Branching processes for multi-curve interest rate modeling

Claudio Fontana[*], Alessandro Gnoatto[†], Guillaume Szulda[‡]

December 10, 2018

Preliminary version: please do not quote or distribute without permission

Key-words: — multiple curves, low interest rate, multiplicative spread, CBI processes.
The first author gratefully acknowledges financial support from the Europlace Institute of Finance.

Abstract

The post-crisis interest rate market exhibits two striking features: the presence of multiple yield curves, corresponding to interbank rates of different tenors, and a persistence of low rates. We introduce a tractable model which takes into account both features. The proposed model is driven by continuous-state branching processes with immigration (CBI processes), a particular class of non-negative affine processes exhibiting self-exciting jumps. CBI processes enable us to capture several facts of post-crisis interest rate markets, including the monotonicity of Euribor-OIS spreads with respect to the tenor's length, and allow for a perfect fit to the initially observed term structures. Furthermore, the model admits semi-closed valuation formulae for caplets and an approximate pricing formula for swaption. The empirical performance of the model is illustrated by calibration to market data.

1 General definition and properties

Let $(\Omega, \mathbb{F}, \mathbb{Q})$ be a filtered probability space where $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is the filtration to which all random processes will be adapted and $\mathbb{Q}$ is a risk-neutral probability measure whose numéraire process B is equal to the OIS bank account process $(\exp(\int_0^t r_s ds))_{t\geq 0}$ where $(r_t)_{t\geq 0}$ stands for the OIS short rate process and whereby all the B-discounted basic traded assets are $\mathbb{Q}$-martingales. More specifically, the *basic traded assets* include:

(i) OIS zero-coupon bonds for all maturities $T \in [0, \mathbb{T}]$;

(ii) Forward Rate Agreements for all maturities $T \in [0, \mathbb{T}]$ and tenors $\{\delta_1, \dots, \delta_I\}$.

Here, $\mathbb{T}$ denotes a given terminal horizon, while $\{\delta_1, \dots, \delta_I\}$ is the set of tenors present in the interbank market (in practice, the traded tenors range between one day and one year), with $\delta_i \leq \delta_{i+1}$ for each $1 \leq i \leq I-1$.

The fundamental modeling quantities of this framework will be the following :

[*]University of Padova(Italy), Department of Mathematics, e-mail: fontana@math.unipd.it
[†]University of Verona (Italy), Department of Economics, e-mail: alessandro.gnoatto@univr.it
[‡]Paris Diderot University (France), LPSM, e-mail: guillaume.szulda@upmc.fr

(i) The *OIS short rate*, given by $(r_t)_{t\in[0,\mathbb{T}]}$.

(ii) For each tenor $\delta \in \{\delta_1, \dots, \delta_I\}$, the logarithm of the *spot multiplicative spread* between the Libor rate and the OIS rate, denoted by $(S^{\delta_i}(t))_{t\in[0,\mathbb{T}]}$.

The framework consists in expressing such quantities by means of deterministic affine functions depending on a vector-valued stochastic process composed of mutually independent *CBI processes*. As we are going to show, this generates exponentially affine functions for OIS zero-coupon bond prices (denoted by $B_t(T)$, $\forall t \le T$ and $\forall T \le \mathbb{T}$) and forward multiplicative spreads (denoted by $S_t^{\delta_i}(T)$ $\forall t \le T$, $\forall T \le \mathbb{T}$ and for each tenor δ_i). This modeling framework is inspired from the affine setup recently proposed in [3].

As a preliminary, let us introduce the concept of *Continuous-state Branching process with immigration* (CBI process, see [7]):

Definition 1.1. *$(X_t)_{t\ge 0}$ is said to be a CBI process if it is a Markov process with state space $\mathbb{R}_+$ such that $\forall p, q \ge 0$ and $\forall t \ge s \ge 0$:*

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-pX_t - q\int_s^t X_u du\right)|\mathcal{F}_s\right] = \exp\left(-X_s v(t-s,p,q) - \int_0^{t-s} \beta v(u,p,q)du\right), \quad (1)$$

where β is a positive constant called the immigration rate and the function $v : \mathbb{R}_+^3 \longrightarrow \mathbb{R}_+$ is the unique solution of the following ordinary differential equation:

$$\frac{\partial v}{\partial t}(t,p,q) = q - \psi(v(t,p,q)), \quad v(0,p,q) = p, \quad (2)$$

where the function ψ is termed branching mechanism and is given as follows, $\forall x \ge 0$:

$$\psi(x) = bx + \frac{1}{2}\sigma^2 x^2 + \int_0^\infty (\exp(-xu\eta) - 1 + x\eta u)\mu(du), \quad (3)$$

where b, σ and η are both positive constants and μ is a Levy measure satisfying the integrability condition $\int_0^{+\infty} \min(u,u^2)\mu(du) < +\infty$.

CBI processes have been recently adopted in financial modeling in [5] in the context of a short-rate model. In particular, CBI processes have been shown to reconcile the presence of large fluctuations with low interest rates on the sovereign bond market. The following theorem (see [4]) shows that a CBI process can be considered as the unique non-negative strong solution of a stochastic differential equation.

Theorem 1.1. *There exists a unique non-negative strong solution to the following stochastic differential equation:*

$$X_t = X_0 + \int_0^t (\beta - bX_s)ds + \sigma\int_0^t\int_0^{X_s} W(ds,du) + \eta\int_0^t\int_0^{X_{s-}}\int_{\mathbb{R}_+} v\tilde{N}(ds,du,dv), \quad (4)$$

where W is a white noise on $\mathbb{R}_+^2$ and $\tilde{N}$ is the compensated Poisson random measure of a standard Poisson random measure N on $\mathbb{R}_+^3$ with intensity $dsdu\mu(dv)$ and independent of W. Furthermore, $(X_t)_{t\ge 0}$ is a CBI process with branching mechanism ψ and immigration rate β.

The above equation makes clear the self-exciting property of a CBI process: the greater the value of the process, the higher the probability of making an upward jump. That is one of the reasons why we have chosen such a stochastic process in order to model multiple yield curves.

The general model can now be defined via the following definition.

Definition 1.2. • *Let d be a positive integer.*

- *Let $X = (X^j)_{1\leq j\leq d}$ be a vector composed of d mutually independent CBI processes such that for each $1 \leq j \leq d$, X^j is defined as the unique non-negative strong solution of the following stochastic differential equation:*

$$X_t^j = X_0^j + \int_0^t (\beta_j - b_j X_s^j) ds + \sigma_j \int_0^t \int_0^{X_s^j} W^j(ds, du) + \eta_j \int_0^t \int_0^{X_{s-}^j} \int_{\mathbb{R}_+} v\tilde{N}^j(ds, du, dv),$$

where W^j is a white noise on $\mathbb{R}_+^2$ and $\tilde{N}^j$ is the compensated Poisson random measure of a standard Poisson random measure N^j on $\mathbb{R}_+^3$, independent of W^j and of intensity $dsdu\mu_j(dv)$ where μ_j is a Levy measure satisfying the following integrability condition : $\int_0^{+\infty} \min(u, u^2)\mu_j(du) < +\infty$.

- *Let $\lambda = (\lambda_j)_{1\leq j\leq d} \in \mathbb{R}_+^d$.*
- *Let $l : [0, \mathbb{T}] \to \mathbb{R}_+$ be a non-negative function such that :*

$$\int_0^{\mathbb{T}} |l(s)| ds < +\infty.$$

- *Let $\gamma = (\gamma^i)_{1\leq i\leq I}$ be a family of vectors such that for each $1 \leq i \leq I$, $\gamma^i \in \mathbb{R}_+^d$.*
- *Let $c = (c_i)_{1\leq i\leq I}$ be a family of non-negative functions such that for each $1 \leq i \leq I$, $c_i : [0, \mathbb{T}] \to \mathbb{R}_+$ and :*

$$\int_0^{\mathbb{T}} |c_i(s)| ds < +\infty.$$

Then, $(X, l, \lambda, c, \gamma)$ is termed a short-rate CBI-driven multi-curve model with, $\forall t \in [0, \mathbb{T}]$, for each $i \leq I$:

$$\begin{aligned} r_t &= l(t) + <\lambda|X_t>; \\ \log S^{\delta_i}(t) &= c_i(t) + <\gamma^i|X_t> . \end{aligned}$$

The model introduced above generates exponentially affine functions of the vector-valued stochastic process X for OIS zero-coupon bond prices and forward multiplicative spreads. We need to introduce a further integrability assumption on the measure μ_j, needed to ensure that the jump measure possesses exponential moments up to some order: for each $1 \leq i \leq I$ and $1 \leq j \leq d$, we assume that

$$\gamma_j^i \in \mathcal{Y}_j = \left\{ y \in \mathbb{R}_+ | \int_{v\geq \frac{1}{\eta_j}} \exp(y\eta_j v)\mu_j(dv) < +\infty \right\}.$$

Indeed, a CBI process is an affine process, therefore we can write for each X^j the affine transform formula on the set $\mathcal{Y}_j$ (see [6], Corollary 2.16 along with Definitions 2.7 and 2.10). This formula will give finiteness to the following conditional expectation and a way to compute it using the CBI characterization (1), for each $1 \leq j \leq d$, $\forall p \in \mathcal{Y}_j$ and $\forall t \geq s \geq 0$:

$$\mathbb{E}^{\mathbb{Q}} \left[\exp(pX_t^j) | \mathcal{F}_s \right]. \tag{5}$$

Therefore, if $\gamma_j^i \in \mathcal{Y}_j$ for each $1 \leq i \leq I$ and $1 \leq j \leq d$, then the following holds, $\forall t \in [0, \mathbb{T}]$:

$$\mathbb{E}^{\mathbb{Q}} \left[\exp\left(<\gamma^i|X_t> \right) \right] < +\infty.$$

Under these hypotheses, the following theorem can then be stated.

Theorem 1.2. *If $(X, l, \lambda, c, \gamma)$ is a short-rate CBI-driven multi-curve model, then OIS zero-coupon bond prices and forward multiplicative spreads can be explicitly computed as follows, for all maturities $T \leq \mathbb{T}$:*

- $B_t(T) = \exp\left(M(t,T) + < N(T-t)|X_t >\right)$, $\forall t \leq T$;
- *For each* $i \leq I$: $S_t^{\delta_i}(T) = \exp\left(M^i(t,T) + < N^i(T-t)|X_t >\right)$, $\forall t \leq T$,

where M, N, M^i and N^i are computed as follows:

$$
\begin{aligned}
M(t,T) &= -\int_t^T l(s)ds - \sum_{j=1}^{d} \beta_j \int_0^{T-t} v^j(s,0,\lambda_j)ds; \\
N(T-t) &= \Big(-v^1(T-t,0,\lambda_1), \cdots, -v^d(T-t,0,\lambda_d)\Big),
\end{aligned}
$$

and for each $i \leq I$:

$$
\begin{aligned}
M^i(t,T) &= c_i(T) + \sum_{j=1}^{d} \beta_j \int_0^{T-t} \left(v^j(s,0,\lambda_j) - v^j(s,-\gamma_j^i,\lambda_j)\right) ds; \\
N^i(T-t) &= \Big(v^1(T-t,0,\lambda_1) - v^1(T-t,-\gamma_1^i,\lambda_1), \cdots, v^d(T-t,0,\lambda_d) - v^d(T-t,-\gamma_d^i,\lambda_d)\Big).
\end{aligned}
$$

The following proposition shows that our model is consistent with the post-crisis interbank market especially regarding the emergence of risk factors:

Proposition 1.1. *If the following conditions are true for each $1 \leq i \leq j \leq I$:*

- $\forall t \geq 0$, $c_i(t) \leq c_j(t)$.
- *For each* $1 \leq l \leq d$, $\gamma_l^i \leq \gamma_l^j$.

Then, it holds that:

- $S^{\delta_i}(T) \geq 1$ *for each* $i \leq I$
- *For every* $i, j \leq I$, *if* $i \leq j$, *then* $S^{\delta_i}(T) \leq S^{\delta_j}(T)$.

Another appealing feature of the model is its ability to fit the initially observed term structures of OIS zero-coupon bonds and multiplicative spreads. The following proposition provides a necessary condition on the functions l and c to achieve a perfect fit, depending on the market data (denoted by M) and the theoretical term structures associated to the model $(X, 0, \lambda, 0, \gamma)$ (denoted by O).

Proposition 1.2. *The short-rate CBI-driven multi-curve model $(X, l, \lambda, c, \gamma)$ achieves an exact fit to the initially observed term structures if and only if its functions l and c_j for each $i \leq I$ are given by $\forall t \leq \mathbb{T}$:*

- $l(t) = f_0^M(t) - f_0^O(t)$.
- *For each* $i \leq I$, $c_i(t) = \log S_0^{M,\delta_i}(t) - \log S_0^{O,\delta_i}(t)$,

where $f_0^{M,O}(T) = -\frac{\partial \log B_0^{M,O}(x)}{\partial x}|_{x=T}$ *stands for the OIS instantaneous forward rates.*

2 Model specification and pricing

2.1 Flow of CBI processes

The specification of the general model that we will implement and calibrate on market data corresponds to modeling spot multiplicative spreads by means of a flow of CBI processes (see [4]). In particular, this specification will allow to generate a non-trivial correlation among the multiplicative spreads.

Definition 2.1. *Consider a non-decreasing function $a : \{\delta_1, \dots, \delta_I\} \to \mathbb{R}_+$ and for each $1 \le i \le I$, the logarithm of the spot multiplicative spread is given by*

$$\log S^{\delta_i}(t) = c_i(t) + Y_t^{\delta_i},$$

where Y^{δ_i} is given by the solution to

$$Y_t^{\delta_i} = Y_0^{\delta_i} + \int_0^t (a(\delta_i) - bY_s^{\delta_i})ds + \sigma \int_0^t \int_0^{Y_s^{\delta_i}} W(ds, du) + \eta \int_0^t \int_0^{Y_{s-}^{\delta_i}} \int_{\mathbb{R}_+} v\tilde{N}(ds, du, dv), \quad (6)$$

$Y = (Y^{\delta_i})_{1 \le i \le I}$ is then a flow of CBI processes of branching mechanism ψ given by (3) and immigration rate $\beta = (\beta_i)_{1 \le i \le I}$ where $\beta_i = a(\delta_i)$ for each $1 \le i \le I$.

We can show that this specification is a particular case of the general model of Definition 1.2 through the following proposition.

Proposition 2.1. *If for each $1 \le i \le j \le I$, $\delta_j \ge \delta_i$ and $Y_0^{\delta_j} \ge Y_0^{\delta_i}$, then $Y_t^{\delta_j} \ge Y_t^{\delta_i} \ge 0$, $\forall t \ge 0$ and with probability 1. Also, there exists a vector-valued stochastic process $X = (X^l)_{1 \le l \le I}$ made up of I mutually independent CBI processes such that for each tenor δ_i and $\forall t \ge 0$:*

$$Y_t^{\delta_i} = \sum_{l \le i} X_t^l,$$

where for each $1 \le l \le I$, $(X_t^l)_{t \ge 0}$ is a CBI process with branching mechanism ψ given by (3) and immigration rate $a(\delta_l) - a(\delta_{l-1})$ with $a(\delta_0) = 0$.

This specification is a particular case of the short-rate CBI-driven multi-curve model of Definition 1.2 by setting the parameter d equal to the number of tenors I, the vector X of Definition 1.2 equal to the one of size I whose existence was stated in Proposition 2.1 and by choosing the square matrix γ equal to an upper triangular matrix in which all the elements located in the upper part are equal to one.

In Theorem 1.2, we need the following condition satisfied for $p = 1$ for each $1 \le i \le I$ and $\forall t \ge 0$:

$$\mathbb{E}^{\mathbb{Q}} \left[\exp \left(p Y_t^{\delta_i} \right) \right] < +\infty, \quad (7)$$

namely p must belong to the following set to remain consistent with the statement of the affine transform formula:

$$p \in \mathcal{Y} = \left\{ y \in \mathbb{R}_+ | \int_{v \ge \frac{1}{\eta}} \exp(y\eta v)\mu(dv) < +\infty \right\}. \quad (8)$$

This constraint can then be reduced to the following integrability condition:

$$\int_{v \ge \frac{1}{\eta}} \exp(v\eta)\mu(dv) < +\infty.$$

Note that the finiteness of this integral depends on the structure of the Levy measure μ, which will be studied in the next section.

2.2 Tempered alpha-stable measure

Henceforth, we will specify further the model introduced in Section 2.1 by choosing a specific structure for the Levy measure μ in the intensity of the Poisson random measure N

Definition 2.2. *$Y = (Y^{\delta_i})_{1 \le i \le I}$ is said to be a flow of tempered alpha-stable CBI processes if its Levy measure μ is given by :*

$$\mu(dx) = -\frac{\mathbb{1}_{x>0}\exp(-\zeta x)}{\cos(\pi\alpha/2)\,\Gamma(-\alpha)x^{1+\alpha}}dx = \exp(-\zeta x)\nu(dx) \tag{9}$$

This Levy measure is called tempered alpha-stable because it is obtained after tempering the standard alpha-stable Levy measure ν.

The set $\mathcal{Y}$ can here be determined in an explicit form due to the structure of the Levy measure:

$$\begin{aligned}\mathcal{Y} &= \left\{y \in \mathbb{R}_+ | \int_{v \ge \frac{1}{\eta}} \exp(y\eta v)\mu(dv) < +\infty\right\} \\ &= \left\{y \in \mathbb{R}_+ | \int_{v \ge \frac{1}{\eta}} \exp((y\eta - \zeta)v)\nu(dv) < +\infty\right\} \\ &= [0, \frac{\zeta}{\eta}].\end{aligned}$$

Consequently, for each $1 \le i \le I$ and $\forall t \ge 0$, the integrability condition $\mathbb{E}^{\mathbb{Q}}[\exp(pY_t^{\delta_i})] < +\infty$ reduces to

$$p \le \frac{\zeta}{\eta}. \tag{10}$$

This specification also enables us to simplify the expressions of OIS zero-coupon bond prices and forward multiplicative spreads by holding the same function v throughout the computations, since the corresponding branching mechanism ψ is always of the same form. However, ψ has to be extended to the domain $D = [-\frac{\zeta}{\eta}, +\infty)$ in the same way as the CBI characterization and the function v. In this setting, the branching mechanism ψ can then be computed explicitly through the following proposition.

Proposition 2.2. *In the context of a flow of tempered alpha-stable CBI processes $Y = (Y^{\delta_i})_{1 \le i \le I}$, the associated branching mechanism ψ is of the following form $\forall x \in D$*

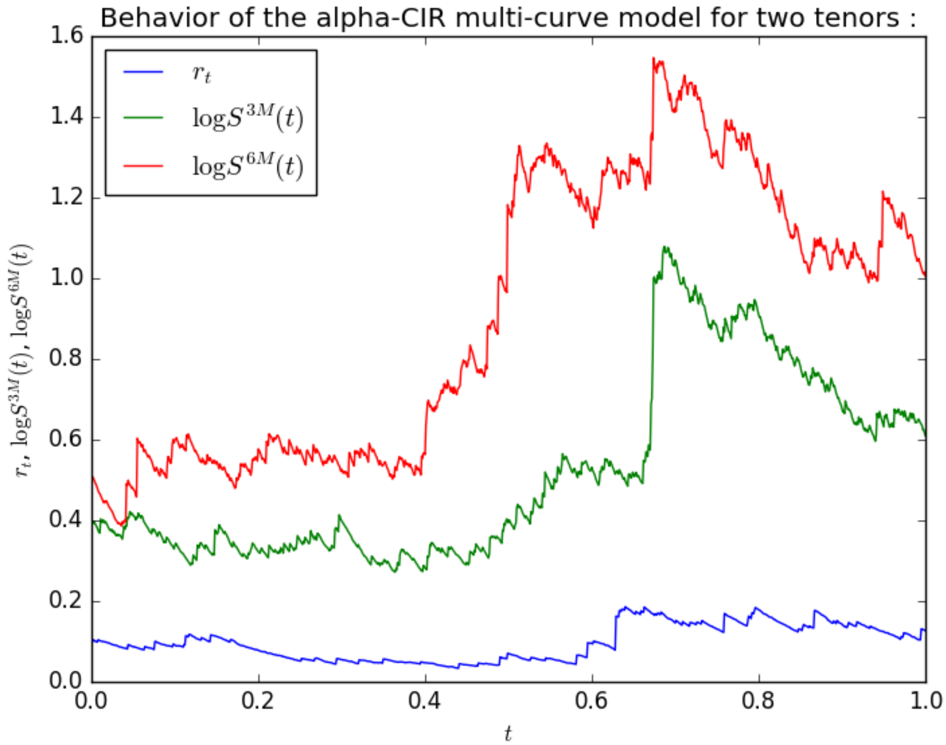
$$\psi(x) = bx + \frac{1}{2}\sigma^2 x^2 + \frac{\zeta^\alpha + x\eta\zeta^{\alpha-1}\alpha - (x\eta + \zeta)^\alpha}{\cos(\pi\alpha/2)}.$$

Besides, under the additional hypothesis:

$$b - \sigma^2\frac{\zeta}{\eta} + \frac{\alpha\eta\zeta^{\alpha-1}}{\cos(\pi\alpha/2)} \ge 0,$$

the associated function v of (2) can also be expressed in a semi-explicit form through the following proposition.

Proposition 2.3. *Let $x_q \in D$ be the unique solution of the equation $\psi(x) = q$, $\forall q \ge 0$ and $\forall x \in D$ (its existence is guaranteed thanks to the properties of ψ over D by means of the above hypothesis), then the function v can be given $\forall t \ge 0$, $\forall p \in D$ and $\forall q \ge 0$, as follows:*



- *If* $p < x_q$*, then* $v(t,p,q) = f_{p,q}^{-1}(t)$ *where* $f_{p,q} : x \in [p, x_q) \mapsto \int_p^x \frac{1}{q-\psi(u)} du$*.*
- *If* $p > x_q$*, then* $v(t,p,q) = f_{p,q}^{-1}(t)$ *where* $f_{p,q} : x \in (x_q, p] \mapsto \int_x^p \frac{1}{\psi(u)-q} du$*.*
- *If* $p = x_q$*, then* $v(t,p,q) = p$*.*

To illustrate the behavior of the model, we report in the above figure a simulated trajectory of the OIS short-rate and two spot multiplicative spreads. In particular, observe the persistence of low values for the OIS short-rate and the self-exciting feature of jumps in the evolution of the spot multiplicative spreads.

3 Interest rate derivative pricing

On the post-crisis interbank market and in the context of the short-rate CBI-driven multi-curve model, all fixed income financial contracts will be assumed to be perfectly collateralized financial contracts whose collateral rate is equal to the OIS short rate. Under this assumption, the clean price of every interest rate derivative can be defined as follows.

Definition 3.1. *Let* X_T *be the* $\mathcal{F}_T$*-measurable payoff of an interest rate derivative at maturity* T*. The clean price* $P(X_T)$ *at* $t = 0$ *is given by :*

$$P(X_T) = \mathbb{E}^{\mathbb{Q}} \left[\frac{X_T}{B_T} \right]. \tag{11}$$

By using the Bayes formula, this price is also given by means of the T*-forward measure* $\mathbb{Q}^T$ *:*

$$P_t(X_T) = B_0(T) \mathbb{E}^{\mathbb{Q}^T} [X_T]. \tag{12}$$

Given these two pricing formulas and since the specification of our model will be calibrated on caplets, which are non-linear interest rate products, we will need three fast pricing routines at our disposal, which are the following :

- Fourier-based valuation formula.
- quantization.

Furthermore, we will employ a Monte-Carlo pricing routine as a benchmark in the calibration analysis.

3.1 Fourier-based valuation formula

We will work under the risk-neutral probability measure $\mathbb{Q}$, under which the caplet price can be written as follows:

$$P^{Cplt}(S,T,K) = \mathbb{E}^{\mathbb{Q}}\left[\frac{B_S(T)}{B_S}\left(\frac{S^{\delta_i}(S)}{B_S(T)} - (1+K\delta_i)\right)^+\right],$$

or equivalently

$$P^{Cplt}(S,T,K) = \mathbb{E}^{\mathbb{Q}}\left[\frac{B_S(T)}{B_S}\left(\exp(Z_S^i) - \exp(k)\right)^+ |\mathcal{F}_t\right], \tag{13}$$

where $Z_S^i = \log\left(\frac{S^{\delta_i}(S)}{B_S(T)}\right)$ and $k = \log(1+K\delta_i)$. Note that this random variable can be computed given as follows:

$$Z_S^i = c_i(S) + <\gamma^i|X_S> - M(S,T) - <N(\delta_i)|X_S>.$$

The caplet price can then be computed by means of a generalized (complex) inverse Fourier transform representation depending on the discounted characteristic function of Z_S^i, which is analytic on the following strip of complex numbers (the simplified form of the strip is due to the structure of the state variable):

$$\left\{z \in \mathbb{C} \quad | \quad -\Im(z) < \frac{\zeta}{\eta} + 1\right\}. \tag{14}$$

On the above, set the discounted characteristic function admits an explicit representation as follows:

$$\begin{aligned}\Pi^{Z_S^i}(z) &= \mathbb{E}^{\mathbb{Q}}\left[\frac{B_S(T)}{B_S}\exp\left(izZ_S^i\right)\right] \\ &= \exp\left(-\int_0^S l(u)du + M(S,T)(1-iz) + izc_i(S) + <\hat{N}(S)|X_0> + \hat{M}(0,S)\right),\end{aligned}$$

where the functions $\hat{N} : \mathbb{R}_+ \to \mathbb{C}^d$ and $\hat{M} : \mathbb{R}_+^2 \to \mathbb{C}$ are computed as follows, $\forall t \leq S$ and $\forall S \leq \mathbb{T}$:

$$\begin{aligned}\hat{N}(S-t) &= (-w(S-t,-u_1,\lambda_1),\ldots,-w(S-t,-u_I,\lambda_I))\,; \\ \hat{M}(t,S) &= -\sum_{l=1}^{I}(a(\delta_l) - a(\delta_{l-1}))\int_0^{S-t} w(s,-u_l,\lambda_l)ds,\end{aligned}$$

where the complex number u_l also depends upon z via the following formula, for each $1 \leq l \leq I$:

$$u_l = (iz-1)v(\delta_i, 0, \lambda_l) + iz\gamma_l^i.$$

It can then be noticed that its real part satisfies

$$\Re u_l < \frac{\zeta}{\eta},$$

which allows us to affirm that u_l belongs to the following strip:

$$\mathcal{S} = \{z \in \mathbb{C} \quad | \quad \Re z \in \mathbb{R}_- \cup \mathcal{Y}^{\text{o}}\},$$

where $\mathcal{Y}^{\text{o}}$ denotes the interior of $\mathcal{Y}$ (See (8)). Note that the inverse of this set is given by

$$\mathcal{IS} = \{z \in \mathbb{C} \quad | \quad -\Re z \in \mathbb{R}_- \cup \mathcal{Y}^{\text{o}}\}.$$

Thanks to [6], Theorem 2.26, the function $w : \mathbb{R}_+ \times \mathcal{IS} \times \mathbb{R}_+ \to \mathcal{IS}$ is defined as the unique function of the following ordinary differential equation (similarly to the function v, it can then be seen as the analytic extension of the function v to the complex set $\mathcal{IS}$):

$$\frac{\partial w}{\partial t}(t, -u_l, \lambda_l) = \lambda_l - R(w(t, -u_j, \lambda_l)), \quad w(0, -u_l, \lambda_l) = -u_l,$$

for each $1 \leq l \leq I$, where R the analytical extension of ψ to the strip $\mathcal{IS}$.

A caplet can then be priced by means of a generalized inverse Fourier transform, as considered in [3]

$$P^{Cplt}(S,T,K) = R(Z_S^i, k, g_i) + \frac{1}{\pi}\int_{0-ig_i}^{+\infty-ig_i} \Re\left(\exp(-izk)\frac{\Pi^{Z_S^i}(z-i)}{-z(z-i)}\right)dz, \qquad (15)$$

where g_i has to satisfy :

$$g_i < \frac{\zeta}{\eta},$$

and where $R(Z_S^i, k, g_i)$ refers to the correction term due to the application of the residue theorem and it has several definitions depending upon the value of g_i:

$$\begin{aligned} R(Z_s^i, k, g_i) &= 0 \quad \text{if} \quad g_i > 0; \\ &= \frac{1}{2}\Pi^{Z_S^i}(-i) \quad \text{if} \quad g_i = 0; \\ &= \Pi^{Z_S^i}(-i) \quad \text{if} \quad -1 < g_i < 0; \\ &= \Pi^{Z_S^i}(-i) - \exp(k)\frac{\Pi^{Z_S^i}(0)}{2} \quad \text{if} \quad g_i = -1; \\ &= \Pi^{Z_S^i}(-i) - \exp(k)\Pi^{Z_S^i}(0) \quad \text{if} \quad g_i > -1. \end{aligned}$$

For further details on the numerical implementation of the contour integral, first we can notice that it can be reduced to an integral computation over the positive real axis by expressing the integration variable as $z = x - ig_i$ where $x \in \mathbb{R}_+$, which yields

$$\frac{e^{-kg_i}}{\pi}\Re\left(\int_0^{+\infty} e^{-ixk}\frac{\Pi^{Z_S^i}(x-i(g_i+1))}{-(x-ig_i)(x-i(g_i+1))}dx\right).$$

We can then apply the Fast Fourier Transform (FFT) to compute this integral as suggested in [2]. To do so, we first have to approximate the integral by means of the trapezoid rule for $x_{j_1} = \theta(j_1 - 1)$, $1 \leq j_1 \leq N$ where N must be a power of 2 to apply the algorithm, which yields

$$\frac{e^{-kg_i}}{\pi}\Re\left(\sum_{j_1=1}^{N} e^{-ix_{j_1}k}\frac{\Pi^{Z_S^i}(x_{j_1} - i(g_i+1))}{-(x_{j_1} - ig_i)(x_{j_1} - i(g_i+1))}\theta\right).$$

Given that the FFT usually takes as an input N complex numbers and returns N other complex numbers, we can take advantage of it by computing directly different caplet prices for several strikes, which is why we set a grid for the log-strikes as follows: $k_{j_2} = -b+\theta^*(j_2-1)$, $1 \leq j_2 \leq N$ where $b = \frac{N\theta^*}{2}$. In order to apply the FFT, we have to set $\theta\theta^* = \frac{2\pi}{N}$, which yields

$$\frac{e^{-k_{j_2}g_i}}{\pi}\Re\left(\sum_{j_1=1}^{N} e^{-i\frac{2\pi}{N}(j_1-1)(j_2-1)}e^{ibx_{j_1}}\frac{\Pi^{Z_S^i}(x_{j_1} - i(g_i+1))}{-(x_{j_1} - ig_i)(x_{j_1} - i(g_i+1))}\theta\right).$$

Then, after adding Simpson's rule weightings into our summation to keep accuracy even for large values of θ, we have the following :

$$\frac{e^{-k_{j_2}g_i}}{\pi}\Re\left(\sum_{j_1=1}^{N} e^{-i\frac{2\pi}{N}(j_1-1)(j_2-1)}e^{ibx_{j_1}}\frac{\Pi^{Z_S^i}(x_{j_1} - i(g_i+1))}{-(x_{j_1} - ig_i)(x_{j_1} - i(g_i+1))}\frac{\theta}{3}(3+(-1)^{j_1} - \mathbb{1}_{j_1=1})\right),$$

which can be totally computed by means of the FFT if N is a power of 2. This finally provides us with the time-0 caplet prices along the considered log-strike grid.

3.2 Pricing by quantization

Henceforth, we will place ourselves under the T-forward probability measure $\mathbb{Q}^T$ so as to reduce our caplet pricing formula to a one-dimensional option pricing problem whose state variable is strictly positive, then the time-0 price of our caplet is given by

$$P^{Cplt}(S,T,K) = B_0(T)\mathbb{E}^{\mathbb{Q}^T}\left[\left(\frac{S^{\delta_i}(S)}{B_S(T)} - (1+K\delta_i)\right)^+\right]. \tag{16}$$

The principle of one-dimensional quantization, here applied to numerical integration in particular to option pricing, intuitively consists in finding the best discrete representation of the continuous distribution of the state variable under consideration, here equal to $\frac{S^{\delta_i}(S)}{B_S(T)}$. The random variable that will represent this discrete distribution will be denoted by $\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}}$ and will take values into a finite set of elements that will be termed a *quantization grid*, or *quantizer*. N-quantizing $\frac{S^{\delta_i}(S)}{B_S(T)}$ then means approximating it by the discrete variate $\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}}$ such that it only takes values in a quantization grid of level N, typically denoted by $\Gamma^N = \{x_1, \ldots, x_N\}$ where Γ^N is of size N with $x_1 < \cdots < x_N$ (the state space is the positive real axis given our state variable).

In our setting, we will work with the Euclidean norm on the positive real axis and $\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}}$ will be defined as the nearest-neighbor projection of the state variable $\frac{S^{\delta_i}(S)}{B_S(T)}$ on $\Gamma^N = \{x_1, \ldots, x_N\}$, computed as follows:

$$\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}} = \sum_{i=1}^{N} x_i \mathbb{1}_{x_i^- \leq \frac{S^{\delta_i}(S)}{B_S(T)} \leq x_i^+},$$

where $x_i^- = \frac{x_{i-1}+x_i}{2}$ and $x_i^+ = \frac{x_{i+1}+x_i}{2}$ for each $1 \le i \le N$ with $x_1^- = 0$ and $x_N^+ = +\infty$. $\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}}$ is then called the *Voronoi* Γ^N quantization of $\frac{S^{\delta_i}(S)}{B_S(T)}$.

Therefore, as soon as we have the Voronoi quantization of the state variable, we can proceed by computing

$$\begin{aligned} P^{Cplt}(S,T,K) &= B_0(T)\mathbb{E}^{\mathbb{Q}^T}\left[\left(\frac{S^{\delta_i}(S)}{B_S(T)} - (1+K\delta_i)\right)^+\right] \\ &\approx B_0(T)\mathbb{E}^{\mathbb{Q}^T}\left[\left(\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}} - (1+K\delta_i)\right)^+\right] \\ &= B_0(T)\sum_{i=1}^{N}(x_i - (1+K\delta_i))^+\,\mathbb{Q}^T\left(\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}} = x_i\right). \end{aligned}$$

where $\mathbb{Q}^T\left(\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}} = x_i\right)$, for each $1 \le i \le N$, are called the companion weights of the quantization and can be computed via the cumulative distribution function of $\frac{S^{\delta_i}(S)}{B_S(T)}$ under $\mathbb{Q}^T$:

$$\begin{aligned} \mathbb{Q}^T\left(\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}} = x_i\right) &= \mathbb{Q}^T\left(x_i^- \le \frac{S^{\delta_i}(S)}{B_S(T)} \le x_i^+\right) \\ &= \mathbb{Q}^T\left(\frac{S^{\delta_i}(S)}{B_S(T)} \le x_i^+\right) - \mathbb{Q}^T\left(\frac{S^{\delta_i}(S)}{B_S(T)} \le x_i^-\right). \end{aligned}$$

Therefore, since the Voronoi quantization can be determined quite easily as soon as we know about the corresponding quantization grid, the real optimal quantization problem at level N consists in reality in finding the best quantization grid or quantizer Γ^N such that the discrete distribution of $\widehat{\frac{S^{\delta_i}(S)}{B_S(T)}}$ over Γ^N approximates the continuous one of $\frac{S^{\delta_i}(S)}{B_S(T)}$ at best in the L^p-sense for $p \ge 1$. Such a grid is termed *optimal* and minimizes the L^p-mean quantization error between the state variable and its quantization. This error is defined below for some not necessarily optimal grid Γ^N:

$$e_{p,N}\left(\frac{S^{\delta_i}(S)}{B_S(T)}, \Gamma^N\right) = \left\|\frac{S^{\delta_i}(S)}{B_S(T)} - \widehat{\frac{S^{\delta_i}(S)}{B_S(T)}}\right\|_{L^p(\mathbb{Q}^T)} = \mathbb{E}^{\mathbb{Q}^T}\left[\min_{1\le j\le N}\left|\frac{S^{\delta_i}(S)}{B_S(T)} - x_j\right|^p\right]^{\frac{1}{p}}.$$

Hence, the optimal quantization grid of size N solves the optimal L^p-mean quantization problem that can be rewritten as follows:

$$e_{p,N}\left(\frac{S^{\delta_i}(S)}{B_S(T)}\right) = \inf_{\Gamma^N \subset \mathbb{R}_+^*} e_{p,N}\left(\frac{S^{\delta_i}(S)}{B_S(T)}, \Gamma^N\right) = \inf_{(x_1,\ldots,x_N)\in(\mathbb{R}_+^*)^N} D_p(x_1,\ldots,x_N)^{\frac{1}{p}}, \quad (17)$$

where the positive-valued multivariate function D_p stands for the L^p-distortion function associated to $\frac{S^{\delta_i}(S)}{B_S(T)}$ and is defined over $(\mathbb{R}_+^*)^N$:

$$\begin{aligned} D_p(x_1,\ldots,x_N) &= e_{p,N}\left(\frac{S^{\delta_i}(S)}{B_S(T)}, \Gamma^N\right)^p \\ &= \mathbb{E}^{\mathbb{Q}^T}\left[\min_{1\le j\le N}\left|\frac{S^{\delta_i}(S)}{B_S(T)} - x_j\right|^p\right] \\ &= \sum_{j=1}^{N}\int_{x_j^-}^{x_j^+}|x - x_j|^p d\mathbb{Q}^T\left(\frac{S^{\delta_i}(S)}{B_S(T)} \le x\right). \end{aligned}$$

Remark 1. Notice that since we deal with a one-dimensional setting and we are only interested in quantization grids of size N with distinct increasing components, then each grid Γ^N in our setting can be represented by a N-tuple $(x_1, \ldots, x_N)$, or a N-dimensional vector, with pairwise distinct components along with its $N!$ representations obtained by permutations of its components given that the L^p-distortion function D_p has multivariate symmetry.

In our one-dimension setting, there exists at least one solution, at least one optimal quantization grid or quantizer, and each of them has full size N. Therefore, computing the optimal quantization grids of size N for the state variable $\frac{S^{\delta_i}(S)}{B_S(T)}$ is equivalent to computing the minimum points of the L^p-distortion function D_p, which are part of its critical points, also called the p-stationary or sub-optimal grids. To this effect, recall that the L^p-distortion function is differentiable at any N-tuple with pairwise distinct components, so that we have to look for the N-tuples that make the gradient of the L^p-distortion function equal to zero to obtain the p-stationary or sub-optimal grids. Moreover, we can perform under $\mathbb{Q}^T$ the Fourier-Quantization algorithm described in [1] given that the T-forward characteristic function of the logarithm of the state variable, which will be denoted by $\Pi_T^{Z_S^i}$ where $Z_S^i = \log\left(\frac{S^{\delta_i}(S)}{B_S(T)}\right)$, is known in closed-form, $\forall u \in \mathbb{R}$:

$$
\begin{aligned}
\Pi_T^{Z_S^i}(u) &= \mathbb{E}^{\mathbb{Q}^T}\left[e^{iuZ_S^i}\right] \\
&= \frac{\Pi^{Z_S^i}(u)}{B_0(T)},
\end{aligned}
$$

where $\Pi^{Z_S^i}$ refer to the (complex) discounted characteristic function.

Therefore, the T-forward probability density function of the state variable along with its cumulative distribution function can be expressed by means of Fourier transforms:

$$
\begin{aligned}
\mathbb{Q}^T\left(\frac{S^{\delta_i}(S)}{B_S(T)} \in dx\right) &= \left(\frac{1}{x\pi}\int_0^{+\infty} \Re\left(e^{-iu\log(x)}\Pi_T^{Z_S^i}(u)\right)du\right)dx \\
\mathbb{Q}^T\left(\frac{S^{\delta_i}(S)}{B_S(T)} \leq x\right) &= \frac{1}{2} - \frac{1}{\pi}\int_0^{+\infty} \Re\left(\frac{e^{-iu\log(x)}\Pi_T^{Z_S^i}(u)}{iu}\right)du.
\end{aligned}
$$

Thus, the above L^p-distortion function can be computed by means of these two quantities as well as its gradient denoted by ∇D_p. Besides, since we are in a one-dimension setting and the state variable is assumed to be continuous and strictly positive, ∇D_p can also be differentiated, which provides us with its Hessian represented by a tridiagonal matrix denoted by $H[D_p]$. Note that we can then perform fast and powerful deterministic zero-search algorithms to find p-stationary grids like the Newton-Raphson one on which the entire Fourier-Quantization algorithm relies.

Furthermore, we also would like to have a database of optimal or sub-optimal quantization grids for the state variable $\frac{S^{\delta_i}(S)}{B_S(T)}$ for all levels from $N = 1$ to a fixed N_{max}. For $N = 1$, it is clear that the optimal quantization grid is reduced to $\left\{\mathbb{E}^{\mathbb{Q}^T}\left[\frac{S^{\delta_i}(S)}{B_S(T)}\right]\right\}$, where $\mathbb{E}^{\mathbb{Q}^T}\left[\frac{S^{\delta_i}(S)}{B_S(T)}\right] = 1 + \delta_i \mathbb{E}^{\mathbb{Q}^T}\left[L_S(S,T)\right] = 1 + \delta_i L_0(S,T)$ with $T = S + \delta_i$ and the martingale feature of the forward Libor rate applied over $[S, T]$ under $\mathbb{Q}^T$.

However, in regard to the general case $N \geq 2$, such a database could be used to initialize the Newton-Raphson algorithm at each level N in the sense that having the grid

$\Gamma^{N-1} = \{x_1^{N-1}, \ldots, x_{N-1}^{N-1}\}$ of size $N-1$ resulting from the convergence of the Newton-Raphson iteration performed at level $N-1$, we can initialize that of level N with $\Gamma_{(0)}^N$ defined as

$$\Gamma_{(0)}^N = \left\{x_1^{N-1}, \ldots, x_j^{N-1}, 1 + \delta_i L_0(S,T), x_{j+1}^{N-1}, \ldots, x_{N-1}^{N-1}\right\},$$

where $x_j^{N-1} < 1 + \delta_i L_0(S,T) < x_{j+1}^{N-1}$. Then, we can perform the Newton-Raphson algorithm of level N at each iteration n :

$$\Gamma_{(n+1)}^N = \Gamma_{(n)}^N - \left(H[D_p](\Gamma_{(n)}^N)\right)^{-1} . \nabla D_p(\Gamma_{(n)}^N).$$

Remark 2. Note that for each iteration n, $\Gamma_{(n)}^N$ is seen as its N-dimensional vector representation such that its coordinates have been sorted so that it has distinct increasing components to be similar to the structure of a quantization grid.

This can effectively improve the level of convergence of the iteration towards an optimal, or sub-optimal, grid at each level N of the database. Also, we can choose the level N for which the L^p-mean quantization error between the state variable and its Voronoi quantization is the smallest, which can allow us to further approach an optimal quantization grid that really minimizes the L^p-distortion function instead of a sub-optimal quantizer that only makes the gradient of the distortion function equal to zero. Nevertheless, either way, after expressing the corresponding companion weights of the law of the associated Voronoi quantization thanks to the cumulative distribution function of the state variable $\frac{S^{\delta_i}(S)}{B_S(T)}$, we can then determine our caplet price (16) by means of the cubature formula resulting from the discrete feature of the quantization:

$$P^{Cplt}(S,T,K) = B_0(T) \sum_{i=1}^{N} (x_i - (1 + K\delta_i))^+ \left(\mathbb{Q}^T \left(\frac{S^{\delta_i}(S)}{B_S(T)} \leq x_i^+ \right) - \mathbb{Q}^T \left(\frac{S^{\delta_i}(S)}{B_S(T)} \leq x_i^- \right) \right).$$

References

[1] Callegaro, G., Fiorin, L., Grasselli, M. (2018), Quantization meets Fourier: a new technology for pricing options, *Annals of Operations Research*, forthcoming.

[2] Carr, P., Madan, D. (1999), Option valuation using the fast Fourier transform, *Journal of Computational Finance*, 2(4): 61–73.

[3] Cuchiero, C., Fontana, C., Gnoatto, A. (2018), Affine multiple yield curve models, *Mathematical Finance*, forthcoming.

[4] Dawson, D.A., Li, Z. (2012), Stochastic equations, flows and measure-valued processes, *Annals of Probability*, 40(2): 813–857.

[5] Jiao, Y., Ma, C., Scotti, S. (2017), Alpha-CIR model with branching processes in sovereign interest rate modeling, *Finance and Stochastics*, 21(3): 789–813.

[6] Keller-Ressel, M., Mayerhofer, E. (2015), Exponential moments of affine processes, *Annals of Applied Probability*, 25(2): 714–752.

[7] Li, Z. (2011), *Measure-Valued Branching Markov Processes*, Springer, Berlin - Heidelberg.