STOCHASTIC VOLATILITY AND MULTI-DIMENSIONAL MODELING IN THE EUROPEAN ENERGY MARKET

LINDA VOS

Thesis submitted for the degree of Ph.d.
Department of mathematics
University of Oslo
2012
Acknowledgement

During my Ph.d. I was given the opportunity to work together with inspiring and highly gifted people. I am grateful for the useful instructions of my supervisor Fred Espen Benth. His guidance was of vital importance for the establishment of this Ph.d. Fred has continuously supported me throughout my Ph.d. also when I faced hard times. I was given the chance to work together with Claudia Klüppelberg. She showed me how to see some things differently. I also would like to thank my other co-author Gernot Müller and the inspiring chats with Robert Stelzer and Peter Hepperger.

I’m grateful for the financial support given to me by the Competence Development Fund of Southern Norway and the Center of Mathematics for Applications (CMA). Moreover I would like to thank CMA for providing good working conditions.

Without the support of my family I would never have managed to finish my Ph.d. My boyfriend, father, brother and sister, uncles and aunts, have been of great support. Especially my mom never stopped believing in me and made any trouble visible as something positive. Seeing this book would have made her unbelievably proud and was something she was keeping on living for.

Linda Vos, Oslo, June 2012
Preface

This thesis is submitted as partial fulfillment of the degree of *philosophiae doctor* (Ph.D.) at the University of Oslo. The work has been performed during my Ph.D. scholarship lasting from August 2008 to August 2011. Most of the work is carried out during my main period in the Center of Mathematics for Applications (CMA) at the University of Oslo under the guidance of my main supervisor Fred Espen Benth. However part of the work has been carried out during 2 visits, to the university of Agder, Kristiansand and the center of advanced studies at the Technical University of Munich. During this stays I got the opportunity to work with Steen Koekebakker (University of Agder) and Claudia Klüppelberg (Technical university of Munich).

This thesis consists of four papers, all of them are submitted. I wrote one of them alone and contributed substantially to the other ones.
## 4 Pricing of forwards and options in a multivariate non-Gaussian stochastic volatility model for energy markets

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>67</td>
</tr>
<tr>
<td>2 A cross-commodity energy spot price model with stochastic volatility</td>
<td>68</td>
</tr>
<tr>
<td>3 Forward pricing</td>
<td>71</td>
</tr>
<tr>
<td>3.1 A class of equivalent probabilities</td>
<td>72</td>
</tr>
<tr>
<td>3.2 Analysis of forward prices</td>
<td>76</td>
</tr>
<tr>
<td>3.3 Shapes of the forward curve</td>
<td>83</td>
</tr>
<tr>
<td>4 Transform-based pricing of options</td>
<td>85</td>
</tr>
<tr>
<td>5 Conclusions</td>
<td>92</td>
</tr>
</tbody>
</table>

## 5 Futures pricing in electricity markets based on stable CARMA spot models

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>97</td>
</tr>
<tr>
<td>2 The spot price dynamics</td>
<td>101</td>
</tr>
<tr>
<td>2.1 The stable CARMA-process</td>
<td>101</td>
</tr>
<tr>
<td>2.2 Dimensionality of CARMA-process</td>
<td>103</td>
</tr>
<tr>
<td>3 The futures price dynamics</td>
<td>104</td>
</tr>
<tr>
<td>3.1 Equivalent measure transforms for Lévy and α-stable processes</td>
<td>107</td>
</tr>
<tr>
<td>4 Fitting the model to German electricity data</td>
<td>109</td>
</tr>
<tr>
<td>4.1 Seasonality function Λ</td>
<td>109</td>
</tr>
<tr>
<td>4.2 Filtering the realization of the non-stationary stochastic process Z</td>
<td>113</td>
</tr>
<tr>
<td>4.3 Estimation of the CARMA parameters</td>
<td>114</td>
</tr>
<tr>
<td>4.4 Estimation of the stable parameters</td>
<td>118</td>
</tr>
<tr>
<td>4.5 Recovering the states</td>
<td>118</td>
</tr>
<tr>
<td>4.6 Risk premium comparison</td>
<td>120</td>
</tr>
<tr>
<td>5 Estimation results</td>
<td>123</td>
</tr>
<tr>
<td>5.1 Distributional properties of the filtered sample path of Z</td>
<td>123</td>
</tr>
<tr>
<td>5.2 Estimation of the CARMA parameters</td>
<td>125</td>
</tr>
<tr>
<td>5.3 Market price of risk and risk premium</td>
<td>126</td>
</tr>
<tr>
<td>6 Conclusion</td>
<td>129</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Since the 1 of July 2004 the energy market in Europe is officially liberalized. In the European Union the liberalization of power markets has been driven by the directive 96/92/EC of the European Parliament and Council. The directive is aimed at opening up the member states’ electricity markets, in order to increase the number of generators and consumers which can negotiate the purchase and sale of electricity. Since then, a large number of electricity exchanges have opened in Europe. England and Wales were the first to open an electricity exchange in 1990. After that Norway followed in 1993 with Statnett Marked (Nasdaq OMX commodities) [19]. The English and the Norwegian markets were already rather reasonably liberalized. The gas and electricity markets are owned by the private sector now in England and Norway. The markets in other countries like France, Italy, Netherlands, Finland and Sweden were owned by companies who had a government monopoly [14]. However, at the end of the 90-ties the Netherlands as well as Finland and Sweden got a liberalized market. The southern European countries took a bit more time before being liberalized.

The liberalization of energy markets gave rise to many regulation issues. Still regulations are changed in order to get a better functioning market. How to regulate the market is a separate field of studies, which will not be considered in this thesis. More interesting for a mathematician is that the energy exchanges give new financial data where lot of things are still unclear and interesting studies can be done. In order to model the energy market new theories can be developed. We will focus on the stochastic modeling of the electricity market.

1.1 Spot market

In this section we will give a short overview of common stochastic models used in order to model the electricity spot market. This will be used to put the papers of which this Ph.d. thesis consists in a context. The energy market distinguishes itself from other commodity markets by the non-storability of the products electricity and gas. This affects how the market is regulated and what
Chapter 1. Introduction

kind of products are sold.

By the energy spot market actually the one-day ahead market is meant. Contracts are sold for delivering 1 MWh (mega Watt hour) electricity during one hour for the different hours of the next day, as well as for 1 MWh electricity for blocks of hours during the next day. In a sense these contracts are forward contracts, since the spot contracts are about electricity which will be delivered in the future (one-day head). The price of these one-day ahead contracts is determined by supply and demand during an auction at noon. Here the difference is with an ordinary forward contract, since that can be traded till delivery starts. The contracts that are sold on the one day-ahead market can be traded till delivery starts on the real-time market. Different real-time markets exists throughout Europe (for instance Elbas (Nordic countries), EEX intraday (Germany)).

What normally is denoted by the base spot price of an electricity market is an index, which is given by the average of all the hourly prices. Or in the case of the peak spot price the average of the hourly prices in the peak hours (between 8 a.m. and 8 p.m). The peak hours are the hours when there is naturally more demand for energy.

Remarkably is that the spot itself is not traded (it is an average of products which are traded). This leads to the possibility of structure in the spot price data. In a way the energy market is comparable to the interest rate market. Also the short-rate is not traded in itself. Moreover the economical relationship between the short rate and the bond-market is comparable to the relationship between the energy spot and the forward contracts. This is why many bond market models are used in the modeling of the energy market. However some adjustments have to be made. Typical features of the energy market are extreme spikes, seasonal behavior and stochastic volatility. These have to be accounted for in the models.

The spikes occur when there are misbalances between supply and demand of energy. Since energy can only be stored to a limited extend, the demand and supply have to match approximately. Both the demand and the supply sides are inelastic. On the supply side it is impossible to turn on a power plant on short notice. Depending on the kind of power plant this can take days (nuclear) or hours (gas, benzine). On the demand side the customer is not aware of the market structure. Most customers have a contract with an energy company and pay one pre-defined price for electricity. They are not aware that the price can change from hour to hour. Moreover electricity is often needed to keep a business running and it can be more expensive to slow down an industry then to pay more for electricity. This in-elasticity on both sides has lead to the possibility of enormous price spikes.

A second feature of the energy market is its seasonal behavior. In many countries there is naturally more demand for energy during the winter than during the summer, because of heating. Therefore the price of energy is often higher during winter than during summer. In some states with a warmer climate, like in California, it is the other way around. Here the demand for energy in order to heat during the winter is not so high. However during the summer air-conditioning demand, lots of energy. To capture this seasonality a seasonal function is introduced in energy
A last feature we will focus on in this thesis is stochastic volatility. Empirical studies by Trolle and Schwartz [16] confirmed stochastic volatility in energy markets. This means that the volatility is changing stochastically over time. Moreover there is evidence of a so-called inverse leverage effect. The volatility tends to increase with the level of power prices, because there is a negative relationship between inventory and prices (see for instance Deaton and Leroque [10]). In order to model stochastic volatility we have chosen to follow the approach of Barndorff-Nielsen and Shephard [1]. The stochastic volatility (SV) model of Barndorff-Nielsen and Shephard is statistically tractable and fits well on the existing literature of energy modeling. In Chapter 2 in this thesis the effect of stochastic volatility in pricing path dependent options is investigated. This is done using the Barndorff-Nielsen and Shephards stochastic volatility model (BNS SV model). Here ordinary stock data is used to estimate the parameters. However similar results are expected when energy data would have been used to estimate the parameters and instead of Asian options the valuation of an electricity line or gas pipe would have been considered.

One of the most common models is to model the spot as a combination of an Ornstein-Uhlenbeck processes and a seasonal function. This can either be done arithmetically or geometrically. A common geometric model is given by

$$S(t) = \Lambda(t) \cdot \exp(X(t) + \sum_i Y_i(t)) \quad \quad (1.1.1)$$
Chapter 1. Introduction

here Λ is a seasonality function. The normal variation is given by $X$, an Ornstein-Uhlenbeck (OU) process with the following stochastic differential equation (SDE)

$$dX(t) = aX(t)\, dt + \sigma \, dW(t)$$

here $a < 0$ and $\sigma > 0$ are constants and $W$ is a Brownian motion. The $Y_i$’s are spike processes modeling the extreme spikes in the electricity market. The $Y_i$’s are Ornstein-Uhlenbeck processes with the following SDE’s.

$$dY_i(t) = b_i Y_i(t)\, dt + \eta_i \, dL_i(t)$$

here $b_i < 0$ and $\eta_i > 0$ are constants. The $L_i$’s are Lévy processes possibly from different distributions.

Or similarly one can put the same variables in an arithmetic model

$$S(t) = \Lambda(t) + X(t) + \sum_i Y_i(t) \quad (1.1.2)$$

There is a discussion going on whether energy spot prices should be modeled geometrically or arithmetically. In financial modeling often geometric models are chosen (as this would be a natural choice). However, working with arithmetic models is computationally easier. Moreover in the energy market also negative prices are observed. For instance, the German power exchange EEX permitted negative price outcomes for the spot auction from autumn 2008. Since then negative prices have occurred. Sometimes it is more expensive to shut down a power plant then to pay somebody to use the electricity. These are of course rare situations. Negative prices can not be modeled by a geometric model. Therefor some argue that an arithmetic model should be preferred. Moreover Lucia and Schwartz [15] have found statistical evidence in the Nordpool market that arithmetic models do a better job in explaining prices then geometric models.

In chapter 3 and 5 in this thesis variations on this model are treated. In chapter 3 a multi-dimensional version of the geometric model (1.1.2) is introduced. Here, instead of static volatility, a stochastic volatility process is chosen. In order to model the normal variation $X$ a multi-dimensional version of the BNS-SV model (Barndorff-Nielsen and Stelzer [2]) was taken.

In chapter 5 the arithmetic model (1.1.1) is chosen with the possibility of a higher auto-regressive order. Instead of an OU-process a continuous auto-regressive moving average (CARMA) process (Brockwell [9]) is chosen. This is a continuous version of a auto-regressive moving average (ARMA) process. The OU-process is a special case of a CARMA process. As a driver of the CARMA process a stable process (Samorodnitsky and Taqqu [18]) is chosen. The Brownian motion is a special case of a stable process. Stable processes are known for their heavy tails and are therefore suitable to model extreme events like price spikes. When working with stable processes it is unnecessary to add several spike processes since one stable process can capture
1.2 Forward market

Unlike more classical commodity markets like agriculture and metals, energy-related futures contracts deliver the underlying spot over a contracted period. There are for example contracts sold which deliver 1 MWh during every hour of the month in the future. Also year contracts and quarter of a year contracts are sold. In some markets also week contracts and 3 year contracts are sold. Contracts can be settled either physically or financially. Forwards are traded continuously till settlement starts.

Different to other commodity markets delivery is done over a period instead of at one point in time. This is due to the difficulties of energy storage. Mathematically this lead to some differences. From the theory of mathematical finance (Duffie [11]), we know that the value of any derivative is given as the present expected value of its payoff, where the expectation is taken with respect to a risk-neutral probability \( Q \). It holds that

\[
F(t, T_1, T_2) = \mathbb{E}_Q \left[ \int_{T_1}^{T_2} w(u, T_1, T_2) S(u) du \mid \mathcal{F}_t \right]
\]

(1.2.1)

here \( Q \) is a risk neutral measure, \( t \) is the time, \( T_1 \) is the beginning of the delivering period and \( T_2 \) is the end of the delivering period. \( w \) is a weight function, which should be chosen according to the market structure. In electricity markets the swap price \( F \) is normally denoted by price per MWh. Therefore \( w(u, T_1, T_2) \) is taken \( \frac{1}{T_2 - T_1} \). However when settlement is done continuously over the delivering period

\[
w(u, T_1, T_2) = \frac{e^{-ru}}{\int_{T_1}^{T_2} e^{-rv} dv}
\]

(1.2.2)

here \( r \) is the interest rate. This to adjust for the advantage of having money at the bank.

In geometric models it is often not possible to calculate the swap \( F \) explicitly using equation (1.2.1). However it is possible to given an expression in the form of an integral which can be approximated numerically. From no arbitrage theory we know that if the forward would have been delivered at one point in time \( \tau \) the forward price \( f \) would have been given by

\[
f(t, \tau) = \mathbb{E}_Q [S(\tau) \mid \mathcal{F}_t]
\]

(1.2.3)

meaning that the forward price is the best risk-neutral prediction at time \( t \) of the spot price \( S(\tau) \) at delivery. The swap \( F \) may be viewed as a continuous flow of forwards \( f \) (see Prop. 4.1 in
Chapter 1. Introduction

Benth, Saltyte Benth and Koekebakker [3]).

\[ F(t, T_1, T_2) = \int_{T_1}^{T_2} w(\tau, T_1, T_2) f(t, \tau) d\tau \quad (1.2.4) \]

In chapter 4 the derivation (1.2.3) is done for the model introduced in Benth and Vos [6]. Moreover some properties of the derived forward curves are analyzed. Examples of possible shapes of the forward curve using this model are given and transform based methods to price options on the forward curves are given. In chapter 5 the derivation (1.2.1) is done for the model introduced in García, Klüppelberg and Müller [12]. Furthermore using parameters estimated on the spot price the risk premium is empirically analyzed.

The risk premium is defined as the difference between the futures price and the predicted spot, that is, in terms of electricity futures contracts,

\[ R_{pr}(t, T_1, T_2) = F(t, T_1, T_2) - \mathbb{E} \left[ \int_{T_1}^{T_2} w(\tau, T_1, T_2) S(\tau) d\tau \mid \mathcal{F}_t \right]. \quad (1.2.5) \]

The risk premium can be used to identify the risk-measure \( Q \). In chapter 5 is found that the risk premium is negative in the long end of the market and positive in the short end of the market. The positive risk premium for contracts close to delivery tells us that the demand side (retailers and consumers) of the market is willing to pay a premium for locking in electricity prices as a hedge against spike risk (see Geman and Vasicek [13]). In the long end of the market the risk premium is negative meaning that on the supply side the energy companies want to hedge their risk and willing to accept a lower price. The energy companies hedge the risk of an uncertain spot price.
Bibliography


Chapter 2

Path dependent options and the effect of stochastic volatility

Linda Vos
Abstract

In modern asset price models, stochastic volatility plays a crucial role explaining several stylized facts of returns. Recently, Barndorff-Nielsen and Shephard [4] introduced a class of stochastic volatility models (the so called BNS SV model) based on superposition of Ornstein-Uhlenbeck processes driven by subordinators. The BNS SV model forms a flexible class, which can easily explain heavy-tails and skewness in returns and the typical time-dependency structures seen in asset return data. In this paper the effect of stochastic volatility on path-dependent options is studied. This is done by simulation studies of comparable models, one with and one without stochastic volatility.

1 Introduction

Lévy processes are popular models for stock price behavior since they allow to take into account jump risk and reproduce the implied volatility smile. Barndorff-Nielsen and Shephard [4] introduced a class of stochastic volatility models (BNS SV model) based on superposition of Ornstein-Uhlenbeck processes driven by subordinators (Lévy processes with only positive jumps and no drift). The distribution of these subordinators will be chosen such that the log-returns of asset prices will be distributed approximately normal inverse Gaussian (NIG) in stationarity. This family of distributions has proven to fit the semi-heavy tails observed in financial time series of various kinds extremely well (see Rydberg [18], or Eberlein and Keller [10]).

We will compare the BNS SV model to a NIG Lévy process model, which has NIG distributed log-returns of asset prices, with the same parameters as in the BNS SV case. Hence in stationarity the BNS-SV model and the NIG model have equal distributed increments. Unlike the BNS SV model, the NIG Lévy process model does not have time-dependency of asset return data. Both models are described and the difference in pricing Asian and barrier options with the two different models will be studied. This is done by a case study with calibrated parameters on stock data of the Amsterdam-Stock Exchange Index (AEX). We chose Asian and barrier options because they are path dependent options. The time-dependency of the asset return data in the BNS SV model leads to a difference in pricing. This difference is visible in the distribution of the prices.

Unlike the Black-Scholes model, closed option pricing formulae are in general not available in exponential Lévy models and one must use either deterministic numerical methods (see e.g. Carr and Madan [9] for the Lévy process model and Benth and Groth [8] for the BNS SV model) or Monte Carlo methods. In this paper we will restrict ourselves to Monte Carlo methods. As described in Benth, Groth and Kettler [7] an efficient way of simulating a NIG Lévy process is by a quasi-Monte Carlo method. We will use a simpler Monte-Carlo method, which needs bigger sample size to reduce the error. Simulating from the BNS SV model involves simulation of an Inverse Gaussian Ornstein-Uhlenbeck (IG-OU) process. The oldest algorithm of simulating a IG-OU process is described in Barndorff-Nielsen and Shephard [4]. This is a quite bothersome
algorithm, since it includes a numerical inversion of the Lévy measure of the Background driving
Lévy process (BDLP). Therefore it has a long processing time, hence we will not deal with this
algorithm.

The most popular algorithm is a random series representation by Rosinski [17]. The special
case of the IG-OU process is described in Barndorff-Nielsen and Shephard [6]. Recently Zhang
& Zhang [21] introduced an exact simulation method of an IG-OU process, using the rejection
method. This algorithm needs less processing time than a series representation. Therefore we
will use it.

2 The models

To price derivative securities, it is crucial to have a good model of the probability distribution of
the underlying asset. The most famous continuous-time model is the celebrated Black-Scholes
model (also called geometric Brownian motion), which uses the Normal distribution to fit the
log-returns of the underlying asset. However the log-returns of most financial assets do not
follow a Normal law. They are skewed and have actual kurtosis higher than that of the Normal
distribution. Other, more flexible distributions are needed. Moreover, to model the behavior
through time we need more flexible stochastic processes. Lévy processes have proven to be good
candidates, since they preserve the property of having stationary and independent increments and
they are able to represent skewness and excessive kurtosis. In this section we will describe the
Lévy process models we use. We also give a method how to fit the models to historical data.

In this text we will work on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) equipped with a filtration \(\{\mathcal{F}_t\}_{t \in [0,T]}\)
satisfying the usual conditions\(^1\), with \(T < \infty\) being the time horizon.

2.1 Distributions

The inverse Gaussian distribution IG\((\delta, \gamma)\) is a distribution in \(\mathbb{R}_+\) given in terms of its density,

\[
f_{\text{IG}}(x; \delta, \gamma) = \frac{\delta}{\sqrt{2\pi}} \exp(\delta \gamma x^{-1}) x^{-\frac{3}{2}} \exp\left\{-\frac{1}{2} \left(\delta^2 x^{-1} + \gamma^2 x\right)\right\}, \quad x > 0
\]

where the parameters \(\delta\) and \(\gamma\) satisfy \(\delta > 0\) and \(\gamma \geq 0\). The IG distribution is infinitely divisible
and self-decomposable. The associated Lévy process is a jump process of finite variation.

The Normal inverse Gaussian (NIG) distribution has values on \(\mathbb{R}\), it is defined by its density

\(^1\)See e.g. Protter [15]
function,

\[ f_{\text{NIG}}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha}{\pi} \exp \left( \delta \sqrt{\alpha^2 - \beta^2} + \beta(x - \mu) \right) K_1 \left( \frac{\delta \alpha}{\sqrt{1 + \left( \frac{x-\mu}{\delta} \right)^2}} \right) \]

where \( K_1 \) is the modified Bessel function of the third kind and index 1. Moreover the parameters are such that \( \mu \in \mathbb{R}, \delta \in \mathbb{R}_+ \) and \( 0 \leq \beta < |\alpha| \). The NIG distribution is infinitely divisible and there exists a NIG Lévy process.

The NIG distribution can be written as a mean-variance mixture (see Barndorff-Nielsen [1]). More specifically, if we take \( \sigma^2 \sim \text{IG}(\delta, \sqrt{\alpha^2 - \beta^2}) \) independently distributed of \( \epsilon \sim \mathcal{N}(0,1) \), then \( x = \mu + \beta \sigma^2 + \sigma \epsilon \) is distributed \( \text{NIG}(\alpha, \beta, \delta, \mu) \). See e.g. Barndorff-Nielsen [2] for an extensive description of the above distributions.

### 2.2 Exponential Lévy process model

As model without stochastic volatility we will use a fitted NIG Lévy process to the log-returns. i.e.

\[ d \log S(t) = dX(t), \tag{2.1} \]

where \( S \) is an arbitrary stock-price and \( X \) is a NIG Lévy process. This model is flexible to model with, but a drawback is that the returns are assumed independently. In the literature this model has also been referred to as exponential Lévy process model, since the solution is of the form,

\[ S(t) = S(0)e^{X(t)}. \]

### 2.3 Barndorf-Nielsen and Shephard stochastic volatility model

In practice there is evidence for volatility clusters in log return data. There seems to be a succession of periods with high return variance and with low variance. This means that large price variations are more likely to be followed by large price variations. This observation motivates the introduction of a model where the volatility itself is stochastic.

The Barndorf-Nielsen and Shephard stochastic volatility (BNS SV) model is an extension of the Black-Scholes model, where the volatility follows an Ornstein-Uhlenbeck (OU) process driven by a subordinator.

In the Black-Scholes model the asset price process \( \{S(t)\}_{t \geq 0} \) is known as a solution to the SDE,

\[ d \log S(t) = \{ \mu + \beta \sigma^2 \} dt + \sigma dB(t) \tag{2.2} \]
where an unusual drift is chosen to be in line with the BNS SV model and \( B(t) \) is a standard Brownian motion. The BNS SV model allows \( \sigma^2 \) to be stochastic. More precisely \( \sigma^2(t) \) is an OU process or a superposition of OU processes i.e.

\[
\sigma^2(t) = \sum_{j=1}^{m} a_j \sigma_j^2(t)
\]

with the weights \( a_j \) all positive, summing up to one and where the processes \( \{\sigma_j^2(t)\}_{t \geq 0} \) are OU processes satisfying,

\[
d\sigma_j^2(t) = -\lambda_j \sigma_j^2(t) dt + dZ_j(\lambda_j t)
\]

(2.3)

where the processes \( Z_j \) are independent subordinators i.e. Lévy processes with positive increments and no drift. Since \( Z_j \) is a subordinator the process \( \sigma_j^2 \) will jump up according to \( Z_j \) and decay exponentially afterwards with a positive rate \( \lambda_j \). As \( Z_j \) is used to drive the OU process, we shall call \( Z_j(t) \) a background driving Lévy process (BDLP).

We can rewrite equation (2.2) into,

\[
d \log S(t) = \left\{ \mu + \beta \sigma^2(t) \right\} dt + \sigma(t) dB(t) := dX(t).
\]

(2.4)

We can reformulate this into,

\[
\log S(t) = \log S(0) + \mu t + \beta \sigma^{2*}(t) + \int_{0}^{t} \sigma(s) dB(s)
\]

where \( \sigma^{2*}(t) \) is the integrated process, i.e.

\[
\sigma^{2*}(t) = \int_{0}^{t} \sigma^2(u) du.
\]

Barndorff-Nielsen and Shephard [5] showed that the tail behavior (or superposition) of an integrated IG process is approximately Inverse Gaussian. Moreover an integrated superposition of IG processes will show similar behavior as just one integrated IG process with the same mean. For \( \sigma^2 \) an IG(\( \delta, \gamma, \lambda \))-OU process, \( \sigma^2(t) \) has stationary marginal law IG(\( \delta, \gamma \)). Hence by taking the \( \sigma^2(t) \) an inverse Gaussian Ornstein-Uhlenbeck (IG-OU) process one gets the same tail behavior as taking \( \sigma^2(t) \) a superposition of IG-OU processes. So the log returns will be approximately NIG distributed, since the NIG is a mean-variance mixture.
2.4 Parameter estimation:

The stationary distribution of $\sigma_j^2$ is independent of $\lambda_j$ (see Barndorff-Nielsen and Shephard [3], Th. 6.1). This allows us to calibrate the log return distribution to the data separately of the $\lambda_j$’s.

In order to get comparable models we will assume that both models have NIG distributed returns with the same parameters. The parameters will be estimated with a maximum likelihood estimation. The maximum-likelihood estimator $(\hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu})$ is the parameter set that maximizes the likelihood function

$$L(\alpha, \beta, \delta, \mu) = \prod_{i=1}^{n} f_{NIG}(x_i; \alpha, \beta, \delta, \mu)$$

where $x_i$ are points of our data-set. Since the NIG distribution is a mean-variance mixture we can choose the parameters of the IG-OU process such that returns will be approximately NIG distributed.

To get reasonable estimates for our simulation study we estimate the models on real stock data. We have chosen a time-series from 12 March 2005 to 7 October 2008 of the Amsterdam stock exchange index (AEX) (see Figure 2.1).

Using maximum likelihood estimation the parameters are given by

$$\alpha = 94.1797 \quad \beta = -16.0141 \quad \delta = 0.0086 \quad \mu = 0.0017$$
Chapter 2. Path dependent options and the effect of stochastic volatility

Figure 2.2: Least square fitting to the auto-covariance function (black) and the empirical auto-covariance (grey).

Least square estimation

In Barndorff-Nielsen and Shephard [4] the analytic auto-covariance function of the squared log returns is calculated and is found to be given by,

$$ \text{cov}(y^2_i, y^2_{i+s}) = \omega^2 \sum_{j=1}^{m} a_j \lambda_j^{-2} (1 - e^{-\lambda_j \Delta t})^2 e^{-\lambda_j \Delta t (s-1)} $$

(2.5)

where $\omega^2$ is equal to the variance of $\sigma^2(t)$ and $a_j$’s are the weights of the different OU processes. Looking at the empirical auto-covariance function of the squared log-returns of financial data it seems that this function fits the data really well (see Figure 2.2). One can see that the empirical auto-covariance function of real financial data looks like a sum of exponentially decaying functions.

We estimated the empirical auto-covariance function $\gamma(s)$ for a data-set $x_1, \ldots, x_n$ by,

$$ \gamma(s) = \frac{1}{n-s} \sum_{i=1}^{n-s} \left( x^2_{i+s} - \frac{1}{n-s} \sum_{k=s}^{n} x^2_k \right) \left( x^2_i - \frac{1}{n-s} \sum_{k=1}^{n-s} x^2_k \right) $$

For the calibration we will use a non-linear least square comparison of the empirical auto-covariance function with the analytic auto-covariance function (2.5). Hence we will minimize,

$$ \sum_s \left\{ \gamma(s) - \text{cov}(y^2_n, y^2_{n+s}) \right\}^2 $$
For a non-linear least square comparison several algorithms are available. We used a standard function of Matlab which is based on the Gauss-Newton method.

Note that the least square comparison is only done on the $\lambda_j$’s and $a_j$’s since $\omega$ is already given from the maximum likelihood estimation.

For our data it is sufficient to let the volatility process $\sigma^2$ exist out of one OU-process. Hence we chose $m = 1$. Using the least-squares estimation procedure on our AEX data we find that $\lambda$ is given by 0.0146.

## 3 Pricing

### 3.1 Asian option

An Asian option is an option where the payoff is not determined by the underlying price at maturity but by the average underlying price over some pre-set period of time. We will use Asian option because it is a path dependent option, where one pretends to see a difference in pricing between our two models.

Let $X(t)$ be a process with càdlàg sample paths, given as in equation (2.1) or (2.4). Consider the following exponential model for the asset price dynamics,

$$S(t) = S(0)e^{X(t)}$$

Now the aim is to price arithmetic Asian call options written on $S(t)$. Consider such an option with exercise time at $T$ and strike price $K$ on the average over the time span up to $T$. The risk-neutral price is (See e.g. [13]),

$$A(0) = e^{-rT}\mathbb{E}_Q\left[\max\left\{\frac{1}{T}\int_0^T S(t)dt - K, 0\right\}\right]$$

Here $r$ is the risk-free interest rate and $Q$ is an equivalent martingale measure (EMM). Now by applying the below described change of measure and approximating the integral with a Riemann sum,

$$A(0) \approx e^{-rT}\mathbb{E}\left[\max\left\{\frac{S(0)}{N}\sum_{i=0}^{N} e^{X(t_i)} \Delta - K, 0\right\}\right]$$

For simplicity we will work with a regular time partition $0 = t_0 < t_1 < \cdots < t_N = T$ with mesh $\Delta$. In the next section we will describe several methods to valuate the expectation in equation (3.2).
3.2 Barrier option

Just as an Asian option a barrier option is a path-dependent option. We will consider a European call barrier option. This option acts as a Vanilla European call option unless the asset price $S$ crosses a barrier $B$ in which case it knocks out and the value of the option is zero. The value of the barrier option stays zero. When the asset price is below the barrier at maturity the value of the barrier option is still zero. In pricing a barrier option we will use the same equivalent martingale measure transforms as described below.

3.3 Equivalent martingale measure

Unfortunately, as in most realistic models, there is no unique equivalent martingale measure; the proposed Lévy models yield an incomplete market. This means there exists an infinite number of EMM’s for both models, hence there is no unique arbitrage free price. In the case of the exponential Lévy process model with NIG distributed increments, the Esscher transform (Shirayev [20]) would be the most natural measure transform. However an Esscher transform for the BNS-SV model is not structure preserving, meaning that under $Q$ the distribution of the returns does not remain in the class of BNS-SV distributed random variables. In order to keep the models comparable, we will use a similar measure transform for both of the models. We will follow the measure transform as described by Nicolato and Vernardos [14].

By Girsanov’s theorem there exists an $\psi$ such that the density process $L(t) = \frac{dQ}{dP}|_{\mathcal{F}_t}$ can be represented as

$$L(t) = \mathcal{E}(\psi \cdot B)$$

where $\mathcal{E}$ is the Doléans-Dade exponential. The process $B^Q$ defined by

$$B^Q(t) = B(t) - \int_0^t \psi(s) ds$$

is a $Q$ Brownian motion. By Ito’s formula it holds that

$$dS(t) = S(t) \left( \left\{ \mu + (\beta + \frac{1}{2})\sigma^2(t) \right\} dt + \sigma(t) dB(t) \right)$$

$$= S(t) \left( \left\{ \mu + (\beta + \frac{1}{2})\sigma^2(t) + \psi(t) \sigma(t) \right\} dt + \sigma(t) dB^Q(t) \right)$$

$\psi$ will be chosen such that the process $e^{-rt}S(t)$ is a martingale under $Q$. Hence

$$\psi(t) = \sigma(t)^{-1} \left( r - \mu - (\beta + \frac{1}{2})\sigma(t)^2 \right)$$
and the dynamics of $S$ in the risk-neutral world $Q$ are given by

$$dS(t) = S(t) \left( r dt + \sigma(t) dQ(t) \right)$$

Using Ito’s formula one can conclude that the dynamics of the logarithmic spot process $X(t) = \log(S(t))$ in the risk-neutral world $Q$ is given by

$$dX(t) = \{ r - \frac{1}{2} \sigma(t)^2 \} dt + \sigma(t) dB^Q(t)$$

The difference between the BNS-SV model and NIG exponential Lévy process model lies in the definition of the volatility process $\sigma(t)$. In the BNS-SV model case $\sigma(t)$ is determined by a (sum of) IG-OU processes where the BDLP is independent of the Brownian motion driving the spot. In the case of a NIG exponential Lévy process model the volatility process $\sigma(t)$ is modeled by an inverse-Gaussian process, which is independent of the Brownian motion which is driving the spot. Due to this independence the measure change of the volatility process $\sigma(t)$ can be done independently. Moreover the martingale property is satisfied regardless of the measure transform of the volatility processes, therefore an equivalent measure transform of the volatility process or the BDLP of the volatility process is sufficient.

Transforming is done by functions $y : \mathbb{R}^+ \mapsto \mathbb{R}^+$ satisfying the following integral condition

$$\int_0^\infty \left( \sqrt{y(x)} - 1 \right)^2 u(x) dx < \infty$$

here $u$ is the Lévy density of the BDLP (in case of BNS-SV model) or the IG-distribution (in case of the exponential Lévy process model). Define the equivalent Lévy density in $Q$ as

$$u^Q(x) = y(x) u(x), \quad x \in \mathbb{R}^+,$$

Moreover we will choose $y$ such that $u^Q$ is a Lévy density in the same class of distributions as the Lévy density $u$. Taking into account the integral condition (3.3) we have the following restriction for $\gamma$, the parameter of the inverse Gaussian distribution.

**Example 1** (Exponential Lévy model). The Lévy density $u$ of the inverse Gaussian distribution is given by

$$u(x) = \frac{\delta}{\sqrt{2\pi}} x^{-\frac{3}{2}} e^{-\frac{1}{2} \gamma^2 x^2} 1_{x > 0}$$

the collection of function $y$ preserving the measure in the Lévy density is given by

$$\{ y : \mathbb{R}^+ \mapsto \mathbb{R}^+ | y(x) = e^{-\frac{1}{2}(\tilde{\gamma}^2 - \gamma^2) x}, \quad \tilde{\gamma} \in \mathbb{R}^+ \}$$
Chapter 2. Path dependent options and the effect of stochastic volatility

Example 2 (BNS-SV model). The Lévy density $u$ of the BDLP of an IG-OU process is given by

$$u(x) = \frac{\delta}{\sqrt{2\pi}} \frac{1}{2} (x^{-1} - \gamma^2) e^{-\frac{1}{2} \gamma^2 x} 1_{x>0}$$

the collection of of function $y$ preserving the measure in the Lévy density is given by

$$\{y : \mathbb{R}^+ \rightarrow \mathbb{R}^+ | y(x) = \frac{1 + \gamma^2 x}{1 + \gamma^2 x} e^{-\frac{1}{2} (\gamma^2 - \gamma^2) x}, \quad \gamma \in \mathbb{R}^+ \}$$

The parameter $\delta$ can not be altered, otherwise the function $(\sqrt{y(x)} - 1)^2$ is bounded away from zero in a neighborhood of zero, implying that the integral condition (3.3) can not be satisfied.

The parameter $\gamma$ of the inverse Gaussian distribution IG$(\delta, \gamma)$ can be chosen in order to fit the data, when information about the risk-neutral world $Q$ is available. This information can be extracted from derivatives priced on the stock $S$.

Remark 1. Note that this measure transform is not structure preserving for the NIG exponential Lévy process model. However it is still tractable to work with.

4 Simulation

There are two trends to simulate Inverse Gaussian random variates and processes. One is by exact simulation using the general rejection method [21] and the other is a series expansion based on path rejection proposed by Rosinski [17]. Earlier Rosinski [16] had a technique using the inverse of the Lévy measure of the BDLP process. In the IG case this measure is not analytically invertible, hence this can only be done numerically. This is time-consuming calculation and is therefore bothersome to work with.

In the algorithms it is assumed that one can simulate from standard distributions.

4.1 Exponential Lévy process model

In the exponential Lévy process model $X(t)$ is simply a NIG Lévy process. Hence we will describe a method to simulate a NIG Lévy process.

Since the NIG distribution is a mean-variance mixture (see Section 2.1) we can simulate a NIG random variate by,

- Sample $\sigma^2$ from IG$(\delta, \sqrt{\alpha^2 - \beta^2})$.
- Sample $\epsilon$ from N$(0, 1)$.
- $X = \mu + \beta \sigma^2 + \sigma \epsilon$. 

20
By summing NIG($\alpha, \beta, \Delta \delta, \mu$) independent distributed increments one gets an NIG Lévy process on discrete points.

- Take $X(0) = 0$.
- For each increment, sample $x_i \sim \text{NIG}(\alpha, \beta, \Delta \delta, \Delta \mu)$.
- Set $X(t_i) = X(t_{i-1}) + x_i$.

To simulate $\sigma^2$ from an IG($\delta, \gamma$) distribution we will use a generator proposed by Michael, Schucany and Haas [12] (from now on referred to as MSH-method).

- Generate a random variate $Y$ with density $\chi_1^2$.
- Set $y_1 = \frac{\delta \gamma}{\delta + \gamma} + \frac{1}{\gamma^2} \sqrt{4 \delta \gamma Y + Y^2}$.
- Generate a uniform $[0, 1]$ random variate $U$ and if $U \leq \frac{\delta}{\delta + \gamma} y_1$, set $\sigma^2 = y_1$.
  If $U > \frac{\delta}{\delta + \gamma} y_1$, set $\sigma^2 = \frac{\delta^2}{\gamma^2} y_1$.

4.2 BNS SV model

By discretising time we may conclude that equation (2.4) can be rewritten into,

$$x(t) := \log S(t + \Delta) - \log S(t) \overset{D}{=} \mu \cdot \Delta + \beta \sigma^2(t) \cdot \Delta + \sqrt{\sigma^2(t) \Delta} \cdot \epsilon$$

where $\epsilon$ is a standard normal distributed random variable. This is based on the fact that for a Brownian motion $B$, $B(t + \Delta) - B(t)$ is equal in distribution to the random variate $\epsilon \sqrt{\Delta}$. We can use the following algorithm to generate log-returns $x$ on discrete points.

- Generate a sample path $\sigma^2_j(t_i)$ from IG-OU($\delta, \sqrt{\alpha^2 - \beta^2}, \lambda_j$) process, for $i = 0, \ldots, n$ and $j = 1, \ldots, m$.
- Set $\sigma^2_j(t_i) = \sum_{j=1}^{m} a_j \sigma^2_j(t_i)$ for $i = 0, \ldots, n$.
- Sample $\{\epsilon_i\}_{i=0}^{n}$ as a sequence i.i.d standard normal variables.
- Set $x_i = \mu \cdot \Delta + \beta \sigma^2_i(t_i) \cdot \Delta + \sigma(t_i) \cdot \epsilon_i \sqrt{\Delta}$, for $i = 0, \ldots, n$.
  Again by summing the increments one gets the process $X$ on discrete points.

We now focus on methods to simulate $\sigma^2$ from a IG-OU($\delta, \sqrt{\alpha^2 - \beta^2}, \lambda$) process. A solution to a SDE of the Ornstein-Uhlenbeck type,

$$d\sigma^2(t) = -\lambda \sigma^2(t) dt + dz(\lambda t)$$

is given by,

$$\sigma^2(t) = e^{-\lambda t} \sigma^2(0) + \int_0^t e^{-\lambda(t-s)} dZ(\lambda s)$$

Moreover, up to indistinguishability, this solution is unique (See Sato [19] and Barndorff-Nielsen [3]). For a IG($\delta, \gamma, \lambda$)-OU process, $\sigma^2(t)$ has stationary marginal law IG($\delta, \gamma$). So $\sigma^2(0)$ is
Chapter 2. Path dependent options and the effect of stochastic volatility

Figure 2.3: Sample path of an IG-OU process.

IG(δ, γ) distributed. Hence the most difficult term to simulate in (4.2) is the integral $\int_0^t e^{-\lambda(t-s)}dZ(\lambda s)$. We will consider two methods to simulate from an IG-OU process.

**Exact simulation**

Take,

$$\sigma^2(\Delta) = \int_0^\Delta e^{-\lambda(\Delta-s)}dZ(\lambda s)$$  \hspace{1cm} (4.3)

As shown in Zhang & Zhang [21], for fixed $\Delta > 0$ the random variable $\sigma^2(\Delta)$ can be represented as the sum of an inverse Gaussian random variable and a compound Poisson random variable in distribution, i.e.,

$$\sigma^2(\Delta) \overset{D}{=} W_0^\Delta + \sum_{i=1}^{N_\Delta} W_i^\Delta$$  \hspace{1cm} (4.4)

where $W_0^\Delta \sim IG(\delta(1 - e^{-\frac{1}{2}\lambda\Delta}), \gamma)$, random variable $N_\Delta$ has a Poisson distribution of intensity $\delta(1 - e^{-\frac{1}{2}\lambda\Delta})\gamma$ and $W_1^\Delta, W_2^\Delta, \ldots$ are independent random variables having a common specified density function,

$$f_{W^\Delta}(w) = \left\{ \begin{array}{ll} \frac{\gamma^{-1}}{\sqrt{2\pi\omega}} w^{-3/2}(e^{\frac{1}{2}\lambda\Delta} - 1)^{-1} & \text{for } w > 0, \\ \text{otherwise} & \end{array} \right.$$

$$e^{-\frac{1}{2}\gamma^2 w} - e^{-\frac{1}{2}\gamma^2 w e^{\lambda\Delta}}$$
Moreover for any \( w > 0 \) the density function \( f_{W^\Delta}(w) \) satisfies,

\[
f_{W^\Delta}(w) \leq \frac{1}{2} \left( 1 + e^{\frac{1}{2} \lambda \Delta} \right) \left( \frac{1}{2} \gamma^2 \right)^{1/2} \frac{1}{\Gamma(\frac{1}{2})} \frac{1}{2} e^{-\frac{1}{2} \gamma^2 w}
\]

Hence we can use the rejection method on a \( \Gamma\left(\frac{1}{2}, \frac{1}{2} \gamma^2\right) \) distribution to simulate random variables \( W^\Delta \) with density function \( f_{W^\Delta}(w) \).

- Generate a \( \Gamma\left(\frac{1}{2}, \frac{1}{2} \gamma^2\right) \) random variate \( Y \).
- Generate a uniform \([0, 1]\) random variate \( U \).
- If \( U \leq \frac{f_{W^\Delta}(w)}{\frac{1}{2} \left( 1 + e^{\frac{1}{2} \lambda \Delta} \right) g(Y)} \), set \( W^\Delta = Y \), where \( g(Y) = \left( \frac{1}{2} \gamma^2 \right)^{1/2} \frac{1}{\Gamma(\frac{1}{2})} Y e^{-\frac{1}{2} \gamma^2 Y} \).
  Otherwise return to the first step.

Since \( \sigma^2 \) is a stationary process we can conclude with equation (4.2), (4.3) and (4.4) that for all \( t > 0 \) we have the following equality in distribution,

\[
\sigma^2(t + \Delta) \overset{D}{=} e^{-\lambda \Delta} \sigma^2(t) + W^\Delta_0 + \sum_{j=1}^{N^\Delta} W^\Delta_j
\]

which can be translated in the following algorithm to generate a random variate \( \sigma^2(t_i) \) given the value of \( \sigma^2(t_{i-1}) \).

- Generate a \( \text{IG}(\delta(1 - e^{-\frac{1}{2} \lambda \Delta}), \gamma) \) random variate \( W^\Delta_0 \).
- Generate a random variate \( N^\Delta \) from the Poisson distribution with intensity \( \delta(1 - e^{-\frac{1}{2} \lambda \Delta}) \gamma \).
- Generate \( W^\Delta_1, W^\Delta_2, \ldots, W^\Delta_{N^\Delta} \) from the density \( f_{W^\Delta}(w) \) as independent identically distributed random variate.
- Set \( \sigma^2(t_i) = e^{-\lambda \Delta} \sigma^2(t_{i-1}) + \sum_{j=0}^{N^\Delta} W^\Delta_j \).

Moreover \( \sigma^2 \) is a stationary process therefore the initial value \( \sigma^2(0) \) can be generated from the density \( \text{IG}(\delta, \gamma) \) using the MSH-method.

## 5 Pricing of path dependent options

We consider the problem of pricing path dependent options written on an asset dynamics given by an exponential NIG-Lévy process resp. the BNS SV model. We will handle a representative case of pricing with calibrated parameters on log-returns of the Amsterdam stock exchange index (AEX). For simplicity we assume that the stock price today is \( S(0) = 100 \) and that the risk-free interest rate is \( r = 3.75\% \) yearly.
Recalling our calibration on a set of daily return data of the AEX-index, as done in section 2.4, the parameters are given by,

\[ \alpha = 94.1797 \quad \beta = -16.0141 \quad \delta = 0.0086 \quad \mu = 0.0017 \]

We will assume that in the risk-neutral world the \( \beta \) parameter is changed to \(-\frac{1}{2}\) as discussed in section 3.3. The other parameters stay unchanged in the risk-neutral world \( Q \), the parameters are given by

\[ \alpha = 94.1797 \quad \hat{\beta} = -\frac{1}{2} \quad \delta = 0.0086 \quad \mu = 0.0017 \]

Our objective is to do a simulation study with reasonable parameters. For ease we will not consider option price data and estimate the volatility parameters \( \delta \) and \( \gamma \) in the risk-neutral world \( Q \). We will assume that realistic parameters of the volatility process are \( \gamma = \sqrt{\alpha^2 - \beta^2} = 92.8083 \) and \( \delta = 0.0086 \).

Recall from section 2.4 that by least square fitting with 1 superposition \( \lambda \) is given by,

\[ \lambda = 0.0146 \]

With the above parameters we priced an Asian option, a barrier option and a Vanilla European option with a common strike \( K = 100 \) and barrier \( B = 120 \) (See Table 2.1). We priced these options for three comparable models, the BNS SV model, the exponential Lévy process model with NIG distributed increments and as an extra comparison we priced a Brownian model with drift \( \hat{\beta}\hat{\sigma}^2 \) and with the mean of the IG processes \( \hat{\sigma} = \frac{\delta}{\gamma} \) as volatility.

Table 2.1: Option prices with strike price \( K = 100 \), barrier \( B = 120 \) and exercise time \( T = 40 \). Comparing the prices under the different models.

<table>
<thead>
<tr>
<th>Model</th>
<th>European Price</th>
<th>European conf. interval</th>
<th>Asian Price</th>
<th>Asian conf. interval</th>
<th>Barrier Price</th>
<th>Barrier conf. interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brownian</td>
<td>2.3854</td>
<td>(2.2836, 2.4886)</td>
<td>1.4047</td>
<td>(1.3504, 1.4657)</td>
<td>2.3758</td>
<td>(2.2754, 2.4768)</td>
</tr>
<tr>
<td>NIG distr.</td>
<td>2.4298</td>
<td>(2.3235, 2.5363)</td>
<td>1.3611</td>
<td>(1.3029, 1.4178)</td>
<td>2.3631</td>
<td>(2.2701, 2.4588)</td>
</tr>
<tr>
<td>BNS SV model</td>
<td>2.3402</td>
<td>(2.2436, 2.4500)</td>
<td>1.3837</td>
<td>(1.3237, 1.4495)</td>
<td>2.1974</td>
<td>(2.1009, 2.2939)</td>
</tr>
</tbody>
</table>

The estimate of the option price in Table 2.1 is taken over 5,000 simulations. The confidence intervals are calculated by bootstrap using 2,000 resamples. A European option is not path dependent, hence as expected the prices of the different models are in each others confidence intervals. Also in pricing a Asian option no significant difference are observed. However in the case of pricing a barrier option, there are significant differences observed between the different models. The price simulated using the BNS-SV model is significantly lower then the price of the other models. This is due to the number of barrier crossings. The simulation of 5000 paths
of the BNS-SV model lead to 35 barrier crossings, while the simulation of 5000 paths using the Exponential Lévy model with NIG distributed increments lead to only 16 barrier crossings. So in the outliers there is a difference in distribution between the BNS-SV model and the exponential Lévy model.

Table 2.2: Option prices with strike price \( K = 100 \), barrier \( B = 110 \) and exercise time \( T = 20 \). Comparing the prices under the different models. Again the estimate of the option price is taken over 5,000 simulations. The confidence intervals are calculated by bootstrap using 2,000 resamples.

<table>
<thead>
<tr>
<th>Model</th>
<th>European Price conf. interval</th>
<th>Asian Price conf. interval</th>
<th>Barrier Price conf. interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brownian model</td>
<td>1.6731 (1.6028, 1.7427)</td>
<td>1.6471 (1.5716, 1.7228)</td>
<td>1.0087 (0.9697, 1.0543)</td>
</tr>
<tr>
<td>NIG distr.</td>
<td>1.6797 (1.6099, 1.7522)</td>
<td>1.4882 (1.4295, 1.5525)</td>
<td>1.0073 (0.9638, 1.0482)</td>
</tr>
<tr>
<td>BNS SV model</td>
<td>1.6634 (1.5920, 1.7388)</td>
<td>1.3671 (1.3092, 1.4267)</td>
<td>0.9596 (0.9165, 1.0026)</td>
</tr>
</tbody>
</table>

When altering the exercise time to 4 weeks \( (T = 20) \), both the Asian option and the barrier option show a significant price difference between the different models. However for the non-path dependent option, the European option, no significant price difference is observed when pricing with the different models (see Table 2.2). Again the number of barrier crossings when pricing with the BNS-SV model, 136, is a lot higher as when pricing is done with the exponential Lévy model 89. This is causing a price difference in pricing a barrier option using the different models. Also in the Asian option case a significant price difference is observed when pricing with the different models. The Asian prices of the 3 models are all significantly distinct.

Table 2.3: Option prices with strike price \( K = 100 \), barrier \( B = 120 \) and exercise time \( T = 60 \). Comparing the prices under the different models. Again the estimate of the option price is taken over 5,000 simulations. The confidence intervals are calculated by bootstrap using 2,000 resamples.

<table>
<thead>
<tr>
<th>Model</th>
<th>European Price conf. interval</th>
<th>Asian Price conf. interval</th>
<th>Barrier Price conf. interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brownian model</td>
<td>2.9670 (2.8412, 3.0919)</td>
<td>2.9272 (2.7940, 3.0406)</td>
<td>1.7375 (1.6692, 1.8174)</td>
</tr>
<tr>
<td>NIG distr.</td>
<td>2.9633 (2.8329, 3.0968)</td>
<td>2.6658 (2.5482, 2.7723)</td>
<td>1.7399 (1.6677, 1.8122)</td>
</tr>
<tr>
<td>BNS SV model</td>
<td>2.8776 (2.7527, 3.0147)</td>
<td>2.4794 (2.3750, 2.5831)</td>
<td>1.6651 (1.6049, 1.7499)</td>
</tr>
</tbody>
</table>

When altering the exercise time to 12 weeks \( (T = 60) \), a similar conclusion as in the case of exercise time \( T = 20 \), can be drawn. There is a significant difference between the Asian option prices, priced on the three different models. Since the Asian is path dependent, one would expect a difference in pricing with the BNS-SV model compared to pricing with the exponential Lévy process model. Due to distributional differences also a significant difference between pricing
with the Brownian-model and pricing with the other models occurs. The path dependent option shows a significant price differences. In the barrier option case this is caused by the number of barrier crossings. The BNS-SV crosses the barrier 98 times in a sample of 5000 paths, while the exponential NIG model only crosses the barrier 70 times in a sample 5000 paths.

Table 2.4: Option prices with strike price $K = 100$, barrier $B = 120$, exercise time $T = 60$ and mean reversion speed of the volatility process given by $\lambda = 2$. Comparing the prices under the different models. Again the estimate of the option price is taken over 5,000 simulations. The confidence intervals are calculated by bootstrap using 2,000 resamples.

<table>
<thead>
<tr>
<th>Model</th>
<th>European Price</th>
<th>European conf. interval</th>
<th>Asian Price</th>
<th>Asian conf. interval</th>
<th>Barrier Price</th>
<th>Barrier conf. interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brownian model</td>
<td>2.9183</td>
<td>(2.7891, 3.0396)</td>
<td>2.8813</td>
<td>(2.7607, 3.0057)</td>
<td>1.7098</td>
<td>(1.6387, 1.7781)</td>
</tr>
<tr>
<td>NIG distr.</td>
<td>2.9371</td>
<td>(2.8150, 3.0560)</td>
<td>2.7114</td>
<td>(2.6037, 2.8282)</td>
<td>1.7200</td>
<td>(1.6438, 1.7900)</td>
</tr>
<tr>
<td>BNS SV model</td>
<td>3.0275</td>
<td>(2.8978, 3.1461)</td>
<td>2.7693</td>
<td>(2.6606, 2.8941)</td>
<td>1.7537</td>
<td>(1.6846, 1.8332)</td>
</tr>
</tbody>
</table>

This difference in pricing is dependent on the mean reversion speed $\lambda$ of the volatility process $\sigma^2$ in the BNS-SV model. The bigger $\lambda$ the faster mean-reverting is the volatility process to its stationary distribution $IG(\delta, \gamma)$. Hence if $\lambda$ increases less path dependency is observed. In the limit the exponential Lévy model with NIG-distributed increments coincide with the BNS-SV model. If $\lambda = 2$ the price differences are vanishing when comparing pricing with underlying the BNS-SV model to pricing with underlying the exponential NIG model (see table 2.4). No significant price difference between pricing a barrier option with the divers models is observed. The number of barrier crossings, 60 in the BNS-SV case and 55 in the exponential NIG model, is close to each other. In pricing a Asian option a significant price difference is observed between pricing with the Brownian model and pricing with the other two models. This is probably due to the distributional differences of the increments of the different models. For big $\lambda$ when there is hardly any time dependency in the volatility process of the BNS-SV model, the increments of the BNS-SV model and the exponential NIG model are approximately the same.

6 Conclusion:

The Barndorff-Nielsen and Shephard stochastic volatility model is a complex but tractable model. It preserves nice properties as having skewed and heavy-tailed distribution of the log-returns. Moreover it models stochastic volatility.

We have done a simulation study comparing pricing with three different models describing the underlying, a Brownian model, an exponential Lévy model with NIG distributed increments and the BNS SV model. The aim was to see the effect of stochastic volatility on the pricing of path-dependent options. To this end we priced three options a Vanilla European option and two
path-dependent options; a Asian option and a Barrier option. The simulation study confirms that there is a significant price difference in pricing path-dependent option (Asian option and barrier option) when describing the underlying with different, but comparable models. Especially in pricing a barrier option the price difference between pricing with the BNS SV model compared to pricing with the other models is significantly big.

In this paper we showed that there is a difference in pricing caused by stochastic volatility, when pricing a path dependent option. A period with low return variance can lead to a more extreme price of the option.
Bibliography


Chapter 3

Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

Fred Espen Benth and Linda Vos
Abstract

Spot prices in energy markets exhibit special features like price spikes, mean-reversion, stochastic volatility, inverse leverage effect and dependencies between the commodities. In this paper a multivariate stochastic volatility model is introduced which captures these features. The second order structure and stationarity of the model are analysed in detail. A simulation method for Monte Carlo generation of price paths is introduced and a numerical example is presented.

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 integrate the power markets. Also, different electricity markets on the continental Europe exchange to a large extent energy across borders. As a reflection of this market integration is the growing need for multivariate price models for power. This includes cross-commodity models for gas and electricity, say, but also models for electricity traded in different but integrated markets. In this paper we propose and analyse a multivariate spot price model for power.

Power market spot prices have several distinct characteristics. Typically, spot prices spike occasionally when there is an imbalance in supply and demand, since the supply curve is inelastic. Further, the market prices are moving with the season, with high prices in winter due to heating, or in summer due to air-conditioning cooling. Prices also naturally mean revert due to demand and supply forces. Partly because of the spikes, the prices observed in gas and electricity markets are to a large extent leptokurtic. In fact, volatility may easily reach above 100%. A discussion of the features of power spot prices can be found in Eydeland and Wolyniec [16] and Geman [17]. There exists many models for spot price dynamics in power markets, and we refer to Benth et al. [9] 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 [15]). Little available inventory means higher and more volatile prices. This is reflected in gas markets where storage facilities play an important role in price determination. There is also evidence for dependence between different commodities. For instance, it is unlikely that the price of gas and electricity in the UK market, say, will drive too far apart, since gas is the dominating fuel for power production. Likewise, since gas can be transported as liquid natural gas (LNG), different gas markets can not have prices which become increasingly different.

In recent years there has been an interest in stochastic volatility models for commodities, and in particular energy. In Hikspoors and Jaimungal [18] we find an analysis of forward pric-
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

ing in commodity markets in the presence of stochastic volatility. Several popular models are treated. More recently, Trolle and Schwartz [30] 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 [7] applied the Barndorff-Nielsen and Shephard 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 dynamic for cross-commodity spot price modelling generalizing the univariate dynamics studied in Benth [7]. The model is flexible enough to capture spikes, mean-reversion and stochastic volatility. Moreover, it includes the possibility to model inverse leverage. Our proposed dynamics can model co- and independent jump behaviour (spikes) in cross-commodity markets. Also, the model allows for analytical forward prices. This issue, along with pricing of derivatives on spots and forwards, are left to the follow-up paper by Benth and Vos [12].

The spot price dynamics we propose are 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 [4] and Barndorff-Nielsen and Stelzer [6]). 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 [3]). 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. We provide an empirical example where the algorithm is applied. Our approach is influenced by the work of Stelzer [29].

The paper is organized as follows. Section 2 introduces the spot model, thereafter the stationary distribution and the probabilistic properties of the various factors of the model are deduced in Section 3. The following section deals with the same properties of the spot price model. Section 5 gives an empirical example and a method to perform Monte-Carlo simulation of the model. Finally, in Section 6 we conclude.

Notation

For the convenience of the reader, we have collected some frequently used notations. We adopt the notation used by Pirgorsch and Stelzer [22]. 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 $S_n$, 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 $\mathbb{R}^n \to M_n(\mathbb{R})$ which creates a diagonal matrix with the vector $v \in \mathbb{R}^n$ on the diagonal, $\text{diag}(A)$ 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 $A \otimes B$. $\text{vec}$ 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 $\text{tr}(A)$ 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. The unit vector with on the $i$-th place a one is denoted $e_i$. For $A \in M_n(\mathbb{R})$, we denote the operator $A$ associated with the matrix $A$ as $A : X \mapsto AX + XA^T$. This operator can be represented as $\text{vec}^{-1} \circ ((A \otimes I_n) + (I_n \otimes A)) \circ \text{vec}$. Its inverse is denoted by $A^{-1}$, which exists whenever $I \otimes A + A \otimes I$ is invertible. In this case, we can represent $A^{-1}$ by $\text{vec}^{-1} \circ ((A \otimes I_n) + (I_n \otimes A))^{-1} \circ \text{vec}$. Remark that $A \otimes I_n + I_n \otimes A$ is equal to the Kronecker sum of the matrix $A$ with itself.

2 The cross-commodity spot price model

Suppose we are given a complete filtered probability space $(\Omega, \mathcal{F}, P)$ equipped with the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual conditions (see e.g. Protter [24]). Assume $m, n \in \mathbb{N}$ with $0 \leq m < n$. Let $\{\tilde{L}_j(t)\}_{t \in \mathbb{R}^+} \in \mathbb{S}_d^+$, $j = 1, \ldots, n$ be $n$ independent matrix-valued subordinators as introduced in Barndorff-Nielsen and Pérez-Abreu [3]. Furthermore, let $L_i$, $i = 1, \ldots, m$ be $\mathbb{R}^d$-valued subordinators. For $i = 1, \ldots, m$ 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 $\tilde{L}_i$ in the positive definite cone $\mathbb{S}_d^+$. The subordinators are assumed to be driftless, and we choose to work with the versions which are right-continuous with left limits (RCLL, for short). 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

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

(2.1)

where $\Lambda : [0, T] \to \mathbb{R}_+^d$ is a vector of bounded measurable seasonality functions, ‘$\cdot$’ denotes coordinate-wise multiplication, and

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

(2.2)

$^1$A multivariate subordinator is a Lévy process which is increasing in each of its coordinates (see Sato [1]).
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

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

for \( i = 1, \ldots, m \). \( A, B_i \)'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. In order to have the Itô integral in (2.2) well-defined, we suppose that

\[ P\left( \int_0^T \text{tr}(\Sigma(t)) \, dt < \infty \right) = 1 . \quad (2.4) \]

Here, \( T < \infty \) is some terminal time for our energy markets. 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 processes as introduced in Barndorff-Nielsen and Stelzer [6],

\[ \Sigma(t) = \sum_{j=1}^{n} \omega_j Z_j(t) , \quad (2.5) \]

with

\[ dZ_j(t) = (C_j Z_j(t) + Z_j(t) C_j^T) dt + d\tilde{L}_j(t) , \quad (2.6) \]

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 [4] for general asset price processes. The commodity spot price dynamics with the BNS SV model as stochastic volatility structure is a generalization of the univariate spot model analysed in Benth [7].

Note that \( Y_i \) and \( \Sigma_i \) for \( i = 1, \ldots, m \) 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 follow from increasing volatility, and vice versa (see Eydeland and Wolyniec [16] and Geman [17] for a discussion on inverse leverage in power markets). We also have \( n - m \) independent stochastic volatility components \( Z_j, j = m + 1, \ldots, n \) 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 variation in the market. The latter is also sometimes referred to as the base component of the price variations.

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:

\[ S(t) = \Lambda(t) \cdot \exp(X(t)) \quad (2.7) \]

where \( X(t) \) is defined in (2.2). The Schwartz model with constant volatility is a mean-reversion
process proposed by Schwartz [28] for spot price dynamics in commodity markets like oil. In order to have spikes being independent of the volatility process $\Sigma(t)$, we can switch off some of the $\omega_j$’s in (2.5), that is choose some $\omega_j = 0$. Then the $L_i$’s from the corresponding $\tilde{L}_j$’s will give rise to independent spike components.

In electricity markets one observes spikes in the spot price dynamics (see e.g. Benth et. al. [9]). 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 $L_i$, $i = 1, \ldots, m$ 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 [20]). Independent increment processes can be thought of as time-inhomogeneous Lévy processes. Our modeling and analysis to come are easily modified to include such processes. To keep matters slightly more simplified, we stick to the time-homogeneous case here. The interested reader is referred to Benth et al [9] for applications of independent increment processes in energy markets.

We assume the following integrability conditions for the subordinators.

$$\mathbb{E} \left[ \log^+ \| \tilde{L}_j(1) \| \right] < \infty, \quad (2.8)$$

where $\log^+(x)$ is defined as $\max(\log(x), 0)$ and $j = 1, \ldots, n$ and $\|A\|^2 = \text{tr}(A^T A)$ is the Frobenius norm of the matrix $A \in M_d(\mathbb{R})$. Note that this condition implies

$$\mathbb{E} \left[ \log^+ |L_i(1)| \right] < \infty, \quad (2.9)$$

for $i = 1, \ldots, m$ and $|\cdot|$ is the Euclidean 2-norm on $\mathbb{R}^d$.

In the next Section we study the probabilistic properties of the factor processes $X$ and $Y_i$. As we shall see, the analysis of the spot price model will depend crucially on the properties of certain operators, which will reflect back to restrictions on the matrices $A$, $B_i$ and $C_j$, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Throughout the rest of the paper, we suppose that $A$, $B_i$ and $C_j$ are invertible for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Furthermore, the matrices $A$ and $C_j$ are commuting, for each $j = 1, \ldots, n$. Finally, the operators $A - C_j$ are invertible for $j = 1, \ldots, n$.

3 Stationarity and probabilistic properties of the factor processes

The processes $X, Y_i$ are Ornstein-Uhlenbeck processes. Applying the multi-dimensional Itô formula (see Ikeda and Watanabe [19]) to the stochastic differential equations yields the following
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

solutions: for \(0 \leq s \leq t\),

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

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

for \(i = 1, \ldots, m\). The matrix exponentials are defined as usual as \(e^A := I + \sum_{i=1}^{\infty} \frac{A^n}{n!}\).

According to Barndorff-Nielsen and Stelzer [6], Sect. 4, the solution of \(Z_j(t), j = 1, \ldots, n\), is given by

\[
Z_j(t) = e^{C_j(t-s)}Z_j(s)e^{C_jT(t-s)} + \int_s^t e^{C_j(t-u)}\tilde{L}_j(u)e^{C_jT(t-u)}. \tag{3.3}
\]

The matrix-valued stochastic integral in the second term of \(Z_j(t)\) is understood as follows. For two \(M_d(\mathbb{R})\)-valued bounded and measurable functions \(E(u)\) and \(F(u)\) on \([t, s]\), the notation \(\int_s^t E(u) d\tilde{L}(u) F(u)\) means the matrix \(G(s, t) \in M_d(\mathbb{R})\) with coordinates defined by

\[
G_{ij}(s, t) = \sum_{k=1}^d \sum_{l=1}^d \int_s^t E_{ik}(u)F_{lj}(u) d\tilde{L}_{kl}(u).
\]

Here, \(\tilde{L}\) is the generic notation for some \(\tilde{L}_j\). We remark that since \(\tilde{L}_j\) are supposed to be RCLL, the processes \(Z_j\) also are RCLL.

Let us first look at the expected values of \(X\) and \(Y_i\). For this, the following Lemma, which is interesting in its own right, is useful:

**Lemma 3.1.** Let \(L\) be an integrable Lévy process in \(\mathbb{R}^d\) and \(f\) a bounded measurable function from \([s, t]\) into \(M_d(\mathbb{R})\) being of bounded variation. Then it holds that

\[
\mathbb{E} \left[ \int_s^t f(u)dL(u) \right] = \int_s^t f(u)du \mathbb{E}[L(1)]. \tag{3.4}
\]

**Proof.** Define the Lévy process \(\hat{L}(u) := L(u) - \mathbb{E}[L(1)]u\), which has expectation zero. From integration by parts (use the multi-dimensional Itô Formula for jump processes in Ikeda and Watanabe [19]), it holds

\[
\int_s^t f(u) d\hat{L}(u) = f(t)\hat{L}(t) - f(s)\hat{L}(s) - \int_s^t \hat{L}(u) df(u).
\]

Now, choosing the right-continuous with left limits version of \(L\) (as we always can do for Lévy
processes), we can apply the Fubini-Tonelli Theorem to conclude that
\[
\mathbb{E} \left[ \int_s^t f(u) \, d\hat{L}(u) \right] = 0 ,
\]
and hence the Lemma follows.

We find the following conditional expectations for the factor processes:

**Lemma 3.2.** Suppose that \( L_i(1) \) are integrable for \( i = 1, \ldots, m \). Then it holds
\[
\mathbb{E}[X(t)|\mathcal{F}_s] = e^{A(t-s)}X(s) ,
\]
\[
\mathbb{E}[Y_i(t)|\mathcal{F}_s] = e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i + B_i^{-1}(\eta_i - e^{B_i(t-s)}\eta_i) \mathbb{E}[L_i(1)] ,
\]
for \( i = 1, \ldots, m \)

**Proof.** The conditional expectation of \( X(t) \) is given by
\[
\mathbb{E}[X(t)|\mathcal{F}_s] = e^{A(t-s)}X(s) + \mathbb{E} \left[ \int_s^t e^{A(t-u)}\Sigma(u)^{1/2} dW(u) \right] ,
\]
\[
= e^{A(t-s)}X(s) + \mathbb{E} \left[ \int_s^t e^{A(t-u)}\Sigma(u)^{1/2} dW(u) | \Sigma(u)_{s \leq u \leq t} \right] ,
\]
\[
= e^{A(t-s)}X(s) .
\]
In the third equality we use that the paths of \( \Sigma(u) \) are right-continuous with left limits, and therefore bounded on \([s, t]\), and hence \( u \mapsto \exp(A(s - u))\Sigma^{1/2}(s - u) \) is Itô integrable on \([t, s]\) in a strong sense. We can thus conclude that the expectation is zero of this Itô integral.

For \( Y_i, i = 1, \ldots, m \), we get
\[
\mathbb{E}[Y_i(t)|\mathcal{F}_s] = e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i + \mathbb{E} \left[ \int_s^t e^{B_i(t-u)}\eta_i dL_i(u)|\mathcal{F}_s \right] ,
\]
\[
= e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i + \int_s^t e^{B_i(t-u)}\eta_i \, du \cdot \mathbb{E}[L_i(1)] ,
\]
\[
= e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i + B_i^{-1}(\eta_i - e^{B_i(t-s)}\eta_i) \mathbb{E}[L_i(1)] .
\]
where we used Lemma 3.1 to obtain the last equality.  \( \Box \)

Since \( A \) and \( B_i \) have eigenvalues with a negative real part, letting \( t \) tend to infinity yields
\[
\lim_{t \to \infty} \mathbb{E}[X(t)|\mathcal{F}_s] = 0 ,
\]
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

\[
\lim_{t \to \infty} \mathbb{E}[Y_i(t) \mid \mathcal{F}_s] = B_i^{-1}(\mu_i + \eta_i \mathbb{E}[	ilde{L}_i(1)]) ,
\]

for \( i = 1, \ldots, m \). Hence, in stationarity, the "base-term" \( X(t) \) will contribute zero in expectation, whereas the "leverage-terms" \( Y_i \) will give a drift imposed from the subordinators and the coefficients \( \mu_i \).

Let us analyse the second-order properties of the factor processes. We have the following result for the variance of the "base component" \( X(t) \):

**Lemma 3.3.** Assume that \( \tilde{L}_j(1) \) is integrable for \( j = 1, \ldots, n \). Then it holds

\[
\text{Var}[X(t) \mid \mathcal{F}_s] = \sum_{j=1}^{n} \omega_j \left( \mathbf{A} - C_j \right)^{-1} \left\{ e^{A(t-s)}\mathbf{Z}_j(s)e^{AT(t-s)} - e^{C_j(t-s)}\mathbf{Z}_j(s)e^{C_j(t-s)} \right\} + \sum_{j=1}^{n} \omega_j C_j^{-1} \left\{ \left( \mathbf{A} - C_j \right)^{-1} \left\{ e^{A(t-s)}\mathbb{E}[\tilde{L}_j(1)]e^{AT(t-s)} - e^{C_j(t-s)}\mathbb{E}[\tilde{L}_j(1)]e^{C_j(t-s)} \right\} \right\}
\]

\[
- \sum_{j=1}^{n} \omega_j A^{-1} \left\{ C_j^{-1} \left\{ e^{A(t-s)}\mathbb{E}[\tilde{L}_j(1)]e^{AT(t-s)} - \mathbb{E}[\tilde{L}_j(1)] \right\} \right\},
\]

for \( 0 \leq s \leq t \).

**Proof.** We compute the conditional variance for the process \( X \) by appealing to the tower property of conditional expectations and the independent increment property of Brownian motion. Letting \( \mathcal{G}_{s,t} \) be the \( \sigma \)-algebra generated by \( \mathcal{F}_s \) and the paths \( \Sigma(u), s \leq u \leq t \), we find,

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

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

\[
= \mathbb{E} \left[ \int_s^t e^{A(t-u)}\Sigma(u)e^{AT(t-u)}du \mathcal{F}_s \right],
\]

\[
= \sum_{j=1}^{n} \omega_j \int_s^t e^{A(t-u)}\mathbb{E}[Z_j(u) \mid \mathcal{F}_s] e^{AT(t-u)}du ,
\]

after appealing to Fubini’s Theorem. From the explicit representation of \( Z_j(t) \) in (3.3), we find

\[
\mathbb{E}[Z_j(u) \mid \mathcal{F}_s] = e^{C_j(u-s)}\mathbf{Z}_j(s)e^{C_j'T(u-s)} + \int_s^u e^{C_j(u-v)}\mathbb{E}[\tilde{L}_j(1)]e^{C_j'T(u-v)}dv
\]

\[
= e^{C_j(u-s)}\mathbf{Z}_j(s)e^{C_j'T(u-s)} + \int_0^{u-s} e^{C_jz}\mathbb{E}[\tilde{L}_j(1)]e^{C_j'Tz}dz
\]

\[
= e^{C_j(u-s)}\mathbf{Z}_j(s)e^{C_j'T(u-s)} + C_j^{-1} \left\{ e^{C_j(t-s)}\mathbb{E}[\tilde{L}_j(1)]e^{C_j'T(t-s)} - \mathbb{E}[\tilde{L}_j(1)] \right\} ,
\]

40
after appealing to Lemma 3.1. Hence, using that $A$ and $C_j$ are commuting for each $j = 1, \ldots, n$, we find

$$
\text{Var}[X(t)|\mathcal{F}_s] = \sum_{j=1}^n \omega_j \int_s^t e^{A(t-u)} e^{C_j(u-s)} Z_j(s) e^{C_j^T(u-s)} e^{A^T(t-u)} du \\
+ \sum_{j=1}^n \omega_j C_j \left\{ \int_s^t e^{A(t-u)} e^{C_j(u-s)} E[\tilde{L}_j(1)] e^{C_j^T(u-s)} e^{A^T(t-u)} du \right\} \\
- \sum_{j=1}^n \omega_j C_j^{-1} \int_s^t e^{A(t-u)} E[\tilde{L}_j(1)] e^{A^T(t-u)} du \\
= \sum_{j=1}^n \omega_j (A - C_j)^{-1} \left\{ e^{A(t-s)} Z_j(s) e^{A^T(t-s)} - e^{C_j(t-s)} Z_j(s) e^{C_j^T(t-s)} \right\} \\
+ \sum_{j=1}^n \omega_j C_j^{-1} \left\{ (A - C_j)^{-1} \left\{ e^{A(t-s)} E[\tilde{L}_j(1)] e^{A^T(t-s)} - e^{C_j(t-s)} E[\tilde{L}_j(1)] e^{C_j^T(t-s)} \right\} \right\} \\
- \sum_{j=1}^n \omega_j A^{-1} \left\{ C_j^{-1} \left\{ e^{A(t-s)} E[\tilde{L}_j(1)] e^{A^T(t-s)} - E[\tilde{L}_j(1)] \right\} \right\}.
$$

The Lemma follows. \hfill \Box

Note that the explicit expression for the variance of the base component is computed under the condition of the matrices $A$ and $C_j$ being commutable. Moreover, we observe that for the Lemma to hold, we must have the imposed conditions of invertibility of the operators $A$, $C_j$ and $A - C_j$. Recalling that the matrices $A$ and $C_j$ have eigenvalues with negative real part, we pass to the limit $t \to \infty$ to find

$$
\lim_{t \to \infty} \text{Var}[X(t)] = \sum_{j=1}^n \omega_j A^{-1} C_j^{-1} E[\tilde{L}_j(1)].
$$

Observe that the stationary limit of the variance depends explicitly on the mean-reversion coefficient matrices $A$ and $C_j$. In fact, from Barndorff-Nielsen and Stelzer [6] we know that the stationary expected value of $Z_j(s)$ is $C_j^{-1} E[\tilde{L}_j(1)]$, so we can write

$$
\lim_{t \to \infty} \text{Var}[X(t)|\mathcal{F}_s] = A^{-1} \lim_{t \to \infty} E[\Sigma(t)]. \tag{3.5}
$$

for the stationary variance of the base component.

We move on and find the variance of $Y_i(t)$:

**Lemma 3.4.** Suppose that $L_i(1)$ are square integrable for $i = 1, \ldots, m$. Then it holds,

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

41
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

\[ - B_i^{-1}(I - e^{B_i (t-s)}) \eta_i \mathbb{E}[L_i(1)] \mathbb{E}[L_i^T(1)] \eta_i^T (I - e^{B_i^T (t-s)}) B_i^{-T}, \]

for \( i = 1, \ldots, m \) and \( 0 \leq s \leq t \).

**Proof.** Fix an \( i = 1, \ldots, n \). By (3.2), we find that the conditional variance of \( Y_i(t) \) given \( \mathcal{F}_s \) is

\[ \text{Var}[Y_i(t)|\mathcal{F}_s] = \text{Var}\left[ \int_s^t e^{B_i(u-s)} \eta_i \, dL_i(u) | \mathcal{F}_s \right]. \]

Moreover, by the independent increment property of Lévy processes it holds

\[ \text{Var}[Y_i(t)|\mathcal{F}_s] = \text{Var}\left[ \int_s^t e^{B_i(u-s)} \eta_i \, dL_i(u) \right]. \]

But, by Itô isometry for Lévy process integrals

\[
\mathbb{E} \left[ \int_s^t e^{B_i(u-s)} \eta_i \, dL_i(u) \int_s^t e^{B_i(u-s)} \eta_i \, dL_i(u)^T \right] \\
= \int_s^t e^{B_i(u-s)} \eta_i \mathbb{E}[L_i(1) L_i^T(1)] \eta_i^T e^{B_i^T(u-s)} \, du \\
= \mathbf{B}^{-1} \left( \eta_i \mathbb{E}[L_i 1(1) L_i^T(1)] \eta_i^T - e^{B_i(t-s)} \eta_i \mathbb{E}[L_i(1) L_i^T(1) \eta_i^T e^{B_i^T(t-s)}] \right)
\]

On the other hand, following from Lemma 3.1

\[
\mathbb{E} \left[ \int_s^t e^{B_i(u-s)} \eta_i \, dL_i(u) \right] = \int_s^t e^{B_i(u-s)} \, du \eta_i \mathbb{E}[L_i(1)] \\
= \mathbf{B}_i^{-1}(I - e^{B_i(t-s)}) \eta_i \mathbb{E}[L_i(1)].
\]

Hence, the Lemma follows. \( \square \)

Note that we have used the standing condition of invertibility of the operators \( \mathbf{B}_i \) in this Lemma. We can also for \( Y_i(t) \) compute an explicit stationary limit for the variance using that the eigenvalues of \( \mathbf{B}_i \) have negative real parts:

\[
\lim_{t \to \infty} \text{Var}[Y_i(t)|\mathcal{F}_s] = \mathbf{B}_i^{-1} \eta_i \mathbb{E}[L_i(1) L_i^T(1)] \eta_i^T - \mathbf{B}_i^{-1} \eta_i \mathbb{E}[L_i(1)] \mathbb{E}[L_i^T(1)] \eta_i^T \mathbf{B}_i^{-T}. \tag{3.6}
\]

This holds for every \( i = 1, \ldots, m \).

From an empirical point of view, the covariance structures between factors in the temporal direction are useful. We compute this in the next Lemma:
Lemma 3.5. Suppose that $L_i(1)$ is square integrable for $i = 1, \ldots, m$. Then, for $0 \leq s \leq t$,

$$\text{Cov}[X(t), Y_i(t)|F_s] = 0 = \text{Cov}[Y_i(t), Y_j(t)|F_s],$$

for $i \neq j$ and $i, j = 1, \ldots, m$. Furthermore, if $\tilde{L}_j(1)$ are integrable for $j = 1, \ldots, n$, then the conditional auto-covariance functions of $X$ and $Y_i$ are given by,

$$\text{acov}_X(s, t, h) := \text{Cov}[X(t + h), X(t)|F_s] = e^{Ah}\text{Var}[X(t)|F_s]$$

$$\text{acov}_{Y_i}(s, t, h) := \text{Cov}[Y_i(t + h), Y_i(t)|F_s] = e^{B_ih}\text{Var}[Y_i(t)|F_s],$$

for $i = 1, \ldots, m, 0 \leq s \leq t$ and $h \geq 0$ a constant (the lag of the auto-covariance).

Proof. First, from (3.2) we find,

$$\text{Cov}[Y_i(t), Y_j(t)|F_s] = \text{Cov} \left[ \int_s^t e^{B_i(t-u)} \eta_i dL_i(u), \int_s^t e^{B_j(t-u)} \eta_j dL_j(u) \right] = 0,$$

for $i \neq j$, since in that case $L_i$ and $L_j$ are independent.

Next, from (3.1) and (3.2) we find for given $i = 1, \ldots, m$,

$$\text{Cov}[X(t), Y_i(t)|F_s] = \text{Cov} \left[ \int_s^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u), \int_s^t e^{B_i(t-u)} \eta_i dL_i(u)|F_s \right].$$

We recall that $\Sigma(t)$ and $W(t)$ are independent. Using the tower property of conditional expectation, where we condition of the $\sigma$-algebra $G_{t,s}$ generated by all paths of $\tilde{L}_j(u); 0 \leq u \leq t$ and $F_s, for \ j = 1, \ldots, n$, we find,

$$\mathbb{E} \left[ \left( \int_s^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u) \right) \left( \int_s^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T |F_s \right]$$

$$= \mathbb{E} \left[ \left( \int_s^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u) \right) \left( \int_s^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T |G_{t,s} |F_s \right],$$

$$= \mathbb{E} \left[ \left( \int_s^t e^{A(t-u)} \Sigma(u)^{1/2} dW(u) \right) |G_{t,s} \right] \left( \int_s^t e^{B_i(t-u)} \eta_i dL_i(u) \right)^T |F_s \right],$$

$$= 0.$$
Next, let us derive the auto-covariance function for $X$. From (3.1), we find for $h \geq 0$

$$X(t + h) = e^{Ah}X(t) + \int_t^{t+h} e^{A(t+h-u)}\Sigma(u)^{1/2}dW(u).$$

Hence,

$$\text{acov}_X(s, t, h) = e^{Ah}\text{Var}[X(t)|F_s] + \text{Cov}[\int_t^{t+h} e^{A(t+h-u)}\Sigma(u)^{1/2}dW(u), X(t)|F_s].$$

By using the same double conditioning argument as above, we see that the second term is equal to zero since Brownian motion has independent increments. This proves the auto-covariance function of $X$. For the case $Y_i$, we use exactly the same argument. Use (3.2) and the independent increment property of Lévy processes to reach the result. The Lemma follows.

From an empirical point of view, the stationary auto-covariance functions are particularly interesting. From (3.5) and (3.6) it follows

$$\lim_{t \to \infty} \text{acov}_X(s, t, h) = e^{Ah} \sum_{j=1}^{n} \omega_j A^{-1} C_j^{-1} \mathbb{E}[\tilde{L}_j(1)],$$

$$\lim_{t \to \infty} \text{acov}_Y(s, t, h) = e^{B_i h} \left( B_i^{-1} \eta_i \mathbb{E}[L_i(1) L_i^T(1)] \eta_i^T - B_i^{-1} \eta_i \mathbb{E}[L_i(1)] \mathbb{E}[L_i^T(1)] \eta_i^T \right).$$

As $A$ and $B_i$ have eigenvalues with negative real parts, we see that the de-seasonalized log-spot prices $\ln S_k(t) - \ln \Lambda_k(t)$ of commodity $k = 1, \ldots, d$ will in stationarity have an auto-correlation function being a sum of exponential functions, with decay rates given by the real parts of the eigenvalues of $A$ and $B_i$, $i = 1, \ldots, n$. This is an empirical feature we often see with energy prices (see for example Benth, Kiesel and Nazarova [10]).

### 3.1 Cumulants and stationary distributions

Under the log integrability conditions (2.8), the processes $Y_i$ and $Z_j$ are stationary (see Sato [27], Thm. 5.2). In the next Proposition the characteristic function of the stationary distributions of $X$, $Y_i$ and $Z_j$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ are calculated in terms of the characteristic function of the matrix-valued processes $\tilde{L}_j$.

Let us first investigate the cumulant and the stationary distribution of $Z_j$ and $Y_i$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

**Proposition 3.6.** For $t \geq s$, the conditional cumulant functions of $Y_i$ and $Z_j$ are, resp.,

$$\phi^{(s,t)}_{Y_i}(z) = i \left( e^{B_i(t-s)} Y_i(s) + iB_i^{-1}(I - e^{B_i(t-s)}) \mu_i \right)^T z + \int_0^{t-s} \phi^{(J_d)}_{Z_j}(J_d(\eta_i^T e^{B_i^T u} z)) du, \quad (3.7)$$
\[ \phi_{Z_i}^{(s,t)}(V) = iV e^{C_j(t-s)}Z_j(s)e^{CT_j(t-s)} + \int_0^{t-s} \phi_{L_j}(e^{C_j(t-s)}Ve^{CT_j(t-s)}) \, du, \quad (3.8) \]

for \( i = 1, \ldots, m \) and \( j = 1, \ldots, n \).

**Proof.** For the cumulants of \( Y_i, i = 1, \ldots, m \) using (3.2) it holds

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

for \( t \geq s \). Hence, by the key formula (see Sato [27]), the conditional cumulant function of \( Y_i(t) \) given \( \mathcal{F}_s \) is

\[ \phi_{Y_i}^{(s,t)}(z) = i(e^{B_i(t-s)}Y_i(s) + iB_i^{-1}(I - e^{B_i(t-s)})\mu_i)^T z + \int_s^t \phi_{L_i}(\eta_i^T e^{B_i(t-u)z}) \, du, \]

\[ = i(e^{B_i(t-s)}Y_i(s) + iB_i^{-1}(I - e^{B_i(t-s)})\mu_i)^T z + \int_0^{t-s} \phi_{L_i}(J_d(\eta_i^T e^{B_i(t-u)z})) \, du. \]

The cumulant functions of the \( Z_j \)'s are computed in Pighorsch and Stelzer [23]. We include the derivation here for the convenience of the reader. By (3.3) and the independent increment property of Lévy processes,

\[ \ln E \left[ e^{iVZ_j(t)} | \mathcal{F}_s \right] = iVe^{C_j(t-s)}Z_j(s)e^{CT_j(t-s)} + \ln \mathbb{E} \left[ e^{iV \int_s^t e^{C_j(t-u)}dL_j(u)e^{CT_j(t-u)}} | \mathcal{F}_s \right] \]

\[ = iVe^{C_j(t-s)}Z_j(s)e^{CT_j(t-s)} + \ln \mathbb{E} \left[ e^{iV \int_s^t e^{C_j(t-u)}dL_j(u)e^{CT_j(t-u)}} \right] \]

\[ = iVe^{C_j(t-s)}Z_j(s)e^{CT_j(t-s)} + \int_s^t \phi_{L_j}(e^{C_j(t-u)}Ve^{CT_j(t-u)}) \, du. \]

Hence, the Lemma follows. \( \square \)

Since \( L(t) \) has finite log moments and \( \sigma(B_i) \subseteq (-\infty, 0) + i\mathbb{R}^+ \), the limit of \( \phi_{Y_i}^{(s,t)} \) for \( t \to \infty \) is well-defined (see Sato [27]) and given by

\[ \lim_{t \to \infty} \phi_{Y_i}^{(s,t)}(z) := \phi_{Y_i}(z) = i\mu_i^T (B_i^T)^{-1}z + \int_0^\infty \phi_{L_i}(J_d(\eta_i^T e^{B_i^T u}z)) \, du, \quad z \in \mathbb{R}^d, \]

for \( i = 1, \ldots, m \). This is the cumulant function of the stationary distribution of \( Y_i \). Similarly, we find the cumulant function of the stationary distribution of the \( Z_j \)'s to be

\[ \lim_{t \to \infty} \phi_{Z_j}^{(s,t)}(z) := \phi_Z(V) = \int_0^\infty \phi_{L_j}(e^{CT_j Ve^{C_j s}}) \, ds, \quad V \in \mathbb{S}_d, \]
Lemma 3.7. Define \( f(s, t) := \int_s^t e^{A(t-u)} \sum(u)e^{AT(t-u)} du \). Then it holds

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

for \( 0 \leq s \leq t \).

Proof. Using (3.3) and the assumption that \( A \) and \( C_j \) commute for \( j = 1, \ldots, n \) it holds

\[
f(s, t) = \int_s^t e^{A(t-u)} \sum_{j=1}^{n} \omega_j \left( e^{C_j(u-s)} Z_j(s) e^{C_j(u-s)} + \int_s^u e^{C_j(u-v)} d\tilde{L}_j(v) e^{C_j(u-v)} \right) e^{AT(t-u)} du
\]

\[
= \sum_{j=1}^{n} \omega_j \int_s^t e^{(C_j-A)u} e^{At-C_j s} \left( Z_j(s) + \int_s^u e^{-C_j v} d\tilde{L}_j(v) e^{-C_j v} \right) e^{AT-C_j s} e^{(C_j-A)T u} du
\]

\[
= \sum_{j=1}^{n} \omega_j (C_j-A)^{-1} \left( e^{C_j(t-s)} Z_j(s) e^{C_j(t-s)} - e^{A(t-s)} Z_j(s) e^{AT(t-s)} \right)
\]

\[
+ \int_s^t \int_s^u \left\{ e^{(C_j-A)u} e^{At} e^{-C_j v} d\tilde{L}_j(v) e^{-C_j v} e^{AT} e^{(C_j-A)T u} \right\} du.
\]

The last integral is interpreted as first integrating with respect to \( d\tilde{L}_j(v) \), and next integrating the obtained expression with respect to \( du \). But, by spelling out the integrals in terms of sums, using the definition of the \( d\tilde{L}_j(v) \) integrals, and invoking the stochastic Fubini theorem (see Protter [24]), we get

\[
\int_s^t \int_s^u \left\{ e^{(C_j-A)u} e^{At} e^{-C_j v} d\tilde{L}_j(v) e^{-C_j v} e^{AT} e^{(C_j-A)T u} \right\} du
\]

\[
= \int_s^t \int_v^u \left\{ e^{(C_j-A)u} e^{At} e^{-C_j v} d\tilde{L}_j(v) e^{-C_j v} e^{AT} e^{(C_j-A)T u} \right\} du.
\]

Here, the right hand side is interpreted as first integrating with respect to \( du \), treating \( d\tilde{L}_j(v) \) as a matrix and not a differential, and next integrating with respect to \( d\tilde{L}_j(v) \) the obtained expression.
Hence, we find
\[ f(s, t) = \sum_{j=1}^{n} C_j(t - s) Z_j(s) \]
\[ + (C_j - A)^{-1} \left( \int_{s}^{t} e^{C_j(t-u)} d\tilde{L}_j(u) e^{C_j^T(t-u)} - \int_{s}^{t} e^{A(t-u)} d\tilde{L}_j(u) e^{A^T(t-u)} \right) . \]

The Lemma follows. \(\square\)

With this result at hand, we can derive the conditional cumulant function of \(X(t)\). This is done in the next Proposition.

**Proposition 3.8.** The conditional cumulant function of the process \(X(t)\) given \(\mathcal{F}_s\) is

\[ \phi^{X}_{s,t}(z) = iX^T(s) e^{A^T(t-s)} z - \frac{1}{2} \sum_{j=1}^{n} z^T C_j(t - s) Z_j(s) z + \sum_{j=1}^{n} \int_{0}^{t-s} \phi_{\tilde{L}_j} \left( \frac{1}{2} C_j^*(u) z z^T \right) du , \]

for every \(0 \leq s \leq t\), and \(z \in \mathbb{R}^d\), where \(C_j^*\) is the adjoint operator of \(C_j\) defined in (3.9).

**Proof.** Let \(\mathcal{G}_{t,s}\) denote the filtration generated by \(\mathcal{F}_s\) and the paths \(\Sigma(u), 0 \leq u \leq t\). By the independence of \(W\) and \(\tilde{L}_j\) for \(j = 1 \ldots n\), and the tower property of conditional expectations, we have that

\[ \phi^{X}_{s,t}(z) = \ln \mathbb{E} \left[ \mathbb{E} \left[ e^{i(z,X(t))} | \mathcal{G}_{t,s} \right] | \mathcal{F}_s \right] \]
\[ = iX^T(s) e^{A^T(t-s)} z + \ln \mathbb{E} \left[ \exp \left( i \left( \int_{s}^{t} \Sigma(u)^{1/2} e^{A(t-u)} dW(u) \right)^T z \right) | \mathcal{G}_{t,s} \right] | \mathcal{F}_s \]
\[ = iX^T(s) e^{A^T(t-s)} z + \ln \mathbb{E} \left[ \exp \left( -\frac{1}{2} z^T \int_{s}^{t} e^{A(t-u)} \Sigma(u) e^{A^T(t-u)} du z \right) | \mathcal{F}_s \right] \]

In the second equality, we used (3.1) and in the third equality we used the Gaussianity of a Wiener integral (note that the integrand is a deterministic function after conditioning on \(\mathcal{G}_{t,s}\)). Appealing to Lemma 3.7, we find

\[ \phi^{X}_{s,t}(z) = iX^T(s) e^{A^T(t-s)} z - \frac{1}{2} \sum_{j=1}^{n} z^T C_j(t - s) Z_j(s) z \]
\[ + \sum_{j=1}^{n} \ln \mathbb{E} \left[ \exp \left( -\frac{1}{2} z^T C_j(t-u) d\tilde{L}_j(u) z \right) | \mathcal{F}_s \right] \]
together with Fubini-Tonelli’s Theorem, we get

\[ iX^T(s)e^{dt(s)}z - \frac{1}{2} \sum_{j=1}^{n} z^T C_j(t - s) Z_j(s)z \]

\[ + \sum_{j=1}^{n} \ln \mathbb{E} \left[ \exp \left( \text{itr} \left( \frac{1}{2} i z z^T \int_{s}^{t} C_j(t - u) d\tilde{L}_j(u) \right) \right) \right]. \]

In the last step, we used the fundamental relation \( z^T U z = \text{tr}(zz^T U) \) for a quadratic matrix \( U \) together with the independent increment property of Lévy processes. Now, observe that the stochastic integral can be expressed as

\[ \int_{s}^{t} C_j(t - u) d\tilde{L}_j(u) = \lim_{|\Delta_t| \to 0} \sum_{i=0}^{n-1} C_j(t - u_i) \Delta \tilde{L}_j(u_i), \]

for partitions \( s = u_0 < \cdots < u_n = t \) with \( \Delta_i := \tilde{L}_j(u_{i+1}) - \tilde{L}_j(u_i) \) and \( \Delta_i := u_{i+1} - u_i \). By independence of increments of a Lévy process, and continuity of the exponential function together with Fubini-Tonelli’s Theorem, we get

\[ \mathbb{E} \left[ \exp \left( \text{itr} \left( \frac{1}{2} i z z^T \int_{s}^{t} C_j(t - u) d\tilde{L}_j(u) \right) \right) \right] \]

\[ = \lim_{|\Delta_t| \to 0} \prod_{i=1}^{n-1} \mathbb{E} \left[ \exp \left( \text{itr} \left( \frac{1}{2} i z z^T C_j(t - u_i) \Delta \tilde{L}_j(u_i) \right) \right) \right]. \]

Now, the linear operators \( C_j(t - u_i) \) can be represented as \( \text{vec}^{-1} \circ K \circ \text{vec} \) for a matrix \( K \in \mathbb{R}^{d^2} \). Hence, since for quadratic matrices \( \text{tr}(VX) = \text{vec}(V)^T \text{vec}(X) \), we find

\[ \text{tr} \left( VC_j(t - u_i) \Delta \tilde{L}_j(u_i) \right) = \text{vec}(V)^T \text{vec} \left( C_j(t - u_i) \Delta \tilde{L}_j(u_i) \right) \]

\[ = \text{vec}(V)^T \text{vec} \left( \text{vec}^{-1} \circ K \circ \text{vec}(\Delta \tilde{L}_j(u_i)) \right) \]

\[ = \text{vec}(V)^T K \text{vec}(\Delta \tilde{L}_j(u_i)) \]

\[ = (K^T \text{vec}(V))^T \text{vec}(\Delta \tilde{L}_j(u_i)). \]

Thus,

\[ \ln \mathbb{E} \left[ \exp \left( \text{itr}(VC_j(t - u_i) \Delta \tilde{L}_j(u_i)) \right) \right] = \ln \mathbb{E} \left[ \exp \left( i (K^T \text{vec}(V))^T \text{vec}(\Delta \tilde{L}_j(u_i)) \right) \right] \]

\[ = \ln \mathbb{E} \left[ \exp \left( \text{itr} \left( \text{vec}^{-1}(K^T \text{vec}(V)) \Delta \tilde{L}_j(u_i) \right) \right) \right] \]

\[ = \phi_{\tilde{L}_j} \left( \text{vec}^{-1} \circ K^T \circ \text{vec}(V) \right) \Delta_i \]

\[ = \phi_{\tilde{L}_j} \left( C_j^*(t - u_i) V \right) \Delta_i. \]
Letting $V = \frac{1}{2}izz^T$, we conclude that
\[
\ln \mathbb{E} \left[ \exp \left( i \text{tr} \left( \frac{1}{2}izz^T C_j(t - u_i) \Delta \tilde{L}_j(u_i) \right) \right) \right] = \int_s^t \phi_{\tilde{L}_j} \left( \frac{1}{2} i C_j^*(t - u) zz^T \right) du
= \int_0^{t-s} \phi_{\tilde{L}_j} \left( \frac{1}{2} i C_j^*(u) zz^T \right) du.
\]
This proves the Proposition.

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

**Proposition 3.9.** The process $X(t)$ is stationary and the cumulant function of the limiting distribution is given by

\[
\lim_{t \to \infty} \phi_X^{(n,t)}(z) := \phi_X(z) = \sum_{j=1}^n \int_0^{\infty} \phi_{\tilde{L}_j} \left( \frac{1}{2} i C_j^*(s) zz^T \right) ds,
\]
where $z \in \mathbb{R}^d$ and the linear operator $C_j(t)$ is defined in (3.9).

**Proof.** By the definition of $C_j(t)$ and the fact that $A$ and $C_j$, $j = 1, \ldots, n$ have eigenvalues with negative real parts, it is straightforward to see that
\[
\lim_{t \to \infty} iX^T(s)e^{A^T(t-s)}z - \frac{1}{2} \sum_{j=1}^n z^T C_j(t-s)Z_j(s)z = 0.
\]
Hence, we must prove that the integral
\[
\int_0^t \phi_{\tilde{L}_j} \left( \frac{1}{2} i C_j^*(s) zz^T \right) ds
\]
converges when $t \to \infty$, for every $j = 1, \ldots, n$. To prove this it is sufficient to show that
\[
\int_s^t C_j(t-u)d\tilde{L}_j(u)
\]
has a stationary solution for each $j = 1, \ldots, n$. Let us fix $j = 1, \ldots, n$, and observe that by definition of $C_j(t)$ and linearity of the $C_j - A$-operator, we have
\[
\int_s^t C_j(t-u)d\tilde{L}_j(u) = \omega_j(C_j - A)^{-1} \left\{ \int_s^t e^{C_j(t-u)}d\tilde{L}_j(u)e^{C_j(t-u)} - \int_s^t e^{A(t-u)}d\tilde{L}_j(u)e^{A(t-u)} \right\}.
\]
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

But the two stochastic integrals are stationary by Sato [27] Theorem 5.2 since \( A \) and \( C_j \) have eigenvalues with negative real parts. Hence, the result follows since any linear combination of stationary processes will in itself be stationary. \( \square \)

We observe that the limiting distribution of \( X \) must be centered and symmetric since its cumulant function satisfies \( \phi_X(z) = \phi_X(-z) \). We discuss the stationary distribution of \( X \) in more detail.

As we now argue, the stationary distribution of \( X \) can be viewed as the convolution of a centered normal and a leptokurtic distribution whenever \( \tilde{L}_j(1) \) are integrable for \( j = 1, \ldots, n \). To show this we introduce the zero-mean matrix valued Lévy process \( \tilde{L}_j(t) \triangleq \tilde{L}_k(t) - \mathbb{E}[\tilde{L}_j(1)]t \), and denote by \( \phi_{\tilde{L}_j}(V) \) its cumulant defined by

\[
\phi_{\tilde{L}_j}(V) = \phi_{\tilde{L}_j}(V) - i \text{tr}(V \mathbb{E}[\tilde{L}_j(1)]) .
\]

The cumulant function of the stationary distribution of \( X(t) \) can henceforth be expressed as

\[
\phi_X(z) = \sum_{j=1}^{n} \left\{ \int_0^\infty \phi_{\tilde{L}_j} \left( \frac{1}{2} i C_j^*(s) zz^T \right) ds + i \int_0^\infty \text{tr} \left( \frac{1}{2} i C_j^*(s) zz^T \mathbb{E}[\tilde{L}_j(1)] \right) ds \right\}.
\]

Using properties of the trace-operator we have

\[
\text{tr} \left( (C_j^*(s) zz^T) \mathbb{E}[\tilde{L}_j(1)] \right) = \text{vec}(C_j^*(s) zz^T)^T \text{vec}(\mathbb{E}[\tilde{L}_j(1)])
\]

\[
= \text{vec} ( \text{vec}^{-1}(K^T \text{vec}(zz^T)))^T \text{vec}(\mathbb{E}[\tilde{L}_j(1)])
\]

\[
= (K^T \text{vec}(zz^T))^T \text{vec}(\mathbb{E}[\tilde{L}_j(1)])
\]

\[
= \text{vec}(zz^T)K \text{vec}(\mathbb{E}[\tilde{L}_j(1)])
\]

\[
= \text{tr} \left( zz^T \text{vec}^{-1}(K \text{vec}(\mathbb{E}[\tilde{L}_j(1)])) \right)
\]

\[
= \text{tr} (zz^T C_j(s) \mathbb{E}[\tilde{L}_j(1)])
\]

\[
= z^T C_j(s) \mathbb{E}[\tilde{L}_j(1)] z .
\]

Here, we have used that the operator \( C_j(s) \) can be represented by the \( \mathbb{R}^{d^2 \times d^2} \)-matrix \( K \) as \( C_j(s) = \text{vec}^{-1} \circ K \circ \text{vec} \). Using Lemma 3.3, we conclude

\[
\phi_X(z) = \sum_{j=1}^{n} \int_0^\infty \phi_{\tilde{L}_j} \left( \frac{1}{2} i C_j^*(s) zz^T \right) ds - \frac{1}{2} z^T \left( \lim_{t \to \infty} \text{Var}[X(t)] \right) z .
\]
The last term is the characteristic function of a centered multivariate normal distribution with variance equal to \( \lim_{t \to \infty} \text{Var}[X(t)] \). We remark that this coincides with the stationary distribution obtained from the multivariate Schwartz model having constant volatility \( \Sigma \in \mathcal{M}(\mathbb{R}) \) given by

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

The first term in \( \phi_X(z) \) will be the characteristic function of a non-Gaussian distribution.

### 4 Analysis of the spot dynamics

Let us look at the dynamics of \( \widetilde{S}(t) \equiv S(t)/\Lambda(t) \), the deseasonalized spot price, where the division is done elementwise.

**Proposition 4.1.** It holds that

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

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). \( \square \)

We see from this result that the dynamics can be interpreted as a mean-reverting process towards a stochastic mean. The mean will be described by the multivariate process \( M(t) \), which will consist of linear combinations of the different "spike" components \( Y_j \). The matrix \( A \) describes the "speed" of mean-reversion, as well as how the different commodities are functionally dependent on each other. Moreover, the stochastic volatility term and the spike contributions are clearly dependent.

We move on analysing the stationary distribution of \( \ln \widetilde{S}(t) \). From Lemma 3.2, we find in stationarity that

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

Furthermore, from Lemma 3.5 we know that in stationarity, the auto-covariance function of \( \ln \widetilde{S}(t) \) is

\[
\text{acov}_{\ln \widetilde{S}}(h) = \text{acov}_X(h) + \text{acov}_{\sum Y_i}(h)
\]

(4.1)
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

\[ \begin{align*}
&= e^{A[h]} \sum_{j=1}^{n} \omega_j A^{-1} C_j^{-1} \mathbb{E}[\tilde{L}_j(1)] \\
&\quad + \sum_{i=1}^{m} e^{B_i[h]} (B_i^{-1} \eta_i \mathbb{E}[L_i(1) L_i^T(1)] (\eta_i)^T - (B_i^{-1} \eta_i \mathbb{E}[L_i(1)] \mathbb{E}[L_i^T(1)] \eta_i^T).
\end{align*} \]

Hence, in stationarity, the auto-covariance function of \( \ln \tilde{S}(t) \) will be a linear combination of exponentially decaying functions due to eigenvalues with a negative real part. This is in line with empirical observations of power prices, as we have earlier noted (see e.g. Benth, Kiesel and Nazarova [10]).

By combining the results of the cumulant functions for the different factors in the dynamics of \( \ln \tilde{S}(t) \) derived in the previous section, we can compute the cumulant of the deseasonalized log-spot prices. This is presented in the next Proposition.

**Proposition 4.2.** The characteristic function of the stationary distribution of \( \ln \tilde{S}(t) \) is given by

\[ \phi_{\ln \tilde{S}}(z) = \sum_{i=1}^{m} i \mu_i^T (B_i^T)^{-1} z + \sum_{j=1}^{n} \int_{0}^{\infty} \phi_{L_j} \left( \frac{1}{2} i C_j^* (u) z z^T \right) du \]

\[ \quad + \sum_{i=1}^{m} \int_{0}^{t-s} \phi_{L_i} \left( \frac{1}{2} i C_i^* (u) z z^T + J_d (\eta_i e^{B_i^T u} z) \right) - \phi_{L_i} \left( \frac{1}{2} i C_i^* (u) z z^T \right) du, \]

for \( z \in \mathbb{R}^d \).

**Proof.** By combining Proposition 3.8 and equation (3.7) in Proposition 3.6 the conditional cumulant function of \( \ln \tilde{S} \) given \( \mathcal{F}_s \) is

\[ \phi_{\ln \tilde{S}}^{s,t}(z) = iX^T(s)e^{AT(t-s)}z - \frac{1}{2} \sum_{j=1}^{n} z^T C_j (t-s) Z_j (s) z \]

\[ \quad + \sum_{i=1}^{m} iY_i^T(s)e^{BT(t-s)}z + i (B_i^{-1} (I - e^{B_i^T (t-s)}) \mu_i)^T z \]

\[ \quad + \sum_{i=1}^{m} \int_{0}^{t-s} \phi_{L_i} \left( \frac{1}{2} i C_i^* (u) z z^T + J_d (\eta_i e^{B_i^T u} z) \right) du \]

\[ \quad + \sum_{j=m+1}^{n} \int_{0}^{t-s} \phi_{L_j} \left( \frac{1}{2} i C_j^* (u) z z^T \right) du. \]

Since a stationary solution exists for \( X \) and all \( Y_i \)'s, there also exists a stationary solution for \( \ln \tilde{S} \). The Proposition follows by taking limits for \( t \to \infty \) using that the real parts of the eigenvalues of the involved matrices are negative. \( \square \)

52
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 leave the analysis of this to the interested reader.

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

$$\ln \tilde{S}(t + \tau) - e^{A \tau} \ln \tilde{S}(t) = X(t + \tau) - e^{A \tau} X(t)$$

$$= \int_{t}^{t+\tau} e^{A(t+s)} \Sigma^{1/2}(s) dW(s)$$

$$\approx e^{A \tau} \Sigma^{1/2}(t) \Delta W(t).$$

Here, $\Delta W(t) \triangleq W(t + \tau) - W(t)$ and $\tau$ is supposed to be sufficiently small in order to make the approximation above reasonable. Hence, we have that “reversion-adjusted” logreturns are approximately distributed according to the multivariate mean-variance mixture model

$$e^{A \tau} \Sigma^{1/2}(t) \Delta W(t) \mid_{\Sigma(t)} \sim \mathcal{N}(0, e^{A \tau} \Sigma(t) e^{A^{T} \tau}).$$

In Benth [7], 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 [8] 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 $Y_i(t), i = 1, \ldots, m$ 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 (4.1) 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.

Let us discuss possible specifications of our spot price model satisfying the fundamental conditions on the operators and matrices in question. First of all, it is easily seen that if either one
or both of the matrices $A$ and $C_j$ are diagonal, then they will commute. In fact, supposing that $A$ is a diagonal matrix could be natural in view of interpreting the speed of mean reversion of each commodity modelled separately (as the corresponding entry on the diagonal), and not imposing any functional cross dependencies between the commodities. In such a model, dependencies will enter via the spike terms and in the stochastic volatility. If $A$ is diagonal, then all the diagonal elements must be negative in order to have negative eigenvalues (eigenvalues are equal to the diagonal elements, of course). It is simple to see that the determinant of $A \otimes I + I \otimes A$ becomes

$$\det(A \otimes I + I \otimes A) = 2^d \prod_{i=1}^{d} a_i \prod_{i \neq j}^{d} (a_i + a_j)^2,$$

which is unequal to zero since all the diagonal elements are strictly less than zero. This means that $A$ is invertible. In fact, if we suppose that $C_j$ also is diagonal, one finds that $A - C_j$ is invertible if and only if $a_i + a_j \neq c_i + c_j$ for $i, j = 1, \ldots, d$. Note also that stationarity of the volatility holds only if all the diagonals of $C_j$ are strictly negative. But this also implies that $C_j$ is invertible.

5 Simulation of matrix-valued subordinators

In this section we discuss simulation of our spot price dynamics, which essentially means to discuss simulation of matrix-valued subordinators.

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 [14]). 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 makes it even more complicated. We introduce a simple approximative algorithm\(^2\) to simulate from matrix-valued compound Poisson, stable, and tempered stable processes with stable or constant jump-size distribution.

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

Suppose that $\nu$ is a Lévy measure on $S_+^d$ of the subordinator $L$, such that it can be decomposed into

$$\nu(dU) = \Gamma(d\Theta)\tilde{\nu}(\Theta, dr), \quad U \in S_+^d,$$

where $\tilde{\nu}(\Theta, dr)$ is a Lévy measure on $\mathbb{R}_+$ and $\Gamma$ is a spectral measure on $SS_{+}^d$ concentrated on a finite number of points $\{\Theta_i\}_{1 \leq i \leq p}$. Note in passing that any measure can be approximated by a

\(^2\)The idea of the algorithm was kindly proposed to us by Robert Stelzer.
measure concentrated on finitely many points. Since \( L \) is a pure-jump subordinator, its cumulant function is given by

\[
\phi_L(1)(V) = \int_{S_d^+ \setminus \{0\}} \left( e^{itr(VU)} - 1 \right) \nu(dU) \\
= \int_{SS_d^+} \int_0^\infty \left( e^{itr(V\Theta)} - 1 \right) \tilde{\nu}(\Theta, dr) \Gamma(d\Theta) \\
= \sum_{i=1}^p \Gamma(\Theta_i) \int_0^\infty \left( e^{itr(V\Theta_i)} - 1 \right) \tilde{\nu}(\Theta_i, dr).
\]

One recognizes this as the cumulant of a weighted sum of \( p \) independent real-valued subordinator processes. This leads to the following simple algorithm to sample \( L \) according to its cumulant function:

1. Find the finite set of points \( \{\Theta\}_{1 \leq i \leq n} \) where \( \Gamma \) is concentrated.
2. Simulate \( p \) independent subordinators \( R_i(t) \) with cumulant function

\[
\phi_{R_i}(1)(tr(V\Theta_i)) = \int_0^\infty \left( e^{itr(V\Theta_i)} - 1 \right) \tilde{\nu}(\Theta_i, dr).
\]

3. Set \( L(t) = \sum_{i=1}^p R_i(t)\Theta_i \).

To make this algorithm operationable, we must be able to sample the \( R_i \)'s, which we now discuss in particular cases which are of interest in energy markets.

First, let us consider a matrix-valued compound process (mCP) with only positive jumps \( L \). This becomes a multivariate compound Poisson process restricted to values in the symmetric positive definite cone. Its cumulant function is

\[
\phi_{L(1)}(V) = \lambda \int_{S_d^+} \left( e^{itr(VU)} - 1 \right) \nu(dU),
\]

where \( \nu \) is the jump size distribution and \( \lambda \) the intensity. Supposing that \( \nu(dU) = \tilde{\nu}(\Theta, dr)\Gamma(d\Theta) \) with \( \tilde{\nu}(\Theta, dr) \) being a probability distribution on \( \mathbb{R}_+ \) and \( \Gamma(d\Theta) \) for a spectral measure \( \Gamma \) on \( SS_d^+ \), concentrated on finitely many points, it holds

\[
\phi_{L(1)}(V)t = \lambda \sum_{i=1}^p \Gamma(\Theta_i) \int_0^\infty \left( e^{itr(V\Theta_i)} - 1 \right) \tilde{\nu}(\Theta_i, dr).
\]

Hence, \( R_i \) for \( i = 1, \ldots, p \) will follow a one-dimensional compound Poisson process with jump intensity \( \lambda \Gamma(\Theta_i) \) and jump-size distribution \( \tilde{\nu}(\Theta_i, dr) \). The mCP(\( \lambda \)) process \( L \) is represented
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

as a linear combination of angles $\Theta_i$ and radius processes being one-dimensional compound Poisson processes $R_i$, i.e. $L(t) = \sum_{i=1}^{p} R_i(t)\Theta_i$.

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 (Barndorff-Nielsen and Pérez-Abreu [2])

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

In case $\alpha = 1$ then $\nu$ is a Lévy measure of a matrix extension of the inverse Gaussian distribution ($mIG$), where $\Delta \in S^+_d$ and $\Gamma$, a finite measure on $SS^+_d$, are parameters. As in the univariate case the inverse Gaussian process is a pure jump process, hence the cumulant function is given by

$$\phi_{L(1)}(V) = \int_{SS^+_d} \Gamma(\Theta_i) \int_{0}^{\infty} (e^{irtr(V\Theta_i)} - 1) e^{-rtr(\Delta \Theta_i)} \frac{dr}{r^{3/2}} + itr(V\mu_0),$$

for $L(1) \sim mIG(\Delta, \Gamma, \mu_0)$, where $\mu_0 \in S^+_d$ is a parameter. Choosing $\Gamma$ such that it is concentrated on finitely many point and decomposing $\mu_0$ in an angle $\Theta_0 \in SS^+_d$ and a radius $r_0 \in \mathbb{R}$ leads to

$$\phi_{L(1)}(V) = \sum_{i=1}^{p} \Gamma(\Theta_i) \int_{0}^{\infty} (e^{irtr(V\Theta_i)} - 1) e^{-rtr(\Delta \Theta_i)} \frac{dr}{r^{3/2}} + irrtr(\Phi\Theta_0).$$

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

$$\phi_G(\zeta) = \frac{\delta}{\gamma} (2N(\gamma) - 1) \zeta + \frac{\delta}{\sqrt{2\pi}} \int_{0}^{\infty} (e^{\zeta x} - 1) e^{-1/2\gamma^2 x} \frac{dx}{x^{3/2}}, \quad \zeta \in \mathbb{R},$$

where $N$ denotes the cumulative normal distribution. We recognize $L$ as a matrix of linear combinations of a finite number of angles $\Theta_i$, $i = 1, \ldots, p$ with coefficients given by one-dimensional inverse Gaussian subordinator processes $R_i(t)$, where $R_i(1)$ is distributed according to the inverse Gaussian distribution $IG(\delta_i, \gamma_i)$, where $\delta_i = \sqrt{2\pi} \Gamma(\Theta_i)$ and $\gamma_i = 2\sqrt{tr(\Delta \Theta_i)}$. 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 $\sum_{i=1}^{p} R_i(t)\Theta_i$.

As an example, consider the case of two spot prices $S_1(t)$ and $S_2(t)$ modelled by our dynamics. For example, we could think of the spot price of electricity in two interconnected markets, or the spot price of gas and electricity. We suppose that the prices are driven by two $M_2(\mathbb{R})$-valued subordinator processes $\tilde{L}_1(t), \tilde{L}_2(t)$. The first process defines the spike component, while the second part is determining the stochastic volatility. We assume that there is one spike compo-
Y(t) ∈ \mathbb{R}^2, while the stochastic volatility process \( \Sigma(t) \) is the equally weighted sum of two processes \( Z_1(t) \) and \( Z_2(t) \), where the dynamics is driven by \( \tilde{L}_1 \) and \( \tilde{L}_2 \), resp. The dynamics of the spike process \( Y(t) \) is driven by the diagonal of \( \tilde{L}_1(t) \). In order to make simulations from the model, we use specifications of the parameters in the model inspired by Vos [31], where the BNS stochastic volatility model was estimated to stock price data observed on the Dutch stock exchange. For simplification, we set the seasonality function equal to one, that is, \( \Lambda_i(t) = 1 \) for \( i = 1, 2 \). Moreover, we choose
\[
A = \begin{pmatrix} -1.4 & -0.3 \\ -0.3 & -1.4 \end{pmatrix} \quad B = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \quad \eta = \begin{pmatrix} 1 \\ 0.5 \\ 0.5 \\ 1 \end{pmatrix}
\]
\[
C_1 = \begin{pmatrix} -0.4 & 0.3 \\ 0.3 & -0.4 \end{pmatrix} \quad C_2 = \begin{pmatrix} -0.045 & 0.03 \\ 0.03 & -0.045 \end{pmatrix}
\]
We let the levels of the spike component \( Y(t) \) be zero, \( \mu_1 = \mu_2 = 0 \).

Next, let us define the two subordinator processes \( \tilde{L}_1(t) \) and \( \tilde{L}_2(t) \). To mimic spikes in the market, we consider a simple Poisson process for \( \tilde{L}_1(t) \). To have a stochastic volatility process which can generate adjusted returns being close to NIG distributed, we suppose that \( \tilde{L}_2(t) \) is mIG. In order to be able to simulate these two processes, we apply the idea above, and define a simple discrete spectral measure on \( S^+ \). It is simple to see that
\[
\Theta = \begin{pmatrix} \theta \\ \pm \sqrt{\theta(1-\theta)} \\ \pm \sqrt{\theta(1-\theta)} \\ 1-\theta \end{pmatrix}
\]
or
\[
\begin{pmatrix} \theta \\ 0 \\ 0 \\ \sqrt{1-\theta^2} \end{pmatrix}
\]
for \( \theta \in (0, 1) \). There are three valid choices of \( \Theta \in S^+_2 \). To this end, we discretize the unit interval with step size 0.1, and choose \( \theta_j = j \times 0.1 \) for \( j = 1, \ldots, 9 \). We choose either one of the three possible matrix structures for \( \Theta \) with given \( \theta_i \), making up a total of 27 matrices \( \Theta_i \). For the Poisson process, we choose the intensity such that \( \lambda \Gamma(\Theta_i) = 3/100 \) and the jump size distribution set fixed to be 1.7, that is, if \( R_i(t) \) is jumping at time \( t \), then \( \Delta R_i(t) = 1.7 \). This will correspond to a change in spot price of a factor \( \exp(1.7) = 5.47 \), which is a rather dramatic price change. As a measure for the mIG part, we set \( \Gamma(\Theta_i) = 1/324\sqrt{2\pi} \) uniformly for all \( 1 \leq i \leq 27 \). Finally, we suppose that the parameter \( \Delta \) of the mIG part is
\[
\Delta = \begin{pmatrix} 50 & 45 \\ 45 & 50 \end{pmatrix}
\]
In Figure 1 the spot price series resulting from our 2-dimensional example is shown, where we have used an Euler scheme to discretize the dynamics in time and standard schemes for the sampling of inverse Gaussian distributions (see Rydberg [25]). One clearly can see the dependency between the two spot prices, in particular, how the spikes follow each other in the two series.
Chapter 3. Cross-commodity spot price modeling with stochastic volatility and leverage for energy markets

Figure 3.1: Simulated spot prices of the two commodities

6 Conclusions

We have proposed a model to describe the spot price dynamics for cross-commodity markets in a multivariate setting. The model captures features like mean-reversion, spikes, stochastic volatility, and inverse leverage effect. The dynamic is a multi-dimensional extension of the Barndorff-Nielsen and Shephard stochastic volatility model embedded into mean-reversion dynamics. This is relevant for commodity price series. The choice of the multi-dimensional extension is influenced by the work of Stelzer [29]. The multivariate spot model is analytically tractable and probabilistic properties can to a large extent 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 [7].

A simple algorithm to simulate from matrix-valued subordinators is introduced. The method is demonstrated on an empirical 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.

No methods exist 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 [5]. 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 of 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 the parameters directly using the characteristic function in the frequency domain.
Bibliography


Bibliography


Chapter 4

Pricing of forwards and options in a multivariate non-Gaussian stochastic volatility model for energy markets

Fred Espen Benth and Linda Vos
Abstract

In Benth and Vos [7] we introduced a multivariate spot price model with stochastic volatility for energy markets which captures characteristic features like price spikes, mean-reversion, stochastic volatility and inverse leverage effect as well as dependencies between commodities. In this paper we derive the forward price dynamics based on our multivariate spot price model, providing a very flexible structure for the forward curves, including contango, backwardation and hump shape. Moreover, a Fourier transform-based method to price options on the forward is described.

1 Introduction

The last decades the energy markets have been liberalized world-wide, resulting in market-places for commodities such as electricity, gas and coal. There are several markets for each of these commodities, geographically spread over the continents. For example in Europe we have markets for power in the UK, Germany, France, and the Nordic countries, to mention a few. There are transmission lines which interconnect these markets for electricity. Furthermore, since coal and gas are used to a large extent as fuels for power production, the prices for these commodities naturally affect the power prices. These markets become more and more integrated, both within one commodity, but also across the commodities. For this reason there is an increasing interest in studying multivariate models for energy markets, including cross-commodity models (like for example for gas, coal and electricity-markets) or multivariate models for the same commodity traded in different, but integrated markets (like for example the power markets in the Nordic countries and Germany).

In Benth and Vos [7] we propose stochastic dynamics for cross-commodity spot price modelling generalizing the univariate dynamics studied in Benth [5]. The model is flexible enough to capture spikes and mean-reversion. Moreover, it includes the possibility to model inverse leverage and stochastic volatility. The proposed dynamics can model co- and independent jump behaviour (spikes) in cross-commodity markets, and is analytically tractable. We apply the multivariate extension of the stochastic volatility model of Barndorff-Nielsen and Shephard [3], analysed in detail by Pigorsch and Stelzer [21]. The mean-reverting features of our spot model require a significant extension of their analysis.

In this paper we derive the forward dynamics using a no-arbitrage pricing. Despite the rather general nature of our spot model, the dynamics of the 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 [12], 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 curve can be achieved. As an implication of the stationary properties of the spot model, the forward prices in the long-end of the forward curve (far until
maturity) will move deterministically. The Samuelson effect can be identified in the forward dynamics as well.

By using Fourier methods, options on spreads between different forward contracts can be represented as integrals which can be computed efficiently. Spread options are traded in various energy markets, mostly over-the-counter. However, such options are also used in valuation of new power plant projects and the construction of interconnecting pipelines between different markets. In fact, the construction of a new pipeline connecting two markets can be viewed as a long term spread option. On the other hand, the value of a gas-fired power plant can be represented as a spread between electricity and gas (so-called spark spread).

The paper is organized as follows. Section 2 recalls the spot model proposed in Benth and Vos [7]. Next, in Section 3, the implied multivariate forward dynamics are derived and properties of the forward curve are analysed. Methods based on the Fourier transform are applied to cross-commodity option pricing in Section 4, including special attention to spread options. Finally, in Section 5, we conclude.

2 A cross-commodity energy spot price model with stochastic volatility

In this section we recall briefly the main aspects of the spot model with stochastic volatility for cross-commodity energy markets introduced in Benth and Vos [7]. We suppose that we are given a complete filtered probability space \((\Omega, \mathcal{F}, P)\) equipped with the filtration \(\{\mathcal{F}_t\}_{t \geq 0}\) satisfying the usual conditions (see e.g. Protter [22]).

Assume that \(m \leq n \in \mathbb{N}\), and for \(d \in \mathbb{N}\), consider the \(d\)-dimensional spot price dynamics as a combination of a seasonality function \(\Lambda\), stochastic processes \(\{Y_i\}_{i=1}^m\) modeling spikes and a stochastic process \(X\) modeling the "normal" variations of the price evolution. Here, the seasonality and the stochastic processes \(X\) and \(\{Y_i\}_{i=1}^m\) are all \(d\)-dimensional. More precisely, we define the spot price dynamics of \(d\) energy commodities as follows:

\[
S(t) = \Lambda(t) \cdot \exp \left( X(t) + \sum_{i=1}^{m} Y_i(t) \right). \tag{2.1}
\]

Here, \(\cdot\) denotes pointwise multiplication, and the seasonality \(\Lambda\) is supposed to be a deterministic bounded measurable function. The stochastic processes \(\{Y_i\}_{i=1}^m\) are \(d\)-dimensional Ornstein-Uhlenbeck processes driven by vector valued subordinators \(\{L_i\}_{i=1}^m\), that is, Lévy processes which are increasing in each of its coordinates (see Barndorff-Nielsen et al. [2]).

\[
dY_i(t) = (\mu_i + B_iY_i(t)) \, dt + \eta_i \, dL_i(t), \tag{2.2}
\]
where $\{\mu_i\}_{i=1}^m$ are vectors in $\mathbb{R}^d$. Furthermore, $\{B_i\}_{i=1}^m$ and $\{\eta_i\}_{i=1}^m$ are elements of $GL_d(\mathbb{R})$, the group of $d \times d$ matrices which are invertible. The entries of $\eta_i$ do not necessarily have to be positive, so although $L_i$ are subordinators the process $Y_i$ may exhibit negative jumps. In electricity markets, say, negative spikes are observed.

The "normal variations" process $X$ is an extension of the Barndorff-Nielsen and Shephard [3] stochastic volatility (BNS SV) model into the multidimensional Ornstein-Uhlenbeck setting. The stochastic process $X$ is defined by the following SDE

$$dX(t) = AX(t)\, dt + \Sigma(t)^{1/2} \, dW(t),$$  \hspace{1cm} (2.3)$$

where $A$ is a matrix in $GL_d(\mathbb{R})$ and $W$ is a standard $d$-dimensional Brownian motion in $\mathbb{R}^d$. The square of the volatility $\Sigma(t)$ is chosen to be a matrix valued stochastic process. More precisely, the stochastic volatility $\Sigma(t)$ is a superposition of positive definite matrix valued Ornstein-Uhlenbeck processes as introduced in Barndorff-Nielsen and Stelzer [4],

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

with

$$dZ_j(t) = (C_j Z_j(t) + Z_j(t)C_j^T)\, dt + d\tilde{L}_j(t),$$  \hspace{1cm} (2.5)$$

and the $\omega_j$'s are positive weights summing up to 1. Moreover, for $j = 1, \ldots, n$, $C_j \in GL_d(\mathbb{R})$ and $\tilde{L}_j$ are independent matrix valued subordinators, that is, independent increment processes with values in $\mathbb{S}_d^+$, the positive definite cone of symmetric $d \times d$ matrices. Naturally, $\tilde{L}_j$ are independent of $W$ for $j = 1, \ldots, n$, and we suppose for convenience that the subordinators are driftless. In order to have the Itô integrals in (2.3) well-defined, we suppose that

$$P\left(\int_0^{\tilde{T}} tr(\Sigma(t)) \, dt < \infty\right) = 1.$$

(2.6)$$

Here, $\tilde{T} < \infty$ is some finite horizon time for our energy markets, and $tr$ is the trace operator on matrices. We assume that the eigenvalues of $C_j$ have negative real parts, a necessary condition for ensuring stationarity of the $Z_j$’s. We denote by $\nu_{\tilde{L}_j}$ the Lévy measure of $\tilde{L}_j$, $j = 1, \ldots, n$.

The processes $X, Y_i$ are Ornstein-Uhlenbeck processes. Applying the multi-dimensional Itô Formula (see Ikeda and Watanabe [16]) yields the following explicit dynamics: for $0 \leq s \leq t$,

$$X(t) = e^{A(t-s)}X(s) + \int_s^t e^{A(t-u)}\Sigma(u)^{1/2} \, dW(u),$$  \hspace{1cm} (2.7)$$

$$Y_i(t) = e^{B_i(t-s)}Y_i(s) + B_i^{-1}(I - e^{B_i(t-s)})\mu_i + \int_s^t e^{B_i(t-u)}\eta_i \, dL_i(u),$$  \hspace{1cm} (2.8)$$
for \(i = 1, \ldots, m\). The matrix exponentials are defined as usual as \(e^{A} := I + \sum_{i=1}^{\infty} \frac{A^n}{n!}\).

According to Barndorff-Nielsen and Stelzer [4], Sect. 4, the solution of \(Z_j(t), j = 1, \ldots, n\), is given by

\[
Z_j(t) = e^{G_j(t-s)}Z_j(s) + \int_s^t e^{G_j(t-u)}d\tilde{L}_j(u)e^{G_j(t-u)}. \tag{2.9}
\]

The matrix-valued stochastic integral in the second term of \(Z_j(t)\) is understood as follows: let \(M_d(\mathbb{R})\) be the space of real \(d \times d\) matrices. For two \(M_d(\mathbb{R})\)-valued bounded and measurable functions \(E(u)\) and \(F(u)\) on \([s, t]\), the notation \(\int_s^t E(u) dL(u)F(u)\) means the matrix \(G(s, t) \in M_d(\mathbb{R})\) with coordinates defined by

\[
G_{ij}(s, t) = \sum_{k=1}^{d} \sum_{l=1}^{d} \int_s^t E_{ik}(u)F_{lj}(u) d\tilde{L}_{kl}(u).
\]

Here, \(\tilde{L}\) is the generic notation for some \(\tilde{L}_j\). We remark that since \(\tilde{L}_j\) are supposed to be RCLL, the processes \(Z_j\) also are RCLL.

In energy markets like gas and electricity it is often observed that a spike and an increase in volatility occur at the same time. This is known as the inverse leverage effect. To model this phenomenon we take the vector valued subordinators \(L_i\) driving the processes \(Y_i, i = 1, \ldots, m\), as the diagonal entries of the \(m\) first matrix valued subordinators \(\tilde{L}_j, j = 1, \ldots, m\). 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 \(\mathbb{S}_d^+\). Such a modelling choice ensures that the volatility jumps simultaneously with a spike in the spot price process. Since \(n \geq m \in \mathbb{N}\), and the volatility process is a weighted sum of \(n\) different volatility processes, there are still \(n-m\) volatility processes \(Z_j, j = m+1, \ldots, n\) which can be freely chosen.

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:

\[
S(t) = \Lambda(t) \cdot \exp(X(t)) \tag{2.10}
\]

where \(X(t)\) is defined in (2.3). The Schwartz model with constant volatility is a mean reversion process proposed by Schwartz [24] for spot price dynamics in commodity markets like oil.

To ensure solutions to the SDE’s (2.2) and (2.3) we impose the following log integrability conditions on the subordinators: for \(j = 1, \ldots, n\), it holds that

\[
\mathbb{E} \left[ \log^+ \|\tilde{L}_j(1)\| \right] < \infty, \tag{2.11}
\]

where \(\log^+(x)\) is defined as \(\max(\log(x), 0)\). We use the Frobenius norm for matrices, \(\|A\| = \text{tr}(A^T A)^{1/2}\), \(A \in M_d(\mathbb{R})\).

For a detailed analysis of this spot price model for cross-commodity energy markets, we refer
3 Forward pricing

In commodity markets, forward contracts are commonly traded on exchanges, including power, gas, oil, coal, etc. In this Section we derive the forward price dynamics based on the multivariate spot price model (2.1).

Appealing to general arbitrage theory, we define the forward price \( F(t, \tau) \) at time \( t \) for contracts delivering the energy commodity at time \( \tau \) by (see e.g. Duffie [9])

\[
F(t, \tau) = \mathbb{E}_Q [S(\tau) \mid \mathcal{F}_t],
\]

where \( Q \) is a risk-neutral probability measure. This definition is valid as long as \( S(\tau) \in L^1(Q) \). Below we give sufficient conditions ensuring integrability of the spot price with respect to a parametric class of pricing measures \( Q \). 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 situation here, however the calculations can be easily adjusted to take this into account (see for example Benth et al. [6] for a discussion).

The stochastic volatility model we are discussing gives rise to an incomplete market, and hence there exists a continuum of equivalent martingale measures \( Q \) that can be used for pricing. Moreover, in energy markets, the underlying spot is in general not tradeable, due to for example 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 forward price fails. As a result, all equivalent measures \( Q \sim P \) may be chosen as pricing measures since the underlying spot is not directly tradeable. In our considerations, we do not require the martingale property under \( Q \) for discounted spot prices. We refer to Benth et al. [6] for more on this.
3.1 A class of equivalent probabilities

A convenient way to define a parametric class of risk-neutral probabilities for Lévy-based models is the Esscher transform (see Benth et al. [6] for applications of the Esscher transform in energy markets). Before introducing the measure transform, we need to introduce some notation and state some conditions: for $V \in \mathbb{S}_d^+$ we let $\phi_{L_j}(V)$ be the cumulant function of $\tilde{L}_j(1)$, that is,

$$
\phi_{L_j}(V) = \ln \mathbb{E} \left[ \exp \left( i \text{tr} (V \tilde{L}_j(1)) \right) \right]. \tag{3.2}
$$

The Esscher transform is defined via the logarithmic moment generating functions of $\tilde{L}_j$, and for this purpose we need to have certain exponential moments existing for $\tilde{L}_j$. Let $\Theta_j \in \mathbb{S}_d^+$, and suppose that $\phi_{L_j}(-i\Theta_j)$ is well-defined. We have that

$$
\phi_{L_j}(-i\Theta_j) = \int_{\mathbb{S}_d^+} \{ e^{tr(\Theta_j U)} - 1 \} \nu_{\tilde{L}_j}(dU),
$$

and therefore, $\phi_{L_j}(-i\Theta_j)$ is well-defined as long as

$$
\int_{\mathbb{S}_d^+} \{ e^{tr(\Theta_j U)} - 1 \} \nu_{\tilde{L}_j}(dU) < \infty. \tag{3.3}
$$

Note that for $U, V \in \mathbb{S}_d^+$, $tr(UV) = \langle U, V \rangle$, the inner product associated with the Frobenius matrix norm $\|A\| := tr(A^T A)^{1/2}$. Hence, we have the inequality $|tr(UV)| \leq \|U\|\|V\|$. Thus, a sufficient condition for (3.3) to hold is that

$$
\int_{\mathbb{S}_d^+} e^{tr(\Theta_j U)} \nu_{\tilde{L}_j}(dU) \leq \int_{\mathbb{S}_d^+} e^{\|\Theta_j\|\|U\|} \nu_{\tilde{L}_j}(dU) < \infty.
$$

Throughout this paper we suppose that there exists a constant $c_j > 0$ such that the following exponential integrability condition holds for $\nu_{\tilde{L}_j}$:

$$
\int_{\mathbb{S}_d^+} e^{c_j\|U\|} \nu_{\tilde{L}_j}(dU) < \infty, \tag{3.4}
$$

for $j = 1, \ldots, n$. This condition implies that $\phi_{L_j}(-i\Theta_j)$ is well-defined for all $\Theta_j \in \mathbb{S}_d^+$ such that $\|\Theta_j\| \leq c_j$.

We move on to define the equivalent probability measure $Q$. For $\Theta_j \in \mathbb{S}_d^+$, such that $\|\Theta_j\| \leq c_j$, define the processes

$$
\mathcal{V}_j(t) = \exp \left( tr(\Theta_j \tilde{L}_j(t)) - \phi_{L_j}^j(-i\Theta_j)t \right), \tag{3.5}
$$
for \( j = 1, \ldots, n \) and \( t \leq \tilde{T} \). Here we recall \( \tilde{T} \) to be a finite time horizon of the market for which all delivery times \( \tau \) of interest are included. Note that \( V_j(t) \) are martingales for \( j = 1, \ldots, m \); in fact, by the exponential moment condition in (3.4) we find that

\[
\mathbb{E}[V_j(t)] = 1,
\]

for every \( j = 1, \ldots, n \). For a vector \( \theta_0 \in \mathbb{R}^d \), introduce the process

\[
V_0(t) = \exp \left( - \int_0^t \theta_0^T \Sigma^{-1/2}(s) dW(s) - \frac{1}{2} \theta_0^T \int_0^t \Sigma^{-1}(s) \theta_0 \, ds \right). \tag{3.6}
\]

We have the following Lemma:

**Lemma 3.1.** For all \( \theta_0 \in \mathbb{R}^d \), the process \( V_0(t) \) for \( t \leq \tilde{T} \) is a martingale.

**Proof.** We show that the Novikov condition holds. From (2.9) we have for every \( j = 1, \ldots, n \) and any \( x \in \mathbb{R}^d \)

\[
x^T Z_j(t) x = x^T e^{C_j t} Z_j(0) e^{C_j^T t} x + x^T \int_0^t e^{C_j (t-u)} dL_j(u) e^{C_j^T (t-u)} x
\]

\[
\geq x^T e^{C_j (t-s)} Z_j(s) e^{C_j^T (t-s)} x
\]

by positive definiteness of the stochastic integral term. Hence,

\[
\Sigma(t) = \sum_{j=1}^n \omega_j Z_j(t) \geq \sum_{j=1}^n \omega_j e^{C_j (t-s)} Z_j(s) e^{C_j^T (t-s)} > 0.
\]

But then, from linear algebra on positive definite matrices,

\[
\Sigma^{-1}(t) \leq \left( \sum_{j=1}^n \omega_j e^{-C_j t} Z_j^{-1}(0) e^{-C_j^T t} \right)^{-1},
\]

which means in particular

\[
\theta_0^T \Sigma^{-1}(t) \theta_0 \leq \theta_0^T \left( \sum_{j=1}^n \omega_j e^{-C_j t} Z_j^{-1}(0) e^{-C_j^T t} \right)^{-1} \theta_0.
\]

As the right-hand side is a continuous function in \( t \) on \( [0, \tilde{T}] \), it follows that

\[
\mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^{\tilde{T}} \theta_0^T \Sigma^{-1}(t) \theta_0 \, dt \right) \right]
\]
Chapter 4. Pricing of forwards and options in a multivariate non-Gaussian stochastic volatility model for energy markets

\[ \leq \mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T \theta_0^T \left( \sum_{j=1}^n \omega_j e^{-C_j \lambda W_j^{-1}(0) e^{-C_j T_j}} \right)^{-1} \theta_0 \, dt \right) \right] < \infty. \]

Hence, by Novikov’s condition, it follows from the Girsanov Theorem that \( \mathcal{V}_0(t) \) is a martingale. □

Thus, the process

\[ V(t) = \mathcal{V}_0(t) \times \mathcal{V}_1(t) \times \ldots \times \mathcal{V}_n(t), \tag{3.7} \]

becomes a martingale for \( t \leq \tilde{T} \) and is the density process of a probability measure \( Q \) equivalent with \( P \), that is,

\[ \frac{dQ}{dP} \bigg|_{\mathcal{F}_t} = V(t). \tag{3.8} \]

From Girsanov’s Theorem we find that

\[ \tilde{dW}(t) = dW(t) - \Sigma^{-1/2}(t)\theta_0 \, dt, \tag{3.9} \]

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

**Lemma 3.2.** Assume \( \Theta_j \in \mathbb{S}^+_d \) such that \( \| \Theta_j \| \leq c_j \) for \( j = 1, \ldots, n \). Then \( \tilde{L}_j(t) \) are subordinators under \( Q \) defined in (3.8) having Lévy measure with respect to \( Q \) given by

\[ \nu^{Q}_{L_j}(dU) = \exp(tr(\Theta_j U))\nu_{L_j}(dU), \]

for \( j = 1, \ldots, n \).

**Proof.** First we prove that \( \tilde{L}_j(t) \) is a matrix-valued subordinator under \( Q \). Consider its conditional cumulant function with respect to \( Q \), \( \phi^{(s,t)}_{L_j}(V) \): for \( 0 \leq s \leq t \) and using Bayes’ Formula for conditional expectations (see Karatzas and Shreve [18])

\[
\phi^{(s,t)}_{L_j}(V) = \ln \mathbb{E}_Q \left[ \exp \left( i tr(V (\tilde{L}_j(t) - \tilde{L}_j(s))) \right) \bigg| \mathcal{F}_s \right] \\
= \ln \mathbb{E} \left[ \exp \left( i tr(V (\tilde{L}_j(t) - \tilde{L}_j(s))) \right) \frac{V(t)}{V(s)} \bigg| \mathcal{F}_s \right] \\
= \ln \mathbb{E} \left[ \exp \left( i tr((V - i\Theta_j)\tilde{L}_j(1))) \bigg| \mathcal{F}_s \right] - \phi_{L_j}(-i\Theta_j)(t-s) \\
= \ln \mathbb{E} \left[ \exp \left( i tr((V - i\Theta_j)\tilde{L}_j(1))) \bigg| \mathcal{F}_s \right] - \phi_{L_j}(-i\Theta_j)(t-s) \\
= \phi_{L_j}(V - i\Theta_j)(t-s) - \phi_{L_j}(-i\Theta_j)(t-s). \]

74
In the second to the last equality we used the independent increment property of \( \tilde{L}_j(t) \). This proves that the increment \( \tilde{L}_j(t) - \tilde{L}_j(s) \) is stationary and independent of \( \mathcal{F}_s \), hence a Lévy process with respect to the probability \( Q \). Moreover, \( \tilde{L}_j(t) \) has values in \( S^+_d \), and therefore it is a subordinator under \( Q \). From the above calculation we find its cumulant under \( Q \) to be

\[
\hat{\phi}_{L_j} := \ln \mathbb{E}_Q \left[ \exp \left( i \text{tr}(V \tilde{L}_j(1)) \right) \right] \\
= \phi_{L_j}(V - i\Theta_j) - \phi_{L_j}(-i\Theta_j) \\
= \int_{S^+_d} \left\{ e^{i\text{tr}(V - i\Theta_j)U} - 1 \right\} \nu_{L_j}(dU) - \int_{S^+_d} \left\{ e^{i\text{tr}(-i\Theta_j)U} - 1 \right\} \nu_{L_j}(dU) \\
= \int_{S^+_d} \left\{ e^{i\text{tr}(VU)} - 1 \right\} e^{i\text{tr}(\Theta_jU)} \nu_{L_j}(dU).
\]

Hence, the Lemma follows. \( \square \)

Since a subordinator is a pure-jump process, we must have that \( \tilde{L}_j \) for \( j = 1, \ldots, m \) are independent of \( \tilde{W} \) with respect to \( Q \), since a Brownian motion has continuous paths.

The parameters \( \theta_0 \) and \( \Theta_j, j = 1, \ldots, n \) may be referred to as the market prices of risk, extending the similar notion in the univariate case (see Benth et al. [6]). Note that the Esscher transform gives an exponential tilting of the Lévy measure of the matrix-valued subordinators \( \tilde{L}_j \). One effect of this is that the probabilities for large jumps are re-scaled, and we may get more or less pronounced large jumps under \( Q \).

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

\[
dX(t) = AX(t) + \Sigma^{1/2}(t) \left( d\tilde{W}(t) + \Sigma^{-1/2}(t)\theta_0 \, dt \right) \\
= (\theta_0 + AX(t)) \, dt + \Sigma^{1/2}(t) \, d\tilde{W}(t).
\] (3.10)

Thus, under \( Q \), the mean-reversion level is shifted from 0 to \( \theta_0 \). If \( e_k^T\theta_0 > 0 \) for a \( k = 1, \ldots, d \) and \( e_k \) being the \( k \)th canonical unit vector of \( \mathbb{R}^d \), then the base component of the \( k \)th commodity mean-reverts towards a higher level under \( Q \) than under \( P \), implying that the market assesses the base component as being more risky under the pricing measure \( Q \). A negative market price of risk \( e_k^T\theta_0 \) will imply less risk loading on the \( k \)th base component. The dynamics of \( Y_i \) and \( Z_j \) are changed in a similar fashion. We have for \( i = 1, \ldots, m \)

\[
dY_i(t) = (\mu_i + B_iY_i(t)) \, dt + \eta_i \, dL_i(t) \\
= (\mu_i + \eta_i \mathbb{E}_Q[L_i(1)] + B_iY_i(t)) \, dt + \eta_i \, dL_i^Q(t),
\] (3.11)

where \( dL_i^Q(t) \triangleq dL_i(t) - \mathbb{E}_Q[L_i(1)] \, dt \) is a \( Q \)-martingale. Hence, the process \( Y_i \) varies around the level \( \mu_i + \eta_i \mathbb{E}_Q[L_i(1)] \) under \( Q \), whereas the level is \( \mu_i + \eta_i \mathbb{E}[L_i(1)] \) under \( P \). Thus, by
appropriately choosing $\Theta_i$, we can obtain a higher or lower mean-reversion level, implying a higher or lower risk loading on the spike processes $Y_i$ under $Q$. Similar considerations hold for the volatility processes $Z_j$. We remark in passing that the market prices of risk $\theta_0, \Theta_1, \ldots, \Theta_n$ will implicitly model the risk premium in the market, being the difference between the forward price and the predicted spot at delivery.

### 3.2 Analysis of forward prices

Before we derive the forward price, we need to introduce some notation and prove an auxiliary result. To this end, let $J_d$ be the linear operator that maps a vector $v \in \mathbb{R}^d$ to a symmetric $d \times d$-matrix $J_d(v)$, consisting of zeros except on the diagonal, which is equal to $v$. On the other hand, $\text{diag}$ is a linear operator mapping a matrix into a vector, where the vector is the diagonal of the matrix.

The family of linear operators $C_j(t)$ for $t \in [0, T]$ are defined as

$$C_j(t) : X \mapsto \omega_j \left[ (C_j - A)^{-1}\left( e^{C_j^t X e^{C_j^T t}} - e^{A^T X e^{A^T t}} \right) \right],$$

for $j = 1, \ldots, n$. For $A$ being an $n \times n$-matrix, we denote the operator $A$ associated with the matrix $A$ as $A : X \mapsto AX + XA^T$. This operator can be represented as $\text{vec}^{-1} \circ ((A \otimes I_n) + (I_n \otimes A)) \circ \text{vec}$, with $I_n$ being the $n \times n$ identity matrix and $\text{vec}$ meaning the operator which stacks the columns of a matrix into a vector. Its inverse is denoted by $A^{-1}$, which exists whenever $I_n \otimes A + A \otimes I_n$ is invertible. In this case, we can represent $A^{-1}$ by $\text{vec}^{-1} \circ ((A \otimes I_n) + (I_n \otimes A))^{-1} \circ \text{vec}$. Remark that $A \otimes I_n + I_n \otimes A$ is equal to the Kronecker sum of the matrix $A$ with itself.

The following auxiliary result is useful in deriving the forward prices, and is proven in Benth and Vos [7].

**Lemma 3.3.** Define $f(s, t) := \int_s^t e^{A(t-u)\Sigma(u)}e^{AT(t-u)}du$. Assume for $j = 1, \ldots, n$ that $A$ and $C_j$ commute and $A - C_j$ are invertible. Then it holds

$$f(s, t) = \sum_{j=1}^n C_j(t-s)Z_j(s) + \int_s^t C_j(t-v)dL_j(v),$$

for $0 \leq s \leq t$.

**Proof.** The proof of this result is found in Benth and Vos [7]. We include it here for the convenience of the reader. Using (2.9) and the assumption that $A$ and $C_j$ commute for $j = 1, \ldots, n$ it holds

$$f(s, t) = \int_s^t e^{A(t-u)} \sum_{j=1}^n \omega_j \left( e^{C_j(u-s)}Z_j(s)e^{C_j^T(u-s)} + \int_s^u e^{C_j(u-v)}dL_j(v)e^{C_j^T(u-v)} \right) e^{A^T(t-u)}du$$

76
Proposition 3.4. For $k = 1, \ldots, d$, suppose $\Theta_j$ are such that

$$\sup_{u \in [0,T]} \frac{1}{2} C^*_j(u)(e_k e_k^T) \| \Theta_j \| \leq c_j$$
Chapter 4. Pricing of forwards and options in a multivariate non-Gaussian stochastic volatility model for energy markets

for \( j = 1, \ldots, n \), and

\[
\sup_{u \in [0, \tilde{T}]} \left\| \frac{1}{2} C_i^*(u)(e_k e_k^T) + J_d(e_k^T e^B_{i u} \eta_i) \right\| + \| \Theta_i \| \leq c_i
\]

for \( i = 1, \ldots, m \). Assume for \( j = 1, \ldots, n \) that \( A \) and \( C_j \) commute and \( A - C_j \) are invertible. Then the forward price at time \( t \geq 0 \) of a contract delivering the \( d \) spots \( S(\tau) \) at time \( \tau \geq t \) is

\[
F(t, \tau) = \Lambda(\tau) \cdot \exp \left( e^{A(\tau-t)} X(t) + \sum_{i=1}^{m} e^{B_i(\tau-t)} Y_i(t) + A^{-1}(I - e^{A(\tau-t)}) \theta_0 \right.
\]

\[
+ \sum_{i=1}^{m} B_i^{-1}(I - e^{B_i(\tau-t)}) \mu_i + \frac{1}{2} \mathrm{diag} \left\{ \sum_{j=1}^{n} C_j(\tau - t) Z_j(t) \right\} \right) \cdot \Psi(\tau - t),
\]

where the \( k \)-th coordinate of \( \Psi(s) \in \mathbb{R}^d \) for \( 0 \leq s \leq \tilde{T} \) is

\[
\ln \Psi_k(s) = \sum_{j=1}^{n} \int_{0}^{s} \left\{ \phi_{L_j} \left( -\frac{1}{2} i C_j^*(u)(e_k e_k^T) - i \Theta_j \right) - \phi_{L_j}(-i \Theta_j) \right\} du
\]

\[
+ \sum_{i=1}^{m} \int_{0}^{s} \left\{ \phi_{L_i} \left( -\frac{1}{2} i C_i^*(u)(e_k e_k^T) - i J_d(e_k^T e^B_{i u} \eta_i) - i \Theta_i \right) \right.
\]

\[
- \phi_{L_i}(-\frac{1}{2} i C_i^*(u)(e_k e_k^T) - i \Theta_i) \} du,
\]

for \( k = 1, \ldots, d \).

Proof. For simplicity, we let \( m = n = 1 \) and defer the subscripts with respect to \( i \) and \( j \). From (2.7) and (2.8) along with the definition of the measure \( Q \), we have

\[
X(\tau) = e^{A(\tau-t)} X(t) + \int_{t}^{\tau} e^{A(\tau-u)} \Sigma^{1/2}(u) dW(u)
\]

\[
= e^{A(\tau-t)} X(t) + \int_{t}^{\tau} e^{A(\tau-u)} \theta_0 du + \int_{t}^{\tau} e^{A(\tau-u)} \Sigma^{1/2}(u) \tilde{dW}(u)
\]

\[
= e^{A(\tau-t)} X(t) + A^{-1}(I - e^{A(\tau-t)}) \theta_0 + \int_{t}^{\tau} e^{A(\tau-u)} \Sigma^{1/2}(u) \tilde{dW}(u),
\]

and

\[
Y(\tau) = e^{B(\tau-t)} Y(t) + B^{-1}(I - e^{B(\tau-t)}) \mu + \int_{t}^{\tau} e^{B(\tau-u)} \eta dL(u).
\]
Hence, using the $\mathcal{F}_t$-adaptedness of $X(t)$ and $Y(t)$, we find

\[
F(t, \tau) = \Lambda(\tau) \cdot \mathbb{E}_Q \left[ \exp(X(\tau) + Y(\tau)) \mid \mathcal{F}_t \right] \\
= \Lambda(\tau) \cdot \exp \left( \mathbf{e}^{A(\tau-t)} X(t) + \mathbf{e}^{B(\tau-t)} Y(t) + A^{-1}(1 - \mathbf{e}^{A(\tau-t)}) \theta_0 + B^{-1}(1 - \mathbf{e}^{B(\tau-t)}) \mu \right) \\
\quad \cdot \mathbb{E}_Q \left[ \exp \left( \int_t^\tau \mathbf{e}^{A(\tau-u)} \Sigma^{1/2}(u) \, d\mathbf{W}(u) + \int_t^\tau \mathbf{e}^{B(\tau-u)} \eta \, dL(u) \right) \mid \mathcal{F}_t \right]
\]

We consider the expectation in the last equality, which we denote by $\hat{F}(t, \tau)$. Let $\mathcal{G}_{t,\tau}$ be the $\sigma$-algebra generated by $\mathcal{F}_t$ and $\hat{L}(u)$ for $t \leq u \leq \tau$. Recalling that under $Q$, $\hat{W}$ and $\hat{L}$ are independent, we find from the tower property of the conditional expectation operator

\[
\hat{F}(t, \tau) = \mathbb{E}_Q \left[ \mathbb{E}_Q \left[ \exp \left( \int_t^\tau \mathbf{e}^{A(\tau-s)} \Sigma^{1/2}(s) \, d\mathbf{W}(s) + \int_t^\tau \mathbf{e}^{B(\tau-s)} \eta \, dL(s) \right) \mid \mathcal{G}_{t,\tau} \right] \mid \mathcal{F}_t \right] \\
= \mathbb{E}_Q \left[ \exp \left( \int_t^\tau \mathbf{e}^{B(\tau-s)} \eta \, dL(s) \right) \right] \cdot \mathbb{E}_Q \left[ \exp \left( \int_t^\tau \mathbf{e}^{A(\tau-s)} \Sigma^{1/2}(s) \, d\mathbf{W}(s) \right) \mid \mathcal{G}_{t,\tau} \right] \mid \mathcal{F}_t \right] \\
= \mathbb{E}_Q \left[ \exp \left( \frac{1}{2} \text{diag} \left[ \int_t^\tau \mathbf{e}^{A(\tau-s)} \Sigma(s) \mathbf{e}^T(\tau-s) \, ds \right] + \int_t^\tau \mathbf{e}^{B(\tau-s)} \eta \, dL(s) \right) \mid \mathcal{F}_t \right].
\]

In the second equality we used that $L$ is measurable with respect to $\mathcal{G}_{t,\tau}$, while in the last equality we applied the facts that the Wiener integral of a deterministic function is independent of $\mathcal{F}_t$ and a Gaussian random variable.

From Lemma 3.3, we find after appealing to the $\mathcal{F}_t$-measurability of $Z(t)$ and the independent increment property of Lévy processes,

\[
\hat{F}(t, \tau) = \mathbb{E}_Q \left[ \exp \left( \frac{1}{2} \text{diag} \left[ C(\tau - t) Z(t) \right] \right) + \frac{1}{2} \text{diag} \left( \int_t^\tau \mathcal{C}(\tau - u) \, d\hat{L}(u) \right) + \int_t^\tau \mathbf{e}^{B(\tau-u)} \eta \, dL(u) \right) \mid \mathcal{F}_t \right] \\
= \exp \left( \frac{1}{2} \text{diag} \left[ C(\tau - t) Z(t) \right] \right) \\
\quad \cdot \mathbb{E}_Q \left[ \exp \left( \frac{1}{2} \text{diag} \left( \int_t^\tau \mathcal{C}(\tau - u) \, d\hat{L}(u) \right) + \int_t^\tau \mathbf{e}^{B(\tau-u)} \eta \, dL(u) \right) \right].
\]

Let us focus on the expectation above, and denote it by $\Psi(t, \tau)$. It is a vector in $\mathbb{R}^d$, and we look at it componentwise. Note that the $k$th coordinate of $\text{diag} \left( \int_t^\tau \mathcal{C}(\tau - u) \, d\hat{L}(u) \right)$ can be expressed as $\mathbf{e}_k^T \int_t^\tau \mathcal{C}(\tau - u) \, d\hat{L}(u) \mathbf{e}_k$, while the $l$th coordinate of $\int_t^\tau \mathbf{e}^{B(\tau-u)} \eta \, dL(u)$ is $\mathbf{e}_l^T \int_t^\tau \mathbf{e}^{B(\tau-u)} \eta \, dL(u)$. Hence, from the fundamental relation $w^k U w = tr(w w^k A)$ for a vector $w$ and a matrix $U$,

\[
\Psi_k(t, \tau) = \mathbb{E}_Q \left[ \exp \left( \frac{1}{2} \mathbf{e}_k^T \int_t^\tau \mathcal{C}(\tau - u) \, d\hat{L}(u) \mathbf{e}_k + \mathbf{e}_k^T \int_t^\tau \mathbf{e}^{B(\tau-u)} \eta \, dL(u) \right) \right] \\
= \mathbb{E}_Q \left[ \exp \left( \frac{1}{2} \text{tr} \left( \frac{1}{2} \mathbf{e}_k \mathbf{e}_k^T \int_t^\tau \mathcal{C}(\tau - u) \, d\hat{L}(u) \right) + \mathbf{e}_k^T \int_t^\tau \mathbf{e}^{B(\tau-u)} \eta \, dL(u) \right) \right]
\]
Chapter 4. Pricing of forwards and options in a multivariate non-Gaussian stochastic volatility model for energy markets

Note that \( e_k^T e^{B(\tau-u)} \eta \) is a \( d \)-dimensional vector. It is simple to see that

\[
\int_t^\tau e_k^T e^{B(\tau-u)} \eta \, dL(u) = tr(\int_t^\tau J_d(e_k^T e^{B(\tau-u)} \eta) \, d\tilde{L}(u)).
\]

Hence,

\[
\Psi_k(t,\tau) = \mathbb{E}_Q \left[ \exp \left( itr \left( -\frac{1}{2} \int_t^\tau e_k^T e_k^T C(\tau-u) \, d\tilde{L}(u) \right) + itr \left( -i \int_t^\tau J_d(e_k^T e^{B(\tau-u)} \eta) \, d\tilde{L}(u) \right) \right) \right] = \mathbb{E} \left[ \exp \left( itr \left( \left\{ -\frac{1}{2} i e_k^T e_k^T C(\tau-u) - i J_d(e_k^T e^{B(\tau-u)} \eta) \right\} \, d\tilde{L}(u) \right) \right) \right]
\times \exp \left( -\phi_{\tilde{E}}(-i\Theta) \right)
\]

Next, observe that the stochastic integral can be expressed as

\[
\int_t^\tau \left\{ \frac{1}{2} e_k^T e^T C(\tau-u) + J_d(e_k^T e^{B(\tau-u)} \eta) \right\} \, d\tilde{L}(u) = \lim_{|\Delta|\to 0} \sum_{i=0}^{n-1} \left\{ \frac{1}{2} e_k^T e_k^T C(\tau-u_i) + J_d(e_k^T e^{B(\tau-u)} \eta) \right\} \Delta\tilde{L}(u_i),
\]

for partitions \( t = u_0 < \cdots < u_n = \tau \) with \( \Delta_i := \tilde{L}(u_{i+1}) - \tilde{L}(u_i) \) and \( \Delta_i := u_{i+1} - u_i \).

By independence of increments of a Lévy process, and continuity of the exponential function together with Fubini-Tonelli’s Theorem, we get

\[
\mathbb{E} \left[ \exp \left( itr \left( \int_t^\tau \left\{ -\frac{1}{2} i e_k^T e_k^T C(\tau-u) - i J_d(e_k^T e^{B(\tau-u)} \eta) \right\} \, d\tilde{L}(u) \right) \right) \right] = \lim_{|\Delta|\to 0} \prod_{i=1}^{n-1} \mathbb{E} \left[ \exp \left( itr \left( \left\{ -\frac{1}{2} i e_k^T e_k^T C_j(\tau-u) - i J_d(e_k^T e^{B(\tau-u)} \eta) \right\} \, \Delta\tilde{L}(u_i) \right) \right) \right].
\]

Now, the linear operators \( C(\tau-u_i) \) can be represented as \( \text{vec}^{-1} \circ \mathcal{K}(\tau-u_i) \circ \text{vec} \) for a matrix \( \mathcal{K} \in \mathbb{R}^{d^2 \times d^2} \). Hence, since for quadratic matrices \( tr(VX) = \text{vec}(V)^T \text{vec}(X) \), we find

\[
tr \left( (e_k^T e_k^T) C(\tau-u_i) \Delta\tilde{L}(u_i) \right) = \text{vec}(e_k^T e_k^T)^T \text{vec} \left( (\tau-u_i) \Delta\tilde{L}(u_i) \right) = \text{vec}(e_k^T e_k^T)^T \text{vec} \left( \text{vec}^{-1} \mathcal{K}(\tau-u_i) \text{vec}(\Delta\tilde{L}(u_i)) \right)
= \text{vec}(e_k^T e_k^T)^T \mathcal{K}(\tau-u_i) \text{vec}(\Delta\tilde{L}(u_i)) = (\mathcal{K}^T (\tau-u_i) \text{vec}(e_k^T e_k^T))^T \text{vec}(\Delta\tilde{L}(u_i))
= tr \left( \text{vec}^{-1} (\mathcal{K}^T (\tau-u_i) \text{vec}(e_k^T e_k^T)) \Delta\tilde{L}(u_i) \right)
\]
\[
= \text{tr} \left( C^* (\tau - u_i) (e_k e_k^T) \Delta \tilde{L}(u_i) \right).
\]

Thus,
\[
\mathbb{E} \left[ \exp \left( i \text{tr} \left\{ \left\{ -\frac{1}{2} i e_k e_k^T C_j (\tau - u_i) - i J_d (e_k^T e B (\tau - u_i) \eta) \right\} \Delta \tilde{L}(u_i) \right\} \right) \right] = \exp \left( \phi_L \left( -\frac{1}{2} i C^* (\tau - u_i) (e_k e_k^T) - i J_d (e_k^T e B (\tau - u_i) \eta) \right) \Delta u_i \right).
\]

Gathering information, we find that
\[
\ln \Psi_k(t, \tau) = \int_t^\tau \left\{ \phi_L \left( -\frac{1}{2} i C^* (\tau - u) (e_k e_k^T) - i J_d (e_k^T e B (\tau - u) \eta) - i \Theta \right) - \phi_L(-i \Theta) \right\} \, du
\]

By changing variables we see that \( \Psi_k \) depends on \( \tau - t \). This completes the proof.

The forward price \( F(t, \tau) \) gives us the joint dynamics of forward prices on each of the spot commodities. Hence, it is a \( d \)-variate process, giving the cross-commodity forward price dynamics. Recall that \(-\) denotes the pointwise product, and that we use the notation for the exponential function interchangeably, in the sense that \( \exp(x) \) means elementwise exponentiation as long as \( x \) is a vector, and the matrix exponential when \( x \) is a matrix.

Note that since \( C_j(0) = 0 \) and \( \Psi_k(0) = 1 \) for \( k = 1, \ldots, d \), it is easily seen that the expression for \( F(t, \tau) \) is equal to \( S(t) \) when \( \tau = t \). This shows that the forward price converges to the spot at maturity, which it should by definition of the forward price as the conditional expectation of the spot at maturity. More interestingly is that the forward price dynamics is explicitly dependent on the stochastic volatility factors \( Z_j(t) \). This has the interesting effect that even in the case of no spike components in the spot dynamics (i.e., when \( m = 0 \)), the forward price dynamics will have jumps. That is, continuous spot price dynamics with stochastic volatility will imply forward price dynamics which jumps according to the jumps in the stochastic volatility.

We state the dynamics of the forward price.

**Proposition 3.5.** Suppose the conditions in Prop. 3.4 holds. Then the dynamics of \( F_k(t, \tau) \) of commodity \( k \) with respect to \( Q \) is

\[
\frac{dF_k(t, \tau)}{F_k(t-, \tau)} = e_k^T e A(\tau-t) \Sigma^{1/2}(t) \, d\tilde{W}(t)
\]

\[
+ \sum_{i=1}^m \int_{S^+_n \setminus \{0\}} \left\{ \exp \left( \frac{1}{2} e_k^T \text{diag}(C_i(\tau - t) V) + e_k^T e B_i(\tau - t) \eta_i \text{diag}(V) \right) - 1 \right\} \tilde{N}_i^Q(dt, dV)
\]

\[
+ \sum_{j=m+1}^n \int_{S^+_n \setminus \{0\}} \left\{ \exp \left( \frac{1}{2} e_k^T \text{diag}(C_j(\tau - t) V) \right) - 1 \right\} \tilde{N}_j^Q(dt, dV).
\]

81
Here, \( \tilde{N}^Q_j(dt,dV) = N_j(dt,dV) - \exp(tr(V\Theta_j))\nu_{L_j}(dV) dt \) and \( N_j \) is the Poisson random measure of \( \tilde{L}_j \), for \( j = 1, \ldots, n \).

Proof. First, let us notice that by definition, the process \( t \mapsto F_k(t,\tau) \) is a martingale for \( t \leq \tau \). From Prop. 3.4, we have in a compact form

\[
F_k(t,\tau) = \Lambda_k(\tau) \exp\left( e_k^T e^{A(\tau-t)} X(t) + e_k^T e^{B(\tau-t)} Y(t) + \frac{1}{2} \text{diag}(C(\tau-t)Z(t)) \right) G_k(\tau-t),
\]

where we have collected all non-random terms into \( G \), being a vector in \( \mathbb{R}^d \). Since \( F_k(t,\tau) \) depends on \( X(t), Y(t) \) and \( Z(t) \), the dynamics of \( F_k \) will necessarily be expressible in terms of the \( Q \)-Wiener process \( \tilde{W} \) and the compensated Possion random measures of \( \tilde{L}_j \) under \( Q \). Hence, when using Itô’s Formula for jump processes (see e.g. Shiryaev [26]), we only need to focus on terms involving \( d\tilde{W} \) and \( \tilde{N}^Q_j(dt,dV) \). To do this, we note that the dynamics of \( Y(t) \) can be written as

\[
dY(t) = (\mu + BY(t)) dt + \eta d(\text{diag}(\tilde{L}(u))).
\]

Moreover, since \( C(\tau-t) \) and \( \text{diag} \) are linear operators on matrices, we have that \( F_k \) is a function of linear combinations of \( Z_{u,v}(t) \) and \( Y_u(t) \), for \( u,v = 1, \ldots, d \). Hence, the dynamics will consist of linear combinations of the elements of the \( \tilde{L}(t) \)-matrix. Applying Itô’s Formula taking into account all these considerations yields the result. \( \square \)

We see that there is a Samuelson effect in the forward price dynamics. The volatility appearing in the \( d\tilde{W} \)-term of the dynamics takes the form \( e_k^T \exp(A(\tau-t))\Sigma^{1/2}(t) \). The contribution from \( e_k^T \exp(A(\tau-t)) \) is an “exponential scaling” of the stochastic spot volatility \( \Sigma^{1/2}(t) \). Moreover, as time to maturity goes to zero, we obtain a convergence of the forward volatility to the spot volatility,

\[
\lim_{\tau \uparrow t} e_k^T e^{A(\tau-t)}\Sigma^{1/2}(t) = \Sigma^{1/2}(t).
\]

This yields a generalization of the Samuelson effect known in the one-dimensional case to cross-commodity forward prices. We remark that the one-dimensional Samuelson effect gives a forward volatility which is exponential dampening (in ‘time-to-maturity’) of the spot volatility. However, in the multi-dimensional case, the shape of \( e_k^T \exp(A(\tau-t))\Sigma^{1/2} \) will be much richer than simply exponential decay in time to maturity towards spot volatility. In fact, one may get situations where the forward volatility is increasing rather than decreasing with time to maturity. For example, choosing \( A \) to be a matrix of CARMA-type (see Benth et al. [6]), we may get this situation, which is in contrast to the classical Samuelson effect. Observe that also the jump-terms in the dynamics of the forward price contributes with a Samuelson effect, however, this is much more complex to analyse.

In the next Proposition we show that the forward price will behave like the seasonal function.
in the long end of the market. To prove this result, we dispense with the restriction that the forward price is only defined up to maturities $\tilde{T} < \infty$, but do an asymptotic consideration of $F$ only focusing on the expression in Prop. 3.4.

**Proposition 3.6.** Let $F(t, \tau)$ be given as in Prop. 3.4 and suppose $\lim_{t \to \infty} \ln \Psi(t)$ exists. Then,

$$\lim_{\tau \to \infty} (\ln F(t, \tau) - \ln \Lambda(\tau)) = A^{-1} \theta_0 + \sum_{i=1}^{m} B_i^{-1} \mu_i + \lim_{\tau \to \infty} \ln \Psi(\tau).$$

Here we understand the operations of the function $\ln$ coordinate-wise.

**Proof.** This result follows immediately from the assumption that the real parts of the eigenvalues of the matrices $A$, $B_i$ and $C_j$ are all negative, $i = 1, \ldots, m$ and $j = 1, \ldots, n$. $\square$

Note that the condition that $\Psi(\tau)$ has a limit is equivalent to the existence of a stationary dynamics of $\int_{t}^{\tau} C_j(t - s) d\tilde{L}_j(s)$ and $\int_{0}^{t} e^{B_i(t-s)} \eta_i dL_i(u)$ under $Q$. If this is the case, then we can interpret $\lim_{\tau \to \infty} \ln \Psi(\tau)$ as the long-term mean value of the market price of risk.

From Prop. 3.6 above, contracts with maturities in the long end of the market will have forward prices which are basically equal to the seasonality function, adjusted by the stationary mean values 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 seasonal mean adjusted by the market prices of risk.

### 3.3 Shapes of the forward curve

Note that we can view the forward price dynamics as 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

$$\ln \Theta(t, \tau) \triangleq \ln \Psi(\tau - t) + \ln \Lambda(\tau) + A^{-1}(I - e^{A(\tau-t)})\theta_0 + \sum_{i=1}^{m} B_i^{-1}(I - e^{B_i(\tau-t)})\mu_i. \quad (3.14)$$

Then, from Prop. 3.4,

$$\ln F(t, \tau) = \ln \Theta(t, \tau) + e^{A(\tau-t)} \ln S(t) + \sum_{i=1}^{m} (e^{B_i(\tau-t)} - e^{A(\tau-t)}) Y_i(t)$$

$$+ \frac{1}{2} \text{diag}\{\sum_{j=1}^{n} C_j(\tau - t) Z_j(t)\}. \quad (3.15)$$
Chapter 4. Pricing of forwards and options in a multivariate non-Gaussian stochastic volatility model for energy markets

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 $Y_i$ and volatility processes $Z_j$ 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. There is also a direct influence from the volatility processes between the forwards both directly, and indirectly via the stochastic volatility $\Sigma(t)$ in the dynamics of $X$.

As noted in Andresen et al. [1], 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 the eigenvalues of $A$ on the diagonal and zero elsewhere (see e.g. Horn and Johnson [14]). Hence, an entry of the vector $e^{A(\tau-t)} X(t)$ can be represented as

$$
\sum_{i=1}^{d} a_1^i e^{\lambda_i(\tau-t)} X_1(t) + \sum_{i=1}^{d} a_2^i e^{\lambda_i(\tau-t)} X_2(t) + \ldots + \sum_{i=1}^{d} a_d^i e^{\lambda_i(\tau-t)} X_d(t),
$$

for some constants $a_{ij} \in \mathbb{R}$ and eigenvalues $\lambda_i, i, j = 1, \ldots, d$. Consider first the Schwartz model with constant volatility, i.e. no contribution of the processes $Y_i$ and $Z_j$ in the forward price. If $X$ is positive in all its components, $\lambda_i$ is real and $a_{ij} \in \mathbb{R}^+$ for all $i, j = 1, \ldots, d$, 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 [12] 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 $a_{ij}$ 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 which rise and decay at different speeds. Due to this the forward may alternate between backwardation and contango and humps may appear (see figure 3.3 (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 3.3 (a)).

A similar analysis can be made for the spike processes $Y_i$. However, since $Y_i$ 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 $Y_i$ may be averaged out by jumps in the volatility process $Z_i$. The processes $Y_i$ and $Z_i$ are driven
Figure 4.1: Paths of $e^{A(\tau-t)}X$ for (a) complex eigenvalues, (b) real eigenvalues of $A$, moreover $X = (1, 2)^T$ is taken constant.

by related subordinators $L_i$ and $\tilde{L}_i$. 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. This is a result of the inverse leverage effect in the spot model, which has a "double" impact on forward prices.

4 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 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 traded.

In this section we will consider pricing of options on a combination of forwards, with the spread as a special case. The dynamics of the forward is given by our multivariate model, which allows for the application of the Fourier method to pricing.

Consider an option written on a combination of the forwards expressed via the payoff function $p : \mathbb{R}^d \rightarrow \mathbb{R}$. At exercise time $T \leq \tau$, the option pays out $p(F(T, \tau))$, with the forwards maturing at time $\tau \geq T$. Supposing that $p(F(T, \tau))$ is integrable with respect to the pricing measure $Q$ defined in the previous section, we have that the option price at time $t \leq T$ becomes

$$C(t) = e^{-r(T-t)}\mathbb{E}_Q [p(F(T, \tau)) | \mathcal{F}_t] ,$$

where the constant $r > 0$ is the risk-free interest rate. As it turns out, the forward price, or
rather its logarithm, has a semi-analytic cumulant function, which opens for applying the Fourier method to option pricing (see Carr and Madan [8] for a general treatment of Fourier methods in pricing of options). We now discuss this in more detail.

First, define the function
\[ g(x) \triangleq p(e^x), \]
and observe that
\[ g(\ln x) = p(x), \]
where we used pointwise exponentials and logarithms. Suppose \( g \in L^1(\mathbb{R}^d) \), the space of integrable functions on \( \mathbb{R}^d \), and recall the \( d \)-dimensional Fourier transform as
\[ \hat{g}(y) = \int_{\mathbb{R}^d} g(x) e^{-i\langle x, y \rangle} dx. \]
(4.3)

If \( \hat{g} \in L^1(\mathbb{R}^d) \), then the inverse Fourier transform becomes
\[ g(x) = \frac{1}{2\pi} \int_{\mathbb{R}^d} \hat{g}(y) e^{i\langle y, x \rangle} dy. \]
(4.4)

See Folland [11] for these definitions. To price options, let us introduce the conditional cumulant function of the log-forward prices under \( Q \): for \( s \leq t \leq \tau \) and \( x \in \mathbb{R}^d \), define
\[ \tilde{\phi}_{\ln F}^{(s,t,\tau)}(x) \triangleq \ln E_Q [e^{i\langle x, \ln F(t,\tau) \rangle} | \mathcal{F}_t] \]
(4.5)

The following pricing relation holds:

**Proposition 4.1.** Suppose that \( g, \hat{g} \in L^1(\mathbb{R}^d) \), where \( g \) is defined in (4.2). Then
\[ C(t) = e^{-r(T-t)} \frac{1}{2\pi} \int_{\mathbb{R}^d} \hat{g}(y) \exp \left( \tilde{\phi}_{\ln F}^{(t,T,\tau)}(y) \right) dy, \]
where \( \tilde{\phi}_{\ln F}^{(t,T,\tau)}(y), t \leq T \leq \tau \) is the conditional characteristic function of \( \ln F(T,\tau) \) defined in (4.5).

**Proof.** Since \( g \in L^1(\mathbb{R}^d) \), using dominated convergence to commute integration and expectation (see Folland [11]), we conclude
\[ C(t) = e^{-r(T-t)} E_Q [p(F(T, \tau)) | \mathcal{F}_t] \]
\[ = e^{-r(T-t)} E_Q [g(\ln F(T, \tau)) | \mathcal{F}_t] \]
\[ = e^{-r(T-t)} E_Q \left[ \frac{1}{2\pi} \int_{\mathbb{R}^d} \hat{g}(y) e^{i\langle y, \ln F(T,\tau) \rangle} dy | \mathcal{F}_t \right] \]
\[ = e^{-r(T-t)} \frac{1}{2\pi} \int_{\mathbb{R}^d} \hat{g}(y) E_Q [e^{i\langle y, \ln F(T,\tau) \rangle} | \mathcal{F}_t] dy \]
\[
= e^{-r(T-t)} \frac{1}{2\pi} \int_{\mathbb{R}^d} \hat{g}(y) \exp \left( \tilde{\phi}_{\ln F}^{(t,T,\tau)}(y) \right) \, dy.
\]

This proves the result. \qed

The two main ingredients in the pricing using Fourier methods are the transform of the payoff function, \( \hat{g} \) and the cumulant of the forward price under the pricing measure \( Q \). We state a semi-analytical expression for the latter.

**Proposition 4.2.** Assume the conditions of Prop. 3.4 hold. Then the conditional cumulant function of \( \ln F(t,\tau) \) for \( s \leq t \leq \tau \) defined in (4.5) is

\[
\tilde{\phi}_{\ln F}^{(s,t,\tau)}(x) = i x^T H(s, t, \tau) + i x^T e^{A(\tau-s)} X(s) + \sum_{i=1}^m i x^T e^{B_i(\tau-s)} Y_i(s)
\]

\[
+ \frac{1}{2} \sum_{j=1}^n i x^T \text{diag} (D_j(t-s, \tau-t) Z_j(s)) - \frac{1}{2} \sum_{j=1}^n x^T C_j(\tau-s) Z_j(s) x + \Xi(s, t, \tau, x),
\]

for \( x \in \mathbb{R}^d \), where

\[
H(s, t, \tau) = \ln \Lambda(\tau) + \ln \Psi(\tau - t) + A^{-1} \left( I - e^{A(\tau-t)} \right) \theta_0 + \sum_{i=1}^m B_i^{-1} \left( I - e^{B_i(\tau-t)} \right) \mu_i
\]

\[
+ A^{-1} \left( e^{A(\tau-t)} - e^{A(\tau-s)} \right) \theta_0 + \sum_{i=1}^m B_i^{-1} \left( e^{B_i(\tau-t)} - e^{B_i(\tau-s)} \right) \mu_i
\]

and

\[
\Xi(s, t, \tau, x)
\]

\[
= \sum_{j=1}^m \int_0^{t-s} \left\{ \phi_{L_j} \left( \frac{1}{2} i C_j^*(\tau - t + v)(x x^T) + \frac{1}{2} D_j^T(v, \tau - t)(J_d(x)) + J_d(x^T e^{B_j(\tau-t+v)} \eta_j) - i \Theta_j \right) 
\]

\[- \phi_{L_j} (-i \Theta_j) \right\} \, dv
\]

\[
+ \sum_{j=m+1}^n \int_0^{t-s} \left\{ \phi_{L_j} \left( \frac{1}{2} i C_j^*(\tau - t + v)(x x^T) + \frac{1}{2} D_j^T(v, \tau - t)(J_d(x)) - i \Theta_j \right) - \phi_{L_j} (-i \Theta_j) \right\} \, dv.
\]

The family of linear operators \( D_j(u, v), (u, v) \in \mathbb{R}^2 \), are defined as

\[
D_j(u, v) X = C_j(v) e^{C_j^T u} X e^{C_j^T u},
\]

for \( j = 1, \ldots, n \) and a matrix \( X \in M_d(\mathbb{R}) \).
Proof. From Prop. 3.4, it holds

\[
\ln F(t, \tau) = \ln \Theta(t, \tau) + e^{A(\tau)} X(t) + \sum_{i=1}^{m} e^{B_i(\tau-t)} Y_i(t) + \frac{1}{2} \text{diag} \{ \sum_{j=1}^{n} C_j(\tau-t) Z_j(t) \},
\]

where we recall the short-hand notation for \( \Theta(t, \tau) \) defined in (3.14). Now, from the explicit solutions of the factors in (2.7), (2.8) and (2.9), together with the Girsanov change of the Brownian motion \( W \), we find by adaptedness that

\[
\bar{\phi}_{ln \ F}^{(s,t,\tau)}(x) = ix^T H(s,t,\tau) + ix^T e^{A(\tau-s)} X(s) + ix^T \sum_{i=1}^{m} e^{B_i(\tau-s)} Y_i(s) + \frac{1}{2} ix^T \text{diag} \{ \sum_{j=1}^{n} D_j(t-s, \tau-t) Z_j(s) \} + \ln \mathbb{E}_Q \left[ \exp \left( ix^T \int_{s}^{t} e^{A(\tau-u) \Sigma_{1/2}(u)} d\tilde{W}(u) + ix^T \sum_{i=1}^{m} \int_{s}^{t} e^{B_i(\tau-u)} \eta_i dL_i(u) 
\right. 
\left. + \frac{1}{2} ix^T \text{diag} \{ \sum_{j=1}^{n} C_j(\tau-t) \int_{s}^{t} e^{C_j(t-u)} dL_j(u) e^{C^T_j(t-u)} \} \right] | \mathcal{F}_s \right].
\]

Define \( \psi(s, t, \tau) \) as the logarithm of the conditional expectation in the expression above. Letting \( \mathcal{G}_{s,t} \) be the \( \sigma \)-algebra generated by \( \mathcal{F}_s \) and the paths of \( \tilde{L}_j(u), s \leq u \leq t \), we find after using the tower property of the conditional expectation operator and the Gaussianity of Itô integrals of deterministic functions

\[
\psi(s, t, \tau) = \ln \mathbb{E}_Q \left[ \mathbb{E}_Q \left[ \exp \left( ix^T \int_{s}^{t} e^{A(\tau-u) \Sigma_{1/2}(u)} d\tilde{W}(u) \right) | \mathcal{G}_{s,t} \right] \right]
\cdot \exp \left( ix^T \sum_{i=1}^{m} \int_{s}^{t} e^{B_i(\tau-u)} \eta_i dL_i(u) \right)
\cdot \exp \left( \frac{1}{2} ix^T \text{diag} \{ \sum_{j=1}^{n} C_j(\tau-t) \int_{s}^{t} e^{C_j(t-u)} dL_j(u) e^{C^T_j(t-u)} \} \right) | \mathcal{F}_t \right]
\]

\[
= \ln \mathbb{E}_Q \left[ \exp \left( -\frac{1}{2} x^T iz^T \int_{s}^{t} e^{A(\tau-u) \Sigma_{1/2}(u)} e^{A^T(\tau-u)} x \right) \right]
\cdot \exp \left( ix^T \sum_{i=1}^{m} \int_{s}^{t} e^{B_i(\tau-u)} \eta_i dL_i(u) \right)
\cdot \exp \left( \frac{1}{2} ix^T \text{diag} \{ \sum_{j=1}^{n} C_j(\tau-t) \int_{s}^{t} e^{C_j(t-u)} dL_j(u) e^{C^T_j(t-u)} \} \right) | \mathcal{F}_t \right]
\]

88
Inspecting the proof of Lemma 3.3, we find

$$\int_s^t e^{A(\tau-u)\Sigma(u)} e^{A^T(\tau-u)} \, du = \sum_{j=1}^n C_j(\tau-s)Z_j(s) + \int_s^t C_j(\tau-u) \, d\tilde{L}_j(u).$$

By $\mathcal{F}_s$-adaptedness and independent increment property of Lévy processes, it holds

$$\psi(s, t, \tau) = \frac{1}{2} \sum_{j=1}^n x^T C_j(\tau-s)Z_j(s) x + \ln \mathbb{E}_Q \left[ \exp \left( -\frac{1}{2} \sum_{j=1}^n x^T \int_s^t C_j(\tau-u) \, d\tilde{L}_j(u) x \right) \right] + \frac{1}{2} ix^T \text{diag} \left\{ \sum_{j=1}^n C_j(\tau-t) \int_s^t e^{C_j(t-u)} \, d\tilde{L}_j(u) e^{C_j^T(t-u)} \right\} + \sum_{i=1}^m \frac{1}{2} x^T \int_s^t e^{B_i(t-u)} \eta_i \, dL_i(u) \right].$$

We focus next on the last term, the logarithm of the expectation, which we denote by $\tilde{\psi}(s, t, \tau)$. Observe first that

$$C_j(\tau-t) \int_s^t e^{C_j(t-u)} \, d\tilde{L}_j(u) e^{C_j^T(t-u)} = \int_s^t D_j(t-u, \tau-t) \, d\tilde{L}_j(u).$$

But since for a matrix $A \in M_d(\mathbb{R})$, $x^T \text{diag}(A) = tr\{J_d(x)A\}$,

$$\frac{1}{2} x^T \text{diag} \left\{ \sum_{j=1}^n \int_s^t D_j(t-u, \tau-t) \, d\tilde{L}_j(u) \right\} = \sum_{j=1}^n \int_s^t D_j(t-u, \tau-t) \, d\tilde{L}_j(u).$$

Furthermore, it holds that

$$x^T \int_s^t e^{B_i(t-u)} \eta_i \, dL_i(u) = \int_s^t x^T e^{B_i(t-u)} \eta_i \, dL_i(u) = tr\left\{ \int_s^t J_d(x) e^{B_i(t-u)} \, d\tilde{L}_i(u) \right\},$$

and

$$-\frac{1}{2} x^T \int_s^t C_j(\tau-u) \, d\tilde{L}_j(u) x = i tr\left\{ \frac{1}{2} ix^T \int_s^t C_j(\tau-u) \, d\tilde{L}_j(u) \right\}.$$

Hence, collecting terms and using that $\tilde{L}_j$ are independent for $j = 1, \ldots, n$, we find

$$\tilde{\psi}(s, t, \tau) = \sum_{j=1}^m \ln \mathbb{E}_Q \left[ \exp \left( i tr\left\{ \frac{1}{2} ix^T \int_s^t C_j(\tau-u) \, d\tilde{L}_j(u) \right\} \right) \right].$$
In the last equality, we used the same argumentation as is found in the proof of Prop. 3.4. After collecting terms, the proposition is proved.

The fast Fourier transform (FFT) algorithm may be used to compute the option price efficiently, as long as we know the Fourier transform of the payoff function \( g \). Note that implementing the FFT algorithm requires some numerical integration routines to evaluate the characteristic function of \( \ln F \).

We consider the specific case of a call option on the spread between two forwards. The payoff function of such a contract is \( p(x) = \max(x_1 - x_2 - K, 0) \), where \( K \) is the strike price. Without loss of generality, we can suppose that \( K = 1 \). The function \( g \) becomes

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

We observe that this function is not integrable on \( \mathbb{R}^2 \). However, following the idea in Carr and Madan [8], we can damp \( g \) by an exponential function. To this end, define for \( \xi = (\xi_1, -\xi_2) \) with \( \xi_1, \xi_2 > 0 \),

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

(4.6)

We show that this becomes an integrable function under natural conditions on the damping factors \( \xi_1, \xi_2 \).

Lemma 4.3. If \( \xi_1 - \xi_2 > 1 \), then \( g_\xi \in L^1(\mathbb{R}^2) \) where \( g_\xi \) is defined in (4.6).

Proof. Note that the function \( g_\xi \) is non-zero whenever \( x_1 > \ln(e^{x_2} + 1) \). Thus, since \( \xi_1 > 1 \),

\[
\int_{\mathbb{R}^2} g_\xi(x) \, dx = \int_{-\infty}^{\infty} e^{\xi_2 x_2} \int_{\ln(e^{x_2} + 1)}^{\infty} e^{-\xi_1 x_1} (e^{x_1} - (e^{x_2} + 1)) \, dx_1 \, dx_2.
\]
If \( x_2 > 0 \), we find that
\[
e^{\xi_2 x_2} (e^{x_2} + 1)^{-\xi_1 - 1} = e^{\xi_2 x_2} e^{-(\xi_1 - x_2)(1 + e^{-x_2}) - (\xi_1 - 1)} \leq e^{\xi_2 - \xi_1} x_2.
\]

By assumption \( \xi_1 - \xi_2 > 1 \), so \( \xi_2 - \xi_1 + 1 < 0 \). If \( x_2 < 0 \), then
\[
e^{\xi_2 x_2} (e^{x_2} + 1)^{-\xi_1 - 1} \leq e^{\xi_2 x_2}.
\]

The Lemma follows. \( \square \)

In the next Lemma we state the Fourier transform of \( g_\xi \).

**Lemma 4.4.** Suppose \( \xi_2 > 0 \) and \( \xi_1 - \xi_2 > 1 \). Then the Fourier transform of \( g_\xi(x) \) defined in (4.6) is
\[
\hat{g}_\xi(y) = \frac{\Gamma(i(y_1 + y_2) - (1 + \xi_1 + \xi_2))\Gamma(-iy_2 + \xi_2 + 2)}{\Gamma(iy_1 + 1 - \xi_1)} ,
\]
where \( \Gamma \) denotes the gamma function.

**Proof.** For the proof we follow the approach of Hurd and Zhou [15] (Theorem 1). When one takes into account the exponential damping of the pay-off function \( g \) by \( e^{\langle \xi, x \rangle} \) then the above result follows. \( \square \)

We have that
\[
g(x) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \hat{g}_\xi(y) e^{i(y-x)\xi} dy .
\]
Thus, the price of a spread option entails in substituting \( y \) with \( y - i\xi \) in the formula for \( C(t) \) in Prop. 4.1, and use \( \hat{g}_\xi \) instead of \( \hat{g} \). In addition comes an exponential integrability condition on \( \ln F(T, \tau) \) in order to take into account the additional contribution from \( \exp(\langle \xi, x \rangle) \).

An alternative approach to the Fourier method is to apply Monte Carlo simulation of the forward price dynamics \( F(T, \tau) \) in the pricing of options. In practice, this means simulating matrix valued subordinators \( \tilde{L}_j \) and a multi-dimensional Wiener processes \( \tilde{W} \) under \( Q \). The latter can be simulated using classical sampling techniques. Efficient simulation methods for matrix-valued subordinators is in general an open question, however, for a specific class of such processes a method is proposed in Benth and Vos [7].
5 Conclusions

Based on the multivariate spot price model with Barndorff-Nielsen and Shephard stochastic volatility introduced in Benth and Vos [7], we derive the multivariate forward price dynamics. These analytical forward prices are calculated 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 has 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 to price cross-commodity options on forwards. The particular case of spread options were analysed in more detail.
Bibliography


Bibliography


Chapter 5

Futures pricing in electricity markets based on stable CARMA spot models

Fred Espen Benth, Claudia Klüppelberg, Gernot Müller and Linda Vos
Abstract

We present a new model for the electricity spot price dynamics, which is able to capture seasonality, low-frequency dynamics and the extreme spikes in the market. Instead of the usual purely deterministic trend we introduce a non-stationary independent increments process for the low-frequency dynamics, and model the large fluctuations by a non-Gaussian stable CARMA process. The model allows for analytic futures prices, and we apply these to model and estimate the whole market consistently. Besides standard parameter estimation, an estimation procedure is suggested, where we fit the non-stationary trend using futures data with long time until delivery, and a robust $L^1$-filter to find the states of the CARMA process. The procedure also involves the empirical and theoretical risk premiums which – as a by-product – are also estimated. We apply this procedure to data from the German electricity exchange EEX, where we split the empirical analysis into base load and peak load prices. We find an overall negative risk premium for the base load futures contracts, except for contracts close to delivery, where a small positive risk premium is detected. The peak load contracts, on the other hand, show a clear positive risk premium, when they are close to delivery, while the contracts in the longer end also have a negative premium.

1 Introduction

In the last decades the power markets have been liberalized world-wide, and there is a large interest for modelling power spot prices and derivatives. Electricity spot prices are known to be seasonally varying and mean-reverting. Moreover, a distinctive characteristic of spot prices is the large spikes that occur due to sudden imbalances in supply and demand, for example, when a large production utility experiences a black-out or temperatures are suddenly dropping. Typically in these markets, different production technologies have big variations in costs, leading to a very steep supply curve. Another characteristic of electricity is the lack of efficient storage possibilities. Many spot price models have been suggested for electricity, and we refer to Eydeland and Wolyniec [21] and Benth, Šaltytė Benth and Koekebakker [3] for a discussion on various models and other aspects of modelling of energy markets.

In this paper we propose a two-factor arithmetic spot price model with seasonality, which is analytically feasible for pricing electricity forward and futures contracts. The spot price model consists of a continuous-time autoregressive moving average factor driven by a stable Lévy process for modelling the stationary short-term variations, and a non-stationary long-term factor given by a Lévy process. We derive futures prices under a given pricing measure, and propose to fit the spot model by a novel optimization algorithm using spot and futures price data simultaneously. We apply our model and estimation procedure on price data observed at the German electricity exchange EEX.
In a seminal paper by Schwartz and Smith [35] a two-factor model for commodity spot prices is proposed. Their idea is to model the short-term logarithmic spot price variations as a mean-reverting Ornstein-Uhlenbeck process driven by a Brownian motion, reflecting the drive in prices towards its equilibrium level due to changes in supply and demand. But, as argued by Schwartz and Smith [35], there may be significant uncertainty in the equilibrium level caused by inflation, technological innovations, scarceness of fuel resources like gas and coal etc.. To account for such long-term randomness in prices, Schwartz and Smith [35] include a second non-stationary factor being a drifted Brownian motion, possibly correlated with the short-term variations. They apply their model to crude oil futures traded at NYMEX, where the non-stationary part is estimated from futures prices, which are far from delivery. Mean-reversion will kill off the short-term effects from the spot on such futures, and they can thus be applied to filter out the non-stationary factor of the spot prices.

This two-factor model is applied to electricity prices by Lucia and Schwartz [28]. Among other models, they fit an arithmetic two-factor model with deterministic seasonality to electricity spot prices collected from the Nordic electricity exchange NordPool. Using forward and futures prices they fit the model, where the distance between theoretical and observed prices are minimized in a least squares sense.

A major criticism against the two-factor model considered in Lucia and Schwartz [28] is the failure to capture spikes in the power spot dynamics. By using Brownian motion driven factors, one cannot explain the sudden large price spikes frequently observed in spot data. Multi-factor models, where one or more factors are driven by jump processes, may mend this. For example, Benth, Kallsen and Meyer-Brandis [2] suggest an arithmetic spot model where normal and spike variations in the prices are separated into different factors driven by pure-jump processes. In this way one may model large price increases followed by fast speed of mean-reversion together with a “base component”-behaviour, where price fluctuations are more slowly varying around a mean level. Such multi-factor models allow for analytic pricing of forward and futures contracts.

A very attractive alternative to these multi-factor models are given by the class of continuous-time autoregressive moving-average processes, also called CARMA processes. These processes incorporate in an efficient way memory effects and mean-reversion, and generalize Ornstein-Uhlenbeck processes in a natural way (see Brockwell [11]). As it turns out, (C)ARMA processes fit power spot prices extremely well, as demonstrated by Bernhardt, Klüppelberg and Meyer-Brandis [7] and Garcia, Klüppelberg and Müller [23]. In [7] an ARMA process with stable innovations, and in [23] a CARMA(2,1)-model driven by a stable Lévy process are suggested and empirically studied on power spot price data collected from the Singapore and German EEX markets, respectively. A CARMA(2,1) process may be viewed on a discrete-time scale as an autoregressive process of order 2, with a moving average order 1. By invoking a stable Lévy process to drive the CARMA model, one is blending spikes and small innovations in prices into one process. We remark in passing that a CARMA dynamics has been applied to model crude oil prices at NYMEX by Paschke and Prokopczuk [30] and interest-rates by Zakamouline, Benth
Fred Espen Benth, Claudia Klüppelberg, Gernot Müller and Linda Vos

and Koekebakker [39].

We propose a generalization of the stable CARMA model of Garcia et al. [23] by including a long-term non-stationary factor being a general Lévy process. The model allows for analytical pricing of electricity futures, based on pricing measures, which preserve the Lévy property of the driving processes. More precisely, we apply an Esscher measure transform to the non-stationary part, and a transform which maps the stationary stable process into a tempered stable. Due to the semi-affine structure of the model, the futures and forward prices becomes implicitly dependent on the states of the CARMA model and the non-stationary factor.

The CARMA-based factor in the spot model accounts for the short-term variations in prices and will be chosen stationary. By a CARMA model with a higher order autoregressive part we may include different mean-reversion speeds, such that we can mimic the behaviour of a multi-factor model accounting for spikes and base variations separately. The moving average part is necessary to model the observed dependence structure. The stable Lévy process may have very big jumps, which then can explain spike behaviour in the prices. The smaller variations of the stable Lévy process model the base signal in power prices. As it turns out from our empirical investigations using market price data from the EEX, the non-stationary long-term behaviour may accurately be modelled using a normal inverse Gaussian (NIG) Lévy process. We filter out the non-stationary part from observing futures prices, which are far from delivery. The influence of the stationary CARMA factor is then not present, and the data shows a significant non-normal behaviour. This is in contrast to the choice suggested by Schwartz and Smith [35] and applied by Lucia and Schwartz [28]. Moreover, we find that a CARMA(2,1) model is accurately explaining the mean-reversion and memory effects in the spot data.

A novelty of our paper apart from the generalizing existing one and two factor models, is our estimation procedure. Lucia and Schwartz [28] propose an iterative algorithm for estimating their two-factor model to NordPool electricity data, where they minimize the least-squares distance between the theoretical and observed forward and futures prices to find the risk-neutral parameters. In order to find the theoretical prices, they must have the states of the two factors in the spot model accessible. Since these are not directly observable, they choose an iterative scheme, where they start with a guess on the parameter values, find the states minimizing distance, update parameters by estimation, find the states minimizing the distance etc. until convergence is reached. We propose a different approach, utilizing the idea in Schwartz and Smith [35] that the non-stationary factor is directly observable, at least approximately, from forward prices, which are far from delivery. We apply this to filter out the non-stationary factor. The CARMA-part is then observable from the spot prices, where seasonality and the non-stationary term is subtracted. Since we work with stable processes, which do not have finite second moments, \( L^2 \)-filters can not be used to find the states of the CARMA-process. We propose a simple \( L^1 \)-filter being more robust with respect to spikes in spot data to do this. The problem we are facing is to determine, what contracts to use for filtering out the non-stationary part. To find an optimal “time-to-maturity” which is sufficiently far from delivery, so that the futures prices behave as the non-stationary fac-
We apply our model and estimation scheme to data from the German EEX (European Energy Exchange) market where we use spot prices as well as futures prices of contracts with a delivery period of one month. Our empirical studies cover both base load and peak load contracts, where base load contracts are settled against the average of all hourly spot prices in the delivery period. Peak load futures contracts are settled against the average of hourly spot price in peak periods of the delivery period. The peak load period is the period between 8 a.m. and 8 p.m., during every working day. As a first summary, we can say that the results for both base and peak load data are in general rather similar. However, the peak load data show a more extreme behaviour. The average risk premium decays when time to maturity increases, and is negative for contracts in the longer end of the futures curve. This points towards a futures market, where producers use the contracts for hedging and in return accept to pay a premium to insure their production, in accordance with the theory of normal backwardation. The risk premium is completely determined by the effect of the long-term factor, which induces a close to linear decay as a function of “time-to-maturity”. We see that for the base load contracts the risk premium in the short end of the curve is only slightly positive. The risk premium is negative for contracts starting to deliver in about two months or later. On the other hand, the peak load contracts have a clear positive risk premium, which turns to a negative one for contracts with delivery in about four months or later. The positive risk premium for contracts close to delivery tells us that the demand side (retailers and consumers) of the market is willing to pay a premium for locking in electricity prices as a hedge against spike risk (see Geman and Vasicek [22]).

Our results are presented as follows. In Section 2 we present the two-factor spot model, and we compute analytical futures prices along with a discussion of pricing measures in Section 3. Section 4 explains in detail the estimation steps and the procedure applied to fit the model to data. The results of this estimation procedure applied to EEX data is presented and discussed in Section 5. We conclude in Section 6.

Throughout we use the following notation. For a matrix $D$ we denote by $D^*$ its transposed, and $I$ is the identity matrix. For $p \in \mathbb{N}$ we denote by $e_p$ the $p$-th unit vector. The matrix exponential $e^{At}$ is defined by its Taylor expansion $e^{At} = I + \sum_{n=1}^{\infty} \frac{(At)^n}{n!}$ with identity matrix $I$. We also denote by $\log^+ x = \max(\log x, 0)$ for $x \in \mathbb{R}$. 
2 The spot price dynamics

In most electricity markets, like the EEX, hourly spot prices for the delivery of 1 MW of electricity are quoted. As is usual in the literature on electricity spot price modeling, one assumes a continuous-time model and estimates it on the discretely observed daily average spot prices. We refer, for instance, to Lucia and Schwartz [28] and Benth, Šaltytė Benth and Koekbakker [3] for more details.

We generalize the $\alpha$-stable (C)ARMA model of Bernhardt, Klüppelberg and Meyer-Brandis [7] and Garcia, Klüppelberg and Müller [23] by adding a non-stationary stochastic component in the trend of the spot dynamics. By modeling the trend as a combination of a stochastic process and a deterministic seasonality function rather than only a deterministic seasonality function, which seems common in most models, we are able to describe the low frequent variations of the spot dynamics quite precisely. As it turns out, this trend will explain a significant part of the futures price variations and lead to an accurate estimation of the risk premium in the EEX market.

A two-factor spot price model for commodities, including a mean-reverting short-time dynamics and a non-stationary long-term variations component was first suggested by Schwartz and Smith [35], and later applied to electricity markets by Lucia and Schwartz [28]. Their models were based on Brownian motion driven stochastic processes, more precisely, the sum of an Ornstein-Uhlenbeck (OU) process with a drifted Brownian motion. We significantly extend this model to include jump processes and higher-order memory effects in the dynamics.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ be a complete filtered probability space satisfying the usual conditions of completeness and right continuity. We assume the spot price dynamics

\[ S(t) = \Lambda(t) + Z(t) + Y(t), \quad t \geq 0, \tag{2.1} \]

where $\Lambda$ is a deterministic trend/seasonality function and $Z$ is a Lévy process with zero mean. The process $Z$ models the low-frequency non-stationary dynamics of the spot, and can together with $\Lambda$ be interpreted as the long-term factor for the spot price evolution. The process $Y$ accounts for the stationary short-term variations. We will assume that $Y$ and $Z$ are independent processes. We follow Garcia et al. [23] and Bernhardt et al. [7] and suppose that $Y$ is a stationary CARMA-process driven by an $\alpha$-stable Lévy process.

2.1 The stable CARMA-process

We introduce stationary CARMA$(p, q)$-Lévy processes (see Brockwell [11]) and discuss its relevant properties.

**Definition 2.1** (CARMA$(p, q)$-Lévy process).

A CARMA$(p, q)$-Lévy process $\{Y(t)\}_{t \geq 0}$ (with $0 \leq q < p$) driven by a Lévy-process $L$ is defined
as the solution of the state space equations

\[ Y(t) = \mathbf{b}^* \mathbf{X}(t) \]

\[ d\mathbf{X}(t) = A\mathbf{X}(t)dt + e_p dL(t), \]

with

\[
\mathbf{b} = \begin{pmatrix}
    b_0 \\
    b_1 \\
    \vdots \\
    b_{p-2} \\
    b_{p-1}
\end{pmatrix}, \quad e_p = \begin{pmatrix}
    0 \\
    0 \\
    \vdots \\
    0 \\
    1
\end{pmatrix}, \quad A = \begin{pmatrix}
    0 & 1 & 0 & \cdots & 0 \\
    0 & 0 & 1 & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & 0 & \cdots & 1 \\
    -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1
\end{pmatrix}.
\]

where \( a_1, \ldots, a_p, b_0, \ldots, b_{p-1} \) are possibly complex-valued coefficients such that \( b_q = 1 \) and \( b_j = 0 \) for \( q < j \) \( \leq p \). For \( p = 1 \) the matrix \( A \) is to be understood as \( A = -a_1 \).

The driving process \( L \) of \( Y \) will be a non-Gaussian \( \alpha \)-stable Lévy process \( \{L(t)\}_{t \geq 0} \) with characteristic function given by

\[ \ln \mathbb{E} e^{izL(t)} = t\phi_L(z) \] for \( z \in \mathbb{R} \), where,

\[
\phi_L(z) = \begin{cases}
-\gamma |z|^\alpha (1 - i\beta (\text{sign } z) \tan \left( \frac{\pi \alpha}{2} \right)) + i\mu z & \text{for } \alpha \neq 1, \\
-\gamma |z|(1 + i\beta \frac{\pi}{2} (\text{sign } z) \log |z|) + i\mu z & \text{for } \alpha = 1.
\end{cases}
\] (2.4)

The sign function is defined by \( \text{sign } z = -1 \) for \( z < 0 \), \( \text{sign } z = 1 \) for \( z > 0 \) and \( \text{sign } 0 = 0 \), respectively. Further, \( \alpha \in (0, 2) \) is the shape parameter, \( \gamma > 0 \) the scale, \( \beta \in [-1, 1] \) the skewness, and \( \mu \) the location parameter. If \( \gamma = 1 \) and \( \mu = 0 \), then \( L \) is called standardized.

The parameter \( \alpha \in (0, 2) \) determines the tail of the distribution function of \( L(t) \) for all \( t \geq 0 \). Moreover, only moments strictly less than \( \alpha \) are finite, so that no second moment exists. This implies also that the autocorrelation function does not exist. For further properties on stable processes and Lévy processes, we refer to the monographs by Samorodnitsky and Taqqu [33] and Sato [34].

The solution of the SDE (2.3) is a \( p \)-dimensional Ornstein-Uhlenbeck (OU) process given by

\[ \mathbf{X}(t) = e^{At-s} \mathbf{X}(s) + \int_s^t e^{A(t-u)} e_p dL(u), \quad 0 \leq s < t, \] (2.5)

where the stable integral is defined as in Ch. 3 of Samorodnitsky and Taqqu [33]. From (2.2) we find that \( Y \) is given by

\[ Y(t) = \mathbf{b}^* e^{At-s} \mathbf{X}(s) + \int_s^t \mathbf{b}^* e^{A(t-u)} e_p dL(u), \quad 0 \leq s < t. \] (2.6)
Equations (2.2) and (2.3) constitute the state-space representation of the formal $p$-th order SDE

$$a(D)Y(t) = b(D)DL(t), \quad t \geq 0,$$

(2.7)

where $D$ denotes differentiation with respect to $t$, and

$$a(z) := z^p + a_1z^{p-1} + \cdots + a_p$$

(2.8)

$$b(z) := b_0 + b_1z + \cdots + b_qz^q$$

(2.9)

are the characteristic polynomials. Equation (2.7) is a natural continuous-time analogue of the linear difference equations, which define an ARMA process (cf. Brockwell and Davis [13]).

Throughout we assume that $Y$ and $X$ are stationary in the sense that all finite dimensional distributions are shift-invariant. Based on Proposition 2.2 of Garcia et al. [23] (which summarizes results by Brockwell and Lindner [15]) we make the following assumptions to ensure this:

**Assumptions 2.2. Stationarity of CARMA-process.**

(i) The polynomials $a(\cdot)$ and $b(\cdot)$ defined in (2.8) and (2.9), resp., have no common zeros.

(ii) $E[\log^+ |L(1)|] < \infty$.

(iii) All eigenvalues of $A$ are distinct and have strictly negative real parts.

Assumption (ii) and (iii) imply that $X$ is a causal $p$-dimensional OU process, hence also $Y$ is causal.

**Remark 2.3.** Our model is a significant generalization of the two-factor dynamics by Schwartz and Smith [35] and Lucia and Schwartz [28]. Among various models, Lucia and Schwartz [28] suggested a two factor dynamics of the spot price evolution based on a short term Gaussian OU process and a long-term drifted Brownian motion. In our framework a Gaussian OU process would correspond to a Gaussian CARMA(1,0) process. It is clear that such a Gaussian model cannot capture the large fluctuations in the spot price, like for example spikes, and jump processes seem to be the natural extension. Based on the studies of Bernhardt et al. [7] and Garcia et al. [23], $\alpha$-stable processes are particularly suitable for the short-term dynamics in the spot price evolution. Furthermore, empirical analysis of electricity spot price data from Singapore and Germany in [7] and [23] show strong statistical evidence for full CARMA processes to capture the dependency structure of the data. As for the long-term baseline trend, we shall see in Section 4 that a normal inverse Gaussian Lévy process is preferable to a Gaussian process in a data study from the German electricity exchange EEX.

## 2.2 Dimensionality of CARMA-processes

A more standard model in electricity is to describe the spot by a sum of several OU processes, where some summands describe the spike behavior and others the baseline dynamics (see for
Chapter 5. Futures pricing in electricity markets based on stable CARMA spot models

example the model by Benth et al. [2]). A CARMA process is in a sense comparable to such models, as we now discuss.

By Assumption 2.2(iii) \( A \) has full rank, i.e. it is diagonalizable with \( e^{At} = U e^{Dt} U^{-1} \). Here, \( D \) is a diagonal matrix with the eigenvalues \( \lambda_1, \ldots, \lambda_p \) of \( A \) on the diagonal and \( U \) is the full rank matrix having the eigenvectors of \( A \) as columns. Since all eigenvalues have negative real parts, all components of \( e^{At} \) are mean reverting. Each component of the vector \( e^{A(t-s)} \mathbf{X}(s) \) from (2.6) will therefore mean revert at its own speed, where the speed of mean reversion is a linear combination of the diagonal elements of \( e^{Dt} \).

As we shall see in a simulation example of a CARMA(2,1)-process in Section 4.5, it captures the situation, where a first component has a slower rate of mean-reversion than the second (see Fig. 5.5). This is similar to a two-factor spot model, where the base and spike components of the spot price evolution are separated into two OU processes with different speeds of mean reversion. The advantage of working with a stable CARMA process, as we propose, is that it is possible to capture the distribution of the small and large jumps in one distribution. Since extreme spikes are rather infrequently observed, it is difficult to calibrate the spike component in a conventional two-factor model; this has been observed in Klüppelberg et al. [25]. With our CARMA-model, we avoid the difficult question of spike identification and filtering.

3 The futures price dynamics

In commodity markets, futures contracts are commonly traded on exchanges, including electricity, gas, oil, and coal. In this section we derive the futures price dynamics based on the \( \alpha \)-stable CARMA spot model (2.1). Appealing to general arbitrage theory (see e.g. Duffie [20], Ch.1), we define the futures price \( f(t, \tau) \) at time \( t \) for a contract maturing at time \( \tau \) by

\[
f(t, \tau) = \mathbb{E}_Q [S(\tau) \mid \mathcal{F}_t], \quad 0 \leq t \leq \tau < \infty,
\]

where \( Q \) is a risk neutral probability measure. This definition is valid as long as \( S(\tau) \in L^1(Q) \). In the electricity market, the spot cannot be traded, and every \( Q \sim P \) will be a risk neutral probability (see Benth et al. [3]). For example, \( Q = P \) is a valid choice of a pricing measure. In that case, the condition \( S(\tau) \in L^1(P) \) is equivalent to a tail parameter \( \alpha \) of the stable process \( L \) being strictly larger than one, and a process \( Z \) with finite expectation. In real markets one expects a risk premium and hence it is natural to use a pricing measure \( Q \neq P \). We will discuss possible choices of risk neutral probability measures \( Q \) in Section 3.1.

Based on our spot price model, we find the following explicit dynamics of the futures price for a given class of risk neutral probability measures:

**Theorem 3.1.** Let \( S \) be the spot dynamics as in (2.1), and suppose that \( Q \sim P \) is such that \( L \) and \( Z \) are Lévy processes under \( Q \). Moreover, assume that the processes \( Z \) and \( L \) have finite
which implies

\[ E \left[ A(t) \right] = b^* e^{A(t)} X(t) + (\tau - t) E[Z(1)] + b^* A^{-1} \left( I - e^{A(t)} \right) e_p E[L(1)]. \]

Proof. Using (3.1), \( f(t, \tau) = E[Z(\tau) | F_t] = E[Z(t) + E[Z(\tau) - Z(t) | F_t]] = Z(t) + (\tau - t) E[Z(1)]. \) Since \( Z \) is a Lévy process under \( Q \), we find

\[ E[Z(\tau) | F_t] = Z(t) + E[Z(\tau) - Z(t) | F_t] = Z(t) + (\tau - t) E[Z(1)]. \]

Now denote by \( M(u) = L(u) - E[Z(1)]u \) for \( t \leq u \leq \tau \), which has zero mean. Then, by partial integration,

\[ E_Q \left[ \int_t^\tau b^* e^{A(t-u)} e_p dM(u) \right] = E_Q \left[ b^* e_p M(\tau) - b^* e^{A(t)} e_p M(t) - \int_t^\tau b^* A e^{A(t-u)} e_p E[Z(1)] du = 0, \right. \]

which implies \( E_Q \left[ \int_t^\tau b^* e^{A(t-u)} e_p dL(u) \right] = E_Q[Z(1)] \int_t^\tau b^* e^{A(t-u)} e_p du. \)

Hence, the CARMA part of the spot dynamics converts to

\[ E[Z(\tau) | F_t] = E_Q \left[ b^* e^{A(t)} X(t) + \int_t^\tau b^* e^{A(t-u)} e_p dL(u) | F_t \right] = b^* e^{A(t)} X(t) + \int_t^\tau b^* e^{A(t-u)} e_p du E[Z(1)] \]

Combining the terms yields the result.

In electricity markets the futures contracts deliver the underlying commodity over a period rather than at a fixed maturity time \( \tau \). For instance, in the German electricity market contracts for delivery over a month, a quarter or a year, are traded. These futures are sometimes referred to as swaps, since during the delivery period a fixed (futures) price of energy is swapped against a floating (uncertain) spot price. The futures price is quoted as the price of 1 MWh of power and, therefore, it is settled against the average spot price over the delivery period. Hence, the futures price \( F(t, T_1, T_2) \) at time \( 0 \leq t \leq T_1 < T_2 \) for a contract with delivery period \( [T_1, T_2] \) is defined as

\[ F(t, T_1, T_2) = E_Q \left[ \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} S(\tau) d\tau \right] F_t, \quad (3.2) \]
where we have assumed that settlement of the contract takes place at the end of the delivery period, \( T_2 \).

Using Theorem 3.1 we derive by straightforward integration the swap price dynamics \( F(t, T_1, T_2) \) from (3.2).

**Corollary 3.2.** Suppose all assumptions of Theorem 3.1 are satisfied. Then,

\[
F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \Lambda(\tau) d\tau + Z(t) + \frac{b^* A^{-1}}{T_2 - T_1} (e^{AT_2} - e^{AT_1}) e^{-\alpha T} X(t) + \Gamma_Q(t, T_1, T_2)
\]

where

\[
\Gamma_Q(t, T_1, T_2) = \left( \frac{1}{2} (T_2 + T_1) - t \right) \mathbb{E}_Q[Z(1)] - \frac{b^* A^{-2}}{T_2 - T_1} (e^{AT_2} - e^{AT_1}) e^{-\alpha T} e_p \mathbb{E}_Q[L(1)]
\]

\[+ b^* A^{-1} e_p \mathbb{E}_Q[L(1)]. \]

**Proof.** By the Fubini Theorem, we find

\[
F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(t, \tau) d\tau.
\]

Applying Theorem 3.1 and integrating yield the desired result. \( \square \)

The risk premium is defined as the difference between the futures price and the predicted spot, that is, in terms of electricity futures contracts,

\[
R_{pr}(t, T_1, T_2) = F(t, T_1, T_2) - \mathbb{E} \left[ \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} S(\tau) d\tau \right] | \mathcal{F}_t]. \tag{3.3}
\]

From Cor. 3.2 we find that the theoretical risk premium for a given pricing measure \( Q \) is

\[
R(t, T_1, T_2) = \Gamma_Q(t, T_1, T_2) - \Gamma_P(t, T_1, T_2)
\]

\[= \left( \frac{1}{2} (T_2 + T_1) - t \right) \mathbb{E}_Q[Z(1)] - \frac{b^* A^{-2}}{T_2 - T_1} (e^{AT_2} - e^{AT_1}) e^{-\alpha T} e_p \left( \mathbb{E}_Q[L(1)] - \mathbb{E}[L(1)] \right)
\]

\[+ b^* A^{-1} e_p \left( \mathbb{E}_Q[L(1)] - \mathbb{E}[L(1)] \right). \tag{3.4}
\]

Here, we used the assumption that \( Z \) has zero mean under \( P \). The first term gives a trend in "time to maturity" implied by the non-stationarity part \( Z \) in the spot price dynamics. "Time to maturity" is here interpreted as the time left to the middle of the delivery period. The two last terms are risk premia contributions from the CARMA short-term spot dynamics. They involve an explicit dependence on the speeds of mean-reversion of the autoregressive parts and the memory...
in the moving-average part. We will apply the risk premium in the empirical analysis of spot and futures data from the EEX.

3.1 Equivalent measure transforms for Lévy and $\alpha$-stable processes

In this subsection we discuss a class of pricing measures that will be used for the specification of the futures price dynamics.

We require from the pricing measure that $Z$ and $L$ preserve their Lévy property and the independence. For this purpose, we consider probability measures $Q = Q_L \times Q_Z$, where $Q_L$ and $Q_Z$ are measure changes for $L$ and $Z$, respectively (leaving the other process unchanged). Provided $Z$ has exponential moments, a standard choice of measure change is given by the Escher transform (see Benth et al. [3], Section 4.1.1). Note that $L$, the $\alpha$-stable process in the CARMA-dynamics, does not have exponential moments.

We define the density process of the Radon-Nikodym derivative of $Q_Z$ as

$$\frac{dQ_Z}{dP} \bigg|_{F_t} = \exp(\theta_Z Z(t) - \phi_Z(\theta_Z)t), \quad t \geq 0,$$

for a constant $\theta_Z \in \mathbb{R}$ and $\phi_Z$ being the log-moment generating function of $Z(1)$ (sometimes called the cumulant function of $Z$). In order to make this density process well-defined, exponential integrability of the process $Z$ up to the order of $\theta_Z$ must be assumed. Under this change of measure, the Lévy measure of $Z$ will be exponentially tilted by $\theta_Z$, that is, if we denote the Lévy measure of $Z$ (under $P$) by $\nu(dx)$, then its Lévy measure under $Q_Z$ becomes $\nu_{Q_Z}(dx) = \exp(\theta_Z)\nu(dx)$ (see Benth et al. [3], Section. 4.1.1-4.1.2 for details).

To choose a risk neutral measure $Q_L$ is a more delicate task. We know from Sato [34], Theorems 33.1 and 33.2, that equivalent measures $Q$ exists for stable processes, however, it seems difficult to construct one which preserves the stable property. As an alternative, one may introduce the class of tempered stable processes (see e.g. Cont and Tankov [17], Chapter 9), and apply standard Esscher transformation on these.

A tempered stable process is a pure jump Lévy process, where the stable-like behavior is preserved for the small jumps. However, the tails are tempered and, therefore, extreme spikes are less likely to be modeled with the tempered stable process. The Lévy measure is given by

$$\nu_{TS}(dx) = \frac{c_+ e^{\theta_L x}}{x^{1+\alpha}} 1_{(0,\infty)}(x) \, dx + \frac{c_- e^{\theta_L |x|}}{|x|^{1+\alpha}} 1_{(-\infty,0)}(x) \, dx.$$

Here, $\theta_L \leq 0$ and $c_-, c_+ \in \mathbb{R}_+$. A consequence of the tempering is that certain exponential moments exist. Tempering of a stable distribution results in a tempered stable distribution, and is analogous to taking an Esscher transform of the stable process using a negative parameter $\theta_L$ on the positive jumps, and a positive parameter $-\theta_L$ on the negative jumps.
In particular, define \( q : \mathbb{R} \rightarrow \mathbb{R} \) as \( q(x) := e^{\theta_L x} 1_{(0,\infty)}(x) + e^{\theta_L |x|} 1_{(-\infty,0)}(x) \) for some constant \( \theta_L < 0 \). Suppose the stable distribution \( L \) has (under \( P \)) the characteristic triplet \((\gamma_L, 0, \nu_L)\), where
\[
\nu_L(dx) = \frac{c_+}{x^{1+\alpha}} 1_{(0,\infty)}(x) \, dx + \frac{c_-}{|x|^{1+\alpha}} 1_{(-\infty,0)}(x) \, dx
\]
is the Lévy measure of our stable process \( L \). The parameters \( c_+, c_- \) can be matched to the parameters in (2.4), using Example 2.3.3 of Samorodnitsky and Taqqu [33]. Then the tempered stable measure \( Q_L \), with characteristic triplet \((\gamma_{TS}, 0, \nu_{TS})\) is equivalent to the physical probability measure \( P \) (see Cont and Tankov [17], Proposition 9.8), with drift parameter \( \gamma_{TS} \) is given by
\[
\gamma_{TS} = \begin{cases} 
\gamma_L & 0 < \alpha < 1 \\
\gamma_L + \int_{(|x|<1)} x(q(x) - 1) \nu_L(dx) & 1 < \alpha < 2
\end{cases}
\]
and the Lévy measure \( \nu_{TS} \) is given by \( \nu_{TS}(dx) = q(x) \nu_L(dx) \). The special case of a Cauchy distribution (\( \alpha = 1 \)) is left out since one is not able to define the Lévy-Kintchine formula using a truncation of the small jumps and the large jumps given by \( 1_{|x|>1} \). For our applications it is of particular value to know the expectations of \( L(1) \) under \( P \) and \( Q_L \).

**Lemma 3.3.** Let \( L \) be an \( \alpha \)-stable Lévy process under \( P \) with \( \alpha \in (1,2) \). Find \( Q_L \) by stable tempering for \( \theta_L < 0 \) as in (3.6). Then the difference in mean of \( L(1) \) under \( Q_L \) and \( P \) is given by
\[
\mathbb{E}_{Q_L}[L(1)] - \mathbb{E}[L(1)] = \Gamma(1-\alpha)(-\theta_L)^{\alpha-1} (c_+ - c_-),
\]
where \( \Gamma \) is the gamma function.

**Proof.** Using (3.7) and the Lévy-Khintchine formula (e.g. Cont and Tankov [17], Prop. 3.13) for \( 1 < \alpha < 2 \) we obtain
\[
\mathbb{E}[L(1)] - \mathbb{E}_{Q_L}[L(1)] = \gamma_L - \gamma_{TS} + \int_{(|x|>1)} x(\nu_L - \nu_{TS})(dx)
\]
\[
= \gamma_L - \gamma_{TS} + \int_{(|x|>1)} x(1 - q(x)) \nu_L(dx)
\]
\[
= c_- \int_{-\infty}^{0} \frac{1 - e^{\theta_L |x|}}{x^{1+\alpha}} \, dx + c_+ \int_{0}^{\infty} \frac{1 - e^{\theta_L x}}{x^{1+\alpha}} \, dx
\]
\[
= -\Gamma(1-\alpha)(-\theta_L)^{\alpha-1} (c_+ - c_-),
\]
where we have used partial integration on the two integrals and l’Hospital’s rule to obtain the last identity. This proves the result. \( \square \)
Remark 3.4. By altering $\theta_L$, one can match any relevant mean change in the risk premium $E_Q[L(1)] - E[L(1)]$, as long as this can be obtained by a negative choice of $\theta_L$. This turns out to be appropriate for our applications.

4 Fitting the model to German electricity data

Our data are daily spot and futures prices from July 1, 2002 to June 30, 2006 (available from http://eex.com). We fitted our model both to base load and peak load data, respectively. The futures contracts considered in this analysis are the Phelix-Base-Month-Futures and the Phelix-Peak-Month-Futures. **Base load futures contracts** are settled against the average of all hourly spot prices in the delivery period. **Peak load futures contracts**, on the other hand, are settled against the average of the hourly spot prices in peak periods of the delivery period. The peak period is counted as the hours between 8 a.m. and 8 p.m. every working day during the delivery period. The time series of daily spot prices used for our combined statistical analysis is taken to match the futures contracts: for the base load contracts we use the full time series consisting of daily observations including weekends (i.e., we have 7 observations per week), while in the case of peak load contracts the weekends are excluded (i.e., we have 5 observations per week).

Figures 5.1 and 5.2 show the spot and futures prices for both base load and peak load. From these plots we can see similar patterns of the base and peak load data, however, peak load data are more extreme. Note that all plots cover the same time period; however, for the base spot data we have 1461 observations, whereas for the peak spot data we have only 1045 observations in the same period, due to the missing weekends.

The estimation procedure for our model consists of several steps, which are explained in the following.

4.1 Seasonality function $\Lambda$

The estimation of the deterministic trend component $\Lambda$ is a delicate question. A mis-specification of the trend has a significant effect on the subsequent analysis, in particular, on the risk premium. Motivated by the seasonality functions used in Bernhardt et al. [7] and Garcia et al. [23], we take the seasonality function of the peak load contracts as a combination of a linear trend and some periodic function

$$
\Lambda_p(t) = c_1 + c_2 t + c_3 \cos \left( \frac{2\pi t}{261} \right) + c_4 \sin \left( \frac{2\pi t}{261} \right).
$$

(4.1)

Note that we choose a slightly simpler seasonality function than Bernhardt et al. [7] and Garcia et al. [23], only taking the mean level, a linear trend and a yearly periodicity (modeling the weather difference between summer and winter) into account. Weekly periodicity in peak load
contracts is not that pronounced, since weekends are not considered in peak load data. Since no
new trading information is entering during the weekend (trading takes place during weekdays),
we will adjust the periodicity to 261 and consider the peak load contracts as a continuous process
on all non-weekend days.

In the base load prices a clear weekly seasonality is visible. Weekend prices are in general
lower than during the rest of the week, and over the week one observes the pattern that Monday
and Friday have prices lower than in the middle of the week. Therefore we include a weekly
term in the base load seasonality function:

$$\Lambda_b(t) = c_1 + c_2 t + c_3 \cos\left(\frac{2\pi t}{365}\right) + c_4 \sin\left(\frac{2\pi t}{365}\right) + c_5 \cos\left(\frac{2\pi t}{7}\right) + c_6 \sin\left(\frac{2\pi t}{7}\right)$$

Since time $t$ is running through the weekends, a yearly periodicity of 365 is chosen.

In the following, we will analyse both data sets, base load data and peak load data (spot and
futures, respectively). For simplicity we will suppress the indices $p$ for peak load and $b$ for base
load.

This seasonality functions can be estimated using a robust least-squares estimate on the data;
see Table 5.1 for the resulting estimates. The first plot of Figure 5.3 shows the 1461 observations of the base data set together with the estimated trend and seasonality curve, the second plot zooms into this plot and shows these curves for the first 200 observations. From this second plot we can clearly see that the daily seasonality is captured by $\Lambda$ quite accurately. For the peak data over the same period (consisting of 1045 observations) we get similar plots (hence, we omit them here), except of the fact that $\Lambda$ does not show a weekly seasonality in this case. The overall growth rate was very small during the data period, justifying a linear term as a first order approximation. Note that we have introduced the non-stationary stochastic process $Z$ to absorb all stochastic small term effects in the seasonality. This term will play a prominent role for the

<table>
<thead>
<tr>
<th></th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
<th>$c_5$</th>
<th>$c_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>19.4859</td>
<td>0.0217</td>
<td>-2.8588</td>
<td>0.6386</td>
<td>-6.7867</td>
<td>2.8051</td>
</tr>
<tr>
<td>Peak</td>
<td>30.7642</td>
<td>0.0349</td>
<td>-2.5748</td>
<td>1.5762</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1: Estimated parameters of the seasonality function $\Lambda(\cdot)$.
futures prices later.

Subtracting the estimated seasonality function from the spot data leaves us with the reduced model $Z(\cdot) + Y(\cdot)$, where we have neglected in our notation the fact that we have subtracted only an estimator of $\Lambda(\cdot)$.

Next we want to estimate both components $Z$ and $Y$ invoking the deseasonalised spot price data and the futures prices. We will exploit the fact that the futures prices far from delivery will have a dynamics approximately given by the non-stationary trend component $Z$. Only relatively close to delivery, large fluctuations in the spot price dynamics are reflected in the futures prices. Since it is not clear how far away from delivery we need to be before the approximation of futures prices by $Z$ works well, we will invoke an optimization routine to find the optimal distance. For this purpose, we introduce the notation $u := \frac{1}{2}(T_1 + T_2) - t$, which will be referred to as “time to maturity”.

Denote by $\hat{u}^*$ the optimal time to maturity (we will define what we understand by “optimal” below), where futures contracts with time to maturity $u \geq \hat{u}^*$ have a dynamics approximately behaving like the non-stationary term. How big to choose $\hat{u}^*$ is not possible to determine \textit{a priori}, since we must analyse the error in an asymptotic consideration of the futures prices (see (4.3) and

Figure 5.3: \textit{Base spot prices and estimated seasonality function. Top: whole period (1461 observations). Bottom: first 200 observations.}

\textit{Figure 5.3: Base spot prices and estimated seasonality function. Top: whole period (1461 observations). Bottom: first 200 observations.}
This error is highly dependent on the parameters in the spot price model, which we do not yet know. In the end, \( \hat{u}^* \) should be chosen so that the error in the risk premium estimation is minimal; cf. Section 4.6.

The estimation will, therefore, be repeated using different values of \( u^* \) (this parameter will sometimes be called threshold in the following); i.e., we choose a subset \( U^* := [u^*_{\text{min}}, u^*_{\text{max}}] \subseteq [v/2, M_f] \), where \( v \) is the average delivery period and \( M_f \) is the maximal time to maturity observed in the futures data set, and perform the steps of the Sections 4.2–4.6 below repeatedly for \( u^* \in U^* \). For each value \( u^* \in U^* \) the error in the risk premium is calculated, and \( \hat{u}^* \) is the value which minimizes this error among all \( u^* \in U^* \). This optimal threshold \( \hat{u}^* \) is then considered as final choice of \( u^* \) for the calculation of all estimates including the processes \( Z \) and \( Y \) and for the CARMA parameters.

One should keep in mind that for too large \( u^* \) there is only few data available with \( u \geq u^* \), which yields unreliable estimates. Since we count the time to maturity as number of trading days until the mid of the delivery period (which has length \( v \)), time to maturity is always at least \( v/2 \); hence we do not consider any \( u^* \) smaller than \( v/2 \). Overall, we decided to choose \( u^*_{\text{min}} = [v/2] \) (which is 16 for the base load and 11 for the peak load data) and \( u^*_{\text{max}} = M_f/2 \) (note that \( M_f \) is 200 days for the base load and 144 days for the peak load contracts). As we will see later, the optimal \( \hat{u}^* \) in our data examples is quite small, so that this choice of \( u^*_{\text{max}} \) is completely satisfying.

Next we want to explain in detail, how we separate \( Z \) and \( Y \) for a given fixed time to maturity. Consequently, we perform the model estimation for all \( \frac{1}{2}(T_1 + T_2) \leq u^* \leq 200(146) \) and take all futures prices for the estimation procedure, whose time to maturity \( u \geq u^* \).

We explain each step in the estimation procedure in detail:

### 4.2 Filtering the realization of the non-stationary stochastic process \( Z \)

Recall the futures price \( F(t, T_1, T_2) \) in Corollary 3.2. Since we assume that the high-frequency CARMA term \( Y \) is stationary, it holds for fixed length of delivery \( T_2 - T_1 \) that

\[
\lim_{T_1, T_2 \to \infty} \frac{b^* A^{-1}}{T_2 - T_1} (e^{AT_2} - e^{AT_1}) e^{-AT} X(t) = 0 \tag{4.3}
\]

\[
\lim_{T_1, T_2 \to \infty} \frac{b^* A^{-2}}{T_2 - T_1} (e^{AT_2} - e^{AT_1}) e^{-AT} e_p E_Q[L(1)] = 0. \tag{4.4}
\]

Hence, in the long end of the futures market, the contribution from \( Y \) to the futures prices may be considered as negligible. In particular, from the futures price dynamics (Corollary 3.2) we find for \( [T_1, T_2] \) far into the future (that is, \( t \) much smaller than \( T_1 \)) that

\[
\bar{F}(t, T_1, T_2) := F(t, T_1, T_2) - \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \Lambda(\tau) d\tau
\]
Chapter 5. Futures pricing in electricity markets based on stable CARMA spot models

\begin{align*}
\approx Z(t) + b^* A^{-1} e_p \mathbb{E}_Q[L(1)] + \left( \frac{1}{2} (T_1 + T_2) - t \right) \mathbb{E}_Q[Z(1)].
\end{align*}

(4.5)

Recalling the notation \( u := \frac{1}{2} (T_1 + T_2) - t \) coined “time-to-maturity”, we slightly abuse the notation and introduce \( \tilde{F}(t, u) := \tilde{F}(t, T_1, T_2) \).

For \( u \geq u^* \), we approximate

\[
\mu_{\tilde{F}}(u) := \mathbb{E}[\tilde{F}(t, u)] \\
\approx \mathbb{E}[Z(t)] + b^* A^{-1} e_p \mathbb{E}_Q[L(1)] + u \mathbb{E}_Q[Z(1)] \\
= b^* A^{-1} e_p \mathbb{E}_Q[L(1)] + u \mathbb{E}_Q[Z(1)] \\
=: C + u \mathbb{E}_Q[Z(1)],
\]

(4.6)

where we have used the zero-mean assumption of \( Z \) under \( P \). This approximative identity can now be used for a robust linear regression on the time to maturity \( u \), in order to estimate the real numbers \( C \) and \( \mathbb{E}_Q[Z(1)] \). Knowing these two parameters enables us to filter out the realization of the process \( Z \). According to Equation (4.5) we obtain

\[
\tilde{Z}(t) = \tilde{Z} \left( \frac{1}{2} (T_1 + T_2) - u \right) = \frac{1}{\text{card} U(t, u^*)} \sum_{(u, T_1, T_2) \in U(t, u^*)} \left[ \tilde{F}(t, T_1, T_2) - \hat{C} - u \hat{\mathbb{E}_Q}[Z(1)] \right],
\]

(4.7)

where

\[
U(t, u^*) := \{(u, T_1, T_2) \in \mathbb{R}^3 | u \geq u^* \text{ and } \exists F(t, T_1, T_2) : \frac{1}{2} (T_1 + T_2) - t = u \}.
\]

Remark 4.1. Note that after estimating the CARMA parameters, we can also find an estimate for \( \mathbb{E}_Q[L(1)] \) simply by taking \( \hat{\mathbb{E}_Q}[L(1)] = \hat{C}(b^* A^{-1} e_p)^{-1} \). \( \Box \)

Remark 4.2. We recall that the futures market at EEX is not open for trade during the weekend. Therefore, using our estimation procedure, we do not get any observations of \( Z \) during weekends. We will assume that \( Z \) is constant and equal to the Friday value over the weekend, when filtering the non-stationary part of the spot in the base load model. One may argue that this strategy could lead to large observed jump of \( Z \) on Monday morning, when all information accumulated over the weekend is subsumed at once. We will return to this question in Section 5.1. \( \Box \)

4.3 Estimation of the CARMA parameters

Recall our spot-model (2.1)

\[
S(t) = \Lambda(t) + Y(t) + Z(t), \quad t \geq 0.
\]

After \( \Lambda(\cdot) \) and \( Z(\cdot) \) have been estimated in Subsections 4.1 and 4.2, respectively, a realization of the CARMA-process \( Y \) can be found by subtracting both from the spot price. Figure 5.4 shows
the estimated processes $Z$ (dotted red) and $Y+Z$ (black) for both the base and the peak load data, for the full period July 1, 2002 to June 30, 2006 and for $u^* = 16$ exemplarily, each. Obviously, the process $Z$ captures the medium-range fluctuations and $Y$ the short-range fluctuations of the detrended and deseasonalized process $Y+Z$.

![Estimated processes Z and Y+Z](image1.png)

![Estimated processes Z and Y+Z](image2.png)

Figure 5.4: Estimated processes $Z$ (red) and $Y+Z$ (black) in the period July 1, 2002, to June 30, 2006, for both the base load data (top) and the peak load data (bottom).

Again we keep in mind that the process $Y$ is the result of some estimation procedure. There exists a number of papers devoted to the estimation of the CARMA parameters in $L^2$ (see for instance Brockwell et al. [14], Tsai and Chan [36]). Methods can be based either directly on the continuous-time process or on a discretised version. The latter relates the continuous-time dynamics to a discrete time ARMA process. The advantage of this method is obvious, since standard packages for the estimation of ARMA processes may be used in order to estimate the parameters of the corresponding CARMA process. Some care, however, is needed since this approach does not work in all cases. Brockwell and collaborators devote several papers to the embedding of ARMA processes in a CARMA process; cf. [10, 12]. Not every ARMA($p,q$) process is embeddable in a CARMA($p,q$) process.

From Assumptions 2.2, it follows by Proposition 2 of Brockwell et al. [14] (cf. also Garcia et al. [23], Prop. 2.5) that every CARMA($p,q$) process $Y$ observed at discrete times can be
Chapter 5. Futures pricing in electricity markets based on stable CARMA spot models

represented as an autoregressive process of order $p$ with a more complex structure than a moving average process for the noise. For such a discretely observed CARMA process $Y$ on a grid with grid size $h$, denoting the sequence of observations by $\{y_n\}_{n \in \mathbb{N}}$; i.e. $y_n \triangleq Y(nh)$, Prop. 3 of Brockwell et al. [14] gives

$$
\prod_{i=1}^{p} (1 - e^{\lambda_i B}) y_n = \varepsilon_n. \quad (4.8)
$$

Here, $B$ is the usual backshift operator and $\{\varepsilon_n\}_{n \in \mathbb{N}}$ is the noise process, which has representation

$$
\varepsilon_n = \sum_{i=1}^{p} \kappa_i \prod_{j \neq i} (1 - e^{\lambda_j h B}) \int_{(n-1)h}^{nh} e^{\lambda_i (nh-u)} dL(u). \quad (4.9)
$$

The constants $\kappa_i$ are given by $\kappa_i := b(\lambda_i)/a'(\lambda_i)$ and $\lambda_1, \ldots, \lambda_p$ are the eigenvalues of $A$. The process $\{\varepsilon_n\}_{n \in \mathbb{N}}$ is $p$-dependent. When $L$ has finite variance, $\varepsilon_n$ has a moving average representation; cf. Brockwell et al. [14], Proposition 3.2.1. However, for the case of infinite variance this is no longer true and causes problems.

Davis [18], Davis and Resnick [19] and Mikosch et al. [29] have proved that ordinary $L^2$-based estimation methods for ARMA parameters may be used for $\alpha$-stable ARMA processes, although they have no finite second moments. Moreover, Davis and Resnick [19] showed that the empirical autocorrelation function of an $\alpha$-stable ARMA process yields a consistent estimator of the linear filter of the model, although the autocorrelation function of the process does not exist. Hence the parameters $a_1, \ldots, a_p$ can be consistently estimated by $L^2$-methods. In [19] it has also been shown that the rate of convergence is faster than in the $L^2$-case.

Already in the analysis performed in Garcia et al. [23] the CARMA(2,1) process has been found to be optimal. Although our model is slightly different, it turns out that this CARMA dynamics is still preferrable for $Y$ based on the AIC model selection criterion. Hence, in the following example we spell out the above equations for the case of a stable CARMA(2,1) model.

**Example 4.3.** [The CARMA(2,1) process]

By applying (4.8) and (4.9) for the case of a CARMA(2,1) process, we find the discrete-time representation for a gridsize $h > 0$,

$$
y_n = (e^{\lambda_1 h} + e^{\lambda_2 h}) y_{n-1} - e^{(\lambda_1 + \lambda_2) h} y_{n-2} + \varepsilon_n,
$$

where $\varepsilon_n$ is given by

$$
\varepsilon_n = \int_{(n-1)h}^{nh} \left( \kappa_1 e^{\lambda_1 (nh-u)} + \kappa_2 e^{\lambda_2 (nh-u)} \right) dL(u)
$$

...
Fred Espen Benth, Claudia Klüppelberg, Gernot Müller and Linda Vos

\[ + \int_{(n-2)h}^{(n-1)h} \left( \kappa_1 e^{\lambda_1 h} e^{\lambda_2 (nh-u)} + \kappa_2 e^{\lambda_1 h} e^{\lambda_2 (nh-u)} \right) dL(u). \]

The two integrals in the noise are independent. It is, however, not possible to recover the noise by simple multiplication and subtraction as in the ARMA case. The actual relation of two successive noise terms \( \varepsilon_n \) and \( \varepsilon_{n+1} \) is based on the continuous realization of \( \{L(t)\}_{t \geq 0} \) in the relevant intervals, which is unobservable.

For the mapping of the estimated ARMA parameters to the corresponding CARMA parameters we observe that equation (4.8) is a complex way to express that \( \{e^{-\lambda_i h}\}_{i=1,\ldots,p} \) are the roots of the autoregressive polynomial \( \phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \) of the ARMA process. We proceed, therefore, as follows for identifying the CARMA parameters from the estimated ARMA process:

- Estimate the coefficients \( \phi_1, \ldots, \phi_p \) of the ARMA process
- Determine the distinct roots \( \xi_i \) for \( i = 1, \ldots, p \) of the characteristic polynomial.
- Set \( \lambda_i = -\log(\xi_i)/h \), where we recall that \( h \) denotes the grid size.

Because of the simple structure of the autoregressive matrix \( A \) of the CARMA process we can calculate the characteristic polynomial \( P \) of the matrix \( A \) as

\[ P(\lambda) = (-1)^p \left( \lambda^p + a_1 \lambda^{p-1} + \cdots + a_p \right). \]

Since the \( \lambda_i \) are the eigenvalues of \( A \), we know that \( P(\lambda_i) = 0 \) for \( i = 1, \ldots, p \). Hence, given the eigenvalues \( \lambda_i \) of the matrix \( A \) we recover the coefficients \( a_1, \ldots, a_p \) by solving a system of \( p \) linear equations.

We estimate the moving average parameters based on the autocorrelation function. For its estimation we apply a least absolute deviation algorithm based on the empirical and theoretical autocorrelation functions of the CARMA process. The theoretical autocorrelation function of \( y \) takes the form

\[ \gamma_y(s) = b^* e^{A|s|} \Sigma b, \quad s > 0, \]

where the matrix \( \Sigma \) is given by

\[ \Sigma = \int_0^\infty e^{Au} e_p^* e^{A^* u} du = -A^{-1} e_p e_p^*. \]

In this representation, \( A^{-1} \) is the inverse of the operator \( A : X \mapsto AX + XA^* \) and can be represented as \( \text{vec}^{-1} \circ ((A \otimes I_p) + (I_p \otimes A))^{-1} \circ \text{vec} \) (see Pigorsch and Stelzer [31]). Using the above procedure for the estimation of the moving average parameter \( b \) is based on second order structure and, therefore, not straightforward to use for stable processes. In practice this procedure works and we can use it to estimate the moving average parameter \( b \).
4.4 Estimation of the stable parameters

After estimating the autoregressive parameters, the noise \( \{\varepsilon_n\}_{n \in \mathbb{N}} \) can be recovered. Recall that, by \( p \)-dependence, the noise terms of lags \( m > p \) are independent. Motivated by results for discrete-time stable ARMA processes, Garcia et al. [23] have applied estimation methods for independent noise variables. They have also shown in a simulation study that one gets quite reliable estimates by treating \( R := \{\varepsilon_n\}_{n \in \mathbb{N}} \) as independent sequence.

By a simple computation we can relate the estimated parameters of the series \( R \) back to an estimate of the \( \alpha \)-stable process \( L \). We show this for the CARMA(2,1) model in the next example:

**Example 4.4.** [Continuation of Example 4.3, cf. [23]]

Using Samorodnitsky and Taqqu [33], Property 3.2.2, a relation in distribution between the \( \alpha \)-stable process \( L \) and the noise process \( \{\varepsilon_n\}_{n \in \mathbb{N}} \) of the ARMA(2,1) model sampled on a grid with grid size \( h \) can be established (for clarification, we will now sometimes write \( (\alpha, \gamma, \beta, \mu) \) for the parameters of the \( \alpha \)-stable process \( L \), which were so far denoted just by \( (\alpha, \gamma, \beta, \mu) \)). In particular, \( \varepsilon_n \) has an \( \alpha \)-stable distribution with parameters \( (\alpha, \gamma, \beta, \mu) \) given by

\[
\alpha = \alpha_L, \quad \gamma = \left( \int_0^h \left[ \kappa_1 e^{\lambda_1(h-u)} + \kappa_2 e^{\lambda_2(h-u)} \right]^\alpha \, du \right)^{1/\alpha} \gamma_L, \\
\beta = \beta_L \frac{\gamma^\alpha}{\gamma} \left( \int_0^h (\kappa_1 e^{\lambda_1(h-u)} + \kappa_2 e^{\lambda_2(h-u)})^\alpha (\kappa_1 e^{\lambda_1(h-u)} + \kappa_2 e^{\lambda_2(h-u)})^\alpha \, du \right)^{1/\alpha}, \\
\mu = \mu_L = \mu \quad \text{for } \alpha \neq 1
\]

Note that, for \( a \) and \( p \) being real numbers, \( a^{(p)} := |a|^p \text{sign}(a) \) denotes the signed power (Samorodnitsky and Taqqu [33], eq. (2.7.1)). Moreover, we can easily see that \( \beta = \beta_L \), if both \( \kappa_1 \) and \( \kappa_2 \) are positive.

4.5 Recovering the states

In order to calculate the theoretical futures prices derived in Corollary 3.2 it is necessary to recover the states \( X \) of the CARMA-process. Brockwell et al. [14] describe a rather ad-hoc method to do this by using an Euler approximation.

In the linear state space model (2.2), the Kalman filter is the best linear predictor provided the driving noise is in \( L^2 \). Since \( \alpha \)-stable Lévy processes for \( \alpha \in (0, 2) \) do not have finite second moments, the Kalman filter will perform unsatisfactorily. One possibility to resolve this is to apply a particle filter, which does not require a finite second moment of the noise process. However, the particle filter requires a density function instead, which poses a new problem for \( \alpha \)-stable processes. Integral approximations of \( \alpha \)-stable densities exist, but they are time consuming
to calculate and simple expressions do not exist. One can use a particle filter by simulating from the \(\alpha\)-stable distribution, but this is also very time consuming. A large number of paths need to be simulated in order to get a reasonable estimation (even when using appropriate variance reducing methods like importance sampling). As an attractive alternative, we introduce a simple \(L^1\)-filter applicable to CARMA processes with finite mean.

Recall from (2.2)-(2.6) that we can work with the following state-space representation of the CARMA process

\[
y_n = b^* x_n, \tag{4.10}
\]
\[
x_n = e^{Ah} x_{n-1} + z_n \quad \text{with} \quad z_n = \int_{(n-1)h}^{nh} e^{A(nh-u)} e_p \, dL(u) \tag{4.11}
\]

Here, \(y_n\) and \(x_n\) are discrete observations of \(Y\) and \(X\), respectively, on a grid with grid size \(h\).

Notice that given \(y_n\) and \(x_{n-1}\) the value of \(b^* z_n\) is determined and given by

\[
b^* z_n \mid y_n, x_{n-1} = y_n - b^* e^{Ah} x_{n-1}. \tag{4.12}
\]

This will come to use in a moment when deriving the filter.

First, we make an "Euler" approximation of the stochastic integral defining \(z_n\) by

\[
z_n \approx \frac{1}{h} \int_{(n-1)h}^{nh} e^{A(nh-u)} e_p \, du \Delta L(n, h) = -A^{-1} (I - e^{Ah}) \frac{\Delta L(n, h)}{h},
\]

where \(\Delta L(n, h) = L(nh) - L((n-1)h)\). Note that a traditional Euler approximation (see Kloeden and Platen [24]) would use the left end-point value of the integrand in the approximation, whereas here we use the average value of the integrand over the integration interval. We find

\[
\mathbb{E}[z_n \mid y_n, x_{n-1}] \approx -\mathbb{E}[\Delta L(n, h)/h \mid y_n, x_{n-1}] A^{-1} (I - e^{Ah}) e_p. \tag{4.13}
\]

Multiplying (4.13) with \(b^*\) and combining it with (4.12) gives

\[
\mathbb{E}[\Delta L(n, h)/h \mid y_n, x_{n-1}] \approx \frac{y_n - b^* e^{Ah} x_{n-1}}{-b^* A^{-1} (I - e^{Ah}) e_p}. \tag{4.14}
\]

By plugging (4.14) into (4.13) we find

\[
\mathbb{E}[z_n \mid y_n, x_{n-1}] \approx -A^{-1} (I - e^{Ah}) e_p \frac{y_n - b^* e^{Ah} x_{n-1}}{-b^* A^{-1} (I - e^{Ah}) e_p}. \tag{4.15}
\]

We can use this as an \(L^1\)-filter for \(z_n\). Applying (4.15), we can filter the states \(X\) of the CARMA-
process. Using the state equation (4.11) we find
\[
\mathbb{E}[x_n \mid y_n, x_{n-1}] \approx e^{Ah}x_{n-1} + \mathbb{E}[z_n \mid y_n, x_{n-1}].
\] (4.16)

We tested the filter on simulated data from a CARMA(2,1) process with the same parameters as we find from our model for the base spot prices (see Table 5.3). The path of the CARMA(2,1) process was simulated based on an Euler scheme on a grid size of 0.01 for \(0 \leq t \leq 1461\), and the \(\alpha\)-stable Lévy process was simulated using the algorithm suggested by Chambers, Mallows and Stuck [16]. The estimation of the states is done on a grid with grid size \(h = 1\). Figure 5.5 shows the estimated states (red curve) for both state components together with the simulated states (black curve). It is clearly visible that the \(L^1\)-filter gives a good approximation of the true states \(X\) driving the \(\alpha\)-stable CARMA process \(Y\).

![Simulated and filtered states](image)

**Figure 5.5:** Estimated states (red) and true states (black) of a simulated CARMA(2, 1) process using the \(L^1\)-filter.

### 4.6 Risk premium comparison

In order to find the optimal threshold \(\hat{u}^*\) for filtering out the non-stationary process \(Z\) from the futures data, we compare the empirically observed risk premium with its theoretical counterpart.
Recall the risk premium $R_{pr}$ in (3.4) implied by the futures price dynamics in Corollary 3.2. By using $v = T_2 - T_1$ and recalling the notation $u = \frac{1}{2}(T_1 + T_2) - t$, we can rewrite $R_{pr}$ for $u \geq \frac{1}{2}(T_2 - T_1)$ and fixed $v$ (being one month in our studies) to

$$R_{pr}(u^*, u, v) = -\frac{1}{v} b^* A^{-2} \left( e^{\frac{1}{2} A_v} - e^{-\frac{1}{2} A_v} \right) e^{A_u} e_p \left( \mathbb{E}_Q[L(1)] - \mathbb{E}[L(1)] \right) + b^* A^{-1} e_p \left( \mathbb{E}_Q[L(1)] - \mathbb{E}[L(1)] \right) + u \mathbb{E}_Q[Z(1)].$$

Note that we can estimate all parameters in (4.17) only depending on a chosen threshold $u^*$. Hence, the risk premium in Equation (4.17) also depends on $u^*$. In order to find an optimal $\hat{u}^*$ we compare the risk premium (4.17), which has been estimated on our model assumptions, with the mean empirical risk premium based on the futures prices given by

$$\tilde{R}_{pr}(u^*, u, v) := \frac{1}{v} b^* A^{-2} \left( e^{\frac{1}{2} A_v} - e^{-\frac{1}{2} A_v} \right) e^{A_u} e_p \mathbb{E}[L(1)] - b^* A^{-1} e_p \mathbb{E}[L(1)]$$

$$+ \frac{1}{\text{card} U(u, v)} \sum_{t, T_1, T_2 \in U(u, v)} \left[ F(t, T_1, T_2) - \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \Lambda(\tau) d\tau \right]$$

$$- \frac{b^* A^{-1}}{T_2 - T_1} \left( e^{A T_2} - e^{A T_1} \right) e^{-A t} X(t) - Z(t).$$

Here,

$$U(u, v) := \left\{ t, T_1, T_2 \in \mathbb{R} : \frac{1}{2}(T_2 + T_1) - t = u, T_2 - T_1 = v \text{ and } F(t, T_1, T_2) \text{ exists} \right\}.$$ 

The dependence on the threshold $u^*$ is only implicit. The estimated sample paths of $Z$ and $Y$ depend on $u^*$, therefore, also the CARMA parameters $A, b$, the stable parameters $(\alpha, \beta, \gamma, \mu)$ and the estimated sample paths of the states $X$ also depend on $u^*$. In order to compute $\tilde{R}_{pr}$ and $R_{pr}$ all these estimated parameters are used. Consequently, by using different thresholds we will get different estimates and different risk premia. We want to choose an optimal threshold $\hat{u}^*$, such that the mean empirical risk premium $\tilde{R}_{pr}$ is as close as possible to the model based risk-premium $R_{pr}$. We invoke a least squares method, i.e. we minimize (for fixed $v$) the mean square error between the two functions $(\tilde{R}_{pr}(u^*, u, v))_{u \geq \frac{1}{2}v}$ and $(R_{pr}(u^*, u, v))_{u \geq \frac{1}{2}v}$ with respect to all chosen thresholds $u^* \in U^*$ for the estimation procedure, cf. Section 4.1.

$$\hat{u}^* = \arg\min_{u^* \in U^*} \sum_{u = v/2}^{M_T} |\tilde{R}_{pr}(u^*, u, v) - R_{pr}(u^*, u, v)|^2.$$
Here the dependence of the error function
\[ f(u^*, v) := \sum_{u=v/2}^{M_f} |\tilde{R}_{pr}(u^*, u, v) - R_{pr}(u^*, u, v)|^2 \quad (u^* \in U^*) \] (4.19)
on \( u^* \) is only implicit. In our data \( v \) corresponds to the average number of days per month (i.e. \( v = 1461/48 = 30.44 \) for the base data and \( v = 1045/48 = 21.77 \) for the peak data), and the number \( M_f \) is the longest time to maturity, which we recall from Section 4.1 being 200 for the base load contracts and 144 for the peak load. In order to calculate this minimum we calculate the values of \( f(u^*, v) \) for all \( u^* \in [v/2, M_f/2] \cap \mathbb{N} \). Figure 5.6 shows the risk premium error function \( f(u^*, v) \) for base load (left) and peak load (right). In both cases, the minimum is attained at \( \hat{u}^* = 16 \). So our estimation procedure considers only base load forward contracts with delivery at least (about) two weeks away, and peak load contracts with delivery at least (about) three weeks away.

Below we present a summary of the estimation algorithm again.

**The algorithm**

Estimate \( \Lambda(\cdot) \) as in (4.2) for the base load model and as in (4.1) for peak load model, and subtract from \( S(\cdot) \).

For each threshold \( u^* \in U^* \):

- Approximate \( \mu_{\mathcal{F}}(u) = C + u \mathbb{E}_Q[Z(1)] \) for \( u \geq u^* \) and estimate \( C, \mathbb{E}_Q[Z(1)] \) by linear regression (4.6);
- filter \( Z \) by (4.7);
• model $Y = S - \Lambda - Z$ as CARMA(2,1) process, estimate the coefficients $a_1, a_2, b_0$ (recall that $b_1 = 1$) and estimate the parameters $(\alpha_L, \gamma_L, \beta_L, \mu_L)$ of $L$;

• estimate $\mathbb{E}_Q[L(1)]$ using (4.7);

• filter states of $X = (X_1, X_2)^*$ using (4.16);

• calculate $R_{pr}(u^*, u, v)$ as in (4.17) invoking the estimated parameters and states from the former steps;

• calculate $\tilde{R}_{pr}(u^*, u, v)$ as in (4.18) invoking the estimated parameters and states from the former steps and the futures data.

Now define the mean square error of the estimated $R_{pr}(u^*, u, v)$ and $\tilde{R}_{pr}(u^*, u, v)$ based on all different thresholds $u^* \in U^*$. The optimal threshold is found to be $\hat{u}^*$.

5 Estimation results

We now report the other results from the estimation procedure, when using the optimal threshold $\hat{u}^* = 16$ both for base and peak load data. In this section we discuss the estimated values and their implications.

5.1 Distributional properties of the filtered sample path of $Z$

For the filtered $Z$ which was found using (4.7) we can derive certain properties. Both for the base and the peak data, the realization of $Z$ shows uncorrelated increments, see Figure 5.7.

![acf for increments of Z (base)](image)

![acf for increments of Z (peak)](image)

Figure 5.7: Empirical autocorrelation functions for the increments of $Z$, for base data (left) and peak data (right).
Chapter 5. Futures pricing in electricity markets based on stable CARMA spot models

<table>
<thead>
<tr>
<th></th>
<th>$\alpha_Z$</th>
<th>$\beta_Z$</th>
<th>$\delta_Z$</th>
<th>$\mu_Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base load</strong></td>
<td>0.6451</td>
<td>0.0998</td>
<td>0.2206</td>
<td>-0.0346</td>
</tr>
<tr>
<td><strong>Peak load</strong></td>
<td>0.2371</td>
<td>-0.0083</td>
<td>0.6582</td>
<td>0.0230</td>
</tr>
</tbody>
</table>

Table 5.2: Estimated parameters of the NIG distribution $Q$ for the increments of $Z$. Since we assume that $E[Z] = 0$, the parameters have been estimated conditionally on this assumption.

Figure 5.8 shows QQ-plots for the increments of $Z$ versus a corresponding normal distribution, both for base data (left) and peak data (right). Note that the empirical variances of the increments are 0.35 and 2.78 for the base and peak data, respectively. From these plots we conclude that for both data sets the increments of $Z$ have heavier tails than the Gaussian distribution, and that this feature is even more pronounced for the peak data.

Kernel density estimates suggest that the increments of $Z$ can be described quite well using a normal inverse Gaussian (NIG) distribution (we refer to Barndorff-Nielsen [1] for a thorough discussion on the NIG distribution and its properties). The red curves in Figure 5.9 show log-density estimates for the increments of $Z$ for the base data (left) and the peak data (right), respectively. For comparison, we also plot the log-density curves of NIG distributions (black solid curves) and normal distributions (black dashed curves) that have been fitted to the increments of $Z$ via maximum likelihood (the parameters for the NIG distributions can be found in Table 5.2). Clearly, the NIG distribution gives a much better fit than the normal distribution. Hence, we identify the non-stationary process $Z$ with a normal inverse Gaussian Lévy process.

**Remark 5.1.** Recall that futures contracts are only traded on weekdays and, therefore, no variability in $Z$ during weekends is observed. For base load contracts we have thus assumed that $Z$ is constant during weekends for filtering purposes, but only considered weekdays data for analysing
the distributional properties of $Z$. As we already mentioned in Section 4.2, this strategy could lead to larger up– or downward movements of $Z$ on Mondays, when all information accumulated over the weekend is subsumed at once. However, we do not find such a behaviour in the estimated increments of $Z$. To see this, we calculated the variance of the increments of $Z$ in the base data set for different weekdays separately. We find a variance of 0.3723 for the increments occurring from Fridays to Mondays, and a variance of 0.3526 for the remaining increments, i.e. Mon/Tue, Tue/Wed, Wed/Thu, Thu/Fri. For comparison, the corresponding variance from Wednesdays to Thursdays only is 0.3202, whereas for the remaining increments, i.e. Mon/Tue, Tue/Wed, Thu/Fri, Fri/Mon, the value is 0.3649. Furthermore, the estimated overall variance of the increments of $Z$ is 0.35. Taking the estimation error for all these variances into account, we do not observe a significantly higher variance for the increments from Friday to Monday in the base data. This supports the idea that no relevant information enters the futures market over the weekends. Maybe this even can be expected, since the futures deal with a product to be delivered quite far in the future; hence, only really influential news with a long-range impact should affect the market.

\[\square\]

### 5.2 Estimation of the CARMA parameters

The autocorrelation and partial autocorrelation function of the data suggest that there are two significant autoregressive lags, but also a relevant moving average component. Also the AIC and BIC criterion confirm that a CARMA(2,1) model leads to the best fit. The estimated parameters of a CARMA(2,1) model are given in Table 5.3.

For the estimated parameters $(\hat{a}_1, \hat{a}_2)$ in the autoregression matrix $A$ the eigenvalues of $A$ are real and strictly negative, being $\lambda_1 = -0.0641, \lambda_2 = -1.4213$ for the base load and $\lambda_1 =$
Chapter 5. Futures pricing in electricity markets based on stable CARMA spot models

<table>
<thead>
<tr>
<th>CARMA parameters</th>
<th>Stable parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>$a_2$</td>
</tr>
<tr>
<td><strong>Base load</strong></td>
<td>1.4854</td>
</tr>
<tr>
<td><strong>Peak load</strong></td>
<td>2.3335</td>
</tr>
</tbody>
</table>

Table 5.3: Estimates of the CARMA parameters and of the parameters of the stable process $L$.

<table>
<thead>
<tr>
<th></th>
<th>$\mathbb{E}_Q[Z(1)]$</th>
<th>$\tilde{C}$</th>
<th>$\mathbb{E}_Q[L(1)]$</th>
<th>$\tilde{\theta}_Z$</th>
<th>$\tilde{\theta}_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base load</strong></td>
<td>$-0.0243$</td>
<td>1.6587</td>
<td>$-0.5282$</td>
<td>$-0.1093$</td>
<td>$-0.0021$</td>
</tr>
<tr>
<td><strong>Peak load</strong></td>
<td>$-0.0382$</td>
<td>3.5678</td>
<td>$-1.3178$</td>
<td>$-0.0168$</td>
<td>$-0.0552$</td>
</tr>
</tbody>
</table>

Table 5.4: Estimates of parameters determining the risk.

$-0.1014, \lambda_2 = -2.2319$ for the peak load. Our parameters satisfy Assumptions 2.2. Hence, the estimated model is stationary.

The estimates of $\alpha_L$ (1.6524 for base and 1.3206 for peak) confirm that extreme spikes are more likely in the peak load data. As we can conclude from the positive signs of the skewness parameter $\beta_L$, positive spikes are more likely to happen than negative spikes for both data sets. Note, however, that a direct comparison of the values of $\beta_L$ for the base and the peak load data is misleading, due to the significantly different parameters $\alpha_L$. Indeed, if we calculate the empirical skewness for the estimated $\varepsilon_n$ (cf. Section 4.4) directly, we get a value of 0.28 for the base load data and 1.59 for the peak load data (cf. the comment on $\beta_\varepsilon = \beta_L$ in Example 4.4; for the base load data $(\kappa_1, \kappa_2) = (0.1636, 0.8364)$, and for the peak load data $(\kappa_1, \kappa_2) = (0.2400, 0.7600)$).

5.3 Market price of risk and risk premium

We next present the results on the risk premium and the parameters for the market price of risk, based on our statistical analysis of base and peak load contracts with threshold $\hat{a}^* = 16$ days in both cases. Estimates of the relevant parameters are presented in Table 5.4.

Recall that $\mathbb{E}_Q[Z(1)]$ and $\tilde{C}$ are found from regression (4.6); using the estimates of the CARMA model from Table 5.3, we derive an estimate of $\mathbb{E}_Q[L(1)]$. Having both $\mathbb{E}_Q[L(1)]$ and $\mathbb{E}_Q[Z(1)]$ estimated, we can compute the parameters in the respective measure transforms of the NIG Lévy process $Z$ and the stable process $L$. For an NIG Lévy process we use the fact that an Esscher transformed NIG($\alpha_Z, \beta_Z, \delta_Z, \mu_Z$) random variable $Z$ is again NIG distributed with parameters $(\alpha_Z, \beta_Z + \theta_Z, \delta_Z, \mu_Z)$ (see e.g. Benth et al. [3], p. 99). Using the mean of an NIG distributed random variable it holds that

$$\mathbb{E}_Q[Z(1)] = \mu_Z + \frac{\delta_Z (\beta_Z + \theta_Z)}{\sqrt{\alpha_Z - (\beta_Z + \theta_Z)^2}}.$$
where $\theta_Z$ is the market price of risk for $Z$. Since estimates for the parameters $\alpha_Z$, $\beta_Z$, $\delta_Z$ and $\mu_Z$ are known from Table 5.2, we can use the estimate for $E_Q[Z(1)]$ together with the above equality to obtain an estimate of $\theta_Z$, which results here in $\hat{\theta}_Z = -0.1093$ for the base load and in $\hat{\theta}_Z = -0.0168$ for the peak load data. Since $\theta_Z$ is estimated negative, more emphasis is given to the negative jumps and less emphasis to the positive jumps of $Z$ in the risk neutral world $Q$. We see from the estimate on the risk-neutral expectation of $Z$ that the contribution from the non-stationarity factor of the spot on the overall risk premium is negative. This is natural from the point of view of the hedging needs of producers. The non-stationary factor induces a long-term risk, which is the risk producers want to hedge using futures contracts.

Lucia and Schwartz [28] also find a negative market price of risk associated to the non-stationary term in their two-factor models, when analysing data from the NordPool market. We recall that they propose a two-factor model, where the non-stationary term is a drifted Brownian motion. The negative market price of risk appears as a negative risk-neutral drift, which corresponds to a contribution to the risk premium similar to our model. We refer to Benth and Sgarra [6] for a theoretical analysis of the Esscher transform in factor models applied to power markets.

Using the relations $\gamma^\alpha = c_+ + c_-\gamma^\alpha$ and $\beta = (c_+ - c_-)/(c_+ + c_-)$ in the stable parameters as calculated in Example 2.3.3 of Samorodnitsky and Taqqu [33] and plugging in the estimated parameters from Table 5.3, we find

$$\hat{c}_+ = \frac{1}{2}(1 + \beta_L)\gamma^\alpha_L \approx 14.9715 \quad \text{and} \quad \hat{c}_- = \frac{1}{2}(1 - \beta_L)\gamma^\alpha_L \approx 6.5532$$

for the base load data, and

$$\hat{c}_+ = \frac{1}{2}(1 + \beta_L)\gamma^\alpha_L \approx 6.3342 \quad \text{and} \quad \hat{c}_- = \frac{1}{2}(1 - \beta_L)\gamma^\alpha_L \approx 5.5587$$

for the peak load data. Then by using (3.9) we can derive an estimate for $\theta_L$ by

$$\hat{\theta}_L = -\left(\frac{E_Q[L(1)] - E[L(1)]}{\Gamma(1 - \alpha_L)(\hat{c}_+ - \hat{c}_-)}\right)^{\frac{1}{\alpha_L - 1}}$$

which leads to $\hat{\theta}_L = -0.0021$ for the base lead data and to $\hat{\theta}_L = -0.0552$ for the peak load data. The market price of risk for the CARMA-factor noise $L$ is also negative, however, unlike the non-stationary factor a negative sign does not necessarily lead to a negative contribution to the risk premium. As we already see in the estimate of the constant $C$ in the regression (4.6), we get a positive contribution to the risk premium. There will also be a term involving time to maturity, which will converge to zero in the long end of the futures curve. This part of the risk premium may contribute both positively or negatively. The CARMA factor is thus giving a positive risk premium for contracts, which start delivering reasonably soon. Since this factor is accounting for
Chapter 5. Futures pricing in electricity markets based on stable CARMA spot models

Figure 5.10: The estimated risk premium $R_{pr}$ (red) and the empirical risk premium $\tilde{R}_{pr}$ (black), for base load data (left) and peak load data (right).

the short term variations, and in particular the spike risk of the spot, we may view this as a result of consumers and retailers hedging their price risk and, therefore, accepting to pay a premium for this. This conclusion is in line with the theoretical considerations of Benth, Cartea and Kiesel [4], who showed – using the certainty equivalence principle – that the presence of jumps in the spot price dynamics will lead to a positive risk premium in the short end of the futures curve. Bessembinder and Lemmon [8] explain the existence of a positive premium in the short end of the futures market by an equilibrium model, where the skewness in spot prices induced by spikes is a crucial driver.

As we know from the third summand in Equation (4.17), the risk premium is dominated by a linear trend for most times to maturity except very short ones. In the latter case, the first summand of Equation (4.17) is dominating, leading to a small exponential decay when time to maturity tends to 0. A plot of the empirical risk premium versus the theoretical one is given in Figure 5.10.

We see that for the base load contracts the positive risk premium in the short end of the curve is not that pronounced, however, it is detectable. The risk premium is negative for contracts starting to deliver in about two months. On the other hand, the peak load contracts have a clear positive risk premium, which changes to a negative one for contracts starting to deliver in about four months. This form of the risk premium is in line with the analysis of Geman and Vasicek [22]. Interesting here is the difference between base and peak load contracts. Base load futures have a longer delivery period than peak loads, since they are settled against more hours. This means that extreme prices are more smoothed out for base load contracts. The sensitivity towards spikes are even more pronounced in peak load contracts, since they concentrate their settlement for the hours, where typically the extreme spikes occur, and ignore to night hours where prices are usually lower and thus would smooth out the spikes. Hence, peak load contracts are much more spike sensitive than base load contracts, which we see reflected in the risk premium having a
bigger and more visible positive part in the short end of the futures curve. The study of Longstaff and Wang [27] on the PJM (Pennsylvania, New Jersey and Maryland) market shows that the risk premium may vary over time, and indeed change sign. Their analysis is performed on hourly prices in the balancing market as being the spot, and the day ahead hourly prices as being futures contracts. Hence, the analysis by Longstaff and Wang [27] is valid for the very short end of the futures curve.

Our findings for the EEX are in line with the empirical studies in Benth et al. [4], which applies a two factor model to analyse the risk premium in the EEX market. Their model consists of two stationary processes, one for the short term variations, and another for stationary variations mean-reverting at a slower speed. Their studies confirm a change in sign of the risk premium as we observe for our model. Moreover, going back to Lucia and Schwartz [28], they find a positive contribution to the risk premium from their short-term variation factor, when applying their analysis to NordPool data. This shows that also in this market there is a tendency towards hedging of spike risk in the short end of the futures curve. On the other hand, our results for the EEX market are at stake with the findings in Kolos and Ronn [26]. They perform an empirical study of many power markets, where they estimate market prices of risk for a two-factor Schwartz and Smith model. In the EEX market, they find that both the short and the long term factors contribute negatively\(^1\) for the case of the EEX market. However, interestingly, the PJM market in the US, which is known to have huge price variations with many spikes observed, they find results similar to ours. We find our results natural in view of the spike risk fully accounted for in the short term factor, and the natural explanation of the hedging pressure from producers in the non-stationary factor. Our statistical analysis also strongly suggest non-Gaussian models for both factors, which is very different to the Gaussian specification of the dynamics in Kolos and Ronn [26].

6 Conclusion

In this paper we suggest a two-factor arithmetic spot model to analyse power futures prices. After removal of seasonality, a non-stationary long term factor is modelled as a Lévy process, while the short term variations in the spot price is assumed to follow a stationary stable CARMA process. An empirical analysis of spot price data from the German power exchange EEX shows that a stable CARMA processes is able to capture the extreme behavior of electricity spot prices, as well as the more normal variations when the market is in a quite period.

As in Lucia and Schwartz [28] we use a combination of a deterministic function and a non-stationary term to model the low frequency long term dynamics of the spot. Empirical data suggests that futures curves and spot prices are driven by a common stochastic trend, and it turns out that this is very well described by a normal inverse Gaussian Lévy process. This leads to

\(^1\)In their paper, the signs are positive due to the choice of parametrization of the market price of risk
realistic predictions of the futures prices. Moreover, a CARMA(2,1) process is statistically the best model for the short term variations in the spot dynamics.

We apply the Esscher transform to produce a parametric class of market prices of risk for the non-stationary term. The \( \alpha \)-stable Lévy process driving the CARMA-factor is transformed into a tempered stable process in the risk neutral setting. The spot price dynamics and the chosen class of risk neutral probabilities allow for analytic pricing of the futures. A crucial insight in the futures price dynamics is that the stationary CARMA effect from the spot price is vanishing for contracts far from delivery, where prices essentially behave like the non-stationary long-term factor.

We propose a statistical method to calibrate the suggested spot and futures model to real data. The calibration is done using spot and futures data together, where we applied futures prices in the far end of the market to filter out the non-stationary factor in the spot. We choose a threshold for what is sufficiently “far out” on the futures curve by minimizing the error in matching the theoretical risk premium to the empirical. In this minimization over thresholds, we need to re-estimate the whole model until the minimum is attained. Since \( \alpha \)-stable processes are not in \( L^2 \), we introduce a robust \( L^1 \)-filter in order to recover the states of the CARMA process required for the estimation of the risk premium.

Our model and calibration technique is used on spot and futures data collected at the EEX. Moreover, in order to gain full insight into the risk premium structure in this market, we study both peak load and base load futures contracts with delivery over one month. The base load futures are settled against the hourly spot price over the whole delivery period, while the peak load contracts only deliver against the spot price in the peak hours from 8 a.m. to 8 p.m. on working days. Our model and estimation technique seem to work well in both situations.

We find that the base load futures contracts have a risk premium which is close to linearly decaying with time to delivery. The risk premium is essentially governed by the long term factor. There is evidence of a positive premium in the short end of the futures curve. For peak load contracts, which are much more sensitive to spikes, the positive premium in the short end is far most distinct, but also here the premium decays close to linearly in the long end of the market. These observations are in line other theoretical and empirical studies of risk premia in electricity markets, which argue that the risk premia in power markets are driven by hedging needs. Our findings also show that we should have an exponential dampening of the premium towards maturity, resulting from the CARMA factor of the spot.

Acknowledgments

The STABLE package is a collection of algorithms for computing densities, distribution functions, quantiles, simulating and estimating for \( \alpha \)-stable distributions. A free basic version is available from academic2.american.edu/~jpnolan; an advanced version is available from www.robustanalysis.com. We are grateful to John Nolan for making the packages available.
FEB gratefully acknowledges the financial support from the project "Energy markets: modelling, optimization and simulation (EMMOS)", funded by the Norwegian Research Council under grant 205328/v30.

CK gratefully acknowledges financial support by the Institute for Advanced Study of the Technische Universität München (TUM-IAS).
Bibliography


Bibliography


[38] EEX Product Specification. Available at http://www.eex.com