Technische Universität München
Fakultät für Mathematik
Lehrstuhl für Finanzmathematik

Optimal Investment Strategies under Affine Markov-Switching Models

Theory, Examples and Implementation

Daniela Neykova

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:		Univ.-Prof. Claudia Czado, Ph.D.
Prüfer:	1.	Univ.-Prof. Dr. Rudi Zagst
	2.	Univ.-Prof. Dr. Ralf Korn, Technische Universität Kaiserslautern
	3.	Prof. Dr. Marcos Escobar, Ryerson University Toronto/Kanada (nur schriftliche Beurteilung)

Die Dissertation wurde am 01.09.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 07.01.2016 angenommen.

Abstract

In a multidimensional affine framework we consider a portfolio optimization problem with finite horizon, where an investor aims to maximize the expected utility of her terminal wealth. We state a very flexible asset price model that incorporates several risk factors modeled both by diffusion processes and by a Markov chain. We apply Merton's approach, because we are dealing with an incomplete market. Based on the semimartingale characterization of Markov chains, we first derive the Hamilton-Jacobi-Bellman (HJB) equations that, in our case, correspond to a system of coupled non-linear partial differential equations (PDE). Exploiting the affine structure of the model we solve the corresponding HJB equations explicitly up to an expectation only over the Markov chain or equivalently up to a system of simple ODEs. Furthermore, general verification theorems are proved. These results are provided both for the constant relative risk-aversion (CRRA) and hyperbolic absolute risk-aversion (HARA) utility functions. The relevance of the presented general model is illustrated on various examples including among others a stochastic short rate model with trading in the bond and the stock market, and a multidimensional stochastic volatility and stochastic correlation model. Precise verification results for all examples are provided. Economic interpretations of the models and results complement the theoretical analysis.

Zusammenfassung

Diese Arbeit befasst sich mit dem Portfoliooptimierungsproblem eines Investors, der den erwarteten Nutzen seines Endvermögens maximieren möchte. Das betrachtete multidimensionale Preismodell berücksichtigt verschiedene Risikofaktoren, die sowohl durch Diffusionsprozesse, als auch durch eine Markovkette modelliert werden. Da es sich um einen unvollständigen Markt handelt, wenden wir Merton's Methode an. Ausgehend von der Semimartingaldarstellung der Markovkette leiten wir zuerst die Hamilton-Jacobi-Bellman Gleichungen her, die in diesem Fall einem System von gekoppelten nichtlinearen partiellen Differentialgleichungen entsprechen. Wir nutzen die affine Struktur des Modells aus und finden explizite Lösungen für das betrachtete Problem in Form von einem Erwartungswert nur über die Markovkette oder äquivalent dazu einem System von einfachen gewöhnlichen Differentialgleichungen. Zudem, beweisen wir allgemeine Verifikationssätze. Die Ergebnisse werden für die *constant relative risk-aversion* (CRRA) und *hyperbolic absolute risk-aversion* (HARA) Nutzenfunktionen hergeleitet. Die Relevanz der dargestellten Ergebnisse wird veranschaulicht durch diverse Beispiele, die unter anderem ein Bond-Aktien-Modell mit stochastischen Zinssätzen und ein multidimensionales Model mit stochastischer Volatilität und stochastischer Korrelation einschließen. Detaillierte Verifikationsergebnisse werden für alle Beispiele bewiesen. Die ökonomische Interpretation des Models und der Ergebnisse rundet die theoretische Analyse ab.

Acknowledgment

First of all, I would like to express my deep gratitude to my supervisors Prof. Dr. Rudi Zagst and Prof. Dr. Marcos Escobar for all their encouragement and support over the last years. I am very thankful to Prof. Zagst for giving me the opportunity to write this thesis and for introducing me to the general topic of my research. I would like to give warm thanks to Prof. Escobar for his commitment and for sharing his research experience with me. My research stays in Toronto were very beneficial to my work. I am thankful to both of them for their valuable advice and feedback in numerous discussions. Our joint work was always a pleasure.

Furthermore, I would like to thank my colleagues at the Chair of Mathematical Finance for the very pleasant and academically productive working atmosphere. Especially, I would like to thank Peter Hieber and Lexuri Fernández for their encouragement and helpful discussions.

I also would like to acknowledge the financial support from the Women for Math Science Program, the German Academic Exchange Service (DAAD) and the TUM Graduate School that enabled my research stays in Toronto and fruitful conference visits.

I am very thankful to my family and friends for their encouragement throughout the last years. Special thanks go to my husband, Andrew, for his caring support in all situations.

Contents

Chapter 1

Introduction

1.1 Literature overview and motivation

In this thesis we derive optimal investment strategies by maximizing the expected utility from terminal wealth under a very flexible model, influenced simultaneously by a (multidimensional) continuous stochastic factor and a Markov chain. We do this for two utility functions: the CRRA, called also power utility, and the HARA utility function. In what follows we embed this problem in the existing literature and motivate its relevance.
The utility maximization problem was first stated and solved in continuous-time for the Black-Scholes model by [80] and [81]. Thereafter, many authors further developed this theory for more sophisticated and realistic market models. One important extension is the stochastic modeling of the asset returns volatility and the risk premium, because the constant Black-Scholes drift and volatility cannot reflect important empirical observations, such as asymmetric fat-tailed stock return distributions and volatility clustering (see, e.g., [91], [18], and [41] for detailed empirical studies). Some of the most popular examples of stochastic volatility models include the affine models by [96] and [58]. Optimal investment rules for one-dimensional continuous-time models with an additional stochastic factor that follows a diffusion process were derived, e.g., in [67], [104], and [19]. [74] proved rigorously their validity in the Heston model. Using martingale theory [63] solved the problem of maximizing expected utility for affine stochastic volatility models with jumps. [20] and [14] motivated in this context a model with a time-varying expected return by the fact that investors cannot observe directly the mean of the returns and their estimate changes as time goes by. As motivated in the next paragraph, in this thesis we extend the existing results for affine models with continuous stochastic factors by Markov switching.
Although the incorporation of stochastic volatility and drift terms makes asset price modeling more realistic, it does not capture long-term macroeconomic developments. Such fundamental factors can be described using Markov chains, wherein each state of the Markov chain describes a different market situation, such as a crisis or a boom-

ing economy. Since the introduction of a Markov-switching autoregressive model by [53], Markov chains have found several applications in financial mathematics. Empirical studies on Markov-modulated models can be found e.g. in [11] and [105] for the Markov-switching Black-Scholes and the Markov-switching Vasicek models, respectively. Pricing of bonds in this context is studied in [37]. The literature on utility maximization considers mainly the Black-Scholes model. [8] solved the problem of maximizing utility from terminal wealth when trading in one risky asset and the bank account, while [95] maximized the utility from consumption over an infinite time horizon in a multiasset environment. [48] added inflation and [76] considered the case with transaction costs. By a fixed-point argument [77] derived an implicit probabilistic representation for the value function when maximizing expected utility from consumption and terminal wealth over a finite time horizon. Furthermore, empirical confirmation for the significantly better performance of models exhibiting a stochastic factor and Markov switching in comparison to models without Markov switching can be found in [30] and [35].
All this motivates why this thesis combines both sources of randomness - a continuous stochastic factor and a Markov chain - in the context of portfolio optimization. To assure the analytic tractability of our model, we assume an affine structure. Such models were recently considered for derivative pricing. [40] derived pricing formulas for volatility swaps and [83] and [87] applied perturbation methods for option pricing under different Markov-switching affine models. We extend the existing literature by solving the problem of maximizing the utility from terminal wealth in this framework.
We derive the portfolio optimization results in great detail first in a one-dimensional model with a bank account and one risky asset. This is done in Chapter 4. Afterwards, in Chapter 5, we direct our attention to the multidimensional context with several traded assets and multiple stochastic factors. Our interest in the multi-asset case is motivated by the fact that when considering portfolio optimization it is crucial to account for the interaction between the single assets. Empirical evidences for the importance of multiasset modeling and stochastic correlation for utility maximization problems can be found in [23]. Furthermore, the existence of various stochastic factors influencing the asset price processes is also supported by numerous empirical findings in the literature. More precisely, [31] confirmed empirically that various stochastic factors drive the stock price volatility and [45] found empirical evidences for the existence of fast and slow mean-reverting factors influencing asset prices. Besides volatility, additional sources of randomness might be stochastic interest rates or factors explaining the expected return of assets, see [93] and [4] for an empirical verification of such models. What is more, multidimensional models allow for simultaneous consideration of several risk factors, for sophisticated modeling of the correlation structure and for analysis of the optimal trading in different asset classes, such as bonds and stocks. All this emphasizes the importance of considering multifactor multiasset models in the context of portfolio optimization.
Utility maximization problems for multidimensional models were solved by [63] and [75] for affine and quadratic models without Markov switching, however without providing explicit verification results for multidimensional examples. [56] and [7]

filled this gap for a Wishart market model. For utility maximization in models with stochastic interest rates and trading in stocks and a bond consult [21] and [92]. In Chapter 5 of this thesis we solve the utility maximization problem in a general affine multidimensional model influenced by Markov switching and also prove easy to apply verification results.
After deriving the results with the power utility function we extend our study to the HARA utility. It includes the power utility as a special case and allows for more flexible modeling of the investor's risk preferences. Optimal portfolio and consumption rules for the HARA utility function were derived by [81] in the context of the Black-Scholes model. Two decades later [15] related the resulting solutions to the so-called constant proportion portfolio insurance (CPPI) strategy. The CPPI strategy assures a lower bound on the terminal wealth by reducing the taken risk at lower wealth levels and increasing it at higher wealth levels. Thus, it provides downside protection in bearish markets but also allows for profit in bullish markets. A lower bound on terminal wealth is a desirable property e.g. for institutional investors such as pension funds that often have to guarantee their customers minimal performance. This motivates additionally the relevance of the HARA utility function. The connection between the HARA utility and the CPPI strategy under the Black-Scholes framework was further investigated by [97] and [103]. Optimal investment rules with the HARA utility for the Markov-modulated Black-Scholes model were derived in discrete time by [25] and in continuous time by [26]. In Chapter 6 we investigate the results in a multidimensional context wherein the asset price processes are influenced by a Markov chain and additional stochastic factors, modeling e.g. a stochastic short rate and stochastic volatility. We are also interested in comparing the results for the general HARA utility function with the results for the power utility function and in verifying the performance of the optimal strategy using real data.

1.2 Summary of the results and contributions to literature

To the best of our knowledge we are the first to apply both sources of randomness: Markov switching and continuous stochastic factors for utility maximization. More precisely, the contributions of this thesis are the following:

i) Based on the semimartingale characterization of Markov chains we formulate the HJB equation for a general class of affine models exhibiting simultaneously a (multidimensional) continuous stochastic factor and a Markov chain (see Section 3.3). It is in this case a system of coupled PDEs.

ii) In the one-dimensional case the optimal portfolio strategy is derived and the corresponding value function is explicitly computed up to an expectation over the paths of the Markov chain (see Section 4.2). Note that we allow for instantaneous correlation between the stochastic factor and the asset price process, which is in accordance with empirical observations (see [41]).

iii) We present a very flexible multidimensional model including continuous stochastic factors and a Markov chain and work out how exactly to include the Markov switching component so that the analytical tractability of the model remains preserved (see Section 3.1 and 5.2). According to our knowledge this is the most general affine model with Markov switching where the utility optimization problem can be solved explicitly. This general framework can be easily used for various applications by specifying the model parameters appropriately.

iv) We derive explicitly both, the optimal portfolio as well as the corresponding value function in the multidimensional context. In the case without leverage effect we allow for all parameters to depend on the Markov chain and obtain a simple probabilistic representation for the value function (see Section 5.2.1). If we assume correlated Brownian motions we apply a separability ansatz to obtain an explicit solution (see Section 5.2.2).

v) We prove a verification result that reduces the case with Markov switching to the one with deterministic coefficients, making it easy to check (see Theorem 3.5).

vi) We state and prove general verification theorems for the Markov-modulated models (see Theorems 4.2, 5.1 and 6.6).

vii) We extend the results from above to the HARA utility function and obtain as optimal solution a CPPI-type strategy for Markov-modulated models with a stochastic factor (see Section 6.2). We prove that the derived optimal strategy ensures a lower bound on the terminal wealth.

viii) We illustrate the flexibility of the proposed framework and the relevance of the results by various examples including:

- Markov-modulated Heston model (see Section 4.3)
- Markov-modulated stock-bond model, where the interest rate evolves according to the Vasicek model extended by Markov switching and the stock follows a Markov-modulated geometric Brownian motion (see Section 5.3.1)
- Markov-modulated model with two assets exhibiting stochastic covariance and correlation (see Section 5.3.2)
- Markov-modulated stock-bond model with stochastic interest rates and stochastic stock volatility (see Section 6.3).

Explicit solutions are derived for all these examples. Furthermore, we prove specific verification theorems for all considered examples that can be directly applied without any additional computations. All these models are relevant not only from a theoretical point of view but are also interesting for practitioners because they combine analytical solutions and desirable model properties such as stochastic interest rates, stochastic correlations and macroeconomic regimes.

ix) We illustrate the obtained results by various numerical studies and interpret them from an economic point of view (see Sections 4.3.4, 5.3.1, 5.3.2 and 6.4.3). For all examples a detailed interpretation of the model parameters and the results is provided. Various numerical studies complete the analysis and highlight the importance of considering both sources of randomness.

x) We exemplarily estimate one of the considered examples using real data and test the performance of the derived strategy. We provide a comparison between the HARA and the CRRA utility functions (see Section 6.4).

xi) We state a version of Itô's formula especially tailored for Markov-modulated Itô diffusions (see Section 2.6.2). It is obtained as a corollary of the general Itô formula for semimartingales.

xii) We state a version of the Feynman-Kac theorem applied to Markov-modulated stochastic processes (see Section 2.7). It contains a list of explicit sufficient conditions for the result to hold, so that it can be used in a straightforward way to solve systems of coupled PDEs in portfolio optimization problems and other applications, such as derivative pricing and risk measures in Markov-modulated models.

1.3 Structure of the thesis

In what follows we give an overview over the structure of the thesis.
Chapter 2 contains some existing and new mathematical results needed for the further derivations. In particular in Section 2.7 we state and prove a version of the Feynman-Kac theorem that is especially tailored for Markov-modulated models.
The next chapter introduces the model we consider throughout the whole thesis and the exact formulation of the optimization problem. Furthermore, we give an overview over the method we apply to solve it and we state the resulting HJB equations. A verification theorem that reduces the problem with Markov switching to one with time-dependent parameters is provided in Section 3.3.
Chapter 4 deals with the one-dimensional case. After deriving the solution in an auxiliary market with deterministic time-dependent parameters, we continue with the Markov-modulated model. The solution for the optimal portfolio and value function as well as general verification theorems are stated in Section 4.2. Here we consider separately the cases with and without correlation between the Brownian motions driving the stock and the stochastic factor. In Section 4.3 we apply the derived results to a Markov-modulated Heston model and analyze them within a numerical study.

Chapter 5 extends the results to multiple dimensions. Again we first solve the problem with time-dependent coefficients and then deal with the Markov-modulated model. The solutions with and without instantaneous correlation between the assets and the stochastic factors are presented separately in Sections 5.2.1 and 5.2.2. Two examples complement the multidimensional analysis in Section 5.3. The first one includes stochastic interest rate and trading both in a bond and a stock. The second one presents a two-dimensional model, where the traded assets exhibit not only stochastic volatility but also stochastic correlation. This framework can be considered as a generalization of the Markov-modulated Heston model from Chapter 4 to multiple dimensions and stochastic correlation between the assets. For both examples we derive the explicit solutions and prove very easy to apply verification results. The derived solutions are illustrated by various numerical computations and are interpreted from an economic point of view.

The results so far are derived with the CRRA utility function. In Section 6 we generalize them to the HARA utility function. To this aim we consider a multidimensional model with a stochastic interest rate modeled by an Ornstein-Uhlenbeck process with a Markov-switching mean-reverting level, and explicitly identify one of traded assets by a bond. The theoretical results are presented in Section 6.2 and a specific example with stochastic volatility is studied in Section 6.3. This model is a generalization of the Markov-modulated Heston model from Chapter 4 to stochastic interest rates and trading in a bond, and extends the first example from Chapter 5 to stochastic volatility. In Section 6.4 we estimate the model parameters for the considered model and illustrate the optimal portfolio using real data.

Chapter 7 summarizes the main findings of this study and concludes.

Lengthy proofs are outsourced in the appendix.

Chapter 2

Mathematical preliminaries

In this chapter we first recall some general results from stochastics (Sections 2.1-2.6). Afterwards, in Section 2.7 we state and prove a Feynman-Kac theorem especially tailored for multidimensional models with Markov switching. To the best of our knowledge, this result is a novel contribution. Finally, in Section 2.8 we recall some basic concepts concerning utility functions.

2.1 General notation

We start with the basic notation and abbreviations used in the whole thesis without any further explanation. They are summarized in Table 2.1.

symbol	explanation
$\mathbb{N}$	the set of natural numbers including 0
$\mathbb{R}$	the set of real numbers
$[0, \infty)$	the set of all non-negative real numbers
$\mathbb{R}_{>0}$	the set of all strictly positive real numbers
$\bar{\mathbb{R}}_{\geq 0}$	the set of all positive real numbers and ∞
$\mathbb{R}^d$	the Euclidean d-dimensional space for any $d \in \mathbb{N}\backslash 0$
$\mathbb{R}^{d_1,d_2}$	the set of all $d_1 \times d_2$-dimensional matrices with elements in $\mathbb{R}$ for any $d_1, d_2 \in \mathbb{N}\backslash 0$
$\bar{y}_i$	the i-th element of a vector $\bar{y} \in \mathbb{R}^d$
$Y_{i,j}$	the element in position (i, j) of a matrix $Y \in \mathbb{R}^{d_1,d_2}$
$diag(\bar{y})$	the $d \times d$ matrix with the elements of $\bar{y} \in \mathbb{R}^d$ on the diagonal and zeroes otherwise
$Tr(Y)$	the trace of a matrix $Y \in \mathbb{R}^{d,d}$
$\bar{y}', Y'$	the transpose of vector $\bar{y} \in \mathbb{R}^d$, respectively matrix $Y \in \mathbb{R}^{d_1,d_2}$
$\|y\|, \|\bar{y}\|$	the absolute value of a real number $y \in \mathbb{R}$; for vectors $\bar{y} \in \mathbb{R}^d$ the absolute value is to be taken element-wise
$\|\|\bar{y}\|\|$	the Euclidean norm of $\bar{y} \in \mathbb{R}^d$

$h(z)$	a function $\mathbb{R}^d \to \mathbb{R}, z \mapsto h(z)$
$h_{z_i}(z)$	the derivative of function h with respect to (w.r.t.) z_i
$h_z(z)$	the gradient of function h
$h_{z_i,z_j}(z)$	the second derivative of function h w.r.t. z_i and z_j
$h_{zz'}(z)$	the Hessian matrix of function h
$h(z)\|_{z_0}$	the value of function $h : z \mapsto h(z)$ at point z_0
$\mathscr{C}^0(D^h)$	the set of all continuous functions $h : D^h \to \mathbb{R}$
$\mathscr{C}^{k_1,\ldots,k_d}(D_1^h \times \ldots \times D_d^h)$	the set of all functions $h : D_1^h \times \ldots \times D_d^h \to \mathbb{R}$, $(x_1, \ldots, x_d) \mapsto h(x_1, \ldots, x_d)$ that are k_i-times continuously differentiable in x_i, for all $i = 1, \ldots, d$; $k_i = 0$ means continuity in x_i
$\mathbb{1}_{\{\ldots\}}$	the indicator function: it has value 1 if the condition $\{\ldots\}$ is fulfilled and 0 otherwise
Ω	state space
$\mathscr{F}$	Σ-algebra on Ω
$\mathbb{F} = \{\mathscr{F}_t\}_{t\in[0,\infty)}$	filtration on $(\Omega, \mathscr{F})$
$\mathbb{P}$	probability measure
$(\Omega, \mathscr{F}, \mathbb{P})$	probability space
$(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$	filtered probability space
$\mathscr{B}(A)$	the Borel-Σ-algebra on A, i.e. the smallest Σ-algebra containing all open subsets of A
$\mathscr{F} \otimes \mathscr{G}$	product Σ-algebra of $\mathscr{F}$ and $\mathscr{G}$, i.e. the smallest Σ-algebra containing all sets $A_1 \times A_2 \in \mathscr{F} \times \mathscr{G}$
$RV : \Omega \to \mathbb{R}$	a random variable on $(\Omega, \mathscr{F}, \mathbb{P})$
$RV \sim Distr$	the random variable RV has distribution $Distr$
$\mathscr{N}(\mu, \Sigma^2)$	the normal distribution with mean $\mu \in \mathbb{R}$ and variance $\Sigma^2 \in \mathbb{R}$
$\mathscr{N}_d(\mu, \Sigma\Sigma')$	the d-dimensional normal distribution with mean vector $\mu \in \mathbb{R}^d$ and covariance matrix $\Sigma\Sigma' \in \mathbb{R}^{d,d}$
$\mathbb{E}_{\mathbb{P}}[RV], \mathbb{E}[RV]$	the expectation of RV under $\mathbb{P}$; whenever the measure is clear from the context, we omit the index
$Var_{\mathbb{P}}(RV)$, $Var(RV)$	the variance of RV under $\mathbb{P}$; whenever the measure is clear from the context, we omit the index
$sd_{\mathbb{P}}(RV)$, $sd(RV)$	the standard deviation of RV under $\mathbb{P}$; whenever the measure is clear from the context, we omit the index
$\mathbb{E}[RV\|A]$	the conditional expectation of RV given event $A \subset \Omega$
$\mathbb{E}[RV\|\mathscr{G}]$	the conditional expectation of RV given the sub-Σ-algebra $\mathscr{G} \subset \mathscr{F}$
$X = \{X(t)\}_{t\in[0,\infty)}$	stochastic process on $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$
$X_{t,x}$	process X started at time t in point x
$\mathbb{F}^X = \{\mathscr{F}_t^X\}_{t\in[0,\infty)}$	the natural filtration of process X
dim.	dimension(al)
HJB	Hamilton-Jacobi-Bellman
MS	Markov switching
ODE	ordinary differential equation
p.	page

PDE	partial differential equation
SDE	stochastic differential equation
w.r.t.	with respect to

Table 2.1: Basic notation and abbreviations.

2.2 A few basic tools

In this section we summarize for the convenience of the reader some very basic definitions from stochastics that we will apply later on. They are cited from [12]. If nothing else is stated all objects are defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Definition 2.1 (Integrable random variable)
A random variable X is called integrable if $\mathbb{E}[|X|] < \infty$.

Definition 2.2 (Uniformly integrable)
A family $\{X(t)\}_{t\in[0,\infty)}$ of real-valued random variables is called uniformly integrable if for any $\varepsilon > 0$, there exists a constant K_ε such that $\mathbb{E}[|X(t)|\mathbb{1}_{|X(t)|>K_\varepsilon}] < \varepsilon$ for all $t \in [0,\infty)$.

Definition 2.3 (Almost sure convergence)
Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables. We say that it converges almost surely (a.s.) to the random variable X and write:

$$a.s. \lim_{n\to\infty} X_n = X,$$

if

$$\mathbb{P}\big(\{\omega \in \Omega | \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\big) = 1.$$

Definition 2.4 (Convergence in mean)
Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables with $\mathbb{E}[|X_n|] < \infty$ for all $n \in \mathbb{N}$. Further, let X be a random variable with $\mathbb{E}[|X|] < \infty$. We say that $\{X_n\}_{n\in\mathbb{N}}$ converges to X in mean if the following holds:

$$\lim_{n\to\infty} \mathbb{E}[|X_n - X|] = 0.$$

2.3 Semimartingales

Now we introduce a very general class of stochastic processes, called semimartingales. Note that in the whole thesis we consider only processes with values in $\mathbb{R}^d$ for some $d \in \mathbb{N}\backslash 0$. They are defined on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$ unless something else is explicitly specified. Recall the following definitions and statements cited from [59]:

Definition 2.5 (Càdlàg/càg process, p. 3)
A stochastic process X is called càdlàg if all its paths are right-continuous and admit left-hand limits. A stochastic process X is called càg if all its paths are left-continuous.

Definition 2.6 (Jumps, p. 3, Equation 1.8)
Let X be a càdlàg process. Then we define two other processes $X^- = \{X^-(t)\}_{t\in[0,\infty)}$ and $\Delta X = \{\Delta X(t)\}_{t\in[0,\infty)}$ by:

$$X^-(0) := X(0), X^-(t) := \lim_{s\uparrow t} X(s) \text{ for } t > 0$$
$$\Delta X(t) := X(t) - X^-(t).$$

Definition 2.7 (Stopped process, p. 3, Definition 1.9)
Let X be a process and τ a mapping $\Omega \to \bar{\mathbb{R}}_{\geq 0}$. We define process X^τ by:

$$X^\tau(t) := X\left(\min(\tau, t)\right),$$

and call it the process stopped at time τ.

Definition 2.8 (Evanescent set, p. 3, Definition 1.10)
A set $A \subset \Omega \times [0,\infty)$ is called evanescent if:

$$\mathbb{P}\big(\big\{\omega \in \Omega | \exists t \geq 0 \text{ with } (\omega, t) \in A\big\}\big) = 0.$$

Two stochastic processes X and Y are called indistinguishable if the set $\{(\omega, t) \in \Omega \times [0,\infty) | X(t,\omega) \neq Y(t,\omega)\}$ is evanescent.

Definition 2.9 (Adapted process, p. 5, Definition 1.20)
A stochastic process $X = \{X(t)\}_{t\in[0,\infty)}$ is called adapted to the filtration $\mathbb{F}$ ($\mathbb{F}$-adapted) if $X(t)$ is $\mathscr{F}_t$-measurable for all $t \in [0,\infty)$.

Definition 2.10 (Optional process, p. 5, Definition 1.20)
The Σ-algebra $\mathscr{O}$ on $\Omega \times [0,\infty)$ generated by all càdlàg $\mathbb{F}$-adapted processes (considered as mappings on $\Omega \times [0,\infty)$) is called the optional Σ-algebra. A process that is $\mathscr{O}$-measurable is called optional.

Definition 2.11 (Localized class, p. 8, Definition 1.33)
Let $\mathscr{C}$ be a class of processes. The corresponding localized class $\mathscr{C}_{loc}$ is defined as follows: a process X belongs to $\mathscr{C}_{loc}$ if and only if there exists an increasing sequence $\{T_n\}_{n\in\mathbb{N}}$ of stopping times (depending on X) with a.s. $\lim_{n\to\infty} T_n = \infty$ such that each stopped process X^{T_n} is in $\mathscr{C}$. The sequence $\{T_n\}_{n\in\mathbb{N}}$ is called a localizing sequence for X (relative to $\mathscr{C}$).

Definition 2.12 (Martingale, p. 10, Definition 1.36)
Let M be an adapted process on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$. Assume that its $\mathbb{P}$-almost all paths[1] *are càdlàg and that every $X(t)$ is integrable. X is called a martingale (resp. submartingale, resp. supermartingale) if for all (s,t) with $s \leq t$ it holds:*

$$X(s) = \mathbb{E}[X(t)|\mathscr{F}_s] \ (resp.\, X(s) \leq \mathbb{E}[X(t)|\mathscr{F}_s], resp.\, X(s) \geq \mathbb{E}[X(t)|\mathscr{F}_s]).$$

Definition 2.13 (Uniformly integrable martingale, p. 10, Definition 1.40)
A martingale M is called uniformly integrable if the family of random variables $\{M(t)\}_{t\in[0,\infty)}$ is uniformly integrable. The class of all uniformly integrable martingales is denoted by $\mathscr{M}$.

Definition 2.14 (Local martingale, p. 11, Definition 1.45)
The processes in the localized class $\mathscr{M}_{loc}$ constructed from $\mathscr{M}$ are called local martingales. The set of all local martingales M such that $M(0) = 0$ is called $\mathscr{L}$.

Definition 2.15 (Square integrable martingale, p. 11, Definition 1.41)
A martingale M is square integrable if $\sup_{t\in[0,\infty)} \mathbb{E}\left[M^2(t)\right] < \infty$. The class of all such processes is denoted by $\mathscr{H}^2$. Its localized class is $\mathscr{H}^2_{loc}$ and the processes in $\mathscr{H}^2_{loc}$ are called locally squared integrable martingales.
For multidimensional processes the definition is to be understood component-wise (see p. 204 in [59]).

Definition 2.16 (Predictable process, p. 16, Definition 2.1)
The Σ-algebra $\mathscr{P}$ on $\Omega \times [0,\infty)$ generated by all càg processes (considered as mappings on $\Omega \times [0,\infty)$) is called the predictable Σ-algebra. A stochastic process is called predictable if it is $\mathscr{P}$-measurable.

Definition 2.17 (Finite variation, p. 27, Definition 3.1)
The set of all real-valued processes X that are càdlàg, adapted, with $X(0) = 0$ such that each path $t \mapsto X(t,\omega)$ has a finite variation over each finite interval $[0,t]$[2] *is denoted by $\mathscr{V}$.*
For a process $X \in \mathscr{V}$ we denote its variation by $\mathcal{V}^X$, i.e. $\mathcal{V}^X(t,\omega)$ is the total variation of the function $s \mapsto X(s,\omega)$ on [0,t].
$\mathscr{V}^+$ denotes the set of all real-valued processes that are càdlàg, adapted, with $X(0) = 0$ and non-decreasing paths. For processes $X \in \mathscr{V}^+$ we define $X(\infty) := \lim_{t\to\infty} X(t) \in \bar{\mathbb{R}}_{\geq 0}$ and call $X(\infty)$ the terminal variable.

Proposition 2.18 (p. 28, Proposition 3.5)
Consider a process $A \in \mathscr{V}$ and let H be an optional process, such that the process $B := \int_0^t H \mathrm{d}A$ is finite-valued for all $t \geq 0$. Then $B \in \mathscr{V}$.

[1] This means up to a set of paths of measure zero under $\mathbb{P}$.
[2] Recall e.g. from [12], p. 37, that a real-valued function h defined on the interval $[a,b]$ is called of finite variation over $[a,b]$ if $\sup_{(z_0,\dots,z_n)\in\mathcal{P}} \sum_{i=1}^n |h(z_i) - h(z_{i-1})| < \infty$, where $\mathcal{P}$ denotes the set of all partitions of the interval $[a,b]$: $\mathcal{P} = \{(z_0,\dots,z_n) : a = z_0 < z_1 < \dots < z_n = b, n \in \mathbb{N}\}$.

Definition 2.19 (Integrable variation, p. 28/29, Definition 3.6/3.7)
$\mathscr{A}^+$ denotes the set of all process $X \in \mathscr{V}^+$ that are integrable, i.e. $\mathbb{E}[X(\infty)] < \infty$. The set of all processes $X \in \mathscr{V}$ with integrable variation, i.e. $\mathbb{E}[\mathcal{V}^X(\infty)] < \infty$, is denoted by $\mathscr{A}$.
$\mathscr{A}^+_{\text{loc}}$ and $\mathscr{A}_{\text{loc}}$ are the localized classes constructed from $\mathscr{A}^+$ and $\mathscr{A}$, respectively. A process in $\mathscr{A}^+_{\text{loc}}$ is called a locally integrable adapted increasing process and a process in $\mathscr{A}_{\text{loc}}$ is called an adapted process with locally integrable variation.

Theorem 2.20 (Predictable quadratic covariation, p. 38, Theorem 4.2)
Let M and N be locally square-integrable martingales, i.e. in $\mathscr{H}^2_{\text{loc}}$. Then there exists a predictable process $\langle M, N\rangle \in \mathscr{V}$, unique up to an evanescent set, such that $MN - \langle M, N\rangle$ is a local martingale. Furthermore, if M and N are square integrable martingales then $MN - \langle M, N\rangle \in \mathscr{M}$. The process $\langle M, N\rangle$ is called the predictable quadratic covariation of the pair (M, N).

Definition 2.21 (Orthogonal local martingales, p. 40, Definition 4.11)
Consider two local martingales M and N. They are called orthogonal if their product MN is a local martingale.

Definition 2.22 (Purely discontinuous, p. 40, Definition 4.11)
Let M be a local martingale. If $M(0) = 0$ and M is orthogonal to all continuous local martingales then it is called a purely discontinuous local martingale.

Lemma 2.23 (p. 41, Lemma 4.14)
A local martingale that belongs to $\mathscr{V}$ is purely discontinuous.

Theorem 2.24 (p. 42, Theorem 4.18)
Any local martingale M admits a unique (up to indistinguishability) representation of the form:

$$M = M_0 + M^c + M^d,$$

where $M^c(0) = M^d(0) = 0$, M^c is a continuous local martingale, and M^d is a purely discontinuous local martingale. M^c is called the continuous part of M and M^d is its purely discontinuous part.

Definition 2.25 (Semimartingale, p. 43, Definition 4.21)
Processes X of the form:

$$X = X_0 + M + A, \tag{2.1}$$

where X_0 is finite-valued and $\mathscr{F}_0$-measurable, $M \in \mathscr{L}$ and $A \in \mathscr{V}$, are called semimartingales. The space of all semimartingales is denoted by $\mathscr{S}$.

Proposition 2.26 (p. 45, Proposition 4.27)
Let X be a semimartingale. There exists a unique (up to indistinguishability) continuous local martingale X^c with $X^c(0) = 0$, such that any decomposition $X = X_0 + M + A$ of type (2.1) meets $M^c = X^c$ (up to indistinguishability again). X^c is called the continuous martingale part of X.

Remark 2.27
For processes $X = (X_1, \ldots, X_d)'$ with values in $\mathbb{R}^d$ we define the sets $\mathscr{V}^d$, $\mathscr{V}^{+,d}$, $\mathscr{A}^d$, $\mathscr{A}^{+,d}$, $\mathscr{A}^d_{loc}$, $\mathscr{A}^{+,d}_{loc}$, $\mathscr{M}^d$, $\mathscr{L}^d$, $\mathscr{H}^{2,d}$, $\mathscr{H}^{2,d}_{loc}$ and $\mathscr{S}^d$ analogously as before, such that the corresponding property holds for each component X_i, $i = 1, \ldots, d$.

Definition 2.28 (Characteristics, p. 76, Definition 2.6)
Let $X = (X_1, \ldots, X_d)$ be a continuous d-dimensional semimartingale with the following semimartingale decomposition:

$$X = X_0 + M + A,$$

where $X_0 \in \mathbb{R}^d$ is $\mathscr{F}_0$-measurable, $M \in \mathscr{L}^d$, and $A \in \mathscr{V}^d$. Define process C with values in $\mathbb{R}^{d,d}$ by:

$$C_{i,j} := \langle M_i, M_j \rangle.$$

Then the pair (A, C) is called the characteristics of the continuous semimartingale X.
Note that this definition is restricted to continuous semimartingales as this complexity suffices for our applications. For the general definitions with jumps, see [59].

Proposition 2.29 (Factorization, p. 77, Proposition 2.9)
Let X be a d-dimensional continuous semimartingale. There exists a version of its characteristics (A, C), such that:

$$\begin{aligned} A_i(t) &= \int_0^t a_i(s) \mathrm{d}F(s) \\ C_{i,j}(t) &= \int_0^t c_{i,j}(s) \mathrm{d}F(s), \end{aligned} \tag{2.2}$$

where

- *$F \in \mathscr{A}^+_{loc}$ is a predictable process.*
- *a is a d-dimensional predictable process.*
- *c is a predictable process with values in the set of all symmetric non-negative $d \times d$ matrices.*

Definition 2.30 ($L^{2,d}(M)$, p. 204, Definition 6.3)
Let $M \in \mathscr{H}^{2,d}$, i.e. M is a d-dimensional process such that each element M_i is a locally square-integrable martingale. Define $C_{i,j} := \langle M_i, M_j \rangle$ and consider a factorization as in (2.2):

$$C_{i,j}(t) = \int_0^t c_{i,j}(s) \mathrm{d}F(s).$$

The set of all d-dimensional predictable processes H such that the increasing process $\{\int_0^t \sum_{i,j=1}^{d} H_i c_{i,j} H_j \mathrm{d}F(s)\}_{t\in[0,\infty)}$ is integrable, respectively locally integrable, is denoted by $L^{2,d}(M)$, respectively $L^{2,d}_{loc}(M)$.

Theorem 2.31 (Square integrable martingales, p. 204f, Theorem 6.4)
Let M be a d-dimensional locally square integrable martingale and $H \in L^{2,d}_{loc}(M)$. Then $\{\int_0^t H \mathrm{d}M\}_{t\in[0,\infty)}$ is a locally square integrable martingale. Further, it is a square integrable martingale if and only if $H \in L^{2,d}(M)$.

Theorem 2.32 (Itô's formula for semimartingales, p. 57, Theorem 4.47)
Let $X = (X_1, \dots, X_d)$ be a d-dimensional semimartingale, and $f : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ be a real-valued function with $f \in \mathscr{C}^{1,2}([0,\infty)\times\mathbb{R}^d)$. Then $f(X)$ is a semimartingale and we have:

$$\begin{aligned} f\big(t, X(t)\big) =& f\big(0, X(0)\big) + \int_0^t f_t\big(s, X^-(s)\big)\mathrm{d}s + \sum_{i=1}^{d} \int_0^t f_{x_i}\big(s, X^-(s)\big)\mathrm{d}X_i(s) \\ &+ \frac{1}{2}\sum_{i,j=1}^{d} f_{x_i,x_j}\big(s, X^-(s)\big)\mathrm{d}\langle X_i^c, X_j^c\rangle(s) \\ &+ \sum_{s\le t}\Big[f\big(s, X(s)\big) - f\big(s, X^-(s)\big) - \sum_{i=1}^{d} \Delta X_i(s) f_{x_i}\big(s, X^-(s)\big)\Big]. \end{aligned} \tag{2.3}$$

Note that the result in Theorem 4.47 from [59] is not explicitly stated with the dependence of function f on t, however Representation (2.3) follows directly by identifying in Theorem 4.47 the first component of the semimartingale with process dt.
Now we cite a result from [62], that states a set of very general sufficient conditions for an exponential of an affine process to be a martingale. Before we present the result,we need to define the concept of differential semimartingale characteristics. Again we do this in the context of continuous processes, the general definition with jumps can be found in [62].

Definition 2.33 (Differential semimartingale characteristics, [62])
Let X be a continuous d-dimensional semimartingale and denote its semimartingale characteristics by (A, C). Furthermore, assume that there exist:

- *a predictable processes* μ *with values in* $\mathbb{R}^d$, *such that* $A(t) = \int_0^t \mu(s)\mathrm{d}s$,
- *a predictable non-negative process* Γ *with values in the set of symmetric matrices in* $\mathbb{R}^{d\times d}$ *with* $C(t) = \int_0^t \Gamma(s)\mathrm{d}s$.

Then, the pair (μ, Γ) *is called the differential characteristics of the continuous semimartingale* X.

Now we cite Corollary 3.4 from [62].

Theorem 2.34 (Exponential affine martingales)
Let $X = (X_1, \ldots, X_d)'$ *be a d-dim continuous semimartingale with affine differential characteristics* $(\mu, \Gamma) \in \mathbb{R}^d \times \mathbb{R}^{d\times d}$, *i.e.*

$$\mu(t) = \begin{pmatrix} \mu_1(t) \\ \vdots \\ \mu_d(t) \end{pmatrix} = \underbrace{\begin{pmatrix} \alpha_1^0(t) \\ \vdots \\ \alpha_d^0(t) \end{pmatrix}}_{=:\alpha^0} + \sum_{i=1}^{d} X_i \underbrace{\begin{pmatrix} \alpha_1^i(t) \\ \vdots \\ \alpha_d^i(t) \end{pmatrix}}_{=:\alpha^i}$$

$$\Gamma(t) = \begin{pmatrix} \Gamma_{1,1}(t) & \ldots & \Gamma_{1,d}(t) \\ \vdots & \vdots & \vdots \\ \Gamma_{d,1}(t) & \ldots & \Gamma_{d,d}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} \beta_{1,1}^0(t) & \ldots & \beta_{1,d}^0(t) \\ \vdots & \vdots & \vdots \\ \beta_{d,1}^0(t) & \ldots & \beta_{d,d}^0(t) \end{pmatrix}}_{=:\beta^0}$$

$$+ \sum_{i=1}^{d} X_i \underbrace{\begin{pmatrix} \beta_{1,1}^i(t) & \ldots & \beta_{1,d}^i(t) \\ \vdots & \vdots & \vdots \\ \beta_{d,1}^i(t) & \ldots & \beta_{d,d}^i(t) \end{pmatrix}}_{=:\beta^i},$$

for some deterministic functions $\alpha_k^i, \beta_{k,l}^i : [0, \infty] \to \mathbb{R}$, $i \in \{0, \ldots, d\}$, $k, l \in \{1, \ldots, d\}$. *Further assume that there exists a number* $p \in \mathbb{N}$, $p \leq d$ *such that for all* $t \in [0, T]$:

i) $\alpha_k^i(t) \geq 0$ *if* $0 \leq i \leq p$, $1 \leq k \leq p$, $k \neq i$,

ii) $\alpha_k^i(t) = 0$ *if* $i \geq p + 1$, $1 \leq k \leq p$,

iii) $\beta_{k,l}^i(t) = 0$ *if* $0 \leq i \leq p$, $1 \leq k, l \leq p$ *unless* $k = l = i$,

iv) $\beta_{k,l}^i(t) = 0$ *if* $i \geq p + 1$, $1 \leq k, l \leq d$.

If additionally $\alpha^i(t)$ *and* $\beta^i(t)$ *are continuous in* $t \in [0, \infty)$, *for all* $0 \leq i \leq d$, *and the following condition holds for some* $1 \leq i \leq d$:

$$\alpha_k^i(t) + \frac{1}{2}\beta_{k,k}^i(t) = 0, \forall\ 0 \leq k \leq d, \tag{2.4}$$

then $\big\{ \exp\big(X_i(t)\big)\big\}_{t\in[0,\infty)}$ *is a martingale.*

2.4 Itô processes

Now we consider a special class of semimartingales called Itô processes. The following short summary is based on [12]. All results are cited from there unless something else is stated. Further details are provided there and in the references cited therein. In order to present the concept of Itô processes we need to introduce a special process called Brownian motion.

Definition 2.35 (Brownian motion, p. 160, Definition 5.3.1.)
An adapted process $W = \{W(t)\}_{t\in[0,\infty)}$ is called a Brownian motion (or a Wiener process), if W satisfies the following properties:

i) $W(0) = 0$, $\mathbb{P}$-*a.s.,*

ii) *W has independent increments, i.e. $W(t+s) - W(t)$ is independent of $\mathscr{F}_t^W$ for $s > 0$,*

iii) *W has Gaussian increments, $W(t+h) - W(t) \sim \mathscr{N}(0,h)$, for all $h > 0$,*

iv) *W has continuous paths, i.e. $t \mapsto W(t,\omega)$ is continuous in t for all $\omega \in \Omega$.*

This definition can be naturally extended to more dimensions.

Definition 2.36 (d-dimensional Brownian motion, p. 160)
W given by $W = \big(W_1,\dots,W_d\big)' = \big\{\big(W_1(t),\dots,W_d(t)\big)'\big\}_{t\in[0,\infty)}$ is called a d-dimensional Brownian motion (or d-dimensional Wiener process) if its components W_j, for $j = 1,\dots,d$ are independent Brownian motions.

Definition 2.37 (Itô process, p. 193, Section 5.6.2)
Let $X = (X_1,\dots,X_{d_1})'$ be a d_1-dimensional stochastic process. It is called an Itô process if it satisfies the following equation:

$$X(t) = X(0) + \int_0^t \mu(s)\mathrm{d}s + \int_0^t \Sigma(s)\mathrm{d}W(s) \tag{2.5}$$

$$= \begin{pmatrix} X_1(0) + \int_0^t \mu_1(s)\mathrm{d}s + \sum_{j=1}^{d_2} \int_0^t \Sigma_{1,j}(s)\mathrm{d}W_j(s) \\ \vdots \\ X_{d_1}(0) + \int_0^t \mu_{d_1}(s)\mathrm{d}s + \sum_{j=1}^{d_2} \int_0^t \Sigma_{d_1,j}(s)\mathrm{d}W_j(s) \end{pmatrix}, \tag{2.6}$$

where:

i) *$X(0) = (X_1(0),\dots,X_{d_1}(0))'$ is d_1-dimensional $\mathscr{F}_0$-measurable random variable,* *ii)* $\mu = (\mu_1, \ldots, \mu_{d_1})'$ *is a* d_1*-dimensional stochastic process, where* μ_i *is adapted and* $\int_0^t |\mu_i(s)|\mathrm{d}s < \infty$*, for all* $i = 1, \ldots, d_1$*,*

iii) $\Sigma = \begin{pmatrix} \Sigma_{1,1} & \cdots & \Sigma_{1,d_2} \\ \vdots & \cdots & \vdots \\ \Sigma_{d_1,1} & \cdots & \Sigma_{d_1,d_2} \end{pmatrix}$ *is a* $d_1 \times d_2$*-dimensional stochastic process, where* $\Sigma_{i,j}$ *is adapted,* $\mathscr{B}([0,\infty)) \otimes \mathscr{F}$*-measurable and* $\int_0^t \mathbb{E}[(\Sigma_{i,j})^2]\mathrm{d}s < \infty$*, for all* $i = 1, \ldots, d_1$ *and* $j = 1, \ldots, d_2$*,*

iv) $W = (W_1, \ldots, W_{d_2})'$ *is a* d_2*-dimensional Brownian motion.*

For convenience, we can rewrite Equation (2.6) in the following form:

$$\mathrm{d}X(t) = \mu(t)\mathrm{d}t + \Sigma(t)\mathrm{d}W(t).$$

We call this expression a stochastic differential equation (SDE) with drift μ *and diffusion term* Σ*.*

The next three results allow us to derive the dynamics of an Itô process under different probability measures.

Lemma 2.38 (Novikov condition, p. 198)
Let $W = (W_1, \ldots, W_d)$ *be a* d*-dimensional Brownian motion and* $\gamma = \{\gamma(t)\}_{t\in[0,T]}$ *a measurable, adapted,* d*-dimensional process with* $\int_0^T \gamma_i(t)\mathrm{d}t < \infty$ $\mathbb{P}$*-a.s., for all* $i = 1, \ldots, d$*. Assume that*

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T ||\gamma(s)||^2\mathrm{d}s\Big)\Big] < \infty. \tag{2.7}$$

Then process $L = \{L(t)\}_{t\in[0,T]}$ *defined as follows:*

$$L(t) := \exp\Big(-\int_0^t \gamma'(s)\mathrm{d}W(s) - \frac{1}{2}\int_0^t ||\gamma(s)||^2\mathrm{d}s\Big) \tag{2.8}$$

is a continuous martingale. Note that process γ *is not required to be continuous.*

Remark 2.39
Under Novikov's condition it holds that:

$$\int_0^t ||\gamma(s)||^2\mathrm{d}s < \infty, \ \mathbb{P}\text{-a.s., for all } t \in [0,T],$$

so that the stochastic integral in (2.8) is well-defined. For a detailed explanation see [102], p. 34.

Theorem 2.40 (Girsanov theorem, p. 199, Theorem 5.7.1)
Let processes W, γ and L be as in Lemma 2.38. Define the equivalent probability measure $\mathcal{Q}$ on $(\Omega, \mathscr{F}_T)$ by:

$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathbb{P}} = L(T), \text{ i.e. } \mathcal{Q}(A) = \mathbb{E}[\mathbb{1}_A\, L(T)] = \int_A L(T)\mathrm{d}\mathbb{P}, \forall A \in \mathscr{F}_T.$$

Then, process $\tilde{W} = \{\tilde{W}(t)\}_{t\in[0,T]} = \{(\tilde{W}_1, \dots, \tilde{W}_d)'\}_{t\in[0,T]}$ given by:

$$\tilde{W}_i(t) = W_i(t) + \int_0^t \gamma_i(s)\mathrm{d}s,$$

for $t \in [0, T]$ and $i = 1, \dots, d$, is a $\mathcal{Q}$-Brownian motion.
Process L is called the density or the change of measure from $\mathbb{P}$ to $\mathcal{Q}$.

2.4.1 CIR Process

In what follows we consider an important example for an Itô process: the so-called Cox-Ingersoll-Ross (CIR) process. We will summarize some of its properties as we will need them later on.

Example 2.41 (CIR process)
The CIR process is given by the following SDE:

$$\mathrm{d}X(t) = \kappa(\theta - X(t))\mathrm{d}t + \chi\sqrt{X(t)}\mathrm{d}W(t),$$

where κ, θ and Σ are positive real constants and W is a one-dimensional Brownian motion. If $X(0) \geq 0$, then $X = \{X(t)\}_{t\in[0,\infty)}$ remains positive, provided that $\theta > \chi^2/2\kappa$ (see [49]).
Although there exists no explicit formula for $X = \{X(t)\}_{t\in[0,\infty)}$ in terms of $W = \{W(t)\}_{t\in[0,\infty)}$, a formula for the transition probability density $p(X(s) = y|X(t) = x)$ can be found. More details are provided by [32].
As shown by [33], the expectation, variance and covariance functions are given by:

$$\begin{aligned}
\mathbb{E}[X(t)] =& \theta\big(1 - e^{-\kappa t}\big) + X(0)e^{-\kappa t}\\
\mathrm{Var}(X(t)) =& \theta\frac{\chi^2}{2\kappa}\big(1 - e^{-\kappa t}\big)^2 + X(0)\frac{\chi^2}{\kappa}\big(e^{-\kappa t} - e^{-2\kappa t}\big)\\
\mathrm{Cov}(X(s), X(t)) =& \theta\frac{\chi^2}{2\kappa}\big(e^{-\kappa(t-s)} - 2e^{-\kappa t} + e^{-\kappa(t+s)}\big) + X(0)\frac{\chi^2}{\kappa}\big(e^{-\kappa t} - e^{-\kappa(t+s)}\big), s \leq t.
\end{aligned}$$

Note that

$$\lim_{t\to\infty} \mathbb{E}[X(t)] = \lim_{t\to\infty} \Big(\theta\big(1 - e^{-\kappa t}\big) + X(0)e^{-\kappa t}\Big) = \theta,$$

i.e. the CIR process is mean-reverting. The invariant distribution of the CIR process is the Gamma distribution with expectation θ and variance $\frac{\chi^2\theta}{2\kappa}$ (see [34]).

Now we state a result on the expectations of exponentials of the CIR process. It is cited from [74], Proposition 5.1.

Lemma 2.42
Let X be a CIR process as defined in Example 2.41. For $\beta \leq \frac{\kappa^2}{2\chi^2}$ and $\alpha \leq \frac{\kappa+a}{\chi^2}$, where $a := \sqrt{\kappa^2 - 2\beta\chi^2}$, the following function is well-defined:

$$\varphi^{\alpha,\beta}(t,T,x) := \mathbb{E}\Big[\exp\Big\{\alpha X(T) + \beta \int_t^T X(s)\mathrm{d}s\Big\}\Big|X(t) = x\Big].$$

More precisely, it is given by:

$$\varphi^{\alpha,\beta}(t,T,x) = \exp\big\{A^{\alpha,\beta}(T-t) + B^{\alpha,\beta}(T-t)x\big\},$$

where for a fixed $T > 0$ functions $A^{\alpha,\beta}(\tau)$ and $B^{\alpha,\beta}(\tau)$ are real-valued, continuously differentiable on $[0,T]$ and satisfy the following system of ordinary differential equations (ODEs):

$$\begin{aligned}
&- B^{\alpha,\beta}_{\tau}(\tau) + \frac{1}{2}\chi^2\big(B^{\alpha,\beta}(\tau)\big)^2 - \kappa B^{\alpha,\beta}(\tau) + \beta = 0, B^{\alpha,\beta}(0) = \alpha \\
&- A^{\alpha,\beta}_{\tau}(\tau) + \kappa\theta B^{\alpha,\beta}(\tau) = 0, A^{\alpha,\beta}(0) = 0.
\end{aligned}$$

For $\beta < \frac{\kappa^2}{2\chi^2}$ and $\alpha < \frac{\kappa+a}{\chi^2}$ they are given by:

$$\begin{aligned}
A^{\alpha,\beta}(\tau) &= \frac{\kappa\theta(\kappa-a)}{\chi^2}\tau - \frac{2\kappa\theta}{\chi^2}\ln\Big\{\frac{1-c\exp(-a\tau)}{1-c}\Big\} \\
B^{\alpha,\beta}(\tau) &= \frac{-c(\kappa+a)\exp(-a\tau) + \kappa - a}{\chi^2\big(1-c\exp(-a\tau)\big)},
\end{aligned}$$

where

$$c := \frac{-\alpha\chi^2 + \kappa - a}{-\alpha\chi^2 + \kappa + a}.$$

For $\beta \leq \frac{\kappa^2}{2\chi^2}$ and $\alpha = \frac{\kappa+a}{\chi^2}$ we obtain:

$$\begin{aligned}
A^{\alpha,\beta}(\tau) &= \kappa\theta\frac{\kappa+a}{\chi^2}\tau \\
B^{\alpha,\beta}(\tau) &= \frac{\kappa+a}{\chi^2}.
\end{aligned}$$

In the following lemma we summarize some properties of functions $A^{\alpha,\beta}$ and $B^{\alpha,\beta}$. The proof is given in Appendix A

Lemma 2.43
Consider the notation from Lemma 2.42 and assume that $\beta < \frac{\kappa^2}{2\chi^2}$ and $\alpha < \frac{\kappa+a}{\chi^2}$. Then it holds that:

i) $B^{\alpha,\beta}(\tau)$ *is monotone in* τ.

ii) $\lim_{\tau\downarrow 0} B^{\alpha,\beta}(\tau) = \alpha$.

iii) $\lim_{\tau\uparrow\infty} B^{\alpha,\beta}(\tau) = \frac{\kappa-a}{\chi^2} \begin{cases} < 0, & \text{for } \beta < 0 \\ = 0, & \text{for } \beta = 0 \\ > 0, & \text{for } \beta > 0 \end{cases}.$

iv) $\frac{\partial}{\partial\tau} A^{\alpha,\beta}(\tau) \begin{cases} \leq 0, & \text{for } \alpha \leq 0 \text{ and } \beta < 0 \\ \geq 0, & \text{for } \alpha \geq 0 \text{ and } \beta > 0 \end{cases}.$

v) *Let* $\beta \geq 0$ *and* $\alpha \geq 0$. *Then*

$$A^{\alpha,\beta}(\tau) \in \big[-2\frac{\kappa\theta}{\chi^2}\ln\{1+T\kappa\}, 3\frac{\kappa^2\theta T}{\chi^2}\big],$$

for all $\tau \in [0,T]$.

vi) *Let* $a \neq 0$. *For all* $c_2 \in (\frac{\kappa-a}{\chi^2}, \frac{\kappa+a}{\chi^2})$ *it holds: if* $\alpha < c_2$, *then* $B^{\alpha,\beta}(\tau) < c_2$.

As a trivial application of Lemma 2.42 we obtain a result on the solution of the so-called Riccati differential equations. We state it explicitly in the next corollary in the form we will apply in later on.

Corollary 2.44
Consider the following Riccati differential equation:

$$B_t(t) + \frac{1}{2}\chi^2 B^2(t) - \kappa B(t) + \beta = 0, B(T) = \alpha,$$

for some constants $\kappa > 0$, $\chi \neq 0$, β *and* α. *It is solvable if* $\beta \leq \frac{\kappa^2}{2\chi^2}$ *and* $\alpha \leq \frac{\kappa+a}{\chi^2}$, *where* $a = \sqrt{\kappa^2 - 2\beta\chi^2}$. *More precisely, for* $\beta < \frac{\kappa^2}{2\chi^2}$ *and* $\alpha < \frac{\kappa+a}{\chi^2}$, *its solution is given by:*

$$B(t) = \frac{-c(\kappa+a)\exp\{-a(T-t)\} + \kappa - a}{\chi^2\big(1 - c\exp\{-a(T-t)\}\big)},$$

where:

$$c := \frac{-\alpha\chi^2 + \kappa - a}{-\alpha\chi^2 + \kappa + a}.$$

For $\beta \leq \frac{\kappa^2}{2\chi^2}$ *and* $\alpha = \frac{\kappa+a}{\chi^2}$, *we obtain:*

$$B(t) = \frac{\kappa+a}{\chi^2}.$$

Remark 2.45 (Half-life of a mean-reverting process)
As already mentioned, the CIR process is a mean-reverting process with mean-reverting level θ. The speed of the mean-reversion is driven by the parameter κ and can be described by the concept of the so-called half-life of the process. This is the average time needed for the process to halve the distance to θ. In order to derive this formally we approximate the increments of the process by their expectation:

$$\mathrm{d}X(t) \doteq \mathbb{E}[\mathrm{d}X(t)] = \kappa(\theta - X(t))\mathrm{d}t.$$

This, together with the initial value $X(0)$ leads to the following expression for $X(t)$:

$$X(t) \doteq \theta\big(1 - e^{-\kappa t}\big) + X(0)e^{-\kappa t}.$$

Now we want to find this t_1, such that:

$$X(t_1) - \theta = 0.5\big(X(0) - \theta\big).$$

It follows that:

$$t_1 = \frac{\ln 2}{\kappa}.$$

So, the expected time that the process needs to revert half the way back to θ is inversely proportional to κ. The quantity $\frac{\ln 2}{\kappa}$ is called the half-life of the process.

2.5 Basic continuous financial models

In this section we introduce some general notation concerning financial models in continuous-time and provide a very brief overview over the definitions of some basic continuous financial models used later on in the thesis. For further details we refer to [12] and to the literature cited for each example.
All definitions are stated on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$, where measure $\mathbb{P}$ is called the real-world measure. We consider a bank account with price process P_0 and N traded risky assets. Their price processes are summarized in the N-dimensional process P. Furthermore, we define the vector of discounted price processes $\tilde{P}$ by:

$$\tilde{P} := \big(\frac{P_1}{P_0}, \ldots, \frac{P_N}{P_0}\big).$$

The log-return of asset P_i over the period from t_0 to t is defined as $\ln \frac{P_i(t)}{P_i(t_0)}$, for all $0 \le t_0 \le t$ and all $i = 1, \ldots, N$. It is often useful to consider the market under another equivalent measure, the so-called risk-neutral measure with some favorable properties. It is precisely defined in the following definition.

Definition 2.46 (Risk-neutral measure)
A probability measure $\mathcal{Q}$ on $(\Omega, \mathscr{F})$ is called a risk-neutral probability measure (or equivalent martingale measure) if it is equivalent to $\mathbb{P}$ and the discounted price process $\tilde{P}$ is a $\mathcal{Q}$-local martingale.

Now we continue with the brief definition of some basic financial models where the price processes are modeled by continuous stochastic processes.

Example 2.47 (Black-Scholes model, [16])
The Black-Scholes model consists of a bank account P_0 and one risky asset with price process P_1. The dynamics under the real-world measure $\mathbb{P}$ are defined as follows:

$$\begin{aligned} \mathrm{d}P_0(t) &= P_0(t) r \mathrm{d}t \\ \mathrm{d}P_1(t) &= P_1(t)\big\{\mu \mathrm{d}t + \Sigma \mathrm{d}W(t)\big\}, \end{aligned}$$

where W is a one-dimensional Brownian motion and r, μ, Σ are constants with $\Sigma \neq 0$.
Application of Itô's formula to $\ln(P_1)$ yields the solution for P_1:

$$P_1(t) = P_1(t_0) \exp\big\{\big(\mu - \frac{1}{2}\Sigma^2\big)(t - t_0) + \Sigma\big(W(t) - W(t_0)\big)\big\},$$

for any $0 \le t_0 \le t$. Thus, the log-returns of the risky asset are normally distributed:

$$\ln \frac{P_1(t)}{P_1(t_0)} \sim \mathscr{N}\Big(\big(\mu - \frac{1}{2}\Sigma^2\big)(t - t_0), \Sigma^2(t - t_0)\Big).$$

The classical Black-Scholes model can be extended to N traded risky assets in a straightforward manner if we take $\mu = (\mu_1, \ldots, \mu_N)' \in \mathbb{R}^N$, $\Sigma = \begin{pmatrix} \Sigma_{1,1} & \ldots & \Sigma_{1,N} \\ \vdots & \ddots & \vdots \\ \Sigma_{N,1} & \ldots & \Sigma_{N,N} \end{pmatrix} \in \mathbb{R}^{N,N}$ invertible and let $W = (W_1, \ldots, W_N)'$ be an N-dimensional Brownian motion:

$$\begin{aligned} \mathrm{d}P_0(t) &= P_0(t) r \mathrm{d}t \\ \mathrm{d}P_n(t) &= P_n(t)\big\{\mu_n \mathrm{d}t + \sum_{m=1}^{N} \Sigma_{n,m} \mathrm{d}W_m(t)\big\}, n = 1, \ldots, N. \end{aligned}$$

Analogously as before it holds:

$$\Big(\ln \frac{P_1(t)}{P_1(t_0)}, \ldots, \ln \frac{P_N(t)}{P_N(t_0)}\Big)' \sim \mathscr{N}_N\Big(\big(\mu - \frac{1}{2} diag(\Sigma\Sigma')\big)(t - t_0), \Sigma\Sigma'(t - t_0)\Big).$$

Example 2.48 (Vasicek model, [99])
The Vasicek model proposes stochastic dynamics for the short rate. One assumes that the short rate develops according to the following SDE under the real-world measure:

$$\mathrm{d}X(t) = \kappa(\theta - X(t))\mathrm{d}t + \chi \mathrm{d}W^X(t), \tag{2.9}$$

where W^X is a Brownian motion. Furthermore, it is assumed that the density for the change of measure form the real-world measure to the risk-neutral measure $\mathcal{Q}$ is given by:

$$L(t) := \exp\Big(\int_0^t \lambda \mathrm{d}W(s) - \frac{1}{2}\int_0^t \lambda^2 \mathrm{d}s\Big),$$

for a constant λ. Then, the dynamics of X under the risk-neutral measure are:

$$\mathrm{d}X(t) = \kappa\Big(\underbrace{\theta + \frac{\chi\lambda}{\kappa}}_{=:\tilde{\theta}} - X(t)\Big)\mathrm{d}t + \chi \mathrm{d}\tilde{W}^X(t),$$

where $\tilde{W}^X$ is a $\mathcal{Q}$-Brownian motion. The price at time point t of a zero-coupon bond with maturity T is:

$$P_1(t) = \mathbb{E}_{\mathcal{Q}}\Big[\exp\Big\{-\int_t^T X(s)\mathrm{d}s\Big\}\Big|\mathscr{F}_t^X\Big] = \exp\big\{-A_1(T-t) - A_2(T-t)X(t)\big\},$$

where for $\tau \in [0, T]$:

$$\begin{aligned} A_1(\tau) &= \big(\tilde{\theta} - \frac{\chi^2}{2\kappa^2}\big)\big(\tau - A_2(\tau)\big) + \frac{\chi^2}{4\kappa}A_2^2(\tau) \\ A_2(\tau) &= \frac{1}{\kappa}\big(1 - \exp\{-\kappa\tau\}\big). \end{aligned}$$

Note that Remark 2.45 can be applied to process (2.9) as well, so that its half-life is given as $\frac{\ln 2}{\kappa}$.

Example 2.49 (Heston model, [58])
The Heston model extends the Black-Scholes framework by stochastic volatility. More precisely, we have:

$$\begin{aligned} \mathrm{d}P_0(t) &= P_0(t) r \mathrm{d}t \\ \mathrm{d}P_1(t) &= P_1(t)\big\{\big(r + \lambda X(t)\big)\mathrm{d}t + \sqrt{X(t)}\mathrm{d}W^P(t)\big\} \\ \mathrm{d}X(t) &= \kappa(\theta - X(t))\mathrm{d}t + \chi\sqrt{X(t)}\mathrm{d}W^X(t) \\ \mathrm{d}\langle W^X, W^P\rangle(t) &= \rho \mathrm{d}t, \end{aligned}$$

where W^P and W^X are one-dimensional Brownian motions with constant correlation $\rho \in (-1, 1)$ and $r, \lambda, \kappa, \theta, \chi$ are constants with $\kappa, \theta, \chi > 0$ and $\theta > \frac{\chi^2}{2\kappa}$. So, the risky asset exhibits stochastic volatility modeled by a CIR process. We introduced the CIR process in more detail in Section 2.4.1.

2.6 Markov chains

Now we continue with an example of semimartingales with jumps: Markov chains. They are one of the basic ingredients for our further derivations, that is why we pay some special attention to them. First, we summarize some basic definitions and properties concerning Markov chains and then we show how to embed them in the general semimartingale framework. Again, all definitions are considered on a filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, \mathbb{P})$.

2.6.1 Basic definitions and properties

Definition 2.50 (Markov process, [66], p. 74, Definition 5.10)
For $d \in \mathbb{N}\backslash\{0\}$ consider an adapted d-dimensional process $X = \{X(t)\}_{t\in[0,\infty)}$. Let ν be a probability measure on $\big(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d)\big)$. X is said to be a Markov process with initial distribution ν if:

i) $\mathbb{P}(X(0) \in B) = \nu(B), \forall B \in \mathscr{B}(\mathbb{R}^d)$,

ii) for $t \geq s \geq 0$ and $B \in \mathscr{B}(\mathbb{R}^d)$,

$$\mathbb{P}\big(X(t) \in B|\mathscr{F}_s\big) = \mathbb{P}\big(X(t) \in B|X(s)\big),\ \mathbb{P} - a.s.$$

Definition 2.51 (Jump process, [101], p. 16, Section 2.3)
A jump process is a right-continuous stochastic process with piece-wise constant sample paths.

Definition 2.52 (Continuous-time Markov chain, [101], p. 16)
Let $MC := \{MC(t)\}_{t\geq 0}$ denote a jump process that takes values in some finite or countable space E. For convenience, the space E is set to either $E = \{1, 2, \dots, I\}$ or $E = \{1, 2, \dots\}$. Then MC is called a continuous-time Markov chain with state space E if and only if it is a Markov process, i.e.

$$\mathbb{P}\big(MC(t) = i|\mathscr{F}_s\big) = \mathbb{P}\big(MC(t) = i|MC(s)\big), t \geq s \geq 0. \tag{2.10}$$

Further, denote the jump times of MC by $\{t_n\}_{n\in\mathbb{N}}$:

$$t_0 = 0, t_n := \inf\{s > t_{n-1}|MC(s) \neq MC^-(s)\}.$$

The random variable $T_i(t) := \int_0^t \mathbb{1}_{\{MC(s)=i\}}\mathrm{d}s, i \in E, t \geq 0$ is called the occupation time of MC in state i up to time t. We denote $T(t) := (T_1(t), \dots, T_I(t))'$.
For the whole thesis we use the convention that Markov chains are càdlàg processes.

Definition 2.53 (Transition matrix, [101], p. 17)
For any $i, j \in E$ and $t \geq s \geq 0$ denote:

$$p_{i,j}(t, s) := \mathbb{P}\big(MC(t) = j|MC(s) = i\big),$$

and $P(t, s) := \{p_{i,j}(t, s)\}_{i,j\in E}$. We name $P(t, s)$ the transition matrix of the Markov chain MC, and postulate that:

$$\lim_{t\downarrow s} P(t, s) = \mathbb{1}, \tag{2.11}$$

where $\mathbb{1}$ denotes the identity matrix.

Proposition 2.54 (Properties, [101], p. 17)
It holds for $t \geq \tau \geq s \geq 0$:

i) $p_{i,j}(t,s) \geq 0$, *for all* $i,j \in E$.

ii) $\sum_{j\in E} p_{i,j}(t,s) = 1$, *for all* $i \in E$.

iii) $p_{i,j}(t,s) = \sum_{k\in E} p_{i,k}(\tau,s) p_{k,j}(t,\tau)$, *for all* $i,j \in E$.

The last identity is usually referred to as the Chapman-Kolmogorov equation.

Definition 2.55 (Stationarity, [101], p. 17)
If the transition probability $\mathbb{P}\big(MC(t) = i | MC(s)\big)$ *depends only on* $t - s$, *then the Markov chain* MC *is stationary. In this case we define*

$$p_{i,j}(h) := p_{i,j}(s+h, s)$$

and $P(h) := \{p_{i,j}(h)\}_{i,j\in E}$ *for any* $h \geq 0$. *The process is non-stationary otherwise.*

Definition 2.56 (q-Property, [101], p. 17, Definition 2.1)
Denote by $Q(t) = \{q_{i,j}(t)\}_{i,j\in E}$ *a matrix-valued function in* t, *for* $t \geq 0$. *It satisfies the q-Property if:*

i) $q_{i,j}$ *is* $\mathscr{B}(\mathbb{R})$*-measurable for all* $i,j \in E$,

ii) $q_{i,j}(t)$ *is uniformly bounded, that is, there exists a constant* K *such that* $|q_{i,j}(t)| \leq K$, *for all* $i,j \in E$ *and* $t \geq 0$,

iii) $q_{i,j}(t) \geq 0$ *for* $i \neq j$ *and* $\sum_{j\in E} q_{i,j}(t) = 0$, *for all* $i \in E$.

Definition 2.57 (Generator matrix, [101], p. 17, Definition 2.2)
A matrix-valued function $Q(t) = \{q_{i,j}(t)\}_{i,j\in E}$, *defined for* $t \geq 0$ *is a generator of the Markov chain* MC *if it satisfies the q-Property and for all bounded, real-valued functions* f *on* E *the following process*

$$f\big(MC(t)\big) - \int_0^t \sum_{j\in E} q_{MC(s),j}(s) f(j) \mathrm{d}s$$

is a martingale w.r.t the filtration generated by MC.

Lemma 2.58 (Generator matrix, [101], p. 18, Lemma 2.4)
Let $E = \{1, \ldots, I\}$. *Then process* MC *with values in* $\mathcal{E}$ *is a Markov chain with state space* E *and generator* $Q(t)$ *if and only if*

$$\begin{pmatrix} \mathbb{1}_{\{MC(t)=1\}} \\ \vdots \\ \mathbb{1}_{\{MC(t)=I\}} \end{pmatrix} - \int_0^t Q'(s) \begin{pmatrix} \mathbb{1}_{\{MC(t)=1\}} \\ \vdots \\ \mathbb{1}_{\{MC(t)=I\}} \end{pmatrix} \mathrm{d}s$$

is a martingale.

Theorem 2.59 (Generated Markov chain, [101], p. 19, Theorem 2.5)
Suppose that the matrix $Q(t) = \{q_{i,j}(t)\}_{i,j \in E}$ satisfies the q-Property for $t \geq 0$. Then, there exists a Markov chain $MC = \{MC(t)\}_{t \in [0,\infty)}$ with state space E and generator matrix $Q(t)$.

i) Its transition matrix $P(t, s)$ satisfies the following forward differential equation:

$$\frac{\partial}{\partial t} P(t,s) = P(t,s)Q(t) \text{ for } t \geq s, P(s,s) = \mathbb{1}. \tag{2.12}$$

If we further assume that $Q(t)$ is continuous in t then it satisfies also the following backwards differential equation:

$$\frac{\partial}{\partial s} P(t,s) = Q(s)P(t,s) \text{ for } t \geq s, P(s,s) = \mathbb{1}. \tag{2.13}$$

These two differential equations are referred to as forward and backward Kolmogorov equations, respectively.

ii) If we again assume that $Q(t)$ is continuous in t then

$$\lim_{\Delta t \downarrow 0} \frac{p_{i,j}(t + \Delta t, t) - \delta_{i,j}}{\Delta t} = q_{i,j}(t),$$

where $\delta_{i,j} = \mathbb{1}_{\{i=j\}}$ denotes the Kronecker symbol.

From now on we will consider stationary continuous-time Markov chains with a finite state space as this is the relevant case in most financial applications. As a direct application of Theorem 2.59 and Theorem 2.1.1 from [85] one obtains the following corollary:

Corollary 2.60 (Stationary generator matrices)
Assume that the constant matrix $Q = \{q_{i,j}\}_{ij \in E}$ satisfies:

i) for $i \neq j$ it holds $q_{i,j} \geq 0$,

ii) $\sum_{j \in E} q_{i,j} = 0$, for all $i \in E$.

Then, it is the generator matrix of a stationary Markov chain $MC = \{MC(t)\}_{t \in [0,\infty)}$ with state space E and transition matrix for $h \geq 0$:

$$P(h) = \exp\{hQ\} = \sum_{k=0}^{\infty} \frac{h^k}{k!} Q^k.$$

Theorem 2.61 (Existence of a generator matrix, [2])
Let $\{P(h)\}_{h \geq 0}$ be the family of transition matrices of the Markov chain MC with state space E. Then the following holds:

i) The limit $q_{i,i} := \lim_{h\downarrow 0} \frac{p_{i,i}(h)-1}{h}$ *exists for all* $i \in E$.

Assume further that $q_{i,i} > -\infty$ *for all* $i \in E$. *Then it holds that:*

ii) The limits $q_{i,j} := \lim_{h\downarrow 0} \frac{p_{i,j}(h)-\delta_{i,j}}{h}$, *for all* $i, j \in E$ *exist and are finite.*

iii) For $i \neq j$, $q_{i,j} \geq 0$. *Further,* $\sum_{j\in E} q_{i,j} = 0$, *for all* $i \in E$.

iv) Seen as a function, $\{P(h)\}_{h\geq 0}$ *solves the following differential equation*

$$\frac{d}{dh}P(h) = P(h)Q, \text{ for } h \geq 0$$
$$P(0) = \mathbb{1},$$

where $Q := \{q_{i,j}\}_{i,j\in E}$. *Thus,* $P(h) = \exp\{hQ\} = \sum_{k=0}^{\infty} \frac{h^k}{k!}Q^k, h \geq 0$.

Proof
Statement i): [2], p. 9, Proposition 2.2.
Statement ii): [2], p. 10, Proposition 2.4.
Statement iii): Follows directly from statement (ii).
Statement iv): [2], p. 13, Proposition 2.7.

□

Remark 2.62
Applying Corollary 2.60, one can identify matrix Q *from Theorem 2.61 with the generator matrix of the Markov chain* MC. *Thus, a Markov chain can be uniquely characterized by its initial value resp. initial distribution* $\{\mathbb{P}(MC(0) = i)\}_{i\in E}$ *and either its generator matrix* Q *or the family of transition matrices* $\{P(h)\}_{h\geq 0}$.

Proposition 2.63 (Distribution properties, [2], p. 16, Proposition 2.8)
The following holds:

$$\mathbb{P}\big(t_1 > t | MC(0) = i\big) = \exp\{q_{i,i}t\} \tag{2.14}$$

$$\mathbb{P}\big(MC(t_1) = j | MC(0) = i\big) = \begin{cases} (1-\delta_{i,j})\frac{q_{i,j}}{-q_{i,i}}, & \text{if } -q_{i,i} > 0 \\ \delta_{i,j}, & \text{if } -q_{i,i} = 0 \end{cases}. \tag{2.15}$$

Recall that t_1 *denotes the first jump time of* MC.

This proposition leads directly to the following corollary:

Corollary 2.64 (Embedded Markov chain / jump chain)
The discrete time process $\{\overline{MC}(n)\}_{n\in\mathbb{N}}$ *defined via* $\overline{MC}(n) := MC(t_n)$ *is a discrete-time Markov chain, i.e. it is a Markov process in discrete time with a finite or countable state space. More precisely, its state space is equal to the state space of* MC, $E = \{1, 2, \ldots, I\}$ *and the transition probabilities are given by*

$$r_{i,j} := \mathbb{P}\big(\overline{MC}(n+1) = j | \overline{MC}(n) = i\big) = \begin{cases} (1-\delta_{i,j})\frac{q_{i,j}}{-q_{i,i}}, & \text{if } -q_{i,i} > 0 \\ \delta_{i,j}, & \text{if } -q_{i,i} = 0 \end{cases}. \tag{2.16}$$

2.6.2 Semimartingale characterization

Now we show that Markov chains are semimartingales and apply some known results from the semimartingale theory to Markov chains.

Remark 2.65
Without loss of generality one can identify a Markov chain MC with state space $E = \{1, 2, \dots, I\}$ and transition matrices $\{P(h)\}_{h\geq 0}$ by a Markov chain $\mathcal{MC}$ with state space $\mathcal{E} = \{e_1, e_2, \dots, e_I\}$ and the same transition probabilities:

$$\mathbb{P}\big(\mathcal{MC}(t) = e_j | \mathcal{MC}(s) = e_i\big) = p_{i,j}(t,s), t \geq s \geq 0, i, j \in E,$$

where $e_i = (0, \dots, 0, 1, 0, \dots, 0)'$ denotes the i^{th} unit vector in $\mathbb{R}^I$. More precisely, one defines:

$$\mathcal{MC}(t) := \begin{pmatrix} \mathbb{1}_{\{MC(t)=1\}} \\ \vdots \\ \mathbb{1}_{\{MC(t)=I\}} \end{pmatrix} \in \mathbb{R}^I. \tag{2.17}$$

Naturally, the jump times and occupation times for both Markov chains coincide.

Notation
From now on throughout the whole study we denote by MC a Markov chain with state space $E = \{1, 2, \dots, I\}$, transition matrices $\{P(h)\}_{h\geq 0}$ and generator matrix Q. $\mathcal{MC}$ stays for a Markov chain with state space $\mathcal{E} = \{e_1, e_2, \dots, e_I\}$, transition matrices $\{P(h)\}_{h\geq 0}$ and generator matrix Q. The corresponding embedded Markov chains are denoted by $\overline{MC}$ and $\overline{\mathcal{MC}}$, respectively. The jump times for both Markov chains are denoted by $\{t_n\}_{n\in\mathbb{N}}$ and $T(t)$ stays for the occupation times. The natural filtration of process $\mathcal{MC}$ is denoted by $\mathbb{F}^{\mathcal{MC}} = \{\mathcal{F}_t^{\mathcal{MC}}\}_{t\in[0,\infty)}$.

Corollary 2.66 (Semimartingale decomposition)
By Corollary 2.58 it follows that the Markov chain $\mathcal{MC}$ can be written in the following form:

$$\mathcal{MC}(t) = \mathcal{MC}(0) + \int_0^t Q'\mathcal{MC}(s)\mathrm{d}s + M(s), \tag{2.18}$$

where M is a martingale w.r.t. the filtration generated by $\mathcal{MC}$. Observe that according to Proposition 2.18 process $\int_0^t Q'\mathcal{MC}(s)\mathrm{d}s$ is of finite variation. Thus, Equation (2.18) is a semimartingale representation of $\mathcal{MC}$.

Using the notation about semimartingales from Section 2.3 we can state the following corollary:

Corollary 2.67
It holds that $\mathcal{MC}^c = M^c = 0$.

Proof
Observe that M is of finite variation, as the difference of two processes with finite variation:

$$M(s) = \underbrace{\mathcal{MC}(t) - \mathcal{MC}(0)}_{\in \mathscr{V}} - \underbrace{\int_0^t Q'\mathcal{MC}(s)\mathrm{d}s}_{\in \mathscr{V}}.$$

Thus, with Lemma 2.23 it follows that M is purely discontinuous, i.e. $M^c = 0$. Further, from Proposition 2.26 we know that $\mathcal{MC}^c = M^c$. The statement follows.
□

Lemma 2.68 (Laplace transform of the occupation times, [38])
For any $\lambda \in \mathbb{R}^I$ it holds:

$$\begin{aligned}\phi_{T(t)}(\lambda) :&= \mathbb{E}\big[\exp\big\{\langle\lambda, T(t)\rangle\big\}\big] = \mathbb{E}\Big[\exp\Big\{\int_0^t \langle\lambda, \mathcal{MC}(s)\rangle \mathrm{d}s\Big\}\Big] \\ &= \big(\exp\big\{\big(Q' + diag(\lambda)\big)t\big\}\mathbb{E}\big[\mathcal{MC}(0)\big]\big)' \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.\end{aligned} \tag{2.19}$$

Lemma 2.69 (Quadratic variation, [39], p. 340, Lemma 1.3)
For the martingale M from Equation (2.18) it holds:

$$\begin{aligned}\langle M, M\rangle(t) =& diag\Big(\int_0^t Q'\mathcal{MC}(s)\mathrm{d}s\Big) - \int_0^t \big(diag(\mathcal{MC}(s))\big)Q\mathrm{d}s \\ &- \int_0^t \big(Q' diag(\mathcal{MC}(s))\big)\mathrm{d}s =: \int_0^t \{v_{i,j}(s)\}_{i,j\in E}\mathrm{d}s.\end{aligned} \tag{2.20}$$

Now we define the class of processes used in this thesis. This class extends Itô processes by Markov switching.

Definition 2.70 (Markov-modulated Itô diffusion)
Let $X := \{X(t)\}_{t\geq 0} =: \big\{\big(X_1(t), \dots, X_{d_1}(t)\big)'\big\}_{t\geq 0}$ be a d_1-dimensional process and $\mathcal{MC} := \{\mathcal{MC}(t)\}_{t\geq 0}$ a continuous-time Markov chain with finite state space $\mathcal{E} = \{e_1, \dots, e_I\}$. Process X is called a Markov-modulated Itô diffusion if it satisfies the

following equation:

$$\begin{aligned}X(t) &= X(0) + \int_0^t \mu\big(s, X(s), \mathcal{MC}(s)\big)\mathrm{d}s + \int_0^t \Sigma\big(s, X(s), \mathcal{MC}(s)\big)\mathrm{d}W(s) \qquad (2.21)\\ &= \begin{pmatrix} X_1(0) \\ \vdots \\ X_{d_1}(0) \end{pmatrix} + \int_0^t \begin{pmatrix} \mu_1\big(s, X(s), \mathcal{MC}(s)\big) \\ \vdots \\ \mu_{d_1}\big(s, X(s), \mathcal{MC}(s)\big) \end{pmatrix} \mathrm{d}s \\ &+ \int_0^t \begin{pmatrix} \Sigma_{1,1}\big(s, X(s), \mathcal{MC}(s)\big) & \dots & \Sigma_{1,d_2}\big(s, X(s), \mathcal{MC}(s)\big) \\ \vdots & \vdots & \vdots \\ \Sigma_{d_1,1}\big(s, X(s), \mathcal{MC}(s)\big) & \dots & \Sigma_{d_1,d_2}\big(s, X(s), \mathcal{MC}(s)\big) \end{pmatrix} \mathrm{d} \begin{pmatrix} W_1(s) \\ \vdots \\ W_{d_2}(s) \end{pmatrix},\end{aligned}$$

where $X(0) = (X_1(0), \dots, X_{d_1}(0))'$ is a d_1-dimensional $\mathscr{F}_0$-measurable random variable, $W := \{W(t)\}_{t\geq 0}$ is a d_2-dimensional Brownian motion and $\mu : [0,\infty) \times \mathbb{R}^{d_1} \times \mathcal{E} \mapsto \mathbb{R}^{d_1}$ and $\Sigma : [0,\infty) \times \mathbb{R}^{d_1} \times \mathcal{E} \mapsto \mathbb{R}^{d_1,d_2}$ are deterministic measurable functions such that:

$$\int_0^t \big|\mu_i\big(s, X(s), \mathcal{MC}(s)\big)\big|\mathrm{d}s < \infty, i = 1, \dots, d_1$$
$$\int_0^t \mathbb{E}\big[\big\{\Sigma_{i,j}\big(s, X(s), \mathcal{MC}(s)\big)\big\}^2\big]\mathrm{d}s < \infty, i = 1, \dots, d_1, j = 1, \dots, d_2.$$

Observe that X is a continuous process as we only have integrals w.r.t. continuous processes.

Remark 2.71 (Existence of Markov-modulated Itô diffusions)
If one assumes that the following process exists for all $i \in E$:

$$X^{(i)}(t) = X(0) + \int_0^t \mu\big(s, X(s), e_i\big)\mathrm{d}s + \int_0^t \Sigma\big(s, X(s), e_i\big)\mathrm{d}W(s)$$

for any starting value $X(0) \in D^X \subset \mathbb{R}^{d_1}$, and that $X^{(i)}(t) \in D^X$, for all $t \geq 0$, then it is clear that process X defined in (2.21) exists as well, as it can be constructed step-wise from processes $X^{(1)}, \dots, X^{(I)}$.

Now we derive the most basic tool for Markov-modulated Itô diffusions: Itô's formula. It is obtained as an application of the general result for semimartingales cited in Theorem 2.32.

Theorem 2.72 (Itô's formula for Markov-modulated Itô diffusions)
Consider process X defined as in Definition 2.70. Let the Markov chain $\mathcal{MC}$ have the following semimartingale decomposition:

$$\mathcal{MC}(t) = \mathcal{MC}(0) + \int_0^t Q'\mathcal{MC}(s)\mathrm{d}s + M(s),$$

with intensity matrix $Q = \{q_{i,j}\}_{i,j \in E}$. Further, a function $f : [0,\infty) \times \mathbb{R}^n \times \mathcal{E} \to \mathbb{R}$, $\big(t, (x_1, \dots, x_n)', e_i\big) \mapsto f\big(t, (x_1, \dots, x_n)', e_i\big)$ is given, which is once continuously differentiable in the first variable and twice continuously differentiable in the second for all $e_i \in \mathcal{E}$. Then $\big\{f\big(t, X(t), \mathcal{MC}(t)\big)\big\}_{t \in [0,\infty)}$ is a semimartingale and we have:

$$\begin{aligned} f\big(t, X(t), \mathcal{MC}(t)\big) &= f(0, X(0), \mathcal{MC}(0)) \\ &+ \int_0^t f_t\big(s, X(s), \mathcal{MC}(s)\big) + \sum_{j=1}^{d_1} f_{x_j}\big(s, X(s), \mathcal{MC}(s)\big)\mu_j\big(s, X(s), \mathcal{MC}(s)\big) \\ &+ \frac{1}{2} \sum_{j,k=1}^{d_1} f_{x_j,x_k}\big(s, X(s), \mathcal{MC}(s)\big) \sum_{r=1}^{d_2} \Sigma_{j,r}\big(s, X(s), \mathcal{MC}(s)\big)\Sigma_{k,r}\big(s, X(s), \mathcal{MC}(s)\big)\mathrm{d}s \\ &+ \int_0^t \sum_{j=1}^{d_1} \Big(f_{x_j}\big(s, X(s), \mathcal{MC}(s)\big) \sum_{r=1}^{d_2} \Sigma_{j,r}\big(s, X(s), \mathcal{MC}(s)\big)\mathrm{d}W_r(s)\Big) \\ &+ \int_0^t \sum_{i=1}^{I} f\big(s, X(s), e_i\big) q_{MC(s),i}\mathrm{d}s + \int_0^t \sum_{i=1}^{I} f(s, X(s), e_i)\mathrm{d}M_i(s). \end{aligned} \tag{2.22}$$

Proof

The statement is an application of the general Itô formula for semimartingales (see Theorem 2.32). Observe that function f is defined in the third variable only on the discrete space $\mathcal{E}$. So, in order to be able to apply Theorem 2.32, we need to extend function f to some open set on which we can define derivatives and assure that the extended function is twice continuously differentiable. This is possible, e.g. via polynomials. As we will see in the following computation, it does not make any difference how this is done, so we do not consider the issue further. Denote by f_m the derivative of this extended function w.r.t. the third variable. Recall from Corollary 2.67 that $\mathcal{MC}^c = 0$. Substitution in Equation (2.3) leads to:

$$\begin{aligned} f\big(t, X(t), \mathcal{MC}(t)\big) &= f(0, X(0), \mathcal{MC}(0)) \\ &+ \int_0^t \Big[f_t\big(s, X(s), \mathcal{MC}(s)\big) + \sum_{j=1}^{d_1} f_{x_j}\big(s, X(s), \mathcal{MC}(s)\big)\mu_j\big(s, X(s), \mathcal{MC}(s)\big) \\ &+ \frac{1}{2} \sum_{j,k=1}^{d_1} f_{x_j,x_k}\big(s, X(s), \mathcal{MC}(s)\big) \sum_{r=1}^{d_2} \Sigma_{j,r}\big(s, X(s), \mathcal{MC}(s)\big)\Sigma_{k,r}\big(s, X(s), \mathcal{MC}(s)\big)\Big]\mathrm{d}s \\ &+ \int_0^t \sum_{j=1}^{d_1} \Big(f_{x_j}\big(s, X(s), \mathcal{MC}(s)\big) \sum_{r=1}^{d_2} \Sigma_{j,r}\big(s, X(s), \mathcal{MC}(s)\big)\mathrm{d}W_r(s)\Big) \\ &+ \int_0^t f_m\big(s, X(s), \mathcal{MC}^-(s)\big)\mathrm{d}\mathcal{MC}(s) \\ &+ \sum_{0 \le s \le t} \Big[f\big(s, X(s), \mathcal{MC}(s)\big) - f\big(s, X(s), \mathcal{MC}(s-)\big)\Big] \end{aligned}$$

$$- \sum_{0 \le s \le t} f_m\big(s, X(s), \mathcal{MC}^-(s)\big) \Delta \mathcal{MC}(s).$$

As $\mathcal{MC}$ is a purely jump process, it holds that:

$$\int_0^t f_m\big(s, X(s), \mathcal{MC}^-(s)\big) \mathrm{d}\mathcal{MC}(s) = \sum_{0 \le s \le t} f_m\big(s, X(s), \mathcal{MC}^-(s)\big) \Delta \mathcal{MC}(s).$$

Further,

$$\sum_{0 \le s \le t} \left[f\big(s, X(s), \mathcal{MC}(s)\big) - f\big(s, X(s), \mathcal{MC}^-(s)\big) \right] = \int_0^t \sum_{i=1}^{I} f\big(s, X(s), e_i\big) \mathrm{d}\mathcal{MC}_i(s)$$
$$= \int_0^t \Big(f\big(s, X(s), e_1\big), \ldots, f\big(s, X(s), e_I\big) \Big) Q' \mathcal{MC}(s) \mathrm{d}s$$
$$+ \int_0^t \Big(f\big(s, X(s), e_1\big), \ldots, f\big(s, X(s), e_I\big) \Big) \mathrm{d}M(s).$$

Substituting the fact that $Q'\mathcal{MC}(s) = \begin{pmatrix} q_{MC(s),1} \\ \vdots \\ q_{MC(s),I} \end{pmatrix}$ yields the statement.

□

Remark 2.73
For notational simplicity we have formulated all statements from above without loss of generality for processes staring at time-point 0. *Naturally they can be rewritten for arbitrary starting point* $t \ge 0$.

2.7 Feynman-Kac theorem for Markov-modulated Itô processes

Another important result about Itô diffusions extended by Markov chains is the connection between the solution of a deterministic PDE and expectations of exponentials. It is presented in the following theorem. As we were not able to find in the literature a formulation of the statement under conditions applicable for our case, we extended the proof for the case without Markov switching (see e.g. [86], p. 137, Theorem 8.2.1.) to Markov-switching parameters. The detailed proof is given in Appendix A.

Theorem 2.74 (Feynman-Kac theorem with Markov switching I)
Let process X*, valued in the set* $D^X \subseteq \mathbb{R}^{d_1}$*, be a* d_1*-dimensional Markov-modulated Itô diffusion given by the following SDE:*

$$\mathrm{d}X(t) = \mu\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}t + \Sigma\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}W(s),$$

where $W = (W_1, \dots, W_{d_2})'$ *is an* d_2*-dimensional Brownian motion and* $\mu : D^X \times \mathcal{E} \to \mathbb{R}^{d_1}$ *and* $\Sigma : D^X \times \mathcal{E} \to \mathbb{R}^{d_1,d_2}$ *are deterministic functions. Consider function* $K : [0,T] \times D^X \times \mathcal{E} \to \mathbb{R}$ *and define function* $k : [0,T] \times D^X \times \mathcal{E} \to \mathbb{R}$ *via:*

$$k(t,x,e_i) := \mathbb{E}\Big[\exp\Big\{-\int_0^t K\big(t-s, X(s), \mathcal{MC}(s)\big)\mathrm{d}s\Big\}\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big].$$

For all $(t,x,e_i) \in [0,T]\times D^X \times \mathcal{E}$*, assume that function* k *is well-defined,* $k(t,x,e_i) < \infty$*, and that the following conditions hold for all* $i = 1, \dots, I$*:*

i) Function k *is twice continuously differentiable in* x*.*

ii) Process $\{N(r)\}_{r\in[0,T]}$ *defined by:*

$$\begin{aligned} N(r) &:= \frac{1}{r}\int_0^r \Big[k_x\big(t, X(s), \mathcal{MC}(s)\big)'\mu\big(X(s), \mathcal{MC}(s)\big) \\ &+ \frac{1}{2}Tr\Big(k_{xx'}\big(t, X(s), \mathcal{MC}(s)\big)\Sigma\big(X(s), \mathcal{MC}(s)\big)\Sigma'\big(X(s), \mathcal{MC}(s)\big)\Big) \\ &+ \sum_{j=1}^{I} q_{MC(s),j}k(t, X(s), e_j)\Big]\mathrm{d}s, \end{aligned}$$

converges in mean to its almost sure limit, for all $t \in [0,T]$*:*

$$\lim_{r\downarrow 0}\mathbb{E}[N(r)|X(0)=x, \mathcal{MC}(0)=e_i] = \mathbb{E}[a.s. \lim_{r\downarrow 0} N(r)|X(0)=x, \mathcal{MC}(0)=e_i].$$

iii) $\mathbb{E}\Big[\int_0^r k_x\big(t, X(s), \mathcal{MC}(s)\big)'\Sigma\big(X(s), \mathcal{MC}(s)\big)\mathrm{d}W(s)\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] = 0, \forall r \in [0,T]$*.*

iv) $\mathbb{E}\Big[\int_0^r \big(k(t, X(s), e_1), \dots, k(t, X(s), e_I)\big)\mathrm{d}M(s)\Big|X(0) = x, \mathcal{MC}(0) = e_i\Big] = 0, \forall r \in [0,T]$*.*

v) For all $r \in [0, T-t]$ *define:*

$$\begin{aligned} Z(t+r) &:= \exp\Big\{-\int_0^{t+r} K\big(t+r-s, X(s), \mathcal{MC}(s)\big)\mathrm{d}s\Big\} \\ Y(r) &:= \exp\Big\{\int_0^r K\big(t+r-s, X(s), \mathcal{MC}(s)\big)\mathrm{d}s\Big\}, \end{aligned}$$

and assume that:

$$\begin{aligned} &\lim_{r\downarrow 0}\mathbb{E}\Big[Z(t+r)\frac{Y(r)-Y(0)}{r}\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] \\ &= \mathbb{E}\Big[a.s. \lim_{r\downarrow 0} Z(t+r)\frac{Y(r)-Y(0)}{r})\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] \\ &= \mathbb{E}[Z(t)\Big|X(0)=x, \mathcal{MC}(0)=e_i]K(t,x,e_i). \end{aligned}$$

Then k is differentiable w.r.t. t and satisfies for all $(t,x) \in [0,T] \times D^X$ the following system:

$$\begin{aligned}
&-k_t(t,x,e_i) - k(t,x,e_i)K(t,x,e_i) + k_x(t,x,e_i)'\mu(x,e_i) \\
&+ \frac{1}{2} Tr\big(k_{xx'}(x,e_i)\Sigma(t,x,e_i)\Sigma(t,x,e_i)'\big) = -\sum_{j=1}^{I} q_{i,j} k(t,x,e_j) \qquad (2.23) \\
&k(0,x,e_i) = 1, \forall i \in \{1,\dots,I\}.
\end{aligned}$$

Further, conditions iii) and iv) and v) can be replaced by the following conditions, respectively:

iii)' $\mathbb{E}\Big[\int_0^r k_x\big(t,X(s),\mathcal{MC}(s)\big)'\Sigma\big(X(s),\mathcal{MC}(s)\big)\Sigma\big(X(s),\mathcal{MC}(s)\big)' \cdot k_x\big(t,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big] < \infty, \forall r \in [0,T].$

iv)' $\mathbb{E}\Big[\int_0^r \big(k(t,X(s),e_j)\big)^2 \mathrm{d}s\Big|X(0) = x, \mathcal{MC}(0) = e_i\Big] < \infty, \forall r \in [0,T],$ $\forall j = 1,\dots,I.$

v)' *$K(t,x,e_i)$ is differentiable in t, $K(t,x,e_i)$ and $K_t(t,x,e_i)$ are continuous in t and x, and*

$$\begin{aligned}
&\lim_{r\downarrow 0}\mathbb{E}\Big[Z(t+r)\frac{Y(r)-Y(0)}{r}\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big] \\
&= \mathbb{E}\Big[a.s.\lim_{r\downarrow 0} Z(t+r)\frac{Y(r)-Y(0)}{r})\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big],
\end{aligned}$$

for all $r \in [0,T-t]$. Note that due to the Dominated convergence theorem (Corollary 6.26 in [68]) a sufficient condition for the equation above to hold is the existence of an integrable bound for $Z(t+r)\frac{Y(r)-Y(0)}{r}$.

Remark 2.75
Some boundedness conditions on k, μ_X, Σ_X and K can easily replace assumptions ii), iii), iv) and v). However these functions are not bounded for important examples, such as the Heston model. That is why we keep the assumptions as general as possible. Furthermore, observe that the result without Markov switching is a special case of the theorem above, so we do not state it explicitly here.

For our applications we need the backwards formulation of the Feynman-Kac result, which follows as an application of the theorem above and is stated in the following corollary:

Corollary 2.76 (Feynman-Kac theorem with Markov switching II)
Consider processes X and $\mathcal{MC}$ as in Theorem 2.74 and some function $H : [0,T] \times D^X \times \mathcal{E} \to \mathbb{R}$. Define function $h : [0,T] \times D^X \times \mathcal{E} \mapsto \mathbb{R}$ via:

$$h(t,x,e_i) := \mathbb{E}\Big[\exp\Big\{-\int_t^T H\big(s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\Big|X(t)=x,\mathcal{MC}(t)=e_i\Big].$$

Denote $K(t,x,e_i) := H(T-t,x,e_i)$ and define function $k(t,x,e_i)$ as in Theorem 2.74. Assume that the conditions of Theorem 2.74 hold for k and K. Then, h is differentiable w.r.t. t and satisfies the following system of coupled PDEs for all $(t,x) \in [0,T] \times D^X$:

$$\begin{aligned}
&h_t(t,x,e_i) - h(t,x,e_i)H(t,x,e_i) + h_x(t,x,e_i)'\mu(x,e_i) \\
&+ \frac{1}{2}Tr\big(h_{xx'}(t,x,e_i)\Sigma(x,e_i)\Sigma(x,e_i)'\big) = -\sum_{j=1}^{I} q_{i,j}h(t,x,e_j) \\
&h(T,x,e_i) = 1, \forall i \in \{1,\ldots,I\}.
\end{aligned}$$

Proof
First recall that the stated assumptions assure that k satisfies System (2.23). Observe further that:

$$\begin{aligned}
h(t,x,e_i) &= \mathbb{E}\Big[\exp\Big\{-\int_t^T H\big(s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\Big|X(t)=x,\mathcal{MC}(t)=e_i\Big] \\
&= \mathbb{E}\Big[\exp\Big\{-\int_t^T K\big(T-s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\Big|X(t)=x,\mathcal{MC}(t)=e_i\Big] \\
&= \mathbb{E}\Big[\exp\Big\{-\int_0^{T-t} K\big(T-t-s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big] \\
&= k(T-t,x,e_i),
\end{aligned}$$

because of the Markov property of $(X,\mathcal{MC})$. It follows that:

$$\begin{aligned}
h_t(t,x,e_i) &= -k_t(T-t,x,e_i), h_x(t,x,e_i) = k_x(T-t,x,e_i) \\
h_{xx'}(t,x,e_i) &= k_{xx'}(T-t,x,e_i), h(T,x,e_i) = k(0,x,e_i) = 1.
\end{aligned}$$

Substitution of the equations above in System (2.23) shows that h satisfies the following system for all $(t,x) \in [0,T] \times D^X$:

$$\begin{aligned}
&h_t(t,x,e_i) - h(t,x,e_i)\underbrace{K(T-t,x,e_i)}_{=H(t,x,e_i)} + h_x(t,x,e_i)'\mu(x,e_i) \\
&+ \frac{1}{2}Tr\big(h_{xx'}(t,x,e_i)\Sigma(x,e_i)\Sigma(x,e_i)'\big) = -\sum_{j=1}^{I} q_{i,j}h(t,x,e_j) \\
&h(T,x,e_i) = 1, \forall i \in \{1,\ldots,I\}.
\end{aligned}$$

□

Remark 2.77
The conditions of Theorem 2.74 for k and K can be trivially rewritten in terms of h and H as follows:

i) Function h is twice continuously differentiable in x.

ii) For the process $\{\tilde{N}(r)\}_{r\in[0,T]}$ defined by

$$\begin{aligned}\tilde{N}(r) :=& \frac{1}{r}\int_0^r h_x\big(t, X(s), \mathcal{MC}(s)\big)'\mu\big(X(s), \mathcal{MC}(s)\big)\\&+\frac{1}{2}Tr\Big(h_{xx'}\big(t, X(s), \mathcal{MC}(s)\big)\Sigma\big(X(s), \mathcal{MC}(s)\big)\Sigma\big(X(s), \mathcal{MC}(s)\big)'\Big)\\&+\sum_{j=1}^{I} q_{MC(s),j}h(t, X(s), e_j)\mathrm{d}s,\end{aligned}$$

it holds that:

$$\lim_{r\downarrow 0}\mathbb{E}[\tilde{N}(r)|X(0)=x, \mathcal{MC}(0)=e_i] = \mathbb{E}[a.s. \lim_{r\downarrow 0}\tilde{N}(r)|X(0)=x, \mathcal{MC}(0)=e_i].$$

This means, it converges in mean to its almost sure limit, for all $t\in[0,T]$.

iii) $\mathbb{E}\Big[\int_0^r h_x\big(t, X(s), \mathcal{MC}(s)\big)'\Sigma\big(X(s), \mathcal{MC}(s)\big)\mathrm{d}W(s)\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] = 0$, *for all $r\in[0,T]$.*

iv) $\mathbb{E}\Big[\int_0^r \big(h(t, X(s), e_1), \dots, h(t, X(s), e_I)\big)\mathrm{d}M(s)\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] = 0$, *for all $r\in[0,T]$.*

v) For

$$\begin{aligned}Z(t+r) :=& \exp\Big\{-\int_0^{t+r} H\big(T-t-r+s, X(s), \mathcal{MC}(s)\big)\mathrm{d}s\Big\}\\Y(r) :=& \exp\Big\{\int_0^{r} H\big(T-t-r+s, X(s), \mathcal{MC}(s)\big)\mathrm{d}s\Big\},\end{aligned}$$

it holds that

$$\begin{aligned}&\lim_{r\downarrow 0}\mathbb{E}\Big[Z(t+r)\frac{Y(r)-Y(0)}{r}\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big]\\&=\mathbb{E}\Big[a.s. \lim_{r\downarrow 0}Z(t+r)\frac{Y(r)-Y(0)}{r})\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big]\\&=\mathbb{E}[Z(t)\Big|X(0)=x, \mathcal{MC}(0)=e_i]H(T-t, x, e_i),\end{aligned}$$

for all $r\in[0,T-t]$.

iii)' $\mathbb{E}\Big[\int_0^r h_x\big(t, X(s), \mathcal{MC}(s)\big)'\Sigma\big(X(s), \mathcal{MC}(s)\big)\Sigma\big(X(s), \mathcal{MC}(s)\big)'$
$\cdot\, h_x\big(t, X(s), \mathcal{MC}(s)\big)\mathrm{d}s\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] < \infty, \forall r\in[0,T]$.

iv)' $\mathbb{E}\Big[\int_0^r \big(h(t, X(s), e_j)\big)^2\mathrm{d}s\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] < \infty, \forall r\in[0,T]$, $\forall j=1,\dots,I$.

v)' $H(t,x,e_i)$ is differentiable in t, $H(t,x,e_i)$ and $H_t(t,x,e_i)$ are continuous in t and x, and

$$\begin{aligned}&\lim_{r\downarrow 0}\mathbb{E}\Big[Z(t+r)\frac{Y(r)-Y(0)}{r}\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big]\\&=\mathbb{E}\Big[a.s.\lim_{r\downarrow 0}Z(t+r)\frac{Y(r)-Y(0)}{r})\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big],\end{aligned}$$

for all $r\in[0,T-t]$.

As an application of the backwards Feynman-Kac result one obtains the following corollary:

Corollary 2.78 (Linear system of coupled ODEs)
Consider again the Markov chain $\mathcal{MC}$ and define function $\xi:[0,T]\times\mathcal{E}\to\mathbb{R}$ as follows:

$$\xi(t,e_i)=\mathbb{E}\Big[\exp\Big\{-\int_t^T\Xi\big(s,\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\Big|\mathcal{MC}(t)=e_i\Big],$$

for some finite function $\Xi:[0,T]\times\mathcal{E}\to\mathbb{R}$. Assume that function ξ is well-defined and $\xi(t,e_i)<\infty,\forall\ (t,e_i)\in[0,T]\times\mathcal{E}$. Further, for an arbitrary but fix $t\in[0,T]$ and all $r\in[0,t]$, define:

$$\begin{aligned}Z(t+r):&=\exp\Big\{-\int_0^{t+r}\Xi\big(T-t-r+s,\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\\Y(r):&=\exp\Big\{\int_0^{r}\Xi\big(T-t-r+s,\mathcal{MC}(s)\big)\mathrm{d}s\Big\},\end{aligned}$$

and assume that

$$\begin{aligned}&\lim_{r\downarrow 0}\mathbb{E}\Big[Z(t+r)\frac{Y(r)-Y(0)}{r}\Big|\mathcal{MC}(0)=e_i\Big]\\&=\mathbb{E}\Big[a.s.\lim_{r\downarrow 0}Z(t+r)\frac{Y(r)-Y(0)}{r})\Big|\mathcal{MC}(0)=e_i\Big]\\&=\mathbb{E}[Z(t)\Big|\mathcal{MC}(0)=e_i]\Xi(T-t,e_i),\end{aligned}\tag{2.24}$$

for all $e_i\in\mathcal{E}$.
Alternatively to assuming (2.24) one can require that $\Xi(t,e_i)$ is differentiable in t, and that $\Xi(t,e_i)$ and $\frac{\partial}{\partial t}\Xi(t,e_i)$ are continuous in t for all $e_i\in\mathcal{E}$.
Then ξ satisfies the following system of ODEs for all $t\in[0,T]$:

$$\begin{aligned}&\frac{\partial}{\partial t}\xi(t,e_i)-\xi(t,e_i)\Xi(t,e_i)=-\sum_{j=1}^{l}q_{i,j}\xi(t,e_j)\\&\xi(T,e_i)=1,\forall i=1,\ldots,I.\end{aligned}$$

Proof
Recall that we have to check the conditions of Theorem 2.74 for $K(t, e_i) := \Xi(T - t, e_i)$ and $k(t, e_i) := \xi(T-t, e_i)$ in order to be able to apply Corollary 2.76. Conditions i) and iii) from Theorem 2.74 are trivially fulfilled, as we do not have dependence on x. The term $N(r)$ from Condition ii) has in this case the following form:

$$N(r) = \frac{1}{r}\int_0^r \sum_{j=1}^{I} q_{MC(s),j}\xi(t, e_j)\mathrm{d}s.$$

By the Dominated convergence theorem (Corollary 6.26 in [68]) it converges in mean to its almost sure limit, as the elements of the intensity matrix $q_{MC(s),j}$ are constrained. Condition iv)' is fulfilled as it holds:

$$\int_0^r \big(\xi(t, e_j)\big)^2 \mathrm{d}s = r\big(\xi(t, e_j)\big)^2 < \infty,$$

for all $r \in [0, t]$ and $e_j \in \mathcal{E}$. Condition v) coincides with (2.24).
It is only left to show that the continuity of $\Xi(t, e_i)$ and $\frac{\partial}{\partial t}\Xi(t, e_i)$ suffices for the convergence in mean in Condition v)'. To this aim observe that for all $s, t \in [0, T]$:

$$-\sum_{j=1}^{I} |\Xi(t, e_j)| \leq \Xi(t, \mathcal{MC}(s)) \leq \sum_{j=1}^{I} |\Xi(t, e_j)|.$$

This implies that:

$$\begin{aligned}
&\Big|Z(t+r)\frac{Y(r)-Y(0)}{r}\Big| = Z(t+r)\frac{|Y(r)-Y(0)|}{r}\\
&= \exp\Big\{\int_0^{t+r} -\Xi\big(T-t-r+s, \mathcal{MC}(s)\big)\mathrm{d}s\Big\}\\
&\frac{\big|\exp\Big\{\int_0^r \Xi\big(T-t-r+s, \mathcal{MC}(s)\big)\mathrm{d}s\Big\} - 1\big|}{r}\\
&\leq \exp\Big\{\int_0^{t+r} |\Xi\big(T-t-r+s, \mathcal{MC}(s)\big)|\mathrm{d}s\Big\}\\
&\frac{\big|\exp\Big\{\int_0^r \Xi\big(T-t-r+s, \mathcal{MC}(s)\big)\mathrm{d}s\Big\} - 1\big|}{r}
\end{aligned}$$

$$
\begin{aligned}
&\le \exp\Big\{\int_0^{t+r}\sum_{j=1}^{I}|\Xi\big(T-t-r+s,e_j\big)|\mathrm{d}s\Big\}\\
&\max\Bigg\{\underbrace{\frac{\big|\exp\Big\{-\int_0^r\sum_{j=1}^{I}|\Xi((T-t-r+s,e_j)|\mathrm{d}s\Big\}-1\big|}{r}}_{\overset{r\downarrow 0}{\to}\sum_{j=1}^{I}|\Xi\left(T-t,e_j\right)|\text{ by the Leibniz rule}},\\
&\underbrace{\frac{\big|\exp\Big\{\int_0^r\sum_{j=1}^{I}|\Xi((T-t-r+s,e_j)|\mathrm{d}s\Big\}-1\big|}{r}}_{\overset{r\downarrow 0}{\to}\sum_{j=1}^{I}|\Xi\left(T-t,e_j\right)|\text{ by the Leibniz rule}}\Bigg\}\\
&\le c_1\big(\sum_{j=1}^{I}|\Xi\big(T-t,e_j\big)|+c_2\big)<c_3,
\end{aligned}
$$

for sufficiently small $r > 0$ and all $t \in [0, T]$, where c_1, c_2 and c_3 are some positive real numbers. The convergence in mean follows again by the Dominated convergence theorem (Corollary 6.26 in [68]).

□

2.8 Utility functions

The risk preferences of the investor are characterized by her utility function $U : D^U \to \mathbb{R}, v \mapsto U(v)$. We adopt the usual assumptions on the utility functions (see e.g. [71]): $U(v)$ is increasing, continuously differentiable and concave in v. These assumptions imply that the investor prefers more to less, is risk averse and that it gets harder and harder to increase her happiness as wealth increases. In what follows we also assume that U is twice continuously differentiable. In order to understand better how utility functions reflect the risk-preferences of the investor we briefly summarize the concepts of risk-premium and risk-aversion. This overview is based on [90].

Definition 2.79 (Risk premium)
Consider an investor with current wealth v and utility function U who is faced with a risk with payoff given by the random variable Z with $\mathbb{E}[Z] = 0$, such that $\mathbb{E}[U(v+Z)]$ is finite. The risk premium is the amount $p^U(v;Z)$ that makes the investor indifferent between taking the risk or paying the deterministic amount $p^U(v;Z)$:

$$U(v - p^U(v;Z)) = \mathbb{E}[U(v + Z)].$$

An investor is called (strictly) risk averse if $p^U(v;Z) \ge 0$ ($p^U(v;Z) > 0$) for all $v \in D^U$ and all Z. The higher the risk premium $p^U(v;Z)$ for the same risk Z, the more risk averse the investor.

Proposition 2.80 ([90], p. 130)
An investor is (strictly) risk averse if and only if U is (strictly) concave.

Definition 2.81 (Arrow-Pratt measure of absolute risk aversion)
The Arrow-Pratt measure of absolute risk aversion for utility function U is defined as follows:

$$AP^U(v) = -\frac{U_{vv}}{U_v}.$$

The next proposition explains why AP^U is considered as a measure of risk aversion.

Proposition 2.82 ([90], p. 125, Equation (5))
Consider the setting from Definition 2.79 and denote $\Sigma^Z := sd(Z)$. Assume that $\mathbb{E}[|Z|^3] = o((\Sigma^Z)^2)$, i.e. $\mathbb{E}[|Z|^3]$ is of smaller order than $(\Sigma^Z)^2$. Then, it holds:

$$p^U(v;Z) = \frac{1}{2}(\Sigma^Z)^2 AP^U(v) + o\big((\Sigma^Z)^2\big).$$

So, AP^U is twice the risk premium per unit of variance.

As stated in the next proposition, AP^U can be used equivalently to p^U to compare the risk aversion of two investors, respectively utility functions.

Proposition 2.83 ([90], p. 128, Theorem 1)
Let U_1 and U_2 be two utility functions with $D^{U_1} = D^{U_2} = D^U$. Then the following two statements are equivalent:

i) $AP^{U_1}(v) \geq AP^{U_2}(v)$, *for all* $v \in D^U$.

ii) $p^{U_1}(v;Z) \geq p^{U_2}(v;Z)$, *for all* $v \in D^U$ *and all* Z.

An alternative way of characterizing the risk aversion is given by the so-called Arrow-Pratt measure of relative risk aversion. It is related to the concept of proportional risk premium:

Definition 2.84 (Proportional risk premium)
If an investor with utility function U is indifferent between carrying a risk $v\tilde{Z}$, for a random variable $\tilde{Z}$ with $\mathbb{E}[\tilde{Z}] = 0$, and paying the deterministic amount $v\tilde{p}^U(v;\tilde{Z})$, then we call $\tilde{p}^U(v;\tilde{Z})$ the proportional risk premium. This means that:

$$U(v - v\tilde{p}^U(v;\tilde{Z})) = \mathbb{E}[U(v + v\tilde{Z})].$$

Thus,

$$\tilde{p}^U(v;\tilde{Z}) = \frac{p^U(v;v\tilde{Z})}{v}.$$

Definition 2.85 (Arrow-Pratt measure of relative risk aversion)
The Arrow-Pratt measure of relative risk aversion is defined as follows:

$$\tilde{AP}^U(v) = -\frac{vU_{vv}}{U_v} = vAP^U(v).$$

Some of its basic properties are summarized in the following two corollaries, obtained directly from Propositions 2.82 and 2.83:

Corollary 2.86
Let $\tilde{Z}$ be a random variable with $\mathbb{E}[\tilde{Z}] = 0$. Denote $\Sigma^{\tilde{Z}} := sd(\tilde{Z})$. Assume that $\mathbb{E}[|\tilde{Z}|^3] = o\big((\Sigma^{\tilde{Z}})^2\big)$. Then it holds for any utility function U:

$$\tilde{p}^U(v;\tilde{Z}) = \frac{1}{2}(\Sigma^{\tilde{Z}})^2\tilde{AP}^U(v) + o\big((\Sigma^{\tilde{Z}})^2\big).$$

So, $\tilde{AP}^U$ is twice the relative risk premium (as a portion of the whole wealth) per unit of variance.

Corollary 2.87
Let U_1 and U_2 be two utility functions with $D^{U_1} = D^{U_2} = D^U$. Then the following two statements are equivalent:

i) $\tilde{AP}^{U_1}(v) \geq \tilde{AP}^{U_2}(v)$*, for all* $v \in D^U$.

ii) $\tilde{p}^{U_1}(v;\tilde{Z}) \geq \tilde{p}^{U_2}(v;\tilde{Z})$*, for all* $v \in D^U$ *and all* $\tilde{Z}$.

Utility functions can be classified by the criteria whether they have constant, decreasing or increasing Arrow-Pratt measure of (relative) risk aversion. One of the most popular utility functions is the constant relative risk aversion (CRRA) utility function, that up to a constant is given as follows:

$$U_P(v) = \frac{v^\delta}{\delta}, v \in D^{U_P} := [0,\infty), \tag{2.25}$$

where $\delta < 1, \delta \neq 0$. In what follows we call this utility function the power utility function due to its structure. Observe that its Arrow-Pratt measure of relative risk aversion is given by:

$$-\frac{v(U_P)_{vv}}{(U_P)_v} = 1 - \delta,$$

so parameter δ describes the risk aversion of the investor: $\delta \to 1$ describes a risk-neutral investor and the smaller δ, the more risk-averse the investor. This is confirmed by Figure 2.1, that shows the utility function for different parameter values. It can be seen that the higher δ, the bigger the weight of high wealth levels and the smaller the weight of low wealth values. This means that an investor with a δ

close to 1 is more willing to take risks compared to an investor with a lower δ. This relationship can be recognized even more clearly for negative values for δ, for which a very high negative weight is assigned to wealth levels close to zero whereas very high wealth levels improve only marginally the utility of the investor. That is why utility functions with a negative parameter δ are used to describe very risk averse investors.
A generalization of the power utility function is the so-called hyperbolic absolute

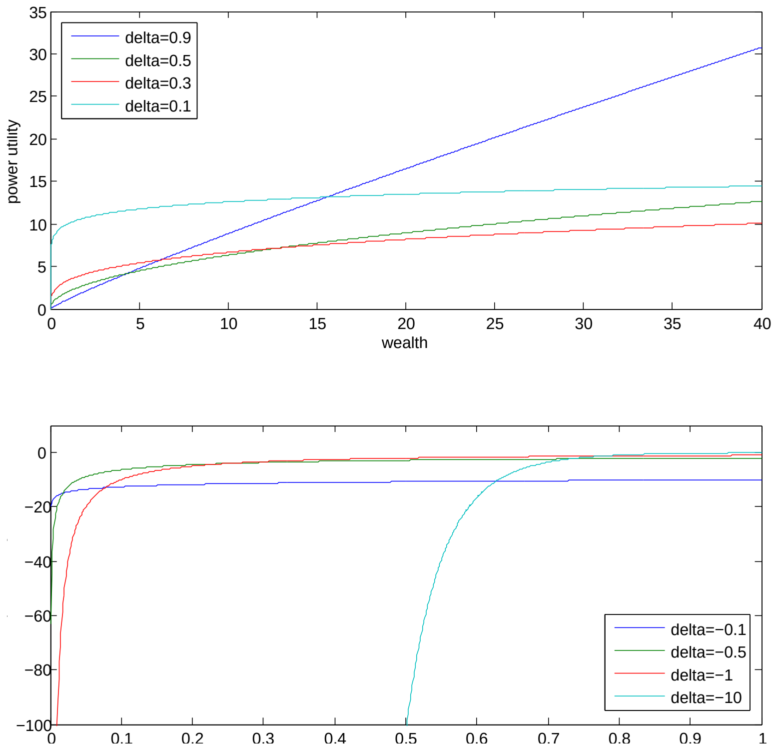


Figure 2.1: Power utility function $U_P(v) = \frac{v^\delta}{\delta}$ for different parameters $\delta \leq 1$.

risk aversion (HARA) utility function (see [81]). It can be expressed by different parametrizations, in the sequel we adopt the following definition:

$$U_H(v) = \frac{1-\delta}{\delta}\alpha\left\{\frac{1}{1-\delta}(v-F)\right\}^\delta, v \in D^{U_H} := [F, \infty), \tag{2.26}$$

where $\delta < 1$, $\delta \neq 0$, $F \geq 0$. In this case the Arrow-Pratt measure of absolute risk aversion is given by:

$$-\frac{(U_H)_{vv}}{(U_H)_v} = (1-\delta)\frac{1}{v-F},$$

so it is decreasing in v and the smaller the difference $v - F$, the more risk-averse the investor. Note that for $F = 0$ and $\alpha = (1-\delta)^{\delta-1}$, $U_H(v)$ corresponds to the power utility function $U_P(v) = \frac{v^\delta}{\delta}$. The HARA utility function is plotted in Figure 2.2 for different values of parameter δ. It basically corresponds to the power utility function shifted to the right by F.

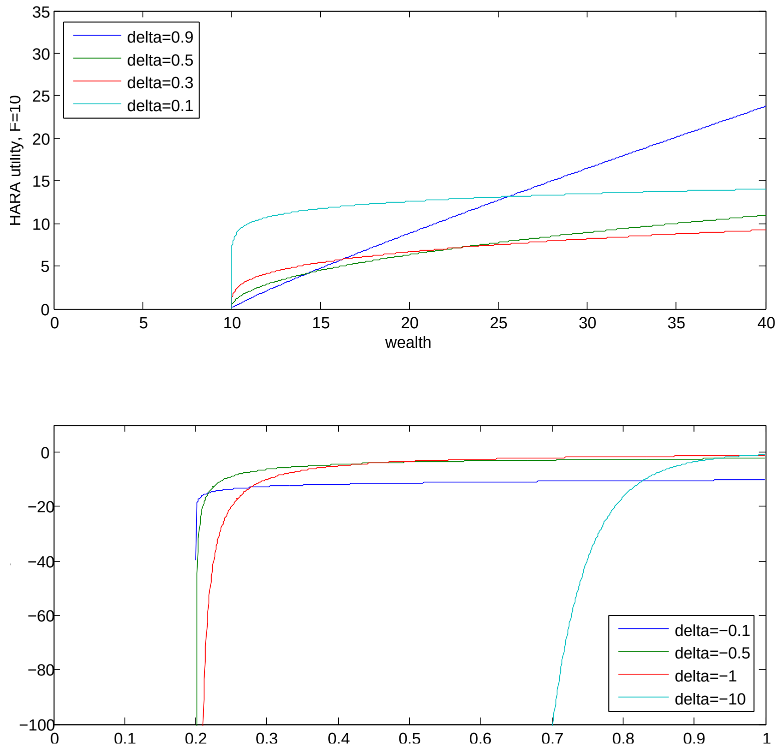


Figure 2.2: HARA utility function $U_H(v) = \frac{1-\delta}{\delta}\alpha\left\{\frac{1}{1-\delta}(v-F)\right\}^{\delta}, v \in D^{U_H} := [F,\infty)$ for different parameters $\delta \leq 1$. In the upper plot $F = 10$ and in the lower $F = 0.2$. For both plots we set $\alpha = (1-\delta)^{\delta-1}$ to ease the comparability with the power utility function.

Chapter 3

Problem formulation and the HJB approach

In this chapter we first present the general model we work with and state formally the optimization problem (see Section 3.1). Then, in Section 3.2 we introduce an auxiliary model without Markov switching, but with deterministic time-dependent model parameters. It is later on used to simplify the results in the case with Markov switching. Finally, in Section 3.3 we give an overview over the method we apply to solve the considered optimization problem. More precisely, we derive the corresponding HJB equations, state a candidate for the optimal control and prove a verification theorem that reduces the case with Markov switching to the one with time-dependent coefficients. The notation introduced in this chapter holds throughout the whole thesis unless something else is stated.

3.1 Model and optimization problem

We start with a formal definition of the considered class of models on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, where $\mathbb{P}$ denotes the real-world measure and $\mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ describes the corresponding information flow. Whenever nothing else is stated expectations are w.r.t. the real-world measure $\mathbb{P}$. As we have a finite investment horizon T we consider all processes over the time interval $[0, T]$.

Definition 3.1 (Multidimensional affine model with Markov switching)
We consider a financial model where investors dispose of a bank account with riskless interest rate r and have the opportunity to invest in N risky securities with price processes P_n, $n = 1, \ldots, N$. The bank account price process is denoted by P_0. We assume that these price processes are influenced both, by a Markov chain $\mathcal{MC}$ and by J stochastic factors modeled by the J-dimensional process X with values in $D^X \subseteq$

$\mathbb{R}^J$. More precisely, we consider the following model dynamics:

$$\begin{aligned}
\mathrm{d}P_0(t) &= P_0(t)r\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}t\\
\mathrm{d}P_n(t) &= P_n(t)\Big[\mu_n\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}t+\Sigma_n\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}W^P(t)\Big], n=1,\dots,N\\
\mathrm{d}X(t) &= \mu^X\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}t+\Sigma^X\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}W^X(t)\\
\mathrm{d}\langle W^X,W^P\rangle(t) &= \rho\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}t,
\end{aligned} \tag{3.1}$$

where W^P is a standard N-dimensional Brownian motion, W^X is a standard J-dimensional Brownian motion, $r(x,e_i): D^X\times\mathcal{E}\to\mathbb{R}$, $\mu_n(x,e_i): D^X\times\mathcal{E}\to\mathbb{R}$, $\Sigma_n(x,e_i): D^X\times\mathcal{E}\to\mathbb{R}^{1,N}$, $\mu^X(x,e_i): D^X\times\mathcal{E}\to\mathbb{R}^J$, $\Sigma^X(x,e_i): D^X\times\mathcal{E}\to\mathbb{R}^{J,J}$, $\rho(x,e_i): D^X\times\mathcal{E}\to\mathbb{R}^{J,N}$. The last equation in (3.1) is to be understood as follows:

$$\mathrm{d}\langle W_j^X,W_n^P\rangle(t)=\rho_{j,n}\big(X(t),\mathcal{MC}(t)\big)\ dt, \forall j=1,\dots,J, n=1,\dots,N.$$

Further we denote:

$$\begin{aligned}
P(t) &:= \big(P_1(t),\dots,P_n(t)\big)'\in\mathbb{R}^N\\
\mu\big(X(t),\mathcal{MC}(t)\big) &:= \Big(\mu_1\big(X(t),\mathcal{MC}(t)\big),\dots,\mu_N\big(X(t),\mathcal{MC}(t)\big)\Big)'\in\mathbb{R}^N\\
\Sigma\big(X(t),\mathcal{MC}(t)\big) &:= \Big(\Sigma_1\big(X(t),\mathcal{MC}(t)\big),\dots,\Sigma_N\big(X(t),\mathcal{MC}(t)\big)\Big)'\in\mathbb{R}^{N,N},
\end{aligned}$$

and assume that $\Sigma(X(t),\mathcal{MC}(t))$ is a.s. invertible for all $t\in[0,T]$ and that $\{\Sigma(X(t),\mathcal{MC}(t))\}_{t\in[0,T]}$ has a.s. finite paths. We call this model a multidimensional affine model with Markov switching if the following conditions are fulfilled for all $(x,e_i)\in D^X\times\mathcal{E}$:

$$r=\varepsilon^{(0)}(e_i)+\bar{\varepsilon}^{(1)}(e_i)'x \tag{3.2}$$

$$\mu^X=\bar{k}^{(0)}(e_i)-K^{(1)}(e_i)x \tag{3.3}$$

$$\Sigma^X(\Sigma^X)'=H^{(0)}(e_i)+\sum_{j=1}^J H^{(1j)}(e_i)x_j \tag{3.4}$$

$$(\mu-r)'\big(\Sigma\Sigma'\big)^{-1}(\mu-r)=h^{(0)}(e_i)+\bar{h}^{(1)}(e_i)'x \tag{3.5}$$

$$\Sigma^X\rho\Sigma^{-1}(\mu-r)=\bar{g}^{(0)}(e_i)+G^{(1)}(e_i)x \tag{3.6}$$

$$\Sigma^X\rho\rho'(\Sigma^X)'-\Sigma^X(\Sigma^X)'=L^{(0)}(e_i)+\sum_{j=1}^J L^{(1j)}(e_i)x_j, \tag{3.7}$$

where $\varepsilon^{(0)},h^{(0)}\in\mathbb{R}$, $\bar{\varepsilon}^{(1)},\bar{k}^{(0)},\bar{h}^{(1)},\bar{g}^{(0)}\in\mathbb{R}^J$, and $K^{(1)},H^{(0)},H^{(1j)},G^{(1)},L^{(0)},L^{(1j)}\in\mathbb{R}^{J,J}$, for all $j=1,\dots,J$. Subtraction of a scalar from a vector is to be understood as subtracting the scalar from every element of the vector, i.e. $\mu-r=(\mu_1-r,\dots,\mu_n-r)'$. Observe that on the right hand side of the expressions above we use small letters for scalars, small letters with an upper bar for vectors and capital letters for matrices. Furthermore, on the left hand side of the equations above the dependence on (x,e_i) is omitted for purposes of clarity.

Remark 3.2
In the case without Markov switching a generalization of the multidimensional affine models to quadratic processes and application to portfolio optimization is provided in [75].

Remark 3.3
The requirement that Σ is invertible is made just to ease the notation in the following derivation. However, it is not absolutely necessary. If suffices to assume that $\Sigma\Sigma'$ is invertible. In this case Condition (3.6) is replaced by the following equation:

$$\Sigma^X \rho \Sigma'(\Sigma\Sigma')^{-1}(\mu - r) = \bar{g}^{(0)}(e_i) + G^{(1)}(e_i)x, \tag{3.8}$$

and Conditions (3.2)-(3.5), (3.7) remain the same.

Note that the affine structure of the terms in Definition 3.1 is the most general one that allows for an exponential affine ansatz for the resulting HJB equation. On the other hand, this model definition is not only analytically tractable but also very flexible and covers various models as illustrated in Sections 4.3, 5.3 and 6.3. In the most general case all parameters are Markov-switching. This setting is tractable when $\rho = 0$. The challenge in the presence of a leverage effect $\rho \neq 0$ is to position the Markov chain in a suitable way so that both, the model flexibility and the existence of explicit solutions are preserved. We propose such frameworks in Sections 4.2.2, 5.2.2 and 6.1.
But first of all let us specify formally the optimization problem. We allow for continuous trading in the risky assets and the cash account, and assume that the investor can observe not only the asset prices but also the value of the stochastic factor X and the state of the Markov chain $\mathcal{MC}$, and makes her investment decision based on all this information. This assumption was made in numerous studies: see e.g. [104] and [74] for the observability of the stochastic factor and [8] and [95] for the Markov chain. So, a trading strategy can be described with an $N+1$-dimensional $\mathbb{F}$-adapted real-valued process $\varphi = \{\varphi(t)\}_{t\in[0,T]} = \{(\varphi_0(t), \varphi_1(t), \ldots, \varphi_N(t))'\}_{t\in[0,T]}$, where at time t, $\varphi_0(t)$ stands for the number of units invested in the cash account and, for $n = 1, \ldots, N$, $\varphi_n(t)$ denotes the number of risky assets P_n in the portfolio of the investor. In terms of this notation we define the wealth process corresponding to strategy φ as:

$$V^{\varphi}(t) := \sum_{i=0}^{N} \varphi_i(t) P_i(t).$$

We consider only self-financing portfolio strategies, which are characterized by the following equation:

$$V^{\varphi}(t) = V^{\varphi}(0) + \int_0^t \varphi'(s) \mathrm{d}P(s), t \in [0, T].$$

However, for the further computations it is more convenient to describe trading strategies not by the absolute portfolio positions but by the relative portfolio process that corresponds to the fraction of wealth invested in the single assets at each time point:

$$\pi_n(t) := \frac{\varphi_n(t)P_n(t)}{V^{\varphi}(t)}, n = 1, \dots, N.$$

From now on we denote the corresponding wealth process by V^{π}. We assume that $\pi := (\pi_1, \dots, \pi_n)'$ has a.s. finite paths on $[0, T]$ and that it is self-financing, so that the SDE for the wealth process is given by:

$$\begin{aligned}
\mathrm{d}V^{\pi}(t) =& \underbrace{V^{\pi}(t)\Big[r\big(X(t), \mathcal{MC}(t)\big) + \pi(t)'\Big(\mu\big(X(t), \mathcal{MC}(t)\big) - r\big(X(t), \mathcal{MC}(t)\big)\Big)\Big]}_{=:\mu^V(V^{\pi}(t), X(t), \mathcal{MC}(t), \pi(t))} \mathrm{d}t \\
&+ \underbrace{V^{\pi}(t)\pi(t)'\Sigma\big(X(t), \mathcal{MC}(t)\big)}_{=:\Sigma^V(V^{\pi}(t), X(t), \mathcal{MC}(t), \pi(t))} \mathrm{d}W^P(t),
\end{aligned} \tag{3.9}$$

where the initial wealth of the investor is $V^{\pi}(0) = v_0$. An application of Theorem 2.72 to $ln(V^{\pi})$ leads to the following solution of SDE (3.9):

$$\begin{aligned}
V^{\pi}(t) =& v_0 \exp\Bigg\{ \int_0^t \Big[r\big(X(t), \mathcal{MC}(t)\big) + \pi(t)'\Big(\mu\big(X(t), \mathcal{MC}(t)\big) - r\big(X(t), \mathcal{MC}(t)\big)\Big) \\
&- \frac{1}{2}\pi(t)'\Sigma\big(X(t), \mathcal{MC}(t)\big)\Sigma\big(X(t), \mathcal{MC}(t)\big)'\pi(t)\Big]\mathrm{d}s \\
&+ \int_0^t \pi(t)'\Sigma\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}W^P(s)\Bigg\},
\end{aligned}$$

which shows that $V^{\pi}(t) > 0$, for all $t \in [0, T]$, if $v_0 > 0$.
The risk preferences of the investor are characterized by her utility function $U : D^U \to \mathbb{R}, v \mapsto U(v)$. In the subsequent sections we consider the power and the HARA utility functions (see Section 2.8). The investor aims at maximizing her expected utility from terminal wealth by dynamically choosing her investment strategy from the set of all admissible portfolio strategy. More, precisely, this set is defined as follows for all $(t, v) \in [0, T] \times [0, \infty)$:

$$\begin{aligned}
\Lambda(t, v) := \Big\{\pi \Big| \pi(s) \in \mathbb{R}^N, V^{\pi}(t) = v, V^{\pi}(s) \geq 0, \forall s \in [t, T], V^{\pi}(T) \in D^U, \\
\mathbb{E}\Big[-\min\big\{U\big(V^{\pi}(T)\big), 0\big\} | \mathscr{F}_t\Big] < \infty\Big\}.
\end{aligned} \tag{3.10}$$

So, $\Lambda(t, v)$ contains strategies starting at time t with wealth v that lead to a non-negative wealth process. Strictly speaking, one should write $\Lambda(t, v, x, e_i)$, however we omit the remaining arguments for better readability. Furthermore, denote:

$$D^{\Lambda} := \big\{(t, v, x, e_i) \in [0, T] \times [0, \infty) \times D^X \times \mathcal{E} \big| \Lambda(t, v, x, e_i) \neq \emptyset\big\}. \tag{3.11}$$

Now we can formally state the optimization problem we consider throughout the whole thesis. For all $(t, v, x, e_i) \in D^\Lambda$ we are interested in the following problem:

$$\begin{aligned} J^{(t,v,x,e_i)}(\pi) &:= \mathbb{E}\Big[U(V^\pi(T))|V^\pi(t) = v, X(t) = x, \mathcal{MC}(t) = e_i\Big] \\ \Phi(t, v, x, e_i) &:= \max_{\pi\in\Lambda(t,v)} J^{(t,v,x,e_i)}(\pi). \end{aligned} \tag{3.12}$$

The maximal expected utility Φ is called the value function. Observe that now the optimization problem has not only one state variable (v) expressing the current wealth level V^π, as in the classical Merton's optimization problem, but two additional (x, e_i), corresponding to the two additional sources of randomness in our model: the stochastic factors X and the Markov chain $\mathcal{MC}$.
Let us have a closer look at the set of admissible portfolio strategies $\Lambda(t, v)$. The last condition in Equation (3.10) requires that the negative part of the terminal utility is integrable, excluding strategies that might lead with a positive probability to infinite negative utility. Note that this condition is trivially fulfilled for the optimal portfolio with the power and the HARA utility functions, as they are either positive or negative on their whole definition sets. So, if one can show that the value function is finite, its negative part is either zero or equals the finite value function. Furthermore, due to the exponential structure of our model, the wealth process is positive for all self-financing trading strategies. Thus, for the power utility function we do not need to check additionally the admissibility of the optimal trading strategies but can perform optimization directly over $\mathbb{R}^N$. For the HARA utility function U_H we just need to assure that $V^\pi(T) \in D^{U_H}$.

3.2 Time-dependent model

In some cases Problem (3.12) can be solved using the results in a simpler auxiliary market with deterministic time-dependent coefficients instead of Markov-switching ones. To this aim we replace in Model (3.1) the Markov chain $\mathcal{MC}$ by a deterministic piece-wise constant function $m : [0, T] \to \mathcal{E}$ with at most a countable number of jumps in $[0, T]$, denoted by $0 < t_1 < \ldots < t_K \leq T$. Set $t_0 = 0$. More precisely, m is given by:

$$m(t) := \begin{cases} m_0 & t \in [t_0, t_1) \\ m_1 & t \in [t_1, t_2) \\ \vdots & \\ m_K & t \in [t_K, T), \end{cases} \tag{3.13}$$

where $m_k := m(t_k) \in \mathcal{E}$ are the corresponding states of m. Let us denote the set of all such functions by $\mathbb{M}$.

Using this notation we define the following model:

$$
\begin{aligned}
&\mathrm{d}P_0^m(t) = P_0^m(t) r\big(X^m(t), m(t)\big)\mathrm{d}t \\
&\mathrm{d}P_n^m(t) = P_n^m(t)\Big[\mu_n\big(X^m(t), m(t)\big)\mathrm{d}t + \Sigma_n\big(X^m(t), m(t)\big)\mathrm{d}W^P(t)\Big], n = 1, \dots, N \\
&\mathrm{d}X^m(t) = \mu^X\big(X^m(t), m(t)\big)\mathrm{d}t + \Sigma^X\big(X^m(t), m(t)\big)\mathrm{d}W^X(t) \\
&\mathrm{d}\langle W^P, W^X\rangle(t) = \rho\big(X^m(t), m(t)\big)\mathrm{d}t,
\end{aligned}
\tag{3.14}
$$

where W^P and W^X, as well as r, $\mu_n, \Sigma_n, \mu^X, \Sigma^X$ are as in the definition of Model (3.1). Analogously to the previous section, P_0^m, P^m, X^m and $V^{m,\pi}$ denote the bank account, the risky assets, the stochastic factors and the wealth process. We can interpret this model as Model (3.1) conditioned on an arbitrary but fixed path m of the Markov chain. That is why we call it the time-dependent model corresponding to Model (3.1) or the time-dependent model induced by m and vice versa, Model (3.1) is the Markov-switching model corresponding to Model (3.14).
For the time-dependent model the optimization problem is stated as follows:

$$
\begin{aligned}
J^{(t,v,x,m)}(\pi) &:= \mathbb{E}\Big[U\big(V^{m,\pi}(T)\big)\Big|V^{m,\pi}(t) = v, X^m(t) = x\Big] \\
\Phi^m(t,v,x) &:= \max_{\pi\in\Lambda^m(t,v)} J^{(t,v,x,m)}(\pi),
\end{aligned}
\tag{3.15}
$$

with

$$
\begin{aligned}
\Lambda^m(t,v) := \Big\{\pi^m\Big|\pi^m(s) \in \mathbb{R}^N, V^{m,\pi}(t) = v, V^{m,\pi}(s) \geq 0, \forall s \in [t,T], V^{m,\pi}(T) \in D^U, \\
\mathbb{E}\Big[-\min\big\{U\big(V^{m,\pi}(T)\big), 0\big\}|\mathscr{F}_t\Big] < \infty\Big\}.
\end{aligned}
$$

In the next section we give an overview over the method that we will apply to solve Problems (3.12) and (3.15) and provide the link between their solutions.

3.3 HJB approach

As we are dealing with an incomplete market we will apply the Hamilton-Jacobi-Bellman (HJB) approach. It is based on Bellman's principle (see [10]):

$$
\Phi(t,v,x,e_i) = \sup_{\pi\in\Lambda(t,v)} \mathbb{E}\Big[\Phi\big(t+h, V^\pi(t+h), X(t+h)\big), \mathcal{MC}(t+h)\Big|\mathscr{F}_t\Big],
$$

for any $h > 0$, where $V^\pi(t) = v$, $X(t) = x$ and $\mathcal{MC}(t) = e_i$.
Applying Itô's formula to the right-hand side of the above equation, taking the limit $h \to 0$ and interchanging the expectation and the limit leads to the following equation, called the HJB equation:

$$
\begin{aligned}
\max_{\pi\in\mathbb{R}^n}\{\mathcal{L}(e_i,\pi)\Phi(t,v,x,e_i)\} &= -\sum_{z=1}^{I} q_{i,z}\Phi(t,v,x,e_z) \\
\Phi(T,v,x,e_i) &= U(v), \forall i \in \{1,\dots,I\},
\end{aligned}
\tag{3.16}
$$

where the differential operator $\mathcal{L}(e_i,\pi)$ is given for each $e_i \in \mathcal{E}$ as follows:

$$\begin{aligned}\mathcal{L}(e_i,\pi)\Phi :=& \Phi_t + \mu^V\Phi_v + (\mu^X)'\Phi_x + \frac{1}{2}\Sigma^V(\Sigma^V)'\Phi_{vv} + \frac{1}{2}Tr\big(\Sigma^X(\Sigma^X)'\Phi_{xx'}\big) \\ &+ \Sigma^V\rho'(\Sigma^X)'\Phi_{vx}\Big|_{(t,v,x,e_i,\pi)}. \end{aligned} \tag{3.17}$$

Note that we have neglected here the admissibility restrictions for π. They will be verified at the end for the derived optimal portfolio. Observe that because of the Markov chain we have to deal with a system of coupled PDEs in contrast to a single PDE for models driven by standard Itô diffusions.
Now, we consider the first-order condition for an interior maximum:

$$\frac{\partial}{\partial\pi}\big\{\mathcal{L}(e_i,\pi)\Phi(t,v,x,e_i)\big\} = 0.$$

By differentiating the left hand-side of Equation (3.16) we obtain:

$$v\Phi_v(\mu - r) + v^2\Phi_{vv}\Sigma\Sigma'\pi + v\Sigma\rho'(\Sigma^X)'\Phi_{vx} = 0,$$

which leads to the following candidate for the optimal investment strategy:

$$\bar{\pi}(t) = -\frac{1}{V^{\bar{\pi}}(t)\Phi_{vv}}\Big\{\Phi_v\big(\Sigma\Sigma'\big)^{-1}(\mu - r) + (\Sigma')^{-1}\rho'(\Sigma^X)'\Phi_{vx}\Big\}\Big|_{(t,V^{\bar{\pi}}(t),X(t),\mathcal{MC}(t))}. \tag{3.18}$$

So, we need the solution for the value function in order to derive the optimal portfolio in an explicit form. Note that substitution of (3.18) in (3.16) yields the following PDE for the value function Φ:

$$\begin{aligned}&\Phi_t + vr\Phi_v - \frac{1}{2}\frac{\Phi_v^2}{\Phi_{vv}}(\mu - r)'(\Sigma\Sigma')^{-1}(\mu - r) + (\mu^X)'\Phi_x - \frac{1}{2}\frac{1}{\Phi_{vv}}\Phi_{vx}'\Sigma^X\rho\rho'(\Sigma^X)'\Phi_{vx} \\ &+ \frac{1}{2}Tr\big(\Sigma^X(\Sigma^X)'\Phi_{xx'}\big) - \frac{\Phi_v}{\Phi_{vv}}(\mu - r)'(\Sigma')^{-1}\rho'(\Sigma^X)'\Phi_{vx} = -\sum_{z=1}^{I} q_{i,z}\Phi(t,v,x,e_z).\end{aligned} \tag{3.19}$$

In order to solve this equation we consider an ansatz for Φ suitably chosen depending on the considered utility function. In Chapters 4 and 5 we will see how it looks like for the power utility function and in Chapter 6 we deal with the HARA utility function. However, a solution for the HJB equation is not necessarily the value function to the considered optimization problem, as the HJB equation is obtained by a heuristic derivation and the necessary technical assumptions have not been proved so far. This is done in the so-called verification theorem. In what follows we state general verification theorems that summarize a set of necessary conditions for a function to be indeed the value function. When applying the verification theorem to a specific model one has to check these conditions, which might be quite technical. We do this for several relevant examples.

Remark 3.4 (Decomposition of the optimal portfolio)
It can be seen in Equation (3.18) that the well-known decomposition of the optimal portfolio in a mean-variance part and a hedging term known from the case with deterministic coefficients is preserved also with Markov switching. The mean-variance part (first summand in (3.18)) is mainly driven by the excess return of the traded risky assets and their volatility. The hedging part (second summand in (3.18)) appears because of the additional stochastic factor X. For the case with deterministic model parameters see e.g. [104]. In the next sections we analyze in detail these two terms.

In some cases the verification theorem for the Markov-switching model (3.1) can be reduced to the corresponding time-dependent case. That is why, here we state also the HJB equation in the time-dependent Model (3.14). It is defined piece-wise by:

$$\begin{aligned} \max_{\pi \in \mathbb{R}^N} \{\mathcal{L}(m(t), \pi)\Phi^m(t, v, x)\} = 0, \forall (t, v, x) \in [t_k, t_{k+1}) \times [0, \infty) \times D^X \\ \Phi^m(T, v, x) = U(v), \end{aligned} \tag{3.20}$$

for all $k = 0, \dots, K$, where the differential operator $\mathcal{L}$ is as given by (3.17) and function Φ^m is required to be continuous. Observe that in this context the HJB equation corresponds to a classical PDE with piece-wise constant coefficients and not to a system of PDEs, as in the case with Markov switching.
As before we use the first-order condition for an interior maximum to obtain a candidate for the optimal investment strategy:

$$\bar{\pi}^m(t) = -\frac{1}{V^{m,\pi}(t)\Phi^m_{vv}} \Big\{ \Phi^m_v \big(\Sigma\Sigma'\big)^{-1}(\mu - r) + \big(\Sigma'\big)^{-1} \rho'(\Sigma^X)'\Phi^m_{vx} \Big\} \Big|_{(t, V^{m,\pi}(t), X^m(t), m(t))}. \tag{3.21}$$

The next theorem links the solutions of the two optimization Problems (3.12) and (3.15).

Theorem 3.5 (Verification via the time-dependent model)
For all $m \in \mathbb{M}$, denote the value function in the time-dependent model induced by m by $\Phi^m(t, v, x) : [0, T] \times \mathbb{R}_{\geq 0} \times D^X \to \mathbb{R}$ and assume that there exists an optimal investment strategy $\bar{\pi}^m$, which depends on the current level of m and the current values of the stochastic processes, but not on the entire path of m on $[0, T]$, i.e. it holds:

$$\bar{\pi}^m(t) = p(t, V^{m,\bar{\pi}^m}(t), X^m(t), m(t)),$$

for some function $p : [0, T] \times [0, \infty) \times D^X \times \mathcal{E} \to \mathbb{R}$. Then, an optimal investment strategy in the corresponding Markov-modulated model is given by $\bar{\pi}(t) = p(t, V^{\bar{\pi}}(t), X(t), \mathcal{MC}(t))$ and for the value function it holds:

$$\Phi(t, v, x, e_i) = \mathbb{E}[\Phi^{\mathcal{MC}}(t, v, x) | \mathcal{MC}(t) = e_i].$$

Proof

First, to ease the readability of the derivations below, we recall the following notation for any n-dimensional process Z: $Z_{t,z}$ denotes the process Z started at time t in point $z \in \mathbb{R}^n$.

Let $\bar{\pi}^m$ and $\bar{\pi}$ be as defined in the statement of the theorem. Note the following equation for an arbitrary but fix $m \in \mathbb{M}$:

$$\big(X^m_{t,x}, V^{m,\bar{\pi}^m}_{t,v,x}, \bar{\pi}^m_{t,v,x}\big) \equiv \big(X_{t,x,m(t)}, V^{\bar{\pi}}_{t,v,x,m(t)}, \bar{\pi}_{t,v,x,m(t)}\big)\big|\{\mathcal{MC}(s) = m(s), \forall s \in [t,T]\}.$$

This implies that:

$$\big(X^{\mathcal{MC}}_{t,x,e_i}, V^{\mathcal{MC},\bar{\pi}^{\mathcal{MC}}}_{t,v,x,e_i}, \bar{\pi}^{\mathcal{MC}}_{t,v,x,e_i}\big) \equiv \big(X_{t,x,e_i}, V^{\bar{\pi}}_{t,v,x,e_i}, \bar{\pi}_{t,v,x,e_i}\big).$$

Pay attention that this is true because the optimal strategy $\bar{\pi}^m$ in the time-dependent model depends only on the current value of the step-wise function m and not on its whole path on $[t,T]$ and because of the independence between the Brownian motions and the Markov chain. By applying this observation and the tower rule for conditional expectations we obtain:

$$\begin{aligned}
\Phi(t,v,x,e_i) &= \mathbb{E}\big[\Phi^{\mathcal{MC}}(t,v,x)\big|\mathcal{MC}(t)=e_i\big] = \mathbb{E}\big[\Phi^{\mathcal{MC}_{t,e_i}}(t,v,x)\big] \\
&= \mathbb{E}\Big[\mathbb{E}\big[\Phi^{\mathcal{MC}_{t,e_i}}(t,v,x)\big|\mathscr{F}^{\mathcal{MC}_{t,e_i}}_T\big]\Big] = \mathbb{E}\Big[\mathbb{E}\big[\frac{(V^{\mathcal{MC},\bar{\pi}^{\mathcal{MC}}}_{t,v,x,e_i}(T))^\delta}{\delta}\big|\mathscr{F}^{\mathcal{MC}_{t,e_i}}_T\big]\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\big[\frac{\big(V^{\bar{\pi}}_{t,v,x,e_i}(T)\big)^\delta}{\delta}\big|\mathscr{F}^{\mathcal{MC}_{t,e_i}}_T\big]\Big] = \mathbb{E}\Big[\frac{\big(V^{\bar{\pi}}_{t,v,x,e_i}(T)\big)^\delta}{\delta}\Big].
\end{aligned}$$

So, function Φ expresses indeed the expected utility corresponding to strategy $\bar{\pi}$. To obtain the optimality of $\bar{\pi}$ consider an arbitrary admissible π and note that for all $m \in \mathbb{M}$, $\pi|\{\mathcal{MC}(s) = m(s), \forall s \in [t,T]\}$ is admissible also in the time dependent model induced by m. Thus, using the optimality of $\bar{\pi}^m$, one can compute:

$$\begin{aligned}
\Phi(t,v,x,e_i) &= \mathbb{E}\Big[\mathbb{E}\big[\frac{\big(V^{\mathcal{MC},\bar{\pi}^{\mathcal{MC}}}_{t,v,x,e_i}(T)\big)^\delta}{\delta}\big|\mathscr{F}^{\mathcal{MC}_{t,e_i}}_T\big]\Big] \\
&\geq \mathbb{E}\Big[\mathbb{E}\big[\frac{\big(V^{\mathcal{MC},\pi}_{t,v,x,e_i}(T)\big)^\delta}{\delta}\big|\mathscr{F}^{\mathcal{MC}_{t,e_i}}_T\big]\Big] = \mathbb{E}\Big[\frac{\big(V^{\pi}_{t,v,x,e_i}(T)\big)^\delta}{\delta}\Big].
\end{aligned}$$

□

Remark 3.6

Note that this theorem allows us to show a verification result for the Markov-switching model without showing explicitly that the value function solves the corresponding HJB system of PDEs. This might save lengthy technical proofs.

For a more detailed study of portfolio optimization problems in continuous-time and further solution methods besides the HJB approach we refer to [71].

After this general introduction to the HJB approach, in what follows we apply it to the considered class of models. In the next chapter we give a very detailed presentation for the one-dimensional case wherein trading is done only in the bank account and one risky asset, and the price processes are influenced by one stochastic factor.

Chapter 4

One-dim. affine Markov-modulated model

In this chapter we consider a special case of Model (3.1) with a riskless investment opportunity (bank account) and one risky asset. As mentioned in the introduction there are various studies in literature considering portfolio optimization either with a stochastic factor or with Markov switching. In what follows we extend this by combining both sources of randomness. The example considered at the end of this chapter is an extension of the well-known Heston model by Markov switching. Note that parts of this chapter have been published in [47].
The price processes of the riskless investment and the risky asset are denoted by $\{P_0(t)\}_{t\in[0,T]}$ and $\{P_1(t)\}_{t\in[0,T]}$, respectively. Their dynamics are influenced by a one-dimensional stochastic factor X with values in $D^X \subseteq \mathbb{R}$ and an observable continuous-time Markov chain $\mathcal{MC}$. More precisely, the considered model is defined as follows:

$$\begin{aligned} \mathrm{d}P_0(t) &= P_0(t) r\big(\mathcal{MC}(t)\big)\mathrm{d}t \\ \mathrm{d}P_1(t) &= P_1(t)\Big[\mu_1\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}t + \Sigma_1\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}W^P(t)\Big] \\ \mathrm{d}X(t) &= \mu^X\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}t + \Sigma^X\big(X(t),\mathcal{MC}(t)\big)\mathrm{d}W^X(t) \\ \mathrm{d}\langle W^P, W^X\rangle(t) &= \rho \mathrm{d}t, \end{aligned} \tag{4.1}$$

where W^P and W^X are two one-dimensional Brownian motions with constant correlation $\rho \in \mathbb{R}$. They are independent of the Markov chain. $r : \mathcal{E} \to \mathbb{R}$, $\mu_1, \Sigma_1, \mu^X, \Sigma^X : D^X \times \mathcal{E} \to \mathbb{R}$ are deterministic real-valued functions. Furthermore, we assume for all $(x, e_i) \in D^X \times \mathcal{E}$ the following conditions, which correspond to Assumptions (3.3)-(3.6) from the general model definition in Chapter 3:

$$\mu^X = \bar{k}^{(0)}(e_i) - K^{(1)}(e_i)x \tag{4.2}$$

$$(\Sigma^X)^2 = H^{(0)}(e_i) + H^{(11)}(e_i)x \tag{4.3}$$

$$\left(\frac{\mu_1 - r}{\Sigma_1}\right)^2 = h^{(0)}(e_i) + \bar{h}^{(1)}(e_i)x \tag{4.4}$$

$$\Sigma^X \rho \frac{\mu_1 - r}{\Sigma_1} = \bar{g}^{(0)}(e_i) + G^{(1)}(e_i)x, \tag{4.5}$$

for some real-valued functions $\bar{k}^{(0)}, K^{(1)}, H^{(0)}, H^{(11)}, h^{(0)}, \bar{h}^{(1)}, \bar{g}^{(0)}, G^{(1)} : \mathcal{E} \to \mathbb{R}$. Note that Condition (3.2) is trivially fulfilled for Model (4.1) and (3.7) follows directly from (4.3) in the one-dimensional case. In the one-dimensional case the SDE for the wealth process corresponding to portfolio π is given by:

$$\begin{aligned} \mathrm{d}V^\pi =& \underbrace{V^\pi(t)\Big[r\big(\mathcal{MC}(t)\big) + \pi(t)\Big(\mu_1\big(\mathcal{MC}(t)\big) - r\big(X(t), \mathcal{MC}(t)\big)\Big)\Big]}_{=:\mu^V} \mathrm{d}t \\ &+ \underbrace{V^\pi(t)\pi(t)\Sigma_1\big(X(t), \mathcal{MC}(t)\big)}_{=:\Sigma^V} \mathrm{d}W^P(t). \end{aligned} \tag{4.6}$$

Remark 4.1
Note that Assumption (4.5) is trivially fulfilled for $\rho = 0$, and for $\rho \neq 0$ it implies that:

$$\Sigma^X(x, e_i) = b(e_i)\frac{\mu_1(x, e_i) - r(e_i)}{\Sigma_1(x, e_i)} \Rightarrow \frac{\mu_1(x, e_i) - r(e_i)}{\Sigma_1(x, e_i)} = \frac{1}{b(e_i)}\Sigma^X(x, e_i), \tag{4.7}$$

for some deterministic function $b : \mathcal{E} \to \mathbb{R}$. So, for the case with correlation the market price of risk $\frac{\mu_1 - r}{\Sigma_1}$ should be a multiple of the volatility of the stochastic factor. Formally, this implies the following:

$$\begin{aligned} H^{(0)}(e_i) =& \big(b(e_i)\big)^2 h^{(0)}(e_i) \\ H^{(11)}(e_i) =& \big(b(e_i)\big)^2 \bar{h}^{(1)}(e_i) \\ \bar{g}^{(0)}(e_i) =& \rho b(e_i) h^{(0)}(e_i) \\ G^{(1)}(e_i) =& \rho b(e_i) \bar{h}^{(1)}(e_i). \end{aligned}$$

Observe that the drift and the diffusion term for the stochastic factor, as well as the squared market price of risk are affine in X. We will see later on that as a consequence of this, all terms in the corresponding HJB equation have affine structure, which allows for an exponentially affine ansatz for its solution. One important example for this class of models can be obtained by adding Markov switching to the famous Heston model. This example will be considered in detail in Section 4.3.
We assume that the risk-preferences of the investor are characterized by the power

utility function U_P as defined in (2.25) and consider Problem (3.12). The general HJB Equation (3.16) takes the following form in the one-dimensional case:

$$\begin{aligned} \max_{\pi\in\mathbb{R}}\{\mathcal{L}(e_i,\pi)\Phi(t,v,x,e_i)\} &= -\sum_{z=1}^{I} q_{i,z}\Phi(t,v,x,e_z) \\ \Phi(T,v,x,e_i) &= \frac{v^\delta}{\delta}, \forall i\in\{1,\dots,I\}, \end{aligned} \tag{4.8}$$

where

$$\mathcal{L}(e_i,\pi)\Phi := \Phi_t + \mu^V\Phi_v + \mu^X\Phi_x + \frac{1}{2}(\Sigma^V)^2\Phi_{vv} + \frac{1}{2}(\Sigma^X)^2\Phi_{xx} + \Sigma^V\rho(\Sigma^X)\Phi_{vx}\Big|_{(t,v,x,e_i,\pi)}, \tag{4.9}$$

and μ^V and Σ^V are as defined in Equation (4.6). As mentioned in Section 3.3 the candidate for the optimal portfolio derived from the first-order condition for a maximum is given by:

$$\bar{\pi}(t) = -\frac{(\mu_1 - r)\Phi_v + \rho\Sigma^X\Sigma_1\Phi_{vx}}{V^\pi(t)(\Sigma_1)^2\Phi_{vv}}\Big|_{\left(t,X(t),\mathcal{MC}(t)\right)}. \tag{4.10}$$

Now we propose an ansatz for the value function:

$$\Phi(t,v,x,e_i) = \frac{v^\delta}{\delta} f(t,x,e_i). \tag{4.11}$$

Substituting (4.11) and the one-dimensional model specifications in (3.19) leads to the following system for function f:

$$\begin{aligned} & f_t(t,x,e_i) + f(t,x,e_i)\underbrace{\delta\Big\{r(e_i) + \frac{1}{2}\frac{1}{1-\delta}\Big(\frac{\mu_1(x,e_i) - r(e_i)}{\Sigma_1(x,e_i)}\Big)^2\Big\}}_{=:g(x,e_i)} \\ & + f_x(t,x,e_i)\underbrace{\Big\{\mu^X(x,e_i) + \frac{\delta}{1-\delta}\rho\Sigma^X(x,e_i)\frac{\mu_1(x,e_i)-r(e_i)}{\Sigma_1(x,e_i)}\Big\}}_{=:\tilde{\mu}^X(x,e_i)} \\ & + \frac{1}{2}f_{xx}(t,x,e_i)\big(\Sigma^X(x,e_i)\big)^2 + \frac{1}{2}\frac{\big(f_x(t,x,e_i)\big)^2}{f(t,x,e_i)}\frac{\delta}{1-\delta}\rho^2\big(\Sigma^X(x,e_i)\big)^2 \\ & = -\sum_{z=1}^{I} q_{i,z} f(t,x,e_z), \\ & f(T,x,e_i) = 1, \forall i\in\{1,\dots,I\}. \end{aligned} \tag{4.12}$$

Furthermore, (4.10) and (4.11) yield the following simplified expression for $\bar{\pi}$:

$$\bar{\pi}(t) = \frac{1}{1-\delta}\Big\{\frac{\mu_1 - r}{(\Sigma_1)^2} + \rho\frac{\Sigma^X}{\Sigma_1}\frac{f_x}{f}\Big\}\Big|_{\left(t,X(t),\mathcal{MC}(t)\right)}. \tag{4.13}$$

So in order to solve the HJB equation and obtain a candidate for the optimal control we need to find function f. Before we do so, we state a general verification result that summarizes a set of sufficient conditions for the HJB solution to be the value function of our optimization problem.

Theorem 4.2 (Verification via a martingale condition)
Consider a real-valued function $\Phi(t, v, x, e_i) : [0, T] \times [0, \infty) \times D^X \times \mathcal{E} \to \mathbb{R}$ *and assume that:*

i) For each $e_i \in \mathcal{E}$, $\Phi(\cdot, \cdot, \cdot, e_i) \in \mathscr{C}^{1,2,2}\big([0, T] \times [0, \infty) \times D^X\big)$, *i.e.* Φ *is once continuously differentiable in* t *and twice continuously differentiable in* v *and* x.

ii) Φ *satisfies the following equation:*

$$\begin{aligned}\mathcal{L}(e_i, \bar{\pi})\Phi(t, v, x, e_i) &= -\sum_{z=1}^{I} q_{i,z}\Phi(t, v, x, e_z)\\ \Phi(T, v, x, e_i) &= U_P(v), \forall i \in \{1, \dots, I\},\end{aligned}$$

where operator $\mathcal{L}$ *is defined in (4.9) and* $\bar{\pi}$ *is given by (4.10).*

iii) $\big\{\Phi\big(t, V^{\bar{\pi}}(t), X(t), \mathcal{MC}(t)\big)\big\}_{t\in[0,T]}$ *is a martingale.*

Then:

$$\mathbb{E}\big[U_P\big(V^{\bar{\pi}}(T)\big)\big|V^{\bar{\pi}}(t) = v, X(t) = x, \mathcal{MC}(t) = e_i\big] = \Phi(t, v, x, e_i).$$

Further, if $\Phi = \frac{v^\delta}{\delta} f(t, x, e_i)$ *for a positive function* f *then* $\bar{\pi}$ *is the optimal solution and* Φ *is the corresponding value function.*

As the same result holds for the general multidimensional case (see Theorem 5.1), we omit the proof here and provide it in Chapter 5 for Theorem 5.1.
So, basically we have two approaches how to deal with the considered optimization problem. Either we solve System (4.12), find the HJB solution and apply Theorem 4.2, or we solve the HJB equation for the corresponding time-dependent model, proof that it is the value function and apply Theorem 3.5. As the second possibility is very convenient in many cases, in what follows we derive the solution for the time-dependent model. Afterwards, in Section 4.2, we will use these results for the solution with Markov switching.

4.1 Time-dependent model

In the one-dimensional case Model (3.14) is stated as follows:

$$\begin{aligned}
\mathrm{d}P_0^m(t) &= P_0^m(t) r\big(m(t)\big)\mathrm{d}t \\
\mathrm{d}P_1^m(t) &= P_1^m(t)\Big[\mu_1\big(X^m(t), m(t)\big)\mathrm{d}t + \Sigma_1\big(X^m(t), m(t)\big)\mathrm{d}W^P(t)\Big] \\
\mathrm{d}X^m(t) &= \mu^X\big(X^m(t), m(t)\big)\mathrm{d}t + \Sigma^X\big(X^m(t), m(t)\big)\mathrm{d}W^X(t) \\
\mathrm{d}\langle W^P, W^X\rangle(t) &= \rho \mathrm{d}t.
\end{aligned} \tag{4.14}$$

The risk preferences of the investor are again described by the power utility function U_P and we consider Problem (3.15). The HJB equation in this case takes the following form:

$$\begin{aligned}
\max_{\pi\in\mathbb{R}}\{\mathcal{L}(m(t), \pi)\Phi^m(t, v, x)\} = 0, \forall(t, v, x) \in [t_k, t_{k+1}) \times [0, \infty) \times D^X \\
\Phi^m(T, v, x) = \frac{v^\delta}{\delta},
\end{aligned} \tag{4.15}$$

for all $k = 0, \dots, K$, where the differential operator $\mathcal{L}$ is defined by (4.9), $t_0 := 0$, and $t_{K+1} := T$. As before, the candidate for the optimal investment strategy looks like this:

$$\bar{\pi}^m(t) = -\frac{(\mu_1 - r)\Phi_v^m + \rho\Sigma^X\Sigma_1\Phi_{vx}^m}{V^{m,\pi}(t)(\Sigma_1)^2\Phi_{vv}^m}\Big|_{\left(t, X^m(t), m(t)\right)}. \tag{4.16}$$

By the ansatz:

$$\Phi^m(t, v, x) = \frac{v^\delta}{\delta} f^m(t, x), \tag{4.17}$$

for some continuous real-valued function $f^m : [0, T] \times D^X \to \mathbb{R}$, we simplify $\bar{\pi}^m$ to:

$$\bar{\pi}^m(t) = \frac{1}{1-\delta}\Big\{\frac{\mu_1 - r}{(\Sigma_1)^2} + \rho\frac{\Sigma^X}{\Sigma_1}\frac{f_x^m}{f^m}\Big\}\Big|_{\left(t, X^m(t), m(t)\right)}. \tag{4.18}$$

Substitution of Equations (4.16) and (4.17) in the HJB equation (4.15) leads to the following PDE in t and x for function f^m defined piece-wise for all $k = 0, \dots, K$:

$$\begin{aligned}
& f_t^m(t, x) + f^m(t, x)\underbrace{\delta\Big\{r\big(m(t)\big) + \frac{1}{2}\frac{1}{1-\delta}\Big(\frac{\mu_1\big(x, m(t)\big) - r\big(m(t)\big)}{\Sigma_1\big(x, m(t)\big)}\Big)^2\Big\}}_{=:g(x, m(t))} \\
& + f_x^m(t, x)\underbrace{\Big\{\mu^X\big(x, m(t)\big) + \frac{\delta}{1-\delta}\rho\Sigma^X\big(x, m(t)\big)\frac{\mu_1\big(x, m(t)\big) - r\big(m(t)\big)}{\Sigma_1\big(x, m(t)\big)}\Big\}}_{=:\tilde{\mu}^X(x, m(t))} \\
& + \frac{1}{2}f_{xx}^m(t, x)\Big(\Sigma^X\big(x, m(t)\big)\Big)^2 + \frac{1}{2}\frac{\big(f_x^m(t, x)\big)^2}{f^m(t, x)}\frac{\delta}{1-\delta}\rho^2\Big(\Sigma^X\big(x, m(t)\big)\Big)^2 \\
& = 0, \forall(t, x) \in [t_k, t_{k+1}) \times D^X \\
& f^m(T, x) = 1.
\end{aligned} \tag{4.19}$$

Note that in this case we obtain a single PDE and not a system of PDEs as in the case with Markov switching. So, if we manage to find a solution to Equation (4.19) we will have also the solution to the HJB PDE (4.15). But first of all, we state in the subsequent proposition a set of sufficient conditions for this solution to be indeed the value function to the considered optimization problem.

Proposition 4.3 (Verification result in the time-dependent model)
Consider a real-valued function $\Phi^m(t,v,x) : [0,T] \times [0,\infty) \times D^X \to \mathbb{R}$ given by $\Phi^m(t,v,x,e_i) = \frac{v^\delta}{\delta} f^m(t,x)$ for a positive function f^m. Assume that:

i) $\Phi^m \in \mathscr{C}^{1,2,2}\big([t_k,t_{k+1}) \times [0,\infty) \times D^X\big)$ *for all* $k = 0,\dots,K$,

ii) $\Phi^m \in \mathscr{C}\big([0,T] \times [0,\infty) \times D^X\big)$,

iii) Φ^m *satisfies the following PDE, defined piece-wise for all* $k = 1,\dots,K$:

$$\mathcal{L}(m(t),\bar{\pi}^m)\Phi^m(t,v,x) = 0, \forall (t,v,x) \in [t_k,t_{k+1}) \times [0,\infty) \times D^X$$
$$\Phi^m(T,v,x) = \frac{v^\delta}{\delta},$$

with $\bar{\pi}^m$ *as given by (4.16).*

iv) $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ *is a martingale.*

Then:

$$\mathbb{E}\big[U_P\big(V^{m,\bar{\pi}^m}(T)\big)\big|V^{m,\bar{\pi}^m}(t) = v, X^m(t) = x\big] = \Phi^m(t,v,x),$$

and

$$\mathbb{E}\big[U_P\big(V^{m,\pi}(T)\big)\big|V^{m,\pi}(t) = v, X^m(t) = x\big] \le \Phi^m(t,v,x),$$

for all $(t,v,x) \in [0,T] \times [0,\infty) \times D^X$ and all admissible portfolio strategies π, i.e. $\bar{\pi}^m$ is an optimal investment strategy and Φ^m is the value function for the considered problem.

The proof is given directly for the multidimensional case in Proposition 5.3.

Remark 4.4
Note that a related result is presented in Corollary 2.5 from [63]. Using a fundamentally different methodology based on martingale theory and semimartingale characteristics the authors show that a trading strategy is optimal if there exists a special semimartingale that fulfills some martingale conditions. This process can be identified in our context with $\{f^m(t,X^m(t))\}_{t\in[0,T]}$ and the required assumptions there correspond to Conditions iii) and iv) from Proposition 4.3. Whereas the required conditions in [63] are rather abstract our derivation based on the HJB equation theory allows us to specify the necessary conditions in an explicit way.

Remark 4.5
If Φ^m is positive then the following two conditions:

i) $\Phi^m(\cdot,\cdot,\cdot) \in \mathscr{C}^{1,2,2}\big([0,T]\times[0,\infty)\times D^X\big)$,

ii) Φ *satisfies Equation (4.15) where the maximum on the left hand side is obtained at* $\bar\pi$ *as given by (4.16),*

suffice to show that:

$$\mathbb{E}\big[U_P\big(V^{m,\pi}(T)\big)\big|V^{m,\pi}(t)=v, X(t)=x\big] \le \Phi^m(t,v,x),$$

for all $(t,v,x)\in[0,T]\times[0,\infty)\times D^X$ *and all portfolio strategies* π*. The derivation can be found in Appendix B.*

Remark 4.6
Assumption iv) in Proposition 4.3 can be replaced by one of the following conditions:

iv)' For all $(v,x)\in[0,\infty)\times D^X$ *and all* $k=1,\dots,K$ *it holds:*

$$\begin{aligned}&\mathbb{E}\Big[\int_{t_k}^{t_{k+1}}\Big(\Phi_v^m\big(s,V^{m,\bar\pi^m}(s),X^m(s)\big)\big)\Sigma^V\big(V^{m,\bar\pi^m}(s),X^m(s),m(s),\bar\pi^m(s)\big)\Big)^2\\&+\Big(\Phi_x^m\big(s,V^{m,\bar\pi^m}(s),X^m(s)\big)\Sigma^X\big(X^m(s),m(s)\big)\Big)^2\mathrm{d}s\\&\Big|V^{m,\bar\pi^m}(t_k)=v,X^m(t_k)=x\Big]<\infty.\end{aligned}$$

iv)" Functions $\bar\pi^m(s,v,x,e_i)$ *and* $\Sigma_1(x,e_i)$ *are continuous in* (s,v,x) *for all* $e_i\in\mathcal{E}$*, and for all* $k=1,\dots,K$ *it holds: for every sequence of stopping times* $\{\theta_n\}_{n\in\mathbb{N}}$ *with* $\theta_n\to t_{k+1}$ *it holds that:*

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\big[\Phi^m(\theta_n,V^{m,\bar\pi^m}(\theta_n),X^m(\theta_n))\big|V^{m,\bar\pi^m}(t_k)=v,X^m(t_k)=x\big]\\&=\mathbb{E}\big[\Phi^m(t_{k+1},V^{m,\bar\pi^m}(t_{k+1}),X^m(t_{k+1}))\big|V^{m,\bar\pi^m}(t_k)=v,X^m(t_k)=x\big].\end{aligned}$$

The proof is given in Appendix B.

Now we can continue with the solution to the HJB equation.

Proposition 4.7 (Solution in the time-dependent model)
Set $\vartheta=\frac{1-\delta}{1-\delta+\delta\rho^2}$ *. Assume that the following equation:*

$$\begin{aligned}&\tilde B_t^m(t)+\frac{1}{2}H^{(11)}\big(m(t)\big)\big(\tilde B^m(t)\big)^2+\big\{\frac{\delta}{1-\delta}G^{(1)}\big(m(t)\big)-K^{(1)}\big(m(t)\big)\big\}\tilde B^m(t)\\&+\frac{1}{2}\frac{1}{\vartheta}\frac{\delta}{1-\delta}\bar h^{(1)}\big(m(t)\big)=0,\tilde B^m(T)=0\end{aligned}\tag{4.20}$$

admits a continuous, piece-wise continuously differentiable solution $\tilde{B}^m \in \mathscr{C}^1\big([t_k, t_{k+1}) \times [0,\infty) \times D^X\big)$, *for all* $k \in \{0,1,2,\dots,K\}$ *and denote* $B^m(t) := \vartheta \tilde{B}^m(t)$. *Furthermore, define function* A^m *as follows:*

$$A^m(t) = \int_t^T \frac{1}{2}\frac{1}{\vartheta} H^{(0)}\big(m(s)\big)\big(B^m(s)\big)^2 + \big\{\frac{\delta}{1-\delta}\bar{g}^{(0)}\big(m(s)\big) + \bar{k}^{(0)}\big(m(s)\big)\big\}B^m(s)\mathrm{d}s. \tag{4.21}$$

Then the solution of the HJB Equation (4.15) is given by:

$$\begin{aligned}
\Phi^m(t,v,x) &= \frac{v^\delta}{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{\vartheta} g\big(\tilde{X}^m(s), m(s)\big)\mathrm{d}s\Big\}\Big|\tilde{X}^m(t) = x\Big]^\vartheta \\
&= \frac{v^\delta}{\delta}\exp\Big\{\int_t^T \delta\Big\{r(m(s)) + \frac{1}{2}\frac{1}{1-\delta}h^{(0)}(m(s))\Big\} + \frac{1}{2}\frac{1}{\vartheta}H^{(0)}\big(m(s)\big)\big(B^m(s)\big)^2 \\
&+ \big\{\frac{\delta}{1-\delta}\bar{g}^{(0)}\big(m(s)\big) + \bar{k}^{(0)}\big(m(s)\big)\big\}B^m(s)\mathrm{d}s\Big\}\exp\{B^m(t)x\} \qquad (4.22) \\
&= \frac{v^\delta}{\delta}\exp\Big\{\int_t^T \delta\Big\{r(m(s)) + \frac{1}{2}\frac{1}{1-\delta}h^{(0)}(m(s))\Big\}\mathrm{d}s\Big\}\exp\{A^m(t) + B^m(t)x\} \\
&=: \frac{v^\delta}{\delta}\xi^m(t)\exp\big\{B^m(t)x\big\}. \qquad (4.23)
\end{aligned}$$

The maximum in (4.15) is obtained at:

$$\bar{\pi}^m(t) = \frac{1}{1-\delta}\Big\{\frac{\mu_1 - r}{(\Sigma_1)^2} + \rho\frac{\Sigma^X}{\Sigma_1}B^m\Big\}\Big|_{(t,X^m(t),m(t))}. \tag{4.24}$$

Trivially, $\Phi^m \in \mathscr{C}^{1,2,2}\big([t_k, t_{k+1}) \times [0,\infty) \times D^X\big)$ *for all* $k \in \{0,1,2,\dots,K\}$ *and it is continuous on the whole interval* $[0,T]$.
If $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ *is a martingale, then* Φ^m *is the value function and* $\bar{\pi}^m$ *is the optimal portfolio.*

Proof
We start with the solution to (4.19). Observe that due to the nonlinear term Feynman-Kac theorem cannot be directly applied to obtain a probabilistic representation of the solution. As in [104] we apply the following transformation to eliminate the nonlinear term:

$$h^m(t,x) := \big(f^m(t,x)\big)^{\frac{1}{\vartheta}},$$

for some $\vartheta \in \mathbb{R}_{>0}$. Then h^m solves the PDE below:

$$\begin{aligned}
&h_t^m(t,x) + h^m(t,x)\frac{1}{\vartheta}g\big(x,m(t)\big) + h_x^m(t,x)\tilde{\mu}^X\big(x,m(t)\big) + \frac{1}{2}h_{xx}^m(t,x)\Big(\Sigma^X\big(x,m(t)\big)\Big)^2 \\
&+ \frac{1}{2}\frac{\big(h_x^m(t,x)\big)^2}{h^m(t,x)}\Big(\Sigma^X\big(x,m(t)\big)\Big)^2\Big\{\vartheta - 1 + \vartheta\frac{\delta}{1-\delta}\rho^2\Big\} = 0, \\
&h^m(T,x) = 1.
\end{aligned}$$

We choose ϑ in such a way that the factor for the nonlinear term disappears, i.e.

$$\vartheta = \frac{1-\delta}{1-\delta+\delta\rho^2}, \tag{4.25}$$

and obtain:

$$\begin{aligned} &h^m_t(t,x) + h^m(t,x)\frac{1}{\vartheta}g(x,m(t)) + h^m_x(t,x)\tilde{\mu}^X(x,m(t)) \\ &+\frac{1}{2}h^m_{xx}(t,x)\Big(\Sigma^X\big(x,m(t)\big)\Big)^2 = 0, \\ &h^m(T,x) = 1. \end{aligned} \tag{4.26}$$

which is to be understood piece-wise. Observe that $\vartheta = 1$ for $\rho = 0$, which is in accordance with the fact that for $\rho = 0$ the PDE for f^m is linear. Under some integrability conditions we can apply Corollary 2.76 to obtain the following probabilistic representation for the solution of Equation (4.26):

$$\begin{aligned} h^m(t,x) =& \mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{\vartheta}g\big(\tilde{X}^m(s),m(s)\big)\Big\}\Big|\tilde{X}^m(t)=x\Big] \\ =& \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta\Big\{r\big(m(s)\big) + \frac{1}{2}\frac{1}{1-\delta}h^{(0)}\big(m(s)\big)\Big\}\mathrm{d}s\Big\} \\ &\cdot\mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{\vartheta}\frac{1}{2}\frac{\delta}{1-\delta}\bar{h}^{(1)}\big(m(s)\big)\tilde{X}^m(s)\mathrm{d}s\Big\}\Big|\tilde{X}^m(t)=x\Big], \end{aligned} \tag{4.27}$$

where the dynamics of process $\tilde{X}^m$ are given for $t\in[0,T]$ by the following SDE:

$$\mathrm{d}\tilde{X}^m(t) = \tilde{\mu}^X\big(\tilde{X}^m(t),m(t)\big)\mathrm{d}t + \Sigma^X\big(\tilde{X}(t),m(t)\big)\mathrm{d}W^X(s).$$

Observe that when the expectation in Representation (4.27) is known in a closed form, we do not need to check the assumptions of Corollary 2.76, as we can just plug in the expectation and verify that it is indeed the solution to the HJB equation. In what follows we will characterize the expression for this expectation up to a Riccati ODE.

First recall the affine definition of our model, which implies that:

$$\begin{aligned} \tilde{\mu}_X(x,e_i) &= \mu_X(x,e_i) + \frac{\delta}{1-\delta}\rho\Sigma^X(x,e_i)\frac{\mu_1(x,e_i)-r(e_i)}{\Sigma_1(x,e_i)} \\ &= \bar{k}^{(0)}(e_i) + \frac{\delta}{1-\delta}\bar{g}^{(0)}(e_i) + \{\frac{\delta}{1-\delta}G^{(1)}(e_i) - K^{(1)}(e_i)\}x \\ \big(\Sigma^X(x,e_i)\big)^2 &= H^{(0)}(e_i) + H^{(11)}(e_i)x. \end{aligned}$$

This specification of the parameters allows for an affine ansatz for function h^m. More precisely, we assume that:

$$h^m(t,x) = \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta\Big\{r(m(s)) + \frac{1}{2}\frac{1}{1-\delta}h^{(0)}\big(m(s)\big)\Big\}\mathrm{d}s\Big\}\exp\{\tilde{A}^m(t)+\tilde{B}^m(t)x\}.$$

Inserting this ansatz in Equation (4.26) and equating the coefficients in front of x^0 and x^1 leads to the following two ODEs for $\tilde{A}^m$ and $\tilde{B}^m$:

$$\begin{aligned} &\tilde{B}^m_t(t) + \frac{1}{2}H^{(11)}\big(m(t)\big)\big(\tilde{B}^m(t)\big)^2 + \Big\{\frac{\delta}{1-\delta}G^{(1)}\big(m(t)\big) - K^{(1)}\big(m(t)\big)\Big\}\tilde{B}^m(t) \\ &+ \frac{1}{2}\frac{1}{\vartheta}\frac{\delta}{1-\delta}\bar{h}^{(1)}\big(m(t)\big) = 0, \tilde{B}^m(T) = 0 \qquad (4.28) \\ &\tilde{A}^m_t(t) + \frac{1}{2}H^{(0)}\big(m(t)\big)\big(\tilde{B}^m(t)\big)^2 + \Big\{\frac{\delta}{1-\delta}\bar{g}^{(0)}\big(m(t)\big) + \bar{k}^{(0)}\big(m(t)\big)\Big\}\tilde{B}^m(t) = 0, \\ &\tilde{A}^m(T) = 0. \qquad (4.29) \end{aligned}$$

So in order to obtain the solution of the HJB equation we only need to solve the Riccati Equation (4.28) and do the integration in Equation (4.29):

$$\tilde{A}^m(t) = \int_t^T \frac{1}{2}H^{(0)}\big(m(s)\big)\big(\tilde{B}^m(s)\big)^2 + \Big\{\frac{\delta}{1-\delta}\bar{g}^{(0)}\big(m(s)\big) + \bar{k}^{(0)}\big(m(s)\big)\Big\}\tilde{B}^m(s)\mathrm{d}s.$$

Then, the expression for Φ^m from (4.23) follows directly.
The verification result follows as an application of Proposition 4.3.

□

Remark 4.8 *Observe that the parameters in the equations for $\tilde{B}^m$ are time-dependent but piece-wise constant. So we are looking for a continuous, piece-wise continuously differentiable solution. If we know the solution for constant parameters and inhomogeneous terminal condition a recursive solution method with finitely many steps is possible. A result on the solvability of Equation (4.20) with constant parameters is summarized in Corollary 2.44. We will apply this result in Section 4.3 to construct step-wise a solution for the Heston model with time-dependent coefficients.*

So, when applying the derived optimization results to a special model of interest, one basically needs to solve Equation (4.20) and to verify that $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ is a martingale. The latter might be in general quite laborious. That is why we present in what follows a special case of the general time-dependent model, for which this can be easily shown. To this aim we additionally assume for Model (4.14) that:

$$H^{(0)} = h^{(0)} = \bar{g}^{(0)} = 0, \bar{k}^{(0)} \geq 0.$$

Then the martingale property and thus the verification result follow easily from a general statement about exponentials of affine processes, given in Corollary 3.4 from [62]. For convenience, this result is summarized in Theorem 2.34. In the following proposition we apply it for the verification result.

Proposition 4.9 (Verification via Theorem 2.34)
Let Φ^m be as given by Expression (4.23). Furthermore, assume that $H^{(0)} = h^{(0)} = \bar{g}^{(0)} = 0$ and $\bar{k}^{(0)} \geq 0$. Then the optimal portfolio strategy is as given in Equation (4.24) and Φ^m is the corresponding value function.

The proof is given in Appendix B.

Remark 4.10
Note that Theorem 2.34 can be applied also if X_1 does not appear in the volatility of the traded asset, nor in the excess return, but just in the riskless interest rate. As in this case we will have stochastic interest rates, we should consider additionally trading in a bond. So, we end up with a multidimensional model. This example is considered in Section 5.3.

To sum up, when we consider a special example for a time-dependent model with $H^{(0)} = h^{(0)} = \bar{g}^{(0)} = 0$ and $\bar{k}^{(0)} \geq 0$, we just need to find a continuous, piece-wise differentiable functions $\tilde{B}^m$ that solves Equations (4.20). Then we can apply the just proved theorem and derive directly the optimal solution. In Section 4.3.1 we will do this for the time-dependent Heston model.

Remark 4.11
Let us summarize the different possibilities for proving a verification result that we have considered so far. After finding a sufficiently differentiable solution to the HJB equation one can either prove the martingale condition directly for the derived value function (see Theorem 4.2) or prove it first for the time-dependent model and then apply Theorem 3.5. In the latter case, Theorem 2.34 can be useful, when applicable (as in Theorem 4.9). As we will see in Chapters 5 and 6, the same methods can be adopted also in the multidimensional case and for the HARA utility function.

After deriving all necessary results for the time-dependent model, we can proceed with the Markov-modulated model. We consider separately the cases with and without correlation between the Brownian motions driving the stock price and the stochastic factor in Sections 4.2.1 and 4.2.2, respectively.

4.2 Markov-modulated model

In this section we consider the Markov-modulated Model (4.1). For the case without leverage (i.e. $\rho = 0$), we provide the solution up to a simple expectation only over the probability measure of the Markov chain, which can be computed very efficiently (see Corollary 4.12). If we assume separability of the value function in the Markov chain and the stochastic volatility we can further simplify this expression and derive an explicit solution even for the case wherein the Brownian motions exhibit instantaneous correlation (see Theorem 4.13).

4.2.1 Solution with no correlation

Let us first consider the case with $\rho = 0$. Observe that although the Brownian motions do not exhibit instantaneous correlation, the two processes are correlated by the joint Markov chain.
In the considered case the PDE system for function f has the following simpler form:

$$\begin{aligned}
&f_t(t,x,e_i) + f(t,x,e_i)\underbrace{\delta\Big\{r(e_i) + \frac{1}{2}\frac{1}{1-\delta}\Big(\frac{\mu_1(x,e_i)-r(e_i)}{\Sigma_1(x,e_i)}\Big)^2\Big\}}_{=:g(x,e_i)} + f_x(t,x,e_i)\mu^X(x,e_i)\\
&+\frac{1}{2}f_{xx}(t,x,e_i)(\Sigma^X)^2(x,e_i) = -\sum_{z=1}^{I} q_{i,z} f(t,x,e_z)\\
&f(T,x,e_i) = 1, \forall\, i \in \{1,\dots,I\}.
\end{aligned} \tag{4.30}$$

As we are dealing with a system of linear PDEs we can apply Corollary 2.76, if its conditions are fulfilled for process X and function $g(x,e_i)$, which is defined in the equation above, and derive the following probabilistic representation for f:

$$\begin{aligned}
f(t,x,e_i) =&\mathbb{E}\Big[\exp\Big\{\int_t^T \delta\Big(r\big(\mathcal{MC}(s)\big)\\
&+\frac{1}{2}\frac{1}{1-\delta}\Big(\frac{\mu_1\big(X(s),\mathcal{MC}(s)\big)-r\big(\mathcal{MC}(s)\big)}{\Sigma_1\big(X(s),\mathcal{MC}(s)\big)}\Big)^2\Big)\mathrm{d}s\Big\}\\
&\Big|X(t)=x,\mathcal{MC}(t)=e_i\Big].
\end{aligned} \tag{4.31}$$

Now transform this expression by the tower rule for conditional expectations as follows:

$$\begin{aligned}
f(t,x,e_i) =&\mathbb{E}\Big[\exp\Big\{\int_t^T \delta\Big(r\big(\mathcal{MC}_{t,e_i}(s)\big)\\
&+\frac{1}{2}\frac{1}{1-\delta}\Big(\frac{\mu_1\big(X_{t,x,e_i}(s),\mathcal{MC}_{t,e_i}(s)\big)-r\big(\mathcal{MC}_{t,e_i}(s)\big)}{\Sigma_1\big(X_{t,x,e_i}(s),\mathcal{MC}_{t,e_i}(s)\big)}\Big)^2\Big)\mathrm{d}s\Big\}\Big]\\
=&\mathbb{E}\Big[\mathbb{E}\Big[\exp\Big\{\int_t^T \delta\Big(r\big(\mathcal{MC}_{t,e_i}(s)\big)\\
&+\frac{1}{2}\frac{1}{1-\delta}\Big(\frac{\mu_1\big(X_{t,x,e_i}(s),\mathcal{MC}_{t,e_i}(s)\big)-r\big(\mathcal{MC}_{t,e_i}(s)\big)}{\Sigma_1\big(X_{t,x,e_i}(s),\mathcal{MC}_{t,e_i}(s)\big)}\Big)^2\Big)\mathrm{d}s\Big\}|\mathscr{F}_T^{\mathcal{MC}}\Big]\Big]\\
=&\mathbb{E}\Big[f^{\mathcal{MC}_{t,e_i}}(t,x)\Big] = \mathbb{E}\Big[f^{\mathcal{MC}}(t,x)\Big|\mathcal{MC}(t)=e_i\Big],
\end{aligned} \tag{4.32}$$

for all $(t,x,e_i) \in [0,T] \times D^X \times \mathcal{E}$, where f^m denotes for all $m \in \mathbb{M}$ the solution of System (4.19) in the time-dependent model induced by m, thus $f^{\mathcal{MC}}(t,x)$ is an $\mathscr{F}_T^{\mathcal{MC}}$-measurable random variable. Observe that here we have used the independence of $\mathcal{MC}$ and W^X. So, the candidate for the value function in the considered model has the following form:

$$\begin{aligned}\Phi(t,v,x,e_i) = \frac{v^\delta}{\delta} f(t,x,e_i) = \frac{v^\delta}{\delta}\mathbb{E}\Big[f^{\mathcal{MC}_{t,e_i}}(t,x)\Big] = \mathbb{E}\Big[\frac{v^\delta}{\delta} f^{\mathcal{MC}_{t,e_i}}(t,x)\Big] \\ = \mathbb{E}[\Phi^{\mathcal{MC}_{t,e_i}}(t,v,x)] = \mathbb{E}[\Phi^{\mathcal{MC}}(t,v,x)|\mathcal{MC}(t)=e_i],\end{aligned} \tag{4.33}$$

where for each $m \in \mathbb{M}$, Φ^m denotes the value function in the time-dependent model. This expression is in accordance with Theorem 3.5. So, instead of proving that the Markov-switching Feynman-Kac theorem is applicable in order to show that Expression (4.32) indeed solves Equation (4.30), we can directly prove that Φ as given by (4.33) is the value function for the considered problem. To this aim Theorem 3.5 can be applied, as stated in the next corollary.

Corollary 4.12 (Solution with no correlation)
Consider Model (4.1) and set $\rho = 0$. Assume that the value function in the corresponding time-dependent model is given by:

$$\begin{aligned}\Phi^m(t,v,x) &= \frac{v^\delta}{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T g(X^m(s),m(s))\mathrm{d}s\Big\}\Big|X^m(t)=x\Big] \\ &=: \frac{v^\delta}{\delta} f^m(t,x),\end{aligned}$$

and the optimal investment strategy by:

$$\bar{\pi}^m(t) = \frac{1}{1-\delta}\frac{\mu_1 - r}{(\Sigma_1)^2}\Big|_{(X^m(t),m(t))}$$

Then, the value function in the Markov-modulated model has the following form:

$$\begin{aligned}\Phi(t,v,x,e_i) =& \mathbb{E}[\Phi^{\mathcal{MC}}(t,v,x)|\mathcal{MC}(t)=e_i] \\ =& \frac{v^\delta}{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T g(X(s),\mathcal{MC}(s))\mathrm{d}s\Big\}\Big|X(t)=x,\mathcal{MC}(t)=e_i\Big] \\ =:& \frac{v^\delta}{\delta} f(t,x,e_i),\end{aligned} \tag{4.34}$$

and the optimal portfolio strategy is

$$\bar{\pi}(t) = \frac{1}{1-\delta}\frac{\mu_1 - r}{(\Sigma_1)^2}\Big|_{(X(t),\mathcal{MC}(t))}$$

Note that in the case of $\rho = 0$ the optimal strategy consists only on the mean-variance portfolio.

Proof
Verify that the optimal portfolio in the time-dependent model does not depend on the whole path of function m and apply Theorem 3.5.

□

After deriving the general solution we are now interested in simplifying the probabilistic representation from Corollary 4.12. This can be done by assuming separability of the value function in the state of the Markov chain and the stochastic factor. As in this case one can even allow for correlation, we directly consider the general model with $\rho \neq 0$ in what follows. For the result without leverage, just set $\rho = 0$ in the subsequent analysis.

4.2.2 Solution with correlation

In this section we consider the general Model (4.1), so function f is characterized by System (4.12). We were not able to find its solutions in general, mainly because of the nonlinear term. Unfortunately a transformation like the one in Section 4.1 does not work in the most general case, because here we have a system of coupled PDEs. What is more, Theorem 3.5 cannot be applied in this case, as in general $\bar{\pi}^m$ depends on the whole path of m. The key to find a solution in this case is to assume a separable exponential ansatz:

$$f(t,x,e_i) = \xi(t,e_i)\exp\big\{B(t)x\big\}, \tag{4.35}$$

for some functions $\xi : [0,T] \times \mathcal{E} \to \mathbb{R}$, $B : [0,T] \to \mathbb{R}$, which of course implies certain restrictions on the model parameters: we assume that $\bar{h}^{(1)}$, $K^{(1)}$, $H^{(11)}$ and $G^{(1)}$ are constants. The solution in this case and a link to the verification results from Theorem 4.2 and Theorem 3.5 are precisely stated in the following theorem.

Theorem 4.13 (Solution with correlation)
Consider Model (4.1) and let $\bar{h}^{(1)}$, $K^{(1)}$, $H^{(11)}$ and $G^{(1)}$ be constants. Set $\vartheta = \frac{1-\delta}{1-\delta+\delta\rho^2}$. Assume that the following equation:

$$B_t + \frac{1}{2}\frac{\delta}{1-\delta}\bar{h}^{(1)} + B\big[\frac{\delta}{1-\delta}G^{(1)} - K^{(1)}\big] + \frac{1}{2}B^2\frac{H^{(11)}}{\vartheta} = 0, B(T) = 0 \tag{4.36}$$

possesses a differentiable solution B. Then the solution of the corresponding HJB equation is given by:

$$\begin{aligned}\Phi(t,v,x,e_i) =& \frac{v^\delta}{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T w(s,\mathcal{MC}(s))\mathrm{d}s\Big\}\Big|\mathcal{MC}(t)=e_i\Big]\exp\{B(t)x\} \\ =:& \frac{v^\delta}{\delta}\xi(t,e_i)\exp\{B(t)x\}, \forall (t,v,x,e_i) \in [0,T]\times[0,\infty)\times D^X \times \mathcal{E},\end{aligned} \tag{4.37}$$

where function w is given by:

$$w(t,e_i) = \delta r(e_i) + \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}(e_i) + B\big[\frac{\delta}{1-\delta}\bar{g}^{(0)}(e_i) + \bar{k}^{(0)}(e_i)\big] + \frac{1}{2}B^2\frac{H^{(0)}(e_i)}{\vartheta}.$$

The maximum point for the HJB equation is given by:

$$\bar{\pi} = \frac{1}{1-\delta}\Big\{\frac{(\mu_1 - r)}{(\Sigma_1)^2} + \rho\frac{\Sigma^X}{\Sigma_1}B\Big\}\Big|_{(t,(X(t),\mathcal{MC}(t))}.$$

Note that $\Phi \in \mathscr{C}^{1,2,2}$ *for all* $e_i \in \mathcal{E}$.
Denote:

$$a := \sqrt{\big(K^{(1)} - \frac{\delta}{1-\delta}G^{(1)}\big)^2 - \frac{\delta}{1-\delta}\bar{h}^{(1)}\frac{H^{(11)}}{\vartheta}}$$
$$c := \frac{K^{(1)} - \frac{\delta}{1-\delta}G^{(1)} - a}{K^{(1)} - \frac{\delta}{1-\delta}G^{(1)} + a},$$

and assume that:

$$\frac{H^{(11)}}{\vartheta} > 0, K^{(1)} - \frac{\delta}{1-\delta}G^{(1)} > 0$$
$$\frac{\delta}{1-\delta}\bar{h}^{(1)} < \frac{\vartheta\big(K^{(1)} - \frac{\delta}{1-\delta}G^{(1)}\big)^2}{H^{(11)}}$$
$$0 \le \vartheta\frac{a + K^{(1)} - \frac{\delta}{1-\delta}G^{(1)}}{H^{(11)}}.$$

Then function B *is given by:*

$$B(t) = \begin{cases} \vartheta\frac{-c(K^{(1)} - \frac{\delta}{1-\delta}G^{(1)} + a)\exp\{-a(T-t)\} + K^{(1)} - \frac{\delta}{1-\delta}G^{(1)} - a}{\bar{H}^{(11)}(1 - c\exp\{-a(T-t)\})} & \text{for } 0 < \vartheta\frac{a+K^{(1)} - \frac{\delta}{1-\delta}G^{(1)}}{H^{(11)}} \\ 0 & \text{for } 0 = \vartheta\frac{a+K^{(1)} - \frac{\delta}{1-\delta}G^{(1)}}{H^{(11)}} \end{cases}.$$

Further, consider an arbitrary but fixed path m *of the Markov chain and consider the time-dependent model associated with* m. *Then, the solution* Φ^m *to its HJB equation is given by:*

$$\Phi^m(t,v,x) = \frac{v^\delta}{\delta}\exp\Big\{\int_t^T w\big(s,m(s)\big)\mathrm{d}s\Big\}\exp\big\{B(t)x\big\}, \tag{4.38}$$

and

$$\bar{\pi}^m = \frac{1}{1-\delta}\Big\{\frac{\mu_1 - r}{(\Sigma_1)^2} + \rho\frac{\Sigma^X}{\Sigma_1}B\Big\}\Big|_{(t,(X^m(t),m(t))}.$$

Assume that one of the following conditions holds:

i) $\{\Phi^m(t, V^{m,\bar{\pi}^m}(t), X^m(t))\}_{t\in[0,T]}$ *is a martingale,*

ii) $\{\Phi(t, V^{\bar{\pi}}(t), X(t), \mathcal{MC}(t))\}_{t\in[0,T]}$ *is a martingale.*

Then, function Φ is the value function for the model with Markov switching and the optimal portfolio is given by $\bar{\pi}$.

Proof
As already mentioned we consider the ansatz $f(t,x,e_i) = \xi(t,e_i)\exp\big\{B(t)x\big\}$ and substitute it in PDE (4.12). This leads to the following system of PDEs for function $\xi(t,e_i)$:

$$\xi(t,e_i)\underbrace{\Big[\delta r(e_i) + \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}(e_i) + B\big[\frac{\delta}{1-\delta}\bar{g}^{(0)}(e_i) + \bar{k}^{(0)}(e_i)\big] + \frac{1}{2}B^2\frac{H^{(0)}(e_i)}{\vartheta}\Big]}_{=w(t,e_i)}$$
$$+\,\xi_t(t,e_i) = -\sum_{z=1}^{I} q_{i,z}\xi(t,e_z), \xi(T,e_i) = 1, \forall i = 1,\dots,I, \tag{4.39}$$

where $\vartheta = \frac{1-\delta}{1-\delta+\delta\rho^2}$, and the following ODE with constant parameters for $B(t)$:

$$B_t + \frac{1}{2}\frac{\delta}{1-\delta}\bar{h}^{(1)} + B\big[\frac{\delta}{1-\delta}G^{(1)} - K^{(1)}\big] + \frac{1}{2}B^2\frac{H^{(11)}}{\vartheta} = 0, B(T) = 0. \tag{4.40}$$

Note that Equation (4.40) can be solved by Corollary 2.44. Furthermore, Corollary 2.78 provides the following probabilistic representation for function $\xi(t,e_i)$:

$$\xi(t,e_i) = \mathbb{E}\Big[\exp\Big\{\int_t^T w(s,\mathcal{MC}(s))\mathrm{d}s\Big\}\Big|\mathcal{MC}(t) = e_i\Big], \forall e_i \in \mathcal{E}, \tag{4.41}$$

as $w(t,e_i)$ and $\frac{\partial}{\partial t}w(t,e_i)$ are continuous in t. Observe that ξ is continuous and differentiable w.r.t t.
For the proof of the HJB solution (4.38) in the corresponding time-dependent model observe that $\tilde{B}(t) := \frac{1}{\vartheta}B(t)$ solves Equation (4.20) for the considered model specification and apply Proposition 4.7. It is easily verified that the term in the integral in Equation (4.22) corresponds to $w(s,m(s))$ with $B = B^m$. So, Φ^m solves the time-dependent HJB equation. Furthermore, Assumption i) implies by Proposition 4.3 that Φ^m is the value function for the time-dependent model. Observe that for function Φ it holds:

$$\Phi(t,v,x,e_i) = \mathbb{E}[\Phi^{\mathcal{MC}}(t,v,x)|\mathcal{MC}(t) = e_i].$$

Further, the optimal strategy in the time-dependent model is given by:

$$\bar{\pi}^m = \frac{1}{1-\delta}\Big\{\frac{\mu_1 - r}{(\Sigma_1)^2} + \rho\frac{\Sigma^X}{\Sigma_1}B\Big\}\Big|_{(t,(X^m(t),m(t))}.$$

As it does not depend on the whole path of m, but only on its current level, we can apply Theorem 3.5 and prove the statement.
Assumption ii) allows us to apply directly Theorem 4.2 and conclude that Φ is the value function for the Markov-switching model and that the optimal portfolio is given by $\bar{\pi}$.

□

Remark 4.14
Alternatively to the probabilistic representation for ξ the solution for System (4.39) can be approximated by the so-called Magnus exponential series, for details see [78]. Applying a numerical solving scheme is also possible. Another possibility is presented in the following lemma for the case of two possible states of the Markov chain, i.e. $I = 2$.

Lemma 4.15
Let $I = 2$ and assume that the Riccati equation:

$$c_t(t) + \big[-w(t,e_1) + w(t,e_2) - q_{1,1} + q_{2,2}\big]c(t) - q_{1,2}c^2(t) + q_{2,1} = 0, c(T) = 1,$$

has a unique solution $c(t)$. Then, the solution of System (4.39) is given by:

$$\begin{aligned}\xi(t,e_1) &= \exp\Big\{\int_t^T w(s,e_1) + q_{1,1} + q_{1,2}c(s)\mathrm{d}s\Big\}\\ \xi(t,e_2) &= c(t)\xi(t,e_1).\end{aligned}$$

Proof
For $I = 2$ we have the following system of two coupled ODEs:

$$\xi_t(t,e_1) + \xi(t,e_1)w(t,e_1) = -q_{1,1}\xi(t,e_1) - q_{1,2}\xi(t,e_2), \xi(T,e_1) = 1 \tag{4.42}$$
$$\xi_t(t,e_2) + \xi(t,e_2)w(t,e_2) = -q_{2,1}\xi(t,e_1) - q_{2,2}\xi(t,e_2), \xi(T,e_2) = 1. \tag{4.43}$$

Denote $c(t) := \frac{\xi(t,e_2)}{\xi(t,e_1)}$ and substitute $\xi(t,e_2) = c(t)\xi(t,e_1)$ in Equation (4.42) to obtain:

$$\xi_t(t,e_1) = -\big[w(t,e_1) + q_{1,1} + q_{1,2}c(t)\big]\xi(t,e_1), \xi(T,e_1) = 1.$$

It follows that:

$$\xi(t,e_1) = \exp\Big\{\int_t^T w(s,e_1) + q_{1,1} + q_{1,2}c(s)\mathrm{d}s\Big\}.$$

Furthermore, compute:

$$\xi_t(t,e_2) = c_t(t)\xi(t,e_1) + c(t)\xi_t(t,e_1) = \Big\{c_t(t) - c(t)\big[w(t,e_1) + q_{1,1} + q_{1,2}c(t)\big]\Big\}\xi(t,e_1),$$

and substitute this result in Equation (4.43). Canceling of $\xi(t,e_1)$ on both sides leads then to:

$$c_t(t) + \big[-w(t,e_1) + w(t,e_2) - q_{1,1} + q_{2,2}\big]c(t) - q_{1,2}c^2(t) + q_{2,1} = 0, c(T) = 1,$$

which is a Riccati equation with one time-dependent parameter and can be solved numerically.

□

Remark 4.16
By analogous computations one can "decouple" System (4.39) also for $I = 3$. In this case, solving the whole system can be reduced to solving just one nonlinear ODE of second grade.

In the next remark we point out two further special cases, where $\xi(t, e)$ can be computed explicitly.

Remark 4.17
Recall that throughout this section it is assumed that $\bar{h}^{(1)}$, $K^{(1)}$, $H^{(11)}$ and $G^{(1)}$ are constants.

i) If additionally we assume that process X does not depend on $\mathcal{MC}$, i.e. $\bar{k}^{(0)}$ and $H^{(0)}$ are constants, and that $\rho = 0$, then function ξ is given for all $(t, e) \in [0, T] \times \mathcal{E}$ by:

$$\begin{aligned}\xi(t, e_i) =& \mathbb{E}\Big[\exp\Big\{\int_t^T w(s, \mathcal{MC}(s))\mathrm{d}s\Big\}\Big|\mathcal{MC}(t) = e_i\Big]\\ =& \exp\Big\{\int_t^T B(s)\bar{k}^{(0)} + \frac{1}{2}B^2(s)H^{(0)}\mathrm{d}s\Big\}\\ & \mathbb{E}\Big[\exp\Big\{\int_t^T \underbrace{\delta r(\mathcal{MC}(s)) + \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}(\mathcal{MC}(s))}_{=:\bar{w}(\mathcal{MC}(s))}\mathrm{d}s\Big\}\Big|\mathcal{MC}(t) = e_i\Big]\\ =& \exp\Big\{\int_t^T B(s)\bar{k}^{(0)} + \frac{1}{2}B^2(s)H^{(0)}\mathrm{d}s\Big\}\\ & \cdot\left\langle \exp\Big\{\big[Q' + diag\big(\bar{w}(e_1), \dots, \bar{w}(e_I)\big)\big](T-t)\Big\}e_i, \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}\right\rangle,\end{aligned}$$

where in the last equation we have used Lemma 2.68.

ii) If we allow for $\rho \neq 0$ and assume that only parameter r depends on the Markov chain, we obtain:

$$\begin{aligned}\xi(t, e_i) =& \mathbb{E}\Big[\exp\Big\{\int_t^T w(s, \mathcal{MC}(s))\mathrm{d}s\Big\}\Big|\mathcal{MC}(t) = e_i\Big]\\ =& \exp\Big\{\int_t^T \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)} + B(s)\big[\frac{\delta}{1-\delta}\bar{g}^{(0)} + \bar{k}^{(0)}\big] + \frac{1}{2}B^2(s)\frac{H^{(0)}}{\vartheta}\mathrm{d}s\Big\}\\ & \mathbb{E}\Big[\exp\Big\{\int_t^T \delta r(\mathcal{MC}(s))\mathrm{d}s\Big\}\Big|\mathcal{MC}(t) = e_i\Big]\end{aligned}$$

$$= \exp\Big\{\int_t^T \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)} + B(s)\big[\frac{\delta}{1-\delta}\bar{g}^{(0)} + \bar{k}^{(0)}\big] + \frac{1}{2}B^2(s)\frac{H^{(0)}}{\vartheta}\mathrm{d}s\Big\}$$
$$\cdot \left\langle \exp\Big\{\big[Q' + diag\big(\delta r(e_1), \ldots, \delta r(e_I)\big)\big](T-t)\Big\}e_i, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\rangle.$$

4.3 Example: Markov-modulated Heston model

In this section we apply the derived results to the famous Heston model (see Example 2.49), where the stochastic factor follows a mean-reverting CIR process and is interpreted as the stochastic volatility of the asset price process. The original model was introduced in [58]. Optimal portfolios under the original Heston model are derived in [74] and [61].
In what follows we extend this framework to Markov-switching parameters. More precisely, we consider the following model:

$$\begin{aligned}
&\mathrm{d}P_0(t) = P_0(t)r\big(\mathcal{MC}(t)\big)\mathrm{d}t \\
&\mathrm{d}P_1(t) = P_1(t)\Big[r\big(\mathcal{MC}(t)\big) + \lambda\big(\mathcal{MC}(t)\big)X(t)\mathrm{d}t + \nu\big(\mathcal{MC}(t)\big)\sqrt{X(t)}\mathrm{d}W^P(t)\Big] \\
&\mathrm{d}X(t) = \kappa\big(\mathcal{MC}(t)\big)(\theta\big(\mathcal{MC}(t)\big) - X(t))\mathrm{d}t + \chi\big(\mathcal{MC}(t)\big)\sqrt{X(t)}\mathrm{d}W^X(t) \\
&\mathrm{d}\langle W^P, W^X\rangle(t) = \rho\mathrm{d}t,
\end{aligned} \tag{4.44}$$

with $r, \lambda, \nu, \kappa, \theta, \chi : \mathcal{E} \to \mathbb{R}$ being deterministic functions with $\kappa(e_i), \theta(e_i), \chi(e_i) > 0$, for all $e_i \in \mathcal{E}$. Furthermore, it is assumed that $2\kappa(e_i)\theta(e_i) \geq \chi^2(e_i)$, for all $e_i \in \mathcal{E}$, in order to assure the positivity of process X^m. Observe that this framework corresponds to the following parameter specifications in terms of the notation from Model (4.1):

$$\begin{aligned}
\Sigma_1(x, e_i) &= \nu(e_i)\sqrt{x} && \\
\frac{\mu_1(x, e_i) - r(e_i)}{\Sigma_1(x, e_i)} &= \frac{\lambda(e_i)}{\nu(e_i)}\sqrt{x} && \Rightarrow h^{(0)}(e_i) = 0, \;\; \bar{h}^{(1)}(e_i) = \frac{\lambda^2(e_i)}{\nu^2(e_i)} \\
\mu^X(x, e_i) &= \kappa(e_i)\big(\theta(e_i) - x\big) && \Rightarrow \bar{k}^{(0)}(e_i) = \kappa(e_i)\theta(e_i), \\
& && \quad K^{(1)}(e_i) = \kappa(e_i) \\
\Sigma^X(x, e_i) &= \chi(e_i)\sqrt{x} && \Rightarrow H^{(0)}(e_i) = 0, \\
& && \quad H^{(11)}(e_i) = \chi^2(e_i) \\
\rho\frac{\mu_1(x, e_i) - r(e_i)}{\Sigma_1(x, e_i)}\Sigma^X(x, e_i) &= \rho\frac{\lambda(e_i)}{\nu(e_i)}\chi(e_i)x && \Rightarrow \bar{g}^{(0)}(e_i) = 0, \\
& && \quad G^{(1)}(e_i) = \rho\frac{\lambda(e_i)}{\nu(e_i)}\chi(e_i),
\end{aligned} \tag{4.45}$$

for all $(t, e_i) \in [0, T] \times \mathcal{E}$. A similar model is used for pricing of volatility swaps in [40].
As before, we first derive the optimal portfolio strategy and the value function in the corresponding time-dependent model (Section 4.3.1). Afterwards, in Section 4.3.2, we present the results for the Markov-modulated model without correlation between the Brownian motion driving the risky asset price process and the one for the volatility. Section 4.3.3 deals with the case with correlation.

4.3.1 Time-dependent Heston model (TDH)

The corresponding time-dependent Heston model is stated as follows:

$$\begin{aligned} \mathrm{d}P_0(t) &= P_0(t) r\big(m(t)\big)\mathrm{d}t, \\ \mathrm{d}P_1(t) &= P_1(t)\Big[r\big(m(t)\big) + \lambda\big(m(t)\big)X^m(t)\mathrm{d}t + \nu\big(m(t)\big)\sqrt{X^m(t)}\mathrm{d}W^P(t)\Big], \\ \mathrm{d}X^m(t) &= \kappa\big(m(t)\big)(\theta\big(m(t)\big) - X^m(t))\mathrm{d}t + \chi\big(m(t)\big)\sqrt{X^m(t)}\mathrm{d}W^X(t), \\ \mathrm{d}\langle W^P, W^X\rangle(t) &= \rho\ \mathrm{d}t, \end{aligned} \tag{4.46}$$

where function m is defined as in Equation (3.13), and $\kappa(e_i), \theta(e_i), \chi(e_i) > 0$ are as in (4.44). A similar model is presented and motivated in the context of calibration and derivatives pricing in [36] and [82].
Recall from Proposition 4.7 that in order to find a solution for the value function we either have to solve Equation (4.20) for B^m or calculate explicitly the following expectation:

$$h^m(t, x) = \mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{\vartheta} g(s, \tilde{X}^m(s), m(s))\mathrm{d}s\Big\}\Big|\tilde{X}^m(t) = x\Big]. \tag{4.47}$$

From Equation (4.19) we know that the drift $\tilde{\mu}_X$ of the modified process $\tilde{X}^m$ is given by:

$$\begin{aligned} \tilde{\mu}_X(\tilde{X}^m(t), m(t)) &= \kappa(m(t))\big(\theta(m(t)) - \tilde{X}^m(t)\big) + \frac{\delta}{1-\delta}\rho\frac{\chi(m(t))\lambda(m(t))}{\nu(m(t))}\tilde{X}^m(t) \\ &=: \tilde{\kappa}(m(t))\big(\tilde{\theta}(m(t)) - \tilde{X}^m(t)\big). \end{aligned}$$

So, $\tilde{X}^m$ is also a CIR process with piece-wise constant coefficients. Thus, the expectation above can be calculated in an explicit form by applying step-wise Lemma 2.42. We will do so in the next proposition to derive the solution to our optimization problem.

Proposition 4.18 (Solution and verification in TDH)
Consider Model (4.46) and assume that $\tilde{\kappa}(e_i) > 0$, for all $e_i \in \mathcal{E}$. Let $e_i \in \mathcal{E}$ be arbitrary but fixed. For any constants α and β, such that:

$$\begin{aligned} \beta &< \frac{\tilde{\kappa}^2(e_i)}{2\chi^2(e_i)} \\ \alpha &< \frac{\tilde{\kappa}(e_i) + \tilde{a}(e_i)}{\chi^2(e_i)}, \end{aligned}$$

with:

$$\tilde{a}(e_i) := \sqrt{\tilde{\kappa}(e_i)^2 - 2\beta\chi(e_i)^2},$$

define the following two functions:

$$\tilde{A}^{\alpha,\beta,e_i}(\tau) := \frac{\tilde{\kappa}(e_i)\tilde{\theta}(e_i)(\tilde{\kappa}(e_i) - \tilde{a}(e_i))}{\chi(e_i)^2}\tau - \frac{2\tilde{\kappa}(e_i)\tilde{\theta}(e_i)}{\chi(e_i)^2}\ln\Big\{\frac{1-\tilde{c}(e_i)\exp(-\tilde{a}(e_i)\tau)}{1-\tilde{c}(e_i)}\Big\}$$
$$\tilde{B}^{\alpha,\beta,e_i}(\tau) := \frac{-\tilde{c}(e_i)(\tilde{\kappa}(e_i)+\tilde{a}(e_i))\exp(-\tilde{a}(e_i)\tau) + \tilde{\kappa}(e_i) - \tilde{a}(e_i)}{\chi(e_i)^2\big(1-\tilde{c}(e_i)\exp(-\tilde{a}(e_i)\tau)\big)},$$

where

$$\tilde{c}(e_i) := \frac{-\alpha\chi(e_i)^2 + \tilde{\kappa}(e_i) - \tilde{a}(e_i)}{-\alpha\chi(e_i)^2 + \tilde{\kappa}(e_i) + \tilde{a}(e_i)}.$$

Assume the following conditions on the model parameters in (4.46):

$$\frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(e_i)\big)^2}{\big(\nu(e_i)\big)^2} < \frac{\tilde{\kappa}^2(e_i)}{2\chi^2(e_i)}, \forall e_i \in \mathcal{E} \tag{4.48}$$

$$\max_{e_i\in\mathcal{E}}\Big\{\frac{\tilde{\kappa}(e_i)-\tilde{a}(e_i)}{\chi^2(e_i)}\Big\} \leq \min_{e_i\in\mathcal{E}}\Big\{\frac{\tilde{\kappa}(e_i)+\tilde{a}(e_i)}{\chi^2(e_i)}\Big\}. \tag{4.49}$$

Then the following expressions are well-defined:

$$A_k(\tau) := \tilde{A}^{\alpha_k,\beta_k,m_k}(\tau), B_k(\tau) := \tilde{B}^{\alpha_k,\beta_k,m_k}(\tau), \forall k = 0,\dots,K,$$

where

$$\beta_k := \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(t_k))\big)^2}{\big(\nu(m(t_k))\big)^2}, \forall k = 0,\dots,K$$
$$\alpha_k := B_{k+1}(\tau_{k+1}),\ \textit{for}\ \tau_k := t_{k+1} - t_k, \forall k = 0,\dots,K-1$$
$$\alpha_K := 0.$$

The solution of the corresponding HJB system is given for all $(t,v,x) \in [0,T) \times \mathbb{R}_{\geq 0} \times D^X$ *by:*

$$\Phi^m(t,v,x) = \frac{v^\delta}{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{\vartheta}g(s,\tilde{X}^m(s),m(s))\mathrm{d}s\Big\}\Big|\tilde{X}(t)=x\Big]^\vartheta$$
$$= \frac{v^\delta}{\delta}\xi^m(t)\exp\big\{B^m(t)x\big\},$$

where:

$$\xi^m(t) := \exp\Big\{\int_t^T \delta r(m(s))\mathrm{d}s + \vartheta\sum_{k=0}^K\Big\{A_k(t_{k+1}-t) + \sum_{z=k+1}^K A_z(\tau_z)\Big\}\mathbb{1}_{t\in[t_k,t_{k+1})}\Big\}$$

$$B^m(t) := \vartheta\sum_{k=0}^K\big\{B_k(t_{k+1}-t)\big\}\mathbb{1}_{t\in[t_k,t_{k+1})}.$$

Further, $\Phi^m(t,v,x)$ is the value function of the optimization problem and the optimal portfolio strategy is given for all $t \in [0,T]$ by:

$$\bar{\pi}^m(t) = \frac{1}{1-\delta}\Big\{\frac{\lambda(m(t))}{\big(\nu(m(t))\big)^2} + \rho\frac{\chi(m(t))}{\nu(m(t))}B^m(t)\Big\}.$$

The proof can be found in Appendix B.

Remark 4.19
Observe that by Lemma 2.43, $B^m > 0$ for $\delta > 0$ and $B^m < 0$ for $\delta < 0$. We will analyze this in detail in Section 4.3.4.

Remark 4.20
An alternative verification theorem for the Heston model with constant coefficients is presented in [74]. It relies on the specific form of the solution in this example. However, it is very technical and requires some more restrictions on the model parameters in the case with time-dependent parameters.

4.3.2 Markov-modulated Heston model with no correlation (MMH^0)

We continue with the Markov-modulated Heston model with no correlation. Consider Model (4.44) and set $\rho = 0$. Based on the results for the time-dependent model, the solution of the HJB equation in the Markov-switching model can be derived as shown in the following proposition.

Proposition 4.21 (Solution and verification in MMH^0)
Consider Model (4.44) with $\rho = 0$. Denote analogously as before for all $e_i \in \mathcal{E}$:

$$\begin{aligned}
a(e_i) &:= \sqrt{\kappa(e_i)^2 + 2\beta\chi(e_i)^2}\\
c(e_i) &:= \frac{\alpha\chi(e_i)^2 + \kappa(e_i) - a(e_i)}{\alpha\chi(e_i)^2 + \kappa(e_i) + a(e_i)}\\
A^{\alpha,\beta,e_i}(\tau) &:= -\frac{\kappa(e_i)\theta(e_i)(\kappa(e_i) - a(e_i))}{\chi(e_i)^2}\tau + \frac{2\kappa(e_i)\theta(e_i)}{\chi(e_i)^2}\ln\Big\{\frac{1 - c(e_i)\exp(-a(e_i)\tau)}{1-c(e_i)}\Big\}\\
B^{\alpha,\beta,e_i}(\tau) &:= -\frac{-c(e_i)(\kappa(e_i) + a(e_i))\exp(-a(e_i)\tau) + \kappa(e_i) - a(e_i)}{\chi(e_i)^2\big(1 - c(e_i)\exp(-a(e_i)\tau)\big)},
\end{aligned}$$

for any parameters α and β, satisfying:

$$\begin{aligned}
\beta &< \frac{\kappa^2(e_i)}{2\chi^2(e_i)}\\
\alpha &< \frac{\kappa(e_i) + a(e_i)}{\chi^2(e_i)}.
\end{aligned}$$

Assume the following:

$$\frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(e_i)\big)^2}{\big(\nu(e_i)\big)^2} < \frac{\kappa^2(e_i)}{2\chi^2(e_i)}, \forall e_i \in \mathcal{E} \tag{4.50}$$

$$\max_{e_i\in\mathcal{E}}\Big\{\frac{\kappa(e_i)-a(e_i)}{\chi^2(e_i)}\Big\} \le \min_{e_i\in\mathcal{E}}\Big\{\frac{\kappa(e_i)+a(e_i)}{\chi^2(e_i)}\Big\}, \tag{4.51}$$

and define recursively backwards the following expressions:

$$A_k(\tau) := A^{\alpha_k,\beta_k,m_k}(\tau), B_k(\tau) := B^{\alpha_k,\beta_k,m_k}(\tau), \forall k = 0,\dots,K,$$

where

$$\begin{aligned}\beta_k &:= \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(t_k))\big)^2}{\big(\nu(m(t_k))\big)^2}, \forall k = 0,\dots,K\\ \alpha_k &:= B_{k+1}(\tau_{k+1}),\ \textit{for}\ \tau_k := t_{k+1}-t_k, \forall k = 0,\dots,K-1\\ \alpha_K &:= 0.\end{aligned}$$

Then the value function Φ *is given by the following equation:*

$$\begin{aligned}\Phi(t,v,x,e_i) =& \frac{v^\delta}{\delta}\mathbb{E}\big[f^{\mathcal{MC}}(t,x)\big|\mathcal{MC}(t)=e_i\big]\\ =& \frac{v^\delta}{\delta}\mathbb{E}\big[\xi^{\mathcal{MC}}(t)\exp\big\{B^{\mathcal{MC}}(t)x\big\}\big|\mathcal{MC}(t)=e_i\big],\end{aligned} \tag{4.52}$$

where for any $m \in \mathbb{M}$*, functions* ξ^m *and* B^m *are given by:*

$$\xi^m(t) = \exp\Big\{\sum_{k=0}^{K}\Big\{\int_t^T \delta r(m(s))\mathrm{d}s + A_k(t_{k+1}-t) + \sum_{z=k+1}^{K} A_z(\tau_z)\Big\}\mathbb{1}_{t\in[t_k,t_{k+1})}\Big\}$$

$$B^m(t) = \sum_{k=0}^{K}\big\{B_k(t_{k+1}-t)\big\}\mathbb{1}_{t\in[t_k,t_{k+1})}.$$

The optimal portfolio is:

$$\bar{\pi}(t) = \frac{1}{1-\delta}\frac{\lambda(\mathcal{MC}(t))}{\big(\nu(\mathcal{MC}(t))\big)^2}.$$

Proof

An application of Proposition 4.18 for $\rho = 0$, i.e. $\vartheta = 1$, shows that the value function in the corresponding time-dependent model is given by:

$$\begin{aligned}\Phi^m(t,v,x) &= \frac{v^\delta}{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T g(s,\tilde{X}^m(s),m(s))\mathrm{d}s\Big\}\Big|\tilde{X}^m(t)=x\Big]\\ &= \frac{v^\delta}{\delta}\xi^m(t)\exp\big\{B^m(t)x\big\},\end{aligned}$$

and for the optimal strategy it holds that:

$$\bar{\pi}^m(t) = \frac{1}{1-\delta}\frac{\lambda(m(t))}{\big(\nu(m(t))\big)^2}.$$

So, Corollary 4.12 can be applied and the statement follows directly.

□

Observe that function Φ can be easily computed by a partial Monte Carlo simulation, where one has to simulate only the path of the Markov chain and not all other processes.
If we assume separability of the value function in e_i and x, the solution can be even further simplified. As in this case an explicit solution can be derived also in the Heston model with leverage, we directly consider the case with general correlation ρ in the next section.

4.3.3 Markov-modulated Heston model with correlation (MMH$^\rho$)

For a tractable framework with leverage we specify our model in such a way that a separable explicit solution can be found. More precisely, we assume the following:

$$\begin{aligned}
\mathrm{d}P_0(t) &= P_0(t)r\big(\mathcal{MC}(t)\big)\mathrm{d}t \\
\mathrm{d}P_1(t) &= P_1(t)\Big[r\big(\mathcal{MC}(t)\big) + \underbrace{d\nu\big(\mathcal{MC}(t)\big)}_{=\lambda(\mathcal{MC}(t))} X(t)\mathrm{d}t + \nu\big(\mathcal{MC}(t)\big)\sqrt{X(t)}\mathrm{d}W^P(t)\Big] \\
\mathrm{d}X(t) &= \kappa\big\{\theta\big(\mathcal{MC}(t)\big) - X(t)\big\}\mathrm{d}t + \chi\sqrt{X(t)}\mathrm{d}W^X(t) \\
\mathrm{d}\langle W^P, W^X\rangle(t) &= \rho\ \mathrm{d}t.
\end{aligned} \tag{4.53}$$

This corresponds to the following specifications in the notation from (4.1):

$$\begin{aligned}
\Big(\frac{\mu_1(x,e_i)-r(e_i)}{\Sigma_1(x,e_i)}\Big)^2 &= d^2x && \Rightarrow h^{(0)}(e_i)=0,\quad \bar{h}^{(1)} = d^2 \\
\mu^X(x,e_i) &= \kappa\theta(e_i) - \kappa x && \Rightarrow \bar{k}^{(0)}(e_i) = \kappa\theta(e_i),\quad K^{(1)} = \kappa \\
\big(\Sigma^X(x,e_i)\big)^2 &= \chi^2 x && \Rightarrow H^{(0)}(e_i)=0,\quad H^{(11)} = \chi^2 \\
\rho\frac{\mu_1(x,e_i)-r(e_i)}{\Sigma_1(x,e_i)}\Sigma^X(x,e_i) &= \rho d\chi x && \Rightarrow \bar{g}^{(0)}(e_i)=0,\quad G^{(1)}(e_i) = \rho d\chi \\
& && \Rightarrow b = \frac{\chi}{|d|},
\end{aligned} \tag{4.54}$$

where b is the parameter defined on (4.7). Proposition 4.9 and Theorem 4.13 lead to the verification result in this case:

Proposition 4.22 (Solution and verification in MMH$^\rho$)
Assume that:

$$0 < \kappa - \frac{\delta}{1-\delta}\rho\chi|d| \tag{4.55}$$

$$\frac{\delta}{1-\delta}d^2 < \frac{\vartheta(\kappa - \frac{\delta}{1-\delta}\rho\chi|d|)^2}{\chi^2}. \tag{4.56}$$

For $a := \sqrt{(\kappa - \frac{\delta}{1-\delta}\rho\chi|d|)^2 - \frac{\delta}{1-\delta}\frac{\chi^2}{\vartheta}d^2}$, *define*

$$B(t) = \frac{\vartheta\big(- c(\kappa - \frac{\delta}{1-\delta}\rho\chi|d| + a)\exp\{-a(T-t)\} + \kappa - \frac{\delta}{1-\delta}\rho\chi|d| - a\big)}{\chi^2(1 - c\exp\{-a(T-t)\})}, \tag{4.57}$$

for all $t \in [0,T]$, *with* $c := \frac{\kappa - \frac{\delta}{1-\delta}\rho\chi|d| - a}{\kappa - \frac{\delta}{1-\delta}\rho\chi|d| + a}$. *Furthermore, set for all* $(t, e_i) \in [0,T] \times \mathcal{E}$:

$$\xi(t,e_i) = \mathbb{E}\Big[\exp\Big\{\int_t^T w(s, \mathcal{MC}(s))\Big\}\mathrm{d}s\Big|\mathcal{MC}(t) = e_i\Big], \tag{4.58}$$

where $w(t,e_i) = \delta r(e_i) + B(t)\kappa\theta(e_i)$. *The value function in Model (4.53) is given for all* $(t,v,x,e_i) \in [0,T] \times [0,\infty) \times D^X \times \mathcal{E}$ *by:*

$$\Phi(t,v,x,e_i) = \frac{v^\delta}{\delta}\xi(t,e_i)\exp\{B(t)x\}. \tag{4.59}$$

The optimal portfolio is:

$$\bar{\pi}(t) = \frac{1}{1-\delta}\Big[\frac{d}{\nu\big(\mathcal{MC}(t)\big)} + \rho\frac{\chi}{\nu\big(\mathcal{MC}(t)\big)}B(t)\Big]. \tag{4.60}$$

As mentioned earlier the first part of the optimal portfolio is called the mean-variance term:

$$\bar{\pi}^{MV}(t) := \frac{1}{1-\delta}\frac{d}{\nu(\mathcal{MC}(t))}, \tag{4.61}$$

and the second one is the hedging term:

$$\bar{\pi}^{H}(t) := \frac{1}{1-\delta}\rho\frac{\chi}{\nu(\mathcal{MC}(t))}B(t). \tag{4.62}$$

Proof
We would like to apply Theorem 4.13. To this aim, observe that Equation (4.36) takes the following form in this case:

$$B_t(t) + \frac{1}{2}\frac{\chi^2}{\vartheta}B^2(t) - \big[\kappa - \frac{\delta}{1-\delta}\rho\chi|d|\big]B(t) + \frac{1}{2}\frac{\delta}{1-\delta}d^2 = 0, B(T) = 0.$$

Conditions (4.55) and (4.56) allow us to conclude by Theorem 4.13 that the solution of the ODE from above is given by function B as defined in Equation (4.57). Furthermore, from Theorem 4.13 it follows that the HJB solution in the time-dependent model is given by:

$$\begin{aligned}\Phi^m(t,v,x) =& \frac{v^\delta}{\delta}\exp\Big\{\int_t^T w(s,m(s))\mathrm{d}s\Big\}\exp\{B(t)x\}\\ =& \frac{v^\delta}{\delta}\exp\Big\{\int_t^T \delta r(m(s)) + B(t)\kappa\theta(m(s))\mathrm{d}s\Big\}\exp\{B(t)x\}.\end{aligned}$$

Finally, observe that by Proposition 4.9, Φ^m is indeed the corresponding value function. The statement follows by applying the verification result from Theorem 4.13. □

Now let us have a closer look at the variance process of the log returns of the risky asset: $Var(t) := \nu^2(\mathcal{MC}(t))X(t)$. We start by deriving explicitly its dynamics.

Corollary 4.23 (Variance of the log returns)
The instantaneous variance of the log asset returns is characterized by the following SDE:

$$\begin{aligned}\mathrm{d}Var(t) =& \hat{\kappa}\big(\mathcal{MC}(t)\big)\big(\hat{\theta}\big(\mathcal{MC}(t)\big) - Var(t)\big)\mathrm{d}t + \hat{\chi}\big(\mathcal{MC}(t)\big)\sqrt{Var(t)}\mathrm{d}W^X(t)\\ &+ Var(t)\sum_{i=1}^I \frac{\nu^2(e_i)}{\nu^2\big(\mathcal{MC}(t)\big)}\mathrm{d}M_i(t),\end{aligned}$$

where:

$$\begin{aligned}\hat{\kappa}\big(\mathcal{MC}(t)\big) &= \kappa - \sum_{i=1}^I \frac{\nu^2(e_i)}{\nu^2\big(\mathcal{MC}(t)\big)} q_{MC(t),i}\\ \hat{\theta}\big(\mathcal{MC}(t)\big) &= \frac{\nu^2\big(\mathcal{MC}(t)\big)\kappa\theta\big(\mathcal{MC}(t)\big)}{\hat{\kappa}}\\ \hat{\chi}\big(\mathcal{MC}(t)\big) &= |\nu\big(\mathcal{MC}(t)\big)|\chi.\end{aligned}$$

With this notation the price process for the risky asset is given by the following SDE:

$$\mathrm{d}P_1(t) = P_1(t)\Big[r\big(\mathcal{MC}(t)\big) + \frac{d}{\nu\big(\mathcal{MC}(t)\big)}Var(t)\Big]\mathrm{d}t + \sqrt{Var(t)}\mathrm{d}W^P(t). \tag{4.63}$$

Proof
Follows directly as an application of Itô's formula for Markov-modulated Itô diffu-

sions, which is stated in Theorem 2.72:

$$\begin{aligned}
\mathrm{d}Var(t) =&\nu^2\big(\mathcal{MC}(t)\big)\kappa\big\{\theta\big(\mathcal{MC}(t)\big) - X(t)\big\}\mathrm{d}t + \nu^2\big(\mathcal{MC}(t)\big)\chi\sqrt{X(t)}\mathrm{d}W^X(t) \\
&+ \sum_{i=1}^{I}\nu^2(e_i)X(t)q_{MC(t),i}\mathrm{d}t + \sum_{i=1}^{I}\nu^2(e_i)X(t)\mathrm{d}M_i(t) \\
=&\nu^2\big(\mathcal{MC}(t)\big)\kappa\theta\big(\mathcal{MC}(t)\big) - Var(t)\underbrace{\left\{\kappa - \sum_{i=1}^{I}\frac{\nu^2(e_i)}{\nu^2\big(\mathcal{MC}(t)\big)}q_{MC(t),i}\right\}}_{=:\hat{\kappa}(\mathcal{MC}(t))}\mathrm{d}t \\
&+ |\nu\big(\mathcal{MC}(t)\big)|\chi\sqrt{Var(t)}\mathrm{d}W^X(t) + Var(t)\sum_{i=1}^{I}\frac{\nu^2(e_i)}{\nu^2\big(\mathcal{MC}(t)\big)}\mathrm{d}M_i(t).
\end{aligned}$$

□

Observe that the variance Var follows a mean-reverting process with jumps according to the Markov chain, where all parameters depend on the Markov chain. This rich stochastic structure makes the considered model very flexible and suitable for describing a wide range of markets. Furthermore, from SDE (4.63) the structure of the mean-variance portfolio becomes intuitively clear: $\frac{d}{\nu(\mathcal{MC}(t))}$ corresponds to the excess return for a unit of variance. A detailed analysis of the single components driving the optimal portfolio is given in the next section.

4.3.4 Numerical implementation and discussion

In this section we illustrate and interpret the results derived above by some numerical examples. First of all we specify the basic parameter set we are working with and show the numerical results for this parameter specification in order to indicate the impact of the Markov switching. We continue with the hedging term and the impact of the stochastic volatility. Thereafter, we study the influence of the risk aversion parameter δ on the behavior of the investor. Then, we extend the analysis on the impact of d and ν, which are the driving parameters of $\bar{\pi}^{MV}$ as one can see in Equation (4.61). Subsequently we discuss the sensitivity of the results to changes of the remaining parameters. Finally we present an alternative parameter specification which can capture an even more complex market behavior, and interpret the obtained results.

In what follows we consider the model presented in Section 4.3.3 and assume that the Markov chain can switch between two states. The first one, e_1, describes a calm market with moderate volatility levels. The second one, e_2 corresponds to a turbulent state with higher volatility and lower ratio $\frac{\mu_1 - r}{Var(t)}$. The investment time horizon is set to $T = 5$ years throughout the whole section. Based on the empirical results from [1][1] we fix the following basic parameter set: $\kappa = 4$, $\theta(e_1) := \theta_1 = 0.02$,

[1]In this paper the parameters of a Heston model are estimated using daily observations of the S&P500 stock index and VIX volatility index over the period from 1990 to 2003. Using their

$\theta(e_2) := \theta_2 = 0.04$, $\chi = 0.35$, $d = 1.7$, $\nu(e_1) := \nu_1 = 1$, $\nu(e_2) := \nu_2 = 1.3$, $r(e_1) := r_1 = 0.03$, $r(e_2) := r_2 = 0.01$, $\rho = -0.8$. So, as mentioned in Remark 2.45 the average half-life of process X is $\frac{\ln(2)}{\kappa} = \frac{\ln(2)}{4} \approx 2$ months. Further, following [11][2], we set the elements of the intensity matrix to $q_{1,1} = -1.0909$, $q_{2,2} = -3.4413$. This means that on average the Markov chain remains one year in the calm state and around 4 months in the turbulent state, as the waiting time the Markov chain spends in state e_i before the next jump is exponentially distributed with parameter $-q_{i,i}$ and expectation $-\frac{1}{q_{i,i}}$ (see Proposition 2.63). We will compare the results for two investors with different risk preferences: the first one has a positive risk aversion parameter $\delta = 0.3$ and for the second one it is negative $\delta = -1$. We denote these parameter specifications by Set 1 and Set 2, respectively. This differentiation is necessary because δ influences strongly the optimal behavior of the investor, as we will see in what follows.
Table 1 contains the optimal expected utility for Set 1 and Set 2 calculated as in Proposition 4.22, where function ξ is computed using a Monte Carlo simulation of the Markov chain with $10,000$ simulations. Additionally, it is compared with the optimal expected utility computed through a full Monte Carlo simulation of the three-dimensional process $(P_1, X, \mathcal{MC})$ with 1 Mio. simulations and 250 steps per year, in other words, trading occurs once a day. The values from the second and fourth columns are quite close to each other. This indicates that trading only once a day, which is sensible in reality, does not lead to a significant loss of utility. Therefore, the derived optimal strategy is applicable in practice. Moreover, many time-consuming Monte Carlo simulations are required to obtain converging results for the full Monte Carlo approach, which confirms the importance of the derived theoretical results in Proposition 4.22. Now we compare the optimal trading strate-

Parameter	Formula	Comp. time	Monte Carlo	Comp. time
Set 1	7.4261	40 sec	7.4260	approx. 2.2 h
Set 2	-0.0802	40 sec	-0.0802	approx. 2.2 h

Table 4.1: Comparison of the optimal expected utility computed as in Proposition 4.22 (second column) and by a full Monte Carlo simulation(fourth column), where $V(0) = 10$, $X(0) = 0.02$, $\mathcal{MC}(0) = e_1$, $T = 5$. The third and fifth columns contain the corresponding computational time.

gies in the two discussed parameter settings, which are presented in Figure 4.1. The main part of the optimal strategy is given by the mean-variance portfolio, which is observed to be positive, as expected, given that the expected asset return exceeds the riskless interest rate for both states of the Markov chain (at the end of this section we consider also model specifications where the excess return is negative for one of the states). One can also recognize that higher δ leads to a higher long position in the risky asset and for $\delta = 0.3$ the exposure even exceeds the investor's wealth. A lower δ results in a more moderate investment in the risky asset. This

results we selected our Markov-modulated parameters in such a manner, that the first state of $\mathcal{MC}$ describes a calm market and the second state describes a volatile one.

[2] In this paper a Markov-modulated Black-Scholes model is estimated for weekly S&P500 prices from 1987 to 2009. We use their result for the intensity matrix of the Markov chain.

observation is in accordance with the interpretation of δ as a risk aversion parameter. Furthermore, the investment in the risky asset is lower in the turbulent state of the Markov chain because it is associated with lower excess return and higher volatility and, thus, higher risk. Because of this difference in the optimal behavior for different states of the Markov chain, it is important for the investor to recognize the true state and to react to the parameter changes.
Let us continue with the hedging term. It follows from Statements i) and ii)

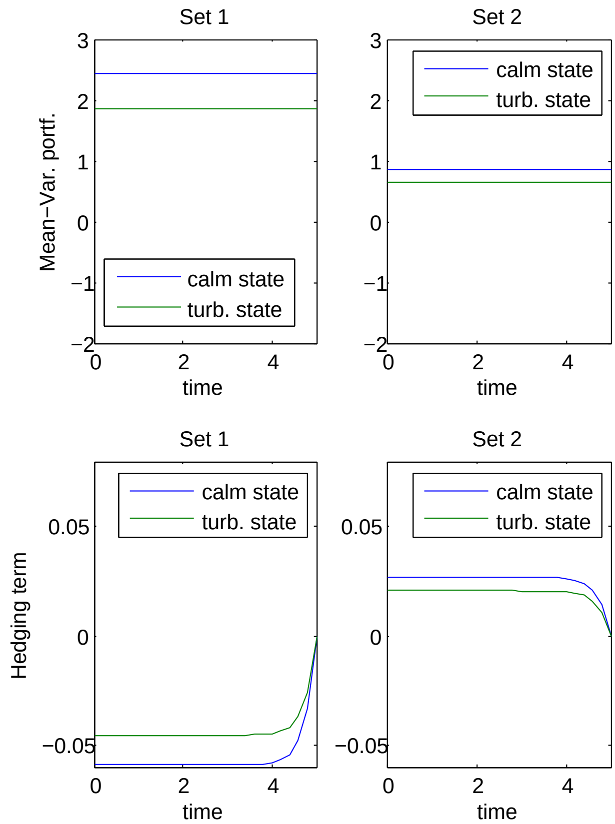


Figure 4.1: Mean-variance portfolio (4.61) (upper row) and hedging term (4.62) (lower row) for the two considered parameter sets. The blue lines represent the optimal investment in the calm state and the green lines the optimal investment in the turbulent state.

in Lemma 2.43 that $sign(B) = sign(\delta)$, so that $sign(\bar{\pi}^H) = sign(B)sign(\rho) = sign(\delta)sign(\rho)$. As we are dealing with a negative correlation $\rho < 0$ the correction term is negative for $\delta > 0$ and positive for $\delta < 0$. Although this might seem surprising at first sight it gets more clear once we clarify the impact of the stochastic factor X and the correlation ρ. By inserting the mean-variance portfolio $\bar{\pi}^{MV}$ in the SDE for the wealth process (4.6) we observe that the market price of risk, which is proportional to $\sqrt{X}$, drives the drift and volatility terms of the wealth process. Therefore, the relevant risk for the investor coming from the randomness of X is the change in the market price of risk. Analogously to the derivation of the value function, it can be shown that the expected utility of the terminal wealth of the mean-variance portfolio is increasing in X. So, the very risk averse investor ($\delta < 0$) would like to hedge his mean-variance portfolio against a falling market price of risk.

This is achieved by an additional long position in the risky asset ($\bar{\pi}^H > 0$) as the negative correlation relates lower values for X, i.e. lower levels of the market price of risk, to a higher asset price, i.e. higher wealth for a positive investment position. On the contrary, the investor with positive δ speculates on an increasing market price of risk and his strategy aims at high wealth in this case, realized by a short position in the risky asset ($\bar{\pi}^H < 0$). The reason for the differentiation between positive and negative values for δ is the form of the utility function: if $\delta < 0$ low wealth levels are heavily penalized, whereas for $\delta > 0$ the emphasis is on the reward for high wealth values. Moreover, observe that the higher the correlation between W^X and W^P, the higher the absolute value of the hedging term as the stochastic factor can be better hedged by the risky asset. Similar interpretations can be found in [67][3] and [24][4].

Now, we deepen our observations of the influence of the risk aversion parameter δ on the optimal portfolio. Figure 4.2 illustrates the density of the terminal wealth of an investor following the derived optimal strategy for different values of δ.

Higher values for δ are clearly recognized as leading to higher probabilities for both very low (close to zero) and very high (even above 120) wealth levels. In contrast, for small δ the probabilities for both high losses and high profits are much lower. This fact is also reflected in the shift of the 5%-quantile to the right and of the 95%-quantile to the left for smaller values for δ. This observation is in accordance with [100] who has proved that extremely risk averse investors replicate with their optimal investment a bond with maturity T.

Figure 4.3 confirms that a smaller δ results in a more conservative mean-variance portfolio. The absolute value of the hedging term also decreases for smaller δ primarily due to the influence of the factor $\frac{1}{1-\delta}$ that drives the mean-variance term as well. What is more, Figure 4.3 illustrates that depending on the risk aversion of the investor, the hedging term can play a significant role in the optimal asset allocation. A similar observation has been made also in [24]. Furthermore, it is clear from Figure 4.3 that in the turbulent state the investor holds less of the risky asset throughout time and for all values for δ.

We proceed with the remaining components of the mean-variance portfolio: d and ν. The influence of d on $\bar{\pi}^{MV}$ is naturally positive if ν is positive because a high d indicates a high market price of risk. The plots in Figure 4.4 illustrate the influence of d on the distribution of the terminal wealth of an investor with $\delta = 0.3$. The higher d, the higher the 95% quantile on the one side and the lower the 5% quantile on the other side. Therefore, the probabilities for both significant gains and significant losses increase. The risk on the downside results from the fact that the investor has a greater exposure for larger values of d, therefore, if the stock price declines, she experiences larger losses. However, the effect on the positive side is stronger because an increasing d leads to high excess returns for the stock and, thus, to the possibility for much greater gains. As the plots in Figure 4.5 illustrate, the same phenomenon

[3] In this paper the authors consider portfolio optimization in a model where a correlated mean-reverting stochastic factor influences the market price of risk.

[4] This paper derives optimal consumption and investment rules in discrete time on an infinite time horizon in a market where the excess return of the traded asset is modeled by an AR(1) process.

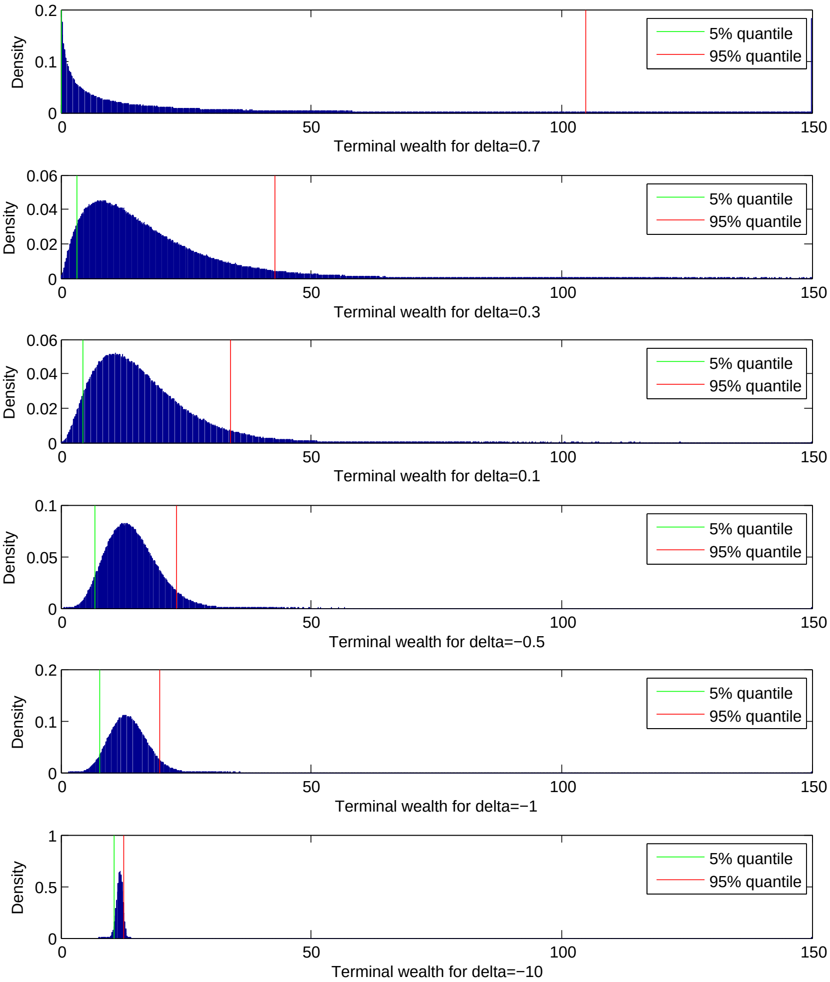


Figure 4.2: Density of the terminal wealth of an investor following the derived optimal strategy for different values of δ, where the remaining parameters are adopted from Set 1. The densities are obtained via Monte Carlo simulations of Model (4.53). For reasons of better comparability all values higher than 150 are summarized in the last bar in the plots.

is also observed for $\delta = -1$; however, the influence of d on the wealth distribution is not as strong because we are dealing with a more risk averse investor who prefers less exposure to the risky asset. Furthermore, the wealth distribution in this case is more symmetric than for $\delta = 0.3$, which reflects once again the risk aversion of the investor. The influence of ν on the investment strategy can be easily derived from the analytical formula for the optimal portfolio. Higher values of ν lead to lower investments in the risky assets because it reduces the excess return for a unit of variance. As $\nu_1 < \nu_2$, exposure in the risky asset is lower in the turbulent state than in the calm state. This observation holds for the mean-variance portfolio and for the hedging term. The larger the difference between ν_1 and ν_2, the more important for the investor is to recognize the Markov-switching character of the market and to adjust her strategy. One can understand the influence of ν on the optimal portfolio even better by recalling the SDE of the wealth process resulting from the optimal

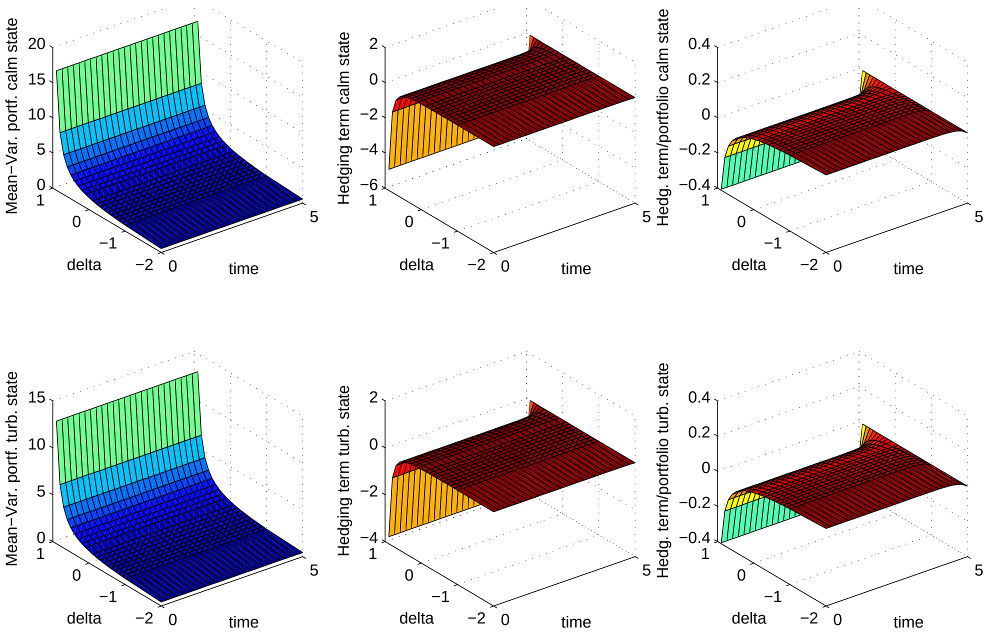


Figure 4.3: Optimal mean-variance portfolio (see Equation (4.61)) and hedging term (see Equation (4.62)) as well as the proportion of the hedging term to the whole portfolio over time for different values of δ.

strategy given in Equation (4.60):

$$\begin{aligned}\mathrm{d}V^{\bar{\pi}}(t) =& V^{\bar{\pi}}(t)\Big\{r\big(\mathcal{MC}(t)\big) + \frac{d}{1-\delta}\big(d+\rho\chi D(t)\big)X(t)\mathrm{d}t \\ &+ \frac{1}{1-\delta}\big(d+\rho\chi D(t)\big)\sqrt{X(t)}\mathrm{d}W^P(t)\Big\}.\end{aligned}$$

Note that the wealth process does not depend on ν. So, the optimal portfolio is chosen in such a way, that the investor is protected against changes in ν.
We continue with the remaining model parameters. Lower values for χ and higher values for κ may be summarized as leading to a lower absolute value of the hedging term because a lower volatility coefficient reduces the risk from the stochastic factor X and a faster mean-reversion makes hedging more difficult. The corresponding plots can be found in Appendix B.1. Finally, we remark that the mean-reversion level θ and the current level of X do not influence directly the optimal policy. The reason is that the stochastic factor X influences proportionally the instantaneous variance and the excess return of the risky asset price.
As announced in the beginning of the section, we now present an additional parameter setting. It describes a market, where the asset return in the turbulent state is on average lower than the riskless interest rate, so we set $\nu_1 = -1$, $\nu_2 = 1.3$, $d = -1.7$. State e_1 is again interpreted as a calm market and the second state e_2 describes a crisis. Note that this parameter specification makes it possible to capture three different patterns in the price process: If process X is at a low level and the Markov chain is in the first state we observe a calm market with positive expected return. If process X reaches high levels because of big positive increments of W^X but the

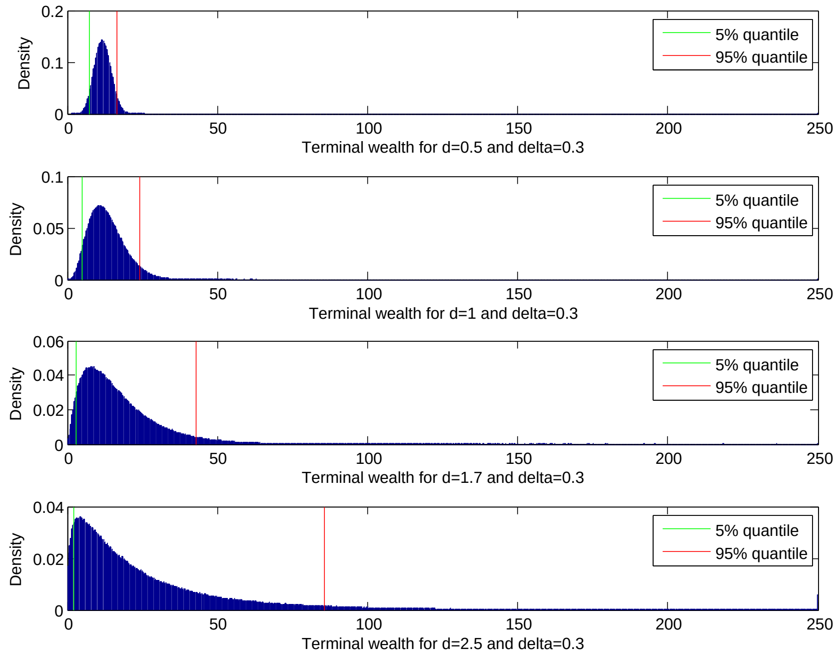


Figure 4.4: Density of the terminal wealth of an investor following the derived optimal strategy for different values of d, where the remaining parameters are adopted from Set 1. The densities are obtained via Monte Carlo simulations of Model (4.53). For reasons of better comparability all values higher than 250 are summarized in the last bar in the plots.

Markov chain remains in the first state, then the ratio $\frac{d}{\nu}$ would remain the same, but the negative coefficient ν_1 together with the negative correlation ρ would lead to positive increments in the diffusion term of the price process. So, in this case we are dealing with a continuously rising market. The second state of the Markov chain can be interpreted as a crisis situation as the volatility takes higher values and the expected return is lower than the riskless interest rate. We call this parameter setting Set 3 if $\delta = 0.3$ and Set 4 if $\delta = -1$. Table 4.2 summarizes these parameter specifications in comparison to Sets 1 and 2. The value function for the 4 parameter

	Set 1	Set 2	Set 3	Set 4
Parameter	$\kappa = 4$, $\theta_1 = 0.02$, $\theta_2 = 0.04$, $\chi = 0.35$, $r_1 = 0.03$, $r_2 = 0.01$, $\rho = -0.8$			
ν_1	1	1	-1	-1
ν_2	1.3	1.3	1.3	1.3
d	1.7	1.7	-1.7	-1.7
δ	0.3	-1	0.3	-1

Table 4.2: Summary of the parameter specifications for the four considered sets.

sets is presented in Table 4.3. Note that changing the signs of ν_1 and d does not influence heavily the optimal expected utility. In contrast, as we can see in Figure 4.6, the optimal portfolio strategies are very different. They are adjusted to the considered market in such a way that the investor takes the most advantage out of it and thus, her expected utility does not change a lot. For Sets 3 and 4 it is optimal

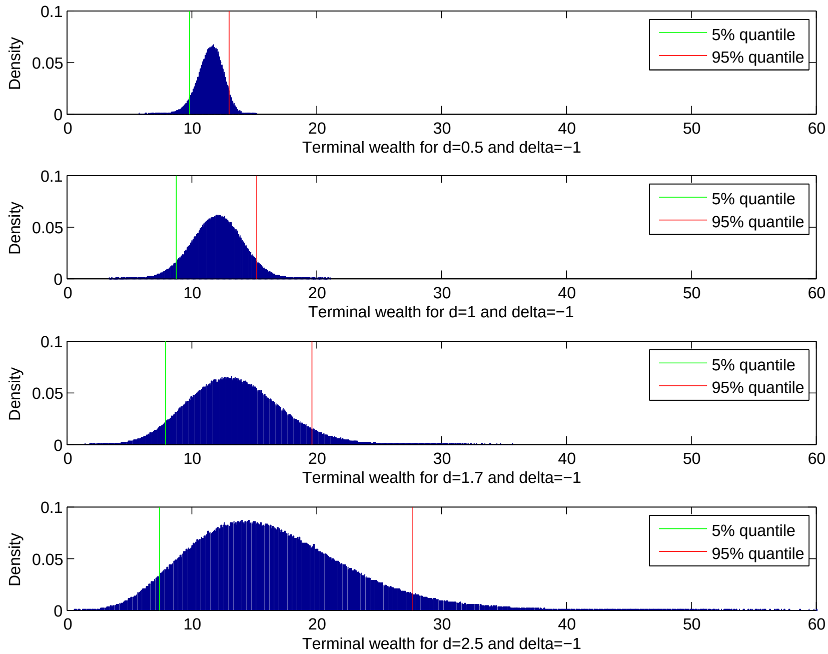


Figure 4.5: Density of the terminal wealth of an investor following the derived optimal strategy for different values of d, where the remaining parameters are adopted from Set 2. The densities are obtained via Monte Carlo simulations of Model (4.53). For reasons of better comparability all values higher than 60 are summarized in the last bar in the plots.

Parameter	Formula	Comp. time	Monte Carlo	Comp. time
Set 1	7.4261	40 sec	7.4260	approx. 2.2 h
Set 2	-0.0802	40 sec	-0.0802	approx. 2.2 h
Set 3	7.4810	40 sec	7.4780	approx. 2.2 h
Set 4	-0.0810	40 sec	-0.0810	approx. 2.2 h

Table 4.3: Comparison of the optimal expected utility computed as in Proposition 4.22 (second column) and by a full Monte Carlo simulation (fourth column), where $V(0) = 10$, $X(0) = 0.02$, $\mathcal{MC}(0) = e_1$, $T = 5$. The third and fifth columns contain the corresponding computational time.

to short sell the risky asset in the turbulent state, as e_2 is clearly interpreted as a crash. So, recognizing the Markov-switching character of the market protects the investor in the case of a crisis.

For Sets 3 and 4 also the sign of the hedging term in the first state changes. The reason is the negative sign of $\nu(e_1)$, as in this case an increase of X due to a positive increment $\mathrm{d}W^X$ is related to an increase in $\nu \mathrm{d}W^P$ as well. So, the speculation of the riskier investor (Set 3) on rising X is realized in state e_1 by a positive investment in the stock, whereas the negative hedging term in Set 4 acts as a protection against negative increments of W^X which might lead to falling stock prices, as $\nu(e_1) < 0$. Note that in spite of the positive correlation between $\mathrm{d}W^X$ and $\nu \mathrm{d}W^P$, the empirically observed tendency that falling stock prices and high volatility occur together is realized by the Markov switching in θ and the excess return of the stock.

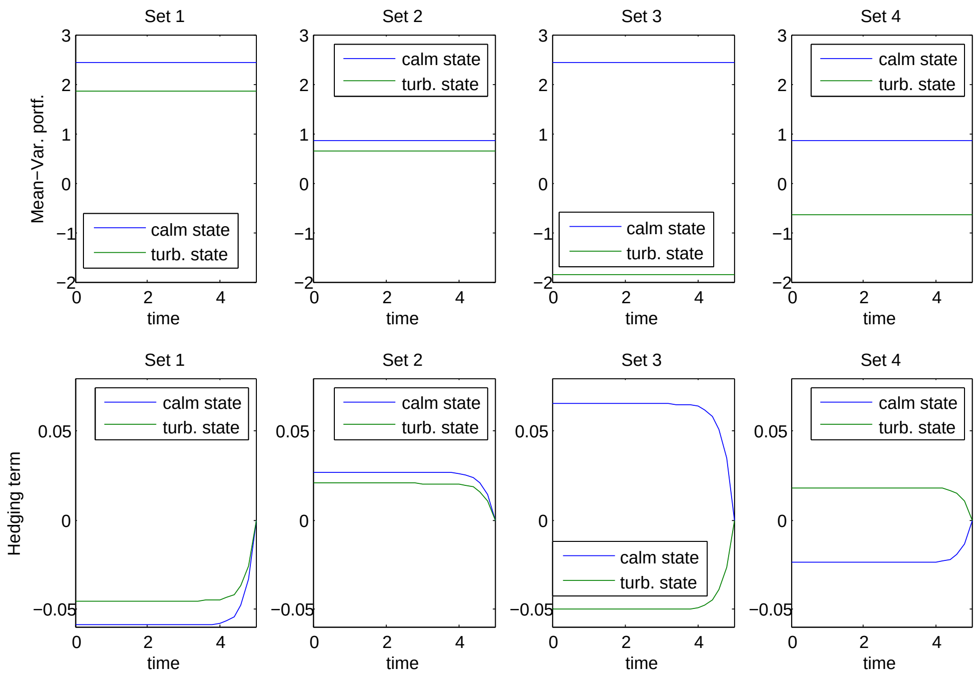


Figure 4.6: Mean-variance portfolio (4.61) (first row) and hedging term (4.62) (second row) for the four considered parameter sets. The blue lines represent the optimal investment in the calm state and the green lines - in the turbulent state.

Chapter 5

Multidim. affine Markov-modulated model

In this chapter we consider the general multidimensional Model (3.1) and Problem (3.12) with the power utility function U_P defined as in (2.25). We extend the literature on utility maximization in multidimensional models listed in the introduction in two directions: first, we add Markov switching to the affine framework and second, we provide easy to apply verification theorems. The example we present in Section 5.3.1 is a generalization of the results presented in [73] to Markov-modulated parameters and the example considered in Section 5.3.2 extends the model presented in [45] to Markov switching and solves the utility maximization problem for this model. Note that parts of this chapter have been published in [84].
As in the one-dimensional case, we propose an ansatz for the value function in accordance with the utility function:

$$\Phi(t, v, x, e_i) = \frac{v^\delta}{\delta} f(t, x, e_i), \tag{5.1}$$

for a function $f : [0, T] \times D^X \times \mathcal{E} \to \mathbb{R}$. Together with (3.18) this leads to the following simplified form for $\bar{\pi}$:

$$\bar{\pi}(t) = \frac{1}{1-\delta}\Big\{(\Sigma\Sigma')^{-1}(\mu - r) + (\Sigma')^{-1}\rho'(\Sigma^X)'\frac{f_x}{f}\Big\}\Big|_{(t,X(t),\mathcal{MC}(t))}. \tag{5.2}$$

Inserting (5.1) and (5.2) in the HJB Equation (3.16) yields the following PDE system for f:

$$\begin{aligned} &f_t + f\delta\{r + \frac{1}{2}\frac{1}{1-\delta}(\mu-r)'\big(\Sigma\Sigma'\big)^{-1}(\mu-r)\} + f_x'\{\mu^X + \frac{\delta}{1-\delta}\Sigma^X\rho\Sigma^{-1}(\mu-r))\} \\ &+ \frac{1}{2}Tr\big(\Sigma^X(\Sigma^X)'f_{xx'}\big) + \frac{1}{2}\frac{\delta}{1-\delta}\frac{1}{f}f_x'\Sigma^X\rho\rho'(\Sigma^X)'f_x\Big|_{(t,x,e_i)} = -\sum_{z=1}^{I} q_{i,z} f(t, x, e_z) \\ &f(T, x, e_i) = 1, \forall i \in \{1, \dots, I\}. \end{aligned} \tag{5.3}$$

So, we are interested in finding the solution to this system. Before we present the explicit solutions for function f, we prove a verification theorem, which assures that if we find a solution to the HJB equation, then it is indeed the value function for the considered optimization problem and $\bar{\pi}$ from (5.2) is the optimal portfolio. The verification result is based on martingale theory.

Theorem 5.1 (Verification via a martingale condition)
Consider a real-valued function $\Phi(t, v, x, e_i) : [0, T] \times [0, \infty) \times D^X \times \mathcal{E} \to \mathbb{R}$ and assume that:

i) For each $e_i \in \mathcal{E}$, $\Phi(\cdot, \cdot, \cdot, e_i) \in \mathscr{C}^{1,2,2}\big([0, T] \times [0, \infty) \times D^X\big)$, i.e. Φ is once continuously differentiable in t and twice continuously differentiable in v and x.

ii) Φ satisfies the following equation:

$$\mathcal{L}(e_i, \bar{\pi})\Phi(t, v, x, e_i) = -\sum_{z=1}^{I} q_{i,z}\Phi(t, v, x, e_z)$$
$$\Phi(T, v, x, e_i) = U_P(v), \forall i \in \{1, \ldots, I\},$$

where operator $\mathcal{L}$ is defined in (3.17) and $\bar{\pi}$ is given by (3.18).

iii) $\big\{\Phi\big(t, V^{\bar{\pi}}(t), X(t), \mathcal{MC}(t)\big)\big\}_{t\in[0,T]}$ is a martingale.

Then:

$$\mathbb{E}\big[U_P\big(V^{\bar{\pi}}(T)\big)\big|V^{\bar{\pi}}(t) = v, X(t) = x, \mathcal{MC}(t) = e_i\big] = \Phi(t, v, x, e_i).$$

Further, if $\Phi = \frac{v^\delta}{\delta} f(t, x, e_i)$ for a positive function f then $\bar{\pi}$ is the optimal solution and Φ is the corresponding value function.

The proof is given in Appendix C.

Similarly to Remark 4.5 one can show the following result:

Remark 5.2
If Φ is positive then just the following two conditions:

i) For each $e_i \in \mathcal{E}$, $\Phi(\cdot, \cdot, \cdot, e_i) \in \mathscr{C}^{1,2,2}\big([0, T] \times [0, \infty) \times D^X\big)$,

ii) Φ satisfies System (3.16), where the maximum on the left hand side is obtained at $\bar{\pi}$ as given by (3.18),

suffice to show that:

$$\mathbb{E}\big[U_P\big(V^{\pi}(T)\big)\big|V^{\pi}(t) = v, X(t) = x, \mathcal{MC}(t) = e_i\big] \le \Phi(t, v, x, e_i),$$

for all $(t, v, x, e_i) \in [0, T] \times [0, \infty) \times D^X \times \mathcal{E}$ and all portfolio strategies π. The derivation can be found in Appendix C.

5.1 Time-dependent model

Alternatively to applying Theorem 5.1, one can first derive the solution in the corresponding time-dependent model and then apply Theorem 3.5, analogously to the one-dimensional case. Thus, in this section we briefly state the basic results for the multidimensional time-dependent Model (3.14).
Consider Problem (3.15) with the power utility function $U_P(v) = \frac{v^\delta}{\delta}$. As in the model with Markov-switching, an ansatz of the form:

$$\Phi^m(t, v, x) = \frac{v^\delta}{\delta} f^m(t, x), \tag{5.4}$$

and the first-order condition for a maximum in the HJB equation (3.21) lead to the following simplified PDE defined piece-wise for all $k = 0, \ldots, K$ by:

$$\begin{aligned} & f_t^m + f^m \delta \big\{ r + \frac{1}{2} \frac{1}{1-\delta} (\mu - r)' \big(\Sigma\Sigma'\big)^{-1} (\mu - r) \big\} \\ & + (f_x^m)' \big\{ \mu^X + \frac{\delta}{1-\delta} \Sigma^X \rho \Sigma^{-1} (\mu - r) \big) \big\} + \frac{1}{2} Tr \big(\Sigma^X (\Sigma^X)' f_{xx'}^m \big) \\ & + \frac{1}{2 f^m} (f_x^m)' \frac{\delta}{1-\delta} \Sigma^X \rho \rho' (\Sigma^X)' f_x^m \Big|_{(t,x,m(t))} = 0, \forall (t, v) \in [t_k, t_{k+1}) \times D^X \\ & f^m(T, x) = 1. \end{aligned} \tag{5.5}$$

So, if we find a solution f^m to this equation, then function $\Phi^m = \frac{v^\delta}{\delta} f^m(t, x)$ solves the HJB PDE (3.20). Before, we go deeper in the derivation of f^m, we prove that if such a solution exists, then $\Phi^m = \frac{v^\delta}{\delta} f^m(t, x)$ is indeed the value function.

Proposition 5.3 (Verification result in the time-dependent model)
Consider a real-valued function $\Phi^m(t, v, x) : [0, T] \times [0, \infty) \times D^X \to \mathbb{R}$ where $\Phi^m(t, v, x, e_i) = \frac{v^\delta}{\delta} f^m(t, x)$ for a positive function f^m. Assume that:

i) $\Phi^m \in \mathscr{C}^{1,2,2}\big([t_k, t_{k+1}) \times [0, \infty) \times D^X\big)$ *for all* $k = 0, \ldots, K$,

ii) $\Phi^m \in \mathscr{C}\big([0, T] \times [0, \infty) \times D^X\big)$,

iii) Φ^m *satisfies the following PDE, defined piece-wise for all* $k = 1, \ldots, K$*:*

$$\begin{aligned} \mathcal{L}(m(t), \bar{\pi}^m) \Phi^m(t, v, x) &= 0, \forall (t, v, x) \in [t_k, t_{k+1}) \times [0, \infty) \times D^X \\ \Phi^m(T, v, x) &= \frac{v^\delta}{\delta}, \end{aligned} \tag{5.6}$$

where $\mathcal{L}$ is given by (3.17) and $\bar{\pi}^m$ is as in (3.21).

iv) $\big\{ \Phi^m \big(t, V^{m, \bar{\pi}^m}(t), X^m(t)\big) \big\}_{t \in [0,T]}$ *is a martingale.*

Then $\bar{\pi}^m$ is the optimal portfolio for Model (3.14) and Φ^m is the corresponding value function.

The proof is given in Appendix C.

Now we state explicitly the value function and the optimal trading strategy.

Proposition 5.4 (Solution in the time-dependent model)
Consider Model (3.14) and Problem (3.15) with $U = U_P$. Assume that B^m is a continuous solution to the following system, which holds piece-wise on $[t_k, t_{k+1})$ for all $k = 0, \dots, K$:

$$\begin{aligned}
&B_t^m(t)'x + \delta\big\{\bar{\varepsilon}^{(1)}\big(m(t)\big)\big\}'x + \frac{1}{2}\frac{\delta}{1-\delta}\big\{\bar{h}^{(1)}\big(m(t)\big)\big\}'x - B^m(t)'K^{(1)}\big(m(t)\big)x \\
&+ \frac{\delta}{1-\delta}B^m(t)'G^{(1)}\big(m(t)\big)x + \frac{1}{2}\frac{\delta}{1-\delta}B^m(t)'\Big[\sum_{j=1}^{J}\big\{L^{(1j)}\big(m(t)\big) \\
&+ H^{(1j)}\big(m(t)\big)\big\}x_j\Big]B^m(t) + \frac{1}{2}B^m(t)'\Big[\sum_{j=1}^{J}H^{(1j)}\big(m(t)\big)x_j\Big]B^m(t) = 0 \\
&B^m(T) = 0,
\end{aligned} \tag{5.7}$$

such that $B^m \in \mathscr{C}^1\big([t_k, t_{k+1})\big)$ for all $k = 0, \dots, K$. Furthermore, let function ξ^m be given by:

$$\xi^m(t) = \exp\Big\{\int_t^T w\big(s, m(s)\big)\mathrm{d}s\Big\}, \tag{5.8}$$

where

$$\begin{aligned}
w(t, e_i) =& \delta\varepsilon^{(0)}(e_i) + \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}(e_i) + \big(B^m(t)\big)'\bar{k}^{(0)}(e_i) + \frac{\delta}{1-\delta}\big(B^m(t)\big)'\bar{g}^{(0)}(e_i) \\
&+ \frac{1}{2}\frac{\delta}{1-\delta}\big(B^m(t)\big)'\big(L^{(0)}(e_i) + H^{(0)}(e_i)\big)B^m(t) + \frac{1}{2}\big(B^m(t)\big)'H^{(0)}(e_i)B^m(t).
\end{aligned}$$

Then, Φ^m given by:

$$\Phi^m(t, v, x) = \frac{v^\delta}{\delta}\xi^m(t)\exp\{(B^m(t))'x\} \tag{5.9}$$

satisfies the HJB Equation (3.20) with:

$$\bar{\pi}^m = \frac{1}{1-\delta}\Big\{(\Sigma\Sigma')^{-1}(\mu - r) + (\Sigma')^{-1}\rho'(\Sigma^X)'B^m\Big\}\Big|_{(t,X(t),m(t))}. \tag{5.10}$$

Furthermore, if $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ is a martingale, then Φ^m is the value function for the considered problem and $\bar{\pi}^m$ is the optimal trading strategy.

Proof
First of all, we solve Equation (5.5). One can apply a similar transformation as in

Proposition 4.7 to eliminate the nonlinear term. However, as we already have the experience from the one-dimensional case, we directly state the following ansatz for f^m:

$$f^m(t,x) = \xi^m(t)\exp\{(B^m(t))'x\}, \tag{5.11}$$

for some functions $\xi^m : [0,\infty) \to \mathbb{R}$ and $B^m(t) = (B_1^m(t),\dots,B_J^m(t))' : [0,\infty) \to \mathbb{R}^J$. Substitution of (5.11) in (5.5) leads to System (5.8) for B^m and the following system of coupled ODEs for ξ^m:

$$\begin{aligned}&\xi_t^m(t) + \xi^m(t)\big[\delta\varepsilon^{(0)}\big(m(t)\big) + \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}\big(m(t)\big) + B^m(t)'\bar{k}^{(0)}\big(m(t)\big)\\ &+ \frac{\delta}{1-\delta}B^m(t)'\bar{g}^{(0)}\big(m(t)\big) + \frac{1}{2}\frac{\delta}{1-\delta}B^m(t)'\big\{L^{(0)}\big(m(t)\big) + H^{(0)}\big(m(t)\big)\big\}B^m(t)\\ &+ \frac{1}{2}B^m(t)'H^{(0)}\big(m(t)\big)B^m(t)\big] = 0, \xi^m(T) = 1.\end{aligned} \tag{5.12}$$

This equation leads directly to the solution for ξ^m given by (5.8). Note that the involved integral over the finite interval $[0,T]$ is well-defined and finite, as function B^m is continuous. So, Φ^m as given by (5.9) solves the HJB PDE in the considered case.
For the verification result we check the conditions of Proposition 5.3. Conditions i) and ii) are trivially fulfilled for function Φ^m. Condition iii) has just been shown and Condition iv) holds as well. Thus, the verification result follows by a direct application of Proposition 5.3.

□

Remark 5.5
The multidimensional Riccati ODE for B^m can be transformed by comparison of the factors in front of the single x_j's to the following system of coupled ODEs:

$$\begin{aligned}&\frac{\partial}{\partial t}B_j^m(t) + \delta\bar{\varepsilon}_j^{(1)}\big(m(t)\big) + \frac{1}{2}\frac{\delta}{1-\delta}\bar{h}_j^{(1)}\big(m(t)\big) - \sum_{d=1}^{J} B_d^m(t)K_{jd}^{(1)}\big(m(t)\big)\\ &+ \frac{\delta}{1-\delta}\sum_{d=1}^{J} B_d^m(t)G_{jd}^{(1)}\big(m(t)\big) + \frac{1}{2}\frac{\delta}{1-\delta}B^m(t)'\big(L^{(1j)}\big(m(t)\big)\\ &+ H^{(1j)}\big(m(t)\big)\big)B^m(t) + \frac{1}{2}B^m(t)'H^{(1j)}\big(m(t)\big)B^m(t) = 0\\ &B_j(T) = 0, \forall j \in \{1,\dots,J\}.\end{aligned} \tag{5.13}$$

In the special case where $K^{(1)}$ and $G^{(1)}$ are diagonal matrices, and $L^{(1j)}$ and $H^{(1j)}$ have a non-zero element only at position (j,j), we have J not coupled Riccati ODEs with piece-wise constant coefficients. So, a solution can be constructed starting at the back and applying in each step Corollary 2.44. Note that in general B^m depends on the whole path $\{m(t)\}_{t\in[0,T]}$.

As in the one-dimensional case, the martingale condition required in the previous proposition can be shown in some special cases by applying Theorem 2.34. This is more precisely stated in the following proposition:

Proposition 5.6 (Verification via Theorem 2.34)
Consider again Model (3.14) and Problem (3.15) with $U = U_P$. Let Φ^m and $\bar{\pi}^m$ be defined as in Proposition 5.4. Define process G by:

$$G(t) := \ln\left(\frac{\Phi^m(t, V^{m,\bar{\pi}^m}(t), X^m(t))}{\Phi^m(0, V^{m,\bar{\pi}^m}(0), X^m(0))}\right).$$

Now consider the $J+1$-dimensional process $Z := (\bar{X}^m, G)'$, where $\bar{X}^m$ is a suitable permutation of X^m. Then its semimartingale characteristics μ^Z, Γ^Z exhibit an affine structure as in Theorem 2.34. Assume that μ^Z, Γ^Z fulfill requirements i)-iv) from Theorem 2.34. Then $\left\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\right\}_{t\in[0,T]}$ is a martingale. Furthermore, Φ^m is the value function to the considered optimization problem and $\bar{\pi}^m$ is the optimal portfolio.

The proof is given in Appendix C.

Now we can continue with the solution for the Markov-modulated model.

5.2 Markov-modulated model

Consider again Model (3.1) and Problem (3.18) with the power utility function. In what follows we derive an explicit solution for the value function and the optimal investment strategy. We start with the case for which the Brownian motions driving the stochastic factors and the ones for the asset prices are independent. Thereafter we derive the solutions for the case with correlation.

5.2.1 Solution with no correlation

Consider Model (3.1) and set $\rho = 0$. Then the non-linear term in Equation (5.3) disappears and under some technical requirements we can apply the Feynman-Kac theorem for Markov-modulated processes (see Corollary 2.76), which yields the following probabilistic representation:

$$\begin{aligned} f(t,x,e_i) =& \mathbb{E}\Big[\exp\Big\{\int_t^T \delta\Big[r + \frac{1}{2}\frac{1}{1-\delta}(\mu - r)'\big(\Sigma\Sigma'\big)^{-1}(\mu - r)\Big]\mathrm{d}s\Big\} \\ &\Big| X(t) = x, \mathcal{MC}(t) = e_i\Big]. \end{aligned} \tag{5.14}$$

Alternatively to showing the conditions necessary for the application of the Feynman-Kac theorem in the Markov-modulated model, one can use Theorem 3.5

and reduce the problem to finding a solution in the corresponding time-dependent model with deterministic piece-wise constant coefficients. As shown in the next corollary, this approach is always possible in the case without leverage, as the optimal solution in the corresponding time-dependent model depends only on the current value of m and not on its entire path.

Corollary 5.7 (Solution with no correlation)
Assume that $\rho = 0$ in Model (3.1). Let B^m be a continuous solution to the following system, which holds piece-wise on each interval $[t_k, t_{k+1})$ for $k = 0, \dots, K$:

$$\begin{aligned} &B^m_t(t)'x + \delta\bar{\varepsilon}^{(1)}\big(m(t)\big)'x + \frac{1}{2}\frac{\delta}{1-\delta}\bar{h}^{(1)}\big(m(t)\big)'x - B^m(t)'K^{(1)}\big(m(t)\big)x \\ &+ \frac{1}{2}B^m(t)'\sum_{j=1}^{J} H^{(1j)}\big(m(t)\big)x_j B^m(t) = 0, B^m(T) = 0, \end{aligned} \tag{5.15}$$

with $B^m \in \mathscr{C}^1\big([t_k, t_{k+1})\big)$ for all $k = 0, \dots, K$. Define function ξ^m by:

$$\xi^m(t) = \exp\Big\{\int_t^T w\big(s, m(s)\big)\mathrm{d}s\Big\},$$

where

$$w(t, e_i) = \delta\varepsilon^{(0)}(e_i) + \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}(e_i) + B^m(t)'\bar{k}^{(0)}(e_i) + \frac{1}{2}B^m(t)'H^{(0)}(e_i)B^m(t). \tag{5.16}$$

Furthermore, define function Φ^m by:

$$\Phi^m(t, v, x) = \frac{v^\delta}{\delta}\exp\Big\{\int_t^T w\big(s, m(s)\big)\mathrm{d}s\Big\}\exp\big\{B^m(t)'x\big\}, \tag{5.17}$$

and assume that $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ is a martingale. Then function Φ given by:

$$\Phi(t, v, x, e_i) = \frac{v^\delta}{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T w\big(s, \mathcal{MC}(s)\big)\mathrm{d}s\Big\}\exp\big\{B^{\mathcal{MC}}(t)'x\big\}\Big|\mathcal{MC}(t) = e_i\Big] \tag{5.18}$$

is the value function for the considered problem and the optimal investment strategy is:

$$\bar{\pi}(t) = \frac{1}{1-\delta}\big(\Sigma\Sigma'\big)^{-1}(\mu - r)\Big|_{(t,X(t),\mathcal{MC}(t))}. \tag{5.19}$$

Proof
It follows from Proposition 5.4 and $\rho = 0$ that Φ^m is the solution for the corresponding time-dependent model and $\bar{\pi}^m$ given by:

$$\bar{\pi}^m(t) = \frac{1}{1-\delta}\big(\Sigma\Sigma'\big)^{-1}(\mu - r)\Big|_{(t,X^m(t),m(t))} \tag{5.20}$$

is the optimal investment strategy. As $\bar{\pi}^m(t)$ depends only on the current value $m(t)$ and not on the whole path $\{m(t)\}_{t\in[0,T]}$, we can apply Theorem 3.5 that yields the statement.

□

Remark 5.8
Equation (5.15) is equivalent to the following system of coupled Riccati ODEs:

$$\begin{aligned}&\frac{\partial}{\partial t}B_j^m(t)+\delta\bar{\varepsilon}_j^{(1)}\big(m(t)\big)+\frac{1}{2}\frac{\delta}{1-\delta}\bar{h}_j^{(1)}\big(m(t)\big)-\sum_{d=1}^{J}B_d^m(t)K_{jd}^{(1)}\big(m(t)\big)\\&+\frac{1}{2}B^m(t)'H^{(1j)}\big(m(t)\big)B^m(t)=0, B_j(T)=0, \forall j\in\{1,\dots,J\}.\end{aligned}\tag{5.21}$$

Note that if $K^{(1)}$ is a diagonal matrix, and $H^{(1j)}$ has a non-zero element only at position (j,j), then in (5.21) we have J not coupled Riccati ODEs with piece-wise constant coefficients. Although at first sight this condition appears quite restrictive, it concerns only the structure of process X and includes many very flexible examples, wherein it is even allowed for correlation between the stochastic factors. E.g. it is fulfilled if for all $j\in\{1,\dots,J\}$, x_j appears linearly in its own drift term μ_j^X but does not influence the drifts of the other stochastic processes, and Σ^X is constant or $(\Sigma_{j,k}^X)^2$ is an affine function of x_j, for some $k\in\{1,\dots,J\}$, and $\Sigma_{i,k}^X=0$ for all $i\neq j$. See Section 5.3 for some examples, such as a bond-stock market model and a principal component stochastic correlation model.

Remark 5.9
Note that the expectation in (5.18) can be easily calculated using a Monte Carlo simulation with a very short computation time, as it involves only the simulation of the Markov chain.

The obtained expression in (5.18) can be further simplified if we assume separability of the value function in the Markov chain and the stochastic factor. In this case we can even allow for leverage. The details are presented in the subsequent analysis.

5.2.2 Solution with correlation

We continue with the case wherein W^P and W^X are correlated. In this case a probabilistic representation as in (5.14) is not possible because of the non-linear term in the PDE for f. An explicit solution can be found by assuming that coefficients $\bar{\varepsilon}^{(1)}$, $K^{(1)}$, $\bar{h}^{(1)}$, $G^{(1)}$, $H^{(1j)}$ and $L^{(1j)}$ for $j=1,\dots,J$ do not depend on the Markov chain and the application of a separable ansatz. The result is stated in the following theorem:

Theorem 5.10 (Solution with correlation)
Consider Model (3.1) with a general matrix ρ and assume that $\bar{\varepsilon}^{(1)}$, $K^{(1)}$, $\bar{h}^{(1)}$, $G^{(1)}$, $H^{(1j)}$ and $L^{(1j)}$ for $j=1,\dots,J$ are constant. Assume that the following equation:

$$\begin{aligned}
&B_t(t)'x + \delta(\bar{\varepsilon}^{(1)})'x + \frac{1}{2}\frac{\delta}{1-\delta}(\bar{h}^{(1)})'x - B(t)'K^{(1)}x + \frac{\delta}{1-\delta}B(t)'G^{(1)}x \\
&+ \frac{1}{2}\frac{\delta}{1-\delta}B(t)'\sum_{j=1}^{J}\big(L^{(1j)} + H^{(1j)}\big)x_j B(t) + \frac{1}{2}B(t)'\sum_{j=1}^{J}H^{(1j)}x_j B(t) = 0 \qquad (5.22)\\
&B(T) = 0
\end{aligned}$$

possesses a continuously differentiable solution B. Define function $\xi : [0,T]\times\mathcal{E}\to\mathbb{R}$ by:

$$\xi(t,e_i) = \mathbb{E}\Big[\exp\Big\{\int_t^T w\big(s,\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\big|\mathcal{MC}(t)=e_i\Big], \tag{5.23}$$

where

$$\begin{aligned}
w(t,e_i) =&\delta\varepsilon^{(0)}(e_i) + \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}(e_i) + B(t)'\bar{k}^{(0)}(e_i) + \frac{\delta}{1-\delta}B(t)'\bar{g}^{(0)}(e_i)\\
&+ \frac{1}{2}\frac{\delta}{1-\delta}B(t)'\big(L^{(0)}(e_i) + H^{(0)}(e_i)\big)B(t) + \frac{1}{2}B(t)'H^{(0)}(e_i)B(t).
\end{aligned} \tag{5.24}$$

Then the HJB system of equations is solved by the following function:

$$\Phi(t,v,x,e_i) = \frac{v^\delta}{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T w\big(s,\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\big|\mathcal{MC}(t)=e_i\Big]\exp\big\{B(t)'x\big\}, \tag{5.25}$$

with

$$\bar{\pi}(t) = \frac{1}{1-\delta}\Big\{\big(\Sigma\Sigma'\big)^{-1}(\mu - r) + \big(\Sigma'\big)^{-1}\rho'(\Sigma^X)'B(t)\Big\}\Big|_{(t,X(t),\mathcal{MC}(t))}. \tag{5.26}$$

Let Φ^m and $\bar{\pi}^m$ be as given by:

$$\begin{aligned}
\Phi^m(t,v,x) =&\frac{v^\delta}{\delta}\exp\Big\{\int_t^T w\big(s,m(s)\big)\mathrm{d}s\Big\}\exp\big\{B(t)'x\big\}\\
\bar{\pi}^m(t) =&\bar{\pi}(t)\big|_{(t,X^m(t),m(t))},
\end{aligned}$$

for any $m \in \mathbb{M}$. If at least one of the following two processes $\{\Phi^m(t,V^{m,\bar{\pi}^m}(t),X^m(t))\}_{t\in[0,T]}$ and $\{\Phi(t,V^{\bar{\pi}}(t),X(t),\mathcal{MC}(t))\}_{t\in[0,T]}$ is a martingale, then Φ is the value function to the considered optimization problem and the optimal portfolio is given by $\bar{\pi}$.

Proof
Analogously to the time-dependent model we consider the following ansatz

$$f(t,x,e_i) = \xi(t,e_i)\exp\big\{B(t)'x\big\},$$

where $\xi(t,e_i):[0,T]\times\mathcal{E}\to\mathbb{R}$, $B(t)=(B_1(t),\dots,B_J(t))':[0,T]\to\mathbb{R}^J$. Inserting this ansatz in Equation (5.3) leads to a system of ODEs for ξ and B. More precisely, for ξ we have the following system:

$$
\begin{aligned}
&\xi_t(t,e_i)+\xi(t,e_i)\big[\delta\varepsilon^{(0)}(e_i)+\frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}(e_i)+B(t)'\bar{k}^{(0)}(e_i)+\frac{\delta}{1-\delta}B(t)'\bar{g}^{(0)}(e_i)\\
&+\frac{1}{2}\frac{\delta}{1-\delta}B(t)'\big(L^{(0)}(e_i)+H^{(0)}(e_i)\big)B(t)+\frac{1}{2}B(t)'H^{(0)}(e_i)B(t)\big]=-\sum_{z=1}^{J}q_{i,z}\xi(t,e_z)\\
&\xi(T,e_i)=1,\forall i=1,\dots,I.
\end{aligned}
\tag{5.27}
$$

As function $w(t,e_i)$ is continuously differentiable in t for all $e_i\in\mathcal{E}$, we can apply the Feynman-Kac theorem for Markov chains from Corollary 2.78 to obtain the representation from (5.23). The equation obtained for B is given by (5.22). So, if we find a solution to System (5.22) then we have found also a solution to the HJB equation. For the verification result we can apply either directly Theorem 5.1 or use the time-dependent model together with Proposition 5.4 and Theorem 3.5. For the latter possibility note that from Theorem 3.5 we obtain the HJB solution in the time-dependent model as follows:

$$
\Phi^m(t,v,x)=\frac{v^\delta}{\delta}\exp\Big\{\int_t^T w\big(s,m(s)\big)\mathrm{d}s\Big\}\exp\big\{B(t)'x\big\}. \tag{5.28}
$$

As this function is sufficiently differentiable, the required martingale property of $\{\Phi^m(t,V^{m,\bar{\pi}^m}(t),X^m(t))\}_{t\in[0,T]}$ yields the verification result.

□

Remark 5.11 *System (5.22) is equivalent to the following system of (in generally) coupled ODEs for the single components of B:*

$$
\begin{aligned}
&\frac{\partial}{\partial t}B_j(t)+\delta\bar{\varepsilon}_j^{(1)}+\frac{1}{2}\frac{\delta}{1-\delta}\bar{h}_j^{(1)}-\sum_{d=1}^{J}B_d(t)K_{jd}^{(1)}+\frac{\delta}{1-\delta}\sum_{d=1}^{J}B_d(t)G_{jd}^{(1)}\\
&+\frac{1}{2}\frac{\delta}{1-\delta}B(t)'\big(L^{(1j)}+H^{(1j)}\big)B(t)+\frac{1}{2}B(t)'H^{(1j)}B(t)=0\\
&B_j(T)=0,\forall j\in\{1,\dots,J\}.
\end{aligned}
\tag{5.29}
$$

In the special case where $K^{(1)}$ and $G^{(1)}$ are diagonal matrices, and $L^{(1j)}$ and $H^{(1j)}$ have a non-zero element only at position (j,j), we have J not coupled Riccati ODEs with constant coefficients, which can be solved by Corollary 2.44.

Note that under some conditions on the model Proposition 5.6 can be applied to show that $\{\Phi^m(t,V^{m,\bar{\pi}^m}(t),X^m(t))\}_{t\in[0,T]}$ is a martingale. We do this for Example 2 in Section 5.3.

Remark 5.12
If we require only that $\Sigma\Sigma'$ is invertible instead of Σ itself being quadratic and invertible, then the PDE System (5.3) for function f is modified as follows:

$$\begin{aligned}
&f_t + f\delta\big\{r + \frac{1}{2}\frac{1}{1-\delta}(\mu - r)'\big(\Sigma\Sigma'\big)^{-1}(\mu - r)\big\} \\
&+ f_x'\big\{\mu^X + \frac{\delta}{1-\delta}\Sigma^X\rho\Sigma'(\Sigma\Sigma')^{-1}(\mu - r))\big\} + \frac{1}{2}Tr\big(\Sigma^X(\Sigma^X)'f_{xx'}\big) \\
&+ \frac{1}{2f}f_x'\frac{\delta}{1-\delta}\Sigma^X\rho\Sigma'(\Sigma\Sigma')^{-1}\Sigma\rho'(\Sigma^X)'f_x\Big|_{(t,x,e_i)} = -\sum_{z=1}^{I} q_{i,z}f(t,x,e_z) \\
&f(T,x,e_i) = 1, \forall i \in \{1,\dots,I\},
\end{aligned} \tag{5.30}$$

and strategy $\bar{\pi}$ has the following form:

$$\bar{\pi}(t) = \frac{1}{1-\delta}(\Sigma\Sigma')^{-1}\Big\{(\mu - r) + \Sigma\rho'(\Sigma^X)'\frac{f_x}{f}\Big\}\Big|_{(t,X(t),\mathcal{MC}(t))}. \tag{5.31}$$

Considering the definition in Remark 3.3, all other calculations and proofs remain valid.

After deriving the general theoretical results we show their relevance by presenting two examples in what follows. The first one includes stochastic interest rates and models simultaneously the stock and bond markets (see Section 5.3.1). The second one covers a multidimensional generalization of the Heston model (see Section 5.3.2).

5.3 Examples

5.3.1 Example 1: Markov-modulated Black-Scholes model with a stochastic short rate

In this case the stochastic factor X models a stochastic riskless interest rate and thus investment is possible not only in the bank account and a stock but also in a bond. The Markov chain is interpreted as the state of the economy. In the numerical example we consider at the end of this section it switches between two states: a calm period and a recession.
To model the stochastic short rate we adopt the Vasicek model (see Example 2.48), i.e. under the risk-neutral measure $\mathcal{Q}$ the SDE for X is given as follows:

$$\mathrm{d}X(t) = \kappa\big(\tilde{\theta} - X(t)\big)\mathrm{d}t + \chi\mathrm{d}\tilde{W}^X(t),$$

where $\tilde{W}^X$ is a $\mathcal{Q}$-Brownian motion, $\kappa, \chi \in [0,\infty)$. We denote the market price of interest rate risk by $\lambda_1\big(\mathcal{MC}(t)\big)$ and allow for dependence on the Markov chain. More precisely, the change from the real-world to the risk-neutral measure is given by $\frac{d\mathcal{Q}}{d\mathbb{P}}\Big|_{\mathscr{F}^X} = L(T)$ where:

$$L(t) = \exp\Big\{\int_0^t \lambda_1\big(\mathcal{MC}(s)\big)\mathrm{d}W^X(s) - \frac{1}{2}\int_0^t \parallel \lambda_1\big(\mathcal{MC}(s)\big) \parallel^2 \mathrm{d}s\Big\}.$$

As the Markov chain can take values in a finite set, Novikov's condition (see Lemma 2.38) is trivially fulfilled for L. Thus, by Girsanov's Theorem 2.40 process W_1^X defined by:

$$W^X(t) = \tilde{W}^X(t) + \int_0^t \lambda_1\big(\mathcal{MC}(s)\big)\mathrm{d}s$$

is a $\mathbb{P}$-Brownian motion. Then, under the real-world measure $\mathbb{P}$, process X is characterized by the following SDE:

$$\mathrm{d}X(t) = \kappa\Big(\underbrace{\tilde{\theta} - \frac{\chi\lambda_1\big(\mathcal{MC}(t)\big)}{\kappa}}_{=\theta(\mathcal{MC}(t))} - X(t)\Big)\mathrm{d}t + \chi\mathrm{d}W^X(t).$$

See [105] and [37] for bond pricing and parameters estimation of short rate Vasicek models extended by Markov switching.
As the interest rate is modeled by a stochastic process, a zero-coupon bond cannot be replicated by the bank account. That is why we allow for the investor to trade also in a zero-coupon bond with maturity $T_1 > T$ additional to its investment in the risky stock and the bank account. The price of this bond at time $t < T_1$ is given by

$$P_1(t, T_1, X(t)) = \mathbb{E}_{\mathcal{Q}}\Big[\exp\Big\{-\int_t^{T_1} X(s)\mathrm{d}s\Big\}\Big|\mathscr{F}_t^X\Big].$$

Thus, by Feynman-Kac theorem (see Corollary 2.76) it satisfies the following PDE:

$$(P_1)_t + (P_1)_x\kappa\big(\tilde{\theta} - X(t)\big) + \frac{1}{2}\chi^2(P_1)_{xx} - P_1 x = 0, P_1(T_1, T_1, x) = 1. \tag{5.32}$$

Recall from [99] that its solution is given by:

$$P_1(t, T_1, X(t)) = \exp\big\{-A_1(T_1 - t) - A_2(T_1 - t)X(t)\big\}, \tag{5.33}$$

where for $\tau \in [0, T_1]$:

$$A_1(\tau) = \big(\tilde{\theta} - \frac{\chi^2}{2\kappa^2}\big)\big(\tau - A_2(\tau)\big) + \frac{\chi^2}{4\kappa}A_2^2(\tau)$$
$$A_2(\tau) = \frac{1}{\kappa}\big(1 - \exp\{-\kappa\tau\}\big).$$

Then, the SDE for $P_1(t) := P_1(t, T_1, X(t))$ can be derived by applying Itô's formula for Markov-modulated processes and substituting PDE (5.32):

$$\begin{aligned}
\mathrm{d}P_1(t) =& \big[(P_1)_t\big(t,T_1,X(t)\big) + (P_1)_x\big(t,T_1,X(t)\big)\kappa\big(\theta - X(t)\big) \\
&+ \frac{1}{2}\chi^2 (P_1)_{xx}\big(t,T_1,X(t)\big)\big]\mathrm{d}t + (P_1)_x\big(t,T_1,X(t)\big)\chi \mathrm{d}W^X \\
=& \big[(P_1)_t\big(t,T_1,X(t)\big) + (P_1)_x\big(t,T_1,X(t)\big)\kappa\big(\tilde{\theta} - X(t)\big) \qquad (5.34)\\
&- (P_1)_x\big(t,T_1,X(t)\big)\chi\lambda_1\big(\mathcal{MC}(t)\big) + \frac{1}{2}\chi^2 (P_1)_{xx}\big(t,T_1,X(t)\big)\big]\mathrm{d}t \\
&+ (P_1)_x\big(t,T_1,X(t)\big)\chi \mathrm{d}W^X \\
\overset{(5.32)}{=}& \big[P_1\big(t,T_1,X(t)\big)X(t) - (P_1)_x\big(t,T_1,X(t)\big)\chi\lambda_1\big(\mathcal{MC}(t)\big)\big]\mathrm{d}t \\
&+ (P_1)_x\big(t,T_1,X(t)\big)\chi \mathrm{d}W^X \\
=& \big[P_1(t)X(t) + P_1(t)A_2(T_1-t)\chi\lambda_1\big(\mathcal{MC}(t)\big)\big]\mathrm{d}t - P_1(t)A_2(T_1-t)\chi \mathrm{d}W^X \\
=& P_1(t)\Big[\big\{X(t) + A_2(T_1-t)\chi\lambda_1\big(\mathcal{MC}(t)\big)\big\}\mathrm{d}t + A_2(T_1-t)\chi \mathrm{d}W_1^P\Big], \qquad (5.35)
\end{aligned}$$

where $W_1^P := -W^X$.
Besides the bond the investor has the opportunity to invest in a stock with price process denoted by P_2. We assume that the stock follows a geometric Brownian motion, where both the market price of risk and the volatility of the stock switch with the Markov chain. Furthermore we allow for regime-switching correlation ρ_{12} between the Brownian motions driving the bond and the stock. This flexibility of the model is in accordance with the empirical observation that the correlation between the bond and stock markets changes between the different states of the economy (see [11]).
To summarize, the model has the following dynamics under the real-world measure $\mathbb{P}$:

$$\begin{aligned}
\mathrm{d}X(t) =& \kappa\big(\theta(\mathcal{MC}(t)) - X(t)\big)\mathrm{d}t + \chi \mathrm{d}W^X(t) \\
\mathrm{d}P_0(t) =& P_0(t)X(t)\mathrm{d}t \\
\mathrm{d}P_1(t) =& P_1(t)\big[\big\{X(t) + \lambda_1\big(\mathcal{MC}(t)\big)\chi A_2(T_1-t)\big\}\mathrm{d}t + \chi A_2(T_1-t)\mathrm{d}W_1^P\big] \\
\mathrm{d}P_2(t) =& P_2(t)\big[X(t) + \lambda_2\big(\mathcal{MC}(t)\big)\nu\big(\mathcal{MC}(t)\big)\mathrm{d}t + \nu\big(\mathcal{MC}(t)\big)\rho_{12}\big(\mathcal{MC}(t)\big)\mathrm{d}W_1^P \\
&+ \nu\big(\mathcal{MC}(t)\big)\sqrt{1-\rho_{12}\big(\mathcal{MC}(t)\big)^2}\mathrm{d}W_2^P\big] \\
\rho\mathrm{d}t =& \big(\mathrm{d}\langle W^X, W_1^P\rangle(t), \mathrm{d}\langle W^X, W_2^P\rangle(t)\big) = (-1,0)\mathrm{d}t,
\end{aligned} \qquad (5.36)$$

where W^X is a one-dimensional Brownian motion, $\mathrm{d}W_1^P = -\mathrm{d}W^X$ and $W^P = (W_1^P, W_2^P)$ is a standard two-dimensional Brownian motion. Utility maximization in a similar model without Markov switching is considered in [73]. In terms of the notation from Model (3.1) we have the following specifications:

$$\Sigma = \begin{pmatrix} \chi A_2(T_1-t) & 0 \\ \nu\big(\mathcal{MC}(t)\big)\rho_{12}\big(\mathcal{MC}(t)\big) & \nu\big(\mathcal{MC}(t)\big)\sqrt{1-\rho_{12}\big(\mathcal{MC}(t)\big)^2} \end{pmatrix}$$
$$\mu - r = \begin{pmatrix} \lambda_1\big(\mathcal{MC}(t)\big)\chi A_2(T_1-t) \\ \lambda_2\big(\mathcal{MC}(t)\big)\nu\big(\mathcal{MC}(t)\big) \end{pmatrix}$$
$$\mu^X = \kappa\theta\big(\mathcal{MC}(t)\big) - \kappa X(t)$$
$$\Sigma^X = \chi,$$

which implies that:

$$\begin{aligned}
&\varepsilon^{(0)} = 0, \bar{\varepsilon}^{(1)} = 1 \\
&\bar{k}^{(0)} = \kappa\theta(e_i), K^{(1)} = \kappa \\
&H^{(0)} = \chi^2, H^{(11)} = 0 \\
&h^{(0)} = \frac{\lambda_1(e_i)^2 - 2\rho_{12}(e_i)\lambda_1(e_i)\lambda_2(e_i) + \lambda_2(e_i)^2}{1-\rho_{12}(e_i)^2}, \bar{h}^{(1)} = 0 \\
&\bar{g}^{(0)} = -\chi\lambda_1(e_i), G^{(1)} = 0 \\
&L^{(0)} = 0, L^{(11)} = 0.
\end{aligned} \tag{5.37}$$

So, the considered model fits in our general framework and allows the application of Theorem 5.10. The result is stated in the following proposition:

Proposition 5.13 (Solution and verification in Example 1)
Consider Model (5.36). The value function for the portfolio optimization problem is given as follows:

$$\Phi(t, v, x, e_i) = \frac{v^\delta}{\delta}\xi(t, e_i)\exp\{B(t)x\}, \tag{5.38}$$

with

$$B(t) = \frac{\delta}{\kappa}\big(1 - \exp\big\{-\kappa(T-t)\big\}\big) \tag{5.39}$$
$$\xi(t, e_i) = \mathbb{E}\Big[\exp\Big\{\int_t^T w(s, \mathcal{MC}(s))\mathrm{d}s\Big\}\Big|\mathcal{MC}(t) = e_i\Big] \tag{5.40}$$
$$\begin{aligned} w(t, e_i) =& \frac{1}{2}\frac{\delta}{1-\delta}\frac{\lambda_1(e_i)^2 - 2\rho_{12}(e_i)\lambda_1(e_i)\lambda_2(e_i) + \lambda_2(e_i)^2}{1-\rho_{12}(e_i)^2} + B(t)\kappa\theta(e_i) \\ &- \frac{\delta}{1-\delta}B(t)\lambda_1(e_i)\chi + \frac{1}{2}\frac{1}{1-\delta}B(t)^2\chi^2. \end{aligned} \tag{5.41}$$

The optimal portfolio is[1]*:*

$$\bar{\pi}(t) = \frac{1}{1-\delta}\begin{pmatrix} \frac{\lambda_1(\mathcal{MC}(t))-\rho_{12}(\mathcal{MC}(t))\lambda_2(\mathcal{MC}(t))}{\chi A_2(T_1-t)(1-\rho_{12}^2(\mathcal{MC}(t)))} - \frac{B(t)}{A_2(T_1-t)} \\ \frac{\lambda_2(\mathcal{MC}(t))-\rho_{12}(\mathcal{MC}(t))\lambda_1(\mathcal{MC}(t))}{\nu(\mathcal{MC}(t))(1-\rho_{12}^2(\mathcal{MC}(t)))} \end{pmatrix}. \tag{5.42}$$

The proof can be found in Appendix C.

Now that we have verified the result from the theoretical point of view, let us have a closer look at the structure of the optimal portfolio (5.42) and especially at the influence of the Markov-switching parameters. As expected the factors driving the excess return of the single assets influence positively the positions in the corresponding assets, i.e. the bigger λ_i, the higher $\bar{\pi}_i$. The contrary holds for the volatility terms. As the two assets are correlated, the excess return of one asset influences also the position in the other one. More precisely, if the correlation ρ_{12} is positive, the two assets (partially) act as substitutes: e.g. if $\bar{\pi}_1$ rises as a result of rising λ_1, whereas all other parameters remain the same, then $\bar{\pi}_2$ is reduced. On the other side, if ρ_{12} is negative, the two assets can be seen as complements and are used for diversification: an increase of $\bar{\pi}_1$ due to a higher λ_1, leads to an increase of $\bar{\pi}_2$.
Let us deepen our analysis of the influence of the Markov-switching correlation factor ρ_{12} on the optimal portfolio. Observe that $sign(\frac{\partial\bar{\pi}_1}{\partial\rho_{12}}) = sign(-\lambda_2 - \lambda_2\rho_{12}^2 + 2\lambda_1\rho_{12})$ and, if $\nu > 0$, $sign(\frac{\partial\bar{\pi}_2}{\partial\rho_{12}}) = sign(-\lambda_1 - \lambda_1\rho_{12}^2 + 2\lambda_2\rho_{12})$. Let us concentrate on the case where $\lambda_i > 0$, $i = 1, 2$, as this corresponds to the classical setting, where assets have positive market price of risk. First consider the case when $\rho_{12} < 0$. Then $\frac{\partial\bar{\pi}_1}{\partial\rho_{12}} < 0$ and $\frac{\partial\bar{\pi}_2}{\partial\rho_{12}} < 0$. So, if ρ_{12} increases, which means that its absolute value decreases, the investor should hold less in both assets. The reason is that assets with negative correlation are used for diversification, so the smaller $|\rho_{12}|$, the less effective the diversification. In the second case, where $\rho_{12} > 0$, the influence of the correlation depends on the relationship between λ_1 and λ_2. One can differentiate the following three cases:

$$\begin{cases} \frac{\lambda_1}{\lambda_2} > \frac{1+\rho_{12}^2}{2\rho_{12}}(> 1) & \Rightarrow \frac{\partial\bar{\pi}_1}{\partial\rho_{12}} > 0, \frac{\partial\bar{\pi}_2}{\partial\rho_{12}} < 0 \\ \frac{1+\rho_{12}^2}{2\rho_{12}} \geq \frac{\lambda_1}{\lambda_2} \geq \frac{2\rho_{12}}{1+\rho_{12}^2} & \Rightarrow \frac{\partial\bar{\pi}_1}{\partial\rho_{12}} \leq 0, \frac{\partial\bar{\pi}_2}{\partial\rho_{12}} \leq 0 \\ (1 >)\frac{2\rho_{12}}{1+\rho_{12}^2} > \frac{\lambda_1}{\lambda_2} & \Rightarrow \frac{\partial\bar{\pi}_1}{\partial\rho_{12}} < 0, \frac{\partial\bar{\pi}_2}{\partial\rho_{12}} > 0. \end{cases}$$

Note that if one of the assets clearly outperforms the other one, i.e. its market price of risk is sufficiently higher that the market price of risk of the other asset, then the

[1]In the case of constant parameters without Markov switching Formula (5.42) leads to the optimal portfolio given by [73]. Note that in the latter study the authors also present a proof of the verification theorem without Markov switching based on a general result in [50], p. 163. They require explicitly the continuity of the model parameters. Instead of loosening this assumption for our proof we prefer to apply directly Theorem 5.10 in a straightforward manner to obtain an alternative proof that is shorter and easy to follow.

higher the positive correlation the bigger the investment in the better performing asset and the smaller the position in the other one. As mentioned before, the two assets function as partial substitutes. Only in a certain interval, where λ_1 and λ_2 are not sufficiently different, increasing correlation leads to a reduction of both positions, because the substitution effect is not strong enough to compensate the risk of higher correlation. However, this interval gets smaller for higher ρ_{12} as the two thresholds $\frac{2\rho_{12}}{1+\rho_{12}^2}$ and $\frac{1+\rho_{12}^2}{2\rho_{12}}$ approach one from below and above respectively, when ρ_{12} increases. So when the correlation becomes high enough, the investor starts expanding her increasing position in the better performing asset because he can hedge it better by her decreasing position in the other one. To summarize, starting at a negative ρ_{12}, both $\bar{\pi}_1$ and $\bar{\pi}_2$ decrease until ρ_{12} reaches some critical level $\hat{\rho}$ given by $\frac{2\hat{\rho}}{1+\hat{\rho}^2} = \frac{\lambda_{n_1}}{\lambda_{n_2}}$, where $\lambda_{n_1} < \lambda_{n_2}$, $n_1, n_2 \in \{1, 2\}$, $n_1 \neq n_2$. If ρ_{12} exceeds $\hat{\rho}$ and increases further, then $\bar{\pi}_{n_1}$ continues decreasing and $\bar{\pi}_{n_2}$ gets bigger.
Now let us pay attention to the second part of $\bar{\pi}_1$: $-\frac{1}{1-\delta}\frac{B(t)}{A_2(T_1-t)}$. This fraction of the wealth is shifted between the bond position and the bank account because of the additional risk coming from the stochastic interest rate, that is why we will call it the hedging term. Observe that $sign(\frac{1}{1-\delta}\frac{B(t)}{A_2(T_1-t)}) = sign(\delta)$. So, for a very risk averse investor with $\delta < 0$ the bond position is increased. The reason is that the relevant risk for the investor comes from a falling short rate as this would reduce the drift of her wealth process (see Equation (3.9)). If this happens the bond price would increase due to its negative correlation with the short rate. So, the investor protects his portfolio against decreasing X with a positive bond hedging term. On the contrary, a less risk averse investor with $\delta > 0$ speculates on an increasing short rate and enters a short bond position in order to profit it this case.
To summarize, it is crucial to consider multidimensional models, as the relation between the assets influences strongly the optimal portfolio. Furthermore, the Markov chain plays an important role in the portfolio choice, as the switching parameters drive the performance of the single assets considered for themselves and in comparison to each other.
In order to illustrate this analysis we fix a realistic set of parameters following the empirical results of [17], [11], and [57] and implement the optimal solution. More precisely we set: $\lambda_1(e_1) = 0.1$, $\lambda_1(e_2) = 0.3$, $\lambda_2(e_1) = 0.26$, $\lambda_2(e_2) = -0.22$, $\nu(e_1) = 0.13$, $\nu(e_2) = 0.20$, $\rho_{12}(e_1) = -0.14$, $\rho_{12}(e_2) = -0.34$, $\chi = 0.03$, $\kappa = 0.15$, $\theta(e_1) = 0.08$, $\theta(e_2) = 0.04$. Clearly, e_1 corresponds to a calm state of the economy with relatively low volatility. The second state e_2 describes an economy in a recession, so the volatility is higher and the expected stock return and the log-term mean of the short rate are lower than in e_1. Note, that the negative correlation between the bond and the stock is stronger in the second state. For the intensity matrix of the Markov chain we set $q_{1,1} = -1.0909$ and $q_{2,2} = -3.4413$, which implies that the calm state lasts on average 1 year and the turbulent one approximately 4 months. We consider four different values for δ, corresponding to investors with different risk preferences: $\delta \in \{0.1, -1, -5, -10\}$, see [22] for some discussion on the topic of how to choose δ. Figure 5.1 illustrates the resulting optimal portfolios. One can recognize the big differences between the two states of the economy: in the second

state the investor shifts weight from the stock to the bond, whose price is positively effected by the turbulent situation in the stock market.

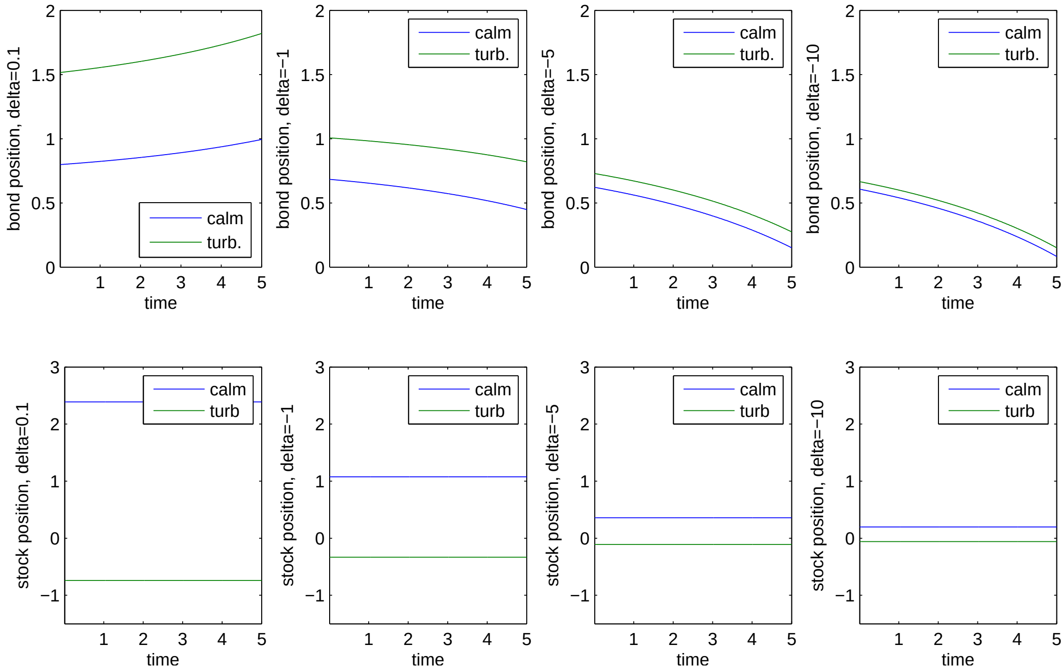


Figure 5.1: Example 1: Fraction of wealth invested in a bond with maturity 15 years (upper row) and in the stock (lower row) for an investor with time horizon $T = 5$ and different risk parameters δ.

As an illustration we also plotted the distribution of the terminal wealth of an investor following the derived optimal strategy for different values of δ (see Figure 5.2). It can be clearly seen that a decrease in δ, thus an increase in the risk aversion, leads to a more conservative investment.

5.3.2 Example 2: Two-dimensional Markov-modulated stochastic correlation model

In the second example we consider a two-dimensional model with stochastic volatility and stochastic correlation between the assets. More precisely, two mean-reverting stochastic factors X_1 and X_2 influence the price processes of two traded assets P_1

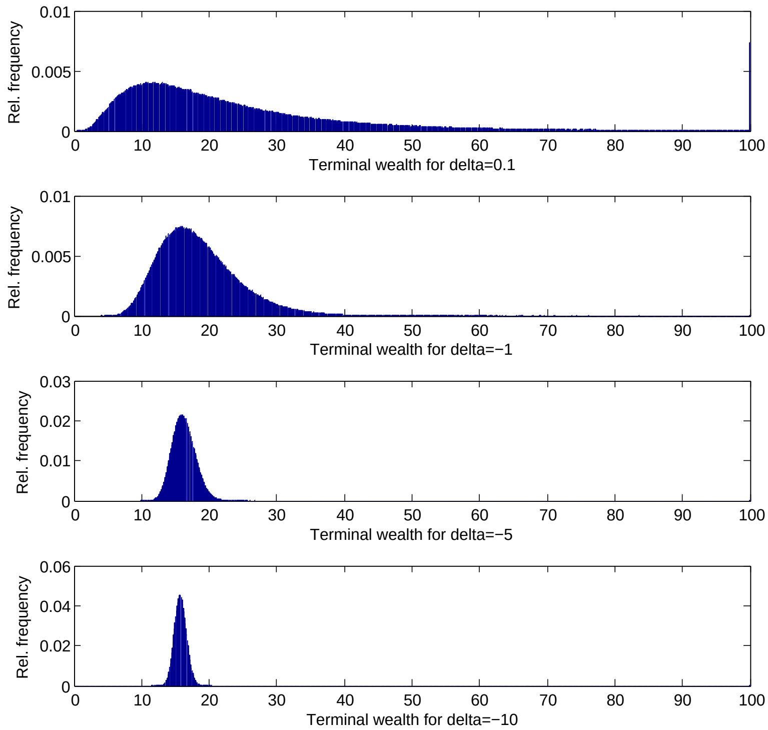


Figure 5.2: Example 1: Distribution of the terminal wealth obtained using the derived optimal strategy (5.42) with time horizon $T = 5$ and different risk parameters δ. For better comparability all values higher than 100 are summarized in the last bar.

and P_2. The formal definition of the model reads as follows:

$$
\begin{aligned}
\mathrm{d}P_0(t) =& P_0(t) r \mathrm{d}t \\
\mathrm{d}P_1(t) =& P_1(t)\Big[\big\{r + c_1 a_{11}\big(\mathcal{MC}(t)\big)\sqrt{X_1(t)}\sigma_1\big(X_1(t)\big) \\
&+ c_2 a_{12}\big(\mathcal{MC}(t)\big)\sqrt{X_2(t)}\sigma_2\big(X_2(t)\big)\big\}\mathrm{d}t \\
&+ a_{11}\big(\mathcal{MC}(t)\big)\sigma_1\big(X_1(t)\big)\mathrm{d}W_1^P(t) + a_{12}\big(\mathcal{MC}(t)\big)\sigma_2\big(X_2(t)\big)\mathrm{d}W_2^P(t)\Big] \\
\mathrm{d}P_2(t) =& P_2(t)\Big[\big\{r + c_1 a_{21}\big(\mathcal{MC}(t)\big)\sqrt{X_1(t)}\sigma_1\big(X_1(t)\big) \\
&+ c_2 a_{22}\big(\mathcal{MC}(t)\big)\sqrt{X_2(t)}\sigma_2\big(X_2(t)\big)\big\}\mathrm{d}t \\
&+ a_{21}\big(\mathcal{MC}(t)\big)\sigma_1\big(X_1(t)\big)\mathrm{d}W_1^P(t) + a_{22}\big(\mathcal{MC}(t)\big)\sigma_2\big(X_2(t)\big)\mathrm{d}W_2^P(t)\Big] \\
\mathrm{d}X(t) =& \begin{pmatrix} \kappa_1\big(\theta_1\big(\mathcal{MC}(t)\big) - X_1(t)\big) \\ \kappa_2\big(\theta_2\big(\mathcal{MC}(t)\big) - X_2(t)\big) \end{pmatrix}\mathrm{d}t + \begin{pmatrix} \chi_1\sqrt{X_1(t)}\mathrm{d}W_1^X(t) \\ \chi_2\sqrt{X_2(t)}\mathrm{d}W_2^X(t) \end{pmatrix} \\
\rho\mathrm{d}t =& \begin{pmatrix} \mathrm{d}\langle W_1^X, W_1^P\rangle(t) & \mathrm{d}\langle W_1^X, W_2^P\rangle(t) \\ \mathrm{d}\langle W_2^X, W_1^P\rangle(t) & \mathrm{d}\langle W_2^X, W_2^P\rangle(t) \end{pmatrix} = \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix}\mathrm{d}t,
\end{aligned}
\tag{5.43}
$$

where $W^P = (W_1^P, W_2^P)$ and $W^X = (W_1^X, W_2^X)$ are standard Brownian motions. We denote $A\big(\mathcal{MC}(t)\big) = \begin{pmatrix} a_{11}\big(\mathcal{MC}(t)\big) & a_{12}\big(\mathcal{MC}(t)\big) \\ a_{21}\big(\mathcal{MC}(t)\big) & a_{22}\big(\mathcal{MC}(t)\big) \end{pmatrix}$. To facilitate the calculations in the subsequent analysis we summarize the model in terms of the notation from (3.1):

$$\begin{aligned}
\begin{pmatrix} \mu_1 - r \\ \mu_2 - r \end{pmatrix} =& A\big(\mathcal{MC}(t)\big) \begin{pmatrix} c_1\sqrt{X_1(t)}\sigma_1(X_1) \\ c_2\sqrt{X_2(t)}\sigma_2(X_2) \end{pmatrix} \\
\Sigma =& \begin{pmatrix} a_{11}\big(\mathcal{MC}(t)\big)\sigma_1\big(X_1(t)\big) & a_{12}\big(\mathcal{MC}(t)\big)\sigma_2\big(X_2(t)\big) \\ a_{21}\big(\mathcal{MC}(t)\big)\sigma_1\big(X_1(t)\big) & a_{22}\big(\mathcal{MC}(t)\big)\sigma_2\big(X_2(t)\big) \end{pmatrix} \\
=& A\big(\mathcal{MC}(t)\big) \begin{pmatrix} \sigma_1\big(X_1(t)\big) & 0 \\ 0 & \sigma_2\big(X_2(t)\big) \end{pmatrix} \\
\mu^X =& \begin{pmatrix} \mu_1^X \\ \mu_2^X \end{pmatrix} = \begin{pmatrix} \kappa_1\big(\theta_1\big(\mathcal{MC}(t)\big) - X_1(t)\big) \\ \kappa_2\big(\theta_2\big(\mathcal{MC}(t)\big) - X_2(t)\big) \end{pmatrix} \\
\Sigma^X =& \begin{pmatrix} \chi_1\sqrt{X_1} & 0 \\ 0 & \chi_2\sqrt{X_2} \end{pmatrix}.
\end{aligned}$$

Using the notation from (3.2)-(3.7) we obtain the following equations:

$$\begin{aligned}
&\varepsilon^{(0)} = r, \varepsilon^{(1)} = 0 \\
&\bar{k}^{(0)}(e_i) = \begin{pmatrix} \kappa_1\theta_1(e_i) \\ \kappa_2\theta_2(e_i) \end{pmatrix}, K^{(1)} = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix} \\
&H^{(0)} = 0, H^{(11)} = \begin{pmatrix} \chi_1^2 & 0 \\ 0 & 0 \end{pmatrix}, H^{(12)} = \begin{pmatrix} 0 & 0 \\ 0 & \chi_2^2 \end{pmatrix} \\
&h^{(0)} = 0, \bar{h}^{(1)} = \begin{pmatrix} c_1^2 \\ c_2^2 \end{pmatrix} \\
&\bar{g}^{(0)} = 0, G^{(1)} = \begin{pmatrix} \chi_1\rho_1 c_1 & 0 \\ 0 & \chi_2\rho_2 c_2 \end{pmatrix} \\
&L^{(0)} = 0, L^{(11)} = \begin{pmatrix} \chi_1^2(\rho_1^2 - 1) & 0 \\ 0 & 0 \end{pmatrix}, L^{(12)} = \begin{pmatrix} 0 & 0 \\ 0 & \chi_2^2(\rho_2^2 - 1) \end{pmatrix}.
\end{aligned}$$

Note that the market price of risk associated to the two Brownian motions W_1^P and W_2^P driving the asset processes has the following form:

$$\Sigma^{-1}(\mu - r) = \begin{pmatrix} c_1\sqrt{X_1} \\ c_2\sqrt{X_2} \end{pmatrix}.$$

So, it resembles the structure of the market price of risk in the one-dimensional Heston model.

This model presents a very flexible framework that covers different known examples as special cases. If the vectors $(a_{11}, a_{21})'$ and $(a_{12}, a_{22})'$ are assumed to be orthonormal then we are dealing with an example of a principle component model (see [43], [44] and [46]). The othonormality assumption might be useful to maintain the number of parameters manageable when the dimension increases, however this is not a necessary assumption for the derivation of the optimal portfolio and we do not imply it in what follows.

Note that if we set $\sigma_1(x_1) = \sqrt{x_1}$ and $\sigma_2(x_2) = \sqrt{x_2}$ the model can be seen as a two-dimensional extension of the Heston Model, where not only the covariance matrix,

but also the correlation between the traded assets is stochastic. The two stochastic factors may also be used to model stochastic factors with different scales of mean-reversion: a slow one and a fast one, see e.g. [51] and [45] for a motivation and discussion of stochastic volatility with slow and fast mean-reverting components. Observe furthermore, that we allow for correlation between the Brownian motions driving X and P, a desirable property observed in reality (see [41] for a discussion of the topic in a stochastic volatility context). Portfolio optimization results for the one-dimensional Heston model with constant parameters are presented in [74] and [63]. By merging ideas from these two papers and applying the results derived in Section 3 we deliver in Proposition 5.14 an easy to follow proof of the verification result for a multidimensional Markov-switching extension of the Heston model.
Another application of Model (5.43) arises if we set σ_1 and σ_2 to be constants and let the stochastic factors influence only the excess return of the stocks. A similar framework has been motivated in the literature on predictable stock returns, see [93] and [4] for a discussion of models wherein the stock returns are driven by the labor income or the dividend yield.
The Markov chain can be again interpreted as the state of the economy. We allow for regime switching in the mean-reverting level of the stochastic factors, the stock volatility and its excess return. As an illustration we will present at the end of this section some numerical results for a stochastic volatility model where the first state of $\mathcal{MC}$ represents a calm state and the second one corresponds to a turbulent market with high volatility levels.
The explicit solution to the optimization problem and the corresponding verification result are shown in the following proposition:

Proposition 5.14 (Solution and verification in Example 2)
Consider Model (5.43) and assume that:

$$0 < \tilde{\kappa}_j \tag{5.44}$$

$$\frac{\delta}{1-\delta} c_j^2 < \frac{\tilde{\kappa}_j^2 \vartheta_j}{\chi_j^2}, \tag{5.45}$$

with $\tilde{\kappa}_j = \kappa_j - \frac{\delta}{1-\delta}\chi_j \rho_j c_j$, $\vartheta_j = \frac{1-\delta}{1-\delta+\delta\rho_j^2}$ *and* $a_j = \sqrt{\tilde{\kappa}_j^2 - \frac{\delta}{1-\delta} c_j^2 \frac{\chi_j^2}{\vartheta_j}}$, *for* $j = 1, 2$. *Then the value function for the considered portfolio optimization problem is given by:*

$$\Phi(t, v, x, e_i) = \frac{v^\delta}{\delta} \xi(t, e_i) \exp\{B_1(t) x_1 + B_2(t) x_2\},$$

where

$$B_j(t) = \frac{\vartheta_j (\tilde{\kappa}_j - a_j)\big[1 - \exp\{-a_j(T-t)\}\big]}{\chi_j^2 \big(1 - b_j \exp\{-a_j(T-t)\}\big)}, j = 1, 2 \tag{5.46}$$

$$\xi(t, e_i) = \mathbb{E}\Big[\exp\Big\{\int_t^T w(s, \mathcal{MC}(s)) \mathrm{d}s\Big\}\Big|\mathcal{MC}(t) = e_i\Big] \tag{5.47}$$

$$w(t, e_i) = \delta r + B_1 \kappa_1 \theta_1(e_i) + B_2 \kappa_2 \theta_2(e_i), \tag{5.48}$$

with $b_j = \frac{\tilde{\kappa}_j - a_j}{\tilde{\kappa}_j + a_j}$.
The optimal investment strategy has the following form (where the dependence of $a_{ij}, i, j = 1, 2$, *on* $\mathcal{MC}$ *is omitted for better readability):*

$$\begin{aligned}
\bar{\pi}(t) =& \frac{1}{1-\delta}\big(A'(\mathcal{MC}(t))\big)^{-1}\left\{\begin{pmatrix} \frac{c_1\sqrt{X_1(t)}}{\sigma_1(X_1(t))} \\ \frac{c_2\sqrt{X_2(t)}}{\sigma_2(X_2(t))}\end{pmatrix} + \begin{pmatrix} \frac{\rho_1\chi_1\sqrt{X_1(t)}B_1(t)}{\sigma_1(X_1(t))} \\ \frac{\rho_2\chi_2\sqrt{X_2(t)}B_2(t)}{\sigma_2(X_2(t))}\end{pmatrix}\right\} \\
=& \frac{1}{(1-\delta)(a_{11}a_{22}-a_{12}a_{21})}\left\{\begin{pmatrix} \frac{a_{22}c_1\sqrt{X_1(t)}}{\sigma_1(X_1(t))} - \frac{a_{21}c_2\sqrt{X_2(t)}}{\sigma_2(X_2(t))} \\ \frac{a_{11}c_2\sqrt{X_2(t)}}{\sigma_2(X_2(t))} - \frac{a_{12}c_1\sqrt{X_1(t)}}{\sigma_1(X_1(t))}\end{pmatrix}\right. \\
&+ \left.\left.\begin{pmatrix} \frac{a_{22}\rho_1\chi_1\sqrt{X_1(t)}B_1}{\sigma_1(X_1(t))} - \frac{a_{21}\rho_2\chi_2\sqrt{X_2(t)}B_2}{\sigma_2(X_2(t))} \\ \frac{a_{11}\rho_2\chi_2\sqrt{X_2(t)}B_2}{\sigma_2(X_2(t))} - \frac{a_{12}\rho_1\chi_1\sqrt{X_1(t)}B_1}{\sigma_1(X_1(t))}\end{pmatrix}\right\}\right|_{(t,\mathcal{MC}(t))}.
\end{aligned} \tag{5.49}$$

The proof can be found in Appendix C.

Again, we call the first summand in (5.49) the mean-variance part and denote it by $\bar{\pi}^{MV}$. It resembles the optimal portfolio if X_1 and X_2 were deterministic. The second summand is the hedging term $\bar{\pi}^{H}$. It accounts for the additional risk coming from the stochastic factors X_1 and X_2. In what follows we analyze these two terms and the sensitivity of the optimal portfolio to the model parameters. We will pay special attention to the Markov-switching parameters.
To ease the exposition we assume that $a_{ij}, c_j \geq 0, i, j = 1, 2$. Furthermore, we set $\sigma_1(x_1) = \sqrt{x_1}$ and $\sigma_2(x_2) = \sqrt{x_2}$, which, as already mentioned, leads to a two-dimensional Heston-type model. Then, the expression for the strategy simplifies to:

$$\begin{aligned}
\bar{\pi}(t) =& \frac{1}{(1-\delta)(a_{11}a_{22}-a_{12}a_{21})}\left\{\begin{pmatrix} a_{22}c_1 - a_{21}c_2 \\ a_{11}c_2 - a_{12}c_1\end{pmatrix} + \right. \\
& \left.\left.\begin{pmatrix} a_{22}\rho_1\chi_1 B_1 - a_{21}\rho_2\chi_2 B_2 \\ a_{11}\rho_2\chi_2 B_2 - a_{12}\rho_1\chi_1 B_1\end{pmatrix}\right\}\right|_{(t,\mathcal{MC}(t))}.
\end{aligned} \tag{5.50}$$

To better understand the structure of the portfolio we define the following quantities:

$$\begin{aligned}
\bar{m}_1(e) :=& c_1 a_{11}(e)\theta_1(e) + c_2 a_{12}(e)\theta_2(e),\ \bar{m}_2(e) := c_1 a_{21}(e)\theta_1(e) + c_2 a_{22}(e)\theta_2(e) \\
(\bar{s}_1(e))^2 :=& (a_{11}(e))^2\theta_1(e) + (a_{12}(e))^2\theta_2(e),\ (\bar{s}_2(e))^2 := (a_{21}(e))^2\theta_1(e) + (a_{22}(e))^2\theta_2(e) \\
\bar{\rho}(e) :=& \frac{a_{11}(e)a_{21}(e)\theta_1(e) + a_{12}(e)a_{22}(e)\theta_2(e)}{\bar{s}_1(e)\bar{s}_2(e)}.
\end{aligned} \tag{5.51}$$

Observe that $\bar{m}_1$ and $\bar{m}_2$ correspond to the "long-term average" excess returns of P_1 and P_2, respectively. They are obtained by replacing processes X_1 and X_2 in $(\mu - r)$ by their mean-reversion levels θ_1 and θ_2, respectively. Analogously, $\bar{s}_1^2$ and

$\bar{s}_2^2$ correspond to the "long-term average" instantaneous variance of the log-returns and $\bar{\rho}$ is obtained from the instantaneous correlation between the two assets. Using this notation it is easily verified that the mean-variance part $\bar{\pi}^{MV}$ of the optimal portfolio can be expressed as follows:

$$\bar{\pi}^{MV} = \frac{1}{1-\delta} \begin{pmatrix} \frac{1}{1-\bar{\rho}^2}\frac{\bar{m}_1}{(\bar{s}_1)^2} - \frac{\bar{\rho}}{1-\bar{\rho}^2}\frac{\bar{m}_2}{\bar{s}_1\bar{s}_2} \\ \frac{1}{1-\bar{\rho}^2}\frac{\bar{m}_2}{(\bar{s}_2)^2} - \frac{\bar{\rho}}{1-\bar{\rho}^2}\frac{\bar{m}_1}{\bar{s}_1\bar{s}_2} \end{pmatrix}.$$

So, if $\bar{\rho} = 0$ the mean-variance portfolio for both assets reduces to the ratio between the "long-term average" excess return and the variance as a measure of risk. This is the well-known myopic portfolio obtained when considering only one risky asset with deterministic coefficients. This term is intuitively clear, as higher excess return leads to higher investment, whereas higher variance, i.e. higher risk, reduces the position in the corresponding asset. If the two assets are correlated they are considered relatively to each other, e.g. if $\bar{\rho}$ is positive, an increase of $\bar{m}_2$ leads not only to a higher position in P_2 but also to a reduction of the investment in P_1, as the performance of the second asset has improved relatively to the first one and the two assets are considered as partial substitutes. Later on we will illustrate the influence of the instantaneous asset correlation using a numerical example.

Now let us have a look at the hedging term. To separate the hedging effect from the interaction between the assets we consider the case when $a_{12} = a_{21} = 0$ and the two stocks are independent. Then the hedging term has the following form:

$$\bar{\pi}^H(t) = \frac{1}{(1-\delta)} \left\{ \begin{pmatrix} \frac{\rho_1\chi_1 B_1}{a_{11}} \\ \frac{\rho_2\chi_2 B_2}{a_{22}} \end{pmatrix} \right\}\Bigg|_{(t,\mathcal{MC}(t))}. \tag{5.52}$$

The sign of the hedging part depends on ρ_j and B_j. It can be shown that $sign(B_j) = sign(\delta)$[2]. Assuming that $\rho_j < 0$, which is related to the well-known leverage effect, it follows that the hedging term is negative if $\delta > 0$ and positive if $\delta < 0$. The reason is that the more risk-averse investor, characterized by $\delta < 0$, would like to reduce the impact of a falling market price of risk due to a decrease in X. Due to $\rho < 0$, when X is decreasing, the asset prices tend to increase. Hence, as the investor is looking for a hedging term with increasing value in this situation, she adds a long position in the risky assets, i.e. $\bar{\pi}^H > 0$. On the contrary, the less risk-averse investor with $\delta > 0$ wants to participate in an increasing market price of risk, when X is increasing and his hedging part consists of a short position in the risky asset. For similar interpretations consult [67] and [24]. For a more detailed analysis of the hedging term without Markov switching we refer the reader to [23]. A numerical illustration of the sensitivity of the hedging part to different model parameters is given at the end of this section.

We continue with an analysis of the optimal strategy on a concrete numerical example. We assume that the Markov chain has two states: e_1 describes a calm

[2] First note that $\vartheta_j > 0$. If $\delta > 0$, then $0 < a_j < |\tilde{\kappa}_j|$. As $\tilde{\kappa}_j + a_j > 0$, it follows that $\tilde{\kappa}_j > 0$. So, $0 < a_j < \tilde{\kappa}_j$. This leads to $0 < b_j < 1$. We can conclude that $B_j > 0$. The case $\delta < 0$ goes analogously.

period corresponding to normal market conditions and e_2 models a turbulent period characterized by lower excess returns and higher volatility for both assets, i.e. we can interpret the second state as a crisis. More precisely, we set: $\bar{m}_1(e_1) = 0.05$, $\bar{m}_2(e_1) = 0.065$, $\bar{s}_1(e_1) = 0.15$, $\bar{s}_2(e_1) = 0.2$ for the calm state and $\bar{m}_1(e_2) = 0.005$, $\bar{m}_2(e_2) = 0.01$, $\bar{s}_1(e_2) = 0.25$, $\bar{s}_2(e_2) = 0.4$ for the volatile one. Furthermore, we choose $\bar{\rho}(e_1) = 0.2$ and $\bar{\rho}(e_2) = 0.9$ to reflect the empirical observation that in a financial crisis the correlation between related assets may increase dramatically (see e.g. [11]). We set $a_{12} = 0$ for both states, so that factor X_1 is interpreted as the common driver for both assets, whereas factor X_2 is specific only for the second asset. This specification allows for an intuitive interpretation of matrix A in terms of the assets correlation: if $a_{21} \gg a_{22}$ the two assets are strongly correlated, whereas if $a_{21} \ll a_{22}$ they are mainly driven by different factors. Using System (5.51) we set $A(e_1) = \begin{pmatrix} 0.68 & 0 \\ 0.18 & 0.74 \end{pmatrix}$, $A(e_2) = \begin{pmatrix} 18.75 & 0 \\ 27 & 10.86 \end{pmatrix}$, $c_1 = 1.5$, $c_2 = 1$, $\theta_1(e_1) = 0.05$, $\theta_1(e_2) = 0.0002$, $\theta_2(e_1) = 0.06$, $\theta_2(e_2) = 0.0003$[3]. Furthermore, $\rho_1 = -0.8$, $\rho_2 = -0.6$, $\chi_1 = \chi_2 = 0.35$, $\kappa_1 = \kappa_2 = 4$. We adopt the same values for Q as in Example 1 and set $r = 0.05$. Call these parameter specifications Set 1.
The optimal investment strategy for different values for δ is presented in Figure 5.3.

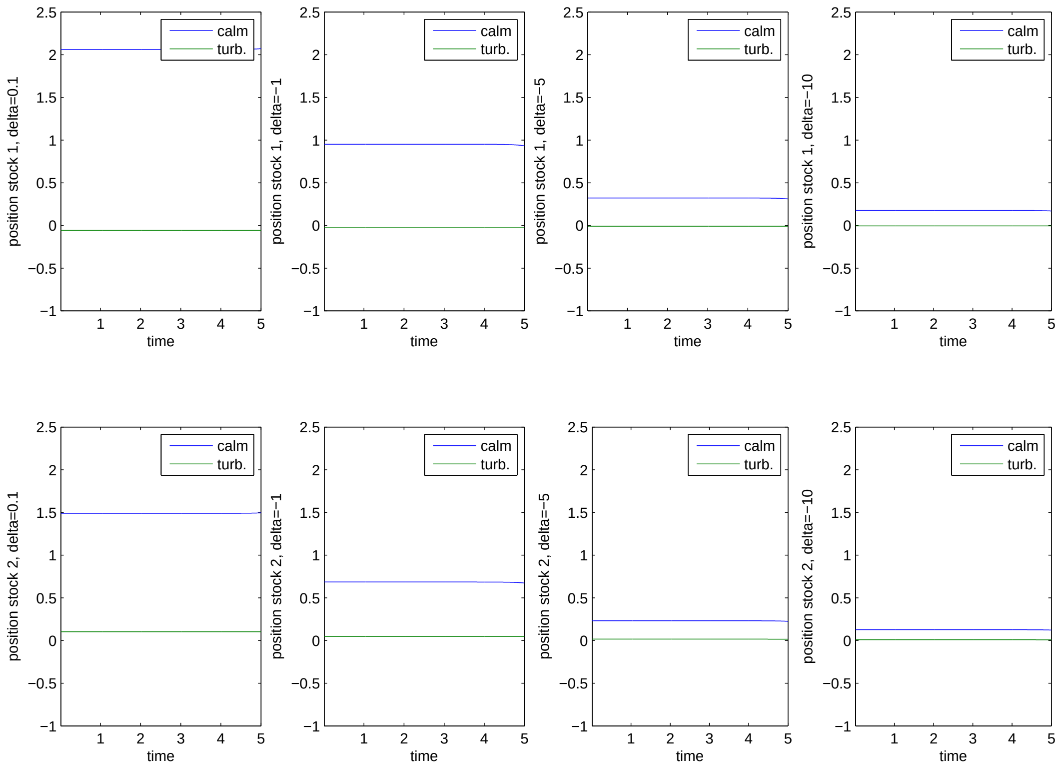


Figure 5.3: Example 2: Fraction of wealth invested in the high-risk stock P_1 (upper row) and in the less risky stock P_2 (low raw) for an investor with time horizon $T = 5$ and different risk parameters δ. All other parameters are as in Set 1.

[3]Note that the increase in the asset variances is achieved by an increase in matrix A, whereas the decrease in the returns is modeled by lower values for θ_1 and θ_2 for the second state.

In the calm state the investor holds high long positions in both assets, whereas in the turbulent state she has almost no investment in the risky assets. This is explained not only by the lower return and higher variance but also by the increased correlation in a crisis. To illustrate this we consider a second example wherein only the instantaneous correlation changes between the states, whereas $\bar{m}_1$, $\bar{m}_2$, $\bar{s}_1$ and $\bar{s}_2$ are the same for all states. We allow for three states: e_1, e_2^* and e_3^*. The parameters for state e_1 are the same as in Set 1. We set $\bar{\rho}(e_2^*) = 0$ and $\bar{\rho}(e_3^*) = 0.9$ so that in the second state the two assets are independent and in the third state they are highly correlated. This is achieved by the following parameter specifications $A(e_2^*) = \begin{pmatrix} 0.68 & 0 \\ 0 & 0.62 \end{pmatrix}$, $A(e_3^*) = \begin{pmatrix} 0.68 & 0 \\ 0.81 & 1.52 \end{pmatrix}$, $\theta_2(e_2^*) = 0.11$ and $\theta_2(e_3^*) = 0.003$. All other parameters remain the same as in state e_1. We call these parameter specifications Set 2. The optimal portfolio strategies are shown in Figure 5.4. It is

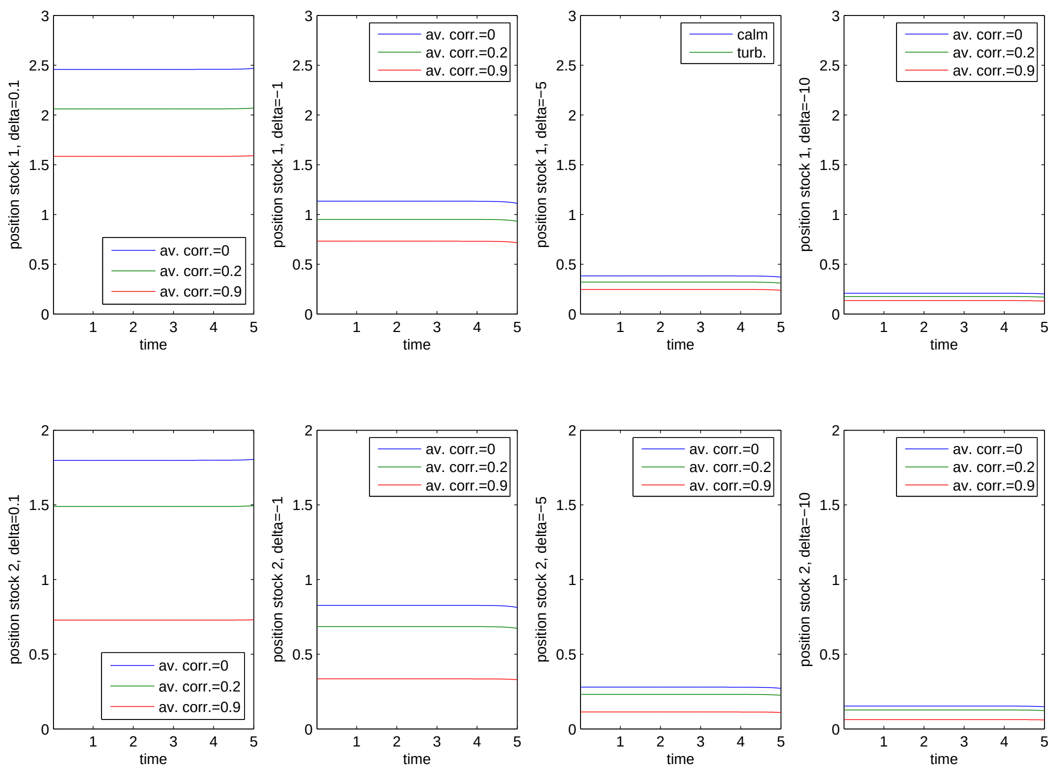


Figure 5.4: Example 2: Fraction of wealth invested in the high risk stock P_1 (upper row) and in the less risky stock P_2 (low raw) for an investor with time horizon $T = 5$ and different risk parameters δ. All other parameters are as in Set 2.

clearly recognizable that in situations with high asset correlation the investor holds less in both assets compared to the independent case. This can be interpreted as a protection against potential extremely high losses if both asset prices decrease.

The influence of δ is also clearly recognizable in Figure 5.3 and Figure 5.4: the smaller δ, thus the more risk-averse the investor, the smaller the absolute value of her positions in the stocks. The effect of the risk aversion is also reflected in the distribution of the terminal wealth illustrated in Figure 5.5: for lower values of δ, its

standard deviation gets smaller, so that both high losses and high gains have lower probability.
Finally, we address the question of the importance of the hedging term. Figure 5.6

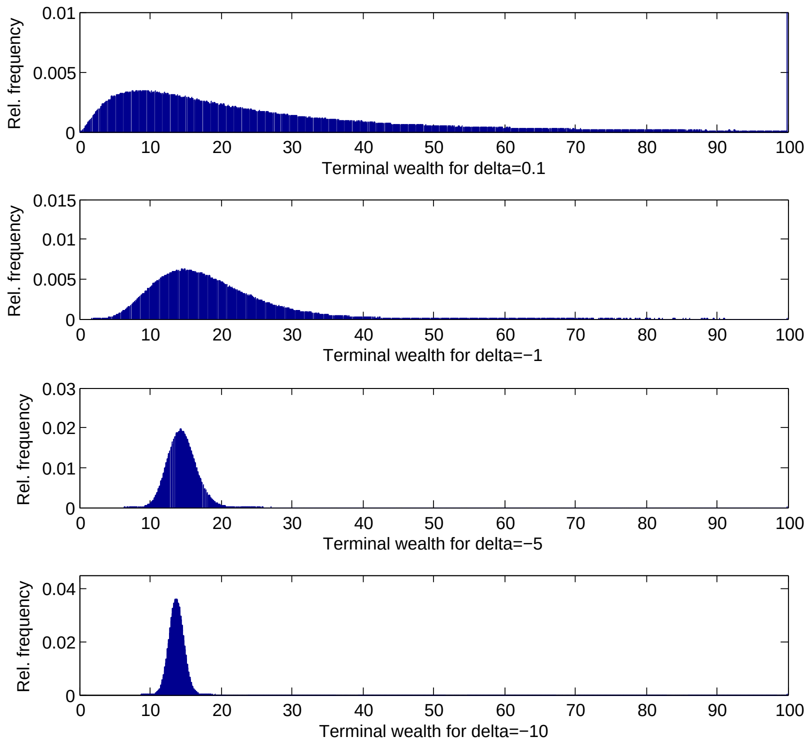


Figure 5.5: Example 2: Distribution of the terminal wealth obtained using the derived optimal strategy (5.50) with time horizon $T = 5$ and different risk parameters δ. All other parameters are as in Set 1. For better comparability all values higher than 100 are summarized in the last bar.

shows the relation between the hedging portfolio and the mean-variance position in stock P_2 as a function of the different parameters influencing the stochastic factor X_2[4]. It can be seen that slow mean-reversion and high correlation with the asset price increase the weight of the hedging term as in this case the stochastic factor can be better hedged using the traded asset. Furthermore, as the risk coming from the stochastic factor increases when its volatility parameter increases, the bigger χ, the higher the ratio $|\frac{\bar{\pi}_2^H}{\bar{\pi}_2^{MV}}|$. We obtained similar results in Section 4.3.4 for the one-dimensional Markov-modulated Heston model.
To summarize, depending on the risk preferences of the investor and the parameter values of the stochastic factors the hedging term may play an important role in the optimal portfolio. Thus, if an investor neglects the randomness of X_1 and X_2

[4]Note that X_2 and W_2^P influence only P_2, whereas X_1 and W_1^P appear in both assets. So, the position in P_2 is chosen in such a way to assure optimal exposure of the whole portfolio to X_2 and W_2^P. On the other side the position in P_2 contains additional exposure also to X_1 and W_1^P. So, the investment in P_1 is adjusted to compensate for this and thus, reflects also the interaction effects between the assets. We illustrate the importance of $\bar{\pi}^H$ using the second stock as it contains the hedging effect separately from the interaction between the assets.

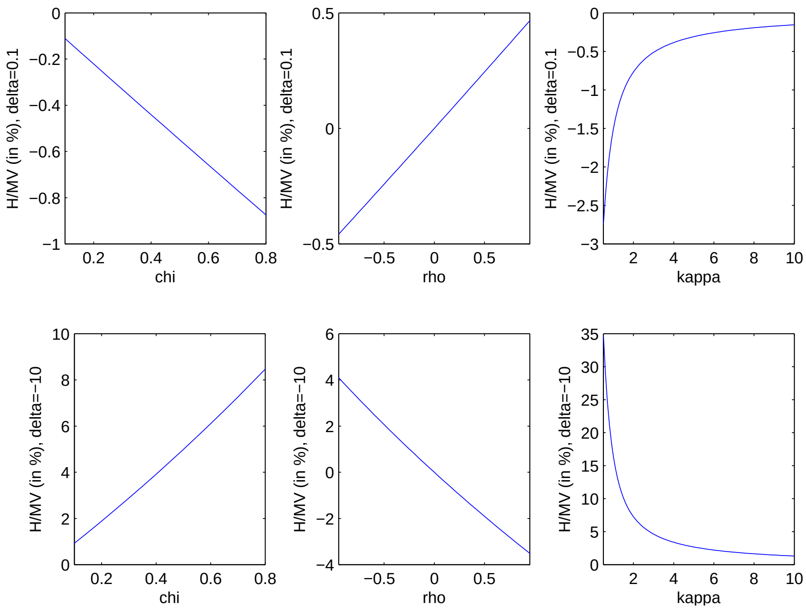


Figure 5.6: Example 2: Ratio (in %) between the hedging term and the mean-variance portfolio invested in stock P_2 for an investor with time horizon $T = 5$ and $\delta = 0.1$ (upper row), resp. $\delta = -10$ (lower row) as a function of χ_2 (first column), ρ_2 (second column), κ_2 (third column). All other parameters are as in Set 1.

and invests only in $\bar{\pi}^{MV}$, significant misallocations in the order of a double-digits percentage may arise.

To conclude the numerical analysis, considering the interaction between the two assets, as well as recognizing the additional stochastic factors and the Markov-switching character of the parameters is crucial for an optimal investment. This supports the relevance of the developed results.

Chapter 6

Extension to the HARA utility function

In this chapter we extend our study from the power utility function to the more general case of the HARA utility function U_H as defined by (2.26). We specify one of the stochastic factors to be the short rate and propose a consistent bond-stock framework with bond-stock correlation. For utility optimization results in a bond-stock market without Markov switching see [21]. There the interest rate is driven by two stochastic factors and the stock price process has constant excess return and volatility. In what follows we use for the interest rate a Markov-modulated Vasicek model and extend the optimization results from [21] to Markov switching and additional stochastic factors such as stochastic volatility and excess return.

6.1 Model and optimization problem

Let us first introduce formally the model we are working with, which is a special case of Model (3.1). As in the general definition we consider N traded risky assets, influenced by a Markov chain $\mathcal{MC}$ and J further sources of randomness modeled by the stochastic process X. Similarly to Section 5.3.1, in this chapter we identify process X_1 as the driver of the short rate $r = r\big(X_1\big)$ and P_1 as the price of a bond with fixed maturity $T_1 > T$. So, additionally to the bank account we also allow for trading in a bond.
We start by specifying the dynamics of X_1 and deriving the SDE for the bond price. Let X_1 be characterized by the following SDE under the risk-neutral measure $\mathcal{Q}$:

$$\mathrm{d}X_1(t) = \tilde{\mu}_1^X\big(X_1(t)\big)\mathrm{d}t + \Sigma_{11}^X\big(X_1(t)\big)\mathrm{d}\tilde{W}_1^X(t),$$

where $\tilde{W}_1^X$ is a one-dimensional $\mathcal{Q}$-Brownian motion. As in Section 5.3.1, the change from the real-world to the risk-neutral measure is given by the following density

$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathbb{P}}\Big|_{\mathscr{F}^{X_1}} = L(T)$ where:

$$L(t) = \exp\Big\{\int_0^t \lambda_1\big(\mathcal{MC}(s)\big)\mathrm{d}W_1^X(s) - \frac{1}{2}\int_0^t \parallel \lambda_1\big(\mathcal{MC}(s)\big) \parallel^2 \mathrm{d}s\Big\}.$$

Note that we allow for the market price of risk λ_1 to depend on the Markov chain. As Novikov's condition is trivially fulfilled for L we can apply Girsanov's Theorem 2.40 to show that process W_1^X defined by:

$$W_1^X(t) = \tilde{W}_1^X(t) + \int_0^t \lambda_1\big(\mathcal{MC}(s)\big)\mathrm{d}s$$

is a $\mathbb{P}$-Brownian motion. Thus, under $\mathbb{P}$ we have the following SDE for X_1:

$$\mathrm{d}X_1(t) = \mu_1^X\big(X_1(t), \mathcal{MC}(t)\big)\mathrm{d}t + \Sigma_{11}^X\big(X_1(t)\big)\mathrm{d}W_1^X(t),$$

where

$$\mu_1^X\big(X_1(t), \mathcal{MC}(t)\big) = \tilde{\mu}_1^X\big(X_1(t)\big) - \Sigma_{11}^X\big(X_1(t)\big)\lambda_1\big(\mathcal{MC}(s)\big). \tag{6.1}$$

The price at time t of a zero-coupon bond with maturity T_1 is then given by:

$$P_1(t, T_1, x_1) = \mathbb{E}_{\mathcal{Q}}\Big[\exp\Big\{-\int_t^{T_1} r\big(X_1(s)\big)\mathrm{d}s\Big\}\Big|X_1(t) = x_1\Big].$$

By the Feynman-Kac theorem, for a fixed T_1 function P_1 satisfies the following PDE:

$$(P_1)_t + (P_1)_{x_1}\tilde{\mu}_1^X + \frac{1}{2}\big(\Sigma_{11}^X\big)^2(P_1)_{x_1x_1} - P_1 r = 0, P_1(T_1, T_1, x_1) = 1. \tag{6.2}$$

As in Section 5.3.1 we adopt a Vasicek model for the short rate: $r(X_1) = X_1$ with

$$\Sigma_{11}^X(x_1) = \chi_1 \tag{6.3}$$

$$\tilde{\mu}_1^X(x_1) = \kappa_1(\tilde{\theta}_1 - x_1), \tag{6.4}$$

thus,

$$\mu_1^X(x_1, e_i) = \kappa_1(\tilde{\theta}_1 - x_1) - \chi_1\lambda_1(e_i) =: \kappa_1(\theta_1(e_i) - x_1). \tag{6.5}$$

Analogously to Equation (5.33), the bond price is known in an explicit form:

$$P_1(t, T_1, x_1) = \exp\big\{-A_1(T_1 - t) - A_2(T_1 - t)x_1\big\}, \tag{6.6}$$

where for $\tau \in [0, T_1]$ functions $A_1(\tau)$ and $A_2(\tau)$ are given by:

$$A_1(\tau) = \big(\tilde{\theta}_1 - \frac{\chi_1^2}{2\kappa_1^2}\big)\big(\tau - A_2(\tau)\big) + \frac{\chi_1^2}{4\kappa_1}A_2^2(\tau) \tag{6.7}$$

$$A_2(\tau) = \frac{1}{\kappa_1}\big(1 - \exp\{-\kappa_1\tau\}\big). \tag{6.8}$$

Remark 6.1
If we assume the Vasicek model for X_1 and an affine function for $r(x_1)$:

$$r(x_1) = \varepsilon^{(0)} + \bar{\varepsilon}_1^{(1)} x_1,$$

we obtain the following SDE for the short rate:

$$\mathrm{d}r(X_1(t)) = \kappa_1(\underbrace{\varepsilon^{(0)} + \bar{\varepsilon}_1^{(1)}\tilde{\theta}_1}_{=:\hat{\theta}_1} - \underbrace{(\varepsilon^{(0)} + \bar{\varepsilon}_1^{(1)} X_1(t))}_{=r(X_1(t))}\,\mathrm{d}t + \underbrace{\bar{\varepsilon}_1^{(1)}\chi_1}_{=:\hat{\chi}_1}\,\mathrm{d}\tilde{W}_1^X(t).$$

So the affine transformation results again in a Vasicek process. As it does not increase the flexibility of the model, we define $r(x_1) = x_1$ in order to avoid unnecessary complication of the notation.

From now on we use the following shorter notation whenever referring to a bond with maturity T_1: $P_1(t) := P_1(t, T_1, X_1(t))$. Now we derive the $\mathbb{P}$-dynamics of the bond price:

$$\begin{aligned}
\mathrm{d}P_1(t) =& \big[(P_1)_t\big(t, T_1, X_1(t)\big) + (P_1)_{x_1}\big(t, T_1, X_1(t)\big)\mu_X\big(X_1(t), \mathcal{MC}(t)\big) \\
&+ \frac{1}{2}\Sigma_{11}^X\Sigma_{11}^X (P_1)_{x_1x_1}\big(t, T_1, X_1(t)\big)\big]\mathrm{d}t + (P_1)_{x_1}\big(t, T_1, X_1(t)\big)\Sigma_{11}^X \mathrm{d}W_1^X \\
=& \big[(P_1)_t\big(t, T_1, X_1(t)\big) + (P_1)_{x_1}\big(t, T_1, X_1(t)\big)\tilde{\mu}_X\big(X_1(t), \mathcal{MC}(t)\big) \\
&- (P_1)_{x_1}\big(t, T_1, X_1(t)\big)\Sigma_{11}^X\lambda_1\big(\mathcal{MC}(t)\big) + \frac{1}{2}\Sigma_{11}^X\Sigma_{11}^X (P_1)_{x_1x_1}\big(t, T_1, X_1(t)\big)\big]\mathrm{d}t \\
&+ (P_1)_{x_1}\big(t, T_1, X_1(t)\big)\Sigma_{11}^X \mathrm{d}W_1^X \\
\overset{(6.2)}{=}& \big[P_1\big(t, T_1, X_1(t)\big)r\big(X_1(t)\big) - (P_1)_{x_1}\big(t, T_1, X_1(t)\big)\Sigma_{11}^X\lambda_1\big(\mathcal{MC}(t)\big)\big]\mathrm{d}t \\
&+ (P_1)_{x_1}\big(t, T_1, X_1(t)\big)\Sigma_{11}^X \mathrm{d}W_1^X \\
=& \big[P_1(t)r\big(X_1(t)\big) + P_1(t)A_2(T_1 - t)\Sigma_{11}^X\lambda_1\big(\mathcal{MC}(t)\big)\big]\mathrm{d}t \\
&- P_1(t)A_2(T_1 - t)\Sigma_{11}^X \mathrm{d}W_1^X \\
=& P_1(t)\Big[\underbrace{\big\{ r\big(X_1(t)\big) + A_2(T_1 - t)\Sigma_{11}^X\lambda_1\big(\mathcal{MC}(t)\big)\big\}}_{=\mu_1(t, X_1(t), \mathcal{MC}(t))}\mathrm{d}t + \underbrace{A_2(T_1 - t)\Sigma_{11}^X}_{=\Sigma_{11}(t)}\mathrm{d}W_1^P\Big],
\end{aligned} \tag{6.9}$$

where $W_1^P := -W_1^X$ and consequently it is a $\mathbb{P}$-Brownian motion.
As in Model (3.1), for the remaining $N-1$ risky assets we assume an affine exponential structure, where the drift and the volatility terms are influenced both by the Markov chain and the stochastic factors. We assume that the remaining stochastic factors are driven by $J-1$ Brownian motions $(W_2^X, \ldots, W_J^X)$, which are independent of W_1^X.
In the next definition we summarize the assumptions so far and state the considered model explicitly in terms of the notation from Model (3.1).

Definition 6.2 (Affine bond-stock model with Markov switching)

We consider the following model:

$$\begin{aligned}
&\mathrm{d}P_0(t) = P_0(t)r\big(X_1(t)\big)\mathrm{d}t \ \textit{(bank account)}\\
&\mathrm{d}P_1(t) = P_1(t)\Big[\mu_1\big(t, X_1(t), \mathcal{MC}(t)\big)\mathrm{d}t + \Sigma_1(t)\mathrm{d}W^P(t)\Big] \ \textit{(bond)}\\
&\mathrm{d}P_n(t) = P_n(t)\Big[\mu_n\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}t + \Sigma_n\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}W^P(t)\Big], n = 2, \dots, N\\
&\mathrm{d}X_1(t) = \mu_1^X\big(X_1(t), \mathcal{MC}(t)\big)\mathrm{d}t + \Sigma_1^X \mathrm{d}W^X(t) \ \textit{(short rate)}\\
&\mathrm{d}X_j(t) = \mu_j^X\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}t + \Sigma_j^X\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}W^X(t), j = 2, \dots, J\\
&\mathrm{d}\langle W^X, W^P\rangle(t) = \rho\big(X(t), \mathcal{MC}(t)\big)\ dt, \ \textit{i.e.}\ \mathrm{d}\langle W_j^X, W_n^P\rangle(t) = \rho_{jn}\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}t,
\end{aligned} \tag{6.10}$$

where $W^P = (W_1^P, \dots, W_N^P) \in \mathbb{R}^N$ *and* $W^X = (W_1^X, \dots, W_J^X) \in \mathbb{R}^J$ *are standard Brownian motions,* $\Sigma_n(x, e_i) : D^X \times \mathcal{E} \to \mathbb{R}^{1,N}$, $\mu_n(x, e_i) : D^X \times \mathcal{E} \to \mathbb{R}$, *for* $n = 2, \dots, N$, $\mu_j^X(x, e_i) : D^X \times \mathcal{E} \to \mathbb{R}$, $\Sigma_j^X(x, e_i) : D^X \times \mathcal{E} \to \mathbb{R}^J$, *for* $j = 2, \dots, J$, $\rho(x, e_i) : D^X \times \mathcal{E} \to \mathbb{R}^{J,N}$, $\rho_{jn}\big(X(t), \mathcal{MC}(t)\big)\mathrm{d}t = \mathrm{d}\langle W_j^X, W_n^P\rangle(t)$ *and the following holds:*

$$\Sigma_1(t) = (\Sigma_{11}(t), 0, \dots, 0) : [0, T] \mapsto \mathbb{R}^N \tag{6.11}$$

$$\Sigma_1^X = (\Sigma_{11}^X, 0, \dots, 0) \in \mathbb{R}^J \tag{6.12}$$

$$\rho_{11} = -1, \rho_{1n} = \rho_{j1} = 0, \ \textit{for}\ n = 2, \dots, N, j = 2, \dots, J. \tag{6.13}$$

Recall that Σ_{11} *and* μ_1 *are defined in (6.9) and* Σ_{11}^X *and* μ_1^X *are given by (6.3) and (6.5), respectively. To ease the exposition, we denote:*

$$\begin{aligned}
\mu(t, x, e_i) &:= (\mu_1(t, x, e_i), \mu_2(x, e_i), \dots, \mu_N(x, e_i))' \in \mathbb{R}^N\\
\Sigma(t, x, e_i) &:= (\Sigma_1(t), \Sigma_2(x, e_i), \dots, \Sigma_N(x, e_i))' \in \mathbb{R}^{N,N}\\
\mu^X(x, e_i) &:= (\mu_1^X(x, e_i), \mu_2^X(x, e_i), \dots, \mu_J^X(x, e_i))' \in \mathbb{R}^J\\
\Sigma^X(x, e_i) &:= (\Sigma_1^X, \Sigma_2^X(x, e_i), \dots, \Sigma_J^X(x, e_i))' \in \mathbb{R}^{J,J}.
\end{aligned}$$

We assume that Σ *is a.s. invertible and that the following conditions hold:*

$$r = x_1 \tag{6.14}$$

$$\mu^X = \bar{k}^{(0)}(e_i) - K^{(1)}x \tag{6.15}$$

$$\Sigma^X(\Sigma^X)' = H^{(0)}(e_i) + \sum_{j=1}^J H^{(1j)}x_j \tag{6.16}$$

$$(\mu - r)'\big(\Sigma\Sigma'\big)^{-1}(\mu - r) = h^{(0)}(e_i) + (\bar{h}^{(1)})'x \tag{6.17}$$

$$\Sigma^X\rho\Sigma^{-1}(\mu - r) = \bar{g}^{(0)}(e_i) + G^{(1)}x \tag{6.18}$$

$$\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)' = L^{(0)}(e_i) + \sum_{j=1}^J L^{(1j)}x_j, \tag{6.19}$$

where $\bar{k}^{(0)}, \bar{g}^{(0)} : \mathcal{E} \mapsto \mathbb{R}^J$, $h^{(0)} : \mathcal{E} \mapsto \mathbb{R}$, $H^{(0)}, L^{(0)} : \mathcal{E} \mapsto \mathbb{R}^{J,J}$, $\bar{\varepsilon}^{(1)}, \bar{h}^{(1)} \in \mathbb{R}^J$, $K^{(1)}, H^{(1j)}, G^{(1)}, L^{(1j)} \in \mathbb{R}^{J,J}$, *for all* $j = 1, \dots, J$. *Recall that* $\bar{k}_1^{(0)} = \kappa_1\theta_1(e_i)$,

$K_{11}^{(1)} = \kappa_1$, $K_{1j}^{(1)} = 0$, *for* $j = 2, \ldots, J$, $H_{11}^{(0)} = \chi_1^2$, $H_{11}^{(1j)} = 0$, *for* $j = 1, \ldots, J$.
Note that the above conditions are a special case of Conditions (3.2)-(3.7) and correspond to the assumptions made in Section 5.2.2. Furthermore, observe that this framework allows for instantaneous correlation between the bond and the stock as the Brownian motion of the bond W_1^P *might influence the dynamics of all assets.*

Remark 6.3
If $\rho = 0$ *one can relax Conditions (6.15)-(6.19) by allowing for* $K^{(1)}$, $H^{(1j)}$, $\bar{h}^{(1)}$, $G^{(1)}$, *and* $L^{(1j)}$ *to depend on the Markov chain.*

For the risk preferences of the investor we assume a HARA utility function U_H as defined by (2.26) and as discussed in Section 3.1, the set of all admissible trading strategies is given by:

$$\Lambda(t, v) := \Big\{\pi \Big| \pi(s) \in \mathbb{R}^N, V^\pi(t) = v, V^\pi(T) \geq F\Big\}.$$

Observe that $V^\pi(T) \geq F$ implies that $V^\pi(t) \geq F P_1(t, T, X_1(t))$ for all $t \in [0, T]$, as otherwise the investor would not have enough capital to buy the necessary amount of bonds to assure that her terminal wealth is higher than F. This is more precisely explained by the fact that the stock price and the value of the bank account at maturity are not bounded from below. Reversely, if $V^\pi(t) \geq F P_1(t, T, X_1(t))$, then investing the whole wealth in bonds with maturity T assures that the terminal wealth is higher or equal to F. Thus,

$$D^\Lambda = \big\{(t, v, x, e_i) \in [0, T] \times [0, \infty) \times D^X \times \mathcal{E} \big| v \geq F P_1(t, T, x)\big\}. \tag{6.20}$$

We consider the optimization problem from (3.12) for the HARA utility function:

$$\begin{aligned} J^{(t,v,x,e_i)}(\pi) &:= \mathbb{E}_{\mathbb{Q}}\Big[U_H(V^\pi(T)) | V^\pi(t) = v, X(t) = x, \mathcal{MC}(t) = e_i\Big] \\ \Phi(t, v, x, e_i) &:= \max_{\pi \in \Lambda(t,v)} J^{(t,v,x,e_i)}(\pi). \end{aligned} \tag{6.21}$$

6.2 General optimization results

As in Chapters 4 and 5 we already discussed in detail the two simpler frameworks: with time-dependent deterministic coefficients and with $\rho = 0$, here we omit the detailed results for these two cases and proceed directly with the more relevant case $\rho \neq 0$. Recall from Equation (3.16) the corresponding HJB equation for all $(t, v, x, e_i) \in D^\Lambda$:

$$\begin{aligned} \max_{\pi \in \Lambda(t,v)} \{\mathcal{L}(e_i, \pi)\Phi(t, v, x, e_i)\} &= -\sum_{z=1}^{I} q_{i,z} \Phi(t, v, x, e_z) \\ \Phi(T, v, x, e_i) &= U_H(v), \forall i \in \{1, \ldots, I\}, \end{aligned} \tag{6.22}$$

where the differential operator $\mathcal{L}(e_i, \pi)$ is given for each $e_i \in \mathcal{E}$ as follows:

$$\begin{aligned}\mathcal{L}(e_i, \pi)\Phi :=& \Phi_t + \mu^V \Phi_v + (\mu^X)'\Phi_x + \frac{1}{2}\Sigma^V(\Sigma^V)'\Phi_{vv} + \frac{1}{2}Tr\big(\Sigma^X(\Sigma^X)'\Phi_{xx'}\big)\\ &+ \Sigma^V \rho'(\Sigma^X)'\Phi_{vx}\Big|_{(t,v,x,e_i,\pi)}.\end{aligned}$$

As in Equation (3.18) the candidate for the optimal portfolio is given by:

$$\bar{\pi}(t) = -\frac{1}{V^{\bar{\pi}}(t)\Phi_{vv}}\Big\{\Phi_v\big(\Sigma\Sigma'\big)^{-1}(\mu - r) + (\Sigma')^{-1}\rho'(\Sigma^X)'\Phi_{vx}\Big\}\Big|_{(t,V^{\bar{\pi}}(t),X(t),\mathcal{MC}(t))}. \tag{6.23}$$

According to the form of the utility function we propose the following ansatz for the value function:

$$\Phi(t, v, x, e_i) = \frac{1-\delta}{\delta}\alpha\Big\{\frac{1}{1-\delta}\big(v - Fd(t, x_1)\big)\Big\}^\delta f(t, x, e_i), \tag{6.24}$$

where $d : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ and $f : [0, T] \times \mathbb{R}^J \times \mathcal{E} \mapsto \mathbb{R}$ are real-valued functions. Together with (6.23) this ansatz leads to the following expression for the candidate optimal strategy $\bar{\pi}$:

$$\begin{aligned}\bar{\pi}(t) =& \frac{1}{1-\delta}\frac{V^{\bar{\pi}}(t) - Fd}{V^{\bar{\pi}}(t)}\Big\{(\Sigma\Sigma')^{-1}(\mu - r) + (\Sigma')^{-1}\rho'(\Sigma^X)'\frac{f_x}{f}\Big\}\\ &+ \frac{Fd}{V^{\bar{\pi}}(t)}(\Sigma')^{-1}\rho'(\Sigma^X)'\frac{d_x}{d}\Big|_{(t,V^{\bar{\pi}}(t),X(t),\mathcal{MC}(t))}.\end{aligned} \tag{6.25}$$

Substitution of (6.23) and (6.24) in (6.22) leads to the following system of PDEs for f and d:

$$\begin{aligned}&v^2\Big\{\frac{L_1}{1-\delta} + rf\Big\} + vF\Big\{ - fL_2 + f_x'\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x - \frac{2}{1-\delta}dL_1\\ &- drf\Big\} + F^2 dfL_2 - F^2 df_x'\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x + \frac{1}{1-\delta}F^2d^2L_1\\ &+ \frac{1-\delta}{2}F^2 fd_x'\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x\Big|_{(t,x,e_i)} = 0, \forall i = 1, \dots, I,\end{aligned} \tag{6.26}$$

where

$$\begin{aligned}L_1 :=& \frac{1-\delta}{\delta}f_t + \frac{1-\delta}{\delta}(\mu^X)'f_x + \frac{1}{2}\frac{1-\delta}{\delta}Tr\big(\Sigma^X(\Sigma^X)'f_{xx'}\big)\\ &+ \frac{1}{2}(\mu - r)'\big(\Sigma\Sigma'\big)^{-1}(\mu - r)f + f_x'\Sigma^X\rho\Sigma^{-1}(\mu - r) + \frac{1}{2}\frac{f_x'}{f}\Sigma^X\rho\rho'(\Sigma^X)'f_x\\ &+ \frac{1-\delta}{\delta}\sum_{z=1}^{I} q_{i,z}f(t, x, e_z)\\ L_2 :=& d_t + (\mu^X)'d_x + \frac{1}{2}Tr\big(\Sigma^X(\Sigma^X)'d_{xx'}\big) - d_x'\Sigma^X\rho\Sigma^{-1}(\mu - r).\end{aligned}$$

The terminal conditions read as follows:

$$\begin{aligned} f(T,x,e_i) &= 1 \\ d(T,x_1) &= 1. \end{aligned}$$

The explicit solutions for f and d and consequently the solution to the HJB equation are stated in the following theorem. We also state the solution in the corresponding time-dependent model, as we will need it later on.

Theorem 6.4 (Solution with the HARA utility function)
Assume that the following system of ODEs:

$$\begin{aligned} &\frac{\partial}{\partial t}B_j(t) + \delta\bar{\varepsilon}_j^{(1)} + \frac{1}{2}\frac{\delta}{1-\delta}\bar{h}_j^{(1)} - \sum_{d=1}^{J} B_d(t)K_{jd}^{(1)} + \frac{\delta}{1-\delta}\sum_{d=1}^{J} B_d(t)G_{jd}^{(1)} \\ &+ \frac{1}{2}\frac{\delta}{1-\delta}B(t)'\big(L^{(1j)} + H^{(1j)}\big)B(t) + \frac{1}{2}B(t)'H^{(1j)}B(t) = 0 \\ &B_j(T) = 0, \forall j \in \{1,\dots,J\}, \end{aligned} \tag{6.27}$$

admits a solution $B(t) = \big(B_1(t),\dots,B_J(t)\big)'$. *Furthermore, define:*

$$\xi(t,e_i) = \mathbb{E}\Big[\exp\Big\{\int_t^T w\big(s,\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\big|\mathcal{MC}(t) = e_i\Big], \tag{6.28}$$

with

$$\begin{aligned} w(t,e_i) =& \frac{1}{2}\frac{\delta}{1-\delta}h^{(0)}(e_i) + B(t)'\bar{k}^{(0)}(e_i) + \frac{\delta}{1-\delta}B(t)'\bar{g}^{(0)}(e_i) \\ &+ \frac{1}{2}\frac{\delta}{1-\delta}B(t)'\big(L^{(0)}(e_i) + H^{(0)}(e_i)\big)B(t) + \frac{1}{2}B(t)'H^{(0)}(e_i)B(t). \end{aligned} \tag{6.29}$$

Then functions f and d given by:

$$f(t,x,e_i) = \xi(t,e_i)\exp\big\{B(t)'x\big\}, \tag{6.30}$$

$$d(t,x_1) = P_1(t,T,x_1) = \exp\big\{-A_1(T-t) - A_2(T-t)x_1\big\} \tag{6.31}$$

satisfy System (6.26). Function Φ given by:

$$\Phi(t,v,x,e_i) = \frac{1-\delta}{\delta}\alpha\Big\{\frac{1}{1-\delta}\big(v - Fd(t,x_1)\big)\Big\}^{\delta} f(t,x,e_i) \tag{6.32}$$

solves Equation (6.22) with $\pi = \bar{\pi}$ given by:

$$\begin{aligned} \bar{\pi}(t) =& \frac{1}{1-\delta}\frac{v-Fd}{v}\Big\{(\Sigma\Sigma')^{-1}(\mu - r) + (\Sigma')^{-1}\rho'(\Sigma^X)'(B_1,B_2,\dots,B_J)'\Big\} \\ &+ \frac{Fd}{v}(\Sigma')^{-1}\rho'(\Sigma^X)'(\frac{d_{x_1}}{d},0,\dots,0)'\Big|_{(t,V^{\bar{\pi}}(t),X(t),\mathcal{MC}(t))}. \end{aligned} \tag{6.33}$$

The HJB solution for the corresponding time-dependent model is given by:

$$\Phi^m = \frac{1-\delta}{\delta}\alpha\Big\{\frac{1}{1-\delta}\big(v - Fd(t,x_1)\big)\Big\}^\delta \exp\Big\{\int_t^T w\big(s,m(s)\big)\mathrm{d}s\Big\}\exp\big\{B(t)'x\big\}, \tag{6.34}$$

and

$$\begin{aligned}\bar{\pi}^m =& \frac{1}{1-\delta}\frac{v-Fd}{v}\Big\{(\Sigma\Sigma')^{-1}(\mu - r) + (\Sigma')^{-1}\rho'(\Sigma^X)'(B_1, B_2, \ldots, B_J)'\Big\} \\ &+ \frac{Fd}{v}(\Sigma')^{-1}\rho'(\Sigma^X)'(\frac{d_{x_1}}{d}, 0, \ldots, 0)'\Big|_{(t,V^{m,\bar{\pi}^m}(t),X^m(t),m(t))}.\end{aligned} \tag{6.35}$$

Proof

We are interested in solving Equation (6.26) in order to find the HJB solution. By comparison of coefficients the terms in front of the powers of v should be zero, which yields the following system:

$$v^2 : \frac{L_1}{1-\delta} + rf = 0 \Leftrightarrow L_1 = -(1-\delta)rf \tag{6.36}$$

$$\begin{aligned}v^1 :& -fL_2 + f'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x - \frac{2}{1-\delta}dL_1 - drf = 0 \\ &\overset{(6.36)}{\Leftrightarrow} -fL_2 + f'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x + \frac{2}{1-\delta}d(1-\delta)rf - drf = 0 \\ &\Leftrightarrow L_2 = \frac{f'_x}{f}\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x + dr\end{aligned} \tag{6.37}$$

$$\begin{aligned}v^0 :& F^2dfL_2 - F^2df'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x + \frac{1}{1-\delta}F^2d^2L_1 \\ &+ \frac{1-\delta}{2}F^2fd'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x = 0 \\ &\overset{6.36}{\underset{6.37}{\Leftrightarrow}} F^2df\left(\frac{f'_x}{f}\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x + dr\right) \\ &- F^2df'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x - \frac{1}{1-\delta}F^2d^2(1-\delta)rf \\ &+ \frac{1-\delta}{2}F^2fd'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x = 0 \\ &\Leftrightarrow F^2df'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x + F^2d^2fr - F^2df'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' \\ &- \Sigma^X(\Sigma^X)'\big)d_x - F^2d^2fr + \frac{1-\delta}{2}F^2fd'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x \\ &\Leftrightarrow fd'_x\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x = 0.\end{aligned} \tag{6.38}$$

Note that $d_x = (d_{x_1}, 0, \ldots, 0)'$ and recall Equations (6.12)and (6.13) from Definition 6.2 to conclude that:

$$\big(\Sigma^X\rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)'\big)d_x = 0.$$

Thus, Equation (6.38) is trivially fulfilled. Considering additionally Equation (6.11) and the fact that function d depends only on x_1, Equation (6.37) simplifies to:

$$d_t + \big(\mu_1^X - \Sigma_{11}^X \rho_{11} \frac{\mu_1 - r}{\Sigma_{11}}\big) d_{x_1} + \frac{1}{2} (\Sigma_{11}^X)^2 d_{x_1 x_1} - dr = 0, \tag{6.39}$$

with terminal condition $d(T, x_1) = 1$. By recalling the definitions of μ_1 and Σ_{11} from Equation (6.9), as well as the fact that $\rho_{11} = -1$ and the definition of $\tilde{\mu}_1^X$ from Equation (6.1), we can follow that:

$$\mu_1^X - \Sigma_{11}^X \rho_{11} \frac{\mu_1 - r}{\Sigma_{11}} = \tilde{\mu}_1^X.$$

Inserting this into (6.39) and comparing with (6.2) we see that d corresponds to the price of a zero-coupon bond with maturity T and analogously to Equation (6.6) is given by:

$$d(t, x_1) = P_1(t, T, x_1) = \exp\big\{ - A_1(T - t) - A_2(T - t) x_1 \big\}. \tag{6.40}$$

Now let us continue with Equation (6.36) and the calculation of its solution f. Note that up to a constant it is the same as Equation (5.3), so from Theorem 5.10 we know its solution and it is exactly as given by the theorem. The solution for the corresponding time-dependent model follows analogously. □

Before we state a verification result, we first check the admissibility of the candidate $\bar{\pi}$ for the optimal portfolio by showing in the next theorem that $V^{\bar{\pi}}(t) \geq Fd(t, X_1(t))$. The proof can be found in Appendix D.

Theorem 6.5 (Lower Bound)
Let f and d be given as in Theorem 6.4 and $\bar{\pi}$ be the strategy given by Equation (6.33). Assume the following lower bound for the initial wealth of the investor:

$$V^{\bar{\pi}}(0) \geq Fd(0, X_1(0)).$$

Then it holds:

$$V^{\bar{\pi}}(t) \geq Fd(t, X_1(t)), \forall t \in (0, T].$$

Now we prove a verification theorem that gives a set of sufficient conditions for the derived function Φ and portfolio $\bar{\pi}$ to be the solution to the considered optimization problem. It is analogous to Theorem 5.1. Although the proof follows the same steps as Theorem 5.1, we provide it in Appendix D as there are some complications due to the more general utility function.

Theorem 6.6 (Verification via a martingale condition)
Consider function d as given by (6.31) and let function $\Phi : D^\Lambda \to \mathbb{R}$, where D^Λ is defined by (6.20) fulfill the following conditions:

i) For each $e_i \in \mathcal{E}$, Φ is once continuously differentiable in t and twice continuously differentiable in v and x.

ii) Φ satisfies System (6.22) with $\pi = \bar{\pi}$ given as in (6.33).

iii) $\{\Phi(t, V^{\bar{\pi}}(t), X(t), \mathcal{MC}(t))\}_{t\in[0,T]}$ is a martingale.

Then $\mathbb{E}\big[U\big(V^{\bar{\pi}}(T)\big)\big|V^{\bar{\pi}}(t) = v, X(t) = x, \mathcal{MC}(t) = e_i\big] = \Phi(t, v, x, e_i)$. Further, if $\Phi(t, v, x, e_i) = \frac{1-\delta}{\delta}\alpha\Big\{\frac{1}{1-\delta}\big(v - Fd(t, x_1)\big)\Big\}^{\delta} f(t, x, e_i)$ with some positive function f then Φ is the value function and $\bar{\pi}$ is the optimal trading strategy to Problem (6.21).

Now, we combine Theorems 6.4, 6.6 and 3.5 in the following corollary.

Corollary 6.7
Adopt the notation from Theorem 6.4 and assume that the conditions required there are true. If at least one of the processes $\{\Phi^m(t, V^{m,\bar{\pi}^m}(t), X^m(t))\}_{t\in[0,T]}$ or $\{\Phi(t, V^{\bar{\pi}}(t), X(t), \mathcal{MC}(t))\}_{t\in[0,T]}$ is a martingale, then Φ is the value function and $\bar{\pi}$ the optimal strategy for Problem 6.21.

Proof
As in Proposition 5.3 one can show that the martingale property of Φ^m leads to the verification result for the time-dependent model. Then, Theorem 3.5 yields the verification result for the Markov-switching model as well. Furthermore, if Φ is a martingale, we can apply Theorem 6.6 to show that Φ is the value function and $\bar{\pi}$ the optimal trading strategy for the considered optimization problem.

□

As in Proposition 5.6, the martingale property of Φ^m can be shown under certain conditions by Theorem 2.34. We state this is the following proposition:

Proposition 6.8 (Verification via Theorem 2.34)
Consider $\Phi^m : D^\Lambda \to \mathbb{R}$ and $\bar{\pi}^m$ given by (6.34) and (6.35), respectively. Define process G by:

$$G(t) := \ln\left(\frac{\Phi^m(t, V^{m,\bar{\pi}^m}(t), X^m(t))}{\Phi^m(0, V^{m,\bar{\pi}^m}(0), X^m(0))}\right).$$

Now consider the $J+1$-dimensional process $Z := (\bar{X}^m, G)'$, where $\bar{X}^m$ is a suitable permutation of X^m. Then its differential semimartingale characteristics μ^Z, Γ^Z exhibit an affine structure as in Theorem 2.34. Assume that μ^Z, Γ^Z fulfill requirements i)-iv) from Theorem 2.34. Then Φ given by (6.32) is the value function to Problem (6.21) and the optimal portfolio is as given in (6.33).

Proof
Substitution of (6.34) and (6.35) in the definition of G leads to the same dynamics for process G as in Proposition 5.6. Thus, the rest of the proof is exactly the same as the proof of Proposition 5.6.

□

Now we would like to have a closer look at the optimal investment strategy (6.33) and understand better the role of its single parts. Note that it exhibits a structure similar to a CPPI strategy: the investment in the risky assets (up to the last term for the bond investment) is proportional to the difference between the current wealth and the discounted floor, which is in this case Fd. This difference is called the cushion. The CPPI strategy was introduced by [88] and [13]. For further details see e.g. [89] and [103]. As we have additional stochastic factors, the multiplier for the risky exposure is not constant but stochastic. This analogon is completed by Theorem 6.5 that proves that F is a lower bound for the terminal wealth. Thus it can be indeed interpreted as a floor.
The additional term in the bond position is a correction for the stochastic short rate and accounts for the fact that the floor cannot be replicated only by the bank account. To better explain this we introduce the following notation:

$$\bar{\pi}(t) = \hat{\pi}(t) + \tilde{\pi}(t) = \frac{V^{\bar{\pi}} - Fd}{V^{\bar{\pi}}}\dot{\pi}(t) + \frac{Fd}{V^{\bar{\pi}}}\ddot{\pi}(t),$$

where $\hat{\pi}$ contains the terms in (6.33) proportional to $\frac{V^{\bar{\pi}}-Fd}{V^{\bar{\pi}}}$ and $\tilde{\pi}$ contains the last term in (6.33), which is proportional to $\frac{Fd}{V^{\bar{\pi}}}$. The optimal investment for the power utility function is obtained by setting $F = 0$ and is denoted by $\dot{\pi}$. So, $\hat{\pi}$ corresponds to the optimal investment for the power utility function scaled by the factor $\frac{V^{\bar{\pi}}-Fd}{V^{\bar{\pi}}} \in (0, 1]$. Note that it exhibits the usual decomposition in a myopic part driven by the stocks excess return and volatility and a hedging part coming from the stochastic factor X. The second part $\tilde{\pi}$ contains an additional investment in the bond due to the stochastic interest rate. In the case of a deterministic interest rate or for $F = 0$ it would be zero. Let us rewrite the SDE for $V^{\bar{\pi}}$ with this new notation:

$$\begin{aligned} \mathrm{d}V^{\bar{\pi}} = & \underbrace{(V^{\bar{\pi}} - Fd)\Big[\big\{r + \dot{\pi}'(\mu - r)\big\}\mathrm{d}t + \dot{\pi}\Sigma \mathrm{d}W^P\Big]}_{=\mathrm{d}(V^{\bar{\pi}}-Fd)} \\ & + \underbrace{Fd\Big[\big\{r + \ddot{\pi}'(\mu - r)\big\}\mathrm{d}t + \ddot{\pi}\Sigma \mathrm{d}W^P\Big]}_{=\mathrm{d}(Fd)}. \end{aligned}$$

So, the optimal investment consists basically of two parts: the first one invests the cushion $V^{\bar{\pi}} - Fd$ according to the optimal portfolio for the power utility function and the second one replicates Fd by an investment in the bank account and the bond. Using the martingale approach for dynamic optimization [72] obtained similar results in a complete market (without additional stochastic factors and without Markov switching).

In the next two sections we apply the derived results to a concrete example and provide a more detailed analysis of the influence of the model parameters on the optimal strategy.

6.3 Example: Markov-modulated Heston model with stochastic interest rates

In what follows we consider an example with a bond and a stock, where the stock exhibits stochastic volatility and follows Markov-switching Heston-type dynamics. Models with similar features have been motivated in various empirical studies (see the references in Section 6.4). We apply the derived results from Section 6.2 to this bond-stock model, study their sensitivity to the model parameters and provide an economic interpretation. More precisely the model under consideration is given by:

$$\begin{aligned}
\mathrm{d}P_0(t) =& P_0(t)X_1(t)\mathrm{d}t \text{ (bank account)}\\
\mathrm{d}P_1(t) =& P_1(t)\big[X_1(t) + \lambda_1\big(\mathcal{MC}(t)\big)A_2(T_1-t)\chi_1\mathrm{d}t + A_2(T_1-t)\chi_1\mathrm{d}W_1^P(t)\big] \text{ (bond)}\\
\mathrm{d}P_2(t) =& P_2(t)\big[X_1(t) + \lambda_1\big(\mathcal{MC}(t)\big)a + \lambda_2\nu\big(\mathcal{MC}(t)\big)X_2\mathrm{d}t + a\mathrm{d}W_1^P(t)\\
&+ \nu\big(\mathcal{MC}(t)\big)\sqrt{X_2(t)}\mathrm{d}W_2^P(t)\big] \text{ (stock)}\\
\mathrm{d}X_1(t) =& \kappa_1\big(\theta_1\big(\mathcal{MC}(t)\big) - X_1(t)\big)\mathrm{d}t + \chi_1\mathrm{d}W_1^X(t) \text{ (short rate)}\\
\mathrm{d}X_2(t) =& \kappa_2\big(\theta_2\big(\mathcal{MC}(t)\big) - X_2(t)\big)\mathrm{d}t + \chi_2\sqrt{X_2(t)}\mathrm{d}W_2^X(t) \text{ (stochastic volatility)}\\
\rho =& \begin{pmatrix} -1 & 0 \\ 0 & \bar{\rho} \end{pmatrix},
\end{aligned} \tag{6.41}$$

where $\lambda_1, \nu, \theta_1, \theta_2 : \mathcal{E} \to \mathbb{R}$ are real-valued functions of the state of the Markov chain and $\chi_1, a, \lambda_2, \kappa_1, \kappa_2, \chi_2 \in \mathbb{R}$ are constants. Recall that $\theta_1(e_i) = \tilde{\theta}_1 - \frac{\chi_1\lambda_1(e_1)}{\kappa_1}$, $\kappa_1, \chi_1 > 0$ and the definition of function A_2 from (6.8). Furthermore, assume that $\kappa_2 > 0$, $\chi_2 > 0$ and $\theta_2(e_i) > \frac{\chi^2}{2\kappa}$ in order to assure the positiveness of process X_2. In terms of the general notation from (6.2) we have:

$$\begin{aligned}
r =& x_1\\
\mu^X =& \begin{pmatrix} \kappa_1\theta_1(e_i) - \kappa_1 x_1 \\ \kappa_2\theta_2(e_i) - \kappa_2 x_2 \end{pmatrix}\\
\Sigma^X =& \begin{pmatrix} \chi_1 & 0 \\ 0 & \chi_2\sqrt{x_2} \end{pmatrix} \Rightarrow \Sigma^X\big(\Sigma^X\big)' = \begin{pmatrix} \chi_1^2 & 0 \\ 0 & \chi_2^2 x_2 \end{pmatrix}\\
\Sigma =& \begin{pmatrix} A_2(T_1-t)\chi_1 & 0 \\ a & \nu(e_i)\sqrt{x_2} \end{pmatrix}\\
\mu - r =& \begin{pmatrix} \lambda_1(e_i)A_2(T_1-t)\chi_1 \\ a\lambda_1 + \lambda_2\nu(e_i)x_2 \end{pmatrix} \Rightarrow \Sigma^{-1}(\mu - r) = \begin{pmatrix} \lambda_1(e_i) \\ \lambda_2\sqrt{x_2} \end{pmatrix}\\
\Sigma^X\rho\Sigma^{-1}(\mu - r) =& \begin{pmatrix} -\chi_1\lambda_1(e_i) \\ \chi_2\bar{\rho}\lambda_2 x_2 \end{pmatrix}
\end{aligned}$$

$$\Sigma^X \rho\rho'(\Sigma^X)' - \Sigma^X(\Sigma^X)' = \begin{pmatrix} 0 & 0 \\ 0 & \chi_2^2(\bar{\rho}^2 - 1)x_2 \end{pmatrix},$$

which implies that:

$$\begin{aligned}
&\bar{k}^{(0)}(e_i) = \begin{pmatrix} \kappa_1\theta_1(e_i) \\ \kappa_2\theta_2(e_i) \end{pmatrix}, K^{(1)} = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix} \\
&H^{(0)} = \begin{pmatrix} \chi_1^2 & 0 \\ 0 & 0 \end{pmatrix}, H^{(11)} = 0, H^{(12)} = \begin{pmatrix} 0 & 0 \\ 0 & \chi_2^2 \end{pmatrix} \\
&h^{(0)} = \lambda_1^2(e_i), \bar{h}^{(1)} = \begin{pmatrix} 0 \\ \lambda_2^2 \end{pmatrix} \\
&\bar{g}^{(0)} = \begin{pmatrix} -\chi_1\lambda_1(e_i) \\ 0 \end{pmatrix}, G^{(1)} = \begin{pmatrix} 0 & 0 \\ 0 & \chi_2\bar{\rho}\lambda_2 \end{pmatrix} \\
&L^{(0)} = 0, L^{(11)} = 0, L^{(12)} = \begin{pmatrix} 0 & 0 \\ 0 & \chi_2^2(\bar{\rho}^2 - 1) \end{pmatrix}.
\end{aligned}$$

So, Model (6.41) fits in our general framework and we can apply Proposition 6.8 to derive the solution to the considered optimization problem. The result is given in the following proposition:

Proposition 6.9 (Solution and verification for a bond-stock model) *Consider Model (6.41) and assume that:*

$$0 < \tilde{\kappa}_2 \tag{6.42}$$

$$\frac{\delta}{1-\delta}\lambda_2^2 < \frac{\tilde{\kappa}_2^2}{\tilde{\chi}_2^2}, \tag{6.43}$$

where $\tilde{\kappa}_2 = \kappa_2 - \frac{\delta}{1-\delta}\chi_2\bar{\rho}\lambda_2$ *and* $\tilde{\chi}_2^2 = \chi_2^2\big(1 + \frac{\delta}{1-\delta}\bar{\rho}^2\big)$*. Then, the value function to the considered optimization problem is given by:*

$$\begin{aligned}
\Phi =& \frac{1-\delta}{\delta}\alpha\Big\{\frac{1}{1-\delta}\big(v - Fd(t,x_1)\big)\Big\}^{\delta}\mathbb{E}\Big[\exp\Big\{\int_t^T w\big(s,\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\big|\mathcal{MC}(t) = e_i\Big] \\
&\cdot \exp\big\{B_1(t)x_1 + B_2(t)x_2\big\},
\end{aligned} \tag{6.44}$$

where

$$\begin{aligned}
d(t,x_1) =& P_1(t,T,x_1) = \exp\big\{-A_1(T-t) - A_2(T-t)x_1\big\} \\
w(t,e_i) =& \frac{1}{2}\frac{\delta}{1-\delta}\lambda_1^2(e_i) + B_1(t)\big(\kappa_1\theta_1(e_i) - \frac{\delta}{1-\delta}\chi_1\lambda_1(e_i)\big) \\
&+ B_2(t)\kappa_2\theta_2(e_i) + \frac{1}{2}\frac{1}{1-\delta}B_1^2(t)\chi_1^2 \\
B_1(t) =& \frac{\delta}{\kappa_1}\big(1 - \exp\big\{-\kappa_1(T-t)\big\}\big) \\
B_2(t) =& \frac{-b(\tilde{\kappa}_2 + c)\exp\{-c(T-t)\} + \tilde{\kappa}_2 - c}{\tilde{\chi}_2^2\big(1 - b\exp\{-c(T-t)\}\big)},
\end{aligned} \tag{6.45}$$

with A_1 and A_2 as in (6.7) and (6.8), respectively, and

$$c := \sqrt{\tilde{\kappa}_2^2 - \frac{\delta}{1-\delta}\lambda_2^2\tilde{\chi}_2^2}$$
$$b := \frac{\tilde{\kappa}_2 - c}{\tilde{\kappa}_2 + c}.$$

The optimal portfolio has the following expression:

$$\bar{\pi} = \frac{V^{\bar{\pi}} - Fd}{V^{\bar{\pi}}} \left\{ \underbrace{\frac{1}{1-\delta} \begin{pmatrix} \frac{\lambda_1}{A_2(T_1-t)\chi_1} - \frac{a\lambda_2}{A_2(T_1-t)\chi_1\nu} \\ \frac{\lambda_2}{\nu} \end{pmatrix}}_{=:\bar{\pi}^{MV}} + \underbrace{\frac{1}{1-\delta} \begin{pmatrix} -\frac{1}{A_2(T_1-t)}\left(B_1 + \frac{a\bar{\rho}\chi_2}{\chi_1\nu}B_2\right) \\ \frac{\bar{\rho}\chi_2}{\nu}B_2 \end{pmatrix}}_{=:\bar{\pi}^H} \right\} + \underbrace{\begin{pmatrix} \frac{Fd}{V^{\bar{\pi}}}\frac{A_2(T-t)}{A_2(T_1-t)} \\ 0 \end{pmatrix}}_{=:\bar{\pi}^F} \Bigg|_{(t,V^{\bar{\pi}}(t),X(t),\mathcal{MC}(t))}. \tag{6.46}$$

For the proof see Appendix D.

Now we are interested in better understanding the optimal portfolio and how the single model parameters influence it. Analogously to the case of the power utility function, $\bar{\pi}^{MV}$ is mainly driven by the ratio between the market price of risk and the volatility for the single assets and is called the mean-variance part. The second term $\bar{\pi}^H$ adjusts it for the additional risk coming from the stochastic factors and is called the hedging part. As already discussed in Section 6.2, the last term $\bar{\pi}^F$ is used for the replication of the floor, that is why we call it the replication part. The mean-variance term is intuitive: the higher the market price of risk for the considered asset and the lower its volatility, the higher the investment in this asset. As the two assets are correlated the excess return for P_2 influences also the position in P_1, e.g. if $a < 0$, i.e. the bond and the stock are negatively correlated, a higher mean-variance position in the stock leads to a higher investment in the bond as well. Economically this can be interpreted as a diversification effect.
The same correlation term proportional to a can be observed in the hedging term as well. The remaining drivers of the hedging term are basically the parameters of the stochastic factors. To better understand the hedging effect, let us assume for a moment that $a = 0$. It becomes clear that the hedging bond position is due to the stochastic interest rate X_1 and the hedging stock position comes from the stochastic volatility driver X_2. It is interesting to explore the sign of $\bar{\pi}^H$. As

in Section 4.3 and 5.3.2 it can be easily shown that $sign(-\frac{B_1}{A_2(T_1-t)}) = -sign(\delta)$ and $sign(\frac{\bar{\rho}\chi_2}{\nu}B_2) = sign(\bar{\rho})sign(\nu)sign(\delta)$. So, if $\bar{\rho} < 0$, which corresponds to the leverage effect, $\nu > 0$, which means a positive stock position, and $\delta < 0$, the hedging term is positive. The reason is that the investor wants to hedge his portfolio against the risk of falling interest rates and market price of risk. A falling short rate implies rising bond price, and decreasing X_2 is related to increasing stock price due to the leverage effect. So, the hedge is achieved by a positive position in both risky assets. If $\delta > 0$. i.e. we are dealing with a less risk averse investor, the hedging term is negative, as the investor speculates on rising X_1 and X_2. The importance of the hedging factor for different values for the relevant parameters is further studied in the next section.
Now let us have a look at the replicating term $\bar{\pi}^F$. It contains just a bond position. It is clear that $\bar{\pi}_1^F > 0$. Furthermore, the higher the floor relative to the current wealth, the higher its importance. It also holds that the closer T_1 to T, the higher $\bar{\pi}_1^F$. This is explained by the fact that then the traded bond replicates better the floor at maturity.

6.4 Application to real data

In this section we illustrate the derived results for Model (6.41) using real data. To this aim we first estimate the model parameters and then study the performance of the derived optimal Strategy (6.46).

6.4.1 Data and introduction

For the parameter estimation we use weekly time series over the period from December, 1st 2005 till December, 30th 2010. The used data is publicly available from the Frankfurt Stock Exchange (www.boerse-frankfurt.de) and the Deutsche Bundesbank (www.bundesbank.de). We identify the stock price process P_2 with the German stock index DAX and approximate its volatility process $\sqrt{a^2 + \nu^2\big(\mathcal{MC}(t)\big)X_2(t)}$ by the corresponding volatility index VDAX. For a discussion of this approximation we refer to [1]. For the traded bond P_1 we consider a 10 years German government zero-coupon bond. Its price is calculated from the yield curve published by the Deutsche Bundesbank. The estimation procedure is based on the weekly excess returns of the stock index and the bond. They are approximated by the difference between the respective weekly log-returns and the one-week Euribor rate. For the estimation we use additionally bond yields with maturities ranging from half a year to 25 years. A documentation of the estimation procedure applied by the Deutsche Bundesbank to extract the yield curve from observed Government bond prices can be found in [94]. After the parameter estimation we perform an out-of-sample analysis of the derived portfolio strategy by applying it to the data over the period from January, 3rd 2011 till February, 2nd 2015. We show how the optimal portfolio is adjusted

to the regime shifts. Furthermore, we compare the performance of the strategy to the case wherein the investor does not consider Markov switching and illustrate the huge misallocation that may arise if the regime switching is neglected.
The observations of the DAX, the VDAX and the 10 years bond price can be seen in Figure 6.1 and the considered interest rates are presented in Figure 6.2.

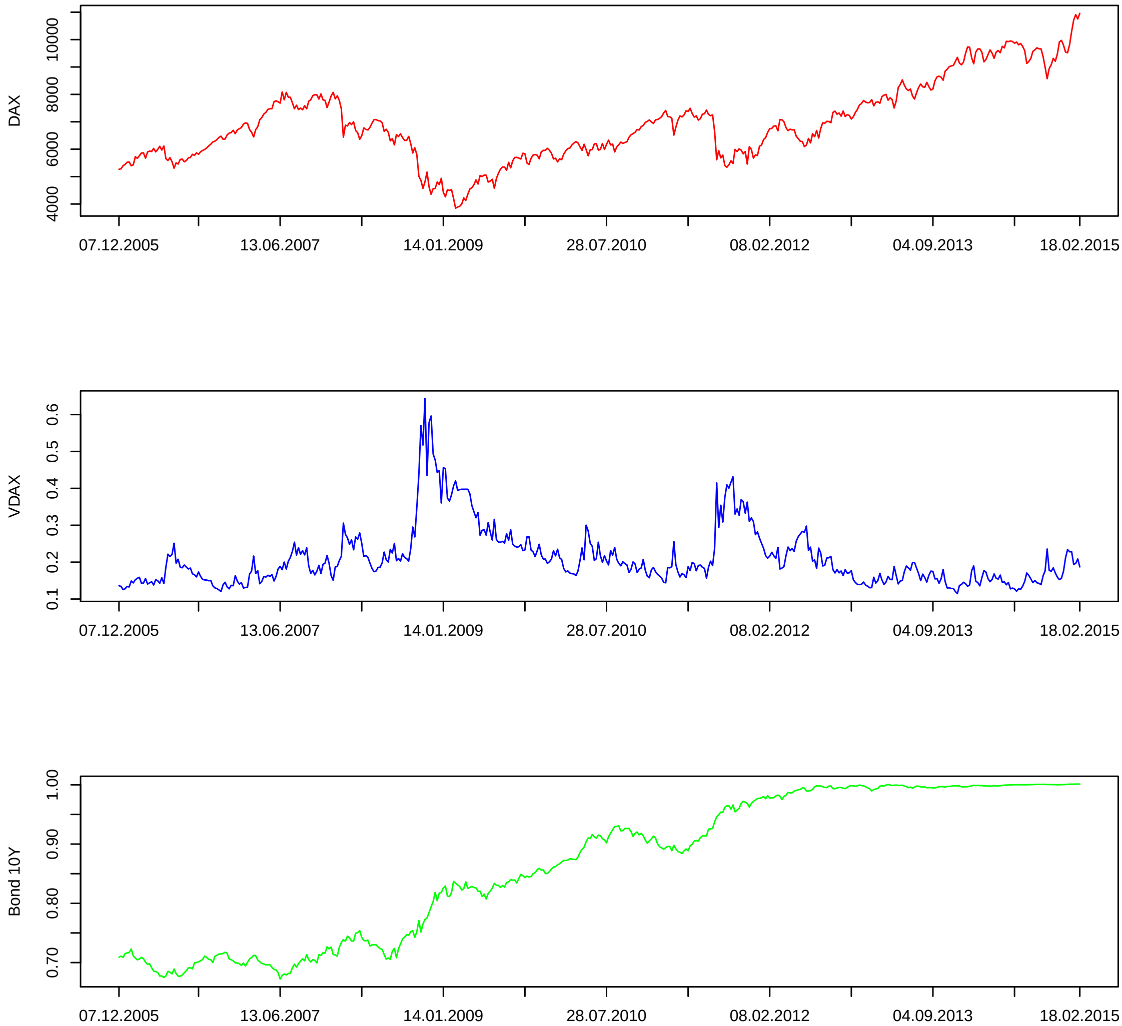


Figure 6.1: Data used for the estimation and the out-of-sample testing: weekly DAX (upper plot) and VDAX (middle plot) time series and price development of a bond with maturity 10 years at the first observation point (lower plot).

6.4.2 Parameter estimation

Before we can test the derived strategies we need to estimate the model parameters using the observed data. To this aim we propose in this section a possible estimation

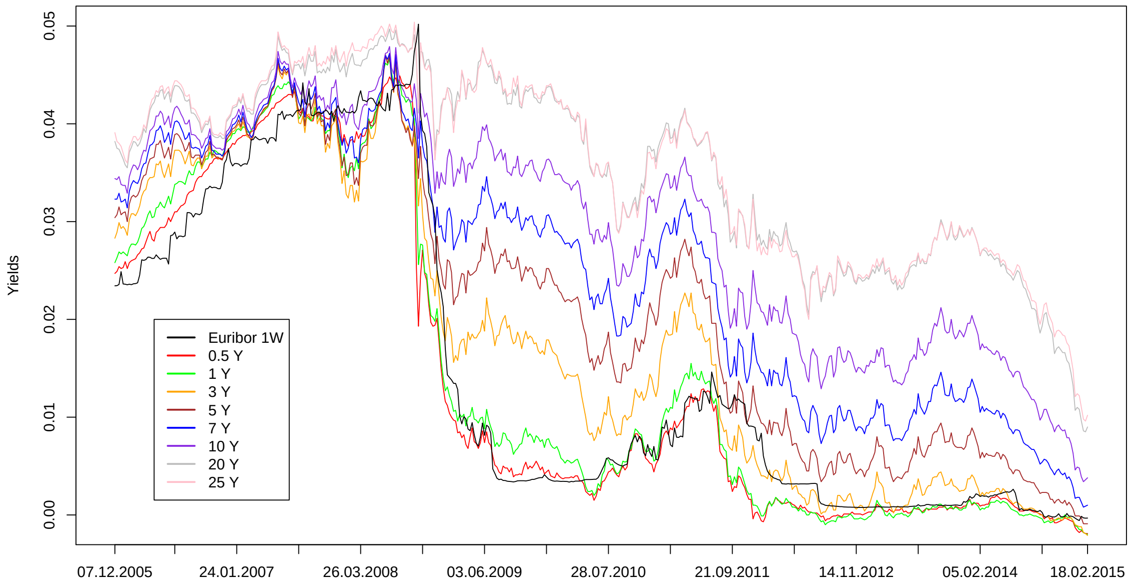


Figure 6.2: Data used for the estimation and out-of-sample testing: 1 week Euribor and German government bond yields for different maturities.

procedure[1]. It consists of 4 steps. First, we use the empirical bond-stock covariance to estimate parameter a. Then, we apply a Baum-Welch Algorithm[2] to the volatility process to estimate its parameters and filter the state of the Markov chain. Afterwards, by using a Kalman Filter[3] for bond yields with different maturities we derive maximum likelihood estimators for the parameters of the short rate and the traded bond price process. Finally, the remaining parameters for the stock price process are obtained via linear regression and the stock-volatility correlation parameter is set to the empirical correlation between the random increments of the two processes. We set the discretization step to $\triangle = \frac{7}{365}$.

Step 1: Stock-bond covariance

As already mentioned, for the estimation we use the excess log-returns of the bond and the stock:

$$\begin{aligned} R_1(t) :=& \ln\left(\frac{P_1(t+\triangle)}{P_1(t)}\right) - X_1(t)\triangle \\ =& \big[\lambda_1\big(\mathcal{MC}(t)\big)A_2(T_1-t)\chi_1 - \frac{1}{2}\big\{A_2(T_1-t)\chi_1\big\}^2\big]\triangle \\ &+ A_2(T_1-t)\chi_1\big(\underbrace{W_1^P(t+\triangle) - W_1^P(t)}_{=:\sqrt{\triangle}\varepsilon_1^P(t+\triangle)}\big) \end{aligned} \tag{6.47}$$

[1]We would like to thank Laslo Bollmann and Andreas Lichtenstern for coding parts of a previous version of the estimation procedure under the close guidance of Daniela Neykova.

[2]For an introduction to the Baum-Welch Algorithm we refer to [9] and [106].

[3]An introduction to the Kalman Filter can be found in [64], [65] and [55].

$$\begin{aligned}
R_2(t) :=& \ln\left(\frac{P_2(t+\triangle)}{P_2(t)}\right) - X_1(t)\triangle \\
=& \big[\lambda_1\big(\mathcal{MC}(t)\big)a + \lambda_2\nu\big(\mathcal{MC}(t)\big)X_2(t) - \frac{1}{2}\left\{a^2 + \nu^2\big(\mathcal{MC}(t)\big)X_2(t)\right\}\big]\triangle \\
&+ a\big(W_1^P(t+\triangle) - W_1^P(t)\big) + \nu\big(\mathcal{MC}(t)\big)\sqrt{X_2}\big(\underbrace{W_2^P(t+\triangle) - W_2^P(t)}_{=:\sqrt{\triangle}\varepsilon_2^P(t+\triangle)}\big),
\end{aligned} \tag{6.48}$$

where $\varepsilon_1^P(t)$ and $\varepsilon_2^P(t)$ are i.i.d. standard-normally distributed for all time points t in the discretization grid. It follows that:

$$\begin{aligned}
&Cov\Big(\frac{R_1(t) - \overbrace{[\lambda_1\big(\mathcal{MC}(t)\big)A_2(T_1-t)\chi_1 - \frac{1}{2}\{A_2(T_1-t)\chi_1\}^2]\triangle}^{\approx\ \text{emp. mean of } R_1}}{\underbrace{A_2(T_1-t)\chi_1}_{\approx\ \text{emp. st. dev. of } R_1/\sqrt{\triangle}}}, \\
&R_2(t) + [\frac{1}{2}\underbrace{\{a^2 + \nu^2\big(\mathcal{MC}(t)\big)X_2(t)\}}_{=VDAX^2}]\triangle\Big) \\
&= Cov\Big(\sqrt{\triangle}\varepsilon_1^P(t+\triangle), \big[\lambda_1\big(\mathcal{MC}(t)\big)a + \lambda_2\nu\big(\mathcal{MC}(t)\big)X_2(t)\big]\triangle \\
&\quad + a\sqrt{\triangle}\varepsilon_1^P(t+\triangle) + \nu\big(\mathcal{MC}(t)\big)\sqrt{X_2}\sqrt{\triangle}\varepsilon_2^P(t+\triangle)\Big) \\
&= Cov\Big(\sqrt{\triangle}\varepsilon_1^P(t+\triangle), a\sqrt{\triangle}\varepsilon_1^P(t+\triangle)\Big) \\
&= a\triangle,
\end{aligned}$$

where in the last but one equation we have used the fact that W_1^P is independent of W_2^X and W_2^P. By applying the above equation we obtain $a = -0.096$. Note that the instantaneous bond-stock correlation ($\frac{a}{\sqrt{a^2+\nu^2\left(\mathcal{MC}(t)\right)X_2(t)}}$) is a stochastic process, however, as an orientation for the average correlation we can replace the stock volatility by its empirical mean and obtain $\frac{a}{mean(VDAX)} = -0.425$. Recent empirical studies on the bond-stock correlation in different markets report as well highly negative bond-stock correlation over the last 15 years, see e.g. [3], [6] and [60]. The average correlation we obtain is in accordance with these findings.

Step 2: Volatility process and Markov chain

Let us now continue with the volatility process. It is clearly visible in Figure 6.1 that the volatility process exhibits short periods with very high levels, followed by longer periods with moderate values. We would like to reflect this behavior by considering a Markov chain with two regimes and estimating its parameters directly from the volatility data. One can also recognize from Figure 6.1 that the periods with high volatility are related to negative DAX returns and falling interest rates. This observation is in accordance to existing literature: [79] and [28] conduct detailed empirical studies to motivate the existence of two market regimes: one

with high stock return and low volatility and one with lower returns and higher volatility. For further motivation of Markov-modulated models consult [98] and [54] who show that Markov-switching models capture well observed properties of stock returns distributions such as heavy tails, volatility clustering and asymmetry. [52] and [11] motivate empirically their application to asset allocation problems. Furthermore, empirical evidences for regime-switching stochastic volatility are presented in [35] and [30]. Motivated by all these empirical studies we consider Markov-switching model parameters with two possible regimes and use the VDAX time series to estimate the parameters of the Markov chain and its state at each time point.
To ease the exposition we introduce the following notations:

$$\begin{aligned} Var_a(t) &:= a^2 + \nu^2\big(\mathcal{MC}(t)\big)X_2(t) = VDAX(t)^2 \\ Var_0(t) &:= \nu^2\big(\mathcal{MC}(t)\big)X_2(t) = VDAX(t)^2 - a^2. \end{aligned}$$

Then, for the Euler discretization of Var_0 conditional on the state of the Markov chain we obtain:

$$\begin{aligned} Var_0(t+\triangle) =& Var_0(t) + \kappa_2\Big(\underbrace{\nu\big(\mathcal{MC}(t)\big)^2\theta_2\big(\mathcal{MC}(t)\big)}_{=:\hat{\theta}_2(\mathcal{MC}(t))} - Var_0(t)\Big)\triangle \\ &+ \underbrace{|\nu\big(\mathcal{MC}(t)\big)|\chi_2}_{=:\hat{\chi}_2(\mathcal{MC}(t))}\sqrt{Var_0(t)}\underbrace{\big(W_2^X(t+\triangle) - W_2^X(t)\big)}_{=:\sqrt{\triangle}\varepsilon_2^X(t+\triangle)}, \end{aligned} \tag{6.49}$$

where $\varepsilon_2^X(t)$ are i.i.d. standard normally distributed for all time points t in the grid. We rewrite it in the form of a linear regression:

$$\begin{aligned} \underbrace{\frac{Var_0(t+\triangle) - Var_0(t)}{\sqrt{Var_0(t)}\sqrt{\triangle}}}_{=:y(t+\triangle)} =& \underbrace{\kappa_2\hat{\theta}_2\big(\mathcal{MC}(t)\big)}_{=:\alpha_1(\mathcal{MC}(t))}\underbrace{\frac{\sqrt{\triangle}}{\sqrt{Var_0(t)}}}_{=:x_1(t)}\underbrace{-\kappa_2}_{=:\alpha_2}\underbrace{\sqrt{Var_0(t)}\sqrt{\triangle}}_{=:x_2(t)} \\ &+ \hat{\chi}_2\big(\mathcal{MC}(t)\big)\varepsilon_2^X(t+\triangle). \end{aligned} \tag{6.50}$$

For comparison reasons we first estimate the parameters without Markov switching using a classical linear regression. Then we obtain the following values: $\kappa_2 = 2.115$, $\hat{\theta}_2 = 0.051$, $\hat{\chi}_2 = 0.505$. The standard errors and levels of significance for the parameters are presented in Table 6.1: The high levels of significance suggest

Coefficient	Value	Standard error	Level of significance
α_1	0.1082	0.0493	0.0291
α_2	-2.11531	1.4146	0.1360

Table 6.1: Results for the linear regression in (6.50) with constant parameters (no Markov switching).

hat the proposed model does not fit well the observed data. The QQ plot of the residuals shown in Figure 6.3 confirms this conclusion. As we will see, this is

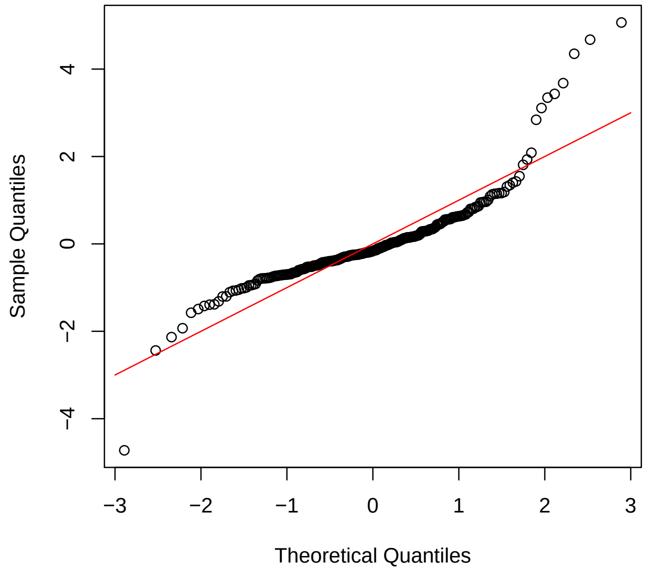


Figure 6.3: Normal QQ plot of the residuals of the linear regression in (6.50) with constant parameters (no Markov switching).

strongly improved by considering Markov switching. To estimate the parameters of the Markov-switching linear regression we use the R-package `MSwM` that contains a suitable implementation of the Baum-Welch Algorithm. Applications of the Baum-Welch Algorithm in similar contexts can be found e.g. in [11] and [57]. The following results are obtained: $\kappa_2 = 5.563$, $\hat{\theta}_2(e_1) = 0.027$, $\hat{\theta}_2(e_2) = 0.285$, $\hat{\chi}_2(e_1) = 0.309$, $\hat{\chi}_2(e_2) = 1.205$, $Q = \begin{pmatrix} -1.109 & 1.109 \\ 11.385 & -11.385 \end{pmatrix}$. So, the algorithm recognizes two regimes with very different parameters: the first one is characterized by a moderate mean-reverting volatility level $\sqrt{a^2 + \hat{\theta}_2(e_1)} = 0.19$ and lower volatility of volatility. On the contrary, in the second state both the average stock volatility $\sqrt{a^2 + \hat{\theta}_2(e_2)} = 0.54$ and the volatility of volatility are much higher. These results coincide with our intuitive expectations based on Figure 6.1. As already mentioned in Section 4.3.4 the average occupation time for e_1 is $-\frac{1}{q_{11}} \approx 11$ months, whereas for e_2 it is $-\frac{1}{q_{22}} \approx 1$ month. Thus, we can identify the first state as a normal stock market situation, whereas the second one describes shorter turbulent periods. As we will see later on, e_2 is also characterized by lower stock returns, which allows to interpret it as a crisis. These results are in accordance with the findings in the studies cited at the beginning of Step 2. Furthermore, note that as explained in Remark 2.45 the average half-life for the mean-reversion without regime switching is $\frac{0.7}{\kappa_2} \approx 3$ months, whereas with the regime differentiation it reduces to approx. 1 months. The reason is that without the Markov chain the regime switch is interpreted as a deviation from the overall mean-reversion level and this biases the estimation for κ_2.

A look at the levels of significance of the parameters in Table 6.2 and the QQ plot in Figure 6.4 shows that the model with Markov switching describes the observed

data much better than the model with constant parameters.
As a byproduct of the Baum-Welch algorithm we obtain also the probabilities

Coefficient	Value	Standard error	Level of significance
$\alpha_1(e_1)$	0.1527	0.0353	$1.520e^{-5}$
$\alpha_1(e_2)$	1.5976	0.6872	0.0209
α_2	-5.5631	1.1823	$2.535e^{-6}$

Table 6.2: Results for the linear regression in (6.50) with Markov-switching parameters.

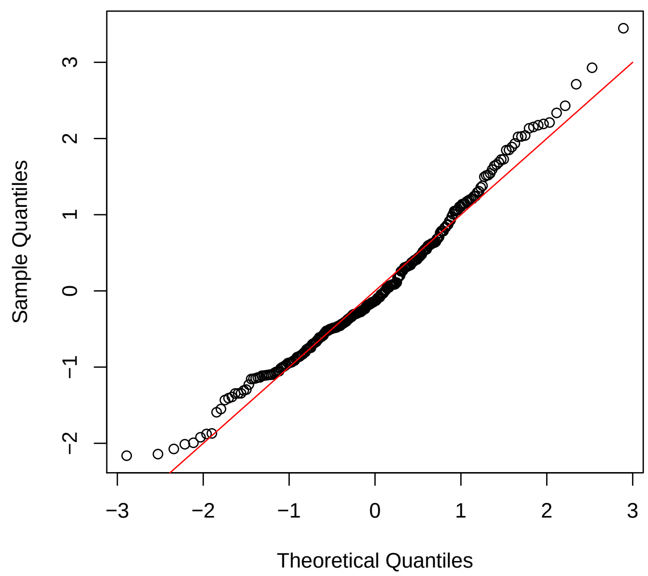


Figure 6.4: Normal QQ plot of the residuals of the linear regression in (6.50) with Markov-switching parameters.

for being in the single states at each time point conditional on the information till this time (so-called filtered probabilities) and conditional on the whole data (smoothed probabilities). They are plotted in Figure 6.5 (first and second plot). We use the smoothed probabilities to determine the state of the Markov chain by a simple intuitive rule: we assume that the Markov chain is in the most probable state according to the smoothed probability. The filtered state of $\mathcal{MC}$ can also be seen in Figure 6.5 (third plot). Figure 6.6 shows the filtered state together with the DAX, VDAX and Euribor time series. The algorithm recognizes the financial crisis of 2008, as well as a few short turbulent periods. These results coincide with the interpretation of the second state as a crisis. Figure 6.7 depicts process Var_0 together with its estimated mean-reverting level for the two regimes. It can be observed that the estimated parameters match well the observed data. Alternatively one can apply a more sophisticated method for the estimation of the state of the Markov chain by considering further macroeconomic factors, see e.g. [57].

Step 3: Short rate and bond price

Now that we have the estimates for the states of the Markov chain, we can continue

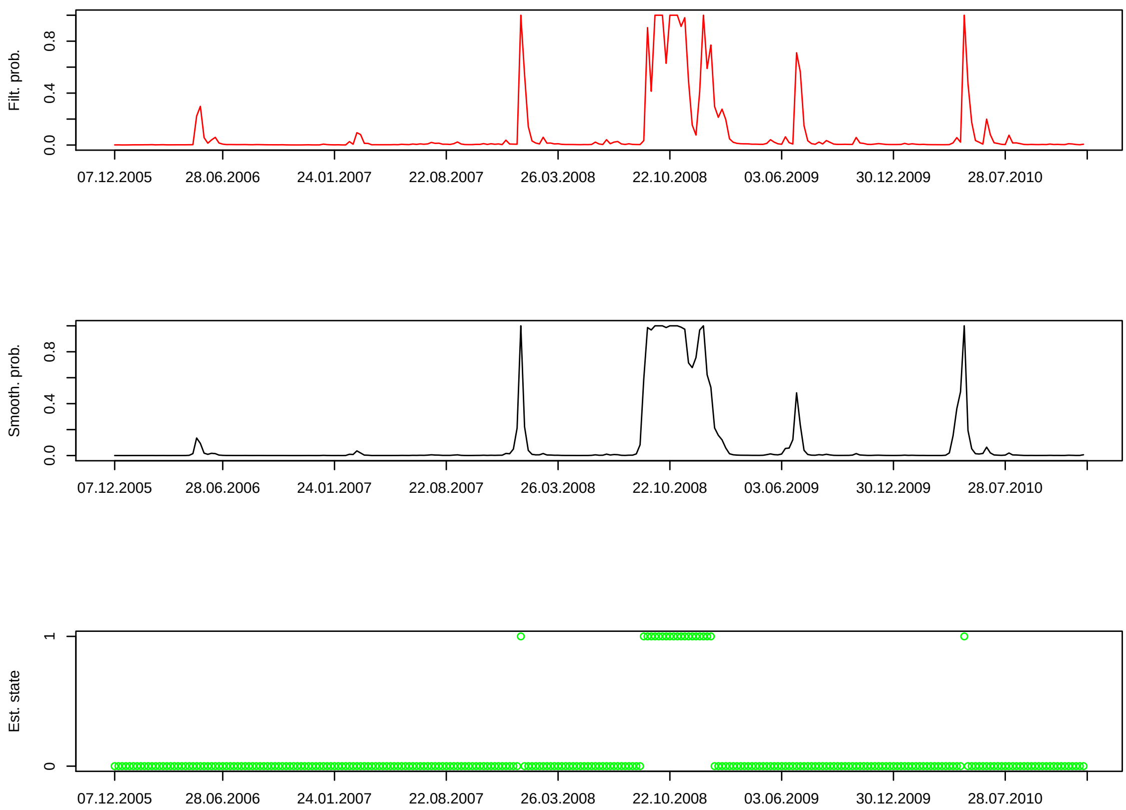


Figure 6.5: Filtered probability for the second state (upper plot), smoothed probabilities for the second state (middle plot) and estimated state using the smoothed probability: "0"$=e_1$, "1"$=e_2$ (lower plot).

with the parameters for the bond and the short rate. One difficulty is that the short rate is not directly observable. That is why we apply a Kalman Filter methodology to bond yields with different maturities. For detailed empirical studies on the Kalman Filter application to short rate models consult e.g. [5] and [27]. [29], [42] and [105] motivate the regime-switching extension of such kind of models.
The first step for the application of the Kalman Filter is to derive the unobservable transition equation from the solution of the SDE for the short rate process. We assume that the Markov chain remains constant on the discretization interval $[t, t + \triangle)$ and obtain the following:

$$X_1(t+\triangle) = \theta_1\big(\mathcal{MC}(t)\big)\big(1 - \exp\{-\kappa_1\triangle\}\big) + X_1(t)\exp\{-\kappa_1\triangle\} \tag{6.51}$$

$$+ \underbrace{\chi_1 \int_t^{t+\triangle} \exp\{\kappa_1(s - t - \triangle)\}\mathrm{d}W_1^X(s)}_{\stackrel{d}{=}\mathcal{N}(0,\frac{\chi_1^2}{2\kappa_1}(1-\exp\{-2\kappa_1\triangle\}))} . \tag{6.52}$$

This equation can be rewritten as follows:

$$X_1(t+\triangle) = \theta_1\big(\mathcal{MC}(t)\big)\big(1 - \exp\{-\kappa_1\triangle\}\big) + X_1(t)\exp\{-\kappa_1\triangle\} + \Sigma_1\varepsilon_1^X(t), \tag{6.53}$$

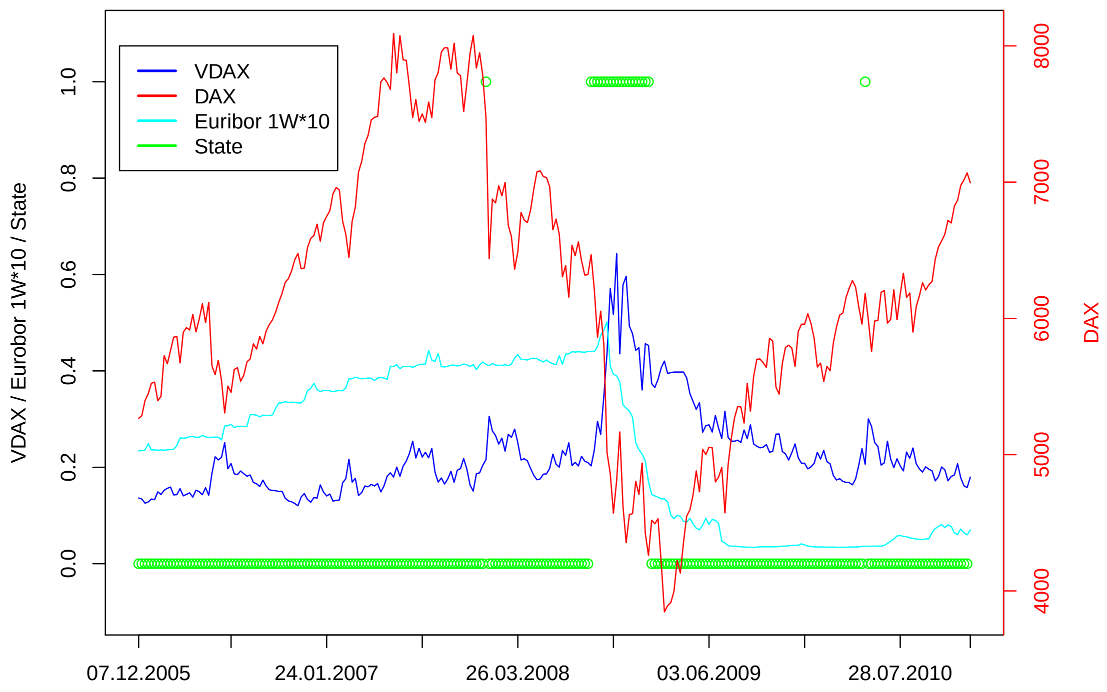


Figure 6.6: On the left y-axis: VDAX time series (blue line), 1-week-Euribor time series scaled by the factor 10 (cyan line), estimated state (green dots: "0"$=e_1$, "1"$=e_2$); on the right y-axis: DAX time series (red line).

where $\theta_1(e_i) = \tilde{\theta}_1 - \frac{\chi_1 \lambda_1(e_1)}{\kappa_1}$, $\Sigma_1^2 := \frac{\chi_1^2}{2\kappa_1}(1 - \exp\{-2\kappa_1 \triangle\})$ and $\varepsilon_1^X(t)$ are i.i.d. standard normally distributed. For the measurement equations we use bond yields with d different times to maturity $\tau_1, \dots, \tau_d$:

$$\begin{aligned}
R_{\tau_1}(t, X_1(t)) :&= -\frac{\ln\left\{P_1(t, t+\tau_1, X_1(t))\right\}}{\tau_1} + \Sigma_1^R \varepsilon_1^R(t) \\
&= \frac{A_1(\tau_1)}{\tau_1} + \frac{A_2(\tau_1)}{\tau_1} X_1(t) + \Sigma_1^R \varepsilon_1^R(t) \\
&\vdots \\
R_{\tau_d}(t, X_1(t)) :&= -\frac{\ln\left\{P_d(t, t+\tau_d, X_1(t))\right\}}{\tau_d} + \Sigma_d^R \varepsilon_d^R(t) \\
&= \frac{A_1(\tau_d)}{\tau_d} + \frac{A_2(\tau_d)}{\tau_d} X_1(t) + \Sigma_d^R \varepsilon_d^R(t),
\end{aligned} \tag{6.54}$$

where $\varepsilon_1^R(t), \dots, \varepsilon_d^R(t)$ are i.i.d. standard normally distributed measurement errors. We need to assume the random measurement errors, so that we can consider simultaneously various maturities without obtaining a non-solvable system. Note that due to the dependence of λ_1 and thus θ_1 on the Markov chain we have to deal with time-dependent parameters in the transition Equation (6.52). To this aim we use the `R`-package `FKF` that offers a Kalman Filter implementation with

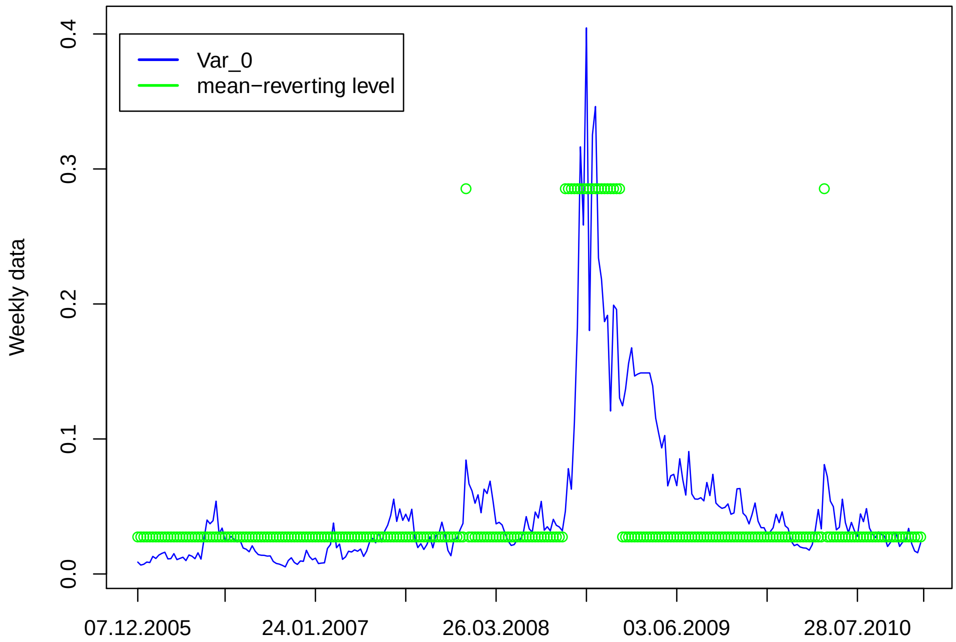


Figure 6.7: Process Var_0 (blue line) together with the estimated mean-reverting levels $\hat{\theta}_2(e_1)$ and $\hat{\theta}_2(e_2)$ for the estimated regimes.

time-dependent model parameters. We apply it to the German government bond yields for maturities 0.5, 1, 2, 3, 4, 5, 7, 10, 20 and 25 years and obtain the following results: $\kappa_1 = 0.266(0.004)$, $\tilde{\theta}_1 = 0.047(0.0002)$, $\chi_1 = 0.013(0.0005)$, $\lambda_1(e_1) = 0.146(0.316)$, $\lambda_1(e_2) = 6.674(1.549)$, where the values in parenthesis correspond to the standard errors of the estimates. First of all, note that there is an immense difference between the values for λ_1 for the two states. This is one more indicator for the existence of two very different market regimes. Moreover, we would like to draw the attention of the reader to the well-known difficulty when estimating λ_1, recognizable in the big standard errors of the estimates for λ_1 (see e.g. [105]). Incorporating expert opinions or applying a more sophisticated estimation method might be useful. However, as the aim of this exemplary estimation is the illustration of the derived optimal strategies we do not go deeper in this issue. Let us have a closer look at the estimated values for λ_1. Substitution in the formula $\theta_1 = \tilde{\theta}_1 - \frac{\chi_1 \lambda_1}{\kappa_1}$ leads to $\theta_1(e_1) = 0.040$ and $\theta_1(e_2) = -0.275$. At first sight the negative mean-reverting level for the second regime might seem surprising. However, it is explained by the fact that during the crisis in 2008 the interest rates dropped extremely fast to very low levels. Furthermore, note that in our case the negative $\theta_1(e_2)$ does not mean that the short rate will become negative, as the average time needed for the mean-reversion is much longer than the average time spent in the second regime. To see this recall that the half-life of process X is given by $\frac{\ln 2}{\kappa_1} \approx 2.6$ years, whereas the expected time spent in regime e_2 is $-\frac{1}{q_{22}} \approx$1 month. The filtered short rate and the reconstructed yields using the derived parameters

(without the measurement errors) are shown in Figure 6.8. It is recognizable that the yields on both extreme ends - very short and very long - are not reconstructed accurately, however the yields for maturities from three to seven years are matched well. Clearly we are dealing with a situation where the one-factor interest rate model meets its limits. However, we prefer to keep this simplicity of the interest rate modeling as it allows us to present the portfolio optimization results in a clearer way. What is more, the investor trades in a bond with maturity of ten years at the initial time point and has a trading horizon of roughly five years. So, she is not strongly affected by the mismatch on the very short and very long ends.
For comparison, the estimated parameters without Markov switching are given by: $\kappa_1 = 0.266(0.004)$, $\tilde{\theta}_1 = 0.047(0.0002)$, $\chi_1 = 0.013(0.0005)$, $\lambda_1 = 0.689(0.284)$, where again the values in parenthesis correspond to the standard errors of the estimates. The values for κ_1, $\tilde{\theta}_1$ and χ_1 are the same as in the case of a Markov-switching λ_1. The reason is that λ_1 does not appear directly in the formulas for the bond yields in (6.54) but only through the state of the unobservable short rate X_1.

Step 4: Stock price

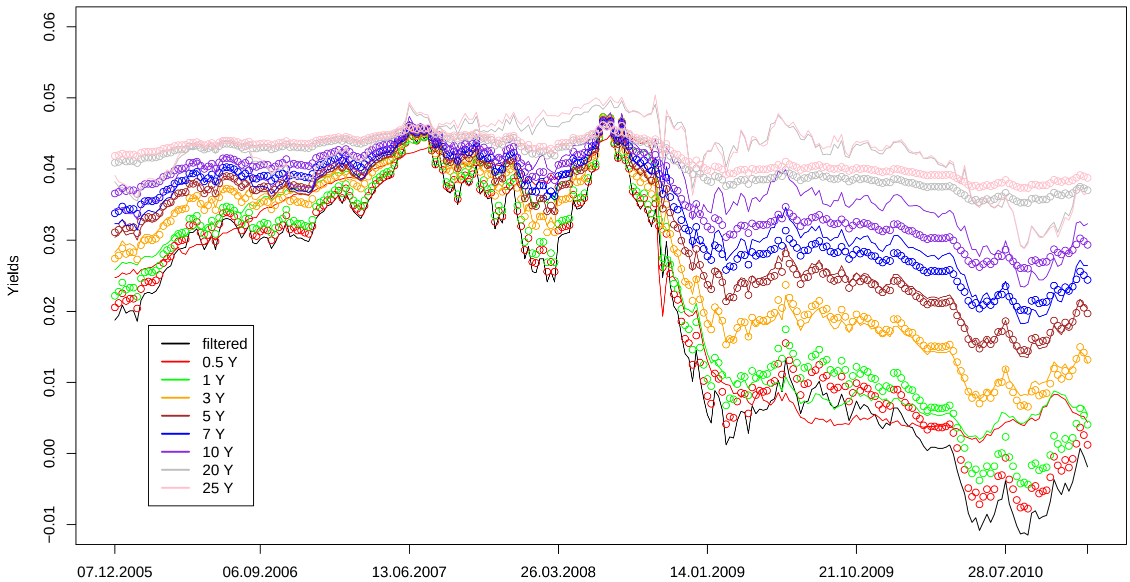


Figure 6.8: 1 week Euribor and German government bond yields for different maturities: real data (lines) and reconstructed yields using the estimated parameters (circles).

What is left to be estimated are the stock parameters and the stock - volatility correlation. For these parameters we again consider the weekly excess return of the stock price process conditional on the state of the Markov chain and rewrite it as a linear regression:

$$\underbrace{\frac{R_2(t) - \left[\lambda_1(\mathcal{MC}(t))a - \frac{1}{2}Var_a(t)\right]\triangle}{\underbrace{\sqrt{Var_a(t)}}_{=VDAX(t)}\sqrt{\triangle}}}_{=:y_2(t+\triangle)} = \underbrace{\frac{\lambda_2}{\nu(\mathcal{MC}(t))}}_{=:\beta(\mathcal{MC}(t))} \underbrace{\frac{\overbrace{\nu^2(\mathcal{MC}(t))X_2}^{=Var_0(t)}}{\sqrt{Var_a(t)}}\sqrt{\triangle}}_{=:x(t)} + \varepsilon_2^P(t+\triangle). \tag{6.55}$$

We will use the states of the Markov chain from Step 2 and a linear regression estimation to obtain the unknown parameters. However, note that $\beta(e_1)$ and $\beta(e_2)$ can not be chosen freely, as from Step 2 we have that $\frac{\beta(e_1)}{\beta(e_2)} = \frac{\nu(e_2)}{\nu(e_1)} = \frac{\hat{\chi}_2(e_2)sign(\nu(e_2))}{\hat{\chi}_2(e_1)sign(\nu(e_1))}$. Thus we cannot run two separate linear regressions for the two states of the Markov chain. To deal with this restriction we transform the x-values corresponding to the second state by multiplying them with $\frac{\hat{\chi}_2(e_1)}{\hat{\chi}_2(e_2)}$. By doing so we obtain a linear regression only in terms of $\beta(e_1)$ over the whole time series. The following values are obtained: $\frac{\lambda_2}{\nu(e_1)} = 5.102$, $\frac{\lambda_2}{\nu(e_2)} = -1.307$. Note that $\frac{\lambda_2}{\nu}$ corresponds to the ratio between the excess return related to the stochastic volatility and the corresponding stochastic part of the variance. There is a clear difference between the two states: this ratio is much higher in the first state, whereas in the second state it even becomes negative. This result confirms the interpretation of the states of the Markov chain as a normal situation and a crisis. To get an intuition on the estimated parameters in Table 6.3 we compare the empirical mean of the stock variance and excess return with the estimated average values for the two regimes. More precisely, we compare the following:

$$\begin{aligned}
\text{emp. mean}(Var_a(t)|\mathcal{MC}(t) = e_i) &\approx a^2 + \hat{\theta}_2(e_i) \\
\text{emp. mean}(R_2(t)|\mathcal{MC}(t) = e_i) &\approx \Big[\lambda_1(e_i)a + \frac{\lambda_2}{\nu(e_i)}\text{emp. mean}(Var_0(t)|\mathcal{MC}(t) = e_i) \\
&\quad - \frac{1}{2}\text{emp. mean}(Var_a(t)|\mathcal{MC}(t) = e_i)\Big]\triangle,
\end{aligned}$$

for $i = 1, 2$. It can be recognized that the estimated average for the variance is lower than the empirical mean in state e_1 and higher in regime e_2. The reason is that at every regime switch the process starts moving towards its long-term mean for the corresponding regime, however it takes some time till it reaches it. For the excess return we observe that the empirical mean is exactly matched in the first regime, in contrast to the second regime, where we observe a difference. This is explained by the fact that the time series contains much more observations for state e_1, so regime e_1 has a bigger weight in the computation of the estimates via the linear regression. Without Markov switching one obtains a value between the two regimes: $\frac{\lambda_2}{\nu} = 2.2304$. Although the QQ plots for both regressions with and without Markov switching are quite similar (see Figures 6.9 and 6.10), allowing for Markov switching improves strongly the level of significance of the considered parameter (compare Tables 6.4 and 6.5). We continue with the estimation of the

state	Excess return		Variance	
	empirical mean	estimated average	empirical mean	estimated average
e_1	0.003	0.003	0.047	0.037
e_2	-0.026	-0.019	0.200	0.295

Table 6.3: Comparison of the empirical mean of the stock excess return and variance with the estimated average values for the two regimes.

Coefficient	Value	Standard error	Level of significance
β	2.234	2.072	0.282

Table 6.4: Results for the linear regression in (6.55) with constant parameters (no Markov switching).

stock-volatility correlation. To this aim we consider the noise terms in the weekly excess return of P_2 and the Euler discretization of Var_0:

$$\varepsilon_2^P(t+\triangle) = \frac{R_2(t) - \left[\lambda_1 a + \frac{\lambda_2}{\nu} Var_0(t) - \frac{1}{2} Var_a(t)\right]\triangle - a\sqrt{\triangle}\varepsilon_1^P(t+\triangle)}{\sqrt{Var_0(t)}\sqrt{\triangle}}$$
$$\varepsilon_2^X(t+\triangle) = \frac{Var_0(t+\triangle) - Var_0(t) - \left[\kappa_2\big(\hat{\theta}_2 - Var_0(t)\big)\right]\triangle}{\hat{\chi}_2\sqrt{Var_0(t)}\sqrt{\triangle}},$$

where for better readability the dependence on the Markov chain is omitted. Note that ε_1^P is obtained from the weekly excess return of P_1 given in (6.47). The empirical correlation of ε_2^P and ε_2^X yields $\bar{\rho} = -0.454$. For the case without Markov switching we get $\bar{\rho} = -0.514$.

The estimated values for the parameters are summarized in Table 6.6. For comparison, we also include the parameter estimates without Markov switching.

6.4.3 Optimal portfolio

Now we are ready to test the derived portfolio strategies. Using the parameters estimated over the period 2005-2010, we perform an out-of sample test over the period 2011-2015. For the estimate of the state of the Markov chain we use the filtered probabilities. So, at every time point, the investor uses only past data for her investment decision and assumes that the Markov chain will stay in the same state for the next short period. The investor trades in the DAX and in a German government bond with maturity 10 years at the beginning. Furthermore, she can borrow and lend money at the 1 week Euribor rate. The investment horizon is set to

Coefficient	Value	Standard error	Level of significance
$\beta(e_1)$	5.102	2.404	0.0348

Table 6.5: Results for the linear regression in (6.55) with Markov-switching parameters.

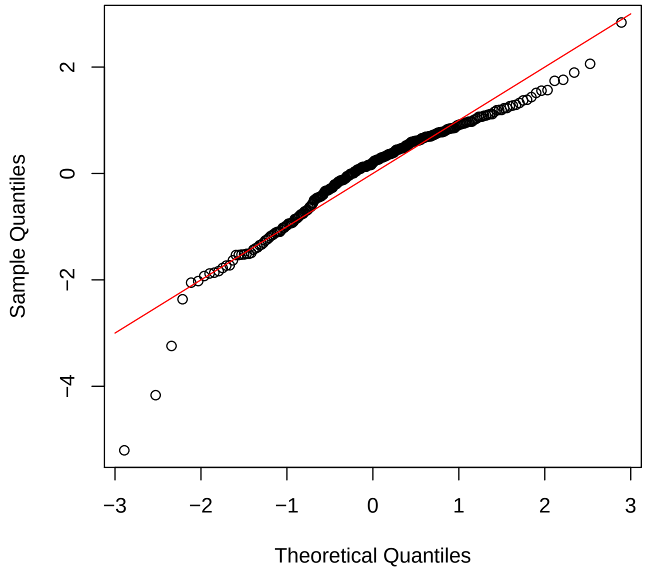


Figure 6.9: Normal QQ plot of the residuals of the linear regression in (6.55) with constant parameters (no Markov switching).

the last day in our time series, February, 20^{th} 2015. The portfolio is restructured on a weekly basis. The risk preferences of the investor are characterized by the HARA utility function (2.26) with the following parameters: $\delta = -10$, $\alpha = (1-\delta)^{\delta-1}$ (to assure comparability with the power utility function) and the floor F is set to 80% of the initial wealth. The resulting wealth path is plotted in Figure 6.11. It is compared to the following three cases:

1. the complete initial wealth is invested in the DAX;
2. the complete initial wealth is invested in the bond;
3. the investor estimates the model without considering Markov switching and follows the resulting optimal strategy.

It is clearly recognizable that the wealth process with the derived MS-strategy is most of the time above the paths in the other three cases and ends up with the highest wealth. The positions in the bond and in the stock are presented in Figure 6.12. Let us first consider the stock investment. One can recognize that by differentiating between two different states and adjusting her portfolio accordingly the investor can profit from the DAX in both states of the economy: she holds a long position in the good state and a short position in a crisis. In contrast, if she accounts only for one state she almost does not invest in the stock due to the averaging between the two states. The effect of the Markov switching is clearly visible also in the bond position: in times of a crisis the investor holds a big long position due to the falling interest rates and increasing bond prices.
Now let us compare the solution to the power utility functions by setting $F = 0$. The wealth path and the optimal portfolio are plotted in Figures 6.13 and 6.14,

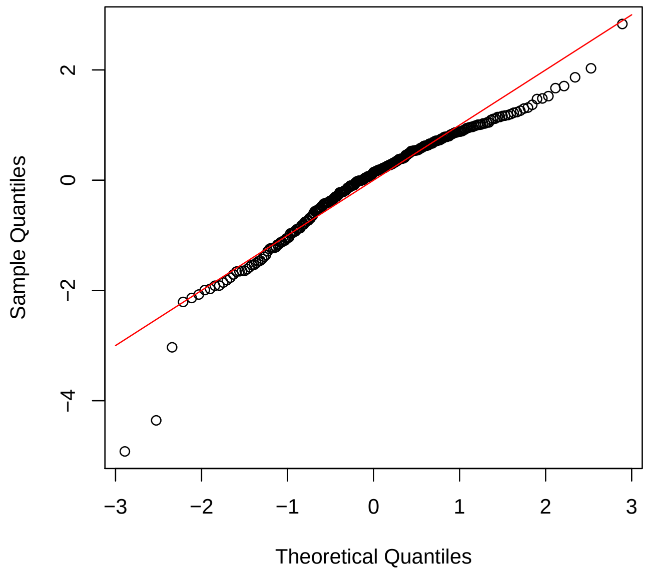


Figure 6.10: Normal QQ plot of the residuals of the linear regression in (6.55) with Markov-switching parameters.

respectively. One can recognize that the minimum of the wealth path is lower than in the previous case of $F = 80\%$ of the initial wealth. This is intuitively clear as in the present case the utility function does not imply a lower bound on the wealth. Thus, the investor is exposed to a higher risk on both the down and up side: the minimum of the wealth is lower, however the terminal wealth is higher than when we impose a floor restriction. The higher absolute value of the investment in the risky assets is in accordance with this interpretation.
We would like to deepen our analysis on the influence of F by studying not only one single path but the whole distribution of the terminal wealth. To this aim we simulate 1100000 paths according to Model (6.41) with the parameter values from Table 6.6 and apply the derived optimal strategy (6.46) for different values for the floor F and the risk aversion parameter δ. The resulting histograms are plotted in Figure 6.15. For comparison reasons we also include the histograms of the terminal wealth when the whole initial wealth is invested only in the stock, respectively only in the bond (see Figure 6.16). It is clearly recognizable that the assumption of a floor restriction reduces the probability on the high-end of the distribution, independently of the risk aversion parameter. This observation is confirmed by the cumulative distribution functions of the simulated terminal wealth shown in Figure 6.17. This is the price to pay for assuring a lower bound on the terminal wealth. On the other side, the floor restriction assures higher quantiles on the very low end. This can be seen in Table 6.7. Furthermore, compared to the stock-only portfolio, it is clear that the quantiles both on the low and high end are improved. The bond-only portfolio represents a very conservative investment with very high quantiles on the low end, however also very low quantiles on the high end.
Figure 6.15 and Table 6.7 illustrate also the influence of δ: the smaller δ, i.e. the more risk averse the investor, the lower the probability for low wealth levels on the

Process	Parameter	No Markov switching	Markov switching
X_1, P_1	κ_1	0.266	0.266
	$\tilde{\theta}_1$	0.047	0.047
	χ_1	0.013	0.013
	$\lambda_1(e_1)$	0.689	0.146
	$\lambda_1(e_2)$	0.689	6.674
X_2	κ_2	2.115	5.563
	$\hat{\theta}_2(e_1)$	0.051	0.027
	$\hat{\theta}_2(e_2)$	0.051	0.285
	$\hat{\chi}_2(e_1)$	0.505	0.309
	$\hat{\chi}_2(e_2)$	0.505	1.205
P_2	$\frac{\lambda_2}{\nu(e_1)}$	2.234	5.102
	$\frac{\lambda_2}{\nu(e_2)}$	2.234	-1.307
	a	-0.096	-0.096
	$\bar{\rho}$	-0.514	-0.454

Table 6.6: Estimated model parameter values for Model (6.41). The third column contains the values without Markov switching, whereas the Markov-switching values are listed in the fourth column.

one side, and high profits on the other.
Now we extend this comparison between the HARA and the power utility functions for different risk aversion parameters by various portfolio risk measures. They are either based on the distribution properties of the yearly log-return of the portfolio: $R^{\bar{\pi}}(T) := \frac{1}{T} \ln \frac{V^{\bar{\pi}}(T)}{V^{\bar{\pi}}(0)}$ or on the distribution of the terminal wealth $V^{\bar{\pi}}(T)$ itself. More precisely, we consider the following characteristics of the investment:

i) expected log-return (exp. return): $\mathbb{E}\left[R^{\bar{\pi}}(T)\right]$

ii) expected excess log-return (exp. ex. return): $\mathbb{E}\left[R^{\bar{\pi}}(T) - \frac{1}{T} \ln \frac{P_0(T)}{P_0(0)}\right]$

iii) standard deviation of the log-returns (st. dev.): $\sqrt{Var\left(R^{\bar{\pi}}(T)\right)}$

iv) skewness of the log-returns: $\frac{\mathbb{E}\left[\{R^{\bar{\pi}}(T) - \mathbb{E}[R^{\bar{\pi}}(T)]\}^3\right]}{\sqrt{Var(R^{\bar{\pi}}(T))}^3}$

v) kurtosis of the log-returns: $\frac{\mathbb{E}\left[\{R^{\bar{\pi}}(T) - \mathbb{E}[R^{\bar{\pi}}(T)]\}^4\right]}{Var(R^{\bar{\pi}}(T))^2}$

vi) Sharpe-ratio of the log-returns: $\frac{\mathbb{E}\left[R^{\bar{\pi}}(T) - \frac{1}{T} \ln \frac{P_0(T)}{P_0(0)}\right]}{\sqrt{Var(R^{\bar{\pi}}(T))}}$

vii) downside probability: $\mathbb{P}\left(R^{\bar{\pi}}(T) < \frac{1}{T} \ln \frac{P_0(T)}{P_0(0)}]\right)$

viii) Omega with the bank account as a benchmark (Omega): $\frac{\mathbb{E}\left[\max\left\{R^{\bar{\pi}}(T) - \frac{1}{T} \ln \frac{P_0(T)}{P_0(0)},\ 0\right\}\right]}{\mathbb{E}\left[\max\left\{\frac{1}{T} \ln \frac{P_0(T)}{P_0(0)} - R^{\bar{\pi}}(T),\ 0\right\}\right]}$,

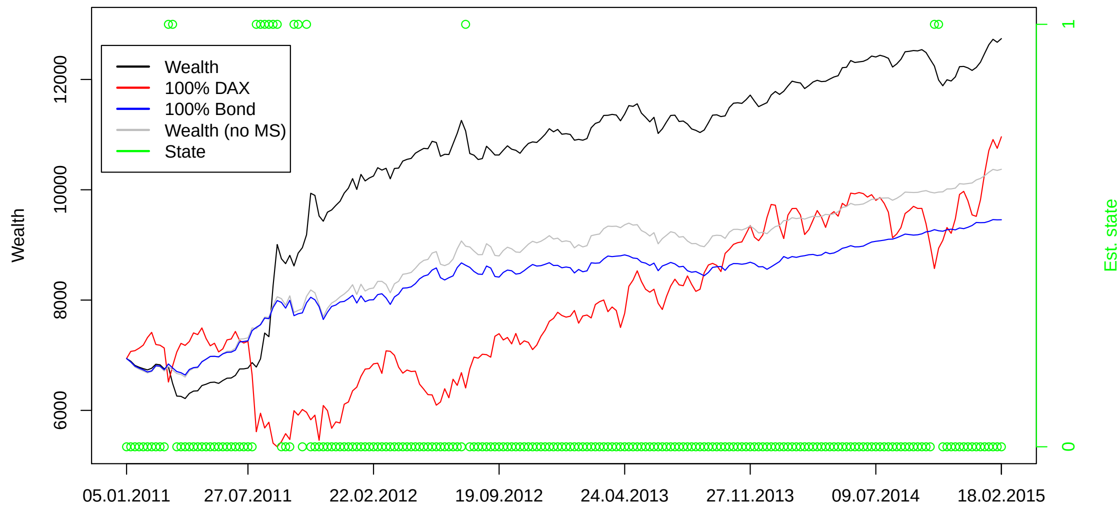


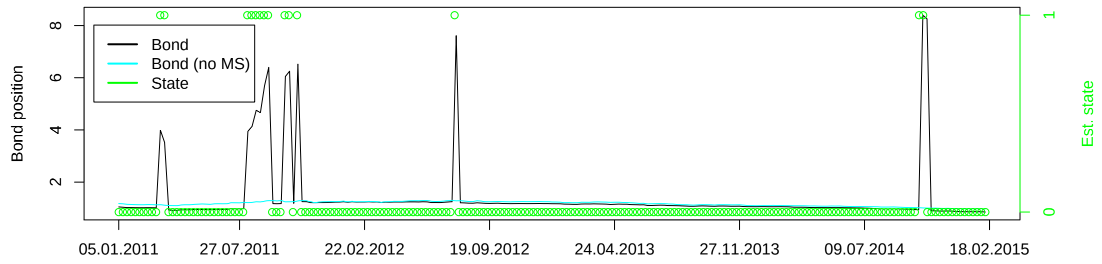
Figure 6.11: On the left y-axis are plotted the paths for the wealth process for the following portfolios: derived optimal Strategy (6.46) considering Markov switching (black line), 100% DAX (red line), 100% bond (blue line), derived optimal Strategy (6.46) without Markov switching (grey line). The green circles correspond to the estimated state of the Markov chain: "0"$=e_1$, "1"$=e_2$ on the right y-axis.

upside part: $\mathbb{E}\left[\max\left\{R^{\bar{\pi}}(T) - \frac{1}{T}\ln\frac{P_0(T)}{P_0(0)},\ 0\right\}\right]$,
downside part: $\mathbb{E}\left[\max\left\{\frac{1}{T}\ln\frac{P_0(T)}{P_0(0)} - R^{\bar{\pi}}(T),\ 0\right\}\right]$

ix) value at risk at the confidence level q (VaR_q):
$\inf\left\{v \in \mathbb{R} : \mathbb{P}\left(V^{\bar{\pi}}(0) - V^{\bar{\pi}}(T) > v\right) \leq 1-q\right\}$
$= \inf\left\{v \in \mathbb{R} : \mathbb{P}\left(V^{\bar{\pi}}(0) - V^{\bar{\pi}}(T) \leq v\right) \geq q\right\}$

x) conditional value at risk at the confidence level q (CVaR_q):
$\mathbb{E}\left[\left(V^{\bar{\pi}}(0) - V^{\bar{\pi}}(T)\right) | V^{\bar{\pi}}(0) - V^{\bar{\pi}}(T) > \text{VaR}_q\right]$.

The results are summarized in Table 6.8. It can be seen that $F > 0$ results in lower expected return however also in lower standard deviation of the returns. The lower bound on the terminal wealth improves basically the distribution of the terminal wealth on the low end. This is reflected in the lower VaR_q and CVaR_q-levels for high q. Furthermore, note that although the downside probability is bigger when we impose a positive floor F, the floor leads to an improvement in the downside of Omega. The price to pay for this improvement is the reduced upside part of Omega. For comparison, in Table 6.9 we also include the values of these risk-measures when investing only in the stock or only in the bond. It is confirmed once again that the bond represents a very conservative investment with very low expected return on the one side and very low standard deviation and value at risk on the other side. Compared to the most conservative of the considered optimal trading strategies (i.e. $\delta = -15$ and $F = 80\%$), it has lower standard deviation, however it is clearly outperformed in the excess return, the value at risk and the conditional value at



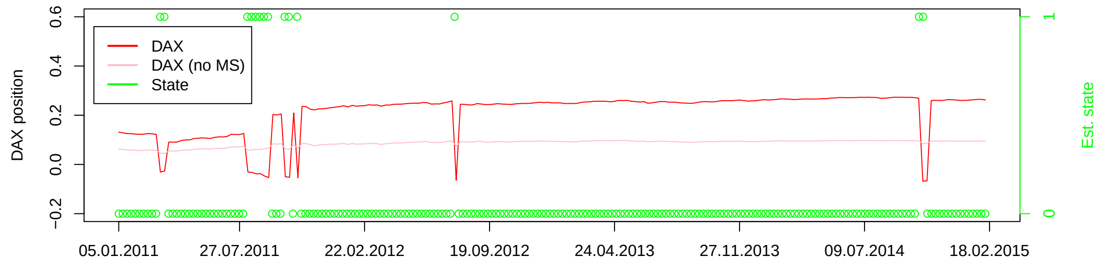


Figure 6.12: On the left y-axis are plotted the portfolio positions following the derived optimal Strategy (6.46). Upper plot: position in the bond with and without Markov switching (black and grey lines, respectively), lower plot: position in the DAX with and without Markov switching (red and pink lines, respectively). The green circles correspond to the estimated state of the Markov chain: "0"$=e_1$, "1"$=e_2$ on the right y-axis.

risk. The pure DAX investment is characterized by higher expected excess return and higher standard deviation compared to the bond investment. Compared to the most conservative of the considered optimal trading strategies (i.e. $\delta = -15$ and $F = 80\%$), it is outperformed both in the expected return, as well as in the value at risk at all considered levels and the conditional value at risk. Note that the higher expected return of the portfolio strategies in Table 6.8 is possible due to the fact that the investor does not have any short-selling or borrowing restrictions. So, in times of a crisis she can short-sell the stock and enter a very long bond position (possibly higher than her wealth), whereas in calm periods she can purchase big amounts of stocks financed by borrowing money from the bank.

Let us now go back to the case with $F = 80\%$ of the initial wealth. Although the portfolio profile from Figure 6.12 corresponds to the theoretical optimal trading strategy, such big shifts as in the bond position might not be desirable in reality due to transaction costs. While introducing transaction costs would exceed the scope of this study, we propose an ad-hoc method to overcome this difficulty by assuming that λ_1 is constant instead of Markov-switching. This might be reasonable also considering the difficulty in the estimation of λ_1 as mentioned above. We still allow for Markov switching in the stock price and volatility processes. Thus, Step 1 and Step 2 in the estimation procedure remain unchanged. For the parameters

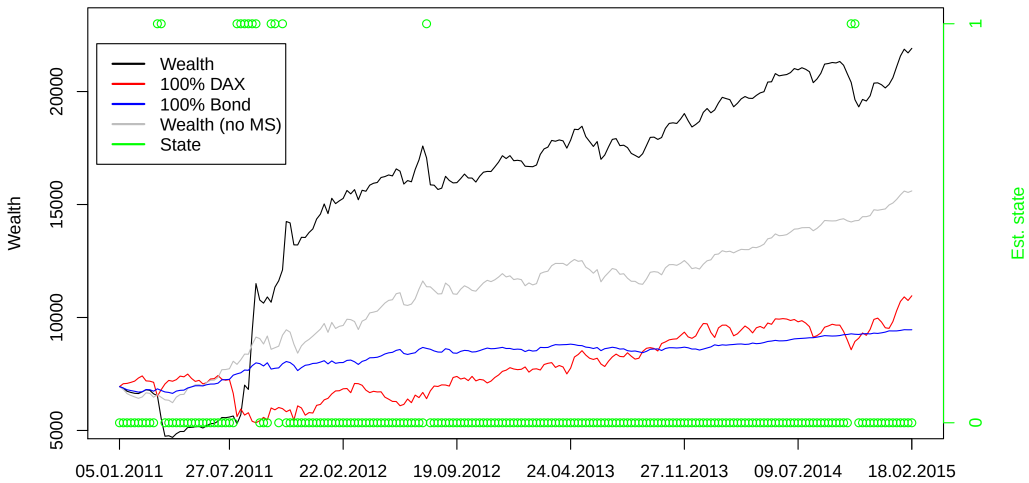
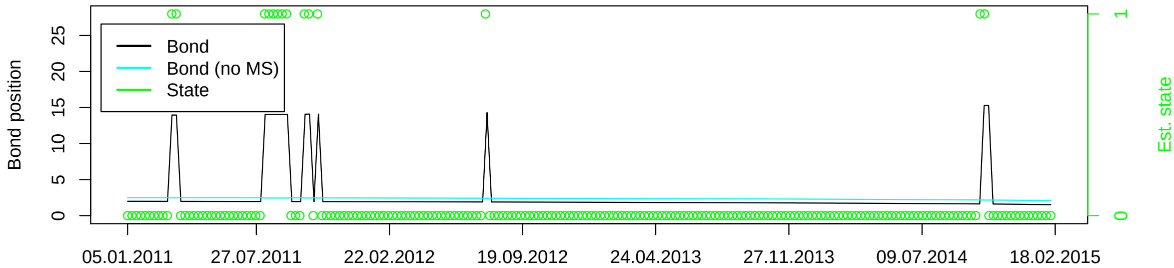


Figure 6.13: On the left y-axis are plotted the paths for the wealth process for the following portfolios: derived optimal Strategy (6.46) with $F = 0$ considering Markov switching (black line), 100% DAX (red line), 100% bond (blue line), derived optimal Strategy (6.46) without Markov switching (grey line). The green circles correspond to the estimated state of the Markov chain: "0"$=e_1$, "1"$=e_2$ on the right y-axis.

for X_1 we obtain the results without Markov switching: $\kappa_1 = 0.266$, $\tilde{\theta}_1 = 0.047$, $\chi_1 = 0.013$, $\lambda_1(e_1) = \lambda_1(e_2) = 0.689$ and in Step 4 we obtain: $\frac{\lambda_2}{\nu(e_1)} = 6.639$, $\frac{\lambda_2}{\nu(e_2)} = -1.701$ and $\bar{\rho} = -0.460$. The corresponding wealth path and portfolio development can be seen in Figure 6.18 and Figure 6.19, respectively. Although the terminal wealth in this case is lower than if we allow for a Markov-switching λ_1, it is still higher compared to the other three portfolios. Furthermore, note that the peaks in the bond position are much lower. They are mainly caused by the Markov switching in the stock process and the lower stock position for state e_2, as the bond is used to hedge the stock due to their negative correlation.

Finally, we would like to study the importance of the hedging term. In Figure 6.20 we plotted the ratio $\frac{\bar{\pi}_1^H}{\bar{\pi}_1^{MV}}$ for different values for κ_1 and χ_1 and Figure 6.21 shows $\frac{\bar{\pi}_2^H}{\bar{\pi}_2^{MV}}$ as a function of $\bar{\rho}$, κ_2 and χ_2. It can be seen, that independently of the state of the economy, the lower the mean-reversion speed and the higher the volatility coefficient for stochastic factors, the higher the influence of the hedging portfolio. Furthermore, the higher the absolute value of the leverage effect, the higher the relation $|\frac{\bar{\pi}_2^H}{\bar{\pi}_2^{MV}}|$. Intuitively this can be explained by the fact that for slower mean reversion and higher correlation the stochastic factor can be better hedged by the risky asset. What is more, the higher the volatility of the stochastic factor, the higher its risk, thus the bigger the hedging term. These results are in accordance with the findings in Sections 4.3.4 and 5.3.2.
We can conclude from the whole empirical and numerical study that the proposed model and investment strategies can be successfully applied in reality and that



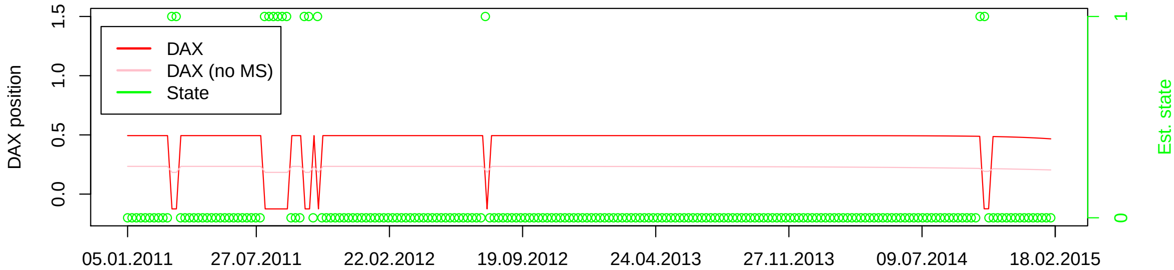


Figure 6.14: On the left y-axis are plotted the portfolio positions following the derived optimal Strategy (6.46) with $F = 0$. Upper plot: position in the bond with and without Markov switching (black and grey lines, respectively), lower plot: position in the DAX with and without Markov switching (red and pink lines, respectively). The green circles correspond to the estimated state of the Markov chain: "0"$=e_1$, "1"$=e_2$ on the right y-axis.

considering Markov switching, various stochastic factors and simultaneous investment in a bond and a stock is very beneficial for the investor and leads to a higher terminal wealth and thus, to higher utility.

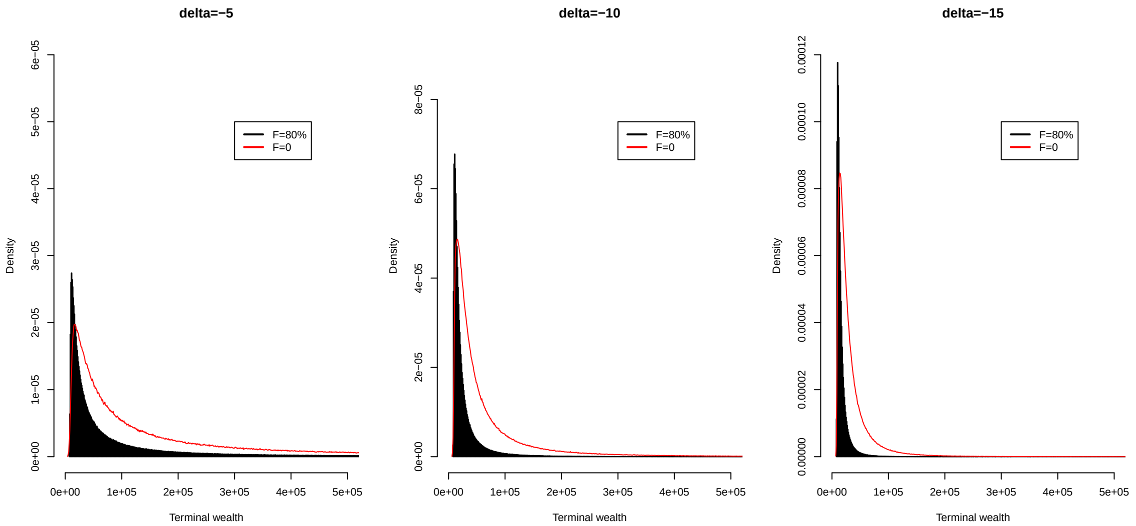


Figure 6.15: Numerically calculated distribution of the terminal wealth using the derived optimal strategy (6.46) with data simulated from Model (6.41) for the parameters from Table 6.6. Initial wealth and maturity are as in the estimation procedure. Three different values were used for δ: -5 (left plot), -10 (middle plot) and -15 (right plot). For F two cases were considered: $F = 80\%$ of the initial wealth (black bars) and $F = 0$ (red lines).

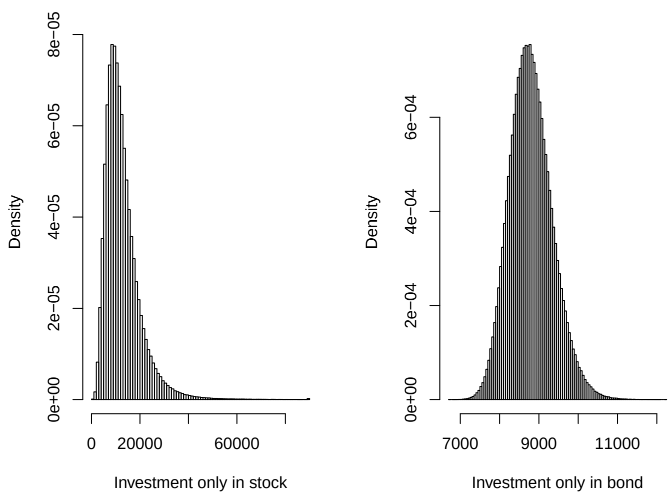


Figure 6.16: Numerically calculated distribution of the terminal wealth when investing only in the stock, resp. bond with data simulated from Model (6.41) for the parameters from Table 6.6. Initial wealth and maturity are as in the estimation procedure.

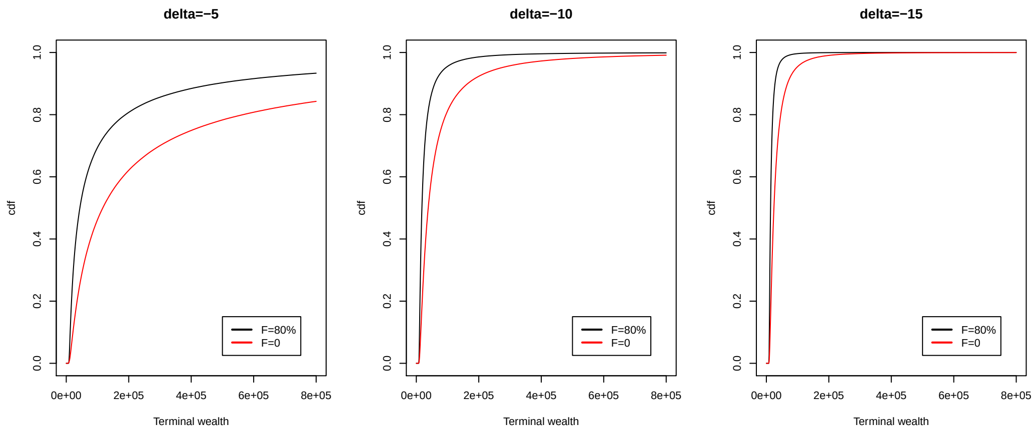


Figure 6.17: Numerically calculated cumulative distribution distribution function of the terminal wealth when using the derived optimal strategy (6.46) with data simulated from Model (6.41) for the parameters from Table 6.6. Initial wealth and maturity are as in the estimation procedure. Three different values were used for δ: -5 (left plot), -10 (middle plot) and -15 (right plot). For F two cases were considered: $F = 80\%$ of the initial wealth (black bars) and $F = 0$ (red lines). For reasons of better comparability, the x-axis is restricted to the interval $[0, 8 \times 10^5]$ for all three plots.

Parameters for the utility function	Quantiles for the terminal wealth 0.01%	0.1%	1%	95%	median
$\delta = -5, F = 0$	5328	7203	10392	3438166	116967
$\delta = -5, F = 80\%$	7283	7894	8931	1123855	43597
$\delta = -10, F = 0$	6642	7799	9521	268887	38159
$\delta = -10, F = 80\%$	7711	8088	8649	93029	17965
$\delta = -15, F = 0$	7141	7969	9140	95597	24204
$\delta = -15, F = 80\%$	7874	8143	8524	36656	13426
only stock	1036	1756	3005	26015	11229
only bond	7114	7359	7672	9749	8767

Table 6.7: Numerically calculated quantiles and median for the terminal wealth using the derived optimal strategy (6.46) with data simulated from Model (6.41) for the parameters from Table 6.6. Initial wealth and maturity are as in the estimation procedure. The quantiles are calculated for different values for δ and F. The last two lines contain the quantiles when the whole initial wealth is invested only in the stock, respectively only in the bond.

	$\delta = -5$		$\delta = -10$		$\delta = -15$	
	$F = 0$	$F = 80\%$	$F = 0$	$F = 80\%$	$F = 0$	$F = 80\%$
exp. return	0.7637	0.5406	0.4594	0.2821	0.3357	0.1917
exp. ex. return	0.7342	0.5112	0.4300	0.2526	0.3062	0.1622
st. dev.	0.4168	0.3705	0.2405	0.1834	0.1692	0.1127
skewness	0.8531	1.1956	0.8575	1.4218	0.8593	1.5087
kurtosis	3.8657	4.6843	3.8714	5.6821	3.8742	6.2502
Sharpe-ratio	1.7615	1.3798	1.7874	1.3776	1.8103	1.4395
downside prob.	0.0026	0.0030	0.0022	0.0037	0.0023	0.0051
Omega	8140	14072	10258	8696	10068	4744
upside part	0.7343	0.5112	0.4300	0.2526	0.3062	0.1622
downside part	9×10^{-5}	4×10^{-5}	4×10^{-5}	3×10^{-5}	3×10^{-5}	3×10^{-5}
$\mathrm{VaR}_{99\%}$	-3452	-1992	-2582	-1709	-2200	-1584
$\mathrm{VaR}_{99.99\%}$	1612	-343	297	-771	-201	-934
$\mathrm{VaR}_{99.999\%}$	2734	20	1123	-504	431	-729
$\mathrm{CVaR}_{99.999\%}$	3044	122	1320	-439	571	-683

Table 6.8: Numerically calculated risk measures for the terminal wealth using the derived optimal strategy (6.46) with data simulated from Model (6.41) for the parameters from Table 6.6. Initial wealth and maturity are as in the estimation procedure. The risk measures are calculated for different values for δ and F.

	only stock	only bond
exp. return	0.1172	0.0585
exp. ex. return	0.0878	0.0290
st. dev.	0.1313	0.0153
skewness	-0.0905	0.1938
kurtosis	3.4470	3.1288
Sharpe-ratio	0.6683	1.9016
downside prob.	0.2325	0.1292
Omega	5.8807	19.7336
upside part	0.1058	0.0306
downside part	0.0180	0.0015
$\mathrm{VaR}_{99\%}$	3935	-732
$\mathrm{VaR}_{99.99\%}$	5904	-174
$\mathrm{VaR}_{99.999\%}$	6253	12
$\mathrm{CVaR}_{99.999\%}$	6381	87

Table 6.9: Numerically calculated risk measures for the terminal wealth when investing only in the stock resp. only in the bonds with data simulated from Model (6.41) for the parameters from Table 6.6. Initial wealth and maturity are as in the estimation procedure.

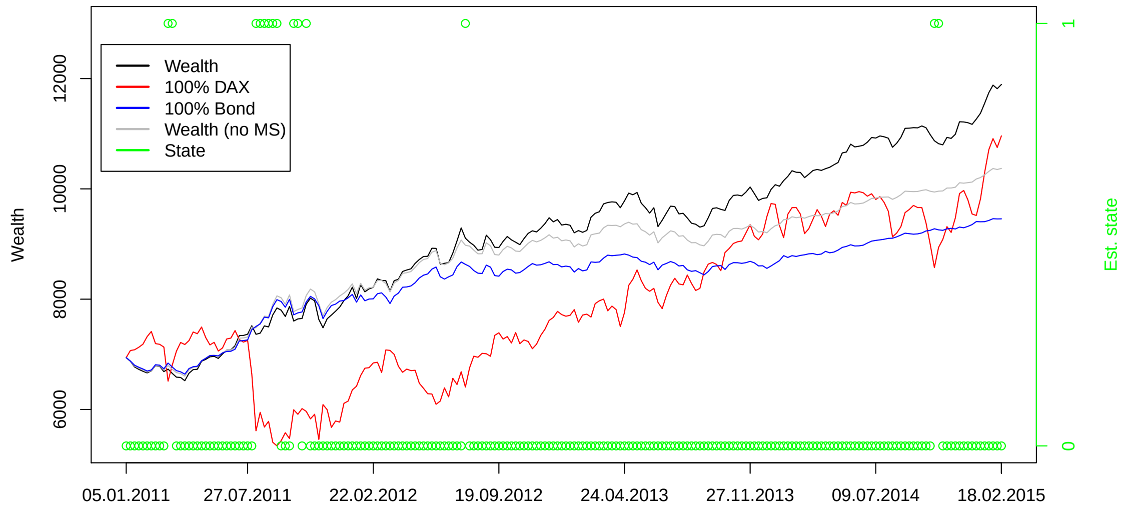


Figure 6.18: On the left y-axis are plotted the paths for the wealth process for the following portfolios: derived optimal Strategy (6.46) with a constant λ_1 but Markov-switching parameters for X_2 and P_2 (black line), 100% DAX (red line), 100% bond (blue line), derived optimal Strategy (6.46) without any Markov switching (grey line). The green circles correspond to the estimated state of the Markov chain: "0"$=e_1$, "1"$=e_2$ on the right y-axis.

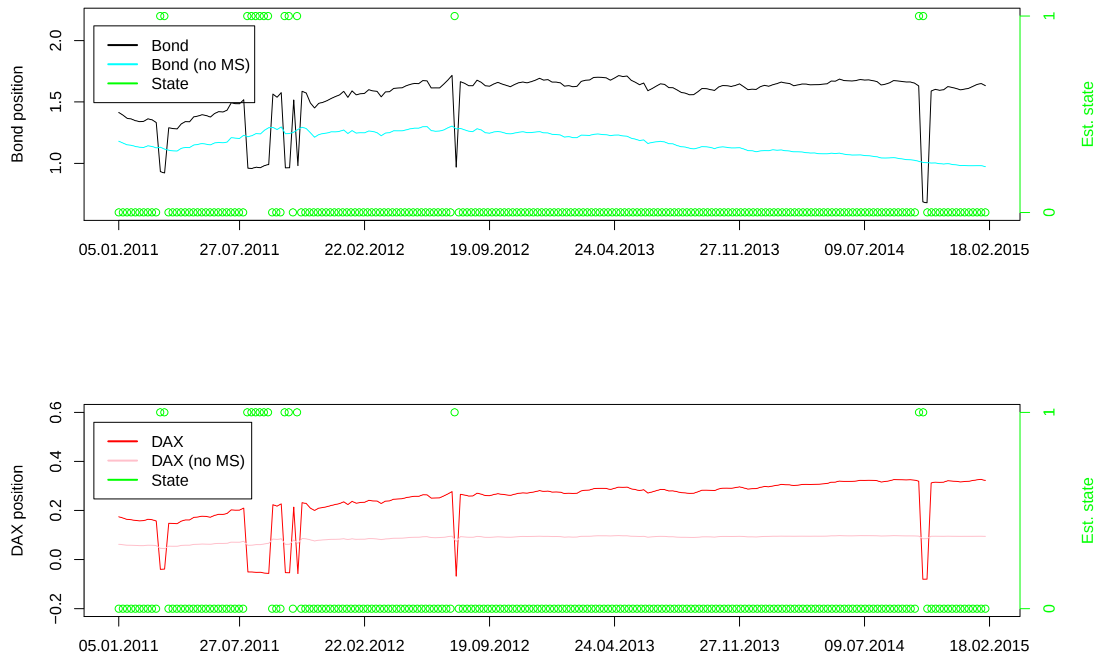


Figure 6.19: On the left y-axis are plotted the portfolio positions following the derived optimal Strategy (6.46). Upper plot: position in the bond with a constant λ_1 but Markov-switching parameters for X_2 and P_2 (black line) and without Markov switching (grey line), lower plot: position in the DAX with a constant λ_1 but Markov-switching parameters for X_2 and P_2 (red line) and without Markov switching (pink line). The green circles correspond to the estimated state of the Markov chain: "0"$=e_1$, "1"$=e_2$ on the right scale.

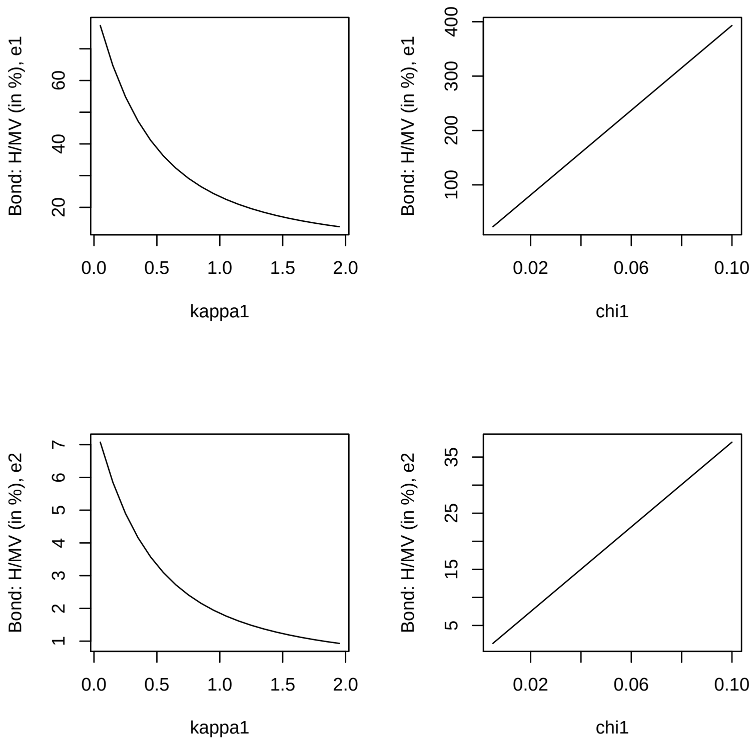


Figure 6.20: Ratio (in %) between the hedging term and the mean-variance portfolio invested in bond P_1 at the beginning of the investment period for an investor with $\delta = -10$ as a function of κ_1 (first column) and χ_1 (second column) for state e_1 (first row), resp. e_2 (second row).

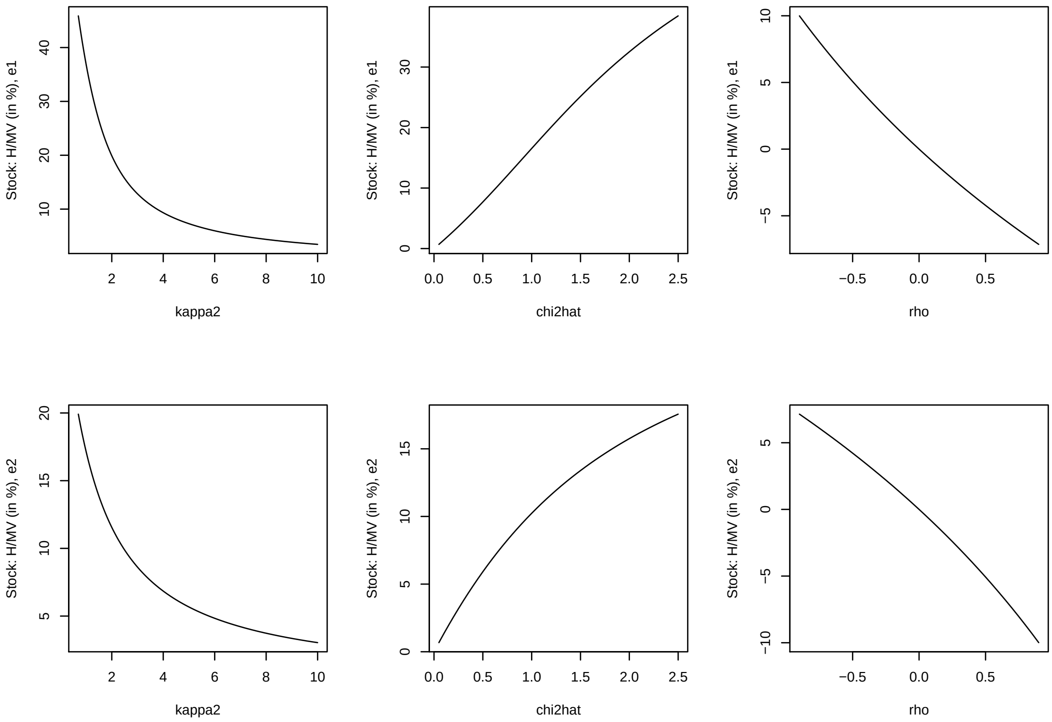


Figure 6.21: Ratio (in %) between the hedging term and the mean-variance portfolio invested in stock P_2 at the beginning of the investment period for an investor with $\delta = -10$ as a function of κ_2 (first column), $\hat{\chi}_2$ (second column), $\bar{\rho}$ (third column) for state e_1 (first row), resp. e_2 (second row).

Chapter 7

Conclusion

In the context of utility maximization we presented a flexible and simultaneously analytically tractable multidimensional framework that allows for incorporating various sources of risk, such as stochastic asset volatility and correlation, stochastic interest rate and macroeconomic regime changes. To the best of our knowledge, this is the most general affine model with Markov switching for which explicit solutions can be derived.

More precisely, in Chapter 3 we stated the corresponding HJB equations, relying on the semimartingale representation of the Markov chain. We discovered that adding a Markov chain does not change the decomposition of the optimal portfolio in a mean-variance part and a correction for the additional risk from the stochastic factor. Furthermore, we proved a verification theorem that reduced the case with Markov switching to a model with time-dependent coefficients (Theorem 3.5).

In Chapter 4 we found explicit solutions for the optimal control with the power utility function when trading in one risky asset. We derived the value function up to an expectation over the Markov chain (Corollary 4.12 and Theorem 4.13). We applied the results to the Heston model extended by Markov switching and stated besides the explicit optimal solution also very easy to check conditions for the verification result (Propositions 4.21 and 4.22). Thus, we extended the results from [8] by stochastic volatility and the results from [74] by Markov switching. Moreover, by various numerical computations we illustrated that the state of the Markov chain strongly influenced the optimal investment decision.

Chapter 5 presented the multidimensional case. We worked through the models with and without correlation between the Brownian motions driving the risky assets and the stochastic factors, and managed to place the dependence on the Markov chain in a suitable way to assure both the flexibility and the analytical tractability of the model. We derived explicit solutions (Corollary 5.7 and Theorem 5.10) and general verification results (Theorem 5.1 and Proposition 5.6). These results can be seen as an extension of [63] and [75] to Markov switching. The presented examples covered a bond-stock market and a two-asset market with stochastic volatility and stochastic correlation. For each of them we derived the solution in an explicit way and proved a verification theorem (Propositions 5.13 and 5.14). These models illustrated the

broad range of possible applications of the presented results.
In Chapter 6 we generalized the study to the HARA utility function (Theorem 6.4, Theorem 6.6, Proposition 6.8). We showed that the HARA utility function leads to a lower bound on the terminal wealth and a CPPI-type strategy with a Markov-switching stochastic multiplier (Theorem 6.5). Thus, we extended the results from [15] to Markov switching and further stochastic factors. Furthermore, we stated a special stock-bond model with stochastic volatility and stochastic interest rate and derived the optimal solution and a verification result (Proposition 6.9). This model can be seen as an extension of [26] that allows for stochastic volatility, stochastic interest rates and trading not only in a stock, but also in a bond. After estimating its parameters we provided some numerical computations and economic interpretations that showed the importance of considering various assets simultaneously and illustrated the influence of incorporating Markov switching parameters and different stochastic factors on the optimal investment strategy. An empirical study based on real data confirmed the practicability and good performance of the derived results.
To sum up, we contributed to literature by solving the optimal investment problem in a realistic model and analyzing the results from theoretical and practical point of view.

Appendix A

Appendix for Chapter 2

Proof of Lemma 2.43
Statement i): It is trivially calculated that:

$$\frac{\partial}{\partial \tau} B^{\alpha,\beta}(\tau) = \underbrace{\frac{2a^2 \exp\{-a\tau\}}{\chi^2 \big[1 - c\exp\{-a\tau\}\big]^2}}_{>0} c.$$

The statement follows directly.
Statement ii): Observe that:

$$\lim_{\tau \downarrow 0} B^{\alpha,\beta}(\tau) = \frac{-c(\kappa + a) + \kappa - a}{\chi^2(1-c)}.$$

Inserting the definition of c leads to the statement.
Statement iii): For the limit result consider that $a > 0$ and for the inequalities recall the definition of a.
Statement iv): First observe that:

$$1 - c = \frac{2a}{-\alpha\chi^2 + \kappa + a}, \tag{A.1}$$

and use that $\alpha < \frac{\kappa+a}{\chi^2}$ and $a > 0$ to conclude that $1 - c > 0$, i.e. $c < 1$. Now calculate:

$$\frac{\partial}{\partial \tau} A^{\alpha,\beta}(\tau) = \frac{\kappa\theta}{\chi^2} \frac{\kappa - a - \kappa c\exp(-a\tau) - ac\exp(-a\tau)}{1 - c\exp(-a\tau)}.$$

Further, insert $c = 1 - \frac{2a}{-\alpha\chi^2+\kappa+a}$ to obtain:

$$\begin{aligned}\frac{\partial}{\partial \tau} A^{\alpha,\beta}(\tau) =& \underbrace{\frac{\kappa\theta}{\chi^2} \frac{1}{1 - c\exp(-a\tau)}}_{>0} \Bigg[\underbrace{2a\exp(-a\tau)}_{>0} \underbrace{\left\{ \frac{\kappa + a}{-\alpha\chi^2 + \kappa + a} - 1 \right\}}_{=:d_1} \\ &+ \underbrace{\big\{1 - \exp(-a\tau)\big\}}_{>0} (\kappa - a) \Bigg].\end{aligned}$$

Now note that:

$$\frac{\kappa+a}{-\alpha\chi^2+\kappa+a}-1\begin{cases}<0, & \text{for } \alpha<0\\ =0, & \text{for } \alpha=0\\ >0, & \text{for } \alpha>0\end{cases}, \quad \kappa-a\begin{cases}<0, & \text{for } \beta<0\\ =0, & \text{for } \beta=0\\ >0, & \text{for } \beta>0\end{cases},$$

which yields the statement.

Statement v): First we rewrite function $A^{\alpha,\beta}(\tau)$ in a convenient way by inserting Equality (A.1):

$$\begin{aligned} A^{\alpha,\beta}(\tau) &= \frac{\kappa\theta}{\chi^2}\Big[(\kappa-a)\tau - 2\ln\Big\{\frac{1-c\exp(-a\tau)}{1-c}\Big\}\Big] \\ &= \frac{\kappa\theta}{\chi^2}\Big[\underbrace{(\kappa-a)}_{\in[0,\kappa]}\tau - 2\ln\Big\{\underbrace{\frac{1-\exp(-a\tau)}{2a}}_{\in[0,\frac{T}{2}]}\underbrace{\big(-\alpha\chi^2+\kappa+a\big)}_{\in[0,2\kappa]}+\underbrace{\exp(-a\tau)}_{\in[\exp(-\kappa T),1]}\Big\}\Big]. \end{aligned}$$

Now observe that $a\in[0,\kappa]$, as $\beta\geq 0$. Thus, $\kappa-a\in[0,\kappa]$ and $\exp(-a\tau)\in[\exp(-\kappa\tau),1]$. Further, $-\alpha\chi^2+\kappa+a\in[0,2\kappa]$, as $\alpha\geq 0$. Now consider the term $\frac{1-\exp(-a\tau)}{2a}$ and prove that it is monotonically decreasing in a by showing that its derivative w.r.t. a is negative:

$$\frac{\partial}{\partial a}\Big(\frac{1-\exp(-a\tau)}{2a}\Big)=\frac{\exp(-a\tau)(1+a\tau)-1}{2a^2},$$

where negativity follows by the general inequality $\exp(x)>1+x, \forall x\in\mathbb{R}$. So, for $a\in[0,\kappa]$, $\frac{1-\exp(-a\tau)}{2a}\in[\frac{1-\exp(-\kappa\tau)}{2\kappa},\lim_{a\downarrow 0}\frac{1-\exp(-a\tau)}{2a}]$. The limit is given by:

$$\lim_{a\downarrow 0}\frac{1-\exp(-a\tau)}{2a}=\lim_{a\downarrow 0}\frac{\exp(-a\tau)\tau}{2}=\frac{\tau}{2}.$$

As $\tau\in[0,T]$, we obtain: $\frac{1-\exp(-a\tau)}{2a}\in[0,\frac{T}{2}]$. Combining the inequalities from above leads to the statement.

Statement vi): In this proof we consider $B^{\alpha,\beta}(\tau)$ as a function in α and fix all other parameters. Computing the first two derivatives shows that $B^{\alpha,\beta}$ is a convex, monotonically increasing function of α:

$$\begin{aligned} \frac{\partial}{\partial\alpha}B^{\alpha,\beta}(\tau) &= \frac{4a^2\exp(-a\tau)}{(1-c\exp(-a\tau))^2(-\alpha\chi^2+\kappa+a)^2}\geq 0 \\ \frac{\partial^2}{\partial\alpha^2}B^{\alpha,\beta}(\tau) &= \frac{8a^2\chi^2\exp(-a\tau)(1-\exp(-a\tau))}{\underbrace{(1-c\exp(-a\tau))}_{>0}{}^3\underbrace{(-\alpha\chi^2+\kappa+a)}_{>0}{}^3}\geq 0. \end{aligned}$$

Further,

$$\lim_{\alpha\uparrow\frac{\kappa+a}{\chi^2}}B^{\alpha,\beta}(\tau)=\lim_{c\uparrow\infty}B^{\alpha,\beta}(\tau)=\frac{\kappa+a}{\chi^2}.$$

Now we would like to find the points where the graph of $B^{\alpha,\beta}$ crosses the graph of function $f(\alpha)=\alpha$. To this aim we solve the following equation:

$$B^{\alpha,\beta}(\tau)=\alpha \Leftrightarrow (\alpha\chi^2-\kappa+a)(\exp(-a\tau)-1)=0.$$

Now assume that $\tau\neq 0$ and $a\neq 0$ and observe that the only solution is given by $\alpha=\frac{\kappa-a}{\chi^2}$. What is more, in this case the first two derivatives are even strictly positive, which means that $B^{\alpha,\beta}$ is strictly monotonically increasing and strictly convex in α. Thus, its graph stays for $\alpha\in\left(\frac{\kappa-a}{\chi^2},\frac{\kappa+a}{\chi^2}\right)$ under the graph of the function $f(\alpha)=\alpha$, crosses it at $\alpha=\frac{\kappa-a}{\chi^2}$ and converges to it for $\alpha\uparrow\frac{\kappa+a}{\chi^2}$. This proves Statement vi) for $\tau\neq 0$ and $a\neq 0$.
Now assume $a>0$ and $\tau=0$. Then $B^{\alpha,\beta}(0)=\alpha$ and Statement vi) follows directly in this case.

□

Proof of Theorem 2.74
The proof follows the main idea from [86], p. 137, Theorem 8.2.1. Here we generalize the necessary assumptions and extend the statement to multidimensional Markov-modulated diffusions and time-dependence of K.
First note that because of the Markov property of $(X,\mathcal{MC})$ it holds for any $t,r\in[0,T]$:

$$\begin{aligned}k(t,x,e_i)&=\mathbb{E}\Big[\exp\Big\{-\int_0^t K\big(t-s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big]\\&=\mathbb{E}\Big[\exp\Big\{-\int_r^{t+r} K\big(t+r-s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}\Big|X(r)=x,\mathcal{MC}(r)=e_i\Big].\end{aligned}$$

Using the equation above and the tower rule for conditional expectations (see Proposition 2.5.1 in [12], p. 48) compute for any arbitrary but fixed $t\in[0,T]$ and $r<t$:

$$LS(r):=\frac{1}{r}\Big\{\mathbb{E}\Big[k\big(t,X(r),\mathcal{MC}(r)\big)\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big]-k(t,x,e_i)\Big\} \tag{A.2}$$

$$\begin{aligned}&=\frac{1}{r}\Big\{\mathbb{E}\Big[\underbrace{\exp\Big\{-\int_0^{t+r} K\big(t+r-s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}}_{=Z(t+r)}\\&\cdot\underbrace{\exp\Big\{\int_0^{r} K\big(t+r-s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}}_{=Y(r)}\\&-\underbrace{\exp\Big\{-\int_0^{t} K\big(t-s,X(s),\mathcal{MC}(s)\big)\mathrm{d}s\Big\}}_{=:Z(t)}\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big]\Big\}\\&=\frac{1}{r}\mathbb{E}\big[Z(t+r)-Z(t)\big|X(0)=x,\mathcal{MC}(0)=e_i\big]\\&+\frac{1}{r}\mathbb{E}\big[Z(t+r)\big(Y(r)-1\big)\big|X(0)=x,\mathcal{MC}(0)=e_i\big]\end{aligned}$$

$$= \frac{1}{r}\big(k(t+r,x,e_i) - k(t,x,e_i)\big)$$
$$+ \mathbb{E}\Big[Z(t+r)\Big(\frac{Y(r)-Y(0)}{r}\Big)\Big|X(0)=x, \mathcal{MC}(0)=i\Big]. \tag{A.3}$$

Now define $\hat{k}^{(t)}(x,e_i) := k(t,x,e_i)$ and apply the Itô's formula for Markov-modulated diffusions (Theorem 2.72) to $\hat{k}^{(t)}\big(X(r),\mathcal{MC}(r)\big)$ (considered as a process in r):

$$\hat{k}^{(t)}\big(X(r),\mathcal{MC}(r)\big) = \hat{k}^{(t)}\big(X(0),\mathcal{MC}(0)\big) + \int_0^r \Big[\hat{k}_x^{(t)}\big(X(s),\mathcal{MC}(s)\big)'\mu\big(X(s),\mathcal{MC}(s)\big)$$
$$+ \frac{1}{2}Tr\Big(\hat{k}_{xx'}^{(t)}\big(X(s),\mathcal{MC}(s)\big)\Sigma\big(X(s),\mathcal{MC}(s)\big)\Sigma\big(X(s),\mathcal{MC}(s)\big)'\Big)$$
$$+ \sum_{j=1}^{I} q_{MC(s),j}\hat{k}^{(t)}\big(X(s),j\big)\Big]\mathrm{d}s$$
$$+ \int_0^r \hat{k}_x^{(t)}\big(X(s),\mathcal{MC}(s)\big)'\Sigma\big(X(s),\mathcal{MC}(s)\big)\mathrm{d}W(s)$$
$$+ \int_0^r \Big(\hat{k}^{(t)}\big(X(s),e_1\big),\ldots,\hat{k}^{(t)}\big(X(s),e_I\big)\Big)\mathrm{d}M(s).$$

This allows us to rewrite the expression in (A.2) as follows:

$$LS(r) = \frac{1}{r}\Big\{\mathbb{E}\Big[k\big(t,X(r),\mathcal{MC}(r)\big)\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big] - k(t,x,e_i)\Big\}$$
$$= \mathbb{E}\Big[\frac{1}{r}\int_0^r \hat{k}_x^{(t)}\big(X(s),\mathcal{MC}(s)\big)'\mu\big(X(s),\mathcal{MC}(s)\big)$$
$$+ \frac{1}{2}Tr\Big(\hat{k}_{xx'}^{(t)}\big(X(s),\mathcal{MC}(s)\big)\Sigma\big(X(s),\mathcal{MC}(s)\big)\Sigma\big(X(s),\mathcal{MC}(s)\big)'\Big)$$
$$+ \sum_{j=1}^{I} q_{MC(s),j}\hat{k}^{(t)}\big(X(s),e_j\big)\big)\mathrm{d}s\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big]$$
$$+ \frac{1}{r}\mathbb{E}\Big[\underbrace{\int_0^r \hat{k}_x^{(t)}\big(X(s),\mathcal{MC}(s)\big)'\Sigma\big(X(s),\mathcal{MC}(s)\big)\mathrm{d}W(s)}_{=:M_1(r)}\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big]$$
$$+ \frac{1}{r}\mathbb{E}\Big[\underbrace{\int_0^r \Big(\hat{k}^{(t)}\big(X(s),e_1\big),\ldots,\hat{k}^{(t)}\big(X(s),e_l\big)\Big)\mathrm{d}M(s)}_{=:M_2(r)}\Big|X(0)=x,\mathcal{MC}(0)=e_i\Big]$$
$$=: \mathbb{E}\big[N(r)\big|X(0)=x,\mathcal{MC}(0)=e_i\big] + \frac{1}{r}\mathbb{E}\big[M_1(r)\big|X(0)=x,\mathcal{MC}(0)=e_i\big]$$
$$+ \frac{1}{r}\mathbb{E}\big[M_2(r)\big|X(0)=x,\mathcal{MC}(0)=e_i\big].$$

Observe the following $\mathbb{P}$.a.s. (point-wise) limit for all $i=1,\ldots,I$:

$$a.s.\lim_{r\downarrow 0} N(r)|\{X(0)=x,\mathcal{MC}(0)=e_i\} = k_x(t,x,e_i)'\mu(x,e_i)$$
$$+ \frac{1}{2}Tr\big(k_{xx'}(t,x,e_i)\Sigma(x,e_i)\Sigma(x,e_i)'\big) + \sum_{j=1}^{I} q_{i,j}k(t,x,e_j).$$

Together with Assumptions ii), iii) and iv) we obtain:

$$\lim_{r\downarrow 0} LS(r) = \lim_{r\downarrow 0} \mathbb{E}[N(r)|X(0)=x, \mathcal{MC}(0)=e_i] = k_x(t,x,e_i)'\mu(x,e_i) + \frac{1}{2}Tr\big(k_{xx'}(t,x,e_i)\Sigma(x,e_i)\Sigma(x,e_i)'\big) + \sum_{j=1}^{I} q_{i,j}k(t,x,j). \tag{A.4}$$

Further, Assumption v) leads to:

$$\lim_{r\downarrow 0} \mathbb{E}\big[Z(t+r)\frac{Y(r)-Y(0)}{r}\big|X(0)=x, \mathcal{MC}(0)=e_i\big] = K(t,x,e_i)\mathbb{E}\big[Z(t)\big|X(0)=x, \mathcal{MC}(0)=e_i\big] = K(t,x,e_i)k(t,x,e_i). \tag{A.5}$$

Combining Equations (A.4), (A.5) and (A.3) gives us:

$$\lim_{r\downarrow 0} \frac{1}{r}\big(k(t+r,x,e_i) - k(t,x,e_i)\big) = k_x(t,x,e_i)'\mu(x,e_i) + \frac{1}{2}Tr\big(k_{xx'}(t,x,e_i)\Sigma(x,e_i)\Sigma(x,e_i)'\big) + \sum_{j=1}^{I} q_{i,j}k(t,x,e_j) - K(t,x,e_i)k(t,x,e_i).$$

By considering $\frac{1}{r}\Big\{\mathbb{E}\Big[k\big(t-r, X(r), \mathcal{MC}(r)\big)\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] - k(t-r,x,e_i)\Big\}$ we obtain analogously:

$$\lim_{r\downarrow 0} \frac{1}{r}\big(k(t,x,e_i) - k(t-r,x,e_i)\big) = k_x(t,x,e_i)'\mu(x,e_i) + \frac{1}{2}Tr\big(k_{xx'}(t,x,e_i)\Sigma(x,e_i)\Sigma(x,e_i)'\big) + \sum_{j=1}^{I} q_{i,j}k(t,x,e_j) - K(t,x,e_i)k(t,x,e_i).$$

Thus, k is continuously differentiable in t for all $(x,e_i) \in D^X \times \mathcal{E}$ and fulfills System (2.23).
Finally, we show that iii)'⇒iii), iv)'⇒iv) and v)'⇒v). For the first statement apply directly Theorem 2.31 to follow that M_1 is a square integrable martingale, thus

$$\mathbb{E}\big[M_1(r)\big|X(0)=x, \mathcal{MC}(0)=e_i\big] = \mathbb{E}\big[M_1(0)\big|X(0)=x, \mathcal{MC}(0)=e_i\big] = 0.$$

For the second statement consider:

$$\int_0^r \Big(k\big(t, X(s), e_1\big), \ldots, k\big(t, X(s), e_I\big)\Big)\mathrm{d}M(s) = \sum_{j=1}^{I} \int_0^r k\big(t, X(s), e_j\big)\mathrm{d}M_j(s),$$

and by Theorem 2.69:

$$\mathbb{E}\Big[\int_0^r \big(k(t,X(s),e_j)\big)^2 \mathrm{d}\langle M_j, M_j\rangle(s)\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big] = \mathbb{E}\Big[\int_0^r \big(k(t,X(s),e_j)\big)^2 v_{j,j}(s)\mathrm{d}s\Big|X(0)=x, \mathcal{MC}(0)=e_i\Big].$$

As v_{ii} is bounded for all $i = 1, \dots, I$, condition iv)' implies again by Lemma 2.31 that M_2 is a martingale. So,

$$\mathbb{E}\big[M_2(r)\big|X(0) = x, \mathcal{MC}(0) = e_i\big] = \mathbb{E}\big[M_2(0)\big|X(0) = x, \mathcal{MC}(0) = e_i\big] = 0.$$

For the last statement we just need to show the a.s. convergence of $Z(t+r)\frac{Y(r)-Y(0)}{r}$ for $r \downarrow 0$. To this aim observe that for any arbitrary but fixed path of X and $\mathcal{MC}$ one can consider Y as a function of r. Furthermore, for any fixed path of $\mathcal{MC}$ there exists a positive number $\varepsilon > 0$ such that $\mathcal{MC}$ remains constant on $[0, \varepsilon]$. So, we can apply the generalized Leibniz integral rule (summarized for convenience in Lemma A.1 after the end of this proof) to obtain the derivative Y_r:

$$Y_r(r) = Y(r)\Big[\int_0^r K_t\big(t + r - s, X(s), \mathcal{MC}(s)\big)\mathrm{d}s + K\big(t + r - r, X(r), \mathcal{MC}(r)\big)\Big].$$

Thus, we get:

$$a.s. \lim_{r\downarrow 0} Z(t + r)\frac{Y(r) - Y(0)}{r} = Z(t)Y_r(0) = Z(t)K\big(t, X(0), MC(0)\big).$$

□

For the convenience of the reader we recall the Generalized Leibniz integral rule in the following lemma:

Lemma A.1 (Generalized Leibniz integral rule))
Consider a real-valued function $f : O \times [a, b] \to \mathbb{R}$, $(x, t) \mapsto f(x, t)$ for an open set $O \subset \mathbb{R}$ and a compact interval $[a, b] \subset \mathbb{R}$. Assume that the following conditions hold:

i) For all $x \in O$, the map $t \mapsto f(x, t)$ is continuous.

ii) For all $t \in [a, b]$, the map $x \mapsto f(x, t)$ is differentiable.

iii) The function $(x, t) \mapsto f_x(x, t)$ is continuous on $O \times [a, b]$.

Let $\tilde{a} : [\underline{x}, \overline{x}] \to \mathbb{R}$ and $\tilde{b} : [\underline{x}, \overline{x}] \to \mathbb{R}$ be two real-valued continuously differentiable functions. Define the following set:

$$D^{\tilde{a},\tilde{b}} = \{(x, t) \in \mathbb{R}^2 | x \in [\underline{x}, \overline{x}], t \in [\tilde{a}(\underline{x}), \tilde{b}(\overline{x})]\}.$$

Further, assume that $D^{\tilde{a},\tilde{b}} \subset O \times [a, b]$ and define the following function:

$$\tilde{F}(x) := \int_{\tilde{a}(x)}^{\tilde{b}(x)} f(x, t)\mathrm{d}t.$$

Then, $\tilde{F}(x)$ is differentiable w.r.t. x and has the following derivative:

$$\tilde{F}_x(x) = \int_{\tilde{a}(x)}^{\tilde{b}(x)} f_x(x, t)\mathrm{d}t + f\big(x, \tilde{b}(x)\big)\tilde{b}_x(x) - f\big(x, \tilde{a}(x)\big)\tilde{a}_x(x).$$

Proof
Follows directly from the fundamental theorem of calculus (see [69], p. 200) and the Leibniz integral rule (see [70], p. 75.)

Appendix B

Appendix for Chapter 4

Proof of Remark 4.5
Let t be an arbitrary time point in $[0,T)$, where $t \in [t_{\tilde{K}}, t_{\tilde{K}+1})$ for some $\tilde{K} \in 0,\dots,K$. Start by applying Itô's formula step-wise to $\Phi^m\big(T, V^{m,\pi}(T), X^m(T)\big)$:

$$
\begin{aligned}
&\Phi^m\big(T, V^{m,\pi}(T), X^m(T)\big) = \Phi^m\big(t_K, V^{m,\pi}(t_K), X^m(t_K)\big)\\
&+ \int_{t_K}^{T} \mathcal{L}^m(m_K,\pi)\Phi^m\big(s, V^{m,\pi}(s), X^m(s)\big)\mathrm{d}s\\
&+ \int_{t_K}^{T} \Phi^m_v\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^V\big(V^{m,\pi}(s), X^m(s), m(s), \pi(s)\big)\mathrm{d}W^P(s)\\
&+ \int_{t_K}^{T} \Phi^m_x\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^X\big(X^m(s), m(s)\big)\mathrm{d}W^X(s)\\
&= \Phi^m\big(t_{K-1}, V^{m,\pi}(t_{K-1}), X^m(t_{K-1})\big)\\
&+ \sum_{k=K-1}^{K} \int_{t_k}^{t_{k+1}} \mathcal{L}^m(m_i,\pi)\Phi^m\big(s, V^{m,\pi}(s), X^m(s)\big)\mathrm{d}s\\
&+ \int_{t_{K-1}}^{T} \Phi^m_v\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^V\big(V^{m,\pi}(s), X^m(s), m(s), \pi(s)\big)\mathrm{d}W^P(s)\\
&+ \int_{t_{K-1}}^{T} \Phi^m_x\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^X\big(X^m(s), m(s)\big)\mathrm{d}W^X(s)\\
&= \dots\\
&= \Phi^m\big(t, V^{m,\pi}(t), X^m(t)\big) + \sum_{k=\tilde{K}+1}^{K} \int_{t_k}^{t_{k+1}} \underbrace{\mathcal{L}^m(m_k,\pi)\Phi^m\big(s, V^{m,\pi}(s), X^m(s)\big)}_{\leq 0}\mathrm{d}s\\
&+ \int_{t}^{t_{\tilde{K}+1}} \underbrace{\mathcal{L}^m(m_{\tilde{K}},\pi)\Phi^m\big(s, V^{m,\pi}(s), X^m(s)\big)}_{\leq 0}\mathrm{d}s\\
&+ \int_{t}^{T} \Phi^m_v\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^V\big(V^{m,\pi}(s), X^m(s), m(s), \pi(s)\big)\mathrm{d}W^P(s)
\end{aligned}
$$

$$
\begin{aligned}
&+ \int_t^T \Phi_x^m\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^X\big(X^m(s), m(s)\big)\mathrm{d}W^X(s) \\
&\leq \Phi^m\big(t, V^{m,\pi}(t), X^m(t)\big) \\
&+ \int_t^T \Phi_v^m\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^V\big(V^{m,\pi}(s), X^m(s), m(s), \pi(s)\big)\mathrm{d}W^P(s) \\
&+ \int_t^T \Phi_x^m\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^X\big(X^m(s), m(s)\big)\mathrm{d}W^X(s) =: Y^m(T).
\end{aligned}
$$

Note that here we have used the continuity and the piece-wise differentiability of function Φ^m. One can show the statement analogously for an arbitrary end-point $\tau \in [t, T]$:

$$
\begin{aligned}
&\Phi^m\big(\tau, V^{m,\pi}(\tau), X^m(\tau)\big) \leq \Phi^m\big(t, V^{m,\pi}(t), X^m(t)\big) \\
&+ \int_t^\tau \Phi_v^m\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^V\big(V^{m,\pi}(s), X^m(s), m(s), \pi(s)\big)\mathrm{d}W^P(s) \\
&+ \int_t^\tau \Phi_x^m\big(s, V^{m,\pi}(s), X^m(s)\big)\Sigma^X\big(X^m(s), m(s)\big)\mathrm{d}W^X(s) =: Y^m(\tau).
\end{aligned}
\tag{B.1}
$$

So, by definition:

$$
Y^m(t) = \Phi^m\big(t, V^{m,\pi}(t), X^m(t)\big).
$$

As $\Phi^m\big(\tau, v, x\big) \geq 0$ for all $(\tau, v, x) \in [0, T] \times [0, \infty) \times D^X$, it follows that $Y^m(\tau) \geq 0$. Furthermore, Y^m is a local martingale, as it has a zero drift and all involved functions are at least piece-wise continuous, so it is a supermartingale. Then it holds that:

$$
\begin{aligned}
&\mathbb{E}\Big[U_P\big(V^{m,\pi}(T)\big)\Big|\mathscr{F}_t\Big] = \mathbb{E}\Big[\frac{\big(V^{m,\pi}(T)\big)^\delta}{\delta}\Big|\mathscr{F}_t\Big] = \mathbb{E}\Big[\Phi^m\big(T, V^{m,\pi}(T), X^m(T)\big)\Big|\mathscr{F}_t\Big] \\
&\leq \mathbb{E}[Y^m(T)|\mathscr{F}_t] \leq Y^m(t) = \Phi^m\big(t, V^{m,\pi}(t), X^m(t)\big),
\end{aligned}
$$

which proves the statement.

□

Proof of Remark 4.6

First assume iv)'. Going through the same calculations as in the proof of Remark 4.5 for $\bar{\pi}^m$ instead of π we obtain " = " instead if " ≤ ":

$$
\begin{aligned}
&\Phi^m\big(\tau, V^{m,\bar{\pi}^m}(\tau), X^m(\tau)\big) = \Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big) \\
&+ \int_t^\tau \Phi_v^m\big(s, V^{m,\bar{\pi}^m}(s), X^m(s)\big)\Sigma_V\big(V^{m,\bar{\pi}^m}(s), X^m(s), m(s), \bar{\pi}^m(s)\big)\mathrm{d}W^P(s) \\
&+ \int_t^\tau \Phi_x^m\big(s, V^{m,\bar{\pi}^m}(s), X^m(s)\big)\Sigma_X\big(X^m(s), m(s)\big)\mathrm{d}W^X(s),
\end{aligned}
\tag{B.2}
$$

for all $0 \leq t \leq \tau \leq T$. So, process $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ is a local martingale. Condition iv)' assures by Theorem 2.31 that $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[t_k,t_{k+1}]}$

is a martingale for all $k \in \{0,\dots,K\}$. As function Φ^m is continuous, we can conclude that $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ is a martingale.
Now assume iv)" and consider an arbitrary but fixed index $k \in \{0,1,\dots,K\}$. Define θ_n as the exit time of $\{\big(t, V^{m,\pi}(t), X^m(t)\big)\}_{t\in[t_k,t_{k+1}]}$ out of the region $[t_k, t_{k+1}] \times [0,n] \times [-n,n]$. Then $\lim_{n\to\infty} \theta_n = t_{k+1}$ and

$$\mathbb{E}\Big[\int_{t_k}^{\theta_n} \Big(\Phi_v^m\big(s, V^{m,\bar{\pi}^m}(s), X^m(s)\big)\Sigma_V\big(V^{m,\bar{\pi}^m}(s), X^m(s), m(s), \bar{\pi}^m(s)\big)\Big)^2 \\ + \Big(\Phi_x^m\big(s, V^{m,\bar{\pi}^m}(s), X^m(s)\big)\Sigma_X\big(X^m(s), m(s)\big)\Big)^2 \mathrm{d}s\Big|\mathscr{F}_{t_k}\Big] < \infty,$$

as all involved functions are continuous in (s,v,x). Together with Equation (B.2) it follows that:

$$\mathbb{E}\big[\Phi^m\big(\theta_n, V^{m,\bar{\pi}^m}(\theta_n), X^m(\theta_n)\big)\big|\mathscr{F}_{t_k}\big] = \Phi^m\big(t_k, V^{m,\bar{\pi}^m}(t_k), X^m(t_k)\big).$$

Condition iv)" implies that:

$$\mathbb{E}\big[\Phi^m\big(t_{k+1}, V^{m,\bar{\pi}^m}(t_{k+1}), X^m(t_{k+1})\big)\big|\mathscr{F}_{t_k}\big] \\ = \lim_{n\to\infty} \mathbb{E}\big[\Phi^m\big(\theta_n, V^{m,\bar{\pi}^m}(\theta_n), X^m(\theta_n)\big)\big|\mathscr{F}_{t_k}\big] = \Phi^m\big(t_k, V^{m,\bar{\pi}^m}(t_k), X^m(t_k)\big).$$

Again, due to the continuity of Φ^m it follows that $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ is a martingale.

□

Proof of Proposition 4.9
As function Φ^m given by Equation (4.23) is obviously continuous and piece-wise sufficiently differentiable, by Proposition 4.3 we only need to show that $\big\{\Phi^m(t)\}_{t\in[0,T]} := \big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\}_{t\in[0,T]}$ is a martingale. We start by writing down the solution of the SDE for $V^{m,\bar{\pi}^m}$:

$$V^{m,\bar{\pi}^m}(t) = v_0 \exp\Big\{\int_0^t r\big(m(s)\big) + \big(\mu_1\big(X^m(s), m(s)\big) - r\big(m(s)\big)\big)\,\bar{\pi}^m(s) \\ - \frac{1}{2}\big(\bar{\pi}^m(s)\big)^2\Big(\Sigma_1\big(X^m(s), m(s)\big)\Big)^2 \mathrm{d}s + \int_0^t \bar{\pi}^m(s)\Sigma_1\big(X^m(s), m(s)\big)\mathrm{d}W^P(s)\Big\},$$

for all $0 \leq t \leq T$, where $v_0 := V^{m,\bar{\pi}^m}(0)$. Then we insert it in the expression for Φ^m:

$$
\begin{aligned}
&\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big) \\
&= \frac{\big(V^{m,\bar{\pi}^m}(t)\big)^\delta}{\delta} \exp\Big\{\int_t^T \delta r\big(m(s)\big)\mathrm{d}s\Big\} \exp\left\{A^m(t) + B^m(t)X^m(t)\right\} \\
&= \underbrace{\frac{v_0^\delta}{\delta} \exp\Big\{\int_0^T \delta r\big(m(s)\big)\mathrm{d}s\Big\} \exp\{A^m(0) + B^m(0)x_0\}}_{=\Phi^m(0,v_0,x_0)} \\
&\cdot \exp\Big\{\int_0^t \delta\left(\mu_1\big(X^m(s), m(s)\big) - r\big(m(s)\big)\right)\bar{\pi}^m(s) - \frac{1}{2}\delta\big(\bar{\pi}^m(s)\big)^2\Big(\Sigma_1\big(X^m(s), m(s)\big)\Big)^2 \\
&+ A_t^m(s) + B_t^m(s)X^m(s) + B^m(s)\mu^X\left(X^m(s), m(s)\right)\mathrm{d}s \\
&+ \int_0^t \underbrace{B^m(s)\Sigma^X\left(X^m(s), m(s)\right)}_{=:\Sigma_1^G(s,X^m(s))} \mathrm{d}W^X(s)\Big\} + \int_0^t \underbrace{\delta\bar{\pi}^m(s)\Sigma_1\big(X^m(s), m(s)\big)}_{=:\Sigma_2^G(s,X^m(s))} \mathrm{d}W^P(s) \\
&=: \Phi^m(0, v_0, x_0) \exp\Big\{\int_0^t \mu^G\big(s, X^m(s)\big)\mathrm{d}s + \int_0^t \Sigma_1^G\big(s, X^m(s)\big)\mathrm{d}W^X(s) \\
&+ \int_0^t \Sigma_2^G\big(s, X^m(s)\big)\mathrm{d}W^P(s)\Big\} =: \Phi^m(0, v_0, x_0) \exp\{G(t)\}.
\end{aligned}
$$

Now we recognize easily the differential semimartingale characteristics $\mu^Z(t) = \begin{pmatrix} \mu_1^Z\big(t, X^m(t)\big) \\ \mu_2^Z\big(t, X^m(t)\big) \end{pmatrix}$ and $\Gamma^Z(t) = \begin{pmatrix} \Gamma_{1,1}^Z\big(t, X^m(t)\big) & \Gamma_{1,2}^Z\big(t, X^m(t)\big) \\ \Gamma_{2,1}^Z\big(t, X^m(t)\big) & \Gamma_{2,2}^Z\big(t, X^m(t)\big) \end{pmatrix}$ of the two-dimensional process $Z := (X^m, G)'$:

$$
\begin{aligned}
\mu_1^Z\big(t, x\big) =& \mu^X(x, m(t)) = \bar{k}^{(0)}\big(m(t)\big) - xK^{(1)}\big(m(t)\big) \\
\mu_2^Z\big(t, x\big) =& \mu^G(t, x) \\
=& \delta\left(\mu_1\big(x, m(t)\big) - r\big(m(t)\big)\right)\bar{\pi}^m(t) - \frac{1}{2}\delta\big(\bar{\pi}^m(t)\big)^2\Big(\Sigma_1\big(x, m(t)\big)\Big)^2 \\
&+ A_t^m + B_t^m x + B^m\mu^X\left(x, m(t)\right) \\
=& A_t^m + B^m\bar{k}^{(0)} + x\Big\{B_t^m - B^mK^{(1)} + \frac{\delta}{1-\delta}\bar{h}^{(1)}\big(1 - \frac{1}{2(1-\delta)}\big) \\
&- B^mG^{(1)}\frac{\delta^2}{(1-\delta)^2} - \frac{1}{2}\frac{\delta}{(1-\delta)^2}(B^m)^2\rho^2H^{(11)}\Big\} \\
\Gamma_{1,1}^Z(t, x) =& \Big(\Sigma^X\big(x, m(t)\big)\Big)^2 = xH^{(11)}\big(m(t)\big) \\
\Gamma_{1,2}^Z(t, x) =& \Gamma_{2,1}^Z(t, x) = \Sigma^X(x, m(t))\big(\Sigma_1^G(t, x) + \rho\Sigma_2^G(t, x)\big) \\
&= x\Big\{\frac{1}{\vartheta}B^m(t)H^{(11)}\big(m(t)\big) + \frac{\delta}{1-\delta}G^{(1)}\big(m(t)\big)\Big\}
\end{aligned}
$$

$$\Gamma^Z_{2,2}(t,x) =\big(\Sigma^G_1(t,x)\big)^2 + \big(\Sigma^G_2(t,x)\big)^2 + 2\rho\Sigma^G_1(t,x)\Sigma^G_2(t,x)$$
$$=x\Big\{(B^m(t))^2 H^{(11)}\big(m(t)\big)\Big(1+\frac{\rho^2\delta^2}{(1-\delta)^2}+2\frac{\rho^2\delta}{1-\delta}\Big)$$
$$+2B^m(t)G^{(1)}\big(m(t)\big)\frac{\delta}{(1-\delta)^2}+\frac{\delta^2}{(1-\delta)^2}\bar{h}^{(1)}\big(m(t)\big)\Big\},$$

where we have omitted the dependence on m, t and x for reasons of better readability and we have substituted the following equations:

$$\bar{\pi}^m(t) =\frac{1}{1-\delta}\Big\{\frac{\big(\mu_1\big(x,m(t)\big)-r\big(m(t)\big)\big)}{\Big(\Sigma_1\big(x,m(t)\big)\Big)^2}$$
$$+\rho\frac{\Sigma^X\big(x,m(t)\big)}{\Sigma_1\big(x,m(t)\big)}B^m(t)\Big\}$$
$$\Big(\Sigma^X\big(x,m(t)\big)\Big)^2 =H^{(11)}\big(m(t)\big)x$$
$$\mu^X\big(x,m(t)\big) =\bar{k}^{(0)}\big(m(t)\big)-K^{(1)}\big(m(t)\big)x$$
$$\rho\Sigma^X\big(x,m(t)\big)\frac{\big(\mu_1\big(x,m(t)\big)-r\big(m(t)\big)\big)}{\Sigma_1\big(x,m(t)\big)} =G^{(1)}\big(m(t)\big)x$$
$$\left(\frac{\big(\mu_1\big(x,m(t)\big)-r\big(m(t)\big)\big)}{\Sigma_1\big(x,m(t)\big)}\right)^2 =\bar{h}^{(1)}\big(m(t)\big)x.$$

Observe that μ^Z and Γ^Z are piece-wise continuous and fulfill conditions i),ii),iii), iv) from Theorem 2.34 with $p=1$. Next we show that $\mu^Z_2(t,x)+\frac{1}{2}\Gamma^Z_{22}(t,x)=0$, which by comparison of coefficients is equivalent to Equation (2.4) for $i=2$. This equation is a direct consequence of the HJB PDE. More precisely, we have:

$$\mathrm{d}\Phi^m = \underbrace{\mathcal{L}^m(\pi)\Phi^m}_{=:\mu^{\Phi^m}=0}\mathrm{d}t + \Phi^m_x\Sigma^X\mathrm{d}W^X + \Phi^m_v\Sigma^V\mathrm{d}W^P$$

On the other side:

$$\Phi^m(t) = \Phi^m(0)\exp\{G(t)\}.$$

Thus,

$$0=\mu^{\Phi^m}=\Phi^m\big(\mu^Z_2+\frac{1}{2}\Gamma^Z_{22}\big).$$

It follows from Theorem 2.34 that process $\{\exp\{G(t)\}\}_{t\in[0,T]}$ and thus also process $\big\{\Phi^m(t)\big\}_{t\in[0,T]} = \big\{\Phi^m\big(t,V^{m,\bar{\pi}^m}(t),X^m(t)\big)\big\}_{t\in[0,T]}$ are martingales. Application of Proposition 4.3 completes the proof.

□

Proof of Proposition 4.18

Consider a fixed point $t\in[t_k,t_{k+1})$ for some $k\in\{0,\dots,K\}$. Recall the probabilistic representation for h^m from Proposition 4.7 and apply Lemma 2.42 step-wise starting at the back:

$$h^m(t,x) = \mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{\vartheta} g(s,\tilde{X}^m(s),m(s))\mathrm{d}s\Big\}\Big|\tilde{X}^m(t)=x\Big]$$

$$= \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(s))\big)^2}{\big(\nu(m(s))\big)^2}\tilde{X}^m(s)\mathrm{d}s\Big\}\Big|\tilde{X}^m(t)=x\Big]$$

$$= \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(s))\big)^2}{\big(\nu(m(s))\big)^2}\tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}\Big]$$

$$= \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\mathbb{E}\Big[\mathbb{E}\Big[\exp\Big\{\int_t^T \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(s))\big)^2}{\big(\nu(m(s))\big)^2}\tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}\Big|\mathscr{F}^{\tilde{X}^m}_{t_K}\Big]\Big]$$

$$= \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\mathbb{E}\Big[\exp\Big\{\int_t^{t_K} \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(s))\big)^2}{\big(\nu(m(s))\big)^2}\tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}$$

$$\cdot\,\mathbb{E}\Big[\exp\Big\{\underbrace{\frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(t_K))\big)^2}{\big(\nu(m(t_K))\big)^2}}_{=\beta_K}\int_{t_K}^T \tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}\Big|\mathscr{F}^{\tilde{X}^m}_{t_K}\Big]\Big]$$

$$\overset{(*)}{=} \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\mathbb{E}\Big[\exp\Big\{\int_t^{t_K} \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(s))\big)^2}{\big(\nu(m(s))\big)^2}\tilde{X}^m_{t,x}(s)\mathrm{d}s$$

$$+\underbrace{\tilde{A}^{0,\beta_K,m_K}(T-t_K)}_{=A_K(\tau_K)}+\underbrace{\tilde{B}^{0,\beta_K,m_K}(T-t_K)}_{=B_K(\tau_K)}\tilde{X}^m_{t,x}(t_K)\}\Big]$$

$$= \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\exp\{A_K(\tau_K)\}\mathbb{E}\Big[\mathbb{E}\Big[\exp\Big\{B_K(\tau_K)\tilde{X}^m_{t,x}(t_K)$$

$$+\int_t^{t_K} \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(s))\big)^2}{\big(\nu(m(s))\big)^2}\tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}\Big|\mathscr{F}^{\tilde{X}^m}_{t_{K-1}}\Big]\Big]$$

$$= \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\exp\{A_K(\tau_K)\}$$

$$\cdot\,\mathbb{E}\Big[\exp\Big\{\int_t^{t_{K-1}} \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(s))\big)^2}{\big(\nu(m(s))\big)^2}\tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}$$

$$\cdot\,\mathbb{E}\Big[\exp\Big\{B_K(\tau_K)\tilde{X}^m_{t,x}(t_K)+\underbrace{\frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(t_{K-1}))\big)^2}{\big(\nu(m(t_{K-1}))\big)^2}}_{=\beta_{K-1}}\int_{t_{K-1}}^{t_K} \tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}\Big|\mathscr{F}^{\tilde{X}^m}_{t_{K-1}}\Big]\Big]$$

$$
\begin{aligned}
&\overset{(*)}{=} \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\exp\{A_K(\tau_K)\}\\
&\cdot \mathbb{E}\Big[\exp\Big\{\int_t^{t_{K-1}} \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(s))\big)^2}{\big(\nu(m(s))\big)^2}\tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}\\
&\cdot \exp\Big\{\underbrace{\tilde{A}^{B_K,\beta_{K-1},m_{K-1}}(t_K-t_{K-1})}_{=A_{K-1}(\tau_{K-1})}+\underbrace{\tilde{B}^{B_K,\beta_{K-1},m_{K-1}}(t_K-t_{K-1})}_{=B_{K-1}(\tau_{K-1})}\tilde{X}^m_{t,x}(t_{K-1})\Big\}\Big]\\
&= \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\exp\{A_K(\tau_K)+A_{K-1}(\tau_{K-1})\}\\
&\cdot \mathbb{E}\Big[\mathbb{E}\Big[\exp\Big\{B_{K-1}(\tau_{K-1})\tilde{X}^m_{t,x}(t_{K-1})\\
&+\underbrace{\frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(t_{K-2}))\big)^2}{\big(\nu(m(t_{K-2}))\big)^2}}_{=:\beta_{K-2}}\int_t^{t_{K-1}}\tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}\Big|\mathscr{F}^{\tilde{X}^m}_{t_{K-2}}\Big]\Big]\\
&= \cdots\\
&= \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\Bigg(\prod_{z=k+1}^{K}\exp\{A_z(\tau_z)\}\Bigg)\mathbb{E}\Big[\exp\Big\{B_{k+1}(\tau_{k+1})\tilde{X}^m_{t,x}(t_{k+1})\\
&+\underbrace{\frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(t_k))\big)^2}{\big(\nu(m(t_k))\big)^2}}_{=:\beta_k}\int_t^{t_{k+1}}\tilde{X}^m_{t,x}(s)\mathrm{d}s\Big\}\Big]\\
&\overset{(*)}{=} \exp\Big\{\int_t^T \frac{1}{\vartheta}\delta r(m(s))\mathrm{d}s\Big\}\Bigg(\prod_{z=k+1}^{K}\exp\{A_z(\tau_z)\}\Bigg)\exp\Big\{\underbrace{\tilde{A}^{B_{k+1},\beta_k,m_k}(t_{k+1}-t)}_{=A_k(t_{k+1}-t)}\Big\}\\
&\cdot \exp\Big\{\underbrace{\tilde{B}^{B_{k+1},\beta_k,m_k}(t_{k+1}-t)}_{=B_k(t_{k+1}-t)}x\Big\}.
\end{aligned}
$$

For the equations marked with $(*)$ we have applied Lemma 2.42, so its assumptions need to be checked. They read as follows:

$$\beta_k = \frac{1}{2\vartheta}\frac{\delta}{1-\delta}\frac{\big(\lambda(m(t_k))\big)^2}{\big(\nu(m(t_k))\big)^2} < \frac{\tilde{\kappa}^2(m_k)}{2\chi^2(m_k)}, \forall k \in \{0,\dots,K\} \tag{B.3}$$

$$\alpha_k := B_{k+1}(\tau_{k+1}) < \frac{\tilde{\kappa}(m_k)+\tilde{a}(m_k)}{\chi^2(m_k)}, \forall k \in \{0,\dots,K-1\} \tag{B.4}$$

$$\alpha_K = 0 < \frac{\tilde{\kappa}(m_K)+\tilde{a}(m_K)}{\chi^2(m_K)}. \tag{B.5}$$

Inequality (B.3) follows directly from Assumption (4.48) and Inequality (B.5) is obvious as $\tilde{\kappa}(e), \tilde{a}(e) > 0, \forall e \in \mathcal{E}$. For Inequality (B.4) we consider two cases: $\delta > 0$ and $\delta < 0$. First assume that $\delta > 0$. It follows that $\beta_i > 0$ and thus $\tilde{a}(m_k) < \tilde{\kappa}(m_k)$ for all $k \in \{0,\dots,K\}$. Then, $\alpha_K = 0 < \max_{e\in\mathcal{E}}\Big\{\frac{\tilde{\kappa}(e)-\tilde{a}(e)}{\chi^2(e)}\Big\} := c_2$. From Assumption

(4.49) we obtain further $c_2 < \frac{\tilde{\kappa}(m_k)+\tilde{a}(m_k)}{\chi^2(m_k)}$ for all $i \in \{0, \ldots, K\}$. Statement vi) from Lemma 2.43 leads to $\alpha_{K-1} = B_K(\tau_K) < c_2 < \frac{\tilde{\kappa}(m_{K-1})+\tilde{a}(m_{K-1})}{\chi^2(m_{K-1})}$. To obtain Condition (B.4) for all k, observe that $\alpha_k = B_{k+1}(\tau_{k+1})$ and continue backwards in an analogous way showing that $B_{k+1} < c_2 < \frac{\tilde{\kappa}(m_k)+\tilde{a}(m_k)}{\chi^2(m_k)}$ for all $k \in \{0, \cdots, K\}$. Now let $\delta < 0$. Then $\beta_k < 0$ for all k. It follows from Statements ii) and iii) in Lemma 2.43 that $B_k(\tau_k) < 0$ for all k. This yields directly Condition (B.4).
So, by Proposition 4.7, Φ^m solves the corresponding HJB equation. The verification result and the optimal portfolio strategy follow as a direct application of Proposition 4.9.

□

B.1 Additional plots for Section 4.3.4

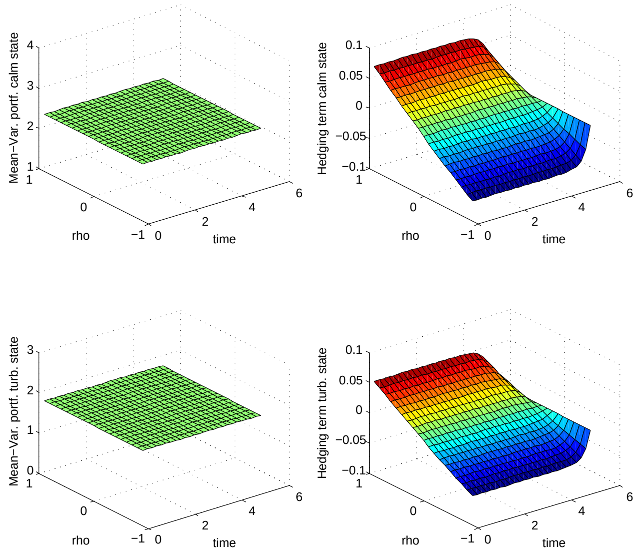


Figure B.1: Optimal mean-variance portfolio (see Equation (4.61)) and hedging term (see Equation (4.62)) over time for $\delta = 0.3$ and different values of ρ.

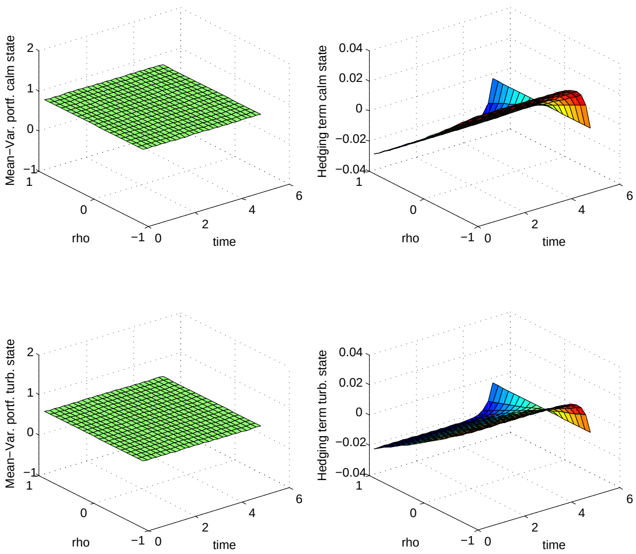


Figure B.2: Optimal mean-variance portfolio (see Equation (4.61)) and hedging term (see Equation (4.62)) over time for $\delta = -1$ and different values of ρ.

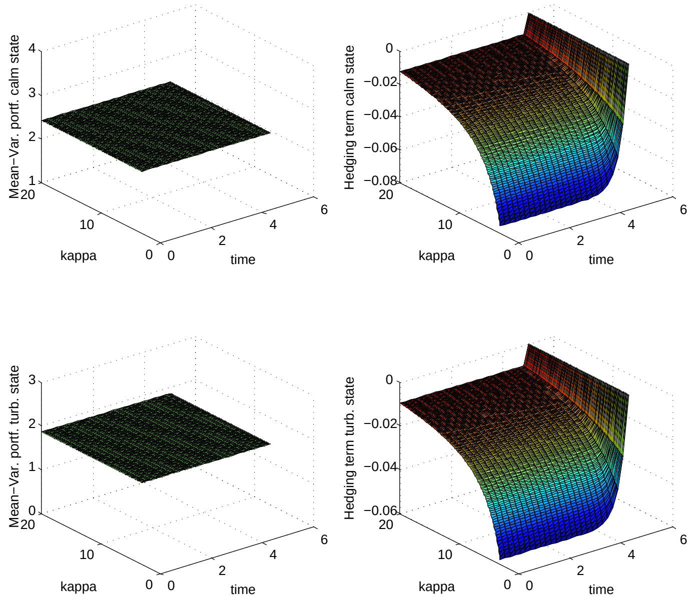


Figure B.3: Optimal mean-variance portfolio (see Equation (4.61)) and hedging term (see Equation (4.62)) over time for $\delta = 0.3$ and different values of κ.

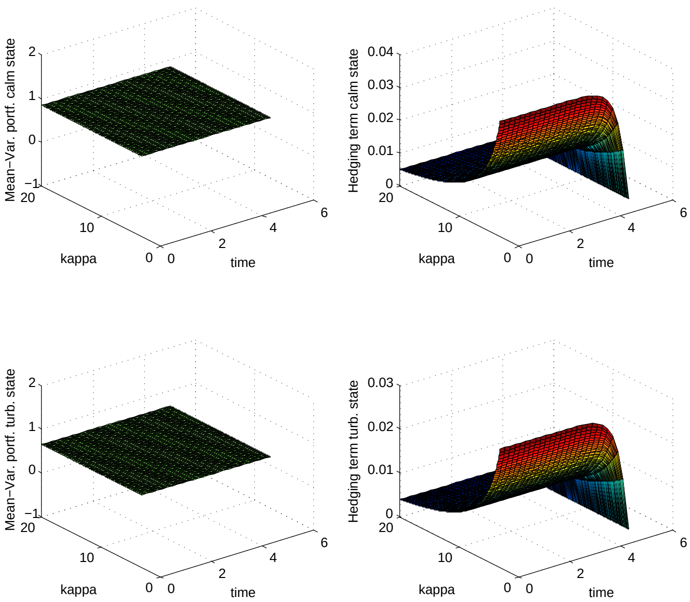


Figure B.4: Optimal mean-variance portfolio (see Equation (4.61)) and hedging term (see Equation (4.62)) over time for $\delta = -1$ and different values of κ.

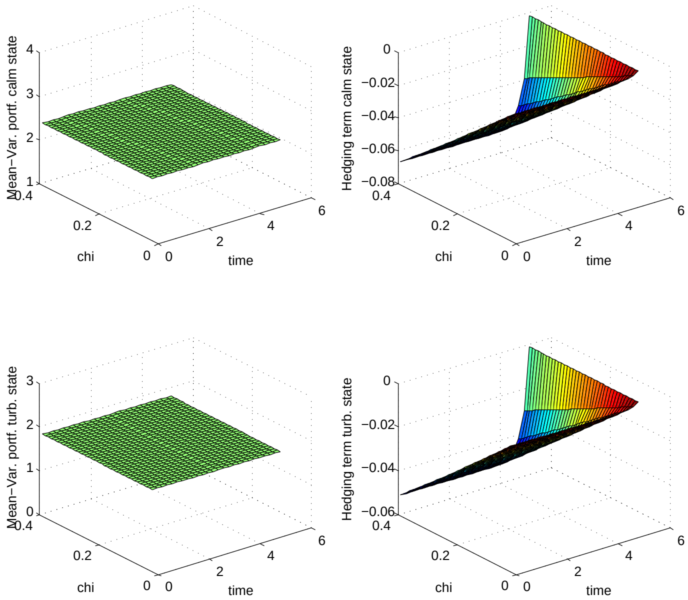


Figure B.5: Optimal mean-variance portfolio (see Equation (4.61)) and hedging term (see Equation (4.62)) over time for $\delta = 0.3$ and different values of χ.

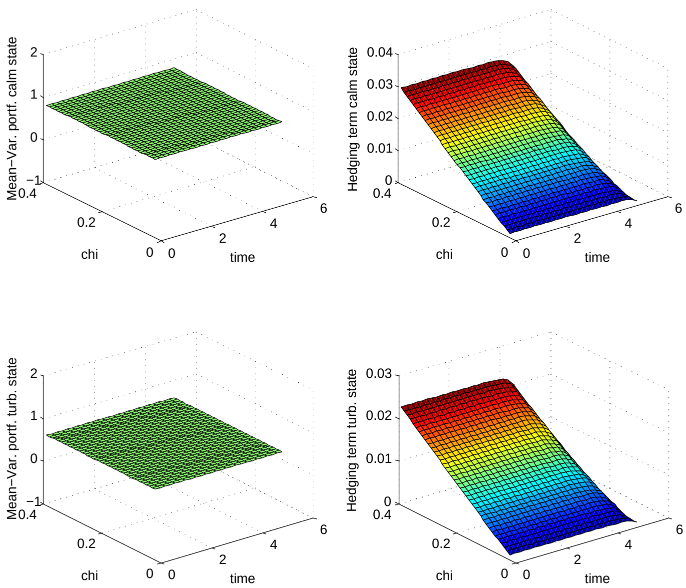


Figure B.6: Optimal mean-variance portfolio (see Equation (4.61)) and hedging term (see Equation (4.62)) over time for $\delta = -1$ and different values of χ.

Appendix C

Appendix for Chapter 5

Proof of Theorem 5.1
Consider an arbitrary but fixed point $(t, v, x, e_i) \in [0,T]\times[0,\infty)\times D^X\times\mathcal{E}$ and assume that the wealth of the investor at time t is v, further $X(t) = x$ and $\mathcal{MC}(t) = e_i$. As Φ is a martingale, it follows:

$$\begin{aligned}\mathbb{E}\Big[U_P\big(V^{\bar{\pi}}(T)\big)\Big|\mathscr{F}_t\Big] =\ & \mathbb{E}\Big[\frac{\big(V^{\bar{\pi}}(T)\big)^{\delta}}{\delta}\Big|\mathscr{F}_t\Big] = \mathbb{E}\Big[\Phi\big(T, V^{\bar{\pi}}(T), X(T), \mathcal{MC}(T)\big)\Big|\mathscr{F}_t\Big]\\ =&\Phi(t, v, x, e_i),\end{aligned}$$

which proves the first statement. For the second one we apply the idea from Proposition 2.3 in [63]: First observe that substitution of the ansatz $\Phi = \frac{v^\delta}{\delta} f$ in (3.16) leads to the following system of PDEs for f:

$$\begin{aligned}&f_t + f\delta\big(r + \bar{\pi}'(\mu - r)\big) + f_x'\mu^X + \frac{1}{2}Tr\big(f_{xx'}\Sigma^X(\Sigma^X)'\big) + f_x'\delta\Sigma^X\rho\Sigma'\bar{\pi}\\ &+ \frac{1}{2}f\delta(\delta-1)\bar{\pi}'\Sigma\Sigma'\bar{\pi}\Big|_{(t,x,e_i)} = -\sum_{z=1}^{I} q_{i,z} f(t, x, e_z),\ f(T, x, e_i) = 1, \forall i \in \{1, \dots, I\},\end{aligned} \tag{C.1}$$

where $\bar{\pi}(t) = \frac{1}{1-\delta}\Big\{\big(\Sigma\Sigma'\big)^{-1}(\mu - r) + \big(\Sigma'\big)^{-1}\rho'(\Sigma^X)'\frac{f_x}{f}\Big\}\Big|_{(t,x,e_i)}$. Let π be an arbitrary admissible strategy and define process $\{L(\tau)\}_{\tau\in[t,T]}$ by

$$L(\tau) := \big\{V^{\bar{\pi}}(\tau)\big\}^{\delta-1}V^{\pi}(\tau)f\big(\tau, X(\tau), \mathcal{MC}(\tau)\big).$$

Apply Itô's formula to obtain its dynamics:

$$\begin{aligned}\mathrm{d}L =&\big\{V^{\bar{\pi}}\big\}^{\delta-1}V^{\pi}\Big[f_t + f_x'\mu^X + \frac{1}{2}Tr\Big(f_{xx'}\Sigma^X(\Sigma^X)'\Big) + f(\delta-1)(r + \bar{\pi}'(\mu - r))\\ &+ f(r + \pi'[\mu - r]) + f_x'(\delta-1)\Sigma^X\rho\Sigma'\bar{\pi} + f_x'\Sigma^X\rho\Sigma'\pi + f(\delta-1)\pi'\Sigma\Sigma'\bar{\pi}\\ &+ \frac{1}{2}f(\delta-1)(\delta-2)\bar{\pi}'\Sigma\Sigma'\bar{\pi} + \sum_{i=1}^{I} q_{MC(\tau),i}f\big(\tau, X(\tau), e_i\big)\Big]\mathrm{d}\tau\end{aligned}$$

$$
\begin{aligned}
&+\{V^{\bar\pi}\}^{\delta-1}V^{\pi}f_x'\Sigma^X \mathrm{d}W^X+\{V^{\bar\pi}\}^{\delta-1}V^{\pi}f\big((\delta-1)\bar\pi'+\pi'\big)\Sigma \mathrm{d}W^P\\
&+\{V^{\bar\pi}\}^{\delta-1}V^{\pi}\big(f(\tau,X(\tau),e_1),\dots,f(\tau,X(\tau),e_I)\big)\mathrm{d}M\\
=:\ &\mu^L \mathrm{d}\tau+\Sigma_1^L \mathrm{d}W^X+\Sigma_2^L \mathrm{d}W^P+\Sigma_3^L \mathrm{d}M,
\end{aligned}
$$

where we have missed out the arguments $(\tau, X(\tau), \mathcal{MC}(\tau))$ for better readability. By a substitution of Equation (C.1) and the definition of $\bar\pi$ it holds:

$$
\begin{aligned}
\mu^L =&\{V^{\bar\pi}\}^{\delta-1}V^{\pi}\big[f_t+f\delta(r+\bar\pi'[\mu-r])+f_x'\mu^X+\frac{1}{2}Tr\big(f_{xx'}\Sigma^X(\Sigma^X)'\big)+f_x'\delta\Sigma^X\rho\Sigma'\bar\pi\\
&+\frac{1}{2}f\delta(\delta-1)\bar\pi'\Sigma\Sigma'\bar\pi+\sum_{i=1}^{I}q_{MC(\tau),i}f(\tau,X(\tau),e_i)+(\pi'-\bar\pi')\big(f[\mu-r]\\
&+\Sigma\rho'(\Sigma^X)'f_x+f(\delta-1)\Sigma\Sigma'\bar\pi\big)\big]=0.
\end{aligned}
$$

It follows that L is a local martingale, as all involved functions are continuous in X, V^{π}, $V^{\bar\pi}$, π and $\bar\pi$ for all $e_i \in \mathcal{E}$. Furthermore, as f is assumed to be positive, process L is positive as well, so it is a supermartingale. Using this together with the concavity of the utility function $U_P(v) = \frac{v^\delta}{\delta}$, the martingale property of $\{\Phi(\tau, V^{\bar\pi}(\tau), X(\tau))\}_{\tau\in[t,T]}$ and $L(T) = 1$ we obtain the following inequality:

$$
\begin{aligned}
&\mathbb{E}\big[U_P\big(V^{\pi}(T)\big)\big|\mathscr{F}_t\big]\le\mathbb{E}\big[U_P\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big]+\mathbb{E}\big[(U_P)_v\big(V^{\bar\pi}(T)\big)\big(V^{\pi}(T)-V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big]\\
&=\mathbb{E}\big[U_P\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big]+\mathbb{E}\big[L(T)\big|\mathscr{F}_t\big]-\mathbb{E}\big[\{V^{\bar\pi}(T)\}^{\delta}\big|\mathscr{F}_t\big]\\
&\le\mathbb{E}\big[U_P\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big]+L(t)-\delta\mathbb{E}\big[\Phi\big(T,V^{\bar\pi}(T),X(T),\mathcal{MC}(T)\big)\big|\mathscr{F}_t\big]\\
&=\mathbb{E}\big[U_P\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big]+v^{\delta}f\big(t,x,e_i\big)-\delta\Phi\big(t,v,x,e_i\big)=\mathbb{E}\big[U_P\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big].
\end{aligned}
$$

Our proof is complete.
Observe that we have not used the exponential structure of our model for this proof. Thus, the result holds for general time-dependent models with a stochastic factor.

□

Derivation of Remark 5.2
Consider an arbitrary but fixed point $(t, v, x, e_i) \in [0, T] \times [0, \infty) \times D^X \times \mathcal{E}$ and assume that the wealth of the investor at time t is v, further $X(t) = x$ and $\mathcal{MC}(t) = e_i$. Let π be an arbitrary admissible strategy. Apply Itô's formula to $\Phi(\tau, V^{\pi}(\tau), X(\tau), \mathcal{MC}(\tau))$, for an arbitrary $\tau \in [t, T]$, and use the HJB equation to obtain:

$$
\begin{aligned}
&\Phi\big(\tau,V^{\pi}(\tau),X(t),\mathcal{MC}(\tau)\big)\le\Phi\big(t,v,x,e_i\big)\\
&+\int_t^{\tau}\Phi_v\big(s,V^{\pi}(s),X(s),\mathcal{MC}(s)\big)\Sigma^V\big(V^{\pi}(s),X(s),\mathcal{MC}(s),\pi(s)\big)\mathrm{d}W^P(s)\\
&+\int_t^{\tau}\Phi_x\big(s,V^{\pi}(s),X(s),\mathcal{MC}(s)\big)'\Sigma^X\big(X(s),\mathcal{MC}(s)\big)\mathrm{d}W^X(s)\\
&+\int_t^{\tau}\sum_{i=1}^{I}\Phi\big(s,V^{\pi}(s),X(s),e_i\big)\mathrm{d}M_i(s)=:Y(\tau).
\end{aligned}
\tag{C.2}
$$

As $\Phi \geq 0$, it follows that $Y(\tau) \geq 0$. Furthermore, Y is a local martingale, as all involved functions are continuous in X, V^π and π, for all $e_i \in \mathcal{E}$. So, Y is a lower bounded local martingale, thus it is a supermartingale. Then it holds that:

$$\begin{aligned}&\mathbb{E}\Big[U_P\big(V^\pi(T)\big)\Big|\mathscr{F}_t\Big] = \mathbb{E}\Big[\frac{\big(V^\pi(T)\big)^\delta}{\delta}\Big|\mathscr{F}_t\Big] = \mathbb{E}\Big[\Phi\big(T, V^\pi(T), X(T), \mathcal{MC}(T)\big)\Big|\mathscr{F}_t\Big]\\ &\leq \mathbb{E}[Y(T)|\mathscr{F}_t] \leq Y(t) = \Phi\big(t, v, x, e_i\big).\end{aligned}$$

□

Proof for Proposition 5.3

The proof goes analogously to the proof of Theorem 5.1. Here we include it for completeness.

Consider an arbitrary but fixed point $(t, v, x) \in [0, T] \times [0, \infty) \times D^X$ and assume that the wealth of the investor at time t is v and $X^m(t) = x$. From the martingale property of $\{\Phi^m\big(\tau, V^{m,\bar{\pi}^m}(\tau), X^m(\tau)\big)\}_{\tau\in[t,T]}$ we obtain directly:

$$\mathbb{E}\Big[\frac{\big(V^{m,\bar{\pi}^m}(T)\big)^\delta}{\delta}\Big|\mathscr{F}_t\Big] = \mathbb{E}\Big[\Phi^m\big(T, V^{m,\bar{\pi}^m}(T), X^m(T)\big)\Big|\mathscr{F}_t\Big] = \Phi^m\big(t, v, x\big).$$

For the second statement substitute expression $\Phi^m = \frac{v^\delta}{\delta} f(t,x)$ in Equation (3.20) to obtain the following PDE for f^m:

$$\begin{aligned}&f^m_t + f^m\delta\big(r + (\bar{\pi}^m)'(\mu - r)\big) + (f^m_x)'\mu^X + \frac{1}{2}Tr\big(f^m_{xx'}\Sigma^X(\Sigma^X)'\big) + (f^m_x)'\delta\Sigma^X\rho\Sigma'\bar{\pi}^m\\ &+ \frac{1}{2}f^m\delta(\delta-1)(\bar{\pi}^m)'\Sigma\Sigma'\bar{\pi}^m\Big|_{\big(t,x,m(t)\big)} = 0, \forall (t,x) \in [t_k, t_{k+1}) \times D^X\\ &f^m(T,x) = 1,\end{aligned} \tag{C.3}$$

for all $k = 0, \ldots, K$, where $\bar{\pi}^m(t) = \frac{1}{1-\delta}\Big\{(\Sigma\Sigma')^{-1}(\mu - r) + (\Sigma')^{-1}\rho'(\Sigma^X)'\frac{f^m_x}{f^m}\Big\}\Big|_{(t,x,m(t))}$.

Now define process $\{L^m(\tau)\}_{\tau\in[t,T]}$ by $L^m(\tau) := \big\{V^{m,\bar{\pi}^m}(\tau)\big\}^{\delta-1}V^{m,\pi}(\tau)f\big(\tau, X^m(\tau)\big)$ for an arbitrary admissible portfolio strategy π. By Itô's formula it follows:

$$\begin{aligned}\mathrm{d}L^m =& \big\{V^{m,\bar{\pi}^m}\big\}^{\delta-1}V^{m,\pi}\Big[f^m_t + (f^m_x)'\mu^X + \frac{1}{2}Tr\Big(f^m_{xx'}\Sigma^X(\Sigma^X)'\Big)\\ &+ f^m(\delta-1)(r + (\bar{\pi}^m)'(\mu - r)) + f^m(r + \pi'(\mu - r)) + (f^m_x)'(\delta-1)\Sigma^X\rho\Sigma'\bar{\pi}^m\\ &+ (f^m_x)'\Sigma^X\rho\Sigma'\pi + f^m(\delta-1)\pi'\Sigma\Sigma'\bar{\pi}^m + \frac{1}{2}f^m(\delta-1)(\delta-2)(\bar{\pi}^m)'\Sigma\Sigma'\bar{\pi}^m\Big]\mathrm{d}\tau\\ &+ \big\{V^{m,\bar{\pi}}\big\}^{\delta-1}V^{m,\pi}(f^m_x)'\Sigma^X\mathrm{d}W^X\\ &+ \big\{V^{m,\bar{\pi}}\big\}^{\delta-1}V^{m,\pi}f^m\big((\delta-1)(\bar{\pi}^m)' + \pi'\big)\Sigma\mathrm{d}W^P\\ =:& \,\mu^{L^m}\mathrm{d}\tau + \Sigma^{L^m}_1\mathrm{d}W^X + \Sigma^{L^m}_2\mathrm{d}W^P,\end{aligned}$$

where we have missed out the arguments $(\tau, X^m(\tau), m(\tau))$ for better readability. It holds:

$$\begin{aligned}\mu^{L^m} =& \big\{V^{m,\bar{\pi}^m}\big\}^{\delta-1} V^{m,\pi}\big[f_t^m + f^m\delta(r + (\bar{\pi}^m)'(\mu - r)) + (f_x^m)'\mu^X + \frac{1}{2}Tr\big(f_{xx'}^m \Sigma^X (\Sigma^X)'\big)\\ &+ (f_x^m)'\delta\Sigma^X\rho\Sigma'\bar{\pi}^m + \frac{1}{2}f^m\delta(\delta-1)(\bar{\pi}^m)'\Sigma\Sigma'\bar{\pi}^m\\ &+ (\pi' - (\bar{\pi}^m)')\big(f(\mu - r) + \Sigma\rho'(\Sigma^X)'f_x^m + f^m(\delta-1)\Sigma\Sigma'\bar{\pi}\big)\big] = 0,\end{aligned}$$

where we have substituted Equation C.3 and the definition of $\bar{\pi}^m$. It follows that L^m is a local martingale. Furthermore, observe that L^m is positive, so it is a supermartingale. Using this together with the concavity of the utility function $U_P(v) = \frac{v^\delta}{\delta}$ and the martingale property of $\big\{\Phi^m\big(\tau, V^{m,\bar{\pi}^m}(\tau), X^m(\tau)\big)\big\}_{\tau\in[t,T]}$ we obtain the following inequality:

$$\begin{aligned}&\mathbb{E}\big[U_P\big(V^{m,\pi}(T)\big)\big|\mathscr{F}_t\big]\\ &\le \mathbb{E}\big[U_P\big(V^{m,\bar{\pi}^m}(T)\big)\big|\mathscr{F}_t\big] + \mathbb{E}\big[(U_P)_v\big(V^{m,\bar{\pi}^m}(T)\big)\big(V^{m,\pi}(T) - V^{m,\bar{\pi}^m}(T)\big)\big|\mathscr{F}_t\big]\\ &= \mathbb{E}\big[U_P\big(V^{m,\bar{\pi}^m}(T)\big)\big|\mathscr{F}_t\big] + \mathbb{E}\big[L^m(T)\big|\mathscr{F}_t\big] - \mathbb{E}\big[\big\{V^{m,\bar{\pi}^m}(T)\big\}^\delta\big|\mathscr{F}_t\big]\\ &\le \mathbb{E}\big[U_P\big(V^{m,\bar{\pi}^m}(T)\big)\big|\mathscr{F}_t\big] + L^m(t) - \delta\mathbb{E}\big[\Phi^m\big(T, V^{m,\bar{\pi}^m}(T), X^m(T)\big)\big|\mathscr{F}_t\big]\\ &= \mathbb{E}\big[U_P\big(V^{m,\bar{\pi}^m}(T)\big)\big|\mathscr{F}_t\big] + v^\delta f\big(t,x\big) - \delta\Phi^m\big(t,v,x\big) = \mathbb{E}\big[U_P\big(V^{m,\bar{\pi}^m}(T)\big)\big|\mathscr{F}_t\big].\end{aligned}$$

Our proof is complete.

□

Proof of Proposition 5.6

We will show that $\{\Phi^m(t)\}_{t\in[0,T]} := \big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ is a martingale and apply Proposition 5.4.

To this aim substitute (5.9) and (5.10) in the definition for G and derive its SDE:

$$\begin{aligned}\mathrm{d}G(t) =& \mu^G\big(t, X^m(t), m(t)\big)\mathrm{d}t + \Sigma_1^G\big(t, X^m(t), m(t)\big)\mathrm{d}W^X(t)\\ &+ \Sigma_2^G\big(t, X^m(t), m(t)\big)\mathrm{d}W^P(t),\end{aligned}$$

with

$$\begin{aligned}\mu^G =& -w + \delta r + \frac{\delta}{1-\delta}(\mu - r)'(\Sigma\Sigma')^{-1}(\mu - r) + \frac{\delta}{1-\delta}(B^m)'\Sigma^X\rho\Sigma^{-1}(\mu - r)\\ &- \frac{1}{2}\frac{\delta}{(1-\delta)^2}\big\{(\mu - r)'(\Sigma\Sigma')^{-1}(\mu - r) + (B^m)'\Sigma^X\rho\rho'(\Sigma^X)'B^m\\ &+ 2(\mu - r)'(\Sigma')^{-1}\rho'(\Sigma^X)'B^m\big\} + (B_t^m)'X^m + (B^m)'\mu^X\\ \Sigma_1^G =& (B^m)'\Sigma^X\\ \Sigma_2^G =& \frac{\delta}{1-\delta}\big\{(\mu - r)'(\Sigma')^{-1} + (B^m)'\Sigma^X\rho\big\},\end{aligned}$$

where the dependence on $(t, X^m(t), m(t))$ is omitted for simplicity. Thus, for the second component of the semimartingale characteristics of G we obtain:

$$\begin{aligned}\Gamma^G =&(B^m)'\Sigma^X(\Sigma^X)'B^m + 2\frac{\delta}{1-\delta}(B^m)'\Sigma^X\rho\Sigma^{-1}(\mu - r)\\&+ 2\frac{\delta}{1-\delta}(B^m)'\Sigma^X\rho\rho'(\Sigma^X)'B^m + \frac{\delta^2}{(1-\delta)^2}\big\{(\mu-r)'(\Sigma\Sigma')^{-1}(\mu-r)\\&+ (B^m)'\Sigma^X\rho\rho'(\Sigma^X)'B^m + 2(\mu-r)'(\Sigma')^{-1}\rho'(\Sigma^X)'B^m\big\}.\end{aligned}$$

For process $\bar{X}^m$ we introduce the following notation:

$$\mathrm{d}\bar{X}^m = \bar{\mu}^X\big(t, X^m(t), m(t)\big)\mathrm{d}t + \bar{\Sigma}^X\big(t, X^m(t), m(t)\big)\mathrm{d}W^X.$$

Together, the semimartingale characteristics of process Z are given as follows:

$$\begin{aligned}\mu^Z &= \begin{pmatrix}\bar{\mu}^X\\ \mu^G\end{pmatrix}\\ \Gamma^Z &= \begin{pmatrix}\bar{\Sigma}^X & \bar{\Sigma}^X\big((\Sigma_1^G)' + \rho(\Sigma_2^G)'\big)\\ \big(\Sigma_1^G + \Sigma_2^G\rho'\big)(\bar{\Sigma}^X)' & \Gamma^G\end{pmatrix}.\end{aligned}$$

The model specifications from (3.2)-(3.7) lead to the affine representations for μ^Z and Γ^Z as stated in the proposition. The required conditions i)-iv) assure that we can apply Theorem 2.34 to process Z. We only need to show that $\mu^G + \frac{1}{2}\Gamma^G = 0$ in order to conclude that $\exp\{G\}$ and thus Φ^m as well are martingales. This condition follows from the fact that Φ^m solves the HJB PDE. More precisely, we have:

$$\mathrm{d}\Phi^m = \underbrace{\mathcal{L}^m(\pi)\Phi^m}_{=:\mu^{\Phi^m}=0}\mathrm{d}t + \Phi_x^m\Sigma^X\mathrm{d}W^X + \Phi_v^m\Sigma^V\mathrm{d}W^P.$$

On the other side:

$$\Phi^m(t) = \Phi^m(0)\exp\{G(t)\}.$$

Thus,

$$0 = \mu^{\Phi^m} = \Phi^m\big(\mu^G + \frac{1}{2}\Gamma^G\big).$$

So, process Φ^m is a martingale. We can conclude by Proposition 5.3 that Φ^m and $\bar{\pi}^m$ are the optimal solution for the time-dependent model.

□

Proof of Proposition 5.13
It follows directly from Theorem 5.10 that function B is the solution of the following ODE:

$$B_t(t) + \delta - B(t)\kappa = 0, B(T) = 0.$$

Thus, we obtain for B the expression given by (5.39). Substitution of the model specifications from (5.37) in (5.24) leads to the solution for w as given by (5.41). So, function Φ as defined in (5.38) is the HJB solution and its maximum is obtained at $\bar{\pi}$ as given in (5.42).
What remains to be demonstrated is that $\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\}_{t\in[0,T]}$ is a martingale, where function Φ^m is given by:

$$\Phi^m(t, v, x) = \frac{v^\delta}{\delta} \exp\Big\{\int_t^T w(s, m(s))\mathrm{d}s\Big\} \exp\{B(t)x\}. \tag{C.4}$$

Start by applying Itô's formula:

$$\begin{aligned}
&\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big) = \Phi\big(0, V^{m,\bar{\pi}^m}(0), X^m(0)\big) \\
&+ \int_0^t \underbrace{\mathcal{L}(m(s), \bar{\pi}^m)\Phi\big(s, V^{m,\bar{\pi}^m}(s), X^m(s)\big)}_{=0} \mathrm{d}s \\
&+ \int_0^t \Phi_v^m\big(s, V^{m,\bar{\pi}^m}(s), X^m(s)\big)\Sigma_1^V(s, V^{m,\bar{\pi}^m}(s), m(s), \bar{\pi}^m(s))\mathrm{d}W_1^P(s) \\
&+ \int_0^t \Phi_v^m\big(s, V^{m,\bar{\pi}^m}(s), X^m(s)\big)\Sigma_2^V(V^{m,\bar{\pi}^m}(s), m(s), \bar{\pi}^m(s))\mathrm{d}W_2^P(s) \\
&+ \int_0^t \Phi_x^m\big(s, V^{m,\bar{\pi}^m}(s), X^m(s), m(s)\big)\chi \mathrm{d}W^X(s),
\end{aligned}$$

for any $t \in [0, T]$. We will apply Theorem 2.31 to prove that the process above is a square integrable martingale. Thus, we need to show that:

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^T \Big(\Phi_v^m(t)\Sigma_1^V(t) - \Phi_x^m(t)\chi\Big)^2 \mathrm{d}t\Big] < \infty \\
&\mathbb{E}\Big[\int_0^T \Big(\Phi_v^m(t)\Sigma_2^V(t)\Big)^2 \mathrm{d}t\Big] < \infty,
\end{aligned}$$

where we substituted $\mathrm{d}W_1^P = -\mathrm{d}W^X$ and omitted the arguments of the functions to improve readability. Recalling from (5.38) and (3.9) the definitions of Φ^m and $\Sigma^V = (\Sigma_1^V, \Sigma_2^V)$ and the fact that $B(t)$, $w(t, e_i)$, $A_2(T_1 - t)$ and $\bar{\pi}^m(t)$ are bounded on $[0, T]$, it becomes clear that it suffices to show the following inequality:

$$\mathcal{A} := \mathbb{E}\Big[\int_0^T (V^{m,\bar{\pi}^m}(t))^{2\delta} \exp\big\{2B(t)X^m(t)\big\}\mathrm{d}t\Big] < \infty. \tag{C.5}$$

Note that the solution for process $V^{m,\bar{\pi}^m}$ is given by:

$$\begin{aligned}
&V^{m,\bar{\pi}^m}(t) = V^{m,\bar{\pi}^m}(0) \exp\Big\{\int_0^t X^m(s) + \bar{\pi}_1^m(s)\lambda_1\chi A_2(T_1 - s) + \bar{\pi}_2^m(s)\lambda_2\nu \\
&- \frac{1}{2}\big[\chi A_2(T_1 - s)\bar{\pi}_1^m(s) + \bar{\pi}_2^m(s)\rho_{12}\nu\big]^2 - \frac{1}{2}\big[\bar{\pi}_2^m(s)\sqrt{1 - \rho_{12}^2}\nu\big]^2 \mathrm{d}s \\
&+ \int_0^t \big[\chi A_2(T_1 - s)\bar{\pi}_1^m(s) + \bar{\pi}_2^m(s)\rho_{12}\nu\big]\mathrm{d}W_1^P(s) + \int_0^t \bar{\pi}_2^m(s)\sqrt{1 - \rho_{12}^2}\nu \mathrm{d}W_2^P(s)\Big\},
\end{aligned}$$

where we again omitted the dependence of the parameters on $m(t)$. Inserting the previous expression and the relation $W_1^P = -W^X$ in the formula for $\mathcal{A}$ from (C.5) leads to:

$$\begin{aligned}\mathcal{A} = v_0^{2\delta}\mathbb{E}\Big[\int_0^T \exp\Big\{2\delta\int_0^t X^m(s) + \bar{\pi}_1^m(s)\lambda_1\chi A_2(T_1-s) + \bar{\pi}_2^m(s)\lambda_2\nu \\ -\frac{1}{2}\big[\chi A_2(T_1-s)\bar{\pi}_1^m(s) + \bar{\pi}_2^m(s)\rho_{12}\nu\big]^2 - \frac{1}{2}\big[\bar{\pi}_2^m(s)\sqrt{1-\rho_{12}^2}\nu\big]^2 \mathrm{d}s \\ + 2\delta\int_0^t -\bar{\pi}_2^m(s)\rho_{12}\nu - \chi A_2(T_1-s)\bar{\pi}_1^m(s)\mathrm{d}W^X(s) \\ + 2\delta\int_0^t \bar{\pi}_2^m(s)\nu\sqrt{1-\rho_{12}^2}\mathrm{d}W_2^P(s) + 2B(t)X^m(t)\Big\}\mathrm{d}t\Big].\end{aligned}$$

By exploiting the boundedness of $\bar{\pi}^m(t)$ and $A_2(T_1-t)$ on $[0,T]$, we obtain:

$$\begin{aligned}\mathcal{A} \leq c_1\mathbb{E}\Big[\int_0^T \exp\Big\{2\delta\int_0^t X^m(s)\mathrm{d}s + 2\delta\int_0^t -\bar{\pi}_2^m(s)\rho_{12}\nu - \chi A_2(T_1-s)\bar{\pi}_1^m(s)\mathrm{d}W^X(s) \\ + 2\delta\int_0^t \bar{\pi}_2^m(s)\nu\sqrt{1-\rho_{12}^2}\mathrm{d}W_2^P(s) + 2B(t)X^m(t)\Big\}\mathrm{d}t\Big],\end{aligned}$$

for some constant $c_1 \in [0,\infty)$. Now insert the following two equations for process X^m:

$$\begin{aligned}\int_0^t X^m(s)\mathrm{d}s = \frac{1}{\kappa}\big(X^m(0) - X^m(t) + \int_0^t \kappa\theta(m(s))\mathrm{d}s + \chi W^X(t)\big) \\ X^m(t) = X^m(0)e^{-\kappa t} + \int_0^t e^{\kappa(s-t)}\kappa\theta(m(s))\mathrm{d}s + \int_0^t \chi e^{\kappa(s-t)}\mathrm{d}W^X(s),\end{aligned}$$

and rewrite the inequality for $\mathcal{A}$:

$$\begin{aligned}\mathcal{A} \leq c_1\mathbb{E}\Big[\int_0^T \exp\Big\{\frac{2\delta}{\kappa}\big(X^m(0) - X^m(t) + \int_0^t \kappa\theta(m(s))\mathrm{d}s + \chi W^X(t)\big) \\ + 2\delta\int_0^t -\bar{\pi}_2^m(s)\rho_{12}\nu - \chi A_2(T_1-s)\bar{\pi}_1^m(s)\mathrm{d}W^X(s) \\ + 2\delta\int_0^t \bar{\pi}_2^m(s)\nu\sqrt{1-\rho_{12}^2}\mathrm{d}W_2^P(s) + 2B(t)X^m(t)\Big\}\mathrm{d}t\Big] \\ = c_1\mathbb{E}\Big[\int_0^T \exp\Big\{\frac{2\delta}{\kappa}\big(X^m(0) + \int_0^t \kappa\theta(m(s))\mathrm{d}s + \chi W^X(t)\big) \\ + 2\delta\int_0^t -\bar{\pi}_2^m(s)\rho_{12}\nu - \chi A_2(T_1-s)\bar{\pi}_1^m(s)\mathrm{d}W^X(s) \\ + 2\big(B(t) - \frac{\delta}{\kappa}\big)\int_0^t \chi e^{\kappa(s-t)}\mathrm{d}W^X(s) + 2\delta\int_0^t \bar{\pi}_2^m(s)\nu\sqrt{1-\rho_{12}^2}\mathrm{d}W_2^P(s) \\ + 2\big(B(t) - \frac{\delta}{\kappa}\big)\big(X^m(0)e^{-\kappa t} + \int_0^t e^{\kappa(s-t)}\kappa\theta(m(s))\mathrm{d}s\big)\Big\}\mathrm{d}t\Big]\end{aligned}$$

$$\leq c_2\mathbb{E}\Big[\int_0^T \exp\Big\{\int_0^t \frac{2\delta}{\kappa}\chi + 2\delta\big(-\bar{\pi}_2^m(s)\rho_{12}\nu - \chi A_2(T_1-s)\bar{\pi}_1^m(s)\big)$$
$$+ 2\big(B(t)-\frac{\delta}{\kappa}\big)\big(\chi e^{\kappa(s-t)}\big)\mathrm{d}W^X(s) + \int_0^t 2\delta\bar{\pi}_2^m(s)\nu\sqrt{1-\rho_{12}^2}\mathrm{d}W_2^P(s)\Big\}\mathrm{d}t\Big]$$
$$=: c_2\mathbb{E}\Big[\int_0^T \exp\Big\{\int_0^t a(s,t)\mathrm{d}W^X(s) + \int_0^t b(s)\mathrm{d}W_2^P(s)\Big\}\mathrm{d}t\Big],$$

where $c_2 \in [0,\infty)$. As a and b are bounded and thus integrable, and W^X and W_2^P are independent we get:

$$\int_0^t a(s,t)\mathrm{d}W^X(s) + \int_0^t b(s)\mathrm{d}W_2^P(s) \overset{d}{=} \mathcal{N}\Big(0, \int_0^t a^2(s,t) + b^2(s)\mathrm{d}s\Big),$$

where $\mathcal{N}(\mu,\Sigma^2)$ denotes the normal distribution with mean μ and variance Σ^2. Now exchange the integration and the expectation:

$$\mathcal{A} \leq c_2\int_0^T \mathbb{E}\Big[\exp\Big\{\int_0^t a(s,t)\mathrm{d}W^X(s) + \int_0^t b(s)\mathrm{d}W_2^P(s)\Big\}\Big]\mathrm{d}t$$
$$= c_2\int_0^T \exp\Big\{\frac{1}{2}\int_0^t a^2(s,t) + b^2(s)\mathrm{d}s\Big\}\mathrm{d}t \leq c_3 < \infty,$$

for a constant $c_3 \in [0,\infty)$. Our proof is complete.
Note that the verification result can be also proved by Proposition 5.6. However, we prefer to provide an alternative approach as it is straightforward, does not rely on more specific concepts and illustrates a possible solution method when Proposition 5.6 is not applicable. □

Proof of Proposition 5.14
Again, we will apply Theorem 5.10. To this aim we first find in explicit form a function Φ that solves the HJB equation. Then we show that process $\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\}_{t\in[0,T]}$ is a martingale by Proposition 5.6.
From Theorem 5.10 it follows directly that functions B_j, for $j=1,2$ satisfy the following ODEs:

$$\frac{\partial}{\partial t}B_j + B_j\big[\frac{\delta}{1-\delta}\chi_j\rho_j c_j - \kappa_j\big] + \frac{1}{2}\big(B_j\big)^2\frac{\chi_j^2}{\vartheta_j} + \frac{1}{2}\frac{\delta}{1-\delta}c_j^2 = 0, \tag{C.6}$$

with terminal condition $B_j(T)=0$. Note that Equation (C.6) corresponds to Equation (5.29), where we have substituted the model specifications from (5.43). Conditions (5.44) and (5.45) allow us to apply Corollary (2.44) and conclude that functions B_j, for $j=1,2$ are as in Equation (5.46). Furthermore, from Theorem 5.10 function w is given as in Equation (5.48). The expression for $\bar{\pi}$ follows as well directly from Theorem 5.10.
What remains to be proved is that $\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\}_{t\in[0,T]}$ is a martingale. We will show this by applying Proposition 5.6. Consider process G as defined there:

$$G(t) := \ln\left(\frac{\Phi^m(t, V^{m,\bar{\pi}^m}(t), X^m(t))}{\Phi^m(0, V^{m,\bar{\pi}^m}(0), X^m(0))}\right).$$

From the poof of Proposition 5.6 we know that the dynamics of process G are given by:

$$\mathrm{d}G(t) = \mu^G\big(s, X^m(s)\big)\mathrm{d}s + \sum_{j=1}^{2} \Sigma_j^G\big(s, X^m(s)\big)\mathrm{d}W_j^X + \sum_{j=1}^{2} \Sigma_{(2+j)}^G\big(s, X^m(s)\big)\mathrm{d}W_j^P,$$

where

$$\begin{aligned}\mu^G\big(s, X^m(s)\big) =& \sum_{j=1}^{2} X_j^m(s)\Big[\frac{\partial}{\partial t}B_j(s) - \kappa_j B_j(s) + \frac{\delta - 2\delta^2}{2(1-\delta)^2}c_j^2 \\ & - \frac{\delta^2}{(1-\delta)^2}B_j\chi_j c_j\rho_j - \frac{\delta}{2(1-\delta)^2}\big(B_j(s)\chi_j\rho_j\big)^2 \\ \Sigma_j^G\big(s, X^m(s)\big) =& B_j(s)\chi_j\sqrt{X_j^m(s)}, j = 1, 2 \\ \Sigma_{(2+j)}^G\big(s, X^m(s)\big) =& \frac{\delta}{1-\delta}\sqrt{X_j^m(s)}\big(c_j + B_j(s)\chi_j\rho_j\big), j = 1, 2.\end{aligned}$$

Thus, the differential semimartingale characteristics of process $Z := (X_1^m, X_2^m, G)$ are given as follows:

$$\begin{aligned}\mu_j^Z(t, x) =& \mu_j^X(x, m(t)) \\ \mu_3^Z(t, x) =& \mu^G(t, x) \\ \Gamma_{jj}^Z(t, x) =& \chi_j^2 x_j \\ \Gamma_{33}^Z(t, x) =& \sum_{j=1}^{2} x_j\Big[\big(B_j(t)\big)^2\chi_j^2 + \frac{\delta^2}{(1-\delta)^2}c_j^2 + \frac{2\delta}{(1-\delta)^2}B_j\chi_j c_j\rho_j \\ & + \frac{2\delta - \delta^2}{(1-\delta)^2}\big(B_j(s)\chi_j\rho_j\big)^2 \\ \Gamma_{12}^Z(t, x) =& \Gamma_{21}^Z(t, x) = 0 \\ \Gamma_{j3}^Z(t, x) =& \Gamma_{3j}^Z(t, x) = \chi_j x_j\Big[B_j(t)\chi_j + \frac{\delta}{1-\delta}\rho_j\big(c_j + B_j(t)\chi_j\rho_j\big)\Big],\end{aligned}$$

for $j = 1, 2$. It is easily checked that μ^Z and Γ^Z satisfy Conditions i)-iv) from Theorem 2.34 with $p = 2$. Thus, by Proposition 5.6, $\big\{\Phi^m\big(t, V^{m,\bar{\pi}^m}(t), X^m(t)\big)\big\}_{t\in[0,T]}$ is a martingale. Note that the martingale condition 2.4 from Theorem 2.34 is already proved in Proposition 5.6. Finally, Theorem 5.10 delivers the verification result for the Markov-modulated model.

□

Appendix D

Appendix for Chapter 6

Proof of Theorem 6.5
First denote:

$$C^{\bar{\pi}}(t) := V^{\bar{\pi}}(t) - Fd(t, X_1(t)), \tag{D.1}$$

for all $t \in [0,T]$. Analogously to SDE (6.9) it can be shown that d exhibits the following dynamics:

$$\begin{aligned} \mathrm{d}d\big(t, X_1(t)\big) =& d\big(t, X_1(t)\big)\Big[\big\{r\big(X_1(t)\big) + A_2(T_1 - t)\Sigma_{11}^X \lambda_1\big(\mathcal{MC}(t)\big)\big\}\mathrm{d}t \\ &+ A_2(T_1 - t)\Sigma_{11}^X \mathrm{d}W_1^P\Big]. \end{aligned} \tag{D.2}$$

Substitution of (D.2), the SDE for $V^{\bar{\pi}}$ from (3.9) and the definitions of process $\bar{\pi}$ from (6.33) in (D.1), leads to the following SDE for process $C^{\bar{\pi}}$:

$$\begin{aligned} \mathrm{d}C^{\bar{\pi}}(t) = &\Big\{V^{\bar{\pi}} X_1 - FdX_1 + \frac{1}{1-\delta} C^{\bar{\pi}}(\mu - r)'(\Sigma\Sigma')^{-1}(\mu - r) \\ &+ \frac{1}{1-\delta} C^{\bar{\pi}} \frac{f_x'}{f} \Sigma^X \rho \Sigma^{-1}(\mu - r) + Fd_x' \Sigma^X \rho \Sigma^{-1}(\mu - r) - FdA_2(T - t)\Sigma_{11}^X \lambda_1\Big\}\mathrm{d}t \\ &+ \Big\{\frac{1}{1-\delta} C^{\bar{\pi}}(\mu - r)'(\Sigma')^{-1} + \frac{1}{1-\delta} C^{\bar{\pi}} \frac{f_x'}{f} \Sigma^X \rho + Fd_x' \Sigma^X \rho \\ &- FdA_2(T - t)\Sigma_{11}^X (1, 0, \ldots, 0)\Big\}\mathrm{d}W^P, \end{aligned}$$

where the dependence on $(t, X(t), \mathcal{MC}(t))$ is omitted for better readability. Using that:

$$\begin{aligned} d_x' \Sigma^X \rho &= \big(dA_2(T - t)\Sigma_{11}^X, 0, \ldots, 0\big) \\ d_x' \Sigma^X \rho \Sigma^{-1}(\mu - r) &= dA_2(T - t)\Sigma_{11}^X \lambda_1, \end{aligned}$$

the SDE for $C^{\bar{\pi}}$ simplifies to:

$$\begin{aligned} \mathrm{d}C^{\bar{\pi}}(t) = C^{\bar{\pi}}(t)\Big[&\Big\{X_1 + \frac{1}{1-\delta}(\mu - r)'(\Sigma\Sigma')^{-1}(\mu - r) + \frac{1}{1-\delta}\frac{f_x'}{f}\Sigma^X \rho \Sigma^{-1}(\mu - r)\Big\}\mathrm{d}t \\ &+ \frac{1}{1-\delta}\Big\{(\mu - r)'(\Sigma')^{-1} + \frac{f_x'}{f}\Sigma^X \rho\Big\}\mathrm{d}W^P\Big]. \end{aligned}$$

From the exponential structure of the above SDE we can follow that if $C^{\bar{\pi}}(0) \geq 0$, then $C^{\bar{\pi}}(t) \geq 0$, for all $t \in (0, T]$.

□

Proof of Theorem 6.6

In what follows we extend the proof of Theorem 5.1 to the HARA utility function. Consider an arbitrary point $(t, v, x, e_i) \in D^{\Lambda}$, where v corresponds to the wealth of the investor at time point t and $X(t) = x, \mathcal{MC}(t) = e_i$. The first statement follows directly by applying the terminal condition from (6.22) and the martingale property of Φ:

$$\mathbb{E}\big[U_H\big(V^{\bar{\pi}}(T)\big)|\mathscr{F}_t\big] = \mathbb{E}\big[\Phi\big(T, V^{\bar{\pi}}(T), X(T), \mathcal{MC}(T)\big)|\mathscr{F}_t\big] = \Phi(t, v, x, e_i).$$

For the second statement consider an arbitrary admissible strategy $\pi \in \Lambda(t, v)$ and define the following process for $\tau \in [t, T]$:

$$L(\tau) := \big\{\underbrace{V^{\bar{\pi}}(\tau) - Fd(\tau, X_1(\tau))}_{=:C^{\bar{\pi}}(\tau)}\big\}^{\delta-1}\big\{\underbrace{V^{\pi}(\tau) - Fd(\tau, X_1(\tau))}_{=:C^{\pi}(\tau)}\big\}f\big(\tau, X(\tau), \mathcal{MC}(\tau)\big).$$

First, we show that process L is a supermartingale. To this aim we derive its dynamics by applying Itô's formula for Markov-modulated diffusions (see Theorem 2.72):

$$\begin{aligned}
\mathrm{d}L(\tau) = \Big[&\big(C^{\bar{\pi}}\big)^{\delta-3}C^{\pi}\Big\{\big(C^{\bar{\pi}}\big)^2\big[fX_1 + f_t + f_x'\mu^X + \frac{1}{2}Tr\{f_{xx'}\Sigma^X(\Sigma^X)'\}\\
&+ \sum_{z=1}^{I} q_{MC,z}f(e_z)\big] + C^{\bar{\pi}}(\delta-1)\big[fV^{\bar{\pi}}[X_1 + \bar{\pi}'(\mu - r)] - fFd[X_1 + A_2\Sigma_{11}^X\lambda_1]\\
&+ f_x'\Sigma^X\rho\Sigma'V^{\bar{\pi}}\bar{\pi} - f_x'\Sigma^X\rho FdA_2\Sigma_{11}^X(1,0,\dots,0)'\big]\\
&+ \frac{1}{2}(\delta-1)(\delta-2)f\big[\big(V^{\bar{\pi}}\big)^2\bar{\pi}'\Sigma\Sigma'\bar{\pi} + \big(FdA_2\Sigma_{11}^X\big)^2 + 2V^{\bar{\pi}}FdA_2\Sigma_{11}^X\bar{\pi}'\Sigma(1,0,\dots,0)'\big]\Big\}\\
&+ \big(C^{\bar{\pi}}\big)^{\delta-2}V^{\pi}\pi'\Big\{C^{\bar{\pi}}f(\mu - r) + (\delta-1)f\Sigma\Sigma'V^{\bar{\pi}}\bar{\pi}\\
&- (\delta-1)fFdA_2\Sigma_{11}^X\Sigma(1,0,\dots,0)' + C^{\bar{\pi}}\Sigma\rho'(\Sigma^X)'f_x\Big\}\\
&+ \big(C^{\bar{\pi}}\big)^{\delta-2}FdA_2\Sigma_{11}^X\Big\{-C^{\bar{\pi}}f\lambda_1 + (\delta-1)fFdA_2\Sigma_{11}^X\\
&- (\delta-1)fV^{\bar{\pi}}\bar{\pi}'\Sigma(1,0,\dots,0)' - C^{\bar{\pi}}f_x'\Sigma^X\rho(1,0,\dots,0)'\Big\}\Big]\mathrm{d}\tau\\
&+ \big(C^{\bar{\pi}}\big)^{\delta-1}C^{\pi}f_x'\Sigma^X\mathrm{d}W^X\\
&+ \Big[(\delta-1)\big(C^{\bar{\pi}}\big)^{\delta-2}C^{\pi}f\big\{V^{\bar{\pi}}\bar{\pi}'\Sigma - FdA_2\Sigma_{11}^X(1,0\dots,0)\big\}\\
&+ \big(C^{\bar{\pi}}\big)^{\delta-1}f\big\{V^{\pi}\pi'\Sigma - FdA_2\Sigma_{11}^X(1,0\dots,0)\big\}\Big]\mathrm{d}W^P\\
&+ \big(C^{\bar{\pi}}\big)^{\delta-1}C^{\pi}\big(f(e_1),\dots,f(e_I)\big)\mathrm{d}M =: \mu^L\mathrm{d}\tau + \Sigma_1^L\mathrm{d}W^X + \Sigma_2^L\mathrm{d}W^P + \Sigma_3^L\mathrm{d}M,
\end{aligned}$$

where $A_2 = A_2(T-\tau)$ and the dependence on $(\tau, X_1(\tau), \mathcal{MC}(\tau))$ is omitted for better readability. Now we show that $\mu^L = 0$. To ease the exposition we introduce the following notation:

$$\mu^L =: \left(C^{\bar\pi}\right)^{\delta-3} C^{\pi}(*) + \left(C^{\bar\pi}\right)^{\delta-2} V^{\pi}\pi'(**) + \left(C^{\bar\pi}\right)^{\delta-2} F d A_2 \Sigma^X_{11}(***).$$

By a substitution of (6.25) in the expressions for $(**)$ and $(***)$, we obtain $(**) = (0,\ldots,0)'$ and $(***) = 0$. It remains to be shown that $(*) = 0$. Observe that by inserting $\Phi(t,v,x,e_i) = \frac{1-\delta}{\delta}\alpha\Big\{\frac{1}{1-\delta}\big(v - Fd(t,x_1)\big)\Big\}^{\delta} f(t,x,e_i)$ in System (6.22) with $\pi = \bar\pi$ we obtain the following System of PDEs for f:

$$\begin{aligned}
&\left(C^{\bar\pi}\right)^2\big[f_t + f_x'\mu^X + \frac{1}{2}Tr\{f_{xx'}\Sigma^X(\Sigma^X)'\} + \sum_{z=1}^{I} q_{MC,z} f(e_z)\big]\\
&+ \delta C^{\bar\pi} V^{\bar\pi}\big[f[X_1 + \bar\pi'(\mu - r)] + \bar\pi'\Sigma\rho'(\Sigma^X)' f_x\big]\\
&- \delta C^{\bar\pi} F\big[f(d_t + \mu_1^X d_{x_1} + \frac{1}{2}(\Sigma^X_{11})^2 d_{x_1x_1}) + f_x'\Sigma^X(\Sigma^X)' d_x\big]\\
&- \delta(1-\delta)\big[\frac{1}{2}\left(V^{\bar\pi}\right)^2 f\bar\pi'\Sigma\Sigma'\bar\pi - V^{\bar\pi} F f \bar\pi'\Sigma\rho'(\Sigma^X)' d_x + \frac{1}{2} f(F\Sigma^X_{11} d_{x_1})^2\big] = 0.
\end{aligned} \tag{D.3}$$

Exploiting (D.3), the PDE for d (6.39) and the definition of $\bar\pi$ (6.25) we obtain $(*) = 0$. Together, $\mu^L = 0$. Thus, L is a local martingale. As it is positive, it is a supermartingale.

Now we apply the concavity of the utility function, the supermartingale property of L and the martingale property of Φ to prove the statement:

$$\begin{aligned}
&\mathbb{E}\big[U_H\big(V^{\pi}(T)\big)\big|\mathscr{F}_t\big] \le \mathbb{E}\big[U_H\big(V^{\bar\pi}(T)\big) + (U_H)_v\big(V^{\bar\pi}(T)\big)\big(V^{\pi}(T) - V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big]\\
&= \mathbb{E}\big[U_H\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big] + \mathbb{E}\big[(U_H)_v\big(V^{\bar\pi}(T)\big)\big(V^{\pi}(T) - FD(T,T,X_1(T))\big)\big|\mathscr{F}_t\big]\\
&- \mathbb{E}\big[(U_H)_v\big(V^{\bar\pi}(T)\big)\big(V^{\bar\pi}(T) - FD(T,T,X_1(T))\big)\big|\mathscr{F}_t\big]\\
&= \mathbb{E}\big[U_H\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big] + \mathbb{E}\big[\frac{\alpha}{(1-\delta)^{\delta-1}} L(T)\big|\mathscr{F}_t\big]\\
&- \mathbb{E}\big[\delta\Phi\big(T, V^{\bar\pi}(T), X(T), \mathcal{MC}(T)\big)\big|\mathscr{F}_t\big]\\
&\le \mathbb{E}\big[U_H\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big] + \frac{\alpha}{(1-\delta)^{\delta-1}} L(t) - \delta\Phi\big(t,v,x,e_i\big)\\
&= \mathbb{E}\big[U_H\big(V^{\bar\pi}(T)\big)\big|\mathscr{F}_t\big].
\end{aligned}$$

□

Proof of Proposition 6.9

We apply Proposition 6.8. Note that for the considered model specifications System (6.27) takes the following form:

$$\begin{aligned}
&\frac{\partial}{\partial t}B_1(t) - B_1(t)\kappa_1 + \delta = 0\\
&\frac{\partial}{\partial t}B_2(t) - B_2\underbrace{\left(\kappa_2 - \frac{\delta}{1-\delta}\chi_2\bar\rho\lambda_2\right)}_{=\tilde\kappa_2} + \frac{1}{2}B_2^2\underbrace{\chi_2^2\big(1 + \frac{\delta}{1-\delta}\bar\rho^2\big)}_{=\tilde\chi_2^2} + \frac{1}{2}\frac{\delta}{1-\delta}\lambda_2^2 = 0,
\end{aligned} \tag{D.4}$$

with $B_1(T) = 0$ and $B_2(T) = 0$. The solution for B_1 is trivial and the solution for B_2 follows from Corollary 2.44, where Conditions (6.42) and (6.43) assure its applicability. Now, adopting the notation from Proposition 6.8, we define $Z := (X_2^m, X_1^m, G)$. A straightforward substitution of the model definition in the expressions for μ^Z and Γ^Z from the proof of Proposition 6.8 and Proposition 5.6 shows that conditions i)-iv) from Theorem 2.34 are fulfilled for $p = 1$. The statement follows directly.

□

List of Tables

List of Figures

Bibliography

[1] Y. Aït-Sahalia and R. Kimmel. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83(2):413–452, 2007.

[2] W. J. Anderson. *Continuous-Time Markov Chains.* Springer-Verlag, New York, 1st edition, 1991.

[3] M. Andersson, E. Krylova, and S. Vähämaa. Why does the correlation between stock and bond returns vary over time? *Applied Financial Economics*, 18(2): 139–151, 2008.

[4] A. Ang and G. Bekaert. Stock return predictability: Is it there? *The Review of Financial Studies*, 20(3):651–707, 2007.

[5] S. H. Babbs and K. B. Nowman. Kalman filtering of generalized Vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34 (1):115–130, 1999.

[6] L. Baele, G. Bekaert, and K. Inghelbrecht. The determinants of stock and bond return comovements. *The Review of Financial Studies*, 23(6):2374–2428, 2010.

[7] N. Bäuerle and Z. Li. Optimal portfolios for financial markets with Wishart volatility. *Journal of Applied Probability*, 50(4):1025–1043, 2013.

[8] N. Bäuerle and U. Rieder. Portfolio optimization with Markov-modulated stock prices and interest rates. *Automatic Control*, 49(3):442–447, 2004.

[9] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, 41(1):164–171, 1970.

[10] R. Bellman. *Dynamic Programming.* Princeton University Press, Princeton, New Jersey, 1st edition, 1957.

[11] G. Bernhart, S. Höcht, M. Neugebauer, M. Neumann, and R. Zagst. Asset correlations in turbulent markets and the impact of different regimes on asset management. *Asia-Pacific Journal of Operational Research*, 28(1):1–23, 2011.

[12] N. Bingham and R. Kiesel. *Risk-Neutral Valuation.* Springer-Verlag, London, 2nd edition, 2004.

[13] F. Black and R. Jones. Simplifying portfolio insurance. *The Journal of Portfolio Management*, 14(1):48–51, 1987.

[14] F. Black and R. Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43, 1992.

[15] F. Black and A. Perold. Theory of constant proportion portfolio insurance. *Journal of Economic Dynamics & Control*, 16(3-4):403–426, 1992.

[16] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[17] D. J. Bolder. Affine term-structure models: Theory and implementation. *Bank of Canada Working Paper No. 2001-15*, 2001.

[18] T. Bollerslev, R. F. Engle, and J. M. Wooldridge. A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, 96(1):116–131, 1988.

[19] S. Brendle. Portfolio selection under incomplete information. *Stochastic Processes and their Applications*, 116(5):701–723, 2006.

[20] M. J. Brennan. The role of learning in dynamic portfolio decisions. *European Finance Review*, 1(3):295–306, 1998.

[21] M. J. Brennan and Y. Xia. Stochastic interest rates and the bond-stock mix. *European Finance Review*, 4(2):197–210, 2000.

[22] D. P. Brown and M. R. Gibbson. A simple econometric approach for utility-based asset pricing models. *Journal of finance*, 40(2):359–381, 1985.

[23] A. Buraschi, P. Porchia, and F. Trojani. Correlation risk and optimal portfolio choice. *The Journal of Finance*, 65(1):393–420, 2010.

[24] J. Y. Campbell and L. M. Viceira. Consumption and portfolio decisions when expected returns are time varying. *The Quarterly Journal of Economics*, 114 (2):433–495, 1999.

[25] E. Çanakoğlu and S. Özekici. Portfolio selection in stochastic markets with HARA utility functions. *European Journal of Operational Research*, 201(2): 520–536, 2010.

[26] E. Çanakoğlu and S. Özekici. HARA frontiers of optimal portfolios in stochastic markets. *European Journal of Operational Research*, 221(1):129–137, 2012.

[27] R. Chen and L. Scott. Multi-factor cox-ingersoll-ross models of the term structure: Estimates and tests from a kalman filter model. *Journal of Real Estate Finance and Economics*, 27(2):143–172, 2003.

[28] S. Chen. Predicting the bear stock market: Macroeconomic variables as leading indicators. *Journal of Banking & Finance*, 33(2):211–223, 2009.

[29] S. Choi. Regime-switching univariate diffusion models of the short-term interest rate. *Studies in Nonlinear Dynamics & Econometrics*, 13(1):1–41, 2009.

[30] S. Choi and D. Yuan. Continuous time stochastic volatility models with regime shifts. *Working paper*, 2013.

[31] R. Cont and J. da Fonseca. Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60, 2002.

[32] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–408, 1985.

[33] G. Deelstra and G. Parker. A covariance equivalent discretisation of the CIR model. *Proceedings of the 5th AFIR International Colloquium*, 2:731–747, 1995.

[34] A. A. Dragulescu and V. M. Yakovenko. Probability of distribution of returns in the Heston model with stochastic volatility. *Quantitative Finance*, 2:443–453, 2002.

[35] G. Durham and Y. Park. Beyond stochastic volatility and jumps in returns and volatility. *Journal of Business & Economic Statistics*, 31(1):107–121, 2013.

[36] A. Elices. Models with time-dependent parameters using transform methods: application to Heston's model. *Working paper*, 2007.

[37] R. Elliott and K. Nishide. Pricing of discount bonds with a Markov switching regime. *Annals of Finance*, 10(3):509–522, 2014.

[38] R. J. Elliott and C. U. Osakwe. Option pricing for pure jump processes with Markov switching compensators. *Finance and stochasticss*, 10(2):250–275, 2006.

[39] R. J. Elliott, L. Aggoun, and J. Moore. *Hidden Markov Models*. Springer-Verlag, New York, 1st edition, 1995.

[40] R. J. Elliott, K. Siu, and L. Chan. Pricing volatility swaps under Heston's stochastic volatility model with regime switching. *Applied Mathematical Finance*, 14(1):41–62, 2007.

[41] R. Engle. What good is a volatility model? *Quantitative Finance*, 1(2): 237–245, 2001.

[42] C. Erlwein and R. Mamon. An online estimation scheme for a Hull-White model with HMM-driven parameters. *Statistical Methods and Applications*, 18(1):87–107, 2009.

[43] M. Escobar and P. Olivares. Pricing of mountain range derivatives under a principal component stochastic volatility model. *Applied Stochastic Models in Business and Industry*, 29(1):31–44, 2013.

[44] M. Escobar, B. Götz, and R. Zagst. Pricing of a CDO on stochastically correlated underlyings. *Quantitative Finance*, 10(3):265–277, 2010.

[45] M. Escobar, B. Götz, D. Neykova, and R. Zagst. Stochastic correlation and volatility mean-reversion - empirical motivation and derivatives pricing via perturbation theory. *Applied Mathematical Finance*, 21(6):555–594, 2014.

[46] M. Escobar, B. Götz, D. Neykova, and R. Zagst. Pricing two-asset barrier options under stochastic correlation via perturbation. *International Journal of Theoretical and Applied Finance*, 18(3):1–44, 2015.

[47] M. Escobar, D. Neykova, and R. Zagst. Portfolio optimization in affine models with Markov switching. *International Journal of Theoretical and Applied Finance*, 18(5):1–46, 2015.

[48] W. Fei. Optimal consumption and portfolio under inflation and Markovian switching. *Stochastics An International Journal of Probability and Stochastic Processes*, 85(2):272–285, 2013.

[49] W. Feller. *An Introduction to Probability Theory and its Applications, Vol I.* Wiley, New York, 1st edition, 1950.

[50] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, New York, 1st edition, 1993.

[51] J. Fouque, G. Papanicolaou, R. Sircar, and K.Solna. *Multiscale Stochastic Volatility for Equity, Interest-Rate and Credit Derivatives.* Cambridge University Press, New York, 1st edition, 2011.

[52] K. Frauendorfer, U. Jacoby, and A. Schwendener. Regime switching based portfolio selection for pension funds. *Journal of Banking & Finance*, 31(8): 2265–2280, 2007.

[53] J. D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.

[54] M. R. Hardy. A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 4(2):41–53, 2001.

[55] A. Harvey. *Forecasting, structural time series models and the Kalman Filter.* Cambridge University Press, Cambridge, 1st edition, 1989.

[56] H. Hata and J. Sekine. Risk-sensitive asset management under a Wishart autoregressive factor model. *Journal of Mathematical Finance*, 3(1A):222–229, 2013.

[57] J. Hauptmann, A. Hoppenkamps, A. Min, F. Ramsauer, and R. Zagst. Forecasting market turbulence using regime-switching models. *Financial Markets and Portfolio Management*, 28(2):139–164, 2014.

[58] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Journal of Finance*, 42:327–343, 1993.

[59] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, Heidelberg, 2nd edition, 2003.

[60] N. Johnson, V. Naik, S. Page, N. Pedersen, and S. Sapra. The stock-bond correlation. *Pimco Quantitative Research (Working paper)*, 2013.

[61] J. Kallsen and A. Cerny. Mean-variance hedging and optimal investment in Heston's model with correlation. *Mathematical Finance*, 18(3):473–492, 2008.

[62] J. Kallsen and J. Muhle-Karbe. Exponentially affine martingales, affine measure changes and exponential moments of affine processes. *Stochastic Processes and their Applications*, 120(2):163–181, 2010.

[63] J. Kallsen and J. Muhle-Karbe. Utility maximization in affine stochastic volatility models. *International Journal of Theoretical and Applied Finance (IJTAF)*, 13(03):459–477, 2010.

[64] R. E. Kalman. New approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[65] R. E. Kalman and R. S. Buchy. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, 83(Series D):95–107, 1961.

[66] I. Karatzas and S. Shreve. *Methods of Mathematical Finance*. Springer-Verlag, New York, 1st edition, 1998.

[67] T. S. Kim and E. Omberg. Dynamic nonmyopic portfolio behavior. *Review of Financial Studies*, 9(1):141–161, 1996.

[68] A. Klenke. *Probability Theory*. Springer-Verlag, London, 2nd edition, 2004.

[69] K. Königsberger. *Analysis 1*. Springer-Verlag, Berlin Heidelberg, 6st edition, 2004.

[70] K. Königsberger. *Analysis 2*. Springer-Verlag, Berlin Heidelberg, 5th edition, 2004.

[71] R. Korn. *Optimal Portfolios: Stochastic Models For Optimal Investment And Risk Management In Continuous Time*. World Scientific, Singapore, 1st edition, 1997.

[72] R. Korn. Optimal portfolios with a positive lower bound on final wealth. *Quantitative finance*, 5(3):315–321, 2005.

[73] R. Korn and H. Kraft. A stochastic control approach to portfolio problems with stochastic interest rates. *SIAM Journal on Control and Optimization*, 40 (4):1250–1269, 2001.

[74] H. Kraft. Optimal portfolios and Heston's stochastic volatility model. *Quantitative Finance*, 5(3):303–313, 2005.

[75] J. Liu. Portfolio selection in stochastic environments. *The review of financial studies*, 20(1):1–39, 2007.

[76] R. H. Liu. Optimal investment and consumption with proportional transaction costs in regime-switching model. *Journal of Optimization Theory and Applications*, 163(2):614–641, 2014.

[77] R. H. Liu. A finite-horizon optimal investment and consumption problem using regime-switching models. *International Journal of Theoretical and Applied Finance*, 17(4), 2014.

[78] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics*, 7(4):649–673, 1954.

[79] J. M. Maheu and T. H. McCurdy. Identifying bull and bear markets in stock returns. *Journal of Business & E onomic Statistics*, 18(1):100–112, 2000.

[80] R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51(3):247–257, 1969.

[81] R. C. Merton. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory*, 3(4):373–413, 1971.

[82] S. Mikhailov and U. Nogel. Heston's stochastic volatility model: Implementation, calibration and some extensions. *Working paper*, 2004.

[83] S. Mitra. Regime switching stochastic volatility option pricing. *Journal of Financial Markets and Derivatives*, 1(2):213–242, 2010.

[84] D. Neykova, M. Escobar, and R. Zagst. Optimal investment in multidimensional Markov-modulated affine models. *Annals of Finance*, 11(3):503–530, 2015.

[85] J. Norris. *Markov chains.* Cambridge University Press, Cambridge, 1st edition, 1997.

[86] B. Øksendal. *Stochastic Differential Equations.* Springer-Verlag, Berlin, 5th edition, 2000.

[87] A. Papanicolaou and R. Sircar. A regime-switching Heston model for VIX and S&P 500 implied volatilities. *Quantitative Finance*, 19(2):1–7, 2013.

[88] A. F. Perold. Constant proportion portfolio insurance. *Harvard Business School (unpublished manuscript)*, 1986.

[89] A. F. Perold and W. F. Sharpe. Dynamic strategies for asset allocation. *Financial Analysts Journal*, 44(1):16–27, 1988.

[90] J. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2): 122–136, 1964.

[91] M. Rubinstein. Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978. *The Journal of Finance*, 40(2):455–480, 1985.

[92] A. Sangvinatsos and J. A. Wachter. Does the failure of the expectations hypothesis matter for long-term investors? *Journal of Finance*, 60(1):179–230, 2005.

[93] T. Santos and P. Veronesi. Labor income and predictable stock returns. *The Review of Financial Studies*, 19(1):1–44, 2006.

[94] B. f. I. Settlements. *Zero-coupon yield curves: technical documentation.* Bank for International Settlements, 2005. URL `http://EconPapers.repec.org/RePEc:bis:bisbps:25`.

[95] L. R. Sotomayor and A. Cadenillas. Explicit solutions of consumption-investment problems in financial markets with regime switching. *Mathematical Finance*, 19(2):251–279, 2009.

[96] E. M. Stein and J. C. Stein. Stock price distributions with stochastic volatility: an analytic approach. *The Review of Financial Studies*, 4(4):727–752, 1991.

[97] L. Teplá. Optimal investment with minimum performance constraints. *Journal of Economic Dynamics & Control*, 25(10):1629–1645, 2001.

[98] A. Timmermann. Moments of Markov switching models. *Journal of Econometrics*, 96(1):75–111, 2000.

[99] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, 1977.

[100] J. Wachter. Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal Financial and Quantitative Analysis*, 37(1):63–91, 2002.

[101] G. Yin and Q. Zhang. *Continuous-Time Markov Chains and Applications.* Springer-Verlag, New York, 1st edition, 1998.

[102] R. Zagst. *Interest rate management*. Springer-Verlag, Berlin, 1st edition, 2002.

[103] R. Zagst and J. Kraus. Stochastic dominance of portfolio insurance strategies. *Annals of Operations Research*, 185(1):75–103, 2011.

[104] T. Zariphopoulou. A solution approach to valuation with unhedgeable risks. *Finance and Stochastics*, 5(1):61–82, 2001.

[105] N. Zhou and R. Mamon. An accessible implementation of interest rate models with Markov-switching. *Expert Systems with Applications*, 39(5):4679–4689, 2012.

[106] W. Zucchini and I. L. MacDonald. *Hidden Markov Models for Time Series: An introduction using R*. Chapman and Hall/CRC Press, Boka Raton, 1st edition, 2009.