Electricity Price Modelling with Stochastic Volatility and Jumps: An Empirical Investigation

This version: Monday 30th April, 2018

Nikolay Gudkov[a,*], Katja Ignatieva[b]

[a]*School of Risk and Actuarial Studies, Business School, UNSW Sydney, NSW 2052, Australia*

[b]*School of Risk and Actuarial Studies, Business School, UNSW Sydney, NSW 2052, Australia*

Abstract

The market of electricity derivatives has experienced a substantial growth in the volume of trade and the diversity of available products over the past few years. This has led to a rich data environment that requests more sophisticated and accurate modelling approaches for electricity spot prices. This paper deals with the analysis of continuous-time stochastic volatility jump-diffusion processes in the context of pricing of futures contracts written on electricity spots. We formulate a model, which aims to capture the most prominent characteristics and stylised facts of the electricity spot market including mean reversion, seasonality, extreme volatility and spikes. The proposed modelling framework extends the already existing models by incorporating mean reversion, stochastic volatility and jumps in both, the underlying spot price process and its volatility. The model parameters are estimated using the Markov Chain Monte Carlo (MCMC) technique for the Australian electricity market, which highly liquid and can be analysed using pricing applications. Using the market price of risk estimated from the futures market, we compute futures prices in a closed or semi-closed form and demonstrate that the model fits data well in-sample and out-of-sample.

Keywords: Power markets, electricity modeling, energy derivatives, jump diffusion models, futures pricing

*Corresponding author

Email addresses: n.gudkov@unsw.edu.au (Nikolay Gudkov), k.ignatieva@unsw.edu.au (Katja Ignatieva)

1. Introduction

Energy derivatives market has increased substantially in the recent years, resulting in significant increase of trading volumes and large variety of offered products. Models for dynamics of electricity spot prices lie at the heart of derivatives pricing and risk management. Accurate valuation energy derivatives contracts and reliable risk management strategies rely heavily on the specified modelling framework, which must capture the most important characteristics and stylised facts of electricity spot prices. These include *mean reversion* (see Schwartz (1997) and Weron (2006)), which is much stronger compared to financial assets or other commodities; *price spikes* (Kaminski (1999) and Weron et al. (2004)), which are short-lived phenomena related to non-storability of electricity, inelasticity of supply and demand, generator capacity constraints and outages; *extreme volatility* whereby small changes in load or generation can cause large changes in price; and seasonality (Pilipovic (1997) and Kaminski (1999)), which is attributed to seasonal patterns observed in electricity prices during the course of a day, week and year.

In order to model the above mentioned characteristics, the literature typically considers the regime-switching models (Huisman and Mahieu, 2003; Haldrup and Nielsen, 2006; Bierbrauer et al., 2007; Janczura and Weron, 2010), or diffusion models where some authors add jumps, stochastic volatility and stochastic equilibrium level as additional risk factors to account for spikes and mean reversion (Cartea and Figueroa, 2005; Geman and Roncoroni, 2006; Kluge et al., 2009). Furthermore, several research works consider additional exogenous variables such as weather or electricity demand (Mount et al., 2006; Huisman, 2008; Kanamura and Ohashi, 2008) which might be helpful to better capture the spiky behaviour of the spot prices. Although jump-diffusion and regime-switching models offer the best alternatives, there is always a trade-off between model parsimony and adequacy in capturing the unique characteristics of electricity prices. If the underlying price process is chosen inappropriately, it will fail to capture the main characteristics of electricity prices, leading to unreliable results. On the other hand, if the model is too complex, the computational burden will prevent its on-line use in trading departments.

The objective of this paper is to extend modelling frameworks proposed in the literature by means of incorporating several factors that have not been considered jointly in the existing literature. Ornstein-Uhlenbeck type of model which accounts for mean reversion of prices has been introduced in Schwartz (1997), and Lucia and Schwartz (2002b) extend the range of these models to the two-factor models which incorporate a deterministic seasonal component. Although these models for the spot price dynamics capture mean reversion in electricity prices, they fail to account for price spikes. A natural extension is to incorporate a jump component, which was first introduced in Merton (2001) to model equity dynamics, has first appeared in Cartea and Figueroa (2005) in relation to electricity spot price modelling. The authors present model which combines mean reversion, jumps and seasonality, and calculate price of the forward contracts in a closed form.

To our knowledge, none of the existing models for electricity spot prices consider all four risk factors simultaneously: mean reversion, jumps, seasonal component and stochastic volatility. Although mean-reverting price process with stochastic volatility and jumps have been proposed in various papers with financial assets as underlying (see (Jacquier et al., 2004; Pan, 2002; Bakshi et al., 1997; Chernov et al., 2003)), none of the papers considers electricity spot prices as underlying assets.

The main contribution of the present paper is twofold. Firstly, we present a model that captures the most important characteristics of electricity spot prices such as mean reversion, extreme volatility, jumps and seasonality. Thereby, the most general model we consider is the mean-reversing model with seasonality, stochastic volatility and jumps. We compare this model to its less heavily parameterised counterparts, namely, the mean-reversing model with seasonality, stochastic volatility without jumps; as well as the model with deterministic volatility with or without jumps. The parameters from the model specifications are estimated using Markov Chain Monte Carlo (MCMC) technique, applied to the electricity spot prices in the Australian electricity market. Secondly, using market price of risk estimated from electricity futures, we calculate an expression for the futures prices in the closed form and show that the model fits well in-samples, as well as when dealing with prediction of futures prices.

The paper is organized as follows. Section 2 considers different model specifications for modeling electricity spot price dynamics and derives closed form formula for futures prices. Section 3 discusses the MCMC estimation approach. Section 4 presents the diagnostic tools for quantification of the model performance. Our empirical results using data from the Australian electricity markets are presented in Section 5. Finally, Section 6 summarizes the findings.

2. Model Specifications

Let $\Xi = \{\Xi_t, t \geq 0\}$ denotes the electricity spot price that can be decomposed into two parts as $\Xi_t = U_t \cdot S_t$, where U_t and S_t are deterministic and stochastic components of the spot price, respectively. In this paper we concentrate on quarterly nearly expiration futures written on the average spot price of electricity over extended periods of time. Specifically, we consider futures contacts with expiration on March 31, June 30, September 30 and December 31 (refer to Section 5.1). The value at time t of the futures contract maturing at time T can be computed as

$$F_t(\Xi_t, T) = \frac{1}{T}\left\{\sum_{s=1}^{t} \Xi_s + \sum_{s=t+1}^{T} \mathbb{E}_t^{\mathbb{Q}}[\Xi_s]\right\} = \frac{1}{T}\left\{\sum_{s=1}^{t} U_s \cdot S_s + \sum_{s=t+1}^{T} U_s \mathbb{E}_t^{\mathbb{Q}}[S_s]\right\}, \tag{2.1}$$

where the expected value is computed under the risk-neutral measure $\mathbb{Q}$. Thus, the futures price is given by the the weighted combination of realised spot prices up to time t, and the the expected spots from $t+1$ to maturity, T, given the information up to time t.

In this section we discuss different model specifications for modelling the stochastic component S_t of the electricity spot price dynamics. These include mean-reverting models with deterministic volatility and with or without jumps, as well as models with stochastic volatility, again with or without jumps. For each of the model specifications we derive the closed form formula for expectations, $\mathbb{E}^{\mathbb{Q}}[S_s]$.

2.1. Mean-Reverting Model with Deterministic Volatility

We assume that under the physical measure $\mathbb{P}$, stochastic component S_t of the electricity spot price follows the following stochastic differential equation (SDE):

$$dS_t = a(S_t)S_t dt + \sigma S_t dW_t^S, \tag{2.2}$$

where $a(S_t)$ is a drift function and dW_t^S is standard Brownian motion. Applying Itô lemma to $X_t = \log(S_t)$, we obtain

$$dX_t = \left(a^X(X_t) - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t^X, \tag{2.3}$$

where $a^X(X_t) = a(e^{X_t})$ and $dW_t^X = dW_t^S$. Since the purpose of our empirical analysis is to price futures contracts we assume that $a^X(X_t)$ is a mean-reverting processes of the form

$$a^X(X_t) = \eta(\mu - X_t), \tag{2.4}$$

where μ is a long-run mean of X_t and η is a speed of mean reversion. Thus, under the physical measure $\mathbb{P}$ the dynamics of X_t reads

$$dX_t = \left(\eta(\mu - X_t) - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t^X = \eta(\hat{\mu} - X_t)dt + \sigma dW_t^X, \tag{2.5}$$

where

$$\hat{\mu} = \mu - \frac{\sigma^2}{2\eta}. \tag{2.6}$$

Since electricity is not a storable commodity, the risk-neutral hedging argument does not apply. In this case, when rewriting the price process under the risk neutral measure, it does not necessarily hold that the expected return corresponds to the risk-free interest rate, which is the case when the underlying is a traded asset. In other words, the expected return of S_t under the risk-neutral measure $\mathbb{Q}$ may differ from r. With a change of the probability measure the standard Brownian motion under the risk-neutral measure becomes

$$d\tilde{W}_t^X = dW_t^X + \lambda dt, \tag{2.7}$$

where λ denotes the market price of risk for the electricity price. Thus, the following dynamics for X_t under the risk-neural measure, $\mathbb{Q}$ can be obtained:

$$dX_t = \left(\eta(\mu - X_t) - \frac{1}{2}\sigma^2 - \sigma\lambda\right) dt + \sigma d\tilde{W}_t^X = \eta(\tilde{\mu} - X_t)dt + \sigma d\tilde{W}_t^X, \tag{2.8}$$

where $\tilde{\mu} = \mu - \frac{1}{\eta}\left(\frac{1}{2}\sigma^2 + \sigma\lambda\right)$.

To derive price of a futures contract, we notice that the process described in Eq. (2.8) is an Ornstein-Uhlenbeck process with a long-run mean $\tilde{\mu}$ and a speed of mean reversion η, which has a solution

$$X_T = e^{-\eta\tau}X_t + \tilde{\mu}(1-e^{-\eta\tau}) + \sigma\int_t^T e^{-\eta(T-s)}d\tilde{W}_s^X, \tag{2.9}$$

where $\tau = T-t$. Thus, under the risk-neutral measure, $\mathbb{Q}$, X_T has a conditional normal distribution with mean

$$\mathbb{E}_t^{\mathbb{Q}}[X_T] = e^{-\eta\tau}X_t + \tilde{\mu}(1-e^{-\eta\tau}) \tag{2.10}$$

and variance

$$\mathbb{V}\mathrm{ar}_t^{\mathbb{Q}}[X_T] = \frac{\sigma^2}{2\eta}(1-e^{-2\eta\tau}),\ \ \eta > 0. \tag{2.11}$$

This leads, using the fact that $S_T = \exp(X_T)$ is log-normally distributed, to:

$$\begin{aligned} \mathbb{E}_t^{\mathbb{Q}}[S_T] &= \exp\left(\mathbb{E}_t^{\mathbb{Q}}[\log(S_T)] + \frac{1}{2}\mathbb{V}\mathrm{ar}_t^{\mathbb{Q}}[\log(S_T)]\right) \\ &= \exp\left(e^{-\eta\tau}\log(S_t) + \tilde{\mu}(1-e^{-\eta\tau}) + \frac{\sigma^2}{4\eta}(1-e^{-2\eta\tau})\right). \end{aligned} \tag{2.12}$$

Plugging the result in Eq. (2.12) to Eq. (2.1), the price of the futures contract can be established.

2.2. Mean-Reverting Model with Deterministic Volatility and Jumps

In this subsection we incorporate random jumps in the stochastic component of the electricity spot price process. Specifically, the dynamics of the stochastic component S_t under the real-world measure $\mathbb{P}$ is given by

$$dS_t = a(S_t)S_t dt + \sigma S_t dW_t + S_t dP_t, \tag{2.13}$$

where P_t is a jump process modelled via the compensated compound Poisson process defined as

$$P_t = \sum_{k=1}^{N_t} J_k - \beta\mu_J t. \tag{2.14}$$

Here, N_t denotes a Poisson process with intensity βt, which represents the number of random jumps in a time interval $[0,t]$; J_k represents k^{th} jump size in a small time interval $[t, t+dt]$ and μ_J is an average jump size. The arrival of one jump in the next small time interval $[t, t+dt]$ occurs with probability $p_1 = e^{-\beta t}\beta dt$, while no jump occurs with probability $p_0 = e^{-\beta t}$. Given that the jump intensity is relatively small, the probability of having more than one jump during a short period of time $[t, t+dt]$ is negligible (see Jacquier et al. (2004)), which implies that $p_1 \approx \beta dt$ and $p_0 \approx 1-\beta dt$. Thus, the expected number of jumps over dt can be approximated by

$$\mathbb{E}^{\mathbb{P}}[dN_t] \approx 1\beta dt + 0(1-\beta dt). \tag{2.15}$$

Further we denote J_t to be the size of a jump occurring in the time interval $[t, t+dt]$. J_t's are assumed to be i.i.d. with

$$\log(1+J_t) \sim N\left(\mu_\xi, \sigma_\xi^2\right), \tag{2.16}$$

where μ_ξ and σ_ξ^2 denote the mean and the variance of $\log(1+J_t)$, respectively. This implies that $\mathbb{E}^{\mathbb{P}}[J_t] = \mu_J = \exp(\mu_\xi + \sigma_\xi^2/2) - 1$ and $\mathbb{V}ar^{\mathbb{P}}[J_t] = \sigma_J^2 = \exp(2\mu_\xi + \sigma_\xi^2)(\exp(\sigma_\xi^2) - 1)$. Moreover, the processes J_t, dN_t and dW_t are assumed to be mutually independent. Note that the jump component P_t is modelled as a compensated process in order to keep the expected value of S_t unchanged and to ensure that there is no excess reward for the risk associated with the random jumps.

After applying the Itô formula to $X_t = \log(S_t)$, we obtain the following dynamics under $\mathbb{P}$:

$$dX_t = \left(a^X(X_t) - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t + dP_t^X, \tag{2.17}$$

where $a^X(X_t) = a(e^{X_t})$ and

$$dP_t^X = \log(1+J_t)\, dN_t - \beta\mu_J dt. \tag{2.18}$$

Now, assuming as above, that the drift $a(X_t)$ corresponds to the one of the mean-reverting processes

$$a^X(X_t) = \eta(\mu - X_t), \tag{2.19}$$

and taking Eq. (2.18) into account, we obtain the following dynamics for X_t under $\mathbb{P}$:

$$\begin{aligned} dX_t &= \left(\eta(\mu - X_t) - \frac{1}{2}\sigma^2 - \beta\mu_J\right) dt + \sigma dW_t + \log(1+J_t)\, dN_t \\ &= \eta(\tilde{\mu} - X_t)dt + \sigma dW_t + \log(1+J_t)\, dN_t, \end{aligned} \tag{2.20}$$

where

$$\tilde{\mu} = \mu - \frac{\sigma^2}{2\eta} - \frac{\beta\mu_J}{\eta}. \tag{2.21}$$

By applying the same change of the probability measure as in Section 2.1,

$$d\tilde{W}_t = dW_t + \lambda dt, \tag{2.22}$$

under the risk-neutral measure $\mathbb{Q}$ we can write

$$\begin{aligned} dX_t &= \left(\eta(\mu - X_t) - \sigma\lambda - \frac{1}{2}\sigma^2 - \beta\mu_J\right) dt + \sigma d\tilde{W}_t + \log(1+J_t)\, dN_t \\ &= \eta(\hat{\tilde{\mu}} - X_t)dt + \sigma d\tilde{W}_t + \log(1+J_t)\, dN_t, \end{aligned} \tag{2.23}$$

where

$$\hat{\tilde{\mu}} = \mu - \frac{\sigma\lambda}{\eta} - \frac{\sigma^2}{2\eta} - \frac{\beta\mu_J}{\eta} = \tilde{\mu} - \frac{\sigma\lambda}{\eta}. \tag{2.24}$$

Eq. (2.23) written under $\mathbb{Q}$ is an Ornstein-Uhlenbeck process with jumps with a long-run mean $\tilde{\mu}$ and a speed of mean reversion η, which has a solution

$$X_T = e^{-\eta\tau} X_t + \hat{\mu}(1 - e^{-\eta\tau}) + \sigma \int_t^T e^{-\eta(T-s)} d\tilde{W}_s + \int_t^T e^{-\eta(T-s)} \log(1 + J_t) dN_s, \tag{2.25}$$

where $\tau = T - t$. Since $S_T = \exp\{X_T\}$ we obtain

$$\begin{aligned} S_T &= \exp\left\{e^{-\eta\tau} X_t + \hat{\mu}(1 - e^{-\eta\tau})\right\} \exp\left\{\sigma \int_t^T e^{-\eta(T-s)} d\tilde{W}_s\right\} \\ &\times \exp\left\{\int_t^T e^{-\eta(T-s)} \log(1 + J_t) dN_s\right\}. \end{aligned} \tag{2.26}$$

The expectation of the stochastic component under $\mathbb{Q}$ is then given by

$$\begin{aligned} \mathbb{E}_t^{\mathbb{Q}}[S_T] &= \exp\left\{e^{-\eta\tau} X_t + \hat{\mu}(1 - e^{-\eta\tau})\right\} \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\sigma \int_t^T e^{-\eta(T-s)} d\tilde{W}_s\right\}\right] \\ &\times \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\int_t^T e^{-\eta(T-s)} \log(1 + J_t) dN_s\right\}\right]. \end{aligned} \tag{2.27}$$

Note that the expression under the first expectation in Eq. (2.27) is log-normally distributed, and thus,

$$\begin{aligned} \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\sigma \int_t^T e^{-\eta(T-s)} d\tilde{W}_s\right\}\right] &= \exp\left\{\frac{1}{2}\mathbb{V}\mathrm{ar}_t^{\mathbb{Q}}\left(\sigma \int_t^T e^{-\eta(T-s)} d\tilde{W}_s\right)\right\} \\ &= \exp\left\{\frac{\sigma^2}{4\eta}\left(1 - e^{-2\eta\tau}\right)\right\}. \end{aligned} \tag{2.28}$$

The second expectation in Eq. (2.27) corresponds to

$$\begin{aligned} &\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\int_t^T e^{-\eta(T-s)} \log(1 + J) dN_s\right\}\right] \\ = &\exp\left[\int_t^T \exp\left\{\left(\log(1 + \mu_J) - \frac{1}{2}\sigma_J^2\right) e^{-\eta(T-s)} + \frac{1}{2}\sigma_J^2 e^{-2\eta(T-s)}\right\} \beta ds - \beta\tau\right], \end{aligned} \tag{2.29}$$

as derived in Appendix 7.1. The proof for the general case is derived in Cartea and Figueroa (2005). Finally, plugging (2.28) and (2.29) into (2.27), we obtain the following expression for the expected spot price:

$$\begin{aligned} &\mathbb{E}_t^{\mathbb{Q}}[S_T] = \exp\left(e^{-\eta\tau} \log(S_t) + \hat{\mu}(1 - e^{-\eta\tau}) + \frac{\sigma^2}{4\eta}(1 - e^{-2\eta\tau})\right) \\ \times &\exp\left[\int_t^T \exp\left\{\left(\log(1 + \mu_J) - \frac{1}{2}\sigma_J^2\right) e^{-\eta(T-s)} + \frac{1}{2}\sigma_J^2 e^{-2\eta(T-s)}\right\} \beta ds - \beta\tau\right]. \end{aligned} \tag{2.30}$$

Note that the first multiplier in (2.30) corresponds to the price of the futures contract under the mean-reverting specification without jumps discussed in Section 2.1. Plugging the result of Eq. (2.30) in Eq. (2.1), futures price under mean-reverting model with jumps is derived.

2.3. Mean-Reverting Model with Stochastic Volatility

In contrast to the previous two models with deterministic volatility, we now assume that the volatility is stochastic. Under the physical measure $\mathbb{P}$, the SDEs of the underlying stochastic component S_t of the electricity price, together with the stochastic variance V_t are given by:

$$\begin{aligned} dS_t &= a(S_t, V_t)S_t dt + \sqrt{V_t} S_t dW_t^S \\ dV_t &= b(V_t)dt + c(V_t)dW_t^V, \end{aligned} \tag{2.31}$$

where the two Wiener processes are correlated via $dW_t^S \cdot dW_t^V = \rho dt$. After applying Itô's lemma to $X_t = \log(S_t)$, we obtain the following dynamics under $\mathbb{P}$:

$$\begin{aligned} dX_t &= \left(a^X(X_t, V_t) - \frac{1}{2}V_t\right) dt + \sqrt{V_t} dW_t^X \\ dV_t &= b(V_t)dt + c(V_t)dW_t^V, \end{aligned} \tag{2.32}$$

where $a^X(X_t, V_t) = a(e^{X_t}, V_t)$ and $dW^X = dW^S$.

We assume that the drift $a^X(X_t, V_t)$ of X_t and the drift $b(V_t)$ of V_t correspond to the ones of the mean-reverting processes:

$$a^X(X_t, V_t) = \eta(\mu - X_t) \text{ and } b(V_t) = \kappa(\theta - V_t). \tag{2.33}$$

Further, we set $c(V_t) = \sigma_v \sqrt{V_t}$. Thus, X_t is modelled via an Ornstein-Uhlenbeck process with a long-run mean μ and a speed of mean reversion η. The variance, V_t, follows a mean-reverting process with square-root diffusion as introduced in Cox et al. (1990), where κ, θ and σ_v are positive constants satisfying the Feller's condition $2\kappa\theta/\sigma_v^2 \geq 1$; κ represents the speed of adjustment, θ is the long-run mean and σ_v is the volatility of volatility.

Hence, the model for X_t written under the real-world measure $\mathbb{P}$ becomes

$$\begin{aligned} dX_t &= \left\{\eta(\mu - X_t) - \frac{1}{2}V_t\right\} dt + \sqrt{V_t} dW_t^X \\ dV_t &= \kappa(\theta - V_t)dt + \sigma_v \sqrt{V_t} dW_t^V. \end{aligned} \tag{2.34}$$

With a change of the probability measure from the real-world, $\mathbb{P}$, to the risk-neutral, $\mathbb{Q}$, we obtain

$$
\begin{aligned}
d\tilde{W}_t^X &= dW_t^X + \lambda_t^X(X_t, V_t)dt \\
d\tilde{W}_t^V &= dW_t^V + \lambda_t^V(V_t)dt,
\end{aligned}
\tag{2.35}
$$

where $\lambda_t^X(X_t, V_t)$ and $\lambda_t^V(V_t)$ are the risk premiums for the log-price and the volatility, respectively. The following dynamics for Eq. (2.31) under $\mathbb{Q}$ can be established:

$$
\begin{aligned}
dX_t &= \left(a^X(X_t, V_t) - \frac{1}{2}V_t - \lambda_t^X(X_t, V_t)\sqrt{V_t}\right) X_t dt + \sqrt{V_t} X_t d\tilde{W}_t^X \\
dV_t &= \left(b(V_t) - \lambda_t^V(V_t)c(V_t)\right) dt + c(V_t)d\tilde{W}_t^V.
\end{aligned}
\tag{2.36}
$$

Following the assumptions in Ait-Sahalia (1996) we assume $\lambda_t^X(X_t, V_t) = \lambda^X\sqrt{V_t}$ and $\lambda_t^V(V_t) = \lambda^V\sqrt{V_t}$ as a market price of risk for the underlying and the volatility, respectively.

Thus under the risk-neutral measure $\mathbb{Q}$, Eq. (2.34) can be written as

$$
\begin{aligned}
dX_t &= \left\{\eta(\mu - X_t) - \tilde{\lambda}^X V_t\right\} dt + \sqrt{V_t} d\tilde{W}_t^X \\
dV_t &= \tilde{\kappa}(\tilde{\theta} - V_t)dt + \sigma_v\sqrt{V_t}d\tilde{W}_t^V,
\end{aligned}
\tag{2.37}
$$

where $\tilde{\lambda}^X = (\lambda^X + \frac{1}{2})$ and $\tilde{\kappa} = \kappa + \sigma_v\lambda^V$ ans $\tilde{\theta} = \frac{\kappa\theta}{\kappa+\sigma_v\lambda^V}$.

For evaluation of futures contracts we require knowledge on the distribution of terminal price S_T under the risk-neutral probability measure, $\mathbb{Q}$. Since the characteristic function is given by the Fourier transform of the density function, it can be used to obtain the distribution function. Fourier inversion approach for option pricing was introduced in Heston (1993), and adopted by many authors, including Bates (1996), Bakshi and Madan (2000) and Duffie et al. (2000).

The characteristic function of the log-spot price X_T at time T is given by

$$
\Gamma(t, T, \phi) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{i\phi X_T}\right], \tag{2.38}
$$

where i is the imaginary unit with $i^2 = -1$ and $\phi \in \mathbb{R}$ is a Fourier parameter.

Thus, the expected spot price can be written using the characteristic function as follows:

$$
\Gamma(t, T, -i) = \mathbb{E}_t^{\mathbb{Q}}[S_T] = \mathbb{E}_t^{\mathbb{Q}}\left[e^{X_T}\right] = \Phi(t, X_t, V_t). \tag{2.39}
$$

Setting parameters from Eq. (2.34) into the fundamental partial differential equation (FPDE) derived in Appendix 7.2 (refer to Eq. (7.33)), we obtain the following FPDE:

$$
\frac{\partial\Phi}{\partial t} + \left\{\eta(\mu - X_t) - \tilde{\lambda}^X V_t\right\}\frac{\partial\Phi}{\partial X} + \tilde{\kappa}(\tilde{\theta} - V_t)\frac{\partial\Phi}{\partial V} + \frac{1}{2}V_t\frac{\partial^2\Phi}{\partial X^2}
$$

$$+\frac{1}{2}\sigma_v^2 V_t \frac{\partial^2 \Phi}{\partial V^2} + \rho V_t \sigma_v \frac{\partial^2 \Phi}{\partial X \partial V} = 0. \tag{2.40}$$

The FPDE in Eq. (2.40) can be transformed into a set of ordinary differential equations (ODEs) with an exponential guess for $\Phi(\cdot)$, see Appendix 7.2. For the system of SDEs with two state variables X_t and V_t as in our case, the guess is exponentially affine:

$$\Phi(t, X_t, V_t) = \exp\left\{i\phi A(\tau) X_t + B(\tau) V_t + C(\tau)\right\}, \tag{2.41}$$

where $\tau = T - t$. Computing corresponding derivatives of Eq. (2.41) and replacing them in Eq. (2.40), and then collecting terms that contain X_t, V_t and constants, leads to the following system of ODEs:

$$\begin{aligned}
\frac{dA(\tau)}{d\tau} &= -\eta A(\tau), \\
\frac{dB(\tau)}{d\tau} &= -\tilde{\lambda}^X i\phi A(\tau) + \frac{1}{2}(i\phi)^2 A^2(\tau) - \tilde{\kappa} B(\tau) + \frac{1}{2}\sigma_v^2 B^2(\tau) + \rho\sigma_v i\phi A(\tau) B(\tau), \\
\frac{dC(\tau)}{d\tau} &= \eta\mu i\phi A(\tau) + \tilde{\kappa}\tilde{\theta} B(\tau).
\end{aligned} \tag{2.42}$$

which can be solved subject to boundary conditions $A(0) = 1$, $B(0) = C(0) = 0$. The solution of the first ODE is given by

$$A(\tau) = \exp\{-\eta\tau\}, \tag{2.43}$$

which leads to a system of two ODEs:

$$\begin{aligned}
\frac{dB(\tau)}{d\tau} &= -\tilde{\lambda}^X i\phi \exp\{-\eta\tau\} + \frac{1}{2}(i\phi)^2 \exp\{-2\eta\tau\} - \tilde{\kappa} B(\tau) + \frac{1}{2}\sigma_v^2 B^2(\tau) \\
&+ \rho\sigma_v i\phi \exp\{-\eta\tau\} B(\tau) \\
\frac{dC(\tau)}{d\tau} &= \eta\mu i\phi \exp\{-\eta\tau\} + \tilde{\kappa}\tilde{\theta} B(\tau).
\end{aligned} \tag{2.44}$$

For $B(\tau)$ in Eq. (2.44) we make the standard substitution for Ricatti equations:

$$D(\tau) = \exp\left\{-\int \frac{1}{2}\sigma_v^2 B(\tau) d\tau\right\}, \tag{2.45}$$

which implies

$$B(\tau) = -\frac{D'(\tau)}{\frac{1}{2}\sigma_v^2 D(\tau)}. \tag{2.46}$$

The first equation in Eq. (2.44) becomes

$$\begin{aligned}
\frac{d^2 D(\tau)}{d\tau^2} &- \rho\sigma_v i\phi \exp\{-\eta\tau\} \frac{dD(\tau)}{d\tau} + \tilde{\kappa}\frac{dD(\tau)}{d\tau} + \frac{1}{4}\sigma_v^2 (i\phi)^2 \exp\{-2\eta\tau\} D(\tau) \\
&- \frac{1}{2}\tilde{\lambda}^X \sigma_v^2 i\phi \exp\{-\eta\tau\} D(\tau) = 0.
\end{aligned} \tag{2.47}$$

Following Lutz (2009) who suggest to use work laid out by Heath et al. (1992) and Collin-Dufresne and Goldstein (2002), we make the substitution $\gamma = i\phi \exp\{-\eta\tau\}$, to arrive at the following second order homogenous equation of the general form $(a_2 x + b_2)\frac{d^2y}{dx^2} + (a_1 x + b_1)\frac{dy}{dx} + (a_0 x + b_0)y = 0$:

$$\gamma \frac{d^2 D(\gamma)}{d\gamma^2} + \frac{dD(\gamma)}{d\gamma}\left\{1 + \frac{\rho\sigma_v}{\eta}\gamma - \frac{\tilde{\kappa}}{\eta}\right\} + \left\{-\frac{1}{2}\left(\frac{\sigma_v \tilde{\lambda}^X}{\eta}\right)^2 + \frac{1}{4}\frac{\sigma_v^2}{\eta^2}\gamma\right\} D(\gamma) = 0. \tag{2.48}$$

For the case when $\rho \neq \pm 1$ and $\tilde{\kappa}/\eta \notin \mathbb{Z}$, which appears to be the correct assumption based on our empirical investigation (refer to Section 5), the solution is given by

$$\begin{aligned} D(\gamma) &= \exp\left\{\frac{\sigma_v}{2\eta}(\sqrt{\rho^2 - 1} - \rho)\gamma\right\} \\ &\times \left[C_1 M\left(a, b, -\gamma \frac{\sigma_v \sqrt{\rho^2 - 1}}{\eta}\right) + C_2 U\left(a, b, -\gamma \frac{\sigma_v \sqrt{\rho^2 - 1}}{\eta}\right)\right], \end{aligned} \tag{2.49}$$

see Polyanin and Zaitsev (2003) (p. 225) and Lutz (2009) (p. 71-72). In Eq. (2.49), C_1 and C_2 are integration constants, determined by the boundary condition $B(0) = 0$, M is the Kummer function and U is the Tricomi function, respectively, with the arguments

$$a = \frac{(1 - \frac{\tilde{\kappa}}{\eta})(\sqrt{\rho^2 - 1} - \rho) - \tilde{\lambda}^X \frac{\sigma_v}{\eta}}{2\sqrt{\rho^2 - 1}} \quad \text{and} \quad b = 1 - \frac{\tilde{\kappa}}{\eta}. \tag{2.50}$$

With an inverse transformation, we obtain the following solutions for the functions $B(\tau)$ and $C(\tau)$ in Eq. (2.41):

$$B(\tau) = \frac{c(\tau)}{\sigma_v}\left[\frac{\rho}{\sqrt{\rho^2 - 1}} - 1 + 2a \frac{\frac{C_1}{C_2} b^{-1} M(a+1, b+1, c(\tau)\frac{\sigma_v}{\eta}) - U(a+1, b+1, c(\tau)\frac{\sigma_v}{\eta})}{\frac{C_1}{C_2} M(a, b, c(\tau)\frac{\sigma_v}{\eta}) + U(a, b, c(\tau)\frac{\sigma_v}{\eta})}\right], \tag{2.51}$$

$$\begin{aligned} C(\tau) &= [c(\tau) - c(0)]\left\{\frac{\mu}{\sqrt{\rho^2 - 1}} + \frac{\tilde{\kappa}\theta}{\sigma_v \eta}\left(1 - \frac{\rho}{\sqrt{\rho^2 - 1}}\right)\right\} \\ &- \frac{2\tilde{\kappa}\theta}{\sigma_v^2} \log\left\{\frac{\frac{C_1}{C_2} M(a, b, c(\tau)\frac{\sigma_v}{\eta}) + U(a, b, c(\tau)\frac{\sigma_v}{\eta})}{\frac{C_1}{C_2} M(a, b, c(0)\frac{\sigma_v}{\eta}) + U(a, b, c(0)\frac{\sigma_v}{\eta})}\right\}, \end{aligned} \tag{2.52}$$

where

$$\frac{C_1}{C_2} = \frac{2aU(1+a, 1+b, c(0)\frac{\sigma_v}{\eta}) + \left(1 - \frac{\rho}{\sqrt{\rho^2 - 1}}\right) U(a, b, c(0)\frac{\sigma_v}{\eta})}{\frac{2a}{b} M(1+a, 1+b, c(0)\frac{\sigma_v}{\eta}) - \left(1 - \frac{\rho}{\sqrt{\rho^2 - 1}}\right) M(a, b, c(0)\frac{\sigma_v}{\eta})} \tag{2.53}$$

and

$$c(\tau) = -i\phi \exp\{-\eta\tau\}\sqrt{\rho^2 - 1}. \tag{2.54}$$

Plugging Eq. (2.43), Eq. (2.51) and Eq. (2.52) into Eq. (2.41) leads to the solution for the characteristic function, which according to Eq. (2.39) for $\phi = -i$ determines $\mathbb{E}_t^{\mathbb{Q}}[S_T]$, which is used in Eq. (2.1) for the computation of the futures price.

2.4. Mean-Reverting Model with Stochastic Volatility and Jumps

We assume that under the physical measure $\mathbb{P}$, the SDE of S_t , the stochastic component of the electricity spot price, and the SDE of the variance V_t are given by

$$\begin{aligned} dS_t &= a(S_t, V_t)S_t dt + \sqrt{V_t} S_t dW_t^S + S_t dP_t^S \\ dV_t &= b(V_t)dt + c(V_t)dW_t^V + dP_t^V, \end{aligned} \tag{2.55}$$

where jump process in underlying spot price P_t^S is defined according to Eq. (2.14) as a compensated compound Poisson process. Jump process in the variance is defined as

$$P_t^V = \sum_{k=1}^{N_t} J_k^V. \tag{2.56}$$

Here, N_t is the same arrival Poisson process as in the process S_t. Further, we assume that the jump size J_k^V is exponentially distributed with parameter γ, and the jump size J_t^S of the compound Poisson process P_t^S is distributed normally, conditional on J_t^V, with the following parameters:

$$\log(1 + J_t^S)|J_t^V \sim N\left(\log(1+\mu_J) - \frac{1}{2}\sigma_J^2 + \rho_J J_t^V, \sigma_J^2\right), \tag{2.57}$$

where μ_J and σ_J^2 denote the mean jump size and the variance of the jump size, respectively.

With the change of measure

$$\begin{aligned} d\tilde{W}_t^S &= dW_t^S + \lambda_t^S(S_t, V_t)dt \\ d\tilde{W}_t^V &= dW_t^V + \lambda_t^V(V_t)dt, \end{aligned} \tag{2.58}$$

where $\lambda_t^S(S_t, V_t)$ and $\lambda_t^V(V_t)$ are the risk premiums for the price and the variance processes, respectively, we obtain the following dynamics for Eq. (2.55) under $\mathbb{Q}$:

$$\begin{aligned} dS_t &= \tilde{a}(S_t, V_t)S_t dt + \sqrt{V_t} S_t d\tilde{W}_t^S + S_t d\tilde{P}_t^S \\ dV_t &= \tilde{b}(V_t)dt + c(V_t)d\tilde{W}_t^V + d\tilde{P}_t^V, \end{aligned} \tag{2.59}$$

where

$$\begin{aligned} \tilde{a}(S_t, V_t) &= a(S_t, V_t) - \lambda_t^S(S_t, V_t)\sqrt{V_t} \\ \tilde{b}(V_t) &= b(V_t) - \lambda_t^V(V_t)c(V_t). \end{aligned} \tag{2.60}$$

After applying Itô formula to $X_t = \log(S_t)$, we obtain the following dynamics for X_t:

$$\begin{aligned} dX_t &= \left(\tilde{a}(X_t, V_t)dt - \frac{1}{2}V_t\right) + \sqrt{V_t}d\tilde{W}_t^X + d\tilde{P}_t^X \\ dV_t &= \tilde{b}(V_t)dt + c(V_t)d\tilde{W}_t^V + d\tilde{P}_t^V, \end{aligned} \tag{2.61}$$

where

$$\begin{aligned} d\tilde{W}^X &= d\tilde{W}^S \\ d\tilde{P}^X &= dN_t \left\{\log\left(S_t(1+\hat{J})\right) - \log(S_t)\right\} = dN_t \log\left(1+\hat{J}\right). \end{aligned} \tag{2.62}$$

Similar to the previous model without jumps, now we assume that the drift $\tilde{a}(X_t, V_t)$ of X_t and the drift $\tilde{b}(V_t)$ of V_t correspond to the ones of the mean-reverting processes:

$$\begin{aligned} \tilde{a}(X_t, V_t) &= \eta(\mu - X_t) \\ \tilde{b}(V_t) &= \kappa(\theta - V_t). \end{aligned} \tag{2.63}$$

Hence, the model written under the risk neutral measure $\mathbb{Q}$ becomes:

$$\begin{aligned} dX_t &= \left\{\eta(\mu - X_t) - \frac{1}{2}V_t\right\} dt + \sqrt{V_t}d\tilde{W}_t^X + d\tilde{P}_t^X \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}d\tilde{W}_t^V + d\tilde{P}_t^V. \end{aligned} \tag{2.64}$$

Here, again, the two Wiener processes are correlated via $d\tilde{W}_t^X \cdot d\tilde{W}_t^V = \rho \cdot dt$. Setting parameters from Eq. (2.64) into the FPDE in Eq. (7.33) derived in Appendix 7.2, we obtain the FPDE of the following form:

$$\begin{aligned} &\frac{\partial \Phi}{\partial t} + \left(\eta(\mu - X_t) - \frac{1}{2}V_t - \lambda^\rho \frac{\mu_J\gamma + \rho_J}{\gamma - \rho_J}\right)\frac{\partial \Phi}{\partial X} + \kappa(\theta - V_t)\frac{\partial \Phi}{\partial V} \\ &\quad + \frac{1}{2}V_t\frac{\partial^2 \Phi}{\partial X^2} + \frac{1}{2}\sigma^2 V_t \frac{\partial^2 \Phi}{\partial V^2} + \rho V_t \sigma \frac{\partial^2 \Phi}{\partial X \partial V} \\ &+ \lambda^\rho \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\left(\Phi(X_t + J^X, V_t + J^V) - \Phi(X_t, V_t)\right) | J^V = J_k^V\right]\right] = 0. \end{aligned} \tag{2.65}$$

Again, with an exponential guess from Eq. (2.41), the ODEs for $A(\tau)$ and $B(\tau)$ are identical to those in Eq. (2.42) for the case of without jumps. The ODE for $C(\tau)$ changes slightly due to the adjustment for the jump component. For general case with jumps in the underlying and the variance, it becomes

$$\begin{aligned} \frac{dC(\tau)}{d\tau} &= \left(\eta\mu - \lambda^\rho \frac{\mu_J\gamma + \rho_J}{\gamma - \rho_J}\right) i\phi e^{-\eta\tau} + \kappa\theta B(\tau) \\ &+ \lambda^\rho \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\exp\{i\phi A(\tau)J_t^X + B(\tau)J^V\} | J^V = J_t^V\right]\right] - \lambda^\rho. \end{aligned} \tag{2.66}$$

To solve the expectation in Eq. (2.66), we set

$$A(\tau) = \exp\{-\eta\tau\} \tag{2.67}$$

and use Eq. (2.57), so that the jump part in Eq. (2.66) becomes

$$\lambda^{\rho}\mathbb{E}^{\mathbb{Q}}\left[\exp\left\{i\phi e^{-\eta\tau}\left(\log(1+\mu_J) - \frac{1}{2}\sigma_J^2 + \rho_J J_k^V\right) - \frac{1}{2}\phi^2\sigma_J^2 e^{-2\eta\tau} + B(\tau)J^V\right\}\right] - \lambda^{\rho}. \tag{2.68}$$

Since the jump size J^V of the compound Poisson process P_t^V is exponentially distributed with parameter γ, Eq. (2.68) becomes

$$\frac{\lambda^{\rho}\gamma\exp\left\{i\phi e^{-\eta\tau}\left(\log(1+\mu_J) - \frac{1}{2}\sigma_J^2 + \rho_J J_k^V\right) - \frac{1}{2}\phi^2\sigma_J^2 e^{-2\eta\tau}\right\}}{\gamma - i\phi e^{-\eta\tau}\rho_J - B(\tau)} - \lambda^{\rho}, \tag{2.69}$$

see Duffie et al. (2000). The ODE for $B(\tau)$ is identical to Eq. (2.42) of the stochastic volatility model without jumps, and the ODE for $C(\tau)$ becomes

$$\begin{aligned}\frac{dC(\tau)}{d\tau} &= \left(\eta\mu - \lambda^{\rho}\frac{\mu_J\gamma + \rho_J}{\gamma - \rho_J}\right) i\phi e^{-\eta\tau} + \kappa\theta B(\tau) \\ &+ \frac{\lambda^{\rho}\gamma\exp\left\{i\phi e^{-\eta\tau}\left(\log(1+\mu_J) - \frac{1}{2}\sigma_J^2 + \rho_J J_k^V\right) - \frac{1}{2}\phi^2\sigma_J^2 e^{-2\eta\tau}\right\}}{\gamma - i\phi e^{-\eta\tau}\rho_J - B(\tau)} - \lambda^{\rho}.\end{aligned} \tag{2.70}$$

3. Markov Chain Monte Carlo Estimation

This section provides a brief overview of Marko Chain Monte Carlo (MCMC) approach used within the Bayesian analysis for the estimation of model parameters and latent variables. We describe the mechanics of MCMC estimation and show how to use MCMC methods to compute the quantities of interest.

3.1. General Principles

The underlying problem setup involves estimation of parameters and latent variables such as jump times and jump sizes. In a Bayesian context each of the unobserved latent variables is treated as a parameter to estimate. This leads to a high dimensional posterior distribution which is not a known distribution. In order to compute the moments of the posterior, we would have to compute a high dimensional integral, which is not available in a closed form. Therefore, we rely on the Markov-Chain Monte-Carlo (MCMC) estimation used in Bayesian analysis to infer the distribution of parameters and latent variables conditional on the observed data. This methodology was introduced in Jacqier et al. (1997, 2004) for modeling equity returns and was shown to outperform several competing estimation methods in a simulation study by Andersen et al. (1997).

MCMC generates samples from a given target distribution, in our case $p(\boldsymbol{\Theta}, \boldsymbol{S}|\boldsymbol{X})$ - the joint distribution of the parameter vector $\boldsymbol{\Theta} = (\eta, \mu, \kappa, \sigma_v, \theta, \mu_J, \sigma_J, \beta, \gamma)^\top$ and the state variables $\boldsymbol{S} = \{\boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J}\}$, given the observed data $\boldsymbol{X}$. In many continuous-time models, $p(\boldsymbol{\Theta}, \boldsymbol{S}|\boldsymbol{X})$ is an extremely complicated, high dimensional distribution which is impossible to sample from directly. However, MCMC solves this problem by first breaking the joint distribution into its complete set of conditionals, which are of lower dimension and, thus, are easier to sample from. In this manner the MCMC algorithms attacks the curse of dimensionality that plagues other methods. The theoretical justification for breaking $p(\boldsymbol{\Theta}, \boldsymbol{S}|\boldsymbol{X})$ into its complete conditional distributions is given by the Hammersley-Clifford theorem, see Hammersley and Clifford (1970) and Besag (1974), which states that under mild regularity conditions, the joint posterior is fully characterized by the complete conditional posteriors. In our case, the joint posterior $p(\boldsymbol{\Theta}, \boldsymbol{S}|\boldsymbol{X})$ is characterized by $p(\boldsymbol{\Theta}|\boldsymbol{S}, \boldsymbol{X})$ - the complete conditional posterior of the parameters conditional on the state variables and the data, and $p(\boldsymbol{S}|\boldsymbol{\Theta}, \boldsymbol{X})$ - the complete conditional posterior conditional on the parameter vector and the data. If these distributions cannot be sampled from directly, then the Hammersley-Clifford theorem can be applied again.

MCMC provides a framework for combining the information in these complete conditional distributions to generate samples from the target distribution $p(\boldsymbol{\Theta}, \boldsymbol{S}|\boldsymbol{X})$. Given two initial values $\boldsymbol{\Theta}^{(0)}$ and $\boldsymbol{S}^{(0)}$, MCMC draws $\boldsymbol{S}^{(1)} \sim p(\boldsymbol{S}|\boldsymbol{\Theta}^{(0)}, \boldsymbol{X})$ and then $\boldsymbol{\Theta}^{(1)} \sim p(\boldsymbol{\Theta}|\boldsymbol{S}^{(1)}, \boldsymbol{X})$. Continuing in this fashion, the algorithm generates a sequence of $\{\boldsymbol{S}^{(g)}, \boldsymbol{\Theta}^{(g)}\}_{g=1}^{G}$. This sequence of random variables forms a Markov Chain which for a large number of draws G converges to $p(\boldsymbol{\Theta}, \boldsymbol{S}|\boldsymbol{X})$, the target distribution. Thus, the principle of breaking up the joint distribution into complete conditional distributions is combined with the principle of a Markov-Chain which starts with arbitrary starting values and converges over time to its stationary distribution.

In order to reduce the influence of the starting point in the sampling procedure and to assure that stationarity is achieved, the general approach is to discard a burn-in period of the first h iterations. The iterations after the burn-in period provide a representative sample from the joint posterior, and averaging over the non-discarded iterations provides an estimate for posterior means of parameters and latent variables.

3.2. Sampling from the Conditional Posterior

To update estimated parameters value in each iteration, MCMC algorithm draws from its conditional posterior distribution conditional on the current values of all other parameters and state variables. Sampling from the conditional posterior can be implemented by either using Gibbs sampler introduced by Geman and Geman (1974) or the Metropolis-Hasting algorithm, see Metropolis et al. (1953). Gibbs sampler is applied if the complete conditional distribution that we want to sample from is known. More precisely, we consider the situation where we want to draw a parameter Θ_i from its conditional posterior. In order to obtain the conditional posterior, the Bayes'

Rule is applied. Hereby all terms that do not involve Θ_i can be ignored, since they are absorbed into the constant, which does not have to be calculated explicitly. The conditional parameter posterior $p(\Theta_i|\boldsymbol{\Theta}_{\backslash i}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J}, \boldsymbol{X})$ is proportional to the joint density $p(\boldsymbol{\Theta}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J}, \boldsymbol{X})$, which is, as a result of additional applications of Bayes' Rule, can be decomposed using hierarchical structure in the following way:

$$\begin{aligned} p(\boldsymbol{\Theta}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J}, \boldsymbol{X}) &= p(\boldsymbol{X}|\boldsymbol{\Theta}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J})p(\boldsymbol{\Theta}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J}) \\ &= p(\boldsymbol{X}|\boldsymbol{\Theta}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J})p(\boldsymbol{J}|\boldsymbol{\Theta}, \boldsymbol{Z})p(\boldsymbol{\Theta}, \boldsymbol{Z})p(\boldsymbol{V}|\boldsymbol{\Theta})p(\boldsymbol{\Theta}) \\ &= p(\boldsymbol{X}|\boldsymbol{\Theta}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J})p(\boldsymbol{J}|\boldsymbol{\Theta}, \boldsymbol{Z})p(\boldsymbol{Z}|\boldsymbol{\Theta})p(\boldsymbol{V}|\boldsymbol{\Theta})p(\boldsymbol{\Theta}). \end{aligned} \tag{3.1}$$

Here $p(\boldsymbol{X}|\boldsymbol{\Theta}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J})$ denotes the likelihood function, $p(\boldsymbol{J}|\boldsymbol{\Theta}, \boldsymbol{Z})$, $p(\boldsymbol{Z}|\boldsymbol{\Theta})$ and $p(\boldsymbol{V}|\boldsymbol{\Theta})$ are the distributions of the latent variables (jump sizes and jump times, respectively) and $p(\boldsymbol{\Theta})$ is the prior distribution. Given Markov property of the model we can write:

$$\begin{aligned} p(\boldsymbol{X}, \boldsymbol{V}|\boldsymbol{\Theta}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{J}) &= \prod_{t=1}^{T} p(X_t, V_t|X_{t-1}, V_{t-1}, J_t^X, J_t^V, Z_t, \boldsymbol{\Theta}), \\ p(\boldsymbol{J}|\boldsymbol{\Theta}, \boldsymbol{Z}) &= \prod_{t=1}^{T} p(J_t^X|J_t^V, \boldsymbol{\Theta}, Z_t)p(J_t^V|\boldsymbol{\Theta}, Z_t), \\ p(\boldsymbol{Z}|\boldsymbol{\Theta}) &= \prod_{t=1}^{T} p(Z_t|\boldsymbol{\Theta}). \end{aligned} \tag{3.2}$$

The above procedure can be applied to conditional state posteriors which include jump times and jump sizes. Thus, MCMC algorithm with the Gibbs step samples iteratively drawing from the complete conditional posteriors:

$$\text{Parameters}: \ p(\Theta_i|\boldsymbol{\Theta}_{\backslash i}, \boldsymbol{Z}, \boldsymbol{J}, \boldsymbol{X}, \boldsymbol{V}), \ i = 1, ..., k$$

$$\text{Jump times}: \ p(Z_t|\boldsymbol{\Theta}, \boldsymbol{Z}_{\backslash t}, \boldsymbol{J}, \boldsymbol{X}, \boldsymbol{V}), \ t = 1, ..., T$$

$$\text{Jump sizes}: \ p(J_t^X, J_t^V|\boldsymbol{\Theta}, \boldsymbol{Z}, \boldsymbol{J}_{\backslash t}, \boldsymbol{X}), \ t = 1, ..., T$$

where Θ_i denotes the i-th element of the parameter vector $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}_{\backslash i}$ is the parameter vector without the i-th element. In the case of stochastic volatility, the conditionals $p(V_t|\boldsymbol{\Theta}, \boldsymbol{Z}, \boldsymbol{J}, \boldsymbol{V}_{\backslash t}, \boldsymbol{X})$ for $t = 1, ..., T$ cannot be sampled from directly and thus, the Metropolis-Hastings algorithm is applied. This algorithm samples a candidate draw from a proposal density and then accepts or rejects the candidate draw based on a certain acceptance criterion. To start the procedure running, we have to specify the prior distributions $p(\boldsymbol{\Theta})$ for all parameters of the model. When possible we assume a so-called conjugate priors which after multiplying with the likelihood lead to a posterior distribution belonging to the same family of distributions as the prior itself. Standard conjugate priors allow to draw from the conditional posteriors directly. The choice of conjugate prior distributions is consistent with Johannes and Polson (2006), and will be discussed below, when we derive the likelihood function and posterior distributions.

3.3. *Derivation of the Likelihood function*

In this section we show how to derive the likelihood of the most general case, the stochastic volatility model with correlated jumps (SVCJ) model specified in Section 2.4. Likelihoods for the models without jumps and/or deterministic volatility can be derived analogously.

We recall that the logarithm of the stochastic component of the electricity spot price is denoted by $X_t = \log(S_t)$ and the instantaneous volatility process is denoted by V_t. X_t and V_t are described by the following system of SDEs:

$$dX_t = \left(\eta(\mu - X_t) - \frac{1}{2}V_t - \beta\mu_J\right)dt + \sqrt{V_t}dW_t^X + J_t^X Z_t, \tag{3.3}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma_v\sqrt{V_t}dW_t^V + J_t^V Z_t. \tag{3.4}$$

Here, we assume that $Z_t \sim Ber(\beta dt)$, $J_t^V \sim \mathcal{E}(\gamma)$ and $J_t^X|J_t^V \sim \mathcal{N}(\mu_\xi + \rho_J J_t^V, \sigma_\xi^2)$[1]. The term $-\beta\mu_J$ is a compensator,= that preserves the martingale property of the spot price S_t process. This means that $\mu_J = E\left[e^{J_t^X} - 1\right] = E\left[E\left[e^{J_t^X} - 1|J_t^V\right]\right] = E\left[e^{\mu_\xi + \rho_J J_t^V + \frac{1}{2}\sigma_\xi^2} - 1\right] = e^{\mu_\xi + \frac{1}{2}\sigma_\xi^2}\frac{\gamma}{\gamma - \rho_J} - 1$. Discretizing the SDEs in Eq. (3.3), we obtain

$$X_t = \left(\hat{\mu}\Delta t - \frac{1}{2}V_{t-1}\Delta t + (1 - \eta\Delta t)X_{t-1} + J_t^X Z_t\right) + \sqrt{V_{t-1}\Delta t}\epsilon_t^X, \tag{3.5}$$

$$V_t = \left(\kappa\theta\Delta t + (1 - \kappa\Delta t)V_{t-1} + J_t^V Z_t\right) + \sigma_v\sqrt{V_{t-1}\Delta t}\epsilon_t^V, \tag{3.6}$$

where $\hat{\mu} = \eta\mu - \beta\mu_J = \eta\mu - \beta\left(e^{\mu_\xi + \frac{1}{2}\sigma_\xi^2}\frac{\gamma}{\gamma - \rho_J} - 1\right)$. Then $(X_t, V_t|X_{t-1}, V_{t-1}, \Theta, Z_t, J_t^X, J_t^V)$ follows bivariate Normal distribution with the following parameters:

$$(X_t, V_t|X_{t-1}, V_{t-1}, \Theta, Z_t, J_t^X, J_t^V) \sim \mathcal{N}(\bar{\mu}_t, \Sigma_t), \tag{3.7}$$

where

$$\bar{\mu}_t = \begin{bmatrix}\mu_t^X \\ \mu_t^V\end{bmatrix} = \begin{bmatrix}(\hat{\mu}\Delta t - \frac{1}{2}V_{t-1}\Delta t + (1 - \eta\Delta t)X_{t-1} + J_t^X Z_t) \\ (\kappa\theta\Delta t + (1 - \kappa\Delta t)V_{t-1} + J_t^V Z_t)\end{bmatrix} \tag{3.8}$$

and

$$\Sigma_t = V_{t-1}\Delta t\begin{bmatrix}1 & \rho\sigma_v \\ \rho\sigma_v & \sigma_v^2\end{bmatrix} = V_{t-1}\Delta t\begin{bmatrix}1 & \psi \\ \psi & \Omega + \psi^2\end{bmatrix} \tag{3.9}$$

[1] Ber, $\mathcal{E}$ and $\mathcal{N}$ denote Bernoulli, Exponential and Normal distributions, respectively.

with $\psi = \rho\sigma_v$ and $\Omega = (1-\rho^2)\sigma_v^2$. Thus, we obtain $\det(\Sigma_t) = (V_{t-1}\Delta t)^2\Omega$ and

$$\Sigma_t^{-1} = \frac{1}{V_{t-1}\Delta t\Omega}\begin{bmatrix}\Omega+\psi^2 & -\psi \\ -\psi & 1\end{bmatrix}.$$

The likelihood function for the SVCJ model is given by

$$p(X,V|\Theta,Z,J) = \prod_{t=1}^{T} p(X_t,V_t|X_{t-1},V_{t-1},\Theta,Z_t,J_t^X,J_t^V) \tag{3.10}$$

$$= \prod_{t=1}^{T}\frac{1}{2\pi\det(\Sigma_t)^{1/2}}\exp\left(-\frac{1}{2}(X_t-\mu_t^X,V_t-\mu_t^V)\Sigma_t^{-1}(X_t-\mu_t^X,V_t-\mu_t^V)'\right) \tag{3.11}$$

$$= \prod_{t=1}^{T}\frac{1}{2\pi V_{t-1}\Delta t\Omega^{1/2}}\exp\left(-\frac{1}{2}(X_t-\mu_t^X,V_t-\mu_t^V)\Sigma_t^{-1}(X_t-\mu_t^X,V_t-\mu_t^V)'\right) \tag{3.12}$$

$$= \Omega^{-T/2}\prod_{t=1}^{T}\left(\frac{1}{V_{t-1}\Delta t}\right)\cdot \tag{3.13}$$

$$\exp\left[-\frac{1}{2\Omega}\sum_{t=1}^{T}\left((\Omega+\psi^2)\frac{(X_t-\mu_t^X)^2}{V_{t-1}\Delta t} - 2\psi\frac{(X_t-\mu_t^X)(V_t-\mu_t^V)}{V_{t-1}\Delta t} + \frac{(V_t-\mu_t^V)^2}{V_{t-1}\Delta t}\right)\right].$$

3.4. Posterior distributions

Assuming that the prior distribution for parameter $\hat{\mu}$, $p(\hat{\mu}) \sim \mathcal{N}(\mu_{\hat{\mu}},\sigma_{\hat{\mu}}^2)$ we can find the posterior distribution for $\hat{\mu}$ as $p(\hat{\mu}|X,V,Z,J,\Theta_{/\hat{\mu}}) \propto p(X,V|\Theta,Z,J)p(\hat{\mu})$. Thus, we obtain

$$p(\hat{\mu}|X,V,Z,J,\Theta_{/\hat{\mu}}) \propto$$
$$\exp\left[-\frac{1}{2\Omega}\sum_{t=1}^{T}\left((\Omega+\psi^2)\frac{(X_t-\mu_t^X)^2}{V_{t-1}\Delta t} - 2\psi\frac{(X_t-\mu_t^X)(V_t-\mu_t^V)}{V_{t-1}\Delta t}\right)\right]\exp\left[-\frac{(\hat{\mu}-\mu_{\hat{\mu}})^2}{2\sigma_{\hat{\mu}}^2}\right] =$$
$$\exp\left[-\frac{1}{2\Omega}\sum_{t=1}^{T}\left((\Omega+\psi^2)\frac{(\hat{\mu}\Delta t-A_t^{\hat{\mu}})^2}{V_{t-1}\Delta t} + 2\psi\frac{(\hat{\mu}\Delta t-A_t^{\hat{\mu}})(V_t-\mu_t^V)}{V_{t-1}\Delta t}\right)\right]\exp\left[-\frac{(\hat{\mu}-\mu_{\hat{\mu}})^2}{2\sigma_{\hat{\mu}}^2}\right].$$

Here $A_t^{\hat{\mu}} = (X_t + \frac{1}{2}V_{t-1}\Delta t - (1-\eta\Delta t)X_{t-1} - J_t^X Z_t)$. Then, rearranging terms we can write

$$p(\hat{\mu}|X,V,Z,J,\Theta_{/\hat{\mu}}) \propto$$
$$\exp\left[-\frac{1}{2\Omega}\sum_{t=1}^{T}\left(\frac{(\Omega+\psi^2)}{V_{t-1}\Delta t}(\hat{\mu}^2\Delta t^2 - 2\hat{\mu}\Delta t A_t^{\hat{\mu}}) + 2\psi\frac{(V_t-\mu_t^V)}{V_{t-1}}\hat{\mu}\Delta t\right) - \frac{1}{2}\left(\frac{\hat{\mu}}{\sigma_{\hat{\mu}}^2} - 2\frac{\mu_{\hat{\mu}}}{\sigma_{\hat{\mu}}^2}\right)\right]$$
$$\exp\left[-\frac{1}{2}\left\{\left(\sum_{t=1}^{T}\frac{(\Omega+\psi^2)\Delta t}{\Omega V_{t-1}} + \frac{1}{\sigma_{\hat{\mu}}^2}\right)\hat{\mu}^2 - 2\left(\sum_{t=1}^{T}\frac{(\Omega+\psi^2)A_t^{\hat{\mu}}}{\Omega V_{t-1}} - \sum_{t=1}^{T}\psi\frac{(V_t-\mu_t^V)}{\Omega V_{t-1}} + \frac{\mu_{\hat{\mu}}}{\sigma_{\hat{\mu}}^2}\right)\hat{\mu}\right\}\right].$$

Thus, we observe that $p(\hat{\mu}|X,V,Z,J,\Theta_{/\hat{\mu}}) \sim \mathcal{N}(\mu^*_{\hat{\mu}}, \sigma^{*2}_{\hat{\mu}})$, where

$$\mu^*_{\hat{\mu}} = \frac{\left(\sum_{t=1}^{T} \frac{(\Omega+\psi^2)(X_t+\frac{1}{2}V_{t-1}\Delta t-(1-\eta\Delta t)X_{t-1}-J_t^X Z_t)}{\Omega V_{t-1}} - \sum_{t=1}^{T}\psi\frac{(V_t-\mu_t^V)}{\Omega V_{t-1}} + \frac{\mu_{\hat{\mu}}}{\sigma^2_{\hat{\mu}}}\right)}{\left(\sum_{t=1}^{T}\frac{(\Omega+\psi^2)\Delta t}{\Omega V_{t-1}} + \frac{1}{\sigma^2_{\hat{\mu}}}\right)} \tag{3.14}$$

and

$$\sigma^{*2}_{\hat{\mu}} = \frac{1}{\left(\sum_{t=1}^{T}\frac{(\Omega+\psi^2)\Delta t}{\Omega V_{t-1}} + \frac{1}{\sigma^2_{\hat{\mu}}}\right)}. \tag{3.15}$$

For the parameter $\Omega = (1-\rho^2)\sigma_v^2$ we assume prior distribution $p(\Omega) \sim \mathcal{IG}(\alpha_0^\Omega, \beta_0^\Omega)$, while for ψ we assume $p(\psi|\Omega) \sim \mathcal{N}(\mu_0^\psi, \frac{\Omega}{p_0})$. Then, the joint posterior distribution can be found as follows:

$$\begin{aligned}
&p(\Omega,\psi|X,V,Z,J,\Theta_{/\{\Omega,\psi\}}) = p(X,V|\Theta,Z,J)p(\psi|\Omega)p(\Omega) = \Omega^{-T/2}\prod_{t=1}^{T}\left(\frac{1}{V_{t-1}\Delta t}\right)\cdot \\
&\exp\left[-\frac{1}{2\Omega}\sum_{t=1}^{T}\left((\Omega+\psi^2)\frac{(X_t-\mu_t^X)^2}{V_{t-1}\Delta t} - 2\psi\frac{(X_t-\mu_t^X)(V_t-\mu_t^V)}{V_{t-1}\Delta t} + \frac{(V_t-\mu_t^V)^2}{V_{t-1}\Delta t}\right)\right]\cdot \\
&\sqrt{\frac{p_0}{\Omega}}\exp\left(-\frac{\psi^2-2\psi\mu_0^\psi+\mu_0^{\psi 2}}{2\frac{\Omega}{p_0}}\right)\cdot\frac{\Omega^{-\alpha_0^\Omega-1}}{\Gamma(\alpha_0^\Omega)}\exp\left(-\frac{\beta_0^\Omega}{\Omega}\right) \propto \\
&\frac{1}{\sqrt{\Omega}}\exp\left[-\frac{1}{2}\left\{\psi^2\left(\frac{\sum_{t=1}^{T}(X_t-\mu_t^X)^2/(V_{t-1}\Delta t)+p_0}{\Omega}\right) - 2\psi\left(\frac{\sum_{t=1}^{T}(X_t-\mu_t^X)(V_t-\mu_t^V)/(V_{t-1}\Delta t)+p_0\mu_0^\psi}{\Omega}\right)\right\}\right] \\
&\frac{\Omega^{-(T/2+\alpha_0^\Omega)-1}}{\Gamma(T/2+\alpha_0^\Omega)}\exp\left[-\frac{1}{\Omega}\left(\frac{1}{2}\frac{\sum_{t=1}^{T}(V_t-\mu_t^V)^2}{V_{t-1}\Delta t} + \frac{1}{2}p_0\mu_0^{\psi 2} + \beta_0^\Omega - \frac{1}{2}\frac{(\sum_{t=1}^{T}(X_t-\mu_t^X)(V_t-\mu_t^V)/(V_{t-1}\Delta t))^2}{\sum_{t=1}^{T}(X_t-\mu_t^X)^2/(V_{t-1}\Delta t)+p_0}\right)\right].
\end{aligned}$$

Therefore, the parameters of the posterior distribution are given by

$$\begin{aligned}
\mu_1^\psi &= \frac{\sum_{t=1}^{T}(X_t-\mu_t^X)(V_t-\mu_t^V)/(V_{t-1}\Delta t)+p_0\mu_0^\psi}{\sum_{t=1}^{T}(X_t-\mu_t^X)^2/(V_{t-1}\Delta t)+p_0}, \\
\sigma_1^{\psi 2} &= \frac{\Omega}{\sum_{t=1}^{T}(X_t-\mu_t^X)^2/(V_{t-1}\Delta t)+p_0}, \\
\alpha_1^\Omega &= \alpha_0^\Omega + \frac{T}{2}, \\
\beta_1^\Omega &= \frac{1}{2}\frac{\sum_{t=1}^{T}(V_t-\mu_t^V)^2}{V_{t-1}\Delta t} + \frac{1}{2}p_0\mu_0^{\psi 2} + \beta_0^\Omega - \frac{1}{2}\frac{(\sum_{t=1}^{T}(X_t-\mu_t^X)(V_t-\mu_t^V)/(V_{t-1}\Delta t)+p_0\mu_0^\psi)^2}{\sum_{t=1}^{T}(X_t-\mu_t^X)^2/(V_{t-1}\Delta t)+p_0}.
\end{aligned}$$

Now, we derive the posterior distributions for the stochastic volatility parameters, θ and κ. The prior for θ is assumed to be $p(\theta) \sim \mathcal{N}(\mu_0^\theta, \sigma_0^{\theta 2})$. The posterior distribution can be obtained as follows:

$$p(\theta|X,V,Z,J,\Theta_{/\{\theta\}}) = p(X,V|\Theta,Z,J)p(\theta) \propto \tag{3.16}$$
$$\exp\left[-\frac{1}{2\Omega}\sum_{t=1}^{T}\left(-2\psi\frac{(X_t-\mu_t^X)(V_t-\mu_t^V)}{V_{t-1}\Delta t}+\frac{(V_t-\mu_t^V)^2}{V_{t-1}\Delta t}\right)\right]\cdot\exp\left[-\frac{\theta^2-2\theta\mu_0^\theta}{\sigma_0^{\theta 2}}\right]\propto$$
$$\exp\left[-\frac{1}{2}\left\{\theta^2\left(\sum_{t=1}^{T}\frac{\kappa^2\Delta t}{\Omega V_{t-1}}+\frac{1}{\sigma_0^{\theta 2}}\right)-2\theta\left(-\sum_{t=1}^{T}\frac{\psi(X_t-\mu_t^X)\kappa}{\Omega V_{t-1}}+\sum_{t=1}^{T}\frac{\kappa A_t^\theta}{\Omega V_{t-1}}+\frac{\mu_0^\theta}{\sigma_0^{\theta 2}}\right)\right\}\right],$$

where $A_t^\theta = (V_t-(1-\kappa\Delta t)V_{t-1}-J_t^V Z_t)$. Thus, the posterior distribution $p(\theta|X,V,Z,J,\Theta_{/\{\theta\}}) \sim \mathcal{N}(\mu_1^\theta,\sigma_1^\theta)$, where

$$\mu_1^\theta = \frac{-\sum_{t=1}^{T}\frac{\psi(X_t-\mu_t^X)\kappa}{\Omega V_{t-1}}+\sum_{t=1}^{T}\frac{\kappa(V_t-(1-\kappa\Delta t)V_{t-1}-J_t^V Z_t)}{\Omega V_{t-1}}+\frac{\mu_0^\theta}{\sigma_0^{\theta 2}}}{\sum_{t=1}^{T}\frac{\kappa^2\Delta t}{\Omega V_{t-1}}+\frac{1}{\sigma_0^{\theta 2}}} \tag{3.17}$$

and

$$\sigma_1^\theta = \frac{1}{\sum_{t=1}^{T}\frac{\kappa^2\Delta t}{\Omega V_{t-1}}+\frac{1}{\sigma_0^{\theta 2}}}. \tag{3.18}$$

The prior for κ is assumed to be $p(\kappa) \sim \mathcal{N}(\mu_0^\kappa,\sigma_0^{\kappa 2})$. The posterior distribution can be written as follows:

$$p(\kappa|X,V,Z,J,\Theta_{/\{\kappa\}}) = p(X,V|\Theta,Z,J)p(\kappa) \propto \tag{3.19}$$
$$\exp\left[-\frac{1}{2\Omega}\sum_{t=1}^{T}\left(-2\psi\frac{(X_t-\mu_t^X)(V_t-\mu_t^V)}{V_{t-1}\Delta t}+\frac{(V_t-\mu_t^V)^2}{V_{t-1}\Delta t}\right)\right]\cdot\exp\left[-\frac{\kappa^2-2\kappa\mu_0^\kappa}{\sigma_0^{\kappa 2}}\right]\propto$$
$$\exp\left[-\frac{1}{2}\left\{\kappa^2\left(\sum_{t=1}^{T}\frac{(\theta-V_{t-1})^2\Delta t}{\Omega V_{t-1}}+\frac{1}{\sigma_0^{\kappa 2}}\right)-2\kappa\left(-\sum_{t=1}^{T}\frac{\psi(X_t-\mu_t^X)(\theta-V_{t-1})}{\Omega V_{t-1}}+\sum_{t=1}^{T}\frac{(\theta-V_{t-1})A_t^\kappa}{\Omega V_{t-1}}+\frac{\mu_0^\kappa}{\sigma_0^{\kappa 2}}\right)\right\}\right],$$

where $A_t^\kappa = (V_t-V_{t-1}-J_t^V Z_t)$. Thus, the posterior distribution $p(\kappa|X,V,Z,J,\Theta_{/\{\kappa\}}) \sim \mathcal{N}(\mu_1^\kappa,\sigma_1^\kappa)$, where

$$\mu_1^\kappa = \frac{-\sum_{t=1}^{T}\frac{\psi(X_t-\mu_t^X)(\theta-V_{t-1})}{\Omega V_{t-1}}+\sum_{t=1}^{T}\frac{(\theta-V_{t-1})(V_t-V_{t-1}-J_t^V Z_t)}{\Omega V_{t-1}}+\frac{\mu_0^\kappa}{\sigma_0^{\kappa 2}}}{\sum_{t=1}^{T}\frac{(\theta-V_{t-1})^2\Delta t}{\Omega V_{t-1}}+\frac{1}{\sigma_0^{\kappa 2}}} \tag{3.20}$$

and

$$\sigma_1^\kappa = \frac{1}{\sum_{t=1}^{T}\frac{(\theta-V_{t-1})^2\Delta t}{\Omega V_{t-1}}+\frac{1}{\sigma_0^{\kappa 2}}} \tag{3.21}$$

In order to derive the posterior distribution for the jump size in the return process, J_t^X, and in the volatility process, J_t^V, we consider two cases:

- No jump, i.e. $Z_t = 0$. In this case there is no additional information from the observations and we simulate with the prior distribution

$$p(J_t^X, J_t^V | X_t, V_t, V_{t-1}, Z_t = 0, \Theta) \propto p(J_t^X | J_t^V) p(J_t^V), \tag{3.22}$$

  where $p(J_t^V) \sim \mathcal{E}(\gamma)$ and $p(J_t^X | J_t^V) \sim \mathcal{N}(\mu_\xi + \rho_J J_t^V, \sigma_\xi^2)$.

- Single jump, $Z_t = 1$.

$$\begin{aligned}
&p(J_t^X, J_t^V | X_t, V_t, V_{t-1}, Z_t = 1, \Theta) \propto p(X_t, V_t | J_t^X, J_t^V, V_{t-1}, Z_t = 1, \Theta) p(J_t^X | J_t^V) p(J_t^V) \propto \\
&\exp\left[-\frac{1}{2\Omega}\left((\Omega+\psi^2)\frac{(X_t-\mu_t^X)^2}{V_{t-1}\Delta t} - 2\psi\frac{(X_t-\mu_t^X)(V_t-\mu_t^V)}{V_{t-1}\Delta t} + \frac{(V_t-\mu_t^V)^2}{V_{t-1}\Delta t}\right)\right] \\
&\cdot\exp\left[-\frac{((J_t^X)^2 - 2J_t^X(\mu_\xi+\rho_J J_t^V) + (\mu_\xi+\rho_J J_t^V)^2)}{2\sigma_\xi^2}\right]\cdot\exp(-\gamma J_t^V) \propto \\
&\exp\left[-\frac{1}{2\Omega}\left((\Omega+\psi^2)\frac{(A_t^X-J_t^X)^2}{V_{t-1}\Delta t} - 2\psi\frac{(A_t^X-J_t^X)(A_t^V-J_t^V)}{V_{t-1}\Delta t} + \frac{(A_t^V-J_t^V)^2}{V_{t-1}\Delta t}\right)\right] \\
&\cdot\exp\left[-\frac{((J_t^X)^2 - 2J_t^X(\mu_\xi+\rho_J J_t^V) + (\mu_\xi+\rho_J J_t^V)^2)}{2\sigma_\xi^2}\right]\cdot\exp(-\gamma J_t^V),
\end{aligned}$$

  where $A_t^X = X_t - \hat{\mu}\Delta t + \frac{1}{2}V_{t-1}\Delta t - (1-\eta\Delta t)X_{t-1}$ and $A_t^V = V_t - \kappa\theta\Delta t - (1-\kappa\Delta t)V_{t-1}$. We obtain

$$\begin{aligned}
&p(J_t^X, J_t^V | X_t, V_t, V_{t-1}, Z_t = 1, \Theta) \propto p(X_t, V_t | J_t^X, J_t^V, V_{t-1}, Z_t = 1, \Theta) p(J_t^X | J_t^V) p(J_t^V) \propto \\
&\exp\left[-\frac{1}{2\Omega}\left((\Omega+\psi^2)\frac{-2A_t^X J_t^X + (J_t^X)^2}{V_{t-1}\Delta t} - 2\psi\frac{-A_t^V J_t^X + J_t^X J_t^V}{V_{t-1}\Delta t}\right)\right] \\
&\cdot\exp\left[-\frac{(J_t^X)^2 - 2J_t^X(\mu_\xi+\rho_J J_t^V)}{2\sigma_\xi^2}\right] \\
&\cdot\exp\left[-\frac{1}{2\Omega}\frac{2(\psi A_t^X - A_t^V)J_t^V + (J_t^V)^2}{V_{t-1}\Delta t} - \frac{2\mu_\xi\rho_J J_t^V + (\rho_J J_t^V)^2}{2\sigma_\xi^2} + \gamma J_t^V\right] \propto \\
&\exp\left[-\frac{1}{2}\left\{(J_t^X)^2\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t} + \frac{1}{\sigma_\xi^2}\right)\right.\right. \\
&-2J_t^X\left(\frac{(\Omega+\psi^2)A_t^X - \psi A_t^V}{\Omega V_{t-1}\Delta t} + \frac{\mu_\xi}{\sigma_\xi^2} + J_t^V\left(\frac{\psi}{\Omega V_{t-1}\Delta t} + \frac{\rho_J}{\sigma_\xi^2}\right)\right) + \\
&\left.\left.\frac{1}{2\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t} + \frac{1}{\sigma_\xi^2}\right)}\left(\frac{(\Omega+\psi^2)A_t^X - \psi A_t^V}{\Omega V_{t-1}\Delta t} + \frac{\mu_\xi}{\sigma_\xi^2} + J_t^V\left(\frac{\psi}{\Omega V_{t-1}\Delta t} + \frac{\rho_J}{\sigma_\xi^2}\right)\right)^2\right\}\right]
\end{aligned}$$

$$
\begin{aligned}
&\exp\left[\frac{1}{2\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t}+\frac{1}{\sigma_\xi^2}\right)}\left(\frac{(\Omega+\psi^2)A_t^X-\psi A_t^V}{\Omega V_{t-1}\Delta t}+\frac{\mu_\xi}{\sigma_\xi^2}+J_t^V\left(\frac{\psi}{\Omega V_{t-1}\Delta t}+\frac{\rho_J}{\sigma_\xi^2}\right)\right)^2\right]\\
&\exp\left[-\frac{1}{2}\left\{(J_t^V)^2\left(\frac{1}{\Omega V_{t-1}\Delta t}+\frac{\rho_J^2}{\sigma_\xi^2}\right)-2J_t^V\left(\frac{-\psi A_t^X+A_t^V}{\Omega V_{t-1}\Delta t}-\frac{\mu_\xi\rho_J}{\sigma_\xi^2}-\gamma\right)\right\}\right]\propto\\
&\exp\left[-\frac{1}{2}\left\{(J_t^X)^2\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t}+\frac{1}{\sigma_\xi^2}\right)\right.\right.\\
&-2J_t^X\left(\frac{(\Omega+\psi^2)A_t^X-\psi A_t^V}{\Omega V_{t-1}\Delta t}+\frac{\mu_\xi}{\sigma_\xi^2}+J_t^V\left(\frac{\psi}{\Omega V_{t-1}\Delta t}+\frac{\rho_J}{\sigma_\xi^2}\right)\right)+\\
&\left.\left.\frac{1}{2\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t}+\frac{1}{\sigma_\xi^2}\right)}\left(\frac{(\Omega+\psi^2)A_t^X-\psi A_t^V}{\Omega V_{t-1}\Delta t}+\frac{\mu_\xi}{\sigma_\xi^2}+J_t^V\left(\frac{\psi}{\Omega V_{t-1}\Delta t}+\frac{\rho_J}{\sigma_\xi^2}\right)\right)^2\right\}\right]\\
&\exp\left[\frac{1}{2\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t}+\frac{1}{\sigma_\xi^2}\right)}\left(\frac{(\Omega+\psi^2)A_t^X-\psi A_t^V}{\Omega V_{t-1}\Delta t}+\frac{\mu_\xi}{\sigma_\xi^2}+J_t^V\left(\frac{\psi}{\Omega V_{t-1}\Delta t}+\frac{\rho_J}{\sigma_\xi^2}\right)\right)^2\right]\\
&\exp\left[-\frac{1}{2}\left\{(J_t^V)^2\left(\frac{1}{\Omega V_{t-1}\Delta t}+\frac{\rho_J^2}{\sigma_\xi^2}-\frac{\left(\frac{\psi}{\Omega V_{t-1}\Delta t}+\frac{\rho_J}{\sigma_\xi^2}\right)^2}{\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t}+\frac{1}{\sigma_\xi^2}\right)}\right)-\right.\right.\\
&\left.\left.2J_t^V\left(\frac{-\psi A_t^X+A_t^V}{\Omega V_{t-1}\Delta t}-\frac{\mu_\xi\rho_J}{\sigma_\xi^2}-\gamma+\frac{\left(\frac{(\Omega+\psi^2)A_t^X-\psi A_t^V}{\Omega V_{t-1}\Delta t}+\frac{\mu_\xi}{\sigma_\xi^2}\right)\left(\frac{\psi}{\Omega V_{t-1}\Delta t}+\frac{\rho_J}{\sigma_\xi^2}\right)}{\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t}+\frac{1}{\sigma_\xi^2}\right)}\right)\right\}\right].
\end{aligned}
$$

Thus, $p(J_t^V|X_t,V_t,V_{t-1},Z_t=1,\Theta)$ is the truncated Normal distribution with mean

$$
\mu_1^{J_t^V}=\frac{\frac{-\psi A_t^X+A_t^V}{\Omega V_{t-1}\Delta t}-\frac{\mu_\xi\rho_J}{\sigma_\xi^2}-\gamma+\frac{\left(\frac{(\Omega+\psi^2)A_t^X-\psi A_t^V}{\Omega V_{t-1}\Delta t}+\frac{\mu_\xi}{\sigma_\xi^2}\right)\left(\frac{\psi}{\Omega V_{t-1}\Delta t}+\frac{\rho_J}{\sigma_\xi^2}\right)}{\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t}+\frac{1}{\sigma_\xi^2}\right)}}{\left(\frac{1}{\Omega V_{t-1}\Delta t}+\frac{\rho_J^2}{\sigma_\xi^2}-\frac{\left(\frac{\psi}{\Omega V_{t-1}\Delta t}+\frac{\rho_J}{\sigma_\xi^2}\right)^2}{\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t}+\frac{1}{\sigma_\xi^2}\right)}\right)}
\tag{3.23}
$$

and variance

$$\left(\sigma_1^{J_t^V}\right)^2 = \frac{1}{\left(\frac{1}{\Omega V_{t-1}\Delta t} + \frac{\rho_J^2}{\sigma_\xi^2} - \frac{\left(\frac{\psi}{\Omega V_{t-1}\Delta t} + \frac{\rho_J}{\sigma_\xi^2}\right)^2}{\left(\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t} + \frac{1}{\sigma_\xi^2}\right)}\right)}. \tag{3.24}$$

The posterior $p(J_t^X|J_t^V, X_t, V_t, V_{t-1}, Z_t = 1, \Theta)$ follows Normal distribution with mean

$$\mu_1^{J_t^X} = \frac{\frac{(\Omega+\psi^2)A_t^X - \psi A_t^V}{\Omega V_{t-1}\Delta t} + \frac{\mu_\xi}{\sigma_\xi^2} + J_t^V\left(\frac{\psi}{\Omega V_{t-1}\Delta t} + \frac{\rho_J}{\sigma_\xi^2}\right)}{\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t} + \frac{1}{\sigma_\xi^2}} \tag{3.25}$$

and variance

$$\left(\sigma_1^{J_t^X}\right)^2 = \frac{1}{\frac{(\Omega+\psi^2)}{\Omega V_{t-1}\Delta t} + \frac{1}{\sigma_\xi^2}}. \tag{3.26}$$

We assume that jumps Z_t arrive with intensity $\beta\Delta t$: $Z_t \sim Ber(\beta\Delta t)$. Then, the posterior distribution is Bernoulli, with the parameter β_1 that can be found as follows

$$\begin{aligned} &p(Z_t = 1|X_t, V_t, V_{t-1}, J, \Theta) \prod p(X_t, V_t|Z_t = 1, V_{t-1}, J, \Theta)p(Z_t = 1) = \\ &\exp\left[-\frac{1}{2}\left\{\frac{(\Omega+\psi^2)(A_t^Z - J_t^X)^2}{\Omega V_{t-1}\Delta t} - 2\frac{\psi(A_t^Z - J_t^X)(B_t^Z - J_t^V)}{\Omega V_{t-1}\Delta t} + \frac{(B_t^Z - J_t^V)^2}{\Omega V_{t-1}\Delta t}\right\}\right] \cdot (\beta\Delta t), \end{aligned} \tag{3.27}$$

$$\begin{aligned} &p(Z_t = 0|X_t, V_t, V_{t-1}, J, \Theta) \prod p(X_t, V_t|Z_t = 0, V_{t-1}, J, \Theta)p(Z_t = 0) = \\ &\exp\left[-\frac{1}{2}\left\{\frac{(\Omega+\psi^2)(A_t^Z)^2}{\Omega V_{t-1}\Delta t} - 2\frac{\psi(A_t^Z)(B_t^Z)}{\Omega V_{t-1}\Delta t} + \frac{(B_t^Z)^2}{\Omega V_{t-1}\Delta t}\right\}\right] \cdot (1 - \beta\Delta t), \end{aligned} \tag{3.28}$$

where $A_t^Z = X_t - \hat{\mu}\Delta t + \frac{1}{2}V_{t-1}\Delta t - (1 - \eta\Delta t)X_{t-1}$ and $B_t^Z = V_t - \kappa\theta\Delta t - (1 - \theta\Delta t)V_{t-1}$. Thus,

$$\beta_1 = \frac{p(X_t, V_t|Z_t = 1, V_{t-1}, J, \Theta)p(Z_t = 1)}{p(X_t, V_t|Z_t = 1, V_{t-1}, J, \Theta)p(Z_t = 1) + p(X_t, V_t|Z_t = 0, V_{t-1}, J, \Theta)p(Z_t = 0)}. \tag{3.29}$$

Let $\hat{\beta} = \beta\Delta t \sim Beta(\alpha_0^{\hat{\beta}}, \beta_0^{\hat{\beta}})$. Therefore, the posterior distribution is given by

$$\begin{aligned} &p(\hat{\beta}|Z) \propto (Z|\hat{\beta})p(\hat{\beta}) = \hat{\beta}^{\sum_{t=1}^T Z_t}(1-\hat{\beta})^{T-\sum_{t=1}^T Z_t}\hat{\beta}^{\alpha_0^{\hat{\beta}}}(1-\hat{\beta})^{\beta_0^{\hat{\beta}}} = \\ &\hat{\beta}^{\sum_{t=1}^T Z_t + \alpha_0^{\hat{\beta}}}(1-\hat{\beta})^{T-\sum_{t=1}^T Z_t + \beta_0^{\hat{\beta}}} \sim Beta(\sum_{t=1}^T Z_t + \alpha_0^{\hat{\beta}}, T - \sum_{t=1}^T Z_t + \beta_0^{\hat{\beta}}). \end{aligned} \tag{3.30}$$

Parameters of the posterior distribution of the jump size $J_t^X|J_t^V$ can be found from regressing

J_t^X on J_t^V. We assume that the prior $p(\mu_\xi) \sim \mathcal{N}(\mu_0^{\mu_\xi}, \sigma_0^{\mu_\xi 2})$. If there is no jump, we can simulate directly from the prior, since no information can be drawn from the sample. The posterior is given by

$$p(\mu_\xi | J^X, J^V, Z, \Theta_{/\mu_\xi}) \propto p(J^X | J^V, Z, \Theta) p(\mu_\xi) \propto \exp\left[-\sum_{t=1}^{T} \mathbb{I}_{Z_t=1} \frac{(J_t^X - \mu_\xi - \rho_J J_t^V)^2}{2\sigma_\xi^2} \right] \cdot$$
$$\exp\left[-\frac{(\mu_\xi - \mu_0^{\mu_\xi})^2}{2\left(\sigma_0^{\mu_\xi}\right)^2} \right]$$
$$\exp\left[-\frac{1}{2} \left\{ \mu_\xi^2 \left(\frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1}}{\sigma_\xi^2} + \frac{1}{\left(\sigma_0^{\mu_\xi}\right)^2} \right) - 2\mu_\xi \left(\sum_{t=1}^{T} \mathbb{I}_{Z_t=1} \frac{(J_t^X - \rho_J J_t^V)}{\sigma_\xi^2} + \frac{\mu_0^{\mu_\xi}}{\left(\sigma_0^{\mu_\xi}\right)^2} \right) \right\} \right].$$

Thus the posterior is normal $\mathcal{N}(\mu_1^{\mu_\xi}, \sigma_1^{\mu_\xi 2})$, where

$$\mu_1^{\mu_\xi} = \frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1} \frac{J_t^X - \rho_J J_t^V}{\sigma_\xi^2} + \frac{\mu_0^{\mu_\xi}}{\left(\sigma_0^{\mu_\xi}\right)^2}}{\frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1}}{\sigma_\xi^2} + \frac{1}{\left(\sigma_0^{\mu_\xi}\right)^2}} \tag{3.31}$$

$$\sigma_1^{\mu_\xi 2} = \frac{1}{\frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1}}{\sigma_\xi^2} + \frac{1}{\left(\sigma_0^{\mu_\xi}\right)^2}}. \tag{3.32}$$

We have to impose a natural restriction on $\rho_J \sim Unif[-1, 1]$.

$$p(\rho_J | J_t^X, J_t^V, \Theta_{/\rho_J}) \propto p(J_t^X | J_t^V, \Theta) p(\rho_J) \propto \exp\left[-\sum_{t=1}^{T} \mathbb{I}_{Z_t=1} \frac{(J_t^X - \mu_\xi - \rho_J J_t^V)^2}{2\sigma_\xi^2} \right] \mathbb{I}_{\{-1 \le \rho_J \le 1\}}$$
$$\propto \exp\left[-\frac{1}{2} \left\{ \rho_J^2 \sum_{t=1}^{T} \mathbb{I}_{Z_t=1} \frac{J_t^{V2}}{\sigma_\xi^2} - 2\rho_J \sum_{t=1}^{T} \mathbb{I}_{Z_t=1} \frac{J_t^V (J_t^X - \mu_\xi)}{\sigma_\xi^2} \right\} \right] \mathbb{I}_{\{-1 \le \rho_J \le 1\}}.$$

Thus, the posterior distribution is the truncated Normal with mean $\mu_1^{\rho_J} = \frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1} J_t^V (J_t^X - \mu_\xi)}{\sum_{t=1}^T \mathbb{I}_{Z_t=1} \left(J_t^V\right)^2}$ and variance $\left(\sigma_1^{\rho_J}\right)^2 = \frac{1}{\sum_{t=1}^T \mathbb{I}_{Z_t=1} \frac{J_t^{V2}}{\sigma_\xi^2}}$.

From the prior $p(\sigma_\xi^2) \sim \mathcal{IG}(\alpha_0^{\sigma_\xi^2}, \beta_0^{\sigma_\xi^2})$, the posterior is given by

$$p(\sigma_\xi^2 | J^X, J^V, \Theta_{/\sigma_\xi^2}) \propto p(J^X | J^V, \Theta) p(\sigma_\xi^2) \propto$$

$$\left(\sigma_\xi^2\right)^{-\frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1}}{2}} \exp\left[-\frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1}(J_t^X-\mu_\xi-\rho_J J_t^V)^2}{2\sigma_\xi^2}\right]\cdot\left(\sigma_\xi^2\right)^{-\alpha_0^{\sigma_\xi^2}-1}\exp\left[-\frac{\beta_0^{\sigma_\xi^2}}{\sigma_\xi^2}\right]$$
$$\sim \mathcal{IG}\left(\frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1}}{2}+\alpha_0^{\sigma_\xi^2}, \frac{\sum_{t=1}^T \mathbb{I}_{Z_t=1}(J_t^X-\mu_\xi-\rho_J J_t^V)^2}{2}+\beta_0^{\sigma_\xi^2}\right).$$

Finally, from the prior $p(\gamma)\sim\mathcal{G}(\alpha_0^\gamma,\beta_0^\gamma)$ we obtain the posterior for γ as

$$p(\gamma|J^V,Z)\propto p(J^V|Z,\gamma)p(\gamma)\propto\gamma^{\sum_{t=1}^T \mathbb{I}_{Z_t=1}}\exp\left[-\gamma\sum_{t=1}^T \mathbb{I}_{Z_t=1}J_t^V\right]\cdot\gamma^{\alpha_0^\gamma-1}\exp\left[-\gamma\beta_0^\gamma\right]$$
$$\sim\mathcal{G}\left(\alpha_0^\gamma+\sum_{t=1}^T \mathbb{I}_{Z_t=1},\beta_0^\gamma+\sum_{t=1}^T \mathbb{I}_{Z_t=1}J_t^V\right).$$

All parameters outlined above can be simulated from the corresponding posterior distributions using the Gibbs sampler. In the next section, we outline the procedure which can be used to simulate the stochastic volatility.

3.5. Estimating stochastic volatility

Stochastic volatility values $\mathbf{V}=\{V_1,...,V_T\}$ that are latent variables, have to be estimated from the data similarly to the set of parameters and latent variables discussed in the previous section. However, there are two issues that make it more challenging to proceed with the simulations. Firstly, the dimensionality of the volatility vector $\mathbf{V}$ is very high and corresponds to the length of the vector of observations $\mathbf{X}$. Secondly, there is no closed form posterior distribution, which could be identified as any well-known distribution to simulate from. This implies that we have to use the Metropolis-Hasting algorithm, approximating the simulation from the posterior distribution by drawing samples from another distribution. This has to be done individually for all values V_t in the vector of volatilities $\mathbf{V}$, leading to a significant increase in computational complexity of the estimation procedure.

The posterior distribution of the single stochastic volatility value V_t for $1\le t\le T$ is given by

$$\begin{aligned}
&\pi(V_t,V_{t+1},V_{t-1})=p(V_t|X,V_{t+1},V_{t-1},Z,J^X,J^V,\Theta)\\
&\propto(X_t,V_t|V_{t-1},Z_t,J_t^X,J_t^V,\Theta)\cdot p(Y_{t+1},V_{t+1}|Y_t,V_t,J_{t+1}^X,J_{t+1}^V,\Theta)\\
&\propto\exp\left[-\frac{1}{2\Omega}\left((\Omega+\psi^2)\frac{(X_t-\mu_t^X)^2}{V_{t-1}\Delta t}-2\psi\frac{(X_t-\mu_t^X)(V_t-\mu_t^V)}{V_{t-1}\Delta t}+\frac{(V_t-\mu_t^V)^2}{V_{t-1}\Delta t}\right)\right]\cdot\\
&\frac{1}{V_t\Delta t}\exp\left[-\frac{1}{2\Omega}\left((\Omega+\psi^2)\frac{(X_{t+1}-\mu_{t+1}^X)^2}{V_t\Delta t}-2\psi\frac{(X_{t+1}-\mu_{t+1}^X)(V_{t+1}-\mu_{t+1}^V)}{V_t\Delta t}+\frac{(V_{t+1}-\mu_{t+1}^V)^2}{V_t\Delta t}\right)\right].
\end{aligned}\tag{3.33}$$

For each iteration n, $n=1,...,N$, the values of the stochastic volatility vector are updated from V_t^n to V_t^{n+1} as described in what follows. Suppose that in the vector of volatilities $\{V_1^{n+1},...,V_{t-1}^{n+1},V_t^n,...,V_T^n\}$ the first $t-1$ values are updated after n iterations, while the other values at $t,...,T$ are yet to be updated. We simulate a proposal for an updated value V_t^* from

$\mathcal{N}(V_t^n, \sigma^2_{MH})$. Here, the parameter of volatility, σ^2_{MH} is chosen exogenously in a such a way that the acceptance rate

$$\alpha(V_t^*, V_t^{n-1}) = \min\left(\frac{\pi(V_t^*, V_{t+1}^n, V_{t-1}^{n+1})}{\pi(V_t^n, V_{t+1}^n, V_{t-1}^{n+1})}), 1\right) \tag{3.34}$$

in a single draw during the Metropolis-Hasting step is approximately 0.5. We accept V_t^* as an update by setting $V_t^{n+1} = V_t^*$ with probability $\alpha(V_t^*, V_t^{n-1})$ and reject the proposal by keeping $V_t^{n+1} = V_t^n$ otherwise.

4. Model Testing

This section aims to present various diagnostic tools which allow to quantify the model performance. Model comparison can be assessed by means of model fit to the data and the complexity of the model as a penalty factor. The model fit is measure by a deviance statistic and the complexity is represented by the number of effective parameters. In a non-Bayesian setting the deviance is used as a quantity which estimates the number of degrees of freedom in the underlying model: It refers to the difference in log-likelihoods between the fitted and the saturated model (that is, the one which yields perfect fit of the data). Obviously, increasing the complexity of the model by, e.g., incorporating stochastic volatility jumps will lead to a better fit of the model to the data. Therefore, one should incorporate a penalty term for complexity.

In analogy, Dempster (1997) and Spiegelhalter et al. (2002)[2] have developed the deviance information criterion (DIC) as a Bayesian model choice criterion. DIC solves the problem of comparing complex hierarchial models when the number of parameters is not clearly defined. The DIC value is computed as a sum of two components: a term $\bar{D}$ that measures goodness of fit and a penalty term p_D which accounts for model complexity:

$$\text{DIC} = \bar{D} + p_D. \tag{4.1}$$

The first term can be calculated as follows:

$$\bar{D} = \mathbb{E}_{\Theta|X}\{D(\Theta)\} = \mathbb{E}_{\Theta|X}\{-2\log f(X|\Theta)\} \tag{4.2}$$

where X denotes the logarithm of the stochastic price component and Θ is a vector of parameters. The better the model fits data, the larger is the likelihood, i.e., smaller values of $\bar{D}$ indicate a better model fit. In fact, since $\bar{D}$ already includes a penalty term p_D, it could be better thought of as a measure of "model adequacy" rather than a measure of fit, although these terms can be used interchangeably. The second component measures the complexity of the model due to the effective

[2]An application of this criterion in financial econometrics can be found in Berg et al. (2004).

number of parameters:

$$p_D = \bar{D} - D(\bar{\Theta}) = \mathbb{E}_{\Theta|X}\{D(\Theta)\} - D\{\mathbb{E}_{\Theta|X}(\Theta)\}, \tag{4.3}$$

it can be rewritten as

$$p_D = \mathbb{E}_{\Theta|X}\{-2\log f(X|\Theta)\} + 2\log f(X|\bar{\Theta}). \tag{4.4}$$

Clearly, since p_D is considered to be the posterior mean of the deviance (average of log-likelihood ratios) minus the deviance evaluated at the posterior mean (likelihood evaluated at average), it can be used to quantify the number of free parameters in the model (the number of degrees of freedom). Further, defining $-2\log f(X|\Theta)$ to be the residual information in data X conditional on Θ, and interpreting it as a logarithmic penalty, or uncertainty, see Kullback and Leibler (1951), Bernardo (1979), p_D can be regarded as the expected excess value of the true over the estimated residual information in the return data X conditional on Θ, and thus, can be thought of as the expected reduction in uncertainty. From Eq. (4.3) we obtain:

$$\bar{D} = D(\bar{\Theta}) + p_D, \tag{4.5}$$

and thus, DIC can be rewritten as the estimate of the fit plus twice the number of effective parameters:

$$DIC = D(\bar{\Theta}) + 2p_D. \tag{4.6}$$

In addition to using the DIC to access model fit, the ability of considered models to price electricity futures will be assessed by means of computing the difference between the actual futures price and the futures price computed using the model.

5. Empirical Results

This section discusses data used in the empirical analysis and presents results for the parameter estimation, model fit and pricing of futures contracts.

5.1. Data Description

In order to perform the empirical analysis of the models presented in Section 2 we consider data for spot electricity prices and futures prices from the state of New South Wales (NSW), Australia. The data set covers time period from January 1, 2006 to December 31, 2015, and consists of 3651 daily observations. As it is evident from Figure 1, the time series of spot electricity prices exhibits seasonal fluctuations, spikes and mean-reverting behaviour. In particular, during the hot months (December to February) and the cold months (June to August) the spot electricity prices are more spiky and volatile compared to prices across other months

Table 1 presents the summary statistics for the spot prices, Ξ_t, the logarithmic prices, $\log(\Xi_t)$, the price changes, $\Delta\Xi_t = S_t - S_{t-1}$, and the logarithmic price changes, $\Delta\log(\Xi_t) = \log(\Xi_t) - \log(\Xi_{t-1})$.

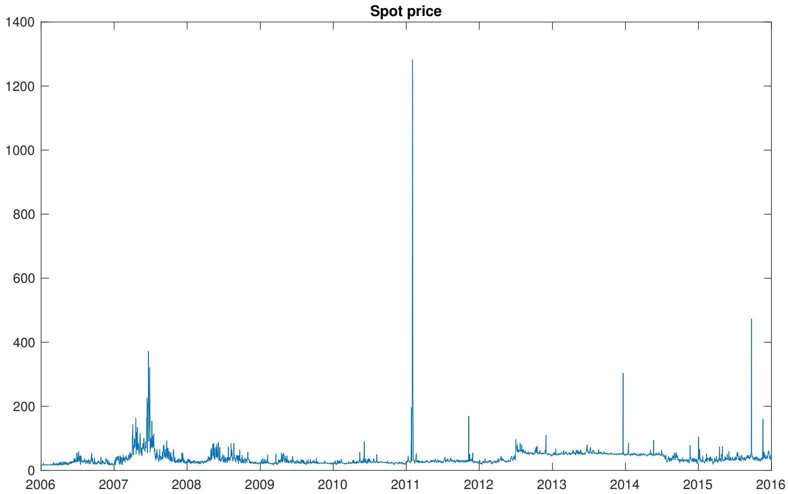

Figure 1: The spot electricity price for the state of New South Wales, Australia for the period from January 1, 2006 to December 31, 2015.

The data display all the characteristics summarised above. In particular, one observes that the range (the difference between the maximal and the minimal price) is large due to spiky behaviour of the spots. The standard deviation of daily logarithmic price changes is 20%, which correspond to annualised volatility of 382%. We also observe that the spot price exhibits a kurtosis of over 680 and a skewness larger than 21, implying a heavy-tailed and right skewed price distribution.

Table 1: Summary statistics for the daily spot price data Ξ_t, the logarithmic prices $\log(\Xi_t)$, the price changes $\Delta\Xi_t$ and the logarithmic price changes, $\Delta\log(\Xi_t))$ for the state of New South Wales, Australia for the time period from January 1, 2006 to December 31, 2015.

Series	Mean	Median	Min	Max	Range	Std Dev	Skew	Kurt
Ξ_t	38.25	31.31	15.20	1281.93	1266.73	36.39	21.74	680.16
$\log(\Xi_t)$	3.53	3.44	2.72	7.16	4.43	0.42	1.24	7.82
$\Delta\Xi_t = \Xi_t - \Xi_{t-1}$	0.01	31.31	-915.72	917.10	1832.83	27.69	0.34	699.35
$\Delta\log(\Xi_t) = \log(\frac{\Xi_t}{\Xi_{t-1}})$	0.00030	-0.0015	-2.099	2.29	4.39	0.20	-0.17	29.78

5.2. *Deterministic Component*

As specified in Section 2 the logarithm of the spot price, $\log(\Xi_t)$, can be decomposed into a sum of two components, namely the logarithm of the deterministic component, $\log(U_t)$, and the logarithm of the stochastic component, $X_t = \log(S_t)$:

$$\log(\Xi_t) = \log(U_t) + X_t. \tag{5.1}$$

The log of the deterministic component, $u_t = \log(U_t)$ combines the long-term linear trend H_1, the long-term sinusoidal (one-year cycle) component H_2, and the short-term seasonal component, H_3

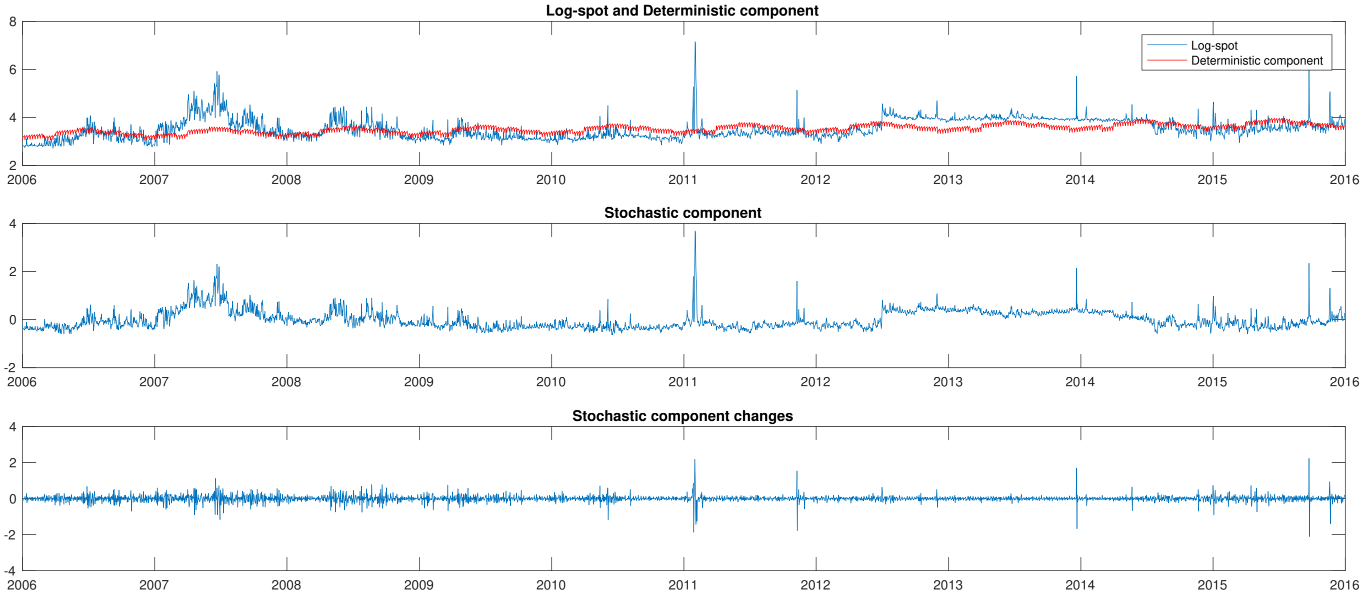

Figure 2: First panel: the logarithmic spot price $\log(\Xi_t)$ together with its deterministic component, u_t; second panel: stochastic component $X_t = \log(\Xi_t) - u_t$; third panel: the differences of the stochastic component $\Delta X_t = X_t - X_{t-1}$.

as follows:

$$u_t = \underbrace{\alpha + \beta \cdot t}_{H_1} + \underbrace{\gamma \cdot \sin\left((t+\tau)\frac{2\pi}{365}\right)}_{H_2} + \underbrace{\phi_d \cdot D_d + \zeta_m \cdot M_m}_{H_3}. \tag{5.2}$$

Here, coefficients α, β, γ, τ, ϕ_d, for $d = \{1, ..., 6\}$, and ζ_m, for $m = \{1, ..., 11\}$, are constant parameters estimated by the means of the non-linear least-squares regression. Dummy variables $D_d = \{0, 1\}$ and $M_m = \{0, 1\}$ are used as indicator variables for the week day and month, respectively. This functional form of the deterministic component combines the ideas proposed in the literature (see, for example, Pilipovic (1997); Lucia and Schwartz (2002a); De Jong (2005); Kosater and Mosler (2006); Ignatieva (2014)).

We estimate parameters of Eq. (5.2) and present the logarithmic spot price $\log(\Xi_t)$ together with its deterministic component in the first panel of Figure 2 for the entire period from January 1, 2006 to December 31, 2015. The stochastic component obtained as a difference between the logarithmic spot price and the deterministic component, $X_t = \log(\Xi_t) - u_t$, is shown in the second panel of Figure 2. Finally, the changes of the stochastic component, $\Delta X_t = X_t - X_{t-1}$, are presented in the bottom panel of Figure 2.

5.3. Estimation Results

Table 2 reports means for parameter estimates obtained for X_t, which is the logarithm of the stochastic component of electricity spot price for different model specifications, which include mean-reverting model (MR); mean-reversing model with jumps in the electricity spot price (MRJ);

mean-reversing model with stochastic volatility but no jumps (MRSV); mean-reversing model with stochastic volatility and jumps in the electricity spot price (MRSVJ); mean-reversing model with stochastic volatility and jumps in the electricity spot price and volatility arriving contemporaneously (MRSVCJ); mean-reversing model with stochastic volatility and jumps in the electricity spot price and volatility arriving independently (MRSVIJ). Table 2 also reports the DIC and the mean squared error (MSE) for spot prices, defined as the average squared difference between the actual spot price and the spot price computed using the model; model ranking is given in parenthesis for both measures.

Table 2: Parameter estimates, the DIC (with model ranking given in parenthesis) and the MSE for the in-sample estimated spot electricity prices relative to the market spot electricity prices for different model specifications. The reported parameter estimates are obtained using spot electricity prices in NSW from 01.01.2006 to 31.12.2015.

	MR	MRJ	MRSV	MRSVJ	MRSVCJ	MRSVIJ
μ	0.1357	0.2687	0.4269	0.6683	0.6260	0.6428
η	0.1202	0.0604	0.0326	0.0266	0.0276	0.0269
σ	0.1797	0.0959	-	-	-	-
λ	-0.1445	-0.0971	-	-	-	-
μ_ξ	-	0.0249	-	0.4044	0.6230	0.2461
σ_ξ	-	0.4307	-	0.1242	0.1102	0.2266
β	-	0.1239	-	0.0021	0.0010	0.0026
κ	-	-	0.0702	0.0712	0.0719	0.0707
θ	-	-	0.0256	0.0275	0.0271	0.0273
σ_v	-	-	0.0467	0.0519	0.0513	0.0512
ρ	-	-	0.7276	0.7978	0.7885	0.7934
ρ_J	-	-	-	-	0.0022	-
γ	-	-	-	-	2.1826	19.1960
β^V	-	-	-	-	-	0.0471
DIC	-2173 (6)	-3738 (5)	-32103 (4)	-32257 (3)	-32522 (1)	-32333 (2)
MSE	0.9988 (6)	0.9601 (5)	0.7623 (4)	0.7524 (3)	0.7515 (1)	0.7521 (2)

Since the results in the table present parameter estimates for the stochastic component of the logarithmic electricity spot price (i.e, after seasonality has been removed) and not the actual log-spot price, it makes some parameters difficult to interpret. However, we notice that the long-run mean μ of the stochastic component is significantly larger for models with stochastic volatility compared to models with deterministic volatility; μ also increases when we add jumps (i.e. μ is larger for MRJ model compared to MR model, and it is also larger for MRSVJ model compared to MRSV model). The speed of mean reversion in the stochastic component of the price, η, on the contrary, decreases when stochastic volatility is incorporated into modelling. Furthermore, for models with stochastic volatility, the long-run mean of the variance θ is approximately 0.027, which corresponds to an annualised long-run volatility $\sqrt{365 \times \theta}$ of 314%, which is consistent with a rough estimate of market volatility obtained using annualised standard deviation of log-changes $\log(\frac{\Xi_t}{\Xi_{t-1}})$ reported in Table 1. The speed of mean reversion of SV κ is consistent across all stochastic volatility models, and corresponds to approximately 0.07.

When comparing model performance in terms of the DIC, we observe that the most complex models MRSVCJ and MRSVIJ that in addition to stochastic volatility contain jumps in the underlying and the volatility, are ranked first and second, respectively. MRJ performs only marginally better compared to the model without jumps (MR model), while stochastic volatility is clearly the most significant factor contributing to the model fit. The last row of Table 2 reports the mean squared error (MSE) for the in-sample estimated spot electricity prices relative to the market spot electricity prices. We observe similar model ranking as the one based on the DIC.

Figure 3 shows the estimated stochastic volatility path $\sqrt{V_t}$ under the mean-reverting stochastic volatility model without jumps (MRSV) in the top panel, and the mean-reverting stochastic volatility model with jumps (MRSVJ) in the bottom panel. We observe that the long-run mean is consistent across both panels, and is in line with the value of θ (of approximately 0.02) reported in Table 2.

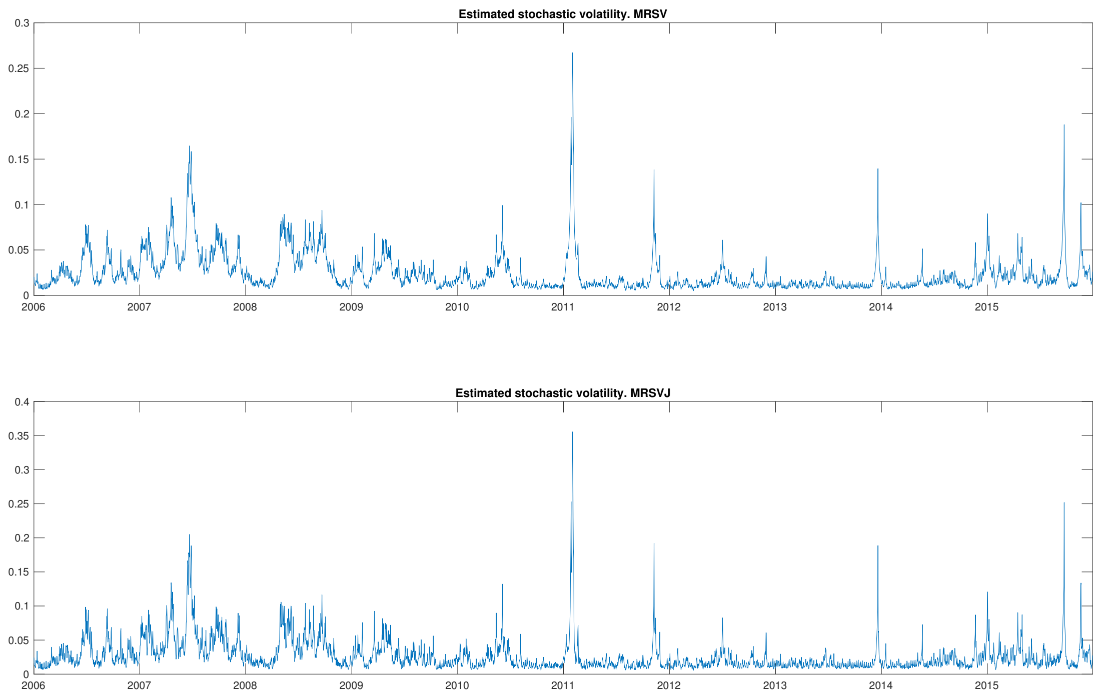

Figure 3: Estimated stochastic volatility paths under the mean-reverting stochastic volatility model without jumps MRSV (top panel) and the mean-reverting stochastic volatility model with jumps MRSVJ (bottom panel).

Figure 4 shows the estimated jump probabilities and jump sizes obtained based on the posterior means under the mean-reverting model with constant volatility MRJ (two panels on the top) and the stochastic volatility MRSVJ (two panels on the bottom).

5.4. *Pricing of futures contracts*

Given the estimated parameters, we can compute prices of futures contracts under various model specifications. In order to access the ability of considered models to price electricity futures,

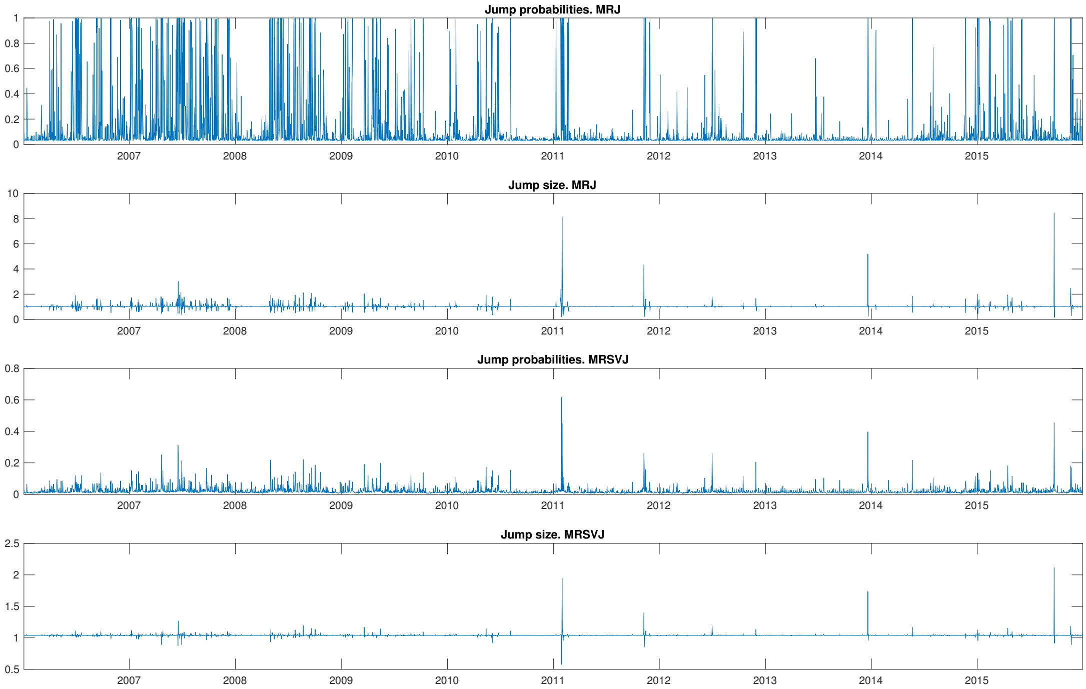

Figure 4: Estimated jump probabilities and jump sizes under the mean-reverting model with constant volatility MRJ (two panels on the top) and with stochastic volatility MRSVJ (two panels on the bottom). The jump probabilities and jump sizes are based on the posterior means.

we will compute the relative pricing error defined as the relative difference (model futures price - futures price)/futures price. We will also report the mean squared error (MSE) for futures prices, which is the average squared difference between the actual futures price and the futures price computed using the model.

5.4.1. Mean - revering model

Figure 5 shows the market price (blue line) and the model price (red line) obtained using the mean-revering model with deterministic volatility and no jumps (MR). We notice that although this model captures largest jumps in the market, if clearly fails to capture small market movements. This results in large pricing errors reported in Table 3 and graphed in Figure 5. Pricing error is defined as the relative difference (model futures price - futures price)/futures price. The distribution of the pricing errors is characterized by the mean, median, standard deviation, 5%-, 25%-, 75%- and 95%-quantiles, skewness and kurtosis. In addition, Figure 6 shows nonparametric distribution of the relative pricing errors. From the table and the figure one observes that the pricing errors exhibit clear seasonalities, where the lowest price difference is typically observed for January (the relative pricing error in January corresponds to 22%) and the highest price difference is typically observed for July (the relative pricing error in January corresponds to 25%). Large negative pricing errors are typically observed in hot summer months December through March when electricity demand is high due to extensive use of air-conditioners or cooling devices. During summer month the model tends to underestimate the market futures price. Note that the estimation bias which results in the underestimation of the market, is likely to be due to the spiky and extreme volatile behaviour of the spot prices used for computation of the futures prices and the fact that market futures prices are considerably smoother than the fitted futures prices. Similarly, the largest positive pricing errors are typically observed in cold winter months July through August, when electricity demand is high due to extensive use of heaters. In these months the model tends to overestimate the market futures price. Again, due to spikes that are present in the data used in the estimation procedure, model prices will experience spikes as well, which will result in the overestimation of futures prices.

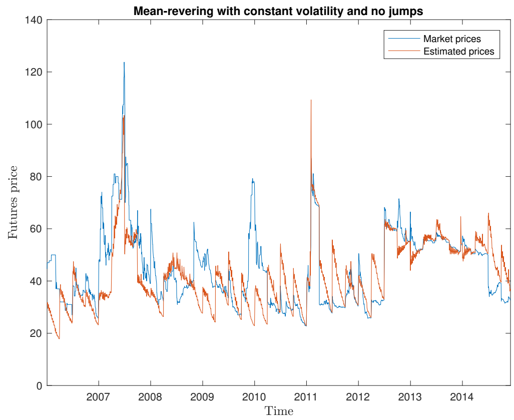

Figure 5: Futures price for the quarterly contracts with nearest maturity expiring on March 31st, June 30th, September 30th, December 31st for NSW. We show market price (blue line) and model price (red line) obtained using the mean-revering model with deterministic volatility and no jumps. The sample covers time period from 03.01.2006 to 05.12.2014.

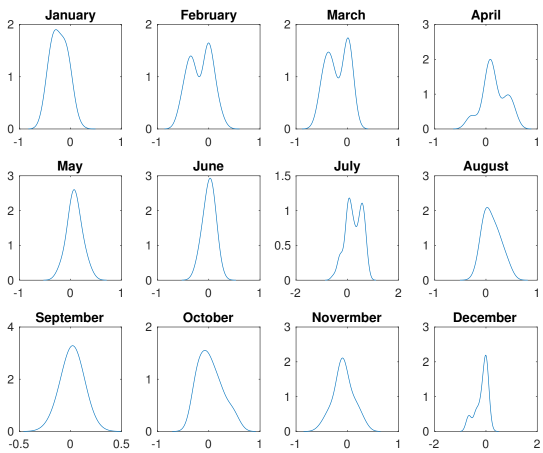

Figure 6: Distribution of in-sample relative futures pricing errors obtained using the mean-revering model with deterministic volatility with no jumps.

Table 3: Summary statistics for in-sample relative futures pricing errors obtained using the mean-revering model with deterministic volatility and no jumps for the state of New South Wales and the time period from January 1, 2006 to December 31, 2015.

Month	Mean	Median	Std. Div	Q5%	Q25%	Q75%	Q95%	Skew	Kurt
January	-0.2212	-0.2387	0.1577	-0.4704	-0.3472	-0.0863	0.0249	0.1276	2.0995
February	-0.1643	-0.0844	0.2191	-0.5137	-0.3530	0.0070	0.1751	-0.0717	1.7913
March	-0.1926	-0.2067	0.2078	-0.5013	-0.3601	0.0101	0.0699	-0.1438	1.4659
April	0.1554	0.1092	0.2286	-0.2969	0.0307	0.3676	0.5181	-0.0621	2.5707
May	0.0824	0.0734	0.1311	-0.1639	0.0171	0.1541	0.3021	0.0850	3.2374
June	0.0061	0.0201	0.0946	-0.1695	-0.0498	0.0699	0.1405	-0.5018	2.7413
July	0.2527	0.2462	0.3196	-0.2991	0.0353	0.5391	0.7013	-0.2671	2.2261
August	0.1120	0.0897	0.1610	-0.1190	-0.0156	0.2209	0.4139	0.4754	2.3384
September	0.0176	0.0185	0.0726	-0.1391	-0.0088	0.0543	0.1308	-0.2447	3.5599
October	0.0304	-0.0213	0.2356	-0.2865	-0.1600	0.1882	0.4744	0.5121	2.4853
November	-0.0753	-0.0882	0.1908	-0.3732	-0.1659	0.0067	0.2683	0.0109	2.9126
December	-0.1581	-0.0387	0.2379	-0.6949	-0.3610	0.0091	0.0915	-1.0686	2.9978

5.4.2. Mean - revering model with jumps

Figure 7 shows market price (blue line) and model price (red line) obtained using the mean-revering model with deterministic volatility and jumps (MRJ). Pricing errors are reported in Table 4 and graphed in Figure 8. Similarly to the MR model without jumps, we observe that the model can only capture large movements in the market, and clearly fails to capture smaller market fluctuations. Similar performance of the MR and MRJ model is consistent with the results for the DIC in Table 2, which shows only marginal improvement when using the MRJ model, compared to the MR model.

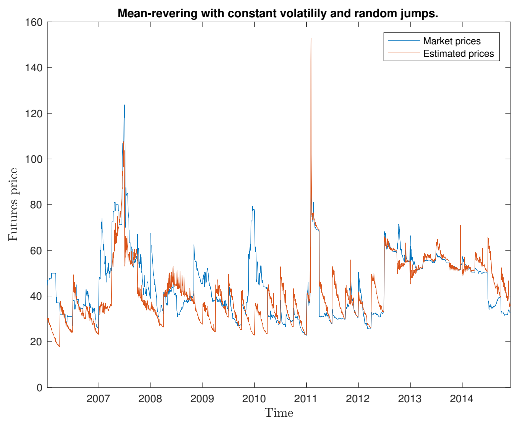

Figure 7: Futures price for the quarterly contracts with nearest maturity expiring on March 31st, June 30th, September 30th, December 31st for NSW. We show market price (blue line) and model price (red line) obtained using the mean-revering model with deterministic volatility and stochastic jumps. The sample covers time period from 03.01.2006 to 05.12.2014.

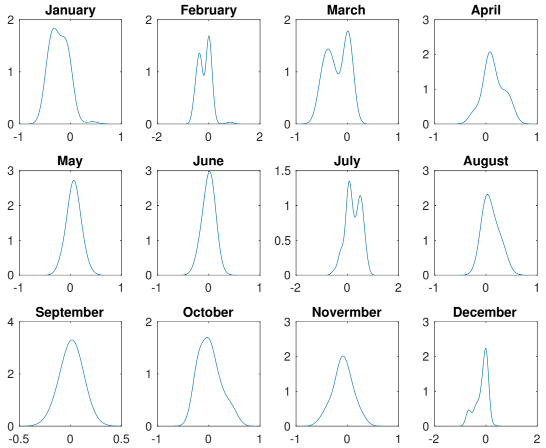

Figure 8: Distribution of in-sample relative futures pricing errors obtained using the mean-revering model with deterministic volatility and stochastic jumps.

Table 4: Summary statistics for in-sample relative futures pricing errors obtained using the mean-revering model with deterministic volatility and stochastic jumps for the state of New South Wales and the time period from January 1, 2006 to December 31, 2015.

Month	Mean	Median	Std. Div	Q5%	Q25%	Q75%	Q95%	Skew	Kurt
January	-0.2241	-0.2502	0.1716	-0.4670	-0.3669	-0.0858	0.0124	0.5885	3.6947
February	-0.1684	-0.0928	0.2426	-0.5395	-0.3601	0.0125	0.1349	0.5375	4.4589
March	-0.1973	-0.2082	0.2091	-0.5088	-0.3643	0.0104	0.0533	-0.1748	1.4480
April	0.1438	0.1055	0.1988	-0.2269	0.0254	0.3012	0.4725	0.1201	2.5639
May	0.0663	0.0627	0.1149	-0.1528	0.0135	0.1370	0.2663	0.0336	3.1385
June	-0.0022	0.0144	0.0934	-0.1852	-0.0633	0.0516	0.1287	-0.4969	2.9794
July	0.2361	0.1943	0.2925	-0.2682	0.0430	0.4855	0.6726	-0.2144	2.3322
August	0.0969	0.0629	0.1471	-0.1127	-0.0095	0.2125	0.3665	0.5587	2.5284
September	0.0098	0.0161	0.0708	-0.1430	-0.0142	0.0433	0.1268	-0.3330	3.5392
October	0.0184	-0.0111	0.2150	-0.2710	-0.1636	0.1518	0.4388	0.5769	2.6750
November	-0.0875	-0.0775	0.1886	-0.3889	-0.1860	0.0148	0.2430	-0.1247	2.9489
December	-0.1625	-0.0248	0.2389	-0.6991	-0.3643	0.0070	0.0636	-1.0590	3.0165

5.4.3. Mean - revering model with stochastic volatility

Figure 9 shows market price (blue line) and model price (red line) obtained using the mean-revering model with stochastic volatility and no jumps (MRSV). Pricing errors are graphed in Figure 10. Similarly to the MR and MRJ model, we observe that the model can only capture large movements in the market, and clearly fails to capture smaller movements.

5.4.4. Model comparison by means of the MSE

In order to draw an overall comparison across models, we report the mean squared error (MSE) for the MR, MRJ and MRSV models in Table 5. The MSE is defined as an average squared difference between the actual futures price and the futures price computed using the model. From the table we observe that the model with jumps (MRJ) performs marginally better compared to the other

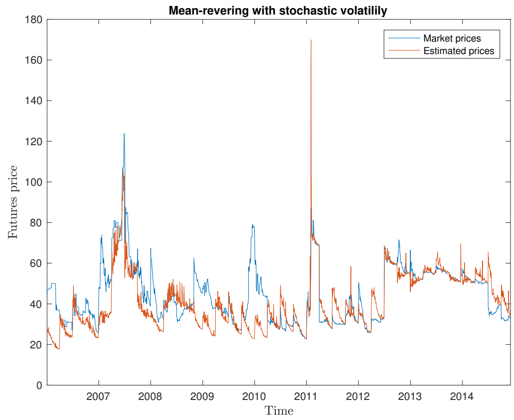

Figure 9: Futures price for the quarterly contracts with nearest maturity expiring on March 31st, June 30th, September 30th, December 31st for NSW. We show market price (blue line) and model price (red line) obtained using the mean-revering model with stochastic volatility and no jumps. The sample covers time period from 03.01.2006 to 05.12.2014.

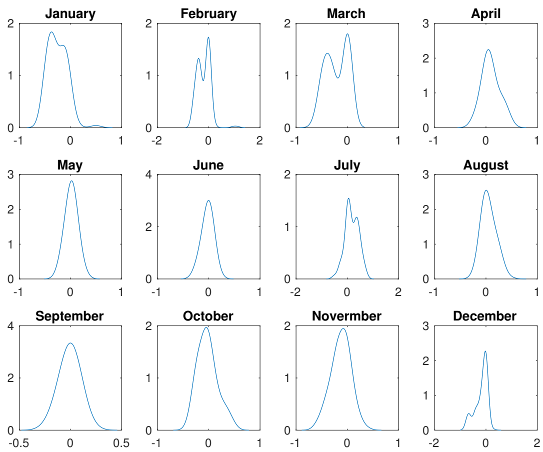

Figure 10: Distribution of in-sample relative futures pricing errors obtained using the mean-revering model with stochastic volatility and no jumps.

two models.

Table 5: Mean Squared Errors (MSE) of the in-sample estimated futures electricity prices and the market futures electricity prices in NSW from 01.01.2006 to 31.12.2015, for different model specifications.

	MR	MRJ	MRSV
MSE	132.9515	130.7440	132.6623

6. Conclusion

This paper provides a comprehensive analysis of continuous-time stochastic volatility jump-diffusion models in context of pricing of futures contracts written on electricity spots. Various parsimonious models and more complex models which include stochastic volatility and jumps in the underlying electricity spot price and volatility are considered. The selected models are specified in such a way that they capture the most prominent characteristics and stylised facts of the electricity spot market including mean reversion, seasonality, extreme volatility and spikes. Estimation of model parameters and latent variables is performed by means of the Markov Chain Monte Carlo (MCMC) technique for the Australian electricity market. Based on the results for the deviance information criterion (DIC) and the mean squared errors (MSE), we conclude that stochastic models with jumps in both, the underlying and its volatility (MRSVCJ and MRSVIJ models) perform best, followed by stochastic volatility models without jumps (MRSV model), while models with deterministic volatility are the worse performing, even if jumps in the underlying electricity spot price are included (MRJ and MR). Finally, as a pricing application, we compute futures prices in a closed or semi-closed form and demonstrate that the model fits data well in-sample and out-of-sample.

References

Ait-Sahalia, Y., 1996. Testing Continuous-Time Models of the Spot Interest Rate. Review of Financial Studies 9, 385–426.

Andersen, T., Chung, H.-J., Sorensen, B. E., 1997. Efficient method of moments estimation of a stochastic volatility model: A monte carlo study. Journal of Econometrics 91, 61Ű87.

Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52, 2003–2049.

Bakshi, G., Madan, D., 2000. Spannig and Derivative Security Valuation. Journal of Finan. Econ. 55, 205–238.

Bates, D., 1996. Jumps and stochastic volatility: Exchange rate process implicit in deutschemark options. Review of Financial Studies 9, 69–108.

Berg, A., Meyer, R., Yu, J., 2004. Deviance information criterion for comparing stochastic volatility models. Journal of Business and Economic Statistics 22, 107–119.

Bernardo, J., 1979. Expected information as expected utility. Annals of Statistics, 686–690.

Besag, J., 1974. Spatial Interaction and the Statistical Analysis of Lattice Systems. Journal of the Royal Statistical Association Series B 36, 192–236.

Bierbrauer, M., Menn, C., Rachev, S., Trück, S., 2007. Spot and Derivative Pricing in the EEX Power Market. Journal of Banking & Finance 31, 3462–3485.

Cartea, A., Figueroa, M., 2005. Pricing in Electricity Markets: a Mean Reverting Jump Diffusion Model with Seasonality. Applied Mathematical Finance 12(4), 313–335.

Chernov, M., Gallant, A., Ghysels, E., 2003. Alternative models for stock price dynamics. Journal of Econometrics 116, 225–257.

Collin-Dufresne, P., Goldstein, R., 2002. Do Bonds Spn The Fixed Income Markets? Theory and Evidence For Unspanned Stochastic Volatility. Journal of Finance 57, 1685–1730.

Cox, J., Ingersoll, J., Ross, S., 1990. A Theory of the Term Structure of Interest Rates. Econometrica 53, 385Ű407.

De Jong, C., 2005. The nature of Power Spikes: A Regime-Switch Approach. ERIM Report Series ERS-2005-052-F&A.

Dempster, A., 1997. The direct use of likelihood for significance testing. Statistics and Computing 7(4), 247–252.

Duffie, D., Pan, J., Singelton, K., 2000. Transform Analysis and Asset Pricing for Affine Jump-Diffusions. Econometrica 68, 1343–1376.

Etheridge, A., 2002. A Course in Financial Calculus. Cambridge University Press, first edition.

Geman, H., Roncoroni, A., 2006. Understanding the fine structure of electricity prices. Journal of Business 79(3), 1225–1262.

Geman, S., Geman, D., 1974. Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. IEEE Trans. on Pattern Analysis and Machine Intelligence 6, 721–741.

Haldrup, N., Nielsen, M., 2006. A regime switching long memory model for electricity prices. Journal of Econometrics 135 (1-2), 349–376.

Hammersley, J., Clifford, P., 1970. Markov Fields on Finite Graphs and Lattices. Unpublished Manuscipt.

Heath, D., Jarrow, R., Morton, A., 1992. Bond Pricing And The Term Structure Of Interest Rates: A New Methodology. Econometrica 60, 70–105.

Heston, S., 1993. A Closed-Form Solution for Options with Stochstic Volatility with Applications to Bond and Curreny Options. Review of Financial Studies 6, 327–343.

Huisman, R., 2008. The Influence of Temperature on Spike Probability in Day-Ahead Power Prices. Energy Economics 30, 2697–2704.

Huisman, R., Mahieu, R., 2003. Regime Jumps in Electricity Prices. Energy Economics 25, 425–434.

Ignatieva, K., 2014. A nonparametric model for spot price dynamics and pricing of futures contracts in electricity markets. Studies in Nonlinear Dynamics and Econometrics 18, 483–505.

Jacqier, E., Polson, N., Rossi, P., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52, 2003Ű2049.

Jacqier, E., Polson, N., Rossi, P., 2004. Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors. Journal of Econometrics 122, 185Ű212.

Jacquier, E., Polson, N., Rossi, P., 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. Journal of Econometrics 122 (Jacquier), 185–212.

Janczura, J., Weron, R., 2010. An empirical comparison of alternate regime-switching models for electricity spot prices. Energy Economics 32(5), 1059–1073.

Johannes, M., Polson, N., 2006. MCMC Methods for Financial Econometrics. Y. Ait-Sahalia and L. Hansen (eds), Elsevier.

Kaminski, V., 1999. Managing Energy Price Risk. Kaminski, V.

Kanamura, T., Ohashi, K., 2008. On transition probabilities of regime switching in electricity prices. Energy Economics 30, 1158–1172.

Kluge, T., Hambly, B., Howison, S., 2009. Modeling Spikes and Pricing Swing Options in Electricity Markets. Quantitative Finance 9(8), 937–949.

Kosater, P., Mosler, K., 2006. Can markov regime-switching models improve power-price forecasts? evidence from german daily power prices. Applied Energy.

Kullback, S., Leibler, R., 1951. On information and sufficiency. Annals of Mathematical Statistics 22, 79–86.

Lucia, J., Schwartz, E., 2002a. Electricity Prices and Power Derivatives: Evidence From the Nordic Power Exchange. Review of Derivatives Research 5, 5–50.

Lucia, J., Schwartz, E. S., 2002b. Electricity prices and power derivatives: Evidence from the nordic power exchange. Review of Derivatives Rewearch 5, 5–50.

Lutz, B., 2009. Pricing of Derivatives on Mean-Reverting Assets. Springer, Berlin.

Merton, R., 2001. Continuous-time finance. Blackwell, first revised edition.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E., 1953. Equations of State Calculations by Fast Computing Machines. Journal of Chemical Physics 21, 1087–1091.

Mount, T., Ning, Y., Cai, X., 2006. Predicting Price Spikes in Electricity Markets Using a Regime-Switching Model with Time-Varying Parameters. Energy Economics 28, 62–80.

Oskendal, B., 2002. Stochastic Differential Equations. 5th edn Springer, Berlin.

Pan, J., 2002. The jump-risk premia implicit in options: Evidence from an integrated time-series study. Journal of Financial Econometrics 63, 3–50.

Pilipovic, D., 1997. Energy Risk: Valuing and Managing Energy Derivatives. McGraw-Hill.

Polyanin, A., Zaitsev, V., 2003. Handbook of Exact Solutions For Ordinary Differential Equations. 2nd Edn. Chapman & Hall/CRC Press, Boca Raton.

Schwartz, E. S., 1997. The stochastic behavior of commodity prices: Implications for valuation and hedging. The Journal of Finance 52, 923–973.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., van der Linde, A., 2002. Bayesian measures of model complexity and fitting. Journal of the Royal Statistical Society B 64 (4), 583–639.

Weron, R., 2006. Modeling and Forecasting Loads and Prices in Deregulated Electricity Markets. ARE Agencja Rynku Energii S.A.

Weron, R., Simonsen, I., Wilman, P., 2004. Modeling highly volatile and seasonal markets: Evidence from the nord pool electricity market. The Application of Econophysics, Proceedings of the Second Nikkei Econophysics Symposium, Yokyp, 182–191.

7. Appendix

7.1. Proof for the Expected Value in Futures Equation

We aim to prove Eq. (2.29), that is, to show that

$$\begin{aligned} & \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\int_t^T e^{-\eta(T-s)}\ln(1+J)dN_s\right\}\right] \\ = \ & \exp\left[\int_t^T \exp\left\{\left(\ln(1+\mu_J)-\frac{1}{2}\sigma_J^2\right)e^{-\eta(T-s)}+\frac{1}{2}\sigma_J^2 e^{-2\eta(T-s)}\right\}\beta ds-\beta\tau\right], \end{aligned} \tag{7.1}$$

where $\tau = T - t$. We use the notation from Cartea and Figueroa (2005) and evaluate

$$\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\int_t^T e^{-\eta(T-s)}\ln(1+J_s)dN_s\right\}\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\int_t^T \alpha_s dN_s\right\}\right] \tag{7.2}$$

where

$$\alpha_s = e^{-\eta(T-s)}\ln(1+J_s). \tag{7.3}$$

We first evaluate (7.2) on the interval $[0,t]$, and then extend calculation to the interval $[t,T]$. We start with defining

$$L_t = \exp\left\{\int_0^t \alpha_s dNs\right\} = e^{m_t} \tag{7.4}$$

where

$$m_t = \int_0^t \alpha_s dNs, \tag{7.5}$$

or, equivalently,

$$dm_t = \alpha_t dN_t. \tag{7.6}$$

In (7.5) and (7.6), m_t denotes a time of a jump which has size α_t such that it holds:

$$m_t = m_{t-} + \alpha_t = \int_0^{t-} \alpha_s dNs + \alpha_t, \tag{7.7}$$

where $t-$ indicates the time just before the jump has occurred.

Following Etheridge (2002) and Cartea and Figueroa (2005), we use generalized Itô lemma to derive the SDE for L_t:

$$dL_t = \frac{\partial L_t(m_{t-})}{\partial m_t}dm_t - \frac{\partial L_t(m_{t-})}{\partial m_t}(m_t - m_{t-})dN_t + (L_t - L_{t-})dN_t, \tag{7.8}$$

where the second derivative is not included since the process in (7.6) is a pure jump process. Using (7.7), we can rewrite (7.4) as follows:

$$L_t = e^{m_{t-}+\alpha_t} = L_{t-}e^{\alpha_t}. \tag{7.9}$$

Thus,

$$\frac{\partial L_t(m_{t-})}{\partial m_t} = L_{t-} \tag{7.10}$$

and plugging (7.6), (7.7), (7.9) and (7.10) into (7.8), we obtain

$$dL_t = L_{t-}(e^{\alpha_t} - 1)dN_t. \tag{7.11}$$

Integrating (7.12), we obtain

$$L_t = 1 + \int_0^t L_s(e^{\alpha_s} - 1)dN_s \text{ with } L_0 = 1. \tag{7.12}$$

By taking $\mathbb{Q}$-expectation of (7.12), we obtain

$$\mathbb{E}_0^{\mathbb{Q}}[L_t] = 1 + \int_0^t \mathbb{E}_0^{\mathbb{Q}}[L_s] \left(\mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_s}] - 1\right) \beta ds, \tag{7.13}$$

where we used $\mathbb{E}_0^{\mathbb{Q}}[dN_t] = \beta dt$ with β denoting the intensity of the Poisson process. Defining $n_t = \mathbb{E}_0^{\mathbb{Q}}[L_t]$, we can rewrite (7.13) as follows:

$$n_t = 1 + \int_0^t n_s \left(\mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_s}] - 1\right) \beta ds, \tag{7.14}$$

or, equivalently,

$$\frac{dn_t}{dt} = n_t \left(\mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_t}] - 1\right) \beta. \tag{7.15}$$

Integrating (7.15), leads to

$$\int_0^t \frac{dn_t}{dt} = \int_0^t \left(\mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_s}] - 1\right) \beta ds, \tag{7.16}$$

that is,

$$\ln(n_t) = \int_0^t \left(\mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_s}] - 1\right) \beta ds. \tag{7.17}$$

Thus, we obtain for n_t:

$$n_t = \exp\left\{\int_0^t \left(\mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_s}] - 1\right) \beta ds\right\}. \tag{7.18}$$

Since $n_t = \mathbb{E}_0^{\mathbb{Q}}[L_t]$, and L_t is given by (7.4), we obtain:

$$\mathbb{E}_0^{\mathbb{Q}}\left[\exp\left\{\int_0^t \alpha_s dNs\right\}\right] = \exp\left\{\int_0^t \left(\mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_s}] - 1\right) \beta ds\right\} \tag{7.19}$$

In order to evaluate (7.19), we start with calculating the expected value:

$$\begin{aligned} \mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_s}] &= \mathbb{E}_0^{\mathbb{Q}}[\exp\{e^{\eta(s-t)}\ln(1+J_s)\}] \\ &= \exp\left\{\mathbb{E}_0^{\mathbb{Q}}\left[e^{\eta(s-t)}\ln(1+J_s)\right] + \frac{1}{2}\mathbb{V}ar_0^{\mathbb{Q}}\left[e^{\eta(s-t)}\ln(1+J_s)\right]\right\} \end{aligned}$$

$$= \exp\left\{e^{\eta(s-t)}\left(\ln(1+\mu_J)-\frac{1}{2}\sigma_J^2\right)+\frac{1}{2}e^{2\eta(s-t)}\sigma_J^2\right\} \tag{7.20}$$

Here, we have assumed that jumps are log-normal with parameters

$$\ln(1+J) \sim N\left(\ln(1+\mu_J)-\frac{1}{2}\sigma_J^2, \sigma_J^2\right), \tag{7.21}$$

where μ_J is the expected jump size and σ_J is the volatility of a jump size. Thus, (7.19) becomes

$$\begin{aligned} & \mathbb{E}_0^{\mathbb{Q}}\left[\exp\left\{\int_0^t \alpha_s dNs\right\}\right] \\ = \; & \exp\left[\int_0^t \exp\left\{\left(\ln(1+\mu_J)-\frac{1}{2}\sigma_J^2\right)e^{\eta(s-t)}+\frac{1}{2}\sigma_J^2 e^{2\eta(s-t)}\right\}\beta ds - \beta t\right]. \end{aligned} \tag{7.22}$$

It is easy to show that on the interval $[t, T]$ (7.19) becomes

$$\mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\int_t^T \alpha_s dNs\right\}\right] = \exp\left\{\int_t^T \left(\mathbb{E}_0^{\mathbb{Q}}[e^{\alpha_s}]-1\right)\beta ds\right\}. \tag{7.23}$$

Thus, we obtain the general formula for (7.24) on the interval $[t, T]$

$$\begin{aligned} & \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\int_t^T \alpha_s dNs\right\}\right] \\ = \; & \exp\left[\int_t^T \exp\left\{\left(\ln(1+\mu_J)-\frac{1}{2}\sigma_J^2\right)e^{-\eta(T-s)}+\frac{1}{2}\sigma_J^2 e^{-2\eta(T-s)}\right\}\beta ds - \beta\tau\right], \end{aligned} \tag{7.24}$$

where $\tau = T - t$.

7.2. *Characteristic Functions and Fundamental PDF*

We define the characteristic function of the logarithm of the stochastic component of the spot price, X_T at time T as follows:

$$\Gamma(t, T, \phi) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{i\phi X_T}\right], \tag{7.25}$$

where i is imaginary unit with $i^2 = -1$ and $\phi \in \mathbb{R}$ is a Fourier parameter.

In order to obtain characteristic function, we need to compute an expectation in (7.25), which is a function of t, X_t and V_t: $\Phi(t, X_t, V_t) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{i\phi X_T}\right]$. By Itô formula, the integral representation for $\Phi(t, X_t, V_t)$ has the following form:

$$\begin{aligned} \Phi(t, X_T, V_T) \; = \; & \Phi(t, X_t, V_t) + \int_t^T \frac{\partial \Phi}{\partial t} ds + \int_t^T \frac{\partial \Phi}{\partial X} dX_s + \frac{1}{2}\int_t^T \frac{\partial^2 \Phi}{\partial X_s^2} dX_s^2 \\ & + \int_t^T \frac{\partial \Phi}{\partial V} dV_s + \frac{1}{2}\int_t^T \frac{\partial^2 \Phi}{\partial V_s^2} dV_s^2 \end{aligned}$$

$$
\begin{aligned}
&+ \int_t^T \frac{\partial^2 \Phi}{\partial X \partial V} dX_s dV_s + \sum_{t<s\leq T} [\Phi(s, X_s, V_s) - \Phi(s, X_{s-}, V_s)] \\
&+ \sum_{t<s\leq T} [\Phi(s, X_s, V_s) - \Phi(s, X_s, V_{s-})] \qquad (7.26)
\end{aligned}
$$

Assuming that the dynamics of the log-price process $X_t = \ln(S_t)$ under the risk-neutral measure $\mathbb{Q}$ take the form

$$
\begin{aligned}
dX_t &= \left(\tilde{a}(X_t, V_t)dt - \frac{1}{2}V_t \right) + \sqrt{V_t} d\tilde{W}_t^X + d\tilde{P}_t^X \\
dV_t &= \tilde{b}(V_t)dt + c(V_t)d\tilde{W}_t^V + d\tilde{P}_t^V, \qquad (7.27)
\end{aligned}
$$

we can substitute the process specifications in (7.27) into the SDE (7.26), which, after simplifying, leads to

$$
\begin{aligned}
\Phi(t, X_T, V_T) &= \Phi(t, X_t, V_t) + \int_t^T \left(\frac{\partial \Phi}{\partial t} + \tilde{A}\Phi \right) ds + \int_t^T \sqrt{V_s} \frac{\partial \Phi}{\partial X} d\tilde{W}_s^X \\
&+ \int_t^T c(V_s) \frac{\partial \Phi}{\partial V} d\tilde{W}_s^V + \sum_{t<s\leq T} [\Phi(s, X_s, V_s) - \Phi(s, X_{s-}, V_s)] \\
&+ \sum_{t<s\leq T} [\Phi(s, X_s, V_s) - \Phi(s, X_s, V_{s-})], \qquad (7.28)
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{A}\Phi = \left(\tilde{a}(X_t, V_t) - \frac{1}{2}V_t \right) \frac{\partial \Phi}{\partial X} + \tilde{b}(V_t) \frac{\partial \Phi}{\partial V} &+ \frac{1}{2} V_t \frac{\partial^2 \Phi}{\partial X^2} + \frac{1}{2} c^2(V_t) \frac{\partial^2 \Phi}{\partial V^2} \\
&+ \rho \sqrt{V_t} c(V_t) \frac{\partial^2 \Phi}{\partial X \partial V} \qquad (7.29)
\end{aligned}
$$

is the so-called risk-neutral generator defined in Oskendal (2002).

If we could show that $\Phi(t, X_t, V_t)$ solves the PDE

$$
\begin{aligned}
\frac{\partial \Phi}{\partial t} + \tilde{A}\Phi &+ \lambda^X \mathbb{E}^{\mathbb{Q}} \left[\Phi(t, X_t + J^X, V_t) - \Phi(t, X_t, V_t) \right] \\
&+ \lambda^V \mathbb{E}^{\mathbb{Q}} \left[\Phi(t, X_t, V_t + J^V) - \Phi(t, X_t, V_t) \right] = 0, \qquad (7.30)
\end{aligned}
$$

where J^X and J^V denote the jump sizes in the log-price and the variance, respectively, equation (7.28) will result in

$$
\Phi(t, X_T, V_T) = \Phi(t, X_t, V_t) + \int_t^T \sqrt{V_s} \frac{\partial \Phi}{\partial X} d\tilde{W}_s^X + \int_t^T c(V_s) \frac{\partial \Phi}{\partial V} d\tilde{W}_s^V, \qquad (7.31)
$$

and, taking the expectation, we obtain

$$
\Phi(t, X_t, V_t) = \mathbb{E}_t^{\mathbb{Q}} \left[\Phi(T, X_T, V_T) | X_t = X_0, V_t = V_0 \right], \qquad (7.32)
$$

where X_0 and V_0 are initial values for X and V, respectively. Thus, taking an expectation in (7.25) is equivalent to solving the PDE (7.30). Equation (7.30), when plugging (7.29), determines the so-called *fundamental PDE* (FPDE):

$$\begin{aligned}\frac{\partial \Phi}{\partial t} &+ \left(\tilde{a}(X_t, V_t) - \frac{1}{2}V_t\right)\frac{\partial \Phi}{\partial X} + \tilde{b}(V_t)\frac{\partial \Phi}{\partial V} + \frac{1}{2}V_t\frac{\partial^2 \Phi}{\partial X^2} + \frac{1}{2}c^2(V_t)\frac{\partial^2 \Phi}{\partial V^2} \\ &+ \rho\sqrt{V_t}c(V_t)\frac{\partial^2 \Phi}{\partial X \partial V} + \lambda^X \mathbb{E}^{\mathbb{Q}}\left[\Phi(t, X_t + J^X, V_t) - \Phi(t, X_t, V_t)\right] \\ &+ \lambda^V \mathbb{E}^{\mathbb{Q}}\left[\Phi(t, X_t, V_t + J^V) - \Phi(t, X_t, V_t)\right] = 0. \end{aligned} \tag{7.33}$$

Noth that if (7.30) holds and the process solving the SDE (7.31) has a bounded variance, that is it a local martingale. Furthermore, if the process has zero drift, than it is a martingale and thus, the price of the future contract can be computed as a discounted expectation under the $\mathbb{Q}$ measure.

The FPDE (7.33) can be solved by taking a guess on the functional form of $\Phi(\cdot)$, which allows to transform it to a set of the ordinary differential equation (ODEs). For a system with two state variables X_t and V_t as in (7.27), the guess is exponentially affine:

$$\Phi(t, X_t, V_t) = \exp\left\{i\phi A(\tau)X_t + B(\tau)V_t + C(\tau)\right\}, \tag{7.34}$$

where $\tau = T - t$. Computing the derivatives $\frac{\partial \Phi}{\partial X}$, $\frac{\partial \Phi}{\partial V}$, $\frac{\partial^2 \Phi}{\partial X^2}$, $\frac{\partial^2 \Phi}{\partial V^2}$ and $\frac{\partial^2 \Phi}{\partial X \partial Y}$, replacing them in (7.33), and collecting terms with constants, X_t, V_t, will lead to a system of ODEs, which can be solved subject to boundary conditions on $A(\cdot)$, $B(\cdot)$ and $C(\cdot)$.