An intensity model for credit risk with switching Lévy processes.

Donatien Hainaut† Olivier Le Courtois‡

January 9, 2012

† ESC Rennes and CREST, France. Email: donatien.hainaut@ensae.fr

‡ EM Lyon, France. Email: lecourtois@em-lyon.com

Abstract

In this paper, we develop a switching regime version of the intensity model for credit risk pricing. The default event is specified by a Poisson process whose intensity is modeled by a switching Lévy process. This model presents multiple interesting features. Firstly, as Lévy processes encompass numerous jump processes, our models can duplicate sudden jumps observed in credit spreads. And due to the presence of jumps, probabilities don’t vanish at very short maturities, contrary to models based on Brownian dynamics. Furthermore, as parameters of the Lévy process are modulated by a hidden Markov process, our approach is well suited to model eventual changes of volatility trends in credit spreads, related to modifications of some unobservable economic factors.

Keywords. Regime-switching model, Markov chain, Lévy process.

1 Introduction.

Assessing correctly the credit risk is a matter of concerns for all institutional lenders. Actually, there exists two main approaches to price the risk of default. The first category of models is called structural and price a defaultable debt by an option theoretic approach wherein the debt raised is a call option on the firm’s value. Merton (1974) pioneered this approach by pricing debt assuming constant interest rate and volatility of firm’s asset. Intensity models are an efficient alternative to structural models. In the intensity models the time of default is modeled directly as the time of the first jump of a Poisson process with random intensity (a Cox Process). In this group of models a striking similarity to default-free interest rate modelling is found. The first models of this type were developed by Jarrow and Turnbull (1995), Madan and Unal (1998) and Duffie and Singleton (1999). Lando (1998) developed the Cox-process methodology with the iterated conditional expectations.

Most of mathematical credit risk models use a Brownian motion as source of uncertainty. It ensures a certain analytical tractability but has also serious drawbacks. In structural and intensity models, the credit spreads vanish for short term corporate bonds, which is not realistic. An efficient way to avoid this feature, is to replace the Brownian motion by a jump process, or more generally by a Lévy process. This approach has been investigated by Kluge (2005) in his PhD thesis or by Cariboni and Schoutens (2009).

Even if credit risk models based on Lévy processes represent a significant advance in research, they are still partly unsatisfactory. In particular, as mentioned in Maalaoui et al. (2010), there exist evidences that credit spreads exhibit changes of trends, that cannot be replicated by Lévy
processes. In particular, the volatility of the credit spread can suddenly switches from a low to a high level, after a rating downgrade or in phase of decline. A number of theoretical papers use regime switches to capture state dependent movements in credit spread dynamics. The contribution of this paper is to explore the ability of switching Lévy processes to model credit risk, in an intensity framework. Switching Lévy processes are Lévy processes whose parameters are modulated by a hidden Markov chain. For a survey of properties of this category of processes, we refer the reader to the working paper of Hainaut (2010).

The outline of this paper is as follows. In the first sections, we develop an affine intensity based model and defines the hidden Markov process modulating the intensity of default. Next, we briefly describe the switching Lévy process driving the default rate. The following sections develop a numerical method to assess the survival probabilities. Finally, after a review of switching versions of popular Lévy processes, we present the econometric procedure of calibration and fit them to historical default intensities of four companies. Finally, we test the ability of this family of models to replicate survival probabilities curves, bootstrapped from CDS quotes.

2 Intensity Model

For a given time horizon $T$, we consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]} P)$ on which the default time of the firm is modeled as a stopping time $\tau$ driven by a non negative intensity process $\lambda$. The filtrations of $\tau$ is denoted by $\mathcal{G}$. We also define $\mathcal{H}$ as the filtration carrying information about $\lambda$ and about all stochastic processes involved in its dynamics (see section 4), such that $\mathcal{F} = \mathcal{G} \vee \mathcal{H}$. The default time is the first jump of a Poisson process, denoted by $N$, whose intensity is $\lambda$. This latter quantity may be seen as the instantaneous failure rate. It is well-known (see for instance Bielecki and Rutkowski (2002)) that, conditionally on the path followed by the intensity until time $T > t$, the probability that a firm is still in activity at time $t$ is given by:

$$P(\tau \geq t \mid \mathcal{H}_T) = e^{-\int_0^t \lambda_s \, ds},$$

whilst the survival probability from time $u < t$ to time $t$ is given by:

$$P(\tau \geq t \mid \mathcal{F}_u) = 1_{\tau > u}E\left(e^{-\int_u^t \lambda_s \, ds \mid \mathcal{H}_u}\right).$$

Furthermore, the process defined for all $t$ by $N_t - \int_0^{t \wedge \tau} \lambda_s \, ds$ is a martingale under the considered measure.

In our approach, the intensity is stochastic and depends on the state of the economy. As mentioned in the introduction, there are evidences that credit spreads exhibit changes of trends that are directly related to the evolution of unobservable economic factors. In periods of economic recession, defaults are more likely and the volatility of credit spreads can be important. However, in periods of economic growth, spreads are smaller and less volatile. To model this phenomenon, we assume that the economic conjuncture can be categorized into a finite number of $N$ states. We consider a Markov process $\alpha$ that contains the information about the effective economic factors but that is not directly observable. This hidden process influences the dynamics of the failure rate. In the remainder of this work, we consider that $\lambda$ admits the following dynamics:

$$d\lambda_t = a(\lambda_t, \alpha_t) \, dt + dX_t^{\alpha_t},$$

where the process defined for all $t$ by $X_t^{\alpha_t}$ is a switching Lévy process (the features of this category of processes are detailed in the next section) and where $a(\lambda_t, \alpha_t)$ is a linear function of $\lambda_t$ given by the following expression:

$$a(\lambda_t, \alpha_t) = a_{1,\alpha_t} + a_2 \lambda_t.$$
Note that, if by convention \( \tilde{a}_2 = -a_2 \) and \( \tilde{a}_{1,\alpha_t} = \frac{a_{1,\alpha_t}}{a_2} \), the intensity can be rewritten as a switching mean-reverting process:

\[
    d\lambda_t = \tilde{a}_2 (\tilde{a}_{1,\alpha_t} - \lambda_t) dt + dX_t^{\alpha_t}.
\]

In this formulation, the long term mean of the failure rate \( a_{1,\alpha_t} \) and the source of uncertainty \( X_t^{\alpha_t} \) both depend on the state of the economy. However, the speed of mean reversion is assumed in our framework independent from \( \alpha_t \). This may be seen as an intrinsic feature of the firm.

### 3 The Markov Process

In this article, we model the source of noise \( X \) in the dynamics of default intensities by a Lévy process whose parameters depend on a certain state of a hidden Markov process. The state indicator, denoted by \( \alpha_t \), is a Markov process that is not directly observable. This approach allows us to model the eventual changes of trends exhibited by credit spreads.

Under the assumption that there exist \( N \) states, \( \alpha \) takes its values in the set \( \mathcal{N} = \{1, 2, \ldots, N\} \) and admits an intensity matrix \( Q \) whose elements, denoted by \( q_{i,j} \), satisfy the following conditions:

\[
q_{i,j} \geq 0 \quad \forall i \neq j \quad \sum_{j=1}^{N} q_{i,j} = 0 \quad \forall i \in \mathcal{N}.
\]

The transition probabilities (under the real measure) between any two times \( t \) and \( u \geq t \) are computed as the (matrix) exponential of \( Q \):

\[
P(t,u) = \exp(\mathbf{Q}(u-t)).
\]

The elements of the matrix \( P(t,u) \) are denoted by \( p_{i,j}(t,u) \) for all \( i, j \in \mathcal{N} \). Indeed, \( p_{i,j}(t,u) \) is the probability of jumping from state \( i \) at time \( t \) to state \( j \) at time \( u \):

\[
p_{i,j}(t,u) = P(\alpha_u = j | \alpha_t = i) \quad i, j \in \mathcal{N}.
\]

The probability of being in state \( i \) at time \( t \), denoted by \( p_i(t) \), can be expressed as a function of the initial probabilities \( p_{k=1..N}(0) \) at time \( t = 0 \) as follows:

\[
p_i(t) = P(\alpha_t = i) = \sum_{k=1}^{N} p_k(0)p_{k,i}(0,t) \quad \forall i \in \mathcal{N}.
\]

When the Markov process has been running for a sufficiently long period of time, it can be shown that this probability is independent from the initial state:

\[
\lim_{t \to +\infty} p_i(t) = p_i \quad \forall i \in \mathcal{N}.
\]

In this framework, we denote by \( \tau_i \) the random time at which the Markov chain \( \alpha \) changes of state for the \( i^{th} \) times.

Among the approaches chosen to model the Markov chain, we adopt the marked point process one for its simplicity. Following Landen (2000), we define a mark space \( E \) which includes all possible regime switches as:

\[
E = \{z = (i,j) : i \in \{1,\ldots,N\}, j \in \{1,\ldots,N\}, i \neq j\}.
\]

The \( \sigma \)-algebra generated by \( E \) is denoted by \( \mathcal{E} \). On \( \mathcal{E} \) we define a marked point process \( \mu(t,.) \). See Bremaud (1981) for an introduction to these processes. If \( A \) is a subset of \( E \), \( \mu(t,A) \) counts
the cumulative number of regime shifts that belong to $A$ during $(0; t]$. The compensator of $\mu(t, \cdot)$ is given by:

$$\gamma(dt, dz) = \sum_{i \neq j} q_{i,j} I(\alpha_t = i) \epsilon(i,j)(dz) dt,$$

where $I(.)$ is the indicator function and $\epsilon(i,j)$ denotes the Dirac measure at point $z = \{i, j\}$. The Markov process $\alpha$ is given by:

$$\alpha_t = \int_0^t \int_E \eta(z) \mu(ds, dz).$$

By definition, $\alpha$ is $\mathcal{E}$-adapted. Furthermore, if we define $q(t, z) = \mu(t, z) - \gamma(t, z)$, then:

$$M_t = \alpha_t - \int_0^t \int_E \eta(z) \gamma(ds, dz) = \int_0^t \int_E \eta(z) q(ds, dz),$$

is a local martingale under the real measure $P$.

4 Switching Lévy Processes

The dynamics of the failure rate is driven by a particular stochastic process $X^\alpha$ which is a Lévy process conditionally on the state of the economy $\alpha$. Recall that a Lévy process is a càdlàg stochastic process, continuous in probability, with independent and stationary increments. We refer the reader to Applebaum (2004) for a detailed presentation. A switching Lévy process may be seen as a piecewise Lévy process. By piecewise, we mean that the process $X^\alpha$ is a Lévy process characterized by a set of parameters that depend on the state of $\alpha$. If between two times $[\tau_1, \tau_2]$ of transition, the Markov chain $\alpha$ is in state $j$, the switching Lévy process is driven by the following SDE:

$$dX^\alpha_t = dX^j_t \quad \alpha_t = j \in \mathbb{N},$$

where each $X^j$ is a Lévy process defined on subfiltrations of $\mathcal{H}$ denoted by $\mathcal{H}^j$. Indeed, each $X^j$ can be split into three components (according to the Lévy-Itô decomposition): a deterministic drift of parameter $\beta_j$, a Brownian motion of unit-time variance $\sigma_j^2$, and a jump process given by $J_{X^j}(t, z)$. The intensity of the latter component is $\nu(j, z)$. This is the Lévy measure of $X$ in state $j$. By Lévy measure, we mean that the probability of observing $k$ jumps of size included in a set $B \subset \mathbb{R}$ between $[\tau_1, \tau_2]$ is given by:

$$P( J_{X^j}([\tau_1, \tau_2] \times B) = k) = e^{-\int_{\tau_1}^{\tau_2} \int_B \nu(j, dz)dt} \left( \int_{\tau_1}^{\tau_2} \int_B \nu(j, dz)dt \right)^k / k!. \quad (4.1)$$

If $W$ designates a standard Brownian motion, the Lévy-Itô decomposition of $X^j$ is given by:

$$dX^j_t = \beta_j dt + \sigma_j dW_t + \int_{|z| > 1} z J_{X^j}(dt, dz) + \int_{|z| \leq 1} z (J_{X^j}(dt, dz) - \nu(j, dz)dt). \quad (4.2)$$

The triplet $(\beta_j, \sigma_j, \nu(j, z))$ fully determines the characteristic function of $X^j$:

$$\phi^j_t(u) = \mathbb{E} \left( \exp \left( iuX^j_t \right) \mid \mathcal{H}_0 \right) = \exp \left( t \left( i \beta_j u - \frac{1}{2} \sigma_j^2 u^2 + \int_{\mathbb{R}} \left( e^{izu} - 1 - iuz \chi_{|z| \leq 1} \nu(j, dz) \right) \right) \right).$$
Hence, the dynamics of $\lambda$ can be rewritten as follows:

$$d\lambda_t = \left(a(\lambda_t, \alpha_t) + \beta_{\alpha_t}\right) dt + \sigma_{\alpha_t} dW_t + \int_{|z| > 1} z J X^{\alpha_t}(dt, dz) + \int_{|z| \leq 1} z (J X^{\alpha_t}(dt, dz) - \nu(\alpha_t, dz)) dt,$$

If we consider finite variation Lévy processes, such that:

$$\int_{|z| \leq 1} z \nu(j, dz) < +\infty \quad j = 1...N,$$

the dynamics of $\lambda$ can be simplified as follows:

$$d\lambda_t = \left(a(\lambda_t, \alpha_t) + \beta'_{\alpha_t}\right) dt + \sigma_{\alpha_t} dW_t + \int_{|z| \leq 1} z J X^{\alpha_t}(dt, dz),$$

where $\beta'_{\alpha_t} = \beta_{\alpha_t} - \int_{|z| \leq 1} z \nu(\alpha_t, dz)$.

We end this paragraph with a remark about filtrations. As mentioned earlier, $\mathcal{F}$ is the filtration on which the process $N$ and the intensity $\lambda$ are defined. We underline the fact that $\mathcal{F}$ is not the smallest filtration including both $\mathcal{E}$ and $\mathcal{H}$ simply because the process $\alpha$ (adapted to $\mathcal{E}$) is not visible. However, the relationship $\mathcal{F}_t = \mathcal{G}_t \lor \mathcal{H}_t \subset \mathcal{E}_t \lor \mathcal{G}_t \lor \mathcal{H}_t$ obviously holds for all $t$. This relationship will play an important role in the forthcoming developments.

## 5 Default Probabilities

This section presents a method to calculate corporate default probabilities when the dynamics of the failure rate is driven by a switching Lévy process. This is particularly useful for pricing defaultable claims such as defaultable zero coupon bonds. For example, let us assume that a company issues a zero coupon bond of nominal $L$ that is fully repaid at time $T$ if no default has occurred. If the firm goes bankrupt before this time horizon, only a fraction $R$ of the nominal is repaid. The pricing of this bond can be done in two steps. First, we determine the bond value under the assumption that both the asset value and state of the Markov chain are visible. These prices are denoted by $B(t, T)$ in the remainder of this work. The bond market value $B(t, T)$ is equal to:

$$B(t, T) = \mathbb{E}\left(e^{-r(T-t)} (L + RL 1_{\tau \leq T}) \mid \mathcal{F}_t\right),$$

or, by the tower property of conditional expectations, to:

$$B(t, T) = \mathbb{E}\left(\mathbb{E}\left(e^{-r(T-t)} (L 1_{\tau > T} + RL 1_{\tau \leq T}) \mid \mathcal{E}_t \lor \mathcal{G}_t \lor \mathcal{H}_t\right) \mid \mathcal{F}_t\right).$$

Then:

$$B(t, T) = \mathbb{E}\left(1_{\tau > t} \mathbb{E}\left(e^{-r(T-t)} (L 1_{\tau > T} + RL 1_{\tau \leq T}) \mid \mathcal{H}_t \lor \mathcal{E}_t\right) \mid \mathcal{F}_t\right).$$

Define:

$$B(t, T, \lambda_t, \alpha_t) = \mathbb{E}\left(e^{-r(T-t)} (L 1_{\tau > T} + RL 1_{\tau \leq T}) \mid \mathcal{H}_t \lor \mathcal{E}_t\right) = e^{-r(T-t)} L (R + (1 - R) P(t, T, \lambda_t, \alpha_t)),$$

where $P(t, T, \lambda_t, \alpha_t)$ is the survival probability from $t$ to $T$ for given $\lambda_t$ and $\alpha_t$:

$$P(t, T, \lambda_t, \alpha_t) = \mathbb{E}\left(e^{-\int_t^T \lambda_s ds} \mid \mathcal{H}_t \lor \mathcal{E}_t\right). \quad (5.1)$$
In addition, we have the following natural boundary conditions:

\[ P(T,T,\lambda_T,\alpha) = 1, \]

and:

\[ \lim_{\lambda_T \to +\infty} P(t,T,\lambda_t,\alpha) = 0. \]

**Proposition 5.1.** Let us denote \( Q \) the matrix of transition probabilities of the Markov process \( \alpha_t \) and \( F(t) \) the \( N \)–vector of:

\[ f(t,j) = -a_{1,j} B(t) + \psi_j(i B(t)). \] (5.2)

The survival probabilities \( P(t,T,\lambda,\alpha) \) are given by the following expression:

\[ P(t,T,\lambda,j) = \exp(A(t,j) - B(t)\lambda), \] (5.3)

where \( B(t) \) is a function of time:

\[ B(t) = \frac{1}{a_2} \left( e^{a_2(T-t)} - 1 \right), \] (5.4)

and where \( \tilde{A}(t) = [e^{A(t,1)}, \ldots, e^{A(t,N)}]' \) is a vector, solution of the ODE system:

\[ \frac{\partial \tilde{A}(t)}{\partial t} + (\text{diag}(F(t)) + Q) \tilde{A}(t) = 0, \] (5.5)

under the terminal boundary condition:

\[ \tilde{A}(T,j) = 1 \quad j = 1 \ldots N. \]

**Proof.** By definition of \( P(t,T,\lambda_t,\alpha_t) \), we have that for all \( u \geq t \):

\[ P(t,T,\lambda_t,\alpha_t) = \mathbb{E} \left( \mathbb{E} \left( e^{-\int_t^T \lambda_t ds} \mid \mathcal{H}_u \lor \mathcal{E}_u \right) \mid \mathcal{H}_t \lor \mathcal{E}_t \right), \]

yielding, thanks to the definition (5.1) of the quantity \( P \):

\[ P(t,T,\lambda_t,\alpha_t) = \mathbb{E} \left( e^{-\int_t^u \lambda_s ds} P(u,T,\lambda_u,\alpha_u) \mid \mathcal{H}_t \lor \mathcal{E}_t \right). \]

Then, by assuming enough regularity to allow one to take the limit within the expectation, the following limit converges to zero:

\[ \lim_{u \to t} \frac{\mathbb{E} \left( e^{-\int_t^u \lambda_s ds} P(u,T,\lambda_u,\alpha_u) \mid \mathcal{H}_t \lor \mathcal{E}_t \right) - P(t,T,\lambda_t,\alpha_t)}{u-t} = 0. \]

If we develop the exponential by its Taylor approximation of first order, we can rewrite this limit as:

\[ \lim_{u \to t} \frac{\mathbb{E} \left( P(u,T,\lambda_u,\alpha_u) \mid \mathcal{H}_t \lor \mathcal{E}_t \right) - P(t,T,\lambda_t,\alpha_t)}{u-t} = \lambda_t P(t,T,\lambda_t,\alpha_t). \]

The right hand term being calculable by the Itô formula for switching Lévy processes, we infer that \( P(t,T,\lambda,\alpha_t) \) is the solution of a system of partial integro-differential equations:

\[ \frac{\partial}{\partial t} P(t,T,\lambda,j) + \mathcal{L} P(t,T,\lambda,j) = \lambda P(t,T,\lambda,j) \quad j = 1 \ldots N, \] (5.6)

where \( \mathcal{L} P(t,T,\lambda,j) \) is the generator of the switching Lévy process:

\[ \mathcal{L} P(t,T,\lambda,j) = (a(\lambda,j) + \beta_j) \frac{\partial P}{\partial \lambda} + \frac{\sigma^2}{2} \frac{\partial^2 P}{\partial \lambda^2} + \sum_{k \neq j} q_{j,k} (P(t,T,\lambda,k) - P(t,T,\lambda,j)) \]

\[ + \int_{\mathbb{R} \setminus \{0\}} \left( P(t,T,\lambda + z,j) - P(t,T,\lambda,j) - z 1_{|z| \leq 1} \frac{\partial P}{\partial \lambda} \right) \nu(j,dz). \] (5.7)
As \( \sum_{k \neq j} q_{j,k} = -q_{j,j} \), this last expression is also equivalent to:

\[
\mathcal{L} P(t, T, \lambda, j) = (a(\lambda, j) + \beta_j) \frac{\partial P}{\partial \lambda} + \frac{\sigma_j^2}{2} \frac{\partial^2 P}{\partial \lambda^2} + \sum_{k=1}^{N} q_{j,k} P(t, T, \lambda, k) + \int_{\mathbb{R}\{0\}} \left( P(t, T, \lambda + z, j) - P(t, T, \lambda, j) - z 1_{|z| \leq 1} \frac{\partial P}{\partial \lambda} \right) \nu(j, dz).
\] (5.8)

If we try a solution of the form:

\[
P(t, T, \lambda, \alpha) = \exp(A(t, \alpha) - B(t) \lambda),
\]

we get the following expressions for the derivatives of \( P \):

\[
\frac{\partial P}{\partial t} = P(t, T, \lambda, j) \left( \frac{\partial A(t, j)}{\partial t} - \frac{\partial B(t)}{\partial t} \lambda \right),
\]

\[
\frac{\partial P}{\partial \lambda} = -P(t, T, \lambda, j) B(t),
\]

and:

\[
\frac{\partial^2 P}{\partial \lambda^2} = P(t, T, \lambda, j) B(t)^2.
\]

The other terms involved in equation (5.6) are:

\[
P(t, T, \lambda + z, k) = P(t, T, \lambda, j) e^{-B(t)z},
\]

and:

\[
P(t, T, \lambda, k) = P(t, T, \lambda, j) e^{A(t,k)-A(t,j)}.
\]

We infer from the previous relationships that equation (5.6) can be reformulated as follows:

\[
\frac{\partial A(t, j)}{\partial t} - \frac{\partial B(t)}{\partial t} \lambda - (a_{1,j} + \beta_j + a_2 \lambda) B(t) + \frac{1}{2} \sigma_j^2 B(t)^2 - \lambda + \sum_{k=1}^{N} q_{j,k} (\exp(A(t,k) - A(t,j))) + \int_{\mathbb{R}\{0\}} (e^{-B(t)z} - 1 + B(t) z 1_{|z| \leq 1}) \nu(j, dz) = 0.
\]

This equation can indeed be split into two ODEs. The first one groups all the terms that are multiplied by \( \lambda \):

\[
\frac{\partial B(t)}{\partial t} + a_2 B(t) = -1,
\]

with the terminal condition \( B(T) = 0 \). Solving, we obtain the expression of \( B(t) \):

\[
B(t) = \frac{1}{a_2} \left( e^{a_2(T-t)} - 1 \right).
\]

The second ODE is:

\[
\frac{\partial A(t, j)}{\partial t} - (a_{1,j} + \beta_j) B(t) + \frac{1}{2} \sigma_j^2 B(t)^2 + \sum_{k=1}^{N} q_{j,k} (\exp(A(t,k) - A(t,j))) + \int_{\mathbb{R}\{0\}} (e^{-B(t)z} - 1 + B(t) z 1_{|z| \leq 1}) \nu(j, dz) = 0 \quad \forall \ j = 1 \ldots N,
\] (5.9)
with the following boundary conditions:

\[ A(T, j) = 0 \quad j = 1\ldots N. \]

The integral term in equation (5.9) can be inferred from the characteristic function of the Lévy process. We know indeed that the characteristic function of \( X_j^t \) is given by:

\[
\phi^j(u) = \exp \left( t \left( i \beta_j u - \frac{1}{2} \sigma_j^2 u^2 + \int_{\mathbb{R}} (e^{iuz} - 1 - i uz I(|z| \leq 1)) \nu(j, dz) \right) \right),
\]

where \( \psi^j(u) \) is called characteristic exponent. Therefore:

\[
\hat{\mathcal{R}}(e^{iB(t)z} - 1 + B(t)z I(|z| \leq 1)) \nu(j, dz) = \psi^j(iB(t)) - i \beta_j B(t) - \frac{1}{2} \sigma_j^2 B(t)^2,
\]

and if we set \( u = iB(t) \), we get that:

\[
\hat{\mathcal{R}}(e^{-B(t)z} - 1 + B(t)z I(|z| \leq 1)) \nu(j, dz) = \psi^j(iB(t)) + \beta_j B(t) - \frac{1}{2} \sigma_j^2 B(t)^2.
\]

Equation (5.9) can therefore be rewritten as:

\[
\frac{\partial A(t, j)}{\partial t} - (a_{1,j} + \beta_j) B(t) + \frac{1}{2} \sigma_j^2 B(t)^2 + \sum_{k=1}^{N} q_{j,k} (\exp(A(t, k) - A(t, j)))
\]

\[
+ \psi^j(iB(t)) + \beta_j B(t) - \frac{1}{2} \sigma_j^2 B(t)^2 = 0 \quad \forall j = 1\ldots N,
\]

or, after simplifications, as:

\[
\frac{\partial A(t, j)}{\partial t} - a_{1,j} B(t) + \psi^j(iB(t)) + \sum_{k=1}^{N} q_{j,k} (\exp(A(t, k) - A(t, j))) = 0 \quad \forall j = 1\ldots N.
\]

Choosing the convention:

\[ f(t, j) = -a_{1,j} B(t) + \psi^j(iB(t)), \]

we can rewrite equation (5.11) as follows:

\[
\frac{\partial A(t, j)}{\partial t} e^{A(t, j)} + f(t, j) e^{A(t, j)} + \sum_{k=1}^{N} q_{j,k} e^{A(t, k)} = 0 \quad \forall j = 1\ldots N,
\]

or equivalently as:

\[
\frac{\partial e^{A(t, j)}}{\partial t} + f(t, j) e^{A(t, j)} + \sum_{k=1}^{N} q_{j,k} e^{A(t, k)} = 0 \quad \forall j = 1\ldots N.
\]

If we define \( \hat{A}(t, j) = e^{A(t, j)} \), the equation can finally be put in matrix form as:

\[
\frac{\partial \hat{A}(t)}{\partial t} + (\text{diag}(F(t)) + Q) \hat{A}(t) = 0,
\]

where \( \text{diag}(F(t)) \) is the diagonal matrix whose components are those of \( F(t) \), and under the boundary condition \( \hat{A}(T, j) = 1 \quad j = 1\ldots N \).

In this article, we solve the system of equations (5.13) numerically by Euler’s method. See Appendix A for an alternative. The Markov process modulating the parameters of the intensity
being hidden, the probability of survival from \( t \) to \( T \), given a certain \( \lambda \) is calculated as a weighted sum:

\[
P(t, T, \lambda) = \mathbb{E}(P(t, T, \lambda, \alpha_t) \mid \mathcal{F}_t)
\]

\[
= 1_{T > t} \sum_{j=1}^{N} p_j(t) P(t, T, \lambda),
\]

where \( p_j(t) \) is the probability of being in state \( N \) at time \( t \). In the following numerical applications, the \( p_j(t) \)'s are replaced by stationary probabilities \( p_i \) as defined by equation (3.5). Finally, we can infer from the previous proposition, the following corollary:

**Corollary 5.2.** The survival probabilities \( P(t, T, \lambda_t, \alpha_t) \) are driven by the following stochastic differential equation on the enlarged filtration \( \mathcal{H}_t \vee \mathcal{E}_t \):

\[
dP(t, T, \lambda_t, \alpha_t) = \lambda_t P(t, T, \lambda_t, \alpha_t) dt - \sigma_{\alpha_t} B(t) P(t, T, \lambda_t, \alpha_t) dW_t
\]

\[
+ \int_{\mathbb{R} \setminus \{0\}} \left( P(t, T, \lambda + z, \alpha_t) - P(t, T, \lambda, \alpha_t) - z 1_{|z| \leq 1} \frac{\partial P}{\partial \lambda} \right) (dJ_{X_{\alpha_t}} - \nu(\alpha_t, dz) dt)
\]

\[
+ \int_{E} (P(t, T, \lambda, \alpha_t + \eta(z)) - P(t, T, \lambda, \alpha_t)) \mu(ds, dz)
\]

\[
+ \frac{1}{2} \sigma_{\alpha_t} ^2 \int_{\mathbb{R} \setminus \{0\}} \frac{\partial^2 P}{\partial \lambda^2} (dJ_{X_{\alpha_t}})
\]

and

\[
\mathbb{E}(dP(t, T, \lambda_t, \alpha_t) \mid \mathcal{H}_t \vee \mathcal{E}_t) = \lambda_t P(t, T, \lambda_t, \alpha_t) dt
\]

**Proof.** The proof is a direct consequence of relations (5.3), (5.6), and of Itô’s lemma, which states that:

\[
dP = \left[ \frac{\partial P}{\partial t} + (\alpha(\lambda_t, \alpha_t) + \beta_{\alpha_t}) \frac{\partial P}{\partial \lambda} + \frac{1}{2} \sigma_{\alpha_t}^2 \frac{\partial^2 P}{\partial \lambda^2} \right] dt + \sigma_{\alpha_t} \frac{\partial P}{\partial \lambda} dW_t
\]

\[
+ \int_{\mathbb{R} \setminus \{0\}} \left( P(t, T, \lambda + z, \alpha_t) - P(t, T, \lambda, \alpha_t) - z 1_{|z| \leq 1} \frac{\partial P}{\partial \lambda} \right) dJ_{X_{\alpha_t}}
\]

\[
+ \int_{E} (P(t, T, \lambda, \alpha_t + \eta(z)) - P(t, T, \lambda, \alpha_t)) \mu(ds, dz)
\]

\[
+ \frac{1}{2} \sigma_{\alpha_t} ^2 \int_{\mathbb{R} \setminus \{0\}} \frac{\partial^2 P}{\partial \lambda^2} (dJ_{X_{\alpha_t}})
\]

\[
= \lambda_t P(t, T, \lambda_t, \alpha_t) dt
\]

6 **Lévy Processes**

This section presents four examples of switching Lévy processes which will be used in the following numerical applications. The first of these processes is the switching Brownian motion. If \( W^j \) denotes a Brownian motion, the instantaneous return of the asset value is ruled by the following SDE:

\[
dX^j_t = \theta_j dt + \sigma_j dW^j_t \quad \forall j \in \mathcal{N},
\]

and its characteristic exponent is equal to:

\[
\psi_j(u) = i \theta_j u - \frac{1}{2} \sigma_j^2 u^2 \quad \forall j \in \mathcal{N}.
\]

In numerical applications, we will set \( \theta_j = 0 \), given that \( \lambda \) already admits a drift term, so that \( \psi_j(iB(t)) = \frac{1}{2} \sigma_j^2 B(t)^2 \). Note that the trajectory of a switching Brownian motion is continuous.

The second process is directly inspired by the popular jump diffusion model developed by Kou (2002). The number of jumps observed in the asset return \( X^j \) is a Poisson process \( N^j \) whose
intensity is $\lambda_j$. The amplitude of jumps, denoted by $Z_j$, has a double exponential distribution. Upward or downward exponentially-distributed jumps are observed, with respective probabilities $p_j$ and $q_j = 1 - p_j$. The parameters of the jump distribution are denoted by $\eta_j^+$ and $\eta_j^-$. The density function of $Z_j$ is therefore:

$$f_{Z_j}(z) = p_je^{\eta_j^+z}1_{\{z \geq 0\}} + q_je^{\eta_j^-z}1_{\{z < 0\}}.$$ (6.1)

The dynamics of the asset evolution are given by the following SDE:

$$dX_t^j = \theta_j dt + \sigma_j dW_t^j + Z_j dN_t^j.$$ (6.2)

The Lévy measure of $X_t^j$ is in this case the product of the frequency of jumps and of the density function describing the amplitude of jumps. From there, the characteristic exponent can be deduced:

$$\psi_j(u) = i \theta_j u - \frac{\sigma_j^2 u^2}{2} + i u \lambda_j \left( \frac{p_j}{\eta_j^+ - i u} - \frac{q_j}{\eta_j^- + i u} \right).$$ (6.3)

As with the switching Brownian motion, we will set $\theta_j = 0$ in numerical applications (because $\theta_j$ is redundant with the drift term of $\lambda$), yielding:

$$\psi_j(iB(t)) = \frac{1}{2} \sigma_j^2 B(t)^2 - B(t) \lambda_j \left( \frac{p_j}{\eta_j^+ + B(t)} - \frac{q_j}{\eta_j^- - B(t)} \right).$$

Note that marginal distributions of the Kou process do not admit closed-form-expressions. As we will see in the next section, this prevents us from fitting it to observed past intensities by an econometric approach. The next two Lévy processes are subordinated Brownian motions. These are Brownian motions observed in a new time scale (sometimes called business time) given by $S$, which is an increasing positive stochastic process. In financial models based upon subordinated Brownian motions, each economic agent assumes that the instantaneous asset return is normal but that trading time is randomly distributed according to $S$, which is referred to as the subordinator. More detailed information about subordinated Brownian motions can be found in Cont and Tankov (2004) or Applebaum (2004).

The Variance Gamma process (VG), used in financial modelling by Madan and Seneta (1990), is a Brownian motion subordinated by a Gamma random variable. If $\theta_j$ and $\sigma_j$ are the drift and variance of the Brownian motion, $X_t^j$ is defined as follows:

$$dX_t^j = \theta_j dS_t^j + \sigma_j dW_{S_t^j} \quad \forall j \in \mathcal{N},$$ (6.4)

where $S_t^j \sim Gamma\left(\frac{1}{\kappa_j}, \frac{1}{\kappa_j}\right)$. In this case, the expectation and variance of $S_t^j$ are respectively equal to $t$ and $\kappa_j t$. The characteristic exponent of this process is

$$\psi_j(u) = -\frac{1}{\kappa_j} \log \left( 1 - i \theta_j \kappa_j u + \frac{1}{2} u^2 \kappa_j \sigma_j^2 \right) \quad \forall j \in \mathcal{N},$$ (6.5)

where, setting $\theta_j = 0$ as in the Kou and Brownian motion models:

$$\psi_j(iB(t)) = -\frac{1}{\kappa_j} \log \left( 1 - \frac{1}{2} B(t)^2 \kappa_j \sigma_j^2 \right) \quad \forall j \in \mathcal{N}.$$
asymmetry. We refer the interested reader to the paper by Barndorff-Nielsen (1998) for a detailed analysis of this process. In this setting, the characteristic function of $X^j$ is known analytically:

$$
\psi_j(u) = \frac{1}{\kappa_j} - \frac{1}{\kappa_j} \sqrt{1 - 2i\kappa_j \theta_j u + u^2 \sigma_j^2 \kappa_j},
$$

(6.6)

where, setting $\theta_j = 0$:

$$
\psi_j(iB(t)) = \frac{1}{\kappa_j} - \frac{1}{\kappa_j} \sqrt{1 - B(t)^2 \sigma_j^2 \kappa_j}.
$$

(6.7)

The Variance Gamma and the Normal Inverse Gaussian processes both have closed-form probability density functions. These expressions are presented in the next section.

7 Econometric Calibration

In this section, we report the results of our fit of mean-reverting switching Lévy processes to historical default intensities. For the sake of illustration, we considered four companies: Volvo, Banco Bilbao, BNP Paribas, and Mittal. Default intensities were inferred from daily quotes (in Euros) of credit default swaps (CDS), of maturity 6 months. The 6 Month CDS premiums have been retrieved from Reuters and run from the 17/12/2007 to the 8/11/2011. In exchange of a premium, expressed as a percentage of the principal, the default swap seller promises to make a payment in the event of default of a reference obligation, which is usually a bond or a loan. In case of default, the CDS pays an amount of money equal to one minus the recovery rate (which is the rate of the company’s debt that is redeemed to debtholders), times the principal. Usually, the recovery rate, noted $R$, is assumed to be 40%. There are 3 main seniorities/tiers: SECDOM Secured Debts, SNRFOR Senior Unsecured Debts, SUBLT2 Subordinated or Lower Tier2 Debts. As SNRFOR debts are the most actively traded, we considered CDS quotes on this category of debts. The CDS premium is calculated as the expected discounted cost of the claim. For a 6 month CDS, if the intensity of default is assumed constant, the premium at time $t$ is given by the product of the default probability times the discounted cost:

$$
CDS_{6M}(t) = \left(1 - e^{-\lambda_{6M}(t)\frac{1}{2}}\right) e^{-r_{6M}(t)\frac{1}{2}} (1 - R),
$$

where $r_{6M}(t)$ and $\lambda_{6M}(t)$ are respectively the 6 months interest rate (in our case, the 6M Euribor) and the intensity of default at time $t$. From this relationship, we can infer $\lambda_{6M}$, used as a proxy for the instantaneous intensity $\lambda$. Figure 7.1 presents the evolution of these intensities.

![Figure 7.1: Intensities of Default, from the 17/12/2007 to the 8/11/2011.](image)
We fitted to these four time series a discrete version of a mean reverting switching Lévy processes, such as introduced in section 2 equation (2.1):

\[ \lambda_{6M}(t + \Delta t) - \lambda_{6M}(t) = \tilde{a}_2 (\tilde{a}_{1,\alpha} - \lambda_{6M}(t)) \Delta t + \Delta X_i^{\alpha_i}, \]

where \( \alpha \) is a 2 states Markov chain \( \mathcal{N} = \{1, 2\} \) whose daily transition matrix shall be denoted by \( P = p_{ij}(t, t + \Delta t) \) in the remainder of this section. Thus, we set \( \Delta t = 1/250 \). The state of \( \alpha \) is not directly observable, but the filtering technique developed by Hamilton (1989) and inspired by the Kalman filter (1960) allows us to retrieve the probabilities of being in a state given previous observations. We briefly summarize this filter. Let us define the probabilities of presence in state \( j \) as:

\[ \Pi_i^j = P(\alpha_t = j | \lambda_{6M}(t), \ldots, \lambda_{6M}(1)). \]

Hamilton proved that the vector \( \Pi_t = (\Pi_t^j)_{j=1}^m \) can be calculated as a function of the probabilities of presence during the previous period. If we denote by \( f_{\lambda}(t, \lambda_{6M}(t)) \) the vector of probability densities of \( \lambda_{6M}(t) \), in state 1 and 2, the vector of presence probabilities is given by:

\[ \Pi_t^j = \frac{f_{\lambda}(t, \lambda_{6M}(t)) \ast (\Pi_{t-\Delta t}^j P)}{\langle f_{\lambda}(t, \lambda_{6M}(t)) \ast (\Pi_{t-\Delta t}^j P), 1 \rangle}, \quad (7.1) \]

where \( 1 = (1, \ldots, 1) \in \mathbb{R}^d \) and \( x \ast y \) is the Hadamard product \( (x_1 y_1, \ldots, x_d y_d) \). To start the recursion, we assume that the Markov processes have reached their stable distribution. \( \Pi_0 \) is then set to the ergodic distribution, which is the eigenvector of the matrix \( P \), coupled to the eigenvalue equal to 1. If the intensity process is observed on \( T \) days, the loglikelihood is:

\[ \ln L(\lambda_{6M}(1) \ldots \lambda_{6M}(T)) = \sum_{t=1}^T \ln \langle f_{\lambda}(t, \lambda_{6M}(t)), (\Pi_{t-\Delta t}^j P) \rangle. \quad (7.2) \]

The most likely parameters are obtained by numerical maximization of (7.2). The Hamilton filter requires a closed form expression for the density of \( \lambda_{6M}(t) \). As the Kou process does not have an analytical expression for its marginal distributions, we limit our study to Brownian, Variance Gamma and Normal Inverse Gaussian distributions, for \( \Delta X_i^{\alpha_i} \). To simplify future calculations, we define the random variable as follows, when \( \alpha_t = j \):

\[ Y_t = \lambda_{6M}(t) - \lambda_{6M}(t - \Delta t) - \tilde{a}_2 (\tilde{a}_{1,j} - \lambda_{6M}(t - \Delta t)) \Delta t. \quad (7.3) \]

This random variable has the same density function as \( \lambda_t \) \( f_{\lambda}(t, \lambda_{6M}(t)) = f_Y(t, y(t)) \). In the Brownian case, \( Y_t \) is distributed as a normal random variable \( \mathcal{N} \left( \theta_j, \sigma_j^2 \sqrt{\Delta t} \right) \). If the intensity is driven by a Variance Gamma process, the density of \( Y_t \) in state \( \alpha_t = j \) is given by

\[ f(t, y, j) = C_j |y|^{\Delta t - \frac{1}{2}} \exp(A_j y) K_{\Delta t - \frac{1}{2}}(B_j |y|), \quad (7.4) \]

where \( A_j, B_j \) and \( C_j \) are constants defined hereafter, and \( K_{\Delta t - \frac{1}{2}}(\cdot) \) is the modified Bessel function of the second kind. Indeed:

\[ A_j = \frac{\theta_j}{\sigma_j^2}, \quad B_j = \frac{1}{\sigma_j^2} \sqrt{\theta_j^2 + 2 \sigma_j^2}, \]

and:

\[ C_j = \left( \frac{\theta_j^2 \kappa_j + 2 \sigma_j^2}{\kappa_j} \right)^{\frac{1}{2}} \frac{2}{\Gamma \left( \frac{\Delta t}{\kappa_j} \right)} \frac{1}{\sqrt{2\pi} \sigma_j \kappa_j^{\frac{1}{2}}}. \]
Finally, if the intensity is driven by a Normal Inverse Gaussian process, the density of $Y_t$ (in state $\alpha_t = j$) is given by:

$$f(t, y, j) = C_j \frac{1}{\sqrt{\delta^2_j + y^2}} \exp(A_j y) K_1 \left( B_j \sqrt{\delta^2_j + y^2} \right),$$

(7.5)

where $A_j$, $\delta_j$, $B_j$ and $C_j$ are constants defined by:

$$A_j = \frac{\theta_j}{\sigma_j^2}, \quad \delta_j = \frac{\sigma_j \Delta_t}{\sqrt{\kappa_j}}, \quad B_j = \frac{1}{\sigma_j^2} \sqrt{\theta_j^2 + \delta_j^2},$$

and:

$$C_j = \Delta_t \frac{1}{\pi \sigma_j \sqrt{\kappa_j}} \sqrt{\theta_j^2 + \delta_j^2} \exp \left( \delta_j \sqrt{B_j^2 - A_j^2} \right).$$

Note that, as mentioned in the previous section, we set $\theta_j = 0$ in numerical applications for the Brownian, VG and NIG dynamics, given that it is redundant with the drift term of $\lambda$. The standard deviations of the VG and NIG processes are equal by construction to $\sigma_j = 1.2 \sqrt{\Delta t}$. The parameters $\kappa_j = 1.2$ control the skew and the kurtosis of the process. Table 7.1 compares loglikelihoods of mean reverting switching processes. Working with VG or NIG processes clearly improves the quality of the fit.

<table>
<thead>
<tr>
<th>Brownian</th>
<th>VG</th>
<th>NIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNP Paribas</td>
<td>5 031</td>
<td>6 643</td>
</tr>
<tr>
<td>Volvo</td>
<td>4 586</td>
<td>6 706</td>
</tr>
<tr>
<td>Banco Bilbao</td>
<td>5 372</td>
<td>6 487</td>
</tr>
<tr>
<td>Mittal</td>
<td>4 111</td>
<td>6 160</td>
</tr>
</tbody>
</table>

Table 7.1: Loglikelihoods, 2 states models

All parameters are available in Appendix B. We see that the speed of mean reversion is quasi null with the Variance Gamma process. For the NIG process, the parameter $\kappa_j$ is small whatever the state. The filter identifies for each model a state in which the failure rate has a low volatility and a state in which the volatility is significantly higher. Exhibit 7.2 emphasizes the influence of the state of the Markov process on the shape of the one year probability density function of $X_{\Delta t}^{\alpha_t}$, involved in the dynamics of spreads in equation 6.4. The distribution plotted is a NIG and parameters are those obtained for Volvo. In state 1, the leptokurticity is accentuated and the default probability is higher than in state 2.

Figure 7.2: Distribution of $X_{\Delta t}^{\alpha_t}$, NIG model.
To conclude this paragraph, we draw the attention of the reader on the fact that the Hamilton filter also yields probabilities of sojourn in each state (the $\Pi^t_j$ such as defined by equation (7.1)). Figure 7.3 presents this for the 2D VG model, fitted to Volvo. This information could be used by traders to anticipate the evolution of CDS spreads (a similar approach has been developed in Hainaut and Maggilechrist, 2010). A daily fit of the model with the Hamilton filter will indeed reveal the probability of being in a period of high or low volatility for default intensities and help traders to take positions.

![Figure 7.3: Probability of being in state 1 and 2, NIG model.](image)

8 Application to Pricing

A default model should not only be justified from an econometric point of view but should also be able to replicate the curve of survival probabilities used by the market to price defaultable claims. If it is not the case, prices of defaultable claims computed with this model are not arbitrage free. This is why we test in this section the ability of the previously defined switching Lévy processes to fit survival probabilities extracted from CDS curves (source Reuters). We conduct this test for the four companies studied in the preceding section. Table 8.1 presents the CDS spreads in bps and the Euro swap curve on the 7/5/2011. The recovery rate chosen to bootstrap survival probabilities is set to 40%.

<table>
<thead>
<tr>
<th>CDS quote in bps</th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
<th>Eur Swap curves</th>
<th>Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>32.1</td>
<td>13.97</td>
<td>113.95</td>
<td>30.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>33.58</td>
<td>20.71</td>
<td>108.22</td>
<td>37.02</td>
<td>1</td>
<td>1.970%</td>
</tr>
<tr>
<td>2</td>
<td>51.88</td>
<td>40.36</td>
<td>140.21</td>
<td>93.19</td>
<td>2</td>
<td>2.307%</td>
</tr>
<tr>
<td>3</td>
<td>65.7</td>
<td>68.17</td>
<td>166.59</td>
<td>138.36</td>
<td>3</td>
<td>2.527%</td>
</tr>
<tr>
<td>4</td>
<td>85.21</td>
<td>81.21</td>
<td>188.98</td>
<td>169.99</td>
<td>4</td>
<td>2.742%</td>
</tr>
<tr>
<td>5</td>
<td>99.08</td>
<td>103.22</td>
<td>215.72</td>
<td>192.34</td>
<td>5</td>
<td>2.913%</td>
</tr>
<tr>
<td>7</td>
<td>108.5</td>
<td>120.11</td>
<td>221.68</td>
<td>210.24</td>
<td>7</td>
<td>3.169%</td>
</tr>
<tr>
<td>10</td>
<td>115.35</td>
<td>135.26</td>
<td>229.38</td>
<td>225.53</td>
<td>10</td>
<td>3.447%</td>
</tr>
<tr>
<td>20</td>
<td>122.47</td>
<td>140.02</td>
<td>232.08</td>
<td>244.48</td>
<td>20</td>
<td>3.828%</td>
</tr>
</tbody>
</table>

Table 8.1: CDS quotes, 17/5/11

Figure 8.1 presents survival probabilities inferred from CDS quotes (detailed figures are provided in table 8.2). According to market data, Mittal and Banco Bilbao can go bankrupt with a high probability, compared with Volvo and BNP Paribas. CDS quotes were linearly interpolated for missing maturities. From these quotes, we bootstrapped 20 default probabilities.
Next, we fitted switching Lévy models to survival probabilities. Having at our disposal only 20 survival probabilities, we limited the number of states of the Markov chain to $N = 2$. The mean error after calibration, defined as:

$$\epsilon = \frac{1}{20} \sum_{i=1}^{20} (\text{Modeled } DP(i) - \text{Market } DP(i))^2,$$

Table 8.2: Estimated survival probabilities

<table>
<thead>
<tr>
<th>Maturity</th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9943</td>
<td>0.9865</td>
<td>0.9820</td>
<td>0.9938</td>
</tr>
<tr>
<td>2</td>
<td>0.9825</td>
<td>0.9863</td>
<td>0.9534</td>
<td>0.9687</td>
</tr>
<tr>
<td>3</td>
<td>0.9667</td>
<td>0.9652</td>
<td>0.9169</td>
<td>0.9300</td>
</tr>
<tr>
<td>4</td>
<td>0.9419</td>
<td>0.9446</td>
<td>0.8739</td>
<td>0.8850</td>
</tr>
<tr>
<td>5</td>
<td>0.9151</td>
<td>0.9110</td>
<td>0.8187</td>
<td>0.8366</td>
</tr>
<tr>
<td>6</td>
<td>0.8933</td>
<td>0.8843</td>
<td>0.7803</td>
<td>0.7950</td>
</tr>
<tr>
<td>7</td>
<td>0.8697</td>
<td>0.8542</td>
<td>0.7408</td>
<td>0.7497</td>
</tr>
<tr>
<td>8</td>
<td>0.8480</td>
<td>0.8262</td>
<td>0.7008</td>
<td>0.7071</td>
</tr>
<tr>
<td>9</td>
<td>0.8254</td>
<td>0.7938</td>
<td>0.6596</td>
<td>0.6620</td>
</tr>
<tr>
<td>10</td>
<td>0.8017</td>
<td>0.7631</td>
<td>0.6173</td>
<td>0.6145</td>
</tr>
<tr>
<td>11</td>
<td>0.7808</td>
<td>0.7396</td>
<td>0.5800</td>
<td>0.5733</td>
</tr>
<tr>
<td>12</td>
<td>0.7597</td>
<td>0.7160</td>
<td>0.5428</td>
<td>0.5310</td>
</tr>
<tr>
<td>13</td>
<td>0.7381</td>
<td>0.6920</td>
<td>0.5054</td>
<td>0.4874</td>
</tr>
<tr>
<td>14</td>
<td>0.7161</td>
<td>0.6679</td>
<td>0.4681</td>
<td>0.4427</td>
</tr>
<tr>
<td>15</td>
<td>0.6938</td>
<td>0.6434</td>
<td>0.4308</td>
<td>0.3964</td>
</tr>
<tr>
<td>16</td>
<td>0.6711</td>
<td>0.6188</td>
<td>0.3935</td>
<td>0.3485</td>
</tr>
<tr>
<td>17</td>
<td>0.6478</td>
<td>0.5936</td>
<td>0.3561</td>
<td>0.2984</td>
</tr>
<tr>
<td>18</td>
<td>0.6241</td>
<td>0.5683</td>
<td>0.3188</td>
<td>0.2457</td>
</tr>
<tr>
<td>19</td>
<td>0.5998</td>
<td>0.5426</td>
<td>0.2815</td>
<td>0.1894</td>
</tr>
<tr>
<td>20</td>
<td>0.5750</td>
<td>0.5165</td>
<td>0.2442</td>
<td>0.1278</td>
</tr>
</tbody>
</table>
is presented in table 8.3. Given that the Kou model is overparametrized (2 times 7 parameters), we assumed that $\eta_j^+ = \eta_j^-$. Even with this assumption, the calibration remains unstable and errors are high. For the considered curves, the most efficient model seems to be the Variance Gamma. The calibrated parameters are provided in Appendix C. For a given model, we note that they are well-behaved and consistent in such a way that all the parameters exhibit stability. Whatever the dynamics, the calibration procedure identifies a state with a low volatility and one with a high volatility. And for most of models, the probabilities of transition between states are low.

<table>
<thead>
<tr>
<th></th>
<th>Brownian</th>
<th>Kou</th>
<th>VG</th>
<th>NIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNP Paribas</td>
<td>0.0004</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0002</td>
</tr>
<tr>
<td>Volvo</td>
<td>0.0007</td>
<td>0.0024</td>
<td>0.0006</td>
<td>0.0009</td>
</tr>
<tr>
<td>Banco Bilbao</td>
<td>0.0004</td>
<td>0.0011</td>
<td>0.0004</td>
<td>0.0004</td>
</tr>
<tr>
<td>Mittal</td>
<td>0.0017</td>
<td>0.0033</td>
<td>0.0004</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

Table 8.3: Average errors, 2 states models

9 Conclusions

This paper explores an extension of the intensity model for credit risk pricing. The default event is specified by a Poisson process whose intensity is modeled by a switching Lévy process. A switching Lévy process is a Lévy process whose parameters are modulated by a hidden Markov process. This category of models is well suited to duplicate the change of credit spread dynamics, observed in markets. In this setting, we show that the probabilities of default can easily be retrieved by solving a system of ordinary differential equations.

Furthermore, if the probability density function of the Lévy process has a closed form expression, we can fit with the Hamilton filter the switching Lévy processes to historical time series. The performed econometric tests reveal that this category of models, and in particular a mean reverting 2D NIG processes, explains relatively well the evolution of past default intensities. Finally, it seems that an intensity model based on 2D VG or NIG processes is well suited for pricing purposes, given that they fit relatively well survival probabilities, bootstrapped from the CDS market.
Appendix A

We recall Equation (5.13):

\[
\frac{\partial \tilde{A}(t)}{\partial t} + \text{diag}(F(t)) \tilde{A}(t) + Q \tilde{A}(t) = 0
\]  

(9.1)

where \( \tilde{A}(t) = (\tilde{A}(t, 1), ..., \tilde{A}(t, N))' \), \( \text{diag}(F(t)) \) is a diagonal matrix components \( \{f(t, j)\}_{j=1}^{N} \), \( G(t) \) is a diagonal matrix of components \( \left( \int_{T}^{t} f(s, j) ds \right)_{j=1}^{N} \), and \( Q = (q_{j,k})_{j=1}^{N}, k=1...N \).

We modify Equation (9.1) as follows:

\[
e^{G(t)} \frac{\partial \tilde{A}(t)}{\partial t} + e^{G(t)} \text{diag}(F(t)) \tilde{A}(t) + e^{G(t)} Q \tilde{A}(t) = 0
\]  

(9.2)

so that:

\[
\frac{\partial e^{G(t)} \tilde{A}(t)}{\partial t} + e^{G(t)} Q \tilde{A}(t) = 0
\]

Note that \( e^{G(t)} \) and \( Q \) may not necessarily commute, preventing us from solving the above equation readily. However, it is possible to write:

\[
\frac{\partial e^{G(t)} \tilde{A}(t)}{\partial t} + e^{G(t)} Q e^{-G(t)} e^{G(t)} \tilde{A}(t) = 0
\]

Defining \( \tilde{V}(t) = e^{G(t)} \tilde{A}(t) \) and \( D(t) = e^{G(t)} Q e^{-G(t)} \), we can write:

\[
\frac{\partial \tilde{V}(t)}{\partial t} + D(t) \tilde{V}(t) = 0
\]

This equation does not admit a closed-form solution because \( D \) is a function of time. However, it admits a semi-closed-form formula, the so-called Magnus expansion. This corresponds to writing:

\[
\tilde{V}(t) = \tilde{V}(T) e^{-M(t)}
\]

where:

\[
M(t) = \sum_{k=1}^{+\infty} M_k(t)
\]

and where:

\[
M_1(t) = \int_{T}^{t} D(u) du
\]

\[
M_2(t) = \frac{1}{2} \int_{T}^{t} \int_{T}^{u} [D(u), D(v)] dudv
\]

with \( [D(u), D(v)] = D(u)D(v) - D(v)D(u) \), and:

\[
M_3(t) = \frac{1}{6} \int_{T}^{t} \int_{T}^{u} \int_{T}^{v} ([D(u), [D(v), D(w)]) + [D(w), [D(v), D(u)])] dudvdw
\]

and so on.
### Appendix B

#### Table 9.1: Parameters Brownian motion

<table>
<thead>
<tr>
<th></th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{12}(0, \Delta t)$</td>
<td>0.0348</td>
<td>0.0849</td>
<td>0.6463</td>
<td>0.0547</td>
</tr>
<tr>
<td>$p_{21}(0, \Delta t)$</td>
<td>0.1063</td>
<td>0.2082</td>
<td>0.2238</td>
<td>0.5116</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.0153</td>
<td>0.0152</td>
<td>0.0001</td>
<td>0.0373</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.0652</td>
<td>0.1577</td>
<td>0.0577</td>
<td>0.8388</td>
</tr>
<tr>
<td>$a_{1,1}$</td>
<td>0.0158</td>
<td>0.0001</td>
<td>0.0436</td>
<td>0.0081</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>0.0430</td>
<td>0.2881</td>
<td>2.0000</td>
<td>0.9371</td>
</tr>
<tr>
<td>$a_2$</td>
<td>4.7156</td>
<td>0.2861</td>
<td>0.0100</td>
<td>0.4746</td>
</tr>
</tbody>
</table>

#### Table 9.2: Parameters Variance Gamma

<table>
<thead>
<tr>
<th></th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{12}(0, \Delta t)$</td>
<td>0.1803</td>
<td>0.0398</td>
<td>0.2269</td>
<td>0.1995</td>
</tr>
<tr>
<td>$p_{21}(0, \Delta t)$</td>
<td>0.7468</td>
<td>0.0092</td>
<td>0.6414</td>
<td>0.4300</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.0416</td>
<td>0.0394</td>
<td>0.0667</td>
<td>0.0781</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.0001</td>
<td>0.1002</td>
<td>0.0001</td>
<td>0.1878</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>0.0100</td>
<td>0.0806</td>
<td>0.0100</td>
<td>0.0100</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>0.0655</td>
<td>0.0318</td>
<td>0.0547</td>
<td>0.0084</td>
</tr>
<tr>
<td>$a_{1,1}$</td>
<td>0.0000</td>
<td>0.0070</td>
<td>0.0909</td>
<td>0.0482</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>0.0186</td>
<td>0.0115</td>
<td>0.0199</td>
<td>0.0203</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

#### Table 9.3: Parameters Normal inverse Gaussian

<table>
<thead>
<tr>
<th></th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{12}(0, \Delta t)$</td>
<td>0.0373</td>
<td>0.2600</td>
<td>0.1032</td>
<td>0.0714</td>
</tr>
<tr>
<td>$p_{21}(0, \Delta t)$</td>
<td>0.0142</td>
<td>0.2786</td>
<td>0.1510</td>
<td>0.0308</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.0048</td>
<td>0.0077</td>
<td>0.0030</td>
<td>0.0121</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.0014</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0023</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>$a_{1,1}$</td>
<td>0.0663</td>
<td>0.1065</td>
<td>0.9731</td>
<td>0.0000</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>0.0178</td>
<td>0.0000</td>
<td>0.0475</td>
<td>0.0335</td>
</tr>
<tr>
<td>$a_2$</td>
<td>2.5188</td>
<td>0.0386</td>
<td>0.0431</td>
<td>0.0642</td>
</tr>
</tbody>
</table>
# Appendix C

<table>
<thead>
<tr>
<th></th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_2$</td>
<td>-0.2913</td>
<td>-0.0043</td>
<td>-0.2689</td>
<td>-0.1171</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>0.0000</td>
<td>0.0251</td>
<td>0.0059</td>
<td>0.0104</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>-0.0095</td>
<td>-0.0035</td>
<td>-0.0100</td>
<td>-0.0460</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>-0.0110</td>
<td>-0.0024</td>
<td>-0.0024</td>
<td>0.0664</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.0348</td>
<td>0.0122</td>
<td>0.0002</td>
<td>0.0376</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.0001</td>
<td>0.2003</td>
<td>0.0420</td>
<td>0.1588</td>
</tr>
<tr>
<td>$p_{11}(0,1)$</td>
<td>0.7996</td>
<td>0.9997</td>
<td>0.9732</td>
<td>0.9358</td>
</tr>
<tr>
<td>$p_{22}(0,1)$</td>
<td>0.9909</td>
<td>0.9967</td>
<td>0.9501</td>
<td>0.9998</td>
</tr>
</tbody>
</table>

Table 9.4: Parameters Brownian motion

<table>
<thead>
<tr>
<th></th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_2$</td>
<td>-0.3765</td>
<td>-0.0074</td>
<td>-0.7369</td>
<td>-0.0374</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>0.0199</td>
<td>0.0509</td>
<td>0.0109</td>
<td>0.0657</td>
</tr>
<tr>
<td>$a_{1,1}$</td>
<td>-0.0300</td>
<td>-0.0132</td>
<td>-0.0439</td>
<td>-0.0103</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>-0.0281</td>
<td>-0.0049</td>
<td>-0.0516</td>
<td>-0.0013</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.0156</td>
<td>0.0025</td>
<td>0.0278</td>
<td>0.0126</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.0236</td>
<td>0.0000</td>
<td>0.0371</td>
<td>0.0550</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.1041</td>
<td>0.0396</td>
<td>0.1372</td>
<td>0.1005</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.2614</td>
<td>0.0518</td>
<td>0.2620</td>
<td>0.3687</td>
</tr>
<tr>
<td>$p_1$</td>
<td>0.2716</td>
<td>0.2708</td>
<td>0.3403</td>
<td>0.2159</td>
</tr>
<tr>
<td>$p_2$</td>
<td>0.5244</td>
<td>0.4367</td>
<td>0.2631</td>
<td>0.1844</td>
</tr>
<tr>
<td>$\eta_1$</td>
<td>22.1097</td>
<td>22.0925</td>
<td>22.1176</td>
<td>21.9651</td>
</tr>
<tr>
<td>$\eta_2$</td>
<td>21.0373</td>
<td>21.0307</td>
<td>21.0115</td>
<td>20.9393</td>
</tr>
<tr>
<td>$p_{11}(0,1)$</td>
<td>0.9733</td>
<td>0.9909</td>
<td>0.9037</td>
<td>0.8297</td>
</tr>
<tr>
<td>$p_{22}(0,1)$</td>
<td>0.7524</td>
<td>0.9909</td>
<td>0.8042</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

Table 9.5: Parameters, Kou’s model.

<table>
<thead>
<tr>
<th></th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_2$</td>
<td>-0.3138</td>
<td>-2.9088</td>
<td>-1.9300</td>
<td>-0.3008</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>0.0456</td>
<td>0.0000</td>
<td>0.0013</td>
<td>0.0165</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>-0.0489</td>
<td>-0.3260</td>
<td>-0.2392</td>
<td>-0.0452</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>-0.0119</td>
<td>-0.1546</td>
<td>-0.0018</td>
<td>-0.0096</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>5.7660</td>
<td>5.6625</td>
<td>5.7014</td>
<td>5.7040</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.1378</td>
<td>0.4668</td>
<td>0.0157</td>
<td>0.0605</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>61.1049</td>
<td>57.6398</td>
<td>64.3565</td>
<td>187.5381</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>3.2855</td>
<td>3.2928</td>
<td>3.2947</td>
<td>3.3198</td>
</tr>
<tr>
<td>$p_{11}(0,1)$</td>
<td>0.8747</td>
<td>0.9824</td>
<td>0.9999</td>
<td>0.9999</td>
</tr>
<tr>
<td>$p_{22}(0,1)$</td>
<td>0.9966</td>
<td>0.9997</td>
<td>0.8367</td>
<td>0.6364</td>
</tr>
</tbody>
</table>

Table 9.6: Parameters Variance Gamma
<table>
<thead>
<tr>
<th></th>
<th>BNP Paribas</th>
<th>Volvo</th>
<th>Banco Bilbao</th>
<th>Mittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_2$</td>
<td>-0.3150</td>
<td>-3.1781</td>
<td>-0.9957</td>
<td>-4.6992</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>0.0038</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>$a_{1,1}$</td>
<td>-0.0028</td>
<td>-0.0952</td>
<td>-0.0305</td>
<td>-0.0000</td>
</tr>
<tr>
<td>$a_{1,2}$</td>
<td>-0.0048</td>
<td>-0.0023</td>
<td>-0.0061</td>
<td>-0.2392</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.0892</td>
<td>2.0697</td>
<td>0.6506</td>
<td>5.2050</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.0000</td>
<td>0.0379</td>
<td>0.0185</td>
<td>6.2577</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>25.2348</td>
<td>122.2019</td>
<td>222.1992</td>
<td>222.1901</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>3.1581</td>
<td>3.2656</td>
<td>3.2663</td>
<td>33.2658</td>
</tr>
<tr>
<td>$p_{11}(0,1)$</td>
<td>0.9159</td>
<td>0.9838</td>
<td>0.9907</td>
<td>0.9999</td>
</tr>
<tr>
<td>$p_{22}(0,1)$</td>
<td>0.9651</td>
<td>0.9996</td>
<td>0.8245</td>
<td>0.9851</td>
</tr>
</tbody>
</table>

Table 9.7: Parameters Normal Inverse Gaussian
References


