Finance Stochast. 7, 1–27 (2003)

Finance and Stochastics

Numerical solution of jump-diffusion LIBOR market models

Paul Glasserman[1], **Nicolas Merener**[2]

[1] 403 Uris Hall, Graduate School of Business, Columbia University, New York, NY 10027, USA (e-mail: pg20@columbia.edu)
[2] Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, USA (e-mail: nm187@columbia.edu)

Abstract. This paper develops, analyzes, and tests computational procedures for the numerical solution of LIBOR market models with jumps. We consider, in particular, a class of models in which jumps are driven by marked point processes with intensities that depend on the LIBOR rates themselves. While this formulation offers some attractive modeling features, it presents a challenge for computational work. As a first step, we therefore show how to reformulate a term structure model driven by marked point processes with suitably bounded state-dependent intensities into one driven by a Poisson random measure. This facilitates the development of discretization schemes because the Poisson random measure can be simulated without discretization error. Jumps in LIBOR rates are then thinned from the Poisson random measure using state-dependent thinning probabilities. Because of discontinuities inherent to the thinning process, this procedure falls outside the scope of existing convergence results; we provide some theoretical support for our method through a result establishing first and second order convergence of schemes that accommodates thinning but imposes stronger conditions on other problem data. The bias and computational efficiency of various schemes are compared through numerical experiments.

Key words: Interest rate models, Monte Carlo simulation, market models, marked point processes

JEL Classification: G13, E43

Mathematics Subject Classification (1991): 60G55, 60J75, 65C05, 90A09

The authors thank Professor Steve Kou for helpful comments and discussions. This research is supported by an IBM University Faculty Award and NSF grant DMS0074637.

Manuscript received: February 2001; final version received: April 2002

1 Introduction

The interest rate modeling approach advanced in Miltersen et al. [22], Brace, Gatarek, and Musiela [4], Musiela and Rutkowski [23], Jamshidian [13] and a large subsequent literature has gained widespread acceptance in the derivatives industry. This *market model* approach emphasizes the use of market observables as model primitives and ease of calibration to market data. This entails modeling the term structure of interest rate through, e.g., simply compounded forward LIBOR rates or forward swap rates rather than the continuously compounded instantaneous forward rates at the heart of the Heath et al. [12] (HJM) framework or through the instantaneous short rate of more traditional models. Ease of calibration to market prices of derivatives requires tractable formulas for liquid instruments like caps or swaptions.

The centerpiece of the market model framework is a class of models ([22,4, 13]) in which the prices of caps or swaptions coincide with the "Black [3] formulas" widely used in industry. Within this class of models, calibration to volatilities implied by the Black formulas is essentially automatic.

However, precisely because these models reproduce Black-formula prices exactly, they cannot generate a *skew* or *smile* in implied volatilities; all caplets of a given maturity must share a common implied volatility regardless of strike. This contradicts empirical evidence that implied volatilities in market prices do vary with strike. Andersen and Andreasen [1] and Zühlsdorff [29] have developed extensions of the basic LIBOR market model combining more general volatility specifications with computational tractability; these extensions produce non-constant implied volatilities. Glasserman and Kou [8] extend the market model to include jumps in interest rates governed by marked point processes and illustrate the variety of implied volatility patterns such a model can produce. This paper addresses computational issues in the extension of the LIBOR market model to include jumps and also some model formulation issues arising from computational considerations.

The potential importance of jumps in financial markets has been widely documented. Their impact is perhaps most pronounced in equity markets, but has been documented in foreign exchange and interest rate markets as well. Jumps play two related but somewhat distinct roles in modeling: one is providing a better fit to time series data and the other is providing greater flexibility in matching derivative prices – i.e., in modeling dynamics under an equivalent martingale measure. Equivalence of physical and martingale measures requires that both admit jumps if either does, but their frequency and magnitudes can be quite different under the two measures. Numerous references to the literature on modeling with jumps are discussed in [8] so here we mention just a few. In empirical work, Das [6] and Johannes [15] argue that the kurtosis in short-term interest rates is incompatible with a pure-diffusion model. Models adding jumps to the HJM framework (and thus focused on derivatives) include Björk et al. [2] (on which Glasserman and Kou [8] build) and Shirakawa [27]. Jamshidian [14] has developed a very general extension of the market model framework in which interest rate dynamics are driven by discontinuous semimartingales.

This paper addresses the numerical solution, through discretization and simulation, of the market models with jumps developed in [8]. Glasserman and Kou show how marked point process intensities can be chosen to produce closed-form expressions for caplets or swaptions, but for pricing general path-dependent interest rate derivatives simulation is necessary. The numerical solution of continuous processes modeled through stochastic differential equations has been studied in depth (see Kloeden and Platen [16] and its many references), but there has been far less work on methods for models with jumps. The models addressed here present special complications arising from the complex form of the intensities describing the dynamics of the marked point processes (MPPs). These intensities are, in general, functions of the current LIBOR term structure, and it is by no means obvious how to simulate an MPP with the required intensity.

The source of the complication can be explained through analogy with the pure-diffusion setting. The key to tractable caplet pricing lies in specifying convenient dynamics (in the simplest case, driftless geometric Brownian motion) for each LIBOR rate under its associated *forward* measure. But when all LIBOR rates are modeled under a single measure, the change of measure introduces a complicated drift term that depends on the current term structure. (In simulation, this drift can be handled using methods in [16] or circumvented using transformations in [11,10].) In the presence of jumps, Glasserman and Kou [8] obtain tractability by ensuring that the jumps in each LIBOR rate follow a compound Poisson process under the forward measure for that rate; this produces pricing formulas of the type in Merton [19] and Kou [17]. However, when all rates are modeled under a single measure, the change of measure introduces a change of intensity; the resulting intensities depend on the current term structure and thus describe processes that are far from Poisson.

In this paper, we show how the required MPPs may nevertheless be constructed through state-dependent thinning of a Poisson random measure. To this end, we first prove a general result showing how to reformulate a model driven by MPPs as one driven by a Poisson random measure, provided the MPP intensities satisfy a boundedness condition and Markovian assumption. We record a no-arbitrage result within the Poisson random measure formulation. We then use the reformulated model as a basis for numerical solution, adapting the general approach of Mikulevicius and Platen [20]. Briefly, the points of the Poisson random measure provide all *potential* jump times of the LIBOR rates; these may be generated without discretization error in advance of the evolution of the LIBOR rates. In between Poisson jumps, the evolution of the LIBOR rates is described by a pure diffusion and may thus be approximated to the desired accuracy using existing methods.

Results on the convergence order of numerical schemes for stochastic differential equations usually impose fairly strong smoothness conditions on the coefficient functions. The analysis of Mikulevicius and Platen [20] requires several orders of continuous differentiability (as do, e.g., the results of [18,26]). However, the thinning procedure we use is intrinsically discontinuous: a potential jump is either accepted or rejected depending on the current state of the LIBOR rates. To address this, we provide a convergence result that accommodates the thinning procedure but

imposes somewhat stronger conditions on other problem data. We also investigate the convergence of various methods numerically.

The rest of the paper is organized as follows. Section 2 reviews LIBOR market models and their extension with jumps described through marked point processes. Section 3 shows how to reformulate a general class of MPP-driven models in terms of a Poisson random measure and presents an associated no-arbitrage condition. Section 4 presents a subclass of models that lead to tractable formulas for caplets and discusses choices of state variables and parameter specifications. Section 5 details the numerical schemes and the jump generation mechanism. The convergence result is in Sect. 6. Section 7 contains numerical results and Sect. 8 concludes the paper.

2 Background on market models

2.1 Forward LIBOR models

Following Miltersen et al. [22], Brace et al. [4], Musiela and Rutkowski [23], and Jamshidian [13] we consider models of the term structure based on discretely compounded forward rates. We start with a discrete tenor structure–a finite set of dates $0 = T_0 < T_1 < \cdots < T_M < T_{M+1}$, with $T_{i+1} - T_i \equiv \delta$. The fixed accrual period δ is expressed as a fraction of a year; for instance, $\delta = 1/2$ represents six months. Each tenor date T_k is the maturity of a zero-coupon bond; $B_k(t)$ denotes the price of that bond at time $t \in [0, T_k]$ and $B_k(T_k) \equiv 1$. Forward LIBOR rates $L_1, \ldots, L_M$ may be defined from the bond prices by setting

$$L_k(t) = \frac{1}{\delta}\left(\frac{B_k(t)}{B_{k+1}(t)} - 1\right), \quad t \in [0, T_k], \quad k = 1, \ldots, M. \tag{1}$$

Each $L_k(t)$ is the forward rate for $[T_k, T_{k+1}]$ as of time $t \le T_k$; we denote by $L_0(0)$ the rate for $[0, T_1]$. Let $\eta(t) = \inf\{j \ge 0 : T_j \ge t\}$ so that $\eta(t)$ is the index of the next maturity as of time t.

We mostly work within the *spot martingale measure* of Jamshidian [13]. This is the measure associated with taking as numeraire

$$M(t) = \frac{B_{\eta(t)}}{B_1(0)} \prod_{j=1}^{\eta(t)-1} \frac{B_j(T_j)}{B_{j+1}(T_j)}.$$

Writing $D_k = B_k/M$ for the discounted (or *deflated*) bond prices, we have

$$D_k(t) = \prod_{j=0}^{k-1} \frac{1}{1 + \delta L_j(t \wedge T_j)}, \qquad L_k(t) = \frac{1}{\delta}\left(\frac{D_k(t)}{D_{k+1}(t)} - 1\right). \tag{2}$$

Absence of arbitrage opportunities by trading in bonds is guaranteed if the discounted bonds $D_k(t)$ are martingales (see Ch. 14 in Musiela and Rutkowski [24]). This requirement and (2) constrain the dynamics of the forward rates. In the pure diffusion case, the specification of the volatility determines the drift as in [4] and [13].

An interest rate caplet for the period $[T_n, T_{n+1}]$ with strike K is a derivative security paying $\delta(L_n(T_n) - K)^+$ at time T_{n+1}. Pricing under the spot martingale measure leads to the representation $C_n(0) = \delta \mathsf{E}[(L_n(T_n) - K)^+ / M(T_{n+1})]$ for the time-0 price of the caplet. This expression simplifies if written in terms of the *forward* or *terminal* measure for maturity T_{n+1}; this is the measure associated with taking the bond B_{n+1} as numeraire. (See, e.g., Ch. 13 of [24] for background.) Writing $\mathsf{E}_{T_{n+1}}$ for expectation under this measure, we have

$$C_n(0) = \delta B_{n+1}(0) \mathsf{E}_{T_{n+1}} \left[(L_n(T_n) - K)^+ \right]. \tag{3}$$

The nth caplet price is thus determined by the dynamics of L_n under the T_{n+1}-forward measure.

2.2 Jump-diffusion models

Björk et al. [2] generalize the no-arbitrage condition of Heath et al. [12] to incorporate jumps modeled through marked point processes in addition to the usual diffusion terms of the HJM framework. Glasserman and Kou [8] build on Björk et al. [2] to derive no-arbitrage conditions for LIBOR market models with jumps. We review the model specification in [8]. Our discussion of MPPs is informal; for mathematical background, see, e.g., [2] and [5].

We describe an MPP through a sequence of pairs of times and marks $\{(\tau_j, X_j), j = 1, 2, \dots\}$, with the interpretation that the mark X_j arrives at τ_j. The τ_j take values in $(0, \infty)$ and are strictly increasing. The marks X_j take values in a general space E^*, which for our purposes may be assumed to be a subset of a Euclidean space. Let N_t be the number of points in $(0, t]$: $N_t = \sup\{j \geq 0 : \tau_j \leq t\}$. From an MPP we construct jump processes by choosing a function $h : E^* \to \Re$ and defining

$$J(t) = \sum_{j=1}^{N_t} h(X_j).$$

The function h transforms the mark X_j into a jump magnitude, and so different jump processes can be generated from one MPP by different choices of h.

The term structure models in [8] are driven by r MPPs $\{(\tau_j^{(i)}, X_j^{(i)}), i = 1, \dots, r, j = 1, 2, \dots\}$ with no common jumps and a $d-$dimensional Brownian motion $W(t)$. The evolution of the rate maturing at T_k takes the form

$$\frac{dL_k^*(t)}{L_k^*(t-)} = \alpha_k(t, L^*(t-))\, dt + \gamma_k(t, L^*(t-))\, dW(t) + dJ_k^*(t), \qquad k = 1, \dots, M$$

for deterministic functions $\alpha_k : [0, \infty) \times \Re^M \to \Re$ and $\gamma_k : [0, \infty) \times \Re^M \to \Re^d$. We take $W(t)$ to be a column vector, $\gamma_k(t)$ a row vector, and we take the L_k to be right-continuous and denote by $L_k(t-)$ the left limit at $t-$. The variable L^* denotes the vector $(L_1^*, \dots, L_M^*)$. The jump term is

$$J_k^*(t) = \sum_{i=1}^{r} \sum_{j=1}^{N^{(i)}(t)} H_{ik}(X_j^{(i)})$$

with $N^{(i)}(t) = \sup\{j \geq 0 : \tau_j^{(i)} \leq t\}$ and deterministic functions $H_{ik} : E^* \to \Re$, $k = 1, \ldots, M$, $i = 1, \ldots, r$. Interpret H_{ik} as the response of the kth forward rate to the ith MPP. Each MPP $\{(\tau_j, X_j^{(i)})\}$ is assumed to admit an intensity process $\nu_i^*(dx, t)$ interpreted as the arrival rate of marks in dx for the i-th MPP, conditional on the history of the MPPs and the Brownian motion $W(t)$ up to $t-$. More precisely, ν_i^* makes the following a martingale in t for all bounded h:

$$\sum_{j=1}^{N^{(i)}(t)} h(X_j, \tau_j) - \int_0^t \int_{E^*} h(x, s)\nu_i^*(dx, s)\, ds.$$

Each MPP can also be described through a random measure $\mu_i^*(dx, dt)$ on the product of the time axis and the mark space assigning unit mass to each point $(\tau_j^{(i)}, X_j^{(i)})$. This representation makes it possible to write

$$\sum_{j=1}^{N^{(i)}(t)} H_{ik}(X_j^{(i)}) = \int_0^t \int_{E^*} H_{ik}(x)\mu_i^*(dx, ds)$$

from which the dynamics of the rates are

$$\frac{dL_k^*(t)}{L_k^*(t-)} = \alpha_k(t, L^*(t-))\, dt + \gamma_k(t, L^*(t-))\, dW(t) + \int_{E^*} \sum_{i=1}^{r} H_{ik}(x)\mu_i^*(dx, dt). \tag{4}$$

Theorem 3.1 in Glasserman and Kou [8] characterizes the class of arbitrage-free models of the form (5) through a restriction on the form of α_k.

Most relevant for practical applications is the case of models driven by MPPs with a "Markovian" property–namely, that $\nu_i^*(dx, t) = \nu_i(dx, L^*(t-), t)$ for some *deterministic* ν_i; i.e., the intensity depends on the history of the process only through the current state L^*. The tractable subclass of models identified in [8] have this form. In these models, each rate-specific jump process J_k^* becomes a compound Poisson process under the corresponding T_{k+1}-forward measure. But, when all LIBOR rates are modeled under a single measure the MPPs must be substantially more complicated than a compound Poisson process, though conveniently the Markovian feature is preserved.

3 Modeling jumps with Poisson random measures

Marked point processes provide a general formalism for constructing models with jumps, but an abstract MPP is difficult to work with computationally. In contrast, a Poisson random measure is easy to simulate, and the literature provides discretization schemes for stochastic differential equations driven by Brownian motion and Poisson random measures. These considerations motivate our next step in which we show how to use Poisson random measures to construct a class of MPPs.

A Poisson random measure $p(dx, dt)$ defines a marked point process with mark space E, that has a deterministic finite intensity $\lambda_P(dx, t)$ such that $\lambda_P(dx, t) = \lambda_0(t)\, f(x)\, dx$, with $f(x)$ a probability density on E. Thus, the arrival times follow a Poisson process with deterministic intensity $\lambda_0(t)$, and the marks are i.i.d. with density f. (Think of $p(dx, dt)$ as assigning unit mass to (x, t) if a mark x arrives at time t.) It will suffice for our purposes to take constant λ_0.

We proceed now to identify a class of MPPs that can be generated from a Poisson random measure through a state-dependent *thinning* mechanism. A thinning function θ_i randomly accepts or rejects the marks of the Poisson process with probability proportional to the value of the state-dependent intensity of the MPP at the moment of the jump. The resulting process of accepted marks has the law of the required MPP.

We begin with the following model for the rate dynamics:

$$\begin{aligned}\frac{dL_k(t)}{L_k(t-)} &= \alpha_k(t, L(t-))\, dt + \gamma_k(t, L(t-))\, dW(t) \\ &\quad + \int_{E^*} \int_0^1 \sum_{i=1}^{r} H_{ik}(y)\, \theta_i(y, u, L(t-), t)\, p(dy \times du, dt)\end{aligned} \tag{5}$$

where $p(dy \times du, dt)$ denotes a Poisson random measure with mark space $E = E^* \times (0, 1)$. This Poisson random measure has intensity $\lambda_P(y, u, t) = \lambda_0 f(y)$, $y \in E^*$, $u \in (0, 1)$. Thus, the marks $y \in E$ are distributed as $f(y)$, with total arrival rate λ_0, and u is uniformly distributed in (0,1). The functions α_k, γ_k, and H_{ik} are as in (5). We assume that the marked point processes μ_i^* in (5) have the following "Markovian" structure: each intensity $\nu_i^*(dx, t)$ can be written as $\nu_i(x, L^*(t-), t)\, dx$ for deterministic nonnegative functions ν_i. We further assume that the functions ν_i satisfy

$$\sum_{i=1}^{r} \nu_i(y, z_1, \ldots, z_M, t) < \lambda_0 f(y) \qquad \text{with} \quad z_k > 0, \quad k = 1, \ldots, M. \tag{6}$$

This allows us to subordinate the MPPs to a Poisson random measure.

From the ν_i we define deterministic thinning functions θ_i; when acting on $L(t)$ they are

$$\theta_i(y, u, L(t-), t) = \begin{cases} 1, \ \dfrac{\sum_{j=1}^{i-1} \nu_j(y, L(t-), t)}{f(y)\lambda_0} \le u \\ \qquad\quad < \dfrac{\sum_{j=1}^{i} \nu_j(y, L(t-), t)}{f(y)\lambda_0} \ \ i = 1, \ldots, r. \\ 0, \text{ otherwise,} \end{cases} \tag{7}$$

The interpretation of each thinning function θ_i is as follows. Associated with each jump time of the Poisson random measure is a mark (y, u). Because of nonnegativity of the functions ν_i and definition (7) there is at most one nonzero θ_i for each jump time. Given y and $L(t-)$, the probability of θ_i being nonzero at a jump time of the

Poisson process is, from (7), $\nu_i(y, L(t-), t)/\lambda_0 f(y)$. Intuitively, the marked point process defined by

$$\mu_i(dy, dt) = \int_0^1 \theta_i\,(y, u, L(t-), t)\, p(dy \times du, dt) \tag{8}$$

has a point in $[t, t + \Delta)$ with mark y with probability $\nu_i(y, L(t-), t)\Delta + o(\Delta)$, given $L(t-)$.

By introducing r thinning functions we are able to generate r MPPs from one Poisson random measure. Using (8), we may rewrite (6) as

$$\begin{aligned}\frac{dL_k(t)}{L_k(t-)} &= \alpha_k(t, L(t-))\, dt + \gamma_k(t, L(t-))\, dW(t) \\ &\quad + \int_{E^*} \sum_{i=1}^{r} H_{ik}(y)\mu_i(dy, dt)\end{aligned} \tag{9}$$

which is formally equivalent to (5). The next result verifies that (5) and (9) describe the same model. This reduces a model driven by abstract MPPs to one driven by a Poisson random measure.

Proposition 3.1 *If the functions ν_i associated to the intensities of the marked point processes μ_i^* satisfy (6) and if the processes $\nu_i(y, L_1(t-), \ldots, L_M(t-), t)$ are left-continuous for each $y \in E^*$ and $i = 1, \ldots, r$, then the marked point process μ_i has intensity $\nu_i(y, L_1(t-), \ldots, L_M(t-), t)$.*

Proof We need to show that for all measurable $A \subset E^*$ the counting process $N_t(A) = \int_0^t \int_A \mu_i(dy, ds)$ has intensity $\nu_i(A, L(t-), t) = \int_A \nu_i(y, L(t-), t)\, dy$ with respect to the history generated by L. Using (8) the counting process can be written as

$$N_t(A) = \int_0^t \int_{A\times[0,1]} \theta_i(y, u, L(s-), s)p(dy \times du, ds).$$

The processes $\nu_j(y, L(t-), t)$, $j = 1, \ldots, r$, are predictable because they are left-continuous; and as the $\theta_i(y, u, L(t-), t)$ are (measurable) deterministic function of the ν_j they too are predictable. Moreover, the θ_i are bounded, so by Theorem 8.T3 and Corollary 8.C4 of Brémaud [5],

$$N_t(A) - \int_0^t \int_{A\times[0,1]} \theta_i(y, u, L(s-), s)\lambda_0 f(y) dy\, du\, ds$$

is a martingale. Rewriting this using (7) we find that

$$N_t(A) - \int_0^t \int_A \nu_i(y, L(s-), s)\, dy\, ds$$

is a martingale and conclude (Theorem II.T9 of [5]) that $N_t(A)$ admits intensity $\nu_i(A, L(t-), t)$. □

Having transformed the underlying jump dynamics, we proceed now to document arbitrage restrictions on jump-diffusion models driven by a Poisson random measure. Our result characterizes the dynamics of forward LIBOR under the spot martingale measure which make the discounted bonds local martingales. This is the weak no-arbitrage condition of Musiela and Rutkowski [24]. Our building blocks are a $d-$dimensional Brownian motion $W(t)$ and a Poisson random measure $p(dx, dt)$ with finite intensity $\lambda_P(dx, t) = \lambda_0(t)f(x)dx$. The marks x take values in a space E.

Theorem 3.1 *For each $k = 1, \ldots, M$ let $\gamma_k : [0, \infty) \times \Re^M \to \Re^d$, $\alpha_k : [0, \infty) \times \Re^M \to \Re$ and $H_k : E \times \Re^M \times [0, \infty) \to \Re$ be deterministic functions. The model for the simple forward rates under the spot measure given by*

$$\frac{dL_k(t)}{L_k(t-)} = \alpha_k(t, L(t-))\, dt + \gamma_k(t, L(t-))\, dW(t) + dJ_k(t), \quad 0 \le t \le T_k, \quad k = 1, \ldots, M, \tag{10}$$

$$dJ_k(t) = \int_E H_k(x, L(t-), t)p(dx, dt) \tag{11}$$

makes the discounted bonds local martingales if

$$\alpha_k(t) = \gamma_k(t) \sum_{n=\eta(t)}^{k} \frac{\delta\gamma_n(t)^\top L_n(t-)}{1 + \delta L_n(t-)} - \int_E H_k(x, L(t-), t) \prod_{n=\eta(t)}^{k} \frac{1 + \delta L_n(t-)}{1 + \delta L_n(t-)(1 + H_n(x, L(t-), t))} \lambda_P(dx, t). \tag{12}$$

In this case, the dynamics of the discounted bonds are given by

$$\frac{dD_n(t)}{D_n(t-)} = \int_E \left[\prod_{j=\eta(t)}^{n-1} \frac{D_j(t-)}{D_j(t-) + (D_j(t-) - D_{j+1}(t-))H_j^*(x, D(t-), t)} - 1 \right] [p(dx, dt) - \lambda_P(dx, t)dt] + \sum_{k=\eta(t)}^{n-1} \left(\frac{D_{k+1}(t-)}{D_k(t-)} - 1 \right) \gamma_k(t)dW(t)$$

where $H_j^*(x, D(t-), t) = H_j(x, L(D(t-)), t)$ and $L_k(D(t-)) = (D_k(t-) - D_{k+1}(t-))/\delta D_{k+1}(t-)$.

Equation (12) is the drift restriction in [8] formulated here in terms of a Poisson random measure. Theorem 3.1 is proved by differentiating the transformation from the L_k to the D_j using an extension of Ito's formula accommodating integrals with respect to the Poisson random measure as well as the Brownian motion. It may also be possible to derive this result as a consequence of general results in Jamshidian [14], but the case of a Poisson random measure is sufficiently interesting and simple to merit separate consideration.

4 A tractable class of models

We focus now on a subclass of models identified in [8] that lead to explicit formulas for caplet pricing. From (3) we know that the nth caplet price is completely determined by the law of L_n under its associated forward measure $P_{T_{n+1}}$. Glasserman and Kou [8] obtain caplet formulas by positing that

$$\frac{dL_n(t)}{L_n(t-)} = -\bar{\lambda}_n(t)m_n\,dt + \gamma_n(t)\,dW_{n+1}(t) + d\left(\sum_{j=1}^{\bar{N}_n(t)}(Y_j^{(n)}-1)\right) \tag{13}$$

where $\gamma_n(t) \in \Re^d$ and $\bar{\lambda}_n(t) \in \Re$ are deterministic and bounded, and where, under $P_{T_{n+1}}$, $W_{n+1}(t)$ is a $d-$dimensional Brownian motion, $\bar{N}_n$ is a Poisson process with rate $\bar{\lambda}_n(t)$, the $Y_j^{(n)} \in (0,\infty)$ are independent with density f_n having mean $1+m_n$, and W_{n+1}, $\bar{N}_n$, and $\{Y_1^{(n)}, Y_2^{(n)}, \dots,\}$ are mutually independent.

While a specification of L_n under its forward measure $P_{T_{n+1}}$ suffices to determine caplet prices, working with a term structure model more generally requires specifying the dynamics of all forward rates simultaneously under a single probability measure, such as the spot measure. This requires the MPP formulation because if the jumps in each L_n are Poisson under the $P_{T_{n+1}}$ (as in (13)), they cannot all simultaneously be Poisson under a single measure. Once we have specified the term structure dynamics with marked point processes, we will however be able to use Proposition 3.1 to construct the MPPs by thinning a Poisson random measure. To get the dynamics of the L_n under the spot martingale measure, we follow the construction of Sect. 3.3 of [8].

4.1 Dynamics

We model the evolution of M rates using M marked point processes with marks in $(0,\infty)$ and intensities ν_i, $i = 1,\dots,M$, and a $d-$dimensional Brownian motion $W(t)$. For each n and t, we introduce a set $I_n(t)$ to be interpreted as the set of MPPs to which L_n is sensitive at time t. The function H_{ni} transforms the abstract marks of the i-th marked point process into jump magnitudes of the n-th rate; with a view towards (13), we choose these to be

$$H_{ni}(x) = \begin{cases} x-1, & \text{for } i \in I_n(t); \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

As shown in Proposition 3.1 of Glasserman and Kou [8], (13) holds if the intensities of the MPPs under the spot martingale measure $\nu_i(dy,t) = \nu_i(y,t)dy$ satisfy

$$\sum_{i\in I_n(t)} \nu_i(y,t) = \prod_{j=\eta(t)}^{n} \frac{1+\delta y L_j(t-)}{1+\delta L_j(t-)}\bar{\lambda}_n(t)f_n(y). \tag{15}$$

A special case, as shown in [8], consists in taking $\bar{\lambda}_n$ constant and f_n lognormal with volatility parameter s_n. The nth caplet is then priced by blending the formulas of Merton [19] and Black [3]:

$$C_n(t) = \delta \sum_{j=0}^{\infty} \exp\left(\bar{\lambda}_n(T_n - t)\right) \frac{(\bar{\lambda}_n(T_n - t))^j}{j!} \times BC\left(L_n^{(j)}(t), T_n - t, K, v_j^2(t), B_{n+1}(t)\right)$$

with $L_n^j(t) = L_n(t) \cdot e^{-\bar{\lambda}_n m_n (T_n - t)} \cdot (1 + m_n)^j$, $v_j^2(t) = (\int_t^{T_n} \|\gamma_k(u)\|^2 du + j s_n^2)/T_n$ and the Black formula [3], $BC(F, T, K, \sigma^2, b)$, with forward price F, maturity T, strike K, volatility parameter σ, and discount factor b. Taking f_n log-Laplace results in caplets priced by formulas of Kou [17].

We consider a slight variation of the special case above to obtain a stationary parameterization. Each forward rate evolves under its own forward measure as in (13), with $f_n(t) = f_{n+1-\eta(t)}$, so that the distribution of jump sizes depends on the number of tenor dates to maturity, and $\bar{\lambda}_n(t) = \bar{\lambda}_{n+1-\eta(t)}$. Coefficients remain constant between tenor dates. We take f_n to be lognormal with $\int_0^\infty y f_n(y) dy = 1 + m_n$. We complete the stationary specification by taking $I_n(t)$ to be dependent on time to maturity,

$$I_n(t) = (n + 1 - \eta(t), n + 2 - \eta(t), \ldots, M). \tag{16}$$

With this choice, the rate that will mature next, $L_{\eta(t)}$, is sensitive to all M marked point processes, and if some rate L_k jumps then all rates maturing earlier than T_k also jump. If we further require that $\gamma_n(t)$ depend on n and t only through $n + 1 - \eta(t)$, then a consequence of this stationary specification is that all rates follow, under their respective forward measures and for a fixed distance to their own maturities, the same stochastic differential equation.

This specification is meaningful only if the intensities $\nu_i(y, t)$ defined by (15) are all nonnegative and this imposes parameter restrictions. Applying the stationary parameterization to (15), we get

$$\nu_i(y,t) = \prod_{j=\eta(t)}^{i+\eta(t)-1} \frac{1 + \delta y L_j(t-)}{1 + \delta L_j(t-)} [\bar{\lambda}_i f_i(y) - \bar{\lambda}_{i+1} f_{i+1}(y) \left(\frac{1 + \delta y L_{i+\eta(t)}(t-)}{1 + \delta L_{i+\eta(t)}(t-)}\right)]. \tag{17}$$

In the lognormal case with f_n having the density of $\exp(N(a_n, s_n^2))$, for the $\nu_i(y, t)$ to be nonnegative it suffices to have

$$\log\left(\frac{s_{n+1}}{s_n}\right) - \frac{1}{2} z^2 \left(\frac{1}{s_n^2} - \frac{1}{s_{n+1}^2}\right) + z \left(\frac{a_n}{s_n^2} - \frac{a_{n+1}}{s_{n+1}^2}\right) - \frac{1}{2}\left(\frac{a_n^2}{s_n^2} - \frac{a_{n+1}^2}{s_{n+1}^2}\right) > \log\left(\frac{\bar{\lambda}_{n+1}}{\bar{\lambda}_n}\right) + \max(0, z). \tag{18}$$

We return now to the specification of the model under the spot martingale measure. The jump component is defined by (14), (16), and (17). Imposing these on Theorem 3.1 of Glasserman and Kou [8], the dynamics of the rates under the spot martingale measure become

$$\begin{aligned}\frac{dL_k(t)}{L_k(t-)} &= [-\bar{\lambda}_{k+1-\eta(t)} m_{k+1-\eta(t)} \\ &\quad + \sum_{j=\eta}^{k} \frac{\delta \gamma_k(t)\gamma_j(t)^\top L_j(t-)}{1+\delta L_j(t-)}]dt + \gamma_k(t)\, dW(t) \\ &\quad + d[\sum_{i=k+1-\eta}^{M} \sum_{j=1}^{N^{(i)}(t)} (Y_j^{(i)} - 1)] \end{aligned} \tag{19}$$

with $N^{(i)}$ the counting process for the ith MPP, which has intensity ν_i as in (17).

Using Proposition 3.1, we can construct this model from a Brownian motion and Poisson random measure. Lognormal densities f_n are of particular interest, but the following holds more generally:

Proposition 4.1 *Let $p(dy \times du, dt)$ be a Poisson random measure on $(0,\infty) \times (0,1) \times [0,\infty)$ with intensity $\lambda_0 f(y)$. Let f_1 have finite first moment $m_1 + 1$. The model defined by (14), (16), (17) and (19) can be written as*

$$\begin{aligned}\frac{dL_k(t)}{L_k(t-)} &= [-\bar{\lambda}_{k+1-\eta(t)} m_{k+1-\eta(t)} \\ &\quad + \sum_{j=\eta(t)}^{k} \frac{\delta \gamma_k(t)\gamma_j(t)^\top L_j(t-)}{1+\delta L_j(t-)}]dt + \gamma_k(t) dW(t) \\ &\quad + \int_0^\infty \int_0^1 (y-1) \sum_{i=k+1-\eta(t)}^{M} \theta_i(y,u,L(t-),t) p(dy \times du, dt) \end{aligned} \tag{20}$$

where

$$\theta_i\,(y,u,L(t-),t) = \begin{cases} 1, & \sum_{j=1}^{i-1} \nu_j(y,L_1,\ldots,L_M,t) \le u\, f(y)\lambda_0 \\ & < \sum_{j=1}^{i} \nu_j(y,L_1,\ldots,L_M,t); \\ 0, & \textit{otherwise} \end{cases} \tag{21}$$

with ν_i as in (17), $\lambda_0 = \bar{\lambda}_1(2+m_1)$ and $f(y) = \frac{f_1(y)+yf_1(y)}{2+m_1}$

Proof We need to check that (6) holds with $r = M$ and with λ_0 and $f(y)$ as given in the statement of the proposition. From (17), we get

$$\sum_{i=1}^{M} \nu_i(y,z_1,\ldots,z_M,t) = \frac{1+\delta y z_1}{1+\delta z_1} \bar{\lambda}_1 f_1(y).$$

with $f_1(y)$ having finite first moment, $yf_1(y)$ is proportional to another density $f^*(y)$, the normalization factor being $\int_0^\infty yf_1(y)dy = 1+m_1$. (If f_1 is lognormal, f^* is too.) Now,

$$\frac{1+\delta y z_1}{1+\delta z_1}\bar{\lambda}_1 f_1(y) = \bar{\lambda}_1\left(\frac{f_1(y)}{1+\delta z_1} + \frac{\delta z_1(1+m_1)f^*(y)}{1+\delta z_1}\right)$$
$$< \bar{\lambda}_1(f_1(y) + (1+m_1)f^*(y)) = \lambda_0 f(y),$$

the inequality following from the fact that $z_1 > 0$. □

This is the class of models we use to develop, test, and analyze numerical procedures. A convenient feature of these models is the availability of easily computed caplet prices for comparison. It should become clear that similar methods apply more generally to intensities satisfying (6).

4.2 Thinning

Through (21), Proposition 4.1 gives an explicit specification of the thinning procedure and identifies a Poisson intensity large enough to dominate the sum of the intensities of the MPPs in the original model. But the notation in (21) obscures the simplicity of the procedure so we supplement it with a more intuitive description.

For the construction in Proposition 4.1, we begin by generating points of a Poisson process with arrival rate λ_0; these are the potential jump times of the forward LIBOR rates. At each Poisson point, we generate a mark from the density f in the proposition. Writing

$$f(y) = \left(\frac{1}{2+m_1}\right) f_1(y) + \left(\frac{1+m_1}{2+m_1}\right)\frac{1}{1+m_1} y f_1(y)$$

makes it evident that $f(y)$ is a mixture of two densities: with probability $1/(2+m_1)$ we generate the mark from $f_1(y)$ and with probability $(1+m_1)/(2+m_1)$ we generate it from $yf_1(y)/(1+m_1)$, both of which are lognormal densities.

The next step is to decide which forward rates (if any) will jump. In (21) we determine which MPP jumps and then (20) translates this into jumps in forward rates. The implementation proceeds as follows. Using (17) and the realized mark y, we compute the partial sums $S_i = \sum_{j=1}^{i}\nu_j$, $i = 1, \ldots, M$. Next, we sample u uniformly from $(0,1)$. If $S_M < uf(y)\bar{\lambda}_0$, all rates remain unchanged. If $S_{i-1} < uf(y)\lambda_0 < S_i$ then the ith MPP is identified as having a point at the current time and, from (16), all rates maturing earlier than $T_{\eta+i}$ jump. The actual jump sizes are given by (20) which for a jumping rate reduces to

$$L_k(t) = L_k(t-) + L_k(t-)(y-1) \;=\; L_k(t-)\, y.$$

Forward rates maturing at $T_{\eta+i}$ and later remain unchanged. Because of (16), this construction is equivalent to having the rate closest to maturity L_η accept the Poisson jump at t with probability

$$\sum_{i\in I_\eta}\nu_i(y,t)/\lambda_0 f(y) = \frac{1+y\delta L_\eta}{1+\delta L_\eta}\frac{\bar{\lambda}_1 f_1(y)}{\lambda_0 f(y)};$$

and then, conditional on $L_{\eta+j}$ jumping, having $L_{\eta+j+1}$ jump with probability

$$\sum_{i\in I_{\eta+j+1}} \nu_i(y,t) \Bigg/ \sum_{i\in I_{\eta+j}} \nu_i(y,t) = \frac{1+y\delta L_{\eta+j+1}}{1+\delta L_{\eta+j+1}} \frac{\bar\lambda_{j+2} f_{j+2}(y)}{\bar\lambda_{j+1} f_{j+1}(y)}. \tag{22}$$

This sequential thinning gives each LIBOR rate the correct arrival rate of jumps and the presence of the densities f_i in the acceptance probabilities gives each LIBOR rate the correct distribution of jump magnitudes conditional on a jump.

Choosing a larger dominating intensity λ_0 would have no effect on the law of the process but would be computationally inefficient because it would result in a higher frequency of rejected jumps. With this in mind, we describe a slight modification of the procedure that results in fewer rejections.

The total intensity of jumps in L_η at time t is

$$\begin{aligned}\int_0^\infty \sum_{i\in I_\eta} \nu_i(y,t)dy &= \int_0^\infty \frac{1+y\delta L_\eta}{1+\delta L_\eta}\bar\lambda_1 f_1(y)dy \\ &= \bar\lambda_1\left(\frac{1}{1+\delta L_\eta} + \frac{\delta L_\eta(1+m_1)}{1+\delta L_\eta}\right) \\ &\le \bar\lambda_1(1+m_1^+), \quad m_1^+ = \max(0,m_1),\end{aligned} \tag{23}$$

which is typically smaller than λ_0; we therefore generate Poisson arrivals at the rate $\bar\lambda_1(1+m_1^+)$. At each point of this Poisson process the rate L_η jumps with probability

$$\frac{1+\delta L_\eta(1+m_1)}{(1+\delta L_\eta)(1+m_1^+)} \quad \text{with mark from} \quad \frac{(1+y\delta L_\eta)f_1(y)}{1+\delta L_\eta(1+m_1)},$$

which is again a mixture of two lognormal densities but now with state-dependent weights. As before, rates maturing later than T_η jump with conditional probabilites (22). The advantage of this method lies in its use of the current level of L_η in generating potential marks, which allows a smaller dominating Poisson rate and a greater frequency of acceptance.

In this construction, the fact that $yf(y)$ is (up to a normalization constant) a lognormal density whenever f is a lognormal density turns out to be convenient. In Kou's [17] model, f would be an asymmetric log-Laplace density and $yf(y)$ would again belong to the same family (after normalization). Indeed, if f is the density of Y and $\log Y$ belongs to an exponential family then (after normalization) $yf(y)$ belongs to the same exponential family.

In our numerical procedures we combine the thinning construction of the jumps with discretization methods for the diffusion terms. In addition to direct simulation of the forward rates we will investigate methods based on other choices of variables. The dynamics of $V_k = \log(L_k)$ are easily derived from those of L_k. The dynamics of the discounted bonds, when the rates evolve as in (20), are given by Theorem 3.1. Minor algebra involving (17) and (2) leads to the following simpler form for the discounted bond dynamics:

$$\frac{dD_n(t)}{D_n(t-)} = \sum_{k=\eta(t)}^{n-1} (\frac{D_{k+1}(t-)}{D_k(t-)} - 1)[-\bar{\lambda}_{k+1-\eta(t)} m_{k+1-\eta(t)} dt + \gamma_k(t) dW]$$

$$+ \int_0^\infty \int_0^1 \left[\prod_{j=\eta}^{n} \frac{D_j(t-)}{D_j(t-) + (D_j(t-) - D_{j+1}(t-)) \sum_{i=j+1-\eta(t)}^{M} (y-1)\theta_i(y,u,L(D),t)} - 1 \right]$$

$$p(dy \times du, dt). \qquad (24)$$

The functions θ_i are as in (21), where $L_k(D) = \frac{1}{\delta}\left(\frac{D_k}{D_{k+1}} - 1\right)$.

4.3 Parameter specification

We are interested in testing numerical procedures in plausible scenarios so we turn our attention to the identification of interesting sets of parameters $\{s_n,\, a_n,\, \bar{\lambda}_n\}$ satisfying (18). First, notice that (18) implies $s_{n+1} < s_n$ for the quadratic term coefficient to be positive. Next we recall that $\int_0^\infty x\, f_n(x)\, dx = \exp(a_n + \frac{1}{2}s_n^2)$, is the expected (multiplicative) jump size under a forward measure. To produce a downward sloping skew in implied volatility (typical of market data) we take a_n negative. We also impose the condition that the probability of a 20% jump (upwards or downwards) on any forward rate, conditional on a jump in the rate, is less than 0.2 . This is a somewhat arbitrary but reasonable cutoff in the physical measure and should be reasonable in the martingale measures. These considerations lead us to the following choice of parameters, which are consistent with (18):

$$A: \{s_{n+1} = 0.9 s_n,\, s_1 = 0.1,\, a_n = -0.1,\, \bar{\lambda}_{n+1} = 0.9\bar{\lambda}_n,\, \bar{\lambda}_1 = 0.5\} \quad (25)$$

$$B: \{s_{n+1} = 0.9 s_n,\, s_1 = 0.1,\, a_n = -0.1,\, \bar{\lambda}_{n+1} = 0.9\bar{\lambda}_n,\, \bar{\lambda}_1 = 5.0\}. \quad (26)$$

This is by no means intended as an exhaustive exploration of permissible parameters, but simply a particular specification for computational and testing purposes.

Among the reasons that motivate the inclusion of jumps in modeling interest rates is the ability to fit a volatility skew. To illustrate, we computed prices and (Black) implied volatilities of caplets maturing at $T = 2$ years, with all initial rates at 6%, $\gamma = 5\%$, and parameter set B. As the strike increases from 3% to 9%, the implied volatility decreases from 0.30 to 0.24.

5 Discretization schemes and implementation

5.1 General treatment

The models (20) and (24) are sets of coupled nonlinear stochastic differential equations. There is no way of generating exact sample paths for these models, nor do there generally exist formulas for expectations of functions of the paths as required for option pricing. Thus, we turn our efforts to the solution of discretized versions

of (20) and (24). We introduce first a general formulation to explain how we handle the combination of jumps and diffusion. Consider the M-dimensional process $X(t),\ t \in [0,T]$ that follows

$$dX(t) = \tilde{a}(X(t))\,dt + b(X(t))\,dW(t) + \int_E c(X(t),z)p(dz,dt) \tag{27}$$

where $p(dz,dt)$ is a Poisson random measure on $E \times [0,T]$ with intensity $\lambda_0\, h(z)$ and vector marks z distributed as $h(z)$. For simplicity, we take W to be a scalar Brownian motion, though the schemes can be easily generalized to the multifactor case. The deterministic functions $\tilde{a}, b,$ and c are M-dimensional vectors with components $\tilde{a}_j, b_j,$ and c_j. An explicit time-dependence in the coefficients of (27) could be accommodated by including time as a component of the vector $X(t)$.

We construct approximate solutions to models of the form (27) at a discrete set of times $\{\tau_i\}$. This set is the superposition of the random jump times of a Poisson process on $[0,T_M]$ and a deterministic grid that includes all maturity dates $T_1,\ldots,T_M$. The random Poisson jump times can be computed without any knowledge of the realized path of (27); this is the main advantage of formulating the MPP construction through thinning of a Poisson random measure.

Mikulevicius and Platen [20] (see also [18,25,26]) introduced explicit schemes that generate approximate solutions $Y(\tau_i)$ of (27) on the grid points τ_i. We briefly review the schemes before applying them to financial models. A scheme $\{Y(\tau_i)\}$ is said to have weak order of convergence ξ if for all sufficiently small ϵ

$$|E(g(X(T_M))) - E(g(Y(T_M)))| \leq \text{ constant} \cdot \epsilon^{\xi}$$

with ϵ the maximum step size in the deterministic grid and g ranging over a class of functions, such as those with $2(\xi+1)$ polynomially bounded derivatives (see p.327 of Kloeden and Platen [16]). Among the simplest schemes is a stochastic Taylor approximation of order one, also called an Euler scheme. The vector $Y(\tau_i)$ is iteratively computed from the initial condition $Y(0)$ using

$$Y(\tau_{i+1}^-) = Y(\tau_i) + f_0(Y(\tau_i))(\tau_{i+1}-\tau_i) + f_1(Y(\tau_i))(W_{\tau_{i+1}} - W_{\tau_i}), \tag{28}$$

$$Y(\tau_{i+1}) = Y(\tau_{i+1}^-) + \int_E c(Y(\tau_{i+1}^-),z)p(dz,\tau_{i+1}) \tag{29}$$

$$f_0(Y(\tau_i)) = \tilde{a}(Y(\tau_i)) \text{ and } f_1(Y(\tau_i)) = b(Y(\tau_i)). \tag{30}$$

At each grid point, (29) computes the magnitude of a jump exactly, conditional on $Y(\tau_{i+1}-)$, if τ_{i+1} is indeed a point of the Poisson random measure (rather than one of the deterministic grid points). Otherwise, the jump term is zero.

The Euler (or higher order) scheme can also be applied to $\log(X(t))$, i.e., with $Y(\tau_i)$ defined as the exponential of a discrete solution of $\log(X(t))$. This is potentially helpful if the coefficients $\tilde{a}, b,$ and c are approximately proportional to $X(t)$. If $\tilde{a}, b,$ and c are linear functions of $X(t)$, this choice of discretization variable gives the exact solution of the model (27) while (28)-(29) does not.

Next we present the generalization of the Milstein [21] scheme proposed by Mikulevicius and Platen [20], a stochastic Taylor approximation of order two. As

in the first-order scheme, jump magnitudes are computed exactly conditional on the state of the system at τ_{i+1}^- and the diffusion is approximated, though more accurately now. The scheme for the continuous part of the path is,

$$\begin{aligned} Y(\tau_{i+1}^-) &= Y(\tau_i) + f_0(Y(\tau_i))(\tau_{i+1} - \tau_i) \\ &+ f_1(Y(\tau_i))Z_i + f_{00}(Y(\tau_i))\frac{1}{2}(\tau_{i+1} - \tau_i)^2 + f_{10}(Y(\tau_i))U_i \\ &+ f_{01}(Y(\tau_i))(Z_i(\tau_{i+1} - \tau_i) - U_i) + f_{11}(Y(\tau_i))\frac{1}{2}(Z_i^2 - (\tau_{i+1} - \tau_i)) \end{aligned} \quad (31)$$

where $U_i = \int_{\tau_i}^{\tau_{i+1}} \int_{\tau_i}^{s_2} dW_{s_1} ds_2 \sim N(0, \frac{1}{3}(\tau_{i+1} - \tau_i))$ and $Z_i = \int_{\tau_i}^{\tau_{i+1}} dW_s \sim N(0, (\tau_{i+1} - \tau_i))$ with $EU_iZ_i = (\tau_{i+1} - \tau_i)^2$ are sampled without error from normal distributions. The updating of the rates at a jump time is as in (29). The M-dimensional functions $\{f_0, f_1, f_{00}, f_{10}, f_{01}, f_{11}\}$ arise in the truncation of the stochastic (Ito calculus) Taylor expansion. The first order coefficients $\{f_0, f_1\}$ are as in (30). Writing ∂_j for a partial derivative with respect to X_j, the others are

$$\begin{aligned} f_{00}(Y) &= \sum_{j=1}^{M} \tilde{a}_j(Y)\partial_j \tilde{a}(Y) + \frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M} b_j(Y)b_k(Y)\partial_{jk}\tilde{a}(Y), \\ f_{11}(Y) &= \sum_{j=1}^{M} b_j(Y)\partial_j b(Y), \\ f_{10}(Y) &= \sum_{j=1}^{M} b_j(Y)\partial_j \tilde{a}(Y), \\ f_{01}(Y) &= \sum_{j=1}^{M} \tilde{a}_j(Y)\partial_j b(Y) + \frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M} b_j(Y)b_k(Y)\partial_{jk} b(Y). \end{aligned} \quad (32)$$

5.2 Discretization of tractable models

We apply these schemes to the forward rates (20), the logarithms of the forward rates, and the discounted bonds (24). For brevity, in this section we detail only the schemes based on the logarithm of the forward rates, which, as it will be clear in Sect. 7, is the choice of variable that minimizes discretization bias. The derivation of schemes based on rates and bonds is straightforward, based on the general treatment in Sect. 5.1 and the corresponding continuous time dynamics for each choice of variable-rates, log rates, or discounted bonds.

We discuss the jump step explicitly only for the Euler scheme because its calculation does not entail discretization. Furthermore, the choice of variables affects only the approximation of the continuous part of the model; jumps are processed exactly regardless of the choice of variables.

To lighten notation we write η for $\eta(t)$ and to simplify the setting we assume that the γ_k are constant. We discretize the logarithms of the forward rates, applying the

Euler scheme and generalized Milstein scheme to the log rates and then recovering the rates by exponentiating. Set

$$\hat{a}_k(\tau_i) = -\bar{\lambda}_{k+1-\eta} m_{k+1-\eta} - \frac{1}{2}\gamma_k^2 + C_k(\tau_i).$$

The first order scheme for the logarithms of rates leads to

$$\begin{aligned}
&\hat{L}_k(\tau_{i+1}^-) = \hat{L}_k(\tau_i)\exp\{\hat{a}_k(\tau_i)(\tau_{i+1}-\tau_i) + \gamma_k(W_{\tau_{i+1}} - W_{\tau_i})\},\\
&\hat{L}_k(\tau_{i+1}) = \hat{L}_k(\tau_{i+1}^-) \qquad (33)\\
&\left[1 + \int_0^\infty \int_0^1 (y-1) \sum_{i=k+1-\eta}^{M} \theta_j(y,u,\hat{L}(\tau_{i+1}^-),\tau_{i+1}^-))\, p(dy\times du, \tau_{i+1})\right].
\end{aligned}$$

The generation of y and calculation of the thinning functions θ_j is exactly as described in Sect. 4.2. In particular, y is drawn from a mixture of two lognormal densities and we then identify, using (21) and a uniformly distributed u, which thinning function θ_i equals 1 (if any). Then the integral in (33) evaluates to $y-1$ if $i \in \{k+1-\eta,\ldots,M\}$ and 0 otherwise. The actual jump increment in each rate is the value of this integral times the rate just before the jump. In the case of bonds, the jump magnitude is not just proportional to $y-1$ but a more complex function because many forward rates contribute to the price of a single bond.

Applying the second order scheme to the logarithms of the rates we get, for the continuous part,

$$\begin{aligned}
\hat{L}_k(\tau_{i+1}^-) = \hat{L}_k(\tau_i)\exp\{&a_k(\tau_i)(\tau_{i+1}-\tau_i) + \gamma_k(W_{\tau_{i+1}} - W_{\tau_i})\\
&+ \sum_{j=\eta}^{k} \frac{\delta\gamma_j^2\gamma_k\hat{L}_j(\tau_i)}{(1+\delta\hat{L}_j(\tau_i))^2} U_i\\
&+ \sum_{j=\eta}^{k} \frac{\delta\gamma_j\gamma_k\hat{L}_j(\tau_i)}{(1+\delta\hat{L}_j(\tau_i))^2}[a_j(\tau_i) + \frac{1}{2}\gamma_j^2\frac{1-\delta\hat{L}_j(\tau_i)}{1+\delta\hat{L}_j(\tau_i)}]\frac{1}{2}(\tau_{i+1}-\tau_i)^2\},
\end{aligned}$$

and the updating at a jump time is as in (33). The implementation is simplified by the fact that the coefficients for $\hat{L}_k$ involve sums of the form $\sum_\eta^k$. Therefore, each of the sums needed to update $\hat{L}_k$ has just one more term than the corresponding sum needed to update $\hat{L}_{k-1}$. So, at each step, the rates are updated in increasing order of maturity. The updating of the diffusion and drift can be implemented in less than ten lines of code in all schemes. Similar discretization methods should be applicable to the swap rate models in [8]. However, we do not consider that case explicitly here.

6 Convergence

We turn now to the issue of convergence of the discretization schemes, working within the general framework of a process $X(t)$, $t \in [0,T]$ as in (27). Recall that

$p(dz, dt)$ in (27) is a Poisson random measure on $E \times [0, T]$ with intensity $\lambda_0\, h(z)$ and vector marks z distributed as $h(z)$. Define

$$a(y) = \tilde{a}(y) + \int_E c(y, z)\, h(z)\, \lambda_0\, dz$$

so the dynamics can be written as

$$dX(t) = a(X(t))\, dt + b(X(t))\, dW + \int_E c(X(t), z)\, q(dz, dt)$$

where $q(dz, dt) = p(dz, dt) - h(z)\, \lambda_0\, dz$ is a Poisson martingale measure on $E \times [0, T]$. In the applications we are considering, $z \in [0, \infty) \times (0, 1)$. We consider the problem of calculating $\mathsf{E}[g(X(T))]$ for some real-valued g which we call the payoff function.

As mentioned in Sect. 5, Mikulevicius and Platen [20] introduced a hierarchy of schemes which, under regularity conditions on a, b, c and the payoff function g, are shown to have arbitrarily high order of weak convergence. In particular, the Euler scheme converges weakly with order one and the Milstein scheme with order two. But the continuous-time models we are considering violate their hypotheses in an important way: the thinning procedure at the heart of our construction makes the function c discontinuous, whereas the analysis in Mikulevicius and Platen [20] requires that this function be several times continuously differentiable. We therefore present an alternative convergence result that allows for discontinuous c, though it imposes stronger requirements on g.

Let $B^\xi(C)$ be the class of $2(\xi+1)$-times continuously differentiable real-valued functions for which the function itself and its partial derivatives up to order $2(\xi+1)$ are uniformly bounded by a constant C. We will need to assume that the payoff g is in some $B^\xi(G)$. Boundedness generally applies to the payoff of a put or floor, though not to a call or cap. The smoothness required is restrictive, but cannot be easily avoided; indeed, the results of Mikulevicius and Platen [20] and nearly all of the literature on the numerical solution of stochastic differential equations requires stronger smoothness conditions than one would like for option pricing. These types of results still provide useful information concerning the level of accuracy one can expect with alternative methods, and these theoretical guides can and should be supplemented with numerical experiments.

For bounded $\psi : \Re^M \to \Re$ let

$$\phi(x) = \int_E \psi(x + c(x, z))\, h(z)\, dz, \tag{34}$$

and let $\bar{\phi}(x) = \int_E (x + c(x, z))\, h(z)\, dz$.

Theorem 6.1 *Fix $\xi \in \{1, 2\}$. Let the payoff function $g : \Re^M \to \Re$ be in $B^\xi(G)$ for some G and let $\{X(t), t \in [0, T]\}$ be as in (27). We assume:*

(i) $\bar{\phi}(x)$ is $2(\xi + 1)$-times continuously differentiable with uniformly bounded derivatives;

(ii) there is a constant K such that if $\psi \in B^\xi(\Psi)$ for some Ψ then $\phi(x) \in B^\xi(K\Psi)$ in (34);

(iii) *a and b are $2(\xi+1)$-times continuously differentiable with uniformly bounded derivatives;*

(iv) *there is a constant K_2 such that any function $f \in S_\xi$ satisfies $|f(y)| \le K_2(1+\|y\|)$, with $S_1 = \{f_0, f_1\}$ as in (30) and $S_2 = \{f_0, f_1, f_{00}, f_{10}, f_{01}, f_{11}\}$ as in (32).*

Then the approximation (29) has weak convergence order one and the approximation (31) has weak convergence order two.

Hypotheses (i) and (ii) replace the assumption in Theorem 3.3 of Mikulevicius and Platen [20] that c is $2(\xi+1)-$times continuously differentiable with bounded derivatives. The result follows from the proof of Theorem 3.3 of Mikulevicius and Platen [20] once we establish that two key properties used in their proof hold in our setting as well: the existence of a stochastic Taylor formula and smoothness of the solution of a backward Kolmogorov equation. Details of the proof of Theorem 6.1 are in [9]. The proof holds, in fact, for the entire hierarchy of schemes proposed in Mikulevicius and Platen [20], which have arbitrarily high orders of convergence. Schemes of order higher than two are constructed using the functions f in S_ξ which are defined in a recursive way in [20]. We have written explicitly the first and second order schemes only. Higher order schemes can be somewhat cumbersome to write out explicitly and to implement.

The conditions of Theorem 6.1 are satisfied for both first and second order schemes applied to the dynamics of the log rates. (The verification entails straightforward but lengthy calculations of derivatives and bounds and is therefore omitted.) More precisely, the conditions are satisfied within each accrual period $[T_i, T_{i+1}]$. Discontinuities in the coefficients at the dates T_i are inherent to the model because of the presence of sums and products whose range begins at $\eta(t)$ and the fact that η increases by one at each tenor date. While it may be possible to choose parameters to interpolate smoothly at the T_i, in practice these types of models are typically calibrated to market data with piecewise constant coefficients. It therefore seems preferable to think of the model as governed by a separate stochastic differential equation over each accrual period, with the terminal value over $[T_{i-1}, T_i]$ determining the initial condition over $[T_i, T_{i+1}]$. This issue is by no means the result of introducing jumps—the same issue arises in pure diffusion market models. In the next section we supplement the theoretical properties of discretization schemes with numerical experiments and these indicate that working with log rates has practical as well as theoretical advantages.

7 Numerical results

We present numerical results generated with the discretization schemes introduced in Sect. 5. These experiments have two purposes. First, we quantify the magnitude of the biases in computed prices arising from time-discretization. Second, we assess the computing demands of alternative schemes. The two issues are related and are heavily dependent on the complexity of the underlying continuous-time model. The schemes we study and their abbreviations are as follows:

Variable	rate	rate	log rate	log rate	bond	log bond
Order	1	2	1	2	1	1
Name	r1	r2	lr1	lr2	b1	lb1

In addition, it is often possible to achieve a higher convergence order from a first order scheme by using Richardson extrapolation as in Sect. 15.3 of Kloeden and Platen [16], Talay and Tubaro [28], and Protter and Talay [26]. Suppose $g(\hat{L}_T^h)$ and $g(\hat{L}_T^{h/N})$ are calculated using time steps h and h/N for some integer N. If the leading term in their biases are proportional to the time step it can be eliminated by combining them in the extrapolated estimate

$$\hat{E}xpt = \frac{N}{N-1} g(\hat{L}_T^{h/N}) - \frac{1}{N-1} g(\hat{L}_T^h).$$

This can achieve weak convergence order two, as shown in Talay and Tubaro [28] and Protter and Talay [26].

7.1 Biases

In our first experiment we calculate discounted bond prices by simulation under the spot martingale measure. In the absence of arbitrage, discounted bond prices must be martingales, therefore

$$\mathsf{E}\left[D_{k+1}(T_{k+1})\right] = \mathsf{E}\left[\prod_{j=0}^{k} \frac{1}{1+\delta L_j(T_j)}\right] = \prod_{j=0}^{k} \frac{1}{1+\delta L_j(0)}, \quad k = 1, \ldots, M. \tag{35}$$

The rightmost expression can be calculated from the initial term structure and either of the expectations on the left can be estimated by simulation. As a check on the accuracy of various schemes, we estimate the relative bias in bond prices – i.e., the bias in estimating (35) divided by the exact expression on the right. Schemes based on discretizing the bond dynamics use the leftmost expectation in (35); schemes based on the rates use the expression in the middle.

In the experiments we take all $L_j(0) = 0.06$, accrual period $\delta = 0.5$ and diffusion volatilities $\gamma_j = 5\%$. We choose a bond expiring at time $T_{11} = 5.5$ years, which requires simulation of rates $L_1(t), \ldots, L_{10}(t)$. The jump parameters are as in (26). We vary the time step in the deterministic grid, but the effective time step is random because the grid includes all jump times. Thus, the nominal time step and the Poisson jump intensity determine the actual time steps.

Figure 1 shows estimated relative biases for various schemes and nominal time steps. The presence of biases means that the discretized model is not strictly arbitrage-free. As demonstrated in [11], this can be avoided through a change of discretization variables in the pure diffusion setting; but in the presence of jumps the dynamics include intensity terms that cannot be discretized exactly, so some bias in bond prices seems unavoidable. Nevertheless, Fig. 1 indicates that this bias can be made very small with even a rather coarse time grid by using a second-order scheme for rates or a first-order scheme for log rates.

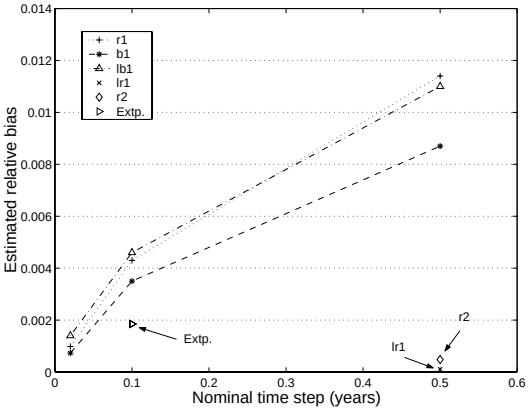

Fig. 1. Bond price biases, $T = 5.5$

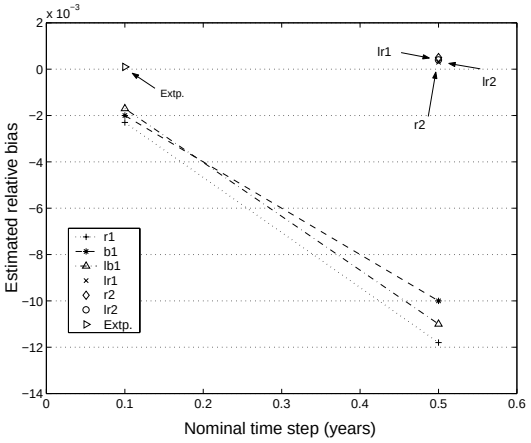

Fig. 2. Caplet price biases, $T = 2$ years, jumps parameters set A

We present next estimated relative biases in caplet prices computed under the spot martingale measure. In order to study biases of the schemes, unbiased caplet prices are simultaneously obtained by Monte Carlo simulation under the forward measure associated with each caplet. Under the T_k forward measure, within each accrual period the drift and diffusion coefficients of L_k are linear functions of L_k (see (13)). Thus, even a first-order scheme for log rates solves this equation exactly and gives an unbiased estimator of the price. In practice, simulation would not be needed for computing caplet prices, but we use the available unbiased prices to test the quality of the methods.

Figures 2 and 3 show estimated relative biases for caplets maturing in 2 years and parameter sets A and B in (25)-(26). These jump specifications differ by a factor of ten in the arrival rate of the jumps, and the biases in Fig. 3 are roughly ten times those in Fig. 2. Using either a second-order scheme for the rates or any

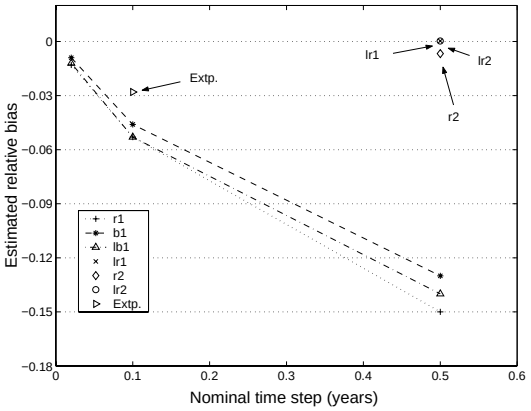

Fig. 3. Caplet price biases, $T = 2$ years, jumps parameters set B

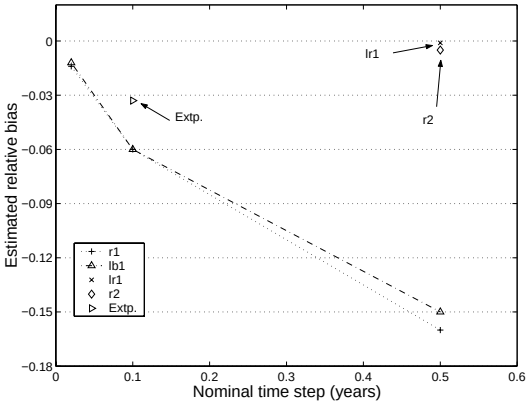

Fig. 4. Caplet price biases, $T = 10$ years, jumps parameters set A

scheme for the log rates removes most of the bias at even the coarsest time step; there do not appear to be appreciable differences among the other methods.

Figures 4 and 5 show biases for caplets maturing in 10 years. By comparing these with the previous figures we find that the relative bias does not significantly increase with maturity. Furthermore, closer comparison of Figs. 3 and 5 reveals a small decrease in the relative bias for the r2 scheme. This is somewhat surprising but, interestingly, similar behavior appears in the experiments of Andersen and Andreasen [1]. As in the 2 year caplets, the biases are an order of magnitude smaller for jump parameters A than for jump parameters B.

In all cases, the smallest biases are achieved by the schemes discretizing the log rates. This is due to the fact that, even under the spot martingale measure, the nonlinear term in the drift is one or two orders of magnitude smaller than the linear term in the drift and diffusion term. This fact is optimally exploited by the lr1 and lr2 schemes, which give the exact solution in the linear coefficients case. In contrast,

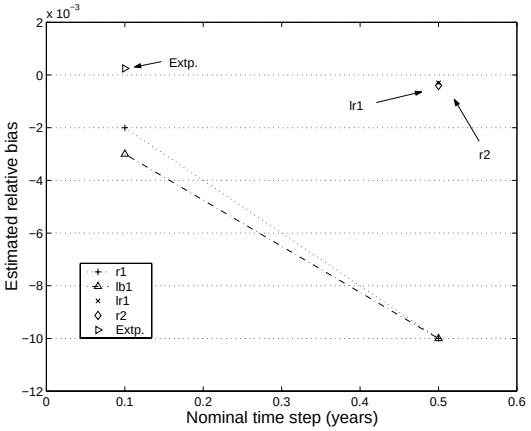

Fig. 5. Caplet price biases, $T = 10$ years, jumps parameters set B

schemes using bonds as computing variables are highly nonlinear, as can be seen from the continuous-time dynamics (24).

7.2 Root mean square error and efficiency

While bias is an important measure of the quality of a scheme, its effect may become evident only in very long computations. For many practical applications, where the computing budget is limited, the dominant effect is the estimation error due to sampling variability. In order to address this issue, we compare the root mean square error of caplet prices computed under various schemes with a fixed computing time. The root mean square error RMS is $(\text{bias}^2 + \text{SE}^2)^{\frac{1}{2}}$, with SE the standard error, estimated as the sample standard deviation divided by the square root of the number of paths. We find empirically that, for a given pricing problem (number of maturities, jump and diffusion parameters), the standard deviation is nearly independent of the scheme and time step. However, with a fixed computing budget the SE still varies across schemes and time steps because the number of paths that can be completed depends on the time required per path. Thus, faster schemes and larger time steps are potentially attractive if biases are much smaller than standard errors. For a theoretical analysis of this tradeoff, see Duffie and Glynn [7].

Table 1 shows estimated relative RMS errors for the pricing of 2-year caplets in 1 second of computing time on a 350MHz Pentium II PC. With parameter set A, several schemes achieve roughly the same RMS, 2%, in one second of computing time. Schemes lr1, r1, and r2 with time step 0.5 are the most competitive, despite their large time step. For parameter set B, we know that biases are roughly ten times larger than for parameter set A. This has an impact on the RMS errors, as in this case the best schemes are r2 and lr1 which have small biases. The errors range from 5% to 15% so the choice of scheme is important.

Table 1. Relative RMS errors in caplet pricing. $T = 2$ years, 1 second computing time

	Jump parameters set A				Jump parameters set B		
Scheme	Time step	Paths	Rel. RMS	Scheme	Time step	Paths	Rel. RMS
r1	0.1	10972	0.023	r1	0.1	3773	0.071
r1	0.5	20338	0.020	r1	0.5	4347	0.155
r2	0.5	14345	0.021	r2	0.5	3285	0.054
lr1	0.5	15288	0.020	lr1	0.5	3557	0.052
lr2	0.5	11678	0.023	lr2	0.5	2793	0.059
b1	0.1	6986	0.030	b1	0.1	2590	0.076
b1	0.5	15274	0.023	b1	0.5	3102	0.137
lb1	0.1	5500	0.034	lb1	0.1	2262	0.082
lb1	0.5	13386	0.024	lb1	0.5	2816	0.153

Table 2. Relative RMS errors in caplet pricing. $T = 10$ years, 1 minute computing time

	Jump parameters set A				Jump parameters set B		
Scheme	Time step	Paths	Rel. RMS	Scheme	Time step	Paths	Rel. RMS
r1	0.1	69718	0.012	r1	0.1	22189	0.072
r1	0.5	155440	0.015	r1	0.5	27272	0.165
r2	0.5	107385	0.009	r2	0.5	19672	0.026
lr1	0.5	76433	0.011	lr1	0.5	17065	0.025
lb1	0.1	22758	0.019	lb1	0.1	8207	0.077
lb1	0.5	59820	0.018	lb1	0.5	10016	0.162

Next we consider longer computations. Table 2 shows estimated relative RMS error for a 10-year caplet and 1 minute of computing time. This is arbitrary; circumstances may demand faster or allow more accurate pricing. By combining the information in the tables (essentially the number of paths per unit time) and the biases of the previous section, it is possible to estimate RMS errors for other budgets. Errors for parameter set A are roughly 1% for all schemes, though r2 and lr1 are slightly better than the rest. Errors for parameter set B again reflect the importance of bias. The best schemes are r2 and lr1, which have RMS around 2%, significantly smaller than the others.

8 Conclusions

We have developed, analyzed, and tested computational procedures for the numerical solution of LIBOR market models with jumps. To carry this out, we have first reformulated a term structure model driven by marked point processes with state-dependent intensities into one driven by a Poisson random measure. This facilitates the development of discretization schemes because the Poisson random measure can be simulated without discretization error. Jumps in LIBOR rates are thinned from the Poisson random measure using state-dependent thinning probabilities. Because of discontinuities inherent to the thinning process, this procedure falls outside the scope of existing convergence results; we provide a measure of theoretical support for our method through a result establishing first and second order convergence of schemes that accommodates thinning but imposes stronger

conditions on other problem data. The results of numerical experiments indicate that the most computationally attractive methods are a second-order scheme for rates and a first-order scheme for log rates.

References

Andersen, L., Andreasen, J.: Volatility skews and extensions of the libor market model. Appl. Math. Finance **7**, 1–32 (2000)

Björk, T., Kabanov, Y., Runggaldier, W.: Bond market structure in the presence of marked point processes. Math. Finance **7**, 211–239 (1997)

Black, F.: The Pricing of Commodity Contracts, J. Financial Econ. **3**, 167–179 (1976)

Brace, A., Gatarek, D., Musiela, M.: The market model of interest rate dynamics. Math. Finance **7**, 127–155 (1997)

Brémaud, P.: Point processes and queues: Martingale dynamics. Springer, Berlin Heidelberg New York (1981)

Das, S.R.: The surprise element: jumps in interest rate diffusions. working paper, Harvard Business Schooly (1999)

Duffie, D., Glynn, P.: Efficient Monte Carlo estimation of security prices. Ann. Appl. Prob. **5**, 897–905 (1995)

Glasserman, P., Kou, S.G.: The term structure of simple forward rates with jump risk. Working paper. New York: Columbia University (1999) (Available at www.paulglasserman.com)

Glasserman, P., Merener, N.: Numerical solution of jump-diffusion LIBOR market models: Addendum, (2001) (Available at www.paulglasserman.com)

Glasserman, P., Wang, H.: Discretization of deflated bond prices. Adv. Appl. Prob. **32**, 540–563 (2000)

Glasserman, P., Zhao, X.: Arbitrage-free discretization of lognormal forward LIBOR and swap rate models. Finance Stochast. **4**, 35–69 (2000)

Heath, D., Jarrow, R., Morton, A.: Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. Econometrica **60**, 77–105 (1992)

Jamshidian, F.: LIBOR swap market models and measures. Finance Stochast. **1**, 293–330 (1997)

Jamshidian, F.: LIBOR market model with semimartingales. Working paper, London: NetAnalytic Ltd (1999)

Johannes, M.S.: Jumps in interest rates: a nonparametric approach. To appear in J. Finance (1999)

Kloeden, P. E., Platen, E.: Numerical solution of stochastic differential equations. Berlin Heidelberg New York: Springer (1992)

Kou, S.G.: A jump diffusion model for option pricing. To appear in Mgmt. Sci. (1999)

Maghsoodi, Y.: Exact solutions and doubly efficient approximations of jump-diffusion ito equations. Stochast. Anal. Appl. **16**, 1049–1072 (1998)

Merton, R.: Option pricing when underlying stock returns are discontinuous. J. Financial Econ. **3**, 125–144 (1976)

Mikulevicius, R., Platen, E.: Time discrete taylor approximations for ito processes with jump component. Math. Nachrichten **138**, 93–104 (1988)

Milstein, G.N.: A method of second-order accuracy integration of stochastic differential equations. Theory Prob. Appl. **19**, 557–562 (1978)

Miltersen, K.R., Sandmann, K, Sondermann, D.: Closed-form solutions for term structure derivatives with lognormal interest rates. J. Finance **52**, 409–430 (1997)

Musiela, M., Rutkowski, M.: Continuous-time term structure models: forward measure approach. Finance Stochast. **1**, 261–292 (1997)

Musiela, M., Rutkowski, M.: Martingale methods in financial modeling. Springer: Berlin Heidelberg New York 1997

Platen, E., Rebolledo, R.: Weak convergence of semimartingales and discretisation methods. Stochast. Proc. Appl. **20**, 41–58 (1985)

Protter, P., Talay, D.: The Euler scheme for Lévy driven stochastic differential equations. Ann. Prob. **25**, 393–423 (1997)

Shirakawa, H.: Interest rate option pricing with poisson-gaussian forward rate curve processes. Math. Finance **1**, 77–94 (1991)

Talay, D. Tubaro, L.: Expansion of the global error for numerical schemes solving stochastic differential equations. Stochast. Anal. Appl. **8**, 483–509 (1990)

Zühlsdorff, C.: Extended LIBOR market models with affine and quadratic volatility. Working paper Bonn: University of Bonn (1999)