Convergence of a Discretization Scheme for Jump-Diffusion Processes with State-Dependent Intensities

Paul Glasserman* and Nicolas Merener†
Columbia University

Abstract

This paper proves a convergence result for a discretization scheme for simulating jump-diffusion processes with state-dependent jump intensities. With a bound on the intensity, the point process of jump times can be constructed by thinning a Poisson random measure using state-dependent thinning probabilities. Between the jump epochs of the Poisson random measure, the dynamics of the constructed process are purely diffusive and may be simulated using standard discretization methods. Under conditions on the coefficient functions of the jump-diffusion process, we show that the weak convergence order of this method equals the weak convergence order of the scheme used for the purely diffusive intervals: the construction of jumps does not degrade the convergence of the method.

1 Introduction

This paper proves a convergence result for a discretization scheme used in the numerical simulation of a class of jump-diffusion processes with state-dependent jump intensities. Under a boundedness condition on the intensity, the point process of jump times can be constructed by thinning a Poisson random measure using state-dependent thinning probabilities. Between the jump epochs of the Poisson random measure, the dynamics of the constructed process are purely diffusive and may be simulated using standard discretization methods. At each jump epoch, a candidate jump is accepted or rejected according to a state-dependent thinning probability evaluated at the approximating discretized process. Under conditions on the coefficient functions of the jump-diffusion process (and the class of admissable "test functions"), we show that the weak convergence order of this method equals the weak convergence order of the scheme used for the purely diffusive intervals. In other words, the jumps do not degrade the convergence of the method.

The method we consider was proposed and numerically tested in Glasserman and Merener [6] in the context of a class of interest rate models with jumps. Numerical simulation is required

*403 Uris Hall, Graduate School of Business, Columbia University, New York, NY 10027, pg20@columbia.edu.
†Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, nm187@columbia.edu

there for the pricing of interest rate derivative securities. This paper therefore provides rigorous support for the numerical results in our earlier paper, proving a convergence result stated without proof in [6].

The analysis in this paper is not limited to the specific model investigated in [6]. However, to provide some motivation for the class of processes we consider, in Section 2 we give a brief overview of stochastic models with jumps arising in mathematical finance and their implications for simulation. Section 3 specifies the general model dynamics we consider. Section 4 explains the thinning construction and Section 5 presents discretization methods. The main convergence result is stated in Section 6 and proved in Section 7.

2 Jumps in Financial Modeling

Among the earliest investigations of continuous-time models of asset prices with jumps is the work of Merton [13]. To model the dynamics of a stock price subject to occasional large moves, Merton proposed the stochastic differential equation

$$\frac{dS(t)}{S(t-)} = (\mu - \lambda m)\, dt + \sigma\, dW(t) + d\left(\sum_{j=1}^{N(t)} (Y_j - 1)\right), \tag{1}$$

in which W is a standard Brownian motion; N is a Poisson process independent of W with constant arrival rate λ; the Y_j are i.i.d. positive random variables with mean $m+1$, independent of N and W; the constant μ is the instantaneous expected rate of return for the stock and σ is a constant volatility parameter. The compound Poisson process

$$J(t) = \sum_{j=1}^{N(t)} (Y_j - 1)$$

has bounded variation so the differential $dJ(t)$ may be understood in the usual Stieltjes sense. We take N and S to be right-continuous; the left limit $S(t-)$ of S at t is the value just before a possible jump at t. If the jth jump of N occurs at τ_j, then the dynamics in (1) specify $S(\tau_j) = S(\tau_j -)Y_j$. Requiring that the Y_j be positive ensures that the stock price does not jump to a negative value.

The term λm in the drift of (1) compensates the jumps in the sense that $J(t) - \lambda m t$ is a martingale. The presence of this term in the drift is important in the financial interpretation of the model. Related but much more general terms arise in the dynamics considered in [6].

The solution to the stochastic differential equation (1) is

$$S(t) = S(0) \exp\left((\mu - \lambda m - \frac{1}{2}\sigma^2)t + \sigma W(t)\right) \prod_{j=1}^{N(t)} Y_j. \tag{2}$$

Merton [13] used this solution to derive a formula for the value of an option on S. From (2) it becomes evident that one may simulate paths of S over a fixed set of dates without discretization error by simulating the increments of W (normal random variables), the increments of N (Poisson random variables), and values of the Y_j. Important special cases for the distribution of the Y_j are the lognormal considered by Merton [13] and the asymmetric log-Laplace considered by Kou [8].

One direction for generalizing Merton's model (1) notes that (2) is a process of the form $S(0)\exp(X(t))$ with X a Lévy process, a process with stationary independent increments. This suggests the possibility of replacing X with a more general Lévy process. As a consequence of the Lévy-Khinchine theorem, the only possible such generalization is to replace the compound Poisson process J with a process having an infinite number of jumps in every interval. Specific models of this type (in which the Brownian component is dropped altogether) have been developed by Barndorff-Nielsen [1] and Madan and Seneta [10]. These admit a fair amount of tractability. Simulation at a fixed set of dates remains feasible given methods for sampling the increments of the Lévy process. This is equivalent to the general problem of sampling from infinitely divisible distributions.

An alternative direction for generalizing (1) keeps the number of jumps finite over finite time intervals and relaxes the requirement that the timing of jumps be independent of the level of the asset price. In the class of affine jump-diffusions (see Duffie, Pan, and Singleton [4] and references there), the jump intensity is an affine function of the state, which may be vector-valued. The affine framework encompasses many specific models proposed in the finance literature. Affine state-dependence leads to a high degree of tractability, primarily through transform inversion. The feasibility of simulation without discretization error depends on the choice of model within this framework. In the general case, simulating state transitions of an affine jump-diffusion entails sampling from distributions known only through their characteristic functions.

The setting developed by Björk, Kabanov, and Runggaldier [2] is among the most general models of asset-price dynamics with jumps and provides part of the motivation leading to the numerical method we consider here and in [6]. Björk et al. [2] model jumps through a random measure $\mu(dz, dt)$ on the product of an abstract mark space E^* and the time axis $[0, \infty)$. Think of $\mu(dz, dt)$ as assigning unit mass to (z, t) if a mark z arrives at time t. Suppose, for example, that jumps arrive at distinct, ordered times $\tau_1 < \tau_2 < \cdots$ with marks $Z_1, Z_2, \ldots$. Let $N(t)$ count the number of jumps in $[0, t]$ and let h denote a real-valued function on E^*. Then within this formalism we can write the counting process

$$\sum_{j=1}^{N(t)} h(Z_j)$$

as

$$\int_0^t \int_{E^*} h(z)\, \mu(dz, dt).$$

A stochastic intensity for $\mu(dz, dt)$ is a measure-valued process $\nu(dz, t)$ that compensates μ in the sense that

$$\int_0^t \int_{E^*} h(z)\, \mu(dz, dt) - \int_0^t \int_{E^*} h(z)\, \nu(dz, t)dt \tag{3}$$

is a martingale (in t) for all bounded h. The compound Poisson process in (1) can be represented within this framework by a Poisson random measure on $\Re \times \Re_+$ and is characterized by having as its intensity $\lambda f(z)\, dz$, with f the common density of the Y_j.

Glasserman and Kou [5] use the framework of Björk et al. [2] to develop a class of interest rate models with jumps. Their use of this very general framework requires comment. An essential tool within both the theory and practice of pricing derivative securities is the ability to describe how asset dynamics transform under certain changes of measure (see, e.g., Musiela and Rutkowski [16]). In a pure diffusion model, the change of measure typically leads to a change of drift; with jumps we also get a change of intensity (see, e.g., [3]). Even if jumps under one measure are Poisson, this property is typically not preserved by the relevant changes of measure, because these transformations introduce state-dependence in the intensity. As in Glasserman and Merener [6], we may restrict ourselves to the case in which the intensity is stochastic only to the extent that depends on the current state because this property *is* preserved by the relevant changes of measure. (The dynamics we consider are made explicit in the next section.) In the general framework of Björk et al. [2] the intensity could depend more generally on, e.g., the past of the process; at the other extreme, a deterministic intensity characterizes the case of a Poisson random measure.

3 Model Dynamics

We now give a concise formulation of the class of jump-diffusion models with state-dependent intensities that are the focus of our attention in this paper. The dynamics of the M-dimensional right-continuous process $X(t), t \in [0, T]$, with fixed initial condition $X(0)$, are given by

$$dX(t) = \tilde{a}(X(t-))dt + b(X(t-))dW(t) + \int_{E^*} H(X(t-), z)\mu(dz, dt). \tag{4}$$

The functions $\tilde{a} : \Re^M \to \Re^M$, $b : \Re^M \to \Re^{M \times d}$, and $H : \Re^M \times E^* \to \Re^M$ are deterministic and will be subject to regularity conditions in Theorem 6.1 below; $W(t) \in \Re^d$ is a vector Brownian motion with independent components, and $\mu(dz, dt)$ a random measure defined on the product of the mark space E^* and the time axis. To avoid inessential complications, we take E^* to be a subset of Euclidean space. As explained in the previous section, think of $\mu(dz, dt)$ as assigning unit mass to (z, t) if a mark z arrives at time t. We assume that the random measure μ has intensity $\nu(z, X(t-), t)\, dz$ for a deterministic non-negative function ν.

Relative to (3), this use of ν represents a relatively minor abuse of notation. The key point is that we now assume that ν is a deterministic function; the stochastic intensity is obtained by evaluating this function at the current state $X(t-)$. Of secondary importance is the assumption that $\nu(\cdot, X(t-), t)$ is absolutely continuous so that we may write $\nu(z, X(t-), t)\, dz$ for $\nu(dz, X(t-), t)$. This is primarily for notational convenience.

A special case of (4) is a model driven by a Poisson random measure $p(dz, dt)$ with a deterministic intensity $\lambda_P(dz, t)$ such that $\lambda_P(dz, t) = \lambda_0(t)\, f(z)\, dz$, with $f(z)$ a probability density on E^*. Thus, the arrival times follow a Poisson process with deterministic (possibly time-varying) intensity $\lambda_0(t)$, and the marks are independent and distributed with density f.

In the Poisson case, the arrival rate is independent of the state of the system. We will see that this leads to important computational gains because the random jumps can be sampled exactly, for each path, independently from the state of the system along the path. This desirable decoupling does not hold in the general case (4) where the arrival rate of the jumps depends on the state of the system. As mentioned in Section 2, state-dependent jump arrival rates may appear in the dynamics of financial variables through the change in intensity produced by a change of measure.

A standard problem in finance is the valuation of a derivative security as an expectation

$$E[g(X(T))], \tag{5}$$

for some discounted payoff function (or "test function") g. Models of the form (4) (even if the driving jump measure is Poisson) are usually not solvable, in the sense that formulas for (5) are not available in closed form. In principle, (5) could be evaluated through numerical integration; however, this requires the distribution of $X(T)$ which is itself intractable for (4).

This leaves Monte Carlo simulation as the only general approach to computing (5). A large number of paths of discrete-time approximations of X are generated numerically, and the expected value in (5) is estimated as the sample average computed on the terminal states of the simulated paths. A crucial element in this approach is the construction of a discrete-time approximation to X, a task that is simplified if the model is driven by a Poisson random measure. In this case, the itensity is deterministic therefore, as noted earlier, an algorithm that generates jump times of the Poisson measure p does not need to know about the discrete solution, and the jump times can be generated exactly.

In order to apply this concept to the solution of a general model (4), we will discuss next how the random measure associated to a state-dependent intensity of interest may nevertheless be constructed through state-dependent thinning of a Poisson random measure.

4 State-Dependent Thinning

Models with state-dependent intensities provide a general formalism for constructing models with jumps, but are 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 the approach in Glasserman and Merener [6] in which a Poisson random measure is used to construct more general models with state-dependent intensities. The idea is to generate a driving jump process associated with a state-dependent intensity from a Poisson random measure and a state-dependent *thinning* mechanism. A thinning function θ randomly accepts or rejects the marks of the Poisson process with probability proportional to the value of the state-dependent intensity at the moment of the jump. The resulting process of accepted marks has the required state-dependent law.

We begin with the following Poisson driven model:

$$dX(t) = \tilde{a}(X(t-))\,dt + b(X(t-))\,dW(t) + \int_{E^*}\int_0^1 H(X(t-),y)\,\theta(y,u,X(t-),t)\,p(dy\times du,dt) \quad (6)$$

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 $\tilde{a}$, b, and H are as in (4). We will use the additional mark component u to implement the acceptance-rejection decision in the thinning construction.

Next, we make a crucial assumption. Let the intensity function ν for (4) satisfy

$$\nu(y,x_1,\ldots,x_M,t) < \lambda_0 f(y) \qquad \text{with} \qquad k = 1,\ldots,M. \quad (7)$$

Under this assumption, we can use ν to define the deterministic thinning function θ; it acts on $X(t)$ as

$$\theta(y,u,X(t-),t) = \begin{cases} 1, & u < \frac{\nu(y,X(t-),t)}{f(y)\lambda_0} \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

The interpretation of the thinning function θ is as follows. Associated with each jump time of the Poisson random measure is a mark (y,u). Because of nonnegativity of the function ν and definition (8), and conditional on y and $X(t-)$, the probability of θ being nonzero at a jump time of the Poisson process is, from (8), $\nu(y, X(t-), t)/\lambda_0 f(y)$. Intuitively, the jump process associated to the random measure

$$\mu(dy, dt) = \int_0^1 \theta\,(y, u, X(t-), t)\, p(dy \times du, dt) \tag{9}$$

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

Using (9), we may rewrite (6) as

$$dX(t) = \tilde{a}(X(t-))\, dt + b(X(t-))\, dW(t) + \int_{E^*} H(X(t-), y)\mu(dy, dt) \tag{10}$$

which is formally equivalent to (4). Proposition 3.1 in Glasserman and Merener [6] verifies that (4) and (10) (and therefore (6)) are indeed the same model under this construction of μ. Thus, under the assumption (7), it is possible to write a model driven by a random measure with state-dependent intensity as a model driven by a Poisson random measure.

5 Discretization Schemes

Now we turn our attention to the numerical solution of (6), for which we draw on Mikulevicius and Platen [14]. In order to make our notation consistent with theirs, we aggregate the effect of H and θ in a function c. We consider then the M-dimensional process $X(t)$, $t \in [0, T]$ that follows

$$dX(t) = \tilde{a}(X(t))\, dt + b(X(t))\, dW(t) + \int_E c(X(t), z)p(dz, dt) \tag{11}$$

where $p(dz, dt)$ is a Poisson random measure on $E \times [0, T]$ with intensity $\lambda_0\, h(z)$ 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 $\tilde{a}, b, \text{and}\, c$ are M-dimensional vectors with components $\tilde{a}_j, b_j, \text{and}\, c_j$. An explicit time-dependence in the coefficients of (11) could be accommodated by, e.g., including time as a component of the vector $X(t)$.

We construct approximate solutions to models of the form (11) at a discrete set of times $\{\tau_i\}$. This set is the superposition of the random jump times of a Poisson process on $[0, T]$ and a deterministic grid $T_1, \ldots, T_M$. As stressed before, the random Poisson jump times can be computed without any knowledge of the realized path of (11). Mikulevicius and Platen [14] (see also [11, 12, 17, 18]) introduced explicit schemes that generate approximate solutions $Y(\tau_i)$ of (11) on the grid points τ_i. The main distinction between our work and theirs is that they imposed smoothness conditions on the function c in (11) which are violated by the discontinuous nature of the thinning construction.

We measure the quality of discretization schemes through a weak convergence criterion, which is appropriate for the computation of derivatives prices (5). A scheme $\{Y(\tau_i)\}$ is said to have weak order of convergence ξ if for all sufficiently small ϵ

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

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 [7]).

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

$$Y(\tau_{i+1}^-) = Y(\tau_i) + f_0(Y(\tau_i))(\tau_{i+1} - \tau_i) + f_1(Y(\tau_i))(W_{\tau_{i+1}} - W_{\tau_i}), \tag{12}$$

$$Y(\tau_{i+1}) = Y(\tau_{i+1}^-) + \int_E c(Y(\tau_{i+1}^-), z)p(dz, \tau_{i+1}) \tag{13}$$

$$f_0(Y(\tau_i)) = \tilde{a}(Y(\tau_i)) \text{ and } f_1(Y(\tau_i)) = b(Y(\tau_i)). \tag{14}$$

At each grid point, (13) 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 integral in (13) entails at most a single evaluation of the function c because $p(dz, \tau_{i+1})$ is a point mass at the mark z that arrives at τ_{i+1} if τ_{i+1} is a jump epoch.)

Next we present the generalization of the Milstein [15] scheme proposed by Mikulevicius and Platen [14], 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

$$\begin{aligned} Y(\tau_{i+1}^-) &= \\ & Y(\tau_i) + f_0(Y(\tau_i))(\tau_{i+1} - \tau_i) + f_1(Y(\tau_i))Z_i + f_{00}(Y(\tau_i))\frac{1}{2}(\tau_{i+1} - \tau_i)^2 \\ + \quad & f_{10}(Y(\tau_i))U_i + f_{01}(Y(\tau_i))(Z_i(\tau_{i+1} - \tau_i) - U_i) + f_{11}(Y(\tau_i))\frac{1}{2}(Z_i^2 - (\tau_{i+1} - \tau_i)) \end{aligned} \tag{15}$$

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 a bivariate normal distribution. The updating of the rates at a jump time is as in (13). 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 (14). Writing ∂_j for a partial derivative with respect to X_j, the others are

$$f_{00}(Y) = \sum_{j=1}^{M} \tilde{a}_j(Y)\partial_j \tilde{a}(Y) + \frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M} b_j(Y)b_k(Y)\partial_{jk}\tilde{a}(Y), \quad f_{11}(Y) = \sum_{j=1}^{M} b_j(Y)\partial_j b(Y),$$

$$f_{10}(Y) = \sum_{j=1}^{M} b_j(Y)\partial_j \tilde{a}(Y), \quad f_{01}(Y) = \sum_{j=1}^{M} \tilde{a}_j(Y)\partial_j b(Y) + \frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M} b_j(Y)b_k(Y)\partial_{jk}b(Y). \quad (16)$$

6 Convergence

Mikulevicius and Platen [14] introduced a hierarchy of schemes which, under regularity conditions on $\tilde{a}, b, c$ 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 [14] 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.

Define

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

so the dynamics (11) can be written (as in [14]) as

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

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, $E = [0, \infty) \times (0, 1)$.

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.

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

$$\phi(x) = \int_E \psi(x + c(x,z))\, h(z)\, dz \quad (17)$$

and let

$$\bar{\phi}(x) = \int_E (x + c(x,z))\, h(z)\, dz. \quad (18)$$

With the notation above, we present a weak convergence result, stated without proof in Glasserman and Merener [6], where the method was tested numerically on a class of interest rate models.

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 (11). 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 (17);

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

(iv) there is a constant K_2 such that, if the functions $f(y)$ in S_ξ are well defined, they satisfy $|f(y)| \leq K_2(1 + \|y\|)$, with $S_1 = \{f_0, f_1\}$ as in (14) and $S_2 = \{f_0, f_1, f_{00}, f_{10}, f_{01}, f_{11}\}$ as in (16).

Then the approximation defined by (12)-(14) has weak convergence order one and the approximation defined by (15)-(16) has weak convergence order two.

Before proceeding with the proof of Theorem 6.1 we briefly discuss its connection with Theorem 3.3 of Mikulevicius and Platen [14] which is also a weak convergence result. They have shown that, under regularity conditions on a, b, c and the payoff function g, the scheme defined by (12)-(14) has weak convergence order one and the scheme defined by (15)-(16) has weak convergence order two. More precisely, Mikulevicius and Platen [14] assume that the payoff g is $2(\xi + 1)$-times continuously differentiable and with polynomial growth, and that the coefficients a, b, and c satisfy:

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

(b) there is a constant K_2 such that the functions $f(y)$ in S_ξ satisfy $|f(y)| \leq K_2(1 + \|y\|)$, with $S_1 = \{f_0, f_1\}$ as in (14) and $S_2 = \{f_0, f_1, f_{00}, f_{10}, f_{01}, f_{11}\}$ as in (16).

We will later show (in the proof of Theorem 6.1) that assumptions (i) and (iii) of Theorem 6.1 guarantee that the functions $f \in S_\xi$ are well defined. This fact, and assumption (iv) in Theorem 6.1, are equivalent to assumption (b) of Mikulevicius and Platen. Also, it is clear that regularity conditions for a and b are identical for both convergence results. The results differ in their requirements for c and the payoff g. Theorem 6.1 allows for discontinuous c, though it imposes stronger requirements on g. Next, we will show explicitly that the requirements for c in Mikulevicius and Platen indeed imply assumptions (i) and (ii) in Theorem 6.1.

PROPOSITION 6.1 *If the function c in (11) is* $2(\gamma+1)$*-times continuously differentiable with derivatives uniformly bounded by a constant* C*, then assumptions (i) and (ii) in Theorem 6.1 are satisfied.*

Proof. We denote by $1_k \in \Re^M$ the vector with kth component equal to one, and the rest equal to zero. For (i) in Theorem 6.1 observe that

$$\frac{\partial \bar{\phi}(x)}{\partial x_k} = \frac{\partial}{\partial x_k}\int_E (x + c(x,z))\, h(z)\, dz = \int_E (1_k + \frac{\partial}{\partial x_k} c(x,z))\, h(z)\, dz \quad k = 1\ldots, M \tag{19}$$

where the boundedness of $\frac{\partial}{\partial x_k} c(x,z)$ has allowed us to invoke the Bounded Convergence Theorem and exchange differentiation and integration to show that $\frac{\partial \bar{\phi}(x)}{\partial x_k}$ exists, and can be written as in the rightmost expression in (19). Furthermore, since $\frac{\partial}{\partial x_k} c(x,z)$ is uniformly bounded, and $h(z)$ is a probability density, then $\frac{\partial \bar{\phi}(x)}{\partial x_k}$ is also uniformly bounded. The same argument applies to show that derivatives of $\bar{\phi}$ up to order $2(\gamma+1)$ exist and are uniformly bounded. Therefore, assumption (i) in Theorem 6.1 holds.

For assumption (ii) in Theorem 6.1 we take $\psi \in B^{\xi}(\Psi)$. That ϕ in (17) is uniformly bounded follows from the fact that ψ is uniformly bounded. Next, we consider the derivatives of ϕ. By the chain rule, we have that

$$\frac{\partial}{\partial x_k}\psi(x + c(x,z)) = \nabla\psi \cdot (1_k + \frac{\partial c(x,z)}{\partial x_k})$$

wich is bounded by $M\Psi(1+C)$. Therefore, we can invoke the Bounded Convergence Theorem to write

$$\frac{\partial \phi(x)}{\partial x_k} = \frac{\partial}{\partial x_k}\int_E \psi(x + c(x,z))\, h(z)\, dz = \int_E \frac{\partial}{\partial x_k}\psi(x + c(x,z))\, h(z)\, dz$$

Furthermore, $|\frac{\partial \phi(x)}{\partial x_k}| < M\Psi(1+C)$. Computation of the bound for higher order derivatives is straightforward. Therefore, assumption (ii) in Theorem 6.1 holds. □

To summarize, we have shown that hypotheses (i) and (ii) in Theorem 6.1 are implied by the regularity condition for c in Theorem 3.3 of Mikulevicius and Platen [14]. But, as mentioned above, Theorem 6.1 is obtained under more restrictive conditions for g than in the result of Mikulevicius and Platen [14].

Next we present a proof of Theorem 6.1. The proof holds in fact for the entire hierarchy of schemes proposed in Mikulevicius and Platen [14], which have arbitrarily high orders of convergence. Schemes of order higher than two are constructed using the funcions f in S_ξ which are defined in a recursive way in [14].

7 Proof of Theorem 6.1

The result follows from the proof of Theorem 3.3 of Mikulevicius and Platen [14] 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 v of a backward Kolmogorov equation and associated functionals.

The stochastic Taylor formula introduced in Section 2 of Mikulevicius and Platen [14] requires that the coefficients functions $f \in A_{\xi+1}$ (in the notation of Section 2 in [14]) are well defined. The functions in $A_{\xi+1}$ are computed through up to $\xi+1$ recursive applications of the differential operators

$$\Pi^0 f(x) \equiv \sum_{j=1}^{M} \tilde{a}_j(x)\partial_j f(x) + 1/2 \sum_{i,r=1}^{M} b_j(x) b_r(x)\partial_{jr} f(x), \quad \text{and} \quad \Pi^1 \equiv \sum_{j=1}^{M} b_j(x)\partial_j f(x), \tag{20}$$

and involve derivatives up to order 2ξ. These coefficient functions include those in S_1 and S_2, which are used in the construction of the truncated Taylor expansion.

The solution of a Kolmogorov equation used in Lemma 4.3 of Mikulevicius and Platen [14] is

$$v(s,x) \equiv E\Big[g(X(T))|X(s)=x\Big],$$

and we will also use the functional

$$\Pi v(s,x) \equiv \int_E \Big[v(s, x + c(x,z)) - v(s,x)\Big]\, h(z)\ \lambda_0\, dz. \tag{21}$$

The following lemma presents sufficient conditions on g, v, Πv and the coefficients functions $f \in A_{\xi+1}$ in order to prove Theorem 6.1 using the proof of Mikulevicius and Platen [14].

LEMMA 7.1 *Fix $\xi \in \{1,2\}$. Let the payoff function $g : \Re^M \to \Re$ be in $B^\xi(G)$ and let $\{X(t), t \in [0,T]\}$ be as in (11). We assume:*

(i) $v(s,x)$ and $\Pi v(s,x)$ are $2(\xi+1)$-times continuously differentiable in the initial condition x;

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

(iii) the functions $f \in A_{\xi+1}$ are well defined;

(iv) there is a constant K_2 such that the functions $f(y)$ in S_ξ satisfy $|f(y)| \leq K_2(1 + \|y\|)$, with $S_1 = \{f_0, f_1\}$ as in (14) and $S_2 = \{f_0, f_1, f_{00}, f_{10}, f_{01}, f_{11}\}$ as in (16).

Then the approximation defined by (12)-(14) has weak convergence order one and the approximation defined by (15)-(16) has weak convergence order two.

Proof of lemma. We check that the assumptions of the lemma guarantee that the proof of Theorem 3.3 in Mikulevicius and Platen also applies to prove this result.

First, $g \in B^\xi(G)$ implies that g is $2(\xi + 1)$-times continuously differentiable and with polynomial growth, as required in Theorem 3.3 of [14].

Next, Assumption (iii) guarantees the existence of a stochastic Taylor formula as introduced in Section 2 of Mikulevicius and Platen [14] and used in Section 5 of [14].

Last, because of (i), ($2(\xi+1)$-times continuous differentiability of v), Lemma 4.3 and Lemma 4.8 in Mikulevicius and Platen [14] hold.

Therefore, Lemma 7.1 is proved by Section 6 of Mikulevicius and Platen [14], which is the proof of the convergence result (Theorem 3.13) in [14], and where assumption (i) ($2(\xi+1)$-times continuous differentiability of v and Πv), is used to apply the Taylor formula in (6.6) and (6.13) of [14]. □

Armed with the lemma above, it is clear that to prove Theorem 6.1 it will suffice to show that assumptions of Theorem 6.1 imply the assumptions of Lemma 7.1. In particular, we need to check that the functions f in $A_{\xi+1}$ (and therefore those in S_ξ) are well defined and that smoothness properties of v and Πv hold. Verification of these two properties divides the proof into two parts.

Part 1 of proof. As mentioned above, the stochastic Taylor formula requires that the coefficients functions $f \in A_{\xi+1}$, computed through the repeated application of Π^0 and Π^1 in (20), are well defined. In order to be able to apply Π^0 and Π^1 it will suffice to show 2ξ-times continuous differentiability of

$$\tilde{a}(x) = a(x) - \int_E c(x, z)\, h(z)\, \lambda_0\, dz.$$

From (18) we have that

$$\int_E c(x, z)\, h(z)\, \lambda_0\, dz = \lambda_0\, \bar{\phi}(x) - \lambda_0\, x$$

leading to

$$\tilde{a}(x) = a(x) - \lambda_0\, \bar{\phi}(x) + \lambda_0\, x\,. \tag{22}$$

Therefore, both $\tilde{a}$ and the right hand side of the equality are $2(\xi + 1)$-times continuously differentiable, the latter by assumptions (i) and (iii) in Theorem 6.1.

This ends the first part of the proof. Next we introduce an auxiliary result that we will later use to prove the $2(\xi+1)$-fold continuous differentiability of the solution v of a Kolmogorov equation and associated functionals (21).

We have assumed in Theorem 6.1 that a, b and $\bar{\phi}$ are $2(\xi+1)$-times continuously differentiable with bounded derivatives, so (22) implies that the derivatives of $\tilde{a}$ are also bounded. We have also assumed in Theorem 6.1 that both $f_0 = \tilde{a}(y)$ and $f_1 = b(y)$ are of linear growth. Therefore, the assumptions of the following auxiliary lemma are implicit in the hypotheses of Theorem 6.1.

LEMMA 7.2 *Let $Z(t) \in \Re^M$ be an Ito process, $t \in [0,T], Z(0) = x$ a.s., with*

$$dZ(t) = \tilde{a}(Z(t))\, dt + b(Z(t))\, dW(t) \tag{23}$$

where the functions $\tilde{a}$, b are $2(\xi+1)$-times continuously differentiable with uniformly bounded partial derivatives. Let a, b be of linear growth; i.e., $\|\tilde{a}(y)\| + \|b(y)\| \leq K_2\,(1+\|y\|)$ for some constant K_2. Let $g : \Re^M \to \Re$ be in $B^\xi(G)$ and define $\phi^D(s,t,x) = E[g(Z(t))|Z(s) = x]$. Then $\phi^D(s,t,\cdot) \in B^\xi(D(\xi)\,G)$ where $D(\xi)$ is independent of g

Proof of lemma. We need to prove that ϕ^D is bounded and that $\phi^D(s,t,\cdot)$ has continuous bounded partial derivatives up to order $2(\xi+1)$. That $|\phi^D| \leq G$ follows from the fact that $|g| \leq G$.

We analyze the derivatives of ϕ^D within the framework of Chapter V of Krylov [9] in which, under technical conditions, it is possible to exchange differentiation and expectation. We introduce the following notation. The vector $\eta^{k_1 \dots k_n}(t) \in \Re^M$ is obtained by formally differentiating $Z(t)$ with respect to the components $x_{k_1}, \dots, x_{k_n}$ of x, $n \in \{1, \dots, 2(\xi+1)\}$. These processes will be derivatives of $Z(t)$ in probability though we will not need to check this explicitly. We need to show that these processes are *solvable by Euler's method in the mean* (SEM) as defined in V.3 of Krylov [9]. The dynamics of the i-th component of the first order derivative process is

$$d\eta_i^{k_1} = (\nabla \tilde{a}_i \cdot \eta^{k_1})\, dt + (\nabla b_i \cdot \eta^{k_1})\, dW \tag{24}$$

with $\eta_{k_1}^{k_1}(0) = 1, \eta_j^{k_1}(0) = 0$ for $j \neq k_1$, $\nabla \tilde{a}_i$, $\nabla b_i \in \Re^M$. The components of the second order derivative processes evolve as

$$d\eta_i^{k_1 k_2} = (\eta^{k_2^\top} \cdot (\nabla\nabla \tilde{a}_i) \cdot \eta^{k_1}) + \nabla \tilde{a}_i \cdot \eta^{k_1 k_2})dt + (\eta^{k_2^\top} \cdot (\nabla\nabla b_i) \cdot \eta^{k_1} + \nabla b_i \cdot \eta^{k_1 k_2})\, dW$$

with $\eta^{k_1 k_2}(0) = 0$ (and $\nabla\nabla$ denoting the Hessian). In general, for each $k_1, \dots, k_n$ through repeated differentiation we define a^* and b^* and then

$$d\eta_i^{k_1 k_2 \dots k_n} = \tilde{a}_i^*\, dt + b_i^*\, dW$$

with $\eta^{k_1 k_2 \dots k_n}(0) = 0$ for derivatives of order higher than two. It is easy to check, as is clear for the lowest two derivatives, that for a derivative of order n, both $\tilde{a}^*$ and b^* are polynomials in the derivative processes of order less than n and affine functions of $\eta^{k_1 k_2 \dots k_n}$. Viewing a^* and b^* as functions of $\eta^{k_1 k_2 \dots k_n}$ with all other derivative processes held fixed, we therefore have

$$\begin{aligned} &\|\tilde{a}^*(\eta^{k_1 k_2 \dots k_n})\| + \|b^*(\eta^{k_1 k_2 \dots k_n})\| \le K_1(1 + \|\eta^{k_1 k_2 \dots k_n}\|), \\ &\|\tilde{a}^*(\eta^{k_1 k_2 \dots k_n}) - \tilde{a}^*(\eta^{k_1 k_2 \dots k_n} + h)\| + \|b^*(\eta^{k_1 k_2 \dots k_n}) - b^*(\eta^{k_1 k_2 \dots k_n} + h)\| \le K_2\|h\| \end{aligned} \tag{25}$$

where K_1, K_2 are independent of $\eta^{k_1 k_2 \dots k_n}$. These are Lipschitz and linear growth conditions. Furthermore, $\tilde{a}^*$ and b^* are continuously differentiable in $\eta^{k_1 k_2 \dots k_n}$. Remark V.7.3 in Krylov [9] and the hypotheses of Lemma 7.2 imply that the system of equations formed by (23) and (24) is SEM. Then, as in Remark V.3.5 of [9], we can inductively add higher order derivatives to the system. These derivatives satisfy the regularity conditions (25) so the expanded system is SEM. Let η be the vector formed by *all* derivatives processes up to order $2(\xi + 1)$. It follows from Remark V.3.2 in [9] that for any positive p there exist positive constants q, M^* such that $E[\|\eta(t)\|^p] \le M^*(1 + \|\eta(0)\|^q)$. The norm of each derivative process at time 0 is bounded by one, so M^* may be chosen to satisfy $E[\|\eta^{k_1 \dots k_n}(t)\|^p] \le M^*$.

We consider now $\partial\phi^D/\partial x_{k_1}$. By Lemma V.7.1 of [9], the hypotheses of Lemma 7.2 and the fact that the derivative processes are SEM, we have that ϕ^D is continuously differentiable and we may exchange differentiation and expectation to get

$$\frac{\partial \phi^D}{\partial x_{k_1}}(t) \quad = \quad E[\nabla g \cdot \eta^{k_1}(t)] \le E[\|\nabla g\| \|\eta^{k_1}(t)\|] \le (E[\|\nabla g\|^2])^{\frac{1}{2}} (E[\|\eta^{k_1}(t)\|^2])^{\frac{1}{2}}$$

where last step is the Cauchy-Schwarz inequality. Also, $(E\|\nabla g\|^2)^{\frac{1}{2}} \le Gd^{\frac{1}{2}}$ because $\nabla g \in \Re^M$, $g \in B^\xi(G)$. We also have that $E\|\eta^{k_1}(t)\|^2$ is bounded as shown above. Therefore $|\partial\phi^D/\partial x_{k_1}(t)| \le GD$ with D independent of g.

Next we consider second order derivatives. These are

$$\frac{\partial^2 \phi^D}{\partial^2 x_{k_1} x_{k_2}} = \frac{\partial}{\partial x_{k_2}} E[\nabla g \cdot \eta^{k_1}]. \tag{26}$$

The quantity between brackets in (26) is of polynomial growth, and the derivative processes $\eta^{k_1}, \eta^{k_1 k_2}$ satisfy the hypotheses of Lemma V.7.1 in [9]. This, again, ensures continuous differentiability of $\partial\phi^D/\partial x_{k_1}$ and allows us to interchange differentiation and expectation to get

$$\begin{aligned} E[\eta^{k_2^\top} \cdot (\nabla\nabla g) \cdot \eta^{k_1} + \nabla g \cdot \eta^{k_1 k_2}] \quad &\le \quad E[\|\eta^{k_2^\top} \cdot (\nabla\nabla g) \cdot \eta^{k_1} + \nabla g \cdot \eta^{k_1 k_2}\|] \\ &\le \quad E[\|\eta^{k_2^\top} \cdot (\nabla\nabla g) \cdot \eta^{k_1}\|] + E[\|\nabla g \cdot \eta^{k_1, k_2}\|]. \end{aligned}$$

All partial derivatives of g have been assumed bounded by G and we have shown before that the norm of the derivative processes have bounded moments for finite time t. Thus, Hölder's inequality and some algebra lead to $|\partial^2\phi^D/\partial^2 x_{k_1}x_{k_2}(t)| \leq GD$ with D independent of g.

We avoid presenting here the cumbersome but straightforward computations that generalize the result to higher derivatives. The proof repeatedly uses the regularity of the derivative processes to apply Lemma V.7.1 in [9]. These computations are analogous to those made explicit above and prove that, for finite ξ, partial derivatives of ϕ^D up to order $2(\xi+1)$ exist, are continuous, and bounded in absolute value by $D(\xi, G) = D(\xi)G$. $\square$

Part 2 of proof. We continue now with the proof of Theorem 6.1. The second issue we need to address to apply the approach of Mikulevicius and Platen [14] is the $2(\xi+1)$-fold continuous differentiability of the solution v of a Kolmogorov equation and associated functionals (21).

That is, we need to show that

$$\begin{aligned} v(s,x) &= E\Big[g(X(T))|X(s)=x\Big], \\ \Pi v(s,x) &= \int_E \Big[v(s, x + c(x,z)) - v(s,x)\Big]\, h(z)\,\lambda_0\, dz \end{aligned}$$

are $2(\xi+1)$-times continuously differentiable in the initial condition x.

We begin with $v(s,x)$. Let N be the number of points in $[s,T]$ of the Poisson random measure in (11), with strictly increasing jump times $\{\tau_1,\ldots,\tau_N\}$, $N<\infty$ a.s. We take the paths of X to be right-continuous and write $X(\tau_j-)$ for $\lim_{t\to\tau_j-} X(t)$. Conditioning on the jump times we define

$$v_n(s, s_1, \ldots, s_n, x) = E\left[g(X(T))|X(s)=x, N=n, \tau_i = s_i, i=1,\ldots,n\right]$$

with $s \leq s_1 < \ldots < s_n$. We show by induction in the number of jumps that $v_n(s, s_1, \ldots, s_n, \cdot)$ is in $B^\xi(K^n D^{n+1} G)$ for all $s, s_1, \ldots, s_n$, with G and K as in Theorem 6.1. For $n=1$ we have

$$v_1(s,s_1,x) = \\ E\left[E\left[E[g(X(T))|X(s_1), N=1, \tau_1=s_1]\Big|X(s_1-), N=1, \tau_1 = s_1\right]\Big|X(s)=x, N=1, \tau_1=s_1\right].$$

The innermost expectation is computed conditional on no jumps in $(s_1, T]$; in the notation of Lemma 7.2, it is $\phi^D(s_1, T, X(s_1))$, which is in $B^\xi(D\,G)$. Thus,

$$v_1(s,s_1,x) = E\left[E\left[\phi^D(s_1,T,X(s_1))|X(s_1-), N=1, \tau_1=s_1\right]|X(s)=x, N=1, \tau_1=s_1\right].$$

Since $\phi^D(s_1, T, \cdot)$ is in $B^\xi(DG)$, by hypothesis (ii) of Theorem 6.1 the inner conditional expectation is in $B^\xi(KDG)$. The outer conditional expectation is computed conditional on no jumps in $[s, s_1)$ so again applying Lemma 7.2 we conclude that $v_1(s, s_1, \cdot) \in B^\xi(KD^2G)$.

For the inductive step define $S_n = \{s_1, \ldots, s_n\}$ and $\Theta_n = \{\tau_1, \ldots, \tau_n\}$. Take as induction hypothesis that

$$v_{n-1}(s, s_1, \ldots, s_{n-1}, x) = E\left[g^*(X(t))|X(s) = x, N = n-1, \Theta_{n-1} = S_{n-1}\right]$$

belongs to $B^\xi(K^{n-1}D^nG^*)$ for any fixed $t \le T$ and $g^* \in B^\xi(G^*)$. Now

$$\begin{aligned} & v_n(s, s_1, \ldots, s_n, x) \\ = \;& E\left[g(X(T))|X(s) = x, N = n, \Theta_n = S_n\right] \\ = \;& E\Big[E\Big[E[g(X(T))|X(s_n), N = n, \tau_n = s_n]\Big|X(s_n-), N = n, \tau_n = s_n\Big] \\ & \quad \Big|X(s) = x, N = n, \Theta_n = S_n\Big]. \end{aligned}$$

The same argument as in the case of one jump applies for the two innermost expectations, allowing us to write

$$v_n(s, s_1, \ldots, s_n, x) = E\left[\phi(s_n, X(s_n-))|X(s) = x, \Theta_{n-1} = S_{n-1}\right]$$

for some $\phi(s_n, \cdot)$ in $B^\xi(KDG)$. For the last expectation we apply the induction hypothesis with $G^* = KDG$ to conclude that $v_n(s, s_1, \ldots, s_n, x) \in B^\xi(K^nD^{n+1}G)$.

Next we integrate over the jump times and write

$$v(s, x, n) = \int \ldots \int q_n(s_1, \ldots, s_n) v_n(s, s_1, \ldots, s_n, x)\, ds_1, \ldots, ds_n,$$

where q_n is the joint density of the jump times in $[s, T]$ of the Poisson random measure, conditional on $N = n$. Because $v(s, s_1, \ldots, s_n, x) \in B^\xi(K^nD^{n+1}G)$, the Bounded Convergence Theorem allows us to interchange differentiation (in x) and integration and conclude that the derivatives of $v(s, x, n)$ up to order $2(\xi + 1)$ exist and are continuous. Furthermore, $v_n(s, s_1, \ldots, s_n, \cdot) \in B^\xi(K^nD^{n+1}G)$ implies that $v(s, x, n) \in B^\xi(K^nD^{n+1}G)$ too.

Finally we treat $v(s, x) = E[g(X(T))|X(s) = x]$. This can be written as

$$v(s, x) = \sum_{n=0}^{\infty} P(N = n) v(s, x, n), \quad \text{with} \quad P(N = n) = \frac{e^{-\lambda_0(T-s)}(\lambda_0(T-s))^n}{n!}.$$

Any series of the form $\sum_{n=0}^{\infty} P(N = n) f_n$ with $|f_n| \le C^n$ for some constant C is absolutely convergent. Therefore $v(s, x)$ is bounded. Notice that $\frac{\partial^m v(s,x,n)}{\partial x_{k_1} \ldots \partial x_{k_m}}$ is continuous and that

$$\begin{aligned} \sum_{n=0}^{\infty} P(N = n) \frac{\partial^m v(s, x, n)}{\partial x_{k_1} \ldots \partial x_{k_m}} &\le \sum_{n=0}^{\infty} |P(N = n) \frac{\partial^m v(s, x, n)}{\partial x_{k_1} \ldots \partial x_{k_m}}| \\ &\le \sum_{n=0}^{\infty} P(N = n) K^n D^{n+1} G = C < \infty. \end{aligned}$$

Then $\sum_{n=0}^{\infty} P(N=n)\frac{\partial^m v(s,x,n)}{\partial x_{k_1}\ldots\partial x_{k_m}}$ converges uniformly and is continuous. Therefore

$$\frac{\partial^m v(s,x)}{\partial x_{k_1}\ldots\partial x_{k_m}} = \sum_{n=0}^{\infty} P(N=n)\frac{\partial^m v(s,x,n)}{\partial x_{k_1}\ldots\partial x_{k_m}}$$

which implies that $\frac{\partial^m v(s,x)}{\partial x_{k_1}\ldots\partial x_{k_m}}$ exists and is continuous. Furthermore, $v(s,x) \in B^{\xi}(C)$.

To complete the proof of the theorem we need to show that

$$\Pi v(s,x) = \lambda_0 \int_E v(s, x + c(x,z))\, h(z)\, dz - \lambda_0\, v(s,x)$$

is $2(\xi+1)$-times continuously differentiable in x. We only need to consider the first term. The function $v(s,\cdot)$ is in $B^{\xi}(C)$ so by hypothesis (ii) of Theorem 6.1 the integral (viewed as a function of x) is in $B^{\xi}(KC)$ and in particular is $2(\xi+1)$-times continuously differentiable. $\square$

References

[1] Barndorff-Nielsen, O.E., Processes of Normal Inverse Gaussian Type, *Finance and Stochastics* 2, 41–68, (1998).

[2] Björk, T., Kabanov, Y., and Runggaldier, W., Bond Market Structure in the Presence of Marked Point Processes, *Mathematical Finance*, 7, 211-239, (1997).

[3] Brémaud, P., *Point Processes and Queues: Martingale Dynamics*, Springer, New York, (1981).

[4] Duffie, D., Pan, J., and Singleton, K., Transform analysis and option pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376, (2000).

[5] Glasserman, P., Kou, S.G., The Term Structure of Simple Forward Rates with Jump Risk, *Mathematical Finance*, to appear.

[6] Glasserman, P., Merener, N., Numerical Solution of Jump-Diffusion Libor Market Models, *Finance and Stochastics*, to appear.

[7] Kloeden, P. E., and Platen, E., *Numerical Solution of Stochastic Differential Equations* Springer-Verlag, Berlin, (1992)

[8] Kou, S.G., A Jump Diffusion Model for Option Pricing, *Management Science*, to appear.

[9] Krylov, N. V., *Introduction to the Theory of Diffusion Processes* Translations of Mathematical Monographs, 142. American Mathematical Society, (1995).

[10] Madan, D.B., and Seneta, E., The Variance Gamma Model for Share Market Returns, *Journal of Business* 63, 511–524 (1990).

[11] Maghsoodi, Y., Exact Solutions and Doubly Efficient Approximations of Jump-Diffusion Ito Equations, *Stochastic Analysis and Applications*, 16, 1049-1072 (1998).

[12] Maghsoodi, Y., and Harris, C.J., In-Probability Approximation and Simulation of Nonlinear Jump-Diffusion Stochastic Differential Equations, *IMA Journal of Mathematical Control and Information*, 4, 65-92 (1987).

[13] Merton, R., Option Pricing when Underlying Stock Returns are Discontinuous, *J. Financial Economics*, 3, 125-144 (1976).

[14] Mikulevicius, R., and Platen, E., Time Discrete Taylor Approximations for Ito Processes with Jump Component, *Mathematische Nachrichten*, 138, 93-104, (1988).

[15] Milstein, G.N., A Method of Second-Order Accuracy Integration of Stochastic Differential Equations, *Theory of Probability and its Applications*, 19, 557-562 (1978).

[16] Musiela, M., and Rutkowski, M., *Martingale Methods in Financial Modeling*, Springer, New York, (1997).

[17] Platen, E., and Rebolledo, R., Weak Convergence of Semimartingales and Discretisation Methods, *Stochastic Processes and Their Applications*, 20, 41-58 (1985).

[18] Protter, P., and Talay, D. The Euler Scheme for Lévy Driven Stochastic Differential Equations, *Annals of Probability*, 25, 393-423 (1997).