TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DIPLOMARBEIT

The Multidimensional Heston Stochastic Volatility Model based on Wishart Processes

Ausgeführt am Institut für
Finanz- und Versicherungsmathematik
der Technischen Universität Wien

unter der Anleitung von
Univ.Prof. Dipl.-Math. Dr.rer.nat. Thorsten Rheinländer
und der Betreuung von
Univ.Ass. Dipl.-Ing. Dr.sc. Christa Cuchiero

durch

Laura-Maria Orth
Brunnengasse 64/11
2230 Gänserndorf

Wien, 10. September 2014 ______________________

Affidavit

With this statement, I declare that this academic thesis:

was written entirely by me, without the use of any sources other than those indicated and without the use of any unauthorized resources;

has never been submitted in any form for evaluation as an examination paper in Austria or any other country.

Acknowledgment

At this point of my thesis, I would like to show my gratitude to Univ.Prof.Dipl.-Math.Dr.rer.nat. Thorsten Rheinländer for his assistance making this thesis possible.
Most of all, I would like to express my thanks to Univ.Ass.Dipl.-Ing.Dr.sc. Christa Cuchiero for her excellent supervisory. She always supported me in a friendly and patient manner.
I deeply appreciate her engagement, efforts and expertise during the learning process of my thesis.

Moreover, I would like to thank my family, friends and colleagues at university for their support, making my years of study unforgettable.

Abstract

In this thesis we study the one-dimensional Heston stochastic volatility model and its multivariate extension, the so-called multidimensional Heston stochastic volatility model.
While the variance process equals the well-known Cox-Ingersoll-Ross process in the one-dimensional setup, the covariance process of the multidimensional Heston stochastic volatility model follows the Wishart process introduced by M.-F. Bru. Both stochastic volatility models belong to the class of affine models whose characterizing property is the exponential affine form of the conditional characteristic function.
A detailed study of the multidimensional Heston stochastic volatility model is given, and interesting relations between the marginals of the multidimensional Heston stochastic volatility model and the one-dimensional Heston stochastic volatility model are derived.
Moreover, small time asymptotics for the implied volatility of a call written on one asset in the multidimensional Heston stochastic volatility model are considered. We show that an expansion of the asymptotic implied volatility proved by M. Forde and A. Jacquier for the one-dimensional Heston stochastic volatility model can be extended to the multidimensional one.

Keywords: *Heston model, multidimensional affine model, Wishart process, stochastic volatility, implied volatility*

Kurzfassung

Diese Diplomarbeit behandelt das eindimensionale stochastische Volatilitätsmodell von Heston („Heston Stochastic Volatility“ Modell genannt) und dessen multivariate Erweiterung, das mehrdimensionale Heston Stochastic Volatility Modell.
Während im eindimensionalen Fall der Varianzprozess durch einen Cox-Ingersoll-Ross Prozess gegeben ist, folgt im mehrdimensionalen Heston Stochastic Volatility Modell der Kovarianzprozess einem Wishart Prozess, der erstmals von M.-F. Bru definiert wurde. Beide stochastischen Volatilitätsmodelle gehören zur Klasse der affinen Modelle, die als charakterisierende Eigenschaft eine bedingte charakteristische Funktion von exponentiell affiner Form aufweisen.
Diese Arbeit umfasst eine detaillierte Analyse des mehrdimensionalen Heston Stochastic Volatility Modells, wobei unter anderem auch interessante Beziehungen zwischen den Marginalen des mehrdimensionalen Heston Stochastic Volatility Modells und des eindimensionalen Heston Stochastic Volatility Modells hergeleitet werden.
Des Weiteren wird die Kurzzeitasymptotik der impliziten Volatilität einer Kaufoption auf ein Asset im eindimensionalen Heston Stochastic Volatility Modell betrachtet. Dabei wird gezeigt, dass die Darstellung der asymptotischen impliziten Volatilität, die von M. Forde und A. Jacquier für das eindimensionale Heston Stochastic Volatility Modell gezeigt wurde, auch auf den mehrdimensionalen Fall erweitert werden kann.

Schlagwörter: *Heston Modell, mehrdimensionales affines Modell, Wishart Prozess, stochastische Volatilität, implizite Volatilität*

Contents

Chapter 0

Introduction

The aim of my thesis is to introduce and study the concept of the multidimensional Heston stochastic volatility model. As the name already indicates, the multidimensional Heston stochastic volatility model is a multivariate generalization of the well-known one-dimensional Heston stochastic volatility model introduced by Steven L. Heston. In a one-dimensional Heston stochastic volatility model only one asset can be considered, in a multidimensional framework however, a vector of different assets and their correlations can be modeled and examined. While the Cox-Ingersoll-Ross process is chosen as variance process in the one-dimensional Heston stochastic volatility model, it is replaced by the so-called Wishart process in the multidimensional Heston stochastic volatility model. The Wishart process can be interpreted as a matrix extension of the CIR process.

In general, stochastic volatility models are used to model the volatility in such a way that it follows or coincides with certain phenomenons in the market: While in the original famous Black-Scholes model, it is assumed that the volatility is given as a constant, one sees, however, that in reality the implied volatility of a derivative depends on its strike and maturity which is presented by the so-called volatility surface: A possible effect is the volatility smile or skew. That is the reason why local and stochastic volatility models were introduced and studied. These models take account of the above mentioned effects.

An import property of the one-dimensional and the multidimensional Heston stochastic volatility model is that they both belong to the class of affine processes. Affine processes are characterized by their special form of their (conditional) characteristic function, namely the characteristic function of an affine process is exponential affine in its initial state. One of the advantage of affine models is that there exist semi-closed-form solutions for pricing of derivatives. Therefore affine processes often appear in financial mathematics.

The structure of my thesis is as follows:

Chapter 1: Affine Processes

At the beginning of this chapter, affine diffusion processes are introduced. As already mentioned, the characterizing property of affine processes is their exponential affine conditional characteristic function. Moreover it is mentioned that the exponent of this characteristic function is determined by solutions of so-called Riccati equations. Then we state that if a stochastic process is affine, its drift and diffusion matrix are also affine. Furthermore, affine processes on the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$ are discussed. In the last section of this chapter, discounting and pricing in affine models on the canonical state space are explained. The main result of this section is a theorem about Fourier pricing in an affine model.

Chapter 2: One-Dimensional Heston Stochastic Volatility Model

The beginning of Chapter 2 deals with the definition and some properties of the one-dimensional Heston stochastic volatility model. It is explained that the variance process in the one-dimensional Heston stochastic volatility model follows a CIR process. Then it is also shown that there exists a link between the CIR process and the Ornstein-Uhlenbeck process.
In the second part of Chapter 2, the Riccati equations for the one-dimensional Heston stochastic volatility model are studied by applying results about Riccati equations of affine processes from Chapter 1. In addition to that, explicit solutions of the Riccati equations are derived for this one-dimensional model. As an application, the joint characteristic function of the asset price and the variance process is given. At the end of this chapter, the pricing of a European call option is explained by applying the theorem about Fourier pricing for affine processes stated in Chapter 1.

Chapter 3: Wishart Processes

Chapter 3 of my thesis can be split into two parts: At first Wishart processes in the sense of Marie-France Bru are introduced. This first part is mainly based on [3] and summarizes its main steps. Then we introduce the kind of Wishart processes studied and examined in this thesis. The Wishart process follows a so-called Wishart distribution which is motivated and introduced in a separate subsection. As a next step the CIR-process as a special case of the Wishart process is considered. Finally, the characteristic function of the Wishart process is studied where its affine property is emphasized.

Chapter 4: Multivariate Heston Stochastic Volatility Model

The beginning of Chapter 4 covers multivariate affine stochastic volatility models and its conditional characteristic function in general. Then the multidimensional Heston stochastic volatility model is defined and its conditional characteristic function is given. Again the affine property of the characteristic function is highlighted.

Chapter 5: Relationships and Results concerning the Multidimensional Heston Stochastic Volatility Model

The main objective of this chapter is to investigate properties of the multidimensional Heston stochastic volatility model. So at first the multidimensional Heston stochastic volatility model is examined component-by-component. To be more precise, the k-th entry of the vector of the log price process in combination with the (k, k)-th entry of the Wishart process represents a "new" model which we call marginal of the multidimensional Heston stochastic volatility model. The second section of this chapter covers the relation between these marginals of the multidimensional Heston stochastic volatility model and the one-dimensional Heston stochastic volatility model. The interesting result is that if we choose the deterministic matrix M to be the null matrix or more generally of diagonal form, the dynamics of the marginals of the multidimensional Heston stochastic volatility model simplify to the one-dimensional Heston stochastic volatility model. Since it is a well-known fact that the drift of a stochastic process can be changed by applying Girsanov's theorem and hence changing the probability measure, we also prove that possible way of obtaining a one-dimensional Heston stochastic volatility model.

Chapter 6: Small-Time Asymptotics for Implied Volatility

In this last chapter of my thesis, the small-time asymptotics for implied volatility are investigated. After repeating the concept of implied volatility in the first section of this chapter, we summarize in the second chapter the results about the small-time asymptotics for implied volatility for the one-dimensional Heston stochastic volatility model published by Martin Forde and Antoine Jacquier in [8]. Finally, in the last section of this chapter we show that these results are also valid for a more general framework, namely the marginals of the multidimensional Heston stochastic volatility model as defined in Chapter 5.

Chapter 1

Affine Processes

The aim of this chapter is to give the reader a short overview of affine diffusion processes which are of major importance in the subsequent chapters and sections. These main results on affine diffusions are taken from “Chapter 10” of the book “Term-Structure Models” published by Damir Filipović (see [5]).
Most of the theory of affine processes given in Chapter 10 of [5] can also be found in the paper [6] written by Damir Filipović and Eberhard Mayerhofer.

1.1 Vector-Valued Affine Processes in General

For the definition of affine processes we follow [5, p.143]:
Denote by $\mathcal{D} \subset \mathbb{R}^d$ a closed state space with non-empty interior, where $d \geq 1$ is some fixed dimension. Furthermore let $b : \mathcal{D} \to \mathbb{R}^d$ be a continuous function and let $\rho : \mathcal{D} \to \mathbb{R}^{d\times d}$ be a measurable function with $a(x) = \rho(x)\rho(x)^\top$ being continuous in $x \in \mathcal{D}$. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a filtered probability space and $W_t = (W_{t,1}, \ldots, W_{t,d})$ be a d-dimensional Brownian motion defined on that probability space. In this framework we assume that for every $x \in \mathcal{D}$ there exists a unique solution $\mathcal{X} = \mathcal{X}^x$ with $\mathcal{X}_0 = x$ of the stochastic differential equation

$$d\mathcal{X}_t = b(\mathcal{X}_t)dt + \rho(\mathcal{X}_t)dW_t. \tag{1.1}$$

Remark 1.1.1. $a(x)$ is called diffusion matrix, whereas $b(x)$ is called drift of the stochastic process.

Now we are ready to give a definition of vector-valued affine processes.

Definition 1.1.2. $\mathcal{X}$ is *affine* if the $\mathcal{F}_t$-conditional characteristic function of $\mathcal{X}_T$ is exponential affine in $\mathcal{X}_t$, for all $t \leq T$. Mathematically this means that there exist $\mathbb{C}$- and $\mathbb{C}^d$- valued functions $\phi(t,u)$ and $\psi(t,u)$ with jointly continuous t-derivatives such that $\mathcal{X} = \mathcal{X}^x$ fulfills

$$\mathbb{E}\left[e^{u^\top \mathcal{X}_T}\middle|\mathcal{F}_t\right] = e^{\phi(T-t,u)+\psi(T-t,u)^\top \mathcal{X}_t} \tag{1.2}$$

$\forall u \in i\mathbb{R}^d, t \leq T$ and $x \in \mathcal{D}$.

As mentioned already above, hereafter important theorems concerning affine processes are stated without proof. Interested readers can find the corresponding proofs in "Chapter 10: Affine Processes" of the book "Term-Structure Models" by Damir Filipović (see [5]).

Theorem 1.1.3. *(see [5, Theorem 10.1]) Suppose $\mathcal{X}$ is affine. Then the diffusion matrix $a(x)$ and drift $b(x)$ are affine in x. That is,*

$$\begin{aligned} a(x) &= a + \sum_{i=1}^{d} x_i \alpha_i, \\ b(x) &= b + \sum_{i=1}^{d} x_i \beta_i = b + \mathcal{B}x \end{aligned} \tag{1.3}$$

for some $d \times d$-matrices a and α_i and d-vectors b and β_i, where we denote by $\mathcal{B} = (\beta_1, \dots, \beta_d)$ the $d \times d$-matrix with i-th column vector $\beta_i, 1 \leq i \leq d$. Moreover, ϕ and $\psi = (\psi_1, \dots, \psi_d)^\top$ solve the system of Riccati equations

$$\begin{aligned} \frac{\partial}{\partial t}\phi(t,u) &= \frac{1}{2}\psi(t,u)^\top a \psi(t,u) + b^\top \psi(t,u), \\ \phi(0,u) &= 0, \\ \frac{\partial}{\partial t}\psi_i(t,u) &= \frac{1}{2}\psi(t,u)^\top \alpha_i \psi(t,u) + \beta_i^\top \psi(t,u), \quad 1 \leq i \leq d, \\ \psi(0,u) &= u. \end{aligned} \tag{1.4}$$

In particular, ϕ is determined by ψ via simple integration:

$$\phi(t,u) = \int_0^t \left(\frac{1}{2}\psi(s,u)^\top a \psi(s,u) + b^\top \psi(s,u)\right) ds.$$

Conversely, suppose the diffusion matrix $a(x)$ and drift $b(x)$ are affine of the form (1.3) *and suppose there exists a solution (ϕ, ψ) of the Riccati equations* (1.4) *such that*
$\phi(t,u) + \psi(t,u)^\top x$ has a nonpositive real part for all $t \geq 0, u \in i\mathbb{R}^d$ and $x \in \mathcal{D}$. Then $\mathcal{X}$ is affine with conditional characteristic function (1.2).

Moreover we state a lemma about the global existence and uniqueness of the system of Riccati equations:
Before stating the lemma, let K be a placeholder for either $\mathbb{R}$ or $\mathbb{C}$.

Lemma 1.1.4. *(see [5, Lemma 10.1]) Consider the system of ordinary differential equations*

$$\begin{aligned} \frac{\partial}{\partial t} f(t,u) &= R(f(t,u)), \\ f(0,u) &= u, \end{aligned} \tag{1.5}$$

where $R : K^d \to K^d$ is a locally Lipschitz continuous function. Then the following holds:

- *For every $u \in K^d$, there exists a life time $t_+(u) \in (0, \infty]$ such that there exists a unique solution $f(\cdot, u) : [0, t_+(u)) \to K \times K^d$ of* (1.5)*.*
- *The domain $\mathcal{G}_K = \{(t,u) \in \mathbb{R}_+ \times K^d | t < t_+(u)\}$ is open in $\mathbb{R}_+ \times K^d$ and maximal in the sense that either $t_+(u) = \infty$ or $\lim_{t \nearrow t_+(u)} \|f(t,u)\| = \infty$ respectively, for all $u \in K^d$.*
- *For every $t \geq 0$, the t-section $\mathcal{G}_K(t) = \{u \in K^d | (t,u) \in \mathcal{G}_K\}$ is open in K^d, and non-expanding in t in the following sense: $K^d = \mathcal{G}_K(0) \supseteq \mathcal{G}_K(t_1) \supseteq \mathcal{G}_K(t_2)$, $0 \leq t_1 \leq t_2$. In fact, we have $f(s, \mathcal{G}_K(t_2)) \subseteq \mathcal{G}_K(t_1)$ for all $s \leq t_2 - t_1$.*

1.2 Vector-Valued Affine Processes on the Canonical State Space

The parameters a, α_i, b, β_i can be further specified according to the state space $\mathcal{D}$:

- the parameters have to fulfill conditions that guarantee that the process $\mathcal{X}$ does not leave the state space $\mathcal{D}$
- the parameters a and α_i determining the diffusion matrix must be defined in such a way that $a(x) = a + \sum_{i=1}^{d} x_i \alpha_i \in \mathcal{S}_d^+, \forall x \in \mathcal{D}$, where $\mathcal{S}_d^+$ denotes the set of symmetric and positive semidefinite matrices

We shall pay particular attention to the so-called canonical state space which is given by

$$\mathcal{D} = \mathbb{R}_+^m \times \mathbb{R}^n,$$

for some integers $m, n \geq 0$ with $m + n = d$.

Hereafter, for the canonical state space, two index sets are defined by $I = \{1, \ldots, m\}$ and $J = \{m+1, \ldots, m+n\}$.
In addition to that, let $\mu_M = (\mu_i)_{i \in M}$ and $\nu_{MN} = (\nu_{ij})_{i \in M, j \in N}$ be the sub-vector and sub-matrix of an arbitrary vector μ or an arbitrary matrix ν for any index sets M or N.

With this setup of the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$ a complete characterization of the parameters of affine processes is possible, which is subject of the main results of the following theorem:

Theorem 1.2.1. *(see [5, Theorem 10.2]) A process of the form* (1.1) *on the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$ is affine if and only if $a(x)$ and $b(x)$ are affine of the form* (1.3) *for parameters*

a, α_i, b, β_i which are admissible in the following sense:

$$\begin{aligned}
& a, \alpha_i \text{ are symmetric positive semidefinite}, \\
& a_{II} = 0 \text{ (and thus } a_{IJ} = a_{JI}^\top = 0), \\
& \alpha_j = 0 \quad \forall j \in J, \\
& \alpha_{i,kl} = \alpha_{i,lk} = 0 \quad \text{for } k \in I \backslash \{i\}, \forall 1 \leq i, l \leq d, \\
& b \in \mathbb{R}_+^m \times \mathbb{R}^n, \\
& \mathcal{B}_{IJ} = 0, \\
& \mathcal{B}_{II} \text{ has nonnegative off-diagonal elements.}
\end{aligned} \tag{1.6}$$

In this case, the corresponding system of Riccati equations (1.4) *simplifies to*

$$\begin{aligned}
\frac{\partial}{\partial t}\phi(t,u) &= \frac{1}{2}\psi_J(t,u)^\top a_{JJ}\psi_J(t,u) + b^\top \psi(t,u), \\
\phi(0,u) &= 0, \\
\frac{\partial}{\partial t}\psi_i(t,u) &= \frac{1}{2}\psi(t,u)^\top \alpha_i \psi(t,u) + \beta_i^\top \psi(t,u), \quad i \in I, \\
\frac{\partial}{\partial t}\psi_J(t,u) &= \mathcal{B}_{JJ}^\top \psi_J(t,u), \\
\psi(0,u) &= u,
\end{aligned} \tag{1.7}$$

and there exists a unique global solution $(\phi(\cdot,u), \psi(\cdot,u)) : \mathbb{R}_+ \to \mathbb{C}_- \times \mathbb{C}_-^m \times i\mathbb{R}^n$ for all initial values $u \in \mathbb{C}_-^m \times i\mathbb{R}^n$. In particular, the equation for ψ_J forms an autonomous linear system with unique global solution $\psi_J(t,u) = e^{\mathcal{B}_{JJ}^\top t} u_J$ for all $u_J \in \mathbb{C}^n$.

We now proceed by formulating an important theorem about the existence of affine processes:

Theorem 1.2.2. *(see [5, Theorem 10.8]) Let a, α_i, b, β_i be admissible parameters. Then there exists a measurable function $\rho : \mathbb{R}_+^m \times \mathbb{R}^n \to \mathbb{R}^{d \times d}$ with $\rho(x)\rho(x)^\top = a + \sum_{i \in I} x_i \alpha_i$ and such that, for any $x \in \mathbb{R}_+^m \times \mathbb{R}^n$, there exists a unique $\mathbb{R}_+^m \times \mathbb{R}^n$-valued solution $\mathcal{X} = \mathcal{X}^x$ of*

$$\begin{aligned}
d\mathcal{X} &= (b + \mathcal{B}\mathcal{X})dt + \rho(\mathcal{X})dW, \\
\mathcal{X}_0 &= x
\end{aligned}$$

and $\rho(x)\rho(x)^\top = a + \sum_{i \in I} x_i \alpha_i$. Moreover, the law of $\mathcal{X}$ is uniquely determined by a, α_i, b, β_i, and does not depend on the particular choice of ρ.

1.3 Discounting and Pricing in Affine Models on the Canonical State Space

This section is mainly based on [5, p.151-155].

Throughout this section it is assumed that the interest rates are given as constants. (In subsequent chapters we set the interest rates to 0.)
With this assumption of deterministic interest rates, the framework is fixed as follows:
Let S be a n-dimensional price process which is a functional of some d-dimensional process $\mathcal{X}$. Hence there exists a function $g : \mathbb{R}^d \to \mathbb{R}^n_+$ such that $g(\mathcal{X}_t) = S_t$. For example, if $g(\mathcal{X}_t) = e^{\mathcal{X}_t}$, $\mathcal{X}_t$ can be interpreted as the log-price of a stock. Moreover, we consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{Q})$, where $\mathbb{Q}$ denotes a risk-neutral probability measure under which the discounted price process S_t is a (local) martingale with respect to $\mathbb{Q}$.

Assume that the process $\mathcal{X}$ is affine on the canonical state space $\mathbb{R}^m_+ \times \mathbb{R}^n$ introduced before with admissible parameters a, α_i, b, β_i specified as in (1.6).

Hereafter, a time horizon $T > 0$ is fixed and a claim with maturity T (called T-claim) is considered. Denote by $f(\mathcal{X}_T)$ the (arbitrary) payoff of the T-claim for a measurable function $f : \mathbb{R}^d \to \mathbb{R}$ which has to fulfill the following integrability condition

$$\mathbb{E}_{\mathbb{Q}}\left[|f(\mathcal{X}_T)|\right] < \infty,$$

where the expectation is taken with respect to the risk-neutral probability measure $\mathbb{Q}$.

Then by risk-neutral valuation, arbitrage-free prices at any time $t \leq T$ of T-claims with payoff $f(\mathcal{X}_T)$ are obtained by

$$\pi(t) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)} f(\mathcal{X}_T) | \mathcal{F}_t\right] = e^{-r(T-t)} \mathbb{E}_{\mathbb{Q}}\left[f(\mathcal{X}_T) | \mathcal{F}_t\right]. \tag{1.8}$$

The goal of the remaining section is now to find an analytical or numerical solvable expression for the pricing formula (1.8).

Some important theorems concerning the pricing of affine processes can now be mentioned (compare [5, p.152-155]):

Before stating the first theorem, some notation has to be introduced:
For any set $U \subset \mathbb{R}^k (k \in \mathbb{N})$, the strip $\mathcal{S}(U)$ is defined by

$$\mathcal{S}(U) = \{z \in \mathbb{C}^k | Re(z) \in U\}$$

in $\mathbb{C}^k$, where $Re(z)$ denotes the real part of z.

Theorem 1.3.1. *(compare [5, Theorem 10.4]) Let $\mathcal{G}_K$ ($K = \mathbb{R}$ or $\mathbb{C}$) denote the maximal domain for the system of Riccati equations in* (1.7)*. Then $\mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}e^{u^T\mathcal{X}_T}\middle|\mathcal{F}_t\right]$ can be expressed as*

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}e^{u^\top\mathcal{X}_T}\middle|\mathcal{F}_t\right] = e^{\phi(T-t,u)-r(T-t)+\psi(T-t,u)^\top\mathcal{X}_t}, \tag{1.9}$$

$\forall u \in \mathcal{S}(\mathcal{G}_{\mathbb{R}}(S)), t \leq T \leq t+S$ *and* $x \in \mathbb{R}_+^m \times \mathbb{R}^n$.

Finally, we state a theorem which makes it possible to price claims analytically, namely the so-called Fourier pricing:

Theorem 1.3.2. *(see [5, Theorem 10.5]) Let $\mathcal{G}_{\mathbb{R}}$ denote the maximal domain for the system of Riccati equations* (1.7)*. Assume that $f : \mathbb{R}^d \to \mathbb{R}$ satisfies*

$$f(x) = \int_{\mathbb{R}^q} e^{(v+iL\lambda)^\top x}\tilde{f}(\lambda)d\lambda \quad dx\text{-}a.s. \tag{1.10}$$

for some $v \in \mathcal{G}_{\mathbb{R}}(T)$ and $d\times q$-matrix L, and some integrable function $\tilde{f} : \mathbb{R}^q \to \mathbb{C}$, for a positive integer $q \leq d$. Then the price (1.8) *is well defined and given by the formula*

$$\pi(t) = \int_{\mathbb{R}^q} e^{\phi(T-t,v+iL\lambda)-r(T-t)+\psi(T-t,v+iL\lambda)^\top\mathcal{X}_t}\tilde{f}(\lambda)d\lambda. \tag{1.11}$$

Proof. The arbitrage-free price can be calculated by

$$\begin{aligned}
\pi(t) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}f(\mathcal{X}_T)|\mathcal{F}_t\right] &= \mathbb{E}_{\mathbb{Q}}\left[\int_{\mathbb{R}^q} e^{-r(T-t)}e^{(v+iL\lambda)^\top\mathcal{X}_T}\tilde{f}(\lambda)d\lambda|\mathcal{F}_t\right] \\
&\overset{Fubini}{=} \int_{\mathbb{R}^q} \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}e^{(v+iL\lambda)^\top\mathcal{X}_T}|\mathcal{F}_t\right]\tilde{f}(\lambda)d\lambda \\
&\overset{(1.9)}{=} \int_{\mathbb{R}^q} e^{\phi(T-t,v+iL\lambda)-r(T-t)+\psi(T-t,v+iL\lambda)^\top\mathcal{X}_t}\tilde{f}(\lambda)d\lambda.
\end{aligned}$$

Fubini can be applied in the third step in the calculations above because

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[|e^{-r(T-t)}f(\mathcal{X}_T)|\middle|\mathcal{F}_t\right] &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}|f(\mathcal{X}_T)|\middle|\mathcal{F}_t\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)}|\int_{\mathbb{R}^q} e^{v^\top\mathcal{X}_T}e^{(iL\lambda)^\top\mathcal{X}_T}\tilde{f}(\lambda)d\lambda|\middle|\mathcal{F}_t\right] \\
&\leq \mathbb{E}_{\mathbb{Q}}\left[\int_{\mathbb{R}^q} e^{-r(T-t)}e^{v^\top\mathcal{X}_T}|\tilde{f}(\lambda)|d\lambda\middle|\mathcal{F}_t\right] < \infty
\end{aligned}$$

by assumption. $\square$

Chapter 2

One-Dimensional Heston Stochastic Volatility Model

2.1 Introduction of the One-Dimensional Heston Stochastic Volatility Model

In this section the famous (one-dimensional) Heston stochastic volatility model by Steven L. Heston is introduced (compare [12]). Briefly, the one-dimensional Heston stochastic volatility model is a generalization of the well-known Black-Scholes model, where the volatility is assumed to be stochastic. This section is mainly based on section "10.3.3 Heston Stochastic Volatility Model" of the book "Term-Structure Models" written by Damir Filipović (see [5, p.166-168]).

The one-dimensional Heston stochastic volatility model belongs to the class of the so-called affine models introduced in the precedent chapter.

Now we give a formal definition of the one-dimensional Heston stochastic volatility model:

Definition 2.1.1. Let $\mathcal{X} = (X, Y)$ be an affine process with state space $\mathbb{R}_+ \times \mathbb{R}$ given by the risk-neutral dynamics

$$\begin{aligned} dX_t &= (k + \kappa X_t)dt + \sigma\sqrt{2X_t}dB_t, \\ dY_t &= (r - X_t)dt + \sqrt{2X_t}\left(\rho dB_t + \sqrt{1-\rho^2}dW_t\right) \end{aligned} \tag{2.1}$$

for some constant parameters $k, \sigma \geq 0, \kappa \in \mathbb{R}$, $\rho \in [-1, 1]$ and two independent Brownian motions B_t and W_t. Then the model (X, Y) determined by its characterizing parameters $(k, \kappa, \sigma, \rho)$ is called *one-dimensional Heston stochastic volatility model.*

In the framework of the one-dimensional Heston stochastic volatility model, it is assumed that interest rates are non-negative constants ($r_t \equiv r \geq 0$) and there exists one risky asset $S_t = e^{Y_t}$. Very often this risky asset can be interpreted as a stock.

Hence Y_t in (2.1) describes the dynamics of the log returns $Y_t = \ln S_t$ of the risky asset S_t while X_t in (2.1) represents the dynamics of the variance of this risky asset.

Remark 2.1.2. Note that if we define the Brownian motion $\mathcal{W}_t := \rho B_t + \sqrt{1-\rho^2} W_t$ (see also Remark 2.2.2), the quadratic covariation can be calculated by

$$\begin{aligned} \langle dB_t, d\mathcal{W}_t \rangle &= \langle dB_t, \rho dB_t + \sqrt{1-\rho^2} dW_t \rangle \\ &= \rho \langle dB_t, dB_t \rangle + \sqrt{1-\rho^2} \underbrace{\langle dB_t, dW_t \rangle}_{=0} \\ &= \rho dt. \end{aligned}$$

2.1.1 The CIR Process as the Variance Process in the One-Dimensional Heston Stochastic Volatility Model

In literature, the variance process in the one-dimensional Heston stochastic volatility model (which is given by the dynamics X_t) is often called CIR process (Cox-Ingersoll-Ross process) after the mathematicians Cox, Ingersoll and Ross who have introduced this square-root process in mathematical finance. Before William Feller has already studied such kind of diffusion processes named Feller processes in genetics.

Often another parametrization for this CIR process is used in literature as for example in the paper "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options" written by Steven L. Heston (compare [12]) or in the paper "Continuous Time Wishart Process for Stochastic Risk" published by Christian Gouriéroux (compare [9]):

The latter is used for the following considerations concerning the link between the CIR process and the OU process (Ornstein-Uhlenbeck process) (see [9, p.178]):

The Cox-Ingersoll-Ross process satisfies the diffusion equation

$$dv_t = \bar{\kappa}(\theta - v_t)dt + \bar{\sigma}\sqrt{v_t} dB_t, \tag{2.2}$$

where (B_t) denotes an univariate Brownian motion and the parameters satisfy the following conditions: $\bar{\sigma} > 0$, $\bar{\kappa}\theta \geq 0$ (the latter condition guarantees the positivity of the stochastic process conditioned that the initial value v_0 is also positive).

In this context the parameters of the CIR process can be interpreted as follows: $\bar{\kappa}$ denotes the mean reversion rate, θ denotes the long run variance and $\bar{\sigma}$ denotes the volatility of the variance.

One can also mention that both the expectation and the variance of the CIR process are affine functions of the current process value v_t since

$$\mathbb{E}[dv_t] = \bar{\kappa}(\theta - v_t)dt \quad \text{and} \quad \mathbb{V}[dv_t] = \bar{\sigma}^2 v_t dt.$$

2.1.2 Link between the Ornstein-Uhlenbeck Process and the CIR Process

This subsection is again based on [9, p.183-184].

Assume that an Ornstein-Uhlenbeck process is given by its dynamics $dx_t = \epsilon x_t dt + \eta dB_t$, where $\epsilon, \eta \in \mathbb{R}$.

By applying Ito's formula, the square of this Ornstein-Uhlenbeck process can be obtained as follows:

$$\begin{aligned} d(x_t^2) &= 2x_t dx_t + \frac{1}{2} 2\langle dx_t \rangle \\ &= 2x_t(\epsilon x_t dt + \eta dB_t) + \langle \eta dB_t \rangle \\ &= 2x_t(\epsilon x_t dt + \eta dB_t) + \eta^2 \langle dB_t \rangle \\ &= 2\epsilon x_t^2 dt + 2x_t \eta dB_t + \eta^2 dt \\ &= (2\epsilon x_t^2 + \eta^2)dt + 2\eta x_t dB_t. \end{aligned} \tag{2.3}$$

If we now define $y_t = x_t^2$, then the dynamics in (2.3) can be rewritten as

$$dy_t = (2\epsilon y_t + \eta^2)dt + 2\eta\sqrt{y_t}dB_t.$$

As a next step we consider n independent Ornstein-Uhlenbeck processes with identical parameters

$$dx_t^i = \epsilon x_t^i dt + \eta dB_t^i \qquad i = 1, \ldots, n,$$

where the Brownian motions B_t^i, $i = 1, \ldots, n$ are independent.

Then we define by z_t the sum of these independent squared Ornstein-Uhlenbeck processes, hence $z_t = \sum_{i=1}^n (x_t^i)^2$.
In addition to that, we define the process $\tilde{B}_t$ by $d\tilde{B}_t = \frac{\sum_{i=1}^n x_t^i dB_t^i}{\sqrt{z_t}}$, which is a Brownian motion. To verify that $\tilde{B}_t$ is a Brownian motion, we apply Levy's characterization theorem. If we can show that $\langle d\tilde{B}_t \rangle = dt$, it implies that $\tilde{B}_t$ is a Brownian motion:

$$\begin{aligned} \langle d\tilde{B}_t \rangle &= \left\langle \frac{\sum_{i=1}^n x_t^i dB_t^i}{\sqrt{z_t}}, \frac{\sum_{j=1}^n x_t^j dB_t^j}{\sqrt{z_t}} \right\rangle \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n x_t^i x_t^j \langle dB_t^i, dB_t^j \rangle}{\sqrt{z_t}\sqrt{z_t}} \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n x_t^i x_t^j \mathbf{1}_{(i=j)} dt}{z_t} \\ &= \frac{\sum_{i=1}^n (x_t^i)^2}{z_t} dt = \frac{z_t}{z_t} dt = dt. \end{aligned} \tag{2.4}$$

Hence we have shown that $\tilde{B}_t$ is a one-dimensional Brownian motion.

By Ito's formula it follows that

$$\begin{aligned}
dz_t &= d(\sum_{i=1}^{n}(x_t^i)^2) = \sum_{i=1}^{n} d((x_t^i)^2) \\
&= \sum_{i=1}^{n}(2\epsilon(x_t^i)^2 + \eta^2)dt + \sum_{i=1}^{n}(2\eta x_t^i dB_t^i) \\
&= \left(2\epsilon \sum_{i=1}^{n}(x_t^i)^2 + \sum_{i=1}^{n}\eta^2\right) dt + 2\eta \sum_{i=1}^{n} x_t^i dB_t^i \\
&= (2\epsilon z_t + n\eta^2)dt + 2\eta\sqrt{z_t}d\tilde{B}_t,
\end{aligned}$$

because $d\tilde{B}_t = \frac{\sum_{i=1}^{n} x_t^i dB_t^i}{\sqrt{z_t}} \Leftrightarrow \sum_{i=1}^{n} x_t^i dB_t^i = \sqrt{z_t}d\tilde{B}_t$ (last step).

Considering these steps and calculations, we can state the following proposition about the link between the CIR process (as defined in (2.2)) and the OU process:

Proposition 2.1.3. *The sum of squares of n independent Ornstein-Uhlenbeck processes with identical parameters ϵ and η is a Cox-Ingersoll-Ross process with parameters*

$$\bar{\kappa} = -2\epsilon, \quad \bar{\sigma} = 2\eta, \quad \bar{\kappa}\theta = n\eta^2 \Rightarrow \theta = -\frac{n\eta^2}{2\epsilon}.$$

2.2 Interesting Results of the One-Dimensional Heston Stochastic Volatility Model

Proposition 2.2.1. *In the one-dimensional Heston stochastic volatility model the risk-neutral stock dynamics are given by*

$$dS_t = S_t r dt + S_t\sqrt{2X_t}d\mathcal{W}_t,$$

where $\mathcal{W}_t$ denotes a linear combination of the two Brownian motions B_t and W_t defined by $\mathcal{W}_t = \rho B_t + \sqrt{1-\rho^2}W_t$.

Proof. By applying Ito's formula, one can show that the risk-neutral stock dynamics are given by

$$
\begin{aligned}
dS_t &= d(e^{Y_t}) = e^{Y_t}dY_t + \frac{1}{2}e^{Y_t}\underbrace{d\langle Y_t, Y_t\rangle}_{=d\langle Y_t\rangle} \\
&= e^{Y_t}\left((r - X_t)dt + \sqrt{2X_t}(\rho dB_t + \sqrt{1-\rho^2}dW_t)\right) + \frac{1}{2}e^{Y_t}2X_t\underbrace{\langle \rho dB_t + \sqrt{1-\rho^2}dW_t\rangle}_{=\rho^2\langle dB_t\rangle + (1-\rho^2)\langle dW_t\rangle} \\
&= S_t\left((r - X_t)dt + \sqrt{2X_t}(\rho dB_t + \sqrt{1-\rho^2}dW_t)\right) + S_tX_t(\rho^2 dt + (1-\rho^2)dt) \\
&= S_t(rdt - X_tdt + \sqrt{2X_t}(\rho dB_t + \sqrt{1-\rho^2}dW_t)) + S_tX_tdt \\
&= S_trdt + S_t\sqrt{2X_t}\underbrace{(\rho dB_t + \sqrt{1-\rho^2}dW_t)}_{=d\mathcal{W}_t} \\
&= S_trdt + S_t\sqrt{2X_t}d\mathcal{W}_t.
\end{aligned}
$$

□

Remark 2.2.2. Clearly $\mathcal{W}_t$ is again a Brownian motion because

$$\langle \mathcal{W}_t\rangle = \langle \rho B_t\rangle + \langle \sqrt{1-\rho^2}W_t\rangle = \rho^2\langle B_t\rangle + (1-\rho^2)\langle W_t\rangle = \rho^2 t + (1-\rho^2)t = t,$$

and then by applying Levy's characterization theorem the assertion follows.

Remark 2.2.3. Note that $\sqrt{2X_t}$ denotes the stochastic volatility of the price process of the risky asset.

Furthermore one can easily show that in the one-dimensional Heston stochastic volatility model the covariation between the risky asset and the variance process can be expressed as follows:

Proposition 2.2.4. *In the one-dimensional Heston stochastic volatility model, the covariation between the risky asset S_t and the variance X_t is given by*

$$d\langle S_t, X_t\rangle = 2\sigma\rho S_tX_tdt.$$

Proof. The quadratic covariation between S_t and X_t can be calculated as follows

$$d\langle S_t, X_t\rangle = S_t\sqrt{2X_t}\sigma\sqrt{2X_t}d\langle \mathcal{W}_t, B_t\rangle = S_t\sigma 2X_t\underbrace{d\langle \rho B_t + \sqrt{1-\rho^2}W_t, B_t\rangle}_{\overset{B_t, W_t \text{indep.}}{=}\rho dt} = 2\sigma\rho S_tX_tdt.$$

□

With the same argumentation as above the covariation between the log return and the variance process is examined in the following proposition:

Proposition 2.2.5. *In the one-dimensional Heston stochastic volatility model, the covariation between the log return Y_t and the variance X_t does not depend on the log return Y_t, it only depends on the variance X_t. The correlation between these two random variables X_t and Y_t is given by*

$$d\langle Y_t, X_t\rangle = 2\sigma\rho X_t dt.$$

Proof. The quadratic covariation between Y_t and X_t is obtained by

$$d\langle Y_t, X_t\rangle = \sqrt{2X_t}\sigma\sqrt{2X_t}\underbrace{d\langle \rho B_t + \sqrt{1-\rho^2}W_t, B_t\rangle}_{\overset{B_t,W_t \text{indep.}}{=} \rho dt} = 2\sigma\rho X_t dt.$$

□

2.3 Riccati Equations in the One-Dimensional Heston Stochastic Volatility Model

2.3.1 Derivation of the Riccati Equations in the One-Dimensional Heston Stochastic Volatility Model

Now the Riccati equations are derived for the special case of the one-dimensional Heston stochastic volatility model.

Theorem 2.3.1. *In the one-dimensional Heston stochastic volatility model, the Riccati equations can be written as*

$$\begin{aligned}
\phi(t,u) &= k\int_0^t \psi_1(s,u)ds + ru_2 t,\\
\frac{\partial}{\partial t}\psi_1(t,u) &= \sigma^2\psi_1^2(t,u) + (2\rho\sigma u_2 + \kappa)\psi_1(t,u) + u_2^2 - u_2,\\
\psi_1(0,u) &= u_1,\\
\psi_2(t,u) &= u_2.
\end{aligned} \tag{2.5}$$

Proof. The Riccati equations for the one-dimensional Heston stochastic volatility model are deduced from Theorem 1.2.1 with $d = 2$ and $I = \{1\}$, $J = \{2\}$.
In matrix notation (2.1) can be rewritten as

$$d\mathcal{X}_t = \begin{pmatrix} dX_t \\ dY_t \end{pmatrix} = \underbrace{\begin{pmatrix} k + \kappa X_t \\ r - X_t \end{pmatrix}}_{=b(\mathcal{X}_t)} dt + \underbrace{\begin{pmatrix} \sigma\sqrt{2X_t} & 0 \\ \rho\sqrt{2X_t} & \sqrt{1-\rho^2}\sqrt{2X_t} \end{pmatrix}}_{=\rho(\mathcal{X}_t)} \cdot \begin{pmatrix} dB_t \\ dW_t \end{pmatrix}.$$

Since $a(\mathcal{X}_t) = \rho(\mathcal{X}_t)\rho(\mathcal{X}_t)^\top$, $a(\mathcal{X}_t)$ can be calculated by

$$\begin{aligned} a(\mathcal{X}_t) &= \begin{pmatrix} \sigma\sqrt{2X_t} & 0 \\ \rho\sqrt{2X_t} & \sqrt{1-\rho^2}\sqrt{2X_t} \end{pmatrix} \cdot \begin{pmatrix} \sigma\sqrt{2X_t} & \rho\sqrt{2X_t} \\ 0 & \sqrt{1-\rho^2}\sqrt{2X_t} \end{pmatrix} \\ &= \begin{pmatrix} \sigma^2 2X_t & \sigma\rho 2X_t \\ \sigma\rho 2X_t & \rho^2 2X_t + (1-\rho^2)2X_t \end{pmatrix} = \begin{pmatrix} 2\sigma^2 X_t & 2\sigma\rho X_t \\ 2\sigma\rho X_t & 2X_t \end{pmatrix}. \end{aligned} \tag{2.6}$$

From Theorem 1.1.3 we know that the diffusion matrix $a(x)$ has also an affine representation and hence comparing (2.6) with (1.3), one can easily see that

$$a = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \alpha_1 = \begin{pmatrix} 2\sigma^2 & 2\sigma\rho \\ 2\sigma\rho & 2 \end{pmatrix}, \qquad \alpha_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Since the drift is also affine, one gets

$$b = \begin{pmatrix} k \\ r \end{pmatrix}, \qquad \beta_1 = \begin{pmatrix} \kappa \\ -1 \end{pmatrix}, \qquad \beta_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \mathcal{B} = \begin{pmatrix} \kappa & 0 \\ -1 & 0 \end{pmatrix}.$$

Clearly, the second and the fifth equation of (1.7) remain also the same in the context of the one-dimensional Heston stochastic volatility model.

Since $\mathcal{B}_{22} = 0$, it follows that $\frac{\partial}{\partial t}\psi_2(t,u) = 0$ and hence

$$\psi_2(t,u) = const. = \psi_2(0,u) = u_2.$$

Plugging $a_{22} = 0$ and $b^\top = (k, r)$ into the first equation of (1.7) leads to

$$\frac{\partial}{\partial t}\phi(t,u) = \begin{pmatrix} k & r \end{pmatrix} \begin{pmatrix} \psi_1(t,u) \\ \psi_2(t,u) \end{pmatrix} = k\psi_1(t,u) + r\underbrace{\psi_2(t,u)}_{=u_2} = k\psi_1(t,u) + ru_2,$$

and by considering that $\phi(0,u) = 0$, it follows that

$$\phi(t,u) = \int_0^t \frac{\partial}{\partial s}\phi(s,u)ds = \int_0^t k\psi_1(s,u)ds + \int_0^t ru_2 ds = k\int_0^t \psi_1(s,u)ds + ru_2 t,$$

which verifies the first equation of (2.5).

Inserting α_1 and β_1 in the third equation of (1.7) and remembering that $\psi_2(t,u) = u_2$, we show that $\frac{\partial}{\partial t}\psi_1(t,u)$ can be calculated from (1.7) as follows:

$$
\begin{aligned}
\frac{\partial}{\partial t}\psi_1(t,u) &= \frac{1}{2}\underbrace{\begin{pmatrix} \psi_1(t,u) & \psi_2(t,u) \end{pmatrix}\begin{pmatrix} 2\sigma^2 & 2\sigma\rho \\ 2\sigma\rho & 2 \end{pmatrix}}_{=(2\sigma^2\psi_1(t,u)+2\sigma\rho\psi_2(t,u)\ ,\ 2\sigma\rho\psi_1(t,u)+2\psi_2(t,u))}\begin{pmatrix} \psi_1(t,u) \\ \psi_2(t,u) \end{pmatrix} + \begin{pmatrix} \kappa & -1 \end{pmatrix}\begin{pmatrix} \psi_1(t,u) \\ \psi_2(t,u) \end{pmatrix} \\
&= \frac{1}{2}\Bigg(2\sigma^2\psi_1^2(t,u) + 2\sigma\rho\psi_1(t,u)\underbrace{\psi_2(t,u)}_{=u_2} + 2\sigma\rho\psi_1(t,u)\underbrace{\psi_2(t,u)}_{=u_2} + 2\underbrace{\psi_2^2(t,u)}_{=u_2^2}\Bigg) \\
&\quad + \kappa\psi_1(t,u) - \underbrace{\psi_2(t,u)}_{=u_2} \\
&= \sigma^2\psi_1^2(t,u) + 2\sigma\rho u_2\psi_1(t,u) + u_2^2 + \kappa\psi_1(t,u) - u_2 \\
&= \sigma^2\psi_1^2(t,u) + (2\sigma\rho u_2 + \kappa)\psi_1(t,u) + u_2^2 - u_2.
\end{aligned}
$$

Hence we have also proved that the second equation of (2.5) is valid. □

2.3.2 Explicit Solutions of the Riccati Equations

The idea of this approach is mainly based on [19].

The aim is to find explicit solutions for the Riccati equations in the one-dimensional Heston stochastic volatility model:

Hence we consider the first two equations of (2.5)

$$
\begin{aligned}
\phi(t,u) &= k\int_0^t \psi_1(s,u)ds + ru_2t, \\
\frac{\partial}{\partial t}\psi_1(t,u) &= \sigma^2\psi_1^2(t,u) + (2\rho\sigma u_2 + \kappa)\psi_1(t,u) + u_2^2 - u_2.
\end{aligned}
\tag{2.7}
$$

Since $\psi_1(t,u)$ is independent of $\phi(t,u)$, we consider the second ordinary differential equation (ODE) at first:

For notational simplicity, we define $\psi_1(t,u) =: \psi_1(t)$ and $\frac{\partial}{\partial t}\psi_1(t,u) = \frac{\partial}{\partial t}\psi_1(t) =: \dot{\psi}_1(t)$ and then the second equation of (2.7) can be rewritten as

$$
\dot{\psi}_1(t) = \frac{\partial}{\partial t}\psi_1(t) = \sigma^2\psi_1^2(t) + v\psi_1(t) + w
$$

with $v = 2\rho\sigma u_2 + \kappa$ and $w = u_2^2 - u_2$.

Firstly, we consider the special case where $0 = \dot{\psi}_1(0) = \sigma^2\psi_1^2(0) + v\psi_1(0) + w$. In this case, we clearly get $\psi_1(t) = \psi_1(0)$ and therefore $\psi_1(t,u) = \psi_1(0,u) = u_1$. Hence in this special case we have $\phi(t,u) = ku_1t + ru_2t$.

In general let ζ_1 be a solution of the equation below:

$$\sigma^2\zeta_1^2 + v\zeta_1 + w = 0.$$

One can easily see that ζ_1 must be of the following form

$$\zeta_1^{\pm} = \frac{-v \pm \sqrt{v^2 - 4\sigma^2 w}}{2\sigma^2} = \frac{-v \pm z}{2\sigma^2} \quad \text{with} \quad z = \sqrt{v^2 - 4\sigma^2 w}.$$

Without loss of generality, we set $\zeta_1 = \zeta_1^+$. As a next step we consider the difference between the stationary solution ζ_1 and the "true" solution $\psi_1(t)$, namely $\Delta(t) = \psi_1(t) - \zeta_1$.

Since $\frac{\partial}{\partial t}\zeta_1 = 0$ (per definition of ζ_1) and remembering that $\sigma^2\zeta_1^2 + v\zeta_1 + w = 0$ and $\zeta_1 = \frac{-v+z}{2\sigma^2} \Leftrightarrow z = 2\sigma^2\zeta_1 + v$, one gets

$$\begin{aligned} \frac{\partial\Delta(t)}{\partial t} &= \frac{\partial(\zeta_1 + \Delta(t))}{\partial t} \\ &= \sigma^2(\zeta_1 + \Delta(t))^2 + v(\zeta_1 + \Delta(t)) + w \\ &= \sigma^2\zeta_1^2 + 2\sigma^2\zeta_1\Delta(t) + \sigma^2\Delta^2(t) + v\zeta_1 + v\Delta(t) + w \\ &= \sigma^2\Delta^2(t) + (2\sigma^2\zeta_1 + v)\Delta(t) = \sigma^2\Delta^2(t) + z\Delta(t) \end{aligned} \tag{2.8}$$

with initial condition $\Delta(0) = \psi_1(0) - \zeta_1$.

Considering Equation (2.8), namely $\frac{\partial\Delta(t)}{\partial t} = \dot{\Delta}(t) = \sigma^2\Delta^2(t) + z\Delta(t)$, we recognize that this differential equation has the form of a Bernoulli differential equation which can be explicitly solved.

Substituting $y(t) = \Delta^{-1}(t)$ leads the following equivalent linear differential equation,

$$\dot{y}(t) = -\Delta^{-2}(t)\dot{\Delta}(t) = -\Delta^{-2}(t)(\sigma^2\Delta^2(t) + z\Delta(t)) = -\sigma^2 - z\Delta^{-1}(t) = -\sigma^2 - zy(t),$$

whose solution can be represented as the sum of the homogeneous solution $y_h(t)$ and a particular solution $y_p(t)$.

To solve the ODE, the homogeneous equation is considered at first:

$$\dot{y}(t) = -zy(t) \quad \Leftrightarrow \quad \int \frac{dy(t)}{y(t)} = \int -z dt \quad \Leftrightarrow \quad \ln|y(t)| = -zt + c \quad (c \in \mathbb{R}).$$

Hence the homogeneous solution is given by $y_h(t) = ce^{-zt}$.

Applying variation of the constant leads to the particular solution $y_p(t) = c(t)e^{-zt}$:
By plugging the particular solution in the linear differential equation above, we get

$$\underbrace{\dot{c}(t)e^{-zt} + c(t)e^{-zt}(-z)}_{\dot{y}_p(t)} = -\sigma^2 - zc(t)e^{-zt}$$

$$\Leftrightarrow \dot{c}(t)e^{-zt} = -\sigma^2$$

$$\Leftrightarrow \dot{c}(t) = -e^{zt}\sigma^2$$

$$\Rightarrow c(t) = -\frac{\sigma^2}{z}e^{zt}.$$

Since only a particular solution is needed, the constant of integrations can be omitted. Hence the particular solution is of the following form

$$y_p(t) = c(t)e^{-zt} = -\frac{\sigma^2}{z}e^{zt}e^{-zt} = -\frac{\sigma^2}{z}.$$

All in all, we get

$$y(t) = y_h(t) + y_p(t) = ce^{-zt} - \frac{\sigma^2}{z}.$$

Resubstituting gives us

$$\Delta(t) = \frac{1}{y(t)} = \frac{1}{ce^{-zt} - \frac{\sigma^2}{z}}. \tag{2.9}$$

To determine the constant, we consider equation (2.9) for the initial value

$$\Delta(0) = \frac{1}{c - \frac{\sigma^2}{z}} = \psi_1(0) - \zeta_1^+ \quad \Rightarrow \quad c = \frac{1}{\psi_1(0) - \zeta_1^+} + \frac{\sigma^2}{z}.$$

For the next step some short auxiliary calculations are needed: Since $\zeta_1^{\pm}$ is defined as $\zeta_1^{\pm} = \frac{-v \pm z}{2\sigma^2}$, it follows that $\zeta_1^+ - \zeta_1^- = \frac{-v+z}{2\sigma^2} - \frac{-v-z}{2\sigma^2} = \frac{z}{\sigma^2}$ and hence $\frac{\sigma^2}{z} = \frac{1}{\zeta_1^+ - \zeta_1^-}$ or equivalent $\frac{\sigma^2}{z}(\zeta_1^+ - \zeta_1^-) = 1$.

With these auxiliary calculations, $\Delta(t)$ in (2.9) becomes

$$
\begin{aligned}
\Delta(t) &= \frac{1}{\left(\frac{1}{\psi_1(0)-\zeta_1^+} + \frac{\sigma^2}{z}\right) e^{-zt} - \frac{\sigma^2}{z}} \\
&= \frac{1}{\left(\frac{1}{\psi_1(0)-\zeta_1^+} + \frac{1}{\zeta_1^+-\zeta_1^-}\right) e^{-zt} - \frac{\sigma^2}{z}} \\
&= \frac{\zeta_1^+ - \zeta_1^-}{\left(\frac{\zeta_1^+-\zeta_1^-}{\psi_1(0)-\zeta_1^+} + 1\right) e^{-zt} - \frac{\sigma^2}{z}(\zeta_1^+ - \zeta_1^-)} \\
&= \frac{\zeta_1^+ - \zeta_1^-}{\left(\frac{\zeta_1^+-\zeta_1^- + \psi_1(0)-\zeta_1^+}{\psi_1(0)-\zeta_1^+}\right) e^{-zt} - 1} \\
&= \frac{\zeta_1^+ - \zeta_1^-}{\left(\frac{\psi_1(0)-\zeta_1^-}{\psi_1(0)-\zeta_1^+}\right) e^{-zt} - 1} \\
&= -\frac{\zeta_1^+ - \zeta_1^-}{\left(\frac{\zeta_1^- - \psi_1(0)}{\psi_1(0)-\zeta_1^+}\right) e^{-zt} + 1} \\
&= \frac{\zeta_1^- - \zeta_1^+}{\left(\frac{\zeta_1^- - \psi_1(0)}{\psi_1(0)-\zeta_1^+}\right) e^{-zt} + 1} \\
&= \frac{\zeta_1^- - \zeta_1^+}{1 - \left(\frac{\zeta_1^- - \psi_1(0)}{\zeta_1^+ - \psi_1(0)}\right) e^{-zt}} \\
&= \frac{\zeta_1^- - \zeta_1^+}{1 - ae^{-zt}},
\end{aligned}
$$

with $a = \frac{\zeta_1^- - \psi_1(0)}{\zeta_1^+ - \psi_1(0)}$.

Hence, $\psi_1(t)$ can be rewritten as

$$
\begin{aligned}
\psi_1(t) &= \zeta_1^+ + \Delta(t) \\
&= \zeta_1^+ + \frac{\zeta_1^- - \zeta_1^+}{1 - ae^{-zt}} + \psi_1(0) - \psi_1(0) \\
&= \psi_1(0) + \frac{\zeta_1^- - \zeta_1^+ + (\zeta_1^+ - \psi_1(0))(1 - ae^{-zt})}{1 - ae^{-zt}} \\
&= \psi_1(0) + \frac{\zeta_1^- - \zeta_1^+ + \zeta_1^+ - \psi_1(0) - ae^{-zt}(\zeta_1^+ - \psi_1(0))}{1 - ae^{-zt}} \\
&= \psi_1(0) + \frac{\zeta_1^- - \psi_1(0) - e^{-zt}(\zeta_1^- - \psi_1(0))}{1 - ae^{-zt}} \\
&= \psi_1(0) + \frac{(\zeta_1^- - \psi_1(0)(1 - e^{-zt}))}{1 - ae^{-zt}}.
\end{aligned}
$$

Inserting $\psi_1(t)$ into $\phi_1(t)$, $\psi_1(t)$ can be rewritten in explicit form as

$$
\begin{aligned}
\phi(t) &= k\int_0^t \psi_1(s)ds + ru_2t \\
&= k\int_0^t \left(\psi_1(0) + \frac{(\zeta_1^- - \psi_1(0))(1-e^{-zs})}{1-ae^{-zs}}\right) ds + ru_2t \\
&= k\psi_1(0)t + k(\zeta_1^- - \psi_1(0))\int_0^t \frac{1-e^{-zs}}{1-ae^{-zs}}ds + ru_2t \\
&= k\psi_1(0)t + k(\zeta_1^- - \psi_1(0))\int_0^t \frac{1-ae^{-zs}+ae^{-zs}-e^{-zs}}{1-ae^{-zs}}ds + ru_2t \\
&= k\psi_1(0)t + k(\zeta_1^- - \psi_1(0))\int_0^t \left(1+\frac{(a-1)e^{-zs}}{1-ae^{-zs}}\right) ds + ru_2t \\
&= k\psi_1(0)t + k(\zeta_1^- - \psi_1(0))\left(t+\int_0^t \frac{(a-1)e^{-zs}}{1-ae^{-zs}}ds\right) + ru_2t \\
&= k\zeta_1^- t + k(\zeta_1^- - \psi_1(0))(a-1)\int_0^t \frac{e^{-zs}}{1-ae^{-zs}}ds + ru_2t \\
&\overset{\text{subs. } p=e^{-zs}}{=} k\zeta_1^- t + k(\zeta_1^- - \psi_1(0))(a-1)\int_1^{e^{-zt}} \frac{p}{1-ap}\left(-\frac{1}{zp}\right) dp + ru_2t \\
&= k\zeta_1^- t - k(\zeta_1^- - \psi_1(0))\frac{a-1}{z}\int_1^{e^{-zt}} \frac{1}{1-ap}dp + ru_2t \\
&= k\zeta_1^- t - k(\zeta_1^- - \psi_1(0))\frac{a-1}{z}\left(-\ln(1-ap)\frac{1}{a}\right)\Bigg|_1^{e^{-zt}} + ru_2t \\
&= k\zeta_1^- t + k(\zeta_1^- - \psi_1(0))\frac{a-1}{az}\ln\left(\frac{1-ae^{-zt}}{1-a}\right) + ru_2t \\
&= k\zeta_1^- t - \frac{k}{\sigma^2}\ln\left(\frac{1-ae^{-zt}}{1-a}\right) + ru_2t.
\end{aligned}
$$

In the last step of the equation above, we have used that

$$
\begin{aligned}
(\zeta_1^- - \psi_1(0))\frac{a-1}{az} &= \frac{\zeta_1^+ - \psi_1(0)}{z}\left(\frac{\zeta_1^- - \psi_1(0)}{\zeta_1^+ - \psi(0)} - 1\right) \\
&= \frac{\zeta_1^- - \psi_1(0) - \zeta_1^+ + \psi_1(0)}{z} \\
&= -\frac{\zeta_1^+ - \zeta_1^-}{z} \\
&= -\frac{1}{\sigma^2}.
\end{aligned}
$$

Resubstituting gives the following explicit solutions for $\phi(t,u)$ and $\psi(t,u)$

$$\begin{aligned}
\psi_1(t,u) &= \psi_1(0,u) + \frac{(\zeta_1^- - \psi_1(0,u))(1 - e^{-(\sqrt{v^2-4\sigma^2 w})t})}{1 - \frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)} e^{-(\sqrt{v^2-4\sigma^2 w})t}} \\
&= \psi_1(0,u) + \frac{(\zeta_1^- - \psi_1(0,u))(1 - e^{-(\sqrt{(2\rho\sigma u_2+\kappa)^2-4\sigma^2(u_2^2-u_2)})t})}{1 - \frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)} e^{-(\sqrt{(2\rho\sigma u_2+\kappa)^2-4\sigma^2(u_2^2-u_2)})t}} \\
\phi(t,u) &= k\zeta_1^- t - \frac{k}{\sigma^2} \ln\left(\frac{1 - \left(\frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)}\right) e^{-(\sqrt{v^2-4\sigma^2 w})t}}{1 - \left(\frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)}\right)} \right) + ru_2 t \\
&= k\zeta_1^- t - \frac{k}{\sigma^2} \ln\left(\frac{1 - \left(\frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)}\right) e^{-(\sqrt{(2\rho\sigma u_2+\kappa)^2-4\sigma^2(u_2^2-u_2)})t}}{1 - \left(\frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)}\right)} \right) + ru_2 t
\end{aligned}$$

with

$$\zeta_1^{\pm} = \frac{-(2\rho\sigma u_2 + \kappa) \pm \sqrt{(2\rho\sigma u_2 + \kappa)^2 - 4\sigma^2(u_2^2 - u_2)}}{2\sigma^2}.$$

These explicit solutions gives rise to the following corollary concerning the Riccati equations in the one-dimensional Heston stochastic volatility model:

Corollary 2.3.2. *In the one-dimensional Heston stochastic volatility model, the Riccati equations in the form of explicit solutions can be written as*

$$\begin{aligned}
\psi_1(t,u) &= \psi_1(0,u) + \frac{(\zeta_1^- - \psi_1(0,u))(1 - e^{-(\sqrt{(2\rho\sigma u_2+\kappa)^2-4\sigma^2(u_2^2-u_2)})t})}{1 - \frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)} e^{-(\sqrt{(2\rho\sigma u_2+\kappa)^2-4\sigma^2(u_2^2-u_2)})t}}, \\
\phi(t,u) &= k\zeta_1^- t - \frac{k}{\sigma^2} \ln\left(\frac{1 - \left(\frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)}\right) e^{-(\sqrt{(2\rho\sigma u_2+\kappa)^2-4\sigma^2(u_2^2-u_2)})t}}{1 - \left(\frac{\zeta_1^- - \psi_1(0,u)}{\zeta_1^+ - \psi_1(0,u)}\right)} \right) + ru_2 t, \\
\psi_1(0,u) &= u_1, \\
\psi_2(t,u) &= u_2.
\end{aligned}$$

with

$$\zeta_1^{\pm} = \frac{-(2\rho\sigma u_2 + \kappa) \pm \sqrt{(2\rho\sigma u_2 + \kappa)^2 - 4\sigma^2(u_2^2 - u_2)}}{2\sigma^2}.$$

2.3.3 Joint Characteristic Function in the One-Dimensional Heston Stochastic Volatility Model

Combining Theorem 1.3.1 and Corollary 2.3.2, we are now prepared to introduce the joint conditional characteristic function of the log price and variance process in the one-dimensional

Heston stochastic volatility model:

Theorem 2.3.3. *In the one-dimensional Heston stochastic volatility model, for all $u_1, u_2 \in i\mathbb{R}$, the joint conditional characteristic function can be expressed as*

$$\mathbb{E}\left[e^{u_1 X_T + u_2 Y_T} \middle| \mathcal{F}_t\right] = e^{\phi(T-t,u)+\psi_1(T-t,u)X_t+u_2 Y_t},$$

where $\phi(T-t,u)$ and $\psi_1(T-t,u)$ are given as in the last Corollary 2.3.2.

Proof. The above theorem follows from equation (1.9) and Corollary 2.3.2 by noting that $\psi_2(t,u) = u_2$. □

2.4 Pricing in the One-Dimensional Heston Stochastic Volatility Model

Now the goal is to show how the price of a European call option can be calculated in the one-dimensional Heston stochastic volatility model (compare [5, p.158-168]).
The calculation is done by Fourier pricing which was introduced in the last chapter for affine models in general.

The price of a European call option at an arbitrary time point $t \leq T$ with strike K and maturity T is given by (risk-neutral valuation formula)

$$\pi(t) = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}[(S_T - K)^+ | \mathcal{F}_t].$$

As a preliminary work, the following lemma is stated and a part of it is also proved.

Lemma 2.4.1. *[5, Lemma 10.2] Let $K > 0$. For any $y \in \mathbb{R}$ the following identities hold:*

$$\frac{1}{2\pi}\int_{\mathbb{R}} e^{(w+i\lambda)y}\frac{K^{-(w-1+i\lambda)}}{(w+i\lambda)(w-1+i\lambda)}d\lambda = \begin{cases} (K-e^y)^+ & \text{if } w < 0, \\ (e^y - K)^+ - e^y & \text{if } 0 < w < 1, \\ (e^y - K)^+ & \text{if } w > 1. \end{cases}$$

The middle case $(0 < w < 1)$ obviously also equals $(K - e^y)^+ - K$.

Now, only one case of the above lemma is proved, namely the case where $w > 1$. This case will be needed in the sequel to derive the price of a European call option in the one-dimensional Heston stochastic volatility model.

Proof. For $w > 1$ we have to prove that

$$(e^y - K)^+ = \frac{1}{2\pi}\int_{\mathbb{R}} e^{(w+i\lambda)y}\frac{K^{-(w-1+i\lambda)}}{(w+i\lambda)(w-1+i\lambda)}d\lambda.$$

Define $h(y) = e^{-wy}(e^y - K)^+$. Then its Fourier transform can be calculated as

$$\begin{aligned}
\hat{h}(\lambda) &= \int_{\mathbb{R}} e^{-(w+i\lambda)y}(e^y - K)^+ dy \\
&= \int_{\ln K}^{\infty} e^{-(w+i\lambda)y}(e^y - K) dy \\
&= \int_{\ln K}^{\infty} e^{-(w-1+i\lambda)y} dy - K \int_{\ln K}^{\infty} e^{-(w+i\lambda)y} dy \\
&= -\frac{e^{-(w-1+i\lambda)y}}{w-1+i\lambda}\bigg|_{\ln K}^{\infty} + K\frac{e^{-(w+i\lambda)y}}{w+i\lambda}\bigg|_{\ln K}^{\infty} \\
&= \frac{K^{-(w-1+i\lambda)}}{w-1+i\lambda} - \frac{K^{-(w-1+i\lambda)}}{w+i\lambda} \\
&= \frac{K^{-(w-1+i\lambda)}(w+i\lambda) - K^{-(w-1+i\lambda)}(w-1+i\lambda)}{(w-1+i\lambda)(w+i\lambda)} \\
&= \frac{K^{-(w-1+i\lambda)}}{(w-1+i\lambda)(w+i\lambda)}.
\end{aligned}$$

Then the fundamental inversion formula from Fourier analysis is applied:
Let $g : \mathbb{R}^q \to \mathbb{C}$ be an integrable function with integrable Fourier transform

$$\hat{g}(\lambda) = \int_{\mathbb{R}^q} e^{-i\lambda^\top y} g(y) dy.$$

Then the inversion formula

$$g(y) = \frac{1}{(2\pi)^q} \int_{\mathbb{R}^q} e^{i\lambda^\top y} \hat{g}(\lambda) d\lambda$$

holds for dy-almost all $y \in \mathbb{R}^q$.

In this context, we get

$$\begin{aligned}
h(y) = e^{-wy}(e^y - K)^+ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\lambda y} \frac{K^{-(w-1+i\lambda)}}{(w-1+i\lambda)(w+i\lambda)} d\lambda \\
\Leftrightarrow (e^y - K)^+ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{(w+i\lambda)y} \frac{K^{-(w-1+i\lambda)}}{(w-1+i\lambda)(w+i\lambda)} d\lambda.
\end{aligned}$$

□

Applying Lemma 2.4.1 we are now prepared to state a theorem about the price of a European call option in the one-dimensional Heston stochastic volatility model.

Theorem 2.4.2. *In the one-dimensional Heston stochastic volatility model, the price of a European call option is given in closed form by*

$$\begin{aligned}
\pi(t) &= e^{-r(T-t)} \mathbb{E}_{\mathbb{Q}}[(S_T - K)^+ | \mathcal{F}_t] \\
&= e^{-r(T-t)} \int_{\mathbb{R}} e^{\phi(T-t,0,w+i\lambda) + \psi_1(T-t,0,w+i\lambda)X_t + (w+i\lambda)Y_t} \tilde{f}(\lambda) d\lambda,
\end{aligned}$$

where

$$\tilde{f}(\lambda) = \frac{1}{2\pi} \frac{K^{-(w-1+i\lambda)}}{(w+i\lambda)(w-1+i\lambda)}$$

and ϕ *and* ψ *are given as in* (2.5)*.*

Proof. Since the exponential function is a strictly monotonic continuous function on $(-\infty, \infty)$ with values in $(0, \infty)$, it holds that $\forall S_T > 0 \; \exists y \in \mathbb{R} : e^y = S_T$.
From Lemma 2.4.1 we know that the integrable function $\tilde{f}(\lambda) = \frac{1}{2\pi} \frac{K^{-(w-1+i\lambda)}}{(w+i\lambda)(w-1+i\lambda)}$ and $f(y) = (e^y - K)^+$ satisfy condition (1.10) of Theorem 1.3.2.
One has also to note that $\psi_2(T-t, 0, w+i\lambda) = w + i\lambda$.

Then by applying Theorem 1.3.2, we obtain

$$\begin{aligned}
\pi(t) &= \int_{\mathbb{R}} e^{\phi(T-t,0,w+i\lambda) - r(T-t) + \psi(T-t,0,w+i\lambda)^\top \mathcal{X}_t} \tilde{f}(\lambda) d\lambda \\
&= e^{-r(T-t)} \int_{\mathbb{R}} e^{\phi(T-t,0,w+i\lambda) + \psi_1(T-t,0,w+i\lambda) X_t + \psi_2(T-t,0,w+i\lambda) Y_t} \tilde{f}(\lambda) d\lambda \\
&= e^{-r(T-t)} \int_{\mathbb{R}} e^{\phi(T-t,0,w+i\lambda) + \psi_1(T-t,0,w+i\lambda) X_t + (w+i\lambda) Y_t} \tilde{f}(\lambda) d\lambda.
\end{aligned}$$

$\square$

Chapter 3

Wishart Processes

The Wishart processes which belong to the class of affine diffusion processes on positive semidefinite matrices are introduced and studied in this chapter. As already discussed some chapters before, the characterizing property of affine processes is that their (conditional) characteristic function is exponential affine.

Henceforward, throughout the remaining chapters, the following notation is used: We denote by $\mathcal{S}_d^{++}$ the set of all symmetric positive definite $d \times d$ matrices, by $\mathcal{S}_d^{+}$ the set of all symmetric positive semidefinite matrices, by $\mathcal{S}_d^{-}$ the set of all symmetric negative semidefinite matrices, by $\mathcal{S}_d^{--}$ the set of all symmetric negative definite $d \times d$ matrices and by $\mathcal{S}_d$ the set of all symmetric matrices with scalar product $\langle x, y \rangle = tr(xy)$. Note that the Euclidean scalarproduct of two $d \times d$ matrices, for example u and X_t, equals the trace of their product, because

$$\langle u, X_t \rangle = \sum_i \sum_j u_{ij}(X_t)_{ji} = \sum_i (uX_t)_{ii} = tr(uX_t).$$

From now on $\sqrt{A}$ denotes the unique symmetric positive semidefinite square root of a matrix A.
The square root of a matrix A can be defined by applying the spectral theorem for symmetric matrices: According to the spectral theorem there exists an orthogonal matrix Q ($Q^\top Q = QQ^\top = \mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix) and a diagonal matrix D such that

$$Q^\top A Q = D$$

and by defining $D^{\frac{1}{2}} = diag(\sqrt{d_1}, \ldots, \sqrt{d_n})$ one gets

$$A^{\frac{1}{2}} = Q D^{\frac{1}{2}} Q^\top.$$

In this chapter, we consider at first the Wishart processes in the form introduced by Marie-France Bru in 1980. Later we study the kind of Wishart processes which is used in this thesis.

3.1 Introduction of the Wishart Process in the sense of Bru

For historical reasons, we start by presenting the approach as given in [3].

3.1.1 Motivation: Wishart Processes as a Generalization of the Famous Squared Bessel Processes

As a starting point we consider an $\mathbb{R}^n$-Brownian motion $B_t = (B_{t,1}, \ldots, B_{t,n})$, where $B_{t,i}$, $i = 1, \ldots, n$ denote independent one-dimensional Brownian motions, hence the quadratic covariation is given by $\langle B_{t,i}, B_{t,j} \rangle = \mathbf{1}_{(i=j)} t$. Then we define a process X_t by $X_t = \|B_t\|^2 = \sum_{i=1}^n B_{t,i}^2$, where $\|\cdot\|$ represents the Euclidean norm.
The stochastic differential equation of X_t can be calculated by applying Ito's formula as follows:

$$\begin{aligned}
dX_t &= \sum_{i=1}^n 2B_{t,i} dB_{t,i} + \frac{1}{2} \sum_{i,j=1}^n 2d \underbrace{\langle B_{t,i}, B_{t,j} \rangle}_{=\mathbf{1}_{(i=j)} t} \\
&= 2 \sum_{i=1}^n B_{t,i} dB_{t,i} + \sum_{i=1}^n 1 dt \\
&= 2 \sum_{i=1}^n B_{t,i} dB_{t,i} + ndt.
\end{aligned} \tag{3.1}$$

To further simplify the last equation, we define $W_t := \sum_{i=1}^n \int_0^t \frac{B_{s,i}}{\|B_t\|} dB_{s,i}$.

The quadratic variation of W_t equals

$$\begin{aligned}
\langle W \rangle_t &= \left\langle \sum_{i=1}^n \int_0^t \frac{B_{s,i}}{\|B_t\|} dB_{s,i} \right\rangle \\
&= \int_0^t \sum_{i=1}^n \frac{B_{s,i} B_{s,i}}{\|B_t\|^2} d\langle B_{s,i} \rangle \\
&= \int_0^t \frac{\overbrace{\sum_{i=1}^n B_{s,i} B_{s,i}}^{=\|B_t\|^2}}{\|B_t\|^2} ds \\
&= \int_0^t 1 ds = t
\end{aligned}$$

and by applying the well-known Levy's characterization theorem, it follows that W_t is a Brownian motion.

Since $X_t = \|B_t\|^2$ and hence $\sqrt{X_t} = \|B_t\|$, W_t can be rewritten as

$$\begin{aligned} W_t &= \sum_{i=1}^{n} \int_0^t \frac{B_{s,i}}{\|B_t\|} dB_{s,i} \\ &= \sum_{i=1}^{n} \int_0^t \frac{B_{s,i}}{\sqrt{X_t}} dB_{s,i}, \end{aligned}$$

which can be in differential notation expressed as

$$dW_t = \sum_{i=1}^{n} \frac{B_{t,i}}{\sqrt{X_t}} dB_{t,i} \qquad \Leftrightarrow \qquad \sqrt{X_t} dW_t = \sum_{i=1}^{n} B_{t,i} dB_{t,i}. \tag{3.2}$$

Hence (3.1) can be rewritten as

$$dX_t = 2 \sum_{i=1}^{n} B_{t,i} dB_{t,i} + ndt = 2\sqrt{X_t} dW_t + ndt. \tag{3.3}$$

As mentioned in [3, p.725], the generator of this SDE $dX_t = 2\sqrt{X_t} dW_t + ndt$ equals $2xD^2 + nD$, where $D = \frac{d}{dx}$.

If we now consider an arbitrary non negative real number $\delta \geq 0$, the process generated by $2xD^2 + \delta D$ is called a squared Bessel process $BESQ(\delta)$.

In literature, for example in the book "Continuous Martingales and Brownian Motion" published by Daniel Revuz and Marc Yor, one can find the following definition of squared Bessel processes (see [16, p.409-410]):

Definition 3.1.1. Let B be a δ-dimensional Brownian motion and define $\rho = \|B\|^2$. For every $\delta \in \mathbb{N}$ and $x \geq 0$, the unique strong solution of the equation

$$Z_t = x + 2 \int_0^t \sqrt{Z_s} d\beta_s + \delta t,$$

where

$$\beta_t = \sum_{i=1}^{\delta} \int_0^t \frac{B_{s,i}}{\rho_s} dB_{s,i}$$

is called the *square of δ-dimensional Bessel process* started at x and is denoted by $BESQ(\delta)$.

According to [3, p.725], the density of a Bessel process $BESQ(\delta)$ with initial value $X_0 = x$ is given by a Bessel function with Laplace transform
$\mathbb{E}\left[e^{-\lambda X_t} | X_0 = x\right] = (1 + 2\lambda t)^{-\frac{\delta}{2}} e^{-\frac{\lambda x}{1+2\lambda t}}$. The density of a Bessel process equals a non-central χ^2-distribution.

Instead of vectors, Marie-France Bru has considered matrices of independent Brownian motions. To be more specific, we now take a look at a sample $(B_1, \ldots, B_n)$ of $\mathbb{R}^d$ Gaussian vectors and denote by B a matrix of dimension $n \times d$ with the vector B_i in the i-th row. Then the matrix $B^\top B$ follows a so-called Wishart distribution[1] which is a natural generalization of the well-known χ^2 distribution. The density of the Wishart distribution equals a Bessel matrix function.

Remark 3.1.2. Note that in this setup, the input of generating a Wishart distribution are Gaussian vectors. Later we will see that there exists also a more general possibility to generate a Wishart process, namely by taking squares of Ornstein-Uhlenbeck processes. Clearly, these Ornstein-Uhlenbeck processes are still kind of Gaussian but they include an additional drift.

That is the intuition why Marie-France Bru has started to examine extensions of squared Bessel processes, to be more precise matrix versions of them.

In her paper she has defined Wishart processes step by step. At first she has considered simple versions of them, and in each section she has generalized them (see [3]).

At the beginning she has defined Wishart processes as follows (see [3, p.726-727]):

A Brownian matrix is a process taking its values in the set of real-valued $n \times d$ matrices whose components are independent Brownian motions.

Let n and d be two integers ≥ 1, and N_t be a $n \times d$ Brownian matrix with initial state $N_0 = C$.

Definition 3.1.3. A *Wishart process in the sense of Bru* of dimension d, index n and initial state x_0 is a matrix process

$$X_t = N_t^\top N_t \qquad \text{with} \qquad x_0 = C^\top C. \tag{3.4}$$

This Wishart process is denoted by $WIS(n, d, x_0)$.

Remark 3.1.4. If we set $d = 1$, we get the special case of a squared Bessel process $WIS(n, 1, x_0)$ of index n and with starting value x_0.

Since we have already motivated the topic of Wishart processes, we start by studying the most general version of the Wishart process introduced in [3]. Later we will also examine less general versions of Wishart processes by introducing them as special cases of the most general one.

3.1.2 The (most general) Wishart Process studied by Bru

In this subsection, a Wishart process determined by five parameters is studied. As already mentioned, this Wishart process presents the most generalized version of the Wishart process, which is examined in section "Five-Parameter Wishart Processes; Square Ornstein-Uhlenbeck Processes - Matrix Case" in [3, p.745-749].

[1] The Wishart distribution was first mentioned and introduced by John Wishart in his paper "The generalised product moment distribution in samples from a normal multivariate population" published in 1928 (compare [18]).

The setting of this Wishart process characterized by five parameters can be summarized as follows:

Denote by V_t a $n \times d$ matrix of stochastic processes which is a solution of the following stochastic differential equation

$$dV_t = dN_t a + V_t b dt, \qquad V_0 = v_0, \tag{3.5}$$

where (N_t) is a $n \times d$ Brownian matrix (matrix of independent one-dimensional Brownian motions) and a and b denote $d \times d$ matrices.

Then we define the matrix-valued process

$$X_t = V_t^\top V_t \tag{3.6}$$

with initial value $x_0 = v_0^\top v_0$.

The dynamics of X_t are examined in the following proposition:

Proposition 3.1.5. *Let $b \in \mathcal{S}_d^-$. The dynamics of the matrix process X_t are given by*

$$dX_t = a^\top dN_t^\top V_t + V_t^\top dN_t a + (bX_t + X_t b)dt + na^\top a dt. \tag{3.7}$$

Proof. To verify (3.7), we examine the dynamics of X_t component-by-component at first: For that we have to note that the (i,j)-th component of X_t is obtained by $(X_t)_{ij} = \sum_{k=1}^n (V_t^\top)_{ik}(V_t)_{kj}$.

Then by applying integration by parts formula, the dynamics of the (i,j)-th component of the matrix-valued process X_t are given by

$$\begin{aligned}
d((X_t)_{ij}) =& d(\sum_{k=1}^n (V_t^\top)_{ik}(V_t)_{kj}) \\
=& \sum_{k=1}^n d((V_t^\top)_{ik}(V_t)_{kj}) \\
=& \sum_{k=1}^n \Big(d(V_t^\top)_{ik}(V_t)_{kj} + (V_t^\top)_{ik} d(V_t)_{kj} + d\langle (V_t^\top)_{ik}, (V_t)_{kj}\rangle \Big) \\
=& \sum_{k=1}^n d(V_t^\top)_{ik}(V_t)_{kj} + \sum_{k=1}^n (V_t^\top)_{ik} d(V_t)_{kj} + \sum_{k=1}^n d\langle (V_t^\top)_{ik}, (V_t)_{kj}\rangle.
\end{aligned} \tag{3.8}$$

Now we do some auxiliary calculations:
Therefore one has to note that the (i,j)-th component of V_t can be written as

$$(dV_t)_{ij} = (dN_t a)_{ij} + (V_t b dt)_{ij} = \sum_{l=1}^d (dN_t)_{il} a_{lj} + \sum_{l=1}^d (V_t)_{il} b_{lj} dt.$$

Moreover, since $V_t^\top$ is determined by $dV_t^\top = a^\top dN_t^\top + b^\top V_t^\top dt$, the (i,j)-th component of $V_t^\top$ is given by $(dV_t^\top)_{ij} = (a^\top dN_t^\top)_{ij} + (b^\top V_t^\top)_{ij}dt = \sum_{l=1}^d (a^\top)_{il}(dN_t^\top)_{lj} + \sum_{l=1}^d (b^\top)_{il}(V_t^\top)_{lj}dt$.

The first term of the last row of (3.8) can further be calculated by

$$\begin{aligned}
\sum_{k=1}^n d(V_t^\top)_{ik}(V_t)_{kj} &= \sum_{k=1}^n \left(\sum_{l=1}^d (a^\top)_{il}(dN_t^\top)_{lk} + \sum_{l=1}^d (b^\top)_{il}(V_t^\top)_{lk}dt\right)(V_t)_{kj} \\
&= \sum_{k=1}^n \sum_{l=1}^d (a^\top)_{il}(dN_t^\top)_{lk}(V_t)_{kj} + \sum_{k=1}^n \sum_{l=1}^d (b^\top)_{il}(V_t^\top)_{lk}(V_t)_{kj}dt \\
&= (a^\top dN_t^\top V_t)_{ij} + (b^\top V_t^\top V_t)_{ij}dt \\
&= (a^\top dN_t^\top V_t)_{ij} + (b^\top X_t)_{ij}dt \\
&= (a^\top dN_t^\top V_t)_{ij} + (bX_t)_{ij}dt,
\end{aligned} \tag{3.9}$$

where the last step follows because $b \in \mathcal{S}_d^-$.

It follows by the same step that for the second term of equation (3.8) we get

$$\begin{aligned}
\sum_{k=1}^n (V_t^\top)_{ik} d(V_t)_{kj} &= \sum_{k=1}^n (V_t^\top)_{ik}\left(\sum_{l=1}^d (dN_t)_{kl}a_{lj} + \sum_{l=1}^d (V_t)_{kl}b_{lj}dt\right) \\
&= \sum_{k=1}^n \sum_{l=1}^d (V_t^\top)_{ik}(dN_t)_{kl}a_{lj} + \sum_{k=1}^n \sum_{l=1}^d (V_t^\top)_{ik}(V_t)_{kl}b_{lj}dt \\
&= (V_t^\top dN_t a)_{ij} + (V_t^\top V_t b)_{ij}dt \\
&= (V_t^\top dN_t a)_{ij} + (X_t b)_{ij}dt.
\end{aligned} \tag{3.10}$$

Considering the last term of equation (3.8), namely the quadratic covariation, yields

$$\begin{aligned}
\sum_{k=1}^n d\langle (V_t^\top)_{ik}, (V_t)_{kj}\rangle &= \sum_{k=1}^n \left\langle \sum_{l=1}^d (a^\top)_{il}(dN_t^\top)_{lk}, \sum_{m=1}^d (dN_t)_{km}a_{mj}\right\rangle \\
&= \sum_{k=1}^n \sum_{l=1}^d \sum_{m=1}^d (a^\top)_{il}a_{mj}\langle \underbrace{(dN_t^\top)_{lk}}_{(dN_t)_{kl}}, (dN_t)_{km}\rangle \\
&= \sum_{k=1}^n \sum_{l=1}^d \sum_{m=1}^d (a^\top)_{il}a_{mj}\underbrace{\langle (dN_t)_{kl}, (dN_t)_{km}\rangle}_{=\mathbf{1}_{(l=m)}dt} \\
&= \sum_{k=1}^n \sum_{l=1}^d (a^\top)_{il}a_{lj}dt \\
&= \sum_{k=1}^n (a^\top a)_{ij}dt \\
&= n(a^\top a)_{ij}dt,
\end{aligned} \tag{3.11}$$

because the deterministic terms (dt-terms) vanish in the quadratic covariation.

Plugging (3.9), (3.10) and (3.11) into (3.8), (3.8) can be rewritten as

$$\begin{aligned}d((X_t)_{ij}) &= (a^\top dN_t^\top V_t)_{ij} + (bX_t)_{ij}dt + (V_t^\top dN_t a)_{ij} + (X_t b)_{ij}dt + n(a^\top a)_{ij}dt \\ &= (a^\top dN_t^\top V_t)_{ij} + (V_t^\top dN_t a)_{ij} + (bX_t + X_t b)_{ij}dt + n(a^\top a)_{ij}dt.\end{aligned}$$

Hence for the $d \times d$ matrix-valued stochastic process X_t it holds that

$$dX_t = a^\top dN_t^\top V_t + V_t^\top dN_t a + (bX_t + X_t b)dt + na^\top a dt.$$

□

As a next step we define a stochastic matrix by $dB_t = \sqrt{X_t^{-1}} V_t^\top dN_t a(\sqrt{a^\top a})^{-1}$ with $n \geq d+1$ and show that this stochastic matrix is a $d \times d$ Brownian matrix. In addition to that, also the SDE fulfilled by the matrix process X_t is derived in the subsequent proposition:

Proposition 3.1.6. *Let X_t be given by* (3.6) *with $n \geq d+1$ and let $b \in \mathcal{S}_d^-$.*
The $d \times d$ matrix process given by

$$dB_t = \sqrt{X_t^{-1}} V_t^\top dN_t a(\sqrt{a^\top a})^{-1}$$

is a $d \times d$ Brownian matrix. With this definition of the Brownian matrix B_t, the matrix process X_t solves the stochastic differential equation

$$dX_t = \sqrt{X_t} dB_t \sqrt{a^\top a} + \sqrt{a^\top a} dB_t^\top \sqrt{X_t} + (bX_t + X_t b)dt + na^\top a dt, \qquad X_0 = x_0. \tag{3.12}$$

Proof. By applying Levy's characterization theorem we show at first that the $d \times d$ matrix B_t equals a Brownian matrix.

By noting that the (i,j)-th entry of the matrix B_t is given by its dynamics

$$dB_{t,ij} = \sum_{m=1}^{n} \sum_{q=1}^{n} (\sqrt{X_t^{-1}} V_t^\top)_{im} (dN_t)_{mq} (a(\sqrt{a^\top a})^{-1})_{qj},$$

we obtain for the quadratic covariation

$$
\begin{aligned}
\langle dB_{t,ij}, dB_{t,kl}\rangle &= \Big\langle \sum_{m=1}^{n}\sum_{q=1}^{n}(\sqrt{X_t^{-1}}V_t^\top)_{im}(dN_t)_{mq}(a(\sqrt{a^\top a})^{-1})_{qj},\\
&\qquad \sum_{u=1}^{n}\sum_{w=1}^{n}(\sqrt{X_t^{-1}}V_t^\top)_{ku}(dN_t)_{uw}(a(\sqrt{a^\top a})^{-1})_{wl}\Big\rangle\\
&= \sum_{m=1}^{n}\sum_{q=1}^{n}\sum_{u=1}^{n}\sum_{w=1}^{n}(\sqrt{X_t^{-1}}V_t^\top)_{im}(\sqrt{X_t^{-1}}V_t^\top)_{ku}(a(\sqrt{a^\top a})^{-1})_{qj}(a(\sqrt{a^\top a})^{-1})_{wl}\,\underbrace{d\langle (N_t)_{mq},(N_t)_{uw}\rangle}_{=\mathbf{1}_{(m,q)=(u,w)}dt}\\
&= \sum_{m=1}^{n}\sum_{q=1}^{n}(\sqrt{X_t^{-1}}V_t^\top)_{im}(\sqrt{X_t^{-1}}V_t^\top)_{km}(a(\sqrt{a^\top a})^{-1})_{qj}(a(\sqrt{a^\top a})^{-1})_{ql}dt\\
&= \sum_{m=1}^{n}\Big((\sqrt{X_t^{-1}}V_t^\top)_{im}(V_t\sqrt{X_t^{-1}})_{mk}\Big)\sum_{q=1}^{n}\Big(((\sqrt{a^\top a})^{-1}a^\top)_{jq}(a(\sqrt{a^\top a})^{-1})_{ql}\Big)\,dt\\
&= (\sqrt{X_t^{-1}}V_t^\top V_t\sqrt{X_t^{-1}})_{ik}((\sqrt{a^\top a})^{-1}a^\top a(\sqrt{a^\top a})^{-1})_{jl}dt\\
&= (X_t^{\frac{1}{2}}X_tX_t^{-\frac{1}{2}})_{ik}((a^\top a)^{-\frac{1}{2}}(a^\top a)(a^\top a)^{-\frac{1}{2}})_{jl}dt\\
&= \mathbf{I}_{ik}\mathbf{I}_{jl}dt\\
&= \mathbf{1}_{(i,j)=(k,l)}dt.
\end{aligned}
$$

To prove the second part of the above proposition, one has to note that $B_t^\top$ is governed by its dynamics

$$
\begin{aligned}
dB_t^\top &= ((\sqrt{a^\top a})^{-1})^\top a^\top dN_t^\top (V_t^\top)^\top(\sqrt{X_t^{-1}})^\top\\
&= (\sqrt{a^\top a})^{-1}a^\top dN_t^\top V_t\sqrt{X_t^{-1}}.
\end{aligned}
$$

Therefore one can verify that (3.12) is true by

$$
\begin{aligned}
dX_t &= \sqrt{X_t}dB_t\sqrt{a^\top a} + \sqrt{a^\top a}dB_t^\top\sqrt{X_t} + (bX_t + X_tb)dt + na^\top a dt\\
&= \sqrt{X_t}\sqrt{X_t^{-1}}V_t^\top dN_t a(\sqrt{a^\top a})^{-1}\sqrt{a^\top a} + \sqrt{a^\top a}(\sqrt{a^\top a})^{-1}a^\top dN_t^\top V_t\sqrt{X_t^{-1}}\sqrt{X_t}\\
&\quad + (bX_t + X_tb)dt + na^\top a dt\\
&= V_t^\top dN_t a + a^\top dN_t^\top V_t + (bX_t + X_tb)dt + na^\top a dt,
\end{aligned}
$$

which was proved in (3.7). □

We can now state an important general existence theorem about the SDE (3.12):

Theorem 3.1.7. *(see [3, Theorem 2″]) If $\alpha \in \Delta_d = \{1,\dots,d-1\}\cup(d-1,+\infty)$, a is in the group of invertible $d\times d$ matrices, $b\in\mathcal{S}_d^-$, x_0 is in $\mathcal{S}_d^+$ and has all its eigenvalues distinct, and (B_t) is a $d\times d$ Brownian matrix, then on $[0,\tau)$ (τ denotes the first eigenvalue collision time),* *the stochastic differential equation*

$$dX_t = \sqrt{X_t}dB_t\sqrt{a^\top a} + \sqrt{a^\top a}dB_t^\top\sqrt{X_t} + (bX_t + X_tb)dt + \alpha\sqrt{a^\top a}dt, \qquad X_0 = x_0 \tag{3.13}$$

has a unique solution (in the sense of probability law) if b and $\sqrt{a^\top a}$ commute.

After stating the last theorem, we can formally give a definition of this class of Wishart processes:

Definition 3.1.8. A matrix process on $\mathcal{S}_d^+$ governed by the stochastic differential equation given in (3.13) with initial value $X_0 = x_0$ is called *Wishart process in the sense of Bru* with index α, dimension d, initial state x_0, and matrix parameters b and a. According to [3], this stochastic process is denoted by $WIS(\alpha, b, a, d, x_0)$.

3.1.3 Characteristic Function of this most general Wishart Process introduced by Bru

In this subsection, the distribution of X_t for fixed t is studied. Since it is a well-known fact that the distribution of a random variable/process is uniquely determined by its characteristic function, the characteristic function is now given in the following theorem:

Theorem 3.1.9. *(see [3, p.749]) If $(X_t) \in WIS(\alpha, b, a, d, x_0)$, where $\alpha \in \Delta_d$, a is in the group of invertible $d \times d$ matrices, $b \in \mathcal{S}_d^{--}$ commutes with $\sqrt{a^\top a}, x_0 \in \mathcal{S}_d^{++}, u \in i\mathcal{S}_d$, the characteristic function of the stochastic matrix-valued process X_t can be expressed as*

$$\begin{aligned}\mathbb{E}[e^{\langle u, X_t\rangle}|X_0 = x_0] &= \mathbb{E}[e^{tr(uX_t)}|X_0 = x_0] \\ &= (\det b^{-1}(b + ua^Ta - ua^Tae^{2bt}))^{-\frac{\alpha}{2}}e^{tr(e^{bt}x_0e^{bt}b(b+ua^Ta-ua^Tae^{2bt})^{-1}u)}.\end{aligned} \tag{3.14}$$

Remark 3.1.10. Since (3.14) can be rewritten as

$$\mathbb{E}[e^{tr(uX_t)}|X_0 = x_0] = e^{ln((\det b^{-1}(b+ua^Ta-ua^Tae^{2bt}))^{-\frac{\alpha}{2}})+tr(e^{bt}x_0e^{bt}b(b+a^Ta-ua^Tae^{2bt})^{-1}u)},$$

one can easily see that the characteristic function is exponential affine in the initial state x_0 and hence, by the definition of affine processes studied some chapters before, it follows that the Wishart process $WIS(\alpha, b, a, d, x_0)$ defined in Definition 3.1.8 is an affine process.

3.1.4 Literature Review of Wishart Processes introduced by Bru

As previously discussed, Marie-France Bru has introduced in her paper [3] Wishart processes step by step, which means that by starting with simple versions, she has defined them more and more generally. We will now give a short overview of the different forms of Wishart processes she has studied but in this thesis we examine them from the opposite direction:

While Marie-France Bru has used an inductive way to introduce Wishart processes, we will now shortly study them from an deductive perspective. In other words, we will examine the less general forms by regarding them as special cases of the most general one.

If we restrict the $d \times d$ matrices a and b to be only one-dimensional parameters γ and β, we are in the case of section "Five Parameter Wishart Processes: Square Ornstein-Uhlenbeck Processes - Real Case" of [3, p.743-745], where the following kind of Wishart processes $WIS(\alpha, \beta, \gamma, d, x_0)$ is studied.

Denote by V_t a $n \times d$ matrix process which fulfills the stochastic differential equation

$$dV_t = \gamma dN_t + \beta V_t dt, \qquad V_0 = v_0, \tag{3.15}$$

where N_t is a $n \times d$ Brownian matrix, v_0 is a $n \times d$ deterministic matrix, $\gamma \in \mathbb{R}$, $\beta \in \mathbb{R}_-$. Then we define the matrix-valued stochastic process X_t by $X_t = V_t^\top V_t$ with initial value $x_0 = v_0^\top v_0$.

Proposition 3.1.11. *Let $n \geq d+1$. The $d \times d$ matrix process B_t governed by*

$$dB_t = \sqrt{X_t^{-1}} V_t^\top dN_t$$

is a $d \times d$ Brownian matrix.
With this definition of the Brownian matrix B_t, the matrix process X_t is a solution of the following stochastic differential equation:

$$dX_t = \gamma(\sqrt{X_t} dB_t + dB_t^\top \sqrt{X_t}) + 2\beta X_t dt + n\gamma^2 \boldsymbol{I} dt, \quad X_0 = x_0. \tag{3.16}$$

Proof. The proposition above follows by considering Proposition 3.1.6 for the special case where the $d \times d$ matrices a and b are replaced by the real-valued parameters γ and β. □

Similar as before, we obtain the following theorem:

Theorem 3.1.12. *(see [3, Theorem 2′]) If $(B_t)_{t \geq 0}$ is a $d \times d$ Brownian matrix, then for all $\gamma \in \mathbb{R}$, $\beta \in \mathbb{R}$, and $x_0 \in \mathcal{S}_d^+$ with distinct eigenvalues labeled $\lambda_1(0) > \ldots > \lambda_d(0) \geq 0$, the stochastic differential equation*

$$dX_t = \gamma(\sqrt{X_t} dB_t + dB_t^\top \sqrt{X_t}) + 2\beta X_t dt + \alpha\gamma^2 \boldsymbol{I} dt \tag{3.17}$$

has

- *a unique solution in $\mathcal{S}_d$ (in the sense of probability law) if $\alpha \in (d-1, d+1)$ and*
- *a unique strong solution in $\mathcal{S}_d^{++}$ if $\alpha \geq d+1$.*

The eigenvalues of such a solution never collide: a.s. for all $t > 0$

$$\lambda_1(t) > \ldots > \lambda_d(t) \geq 0, \quad \lambda_d(t) > 0 \quad \textit{if} \quad \alpha \geq d+1$$

and satisfy the stochastic differential system

$$d\lambda_i = 2\sqrt{\lambda_i}dv_i + \alpha\gamma^2 dt + 2\beta\lambda_i dt + \gamma^2 \sum_{k\neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} dt,$$

where $v_1(t), \dots, v_d(t)$ are independent Brownian motions.

In such a way the Wishart process in the sense of Bru with index α, dimension d, initial state x_0, and real parameters β and γ, denoted by $WIS(\alpha, \beta, \gamma, d, x_0)$, is defined.

As in the previous subsection, the distribution of X_t for fixed t can be determined by its characteristic function as stated in the following corollary:

Corollary 3.1.13. *(see [3, p.749]) If $(X_t) \in WIS(\alpha, \beta, \gamma, d, x_0)$ where $\alpha \in \Delta_d$, $\beta \in \mathbb{R}_-, \gamma \in \mathbb{R}$, $x_0 \in \mathcal{S}_d^{++}, u \in i\mathcal{S}_d$, the characteristic function of the stochastic matrix-valued process X_t is given by*

$$\begin{aligned}\mathbb{E}[e^{\langle u, X_t\rangle}|X_0 = x_0] &= \mathbb{E}[e^{tr(uX_t)}|X_0 = x_0] \\ &= \left(\det\left(\frac{\beta\boldsymbol{I} + \gamma^2 u - e^{2\beta t}\gamma^2 u}{\beta}\right)\right)^{-\frac{\alpha}{2}} e^{\beta e^{2\beta t} tr(x_0(\beta\boldsymbol{I}+\gamma^2 u - e^{2\beta t}\gamma^2 u)^{-1}u)}.\end{aligned} \tag{3.18}$$

If we consider the special case where the drift of (3.5) vanishes and the volatility is assumed to be the identity matrix, which means that we consider a purely Gaussian case, we are in section "Generalization. The $WIS(\alpha, d, x_0)$ Process" of [3].

Hence by choosing the $d \times d$ matrix a to be the identity matrix $\mathbf{I}$ and the $d \times d$ matrix b to be the null matrix, (3.5) simplifies to

$$dV_t = dN_t, \qquad V_0 = v_0.$$

Therefore in this framework, X_t simplifies to $X_t = N_t^\top N_t, \qquad V_0 = v_0$.

Remark 3.1.14. Clearly, setting $\gamma = 1$ and $\beta = 0$ in Proposition 3.1.11 and Theorem 3.1.12, these statements are also valid for this setup.

Remark 3.1.15. As already mentioned before, Marie-France Bru has started by introducing the Wishart process for integer-valued index n, later she has replaced the integer-valued index n by a real-valued index α. If we set $\gamma = 1$ and $\beta = 0$ and assume that the index α is integer-valued in (3.17), we are exactly in the case where the matrix-valued process X_t is of form (3.4).

There exists extensive literature on the subject of Wishart processes. Some authors have also tried to introduce and examine Wishart processes from a point of view strongly related to affine processes which has become more and more popular.

In the subsequent section we state the definition of the Wishart process which will be used during this thesis.

3.2 The kind of Wishart Processes studied and examined in this thesis

According to Eberhard Mayerhofer for example (see [14, p.6]), the concept of Wishart processes can be introduced and defined as follows:

Definition 3.2.1. Let $\sqrt{X}$ denote the unique, positive semidefinite matrix square root on the space of symmetric positive semidefinite $d \times d$ matrices $\mathcal{S}_d^+$. Let Σ, M be real valued $d \times d$ matrices, $\delta \geq d-1$. As *Wishart process* we define the stochastic process given by the stochastic differential equation

$$dX_t = \sqrt{X_t} dB_t \Sigma + \Sigma^\top dB_t^\top \sqrt{X_t} + (\delta \Sigma^\top \Sigma + M X_t + X_t M^\top) dt, \qquad X_0 = x \in \mathcal{S}_d^+, \qquad (3.19)$$

where B is a standard $d \times d$ Brownian motion matrix.
Hence a Wishart process is characterized by its parameters (δ, M, Σ).

Remark 3.2.2. Comparing (3.19) with (3.13), we see that these two definitions of Wishart processes are quite similar. If we set $\Sigma = \sqrt{a^\top a}$ and $M = b$ in (3.19), only the constant part of the drift differs in these two dynamics. But there is also a difference in the domain of definition of the characterizing matrices. While Marie-France Bru has assumed that $b \in \mathcal{S}_d^-$ and b and $\sqrt{a^\top a}$ commute, we do not make such restrictions on the corresponding matrices M and Σ in the Wishart process (3.19) used in this thesis.
Hence the Wishart process defined in (3.19) presents an even more general form than the Wishart process in the sense of Bru studied in the last section.

As a next step we examine the uniqueness of the solution of the Wishart process given in (3.19) (compare [14, p.15]): For that we suppose that $X_0 = x$ is positive definite in (3.19). Then, according to [14], there exists a unique strong solution of the Wishart process as long as X_t does not hit the boundary. This first hitting time of the boundary is defined by

$$T_x := \inf\{t > 0 | \det(X) = 0\}.$$

If the boundary is never reached by the stochastic process, which means that $T_x = \infty$, unique strong solutions of the Wishart process always exist.

There exists an important relation between the first hitting time and the dimension d and the constant part of the drift δ, which is mentioned in the subsequent theorem:

Theorem 3.2.3. *[14, Theorem 3.1.] Suppose $\delta \geq d+1$. Then $T_x = \infty$ almost surely.*

Summarizing all these facts leads to the following important corollary:

Corollary 3.2.4. *If $\delta \geq d+1$, then there exists a unique strong solution of the Wishart process as given in* (3.19).

In the next proposition the distribution of the Wishart process in (3.19) is stated:

Proposition 3.2.5. *The Wishart process as defined in* (3.19) *follows a non-central Wishart distribution* [2]*.*

3.2.1 Motivation and Introduction of the Wishart Distribution

At first, a motivation of the Wishart distribution is given (compare: [14, p.2-3]):

Denote by $\xi_1, \ldots, \xi_k$ a sequence of independent $\mathbb{R}^d$-valued random variables which are normally distributed with mean vector $\mu_i \in \mathbb{R}^d$ and covariance matrix Σ.

Then the random variable defined by $\Xi := \xi_1 \xi_1^\top + \ldots + \xi_k \xi_k^\top$ follows a Wishart distribution, i.e. $\Gamma(p, w; \sigma)$, with scale parameter $p := \frac{k}{2}$, shape parameter $\sigma := 2\Sigma$ and parameter of non-centrality $w := \sum_{i=1}^k \mu_i \mu_i^\top$.

Before a formal definition of the Wishart distribution is given, the Laplace transform is calculated for a special case of the Wishart distribution, namely for the case of one-dimensional random variables ($d = 1$).
In this case the Laplace transform can be easily calculated as follows:

$$
\begin{aligned}
\mathbb{E}[e^{-u\xi_j^2}] &= \frac{1}{\sqrt{2\pi\Sigma}} \int_{\mathbb{R}} e^{-u\eta^2 - \frac{(\eta-\mu_j)^2}{2\Sigma}} d\eta \\
&= \frac{1}{\sqrt{2\pi\Sigma}} \int_{\mathbb{R}} e^{\frac{-(1+2\Sigma u)\eta^2 + 2\eta\mu_j - \mu_j^2}{2\Sigma}} d\eta \\
&= \frac{1}{\sqrt{2\pi\Sigma}} \int_{\mathbb{R}} e^{\frac{-(1+2\Sigma u)^2\eta^2 + 2(1+2\Sigma u)\eta\mu_j - (1+2\Sigma u)\mu_j^2}{2\Sigma(1+2\Sigma u)}} d\eta \\
&= \frac{1}{\sqrt{2\pi\Sigma}} \int_{\mathbb{R}} e^{\frac{-((1+2\Sigma u)\eta - \mu_j)^2 - 2\Sigma u\mu_j^2}{2\Sigma(1+2\Sigma u)}} d\eta \\
&= \frac{1}{\sqrt{2\pi\Sigma}} \int_{\mathbb{R}} e^{-\frac{1+2\Sigma u}{2\Sigma}\left(\frac{\eta(1+2\Sigma u)-\mu_j}{1+2\Sigma u}\right)^2 - \frac{u\mu_j^2}{1+2\Sigma u}} d\eta \\
&= \frac{1}{\sqrt{2\pi\Sigma}} \int_{\mathbb{R}} e^{-\frac{1+2\Sigma u}{2\Sigma}\left(\eta - \frac{\mu_j}{1+2\Sigma u}\right)^2} e^{-\frac{u\mu_j^2}{1+2\Sigma u}} d\eta \\
&= e^{-\frac{u\mu_j^2}{1+2\Sigma u}} \frac{1}{\sqrt{1+2\Sigma u}} \underbrace{\frac{1}{\sqrt{2\pi\frac{\Sigma}{1+2\Sigma u}}} \int_{\mathbb{R}} e^{-\frac{1+2\Sigma u}{2\Sigma}\left(\eta - \frac{\mu_j}{1+2\Sigma u}\right)^2} d\eta}_{(\star)=1} \\
&= \frac{e^{-(u\mu_j^2)(1+2\Sigma u)^{-1}}}{(1+2\Sigma u)^{\frac{1}{2}}}, \qquad u \geq 0.
\end{aligned}
\tag{3.20}
$$

($\star$) integral over the density of a standard normal distributed random variable $\mathcal{N}(\frac{\mu_j}{1+2\Sigma u}, \frac{\Sigma}{1+2\Sigma u})$

[2] This so-called non-central Wishart distribution is introduced and studied in the subsequent subsection

Under the assumption that $\xi_1, \ldots, \xi_k$ are independent, one gets

$$\begin{aligned}\mathbb{E}[e^{-u\Xi}] &= \mathbb{E}[e^{-u\sum_{i=1}^k \xi_i^2}] = \prod_{i=1}^k \mathbb{E}[e^{-u\xi_i^2}] \\ &= \prod_{i=1}^k \frac{e^{-(u\mu_i^2)(1+2\Sigma u)^{-1}}}{(1+2\Sigma u)^{\frac{1}{2}}} = \frac{e^{-u(1+2\Sigma u)^{-1}\sum_{i=1}^k \mu_i^2}}{(1+2\Sigma u)^{\frac{1}{2}}} = \frac{e^{-u(1+\sigma u)^{-1}w}}{(1+\sigma u)^{\frac{1}{2}}},\end{aligned}$$

where we used that $\sigma = 2\Sigma$ and $w = \sum_{i=1}^k \mu_i \mu_i^\top$ as defined before for the case where $d = 1$.

If we now assume that $d > 1$, the positive random variable Ξ becomes a positive semidefinite symmetric $d \times d$ matrix. Then a possible formal definition of the Wishart distribution is the following (see [13, p.1]):

Definition 3.2.6. The *general non-central Wishart distribution* $\Gamma(p, w; \sigma)$ on the cone $\mathcal{S}_d^+$ of symmetric positive semidefinite $d \times d$ matrices is defined (whenever it exists) by its Laplace transform

$$\mathcal{L}(\Gamma(p, w; \sigma))(u) = (\det(\mathbf{I} + \sigma u))^{-p} e^{-tr(u(\mathbf{I}+\sigma u)^{-1}w)}, \quad u \in \mathcal{S}_d^+, \tag{3.21}$$

where $p \geq 0$ denotes its shape parameter, $\sigma \in \mathcal{S}_d^+$ is the scale parameter and the parameter of non-centrality equals $w \in \mathcal{S}_d^+$.
In the special case where $w = 0$, $\Gamma(p; \sigma) := \Gamma(p, 0; \sigma)$ is called the central Wishart distribution.

Remark 3.2.7. For the setup of the Wishart process as defined in (3.19), we have to set $p = \frac{\delta}{2}$.

Remark 3.2.8. The central Wishart distribution was first mentioned in "The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population" (see [18]) written by J. Wishart in 1928.

Remark 3.2.9. Comparing (3.20) with (3.21), one can again see that the first equation is a special case of the second equation.

At the end of this subsection about the Wishart distribution, we take a brief look at the existency of the Wishart distribution:

Proposition 3.2.10. *[14, p.3-4] For invertible σ, the central Wishart distributions $\Gamma(p; \sigma) := \Gamma(p, w = 0; \sigma)$ exist if and only if p belongs to the Gindikin ensemble, which equals the set*

$$\Lambda_d := \left\{0, \frac{1}{2}, \ldots, \frac{d-1}{2}\right\} \cup \left(\frac{d-1}{2}, \infty\right).$$

Equivalently, the right hand side of (3.21) *is the Laplace transform of a distribution on $\mathcal{S}_d^+$ if and only if $p \in \Lambda_d$.*

The following properties concerning the non-central Wishart distribution are taken from [13, p.3-4]:

Before stating this proposition, some notation has to be introduced: We denote by $\partial\mathcal{S}_d^+$ the boundary $\mathcal{S}_d^+ \setminus \mathcal{S}_d^{++}$. Then for $k = 1, 2, \ldots, d-1$, $D_k \subseteq \partial\mathcal{S}_d^+$ describes the $d \times d$ matrices whose rank is less or equal to k.

Proposition 3.2.11. *[13, Lemma 2.2.] Let $p \in \Lambda_d, \sigma \in \mathcal{S}_d^+$ and $w \in \mathcal{S}_d^+$. We have*

- *Suppose $w = mm^\top$ for $m \in \mathbb{R}^d$ and set $\Sigma := \frac{\sigma}{2}$. If $Y \sim \mathcal{N}(m, \Sigma)$, then $X := YY^\top \sim \Gamma(\frac{1}{2}, w; \sigma)$ is supported on D_1.*
- *If $p < \frac{d-1}{2}$ and $rank(w) \leq 2p$, then the right hand side (3.21) is the Laplace transform of a probability measure supported on D_{2p}.*
- *If $p \geq \frac{d-1}{2}$, then the right side of (3.21) is the Laplace transform of a probability measure $\Gamma(p, w; \sigma)$ on $\mathcal{S}_d^+$.*
- *In particular, if $p > \frac{d-1}{2}$ and if σ is invertible, then the density of $\Gamma(p, w; \sigma)$ exists.*

3.2.2 The CIR Process as a Special Case of the Wishart Process

This subsection deals with an interesting relationship between the Wishart process and the CIR process. But before deducing this relation, the dynamic of the CIR process is given here once more:

The Cox-Ingersoll-Ross process (CIR process) satisfies the diffusion equation

$$dv_t = \bar{\kappa}(\theta - v_t)dt + \bar{\sigma}\sqrt{v_t}dB_t,$$

where (B_t) denotes an univariate Brownian motion and the parameters satisfy the following conditions: $\bar{\sigma} > 0$, $\bar{\kappa}\theta \geq 0$.

We now consider a Wishart process with diagonal matrices M and Σ. The result will be that this special case of a Wishart process can be seen as a multidimensional CIR process. The idea of these considerations is mainly based on [2, p.6]:

If the matrices M and Σ in the Wishart dynamics (3.19) are restricted to be diagonal matrices (hence M can be written as $M = diag(M_{11}, \ldots, M_{dd})$ while Σ can be written as $\Sigma = diag(\Sigma_{11}, \ldots, \Sigma_{dd})$), the dynamics of the diagonal elements given in (3.19) simplify to

$$
\begin{aligned}
(dX_t)_{ii} = dX_{t,ii} &= (\sqrt{X_t} dB_t)_{ii}\Sigma_{ii} + (\Sigma^\top)_{ii}(dB_t^\top \sqrt{X_t})_{ii} \\
&\quad + \Big(\delta(\Sigma^\top\Sigma)_{ii} + M_{ii}(X_t)_{ii} + (X_t)_{ii}(M^\top)_{ii}\Big)\, dt \\
&= (\sqrt{X_t} dB_t)_{ii}\Sigma_{ii} + (\Sigma)_{ii}(\underbrace{\sqrt{X_t}^\top}_{=\sqrt{X_t}} dB_t)_{ii} + \big(\delta(\Sigma^2)_{ii} + M_{ii}(X_t)_{ii} + (X_t)_{ii}(M)_{ii}\big)\, dt \\
&= 2\Sigma_{ii}(\sqrt{X_t} dB_t)_{ii} + (\delta(\Sigma^2)_{ii} + 2M_{ii}(X_t)_{ii})dt \\
&= 2\Sigma_{ii}\left(\sum_{k=1}^{d} \underbrace{(\sqrt{X_t})_{ik}}_{=(\sqrt{X_t})_{ki}}(dB_t)_{ki}\right) + (\delta(\Sigma^2)_{ii} + 2M_{ii}(X_t)_{ii})dt \\
&= 2\Sigma_{ii}\left(\sum_{k=1}^{d} \sqrt{X_t}_{ki}(dB_t)_{ki}\right) + (\delta(\Sigma^2)_{ii} + 2M_{ii}(X_t)_{ii})dt.
\end{aligned} \tag{3.22}
$$

Define a vector of d independent Brownian motions $(Z_t) = (Z_{t,1}, Z_{t,2}, \dots, Z_{t,d})$ by

$$
(dZ_t)_i = dZ_{t,i} = \sqrt{(X_t)_{ii}}^{-1} \sum_{k=1}^{d} (\sqrt{X_t}_{ki}(dB_t)_{ki}).
$$

To verify that (Z_t) is a vector of d independent Brownian motions, we show by applying Levy's characterization theorem that its quadratic covariation $\langle (dZ_t)_i, (dZ_t)_j\rangle = \mathbf{1}_{(i=j)}dt$:

$$
\begin{aligned}
\langle (dZ_t)_i, (dZ_t)_j\rangle &= \langle \sqrt{(X_t)_{ii}}^{-1} \sum_{k=1}^{d} (\sqrt{X_t}_{ki}(dB_t)_{ki}), \sqrt{(X_t)_{jj}}^{-1} \sum_{m=1}^{d} (\sqrt{X_t}_{mj}(dB_t)_{mj})\rangle \\
&= \sqrt{(X_t)_{ii}}^{-1}\sqrt{(X_t)_{jj}}^{-1} \sum_{k=1}^{d}\sum_{m=1}^{d} \sqrt{X_t}_{ki}\sqrt{X_t}_{mj} \underbrace{\langle (dB_t)_{ki}), (dB_t)_{mj})\rangle}_{=\mathbf{1}_{(k,i)=(m,j)}dt} \\
&= \sqrt{(X_t)_{ii}}^{-1}\sqrt{(X_t)_{ii}}^{-1} \sum_{k=1}^{d} \sqrt{X_t}_{ki}\sqrt{X_t}_{ki}\mathbf{1}_{(i=j)}dt \\
&= \sqrt{(X_t)_{ii}}^{-2} \sum_{k=1}^{d} \sqrt{X_t}_{ik}\sqrt{X_t}_{kj}\mathbf{1}_{(i=j)}dt \\
&= ((X_t)_{ii})^{-1}(\sqrt{X_t}\sqrt{X_t})_{ij}\mathbf{1}_{(i=j)}dt \\
&= ((X_t)_{ii})^{-1}(X_t)_{ij}\mathbf{1}_{(i=j)}dt \\
&= \mathbf{1}_{(i=j)}dt.
\end{aligned}
$$

Hence (Z_t) denotes in fact a vector of independent Brownian motions.

Then note that

$$(dZ_t)_i = \sqrt{(X_t)_{ii}}^{-1} \sum_{k=1}^{d} (\sqrt{X_t}_{ki}(dB_t)_{ki}) \Leftrightarrow \sqrt{(X_t)_{ii}}(dZ_t)_i = \sum_{k=1}^{d} (\sqrt{X_t}_{ki}(dB_t)_{ki}). \tag{3.23}$$

With this definition of the Brownian motion vector (Z_t) and (3.23), (3.22) can be rewritten as

$$\begin{aligned}(dX_t)_{ii} &= 2\Sigma_{ii}\left(\sum_{k=1}^{d} \sqrt{X_t}_{ki}(dB_t)_{ki}\right) + (\delta(\Sigma^2)_{ii} + 2M_{ii}(X_t)_{ii})dt \\ &= 2\Sigma_{ii}\sqrt{(X_t)_{ii}}(dZ_t)_i + (\delta(\Sigma^2)_{ii} + 2M_{ii}(X_t)_{ii})dt.\end{aligned} \tag{3.24}$$

So, considering the dynamics of the diagonal components of the Wishart process with diagonal matrices M and Σ, we see that each of them can be interpreted as the dynamics of independent Cox-Ingersoll-Ross processes with parameters $\bar{\sigma} = 2\Sigma_{ii}$, $\bar{\kappa} = -2M_{ii}$ and $\theta = -\frac{\delta(\Sigma^2)_{ii}}{2M_{ii}}$.

Hence this important relation is summarized in the following proposition:

Proposition 3.2.12. *In the simple case where the matrices M and Σ are restricted to diagonal matrices in the Wishart process, the diagonal components of the Wishart process become independent Cox-Ingersoll-Ross processes characterized by the following parameters*

$$\bar{\sigma} = 2\Sigma_{ii,} \qquad \bar{\kappa} = -2M_{ii} \quad \textit{and} \quad \theta = -\frac{\delta(\Sigma^2)_{ii}}{2M_{ii}}.$$

Hence the Wishart process is a direct multivariate extension of the well-known CIR process which has been introduced for modeling the variance of the one-dimensional Heston stochastic volatility model; with the concept and setup of a Wishart process it is possible to increase the dimensionality of risk. In the subsequent sections and chapters, one of our aims is to replace the one-dimensional volatility (for example in the one-dimensional Heston stochastic volatility model) by a volatility-covolatility matrix. This step allows us to define so-called multivariate affine stochastic volatility models, for example the multidimensional Heston stochastic volatility model. One of the advantages of these models is that there still exist closed-form solutions.

But before doing so, we will state the characteristic function of the Wishart processes used in this thesis:

3.2.3 The Characteristic Function of Wishart Processes

This section is mainly based on Section 2.2 of the paper "On the existence of non-central Wishart distributions" written by Eberhard Mayerhofer (see [13, p.4-5]).

As already mentioned in Chapter 1 for the case of vector-valued stochastic processes, a process is called affine, if its characteristic function is exponential affine in the state variable. Expressed more mathematically, one has the following definition of multidimensional affine stochastic processes (which may be a repetition for most of the readers):

Definition 3.2.13. A stochastically continuous (Markov) process X on $\mathcal{S}_d^+$ is called *affine*, if its conditional characteristic function can be written as

$$\mathbb{E}[e^{\langle u, X_t \rangle} | X_0 = x] = e^{\phi(t,u) + \langle \psi(t,u), x \rangle} \tag{3.25}$$

with $t \in \mathbb{R}_+$, $x \in \mathcal{S}_d^+$ and $u \in i\mathcal{S}_d$.

From this definition we can draw the following conclusion:

Proposition 3.2.14. *The so-called characteristic exponents ϕ and ψ in the above definition, satisfy a system of generalized Riccati equations*

$$\begin{aligned} \dot{\phi}(t,u) &= F(\psi(t,u)), \quad \phi(0,u) = 0, \\ \dot{\psi}(t,u) &= R(\psi(t,u)), \quad \psi(0,u) = u \end{aligned}$$

with F, R being of a specific Lévy-Khintchine form.

Theorem 3.2.15. *(see [13, Definition 2.3.]) An affine process X is a Wishart process on $\mathcal{S}_d^+$ characterized by its parameters (δ, Σ, M), if its characteristic exponents (ϕ, ψ) satisfy the following Riccati equations*

$$\begin{aligned} \dot{\phi}(t,u) &= \delta \langle \Sigma^\top \Sigma, \psi(t,u) \rangle, \qquad \phi(0,u) = 0, \\ \dot{\psi}(t,u) &= -2\psi(t,u)\Sigma^\top \Sigma \psi(t,u) + \psi(t,u)M + M^\top \psi(t,u), \qquad \psi(0,u) = u. \end{aligned}$$

Before stating the theorem about the characteristic function of Wishart processes studied in my thesis, according to [13, p.5], two functions have to be introduced:

- let w_t^M be the flow of the vector field $Mx + xM^\top$ defined by

$$w^M : \mathbb{R} \times \mathcal{S}_d^+ \to \mathcal{S}_d^+, \quad w_t^M(x) := e^{Mt} x e^{M^\top t}$$

- the corresponding integral $\sigma_t^M : \mathcal{S}_d^+ \to \mathcal{S}_d^+$ for $t \geq 0$ is defined by

$$\sigma^M : \mathbb{R}_+ \times \mathcal{S}_d^+ \to \mathcal{S}_d^+, \quad \sigma_t^M(x) = 2 \int_0^t w_s^M(x) ds.$$

We are now prepared to state the characteristic function of the Wishart process:

Theorem 3.2.16. *(see [13, Proposition 2.5.]) Let X be a Wishart process. Then the charac-*

teristic exponents ϕ, ψ *take the form*

$$\phi(t,u) = \frac{\delta}{2} \log \det(\boldsymbol{I} + u\sigma_t^M(\Sigma^\top \Sigma)),$$
$$\psi(t,u) = e^{M^\top t}(u^{-1} + \sigma_t^M(\Sigma^\top \Sigma))^{-1} e^{Mt}.$$

Consequently, the characteristic function of X *is given by*

$$\mathbb{E}[e^{\langle u, X_t \rangle} | X_0 = x] = \Big(\det(\boldsymbol{I} - \sigma_t^M(\Sigma^\top \Sigma)u)\Big)^{-\frac{\delta}{2}} e^{tr\left(u(\boldsymbol{I} - \sigma_t^M(\Sigma^\top \Sigma)u)^{-1} w_t^M(x)\right)}, \qquad \forall u \in i\mathcal{S}_d. \quad (3.26)$$

Chapter 4

Multidimensional Heston Stochastic Volatility Model

After introducing Wishart processes in the last chapter we are now prepared to introduce the multidimensional Heston stochastic volatility model.
In the setup of a multidimensional Heston stochastic volatility model, the corresponding stochastic covariance process follows a Wishart process. Hence, in other words, the CIR process as variance process in the one-dimensional Heston stochastic volatility model is now replaced by the Wishart process as covariance process in the multidimensional Heston stochastic volatility model.
This stochastic covariance process as a process of $\mathcal{S}_d^+$ should give the possibility to reflect the stylized facts of financial data, like the volatility smile, in the model. In addition to that it should also enable to capture the dependence structure of different assets.

Since the joint conditional characteristic function of the multidimensional Heston stochastic volatility model is exponential affine, the multidimensional Heston stochastic volatility model belongs to the class of affine processes.

But before studying the multidimensional Heston stochastic volatility model, multivariate affine stochastic volatility models will be introduced and defined in general.

4.1 Introduction of Multivariate Affine Stochastic Volatility Models

In this section matrix-valued affine processes on $\mathcal{S}_d^+$ will be introduced and studied; to be more concrete, so-called multivariate affine stochastic volatility models will be defined.

This section is mainly based on "Chapter 5: Multivariate Affine Stochastic Volatility Models" in [4] written by Christa Cuchiero.

Consider a model for a d-dimensional logarithmic price process whose risk-neutral dynamics are given by

$$dY_t = \left(r\mathbb{I} - \frac{1}{2} diag(X_t) \right) dt + \sqrt{X_t} dW_t, \qquad Y_0 = y,$$

where W_t denotes a standard d-dimensional Brownian motion, r the constant interest rate, $\mathbb{I}$ the vector whose entries are all equal to one and $diag(X_t)$ the vector containing the diagonal entries of X.

Then we consider the i-th component of the logarithmic price process, denoted by $Y_{t,i} := (Y_t)_i$ and which can be written as

$$\begin{aligned} dY_{t,i} = (dY_t)_i &= \left((r\mathbb{I})_i - \frac{1}{2} X_{t,ii} \right) dt + (\sqrt{X_t} dW_t)_i \\ &= r - \frac{1}{2} X_{t,ii} dt + \sum_{j=1}^{d} \sqrt{X_t}_{ij} dW_{t,j}. \end{aligned}$$

Then component-by-component, for all $i, j \in \{1, \ldots, d\}$, the quadratic covariation of this process can be calculated as

$$\begin{aligned} d\langle Y_{t,i}, Y_{t,j} \rangle &= \langle \sum_{m=1}^{d} \sqrt{X_t}_{im} dW_{t,m}, \sum_{n=1}^{d} \sqrt{X_t}_{jn} dW_{t,n} \rangle \\ &= \sum_{m=1}^{d} \sum_{n=1}^{d} \sqrt{X_t}_{im} \sqrt{X_t}_{jn} \underbrace{\langle dW_{t,m}, dW_{t,n} \rangle}_{=\mathbf{1}_{(m=n)} dt} \\ &= \sum_{j=1}^{d} \sqrt{X_t}_{im} \underbrace{\sqrt{X_t}_{jm}}_{=\sqrt{X_t}_{mj}} dt \\ &= (\sqrt{X_t} \sqrt{X_t})_{ij} dt \\ &= X_{t,ij} dt. \end{aligned}$$

So we have now shown that $\langle Y_{t,i}, Y_{t,j} \rangle = X_{t,ij}$ and therefore for the stochastic covariation process it holds that

$$\langle Y, Y \rangle = X.$$

In our framework a multivariate affine stochastic volatility model is defined and characterized by its joint characteristic function of the logarithmic price process and its covariation process: (compare [4, p.144-145]):

It is assumed that the d-dimensional asset price process $(S_t)_{t\geq 0}$ is given by

$$S_t = e^{rt\mathbb{I}+Y_t}, \qquad t \geq 0,$$

where r denotes the constant nonnegative interest rate and $(Y_t)_{t\geq 0}$ denotes the d-dimensional discounted logarithmic price process which starts at $Y_0 = y \in \mathbb{R}^d$ a.s. Then, clearly the discounted price process equals $(e^{Y_t})_{t\geq 0}$. Henceforward, without loss of generality and for ease of notation, we assume that the interest rates are 0. As already mentioned, we denote by $(X_t)_{t\geq 0}$ the stochastic covariation process taking its values in $\mathcal{S}_d^+$ and starting at $X_0 = x \in \mathcal{S}_d^+$ a.s. .

According to [4, p.145], in the context of a multivariate affine stochastic volatility model, the joint process $(X_t, Y_t)_{t\geq 0}$ with state space $D := \mathcal{S}_d^+ \times \mathbb{R}^d$ has to fulfill the following two important assumptions:

- $(X_t, Y_t)_{t\geq 0}$ is a stochastically continuous time-homogeneous Markov process on $D := \mathcal{S}_d^+ \times \mathbb{R}^d$
- The characteristic function of (X_t, Y_t) has exponential affine dependence on the initial states (x, y), that is, there exist functions $(t, u, v) \to \Phi(t, u, v)$ and $(t, u, v) \to \Psi(t, u, v)$ such that

$$\mathbb{E}[e^{tr(uX_t)+v^\top Y_t} | X_0 = x, Y_0 = y] = \Phi(t, u, v) e^{tr(\Psi(t,u,v)x)+v^\top y} \tag{4.1}$$

  $\forall (x, y) \in D$ and $\forall (u, v) \in i\mathcal{S}_d \times i\mathbb{R}^d$

Remark 4.1.1. Considering (4.1), we see that in the above definition we have $\Phi(t, u, v)$ instead of $e^{\phi(t,u,v)}$ as defined before. The reason for this is that it is still an open question if $\Phi(t, u, v)$ may also take the value 0. Therefore in literature one can often also find this slightly "different parametrization" of the characteristic function.

At the end of this section we state a theorem about important properties of multivariate affine stochastic volatility models (see [4, p.146,Theorem 5.1.2.])

Theorem 4.1.2. *Let $(\tau, u, v) \in \mathbb{R}_+ \times i\mathcal{S}_d \times i\mathbb{R}^d$ for some $\tau \geq 0$ and suppose that*

$$\mathbb{E}[e^{tr(uX_\tau)+v^\top Y_\tau} | X_0 = 0, Y_0 = 0] \neq 0.$$

Then, for $t, s \geq 0$ such that $t + s = \tau$, we have $(\tau, u, v) \in \mathbb{R}_+ \times i\mathcal{S}_d \times i\mathbb{R}^d$, $(s, \Psi(t, u, v), v) \in \mathbb{R}_+ \times i\mathcal{S}_d \times i\mathbb{R}^d$ and

$$\mathbb{E}[e^{tr(uX_t)+v^\top Y_t} | X_0 = 0, Y_0 = 0] \neq 0 \qquad \textit{and} \qquad \mathbb{E}[e^{tr(\Psi(t,u,v)X_s)+v^\top Y_s} | X_0 = 0, Y_0 = 0] \neq 0.$$

Moreover, the functions Φ *and* Ψ *satisfy the semiflow equations, that is*

$$\begin{aligned}\Phi(t+s,u,v) &= \Phi(t,u,v)\Phi(s,\Psi(t,u,v),v),\\ \Psi(t+s,u,v) &= \Psi(s,\Psi(t,u,v),v),\end{aligned}$$

and the derivatives

$$F(u,v) := \frac{\partial\Phi(t,u,v)}{\partial t}\Big|_{t=0} \qquad \textit{and} \qquad R(u,v) := \frac{\partial\Psi(t,u,v)}{\partial t}\Big|_{t=0}$$

exist and are continuous in (u,v). *Furthermore, for* $t\in[0,\tau)$, Φ *and* Ψ *satisfy the generalized Riccati equations*

$$\begin{aligned}\frac{\partial}{\partial t}\Phi(t,u,v) &= \Phi(t,u,v)F(\Psi(t,u,v),v), \qquad && \Phi(0,u,v)=1,\\ \frac{\partial}{\partial t}\Psi(t,u,v) &= R(\Psi(t,u,v),v), && \Psi(0,u,v)=u.\end{aligned} \tag{4.2}$$

Remark 4.1.3. If we consider $\phi(t,u,v) = \log\Phi(t,u,v)$, the first generalized Riccati equation in (4.2) turns into

$$\begin{aligned}\frac{\partial}{\partial t}\phi(t,u,v) &= \frac{\partial}{\partial t}(\log\Phi(t,u,v)) = \frac{1}{\Phi(t,u,v)}\frac{\partial}{\partial t}\Phi(t,u,v)\\ &= \frac{1}{\Phi(t,u,v)}\Phi(t,u,v)F(\Psi(t,u,v),v) = F(\Psi(t,u,v),v)\end{aligned} \tag{4.3}$$

with $\phi(0,u,v) = \ln\underbrace{\Phi(0,u,v)}_{=1} = 0$.

4.2 Introduction of the Multidimensional Heston Stochastic Volatility Model

In this section the multidimensional Heston stochastic volatility model used in this thesis is introduced.

As the following definition shows, the multidimensional Heston stochastic volatility model is characterized by the dynamics of a d-dimensional logarithmic price process and by the dynamics of the covariance process which follows a Wishart process.

Definition 4.2.1. Let W denote a standard d-dimensional Brownian motion, denote by $\mathbb{I}$ a vector, whose entries are all equal to 1, and by $diag(X)$ a vector containing the diagonal entries of X. In addition to that let B be a standard $d\times d$ Brownian matrix (a $d\times d$- matrix which contains independent one-dimensional Brownian motions as its entries), M a $d\times d$ matrix and Σ an invertible $d\times d$ matrix and let $\delta\geq d+1$. Then the dynamics in the *multidimensional*

Heston stochastic volatility model are given by

$$
\begin{aligned}
dY_t &= \left(-\frac{1}{2}diag(X_t)\right)dt + \sqrt{X_t}dW_t, \qquad Y_0 = y,\\
dX_t &= \left(\delta\Sigma^\top\Sigma + MX_t + X_tM^\top\right)dt + \sqrt{X_t}dB_t\Sigma + \Sigma^\top dB_t^\top\sqrt{X_t},
\end{aligned}
\tag{4.4}
$$

where

$$dW_t = dB_t\rho + dZ_t(1-\rho^\top\rho)$$

with Z being an $\mathbb{R}^d$-valued Brownian motion, independent of B and ρ being an $\mathbb{R}^d$-vector. In addition to that we assume that without loss of generality for the deterministic matix Σ it holds that $\Sigma = \sqrt{\Sigma^\top\Sigma}$ which means that the matrix Σ is symmetric.

Remark 4.2.2. The multidimensional Heston stochastic volatility model defined above is exactly the “Wishart Affine Stochastic Correlation Model” which has been introduced by José Da Fonesca and Claudio Tebaldi in their paper [7] published in 2007 (see Appendix A).
Three years before, Christian Gouriéroux and Razvan Sufana were one of the first mathematicians who have defined and investigated multidimensional stochastic volatility models (see also Appendix A). The main difference between the multidimensional stochastic volatility model introduced by Fonesca and Tebaldi and that introduced by Gouriéroux and Razvan is the correlation between the Brownian motions: While Gouriéroux and Sufana assume that the Brownian motion of the log return is independent of the Brownian motion of the covariance process, there exists a correlation between those in the model studied by Da Fonesca and Tebaldi.

4.3 The Characteristic Function of the Multidimensional Heston Stochastic Volatility Model

This section is based on [4, p.163-164].

In the multidimensional Heston stochastic volatility model the solutions of the Riccati equations

$$
\begin{aligned}
\frac{\partial\Phi(t,u,v)}{\partial t} &= \Phi(t,u,v)tr(\delta\Sigma^\top\Sigma\Psi(t,u,v)),\\
\frac{\partial\Psi(t,u,v)}{\partial t} &= 2\Psi(t,u,v)\Sigma^\top\Sigma\Psi(t,u,v) + \frac{1}{2}vv^\top + \Psi(t,u,v)(\Sigma^\top\rho v^\top + M)\\
&\quad + (v\rho^\top\Sigma + M^\top)\Psi(t,u,v) - \frac{1}{2}\sum_{i=1}^d v_i(e_ie_i^\top),
\end{aligned}
\tag{4.5}
$$

are given by

$$
\begin{aligned}
\Psi(t,u,v) &= (u\Psi_{12}(t,v) + \Psi_{22}(t,v))^{-1}(u\Psi_{11}(t,v) + \Psi_{21}(t,v)),\\
\Phi(t,u,v) &= \exp\left(\int_0^t tr(\delta\Sigma^\top\Sigma\Psi(s,u,v))ds\right),
\end{aligned}
\tag{4.6}
$$

where

$$
\begin{pmatrix} \Psi_{11}(t,v) & \Psi_{12}(t,v) \\ \Psi_{21}(t,v) & \Psi_{22}(t,v) \end{pmatrix} = \exp\left(t\begin{pmatrix} \Sigma^\top\rho v^\top + M & -2\Sigma^\top\Sigma \\ \frac{1}{2}(vv^\top - \sum_{i=1}^d v_i(e_ie_i^\top)) & -(v\rho^\top\Sigma + M^\top) \end{pmatrix}\right).
$$

After having stated the solutions of the Riccati equations in the multidimensional Heston stochastic volatility model, we are now able to mention the theorem about the characteristic function in this model:

Theorem 4.3.1. *In the multidimensional Heston stochastic volatility model, the characteristic function of* (X_t, Y_t) *can be written as*

$$
\begin{aligned}
\mathbb{E}[e^{\langle u,X_t\rangle + v^\top Y_t}|X_0 = x, Y_0 = y] &= \mathbb{E}[e^{tr(uX_t)+v^\top Y_t}|X_0 = x, Y_0 = y]\\
&= \Phi(t,u,v)e^{tr(\Psi(t,u,v)x)+v^\top y},
\end{aligned}
\tag{4.7}
$$

$\forall(x,y) \in D$ *and* $\forall(t,u,v) \in \mathbb{R}_+ \times i\mathcal{S}_d \times i\mathbb{R}^d$*, where the explicit solutions of the Riccati equations are given as in* (4.6)*.*

Remark 4.3.2. If we set

$$
v = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}
$$

in (4.7), we obtain the characteristic function of the Wishart process as stated in (3.26).

Remark 4.3.3. Comparing (4.2) with (4.5), it follows that in the multidimensional Heston stochastic volatility model

$$
R(u,v) := 2u\Sigma^\top\Sigma u + \frac{1}{2}vv^\top + u(\Sigma^\top\rho v^\top + M) + (v\rho^\top\Sigma + M^\top)u - \frac{1}{2}\sum_{i=1}^d v_i(e_ie_i^\top)
$$

and

$$
F(u,v) := tr(\delta\Sigma^\top\Sigma u).
$$

These two functions will play an important role in Chapter 6 of this thesis.

Chapter 5

Relationships and Results concerning the Multidimensional Heston Stochastic Volatility Model

In this chapter we will focus on interesting properties of the multidimensional Heston stochastic volatility model. A relation between the multidimensional Heston stochastic volatility model and the one-dimensional Heston stochastic volatility model is also derived.

5.1 The Multidimensional Heston Stochastic Volatility Model Considered Componentwise

Let us recall the notion of the multidimensional Heston stochastic volatility model introduced in the last chapter:

In this chapter we assume again that the interest rate equals 0. Then the dynamics in the multidimensional Heston stochastic volatility model are given by

$$\begin{aligned} dY_t &= -\frac{1}{2} diag(X_t)dt + \sqrt{X_t} dW_t, \qquad Y_0 = y, \\ dX_t &= \left(\delta \Sigma^\top \Sigma + M X_t + X_t M^\top\right) dt + \sqrt{X_t} dB_t \Sigma + \Sigma^\top dB_t^\top \sqrt{X_t}, \end{aligned} \tag{5.1}$$

where

$$dW_t = dB_t \rho + dZ_t (1 - \rho^\top \rho) \tag{5.2}$$

with Z being an $\mathbb{R}^d$-valued Brownian motion, independent of B, and ρ being an $\mathbb{R}^d$-vector. Without loss of generality we assume again that $\Sigma = \sqrt{\Sigma^\top \Sigma}$.

Before stating the first proposition of this section, two random variables B_t^k, W_t^k are defined in the following way:
Let e_k denote the k-th unit vector. In addition to that let $\delta \geq d+1$ and Σ be invertible such that Σ^{-1} and X^{-1} are well-defined.
Then define B_t^k by

$$\begin{aligned} B_t^k &:= \int_0^t \sqrt{X_{s,kk}}^{-1} e_k^\top \sqrt{X_s} dB_s \sqrt{\Sigma^\top \Sigma}\ e_k \ \sqrt{(\Sigma^\top \Sigma)_{kk}}^{-1} \\ &= \int_0^t \sqrt{X_{s,kk}}^{-1} (\sqrt{X_s} dB_s \sqrt{\Sigma^\top \Sigma})_{kk} \ \sqrt{(\Sigma^\top \Sigma)_{kk}}^{-1} \end{aligned}$$

and W_t^k by

$$\begin{aligned} W_t^k &:= \int_0^t \sqrt{e_k^\top X_s e_k}^{-1} e_k^\top \sqrt{X_s} dW_s \\ &= \int_0^t \sqrt{X_{s,kk}}^{-1} (\sqrt{X_s} dW_s)_k. \end{aligned}$$

Proposition 5.1.1. *Let $\delta \geq d+1$ and Σ be an invertible matrix.*
Then for all $k, l \in \{1, \ldots, d\}$ the quadratic covariation of B_t^k and B_t^l is given by

$$\langle dB_t^k, dB_t^l \rangle = \sqrt{X_{t,kk} X_{t,ll}}^{-1} \sqrt{(\Sigma^2)_{kk} (\Sigma^2)_{ll}}^{-1} X_{t,kl} (\Sigma^2)_{kl} dt. \tag{5.3}$$

For the special case where Σ is assumed to be of diagonal form, B_t^k and B_t^l are uncorrelated. Moreover, the quadratic covariation of W_t^k and W_t^l is obtained by

$$\langle dW_t^k, dW_t^l \rangle = \sqrt{X_{t,kk} X_{t,ll}}^{-1} X_{t,kl} dt. \tag{5.4}$$

Remark 5.1.2. We assume that $\delta \geq d+1$, because then a unique strong solution of the Wishart process is guaranteed (see Corollary 3.2.4), which does not hit the boundary (see Theorem 3.2.3).

Proof. In differential notation, we have

$$\begin{aligned} dB_t^k &= \sqrt{X_{t,kk}}^{-1} (\sqrt{X_t} dB_t \sqrt{\Sigma^\top \Sigma})_{kk} \sqrt{(\Sigma^\top \Sigma)_{kk}}^{-1} \\ &= \sqrt{X_{t,kk}}^{-1} \left(\sum_{i=1}^d \sum_{m=1}^d \sqrt{X_t}_{ki} dB_{t,im} \sqrt{\Sigma^\top \Sigma}_{mk} \right) \sqrt{(\Sigma^\top \Sigma)_{kk}}^{-1} \end{aligned}$$

and

$$\begin{aligned} dW_t^k &= \sqrt{X_{t,kk}}^{-1} (\sqrt{X_t} dW_t)_k \\ &= \sqrt{X_{t,kk}}^{-1} \sum_{i=1}^d \sqrt{X_t}_{ki} dW_{t,i}. \end{aligned}$$

At first, the quadratic covariation of B_t^k and B_t^l is studied

$$
\begin{aligned}
\langle dB_t^k, dB_t^l \rangle = \Big\langle & \sqrt{X_{t,kk}}^{-1} \left(\sum_{i=1}^d \sum_{m=1}^d \sqrt{X_t}_{ki} dB_{t,im} \sqrt{\Sigma^\top \Sigma}_{mk} \right) \sqrt{(\Sigma^\top \Sigma)_{kk}}^{-1}, \\
& \sqrt{X_{t,ll}}^{-1} \left(\sum_{j=1}^d \sum_{q=1}^d \sqrt{X_t}_{lj} dB_{t,jq} \sqrt{\Sigma^\top \Sigma}_{ql} \right) \sqrt{(\Sigma^\top \Sigma)_{ll}}^{-1} \Big\rangle \\
= & \sqrt{X_{t,kk}}^{-1} \sqrt{X_{t,ll}}^{-1} \sqrt{(\Sigma^\top \Sigma)_{kk}}^{-1} \sqrt{(\Sigma^\top \Sigma)_{ll}}^{-1} \\
& \sum_{i=1}^d \sum_{m=1}^d \sum_{j=1}^d \sum_{q=1}^d \sqrt{X_t}_{ki} \sqrt{X_t}_{lj} \sqrt{\Sigma^\top \Sigma}_{mk} \sqrt{\Sigma^\top \Sigma}_{ql} \underbrace{\langle dB_{t,im}, dB_{t,jq} \rangle}_{= \mathbf{1}_{(i,m)=(j,q)} dt} \\
= & \sqrt{X_{t,kk} X_{t,ll}}^{-1} \sqrt{(\Sigma^\top \Sigma)_{kk} (\Sigma^\top \Sigma)_{ll}}^{-1} \sum_{i=1}^d \sum_{m=1}^d \sqrt{X_t}_{ki} \sqrt{X_t}_{li} \sqrt{\Sigma^\top \Sigma}_{mk} \sqrt{\Sigma^\top \Sigma}_{ml} dt \\
= & \sqrt{X_{t,kk} X_{t,ll}}^{-1} \sqrt{(\Sigma^\top \Sigma)_{kk} (\Sigma^\top \Sigma)_{ll}}^{-1} \sum_{i=1}^d \left(\sqrt{X_t}_{ki} \underbrace{\sqrt{X_t}_{li}}_{= \sqrt{X_t}_{il}} \right) \sum_{m=1}^d \left(\underbrace{\sqrt{\Sigma^\top \Sigma}_{mk}}_{= \sqrt{\Sigma^\top \Sigma}_{km}} \sqrt{\Sigma^\top \Sigma}_{ml} \right) dt \\
= & \sqrt{X_{t,kk} X_{t,ll}}^{-1} \sqrt{(\Sigma^\top \Sigma)_{kk} (\Sigma^\top \Sigma)_{ll}}^{-1} (\sqrt{X_t} \sqrt{X_t})_{kl} (\sqrt{\Sigma^\top \Sigma} \sqrt{\Sigma^\top \Sigma})_{kl} dt \\
= & \sqrt{X_{t,kk} X_{t,ll}}^{-1} \sqrt{(\Sigma^\top \Sigma)_{kk} (\Sigma^\top \Sigma)_{ll}}^{-1} (X_t)_{kl} (\Sigma^\top \Sigma)_{kl} dt \\
= & \sqrt{X_{t,kk} X_{t,ll}}^{-1} \sqrt{(\Sigma^2)_{kk} (\Sigma^2)_{ll}}^{-1} X_{t,kl} (\Sigma^2)_{kl} dt.
\end{aligned}
\tag{5.5}
$$

If Σ is restricted to be a diagonal matrix, (5.5) simplifies to

$$
\langle dB_t^k, dB_t^l \rangle = \sqrt{X_{t,kk} X_{t,ll}}^{-1} \sqrt{(\Sigma^2)_{kk} (\Sigma^2)_{ll}}^{-1} X_{t,kl} \underbrace{(\Sigma^2)_{kl}}_{=0} dt = 0 \qquad \forall k \neq l,
$$

because Σ^2 remains a diagonal matrix and therefore $(\Sigma^2)_{kl} = 0$.

Then the quadratic covariation of W_t^k and W_t^l for all $k, l \in \{1, \dots, d\}$ is examined

$$\begin{aligned}
\langle dW_t^k, dW_t^l \rangle &= \left\langle \sqrt{X_{t,kk}}^{-1} \sum_{i=1}^{d} \sqrt{X_t}_{ki} dW_{t,i}, \sqrt{X_{t,ll}}^{-1} \sum_{j=1}^{d} \sqrt{X_t}_{lj} dW_{t,j} \right\rangle \\
&= \sqrt{X_{t,kk}}^{-1} \sqrt{X_{t,ll}}^{-1} \sum_{i=1}^{d} \sum_{j=1}^{d} \sqrt{X_t}_{ki} \sqrt{X_t}_{lj} \underbrace{\langle dW_{t,i}, dW_{t,j} \rangle}_{= \mathbf{1}_{(i=j)} dt} \\
&= \sqrt{X_{t,kk} X_{t,ll}}^{-1} \sum_{i=1}^{d} \sqrt{X_t}_{ki} \underbrace{\sqrt{X_t}_{li}}_{= \sqrt{X_t}_{il}} dt \\
&= \sqrt{X_{t,kk} X_{t,ll}}^{-1} (\sqrt{X_t} \sqrt{X_t})_{kl} dt \\
&= \sqrt{X_{t,kk} X_{t,ll}}^{-1} X_{t,kl} dt.
\end{aligned}$$

□

Corollary 5.1.3. *The random variables B_t^k and W_t^k are both one-dimensional Brownian motions if $\delta \geq d + 1$ and Σ is invertible.*

Proof. The property of the two random variables B_t^k and W_t^k being both one-dimensional Brownian motions can be concluded from equation (5.3) and (5.4):

If we can show that the quadratic variation $\langle B_t^k \rangle = t$, by Levy's characterization theorem it follows that B_t^k is a one-dimensional Brownian motion (the same holds for W_t^k).

Hence, to get the quadratic variation of B_t^k, we set $k = l$ in (5.3):

$$\begin{aligned}
\langle dB_t^k, dB_t^k \rangle = \langle dB_t^k \rangle &= \sqrt{X_{t,kk} X_{t,kk}}^{-1} \sqrt{(\Sigma^2)_{kk} (\Sigma^2)_{kk}}^{-1} X_{t,kk} (\Sigma^2)_{kk} dt \\
&= (X_{t,kk})^{-1} ((\Sigma^2)_{kk})^{-1} X_{t,kk} (\Sigma^2)_{kk} dt \\
&= dt.
\end{aligned}$$

Therefore we have now shown that B_t^k is a one-dimensional Brownian motion.

Then the quadratic covariation of W_t^k is examined: Again setting $k = l$ in (5.4) yields

$$\begin{aligned}
\langle dW_t^k, dW_t^k \rangle = \langle dW_t^k \rangle &= \sqrt{X_{t,kk} X_{t,kk}}^{-1} X_{t,kk} dt \\
&= (X_{t,kk})^{-1} X_{t,kk} dt \\
&= dt.
\end{aligned}$$

Since $\langle W_t^k \rangle = t$, it follows by the same step that also W_t^k is a one-dimensional Brownian motion. □

As a next step we are now prepared to consider the multidimensional Heston stochastic volatility model component-wise:

Theorem 5.1.4. *Let $\delta \geq d+1$ and Σ be invertible. Then componentwise, the dynamics in the multidimensional Heston stochastic volatility model can be written as*

$$\begin{aligned} dX_{t,kk} := (dX_t)_{kk} &= \left(\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk}\right) dt \\ &\quad + \sqrt{X_{t,kk}}dB_t^k\sqrt{(\Sigma^\top\Sigma)_{kk}} + \sqrt{(\Sigma^\top\Sigma)_{kk}}dB_t^k\sqrt{X_{t,kk}} \\ &= \left(\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk}\right) dt + 2\sqrt{(\Sigma^\top\Sigma)_{kk}}\sqrt{X_{t,kk}}dB_t^k \end{aligned} \tag{5.6}$$

and

$$dY_{t,k} := (dY_t)_k = -\frac{1}{2}X_{t,kk}dt + \sqrt{X_{t,kk}}dW_t^k \tag{5.7}$$

with

$$\begin{aligned} B_t^k : &= \int_0^t \sqrt{X_{s,kk}}^{-1} e_k^\top \sqrt{X_s}dB_s\sqrt{\Sigma^\top\Sigma}\ e_k\ \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} \\ &= \int_0^t \sqrt{X_{s,kk}}^{-1}\left(\sqrt{X_s}dB_s\sqrt{\Sigma^\top\Sigma}\right)_{kk}\ \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} \end{aligned}$$

and

$$W_t^k := \int_0^t \sqrt{e_k^\top X_s e_k}^{-1} e_k^\top\sqrt{X_s}dW_s = \int_0^t \sqrt{X_{s,kk}}^{-1}(\sqrt{X_s}dW_s)_k$$

being both one-dimensional Brownian motions.

Proof. Clearly, to get the (k,k)-th component of the Wishart matrix X_t, one has to multiply it with two unit vectors, once from the left and once from the right:

$$\begin{aligned} (dX_t)_{kk} &= e_k^\top dX_t e_k \\ &= e_k^\top\left(\delta\Sigma^\top\Sigma + MX_t + X_tM^\top\right)e_kdt + e_k^\top\sqrt{X_t}dB_t\Sigma e_k + e_k^\top\Sigma^\top dB_t^\top\sqrt{X_t}e_k \\ &= (\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk})dt + (\sqrt{X_t}dB_t\Sigma)_{kk} + (\Sigma^\top dB_t^\top\sqrt{X_t})_{kk} \\ &= (\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk})dt + (\sqrt{X_t}dB_t\Sigma)_{kk} + (\sqrt{X_t}dB_t\Sigma)_{kk} \\ &= (\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk})dt + 2(\sqrt{X_t}dB_t\Sigma)_{kk} \end{aligned} \tag{5.8}$$

The penultimate step follows since

$$(\Sigma^\top dB_t^\top\sqrt{X_t})_{kk} = (\Sigma^\top dB_t^\top\sqrt{X_t}^\top)_{kk} = ((\sqrt{X_t}dB_t\Sigma)^\top)_{kk} = (\sqrt{X_t}dB_t\Sigma)_{kk}.$$

Plugging the Brownian motion B_t^k (B_t^k is given by $dB_t^k = \sqrt{X_{t,kk}}^{-1}(\sqrt{X_t}dB_t\sqrt{\Sigma^\top\Sigma})_{kk}\sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1}$) into (5.6), one obtains

$$
\begin{aligned}
dX_{t,kk} &= \Big(\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk}\Big)\,dt + \sqrt{X_{t,kk}}dB_t^k\sqrt{(\Sigma^\top\Sigma)_{kk}} + \sqrt{(\Sigma^\top\Sigma)_{kk}}dB_t^k\sqrt{X_{t,kk}} \\
&= \Big(\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk}\Big)\,dt \\
&\quad + \sqrt{X_{t,kk}}\sqrt{X_{t,kk}}^{-1}(\sqrt{X_t}dB_t\sqrt{\Sigma^\top\Sigma})_{kk}\sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1}\sqrt{(\Sigma^\top\Sigma)_{kk}} \\
&\quad + \sqrt{(\Sigma^\top\Sigma)_{kk}}\sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1}(\sqrt{X_t}dB_t\sqrt{\Sigma^\top\Sigma})_{kk}\sqrt{X_{t,kk}}^{-1}\sqrt{X_{kk}} \\
&= \Big(\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk}\Big)\,dt + (\sqrt{X_t}dB_t\underbrace{\sqrt{\Sigma^\top\Sigma}}_{=\Sigma})_{kk} + (\sqrt{X_t}dB_t\underbrace{\sqrt{\Sigma^\top\Sigma}}_{=\Sigma})_{kk} \\
&= \Big(\delta(\Sigma^\top\Sigma)_{kk} + (MX_t)_{kk} + (X_tM^\top)_{kk}\Big)\,dt + 2(\sqrt{X_t}dB_t\Sigma)_{kk}.
\end{aligned}
\tag{5.9}
$$

The first equation (5.6) of the theorem is now verified since (5.8) and (5.9) are equivalent.

Similarly as before, to get the k-th component of Y_t, one has to multiply it with the appropriate unit vector

$$
\begin{aligned}
(dY_t)_k &= e_k^\top dY_t = e_k^\top(-\frac{1}{2}diag(X_t)dt) + e_k^\top(\sqrt{X_t}dW_t) \\
&= -\frac{1}{2}(diag(X_t))_k dt + (\sqrt{X_t}dW_t)_k \\
&= -\frac{1}{2}X_{t,kk}dt + (\sqrt{X_t}dW_t)_k.
\end{aligned}
\tag{5.10}
$$

Plugging in the Brownian motion W_t^k (dW_t^k is given by $dW_t^k = \sqrt{X_{t,kk}}^{-1}(\sqrt{X_t}dW_t)_k$) into (5.7), one gets

$$
\begin{aligned}
dY_{t,k} &= -\frac{1}{2}X_{t,kk}dt + \sqrt{X_{t,kk}}dW_t^k \\
&= -\frac{1}{2}X_{t,kk}dt + \sqrt{X_{t,kk}}\sqrt{X_{t,kk}}^{-1}(\sqrt{X_t}dW_t)_k \\
&= -\frac{1}{2}X_{t,kk}dt + (\sqrt{X_t}dW_t)_k.
\end{aligned}
\tag{5.11}
$$

Since (5.10) and (5.11) are equivalent, we have now also proved the second equation (5.7). □

We can now proceed by stating the following definition:

Definition 5.1.5. We call the one-dimensional model determined by the dynamics in (5.6) and (5.7) the k-th marginals of the multidimensional Heston stochastic volatility model. In other words, each of the diagonal element of the Wishart process $X_{t,kk}$ in combination with the corresponding component of the logarithmic price process $Y_{t,k}$ represents the k-th marginals $(X_{t,kk}, Y_{t,k})$ of the multidimensional Heston stochastic volatility model.

As a next step the correlation between $X_{t,kk}$ and $Y_{t,l}$ is examined:
The result will be that the l-th component of the d-dimensional logarithmic price process and the (k,k)-th diagonal element of the covariance process is independent of any component of the logarithmic price process; it only depends on the stochastic covariance process which is in our case given by the Wishart process.
At this point we should mention that this property is valid for affine models in general.

But before stating this important theorem, as an intermediate step the correlation between B_t^k and W_t^l is considered:

Proposition 5.1.6. *The correlation between the Brownian motion B_t^k and the Brownian motion W_t^l is determined by*

$$\langle dB_t^k, dW_t^l \rangle = \sqrt{X_{t,kk}X_{t,ll}}^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} X_{t,kl}(\Sigma\rho)_k dt. \tag{5.12}$$

Especially, the correlation between the Brownian motion B_t^k and the Brownian motion W_t^k does not depend on $X_{t,kl}$ and is given by

$$\langle dB_t^k, dW_t^k \rangle = \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} (\Sigma\rho)_k dt. \tag{5.13}$$

Proof. The quadratic covariation between B_t^k and W_t^l can be calculated as

$$\begin{aligned}
\langle dB_t^k, dW_t^l \rangle &= \left\langle \sqrt{X_{t,kk}}^{-1} (\sqrt{X_t} dB_t \sqrt{\Sigma^\top\Sigma})_{kk} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1}, \sqrt{X_{t,ll}}^{-1} (\sqrt{X_t} dW_t)_l \right\rangle \\
&= \left\langle \sqrt{X_{t,kk}}^{-1} \sum_{i=1}^d \sum_{m=1}^d (\sqrt{X_t}_{ki} (dB_t)_{im} \sqrt{\Sigma^\top\Sigma}_{mk}) \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1}, \sqrt{X_{t,ll}}^{-1} \sum_{q=1}^d (\sqrt{X_t}_{lq} (dW_t)_q) \right\rangle \\
&= \sqrt{X_{t,kk}}^{-1} \sqrt{X_{t,ll}}^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} \sum_{i=1}^d \sum_{m=1}^d \sum_{q=1}^d \sqrt{X_t}_{ki} \sqrt{\Sigma^\top\Sigma}_{mk} \sqrt{X_t}_{lq} \langle (dB_t)_{im}, (dW_t)_q \rangle.
\end{aligned} \tag{5.14}$$

Now as an auxiliary calculation we focus on the last term of the above equation.
In the third step of the subsequent calculations one has to remember that, per definition, Z_t is independent of B_t:

$$
\begin{aligned}
\langle (dB_t)_{im}, (dW_t)_q \rangle &= \langle (dB_t)_{im}, (dB_t\rho)_q + (dZ_t(1-\rho^\top\rho))_q \rangle \\
&= \langle (dB_t)_{im}, (dB_t\rho)_q \rangle + \underbrace{\langle (dB_t)_{im}, (dZ_t(1-\rho^\top\rho))_q \rangle}_{=0} \\
&= \langle (dB_t)_{im}, (dB_t\rho)_q \rangle \\
&= \langle (dB_t)_{im}, \sum_{j=1}^{d} (dB_t)_{qj}\rho_j \rangle \\
&= \rho_j \sum_{j=1}^{d} \underbrace{\langle (dB_t)_{im}, (dB_t)_{qj} \rangle}_{=\mathbf{1}_{(i,m)=(q,j)dt}}.
\end{aligned}
$$

Therefore (5.14) can be simplified to

$$
\begin{aligned}
\langle dB_t^k, dW_t^l \rangle &= \sqrt{X_{t,kk}}^{-1} \sqrt{X_{t,ll}}^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} \\
&\qquad \sum_{i=1}^{d}\sum_{m=1}^{d}\sum_{q=1}^{d} \sqrt{X_t}_{ki} \sqrt{\Sigma^\top\Sigma}_{mk} \sqrt{X_t}_{lq} \rho_j \sum_{j=1}^{d} \underbrace{\langle (dB_t)_{im}, (dB_t)_{qj} \rangle}_{=\mathbf{1}_{(i,m)=(q,j)}dt} \\
&= \sqrt{X_{t,kk}X_{t,ll}}^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} \sum_{i=1}^{d}\sum_{m=1}^{d} \sqrt{X_t}_{ki} \sqrt{\Sigma^\top\Sigma}_{mk} \sqrt{X_t}_{li} \rho_m dt \\
&= \sqrt{X_{t,kk}X_{t,ll}}^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} \sum_{i=1}^{d} (\sqrt{X_t}_{ki} \underbrace{\sqrt{X_t}_{li}}_{=\sqrt{X_t}_{il}}) \sum_{m=1}^{d} (\underbrace{\sqrt{\Sigma^\top\Sigma}_{mk}}_{=\sqrt{\Sigma^\top\Sigma}_{km}} \rho_m) dt \\
&= \sqrt{X_{t,kk}X_{t,ll}}^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} (\sqrt{X_t}\sqrt{X_t})_{kl} (\sqrt{\Sigma^\top\Sigma}\rho)_k dt \\
&= \sqrt{X_{t,kk}X_{t,ll}}^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} X_{t,kl} (\Sigma\rho)_k dt.
\end{aligned}
$$

For the special case where the quadratic covariation of B_t^k and W_t^k is considered, we obtain by setting $k = l$:

$$
\begin{aligned}
\langle dB_t^k, dW_t^k \rangle &= \sqrt{X_{t,kk}X_{t,kk}}^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} X_{t,kk} (\Sigma\rho)_k dt \\
&= (X_{t,kk})^{-1} \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} X_{t,kk} (\Sigma\rho)_k dt \\
&= \sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1} (\Sigma\rho)_k dt.
\end{aligned}
$$

□

After formulating the last proposition we are now prepared to state the following important theorem:

Theorem 5.1.7. *The quadratic covariation between $X_{t,kk}$ and $Y_{t,l}$ is given by*

$$\langle dX_{t,kk}, dY_{t,l}\rangle = 2X_{t,kl}(\Sigma\rho)_k dt,$$

which means that it does not depend on the logarithmic price process $Y_{t,l}$, it only depends on the stochastic covariance process $X_{t,kk}$. In particular, for $k = l$:

$$\langle dX_{t,kk}, dY_{t,k}\rangle = 2X_{t,kk}(\Sigma\rho)_k dt. \tag{5.15}$$

Proof. Considering that the "dt-terms" (deterministic terms) vanish under the quadratic covariation, the correlation is given by

$$\begin{aligned}
\langle dX_{t,kk}, dY_{t,l}\rangle &= \left\langle 2\sqrt{X_{t,kk}}\sqrt{(\Sigma^\top\Sigma)_{kk}}dB_t^k, \sqrt{X_{t,ll}}dW_t^l\right\rangle \\
&= 2\sqrt{X_{t,kk}}\sqrt{X_{t,ll}}\sqrt{(\Sigma^\top\Sigma)_{kk}}\langle dB_t^k, dW_t^l\rangle \\
&\overset{(5.12)}{=} 2\sqrt{X_{t,kk}X_{t,ll}}\sqrt{(\Sigma^\top\Sigma)_{kk}}\sqrt{X_{t,kk}X_{t,ll}}^{-1}\sqrt{(\Sigma^\top\Sigma)_{kk}}^{-1}X_{t,kl}(\Sigma\rho)_k dt \\
&= 2X_{t,kl}(\Sigma\rho)_k dt.
\end{aligned} \tag{5.16}$$

Especially, setting $k = l$ in (5.16), the quadratic covariation between $X_{t,kk}$ and $Y_{t,k}$ can be determined by

$$\langle dX_{t,kk}, dY_{t,k}\rangle = 2X_{t,kk}(\Sigma\rho)_k dt.$$

□

5.2 The One-Dimensional Heston Stochastic Volatility Model as the Marginals of the Multidimensional Heston Stochastic Volatility Model

In this section we choose the matrix M appropriately in the dynamics of the marginals of the multidimensional Heston stochastic volatility model given by (5.6) and (5.7) such that we get the special case of a one-dimensional Heston stochastic volatility model.

But before doing so, we set the dimension $d = 1$ in (5.1). Then we obtain the one-dimensional Heston stochastic volatility model with a slightly different parametrization as introduced some chapters before: To be more precise, under the assumption that $d = 1$, the dynamics given in

(5.1) simplify to

$$
\begin{aligned}
dY_t &= -\frac{1}{2}X_t dt + \sqrt{X_t} dW_t, \\
dX_t &= (\delta\Sigma^2 + 2MX_t)dt + 2\Sigma\sqrt{X_t} dB_t,
\end{aligned}
\tag{5.17}
$$

where $dW_t = \rho dB_t + (1-\rho^2)dZ_t$.

In addition to that, if $d = 1$, the quadratic covariation of W_t and B_t is determined by

$$
\begin{aligned}
dW_t dB_t &= \langle dW_t, dB_t \rangle \\
&= \langle \rho dB_t + (1-\rho^2)dZ_t, dB_t \rangle \\
&= \langle \rho dB_t, dB_t \rangle + \langle (1-\rho^2)dZ_t, dB_t \rangle \\
&= \rho \langle dB_t, dB_t \rangle + (1-\rho^2) \underbrace{\langle dZ_t, dB_t \rangle}_{=0} \\
&= \rho dt.
\end{aligned}
\tag{5.18}
$$

For notational convenience we set $\sigma = \Sigma$, $\tilde{\rho} = \rho$, $k = \delta\Sigma^2$ and $\kappa = 2M$.

Then (5.17) and (5.18) can be rewritten as

$$
\begin{aligned}
dX_t &= (k + \kappa X_t)dt + 2\sigma\sqrt{X_t} dB_t, \\
dY_t &= -\frac{1}{2}X_t dt + \sqrt{X_t} dW_t, \\
dW_t dB_t &= \tilde{\rho} dt
\end{aligned}
\tag{5.19}
$$

with $k, \sigma \geq 0, \kappa \in \mathbb{R}$.

The one-dimensional model (X, Y) given by (5.19) and determined by its characterizing parameters $(k, \kappa, \sigma, \tilde{\rho})$ is called one-dimensional Heston stochastic volatility model.

Remark 5.2.1. This parametrization of a one-dimensional Heston stochastic volatility model as given in (5.19) will be used in the remaining sections and chapters.

In the one-dimensional Heston stochastic volatility model determined by (5.19), the covariation between the log return Y_t and the variance X_t is given by

$$
d\langle Y_t, X_t \rangle = 2\sigma\sqrt{X_t}\sqrt{X_t}\langle dB_t, dW_t \rangle = 2\sigma X_t \tilde{\rho} dt
\tag{5.20}
$$

and does not depend on the log return Y_t.

Now denote by $\mathbb{O}$ the zero matrix (matrix which contains 0 in each component). If in equation (5.6) the deterministic matrix M is chosen to be the zero matrix ($M = \mathbb{O}$), we are in the case of the one-dimensional Heston stochastic volatility model, which is the idea of the following theorem:

Theorem 5.2.2. *Let $\delta \geq d+1$ and Σ be invertible. Setting the deterministic matrix $M = \mathbb{O}$ in (5.6), the marginals (X_{ii}, Y_i) of the multidimensional Heston stochastic volatility model simplify to the well-known one-dimensional Heston stochastic volatility model. Hence each component (in the diagonal) corresponds to a one-dimensional Heston stochastic volatility model with the following characterizing parameters*

$$k_i = \delta(\Sigma^\top\Sigma)_{ii}, \quad \kappa_i = 0, \quad \sigma_i = \sqrt{(\Sigma^\top\Sigma)_{ii}} \text{ and } \tilde{\rho}_i = \frac{(\Sigma\rho)_i}{\sqrt{(\Sigma^\top\Sigma)_{ii}}}.$$

Proof. If M equals the zero matrix, the dynamics of the (i,i)-th component of the multidimensional Heston stochastic volatility model (5.6) can be simplified to

$$\begin{aligned} dX_{t,ii} &= \left(\delta(\Sigma^\top\Sigma)_{ii} + (\mathbb{O}X_t)_{ii} + (X_t\mathbb{O}^\top)_{ii}\right) dt + \sqrt{X_{t,ii}}dB_t^i\sqrt{(\Sigma^\top\Sigma)_{ii}} + \sqrt{(\Sigma^\top\Sigma)_{ii}}dB_t^i\sqrt{X_{t,ii}} \\ &= \left(\delta(\Sigma^\top\Sigma)_{ii}\right) dt + 2\sqrt{(\Sigma^\top\Sigma)_{ii}}\sqrt{X_{t,ii}}dB_t^i. \end{aligned} \tag{5.21}$$

Comparing (5.19) with (5.21), the parameters of the one-dimensional Heston stochastic volatility model can be determined as follows: $k_i = \delta(\Sigma^\top\Sigma)_{ii}$, $\kappa_i = 0$ and $2\sigma_i = 2\sqrt{(\Sigma^\top\Sigma)_{ii}} \Rightarrow$
$\sigma_i = \sqrt{(\Sigma^\top\Sigma)_{ii}}$.

And finally, comparing (5.20) with (5.15), we see that

$$2X_{t,ii}(\Sigma\rho)_i dt = 2\sigma_i\tilde{\rho}_i X_{t,ii}dt \quad \Leftrightarrow \quad \sigma_i\tilde{\rho}_i = (\Sigma\rho)_i,$$

and hence by noting that $\sigma_i = \sqrt{(\Sigma^\top\Sigma)_{ii}}$, it follows that

$$\tilde{\rho}_i = \frac{(\Sigma\rho)_i}{\sigma_i} = \frac{(\Sigma\rho)_i}{\sqrt{(\Sigma^\top\Sigma)_{ii}}}.$$

□

The statement about the last theorem is still true if we consider the more general case where the deterministic matrix M is assumed to be of diagonal form:

Theorem 5.2.3. *Let $\delta \geq d+1$ and Σ be invertible. If the matrix M is restricted to be a diagonal matrix ($M = diag(M_1, M_2, \ldots, M_d) = diag(M_{11}, M_{22}, \ldots, M_{dd})$), the multidimensional Heston stochastic volatility model considered component-by-component still simplifies to the one-dimensional Heston stochastic volatility model with the following characterizing parameters*

$$k_i = \delta(\Sigma^\top\Sigma)_{ii} \text{ and } \kappa_i = 2M_{ii} \text{ and } \sigma_i = \sqrt{(\Sigma^\top\Sigma)_{ii}} \text{ and } \tilde{\rho}_i = \frac{(\Sigma\rho)_i}{\sqrt{(\Sigma^\top\Sigma)_{ii}}}.$$

Proof. Plugging $M = diag(M_{11}, M_{22}, \ldots, M_{dd})$ into the dynamics of the (i,i)-th component of the multidimensional Heston stochastic volatility model (5.6), we get

$$\begin{aligned} dX_{t,ii} &= \left(\delta(\Sigma^\top\Sigma)_{ii} + M_{ii}X_{t,ii} + X_{t,ii}M_{ii}\right) dt + \sqrt{X_{t,ii}}dB_t^i\sqrt{(\Sigma^\top\Sigma)_{ii}} + \sqrt{(\Sigma^\top\Sigma)_{ii}}dB_t^i\sqrt{X_{t,ii}} \\ &= \left(\delta(\Sigma^\top\Sigma)_{ii} + 2M_{ii}X_{t,ii}\right) dt + 2\sqrt{(\Sigma^\top\Sigma)_{ii}}\sqrt{X_{t,ii}}dB_t^i. \end{aligned}$$

As before, the parameters can be determined: $k_i = \delta(\Sigma^\top\Sigma)_{ii}$, $\kappa_i = 2M_{ii}$ and σ_i and $\tilde{\rho}_i$ stay the same as before, which means that $\sigma_i = \sqrt{(\Sigma^\top\Sigma)_{ii}}$ and $\tilde{\rho}_i = \frac{(\Sigma\rho)_i}{\sqrt{(\Sigma^\top\Sigma)_{ii}}}$. □

Remark 5.2.4. This interesting relationship between the multidimensional Heston stochastic volatility model and the one-dimensional Heston stochastic volatility model can also be derived by considering their characteristic functions instead of the dynamics as we have done it before: At first we choose u and v to be of the following form

$$u = \begin{pmatrix} \mathcal{O} & \mathcal{O} & \mathcal{O} \\ \mathcal{O} & u_{ii} & \mathcal{O} \\ \mathcal{O} & \mathcal{O} & \mathcal{O} \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} \mathcal{O} \\ v_i \\ \mathcal{O} \end{pmatrix}$$

(so u denotes the $d \times d$ matrix which has u_{ii} as its (i,i)-th entry and otherwise only 0; v denotes a d-dimensional null vector with the exception that its i-th entry equals v_i). As a next step we plug them into the equation of the characteristic function of the multidimensional Heston stochastic volatility model as defined in (4.7). In other words, we have to consider $E[e^{u_{ii}X_{t,ii}+v_iY_{t,i}}|X_{0,ii} = x_{ii}, Y_{0,i} = y_i]$. In addition to that the deterministic matrix M has to be chosen appropriately as before, which means that in the first case we set $M = \mathbb{O}$ and then in the second less restricted case we set $M = diag(M_{11}, M_{22}, \ldots, M_{dd})$. If we then simplify the right hand side of (4.7) and compare this term with the characteristic function of the one-dimensional Heston stochastic volatility model, the parameters as specified in Theorem 5.2.2 or Theorem 5.2.3 could be determined as well.
So to summarize, there exist two different ways to derive the relationship between the marginals of the multidimensional Heston stochastic volatility model and the one-dimensional Heston stochastic volatility model, either by studying their dynamics or their characteristic functions.

As we have seen in the previous examinations, vanishing a linear drift ($M = \mathbb{O}$) or choosing the matrix M to be only of diagonal form is a possibility to get a one-dimensional Heston stochastic volatility model out of the marginals of the multidimensional Heston stochastic volatility model.

It is a known fact that if one changes the probability measure under Girsanov's theorem, the drift of the stochastic process changes. Hence, by applying Girsanov's theorem, we get the desired drift:

This result is stated in the following theorem proved by Eberhard Mayerhofer in [14, p.17-18]:

Theorem 5.2.5. *(see [14, Theorem 4.1.]) Suppose X_t is a Wishart process with parameters (δ, M, Σ), where Σ is invertible, and let $X_0 = x \in \mathcal{S}_d^+$. For $\delta^{\tilde{\mathbb{Q}}} \in \mathbb{R}$ and a $d \times d$ matrix $M^{\tilde{\mathbb{Q}}}$, we set*

$$\gamma_t := \sqrt{X_t}(M^\top - (M^{\tilde{\mathbb{Q}}})^\top)\Sigma^{-1} + \left(\frac{\delta}{2} - \frac{\delta^{\tilde{\mathbb{Q}}}}{2}\right)\sqrt{X_t}^{-1}\Sigma^\top$$

and

$$Z_t := \mathcal{E}\left(-\int_0^t tr(\gamma_s dB_s)\right),$$

If $\min(\frac{\delta}{2}, \frac{\delta^{\tilde{\mathbb{Q}}}}{2}) \geq \frac{d+1}{2}$, then Z_t is a martingale on $[0,T]$, and $B_t^{\tilde{\mathbb{Q}}} := \gamma_t + B_t$ is a $\tilde{\mathbb{Q}}$-Brownian motion on $[0,T]$. Furthermore X_t is a Wishart process with parameters $(\delta^{\tilde{\mathbb{Q}}}, M^{\tilde{\mathbb{Q}}}, \Sigma)$ under $\tilde{\mathbb{Q}}$.

Remark 5.2.6. Clearly $\tilde{\mathbb{Q}}$ is defined by $d\tilde{\mathbb{Q}} = \mathcal{E}\left(-\int_0^T tr(\gamma_s dB_s)\right) d\mathbb{Q} = Z_T d\mathbb{Q}$.

Hence by applying Theorem 5.2.5, under the equivalent probability measure $\tilde{\mathbb{Q}}$, the dynamics of (3.19) can be written as

$$dX_t = \sqrt{X_t} dB_t^{\tilde{\mathbb{Q}}}\Sigma + \Sigma^\top (dB^{\tilde{\mathbb{Q}}})_t^\top \sqrt{X_t} + (\delta^{\tilde{\mathbb{Q}}}\Sigma^\top\Sigma + M^{\tilde{\mathbb{Q}}}X_t + X_t(M^{\tilde{\mathbb{Q}}})^\top)dt. \tag{5.22}$$

By noting that $\gamma_t^\top = (\Sigma^{-1})^\top(M - M^{\tilde{\mathbb{Q}}})\sqrt{X_t} + \Sigma\sqrt{X_t}^{-1}(\frac{\delta}{2} - \frac{\delta^{\tilde{\mathbb{Q}}}}{2})$, expression (5.22) can be verified because

$$\begin{aligned}
dX_t &= \sqrt{X_t} dB_t^{\tilde{\mathbb{Q}}}\Sigma + \Sigma^\top (dB^{\tilde{\mathbb{Q}}})_t^\top \sqrt{X_t} + (\delta^{\tilde{\mathbb{Q}}}\Sigma^\top\Sigma + M^{\tilde{\mathbb{Q}}}X_t + X_t(M^{\tilde{\mathbb{Q}}})^\top)dt \\
&= \sqrt{X_t}(\gamma_t dt + dB_t)\Sigma + \Sigma^\top(\gamma_t^\top dt + dB_t^\top)\sqrt{X_t} + (\delta^{\tilde{\mathbb{Q}}}\Sigma^\top\Sigma + M^{\tilde{\mathbb{Q}}}X_t + X_t(M^{\tilde{\mathbb{Q}}})^\top)dt \\
&= \sqrt{X_t}\left(\left(\sqrt{X_t}(M^\top - (M^{\tilde{\mathbb{Q}}})^\top)\Sigma^{-1} + \left(\frac{\delta}{2} - \frac{\delta^{\tilde{\mathbb{Q}}}}{2}\right)\sqrt{X_t}^{-1}\Sigma^\top\right)dt + dB_t\right)\Sigma \\
&\quad + \Sigma^\top\left(\left((\Sigma^{-1})^\top(M - M^{\tilde{\mathbb{Q}}})\sqrt{X_t} + \Sigma\sqrt{X_t}^{-1}\left(\frac{\delta}{2} - \frac{\delta^{\tilde{\mathbb{Q}}}}{2}\right)\right)dt + dB_t^\top\right)\sqrt{X_t} \\
&\quad + (\delta^{\tilde{\mathbb{Q}}}\Sigma^\top\Sigma + M^{\tilde{\mathbb{Q}}}X_t + X_t(M^{\tilde{\mathbb{Q}}})^\top)dt \\
&= X_t(M^\top - (M^{\tilde{\mathbb{Q}}})^\top)dt + \left(\frac{\delta}{2} - \frac{\delta^{\tilde{\mathbb{Q}}}}{2}\right)\Sigma^\top\Sigma dt + \sqrt{X_t}dB_t\Sigma \\
&\quad + (M - M^{\tilde{\mathbb{Q}}})X_t dt + \Sigma^\top\Sigma\left(\frac{\delta}{2} - \frac{\delta^{\tilde{\mathbb{Q}}}}{2}\right)dt + \Sigma^T dB_t^\top\sqrt{X_t} + (\delta^{\tilde{\mathbb{Q}}}\Sigma^\top\Sigma + M^{\tilde{\mathbb{Q}}}X_t + X_t(M^{\tilde{\mathbb{Q}}})^\top)dt \\
&= \sqrt{X_t}dB_t\Sigma + \Sigma^\top dB_t^\top\sqrt{X_t} + (\delta\Sigma^\top\Sigma + MX_t + X_tM^\top)dt,
\end{aligned}$$

which equals exactly the Wishart process given in (3.19).

Remark 5.2.7. Clearly, Theorem 5.2.5 is also valid if we restrict $M^{\tilde{\mathbb{Q}}}$ to be only of diagonal form which we denote by $M_{diag}^{\tilde{\mathbb{Q}}}$.

Hence by applying Theorem 5.2.5, we get the desired diagonal form of $M^{\tilde{\mathbb{Q}}}$ (denoted by $M^{\tilde{\mathbb{Q}}}_{diag}$), which yields to a one-dimensional Heston stochastic volatility model.

This important result is stated in the following theorem:

Theorem 5.2.8. *Define γ_t by*

$$\gamma_t := \sqrt{X_t}(M^\top - (M^{\tilde{\mathbb{Q}}}_{diag})^\top)\Sigma^{-1} + \left(\frac{\delta}{2} - \frac{\delta^{\tilde{\mathbb{Q}}}}{2}\right)\sqrt{X_t}^{-1}\Sigma^\top.$$

Considering the dynamics of (3.19) *under the equivalent probability measure $\tilde{\mathbb{Q}}$ given by*

$$d\tilde{\mathbb{Q}} = \mathcal{E}\left(-\int_0^T tr(\gamma_s dB_s)\right) d\mathbb{Q},$$

the marginals $(X_{t,ii}, Y_{t,i})$ of the corresponding multidimensional Heston stochastic volatility model belong to a one-dimensional Heston stochastic volatility model with characterizing parameters

$$k_i^{\tilde{\mathbb{Q}}} = \delta^{\tilde{\mathbb{Q}}}(\Sigma^\top\Sigma)_{ii}, \quad \kappa_i^{\tilde{\mathbb{Q}}} = 2M^{\tilde{\mathbb{Q}}}_{diag,ii}, \quad \sigma_i = \sqrt{(\Sigma^\top\Sigma)_{ii}} \quad \text{and} \quad \tilde{\rho}_i = \frac{(\Sigma\rho)_i}{\sqrt{(\Sigma^\top\Sigma)_{ii}}}.$$

Remark 5.2.9. Clearly, the multidimensional Heston stochastic volatility model under the equivalent probability measure $\tilde{\mathbb{Q}}$ is given by

$$\begin{aligned} dY_t &= \left(-\frac{1}{2}diag(X_t)\right)dt + \sqrt{X_t}dW_t^{\tilde{\mathbb{Q}}}, \qquad Y_0 = y \\ dX_t &= \left(\delta^{\tilde{\mathbb{Q}}}\Sigma^\top\Sigma + M^{\tilde{\mathbb{Q}}}_{diag}X_t + X_t(M^{\tilde{\mathbb{Q}}}_{diag})^\top\right)dt + \sqrt{X_t}dB_t^{\tilde{\mathbb{Q}}}\Sigma + \Sigma^\top(dB^{\tilde{\mathbb{Q}}})_t^\top\sqrt{X_t}, \end{aligned}$$

where

$$dW_t^{\tilde{\mathbb{Q}}} = dB_t^{\tilde{\mathbb{Q}}}\rho + dZ_t^{\tilde{\mathbb{Q}}}(1 - \rho^\top\rho)$$

with $Z^{\tilde{\mathbb{Q}}}$ being an $\mathbb{R}^d$-valued $\tilde{\mathbb{Q}}$-Brownian motion, independent of $B^{\tilde{\mathbb{Q}}}$ and ρ being an $\mathbb{R}^d$-vector.

Remark 5.2.10. If we start under the physical measure with a multidimensional Heston stochastic volatility model, there always exists some equivalent measure change such that we end up in the setting of Theorem 5.2.8 and have risk neutral one-dimensional Heston stochastic volatility models for each marginal.

Chapter 6

Small-Time Asymptotics for Implied Volatility

In this chapter the small-time asymptotics for implied volatility in different stochastic volatility models are studied. At first a one-dimensional Heston stochastic volatility model is considered and then the marginals of the multidimensional Heston stochastic volatility model are regarded.

Martin Forde and Antoine Jacquier have examined the small-time asymptotics for implied volatility under the one-dimensional Heston stochastic volatility model in their paper [8]. In the second section of this chapter we summarize important results concerning this topic. And finally, in the third section of this chapter, we show that these statements and theorems are also valid in a more general framework, namely for the marginals of the multidimensional Heston stochastic volatility model defined in (5.6) and (5.7). In other words, we assert that the results proved in [8] by Martin Forde and Antoine Jacquier are independent of any choice of the deterministic matrix M in the dynamics (5.6) of the multidimensional Heston stochastic volatility model regarded component-wise.
Note that also Elisa Alòs has studied small-time asymptotics of option prices for the Heston stochastic volatility model by using Malliavin calculus (see [1]).

Now in the first section of this chapter we define the implied volatility:

6.1 Short Introduction of the Implied Volatility

It is a well-known fact that a popular possibility to price a European call or put option is to use the famous Black-Scholes formula. So if the Black-Scholes formula is applied for pricing purposes, one has to plug in the interest rate, the strike, the current value of the underlying, the maturity of the option and the volatility to get the price. The four former quantities can be easily observed in the market.

In reality, the price of a European call or put option is quoted in the market. It may now arise the question which volatility has been used to get the quoted market price. This input is called implied volatility. So in financial mathematics, "the implied volatility is the volatility of the underlying which when substituted into the Black-Scholes formula gives a theoretical price equal to the market price." (see, [17, p.130])

More mathematically, the implied volatility can be defined as follows in the below-mentioned framework (see [15, p.220]):

Let us consider a European call option with exercise price K, maturity T and time $\tau = T - t$ to the maturity T. Assume that the current market price C_t^m can be taken as an input due to the fact that it can be observed in the market.

Definition 6.1.1. The *implied volatility* at any time point $t \leq T$, denoted by $\hat{\sigma}_t$, is derived from the non linear equation (the well-known Black-Scholes equation)

$$C_t^m = S_t \Phi(d_1(S_t, \tau, \hat{\sigma}_t, K, r)) - K e^{-r\tau} \Phi(d_2(S_t, \tau, \hat{\sigma}_t, K, r)),$$

with

$$d_1(s, \tau) = \frac{\ln(\frac{S}{K}) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

and

$$d_2(s, \tau) = d_1(s, t) - \sigma\sqrt{\tau},$$

where the only unknown quantity is $\hat{\sigma}_t$.
This means that the implied volatility $\hat{\sigma}_t$ is the value that, when put in the Black-Scholes formula, results in a model price equal to the current market price of a call option.

Remark 6.1.2. The implied volatility can equivalently be defined by the corresponding price of a European put option. Due to the put-call parity relationship, the implied volatility of a European call option coincides with that of a European put option. Very small differences may be caused by the bid-ask spread.

6.2 Small-Time Asymptotics for Implied Volatility under the One-Dimensional Heston Stochastic Volatility Model

This section which covers the small-time asymptotics for implied volatility under the one-dimensional Heston stochastic volatility model is mainly based on the paper [8] published by Martin Forde and Antoine Jacquier.

The most important results of their studies are summarized:

Theorem 6.2.1. *(see [8, Theorem 1.1]) Consider the one-dimensional Heston stochastic volatility model given by the following dynamics*

$$\begin{aligned} dX_t &= (k + \kappa X_t)dt + 2\sigma\sqrt{X_t}dB_t, \\ dY_t &= -\frac{1}{2}X_t dt + \sqrt{X_t}dW_t, \\ dW_t dB_t &= \tilde{\rho}dt, \end{aligned}$$

with $X_0 = x$, $Y_0 = y$, $\kappa < 0$, $k > 0$, $\sigma > 0$, $|\tilde{\rho}| < 1$ and $k > 2\sigma^2$, so that $X = 0$ is an unattainable barrier, where W_t and B_t are two correlated Brownian motions. Then $Y_t - y$ satisfies a Large deviation principle (LDP) as $t \to 0$, with rate function $\Lambda^\star(x)$ equal to the Legendre transform of the continuous function $\Lambda : \mathbb{R} \to \mathbb{R}^+ \cup \{\infty\}$ given by

$$\begin{aligned} \Lambda(p) &= \frac{xp}{2\sigma(\sqrt{1-\tilde{\rho}^2}\cot(\sigma p\sqrt{1-\tilde{\rho}^2}) - \tilde{\rho})} \qquad \text{for} \qquad p \in (p_-, p_+) \\ &= \infty \qquad \text{for} \qquad p \notin (p_-, p_+), \end{aligned} \tag{6.1}$$

where the following table shows how to compute the values of p_- and p_+:

$\tilde{\rho}$	p_-	p_+
< 0	$\frac{\arctan\left(\frac{\sqrt{1-\tilde{\rho}^2}}{\tilde{\rho}}\right)}{\sigma\sqrt{1-\tilde{\rho}^2}}$	$\frac{\pi+\arctan\left(\frac{\sqrt{1-\tilde{\rho}^2}}{\tilde{\rho}}\right)}{\sigma\sqrt{1-\tilde{\rho}^2}}$
$= 0$	$-\frac{\pi}{2\sigma}$	$\frac{\pi}{2\sigma}$
> 0	$\frac{-\pi+\arctan\left(\frac{\sqrt{1-\tilde{\rho}^2}}{\tilde{\rho}}\right)}{\sigma\sqrt{1-\tilde{\rho}^2}}$	$\frac{\arctan\left(\frac{\sqrt{1-\tilde{\rho}^2}}{\tilde{\rho}}\right)}{\sigma\sqrt{1-\tilde{\rho}^2}}$

Remark 6.2.2. The Fenchel-Legendre transform $\Lambda^\star : \mathbb{R} \to \mathbb{R}$ of the function Λ is defined by

$$\Lambda^\star(x) := \sup_{p \in (p_-, p_+)} \{px - \Lambda(p)\}, \qquad \forall x \in \mathbb{R}.$$

Martin Forde and Antoine Jacquier have proved the above theorem in their paper [8] by applying Gärtner-Ellis Theorem. (An explanation of the Gärtner-Ellis Theorem and the corresponding relevant definitions can be found for example in [8, Appendix A].)

The important issue is to show that

$$\lim_{t\to 0} t \log \mathbb{E}[e^{\frac{p}{t}(Y_t - y)}] = \Lambda(p), \tag{6.2}$$

Remark 6.2.3. Note that (6.2) is equivalent to show that

$$\lim_{t\to 0} t \log \mathbb{E}[e^{\frac{p}{t}Y_t}] = \Lambda(p) + py,$$

because

$$\Lambda(p) = \lim_{t\to 0} t \log \mathbb{E}[e^{\frac{p}{t}(Y_t - y)}] = \lim_{t\to 0} t \log \left(e^{-\frac{p}{t}y}\mathbb{E}[e^{\frac{p}{t}Y_t}]\right) = \lim_{t\to 0} t \left(\log(e^{-\frac{p}{t}y}) + \log(\mathbb{E}[e^{\frac{p}{t}Y_t}])\right)$$
$$= \lim_{t\to 0} t \left(-\frac{p}{t}y + \log \mathbb{E}[e^{\frac{p}{t}Y_t}]\right) = -py + \lim_{t\to 0} t \log \mathbb{E}[e^{\frac{p}{t}Y_t}].$$

Then one can show that $Y_t - y$ satisfies a Large deviation principle as $t \to 0$ with rate function equal to the Legendre transform of the continuous function Λ. Equation (6.2) will play an important role in the subsequent section.

As a next step we study an application of the last theorem, namely the pricing of out-of-the-money call and put options of small maturity.
But before doing so, some terms used in the subsequent theorems and corollaries are repeated: the moneyness and out-of-the-money/ in-the-money/ at-the-money call and put options

A call option is called out-of-the-money (in-the-money) if its strike price K is greater (smaller) than the market price of the underlying asset; therefore for out-of-the-money call options we get $K \geq S_t \Rightarrow \frac{K}{S_t} \geq 1 \Rightarrow \log(\frac{K}{S_t}) \geq 0$.

Contrary, a put option is called out-of-the-money (in-the-money) if its strike price K is smaller (greater) than the market price of the underlying asset; therefore for out-of-the-money put options we have $K \leq S_t \Rightarrow \frac{K}{S_t} \leq 1 \Rightarrow \log(\frac{K}{S_t}) \leq 0$.

Moreover, an option (put or call) is called at-the-money if the strike price equals the market price of the underlying asset, hence $K = S_t$.

We are now prepared to state the following corollaries about the pricing of out-of-the-money call and put options (compare [8, Corollary 2.1 & 2.2]):

Corollary 6.2.4. *The small-time behaviour for out-of-the-money call options on $S_t = e^{Y_t}$ can be written as*

$$-\lim_{t\to 0} t \log \mathbb{E}[S_t - K]^+ = \Lambda^\star(m),$$

where $m = \log(\frac{K}{S_0}) \geq 0$ is the log-moneyness.

Corollary 6.2.5. *The small-time behaviour for out-of-the-money put options on $S_t = e^{Y_t}$ can be written as*

$$-\lim_{t\to 0} t \log \mathbb{E}[K - S_t]^+ = \Lambda^\star(m),$$

where $m = \log(\frac{K}{S_0}) \leq 0$ is the log-moneyness.

Knowing the rate function $\Lambda^\star(m)$, according to Martin Forde and Antoine Jacquier, one can also determine the asymptotic implied volatility:

First we state a theorem which covers the implied volatility of a European call option which is not at-the-money (hence in-the-money or out-of-the money). Secondly, a theorem about the small-time implied volatility of at-the-money options is mentioned.

Theorem 6.2.6. *(see [8, Theorem 2.4]) Let $m = \log\left(\frac{K}{S_0}\right)$ be the log-moneyness. We have the following asymptotic behaviour for the implied volatility $\sigma_t = \sigma_t(m)$ of a European call option on $S_t = e^{Y_t}$ with strike $K = S_0 e^m$ and $m \in \mathbb{R}, m \neq 0$, as $t \to 0$*

$$I(m) = \lim_{t\to 0} \sigma_t(m) = \frac{m}{\sqrt{2\Lambda^\star(m)}}. \tag{6.3}$$

Theorem 6.2.7. *(see [8, Theorem 2.5]) Let $m = \log\left(\frac{K}{S_0}\right)$ be the log-moneyness. The asymptotic implied volatility $I(m)$ has the following expansion around $m = 0$*

$$I(m) = \sqrt{x}\left(1 + \frac{1}{4}\tilde{\rho}z + \left(\frac{1}{24} - \frac{5}{48}\tilde{\rho}^2\right)z^2 + O(z^3)\right), \tag{6.4}$$

where $z = \frac{2\sigma m}{x}$.

Remark 6.2.8. Considering expression (6.3) and (6.4), the interesting property is that in the representation respectively expansion of the asymptotic implied volatility of the European options k and κ do not play a role. Hence these two expressions, namely (6.3) and (6.4), do not depend on the drift of the considered model.

6.3 Small-Time Asymptotics for Implied Volatility under the Marginals of the Multidimensional Heston Stochastic Volatility Model

As previously discussed, the aim of this section is to prove that for the marginals of the multidimensional Heston stochastic volatility model, as given in (5.7) and (5.6), it holds that

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}[e^{\epsilon^{-1} p Y_{\epsilon,i}}] = \Lambda(p) + p y_i, \tag{6.5}$$

with $\Lambda(p)$ as defined in (6.1) and y_i being the initial value of the i-th component of Y_t, where the parameters are matched as already shown in the last chapter.

Remark 6.3.1. To verify that (6.5) is valid, we will at first consider and examine the following expression

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}[e^{\langle \epsilon^{-1} u, X_\epsilon\rangle + \epsilon^{-1} v^\top Y_\epsilon}] \tag{6.6}$$

in the setup of the multidimensional Heston stochastic volatility model.

By choosing u to be the $d \times d$ null matrix ($u = \mathbb{O}$) and v to be a d-dimensional vector whose entries are all equal to 0 except for the i-th component which equals p, expression (6.6) simplifies to the left hand side of (6.5).

If we can show that (6.5) is valid, the results proved by Martin Forde and Antoine Jacquier which we have summarized in the last section are also applicable for the marginals of the multidimensional Heston stochastic volatility model as defined in (5.6) and (5.7).

To show that (6.5) is true in this more general framework, we use some lemmas and statements, Archil Gulisashvili and Josef Teichmann have studied in their paper [11] in Section 4 called "Homogenization Procedure". To be more precise, while they both have stated and proved the following lemmas and statements for affine processes on the canonical state space $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^n$, we will consider an adapted version of these lemmas and statements, namely in a setup for (the marginals of) the multidimensional Heston stochastic volatility models.

At first let us define the set $\mathcal{Q}$ and $\mathcal{V}$ by

$$\mathcal{Q} = \Big\{ (t,u,v) \in \mathbb{R}_+ \times \mathcal{S}_d + i\mathcal{S}_d \times \mathbb{C}^d \Big| \\ \mathbb{E}\left[|e^{tr(uX_t)+v^\top Y_t}| \big| X_0 = x, Y_0 = y \right] = \mathbb{E}\left[e^{tr(Re(u)X_t)+Re(v)^\top Y_t} \big| X_0 = x, Y_0 = y \right] < \infty \Big\}$$

and

$$\mathcal{V} = \{(u,v) \in \mathcal{S}_d + i\mathcal{S}_d \times \mathbb{C}^d | \exists t > 0 \text{ such that } (t,u,v) \in \mathcal{Q}\}.$$

Then the following two lemmas are stated for the solutions of the generalized Riccati equations of multivariate affine stochastic processes:

Lemma 6.3.2. *Let Ψ be the unique solution of the generalized Riccati equation as given in the second equation of (4.2), namely*

$$\frac{\partial}{\partial t}\Psi(t,u,v) = R(\Psi(t,u,v),v), \qquad \Psi(0,u,v) = u.$$

Then, for every $\epsilon > 0$, the function

$$\Psi^\epsilon(t,u,v) := \epsilon\Psi(\epsilon t, \epsilon^{-1}u, \epsilon^{-1}v)$$

solves the equation

$$\frac{\partial}{\partial t}\Psi^\epsilon(t,u,v) = R^\epsilon(\Psi^\epsilon(t,u,v),v), \qquad \Psi^\epsilon(0,u,v) = u \tag{6.7}$$

with $R^\epsilon(u,v) := \epsilon^2 R(\epsilon^{-1}u, \epsilon^{-1}v)$, $\forall (u,v) \in \mathcal{V}$.

Similar, let $\phi := \log\Phi$ be the unique solution of the generalized Riccati equation as given in the *equation* (4.3)*, namely*

$$\frac{\partial}{\partial t}\phi(t,u,v) = F(\Psi(t,u,v),v), \qquad \phi(0,u,v) = 0.$$

Then, for every $\epsilon > 0$*, the function*

$$\phi^\epsilon(t,u,v) := \epsilon\phi(\epsilon t, \epsilon^{-1}u, \epsilon^{-1}v)$$

solves the equation

$$\frac{\partial}{\partial t}\phi^\epsilon(t,u,v) = F^\epsilon(\Psi^\epsilon(t,u,v),v), \qquad \phi^\epsilon(0,u,v) = 0 \tag{6.8}$$

with $F^\epsilon(u,v) = \epsilon^2 F(\epsilon^{-1}u, \epsilon^{-1}v)$*,* $\forall(u,v) \in \mathcal{V}$*.*

Proof. At first we show that (6.7) is true:

$$\begin{aligned}
\frac{\partial}{\partial t}\Psi^\epsilon(t,u,v) &= \frac{\partial}{\partial t}\left(\epsilon\Psi\left(\epsilon t, \epsilon^{-1}u, \epsilon^{-1}v\right)\right) \\
&= \epsilon\frac{\partial}{\partial t}\Psi\left(\epsilon t, \epsilon^{-1}u, \epsilon^{-1}v\right) \\
&= \epsilon R\left(\Psi\left(\epsilon t, \epsilon^{-1}u, \epsilon^{-1}v\right), \epsilon^{-1}v\right)\epsilon \\
&= \epsilon^2 R\left(\Psi\left(\epsilon t, \epsilon^{-1}u, \epsilon^{-1}v\right), \epsilon^{-1}v\right) \\
&= \epsilon^2 R\left(\epsilon^{-1}\epsilon\Psi\left(\epsilon t, \epsilon^{-1}u, \epsilon^{-1}v\right), \epsilon^{-1}v\right) \\
&= \epsilon^2 R(\epsilon^{-1}\Psi^\epsilon(t,u,v), \epsilon^{-1}v) \\
&= R^\epsilon(\Psi^\epsilon(t,u,v),v).
\end{aligned}$$

With exactly the same steps (6.8) can be verified.

□

Lemma 6.3.3. *Under the previous assumptions, the limit* $\lim_{\epsilon\to 0}\Psi^\epsilon = \Psi^{(0)}$ *exists uniformly on compact sets in* $\mathbb{R}_+ \times \mathcal{V}$*. Furthermore*

$$\Psi^\epsilon(t,u,v) = \Psi^{(0)}(t,u,v) + \epsilon\Psi^{(1)}(t,u,v) + \sum_{n\geq 2}\epsilon^n\Psi^{(n)}(t,u,v) \tag{6.9}$$

is a convergent power series expansion for small $\epsilon > 0$*. The coefficient functions in* (6.9) *satisfy certain ordinary differential equations, i.e. in particular*

$$\frac{\partial}{\partial t}\Psi^{(0)}(t,u,v) = R^{(0)}(\Psi^{(0)}(t,u,v),v), \qquad \Psi^{(0)}(0,u,v) = u.$$

and

$$\frac{\partial}{\partial t}\Psi^{(1)}(t,u,v) = \frac{\partial}{\partial \epsilon}\bigg|_{\epsilon=0} R^{\epsilon}(\Psi^{(0)}(t,u,v),v)\Psi^{(1)}(t,u,v) \qquad \Psi^{(1)}(0,u,v) = 0.$$

For $n \geq 2$, the equations for the coefficient functions involve higher order derivatives. In complete analogy, the limit $\lim_{\epsilon\to 0}\phi^{\epsilon} = \phi^{(0)}$ exists uniformly on compact sets in $\mathbb{R}_+ \times \mathcal{V}$. Furthermore

$$\phi^{\epsilon}(t,u,v) = \phi^{(0)}(t,u,v) + \epsilon\phi^{(1)}(t,u,v) + \sum_{n\geq 2}\epsilon^n\phi^{(n)}(t,u,v)$$

for small enough values of ϵ.

As a next step we define by $\hat{\Lambda}^{(i)}, i \geq 0$, the functions appearing in the following power series expansion in ϵ (compare [11, p.15]):

$$\hat{\Lambda}^{(0)}(u,v) + \epsilon\hat{\Lambda}^{(1)}(u,v) + \ldots := \phi^{\epsilon}(1,u,v) + \langle x, \Psi^{\epsilon}(1,u,v)\rangle + v^{\top}y,$$

where (x,y) denotes the initial value of (X,Y).

By applying the homogenization procedure for continuous multivariate affine processes, we obtain

$$\mathbb{E}[e^{\langle \epsilon^{-1}u, X_{\epsilon}\rangle + \epsilon^{-1}v^T Y_{\epsilon}}] = e^{\frac{\hat{\Lambda}^{(0)}(u,v)}{\epsilon} + \hat{\Lambda}^{(1)}(u,v) + \ldots}, \tag{6.10}$$

where (u,v) is such that the expressions on both sides of (6.10) are finite for small enough values of ϵ (compare [11, Remark 4.4.]).

Considering the left hand side of (6.10), we see that we have obtained a similar expression as (6.6). If we take the logarithm, multiply the equation with ϵ and take the limit, we obtain exactly expression (6.6):

$$\begin{aligned}
\mathbb{E}[e^{\langle \epsilon^{-1}u, X_{\epsilon}\rangle + \epsilon^{-1}v^T Y_{\epsilon}}] &= e^{\frac{\hat{\Lambda}^{(0)}(u,v)}{\epsilon} + \hat{\Lambda}^{(1)}(u,v) + \ldots} \\
&\Leftrightarrow \quad \log \mathbb{E}[e^{\langle \epsilon^{-1}u, X_{\epsilon}\rangle + \epsilon^{-1}v^T Y_{\epsilon}}] = \frac{\hat{\Lambda}^{(0)}(u,v)}{\epsilon} + \hat{\Lambda}^{(1)}(u,v) + \ldots \\
&\Leftrightarrow \quad \epsilon \log \mathbb{E}[e^{\langle \epsilon^{-1}u, X_{\epsilon}\rangle + \epsilon^{-1}v^T Y_{\epsilon}}] = \hat{\Lambda}^{(0)}(u,v) + \epsilon\hat{\Lambda}^{(1)}(u,v) + \ldots
\end{aligned}$$

The last expression can further be rewritten as

$$
\begin{aligned}
\epsilon \log \mathbb{E}[e^{\langle \epsilon^{-1}u, X_\epsilon\rangle + \epsilon^{-1}v^\top Y_\epsilon}] &= \hat{\Lambda}^{(0)}(u,v) + \epsilon\hat{\Lambda}^{(1)}(u,v) + \dots \\
&= \phi^\epsilon(1,u,v) + \langle x, \Psi^\epsilon(1,u,v)\rangle + v^\top y \\
&= \phi^{(0)}(1,u,v) + \epsilon\phi^{(1)}(1,u,v) + \sum_{n\geq 2}\epsilon^n\phi^{(n)}(1,u,v) \\
&\quad + \left\langle x, \Psi^{(0)}(1,u,v) + \epsilon\Psi^{(1)}(1,u,v) + \sum_{n\geq 2}\epsilon^n\Psi^{(n)}(1,u,v)\right\rangle + v^\top y.
\end{aligned} \tag{6.11}
$$

If we take the limit on both sides of the last equation, (6.11) can be written as

$$
\lim_{\epsilon\to 0}\epsilon \log \mathbb{E}[e^{\langle \epsilon^{-1}u, X_\epsilon\rangle + \epsilon^{-1}v^\top Y_\epsilon}] = \phi^{(0)}(1,u,v) + \langle x, \Psi^{(0)}(1,u,v)\rangle + v^\top y, \tag{6.12}
$$

where from Lemma 6.3.3 we know that
$\Psi^{(0)}(t,u,v)$ is a solution of $\frac{\partial}{\partial t}\Psi^{(0)}(t,u,v) = R^{(0)}(\Psi^{(0)}(t,u,v),v)$ and $\phi^{(0)}(t,u,v)$ is a solution of $\frac{\partial}{\partial t}\phi^{(0)}(t,u,v) = F^{(0)}(\Psi^{(0)}(t,u,v),v)$.

The results we have stated and proved so far are valid for affine processes in general.

We will now derive $R^{(0)}$ and $F^{(0)}$ for the multidimensional Heston stochastic volatility model:

Remark 6.3.4. As already stated in Remark 4.3.3 for the case of the multidimensional Heston stochastic volatility model, $R(u,v)$ is given by

$$
R(u,v) = 2u\Sigma^\top\Sigma u + \frac{1}{2}vv^\top + u(\Sigma^\top\rho v^\top + M) + (v\rho^\top\Sigma + M^\top)u - \frac{1}{2}\sum_{i=1}^d v_i(e_ie_i^\top).
$$

Since $R^\epsilon(u,v) := \epsilon^2 R(\epsilon^{-1}u, \epsilon^{-1}v)$, it holds that

$$
\begin{aligned}
R^\epsilon(u,v) &:= \epsilon^2 R(\epsilon^{-1}u, \epsilon^{-1}v) \\
&= \epsilon^2 2\epsilon^{-1}u\Sigma^\top\Sigma\epsilon^{-1}u + \epsilon^2\frac{1}{2}\epsilon^{-1}v\epsilon^{-1}v^\top + \epsilon^2\epsilon^{-1}u(\Sigma^\top\rho\epsilon^{-1}v^\top + M) \\
&\quad + \epsilon^2(\epsilon^{-1}v\rho^\top\Sigma + M^\top)\epsilon^{-1}u - \epsilon^2\frac{1}{2}\sum_{i=1}^d\epsilon^{-1}v_i(e_ie_i^\top) \\
&= 2u\Sigma^\top\Sigma u + \frac{1}{2}vv^\top + u\Sigma^\top\rho v^\top + v\rho^\top\Sigma u + \epsilon uM + \epsilon M^\top u - \epsilon\frac{1}{2}\sum_{i=1}^d v_i(e_ie_i^\top) \\
&=: R^{(0)}(u,v) + \epsilon R^{(1)}(u,v),
\end{aligned}
$$

where

$$
R^{(0)}(u,v) := 2u\Sigma^\top\Sigma u + \frac{1}{2}vv^\top + u\Sigma^\top\rho v^\top + v\rho^\top\Sigma u
$$

and

$$R^{(1)}(u,v) := uM + M^\top u - \frac{1}{2}\sum_{i=1}^{d} v_i(e_i e_i^\top).$$

Considering Remark 4.3.3, we see that $F(u,v)$ can be written as

$$F(u,v) = tr(\delta\Sigma^\top\Sigma u).$$

Therefore $F^\epsilon(u,v)$ can be rewritten as

$$\begin{aligned} F^\epsilon(u,v) &:= \epsilon^2 F(\epsilon^{-1}u, \epsilon^{-1}v) \\ &= \epsilon^2 tr(\delta\Sigma^\top\Sigma\epsilon^{-1}u) \\ &= \epsilon^2\epsilon^{-1} tr(\delta\Sigma^\top\Sigma u) \\ &= \epsilon tr(\delta\Sigma^\top\Sigma u) \\ &=: F^{(0)}(u,v) + \epsilon F^{(1)}(u,v), \end{aligned}$$

where

$$F^{(0)}(u,v) := 0$$

and

$$F^{(1)}(u,v) := tr(\delta\Sigma^\top\Sigma u).$$

Using expression (6.12) and applying Remark 6.2.3 and Remark 6.3.4, we are prepared to state the following proposition:

Proposition 6.3.5. *For the marginals of the multidimensional Heston stochastic volatility model as given in* (5.6) *and* (5.7) *it holds that*

$$\lim_{\epsilon\to 0}\epsilon\log\mathbb{E}[e^{\epsilon^{-1}pY_{\epsilon,i}}] = \Lambda(p) + py_i \qquad \Leftrightarrow \qquad \lim_{\epsilon\to 0}\epsilon\log\mathbb{E}[e^{\epsilon^{-1}p(Y_{\epsilon,i}-y_i)}] = \Lambda(p), \tag{6.13}$$

where $\Lambda(p)$ *is given in* (6.1) *with* $\sigma_i = \sqrt{(\Sigma^\top\Sigma)_{ii}}$ *and* $\tilde{\rho}_i = \frac{(\Sigma\rho)_i}{\sqrt{(\Sigma^\top\Sigma)_{ii}}}$.

Proof. Let (X,Y) be a multidimensional Heston stochastic volatility model and let $(\widetilde{X},\widetilde{Y})$ denote a multidimensional Heston stochastic volatility model with the same parameters but with $M = \mathbb{O}$ and the same starting values (x,y).
Then we know by equation (6.12) that

$$\lim_{\epsilon\to 0}\epsilon\mathbb{E}\left[e^{\langle\epsilon^{-1}u,\widetilde{X}_\epsilon\rangle+\epsilon^{-1}v^\top\widetilde{Y}_\epsilon}\right] = \phi^{(0)}(1,u,v) + \langle x, \Psi^{(0)}(1,u,v)\rangle + v^\top y = \langle x, \Psi^{(0)}(1,u,v)\rangle + v^\top y, \tag{6.14}$$

where $\phi^{(0)} = 0$ and $\Psi^{(0)}$ are equal to the same expressions for (X, Y). Since the marginal distribution of $\widetilde{Y}_{\epsilon,i}$ is the same as the one of the one-dimensional Heston stochastic volatility Y_ϵ^{hest} where the parameters are matched accordingly (see Theorem 5.2.2), we deduce

$$\lim_{\epsilon \to 0} \epsilon \mathbb{E}\left[e^{\epsilon^{-1} p \widetilde{Y}_{\epsilon,i}}\right] = \lim_{\epsilon \to 0} \epsilon \mathbb{E}\left[e^{\epsilon^{-1} p Y_\epsilon^{hest}}\right] = \Lambda(p) + p y_i.$$

Since by equation (6.14), we have

$$\lim_{\epsilon \to 0} \epsilon \mathbb{E}\left[e^{\epsilon^{-1} p \widetilde{Y}_{\epsilon,i}}\right] = \lim_{\epsilon \to 0} \epsilon \mathbb{E}\left[e^{\epsilon^{-1} p Y_{\epsilon,i}}\right],$$

it follows that

$$\lim_{\epsilon \to 0} \epsilon \mathbb{E}\left[e^{\epsilon^{-1} p Y_{\epsilon,i}}\right] = \Lambda(p) + p y_i.$$

□

Remark 6.3.6. The intuition why (6.13) is valid is that the linear drift determined by the matrix M does not appear in the above mentioned asymptotic.

That is the reason why the results and statements of Section 6.2 are also valid for the marginals of the multidimensional Heston stochastic volatility model.

Therefore, at this point we can state the following two theorems similar as in the last section (compare Theorem 6.2.6 and Theorem 6.2.7); the only difference is that we choose now the parameters appropriately for the setup of the marginals of the multidimensional Heston stochastic volatility model.

So the first theorem covers the implied volatility of a European call option which is not at-the-money in the setup of a multidimensional Heston stochastic volatility model:

Theorem 6.3.7. *Define $S_t = e^{Y_{t,i}}$ with $Y_{t,i}$ being the i-th component of the log-price in the multidimensional Heston stochastic volatility model and let $m = \log\left(\frac{K}{S_0}\right)$ be the log-moneyness. For the marginals of the multidimensional Heston stochastic volatility model we have the following asymptotic behaviour for the implied volatility $\sigma_t = \sigma_t(m)$ of a European call option on $S_t = e^{Y_{t,i}}$ with strike $K = S_0 e^m$ and $m \in \mathbb{R}, m \neq 0$, as $t \to 0$*

$$I(m) = \lim_{t \to 0} \sigma_t(m) = \frac{m}{\sqrt{2\Lambda^\star(m)}}$$

with $\sigma_i = \sqrt{(\Sigma^\top \Sigma)_{ii}}$ and $\tilde{\rho}_i = \frac{(\Sigma\rho)_i}{\sqrt{(\Sigma^\top \Sigma)_{ii}}}$.

In the second theorem we state an expansion for the asymptotic implied volatility of at-the-money European call options in the setup of the multidimensional Heston stochastic volatility model:

Theorem 6.3.8. *Define $S_t = e^{Y_{t,i}}$ with $Y_{t,i}$ being the i-th component of the log-price in the multidimensional Heston stochastic volatility model and let $m = \log\left(\frac{K}{S_0}\right)$ be the log-moneyness. For the marginals of the multidimensional Heston stochastic volatility model $(X_{t,ii}, Y_{t,i})$ the asymptotic implied volatility $I(m)$ has the following expansion around $m = 0$*

$$I(m) = \sqrt{x_{ii}}\left(1 + \frac{1}{4}\tilde{\rho}_i z_i + \left(\frac{1}{24} - \frac{5}{48}\tilde{\rho}_i^2\right) z_i^2 + O(z_i^3)\right),$$

where $z_i = \frac{2\sqrt{(\Sigma^\top \Sigma)_{ii}}m}{x_{ii}}$ with $X_{0,ii} = x_{ii}$ and $\tilde{\rho}_i = \frac{(\Sigma\rho)_i}{\sqrt{(\Sigma^\top \Sigma)_{ii}}}$.

Chapter 7

Conclusion

In the present thesis a detailed study of the one-dimensional Heston stochastic volatility model determined by the dynamics in (2.1) and its multivariate extension, the so-called multidimensional Heston stochastic volatility model defined by the dynamics in (4.4) was given.
While the variance process equals the well-known Cox-Ingersoll-Ross process in the one-dimensional setup, the covariance process of the multidimensional Heston stochastic volatility model follows the so-called Wishart process introduced by Marie-France Bru (see [3]).

Since both stochastic volatility models belong to the class of affine models, we have shortly introduced and studied the concept of affine processes in the beginning of this thesis. To summarize, the characterizing property of affine processes is the exponential affine form of the conditional characteristic function and the exponent of the conditional characteristic function is determined by two solutions of so-called Riccati equations. Moreover, we have also shown a possible way to derive explicit solutions of the above mentioned Riccati equations for the one-dimensional Heston stochastic volatility model.

As we have already given a detailed overview of this thesis in the Introduction, we will only summarize the most important results at this point:

In the first part of Chapter 5, we have derived the so-called k-th marginal of the multidimensional Heston stochastic volatility model. The important result is that if we choose the matrix determining the linear drift appropriately, the marginals of the multidimensional Heston stochastic volatility model simplify to the well-known one-dimensional Heston stochastic volatility model. Appropriate forms of the above mentioned matrix are a null matrix or a diagonal matrix. Additionally, we have shown that by applying Girsanov's theorem we also get a one-dimensional Heston stochastic volatility model out of the marginals of the multidimensional one.

Moreover, in Chapter 6 of this thesis, we have considered the small-time asymptotics for the implied volatility for a European call option written on one asset in the multidimensional Heston stochastic volatility model. Thereby, we have proved that an expansion of the asymptotic implied volatility proved by Martin Forde and Antoine Jacquier in [8] for the one-dimensional Heston stochastic volatility model can be extended to the multidimensional one.

Appendix A

Multivariate Stochastic Volatility Models in Literature

In this chapter of the Appendix two famous examples of multivariate stochastic volatility models, where the variance follows a Wishart process are stated. In the definition of the models the respective notation of the paper is used.

A.1 The Multidimensional Heston Stochastic Volatility Model in the sense of Gouriéroux and Sufana

Christian Gouriéroux and Razvan Sufana have considered a market which consists of n risky assets and one riskfree asset. Their model is determined by infinitesimal geometric returns of these n risky assets which are presented in a n-dimensional vector (see [10]):

The joint dynamics of $\log S_t$ and Σ_t are given by the stochastic differential system:

$$d\log S_t = \left[\mu + \begin{pmatrix} tr(D_1\Sigma_t) \\ \vdots \\ tr(D_n\Sigma_t) \end{pmatrix}\right] dt + \Sigma_t^{\frac{1}{2}} dW_t^S$$

$$d\Sigma_t = (\Omega\Omega^\top + M\Sigma_t + \Sigma_t M^\top)dt + \Sigma_t^{\frac{1}{2}} dW_t^\sigma Q + Q^\top (dW_t^\sigma)^\top \Sigma_t^{\frac{1}{2}},$$

where W_t^S denotes a n-dimensional vector-valued Brownian motion and W_t^σ denotes a $d \times d$ matrix-valued Brownian motion; μ equals a deterministic n-dimensional vector and $D_i, i = 1, \dots, n, \Omega, M, Q$ are $n \times n$ matrices with Ω assumed to be invertible.

The authors also point out in their paper that the joint process $(\log S_t, \Sigma_t)$ is an affine process, which means that the drift and volatility functions are affine functions of $\log S_t$ and Σ_t (compare

[10, p.5-10]).

A.2 The Multidimensional Heston Stochastic Volatility Model in the sense of Fonesca, Grasselli and Tebaldi

José Da Fonesca, Martino Grasselli and Claudio Tebaldi have stated the following assumptions concerning the Wishart Affine Stochastic Volatility Model in [7]:

Assumption 1*:* The continuous time diffusive Factor Model is considered to be Affine in the terminology of Duffie and Kan (1996).

Assumption 2*:* The evolution of asset returns is conditionally Gaussian while the stochastic covariance matrix follows a Wishart process.

The authors assume that the n-dimensional risky asset S_t is given by its risk-neutral dynamics

$$dS_t = diag(S_t)(r\mathbb{I}dt + \sqrt{\Sigma_t}dZ_t)$$

where $\mathbb{I} = (1, \ldots, 1)^\top$ and $Z_t \in \mathbb{R}^n$ denotes an n-dimensional vector Brownian motion. In addition to that they assume that the quadratic variation of the risky assets is given by a matrix analogue of the square root-mean reverting process

$$d\Sigma_t = (\Omega\Omega^\top + M\Sigma_t + \Sigma_t M^\top)dt + \sqrt{\Sigma}dW_t Q + Q^\top(dW_t)^\top\sqrt{\Sigma_t} \tag{A.1}$$

with $\Omega, M, Q \in M_n$ (M_n denotes the set of square matrices), Ω invertible and $W_t \in M_n$ denotes a matrix Brownian motion.

Remark A.2.1*.* The dynamic given in (A.1) is exactly the variance process which Christian Gouriéroux and Razvan Sufana have also used in their model - namely the Wishart process which has been introduced by Marie-France Bru.

To guarantee the strict positivity and the mean-reverting behaviour of the volatility, it is assumed that M is a negative semidefinite matrix and Ω fulfills

$$\Omega\Omega^\top = \beta Q^\top Q$$

with the real parameter $\beta > n - 1$.
The authors interpret the parameters as follows: the matrix-product $\Omega\Omega^\top$ is related to the expected long-term variance-covariance matrix Σ_∞ through the solution to the following linear equation:

$$-\Omega\Omega^\top = M\Sigma_\infty + \Sigma_\infty M^\top.$$

The volatility of volatility matrix Q accounts for the variance-covariance fluctuations.

Assumption 3*:* The Brownian motions of the assets' returns and those driving the covariance matrix are linearly correlated.

A possible way to correlate these Brownian motions is to introduce n real matrices $R_k \in M_n, k = 1, \ldots, n$ such that

$$dZ_t^k = \sqrt{1 - tr(R_k R_k^\top)} dB_t^k + tr(R_k dW_t^\top), \quad k = 1, \ldots, n,$$

where the (vector) Brownian motion B is independent of W.

Bibliography

[1] E. Alòs. A decomposition formula for option prices in the Heston model and applications to option pricing approximation. *Finance and Stochastics*, 16:403–422, 2012.

[2] A. Benabid, H. Bensusan, and N. E. Karoui. Wishart Stochastic Volatility: Asymptotic Smile and Numerical Framework. 2:1–48, 2010.

[3] M.-F. Bru. Wishart Processes. *J. Theoret. Probab.*, 4:725–751, 1991.

[4] C. Cuchiero. Affine and Polynomial Processes. *Diss. ETH No. 19652*, 1:1–200, 2011.

[5] D. Filipovic. *Term-Structure Models: A Graduate Course*. Springer Finance. Springer, 2009.

[6] D. Filipovic and E. Mayerhofer. Affine Diffusion Processes: Theory and Applications. *SSRN eLibrary*, 2009.

[7] J. Fonseca, M. Grasselli, and C Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 10:151–180, 2007.

[8] M. Forde and A. Jacquier. Small-time asymptotics for implied volatility under the heston model. *International Journal of Theoretical and Applied Finance (IJTAF)*, 12(06):861–876, 2009.

[9] C. Gourieroux. Continuous Time Wishart Process for Stochastic Risk. *Econometric Reviews*, 25(2):177–217, 2006.

[10] C. Gourieroux and R. Sufana. Derivative Pricing with Multivariate Stochastic Volatility: Application to Credit Risk. *Les Cahiers du CREF of HEC Montréal Working Paper No. CREF 04-09*, 2004.

[11] A. Gulisashvili and J. Teichmann. The Gaertner-Ellis Theorem, Homogenization, and Affine Processes. *Working Paper*, 2014.

[12] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

[13] E. Mayerhofer. On the Existence of Non-Central Wishart Distributions. *Cornell University Library*, 114:1–11, 2012.

[14] E. Mayerhofer. Stochastic Analysis Methods in Wishart Theory: Part II: Wishart Processes and Wishart Distributions: An Affine Processes Point of View. *Journal of Multivariate Analysis*, CIMPA Workshop:2–31, 2012.

[15] M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling.* Stochastic Modelling and Applied Probability. Springer, 2006.

[16] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Grundlehren der mathematischen Wissenschaften; A Series of Comprehensive Studies in Mathematics. Springer, 1999.

[17] P. Wilmott. *Paul Wilmott on Quantitative Finance.* Number Bd. 1-3 in Paul Wilmott on Quantitative Finance. Wiley, 2006.

[18] J. Wishart. *The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population.* Biometrika, 1928.

[19] L. Wu. Lecture notes: Chapter 04. stochastic volatility models. *http://www.math.ust.hk/~malwu/math6380a/Chapter%2004.pdf*, downloaded 06.06.2014.