A Bivariate Mutually-Excited Switching Jump Diffusion (BMESJD) for asset prices

Donatien Hainaut*
ISBA, Université Catholique de Louvain, Belgium
Griselda Deelstra†
Department of Mathematics, Université libre de Bruxelles, Belgium

May 25, 2018

Abstract

We propose a new approach for bivariate financial time series modelling which allows for mutual excitation between shocks. Jumps are triggered by changes of regime of a hidden Markov chain whose matrix of transition probabilities is constructed in order to approximate a bivariate Hawkes process. This model, called the Bivariate Mutually-Excited Switching Jump Diffusion (BMESJD) presents several interesting features. Firstly, compared to alternative approaches for modelling the contagion between jumps, the calibration is easier and performed with a modified Hamilton's filter. Secondly, the BMESJD allows for simultaneous jumps when markets are highly stressed. Thirdly, a family of equivalent probability measures under which the BMESJD dynamics are preserved, is well identified. Finally, the BMESJD is a continuous time process that is well adapted for pricing options with two underlying assets.

KEYWORDS : Switching process, Self-Excited process, Jump-diffusions
Classification: 60G46, 60G55, 91G40

1 Introduction

In the last century, financial markets became increasingly interconnected. The consequence of this globalization is that violent shocks to stock markets tend to propagate across the planet. Furthermore, shocks seem to increase the probability of observing new successive jumps, in the original and other markets. In continuous time finance, modelling this mutual excitation between large price moves is a challenging task. Models only based on Brownian motion fail to replicate sudden jumps, whereas jump diffusion models do not capture the interplay between shocks across different markets.

A way to deal with the contagion of jumps between two markets, is to decompose the price variability into two components - a bivariate Brownian process and a mutually-excited bivariate

*Postal address: Voie du Roman Pays 20, 1348 Louvain-la-Neuve (Belgium). E-mail to: donatien.hainaut(at)uclouvain.be
†Postal address: CP210, boulevard du Triomphe, 1050 Bruxelles (Belgium). E-mail to: griselda.deelstra(at)ulb.ac.be

jump process, called Hawkes process[1]. In a univariate setting, Chen and Poon (2013), Boswijk et al. (2015), and Carr and Wu (2016) investigate some self-excited jump diffusion processes for modelling stock index returns. Stabile and Torrisi (2010) consider a risk process with non-stationary Hawkes claims arrivals. Aït-Sahalia et al. (2015) used a bivariate mutually-excited jump diffusion to evaluate the interdependence between worldwide stock markets. Dong et al. (2016) consider a two-dimensional reduced form model for credit risk with regime-switching shot noise intensities However, the main drawback of Hawkes based models is the absence of a robust procedure for parameter estimation as underlined by Rasmussen (2013). Ait-Sahalia et al. (2015) use a generalized method of moments based on local approximations. Embrechts et al. (2011) or Hainaut and Moraux (2017 a) propose a peak over threshold procedure. Chen and Poon (2013) or Hainaut (2017 b) instead use a particle Markov Chain Monte-Carlo (PMCMC) method to estimate parameters. This last approach is computationally intensive and sensitive to the choice of prior and transition density functions.

In Hainaut and Deelstra (2017 c), a model has been studied as a substitute of univariate Hawkes processes. In this article, we propose an alternative to bivariate Hawkes processes which is easy to calibrate, allows for simultaneous jumps and is applicable to bivariate option pricing. In our model, called the Bivariate Mutually-Excited Switching Jump Diffusion (BMESJD), large moves of prices are triggered by the change of regime of a hidden Markov chain. The states of this chain are ranked by increasing levels of mutual-excitation. The matrix of transition probabilities is constructed such that after an upward move of the chain, the probability of climbing again the scale of states increases instantaneously. Our model belongs to the family of regime switching processes. In classic regime switching models as in Honda (2003), Guidolin and Timmermann (2008) or Hainaut and Macgilchrist (2012), the Markov chain modulates the parameters of a diffusion process. In Chourdakis (2005) and Hainaut and Colwell (2016), jumps are introduced and synchronized to the Markov chain transitions in a two or three regimes model. The length of the stay in each regime being memory-less exponential random variables, these models cannot explain the mutual excitation between shocks in financial markets. The BMESJD differs from these previous approaches in several directions. First, shocks are exclusively observed when the Markov chain changes of regime. Second, compared to the existing literature, we consider a large number of regimes. But the parsimony of the model is preserved given that the matrix of transition is parameterized by a small number of variables. Third, the transition probabilities are designed in order to replicate the mutual-excitation behaviour of a bivariate Hawkes process. However, the BMESJD presents an additional feature. Contrary to Hawkes processes, simultaneous jumps may occur with a non-zero probability when markets are stressed. Fourth, we propose a modified version of the Hamilton's filter (1989) to calibrate this new process. Finally, we construct a family of equivalent probability measures under which the dynamics of the BMESJD are preserved. We present the conditions necessary to obtain risk neutral martingale measures and we price exchange options as an example of derivative pricing in this model.

We proceed as follows to introduce the BMESJD. In section 2, we briefly remind the features of bivariate Hawkes processes and their intensity processes. We construct a Markov chain that serves to replicating their behaviour and present the moment generating function of some related jump counting processes. In section 3, we detail the construction of the BMESJD and we provide the moment generating function of the marginal returns. In section 4, we establish the statistical distribution of the sum of bivariate normal and exponential random variables. To the best of our

[1]The very first process, developed by Hawkes (1971), has been used in seismology to model the frequency of earthquakes and aftershocks.

knowledge, this result is new and is required for the implementation of the modified Hamilton's filter. As numerical illustration, we fit the BMESJD to S&P 500 and Euronext 100 time series, over a period from 2005 to 2017. The evolution of the hidden Markov chain allows us to detect periods during which levels of mutual excitations between European and US markets were particularly high. In section 5, we provide a detailed analysis of the change of measure approach in this model and in particular of the martingale condition to be fulfilled by the BMESJD under the risk neutral measure. The last section focuses on the pricing of exchange options and studies the impact of contagion on these derivatives.

2 A bivariate Markov chain approximation for mutually excited processes.

This article proposes a new approach to introduce mutual excitation in the dynamics of large moves of two financial assets or markets. It consists in approaching the intensity processes of a bivariate Hawkes process by using a continuous Markov chain, with a finite number of states, ranked by amplitudes of mutual excitation. The states of the Markov chain with a low and high index correspond respectively to a low and high frequency of mutual jumps. A shock of prices occurs when the Markov chain climbs the scale of states. The transition probabilities of the Markov chain are designed such that after an upward change of state, the probability that the chain climbs again, increases. However if the chain does not jump to a higher state, it moves back with a high probability to a lower state. By the use of this Markov chain, the dynamics of a bivariate Hawkes process can be approached. We first recall the definition of a bivariate Hawkes process and its main features. Next, we construct the Markov chain driving the jump processes.

Let us consider 2 processes $\left(N_t^k\right)_{t\geq 0}$ for $k=1,2$, counting the number of shocks hitting two financial markets or two price dynamics. Their intensity is a stochastic process denoted by $\left(\lambda_t^k\right)_{t\geq 0}$, depending upon the history of the processes. λ_t^k reverts to a level θ_k at a speed α_k and increases of $\eta_{k,1}$ or $\eta_{k,2}$ $(\alpha_k, \theta_k, \eta_{k,1}, \eta_{k,2} \in \mathbb{R}^+)$ when a jump occurs. Intensities are driven by the next system of stochastic differential equations (SDE):

$$\begin{aligned} d\lambda_t^1 &= \alpha_1\left(\theta_1-\lambda_t^1\right)dt+\eta_{11}dN_t^1+\eta_{12}dN_t^2, \\ d\lambda_t^2 &= \alpha_2\left(\theta_2-\lambda_t^2\right)dt+\eta_{21}dN_t^1+\eta_{22}dN_t^2. \end{aligned} \tag{1}$$

These dynamics introduce both contagion and spillover effects between shocks. A jump in one market immediately raises the probability of observing a new jump in the same market and a jump in the other market. The volatility parameters η_{11} and η_{22} tune the levels of self-excitation whereas the other volatility parameters η_{12} and η_{21} define the levels of contagion.

In the remainder, we approximate the dynamics of $(\lambda_t^1, \lambda_t^2)$ by a continuous time Markov chain $(\delta_t)_{t\geq 0}$, for which each state corresponds to a couple of values $(\tilde{\lambda}_t^1, \tilde{\lambda}_t^2)$ approaching $(\lambda_t^1, \lambda_t^2)$. In order to build this process, we have to remind that if Δ_t represents a very small interval of time, then λ_t may be related to $\lambda_{t-\Delta_t}$ as follows

$$\lambda_t^k \approx \lambda_{t-\Delta_t}^k-\alpha_k\left(\lambda_{t-\Delta_t}^k-\theta_k\right)\Delta_t+\sum_{j=1}^{2}\eta_{kj}\left(N_t^j-N_{t-\Delta_t}^j\right). \tag{2}$$

Moreover, conditionally to the sample path of λ_t^k, N_t^k is a heterogeneous Poisson process and therefore the probability that N_t^k jumps twice over Δ_t is nearly null. We will use this feature later to construct the matrix of instantaneous transition probabilities of the Markov chain δ_t. But before, we introduce some additional notations. Given that the number of states of δ_t is finite, we discretize and bound the domain on which $(\tilde{\lambda}_t^1, \tilde{\lambda}_t^2)$ are defined. We assume that $(\delta_t)_{t\geq 0}$ is defined on a probability space Ω and takes its values from a set of $\mathbb{R}^{(n+1)^2}$-valued unit vectors $E_0 = \{e_0, \ldots, e_{(n+1)^2-1}\}$ where $e_j = (0, \ldots, 1, \ldots, 0)^\top$ with the $(j+1)^{\text{th}}$ component equal to 1. The filtration generated by $(\delta_t)_{t\geq 0}$ is denoted by $\{\mathcal{G}_t\}_{t\geq 0}$. We also introduce a parameter Δ_t that is involved later in the construction of the transition probabilities of $(\delta_t)_{t\geq 0}$. If we define $\Delta_{\lambda^k} = \frac{\eta_{kk}}{m}$ and let n be a multiple of m, $(\tilde{\lambda}_t^1, \tilde{\lambda}_t^2)$ take their values in the following set of vectors

$$\{(\theta_1 + i\Delta_{\lambda^1}, \theta_2 + j\Delta_{\lambda^2}) \mid i, j = 0, \ldots, n\}.$$

The next step consists in building the matrix of instantaneous probabilities of transition for δ_t, such that $\tilde{\lambda}_t^1$ and $\tilde{\lambda}_t^2$ have a similar behaviour as the intensity processes of Hawkes processes, λ_t^1 and λ_t^2, as defined by equation (1). The matrix of transition probabilities over a time interval $[t, s]$ is denoted as $P(t, s)$. The elements $p_{i,j}(t, s)$ of this matrix are defined by

$$p_{i,j}(t,s) = P(\delta_s = e_j \mid \delta_t = e_i), \quad i, j \in \{0, ..., (n+1)^2 - 1\}, \tag{3}$$

and represent the probabilities of switching from state i at time t to state j at time s. The probability of the chain being in state i at time t, denoted by $p_i(t)$, depends upon the initial probabilities $p_k(0)$ at time $t = 0$ and the transition probabilities $p_{k,i}(0, t)$, where $k = 0, 1, \ldots, (n+1)^2 - 1$, as follows:

$$p_i(t) = P(\delta_t = e_i) = \sum_{k=0}^{(n+1)^2-1} p_k(0) p_{k,i}(0, t), \quad \forall i \in \{0, ..., (n+1)^2 - 1\}. \tag{4}$$

The matrix $P(t, s)$ of transition probabilities over the time interval $[t, s]$ is the matrix exponential of the generator matrix, denoted by $Q_0 := [q_{i,j}]_{i,j=0:(n+1)^2-1}$, times the length of the time interval:

$$P(t, s) = \exp\left(Q_0(s - t)\right), \quad s \geq t. \tag{5}$$

The elements of the generator matrix Q_0 satisfy the following standard conditions:

$$q_{i,j} \geq 0, \quad \forall i \neq j, \quad \text{and} \quad \sum_{j=0}^{(n+1)^2-1} q_{i,j} = 0, \quad \forall i \in \{0, ..., (n+1)^2 - 1\}. \tag{6}$$

For $i \neq j$, $q_{i,j}$ is the instantaneous probability that the Markov chain transits from state i to state j.

As mentioned above, each regime of the Markov chain corresponds to a couple of intensities $\left(\tilde{\lambda}_t^1, \tilde{\lambda}_t^2\right)$. We adopt the following convention to link the index of the state i for $i = 0, ..., (n+1)^2 - 1$ to values of $\left(\tilde{\lambda}_t^1, \tilde{\lambda}_t^2\right)$: If $\delta_t = e_i$, then

$$\begin{cases} \tilde{\lambda}_t^1 & := \lambda_i^1 = \theta_1 + (i \mod (n+1))\, \Delta_{\lambda^1} \\ \tilde{\lambda}_t^2 & := \lambda_i^2 = \theta_2 + \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_{\lambda^2} \end{cases}$$

As a consequence, $\tilde{\lambda}_t^1$ and $\tilde{\lambda}_t^2$ evolve within our approach in corridors delimited by θ_1 and $\theta_1 + n\Delta_{\lambda^1}$ and θ_2 and $\theta_2 + n\Delta_{\lambda^2}$. As mentioned earlier, if Δ_t is small enough, the probability of switching from the state i to j is $q_{i,j}\Delta_t$ for $i \neq j$. To approach the dynamics of λ_t^1 and λ_t^2 over a period Δ_t, $\delta_{t-\Delta_t}$ should transit from a state in which intensities are equal to $\tilde{\lambda}_{t-\Delta_t}^k$ to another state in which

$$\tilde{\lambda}_t^k = \tilde{\lambda}_{t-\Delta_t}^k - \left\lceil \alpha \left(\tilde{\lambda}_{t-\Delta_t}^k - \theta \right) \frac{\Delta_t}{\Delta_{\lambda^k}} \right\rceil \Delta_{\lambda^k} \quad k = 1, 2 \tag{7}$$

when N_t^1 and N_t^2 do not jump. If one of these processes jumps, $\delta_{t-\Delta_t}$ should transit over a period Δ_t, from a state in which intensities are equal to $\tilde{\lambda}_{t-\Delta_t}^k$ to another state in which

$$\tilde{\lambda}_t^k = \tilde{\lambda}_{t-\Delta_t}^k + \sum_{j=1}^{2} \eta_{kj} \left(N_t^j - N_{t-\Delta_t}^j \right) \quad k = 1, 2. \tag{8}$$

So as to build the generator Q_0 of δ_t such that the dynamics of $\left(\tilde{\lambda}_t^1, \tilde{\lambda}_t^2\right)$ approach the dynamics of λ_t^1 and λ_t^2 over a period Δ_t, we analyze three scenarii:

1. N_t^1 and N_t^2 do not jump,
2. N_t^1 jumps and N_t^2 do not jump,
3. N_t^1 do not jump and N_t^2 jump.

Each scenario will be associated to a transition of δ_t from a state i to a state j. In order to construct the matrix Q_0, we again consider the discrete version of the dynamics of λ_t^1 and λ_t^2, presented in equation (2).

First scenario: No jump for N_t^1 and N_t^2 between $t - \Delta_t$ and t.
From relation (7), if the chain is in state i at time $t - \Delta_t$, we infer that $\tilde{\lambda}_t^1$ must be equal to

$$\begin{aligned}
\tilde{\lambda}_t^1 &= \theta_1 + (i \mod (n+1)) \Delta_{\lambda^1} - \left\lceil \alpha_1 (\theta_1 + (i \mod (n+1)) \Delta_{\lambda^1} - \theta_1) \frac{\Delta_t}{\Delta_{\lambda^1}} \right\rceil \Delta_{\lambda^1} \\
&= \theta_1 + (i \mod (n+1)) \Delta_{\lambda^1} - \lceil \alpha_1 (i \mod (n+1)) \Delta_t \rceil \Delta_{\lambda^1}.
\end{aligned}$$

Similar, if the chain is in state i at time $t - \Delta_t$, $\tilde{\lambda}_t^2$ must be

$$\begin{aligned}
\tilde{\lambda}_t^2 &= \theta_2 + \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_{\lambda^2} - \left\lceil \alpha_2 \left(\theta + \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_{\lambda^2} - \theta_2 \right) \Delta_t \right\rceil \\
&= \theta_2 + \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_{\lambda^2} - \left\lceil \alpha_2 \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_t \right\rceil \Delta_{\lambda^2}.
\end{aligned}$$

Therefore, if N_t^1 and N_t^2 do not jump (and by ignoring possible constraints), the Markov chain δ_t should switch from state i at time $t - \Delta_t$ to state j given by

$$j = i - \lceil \alpha_1 (i \mod (n+1)) \Delta_t \rceil - \left\lceil \alpha_2 \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_t \right\rceil (n+1)$$

at time t and the probability of this movement is equal to $1 - \lambda_i^1 \Delta_t - \lambda_i^2 \Delta_t$. However since the size of the matrix is finite and the different values of $\tilde{\lambda}_t^1 \geq \theta_1$, the index arrival state is bounded by the state given by the sum of the following two maxima[2]:

$$\max\left(\left(i \mod (n+1)\right) - \lceil \alpha_1 \left(i \mod (n+1)\right) \Delta_t \rceil \,;\, 0\right)$$
$$+ \max\left(\left\lfloor \frac{i}{n+1} \right\rfloor (n+1) - \left\lceil \alpha_2 \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_t \right\rceil (n+1)\,;\, 0\right).$$

Second scenario: Only a jump of N_t^1 between $t - \Delta_t$ and t.
In the discrete framework, N_t^1 jumps with a probability $\lambda_t^1 \Delta_t$. When $\delta_{t-\Delta_t} = e_i$ and if we ignore the drift term, the arrival state of δ_t corresponds to the following values for $\tilde{\lambda}_t^1$ and $\tilde{\lambda}_t^2$

$$\tilde{\lambda}_t^1 = \theta_1 + \left[\left(i \mod (n+1)\right) + m\right] \Delta_{\lambda^1}$$
$$\tilde{\lambda}_t^2 = \theta_2 + \left[\left\lfloor \frac{i}{n+1} \right\rfloor + \left\lceil \frac{\eta_{21}}{\Delta_{\lambda^2}} \right\rceil\right] \Delta_{\lambda^2}$$

which corresponds at first sight to the regime $i + m + \left\lceil \frac{\eta_{21}}{\Delta_{\lambda^2}} \right\rceil (n+1)$ of δ_t. However, the size of the matrix being finite, the arrival state must be bounded by the state given by the sum of the following two minima:

$$\min\left(i \mod (n+1) + m\,;\, n\right)$$
$$+ \min\left(\left\lfloor \frac{i}{n+1} \right\rfloor (n+1) + \left\lceil \frac{\eta_{21}}{\Delta_{\lambda^2}} \right\rceil (n+1)\,;\, n\,(n+1)\right).$$

Third scenario: Only a jump of N_t^2 between $t - \Delta_t$ and t.
By construction, N_t^2 jumps with a probability $\lambda_i^2 \Delta_t$ when $\delta_t = e_i$. If we ignore the drift term, the arrival state of δ_t corresponds to the following values for $\tilde{\lambda}_t^1$ and $\tilde{\lambda}_t^2$

$$\tilde{\lambda}_t^1 = \theta_1 + \left(i \mod (n+1)\right) \Delta_{\lambda^1} + \left\lceil \frac{\eta_{12}}{\Delta_{\lambda^1}} \right\rceil \Delta_{\lambda^1}$$
$$\tilde{\lambda}_t^2 = \theta_2 + \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_{\lambda^2} + m\,\Delta_{\lambda^2}$$

which corresponds at first sight to the regime $i + \left\lceil \frac{\eta_{12}}{\Delta_{\lambda^1}} \right\rceil + m(n+1)$. However, the size of the transition matrix being finite, the arrival state is bounded by the state given by the sum of the following two minima:

$$\min\left(i \mod (n+1) + \left\lceil \frac{\eta_{12}}{\Delta_{\lambda^1}} \right\rceil \,;n\right)$$
$$+ \min\left(\left\lfloor \frac{i}{n+1} \right\rfloor (n+1) + m(n+1)\,;\, n\,(n+1)\right).$$

In order to define later the jump processes $\left(\tilde{N}_t^1, \tilde{N}_t^2\right)$ that jump according to the intensities $\left(\tilde{\lambda}_t^1, \tilde{\lambda}_t^2\right)$, we also need to define three $(n+1)^2 \times (n+1)^2$ matrix: $H^1 = \left(h_{i,j}^1\right)_{i,j=0:(n+1)^2-1}$, $H^2 = \left(h_{i,j}^2\right)_{i,j=0:(n+1)^2-1}$ and $H^3 = \left(h_{i,j}^3\right)_{i,j=0:(n+1)^2-1}$. The matrix H^1 (*resp.* H^2) informs us about arrival states of transitions between regimes of δ_t that correspond in a unique way to a

[2]Notice that we use the relation $i = i \mod (n+1) + \left\lfloor \frac{i}{n+1} \right\rfloor (n+1)$

jump of N_t^1 (*resp.* N_t^2). As the domain of $\left(\tilde{\lambda}_t^1, \tilde{\lambda}_t^2\right)$ is bounded, the highest regimes of δ_t (e.g. $\delta_t = e_{(n+1)^2-1}$) may be reached after either a jump of N_t^1 or a jump of N_t^2. These states are identified by the matrix H^3 and will play a particular role in the construction of jump processes $\left(\tilde{N}_t^1, \tilde{N}_t^2\right)$ on which the BMESJD is based.

Algorithm 1 Construction of Q_0, H^1, H^2 and H^3.

Initialization : $(q_{i,j})_{i,j=0:n} = 0$, $\left(h^1_{i,j}\right)_{i,j=0:n} = 0$, $\left(h^2_{i,j}\right)_{i,j=0:n} = 0$, $\left(h^3_{i,j}\right)_{i,j=0:n} = 0$

$$\begin{pmatrix} \lambda_i^1 \\ \lambda_i^2 \end{pmatrix} = \begin{pmatrix} \theta_1 + (i \mod (n+1))\,\Delta_{\lambda^1} \\ \theta_2 + \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_{\lambda^2} \end{pmatrix} \quad \text{for } i = 0, ..., (n+1)^2 - 1.$$

For $i = 0$ to $(n+1)^2 - 1$

For $j = 0$ to $(n+1)^2 - 1$

$$q_{i,j} = \begin{cases} q_{i,j} + \lambda_i^1 & if\ j = \min\left(i \mod (n+1) + m\,;\, n\right) \\ & \quad + \min\left(\left\lfloor \frac{i}{n+1} \right\rfloor (n+1) + \left\lceil \frac{\eta_{21}}{\Delta_{\lambda^2}} \right\rceil (n+1)\,;\, n\,(n+1)\right) \\ \\ q_{i,j} + \lambda_i^2 & if\ j = \min\left(i \mod (n+1) + \left\lceil \frac{\eta_{12}}{\Delta_{\lambda^1}} \right\rceil\,; n\right) \\ & \quad + \min\left(\left\lfloor \frac{i}{n+1} \right\rfloor (n+1) + m(n+1)\,;\, n(n+1)\right) \\ \\ \frac{1}{\Delta_t} - \lambda_i^1 - \lambda_i^2 & if\ j = \max\left((i \mod (n+1)) - \lceil \alpha_1 (i \mod (n+1))\,\Delta_t \rceil\,;\, 0\right) \\ & \quad + \max\left(\left\lfloor \frac{i}{n+1} \right\rfloor (n+1) - \left\lceil \alpha_2 \left\lfloor \frac{i}{n+1} \right\rfloor \Delta_t \right\rceil (n+1)\,;\, 0\right) \\ \\ 0 & \text{else} \end{cases} \tag{9}$$

$$\begin{aligned} h^1_{i,j} &= 1\ if\ j = \min\left(i \mod (n+1) + m\,;\, n\right) \\ &\qquad + \min\left(\left\lfloor \frac{i}{n+1} \right\rfloor (n+1) + \left\lceil \frac{\eta_{21}}{\Delta_{\lambda^2}} \right\rceil (n+1)\,;\, n\,(n+1)\right) \\ h^2_{i,j} &= 1\ if\ j = \min\left(i \mod (n+1) + \left\lceil \frac{\eta_{12}}{\Delta_{\lambda^1}} \right\rceil\,; n\right) \\ &\qquad + \min\left(\left\lfloor \frac{i}{n+1} \right\rfloor (n+1) + m(n+1)\,;\, n(n+1)\right) \end{aligned} \tag{10}$$

If $h^1_{i,j} = h^2_{i,j} = 1$ **then** $h^3_{i,j} = 1$ **and** $h^1_{i,j} = h^2_{i,j} = 0$

End loop on j

$q_{i,i} = -\sum_{j \neq i} q_{i,j}$

End loop on i

The analysis of the three scenarii suggests the following algorithm 1 for the construction of

Q_0, H^1, H^2 and H^3. Notice that $(\delta_t)_{t\geq 0}$ is a continuous Markov chain. Therefore, Δ_t must be seen as a parameter of Q_0 and no more as a time interval (even if it was intuitively its origin for constructing the transition probability matrix). For such a generator, the probability of switching from a state i to j over a small time-interval Δ_t (equal to $q_{i,j}\Delta_t$) is either null, either equal to $\lambda_i^1\Delta_t$, $\lambda_i^2\Delta_t$, $\left(\lambda_i^1+\lambda_i^2\right)\Delta_t$ or $1-\left(\lambda_i^1+\lambda_i^2\right)\Delta_t$. The probability of staying in the same state over Δ_t is nearly null. To ensure the positiveness of non-diagonal elements of the generator, we impose that $\frac{1}{\theta_2+\theta_2+n(\Delta_{\lambda_1}+\Delta_{\lambda_2})}>\Delta_t$.

The remainder of this section studies the properties of the Markov Chain δ_t defined by this generator. Firstly, we define new point processes counting the number of transitions between states. To each pair of distinct states i and j in the state space of the Markov chain δ_t, we define a point process $N_{i,j}(t)$ as follows

$$N_{i,j}(t) := \sum_{0<s\leq t} \boldsymbol{I}_{\{\delta_{s-}=e_i\}}\boldsymbol{I}_{\{\delta_s=e_j\}}\,,$$

where $\boldsymbol{I}$ is the indicator function. $N_{i,j}(t)$ is $\mathcal{G}_t$-adapted and counts the number of transitions from states i to j up to time t. We further define the following intensity process

$$\lambda_{i,j}(t) := q_{i,j}\boldsymbol{I}_{\{\delta_{t-}=e_i\}}\,.$$

Compensating the counting process $N_{i,j}(t)$ by the integral of $\lambda_{i,j}(.)$, the resulting process

$$M_{i,j}(t) := N_{i,j}(t) - \int_0^t \lambda_{i,j}(s)ds\,,$$

is a martingale.

The next step consists in defining $\mathcal{G}_t$-adapted point processes $\tilde{N}_t^1$ and $\tilde{N}_t^2$, counting the number of transitions between states assimilated to jumps of N_t^1 and N_t^2. Let us define

$$\tilde{N}_t^k := \sum_{i=0}^{(n+1)^2-1}\sum_{j=0,j\neq i}^{(n+1)^2-1}\int_0^t e_i^\top\left(H^k+H^3\right)e_j\,dN_{i,j}(s) \tag{11}$$

for $k=1,2$. By construction, the intensity processes are for $k=1,2$ equal to:

$$\begin{aligned}\tilde{\lambda}_t^k &:= \sum_{i=0}^{(n+1)^2-1}\sum_{j=0,j\neq i}^{(n+1)^2-1}\lambda_{i,j}(t)e_i^\top\left(H^k+H^3\right)e_j \\ &= \sum_{i=0}^{(n+1)^2-1}\sum_{j=0,j\neq i}^{(n+1)^2-1} q_{i,j}\delta_{t-}^\top\left(H^k+H^3\right)e_j.\end{aligned} \tag{12}$$

Compensating the counting processes $\tilde{N}_t^1$ and $\tilde{N}_t^2$ by *resp.* the integral of $\tilde{\lambda}_t^1$ and of $\tilde{\lambda}_t^2$, the resulting processes

$$M_t^k := \tilde{N}_t^k - \int_0^t \tilde{\lambda}_s^k ds\,,$$

for $k=1,2$, are martingales. We mentioned earlier that states of the Markov chain identified by the matrix H_3 correspond to either a jump of N_t^1 or N_t^2 (and it cannot be distinguished which one). Such states exist only because the Markov chain counts a finite number of states. In our framework, it is clear that if δ_t transits from states i to j for which $h_{i,j}^3=1$, $\tilde{N}_t^1$ and $\tilde{N}_t^2$ jump at the same time. We think that it is a particularly interesting feature of our bivariate process. When the level of mutual excitation is low, $\tilde{N}_t^1$ and $\tilde{N}_t^2$ never jump simultaneously. On the contrary, when the mutual excitation is high (in other words when δ_t reaches its highest states), $\tilde{N}_t^1$ and $\tilde{N}_t^2$ can jump at the same time with a non-zero probability. The next proposition presents the moment generating function (mgf) of the bivariate jump process.

Proposition 2.1. *The joint mgf of $\tilde{N}_s^1$ and $\tilde{N}_s^2$ for $s \geq t$ with $\omega_1, \omega_2 \in \mathbb{C}_-$, is given by the following expression*

$$\mathbb{E}\left(e^{\omega_1 \tilde{N}_s^1 + \omega_2 \tilde{N}_s^2} \,|\, \mathcal{F}_t\right) = \exp\left(A(\omega_1, \omega_2, t, s, \delta_t) + \omega_1 \tilde{N}_t^1 + \omega_2 \tilde{N}_t^2\right) \qquad u = 1, 2\,.$$

where

$$\begin{aligned}\tilde{A}(\omega_1, \omega_2, t, s) &= \left[e^{A(\omega_1, \omega_2, t, s, e_0)}, \,...,\, e^{A(\omega_1, \omega_2, t, s, e_{(n+1)^2-1})}\right]^\top \\ &= \left[\tilde{A}(\omega_1, \omega_2, t, s, e_0), \,...,\, \tilde{A}(\omega_1, \omega_2, t, s, e_{(n+1)^2-1})\right]^\top\end{aligned}$$

is a $(n+1)^2$-vector of functions, solution of the system of ODE's

$$\begin{aligned} 0 \;=\; & \frac{\partial}{\partial t}\tilde{A}(.) + \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^1 e_j \left(\tilde{A}(., e_j)\, e^{\omega_1} - \tilde{A}(., e_k)\right) \\ & + \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^2 e_j \left(\tilde{A}(., e_j)\, e^{\omega_2} - \tilde{A}(., e_k)\right) \\ & + \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(\tilde{A}(., e_j)\, e^{\omega_1 + \omega_2} - \tilde{A}(., e_k)\right), \end{aligned} \tag{13}$$

for $k = 0, \ldots, (n+1)^2 - 1$, under the terminal boundary conditions

$$\tilde{A}(\omega_1, \omega_2, s, s, e_k) = 1 \quad k = 0, \ldots, (n+1)^2 - 1.$$

Proof of Proposition 2.1. If we denote $f(t, \tilde{N}_t^1, \tilde{N}_t^2, \delta_t) = \mathbb{E}\left(e^{\omega_1 \tilde{N}_s^1 + \omega_2 \tilde{N}_s^2} \,|\, \mathcal{F}_t\right)$, then f is by Itô's lemma solution of the equation:

$$\mathcal{A} f(t, \tilde{N}_t^1, \tilde{N}_t^2, \delta_t) = 0,$$

with the infinitesimal generator $\mathcal{A}$ for $\delta_t = e_k$ ($k \in \{0, \ldots, (n+1)^2 - 1\}$) equal to

$$\begin{aligned} & \mathcal{A} f(t, \tilde{N}_t^1, \tilde{N}_t^2, e_k) \\ &= \frac{\partial f}{\partial t} + \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^1 e_j \left(f(t, \tilde{N}_t^1 + 1, \tilde{N}_t^2, e_j) - f(t, \tilde{N}_t^1, \tilde{N}_t^2, e_k)\right) \\ &+ \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^2 e_j \left(f(t, \tilde{N}_t^1, \tilde{N}_t^2 + 1, e_j) - f(t, \tilde{N}_t^1, \tilde{N}_t^2, e_k)\right) \\ &+ \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(f(t, \tilde{N}_t^1 + 1, \tilde{N}_t^2 + 1, e_j) - f(t, \tilde{N}_t^1, \tilde{N}_t^2, e_k)\right). \end{aligned} \tag{14}$$

Let us assume that f is an exponential affine function of $\tilde{N}_t^1$ and $\tilde{N}_t^2$ as follows:

$$f = \exp\left(A(\omega_1,\omega_2,t,s,e_k) + B^1(\omega_1,t,s)\tilde{N}_t^1 + B^2(\omega_2,t,s)\tilde{N}_t^2\right),$$

where $A(\omega_1,\omega_2,t,s,e_k)$ for $k=0,\ldots,(n+1)^2-1$ and $B^1(\omega_1,t,s)$, $B^2(\omega_2,t,s)$ are time dependent functions with terminal conditions given by $A(\omega_1,\omega_2,s,s,e_k)=0$, $B^1(\omega_1,s,s)=\omega_1$ and $B^2(\omega_2,s,s)=\omega_2$. The partial derivatives of f are then given by:

$$f_t = \left(\frac{\partial}{\partial t}A(\omega_1,\omega_2,t,s,e_k) + \frac{\partial}{\partial t}B^1(\omega_1,t,s)\tilde{N}_t^1 + \frac{\partial}{\partial t}B^2(\omega_2,t,s)\tilde{N}_t^2\right) f,$$

whereas the sums in equation (14) are equal to respectively

$$\sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^1 e_j \left(f(t,\tilde{N}_t^1+1,\tilde{N}_t^2,e_j) - f(t,\tilde{N}_t^1,\tilde{N}_t^2,e_k)\right)$$
$$= f \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^1 e_j \left(e^{A(.,e_j)-A(.,e_k)+B^1(.)} - 1\right),$$

$$\sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^2 e_j \left(f(t,\tilde{N}_t^1,\tilde{N}_t^2+1,e_j) - f(t,\tilde{N}_t^1,\tilde{N}_t^2,e_k)\right)$$
$$= f \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^2 e_j \left(e^{A(.,e_j)-A(.,e_k)+B^2(.)} - 1\right)$$

and

$$\sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(f(t,\tilde{N}_t^1+1,\tilde{N}_t^2+1,e_j) - f(t,\tilde{N}_t^1,\tilde{N}_t^2,e_k)\right)$$

$$= f \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(e^{A(.,e_j)-A(.,e_k)+B^1(.)+B^2(.)} - 1\right).$$

Substituting these expressions into equation (14), leads to the following relation:

$$\begin{aligned} 0 = & \left(\frac{\partial}{\partial t}A(.) + \frac{\partial}{\partial t}B^1(.)\tilde{N}_t^1 + \frac{\partial}{\partial t}B^2(.)\tilde{N}_t^2\right) e^{A(.,e_k)} \\ & + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^1 e_j \left(e^{A(.,e_j)+B^1(.)} - e^{A(.,e_k)}\right) \\ & + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^2 e_j \left(e^{A(.,e_j)+B^2(.)} - e^{A(.,e_k)}\right) \\ & + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(e^{A(.,e_j)+B^1(.)+B^2(.)} - e^{A(.,e_k)}\right), \end{aligned}$$

from which we deduce that $B^1(\omega_1, t, s) = \omega_1$ and $B^2(\omega_2, t, s) = \omega_2$. Regrouping terms allows to infer that the functions A are solutions to the following system of ODE's:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial t} A(.)e^{A(.,e_k)} + \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^1 e_j \left(e^{A(.,e_j)+B^1(.)} - e^{A(.,e_k)} \right) \\
&+ \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^2 e_j \left(e^{A(.,e_j)+B^2(.)} - e^{A(.,e_k)} \right) \\
&+ \sum_{j \neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(e^{A(.,e_j)+B^1(.)+B^2(.)} - e^{A(.,e_k)} \right).
\end{aligned}
$$

Defining $\tilde{A}(\omega_1, \omega_2, t, s) = (e^{A(\omega_1,\omega_2,t,s,e_k)})_{k=0,\ldots,(n+1)^2-1}$, leads to the relation (13).

■

3 The Bivariate Mutually-Excited Switching Jump Diffusion (BMESJD) model

In this section, we construct a bivariate price process $\left(S_t^1, S_t^2\right)$ for financial assets with jumps induced by the jump processes $\tilde{N}_t^1$ and $\tilde{N}_t^2$ defined in (11) and modelled by the use of random variables J_1 and J_2, which are assumed to be independent of the processes $\tilde{N}_t^i$ for $i = 1, 2$. This bivariate process $\left(S_t^1, S_t^2\right)$ is defined on Ω and the filtration generated by the asset prices is denoted by $\{\mathcal{H}_t\}_{t \geq 0}$. We recall that the information about the Markov Chain $(\delta_t)_{t\geq 0}$ is contained in the filtration $\{\mathcal{G}_t\}_{t\geq 0}$. The augmented filtration gathering information about $\left(S_t^1, S_t^2, \delta_t\right)$ is denoted by $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t$. We assume that $(W_t)_t$ is a standard Brownian motion under $\mathbb{P}$ independent of J_i and $\tilde{N}_t^i$ for $i = 1, 2$; and that the instantaneous return of the asset price process is modelled by the following sum of a drift term, a Brownian motion term, and a compensated jumps part:

$$
\begin{pmatrix} \frac{dS_t^1}{S_{t-}^1} \\ \frac{dS_t^2}{S_{t-}^2} \end{pmatrix} = \underbrace{\begin{pmatrix} \mu_t^1 \\ \mu_t^2 \end{pmatrix}}_{\mu_t} dt + \underbrace{\begin{pmatrix} \sigma_t^{11} & 0 \\ \sigma_t^{21} & \sigma_t^{22} \end{pmatrix}}_{\Sigma_t} \underbrace{\begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix}}_{dW_t} + \begin{pmatrix} \left(e^{J_1} - 1\right) d\tilde{N}_t^1 \\ \left(e^{J_2} - 1\right) d\tilde{N}_t^2 \end{pmatrix} - \begin{pmatrix} \tilde{\lambda}_t^1 \mathbb{E}\left(e^{J_1} - 1\right) \\ \tilde{\lambda}_t^2 \mathbb{E}\left(e^{J_2} - 1\right) \end{pmatrix} dt \tag{15}
$$

The drift rates $\mu_t = \left(\mu_t^1, \mu_t^2\right)$, and Brownian volatilities $\left(\sigma_t^{11}, \sigma_t^{21}, \sigma_t^{22}\right)$ are modulated by the Markov chain δ. That is, for $j = 1, 2$, $\mu_t^j = \delta_t^\top \bar{\mu}^j$ where

$$
\bar{\mu}^j = \left(\mu_0^j, \ldots, \mu_{(n+1)^2-1}^j\right)^\top \in \mathbb{R}^{(n+1)^2}
$$

and for the pairs $(i\,j) = (1\,2)\,, \,(2\,1)\,, \,(2\,2)$, $\sigma_t^{ij} = \delta_t^\top \bar{\sigma}^{ij}$ where

$$
\bar{\sigma}^{ij} = \left(\sigma_0^{ij}, \ldots, \sigma_{(n+1)^2-1}^{ij}\right)^\top \in \mathbb{R}_+^{(n+1)^2}.
$$

The matrix of covariance of the Brownian part is then stochastic and equal to $\Sigma\Sigma^{\top}$.

We call this model the bivariate mutually-excited switching jump diffusion (BMESJD) model. Applying Itô's lemma to $\ln S_t^j$ leads to the alternative representation:

$$\begin{pmatrix} d\ln S_t^1 \\ d\ln S_t^2 \end{pmatrix} = \left(\mu_t - \frac{1}{2}diag\left(\Sigma_t\Sigma_t^{\top}\right)\right)dt + \Sigma_t \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix} + \begin{pmatrix} J_1 d\tilde{N}_t^1 \\ J_2 d\tilde{N}_t^2 \end{pmatrix} - \begin{pmatrix} \tilde{\lambda}_t^1 \mathbb{E}\left(e^{J_1}-1\right) \\ \tilde{\lambda}_t^2 \mathbb{E}\left(e^{J_2}-1\right) \end{pmatrix} dt \tag{16}$$

from which we deduce that $\left(S_t^1, S_t^2\right)$ are equal to the following exponential processes:

$$S_t^j = \exp\left(\int_0^t e_j^{\top}\left(\mu_s - \frac{1}{2}diag\left(\Sigma_s\Sigma_s^{\top}\right)\right)ds + \int_0^t e_j^{\top}\Sigma_s dW_s + \int_0^t J_j d\tilde{N}_s^j - \int_0^t \tilde{\lambda}_s^j \mathbb{E}\left(e^{J_j}-1\right)ds\right) \qquad j=1,2 \tag{17}$$

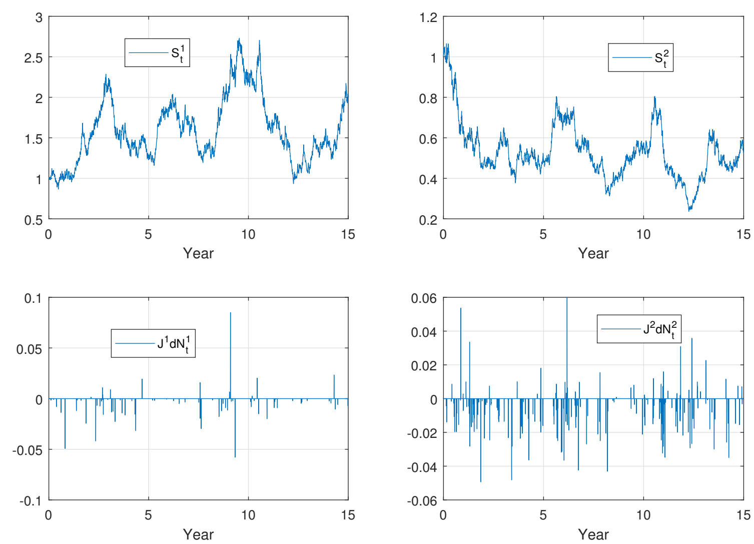

Figure 1: Upper graphs: simulated sample path for S_t^1 and S_t^2. Lower graphs: jump components of S_t^1 and S_t^2.

We assume that J_1 and J_2 are i.i.d. copies of double exponential distributions. Their probability density function (defined on $\mathbb{R}$) is given by

$$\nu_j(z) = p_j \rho_j^+ e^{-\rho_j^+ z} \boldsymbol{I}_{\{z\geq 0\}} - (1-p_j)\rho_j^- e^{-\rho_j^- z} \boldsymbol{I}_{\{z<0\}}, \tag{18}$$

while the associated cumulative distribution function equals

$$\mathbb{P}\left[J_j \leq z\right] = (1-p_j) e^{-\rho_j^- z} \boldsymbol{I}_{\{z\leq 0\}} + \left[(1-p_j) + p_j\left(1-e^{-\rho_j^+ z}\right)\right] \boldsymbol{I}_{\{z>0\}}.$$

This distribution of J_j for $j = 1, 2$ depends on three parameters: $\rho_j^+ \in \mathbb{R}^+$, $\rho_j^- \in \mathbb{R}^-$, and $p_j \in (0, 1)$, where p_j (*resp.* $(1 - p_j)$) denotes the probability of observing an upward (*resp.* downward) exponential jump, and $\frac{1}{\rho_j^+}$ (*resp.* $\frac{1}{\rho_j^-}$) gives the size of an average positive (*resp.* negative) jump. The expected value of the size of the jumps is the weighted sum of these average sizes; $\mathbb{E}(J_j) = p_j \frac{1}{\rho_j^+} + (1 - p_j) \frac{1}{\rho_j^-}$. The moment generating function of J_j is given by

$$\psi_j(\omega) = \mathbb{E}\left(e^{\omega J_j}\right) = p_j \frac{\rho_j^+}{\rho_j^+ - \omega} + (1 - p_j) \frac{\rho_j^-}{\rho_j^- - \omega} \quad j = 1, 2. \tag{19}$$

By construction, the jump processes are mutully-excited. Figure 1 below illustrates this point by showing a simulation of sample paths of S_t^1 and S_t^2. The lower graphs report the jumps hitting each time series. We clearly observe a clustering of jumps caused by the self-excitation and contagion between shocks. In the next section, we propose an algorithm to estimate parameters by using historical financial time series.

We conclude this section by studying the moment generating function of the log-return of $\left(S_t^1, S_t^2\right)$. Hereto we introduce some new notations. First, the drift of $\left(\ln S_t^1, \ln S_t^2\right)$ is denoted by $\tilde{\mu}_t$:

$$\tilde{\mu}_t \quad := \quad \begin{pmatrix} \tilde{\mu}_t^1 \\ \tilde{\mu}_t^2 \end{pmatrix} = \begin{pmatrix} \mu_t^1 \\ \mu_t^2 \end{pmatrix} - \frac{1}{2} diag\left(\Sigma_t \Sigma_t^\top\right) - \begin{pmatrix} \tilde{\lambda}_t^1 \mathbb{E}\left(e^{J_1} - 1\right) \\ \tilde{\lambda}_t^2 \mathbb{E}\left(e^{J_2} - 1\right) \end{pmatrix} \tag{20}$$

such that the log-return $X_t^j := \ln \frac{S_t^j}{S_0^j}$ is given by:

$$X_t^j \quad = \quad \exp\left(\int_0^t e_j^\top \tilde{\mu}_s^j ds + \int_0^t e_j^\top \Sigma_s dW_s + \int_0^t J_j d\tilde{N}_s^j\right). \qquad j = 1, 2 \tag{21}$$

By construction, $\tilde{\mu}_t^j$ takes its value in a $\mathbb{R}^{(n+1)^2}$ vector $\tilde{\mu^j} = (\tilde{\mu}_0^j, ..., \tilde{\mu}_{(n+1)^2-1}^j)$ and is such that $\tilde{\mu}_t^j = \delta_t \tilde{\mu}^j$, for $j = 1, 2$. According to the Itô's lemma for semi-martingales, any function $f(t, X_t^1, X_t^2, \delta_t)$: $\mathbb{R}^+ \times \mathbb{R}^2 \times E_0 \to \mathbb{R}$ that is C^1 with respect to time and C^2 with respect to X_t^1 and X_t^2 admits the following relation for $X_t^1 = x_1$, $X_t^2 = x_2$ and $\delta_t = e_k$:

$$\begin{aligned}
df(t, x_1, x_2, e_k) &= \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x_1}\tilde{\mu}_k^1 + \frac{\partial f}{\partial x_2}\tilde{\mu}_k^2\right) dt \\
&+ \frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}\left(\sigma_k^{11}\right)^2 + \frac{\partial^2 f}{\partial x_2^2}\left(\left(\sigma_k^{21}\right)^2 + \left(\sigma_k^{22}\right)^2\right) + 2\frac{\partial^2 f}{\partial x_2 \partial x_1}\sigma_k^{11}\sigma_k^{21}\right) dt \\
&+ \left(\frac{\partial f}{\partial x_1}\sigma_k^{11} + \frac{\partial f}{\partial x_2}\sigma_k^{21}\right) dW_t^1 + \frac{\partial f}{\partial x_2}\sigma_k^{22} dW_t^2 \\
&+ \sum_{j \neq k}^{(n+1)^2-1} e_k^\top H^1 e_j \left(f(t, x_1 + J_1, x_2, e_j) - f(t, x_1, x_2, e_k)\right) dN_{kj}(t) \\
&+ \sum_{j \neq k}^{(n+1)^2-1} e_k^\top H^2 e_j \left(f(t, x_1, x_2 + J_2, e_j) - f(t, x_1, x_2, e_k)\right) dN_{kj}(t) \\
&+ \sum_{j \neq k}^{(n+1)^2-1} e_k^\top H^3 e_j \left(f(t, x_1 + J_1, x_2 + J_2, e_j) - f(t, x_1, x_2, e_k)\right) dN_{kj}(t).
\end{aligned} \tag{22}$$

The last three terms of this equation correspond respectively to the variation of $f(.)$ caused by a jump of $\tilde{N}_t^1$, a jump of $\tilde{N}_t^2$ and simultaneous jumps of $\tilde{N}_t^1$ and $\tilde{N}_t^2$. The infinitesimal generator $\mathcal{A}f(t,x_1,x_2,e_k)$ is equal to:

$$
\begin{aligned}
\mathcal{A}f(t,x_1,x_2,e_k) &= \frac{\partial f}{\partial t}+\frac{\partial f}{\partial x_1}\tilde{\mu}_k^1+\frac{\partial f}{\partial x_2}\tilde{\mu}_k^2 \\
&+\frac{1}{2}\frac{\partial^2 f}{\partial x_1^2}\left(\sigma_k^{11}\right)^2+\frac{1}{2}\frac{\partial^2 f}{\partial x_2^2}\left(\left(\sigma_k^{21}\right)^2+\left(\sigma_k^{22}\right)^2\right)+\frac{\partial^2 f}{\partial x_2 \partial x_1}\sigma_k^{11}\sigma_k^{21} \\
&+\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^1 e_j\int\left(f(t,x_1+z,x_2,e_j)-f(t,x_1,x_2,e_k)\right)\nu_1(z)dz \\
&+\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^2 e_j\int\left(f(t,x_1,x_2+z,e_j)-f(t,x_1,x_2,e_k)\right)\nu_2(z)dz \\
&+\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^3 e_j\int\int\left(\left(f(t,x_1+z_1,x_2+z_2,e_j)\right.\right. \\
&\qquad\qquad \left.-f(t,x_1,x_2,e_k)\right)\nu_1(z_1)\nu_2(z_2)dz_1dz_2
\end{aligned}
$$

We use these results to infer the moment generator funcion (mgf) of $\left(X_s^1,X_s^2\right)$.

Proposition 3.1. *Let us define the* 2×2 *matrix*

$$
\Sigma_k^{\omega_1,\omega_2} = \begin{pmatrix}\omega_1\sigma_k^{11} & 0\\ \omega_2\sigma_k^{21} & \omega_2\sigma_k^{22}\end{pmatrix} \quad k=0,...,(n+1)^2-1\,,
$$

and $\mathbf{1}=(1,1)^\top$. *The mgf of* $\left(X_s^1,X_s^2\right)$ *for* $s\geq t$ *and* $(\omega_1,\omega_2)\in\mathbb{C}_-^2$, *is given by the following expression*

$$
\mathbb{E}\left(e^{\omega_1 X_s^1+\omega_2 X_s^2}\,|\,\mathcal{F}_t\right) = \left(\frac{S_t^1}{S_0^1}\right)^{\omega_1}\left(\frac{S_t^2}{S_0^2}\right)^{\omega_2}\exp\left(A(\omega_1,\omega_2,t,s,\delta_t)\right), \tag{23}
$$

where

$$
\begin{aligned}
\tilde{A}(\omega_1,\omega_2,t,s) &= \left[e^{A(\omega_1,\omega_2,t,s,e_0)},\,...,\,e^{A(\omega_1,\omega_2,t,s,e_{(n+1)^2-1})}\right]^\top \\
&= \left[\tilde{A}(\omega_1,\omega_2,t,s,e_0),\,...,\,\tilde{A}(\omega_1,\omega_2,t,s,e_{(n+1)^2-1})\right]^\top
\end{aligned}
$$

is a $(n+1)^2-1$ *vector of functions, solution of the following ODE system*

$$
\begin{aligned}
0 &= \frac{\partial}{\partial t}\tilde{A}+\left(\omega_1\tilde{\mu}_k^1+\omega_2\tilde{\mu}_k^2+\frac{1}{2}\mathbf{1}^\top\Sigma_k^{\omega_1,\omega_2}\Sigma_k^{\omega_1,\omega_2\top}\mathbf{1}\right)\tilde{A}+ \\
&+\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^1 e_j\left(\tilde{A}(.,e_j)\psi_1(\omega_1)-\tilde{A}(.,e_k)\right) \\
&+\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^2 e_j\left(\tilde{A}(.,e_j)\psi_2(\omega_2)-\tilde{A}(.,e_k)\right) \\
&+\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^3 e_j\left(\tilde{A}(.,e_j)\psi_1(\omega_1)\psi_2(\omega_2)-\tilde{A}(.,e_k)\right)
\end{aligned}
$$

under the terminal boundary condition:

$$\tilde{A}(\omega_1,\omega_2,s,s,e_k)=1 \quad \textit{for } k=0,\dots,(n+1)^2-1.$$

Proof of Proposition 3.1. Let us denote $f(t,X_t^1,X_t^2,\delta_t)=\mathbb{E}\left(e^{\omega_1 X_s^1+\omega_2 X_s^2}\,|\,\mathcal{F}_t\right)$. If $\delta_t=e_k$, this function is solution of the following equation, implied by the usual argument based on Itô's lemma:

$$\begin{aligned}
0 &= f_t + f_{X_1}\tilde{\mu}_k^1 + f_{X_2}\tilde{\mu}_k^2 + \frac{1}{2}f_{X_1X_1}\left(\sigma_k^{11}\right)^2 + \frac{1}{2}f_{X_2X_2}\left(\left(\sigma_k^{21}\right)^2+\left(\sigma_k^{22}\right)^2\right) + f_{X_1,X_2}\sigma_k^{11}\sigma_k^{21}\\
&+ \sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^1 e_j \int \left(f(t,x_1+z,x_2,e_j)-f(t,x_1,x_2,e_k)\right)\nu_1(z)dz \qquad (24)\\
&+ \sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^2 e_j \int \left(f(t,x_1,x_2+z,e_j)-f(t,x_1,x_2,e_k)\right)\nu_2(z)dz\\
&+ \sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^3 e_j \int\int \left(f(t,x_1+z_1,x_2+z_2,e_j)-f(t,x_1,x_2,e_k)\right)\nu_1(z_1)\nu_2(z_2)dz_1dz_2
\end{aligned}$$

Let us further assume that f is an exponential affine function of $\left(X_t^1,X_t^2\right)$:

$$f \;=\; \exp\left(A(\omega_1,\omega_2,t,s,e_k)+B_1(\omega_1,t,s)X_t^1+B_2(\omega_2,t,s)X_t^2\right),$$

where $A(.,e_k)$ (for $k=0,\dots,n$), $B_1(.)$ and $B_2(.)$ are time dependent functions with terminal conditions $A(\omega_1,\omega_2,s,s,e_k)=0$, $B_1(\omega_1,s,s)=\omega_1$ and $B_2(\omega_2,s,s)=\omega_2$. The partial derivatives of f with respect to the state variables are given by:

$$f_t \;=\; \left(\frac{\partial}{\partial t}A+\frac{\partial}{\partial t}B_1X_t^1+\frac{\partial}{\partial t}B_2X_t^2\right)f,$$

$$\begin{aligned}
f_{X_1} &= B_1 f & f_{X_1X_1} &= \left(B_1\right)^2 f\,,\\
f_{X_2} &= B_2 f & f_{X_2X_2} &= \left(B_2\right)^2 f\,,\\
f_{X_1X_2} &= B_1B_2\, f
\end{aligned}$$

The last terms in equation (24) can be developped as respectively

$$\begin{aligned}
&\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^1 e_j \int \left(f(t,x_1+z,x_2,e_j)-f(t,x_1,x_2,e_k)\right)\nu_1(z)dz\\
&= e^{B_1X_t^1+B_2X_t^2}\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^1 e_j\left(e^{A(.,e_j)}\psi_1(B_1)-e^{A(.,e_k)}\right),
\end{aligned}$$

$$\begin{aligned}
&\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^2 e_j \int \left(f(t,x_1,x_2+z,e_j)-f(t,x_1,x_2,e_k)\right)\nu_2(z)dz\\
&= e^{B_1X_t^1+B_2X_t^2}\sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^2 e_j\left(e^{A(.,e_j)}\psi_2(B_2)-e^{A(.,e_k)}\right)
\end{aligned}$$

and

$$\sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \iint \left(f(t,x_1+z_1,x_2+z_2,e_j) - f(t,x_1,x_2,e_k)\right) \nu_1(z_1)\nu_2(z_2) dz_1 dz_2$$

$$= e^{B_1 X_t^1 + B_2 X_t^2} \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(e^{A(.,e_j)} \psi_1(B_1)\psi_2(B_2) - e^{A(.,e_k)}\right).$$

Injecting these expressions into equation (24), leads to the following relation:

$$\begin{aligned}
0 = & \left(\frac{\partial}{\partial t} A + \frac{\partial}{\partial t} B_1 X_t^1 + \frac{\partial}{\partial t} B_2 X_t^2\right) e^A + \left(B_1 \tilde{\mu}_k^1 + B_2 \tilde{\mu}_k^2\right) e^A \\
& + \frac{1}{2}\left(\left(B_1 \sigma_k^{11}\right)^2 + \left(\left(B_2 \sigma_k^{21}\right)^2 + \left(B_2 \sigma_k^{22}\right)^2\right) + 2 B_1 B_2 \sigma_k^{11} \sigma_k^{21}\right) e^A \\
& + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^1 e_j \left(e^{A(.,e_j)} \psi_1(B_1) - e^{A(.,e_k)}\right) \\
& + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^2 e_j \left(e^{A(.,e_j)} \psi_2(B_2) - e^{A(.,e_k)}\right) \\
& + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(e^{A(.,e_j)} \psi_1(B_1)\psi_2(B_2) - e^{A(.,e_k)}\right)
\end{aligned} \tag{25}$$

from which we infer that $B_1(\omega_1, t, s) = \omega_1$ and $B_2(\omega_2, t, s) = \omega_2$. This fact allows to conclude that $A(\omega_1, \omega_2, t, s, e_k)$ for $k = 0, \ldots, n$ are solutions of the following system of ODE's:

$$\begin{aligned}
0 = & \frac{\partial}{\partial t} A e^A + \left(\omega_1 \tilde{\mu}_k^1 + \omega_2 \tilde{\mu}_k^2\right) e^A \\
& + \frac{1}{2}\left(\left(\omega_1 \sigma_k^{11}\right)^2 + \left(\left(\omega_2 \sigma_k^{21}\right)^2 + \left(\omega_2 \sigma_k^{22}\right)^2\right) + 2 \omega_1 \omega_2 \sigma_k^{11} \sigma_k^{21}\right) e^A \\
& + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^1 e_j \left(e^{A(.,e_j)} \psi_1(\omega_1) - e^{A(.,e_k)}\right) \\
& + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^2 e_j \left(e^{A(.,e_j)} \psi_2(\omega_2) - e^{A(.,e_k)}\right) \\
& + \sum_{j\neq k}^{(n+1)^2-1} q_{k,j} e_k^\top H^3 e_j \left(e^{A(.,e_j)} \psi_1(\omega_1)\psi_2(\omega_2) - e^{A(.,e_k)}\right).
\end{aligned} \tag{26}$$

If we define the matrix $\Sigma_k^{\omega_1,\omega_2}$ as follows

$$\Sigma_k^{\omega_1,\omega_2} = \begin{pmatrix} \omega_1 \sigma_k^{11} & 0 \\ \omega_2 \sigma_k^{21} & \omega_2 \sigma_k^{22} \end{pmatrix}$$

and $\tilde{A}(t,s) = (e^{A(\omega_1,\omega_2,t,s,e_i)})_{i=0,\dots,n}$, equation (26) can easily be rewritten as follows:

$$\begin{aligned} 0 &= \frac{\partial}{\partial t}\tilde{A} + \left(\omega_1\tilde{\mu}_k^1 + \omega_2\tilde{\mu}_k^2 + \frac{1}{2}\mathbf{1}^\top\Sigma_k^{\omega_1,\omega_2}\Sigma_k^{\omega_1,\omega_2\top}\mathbf{1}\right)\tilde{A} \\ &+ \sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^1 e_j\left(\tilde{A}(.,e_j)\psi_1(\omega_1) - \tilde{A}(.,e_k)\right) \\ &+ \sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^2 e_j\left(\tilde{A}(.,e_j)\psi_2(\omega_2) - \tilde{A}(.,e_k)\right) \\ &+ \sum_{j\neq k}^{(n+1)^2-1} q_{k,j}e_k^\top H^3 e_j\left(\tilde{A}(.,e_j)\psi_1(\omega_1)\psi_2(\omega_2) - \tilde{A}(.,e_k)\right). \end{aligned}$$

■

The bivariate density of (S_t^1, S_t^2) may be obtained by inverting the mgf in a numerically way, e.g. by using a discrete Fourier transform algorithm. In theory, we can then use this probability density (pdf) to estimate parameters from time series by log-likelihood maximization techniques. However, this approach is computationally intensive and may be inaccurate due to numerical errors. For these reasons, we propose an alternative estimation method in the next section.

4 BMESJD Parameters estimation with a modified Hamilton filter

At the best of our knowledge, fitting a bivariate Hawkes jump diffusion process is a challenging task. Aït-Sahalia et al. (2015) estimate parameters with a generalized moment matching method. However this approach is based on some approximations of moments. An alternative method consists in implementing a particle Markov Chain Monte Carlo procedure but the speed of convergence to reliable estimates depends upon the choice of the prior distribution of the parameters. The BMESJD model does not present this drawback and can be fitted to a time series with an enhanced version of the Hamilton filter (see Hamilton, 1989). This procedure requires the following new result about the probability density function of the sum of a bivariate normally distributed random variable with independent exponentially distributed random variables.

Proposition 4.1. *Let us denote the elements of the inverse matrix* $\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}$ *by*

$$\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1} = \begin{pmatrix} v_1 & v_{12} \\ v_{12} & v_2 \end{pmatrix}, \qquad (27)$$

and $\Phi_{\Delta\Sigma_i\Sigma_i^\top}(z_1,z_2)$ *the cdf of a bivariate normally distributed random variable, with zero mean and* $\Delta\Sigma_i\Sigma_i^\top$ *as its covariation matrix, which is independent of* J_1 *and* J_2.

1. *The probability density function of the sum* $\begin{pmatrix} J_1 \\ 0 \end{pmatrix} + \Sigma_i \begin{pmatrix} W_\Delta^1 \\ W_\Delta^2 \end{pmatrix}$ *is equal to*

$$\begin{aligned} & g^{J_1}(z_1, z_2 \,|\, \delta_t = e_i) \\ &= p_1 \rho_1^+ e^{-\left(z_1\rho_1^+ - \frac{1}{2}\left(\rho_1^+\right)^2 v_1^{-1}\Delta\right) - \frac{v_{12}}{v_1} z_2 \rho_1^+} \frac{\partial}{\partial z_2} \Phi_{\Delta\Sigma_i\Sigma_i^\top} \left(z_1 - \frac{\rho_1^+}{v_1}\Delta \,,\, z_2\right) \\ &\quad - (1-p_1)\, \rho_1^- e^{-\left(z_1\rho_1^- - \frac{1}{2}\left(\rho_1^-\right)^2 v_1^{-1}\Delta\right) - \frac{v_{12}}{v_1} z_2 \rho_1^-} \times \\ &\qquad \times \left(\frac{\partial}{\partial z_2} \Phi_{\Delta\Sigma_i\Sigma_i^\top} \left(\infty \,,\, z_2\right) - \frac{\partial}{\partial z_2} \Phi_{\Delta\Sigma_i\Sigma_i^\top} \left(z_1 - \frac{\rho_1^-}{v_1}\Delta \,,\, z_2\right)\right) . \end{aligned} \tag{28}$$

2. *The probability density function of the sum* $\begin{pmatrix} 0 \\ J_2 \end{pmatrix} + \Sigma_i \begin{pmatrix} W_\Delta^1 \\ W_\Delta^2 \end{pmatrix}$ *is given by the relation:*

$$\begin{aligned} & g^{J_2}(z_1, z_2 \,|\, \delta_t = e_i) \\ &= p_2 \rho_2^+ e^{-\left(z_2\rho_2^+ - \frac{1}{2}\left(\rho_2^+\right)^2 v_2^{-1}\Delta\right) - \frac{v_{12}}{v_2} z_1 \rho_2^+} \frac{\partial}{\partial z_1} \Phi_{\Delta\Sigma_i\Sigma_i^\top} \left(z_1 \,,\, z_2 - \frac{\rho_2^+}{v_2}\Delta\right) \\ &\quad - (1-p_2)\, \rho_2^- e^{-\left(z_2\rho_2^- - \frac{1}{2}\left(\rho_2^-\right)^2 v_2^{-1}\Delta\right) - \frac{v_{12}}{v_2} z_1 \rho_2^-} \times \\ &\qquad \times \left(\frac{\partial}{\partial z_1} \Phi_{\Delta\Sigma_i\Sigma_i^\top} \left(z_1 \,,\, \infty\right) - \frac{\partial}{\partial z_1} \Phi_{\Delta\Sigma_i\Sigma_i^\top} \left(z_1 \,,\, z_2 - \frac{\rho_2^-}{v_2}\Delta\right)\right) . \end{aligned} \tag{29}$$

3. *Let us define for* $S1 \in \{+,-\}$ *and* $S2 \in \{+,-\}$

$$\alpha_1^{S1\,S2} = \frac{\Delta\left(v_2\rho_1^{S1} - v_{12}\rho_2^{S2}\right)}{v_1 v_2 - \left(v_{12}\right)^2} \quad \textit{and} \quad \alpha_2^{S1\,S2} = \frac{\Delta\left(v_1\rho_2^{S2} - v_{12}\rho_1^{S1}\right)}{v_1 v_2 - \left(v_{12}\right)^2} \,,$$

and

$$\begin{aligned} \gamma^{S1\,S2}(z_1, z_2) &= 2\alpha_1^{S1\,S2}\left(v_1 z_1 + v_{12} z_2\right) + 2\alpha_2^{S1\,S2}\left(v_{12} z_1 + v_2 z_2\right) \\ &\quad - \left(\alpha_1^{S1\,S2}\right)^2 v_1 - 2\alpha_1^{S1\,S2}\alpha_2^{S1\,S2} v_{12} - \left(\alpha_2^{S1\,S2}\right)^2 v_2 \,. \end{aligned}$$

The probability density function of the sum $\begin{pmatrix} J_1 \\ J_2 \end{pmatrix} + \Sigma_i \begin{pmatrix} W_\Delta^1 \\ W_\Delta^2 \end{pmatrix}$ *is then given by the following expression*

$$\begin{aligned} & g^{J_1 J_2}(z_1, z_2 \,|\, \delta_t = e_i) \\ &= p_1 p_2 \rho_1^+ \rho_2^+ e^{-\frac{1}{2\Delta}\gamma^{++}(z_1,z_2)} \Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(z_1 - \alpha_1^{++},\, z_2 - \alpha_2^{++}\right) \\ &\quad - p_1(1-p_2)\rho_1^+\rho_2^- e^{-\frac{1}{2\Delta}\gamma^{+-}(z_1,z_2)} \Big(\Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(z_1 - \alpha_1^{+-}, \infty\right) \\ &\qquad\qquad - \Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(z_1 - \alpha_1^{+-}, z_2 - \alpha_2^{+-}\right)\Big) \\ &\quad - p_2\,(1-p_1)\,\rho_1^-\rho_2^+ e^{-\frac{1}{2\Delta}\gamma^{-+}(z_1,z_2)} \Big(\Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(\infty,\, z_2 - \alpha_2^{-+}\right) \\ &\qquad\qquad - \Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(z_1 - \alpha_1^{-+},\, z_2 - \alpha_2^{-+}\right)\Big) \\ &\quad + (1-p_1)\,(1-p_2)\,\rho_1^-\rho_2^- e^{-\frac{1}{2\Delta}\gamma^{--}(z_1,z_2)} \Big(1 - \Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(\infty, z_2 - \alpha_2^{--}\right) \\ &\qquad\qquad - \Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(z_1 - \alpha_1^{--}, \infty\right) + \Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(z_1 - \alpha_1^{--}, z_2 - \alpha_2^{--}\right)\Big) . \end{aligned} \tag{30}$$

Proof of Proposition 4.1. $\Sigma_i \begin{pmatrix} W_\Delta^1 \\ W_\Delta^2 \end{pmatrix}$ is a bivariate normal distribution with zero mean and its covariance matrix equal to $\Delta\Sigma_i\Sigma_i^\top$. The probability density function of this bivariate Gaussian random variable is denoted by:

$$f(z_1, z_2) = det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{-\frac{1}{2}} e^{-\frac{1}{2}(z_1,z_2)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}(z_1,z_2)^\top}.$$

1. The density function $g^{J_1}(z_1, z_2 \,|\, \delta_t = e_i)$ of $(J_1, 0)^\top + \Sigma_i \left(W_\Delta^1, W_\Delta^2\right)^\top$ is equal to the convolution of ν_1 and f:

$$\begin{aligned} g^{J_1}(z_1, z_2 \,|\, \delta_t = e_i) &= \int_{-\infty}^{+\infty} \nu_1(u) f(z_1 - u, z_2) du \\ &= \frac{p_1\rho_1^+}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}} \int_0^{+\infty} e^{-\frac{1}{2\Delta}\left(2\Delta\rho_1^+ u + (z_1-u,z_2)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}(z_1-u,z_2)^\top\right)} du \\ &\quad - \frac{(1-p)\,\rho_1^-}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}} \int_{-\infty}^{0} e^{-\frac{1}{2\Delta}\left(2\Delta\rho_1^- u + (z_1-u,z_2)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}(z_1-u,z_2)^\top\right)} du\,. \end{aligned}$$

Since the following equality holds (and its analogue for ρ_1^-)

$$\begin{aligned} &-\frac{1}{2\Delta}\left(2\Delta\rho_1^+ u + v_1 (z_1 - u)^2 + 2v_{12} z_2 (z_1 - u) + v_2 z_2^2\right) \\ &\quad = -\frac{1}{2\Delta}\left(\left(z_1 - u - \rho_1^+ v_1^{-1}\Delta\,,\, z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(z_1 - u - \rho_1^+ v_1^{-1}\Delta\,,\, z_2\right)^\top\right) \\ &\quad\quad - \left(z_1\rho_1^+ - \frac{1}{2}\left(\rho_1^+\right)^2 v_1^{-1}\Delta\right) - \frac{v_{12}}{v_1} z_2 \rho_1^+, \end{aligned}$$

the density function $g^{J_1}(z_1, z_2 \,|\, \delta_t = e_i)$ may be rewritten as follows

$$\begin{aligned} &g^{J_1}(z_1, z_2 \,|\, \delta_t = e_i) \\ &= \frac{p_1\rho_1^+ e^{-\left(z_1\rho_1^+ - \frac{1}{2}\left(\rho_1^+\right)^2 v_1^{-1}\Delta\right) - \frac{v_{12}}{v_1} z_2\rho_1^+}}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}} \\ &\quad\quad \times \int_0^{+\infty} e^{-\frac{1}{2\Delta}\left(\left(z_1-u-\rho_1^+ v_1^{-1}\Delta\,,\, z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(z_1-u-\rho_1^+ v_1^{-1}\Delta\,,\, z_2\right)^\top\right)} du \\ &\quad - \frac{(1-p)\,\rho_1^- e^{-\left(z_1\rho_1^- - \frac{1}{2}\left(\rho_1^-\right)^2 v_1^{-1}\Delta\right) - \frac{v_{12}}{v_1} z_2\rho_1^-}}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}} \\ &\quad\quad \times \int_{-\infty}^{0} e^{-\frac{1}{2\Delta}\left(\left(z_1-u-\rho_1^- v_1^{-1}\Delta\,,\, z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(z_1-u-\rho_1^- v_1^{-1}\Delta\,,\, z_2\right)^\top\right)} du\,, \end{aligned} \tag{31}$$

Using the substitution $y_1 = z_1 - u - \rho_1^+ v_1^{-1}\Delta$ implies that $u = z_1 - y_1 - \rho_1^+ v_1^{-1}\Delta$ and $du = -dy_1$. Moreover, if $u = 0$ then $y_1 = z_1 - \rho_1^+ v_1^{-1}\Delta$ and if $u = +\infty$, then $y_1 = -\infty$. As a consequence, the first integral in equation (31) becomes:

$$\begin{aligned}
&\int_0^{+\infty} \exp\left(-\frac{1}{2\Delta}\left(\left(z_1-u-\rho_1^+v_1^{-1}\Delta\,,\,z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(z_1-u-\rho_1^+v_1^{-1}\Delta\,,\,z_2\right)^\top\right)\right)du \\
&\quad = -\int_{z_1-\rho_1^+v_1^{-1}\Delta}^{-\infty} \exp\left(-\frac{1}{2\Delta}\left(\left(y_1\,,\,z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(y_1\,,\,z_2\right)^\top\right)\right)dy_1 \\
&\quad = \int_{-\infty}^{z_1-\rho_1^+v_1^{-1}\Delta} \exp\left(-\frac{1}{2\Delta}\left(\left(y_1\,,\,z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(y_1\,,\,z_2\right)\right)\right)dy_1.
\end{aligned} \tag{32}$$

The second integral in equation (31) turns out to be equal to

$$\begin{aligned}
&\int_{-\infty}^{0} \exp\left(-\frac{1}{2\Delta}\left(\left(z_1-u-\rho_1^-v_1^{-1}\Delta\,,\,z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(z_1-u-\rho_1^-v_1^{-1}\Delta\,,\,z_2\right)^\top\right)\right)du \\
&\quad = -\int_{+\infty}^{z_1-\rho_1^-v_1^{-1}\Delta} \exp\left(-\frac{1}{2\Delta}\left(\left(y_1\,,\,z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(y_1\,,\,z_2\right)^\top\right)\right)dy_1 \\
&\quad = \int_{z_1-\rho_1^-v_1^{-1}\Delta}^{+\infty} \exp\left(-\frac{1}{2\Delta}\left(\left(y_1\,,\,z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(y_1\,,\,z_2\right)^\top\right)\right)dy_1 \\
&\quad = \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\Delta}\left(\left(y_1\,,\,z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(y_1\,,\,z_2\right)^\top\right)\right)dy_1 \\
&\qquad - \int_{-\infty}^{z_1-\rho_1^-v_1^{-1}\Delta} \exp\left(-\frac{1}{2\Delta}\left(\left(y_1\,,\,z_2\right)\left(\Delta\Sigma_i\Sigma_i^\top\right)^{-1}\left(y_1\,,\,z_2\right)^\top\right)\right)dy_1\,.
\end{aligned} \tag{33}$$

From equations (31)-(33), we conclude that $g^{J_1}(z_1,z_2\,|\,\delta_t=e_i)$ equals expression (28).

2. The density function $g^{J_2}(z_1,z_2\,|\,\delta_t=e_i)$ is obtained in the same way.

3. The density function $g^{J_1J_2}(z_1,z_2\,|\,\delta_t=e_i)$ is the double convolution of ν_1, ν_2 and f

$$g^{J_1J_2}(z_1,z_2\,|\,\delta_t=e_i) \;=\; \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\nu_1(u_1)\nu_2(u_2)f(z_1-u_1,z_2-u_2)du_1du_2\,.$$

First, we notice that the product of the jump densities equals

$$\begin{aligned}
\nu_1(u_1)\nu_2(u_2) \;=\; & p_1p_2\rho_1^+\rho_2^+e^{-\rho_1^+u_1-\rho_2^+u_2}\boldsymbol{I}_{\{u_1\geq 0,u_2\geq 0\}} \\
& -p_1(1-p_2)\rho_1^+\rho_2^-e^{-\rho_1^+u_1-\rho_2^-u_2}\boldsymbol{I}_{\{u_1\geq 0,u_2<0\}} \\
& -p_2\,(1-p_1)\,\rho_2^+\rho_1^-e^{-\rho_1^-u_1-\rho_2^+u_2}\boldsymbol{I}_{\{u_1<0,u_2\geq 0\}} \\
& +(1-p_1)\,(1-p_2)\,\rho_1^-\rho_2^-e^{-\rho_1^-u_1-\rho_2^-u_2}\boldsymbol{I}_{\{u_1<0,u_2<0\}}\,,
\end{aligned}$$

Therefore, $g^{J_1J_2}(.)$ can be developped as the sum of four integrals:

$$\begin{aligned}
&g^{J_1J_2}(z_1,z_2\,|\,\delta_t=e_i) \\
&\quad = \frac{p_1p_2\rho_1^+\rho_2^+}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}}I_1(z_1,z_2) - \frac{p_1(1-p_2)\rho_1^+\rho_2^-}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}}I_2(z_1,z_2) \\
&\quad - \frac{p_2\,(1-p_1)\,\rho_2^+\rho_1^-}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}}I_3(z_1,z_2) + \frac{(1-p_1)\,(1-p_2)\,\rho_2^-\rho_1^-}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}}I_4(z_1,z_2)\,,
\end{aligned}$$

where

$$I_1(.) = \int_0^{+\infty}\int_0^{+\infty} e^{-\frac{1}{2\Delta}\left[2\Delta\rho_1^{+}u_1+2\Delta\rho_2^{+}u_2+(z_1-u_1,z_2-u_2)\left(\Delta\Sigma_i\Sigma_i^{\top}\right)^{-1}(z_1-u_1,z_2-u_2)^{\top}\right]}du_1du_2,$$

$$I_2(.) = \int_{-\infty}^{0}\int_0^{+\infty} e^{-\frac{1}{2\Delta}\left[2\Delta\rho_1^{+}u_1+2\Delta\rho_2^{-}u_2+(z_1-u_1,z_2-u_2)\left(\Delta\Sigma_i\Sigma_i^{\top}\right)^{-1}(z_1-u_1,z_2-u_2)^{\top}\right]}du_1du_2,$$

$$I_3(.) = \int_0^{+\infty}\int_{-\infty}^{0} e^{-\frac{1}{2\Delta}\left[2\Delta\rho_1^{-}u_1+2\Delta\rho_2^{+}u_2+(z_1-u_1,z_2-u_2)\left(\Delta\Sigma_i\Sigma_i^{\top}\right)^{-1}(z_1-u_1,z_2-u_2)^{\top}\right]}du_1du_2,$$

$$I_4(.) = \int_{-\infty}^{0}\int_{-\infty}^{0} e^{-\frac{1}{2\Delta}\left[2\Delta\rho_1^{-}u_1+2\Delta\rho_2^{-}u_2+(z_1-u_1,z_2-u_2)\left(\Delta\Sigma_i\Sigma_i^{\top}\right)^{-1}(z_1-u_1,z_2-u_2)^{\top}\right]}du_1du_2\,.$$

If $S_1 \in \{+,-\}$ and $S_2 \in \{+,-\}$, the following relation holds

$$\begin{aligned} &2\Delta\rho_1^{S1}u_1 + 2\Delta\rho_2^{S2}u_2 + (z_1-u_1, z_2-u_2)\left(\Delta\Sigma\Sigma^{\top}\right)^{-1}(z_1-u_1,z_2-u_2)^{\top} \\ &\quad = (z_1-u_1-\alpha_1^{S1\,S2}, z_2-u_2-\alpha_2^{S1\,S2})\left(\Delta\Sigma\Sigma^{\top}\right)^{-1} \\ &\quad\quad \times (z_1-u_1-\alpha_1^{S1\,S2}, z_2-u_2-\alpha_2^{S1\,S2})^{\top} + \gamma^{S1\,S2}(z_1,z_2) \end{aligned}$$

where

$$\begin{aligned} \alpha_1^{S1\,S2} &= \frac{\Delta\left(v_2\rho_1^{S1} - v_{12}\rho_2^{S2}\right)}{v_1v_2 - (v_{12})^2}, \\ \alpha_2^{S1\,S2} &= \frac{\Delta\left(v_1\rho_2^{S2} - v_{12}\rho_1^{S1}\right)}{v_1v_2 - (v_{12})^2}, \end{aligned}$$

and

$$\begin{aligned} \gamma^{S1\,S2}(z_1,z_2) &= 2\alpha_1^{S1\,S2}\left(v_1z_1 + v_{12}z_2\right) + 2\alpha_2^{S1\,S2}\left(v_{12}z_1 + v_2z_2\right) \\ &\quad - \left(\alpha_1^{S1\,S2}\right)^2 v_1 - 2\alpha_1^{S1\,S2}\alpha_2^{S1\,S2}v_{12} - \left(\alpha_2^{S1\,S2}\right)^2 v_2\,. \end{aligned}$$

$g^{J_1 J_2}(.)$ can then be rewritten as

$$\begin{aligned} &g^{J_1 J_2}(z_1, z_2 \,|\, \delta_t = e_i) \\ &\quad = \frac{p_1p_2\rho_1^{+}\rho_2^{+}e^{-\frac{1}{2\Delta}\gamma^{++}(z_1,z_2)}}{det(2\pi\Delta\Sigma_i\Sigma_i^{\top})^{\frac{1}{2}}}I_1^b(z_1,z_2) - \frac{p_1(1-p_2)\rho_1^{+}\rho_2^{-}e^{-\frac{1}{2\Delta}\gamma^{+-}(z_1,z_2)}}{det(2\pi\Delta\Sigma_i\Sigma_i^{\top})^{\frac{1}{2}}}I_2^b(z_1,z_2) \\ &\quad\quad - \frac{p_2\left(1-p_1\right)\rho_1^{-}\rho_2^{+}e^{-\frac{1}{2\Delta}\gamma^{-+}(z_1,z_2)}}{det(2\pi\Delta\Sigma_i\Sigma_i^{\top})^{\frac{1}{2}}}I_3^b(z_1,z_2) \\ &\quad\quad + \frac{\left(1-p_1\right)\left(1-p_2\right)\rho_1^{-}\rho_2^{-}e^{-\frac{1}{2\Delta}\gamma^{--}(z_1,z_2)}}{det(2\pi\Delta\Sigma_i\Sigma_i^{\top})^{\frac{1}{2}}}I_4^b(z_1,z_2) \end{aligned} \tag{34}$$

where

$$I_1^b(.) = \int_0^{+\infty}\int_0^{+\infty} e^{-\frac{1}{2\Delta}\left[(z_1-u_1-\alpha_1^{++},z_2-u_2-\alpha_2^{++})\left(\Delta\Sigma\Sigma^\top\right)^{-1}(z_1-u_1-\alpha_1^{++},z_2-u_2-\alpha_2^{++})^\top\right]}du_1du_2,$$
$$I_2^b(.) = \int_{-\infty}^{0}\int_0^{+\infty} e^{-\frac{1}{2\Delta}\left[(z_1-u_1-\alpha_1^{+-},z_2-u_2-\alpha_2^{+-})\left(\Delta\Sigma\Sigma^\top\right)^{-1}(z_1-u_1-\alpha_1^{+-},z_2-u_2-\alpha_2^{+-})^\top\right]}du_1du_2,$$
$$I_3^b(.) = \int_0^{+\infty}\int_{-\infty}^{0} e^{-\frac{1}{2\Delta}\left[(z_1-u_1-\alpha_1^{-+},z_2-u_2-\alpha_2^{-+})\left(\Delta\Sigma\Sigma^\top\right)^{-1}(z_1-u_1-\alpha_1^{-+},z_2-u_2-\alpha_2^{-+})^\top\right]}du_1du_2,$$
$$I_4^b(.) = \int_{-\infty}^{0}\int_{-\infty}^{0} e^{-\frac{1}{2\Delta}\left[(z_1-u_1-\alpha_1^{--},z_2-u_2-\alpha_2^{--})\left(\Delta\Sigma\Sigma^\top\right)^{-1}(z_1-u_1-\alpha_1^{--},z_2-u_2-\alpha_2^{--})^\top\right]}du_1du_2.$$

Using the substitutions $y_1 = z_1 - u_1 - \alpha_1^{++}$, $y_2 = z_2 - u_2 - \alpha_2^{++}$ implies that $u_1 = z_1 - y_1 - \alpha_1^{++}$, $u_2 = z_2 - y_2 - \alpha_2^{++}$ and that $du_1 = -dy_1$, $du_2 = -dy_2$. Moreover, if $u_1 = u_2 = 0$ then $y_1 = z_1 - \alpha_1^{++}$, $y_2 = z_2 - \alpha_2^{++}$ and if $u_1 = u_2 = +\infty$, then $y_1 = y_2 = -\infty$. As a consequence, the first term in equation (34) becomes

$$\frac{p_1p_2\rho_1^+\rho_2^+e^{-\frac{1}{2\Delta}\gamma^{++}(z_1,z_2)}}{det(2\pi\Delta\Sigma_i\Sigma_i^\top)^{\frac{1}{2}}}\int_{-\infty}^{z_2-\alpha_2^{++}}\int_{-\infty}^{z_1-\alpha_1^{++}} e^{-\frac{1}{2\Delta}\left[(y_1,y_2)\left(\Delta\Sigma\Sigma^\top\right)^{-1}(y_1,y_2)^\top\right]}dy_1dy_2$$
$$= p_1p_2\rho_1^+\rho_2^+e^{-\frac{1}{2\Delta}\gamma^{++}(z_1,z_2)}\Phi_{\Delta\Sigma_i\Sigma_i^\top}\left(z_1-\alpha_1^{++}, z_2-\alpha_2^{++}\right)$$

Applying the same substitutions in the three last integrals of equation (34) leads to the expression (30). ■

Figure 2 shows the three probability density functions (pdf) introduced in proposition (2). We clearly observe the influence of jumps on the bivariate normal random variable: jumps distort the pdf in the predominant direction of the shocks.

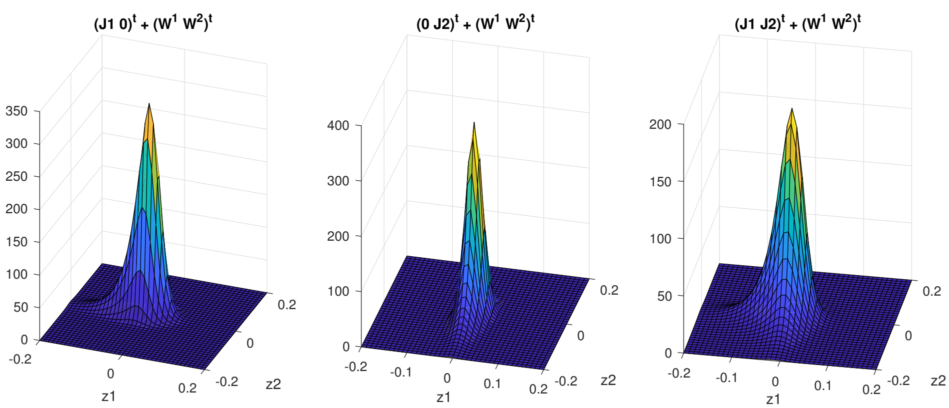

Figure 2: Examples of bivariate distributions. Σ is set to the identity matrix.

In the rest of this section, we denote by $\left(x_1^1, x_1^2\right)$, $\left(x_2^1, x_2^2\right)$, ..., $\left(x_T^1, x_T^2\right)$, the bivariate time series of log-returns of two financial assets or indices, measured at times $t_1, \ldots, t_T$ equally spaced by Δ

(which is not necessary equal to Δ_t involved in the definition of Q_0 for δ_t):

$$x_i^j = \ln\left(\frac{S^j_{t_{i-1}+\Delta}}{S^j_{t_{i-1}}}\right) \quad i=1,...,T\,,\; j=1,2.$$

We assume that the Markov chain δ_t only changes of regime at times t_i for $i=1,\dots,T$. Hence, if the economy stays in the j^{th} state over the period of time $[t_{i-1},t_i]$ with no jumps, then the log-returns are distributed according to a bivariate normal distribution $\left(X_i^1,X_i^2\right)\sim N\left(\left(\tilde{\mu}_j^1\Delta,\tilde{\mu}_j^2\Delta\right)^\top,\Delta\Sigma_j\Sigma_j^\top\right)$. When the Markov chain switches by a jump from regime i to j, the density of the log-return can be found by using proposition 4.1. The set of parameters of the BMESJD is denoted by

$$\Theta=\left(\begin{array}{c}\bar{\mu}^1,\bar{\mu}^2,\bar{\sigma}^{11},\bar{\sigma}^{12},\bar{\sigma}^{22},\alpha_1,\alpha_2,\theta_1,\theta_2,\\ \eta_{11},\eta_{12},p_1,p_2,\rho_1^+,\rho_2^+,\rho_1^-,\rho_1^+\end{array}\right).$$

Δ_t, m and n are not considered as parameters and are chosen a priori.
Using the Bayes' rule, we reformulate the log-likelihood of observed returns as follows:

$$\begin{aligned}\log f\left(\left(x_1^1,x_1^2\right),...,\left(x_T^1,x_T^2\right)|\Theta\right) &= \log f\left(\left(x_1^1,x_1^2\right)|\Theta\right)+\log f\left(\left(x_2^1,x_2^2\right)|\Theta,\left(x_1^1,x_1^2\right)\right)\\ &+\ldots+\log f\left(\left(x_T^1,x_T^2\right)|\Theta,\left(x_1^1,x_1^2\right),\ldots,\left(x_{T-1}^1,x_{T-1}^2\right)\right)\end{aligned}$$

where $f\left(\left(x_k^1,x_k^2\right)|\Theta,\left(x_1^1,x_1^2\right),\ldots\left(x_{k-1}^1,x_{k-1}^2\right)\right)$ is the density function of the return on the k^{th} period, for parameters Θ and conditionally to previous observations $\left(x_1^1,x_1^2\right),\ldots\left(x_{k-1}^1,x_{k-1}^2\right)$. The parameters are estimated by maximizing this log-likelihood function.

Therefore, we concentrate upon the terms in the right-hand side of this log-likelihood. Conditioning upon the state of δ_k allows us to infer that

$$\begin{aligned}&f\left(\left(x_k^1,x_k^2\right)|\Theta,\left(x_1^1,x_1^2\right),\ldots\left(x_{k-1}^1,x_{k-1}^2\right)\right)\\ &=\sum_{i=0}^{(n+1)^2-1}\sum_{j=0}^{(n+1)^2-1}p_i\left(t_{k-1}|\Theta,\left(x_1^1,x_1^2\right),\ldots\left(x_{k-1}^1,x_{k-1}^2\right)\right)\,p_{i,j}(t_{k-1},t_k|\Theta)\\ &\quad\times f\left(\left(x_k^1,x_k^2\right)|\Theta,\delta_{t_k}=e_j\,,\,\delta_{t_{k-1}}=e_i\right)\end{aligned}$$

where

- $f\left(\left(x_k^1,x_k^2\right)|\Theta,\delta_{t_k}=e_j\,,\,\delta_{t_{k-1}}=e_i\right)$ is
  - either the bivariate Gaussian density $N\left(\left(\tilde{\mu}_i^1\Delta,\tilde{\mu}_i^2\Delta\right)^\top,\Delta\Sigma_i\Sigma_i^\top\right)$ in state i, if $i\geq j$,
  - or either
    * $g^{J_1}(z_1-\tilde{\mu}_i^1\Delta,\,z_2-\tilde{\mu}_i^2\Delta\,|\,\delta_t=e_i)$ if $i\neq j$ and $h_{ij}^1=1$, where $g^{J_1}(.)$ is defined by equation (28),
    * $g^{J_2}(z_1-\tilde{\mu}_i^1\Delta,\,z_2-\tilde{\mu}_i^2\Delta\,|\,\delta_t=e_i)$ if $i\neq j$ and $h_{ij}^2=1$, where $g^{J_2}(.)$ is defined by equation (29),
    * $g^{J_1J_2}(z_1-\tilde{\mu}_i^1\Delta,\,z_2-\tilde{\mu}_i^2\Delta\,|\,\delta_t=e_i)$ if $i\neq j$ and $h_{ij}^3=1$, where $g^{J_1J_2}(.)$ is defined by equation (30).
- $p_{i,j}(t_{k-1},t_k|\Theta)$ is the probability of transition, as defined by eq. (3), from state i at time t_{k-1} to state j at time t_k for the set of parameters Θ.

- $p_i\left(t_{k-1}|\Theta,\left(x_1^1,x_1^2\right),\dots\left(x_{k-1}^1,x_{k-1}^2\right)\right)$ is the probability of presence in state i at time t_{k-1}, conditionally to observations up to t_{k-1}.

Using again the Bayes' rule, the probability $p_i\left(t_{k-1}|\Theta,\left(x_1^1,x_1^2\right),\dots\left(x_{k-1}^1,x_{k-1}^2\right)\right)$ may be inferred recursively from $f\left(\left(x_k^1,x_k^2\right)|\Theta,\delta_{t_k}=e_j\,,\,\delta_{t_{k-1}}=e_i\right)$ as follows:

$$\begin{aligned}
&p_i\left(t_{k-1}|\Theta,\left(x_1^1,x_1^2\right),\dots\left(x_{k-1}^1,x_{k-1}^2\right)\right) \\
&\quad= f\left(\left(x_{k-1}^1,x_{k-1}^2\right)|\Theta,\left(x_1^1,x_1^2\right),\dots\left(x_{k-2}^1,x_{k-2}^2\right)\right)^{-1} \\
&\qquad\times\sum_{j=0}^{(n+1)^2-1} p_j\left(t_{k-2}|\Theta,\left(x_1^1,x_1^2\right),\dots\left(x_{k-2}^1,x_{k-2}^2\right)\right) \\
&\qquad\times p_{j,i}(t_{k-2},t_{k-1}|\Theta)\, f\left(\left(x_{k-1}^1,x_{k-1}^2\right)|\Theta,\delta_{t_{k-2}}=e_j\,,\,\delta_{t_{k-1}}=e_i\right)
\end{aligned} \tag{35}$$

To initialize the filter, we need to determine $f\left(\left(x_1^1,x_1^2\right)|\Theta\right)$. If the Markov chain has been running for a sufficiently long enough period of time, we assume that the probability of presence in state i is equal to its stationary probability, $p_i(\Theta)$. Then, we infer that:

$$f\left(\left(x_1^1,x_1^2\right)|\Theta\right)=\sum_{i=0}^{(n+1)^2-1}\sum_{j=0}^{(n+1)^2-1} p_i(\Theta)p_{i,j}(t_0,t_1|\Theta)\, f\left(\left(x_1^1,x_1^2\right)|\Theta,\delta_{t_0}=e_i\,,\,\delta_{t_1}=e_j\right).$$

Therefore, the log-likelihood is computed by recursion and maximized numerically to estimate parameters. After this calibration, we filter the states through which the Markov chain transits by using the relation:

$$\mathbb{E}\left(\delta_{t_k}^{\top}\begin{pmatrix}0\\ \vdots\\ (n+1)^2-1\end{pmatrix}|\,\mathcal{F}_{t_k}\right)=\sum_{i=0}^{(n+1)^2-1} p_i(t_k|\Theta,\left(x_1^1,x_1^2\right),\dots\left(x_k^1,x_k^2\right))\, i.$$

To illustrate our developments, we fit the BMESJD model to time series of the S&P 500 and Euronext 100 (EN 100) stock index, containing daily returns from the 6/9/05 to 5/9/17 (3131 observations). In order to limit the number of parameters to estimate, the drifts μ_t^1, μ_t^2 and the matrix Σ_t are assumed to be constant. This choice also allows us to clearly evaluate the periods of self and mutual excitation between the US and European markets. The parameter of discretization m is either equal to 2 or 4. n is assumed to be equal to $2m$, $3m$ or $4m$ whereas $\Delta_t=\frac{1}{200}$ and Δ is chosen to be equal to one trading day ($\Delta=\frac{1}{252}$). Table 1 reports the log-likelihood and AIC values of several tested models. As we could expect, the goodness of fit is better for the BMESJD model than for a pure diffusion process. Moreover, increasing m and therefore the number of states of δ_t, improves the log-likelihood. Table 2 presents the parameter estimates for $m=4$ and $n=12$. The S&P 500 and the EN 100 turn out to have downward jumps with a probability of respectively 56% and 64%. A comparison of η_{11} and η_{22} indicates that the self-excitation risk is more important in the US than in the European market; whereas the levels of mutual contagion, η_{12} and η_{21}, are comparable. The volatilities of diffusion parts are around 14.19% for the S&P 500 and 17.08% for the EN 100. The higher volatility of the EN 100 can be explained by a smaller diversification of the European index compared to the S&P 500. Figure 3 compares the filtered values for $\tilde{\lambda}_t^1$ (S&P 500) and $\tilde{\lambda}_t^2$ (EN 100) log-returns. Both processes climb the scale of states during the periods of high volatility: from September 2008 to the end 2009 (the US credit crunch period), from September

2011 to February 2012 (the second period of the double-dip recession) or the first months of 2016 (the fear of deflation). We will use the set of parameters reported in Table 2 for further numerical illustrations in the next sections.

Model	Log-likelihood	AIC	Number of states
Brownian Motion	19 275	-38 538	
$m = 2$, $n = 6$	19 900	-39 766	49
$m = 2$, $n = 8$	19 890	-39 745	91
$m = 4$, $n = 8$	19 925	-39 817	91
$m = 4$, $n = 12$	19 933	-39 832	169

Table 1: The first line presents the log-likelihood and AIC for a diffusion with a drift fitted to S&P 500. The other lines present the log-likelihood and AIC for the BMESJD model, with different levels of discretization.

	S&P 500			Euronext 100	
	Values	St.dev.		Values	St.dev.
α_1	16.4878	0.2141	α_2	12.6418	0.1823
θ_1	0.4034	0.0012	θ_2	0.1081	0.0025
η_{11}	32.6738	0.0490	η_{21}	5.8323	0.0012
η_{12}	5.7346	0.0016	η_{22}	11.4693	0.0029
p_1	0.4392	0.0011	p_2	0.3597	0.0021
ρ_1^+	37.6933	0.0759	ρ_2^+	41.9427	0.1148
ρ_1^-	-40.3355	0.0832	ρ_2^-	-39.0573	0.1075
μ_1	0.0561	0.0002	μ_2	0.0220	0.0003
σ^{11}	0.1419	0.0003	σ^{21}	0.1061	0.0003
			σ^{22}	0.1339	0.0004

Table 2: Parameter estimates for $m = 4$, $n = 3$ and $\Delta_t = \frac{1}{200}$.

Notice that it is possible to enhance the fit of our model by considering a regime dependent matrix Σ_t instead of a constant one. To illustrate this, we calibrate a model with $m = 2$ and $n = 6$ and in which Σ_t is either equal to Σ_1 if $\tilde{\lambda}_t^1 \tilde{\lambda}_t^1 \geq g$ or to Σ_2 if $\tilde{\lambda}_t^1 \tilde{\lambda}_t^1 < g$. Here, g is a threshold parameters that we fit by likelihood maximisation. The log-likelihood and AIC for this model are respectively equal to 20 515 and -40 988. A comparison of these results with figures in absence of switching covariance (see Table 1: Log-lik.= 19 900, AIC=-39 766) clearly emphasizes the improvement of the fit caused by the introduction of a switching covariance matrix.

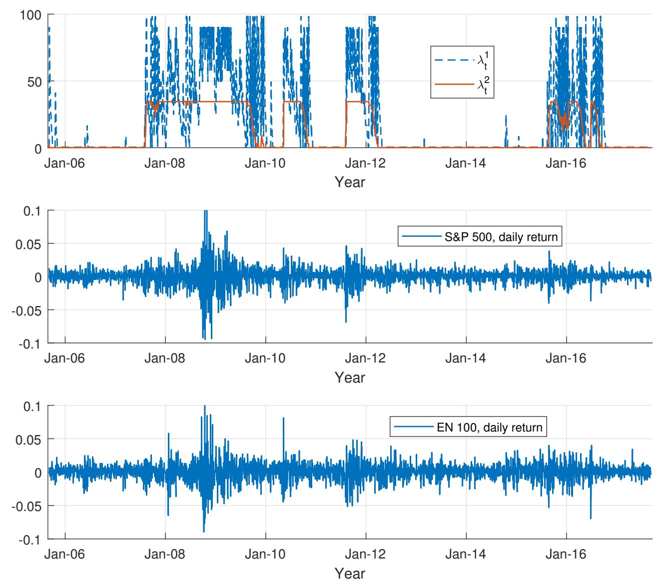

Figure 3: The upper graph shows the filtered sample paths for $\tilde{\lambda}_t^1$ and $\tilde{\lambda}_t^2$. The second and third graph present daily log-returns of resp. the S&P 500 and the EN 100 from September 2005 to 2017.

5 Change of measure

Previous sections focus on features of the pair $\left(S_t^1, S_t^2\right)$ under the real measure $\mathbb{P}$. As option pricing is performed under a risk neutral measure $\mathbb{Q}$, we first need to define the set of eligible probability measures equivalent to $\mathbb{P}$. During this analysis, we will underline the role of the simultaneous jumps under a change of measure. Secondly, we infer the conditions that ensure that discounted stock prices are martingales under the risk neutral measure. Given the nature of $\left(S_t^1, S_t^2\right)$, the market is incomplete and the set of equivalent measures is not finite.
Let $\nu_1^b(.)$ and $\nu_2^b(.)$ denote the pdf of the jumps J_1 and J_2 under $\mathbb{Q}$. We assume that they are defined on the same domain as $\nu_1(.)$ and $\nu_2\,(.)$, the pdf's of the jumps under $\mathbb{P}$. Next, we define the log-ratios:

$$\phi_j\,(b,u) := \ln\left(b\frac{\nu_j^b(u)}{\nu_j(u)}\right) \quad j=1,2 \tag{36}$$

where u is in the support of $\nu_j(.)$ for j=1,2; and $b \in \mathbb{R}^+$ such that the logarithm in equation (36) is well defined. We further take $b_{i,j} \in \mathbb{R}^+$, for $i,j = 0,...,(n+1)^2-1$, such that the following compensated jump processes are well defined and are martingales by construction:

$$\begin{aligned} M_{i,j}(t) =& \sum_{k=1}^{N_{i,j}(t)} \left(e^{\left(h_{ij}^1+h_{ij}^3\right)\phi_1(b_{i,j},J_{1\,k})+\left(h_{ij}^2+h_{ij}^3\right)\phi_2(b_{i,j},J_{2\,k})} - 1 \right) \\ & - \int_0^t \lambda_{i,j}(s)\mathbb{E}\left(e^{\left(h_{ij}^1+h_{ij}^3\right)\phi_1(b_{i,j},J_1)+\left(h_{ij}^2+h_{ij}^3\right)\phi_2(b_{i,j},J_2)} - 1 \,|\, \mathcal{F}_0 \right) ds \\ =& \sum_{k=1}^{N_{i,j}(t)} \left(e^{\left(h_{ij}^1+h_{ij}^3\right)\phi_1(b_{i,j},J_{1\,k})+\left(h_{ij}^2+h_{ij}^3\right)\phi_2(b_{i,j},J_{2\,k})} - 1 \right) \\ & - \int_0^t \lambda_{i,j}(s) \left((b_{i,j})^{h_{ij}^1+h_{ij}^2+2h_{ij}^3} - 1 \right) ds. \end{aligned} \tag{37}$$

We will see later that $b_{i,j}$ is involved in the definition of the transition probabilities of δ_t under the risk neutral measure. Notice that one could consider the generalization that $b_{i,j}$ is a $\mathcal{F}_t-$adapted process but, in this case, the process δ_t would no longer be a Markov chain under equivalent measures.

In the next proposition, we will define an interesting family of equivalent probability measures and the law under the new measures of some point processes, which are used to construct the BMESJD. This proposition follows from Girsanov's theorem for semi-martingales, but we include the proof for comprehension in this specific case.

Proposition 5.1. *Let $Z_{i,j}^1(t)$ and $Z_{i,j}^2(t)$ be point processes defined under $\mathbb{P}$ by*

$$\begin{aligned} Z_{i,j}^1(t) &:= e_i^\top \left(H^1 + H^3\right) e_j \left(\sum_{k=1}^{N_{i,j}(t)} J_{1\,k} \right) \\ Z_{i,j}^2(t) &:= e_i^\top \left(H^2 + H^3\right) e_j \left(\sum_{k=1}^{N_{i,j}(t)} J_{2\,k} \right) \end{aligned} \tag{38}$$

for $i,j = 0,...,(n+1)^2-1$. Under the regularity conditions mentioned above on $b_{i,j}$, the processes $L_{i,j}(t)$ defined as follows

$$\begin{aligned} L_{i,j}(t) =& \exp\left(\int_0^t \left(\left(h_{ij}^1 + h_{ij}^3\right) \phi_1\left(b_{i,j}, J_{1\,s}\right) + \left(h_{ij}^2 + h_{ij}^3\right) \phi_2\left(b_{i,j}, J_{2\,s}\right) \right) dN_{i,j}(s) \right. \\ & \left. - \int_0^t \lambda_{i,j}(s) \left((b_{i,j})^{h_{ij}^1+h_{ij}^2+2h_{ij}^3} - 1 \right) ds \right) \end{aligned} \tag{39}$$

for $i,j = 0,...,(n+1)^2-1$, are Radon-Nikodym derivatives $\frac{d\mathbb{P}^b}{d\mathbb{P}}$ from the real measure $\mathbb{P}$ to a new probability measure $\mathbb{P}^b$. Under $\mathbb{P}^b$, $Z_{i,j}^1(t)$ and $Z_{i,j}^2(t)$ are still point processes but their dynamics are given by

$$\begin{aligned} Z_{i,j}^1(t) &= e_i^\top \left(H^1 + H^3\right) e_j \left(\sum_{k=1}^{N_{i,j}^b(t)} J_{1\,k}^b \right), \\ Z_{i,j}^2(t) &= e_i^\top \left(H^2 + H^3\right) e_j \left(\sum_{k=1}^{N_{i,j}^b(t)} J_{2\,k}^b \right), \end{aligned} \tag{40}$$

where $J^b_{1\,k}$ and $J^b_{2\,k}$ are i.i.d. jumps with pdf $\nu^b_1(u)$ and $\nu^b_2(u)$; and where $N^b_{i,j}(t)$ is a counting process of intensity $\lambda_{i,j}(t)b_{i,j}$ if $h^1_{i,j}=1$ or $h^2_{i,j}=1$, and a counting process of intensity $\lambda_{i,j}(t)b^2_{i,j}$ if $h^3_{i,j}=1$.

Proof of Proposition 5.1. From equation (37), $M_{i,j}(t)$ are martingales satisfying the SDE:

$$
\begin{aligned}
dM_{i,j}(t) &= \left(e^{\left(h^1_{ij}+h^3_{ij}\right)\phi_1(b_{i,j},J_1)+\left(h^2_{ij}+h^3_{ij}\right)\phi_2(b_{i,j},J_2)}-1\right)dN_{i,j}(t) \\
&\quad -\lambda_{i,j}(t)\left((b_{i,j})^{h^1_{ij}+h^2_{ij}+2h^3_{ij}}-1\right)dt\,.
\end{aligned}
$$

Then, we can construct the martingale $L_{i,j}(t)$ with geometric dynamics given by:

$$
\begin{aligned}
dL_{i,j}(t) &= L_{i,j}(t)\,dM_{i,j}(t) \\
&= L_{i,j}(t)\left(e^{\left(h^1_{ij}+h^3_{ij}\right)\phi_1(b_{i,j},J_1)+\left(h^2_{ij}+h^3_{ij}\right)\phi_2(b_{i,j},J_2)}-1\right)dN_{i,j}(t) \\
&\quad -L_{i,j}(t)\lambda_{i,j}(t)\left((b_{i,j})^{h^1_{ij}+h^2_{ij}+2h^3_{ij}}-1\right)dt
\end{aligned}
$$

By applying Itô's lemma, it is easy to see that the differential of $\ln\,L_{i,j}(t)$ is equal to

$$
\begin{aligned}
d\,\ln\,L_{i,j}(t) &= \left(\left(h^1_{ij}+h^3_{ij}\right)\phi_1\left(b_{i,j},J_1\right)+\left(h^2_{ij}+h^3_{ij}\right)\phi_2\left(b_{i,j},J_2\right)\right)dN_{i,j}(t) \\
&\quad -\lambda_{i,j}(t)\left((b_{i,j})^{h^1_{ij}+h^2_{ij}+2h^3_{ij}}-1\right)dt\,.
\end{aligned}
$$

Equation (39) then easily follows.
We now prove relations (40) by identifying the dynamics of $Z^1_{ij}(t)$ and $Z^2_{ij}(t)$ by their moment generating functions (mgf). The mgf of $Z^1_{i,j}(t)$ under the measure $\mathbb{P}^b$ defined by the Radon-Nikodym derivative $L_{i,j}(t)$, is given by

$$
\begin{aligned}
\mathbb{E}^{P^b}\left(e^{uZ^1_{i,j}(t)}\right) &= \mathbb{E}\left(e^{\int_0^t\left(\left(h^1_{ij}+h^3_{ij}\right)(uJ_{1\,s}+\phi_1(b_{i,j},J_{1\,s}))+\left(h^2_{ij}+h^3_{ij}\right)\phi_2(b_{i,j},J_{2\,s})\right)dN_{i,j}(s)}\right. \\
&\qquad \left.\times\; e^{-\int_0^t\lambda_{i,j}(s)\left((b_{i,j})^{h^1_{ij}+h^2_{ij}+2h^3_{ij}}-1\right)ds}\right).
\end{aligned}
\tag{41}
$$

In this proof, we respectively denote the filtrations of $N_{i,j}(t)$ and $\lambda_{i,j}(t)$ by $\mathcal{G}^{i,j}_t\subset\mathcal{F}_t$ and $\mathcal{E}^{i,j}_t\subset\mathcal{F}_t$. Using nested expectations allows to rewrite the expectation (41) as follows:

$$
\begin{aligned}
&\mathbb{E}^{P^b}\left(e^{uZ^1_{i,j}(t)}\right)= \\
&=\mathbb{E}\left[e^{-\int_0^t\lambda_{i,j}(s)\left(b_{i,j}^{h^1_{ij}+h^2_{ij}+2h^3_{ij}}-1\right)ds}\times\right. \\
&\qquad\left.\times\;\mathbb{E}\left(\prod_{k=1}^{N_{i,j}(t)}\mathbb{E}\left(e^{\left(\left(h^1_{ij}+h^3_{ij}\right)(uJ_{1\,k}+\phi_1(b_{i,j},J_{1\,k}))+\left(h^2_{ij}+h^3_{ij}\right)\phi_2(b_{i,j},J_{2\,k})\right)}|\mathcal{G}^{i,j}_t\vee\mathcal{E}^{i,j}_t\right)|\mathcal{E}^{i,j}_t\right)\right]
\end{aligned}
\tag{42}
$$

We assume first that $h^1_{i,j}=1$, which has as a consequence that $h^2_{i,j}=h^3_{i,j}=0$. The expectation embedded into equation (42) then becomes

$$\begin{aligned}
&\mathbb{E}\left(e^{\left(\left(h^1_{ij}+h^3_{ij}\right)(uJ_1+\phi_1(b_{i,j},J_1))+\left(h^2_{ij}+h^3_{ij}\right)\phi_2(b_{i,j},J_1)\right)}|\mathcal{G}^{i,j}_t \vee \mathcal{E}^{i,j}_t\right) \\
&= \mathbb{E}\left(e^{(uJ_1+\phi_1(b_{i,j},J_1))}|\mathcal{G}^{i,j}_t \vee \mathcal{E}^{i,j}_t\right) \\
&= \int b_{i,j} e^{uz_1}\nu^b_1(z_1)dz_1 = b_{i,j}\mathbb{E}\left(e^{uJ^b_1}\right) .
\end{aligned} \tag{43}$$

Similarly, if $h^2_{i,j}=1$ (and thus $h^1_{i,j}=h^3_{i,j}=0$), then

$$\begin{aligned}
&\mathbb{E}\left(e^{\left(\left(h^1_{ij}+h^3_{ij}\right)(uJ_1+\phi_1(b_{i,j},J_1))+\left(h^2_{ij}+h^3_{ij}\right)\phi_2\left(b_{i,j},J^2\right)\right)}|\mathcal{G}^{i,j}_t \vee \mathcal{E}^{i,j}_t\right) \\
&= \mathbb{E}\left(e^{(\phi_2(b_{i,j},J_2))}|\mathcal{G}^{i,j}_t \vee \mathcal{E}^{i,j}_t\right) = b_{i,j} .
\end{aligned} \tag{44}$$

Finally, if $h^3_{i,j}=1$ ($h^2_{i,j}=h^3_{i,j}=0$), given that

$$\begin{aligned}
\phi_1\left(b_{i,j},z_1\right)+\phi_2\left(b_{i,j},z_2\right) &= \ln\left(b_{ij}\frac{\nu^b_1(z_1)}{\nu_1(z_1)}\right)+\ln\left(b_{ij}\frac{\nu^b_2(z_2)}{\nu_2(z_2)}\right) \\
&= \ln\left(\left(b_{ij}\right)^2\frac{\nu^b_1(z_1)}{\nu_1(z_1)}\frac{\nu^b_2(z_2)}{\nu_2(z_2)}\right) ,
\end{aligned}$$

the expectation embedded into equation (42) equals in this case

$$\mathbb{E}\left(e^{(uJ_1+\phi_1(b_{i,j},J_1)+\phi_2(b_{i,j},J_2))}|\mathcal{G}^{i,j}_t \vee \mathcal{E}^{i,j}_t\right) = \left(b_{ij}\right)^2\mathbb{E}\left(e^{uJ^b_1}\right) . \tag{45}$$

Combining equations (43), (44) and (45) allows us to infer that

$$\begin{aligned}
&\mathbb{E}\left(\prod_{k=1}^{N_{i,j}(t)}\mathbb{E}\left(e^{\left(\left(h^1_{ij}+h^3_{ij}\right)(uJ_1+\phi_1(b_{i,j},J_1))+\left(h^2_{ij}+h^3_{ij}\right)\phi_2(b_{i,j},J_2)\right)}|\mathcal{G}^{i,j}_t \vee \mathcal{E}^{i,j}_t\right)|\mathcal{E}^{i,j}_t\right) \\
&= \mathbb{E}\left(\prod_{k=1}^{N_{i,j}(t)}\left(h^1_{i,j}b_{i,j}\mathbb{E}\left(e^{uJ^b_1}\right)+h^2_{i,j}b_{i,j}+h^3_{i,j}\left(b_{i,j}\right)^2\mathbb{E}\left(e^{uJ^b_1}\right)\right)|\mathcal{E}^{i,j}_t\right) .
\end{aligned}$$

Conditionally to $\mathcal{E}^{i,j}_t$, $N_{i,j}(t)$ is an inhomogeneous Poisson process. Using the mgf of an inhomogeneous Poisson law, one easily finds that if $h^1_{i,j}=1$

$$\begin{aligned}
\mathbb{E}\left(\prod_{k=1}^{N_{i,j}(t)} b_{i,j}\mathbb{E}\left(e^{uJ^b_1}\right)|\mathcal{E}^{i,j}_t\right) &= \mathbb{E}\left(e^{N_{i,j}(t)\,\ln\left[b_{i,j}\,\mathbb{E}\left(e^{uJ^b_1}\right)\right]}|\mathcal{E}^{i,j}_t\right) \\
&= \exp\left(\int_0^t \lambda_{i,j}(s)\left(b_{i,j}\mathbb{E}\left(e^{uJ^b_1}\right)-1\right)ds\right) .
\end{aligned}$$

Similarly, if $h^2_{i,j}=1$, we obtain that

$$\begin{aligned}
\mathbb{E}\left(\prod_{k=1}^{N_{i,j}(t)} b_{i,j}|\mathcal{E}^{i,j}_t\right) &= \mathbb{E}\left(e^{N_{i,j}(t)\,\ln[b_{i,j}]}|\mathcal{E}^{i,j}_t\right) \\
&= \exp\left(\int_0^t \lambda_{i,j}(s)\left(b_{i,j}-1\right)ds\right) .
\end{aligned}$$

Finally, if $h^3_{i,j} = 1$, we deduce that

$$\mathbb{E}\left(\prod_{k=1}^{N_{i,j}(t)} b^2_{i,j}\mathbb{E}\left(e^{uJ_1^b}\right)|\mathcal{E}_t^{i,j}\right) \quad = \quad \exp\left(\int_0^t \lambda_{i,j}(s)\left((b_{i,j})^2\,\mathbb{E}\left(e^{uJ_1^b}\right)-1\right)ds\right).$$

Combining these last results leads to the fact that expectation (42) is equal to

$$\mathbb{E}^{P^b}\left(e^{uZ^1_{i,j}(t)}\right) =$$
$$= \mathbb{E}\left(e^{-\int_0^t \lambda_{i,j}(s)\left(b_{i,j}^{h^1_{ij}+h^2_{ij}+2h^3_{ij}}-1\right)ds+\int_0^t \lambda_{i,j}(s)\left(b_{i,j}^{h^1_{ij}+h^2_{ij}+2h^3_{ij}}\mathbb{E}\left(e^{uJ_1^b}\right)^{(h^1_{ij}+h^3_{ij})}-1\right)ds}\right)$$

which turns out to be the moment generating function of $Z^1_{i,j}(t)$ under the equivalent measure $\mathbb{P}^b$. The same result holds for $Z^2_{i,j}(t)$. ■

It is well-known that arbitrage opportunities are avoided by pricing financial derivatives under an equivalent martingale measure under which discounted (non-dividend paying) asset prices are martingales. In the remainder of this section, we consider a financial market composed of three assets: a risk free cash account and two stocks, (S^1_t, S^2_t). The interest rate depends on the Markov chain δ_t and is defined as $r_t = \delta_t\bar{r}^\top$ where $\bar{r} = \left(r_0, ..., r_{(n+1)^2-1}\right)^\top \in \mathbb{R}^{(n+1)^2-1}$. The stock prices S^1_t and S^2_t follow a BMESJD, defined by equation (15). By construction, the risk neutral measure is not unique. We consider Radon-Nikodym derivatives of the following form

$$\begin{aligned} L_t = \prod_{i,j=0}^{(n+1)^2-1} & \exp\left(\int_0^t \left(\left(h^1_{ij}+h^3_{ij}\right)\phi_1\left(b_{i,j}, J_{1\,s}\right)+\left(h^2_{ij}+h^3_{ij}\right)\phi_2\left(b_{i,j}, J_{2\,s}\right)\right)dN_{i,j}(s)\right) \\ & \times \exp\left(-\int_0^t \lambda_{i,j}(s)\left((b_{i,j})^{h^1_{ij}+h^2_{ij}+2h^3_{ij}}-1\right)ds\right) \\ & \times \exp\left(-\frac{1}{2}\int_0^t |\beta_s|^2 ds+\int_0^t \beta_s dW_s\right) \end{aligned} \tag{46}$$

where $\beta_t = \left(\beta^1_t, \beta^2_t\right)^\top$ is a bivariate $\mathcal{G}_t$ measurable process such that $\beta^1_t = \delta_t^\top\beta^1$ and $\beta^2_t = \delta_t^\top\beta^2$ where

$$\begin{aligned} \beta^1 &= \left(\beta^1_0, \beta^1_0, \ldots, \beta^1_{(n+1)^2-1}\right)^\top, \\ \beta^2 &= \left(\beta^2_0, \beta^2_0, \ldots, \beta^2_{(n+1)^2-1}\right)^\top. \end{aligned}$$

The last factor in the definition of the Radon-Nikodym derivative L_t implies that

$$\begin{pmatrix} dW_t^{1\,\beta} \\ dW_t^{2\,\beta} \end{pmatrix} = \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix} + \begin{pmatrix} \beta^1_t \\ \beta^2_t \end{pmatrix} dt$$

is a Brownian motion under the equivalent measure. The next proposition establishes the dynamics of S^1_t and S^2_t under such a new martingale measure, which will be denoted by $\mathbb{Q}$. It is interesting to remark that jump intensities are proportional to the square of $b_{i,j}$ for transitions from i to j causing simultaneous jumps (i.e. the case that $h^3_{i,j} = 1$).

Proposition 5.2. *The dynamics of the asset price under the equivalent measure $\mathbb{Q}$ defined by the Radon-Nikodym derivative (46) equals*

$$\begin{pmatrix} \frac{dS_t^1}{S_{t^-}^1} \\ \frac{dS_t^2}{S_{t^-}^2} \end{pmatrix} = (\mu_t - \Sigma_t \beta_t)\, dt + \Sigma_t dW_t^\beta \tag{47}$$
$$+ \begin{pmatrix} \left(e^{J_1^b} - 1\right) d\tilde{N}_t^1 - \tilde{\lambda}_t^{1\,b} \mathbb{E}\left(e^{J_1^b} - 1\right) dt \\ \left(e^{J_2^b} - 1\right) d\tilde{N}_t^2 - \tilde{\lambda}_t^{2\,b} \mathbb{E}\left(e^{J_2^b} - 1\right) dt \end{pmatrix}$$
$$+ \begin{pmatrix} \tilde{\lambda}_t^{1\,b} \mathbb{E}\left(e^{J_1^b} - 1\right) - \tilde{\lambda}_t^1 \mathbb{E}\left(e^{J_1} - 1\right) \\ \tilde{\lambda}_t^{2\,b} \mathbb{E}\left(e^{J_2^b} - 1\right) - \tilde{\lambda}_t^2 \mathbb{E}\left(e^{J_2} - 1\right) \end{pmatrix} dt$$

where $\tilde{N}_t^1$ and $\tilde{N}_t^2$ are point processes with respective intensities:

$$\tilde{\lambda}_t^{1\,b} := \sum_{\substack{i\,,\,j=0 \\ i \neq j}}^{(n+1)^2-1} \lambda_{i,j}(t) \left(h_{i,j}^1\, b_{i,j} + h_{i,j}^3\, (b_{ij})^2\right)$$
$$\tilde{\lambda}_t^{2\,b} := \sum_{\substack{i\,,\,j=0 \\ i \neq j}}^{(n+1)^2-1} \lambda_{i,j}(t) \left(h_{i,j}^2\, b_{i,j} + h_{i,j}^3\, (b_{ij})^2\right)$$

under $\mathbb{Q}$.

Proof of Proposition 5.2. In the following expressions, the double sums over e.g. i and j are such that the term $i = j$ is excluded. For notational use, we define the set

$$\mathcal{I} := \left\{(i,j) \mid i,j \in \left(0, 1, ..., (n+1)^2 - 1\right),\ i \neq j\right\}$$

Notice that by defining $dY_t := \left(d\ln S_t^1, d\ln S_t^2\right)$ equation (16) can be rewritten as follows under $\mathbb{Q}$

$$\begin{pmatrix} dY_t^1 \\ dY_t^2 \end{pmatrix} = \left(\mu_t - \frac{1}{2} diag\left(\Sigma_t \Sigma_t^\top\right) - \Sigma_t \beta_t\right) dt + \Sigma_t \left(\begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix} + \beta_t dt\right)$$
$$+ \begin{pmatrix} \sum_{(i,j)\in\mathcal{I}} dZ_{i,j}^1(t) \\ \sum_{(i,j)\in\mathcal{I}} dZ_{i,j}^2(t) \end{pmatrix} - \begin{pmatrix} \sum_{(i,j)\in\mathcal{I}} \lambda_{i,j}(t) \left(h_{i,j}^1\, b_{i,j} + h_{i,j}^3\, b_{i,j}^2\right) \mathbb{E}\left(e^{J_1^b} - 1\right) \\ \sum_{(i,j)\in\mathcal{I}} \lambda_{i,j}(t) \left(h_{i,j}^2\, b_{i,j} + h_{i,j}^3\, b_{i,j}^2\right) \mathbb{E}\left(e^{J_2^b} - 1\right) \end{pmatrix} dt$$
$$+ \begin{pmatrix} \sum_{(i,j)\in\mathcal{I}} \lambda_{i,j}(t) \left(\left(h_{i,j}^1\, b_{i,j} + h_{i,j}^3\, b_{i,j}^2\right) \mathbb{E}\left(e^{J_1^b} - 1\right) - \left(h_{i,j}^1 + h_{i,j}^3\right) \mathbb{E}\left(e^{J_1} - 1\right)\right) \\ \sum_{(i,j)\in\mathcal{I}} \lambda_{i,j}(t) \left(\left(h_{i,j}^2\, b_{i,j} + h_{i,j}^3\, b_{i,j}^2\right) \mathbb{E}\left(e^{J_2^b} - 1\right) - \left(h_{i,j}^2 + h_{i,j}^3\right) \mathbb{E}\left(e^{J_2} - 1\right)\right) \end{pmatrix} dt$$

Applying Itô's lemma to the function $f(Y_t) = e^{Y_t}$ leads to the dynamics under $\mathbb{Q}$:

$$\begin{pmatrix} \frac{dS_t^1}{S_{t-}^1} \\ \frac{dS_t^2}{S_{t-}^2} \end{pmatrix} = (\mu_t - \Sigma_t \beta_t)\, dt + \Sigma_t dW_t^\beta + \begin{pmatrix} \left(e^{J_1^b} - 1\right) d\tilde{N}_t^1 - \tilde{\lambda}_t^{1\,b} \mathbb{E}\left(e^{J_1^b} - 1\right) dt \\ \left(e^{J_2^b} - 1\right) d\tilde{N}_t^2 - \tilde{\lambda}_t^{2\,b} \mathbb{E}\left(e^{J_2^b} - 1\right) dt \end{pmatrix} + \begin{pmatrix} \tilde{\lambda}_t^{1\,b} \mathbb{E}\left(e^{J_1^b} - 1\right) - \tilde{\lambda}_t^1 \mathbb{E}\left(e^{J_1} - 1\right) \\ \tilde{\lambda}_t^{2\,b} \mathbb{E}\left(e^{J_2^b} - 1\right) - \tilde{\lambda}_t^2 \mathbb{E}\left(e^{J_2} - 1\right) \end{pmatrix} dt \tag{48}$$

■

Given that under the risk neutral measure, all assets earn on average the risk free rate, one easily obtains the condition that ensures that L_t defines a pricing measure:

Corollary 5.3. *An equivalent probability measure defined by the Radon-Nikodym derivative (46) is a risk neutral measure if and only if the following constraints are satisfied:*

$$r_i = (\mu_i - (\Sigma_t \beta_t)_i) + \begin{pmatrix} \sum_{\substack{j=0 \\ i \neq j}}^{(n+1)^2-1} q_{i,j} \left(\left(h_{i,j}^1\, b_{i,j} + h_{i,j}^3\, (b_{i,j})^2\right) \mathbb{E}\left(e^{J_1^b} - 1\right) - \left(h_{i,j}^1 + h_{i,j}^3\right) \mathbb{E}\left(e^{J_1} - 1\right)\right) \\ \sum_{\substack{j=0 \\ i \neq j}}^{(n+1)^2-1} q_{i,j} \left(\left(h_{i,j}^2\, b_{i,j} + h_{i,j}^3\, (b_{i,j})^2\right) \mathbb{E}\left(e^{J_2^b} - 1\right) - \left(h_{i,j}^2 + h_{i,j}^3\right) \mathbb{E}\left(e^{J_2} - 1\right)\right) \end{pmatrix} \tag{49}$$

If the no-arbitrage condition (49) is fulfilled, then the dynamics of $\left(S_t^1, S_t^2\right)$ are equivalent to

$$\begin{pmatrix} \frac{dS_t^1}{S_{t-}^1} \\ \frac{dS_t^2}{S_{t-}^2} \end{pmatrix} = r_t dt + \Sigma_t dW_t^\beta + \begin{pmatrix} \left(e^{J_1^b} - 1\right) d\tilde{N}_t^1 - \tilde{\lambda}_t^{1\,b} \mathbb{E}\left(e^{J_1^b} - 1\right) dt \\ \left(e^{J_2^b} - 1\right) d\tilde{N}_t^2 - \tilde{\lambda}_t^{2\,b} \mathbb{E}\left(e^{J_2^b} - 1\right) dt \end{pmatrix}$$

where $\tilde{\lambda}_t^{1\,b}$ and $\tilde{\lambda}_t^{2\,b}$ are driven by a Markov chain with transition probabilities

$$q_{i,j}^b \quad := \quad q_{i,j} \left(h_{i,j}^1\, b_{i,j} + h_{i,j}^2 b_{i,j} + h_{i,j}^3\, (b_{i,j})^2\right) .$$

It is worth to notice that the transition probabilities corresponding to simultaneous jumps are multiplied by the square of $b_{i,j}$ under $\mathbb{Q}$, whereas other probabilities are only multiplied by $b_{i,j}$.

6 Pricing of exchange options

The purpose of this section is multiple. Firstly, we illustrate that our model is well adapted for option pricing. Secondly, we apply the technique of change of numeraire to the BMESJD model. Finally, we illustrate numerically the influence of contagion on exchange option prices. We assume for the sake of simplicity that the risk free rate is constant $r_t = r \in \mathbb{R}$. We consider an European option that allows to exchange γS_T^1 with $\gamma \in \mathbb{R}^+$ against S_T^2, at expiry T. The option price at time t is the discounted value of the payoff under a risk neutral measure:

$$C(t, \delta_t) = \mathbb{E}^{\mathbb{Q}}\left(e^{-r(T-t)}\left(S_T^2 - \gamma S_T^1\right)_+ \mid \mathcal{F}_t\right) .$$

The joint-moment generating function of (S_T^2, S_T^2) is known (see proposition 3.1). In theory, we can then invert this mgf numerically with a bivariate Discrete Fourier transform, in order to retrieve the bivariate pdf of (S_T^2, S_T^2) under $\mathbb{Q}$. However this approach is more computationally intensive and less elegant than performing a change of numeraire which has already been applied in a lot of other settings (see e.g. Appendix 1). Let us denote by B_t, the cash account $B_t = B_0 e^{rt}$. If we choose S_t^1 as numeraire, the change of measure from $\mathbb{Q}$ to $\mathbb{Q}^{S_1}$ is defined by:

$$\frac{d\mathbb{Q}^{S_1}}{d\mathbb{Q}} = \frac{S_T^1}{S_0^1}\frac{1}{e^{rT}} .$$

Under the measure $\mathbb{Q}^{S_1}$, ratios of assets or derivative payoffs on S_t^1 are martingales. It is easy to check that $\frac{d\mathbb{Q}^{S_1}}{d\mathbb{Q}}$ is a martingale under $\mathbb{Q}$ and that:

$$\mathbb{E}^{\mathbb{Q}}\left(\frac{d\mathbb{Q}^{S_1}}{d\mathbb{Q}}\left(\frac{S_T^2}{S_T^1}-\gamma\right)_+ |\mathcal{F}_t\right) = \mathbb{E}^{\mathbb{Q}}\left(\frac{e^{-rT}}{S_0^1}\left(S_T^2-\gamma S_T^1\right)_+ |\mathcal{F}_t\right) .$$

We can then deduce the following equality:

$$\mathbb{E}^{\mathbb{Q}^{S_1}}\left(\left(\frac{S_T^2}{S_T^1}-\gamma\right)_+ |\mathcal{F}_t\right) = \frac{\mathbb{E}^{\mathbb{Q}}\left(e^{-r(T-t)}\left(S_T^2-\gamma S_T^1\right)_+ |\mathcal{F}_t\right)}{S_t^1} .$$

The value of the exchange option is then equal to the following expectation under the measure $\mathbb{Q}^{S^1}$:

$$C(t,\delta_t) = S_t^1\mathbb{E}^{\mathbb{Q}^{S_1}}\left(\left(\frac{S_T^2}{S_T^1}-\gamma\right)_+ |\mathcal{F}_t\right) .$$

The next proposition establishes the moment generating function of $\frac{S_t^2}{S_t^1}$ under the measure $\mathbb{Q}^{S_1}$. Inverting this mgf by a one dimensional discrete Fourier transform will allow us to price the exchange option in an efficient way.

Proposition 6.1. *The mgf of* $\frac{S_s^2}{S_s^1}$ *for* $s \geq t$ *under the measure* $\mathbb{Q}^{S_1}$*, is given by the following expression*

$$\mathbb{E}^{\mathbb{Q}^{S_1}}\left(e^{\omega\frac{S_s^2}{S_s^1}} |\mathcal{F}_t\right) = e^{-r(s-t)}\left(\frac{S_t^2}{S_t^1}\right)^{\omega}\exp\left(A(1-\omega,\omega,t,s,\delta_t)\right) \tag{50}$$

where $A(\omega, 1-\omega, t, s, \delta_t)$ *is defined in Proposition 3.1, in which parameters under* $\mathbb{P}$ *are replaced by these under* $\mathbb{Q}$*.*

Proof of Corollary 6.1. In order to prove this result, we develop the next expectation:

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{S_1}}\left(e^{\omega \ln \frac{S_s^2}{S_s^1}} \mid \mathcal{F}_t\right) &= \frac{\mathbb{E}^{\mathbb{Q}}\left(e^{-rs}\frac{S_s^1}{S_0^1} e^{\omega \ln \frac{S_s^2/S_0^2}{S_s^1/S_0^1}} \mid \mathcal{F}_t\right)}{e^{-rt}\frac{S_t^1}{S_0^1}} \left(\frac{S_0^2}{S_0^1}\right)^{\omega} \\
&= e^{-r(s-t)}\frac{S_0^1}{S_t^1}\mathbb{E}^{\mathbb{Q}}\left(e^{\left(\omega \ln \frac{S_s^2}{S_0^2} - (\omega-1)\ln \frac{S_s^1}{S_0^1}\right)} \mid \mathcal{F}_t\right)\left(\frac{S_0^2}{S_0^1}\right)^{\omega} \\
&= e^{-r(s-t)}\frac{S_0^1}{S_t^1}\mathbb{E}^{\mathbb{Q}}\left(e^{\left(\omega X_s^2 + (1-\omega)X_s^1\right)} \mid \mathcal{F}_t\right)\left(\frac{S_0^2}{S_0^1}\right)^{\omega}.
\end{aligned}$$

Since proposition 3.1 provides the mgf of $\left(X_s^1, X_s^2\right)$, we immediately infer that

$$\mathbb{E}^{\mathbb{Q}}\left(e^{(1-\omega)X_s^1 + \omega X_s^2} \mid \mathcal{F}_t\right) = \left(\frac{S_t^1}{S_0^1}\right)^{1-\omega}\left(\frac{S_t^2}{S_0^2}\right)^{\omega} \exp\left(A(1-\omega,\omega,t,s,\delta_t)\right), \tag{51}$$

and finally,

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{S_1}}\left(e^{\omega \ln \frac{S_s^2}{S_s^1}} \mid \mathcal{F}_t\right) &= e^{-r(s-t)}\frac{S_0^1}{S_t^1}\left(\frac{S_t^1}{S_0^1}\right)^{1-\omega}\left(\frac{S_t^2}{S_0^2}\right)^{\omega}\left(\frac{S_0^2}{S_0^1}\right)^{\omega} \exp\left(A(1-\omega,\omega,t,s,\delta_t)\right) \\
&= e^{-r(s-t)}\left(\frac{S_t^2}{S_t^1}\right)^{\omega} \exp\left(A(1-\omega,\omega,t,s,\delta_t)\right)
\end{aligned}$$

which allows us to conclude. ∎

In the remainder of this section, we denote the log-return of $\frac{S_T^2}{S_T^1}$ by $X_T^{21} = \ln \frac{S_T^2/S_0^2}{S_T^1/S_0^1}$. This log-return is such that $\frac{S_T^2}{S_T^1} = \frac{S_0^2}{S_0^1}e^{X_T^{21}}$. If the density at time $t \leq T$ of the log-return $X_T^{21}|\mathcal{F}_t$ under the measure $\mathbb{Q}^{S_1}$, is denoted by $f_{t,T}(x,\delta_t)$ then the value of the exchange option is equal:

$$\begin{aligned}
C(t,\delta_t) &= S_t^1\mathbb{E}^{\mathbb{Q}^{S_1}}\left(\left(\frac{S_0^2}{S_0^1}e^{X_T^{21}} - \gamma\right)_+ \mid \mathcal{F}_t\right) \\
&= \frac{S_t^1 S_0^2}{S_0^1}\mathbb{E}^{\mathbb{Q}^{S_1}}\left(\left(e^{X_T^{21}} - \gamma\frac{S_0^1}{S_0^2}\right)_+ \mid \mathcal{F}_t\right) \\
&= \frac{S_t^1 S_0^2}{S_0^1}\int_{\ln \gamma \frac{S_0^1}{S_0^2}}^{+\infty}\left(e^x - \gamma\frac{S_0^1}{S_0^2}\right) f_{t,T}(x,\delta_t)\, dx\,,
\end{aligned}$$

where $r \in \mathbb{R}^+$ is assumed to be a constant risk free rate.

Proposition 6.2. *Let M be the number of steps used in the Discrete Fourier Transform (DFT) and $\Delta_x = \frac{2x_{max}}{M-1}$ be this step of discretization. Let us denote $\Delta_z = \frac{2\pi}{M\,\Delta_x}$ and*

$$z_j = (j-1)\Delta_z,$$

for $j = 1...M$.
Let $\psi_{t,T}^{21}$ be the mgf of $X_T^{21}|\mathcal{F}_t$, under $\mathbb{Q}^{S_1}$, namely $\psi_{t,T}^{21}(i\,z_j) = \mathbb{E}^{\mathbb{Q}^{S_1}}\left(e^{i\,z_j X_T^{21}}|\mathcal{F}_t\right)$, such as presented in proposition 6.1, with $i = \sqrt{-1}$.

The values of $f_{t,T}(x,\delta_t)$ at points $x_k = -\frac{M}{2}\Delta_x + (k-1)\Delta_x$ are approached by the sum:

$$f_{t,T}(x_k,\delta_t) \approx \frac{2}{M\,\Delta_x} Re\left(\sum_{j=1}^{M} \Upsilon_j \psi_{t,T}^{21}(i\,z_j)\,(-1)^{j-1} e^{-i\frac{2\pi}{M}(j-1)(k-1)}\right). \tag{52}$$

where $\Upsilon_j = \frac{1}{2}1_{\{j=1\}} + 1_{\{j\neq 1\}}$.

Proof. The density of $X_T^{21}|\mathcal{F}_t$ is retrieved by calculating the Fourier transform of $\psi_{t,T}^{21}(iz)$ as follows:

$$\begin{aligned} f_{t,T}(x_k,\delta_t) &= \frac{1}{2\pi}\mathcal{F}[\psi_{t,T}^{21}(iz)](x_k) \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \psi_{t,T}^{21}(iz)\, e^{-i\,x_k\,z} dz \\ &= \frac{1}{\pi} Re(\int_{0}^{+\infty} \psi_{t,T}^{21}(iz) e^{-i\,x_k\,z} dz) \end{aligned}$$

where the last equality comes from the fact that $\psi_{t,T}^{21}(z)$ and $\psi_{t,T}^{21}(-z)$ are complex conjugate. By using the points $x_k = -\frac{M}{2}\Delta_x + (k-1)\Delta_x$, this last integral can be approached with the trapezoid rule and this leads to the following estimate for $f_{t,T}(.)$:

$$\begin{aligned} f_{t,T}(x_k,\delta_t) &\approx \frac{1}{\pi} Re\left(\sum_{j=1}^{M} \Upsilon_j \psi_{t,T}^{21}(iz_j) e^{-i\,x_k z_j}\Delta_z\right) \\ &\approx \frac{1}{\pi} Re\left(\sum_{j=1}^{M} \Upsilon_j \psi_{t,T}^{21}(iz_j)(-1)^{j-1} e^{-i\frac{2\pi}{M}(j-1)(k-1)}\Delta_z\right). \end{aligned}$$

□

We conclude this section with a numerical illustration. We evaluate European exchange options of maturities one and six months. We assume that $S_0^1 = S_0^2 = 100$. The number of DFT steps is chosen to be equal to $M = 2^8$. The risk free rate is set to $r = 2\%$. The other parameters are assumed to be those as presented in table 2. The upper graphs of figure 4 show the probability density functions of $\frac{S_T^2}{S_T^1}$ computed numerically for T equal one and six months. As we could expect, the variance of these densities is proportional to the intensities of the jump processes. In a Gaussian framework, exchange options are evaluated analytically with the Margrabe's formula reported in appendix 1. As in the Black & Scholes formula, the key parameter is the volatility of $\frac{S_T^2}{S_T^1}$ under $\mathbb{Q}^{S_1}$. If we ignore the jumps, this volatility is constant and equal to $\sqrt{\left((\sigma^{21}-\sigma^{11})^2 + (\sigma^{22})^2\right)} = 13.86\%$ based on the estimates in table 2. The lower graphs of figure 4 present the implied volatilities of $\frac{S_T^2}{S_T^1}$, matching option values obtained with our model to Margrabe's prices. For short-term options and when the intensities $(\lambda_t^1, \lambda_t^2)$ are low, the smile of volatilities is a convex function of γ. For longer maturities, the smile of the volatilities become strictly decreasing. These graphs confirm that the self and mutual excitation of jumps can partly explain the smile of volatilities observed for exchange options.

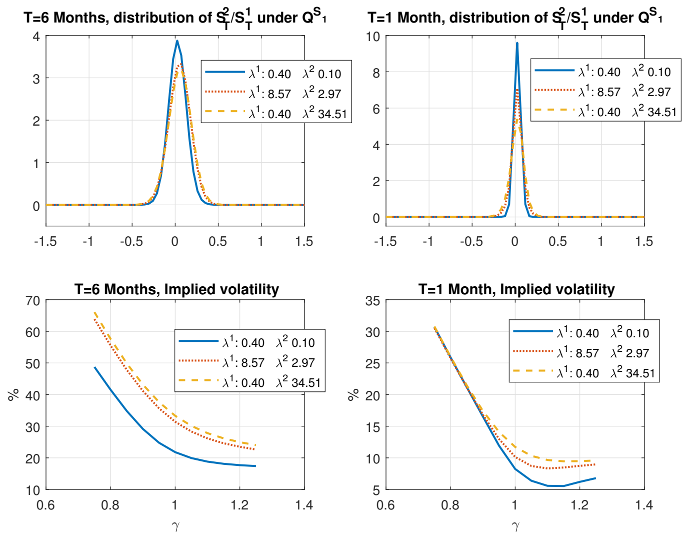

Figure 4: These graphs report the densities of $\frac{S_T^2}{S_T^1}$ and implied volatilities of exchange options for different values of γ and levels of excitation. Two maturities are considered: 1 and 6 months.

7 Conclusions

The Bivariate Mutually-Excited Switching Jump Diffusion (BMESJD) is an alternative to Hawkes processes, capable to explain the mutual-contagion between shocks observed in financial time series. Jumps of prices are triggered by changes of regime of a hidden Markov chain. The excitation mechanism is embedded in the matrix of instantaneous transition probabilities. The BMESJD presents several interesting features. Firstly, contrary to a bivariate Hawkes Diffusion, simultaneous jumps may be observed when financial markets are under pressure. Secondly, the BMESJD belongs to the family of switching processes but the model is parsimonious. Thirdly, the BMESJD is easy to calibrate when jumps are double exponential random variables. We establish in this particular case the statistical distribution of the sum of a bivariate normal and exponential random variable. Next, we propose an enhanced version of the Hamiton filter to estimate parameters. Finally, the set of equivalent risk measures is explicitly defined.

As numerical illustration, we fit the BMESJD to S&P 500 and Euronext 100 time series, from 2005 to 2017. We observe that filtered intensities are good indicators of market stress and reach their highest values during periods of financial turmoil, like the credit crunch of 2008, or the European sovereign debts crisis. Finally, we use a change of numeraire in order to price exchange options. We show that the mutual excitation between two time series may explain the curvature of the smile of volatilities for such options.

Appendix: pricing of exchange options in a Brownian setting

In this appendix, we remind the formula for exchange option pricing when underlying assets are ruled by geometric Brownian motions,

$$\begin{pmatrix} \frac{dS_t^1}{S_{t-}^1} \\ \frac{dS_t^2}{S_{t-}^2} \end{pmatrix} = \begin{pmatrix} r \\ r \end{pmatrix} dt + \begin{pmatrix} \sigma^{11} & 0 \\ \sigma^{21} & \sigma^{22} \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix}, \tag{53}$$

under the risk neutral measure $\mathbb{Q}$. This formula was initially developped by Margrabe (1978) and we briefly redevelop it in our framework. We consider the following change of measure

$$\frac{d\mathbb{Q}^{S_t^1}}{d\mathbb{Q}} = \frac{S_T^1}{S_0^1}\frac{1}{e^{rT}} = \exp\left(-\int_0^T \frac{1}{2}\left(\sigma^{11}\right)^2 ds + \int_0^T \sigma^{11} dW_s^1\right),$$

from $\mathbb{Q}$ to $\mathbb{Q}^{S_1}$, a measure with S_t^1 as numeraire. According to the Girsanov theorem, the quantity

$$\left(W_t^{S_1,1} := W_t^1 - \sigma_t^{11} t\ ,\ W_t^2\right)^\top$$

is a standard Brownian motion under $\mathbb{Q}^{S_1}$. Then, we infer from this last relation that the pair $\left(S_t^1, S_t^2\right)$ is driven by the next dynamics under $\mathbb{Q}^{S_1}$:

$$\begin{pmatrix} \frac{dS_t^1}{S_{t-}^1} \\ \frac{dS_t^2}{S_{t-}^2} \end{pmatrix} = \begin{pmatrix} r + \left(\sigma^{11}\right)^2 \\ r + \sigma^{11}\sigma^{21} \end{pmatrix} dt + \begin{pmatrix} \sigma^{11} & 0 \\ \sigma^{21} & \sigma^{22} \end{pmatrix} \begin{pmatrix} dW_t^{S_1,1} \\ dW_t^2 \end{pmatrix}.$$

Applying Itô's lemma to the logarithm of asset prices leads to the following expressions:

$$\begin{aligned} d\ln S_t^1 &= \left(r + \frac{1}{2}\left(\sigma^{11}\right)^2\right) dt + \sigma^{11} dW_t^{S_1,1}, \\ d\ln S_t^2 &= \left(r + \sigma^{11}\sigma^{21} - \frac{1}{2}\left(\left(\sigma^{21}\right)^2 + \left(\sigma^{22}\right)^2\right)\right) dt + \sigma^{21} dW_t^{S_1,1} + \sigma^{22} dW_t^2, \end{aligned}$$

from which we infer by direct integration that the ratio $\frac{S_t^2}{S_t^1}$ is given under $\mathbb{Q}^{S_1}$ by

$$\begin{aligned} \frac{S_T^2}{S_T^1} = \frac{S_t^2}{S_t^1} \exp\Bigg(&\int_t^T \left(\sigma^{11}\sigma^{21} - \frac{1}{2}\left(\sigma^{11}\right)^2 - \frac{1}{2}\left(\sigma^{21}\right)^2 - \frac{1}{2}\left(\sigma^{22}\right)^2\right) ds \\ &+ \int_t^T \left(\sigma^{21} - \sigma^{11}\right) dW_s^{S_1,1} + \int_t^T \sigma^{22} dW_s^2\Bigg) \end{aligned}$$

and is a lorgnormal random variable. More precisely, $\ln \frac{S_T^2/S_t^2}{S_T^1/S_t^1}$ is a Gaussian random variable with mean and volatility respectively given by

$$\begin{aligned} \sigma^M(t,T) &= \sqrt{\left(\left(\sigma^{21} - \sigma^{11}\right)^2 + \left(\sigma^{22}\right)^2\right)(T-t)} \\ \mu^M(t,T) &= -\frac{1}{2}\left(-2\sigma^{11}\sigma^{21} + \left(\sigma^{11}\right)^2 + \left(\sigma^{21}\right)^2 + \left(\sigma^{22}\right)^2\right)(T-t) \\ &= -\frac{1}{2}\left(\sigma^M(t,T)\right)^2 \end{aligned}$$

The exchange option value is therefore given by

$$\begin{aligned} C(t,\delta_t) &= \mathbb{E}^{\mathbb{Q}}\left(e^{-r(T-t)}\left(S_T^2-\gamma S_T^1\right)_+ \mid \mathcal{F}_t\right) \\ &= S_t^1\mathbb{E}^{\mathbb{Q}^{S_1}}\left(\left(\frac{S_T^2}{S_T^1}-\gamma\right)_+ \mid \mathcal{F}_t\right) \\ &= S_t^2\Phi\left(-d_1\right)-\gamma S_t^1\Phi\left(-d_2\right) \end{aligned}$$

where d_1 and d_2 are defined as follows:

$$\begin{aligned} d_2 &= \frac{\ln\left(\frac{\gamma}{S_t^2/S_t^1}\right)+\frac{1}{2}\left(\sigma^M(t,T)\right)^2}{\sigma^M(t,T)}, \\ d_1 &= d_2-\sigma^M(t,T), \end{aligned}$$

and where $\Phi(.)$ is the cumulative distribution of a standard normal random variable. In this setting, we call the implied volatility, the volatility

$$\sigma^M(0,1)=\sqrt{\left(\left(\sigma^{21}-\sigma^{11}\right)^2+\left(\sigma^{22}\right)^2\right)}$$

such that prices computed with the Margrabe's formula match market prices.

References

[1] Aït-Sahalia Y., Cacho-Diaz J., Laeven R.J.A. 2015. Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117, 585-606.

[2] Boswijk P., Laeven R.J.A, Lalu A. 2015. Asset returns with self-exciting jumps: option pricing and estimation with a continuum of moments. *Working paper.*

[3] Carr P., Wu L. 2016. Leverage effect, volatility feedback, and self-exciting market disruptions. *Forthcoming in Journal of Financial and Quantitative Analysis.*

[4] Chen K., Poon S.-H. 2013. Variance swap premium under stochastic volatility and self-exciting jumps. *working paper* SSRN-id2200172.

[5] Chourdakis K. 2005. Switching Lévy models in continuous time: Finite distributions and option pricing. *working paper* SSRN: 838924.

[6] Dong Y., Yuen K.C., Wang G., Wu C. A reduced-form model for correlated defaults with regime-switching shot noise intensities. Methodology and Computing in Applied Probability. 18(2), 459-486.

[7] Embrechts P., Liniger T., Lu L., 2011, Multivariate Hawkes processes: an application to financial data. *Journal of Applied Probability.* 48 (A), 367–378.

[8] Ficura M. 2015. Modelling jump clustering in the four major foreign exchange rates using high-frequency returns and cross-exciting jump processes. *Procedia Economics and Finance.* 25 208 – 219

[9] Guidolin M., Timmermann A. 2008. International asset allocation under regime switching, Skew, and Kurtosis Preferences. *Review of Financial Studies.* 21 (2), 889-935.

[10] Hainaut D., Colwell D. 2016. A structural model for credit risk with switching processes and synchronous jumps. *The European Journal of Finance*. 22(11), 1040-1062.

[11] Hainaut D., Moraux F. 2017 a. A switching self-exciting jump diffusion process for stock prices. *UCL working paper*

[12] Hainaut D., Goutte S. 2017 b. A switching microstructure model for stock prices. *UCL working paper*

[13] Hainaut D., Deelstra G. 2017 c. A self-exciting switching jump diffusion (SESJD): properties, calibration and hitting time. *UCL working paper*

[14] Hainaut D., Macgilchrist R. 2012. Strategic asset allocation with switching dependence. *Annals of Finance*. 8(1), 75-96.

[15] Hamilton J.D. 1989. A new approach to the economic analysis of non stationary time series and the business cycle. *Econometrica*. 57(2), 357-384.

[16] Hawkes A. 1971. Sprectra of some mutually exciting point processes. *Journal of the Royal Statistical Society Series B*, 33, 438-443. DOI : 10.1093/biomet/58.1.83

[17] Honda, T., 2003. Optimal portfolio choice for unobservable and regime-switching mean returns. *Journal of Economic Dynamics and Control*. 28, 45-78.

[18] Margrabe W. 1978. The Value of an option to exchange one asset for another. *Journal of Finance.* 33(1), 177-186.

[19] Rasmussen J.G. 2013. Bayesian inference for Hawkes processes. Methodology and Computing in Applied Probability. 15(3), 623-642.

[20] Stabile G., Torrisi G. L. 2010. Risk processes with non-stationary Hawkes claims arrivals. Methodology and Computing in Applied Probability. 12(3), 623-642.