in [t,T]$, $\mathbb{P}$-a.s. This is true since they both satisfy equation (28) for the control φ_s^*.

We take the same stopping times, θ_n, as in last Lemma and in an analogous argumentation, we may take Itô's formula on the process $w(\theta_n, X_{\theta_n}^{t,x,\varphi^*})$ to obtain

$$\mathbb{E}\left[w(\theta_n, X_{\theta_n}^{t,x,\varphi^*})\right] = w(t,x),$$

where this time we have equality, by the stronger conditions on the function w, and the fact that $\alpha(t,x)$ satisfies condition (37). Now since

$$\lim_{n\to\infty}\theta_n = T, \quad a.s. - \mathbb{P},$$

and $\left\{w(\theta_n, \hat{X}_{\theta_n}^{t,x})\right\}_{n\in\mathbb{N}}$ is uniformly integrable by hypothesis, we have

$$v(t,x) \geq \mathbb{E}\left[U(X_T^{t,x,\varphi^*})\right] = \mathbb{E}\left[w(T, X_T^{t,x,\varphi^*})\right] = \lim_{n\to\infty}\mathbb{E}\left[w(\theta_n, X_{\theta_n}^{t,x,\varphi^*})\right] = w(t,x). \tag{40}$$

Since $w \geq v$ on $[0,T] \times \mathbb{R}^+$ we have then

$$w(t,x) = v(t,x)$$

and since also $w(t,x) = \mathbb{E}\left[U(X_T^{t,x,\varphi^*})\right]$, $\alpha(s,\hat{X}_s^{t,x})$ is an optimal control. $\square$

We remind that to prove that for any sequence of stopping times, $(\tau_n)_{n\in\mathbb{N}}$, the collection

$$\left\{w(\tau_n, \hat{X}^{t,x}_{\tau_n})\right\}_{n\in\mathbb{N}},$$

is uniformly integrable we can simply prove

$$\mathbb{E}\left[\sup_{s\in[t,T]} |w(s, \hat{X}^{t,x}_s)|\right] < \infty, \quad \text{or} \quad \mathbb{E}\left[w(\tau_n, \hat{X}^{t,x}_{\tau_n})^q\right] \leq C,$$

for some $q > 1$ and $C > 0$, which should be easier.

2.3 Power utility functions

In this section we investigate a special type of utility functions.

Definition 2.3.1. *An utility function, U, is said to be of power utility type if there exists $\gamma \in]0,1[$ such that*

$$U(x) = \frac{x^\gamma}{\gamma}$$

There's meaning for this type of utilities and we refer to Pham [18] for a discussion (they have what is called constant relative risk aversion).

These utilities are interesting because the HJB equation, in the Merton problem, admits a smooth solution, and this solution can be proved to be optimal using the verification theorem.

The HJB equation is given by

$$\begin{cases} -\dfrac{\partial w}{\partial t}(t,x) - \sup\limits_{a\in\mathbb{R}}\left[\dfrac{\partial w}{\partial x}(t,x)\, x\,(r + a(\mu - r)) + \dfrac{1}{2}\dfrac{\partial^2 w}{\partial x^2}(t,x) x^2\sigma^2 a^2\right] = 0, \ \text{on } [0,T]\times\mathbb{R}^+, \\ \\ w(T,x) = x^\gamma/\gamma. \end{cases} \tag{41}$$

Suppose now that this equation has a smooth solution, w, which admits a decomposition of the form $w(t,x) = h(t)\frac{x^\gamma}{\gamma}$, with $h(t) > 0$ for all $t \in [t,T]$. Substituting in the equation we get

$$\begin{cases} -\dfrac{\partial h}{\partial t}(t,x)\dfrac{x^\gamma}{\gamma} - \sup\limits_{a\in\mathbb{R}}\left[h(t)x^\gamma\,(r + a(\mu - r)) + \dfrac{1}{2}h(t)x^\gamma(\gamma-1)\sigma^2 a^2\right] = 0, \ \text{on } [0,T]\times\mathbb{R}^+, \\ \\ h(T) = 1. \end{cases} \tag{42}$$

Now the supremum is taken over a second order polynomial, so it achieves its maximum if its highest order coefficient is negative. We can see this is the case since $h(t) > 0$ and $\gamma \in]0,1[$. Taking the derivative we can find that this polynomial achieves it's maximum in

$$a^* = \frac{\mu - r}{(1-\gamma)\sigma^2}.$$

Dividing the equation by x^γ/γ and using this maximum, the equation simplifies to

$$\begin{cases} -\dfrac{\partial h}{\partial t}(t,x) - h(t)\gamma\left(r + \dfrac{1}{2}\dfrac{(\mu - r)^2}{(1-\gamma)\sigma^2}\right) = 0, \ \text{on } [0,T]\times\mathbb{R}^+, \\ \\ h(T) = 1. \end{cases} \tag{43}$$

Denoting by $\eta = \gamma \left(r + \frac{1}{2} \frac{(\mu - r)^2}{(1-\gamma)\sigma^2} \right)$ we conclude that

$$h(t) = e^{\eta(T-t)}.$$

This suggests that in this problem $v(t,x) = e^{\eta(T-t)} \frac{x^\gamma}{\gamma}$. We shall use the verification theorem to prove this.

Proposition 2.3.2. *For the utility $U(x) = x^\gamma/\gamma$, the value function of the associated Merton control problem is given by*

$$w(t,x) = e^{\eta(T-t)} \frac{x^\gamma}{\gamma},$$

where $\eta = \gamma \left(r + \frac{1}{2} \frac{(\mu - r)^2}{(1-\gamma)\sigma^2} \right)$, and an optimal control is given by

$$\varphi^*(s,\omega) = \frac{\mu - r}{(1-\gamma)\sigma^2}$$

for any starting conditions $t \in [0,T]$, $x \in \mathbb{R}^+$.

Proof. Firstly, to prove that w satisfies conditions *1,2* and *3* of the verification theorem simply follow the steps we did before, backwards. One can also check that

$$\frac{\mu - r}{(1-\gamma)\sigma^2} \in \arg\max_{a \in \mathbb{R}} \left\{ \frac{\partial w}{\partial x}(t,x)\, x \left(r + a(\mu - r) \right) + \frac{1}{2} \frac{\partial^2 w}{\partial x^2}(t,x) x^2 \sigma^2 a^2 \right\},$$

for any $(t,x) \in [0,T] \times \mathbb{R}^+$. Define $\alpha(t,x) = \frac{\mu - r}{(1-\gamma)\sigma^2}$ on $[0,T] \times \mathbb{R}^+$.

Fix starting conditions $(t,x) \in [0,T] \times \mathbb{R}^+$. Then, $\alpha(s,x)$ being constant, equation (38) always has a solution, $\hat{X}_s^{t,x}$ and $\alpha(s, \hat{X}_s^{t,x}) \in \mathcal{A}(t)$. Now lets check that

$$\mathbb{E} \left[\sup_{s \in [t,T]} |w(s, \hat{X}_s^{t,x})| \right] < \infty,$$

which, as we have noted, will imply the condition in the verification theorem. Firstly we note the following. For a Brownian motion, $W_s - W_t$ (starting at t), the process $e^{|C(W_s - W_t)|}$ is a non-negative submartingale for any $C \in \mathbb{R}$, since $x \mapsto e^{|Cx|}$ is convex and the process is integrable. Therefore we can use Doob's submartingale inequality to conclude

$$\mathbb{E} \sup_{s \in [t,T]} e^{|C(W_s - W_t)|} \leq \left(\mathbb{E} \sup_{s \in [t,T]} \left[e^{|C(W_s - W_t)|} \right]^2 \right)^{1/2} \leq 2 \left(\mathbb{E} \left[e^{|C(W_T - W_t)|} \right]^2 \right)^{1/2} < \infty.$$

Denote $\alpha(t,x)$ simply by α. We have

$$\begin{aligned} w(s, \hat{X}_s^{t,x}) &= e^{\eta(T-s)} \frac{(\hat{X}_s^{t,x})^\gamma}{\gamma} \\ &= e^{\eta(T-s)} \frac{\left(x \exp\left\{ r(s-t) + \alpha(\mu - r)(s-t) - \frac{1}{2}\sigma^2 \alpha^2 (s-t) + \sigma\alpha(W_s - W_t) \right\} \right)^\gamma}{\gamma} \\ &\leq \frac{x^\gamma}{\gamma} e^{|C(T-t)|} e^{|\gamma\sigma\alpha(W_s - W_t)|} \end{aligned} \tag{44}$$

where C is some constant. It follows that

$$\mathbb{E} \left[\sup_{s \in [t,T]} |w(s, \hat{X}_s^{t,x})| \right] \leq \frac{x^\gamma}{\gamma} e^{|C(T-t)|}\, \mathbb{E} \left[\sup_{s \in [t,T]} e^{|\gamma\sigma\alpha(W_s - W_t)|} \right] < \infty.$$

Therefore, using the verification theorem, $w(t,x) = v(t,x)$ and α is an optimal control. □

3 Stochastic Volatility

3.1 Stochastic volatility models

In this section we fix a 2-dimensional reference space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W)$ and a constant $\rho \in [-1, 1]$. We define a new Brownian motion $W_t^* = \rho W_t^1 + \sqrt{1-\rho^2} W_t^2$, so that the correlation between W_t^* and W_t^1 is ρ.

We now generalize the Black-Scholes model by adding an extra stochastic process, and letting the mean and the volatility of the stock be functions of it. For the extra stochastic differential equation we shall always consider its domain $D = \mathbb{R}$. Take continuous mappings $A, B : \mathbb{R} \to \mathbb{R}$. Suppose that for these mappings the stochastic differential equation

$$dZ_s = B(Z_s)\,ds + A(Z_s)\,dW_s^*, \quad Z_t = z, \tag{45}$$

has a solution for any starting time $t \in [0, T]$ and initial condition $z \in \mathbb{R}$. Our new *stochastic volatility* market model will be the following system of stochastic differential equations

$$\begin{cases} dB_s = rB_s\,ds, \\ dS_s = \mu(Z_s)S_s\,ds + e^{Z_s}S_s\,dW_s^1, \\ dZ_s = B(Z_s)\,ds + A(Z_s)\,dW_s^*, \end{cases} \tag{46}$$

where $r \geq 0$ and $\mu : \mathbb{R} \to \mathbb{R}$ is a continuous function.

We have generalized the notion of a stochastic volatility model so the volatility process is always the exponential of the extra process. This is always possible if one does the appropriate transformations. For instance if the volatility was of the form $\sigma(Y_s)$, for some process Y_s, with an associated stochastic differential equation, and σ, a bijective positive function, we can always take $Z_s = \log(\sigma(Y_s))$ and find the equation for Z_s by applying Itô's formula on $\log(\sigma(Y_s))$ and substituting Y_s by $\sigma^{-1}(e^{Z_s})$.

There are many examples of stochastic volatility models. For instance there is the popular Heston model, which corresponds to the case

$$\begin{cases} dB_t = rB_t\,dt, \\ dS_t = \mu(Y_t)S_t\,dt + \sqrt{Y_t}S_t\,dW_t^1, \\ dY_t = \kappa(\theta - Y_t)\,dt + \alpha\sqrt{Y_t}\,dW_t^*, \end{cases} \tag{47}$$

where μ is some continuous function, and κ, θ, α are non-negative constants satisfying

$$2\kappa\theta > \alpha^2,$$

the so called Feller condition. The equation of Y_t is over $\mathbb{R}^+$, because we have assumed the Feller condition, so to get to the form we have presented before we take $Z_s = \log(\sqrt{Y_s})$ which if one applies the Itô formula and substitutes Y_t by e^{2Z_t} we have the stochastic differential equation

$$dZ_t = \left[\frac{1}{4}(2\kappa\theta - \alpha^2)e^{-2Z_t} - \frac{1}{2}\kappa\right]dt + \frac{1}{2}\alpha e^{-Z_t}\,dW_t^*.$$

We end this section by giving a rigorous definition of what is a stochastic volatility model.

Definition 3.1.1. *A stochastic volatility model is an ordered vector (A, B, μ), where $\mu : \mathbb{R} \to \mathbb{R}$ is a continuous mapping, and $A, B : \mathbb{R} \to \mathbb{R}$ are continuous mappings, such that $A(z) \neq 0$ for all $z \in \mathbb{R}$, and the stochastic differential equation*

$$dZ_s = B(Z_s)\,ds + A(Z_s)\,dW_s^*, \quad Z_t = z, \tag{48}$$

has an unique solution over $[t, T]$, for any starting conditions $(t, z) \in [0, T] \times \mathbb{R}$.

We could take $A(z)$ to be always positive since if one finds a model where $A(z)$ is negative we can take a different Brownian motion where the correlation is $-\rho$. To avoid having to make these transformations we present this definition this way.

3.2 Formulation of the problem and the verification theorem

We fix now a stochastic volatility model (A, B, μ). For the formulation of the control problem, we can follow a similar approach as the one we did in the Merton problem. Fix $(t, z) \in [0, T] \times \mathbb{R}$ and Z_s the solution to (48) for these starting conditions. Following the same steps as in the Merton problem, considering only self-financed portfolios where the wealth process remains positive and representing by φ_s the fraction of our wealth invested in the stock, we get the stochastic differential equation for the wealth process

$$dX_s = (r + (\mu(Z_s) - r)\varphi_s)X_s\, ds + e^{Z_s}\varphi_s X_s\, dW_s^1, \quad X_t = x, \tag{49}$$

for starting condition $x > 0$. Since μ is continuous and so is Z_s, this linear stochastic differential equation has a solution has long as

$$\int_t^T \varphi(s, \omega)\, ds < \infty, \quad a.s. - \mathbb{P}$$

So it makes sense to define the set of admissible controls in the same way as in the Merton problem. We give the definition again for easy reference .

Definition 3.2.1. *We define the set of admissible controls, $\mathcal{A}(t)$, defined on $t \in [0, T]$, to be the set of adapted stochastic processes, φ, with domain $[t, T] \times \Omega$, such that*

$$\int_t^T |\varphi_s|^2\, ds < \infty, \quad a.s. \tag{50}$$

One can note that by supposing that the admissible controls are adapted stochastic processes, they can be processes of the form $\alpha(s, Z_s)$, and so we're in a sense supposing that the volatility is perfectly observable, which is not entirely realistic. Yet one can use implied volatilities wich is actually observable and infer the value of the instantaneous volatility. Another possibility can be to use high frequency data (we refer to [1] for this).

We use the same concept for utility functions as before

Definition 3.2.2. *An utility function is a function $U \in C^2(\mathbb{R}^+)$ such that*

$$U(x) \geq 0, \quad U'(x) \geq 0, \quad \textit{and} \quad U''(x) \leq 0, \quad \textit{for all } x \in \mathbb{R}^+$$

We can now give the definitions of the Payoff function and the value function, which are slightly different then the Merton problem.

Definition 3.2.3. *For each $(t, z) \in [0, T] \times \mathbb{R}$ we correspond a process, $Z_s^{t,z}$, which is a solution of equation (48) for starting conditions (t, z). Consider now the set*

$$\mathcal{G} = \left\{ (t, x, z, \varphi) \,\middle|\, t \in [0, T],\, x \in \mathbb{R}^+,\, z \in \mathbb{R},\, \varphi \in \mathcal{A}(t) \right\}. \tag{51}$$

For each $(t, x, z, \varphi) \in \mathcal{G}$ we correspond a process, $X_s^{t,x,z,\varphi}$, which is a solution of the equation

$$dX_s^{t,x,z,\varphi} = (r + (\mu(Z_s^{t,z}) - r)\varphi_s)X_s^{t,x,z,\varphi}\, ds + e^{Z_s^{t,z}}\varphi_s X_s^{t,x,z,\varphi}\, dW_s^1, \quad X_t^{t,x,z,\varphi} = x, \tag{52}$$

Definition 3.2.4. *We define the payoff function, $P : \mathcal{G} \to \mathbb{R} \cup \{+\infty\}$, for the utility function U, to be*

$$P(t, x, z, \varphi) = \mathbb{E}\left[U(X_T^{t,x,z,\varphi}) \right], \tag{53}$$

and the value function $V : [0, T] \times \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ to be

$$V(t, x, z) = \sup_{\varphi \in \mathcal{A}(t)} P(t, x, z, \varphi). \tag{54}$$

We also denote the value function V by the lower case v.

Once again, one can prove that the payoff is uniquely defined for any of the mappings we take, since the solutions of the stochastic differential equations are pathwise unique.

Our objective is, for each $(t,x,z) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R}$, find a control $\varphi_s \in \mathcal{A}(t)$ such that $P(t,x,z,\varphi) = V(t,x,z)$. For this we develop the HJB equations. We shall now prove the Verification theorem, as was done for the Merton problem. We start with a Lemma.

Lemma 3.2.5. *Let $w \in C^{1,2}([0,T] \times (\mathbb{R}^+ \times \mathbb{R}))$ be a non-negative function, and let U be an utility function. Suppose that*

1. *for fixed $(t,x,z) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R}$, the mapping*
$$a \to \left[\frac{\partial w}{\partial x} x(r + a(\mu(z) - r)) + \frac{\partial^2 w}{\partial x \partial z} a e^z x A(z) \rho + \frac{1}{2}\frac{\partial^2 w}{\partial x^2} e^{2z} a^2 x^2\right],$$
is bounded above for $a \in \mathbb{R}$;

2. *w satisfies the inequality*
$$\begin{aligned} &-\frac{\partial w}{\partial t} - \frac{\partial w}{\partial z} B(z) - \frac{1}{2}\frac{\partial^2 w}{\partial z^2}(t,x,z) A(z)^2 \\ &\qquad - \sup_{a \in \mathbb{R}} \left[\frac{\partial w}{\partial x} x(r + a(\mu(z) - r)) + \frac{\partial^2 w}{\partial x \partial z} a e^z x A(z) \rho + \frac{1}{2}\frac{\partial^2 w}{\partial x^2} e^{2z} a^2 x^2\right] \geq 0 \end{aligned} \tag{55}$$
in $[0,T] \times \mathbb{R}^+ \times \mathbb{R}$;

3. *$w(T,x,z) \geq U(x), \quad \forall (x,z) \in \mathbb{R}^+ \times \mathbb{R}$.*

Then $w \geq v$ on $[0,T] \times \mathbb{R}^+ \times \mathbb{R}$.

Proof. Let $(t,x,z,\varphi) \in \mathcal{G}$. Take $Z_s^{t,z}$, the solution to equation (48) for starting conditions (t,z), and $X_s^{t,x,z,\varphi}$, as in definition 3.2.3. Consider a stopping time θ with values in $[t,T]$. Since $w \in C^{1,2}([0,T] \times (\mathbb{R}^+ \times \mathbb{R}))$ we may apply Itô's formula on $w(\theta, X_\theta^{t,x,z,\varphi}, Z_\theta^{t,z})$ and obtain

$$\begin{aligned} w(\theta, X_\theta^{t,x,z,\varphi}, Z_\theta^{t,z}) = w(t,x,z) &+ \int_t^T 1_{[t,\theta]}(s) \frac{\partial w}{\partial t}(\cdots)\, ds \\ &+ \int_t^T 1_{[t,\theta]}(s) \frac{\partial w}{\partial x}(\cdots) X_s^{t,x,z,\varphi}(r + (\mu(Z_s^{t,z}) - r)\varphi_s)\, ds \\ &+ \int_t^T 1_{[t,\theta]}(s) \frac{\partial w}{\partial z}(\cdots) B(Z_s^{t,z})\, ds \\ &+ \frac{1}{2}\int_t^T 1_{[t,\theta]}(s) \frac{\partial^2 w}{\partial x^2}(\cdots) \left(X_s^{t,x,\varphi}\right)^2 \left(\varphi_s e^{Z_s^{t,z}}\right)^2 ds \\ &+ \int_t^T 1_{[t,\theta]}(s) \frac{\partial^2 w}{\partial x \partial z}(\cdots) X_s^{t,x,\varphi} e^{Z_s^{t,z}} \varphi_s\, A(Z_s^{t,z})\, \rho\, ds \\ &+ \frac{1}{2}\int_t^T 1_{[t,\theta]}(s) \frac{\partial^2 w}{\partial z^2}(\cdots) \left(A(Z_s^{t,z})\right)^2 ds \\ &+ \int_t^T 1_{[t,\theta]}(s) \frac{\partial w}{\partial x}(\cdots) X_s^{t,x,\varphi} e^{Z_s^{t,z}} \varphi_s\, dW_s^1 \\ &+ \int_t^T 1_{[t,\theta]}(s) \frac{\partial w}{\partial z}(\cdots) A(Z_s^{t,z})\, dW_s^* \end{aligned} \tag{56}$$

a.s.-$\mathbb{P}$, where the "$(\cdots)$" are abbreviations for $(s, X_s^{t,x,z,\varphi}, Z_s^{t,z})$. We define the stopping time, θ_n, for $n \in \mathbb{N}$, and valued in $[t,T]$, by

$$\theta_n(\omega) = \begin{cases} \inf\Big\{ s \in [t,T] \,\Big|\, \int_t^s |\frac{\partial w}{\partial x}(\cdots) X_u^{t,x,z,\varphi} e^{Z_u^{t,z}} \varphi_u|^2\, du \geq n \\ \qquad \vee \int_t^s |\frac{\partial w}{\partial z}(\cdots) A(Z_u^{t,z})|^2\, du \geq n \Big\}, & if\ \{s \in [t,T] \,|\, \cdots\} \neq \emptyset; \\ T, & if \{s \in [t,T] \,|\, \cdots\} = \emptyset. \end{cases}$$

Then for these stopping times, the stochastic integrals terms in (56) have expectation 0. On the other hand we have, by condition *2*

$$
\begin{aligned}
&\frac{\partial w}{\partial t}(\cdots) + \frac{\partial w}{\partial x}(\cdots) X_s^{t,x,z,\varphi}(r + (\mu(Z_s^{t,z}) - r)\varphi_s) + \frac{\partial w}{\partial z}(\cdots) B(Z_s^{t,z}) \\
&\qquad + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(\cdots)\left(X_s^{t,x,\varphi}\right)^2 \left(\varphi_s e^{Z_s^{t,z}}\right)^2 + \frac{\partial^2 w}{\partial x \partial z}(\cdots) X_s^{t,x,\varphi} \varphi_s e^{Z_s^{t,z}} A(Z_s^{t,z})\rho \\
&\qquad + \frac{1}{2}\frac{\partial^2 w}{\partial z^2}(\cdots)\left(A(Z_s^{t,z})\right)^2 \\
\leq\ & \frac{\partial w}{\partial t}(\cdots) + \frac{\partial w}{\partial z}(\cdots) B(Z_s^{t,z}) + \frac{1}{2}\frac{\partial^2 w}{\partial z^2}(\cdots)\left(A(Z_s^{t,z})\right)^2 \\
&\qquad + \sup_{a \in \mathbb{R}} \left\{ \frac{\partial w}{\partial x}(\cdots) X_s^{t,x,z,\varphi}(r + (\mu(Z_s^{t,z}) - r)\, a) + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}(\cdots)\left(X_s^{t,x,\varphi}\right)^2 a^2 e^{2Z_s^{t,z}} \right. \\
&\qquad\qquad\qquad \left. + \frac{\partial^2 w}{\partial x \partial z}(\cdots)\, a\, X_s^{t,x,\varphi} e^{Z_s^{t,z}} A(Z_s^{t,z})\rho \right\} \\
\leq\ & 0
\end{aligned} \tag{57}
$$

for each fixed $s \in [t, T]$ and $\omega \in \Omega$.

Given all these observations, taking the expectation in (56) for θ_n, we have

$$
\mathbb{E}\left[w(\theta_n, X_{\theta_n}^{t,x,z,\varphi}, Z_{\theta_n}^{t,z})\right] \leq w(t,x,z). \tag{58}
$$

Since w is a continuous non-negative function and $\lim_{n\to\infty} \theta_n = T$, a.s.-$\mathbb{P}$, we may use Fatou's Lemma to obtain

$$
\mathbb{E}\left[U(X_T^{t,x,z,\varphi})\right] \leq \mathbb{E}\left[w(T, X_T^{t,s,z,\varphi}, Z_T^{t,z})\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[w(\theta_n, X_{\theta_n}^{t,x,z,\varphi}, Z_{\theta_n}^{t,z})\right] \leq w(t,x,z). \tag{59}
$$

This holds for all controls, so by taking the sup in $\varphi \in \mathcal{A}(t)$, we have

$$
V(t,x,z) \leq w(t,x,z),
$$

□

We present now the verification theorem.

Theorem 3.2.6. *(Verification theorem) Let $w \in C^{1,2}([0,T] \times (\mathbb{R}^+ \times \mathbb{R}))$ be a non-negative function and let U be an utility function. Suppose that*

1. *for fixed $(t,x,z) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R}$, the mapping*

$$
a \to \left[\frac{\partial w}{\partial x} x(r + a(\mu(z) - r)) + \frac{\partial^2 w}{\partial x \partial z} a e^z x A(z)\rho + \frac{1}{2}\frac{\partial^2 w}{\partial x^2} e^{2z} a^2 x^2\right],
$$

   is bounded above for $a \in \mathbb{R}$;

2. *w satisfies the equality*

$$
\begin{aligned}
&-\frac{\partial w}{\partial t} - \frac{\partial w}{\partial z} B(z) - \frac{1}{2}\frac{\partial^2 w}{\partial z^2}(t,x,z) A(z)^2 \\
&\qquad - \sup_{a \in \mathbb{R}} \left[\frac{\partial w}{\partial x} x(r + a(\mu(z) - r)) + \frac{\partial^2 w}{\partial x \partial z} a e^z x A(z)\rho + \frac{1}{2}\frac{\partial^2 w}{\partial x^2} e^{2z} a^2 x^2\right] = 0
\end{aligned} \tag{60}
$$

   in $[0,T] \times \mathbb{R}^+ \times \mathbb{R}$;

3. *$w(T,x,z) = U(x), \quad \forall (x,z) \in \mathbb{R}^+ \times \mathbb{R}$.*

Let $\alpha : [0,T] \times \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ be a function such that

$$\alpha(t,x,z) \in \arg\max_{a\in\mathbb{R}} \left\{ \frac{\partial w}{\partial x} x(r + a(\mu(z) - r)) + \frac{\partial^2 w}{\partial x \partial z} a e^z x A(z)\rho + \frac{1}{2}\frac{\partial^2 w}{\partial x^2} e^{2z} a^2 x^2 \right\}, \tag{61}$$

on $[0,T] \times \mathbb{R}^+ \times \mathbb{R}$ (such a function always exists by condition 1).

For $(t,x,z) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R}$ consider $Z_s^{t,z}$, as in definition 3.2.3, and suppose that the equation

$$d\hat{X}_s^{t,x,z} = (r + (\mu(Z_s^{t,z}) - r)\alpha(t, \hat{X}_s^{t,x,z}, Z_s^{t,z}))\hat{X}_s^{t,x,z}\, ds + e^{Z_s^{t,z}} \alpha(t, \hat{X}_s^{t,x,z}, Z_s^{t,z})\hat{X}_s^{t,x,z}\, dW_s^1, \tag{62}$$

has a solution, $\hat{X}_s^{t,x,z}$, on $[t,T]$ with starting value x. If $\alpha(t, \hat{X}_s^{t,x,z}, Z_s^{t,z}) \in \mathcal{A}(t)$ and for any sequence, $(\tau_n)_{n\in\mathbb{N}}$, of stopping times in $[t,T]$ the collection

$$\left\{ w(\tau_n, \hat{X}_{\tau_n}^{t,x,z}, Z_{\tau_n}^{t,z}) \right\}_{n\in\mathbb{N}} \tag{63}$$

is uniformly integrable, we have

$$w(t,x,z) = P(t,x,z,\alpha(s, \hat{X}_s^{t,x,z}, Z_s^{t,z})) = v(t,x,z).$$

Proof. By the last Lemma it's immediate that $w \geq v$ on $[0,T] \times \mathbb{R}^+ \times \mathbb{R}$. Fix now $(t,x,z) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R}$. Denote by $\varphi_s^* = \alpha(s, \hat{X}_s^{t,x,z}, Z_s^{t,z})$. It should be immediate that $\hat{X}_s^{t,x,z} = X_s^{t,x,z,\varphi^*}$ for all $s \in [t,T]$, a.s.-$\mathbb{P}$, using the pathwise uniqueness property of linear stochastic differential equations.

In an analogous argumentation as in last Lemma, we may take Itô's formula on the process $w(\theta_n, X_{\theta_n}^{t,x,z,\varphi^*}, Z_{\theta_n}^{t,z})$, where θ_n is the same stopping time as in the Lemma, and obtain

$$\mathbb{E}\left[w(\theta_n, X_{\theta_n}^{t,x,z,\varphi^*}, Z_{\theta_n}^{t,z}) \right] = w(t,x,z),$$

for $s \in [t,T]$, where this time we have the equality by the stronger conditions on the function w and condition (61) on α.

Now since

$$\lim_{n\to\infty} \theta_n = T, \quad a.s. - \mathbb{P},$$

and by hypothesis the collection

$$\left\{ w(\theta_n, \hat{X}_{\theta_n}^{t,x,z}, Z_{\theta_n}^{t,z}) \right\}_{n\in\mathbb{N}}$$

is uniformly integrable, we have

$$\mathbb{E}\left[U(X_T^{t,x,z,\varphi^*}) \right] = \mathbb{E}\left[w(T, X_T^{t,x,z,\varphi^*}, Z_T^{t,z}) \right] = \lim_{n\to\infty} \mathbb{E}\left[w(\theta_n, X_{\theta_n}^{t,x,z,\varphi^*}, Z_{\theta_n}^{t,z}) \right] = w(t,x,z). \tag{64}$$

Since we have always $v(t,x,z) \geq P(t,x,z,\varphi^*)$ it follows that $v(t,x,z) \geq w(t,x,z)$. But we also have $w \geq v$ on $[0,T] \times \mathbb{R}^+ \times \mathbb{R}$, therefore

$$w(t,x,z) = v(t,x,z)$$

and since $\mathbb{E}\left[U(X_T^{t,x,z,\varphi^*}) \right] = w(t,x,z)$ we have that $\alpha(s, \hat{X}_s^{t,x,z}, Z_s^{t,z})$ is an optimal control. □

Our proof of the verification theorems is inspired by the one done by Pham [19] for the case where the control equation is Lipschitz. Kraft [13] proves the Lemma in a quicker way using properties of local martingales, but for the full Verification theorem he needs to use stopping times like here. We have decided to present the proof this way because it seems simpler.

3.3 Power utilities in stochastic volatility models

We shall now consider power utilities, i.e., the case $U(x) = \frac{x^\gamma}{\gamma}$ for some $\gamma \in]0,1[$. For these utilities the HJB equation will simplify to a linear equation. This result is due to Zariphopoulou [23].

First we note the following

$$v(t,x,z) = \sup_{\varphi\in\mathcal{A}(t)} \mathbb{E}\left[\left(X_T^{t,x,z,\varphi}\right)^\gamma \Big/ \gamma\right] = \frac{1}{\gamma} \sup_{\varphi\in\mathcal{A}(t)} \mathbb{E}\left[\left(xe^{(\cdots)}\right)^\gamma\right] = \frac{x^\gamma}{\gamma} \sup_{\varphi\in\mathcal{A}(t)} \mathbb{E}\left[\left(e^{(\cdots)}\right)^\gamma\right],$$

i.e., the value function for these utilities always has the form $\frac{x^\gamma}{\gamma}G(t,z)$, where $G : [0,T]\times\mathbb{R} \to \mathbb{R}^+$ is a positive function (we suppose the above sup is finite).

The HJB equation for the stochastic volatility problem is

$$\begin{aligned} &-\frac{\partial w}{\partial t} - \frac{\partial w}{\partial z}B(z) - \frac{1}{2}\frac{\partial^2 w}{\partial z^2}A(z)^2 \\ &\quad - \sup_{a\in\mathbb{R}}\left[\frac{\partial w}{\partial x}x(r + a(\mu(z)-r)) + \frac{\partial^2 w}{\partial x \partial z}ae^z xA(z)\rho + \frac{1}{2}\frac{\partial^2 w}{\partial x^2}e^{2z}a^2x^2\right] = 0, \end{aligned} \tag{65}$$

on $[0,T]\times\mathbb{R}^+\times\mathbb{R}$, with final condition $w(T,x,z) = x^\gamma/\gamma$ on $\mathbb{R}^+\times\mathbb{R}$. Suppose that this equation has smooth solution, w, and that it admits a decomposition of the form $w = G(t,z)\,x^\gamma/\gamma$ where $G(t,z) \in C^{1,2}([0,T]\times\mathbb{R})$ is a positive function. Rewriting the above equation in terms of $G(t,z)$ we get

$$\begin{aligned} &-\frac{\partial G}{\partial t} - \frac{\partial G}{\partial z}B(z) - \frac{1}{2}\frac{\partial^2 G}{\partial z^2}A(z)^2 \\ &\quad - \gamma\sup_{a\in\mathbb{R}}\left[G(t,z)(r + a(\mu(z)-r)) + \frac{\partial G}{\partial z}ae^zA(z)\rho + \frac{1}{2}G(t,z)(\gamma-1)e^{2z}a^2\right] = 0, \end{aligned} \tag{66}$$

on $[0,T]\times\mathbb{R}$, with $G(T,z) = 1$ on $\mathbb{R}$. Since $G(t,z) > 0$ the above supremum always achieves its maximum, because the coefficient of a^2 will always be negative. Applying the derivative on a we find this maximum to be

$$a^* = \frac{\mu(z)-r}{(1-\gamma)e^{2z}} + \frac{A(z)\rho}{(1-\gamma)e^z}\frac{\partial G}{\partial z}\frac{1}{G(t,z)}$$

for each $t\in[0,T]$, $z\in\mathbb{R}$. So the equation simplifies to

$$\begin{aligned} &\frac{\partial G}{\partial t} + \frac{\partial G}{\partial z}B(z) + \frac{1}{2}\frac{\partial^2 G}{\partial z^2}A(z)^2 + G(t,z)\,r\,\gamma \\ &\quad + \frac{1}{2}G(t,z)\frac{\gamma(\mu(z)-r)^2}{(1-\gamma)e^{2z}} + \frac{\partial G}{\partial z}\frac{\gamma(\mu(z)-r)A(z)\rho}{(1-\gamma)e^z} + \frac{1}{2}\frac{\gamma\,(A(z))^2\,\rho^2}{(1-\gamma)}\left(\frac{\partial G}{\partial z}\right)^2\frac{1}{G(t,z)} = 0. \end{aligned} \tag{67}$$

This, being non-linear, is not easily solvable. We shall transform it into a linear equation by considering a transformation of the form $G(t,z) = H(t,z)^\eta$, for some $\eta > 0$. Writing the equation in terms of $H(t,z)$ we have

$$\begin{aligned} &\eta H^{\eta-1}\frac{\partial H}{\partial t} + \eta H^{\eta-1}\frac{\partial H}{\partial z}B(z) + \frac{1}{2}\left(\eta(\eta-1)H^{\eta-2}\left(\frac{\partial H}{\partial z}\right)^2 + \eta H^{\eta-1}\frac{\partial^2 H}{\partial z^2}\right)A(z)^2 + \gamma r H^\eta \\ &\quad + \frac{1}{2}H^\eta\frac{\gamma(\mu(z)-r)^2}{(1-\gamma)e^{2z}} + \eta H^{\eta-1}\frac{\partial H}{\partial z}\frac{\gamma(\mu(z)-r)A(z)\rho}{(1-\gamma)e^z} + \frac{1}{2}\eta^2H^{\eta-2}\left(\frac{\partial H}{\partial z}\right)^2\frac{\gamma A(z)^2\rho^2}{(1-\gamma)} = 0. \end{aligned} \tag{68}$$

Adding the terms with $H^{\eta-2}\left(\frac{\partial H}{\partial z}\right)^2$ we get

$$\frac{1}{2}H^{\eta-2}\left(\frac{\partial H}{\partial z}\right)^2 A(z)^2\left(\eta(\eta-1)+\eta^2\frac{\gamma\rho^2}{(1-\gamma)}\right).$$

This suggest taking η so that $\eta(\eta-1)+\eta^2\frac{\gamma\rho^2}{(1-\gamma)}=0$, which corresponds to

$$\eta=\frac{1-\gamma}{1-\gamma+\gamma\rho^2}.$$

So, taking this η in equation (68) and multiplying by $\frac{1}{\eta}H^{1-\eta}$ we have

$$\begin{aligned}\frac{\partial H}{\partial t}+\frac{\partial H}{\partial z}B(z)+\frac{1}{2}\frac{\partial^2 H}{\partial z^2}A(z)^2+\frac{\gamma r}{\eta}H\\ +\frac{1}{2}H\frac{\gamma(\mu(z)-r)^2}{\eta(1-\gamma)e^{2z}}+\frac{\partial H}{\partial z}\frac{\gamma(\mu(z)-r)A(z)\rho}{(1-\gamma)e^z}=0.\end{aligned}\tag{69}$$

Finally, we introduce the transformation

$$H(t,z)=e^{\frac{\gamma r}{\eta}(T-t)}F(t,z)$$

which, writing the above equation in terms of F, one gets

$$\frac{\partial F}{\partial t}+\frac{\partial F}{\partial z}B(z)+\frac{1}{2}\frac{\partial^2 F}{\partial z^2}A(z)^2+\frac{1}{2}F\frac{\gamma(\mu(z)-r)^2}{\eta(1-\gamma)e^{2z}}+\frac{\partial F}{\partial z}\frac{\gamma(\mu(z)-r)A(z)\rho}{(1-\gamma)e^z}=0. \tag{70}$$

It follows that finding a solution to the HJB equation is equivalent to finding a solution to this linear differential equation. We have then a simpler Verification theorem for these utility functions.

We present now the Verification theorem for these utilities.

Theorem 3.3.1. *(Verification theorem for power utility functions) Let $\gamma\in]0,1[$, and define*

$$\eta=\frac{1-\gamma}{1-\gamma+\gamma\rho^2}.$$

Let $F\in C^{1,2}([0,T]\times\mathbb{R})$ be a positive function satisfying

$$\frac{\partial F}{\partial t}+\frac{\partial F}{\partial z}B(z)+\frac{1}{2}\frac{\partial^2 F}{\partial z^2}A(z)^2+\frac{1}{2}F\frac{\gamma(\mu(z)-r)^2}{\eta(1-\gamma)e^{2z}}+\frac{\partial F}{\partial z}\frac{\gamma(\mu(z)-r)A(z)\rho}{(1-\gamma)e^z}=0, \tag{71}$$

on $[0,T]\times\mathbb{R}$ and $F(T,z)=1$ on $\mathbb{R}$.

Let $\alpha:[0,T]\times\mathbb{R}\to\mathbb{R}$ be the function defined by

$$\alpha(t,z)=\frac{\mu(z)-r}{(1-\gamma)e^{2z}}+\frac{\eta\rho A(z)}{(1-\gamma)e^z}\frac{1}{F(t,z)}\frac{\partial F}{\partial z}(t,z). \tag{72}$$

For $(t,x,z)\in[0,T]\times\mathbb{R}^+\times\mathbb{R}$ consider the process $Z_s^{t,z}$, as in definition 3.2.3, and let $\hat{X}_s^{t,x,z}$ be the solution of the linear stochastic differential equation

$$d\hat{X}_s^{t,x,z}=(r+(\mu(Z_s^{t,z})-r)\alpha(t,Z_s^{t,z}))\hat{X}_s^{t,x,z}\,ds+e^{Z_s^{t,z}}\alpha(t,Z_s^{t,z})\hat{X}_s^{t,x,z}\,dW_s^1,\quad \hat{X}_t^{t,x,z}=x. \tag{73}$$

Then, for the utility function $U(x)=\frac{x^\gamma}{\gamma}$, the function

$$w(t,x,z)=\frac{x^\gamma}{\gamma}e^{\gamma r(T-t)}F(t,z)^\eta$$

is a positive function satisfying conditions 1,2 and 3 of the Verification theorem ($Theorem$ 3.2.6)*, and if for any sequence,* $(\tau_n)_{n\in\mathbb{N}}$*, of stopping times in* $[t,T]$ *the collection*

$$\left\{w(\tau_n, \hat{X}^{t,x,z}_{\tau_n}, Z^{t,z}_{\tau_n})\right\}_{n\in\mathbb{N}}$$

is uniformly integrable, we have

$$w(t,x,z) = P(t,x,z,\alpha(s,Z^{t,z}_s)) = v(t,x,z)$$

Proof. Firstly, to check that w is a positive function satisfying conditions *1,2* and *3* of the Verification theorem simply follow the steps we did backwards. Consider this w now and lets apply the Verification theorem under our suppositions.

Lets start by finding a function $\overline{\alpha}$ satisfying condition (61), i.e.,

$$\overline{\alpha}(t,x,z) \in \arg\max_{a\in\mathbb{R}} \left\{ \frac{\partial w}{\partial x} x(r + a(\mu(z) - r)) + \frac{\partial^2 w}{\partial x \partial z} a e^z x A(z)\rho + \frac{1}{2}\frac{\partial^2 w}{\partial x^2} e^{2z} a^2 x^2 \right\}. \tag{74}$$

Since w is a strictly concave function in x, this max is unique. Applying the derivative on a on the above expression and writing the derivatives of w in terms of F we have

$$x^\gamma e^{\gamma r(T-t)} F(t,z)^\eta (\mu(z)-r) + x^\gamma e^{\gamma r(T-t)} \eta \frac{\partial F}{\partial z} F(t,z)^{\eta-1} e^z A(z)\rho + (\gamma-1) x^\gamma e^{\gamma r(T-t)} F(t,z)^\eta e^{2z} a,$$

so finding the 0 in a we get

$$\overline{\alpha}(t,x,z) = \frac{\mu(z) - r}{(1-\gamma)e^{2z}} + \frac{\eta\rho A(z)}{(1-\gamma)e^z} \frac{1}{F(t,z)} \frac{\partial F}{\partial z}.$$

It follows that the function $\overline{\alpha}$ that satisfies condition (61) of the original Verification theorem is the α we have announced in this Theorem (technically $\overline{\alpha}$ is a function of one more variable but this makes no difference). Therefore equation (73) is the same equation as (62) in the Verification theorem.

Now since $\alpha(t, Z^{t,z}_s) \in \mathcal{A}(t)$ always, being continuous, and the uniform integrability condition is assumed we may apply the Verification theorem and conclude

$$w(t,x,z) = P(t,x,z,\alpha(s,Z^{t,z}_s)) = v(t,x,z).$$

□

We finish this section with a case where the solution is very simple.

Example 3.3.1. *Take* $\mu(z) = r + \lambda e^z$*, for some* $\lambda > 0$*. In this case, equation* (71) *simplifies to*

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial z}B(z) + \frac{1}{2}\frac{\partial^2 F}{\partial z^2}A(z)^2 + \frac{1}{2}F\frac{\gamma}{\eta(1-\gamma)}\lambda^2 + \frac{\partial F}{\partial z}\frac{\gamma A(z)\rho}{(1-\gamma)}\lambda = 0.$$

on $[0,T]\times\mathbb{R}$ *and* $F(T,z) = 1$ *on* $\mathbb{R}$.

Considering the transformation $F(t,z) = \exp\left\{\frac{\gamma}{\eta(1-\gamma)}\lambda^2(T-t)\right\}\overline{F}(t,z)$*, one gets the equation*

$$\frac{\partial \overline{F}}{\partial t} + \frac{\partial \overline{F}}{\partial z}B(z) + \frac{1}{2}\frac{\partial^2 \overline{F}}{\partial z^2}A(z)^2 + \frac{\partial \overline{F}}{\partial z}\frac{\gamma A(z)\rho}{(1-\gamma)}\lambda = 0.$$

on $[0,T]\times\mathbb{R}$ *and* $\overline{F}(T,z) = 1$ *on* $\mathbb{R}$*. This equation has the simple solution* $\overline{F} = 1$*, has one can check. Therefore* $F(t,z) = \exp\left\{\frac{\gamma}{\eta(1-\gamma)}\lambda^2(T-t)\right\}$ *is a solution of the initially considered* *equation. Lets prove now the other conditions of the Verification theorem for this F. The function $\alpha : [0,T] \times \mathbb{R} \to \mathbb{R}$ given by (72) is*

$$\alpha(t,x,z) = \frac{\lambda}{(1-\gamma)e^z}.$$

Fix now $(t,x,z) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R}$. For this α we can check that the equation for $\hat{X}_s^{t,x,z}$ is simply

$$d\hat{X}_s^{t,x,z} = \left(r + \frac{\lambda^2}{1-\gamma}\right)\hat{X}_s^{t,x,z}\,ds + \frac{\lambda}{1-\gamma}\hat{X}_s^{t,x,z}\,dW_s^1,$$

which has a geometric Brownian motion solution. It follows, that

$$w(t,x,z) = \frac{x^\gamma}{\gamma}e^{\gamma r(T-t)}e^{\frac{\gamma}{1-\gamma}\lambda^2(T-t)}$$

satisfies condition 1,2 and 3 of the Verification theorem and since

$$\mathbb{E}\left[\sup_{s\in[t,T]}\left|\left(\hat{X}_s^{t,x,z}\right)^\gamma e^{\gamma r(T-t)}\left(F(s,Z_s^{t,z})\right)^\eta\right|\right] \leq C\,\mathbb{E}\left[\sup_{s\in[t,T]}\left(\hat{X}_s^{t,x,z}\right)^\gamma\right] < \infty$$

for some $C > 0$, where the expected value is finite because $\hat{X}_s^{t,x,z}$ is a geometric Brownian motion, the uniform integrability of the Verification theorem is verified. It follows by the Verification theorem that $w(t,x,z) = v(t,x,z)$ and

$$\alpha(s,Z_s^{t,x}) = \frac{\lambda}{(1-\gamma)e^{Z_s^{t,z}}}$$

is an optimal control for this starting condition (t,x,z).

4 Models based on the CIR process

4.1 Solving the HJB equation

In this section we suppose that $\rho \neq 0$. The stochastic volatility models we shall look at now, are models based on the Cox–Ingersoll–Ross process. These were proposed by Kraft [13], as a generalization of the model found by Liu, in [15], to have an explicit solution to the HJB equations. The CIR process is the solution to the stochastic differential equation

$$dY_s = \kappa(\theta - Y_s)\,ds + \delta\sqrt{Y_s}\,dW_s^*, \quad Y_t = y, \tag{75}$$

where $\kappa, \theta, \delta > 0$, are positive constants and $(t, y) \in [0, T] \times \mathbb{R}^+$ are the initial conditions. Under the condition $2\kappa\theta > \delta^2$, the so called Feller condition which we shall assume, the process Y_s is always positive, i.e., $Y_s > 0$ for all $s \in [t, T]$, almost surely.

Take $\nu \in \mathbb{R}\backslash\{0\}$. We shall consider a model where the risky asset is modelled by the stochastic differential equation

$$dS_s = S_s(r + \overline{\lambda}(Y_s)^{\frac{\nu/2+1}{\nu}})\,ds + S_s(Y_s)^{\frac{1}{\nu}}\,dW_s^1$$

where $\overline{\lambda} > 0$ is some constant. So for instance for $\nu = 2$ we get the model discussed by Liu [15], an Heston model with the drift of the stock to be linear on the square of the volatility, while for $\nu = -2$ we get the model proposed by Chakko and Viceira [2], where in this case we get a constant drift.

Lets now change the volatility process so it's like in our formulation. For this we want a process Z_s, satisfying $e^{Z_s} = (Y_s)^{1/\nu}$. Applying Ito's formula we have

$$d\left(\frac{1}{\nu}\log(Y_s)\right) = \left(\frac{1}{Y_s}\frac{1}{2\nu}(2\kappa\theta - \delta^2) - \frac{\kappa}{\nu}\right)\,ds + \frac{\delta}{\nu}\frac{1}{\sqrt{Y_s}}\,dW_s^*.$$

We can then rewrite this equation in terms of Z_s to obtain

$$dZ_s = \left(e^{-\nu Z_s}\frac{1}{2\nu}(2\kappa\theta - \delta^2) - \frac{\kappa}{\nu}\right)\,ds + \frac{\delta}{\nu}e^{-\frac{\nu}{2}Z_s}\,dW_s^*. \tag{76}$$

In terms of Z_s, the equation for the stock becomes

$$dS_s = S_s(r + \overline{\lambda}e^{(\frac{\nu}{2}+1)Z_s})\,ds + S_s e^{Z_s}\,dW_s^1.$$

Writing as in our formulation, it follows that this is a stochastic volatility model such that

$$B(z) = e^{-\nu z}\frac{1}{2\nu}(2\kappa\theta - \delta^2) - \frac{\kappa}{\nu},\ A(z) = \frac{\delta}{\nu}e^{-\frac{\nu}{2}z},\ \textit{and}\ \mu(z) = r + \overline{\lambda}e^{(\frac{\nu}{2}+1)z}.$$

That there is an unique solution of the associated equation, using A and B, follows from the fact that it's a transformation of the CIR process equation, and that the coefficients are locally Lipschitz.

For ease of notation take $b = \frac{1}{2}(2\kappa\theta - \delta^2)$, which is always positive by our supposition. In this model, equation (71) for F is then

$$\frac{\partial F}{\partial t} + \left(\frac{1}{\nu}be^{-\nu z} - \frac{\kappa}{\nu}\right)\frac{\partial F}{\partial z} + \frac{1}{2}\frac{\delta^2}{\nu^2}e^{-\nu z}\frac{\partial^2 F}{\partial z^2} + \frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2 e^{\nu z}F + \frac{\gamma\rho}{1-\gamma}\frac{\delta\overline{\lambda}}{\nu}\frac{\partial F}{\partial z} = 0 \tag{77}$$

on $[0, T] \times \mathbb{R}$ and $F(T, z) = 1$ on $\mathbb{R}$. We shall now make a change on the time variable, $s = T - t$. Defining $\overline{F}(s, z) = F(T - s, z)$ on $[0, T] \times \mathbb{R}$ we have the equation for $\overline{F}$

$$-\frac{\partial \overline{F}}{\partial s} + \left(\frac{1}{\nu}be^{-\nu z} - \frac{\kappa}{\nu}\right)\frac{\partial \overline{F}}{\partial z} + \frac{1}{2}\frac{\delta^2}{\nu^2}e^{-\nu z}\frac{\partial^2 \overline{F}}{\partial z^2} + \frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2 e^{\nu z}\overline{F} + \frac{\gamma\rho}{1-\gamma}\frac{\delta\overline{\lambda}}{\nu}\frac{\partial \overline{F}}{\partial z} = 0 \tag{78}$$

on $[0,T]\times\mathbb{R}$ and $\overline{F}(0,z)=1$. Now we shall prove that there's a solution to this equation of the form

$$\overline{F}(s,z)=e^{f(s)+g(s)e^{\nu z}}$$

where $f,g\in C^1([0,T])$ are functions to be determined. Given the initial condition $\overline{F}(0,z)=1$ we impose the condition $f(0)=g(0)=0$. Now note that

$$\frac{\partial\overline{F}}{\partial s}=(f'(s)+g'(s)e^{\nu z})e^{f(s)+g(s)e^{\nu z}},\quad \frac{\partial\overline{F}}{\partial z}=g(s)\nu e^{\nu z}e^{f(s)+g(s)e^{\nu z}},$$

and

$$\frac{\partial^2\overline{F}}{\partial z^2}=(g(s)\nu^2e^{\nu z}+g(s)^2\nu^2e^{2\nu z})e^{f(s)+g(s)e^{\nu z}}.$$

Plugging these expressions in the equation we obtain

$$\begin{aligned}&-f'(s)-g'(s)e^{\nu z}+\frac{1}{\nu}bg(s)\nu-\frac{\kappa}{\nu}g(s)\nu e^{\nu z}+\frac{1}{2}\frac{\delta^2}{\nu^2}g(s)\nu^2\\&+\frac{1}{2}\frac{\delta^2}{\nu^2}g(s)^2\nu^2e^{\nu z}+\frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2e^{\nu z}+\frac{\gamma\rho}{1-\gamma}\frac{\delta\overline{\lambda}}{\nu}g(s)\nu e^{\nu z}=0,\end{aligned}\tag{79}$$

which can be simplified to

$$\begin{aligned}&-f'(s)+bg(s)+\frac{1}{2}\delta^2g(s)\\&e^{\nu z}\left(-g'(s)-\left(\kappa-\frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)g(s)+\frac{1}{2}\delta^2g(s)^2+\frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2\right)=0.\end{aligned}\tag{80}$$

It follows that if we find g and f satisfying

$$-g'(s)-\left(\kappa-\frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)g(s)+\frac{1}{2}\delta^2g(s)^2+\frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2=0,\tag{81}$$

$$-f'(s)+bg(s)+\frac{1}{2}\delta^2g(s)=0,\tag{82}$$

with $g(0)=f(0)=0$, we will have found a solution to the initial equation. The first equation is a Riccati type ODE. One can only solve it under certain conditions. For this suppose there is $h\in C^2([0,T])$ with $h(s)>0$ on $[0,T]$ and such that

$$g(s)=-\frac{2}{\delta^2}\frac{h'(s)}{h(s)},\quad \textit{on } [0,T].$$

Plugging this equality on the above equation and simplifying, one obtains the equation for h

$$h''(s)+h'(s)\left(\kappa-\frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)+\frac{1}{4}\frac{\delta^2\gamma}{\eta(1-\gamma)}\overline{\lambda}^2h(s)=0.$$

Since we need h to be always positive, we see that if the characteristic polynomial of the equation has no roots, we have no good solution, since these solutions always reach 0 eventually. If it has one, and only one, root, it might be possible, but we ignore this case since it only happens on a very specific case for our parameters. We suppose then that there are two roots to the characteristic polynomial. This is of course equivalent to supposing that the discriminant is positive

$$\left(\kappa-\frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)^2-\frac{\delta^2\gamma}{\eta(1-\gamma)}\overline{\lambda}^2>0.$$

Noting that $\eta = \frac{1-\gamma}{1-\gamma+\gamma\rho^2}$ and developing the square, we can simplify this condition to be

$$\kappa^2 > \frac{\delta^2\gamma}{1-\gamma}\overline{\lambda}^2 + 2\kappa\frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}. \tag{83}$$

Note that $\frac{\delta^2\gamma}{1-\gamma}\overline{\lambda}^2 > 0$ and $\kappa > 0$ so under the above condition we have

$$\frac{1}{2}\kappa > \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}$$

since $\kappa > 0$ we also have $\kappa > \frac{1}{2}\kappa$ so

$$\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda} > 0,$$

i.e., above, the term multiplying by $h'(s)$ is always positive.

Denote now by r_1 and r_2 the roots of the polynomial, i.e.,

$$r_1, r_2 = -\frac{1}{2}\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right) \pm \frac{1}{2}\sqrt{\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)^2 - \frac{\delta^2\gamma}{\eta(1-\gamma)}\overline{\lambda}^2}.$$

With the previous observations we see that both of these roots are always negative. Lets take the roots such that $r_1 < r_2 < 0$. The general solution of the second order ODE for h is given by

$$h(s) = C_1 e^{r_1 s} + C_2 e^{r_2 s}$$

where $C_1, C_2 \in \mathbb{R}$ are some constants. We need that $h'(0) = 0$ so that $g(0) = 0$. Therefore we need the equality

$$C_1 r_1 + C_2 r_2 = 0$$

Take then $C_1 = -1$ and $C_2 = \frac{r_1}{r_2}$. We note that since $r_1 < r_2 < 0$ we have $\frac{r_1}{r_2} > 1$. Therefore

$$-e^{r_1 s} + \frac{r_1}{r_2}e^{r_2 s} > 0, \quad for\ all\ s \geq 0.$$

It follows that the function

$$h(s) = -e^{r_1 s} + \frac{r_1}{r_2}e^{r_2 s}, \quad s \in [0, T]$$

is always positive, satisfies $h'(0) = 0$ and satisfies the ODE. Therefore we have that $g(s) = -\frac{2}{\delta^2}\frac{h'(s)}{h(s)}$ satisfies equation (81) and $g(0) = 0$. For f, noting equation (82) and condition $f(0) = 0$, we have

$$f(s) = (b + \frac{1}{2}\delta^2)\int_0^s g(\tau)\,d\tau = -\frac{2}{\delta^2}(b + \frac{1}{2}\delta^2)\Big(\log(h(s)) - \log(h(0))\Big), \quad s \in [0, T]. \tag{84}$$

It follows that g and f are solutions of equations (81) and (82) respectively, and $g(0) = f(0) = 0$. Therefore $\overline{F}(s,z) = e^{f(s)+g(s)e^{\nu z}}$ is a solution to equation (78), and noting our transformation, $F(t,z) = e^{f(T-t)+g(T-t)e^{\nu z}}$ is a solution of (71) for this model. We found then an explicit solution to the HJB equations. We synthesize these results in the following Proposition.

Proposition 4.1.1. *Let $\nu \in \mathbb{R}\backslash\{0\}$ and $\kappa, \theta, \delta > 0$ be constants such that $2\kappa\theta > \delta^2$. Suppose further that*

$$\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)^2 - \frac{\delta^2\gamma}{\eta(1-\gamma)}\overline{\lambda}^2 > 0, \tag{85}$$

where $\gamma \in]0,1[$ *and* $\eta = \frac{1-\gamma}{1-\gamma+\gamma\rho^2}$.

Then the linear differential equation

$$\frac{\partial F}{\partial t} + \left(\frac{1}{\nu}be^{-\nu z} - \frac{\kappa}{\nu}\right)\frac{\partial F}{\partial z} + \frac{1}{2}\frac{\delta^2}{\nu^2}e^{-\nu z}\frac{\partial^2 F}{\partial z^2} + \frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2 e^{\nu z}F + \frac{\gamma\rho}{1-\gamma}\frac{\delta\overline{\lambda}}{\nu}\frac{\partial F}{\partial z} = 0 \tag{86}$$

on $[0,T]\times\mathbb{R}$ *with* $F(T,z)=1$, *admits a solution of the form*

$$F(t,z) = e^{f(T-t)+g(T-t)e^{\nu z}},$$

where $f,g \in C^1([0,T])$ *are functions satisfying*

$$-g'(s) - \left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)g(s) + \frac{1}{2}\delta^2 g(s)^2 + \frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2 = 0, \quad \text{on } [0,T], \tag{87}$$

$$-f'(s) + bg(s) + \frac{1}{2}\delta^2 g(s) = 0, \quad \text{on } [0,T], \tag{88}$$

and $f(0)=g(0)=0$, *given explicitly by*

$$g(s) = -r_1\frac{2}{\delta^2}\frac{e^{(r_2-r_1)s}-1}{\frac{r_1}{r_2}e^{(r_2-r_1)s}-1}, \qquad f(s) = -\frac{2}{\delta^2}\kappa\theta\log\left(\frac{r_2e^{r_1 s}-r_1e^{r_2 s}}{r_2-r_1}\right),$$

where r_1, r_2, *are the constants*

$$r_1, r_2 = -\frac{1}{2}\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right) \pm \frac{1}{2}\sqrt{\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)^2 - \frac{\delta^2\gamma}{\eta(1-\gamma)}\overline{\lambda}^2}$$

(we're considering $r_1 < r_2$*).*

4.2 Proving optimality

Now that we have a solution to the HJB equation, we would like to check that this is in fact the value function, using the Verification theorem.

Lets start by noting that

$$\frac{\partial F}{\partial z}\frac{1}{F(t,z)} = \nu g(T-t)e^{\nu z}.$$

The function α in this model, noting (72), is given by

$$\alpha(t,z) = e^{\left(\frac{\nu}{2}-1\right)z}\left(\frac{\overline{\lambda}}{1-\gamma} + \frac{\rho\eta}{1-\gamma}\delta g(T-t)\right).$$

For ease of notation lets define

$$\overline{g}(s) = \frac{\overline{\lambda}}{1-\gamma} + \frac{\rho\eta}{1-\gamma}\delta g(s), \quad \forall s \in [0,T],$$

so that $\alpha(t,z) = e^{\left(\frac{\nu}{2}-1\right)z}\overline{g}(T-t)$. Given all these observations, the solution to equation (73), $\hat{X}_s^{t,x,z}$, is

$$\begin{aligned}\hat{X}_s^{t,x,z} = x\exp\Big\{ & r(s-t) + \overline{\lambda}\int_t^s e^{\nu Z_u^{t,z}}\overline{g}(T-u)\,du \\ & - \frac{1}{2}\int_t^s e^{\nu Z_u^{t,z}}\overline{g}(T-u)^2\,du + \int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}\overline{g}(T-u)\,dW_u^1\Big\}.\end{aligned} \tag{89}$$

Define w as in the Verification theorem. Take now a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ in $[t,T]$. Our objective is to prove that the collection

$$\left\{w(\tau_n, \hat{X}^{t,x,z}_{\tau_n}, Z^{t,z}_{\tau_n})\right\}_{n\in\mathbb{N}}$$

is uniformly integrable. We shall do this by proving that there exists $C>0$ and $q>1$ such that

$$\mathbb{E}\left[w(\tau_n, \hat{X}^{t,x,z}_{\tau_n}, Z^{t,z}_{\tau_n})^q\right] \leq C.$$

Firstly, we shall define

$$\overline{W}_s = \sqrt{1-\rho^2}W^1_s - \rho W^2_s.$$

Noting the definition of $W^*_s = \rho W^1_s + \sqrt{1-\rho^2}W^2_s$, we see that $\overline{W}_s$ is independent to W^*_s (their covariance is always zero). We may in fact write W^1_s and W^2_s in terms of these new Brownian motions

$$W^1_s = \rho W^*_s + \sqrt{1-\rho^2}\,\overline{W}_s, \quad \textit{and} \quad W^2_s = \sqrt{1-\rho^2}\,W^*_s - \rho\overline{W}_s.$$

Writing $w(s, \hat{X}^{t,x,z}_s, Z^{t,z}_s)^q$ explicitly we have

$$\begin{aligned} w(s, \hat{X}^{t,x,z}_s, Z^{t,z}_s)^q = \frac{x^{\gamma q}}{\gamma^q}\exp\Big\{&q\gamma r(T-t) + q\gamma\overline{\lambda}\int_t^s e^{\nu Z^{t,z}_u}\overline{g}(T-u)\,du \\ &-\frac{1}{2}\gamma q\int_t^s e^{\nu Z^{t,z}_u}\overline{g}(T-u)^2\,du + \gamma q\int_t^s e^{\frac{\nu}{2}Z^{t,z}_u}\overline{g}(T-u)\,dW^1_u \\ &+ q\eta f(T-s) + q\eta g(T-s)e^{\nu Z^{t,z}_s}\Big\}. \end{aligned} \tag{90}$$

Now note that

$$\int_t^s e^{\frac{\nu}{2}Z^{t,z}_u}\overline{g}(T-u)\,dW^1_u = \rho\int_t^s e^{\frac{\nu}{2}Z^{t,z}_u}\overline{g}(T-u)\,dW^*_u + \sqrt{1-\rho^2}\int_t^s e^{\frac{\nu}{2}Z^{t,z}_u}\overline{g}(T-u)\,d\overline{W}_u$$

and

$$\begin{aligned} e^{\nu Z^{t,z}_s}\overline{g}(T-s) = e^{\nu z}\overline{g}(T-t) &- \int_t^s e^{\nu Z^{t,z}_u}\overline{g}\,'(T-u)\,du + \kappa\int_t^s(\theta - e^{\nu Z^{t,z}_u})\overline{g}(T-u)\,du \\ &+ \delta\int_t^s e^{\frac{\nu}{2}Z^{t,z}_u}\overline{g}(T-u)\,dW^*_u, \end{aligned} \tag{91}$$

which implies that the stochastic integral above is

$$\begin{aligned} \gamma q\int_t^s e^{\frac{\nu}{2}Z^{t,z}_u}\overline{g}(T-u)\,dW^1_u = \gamma q\sqrt{1-\rho^2}&\int_t^s e^{\frac{\nu}{2}Z^{t,z}_u}\overline{g}(T-u)\,d\overline{W}_u + \frac{\gamma q\rho}{\delta}e^{\nu Z^{t,z}_s}\overline{g}(T-s) \\ &- \frac{\gamma q\rho}{\delta}e^{\nu z}\overline{g}(T-t) - \frac{\gamma q\rho}{\delta}\kappa\theta\int_t^s \overline{g}(T-u)\,du \\ &+ \frac{\gamma q\rho}{\delta}\int_t^s e^{\nu Z^{t,z}_u}\left(\overline{g}'(T-u) + \kappa\overline{g}(T-u)\right)\,du. \end{aligned} \tag{92}$$

Therefore we have the inequality

$$\begin{aligned} w(s, \hat{X}^{t,x,z}_s, Z^{t,z}_s)^q \leq M_1\exp\Big\{&\int_t^s e^{\nu Z^{t,z}_u}\gamma q\Big(\frac{\rho}{\delta}\overline{g}\,'(T-u) + \Big(\frac{\rho}{\delta}\kappa + \overline{\lambda}\Big)\overline{g}(T-u) - \frac{1}{2}\overline{g}(T-u)^2\Big)\,du \\ &+ q\left(\eta g(T-s) + \frac{\gamma\rho}{\delta}\overline{g}(T-s)\right)e^{\nu Z^{t,z}_s} \\ &+ \gamma q\sqrt{1-\rho^2}\int_t^s e^{\frac{\nu}{2}Z^{t,z}_u}\overline{g}(T-u)\,d\overline{W}_u\Big\} \end{aligned} \tag{93}$$

where $M_1>0$ is some constant that bounds the deterministic functions.

Now we prove the following Lemma

Lemma 4.2.1. *The function $\overline{g}(s)$ satisfies*

$$-\overline{g}'(s) - \left(\kappa + \delta\frac{\overline{\lambda}}{\rho}\right)\overline{g}(s) + \frac{1}{2}\delta\frac{1-\gamma}{\rho\eta}\overline{g}(s)^2 + \kappa\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}\delta\frac{\overline{\lambda}^2}{\rho(1-\gamma)} = 0. \tag{94}$$

on $[0,T]$.

Also we have the inequality

$$\eta g(T-s) + \frac{\gamma\rho}{\delta}\overline{g}(T-s) < \frac{\kappa}{\delta^2}. \tag{95}$$

Proof. We know that g satisfies the equation

$$-g'(s) - \left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)g(s) + \frac{1}{2}\delta^2 g(s)^2 + \frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2 = 0,$$

and $\overline{g}(s) = \frac{\overline{\lambda}}{1-\gamma} + \frac{\rho\eta}{1-\gamma}\delta g(s)$, so

$$g(s) = \frac{1-\gamma}{\rho\eta\delta}\left(\overline{g}(s) - \frac{\overline{\lambda}}{1-\gamma}\right)$$

and we may write the above equation in terms of $\overline{g}$

$$\begin{aligned}-\frac{1-\gamma}{\rho\eta\delta}\,\overline{g}\,'(s) - \frac{1-\gamma}{\rho\eta\delta}\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)\overline{g}(s) + \frac{1-\gamma}{\rho\eta\delta}\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)\frac{\overline{\lambda}}{1-\gamma} \\ + \frac{1}{2}\delta^2\left(\frac{1-\gamma}{\rho\eta\delta}\right)^2\left(\overline{g}(s) - \frac{\overline{\lambda}}{1-\gamma}\right)^2 + \frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\overline{\lambda}^2 = 0.\end{aligned}$$

Multiplying by $\frac{\rho\eta\delta}{1-\gamma}$ and simplifying

$$\begin{aligned}-\overline{g}\,'(s) + \left(-\kappa + \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda} - \frac{\delta\overline{\lambda}}{\rho\eta}\right)g(s) + \frac{1}{2}\delta\frac{1-\gamma}{\rho\eta}\overline{g}(s)^2 \\ + \left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}\delta\frac{\overline{\lambda}^2}{(1-\gamma)\rho\eta} + \frac{1}{2}\frac{\gamma}{(1-\gamma)^2}\overline{\lambda}^2\rho\delta = 0.\end{aligned} \tag{96}$$

Now using $\eta = \frac{1-\gamma}{1-\gamma+\gamma\rho^2}$, note that

$$-\kappa + \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda} - \frac{\delta\overline{\lambda}}{\rho\eta} = -\kappa - \delta\frac{\overline{\lambda}}{\rho},$$

and

$$\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}\delta\frac{\overline{\lambda}^2}{(1-\gamma)\rho\eta} + \frac{1}{2}\frac{\gamma}{(1-\gamma)^2}\overline{\lambda}^2\rho\delta = \kappa\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}\delta\frac{\overline{\lambda}^2}{\rho(1-\gamma)}$$

so we have that $\overline{g}$ satisfies

$$-\overline{g}'(s) - \left(\kappa + \delta\frac{\overline{\lambda}}{\rho}\right)\overline{g}(s) + \frac{1}{2}\delta\frac{1-\gamma}{\rho\eta}\overline{g}(s)^2 + \kappa\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}\delta\frac{\overline{\lambda}^2}{\rho(1-\gamma)} = 0. \tag{97}$$

Now lets prove the inequality. Firstly, noting the definition of g as given in the Proposition of the last section, we have for $s > 0$

$$\begin{aligned}g(s) &= -r_1\frac{2}{\delta^2}\frac{e^{(r_2-r_1)s}-1}{\frac{r_1}{r_2}e^{(r_2-r_1)s}-1} = -\frac{2}{\delta^2}r_1r_2\frac{e^{(r_2-r_1)s}-1}{r_1e^{(r_2-r_1)s}-r_2} = -\frac{2}{\delta^2}r_1r_2\frac{e^{(r_2-r_1)s}-1}{r_1e^{(r_2-r_1)s}-r_1+r_1-r_2} \\ &= \frac{2}{\delta^2}r_1r_2\frac{1}{-r_1+\frac{r_2-r_1}{e^{(r_2-r_1)s}-1}} \leq -\frac{2}{\delta^2}r_2 \leq \frac{1}{\delta^2}\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right),\end{aligned} \tag{98}$$

and since $g(0) = 0$ this inequality holds for all $s \geq 0$. Now note

$$\begin{aligned}\eta g(T-s) + \frac{\gamma\rho}{\delta}\overline{g}(T-s) &= \eta g(T-s) + \frac{\gamma\rho}{\delta}\frac{\overline{\lambda}}{1-\gamma} + \gamma\rho\frac{\rho\eta}{1-\gamma}g(T-s) \\ &= \frac{\overline{\lambda}\gamma\rho}{\delta(1-\gamma)} + \eta\left(1 + \frac{\gamma\rho^2}{1-\gamma}\right)g(T-s) = \frac{\overline{\lambda}\gamma\rho}{\delta(1-\gamma)} + g(T-s),\end{aligned} \tag{99}$$

and using our previous inequality

$$\eta g(T-s) + \frac{\gamma\rho}{\delta}\overline{g}(T-s) \leq \frac{\overline{\lambda}\gamma\rho}{\delta(1-\gamma)} + \frac{1}{\delta^2}\left(\kappa - \frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right) = \frac{\kappa}{\delta^2}. \tag{100}$$

□

Using the equation for $\overline{g}$ we just found, we can note that

$$\frac{\rho}{\delta}\overline{g}'(T-u) + \left(\frac{\rho}{\delta}\kappa + \overline{\lambda}\right)\overline{g}(T-u) = \frac{1}{2}\frac{1-\gamma}{\eta}\overline{g}(T-u)^2 + \kappa\frac{\rho}{\delta}\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}\frac{\overline{\lambda}^2}{1-\gamma}.$$

Using this and the inequality proved in the Lemma, we can further inequality (93) to be

$$\begin{aligned}w(s, \hat{X}_s^{t,x,z}, Z_s^{t,z})^q \leq M_1 \exp\bigg\{ &\int_t^s e^{\nu Z_u^{t,z}}\gamma q\bigg(\frac{1}{2}\left(\frac{1-\gamma}{\eta} - 1\right)\overline{g}(T-u)^2 + \kappa\frac{\rho}{\delta}\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}\frac{\overline{\lambda}^2}{1-\gamma}\bigg)\,du \\ &+ q\frac{\kappa}{\delta^2}e^{\nu Z_s^{t,z}} + \gamma q\sqrt{1-\rho^2}\int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}\overline{g}(T-u)\,d\overline{W}_u\bigg\}.\end{aligned} \tag{101}$$

Now note

$$e^{\nu Z_s^{t,z}} = e^{\nu z} + \kappa\int_t^s (\theta - e^{\nu Z_u^{t,z}})\,du + \delta\int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}\,dW_u^*,$$

so we can further the above inequality once again to be

$$\begin{aligned}w(s, \hat{X}_s^{t,x,z}, Z_s^{t,z})^q \leq M_2 \exp\bigg\{ &\int_t^s e^{\nu Z_u^{t,z}}\gamma q\bigg(\frac{1}{2}\left(\frac{1-\gamma}{\eta} - 1\right)\overline{g}(T-u)^2 + \kappa\frac{\rho}{\delta}\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}\frac{\overline{\lambda}^2}{1-\gamma}\bigg)\,du \\ &- q\frac{\kappa^2}{\delta^2}\int_t^s e^{\nu Z_u^{t,z}}\,du + q\frac{\kappa}{\delta}\int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}\,dW_u^* \\ &+ \gamma q\sqrt{1-\rho^2}\int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}\overline{g}(T-u)\,d\overline{W}_u\bigg\},\end{aligned} \tag{102}$$

where M_2 is a new constant that bounds the deterministic functions. Now we shall add and subtract the terms

$$\frac{1}{2}\int_t^s e^{\nu Z_u^{t,z}}\left(q^2\frac{\kappa^2}{\delta^2} + \gamma^2 q^2(1-\rho^2)\overline{g}(T-u)^2\right)du,$$

to obtain

$$\begin{aligned}w(s, \hat{X}_s^{t,x,z}, Z_s^{t,z})^q \leq M_2 \exp\bigg\{ &\int_t^s e^{\nu Z_u^{t,z}}\bigg[\left(q\gamma\frac{1}{2}\left(\frac{1-\gamma}{\eta} - 1\right) + \frac{1}{2}\gamma^2 q^2(1-\rho^2)\right)\overline{g}(T-u)^2 \\ &+ q\gamma\kappa\frac{\rho}{\delta}\frac{\overline{\lambda}}{1-\gamma} + \frac{1}{2}q\frac{\gamma\overline{\lambda}^2}{1-\gamma} - q\frac{\kappa^2}{\delta^2} + \frac{1}{2}q^2\frac{\kappa^2}{\delta^2}\bigg]\,du\bigg\} \\ \exp\bigg\{ &-\frac{1}{2}\int_t^s e^{\nu Z_u^{t,z}}\left(q^2\frac{\kappa^2}{\delta^2} + \gamma^2 q^2(1-\rho^2)\overline{g}(T-u)^2\right)du \\ &+ q\frac{\kappa}{\delta}\int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}\,dW_u^* + \gamma q\sqrt{1-\rho^2}\int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}\overline{g}(T-u)\,d\overline{W}_u\bigg\}.\end{aligned} \tag{103}$$

Now note, using the definition of η and simplifying we have

$$q\gamma\frac{1}{2}\left(\frac{1-\gamma}{\eta}-1\right)+\frac{1}{2}\gamma^2q^2(1-\rho^2)=\frac{1}{2}q\gamma^2(1-\rho^2)(q-1)$$

On the other hand, noting inequality (83), we have

$$\gamma\kappa\frac{\rho}{\delta}\frac{\overline{\lambda}}{1-\gamma}+\frac{1}{2}\frac{\gamma\overline{\lambda}^2}{1-\gamma}<\frac{\kappa^2}{2\delta^2},$$

where this inequality is strict so we can take $\epsilon>0$ such that

$$\gamma\kappa\frac{\rho}{\delta}\frac{\overline{\lambda}}{1-\gamma}+\frac{1}{2}\frac{\gamma\overline{\lambda}^2}{1-\gamma}=\frac{\kappa^2}{2\delta^2}-\epsilon,$$

where this ϵ does not depend on q. It follows that

$$q\gamma\kappa\frac{\rho}{\delta}\frac{\overline{\lambda}}{1-\gamma}+\frac{1}{2}q\frac{\gamma\overline{\lambda}^2}{1-\gamma}-q\frac{\kappa^2}{\delta^2}+\frac{1}{2}q^2\frac{\kappa^2}{\delta^2}=-\epsilon q-q\frac{1}{2}\frac{\kappa^2}{\delta^2}+\frac{1}{2}q^2\frac{\kappa^2}{\delta^2}=\frac{1}{2}\frac{\kappa^2}{\delta^2}q(q-1)-\epsilon q \tag{104}$$

so the first exponential of the above inequality (103) simplifies to

$$\exp\left\{\int_t^s e^{\nu Z_u^{t,z}}\left[\frac{1}{2}q\gamma^2(1-\rho^2)(q-1)\overline{g}(T-u)^2+\frac{1}{2}\frac{\kappa^2}{\delta^2}q(q-1)-\epsilon q\right]du\right\}.$$

Since $\overline{g}$ is bounded, the above expression will approach $-\epsilon$, as $q\to 1$. It follows that we can choose $q>1$ such that the above expression is negative, and so this exponential is smaller then 1 for this choice. Therefore, for this choice of q

$$\begin{aligned} w(s,\hat{X}_s^{t,x,z},Z_s^{t,z})^q\leq M_2\exp\bigg\{&-\frac{1}{2}\int_t^s e^{\nu Z_u^{t,z}}\left(q^2\frac{\kappa^2}{\delta^2}+\gamma^2q^2(1-\rho^2)\overline{g}(T-u)^2\right)du\\ &+q\frac{\kappa}{\delta}\int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}dW_u^*+\gamma q\sqrt{1-\rho^2}\int_t^s e^{\frac{\nu}{2}Z_u^{t,z}}\overline{g}(T-u)\,d\overline{W}_u\bigg\}. \end{aligned} \tag{105}$$

This process that bounds $w(s,\hat{X}_s^{t,x,z},Z_s^{t,z})^q$ is a non-negative local martingale, with starting value M_2, so it's a supermartingale. Denoting by D_s the above supermartingale process that bounds $w(s,\hat{X}_s^{t,x,z},Z_s^{t,z})^q$, we have that for any sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times in $[t,T]$

$$\mathbb{E}\left[w(\tau_n,\hat{X}_{\tau_n}^{t,x,z},Z_{\tau_n}^{t,z})^q\right]\leq\mathbb{E}\Big[D_{\tau_n}\Big]\leq M_2 \tag{106}$$

since D_s is a supermartingale and $D_t=M_2$. Therefore the collection

$$\left\{w(\tau_n,\hat{X}_{\tau_n}^{t,x,z},Z_{\tau_n}^{t,z})^q\right\}_{n\in\mathbb{N}}$$

is uniformly integrable, and so we can use the Verification theorem to prove that the solution we found to the HJB equation is indeed the value function. We synthesize the results proved in this section in the following Theorem.

Theorem 4.2.2. *Let (A,B,μ) be a stochastic volatility model such that*

$$B(z)=e^{-\nu z}\frac{1}{2\nu}(2\kappa\theta-\delta^2)-\frac{\kappa}{\nu},\quad A(z)=\frac{\delta}{\nu}e^{-\frac{\nu}{2}z},\quad \mu(z)=r+\overline{\lambda}e^{(\frac{\nu}{2}+1)z},\quad z\in\mathbb{R},$$

where $\nu\in\mathbb{R}\backslash\{0\}$, and $\kappa,\theta,\delta,\overline{\lambda}\in\mathbb{R}^+$ are constants such that $2\kappa\theta>\delta^2$. Suppose further that

$$\left(\kappa-\frac{\gamma\rho}{1-\gamma}\delta\overline{\lambda}\right)^2-\frac{\delta^2\gamma}{\eta(1-\gamma)}\overline{\lambda}^2>0, \tag{107}$$

where $\gamma \in]0,1[$ *is some constant and* $\eta = \frac{1-\gamma}{1-\gamma+\gamma\rho^2}$. *Then the value function for the problem of the above stochastic volatility model and utility* $U(x) = x^\gamma/\gamma$, $x \in \mathbb{R}^+$, *is given by*

$$v(t,x,z) = \frac{x^\gamma}{\gamma} e^{\gamma r(T-t)} F(t,z)^\eta$$

where F *is the function presented in Proposition* 4.1.1.

Also the function α *presented in the Verification theorem for power utility functions* 3.3.1 *is, for this model, given by*

$$\alpha(t,z) = e^{\left(\frac{\nu}{2}-1\right)z}\left(\frac{\overline{\lambda}}{1-\gamma} + \frac{\rho\eta}{1-\gamma}\delta g(T-t)\right),$$

and for each starting conditions $(t,x,z) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R}$ *the optimal control is* $\alpha(s, Z_s^{t,z})$, *i.e.*

$$v(t,x,z) = P(t,x,z,\alpha(s,Z_s^{t,z})).$$

5 A general approach to solving the power utility case

5.1 Simplification of the equation

For the linear differential equation found for the power utility function, one can find a candidate of the solution using a Feynman-Kac representation. Noting the equation

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial z}B(z) + \frac{1}{2}\frac{\partial^2 F}{\partial z^2}A(z)^2 + \frac{1}{2}F\frac{\gamma}{\eta(1-\gamma)}\lambda(z)^2 + \frac{\partial F}{\partial z}\frac{\gamma\rho A(z)}{(1-\gamma)}\lambda(z) = 0, \tag{108}$$

with the final condition $F(T,z) = 1$, where we're denoting $\lambda(z) = \frac{\mu(z)-r}{e^z}$, to find the corresponding Feynman-Kac representation of this equation we need a solution to the stochastic differential equation

$$dY_s = \left(B(Y_s) + \frac{\gamma\rho}{1-\gamma}\lambda(Y_s)A(Y_s)\right) ds + A(Y_s)\, dW_s^*, \quad Y_t = z,$$

for any $(t,z) \in [0,T] \times \mathbb{R}$. Denoting by $Y_s^{t,z}$ a solution to the above stochastic differential equation with starting conditions (t,z) (if this solution exists of course) the Feynman-Kac representation is given by

$$\mathbb{E}\left[\exp\left\{\frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\int_t^T \lambda(Y_s^{t,z})^2\, ds\right\}\right]. \tag{109}$$

This representation as a function of $(t,z) \in [0,T] \times \mathbb{R}$ is then candidate to be a solution of the above linear equation. For the full proof of this we would need the backwards implication of the Feynman-Kac formula theorem, which we do not know to be true for our conditions. We would like to at least be able to know that the representation we find is finite.

In the model discussed in section 4 proving that the above expectation is finite using this approach is actually possible, and is in fact how Kraft [13] approaches the problem. This is because, this process we need to use for the Feynman-Kac formula remains of the same type, since in this case

$$B(z) + \frac{\gamma\rho}{1-\gamma}\lambda(z)A(z) = e^{-\nu z}\frac{1}{2\nu}(2\kappa\theta - \delta^2) - \frac{\kappa}{\nu} + \frac{\gamma\rho}{1-\gamma}\overline{\lambda}\frac{\delta}{\nu}$$

and since under condition (85) the above constant term has the sign of $-\frac{1}{\nu}$, we have that the corresponding process $e^{\nu Y_s}$ is still CIR process. The above expectation (109) can be solved explicitly in this case and the solution is the same as the one we have obtained.

Yet in some other cases this might not be so nice, in fact remaining in the same example, if we consider other functions for $\lambda(z)$ we may not be sure that equation (108) has a solution, or maybe for the solution it may not be easy to prove that the Feynman-Kac representation is finite.

It might be useful then if we could do some transformation on the equation, such that the infinitesimal generator became the one of $Z_s^{t,z}$, the process of the stochastic volatility model. We would then obtain a different Feynman-Kac representation, but one that uses a process that should be easier to work with. This can indeed be done by considering a transformation of the form

$$\overline{F}(t,z) = e^{-g(z)}F(t,z), \quad on\ [0,T] \times \mathbb{R},$$

where F is a function satisfying equation (71) and $g \in C^2(\mathbb{R})$ is a function to be determined. For this transformation we have the equalities

$$\frac{\partial F}{\partial t} = e^{g(z)}\frac{\partial \overline{F}}{\partial t}, \quad \frac{\partial F}{\partial z} = g'(z)e^{g(z)}\overline{F} + e^{g(z)}\frac{\partial \overline{F}}{\partial z},$$

and
$$\frac{\partial^2 F}{\partial z^2} = (g''(z) + (g'(z))^2)e^{g(z)}\overline{F} + 2g'(z)e^{g(z)}\frac{\partial \overline{F}}{\partial z} + e^{g(z)}\frac{\partial^2 \overline{F}}{\partial z^2}.$$
To simplify notation we define $\lambda : \mathbb{R} \to \mathbb{R}$ as
$$\lambda(z) = \frac{\mu(z) - r}{e^z}.$$
We have then the equation for $\overline{F}$,
$$\begin{aligned}&\frac{\partial \overline{F}}{\partial t} + g'(z)B(z)\overline{F} + B(z)\frac{\partial \overline{F}}{\partial z} + \frac{1}{2}A(z)^2(g''(z) + (g'(z))^2)\overline{F} + A(z)^2 g'(z)\frac{\partial \overline{F}}{\partial z} \\ &+ \frac{1}{2}A(z)^2\frac{\partial^2 \overline{F}}{\partial z^2} + \frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\lambda(z)^2\overline{F} + g'(z)\frac{\gamma\rho}{1-\gamma}A(z)\lambda(z)\overline{F} + \frac{\gamma\rho}{1-\gamma}A(z)\lambda(z)\frac{\partial \overline{F}}{\partial z} = 0.\end{aligned} \tag{110}$$
with final condition $\overline{F}(T, z) = e^{-g(z)}$. This suggest we take g such that
$$A(z)^2 g'(z) + \frac{\gamma\rho}{1-\gamma}A(z)\lambda(z) = 0.$$
For instance we may take
$$g(z) = -\frac{\gamma\rho}{1-\gamma}\int_0^z \frac{\lambda(y)}{A(y)}\,dy.$$
Since we have supposed $A(z) \neq 0$ for all $z \in \mathbb{R}$ this is always possible. We also want this function to be in $C^2(\mathbb{R})$ so we suppose that $\lambda(z)$ and $A(z)$ are $C^1(\mathbb{R})$ functions.

Taking this g, the equation for $\overline{F}$ simplifies to
$$\frac{\partial \overline{F}}{\partial t} + B(z)\frac{\partial \overline{F}}{\partial z} + \frac{1}{2}A(z)^2\frac{\partial^2 \overline{F}}{\partial z^2} + h(z)\overline{F} = 0,$$
where h is the function defined by
$$\begin{aligned}h(z) =& - B(z)\frac{\gamma\rho}{1-\gamma}\frac{\lambda(z)}{A(z)} + \frac{1}{2}A(z)^2\left(-\frac{\gamma\rho}{1-\gamma}\frac{\lambda'(z)}{A(z)} + \frac{\gamma\rho}{1-\gamma}\frac{\lambda(z)}{A(z)^2}A'(z) + \frac{\gamma^2\rho^2}{(1-\gamma)^2}\frac{\lambda(z)^2}{A(z)^2}\right) \\ &+ \frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\lambda(z)^2 - \frac{\gamma^2\rho^2}{(1-\gamma)^2}\lambda(z)^2.\end{aligned} \tag{111}$$

This function can be expressed in a simpler form. Firstly we remind that $\eta = \frac{1-\gamma}{1-\gamma+\gamma\rho^2}$, so the term above with this constant is
$$\frac{1}{2}\frac{\gamma}{\eta(1-\gamma)}\lambda(z)^2 = \frac{1}{2}\frac{\gamma}{1-\gamma}\lambda(z)^2 + \frac{1}{2}\frac{\gamma^2\rho^2}{(1-\gamma)^2}\lambda(z)^2,$$
using this and simplifying, we obtain
$$\begin{aligned}h(z) =& - B(z)\frac{\gamma\rho}{1-\gamma}\frac{\lambda(z)}{A(z)} - \frac{1}{2}\frac{\gamma\rho}{1-\gamma}A(z)\lambda'(z) + \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\lambda(z)A'(z) + \frac{1}{2}\frac{\gamma^2\rho^2}{(1-\gamma)^2}\lambda(z)^2 \\ &+ \frac{1}{2}\frac{\gamma}{1-\gamma}\lambda(z)^2 + \frac{1}{2}\frac{\gamma^2\rho^2}{(1-\gamma)^2}\lambda(z)^2 - \frac{\gamma^2\rho^2}{(1-\gamma)^2}\lambda(z)^2,\end{aligned} \tag{112}$$
or simply
$$h(z) = -B(z)\frac{\gamma\rho}{1-\gamma}\frac{\lambda(z)}{A(z)} - \frac{1}{2}\frac{\gamma\rho}{1-\gamma}A(z)\lambda'(z) + \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\lambda(z)A'(z) + \frac{1}{2}\frac{\gamma}{1-\gamma}\lambda(z)^2. \tag{113}$$
We synthesize this discussion in the following Proposition.

Proposition 5.1.1. *Let (A, B, μ) be a stochastic volatility model such that $A, \mu \in C^1(\mathbb{R})$. Define $\lambda : \mathbb{R} \to \mathbb{R}$ to be the function*

$$\lambda(z) = \frac{\mu(z) - r}{e^z}.$$

If $\overline{F} \in C^{1,2}([0,T] \times \mathbb{R})$ is a function satisfying

$$\frac{\partial \overline{F}}{\partial t} + B(z)\frac{\partial \overline{F}}{\partial z} + \frac{1}{2}A(z)^2\frac{\partial^2 \overline{F}}{\partial z^2} + \overline{F}\left[-B(z)\frac{\gamma\rho}{1-\gamma}\frac{\lambda(z)}{A(z)} - \frac{1}{2}\frac{\gamma\rho}{1-\gamma}A(z)\lambda'(z) + \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\lambda(z)A'(z) + \frac{1}{2}\frac{\gamma}{1-\gamma}\lambda(z)^2\right] = 0 \tag{114}$$

on $[0,T] \times \mathbb{R}$ and $\overline{F}(T,z) = \exp\left\{\frac{\gamma\rho}{1-\gamma}\int_0^z \frac{\lambda(y)}{A(y)}\,dy\right\}$ on $\mathbb{R}$, we have that the function defined by

$$F(t,z) = \exp\left\{-\frac{\gamma\rho}{1-\gamma}\int_0^z \frac{\lambda(y)}{A(y)}\,dy\right\}\overline{F}(t,z), \quad \text{on } [0,T] \times \mathbb{R}$$

satisfies

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial z}B(z) + \frac{1}{2}\frac{\partial^2 F}{\partial z^2}A(z)^2 + \frac{1}{2}F\frac{\gamma}{\eta(1-\gamma)}\lambda(z)^2 + \frac{\partial F}{\partial z}\frac{\gamma A(z)\rho}{(1-\gamma)}\lambda(z) = 0, \tag{115}$$

on $[0,T] \times \mathbb{R}$ and $F(T,z) = 1$ on $\mathbb{R}$.

Following this transformation we have the Feynman-Kac representation for $\overline{F}$

$$\mathbb{E}\Bigg[\exp\Bigg\{\frac{\gamma\rho}{1-\gamma}\int_0^{Z_T^{t,z}}\frac{\lambda(y)}{A(y)}\,dy - \frac{\gamma\rho}{1-\gamma}\int_t^T B(Z_s^{t,z})\frac{\lambda(Z_s^{t,z})}{A(Z_s^{t,z})}\,ds - \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\int_t^T A(Z_s^{t,z})\lambda'(Z_s^{t,z})\,ds + \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\int_t^T \lambda(Z_s^{t,z})A'(Z_s^{t,z})\,ds + \frac{1}{2}\frac{\gamma}{1-\gamma}\int_t^T \lambda(Z_s^{t,z})^2\,ds\Bigg\}\Bigg], \tag{116}$$

and for F, noting the transformation we have made, we have the representation

$$\mathbb{E}\Bigg[\exp\Bigg\{\frac{\gamma\rho}{1-\gamma}\int_z^{Z_T^{t,z}}\frac{\lambda(y)}{A(y)}\,dy - \frac{\gamma\rho}{1-\gamma}\int_t^T B(Z_s^{t,z})\frac{\lambda(Z_s^{t,z})}{A(Z_s^{t,z})}\,ds - \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\int_t^T A(Z_s^{t,z})\lambda'(Z_s^{t,z})\,ds + \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\int_t^T \lambda(Z_s^{t,z})A'(Z_s^{t,z})\,ds + \frac{1}{2}\frac{\gamma}{1-\gamma}\int_t^T \lambda(Z_s^{t,z})^2\,ds\Bigg\}\Bigg], \tag{117}$$

5.2 2-hypergeometric model

The $\alpha-$*hypergeometric* model is a recent model developed in [4] which "has been designed to make up for the numerous flaws observed when implementing the Heston model (or any other affine model)". For instance, in the Heston model, when doing a statistical estimation of the parameters one might not necessarily obtain $2\kappa\theta > \alpha^2$ which is a required property for the strict positivity of the process.

This model is given by the equation

$$dZ_s = (a - be^{\alpha Z_s})\,ds + \delta\,dW_s^*, \quad Z_t = z, \tag{118}$$

where $t \in [0,T]$, $z \in \mathbb{R}$, $b, \delta, \alpha > 0$ and $a \in \mathbb{R}$. The volatility of the stock is then given by e^{Z_s}. In our formulation this will correspond to $A(z) = \delta$ and $B(z) = a - be^{\alpha z}$.

Here we shall consider the $2-hypergeometric$ model, i.e., the case $\alpha = 2$. One can question whether this equation has an unique solution. We follow the approach found in [4] to prove this.

Defining $H(s) = \int_t^s e^{2Z_u}\,du$ we have

$$\frac{\partial H}{\partial s} = e^{2Z_s}, \quad \frac{1}{2}\log\left(\frac{\partial H}{\partial s}\right) = Z_s \tag{119}$$

Therefore writing the equation in terms of $H(s)$ we get

$$\frac{1}{2}\log\left(\frac{\partial H}{\partial s}\right) = z + a(s-t) - bH(s) + \delta(W_s^* - W_t^*)$$

so multiplying everything by 2 and applying an exponential we have

$$\frac{\partial H}{\partial s}e^{2bH(s)} = e^{2z+2a(s-t)+2\delta(W_s^*-W_t^*)},$$

therefore applying the integral between t and s

$$\frac{1}{2b}(e^{2bH(s)} - 1) = e^{2z}\int_t^s e^{2a(r-t)+2\delta(W_r^*-W_t^*)}\,dr,$$

finally we obtain

$$2bH(s) = \log\left(1 + 2be^{2z}\int_t^s e^{2a(r-t)+2\delta(W_r^*-W_t^*)}\,dr\right).$$

This way we may obtain Z_s using relations (119)

$$Z_s = \frac{1}{2}\log\left(\frac{\partial H}{\partial s}\right) = z + a(s-t) + \delta(W_s^* - W_t^*) - \frac{1}{2}\log\left(1 + 2be^{2z}\int_t^s e^{2a(r-t)+2\delta(W_r^*-W_t^*)}\,dr\right).$$

This is indeed a solution to (118) as one may verify, and by locally Lipschitz coefficients one concludes that it is the unique solution to (118).

Now lets prove prove that $\mathbb{E}\left[e^{\frac{C}{2}\int_t^T e^{2Z_s}\,ds}\right] < \infty$, for every $C > 0$. We note that $\int_t^T e^{2Z_s}\,ds = H(T)$ as previously defined, therefore using the previous equality

$$\int_t^T e^{2Z_s}\,ds = \log\left(1 + 2be^{2z}\int_t^T e^{2a(r-t)+2\delta(W_r^*-W_t^*)}\,dr\right)^{\frac{1}{2b}},$$

so

$$e^{\frac{C}{2}\int_t^T e^{2Z_s}\,ds} = \left(1 + 2be^{2z}\int_t^T e^{2a(r-t)+2\delta(W_r^*-W_t^*)}\,dr\right)^{\frac{C}{4b}}. \tag{120}$$

We have that

$$\sup_{r\in[t,T]} e^{2\delta|W_r^*-W_t^*|} \geq 1,$$

so taking the sup in the integral and using the above inequality we get

$$e^{\frac{C}{2}\int_t^T e^{2Z_s}\,ds} \leq \left(\sup_{r\in[t,T]} e^{2\delta|W_r^*-W_t^*|}\right)^{\frac{C}{4b}}\left(1 + 2be^{2z}\int_t^T e^{2a(r-t)}\,dr\right)^{\frac{C}{4b}}.$$

Since $e^{2\delta|W_r^*-W_t^*|}$ is always positive the sup *exchanges* with the exponent, and the other term is some positive constant, so we have the inequality

$$e^{\frac{C}{2}\int_t^T e^{2Z_s}\,ds} \leq D \sup_{r\in[t,T]} e^{\frac{C\delta}{2b}|W_r^*-W_t^*|} \tag{121}$$

for some $D > 0$. Given that $W_r^* - W_t^*$ is a martingale on $[t,T]$, and that $\mathbb{E}\left[e^{\frac{C\delta}{4b}|W_r^*-W_t^*|}\right] < \infty$ (the density of the Brownian motion has decay of type e^{-x^2}), we have by Jensen's inequality of conditional expectation that $e^{\frac{C\delta}{4b}|W_r^*-W_t^*|}$ is a submartingale (since $e^{\frac{C\delta}{4b}|x|}$ is convex). We can then use Doob's submartingale inequality on this process to obtain

$$\mathbb{E}\left[\sup_{r\in[t,T]} e^{\frac{C\delta}{2b}|W_r^*-W_t^*|}\right] \leq 4\,\mathbb{E}\left[e^{\frac{C\delta}{2b}|W_T^*-W_t^*|}\right] < \infty, \tag{122}$$

where we used Theorem 1.2.6 with $p = 2$. It then follows that $\mathbb{E}\left[e^{\frac{C}{2}\int_t^T \sigma(Z_s)^2\,ds}\right] < \infty$, for every $C > 0$.

We synthesize all our previous results into the following Proposition.

Proposition 5.2.1. *The stochastic differential equation*

$$dZ_s = (a - be^{2Z_s})\,ds + \delta\,dW_s^*, \quad Z_t = z, \tag{123}$$

with given starting time $t \in [0,T]$, starting condition $z \in \mathbb{R}$, and parameters $b > 0$, $\delta, a \in \mathbb{R}$, has an unique solution given by

$$Z_s = z + a(s-t) + \delta(W_s^* - W_t^*) - \frac{1}{2}\log\left(1 + 2be^{2z}\int_t^s e^{2a(r-t)+2\delta(W_r^*-W_t^*)}\,dr\right). \tag{124}$$

Also, Z_s satisfies

$$\mathbb{E}\left[e^{\frac{C}{2}\int_t^T e^{2Z_s}\,ds}\right] < \infty,$$

for all $C > 0$.

We would now wish to prove that the Feynman-Kac representation for the function $\overline{F}$ in this model is finite. If one assumes that $\lambda(z) = \overline{\lambda}e^{mz}$ for $\overline{\lambda} > 0$, $m \in \mathbb{R}$, looking at the equation for $\overline{F}$ there is one term of the form $\frac{1}{2}\frac{\gamma}{1-\gamma}\lambda(z)^2\overline{F}$ which in the Feynman-Kac formula will correspond to the term

$$\exp\left\{\frac{1}{2}\frac{\gamma}{1-\gamma}\overline{\lambda}^2\int_t^T e^{2mZ_s}\,ds\right\}.$$

Given the previous Proposition we know that for $m \in [0,1]$ this has finite expectation, but for other m we're unsure. For $\mu(z)$ constant corresponds $m = -1$, so we don't know if this term as finite expectation and looking at the explicit formula for Z_s the expectation seems to be infinite.

We shall then follow the approach of Kraft and Liu discussed in the last section and consider the case where the market price of risk, $\lambda(z)$, grows with volatility. We consider $m \in]0,1]$. The case $m = 0$ is trivial and given in example 3.3.1.

The Feynman-Kac representation for this model is

$$\begin{aligned}\overline{F}(t,z) = \mathbb{E}\,\exp\Bigg\{&\frac{\gamma\rho\overline{\lambda}}{m\delta(1-\gamma)}(e^{mZ_T^{t,z}} - 1) - \overline{\lambda}\frac{\gamma\rho}{1-\gamma}\left(\frac{a}{\delta} + \frac{1}{2}\delta\right)\int_t^T e^{mZ_s^{t,z}}ds \\ &+ b\overline{\lambda}\frac{\gamma\rho}{\delta(1-\gamma)}\int_t^T e^{(2+m)Z_s^{t,z}}ds + \frac{1}{2}\frac{\gamma}{1-\gamma}\overline{\lambda}^2\int_t^T e^{2mZ_s^{t,z}}ds\Bigg\},\end{aligned} \tag{125}$$

If we assume $\rho < 0$, which is usually what's observed in the market, we can see that this integral is finite. This is because

$$
\begin{aligned}
&\exp\left\{\frac{\gamma\rho\overline{\lambda}}{m\delta(1-\gamma)}(e^{mZ_T^{t,z}}-1) - \overline{\lambda}\frac{\gamma\rho}{1-\gamma}\left(\frac{a}{\delta}+\frac{1}{2}\delta\right)\int_t^T e^{mZ_s^{t,z}}ds \right.\\
&\qquad\qquad\left. + b\overline{\lambda}\frac{\gamma\rho}{\delta(1-\gamma)}\int_t^T e^{(2+m)Z_s^{t,z}}ds + \frac{1}{2}\frac{\gamma}{1-\gamma}\overline{\lambda}^2\int_t^T e^{2mZ_s^{t,z}}ds\right\}\\
\le\ &\exp\left\{-\frac{\gamma\rho\overline{\lambda}}{m\delta(1-\gamma)} - \overline{\lambda}\frac{\gamma\rho}{1-\gamma}\left(\frac{a}{\delta}+\frac{1}{2}\delta\right)\int_t^T (1+e^{2Z_s^{t,z}})ds + \frac{1}{2}\frac{\gamma}{1-\gamma}\overline{\lambda}^2\int_t^T (1+e^{2Z_s^{t,z}})ds\right\}
\end{aligned}
$$

and this last expression is integrable (we have used the property $x^\beta \le 1 + x^2$ for all $\beta \in [0, 2]$). If $\overline{F}$ is given this way we can write F in the form of the expected value

$$
\begin{aligned}
F(t,z) = \mathbb{E}\ \exp&\left\{\frac{\gamma\rho\overline{\lambda}}{m\delta(1-\gamma)}(e^{mZ_T^{t,z}}-e^{mz}) - \overline{\lambda}\frac{\gamma\rho}{1-\gamma}\left(\frac{a}{\delta}+\frac{1}{2}\delta\right)\int_t^T e^{mZ_s^{t,z}}ds \right.\\
&\left. + b\overline{\lambda}\frac{\gamma\rho}{\delta(1-\gamma)}\int_t^T e^{(2+m)Z_s^{t,z}}ds + \frac{1}{2}\frac{\gamma}{1-\gamma}\overline{\lambda}^2\int_t^T e^{2mZ_s^{t,z}}ds\right\},
\end{aligned}
\tag{126}
$$

Now, to use the Verification theorem we would need first to prove that this representation is smooth and that it solves the linear equation. This is the reverse implication of the Feynman-Kac formula theorem that we have proved beginning of the thesis, which we don't know to be true for our conditions. If this is verified, then we need to prove the uniform integrability condition of the Verification theorem. Noting that for this model the α function is given by

$$
\alpha(s,z) = \frac{1}{1-\gamma}\overline{\lambda}e^{(m-1)z} + \frac{\eta\delta\rho}{1-\gamma}\frac{\partial F}{\partial z}\frac{1}{e^zF(t,z)}
\tag{127}
$$

and that

$$
\begin{aligned}
w(s,\hat{X}_s^{t,x,z},Z_s^{t,z}) = \frac{x^\gamma}{\gamma}\exp&\left\{\gamma r(T-t) + \overline{\lambda}\int_t^s e^{(m+1)Z_u^{t,z}}\alpha(u,Z_u^{t,z})\,du\right.\\
&\left. -\frac{1}{2}\int_t^s e^{2Z_u^{t,z}}\alpha(u,Z_u^{t,z})^2\,du + \int_t^s e^{Z_u^{t,z}}\alpha(u,Z_u^{t,z})\,dW_u^1\right\}\ \left(F(s,Z_s^{t,z})\right)^\eta,
\end{aligned}
\tag{128}
$$

one approach is to prove that for all sequences of stopping times $(\theta_n)_{n\in\mathbb{N}}$

$$
\mathbb{E}\left[w(\theta_n,\hat{X}_{\theta_n}^{t,x,z},Z_{\theta_n}^{t,z})^q\right] \le C, \qquad \textit{for some } C>0,\ q>1.
$$

5.3 A numerical method for the Feynman-Kac representation

The Feynman-Kac representation suggests a Monte Carlo approximation scheme for the approximation of the value function. For a general reference for these methods we refer to [7].

Considering a general stochastic volatility model (A, B, μ) we want to approximate the expectation

$$
\begin{aligned}
\mathbb{E}\Bigg[\exp&\left\{\frac{\gamma\rho}{1-\gamma}\int_z^{Z_T^{t,z}}\frac{\lambda(y)}{A(y)}\,dy - \frac{\gamma\rho}{1-\gamma}\int_t^T B(Z_s^{t,z})\frac{\lambda(Z_s^{t,z})}{A(Z_s^{t,z})}\,ds - \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\int_t^T A(Z_s^{t,z})\lambda'(Z_s^{t,z})\,ds\right.\\
&\left. + \frac{1}{2}\frac{\gamma\rho}{1-\gamma}\int_t^T \lambda(Z_s^{t,z})A'(Z_s^{t,z})\,ds\ + \frac{1}{2}\frac{\gamma}{1-\gamma}\int_t^T \lambda(Z_s^{t,z})^2\,ds\right\}\Bigg],
\end{aligned}
\tag{129}
$$

where $\lambda(z) = \frac{\mu(z)-r}{e^z}$ and $Z_s^{t,z}$ is the solution of the stochastic differential equation

$$dZ_s^{t,z} = B(Z_s^{t,z})\,ds + A(Z_s^{t,z})\,dW_s, \quad Z_t^{t,z} = z$$

where $(t,z) \in [0,T] \times \mathbb{R}$ and W_s is some Brownian motion. This is the Feynman-Kac representation for F, solution of equation (108).

Usually the functions A, B, λ take some exponential form, as was the case of the models considered here, so simulating the above stochastic differential equation directly through an Euler method may be unwise. Instead we propose we simulate the process $Y_s^{t,z} = e^{d_0 Z_s^{t,z}}$ where $d_0 \in \mathbb{R}\backslash\{0\}$. Applying the Itô formula formula we get the equation for $Y_s^{t,z}$

$$dY_s^{t,z} = \left(d_0 Y_s^{t,z} B\left(\frac{1}{d_0}\log(Y_s^{t,z})\right) + \frac{1}{2} d_0^2\, Y_s^{t,z} A\left(\frac{1}{d_0}\log(Y_s^{t,z})\right)^2 \right) ds + d_0 Y_s^{t,z} A\left(\frac{1}{d_0}\log(Y_s^{t,z})\right) dW_s, \tag{130}$$

Since our equations are homogeneous in time, i.e., B and A don't depend on the time variable, simulating $Y_s^{t,z}$ in $[t,T]$ is the same as simulating $Y_s^{0,z}$ in $[0, T-t]$, so we only need to simulate the process $Y_s^{0,z}$ in $[0,T]$.

Fix a starting value $z \in \mathbb{R}$ and a discretization of time $0 = t_0 < t_1 < \cdots < t_N = T$, where we suppose $t_i - t_{i-1} = h$ for some $h > 0$, $i = 1, \ldots, N$. We simulate the process $Y_s^{t,z}$ by an Euler method. Take $y = e^{d_0 z}$ and consider the scheme

$$\begin{cases} \hat{Y}_0^y = y; \\ \hat{Y}_{i+1}^y = \left(d_0 \hat{Y}_i^y B\left(\frac{1}{d_0}\log(\hat{Y}_i^y)\right) + \frac{1}{2} d_0^2\, \hat{Y}_i^y A\left(\frac{1}{d_0}\log(\hat{Y}_i^y)\right)^2 \right) h + d_0 \hat{Y}_i^y A\left(\frac{1}{d_0}\log(\hat{Y}_i^y)\right)\sqrt{h}\mathcal{N}_{i+1}. \end{cases} \tag{131}$$

Here $\mathcal{N}_i$, $i = 1, \ldots, N$, are independent random variables with normal distribution, mean 0 and variance 1. Using this method, the random variable $\hat{Y}_i^y$ is an approximated simulation of $Y_{t_i}^{0,z}$.

Finally we need to approximate the integrals, for example the integral $\frac{1}{2}\frac{\gamma}{1-\gamma}\int_t^T \lambda(Z_s^{t,z})^2\,ds$. Consider some continuous function $f \in C(\mathbb{R})$. Using a right endpoint approximation of the integral we have for the starting value $t = 0$

$$\int_0^T f(Z_s^{0,z})\,ds = \int_0^T f\left(\frac{1}{d_0}\log(Y_s^{0,z})\right) ds \approx h\sum_{i=1}^N f\left(\frac{1}{d_0}\log(Y_{t_i}^{0,z})\right) \approx h\sum_{i=1}^N f\left(\frac{1}{d_0}\log(\hat{Y}_i^y)\right).$$

On the other hand, noting that the random variables $Y_{t_i}^{t_i,z}, Y_{t_{i+1}}^{t_1,z}, \ldots, Y_{t_N}^{t_i,z}$, are also approximately simulated by $\hat{Y}_0^y, \hat{Y}_1^y, \ldots, \hat{Y}_{N-i}^y$, since the stochastic differential equation is homogeneous as already noted, we also have

$$\int_{t_i}^T f(Z_s^{t_i,z})\,ds = \int_{t_i}^T f\left(\frac{1}{d_0}\log(Y_s^{t_i,z})\right) ds \approx h\sum_{k=1}^{N-i} f\left(\frac{1}{d_0}\log(Y_{t_{i+k}}^{t_i,z})\right) \approx h\sum_{k=1}^{N-i} f\left(\frac{1}{d_0}\log(\hat{Y}_k^y)\right).$$

Given all these observations this suggests the following method.

1. Discretize the time variable by taking $0 = t_0 < t_1 < \cdots t_N = T$, such that $t_i - t_{i-1} = h$ for some $h > 0$, for all $i = 1, \ldots, N$.

2. Discretize the space variable by some values $0 < y_0 < y_1 < \cdots < y_M$. We approximate F in the points $F(t_i, \frac{1}{d_0}\log(y_j))$, for all $i = 0, \ldots, N-1$, $j = 0, \ldots, M$.

3. Choose a value $L \in \mathbb{N}$ which corresponds to the number of simulations of the paths. Then for $l = 1, \ldots, L$ and $j = 0, \ldots, M$ calculate recursively

$$\begin{cases} \hat{Y}_0^{j,l}, = y_j; \\ \hat{Y}_{i+1}^{j,l} = \left(d_0 \hat{Y}_i^{j,l} B\left(\frac{1}{d_0}\log(\hat{Y}_i^{j,l})\right) + \frac{1}{2} d_0^2\, \hat{Y}_i^{j,l} A\left(\frac{1}{d_0}\log(\hat{Y}_i^{j,l})\right)\right) h \\ \qquad\qquad + d_0 \hat{Y}_i^{j,l} A\left(\frac{1}{d_0}\log(\hat{Y}_i^{j,l})\right)\sqrt{h}\mathcal{N}_{i+1}^{j,l}. \end{cases}$$

for $i = 0, \dots, N-1$, where $\mathcal{N}_i^{j,l}$, $l = 1, \dots, L$, $i = 1, \dots, N$, $j = 0, \dots, M$, are independent random variables with normal distribution, mean 0 and variance 1, we need to simulate.

4. For $i = 0, 1, \dots, N-1$ and $j = 0, \dots, M$ do the following

   (a) For $l = 1, \dots, L$ calculate the sums $S_n^{l,i,j} = h \sum_{k=1}^{N-i} f_n(\hat{Y}_k^{j,l})$ for $n = 1, 2, 3, 4$ where

   - $f_1(y) = B\left(\frac{1}{d_0}\log(y)\right)\frac{\lambda(\frac{1}{d_0}\log(y))}{A(\frac{1}{d_0}\log(y))}$,
   - $f_2(y) = A\left(\frac{1}{d_0}\log(y)\right)\lambda'\left(\frac{1}{d_0}\log(y)\right)$,
   - $f_3(y) = A'\left(\frac{1}{d_0}\log(y)\right)\lambda\left(\frac{1}{d_0}\log(y)\right)$,
   - $f_4(y) = \lambda\left(\frac{1}{d_0}\log(y)\right)^2$.

   These are approximations of the time integrals in the Feynman-Kac representation.

   (b) Calculate

$$\begin{aligned}\hat{F}_{i,j} = \frac{1}{L}\sum_{l=1}^{L}\exp\Bigg(&\frac{\gamma\rho}{1-\gamma}\int_{\frac{1}{d_0}\log(y_j)}^{\frac{1}{d_0}\log(\hat{Y}_{N-i}^{j,l})}\frac{\lambda(y)}{A(y)}\,dy - \frac{\gamma\rho}{1-\gamma}S_1^{l,i,j} \\ &- \frac{1}{2}\frac{\gamma\rho}{1-\gamma}S_2^{l,i,j} + \frac{1}{2}\frac{\gamma\rho}{1-\gamma}S_3^{l,i,j} + \frac{1}{2}\frac{\gamma}{1-\gamma}S_4^{l,i,j}\Bigg).\end{aligned} \tag{132}$$

   Here we're supposing that the primitive $\lambda(y)/A(y)$ is easy to calculate so the above integral is easy to calculate also. The above expression is an approximation of F in the point $F(t_i, \frac{1}{d_0}\log(y_j))$.

We note that the sums $S_n^{l,i,j}$ are related by $S_n^{l,i,j} = S_n^{l,i+1,j} + hf_n(\hat{Y}_{N-i}^{j,l})$ so one can speed up the computations using this relation.

5.4 Numerical simulations

Since from section 4 we have an explicit solution, we can apply this approach there and compare the results to see if this method is working. For this we consider the case $\nu = 2$, so the Heston model considered by Liu [15].

In this case $e^{2Z_s^{t,z}}$ is a CIR process, so it makes sense to take $d_0 = 2$. For $Y_s^{t,z} = e^{2Z_s^{t,z}}$ we have the equation

$$dY_s^{t,z} = \kappa(\theta - Y_s^{t,z})\,ds + \delta\sqrt{Y_s^{t,z}}\,dW_s$$

which we can discretize using the Euler method. Also the functions f_1, f_2, f_3, f_4 are given by

$$f_1(y) = \frac{1}{2}(2\kappa\theta - \delta^2)\frac{\overline{\lambda}}{\delta} - \frac{\kappa\overline{\lambda}}{\delta}y, \quad f_2(y) = \frac{1}{2}\delta\overline{\lambda}, \quad f_3(y) = -\frac{1}{2}\delta\overline{\lambda}, \quad f_4(y) = \overline{\lambda}^2 y.$$

Since the error of the Euler method is, in general, of weak order $1/N$ and since the error of the Monte Carlo method is of order $1/\sqrt{L}$ we can consider $L = N^2$, so that the error is of order $1/N$.

In figure 1 we take parameters $\theta = 0.4,\ \gamma = 0.3,\ \delta = 0.1,\ \kappa = 0.3,\ \overline{\lambda} = 0.4,\ \rho = -0.5,\ T = 1,\ N = 50,\ L = 2500$ and approximate in an uniform interval of $y \in [0, 1]$, with $M = 20$. We notice that the stationary point, θ, is 0.4 so it makes sense to consider $y \in [0, 1]$. In the next figures the red line is the explicit solution found in section 4, and the points are the approximations obtained with the proposed method.

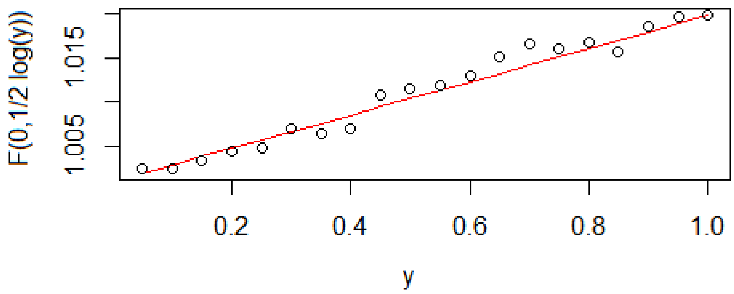

Figure 1: Approximation of $F(0, 1/2\log(y))$ in the Heston model, $\gamma = 0.3$ and $N = 50$.

Next we consider $N = 100$ and $L = 10000$ (the other parameters are the same). We present the results in figure 2.

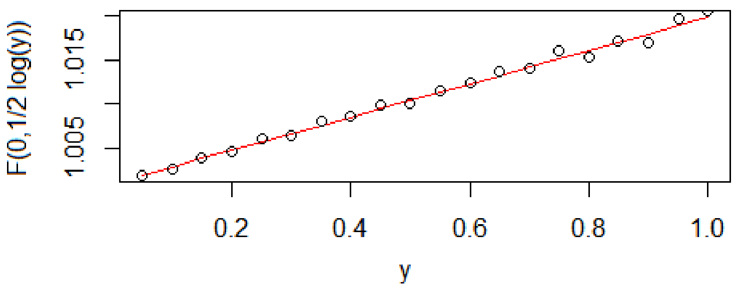

Figure 2: Approximation of $F(0, 1/2\log(y))$ in the Heston model, $\gamma = 0.3$ and $N = 100$.

We observe that the error is smaller for values of t closer to T, so for illustrative purposes we present the estimatives for $t = 0$ since these are ones with highest error.

We measure the RMS error, given by

$$\sqrt{\frac{1}{M}\sum_{j=1}^{M}\left(\hat{F}_{0,j} - F(0, \frac{1}{2}\log(y_j))\right)^2}.$$

For $N = 50$ we obtained an error of 0.0010729 and for $N = 100$ we have obtained an error of 0.0004636. We observe that in the second case the order of the error is less then in the first one.

Lets now consider a different value of γ. Taking into account the Monte Carlo expression (132) for the approximation of F, we see that every term is multiplied by the constant $\frac{\gamma}{1-\gamma}$. As γ goes to 1, $\frac{\gamma}{1-\gamma}$ goes to infinity, as such we expect the error in the approximation of F to grow for higher values of γ. Also as γ goes to 1 we approach a risk-neutral utility, and because the risk asset has a higher return then the non-risk asset, the optimal strategy will be to short the non-risk asset as much as possible, making the value function go to infinity, as such we expect higher values for F with higher values of γ.

Consider now $\gamma = 0.9$ and the same parameters. Take $N = 50$ and $L = 2500$. We present the results in figure 3.

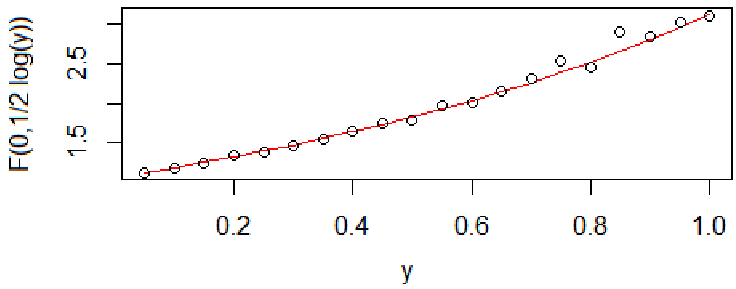

Figure 3: Approximation of $F(0, 1/2\log(y))$ in the Heston model, $\gamma = 0.9$ and $N = 50$.

Now consider $N = 100$ and $L = 10000$ (the other parameters are the same). We present the results in figure 4.

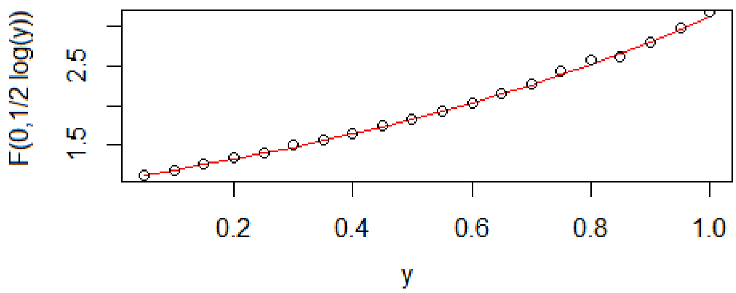

Figure 4: Approximation of $F(0, 1/2\log(y))$ in the Heston model, $\gamma = 0.9$ and $N = 100$.

The RMS here, is of 0.067 for $N = 50$ and $L = 2500$, and of 0.029 for $N = 100$ and $L = 10000$. Comparing with the error in the case of $\gamma = 0.3$, we have a different order of error, which is explained by the effect of the value of γ on F, as noted earlier.

Now lets apply this method on the 2-hypergeometric model. For this we take $d_0 = 1$. The equation for $Y_s^{t,z} = e^{Z_s^{t,z}}$ is given by

$$dY_s^{t,z} = \left(\left(a + \frac{1}{2}\delta^2 \right) Y_s^{t,z} - b \left(Y_s^{t,z} \right)^3 \right) ds + \delta Y_s^{t,z}\, dW_s,$$

and we may discretize using the Euler method. As for the functions f_1, f_2, f_3, f_4 we have

$$f_1(y) = a\frac{\overline{\lambda}}{\delta} y^m - b\frac{\overline{\lambda}}{\delta} y^{2+m}, \quad f_2(y) = \delta\overline{\lambda}m\, y^m, \quad f_3(y) = 0, \quad f_4(y) = \overline{\lambda}^2 y^{2m}.$$

We discretize now the interval $y \in [0,2]$ by taking $0 = y_0 < y_1, \cdots, y_M = 2$, such that $y_{j+1} - y_j$ is constant. We approximate the function F in the points $F(t_i, \log(y_j))$ using the above method. For the parameters $m = 1$, $T = 1$, $a = b = 0.3$, $\gamma = 0.3$, $\delta = 0.1$, $\overline{\lambda} = 0.3$, $\rho = -0.5$, $N = 100$ and $M = 20$ we have obtained an approximation with $L = 10000$ simulations illustrated in the following figure.

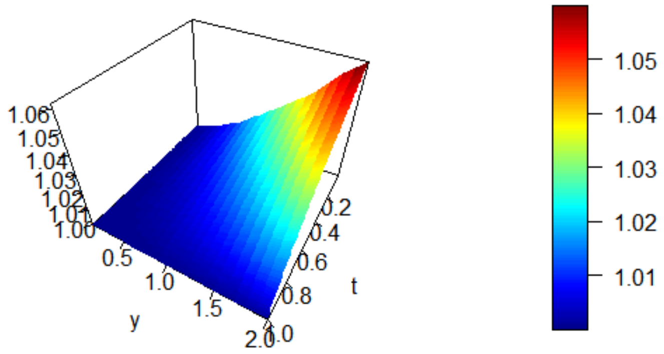

Figure 5: Approximation of $F(t, \log(y))$ in the 2-hypergeometric model, $\gamma = 0.3$

This seems to make sense with our model. Looking at the candidate for the value function

$$w(t,x,z) = \frac{x^\gamma}{\gamma} e^{\gamma r(T-t)} F(t,x,z)^\eta$$

we see that F acts as an extra value we get by also investing in the risk asset. For $F = 1$ we get the return of investing only on the non-risk asset and for higher values of F will correspond that extra value. In our model we have higher returns with higher volatility, and for low volatility the risk asset approaches the non-risk asset, so for low volatility, F, should be close to 1 and for high volatility we expect a higher value of F.

As already noted, as γ goes to 1, the error and values of F should increase. Taking $\gamma = 0.9$ (the other parameters remaining the same) we have obtained an approximation of F illustrated in the following figure.

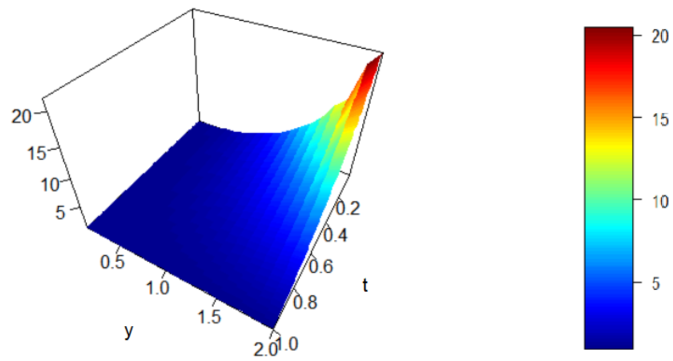

Figure 6: Approximation of $F(t, \log(y))$ in the 2-hypergeometric model, $\gamma = 0.9$

As we can see the maximum value is now 20, while in the previous case it is 1.06, so the values of F have increased as predicted. We should also expect a higher order of error for this case.

6 Concluding Remarks and Future Work

For the portfolio optimization problem, and when considering unbounded utilities, as is the case of the power utilities, there is no guarantee that the value function will be finite, and there can be portfolios that generate an infinite payoff, making the problem degenerate and impossible to solve using the HJB equations. This indeed seems to happen in many cases of stochastic volatility models. For the model considered in section 4, Korn and Kraft in [12] prove that if the condition (85) is not satisfied in the case $\nu = 2$, the problem has portfolios with infinite payoff. Also if we approach the HJB equations though a Feynman-Kac representation there are many instances where the expected value seems to be infinite.

Considering some utility function we want to use, one solution could be to transform this function to be constant after some very high value, making it bounded, and so making the value function bounded as well. In a way, since there are high values for the wealth process that can be nonsensical, having more wealth then that may not have more *utility*, so this seems like a realistic approach. If one checks our definition of utilities, one sees that we have tried to include these transformations, by allowing the derivative and second derivative of the utility function to be 0, where usually in the literature one finds strict inequalities (one needs an interpolation to keep the smoothness of the utility function). In this case, if one can find a bounded function that solves the HJB equation, this function satisfies automatically the uniform integrability condition, and so it is the value function.

Another issue with the way we have formulated the portfolio control problem, is that we have allowed for the portfolio to take any values in $\mathbb{R}$, effectively allowing for unlimited short sales, which can be unrealistic. Bounding the values that the portfolio can take, one may then be able to prove that the value function is finite. Still by doing this, the value function might not be smooth, and so we can't solve the problem through the strategies discussed here. On the other hand, it might be possible to solve the problem using viscosity solutions (see [3, 22]).

Numerical methods for general utilities seem to be scarce in the stochastic volatility portfolio problem, when optimizing utility from terminal wealth. The best one, as far as we know, seems to be presented in [5], where the authors develop a numerical approximation scheme using asymptotic methods. Yet, as we have noted, the problem can be not well posed, in the sense that the value function is infinite, and in this case this approach will not work, so one has to be careful as to the model one considers. Also the authors only consider the case when the parameters are either large or small (slowly varying and fast varying stochastic volatility) and we are unsure if this always applies.

In future work we would like to be able to approximate the optimal control. Continuing the work of section 5.3, and using the same notations, take points $0 = t_0 < t_1 < \cdots < t_N = T$ and $z_0 < z_1 < \cdots < z_M$ and suppose we have an approximation for F in the points (t_i, z_j), for some general stochastic volatility model (A, B, μ). Since the function α that determines the optimal controls is given by

$$\alpha(t,z) = \frac{\mu(z) - r}{(1-\gamma)e^{2z}} + \frac{\eta\rho A(z)}{(1-\gamma)e^{z}} \frac{1}{F(t,z)} \frac{\partial F}{\partial z}(t,z),$$

to approximate the optimal control we need the derivative of F in z. We can use a finite difference to give an approximation of the derivative of the function F, i.e., if we have a discretization of points given by $z_0 < z_1 < \cdots < z_M$ we can approximate the derivative of F in the points (t_i, z_j), $i = 0, 1, \ldots, N$ $j = 0, \ldots, M-1$, by

$$\frac{\partial F}{\partial z}(t_i, z_j) \approx \frac{\hat{F}(t_i, z_{j+1}) - \hat{F}(t_i, z_j)}{z_{j+1} - z_j}.$$

where $\hat{F}(t_i, z_i)$ correspond to our approximation of F in these points. We can therefore give an estimative for $\alpha(t_i, z_i)$ by

$$\alpha(t_i, z_j) \approx \frac{\mu(z_j) - r}{(1-\gamma)e^{2z_j}} + \frac{\eta\rho A(z_j)}{(1-\gamma)e^{z_j}} \frac{1}{\hat{F}(t_i, z_j)} \frac{\hat{F}(t_i, z_{j+1}) - \hat{F}(t_i, z_j)}{z_{j+1} - z_j}.$$

Here we have used a forward difference, but we can also use a backward or central difference. However this method may not give a good approximation, due to accumulation of errors and well known instabilities of finite difference methods.

To overcome these difficulties, we can apply the Malliavin Calculus and write $\frac{\partial F}{\partial z}$ as an expectation of a stochastic functional, and then find an approximation using a Monte Carlo method. This stochastic representation may also be useful to establish the Verification theorem.

References

[1] Aït-Sahalia, Y., & Jacod, J. (2014), High-Frequency Financial Econometrics. Princeton University Press.

[2] Chacko, G., & Viceira, L. (2005), Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets, The Review of Financial Studies, Volume 18, Issue 4, Pages 1369–1402.

[3] Fleming, W., & Soner, H. (2006), Controlled Markov Processes and Viscosity Solutions, Springer.

[4] Fonseca, J., & Martini, C. (2014), The α-Hypergeometric Stochastic Volatility Model, SSRN Electronic Journal.

[5] Fouque, J., & Sircar, R., & Zariphopoulou, T. (2013), Portfolio Optimization & Stochastic Volatility Asymptotics. SSRN Electronic Journal. 10.2139/ssrn.2473289.

[6] Friedman, A. (1975), Stochastic differential equations and applications, Volume 1, Academic Press.

[7] Glasserman, P., (2003), Monte Carlo Methods in Financial Engineering, Springer-Verlag New York.

[8] Heston, S., (1993), A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies, 6, 327–344.

[9] Karatzas, I., & Shreve, S., (1998), Brownian Motion and Stochastic Calculus, Springer-Verlag New York.

[10] Karatzas, I., & Shreve, S., (1998), Methods of Mathematical Finance, Springer-Verlag New York.

[11] Klenke, A., (2014), Probability Theory, A Comprehensive Course, Springer-Verlag London.

[12] Korn, R. & Kraft, H. (2004), On the Stability of Continuous-Time Portfolio Problems with Stochastic Opportunity Set. Mathematical Finance, Vol. 14, No. 3, pp. 403-414.

[13] Kraft, H. (2005), Optimal portfolios and Heston's stochastic volatility model: an explicit solution for power utility, Quantitative Finance, 5:3, 303-313.

[14] Kuo, H. (2006), Introduction to Stochastic Integration, Springer.

[15] Liu, J. (2007), Portfolio Selection in Stochastic Environments. The Review of Financial Studies, 20(1), 1-39.

[16] R. C. Merton (1969), Lifetime portfolio selection under uncertainty: The continuous-time case. Review of Economics and statistics, 51:247–257.

[17] R. C. Merton (1971), Optimum consumption and portfolio rules in a continuous-time model. Journal of economic theory, 3(4):373–413.

[18] Pham, H. (2007), Optimization methods in portfolio management and option hedging. 3rd cycle. Hanoi (Vietnam), pp.27.

[19] Pham, H. (2009), Continuous-time Stochastic Control and Optimization with Financial Applications, Springer.

[20] Privault, N., & She, Q. (2015) Option pricing and implied volatilities in a 2-hypergeometric stochastic volatility model, Applied Mathematics Letters 53.

[21] Sousa, R., & Cruzeiro, A., & Guerra, M. (2016), Barrier Option Pricing under the 2-Hypergeometric Stochastic Volatility Model, Journal of Computational and Applied Mathematics.

[22] Touzi, N. (2013), Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE, Springer.

[23] Zariphopoulou, T. (2001), A solution approach to valuation with unhedgeable risks, Finance and and Stochastics, 5:61.

[24] Zhu, J. (2010), Applications of Fourier Transform to Smile Modeling, Springer.

[25] Zhu, Y., & Avellaneda, M. (1997), A risk-neutral stochastic volatility model, International Journal of Theoretical and Applied Finance, Vol. 1, No. 2, 289–310.