

# A multivariate non-Gaussian stochastic volatility model with leverage for energy markets

Fred Espen Benth and Linda Vos \*

## Abstract

Spot prices in energy markets exhibit special features like price spikes, mean-reversion inverse, stochastic volatility, inverse leverage effect and co-integration between the different commodities. In this paper a multivariate stochastic volatility model is introduced which captures these features. Second order structure and stationary issues of the model are analysed. Moreover the implied multivariate forward model is derived. Due to the flexibility of the model stylized facts of the forward curve as contango, backwardation and humps are explained. Moreover, a transformed-based method to price options on the forward is described, where fast and precise algorithms for price computations can be implemented. A simulation method for Monte Carlo generation of price paths is introduced.

## 1 Introduction

Energy markets world-wide have been liberalized over the last decades to create liquid trading arenas for power commodities like electricity, gas, and coal. The markets are continuously developing, and in recent years gradually becoming more and more connected. For instance, interconnectors between UK, Scandinavia and continental Europe make the various power markets more and more integrated. Also, different electricity markets on the continental Europe exchange to a large extent energy across borders. A reflection of this market integration is the growing interest for multivariate price models for power. This includes cross-commodity models for gas and electricity, say, but also models for the same commodity traded in different but integrated markets.

Power market prices have by now well-known characteristics like distinct price spikes in periods with supply and demand imbalances due to the inelasticity of the supply curve. Further, the markets are typically varying by season, with high prices in cold periods due to heating, or, in warmer climate zones, high prices in summer due to air-conditioning cooling. Prices also naturally mean-revert due to demand and supply. Partly because of the large spikes, the prices observed in markets like gas and electricity are to a large extent leptokurtic. In fact, volatility may easily reach above 100%. A discussion of these features of power spot prices can be found in Geman [22] and Eydeland and Wolyniec [19]. There exists many models for spot price dynamics in power markets, and we refer to Benth *et al.* [10] for an overview and analysis.

In energy markets there is evidence of a so-called inverse leverage effect. The volatility tends to increase with the level of power prices, since there is a negative relationship between inventory and prices (see for instance Deaton and Leroque [16]). Little available inventory means higher and more volatile prices. There is also evidence for co-integration between different commodities. For instance, it is unlikely that the price of gas and electricity in the UK will drive apart since gas is a major fuel for power production. Likewise, since gas can be transported as LNG, different gas markets can not have prices which become increasingly different.

In recent years there have been a concern about stochastic volatility models for commodities, and in particular energies. In Hiksipoors and Jaimungal [23] we find an analysis of forward pricing in commodity markets in the presence of stochastic volatility. Several popular models are treated. More recently, Schwartz

---

\*Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, N-0316 Oslo Norway and Department of Economics, University of Agder, Serviceboks 422, N-4604 Kristiansand, Norway, email: fredb@math.uio.no, linda.vos@cma.uio.no, http://www.math.uio.no/

and Trolle [38] introduced the notion of unspanned volatility, and analysed this in power markets. Their statistical analysis confirms the presence of stochastic volatility in commodity markets. Benth [8] applied the Barndorff-Nielsen and Shepard stochastic volatility model in commodity markets, and derived forward prices based on this. An empirical study on UK gas prices was performed.

In this paper we propose a stochastic dynamics for cross-commodity spot price modelling generalizing the univariate dynamics studied in Benth [8]. The model is flexible enough to capture spikes, mean-reversion and stochastic volatility. Moreover, it includes the possibility to model inverse leverage and stochastic volatility. Our proposed dynamics can model co- and independent jump behaviour (spikes) in cross-commodity markets. Despite its flexible and complex nature, the dynamics of the curve of forward prices is analytically computable. It turns out that the implied forward curves can be in contango and backwardation, as well as having humps. As has been pointed out by Geman [22], hump-shaped forward curves have been observed in for instance the oil market. Due to the flexibility of the multivariate model, even an oscillation of the forward price to the spot can be achieved. Moreover, by using Fourier methods options on spreads can be efficiently computed numerically. In some degenerate cases we can link our proposed model to co-integration in continuous time as proposed by Kessler and Rahbeck [27].

The dynamics we are considering is based on Ornstein-Uhlenbeck processes driven by multivariate subordinators. The logarithmic price dynamics are defined by multi-factor processes and seasonal functions to account for deterministic variability over a year. The stochastic volatility processes are multi-variate as well, so that we can incorporate second-order dependencies between commodities. The volatility model is adopted from the so-called Barndorff-Nielsen and Shephard model (BNS for short), extended to a multivariate setting (see Barndorff-Nielsen and Shephard [5] and Barndorff-Nielsen and Stelzer [7]). As for the multi-dimensional extension, the volatility is modeled with a matrix-valued Ornstein-Uhlenbeck process driven by a positive definite matrix-valued subordinator (see Barndorff-Nielsen and Pérez-Abreu [4]). We prove that the spot prices are stationary, and compute the characteristic function of the stationary distribution. Several other probabilistic features of the model are presented and discussed, demonstrating its flexibility in modelling prices and its analytical tractability. From a more practical point of view, a method for simulating the prices is presented. Various special cases with interest in power markets are analysed, and in particular we provide an empirical example where the algorithm is applied. Our approach is influenced by the work of Stelzer [40].

The paper is organized as follows. Section 2 introduces the spot model, thereafter the stationary distribution and the probabilistic properties of the model are deduced in Section 3. Section 4 gives an empirical example and a method to perform Monte-Carlo simulation of the model. Moreover the issue of incorporating co-integration in the model is discussed. In Section 4 the implied multivariate forward dynamics derived and properties of the forward curve are described. In Section 6 a transform based method to price spread options is given. Section 7 concludes. In the Appendix some proofs are collected.

## Notation

For the convenience of the reader, we have collected some frequently used notations. We adopt the notation used by Pirgorsch and Stelzer [32]. Throughout this paper we write  $\mathbb{R}^+$  for the positive real numbers and we denote the set of real  $n \times n$  matrices by  $M_n(\mathbb{R})$ . We denote the group of invertible matrices by  $GL_n(\mathbb{R})$ , the linear subspace of symmetric matrices by  $\mathbb{S}_n$ , the positive definite cone of symmetric matrices by  $\mathbb{S}_n^+$ .  $I_n$  stands for the  $n \times n$  identity matrix,  $J_n(v)$  is an operator  $\mathbb{R}^n \rightarrow M_n(\mathbb{R})$  which creates a diagonal matrix with the vector  $v \in \mathbb{R}^n$  on the diagonal,  $diag(A)$  is a vector in  $\mathbb{R}^n$  consisting of the diagonal of the matrix  $A \in M_n(\mathbb{R})$ ,  $\sigma(A)$  denotes the spectrum (the set of all eigenvalues) of a matrix  $A \in M_n(\mathbb{R})$ . 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. Furthermore, we denote  $tr(A)$  for the trace of the matrix  $A \in M_n(\mathbb{R})$ , which is the sum of the elements on the diagonal. The transpose of the matrix  $A \in M_n(\mathbb{R})$  is denoted  $A^T$  while  $A_{ij}$  is the element of  $A$  in the  $i$ -th row and  $j$ -th column. This notation is extended to the processes in a natural way. Finally, the unit vector with one on the  $i$ -th place is denoted  $e_i$ .

Regarding all random variables and processes we assume that they are defined on a given complete

filtered probability space  $(\Omega, \mathcal{F}, P)$  equipped with the filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  satisfying the usual conditions.<sup>1</sup> Furthermore, we employ an intuitive notation for the stochastic integration with respect to a matrix-valued integrator: let  $\{L(t)\}_{t \in \mathbb{R}^+}$  in  $M_n(\mathbb{R})$  be a semimartingale and  $\{A(t)\}_{t \in \mathbb{R}^+}, \{B(t)\}_{t \in \mathbb{R}^+} \in M_n(\mathbb{R})$  be adapted integrable with respect to the  $L$  process. Then  $C(t) := \int_0^t A_s dL_s B_s$  is a matrix in  $M_n(\mathbb{R})$  which has  $ij$ -th element  $C_{ij}(t) = \sum_{k=1}^n \sum_{l=1}^n \int_0^t A_{ik} B_{j,s} dL_{kl,s}$ . The logarithmic characteristic function of a stochastic process  $A_i(t)$  is denoted by  $\phi_A^i$ . Suppose  $A \in M_n(\mathbb{R})$ , then with bold face we denote the operator  $\mathbf{A}$  associated with the matrix  $A$  and defined as  $\mathbf{A} : X \mapsto AX + XA^T$ . Its inverse is denoted by  $\mathbf{A}^{-1}$  and can be represented as  $\text{vec}^{-1} \circ ((A \otimes I_n) + (I_n \otimes A))^{-1} \circ \text{vec}$ .

## 2 The stochastic volatility model

Assume  $m, n \in \mathbb{N}$  with  $0 \leq m < n$ . Let  $\{\tilde{L}_j(t)\}_{t \in \mathbb{R}^+} \in \mathbb{S}_d^+, j = 1, \dots, n$  be  $n$  independent matrix-valued subordinators as introduced in Barndorff-Nielsen and Pérez-Abreu [4]. Furthermore, let  $L_i, i = 1, \dots, m$  be  $\mathbb{R}^d$ -valued subordinators<sup>2</sup>. For  $i = 1, \dots, m$  the vector-valued subordinators  $L_i$  are formed by taking the diagonal of the matrix-valued subordinators  $\tilde{L}_i(t)$ . This implies that  $L_i$  will jump whenever  $\tilde{L}_i$  does. If one of the off-diagonal elements jumps, also the diagonal element has to jump in order to keep the volatility process  $\Sigma(t)$  in the positive definite cone  $\mathbb{S}_d^+$ . The subordinators are assumed to be driftless. Moreover, let  $W$  be a standard  $\mathbb{R}^d$ -valued Brownian motion independent of the subordinators.

We define the spot price dynamics of  $d$  commodities as follows. Let

$$S(t) = \Lambda(t) \cdot \exp \left( X(t) + \sum_{i=1}^m Y_i(t) \right), \quad (2.1)$$

where  $\Lambda : [0, T] \mapsto \mathbb{R}^d$  is a vector of bounded measurable seasonality functions,  $\cdot$  denotes point-wise multiplication, and

$$dX(t) = AX(t) dt + \Sigma(t)^{1/2} dW(t), \quad (2.2)$$

$$dY_i(t) = (\mu_i + B_i Y_i(t)) dt + \eta_i dL_i(t), \quad (2.3)$$

for  $i = 1, \dots, m$ .  $A, B_i$ 's and  $\eta_i$  are in  $GL_d(\mathbb{R})$  and  $\mu_i$  is a vector in  $\mathbb{R}^d$ . To ensure the existence of stationary solutions we assume that the eigenvalues of the matrices  $A, B_i$  have negative real-parts. The entries of  $\eta_i$  can be negative. So although  $L_i$  is a  $\mathbb{R}^d$ -valued subordinator, there can be negative jumps in the spot-price process.

The stochastic volatility process  $\Sigma(t)$  is a superposition of positive-definite matrix valued Ornstein-Uhlenbeck process as introduced in Barndorff-Nielsen and Stelzer [7],

$$\Sigma(t) = \sum_{j=1}^n \omega_j Z_j(t), \quad (2.4)$$

with

$$dZ_j(t) = (C_j Z_j(t) + Z_j(t) C_j^T) dt + d\tilde{L}_j(t), \quad (2.5)$$

and the  $\omega_j$ 's are weights summing up to 1. Moreover,  $\{C_j\}_{1 \leq j \leq n} \in GL_d(\mathbb{R})$ . To ensure a stationary solution we will assume that the eigenvalues of  $C_j$  have negative real-parts. This stochastic volatility model is a multivariate extension of the so called BNS SV model introduced by Barndorff-Nielsen and Shephard [5] for general asset price processes. The commodity spot price dynamics with the BNS SV model as stochastic volatility structure is a generalization to cross-commodity markets of the univariate spot model analysed in Benth [8].

Note that  $Y_i$  and  $\Sigma_i$  for  $i = 1, \dots, m$  have related subordinators  $L$  and  $\tilde{L}$  driving the noise. Thus, when the volatility component  $\Sigma$ , jumps, we observe simultaneously a change in the spot price. Hence, we can have an inverse leverage effect since increasing prices follows from increasing volatility, and vice versa.

<sup>1</sup> see e.g. Protter [34]

<sup>2</sup> A multivariate subordinator is a Lévy process which is increasing in each of its coordinates (see Sato [2]).

We also have  $n - m$  independent stochastic volatility components  $Z_j, j = m + 1, \dots, n$  that do not directly influence the price process paths but have a second order effect. The processes  $Y_i$  can be interpreted as modeling the spikes, whereas  $X$  is the normal variations in the market.

By turning off the processes  $Y_i$  (choose  $\mu_i = \eta_i = 0$  and  $B_i = 0$  for all  $i$ ), we obtain a multivariate extension of the Schwartz model with stochastic volatility and stock-price dynamics:

$$S(t) = \Lambda(t) \cdot \exp(X(t)) \quad (2.6)$$

where  $X(t)$  is defined in (2.2). The Schwartz model with constant volatility is a mean-reversion process proposed by Schwartz [37] for spot price dynamics in commodity markets like oil.

In electricity markets one observes spikes in the spot price dynamics (see e.g. Benth *et al.* [10]). These spikes often occur seasonally. In the Nordic electricity market Nord-Pool, price spikes occur in the winter time when demand is high. We therefore may wish the jump frequency of the subordinators  $L_i, i = 1, \dots, m$  to be time-dependent. This is not possible when working with Lévy processes, but we can generalize to independent increment processes instead (see Jacod and Shiryaev [26]). Independent increment processes are time-inhomogeneous Lévy processes. Our modeling and analysis to come are easily modified to include such processes. To keep matters slightly more simplified, we keep to the time-homogeneous case here. The interested reader is referred to Benth *et al* [10] for applications of independent increment processes in energy markets.

We assume the following log integrability conditions for the subordinators.

$$\mathbb{E} \left[ \log^+ \|\tilde{L}_j(1)\| \right] < \infty, \Rightarrow \mathbb{E} \left[ \log^+ \|L_j(1)\| \right] < \infty. \quad (2.7)$$

Where  $\log^+(x)$  is defined as  $\max(\log(x), 0)$ .

### 3 Stationarity and probabilistic properties of the spot price

The processes  $X, Y_i$  are Ornstein-Uhlenbeck processes. By diagonalization of the matrices  $A$  and  $B_i, i = 1, \dots, m$ , one can reduce the stochastic differential equations (2.2) and (2.3) to one-dimensional problems. Applying the one-dimensional Itô formula to these problems and converting back to matrices yields the following solution, for  $0 \leq s \leq t$ ,

$$X(t) = e^{A(t-s)} X(s) + \int_s^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u), \quad (3.1)$$

$$Y_i(t) = e^{B_i(t-s)} Y_i(s) + B_i^{-1} (I - e^{B_i(t-s)}) \mu_i + \int_s^t e^{B_i(t-u)} \eta_i dL_i(u), \quad (3.2)$$

for  $i = 1, \dots, m$ . The matrix exponentials are defined as usual as  $e^A := I + \sum_{i=1}^{\infty} \frac{A^i}{i!}$ . According to Barndorff-Nielsen and Stelzer [7], Sect. 4, the solution of  $Z_j(t), j = 1, \dots, n$ , is given by

$$Z_j(t) = e^{C_j(t-s)} Z_j(s) e^{C_j^T(t-s)} + \int_s^t e^{C_j(t-u)} d\tilde{L}_j(u) e^{C_j^T(t-u)}.$$

We find the following conditional expectations for the processes involved:

**Lemma 3.1.** *Suppose that all processes are in  $L^1$ , then it holds*

$$\begin{aligned} \mathbb{E}[X(t)|\mathcal{F}_s] &= e^{A(t-s)} X(s), \\ \mathbb{E}[Y(t)|\mathcal{F}_s] &= e^{B_i(t-s)} Y_i(s) + B_i^{-1} (I - e^{B_i(t-s)}) \mu_i + B_i^{-1} \left( \eta_i - e^{B_i(t-s)} \eta_i \right) \mathbb{E}[L_i(1)]. \end{aligned}$$

*Proof.* The proof of Lemma 3.1 is given in the Appendix in section A.1. □

The second-order characteristics of the processes involved are given in the next Lemma:

**Lemma 3.2.** *Assume that all the processes are square-integrable, then it holds*

$$\begin{aligned}\text{Var}[X(t)|\mathcal{F}_s] &= \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} \left( e^{C_j(t-s)} Z_j(s) e^{C_j^T(t-s)} - e^{A(t-s)} Z_j(s) e^{A^T(t-s)} \right. \\ &\quad \left. + \mathbf{C}_j^{-1} \left[ \mathbb{E}[\tilde{L}_j(1)] - e^{C_j(t-s)} \mathbb{E}[\tilde{L}_j(1)] e^{C_j^T(t-s)} \right] \right. \\ &\quad \left. - \mathbf{A}^{-1} \left[ \mathbb{E}[\tilde{L}_j(1)] - e^{A(t-s)} \mathbb{E}[\tilde{L}_j(1)] e^{A^T(t-s)} \right] \right), \\ \text{Var}[Y_i(t)|\mathcal{F}_s] &= \mathbf{B}_i^{-1} \left( \eta_i \text{Var}[L(1)] \eta_i^T - e^{B_i(t-s)} \eta_i \text{Var}[L(1)] \eta_i^T e^{B_i^T(t-s)} \right) \\ &\quad - \left( B_i^{-1} \left( \eta_i - e^{B_i(t-s)} \eta_i \right) \mathbb{E}[L_i(1)] \right)^2.\end{aligned}$$

Hence, in stationarity,

$$\begin{aligned}\lim_{t \rightarrow \infty} \text{Var}[X(t)] &= \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} [\mathbf{C}_j^{-1} - \mathbf{A}^{-1}] \mathbb{E}[\tilde{L}(1)], \\ \lim_{t \rightarrow \infty} \text{Var}[Y_i(t)] &= \mathbf{B}_i^{-1} \eta_i \text{Var}[L(1)] \eta_i^T - (B_i^{-1} \eta_i \mathbb{E}[L_i(1)])^2.\end{aligned}$$

Furthermore, we have that  $\text{Cov}[X, Y_i] = 0$  and  $\text{Cov}[Y_i, Y_j]$  for  $i \neq j$ , and, in stationarity, the auto-covariance functions are given by,

$$\begin{aligned}acov_X(h) &= \lim_{t \rightarrow \infty} e^{A|h|} \text{Var}[X(t)] = e^{A|h|} \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} [\mathbf{C}_j^{-1} - \mathbf{A}^{-1}] \mathbb{E}[\tilde{L}(1)] \\ acov_{\sum Y_i}(h) &= \lim_{t \rightarrow \infty} \sum_i e^{B_i|h|} \text{Var}[Y_i(t)] = \sum_{i=1}^m e^{B_i|h|} (\mathbf{B}_i^{-1} \eta_i \text{Var}[L(1)] \eta_i^T - (B_i^{-1} \eta_i \mathbb{E}[L_i(1)])^2).\end{aligned}$$

*Proof.* The proof of Lemma 3.2 is given in the Appendix in section A.2.  $\square$

Note that the conditional quantities are exponentially fast converging to their stationary limits. Moreover the conditional variance is a stochastic variable depending on the volatility process  $\Sigma(t)$ . This reflects the property of stochastic volatility. The auto-covariance function are decaying exponentially fast to zero, which confirms the mean-reverting property of the spot model.

Under the log integrability conditions (2.7), the processes  $Y_i$  and  $Z_j$  are stationary (see Sato [36], Thm. 5.2). In the next Proposition the characteristic function of the stationary distributions are calculated in terms of the characteristic function of the matrix-valued processes  $\tilde{L}_j$ .

For notational simplicity we define the linear operator  $\mathcal{E}_j$ ,

$$\mathcal{E}_j(t) : X \mapsto \omega_j \left[ (\mathbf{C}_j - \mathbf{A})^{-1} \left( e^{C_j t} X e^{C_j^T t} - e^{A t} X e^{A^T t} \right) \right]. \quad (3.3)$$

**Proposition 3.3.** *Suppose the matrices  $A$  and  $C_j$  commute for  $j = 1, \dots, n$ . The stationary distribution of the process  $X$  has characteristic function*

$$\phi_X(w) = \sum_{j=1}^n \int_0^\infty \phi_{\tilde{L}}^j \left( \frac{1}{2} w w^T \mathcal{E}_j(s) I_d \right) ds,$$

Furthermore, the processes  $Y_i$  exist for  $i = 1, \dots, m$  and have characteristic function for the stationary distribution given by

$$\phi_Y^i(w) = i \mu_i^T (B_i^T)^{-1} w + \int_0^\infty \phi_{\tilde{L}}^i (J_d (\eta_i^T e^{B_i^T u} w)) du, \quad w \in \mathbb{R}^d.$$

Moreover, the characteristic function of the stationary distribution of  $Z_j$ ,  $j = 1, \dots, n$  is given by,

$$\phi_Z^j(V) = \int_0^\infty \phi_L^j \left( e^{C_j^T s} V e^{C_j s} \right) ds, \quad V \in \mathbb{S}_d.$$

*Proof.* The proof of proposition 3.3 is given in the Appendix in section A.3.  $\square$

We observe that the limiting distribution of  $X$  must be centered and symmetric since  $\phi^X(w) = \phi^X(-w)$ . Also, note that the limiting distribution of  $X$  is in fact the limiting distribution of  $\ln \tilde{S}(t)$  for the multi-variate Schwartz model (where  $m = 0$ ). We discuss the stationary distribution of  $X$  in more detail.

The stationary distribution of  $X$  can be viewed as the convolution of a centered normal and a leptokurtic distribution. Since the subordinators considered are assumed driftless, the cumulant is

$$\phi_L^j(V) = \int_{\mathbb{S}_d^+ \setminus \{0\}} \{e^{tr(XV)} - 1\} \nu_L^j(dX).$$

Then, introduce the function

$$\tilde{\phi}_L^j(V) \triangleq \phi_L^j(V) - tr \left( V \int_{\mathbb{S}_d^+ \setminus \{0\}} X \nu^j(dX) \right),$$

which becomes the characteristic function of the zero-mean matrix valued Lévy process  $\hat{L}(t) \triangleq \tilde{L}(t) - \mathbb{E}[\tilde{L}(t)]$ . The characteristic function of the stationary distribution of  $X(t)$  can henceforth be expressed as

$$\begin{aligned} \phi^X(w) &= \sum_{j=1}^n \int_0^\infty \tilde{\phi}_L^j \left( \frac{1}{2} w w^T \mathcal{E}_j(s) I_d \right) ds + tr \left( \frac{1}{2} w w^T \int_0^\infty \mathcal{E}_j(s) \mathbb{E}[\tilde{L}(1)] ds \right) \\ &= \sum_{j=1}^n \int_0^\infty \tilde{\phi}_L^j \left( \frac{1}{2} w w^T \mathcal{E}_j(s) I_d \right) ds + \sum_{j=1}^n \frac{\omega_j}{2} tr \left( w w^T (\mathbf{C}_j - \mathbf{A})^{-1} [(\mathbf{C}_j^{-1} - \mathbf{A}^{-1}) \mathbb{E}[\tilde{L}(1)]] \right) \\ &= \sum_{j=1}^n \int_0^\infty \tilde{\phi}_L^j \left( \frac{1}{2} w w^T \mathcal{E}_j(s) I_d \right) ds + \frac{1}{2} tr \left( w w^T \lim_{t \rightarrow \infty} \text{Var}[X(t)] \right). \end{aligned}$$

Where we used Lemma 3.2 to establish the last equality. The last term is the characteristic function of a centered multivariate normal distribution with variance equal to  $\lim_{t \rightarrow \infty} \text{Var}[X(t)]$ , which coincides with the stationary distribution obtained from the multivariate Schwartz model having constant volatility  $\Sigma \in M_d(\mathbb{R})$  given by

$$\Sigma \triangleq \lim_{t \rightarrow \infty} \text{Var}[X(t)].$$

The first term in  $\phi^X(x)$  will be the characteristic function of a non-Gaussian distribution.

Let us look at the dynamics of  $\tilde{S}(t) \triangleq S(t)/\Lambda(t)$ , the deseasonalized spot price, where the division is done elementwise. From Lemma 3.1, we find in stationarity that

$$\lim_{t \rightarrow \infty} \mathbb{E}[\ln(\tilde{S}(t))] = \lim_{t \rightarrow \infty} \mathbb{E}[X(t)] + \sum_{i=1}^m \mathbb{E}[Y_i(t)] = \sum_{i=1}^m B_i^{-1} (\mu_i + \eta_i \mathbb{E}[L_i(1)]).$$

Furthermore, from Lemma 3.2 we have that in stationarity, the auto-covariance function of  $\ln \tilde{S}(t)$  is

$$\begin{aligned} acov_{\ln \tilde{S}}(h) &= acov_X(h) + acov_{\sum Y_i}(h) \\ &= e^{A|h|} \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} [\mathbf{C}_j^{-1} - \mathbf{A}^{-1}] \mathbb{E}[\tilde{L}(1)] \\ &\quad + \sum_{i=1}^m e^{B_i|h|} (\mathbf{B}_i^{-1} \eta_i \text{Var}[L(1)] (\eta_i)^T - (B_i^{-1} \eta_i \mathbb{E}[L_i(1)])^2). \end{aligned} \tag{3.4}$$

Moreover, we can find the dynamics of  $\ln \tilde{S}(t)$ .

**Proposition 3.4.** *It holds that*

$$d \ln \tilde{S}(t) = \left( M(t) + A \ln \tilde{S}(t) \right) dt + \Sigma(t)^{1/2} dW(t) + \sum_{i=1}^m \eta_i dL_i(t),$$

where

$$M(t) = \sum_{i=1}^m \mu_i + (-A + B_j) Y_j(t).$$

*Proof.* This follows from rewriting the equations in (2.2) and (2.3).  $\square$

From the proposition we see that the logarithm of the deseasonalized spot prices is a mean-reverting jump-diffusion model with stochastic volatility in the diffusion term. Moreover, the mean level is stochastic and defined by  $M(t)$ . The process  $M(t)$  is a linear combination of the spike processes  $Y_i$ , which are all mean-reverting to the level  $\mu_i$  and stationary in the limit, this implies that the mean level  $M$  itself is stationary in the limit.

**Proposition 3.5.** *The characteristic function of the stationary distribution of  $\ln \tilde{S}(t)$  is given by*

$$\begin{aligned} \phi_{\ln \tilde{S}}(w) &= \sum_{i=1}^m i \mu_i^T (B_i^T)^{-1} w + \sum_{j=1}^n \int_0^\infty \phi_{\tilde{L}}^j \left( \frac{1}{2} w w^T \mathcal{E}_j^T(u) I_d \right) du \\ &\quad + \sum_{i=1}^m \int_0^\infty \phi_{\tilde{L}}^i \left( \frac{1}{2} w w^T \mathcal{E}_i^T(u) I_d + J_d(\eta_i^T e^{B_i^T u} w) \right) - \phi_{\tilde{L}}^i \left( \frac{1}{2} w w^T \mathcal{E}_i^T(u) I_d \right) du. \end{aligned}$$

*Proof.* By combining Proposition A.3, equation (A.2) the conditionally characteristic function of  $\ln \tilde{S}$  given  $\mathcal{F}_s$  is

$$\begin{aligned} \phi_{\ln \tilde{S}}^{s,t}(w) &= iX^T(s) e^{A^T(t-s)} w + \sum_{j=1}^n \text{itr} \left( \frac{1}{2} w w^T \mathcal{E}_j(t-s) Z_j(s) \right) \\ &\quad + \sum_{i=1}^m iY_i^T(s) e^{B_i^T(t-s)} w + i(B_i^{-1}(I - e^{B_i(t-s)}) \mu_i)^T w \\ &\quad + \sum_{k=1}^m \ln \mathbb{E} \left[ e^{i \text{tr} \left( \left( \frac{1}{2} w w^T \int_s^t \mathcal{E}_k(t-u) d\tilde{L}_k(u) \right)^T I_d \right) + i \text{tr} \left( \left( \int_s^t J_d(\eta_k^T e^{B_k^T u} w) d\tilde{L}_k(u) \right)^T I_d \right)} \right] \\ &\quad + \sum_{j=m_1+1}^n \int_0^{t-s} \phi_{\tilde{L}}^j \left( \frac{1}{2} w w^T \mathcal{E}_j^T(u) I_d \right) \\ &= iX^T(s) e^{A^T(t-s)} w + \sum_{j=1}^n \text{itr} \left( \frac{1}{2} w w^T \mathcal{E}_j(t-s) Z_j(s) \right) \\ &\quad + \sum_{i=1}^m iY_i^T(s) e^{B_i^T(t-s)} w + i(B_i^{-1}(I - e^{B_i(t-s)}) \mu_i)^T w \\ &\quad + \sum_{i=1}^m \int_0^{t-s} \phi_{\tilde{L}}^i \left( \frac{1}{2} w w^T \mathcal{E}_i^T(u) I_d + J_d(\eta_i^T e^{B_i^T u} w) \right) du \\ &\quad + \sum_{j=m_1+1}^n \int_0^{t-s} \phi_{\tilde{L}}^j \left( \frac{1}{2} w w^T \mathcal{E}_j^T(u) I_d \right). \end{aligned}$$

Since a stationary solution exists for  $X$  and all  $Y_i$ 's, there also exists a stationary solution for  $\ln \tilde{S}$ . The result follows by taking limits for  $t \rightarrow \infty$ .  $\square$

Note that the sum over  $j$  in the expression for  $\phi_{\ln \tilde{S}}$  is stemming from the stationary cumulant of  $X$ , and therefore is from a symmetric centered random variable. Stationarity is a desirable feature in commodity markets being a reflection of supply and demand-driven prices. However, many studies argue for non-stationary effects (like for example Burger *et al.* [13] studying German electricity spot prices). We can easily extend our model to include non-stationary factors, like for instance choosing one or more of the  $Y$ 's to be drifted Brownian motions rather than Ornstein-Uhlenbeck processes. We shall not discuss these modelling issues further here, but concentrate on stationary models.

In the special case of a multivariate stochastic volatility Schwartz model (i.e.  $m = 0$ ) the “reversion-adjusted” logreturns are approximately distributed according to a multivariate mean-variance mixture model. Considering the “reversion-adjusted” logreturns over the time interval  $[t, t + \tau]$ , we find

$$\begin{aligned} \ln \tilde{S}(t + \tau) - e^{A\tau} \ln \tilde{S}(t) &= X(t + \tau) - e^{A\tau} X(t) \\ &= \int_t^{t+\tau} e^{A(t+\tau-s)} \Sigma^{1/2}(s) dW(s) \\ &\approx e^{A\tau} \Sigma^{1/2}(t) \Delta_\tau W(t). \end{aligned}$$

Here,  $\Delta_\tau W(t) \triangleq W(t+\tau) - W(t)$ . Hence, we have that “reversion-adjusted” logreturns are approximately distributed according to the multivariate mean-variance mixture model

$$e^{A\tau} \Sigma^{1/2}(t) \Delta_\tau W(t) \Big|_{\Sigma(t)} \sim \mathcal{N}(0, e^{A\tau} \Sigma(t) e^{A^T \tau}).$$

In Benth [8], this was discussed in the univariate case, showing that we can choose stochastic volatility models yielding for instance normal inverse Gaussian distributed “reversion-adjusted” returns. We refer to Benth and Saltyte-Benth [9] for a study of gas and oil prices where the normal inverse Gaussian distribution has been applied to model “reversion-adjusted” returns. We further note that the conditional Gaussian structure of the “reversion-adjusted” returns implies that the covariance is determining the cross-commodity dependency. In this case it is given explicitly by the stochastic volatility model  $\Sigma(t)$ , introducing a time-dependency in the covariance between commodities. In addition, the common factors  $Y_i(t)$ ,  $i = 1, \dots, m$  will give co-dependent paths determined by common jump paths. Hence, we can mix rather complex dependency into the modelling. The auto-covariance function of the de-seasonalized logarithmic spot (3.4) gives explicit formulation to this dependence in terms of second order structure. For  $h = 0$  the auto-covariance of de-seasonalized logarithmic spots gives the covariance matrix of the de-seasonalized logarithmic spots.

## 4 Simulation of matrix-valued subordinators and co-integration

In this section we discuss simulation of our spot price dynamics, which essentially means to discuss simulation of matrix-valued subordinators. In addition, we analyse the connection to co-integration for our spot model. Co-integration is a popular tool for modelling multivariate dynamics in commodities, and we wish to link our approach to this.

### 4.1 Simulation

Limited literature is available on the simulation of matrix-valued subordinators. One possible approach could be to apply existing methods to sample multivariate Lévy processes based on their Lévy measures by iterative sampling from the conditional marginals (see *e.g.* Cont and Tankov [41]). However, the marginal distribution functions are required, which are not always available in a simple form. Moreover, in case of matrix-valued subordinators, the restriction of the domain to the positive definite cone make matters even more complicated. We introduce a simple approximate algorithm<sup>3</sup> to simulate from matrix-valued compound Poisson, stable, and tempered stable processes with stable or constant jump-size distribution. A simulation study confirms the stylized facts of the model previously discussed.

<sup>3</sup>The idea of the algorithm was kindly proposed to us by Robert Stelzer.



For any  $X \in \mathbb{S}_d^+$  one can make a polar decomposition in a ray  $r = \|X\| = \text{tr}(XX)^{1/2}$  and angle  $\Theta = X/r$ , so that  $X = r\Theta$ . Moreover  $\Theta$  is situated on the unit sphere  $\mathbf{S}$  of  $\mathbb{R}^{d \times d}$  intersected with the positive definite cone, i.e.  $\Theta \in \mathbf{SS}_d^+ \triangleq \text{vec}^{-1}\mathbf{S} \cap \mathbb{S}_d^+$ .

Suppose that  $\nu$  is a Lévy measure on  $\mathbb{S}_d^+$  of the pure-jump subordinator  $R$ , such that it can be decomposed into

$$\nu(dX) = f(r, \Theta) dr \Gamma(d\Theta), \quad X \in \mathbb{S}_d^+,$$

where  $f$  is a continuous function and  $\Gamma$  is a spectral measure on  $\mathbf{SS}_d^+$  concentrated on a finite number of points  $\{\Theta_i\}_{1 \leq i \leq n}$ . Note in passing that any measure can be approximated by a measure concentrated on finitely many points. Since  $L$  is a pure-jump subordinator its characteristic function is given by

$$\begin{aligned} \phi_L(\Phi) &= t \int_{\mathbb{S}_d^+ \setminus \{0\}} \left( e^{i \text{tr}(\Phi X)} - 1 \right) \nu(dX), \\ &= t \int_{\mathbf{SS}_d^+} \int_0^\infty \left( e^{i r \text{tr}(\Phi \Theta)} - 1 \right) f(r, \Theta) dr \Gamma(d\Theta) \\ &= t \sum_{i=1}^n \Gamma(\Theta_i) \int_0^\infty \left( e^{i r \text{tr}(\Phi \Theta_i)} - 1 \right) f(r, \Theta_i) dr. \end{aligned}$$

One recognizes this as the characteristic function of a sum of independent random variables. This leads to the following simple algorithm to sample  $L$  according to its characteristic function:

- Find the finite set of points  $\{\Theta\}_{1 \leq i \leq n}$  where  $\Gamma$  is concentrated.
- Sample  $n$  independent random variables  $r_i$ 's with characteristic function

$$\phi_{r_i}(\text{tr}(\Phi \Theta_i)) = t \int_0^\infty \left( e^{i r \text{tr}(\Phi \Theta_i)} - 1 \right) f(r, \Theta_i) dr.$$

- Set  $L = \sum_{i=1}^n r_i \Theta_i$ .

To make this algorithm operationable, we must be able to sample the  $r$ 's, which we now discuss in particular cases.

#### 4.1.1 Compound Poisson matrix-valued subordinator

Matrix-valued compound process ( $mCP$ ) with only positive jumps is a multivariate compound Poisson process restricted to values in the symmetric positive definite cone. The characteristic function is given by

$$\phi_L(\Phi) = t\lambda \int_{\mathbb{S}_d^+} \left( e^{i \text{tr}(\Phi X)} - 1 \right) g(dX),$$

for  $L \sim mCP(\lambda)$ . Here,  $\lambda$  is the intensity and  $g$  is the jump-size distribution. Choosing the jump-size distribution such that it can be decomposed in  $g(dX) = \tilde{g}(dr)\Gamma(d\Theta)$  for a spectral measure  $\Gamma$  on  $\mathbf{SS}_d^+$ , concentrated on finitely many points, and a one-dimensional jump-size distribution  $\tilde{g}$ , it holds

$$\begin{aligned} \phi_L(\Phi) &= t\lambda \int_{\mathbf{SS}_d^+} \left( e^{i r \text{tr}(\Phi \Theta)} - 1 \right) \tilde{g}(dr)\Gamma(d\Theta), \\ &= t\lambda \sum_{i=1}^n \Gamma(\Theta_i) \int_0^\infty \left( e^{i r \text{tr}(\Phi \Theta_i)} - 1 \right) \tilde{g}(dr). \end{aligned}$$

Hence,  $r_i$ 's will follow a one-dimensional compound Poisson process with jump intensity  $\lambda\Gamma(\Theta_i)$  and jump-size distribution  $\tilde{g}$ . A random variable  $L \sim mCP(\lambda)$  is equally distributed as a linear combination of angle's  $\Theta_i$  and radius  $r_i$ , i.e.  $L \stackrel{d}{=} \sum r_i \Theta_i$ .

#### 4.1.2 Inverse Gaussian matrix-valued subordinator

By exponential tilting of matrix-valued  $\alpha$ -stable laws a multivariate extension of tempered stable laws can be made. The inverse Gaussian distribution is a special case of this class of functions. The polar decomposition of the Lévy measure  $\nu$  of a matrix-valued tempered  $\alpha/2$ -stable law is given by (see Barndorff-Nielsen and Pérez-Abreu [3])

$$\nu(dX) = \frac{e^{-r\text{tr}(\Delta\Theta)}}{r^{1+\alpha/2}} dr \Gamma(d\Theta),$$

where  $X \in \mathbb{S}_d^+$ ,  $r = \|X\| = \text{tr}(XX)^{1/2}$  and  $\Theta = X/r$ . In case  $\alpha = 1$  then  $\nu$  is a Lévy measure of a matrix extension of the inverse Gaussian distribution (mIG), where  $\Delta \in \mathbb{S}_d^+$  and  $\Gamma$ , a finite measure on  $\mathbb{S}_d^+$ , are parameters. As in the univariate case the inverse Gaussian process is a pure jump process, hence the characteristic function is given by

$$\phi_L(\Phi) = t \int_{\mathbb{S}_d^+} \int_0^\infty (e^{ir\text{tr}(\Phi\Theta)} - 1) e^{-r\text{tr}(\Delta\Theta)} \frac{dr}{r^{3/2}} \Gamma(d\Theta) + i\text{tr}(\Phi\mu_0),$$

for  $L \sim mIG(\Delta, \Gamma, \mu_0)$ , where  $\mu_0 \in \mathbb{S}_d^+$  is a parameter. Choosing  $\Gamma$  such that it is concentrated on finitely many point and decomposing  $\mu_0$  in an angle  $\Theta_0 \in \mathbb{S}_d^+$  and a radius  $r_0 \in \mathbb{R}$  leads to

$$\phi_L(\Phi) = t \sum_{i=1}^n \Gamma(\Theta_i) \int_0^\infty (e^{ir\text{tr}(\Phi\Theta_i)} - 1) e^{-r\text{tr}(\Delta\Theta_i)} \frac{dr}{r^{3/2}} + ir_0 \text{tr}(\Phi\Theta_0).$$

One can compare this with the characteristic function of an one-dimensional inverse Gaussian random variable  $y$ , for which the characteristic function is given by

$$\phi_y(\zeta) = i \frac{\delta}{\gamma} (2\mathcal{N}(\gamma) - 1) \zeta + \frac{\delta}{\sqrt{2\pi}} \int_0^\infty (e^{i\zeta x} - 1) e^{-1/2\gamma^2 x} \frac{dx}{x^{3/2}} \quad \zeta \in \mathbb{R}.$$

where  $\mathcal{N}$  denotes the cumulative normal distribution. Again one can recognize  $L$  as a matrix of linear combinations of a finite number of angles  $\Theta_i$ ,  $i = 1, \dots, d$  with as coefficients one-dimensional inverse Gaussian random variables  $r_i$ , distributed according to inverse Gaussian distribution  $IG(\delta_i, \gamma_i)$ , with  $\delta_i = \sqrt{2\pi} \Gamma(\Theta_i)$  and  $\gamma_i = 2 \sqrt{\text{tr}(\Delta\Theta_i)}$ . Moreover the drift parameter  $\mu_0$  of the multivariate inverse Gaussian distribution is by default chosen such that the drift term of the mIG distribution equals the drift term of  $\sum_i r_i \Theta_i$ .

In order to define a measure  $\Gamma$  on the symmetric positive definite cone we have to know what  $\Theta \in \mathbb{S}_d^+$  looks like. Below we will calculate this for the simple case that  $d = 2$ .

#### 4.2 Possible values of $\Theta \in \mathbb{S}_2^+$

When working in two dimensions the restriction of  $\Theta$  to the symmetric positive definite cone intersected with the unit sphere  $\mathbb{S}_2$  is restrictive. We will derive the shape of  $\Theta \in \mathbb{S}_2^+$ .

In the 2-dimensional case the following 3 properties hold for  $\Theta \in \mathbb{S}_2^+$ ;

1.  $\Theta$  is symmetric.
2.  $w^T \Theta w \geq 0 \quad \forall w \in \mathbb{R}^2 \Rightarrow \theta_{11} w_1^2 + \theta_{12} w_1 w_2 + \theta_{21} w_1 w_2 + \theta_{22} w_2^2 \geq 0 \Leftrightarrow \theta_{12} \stackrel{(1)}{=} \theta_{21} = \pm \sqrt{|\theta_{11}| |\theta_{22}|}$  or zero and  $\theta_{11}, \theta_{22} \geq 0$ .
3. If  $\theta_{12} = \theta_{21} \neq 0$  then

$$\begin{aligned} \|\Theta\| = 1 &\Rightarrow \theta_{11}^2 + \theta_{12}^2 + \theta_{21}^2 + \theta_{22}^2 = 1 \\ &\stackrel{(1)}{=} \theta_{11}^2 + 2\theta_{12}^2 + \theta_{22}^2 \stackrel{(2)}{=} \theta_{11}^2 + 2\theta_{11}\theta_{22} + \theta_{22}^2 \\ &= (\theta_{11} + \theta_{22})^2 \Rightarrow \theta_{11} + \theta_{22} = \pm 1 \end{aligned}$$

If  $\theta_{12} = \theta_{21} = 0$  then

$$\|\Theta\| = 1 \Rightarrow \theta_{11}^2 + \theta_{22}^2 = 1 \Rightarrow \theta_{22} = \sqrt{1 - \theta_{11}^2}$$

This means that  $\Theta$  must be of the form

$$\Theta = \begin{pmatrix} \theta_{11} & \pm\sqrt{\theta_{11}(1-\theta_{11})} \\ \pm\sqrt{\theta_{11}(1-\theta_{11})} & 1-\theta_{11} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \theta_{11} & 0 \\ 0 & \sqrt{1-\theta_{11}^2} \end{pmatrix}.$$

Similar calculations can be done in higher dimensions.

### 4.3 Empirical example

Consider a 2-dimensional example of our model i.e.

$$S(t) = \begin{pmatrix} S_1(t) \\ S_2(t) \end{pmatrix} = \begin{pmatrix} \Lambda_1(t) \\ \Lambda_2(t) \end{pmatrix} \cdot \exp \left( \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} + \sum_{i=1}^m \begin{pmatrix} Y_1^i(t) \\ Y_2^i(t) \end{pmatrix} \right),$$

and

$$\begin{aligned} dX(t) &= \begin{pmatrix} dX_1(t) \\ dX_2(t) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_{11}(t) & \sigma_{12}(t) \\ \sigma_{21}(t) & \sigma_{22}(t) \end{pmatrix}^{1/2} \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \\ &= AX(t)dt + \Sigma(t)^{1/2}dW(t), \end{aligned}$$

where  $\Sigma(t)$  is a weighted sum of matrix-valued Orstein-Uhlenbeck processes  $Z^j(t)$ , i.e.  $\Sigma(t) = \sum_j \omega_j Z^j(t)$  with

$$\begin{aligned} dZ^j(t) &= d \begin{pmatrix} Z_{11}^j(t) & Z_{12}^j(t) \\ Z_{21}^j(t) & Z_{22}^j(t) \end{pmatrix} = \mathbf{C}_j \begin{pmatrix} Z_{11}^j(t) & Z_{12}^j(t) \\ Z_{21}^j(t) & Z_{22}^j(t) \end{pmatrix} dt + d \begin{pmatrix} L_{11}^j(t) & L_{12}^j(t) \\ L_{21}^j(t) & L_{22}^j(t) \end{pmatrix} \\ &= \mathbf{C}^j Z^j(t) + d\tilde{L}^j(t). \end{aligned}$$

Moreover,

$$\begin{aligned} dY^i(t) &= \begin{pmatrix} dY_1^i(t) \\ dY_2^i(t) \end{pmatrix} = \left( \begin{pmatrix} \mu_1^i \\ \mu_2^i \end{pmatrix} + \begin{pmatrix} b_{11}^i & b_{12}^i \\ b_{21}^i & b_{22}^i \end{pmatrix} \begin{pmatrix} Y_1^i(t) \\ Y_2^i(t) \end{pmatrix} \right) dt + \begin{pmatrix} \eta_{11}^i & \eta_{12}^i \\ \eta_{21}^i & \eta_{22}^i \end{pmatrix} \begin{pmatrix} dL_{11}^i(t) \\ dL_{22}^i(t) \end{pmatrix} \\ &= \{\mu^i + B^i Y^i(t)\}dt + \eta^i dL^i(t). \end{aligned}$$

In particular consider the case of one spike process and two volatility processes with the following parameters,

$$\begin{aligned} A &= \begin{pmatrix} -1.4 & -0.2 \\ -0.3 & -1.6 \end{pmatrix} & B &= \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} & \eta &= \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \\ C_1 &= \begin{pmatrix} -0.4 & 0.3 \\ 0.2 & -0.4 \end{pmatrix} & C_2 &= \begin{pmatrix} -0.045 & 0.03 \\ 0.03 & -0.065 \end{pmatrix} \end{aligned}$$

Moreover

$$\begin{pmatrix} \Lambda_1(t) \\ \Lambda_2(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The process  $\tilde{L}^1$  is a matrix-valued Poisson process with intensity  $\lambda\Gamma(\Theta_i) = 3/100$  and constant jump-size 1.7 for the radius  $r_i$ . Moreover,  $\tilde{L}^2$  are matrix valued inverse Gaussian processes with parameters  $\Delta = \begin{pmatrix} 50 & 45 \\ 45 & 50 \end{pmatrix}$  and  $\Gamma(\Theta_i) = \frac{1}{12n\sqrt{2\pi}}$  for all  $1 \leq i \leq n$ . As a finite set of angles we let  $\theta_{11}$  take all values between zero and one with a stepsize 0.1. For each value of  $\theta_{11}$  we make the 3 possible matrixes. So in total we get  $n = 3/0.1 = 30$  angle matrices  $\Theta$ . In Figure 1(a) the spot price series resulting from

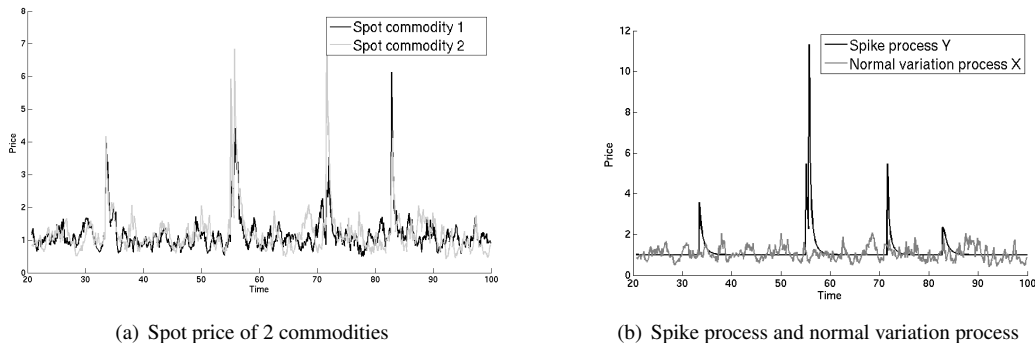


Figure 1: Spot price

our 2-dimensional example is shown, where we have used an Euler scheme to discretize the dynamics in time. One clearly can see the dependency between the two spot prices. In Figure 1(b) the contribution of the spike component  $Y$  and the normal variation component  $X$  are visualized. The spike component contributes to rare spikes and  $X$  gives the normal fluctuation of the spot price. We note that the chosen parameter specifications here are inspired by the estimated stochastic volatility model in Vos [42], where the BNS stochastic volatility model was estimated to stock price data observed on the Dutch stock exchange.

#### 4.4 Co-integration

Co-integration seems to be a popular tool in discrete-time financial multivariate modelling. Statistical evidence of co-integration has been reported for interest-rates by Engle and Granger [18] and for foreign exchange rates by Bailie and Bollerslev [1]. Also in energy markets there is interest to try co-integration techniques for cross-commodity modelling.

A multi-variate discrete-time series  $X$  is said to be co-integrated if each of the time-series taken individually is integrated of order 1, that is, non-stationary with a unit root, whereas some linear combination of the time series,  $b'X$ , is stationary for a vector  $b$ . The implication of being co-integrated is stronger than simply being correlated. Correlated time series can drive apart, while co-integrated time series capture a restriction about a value of a linear combination of the time series which can not alter too much due to its stationarity. Several extensions to continuous time have been proposed in the literature. Comte [15] defines co-integration in terms of the mean-square derivatives. It is, however, not possible directly to invoke this definition in the case of an Ornstein-Uhlenbeck process. This would lead to a nilpotent Ornstein-Uhlenbeck matrix in which case no solution exists. Duan and Pliska [17] take the weak limit of a co-integrated discrete time system. This is a tight definition since the weak limit as introduced by Nelson [30] is only defined for Brownian motions. We will follow the approach of Kessler and Rahbeck [27], who define co-integration in continuous time for multi-variate Ornstein-Uhlenbeck processes (OU-process). They consider the case of Brownian-driven OU-processes, however, their method is general and can be applied to any multi-variate OU-process.

Consider the following  $d$ -dimensional multi-variate OU-process.

$$dX(t) = \{AX(t) + \mu\}dt + \Sigma(t)dW(t) \quad (4.1)$$

According to Kessler and Rahbeck [28], a multi-variate Gaussian OU-process  $X$  is co-integrated if the OU matrix  $A \in M_d(\mathbb{R})$  is of reduced rank  $r < d$  and the last  $r - d$  entries of  $U^{-1}\mu$  are zero. Where  $U$  is the basis of eigenvalues of  $A$ .  $A$  has exactly  $r$  eigenvalues not equal to zero and  $d - r$  zero eigenvalues. There exist exactly  $r$  linearly independent combinations,  $Z_i(t) = b_i'X(t)$   $i = 1, \dots, r$ , of the elements of  $X$  which form a one dimensional OU-process. In the full rank case any component of a  $d$ -dimensional OU-process can be written as a linear combination of  $d$  linear independent OU-processes. In the reduced rank case the components are linear combinations of  $r$  linear independent OU-processes,  $d$  Brownian motions and the initial value of  $X$ . It is important to note that since we work with a reduced rank matrix  $A$ , it is not

invertible. Hence, many of the results deduced in Section 3 do not apply in the case of co-integrated spot prices.

To this end, introduce the following notation  $I_m 0_n \in M_{m \times (m+n)}(\mathbb{R})$  formed by pasting together an  $m \times m$  identity matrix followed by an  $m \times n$  zero matrix, i.e.,

$$I_m 0_n = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & & & 0 & \ddots & \\ \vdots & & \ddots & & \vdots & & \end{pmatrix}.$$

Similarly,  $0_m I_n \in M_{m \times (m+n)}(\mathbb{R})$  is defined by pasting together an  $m \times m$  zero matrix with an  $m \times n$  identity matrix. We define the following two matrices;

$$A = U(0_r I_{d-r})^T 0_r I_{d-r} U^{-1} \quad (4.2)$$

$$B = U(I_r 0_{d-r})^T. \quad (4.3)$$

The following is stated in Kessler and Rahbeck [27]. Since no proof is explicitly given in the paper we provide one here.

**Theorem 4.1** (Kessler and Rahbeck [27], Theorem 1). *Suppose  $X \in \mathbb{R}^d$  is a multivariate Gaussian OU-process as in (4.1) which is co-integrated in the sense  $A \in M_d(\mathbb{R})$  has reduced rank  $r < d$  and is diagonalizable and the last  $r - d$  entries of  $U^{-1}\mu$  are zero, with  $U$  being the basis of eigenvalues of  $A$ . Then, solution of  $X$  is given by*

$$X(t) = \mathcal{A}(\Sigma W(t) + X_0) + \mathcal{B}Z(t),$$

where  $\mathcal{A}$  and  $\mathcal{B}$  are defined in (4.2) and (4.3), resp. The  $r$ -dimensional process  $Z(t)$  has one-dimensional OU-processes as its entries.

*Proof.* Since  $A$  is diagonalizable we have  $A = U\Lambda U^{-1}$ , where  $\Lambda$  is a diagonal matrix with the eigenvalues of  $A$  on the diagonal. Since  $A$  has reduced rank  $r < d$ ,  $r - d$  eigenvalues will be zero. Organize the diagonalization such that the last  $r - d$  rows of  $\Lambda$  are zero. Define the process  $Z(t) = I_r 0_{d-r} U^{-1} X(t)$ . Note that  $I_r 0_{d-r} U^{-1}$  will exactly consist of the  $r$  independent co-integration relations. It holds

$$dX(t) = AX(t) + \mu dt + \Sigma dW(t) = \{U\Lambda U^{-1}X(t) + \mu\}dt + \Sigma dW(t).$$

Multiply both sides with  $U^{-1}$  to get,

$$dU^{-1}X(t) = \{\Lambda U^{-1}X(t) + U^{-1}\mu\}dt + U^{-1}\Sigma dW(t).$$

Now since the last  $d - r$  rows of  $\Lambda$  are zero it holds

$$\begin{aligned} dZ(t) &\triangleq I_r 0_{d-r} U^{-1} dX(t) = I_r 0_{d-r} (\{\Lambda U^{-1}X(t) + U^{-1}\mu\}dt + U^{-1}\Sigma dW(t)), \\ &= \{\Lambda_r Z(t) + \mu'\}dt + I_r 0_{d-r} U^{-1}\Sigma dW(t), \end{aligned}$$

where  $\Lambda_r$  is the  $r \times r$  diagonal matrix with the non-zero eigenvalues on the diagonal and  $\mu' = I_r 0_{d-r} U^{-1}\mu$ .  $Z$  is an  $r$ -dimensional process of independent one-dimensional OU-processes. Again since the last  $d - r$  eigenvalues of  $A$  are zero and  $0_r I_{d-r} U^{-1}\mu = 0$ ,

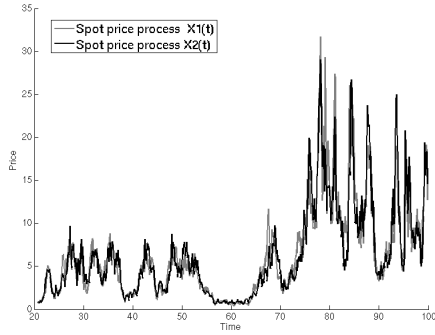
$$0_r I_{d-r} U^{-1} dX(t) = 0_r I_{d-r} U^{-1} \Sigma dW(t),$$

from which follows

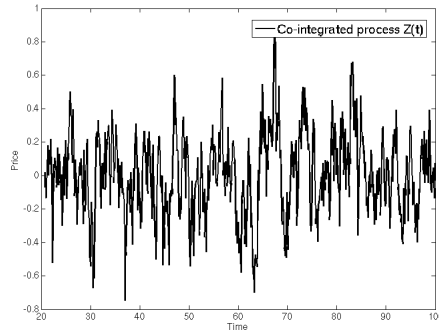
$$0_r I_{d-r} U^{-1} X(t) = 0_r I_{d-r} U^{-1} (\Sigma dW(t) + X(0)).$$

Combining all the results

$$U^{-1} dX(t) = (I_r 0_{d-r})^T I_r 0_{d-r} U^{-1} dX(t) + (0_r I_{d-r})^T 0_r I_{d-r} U^{-1} dX(t)$$



(a) Spot price of two co-integrated spot price series



(b) Dynamics of co-integrated process  $Z(t)$ .

Figure 2: Co-integration

$$= (I_r 0_{d-r})^T dZ(t) + (0_r I_{d-r})^T 0_r I_{d-r} U^{-1} dX(t).$$

Since  $U$  is a basis of  $\mathbb{R}^d$ ,  $Z$  and  $0_r I_{d-r} U^{-1} X(t)$  are independent. Hence we can solve them separately. Solving the SDE's and multiplying both sides with  $U$  yields

$$X(t) = U(I_r 0_{d-r})^T Z(t) + U(0_r I_{d-r})^T 0_r I_{d-r} U^{-1} (\Sigma dW(t) + X(0)), \quad (4.4)$$

$$\triangleq \mathcal{B}Z(t) + \mathcal{A}(\Sigma dW(t) + X(0)). \quad (4.5)$$

Hence, the result follows.  $\square$

The increments  $dX(t)$  are not stationary nor the linear combination  $b'X(t)$  which describe the co-integration relation. The latter is described by a one-dimensional OU-process, so they exhibit in the limit a stationary solution. This is sufficient to describe the explicit dependence of certain linear combinations of components of  $X$ . The total non-stationary part is given by a Brownian motion which is not stationary in the limit. This will give a Brownian behavior in the components of  $X$  viewed independently. Noteworthy is that the whole sample path of  $X$  is dependent on its initial value.

## 4.5 Example

We present an example to illustrate co-integration for the processes discussed in this section. Consider a 2-dimensional example of our model with no spike processes  $Y_i$  (*i.e.*  $m = 0$ ). Moreover the Ornstein-Uhlenbeck matrix  $A$  is given by,

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix},$$

and the volatility is constant

$$\Sigma = \begin{pmatrix} 0.03 & 0.015 \\ 0.015 & 0.03 \end{pmatrix}.$$

The rest of the parameters are as in Example 4.3. The co-integration vector is  $b = (\sqrt{0.5} \quad -\sqrt{0.5}) \in \mathbb{R}^2$ . Note that the co-integrated process  $Z(t) = b'X(t)$  is given by the one-dimensional Ornstein-Uhlenbeck process,

$$\begin{aligned} dZ(t) &= \sqrt{0.5}d(X_1(t) - X_2(t)), \\ &= -\sqrt{2}(X_1(t) - X_2(t))dt - \sqrt{0.5}(\sigma_{11}(t) - \sigma_{21}(t))dW_1(t) - \sqrt{0.5}(\sigma_{12}(t) - \sigma_{22}(t))dW_2(t), \\ &= -\sqrt{2}Z(t)dt + b'\Sigma(t)dW(t), \end{aligned}$$

which will mean-revert to zero with exponential speed  $e^{-\sqrt{2}t}$ . This ensures that  $X_1$  and  $X_2$  can never drive apart although the marginals have Brownian components in the dynamics. The solution of  $X(t)$  is given by

$$X(t) = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} (\Sigma W(t) + X(0)) + \begin{pmatrix} \sqrt{0.5} \\ -\sqrt{0.5} \end{pmatrix} Z(t).$$

So  $X(t)$  is a linear combination of the initial value  $X(0)$ , two Brownian motions  $W_1$  and  $W_2$ , and an OU-process  $Z$  driven by the same Brownian motions. The initial value of  $X$  will influence the whole price curve of the process.

In Figure 2(a) we depict a sample path of the co-integrated two-dimensional spot process, as well as the co-integration process  $Z$  (Figure 2(b)). Clearly visible is that the process  $Z$  only fluctuates around zero, whereas the spot prices are highly non-stationary and exhibit large variations even though the volatility is constant.

One can allow for additional spike processes  $Y$  in the dynamics, however the Ornstein-Uhlenbeck matrix  $B$  can not be of reduced rank. Since these spikes will be rare events, they will have a negligible effect on the co-integration assumption.

## 5 Forward pricing

In this Section we derive the forward price dynamics based on our multivariate spot price model. In commodity markets, forward contracts are commonly traded on exchanges, including power, gas, oil, coal, etc. It is important for risk management and derivatives pricing to understand how the forward prices depend between various products.

Appealing to general arbitrage theory, we define the forward price  $F(t, T)$  at time  $t$  for contracts maturing at time  $\tau$  by (see e.g. Duffie [20])

$$F(t, \tau) = \mathbb{E}_Q [S(\tau) | \mathcal{F}_t], \quad (5.1)$$

where  $Q$  is a risk-neutral probability. This definition is valid as long as  $S(\tau) \in L^1(Q)$ . We will below give conditions ensuring the well-definedness of the forward price. In particular, since the spot price is an adapted process, we obtain the well-known convergence of spot and forward prices at maturity, *i.e.*,

$$F(\tau, \tau) = S(\tau).$$

It is worth noticing that in some energy markets the forward contracts deliver the underlying commodity over a period rather than at a fixed maturity time  $\tau$ . This includes gas and electricity, but also more exotic markets like temperature. In these markets, the forward prices can be represented as some functional of  $F(t, \tau)$ , usually the average of  $F(t, \tau)$  over  $\tau$ , taken over the delivery period of the forward contract. We will not consider this here right now, however the calculations can be easily adjusted to take this into account.

The stochastic volatility model we are discussing gives rise to an incomplete market, and hence the existence of an infinite number of equivalent martingale measures  $Q$  that can be used for pricing. Moreover, in energy markets, the underlying spot is not tradeable in the classical sense, due to high storage costs, illiquidity and other frictions like transportation for delivery. In the extreme case of electricity, it is impossible to trade the underlying spot by the very nature of the commodity. Hence, the classical buy-and-hold hedging argument to pin down a price the forward fails. But even more, *all* equivalent measures  $Q \sim P$  may be chosen as pricing measures since the underlying spot is not directly tradeable, and hence we do not require the martingale property under  $Q$  for discounted prices. We refer to Benth *et al.* [10] for more on this.

A convenient way to define a parametric class of risk-neutral probabilities is the Esscher transform. This is a particular choice of a measure transform convenient to use for Lévy-based models (see Benth *et al.* [10] for applications in energy markets), parametrizing the risk-neutral probabilities  $Q$ . For a vector  $\theta_0 \in \mathbb{R}^d$ , introduce the processes

$$\mathcal{V}_0(t) = \exp \left( - \int_0^t \Sigma^{-1/2}(s) \theta_0 dW(s) - \frac{1}{2} (\theta_0)^T \int_0^t \Sigma^{-1}(s) \theta_0 ds \right), \quad (5.2)$$

and for  $\theta_1, \dots, \theta_m \in \mathbb{S}_d^+$

$$\mathcal{V}_j(t) = \exp \left( -i \text{tr}(\theta_j \tilde{L}_j(t)) - \phi_{\tilde{L}}^j(-i\theta_j)t \right), \quad (5.3)$$

for  $j = 1, \dots, n$  and  $t \leq \tilde{T}$ . Here  $\tilde{T}$  is a finite horizon of the market for which all delivery times  $\tau$  of interest are included. We assume that  $\tilde{L}$  has exponential moments up to a certain order, and we implicitly consider only  $\theta^1, \dots, \theta^m$  for which  $\phi_{\tilde{L}}^j(-i\theta^j)$ ,  $j = 1, \dots, m$  are well-defined. Note that  $\mathcal{V}_j(t)$  are martingales for  $j = 1, \dots, m$ . Moreover, from Girsanov's theorem, one can conclude that also  $\mathcal{V}_0(t)$  is a martingale. Thus, the process

$$\mathcal{V}(t) = \mathcal{V}_0(t) \times \mathcal{V}_1(t) \times \dots \times \mathcal{V}_n(t), \quad (5.4)$$

becomes a martingale for  $t \leq \tilde{T}$ . This process is the density process of a probability  $Q$  equivalent with  $P$ , that is,

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_t} = \mathcal{V}(t).$$

From Girsanov's Theorem we find that

$$d\widehat{W}(t) = dW(t) - \Sigma^{-1/2}(t)\theta_0 dt, \quad (5.5)$$

is an  $\mathbb{R}^d$ -valued Brownian motion with respect to  $Q$  on  $t \in [0, \tilde{T}]$ . Furthermore,  $\tilde{L}_j(t)$  is still a subordinator with respect to  $Q$ , with characteristics stated in the following Lemma:

**Lemma 5.1.** *Assume  $\tilde{L}$  exhibits exponential moments up to a certain order and  $\theta_j$  is such that  $\phi_{\tilde{L}}^j(-i(V + \theta_j))$  and  $\phi_{\tilde{L}}^j(-i\theta_j)$  are well-defined for all  $j = 1, \dots, m$ , then the cumulant of  $\tilde{L}_j$  under  $Q$  is given as*

$$\phi_j^Q(-iV) = \phi_{\tilde{L}}^j(-i(V + \theta_j)) - \phi_{\tilde{L}}^j(-i\theta_j), \quad \text{for } V \in \mathbb{S}_d.$$

Moreover,  $\tilde{L}_j^Q$  is a subordinator with Lévy measure  $\exp(\text{tr}(\theta_j x))\nu_{\tilde{L}}^j(dx)$ .

*Proof.* The proof of Lemma 5.1 is given in the Appendix in section A.4.  $\square$

We note that the Esscher transform will re-scale the jumps by an exponential factor. Consider entry  $\theta_{ik}^j$  of the matrix  $\theta_j \in \mathbb{S}_d^+$ . If  $\theta_{ik}^j > 0$ , the larger jumps of entry  $\tilde{L}_{ik}$  of the process  $\tilde{L} \in \mathbb{S}_d^+$  become more pronounced under  $Q$  than under  $P$ , whereas  $\theta_{ik}^j < 0$  creates a lighter tail of  $\nu_{\tilde{L}}^j(dz)$  and therefore less risk is put into large jumps. The parameters  $\theta_0$  and  $\theta_j$ ,  $j = 1, \dots, n$  are frequently called *the market prices of risk*.

The dynamics of  $X(t)$  under  $Q$  is given by

$$\begin{aligned} dX(t) &= AX(t) + \Sigma^{1/2}(t) \left( d\widehat{W}(t) + \Sigma^{-1/2}(t)\theta_0 dt \right) \\ &= (\theta_0 + AX(t)) dt + \Sigma^{1/2}(t) d\widehat{W}(t). \end{aligned} \quad (5.6)$$

Thus, under  $Q$ , the mean-reversion level is  $\theta_0$  rather than 0 as it is under  $P$ . The dynamics of  $Y_i$  and  $Z_j$  are changed in a similar fashion. We have that for  $i = 1, \dots, m$

$$\begin{aligned} dY_i(t) &= (\mu_i + B_i Y_i(t)) dt + \eta_i dL_i(t) \\ &= (\mu_i + \eta_i \mathbb{E}_Q[L(1)] + B_i Y_i(t)) dt + \eta_i d\tilde{L}_i^Q(t), \end{aligned} \quad (5.7)$$

where  $d\tilde{L}_i^Q(t) \triangleq d\tilde{L}_i(t) - \mathbb{E}_Q[\tilde{L}(1)] dt$  is a  $Q$ -martingale. Hence, the process  $Y_i$  varies around the level  $\mu_i + \eta_i \mathbb{E}_Q[L(1)]$  under  $Q$ , whereas the level is  $\mu_i + \eta_i \mathbb{E}_P[L(1)]$  under  $P$ . Similar considerations hold for the volatility processes  $Z_j$ . We observe that in the case of positive premium  $\{\theta^0\}_k > 0$ , the mean-level for the  $k$ -th entry of process  $X(t)$  is moved upwards compared to the  $P$ -dynamics. The same holds for the processes  $Y_i$  and  $Z_j$  when there are entries of  $\theta^i$  that are positive, with opposite conclusions for negative entries of  $\theta^i$  (i.e., negative market prices of risk).



Define the matrix  $\mathcal{D}_{k,j}$  such that  $\text{tr} \left( \mathcal{D}_{k,j}(t)X \right)$  is equal to the  $k$ -th component of the diagonal from  $\mathcal{E}_j(t)X$  for any  $X \in \mathbb{S}_d^+$ .  $\mathcal{D}_{k,j}(t)$  is given by

$$\omega_j \text{vec}^{-1} \left[ e_{(k-1)d+k}^T \left( (C_j - A) \otimes I + I \otimes (C_j - A) \right)^{-1} \left( e^{(C_j \otimes I + I \otimes C_j)t} - e^{(A \otimes I + I \otimes A)t} \right) \right], \quad (5.8)$$

where  $e_j$  is the unit vector with the one on the  $j$ -th place. Moreover define the linear operator  $\mathcal{E}_k$  such that the  $k$ -th row of the matrix where it is working on is mapped to the diagonal. It can be represented as

$$\mathcal{E}_k : X \mapsto X^T \Lambda_{i;1,\dots,d} \sum_{i=1}^d \Lambda_{ii} X^T \Lambda_{1i}, \quad (5.9)$$

where  $\Lambda_{ij}$  is a zero matrix with on the  $i$ -th row and  $j$ -th column an one.  $\Lambda_{i;1,\dots,d}$  has ones on the whole row  $i$ .

We are now in the position to state the forward price.

**Proposition 5.2.** *Suppose  $\theta^j$  is such that the conditions of Lemma 5.1 are satisfied. Then the forward price  $F(t, \tau)$  for  $t \leq \tau \leq T$  is given by*

$$F(t, \tau) = \Lambda(\tau) \cdot \exp \left( e^{A(\tau-t)} X(t) + \sum_{i=1}^m e^{B_i(\tau-t)} Y_i(t) + A^{-1} (I - e^{A(\tau-t)}) \theta_0 \right. \\ \left. + \sum_{i=1}^m B_i^{-1} (I - e^{B_i(\tau-t)}) \mu_i \right) \exp \left( \sum_{i=1}^n \Psi^i(t, \tau, Z(t)) \right) \Theta^{\text{vol}}(\tau - t) \Theta^{\text{spike}}(\tau - t), \quad (5.10)$$

where  $\Psi^j(t, \tau, X) \in \mathbb{R}^d$  with entries

$$\Psi_k^j(t, \tau, X) = \frac{1}{2} \text{tr} \left( \mathcal{D}_{k,j}(\tau - t) X \right),$$

$\Theta^{\text{vol}}$  is a vector in  $\mathbb{R}^d$  with entries

$$\ln \Theta_k^{\text{vol}}(v) \triangleq - \sum_{i=1}^m \int_0^v \phi_L^i \left( \frac{-i}{2} \mathcal{D}_{k,i}^T(s) I_d - i \mathcal{E}_k \left[ e^{B_i(s)} \eta_i \right] I_d - i \theta^i \right) - \phi_L^i \left( \frac{-i}{2} \mathcal{D}_{k,i}^T(s) I_d - i \theta^i \right) ds,$$

and  $\Theta^{\text{spike}} \in \mathbb{R}^d$  with entries,

$$\ln \Theta_k^{\text{spike}}(v) \triangleq - \sum_{j=1}^n \int_0^v \phi_L^j \left( \frac{-i}{2} \mathcal{D}_{k,j}^T(s) I_d - i \theta_j \right) - \phi_L^j(-i \theta_j) ds.$$

*Proof.* The proof of Prop. 5.2 is given in the Appendix in section A.5. □

We have distinguished the contributions to the market price of risk from spikes and volatility in the terms  $\Theta^{\text{spikes}}$  and  $\Theta^{\text{vol}}$ , resp. However, they can obviously be collected and simplified into one term, which should be done in applications of the price representation.

Since  $C_j$  and  $A_i$  have eigenvalues with negative real-parts one can see from the representation of  $\mathcal{D}$  in (5.8) that  $\mathcal{D}_k(\tau - t)$  converges to zero as  $t \rightarrow \tau$ . This reconfirms that forward price coincides with the spot at maturity.

We state the dynamics of the forward price.

**Proposition 5.3.** *Suppose the conditions in Prop. 5.2 holds. Then the dynamics of  $F_k(t, \tau)$  of commodity  $k$  with respect to  $Q$  is*

$$\frac{dF_k(t, \tau)}{F_k(t-, \tau)} = e^{A(\tau-t)} \Sigma^{1/2}(t) dW(t) \times e_k$$

$$\begin{aligned}
& + \sum_{i=1}^m \int_{\mathbb{S}_d^+ \setminus \{0\}} \left\{ e^{\text{tr}(\mathcal{E}_k [e^{B_i(\tau-t)} \eta_i] z + \frac{1}{2} \mathcal{D}_{k,i}(\tau-t) z) - 1} \right\} \tilde{N}_i(dt, dz) \\
& + \sum_{j=m+1}^n \int_{\mathbb{S}_d^+ \setminus \{0\}} \left\{ e^{\text{tr}(\frac{1}{2} \mathcal{D}_{k,i}(\tau-t) z) - 1} \right\} \tilde{N}_j(dt, dz).
\end{aligned}$$

Here,  $\tilde{N}_j(dt, dz) = N_j(dt, dz) - \exp(\text{tr}(\theta^j z)) \nu_{\tilde{L}}^j(dz) dt$  and  $N_j$  is the Poisson random measure of  $\tilde{L}_j$ ,  $j = 1, \dots, n$ .

*Proof.* This follows by the one-dimensional Itô Formula for jump processes (see e.g. Shiryayev [39]), used on each entry of the vector  $F(t, \tau)$  separately. The utilization of this can be considerably simplified by the knowledge that  $F$  is a  $Q$ -martingale process, and therefore only terms with respect to  $W$  and  $\tilde{N}_j$  matter.  $\square$

We see that there is a so-called Samuelson effect in the  $dW$  term: the stochastic volatility is multiplied with exponentially decaying functions  $f_i$   $i = 1, \dots, d$ , and when  $t \rightarrow \tau$ , the volatility of these terms converge to the spot volatility  $\Sigma(t)$ . Note further the explicit appearance in the forward price of the jump processes driving the stochastic volatility. Even in the case when the spot price dynamics have continuous paths (the Schwartz model,  $m = 0$ ) the jumps are explicitly present in forward dynamics. Remark that for the SV Schwartz model ( $m = 0$ ), the forward price still have jumps explicitly present in its dynamics. Thus, even though the spot price will have continuous sample paths, the forward price will jump as a reaction to jumps in the volatility. This is a distinct feature coming from the choice of modelling the volatility using jump processes instead of Brownian motion, which in the latter case would yield a continuous-path forward dynamics.

In the next Proposition we show that the forward price will behave like the seasonal function in the long end of the market.

**Proposition 5.4.** *Suppose  $\theta^j$  is such that the assumptions of Lemma 5.1 are satisfied. We have that*

$$\lim_{\tau \rightarrow \infty} F(t, \tau) / \Lambda(\tau) = \exp \left( A^{-1} \theta_0 + \sum_{i=1}^m B_i^{-1} \mu_i \right) \Theta^{vol}(\infty) \Theta^{spike}(\infty),$$

where the division on the left-hand side is pointwise.

*Proof.* Using representation (5.8) one can see that  $\mathcal{D}_j(\tau-t)X$  tends to zero when  $\tau \rightarrow \infty$  for all  $X \in \mathbb{S}_d^+$ . So,  $\lim_{\tau \rightarrow \infty} \sum_{i=1}^n \Psi^i(t, \tau, Z(t)) = 0$ . From Prop. 5.2 we see that

$$\lim_{\tau \rightarrow \infty} \left( \ln \left( \frac{F(t, \tau)}{\Lambda(\tau)} \right) - \ln \Theta^{vol}(\tau-t) - \ln \Theta^{spike}(\tau-t) - \sum_{j=1}^n \Psi^j(t, \tau, Z_j(t)) \right) = A^{-1} \theta_0 + \sum_{i=1}^m B_i^{-1} \mu_i.$$

It remains to show that the limits  $\lim_{T \rightarrow \infty} \Theta^{spike, vol}(\tau-t) \triangleq \Theta^{spike, vol}(\infty)$  exist. However, based on the same arguments as in Prop. 3.3 and Prop. A.4, the characteristic functions of  $\int_0^\infty \mathcal{E}(s) d\tilde{L}(s)$  and  $\int_0^\infty e^{B_i(s)} \eta_i dL_i(u)$  exist under  $Q$ , and so also the characteristic functions of the sum of these random variables. Due to our assumption on  $\theta^j$ ,  $j = 1, \dots, n$ , also the cumulants exist.  $\Theta^{vol, spike}(\infty)$  is a combination of these cumulants, hence,  $\Theta^{vol, spike}(\infty)$  exist.  $\square$

Hence, from the Proposition above, in the long end of the market the forward prices are basically equal to the seasonality function adjusted by the long-term means of  $Y_i$  and  $Z_j$  and the market prices of risk, that is

$$F(t, \tau) \sim \text{const.} \cdot \Lambda(\tau).$$

As a result of mean reversion of the spot prices, the forward prices are not reacting to changes in the spot in the long end but only following the mean adjusted by the market prices of risk.

Note that the forward price dynamics is a regression on the spot price, leverage terms and the volatility processes. Introducing the shorthand notation  $\Theta(t, \tau) \in \mathbb{R}^d$  given by

$$\begin{aligned} \ln \Theta(t, \tau) \triangleq & \ln \Theta^{vol}(\tau - t) + \ln \Theta^{spike}(\tau - t) + \ln \Lambda(\tau) - e^{A(\tau-t)} \ln \Lambda(t) \\ & + A^{-1}(I - e^{A(\tau-t)})\theta_0 + \sum_{i=1}^m B_i^{-1}(I - e^{B_i(\tau-t)})\mu_i. \end{aligned} \quad (5.11)$$

Then, from Prop. 5.2,

$$\begin{aligned} \ln F(t, \tau) &= \ln \Theta(t, \tau) + e^{A(\tau-t)} X(t) + \sum_{i=1}^m e^{B_i(\tau-t)} Y_i(t) + \sum_{j=1}^n \Psi^j(t, \tau, Z_j(t)) \\ &= \ln \Theta(t, \tau) + e^{A(\tau-t)} \ln S(t) + \sum_{i=1}^m \left( e^{B_i(\tau-t)} - e^{A(\tau-t)} \right) Y_i(t) + \sum_{j=1}^n \Psi^j(t, \tau, Z_j(t)). \end{aligned} \quad (5.12)$$

Here,  $\Theta$  is a level adjustment function. The impact of the various factors on the forward price  $F(t, \tau)$  goes through the matrix exponentials. In fact, the forward price of one commodity depends on the normal variation processes  $X$ , spike processes  $Y$  and volatility processes  $Z$  of all the commodities modelled. Hence, for example, if one of the commodities has a spike, then the forward prices of all the other commodities will be influenced. As noted in Benth *et al.* [12], the mean-reverting structure represented by a matrix exponential has a richer structure than in the one-dimensional case, and we may include hump structures in the forward curve. We discuss the potential shapes of  $\tau \mapsto F(t, \tau)$  in more detail.

Since  $A \in GL_d(\mathbb{R})$ , it is diagonalizable. So it holds that  $e^{A(\tau-t)} = U e^{\Lambda(\tau-t)} U^{-1}$ , where  $U$  is a basis of eigenvectors and  $\Lambda$  is matrix with on the diagonal the eigenvalues of  $A$  (see *e.g.* Horn and Johnson [24]). Hence, an entry of the vector  $e^{A(\tau-t)} X(t)$  can be represented as

$$\sum_{i=1}^d a_{1i} e^{\lambda_i(\tau-t)} X_1(t) + \sum_{i=1}^d a_{2i} e^{\lambda_i(\tau-t)} X_2(t) + \dots + \sum_{i=1}^d a_{di} e^{\lambda_i(\tau-t)} X_d(t),$$

for some constants  $a_{ij} \in \mathbb{R}$  and eigenvalues  $\lambda_i, i, j = 1, \dots, d$ . Consider first the Schwartz model with constant volatility, *i.e.* no contribution of the processes  $Y$  and  $Z$  in the forward price. If  $X$  is positive in all its components,  $\lambda_i$  is real and  $a_{ij} \in \mathbb{R}^+$  for all  $i, j = 1, \dots, d$ , then the forward is in backwardation since the eigenvalues have negative real-parts. The opposite conclusion (*i.e.* forward prices in contango) can be taken when  $X$  is negative in all its components. A more realistic situation with this model is the case where there are humps in the forward curve and where the forward is changing between backwardation and contango over time. This behavior has been observed for real market prices. For example, on page 216 in Geman [22] the forward curve of WTI oil is plotted together with the spot price. The shape of the curve varies over time from contango to backwardation, including positive humps in the short end. When the constants  $a_{ji}$  for fixed  $j$  are not all of the same sign and the entries of  $X$  have all a positive sign, then an entry of  $e^{A(\tau-t)} X(t)$  is given by a linear combination of increasing and decreasing exponentials who rise and decay at different speeds. Due to this the forward may alternate between backwardation and contango and humps may appear (see figure 3 (b)). Another possibility is the case of complex eigenvalues. This leads to an oscillating structures in the forward curve. So a change upward in the  $i$ -th component of  $X$  may cause a rise or fall of the forward depending on the time to maturity (see figure 3 (a)). A similar analysis can be done for the spike process  $Y$ . However, since  $Y$  is a pure jump process it will contribute to sudden changes in the forward curve. These humps may be upward or downward pointing depending on the time to maturity. The jumps caused by the spike process  $Y$  may be averaged out by jumps in the volatility process  $Z$ . The processes  $Y$  and  $Z$  are driven by related subordinators  $L$  and  $\tilde{L}$ . Hence  $Y_i$  and  $Z_i$  may have simultaneous jumps, however depending on the value of the matrices  $A, B_i$  and  $C_j$  an upward jump caused by the volatility process  $Z_i$  may simultaneously have a downward jump caused by the spike process  $Y_i$ . Hence the jumps may average out. Conversely, depending on the parameters  $A, B_i$  and  $C_j$ , the jumps in  $Y_i$  and  $Z_i$  may enlarge each other and lead to a big jump in the forward curve.

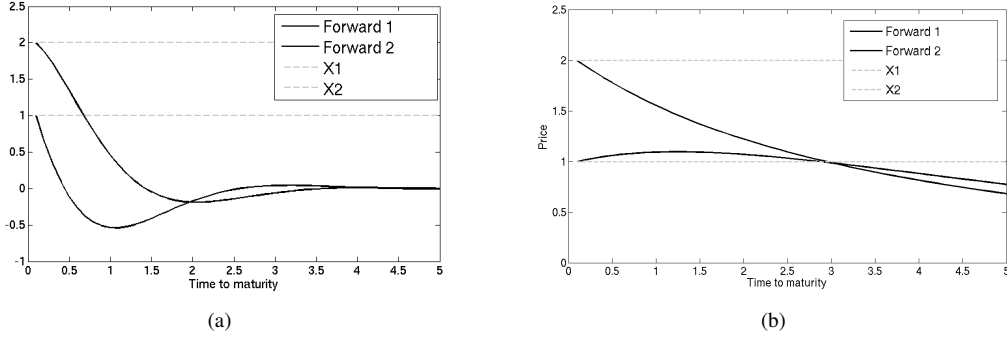


Figure 3: Paths of  $e^{A(\tau-t)}X$  for (a) complex eigenvalues, (b) real eigenvalues of  $A$ , moreover  $X = (1, 2)^T$  is taken constant.

## 6 Transform-based pricing of options

Spread options are popular derivatives in the energy market to hedge price differences. For instance, spread options are traded on the difference in electricity spot/forward prices in neighboring markets, or on the difference between electricity and one of its fuels including spark (electricity and gas) and dark (electricity and coal) spreads. On New York Mercantile Exchange (NYMEX) options on spreads between forwards on different refined oils are offered for trade.

In this section we will consider the case of pricing a spread option on two spots or forwards. The dynamics of the spots and forwards are given by our multidimensional model in the bivariate case, that is,  $d = 2$ . Our method is general and can easily be adapted to other kind of options on the same dynamics including basket options.

### 6.1 Option on the spots

Consider an option written on the spots, with strike one and exercise time at  $\tau < \infty$ . The payoff from the option is expressed through a function  $f(x)$  for a vector  $x = (x_1, x_2)^T \in \mathbb{R}^2$ . The price of the spread option on the spot becomes

$$C_s(t) = e^{-r(\tau-t)} \mathbb{E}_Q [f(S(\tau)) | \mathcal{F}_t], \quad \tau \geq t, \quad (6.1)$$

where  $Q$  is the pricing measure and  $r > 0$  is the risk-free interest rate. As noted already when pricing forward contracts in Section 5, the measure  $Q$  is not unique and has to be chosen. The same discussion applies here, except that *if* we have priced forward contracts, then we must choose the same  $Q$  when pricing the options in this market. The option price  $C$  can be computed by Monte Carlo simulation as described in Section 4 as long as we have efficient methods for simulating the processes  $Y_i$  and  $Z_j$ . This depends on the choice of  $\tilde{L}_j$ 's. An alternative, and maybe more convenient method in the present context is to base the option pricing on the Fourier transform, leading to expressions readily computable by the fast Fourier transform (FFT) algorithm. We now discuss this in more detail.

First, define the function

$$g(x) \triangleq f(e^x), \quad (6.2)$$

and observe that

$$g(\ln x) = f(x),$$

where we used pointwise exponentials and logarithms. Suppose  $g \in L^1(\mathbb{R}^2)$ , the space of absolute integrable functions on  $\mathbb{R}^2$  and recall the two-dimensional Fourier transform as

$$\hat{g}(y) = \int_{\mathbb{R}^2} g(x) e^{-i\langle x, y \rangle} dx. \quad (6.3)$$

The inverse Fourier transform becomes

$$g(x) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{g}(y) e^{i\langle y, x \rangle} dy. \quad (6.4)$$

See Folland [21] for these definitions. Note that there is a different sign in the exponentials than what is usual. We remark that for spread options, we generally have that  $g \notin L^1(\mathbb{R}^2)$ , however by exponential dampening of the pay-off function  $g$  the spread option can be fitted into the framework, as we shall discuss in a moment.

**Proposition 6.1.** *Suppose that  $g \in L^1(\mathbb{R}^2)$ , where  $g$  is defined in (6.2). Then*

$$C_s(t) = e^{-r(\tau-t)} \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{g}(y) e^{i\langle y, \ln \Lambda(\tau) \rangle} \exp \widetilde{\phi}_{\ln \widetilde{S}}^{t, \tau}(y) dy,$$

where  $\widetilde{\phi}_{\ln \widetilde{S}}^{t, \tau}(y)$  is the conditional characteristic function of  $\ln \widetilde{S}$  under  $Q$  and is stated in the Appendix in section A.6.

*Proof.* Since  $g \in L^1(\mathbb{R}^2)$ , using dominated convergence to commute integration and expectation, we conclude

$$\begin{aligned} C_s(t) &= e^{-r(\tau-t)} \mathbb{E}_Q [g(\ln S(\tau)) | \mathcal{F}_t] = e^{-r(\tau-t)} \mathbb{E}_Q [g(\ln S(\tau)) | \mathcal{F}_t], \\ &= e^{-r(\tau-t)} \mathbb{E}_Q \left[ \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{g}(y) e^{i\langle y, \ln S(\tau) \rangle} dy | \mathcal{F}_t \right], \\ &= e^{-r(\tau-t)} \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{g}(y) \mathbb{E}_Q \left[ e^{i\langle y, \ln S(\tau) \rangle} | \mathcal{F}_t \right] dy, \\ &= e^{-r(\tau-t)} \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{g}(y) e^{i\langle y, \ln \Lambda(\tau) \rangle} \exp \left( \widetilde{\phi}_{\ln \widetilde{S}}^{t, \tau}(y) \right) dy. \end{aligned}$$

This proves the result.  $\square$

As noted above, spread options fail to satisfy  $g \in L^1(\mathbb{R}^2)$  in general. We consider the specific case of a call option on the spread between two commodities, with payoff function  $f(x) = \max(x_1 - x_2 - K, 0)$ . Here,  $K$  is the strike price, and by scaling we may focus on  $K = 1$ . To include such options into the framework above, we dampen the payoff function as proposed by Carr and Madan [14]. We consider spread options, which corresponds to

$$g(x) = \max(e^{x_1} - e^{x_2} - 1, 0).$$

Define for a vector  $\xi = (\xi_1, \xi_2)$ ,

$$g_\xi(x) = e^{-\langle \xi, x \rangle} \max(e^{x_1} - e^{x_2} - 1, 0). \quad (6.5)$$

Then, as long as  $\xi_2 < 0$  and  $\xi_1 + \xi_2 > 1$ ,  $g_\xi \in L^1(\mathbb{R}^2)$ , since  $x_1 > x_2$  on the domain that  $g_\xi > 0$ . Moreover, we have that

$$g(x) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{g}_\xi(y) e^{i\langle (y-i\xi), x \rangle} dy.$$

Thus, to price a spread option entails in substituting  $y$  with  $y - i\xi$  in the formula for  $C_s(t)$  in Prop. 6.1, and use  $\widehat{g}_\xi$  instead of  $\widehat{g}$ . The exponential integrability condition must be modified in order to take into account the additional contribution from  $\exp(\langle \xi, x \rangle)$ . In the next Lemma we stated the Fourier transform of  $g_\xi$ .

**Lemma 6.2.** *Suppose  $\xi_2 < 0$  and  $\xi_1 + \xi_2 > 1$ . Then the Fourier transform of  $g_\xi(x)$  defined in (6.5) is*

$$\widehat{g}_\xi(y) = \frac{\Gamma(i(y_1 + y_2) - (1 + \xi_1 + \xi_2)) \Gamma(-iy_2 + \xi_2 + 2)}{\Gamma(iy_1 + 1 - \xi_1)}, \quad (6.6)$$

where  $\Gamma$  denotes the gamma function.

*Proof.* For the proof we follow the approach of Hurd and Zhou [25] (Theorem 1). When one takes in account the exponential lightning of the pay-off function  $g$  by  $e^{(\xi, x)}$  then the above result follows.  $\square$

Combining the above discussion it holds that a spread option priced on the spot is given by

$$C_s(t) = e^{-r(\tau-t)} \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{g}_\xi(y) e^{i\langle (y-i\xi), \ln \Lambda(\tau) \rangle} \exp(\widetilde{\phi}_{\ln \widetilde{S}}^{t, \tau}(y - i\xi)) dy$$

where  $\widehat{g}_\xi$  can be evaluated using (6.6) and  $\widetilde{\phi}_{\ln \widetilde{S}}^{t, \tau}$  is stated in the Appendix in section A.6. The FFT algorithm may be used to evaluate the analytic expression. Note that implementing this requires some numerical integration routines to evaluate the characteristic function of  $\ln \widetilde{S}$ .

## 6.2 Spread option on the forwards

Similarly a spread option on the forwards can be priced. Consider a spread option written on a forward, where the option has exercise time at  $\tau_1$  and the forward matures  $\tau_2 \geq \tau_1$ . Since we consider the same type of option the payoff function is still the same. Only the underlying dynamics differs. Hence the price of a spread option on the forwards will become

$$C_f(t) = e^{-r(\tau_1-t)} \mathbb{E}_Q [f(F(\tau_1, \tau_2)) | \mathcal{F}_t], \quad \tau_2 \geq \tau_1 \geq t.$$

Introduce the following notation  $\widetilde{F}(t, \tau) = F(t, \tau)/\Theta(t, \tau)$ , where the division is done point-wise. Using the Fourier transform method described above it holds

$$C_f(t) = e^{-r(\tau_1-t)} \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{g}_\xi(y) e^{i\langle (y-i\xi), \ln \Theta(\tau_1, \tau_2) \rangle} \exp(\widetilde{\phi}_{\ln \widetilde{F}}^{t, \tau_1, \tau_2}(y - i\xi)) dy,$$

where  $\widetilde{\phi}_{\ln \widetilde{F}}^{t, \tau_1, \tau_2}(y)$  is the conditional characteristic function of  $\ln \widetilde{F}$  under  $Q$ . An analytic expression of  $\widetilde{\phi}_{\ln \widetilde{F}}^{t, \tau_1, \tau_2}(y)$  is given in the Appendix in section A.7. Note that except of the inverse Fourier transform the level adjustment functions  $\Theta^{vol, spike}$ , represented in  $\Theta$ , need to be evaluated numerically in most cases.

## 7 Conclusions

We have proposed a model to describe the spot prices dynamics for cross-commodity markets in a multivariate setting. The model captures features like mean-reversion, spikes, stochastic volatility, inverse leverage effect and in some special cases even co-integration in continuous time. The dynamics is a multi-dimensional extension of the Barndorff-Nielsen and Shephard stochastic volatility model embedded into a mean-reversion dynamics relevant for commodity price series. The choice of the multi-dimensional extension is influenced by the work of Stelzer [40]. The multivariate spot model is analytically tractable and probabilistic properties can to a large degree be explicitly computed. We have derived various characteristics like stationary distributions and covariance functions. The model is a multivariate extension of the one-dimensional spot price dynamics analysed in Benth [8].

Analytical forward prices can be derived based on a combined Esscher-Girsanov change of measure where the risk premium is parametrized into a spike and volatility premium. Although the spot price has continuous sample paths in absence of a spike process, the implied forward curve will still exhibit jumps inherited from the stochastic volatility process. In the long end of the market the forward prices are basically equal to the seasonality function adjusted by the long-term means of the spike processes and volatility process and the market prices of risk. Since the mean-reverting structure of the involved matrix exponentials have a richer structure than in the one-dimensional case, the implied forward curve can alternate between backwardation and contango and humps may appear. Depending on the time to maturity a change in the spot can lead to various changes in the forward curve. We also discuss how a transform-based method can be used in order to price to spread options on the spot and the forward. This shows the tractability of our spot price model.

A simple approximate algorithm to simulate from matrix-valued subordinators is introduced. The method is presented and tested on a relevant example. However, further research has to be done to generate matrix-valued Lévy processes in a more general setting, a study we leave for the future.

As far as we know, there exists no methods to estimate the model based on spot price data. It is obviously of crucial interest for the applicability of the model to understand how to fit the parameters to data. Methods are available to estimate the model in the diffusion case on the quadratic covariation [6]. However these methods require high frequency data, which does not exist in the energy market. Another alternative is to adopt the methods already available for filtering spike data from price series into a multidimensional setting. If this is possible then the estimation on the spike process can be treated separately from the diffusion part, and the diffusion part can be estimated conditionally on the spike parameters. Before this can be implemented further research has to be done on the validity of these methods. Another possibility is to estimate directly on the characteristic function in the frequency domain. However the integrals appearing in the characteristic function (Prop. 3.5) are rarely explicit calculable. Already in simple one dimensional examples explicit solutions do not exist. Numerical solutions however do exist, but this is time consuming and might cause an bias. Moreover if one estimates solely on the characteristic function problems with the identifiability of the parameters will appear. Especially if one allows for several volatility and spike processes. Further research has to be done on the calibration issue.

## A Proofs and auxiliary results

In this appendix we state the proofs of various technical results in the article. We will start the Appendix with two useful results.

**Lemma A.1.** *Define  $f(s, t) := \int_s^t e^{A(t-u)} \Sigma(u) e^{A^T(t-u)} du$ . Assume that  $A$  and  $C_j$  commute for  $j = 1, \dots, n$ . Then it holds*

$$f(s, t) = \sum_{j=1}^n \mathcal{E}_j(t-s) Z_j(s) + \int_s^t \mathcal{E}_j(t-v) d\tilde{L}_j(v),$$

where  $\mathcal{E}_j(t)$  is a linear operator defined as in (3.3).

*Proof.* Using  $Z_j(t) = e^{C_j(t-s)} Z_j(s) e^{C_j^T(t-s)} + \int_s^t e^{C_j(t-v)} d\tilde{L}_j(v) e^{C_j^T(t-v)}$  and the assumption that  $A$  and  $C_j$  commute for  $j = 1, \dots, n$  it holds

$$\begin{aligned} f(s, t) &= \int_s^t e^{A^T(t-u)} \sum_{j=1}^n \omega_j \left( e^{C_j(u-s)} Z_j(0) e^{C_j^T(u-s)} + \int_s^u e^{C_j(u-v)} d\tilde{L}_j(v) e^{C_j^T(u-v)} \right) e^{A^T(t-u)} du \\ &= \sum_{j=1}^n \omega_j \int_s^t e^{(C_j-A)u} e^{At-C_j s} \left( Z_j(s) + \int_s^u e^{-C_j v} d\tilde{L}_j(v) e^{-C_j^T v} \right) e^{A^T t - C_j^T s} e^{(C_j-A)^T u} du \\ &= \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} \left( e^{(C_j-A)t} e^{At-C_j s} Z_j(s) e^{A^T t - C_j^T s} e^{(C_j-A)^T t} \right. \\ &\quad \left. - e^{(C_j-A)s} e^{At-C_j s} Z_j(s) e^{A^T t - C_j^T s} e^{(C_j-A)^T s} \right) \\ &\quad + \int_s^t \int_s^u e^{(C_j-A)u} e^{At} e^{-C_j v} d\tilde{L}_j(v) e^{-C_j^T v} e^{A^T t} e^{(C_j-A)^T u} du. \end{aligned}$$

By Fubini's theorem,

$$\begin{aligned} f(s, t) &= \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} \left( e^{C_j(t-s)} Z_j(s) e^{C_j^T(t-s)} - e^{A(t-s)} Z_j(s) e^{A^T(t-s)} \right) \\ &\quad + \int_s^t \int_v^t e^{(C_j-A)u} e^{At} e^{-C_j v} du d\tilde{L}_j(v) e^{-C_j^T v} e^{A^T t} e^{(C_j-A)^T u} \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} \left( e^{C_j(t-s)} Z_j(0) e^{C_j^T(t-s)} - e^{A(t-s)} Z_j(0) e^{A^T(t-s)} \right) \\
&\quad + (\mathbf{C}_j - \mathbf{A})^{-1} \left( \int_s^t e^{(C_j-A)t} e^{At} e^{-C_j v} d\tilde{L}_j(v) e^{-C_j^T v} e^{A^T t} e^{(C_j-A)^T t} \right) \\
&\quad - (\mathbf{C}_j - \mathbf{A})^{-1} \left( \int_s^t e^{(C_j-A)v} e^{Av} e^{-C_j v} d\tilde{L}_j(v) e^{-C_j^T v} e^{A^T t} e^{(C_j-A)^T v} \right) \\
&= \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} \left( e^{C_j(t-s)} Z_j(s) e^{C_j^T(t-s)} - e^{A(t-s)} Z_j(s) e^{A^T(t-s)} \right) \\
&\quad + (\mathbf{C}_j - \mathbf{A})^{-1} \left( \int_s^t e^{C_j(t-v)} d\tilde{L}_j(v) e^{C_j^T(t-v)} - \int_s^t e^{A(t-v)} d\tilde{L}_j(v) e^{A^T(t-v)} \right) \\
&= \sum_{j=1}^n \mathcal{E}_j(t-s) Z_j(s) + \int_s^t \mathcal{E}_j(t-v) d\tilde{L}_j(v).
\end{aligned}$$

□

**Lemma A.2.** *Let  $L$  be a Lévy process in  $\mathbb{R}^d$  satisfying  $\mathbb{E} \|L(1)\|^r < \infty$  for some  $r > 1$ . If  $f \in M_d$  is a measurable deterministic function which is bounded on  $[s, t]$  then the expectation of the integral  $\int_s^t f(u) dL(u)$  satisfies,*

$$\mathbb{E} \left[ \int_s^t f(u) dL(u) \right] = \int_s^t f(u) du \mathbb{E}[L(1)]. \quad (\text{A.1})$$

*Proof.* For a partition  $s = u_0 < \dots < u_N = t$  of the interval  $[s, t]$ , by the independence and stationarity of the increments of  $L$ , we have

$$\begin{aligned}
\mathbb{E} \left[ \sum_{k=0}^{N-1} f(u_k) (L(u_{k+1}) - L(u_k)) \right] &= \sum_{k=0}^{N-1} \mathbb{E}[f(u_k) (L(u_{k+1}) - L(u_k))] \\
&= \sum_{k=0}^{N-1} \mathbb{E}[f(u_k) (L(u_{k+1} - u_k))] \\
&= \sum_{k=0}^{N-1} f(u_k) \cdot (u_{k+1} - u_k)
\end{aligned}$$

If the mesh of the partition goes to zero, the right-hand side converges to  $\int_s^t f(u) du$ . Since  $f$  is bounded on  $[s, t]$  there exist a partition such that  $X_N \triangleq \sum_{k=0}^{N-1} f(u_k) (L(u_{k+1}) - L(u_k))$  converges in probability to  $\int_s^t f(u) dL(u)$  (Sato [36], prop. 4.5). Moreover  $\|X_N\|$  is in  $L^r$  for some  $r > 1$ , hence the family  $\{X_N\}_{N \in \mathbb{N}}$  is uniformly integrable. Therefor convergence in probability implies convergence in  $L^1$  (see Williams [43], Theorem 13.7). So we may conclude that

$$\mathbb{E} \left[ \int_s^t f(u) dL(u) \right] = \int_s^t f(u) du \mathbb{E}[L(1)].$$

□

### A.1 Proof of Lemma 3.1

*Proof.* The conditional expectations are given by

$$\mathbb{E}[X(t) | \mathcal{F}_s] = e^{A(t-s)} X(s) + \mathbb{E} \left[ \int_s^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u) \right],$$



$$\begin{aligned}
&= e^{A(t-s)}X(s) + \mathbb{E} \left[ \mathbb{E} \left[ \int_s^t e^{A(s-u)}\Sigma(u)^{1/2} dW(u) | \Sigma(u)_{s \leq u \leq t} \right] \right], \\
&= e^{A(t-s)}X(s) \\
\mathbb{E}[Y_i(t)|\mathcal{F}_s] &= e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i + \mathbb{E} \left[ \int_s^t e^{B_i(t-u)}\eta_i dL_i(u) | \mathcal{F}_s \right], \\
&= e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i + \int_s^t e^{B_i(t-u)}\eta_i du \cdot \mathbb{E}[L_i(1)], \\
&= e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i + B_i^{-1} \left( \eta_i - e^{B_i(t-s)}\eta_i \right) \mathbb{E}[L_i(1)].
\end{aligned}$$

where we used Lemma A.2 to obtain the last equality.  $\square$

## A.2 Proof of Lemma 3.2

*Proof.* The conditional variance for the process  $X$  is given by,

$$\begin{aligned}
\text{Var}[X(t)|\mathcal{F}_s] &= \mathbb{E} \left[ \left( e^{A(t-s)}X(s) + \int_s^t e^{A(s-u)}\Sigma(u)^{1/2} dW(u) \right)^2 | \mathcal{F}_s \right] - \mathbb{E}[X(t)|\mathcal{F}_s]^2, \\
&= 2e^{A(t-s)}X(s)\mathbb{E} \left[ \int_s^t e^{A(s-u)}\Sigma(u)^{1/2} dW(u) | \Sigma(u)_{s \leq u \leq t} \right] \\
&\quad + \mathbb{E} \left[ \mathbb{E} \left[ \left( \int_s^t e^{A(s-u)}\Sigma(u)^{1/2} dW(u) \right)^2 | \Sigma(u)_{s \leq u \leq t} \right] \right], \\
&= \mathbb{E} \left[ \int_s^t e^{A(t-u)}\Sigma(u)^{1/2} (\Sigma(u)^{1/2})^T e^{A^T(t-u)} du \right] \\
&= \mathbb{E}[f(s, t)],
\end{aligned}$$

with  $f$  as in Lemma A.1. Hence,

$$\begin{aligned}
\text{Var}[X(t)|\mathcal{F}_s] &= \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} \left( e^{C_j(t-s)}Z_j(s)e^{C_j^T(t-s)} - e^{A(t-s)}Z_j(s)e^{A^T(t-s)} \right. \\
&\quad \left. + \mathbb{E} \left[ \int_s^t e^{C_j(t-v)}d\tilde{L}_j(v)e^{C_j^T(t-v)} - \int_s^t e^{A(t-v)}d\tilde{L}_j(v)e^{A^T(t-v)} \right] \right), \\
&= \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} \left( e^{C_j(t-s)}Z_j(s)e^{C_j^T(t-s)} - e^{A(t-s)}Z_j(s)e^{A^T(t-s)} \right. \\
&\quad + \mathbf{C}_j^{-1} \left[ \mathbb{E}[\tilde{L}_j(1)] - e^{C_j(t-s)}\mathbb{E}[\tilde{L}_j(1)]e^{C_j^T(t-s)} \right] \\
&\quad \left. - \mathbf{A}^{-1} \left[ \mathbb{E}[\tilde{L}_j(1)] - e^{A(t-s)}\mathbb{E}[\tilde{L}_j(1)]e^{A^T(t-s)} \right] \right),
\end{aligned}$$

where we vectorized the integrals and applied Lemma A.2 to obtain the last equality. Passing to the limit when  $t \rightarrow \infty$  establishes that

$$\text{Var}[X(t)|\mathcal{F}_s] \rightarrow \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} [\mathbf{C}_j^{-1} - \mathbf{A}^{-1}] \mathbb{E}[\tilde{L}(1)].$$

We calculate similarly for the conditional variance of the processes  $Y_i$ ,

$$\text{Var}[Y_i(t)|\mathcal{F}_s] = \left( e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i \right) \left( e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i \right)^T$$

$$\begin{aligned}
& + \left( e^{B_i(t-s)} Y_i(s) + B_i^{-1} (I - e^{B_i(t-s)}) \mu_i \right) \mathbb{E} \left[ \left( \int_s^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T \right] \\
& + \mathbb{E} \left[ \left( \int_s^t e^{B_i(t-u)} \eta_i dL_i(u) \right) \right] \left( e^{B_i(t-s)} Y_i(s) + B_i^{-1} (I - e^{B_i(t-s)}) \mu_i \right)^T \\
& + \mathbb{E} \left[ \left( \int_s^t e^{B_i(t-u)} \eta_i dL_i(u) \right) \left( \int_s^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T \right] - \mathbb{E} [Y_i(t) | \mathcal{F}(s)]^2, \\
& \stackrel{*}{=} \int_s^t e^{B_i(t-u)} \eta_i \text{Var}[L_i(1)] (\eta_i)^T e^{B_i^T(t-u)} du - \mathbb{E} \left[ \int_s^t e^{B_i(t-u)} \eta_i dL_i(u) | \mathcal{F}_s \right]^2, \\
& = \mathbf{B}_i^{-1} \left( \eta_i \text{Var}[L(1)] (\eta_i)^T - e^{B_i(t-s)} \eta_i \text{Var}[L_i(1)] (\eta_i)^T e^{B_i^T(t-s)} \right) \\
& \quad - \left( B_i^{-1} \left( \eta_i - e^{B_i(t-s)} \eta_i \right) \mathbb{E} [L_i(1)] \right)^2.
\end{aligned}$$

Passing to the limit gives,

$$\text{Var}[Y_i(t) | \mathcal{F}_s] \rightarrow \mathbf{B}_i^{-1} \eta_i \text{Var}[L_i(1)] (\eta_i)^T - \left( B_i^{-1} \eta_i \mathbb{E} [L_i(1)] \right)^2,$$

as  $t \rightarrow \infty$ . For \* see Marquadt and Stelzert [29]. Note that  $L$  denotes the vector-valued process, which appears in the dynamics of  $Y$ .

From the independence of  $(\Sigma_t)_{t \in \mathbb{R}^+}$  and  $(W_t)_{t \in \mathbb{R}^+}$  the covariance is given by,

$$\begin{aligned}
\text{Cov}[X(t), Y_i(t)] &= \mathbb{E} \left[ \left( \int_0^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u) \right) \left( \int_0^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T \right], \\
&= \mathbb{E} \left[ \mathbb{E} \left[ \left( \int_0^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u) \right) \left( \int_0^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T \mid (L_i(s))_{s \in [0, t]} \right] \right], \\
&= \mathbb{E} \left[ \mathbb{E} \left[ \left( \int_0^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u) \right) \mid (L_i(s))_{s \in [0, t]} \right] \left( \int_0^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T \right], \\
&= \mathbb{E} \left[ 0 \cdot \left( \int_0^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T \right] = 0.
\end{aligned}$$

Moreover, for  $i \neq j$  we have,

$$\begin{aligned}
\text{Cov}[Y_i(t), Y_j(t)] &= \mathbb{E} \left[ \left( \int_0^t e^{B_i(t-u)} \eta_i dL_i(u) \right) \left( \int_0^t e^{B_j(t-u)} \eta_j dL_j(u) \right)^T \right] \\
&\quad - \mathbb{E} \left[ \int_0^t e^{B_i(t-u)} \eta_i dL_i(u) \right] \mathbb{E} \left[ \left( \int_0^t e^{B_j(t-u)} \eta_j dL_j(u) \right)^T \right], \\
&= 0,
\end{aligned}$$

since  $L_i$  and  $L_j$  are independent. With similar calculations as above one can show that  $\text{Cov}[X(t+h), Y_i(t)] = 0$  and  $\text{Cov}[Y_j(t+h), Y_i(t)] = 0$  for  $h \in \mathbb{R}$ . So we conclude that  $\text{acov}_{\text{Im}(\bar{s})}(h) = \text{acov}_X(h) + \text{acov}_{\Sigma Y_i}(h)$ . Moreover it holds

$$\begin{aligned}
\text{acov}_X(h) &\stackrel{*}{=} e^{A|h|} \text{Var}[X(t)] \rightarrow e^{A|h|} \sum_{j=1}^n \omega_j (\mathbf{C}_j - \mathbf{A})^{-1} [\mathbf{C}_j^{-1} - \mathbf{A}^{-1}] \mathbb{E}[\tilde{L}_j(1)] \quad \text{as } t \rightarrow \infty, \\
\text{acov}_{\Sigma Y_i}(h) &= \sum_i e^{B_i|h|} \text{Var}[Y_i(t)] \rightarrow \sum_i e^{B_i|h|} \mathbf{B}_i^{-1} \eta_i \text{Var}[L_i(1)] (\eta_i)^T \quad \text{as } t \rightarrow \infty.
\end{aligned}$$

Here, \* follows from Marquadt and Stelzert [29], Prop. 3.13. Moreover,  $\text{Var}[L_i(1)] = \int_{\mathbb{R}^d} x x^T \nu_L^i(dx)$  (See Sato [35] p. 163).  $\square$

### A.3 Proof of proposition 3.3

We will first provide the characteristic function of the stationary distribution of the processes  $Y_i$ ,  $i = 1, \dots, m$  and  $Z$ . Thereafter we will in several steps calculate the characteristic function of the stationary distribution of  $X$ .

We have that

$$Y_i(t) = e^{B_i(t-s)} Y_i(s) + B_i^{-1} (I - e^{B_i(t-s)}) \mu_i + \int_s^t e^{B_i(t-u)} \eta_i dL_i(u),$$

hence, by the key-formula (See Sato [36]) the conditional characteristic function of  $Y_i(t)$  given  $\mathcal{F}_s$  is

$$\begin{aligned} \phi_Y^{i,(s,t)}(w) &= i \left( e^{B_i(t-s)} Y_i(s) + i B_i^{-1} (I - e^{B_i(t-s)}) \mu_i \right)^T w + \int_s^t \phi_L^i(\eta_i^T e^{B_i^T(t-u)} w) du, \\ &= i \left( e^{B_i(t-s)} Y_i(s) + i B_i^{-1} (I - e^{B_i(t-s)}) \mu_i \right)^T w + \int_0^{t-s} \phi_{\tilde{L}}^i(J_d(\eta_i^T e^{B_i^T u} w)) du. \end{aligned} \quad (\text{A.2})$$

This is well-defined for  $w \in \mathbb{R}^d$ . Since  $L(t)$  has finite log moments and  $\sigma(B_i) \subseteq (-\infty, 0] + i\mathbb{R}^+$  the limit for  $t \rightarrow \infty$  is well-defined (see Sato [36]) and is given by

$$\phi_Y^i(w) = i \mu_i^T (B_i^T)^{-1} w + \int_0^\infty \phi_{\tilde{L}}^i(J_d(\eta_i^T e^{B_i^T u} w)) du, \quad w \in \mathbb{R}^d.$$

For the characteristic function of  $Z_i$ 's see Pigorsch and Stelzer [33].

Now we will analyze the stationarity of the process  $X(t)$  and derive the characteristic function of the stationary distribution. Using Lemma A.1 we can derive the conditional characteristic function of  $X$  given a filtration up to time  $s$ ,  $\phi_X^{s,t}(w) \triangleq \mathbb{E}[e^{i\langle w, X(t) \rangle} | \mathcal{F}_s]$ .

**Proposition A.3.** *Suppose that  $C_j$  and  $A$  commute for  $j = 1, \dots, n$ . Then the conditional characteristic function of the process  $X(t)$  given  $\mathcal{F}_s$  is*

$$\phi_X^{s,t}(w) = i X^T(s) e^{A^T(t-s)} w + i \sum_{j=1}^n \text{tr} \left( \frac{1}{2} w w^T \mathcal{E}_j(t-s) Z_j(s) \right) + \sum_{j=1}^n \int_0^{t-s} \phi_{\tilde{L}}^j \left( \frac{1}{2} w w^T \mathcal{E}_j^T(u) I_d \right) du, \quad (\text{A.3})$$

where  $\mathcal{E}_j$  is as defined in (3.3).

*Proof.* By Girsanov's theorem and the independence of  $W$  and  $\tilde{L}_j$  for  $j = 1 \dots n$  we have that,

$$\begin{aligned} \phi_X^{s,t}(w) &= \ln \mathbb{E} \left[ e^{i\langle w, X(t) \rangle} | \mathcal{F}_s \right] = \mathbb{E} \left[ \mathbb{E} \left[ e^{i\langle w, X(t) \rangle} | (\Sigma(t))_{t \in \mathbb{R}^+} \right] | \mathcal{F}_s \right] \\ &= i X^T(s) e^{A^T(t-s)} w + \ln \mathbb{E} \left[ \mathbb{E} \left[ \exp \left( i \left( \int_s^t \Sigma(u)^{1/2} e^{A(t-u)} dW(u) \right)^T w \right) | (\Sigma(t))_{t \in \mathbb{R}^+} \right] \right] \\ &= i X^T(s) e^{A^T(t-s)} w + \ln \mathbb{E} \left[ \mathbb{E} \left[ \exp \left( i \frac{1}{2} w^T \int_s^t e^{A(t-u)} \Sigma(u) e^{A^T(t-u)} du w \right) | (\Sigma(t))_{t \in \mathbb{R}^+} \right] \right] \\ &= i X^T(s) e^{A^T(t-s)} w + \ln \mathbb{E} \left[ \exp \left( i \frac{1}{2} w^T \int_s^t e^{A(t-u)} \Sigma(u) e^{A^T(t-u)} du w \right) \right] \\ &= i X^T(s) e^{A^T(t-s)} w + \ln \mathbb{E} \left[ \exp \left( i \frac{1}{2} w^T f(s, t) w \right) \right] \quad \text{Here } f \text{ is as in Lemma A.1.} \\ &= i X^T(s) e^{A^T(t-s)} w + \sum_{j=1}^n i \text{tr} \left( \frac{1}{2} w^T \mathcal{E}_j(t-s) Z_j(s) w \right) \\ &\quad + \sum_{j=1}^n \ln \mathbb{E} \left[ e^{i \text{tr} \left( \left( \frac{1}{2} \int_s^t w w^T \mathcal{E}_j(t-u) d\tilde{L}_j(u) \right)^T I_d \right)} \right]. \end{aligned}$$

From the usual formula for the Fourier transform of an integral with a deterministic integrand with respect to a Lévy process (see Marquadt and Stelzer [29] for a brief review) we have that

$$\mathbb{E} \left[ e^{\text{itr} \left( \left( \frac{1}{2} \int_s^t ww^T \mathcal{E}_j(t-u) d\tilde{L}_j(u) \right)^T I_d \right)} \right] = e^{\int_s^t \phi_L^j \left( \frac{1}{2} ww^T \mathcal{E}_j^T(t-u) I_d \right) du} = e^{\int_0^{t-s} \phi_L^j \left( \frac{1}{2} ww^T \mathcal{E}_j^T(u) I_d \right) du}.$$

Hence

$$\begin{aligned} \phi_X(w) &= iX^T(s)e^{A^T(t-s)}w + \sum_{j=1}^n \text{itr} \left( \frac{1}{2} ww^T \mathcal{E}_j(t-s) Z_j(s) \right) \\ &\quad + \sum_{j=1}^n \ln \mathbb{E} \left[ e^{\text{itr} \left( \left( \frac{1}{2} ww^T \int_s^t \mathcal{E}_j(t-u) d\tilde{L}_j(u) \right)^T I_d \right)} \right], \\ &= iX^T(s)e^{A^T(t-s)}w + \sum_{j=1}^n \text{itr} \left( \frac{1}{2} ww^T \mathcal{E}_j(t-s) Z_j(s) \right) \\ &\quad + \sum_{j=1}^n \int_0^{t-s} \phi_L^j \left( \frac{1}{2} ww^T \mathcal{E}_j^T(u) I_d \right) du. \end{aligned}$$

□

We can prove the stationarity of  $X(t)$  and derive the characteristic function for the limiting distribution.

**Proposition A.4.** *Suppose the matrices  $A$  and  $C_j$  commute for  $j = 1, \dots, n$ . The process  $X(t)$  is stationary and the characteristic function of the limiting distribution is given by*

$$\phi_X(w) = \sum_{j=1}^n \int_0^\infty \phi_L^j \left( \frac{1}{2} ww^T \mathcal{E}_j^T(s) I_d \right) ds,$$

where the linear operator  $\mathcal{E}_j$  is defined in (3.3).

*Proof.* It holds  $\lim_{t \rightarrow \infty} \text{tr} \left( \mathcal{E}_j(t) V \right) = 0$  for  $\forall V \in \mathbb{S}_d^+$ . Using the vec transformation we have that,

$$\begin{aligned} \text{tr} \left( \mathcal{E}_j(t) V \right) &= \left( (I_d \otimes (C_j - A)) + ((C_j - A) \otimes I_d) \right)^{-1} \left( \text{vec} \left( e^{C_j t} V e^{C_j^T t} \right) - \text{vec} \left( e^{A t} V e^{A^T t} \right) \right), \\ &= \left( (I_d \otimes (C_j - A)) + ((C_j - A) \otimes I_d) \right)^{-1} \\ &\quad \times \left( e^{((I_d \otimes C_j) + (C_j \otimes I_d))t} - e^{((I_d \otimes A) + (A \otimes I_d))t} \right) \text{vec}(V) \\ &\rightarrow 0, \end{aligned}$$

as  $t \rightarrow \infty$ . The convergence follows from the fact that for all  $A \in M_d(\mathbb{R})$  with  $\sigma(A) \subseteq (-\infty, 0] + i\mathbb{R}^+$ ,

$$\sigma(A \otimes I_d + I_d \otimes A) \subseteq \sigma(A \otimes I_d) + \sigma(I_d \otimes A) = \sigma(A) + \sigma(A) \subseteq (-\infty, 0] + i\mathbb{R}^+$$

by Horn and Johnson [24], Thm. 2.4.9. The addition on the right hand side is the usual set addition. We conclude that the second term in (A.3) converges to zero as  $t \rightarrow \infty$ . Since the eigenvalues of  $A$  have negative real-parts also the first term in (A.3) converges to zero.

We must show that the integrals in the third term of (A.3) are converging. To prove this it is sufficient to show that  $\text{vec} \left( \frac{1}{2} ww^T \int_s^t \mathcal{E}_j(t-u) d\tilde{L}_j(u) \right)$  has a stationary solution. Note that for  $D \in M_d(\mathbb{R})$  with  $\sigma(D) \subseteq (-\infty, 0] + i\mathbb{R}^+$ , the integral  $I_t(D) := \int_s^t e^{((I_d \otimes D) + (D \otimes I_d))(t-u)} d\text{vec}(\tilde{L}(u))$  has a stationary solution by Sato [36] Theorem 5.2. Moreover it holds

$$\begin{aligned} \text{vec} \left( \int_s^t \mathcal{E}_j(t-u) d\tilde{L}_j(u) \right) &= \frac{\omega_j}{2} \text{vec}(I \otimes ww^T) \\ &\quad \times \left( (I_d \otimes (C_j - A)) + ((C_j - A) \otimes I_d) \right)^{-1} (I_t(C_j) - I_t(A)). \end{aligned}$$

Hence, we have a linear combination of two stationary processes, which is therefore stationary. □

#### A.4 Proof of Lemma 5.1

*Proof.* For simplicity we avoid notating indices  $j$  in  $\nu_{\tilde{L}}^j$ . The proof is equivalent for all  $j$ . We assumed the subordinators  $\tilde{L}$  to be driftless. So

$$\begin{aligned}\mathbb{E}_Q[e^{tr(V\tilde{L})}] &= \phi_{\tilde{L}}(-i(V+\theta)) - \phi_{\tilde{L}}(-i\theta) \\ &= \int_{\mathbb{S}_d^+ \setminus \{0\}} \left( e^{tr((V+\theta)x)} - 1 \right) \nu_{\tilde{L}}(dx) - \int_{\mathbb{S}_d^+ \setminus \{0\}} \left( e^{tr(\theta x)} - 1 \right) \nu_{\tilde{L}}(dx), \\ &= \int_{\mathbb{S}_d^+ \setminus \{0\}} \left( e^{tr(Vx)} - 1 \right) e^{tr(\theta x)} \nu_{\tilde{L}}(dx) \\ &= \phi_{\tilde{L}}^Q(-iV) \quad \forall V \in \mathbb{S}_d^+.\end{aligned}$$

We still need to prove that  $\nu_{\tilde{L}}^Q(dx) = \exp(tr(\theta x))\nu_{\tilde{L}}(dx)$  is a Lévy measure of a subordinator, *i.e.*,  $\nu_{\tilde{L}}^Q(\{0\}) = 0$ ,  $\int_{\mathbb{S}_d^+} (|x|^2 \wedge 1) \nu_{\tilde{L}}^Q(dx) < \infty$  and  $\int_{\mathbb{S}_d^+ \cap (0,1]^d} x \nu_{\tilde{L}}^Q(dx) < \infty$ . Here,  $(0,1]^d$  is a shorthand notation for  $(0,1] \times (0,1] \times \dots \times (0,1] \subset \mathbb{R}^d$ .

Since  $x \mapsto e^{tr(\theta x)}$  is a continuous function, it will reach a maximum on a compact set. Therefore

$$\int_{\mathbb{S}_d^+ \cap (0,1]^d} |x|^2 \nu_{\tilde{L}}^Q(dx) \leq \max(e^{tr(\theta x)}) \int_{\mathbb{S}_d^+ \cap (0,1]^d} |x|^2 \nu_{\tilde{L}}(dx) < \infty,$$

since  $\nu_{\tilde{L}}$  is a Lévy measure. Similarly, since  $\nu_{\tilde{L}}$  is the Lévy measure of a subordinator and  $z \mapsto e^{tr(\theta z)}$  will attain a maximum on a compact set

$$\int_{\mathbb{S}_d^+ \cap (0,1]^d} x \nu_{\tilde{L}}^Q(dx) \leq \max(e^{tr(\theta x)}) \int_{\mathbb{S}_d^+ \cap (0,1]^d} x \nu_{\tilde{L}}(dx) < \infty.$$

Moreover by assumption on  $\theta$ ,  $\int_{\mathbb{S}_d^+ \cap (1,\infty)^d} e^{tr(\theta x)} \nu_{\tilde{L}}(dx) < \infty$ . Therefore

$$\int_{\mathbb{S}_d^+} (|x|^2 \wedge 1) \nu_{\tilde{L}}^Q(dx) \leq \int_{\mathbb{S}_d^+ \cap (0,1]^d} |x|^2 \nu_{\tilde{L}}^Q(dx) + \int_{\mathbb{S}_d^+ \cap (1,\infty)^d} \nu_{\tilde{L}}^Q(dx) < \infty$$

□

#### A.5 Proof of Lemma 5.2

*Proof.* For simplicity, we let  $m = n = 1$ . The general case follows readily, by appealing to independence of the different subordinators. Further, we skip indices when referring to  $Y_1$  and  $Z_1$ . Using Lemma 3.1, the  $Q$ -Brownian motion  $W$  and the adaptedness of  $X(t)$  and  $Y(t)$ , we get

$$\begin{aligned}F(t, \tau) &= \mathbb{E}_Q[\Lambda(\tau) \cdot \exp(X(\tau) + Y(\tau)) \mid \mathcal{F}_t] \\ &= \Lambda(\tau) \cdot \exp\left(e^{A(\tau-t)} X(t) + e^{B(\tau-t)} Y(t) + A^{-1}(I - e^{A(\tau-t)})\theta_0 + B^{-1}(I - e^{B(\tau-t)})\mu\right) \\ &\quad \times \mathbb{E}_Q\left[\exp\left(\int_t^\tau e^{A(\tau-s)} \Sigma^{1/2}(s) dW(s) + \int_t^\tau e^{B(\tau-s)} \eta dL(s)\right) \mid \mathcal{F}_t\right].\end{aligned}$$

We consider the expectation in the last equality. Introduce the process  $\tilde{V}(t)$  by

$$\tilde{V}(t) = \exp\left(\int_0^t e^{A(\tau-s)} \Sigma^{1/2}(s) dW(s) - \frac{1}{2} \text{diag} \left[ \int_0^t e^{A(\tau-s)} \Sigma(s) e^{A^T(\tau-s)} ds \right]\right),$$

which by Novikov's condition is a martingale. It will be the density process on  $[0, \tau]$  for a probability measure  $\tilde{Q} \sim Q$ . From Shiryaev [39], the characteristics of  $\tilde{L}_j$ ,  $j = 1, \dots, n$  will not change, only the  $Q$ -Brownian motion  $W$  will get a drift added under  $\tilde{Q}$  by Girsanov's Theorem. Hence, we calculate

$$\hat{F}(t, \tau) \triangleq \mathbb{E}_Q\left[\exp\left(\int_t^\tau e^{A(\tau-s)} \Sigma^{1/2}(s) dW(s) + \int_t^\tau e^{B(\tau-s)} \eta dL(s)\right) \mid \mathcal{F}_t\right]$$

$$\begin{aligned}
&= \mathbb{E}_{\tilde{Q}} \left[ \exp \left( \frac{1}{2} \text{diag} \left[ \int_t^\tau e^{A(\tau-s)} \Sigma(s) e^{A^T(\tau-s)} ds \right] + \int_t^\tau e^{B(\tau-s)} \eta dL(s) \right) \mid \mathcal{F}_t \right] \\
&= \mathbb{E}_Q \left[ \exp \left( \frac{1}{2} \text{diag} \left[ \int_t^\tau e^{A(\tau-s)} Z(s) e^{A^T(\tau-s)} ds \right] + \int_t^\tau e^{B(\tau-s)} \eta dL(s) \right) \mid \mathcal{F}_t \right].
\end{aligned}$$

Using Lemma A.1,

$$\begin{aligned}
\hat{F}(t, \tau) &= \mathbb{E}_Q \left[ \exp \left( \int_t^\tau e^{A(\tau-s)} \Sigma^{1/2}(s) dW(s) + \int_t^\tau e^{B(\tau-s)} \eta dL(s) \right) \mid \mathcal{F}_t \right] \\
&= \exp \left( \frac{1}{2} \text{diag} \left[ \mathcal{E}(\tau - t) Z(t) \right] \right) \\
&\quad \times \mathbb{E}_Q \left[ \exp \left( \text{diag} \left[ \frac{1}{2} \int_t^\tau \mathcal{E}(\tau - s) d\tilde{L}(s) \right] + \int_t^\tau e^{B(\tau-s)} \eta dL(s) \right) \mid \mathcal{F}_t \right] \\
&= \exp \left( \frac{1}{2} \text{diag} \left[ \mathcal{E}(\tau - t) Z(t) \right] \right) \\
&\quad \times \mathbb{E}_Q \left[ \exp \left( \text{diag} \left[ \frac{1}{2} \int_t^\tau \mathcal{E}(\tau - s) d\tilde{L}(s) \right] + \int_t^\tau e^{B(\tau-s)} \eta dL(s) \right) \right].
\end{aligned}$$

The last equality follows from the independent increment property of subordinators.  $F(t, \tau)$  is a vector in  $\mathbb{R}^d$ . We will calculate its entries separately. It holds that  $k$ -th entry of  $\hat{F}$  is given by,

$$\begin{aligned}
\hat{F}_k(t, \tau) &= \exp \left( \frac{1}{2} \text{tr} \left( \mathcal{D}_k(\tau - t) Z(t) \right) \right) \\
&\quad \times \mathbb{E}_Q \left[ \exp \left( \text{tr} \left( \int_t^\tau \frac{1}{2} \mathcal{D}_k(\tau - s) d\tilde{L}(s) + \mathcal{E}_k \left[ e^{B(\tau-s)} \eta \right] d\tilde{L}(s) \right) \right) \right].
\end{aligned}$$

Hence, using Lemma 5.1 we find that

$$\begin{aligned}
\hat{F}_k(t, \tau) &= \mathbb{E}_Q \left[ \exp \left( \int_t^\tau e^{A(\tau-s)} \Sigma^{1/2}(s) dW(s) + \int_t^\tau e^{B(\tau-s)} \eta dL(s) \right) \mid \mathcal{F}_t \right] \\
&= \exp \left( \frac{1}{2} \text{tr} \left( \mathcal{D}_k(\tau - t) Z(t) \right) \right) \\
&\quad \times \mathbb{E}_Q \left[ \exp \left( \int_t^\tau \phi_{\tilde{L}} \left( \frac{-i}{2} \mathcal{D}_k^T(\tau - s) I_d - i \mathcal{E}_k \left[ e^{B(\tau-s)} \eta \right] I_d - i\theta \right) - \phi_{\tilde{L}}(-i\theta) ds \right) \right].
\end{aligned}$$

This completes the proof.  $\square$

## A.6 Conditional characteristic function of $\ln \tilde{S}$ under $Q$

*Proof.* The conditional characteristic function of  $\ln \tilde{S}$  under  $P$  as stated in the proof of Prop. 3.5 needs only a small adjustment in order to take in account the dynamics under  $Q$ . Using the dynamics of  $X$  under  $Q$  in (5.6) and plugging in the characteristic function of  $\tilde{L}$  under  $Q$  we get

$$\begin{aligned}
\tilde{\phi}_{\ln \tilde{S}}^{s,t}(w) &= iX^T(s) e^{A^T(t-s)} w + iA^{-1}(I - e^{A(t-s)})\theta_0 + \sum_{i=1}^n \text{itr} \left( \frac{1}{2} w w^T \mathcal{E}_i(t-s) Z_i(s) \right) \\
&\quad + \sum_{j=1}^m iY_j^T(s) e^{B_j^T(t-s)} w + i(B_j^{-1}(I - e^{B_j(t-s)})\mu^j)^T w \\
&\quad + \sum_{k=1}^m \int_0^{t-s} \phi_{\tilde{L}}^k \left( \frac{1}{2} w w^T \mathcal{E}_k^T(u) I_d + iJ_d(\eta_k^T e^{B_k^T u} w) \right) du
\end{aligned}$$

$$+ \sum_{j=m_1+1}^n \int_0^{t-s} \phi_{\widehat{L}}^j \left( \frac{1}{2} w w^T \mathcal{E}_j^T(u) I_d \right) du.$$

□

## A.7 Conditional characteristic function of $\ln \widetilde{F}$ under $Q$

*Proof.* For simplicity, we let  $m = n = 1$ . Using (5.12) we have that

$$\ln \widetilde{F}(\tau_1, \tau_2) = e^{A(\tau_2-\tau_1)} X(\tau_1) + e^{B(\tau_2-\tau_1)} Y(\tau_1) + \Psi(\tau_1, \tau_2, Z(\tau_1)).$$

Using the explicit solutions from Lemma 3.1 and the dynamics under  $Q$  (5.6), (5.7) it holds

$$\begin{aligned} e^{A(\tau_2-\tau_1)} X(\tau_1) &= e^{A(\tau_2-t)} X(t) + e^{A(\tau_2-\tau_1)} A^{-1} (I - e^{A(\tau_1-t)}) \theta_0 + \int_t^{\tau_1} e^{A(\tau_2-s)} \Sigma^{1/2}(s) d\widehat{W}(s), \\ e^{B(\tau_2-\tau_1)} Y(\tau_1) &= e^{B(\tau_2-t)} Y(t) + e^{B(\tau_2-\tau_1)} B^{-1} (I - e^{B(\tau_1-t)}) \mu + \int_t^{\tau_1} e^{B(\tau_2-s)} \eta dL_Q(s), \\ Z(\tau_1) &= e^{C(\tau_1-t)} Z(t) e^{C^T(\tau_1-t)} + \int_t^{\tau_1} e^{C(\tau_1-s)} d\widehat{L}(s) e^{C^T(\tau_1-s)}. \end{aligned}$$

Define the following operators, using analogue considerations as in Prop 5.2

$$\widehat{\mathcal{E}}_i(t, \tau_1, \tau_2) : X \mapsto \omega_i \left[ (C_i - A)^{-1} \left( e^{C_i(\tau_1-t) + A(\tau_2-\tau_1)} X e^{C_i^T(\tau_1-t) + A^T(\tau_2-\tau_1)} - e^{A(\tau_2-t)} X e^{A^T(\tau_2-t)} \right) \right],$$

and

$$\begin{aligned} \widehat{\mathcal{D}}_{k,i}(t, \tau_1, \tau_2) : X \mapsto \omega_i e_{(k-1)d+k}^T &\left[ ((C_i - A) \otimes I + I \otimes (C_i - A))^{-1} \right. \\ &\left. \left( e^{(A \otimes I + I \otimes A)(\tau_2-\tau_1) + (C_i \otimes I + I \otimes C_i^T)(\tau_1-t)} - e^{(A \otimes I + I \otimes A)(\tau_2-t)} \right) \right]. \end{aligned}$$

For a vector  $w \in \mathbb{R}^d$  it holds

$$\begin{aligned} \widetilde{\phi}_{\ln \widetilde{F}}^{t, \tau_1, \tau_2}(w) &= \ln \mathbb{E} \left[ e^{i \langle w, \ln \widetilde{F}(\tau_1, \tau_2) \rangle} | \mathcal{F}_t \right], \\ &= i X^T(t) e^{A^T(\tau_2-t)} w + i Y^T(t) e^{B^T(\tau_2-t)} w \\ &\quad + i \left( e^{A(\tau_2-\tau_1)} A^{-1} (I - e^{A(\tau_1-t)}) \theta_0 + e^{B(\tau_2-\tau_1)} B^{-1} (I - e^{B(\tau_1-t)}) \mu \right)^T w \\ &\quad + \ln \mathbb{E} \left[ \exp \left( i \left( \int_t^{\tau_1} e^{A(\tau_2-s)} \Sigma^{1/2}(s) d\widehat{W}(s) \right)^T w + i \left( \int_t^{\tau_1} e^{B(\tau_2-s)} \eta dL_Q(s) \right)^T w \right. \right. \\ &\quad \left. \left. + i \Psi(t, \tau_2, Z(\tau_1))^T w \right) | \mathcal{F}_t \right]. \end{aligned}$$

To this end, define  $\widehat{\mathcal{F}}(t, \tau_1, \tau_2)$ ,

$$\begin{aligned} \widehat{\mathcal{F}}(t, \tau_1, \tau_2) &\triangleq \ln \mathbb{E} \left[ \exp \left( i \left( \int_t^{\tau_1} e^{A(\tau_2-s)} \Sigma^{1/2}(s) d\widehat{W}(s) \right)^T w + i \left( \int_t^{\tau_1} e^{B(\tau_2-s)} \eta dL_Q(s) \right)^T w \right. \right. \\ &\quad \left. \left. + i \Psi(t, \tau_2, Z(\tau_1))^T w \right) | \mathcal{F}_t \right]. \end{aligned}$$

By a direct calculation one can conclude

$$\widehat{\mathcal{F}}(t, \tau_1, \tau_2) = \ln \mathbb{E} \left[ \exp \left( i \operatorname{tr} \left( \left( \frac{1}{2} w w^T \widehat{\mathcal{A}}(t, \tau_1, \tau_2) + \sum_{j=1}^d w_j \widehat{\mathcal{D}}_j(t, \tau_1, \tau_2) \right) Z(t) \right) \right) \right]$$

$$\begin{aligned}
& + \text{itr} \left( \int_t^{\tau_1} \left( \frac{1}{2} w w^T \widehat{\mathcal{E}}(s, \tau_1, \tau_2) + \sum_{j=1}^d w_j \widehat{\mathcal{D}}_j(s, \tau_1, \tau_2) + J_d(\eta^T e^{B^T(\tau_2-s)} w) \right) d\widehat{L}(s) \right) \Big], \\
& = \text{itr} \left( \left( \frac{1}{2} w w^T \widehat{\mathcal{E}}(t, \tau_1, \tau_2) + \sum_{j=1}^d w_j \widehat{\mathcal{D}}_j(t, \tau_1, \tau_2) \right) Z(t) \right) \\
& + \int_t^{\tau_1} \phi_{\widehat{L}}^Q \left( \left( \frac{1}{2} w w^T \widehat{\mathcal{E}}^T(s, \tau_1, \tau_2) + \sum_{j=1}^d w_j \widehat{\mathcal{D}}_j^T(s, \tau_1, \tau_2) + J_d(\eta^T e^{B^T(\tau_2-s)} w) \right) I_d \right).
\end{aligned}$$

So,

$$\begin{aligned}
\tilde{\phi}_{\ln \widehat{F}}^{t, \tau_1, \tau_2}(w) & = iX^T(t) e^{A^T(\tau_2-t)} w + iY^T(t) e^{B^T(\tau_2-t)} w \\
& + i \left( e^{A(\tau_2-\tau_1)} A^{-1} (I - e^{A(\tau_1-t)}) \theta_0 + e^{B(\tau_2-\tau_1)} B^{-1} (I - e^{B(\tau_1-t)}) \mu \right)^T w \\
& + \text{itr} \left( \left( \frac{1}{2} w w^T \widehat{\mathcal{E}}(t, \tau_1, \tau_2) + \sum_{j=1}^d w_j \widehat{\mathcal{D}}_j(t, \tau_1, \tau_2) \right) Z(t) \right) \\
& + \int_t^{\tau_1} \phi_{\widehat{L}}^Q \left( \left( \frac{1}{2} w w^T \widehat{\mathcal{E}}(s, \tau_1, \tau_2) + \sum_{j=1}^d w_j \widehat{\mathcal{D}}_j^T(s, \tau_1, \tau_2) + J_d(\eta^T e^{B^T(\tau_2-s)} w) \right) I_d \right).
\end{aligned}$$

□

## References

- [1] Baillie, R., and Bollerslev, T. (1989). Common stochastic trends in a system of exchange rates. *J. Finance*, **44**, pp. 137-151.
- [2] Barndorff-Nielsen, O. E., Pedersen, J., and Sato, K.I. (2001). Multivariate subordination, self-decomposability and stability. *Adv. Appl. Prob.*, **33(1)**, pp. 160-187.
- [3] Barndorff-Nielsen, O.E., and Pérez-Abreu, V.(2002). Extension of type  $G$  and marginal infinite divisibility. *Theory Probab. Appl.*, **47(2)**, pp. 202-218.
- [4] Barndorff-Nielsen, O.E., and Pérez-Abreu, V. (2008). Matrix subordinators and related Upsilon transformations. *Theory Probab. Appl.*, **52(1)**, pp. 1-23.
- [5] Barndorff-Nielsen, O., and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in economics. *J. R. Statist. Soc. B*, **63(2)**, pp. 167–241 (with discussion).
- [6] Barndorff-Nielsen, O., and Shephard, N. (2001). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, **72(3)**, pp. 885-925.
- [7] Barndorff-Nielsen, O., and Stelzer, R. (2007). Positive-definite matrix processes of finite variation. *Probab. Math. Statist.*, **27**, pp. 3-43.
- [8] Benth, F. E. (2009). The stochastic volatility model of Barndorff-Nielsen and Shephard in commodity markets. To appear in *Math. Finance*.
- [9] Benth, F. E., and Šaltytė-Benth, J. (2004). The normal inverse Gaussian distribution and spot price modeling in energy markets. *Intern. J. Theor. Appl. Finance*, **7(2)**, pp. 177–192.
- [10] Benth, F. E., Saltyte Benth, J., and Koekebakker, S. (2008). *Stochastic Modelling of Electricity and Related Markets*. World Scientific.



- [11] Benth, F. E., Klüppelberg, C., and Vos, L. (2009), The CARMA energy model, *Manuscript in preparation*.
- [12] Benth, F.E., Koekebakker, S. and Zakamouline, V.(2008). The CARMA interest rate model. *Manuscript in preparation*.
- [13] Burger, M., Klar, B., Müller, A., and Schindlmayer, G. (2004). A spot market model for pricing derivatives in electricity markets. *Quantit. Finance*, **4**, pp. 109–122.
- [14] Carr, P. and Madan, D. B. (1998). Option valuation using the fast Fourier transform. *J. Comp. Finance*, **2**, pp. 61–73.
- [15] Comte, F. (1995). Discrete and continuous time cointegration. *J. Econometrics*, **88**, pp. 207–226.
- [16] Deaton, A., and Leroque, G. (1992). On the behavior of commodity prices. *Rev. Financ. Econ. Stud.*, **59**, pp. 1–23.
- [17] Duan, J.C. and Pliska, S.R. (2004), Option valuation with co-integrated asset prices. *J. Econ. Dyn. Control*, **28**, pp. 727–754.
- [18] Engle, R., and Granger, C. (1987). Co-integration and error correction: representation, estimation and testing. *Econometrica*, **55**, pp. 251–276.
- [19] Eydeland, A., and Wolyniec, K. (2003). *Energy Power and Risk Management*. Wiley-Finance.
- [20] Duffie, D. (1992). *Dynamic Asset Pricing Theory*. Princeton University Press.
- [21] Folland, G. B. (1984). *Real Analysis – Modern Techniques and their Applications*. John Wiley & Sons.
- [22] Geman, H. (2005). *Commodities and Commodity Derivatives*. Wiley-Finance.
- [23] Hiksloops, S. and Jaimungal, S. (2008). Asymptotic pricing of commodity derivatives for stochastic volatility spot models. *Appl. Math. Finance*, **15**(5&6), pp. 449–467.
- [24] Horn, R. A., and Johnson, C.R. (1985). *Matrix analysis*. Cambridge University Press.
- [25] Hurd, T.R. and Zhou, Z. (2009). A Fourier transform method for spread option pricing, working paper available on <http://www.math.mcmaster.ca/tom/tom.html>.
- [26] Jacod, J. and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer-Verlag.
- [27] Kessler, M., and Rahbeck, A. (2001), Asymptotic likelihood based inference for co-integrated homogenous Gaussian diffusions. *Scand. J. Stat.*, **28**, pp, 455–470.
- [28] Kessler, M., and Rahbeck, A. (2004), Identification and inference for multivariate cointegrated and ergodic Gaussian diffusions. *Stat. Infer. Stoch. Proc.*, **7**, pp, 137–151.
- [29] Marquardt, T., and Stelzer, R. (2006). Multivariate CARMA processes. *Stoch. Proc. Appl.*, **117**, pp. 96–120.
- [30] Nelson, D.B. (1990). ARCH models as diffusion approximations. *J Econometrics*, **45**, pp. 7–38.
- [31] Nicolato, E., and Venardos, E. (2003). Option pricing in stochastic volatility models of the Ornstein-Uhlenbeck type. *Math. Finance*, **13**(4), pp. 445–466.
- [32] Pigorsch, C., and Stelzer, R.(2008). A multivariate generalization of the Ornstein-Uhlenbeck stochastic volatility model. *E-print*, [www-m4.ma.tum.de](http://www-m4.ma.tum.de).
- [33] Pigorsch, C., and Stelzer, R.(2008). On the definition, stationarity and second order structure of positive semi-definite Ornstein-Uhlenbeck type processes. To appear in *Bernoulli*.

- [34] Protter, Ph. (1990). *Stochastic Integration and Differential Equations*. Springer-Verlag.
- [35] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
- [36] Sato, K.-I. (2004). Stochastic integrals in additive processes and application to semi-Lévy processes. *Osaka J.Math*, **41**, pp. 211–236.
- [37] Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *J. Finance*, **LII**, **3**, pp. 923–973.
- [38] Schwartz, E.S., and Trolle, A.B. (2006). Unspanned stochastic volatility and the pricing of commodity derivatives. NBER Working Paper No. 12744 available on <http://www.nber.org/papers/w12744.pdf>.
- [39] Shiryaev, A. N. (1999). *Essentials of Stochastic Finance*. World Scientific.
- [40] Stelzer, R (2007). Multivariate continuous time stochastic volatility models driven by a Lévy process. Dissertation, TU Munchen, available on <http://www-m4.ma.tum.de/pers/stelzer/Publications.html>.
- [41] Tankov, P. Simulation and option pricing in Lévy copula model. Working paper available on <http://people.math.jussieu.fr/tankov/>.
- [42] Vos, L. (2009). Asian options and the effect of a non-Gaussian stochastic volatility model, *Manuscript in preperation*.
- [43] Williams, D (1991). *Probability with martingales*. Cambridge university press.