Technische Universität München

Zentrum Mathematik

Wishart Processes

Projektarbeit

von

Oliver Pfaffel

Themensteller/in: Prof. Dr. Claudia Klüppelberg
Betreuer/in: Dr. Robert Stelzer
Abgabetermin: 08. September 2008

Hiermit erkläre ich, dass ich die Diplomarbeit selbstständig angefertigt und nur die angegebenen Quellen verwendet habe.

Garching, den 08. September 2008

Acknowledgments

Taking this opportunity I would like to thank Prof. Dr. Claudia Klüppelberg for making this thesis possible.

I also would like to thank my supervisor Dr. Robert Stelzer for all his helpful advice. This thesis would never have existed without his support and encouragement. We had a lot of fruitful and interesting discussions. To him, I offer my most sincere gratitude.

Contents

Chapter 1

Introduction

In this thesis we consider a matrix variate extension of the Cox-Ingersoll-Ross process (see Cox et al. [1985]), i.e. a solution of a one-dimensional stochastic differential equation of the form

$$dS_t = \sigma\sqrt{S_t}\,dB_t + a(b - S_t)\,dt, \quad S_0 = s_0 \in \mathbb{R} \tag{1.1}$$

with positive numbers a, b, σ and a one-dimensional Brownian motion B. Our extension is defined by a solution of the $p \times p$-dimensional stochastic differential equation of the form

$$dS_t = \sqrt{S_t}\,dB_t\,Q + Q^{\mathrm{T}}dB_t^{\mathrm{T}}\sqrt{S_t} + (S_tK + K^{\mathrm{T}}S_t + \alpha Q^{\mathrm{T}}Q)\,dt, \quad S_0 = s_0 \in \mathcal{M}_p(\mathbb{R}) \tag{1.2}$$

where Q and K are real valued $p \times p$-matrices, α a non-negative number and B a $p \times p$-dimensional Brownian motion (that is, a matrix of p^2 independent one-dimensional Brownian motions).
While it is well-known that solutions of (1.1), called CIR processes, always exist, the situation for (1.2) is more difficult. As we will see in this thesis, it is crucial to choose the parameter α in the right way, to guarantee the (strong) existence of unique solutions of (1.2), that are then called Wishart processes. We will derive that it is sufficient to choose the parameter α larger or equal to $p + 1$. That is similar to a result given by Bru [1991].

The characteristic fact of (1.1) is that this process remains positive for a certain choice of b. This makes it suitable for modeling for example an interest rate, which should always be positive because of economic reasons. Hence, this is an approach for pricing bonds (see the section 5.1). If we want to consider some (corporate) bonds jointly, e.g. because they are correlated, the need for a multidimensional extension comes up. See section 5.3 of this thesis or Gourieroux [2007, 3.5.2.] for a discussion of this topic. For the Wishart processes, we have in the case $\alpha \geq p + 1$ that the unique solution of (1.2) remains positive definite for all times.
Another well-known fact is that the conditional(on s_0) distribution of the CIR process at a certain point in time is noncentral chi-square. We will see that the conditional distribution of the Wishart process S_t at time t is a matrix variate extension of the noncentral chi-square distribution, that is called noncentral Wishart distribution.

An application for matrix variate stochastic processes can be found in Fonseca et al. [2008], which model the dynamics of a p risky assets X by

$$dX_t = \operatorname{diag}(X_t)[(r\mathbf{1} + \lambda_t)\,dt + \sqrt{S_t}\,dW_t] \tag{1.3}$$

where r is a positive number, $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^p$, λ_t a p-dimensional stochastic process, interpreted as the risk premium, and Z a p-dimensional Brownian motion. The volatility process S is the Wishart process of (1.2).

In chapter 2, we introduce some notations and give the necessary background for understanding the following chapters. In chapter 3, a review of fundamental terms and results of matrix variate stochastics and the theory of stochastic differential equations is given, and in section 3.5 some results are derived concerning a matrix variate extension of the Ornstein-Uhlenbeck processes. The main work on the theory of Wishart processes will be done in chapter 4, where we give a general theorem about the existence and uniqueness of Wishart processes in section 4.2. In section 4.3, we show that some soutions of (1.2) can be expressed in terms of the matrix variate Ornstein-Uhlenbeck process from section 3.5 and in section 4.4 we give an algorithm to simulate Wishart processes. Finally, we give an outlook on the applications of Wishart processes in mathematical finance in chapter 5.

For further readings about Wishart processes, one could have a look at the paper of Bru [1991] and for more financial applications, one could consider Gourieroux [2007] or Fonseca et al. [2008], for example.

Chapter 2

Preliminaries

In this section we summarize the technical prerequisites which are necessary for matrix variate stochastics.

2.1 Matrix Algebra

Definition 2.1.

(i) *Denote by* $\mathcal{M}_{m,n}(\mathbb{R})$ *the set of all* $m \times n$ *matrices with entries in* $\mathbb{R}$.
If $m = n$, *we write* $\mathcal{M}_n(\mathbb{R})$ *instead.*

(ii) *Write* $GL(p)$ *for the group of all invertible elements of* $\mathcal{M}_p(\mathbb{R})$.

(iii) *Let* $\mathcal{S}_p$ *denote the linear subspace of all symmetric matrices of* $\mathcal{M}_p(\mathbb{R})$.

(iv) *Let* $\mathcal{S}_p^+$ ($\mathcal{S}_p^-$) *denote the set of all symmetric positive (negative) definite matrices of* $\mathcal{M}_p(\mathbb{R})$.

(v) *Denote by* $\overline{\mathcal{S}_p^+}$ *the closure of* $\mathcal{S}_p^+$ *in* $\mathcal{M}_p(\mathbb{R})$, *that is the set of all symmetric positive semidefinite matrices of* $\mathcal{M}_p(\mathbb{R})$.

Let us review some characteristics of positive (semi)definite matrices.

Theorem 2.2 (Positive definite matrices)**.**

(i) $A \in \mathcal{S}_p^+$ *if and only if* $x^{\mathrm{T}}Ax > 0 \;\forall\, x \in \mathbb{R}^p : x \neq 0$

(ii) $A \in \mathcal{S}_p^+$ *if and only if* $x^{\mathrm{T}}Ax > 0 \;\forall\, x \in \mathbb{R}^p : ||x|| = 1$

(iii) $A \in \mathcal{S}_p^+$ *if and only if* A *is orthogonally diagonalizable with positive eigenvalues, i.e. there exists an orthogonal matrix* $U \in \mathcal{M}_p(\mathbb{R})$, $UU^{\mathrm{T}} = I_p$, *such that* $A = UDU^{\mathrm{T}}$ *with a diagonal matrix* $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, *where* $\lambda_i > 0$, $i = 1, \ldots, p$, *are the positive eigenvalues of* A

(iv) *If* $A \in \mathcal{S}_p^+$ *then* $A^{-1} \in \mathcal{S}_p^+$

(v) *Let* $A \in \mathcal{S}_p^+$ *and* $B \in \mathcal{M}_{q,p}(\mathbb{R})$ *with* $q \leq p$ *and rank* r. *Then* $BAB^{\mathrm{T}} \in \overline{\mathcal{S}_q^+}$ *and, if* B *has full rank, i.e.* $r = q$, *then* BAB^{T} *is even positive definite,* $BAB^{\mathrm{T}} \in \mathcal{S}_q^+$.

(vi) $A^{\mathrm{T}}A \in \mathcal{S}_p^+$ *for all* $A \in GL(p)$

(vii) $A \in \overline{\mathcal{S}_p^+}$ *if and only if* $x^{\mathrm{T}}Ax \geq 0 \; \forall \, x \in \mathbb{R}^p : x \neq 0$

(viii) $A \in \overline{\mathcal{S}_p^+}$ *if and only if* A *is orthogonally diagonalizable with non-negative eigenvalues*

(ix) $M^{\mathrm{T}}M \in \overline{\mathcal{S}_p^+}$ *for all* $M \in \mathcal{M}_p(\mathbb{R})$

Proof. See Muirhead [2005, Appendix A8] or Fischer [2005]. □

On $\overline{\mathcal{S}_p^+}$ we are able to define a matrix valued square root function by

Definition 2.3 (Square root of positive semidefinite matrices)**.**
Let $A \in \overline{\mathcal{S}_p^+}$. *According to Theorem 2.2 there exists an orthogonal matrix* $U \in \mathcal{M}_p(\mathbb{R})$, $UU^{\mathrm{T}} = I_p$, *such that* $A = UDU^{\mathrm{T}}$ *with* $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, *where* $\lambda_i \geq 0$, $i = 1, \ldots, p$. *Then we define the square root of* A *by* $\sqrt{A} = U\mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_p})U^{\mathrm{T}}$, *what is a positive semidefinite matrix, too.*

The square root $\sqrt{A}$ is well-defined and independent of U, as can be seen in Fischer [2005], for example.

Remark 2.4. *If* $A \in \mathcal{S}_p^+$ *then* $\sqrt{A} \in \mathcal{S}_p^+$.

Now, we talk about the differentiation of matrix valued functions.

Definition 2.5. *Let* $S \in \mathcal{M}_p(\mathbb{R})$. *We define the differential operator* $D = (\frac{\partial}{\partial S_{ij}})_{i,j}$ *for all real valued, differentiable functions* $f : \mathcal{M}_p(\mathbb{R}) \to \mathbb{R}$ *as the matrix of all partial derivations* $\frac{\partial}{\partial S_{ij}} f(S)$ *of* $f(S)$.

The following calculation rules are going to be helpful in the next chapters, so we state them here. In order not to lengthen this chapter, we only give outlines of the proofs.

Lemma 2.6 (Calculation rules for determinants)**.**
For all $A, B \in \mathcal{M}_p(\mathbb{R})$, $S \in GL(p)$ *and* $H_t : \mathbb{R} \to GL(p)$ *differentiable it holds that*

(i) $\det(AB) = \det(A)\det(B)$ *and* $\det(\alpha A) = \alpha^p \det(A) \quad \forall \, \alpha \in \mathbb{R}$.

(ii) *If* $A \in GL(p)$ *or* $B \in GL(p)$ *then* $\det(I_p + AB) = \det(I_p + BA)$

(iii) $\frac{d}{dt}\det(H_t) = \det(H_t)\mathrm{tr}(H_t^{-1}\frac{d}{dt}H_t)$

(iv) $D(\det(S)) = \det(S)(S^{-1})^{\mathrm{T}}$

(v) $\det(A)$ *is the product of the eigenvalues of* A.

If S *is furthermore symmetric then*

(vi) $D(\det(S)) = \det(S)S^{-1}$

(vii) $\frac{\partial^2}{\partial S_{ij}\,\partial S_{kl}}(\det(S)) = \det(S)[(S^{-1})_{kl}(S^{-1})_{ij} - (S^{-1})_{ik}(S^{-1})_{lj}]$
where $(S^{-1})_{ij}$ *denotes the* i,j*-th entry of* S^{-1}

Proof. (i) and (v) can be found in every Linear Algebra book as e.g. in Fischer [2005],
(ii) follows from (i).
(iii) can be proven using the Laplace expansion and the property $\operatorname{adj}(A) = \det(A)A^{-1}$ for the adjugate matrix $\operatorname{adj}(A)$.
(iv) follows by using the Leibniz formula for determinants and again the adjugate property,
(vi) is just a special case of (iv).
Finally, (vii) follows from (vi) using $\frac{\partial}{\partial S_{kl}}D(\det(S)) = \det(S)[(S^{-1})_{kl}S^{-1} + \frac{\partial}{\partial S_{kl}}S^{-1}]$ and $\frac{\partial}{\partial S_{kl}}S^{-1} = -S^{-1}(\frac{\partial}{\partial S_{kl}}S)S^{-1}$. □

Lemma 2.7 (Calculation rules for trace)**.** *For all* $A, B, S \in \mathcal{M}_p(\mathbb{R})$ *it holds that*

(i) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ *and* $\operatorname{tr}(\alpha A + B) = \alpha\operatorname{tr}(A) + \operatorname{tr}(B) \quad \forall\, \alpha \in \mathbb{R}$

(ii) $\operatorname{tr}(A)$ *is the sum of the eigenvalues of* A*.*

Proof. See Fischer [2005]. □

Lemma 2.8 (Calculation rules for adjugate matrices)**.** *For all* $A \in GL(p)$ *it holds that*

(i) $\operatorname{adj}(A) = \det(A)A^{-1}$

(ii) $\operatorname{tr}(\operatorname{adj}(A)) = \det(A)\operatorname{tr}(A^{-1})$

Proof. For (i) see Fischer [2005], (ii) is a trivial consequence of (i). □

The next Definition gives us a one-to-one relationship between vectors and matrices. The idea is the following: Suppose we want to transfer a theorem that holds for multivariate stochastic processes to one that holds for matrix variate stochastic processes S. Then we can apply the theorem to $\operatorname{vec}(S)$ and, if necessary, apply vec^{-1} to the resulting 'multivariate processes' to get the theorem in a matrix variate version.

Definition 2.9. *Let* $A \in \mathcal{M}_{m,n}(\mathbb{R})$ *with columns* $a_i \in \mathbb{R}^m$, $i = 1, \ldots, n$*. Define the function* $\operatorname{vec} : \mathcal{M}_{m,n}(\mathbb{R}) \to \mathbb{R}^{mn}$ *via*

$$\operatorname{vec}(A) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Sometimes, we will also consider $\operatorname{vec}(A)$ as an element of $\mathcal{M}_{mn,1}(\mathbb{R})$.

Remark 2.10.

$$vec(A^{\mathrm{T}}) = \begin{pmatrix} \tilde{a}_1^{\mathrm{T}} \\ \vdots \\ \tilde{a}_m^{\mathrm{T}} \end{pmatrix}$$

where $\tilde{a}_j \in \mathbb{R}^n$, $j = 1, \ldots, m$ *denote the rows of* A*.*

Lemma 2.11. *(Cf. Gupta and Nagar [2000][Theorem 1.2.22])*

(i) For $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$ *it holds that* $\mathrm{tr}(A^{\mathrm{T}}B) = \mathrm{vec}(A)^{\mathrm{T}}\mathrm{vec}(B)$

(ii) Let $A \in \mathcal{M}_{p,m}(\mathbb{R})$*,* $B \in \mathcal{M}_{m,n}(\mathbb{R})$ *and* $C \in \mathcal{M}_{n,q}(\mathbb{R})$*. Then we have*

$$\mathrm{vec}(AXB) = (B^{\mathrm{T}} \otimes A)\mathrm{vec}(X)$$

Before we continue, we introduce a new notation: For every linear operator $\mathcal{A}$ on a finite dimensional space we denote by $\sigma(\mathcal{A})$ the *spectrum* of $\mathcal{A}$, that is the set of all eigenvalues of $\mathcal{A}$.

Lemma 2.12. *Let* $A \in \mathcal{M}_p(\mathbb{R})$ *be a matrix such that* $0 \notin \sigma(A) + \sigma(A)$*. Define the linear operator*

$$\mathcal{A} : \mathcal{S}_p \to \mathcal{S}_p,\ X \mapsto AX + XA^{\mathrm{T}}$$

Then the inverse of $\mathcal{A}$ *is given by*

$$\mathcal{A}^{-1} = \mathrm{vec}^{-1} \circ (I_p \otimes A + A \otimes I_p)^{-1} \circ \mathrm{vec}$$

Proof. Can be shown with Lemma 2.11 (ii), for details see Stelzer [2007, p. 66] □

For symmetric matrices there also exists another operator which transfers a matrix into a vector.

Definition 2.13. *(Cf. Gupta and Nagar [2000, Definition 1.2.7.]) Let* $S \in \mathcal{S}_p$*. Define the function* $\mathrm{vech} : \mathcal{S}_p \to \mathbb{R}^{\frac{p(p+1)}{2}}$ *via*

$$\mathrm{vech}(S) = \begin{pmatrix} S_{11} \\ S_{12} \\ S_{22} \\ \vdots \\ S_{1p} \\ \vdots \\ S_{pp} \end{pmatrix}$$

such that $\mathrm{vech}(S)$ *is a vector consisting of the elements of* S *from above and including the diagonal, taken columnwise.*

Compared to vec, the operator vech only takes the $\frac{p(p+1)}{2}$ distinct elements of a symmetric $p \times p$-matrix.

2.2 Some Functions of Matrices

In this section we give a brief introduction to matrix variate functions that arise in the probability density function of the noncentral Wishart distribution.

Definition 2.14. *(Borel-σ-algebra, cf. Jacod and Protter [2004, p. 48]) Let $(X, \mathcal{T})$ be a topological space. The Borel-σ-algebra on X is then given by the smallest σ-algebra that contains $\mathcal{T}$ and will be denoted by $\mathcal{B}(X)$.*

If we work with the spaces $\mathbb{R}$, $\mathbb{R}^n$ or $\mathcal{M}_{m,n}(\mathbb{R})$ we assume that they have the natural topology, which is the set of all possible unions of open balls, given the euclidean metric. See B.v.Querenburg [2001, p.22] for details on topological spaces. We write $\mathcal{B}$ instead of $\mathcal{B}(\mathbb{R})$, $\mathcal{B}^n$ instead of $\mathcal{B}(\mathbb{R}^n)$ and $\mathcal{B}^{m,n}$ instead of $\mathcal{B}(\mathcal{M}_{m,n}(\mathbb{R}))$.

Definition 2.15 (Integration)**.** *Let $f : \mathcal{M}_{m,n}(\mathbb{R}) \to \mathbb{R}$ be a $\mathcal{B}^{m,n}$-$\mathcal{B}$-measurable function and $M \in \mathcal{B}^{m,n}$ a measurable subset of $\mathcal{M}_{m,n}(\mathbb{R})$ and let λ denote the Lebesgue-measure on $(\mathbb{R}^{mn}, \mathcal{B}^{mn})$. The integral of f over M is then defined by*

$$\int_M f(X)\, dX := \int_M f(X)\, d(\lambda \circ \mathrm{vec})(X) = \int_{\mathrm{vec}(M)} f \circ \mathrm{vec}^{-1}(x)\, d\lambda(x)$$

We call $\lambda \circ \mathrm{vec}$ the Lebesgue-measure on $(\mathcal{M}_{m,n}(\mathbb{R}), \mathcal{B}^{m,n})$.

Because $\mathcal{S}_p$ is isomorphic to $\mathbb{R}^{\frac{p(p+1)}{2}}$ we know that for $p \geq 2$ that $\mathcal{S}_p$ is a real subspace of $\mathcal{M}_p(\mathbb{R})$ and hence a $\lambda \circ \mathrm{vec}$-Null set. Thus, we define another Lebesgue measure on the subspace of all symmetric matrices $\mathcal{S}_p$. This integral is always meant if we integrate on a subset of $\mathcal{S}_p$, like e.g. $\mathcal{S}_p^+$, because an integral w.r.t. $\lambda \circ \mathrm{vec}$ would always be zero.

Definition 2.16 (Integration II)**.** *Let $f : \mathcal{S}_p \to \mathbb{R}$ be a $\mathcal{B}(\mathcal{S}_p)$-$\mathcal{B}$-measurable function and $M \in \mathcal{B}(\mathcal{S}_p)$ a Borel measurable subset of $\mathcal{S}_p$ and let λ denote the Lebesgue-measure on $(\mathbb{R}^{\frac{p(p+1)}{2}}, \mathcal{B}^{\frac{p(p+1)}{2}})$. The integral of f over M is then defined by*

$$\int_M f(X)\, dX := \int_M f(X)\, d(\lambda \circ \mathrm{vech})(X) = \int_{\mathrm{vech}(M)} f \circ \mathrm{vech}^{-1}(x)\, d\lambda(x)$$

We call $\lambda \circ \mathrm{vech}$ the Lebesgue-measure on $(\mathcal{S}_p, \mathcal{B}(\mathcal{S}_p))$.

The following definition is just for convenience and can also be found in Gupta and Nagar [2000].

Definition 2.17. *For $A \in \mathcal{M}_p(\mathbb{R})$ define* $\mathrm{etr}(A) := e^{\mathrm{tr}(A)}$.

Definition 2.18 (Matrix Variate Gamma Function)**.**

$$\Gamma_p(a) := \int_{\mathcal{S}_p^+} \mathrm{etr}(-A) \det(A)^{a - \frac{1}{2}(p+1)}\, dA \quad \forall\ a > \frac{p-1}{2}$$

Gupta and Nagar [2000, Theorem 1.4.1.] show that, for $a > \frac{p-1}{2}$, the matrix variate gamma function can be expressed as a finite product of ordinary gamma functions. Thus, we do not need to worry about the existence of the matrix variate gamma function.

The definition of the Hypergeometric Function is a little bit cumbersome and needs further explanations. Like in Muirhead [2005] by a symmetric *homogeneous polynomial* of degree k in $y_1, \dots, y_m$ we mean a polynomial which is unchanged by a permutation of the subscripts and such that every term in the polynomial has degree k. Denote by V_k the space of all symmetric homogeneous polynomials of degree k in the $\frac{p(p+1)}{2}$ distinct elements of $S \in \mathcal{S}_p^+$. Then, $\operatorname{tr}(S)^k = (S_{11} + \ldots + S_{pp})^k$ is an element of V_k. According to Gupta and Nagar [2000] V_k can be decomposed into a direct sum of irreducible invariant subspaces V_κ where κ is a partition of k. With a *partition* κ of k we mean a p-tuple $\kappa = (k_1, \dots, k_p)$ such that $k_1 \geq \ldots \geq k_p \geq 0$ and $k_1 + \ldots + k_p = k$.

Definition 2.19 (Zonal Polynomials)**.** *The zonal polynomial $C_\kappa(S)$ is the component of $\operatorname{tr}(S)^k$ in the subspace V_κ.*

The Definition implies that $\operatorname{tr}(S)^k = \sum_\kappa C_\kappa(S)$ (according to Gupta and Nagar [2000]). Finally, we are able to state

Definition 2.20 (Hypergeometric Function of matrix argument)**.**
The Hypergeometric Function of matrix argument is defined by

$$ {}_mF_n(a_1, \dots, a_m; b_1, \dots, b_n; S) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_m)_\kappa C_\kappa(S)}{(b_1)_\kappa \cdots (b_n)_\kappa k!} \tag{2.1} $$

where a_i, $b_j \in \mathbb{R}$, S is a symmetric $p \times p$-matrix and $\sum_\kappa$ the summation over all partitions κ of k and $(a)_\kappa = \prod_{j=1}^{p} \left(a - \frac{1}{2}(j-1)\right)_{k_j}$ denotes the generalized hypergeometric coefficient, with $(x)_{k_j} = x(x+1)\cdots(x+k_j-1)$. See Gupta and Nagar [2000, p.30].

Gupta and Nagar [2000, p.34] discuss conditions for the convergence and thus well-definedness of (2.1). A sufficient condition is $m < n+1$.

Remark 2.21. ${}_nF_n(a_1, \dots, a_n; a_1, \dots, a_n; S) = \sum_{k=0}^{\infty} \frac{(\operatorname{tr}(S))^k}{k!} = \operatorname{etr}(S)$

The following Lemma eases later on the calculation of expectations of functions of noncentral Wishart distributed random matrices.

Lemma 2.22. *Let $Z, T \in \mathcal{S}_p^+$. Then*

$$ \int_{\mathcal{S}_p^+} \operatorname{etr}(-ZS) \det(S)^{a - \frac{p+1}{2}} {}_mF_n(a_1, \dots, a_m; b_1, \dots, b_n; ST)\, dS $$
$$ = \Gamma_p(a) \det(Z)^{-a} {}_{m+1}F_n(a_1, \dots, a_m, a; b_1, \dots, b_n; Z^{-1}T) $$

$\forall\, a > \frac{p-1}{2}$.

Proof. This is a special case of Gupta and Nagar [2000, Theorem 1.6.2] □

Chapter 3

Matrix Variate Stochastics

Definition 3.1. *A quadruple* $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\in\mathbb{R}_+}, Q)$ *is called a* filtered probability space *if* Ω *is a set,* $\mathcal{G}$ *is a* σ*-field on* Ω*,* $(\mathcal{G}_t)_{t\in\mathbb{R}_+}$ *is an increasing family of sub-*σ*-fields of* $\mathcal{G}$ *(a* filtration*) and* Q *is a probability measure on* $(\Omega, \mathcal{G})$*. The filtered probability space is said to satisfy the* usual conditions *if*

(i) the filtration is right continuous, i.e. $\bigcap_{s>t} \mathcal{G}_s = \mathcal{G}_t$ *for every* $t \geq 0$*, and*

(ii) if $(\mathcal{G}_t)_{t\in\mathbb{R}_+}$ *is complete, i.e.* $\mathcal{G}_0$ *contains all sets from* $\mathcal{G}$ *having* Q*-probability zero.*

Throughout this thesis we assume $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$ to be a filtered probability space satisfying the usual conditions.

3.1 Distributions

Now we summarize a few results and definitions from Gupta and Nagar [2000].

Definition 3.2 (Random Matrix)**.** *A* $m \times n$*-random matrix* X *is a measurable function* $X : (\Omega, \mathcal{F}) \to (\mathcal{M}_{m,n}(\mathbb{R}), \mathcal{B}^{m,n})$*.*

Definition 3.3 (Probability Density Function)**.** *A nonnegative measureable function* f_X *such that*

$$P(X \in M) = \int_M f_X(A)\, dA \quad \forall\, M \in \mathcal{B}^{m\times n}$$

defines the probability density function(p.d.f.) *of a* $m \times n$*-random matrix X.*

Recall that the *Radon-Nikodym* theorem (see Jacod and Protter [2004, Theorem 28.3]) says that the existence of a p.d.f. of X is equivalent to saying that the distribution P^X, of X under P, is absolutely continuous w.r.t the Lebesgue-measure on $(\mathcal{M}_{m,n}(\mathbb{R}), \mathcal{B}^{m,n})$.

Definition 3.4 (Expectation)**.** *Let* X *be a* $m \times n$*-random matrix. For every function* $h = (h_{ij})_{i,j} : \mathcal{M}_{m,n}(\mathbb{R}) \to \mathcal{M}_{r,s}(\mathbb{R})$ *with* $h_{ij} : \mathcal{M}_{m,n}(\mathbb{R}) \to \mathbb{R}$, $1 \leq i \leq r, 1 \leq j \leq s$*, the expected value* $E(h(X))$ *of* $h(X)$ *is an element of* $\mathcal{M}_{r,s}(\mathbb{R})$ *with elements*

$$E(h(X))_{ij} = E(h_{ij}(X)) = \int_{\mathcal{M}_{m,n}(\mathbb{R})} h_{ij}(A)\, P^X(dA)$$

As a matter of fact, $E(h(X))_{ij} = \int_{\mathcal{M}_{m,n}(\mathbb{R})} h_{ij}(A) f_X(A)\, dA$ if X has p.d.f. f_X.

Definition 3.5 (Characteristic Function)**.**
The characteristic function *of a $m \times n$-random matrix X with p.d.f. f_X is defined by*

$$E[\mathrm{etr}(iXZ^{\mathrm{T}})] = \int_{\mathcal{M}_{m,n}(\mathbb{R})} \mathrm{etr}(iAZ^{\mathrm{T}}) f_X(A)\, dA \tag{3.1}$$

for every $Z \in \mathcal{M}_p(\mathbb{R})$.

Remark 3.6.

- *Because of $|exp(ix)| = 1 \quad \forall x \in \mathbb{R}$ the above integral always exists.*
- *If the distribution of X under P is denoted by P^X, then (3.1) is the* Fourier transform *of the measure P^X at point Z and will be denoted by $\widehat{P^X}(Z)$.*

From now on, with the term "X in $\mathcal{S}_p^+$ is a random matrix" we mean that X is a random matrix with $X(\omega) \in \mathcal{S}_p^+$ for almost all $\omega \in \Omega$.

Definition 3.7 (Laplace transform)**.**
The Laplace transform *of a $p \times p$-random matrix X in $\mathcal{S}_p^+$ with p.d.f. f_X is defined by*

$$E[\mathrm{etr}(-UX)] = \int_{\mathcal{S}_p^+} \mathrm{etr}(-UA) f_X(A)\, dA \tag{3.2}$$

for every $U \in \mathcal{S}_p^+$.

Remark 3.8.

- *From Lemma 2.11 we see that* $\mathrm{tr}(UA) = \mathrm{vec}(U)^{\mathrm{T}}\mathrm{vec}(A)$ *represents an inner product on $\mathcal{M}_p(\mathbb{R})$.*
- *For $A, U \in \mathcal{S}_p^+$ we have that* $\mathrm{tr}(-UA) = -\mathrm{tr}(\sqrt{U}A\sqrt{U}) < 0$, *because $\sqrt{U}A\sqrt{U}$ is positive definite. Thus, the integral in (3.2) is well-defined.*

Basically, Levy's Continuity Theorem says that weak convergence of probability measures is equivalent to the pointwise convergence of their respective Fourier transforms.

Theorem 3.9 (Levy's Continuity Theorem)**.** *Let $(\mu_n)_{n\geq 1}$ be a sequence of probability measures on $\mathcal{M}_p(\mathbb{R})$, and let $(\widehat{\mu_n})_{n\geq 1}$ denote their Fourier transforms.*

(i) If μ_n converges weakly to a probability measure μ, $\mu_n \overset{weak}{\to} \mu$, then $\widehat{\mu_n}(Z) \overset{n\to\infty}{\to} \widehat{\mu}(Z)$ pointwise for all $Z \in \mathcal{M}_p(\mathbb{R})$.

(ii) If $f : \mathcal{M}_p(\mathbb{R}) \to \mathbb{R}$ is continuous at 0 and $\widehat{\mu_n}(Z) \overset{n\to\infty}{\to} f(Z)$ for all $Z \in \mathcal{M}_p(\mathbb{R})$, then there exists a probability measure μ on $\mathcal{M}_p(\mathbb{R})$ such that $f(Z) = \widehat{\mu}(Z)$, and $\mu_n \overset{weak}{\to} \mu$.

Proof. See Jacod and Protter [2004, Theorem 19.1] for a proof of a multivariate version that can be extended to the matrix variate case easily. □

With the term analytic function we mean a function that is locally given by a convergent power series.

Lemma 3.10. *(Cf. Jurek and Mason [1993, Lemma 3.8.4.])*
Let f be a real-valued analytic funtion on $\mathcal{M}_p(\mathbb{R})$ which is not identically equal to zero. Then the set $\{x \in \mathcal{M}_p(\mathbb{R}) : f(x) = 0\}$ has p^2-dimensional Lebesgue measure zero.

Definition 3.11 (Covariance Matrix)**.** *Let X be a $m \times n$-random matrix and Y be a $p \times q$-random matrix. The $mn \times pq$ covariance matrix is defined as*

$$\begin{aligned} \mathrm{cov}(X,Y) &= \mathrm{cov}(\mathrm{vec}(X^{\mathrm{T}}), \mathrm{vec}(Y^{\mathrm{T}})) \\ &= E[\mathrm{vec}(X^{\mathrm{T}})\mathrm{vec}(Y^{\mathrm{T}})^{\mathrm{T}}] - E[\mathrm{vec}(X^{\mathrm{T}})]E[\mathrm{vec}(Y^{\mathrm{T}})]^{\mathrm{T}} \end{aligned}$$

i.e. $\mathrm{cov}(X,Y)$ *is a $m \times p$ block matrix with blocks* $\mathrm{cov}(\tilde{x}_i{}^{\mathrm{T}}, \tilde{y}_j{}^{\mathrm{T}}) \in \mathcal{M}_{n,q}(\mathbb{R})$ *where $\tilde{x}_i$ (or resp. $\tilde{y}_i$) denote the rows of X (or resp. Y).*

Definition 3.12 (Matrix Variate Normal Distribution)**.** *A $p \times n$-random matrix is said to have a* matrix variate Normal distribution *with mean $M \in \mathcal{M}_{p,n}(\mathbb{R})$ and covariance $\Sigma \otimes \Psi$ where $\Sigma \in \mathcal{S}_p^+$, $\Psi \in \mathcal{S}_n^+$, if* $\mathrm{vec}(X^{\mathrm{T}}) \sim \mathcal{N}_{pn}(\mathrm{vec}(M^{\mathrm{T}}), \Sigma \otimes \Psi)$ *where $\mathcal{N}_{pn}$ denotes the multivariate Normal distribution on $\mathbb{R}^{pn}$ with mean* $\mathrm{vec}(M^{\mathrm{T}})$ *and covariance $\Sigma \otimes \Psi$. We will use the notation $X \sim \mathcal{N}_{p,n}(M, \Sigma \otimes \Psi)$.*

Lemma 3.13. *(Cf. Gupta and Nagar [2000, Theorem 2.3.1.]) If $X \sim \mathcal{N}_{p,n}(M, \Sigma \otimes \Psi)$, then $X^{\mathrm{T}} \sim \mathcal{N}_{n,p}(M^{\mathrm{T}}, \Psi \otimes \Sigma)$*

Theorem 3.14 (Characteristic Function of the Matrix Variate Normal Distribution)**.** *Let $X \sim \mathcal{N}_{p,n}(M, \Sigma \otimes \Psi)$. Then the characteristic function of X is given by*

$$E[\mathrm{etr}(iXZ^{\mathrm{T}})] = \mathrm{etr}\left(iZ^{\mathrm{T}}M - \frac{1}{2}Z^{\mathrm{T}}\Sigma Z\Psi\right) \tag{3.3}$$

Proof. Cf. Gupta and Nagar [2000, Theorem 2.3.2] □

If Σ and Ψ are not positive definite, but still positive semidefinite we will use (3.3) as a (generalized) definition for the matrix variate Normal distribution.

Theorem 3.15. *Let $X \sim \mathcal{N}_{p,n}(M, \Sigma \otimes \Psi)$, $A \in \mathcal{M}_{m,q}(\mathbb{R})$, $B \in \mathcal{M}_{m,p}(\mathbb{R})$ and $C \in \mathcal{M}_{n,q}(\mathbb{R})$. Then $A + BXC \sim \mathcal{N}_{m,q}(A + BMC, (B\Sigma B^{\mathrm{T}}) \otimes (C^{\mathrm{T}}\Psi C))$.*

Proof. Follows from Theorem 3.14:

$$\begin{aligned} E[\mathrm{etr}(i(A+BXC)Z^{\mathrm{T}})] &= \mathrm{etr}(iAZ^{\mathrm{T}})E[\mathrm{etr}(iX(CZ^{\mathrm{T}}B))] \\ &= \mathrm{etr}(iAZ^{\mathrm{T}})\mathrm{etr}\left(iCZ^{\mathrm{T}}BM - \frac{1}{2}CZ^{\mathrm{T}}B\Sigma B^{\mathrm{T}}ZC^{\mathrm{T}}\Psi\right) \\ &= \mathrm{etr}\left(iZ^{\mathrm{T}}(A+BMC) - \frac{1}{2}Z^{\mathrm{T}}(B\Sigma B^{\mathrm{T}})Z(C^{\mathrm{T}}\Psi C)\right) \end{aligned}$$

□

Definition 3.16 (Noncentral Wishart Distribution). *A $p \times p$-random matrix X in $\mathcal{S}_p^+$ is said to have a* noncentral Wishart distribution *with parameters $p \in \mathbb{N}$, $n \geq p$, $\Sigma \in \mathcal{S}_p^+$ and $\Theta \in \mathcal{M}_p(\mathbb{R})$ if its p.d.f is given by*

$$f_X(S) = \left(2^{\frac{1}{2}np}\,\Gamma_p(\frac{n}{2})\det(\Sigma)^{\frac{n}{2}}\right)^{-1} \operatorname{etr}\left(-\frac{1}{2}(\Theta + \Sigma^{-1}S)\right) det(S)^{\frac{1}{2}(n-p-1)} {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\Theta\Sigma^{-1}S\right) \tag{3.4}$$

where $S \in \mathcal{S}_p^+$ and ${}_0F_1$ is the hypergeometric function. We write $X \sim \mathcal{W}_p(n, \Sigma, \Theta)$.

Remark 3.17.

- *The requirement $n \geq p$ assures that the matrix variate gamma function is well-defined. For the case $p-1 \leq n \leq p$ or resp. $n \in \{1, \ldots, p-1\}$ one could use the Laplace transform (3.5) or resp. Lemma 3.20 to define the Wishart distribution for this case. However, Olkin and Rubin [1961, Appendix] have shown for non-integer $n < p-1$ that (3.5) is not the Laplace transform of a probability distribution anymore.*

- *If $\Sigma \in \overline{\mathcal{S}_p^+} \backslash \mathcal{S}_p^+$ the p.d.f. does not exist, but we can define the noncentral Wishart distribution using the characteristic function from Theorem 3.19*

- *If $\Theta = 0$, X is said to have (central) Wishart distribution with parameters p, n and $\Sigma \in \mathcal{S}_p^+$ and its p.d.f. is given by*

$$\left(2^{\frac{1}{2}np}\,\Gamma_p(\frac{n}{2})\det(\Sigma)^{\frac{n}{2}}\right)^{-1} \operatorname{etr}\left(-\frac{1}{2}\Sigma^{-1}S\right) det(S)^{\frac{1}{2}(n-p-1)}$$

  where $S \in \mathcal{S}_p^+$ and $n \geq p$ (see Gupta and Nagar [2000, Definition 3.2.1]).

Theorem 3.18 (Laplace Transform of the Noncentral Wishart Distribution).
Let $S \sim \mathcal{W}_p(n, \Sigma, \Theta)$. Then the Laplace transform of S is given by

$$E(\operatorname{etr}(-US)) = \det(I_p + 2\Sigma U)^{-\frac{n}{2}} \operatorname{etr}[-\Theta(I_p + 2\Sigma U)^{-1}\Sigma U] \tag{3.5}$$

with $U \in \mathcal{S}_p^+$

Proof.

$$\begin{aligned} E(\operatorname{etr}{-US}) &= \int_{\mathcal{S}_p^+} \operatorname{etr}(-US) f_S(S)\, dS = \left(2^{\frac{1}{2}np}\,\Gamma_p(\frac{n}{2})\det(\Sigma)^{\frac{n}{2}}\right)^{-1} \operatorname{etr}\left(-\frac{1}{2}\Theta\right) \\ &\quad \times \int_{\mathcal{S}_p^+} \operatorname{etr}\left(-US - \frac{1}{2}\Sigma^{-1}S\right) det(S)^{\frac{1}{2}(n-p-1)} {}_0F_1\left(\frac{n}{2}; \frac{1}{4}\Theta\Sigma^{-1}S\right) dS \end{aligned}$$

With Lemma 2.22 and Remark 2.21 we get

$$\begin{aligned} &\int_{\mathcal{S}_p^+} \det(S)^{\frac{1}{2}(n-p-1)} \operatorname{etr}(-US - \frac{1}{2}\Sigma^{-1}S) {}_0F_1(\frac{n}{2}; \frac{1}{4}\Theta\Sigma^{-1}S)\, dS \\ =\ & \Gamma_p(\frac{n}{2}) \det(U + \frac{1}{2}\Sigma^{-1})^{-\frac{n}{2}} {}_1F_1(\frac{n}{2}, \frac{n}{2}; \frac{1}{4}(U + \frac{1}{2}\Sigma^{-1})^{-1}\Theta\Sigma^{-1}) \\ =\ & \Gamma_p(\frac{n}{2}) \det(\frac{1}{2}\Sigma^{-1}(I_p + 2\Sigma U))^{-\frac{n}{2}} {}_1F_1(\frac{n}{2}, \frac{n}{2}; \frac{1}{2}(I_p + 2\Sigma U)^{-1}\Theta) \\ =\ & 2^{\frac{1}{2}np}\,\Gamma_p(\frac{n}{2}) \det(\Sigma)^{\frac{n}{2}} \det(I_p + 2\Sigma U)^{-\frac{n}{2}} \operatorname{etr}(\frac{1}{2}(I_p + 2\Sigma U)^{-1}\Theta) \end{aligned}$$

Finally,

$$\begin{aligned}
& \operatorname{etr}(-\frac{1}{2}\Theta + \frac{1}{2}(I_p + 2\Sigma U)^{-1}\Theta) \\
= & \operatorname{etr}(-\frac{1}{2}\Theta(I_p - (I_p + 2\Sigma U)^{-1})) \\
= & \operatorname{etr}(-\frac{1}{2}\Theta(I_p + 2\Sigma U)^{-1}(I_p + 2\Sigma U - I_p) \\
= & \operatorname{etr}(-\Theta(I_p + 2\Sigma U)^{-1}\Sigma U)
\end{aligned}$$

and altogether

$$\begin{aligned}
E(\operatorname{etr}{-US}) &= \int_{\mathcal{S}_p^+} \operatorname{etr}(-US) f_S(S)\, dS = (2^{\frac{1}{2}np}\,\Gamma_p(\frac{n}{2})\det(\Sigma)^{\frac{n}{2}})^{-1}\operatorname{etr}(-\frac{1}{2}\Theta) \\
&\quad \times \int_{\mathcal{S}_p^+} \operatorname{etr}(-US - \frac{1}{2}\Sigma^{-1}S)\det(S)^{\frac{1}{2}(n-p-1)}{}_0F_1(\frac{n}{2};\frac{1}{4}\Theta\Sigma^{-1}S)\, dS \\
&= \det(I_p + 2\Sigma U)^{-\frac{n}{2}}\operatorname{etr}[-\Theta(I_p + 2\Sigma U)^{-1}\Sigma U]
\end{aligned}$$

□

Theorem 3.19. *(Characteristic Function of the Noncentral Wishart Distribution, cf. Gupta and Nagar [2000, Theorem 3.5.3.])*
Let $S \sim \mathcal{W}_p(n, \Sigma, \Theta)$. Then the characteristic function of S is given by

$$E(\operatorname{etr}(iZS)) = \det(I_p - 2i\Sigma Z)^{-\frac{n}{2}} etr[i\Theta(I_p - 2i\Sigma Z)^{-1}\Sigma Z] \tag{3.6}$$

with $Z \in \mathcal{M}_p(\mathbb{R})$

The next Lemma shows that the noncentral Wishart distribution is the square of a matrix variate normal distributed random matrix. Hence, it is the matrix variate extension of the noncentral chi-square distribution.

Lemma 3.20. *(Cf. Gupta and Nagar [2000, Theorem 3.5.1.])*
Let $X \sim \mathcal{N}_{p,n}(M, \Sigma \otimes I_n)$, $n \in \{p, p+1, \ldots\}$. Then $XX^{\mathrm{T}} \sim \mathcal{W}_p(n, \Sigma, \Sigma^{-1}MM^{\mathrm{T}})$.

Consider $S \sim \mathcal{W}_p(n, \Sigma, \Theta)$. With the foregoing Lemma we can interpret the parameter Σ as a scale and the parameter Θ as a location parameter for S. Especially, a central Wishart distributed matrix may be thought of a matrix square of normally distributed matrices with zero mean.

3.2 Processes and Basic Stochastic Analysis

First, we make a general definition.

Definition 3.21 (Matrix Variate Stochastic Process)**.**
A measurable function $X : \mathbb{R}_+ \times \Omega \to \mathcal{M}_{m,n}(\mathbb{R}), (t, \omega) \mapsto X(t, \omega) = X_t(\omega)$ is called a (matrix variate) stochastic process *if $X(t, \omega)$ is a random matrix for all $t \in \mathbb{R}_+$.*
Moreover, X is called a stochastic process in $\overline{\mathcal{S}_p^+}$ if $X : \mathbb{R}_+ \times \Omega \to \overline{\mathcal{S}_p^+}$.

With $\mathbb{R}_+$ we always mean the interval of all non-negative real numbers, i.e $[0, \infty)$. We can transfer most concepts from real stochastics to matrix variate stochastics, if we say a matrix variate stochastic process X has property P if each component of X has property P. We explicitly write this down for only a few cases:

Definition 3.22 (Local Martingale)**.** *A matrix variate stochastic process X is called a* local martingale*, if each component of X is a local martingale, i.e. if there exists a sequence of strictly monotonic increasing stopping times $(T_n)_{n\in\mathbb{N}}$,$T_n \overset{\text{a.s.}}{\to} \infty$, such that $X_{\min\{n,T_n\},ij}$ is a martingale for all i, j.*

Definition 3.23 (Matrix Variate Brownian Motion)**.**
A matrix variate Brownian motion *B in $\mathcal{M}_{n,p}(\mathbb{R})$ is a matrix consisting of independent, one-dimensional Brownian motions, i.e. $B = (B_{ij})_{i,j}$ where B_{ij} are independent one-dimensional Brownian motions, $1 \le i \le n$, $1 \le j \le p$. We write $B \sim \mathcal{BM}_{n,p}$ (and $B \sim \mathcal{BM}_n$ if $p = n$).*

Remark 3.24. $B_t \sim \mathcal{N}_{n,p}(0, tI_{np})$

Theorem 3.25. *Let $W \sim \mathcal{BM}_{n,p}$, $A \in \mathcal{M}_{m,q}(\mathbb{R})$, $B \in \mathcal{M}_{m,n}(\mathbb{R})$ and $C \in \mathcal{M}_{p,q}(\mathbb{R})$. Then $A + BW_tC \sim \mathcal{N}_{m,q}(A, t(BB^{\mathrm{T}}) \otimes (C^{\mathrm{T}}C))$.*

Proof. Follows from Theorem 3.15 and the fact that $I_{np} = I_n \otimes I_p$ □

Definition 3.26 (Semimartingale)**.** *A matrix variate stochastic process X is called a* semimartingale *if X can be decomposed into $X = X_0+M+A$ where M is a local martingale and A an adapted process of finite variation.*

We will only consider continuous semimartingales in this thesis.

For a $n\times p$-dimensional Brownian motion $W \sim \mathcal{BM}_{n,p}$, stochastic processes X resp. Y in $\mathcal{M}_{m,n}(\mathbb{R})$ or resp. $\mathcal{M}_{p,q}(\mathbb{R})$ and a stopping time T the matrix variate *stochastic integral* on $[0, T]$ is meant to be a matrix with entries

$$\left(\int_0^T X_t \, dW_t \, Y_t\right)_{i,j} = \sum_{k=1}^{n}\sum_{l=1}^{p}\int_0^T X_{t,ik}\, Y_{t,lj} dW_{t,kl} \quad \forall\, 1 \le i \le m,\, 1 \le j \le q$$

Theorem 3.27 (Matrix Variate Itô Formula on Open Subsets)**.** *Let $U \subseteq \mathcal{M}_{m,n}(\mathbb{R})$ be open, X be a continuous semimartingale with values in U and let $f : U \to \mathbb{R}$ be a twice continuously differentiable function. Then $f(X)$ is a continuous semimartingale and*

$$f(X_t) = f(X_0) + \operatorname{tr}\left(\int_0^t Df(X_s)^{\mathrm{T}}\, dX_s\right) + \frac{1}{2}\int_0^t \sum_{j,l=1}^{n}\sum_{i,k=1}^{m} \frac{\partial^2}{\partial X_{ij}\,\partial X_{kl}} f(X_s)\, d[X_{ij}, X_{kl}]_s \tag{3.7}$$

with $D = (\frac{\partial}{\partial X_{ij}})_{i,j}$

Proof. Follows from Revuz and Yor [2001, Theorem 3.3 and Remark 2, Chapter IV] and Lemma 2.11 □

Corollary 3.28. *Let X be a continuous semimartingale on a stochastic interval $[0, T)$ with $T = \inf\{t : X_t \notin U\} > 0$ for an open set $U \subseteq \mathcal{M}_{m,n}(\mathbb{R})$ and let $f : U \to \mathbb{R}$ be a twice continuously differentiable function.*
Then $(f(X))_{t\in[0,T)}$ is a continuous semimartingale and (3.7) holds for $t \in [0, T)$.

Proof. For any subset $A \subseteq \mathcal{M}_p(\mathbb{R})$, define the distance between A and any point $x \in \mathcal{M}_p(\mathbb{R})$ by $d(x, A) := \inf_{z\in A} d(x, z)$. Clearly, it exists an $N = N(\omega) \in \mathbb{N}$ such that $d(X_0(\omega), \partial U) \leq \frac{1}{N}$. The continuity of X, $X_0 \in U$ and $T > 0$ a.s imply $N < \infty$ a.s. As X is adapted, $(T_n)_{n\geq N}$ with

$$T_n := \inf\{t \geq 0 : \; d(X_t, \partial U) \leq \frac{1}{n}\}$$

defines a sequence of stopping times. For every $n \geq N$, the stopped process $X^{T_n} := X_{t\wedge T_n}$ is a continuous semimartingale with values in U, thus Theorem 3.27 can be applied to see that $f(X^{T_n})$ is a continuous semimartingale and (3.7) holds for X^{T_n}. Because T_n converges to $T = \inf\{t \geq 0 : \; X_t \in \partial U\}$, we can conclude that $(f(X))_{t\in[0,T)}$ is a continuous semimartingale and (3.7) holds for $(X)_{t\in[0,T)}$. □

With the Definition of a ‘Matrix Quadratic Covariation’ we are able to state the matrix variate partial integration formula in a handy way.

Definition 3.29 (Matrix Quadratic Covariation). *For two semimartingales $A \in \mathcal{M}_{d,m}(\mathbb{R})$, $B \in \mathcal{M}_{m,n}(\mathbb{R})$ the matrix variate* quadratic covariation *is defined by*

$$[A, B]^M_{t,ij} = \sum_{k=1}^{m} [A_{ik}, B_{kj}]_t \in \mathcal{M}_{d,n}(\mathbb{R})$$

Theorem 3.30 (Matrix Variate Partial Integration). *(Cf. Barndorff-Nielsen and Stelzer [2007, Lemma 5.11.]) Let $A \in \mathcal{M}_{d,m}(\mathbb{R})$, $B \in \mathcal{M}_{m,n}(\mathbb{R})$ be two semimartingales. Then the matrix product $A_tB_t \in \mathcal{M}_{d,n}(\mathbb{R})$ is a semimartingale and*

$$A_tB_t = \int_0^t A_{t-} \, dB_t + \int_0^t dA_t \, B_{t-} + [A, B]^M_t$$

where A_{t-} denotes the limit from the left of A_t.

3.3 Stochastic Differential Equations

When we talk about (matrix variate) stochastic differential equations, we can classify between two different kind of solutions: Weak and strong solutions. Intuitively, a strong solution is constructed from a given Brownian motion and hence a ‘function’ of that Brownian motion.

Definition 3.31. *(Weak and Strong Solutions, cf. Revuz and Yor [2001, Chapter IX, Definition 1.5]) Let $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\in\mathbb{R}_+}, Q)$ be a filtered probability space satisfying the usual conditions and consider the stochastic differential equation*

$$dX_t = b(t, X_t) \, dt + \sigma(t, X_t) \, dW_t, \quad X_0 = x_0 \tag{3.8}$$

where $b : \mathbb{R}_+ \times \mathcal{M}_{m,n}(\mathbb{R}) \to \mathcal{M}_{m,n}(\mathbb{R})$ *and* $\sigma : \mathbb{R}_+ \times \mathcal{M}_{m,n}(\mathbb{R}) \to \mathcal{M}_{m,p}(\mathbb{R})$ *are measurable functions,* $x_0 \in \mathcal{M}_{m,n}(\mathbb{R})$ *and* W *is a* $p \times n$*-dimensional Brownian motion.*

- *A pair* (X, W) *of* $\mathcal{F}_t$*-adapted continuous processes defined on* $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\in\mathbb{R}_+}, Q)$ *is called*
  a solution *of (3.8) on* $[0, T)$*,* $T > 0$*, if* W *is a* $(\mathcal{G}_t)_{t\in\mathbb{R}_+}$*-Brownian motion and*

$$X_t = x_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s \quad \forall\, t \in [0, T)$$

- *Moreover, the pair* (X, W) *is said to be a* strong solution *of (3.8), if* X *is adapted to the filtration* $(\mathcal{G}_t^W)_{t\in\mathbb{R}_+}$*, where* $\mathcal{G}_t^W = \sigma_c(W_s, s \leq t)$ *is the* σ*-algebra generated from* W_s*,*$s \leq t$*, that is completed with all* Q*-null sets from* $\mathcal{G}$*.*

- *A solution* (X, W) *which is not strong will be termed a* weak solution *of (3.8).*

To define the *probability law* P^X *of a stochastic process* $X : \mathbb{R}_+ \times \Omega \to \mathcal{M}_{m,n}(\mathbb{R})$, we consider, according to Øksendal [2000], X as a random matrix on the functional space $(\mathcal{M}_{m,n}(\mathbb{R})^{\mathbb{R}_+}, \widehat{\mathcal{F}})$ with σ-algebra $\widehat{\mathcal{F}}$ that is generated by the cylinder sets

$$\{\omega \in \Omega : \; X_{t_1}(\omega) \in F_1, \ldots, X_{t_k}(\omega) \in F_k\}$$

where $F_i \subset \mathcal{M}_{m,n}(\mathbb{R})$ are Borel sets and $k \in \mathbb{N}$.
$X : (\Omega, \mathcal{F}, P) \to (\mathcal{M}_{m,n}(\mathbb{R})^{\mathbb{R}_+}, \widehat{\mathcal{F}})$ is then a measurable function with law P^X.

Definition 3.32. *(Uniqueness, Revuz and Yor [2001, Chapter IX, Definition 1.3]) Consider again the stochastic differential equation (3.8).*

- *It is said that* pathwise uniqueness *holds for (3.8), if for every two solutions* (X, W) *and* (X', W') *defined on the same filtered probability space,* $X_0 = X_0'$ *and* $W = W'$ *a.s. implies that* X *and* X' *are* indistinguishable*, i.e. for* P*-almost all* ω *it holds that* $X_t(\omega) = X_t'(\omega)$ *for every* t*, or equivalently*

$$P\left(\sup_{t\in[0,\infty)} |X_t - X_t'| > 0\right) = 0$$

- *There is* uniqueness in law *for (3.8), if whenever* (X, W) *and* (X', W') *are two solutions with possibly different Brownian motions* W *and* W' *(in particular if* (X, W) *and* (X', W') *are defined on two different filtered probability spaces) and* $X_0 \overset{\mathscr{D}}{=} X_0'$*, then the laws* P^X *and* $P^{X'}$ *are equal. In other words,* X *and* X' *are two versions of the same process, i.e. they have the same finite dimensional distributions (see Revuz and Yor [2001, Chapter I, Definition 1.6]).*

Remark 3.33.

(i) As Yamada and Watanabe [1971a, Proposition 1] have shown, pathwise uniqueness implies uniqueness in law, which is not true conversely.

(ii) If pathwise uniqueness holds for (3.8), then every solution of (3.8) is strong (see Revuz and Yor [2001, Theorem 1.7]).

(iii) The definition of pathwise uniqueness implies that there exists at most one strong solution for (3.8) up to indistinguishability.

(iv) According to Skorohod [1965, p. 59f.], the stochastic differential equation (3.8) always has a weak (but not necessarily unique) solution if b and σ are continuous functions. If in this situation pathwise uniqueness holds for (3.8), then there exists a unique strong solution up to indistinguishability.

Similar to the theory of ordinary differential equations, the function on the right hand side of a stochastic differential equation being locally Lipschitz is sufficient in order to guarantee (strong) existence on a nonempty (stochastic) interval and (pathwise) uniqueness.

Theorem 3.34. *(Existence of Solutions of SDEs driven by a continuous Semimartingale, cf. Stelzer [2007, Theorem 6.7.3.]) Let U be an open subset of $\mathcal{M}_{d,n}(\mathbb{R})$ and $(U_n)_{n\in\mathbb{N}}$ a sequence of convex closed sets such that $U_n \subset U$, $U_n \subseteq U_{n+1} \forall\, n \in \mathbb{N}$ and $\bigcup_{n\in\mathbb{N}} U_n = U$. Assume that $f : U \to \mathcal{M}_{d,m}(\mathbb{R})$ is a locally Lipschitz function and Z in $\mathcal{M}_{m,n}(\mathbb{R})$ is a continuous semimartingale. Then for each U-valued $\mathcal{F}_0$-measurable initial value X_0 there exist a stopping time T and a* unique U-valued strong solution *X to the stochastic differential equation*

$$dX_t = f(X_t)dZ_t \tag{3.9}$$

up to the time $T > 0$ a.s., i.e. on the stochastic interval $[0, T)$. On $T < \infty$ we have that either X hits the boundary ∂U of U at T, i.e. $X_T \in \partial U$ or explodes, i.e. $\limsup_{t\to T, t<T} ||X_t|| = \infty$. If f satifies the linear growth condition

$$||f(X)||^2 \le K(1 + ||X||^2)$$

with some constant $K \in \mathbb{R}_+$, then no explosion can occur.

Remark 3.35. *With the term unique it is meant that there holds pathwise uniqueness for (3.9). In other words, every two solutions (X, Z) and (X', Z) of (3.9) defined on the same probability space and with the same continuous semimartingale Z and the same initial value are indistinguishable.*

Definition 3.36. *(Local Lipschitz, cf. Stelzer [2007, Definition 6.7.1.]) Let $(U, ||.||_U)$, $(V, ||.||_V)$ be two normed spaces and $W \subseteq U$ be open. Then a function $f : W \to V$ is called locally Lipschitz, if for every $x \in W$ there exists an open neighbourhood $\mathcal{U}(x) \subset W$ and a constant $C(x) \in \mathbb{R}_+$ such that*

$$||f(z) - f(y)||_V \le C(x)||z - y||_U \quad \forall\, z, y \in \mathcal{U}(x) \tag{3.10}$$

$C(x)$ is said to be a local Lipschitz coefficient. If there is a $K \in \mathbb{R}_+$ such that $C(x) = K$ can be chosen for all $x \in W$, f is called globally Lipschitz.

If the process Z from Theorem 3.34 is a continous Lévy process (Brownian motion with drift), it can be shown that the solution X of (3.9) is a Markov process. Before we state this in a theorem, we give a general definition of the term 'Markov process' and a necessary technical condition on our probability space.

Definition 3.37. *(Markov Process, cf. Stelzer [2007, Definition 6.7.7.]) Let $U \subseteq \mathcal{M}_p(\mathbb{R})$ be open and Z be a process with values in U which is adapted to a filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.*

(i) Z is called a Markov process *with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, if*

$$E(g(Z_u)|\mathcal{F}_t) = E(g(Z_u)|Z_t)$$

for all $t \in \mathbb{R}_+$, $u \geq t$ and $g : U \to \mathbb{R}$ bounded and Borel measurable.

(ii) Let Z be a Markov process and define for all $s, t \in \mathbb{R}_+, s \leq t$ the transition functions *$P_{s,t}(Z_s, g) := E(g(Z_t)|Z_s)$ with $g : U \to \mathbb{R}$ bounded and Borel measurable. If*

$$P_{s,t} = P_{0,t-s} =: P_{t-s} \text{ for all } s, t \in \mathbb{R}_+, s \leq t$$

then Z is said to be a time homogeneous Markov process*.*

(iii) A time homogeneous Markov process is called a strong Markov *process, if*

$$E(g(Z_{T+s})|\mathcal{F}_T) = P_s(Z_T, g) = E(g(Z_{T+s})|Z_T)$$

for all $g : U \to \mathbb{R}$ bounded and Borel measurable and a.s. finite stopping times T.

At the beginning of this chapter we assumed $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ to be a filtered probability space where $(\mathcal{F}_t)_{t\geq 0}$ is right continuous and all $\mathcal{F}_t$ are complete. Now, we need to *enlarge* our given probability space in order to get arbitrary initial values.

Definition 3.38. *(Enlargement of a probability space, cf. Protter [2004])*
Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ be a filtered probability space where $(\mathcal{F}_t)_{t\geq 0}$ is right continuous, i.e. $\bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$ for every $t \geq 0$, and all $\mathcal{F}_t$ are completed with sets from $\mathcal{F}$ having P-probability zero and $U \subseteq \mathcal{M}_p(\mathbb{R})$ be an open subset of $\mathcal{M}_p(\mathbb{R})$.
The probability space $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_t)_{t\geq 0}, (\overline{P}^y)_{y\in U})$ with

$$\overline{\Omega} = U \times \Omega, \quad \overline{\mathcal{F}}_t = \bigcap_{u>t} \sigma(\mathcal{B}(U) \times \mathcal{F}_u), \quad \overline{\mathcal{F}} = \sigma(\mathcal{B}(U) \times \mathcal{F}), \quad \overline{P}^y = \delta_y \times P$$

is called the Enlargement *of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. The measure δ_y is the Dirac measure w.r.t. $y \in U$, i.e. for $A \subseteq U$ it holds that $\delta_y(A) = 1$ if and only if $y \in A$. A random matrix $Z : \Omega \to \mathcal{M}_p(\mathbb{R})$ is* extended *to $\overline{\Omega}$ by setting $Z((y, \omega)) = Z(\omega)$ for all $(y, \omega) \in \overline{\Omega}$.*

Eventually we are able to state

Theorem 3.39. *(Markov Property of Solutions of SDEs driven by a Brownian Motion, cf. Stelzer [2007, Theorem 6.7.8.]) Recall the Assumptions of Theorem 3.34 (without the linear growth condition), but let this time Z be a Brownian motion with drift in $\mathcal{M}_{m,n}(\mathbb{R})$, i.e. $Z \sim \mathcal{BM}_{m,n}$. Suppose further that $T = \infty$ for every initial value $x_0 \in U$. Consider* *the enlarged probability space* $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_t)_{t\geq 0}, (\overline{P}^y)_{y\in U})$ *and define* X_0 *by* $X_0((y,\omega)) := y$ *for all* $y \in U$*. Then the unique strong solution* X *of*

$$dX_t = f(X_t)dZ_t \tag{3.11}$$

is a time homogeneous strong Markov process on U *under every probability measure of the family* $(\overline{P}^y)_{y\in U}$*.*

The Girsanov Theorem is a powerful tool that gives us the opportunity to construct a new probability measure $\widehat{Q}$ such that a drift changed P-Brownian motion (that is not a Brownian motion under the probability measure P anymore) is a Brownian motion under the new measure $\widehat{Q}$. Before we state the Girsanov Theorem we make

Definition 3.40 (Stochastic Exponential)**.** *Let* X *be a stochastic process. The unique strong solution* $Z = \mathcal{E}(X)$ *of*

$$dZ_t = Z_t\, dX_t, \quad Z_0 = 1 \tag{3.12}$$

is called stochastic exponential of X*.*

From Theorem 3.34 we get immediately that (3.12) has a unique strong solution.

Theorem 3.41 (Matrix Variate Girsanov Theorem)**.** *Let* $T > 0$*,* $B \sim \mathcal{BM}_p$ *and* U *be an adapted, continuous stochastic process with values in* $\mathcal{M}_p(\mathbb{R})$ *such that*

$$\left(\mathcal{E}\left(\operatorname{tr}\left(-\int_0^t U_s^{\mathrm{T}}\, dB_s\right)\right)\right)_{t\in[0,T]} \tag{3.13}$$

is a martingale, or, what is sufficient for (3.13), but not necessary, such that the Novikov condition *is satisfied*

$$E\left(\operatorname{etr}\left(\frac{1}{2}\int_0^T U_t^{\mathrm{T}} U_t\, dt\right)\right) < \infty \tag{3.14}$$

Then

$$\widehat{Q} = \int \mathcal{E}\left(\operatorname{tr}\left(-\int_0^T U_t^{\mathrm{T}}\, dB_t\right)\right) dP \tag{3.15}$$

is an equivalent probability measure, and

$$\widehat{B_t} = \int_0^t U_s\, ds + B_t \tag{3.16}$$

is a $\widehat{Q}$*-Brownian motion on* $[0,T)$*.*

Proof. We use the multivariate Girsanov Theorem of Kallenberg [1997, Corollary 16.25]. The process $V := \operatorname{vec}(-U) \in \mathbb{R}^{p^2}$ is according to Revuz and Yor [2001, Proposition 4.8, Chapter I] progressively measurable and $B^v := \operatorname{vec}(B)$ is a Brownian motion on $\mathbb{R}^{p^2}$. The Novikov condition is then

$$\begin{aligned} E\left(exp\left(\frac{1}{2}\int_0^T \sum_{i=1}^{p^2} V_{t,i}^2\, dt\right)\right) &= E\left(exp\left(\frac{1}{2}\int_0^T \operatorname{tr}(U_t^{\mathrm{T}} U_t)\, dt\right)\right) \\ &= E\left(\operatorname{etr}\left(\frac{1}{2}\int_0^T U_t^{\mathrm{T}} U_t\, dt\right)\right) < \infty \end{aligned}$$

With Kallenberg [1997, Corollary 16.25] the new measure Q is then

$$\widehat{Q} = \int \mathcal{E}\left(\int_0^T V_t^{\mathrm{T}}\, dB_t^v\right) dP = \int \mathcal{E}\left(\mathrm{tr}\left(-\int_0^T U_t^{\mathrm{T}}\, dB_t\right)\right) dP$$

and

$$\widetilde{B_t^v} = B_t^v - \int_0^t V_s\, ds$$

is a $\widehat{Q}$-Brownian motion for $t \in [0, T]$ and with values in $\mathbb{R}^{p^2}$. Hence,

$$\widehat{B_t} = \mathrm{vec}^{-1}(\widetilde{B_t^v}) = \int_0^t U_s\, ds + B_t$$

is a $\widehat{Q}$-Brownian motion for $t \in [0, T]$ and with values in $\mathcal{M}_p(\mathbb{R})$. □

3.4 McKean's Argument

Lévy's Theorem gives us an easy way to decide whether a continuous local martingale is a Brownian motion.

Theorem 3.42 (Lévy's Theorem).
Let B be a $p \times p$ dimensional continuous local martingale such that

$$[B_{ij}, B_{kl}]_t = \begin{Bmatrix} t, & \text{if } i = k \text{ and } j = l \\ 0, & \text{else} \end{Bmatrix}$$

for all $i, j, k \in \{1, \dots, p\}$. Then B is a $p \times p$ dimensional Brownian motion, $B \sim \mathcal{BM}_p$.

Proof. It suffices to show that $X := \mathrm{vec}(B)$ is a p^2-dimensional Brownian motion.
Let $a, b \in \{1, \dots, p^2\}$. Then, there exist $i, j, k, l \in \{1, \dots, p\}$ such that $a = p(j-1) + i$ and $b = p(l-1) + k$. Clearly, $i = k$ and $j = l$ imply $a = b$, and conversely, if we assume $a = b$, then $i \neq k$ or resp. $j \neq l$ imply contradictions. Thus,

$$[X_a, X_b] = [B_{ij}, B_{kl}] = t\mathbf{1}_{\{i=k\}}\mathbf{1}_{\{j=l\}} = t\mathbf{1}_{\{a=b\}}$$

With Protter [2004, Theorem 40,Chapter 2] we can conclude that X is a p^2-dimensional Brownian motion. □

For every finite stopping time τ_0 we say M is a local martingale on the (stochastic) interval $[0, \tau_0]$ if the stopped process $(M_{\inf\{t,\tau_0\}})_{t\in\mathbb{R}_+}$ is a local martingale. The next Theorem shows us that every stopped local martingale is a stopped, time-changed Brownian motion. There also exists a version of Theorem 3.43 without stopping times but with the additional assumption that $\lim_{t\to\infty}[M, M]_t = \infty$, see the Dambis-Dubins-Schwarz-Theorem in Revuz and Yor [2001, Theorem 1.6, Chapter V].

Theorem 3.43. *Let $\tau_0 < \infty$ be a stopping time and M be a continuous local martingale on the interval $[0, \tau_0]$. If we set*

$$T_t := \inf\{s : [M,M]_s > t\} \text{ (with the convention } \inf\{\emptyset\} = \infty)$$

then $B_t := M_{T_t}$ is a stopped $\mathcal{F}_{T_t}$-Brownian motion on the interval $[0, [M,M]_{\tau_0})$, i.e B is a Brownian motion w.r.t. the σ-algebra

$$\mathcal{F}_{T_t} = \{A \in \mathcal{F} : A \cap \{T_t \leq t\} \in \mathcal{F}_t \quad t \in \mathbb{R}_+\}$$

that is stopped at $[M,M]_{\tau_0}$. Furthermore, it holds that $M_t = B_{[M,M]_t}$, i.e. M is a stopped, time-changed Brownian motion.

Proof. Observe that T_t is a generalized inverse function of $[M,M]_t$. Because M is continuous, and hence $[M,M]$, too, it holds that $[M,M]_{T_t} = t$, but still we only have $T_{[M,M]_t} \geq t$ in general (and strict equality for every t if and only if M was strict monotonic on the entire interval). See Revuz and Yor [2001, p.7-8] for details.
The set $\{T_t\}_{t \in \mathbb{R}_+}$ is a family of stopping times, because M is adapted. Observe that

$$[M,M]_{\tau_0} > t \Rightarrow T_t \leq \tau_0$$

Hence $T_t < \infty$ as $\tau_0 < \infty$. From Revuz and Yor [2001, Proposition 4.8 and 4.9, Chapter I] we know that the stopped process M_{T_t} is $\mathcal{F}_{T_t}$-measurable. Furthermore,

$$[B,B]_t = [M,M]_{T_t} = t$$

as mentioned above. With Theorem 3.42 we can can conclude that B is a stopped $\mathcal{F}_{T_t}$-Brownian motion on $[0, [M,M]_{\tau_0})$.
To prove that M is a time-changed Brownian motion, observe that $T_{[M,M]_t} > t$ if and only if $[M,M]$ is constant on $[t, T_{[M,M]_t}]$. As M and $[M,M]$ are constant on the same intervals, we have $M_{T_{[M,M]_t}} = M_t$ and thus $B_{[M,M]_t} = M_t$. □

With the foregoing Theorem we can form an argument that will allow us in turn to proof some result on the existence of the Wishart process in the next chapter.

Theorem 3.44. *(McKean's Argument, cf. McKean [1969, p.47,Problem 7])*
Let r be a real-valued stochastic process such that $P(r_0 > 0) = 1$ and $h : \mathbb{R}_+ \to \mathbb{R}$ such that

- *$h(r)$ is a continuous local martingale on the interval $[0, \tau_0)$ for $\tau_0 := \inf\{s : r_s = 0\}$*
- $\lim_{x \to 0, x>0} h(x) = \infty$ *or resp.* $\lim_{x \to 0, x>0} h(x) = -\infty$

Then $\tau_0 = \infty$ almost surely, i.e. $r_t(\omega) > 0 \quad \forall t \in \mathbb{R}_+$ for almost all ω.

Remark 3.45. *If $h(r_t) = \int_0^t f(s, r_s)\, dW_s$ with a Brownian motion W, then for $h(r)$ being a continuous local martingale on $[0, \tau_0)$ it is sufficient that f is square integrable on $[0, \tau_0)$.*

Proof. Suppose $\tau_0 < \infty$.
Define $M := h(r)$ and T_t as in Theorem 3.43. M is a continuous local martingale on $[0, \tau_0]$ and with Theorem 3.43 we can conclude that $B_t := M_{T_t}$ is a stopped Brownian motion on $[\,0, [M, M]_{\tau_0})$. Observe, that $r_0 > 0$ a.s. implies $\tau_0 > 0$.
Consider the case $\lim_{x\to 0, x>0} h(x) = \infty$. On the interval $[0, \tau_0)$, the process $h(r_t)$ takes every value in $[h(r_0), \infty)$, especially $[h(r), h(r)]_{\tau_0} = [M, M]_{\tau_0} > 0$. For $\tau_0 < \infty$ we have

$$t \to [M, M]_{\tau_0} \Rightarrow T_t \to \tau_0 \Rightarrow r_{T_t} \to 0 \Rightarrow B_t = h(r_{T_t}) \to \infty$$

There are two cases:

- $[M, M]_{\tau_0} < \infty$, what is impossible because a Brownian motion can not go to infinity in finite time (and without infinite oscillations), see Revuz and Yor [2001, Law of the iterated logarithm, Corollary 1.12, Chapter II].
- $[M, M]_{\tau_0} = \infty$, which implies that B is a Brownian motion on $\mathbb{R}_+$ that converges to ∞ almost surely and that is a contradiction because,

$$P(B_t \leq 0 \text{ infinitely often for } t \to \infty) = 1$$

As both cases imply contradictions, we conclude $\tau_0 = \infty$.
The case for $\lim_{x\to 0, x>0} h(x) = -\infty$ is analogous. □

3.5 Ornstein-Uhlenbeck Processes

We give the following definition according to Bru [1991] such that we are later able to show that some solutions of the Wishart SDE can be constructed out of a matrix variate Ornstein-Uhlenbeck process.

Definition 3.46 (Matrix Variate Ornstein-Uhlenbeck Process)**.**
Let $A, B \in \mathcal{M}_p(\mathbb{R})$, $x_0 \in \mathcal{M}_{n,p}(\mathbb{R})$ a.s. and $W \sim \mathcal{BM}_{n,p}$. A solution X of

$$dX_t = X_t B\, dt + dW_t\, A, \quad X_0 = x_0 \tag{3.17}$$

is called $n \times p$-dimensional Ornstein-Uhlenbeck process. We write $X \sim \mathcal{OUP}_{n,p}(A, B, x_0)$.

As $X \mapsto XB$ and $X \mapsto A$ are trivially globally Lipschitz and satisfy the linear growth condition we know from Theorem 3.34 that (3.17) has a unique strong solution on the entire interval $[0, \infty)$. Luckily, we are even able to give an explicit formula for the solution of (3.17).

Theorem 3.47 (Existence and Uniqueness of the Ornstein-Uhlenbeck Process)**.**
For a Brownian motion $W \sim \mathcal{BM}_{n,p}$, the unique strong solution of (3.17) is given by

$$X_t = x_0 e^{Bt} + \left(\int_0^t dW_s\, A e^{-Bs} \right) e^{Bt} \tag{3.18}$$

Proof.

$$
\begin{aligned}
dX_t &= d(x_0 e^{Bt}) + d\left(\left(\left(\int_0^t dW_s\, Ae^{-Bs}\right) e^{Bt}\right)\right) \\
&= x_0 e^{Bt} B\, dt + dW_t\, Ae^{-Bt} e^{Bt} + \left(\int_0^t dW_s\, Ae^{-Bs}\right) e^{Bt} B\, dt + \underbrace{d\left[\int_0^{\cdot} dW_s A e^{-Bs}, e^{B\cdot}\right]_t^M}_{=0} \\
&= \left(x_0 e^{Bt} + \left(\int_0^t dW_s\, Ae^{-Bs}\right) e^{Bt}\right) B\, dt + dW_t\, A \\
&= X_t B\, dt + dW_t\, A
\end{aligned}
$$

Furthermore, (3.18) is a strong solution by construction. □

The following Lemma will be useful for determining the conditional distribution of the Ornstein-Uhlenbeck process.

Lemma 3.48. *Let $W \sim \mathcal{BM}_{n,p}$ and $X : \mathbb{R}_+ \to \mathcal{M}_{p,m}(\mathbb{R})$, $t \mapsto X_t$ be a square integrable, deterministic function. Then*

$$\int_0^t dW_s\, X_s \sim \mathcal{N}_{n,m}(0, I_n \otimes \int_0^t X_s^{\mathrm{T}} X_s\, ds)$$

Proof. $H_t := \int_0^t dW_s\, X_s$. $H_t \in \mathcal{M}_{n,m}(\mathbb{R})$ is a $\mathcal{L}^2$-limit of normal distributed, independent, random variables and therefore normally distributed (Cf. Øksendal [2000, Proof of Theorem 5.2.1 and Theorem A.7.]). Furthermore, H_t is a martingale with a.s. initial value zero and thus expectation zero.
$\mathrm{cov}(H_t, H_t)$ is a block diagonal matrix, because different rows of W are independent. Hence, $\mathrm{cov}(H_t, H_t)$ is a tensor product of I_n with another m-dimensional matrix. Furthermore, all rows of W are identically distributed such that we only need to consider row 1 of H_t, i.e. the first block matrix D_t^1 of $\mathrm{cov}(H_t, H_t)$. We have that

$$H_{t,1i} = \sum_{k=1}^{p} \int_0^t dW_{s,1k}\, X_{s,ki}$$

and thus

$$
\begin{aligned}
D_{t,ij}^1 &= \mathrm{cov}(H_{t,1i}, H_{t,1j}) \\
&= \sum_{k_i,k_j=1}^{p} \mathrm{cov}(\int_0^t dW_{s,1k_i}\, X_{s,k_i i}, \int_0^t dW_{s,1k_j}\, X_{s,k_j j}) \\
&= \sum_{k_i,k_j=1}^{p} \int_0^t X_{s,k_i i} X_{s,k_j j}\, ds \mathbf{1}_{\{k_i=k_j\}} \\
&= \sum_{k=1}^{p} \int_0^t X_{s,ki} X_{s,kj}\, ds = \int_0^t (X_s^{\mathrm{T}} X_s)_{ij}\, ds
\end{aligned}
$$

where we used the fact that $\mathrm{cov}(\int f_1(s)\, dW_{s,1k_i}, \int f_2(s)\, dW_{s,1k_j}) = \int f_1(s) f_2(s)\, ds\, \mathbf{1}_{\{k_i=k_j\}}$ for deterministic, square integrable real functions f_1, f_2. □

Now we are able to show that the (matrix variate) Ornstein-Uhlenbeck process is (matrix variate) normally distributed. Later we will see that some Wishart processes are therefore 'square' normally distributed, that is by Lemma 3.20 a noncentral Wishart distribution.
To clarify terms, with *distribution of a stochastic process* X we always mean from now on the distribution of X_t at time t conditional on the initial value X_0. Although it may be confusing, the terms distribution of a stochastic process and law of a stochastic process have a different meaning in this thesis.

Theorem 3.49 (Distribution of the Matrix Variate Ornstein-Uhlenbeck Process)**.**
Let $X \sim \mathcal{OUP}_{n,p}(A, B, x_0)$ *with a matrix* $B \in \mathcal{M}_p(\mathbb{R})$ *that satisfies* $0 \notin -\sigma(B) - \sigma(B)$*. Then the distribution of* X *is given by*

$$X_t|x_0 \sim \mathcal{N}_{n,p}(x_0 e^{Bt}, I_n \otimes [\mathcal{A}^{-1}(A^{\mathrm{T}}A) - \mathcal{A}^{-1}(e^{B^{\mathrm{T}}t}A^{\mathrm{T}}Ae^{Bt})])$$

where $\mathcal{A}^{-1}$ *is the inverse of the linear operator* $\mathcal{A} : \mathcal{S}_p \to \mathcal{S}_p, X \mapsto -B^{\mathrm{T}}X - XB$*.*

Proof. Define $H_t = \int_0^t dW_s A e^{-Bs}$. Lemma 3.48 shows us that

$$H_t \sim \mathcal{N}_{n,p}(0, I_n \otimes \int_0^t e^{-B^{\mathrm{T}}s}A^{\mathrm{T}}Ae^{-Bs}\, ds)$$

Define the operator $\mathcal{A} : X \mapsto -B^{\mathrm{T}}X - XB$ as in Lemma 2.12. Then we have

$$\frac{d}{ds}(e^{-B^{\mathrm{T}}s}A^{\mathrm{T}}Ae^{-Bs}) = \mathcal{A}(e^{-B^{\mathrm{T}}s}A^{\mathrm{T}}Ae^{-Bs})$$

and hence

$$\int_0^t e^{-B^{\mathrm{T}}s}A^{\mathrm{T}}Ae^{-Bs}\, ds = \mathcal{A}^{-1}(e^{-B^{\mathrm{T}}s}A^{\mathrm{T}}Ae^{-Bs})|_{s=0}^{t} \tag{3.19}$$

The equality $X_t = x_0 e^{Bt} + H_t e^{Bt}$ and Theorem 3.15 gives us

$$X_t|x_0 \sim \mathcal{N}_{n,p}(x_0 e^{Bt}, I_n \otimes e^{B^{\mathrm{T}}t}\mathcal{A}^{-1}(e^{-B^{\mathrm{T}}s}A^{\mathrm{T}}Ae^{-Bs})|_{s=0}^{t} e^{Bt})$$

Because $\mathcal{A}$ is a linear operator, $\mathcal{A}^{-1}$ is also linear and we can simplify

$$\mathcal{A}^{-1}(e^{-B^{\mathrm{T}}s}A^{\mathrm{T}}Ae^{-Bs})|_{s=0}^{t} = \mathcal{A}^{-1}(e^{-B^{\mathrm{T}}t}A^{\mathrm{T}}Ae^{-Bt}) - \mathcal{A}^{-1}(A^{\mathrm{T}}A)$$

and, because $\mathcal{A}^{-1}$ is the integral from (3.19),

$$e^{B^{\mathrm{T}}t}(\mathcal{A}^{-1}(e^{-B^{\mathrm{T}}t}A^{\mathrm{T}}Ae^{-Bt}) - \mathcal{A}^{-1}(A^{\mathrm{T}}A))e^{Bt} = \mathcal{A}^{-1}(A^{\mathrm{T}}A) - \mathcal{A}^{-1}(e^{B^{\mathrm{T}}t}A^{\mathrm{T}}Ae^{Bt})$$

□

Theorem 3.50 (Stationary Distribution of the Ornstein-Uhlenbeck Process)**.**
Let $X \sim \mathcal{OUP}_{n,p}(A, B, x_0)$ *with a matrix* $B \in \mathcal{M}_p(\mathbb{R})$ *such that all eigenvalues of* B *have a negative real part, i.e.* $\mathrm{Re}(\sigma(B)) \subseteq (-\infty, 0)$*.*
Then the Ornstein-Uhlenbeck process X *has a stationary limiting distribution, that is*

$$\mathcal{N}_{n,p}\left(0, I_n \otimes \mathcal{A}^{-1}(A^{\mathrm{T}}A)\right)$$

Proof. Observe that $\mathrm{Re}(\sigma(B)) \subseteq (-\infty, 0)$ implies $0 \notin -\sigma(B) - \sigma(B)$ and we can use the foregoing Theorem. With Theorem 3.14 we get for $t \in (0, \infty)$ that X_t given x_0 has characteristic function

$$\widehat{P^{X_t}}(Z) = \mathrm{etr}\left(iZ^{\mathrm{T}} x_0 e^{Bt} - \frac{1}{2} Z^{\mathrm{T}} Z \Psi_t\right) \tag{3.20}$$

with

$$\Psi_t = \mathcal{A}^{-1}(A^{\mathrm{T}} A) - \mathcal{A}^{-1}(e^{B^{\mathrm{T}} t} A^{\mathrm{T}} A e^{Bt}) \tag{3.21}$$

Now, we write the matrix B in its *Jordan normal form*:
There exists a matrix $T \in GL(p)$ such that

$$B = T \begin{pmatrix} J_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_m \end{pmatrix} T^{-1}$$

with block matrices, called Jordan blocks, $J_i \in \mathcal{M}_{l_i}(\mathbb{R})$ such that $l_1 + \ldots + l_m = p$. With the calculation rules for the matrix exponential we get

$$\exp(Bt) = T \exp\left(\begin{pmatrix} J_1 t & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_m t \end{pmatrix}\right) T^{-1} = T \mathrm{diag}(e^{J_1 t}, \ldots, e^{J_m t}) T^{-1} \tag{3.22}$$

Without loss of generality, consider only the first Jordan block

$$J := J_1 = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & \lambda \end{pmatrix} \in \mathcal{M}_l(\mathbb{R})$$

with $l := l_1$, $\lambda = a + bi$ an eigenvalue of B with $a < 0$.
We can separate J into a sum of two matrices, $J = D + N$, with a diagonal matrix $D = \mathrm{diag}(\lambda, \ldots, \lambda) \in \mathcal{M}_l(\mathbb{R})$ and a matrix

$$N = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix} \in \mathcal{M}_l(\mathbb{R})$$

Because D and N commute, $DN = ND$, we have

$$\exp(Jt) = \exp(Dt + Nt) = \exp(Dt) \exp(Nt)$$

and thus

$$\exp(Jt) = \operatorname{diag}(e^{\lambda t}, \dots, e^{\lambda t}) \exp(Nt) = e^{\lambda t} \exp(Nt) = e^{ibt} e^{at} \exp(Nt)$$

The matrix N is nilpotent, as $N^l = 0$. Hence, $\exp(Nt)$ is the finite sum

$$\exp(Nt) = \sum_{k=0}^{l-1} \frac{1}{k!} N^k = \begin{pmatrix} 1 & t & \dots & \dots & \frac{t^{l-1}}{(l-1)!} \\ 0 & 1 & t & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & t \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} \in \mathcal{M}_l(\mathbb{R})$$

For $a < 0$ and $t \to \infty$ we know that $e^{at} p_t$ converges to zero for every polynomial p_t in t. Hence, $\lim_{t\to\infty} e^{at} \exp(Nt) = 0$. As e^{ibt} always has absolute value one, we also have that $\lim_{t\to\infty} \exp(Jt) = 0$. With (3.22) it is now obvious that

$$\lim_{t\to\infty} \exp(Bt) = 0$$

The next step is to observe that $\mathcal{S}_p$ is a finite dimensional and normed space with the norm induced by the inner product trace. Hence, the linear operator $\mathcal{A}^{-1} : \mathcal{S}_p \to \mathcal{S}_p$ is continuous and with the above can conclude that Ψ_t from (3.21) converges pointwise to

$$\lim_{t\to\infty} \Psi_t = \mathcal{A}^{-1}(A^{\mathrm{T}} A)$$

and thus

$$\lim_{t\to\infty} \widehat{P^{X_t}}(Z) = \operatorname{etr}\left(-\frac{1}{2} Z^{\mathrm{T}} Z\, \mathcal{A}^{-1}(A^{\mathrm{T}} A)\right) := f(Z) \tag{3.23}$$

Clearly, f is continuous at $Z = 0$. With Lévy's Continuity Theorem we know that there exists a probability measure μ on $\mathcal{M}_p(\mathbb{R})$ such that $f(Z) = \widehat{\mu}(Z)$, and $P^{X_t} \overset{weak}{\to} \mu$.
The function f is the characteristic function of a normal distributed random matrix. That implies by Theorem 3.14 that X_t converges in distribution to a random matrix with normal distribution μ,

$$\mu = \mathcal{N}_{n,p}(0, I_n \otimes \mathcal{A}^{-1}(A^{\mathrm{T}} A)) \tag{3.24}$$

and the limit is independent of the initial value x_0. Thus, and by the fact that X is a Markov process according to Theorem 3.39, the limit distribution in (3.24) is a stationary distribution. □

Corollary 3.51. *Let $X \sim \mathcal{OUP}_{n,p}(A, B, x_0)$ with a matrix $B \in \mathcal{M}_p(\mathbb{R})$ such that $B \in \mathcal{S}_p^-$ and $A^{\mathrm{T}} A$ and B commute. Then the stationary distribution of X is given by*

$$\mathcal{N}_{n,p}\left(0, I_n \otimes -\frac{1}{2} A^{\mathrm{T}} A B^{-1}\right)$$

Proof.

$$\mathcal{A}^{-1}(A^{\mathrm{T}} A) = -\frac{1}{2} A^{\mathrm{T}} A B^{-1} \Leftrightarrow A^{\mathrm{T}} A = \mathcal{A}(-\frac{1}{2} A^{\mathrm{T}} A B^{-1}) = \frac{1}{2}(B A^{\mathrm{T}} A B^{-1} + A^{\mathrm{T}} A) = A^{\mathrm{T}} A$$

□

Chapter 4

Wishart Processes

In this chapter the main work on the theory of Wishart processes will be done. As an introduction we begin with the one dimensional case, called the square Bessel process. Afterwards, we focus on the general Wishart process and give a theorem about the existence and uniqueness of Wishart processes at the end of section 4.2. In section 4.3, we show that some Wishart processes can be expressed as matrix squares of the matrix variate Ornstein-Uhlenbeck process from section 3.5. Eventually, in section 4.4 we give an algorithm to simulate Wishart processes by using an Euler approximation.

4.1 The one dimensional case: Square Bessel Processes

The square Bessel process is the one-dimensional case of the Wishart process - in fact, even a special case of the one-dimensional Wishart process. We consider it separately, because it is much more elementary than for higher dimensions.

Definition 4.1 (Square Bessel Process)**.** *Let $\alpha \geq 0$, β be a one-dimensional Brownian motion and $x_0 \geq 0$. A strong solution of*

$$dX_t = 2\sqrt{X_t}\, d\beta_t + \alpha\, dt, \quad X_0 = x_0 \tag{4.1}$$

is called a square Bessel process with parameter α and denoted by $X \sim \mathrm{BESQ}(\alpha, x_0)$.

Remark 4.2. *Let B denote a Brownian motion on $\mathbb{R}^n$ (that is an $\mathbb{R}^n$-vector of n independent one-dimensional Brownian motions). The Bessel process is then the Euclidean norm of B, that is $\sqrt{B^{\mathrm{T}}B}$. Hence, the process X with $X_t = B_t^{\mathrm{T}} B_t$ is often called square Bessel process in the literature. Using Theorem 4.9 and Theorem 3.30 one can show that $X \sim \mathrm{BESQ}(n, x_0)$. Hence, using (4.1) for the definition of the square Bessel process is more general as it allows $n = \alpha$ to be any real, non-negative number.*

Before we give a comprehensive theorem about existence, uniqueness and non-negativity of the square Bessel process, we state the following results for one-dimensional stochastic differential equations:

Theorem 4.3. *(Pathwise Uniqueness, cf. Yamada and Watanabe [1971a, Theorem 1]) Let*

$$dX_t = \sigma(X_t)\, d\beta_t + b(X_t)\, dt \tag{4.2}$$

be a one-dimensional stochastic differential equation where $b, \sigma : \mathbb{R} \to \mathbb{R}$ are continuous functions and β a one-dimensional Brownian motion. Suppose there exist increasing functions $\rho, \kappa : (0,\infty) \to (0,\infty)$ such that

$$|\sigma(\xi) - \sigma(\eta)| \leq \rho(|\xi - \eta|) \quad \forall\, \xi, \eta \in \mathbb{R} \\ \text{with} \int_0^1 \rho^{-2}(u)\, du = \infty \tag{4.3}$$

and

$$|b(\xi) - b(\eta)| \leq \kappa(|\xi - \eta|) \quad \forall\, \xi, \eta \in \mathbb{R} \\ \text{with} \int_0^1 \kappa^{-1}(u)\, du = \infty \tag{4.4}$$

Then pathwise uniqueness holds for (4.2).

Theorem 4.4. *(Comparison Theorem, Revuz and Yor [2001, Chapter IX, Theorem 3.7]) Consider two stochastic differential equations*

$$dX_t^1 = \sigma(X_t)\, d\beta_t + b^1(X_t)\, dt, \quad dX_t^2 = \sigma(X_t)\, d\beta_t + b^2(X_t)\, dt$$

that both fulfill pathwise uniqueness and let b^1,b^2 be two bounded Borel functions such that $b^1 \geq b^2$ everywhere and one of them is globally Lipschitz. If (X^1, β) is a solution of the first, and (X^2, β) a solution of the second stochastic differential equation (X^1,X^2 defined on the same probability space), w.r.t the same Brownian motion β and if $X_0^1 \geq X_0^2$ a.s., then

$$P[X_t^1 \geq X_t^2 \text{ for all } t \in \mathbb{R}_+] = 1 \tag{4.5}$$

Theorem 4.5 (Existence, Uniqueness and Non-Negativity of the Square Bessel Process)**.** *For $\alpha \geq 0$ and $x_0 \geq 0$ there exists an unique strong and non-negative solution X of*

$$X_t = x_0 + 2\int_0^t \sqrt{X_s}\, d\beta_s + \alpha t \tag{4.6}$$

on the entire interval $[0, \infty)$.
Moreover, if $\alpha \geq 2$ and $x_0 > 0$ the solution X is positive a.s. on $[0, \infty)$.

Proof. We show with Theorem 4.3 that pathwise uniqueness holds for the stochastic differential equation

$$dX_t = 2\sqrt{|X_t|}\, d\beta_t + \alpha\, dt, \quad X_0 = x_0 \tag{4.7}$$

like in Revuz and Yor [2001, p. 439]. Since $|\sqrt{z} - \sqrt{z'}| \leq \sqrt{|z - z'|}$ for all $z, z' \geq 0$, we have that

$$2\left|\sqrt{|\xi|} - \sqrt{|\eta|}\right| \leq 2\sqrt{|\,|\xi| - |\eta|\,|} \leq 2\sqrt{|\xi - \eta|} = \rho(|\xi - \eta|) \quad \forall\, \xi, \eta \in \mathbb{R}$$

with $\rho(u) = 2\sqrt{u}$. Clearly,

$$\int_0^1 \rho^{-2}(u)\,du = \frac{1}{4}\int_0^1 \frac{1}{u}\,du = \infty \tag{4.8}$$

With Remark 3.33 we conclude that there exists a unique strong solution for (4.7).
Next, we show that our solution never becomes negative: First, consider the case $x_0^1 = 0$ and $\alpha^1 = 0$. Obviously, $X^1 \equiv 0$ is our unique strong solution in this case. Secondly, consider an arbitrary $x_0^2 \geq 0$ and $\alpha^2 \geq 0$. From the above, we have a unique strong solution X^2. The comparison theorem implies that

$$P[X_t^2 \geq X_t^1 \textit{ for all } t \in \mathbb{R}_+] = P[X_t^2 \geq 0 \textit{ for all } t \in \mathbb{R}_+] = 1$$

and thus X_t^2 is non-negative for all t almost surely. Hence, we can discard the $|.|$ in (4.7) and X^2 also is the unique strong solution of (4.6).
Finally, Revuz and Yor [2001, Proposition 1.5, Chapter XI] has shown that for $\alpha \geq 2$, the set $\{0\}$ is polar, i.e. $P(\inf\{s : X_s = 0\} < \infty) = 0$ for all initial values $x_0 > 0$. Hence, in this case the unique strong solution X is positive for all $t \in \mathbb{R}_+$. □

Remark 4.6. *For $\alpha \geq 2$, one could also use McKean's argument (Theorem 3.44) to show that the square Bessel process is positive, analogously to the Proof of Theorem 4.14.*

4.2 General Wishart Processes: Definition and Existence Theorems

Definition 4.7 (Wishart Processes)**.** *Let B be a $p \times p$-dimensional Brownian motion, $B \sim \mathcal{BM}_p$, $Q \in \mathcal{M}_p(\mathbb{R})$ and $K \in \mathcal{M}_p(\mathbb{R})$ be arbitrary matrices, $s_0 \in \overline{\mathcal{S}_p^+}$ the initial value and $\alpha \geq 0$ a non-negative number. Then we call the stochastic differential equation*

$$dS_t = \sqrt{S_t}\,dB_t\,Q + Q^{\mathrm{T}} dB_t^{\mathrm{T}} \sqrt{S_t} + (S_t K + K^{\mathrm{T}} S_t + \alpha Q^{\mathrm{T}} Q)\,dt, \quad S_0 = s_0 \tag{4.9}$$

the Wishart SDE.
A strong solution S of (4.9) in $\overline{\mathcal{S}_p^+}$ is said to be a ($p \times p$-dimensional) Wishart process *with parameters Q, K, α, s_0, written $S \sim \mathcal{WP}_p(Q, K, \alpha, s_0)$.*

To understand (4.9) intuitively, it may help to have a look at a approximation of the form

$$S_{t+h} \approx S_t + \sqrt{S_t}\int_t^{t+h} dB_s\,Q + \int_t^{t+h} Q^{\mathrm{T}} dB_s^{\mathrm{T}} \sqrt{S_t} + (S_t K + K^{\mathrm{T}} S_t + \alpha Q^{\mathrm{T}} Q)h$$

Here we can see that Q controls the covariance of our normal distributed fluctuations, and that this fluctuations are proportional to the square root of our process:

$$\sqrt{S_t}\int_t^{t+h} dB_s\,Q | S_t \sim \sqrt{S_t} \cdot \mathcal{N}_{p,p}(0, I_p \otimes Q^{\mathrm{T}} Q h)$$

Hence, the fluctuations decrease quickly if our process goes to zero.
As $\alpha Q^{\mathrm{T}}Q$ is positive semidefinite, the parameter α determines how large the drift away from zero is.
Just like in the one-dimensional case, the solution of (4.9) has a *mean reverting* property. To understand that, we consider the deterministic equivalent of (4.9), that is the ordinary differential equation

$$\frac{dS_t}{dt} = S_t K + K^{\mathrm{T}} S_t + \alpha Q^{\mathrm{T}} Q, \quad S_0 = s_0$$

To simplify the notation, we define the operator $\mathcal{C} : \mathcal{S}_p \to \mathcal{S}_p, X \mapsto XK + K^{\mathrm{T}}X$ and set $M = \alpha Q^{\mathrm{T}}Q$. Then

$$\frac{dS_t}{dt} = \mathcal{C}S_t + M, \quad S_0 = s_0 \tag{4.10}$$

From the theory of ordinary linear differential equations (see Timmann [2005, p. 83], for example), we know that the solution of (4.10) is given by

$$S_t = e^{\mathcal{C}t}\left(s_0 + \int_0^t e^{-\mathcal{C}u}\, duM\right) \tag{4.11}$$

Indeed, by partial integration follows

$$\frac{dS_t}{dt} = \mathcal{C}\underbrace{\left(e^{\mathcal{C}t}\left(s_0 + \int_0^t e^{-\mathcal{C}u}\, duM\right)\right)}_{=S_t} + \underbrace{e^{\mathcal{C}t}e^{-\mathcal{C}t}}_{=I_p} M$$

Evaluating the integral in (4.11) gives

$$S_t = e^{\mathcal{C}t}((s_0 + \mathcal{C}^{-1}M) - \mathcal{C}^{-1}M \tag{4.12}$$

If the eigenvalues of K only have negative real parts, $Re(\sigma(K)) \subseteq (-\infty, 0)$, then, because of $\sigma(\mathcal{C}) = \sigma(K) + \sigma(K)$, we also have $Re(\sigma(\mathcal{C})) \subseteq (-\infty, 0)$. From the proof of Theorem 3.50 we get that $\lim_{t\to\infty} e^{\mathcal{C}t} = 0$ and thus from (4.12)

$$\lim_{t\to\infty} S_t = -\mathcal{C}^{-1}M$$

Hence, the deterministic solution of (4.10) converges to $-\mathcal{C}^{-1}M$. Thus, the stochastic solution of (4.9) will fluctuate around $-\mathcal{C}^{-1}M$.

Considering the fact that the fluctuations decrease quickly if our processes goes to zero and that we have a non-negative definite drift $\alpha Q^{\mathrm{T}}Q$ away from zero one may think that this process never leaves $\overline{\mathcal{S}_p^+}$. As we have shown in the last section, for the one dimensional case (square Bessel process) we have a unique strong solution for all t that never becomes negative. Unlike in this case, we can not show for $p \geq 2$ that there still exists a unique strong solution after the process hits the boundary of $\mathcal{S}_p^+$ (that means $S_t \in \overline{\mathcal{S}_p^+}\backslash\mathcal{S}_p^+$) the first time. We are only able to give sufficient conditions such that the process stays in the set of all positive definite matrices, and then we have a unique strong solution for all t.

One reason why we cannot transfer the proof of Theorem 4.5 to the matrix variate case is that Theorem 4.3 cannot be generalized to matrix variate stochastic differential equations. In the matrix variate version of Theorem 4.3 as stated in Yamada and Watanabe [1971b, Theorem1], the constraint (4.3) turns to

$$\int_{\{U\in\overline{\mathcal{S}_p^+}:||U||\leq 1\}} \rho^{-2}(U)U\,dU = \infty \tag{4.13}$$

$$U \mapsto \rho^2(U)U^{-1} \text{ is concave} \tag{4.14}$$

but

$$\int_{\{U\in\overline{\mathcal{S}_p^+}:||U||\leq 1\}} \rho^{-2}(U)U\,dU = \int_{\{U\in\overline{\mathcal{S}_p^+}:||U||\leq 1\}} I_p\,dU < \infty$$

for $\rho(U) = \sqrt{U}$. Hence the theorem can not be applied to our case.
Furthermore, Yamada and Watanabe [1971b, Remark 2] have also shown that (4.13) is, for $p \geq 3$, nearly best possible in the sense that, if $\int_{U\in\overline{\mathcal{S}_p^+}:||U||\leq 1} \rho^{-2}(U)U\,du < \infty$ and ρ is subadditive then, a stochastic differential equation can be constructed that has two solutions, and thus pathwise uniqueness cannot hold. But the fact the we cannot use Yamada and Watanabe [1971b, Theorem1] does not give any evidence whether pathwise uniqueness holds for the Wishart SDE or not.

Before we can begin our mathematical analysis of (4.9), we need a few auxiliary results:

Lemma 4.8 (Quadratic Variation of the Wishart Process)**.** *Let $S \sim \mathcal{WP}_p(Q,K,\alpha,s_0)$. Then*

$$d[S_{ij},S_{kl}]_t = S_{t,ik}(Q^{\mathrm{T}}Q)_{jl}\,dt + S_{t,il}(Q^{\mathrm{T}}Q)_{jk}\,dt + S_{t,jk}(Q^{\mathrm{T}}Q)_{il}\,dt + S_{t,jl}(Q^{\mathrm{T}}Q)_{ik}\,dt$$

as long as S exists.

Proof. Define $H_t := \sqrt{S_t}dB_tQ + Q^{\mathrm{T}}dB_t^{\mathrm{T}}\sqrt{S_t}$. Then $[S_{ij},S_{kl}] = [H_{ij},H_{kl}]$.

$$H_{t,ij} = \sum_{m,n}(\sqrt{S_t})_{in}dB_{t,nm}Q_{mj} + Q_{mi}dB_{t,nm}(\sqrt{S_t})_{nj}$$

The first summand of $d[H_{ij},H_{kl}]_t$ is equal to

$$\sum_{m,n}(\sqrt{S_t})_{in}(\sqrt{S_t})_{kn}Q_{mj}Q_{ml}\,dt$$

Because S_t is symmetric this term can be simplified to

$$S_{t,ik}(Q^{\mathrm{T}}Q)_{jl}\,dt$$

The other three summands can be evaluated in the same way. □

Lemma 4.9. *Let $S \in \mathcal{S}_p^+$ be a stochastic process, $B \sim \mathcal{BM}_p$ and $h : \mathcal{M}_p(\mathbb{R}) \to \mathcal{M}_p(\mathbb{R})$. Then there exists a one dimensional Brownian motion β^h such that*

$$\mathrm{tr}\left(\int_0^t h(S_u)\,dB_u\right) = \int_0^t \sqrt{\mathrm{tr}(h(S_u)^{\mathrm{T}}h(S_u))}\,d\beta_u^h$$

Proof. Define

$$\beta_T^h := \sum_{i,n=1}^{p} \int_0^T \frac{h(S_t)_{in}}{\sqrt{\mathrm{tr}(h(S_t)^{\mathrm{T}} h(S_t))}} \, dB_{t,ni}$$

Observe that the denominator is zero if and only if the numerator is zero, so we make the convention $\frac{0}{0} := 1$. Then we have by definition

$$\mathrm{tr}(h(S_t)\, dB_t) = \sum_{i,n=1}^{p} h(S_t)_{in} \, dB_{t,ni} = \sqrt{\mathrm{tr}(h(S_t)^{\mathrm{T}} h(S_t))} \, d\beta_t^h$$

For every $i, n = 1, \ldots, p$ we have

$$\begin{aligned} \int_0^T \left(\frac{h(S)_{in}}{\sqrt{\mathrm{tr}(h(S)^{\mathrm{T}} h(S))}} \right)^2 dt &= \int_0^T \frac{h(S)_{in}^2}{\mathrm{tr}(h(S)^{\mathrm{T}} h(S))} \, dt \\ &\leq \int_0^T \sum_{i,n=1}^{p} \frac{h(S)_{in}^2}{\mathrm{tr}(h(S)^{\mathrm{T}} h(S))} \, dt = T < \infty \end{aligned}$$

and therefore β^h is a sum of continuous local martingales and thus a continuous local martingale itself. Furthermore

$$[\beta^h, \beta^h]_T = \int_0^T \sum_{i,n} \frac{h(S_t)_{in}^2}{\mathrm{tr}(h(S_t)^{\mathrm{T}} h(S_t))} \, dt = \int_0^T dt = T$$

and with Lévy's Theorem the proof is complete. ☐

Lemma 4.10. *The matrix variate square root function is locally Lipschitz on $\mathcal{S}_p^+$.*

Proof. See Stelzer [2007]. ☐

Theorem 4.11 (Existence and Uniqueness of the Wishart Process I)**.**
For every initial value $s_0 \in \mathcal{S}_p^+$ there exists a unique strong solution S of the Wishart SDE (4.9) in the cone $\mathcal{S}_p^+$ of all positive definite matrices up to the stopping time

$$T = \inf\{s : \det(S_s) = 0\} > 0 \ \text{a.s.}$$

Proof. We check that the assumptions of Theorem 3.34 are fulfilled. Observe, that the set $U = \mathcal{S}_p^+$ is open and that there exists a sequence of convex closed subsets $(U_n)_{n\in\mathbb{N}}$ of U that are increasing w.r.t. $\subseteq$ and $\bigcup_{n\in\mathbb{N}} U_n = U$. Indeed, observe that the function that maps every Matrix $M \in \mathcal{S}_p^+$ to its smallest eigenvalue,

$$\lambda_{\min} : \mathcal{S}_p^+ \to (0, \infty), M \mapsto \lambda_{\min}(M) = \min_{||v||=1} v^{\mathrm{T}} M v$$

is continuously differentiable, see Deuflhard and Hohmann [2002, Lemma 5.1]. Hence,

$$U_n := \{M \in \mathcal{S}_p^+ : \lambda_{\min}(M) \geq \frac{1}{n}\} = \lambda_{\min}^{-1}\left(\underbrace{\left[\frac{1}{n}, \infty\right)}_{closed} \right)$$

is a closed set; and also convex, as for all $M_1, M_2 \in U_n, \alpha \in [0,1]$:

$$\begin{aligned}
\lambda_{\min}(\alpha M_1 + (1-\alpha)M_2) &= \min_{||v||=1} v^{\mathrm{T}}(\alpha M_1 + (1-\alpha)M_2)v \\
&\geq \alpha \min_{||v||=1} v^{\mathrm{T}} M_1 v + (1-\alpha) \min_{||v||=1} v^{\mathrm{T}} M_2 v \\
&\geq \alpha \frac{1}{n} + (1-\alpha)\frac{1}{n} = \frac{1}{n}
\end{aligned}$$

Next, we define for $S \in \mathcal{S}_p^+$ the linear operator by

$$\mathcal{Z}_S = \mathcal{Z}(S) : \mathcal{M}_p(\mathbb{R}) \to \mathcal{S}_p, X \mapsto \sqrt{S}XQ + Q^{\mathrm{T}}X^{\mathrm{T}}\sqrt{S}$$

and as before

$$\mathcal{C} : \mathcal{S}_p \to \mathcal{S}_p, X \mapsto XK + K^{\mathrm{T}}X$$

Then we can write (4.9) in the form

$$\begin{aligned}
S_t &= \int_0^t \mathcal{Z}_{S_u} dB_u + \int_0^t (\mathcal{C}S_u + \alpha Q^{\mathrm{T}}Q)\, du \\
&= \int_0^t \underbrace{\begin{pmatrix} \mathcal{Z}(S_u) \\ \mathcal{C}S_u + \alpha Q^{\mathrm{T}}Q \end{pmatrix}^{\mathrm{T}}}_{:=f(S_u)} d \underbrace{\begin{pmatrix} B_u \\ uI_p \end{pmatrix}}_{=:Z_u} \\
&= \int_0^t f(S_u)\, dZ_u
\end{aligned}$$

where Métivier and Pellaumail [1980b] give a formal justification why we can integrate w.r.t $\mathcal{Z}_{S_u} dB_u$. Obviously, Z is a continuous semimartingale. We still have to show that the function f is locally Lipschitz. For any norm given on $\mathcal{M}_p(\mathbb{R})$, we define a norm on $\mathcal{M}_{2p,p}(\mathbb{R})$ by

$$||(X,Y)^{\mathrm{T}}||_{\mathcal{M}_{2p,p}(\mathbb{R})} = ||X||_{\mathcal{M}_p(\mathbb{R})} + ||Y||_{\mathcal{M}_p(\mathbb{R})} \quad \forall\, X, Y \in \mathcal{M}_p(\mathbb{R})$$

but discard the subscripts as it should be obvious which norm to use. Then, for all $S, R \in \mathcal{S}_p^+$ we have

$$||f(S) - f(R)|| = ||\mathcal{Z}_S - \mathcal{Z}_R|| + ||\mathcal{C}(S-R)||$$

Because $\mathcal{C}$ is a bounded linear operator ($\dim \mathcal{S}_p < \infty$), we have

$$||\mathcal{C}(S-R)|| \leq ||\mathcal{C}||\, ||S-R|| \quad \forall\, S, R \in \mathcal{S}_p^+ \text{ with } ||\mathcal{C}|| < \infty$$

Let be $S, R \in \mathcal{S}_p^+$. For $X \in \mathcal{M}_p(\mathbb{R})$ arbitrary we have

$$\begin{aligned}
||(\mathcal{Z}_S - \mathcal{Z}_R)X|| &= ||(\sqrt{S} - \sqrt{R})XQ + Q^{\mathrm{T}}X^{\mathrm{T}}(\sqrt{S} - \sqrt{R})|| \\
&\leq 2||\sqrt{S} - \sqrt{R}||\, ||X||\, ||Q|| \\
&\Rightarrow ||\mathcal{Z}_S - \mathcal{Z}_R|| = \sup_{X \in \mathcal{M}_p(\mathbb{R}) \setminus \{0\}} \frac{||(\mathcal{Z}_S - \mathcal{Z}_R)X||}{||X||} \leq 2||\sqrt{S} - \sqrt{R}||\, ||Q||
\end{aligned}$$

Let now be $Y \in \mathcal{S}_p^+$. As the matrix variate square root function is locally Lipschitz, there exists an open neighbourhood $\mathcal{U}(Y)$ of Y and a constant $C(Y)$, such that for all $S, R \in \mathcal{U}(Y)$

$$||\sqrt{S} - \sqrt{R}|| \leq C(Y)\, ||S - R||$$

Hence, we have for all $S, R \in \mathcal{U}(Y)$

$$||\mathcal{Z}_S - \mathcal{Z}_R|| \leq \underbrace{2\, C(Y)\, ||Q||}_{=:C'(Y)}\, ||S - R|| = C'(Y)\, ||S - R||$$

i.e. $\mathcal{Z}$ is locally Lipschitz on $\mathcal{S}_p^+$. At all f is locally Lipschitz because

$$||f(S) - f(R)|| \leq \underbrace{(C'(Y) + ||\mathcal{C}||)}_{=:K}\, ||S - R|| = K\, ||S - R|| \quad \forall\, S, R \in \mathcal{U}(Y) \tag{4.15}$$

Hence, we know that there exists a non-zero stopping time $T > 0$ such that there exists a unique $\mathcal{S}_p^+$-valued strong solution S of (4.9) for $t \in [0, T)$. If $T < \infty$, S_T either hits the boundary of $\mathcal{S}_p^+$ or explodes. We show that the latter cannot happen. Fix $R \in \mathcal{U}(Y)$ and set $S = Y$. Then we from (4.15)

$$\begin{aligned} ||f(Y) - f(R)|| &\leq K\, (||Y|| + ||R||) \\ \Rightarrow ||f(Y) - f(R)||^2 &\leq K^2\, (||Y||^2 + 2||Y||\, ||R|| + ||R||^2) \leq L\, (1 + ||Y|| + ||Y||^2) \end{aligned}$$

with $L = \max\{K^2||R||^2, 2K^2||R||, K^2\}$. Because of $||Y|| \leq 1 + ||Y||^2$

$$||f(Y) - f(R)||^2 \leq 2L\, (1 + ||Y||^2)$$

Hence, $Y \mapsto f(Y) - f(R)$ satisfies the linear growth condition, and so does $f : Y \mapsto f(Y)$.

Thus, if $T < \infty$, we know that S_T hits the boundary of $\mathcal{S}_p^+$ the first time, i.e. $S_T \in \overline{\mathcal{S}_p^+}$, but $S_T \notin \mathcal{S}_p^+$, and $S_t \in \mathcal{S}_p^+$ for all $t < T$. Hence

$$T = \inf\{u :\, S_u \notin \mathcal{S}_p^+\}$$

or

$$T = \inf\{s :\, \det(S_s) = 0\}$$

□

In other words: A unique strong solution S of (4.9) exists as long as S stays in the interior of $\overline{\mathcal{S}_p^+}$.
As a consequence, in order to show that there exists a unique strong solution S of (4.9) in $\mathcal{S}_p^+$ on the entire interval $[0, \infty)$ we only need to show that $S_t \in \mathcal{S}_p^+$ for all $t \in \mathbb{R}_+$.
Hence, the next step is to give sufficient conditions that guarantee that $S_t \in \mathcal{S}_p^+$ for all $t \in \mathbb{R}_+$. First, we do this for the case with a zero drift, $K = 0$. Later, we can generalize our results for $K \neq 0$ using a Girsanov transformation.

Theorem 4.12. *For every initial value $s_0 \in \mathcal{S}_p^+$ there exists a stopping time $T > 0$ and a unique strong solution of the Wishart SDE (4.9) on $[0, T)$. Suppose $T < \infty$. Then there exists a unique strong solution $(S_t)_{t\in[0,T]}$ of the Wishart SDE (4.9) on $[0, T)$ such* *that $S_t \in \mathcal{S}_p^+$ for all $t \in [0,T)$ and $S_T \in \partial\mathcal{S}_p^+ = \overline{\mathcal{S}_p^+}\backslash\mathcal{S}_p^+$. Then, for every $x \in \mathbb{R}^p$ with $x^{\mathrm{T}}Q^{\mathrm{T}}Qx = 1$ the process $(x^{\mathrm{T}}S_tx)_{t\in[0,T]}$ is a square Bessel process with parameter α and initial value $x^{\mathrm{T}}s_0x$, $(x^{\mathrm{T}}S_tx)_{t\in[0,T]} \sim \mathrm{BESQ}(\alpha, x^{\mathrm{T}}s_0x)$. Moreover, if $\alpha \geq 2$ then the process $(x^{\mathrm{T}}S_tx)_{t\in[0,T]}$ remains positive almost surely. If furthermore $Q \in GL(p)$ then:*

$$\text{For every } y \in \mathbb{R}^p, y \neq 0, \text{ it holds that } y^{\mathrm{T}}S_ty > 0 \text{ for all } t \in [0,T] \text{ a.s.}$$

Proof. We get the required solution $(S_t)_{t\in[0,T]}$ of (4.9) by Theorem 4.11 if we attach the one point S_T to the solution on $[0,T)$ of (4.9).
Let $x \in \mathbb{R}^p$ be an arbitrary vector with $x^{\mathrm{T}}Q^{\mathrm{T}}Qx = 1$. The matrix of all partial derivatives of the map $\mathcal{M}_p(\mathbb{R}) \ni S \mapsto x^{\mathrm{T}}Sx \in \mathbb{R}$ is equal to

$$D(x^{\mathrm{T}}Sx) = (x_ix_j)_{i,j} = xx^{\mathrm{T}}$$

and hence all second derivatives are zero.
With Itô's formula (Theorem 3.27) we get

$$\begin{aligned} d(x^{\mathrm{T}}S_tx) &= \mathrm{tr}(xx^{\mathrm{T}}\,dS_t) \\ &= \mathrm{tr}(Qxx^{\mathrm{T}}\sqrt{S_t}\,dB_t) + \mathrm{tr}(\sqrt{S_t}xx^{\mathrm{T}}Q^{\mathrm{T}}\,dB_t^{\mathrm{T}}) + \mathrm{tr}(xx^{\mathrm{T}}\alpha Q^{\mathrm{T}}Q\,dt) \\ &= 2\,\mathrm{tr}(Qxx^{\mathrm{T}}\sqrt{S_t}\,dB_t) + \mathrm{tr}(xx^{\mathrm{T}}\alpha Q^{\mathrm{T}}Q\,dt) \\ &= 2\sqrt{\mathrm{tr}(S_txx^{\mathrm{T}}Q^{\mathrm{T}}Qxx^{\mathrm{T}})}\,d\beta_t + \alpha\mathrm{tr}(x^{\mathrm{T}}Q^{\mathrm{T}}Qx)\,dt \\ &= 2\sqrt{\mathrm{tr}(S_txx^{\mathrm{T}})}\,d\beta_t + \alpha\,dt \\ &= 2\sqrt{x^{\mathrm{T}}S_tx}\,d\beta_t + \alpha\,dt \end{aligned}$$

where we used Lemma 4.9. Hence, $(x^{\mathrm{T}}S_tx)_{t\in[0,T]} \sim \mathrm{BESQ}(\alpha, x^{\mathrm{T}}s_0x)$.
If $\alpha \geq 2$ we know from Theorem 4.5 that the process $(x^{\mathrm{T}}S_tx)_{t\in[0,T]}$ is strictly positive for all $t \in [0,T]$ a.s., because the initial value is positive, $x^{\mathrm{T}}s_0x > 0$.
Now suppose that $Q \in GL(p)$. Let $y \in \mathbb{R}^p, y \neq 0$, then for $x := (y^{\mathrm{T}}Q^{\mathrm{T}}Qy)^{-\frac{1}{2}}y$ it holds that $x^{\mathrm{T}}Q^{\mathrm{T}}Qx = 1$ and by the above that $x^{\mathrm{T}}S_tx > 0$ for all $t \in [0,T]$ a.s. Thus, we also have $y^{\mathrm{T}}S_ty > 0$ for all $t \in [0,T]$ a.s. □

First it may sound surprisingly that for fixed $y \neq 0$, the process $(y^{\mathrm{T}}S_ty)_{t\in[0,T]}$ is almost surely positive at T, $y^{\mathrm{T}}S_Ty > 0$ a.s., even though that there exists an $z \in \mathbb{R}^p, z \neq 0$, such that $z^{\mathrm{T}}S_Tz = 0$ a.s. But this just tells us, that it is 'unlikely' to find such a vector z in advance. Before we continue with the next theorem, we state

Theorem 4.13. *Let $S \sim \mathcal{WP}_p(Q, 0, \alpha, s_0)$. Then, for $\xi \neq 0$*

$$d(\det(S_t)) = 2\det(S_t)\sqrt{\mathrm{tr}(Q^{\mathrm{T}}QS_t^{-1})}d\beta_t + \det(S_t)(\alpha+1-p)\mathrm{tr}(Q^{\mathrm{T}}QS_t^{-1})\,dt \quad (4.16)$$

$$d(\det(S_t)^{\xi}) = 2\xi\det(S_t)^{\xi}\left[\sqrt{\mathrm{tr}(Q^{\mathrm{T}}QS_t^{-1})}\,d\beta_t + \mathrm{tr}(Q^{\mathrm{T}}QS_t^{-1})(\frac{\alpha-1-p}{2}+\xi)dt\right] \quad (4.17)$$

$$d(\ln(\det(S_t))) = 2\sqrt{\mathrm{tr}(Q^{\mathrm{T}}QS_t^{-1})}\,d\beta_t + (\alpha-p-1)\mathrm{tr}(Q^{\mathrm{T}}QS_t^{-1})\,dt$$

$$d(\ln(\det(S_t))) = 2\sqrt{\mathrm{tr}(Q^{\mathrm{T}}QS_t^{-1})}\,d\beta_t \text{ for } \alpha = p+1 \quad (4.18)$$

for $t \in [0, T)$ with $T = \inf\{s : \det(S_s) = 0\}$, where β is a one-dimensional Brownian motion.

These results can also be found in Bru [1991, p. 747] without proof.

Proof. We consider the SDE

$$dS_t = \sqrt{S_t} dB_t Q + Q^{\mathrm{T}} dB_t^{\mathrm{T}} \sqrt{S_t} + \alpha Q^{\mathrm{T}} Q dt, \quad S_0 = s_0$$

First, we prove Equation (4.16):
According to Itô's formula, we have

$$d(\det(S_t)) = \mathrm{tr}(D(\det(S_t))\, dS_t) + \frac{1}{2} \sum_{i,j,k,l=1}^{p} \frac{\partial^2}{\partial S_{t,ij}\, \partial S_{t,kl}} \det(S_t)\, d[S_{ij}, S_{kl}]_t$$

Using Lemma 2.6 we get

$$\begin{aligned} \mathrm{tr}(D(\det(S))\, dS_t) &= \det(S_t)\mathrm{tr}(S_t^{-1}\, dS_t) \\ &= \det(S_t)\mathrm{tr}(S_t^{-1}(\sqrt{S_t} dB_t Q + Q^{\mathrm{T}} dB_t^{\mathrm{T}} \sqrt{S_t} + \alpha Q^{\mathrm{T}} Q dt)) \\ &= \det(S_t)[\mathrm{tr}(Q S_t^{-\frac{1}{2}}\, dB_t) + \mathrm{tr}(S_t^{-\frac{1}{2}} Q^{\mathrm{T}}\, dB_t^{\mathrm{T}}) + \alpha \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})] \\ &= \det(S_t)[2\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + \alpha \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})] \end{aligned}$$

where we used Lemma 4.9 in the last equation. For the second order term, we get

$$\begin{aligned} & \frac{1}{2} \sum_{i,j,k,l=1}^{p} \frac{\partial^2}{\partial S_{t,ij}\, \partial S_{t,kl}} \det(S_t)\, d[S_{ij}, S_{kl}]_t \\ = & \frac{1}{2} \sum_{i,j,k,l=1}^{p} \det(S)[(S_t^{-1})_{kl}(S_t^{-1})_{ij} - (S_t^{-1})_{ik}(S_t^{-1})_{lj}] d[S_{ij}, S_{kl}]_t \\ = & \frac{1}{2} \sum_{i,j,k,l=1}^{p} \det(S)[(S_t^{-1})_{kl}(S_t^{-1})_{ij} - (S_t^{-1})_{ik}(S_t^{-1})_{lj}][S_{t,ik}(Q^{\mathrm{T}}Q)_{jl}\, dt + S_{t,il}(Q^{\mathrm{T}}Q)_{jk}\, dt \\ & \qquad + S_{t,jk}(Q^{\mathrm{T}}Q)_{il}\, dt + S_{t,jl}(Q^{\mathrm{T}}Q)_{ik}\, dt] \\ = & \det(S_t)[(1-p)\mathrm{tr}(Q Q^{\mathrm{T}} S_t^{-1})\, dt] \end{aligned}$$

where we used Lemma 4.8 and Lemma 2.6, again. At all, we get equation (4.16)

$$d(\det(S_t)) = 2\det(S_t)\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + \det(S_t)(\alpha + 1 - p)\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt$$

For Equation (4.17), we observe that

$$dX_t^{\xi} = \xi X_t^{\xi-1}\, dX_t + \frac{1}{2}\xi(\xi - 1) X_t^{\xi-2} d[X, X]_t \tag{4.19}$$

If we set

$$X_t := det(S_t)$$

then (4.16) is equal to

$$dX_t = 2X_t \left(\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + \frac{\alpha + 1 - p}{2} \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \right) \tag{4.20}$$

and

$$d[X, X]_t = 4X_t^2 \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt$$

If we insert (4.20) into (4.19) we get (4.17):

$$\begin{aligned}
dX_t^\xi &= \xi X_t^{\xi-1} 2X_t \left(\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + \frac{\alpha + 1 - p}{2} \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \right) \\
&+ \frac{1}{2} \xi(\xi - 1) X_t^{\xi-2} 4X_t^2 \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \\
&= 2\xi X_t^\xi \left[\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + \frac{\alpha + 1 - p}{2} \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt + (\xi - 1)\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \right] \\
&= 2\xi X_t^\xi \left[\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1}) \left(\frac{\alpha - 1 - p}{2} + \xi \right) dt \right]
\end{aligned}$$

To prove the last equation, observe that

$$d(\ln(X_t)) = X_t^{-1}\, dX_t - \frac{1}{2} X_t^{-2}\, d[X, X]_t \tag{4.21}$$

and insert (4.20) into (4.21):

$$\begin{aligned}
d(\ln(X_t)) &= X_t^{-1} 2X_t \left(\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + \frac{\alpha + 1 - p}{2} \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \right) \\
&- \frac{1}{2} X_t^{-2} 4X_t^2 \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \\
&= 2 \left(\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + \frac{\alpha + 1 - p}{2} \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \right) - 2\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \\
&= 2\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + (\alpha + 1 - p - 2)\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt \\
&= 2\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t + (\alpha - p - 1)\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt
\end{aligned}$$

which equals (4.18), if $\alpha = p + 1$. □

Theorem 4.14 (Existence and Uniqueness of the Wishart Process II).
Let $K = 0$, $s_0 \in \mathcal{S}_p^+$ and $\alpha \geq p + 1$. Then there exists a unique strong solution in $\mathcal{S}_p^+$ of the Wishart SDE (4.9) on $[0, \infty)$.

Proof. We consider the SDE

$$dS_t = \sqrt{S_t} dB_t Q + Q^{\mathrm{T}} dB_t^{\mathrm{T}} \sqrt{S_t} + \alpha Q^{\mathrm{T}} Q dt, \quad S_0 = s_0$$

and show that

$$T = \inf\{s : \det(S_s) = 0\} = \infty$$

where we adopt the idea from Bru [1991, p.734] to use McKean's argument. As a matrix norm we choose

$$||A|| := \max_{k=1:p} \sum_{j=1}^{p} |A_{jk}| \tag{4.22}$$

Observe that then $|\mathrm{tr}(A)| \leq p||A||$ for every matrix $A \in \mathcal{M}_p(\mathbb{R})$.

Let us assume that $T < \infty$.

First consider the case $\alpha = p + 1$. Then we have from (4.18)

$$d(\ln(\det(S_t))) = 2\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t$$

We can define an increasing sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ with

$$T_n := \inf\{t \in \mathbb{R}_+ :\ ||S_t^{-1}|| = n\}$$

that converges to T such that $\ln(\det(S_{\min\{t,T_n\}}))$ is a martingale:

$$\begin{aligned} E\left(\int_0^{T_n} \left(2\sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\right)^2 dt\right) &= 4E\left(\int_0^{T_n} \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})\, dt\right) \\ &\leq 4E\left(\int_0^{T_n} p||Q^{\mathrm{T}} Q||n\, dt\right) \\ &= 4T_n p||Q^{\mathrm{T}} Q||n < \infty \end{aligned}$$

where we used that

$$0 \leq \mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1}) \leq p||Q^{\mathrm{T}} Q S_t^{-1}|| \leq p||Q^{\mathrm{T}} Q||\, ||S_t^{-1}|| \leq p||Q^{\mathrm{T}} Q||n$$

for $t \in [0, T_n)$. That means by definition that $\ln(\det(S_t))$ is a local martingale on $[0, T)$. Now we can apply McKean's argument, Theorem 3.44, with $r_t = \det(S_t)$ and $h \equiv ln$. By assumption we know $\det(S_0) > 0$ and we have shown above that $h(r_t) = \ln(\det(S_t))$ is a local martingale on $[0, T)$. Obviously, $\ln(\det(S_t))$ converges to -∞ for $t \to T$ and hence McKean's argument implies $T = \infty$. That is a contradiction to our assumption. Logically consistent, we can conclude $T = \infty$.

In the case $\alpha > p + 1$, we set $\xi = \frac{p+1-\alpha}{2} < 0$ and

$$T_n := \inf\{t \in \mathbb{R}_+ :\ ||S_t^{-1}|| = n\} \wedge \inf\{t \in \mathbb{R}_+ :\ \det(S_t^{-1}) \geq n\}$$

where $a \wedge b := \inf\{a, b\}$. Again, $(T_n)_{n\in\mathbb{N}}$ is a sequence of stopping times that converges to T. From (4.17) we know that

$$d(\det(S_t)^{\xi}) = 2\xi \det(S_t)^{\xi} \sqrt{\mathrm{tr}(Q^{\mathrm{T}} Q S_t^{-1})}\, d\beta_t$$

We show again that $\det(S_{t\wedge T_n})^\xi$ is a martingale for every $n \in \mathbb{N}$:

$$\begin{aligned} E\left(\int_0^{T_n}\left(2\xi\det(S_t)^\xi\sqrt{\operatorname{tr}(Q^{\mathrm{T}}QS_t^{-1})}\right)^2 dt\right) &= 4\xi^2 E\left(\int_0^{T_n}\det(S_t)^{2\xi}\operatorname{tr}(Q^{\mathrm{T}}QS_t^{-1})\,dt\right) \\ &\leq 4\xi^2 E\left(\int_0^{T_n} n^{-2\xi}p||Q^{\mathrm{T}}Q||n\,dt\right) \\ &= 4\xi^2 T_n p||Q^{\mathrm{T}}Q||n^{1-2\xi} < \infty \end{aligned}$$

Hence, we can apply McKean's argument for the local martingale $\det(S_t)^\xi$ on $[0,T)$, because $\det(S_t)^\xi$ converges to infinity for $t \to T$. The same reasoning as above implies the contradiction $T = \infty$.
Finally, Theorem 4.11 proves the statement.
Observe that, for $\alpha < p+1$, we have $\xi > 0$ and thus $\det(S_t)^\xi$ does converge to zero, so we cannot apply McKean's argument in this case. □

Eventually, we are able to state the final theorem about the existence and uniqueness of the Wishart process, which is the main achievement of this section.

Theorem 4.15 (Existence and Uniqueness of the Wishart Process III).
Let $Q \in GL(p)$, $K \in \mathcal{M}_p(\mathbb{R})$, $s_0 \in \mathcal{S}_p^+$ and the parameter $\alpha \geq p+1$. For $\widehat{B} \sim \mathcal{BM}_p$ consider the stochastic differential equation

$$d\widehat{S}_t = \sqrt{\widehat{S}_t}\,d\widehat{B}_t\,Q + Q^{\mathrm{T}}d\widehat{B}_t^{\mathrm{T}}\sqrt{\widehat{S}_t} + \alpha\,Q^{\mathrm{T}}Q\,dt, \quad S_0 = s_0 \tag{4.23}$$

From Theorem 4.14 we know that there exists an unique strong solution $(\widehat{S},\widehat{B})$ of (4.23) on $[0,T^{\widehat{S}})$ with $T^{\widehat{S}} = \inf\{s : \det(\widehat{S}_s) = 0\} = \infty$. Define the process

$$U_t := -\sqrt{\widehat{S}_t}KQ^{-1}$$

and suppose that

$$\left(\mathcal{E}\left(\operatorname{tr}\left(-\int_0^t U_s^{\mathrm{T}}\,dB_s\right)\right)\right)_{t\in[0,\infty)} \tag{4.24}$$

is a martingale .
Then there exists an unique strong solution $(S,B) = ((S_t)_{t\in\mathbb{R}_+},(B_t)_{t\in\mathbb{R}_+})$, $B \sim \mathcal{BM}_p$, in the cone of all positive definite matrices $\mathcal{S}_p^+$ of the Wishart SDE

$$dS_t = \sqrt{S_t}\,dB_t\,Q + Q^{\mathrm{T}}dB_t^{\mathrm{T}}\sqrt{S_t} + (S_tK + K^{\mathrm{T}}S_t + \alpha\,Q^{\mathrm{T}}Q)\,dt, \quad S_0 = s_0 \tag{4.25}$$

on the entire interval $[0,\infty)$.

Proof. In this proof, with solution we always mean an $\mathcal{S}_p^+$-valued solution.

With Girsanov's Theorem (Theorem 3.41) we are able to conclude that

$$B_t := \int_0^t U_s\,dt + \widehat{B_t} = -\int_0^t \sqrt{\widehat{S}_s}KQ^{-1}\,dt + \widehat{B_t} \tag{4.26}$$

defines a Brownian motion w.r.t. the equivalentprobability measure $\widehat{Q}$ as defined in (3.15).

Some calculation shows that

$$\begin{aligned} d\widehat{S}_t &= \sqrt{\widehat{S}_t}\, d\widehat{B_t}\, Q + Q^{\mathrm{T}} d\widehat{B_t}^{\mathrm{T}} \sqrt{\widehat{S}_t} + \alpha Q^{\mathrm{T}} Q\, dt \\ &= \sqrt{\widehat{S}_t}\, (dB_t + \sqrt{\widehat{S}_t} K Q^{-1}\, dt)\, Q + Q^{\mathrm{T}} d\widehat{B_t}^{\mathrm{T}} \sqrt{\widehat{S}_t} + \alpha Q^{\mathrm{T}} Q\, dt \\ &= \sqrt{\widehat{S}_t}\, dB_t\, Q + Q^{\mathrm{T}} d\widehat{B_t}^{\mathrm{T}} \sqrt{\widehat{S}_t} + \widehat{S}_t K\, dt + \alpha Q^{\mathrm{T}} Q\, dt \\ &= \sqrt{\widehat{S}_t}\, dB_t\, Q + Q^{\mathrm{T}} dB_t^{\mathrm{T}} \sqrt{\widehat{S}_t} + (\widehat{S}_t K + K^{\mathrm{T}} \widehat{S}_t + \alpha Q^{\mathrm{T}} Q)\, dt \end{aligned} \tag{4.27}$$

Hence $(\widehat{S}, B)$ is a solution of (4.25) on $[0, \infty)$.

From Theorem 4.11 we know that there exists a unique strong solution S of (4.27) on the interval $[0, T^S)$ with $T^S = \inf\{s : \det(S_s) = 0\} > 0$. The pathwise uniqueness implies uniqueness in law, thus S and $\widehat{S}$ have the same distribution, and so have T^S and $T^{\widehat{S}}$. Hence, $T^S = T^{\widehat{S}} = \infty$ and (S, B) is an unique strong solution S of (4.25) on the interval $[0, \infty)$. □

Remark 4.16. *In the case that the matrices $Q^{\mathrm{T}}Q$ and K commute, Bru [1991, p. 748] has shown that (4.24) is a martingale by extending the methods of Pitman and Yor [1982] for the one-dimensional case.*

Assumption 4.17. *For the rest of this thesis, we will assume that (4.24) is a martingale.*

Before we end this chapter, we want to compare our results to the one stated in Bru [1991]:

Theorem 4.18. *(Cf. Bru [1991, Theorem 2"]) If $\alpha \in \Delta_p := \{1, \dots, p-1\} \cup (p-1, +\infty)$, $A \in GL(p)$, $B \in \mathcal{S}_p^-$, $s_0 \in \mathcal{S}_p^+$ and has all its Eigenvalues distinct, and M is a $p \times p$-dimensional Brownian motion, then the stochastic differential equation*

$$dS_t = S_t^{\frac{1}{2}}\, dM_t\, (A^{\mathrm{T}}A)^{\frac{1}{2}} + (A^{\mathrm{T}}A)^{\frac{1}{2}}\, dM_t^{\mathrm{T}}\, S_t^{\frac{1}{2}} + (BS_t + S_t B)\, dt + \alpha A^{\mathrm{T}} A\, dt, S_0 = s_0 \tag{4.28}$$

has a unique solution on $[0, \tau)$ if B and $(A^{\mathrm{T}}A)^{\frac{1}{2}}$ commute, whereas τ denotes the first time of collision, i.e. the first time that two eigenvalues of S become equal. With the term unique solution is meant a weak solution that is unique in law.

Compared to (4.28), our definition of the Wishart SDE (4.9) is more general, as we allow the drift matrix K to be an arbitrary matrix, whereas in (4.28) the matrix B has to be symmetric negative definite. In the case $\alpha \in [p+1, \infty)$, the result in Theorem 4.15 extends the one in Theorem 4.18 because we prove the existence of a strong solution, that is unique up to indistinguishability and takes almost surely values in the cone of all symmetric positive definite matrices. Furthermore, this solution has infinite lifetime independently of any collision of the eigenvalues of our solution.

4.3 Square Ornstein-Uhlenbeck Processes and their Distributions

Theorem 4.19.
Let $n \in \{p+1, p+2, \ldots\}$, $A \in GL(p)$, $B \in \mathcal{M}_p(\mathbb{R})$, $s_0 \in \mathcal{S}_p^+$ and $X \sim \mathcal{OUP}_{n,p}(A, B, x_0)$ be an Ornstein-Uhlenbeck process with $x_0^{\mathrm{T}} x_0 = s_0$. Then there exists a Brownian motion $M \sim \mathcal{BM}_p$ such that the unique strong solution (S, M) in $\mathcal{S}_p^+$ of the stochastic differential equation

$$dS_t = S_t^{\frac{1}{2}}\, dM_t\, (A^{\mathrm{T}}A)^{\frac{1}{2}} + (A^{\mathrm{T}}A)^{\frac{1}{2}}\, dM_t^{\mathrm{T}}\, S_t^{\frac{1}{2}} + (B^{\mathrm{T}} S_t + S_t B)\, dt + nA^{\mathrm{T}}A\, dt, \quad S_0 = s_0 \quad (4.29)$$

is given by the square Ornstein-Uhlenbeck process $S = X^{\mathrm{T}} X = (X_t^{\mathrm{T}} X_t)_{t \in \mathbb{R}_+}$.

Note that the class of stochastic differential equations of the form (4.29) is exactly the class of Wishart SDEs (4.9) with $Q = \sqrt{A^{\mathrm{T}}A} \in \mathcal{S}_p^+$, $K = B \in \mathcal{M}_p(\mathbb{R})$ and $\alpha = n \in \{p+1, p+2, \ldots\}$. Thus, everything we have established in the foregoing section is still valid for the subclass of SDEs (4.29).

Proof. We define

$$S_t := X_t^{\mathrm{T}} X_t \quad \forall\, t \in \mathbb{R}_+$$

and

$$M_t := \int_0^t \sqrt{S_s^{-1}} X_s^{\mathrm{T}}\, dW_s\, A(\sqrt{A^{\mathrm{T}}A})^{-1} \in \mathcal{M}_p(\mathbb{R}) \quad \forall\, t \in \mathbb{R}_+$$

The matrix square root $\sqrt{A^{\mathrm{T}}A}$ is positive definite and therefore invertible. Now we show that M is a Brownian motion. Because of

$$\begin{aligned}
& E\left(\int_0^t (\sqrt{S_s^{-1}} X_s^{\mathrm{T}} A (A^{\mathrm{T}}A)^{-\frac{1}{2}})^{\mathrm{T}} (\sqrt{S_s^{-1}} X_s^{\mathrm{T}} A (A^{\mathrm{T}}A)^{-\frac{1}{2}})\, ds\right) \\
= \;& E\left(\int_0^t (A^{\mathrm{T}}A)^{-\frac{1}{2}} A^{\mathrm{T}} X_s S_s^{-1} X_s^{\mathrm{T}} A (A^{\mathrm{T}}A)^{-\frac{1}{2}}\, ds\right) \\
= \;& \int_0^t (A^{\mathrm{T}}A)^{-\frac{1}{2}} A^{\mathrm{T}} A (A^{\mathrm{T}}A)^{-\frac{1}{2}}\, ds \\
= \;& tI_p < \infty \text{ a.s.}
\end{aligned}$$

M_t is a local martingale. Furthermore, observe that

$$dM_{t,ij} = \sum_{m,n} (\sqrt{S_t^{-1}} X_t)_{im}\, dW_{t,mn}\, (A(\sqrt{A^{\mathrm{T}}A})^{-1})_{nj}$$

and

$$\begin{aligned}
d[M_{ij}, M_{kl}]_t &= \sum_{m,n} (\sqrt{S_t^{-1}} X_t)_{im} (\sqrt{S_t^{-1}} X_t)_{km} (A(\sqrt{A^{\mathrm{T}}A})^{-1})_{nj} (A(\sqrt{A^{\mathrm{T}}A})^{-1})_{nl}\, dt \\
&= (\sqrt{S_t^{-1}} X_t X_t^{\mathrm{T}} \sqrt{S_t^{-1}})_{ik} ((\sqrt{A^{\mathrm{T}}A})^{-1} A^{\mathrm{T}} A (\sqrt{A^{\mathrm{T}}A})^{-1})_{jl} \\
&= (I_p)_{ik} (I_p)_{jl}\, dt \\
&= \mathbf{1}_{\{i=k\}} \mathbf{1}_{\{i=l\}}\, dt
\end{aligned}$$

where we used that

$$d[W_{mn}, W_{m',n'}]_t = dt \Leftrightarrow m = m',\ n = n' \text{ , and zero otherwise}$$

With Theorem 3.42 we con conclude that M is a Brownian motion.
Finally, using the partial integration formula shows us that our pair (S, M) is a solution of the stochastic differential equation (4.29).

$$\begin{aligned}
dS_t &= d(X_t^{\mathrm{T}} X_t) = (dX_t)^{\mathrm{T}} X_t + X_t^{\mathrm{T}} (dX_t) + d[X^{\mathrm{T}}, X]_t^M \\
&= (B^{\mathrm{T}} X_t^{\mathrm{T}} \, dt + A^{\mathrm{T}} \, dW_t^{\mathrm{T}}) X_t + X_t^{\mathrm{T}} (X_t B \, dt + dW_t \, A) + A^{\mathrm{T}} \, d[W_t^{\mathrm{T}}, W_t]_t^M A \\
&= B^{\mathrm{T}} S_t \, dt + A^{\mathrm{T}} \, dW_t \, X_t + S_t B \, dt + X_t^{\mathrm{T}} \, dW_t \, A + A^{\mathrm{T}} \, d[W_t^{\mathrm{T}}, W_t]_t^M A \\
&= (B^{\mathrm{T}} S_t + S_t B) \, dt + A^{\mathrm{T}} \, dW_t \, X_t + X_t^{\mathrm{T}} \, dW_t \, A + A^{\mathrm{T}} \, d[nI_p t]^M A \\
&= (B^{\mathrm{T}} S_t + S_t B) \, dt + (A^{\mathrm{T}} A)^{\frac{1}{2}} \, dM_t^{\mathrm{T}} \, S_t^{\frac{1}{2}} + S_t^{\frac{1}{2}} \, dM_t \, (A^{\mathrm{T}} A)^{\frac{1}{2}} + nA^{\mathrm{T}} A \, dt \\
&= S_t^{\frac{1}{2}} \, dM_t \, (A^{\mathrm{T}} A)^{\frac{1}{2}} + (A^{\mathrm{T}} A)^{\frac{1}{2}} \, dM_t^{\mathrm{T}} \, S_t^{\frac{1}{2}} + (B^{\mathrm{T}} S_t + S_t B) \, dt + nA^{\mathrm{T}} A \, dt
\end{aligned}$$

where we used that

$$d[W_t^{\mathrm{T}}, W_t]_{t,ij}^M = \sum_{k=1}^{n} d[W_{t,ki}, W_{t,kj}]_t = n\mathbf{1}_{\{i=j\}} dt$$

So far we have shown that (S, M) is a weak solution of (4.29). From Theorem 4.15 we know that pathwise uniqueness holds for (4.29), and thus by Remark 3.33 (ii) we have that (S, M) is also a strong solution of (4.29). □

Theorem 4.20 (Conditional Distribution of the Square Ornstein-Uhlenbeck Process)**.** *Let $n \in \{p+1, p+2, \ldots\}$, $A \in GL(p)$, $s_0 \in \mathcal{S}_p^+$ and $B \in \mathcal{M}_p(\mathbb{R})$ with $0 \notin -\sigma(B) - \sigma(B)$. The solution of (4.29) has the conditional distribution*

$$S_t | s_0 \sim \mathcal{W}_p(n, \Sigma_t, \Sigma_t^{-1} e^{B^{\mathrm{T}} t} s_0 e^{Bt}) \tag{4.30}$$

with

$$\Sigma_t = \mathcal{A}^{-1}(A^{\mathrm{T}} A) - \mathcal{A}^{-1}(e^{B^{\mathrm{T}} t} A^{\mathrm{T}} A e^{Bt}) \tag{4.31}$$

where $\mathcal{A}^{-1}$ is the inverse of $\mathcal{A} : \mathcal{S}_p \to \mathcal{S}_p, X \mapsto -B^{\mathrm{T}} X - XB$.

Proof. The solution $S = X^{\mathrm{T}} X$ is given by a square Ornstein-Uhlenbeck process. Theorem 3.49 shows that

$$X_t | x_0 \sim \mathcal{N}_{n,p}(x_0 e^{Bt}, I_n \otimes \Sigma_t)$$

with

$$\Sigma_t = \mathcal{A}^{-1}(A^{\mathrm{T}} A) - \mathcal{A}^{-1}(e^{B^{\mathrm{T}} t} A^{\mathrm{T}} A e^{Bt})$$

From Lemma 3.13 we know

$$X_t^{\mathrm{T}} | x_0 \sim \mathcal{N}_{p,n}(e^{B^{\mathrm{T}} t} x_0^{\mathrm{T}}, \Sigma_t \otimes I_n)$$

Using Lemma 3.20 we achieve

$$X_t^{\mathrm{T}} X_t | x_0 = X_t^{\mathrm{T}} (X_t^{\mathrm{T}})^{\mathrm{T}} | x_0 \sim \mathcal{W}_p(n, \Sigma_t, \Sigma_t^{-1} e^{B^{\mathrm{T}} t} x_0^{\mathrm{T}} x_0 e^{Bt})$$

i.e.

$$S_t | s_0 \sim \mathcal{W}_p(n, \Sigma_t, \Sigma_t^{-1} e^{B^{\mathrm{T}} t} s_0 e^{Bt})$$

□

Theorem 4.21 (Stationary Distribution of the Square Ornstein-Uhlenbeck Process). *Let $n \in \{p+1, p+2, \ldots\}$, $A \in GL(p)$, $s_0 \in \mathcal{S}_p^+$ and $B \in \mathcal{M}_p(\mathbb{R})$ with $Re(\sigma(B)) \subseteq (-\infty, 0)$. Then the solution of (4.29) has a stationary limiting distribution, that is*

$$\mathcal{W}_p(n, \mathcal{A}^{-1}(A^{\mathrm{T}}A), 0) \tag{4.32}$$

where $\mathcal{A}^{-1}$ is the inverse of $\mathcal{A} : \mathcal{S}_p \to \mathcal{S}_p, X \mapsto -B^{\mathrm{T}}X - XB$.

Proof. From (4.30) and Theorem 3.19 we know for the solution S of (4.29) that S_t given S_0 has characteristic function

$$\widehat{P^{S_t}} = \det(I_p - 2i\Sigma_t Z)^{-\frac{n}{2}} etr[i\Theta_t(I_p - 2i\Sigma_t Z)^{-1}\Sigma_t Z] \tag{4.33}$$

with

$$\Sigma_t = \mathcal{A}^{-1}(A^{\mathrm{T}}A) - \mathcal{A}^{-1}(e^{B^{\mathrm{T}}t}A^{\mathrm{T}}Ae^{Bt})$$

and

$$\Theta_t = \Sigma_t^{-1}e^{B^{\mathrm{T}}t}s_0e^{Bt}$$

From the proof of Theorem 3.50 we know that

$$Re(\sigma(B)) \subseteq (-\infty, 0) \Rightarrow \lim_{t\to\infty} \exp(Bt) = 0$$

and

$$\lim_{t\to\infty} \Sigma_t = \mathcal{A}^{-1}(A^{\mathrm{T}}A)$$

Thus

$$\lim_{t\to\infty} \Theta_t = 0$$

and

$$\lim_{t\to\infty} \widehat{P^{S_t}} = \det(I_p - 2i\,\mathcal{A}^{-1}(A^{\mathrm{T}}A)\,Z)^{-\frac{n}{2}} =: f(Z)$$

With Lévy's Continuity Theorem it exist a probability measure μ such that $\widehat{\mu}(Z) = f(Z)$ for all $Z \in \mathcal{M}_p(\mathbb{R})$ and $P^{S_t} \overset{weak}{\to} \mu$. According to Remark 3.17 and Theorem 3.19, f is the characteristic function of a random matrix with central Wishart distribution $\mathcal{W}_p(n, \mathcal{A}^{-1}(A^{\mathrm{T}}A), 0)$. Hence, S has limit distribution $\mathcal{W}_p(n, \mathcal{A}^{-1}(A^{\mathrm{T}}A), 0)$ regardless of any initial value s_0. As S is a Markov process (Theorem 3.39) we conclude that $\mathcal{W}_p(n, \mathcal{A}^{-1}(A^{\mathrm{T}}A), 0)$ is its stationary distribution. □

In the case where $B \in \mathcal{S}_p^-$, and the matrices $A^{\mathrm{T}}A$ and B commute, it holds that $\mathcal{A}^{-1}(A^{\mathrm{T}}A) = -\frac{1}{2}A^{\mathrm{T}}AB^{-1}$ and the stationary limiting distribution is $\mathcal{W}_p(n, -\frac{1}{2}A^{\mathrm{T}}AB^{-1}, 0)$.

4.4 Simulation of Wishart Processes

Recall the stochastic differential equation of the Wishart process, that is

$$dS_s = \sqrt{S_s}\, dB_s\, Q + Q^{\mathrm{T}}\, dB_s^{\mathrm{T}}\, \sqrt{S_s} + (S_s K + K^{\mathrm{T}} S_s + \alpha Q^{\mathrm{T}} Q)\, ds, \quad S_0 = s_0$$

For every $t \in \mathbb{R}_+$, $h > 0$ integration over the interval $[t, t+h]$ yields

$$S_{t+h} = S_t + \int_t^{t+h} \sqrt{S_s}\, dB_s\, Q + \int_t^{t+h} Q^{\mathrm{T}}\, dB_s^{\mathrm{T}}\, \sqrt{S_s} + \int_t^{t+h} S_s K\, ds + \int_t^{t+h} K^{\mathrm{T}} S_s\, ds + \alpha Q^{\mathrm{T}} Q h \tag{4.34}$$

We now try to approximate the stochastic integral above to make the Wishart process suitable for numerical simulation. An easy way of doing this, is the Euler-Maruyama method (see Kloeden and Platen [1999, p. 340]):

$$\widehat{S}_{t+h} = \widehat{S}_t + \sqrt{\widehat{S}_t}(B_{t+h} - B_t)Q + Q^{\mathrm{T}}(B_{t+h}^{\mathrm{T}} - B_t^{\mathrm{T}})\sqrt{\widehat{S}_t} + (\widehat{S}_t K + K^{\mathrm{T}} \widehat{S}_t + \alpha Q^{\mathrm{T}} Q)h \tag{4.35}$$

We call $\widehat{S}$ the discretized Wishart process. As the Brownian motion has stationary, independent increments, we know that the distribution of $B_{t+h} - B_t$ is $\mathcal{N}_p(0, hI_{2p})$ and is independent of all previous increments. Hence, we can use (4.35) to simulate $\widehat{S}$ for any fixed step size $h > 0$. Then we get a process on the mesh $(0, h, 2h, \ldots, T)$ for any $T \in \mathbb{R}_+$, that is $(\widehat{S}_0, \widehat{S}_h, \widehat{S}_{2h}, \ldots, \widehat{S}_T)$.
However, even under the Assumptions of Theorem 4.15, this discretized Wishart process can become negative definite such that we have to stop our simulation before we reach T, because the square root in (4.35) is not well-defined anymore.
To solve this , we observe that $\widehat{S} \to S$ for $h \to 0$. Thus, we can expect $\widehat{S}$ to remain positive semidefinite as long as h is sufficiently small. Hence, we introduce a variable step size to our algorithm:
Suppose we have already given the discretized Wishart process $(\widehat{S}_0, \widehat{S}_h, \widehat{S}_{2h}, \ldots, \widehat{S}_t)$ and the discretized Brownian motion $(0, B_h, B_{2h}, \ldots, B_t)$ up to a time $t < T - h$. Suppose further that we calculate $\widehat{S}_{t+h}$ (and thus B_{t+h}) according to (4.35) and that $\widehat{S}_{t+h}$ is negative definite, i.e. (at least) its smallest eigenvalue becomes negative. Then, we cut our step size by half, calculate $\widehat{S}_{t+\frac{h}{2}}$ and check again if the smallest eigenvalue of $\widehat{S}_{t+\frac{h}{2}}$ is negative. We continue this iteratively, until (hopefully) $\widehat{S}_{t+\frac{h}{2^n}}$ is positive semidefinite or, the step size falls under a certain value, say $\approx 2.2 \times 10^{-16}$ (that is the constant *eps* in MATLAB). In the last case, the step size converges to zero.
In order to get the value $\widehat{S}_{t+\frac{h}{2}}$, we need to draw $B_{t+\frac{h}{2}}$ conditionally on the given values $(0, B_h, B_{2h}, \ldots, B_t, B_{t+h})$. Because of the Markov property of the Brownian motion (see Theorem 3.39), this is the same as drawing $B_{t+\frac{h}{2}}$ conditionally on (B_t, B_{t+h}). For now, we only consider the $i, j-th$ entry B_{ij} of B. From Glasserman [2004, p.84] we know that

$$\begin{aligned} B_{ij,t+\frac{h}{2}} | (B_{ij,t} = x, B_{ij,t+h} = x + y) &\stackrel{\mathscr{D}}{=} \frac{\frac{h}{2}x + \frac{h}{2}(x+y)}{h} + \sqrt{\frac{\frac{h}{2}\frac{h}{2}}{h}} Z_{ij} \\ &= x + \frac{1}{2}y + \frac{\sqrt{h}}{2} Z_{ij} \end{aligned}$$

where $Z_{ij} \sim \mathcal{N}_1(0,1)$ and $\overset{\mathscr{D}}{=}$ denotes distributional equivalence. Written in matrix notation, we get for the increment of B

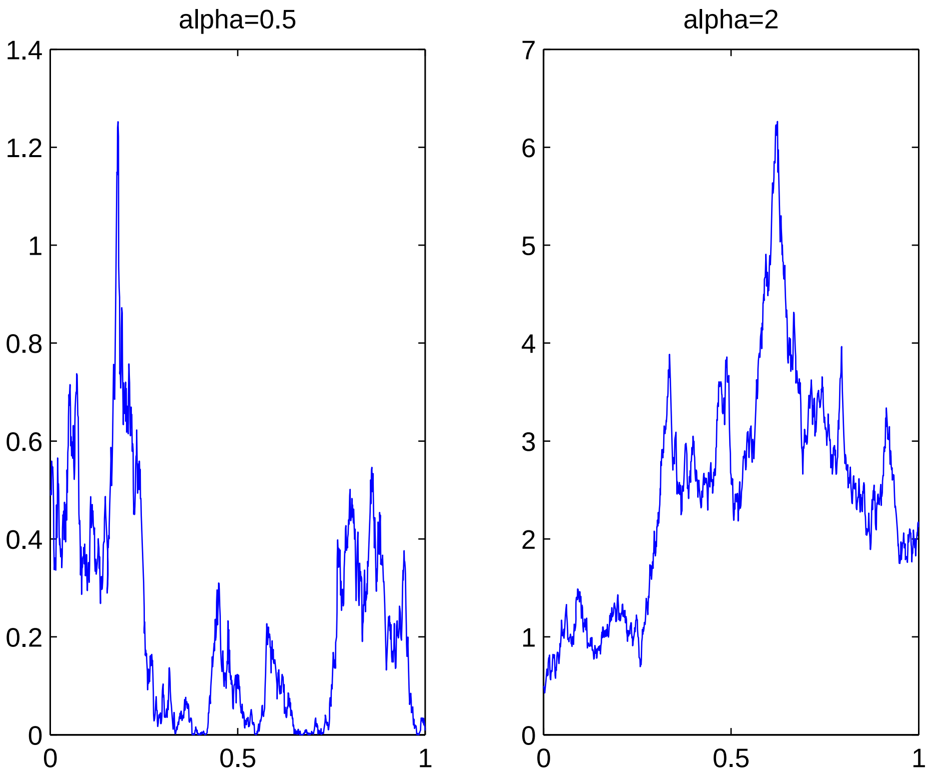
$$(B_{t+\frac{h}{2}} - B_t)|(B_t = X, B_{t+h} = X+Y) \overset{\mathscr{D}}{=} \frac{1}{2}Y + \frac{\sqrt{h}}{2}Z \tag{4.36}$$

where $Z \sim \mathcal{N}_{p,p}(0, I_{2p})$. A sample implementation can be found in the appendix. Now we give a few examples.

Example 1. *First we begin with the one-dimensional square Bessel process, i.e. a solution of*

$$dX_t = 2\sqrt{X_t}\,d\beta_t + \alpha\,dt, \quad X_0 = x_0 := 0.5$$

where β denotes a one-dimensional Brownian motion. From Theorem 4.5 we know that this process never becomes negative for any choice of $\alpha, x_0 \geq 0$ and remains positive for $x_0 > 0$ and $\alpha \geq 2$. Here are two sample paths over the interval $[0,1]$ and initial step size $h = 10^{-3}$:

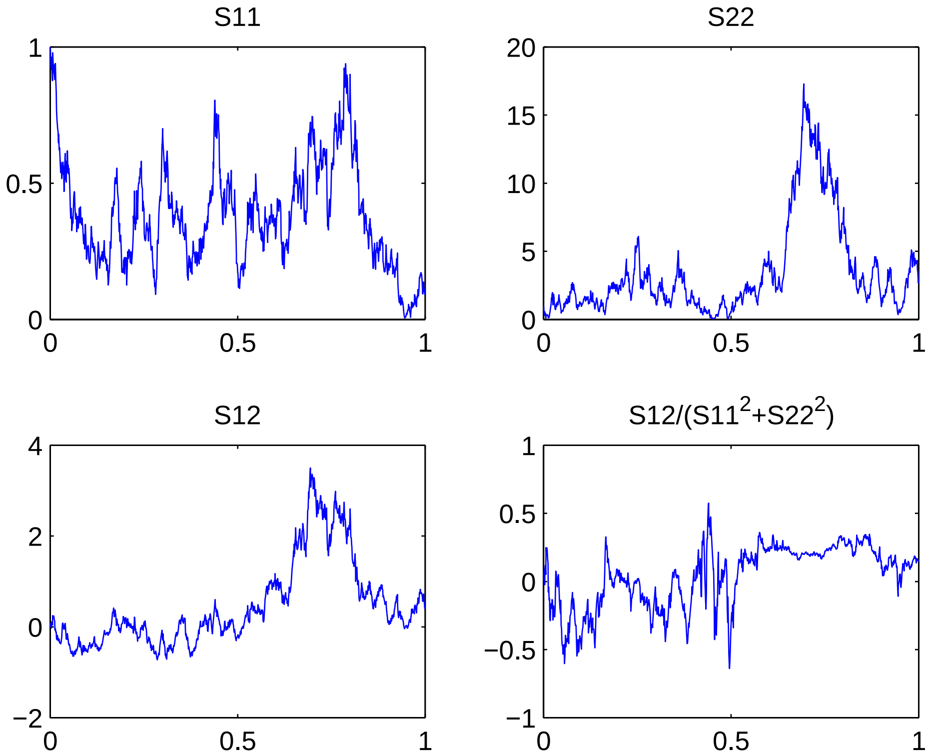


In the case $\alpha = 0.5$, the algorithm had to reduce the initial step size to $1.5625 \cdot 10^{-5}$ at at least one point in order to guarantee non-negativity, whereas for $\alpha = 2$ this was not necessary. Observe that in the case $\alpha = 0.5 < 2$, the square Bessel process can become zero, but gets reflected instantaneously.

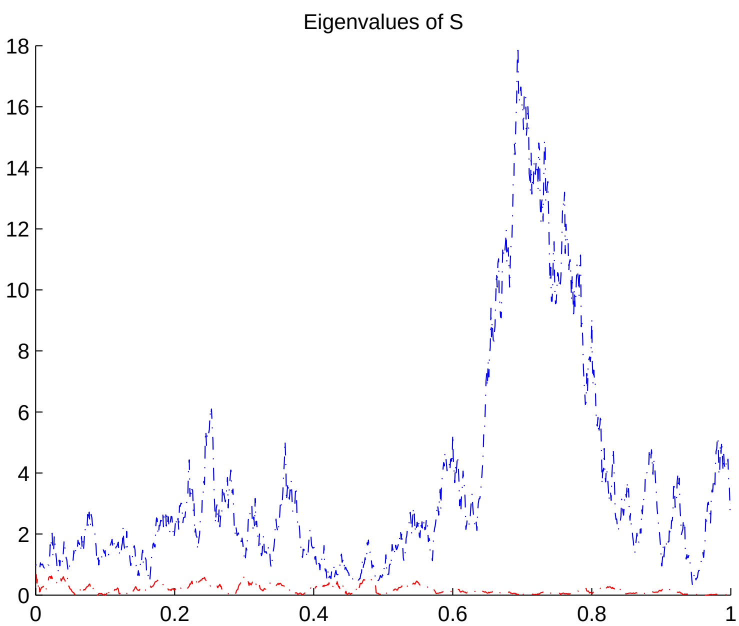
Example 2. *Our next example is the 2-dimensional Wishart process*

$$dS_t = \sqrt{S_t}\,dB_t\,Q + Q^{\mathrm{T}}dB_t^{\mathrm{T}}\sqrt{S_t} + (S_tK + K^{\mathrm{T}}S_t + \alpha Q^{\mathrm{T}}Q)\,dt, \quad S_0 = s_0$$

with a 2-dimensional Brownian motion B, $\alpha = 3$, $Q = \begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix}$, $K = -4I_2$ and $s_0 = I_2$. We show sample paths for the three different entries of S, that are S_{11}, S_{22} and S_{12}, and for $\frac{S_{12}}{S_{11}^2+S_{22}^2}$, what would be the correlation in a stochastic volatility model.



The algorithm had to reduce the initial step size of $h = 10^{-3}$ by half at some points. We have

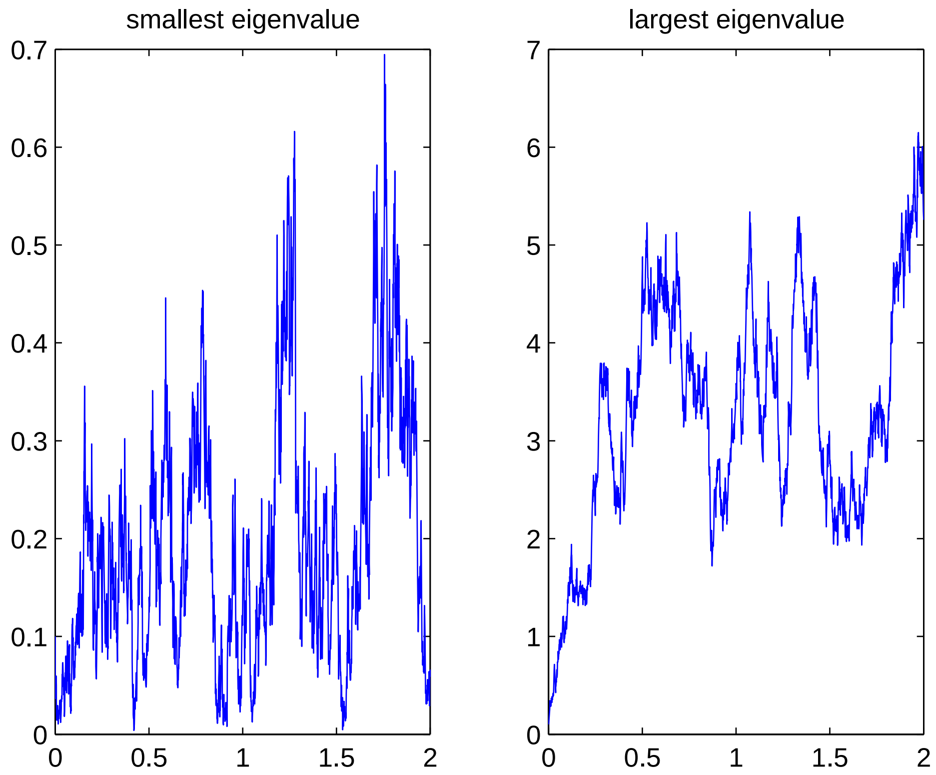


for the eigenvalues of S.

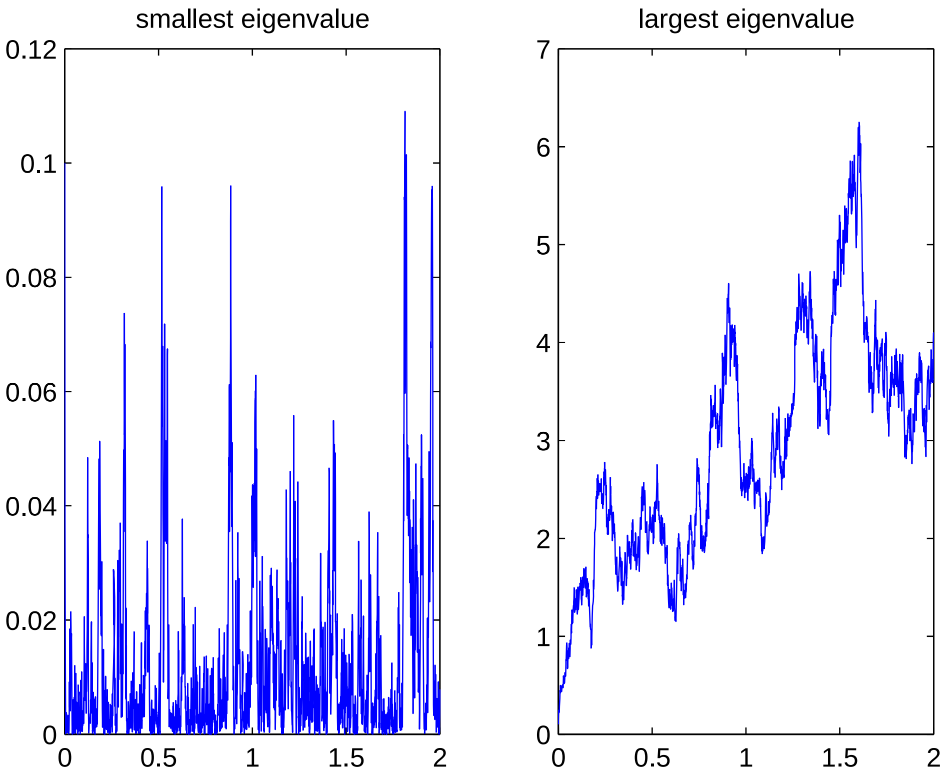
Example 3. *The last example are two 5-dimensional solutions of the stochastic differential equation*

$$dS_t = \sqrt{S_t}\,dB_t + dB_t^{\mathrm{T}}\sqrt{S_t} + (-4S_t + \alpha_i)\,dt, \quad S_0 = 0.1 \cdot I_5, i = 1,2 \tag{4.37}$$

For the first one, we have $\alpha_1 = 7.2$, and the smallest and the largest eigenvalue of S look like



In the second case, $\alpha_2 = 3.5$, we don't know if there exists any solution of (4.37) for $i = 2$. Thus, we do not know whether our simulated process corresponds to a solution of (4.37) for $i = 2$, and if it does, it does not have to be a Wishart process by Definition 4.7 (because we demand the Wishart process to be a strong solution). It may be noted, that the algorithm had to reduce the initial step size of $h = 10^{-3}$ to $7.8125 \cdot 10^{-6}$ at some points. Again, we shall have a look at the eigenvalues of our simulated process:


smallest eigenvalue
0.12
0.1
0.08
0.06
0.04
0.02
0
0
0.5
1
1.5
2
largest eigenvalue
7
6
5
4
3
2
1
0
0
0.5
1
1.5
2

Chapter 5

Financial Applications

We first focus on two well-known financial models that are based on the one-dimensional Wishart processes, that is in fact a generalized squared Bessel process. Later, we consider the actual matrix variate case.

5.1 CIR Model

According to Theorem 4.15, we get for $p = 1$, $q > 0$, $k \in \mathbb{R}$, $\alpha \geq p + 1 = 2$, $s_0 > 0$ and an one-dimensional Brownian motion B that the stochastic differential equation

$$dS_t = 2q\sqrt{S_t}\,dB_t + (2kS_t + \alpha q^2)\,dt, \quad S_0 = s_0$$

has a unique, strong and positive solution in $(0, \infty)$. For $k < 0$ and the new parametrization $\sigma = 2q$, $a = -2k$ and $b = \frac{-\alpha q^2}{2k}$ the stochastic differential equation gets the form

$$dS_t = \sigma\sqrt{S_t}\,dB_t + a(b - S_t)\,dt \tag{5.1}$$

that is the stochastic differential equation of the CIR process. It is mean reverting as $a > 0$ with $b \geq \frac{-q^2}{k}$.
As Glasserman [2004, p.122] has shown, $S_t|s_0$ is distributed as $\frac{\sigma^2(1-\exp(-at))}{4a}$ times a noncentral chi-square random variable with $\frac{4ab}{\sigma^2}$ degrees of freedom and noncentrality parameter $\frac{4a\exp(-at)}{\sigma^2(1-\exp(-at))}s_0$.
From now on we follow Gourieroux [2007, p.184f.]. Let's assume that the interest rate r follows a stochastic differential equation of the form (5.1),

$$dr_t = \sigma\sqrt{r_t}\,dW_t + a(b - r_t)\,dt \tag{5.2}$$

where W is a Brownian motion under a risk-neutral probability measure Q. Then, the price of a zero-coupon bond at time t with time to maturity h is given by

$$B(t, t+h) = E_t^Q\left[\exp\left(-\int_t^{t+h} r_\tau\,d\tau\right)\right] \tag{5.3}$$

where E_t^Q denotes the conditional expectation $E^Q[\cdot|\sigma\{r_s : s \leq t\}]$ under the measure Q. Then Cox et al. [1985] have shown that

$$B(t, t+h) = \exp(-f(h)r_t - g(h)]$$

with functions

$$\begin{aligned} f(h) &= \frac{2}{c+a} - \frac{4c}{(c+a)[(c+a)\exp(ch) + c - a]} \\ g(h) &= -\frac{ab(c+a)h}{\sigma^2} + \frac{2ab}{\sigma^2} \ln\left(\frac{(c+a)\exp(ch) + c - a}{2c}\right) \end{aligned}$$

with $c := \sqrt{a^2 + 2\sigma^2}$.
Hence, we have a closed form solution for $B(t, t+h)$ that is exponential affine in r_t.

5.2 Heston Model

A process S of the form (5.1) can also be used to model the volatility in a stochastic volatility Black-Scholes model according to Heston

$$d(\ln(X_t)) = \mu_t \, dt + \sqrt{S_t} \, dW_t$$

where the stock price at time t is denoted by X_t. We refer to Gourieroux [2007, Chapter 2.2.2.] for details.

5.3 Factor Model for Bonds

In contrast to section 5.1 on the preceding page where the risk-free rate r followed a stochastic differential equation of the form (5.2), we now want to use a factor model to consider corporate bonds jointly with a long term government bond (e.g. T-bond). Again, we summarize the results of Gourieroux [2007, chapter 3.5.2.].
Denote by $\lambda_{i,t}$ the default intensity for firm i, $i = 1, \ldots, K$. Let us assume that

$$\begin{aligned} r_t &= c + \operatorname{tr}(CS_t) && (5.4) \\ \lambda_{i,t} &= d_i + \operatorname{tr}(D_i S_t) \quad \forall\, i = 1, \ldots, K && (5.5) \end{aligned}$$

where $c, d_i \geq 0$ are nonnegative, $C, D_i \in \mathcal{S}_p^+$ and $S \sim \mathcal{WP}_p(Q, K, \alpha, s_0)$, i.e.

$$dS_t = \sqrt{S_t} dB_t Q + Q^{\mathrm{T}} dB_t^{\mathrm{T}} \sqrt{S_t} + (KS_t + S_t K^{\mathrm{T}} + \alpha Q^{\mathrm{T}} Q)dt, \quad S_0 = s_0 \qquad (5.6)$$

under the conditions of Theorem 4.15.
Equation (5.4) is the factor representation for the risk-free rate and equation (5.5) the factor representation for the corporate bonds.
Observe, that $\operatorname{tr}(CS_t) > 0$ for all t. Indeed, as $C \in \mathcal{S}_p^+$ there exists an orthogonal matrix

such that $C = UDU^{\mathrm{T}}$ where D is a diagonal matrix of eigenvalues $\mu_k \geq 0$. Denote by $u_i \in \mathbb{R}^p$, $i = 1, \ldots, p$, the columns of U. Then we have $C = \sum_{k=1}^{p} \mu_k u_k u_k^{\mathrm{T}}$ and thus

$$\mathrm{tr}(CS_t) = \sum_{k=1}^{p} \mu_k \mathrm{tr}(u_k u_k^{\mathrm{T}} S_t) = \sum_{k=1}^{p} \mu_k u_k^{\mathrm{T}} S_t u_k \geq 0$$

With the convention $\lambda_{0,t} = 0$ we get for the price B_0 of a zero-coupon bond and the prices of corporate bonds B_i the formula

$$B_i(t, t+h) = E_t^Q \left[\exp\left(-\int_t^{t+h} r_\tau + \lambda_{i,\tau}\, d\tau \right) \right] \quad \forall\, i = 0, \ldots, K \tag{5.7}$$

under a risk neutral measure Q. If we insert (5.4) and (5.5) into (5.7) we get

$$B_i(t, t+h) = \exp(-h(c + d_i)) E_t^Q \left[\mathrm{etr}\left(-(C + D_i) \int_t^{t+h} S_\tau\, d\tau \right) \right] \quad \forall\, i = 0, \ldots, K \tag{5.8}$$

with the convention $d_0, D_0 \equiv 0$. Gourieroux [2007] shows that there exists a closed form expression for (5.8).

5.4 Matrix Variate Stochastic Volatility Models

An application for matrix variate stochastic processes can be found in Fonseca et al. [2008], which model the dynamics of p risky assets by

$$dX_t = \mathrm{diag}(X_t)[(r\mathbf{1} + \lambda_t)\, dt + \sqrt{S_t}\, dW_t] \tag{5.9}$$

where r is a positive number, $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^p$, λ_t a p-dimensional stochastic process, interpreted as the risk premium, and Z a p-dimensional Brownian motion. The volatility process S is the Wishart process of (5.6).

In contrast to a continous model of the form (1.3), another multivariate stochastic volatility model for the logarithmic stock price process can be given by

$$dY_t = (\mu + \Sigma_t \beta)\, dt + \Sigma_t^{\frac{1}{2}}\, dW_t, \quad Y_0 = 0 \tag{5.10}$$

where $\mu, \beta \in \mathbb{R}^p$, W denotes a p-dimensional Brownian motion and Σ is given by a Lévy-driven positive semidefinite OU type process, see Stelzer [2007] for details. In this case, the volatility is not continuous anymore and has jumps.

If one wants to model the volatility with a time continuous stochastic process, e.g. for economic reasons, one could also suggest the model

$$dY_t = (\mu + S_t \beta)\, dt + S_t^{\frac{1}{2}}\, dW_t, \quad Y_0 = 0 \tag{5.11}$$

where the process Σ was substituted by the Wishart process of (5.6).

Appendix A

MATLAB Code for Simulating a Wishart Process

This is an implementation in MATLAB of the algorithm described in section 4.4 on page 44.

```
function [S,eigS,timestep,minh]=Wishart(T,h,p,alpha,Q,K,s_0)
%
% Simulates the p-dimensional Wishart process on the interval [0,T]
% that follows the stochastic differential equation
%   dS_t=sqrt{S_t}*dB_t*Q+Q'*dB_t'*sqrt{S_t}+(S_t*K+K'*S_t+alpha*Q'*Q)dt
% with initial condition S_0=s_0.
%
% Method of discretization: Euler-Maruyama
% Starting step size: h
% In order to guarantee positive semidefiniteness of S, the step size will be
% reduced iteratively if necessary.
%
% Output:
% S is a three dimensional array of the discretized Wishart process.
% eigS is a matrix consisting the eigenvalues of S
% timestep is the vector of all timesteps in [0,T]
% minh is the smallest step size used, i.e. minh=min(diff(timestep))
%
% Author: Oliver Pfaffel
% Email: O.Pfaffel@gmx.de
% June 17, 2008
%
%--------------------------------------------------------------------

horg=h;
minh=h;
timestep=0;
```

```
[V_0,D_0]=eig(s_0);
eigS=sort(diag(D_0)');
drift_fix=alpha*Q'*Q;

B_old=0;
B_inc=sqrt(h)*normrnd(0,1,p,p);
vola=V_0*sqrt(D_0)*V_0'*B_inc*Q;
drift=s_0*K;
S_new = s_0+vola+vola'+(drift+drift'+drift_fix)*h;

[V_new,D_new]=eig(S_new);
eigS=[eigS;sort(diag(D_new)')];

S=cat(3,s_0,S_new);
t=h;
timestep=[timestep;t];

flag=0;

while t+h<T,

    B_old=B_old+B_inc;
    B_inc=sqrt(h)*normrnd(0,1,p,p);

    S_t=S_new; V_t=V_new; D_t=D_new;

    sqrtm_S_t=V_t*sqrt(D_t)*V_t';
    vola=sqrtm_S_t*B_inc*Q;
    drift=S_t*K;
    S_new = S_t+vola+vola'+(drift+drift'+drift_fix)*h;

    [V_new,D_new]=eig(S_new);

    mineig=min(diag(D_new));

    while mineig<0,

        h=h/2;
        minh=min(minh,h);
        flag=1;

        if h<eps, error('Step size converges to zero'), return, end

        B_inc=0.5*B_inc+sqrt(h/2)*normrnd(0,1,p,p);
```

```
        vola=sqrtm_S_t*B_inc*Q;
        drift=S_t*K;
        S_new = S_t+vola+vola'+(drift+drift'+drift_fix)*h;

        mineig=min(eig(S_new));

    end

    if flag==0,
        eigS=[eigS;sort(diag(D_new)')];
        S=cat(3,S,S_new);
        t=t+h;
        timestep=[timestep;t];
    end

    if flag==1,
        [V_new,D_new]=eig(S_new);
        eigS=[eigS;sort(diag(D_new)')];
        S=cat(3,S,S_new);
        t=t+h;
        timestep=[timestep;t];
        flag=0;
        h=horg;
    end

end

h_end=T-t;

if h_end>0,

    B_inc=sqrt(h_end)*normrnd(0,1,p,p);

    S_t=S_new; V_t=V_new; D_t=D_new;

    vola=V_t*sqrt(D_t)*V_t'*B_inc*Q;
    drift=S_t*K;
    S_new = S_t+vola+vola'+(drift+drift'+drift_fix)*h;

    S=cat(3,S,S_new);
    eigS=[eigS;sort(eig(S_new)')];
    timestep=[timestep;T];

end
```

Bibliography

Ole Eiler Barndorff-Nielsen and Robert Stelzer. Positive-definite matrix processes of finite variation. *Probability and Mathematical Statistics*, 27:3–43, 2007.

Marie-France Bru. Wishart processes. *Journal of Theoretical Probability*, 4:725 – 751, 1991.

B.v.Querenburg. *Mengentheoretische Topologie*. Springer, 2001.

J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

P. Deuflhard and A. Hohmann. *Numerische Mathematik I*. de Gruyter Lehrbuch, 2002.

Gerd Fischer. *Lineare Algebra*. Vieweg, 2005.

J. Da Fonseca, M. Grasselli, and C. Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Springer*, 2008.

Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, 2004.

C. Gourieroux. Continuous time Wishart process for stochastic risk. *Econometric Reviews*, 25:2:177 – 217, 2007.

A. K. Gupta and D. K. Nagar. *Matrix variate distributions*. Chapman & Hall/CRC, 2000.

J. Jacod and P. Protter. *Probability Essentials*. Springer-Verlag, 2004.

Zbigniew J. Jurek and J. David Mason. *Operator-Limit Distributions in Probability Theory*. Wiley Series in Probability and Mathematical Statistics, 1993.

Olav Kallenberg. *Foundations of Modern Probability*. Springer-Verlag, 1997.

Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 1999.

H. P. McKean. *Stochastic Integrals*. Academic Press, 1969.

M. Métivier and J. Pellaumail. *Stochastic Integration*. Academic Press, 1980b.

R. J. Muirhead. *Aspects of Multivariate Statistical Theory*. Wiley, 2005.

Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer-Verlag, 2000.

Ingram Olkin and Herman Rubin. A characterization of the Wishart distribution. *The Annals of Mathematical Statistics*, pages 1272–1280, 1961.

Jim Pitman and Marc Yor. A decomposition of Bessel bridges. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 59:425–457, 1982.

Philip E. Protter. *Stochastic Integration and Differential Equations.* Springer-Verlag Berlin Heidelberg, 2004.

Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion.* Springer-Verlag, 2001.

A. V. Skorohod. *Studies in the theory of random processes.* Addison-Wesley, 1965.

Robert Stelzer. *Multivariate Continuous Time Stochastic Volatility Models Driven by a Lévy Process.* PhD thesis, Centre for Mathematical Sciences, Munich University of Technology, 2007.

Steffen Timmann. *Repetitorium der Gewöhnlichen Differentialgleichungen.* Binomi, 2005.

T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11-1:155–167, 1971a.

T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations II. *J. Math. Kyoto Univ.*, 11-3:553–563, 1971b.