ORIGINAL RESEARCH

CBI-time-changed Lévy processes for multi-currency modeling

Claudio Fontana[1] · Alessandro Gnoatto[2] · Guillaume Szulda[1,3]

Abstract

We develop a stochastic volatility framework for modeling multiple currencies based on CBI-time-changed Lévy processes. The proposed framework captures the typical risk characteristics of FX markets and is coherent with the symmetries of FX rates. Moreover, due to the self-exciting behavior of CBI processes, the volatilities of FX rates exhibit self-exciting dynamics. By relying on the theory of affine processes, we show that our approach is analytically tractable and that the model structure is invariant under a suitable class of risk-neutral measures. A semi-closed pricing formula for currency options is obtained by Fourier methods. We propose two calibration methods, also by relying on deep-learning techniques, and show that a simple specification of the model can achieve a good fit to market data on a currency triangle.

Keywords FX market · Multi-currency market · Branching process · Self-exciting process · Time-change · Stochastic volatility · Deep calibration · Affine process

Mathematics Subject Classification 60G51 · 60J85 · 91G20 · 91G30 · 91G60

JEL Classification C02 · C60 · G13 · G15

1 Introduction

The Foreign-Exchange (FX) market is one of the largest in the world (see, e.g., Bank for International Settlements, 2019; Wooldridge, 2019). From the perspective of quantitative

✉ Claudio Fontana
  fontana@math.unipd.it

  Alessandro Gnoatto
  alessandro.gnoatto@univr.it

  Guillaume Szulda
  szulda.guillaume@gmail.com

1   Department of Mathematics "Tullio Levi Civita", University of Padova, Padua, Italy

2   Department of Economics, University of Verona, Verona, Italy

3   Laboratoire de Probabilités, Statistique et Modélisation, Université de Paris Cité, Paris, France

⌔ Springer

finance, modeling the FX market poses several challenges. First, multi-currency models must respect the symmetric structure of FX rates. To illustrate this aspect, let $S^{d,f}$ represent the value of one unit of a foreign currency f measured in units of the domestic currency d. In a multi-currency model, the following symmetric relations must hold:

- $S^{f,d} = 1/S^{d,f}$: the reciprocal of $S^{d,f}$ must coincide with $S^{f,d}$, representing the value of one unit of currency d measured in units of currency f. This is referred to as *inversion*.
- $S^{d,f} = S^{d,e} \times S^{e,f}$, for any other foreign currency e. In other words, the FX rate $S^{d,f}$ must be inferred from $S^{d,e}$ and $S^{e,f}$ through multiplication. This is referred to as *triangulation*.

Besides these symmetric relations, the FX market presents some specific risk characteristics that should be properly reflected in a multi-currency model. First, FX markets are affected by stochastic volatility and jump risk, similarly to the case of equity markets. This has led to the application to FX markets of well-known models initially conceived for stock returns, such as the Heston model (see also Sect. 1.1 below). Second, the dependence between FX rates is typically stochastic and, in particular, shows evidence of unpredictable changes over time, thus generating correlation risk. Third, the skew of the FX volatility smile exhibits a stochastic behavior. This fact has been documented in Carr and Wu (2007) by analyzing the time series of risk-reversals[1], showing that their values vary significantly over time and exhibit repeated sign changes. Finally, FX rates are affected by volatility clustering effects, similarly to many other asset classes (see, e.g., Cont and Tankok, 2004, Chapter 7). This phenomenon can be observed in Fig. 1, which displays the time series of the USD-JPY exchange rate over the period 01/01/2012 - 31/12/2015.

In this paper, we develop a modeling framework for multiple currencies that is consistent with the symmetric structure of FX rates and captures all risk characteristics described above, including stochastic dependence among FX rates as well as between FX rates and their volatilities. We consider models driven by *CBI-time-changed Lévy processes* (CBITCL processes), a broad and flexible class of processes that allows for self-exciting jump dynamics, with stochastic volatility and mean-reversion (we refer to Fontana et al., 2022 and Szulda, 2021, for a thorough analysis of CBITCL processes). The proposed approach is fully analytically tractable, due to the fact that CBITCL processes are affine processes and, therefore, their characteristic function can be explicitly characterized. Moreover, we will show that CBITCL processes are *coherent* in the sense of Gnoatto (2017), meaning that if an FX rate is modeled by a CBITCL process, then its reciprocal also belongs to the same model class.

We construct our modeling framework by adopting an *artificial currency approach* (see also Sect. 1.1 below), which consists in modeling each FX rate as the ratio of two primitive processes, with one primitive process associated to each currency. FX rates then satisfy the inversion and triangulation symmetries by construction and the model formulation reduces to modeling all primitive processes by means of a common family of CBITCL processes. By relying on a Girsanov-type result for CBITCL processes, we characterize a class of risk-neutral measures that leave invariant the structure of the model. In particular, this allows preserving the CBITCL property under equivalent changes of probability, which enables us to derive an efficient pricing formula for currency options by means of Fourier techniques.

We analyze the empirical performance of a two-dimensional specification of our framework, driven by tempered α-stable CBI processes (as recently introduced in Fontana et al., 2021) and CGMY processes (see Carr et al., 2002). We perform a calibration of the model with respect to an FX triangle consisting of three major currency pairs (USD-JPY, EUR-JPY,

---

[1] We recall that a risk-reversal, in the context of FX options, measures the difference in implied volatility between an OTM call option and a put option with the same characteristics and a symmetric delta.
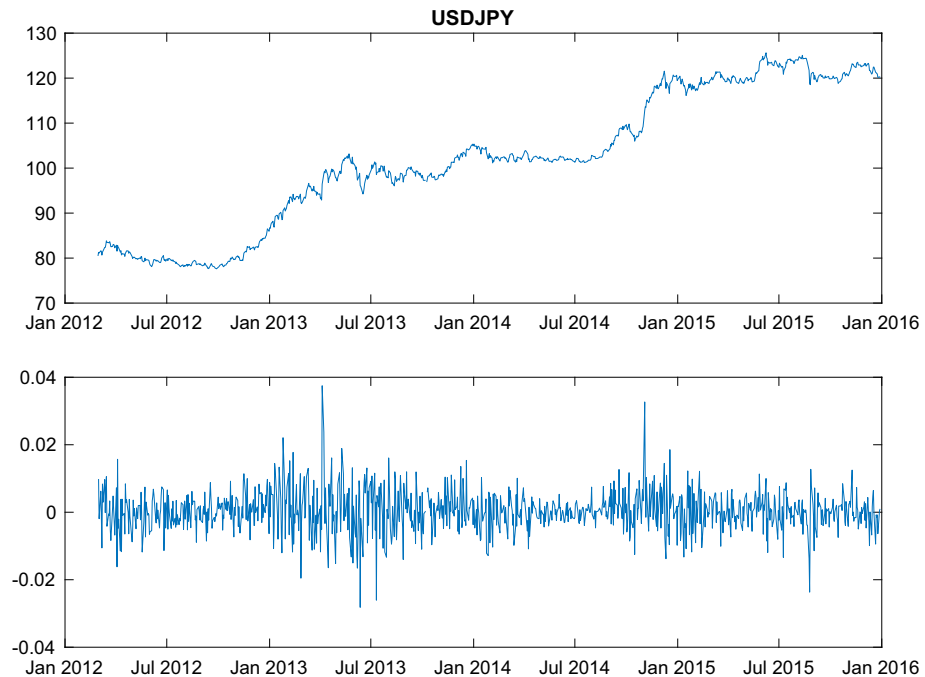
Fig. 1 Time series of the USD-JPY exchange rate (upper panel) and its return (lower panel), period 01/01/2012 - 31/12/2015 (source: Bloomberg)

EUR-USD). We propose two calibration methods: a *standard* calibration algorithm and a *deep* calibration algorithm, inspired by the deep learning techniques recently developed in Horvath et al. (2021) and applied here for the first time in a multi-currency setting. We assess the importance of jumps by showing that our CBITCL model achieves a superior calibration performance with respect to an analogous continuous-path model.

1.1 Related literature

As mentioned above, our framework is based on an *artificial currency approach*, which goes back to the works of Flesaker and Hughston (2000) and Doust (2007) (*intrinsic currency approach*, in the terminology of Doust, 2007). Relying on this approach, several stochastic volatility models for multiple currencies have been developed, see e.g. (Doust, 2012; De Col et al., 2013; Gnoatto and Grasselli, 2014; Baldeaux et al., 2015; Gnoatto et al., 2021) in a Brownian setting. In the latter works, the symmetries of FX rates are respected and several sources of risk can be adequately represented, with the important exceptions of jump risk, volatility clustering and self-excitation effects, which will play an important role in our framework.

Most of the works mentioned in the previous paragraph can be regarded as multi-currency extensions of the Heston model (to this effect, see also Janek et al., 2011). This is motivated by the fact that the Heston model is known to be stable under inversion (see, e.g., Rollin, 2008). The property that a certain model class is invariant under inversion has been termed *coherence* in Gnoatto (2017), where it has been shown that general affine stochastic volatility

models are coherent. As will be shown below, our modeling framework is coherent in this sense. The coherence property has been recently studied by Graceffa et al. (2020) (under the name of *consistency*) in the class of jump-diffusion models.

Beyond the Brownian setting, some important contributions to the modeling of FX rates include the work of Eberlein and Koval (2006) based on time-inhomogeneous Lévy processes and the work of Carr and Wu (2007) based on time-changed Lévy processes with CIR-type activity rates. More recently, Ballotta et al. (2017) have developed a multi-currency framework driven by a multi-dimensional Lévy process with dependent components. In this work, only standard Lévy processes are considered and, therefore, stochastic volatility is not explicitly modeled. Ballotta and Morico (2017) suggest however the possible use of time change methods. This line of research is pursued in Ballotta and Morico (2018) and further expanded in our work. In particular, Ballotta and Morico (2018) propose a model based on time-changed Lévy processes where the activity rate can have jumps and self-excitation. The presence of common jumps between the Lévy process and the activity rate induces a nontrivial dependence between the FX rate and its volatility. The model of Ballotta and Morico (2018) is however limited to a single FX rate, in which case the FX symmetries do not play any role. In contrast, we propose a coherent multi-currency framework that satisfies the FX symmetries and exhibits rich and flexible stochastic dynamics for all FX rates and their volatilities, while preserving an analytical tractability comparable to the model of Ballotta and Morico (2018).

1.2 Outline of the paper

The paper is structured as follows. In Sect. 2 we recall some basic results on CBITCL processes. In Sect. 3 we develop the multi-currency modeling framework, while in Sect. 4 we present two calibration methods and analyze the empirical fit to market data on an FX triangle. Finally, Sect. 5 concludes the paper.

2 CBI-time-changed Lévy processes

In this section, we present some fundamental results on *CBI-time-changed Lévy (CBITCL) processes*, referring to Fontana et al. (2022) for a complete theoretical analysis of this class of processes, detailed proofs and additional results (see also Szulda 2021, Chapter 4). We work on a filtered stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, where $\mathbb{F}$ is a filtration satisfying the usual conditions.

2.1 Definition and characterization of CBITCL processes

Let us start by recalling the definition of a *Continuous-state Branching processes with Immigration* (see Li 2020, Section 5).

- Let the function $\Psi : \mathbb{R}_- \to \mathbb{R}$ be given by

$$\Psi(x) := \beta x + \int_0^{+\infty} (e^{xz} - 1) \, \nu(\mathrm{d}z), \qquad \forall x \in \mathbb{R}_-, \qquad (2.1)$$

where $\beta \geq 0$ and ν is a Lévy measure on $(0, +\infty)$ such that $\int_0^1 z \, \nu(\mathrm{d}z) < +\infty$;

- Let the function $\Phi : \mathbb{R}_- \to \mathbb{R}$ be given by

$$\Phi(x) := -bx + \frac{\sigma^2}{2} x^2 + \int_0^{+\infty} (e^{xz} - 1 - xz) \, \pi(\mathrm{d}z), \qquad \forall x \in \mathbb{R}_-, \qquad (2.2)$$

where $b \in \mathbb{R}, \sigma \in \mathbb{R}$ and π is a Lévy measure on $(0, +\infty)$ such that $\int_1^{+\infty} z \, \pi(\mathrm{d}z) < +\infty$.

Definition 2.1 A Markov process $X = (X_t)_{t \geq 0}$ with initial value X_0 and state space $\mathbb{R}_+$ is said to be a *Continuous-state Branching process with Immigration* (CBI) with immigration mechanism Ψ and branching mechanism Φ if its Laplace transform is given by

$$\mathbb{E}[e^{uX_T}] = \exp\left(\int_0^T \Psi\big(\mathcal{V}(s, u)\big) \, \mathrm{d}s + \mathcal{V}(T, u) X_0 \right),$$

for all $u \in \mathbb{R}_-$ and $T \in \mathbb{R}_+$, where the function $\mathcal{V}(\cdot, u) : \mathbb{R}_+ \to \mathbb{R}_-$ is the unique solution to

$$\frac{\partial \mathcal{V}}{\partial t}(t, u) = \Phi\big(\mathcal{V}(t, u)\big), \qquad \mathcal{V}(0, u) = u.$$

Definition 2.1 corresponds to a conservative stochastically continuous CBI process in the sense of Kawazu and Watanabe (1971). Note that CBI processes are non-negative, strongly Markov (Feller) and with càdlàg trajectories. As a consequence, the path integral $Y := \int_0^{\cdot} X_s \, \mathrm{d}s$ of a CBI process X is always well defined as a non-decreasing process. It can therefore be used as a finite continuous time-change. This motivates the following definition.

Definition 2.2 A process $(X, Z) = ((X_t, Z_t))_{t \geq 0}$ is said to be a *CBI-time-changed Lévy process* (CBITCL process) if

(i) X is a CBI process, and
(ii) $Z = L_Y$, where $L = (L_t)_{t \geq 0}$ is a Lévy process independent of X and $Y = (Y_t)_{t \geq 0}$ denotes the process defined by $Y_t := \int_0^t X_s \, \mathrm{d}s$, for all $t \in \mathbb{R}_+$.

The Lévy exponent Ξ of the Lévy process L admits the Lévy-Khintchine representation

$$\Xi(u) := b_Z u + \frac{\sigma_Z^2}{2} u^2 + \int_{\mathbb{R} \setminus \{0\}} \left(e^{zu} - 1 - zu \mathbf{1}_{\{|z| < 1\}} \right) \gamma_Z(\mathrm{d}z), \qquad \forall u \in i\mathbb{R},$$

where $(b_Z, \sigma_Z, \gamma_Z)$ is the Lévy triplet of L, with $b_Z \in \mathbb{R}, \sigma_Z \in \mathbb{R}$ and γ_Z a Lévy measure on $\mathbb{R}$. In the following, we shall write CBITCL(X_0, Ψ, Φ, Ξ) to denote that a process (X, Z) is a CBI-time-changed Lévy process in the sense of Definition 2.2, where Ψ and Φ denote respectively the immigration and branching mechanisms of the CBI process X and Ξ the Lévy exponent of L.

CBI processes admit a characterization in terms of a Lamperti representation (see Caballero et al., 2013, and Szulda, 2021, Chapter 2). However, it turns out that there exists an equivalent representation that is better suited to our purposes, in terms of solutions to certain stochastic integral equations of the Dawson-Li type (see Dawson and Li, 2006). To this effect, let us introduce the following objects:

- two Brownian motions $B^1 = (B_t^1)_{t \geq 0}$ and $B^2 = (B_t^2)_{t \geq 0}$;
- a Poisson random measure $N_0(\mathrm{d}t, \mathrm{d}x)$ on $(0, +\infty)^2$ with compensator $\mathrm{d}t \, \nu(\mathrm{d}x)$ and compensated measure $\widetilde{N}_0(\mathrm{d}t, \mathrm{d}x) := N_0(\mathrm{d}t, \mathrm{d}x) - \mathrm{d}t \, \nu(\mathrm{d}x)$;
- a Poisson random measure $N_1(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x)$ on $(0, +\infty)^3$ with compensator $\mathrm{d}t \, \mathrm{d}u \, \pi(\mathrm{d}x)$ and compensated measure $\widetilde{N}_1(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) := N_1(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) - \mathrm{d}t \, \mathrm{d}u \, \pi(\mathrm{d}x)$;

- a Poisson random measure $N_2(dt, du, dx)$ on $(0, +\infty)^2 \times \mathbb{R}$ with compensator $dt\, du\, \gamma_Z(dx)$ and compensated measure $\widetilde{N}_2(dt, du, dx) := N_2(dt, du, dx) - dt\, du\, \gamma_Z(dx)$.

We furthermore assume that B^1, B^2, N_0, N_1 and N_2 are mutually independent. For any $X_0 \in \mathbb{R}_+$, let us consider the following stochastic integral equations:

$$X_t = X_0 + \int_0^t (\beta - b\, X_s)\, ds + \sigma \int_0^t \sqrt{X_s}\, dB_s^1$$
$$+ \int_0^t \int_0^{+\infty} x\, N_0(ds, dx) + \int_0^t \int_0^{X_{s-}} \int_0^{+\infty} x\, \widetilde{N}_1(ds, du, dx), \qquad (2.3)$$

$$Z_t = b_Z \int_0^t X_s\, ds + \sigma_Z \int_0^t \sqrt{X_s}\, dB_s^2 + \int_0^t \int_0^{X_{s-}} \int_{|x| \geq 1} x\, N_2(ds, du, dx)$$
$$+ \int_0^t \int_0^{X_{s-}} \int_{|x| < 1} x\, \widetilde{N}_2(ds, du, dx). \qquad (2.4)$$

The connection between Definition 2.2 and the stochastic integral equations (2.3)-(2.4) is given in the following proposition (see Fontana et al., 2022, Theorem 2.3).

Proposition 2.3 *A process (X, Z) with initial value $(X_0, 0)$ is a CBITCL(X_0, Ψ, Φ, Ξ) process if and only if it is a weak solution to the stochastic integral equations (2.3)-(2.4).*

On a given stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, there exists a unique strong solution to (2.3)-(2.4). Indeed, there is a unique strong solution $X = (X_t)_{t \geq 0}$ to (2.3), which corresponds to the *Dawson-Li representation* of a CBI process (see Dawson and Li, 2006, Theorems 5.1 and 5.2). In turn, since the right-hand side of (2.4) does depend only on the process $X = (X_t)_{t \geq 0}$ and not on $Z = (Z_t)_{t \geq 0}$, this obviously implies the existence of a unique strong solution $Z = (Z_t)_{t \geq 0}$ to (2.4) as well. In the following, if a CBITCL process (X, Z) is directly defined as the unique strong solution to (2.3)-(2.4), we will say that (X, Z) is defined through its *extended Dawson-Li representation* (see Fontana et al., 2022, Section 2.1).

Remark 2.4 The system of stochastic integral equations (2.3)-(2.4) makes evident the self-exciting behavior of a CBITCL process. More specifically, we can observe the following:

- For the CBI process X, the domain of integration of the integral with respect to $\widetilde{N}_1$ depends on the value of the process itself. This generates a self-exciting effect since, when a jump occurs, the jump intensity of X increases. In turn, this increases the likelihood of subsequent jumps of X, thereby generating jump clustering phenomena.
- The volatility components of Z depend on the value of the process X. Therefore, large values of X increase the volatility of the process Z, thereby increasing the likelihood of volatility clusters in the dynamics of Z as well as joint clusters between X and Z.

2.2 Affine property and changes of probability

The next proposition shows that CBITCL processes are affine and provides an explicit characterization of the Laplace-Fourier transform.

Proposition 2.5 *Let (X, Z) be a CBITCL(X_0, Ψ, Φ, Ξ) process and consider the joint process (X, Y, Z), where $Y := \int_0^\cdot X_s\, ds$. Then, the process (X, Y, Z) is an affine process on the state space $\mathbb{R}_+^2 \times \mathbb{R}$ with conditional Laplace-Fourier transform given by*

$$\mathbb{E}\big[e^{u_1 X_T + u_2 Y_T + u_3 Z_T} \mid \mathcal{F}_t\big] = \exp\Big(\mathcal{U}(T-t, u_1, u_2, u_3) + \mathcal{V}(T-t, u_1, u_2, u_3)\, X_t$$
$$+ u_2\, Y_t + u_3\, Z_t\Big),$$

for all $(u_1, u_2, u_3) \in \mathbb{C}_-^2 \times i\mathbb{R}$ *and* $0 \le t \le T < +\infty$, *where the functions* $\mathcal{U}(\cdot, u_1, u_2, u_3) :$ $\mathbb{R}_+ \to \mathbb{C}$ *and* $\mathcal{V}(\cdot, u_1, u_2, u_3) : \mathbb{R}_+ \to \mathbb{C}_-$ *are solutions to*

$$\mathcal{U}(t, u_1, u_2, u_3) = \int_0^t \Psi\big(\mathcal{V}(s, u_1, u_2, u_3)\big)\, \mathrm{d}s, \tag{2.5}$$

$$\frac{\partial \mathcal{V}}{\partial t}(t, u_1, u_2, u_3) = \Phi\big(\mathcal{V}(t, u_1, u_2, u_3)\big) + u_2 + \Xi(u_3), \qquad \mathcal{V}(0, u_1, u_2, u_3) = u_1, \tag{2.6}$$

where $\Psi : \mathbb{C}_- \to \mathbb{C}$ *and* $\Phi : \mathbb{C}_- \to \mathbb{C}$ *denote the analytic extensions to* $\mathbb{C}_-$ *of the corresponding functions defined in* (2.1) *and* (2.2), *respectively.*

Proof By Duffie et al. (2003), Corollary 2.10, the process X is an affine process. The affine property of (X, Y, Z) and the characterization of its conditional Laplace-Fourier transform then follow by an application of Keller-Ressel (2009), Theorems 4.10 and 4.16, (see Fontana et al., 2022, Section 2.2 for additional details).

Remark 2.6 In view of Proposition 2.5, CBITCL processes can be viewed as *affine stochastic volatility models* in the sense of Keller-Ressel (2009), Chapter 5. In the context of FX modeling, such models have been proved to be *coherent* by Gnoatto (2017), as mentioned in the Introduction. This fact will play an important role in the construction of our multi-currency framework in Sect. 3. We refer to Fontana et al. (2022), Section 2.2, for a detailed analysis of the relation between CBITCL processes and affine stochastic volatility models.

We close this section by describing a class of equivalent changes of probability of Esscher type that leave invariant the CBITCL structure. Let us first define the convex set $\mathcal{D}_X$ as follows:

$$\mathcal{D}_X := \left\{ x \in \mathbb{R} : \int_1^{+\infty} e^{xz}(\nu + \pi)(\mathrm{d}z) < +\infty \right\}. \tag{2.7}$$

The set $\mathcal{D}_X$ is the effective domain of the functions Ψ and Φ, which can be extended as finite-valued convex functions on $\mathcal{D}_X$. Note that $\mathcal{D}_X$ also represents the extended domain of the Laplace transform of the CBI process X (see Fontana et al. 2021, Theorem 2.6). Let us also introduce the convex set

$$\mathcal{D}_Z := \left\{ x \in \mathbb{R} : \int_{|z| \ge 1} e^{xz} \gamma_Z(\mathrm{d}z) < +\infty \right\}, \tag{2.8}$$

which represents the effective domain of the Lévy exponent Ξ when restricted to real arguments. Let us fix $\zeta \in \mathbb{R}$ and $\lambda \in \mathbb{R}$ and consider the process $\mathcal{W} = (\mathcal{W}_t)_{t \ge 0}$ defined by

$$\mathcal{W}_t := \zeta\, (X_t - X_0) + \lambda\, Z_t, \qquad \text{for all } t \in \mathbb{R}_+. \tag{2.9}$$

By Jacod and Shiryaev, 2003, Proposition II.8.26, it can be checked that $\mathcal{W}$ is an exponentially special semimartingale if and only if $\zeta \in \mathcal{D}_X$ and $\lambda \in \mathcal{D}_Z$. In this case, $\mathcal{W}$ admits a unique exponential compensator, i.e., a predictable process of finite variation, denoted by $\mathcal{K} = (\mathcal{K}_t)_{t \ge 0}$, such that $\exp(\mathcal{W} - \mathcal{K})$ is a local martingale (see Kallsen and Shiryaev, 2002). The following lemma provides the explicit expression of $\mathcal{K}$.

Table 1 Parameter transformations from $\mathbb{Q}$ to $\mathbb{Q}'$ for the CBITCL process (X, Z)	CBITCL parameters under $\mathbb{Q}'$
	$\beta' := \beta$
	$\nu'(\mathrm{d}z) := e^{\zeta z} \nu(\mathrm{d}z)$
	$b' := b - \zeta \sigma^2 - \int_0^{+\infty} z(e^{\zeta z} - 1)\pi(\mathrm{d}z)$
	$\sigma' := \sigma$
	$\pi'(\mathrm{d}z) := e^{\zeta z}\pi(\mathrm{d}z)$
	$b_Z' := b_Z + \lambda \sigma_Z^2 + \int_{|z|<1} z(e^{\lambda z} - 1)\,\gamma_Z(\mathrm{d}z)$
	$\sigma_Z' := \sigma_Z$
	$\gamma_Z'(\mathrm{d}z) := e^{\lambda z}\gamma_Z(\mathrm{d}z)$

Lemma 2.7 *Let (X, Z) be a $\mathrm{CBITCL}(X_0, \Psi, \Phi, \Xi)$ process. Consider the process $\mathcal{W}$ defined in* (2.9), *with $\zeta \in \mathcal{D}_X$ and $\lambda \in \mathcal{D}_Z$. Then, the exponential compensator $\mathcal{K}$ of $\mathcal{W}$ is given by*

$$\mathcal{K}_t = t\,\Psi(\zeta) + Y_t\big(\Phi(\zeta) + \Xi(\lambda)\big), \qquad for\ all\ t \in \mathbb{R}_+. \tag{2.10}$$

Proof For brevity of presentation, we only give a sketch of the proof, referring to Fontana et al. (2022), Lemma 4.1 for full details. Since CBITCL processes are quasi-left-continuous, the exponential compensator $\mathcal{K}$ can be explicitly computed in terms of the semimartingale differential characteristics of (X, Z) (see Kallsen and Shiryaev, 2002). In view of Proposition 2.3, the differential semimartingale characteristics of (X, Z) can be easily obtained from (2.3)-(2.4). Representation (2.10) then follows by standard computations. □

Fixing a time horizon $\mathcal{T} < +\infty$, we can state the following Girsanov-type result for CBITCL processes, which will play a central role in the modeling framework developed in the next section. In the following statement, we denote by $\mathcal{D}_X^\circ$ and $\mathcal{D}_Z^\circ$ the interior of sets $\mathcal{D}_X$ and $\mathcal{D}_Z$, respectively.

Theorem 2.8 *Let (X, Z) be a $\mathrm{CBITCL}(X_0, \Psi, \Phi, \Xi)$ process. Consider the process $\mathcal{W}$ defined in* (2.9), *with $\zeta \in \mathcal{D}_X^\circ$ and $\lambda \in \mathcal{D}_Z^\circ$, and its exponential compensator $\mathcal{K}$ given by* (2.10). *Then, the process $(\exp(\mathcal{W}_t - \mathcal{K}_t))_{t \in [0, \mathcal{T}]}$ is a martingale. Moreover, setting*

$$\frac{\mathrm{d}\mathbb{Q}'}{\mathrm{d}\mathbb{Q}} := e^{\mathcal{W}_\mathcal{T} - \mathcal{K}_\mathcal{T}}, \tag{2.11}$$

defines a probability measure $\mathbb{Q}' \sim \mathbb{Q}$ under which (X, Z) remains a CBITCL process up to time $\mathcal{T}$ with parameters β', ν', b', σ', π', b_Z', σ_Z' and γ_Z' given in Table 1.

Proof By definition of the exponential compensator, the process $\exp(\mathcal{W} - \mathcal{K})$ is a strictly positive local martingale. The true martingale property of the process $\exp(\mathcal{W}_t - \mathcal{K}_t)_{t \in [0, \mathcal{T}]}$ follows similarly as in Keller-Ressel and Mayerhofer (2015), Theorem 3.2, since $\zeta \in \mathcal{D}_X^\circ$ and $\lambda \in \mathcal{D}_Z^\circ$. The fact that (X, Z) is a CBITCL process up to time $\mathcal{T}$ with parameters given in Table 1 under $\mathbb{Q}'$ is a consequence of Girsanov's theorem together with Proposition 2.3 (see Fontana et al., 2022, Theorem 4.2 for full details). □

3 Modeling of multiple currencies via CBITCL processes

In this section, we present our modeling framework for a multi-currency market. In Sect. 3.1, we introduce the main quantities to be modeled together with the specific requirements

induced by absence of arbitrage and by the FX symmetries discussed in the Introduction. Sect. 3.2 contains the construction of the framework and the description of its most relevant features. In Sect. 3.3, Fourier techniques are applied to the pricing of currency options. We work on a filtered stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ and consider models defined up to a time horizon $\mathcal{T} < +\infty$.

3.1 Definition of the multiple currency market

The FX market involves different economies, each of them associated to a specific currency. The ith and j^{th} currencies are related by the spot FX rate process $S^{i,j}$, representing the value of one unit of currency j measured in units of currency i. Our definition of the multiple currency market will make use of the following ingredients, denoting by $N \in \mathbb{N}$, with $N \geq 2$, the number of economies (i.e., currencies) considered:

(i) $\boldsymbol{D} = \{D^i; i = 1, \ldots, N\}$ is an $\mathbb{R}^N_{>0}$-valued process, with D^i representing the bank account of the ith economy, for $i = 1, \ldots, N$;

(ii) $\boldsymbol{S} = \{S^{i,j}; i, j = 1, \ldots, N\}$ is an $\mathbb{R}^{N \times N}_{>0}$-valued process representing the spot FX rates between the N different currencies and such that $S^{i,i} \equiv 1$, for all $i = 1, \ldots, N$.

Definition 3.1 We say that the pair $(\boldsymbol{D}, \boldsymbol{S})$ represents a *multiple currency market* if, for every $i = 1, \ldots, N$, the following assets are traded in the ith economy:

- the bank account D^i;
- for every $j = 1, \ldots, N$ with $j \neq i$, the bank account of the jth economy denominated in units of the ith currency, namely $S^{i,j} D^j$.

We aim at constructing models for multiple currency markets that respect the FX symmetries mentioned in the Introduction and satisfy absence of arbitrage in the sense of *no free lunch with vanishing risk* (NFLVR). Since $(\boldsymbol{D}, \boldsymbol{S})$ are strictly positive processes, Delbaen and Schachermayer (1998), Theorem 1.1, implies that, for each $i = 1, \ldots, N$, NFLVR holds in the ith economy if and only if there exists a risk-neutral measure $\mathbb{Q}^i$, i.e., a probability measure $\mathbb{Q}^i$ equivalent to $\mathbb{Q}$ such that $S^{i,j} D^j / D^i$ is a local martingale under $\mathbb{Q}^i$. We then formulate the following definition, which extends Definition 1 of Escobar and Gschnaidtner (2018) to an FX market consisting of an arbitrary number N of currencies.

Definition 3.2 The multiple currency market $(\boldsymbol{D}, \boldsymbol{S})$ is said to be *well-posed* if the following hold:

(i) *no direct arbitrage*: $S^{j,i} = 1/S^{i,j}$, for all $i, j = 1, \ldots, N$;

(ii) *no triangular arbitrage*: $S^{i,j} = S^{i,k} \times S^{k,j}$, for all $i, k, j = 1, \ldots, N$;

(iii) there exists a risk-neutral measure $\mathbb{Q}^i$ for the ith economy, for all $i = 1, \ldots, N$.

Besides the requirement of well-posedness, we are interested in multi-currency models that are *coherent* in the sense of Gnoatto (2017). This means that, if the FX rate process $S^{i,j}$ belongs to a certain model class under $\mathbb{Q}^i$, then also its reciprocal $S^{j,i}$ belongs to the same model class under $\mathbb{Q}^j$, where $\mathbb{Q}^i$ and $\mathbb{Q}^j$ are risk-neutral measures for the ith and j^{th} economy, respectively. Obviously, *coherence* is a desirable property from the modeling perspective, since it ensures that the model retains its analytical tractability in all N different economies.

To achieve a well-posed as well as coherent multiple currency market, we will proceed as follows:

(1) Adopting the *artificial currency approach*, we express each currency with respect to an artificial currency indexed by 0 and construct the artificial FX rates $S^{0,i}$, for $i = 1, \ldots, N$.

(2) We define the FX rates $S^{i,j}$, for all $i, j = 1, \ldots, N$, by taking suitable ratios of the artificial FX rates $S^{0,i}$, $i = 1, \ldots, N$. Parts (i)-(ii) of Definition 3.2 are then satisfied by construction.

(3) By relying on Theorem 2.8, we construct a risk-neutral measure $\mathbb{Q}^i$, for every $i = 1, \ldots, N$, under which the driving processes remain CBITCL processes, thus ensuring part (iii) of Definition 3.2 as well as the stability of the structure of the model (coherence).

3.2 Construction of the modeling framework

The construction of our modeling framework starts by modeling the N artificial FX rates $S^{0,i}$, for $i = 1, \ldots, N$, by means of a common family of CBITCL processes. To this effect, we assume that the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ supports d mutually independent CBITCL processes (X^k, Z^k), $k = 1, \ldots, d$, defined through the corresponding extended Dawson-Li representations (2.3)-(2.4).[2] For each $k = 1, \ldots, d$, we denote by $\mathcal{D}_{X^k}$ and $\mathcal{D}_{Z^k}$ the sets (2.7) and (2.8), respectively, associated to the CBITCL process (X^k, Z^k).

For each $i = 1, \ldots, N$, we introduce the following parameters:

- $r^i \in \mathbb{R}$, representing the risk-free short rate in the ith economy and generating the bank account $D_t^i := \exp(r^i t)$, for all $t \in [0, \mathcal{T}]$;
- $\zeta^i = (\zeta_1^i, \ldots, \zeta_d^i) \in \mathbb{R}^d$ such that $\zeta_k^i \in \mathcal{D}_{X^k}^\circ$, for all $k = 1, \ldots, d$;
- $\lambda^i = (\lambda_1^i, \ldots, \lambda_d^i) \in \mathbb{R}^d$ such that $\lambda_k^i \in \mathcal{D}_{Z^k}^\circ$, for all $k = 1, \ldots, d$.

In addition, for each $i = 1, \ldots, N$ and $k = 1, \ldots, d$, we denote by $\mathcal{K}^{i,k} = (\mathcal{K}_t^{i,k})_{t \in [0,\mathcal{T}]}$ the exponential compensator of the process $(\zeta_k^i X^k + \lambda_k^i Z^k)$, as characterized in Lemma 2.7.

Remark 3.3 We point out that the modeling framework developed in this section can be easily generalized to the case of stochastic interest rates. In particular, by allowing the interest rates r^i, for $i = 1, \ldots, N$, to be driven by the common family of CBITCL processes (X^k, Z^k), $k = 1, \ldots, d$, one can introduce dependence between the interest rates, the FX rates and their volatilities.

For each $i = 1, \ldots, N$, we specify the artificial FX rate $S^{0,i}$ as follows:

$$S_t^{0,i} := S_0^{0,i} e^{-r^i t} \prod_{k=1}^{d} e^{\zeta_k^i X_t^k + \lambda_k^i Z_t^k - \mathcal{K}_t^{i,k}}, \qquad \text{for all } t \in [0, \mathcal{T}]. \tag{3.1}$$

The artificial FX rates are modeling quantities that cannot be observed in reality. However, the parameters ζ_k^i and λ_k^i will have a specific role in the dynamics of the actual FX rates. Indeed, λ_k^i will measure the relative importance of the risk arising from the kth time-changed Lévy process Z^k, while ζ_k^i will measure the dependence between the kth CBI process X^k and the ith FX rate, as will become clear from Lemma 3.8 below and the following discussion.

Lemma 3.4 *For each $i = 1, \ldots, N$, the process $S^{0,i} = (S_t^{0,i})_{t \in [0,\mathcal{T}]}$ satisfies the following dynamics:*

$$\frac{dS_t^{0,i}}{S_{t-}^{0,i}} = -r^i \, dt + \sum_{k=1}^{d} \left(\sqrt{X_t^k} \left(\zeta_k^i \, \sigma^k \, dB_t^{k,1} + \lambda_k^i \, \sigma_Z^k \, dB_t^{k,2} \right) + \int_0^{+\infty} (e^{\zeta_k^i x} - 1) \widetilde{N}_0^k(dt, dx) \right)$$

---

[2] In the following, we shall use the superscript k to denote all parameters and driving processes appearing in the extended Dawson-Li representation (2.3)-(2.4) of the CBITCL process (X^k, Z^k) on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, for $k = 1, \ldots, d$.

$$+ \sum_{k=1}^{d} \int_{0}^{X_{t-}^{k}} \left(\int_{0}^{+\infty} (e^{\zeta_k^i x} - 1) \widetilde{N}_1^k (\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) \right.$$

$$\left. + \int_{\mathbb{R}} (e^{\lambda_k^i x} - 1) \widetilde{N}_2^k (\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) \right). \tag{3.2}$$

Moreover, the process $S^{0,i} D^i = (S_t^{0,i} D_t^i)_{t \in [0,\mathcal{T}]}$ *is a martingale on* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, *for all* $i = 1, \dots, N$.

Proof Using specification (3.1) of $S^{0,i}$, equation (3.2) follows from an application of Itô's formula together with the extended Dawson-Li representation (2.3)-(2.4) of the process (X^k, Z^k), for $k = 1, \dots, d$. The martingale property of $S^{0,i} D^i$ follows from the independence of the CBITCL processes (X^k, Z^k), for $k = 1, \dots, d$, together with the martingale property stated in Theorem 2.8. □

We define the FX rate process $S = \{S^{i,j}; i, j = 1, \dots, N\}$ as follows:

$$S_t^{i,j} := \frac{S_t^{0,j}}{S_t^{0,i}}, \qquad \text{for all } i, j = 1, \dots, N \text{ and } t \in [0, \mathcal{T}]. \tag{3.3}$$

This specification of S ensures that the inversion and triangulation symmetries of FX rates (corresponding respectively to parts (i) and (ii) of Definition 3.2) are satisfied by construction, thereby completing steps (1) and (2) of the model construction outlined at the end of Sect. 3.1.

The next corollary describes a class of probability measures that leave invariant the structure of our multi-currency model driven by CBITCL processes. This result plays a crucial role in ensuring absence of arbitrage and coherence of our framework.

Corollary 3.5 *For each* $i = 1, \dots, N$, *setting*

$$\frac{\mathrm{d}\mathbb{Q}^i}{\mathrm{d}\mathbb{Q}} := \frac{S_{\mathcal{T}}^{0,i} D_{\mathcal{T}}^i}{S_0^{0,i}}, \tag{3.4}$$

defines a probability measure $\mathbb{Q}^i \sim \mathbb{Q}$ *under which* (X^k, Z^k), *for* $k = 1, \dots, d$, *remain mutually independent CBITCL processes (up to time* $\mathcal{T}$*) with parameters given in Table 2. Moreover, for each* $i = 1, \dots, N$, *the probability measure* $\mathbb{Q}^i$ *is a risk-neutral measure for the* i^{th} *economy.*

Proof In view of Lemma 3.4, $\mathbb{Q}^i$ is well-defined by (3.4) as a probability measure equivalent to $\mathbb{Q}$, for each $i = 1, \dots, N$. The fact that (X^k, Z^k), for all $k = 1, \dots, d$, remains a CBITCL process under $\mathbb{Q}^i$ with parameters given in Table 2 follows from Theorem 2.8 together with the independence of the processes (X^k, Z^k), for $k = 1, \dots, d$. Moreover, again the independence of the processes (X^k, Z^k), for $k = 1, \dots, d$, under $\mathbb{Q}$ together with the structure of the probability $\mathbb{Q}^i$ defined in (3.4) implies that the mutual independence is preserved under $\mathbb{Q}^i$. Finally, for all $i, j = 1, \dots, N$, in view of (3.3) and (3.4), the process $S^{i,j} D^j / D^i$ is a local martingale under $\mathbb{Q}^i$ if and only if $S^{0,j} D^j$ is a local martingale under $\mathbb{Q}$. Since the latter property holds by Lemma 3.4, the proof is complete. □

Remark 3.6 Financial models driven by CBITCL processes are inherently incomplete and, therefore, there exist infinitely many risk-neutral measures beyond the probability measures considered in Corollary 3.5. Our approach is motivated by the preservation of the structure of

Table 2 Parameter transformations from $\mathbb{Q}$ to $\mathbb{Q}^i$ for the CBITCL process (X^k, Z^k)

CBITCL parameters under $\mathbb{Q}^i$
$\beta^{i,k} := \beta^k$
$\nu^{i,k}(\mathrm{d}z) := e^{\zeta_k^i z} \nu^k(\mathrm{d}z)$
$b^{i,k} := b^k - \zeta_k^i (\sigma^k)^2 - \int_0^{+\infty} z(e^{\zeta_k^i z} - 1)\pi^k(\mathrm{d}z)$
$\sigma^{i,k} := \sigma^k$
$\pi^{i,k}(\mathrm{d}z) := e^{\zeta_k^i z}\pi^k(\mathrm{d}z)$
$b_Z^{i,k} := b_Z^k + \lambda_k^i (\sigma_Z^k)^2 + \int_{|z|<1} z(e^{\lambda_k^i z} - 1)\gamma_Z^k(\mathrm{d}z)$
$\sigma_Z^{i,k} := \sigma_Z^k$
$\gamma_Z^{i,k}(\mathrm{d}z) := e^{\lambda_k^i z}\gamma_Z^k(\mathrm{d}z)$

the model under each risk-neutral measure $\mathbb{Q}^i$, for $i = 1, \ldots, N$. In line with the martingale approach to financial modeling, the parameters characterizing the family $\{\mathbb{Q}^i; i = 1, \ldots, N\}$ are determined by calibration to market data (see Sect. 4 for a specific application). We refer to Eberlein and Kallsen (2020) for an overview of several well-known hedging approaches in incomplete markets driven by jump processes.

By combining (3.1) and (3.3), we obtain the following representation of FX rates:

$$S_t^{i,j} = S_0^{i,j} e^{(r^i - r^j)t} \prod_{k=1}^d e^{(\zeta_k^j - \zeta_k^i)X_t^k + (\lambda_k^j - \lambda_k^i)Z_t^k - (\mathcal{K}_t^{j,k} - \mathcal{K}_t^{i,k})}, \qquad \text{for all } t \in [0, \mathcal{T}]. \quad (3.5)$$

In particular, note that all FX rates $S^{i,j}$ share the same modeling structure, for all $i, j = 1, \ldots, N$, and the driving processes (X^k, Z^k), for $k = 1, \ldots, d$, remain mutually independent CBITCL processes under each risk-neutral measure $\mathbb{Q}^i$, as a consequence of Corollary 3.5. In particular, the functional form of the process $S^{i,j}$ under $\mathbb{Q}^i$ is identical to that of its reciprocal $S^{j,i}$ under $\mathbb{Q}^j$.

We have thus proved the next theorem, which shows that we have constructed a well-posed and coherent multiple currency market, in line with the modeling objectives set in Sect. 3.1.

Theorem 3.7 *The multiple currency market* (D, S) *is well-posed and coherent.*

For each $i = 1, \ldots, N$, the Radon-Nikodym density $\mathrm{d}\mathbb{Q}^i/\mathrm{d}\mathbb{Q}$ defined in (3.4) admits the following representation in terms of the sources of randomness driving the extended Dawson-Li representation (2.3)-(2.4) of the CBITCL processes (X^k, Z^k), for $k = 1, \ldots, d$:

$$\frac{\mathrm{d}\mathbb{Q}^i}{\mathrm{d}\mathbb{Q}} = \prod_{k=1}^d \mathcal{E}\left(\zeta_k^i \sigma^k \int_0^\cdot \sqrt{X_s^k}\,\mathrm{d}B_s^{k,1}\right.$$

$$\left. + \lambda_k^i \sigma_Z^k \int_0^\cdot \sqrt{X_s^k}\,\mathrm{d}B_s^{k,2} + \int_0^\cdot \int_0^{+\infty} (e^{\zeta_k^i x} - 1)\widetilde{N}_0^k(\mathrm{d}s, \mathrm{d}x)\right)_{\mathcal{T}}$$

$$\times \prod_{k=1}^d \mathcal{E}\left(\int_0^\cdot \int_0^{X_{s-}^k} \int_0^{+\infty} (e^{\zeta_k^i x} - 1)\widetilde{N}_1^k(\mathrm{d}s, \mathrm{d}u, \mathrm{d}x)\right.$$

$$\left. + \int_0^\cdot \int_0^{X_{s-}^k} \int_{\mathbb{R}} (e^{\lambda_k^i x} - 1)\widetilde{N}_2^k(\mathrm{d}s, \mathrm{d}u, \mathrm{d}x)\right)_{\mathcal{T}}.$$

By Girsanov's theorem, the processes $B^{i,k,1} = (B_t^{i,k,1})_{t \in [0, \mathcal{T}]}$ and $B^{i,k,2} = (B_t^{i,k,2})_{t \in [0, \mathcal{T}]}$ defined by

$$
\begin{aligned}
B_t^{i,k,1} &:= B_t^{k,1} - \zeta_k^i \, \sigma^k \int_0^t \sqrt{X_s^k} \, \mathrm{d}s, \\
B_t^{i,k,2} &:= B_t^{k,2} - \lambda_k^i \, \sigma_Z^k \int_0^t \sqrt{X_s^k} \, \mathrm{d}s,
\end{aligned}
\tag{3.6}
$$

are independent Brownian motions under $\mathbb{Q}^i$. Moreover, $N_0^k(\mathrm{d}t, \mathrm{d}x)$, $N_1^k(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x)$, $N_2^k(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x)$ are Poisson random measures under $\mathbb{Q}^i$ with compensated measures

$$
\begin{aligned}
\widetilde{N}_0^{i,k}(\mathrm{d}t, \mathrm{d}x) &:= N_0^k(\mathrm{d}t, \mathrm{d}x) - \mathrm{d}t \, e^{\zeta_k^i x} \, \nu^k(\mathrm{d}x), \\
\widetilde{N}_1^{i,k}(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) &:= N_1^k(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) - \mathrm{d}t \, \mathrm{d}u \, e^{\zeta_k^i x} \, \pi^k(\mathrm{d}x), \\
\widetilde{N}_2^{i,k}(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) &:= N_2^k(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) - \mathrm{d}t \, \mathrm{d}u \, e^{\lambda_k^i x} \, \gamma_Z^k(\mathrm{d}x).
\end{aligned}
\tag{3.7}
$$

Lemma 3.8 *For each $i, j = 1, \ldots, N$, the process $S^{i,j}$ satisfies the following dynamics under $\mathbb{Q}^i$:*

$$
\begin{aligned}
\frac{\mathrm{d}S_t^{i,j}}{S_{t-}^{i,j}} &= (r^i - r^j)\mathrm{d}t + \sum_{k=1}^d \sqrt{X_t^k} \left(\sigma^k (\zeta_k^j - \zeta_k^i) \, \mathrm{d}B_t^{i,k,1} + \sigma_Z^k (\lambda_k^j - \lambda_k^i) \, \mathrm{d}B_t^{i,k,2} \right) \\
&\quad + \sum_{k=1}^d \int_0^{+\infty} \left(e^{(\zeta_k^j - \zeta_k^i)x} - 1 \right) \widetilde{N}_0^{i,k}(\mathrm{d}t, \mathrm{d}x) \\
&\quad + \sum_{k=1}^d \int_0^{X_{t-}^k} \int_0^{+\infty} \left(e^{(\zeta_k^j - \zeta_k^i)x} - 1 \right) \widetilde{N}_1^{i,k}(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x) \\
&\quad + \sum_{k=1}^d \int_0^{X_{t-}^k} \int_{\mathbb{R}} \left(e^{(\lambda_k^j - \lambda_k^i)x} - 1 \right) \widetilde{N}_2^{i,k}(\mathrm{d}t, \mathrm{d}u, \mathrm{d}x),
\end{aligned}
\tag{3.8}
$$

with the processes $B^{i,k,1}$, $B^{i,k,2}$ and the random measures $\widetilde{N}_0^{i,k}$, $\widetilde{N}_1^{i,k}$, $\widetilde{N}_2^{i,k}$ defined in (3.6)-(3.7).

Proof The claim follows from (3.3) by applying Itô's product rule together with the dynamics (3.2) of the artificial FX rates $S^{0,i}$ and $S^{0,j}$, making use of the notation introduced above. □

In particular, we can notice that the dynamics of $S^{i,j}$ under $\mathbb{Q}^i$ are functionally symmetric with respect to the dynamics of $S^{j,i}$ under $\mathbb{Q}^j$, for all $i, j = 1, \ldots, N$. This is a further evidence of the coherence of our modeling framework, in the sense of Gnoatto (2017).

As can be seen from equation (3.8), in our framework FX rates possess stochastic dynamics that can capture the most relevant risk characteristics of the FX market, as discussed in the Introduction. More specifically, we can remark the following features (for simplicity of presentation, in the following discussion we consider a one-dimensional CBI process X):

Stochastic volatility: for all FX rates, both the diffusive volatility and the jump volatility are stochastic and depend on the current level of the CBI process X. In particular, since X is a self-exciting process, this induces volatility clustering effects in the FX rates.

Jump risk: all FX rates are affected by three different jump terms, corresponding to the three integrals appearing on the right-hand side of (3.8): the first integral results from the

immigration of the CBI process X, the second integral is related to the branching property of X and the third integral is generated by the Lévy process defining the process Z, with a jump intensity proportional to X. In particular, the last two integrals are affected by the self-exciting property of X and can generate jump clusters in the dynamics of FX rates. The parameters $\zeta^j - \zeta^i$ and $\lambda^j - \lambda^i$ control the magnitude of these effects.

Stochastic dependence of FX rates: the quadratic covariation among different FX rates exhibits a rich stochastic structure, with the presence of common jumps and with a jump intensity which is related to the current level of the self-exciting CBI process X.

Stochastic skewness: the quadratic covariation $[S^{i,j}, X]$ also has a rich stochastic structure. In turn, this induces a stochastic instantaneous correlation between each FX rate and its stochastic volatility. As explained in Christoffersen et al. (2009), Da Fonseca and Grasselli (2011), the presence of stochastic correlation is responsible for the existence of stochastic variations in the skew of FX implied volatilites.

3.3 Currency option pricing

The modeling framework constructed in Sect. 3.2 is coherent and, therefore, retains full analytical tractability under all risk-neutral measures considered in Corollary 3.5. In particular, we can derive an explicit representation of the characteristic function of each FX rate. In the following, we denote by $\mathbb{E}^i$ the expectation under $\mathbb{Q}^i$, for each $i = 1, \ldots, N$.

Lemma 3.9 *For all i, $j = 1, \ldots, N$, the characteristic function of the process $(\log S_t^{i,j})_{t \in [0,\mathcal{T}]}$ under $\mathbb{Q}^i$ is given by*

$$\mathbb{E}^i\left[e^{iu \log S_t^{i,j}}\right] = e^{iu(\log S_0^{i,j} + (r^i - r^j)t)} \prod_{k=1}^{d} e^{iu(\Psi^k(\zeta_k^i) - \Psi^k(\zeta_k^j))t + \mathcal{U}^{i,k}(t,u_1^k,u_2^k,u_3^k) + \mathcal{V}^{i,k}(t,u_1^k,u_2^k,u_3^k)X_0^k},$$

(3.9)

for all $(u, t) \in \mathbb{R} \times [0, \mathcal{T}]$, where $(\mathcal{U}^{i,k}(\cdot, u_1^k, u_2^k, u_3^k), \mathcal{V}^{i,k}(\cdot, u_1^k, u_2^k, u_3^k))$ is the unique solution to system (2.5)-(2.6) associated to (X^k, Z^k) under $\mathbb{Q}^i$ with

$$u_1^k = iu(\zeta_k^j - \zeta_k^i), \quad u_2^k = iu\left(\Phi^k(\zeta_k^i) + \Xi_Z^k(\lambda_k^i) - \Phi^k(\zeta_k^j) - \Xi_Z^k(\lambda_k^j)\right) \quad and \quad u_3^k = iu(\lambda_k^j - \lambda_k^i).$$

Proof In view of (3.5) and (2.10), we have that

$$\mathbb{E}^i\left[e^{iu \log S_t^{i,j}}\right] = e^{iu(\log S_0^{i,j} + (r^i - r^j)t)} \prod_{k=1}^{d} e^{iu(\Psi^k(\zeta_k^i) - \Psi^k(\zeta_k^j))t}$$

$$\times \prod_{k=1}^{d} \mathbb{E}^i\left[e^{iu(\zeta_k^j - \zeta_k^i)X_t^k + iu(\Phi^k(\zeta_k^i) + \Xi_Z^k(\lambda_k^i) - \Phi^k(\zeta_k^j) - \Xi_Z^k(\lambda_k^j))Y_t^k + iu(\lambda_k^j - \lambda_k^i)Z_t^k}\right],$$

where we have used the independence of the CBITCL processes (X^k, Z^k), for $k = 1, \ldots, d$, under $\mathbb{Q}^i$ (see Corollary 3.5) and $Y_t^k = \int_0^t X_s^k \, \mathrm{d}s$, for all $k = 1, \ldots, d$. Formula (3.9) then follows from an application of the affine transform formula given in Proposition 2.5. □

The availability of an explicit description of the characteristic function of each FX rate allows for currency option pricing via Fourier techniques. We adopt the COS method of Fang and Oosterlee (2009), which presents the advantage of utilizing only the characteristic function of the process, without requiring any domain extensions as in other Fourier pricing

methods. In our setting, such domain extensions would necessitate additional constraints on the parameters, potentially affecting the calibration results. We consider a European Call option in the ith economy written on the FX rate $S^{i,j}$, with maturity $T \leq \mathcal{T}$ and strike $K > 0$. We assume that the distribution of $\log S_T^{i,j}$ under $\mathbb{Q}^i$ admits a density[3]. Since the multiple currency market is well-posed, we can apply risk-neutral valuation under $\mathbb{Q}^i$ to compute the arbitrage-free price $C(T, K)$ at $t = 0$ of the option:

$$C(T, K) = e^{-r^i T} \, \mathbb{E}^i \left[(S_T^{i,j} - K)^+ \right] = e^{-r^i T} K \int_{\mathbb{R}} (e^x - 1)^+ f_T^{i,j}(x) \, \mathrm{d}x, \qquad (3.10)$$

where $f_T^{i,j}$ represents the density function of $\log(S_T^{i,j}/K)$ under $\mathbb{Q}^i$. To compute the integral in (3.10), we introduce a suitably chosen truncation range $[a, b] \subset \mathbb{R}$ such that $C(T, K)$ can be approximated with good accuracy by

$$C(T, K) \approx e^{-r^i T} K \int_a^b (e^x - 1)^+ f_T^{i,j}(x) \, \mathrm{d}x.$$

The resulting pricing formula is stated in the next proposition, which follows by the same arguments presented in Section 2.1 of Fang and Oosterlee (2009).

Proposition 3.10 *The arbitrage-free price $C(T, K)$ of a European call option written on the FX rate $S^{i,j}$, with maturity $T \leq \mathcal{T}$ and strike $K > 0$, can be approximated by*

$$C(T, K) \approx e^{-r^i T} K \sum_{k=0}^{M-1} \left(1 - \frac{\delta_0(k)}{2} \right) \Re \left(e^{i \frac{k\pi}{a-b}(a + \log K)} \mathbb{E}^i \left[e^{i \frac{k\pi}{b-a} \log S_T^{i,j}} \right] \right) B_k, \quad (3.11)$$

where δ_0 denotes the Kronecker delta at 0, $M \in \mathbb{N}$, $B_0 = (e^b - 1 - b)/(b - 1)$, and where

$$B_k = \frac{2}{b-a} \left(\frac{1}{1 + \left(\frac{k\pi}{b-a} \right)^2} \left((-1)^k e^b - \cos \left(\frac{k\pi a}{b-a} \right) + \frac{k\pi}{b-a} \sin \left(\frac{k\pi a}{b-a} \right) \right) \right.$$

$$\left. - \frac{b-a}{k\pi} \sin \left(\frac{k\pi a}{b-a} \right) \right).$$

Remark 3.11 (1) In order to ensure the accuracy of formula (3.11), one needs to specify the truncation range $[a, b]$ properly. Following Section 5.1 of Fang and Oosterlee (2009), a suitable specification is the following:

$$[a, b] = \left[c_1 - L \sqrt{c_2 + \sqrt{c_4}}, \quad c_1 + L \sqrt{c_2 + \sqrt{c_4}} \right],$$

with $L = 10$ and where c_n, for $n = 1, 2, 4$, represents the nth cumulant of $\log(S_T^{i,j}/K)$. In our framework, the cumulants are not available in closed form. However, they can be approximated by using finite differences since, by definition, they are given by the derivatives at zero of the cumulant-generating function of $\log(S_T^{i,j}/K)$ (see Fang and Oosterlee 2009, Appendix A for further details).

---

[3] By (Williams 1991, Theorem 16.6), the random variable $\log S_T^{i,j}$ admits a density under $\mathbb{Q}^i$ if $\int_{\mathbb{R}} |\mathbb{E}^i[e^{iu \log S_T^{i,j}}]| \mathrm{d}u < +\infty$. Under this assumption, the density can be recovered by Fourier inversion from the characteristic function of $\log S_T^{i,j}$ given in Lemma 3.9.

(2) As explained in Section 3.3 of Fang and Oosterlee (2009), formula (3.11) can be readily extended to a multi-strike setting, which is practically important when one needs to price several options with the same maturity but associated to different strikes (e.g., for model calibration). We refer to Szulda (2021), Remark 5.13, for a description of the multi-strike implementation of formula (3.11).

4 Model calibration

In this section, we calibrate a simple specification of our modeling framework to market data on a currency triangle. We consider a model driven by tempered α-stable CBI processes and CGMY Lévy processes (see Fontana et al., 2021) and propose two different calibration methods, one based on standard techniques and one relying on a deep learning algorithm. The market data are described in Sect. 4.1, while the two calibration methods are presented in Sect. 4.2. Sect. 4.3 contains a description of the model specification and in Sect. 4.4 we report the calibration results.

4.1 FX market data

We consider market data on three FX implied volatility surfaces: EUR-USD, EUR-JPY and USD-JPY (according to the FOR-DOM convention, the second currency of each pair represents the domestic currency). The quoting convention for FX implied volatilities differs from the case of equity markets, since implied volatilities are quoted in terms of deltas and maturities instead of strikes and maturities. Moreover, excluding ATM options, individual volatilities are not directly quoted: the market practice consists in quoting certain combinations of contracts (risk-reversals and butterflies) from which implied volatilities for single contracts in terms of maturities and deltas have to be recovered.

For the three volatility surfaces, we consider a common set of maturities, ranging from one week to one year (1, 2 weeks, 1, 3, 6 months, and 1 year, representing the most liquid part of the implied volatility surface). We retrieved from Bloomberg the following market quotes as of April 15, 2020: ATM implied volatility, 10Δ and 25Δ risk-reversals[4] and butterflies. For 25Δ, we have

$$RR_{25\Delta} = \sigma_{25\Delta Call} - \sigma_{25\Delta Put}, \quad \text{and} \quad BF_{25\Delta} = \frac{\sigma_{25\Delta Call} + \sigma_{25\Delta Put}}{2} - \sigma_{ATM},$$

from which we deduce

$$\sigma_{25\Delta Call} = \sigma_{ATM} + \frac{1}{2} RR_{25\Delta} + BF_{25\Delta} \quad \text{and} \quad \sigma_{25\Delta Put} = \sigma_{ATM} - \frac{1}{2} RR_{25\Delta} + BF_{25\Delta},$$

and similarly for 10Δ. For each currency pair and for each maturity, we have the implied volatilities of 5 contracts at our disposal. Market data not corresponding to the 5 points is interpolated.

In order to reconstruct observed market prices, we also retrieved from Bloomberg FX spots and FX forward points, which enable us to build FX forward curves by adding the spot and the forward points. Equipped with such data, we have all the information needed

---

[4] By 25Δ risk-reversal, we mean an OTM Call option with a delta of 25% and a Put option with a delta of -25%.

to convert deltas into strikes and recover implied volatilities for single contracts in terms of maturities and strikes.[5]

4.2 Two calibration methods

Let p denote a vector of model parameters, belonging to some set of admissible parameters $\mathcal{P}$. Let $\#T$ be the number of maturities and $\#K$ be the number of strikes that we consider. For simplicity of presentation, we assume that all volatility surfaces have the same strike range and the same number of strikes. In general, a calibration to the implied volatilites on a set of N currencies consists in solving the following minimization problem:

$$\min_{p \in \mathcal{P}} \sum_{u=1}^{N} \sum_{i=1}^{\#T} \sum_{j=1}^{\#K} \left(\sigma_{imp}^{mkt}(u, T_i, K_j) - \sigma_{imp}^{mod(p)}(u, T_i, K_j) \right)^2, \qquad (4.1)$$

where $\sigma_{imp}^{mkt}(u, T_i, K_j)$ denotes the implied volatility observed on the market for currency u, maturity T_i, and strike K_j, while $\sigma_{imp}^{mod(p)}(u, T_i, K_j)$ denotes its model-implied counterpart for a given vector of parameters $p \in \mathcal{P}$.

We now present two calibration methods. The first one, to which we refer as *standard calibration*, utilizes pricing formula (3.11) to compute model prices for a given choice of model parameters. Such prices are then converted into model-implied volatilities and inserted into (4.1). This gives rise to a multi-dimensional function $\Sigma : \mathcal{P} \rightarrow \mathbb{R}^{N \times \#T \times \#K}$ such that, for all $p \in \mathcal{P}$, we have $\Sigma(p)_{(u,i,j)} = \sigma_{imp}^{mod(p)}(u, T_i, K_j)$, for every $u = 1, \ldots, N$, $i = 1, \ldots, \#T$, and $j = 1, \ldots, \#K$.

The second calibration method, to which we refer as *deep calibration*, adopts the two-step approach developed by Horvath et al. (2021) for the solution of (4.1). We proceed as follows.

Grid-based implicit training: the purpose of this step is to approximate the non-linear function Σ by a fully-connected feed-forward neural network $\mathcal{N}^w : \mathcal{P} \rightarrow \mathbb{R}^{N \times \#T \times \#K}$ (see Horvath et al. 2021, Definition 1), where w denotes a vector of network parameters (typically weights and biases). We divide this step into two sub-steps:

(1) We generate a training set $\{(p_n, \Sigma(p_n))\}_{n=1,\ldots,N_{train}}$ of size N_{train}, where each vector of parameters p_n is generated randomly by means of a standard random generator (suitable adjustments can be made to guarantee that parameter restrictions are satisfied), and where we have fixed the grid (u, T_i, K_j), $u = 1, \ldots, N$, $i = 1, \ldots, \#T$, and $j = 1, \ldots, \#K$, throughout the generation (hence the term "grid-based").

(2) We solve the following minimization problem called "training" of the neural network:

$$\min_{w} \sum_{n=1}^{N_{train}} \sum_{u=1}^{N} \sum_{i=1}^{\#T} \sum_{j=1}^{\#K} \left(\Sigma(p_n)_{(u,i,j)} - \mathcal{N}^w(p_n)_{(u,i,j)} \right)^2, \qquad (4.2)$$

whose solution is an optimal vector of network parameters $\widehat{w}$ such that the neural network $\mathcal{N} := \mathcal{N}^{\widehat{w}}$ best approximates the observations $\{\Sigma(p_n)\}_{n=1,\ldots,N_{train}}$. Notice that $\widehat{w}$ depends on the grid that we have fixed, thus explaining the term "implicit".

---

[5] We performed these tasks by using the open-source Java library Strata by OpenGamma, available at https://github.com/OpenGamma/Strata.
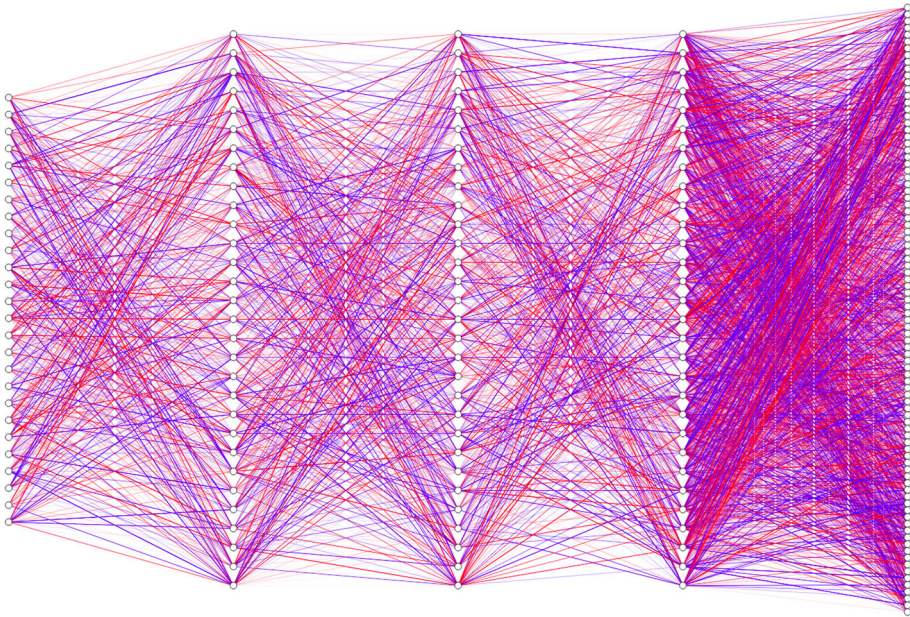
Fig. 2 Illustration of the chosen neural network architecture, where all weights have been generated randomly. The width of an edge is proportional to its weight. The color of an edge defines the sign of the weight (red if positive, blue if negative). The figure has been generated using the method of LeNail (2019)

Deterministic calibration: we rewrite (4.1) with the trained neural network $\mathcal{N}$ as follows:

$$\min_{p \in \mathcal{P}} \sum_{u=1}^{N} \sum_{i=1}^{\#T} \sum_{j=1}^{\#K} \left(\sigma_{imp}^{mkt}(u, T_i, K_j) - \mathcal{N}(p)_{(u,i,j)} \right)^2. \tag{4.3}$$

Following Horvath et al. (2021), we adopt the following neural network architecture:

- 3 hidden layers with 30 nodes on each;
- $N = 3$ surfaces, all sharing the same maturity range of size $\#T = 6$ and the same number of strikes $\#K = 5$. This yields an output layer of $3 \times 6 \times 5 = 90$ nodes. The size of the input layer is simply the number of model parameters;
- On the input and hidden layers, we employ the Exponential Linear Unit (ELU) activation function. The output layer is in turn equipped with the Sigmoid function.

Figure 2 provides a visualization of the neural network architecture.

For the training step, we start with the random generation of a training set of size $N_{train} = 10.000$. After normalization, we proceed with the training of the neural network by solving (4.2). The common practice is to use a stochastic optimization algorithm based on "mini-batch" gradient descent (see Goodfellow et al., 2016), whose updater can be specified following the *Adam* scheme (see Kingma and Ba, 2017). We set the mini-batch size to 32 and the number of epochs to 150 with potential early stopping[6].

---

[6] We rely on the open-source Java library Eclipse Deeplearning4j, available at http://deeplearning4j.org.

4.3 CBITCL specification

We consider a specification of the modeling framework described in Sect. 3.2 driven by two independent CBITCL processes (X^k, Z^k), $k = 1, 2$, where

(i) X^1 and X^2 are tempered α-stable CBI processes, as defined in Fontana et al. (2021);
(ii) the Lévy processes L^1 and L^2 generating the processes Z^1 and Z^2 (see Definition 2.2), respectively, are CGMY processes, as introduced in Carr et al. (2002).

We recall from Fontana et al. (2021) that a CBI process $X = (X_t)_{t \in [0,\mathcal{T}]}$ is said to be *tempered α-stable* if its immigration mechanism (2.1) reduces to $\Psi(x) = \beta x$ and the measure π in the branching mechanism (2.2) corresponds to the Lévy measure of a spectrally positive tempered α-stable compensated Lévy process. More specifically, we set

$$\pi(\mathrm{d}z) = C_\alpha \, \eta^\alpha \frac{e^{-\frac{\theta}{\eta}z}}{z^{1+\alpha}} \mathbf{1}_{\{z>0\}} \mathrm{d}z,$$

where $\eta > 0, \theta \geq 0, \alpha \in (-\infty, 2)$ (restricted to $\alpha \in (1, 2)$ if $\theta = 0$) and C_α is a normalization constant. This family of processes represents the tempered version of α-stable CBI processes, which have been successfully applied in finance in recent years (see, e.g., Jiao et al., 2017; 2019; 2021). The parameter α is referred to as the *stability index* and determines the jump behavior of the process X (see Fontana et al. 2021, Sect. 3.1):

- if $\alpha < 0$, then X has jumps of finite activity and finite variation;
- if $\alpha \in [0, 1)$, then X has jumps of infinite activity and finite variation;
- if $\alpha \in [1, 2)$, then X has jumps of infinite activity and infinite variation.

In the present setting, we shall consider the case $\alpha \in (1, 2)$ and specify the normalization constant as $C_\alpha = 1/\Gamma(-\alpha)$. The jumps of the process X are tempered exponentially depending on the value of the parameter θ, while the parameter η serves as a volatility coefficient controlling the jump volatility of the process X. The following lemma provides the explicit representation of the branching mechanism Φ of a tempered α-stable CBI process X (we refer to Szulda, 2021, for a proof). We denote by Γ the Gamma function extended to $\mathbb{R} \setminus \mathbb{Z}_-$ (see Lebedev, 1972).

Lemma 4.1 *For a tempered α-stable CBI process X with $\eta > 0$, $\theta \geq 0$, $C_\alpha = 1/\Gamma(-\alpha)$ and $\alpha \in (1, 2)$, the set $\mathcal{D}_X$ defined in (2.7) is given by $\mathcal{D}_X = (-\infty, \theta/\eta]$. Moreover, the branching mechanism Φ is given by*

$$\Phi(x) = -bx + \frac{1}{2}(\sigma x)^2 + (\theta - \eta x)^\alpha - \theta^\alpha + \alpha \theta^{\alpha-1} \eta x, \qquad \text{for all } x \leq \theta/\eta.$$

Let us also recall from Carr et al. (2002) that a Lévy process $L = (L_t)_{t \in [0,\mathcal{T}]}$ is of CGMY type if its Lévy measure γ is given by

$$\gamma(\mathrm{d}z) = C_L\big(z^{-1-Y} e^{-Mz} \mathbf{1}_{\{z>0\}} + |z|^{-1-Y} e^{-G|z|} \mathbf{1}_{\{z<0\}}\big)\mathrm{d}z,$$

where we fix $C_L = 1/\Gamma(-Y)$. The parameter $G > 0$ tempers the downward jumps of L, while $M > 0$ tempers the upward jumps, and $Y \in (1, 2)$ controls the local behavior of L in a similar way to the parameter α above. We recall that the Lévy exponent Ξ of a CGMY process is of the form

$$\Xi(u) := \beta u + \int_{\mathbb{R}} (e^{zu} - 1 - zu)\gamma(\mathrm{d}z), \qquad \text{for all } u \in i\mathbb{R}, \tag{4.4}$$

for $\beta \in \mathbb{R}$. It can be easily checked that, in the case of a CGMY process, the set $\mathcal{D}_Z$ defined in (2.8) is given by $\mathcal{D}_Z = [-G, M]$. The Lévy exponent (4.4) then takes the following explicit form:

$$\Xi(u) = \beta u + (M - u)^Y - M^Y + (G + u)^Y - G^Y$$
$$+ u\, Y\, (M^{Y-1} - G^{Y-1}), \quad \text{for all } u \in [-G, M].$$

As discussed in Sect. 3.3, the pricing of currency option requires the transformation of the model under $\mathbb{Q}^i$, for $i = 1, 2, 3$, where each $\mathbb{Q}^i$ represents the risk-neutral measure associated to the ith economy and is given by (3.4). In the present model specification, Theorem 3.5 directly implies the following result, which shows that not only the general CBITCL structure, but also the tempered α-stable property of X^k and the CGMY structure of L^k, for $k = 1, 2$, is preserved.

Corollary 4.2 *Under the model specification considered in this section, let $\mathbb{Q}^i$ be the probability measure defined in (3.4), for each $i = 1, \ldots, N$. Then, the processes (X^k, Z^k), $k = 1, 2$, remain independent CBITCL processes under $\mathbb{Q}^i$ and such that*

(i) X^k is a tempered α-stable CBI process with tempering parameter $\theta^{i,k} = \theta^k - \zeta_k^i \eta^k$;
(ii) $Z^k = L^k_{Y^k}$, where L^k is a CGMY process with tempering parameters $G^{i,k} = G^k + \lambda_k^i$ and $M^{i,k} = M^k - \lambda_k^i$.

Moreover, the drift term $b^{i,k}$ and $b_Z^{i,k}$ given in Table 2 can be explicitly computed as follows (see Szulda, 2021, Chapter 2 for further details), for $i = 1, \ldots, N$ and $k = 1, 2$:

$$b^{i,k} = b_k - \zeta_k^i \sigma_k^2 - \alpha_k \eta_k^{\alpha_k} \left(\theta_k^{\alpha_k - 1} - (\theta_k - \zeta_k^i \eta_k)^{\alpha_k - 1} \right),$$
$$\beta_Z^{i,k} = \beta_Z^k + Y^k \left((M^k)^{Y^k - 1} - (M^k - \lambda_k^i)^{Y^k - 1} + (G^k)^{Y^k - 1} - (G^k + \lambda_k^i)^{Y^k - 1} \right).$$

4.4 Calibration results

For the resolution of (4.1) and (4.3), we use the Levenberg–Marquardt optimizer of the open-source Java library Finmath[7]. We perform standard and deep calibrations. For the standard one, we obtain a root-mean-square error of 0.07557 in 709.977 seconds. Fig. 3 shows a satisfactory fit that slightly worsens for longer maturities. The deep calibration outperforms the standard one, achieving a root-mean-square error of 0.04092 in 0.269 seconds. The better quality of the fit can be seen from Fig. 4, where we can observe an improvement for longer maturities. Moreover, the execution time of the deep calibration is much smaller. However, one should take into account the time required for the training step, which may last up to several hours.

The calibrated values of the parameters are reported in Table 3. In particular, we can notice that the calibrated values of α_1 and α_2 are rather close to 1, indicating the potential presence of jump clustering phenomena (see Fontana et al. 2021, Section 3.1 for a detailed discussion of this aspect and for an analogous empirical evidence in multi-curve interest rate markets). We can also observe that the differences $\zeta_{\text{EUR}} - \zeta_{\text{USD}}$, $\zeta_{\text{EUR}} - \zeta_{\text{JPY}}$, $\zeta_{\text{USD}} - \zeta_{\text{JPY}}$ show evidence of moderate dependence between the FX rates considered and their volatility, in line with the findings of Ballotta and Morico (2018). By inspecting the calibrated values

---

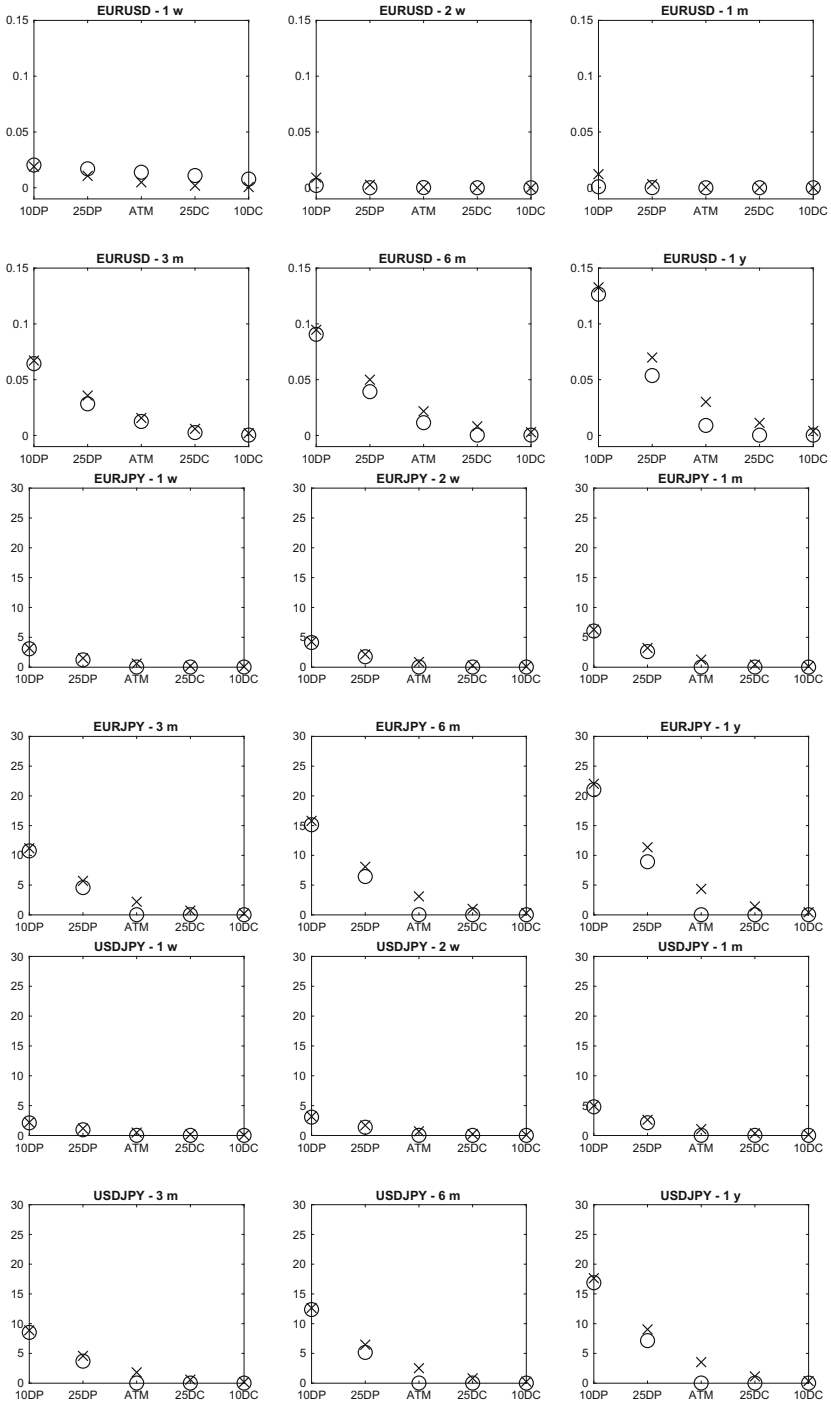[7] Available at https://www.finmath.net/finmath-lib.

Fig. 3 Calibration results obtained via the standard calibration. Market prices are denoted by crosses, model prices are denoted by circles. Moneyness levels follow the standard Delta quoting convention in the FX option market. DC and DP stand for "delta call" and "delta put", respectively
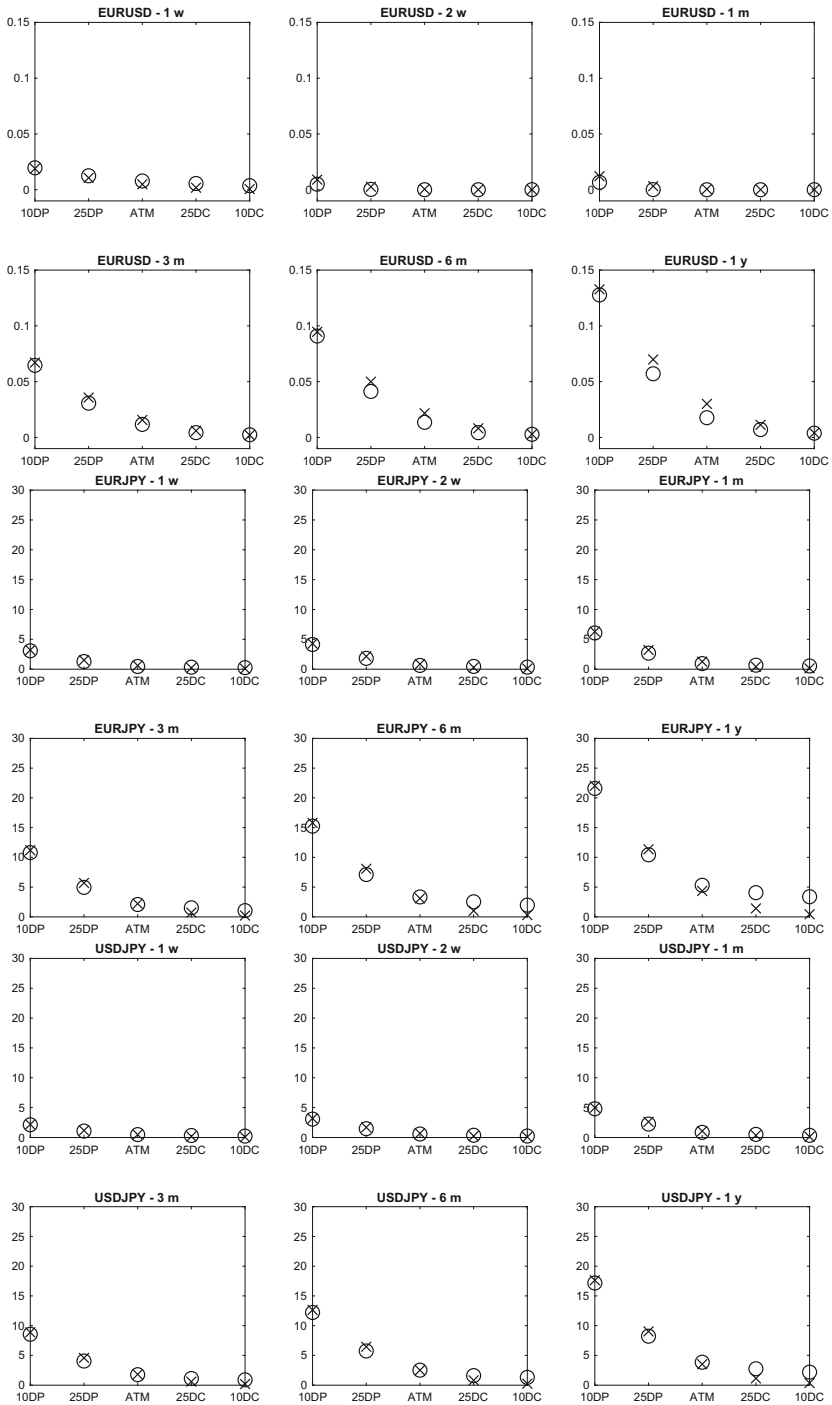
Fig. 4 Calibration results obtained via the deep calibration. Market prices are denoted by crosses, model prices are denoted by circles. Moneyness levels follow the standard Delta quoting convention in the FX option market. DC and DP stand for "delta call" and "delta put", respectively

Table 3 Calibrated values of the model parameters

	Standard	Deep		Standard	Deep
X_0^1	1.1040	1.1106	X_0^2	0.19652	0.18549
β^1	0.37721	0.65766	β^2	1.7524	1.7782
b^1	0.43082	0.43082	b^2	-0.73467	-0.73467
σ^1	2.1473	2.1473	σ^2	1.1174	1.1174
η^1	1.7208	1.7208	η^2	2.1855	2.1855
θ^1	1.9338	1.9338	θ^2	0.65273	0.65273
α^1	1.1697	1.1697	α^2	1.1122	1.1122
β_Z^1	-0.16220	-0.16220	β_Z^2	0.88065	0.88065
G^1	3.0313	3.0313	G^2	0.59711	0.59711
M^1	0.79529	0.79529	M^2	0.22821	0.22821
Y^1	1.7675	1.7675	Y^2	1.2390	1.2390
$\zeta_{JPY,1}$	1.12323	1.12366	$\zeta_{JPY,2}$	0.232636	0.232636
$\zeta_{USD,1}$	0.27244	0.27244	$\zeta_{USD,2}$	0.092184	0.060470
$\zeta_{EUR,1}$	0.089747	0.097352	$\zeta_{EUR,2}$	0.025973	0.024422
$\lambda_{JPY,1}$	0.39764	0.39764	$\lambda_{JPY,2}$	0.11410	0.11410
$\lambda_{USD,1}$	0.32863	0.32863	$\lambda_{USD,2}$	-0.014839	-0.014839
$\lambda_{EUR,1}$	0.16260	0.16260	$\lambda_{EUR,2}$	0.040496	0.040496

of the parameters of the tempered α-stable CBI processes, we can also notice a non-trivial contribution from the self-exciting jumps.

It is interesting to remark that the calibrated values of the parameters are stable across the two types of calibration. This is in accordance with the order of the two calibration exercises: first, we performed the deep calibration, where the initial guess was generated randomly by employing the same random generator that we used for the generation of the training set. We then used the optimal set of parameters obtained from this first calibration as the initial guess of the standard calibration. The fact that the output of the standard calibration is in line with that of the deep calibration provides us with a way of validating, in the present context, the deep calibration.

In order to assess the importance of allowing for jumps in the CBITCL specification, we compared the specification described in Sect. 4.3 with a continuous-path model where the Lévy processes L^1 and L^2 are simply given by two independent Brownian motions and the CBI processes X^1 and X^2 are standard square-root diffusions. This simplified model results in a Heston-type model. For the comparison, we calibrated both models to the same set of market implied volatilites, employing the standard calibration technique described in Sect. 4.2. Due to the much simpler structure of the model (in particular, of the associated Riccati equations), the calibration of the Heston-type model requires 311 seconds, while the calibration of the CBITCL model described in Sect. 4.3 required 703 seconds. However, the Heston-type model exhibits a worse fit to market data, achieving a RMSE of 0.1236. We observe that, in spite of the exclusion of the jump components, the calibrated values of the remaining parameters are quite similar across the two different specifications. These findings suggest that the jump components of our CBITCL specification capture some features of market data that cannot be adequately reproduced by continuous-path models.

5 Conclusions

We have proposed a stochastic volatility modeling framework for multiple currencies based on CBI-time-changed Lévy processes (CBITCL processes). The proposed approach combines full analytical tractability with consistency with the symmetric structure and the most relevant risk characteristics of FX markets. In particular, the self-exciting behavior of CBI processes allows capturing jump and volatility clustering effects. We have characterized a class of risk-neutral measures that leave invariant the structure of the model and allow for the derivation of a semi-closed pricing formula for currency options. Considering a specification driven by tempered α-stable CBI processes and CGMY Lévy processes, we have successfully calibrated the model to an FX triangle, using standard as well as deep learning techniques. The calibrated values of the parameters support the relevance of self-excitation and clustering phenomena.

Among the possible directions for further research, the modeling framework can be extended by considering stochastic interest rates in the different economies, possibly stochastically correlated with the FX rates. Moreover, we believe that CBITCL processes represent a flexible tool that can be successfully applied to other asset classes where stochastic volatility plays a relevant role.

References

Ballotta, L., Deelstra, G., & Rayée, G. (2017). Multivariate FX models with jumps: Triangles, quantos and implied correlation. *European Journal of Operational Research, 260*(3), 1181–1199.

Baldeaux, J., Grasselli, M., & Platen, E. (2015). Pricing currency derivatives under the benchmark approach. *Journal of Banking and Finance, 53*, 34–48.

Ballotta, L., & Morico, A. (2018). Hidden correlations: a self-exciting tale from the FX world. Working paper (available at https://ssrn.com/abstract=3245149).

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2002). The fine structure of asset returns: An empirical investigation. *The Journal of Business, 75*(2), 305–332.

Christoffersen, P., Heston, S., & Jacobs, K. (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science, 55*(12), 1914–1932.

Caballero, M. E., Pérez Garmendia, J. L., & Bravo, G. Uribe. (2013). A lamperti-type representation of continuous-state branching processes with immigration. *Annals of Probability, 41*(3), 1585–1627.

Cont, R., & Tankov, P. (2004). *Financial modelling with jump processes*. London: Chapman and Hall CRC.

Carr, P., & Wu, L. (2007). Stochastic skew in currency options. *Journal of Financial Economics, 86*(1), 213–247.

Rollin, S. Del Baño. (2008). Spot inversion in the Heston model. Working paper (available at https://core.ac.uk/display/13283041)

De Col, A., Gnoatto, A., & Grasselli, M. (2013). Smiles all around: FX joint calibration in a multi-Heston model. *Journal of Banking and Finance, 37*(10), 3799–3818.

Da Fonseca, J., & Grasselli, M. (2011). Riding on the smiles. *Quantitative Finance, 11*(11), 1609–1632.

Duffie, D., Filipović, D., & Schachermayer, W. (2003). Affine processes and applications in finance. *Annals of Applied Probability, 13*(3), 984–1053.

Dawson, D. A., & Li, Z. (2006). Skew convolution semigroups and affine Markov processes. *Annals of Probability, 34*(3), 1103–1142.

Doust, P. (2007). The intrinsic currency valuation framework. *Risk Magazine*, March:76–81 .

Doust, P. (2012). The stochastic intrinsic currency volatility model: A consistent framework for multiple FX rates and their volatilities. *Applied Mathematical Finance, 19*(5), 381–445.

Delbaen, F., & Schachermayer, W. (1998). The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen, 312*, 215–250.

Escobar, M., & Gschnaidtner, C. (2018). A multivariate stochastic volatility model with applications in the foreign exchange market. *Review of Derivatives Research, 21*(1), 1–43.

Eberlein, E., & Koval, N. (2006). A cross-currency Lévy market model. *Quantitative Finance, 6*(6), 465–480.

Eberlein, E., & Kallsen, J. (2020). *Mathematical Finance*. Springer finance: Springer.

Fontana, C., Gnoatto, A., & Szulda, G. (2021). Multiple yield curve modeling with CBI processes. *Mathematics and Financial Economics, 15*(2), 579–610.

Fontana, C., Gnoatto, A., Szulda,G. (2022). CBI-time-changed Lévy processes. Working paper (available at arXiv:2205.12355) .

Flesaker, B., & Hughston, L. P. (1997) International models for interest rates and foreign exchange. *Net Exposure, 3*: 55–79 Reprinted as Chapter 13 in: Hughston, L.P. (ed.) The New Interest Rate Models. Risk Publications, pp. 217–235.

Bank for International Settlements. (2019). *BIS Triennial Central Bank Survey: foreign exchange turnover in 2019*. BIS, Monetary and Economic Department: Technical report.

Fang, F., & Oosterlee, C. W. (2009). A novel pricing method for european options based on fourier-cosine series expansions. *SIAM Journal on Scientific Computing, 31*(2), 826–848.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Graceffa, F., Brigo, D., & Pallavicini, A. (2020). On the consistency of jump-diffusion dynamics for FX rates under inversion. *International Journal of Financial Engineering, 7*(4), 2050046.

Gnoatto, A., & Grasselli, M. (2014). An affine multi-currency model with stochastic volatility and stochastic interest rates. *SIAM Journal on Financial Mathematics, 5*(1), 493–531.

Gnoatto, A., Grasselli, M., & Platen, E. (2021). Calibration to FX triangles of the 4/2 model under the benchmark approach. *Decisions in Economics and Finance* forthcoming, 1-34

Gnoatto, A. (2017). Coherent foreign exchange market models. *International Journal Theoretical Applied Finance, 20*(1), 1750007.

Horvath, B., Muguruza, A., & Tomas, M. (2021). Deep learning volatility. *Quantitative Finance, 21*(1), 11–27.

Janek, A., Kluge, T., Weron, R., & Wystup, U. (2011). FX smile in the Heston model. In P. Cizek, W. K. Härdle, & R. Weron (Eds.), *Statistical Tools for Finance and Insurance* (pp. 133–162). Berlin-Heidelberg: Springer.

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.

Jiao, Y., Ma, C., Scotti, S., & Sgarra, C. (2019). A branching process approach to power markets. *Energy Economics, 79*, 144–156.

Jiao, Y., Ma, C., Scotti, S., & Zhou, C. (2021). The Alpha-Heston stochastic volatility model. *Mathematical Finance, 31*(3), 943–978.

Jacod, J., & Shiryaev, A. (2003). *Limit Theorems for Stochastic Processes*. Berlin-Heidelberg-New York, second edition Springer.

Kingma, D. P., & Ba, J. (2017). Adam: a method for stochastic optimization. Working paper (available at arxiv:1412.6980).

Keller-Ressel, M. (2009). Affine Processes - Theory and Applications in Finance. PhD thesis, Vienna University of Technology .

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

Kallsen, J., & Shiryaev, A. (2002). The cumulant process and esscher's change of measure. *Finance and Stochastics, 6*, 397–428.

Kawazu, K., & Watanabe, S. (1971). Branching processes with immigration and related limit theorems. *Theory Probability and its Applications, 16*(1), 36–54.

Lebedev, N. N. (1972). *Special Functions and their Applications*. Prentice-Hall, Englewood Cliffs (N.J.).

LeNail, A. (2019). NN-SVG: publication-ready neural network architecture schematics. *Journal of Open Source Software, 4*(33), 747.

Li, Z. (2020). Continuous-state branching processes with immigration. In Y. Jiao (Ed.), *From probability to finance - lecture notes of BICMR summer school on financial mathematics* (pp. 1–70). Singapore: Springer.

Szulda, G. (2021). Branching Processes and Multiple Term Structure Modeling. PhD thesis, Université de Paris.

Williams, D. (1991). *Probability with Martingales*. Cambridge: Cambridge University Press.

Wooldridge, P. (December 2019). FX and OTC derivatives markets through the lens of the Triennial Survey. *BIS Quarterly Review*