A Multivariate Ornstein-Uhlenbeck Type Stochastic Volatility Model

Christian Pigorsch* Robert Stelzer†

Using positive semidefinite processes of Ornstein-Uhlenbeck type a multivariate Ornstein-Uhlenbeck (OU) type stochastic volatility model is introduced. We derive many important statistical and probabilistic properties, e.g. the complete second order structure and a state-space representation. Noteworthy, many of our results are shown to be valid for the more general class of multivariate stochastic volatility models, which are driven by a stationary and square-integrable covariance matrix process. For the OU type stochastic volatility our results enable estimation and filtering of the volatility which we finally demonstrate with a short empirical illustration of our model.

Keywords: Barndorff-Nielsen-Shephard (BNS) stochastic volatility model; Lévy process; matrix subordinator; multivariate stochastic volatility; Ornstein-Uhlenbeck type process; state-space representation

1 Introduction

A wide range of different univariate continuous-time stochastic volatility models has been developed in the financial econometrics and statistics literature aiming at capturing the most distinct features of the price process of a single financial asset, see e.g. Barndorff-Nielsen and Shephard (2001) and Chernov, Gallant, Ghysels, and Tauchen (2003).

In contrast to this, the existing literature on multivariate (continuous-time) stochastic volatility models is rather small. One reason is that in a multivariate context modeling becomes even more challenging. In particular, apart from capturing the individual dynamics the model also needs to adequately reproduce the comovements and spill-over effects across different assets, as the knowledge of the dependence structure is crucial for financial decision-making, such as portfolio risk management, asset allocation or the pricing of multi-asset options, whose importance has increased tremendously in recent years. In addition to those requirements, there also arise some challenging mathematical issues in the multivariate setting. One is given by the necessity of a positive semidefinite covariance matrix. For continuous-time stochastic volatility models this implies that the instantaneous covariance should be specified as a positive semidefinite matrix process. Moreover, if the dimension of the return vector increases the number of parameters in

*Institute of Econometrics and Operations Research, Department of Economics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany. *Email:* `christian.pigorsch@uni-bonn.de`, `http://www.wiwi.uni-bonn.de/cpigorsch/`

†TUM Institute for Advanced Study & Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, D-85747 Garching, Germany. *Email:* `rstelzer@ma.tum.de`, `http://www-m4.ma.tum.de`

the model should not explode. Hence, a parsimonious but at the same time accurate and flexible specification is needed.

Given these challenges the theoretical literature on multivariate stochastic volatility models has just developed over the last few years, where the main focus was on discrete-time models as an alternative to the multivariate GARCH models, see e.g. Asai, McAleer, and Yu (2006) and Harvey, Ruiz, and Shephard (1994). A continuous-time specification, however, provides several advantages. In contrast to the discrete-time models a continuous-time specification allows to infer the implied dynamics and properties of the estimated model at various frequencies differing from the one used in the estimation. This is important, inter alia, for forecasting the covariance matrix over short term intervals, where the estimates can be based on less frequent data. Moreover, if irregularly spaced data is observed, the continuous-time specification is very advantageous. Likewise, continuous-time modeling is clearly preferable when it comes to derivative pricing.

Our general d-dimensional continuous-time stochastic volatility model is given by

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

where $(Y_t)_{t \in \mathbb{R}^+}$ denotes the d-dimensional logarithmic stock price process, $\mu, \beta \in \mathbb{R}^d$ are the instantaneous drift and risk premium parameters, respectively, $(W_t)_{t \in \mathbb{R}^+}$ denotes a d-dimensional standard Brownian motion, and $(\Sigma_t)_{t \in \mathbb{R}^+}$ is an adapted, stationary and square-integrable stochastic volatility process with values in the positive semidefinite matrices being independent of $(W_t)_{t \in \mathbb{R}^+}$. The usability and applicability of the stochastic volatility model mainly depends on the specification of the stochastic volatility process $(\Sigma_t)_{t \in \mathbb{R}^+}$. On the one hand the process has to reproduce the stylized facts of financial data, to fulfill the technical requirements mentioned above and should be general enough to describe a large number of different datasets adequately. On the other hand statistical inference of the model should be feasible, usually involving the derivation of specific properties of the model.

In this paper we propose a model where the stochastic volatility process $(\Sigma_t)_{t \in \mathbb{R}^+}$ is given by a Lévy-driven positive semidefinite Ornstein-Uhlenbeck (OU hereafter) type process

$$d\Sigma_t = (A\Sigma_{t-} + \Sigma_{t-}A^T)dt + dL_t, \tag{2}$$

which was recently introduced by Barndorff-Nielsen and Stelzer (2007). Hence, our model is a multivariate extension of the univariate non-Gaussian OU type stochastic volatility model proposed by Barndorff-Nielsen and Shephard (2001, 2002) and used and studied heavily since then (see e.g. Griffin and Steel, 2006, Rheinländer and Steiger, 2006, Roberts, Papaspiliopoulos, and Dellaportas, 2004 and references therein). We therefore call our model the "multivariate Ornstein-Uhlenbeck stochastic volatility model". We show that it is very flexible, while its estimation is nevertheless feasible and that the important stylized facts of financial data (stochastic volatility, jumps in volatility, heavy tails, dependence without correlation) are reproduced.

The model's flexibility is obtained via the specification of the mean reversion coefficient A and via the specification of the driving matrix subordinator $(L_t)_{t \in \mathbb{R}^+}$. We show that the finiteness of the moments of the stationary distribution of the OU type process is completely characterized by the matrix subordinator such that, for instance, we may have stationary distributions, which have finite moments of all orders, as well as such having only some finite moments, maybe not even a first finite moment. Moreover, it is also possible to consider matrix subordinators with elements jumping only together, jumping never together or a combination thereof. This allows for a variety of different types of correlation structures.

Importantly, despite its flexibility the model can also be estimated in several ways. In particular, we derive the conditional characteristic function of the joint process $(Y_t, \Sigma_t)_{t \in \mathbb{R}^+}$ given its initial values, the second order moments of the return process and the vectorized outer product of the returns (the "squared returns" in a multivariate setting) and a state-space representation for the joint process of the returns and their squares. Hence, estimation via the generalized method of moments, the quasi maximum likelihood method with the Kalman filter and via the characteristic function becomes feasible. The Kalman filter is particularly important, since it allows to filter the unobserved volatility process. Although this paper primarily focuses on the multivariate OU type stochastic volatility model we also derive the second order properties of the return process and its "square" for the d-dimensional continuous-time stochastic volatility model driven by a general stationary and square integrable stochastic process. This general model has been stated in several papers (e.g. Barndorff-Nielsen and Shephard, 2001, 2009, Barndorff-Nielsen, Nicolato, and Shephard, 2002 and Lindberg, 2005), but a detailed analysis of it has not been undertaken yet.

A further interesting property of our multivariate model is its implied marginal volatility dynamics (the behavior of the individual components of the positive definite OU type process) which is given by a linear combination (superposition) of (dependent) univariate OU type processes. This is an interesting new result, as several studies have shown that the high persistence in volatility can be reproduced using a superposition of short memory processes. As such the need to consider superposition in our multivariate model is not as important as in the univariate context. Nevertheless, we provide properties of the multivariate superposition model, such as second order moments and a state-space representation, which make statistical inference feasible.

The appearance of jumps in the volatility process, the amenability to estimation and the above mentioned flexibility regarding the matrix subordinator constitute the major differences of our model to the existing multivariate continuous-time stochastic volatility models. Among these, the Wishart model proposed by Gouriéroux (2006) and Gouriéroux, Jasiak, and Sufana (2009), is at least to our knowledge, the only model that also specifies the dynamics of the stochastic volatility process directly in the space of positive semidefinite matrices. Being the multivariate extension of the Cox-Ingersoll-Ross (CIR) process, their model shares the limitations and advantages of the univariate CIR model which are due to the specification of the driving processes as Brownian motions. The probably empirically most relevant shortcomings are that the unconditional distribution is restricted to Wishart distributions having exponential moments (and thus no heavy tails) and that there are no jumps in the volatility process. The importance and empirical evidence of jumps in volatility is discussed in e.g. Todorov and Tauchen (2008) and references therein. Furthermore, although at a first glance Brownian motion based models appear to be easier to understand mathematically than Lévy-driven ones, this is not necessarily the case here. For the stochastic differential equation defining our positive semidefinite OU type processes it is straightforward to establish with global Lipschitz and pathwise arguments that there exists a unique solution for all times $t \in [0, \infty)$ which is positive semi-definite at all times (and even strictly positive definite if the initial value is so). For Wishart processes this is a very intricate issue (see Bru, 1991) and unique strong solutions for all times are only known to exist if the parameter α (notation as in Bru, 1991, in Gouriéroux, 2006 this parameter is called K) is greater than the dimension plus one. Unfortunately this problem is not adequately mentioned in the above cited econometric literature using Wishart processes, where it is argued that, since fixed quadratic forms of the Wishart process cannot get negative almost surely, the Wishart process cannot leave the positive definite matrices and therefore exists for all positive values of α. Yet, one has to consider all possible quadratic forms. As these are uncountably many and uncountable unions of null sets do not have to be null sets, this argument is not rigorous.

Alternative ways to define multivariate stochastic volatility models make use of a factor structure, see e.g Hubalek and Nicolato (2009) and Lindberg (2005), who adopt an approach in which the volatility factors are independent and follow univariate positive non-Gaussian OU type processes. The flexibility of these models, however, is accompanied by the difficulty to achieve identification, which complicates the empirical application of these models. As the identifiability is especially important in the multivariate case, we provide natural conditions for our model, such that the parameters are uniquely identified via the second order properties of the returns.

Before outlining the structure of the paper we introduce our notation. Throughout this paper we write $\mathbb{R}^+$ for the positive real numbers including zero, $\mathbb{R}^{++}$ when zero is excluded and we denote the set of real $m \times n$ matrices by $M_{m,n}(\mathbb{R})$. If $m = n$ we simply write $M_n(\mathbb{R})$ and denote the group of invertible $n \times n$ matrices by $GL_n(\mathbb{R})$, the linear subspace of symmetric matrices by $\mathbb{S}_n$, the (closed) positive semidefinite cone by $\mathbb{S}_n^+$ and the open (in $\mathbb{S}_n$) positive definite cone by $\mathbb{S}_n^{++}$. I_n stands for the $n \times n$ identity matrix, $\sigma(A)$ for the spectrum (the set of all eigenvalues) of a matrix $A \in M_n(\mathbb{R})$ and $\rho(A)$ for its spectral radius. The natural ordering on the symmetric $n \times n$ matrices is denoted by $\leq$, i.e. for $A, B \in \mathbb{S}_n$ we have that $A \leq B$, if and only if $B - A \in \mathbb{S}_n^+$. The tensor (Kronecker) product of two matrices A, B is written as $A \otimes B$. vec denotes the well-known vectorization operator that maps the $n \times n$ matrices to $\mathbb{R}^{n^2}$ by stacking the columns of the matrices below one another. For more information regarding the tensor product and vec operator we refer to Horn and Johnson (1991, Chapter 4). Likewise vech : $\mathbb{S}_n \to \mathbb{R}^{n(n+1)/2}$ denotes the "vector-half" operator that stacks the columns of the lower triangular part of a symmetric matrix below another. Finally, A^T is the transpose of a matrix $A \in M_n(\mathbb{R})$. For a matrix A we denote by A_{ij} the element in the i-th row and j-th column and this notation is extended to processes in a natural way.

Regarding all random variables and processes we assume that they are defined on a given appropriate filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$ satisfying the usual hypotheses. For random functions and measures we usually do not state the dependence on $\omega \in \Omega$ explicitly.

Norms of vectors or matrices are denoted by $\|\cdot\|$. If the norm is not specified, then it is irrelevant which particular norm is used.

Furthermore, we employ an intuitive notation with respect to the (stochastic) integration with a matrix-valued integrator referring to any of the standard texts (e.g. Protter, 2005 or Øksendal, 2003 regarding Brownian motion) for a comprehensive treatment of the theory of stochastic integration. Let $(L_t)_{t \in \mathbb{R}^+}$ in $M_{n,r}(\mathbb{R})$ be a semimartingale and $(A_t)_{t \in \mathbb{R}^+}$ in $M_{m,n}(\mathbb{R})$, $(B_t)_{t \in \mathbb{R}^+}$ in $M_{r,s}(\mathbb{R})$ be adapted integrable (w.r.t. L) processes. Then we denote by $\int_0^t A_s dL_s B_s$ the matrix C_t in $M_{m,s}(\mathbb{R})$ which has ij-th element $C_{ij,t} = \sum_{k=1}^n \sum_{l=1}^r \int_0^t A_{ik,s} B_{lj,s} dL_{kl,s}$. Equivalently such an integral can be understood in the sense of Métivier and Pellaumail (1980), resp. Métivier (1982), by identifying it with the integral $\int_0^t \mathbf{A}_s dL_s$ with $\mathbf{A}_t$ being for each fixed t the linear operator $M_{n,r}(\mathbb{R}) \to M_{m,s}(\mathbb{R})$, $X \mapsto A_t X B_t$. Moreover, we always denote by $\int_a^b$ with $a \in \mathbb{R} \cup \{-\infty\}, b \in \mathbb{R}$ the integral over the half-open interval $(a, b]$ for notational convenience. If $b = \infty$ the integral is understood to be over (a, b). The function $\log^+$ is defined as $\max(\log(x), 0)$ and $\imath = \sqrt{-1}$ is the imaginary unit.

The remainder of the paper is structured as follows. As our model builds on the positive semidefinite matrix process of OU type, Section 2 presents a brief review of them and derives further interesting properties. Section 3 introduces our multivariate OU type stochastic volatility model, derives important properties and provides a detailed model analysis. Section 4 presents an empirical illustration, and Section 5 finally concludes.

2 Positive Semidefinite Matrix Processes of Ornstein-Uhlenbeck Type

In this section we first briefly review the positive semidefinite OU type processes introduced in Barndorff-Nielsen and Stelzer (2007) where detailed proofs have been given. Analyzing this class of processes in more detail, we study some further properties, which turn out to have important implications for our multivariate OU type stochastic volatility model.

2.1 Definition and Probabilistic Properties

The construction of positive semidefinite OU type processes builds on a special type of matrix-valued Lévy process studied in detail in Barndorff-Nielsen and Pérez-Abreu (2008). For the relevant background on Lévy processes we refer to any of the standard references, for instance, Sato (1999).

Definition 2.1. *An $\mathbb{S}_d$-valued Lévy process $L = (L_t)_{t\in\mathbb{R}^+}$ is said to be a* matrix subordinator*, if $L_t - L_s \in \mathbb{S}_d^+$ for all $s, t \in \mathbb{R}^+$ with $t > s$.*

The characteristic function of a matrix subordinator L in $\mathbb{S}_d^+$ at time $t \in \mathbb{R}^+$ is given by

$$\mu_{L_t}(Z) = \exp\left(t\left(i\mathrm{tr}(\gamma_L Z) + \int_{\mathbb{S}_d^+\backslash\{0\}} (e^{i\mathrm{tr}(XZ)} - 1)\nu_L(dX)\right)\right), \ Z \in \mathbb{S}_d,$$

where ν_L is a Lévy measure on $\mathbb{S}_d^+\backslash\{0\}$ satisfying $\int_{\|x\|\le 1} \|x\|\nu_L(dx) < \infty$, $\gamma_L \in \mathbb{S}_d^+$ is referred to as the drift and tr denotes the trace of a matrix.

Matrix subordinators are a generalization of the concept of univariate Lévy subordinators to the matrix case, in particular, they are simply the same as Lévy subordinators for $d = 1$. As in the univariate case there are a lot of very different concrete examples of matrix subordinators. Barndorff-Nielsen and Pérez-Abreu (2008), for example, discuss matrix subordinators which are generalizations of stable, tempered stable and Gamma subordinators. These are examples having infinite activity. Of course, compound Poisson, i.e. finite activity, matrix subordinators can easily be constructed using any probability distribution on $\mathbb{S}_d^+$ (see e.g. Gupta and Nagar, 2000 for some examples) for the jumps. In this context it should be noted that the outer product of any vector random variable is positive semidefinite. Specifying the diagonal elements of the matrix process as (possibly dependent) univariate subordinators forming together a d-dimensional Lévy process and setting the off-diagonal elements to zero leads to another simple example of a matrix subordinator (referred to as a diagonal matrix subordinator). For a more detailed discussion of some specific matrix subordinators and their covariance structure we also refer to Pigorsch and Stelzer (2009, Section 5).

The existence of OU type processes assuming values in the positive semidefinite matrices is ensured by the following theorem, where the Lévy process $(L_t)_{t\in\mathbb{R}^+}$ is extended to a Lévy process $(L_t)_{t\in\mathbb{R}}$ starting in the infinite past in the usual way.

Theorem 2.2 (Barndorff-Nielsen and Stelzer (2007, Theorem 4.5))**.** *Let L be a matrix subordinator with $E(\log^+ \|L_1\|) < \infty$ and $A \in M_d(\mathbb{R})$ such that $\sigma(A) \subset (-\infty, 0) + i\mathbb{R}$. Then the stochastic differential equation of OU type*

$$d\Sigma_t = (A\Sigma_{t-} + \Sigma_{t-}A^T)dt + dL_t$$

has a unique stationary solution

$$\Sigma_t = \int_{-\infty}^t e^{A(t-s)} dL_s e^{A^T(t-s)}$$

or, in vectorial representation,

$$\mathrm{vec}(\Sigma_t) = \int_{-\infty}^t e^{(I_d\otimes A + A\otimes I_d)(t-s)} d\mathrm{vec}(L_s).$$

Moreover, $\Sigma_t \in \mathbb{S}_d^+$ for all $t \in \mathbb{R}$.

The second order properties below as well as the characteristic function of this OU type process have been derived in Barndorff-Nielsen and Stelzer (2007).

Proposition 2.3 (Barndorff-Nielsen and Stelzer (2007, Proposition 4.7))**.** *Assume that the driving Lévy process is square-integrable. Then the second order moment structure is given by*

$$E(\Sigma_t) = -\mathbf{A}^{-1}E(L_1), \qquad \mathrm{var}(\mathrm{vec}(\Sigma_t)) = -\mathcal{A}^{-1}\mathrm{var}(\mathrm{vec}(L_1)) \tag{3}$$

$$\mathrm{cov}(\mathrm{vec}(\Sigma_{t+h}), \mathrm{vec}(\Sigma_t)) = e^{(A\otimes I_d + I_d \otimes A)h} \mathrm{var}(\mathrm{vec}(\Sigma_t)), \tag{4}$$

where $t \in \mathbb{R}$, $h \in \mathbb{R}^+$, $\mathbf{A}$ is the linear operator $\mathbf{A} : M_d(\mathbb{R}) \to M_d(\mathbb{R})$, $X \mapsto AX + XA^T$ which can be represented as $\mathrm{vec}^{-1} \circ ((I_d \otimes A) + (A \otimes I_d)) \circ \mathrm{vec}$ and $\mathcal{A} : M_{d^2}(\mathbb{R}) \to M_{d^2}(\mathbb{R})$, $X \mapsto (A \otimes I_d + I_d \otimes A)X + X(A^T \otimes I_d + I_d \otimes A^T)$. The linear operator $\mathcal{A}$ can be represented as

$$\mathrm{vec}^{-1} \circ ((I_{d^2} \otimes (A \otimes I_d + I_d \otimes A)) + ((A \otimes I_d + I_d \otimes A) \otimes I_{d^2})) \circ \mathrm{vec}.$$

In general the finiteness of the moments of the stationary distribution of the OU type process is completely characterized by the driving Lévy process.

Proposition 2.4. *Let $(\Sigma_t)_{t\in\mathbb{R}}$ be a strictly stationary OU type process in $\mathbb{S}_d^+$ with driving matrix subordinator L which has Lévy measure ν_L and be $r \in \mathbb{R}^{++}$. Then $E(\|\Sigma_0\|^r) < \infty$, if and only if $E(\|L_1\|^r) < \infty$ or equivalently $\int_{\mathbb{S}_d^+, \|x\|\geq 1} \|x\|^r \nu_L(dx) < \infty$.*

Proof. Follows by a straightforward adaptation of the proof of Marquardt and Stelzer (2007, Proposition 3.30) to the matrix case. □

By choosing appropriate Lévy processes, one obtains thus very different possible tail behaviors for the volatility process Σ.

Noteworthy, the finiteness of some moment of the Lévy process also ensures that the stationary OU type process exhibits a very nice dependence structure. For the definition of the relevant mixing properties see Davydov (1973) or Doukhan (1994), for instance.

Proposition 2.5. *Let Σ be an OU type process in $\mathbb{S}_d^+$. Then Σ is a temporally homogeneous strong Markov process.*

If Σ is stationary and the driving Lévy process L with Lévy measure ν_L satisfies additionally $\int_{\mathbb{S}_d^+, \|x\|\geq 1} \|x\|^r \nu_L(dx) < \infty$ for some $r > 0$, then the stationary OU type process Σ is β-mixing with mixing coefficients $\beta_l = O(e^{-al})$ for some $a > 0$. In particular, Σ is strongly (or α-)mixing with exponential rate and ergodic.

Proof. Follows from Protter (2005, Theorem V.32) and Masuda (2004, Theorem 4.3). □

From a financial point of view the integrated process is of major importance, as it corresponds to the integrated volatility, which is a main variable of interest in financial applications.

Proposition 2.6 (Barndorff-Nielsen and Stelzer (2007, Proposition 4.10))**.** *Let Σ be a positive semidefinite OU type process with initial value $\Sigma_0 \in \mathbb{S}_d^+$ and driven by the Lévy process L. Then the integrated OU type process Σ^+ is given by*

$$\Sigma_t^+ := \int_0^t \Sigma_t dt = \mathbf{A}^{-1} (\Sigma_t - \Sigma_0 - L_t)$$

for $t \in \mathbb{R}^+$, where $\mathbf{A}$ is the linear operator defined in Proposition 2.3.

2.2 Marginal Dynamics

Deriving the marginal dynamics, i.e. the behavior of the individual components $\Sigma_{ij} = (\Sigma_{ij,t})_{t\in\mathbb{R}^+}$ of a positive semidefinite OU type process Σ, deepens our understanding of these processes and facilitates especially the comparison with the univariate OU type processes. To this end, we assume that A is real diagonalizable and $\sigma(A) = \{\lambda_1, \ldots, \lambda_d\}$. Let $U \in GL_d(\mathbb{R})$ be such that

$$UAU^{-1} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_d \end{pmatrix} := D. \tag{5}$$

Denoting $(U^T)^{-1} = (U^{-1})^T$ by U^{-T}, it follows that

$$\Sigma_t \quad = \quad U^{-1}\left(\int_{-\infty}^{t} e^{D(t-s)} d(UL_sU^T) e^{D^T(t-s)}\right) U^{-T}. \tag{6}$$

Defining $\widetilde{L}_t = UL_tU^T$ for $t \in \mathbb{R}$ we have that $\widetilde{L}$ is again a Lévy process in $M_d(\mathbb{R})$, or more specifically it is even a matrix subordinator. Moreover, one obtains that

$$\left(\int_{-\infty}^{t} e^{D(t-s)} d\widetilde{L}_s e^{D^T(t-s)}\right)_{ij} = \int_{-\infty}^{t} e^{(\lambda_i+\lambda_j)(t-s)} d\widetilde{L}_{ij,s} \text{ for } i,j = 1,\ldots,d,$$

which obviously shows that the individual components of $U\Sigma_tU^T =: \widetilde{\Sigma}_t$ are stationary one-dimensional Ornstein-Uhlenbeck type processes with associated stochastic differential equations

$$d\widetilde{\Sigma}_{ij,t} = (\lambda_i + \lambda_j)\widetilde{\Sigma}_{ij,t}dt + d\widetilde{L}_{ij,t}. \tag{7}$$

Note further that $\widetilde{L}_{ii}$ for $1 \le i \le d$ is necessarily a subordinator and $\widetilde{\Sigma}_{ii}$ has to be a positive OU type process.

Together with (6) the above considerations show that the individual components Σ_{ij} of Σ are superpositions of (at most d^2) univariate OU type processes. However, unlike in the univariate superposition model, see Barndorff-Nielsen and Shephard (2001), the individual OU-processes superimposed are in general not independent. Actually, they can only be independent when the Lévy measure of $\widetilde{L}$ is concentrated on the diagonal matrices.

With the obvious modifications the above results hold also true for general diagonalizable $A \in M_d(\mathbb{R})$. Then X^T has to be replaced by the Hermitian of a matrix $X \in M_d(\mathbb{C})$ and $\widetilde{\Sigma}$ is an OU type processes in the positive semidefinite complex matrices. Note that $\widetilde{\Sigma}_{ii}$ still have to be real (even positive) and $\widetilde{L}_{ii}$ a real subordinator. Furthermore, (7) becomes $d\widetilde{\Sigma}_{ij,t} = (\lambda_i + \overline{\lambda}_j)\widetilde{\Sigma}_{ij,t}dt + d\widetilde{L}_{ij,t}$.

This result adds important insight regarding the behavior of the autocovariance functions of the volatility of the individual assets. In particular, in order to obtain a more realistic decay (compared to using a single univariate OU type process in a univariate model) of these functions it is no longer necessary to consider superpositions of different OU type processes. So, although it is possible to build superpositions of positive semidefinite OU type processes (see Section 3.5), we expect them to be less important for financial applications as in the univariate case, where it has been shown that sufficiently realistic patterns of the autocorrelation functions can only be obtained by superpositions of OU type processes. As such, the multivariate specification obviously introduces more flexibility.

3 The Multivariate Ornstein-Uhlenbeck type Stochastic Volatility Model

Based on the above results for the positive semidefinite OU type process we can now introduce our multivariate stochastic volatility model. The general d-dimensional stochastic volatility stock price model is given by

$$dY_t = (\mu + \Sigma_t\beta)dt + \Sigma_t^{1/2}dW_t, \; Y_0 = 0, \tag{8}$$

where $(Y_t)_{t\in\mathbb{R}^+}$ denotes the d-dimensional logarithmic stock price process, $\mu, \beta \in \mathbb{R}^d$ are the drift and so-called risk premium parameters, respectively, $(W_t)_{t\in\mathbb{R}^+}$ denotes a d-dimensional standard Brownian motion and $(\Sigma_t)_{t\in\mathbb{R}^+}$ is an adapted, stationary and square integrable stochastic process with values in $\mathbb{S}_d^+$ and independent of $(W_t)_{t\in\mathbb{R}^+}$. Moreover, $(t,\omega) \mapsto \Sigma_t(\omega)$ is assumed to be $\mathcal{B}(\mathbb{R}^+) \times \mathcal{F}$-measurable with $\mathcal{B}(\mathbb{R}^+)$ denoting the Borel-σ-algebra over $\mathbb{R}^+$. As common in the finance literature, $(\Sigma_t)_{t\in\mathbb{R}^+}$ represents the stochastic volatility or instantaneous covariance process.

In this paper we mainly focus on a specification in which the volatility process is given by a Lévy-driven positive semidefinite OU type process where the driving Lévy process $(L_t)_{t\in\mathbb{R}^+}$ and the Brownian Motion of the price process are independent. We refer to this model as the "multivariate Ornstein-Uhlenbeck stochastic volatility model". However, whenever possible we state our results for the general model given in (8).

Furthermore we presume $Y_0 = 0$, which is no real constraint as it just corresponds to a normalization of the prices at time zero. In the OU type stochastic volatility model we extend the driving Lévy process to one defined on the whole real line and write

$$\Sigma_t = \int_{-\infty}^{t} e^{A(t-s)} dL_s e^{A^T(t-s)}. \tag{9}$$

Note that this corresponds to starting the OU type process at time zero with Σ_0 having the stationary distribution and being independent of $(L_t)_{t\in\mathbb{R}^+}$.

The subsequent returns over time intervals of length $\Delta \in \mathbb{R}^{++}$ are denoted by $\mathbf{Y} = (\mathbf{Y}_n)_{n\in\mathbb{N}}$. In many financial applications this time interval, i.e. $[(n-1)\Delta, n\Delta]$ with $n \in \mathbb{N}$, will represent a trading day, for example. So, the logarithmic price increments are defined by

$$\mathbf{Y}_n := Y_{n\Delta} - Y_{(n-1)\Delta} = \int_{(n-1)\Delta}^{n\Delta} (\mu + \Sigma_t\beta)dt + \int_{(n-1)\Delta}^{n\Delta} \Sigma_t^{1/2} dW_t, \;\; n \in \mathbb{N}.$$

As already stated in Barndorff-Nielsen and Shephard (2001) it is easy to see that

$$\mathbf{Y}_n \,|\mathbf{\Sigma}_n \sim N_d\left(\mu\Delta + \mathbf{\Sigma}_n\beta, \mathbf{\Sigma}_n\right), \;\text{ where }\; \mathbf{\Sigma}_n := \int_{(n-1)\Delta}^{n\Delta} \Sigma_t dt = \Sigma_{n\Delta}^+ - \Sigma_{(n-1)\Delta}^+ \tag{10}$$

is the integrated volatility over the unit time interval and $N_d(m,s)$ denotes the d-dimensional normal distribution with mean m and covariance matrix s.

Note that – like in the univariate model – the multivariate OU type stochastic volatility model can easily be extended to account for the leverage effect by specifying

$$dY_t = (\mu + \Sigma_t\beta)dt + \Sigma_t^{1/2} dW_t + \psi dL_t,$$

with ψ being a linear operator from $\mathbb{S}_d$ to $\mathbb{R}^d$. This is a straightforward generalization of the univariate OU type models with leverage effect. However, as the derivation of the properties of the OU type models is markedly complicated by the inclusion of a leverage effect – even in the univariate case, where only very little is known (see Barndorff-Nielsen and Shephard, 2001) –, we solely focus here on the model without leverage effect. Nevertheless, just as for the univariate (OU type) stochastic volatility models with leverage a simulation based Bayesian analysis may be a promising estimation strategy here, see e.g Griffin and Steel (2006) and Omori, Chib, Shephard, and Nakajima (2007). As the present paper introduces the model and its properties, we leave such a Bayesian analysis for future research and rather prefer to focus on more straightforward and easy to implement estimators. Moreover, stochastic volatility models are also heavily used for other assets, e.g. exchange rates like in Section 4, where the leverage is not an issue.

3.1 Characteristic Function of the Markov Transition Kernels

Using general results on SDEs (see e.g. Protter, 2005, Ch. V), the process $(Y_t, \Sigma_t)_{t\in\mathbb{R}^+}$ in the positive semidefinite OU type SV model is easily seen to be a strong Markov process on $\mathbb{R} \times \mathbb{S}_d^+$ having the weak Feller property. The explicit representation of the integrated volatility allows us to explicitly compute the characteristic function of the Markov transition kernels. To this end we use the scalar product $< \cdot, \cdot >$ on $\mathbb{R}^d \times M_d(\mathbb{R})$ given by $(\mathbb{R}^d \times M_d(\mathbb{R}))^2 \to \mathbb{R}$, $(x_1, y_1), (x_2, y_2) \mapsto x_2^T x_1 + \mathrm{tr}(y_2^T y_1)$ (which can be extended to a complex setting by replacing the transpositions with taking the Hermitians). Note moreover that the positive semidefinite OU type volatility process alone is also Markovian and characterizations of its Markov transition kernels can be obtained by straightforward adaptations of the results of Sato and Yamazato (1984).

Due to technical reasons we sometimes need to resort to complex matrices in the following proposition. For a complex matrix X we will denote its Hermitian by X^* and for a linear operator Z we will likewise denote its adjoint operator by Z^*.

Proposition 3.1. *Consider the positive semidefinite OU type SV model and assume the driving matrix subordinator L has characteristic exponent ψ_L, i.e. $E(e^{i\mathrm{tr}(L_t z)}) = e^{t\psi_L(z)}$ for all $z \in M_d(\mathbb{R}) + i\mathbb{S}_d^+$. Let $(Y_0, \Sigma_0) \in \mathbb{R}^d \times \mathbb{S}_d^+$. Then the conditional characteristic function of (Y_t, Σ_t) given (Y_0, Σ_0) is for every $t \in \mathbb{R}^+$ and $(y, z) \in \mathbb{R}^d \times M_d(\mathbb{R})$*

$$
\begin{aligned}
E\left(e^{i\langle (y,z),(Y_t,\Sigma_t)\rangle} \,\middle|\, \Sigma_0, Y_0 \right) &= \exp\Big\{ i(Y_0 + \mu t)^T y + i\mathrm{tr}\left(\Sigma_0 e^{A^T t} z e^{At} \right) \\
&+ i\mathrm{tr}\left(\Sigma_0 e^{A^T t}\left[\mathbf{A}^{-*}\left(y\beta^T + \frac{i}{2} y y^T \right) \right] e^{At} - \Sigma_0 \left[\mathbf{A}^{-*}\left(y\beta^T + \frac{i}{2} y y^T \right) \right] \right) \\
&+ \int_0^t \psi_L\left(e^{A^T s} z e^{As} + e^{A^T s}\left[\mathbf{A}^{-*}\left(y\beta^T + \frac{i}{2} y y^T \right) \right] e^{As} - \mathbf{A}^{-*}\left(y\beta^T + \frac{i}{2} y y^T \right) \right) ds \Big\}
\end{aligned} \tag{11}
$$

with $\mathbf{A}^{-*}$ denoting the inverse of the adjoint of $\mathbf{A}$, i.e. $\mathbf{A}^{-*}$ is the inverse of the linear operator $\mathbf{A}^*$ given by $X \mapsto A^T X + XA$.

Proof. The proof of Proposition 3.1 is given in Appendix A.1. □

$E(e^{i\mathrm{tr}(L_t z)})$ is well-defined for all $z \in M_d(\mathbb{R}) + i\mathbb{S}_d^+$, since L_t is concentrated on $\mathbb{S}_d^+$ and $\mathrm{tr}(xz) \geq 0$ for all $x, z \in \mathbb{S}_d^+$. This function extends the usual characteristic function and is often referred to as the Fourier-Laplace transform. Note that $M_d(\mathbb{R}) + i\mathbb{S}_d^+$ needs not be the maximal set on which it can be defined, but it suffices for our purposes and it does not depend on the matrix subordinator used.

The importance of the above result is, of course, that it completely characterizes the conditional distribution of the log-returns and their stochastic volatility given the initial values. Moreover, the above conditional characteristic function is obviously exponentially affine in (Y_0, Σ_0). The exponential affinity of conditional characteristic functions is one of the equivalent definitions for affine processes in $\mathbb{R}^m \times (\mathbb{R}^+)^n$ in Duffie, Filipović, and Schachermayer (2003, Def. 2.1). So intuitively (Y, Σ) is an affine process, but, strictly speaking, not in the sense of Duffie et al. (2003, Def. 2.1) due to a different state-space. Whether the results of Duffie et al. (2003, Def. 2.1) can be extended to cover our setting, appears to be an intricate question beyond the scope of the present paper.

Under sufficient technical conditions the above result should be extendible to the existence of the Fourier-Laplace transform in a neighborhood of zero, which allows for derivative pricing via Fourier-Laplace inversion techniques. This is currently under investigation and will be reported elsewhere.

3.2 Second Order Structure

In this section we study the second order moments of the multivariate stochastic volatility model. The results to be established turn out to be very useful for the estimation of our model as well as for forecasting. Additionally, they show that the model reflects the stylized fact of dependent returns with vanishing autocorrelations

Henceforth we make the following assumption:

Assumption 3.1. *The stationary stochastic volatility process $(\Sigma_t)_{t\in\mathbb{R}^+}$ has a finite second moment.*

Before moving on to the second order properties of our model, we first introduce some notation regarding the autocovariance function. If $(X_t)_{t\in\mathbb{T}}$ (with $\mathbb{T}$ being either $\mathbb{N}_0$ or $\mathbb{R}^+$) is a second order stationary process with values in $\mathbb{R}^d$, the autocovariance function $\mathrm{acov}_X : \mathbb{T}\cup(-\mathbb{T}) \mapsto M_d(\mathbb{R})$ of X is given by $\mathrm{acov}_X(h) = \mathrm{cov}(X_h, X_0) = E(X_h X_0^T) - E(X_0)E(X_0)^T$ for $h \geq 0$ and by $\mathrm{acov}_X(h) = \mathrm{acov}_X(-h)^T$ for $h < 0$. If $(X_t)_{t\in\mathbb{T}}$ is a second order stationary process with values in $M_d(\mathbb{R})$ (or $\mathbb{S}_d$), then we set $\mathrm{acov}_X := \mathrm{acov}_{\mathrm{vec}(X)}$. As the twice integrated autocovariance function of the stationary volatility process Σ will be of particular importance, we define

$$r^+(t) := \int_0^t \mathrm{acov}_\Sigma(u)du \quad \text{and} \quad r^{++}(t) := \int_0^t r^+(u)du. \tag{12}$$

Theorem 3.2. *For the* ***general stochastic volatility model*** *with $(\Sigma_t)_{t\in\mathbb{R}^+}$ being stationary and square-integrable it holds that the increments of the integrated volatility $(\mathbf{\Sigma}_n)_{n\in\mathbb{N}}$ are stationary and square-integrable. We have:*

$$E(\mathbf{\Sigma}_1) = \Delta E(\Sigma_0), \quad \mathrm{var}(\mathrm{vec}(\mathbf{\Sigma}_1)) = r^{++}(\Delta) + r^{++}(\Delta)^T, \tag{13}$$

$$\mathrm{acov}_{\mathbf{\Sigma}}(h) = r^{++}(h\Delta+\Delta) - 2r^{++}(h\Delta) + r^{++}(h\Delta-\Delta),\ h \in \mathbb{N}. \tag{14}$$

Likewise the discretely observed log-price increments $(\mathbf{Y}_n)_{n\in\mathbb{N}}$ are stationary and square-integrable with

$$E(\mathbf{Y}_1) = (\mu + E(\Sigma_0)\beta)\Delta, \quad \mathrm{var}(\mathbf{Y}_1) = E(\Sigma_0)\Delta + (\beta^T \otimes I_d)\mathrm{var}(\mathrm{vec}(\mathbf{\Sigma}_1))(\beta \otimes I_d) \tag{15}$$

$$\mathrm{acov}_{\mathbf{Y}}(h) = (\beta^T \otimes I_d)\mathrm{acov}_{\mathbf{\Sigma}}(h)(\beta \otimes I_d),\ h \in \mathbb{N}. \tag{16}$$

If Σ is a ***positive semidefinite OU type process*** *with driving matrix subordinator L, then*

$$E(\Sigma_0) = -\mathbf{A}^{-1}E(L_1) \tag{17}$$

$$r^{++}(t) = -\left(\mathscr{A}^{-2}\left(e^{\mathscr{A}t} - I_{d^2}\right) - \mathscr{A}^{-1}t\right)\mathcal{A}^{-1}\mathrm{var}(\mathrm{vec}(L_1)) \tag{18}$$

$$\mathrm{acov}_{\mathbf{\Sigma}}(h) = -e^{\mathscr{A}\Delta(h-1)}\mathscr{A}^{-2}\left(I_{d^2} - e^{\mathscr{A}\Delta}\right)^2\mathcal{A}^{-1}\mathrm{var}(\mathrm{vec}(L_1)),\ h \in \mathbb{N}, \tag{19}$$

where $\mathbf{A}$ and $\mathcal{A}$ are defined in Proposition 2.3 and $\mathscr{A} := A \otimes I_d + I_d \otimes A$. Observe that $\mathscr{A}$ and $\mathcal{A}$ commute (as linear operators over $M_{d^2}(\mathbb{R})$).

Proof. The proof of Theorem 3.2 is given in Appendix A.2. □

The above formulae imply that for $\beta = 0$ the log-price increments $(\mathbf{Y}_n)_{n\in\mathbb{N}}$ form an uncorrelated sequence and are thus white noise (in the second order sense).

Note that for a second order stationary causal m-dimensional ARMA(1,1) process $(X_t)_{t\in\mathbb{Z}}$ given by $X_t - \Phi X_{t-1} = Z_t + \Theta Z_{t-1}$ with $\Phi, \Theta \in M_m(\mathbb{R})$ and $(Z_t)_{t\in\mathbb{Z}}$ being m-dimensional white noise with covariance matrix Σ_Z the following autocovariance function can be obtained using the general formulae of Brockwell and Davis (1991, p. 420):

$$\mathrm{acov}_X(0) = \Sigma_Z + (I_m - B)^{-1}(\Phi+\Theta)\Sigma_Z(\Phi+\Theta)^T$$

$$\mathrm{acov}_X(1) = (\Phi + \Theta)\Sigma_Z + \Phi(I_m - B)^{-1}(\Phi + \Theta)\Sigma_Z(\Phi + \Theta)^T$$
$$\mathrm{acov}_X(h) = \Phi^{h-1}\mathrm{acov}_X(1),\ \ h \geq 1,$$

where $B : M_m(\mathbb{R}) \to M_m(\mathbb{R})$, $X \mapsto \Phi X \Phi^T$. Since we consider a stationary causal ARMA process, $\rho(\Phi) < 1$. Hence, as $\mathrm{vec}(BX) = (\Phi \otimes \Phi)\mathrm{vec}(X)$, it is obvious that $\rho(B) < 1$ and thus $I_m - B$ is invertible.

Therefore, comparing equation (19) with the general autocovariance function of an ARMA(1,1) process, immediately reveals that in the positive semidefinite OU type stochastic volatility model the process $(\mathrm{vec}(\mathbf{\Sigma}_n))_{n\in\mathbb{N}}$ is an ARMA(1,1) process with autoregressive parameter $e^{\mathscr{A}\Delta}$.

Moreover, since we assume $\sigma(A) \subset (-\infty, 0) + i\mathbb{R}$, we have from Horn and Johnson (1991, Theorem 4.4.5) that $\sigma(\mathscr{A}) \subset (-\infty, 0) + i\mathbb{R}$ and thus all elements of $\sigma(e^{\mathscr{A}\Delta})$ are less than one in modulus, which implies that this ARMA(1,1) process is causal.

The ARMA(1,1) structure of $\mathbf{\Sigma}$ might seem to provide a natural starting point for making inference on the OU type stochastic volatility model. However, usually $\mathbf{\Sigma}$ is unobservable and so inference can only be based on the observed returns $\mathbf{Y}$. But the second order structure of the returns obviously does not allow for an in-depth analysis of the latent stochastic volatility model. Yet, the squared log-price increments $\mathbf{Y}\mathbf{Y}^T := (\mathbf{Y}_n\mathbf{Y}_n^T)_{n\in\mathbb{N}}$ are not only observable, but also exhibit a useful second order structure.

Theorem 3.3. *In the* ***general stochastic volatility model*** *with $\mu = \beta = 0$ the second order structure of the squared log returns $(\mathbf{Y}_n\mathbf{Y}_n^T)_{n\in\mathbb{N}}$ is given by*

$$E(\mathbf{Y}_1\mathbf{Y}_1^T) = \mathrm{var}(\mathbf{Y}_1) + E(\mathbf{Y}_1)E(\mathbf{Y}_1^T) = E(\Sigma_0)\Delta \tag{20}$$
$$\mathrm{var}(\mathrm{vec}(\mathbf{Y}_1\mathbf{Y}_1^T)) = (I_{d^2} + \mathbf{Q} + \mathbf{PQ})\left(r^{++}(\Delta) + r^{++}(\Delta)^T\right) + (I_{d^2} + \mathbf{P})\left(E(\Sigma_0) \otimes E(\Sigma_0)\right)\Delta^2 \tag{21}$$
$$\mathrm{acov}_{\mathbf{Y}\mathbf{Y}^T}(h) = \mathrm{acov}_{\mathbf{\Sigma}}(h)\ \ \textit{for}\ \ h \in \mathbb{N} \tag{22}$$

where

$$\mathbf{P} : M_{d^2}(\mathbb{R}) \to M_{d^2}(\mathbb{R}),\ (\mathbf{P}X)_{i,(p-1)d+q} = X_{i,(q-1)d+p}\ \textit{for all}\ i = \{1, 2, \ldots, d^2\},\ p, q = \{1, 2, \ldots d\}$$
$$\mathbf{Q} : M_{d^2}(\mathbb{R}) \to M_{d^2}(\mathbb{R}),\ (\mathbf{Q}X)_{(k-1)d+l,(p-1)d+q} = X_{(k-1)d+p,(l-1)d+q}\ \textit{for all}\ \ k, l, p, q = \{1, 2, \ldots d\}$$

are linear operators. Obviously $\mathbf{P}^{-1} = \mathbf{P}$, $\mathbf{Q}^{-1} = \mathbf{Q}$ and $\mathbf{P}$ is representable as $X \mapsto XP$ with $P \in M_{d^2}(\mathbb{R})$ being a permutation matrix. Moreover, $\mathbf{Q}(\mathrm{vec}(X)\mathrm{vec}(Z)^T) = X \otimes Z$ for all $X, Z \in \mathbb{S}_d$.

Component-wise we have for the variance

$$\begin{aligned}\mathrm{cov}\left(\mathbf{Y}_{i,1}\mathbf{Y}_{j,1}, \mathbf{Y}_{k,1}\mathbf{Y}_{l,1}\right) &= \mathrm{var}(\mathrm{vec}(\mathbf{Y}_1\mathbf{Y}_1^T))_{(j-1)d+i,(l-1)d+k} \\ &= \int_0^\Delta \int_0^z \left(\mathrm{cov}(\Sigma_{ij,z}, \Sigma_{kl,u}) + \mathrm{cov}(\Sigma_{ij,u}, \Sigma_{kl,z})\right) du dz \\ &\quad + \int_0^\Delta \int_0^z \left(E(\Sigma_{jl,z}\Sigma_{ik,u}) + E(\Sigma_{jl,u}\Sigma_{ik,z}) + E(\Sigma_{jk,z}\Sigma_{il,u}) + E(\Sigma_{jk,u}\Sigma_{il,z})\right) du dz\end{aligned} \tag{23}$$

In the OU type stochastic volatility model $(\mathrm{vec}(\mathbf{Y}_n\mathbf{Y}_n^T))_{n\in\mathbb{N}}$ is thus a causal ARMA(1,1) process with autoregressive parameter $e^{\mathscr{A}\Delta}$.

Proof. The proof of Theorem 3.3 is given in Appendix A.3. □

Note that the ARMA(1,1) structure of $\mathrm{vec}(\mathbf{Y}\mathbf{Y}^T)$, of course, means that $\mathbf{Y}\mathbf{Y}^T$ itself is an ARMA(1,1) process. Its autoregressive coefficient is given by the linear operator $\mathbb{S}_d \to \mathbb{S}_d$, $X \mapsto e^{A\Delta}Xe^{A\Delta}$.

In order to obtain consistency results and central limit theorems for the estimation of the multivariate OU type model based on the moments of $\mathbf{Y}$ and $\mathbf{Y}\mathbf{Y}^T$, we need to show that the discretely observed stationary log-returns $\mathbf{Y}$ form a strongly mixing and, thus, ergodic sequence. For $\mu = \beta = 0$ strong mixing in the univariate OU type stochastic volatility model has been obtained in Genon-Catalot, Jeantheau, and Larédo (2000) and Sørensen (2000). For details on mixing we refer again to Doukhan (1994) and regarding ergodicity to Ash and Gardner (1975) or Krengel (1985). In our setup the most important implication of ergodicity is that the usual empirical moments converge almost surely (and in L^1) to the true moments (provided they are finite) as the number of observations goes to infinity.

Proposition 3.4. *(i) Assume that in the* ***general stochastic volatility model*** *the stationary and square integrable process* Σ *is strongly mixing with mixing coefficients* $(\alpha_k(\Sigma))_{k\in\mathbb{N}}$. *Then the process* $\mathbf{Y}$ *is strongly mixing with mixing coefficients* $\alpha_k(\mathbf{Y}) \le \alpha_k(\Sigma)$ *for all* $k \in \mathbb{N}$. *Thus* $\mathbf{Y}$ *is ergodic.*

(ii) In the ***positive semidefinite OU type stochastic volatility model*** *the process* $\mathbf{Y}$ *is always strongly mixing with mixing coefficients* $(\alpha_k)_{k\in\mathbb{N}}$ *decaying at least at an exponential rate. Thus* $\mathbf{Y}$ *is ergodic.*

Proof. Part (i) follows from an immediate adaptation of the proof of Sørensen (2000, Lemma 6.3) to the multivariate case and the case $\mu, \beta \neq 0$. (ii) results from combining (i) with Proposition 2.5. □

Based on these results and the closed form expressions for the second moments a moment matching estimator along the lines of the generalized method of moment (GMM) estimation can be implemented. The resulting estimator is consistent and asymptotically normal, see Hansen (1982) or Hall (2005) for an overview of the GMM method. Yet, this estimation procedure does not allow to filter the current volatility states. To overcome these problems we derive a state-space representation for the joint series of the returns and squared returns, which allows us to use the Kalman recursions for estimation and filtering.

3.3 State-Space Representation

The aim of this section is to establish a state-space representation for the joint process of the returns and squared returns $(\mathbf{Y}_n, \mathbf{Y}_n\mathbf{Y}_n^T)_{n\in\mathbb{N}}$. Throughout we assume $\beta = 0$. As before, we first analyze the general stochastic volatility model and then focus on the OU type specification.

Recall that $\mathbf{Y}_n = \Delta\mu + \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s$, which immediately implies

$$\begin{aligned}\mathbf{Y}_n\mathbf{Y}_n^T =& \Delta^2\mu\mu^T + \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2} \\ &+ \Delta \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \mu^T + \Delta\mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}.\end{aligned}$$

Setting

$$u_n = \begin{pmatrix} u_{1,n} \\ u_{2,n} \end{pmatrix}, \quad u_{1,n} = \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \text{ and}$$

$$u_{2,n} = \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2} + \Delta \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \mu^T + \Delta\mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2} - \mathbf{\Sigma}_n,$$

it follows that

$$\mathbf{Y}_n = \Delta\mu + u_{1,n}, \qquad \mathbf{Y}_n\mathbf{Y}_n^T = \Delta^2\mu\mu^T + \mathbf{\Sigma}_n + u_{2,n}. \tag{24}$$

The martingale property in the following theorem is, of course, understood w.r.t. the filtration $(\mathcal{F}_t)$ which we assume to be given. Recall that all processes (in particular, $L_t, \Sigma_t, W_t, Y_t, \mathbf{Y}_n, \mathbf{\Sigma}_n$) are adapted with respect to this filtration. Moreover, $(W_s - W_t)_{s\geq t}$ is independent of $(\Sigma_s)_{s\in\mathbb{R}^+}$ as well as of $\mathcal{F}_t$ for all $t \in \mathbb{R}^+$. For technical reasons this is, however, not fully sufficient. Thus, we henceforth assume:

Assumption 3.2. *$(W_s - W_t)_{s\geq t}$ is independent of $\boldsymbol{\sigma}(\mathcal{F}_t, (\Sigma_s)_{s\in\mathbb{R}^+})$ for all $t \in \mathbb{R}^+$ ($\boldsymbol{\sigma}(\cdot)$ denoting the generated σ-algebra).*

In the OU type stochastic volatility model this assumption is satisfied if the σ-algebras $\mathcal{F}_t$, $\boldsymbol{\sigma}((L_s - L_t)_{s\geq t})$ and $\boldsymbol{\sigma}((W_s - W_t)_{s\geq t})$ are independent for all $t \in \mathbb{R}^+$ and not only pairwise independent. Clearly the last condition will usually be satisfied. In particular it is satisfied when the pair (L, W) of the driving Lévy process and Wiener process forms a Lévy process in $\mathbb{S}_d^+ \times \mathbb{R}^d$.

Proposition 3.5. *The sequence $(u_n)_{n\in\mathbb{N}}$ is a (second order) stationary zero-mean martingale difference sequence w.r.t. the filtration $(\mathcal{G}_n)_{n\in\mathbb{N}} := (\mathcal{F}_{n\Delta})_{n\in\mathbb{N}}$ and thus in particular white noise. It holds that*

$$\operatorname{var}(u_{1,n}) = E(\mathbf{\Sigma}_n) = E(\Sigma_0)\Delta \tag{25}$$

$$\begin{aligned}\operatorname{var}(\operatorname{vec}(u_{2,n})) =& \Delta^3\big(E(\Sigma_0)\otimes(\mu\mu^T) + (\mu\mu^T)\otimes E(\Sigma_0) + \mu^T\otimes E(\Sigma_0)\otimes\mu \\ &+ \mu\otimes E(\Sigma_0)\otimes\mu^T\big) + \Delta^2(I_{d^2}+\mathbf{P})\left(E(\Sigma_0)\otimes E(\Sigma_0)\right) \\ &+ (\mathbf{Q}+\mathbf{PQ})\left(r^{++}(\Delta) + (r^{++}(\Delta))^T\right)\end{aligned} \tag{26}$$

$$\operatorname{cov}(u_{1,n}, \operatorname{vec}(u_{2,n})) = \Delta^2\left(E(\Sigma_0)\otimes\mu^T + \mu^T\otimes E(\Sigma_0)\right) \tag{27}$$

Proof. The proof of Proposition 3.5 is given in Appendix A.4. □

Remark 3.6. *For the commutation matrix K_d as defined in Magnus and Neudecker (1979), for instance, it can be shown that $K_d(X\otimes Z) = \mathbf{P}(X\otimes Z)$ for all $X, Z \in \mathbb{S}_d$. Thus, the operator $\mathbf{P}$ can be replaced by the commutation matrix K_d in Theorem 3.3 and Proposition 3.5. Note, however, that for $X \in M_{d^2}(\mathbb{R})$ multiplication by K_d is in general not the same as applying $\mathbf{P}$.*

Although the noise u_n is independent of $\mathbf{\Sigma}_n$, we are still confronted with the problem that inference is infeasible, as the latent process $(\mathbf{\Sigma}_n)_{n\in\mathbb{N}}$ still appears in the equations for our observables (24). Clearly, it would be desirable for the process $(\mathbf{\Sigma}_n)_{n\in\mathbb{N}}$ to be also representable as a linear process, preferably with a noise sequence that is uncorrelated with $(u_n)_{n\in\mathbb{N}}$, because then the equations (24) could be extended to a state-space model (for a detailed treatment see e.g. Brockwell and Davis, 1991, Chapter 12) and all the tools developed for these models would be available.

In the following we show that at least for the OU type stochastic volatility model such a state-space representation is indeed available.

To this end, define

$$\eta_{1,n} := \int_{(n-1)\Delta}^{n\Delta} e^{A(n\Delta-s)}dL_s e^{A^T(n\Delta-s)}, \qquad \eta_{2,n} := \int_{(n-1)\Delta}^{n\Delta} dL_s = L_{n\Delta} - L_{(n-1)\Delta}$$

and $\eta_n := (\eta_{1,n}, \eta_{2,n})$. Then for all $n \in \mathbb{N}$ it is obvious that

$$\Sigma_{n\Delta} = e^{A\Delta}\Sigma_{(n-1)\Delta}e^{A^T\Delta} + \eta_{1,n} \text{ and } L_{n\Delta} = L_{(n-1)\Delta} + \eta_{2,n}.$$

Before showing that this leads to a helpful state-space representation, we first study the properties of the noise sequence $(\eta_n)_{n\in\mathbb{N}}$.

Proposition 3.7. *The sequence of random variables $(\eta_n)_{n\in\mathbb{N}}$ is i.i.d. and uncorrelated with $(u_n)_{n\in\mathbb{N}}$. Moreover, it has finite second moments and*

$$E(\eta_{1,n}) = -\mathbf{A}^{-1}\left(E(L_1) - e^{A\Delta}E(L_1)e^{A^T\Delta}\right), \quad E(\eta_{2,n}) = \Delta E(L_1) \tag{28}$$

$$\mathrm{var}(\mathrm{vec}(\eta_{1,n})) = -\mathcal{A}^{-1}\left(\mathrm{var}(\mathrm{vec}(L_1)) - e^{\mathscr{A}\Delta}\mathrm{var}(\mathrm{vec}(L_1))e^{\mathscr{A}^T\Delta}\right) \tag{29}$$

$$\mathrm{var}(\mathrm{vec}(\eta_{2,n})) = \Delta\mathrm{var}(\mathrm{vec}(L_1)) \tag{30}$$

$$\mathrm{cov}\left(\mathrm{vec}(\eta_{1,n}), \mathrm{vec}(\eta_{2,n})\right) = -\mathscr{A}^{-1}\left(\mathrm{var}(\mathrm{vec}(L_1)) - e^{\mathscr{A}\Delta}\mathrm{var}(\mathrm{vec}(L_1))\right). \tag{31}$$

Proof. The proof of Proposition 3.7 is given in Appendix A.5. □

Proposition 2.6 implies $\mathbf{\Sigma}_n = \Sigma^+_{n\Delta} - \Sigma^+_{(n-1)\Delta} = \mathbf{A}^{-1}\left(\Sigma_{n\Delta} - \Sigma_{(n-1)\Delta} - L_{n\Delta} + L_{(n-1)\Delta}\right)$ for all $n \in \mathbb{N}$. Recalling the definition of η_n one thus obtains

$$\mathbf{A}\mathbf{\Sigma}_n = e^{A\Delta}\Sigma_{(n-1)\Delta}e^{A^T\Delta} - \Sigma_{(n-1)\Delta} + \eta_{1,n} - \eta_{2,n}.$$

Combining this with the representation (24) of the observable log price $(\mathbf{Y}_n)_{n\in\mathbb{N}}$ and its "square" $(\mathbf{Y}_n\mathbf{Y}_n^T)_{n\in\mathbb{N}}$ and setting $\alpha_{1,n} = \mathbf{A}\mathbf{\Sigma}_n$ and $\alpha_{2,n} = \Sigma_{n\Delta}$ yields the desired state-space representation:

$$\mathbf{Y}_n = \Delta\mu + u_{1,n}, \quad \mathbf{Y}_n\mathbf{Y}_n^T = \Delta^2\mu\mu^T + \mathbf{A}^{-1}\alpha_{1,n} + u_{2,n}, \tag{32}$$

where

$$\alpha_{1,n} = e^{A\Delta}\alpha_{2,n-1}e^{A^T\Delta} - \alpha_{2,n-1} + \eta_{1,n} - \eta_{2,n}, \quad \alpha_{2,n} = e^{A\Delta}\alpha_{2,n-1}e^{A^T\Delta} + \eta_{1,n} \tag{33}$$

or in pure vector notation with $\alpha_n := \begin{pmatrix}\mathrm{vec}(\alpha_{1,n})\\ \mathrm{vec}(\alpha_{2,n})\end{pmatrix}$

$$\begin{pmatrix}\mathbf{Y}_n\\ \mathrm{vec}(\mathbf{Y}_n\mathbf{Y}_n^T)\end{pmatrix} = \begin{pmatrix}\Delta\mu\\ \Delta^2(\mu\otimes\mu)\end{pmatrix} + \begin{pmatrix}0_{M_{d,d^2}(\mathbb{R})} & 0_{M_{d,d^2}(\mathbb{R})}\\ \mathscr{A}^{-1} & 0_{M_{d^2,d^2}(\mathbb{R})}\end{pmatrix}\alpha_n + \begin{pmatrix}u_{1,n}\\ \mathrm{vec}(u_{2,n})\end{pmatrix} \tag{34}$$

where

$$\alpha_n = \begin{pmatrix}0_{M_{d^2}(\mathbb{R})} & e^{A\Delta}\otimes e^{A\Delta} - I_{d^2}\\ 0_{M_{d^2}(\mathbb{R})} & e^{A\Delta}\otimes e^{A\Delta}\end{pmatrix}\alpha_{n-1} + \begin{pmatrix}\mathrm{vec}(\eta_{1,n} - \eta_{2,n})\\ \mathrm{vec}(\eta_{1,n})\end{pmatrix}. \tag{35}$$

Observe that $0_{M_{d,d^2}(\mathbb{R})}$ denotes the zero matrix in $M_{d,d^2}(\mathbb{R})$.

As regards the noise terms, we have that $(u_n)_{n\in\mathbb{N}}$ and $(\eta_n)_{n\in\mathbb{N}}$ are uncorrelated. Furthermore, $(\eta_n)_{n\in\mathbb{N}}$ is an i.i.d. sequence and $(u_n)_{n\in\mathbb{N}}$ a martingale difference sequence. This state-space representation can be used to conduct model inference. Moreover, the volatility states can be inferred straightforwardly using the Kalman filter, which gives the best estimate in an L^2 sense.

3.4 Identifiability

In this section we discuss conditions ensuring that the second order structure of $\mathbf{Y}\mathbf{Y}^T$ uniquely identifies $A, E(L_1), \mathrm{var}(\mathrm{vec}(L_1))$ in the OU type stochastic volatility model with $\mu = \beta = 0$. Note that the identification can already be obtained by using only $E(\mathbf{Y}_1\mathbf{Y}_1^T), \mathrm{acov}_{\mathbf{Y}\mathbf{Y}^T}(1)$ and $\mathrm{acov}_{\mathbf{Y}\mathbf{Y}^T}(2)$.

In the following we consider $\mathrm{vech}(\Sigma)$ and $\mathrm{vech}(\mathbf{Y}\mathbf{Y}^T)$ rather than $\mathrm{vec}(\Sigma)$ and $\mathrm{vec}(\mathbf{Y}\mathbf{Y}^T)$, which was so far preferable owing to notational advantages, because otherwise we would necessarily be dealing with singular covariance matrices, as the symmetric $d \times d$ matrices are a proper subspace of $M_d(\mathbb{R})$.

Defining $\mathbf{A}_{\text{vech}} := \text{vech} \circ \mathbf{A} \circ \text{vech}^{-1}$ with $\mathbf{A} : \mathbb{S}_d \to \mathbb{S}_d,\ X \mapsto AX + XA^T$ it is easy to see that

$$d\text{vech}(\Sigma_t) = \mathbf{A}_{\text{vech}}\text{vech}(\Sigma_t)dt + d\text{vech}L_t, \tag{36}$$

with the stationary solution being given by $\text{vech}(\Sigma_t) = \int_{-\infty}^{t} e^{\mathbf{A}_{\text{vech}}(t-s)} d\text{vech}L_s$. Observe, moreover, that, if $\mathbf{D}$ denotes the $d^2 \times d(d+1)/2$ duplication matrix and $\mathbf{E}$ the $d(d+1)/2 \times d^2$ elimination matrix by $\mathbf{E}$ (see Magnus, 1988, for instance), then $\mathbf{A}_{\text{vech}} = \mathbf{E}\left(A \otimes I_d + I_d \otimes A\right)\mathbf{D}$ and $e^{\mathbf{A}_{\text{vech}}t} = \mathbf{E}\left(e^{At} \otimes e^{At}\right)\mathbf{D}$. Furthermore, Pigorsch and Stelzer (2009, Proposition 3.1) have shown that the linear operator $\mathbf{A}$ on $\mathbb{S}_d$ uniquely identifies A. Thus, there is a one-to-one correspondence between A and $\mathbf{A}_{\text{vech}}$.

Note that basically all formulae obtained so far in this paper can be rewritten using vech instead of vec in a straightforward manner. For example, the second order structure of $\mathbf{\Sigma}$ can also be expressed using vech:

$$\begin{aligned}
\text{var}(\text{vech}(\Sigma_0)) =& -\mathcal{A}_{\text{vech}}^{-1}\text{var}(\text{vech}(L_1)) \\
&\quad \text{where } \mathcal{A}_{\text{vech}} : M_{d(d+1)/2}(\mathbb{R}) \to M_{d(d+1)/2}(\mathbb{R}),\ X \mapsto \mathbf{A}_{\text{vech}}X + X\mathbf{A}_{\text{vech}}^T \\
\text{acov}_{\text{vech}(\mathbf{\Sigma})}(h) =& \text{acov}_{\text{vech}(\mathbf{YY}^T)}(h) = e^{\mathbf{A}_{\text{vech}}\Delta(h-1)}\mathbf{A}_{\text{vech}}^{-2}\left(I_{d(d+1)/2} - e^{\mathbf{A}_{\text{vech}}\Delta}\right)^2 \text{var}(\text{vech}(\Sigma_0)),\ h \in \mathbb{N}.
\end{aligned}$$

Proposition 3.8. *Assume that the OU type stochastic volatility model with $\mu = \beta = 0$ and $\Delta \in \mathbb{R}^{++}$ is given and that the possible $A \in M_d(\mathbb{R})$ and matrix subordinators L are restricted such that:*

(a) $\sigma(A) \subset (-\infty, 0) + i\mathbb{R}$.

(b) $e^{\mathbf{A}_{\text{vech}}\Delta}$ *uniquely identifies* $\mathbf{A}_{\text{vech}}$.

(c) $\text{var}(\text{vech}(\Sigma_0)) = -\mathcal{A}_{\text{vech}}^{-1}\text{var}(\text{vech}(L_1)) \in GL_{d(d+1)/2}(\mathbb{R})$.

Then $E(\mathbf{Y}_1\mathbf{Y}_1^T), \text{acov}_{\mathbf{YY}^T}(1)$ *and* $\text{acov}_{\mathbf{YY}^T}(2)$ *uniquely identify* $A, E(L_1)$ *and* $\text{var}(\text{vech}(L_1))$.

Proof. By construction $\sigma(\mathbf{A}_{\text{vech}}) \subseteq \sigma(\mathbf{A}) = \sigma(A) + \sigma(A)$ and thus Assumption (a) ensures that $\mathbf{A}_{\text{vech}}^{-2}\left(I_{d(d+1)/2} - e^{\mathbf{A}_{\text{vech}}\Delta}\right)^2$ is invertible. Using also (c) this gives that $\text{acov}_{\text{vech}(\mathbf{YY}^T)}(1)$ is invertible and that $e^{\mathbf{A}_{\text{vech}}\Delta} = \text{acov}_{\text{vech}(\mathbf{YY})}(2)(\text{acov}_{\text{vech}(\mathbf{YY})}(1))^{-1}$. Using (b) $\mathbf{A}_{\text{vech}}$ is therefore identified from $\text{acov}_{\text{vech}(\mathbf{YY})}(1)$ and $\text{acov}_{\text{vech}(\mathbf{YY})}(2)$. Hence, $\mathbf{A}_{\text{vech}}$ and A can be treated as known and so

$$\text{var}(\text{vech}(L_1)) = -\mathcal{A}_{\text{vech}}\left(I_{d(d+1)/2} - e^{\mathbf{A}_{\text{vech}}\Delta}\right)^{-2}\mathbf{A}_{\text{vech}}^2\text{acov}_{\text{vech}(\mathbf{YY}^T)}(1)$$

and $E(L_1) = -\mathbf{A}\Delta^{-1}E(\mathbf{Y}_1\mathbf{Y}_1^T)$ conclude. □

The assumption (b) from above is crucial for the identifiability of the OU type stochastic volatility model. It requests that $\exp(\mathbf{A}_{\text{vech}}\Delta)$ has a unique real logarithm. The following results deal with criteria ensuring the existence of a unique real logarithm of $\exp(\mathbf{A}_{\text{vech}}\Delta)$.

Lemma 3.9. *Assume that A is required to satisfy $\sigma(\mathbf{A}_{\text{vech}}\Delta) \subseteq (-\infty, 0) + i(-\pi, \pi)$. Then $e^{\mathbf{A}_{\text{vech}}\Delta}$ uniquely identifies $\mathbf{A}_{\text{vech}}$.*

Proof. This follows immediately from Horn and Johnson (1991, Section 6.4). □

Lemma 3.10. *Assume that A is required to satisfy $\sigma(\mathbf{A}_{\text{vech}}) \subseteq (-\infty, 0)$ and that all Jordan blocks belonging to the same eigenvalue of $\mathbf{A}_{\text{vech}}$ have to be of a different size. Then $e^{\mathbf{A}_{\text{vech}}\Delta}$ uniquely identifies $\mathbf{A}_{\text{vech}}$ for all $\Delta \in \mathbb{R}^{++}$.*

Proof. Combine Culver (1966, Theorem 2), $\sigma(e^{\mathbf{A}_{\text{vech}}\Delta}) = e^{\sigma(\mathbf{A}_{\text{vech}})\Delta}$ and the fact that the matrix exponential preserves the Jordan block structure (Horn and Johnson, 1991, Theorem 6.2.25). □

The last lemma is comparable to the identifiability restrictions of Kessler and Rahbek (2004) and Bladt and Sørensen (2005). However, it appears to be preferable to have a condition involving only restrictions on A, such that $\mathbf{A}_{\text{vech}}$ does not have to be computed first. To this end we give the following purely linear algebraic lemma stated for real diagonalizable matrices. Its generalization to diagonalizable matrices is straightforward. Below we denote by $\mathbb{SK}_d$ the $d \times d$ skew-symmetric matrices (i.e. the matrices $X \in M_d(\mathbb{R})$ with $X^T = -X$).

Lemma 3.11. *Let $A \in M_d(\mathbb{R})$ be real diagonalizable with (not necessarily distinct) eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_d$.*

Then the linear operator $\mathbf{A} : M_d(\mathbb{R}) \to M_d(\mathbb{R})$, $X \mapsto AX + XA^T$ satisfies $\mathbf{A}(\mathbb{S}_d) \subseteq \mathbb{S}_d$ and $\mathbf{A}(\mathbb{SK}_d) \subseteq \mathbb{SK}_d$. Moreover, $\mathbf{A}$ has $d(d+1)/2$ linearly independent eigenvectors in $\mathbb{S}_d$ with associated eigenvalues $\{\lambda_i + \lambda_j : i = 1, \ldots, d;\ j = 1, \ldots, i\}$ and $d(d-1)/2$ linearly independent eigenvectors in $\mathbb{SK}_d$ with associated eigenvalues $\{\lambda_i + \lambda_j : i = 1, \ldots, d;\ j = 1, \ldots, i-1\}$, which are also linearly independent of the eigenvectors in $\mathbb{S}_d$. Hence, every eigenvalue has an eigenvector in $\mathbb{S}_d \cup \mathbb{SK}_d$.

Proof. That $\mathbf{A}$ preserves (skew-)symmetry is trivial. Assume that $U \in GL_d(\mathbb{R})$ is such that $U^{-1}AU =: D$ is diagonal. Then $AX + XA^T = U\left(DU^{-1}XU^{-T} + U^{-1}XU^{-T}D^T\right)U^T$ and $M_d(\mathbb{R}) \to M_d(\mathbb{R})$, $X \mapsto U^{-1}XU^{-T}$ is an invertible linear map on $M_d(\mathbb{R})$ preserving both $\mathbb{S}_d$ and $\mathbb{SK}_d$ and having inverse $M_d(\mathbb{R}) \to M_d(\mathbb{R})$, $X \mapsto UXU^T$. This implies that we can without loss of generality take A to be diagonal, i.e.

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_d \end{pmatrix}.$$

Let $\{E_{ij}\}_{i,j=1,\ldots,d}$ be the standard basis of $M_d(\mathbb{R})$, i.e. E_{ij} is a matrix having only zero entries except for one entry of one in the i-th row and j-th column. We then have $\mathbf{A}E_{ii} = 2\lambda_i$ for $i = 1, \ldots, d$ and $\mathbf{A}(E_{ij} + E_{ji}) = (\lambda_i + \lambda_j)(E_{ij} + E_{ji})$ for $i = 1, \ldots, d$ and $j = 1, \ldots, i-1$. These are $d(d+1)/2$ linearly independent eigenvectors in $\mathbb{S}_d$. Likewise $\mathbf{A}(E_{ij} - E_{ji}) = (\lambda_i + \lambda_j)(E_{ij} - E_{ji})$ for $i = 1, \ldots, d$ and $j = 1, \ldots, i-1$ gives $d(d-1)/2$ linearly independent eigenvectors in $\mathbb{SK}_d$.

Since $\mathbf{A}$ has d^2 eigenvalues, the fact that $\mathbb{S}_d \cap \mathbb{SK}_d = \{0\}$ and both are linear subspaces of $M_d(\mathbb{R})$ implies that every eigenvalue has an eigenvector in $\mathbb{S}_d \cup \mathbb{SK}_d$. □

Lemma 3.12. *Assume that A is required to be real diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_d$ and that the set $\{\lambda_i + \lambda_j : i = 1, \ldots, d, j = 1, \ldots, i\}$ has to consist of $d(d+1)/2$ pairwise distinct elements. Then $e^{\mathbf{A}_{\text{vech}}\Delta}$ uniquely identifies $\mathbf{A}_{\text{vech}}$ for all $\Delta \in \mathbb{R}^{++}$.*

Proof. $\mathbf{A}_{\text{vech}}$ corresponds to the restriction of $\mathbf{A} : M_d(\mathbb{R}) \to M_d(\mathbb{R})$, $X \mapsto AX + XA^T$ to $\mathbb{S}_d$. Hence, Lemma 3.11 implies that the $d(d+1)/2$ real eigenvalues of $\mathbf{A}_{\text{vech}}$ are pairwise distinct. Thus Culver (1966, Theorem 2) gives that $e^{\mathbf{A}_{\text{vech}}\Delta}$ uniquely identifies $\mathbf{A}_{\text{vech}}$ for all $\Delta \in \mathbb{R}^{++}$. □

Thus, if Lemma 3.12 holds, the assumption (b) for the identifiability of our model is fulfilled.

3.5 Superpositions of Ornstein-Uhlenbeck Type Processes

In this section we indicate a possible extension of our multivariate OU type stochastic volatility model by using a superposition of independent positive semidefinite OU type processes for the volatility process. As in the univariate case this makes the model more flexible without loosing much of its tractability. Yet, in our multivariate setting, this extension is apparently less important, since a superposition of univariate OU type processes is automatically present in the individual components (see Section 2.2).

Let K be a natural number, and $(L^{(1)})_{t\in\mathbb{R}}, (L^{(2)})_{t\in\mathbb{R}}, \ldots, (L^{(K)})_{t\in\mathbb{R}}$ be independent matrix subordinators which are jointly independent of the Brownian motion $(W_t)_{t\in\mathbb{R}^+}$ of our general stochastic volatility model and let $A^{(1)}, A^{(2)}, \ldots, A^{(K)} \in M_d(\mathbb{R})$ be matrices with all eigenvalues in $(-\infty, 0) + i\mathbb{R}$. Then we define K independent stationary positive semidefinite OU type processes by $\Sigma_t^{(i)} = \int_{-\infty}^t e^{A^{(i)}(t-s)} dL_s^{(i)} e^{A^{(i)^T}(t-s)}$ with $i = 1, 2, \ldots, K$ and the stationary stochastic volatility process $\Sigma_t = \sum_{i=1}^K \Sigma_t^{(i)}$. Due to the independence it is clear that the expected value, variance, autocovariance function and integrated volatility of Σ_t are simply the sum over the respective quantities of the individual processes $\Sigma_t^{(i)}$. Thus closed form formulae for these quantities follow immediately from the results of Section 2.

Moreover, also the results on the second order structure of the increments of the integrated volatility generalize.

Proposition 3.13. *Define $r^{(i)+}(t) = \int_0^t \mathrm{acov}_{\Sigma^{(i)}}(u)du$, $r^{(i)++}(t) = \int_0^t r^{(i)+}(u)du$, $\mathbf{A}^{(i)} : M_d(\mathbb{R}) \to M_d(\mathbb{R})$, $X \mapsto A^{(i)}X + XA^{(i)^T}$, $\mathscr{A}^{(i)} = (A^{(i)} \otimes I_d) + (I_d \otimes A^{(i)})$ and $\mathcal{A}^{(i)} : M_{d^2}(\mathbb{R}) \to M_{d^2}(\mathbb{R})$, $X \mapsto \mathscr{A}^{(i)}X + X\mathscr{A}^{(i)^T}$ for $i = 1, 2, \ldots, K$. Then we have for the stochastic volatility model with a superposition of positive semidefinite OU type processes as volatility process:*

$$E(\Sigma_0) = -\sum_{i=1}^K (\mathbf{A}^{(i)})^{-1} E(L_1^{(i)}), \quad \mathrm{var}(\mathrm{vec}(\Sigma_0)) = -\sum_{i=1}^K (\mathcal{A}^{(i)})^{-1} \mathrm{var}(\mathrm{vec}(L_1^{(i)}))$$

$$r^{++}(t) = -\sum_{i=1}^K \left((\mathscr{A}^{(i)})^{-2} \left(e^{\mathscr{A}^{(i)}t} - I_{d^2} \right) - (\mathscr{A}^{(i)})^{-1} t \right) (\mathcal{A}^{(i)})^{-1} \mathrm{var}(\mathrm{vec}(L_1^{(i)}))$$

$$\mathrm{acov}_{\mathbf{\Sigma}}(h) = -\sum_{i=1}^K e^{\mathscr{A}^{(i)}\Delta(h-1)} (\mathscr{A}^{(i)})^{-2} \left(I_{d^2} - e^{\mathscr{A}^{(i)}\Delta} \right)^2 (\mathcal{A}^{(i)})^{-1} \mathrm{var}(\mathrm{vec}(L_1^{(i)})), \ h \in \mathbb{N}.$$

Thus we obtain very explicit formulae whenever they are also available in the simple multivariate OU type stochastic volatility model. However, the volatility increments $\mathbf{\Sigma}$ and the squared logarithmic prices $\mathbf{YY}^T$ are no longer multivariate ARMA(1,1) processes. However, the state-space representation presented in Section 3.3 can be extended to the case of superposition. We have

$$\mathbf{Y}_n = \Delta\mu + u_{1,n}, \qquad \mathbf{YY}_n^T = \Delta^2 \mu\mu^T + \mathbf{\Sigma}_n + u_{2,n} = \Delta^2 \mu\mu^T + \sum_{i=1}^K \mathbf{\Sigma}_n^{(i)} + u_{2,n}.$$

This immediately gives rise to a state-space representation of

$$\left(\mathbf{Y}_n, \mathbf{YY}_n^T\right), \left(\mathbf{A}^{(1)}\mathbf{\Sigma}_n^{(1)}, \Sigma_{n\Delta}^{(1)}\right), \left(\mathbf{A}^{(2)}\mathbf{\Sigma}_n^{(2)}, \Sigma_{n\Delta}^{(2)}\right), \ldots, \left(\mathbf{A}^{(K)}\mathbf{\Sigma}_n^{(K)}, \Sigma_{n\Delta}^{(K)}\right)$$

using equations (33) to obtain independent recursions for $(\mathbf{A}^{(i)}\mathbf{\Sigma}_n^{(i)}, \Sigma_{n\Delta}^{(i)})$.

As in the univariate case (cf. Barndorff-Nielsen, 2001 and Barndorff-Nielsen and Shephard, 2001), one can also model long-range dependence by superimposing infinitely (but countably) many appropriate positive semidefinite OU type processes. This follows from a straightforward generalization of the arguments in Barndorff-Nielsen (1998, Section 4).

4 Empirical Illustration

In this section we provide a small empirical application of the multivariate OU type stochastic volatility model using a quadrivariate exchange rate data set and a bivariate stock price data set

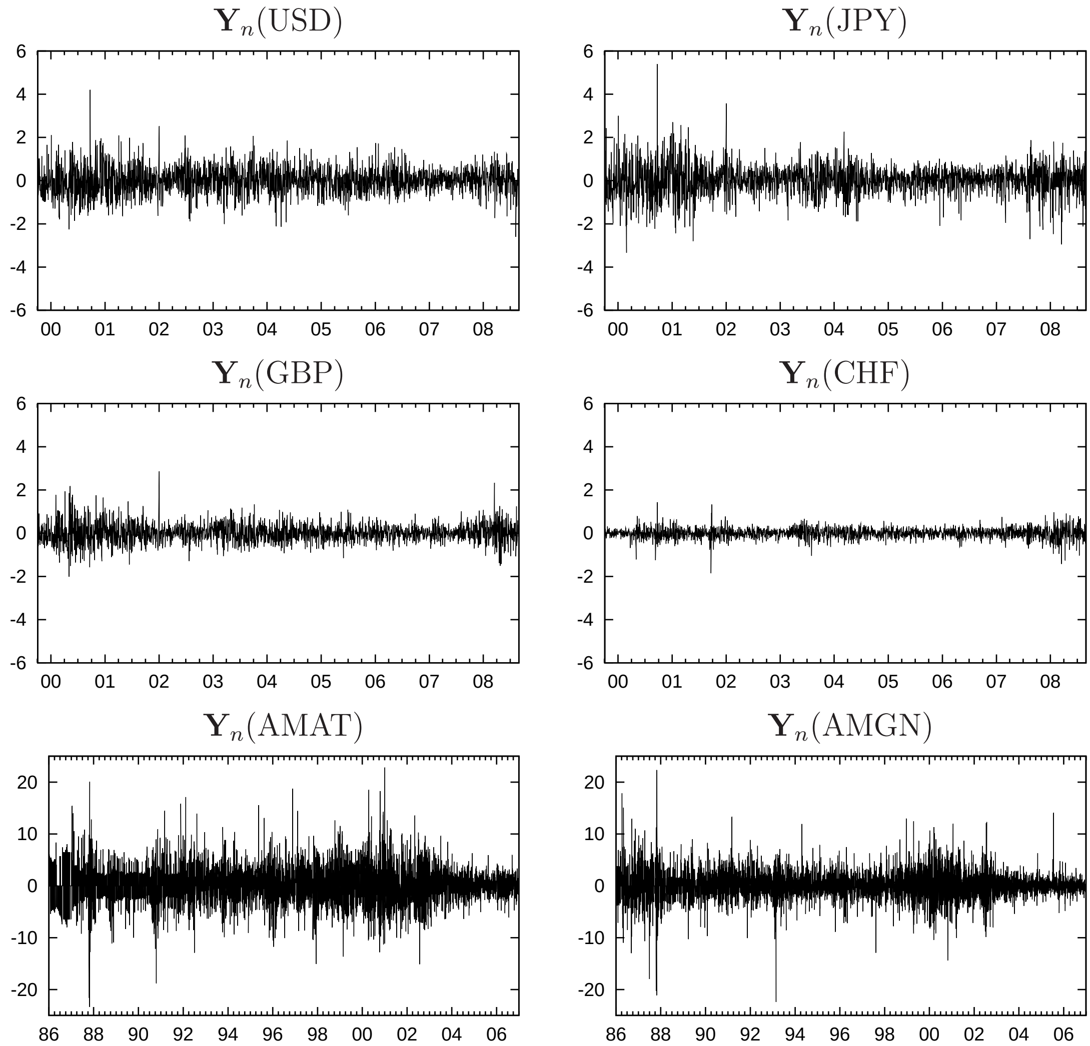

Figure 1: **Daily returns.** Time series plots of the daily percentage logarithmic returns. The four upper panels show the evolvement of the returns of the exchange rates (January 4nd, 1999 until September 30th, 2008) and the bottom two panels show the returns for the stock price data (January 2nd, 1985 until December 29th, 2006).

in order to illustrate the features of the model and to show that it provides an accurate description of the data.

4.1 Data

The first data set consists of 2495 daily observations (ranging from January 1st, 1999 to September 30th, 2008) for the spot exchange rates of the United States dollar (USD), the Japanese yen (JPY), the pound sterling (GBP) and the Swiss franc (CHF) all quoted in terms of the Euro. The second data set is given by two US stocks, viz. Applied Materials Inc. (AMAT) and Amgen Inc. (AMGN). The sample of the second data set covers the period from January 2nd, 1985 to December 29th, 2006, resulting in 5550 observations. Note that splits and dividends are incorporated into the stock prices. Moreover, the continuously compounded returns are mean-adjusted. Figure 1 presents the time evolvement of the corresponding time series exhibiting the usual empirical characteristics, such as volatility clustering.

4.2 Estimation Methods

The estimation of continuous-time stochastic volatility models is complicated by the unavailability of the likelihood function. However, based on the theoretical results derived in Section 3, the multivariate OU type stochastic volatility model can be estimated in several ways, either by using the characteristic function, the second order dependence structure of the squared returns or by exploiting its state-space representation. Among these methods the approach based on the characteristic function is rather inconvenient. In particular, as discussed in Knight and Yu (2002) or Singleton (2001) the characteristic function method requires a fully specified parametric matrix subordinator and is computationally demanding as the unobserved stochastic volatility process needs to be integrated out. As the estimation via the second order moments, i.e. using GMM, or via a state-space representation are not subject to such requirements, we focus on these methods in the following.

4.2.1 Estimation via the Second Order Dependence Structure

Based on the ergodicity of the discretely sampled returns in our model a simple estimator can be obtained by matching a set of empirical moments to their model implied counterparts given in Theorem 3.3. Using the results from Section 3.4 natural candidates are $\text{vech}(E\left(\mathbf{YY}^T\right)$, $\text{acov}_{\text{vech}(\mathbf{YY}^T)}(1)$ and $\text{acov}_{\text{vech}(\mathbf{YY}^T)}(2)$, which ensure in connection with Lemma 3.9 the identification of the model. However, to gain efficiency we consider additional lags of the autocovariance function which yields the objective function

$$SSR(\mathcal{L}) \;=\; \left\|m - E\left(\text{vech}(\mathbf{YY}^T)\right)\right\|_F + \sum_{l\in\mathcal{L}} \left\|a_l - \text{acov}_{\text{vech}(\mathbf{YY}^T)}(l)\right\|_F, \tag{37}$$

with $\mathcal{L}$ denoting the set of selected lags of the autocovariance function and m and a_l for $l \in \mathcal{L}$ the empirical mean and autocovariance function of $\text{vech}(\mathbf{YY}^T)$, respectively. Although several matrix norms can be considered we use here the Frobenius norm, i.e.

$$\|C\|_F = \sum_{i=1}^{d}\sum_{j=1}^{d} c_{ij}^2 = \text{tr}(CC^T),$$

such that the objective function (37) poses a non-linear least squares problem. As noted in Barndorff-Nielsen and Shephard (2001, Section 5.3), who apply this procedure for the estimation of a univariate OU type stochastic volatility model, the estimator is independent of the assumption of a particular OU type process. More precisely, based on the results discussed in Section 3.4, we know, that this estimator generally identifies the mean and the variance of the Lévy process L driving the OU type process (referred to as the background-driving Lévy process or BDLP, for short, in the following), as well as the matrix A. So rather than assuming a specific parametric model for the BDLP, we optimize over the first and second moments of L_1. Often the parameters of a specific BDLP can be identified solely by the mean and the variance of the BDLP. In this case the autocovariance fit also identifies these parameters.

This estimator is computationally very fast but it lacks from the perspective of optimal weighting, especially the autocovariance terms usually have a higher variance as the empirical mean of the time series, which is not incorporated in the estimation by minimizing the objective function (37). We account for these effects by applying the GMM estimation method proposed by Hansen (1982) using the previous estimates based on (37) as starting values. For our weighting matrix we use a HAC estimator of the long run covariance matrix with Parzen kernel and a lag of 25, which is continuously updated. For a detailed discussion of these concepts see Hall (2005).

Although a large number of moment conditions would improve the asymptotic efficiency of the estimator it lacks from the fact that the weighting matrix is estimated with less precision. Given this trade-off in a finite sample setup our moment conditions are given by the mean and the autocovariances of the time series $\mathrm{vech}(\mathbf{YY}^T)$ at lags $(1, 3, 5, 7, 9, 11, 20, 30, 60, 100)$.

4.2.2 Estimation via the State-Space Representation

Based on the state-space representation for the squared returns (see Section 3.3) the Kalman recursions can be used to obtain the quasi-likelihood function of the model, cf. Harvey (1991). This estimation approach has also been considered in Barndorff-Nielsen and Shephard (2001) for the univariate OU type stochastic volatility model. In that paper they also show that the Kalman filter is suboptimal but provides consistent and asymptotically normal estimators. The state-space representation also illustrates once more that the multivariate model can again be estimated without the need of a prespecified parametric BDLP. Thus, estimates of the mean and variance of the BDLP as well as of the matrix A can straightforwardly be obtained by the Kalman filter combined with quasi-maximum likelihood estimation. As in the univariate case the consistency and asymptotic normality of these estimators follows from Dunsmuir (1979). Moreover, in contrast to the above GMM approach, the states of the (co)volatilities can be deduced.

4.3 Estimation Results

In the following we estimate the multivariate OU type stochastic volatility model as given in equations (8) and (9), where we follow Barndorff-Nielsen and Shephard (2001) by assuming that $\mu = \beta = 0$, i.e. in accordance to our use of mean-adjusted returns, the means of the returns are set to zero. Moreover, we assume that the off-diagonal elements of the matrix subordinator are zero. The vector of the diagonal components $(L_1, L_2, \ldots, L_d)^T$ of L to be denoted by $\mathrm{diag}(L)$ is a Lévy process in $\mathbb{R}^d$ with all components being univariate subordinators. Hence, the d elements of $\mathrm{diag}(L)$ only correlate positively (see Pigorsch and Stelzer, 2009, Prop. 5.2) and the correlation between the variances of the different assets in this model is determined by both, the correlation structure of $\mathrm{diag}(L)$ as well as by A. Although this specification seems rather restrictive, it turns out that even this model can quite adequately describe the joint dynamics.

Table 1 presents the estimation results for the two data sets based on the GMM estimation procedure. Although the estimates of A are difficult to interpret, the corresponding eigenvalues show the expected behavior, especially in the exchange rate data set. The matrix A has some very large and very small eigenvalues in absolute magnitude, indicating short term and long term components, respectively. Using this in connection with results from Section 2.2 we have the marginal dynamics as a superposition of very different univariate OU type processes. This can also be observed for univariate OU type processes, where realistic modeling seems to require a superposition of OU type processes, see e.g. Barndorff-Nielsen and Shephard (2001). The presence of very different eigenvalues of A in the exchange rate data set shows that superpositions in the multivariate model are not as essential as in the univariate case.

Figure 2 depicts the autocorrelation functions for the bivariate data set. In particular, the upper two panels show the estimated model-implied autocorrelations of the squared daily returns along with the empirical one (given by dots) for AMAT and AMGN, respectively. The solid line refers to the daily autocovariance function implied by a univariate model and the dotted lines depict the autocovariance function implied by a superposition model with two processes superimposed.[1] Obviously, the superposition model provides a better fit than the simple model. More precisely, in contrast to the simple model its autocovariance function decreases faster for short lags (up to the 10th or 20th lag, depending on the asset) and slower for longer lags as it is the weighted sum of two exponentials decreasing at different rates. These results are in line with the findings of

[1] Estimation results for the univariate and univariate superposition model are not reported here but are available upon request.

Table 1: Estimation results

parameter	quadrivariate data set	bivariate data set
A	$\begin{pmatrix} -0.1702 & 0.3388 & -0.5366 & -0.8994 \\ 0.0662 & -0.3769 & 0.7903 & 1.0016 \\ 4.5057 & -0.0778 & -7.1922 & -0.6793 \\ -1.3457 & 0.5289 & 0.9329 & -1.2028 \end{pmatrix}$	$\begin{pmatrix} -0.0527 & 0.1320 \\ -0.0921 & 0.0230 \end{pmatrix}$
$\mathrm{eig}(A)$	$\begin{pmatrix} -6.7859 & -2.1439 & -0.0015 & -0.0108 \end{pmatrix}^T$	$\begin{pmatrix} -0.0148 + 0.1036\imath & -0.0148 - 0.1036\imath \end{pmatrix}^T$
$E\left(\mathrm{diag}\left(L_1\right)\right)$	$\begin{pmatrix} 0.0007 & 0.0056 & 0.0001 & 0.0052 \end{pmatrix}^T$	$\begin{pmatrix} 0.7838 & 0.0002 \end{pmatrix}^T$
$\mathrm{var}\left(\mathrm{diag}\left(L_1\right)\right)$	$\begin{pmatrix} 0.0006 & 0.0423 & 0.0017 & 0.0000 \\ 0.0001 & 0.0060 & 0.9931 & 0.9415 \\ 0.0000 & 0.0083 & 0.0117 & 0.9502 \\ 0.0000 & 0.0016 & 0.0022 & 0.0005 \end{pmatrix}$	$\begin{pmatrix} 2.5501 & 0.7026 \\ 2.3626 & 4.4340 \end{pmatrix}$

The table presents the estimation result for the quadrivariate exchange rate and bivariate stock price data sets. The first column states the name of the estimated parameter and the second and third column shows the estimates for the quadrivariate and bivariate data set, respectively. The parameter in the first row is the matrix A, in the second row the implied eigenvalues of the matrix A are shown and for the third and fourth row the expectation and variance of the driving Lévy process L at time one is given, respectively. Note that the upper non-diagonal elements of the covariance matrices are normalized in such a way that the entries give the corresponding correlations (based on the non-rounded covariance matrix).

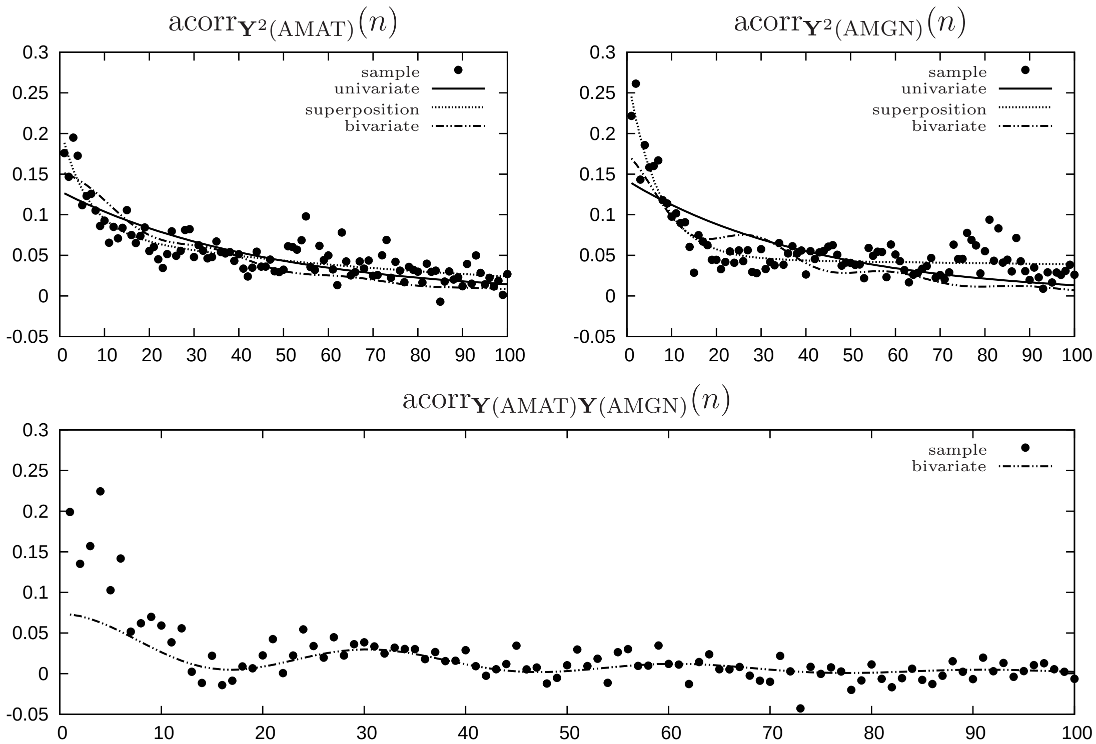

Figure 2: **Autocorrelations for the stock return data set.** Estimated and empirical (black dots) daily autocorrelation functions of the squared returns (upper panels) and of the cross products of the returns (bottom panels). The autocorrelation functions based on the different models are characterized by different line styles: the continuous line refers to the fit of a univariate model, the dotted line corresponds to a univariate superposition model with two OU type processes, and the dashed line corresponds to the autocorrelation based on the bivariate system.

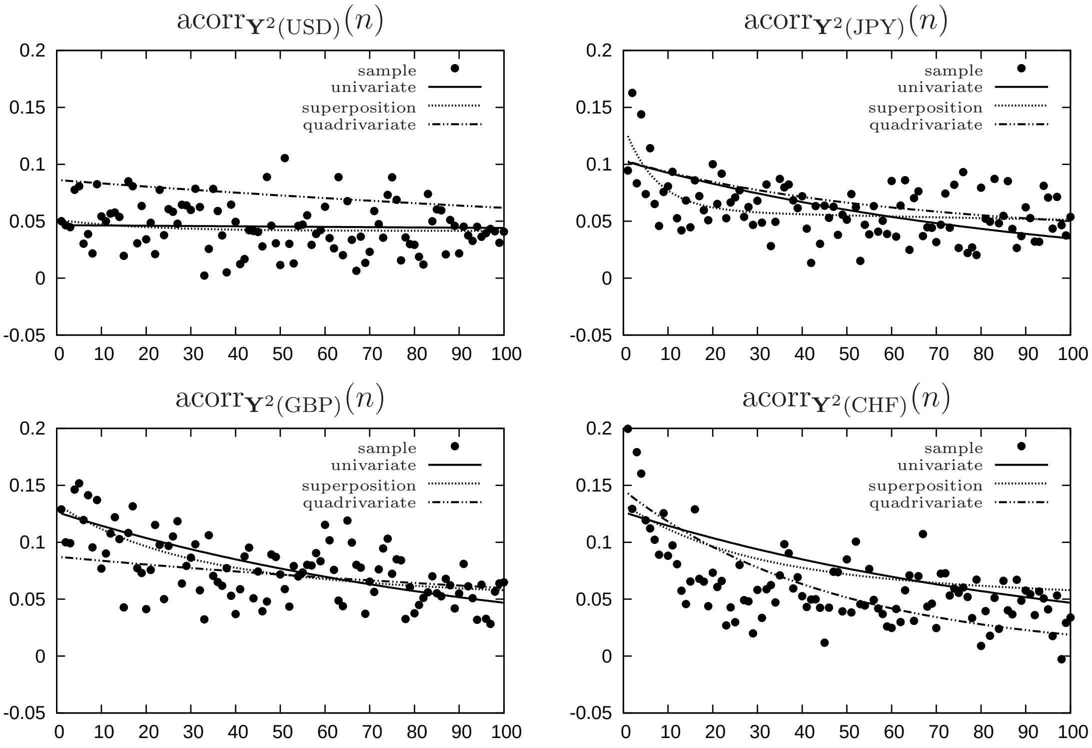

Figure 3: **Autocorrelations of squared returns for the exchange rate data set.** Estimated and empirical (black dots) daily autocorrelation functions of the squared returns. The autocorrelation functions based on the different models are characterized by different line styles: the continuous line refers to the fit of a univariate model, the dotted line corresponds to a univariate superposition model with two OU type processes, and the dashed line corresponds to the autocorrelation based on the quadrivariate system.

Barndorff-Nielsen and Shephard (2001, 2002), who encourage the use of superposition models in the univariate case.

The autocorrelation functions of the bivariate data set in Figure 2 highlight an advantageous property of our model, which is induced by the eigenvalues of A with non-vanishing complex part. These eigenvalues lead to an exponentially damped sinusoidal behavior of the matrix exponential and allows for a non monotonically decreasing autocorrelation function as depicted. Although such a behavior is empirically rather infrequently observed, it emphasizes the flexibility of the multivariate OU stochastic volatility model, even for the marginal dynamics. Note that even for the univariate superposition models the autocorrelation function is necessarily monotonically decreasing. To obtain such a flexible behavior in the univariate case extensions such as the class of CARMA models, proposed by Brockwell (2001) with appropriate positivity restrictions can be applied, see e.g. Todorov and Tauchen (2006).

Figures 3 and 4 show the empirical and estimated autocorrelation function for the squared returns and cross products of the returns from the exchange rate data set, respectively. Also in this medium sized quadrivariate system the model fits the autocorrelation very well.

Note that these properties of our model already emerge empirically within a rather simple specification of the matrix subordinator, viz. a diagonal subordinator. But we conjecture that the flexibility of our model can even be further improved by considering more sophisticated specifications of the background-driving Lévy process or by allowing for superpositions.

The distribution of past and current volatility given all available information is helpful for

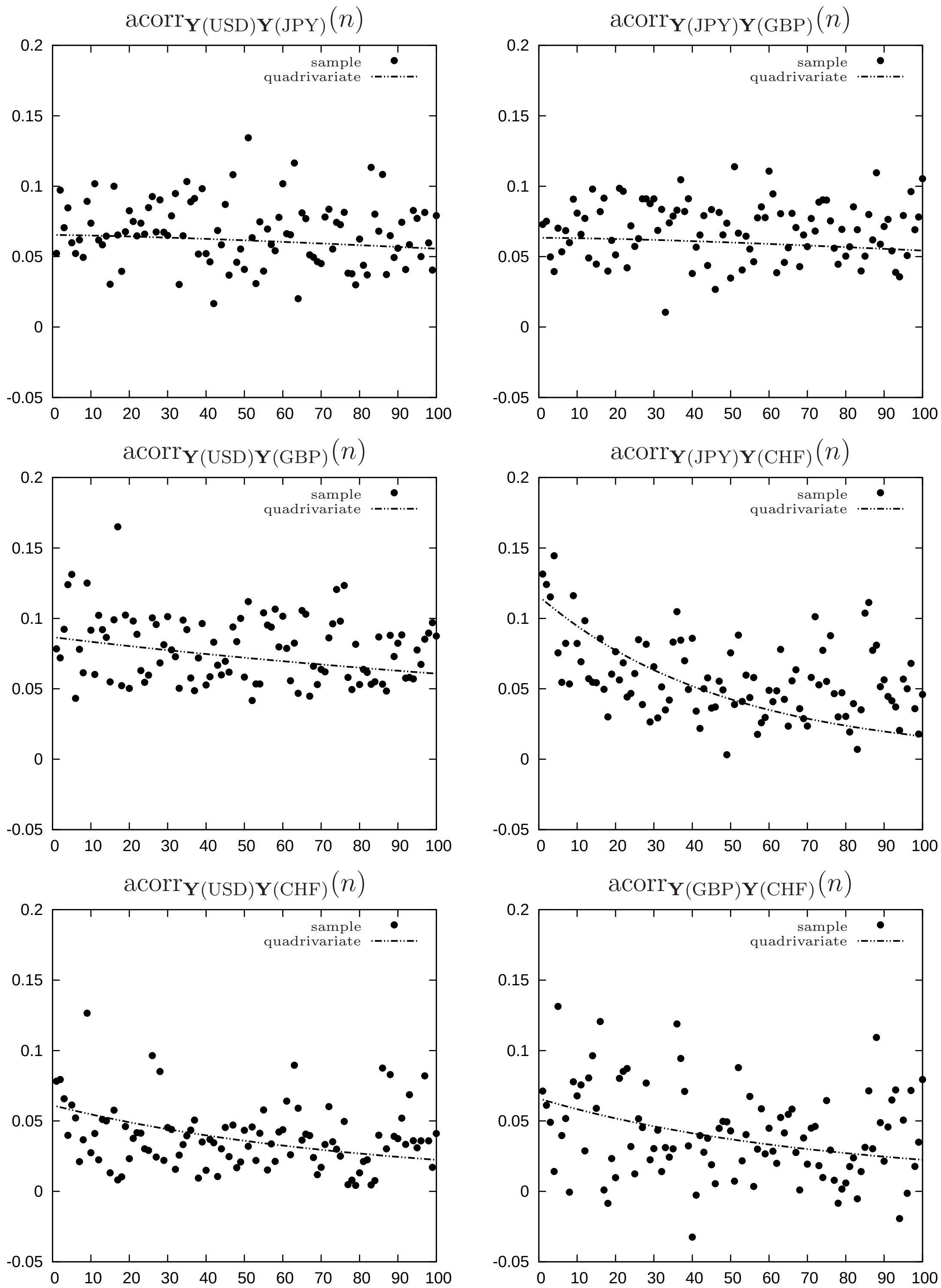

Figure 4: **Autocorrelations of cross products for the exchange rate data set.** Estimated (dashed line) and empirical (black dots) daily autocorrelation functions for the cross products of the returns from the quadrivariate system.

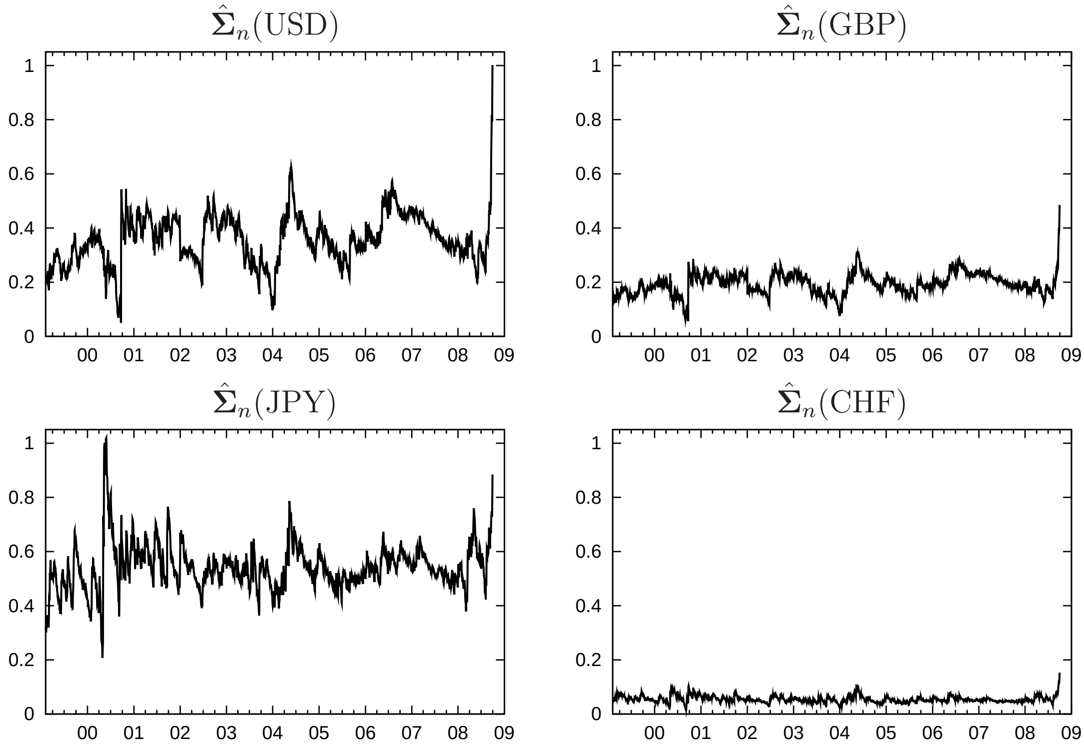

Figure 5: **Filtered daily integrated variances of the exchange rate series.** The filtered values are based on the parameter estimates from the Kalman recursions of the state-space representation of our model.

empirical analysis and crucial for forecasting. Using the state-space representation from Section 3.3 it is straightforward to use the Kalman recursions to estimate the model and filter [2] and forecast the daily integrated volatility $\mathbf{\Sigma}$.[3] Figures 5 and 6 depict the filtered variances and correlations (based on the parameter estimates from the Kalman recursions) of the exchange rate series, respectively.

As one expects due to the financial crisis the volatility is heavily increasing at the end of the sample for all four exchange rates. Figure 7 presents the corresponding results for the bivariate stock price data set. Obviously, in both systems the volatilities tend to move together. Moreover, the correlations, especially the one between AMAT and AMGN, show a relatively constant mean. Note that the volatilities and correlation of AMAT and AMGN are themselves quite volatile.

5 Conclusion

Given the relevance of jointly modeling the dynamics of multiple assets for portfolio and risk management decisions, we have generalized the non-Gaussian OU type stochastic volatility model proposed by Barndorff-Nielsen and Shephard (2001) to the multivariate case. It turns out, that

[2] As the Kalman filter is an L^2 projection on a linear space, it does not necessarily give positive semi-definite matrices. However, in almost all cases in our data sets the filtered covariance matrix was positive definite. In the very few exceptional cases, which only occurred in the foreign exchange data, it returned to the positive semi-definite matrices extremely fast. In Figure 5 we set the correlation to 1 or -1 whenever the filtered covariance matrix was not positive semi-definite. Note, however, that we have not adjusted the values used within the Kalman filter.

[3] Note that we report in Table 1 only the estimates obtained via the GMM method. The corresponding quasi-maximum likelihood estimates based on the Kalman filter are nearly the same and we therefore do not report them here.

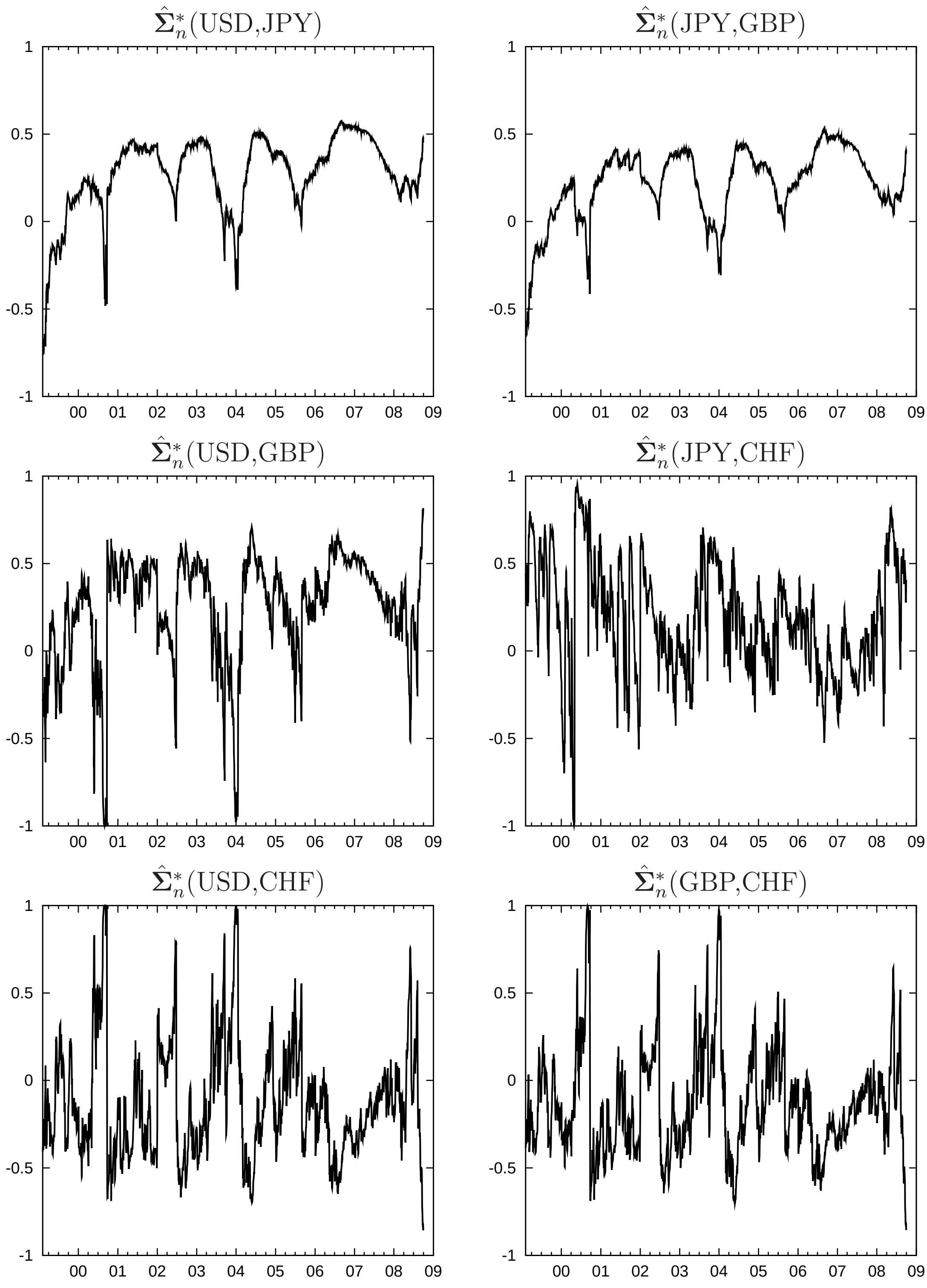

Figure 6: **Filtered daily integrated correlations for the exchange rate data set.** Note that $\hat{\mathbf{\Sigma}}_n^*$(A,B) denotes the filtered correlation instead of the filtered covariance between asset A and B and is calculated as the filtered daily integrated covariance divided by the square root of the product of the filtered daily integrated variances. The filtered values are based on the parameter estimates from the Kalman recursions of the state-space representation of our model.

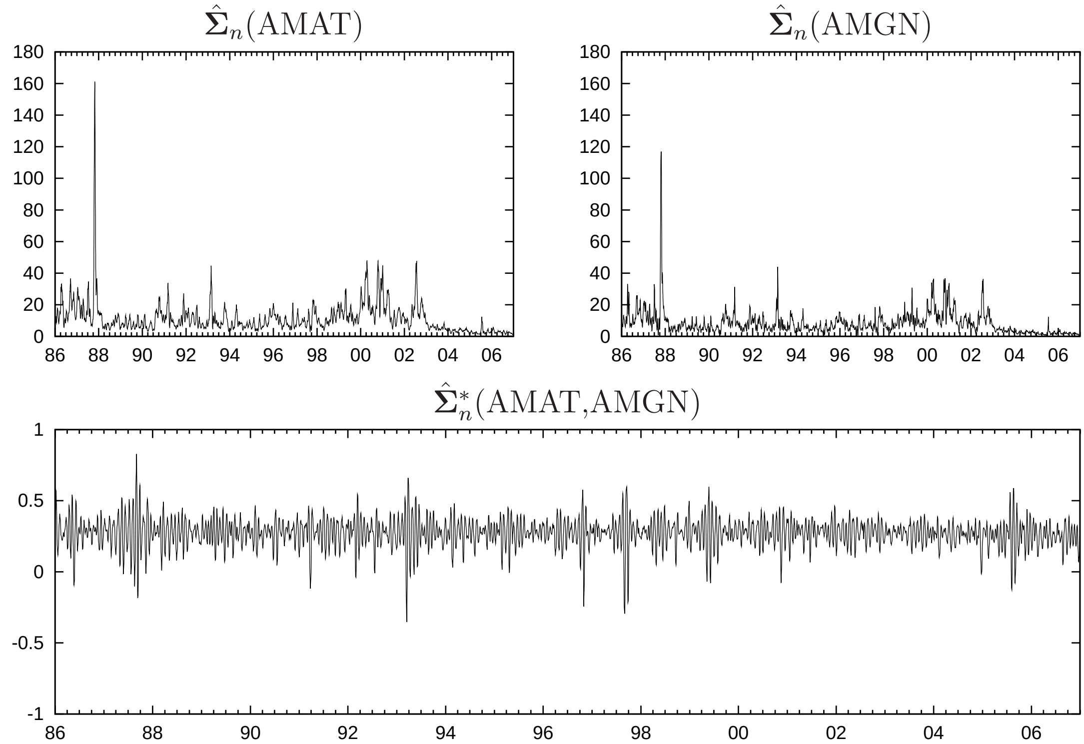

Figure 7: **Filtered daily integrated variances and correlation for the stock return data set.** Note that $\hat{\mathbf{\Sigma}}^*_n$(A,B) denotes the filtered correlation instead of the filtered covariance between asset A and B and is calculated as the filtered daily integrated covariance divided by the square root of the product of the filtered daily integrated variances. The filtered values are based on the parameter estimates from the Kalman recursions of the state-space representation of our model.

our model possesses many attractive features which are mainly a result of our stochastic volatility specification.

Specifying the stochastic volatility by Lévy-driven positive semidefinite OU type processes provides a flexible dependence structure for the volatility. In particular, we show that the increments of the integrated covariance and the outer product of the returns ("squared returns") of a stochastic volatility model based on a single positive semidefinite OU type process follow an ARMA(1,1) processes. Furthermore, closed form expressions are given for the first and second moments of these variables. These results facilitate the implementation of financial decisions, such as the choice of e.g. a minimum-variance portfolio or other types of risk assessment, and the estimation of our model. Moreover, we derived a state-space representation for the joint process of the returns and the outer product of the returns, which provides an additional approach for the estimation of the model as well as for the estimation and forecasting of the volatility states using the Kalman recursions.

Since our model is defined in terms of a matrix subordinator its particular specification may depend on the application at hand. In the empirical part of this paper we focused on models with a simple diagonal matrix subordinator, which already exhibit some nice properties, see Section 4.3. However, studying the empirical relevance of alternative matrix subordinators deserves more attention in future research.

Further improvements in the estimation of the model may be obtained by incorporating the high-frequency based and, thus, more informative realized covariation measure. Similarly to the univariate case (see Barndorff-Nielsen and Shephard, 2002) the derivation of the second order properties and a state-space representation for the realized covariance matrix and the assessment of its usefulness in the estimation of our model is desirable. We are currently working on this extension.

The focus of this paper is on the statistical analysis of our model and the estimation based on historical observations. Another very important question – beyond the scope of the present paper – is its use for derivative pricing (structure preserving changes of measure to equivalent martingale measures, derivative pricing via Fourier-Laplace transform methods, ...). Currently we are looking at these issues and it is planned to report the results elsewhere.

Acknowledgments

Robert Stelzer acknowledges financial support by the Deutsche Forschungsgemeinschaft through the graduate program “Angewandte Algorithmische Mathematik” at the Technische Universität München. Christian Pigorsch acknowledges financial support by the Deutsche Forschungsgemeinschaft through the SFB 386.

A Proofs

A.1 Proof of Theorem 3.1

Proof. Using $\Sigma_t = e^{At}\Sigma_0 e^{A^Tt} + \int_0^t e^{A(t-s)}dL_s e^{A^T(t-s)}$, (8), (10), Proposition 2.6, the invariance of tr under cyclic permutations and the independence of W and L, we obtain

$$E\left(e^{i\langle (y,z),(Y_t,\Sigma_t)\rangle}\,\middle|\,\Sigma_0, Y_0\right) = E\left(e^{i\left(\mathrm{tr}(\Sigma_t z)+(Y_0+\mu t)^T y+\beta^T\Sigma_t^+ y\right)-\frac{1}{2}y^T\Sigma_t^+ y}\,\middle|\,\Sigma_0, Y_0\right)$$
$$= e^{i\left((Y_0+\mu t)^T y+\mathrm{tr}\left(\Sigma_0 e^{A^Tt} z e^{At}\right)+\mathrm{tr}\left(\left[\mathbf{A}^{-1}(\mathscr{B}(t)\Sigma_0)\right](y\beta^T+\frac{i}{2}yy^T)\right)\right)} E\left(e^{i\mathrm{tr}\left(\left(\int_0^t \mathscr{C}_{y,z}(t-s)dL_s\right)^* I_d\right)}\right)$$

with linear operators

$$\mathscr{B}(t) : X \mapsto e^{At}Xe^{A^Tt} - X$$
$$\mathscr{C}_{y,z}(t) : X \mapsto (\beta y^T - \frac{i}{2}yy^T)\left[\mathbf{A}^{-1}\left(e^{At}Xe^{A^Tt} - X\right)\right] + z^T e^{At}Xe^{A^Tt}$$

on $M_d(\mathbb{C})$. Note also that the existence of the expectations is ensured by $e^{-y^T\Sigma_t^+ y} \in (0, 1]$ and observe that $\mathbf{A}^{-1}(\mathbb{S}_d) \subseteq \mathbb{S}_d$.

From the usual formula for the Fourier transform of an integral with a deterministic integrand with respect to a Lévy process (see Marquardt and Stelzer, 2007, Section 2.2, for a brief review) we have that

$$E\left(e^{i\mathrm{tr}\left(\left(\int_0^t \mathscr{C}_{y,z}(t-s)dL_s\right)^* I_d\right)}\right) = e^{\int_0^t \psi_L\left(\mathscr{C}^*_{y,z}(t-s)I_d\right)ds} = e^{\int_0^t \psi_L\left(\mathscr{C}^*_{y,z}(s)I_d\right)ds}.$$

From the definition of the adjoint of a linear operator it follows easily that the adjoints of $\mathbf{A}$ and $\mathscr{C}_{y,z}(t)$ are the following linear operators on $M_d(\mathbb{C})$

$$\mathbf{A}^* : X \mapsto A^TX + XA$$
$$\mathscr{C}^*_{y,z}(t) : X \mapsto e^{A^Ts}zXe^{As} + e^{A^Ts}\left[\mathbf{A}^{-*}\left(\left(y\beta^T + \frac{i}{2}yy^T\right)X\right)\right]e^{As} - \mathbf{A}^{-*}\left(\left(y\beta^T + \frac{i}{2}yy^T\right)X\right).$$

Moreover, $(\mathbf{A}^*)^{-1} = (\mathbf{A}^{-1})^*$ is true for any linear operator.

Combining the above results gives the formula stated in the theorem. Finally, note that $e^{A^Ts}ze^{As}$ $+e^{A^Ts}\left[\mathbf{A}^{-*}\left(y\beta^T+\frac{i}{2}yy^T\right)\right]e^{As}-\mathbf{A}^{-*}\left(y\beta^T+\frac{i}{2}yy^T\right)\in M_d(\mathbb{R})+i\mathbb{S}_d^+$ for all $s\in\mathbb{R}^+$, because

$$e^{A^Ts}\mathbf{A}^{-*}(yy^T)e^{As}-\mathbf{A}^{-*}(yy^T)=\int_0^s e^{A^Tu}yy^Te^{Au}du\in\mathbb{S}_d^+.$$

□

A.2 Proof of Theorem 3.2

Proof. It is immediate from the definitions that the stationarity of $(\Sigma_t)_{t\in\mathbb{R}^+}$ implies the stationarity of $(\mathbf{\Sigma}_n)_{n\in\mathbb{N}}$ and $(\mathbf{Y}_n)_{n\in\mathbb{N}}$. So it remains to verify the above stated formulae.

The first equation of (13) follow immediately from the integral representation $\Sigma_t^+=\int_0^t\Sigma_t dt$ and the second one can be obtained by using a Fubini argument:

$$\begin{aligned}\mathrm{var}(\mathrm{vec}(\mathbf{\Sigma}_1))&=E\left(\int_0^\Delta\mathrm{vec}(\Sigma_s)ds\int_0^\Delta\mathrm{vec}(\Sigma_u)^Tdu\right)-E\left(\int_0^\Delta\mathrm{vec}(\Sigma_s)ds\right)E\left(\int_0^\Delta\mathrm{vec}(\Sigma_u)^Tdu\right)\\&=\int_0^\Delta\int_0^\Delta\mathrm{acov}_\Sigma(s-u)duds=r^{++}(\Delta)+r^{++}(\Delta)^T.\end{aligned}$$

$\mathbf{Y}_1=\int_0^\Delta(\mu+\Sigma_s\beta)ds+\int_0^\Delta\Sigma_s^{1/2}dW_s$ and $E\left(\int_0^\Delta\Sigma_s^{1/2}dW_s\right)=0$ give the first equation in (15). Regarding the second one elementary arguments imply

$$\begin{aligned}\mathrm{var}(\mathbf{Y}_1)=&\mathrm{var}\left(\int_0^\Delta\Sigma_s\beta ds\right)+\mathrm{var}\left(\int_0^\Delta\Sigma_s^{1/2}dW_s\right)+\mathrm{cov}\left(\int_0^\Delta\Sigma_s\beta ds,\int_0^\Delta\Sigma_s^{1/2}dW_s\right)\\&+\mathrm{cov}\left(\int_0^\Delta\Sigma_s\beta ds,\int_0^\Delta\Sigma_s^{1/2}dW_s\right)^T.\end{aligned}$$

The standard Itô isometry implies $\mathrm{var}\left(\int_0^\Delta\Sigma_s^{1/2}dW_s\right)=E(\mathbf{\Sigma}_1)=\Delta E(\Sigma_0)$.

For all matrices $A\in M_{m,n}(\mathbb{R})$, $B\in M_{n,p}(\mathbb{R})$ and $C\in M_{p,q}(\mathbb{R})$ with arbitrary $m,n,p,q\in\mathbb{N}$ it holds that $\mathrm{vec}(ABC)=(C^T\otimes A)\mathrm{vec}(B)$ (see Horn and Johnson, 1991, Lemma 4.3.1). Thus

$$\mathrm{var}\left(\int_0^\Delta\Sigma_s\beta ds\right)=\mathrm{var}\left(\int_0^\Delta\mathrm{vec}(\Sigma_s\beta)ds\right)=(\beta^T\otimes I_d)\mathrm{var}(\mathrm{vec}(\mathbf{\Sigma}_1))(\beta\otimes I_d).$$

Moreover, the independence of $(\Sigma_t)_{t\in\mathbb{R}^+}$ and $(W_t)_{t\in\mathbb{R}^+}$ gives

$$\mathrm{cov}\left(\int_0^\Delta\Sigma_s\beta ds,\int_0^\Delta\Sigma_s^{1/2}dW_s\right)=E\left(\int_0^\Delta\Sigma_s\beta dsE\left(\int_0^\Delta dW_s^T\Sigma_s^{1/2}\middle|(\Sigma_s)_{s\in[0,\Delta]}\right)\right)=0.$$

Formula (14): For $h\in\mathbb{N}$ we have applying Fubini

$$\begin{aligned}\mathrm{acov}_{\mathbf{\Sigma}}(h)&=\int_{h\Delta}^{(h+1)\Delta}\int_0^\Delta\left(E\left(\mathrm{vec}(\Sigma_s)\mathrm{vec}(\Sigma_u)^T\right)-E\left(\mathrm{vec}(\Sigma_s)\right)E\left(\mathrm{vec}(\Sigma_u)^T\right)\right)duds\\&=\int_{h\Delta}^{(h+1)\Delta}\int_0^\Delta\mathrm{acov}_\Sigma(s-u)duds=r^{++}(h\Delta+\Delta)-2r^{++}(h\Delta)+r^{++}(h\Delta-\Delta).\end{aligned}$$

Formula (16): Arguments analogous to the ones given for (15) imply

$$\mathrm{acov}_{\mathbf{Y}}(h)=\mathrm{cov}\left(\int_{h\Delta}^{(h+1)\Delta}\Sigma_t\beta dt,\int_0^\Delta\Sigma_t\beta dt\right)=(\beta^T\otimes I_d)\mathrm{acov}_{\mathbf{\Sigma}}(h)(\beta\otimes I_d).$$

Turning to the positive semidefinite OU type stochastic volatility model (17) has already been given in (3). For the twice integrated covariance $r^{++}(t)$ formula (4) and elementary integration give:

$$
\begin{aligned}
r^{+}(t) &= \int_0^t e^{\mathscr{A}s}\mathrm{var}(\mathrm{vec}(\Sigma_0))ds = \mathscr{A}^{-1}\left(e^{\mathscr{A}t} - I_{d^2}\right)\mathrm{var}(\mathrm{vec}(\Sigma_0)) \\
r^{++}(t) &= \mathscr{A}^{-1}\int_0^t \left(e^{\mathscr{A}s} - I_{d^2}\right)\mathrm{var}(\mathrm{vec}(\Sigma_0))ds = \left(\mathscr{A}^{-2}\left(e^{\mathscr{A}t} - I_{d^2}\right) - \mathscr{A}^{-1}t\right)\mathrm{var}(\mathrm{vec}(\Sigma_0)).
\end{aligned}
$$

Together with (3) this shows (18).

Regarding equation (19) a combination of (14) and (18) implies

$$
\begin{aligned}
\mathrm{acov}_{\boldsymbol{\Sigma}}(h) &= \mathscr{A}^{-2}e^{\mathscr{A}\Delta(h-1)}\left(e^{2\mathscr{A}\Delta} - 2e^{\mathscr{A}\Delta} + I_{d^2}\right)\mathrm{var}(\mathrm{vec}(\Sigma_0)) \\
&= e^{\mathscr{A}\Delta(h-1)}\mathscr{A}^{-2}\left(I_{d^2} - e^{\mathscr{A}\Delta}\right)^2\mathrm{var}(\mathrm{vec}(\Sigma_0)).
\end{aligned}
$$

This shows (19) using (3).

Finally observe that

$$
\mathscr{A}\mathcal{A}X = ((A\otimes I_d) + (I_d\otimes A))^2 X + ((A\otimes I_d) + (I_d\otimes A))X\left((A^T\otimes I_d) + (I_d\otimes A^T)\right) = \mathcal{A}\mathscr{A}X
$$

for all $X \in M_{d^2}(\mathbb{R})$ and thus $\mathcal{A}$ and $\mathscr{A}$ commute. □

A.3 Proof of Theorem 3.3

Proof. The first equation in (20) is standard and the second then follows immediately from (15).

Turning to the proof of (21) we have from Barndorff-Nielsen and Stelzer (2007, Lemma 5.11) that

$$
\begin{aligned}
\mathbf{Y}_1\mathbf{Y}_1^T &= \int_0^\Delta \left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2} + \int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right) \\
&\quad + \left[\int_0^\Delta \Sigma_s^{1/2}dW_s, \int_0^\Delta dW_s^T\Sigma_s^{1/2}\right]_t^M \\
&= \int_0^\Delta \left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2} + \int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right) + \boldsymbol{\Sigma}_1 \qquad (38)
\end{aligned}
$$

referring to Barndorff-Nielsen and Stelzer (2007) for the definition of the matrix covariation $[\cdot,\cdot]^M$ and observing that we do not have to take left limits as integrals with respect to Brownian motion are necessarily continuous.

This stochastic integral representation implies that

$$
\begin{aligned}
\mathrm{var}(\mathrm{vec}(\mathbf{Y}_1\mathbf{Y}_1^T)) &= \mathrm{var}(\mathrm{vec}(\boldsymbol{\Sigma}_1)) + \mathscr{D}\mathrm{cov}\left(\mathrm{vec}(\boldsymbol{\Sigma}_1), \mathrm{vec}\left(\int_0^\Delta\left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right)\right) \\
&+ \mathscr{D}\mathrm{cov}\left(\mathrm{vec}(\boldsymbol{\Sigma}_1), \mathrm{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right) \qquad (39) \\
&+ \mathscr{D}\mathrm{cov}\left(\mathrm{vec}\left(\int_0^\Delta\left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right), \mathrm{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right) \\
&+ \mathrm{var}\left(\mathrm{vec}\left(\int_0^\Delta\left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right)\right) + \mathrm{var}\left(\mathrm{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right)
\end{aligned}
$$

setting $\mathscr{D} : M_{d^2}(\mathbb{R}) \to M_{d^2}(\mathbb{R}),\ X \mapsto X + X^T$

In the following we will now calculate the individual summands above in order to obtain an explicit expression for the variance of $\mathbf{Y}_1\mathbf{Y}_1^T$. To this end we first of all note that

$$E\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2} dW_u\right) dW_s^T \Sigma_s^{1/2}\right) = E\left(\int_0^\Delta \Sigma_s^{1/2} dW_s \left(\int_0^s dW_u^T \Sigma_u^{1/2}\right)\right) = 0.$$

Moreover, we already know from (13) that $\mathrm{var}(\mathrm{vec}(\mathbf{\Sigma}_1)) = r^{++}(\Delta) + r^{++}(\Delta)^T$.

Next we use the independence of $(\Sigma_t)_{t\in\mathbb{R}^+}$ and $(W_t)_{t\in\mathbb{R}^+}$ to obtain

$$\begin{aligned} &E\left(\mathrm{vec}(\mathbf{\Sigma}_1)\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2} dW_u\right) dW_s^T \Sigma_s^{1/2}\right)^T\right) \\ &\quad = E\left(\mathrm{vec}(\mathbf{\Sigma}_1) E\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2} dW_u\right) dW_s^T \Sigma_s^{1/2} \middle| (\Sigma_s)_{s\in[0,\Delta]}\right)^T\right) = 0. \end{aligned} \tag{40}$$

Thus

$$\mathrm{cov}\left(\mathrm{vec}(\mathbf{\Sigma}_1), \mathrm{vec}\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2} dW_u\right) dW_s^T \Sigma_s^{1/2}\right)\right) = 0$$

and likewise

$$\mathrm{cov}\left(\mathrm{vec}(\mathbf{\Sigma}_1), \mathrm{vec}\left(\int_0^\Delta \Sigma_s^{1/2} dW_s \left(\int_0^s dW_u^T \Sigma_u^{1/2}\right)\right)\right) = 0.$$

In order to calculate the remaining covariances we have to study the individual entries. In the following let $k, l, m, n \in \{1, 2, \cdots d\}$, $g := (k-1)d + l, g := (m-1)d + n$ and we write moreover $\Sigma_{ij,s}^{1/2}$ for $(\Sigma_s^{1/2})_{ij}$.

$$\begin{aligned} &E\left(\mathrm{vec}\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2} dW_u\right) dW_s^T \Sigma_s^{1/2}\right) \mathrm{vec}\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2} dW_u\right) dW_s^T \Sigma_s^{1/2}\right)^T\right)_{g,h} \\ &= E\left(\int_0^\Delta \int_0^s \sum_{p=1}^d \sum_{q=1}^d \Sigma_{pk,s}^{1/2} \Sigma_{lq,u}^{1/2} dW_{q,u} dW_{p,s} \int_0^\Delta \int_0^s \sum_{a=1}^d \sum_{b=1}^d \Sigma_{am,s}^{1/2} \Sigma_{nb,u}^{1/2} dW_{b,u} dW_{a,s}\right) \\ &\overset{(*)}{=} \int_0^\Delta \sum_{a=1}^d E\left(\int_0^s \sum_{q=1}^d \Sigma_{ak,s}^{1/2} \Sigma_{lq,u}^{1/2} dW_{q,u} \int_0^s \sum_{b=1}^d \Sigma_{am,s}^{1/2} \Sigma_{nb,u}^{1/2} dW_{b,u}\right) ds \\ &\overset{(*)}{=} \int_0^\Delta \int_0^s \sum_{a=1}^d \sum_{b=1}^d E\left(\Sigma_{ak,s}^{1/2} \Sigma_{lb,u}^{1/2} \Sigma_{am,s}^{1/2} \Sigma_{nb,u}^{1/2}\right) duds \\ &= \int_0^\Delta \int_0^s \sum_{a=1}^d \sum_{b=1}^d E\left(\Sigma_{ka,s}^{1/2} \Sigma_{lb,u}^{1/2} \Sigma_{am,s}^{1/2} \Sigma_{bn,u}^{1/2}\right) duds \\ &= \int_0^\Delta \int_0^s E\left((\Sigma_s^{1/2} \otimes \Sigma_u^{1/2})(\Sigma_s^{1/2} \otimes \Sigma_u^{1/2})\right)_{(k-1)d+l,(m-1)d+n} duds \\ &= \int_0^\Delta \int_0^s E\left(\Sigma_s \otimes \Sigma_u\right)_{(k-1)d+l,(m-1)d+n} duds. \end{aligned}$$

The $(*)$ above indicates that we have used the Itô isometry and the fact that stochastic integrals with respect to two independent Brownian motions are uncorrelated.

Thus we have established that

$$\operatorname{var}\left(\operatorname{vec}\left(\int_0^\Delta\left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right)\right)=\int_0^\Delta\int_0^s E\left(\Sigma_s\otimes\Sigma_u\right)duds$$

and calculations analogous to the above ones give

$$\operatorname{var}\left(\operatorname{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right)=\int_0^\Delta\int_0^s E\left(\Sigma_u\otimes\Sigma_s\right)duds.$$

Using again the Itô isometry twice and the uncorrelatedness of integrals with respect to independent Brownian motions we have with $g:=(k-1)d+l,\ h:=(m-1)d+n$:

$$\begin{aligned}
&E\left(\operatorname{vec}\left(\int_0^\Delta\left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right)\operatorname{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)^T\right)_{g,h}\\
&=\int_0^\Delta\int_0^s\sum_{a=1}^d\sum_{b=1}^d E\left(\Sigma_{ak,s}^{1/2}\Sigma_{lb,u}^{1/2}\Sigma_{na,s}^{1/2}\Sigma_{bm,u}^{1/2}\right)duds\\
&=\int_0^\Delta\int_0^s\sum_{a=1}^d\sum_{b=1}^d E\left(\Sigma_{ka,s}^{1/2}\Sigma_{lb,u}^{1/2}\Sigma_{an,s}^{1/2}\Sigma_{bm,u}^{1/2}\right)duds\\
&=\int_0^\Delta\int_0^s\left(\mathbf{P}E\left(\Sigma_s\otimes\Sigma_u\right)\right)_{(k-1)d+l,(m-1)d+n}duds.
\end{aligned}$$

Thus

$$\begin{aligned}
&\operatorname{cov}\left(\operatorname{vec}\left(\int_0^\Delta\left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right),\operatorname{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right)\\
&=\mathbf{P}\int_0^\Delta\int_0^s E\left(\Sigma_s\otimes\Sigma_u\right)duds.
\end{aligned}$$

Observing that $(\mathbf{P}(A\otimes B))^T_{(k-1)d+l,(m-1)d+n}=a_{ml}b_{nk}=a_{lm}b_{kn}=(\mathbf{P}(B\otimes A))_{(k-1)d+l,(m-1)d+n}$, i.e. $(\mathbf{P}(A\otimes B))^T=\mathbf{P}(B\otimes A)$, for all $A=(a_{ij}),B=(b_{ij})\in\mathbb{S}_d$, we finally obtain

$$\begin{aligned}
&\left(\operatorname{cov}\left(\operatorname{vec}\left(\int_0^\Delta\left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right),\operatorname{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right)\right)^T\\
&=\left(\mathbf{P}\int_0^\Delta\int_0^s E\left(\Sigma_s\otimes\Sigma_u\right)duds\right)^T=\mathbf{P}\int_0^\Delta\int_0^s E\left(\Sigma_u\otimes\Sigma_s\right)duds.
\end{aligned}$$

Inserting all the obtained results into (39) leads to

$$\begin{aligned}
\operatorname{var}\left(\operatorname{vec}\left(\mathbf{Y}_1\mathbf{Y}_1^T\right)\right)&=r^{++}(\Delta)+r^{++}(\Delta)^T\\
&+\left(I_{d^2}+\mathbf{P}\right)\left(\int_0^\Delta\int_0^s E\left(\Sigma_s\otimes\Sigma_u\right)duds+\int_0^\Delta\int_0^s E\left(\Sigma_u\otimes\Sigma_s\right)duds\right).
\end{aligned}\tag{41}$$

Component-wise this gives (23).

As we have that for any $X=(x_{ij}),Z=(z_{ij})\in\mathbb{S}_d$

$$\begin{aligned}
\left(\mathbf{Q}(\operatorname{vec}(X)\operatorname{vec}(Z)^T)\right)_{(k-1)d+l,(m-1)d+n}&=(\operatorname{vec}(X)\operatorname{vec}(Z)^T)_{(k-1)d+m,(l-1)d+n}\\
&=x_{mk}y_{nl}=x_{km}y_{ln}=(X\otimes Y)_{(k-1)d+l,(m-1)d+n},
\end{aligned}$$

it is immediate that $\mathbf{Q}(\text{vec}(X)\text{vec}(Y)^T) = X \otimes Y$ for all $X, Y \in \mathbb{S}_d$. Using this we get from (41):

$$\begin{aligned}
\text{var}\left(\text{vec}\left(\mathbf{Y}_1\mathbf{Y}_1^T\right)\right) &= r^{++}(\Delta) + r^{++}(\Delta)^T + (I_{d^2} + \mathbf{P})\left(E(\Sigma_0)\otimes E(\Sigma_0)\right)\Delta^2 \\
&\quad + (\mathbf{Q}+\mathbf{PQ})\left(\int_0^\Delta\int_0^s \left(E\left(\text{vec}(\Sigma_s)(\text{vec}(\Sigma_u))^T\right) - E(\text{vec}(\Sigma_0))E(\text{vec}(\Sigma_0))^T\right) duds\right. \\
&\quad \left. + \int_0^\Delta\int_0^s \left(E\left(\text{vec}(\Sigma_u)(\text{vec}(\Sigma_s))^T\right) - E(\text{vec}(\Sigma_0))E(\text{vec}(\Sigma_0))^T\right) duds\right) \\
&= r^{++}(\Delta) + r^{++}(\Delta)^T + (\mathbf{Q}+\mathbf{PQ})\left(\int_0^\Delta\int_0^s \left(\text{acov}_\Sigma(s-u) + \text{acov}_\Sigma(u-s)\right) duds\right) \\
&\quad + (I_{d^2} + \mathbf{P})\left(E(\Sigma_0)\otimes E(\Sigma_0)\right)\Delta^2
\end{aligned}$$

Together with (12) this finally shows (21).

It remains to show (22). Applying (38) we have for $h \in \mathbb{N}$:

$$\begin{aligned}
\text{acov}_{\mathbf{YY}^T}(h) &= \text{cov}\left(\text{vec}\left(\mathbf{Y}_{h+1}\mathbf{Y}_{h+1}^T\right), \text{vec}\left(\mathbf{Y}_1\mathbf{Y}_1^T\right)\right) \\
&= \text{cov}\left(\text{vec}\left(\int_{h\Delta}^{(h+1)\Delta}\Sigma_s^{1/2}dW_s\left(\int_{h\Delta}^s dW_u^T\Sigma_u^{1/2}\right)\right), \text{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right) \\
&\quad + \text{cov}\left(\text{vec}\left(\int_{h\Delta}^{(h+1)\Delta}\Sigma_s^{1/2}dW_s\left(\int_{h\Delta}^s dW_u^T\Sigma_u^{1/2}\right)\right), \text{vec}\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right)\right) \\
&\quad + \text{cov}\left(\text{vec}\left(\int_{h\Delta}^{(h+1)\Delta}\Sigma_s^{1/2}dW_s\left(\int_{h\Delta}^s dW_u^T\Sigma_u^{1/2}\right)\right), \text{vec}\left(\mathbf{\Sigma}_1\right)\right) \\
&\quad + \text{cov}\left(\text{vec}\left(\int_{h\Delta}^{(h+1)\Delta}\left(\int_{h\Delta}^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right), \text{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right) \\
&\quad + \text{cov}\left(\text{vec}\left(\int_{h\Delta}^{(h+1)\Delta}\left(\int_{h\Delta}^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right), \text{vec}\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right)\right) \\
&\quad + \text{cov}\left(\text{vec}\left(\int_{h\Delta}^{(h+1)\Delta}\left(\int_{h\Delta}^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right), \text{vec}\left(\mathbf{\Sigma}_1\right)\right) \\
&\quad + \text{cov}\left(\text{vec}\left(\mathbf{\Sigma}_{h+1}\right), \text{vec}\left(\int_0^\Delta \Sigma_s^{1/2}dW_s\left(\int_0^s dW_u^T\Sigma_u^{1/2}\right)\right)\right) \\
&\quad + \text{cov}\left(\text{vec}\left(\mathbf{\Sigma}_{h+1}\right), \text{vec}\left(\int_0^\Delta \left(\int_0^s \Sigma_u^{1/2}dW_u\right)dW_s^T\Sigma_s^{1/2}\right)\right) + \text{cov}\left(\text{vec}\left(\mathbf{\Sigma}_{h+1}\right), \text{vec}\left(\mathbf{\Sigma}_1\right)\right).
\end{aligned}$$

The independence of the increments of Brownian motion over distinct time intervals imply that the first, second, fourth and fifth covariance terms above vanish. Likewise conditioning on $(\Sigma_t)_{t\in\mathbb{R}^+}$ and using the independence of $(\Sigma_t)_{t\in\mathbb{R}^+}$ and $(W_t)_{t\in\mathbb{R}^+}$ (i.e. arguing basically as in (40)) show that the third, sixth, seventh and eighth covariance term are actually zero. Thus only the last term remains which gives $\text{acov}_{\mathbf{YY}^T}(h) = \text{cov}\left(\text{vec}\left(\mathbf{\Sigma}_{h+1}\right), \text{vec}\left(\mathbf{\Sigma}_1\right)\right) = \text{acov}_{\mathbf{\Sigma}}(h)$.

Combining this with the ARMA(1,1) property of the process $(\text{vec}(\mathbf{\Sigma}_n))_{n\in\mathbb{N}}$ we see that in the OU type stochastic volatility model the process $(\text{vec}(\mathbf{Y}_n\mathbf{Y}_n^T))_{n\in\mathbb{N}}$ has a causal ARMA(1,1) structure with autoregressive coefficient $e^{\mathscr{A}\Delta}$. □

A.4 Proof of Theorem 3.5

Proof. The stationarity follows immediately from the stationarity of the processes involved in the definition. For the martingale difference sequence property observe

$$E(u_{1,n}|\mathcal{G}_{n-1}) = E\left(E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \middle| \boldsymbol{\sigma}\left(\mathcal{G}_{n-1}, (\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}\right)\right) \middle| \mathcal{G}_{n-1}\right) = 0,$$

since obviously $\boldsymbol{\sigma}\left(\mathcal{G}_{n-1}, (\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}\right) \subseteq \boldsymbol{\sigma}\left(\mathcal{G}_{n-1}, (\Sigma_s)_{s\in\mathbb{R}^+}\right)$ and thus the Brownian increments $(W_s - W_{(n-1)\Delta})_{s\in[(n-1)\Delta,n\Delta]}$ are independent of the σ-algebra we are conditioning upon, and likewise

$$E(u_{2,n}|\mathcal{G}_{n-1}) = E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s ds + 0 + 0 - \boldsymbol{\Sigma}_n \middle| \mathcal{G}_{n-1}\right) = 0. \tag{42}$$

Taking unconditional expectations above gives that u_n has mean zero. The Itô isometry immediately implies (25).

Turning to (27) we observe that

$$\begin{aligned}\operatorname{cov}(u_{1,n}, \operatorname{vec}(u_{2,n})) = E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \operatorname{vec}\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}\right)^T\right) \quad (43)\\ + E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \operatorname{vec}\left(\Delta \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \mu^T\right)^T\right)\\ + E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \operatorname{vec}\left(\Delta\mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}\right)^T\right) - E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \operatorname{vec}\left(\boldsymbol{\Sigma}_n\right)^T\right).\end{aligned}$$

Conditioning upon $(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}$, it is once again easy to see that the last expectation in (43) vanishes. Set now for a moment $X = (x_i)_{1\le i\le d} := \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s$. Then X is conditionally upon $(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}$ a d-dimensional normal random variable and assume for a moment that the components of $X|(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}$ are moreover independent. For any $i,k,l,\in \{1,2,\ldots,d\}$ we have that $\left(X\operatorname{vec}(XX^T)^T\right)_{i,(k-1)d+l} = x_i x_k x_l$ thus clearly has zero expectation conditional upon $(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}$ due to the conditional independence and Gaussianity. Hence $E(X\operatorname{vec}(XX^T)^T|(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}) = 0$. If the components of $X|(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}$ are dependent, then due the conditional Gaussianity there is a $(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}$-measurable random matrix $C \in M_d(\mathbb{R})$ such that $CX|(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]}$ has independent components. Therefore

$$\begin{aligned}&E(X\operatorname{vec}(XX^T)^T|(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]})\\ &\quad = C^{-1}E(CX\operatorname{vec}(CX(CX)^T)^T|(\Sigma_s)_{s\in[(n-1)\Delta,n\Delta]})(C^{-T}\otimes C^{-T}) = C^{-1}0(C^{-T}\otimes C^{-T}) = 0.\end{aligned}$$

Thus also the first term in (43) above vanishes.

Let us now calculate the remaining expectations:

$$\begin{aligned}&E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \operatorname{vec}\left(\Delta \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \mu^T\right)^T\right)\\ &\quad = \Delta E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s\right)^T\right)(\mu^T \otimes I_d)\\ &\quad = \Delta E\left(\boldsymbol{\Sigma}_n\right)(\mu^T\otimes I_d) = \Delta\left(1\otimes E\left(\boldsymbol{\Sigma}_n\right)\right)(\mu^T\otimes I_d) = \Delta\left(\mu^T \otimes E\left(\boldsymbol{\Sigma}_n\right)\right) \qquad \text{and similar}\\ &E\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \operatorname{vec}\left(\Delta\mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}\right)^T\right) = \Delta\left(E\left(\boldsymbol{\Sigma}_n\right)\otimes\mu^T\right).\end{aligned}$$

Combining these formulae with (43) and (13) establishes (27).

It remains to show (26). Let $\mathscr{D}$ be defined as in the proof of Proposition 3.3

$$\begin{aligned}
\operatorname{var}(\operatorname{vec}(u_{2,n})) = \operatorname{var}&\left(\operatorname{vec}\left(\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right)\right)+\operatorname{var}\left(\operatorname{vec}\left(\mathbf{\Sigma}_n\right)\right)\\
&+\operatorname{var}\left(\operatorname{vec}\left(\Delta\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\mu^T\right)\right)+\operatorname{var}\left(\operatorname{vec}\left(\Delta\mu\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right)\right)\\
&+\mathscr{D}\operatorname{cov}\left(\operatorname{vec}\left(\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right),\operatorname{vec}\left(\Delta\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\mu^T\right)\right)\\
&+\mathscr{D}\operatorname{cov}\left(\operatorname{vec}\left(\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right),\operatorname{vec}\left(\Delta\mu\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right)\right)\\
&-\mathscr{D}\operatorname{cov}\left(\operatorname{vec}\left(\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right),\operatorname{vec}\left(\mathbf{\Sigma}_n\right)\right)\\
&+\mathscr{D}\operatorname{cov}\left(\operatorname{vec}\left(\Delta\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\mu^T\right),\operatorname{vec}\left(\Delta\mu\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right)\right)\\
&-\mathscr{D}\operatorname{cov}\left(\operatorname{vec}\left(\Delta\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\mu^T\right),\operatorname{vec}\left(\mathbf{\Sigma}_n\right)\right)\\
&-\mathscr{D}\operatorname{cov}\left(\operatorname{vec}\left(\Delta\mu\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right),\operatorname{vec}\left(\mathbf{\Sigma}_n\right)\right). \qquad (44)
\end{aligned}$$

The only term we already know above is $\operatorname{var}(\operatorname{vec}(\mathbf{\Sigma}_n)) = r^{++}(\Delta) + r^{++}(\Delta)^T$. Therefore we will now calculate the remaining covariances. As we assumed $\mu = 0$ in Theorem 3.3, we can use (21) and obtain

$$\begin{aligned}
\operatorname{var}\left(\operatorname{vec}\left(\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right)\right) =&(I_{d^2}+\mathbf{Q}+\mathbf{PQ})\operatorname{var}(\operatorname{vec}(\mathbf{\Sigma}_n))\\
&+\left(I_{d^2}+\mathbf{P}\right)\left(E(\mathbf{\Sigma}_n)\otimes E(\mathbf{\Sigma}_n)\right).
\end{aligned}$$

Using the Itô isometry once again gives:

$$\begin{aligned}
\operatorname{var}\left(\operatorname{vec}\left(\Delta\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\mu^T\right)\right) &= \Delta^2\operatorname{var}\left((\mu\otimes I_d)\operatorname{vec}\left(\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\right)\right)\\
&= \Delta^2(\mu\mu^T)\otimes E(\mathbf{\Sigma}_n)
\end{aligned}$$

and

$$\operatorname{var}\left(\operatorname{vec}\left(\Delta\mu\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right)\right) = \Delta^2 E(\mathbf{\Sigma}_n)\otimes(\mu\mu^T).$$

Defining X as in the discussion of the first summand in (43) and applying the result obtained there we get

$$\operatorname{cov}\left(\operatorname{vec}\left(\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\int_{(n-1)\Delta}^{n\Delta}dW_s^T\Sigma_s^{1/2}\right),\operatorname{vec}\left(\Delta\int_{(n-1)\Delta}^{n\Delta}\Sigma_s^{1/2}dW_s\mu^T\right)\right)$$

$$= \Delta E(\mathrm{vec}(XX^T)((\mu \otimes I_d)X)^T) = \Delta E(\mathrm{vec}(XX^T)X^T)(\mu^T \otimes I_d) = 0$$

and likewise

$$\mathrm{cov}\left(\mathrm{vec}\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}\right), \mathrm{vec}\left(\Delta\mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}\right)\right) = 0.$$

Furthermore, the standard conditioning argument and the independence of $(\Sigma_t)_{t\in\mathbb{R}^+}$ and $(W_t)_{t\in\mathbb{R}^+}$ imply

$$\begin{aligned}&\mathrm{cov}\left(\mathrm{vec}\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}\right), \mathrm{vec}\left(\mathbf{\Sigma}_n\right)\right)\\ &\quad = E(\mathrm{vec}(\mathbf{\Sigma}_n)(\mathrm{vec}(\mathbf{\Sigma}_n))^T) - E(\mathrm{vec}(\mathbf{\Sigma}_n))E(\mathrm{vec}(\mathbf{\Sigma}_n))^T = \mathrm{var}(\mathrm{vec}(\mathbf{\Sigma}_n)).\end{aligned}$$

Next we observe that

$$\begin{aligned}&\mathrm{cov}\left(\mathrm{vec}\left(\Delta \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \mu^T\right), \mathrm{vec}\left(\Delta\mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}\right)\right)\\ &\quad = \Delta^2(\mu \otimes I_d)\mathrm{var}\left(\int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s\right)(I_d \otimes \mu^T) = \Delta^2 \mu \otimes E(\mathbf{\Sigma}_n) \otimes \mu^T,\end{aligned}$$

since $\left((\mu \otimes I_d)E(\mathbf{\Sigma}_n)(I_d \otimes \mu^T)\right)_{(i-1)d+j,(k-1)d+l} = \left(\mu \otimes E(\mathbf{\Sigma}_n) \otimes \mu^T\right)_{(i-1)d+j,(k-1)d+l}$ $= \mu_i E(\mathbf{\Sigma}_{jk,n})\mu_l$ for all $i,j,k,l \in \{1,2,\ldots,d\}$.

Finally using again the conditioning and independence it is immediate that

$$\begin{aligned}&\mathrm{cov}\left(\mathrm{vec}\left(\Delta \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \mu^T\right), \mathrm{vec}\left(\mathbf{\Sigma}_n\right)\right) = 0 \quad \text{and}\\ &\mathrm{cov}\left(\mathrm{vec}\left(\Delta\mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}\right), \mathrm{vec}\left(\mathbf{\Sigma}_n\right)\right) = 0.\end{aligned}$$

Inserting all these results into (44) gives

$$\begin{aligned}\mathrm{var}(\mathrm{vec}(u_{2,n})) =&(\mathbf{Q} + \mathbf{PQ})\mathrm{var}(\mathrm{vec}(\mathbf{\Sigma}_n)) + (I_{d^2} + \mathbf{P})E(\mathbf{\Sigma}_n) \otimes E(\mathbf{\Sigma}_n)\\ &+ \Delta^2\left(E(\mathbf{\Sigma}_n) \otimes (\mu\mu^T) + (\mu\mu^T) \otimes E(\mathbf{\Sigma}_n) + \mu \otimes E(\mathbf{\Sigma}_n) \otimes \mu^T + \mu^T \otimes E(\mathbf{\Sigma}_n) \otimes \mu\right).\end{aligned}$$

□

A.5 Proof of Proposition 3.7

Proof. The i.i.d. property immediately follows from the i.i.d. property of the increments of a Lévy process over disjoint intervals of common length. To see that $(\eta_n)_{n\in\mathbb{N}}$ and $(u_n)_{n\in\mathbb{N}}$ are uncorrelated, it suffices to note that

$$E(u_n|(L_s)_{s\in\mathbb{R}^+}) = 0 \text{ for all } n \in \mathbb{N}$$

and $(\eta_n)_{n\in\mathbb{N}}$ is measurable with respect to the σ-algebra generated by $(L_s)_{s\in\mathbb{R}^+}$.

Regarding the first formula in (28) we have

$$E(\eta_{1,n}) = \int_{(n-1)\Delta}^{n\Delta} e^{A(n\Delta-s)}E(L_1)e^{A^T(n\Delta-s)}ds = -\mathbf{A}^{-1}\left(E(L_1) - e^{A\Delta}E(L_1)e^{A^T\Delta}\right)$$

and the second one follows analogously. Turning to (29) we obtain

$$\begin{aligned}\operatorname{var}(\operatorname{vec}(\eta_1)) &= \operatorname{var}\left(\int_{(n-1)\Delta}^{n\Delta} e^{\mathscr{A}(n\Delta-s)} d\operatorname{vec}(L_s)\right) = \int_{(n-1)\Delta}^{n\Delta} e^{\mathscr{A}(n\Delta-s)} \operatorname{var}(\operatorname{vec}(L_1)) e^{\mathscr{A}^T(n\Delta-s)} \\ &= -\mathcal{A}^{-1}\left(\operatorname{var}(\operatorname{vec}(L_1)) - e^{\mathscr{A}\Delta}\operatorname{var}(\operatorname{vec}(L_1))e^{\mathscr{A}^T\Delta}\right)\end{aligned}$$

and (30) is again shown along the same lines. Finally, (31) follows from

$$\begin{aligned}\operatorname{cov}(\operatorname{vec}(\eta_{1,n}), \operatorname{vec}(\eta_{2,n})) &= \operatorname{cov}\left(\int_{(n-1)\Delta}^{n\Delta} e^{\mathscr{A}(n\Delta-s)} d\operatorname{vec}(L_s), \int_{(n-1)\Delta}^{n\Delta} d\operatorname{vec}(L_s)\right) \\ &= \int_{(n-1)\Delta}^{n\Delta} e^{\mathscr{A}(n\Delta-s)} \operatorname{var}(\operatorname{vec}(L_1)) ds.\end{aligned}$$

□

References

Asai, M., M. McAleer, and J. Yu (2006). Multivariate stochastic volatility: A review. *Econometric Reviews 25*, 145–175. [2]

Ash, R. B. and M. F. Gardner (1975). *Topics in Stochastic Processes*. New York: Academic Press. [12]

Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance and Stochastics 2*, 41–68. [17]

Barndorff-Nielsen, O. E. (2001). Superposition of Ornstein–Uhlenbeck type processes. *Theory of Probability and its Applications 45*, 175–194. [17]

Barndorff-Nielsen, O. E., E. Nicolato, and N. Shephard (2002). Some recent developments in stochastic volatility modelling. *Quantitative Finance 2*, 11–23. [3]

Barndorff-Nielsen, O. E. and V. Pérez-Abreu (2008). Matrix subordinators and related Upsilon transformations. *Theory of Probability and its Applications 52*, 1–23. [4, 5]

Barndorff-Nielsen, O. E. and N. Shephard (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B 63*, 167–241. [1, 2, 3, 7, 8, 17, 19, 20, 23, 25]

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*, 253–280. [2, 23, 28]

Barndorff-Nielsen, O. E. and N. Shephard (2009). *Financial Volatility in Continuous Time*. Cambridge: Cambridge University Press. forthcoming. [3]

Barndorff-Nielsen, O. E. and R. Stelzer (2007). Positive-definite matrix processes of finite variation. *Probability and Mathematical Statistics 27*, 3–43. [2, 4, 5, 6, 30]

Bladt, M. and M. Sørensen (2005). Statistical inference for discretely observed Markov jump processes. *Journal of the Royal Statistical Society, Series B 67*, 395–410. [16]

Brockwell, P. (2001). Lévy-driven CARMA processes. *Annals of the Institute of Statistical Mathematics 53*, 113–124. [23]

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods* (2nd ed.). New York: Springer. [10, 13]

Bru, M.-F. (1991). Wishart processes. *Journal of Theoretical Probability 4*, 725– 751. [3]

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003). Alternative models for stock price dynamics. *Journal of Econometrics 116*, 225–257. [1]

Culver, W. J. (1966). On the existence and uniqueness of the real logarithm of a matrix. *Proceedings of the American Mathematical Society 17*, 1146–1151. [16]

Davydov, Y. A. (1973). Mixing conditions for Markov chains. *Theory of Probability and its Applications 18*, 312–328. [6]

Doukhan, P. (1994). *Mixing.* New York: Springer. [6, 12]

Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine processes and applications in finance. *The Annals of Applied Probability 13*, 984–1053. [9]

Dunsmuir, W. (1979). A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. *Annals of Statistics 7*, 490–506. [20]

Genon-Catalot, V., T. Jeantheau, and C. Larédo (2000). Stochastic volatility as hidden Markov models and statistical applications. *Bernoulli 6*, 1051–1079. [12]

Gouriéroux, C. (2006). Continuous time Wishart process for stochastic risk. *Econometric Reviews 25*, 177–217. [3]

Gouriéroux, C., J. Jasiak, and R. Sufana (2009). The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics forthcoming.* [3]

Griffin, J. E. and M. F. J. Steel (2006). Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility. *Journal of Econometrics 134*, 605–644. [2, 8]

Gupta, A. K. and D. K. Nagar (2000). *Matrix Variate Distributions.* Boca Raton: Chapman and Hall. [5]

Hall, A. R. (2005). *Generalized Method of Moments.* Oxford: Oxford University Press. [12, 19]

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*, 1029–1054. [12, 19]

Harvey, A. C. (1991). *Forecasting, Structural Time Series and the Kalman Filter.* Cambridge: Cambridge University Press. [20]

Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies 61*, 247–264. [2]

Horn, R. A. and C. R. Johnson (1991). *Topics in Matrix Analysis.* Cambridge: Cambridge University Press. [4, 11, 15, 16, 29]

Hubalek, F. and E. Nicolato (2009). On multivariate extensions of Lévy driven Ornstein-Uhlenbeck type stochastic volatility models and multi-asset options. In preparation. [3]

Kessler, M. and A. Rahbek (2004). Identification and inference for multivariate cointegrated and ergodic Gaussian diffusions. *Statistical Inference for Stochastic Processes 7*, 137–151. [16]

Knight, J. L. and J. Yu (2002). Empirical characteristic function in time series estimation. *Econometric Theory 18*, 691–721. [19]

Krengel, U. (1985). *Ergodic Theorems*. Berlin: Walter de Gruyter. [12]

Lindberg, C. (2005). *Portfolio Optimization and Statistics in Stochastic Volatility Markets*. PhD thesis, Department of Mathematical Sciences, Division of Mathematical Statistics, Chalmers University of Technology and Göteborg University, Göteborg, Sweden. [3]

Magnus, J. R. (1988). *Linear Structures*. New York: Oxford University Press. [15]

Magnus, J. R. and H. Neudecker (1979). The commutation matrix: Some properties and applications. *Annals of Statistics 7*, 381–394. [13]

Marquardt, T. and R. Stelzer (2007). Multivariate CARMA processes. *Stochastic Processes and their Applications 117*, 96–120. [6, 28]

Masuda, H. (2004). On multidimensional Ornstein-Uhlenbeck processes driven by a general Lévy process. *Bernoulli 10*, 97–120. [6]

Métivier, M. (1982). *Semimartingales: a Course on Stochastic Processes*. Berlin: Walter de Gruyter. [4]

Métivier, M. and J. Pellaumail (1980). *Stochastic Integration*. New York: Academic Press. [4]

Øksendal, B. (2003). *Stochastic Differential Equations* (6th ed.). New York: Springer. [4]

Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics 140*, 425–449. [8]

Pigorsch, C. and R. Stelzer (2009). On the definition, stationary distribution and second order structure of positive semi-definite Ornstein-Uhlenbeck type processes. *Bernoulli forthcoming*. [5, 15, 20]

Protter, P. E. (2005). *Stochastic Integration and Differential Equations* (2nd ed.). New York: Springer. [4, 6, 9]

Rheinländer, T. and G. Steiger (2006). The minimal entropy martingale measure for general Barndorff-Nielsen/Shephard models. *The Annals of Applied Probability 16*, 1319–1351. [2]

Roberts, G. O., O. Papaspiliopoulos, and P. Dellaportas (2004). Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. *Journal of the Royal Statistical Society: Series B 66*, 369–393. [2]

Sato, K. and M. Yamazato (1984). Operator-selfdecomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Processes and their Applications 17*, 73–100. [9]

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press. [5]

Singleton, K. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics 102*, 111–141. [19]

Sørensen, M. (2000). Prediction-based estimating functions. *Econometrics Journal 3*, 123–147. [12]

Todorov, V. and G. Tauchen (2006). Simulation methods for Lévy-driven CARMA stochastic volatility models. *Journal of Business & Economic Statistics 24*, 455–469. [23]

Todorov, V. and G. Tauchen (2008). Volatility jumps. Working Paper. [3]