Multivariate option pricing models: some extensions of the $\alpha$VG model

Guillaume, Florence
Eurandom, Department of Mathematics, T.U.Eindhoven
P.O.Box 513
5600 MB Eindhoven, the Netherlands
E-mail: f.m.y.guillaume@tue.nl

1 Introduction

The use of a time-changed Brownian motion in finance was first proposed by Clark to model cotton future prices [8]. His pioneer work was motivated by the fact that the information flow directly affects the evolution of the price through time. More precisely, when the amount of available information is low, the trading is slow and the price process evolves slowly and the other way around. Since then the concept of business clock has been widely considered in the financial literature, first to model univariate stock price processes ([1], [5], [15]), before being extended to the multivariate setting. Madan and Senata [15] first proposed to subordinate a multivariate Brownian motion by an univariate Gamma time change. However, the uniqueness of the business clock makes impossible to capture independency of the stock log-returns. Hence, Semeraro [16] proposed the so-called $\alpha$VG model which rests on a multivariate subordinator process composed of the weighted sum of two independent Gamma processes: an idiosyncratic and a common component. Later, Luciano and Semeraro [14] extended the $\alpha$VG model to other Lévy distributions by considering other subordinators. This class of multivariate models was motivated by the empirical work of Lo and Wang [12] which gives evidence for the presence of a significant common component in the trading volume and by the study of Harris [9] which shows that the distribution of the information flow is not identical for all securities. In the original setting, Luciano and Semeraro imposed some restrictions on the subordinator parameters such that the subordinator follows the same distribution as its two components, leading to marginal log-return processes of a particular Lévy type. Under this restricted setting, the marginal characteristic functions become independent of the common subordinator setting which affects only the dependence structure of the asset log-returns. This might lead to two undesired features in practice. First, the risk-neutral calibration of the common subordinator parameters requires liquid multivariate derivative quotes which are often unavailable. Secondly, the variance and therefore the volatility of the asset log-returns turn out to be independent of the common subordinator setting. Since the volatility level is directly related to the trading activity, the conditions imposed on the time change parameters imply that the trading activity does not depend on the common component of the business clock, but only on the idiosyncratic one.

If the marginal class is not a desired feature, the model can first be extended by relaxing the constraints imposed on the subordinator parameters. The such obtained generalized $\alpha$VG model ([10]) belongs to the class of exponential Lévy models, although the particular underlying Lévy distribution is not known anymore. We will show that the marginal characteristic functions and consequently also the volatility of the asset log-returns then depend on both the idiosyncratic and common subordinator settings and more specifically on the whole set of parameters. Hence the generalized model is more in line with the empirical evidence of the presence of both a common and an individual component in the business clock and its calibration does not require the existence of actively traded multivariate derivatives anymore.

Nevertheless, given the Lévy margins, the original and generalized $\alpha$VG models are usually
not able to replicate quoted option prices in both the strike and time to maturity dimensions with enough precision, especially during investor’s fear periods ([10]). Hence, we will propose a new class of models, the so-called Sato two factor models and study the particular case of the VG Sato model. These models are obtained by replacing the Lévy time changed Brownian motions in the setting of Luciano and Semeraro by Sato time changed Brownian motions and lead to marginal characteristic functions of the Sato type.

2 The $\alpha$VG Lévy and Sato models

Under the $\alpha$VG models, the N-dimensional stock return is modelled by the exponential of a multivariate time-changed Brownian motion:

$$S_t = \begin{pmatrix} S^{(1)}_t \\ S^{(2)}_t \\ \vdots \\ S^{(N)}_t \end{pmatrix} = \begin{pmatrix} S^{(1)}_0 \exp\left((r-q_t)t + Y^{(1)}_t\right) \\ S^{(2)}_0 \exp\left((r-q_t)t + Y^{(2)}_t\right) \\ \vdots \\ S^{(N)}_0 \exp\left((r-q_t)t + Y^{(N)}_t\right) \end{pmatrix},$$

where $S^{(i)}_0$ is the spot price of the $i$th underlying, $r$ is the risk-free interest rate, $q_t$ denotes the dividend yield of the $i$th stock and $Y = \{Y_t, t \geq 0\}$ is a N-dimensional time-changed Brownian motion. More particularly, under the Lévy settings, the process $Y$ is given by

$$(1) \quad Y_t = \begin{pmatrix} Y^{(1)}_t \\ Y^{(2)}_t \\ \vdots \\ Y^{(N)}_t \end{pmatrix} = \begin{pmatrix} \theta_1 G^{(1)}_t + \sigma_1 W^{(1)}_{G^{(1)}_t} \\ \theta_2 G^{(2)}_t + \sigma_2 W^{(2)}_{G^{(2)}_t} \\ \vdots \\ \theta_N G^{(N)}_t + \sigma_N W^{(N)}_{G^{(N)}_t} \end{pmatrix},$$

where $W^{(i)}$, $i = 1, \ldots, N$ are independent standard Brownian motions and where the subordinators $G^{(i)}_t$'s are the weighted sum of two Gamma processes, one idiosyncratic and one common process:

$$G_t = \begin{pmatrix} G^{(1)}_t \\ G^{(2)}_t \\ \vdots \\ G^{(N)}_t \end{pmatrix} = \begin{pmatrix} X^{(1)}_t + \alpha_1 Z_t \\ X^{(2)}_t + \alpha_2 Z_t \\ \vdots \\ X^{(N)}_t + \alpha_N Z_t \end{pmatrix},$$

where $\alpha_i > 0$, $Z_1 \sim \text{Gamma}(c_1, c_2), c_1, c_2 > 0$ and $X^{(i)}_1 \sim \text{Gamma}(a_i, b_i)$, $a_i, b_i > 0$ are independent random variables and are independent on the $W^{(i)}$'s.

On the other hand, under the Sato settings, the log-asset returns are built by space scaling the time-changed Brownian motions taken at unit time:

$$(2) \quad Y_t = \begin{pmatrix} Y^{(1)}_t \\ Y^{(2)}_t \\ \vdots \\ Y^{(N)}_t \end{pmatrix} = \begin{pmatrix} \theta_1 t^\gamma_1 G^{(1)}_t + \sigma_1 t^\gamma_1 W^{(1)}_{G^{(1)}_t} \\ \theta_2 t^\gamma_2 G^{(2)}_t + \sigma_2 t^\gamma_2 W^{(2)}_{G^{(2)}_t} \\ \vdots \\ \theta_N t^\gamma_N G^{(N)}_t + \sigma_N t^\gamma_N W^{(N)}_{G^{(N)}_t} \end{pmatrix},$$
where

\[ G = \begin{pmatrix} G^{(1)} \\ G^{(2)} \\ \vdots \\ G^{(N)} \end{pmatrix} = \begin{pmatrix} X^{(1)} + \alpha_1 Z \\ X^{(2)} + \alpha_2 Z \\ \vdots \\ X^{(N)} + \alpha_N Z \end{pmatrix}. \]

- The Gamma process

The Gamma process is a Lévy process built on the Gamma distribution and is defined on the positive real line. The characteristic function of the Gamma distribution \( \text{Gamma}(a, b) \) with parameters \( a > 0, b > 0 \) is given by:

\[ \phi_{\text{Gamma}}(u; a, b) = \left( 1 - \frac{iu}{b} \right)^{-a}. \]

The Gamma process \( X = \{X_t, t \geq 0\} \) is a Lévy process such that \( X_t \) follows a \( \text{Gamma}(at, b) \) distribution. The Gamma distribution satisfies the following scaling property: if \( X \sim \text{Gamma}(a, b) \) then \( cX \sim \text{Gamma}(a, b/c), c > 0 \). Moreover, the sum of independent Gamma random variables with the same parameter \( b \) is also a Gamma random variable: if \( X_i \sim \text{Gamma}(a_i, b) \), \( i = 1, \ldots, N \) are \( N \) independent random variables then \( \sum_{i=1}^{N} X_i \sim \text{Gamma}(\sum_{i=1}^{N} a_i, b) \). The first four moments of the Gamma distribution are given in Table 1.

<table>
<thead>
<tr>
<th>Characteristics of the Gamma distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
</tr>
<tr>
<td>variance</td>
</tr>
<tr>
<td>skewness</td>
</tr>
<tr>
<td>kurtosis</td>
</tr>
</tbody>
</table>

- The Variance Gamma (VG) process

The Variance Gamma process is a Lévy process built on the Variance Gamma distribution. The characteristic function of the Variance Gamma distribution \( \text{VG}(\sigma, \nu, \theta) \) with parameters \( \sigma > 0, \nu > 0 \) and \( \theta \in \mathbb{R} \) is given by:

\[ \phi_{\text{VG}}(u; \sigma, \nu, \theta) = \left( 1 - iu\theta\nu + \frac{u^2\sigma^2\nu}{2} \right)^{-\frac{1}{\nu}}, \quad u \in \mathbb{R}. \]

The Variance Gamma process \( X = \{X_t, t \geq 0\} \) is a Lévy process such that \( X_t \) follows a \( \text{VG}(\sqrt{t}\sigma, \frac{\nu}{\sqrt{t}}, \theta t) \) distribution. The VG distribution satisfies the following scaling property: if \( X \sim \text{VG}(\sigma, \nu, \theta) \) then, for \( c > 0 \), \( cX \sim \text{VG}(c\sigma, \nu, c\theta) \). The first four moments of the VG distribution are given in Table 2. A parameter \( \theta \) equal to zero indicates a symmetric distribution around zero whereas negative and positive values of \( \theta \) lead to negative and positive skewness, respectively. The parameter \( \nu \) primarily controls the kurtosis (see Table 2).

The VG distribution is sometimes defined with respect to another parameter set \( \{C > 0, G > 0, M > 0\} \). This parametrization can be inferred from the \( \{\sigma, \nu, \theta\} \) parametrization by

\[
\begin{align*}
C &= \frac{1}{\nu} \\
G &= \left( \frac{\theta^2\nu^2}{4} + \frac{\sigma^2\nu}{2} - \frac{\theta\nu}{2} \right)^{-1} \\
M &= \left( \frac{\theta^2\nu^2}{4} + \frac{\sigma^2\nu}{2} + \frac{\theta\nu}{2} \right)^{-1}.
\end{align*}
\]
<table>
<thead>
<tr>
<th></th>
<th>VG(σ, ν, θ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>θ</td>
</tr>
<tr>
<td>variance</td>
<td>σ² + νθ²</td>
</tr>
<tr>
<td>skewness</td>
<td>\frac{θν(3σ²+2νθ²)}{(σ²+νθ²)\frac{3}{2}}</td>
</tr>
<tr>
<td>kurtosis</td>
<td>3\left(1 + 2ν - \frac{νσ²}{(σ²+νθ²)^2}\right)</td>
</tr>
</tbody>
</table>

Table 2: Characteristics of the Variance Gamma distribution.

The characteristic function then becomes

\[ \phi_{VG}(u; C, G, M) = \left( \frac{GM}{GM + (M - G)iu + u^2} \right)^C. \]

The VG(C, G, M) process can be expressed as the difference of two Gamma processes: \( X_t = G_t^{(1)} - G_t^{(2)} \) where \( G^{(1)} = \{G_t^{(1)}, t \geq 0\} \), \( G^{(1)} \sim \text{Gamma}(Ct, M) \) and \( G^{(2)} = \{G_t^{(2)}, t \geq 0\} \), \( G_t^{(2)} \sim \text{Gamma}(Ct, G) \) are two independent Gamma processes. The Lévy measure of the VG process is given by (see [4])

\[ \nu_{VG}(dx) = \begin{cases} \frac{C\exp(G x)}{|x|} dx & x < 0, \\ \frac{C\exp(-M x)}{|x|} dx & x > 0. \end{cases} \]

Hence, Variance Gamma processes are of finite variation and infinite activity.

Moreover, a VG(σ, ν, θ) process can be seen as a Gamma time-changed Brownian motion with drift:

\[ X_t^{VG} = θG_t + σW_t, \]

where \( G = \{G_t, t \geq 0\} \) is a Gamma process with parameters \( a = b = \frac{1}{ν} \) and \( W = \{W_t, t \geq 0\} \) is a standard Brownian motion independent of \( G \).

• The VG Sato process

**Definition 1.** The distribution of a random variable \( X \) is **self-decomposable** if, for any constant \( c \), \( 0 < c < 1 \), \( X \) has the same probability law as the sum of a down-scaled version of itself and an independent random variable \( X_c \):

\[ X \overset{d}{=} cX + X_c. \]

Self-decomposable distributions are a sub-class of infinitely divisible distributions with a Lévy-Khintchine representation of the form

\[ Ψ_X(u) = iγu - \frac{σ²}{2}u^2 + \int_{-∞}^{+∞} (\exp(iux) - 1 - iux1_{|x|<1}) \frac{h(x)}{|x|} dx \]

where \( h(x) \geq 0 \) is decreasing for positive \( x \) and increasing for negative \( x \). Hence, self-decomposable laws are necessarily of infinite activity.

A Sato process can be constructed from any self-decomposable distribution as follow: the probability law of the Sato process at time \( t \) is obtained by scaling the self-decomposable law of \( X \) at unit time (see [6]):

\[ X_t \overset{d}{=} tγX, \]
where \( \gamma \) is the self-similar exponent. Sato processes are thus processes with independent but
time inhomogeneous increments.

From Equation (5) and the Lévy density of the VG process (3), it is clear that the VG proba-
\[ \phi(8) \]
bility law at unit time is self-decomposable for all acceptable VG parameter sets \( \{\sigma, \nu, \theta\} \). The char-
\[ \phi(7) \]
acteristic function of the VG Sato process at time \( t \) is thus given by
\[ \phi_{\text{VG Sato}}(u, t; \sigma, \nu, \theta, \gamma) = \phi_{\text{VG}}(u, 1; t^\gamma \sigma, \nu, t^\gamma \theta) \]
\[ = \left(1 - iu^\theta t^\gamma + \frac{\sigma^2 \nu t^2 u^2}{2}\right)^{-\frac{1}{\nu}}. \]

3 Model characteristics

3.1 The Generalized \( \alpha \) VG model

The characteristic function of the process \( Y_t \) (1) is given by:
\[ \phi_Y(u, t) = E \left[ \exp(iu'Y_t) \right] = \prod_{i=1}^{N} \phi_{X^{(i)}_t} \left( u_i \theta_i + \frac{1}{2} \sigma_i^2 u_i^2, t \right) \]
\[ \phi_{Z_t} \left( \sum_{i=1}^{N} \alpha_i \left( u_i \theta_i + \frac{1}{2} \sigma_i^2 u_i^2 \right), t \right) . \]

Proof. We have
\[ E \left[ \exp(iu'Y_t) \right] = E \left[ \exp(iu'Y_t) | G_t^{(i)}, i = 1, \ldots, N \right] \]
\[ = E \left[ \prod_{i=1}^{N} \exp(iu_i Y^{(i)}_t) | G_t^{(i)}, i = 1, \ldots, N \right] . \]

Since \( \theta_i G_t^{(i)} + \sigma_i W_{G_t^{(i)}}^{(i)} | G_t^{(i)} \sim \text{Normal}(\theta_i G_t^{(i)}, \sigma_i^2 G_t^{(i)}) \) are independent and since \( \phi_{\text{Normal}}(u; \mu, \sigma^2) = \exp \left( iu\mu - \frac{1}{2} \sigma^2 u^2 \right) \), we have
\[ \phi_Y(u, t) = E \left[ \prod_{i=1}^{N} \exp(iu_i Y^{(i)}_t) | G_t^{(i)} \right] = E \left[ \prod_{i=1}^{n} \exp \left( i \left( u_i \theta_i + \frac{1}{2} \sigma_i^2 u_i^2 \right) \left( X^{(i)}_t + \alpha_i Z_t \right) \right) \right] . \]

Given the independence of the \( X^{(i)}_t \)'s, \( i = 1, \ldots, N \) and \( Z_t \), we finally obtain Equation (6).

The marginal characteristic functions are directly obtained from (6):
\[ \phi_{Y^{(i)}}(u, t) = E \left[ \exp(iuY^{(i)}_t) \right] = \left(1 - i \frac{u \theta_i + i \frac{1}{2} \sigma_i^2 u^2}{b_i} \right)^{-\frac{1}{\nu} \alpha_i t} \left(1 - i \frac{\alpha_i}{c_2} \left( u \theta_i + i \frac{1}{2} \sigma_i^2 u^2 \right) \right)^{-c_i t}. \]

From the marginal characteristic function (7), it is clear that each process \( Y^{(i)} \) = \( \{Y^{(i)}_t, t \geq 0\}, i = 1, \ldots, N \) is a Lévy process (although not necessarily VG) since the marginal characteristic function can be rewritten as \( \phi_{Y^{(i)}}(u, t) = (\phi_{Y^{(i)}}(u, 1))^t \).

The linear correlation between the processes \( Y^{(i)}_t \) and \( Y^{(j)}_t \) is time independent:
\[ \rho_{ij} = \frac{\text{Cov} \left( Y^{(i)}_t, Y^{(j)}_t \right)}{\sqrt{\text{Var} \left( Y^{(i)}_t \right) \text{Var} \left( Y^{(j)}_t \right)}}. \]
where
\[
\text{Cov} \left( Y_t^{(i)}, Y_t^{(j)} \right) = \theta_i \theta_j \alpha_i \alpha_j \frac{c_1}{c_2} t
\]
and
\[
\text{Var} \left[ Y_t^{(i)} \right] = \left( \theta_i^2 \left( \frac{a_i}{b_i} + \alpha_i^2 \frac{c_1}{c_2} \right) + \sigma_i^2 \left( \frac{a_i}{b_i} + \alpha_i \frac{c_1}{c_2} \right) \right) t.
\]

**Proof.** We have
\[
\text{Var} \left[ Y_t^{(i)} \right] = \text{Var} \left[ \theta_i G_t^{(i)} + \sigma_i W_t^{(i)} \right] = \theta_i^2 \text{Var} \left[ G_t^{(i)} \right] + \sigma_i^2 \text{Var} \left[ W_t^{(i)} \right]
\]
since the covariance term is equal to zero due to the zero expectation of the Brownian motion \( W_t^{(i)} \). Moreover,
\[
\text{Var} \left[ Y_t^{(i)} \right] = \theta_i^2 \left( \text{Var} \left[ X_t^{(i)} + \alpha_i Z_t \right] \right) + \sigma_i^2 \left( \mathbb{E} \left[ \text{Var} \left[ W_t^{(i)} \mid G_t^{(i)} \right] \right] + \text{Var} \left[ \mathbb{E} \left[ W_t^{(i)} \mid G_t^{(i)} \right] \right] \right)
\]
\[
= \theta_i^2 t \left( \frac{a_i}{b_i} + \alpha_i^2 \frac{c_1}{c_2} \right) + \sigma_i^2 \mathbb{E} \left[ G_t^{(i)} \right]
\]
\[
= \theta_i^2 t \left( \frac{a_i}{b_i} + \alpha_i^2 \frac{c_1}{c_2} \right) + \sigma_i^2 \mathbb{E} \left[ X_t^{(i)} + \alpha_i Z_t \right]
\]
\[
= \theta_i^2 t \left( \frac{a_i}{b_i} + \alpha_i^2 \frac{c_1}{c_2} \right) + \sigma_i^2 t \left( \frac{a_i}{b_i} + \alpha_i \frac{c_1}{c_2} \right).
\]

For the covariance, we have
\[
\text{Cov} \left( Y_t^{(i)}, Y_t^{(j)} \right) = \text{Cov} \left( \theta_i \left( X_t^{(i)} + \alpha_i Z_t \right), \theta_j \left( X_t^{(j)} + \alpha_j Z_t \right) \right)
\]
\[
= \theta_i \theta_j \alpha_i \alpha_j \text{Var} \left[ Z_t \right],
\]
where we used the property \( \text{Cov} \left( \sum_{i=1}^N A_i, \sum_{j=1}^M B_j \right) = \sum_{i=1}^N \sum_{j=1}^M \text{Cov} \left( A_i, B_j \right) \).

The parameter set of the generalized \( \alpha \)VG model is \( \{ \theta_i; \sigma_i; \alpha_i; a_i; b_i, i = 1, \ldots, N; c_1, c_2 \} \) leading to a number of parameters amounting to \( 5N + 2 \). However, from the space scaling property of a Gamma random variable, we clearly see that the model has one redundant parameter. Indeed, we can scale the parameter \( c_2 \) to 1 since multiplying \( c_2 \) by a constant \( c \) is equivalent to dividing the parameters \( \alpha_i \)'s by \( c \). Moreover, for the sake of coherence, we will impose that the business time \( G_t^{(i)} \) increases on average as the real time \( t \), i.e. we impose that \( \mathbb{E} \left[ G_t^{(i)} \right] = \left( \frac{a_i}{b_i} + \alpha_i \frac{c_1}{c_2} \right) t = t \) which is equivalent to

\[
\frac{a_i}{b_i} = 1 - \alpha_i \frac{c_1}{c_2}.
\]

Hence the number of independent parameters is reduced to \( 4N + 1 \): \( \{ \theta_i; \sigma_i; \alpha_i; b_i, i = 1, \ldots, N; c_1 \} \).

We note that Equation (9) implies the following constraints on the model parameters

\[
b_i \left( 1 - \alpha_i \frac{c_1}{c_2} \right) > 0, \quad i = 1, \ldots, N
\]
to ensure the positivity of the parameters \( a_i \)'s. If we do not impose any other restrictions, the marginal characteristic functions (7) depend on all the model parameters which makes impossible the decoupling of the univariate implied volatility surface calibration and the correlation calibration. Indeed, once the calibration of the option surfaces is performed, there is no parameter left to calibrate the dependence.
structure. Hence, we can either use only univariate derivatives in the calibration procedure or take into account a penalty in the option surface calibration which measures the correlation goodness of fit. However, some additional conditions can be imposed to make the marginal characteristic functions independent on the model parameter $c_1$. This will lead to the original $\alpha$VG model proposed by Semeraro [16].

3.2 The original $\alpha$VG model

The $\alpha$VG model proposed by Luciano and Semeraro ([14], [16]) is obtained by imposing the equality

$$ b_i = \frac{c_2}{\alpha_i} \forall i = 1, \ldots, N $$

such that the Gamma subordinator $G^{(i)}$ is Gamma distributed: $G_1^{(i)} \sim \text{Gamma}(a_i + c_1, \frac{c_2}{\alpha_i})$. The condition (9) then becomes $a_i = \frac{c_2}{\alpha_i} - c_1$ and the marginal characteristic functions become independent on $c_1$:

$$ \phi_{Y^{(i)}}(u, t) = \left(1 - i \frac{\alpha_i}{c_2} \left(u \theta_i + \frac{1}{2} \sigma_i^2 u^2 \right) \right)^{-\frac{c_2}{\alpha_i} t}. $$

The unitary time change associated to the $i$th underlying stock, $G^{(i)}$, is then Gamma distributed and the $i$th asset log-return follows a VG($\sigma_i, \alpha_i/c_2, \theta_i$) process. The number of free parameters amounts then to $3N + 1$ ($\{\theta_i; \sigma_i; \alpha_i, i = 1, \ldots, N; c_1\}$).

Under the reduced setting, the linear correlation between the asset-log returns can be rewritten as:

$$ \rho_{ij} = \frac{\theta_i \theta_j \alpha_i \alpha_j}{\sqrt{\left(\frac{\sigma_i^2}{\theta_i} + \sigma_i^2\right) \left(\frac{\theta_j^2}{\theta_j} + \sigma_j^2\right)}} c_1 \propto c_1 $$

and is directly proportional to the common parameter $c_1$.

3.3 The generalized Sato $\alpha$VG model

Following the same methodology as in Section 3.1, the characteristic function of the process $Y_t$ (2) is given by:

$$ \phi_{Y}(u, t) = E \left[\exp(iu'Y_t)\right] = \prod_{i=1}^{N} \phi_{X^{(i)}} \left(u_i \theta_i t^{\gamma_i} + \frac{1}{2} \sigma_i^2 t^{2\gamma_i} u_i^2 \right) $$

(13)

$$ \phi_{Z_1} \left(\sum_{i=1}^{N} \alpha_i \left(u_i \theta_i t^{\gamma_i} + \frac{1}{2} \sigma_i^2 t^{2\gamma_i} u_i^2 \right) \right). $$

The marginal characteristic functions are directly obtained from (13):

$$ \phi_{Y^{(i)}}(u, t) = \left(1 - \frac{u \theta_i t^{\gamma_i} + \frac{1}{2} \sigma_i^2 t^{2\gamma_i} u_i^2}{b_i} \right)^{-\alpha_i} \left(1 - \frac{\alpha_i}{c_2} \left(u \theta_i t^{\gamma_i} + \frac{1}{2} \sigma_i^2 t^{2\gamma_i} u_i^2 \right) \right)^{-c_1} i = 1, \ldots, N. $$

The linear correlation between the asset log-return processes $Y_t^{(i)}$ and $Y_t^{(j)}$ is time independent and equal to the correlation under the corresponding Lévy models:

$$ \rho_{ij} = \frac{\text{Cov} \left(Y_t^{(i)}, Y_t^{(j)}\right)}{\sqrt{\text{Var} \left[Y_t^{(i)}\right]} \text{Var} \left[Y_t^{(j)}\right]}, $$

(15)
where
\[
\text{Cov} \left( Y_t^{(i)}, Y_t^{(j)} \right) = \theta_i \theta_j \alpha_i \alpha_j \frac{c_1}{c_2^2} t^{\gamma_i + \gamma_j}
\]
and
\[
\text{Var} \left[ Y_t^{(i)} \right] = \left( \phi_t^2 \left( \frac{a_i}{b_i} + \alpha_i^2 \frac{c_1}{c_2^2} \right) + \sigma_i^2 \left( \frac{a_i}{b_i} + \alpha_i \frac{c_1}{c_2} \right) \right) t^{2\gamma_i}.
\]

As under the Lévy two factor models, we can scale the parameter \(c_2\) to 1 and, for the sake of coherence, we will impose that the unit business time \(G^{(i)}\) has a unit expectation, i.e. we impose that
\[
E \left[ G^{(i)} \right] = \left( \frac{a_i}{b_i} + \alpha_i \frac{c_1}{c_2} \right) = 1, \text{ leading to condition (9)}.
\]

We can easily show that under the generalized Sato setting, the asset log-return process \(Y_t^{(i)}\) follows a Sato process, although the self similar distribution on which it is built is not necessarily VG, but has the following characteristic function:
\[
\phi_X(t) = \left( 1 - i \frac{u \theta_i + i \frac{1}{2} \sigma_i^2 u^2}{b_i} \right)^{-a_i} \left( 1 - i \frac{\alpha_i}{c_2} \left( u \theta_i + i \frac{1}{2} \sigma_i^2 u^2 \right) \right)^{-c_1 t}.
\]

**Proof.** To show that the processes \(Y_t^{(i)} = t^{\gamma_i} \left( \theta_i G^{(i)} + \sigma_i W_G^{(i)} \right)\) is a Sato process, we have to show that \(V^{(i)} = Y_1^{(i)} = \theta_i G^{(i)} + \sigma_i W_G^{(i)}\) is a self decomposable random variable. The characteristic function of \(V^{(i)}\) can be rewritten as the product of two VG characteristic functions:
\[
\phi_{V^{(i)}}(u) = \left( 1 - i \frac{u \theta_i + i \frac{1}{2} \sigma_i^2 u^2}{b_i} \right)^{-a_i} \left( 1 - i \frac{\alpha_i}{c_2} \left( u \theta_i + i \frac{1}{2} \sigma_i^2 u^2 \right) \right)^{-c_1} = \phi_{\text{VG}} \left( u; \sqrt{\frac{a_i}{b_i}} \sigma_i, \frac{1}{a_i} \frac{c_1}{c_2} \theta_i \right) \phi_{\text{VG}} \left( u; \sqrt{\frac{c_1}{c_2} \alpha_i \sigma_i}, \frac{1}{c_1}, \frac{c_1}{c_2} \alpha_i \theta_i \right)
\]

Hence \(V^{(i)} \overset{d}{=} U_1^{(i)} + U_2^{(i)}, \) where \(U_1^{(i)} \sim \text{VG} \left( \sqrt{\frac{a_i}{b_i}} \sigma_i, \frac{1}{a_i} \frac{c_1}{c_2} \theta_i \right)\) and \(U_2^{(i)} \sim \text{VG} \left( \sqrt{\frac{c_1}{c_2} \alpha_i \sigma_i}, \frac{1}{c_1}, \frac{c_1}{c_2} \alpha_i \theta_i \right)\) are two independent VG random variables. Hence, by (4), the \(V^{(i)}\)'s are self decomposable since they are the sum of two self decomposable random variables.

\[\square\]

### 4 The original Sato \(\alpha\)VG model

By following the idea proposed by Semeraro and Luciano, we define a reduced (or original) Sato model. Imposing the constraints \(b_i = \frac{c_2}{a_i}, \forall i,\) the constraint (9) becomes \(a_i = b_i - c_1\) and therefore the marginal characteristic function (14) becomes
\[
\phi_{Y_t^{(i)}}(u,t) = \left( 1 - i \frac{u \theta_i t^{\gamma_i} + i \frac{1}{2} \sigma_i^2 t^{2\gamma_i} u^2}{b_i} \right)^{-b_i} i = 1, \ldots, N.
\]
and the process \(Y_t^{(i)}\) then follows a VG Sato process with parameters \(\left\{ \sigma_i, \frac{1}{b_i}, \theta_i, \gamma_i \right\} \). Note that this property can be directly inferred from the subordinator distribution under the reduced setting. Indeed, since we then have \(X^{(i)} \sim \text{Gamma}(b_i - c_1, b_i)\) and \(\alpha_i Z \sim \text{Gamma}(c_1, \frac{c_2}{a_i})\), \(G^{(i)}\) is Gamma\((b_i, b_i)\)-distributed and the process \(Y_t^{(i)}\) is actually a VG Sato process written under its time-changed Brownian motion decomposition.
5 Calibration

5.1 Calibration instruments

• Underlying stocks

For the numerical study, we compare the calibration performance of the Sato VG and Lévy VG two factor models for a period extending from the second of June 2008 until the 30th of October 2009, including thus the market turmoil period of the end of 2008. We consider a basket composed of four major stocks included in the S&P500 index, namely apple, Exxon, Microsoft and Intl. For the calibration we take into account all the liquid quoted call options, i.e. for the whole sets of maturities and strikes. Since the stocks under investigation are typically dividend paying stocks and since the options written on them are of American type, we first check whether it is optimal or not to early exercise the different call options. It is never optimal to early exercise an American call option with maturity \( T \) and strike price \( K \) written on a dividend paying stock if

\[
D_i \leq K (1 - \exp(-r(t_{i+1} - t_i))) , \quad i = 1, 2, \ldots, n - 1
\]

and

\[
D_n \leq K (1 - \exp(-r(T - t_n)))
\]

where \( D_i \) denotes the dividend corresponding to the \( i \)th ex-dividend date \( t_i , i = 1, \ldots, n, t_0 < t_1 < \ldots < t_n < T \).

If the inequalities (16) and (17) hold, the American call price \( AC \) is equal to the corresponding European call price \( EC \) and we can directly apply standard Fourier transform methods such as the Carr-Madan formula [3] to compute the model call option prices. On the other hand, if inequality (16) and/or inequality (17) are/is violated, we can approximate the American call option price in terms of European call option prices written on non dividend paying stock by making use of Black’s approximation [2]

\[
AC(S_0, K, T, D_1, t_1) = \max \left( EC \left( S_0 - \sum_{i=1}^{n} D_i \exp(-r t_i), K, T \right), EC \left( S_0 - \sum_{i=1}^{n-1} D_i \exp(-r t_i), K, t_n \right) \right).
\]

This approximation rests on the fact that the inequality (16) usually holds, in the contrary to inequality (17). In other words, if it is optimal to early exercise an American call option, the early exercise date typically coincides with the last ex-dividend payment date \( t_n \).

For each stock we consider, both inequality (16) and inequality (17) are satisfied on the whole period of time under investigation for all the quoted call options. Actually, we select the four major S&P500 components as measured by market capitalization which satisfy the non early exercise conditions, such that we do not need any approximation or more advanced and time consuming numerical method to compute American call option prices.

• Correlation instruments

The market-implied calibration of the dependence structure usually requires the existence of a liquid market for multivariate derivatives which is nowadays pretty rare. Hence, the linear correlations are usually calibrated on the basis of time series estimates (see for instance [13] or [14]). Nevertheless, this calibration procedure is characterized by two shortcomings: first the marginals are calibrated under the risk-neutral measure whereas the dependence structure is calibrated under the historical measure and secondly the historical correlations turn out to strongly depend on the estimation technique and/or on the length of the time series window.
However, for any actively traded index, we can infer a correlation index by following the Chicago Board Options Exchange (CBOE) methodology [7]. Indeed, in July 2009, CBOE started to issue correlation indices on the S&P500 which are a measure of the expected average correlation for different time horizons and are computed by

\[
\rho_{\text{CBOE}} = \frac{\sigma_{\text{Index}}^2 - \sum_{i=1}^{P} w_i^2 \sigma_i^2}{2 \sum_{i=1}^{P-1} \sum_{j>i}^{P} w_i w_j \sigma_i \sigma_j},
\]

where \(\sigma_{\text{Index}}\) and \(\sigma_i\) denote the volatility of the index and of the \(i\)th index component, respectively and where \(w_i\) is the weight of the \(i\)th index component and is equal to

\[
w_i = \frac{P_i S_i}{\sum_{i=1}^{P} P_i S_i},
\]

\(P_i\) denoting the price of the \(i\)th index component and \(S_i\) the shares outstanding of the \(i\)th index component. Here we propose to calibrate the dependence structure of the subset of the S&P500 components by using the S&P500 implied correlation indices defined by (18). We note that the methodology can be applied to any index as long as options written on the index and on its components are largely traded.

Figure 1 shows that the correlation index depends on the time horizon, especially during high volatility regime periods. This term structure of the market correlation is not taken into account in the proposed models. Nevertheless, to accommodate for this market characteristic, we can extend the models by considering regime switching between a high and a low level of correlation.

5.2 Calibration procedure

For the calibration of the original models, we follow the same procedure as in [11] and [14] since we can then dissociate the calibration of the univariate option surfaces and the calibration of the correlations. On the other hand, the generalized models cannot be calibrated by following this methodology since the marginal characteristic functions (7) and (14) depend on the whole parameter set. Hence, we can either perform the calibration of the option surfaces and the correlations simultaneously or calibrate...
the whole parameter set on univariate derivatives only. In the former procedure, we will introduce a penalty function in the objective function to measure the fit of the market correlations. In the following we describe in details the two calibration procedures.

5.3 The decoupling calibration

The decoupling calibration procedure proposed by Leoni and Schoutens in [11] can be applied for any multivariate model as long as the marginal characteristic functions are independent on at least one model parameter since the methodology consists of dissociating the univariate option surface calibration from the correlation calibration. Hence the calibration might be performed in two successive steps:

1. calibration of the univariate option surfaces

We first perform a simultaneous calibration of each option surface by using fast Fourier transform techniques such as the Carr-Madan formula ([3]). For a particular choice of the common parameters $\vec{p}^c$ (i.e. the parameters which are included in more than one marginal characteristic function), we calibrate the idiosyncratic parameters $\vec{p}^i$ (i.e. the parameters which only appear in one marginal characteristic function). We then repeat the procedure for a wide range of the common parameters. The optimal marginal parameter set $\vec{p}^m = \{\vec{p}^c, \vec{p}^i\}$ (i.e. the set of both the common and idiosyncratic parameters) is the parameter set which leads to the best fit of all the option surfaces.

For the calibration of the marginal distributions, we consider a straightforward multidimensional extension of the widely used one dimensional root mean square error objective function by taking the mean of the marginal RMSE functionals:

\[
\text{MRMSE} = \frac{1}{N} \sum_{i=1}^{N} \frac{\text{RMSE}^{(i)}}{N} = \frac{1}{N} \sqrt{\sum_{i=1}^{N} \frac{\left( \sum_{j=1}^{M^{(i)}} \left( P^{(i)}_j - \hat{P}^{(i)}_j \right)^2 \right)}{M^{(i)}}},
\]

where $N$ is the number of underlying stocks, $M^{(i)}$ is the number of quoted options for the $i$th stock and $P^{(i)}_j$ and $\hat{P}^{(i)}_j$ denote the $j$th market and model option prices of the $i$th stock, respectively. The multivariate weighted RMSE objective function, MRMSE allows to calibrate separately each option surface. Indeed, we can minimize separately $\text{RMSE}^{(i)} = \sum_{j=1}^{M^{(i)}} \left( P^{(i)}_j - \hat{P}^{(i)}_j \right)^2$, where $\vec{p}^i = \{\theta_i, \sigma_i, \alpha_i\}$ and $\vec{p}^i = \{\theta_i, \sigma_i, \alpha_i, \gamma_i\}$ denotes the idiosyncratic parameter set of the $i$th underlying under the Lévy and Sato models, respectively. Hence opting for the MRMSE objective function might turn out to significantly reduce the calibration time, especially for a large number of underlyings. In the particular case of the original models we consider, the MRMSE actually reduces to $N$ univariate VG calibrations since the marginal characteristic functions do not share any common parameter $\vec{p}^c$.

2. calibration of the dependence structure

We fix the marginal parameters $\vec{p}^m$ to their optimal value according to the first step and we calibrate the correlation parameters $\vec{p}^d$ (i.e. the parameters which do not influence any marginal characteristic function, in the present case, $\vec{p}^d = c_1$) on the market implied correlations by minimizing a root mean squared objective function:

\[
\text{RMSE}^\rho = \sqrt{\frac{1}{N^2 - N} \sum_{i,j \neq i}^N \left( \rho_{ij} - \hat{\rho}_{ij} \right)^2}
\]
where $\rho_{ij}$ and $\hat{\rho}_{ij}$ denote the market implied and the model correlations between the $i$th and $j$th asset log-returns, respectively. The model correlation $\hat{\rho}_{ij}$ is directly inferred by Equation (8).

### 5.4 The joint calibration

If no reliable estimate of the dependence structure can be inferred from liquid market quotes, we can then calibrate the whole parameter set of the generalized models on the univariate option surfaces only by following the procedure described in the option surface calibration phase of the decoupling calibration procedure. In other words, we can successively minimize $\text{RMSE}^{(i)} = \text{RMSE}^{(i)}(\theta_i, \sigma_i, \alpha_i, b_i | c_1)$ and $\text{RMSE}^{(i)} = \text{RMSE}^{(i)}(\theta_i, \sigma_i, \alpha_i, \gamma_i, b_i | c_1)$ under the Lévy and Sato models, respectively and repeat the procedure for different values of the common marginal parameter $c_1$.

On the other hand, a joint calibration procedure of the univariate option surfaces and the correlations is required when the marginal characteristic functions depend on the whole model parameter set if the correlation matching is a desired feature. It requires an adequate specification of the penalty function to take into account the correlation matching in the calibration procedure of the option surfaces. We propose to minimize the following objective function:

\begin{equation}
(21) \quad \text{MRMSE} = \sum_{i=1}^{N} \frac{\text{RMSE}^{(i)}}{N} + \alpha^\rho \text{MRMSE}^* \sqrt{\frac{1}{N^2 - N} \sum_{j,k}^{N} (\rho_{jk} - \hat{\rho}_{jk})^2},
\end{equation}

where $\rho_{jk}$ and $\hat{\rho}_{jk}$ denote the market implied and the model correlations between the $j$th and $k$th log-returns, respectively and where $\text{MRMSE}^*$ is the optimal value of the multivariate root mean square errors obtained by fitting the option surfaces only. The scaling of the correlation goodness of fit by this factor ensures that both terms of Equation (21) are of the same magnitude order. The parameter $\alpha^\rho \geq 0$ allows the user to specify the relative importance of the correlation matching; a parameter $\alpha^\rho$ equal to 0 indicating that the correlation calibration is not a desired feature and that the model is calibrated on the univariate option surfaces only.

### 6 Numerical results

The calibration of the original and generalized models is performed for a time period ranging from the 2nd of June 2008 until the 30th of October 2009 with weekly quotes and therefore including the recent credit crunch. The original models are calibrated by performing the decoupling calibration procedure described in Section 5.3 whereas the generalized models are calibrated on the univariate option surfaces only or by including a penalty term which assesses the correlation goodness of fit (referred to as step 2) (see Section 5.4).

#### 6.1 The option surface goodness of fit

The multivariate RMSE (19) which assesses the univariate option surfaces goodness of fit as well as the VIX volatility index which measures the future expected market volatility over the next 30 calendar days are shown on Figure 2. We observe that the Lévy models and to a larger extent the Sato models lead to a better fit of the univariate option surfaces than the multivariate Black-Scholes model. Nevertheless the precision gain obtained by considering more advanced Lévy models than the Black-Scholes model does not turn out to be significant during the panic wave period characterized by a high value of the VIX which occurred in the aftermaths of the bankruptcy of Lehman Brothers, namely from October 2008 until December 2008. On the other hand, the Sato models systematically outperform the Black-Scholes model (and the Lévy models) in terms of the MRMSE whatever the level of investor’s fear. These results are thus in line with the fact highlighted by Carr et al. [6] that
Sato models, unlike Lévy models, are able to reproduce option prices in both the time to maturity and the strike dimensions. Figure 3 indicates that the generalized models are characterized by a slightly lower MRMSE than the original models when the former models are calibrated on univariate option surfaces only. Moreover, taking into account the correlation goodness of fit in the calibration of the generalized models leads to an option surface fit of roughly the same quality as the original and generalized models when those are calibrated on option surfaces only.

![Option surface calibration performance](image)

**Figure 2:** Evolution of the global option surface calibration performance of the Lévy models (upper) and of the Sato models (center) and evolution of the VIX volatility index (lower) through time.

### 6.2 The correlation goodness of fit

Figure 4 shows the correlation RMSE (20) under the original and generalized Lévy and Sato models. It is clear that although the original models, unlike the generalized models, have one degree of freedom to match the market implied correlation, i.e. $c_1$, they are not really able to fit the dependence structure.
Figure 3: Evolution of the global option surface calibration performance of the Lévy models (upper) and of the Sato models (lower) through time.
Indeed, the correlation RMSE, $\text{RMSE}^\rho$, varies between 37.31 and 77.12 percent under the original $\alpha$VG model and between 30.06 and 78.04 percent under the original $\alpha$VG Sato model for the time period considered. This gives some evidence against the use of the decoupling procedure to calibrate the original models and might be explained by two reasons: first there exists only one single parameter to fit the $\frac{N^2-N}{2}$ linear correlations between the $N$ underlyings and secondly, imposing the constraint (9) that on average the business clock grows as the real time, implies some additional constraints on the subordinator parameters. Indeed, in order to ensure the positivity of the parameter $a_i$’s of the idiosyncratic Gamma subordinators, the following inequalities have to hold:

$$1 - \frac{c_1}{c_2} > 0 \quad \forall i = 1, \ldots, N,$$

leading to an upper bound for the set of admissible values for $c_1$: $c_1 \in (0, \frac{1}{\max \alpha_i})$. In the decoupling calibration, the $\alpha_i$’s are determined in the first step, i.e. during the option surfaces calibration phase. Hence, the upper bound on the parameter $c_1$ might turn out to severely restrict the range of admissible values of $c_1$. In order to have some insight in the limitation of the decoupling procedure for the calibration of the original models, we have a look at the optimal value of the common subordinator parameter $c_1$ (see Figure 5). We see that, for both the Lévy and Sato original models and for most of the quoting dates under investigation, the optimal value of the parameter $c_1$ is set equal to the upper bound $\frac{1}{\max \alpha_i}$ which in turn limits severely the range of attainable correlations since the linear correlations are proportional to $c_1$ (see Equation (12) and Figure 7 and Figure 8).

Figure 6 shows the maximal attainable correlation which can be reached by the decoupling calibration procedure for each pair of asset log-returns. The upper bound on the value of the correlation gives evidence against the use of the decoupling procedure to calibrate the original models for strongly correlated underlyings. In this particular case, a more suited calibration methodology would consist of a joint calibration where the parameters $\alpha_i$’s and $c_1$ are determined simultaneously to minimize an objective function of the form (21).

Moreover, as it can be seen from Figure 4, including a penalty term assessing the correlation goodness of fit in the objective function (see Equation (21)) usually allows to significantly improve the correlation goodness of fit of the generalized models. Indeed, the $\text{RMSE}^\rho$ amounts to less than 5 percent for more than 79 and 71 percent of the quoting dates under the generalized Lévy and Sato models, respectively. Nevertheless, for some of the trading days, the correlation RMSE remains significant. This typically occurs more often under the Sato model which might be explained by the fact that the weight of the correlation term, MRMSE$^\ast$ is significantly lower under the generalized Sato model since they lead to a better fit of the univariate option surfaces than the generalized Lévy models (see Figure 2). Nevertheless, the correlation goodness of fit can then be improved by assessing more weight to $\text{RMSE}^\rho$ by increasing the value of the parameter $\alpha^\rho$ (see Figure 9).

### 7 Conclusion

This paper proposes a class of multivariate Sato models for option pricing built upon a Sato time change Brownian motion where the time change consists of a weighted sum of an idiosyncratic and a common component. We consider the particular case of Gamma subordinators and we distinguish in between a reduced model where the asset log-return margins are of VG-Sato type and a generalized model where the margins remain Sato-distributed but not necessarily VG-Sato distributed anymore. These models can be seen as an extension of the $\alpha$VG model [16] and of the generalized $\alpha$VG model [10] where the Lévy margins are replaced by Sato margins while keeping the same dependence structure. The numerical study has shown that the proposed Sato models are able to fit univariate option surfaces quotes both for low and high volatility regime periods and consequently outperform both the
Figure 4: Evolution of the correlation calibration performance of the Lévy models (upper) and of the Sato models (lower) through time.
Figure 5: Evolution of the common subordinator parameter $c_1$ of the Lévy models (upper) and of the Sato models (lower) through time.
Figure 6: Evolution of the maximal attainable correlations of the original Lévy model (upper) and of the original Sato model (lower) through time.
Figure 7: Evolution of the linear correlations under the Lévy models through time.
Figure 8: Evolution of the linear correlations under the Sato models through time.

Figure 9: Influence of $\alpha^\rho$ for the generalized Lévy model.
multivariate Black-Scholes model and the original and generalized αVG models. Moreover, we have shown that the generalized models can adequately reproduce the market implied correlations when a penalty term which assesses the correlation goodness of fit is included into the calibration surface optimizer whereas the original models usually fail to reproduce the market correlations if they are calibrated by using the decoupling calibration procedure.

References


