Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Numerical Mathematics 53 (2005) 1–18
APPLIED NUMERICAL MATHEMATICS
www.elsevier.com/locate/apnum

Numerical valuation of options with jumps in the underlying

Ariel Almendral [a,*], Cornelis W. Oosterlee [b]

[a] *Norwegian Computing Center, Gaustadalleen 23, Postbox 114, Blindern, N-0314, Oslo, Norway*
[b] *Faculty of Information Technology Systems, Department of Applied Mathematical Analysis, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands*

Abstract

A jump-diffusion model for a single-asset market is considered. Under this assumption the value of a European contingency claim satisfies a general partial integro-differential equation (PIDE). The equation is localized and discretized in space using finite differences and finite elements and in time by the second order backward differentiation formula (BDF2). The resulting system is solved by an iterative method based on a simple splitting of the matrix. Using the fast Fourier transform, the amount of work per iteration may be reduced to $O(n \log_2 n)$ and only $O(n)$ entries need to be stored for each time level. Numerical results showing the quadratic convergence of the methods are given for Merton's model and Kou's model.

Keywords: Option pricing; Jump-diffusion processes; Finite differences; Finite elements; Fast Fourier transform; Integro-differential equations

1. Introduction

Based on the model by Samuelson [29], and under general equilibrium assumptions on the market, Black and Scholes [4] derived a differential equation for option prices. Numerical inversion of the Black–Scholes equation based on data from different strikes and fixed maturity produces the so-called volatility skew or smile, contrary to the model's assumption on constant volatility. Empirical studies have revealed that the normality of the log-returns cannot capture features like heavy tails and asymmetries.

* Corresponding author.
E-mail addresses: ariel@math.uio.no (A. Almendral), c.w.oosterlee@math.tudelft.nl (C.W. Oosterlee).

0168-9274/$30.00 © 2004 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.apnum.2004.08.037

To overcome these problems, a number of alternative models have appeared in the financial literature: Stochastic volatility models [20,18]; deterministic local volatility functions [9,13]; jump-diffusion models [22,23,26]; Lévy models [2,7,14,25,28] amongst others. Jump-diffusion models and Lévy based models are attractive because they explain the jump patterns exhibited by some stocks. Some studies also reveal that Lévy models are realistic when pricing options close to maturity [10]. However, they are more difficult to handle numerically and, in contrast to the basic Black–Scholes model, it is not immediately obvious which hedging strategy leads to an instantaneous risk free portfolio.

In this paper we choose the jump-diffusion approach with constant coefficients and we find numerically the value of European Vanilla options. More precisely, we solve the PIDE for two models: the classical Merton's model [26] and Kou's model [22]. For both models analytical formulas for the solution exist, either as an infinite sum or in terms of an integral; the last one is obtained as an application of Fourier analysis [5,8,24]. We will benefit from these formulas to draw conclusions on the accuracy of our numerical schemes. This research is intended to create the foundations for the numerical solution of more complicated products like American options, path-dependent options, or a combination of different models like deterministic local volatility plus jump-diffusion [1].

There exist several papers dealing with the numerical valuation of jump-diffusion processes. In [27], the case of American options with Poisson jumps is treated numerically by a method of lines. More general models based on Lévy processes are also solved numerically in [1] by the ADI finite difference method combined with the fast Fourier transform and in [25] by a finite element method that gives a compressed sparse matrix in a convenient wavelet basis. An explicit method was used in [6] to solve Merton's model and a convergence theory for explicit schemes and CFL conditions were given for a general family of integro-differential Cauchy problems. Recently, we came across [12], where the value of American options using Merton's model is found implicitly by the penalty method. Here, we intend to simplify some ideas from [25] for the jump-diffusion case, while keeping the algorithm fast.

The paper is organized as follows. Section 2 is a short introduction into Lévy driven assets. In Section 3 we show how, knowing the Lévy triplet, we may write down the equation governing option prices. Section 4 deals with the discretization by finite differences and BDF2 and Section 5 with the discretization by FEM and BDF2. In Section 6 we explain the solver based on a splitting of the resulting dense matrix and how the fast Fourier transform helps to speed up the method. In the last Section 7 we carry out some numerical tests to verify the second order accuracy of the numerical schemes.

2. The market modeled by Lévy processes

A one-dimensional stochastic process $\{L_t\}_{t\geqslant 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a Lévy process if:

(1) $L_0 = 0$ almost surely (a.s.);
(2) it has independent increments, that is, if for any $0 \leqslant s < t < s' < t'$ the random variables $L_t - L_s$ and $L_{t'} - L_{s'}$ are independent;
(3) it has stationary increments, that is, for any $0 \leqslant s < t < \infty$, the law of $L_t - L_s$ only depends on the increments $t - s$;
(4) it is stochastically continuous, that is, for any $\varepsilon > 0$, $P[|L_{s+t} - L_s| > \varepsilon] \to 0$ as $t \to 0$;
(5) the sample paths are right continuous with left limits a.s.

Examples of Lévy processes are the Brownian motion, the Poisson processes and also its extension to a Compound Poisson Process. These are so-called finite activity processes, where the paths consist of a continuous Brownian component and a jump component; the jumps occurring a finite number of times on each finite interval. However, there exist processes whose paths jump infinitely many times on each finite interval. Here, we focus only on the first family of processes.

Once we have defined the Lévy process, we may model the asset value by the following Geometric Lévy process

$$S_t = S_0 e^{L_t},$$

on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where $\{\mathcal{F}_t\}$ is the filtration generated by the Lévy process $\{L_t\}_{t\geqslant 0}$.

2.1. Lévy–Khintchine representation

The large family of Lévy processes is characterized by the following fundamental result [30]:

Theorem 2.1 (Lévy–Khintchine representation). *For all $z \in \mathbb{R}$ and $t \geqslant 0$,*

$$\mathbf{E}\big(e^{izL_t}\big) = \exp\left[t\left(-\frac{a}{2}z^2 + i\gamma z + \int_{\mathbb{R}} \big(e^{izx} - 1 - izx\mathbb{1}_{\{|x|\leqslant 1\}}\big)\,\mathrm{d}\nu(x)\right)\right], \tag{1}$$

where a is a non-negative real number, γ is real and ν is a measure on $\mathbb{R}$ satisfying $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}} \min(1, x^2)\,\mathrm{d}\nu(x) < \infty$.

The notation $\mathbb{1}_{\Omega}$ stands for the indicator function of the set Ω. The set of three parameters (a, γ, ν) is commonly known as the *generating Lévy triplet*. The first parameter a is called *Gaussian variance*, since it is associated with the Brownian part of the Lévy process, the third quantity ν is called *Lévy measure*. If $\nu = 0$ then L_t is a drifted Brownian motion. If $a = 0$ then L_t is said to be purely non-Gaussian. For further details on Lévy processes see [3,30].

In this paper, we will concentrate on the family of processes for which ν satisfies

$$\int_{\mathbb{R}} |x|\,\mathrm{d}\nu(x) < \infty. \tag{2}$$

In that case, (1) may be written as

$$\mathbf{E}\big[e^{izL_t}\big] = \exp\left[t\left(-\frac{a}{2}z^2 + ibz + \int_{\mathbb{R}} \big(e^{izx} - 1 - izx\big)\,\mathrm{d}\nu(x)\right)\right]. \tag{3}$$

We will refer to (a, b, ν) as the *reduced Lévy triplet* and to the function in (3)

$$\psi(z) = -\frac{a}{2}z^2 + ibz + \int_{\mathbb{R}} \big(e^{izx} - 1 - izx\big)\,\mathrm{d}\nu(x), \tag{4}$$

as the *reduced Lévy–Khintchine exponent*. This is a convenient reformulation because it gives a direct link between the coefficients in the PIDE satisfied by the option's price and the reduced triplet; see Theorem 3.1.

Example 1. Drifted Brownian motion (classical Samuelson model [29])

$$L_t = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t,$$

where W_t is a Brownian motion in $(\Omega, \mathcal{F}, \mathbb{Q})$. It is well known [21] that in this case there is only one martingale measure, and since the process $\{e^{-\frac{\sigma^2}{2}t+\sigma W_t}\}_{t\geqslant 0}$ is a $\mathbb{Q}$-martingale we have $\mu = r$. The reduced Lévy–Khintchine exponent is

$$\psi(z) = -\frac{\sigma^2}{2}z^2 + \left(r - \frac{\sigma^2}{2}\right)iz,$$

so that the Lévy (and reduced) triplet results in $(\sigma^2, r - \sigma^2/2, 0)$.

Example 2. Drifted Brownian motion with finite number of jumps

$$L_t = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t + \sum_{i=1}^{N_t} Y_i,$$

where $\{Y_i\}$ are independent, identically distributed random variables with common law F; the Poisson process $\{N_t\}$ has intensity λ, and $\{N_t\}$, $\{Y_i\}$ and $\{W_t\}$ are mutually independent.

The "EMM condition" [5] implies the restriction on the asset mean return: $\mu = r - \lambda\zeta$, where ζ is the following constant:

$$\zeta = \int_{\mathbb{R}} \left(e^x - 1\right) \mathrm{d}F(x). \tag{5}$$

Letting $\tilde{\zeta} = \zeta - \mathbf{E}[Y]$, the reduced Lévy triplet results in

$$\left(\sigma^2, r - \lambda\tilde{\zeta} - \sigma^2/2, \lambda F\right).$$

Special cases:

- Classical Merton's model [26]: Y_i are normally distributed, with mean μ_J and standard deviation σ_J. That is, $\mathrm{d}F(x) = f_m(x)\,\mathrm{d}x$, where

$$f_m(x) := \frac{1}{\sqrt{2\pi}\sigma_J} e^{-(x-\mu_J)^2/2\sigma_J^2}. \tag{6}$$

  We may compute then (cf. (5)):

$$\zeta = e^{\mu_J + \sigma_J^2/2} - 1. \tag{7}$$

- Non-symmetric double exponential (Kou's model [22,23])

$$f_{kw}(x) = p\alpha_1 e^{-\alpha_1 x}\mathbb{1}_{\{x\geqslant 0\}} + q\alpha_2 e^{\alpha_2 x}\mathbb{1}_{\{x<0\}}, \tag{8}$$

  where p, q are positive real numbers such that $p + q = 1$. In order to integrate e^x over the real line we must have $\alpha_1 > 1$ and $\alpha_2 > 0$. We then obtain the expression

$$\zeta = \frac{p\alpha_1}{\alpha_1 - 1} + \frac{q\alpha_2}{\alpha_2 + 1} - 1. \tag{9}$$

2.2. Option pricing via equivalent martingale measures

Lévy markets are in general incomplete, i.e., not every contingent claim can be replicated, or in other words, a perfect hedging strategy is impossible. Therefore, in contrast to the classical Black and Scholes model, option prices cannot be obtained by replication. However, a price can still be found based on the no-arbitrage assumption.

An Equivalent Martingale Measure (EMM) $\mathbb{Q}$ is a measure on $\mathcal{F}$ with the same null sets as $\mathbb{P}$, such that the discounted price process $M_t = e^{-rt}S_t$ becomes a $\mathbb{Q}$-martingale.[1] It is known from [11] that the existence of such a measure is in some sense equivalent to the no-arbitrage assumption. For earlier related works see [16,17]. Hence, one of the main problems in option pricing is that of finding a reasonable EMM. Several ideas to find an EMM have been suggested in the financial literature. We briefly outline one method to find an EMM that has been known for long in actuarial sciences, based on the Esscher transform [15,28].

In this paper we deal with the following type of Lévy process (compare with Example 2):

$$\widetilde{L}_t = \left(\mu - r - \frac{\sigma^2}{2}\right)t + \sigma W_t + \sum_{i=1}^{N_t} Y_i. \tag{10}$$

The new measure $\mathbb{Q}$ is defined by the relation

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{\exp(\theta \widetilde{L}_t)}{\mathrm{mgf}(\theta)^t}, \tag{11}$$

where $\mathrm{mgf}(\cdot)$ denotes the moment generating function of $\widetilde{L}_1$

$$\mathrm{mgf}(u) := \mathbf{E}\big[\exp\big(u\widetilde{L}_1\big)\big], \tag{12}$$

and θ is a constant satisfying the equality:

$$\mathrm{mgf}(\theta + 1) = \mathrm{mgf}(\theta). \tag{13}$$

It is proven in [28] that, if Eq. (13) has a solution, then $\mathbb{Q}$ defines an EMM and $\widetilde{L}_t$ becomes a Lévy process under $\mathbb{Q}$.

3. The partial integro-differential equation for option prices and the reduced Lévy triplet

In this section we explain the connection between the martingale approach to option pricing and the PDE approach. From now on we assume that we have already chosen one possible EMM.

Let $\{w(t, S_t)\}_{t\geqslant 0}$ be the value process of a European contingency claim on the asset S_t and let $g(s) := w(T, s)$. For example, for a European call we have $g(s) = (s - K)^+$, where K is the strike price. If we assume that the discounted value process $\{e^{-rt}w(t, S_t)\}_{t\geqslant 0}$ is a martingale, then we have $e^{-rt}w(t, S_t) = \mathbf{E}_{\mathbb{Q}}[e^{-rT}w(T, S_T)|\mathcal{F}_t]$. This yields the well-known formula for the claim,

$$w(t, s) = e^{-r(T-t)}\mathbf{E}_{\mathbb{Q}}\big[g\big(se^{L_{T-t}}\big)\big]. \tag{14}$$

[1] A martingale under $\mathbb{Q}$ is a process satisfying $\mathbf{E}_{\mathbb{Q}}[M_t|\mathcal{F}_s] = M_s$, where $\mathbf{E}_{\mathbb{Q}}[\,\cdot\,|\mathcal{F}]$ is the conditional expectation operator.

In order to have a full characterization of the equation satisfied by w, it is convenient to change variables $x = \ln s$ and consider the process $v(t, X_t) := w(t, e^{X_t})$. This leads to

$$v(t,x) = e^{-r(T-t)}\mathbf{E}_{\mathbb{Q}}\big[g\big(e^{x+L_{T-t}}\big)\big]. \tag{15}$$

Then, the following theorem exists (see, e.g., [28]):

Theorem 3.1. *Let* $v(t,x) \in C^{1,2}([0,T)\times\mathbb{R}) \cap C^0([0,T]\times\mathbb{R})$ *and assume* (2). *Then* v *satisfies the following partial integro-differential equation* (*PIDE*)

$$\begin{aligned} &v_t + \frac{a}{2}v_{xx} + bv_x - rv + \int_{\mathbb{R}} \big(v(t,x+y) - v(t,x) - v_x(t,x)y\big)\,\mathrm{d}\nu(y) = 0, \\ &\quad \forall (t,x) \in [0,T)\times\mathbb{R}, \end{aligned} \tag{16}$$

with final condition

$$v(T,x) = \tilde{g}(x) := g\big(e^x\big), \quad \forall x \in \mathbb{R}. \tag{17}$$

The parameters involved are the risk-free interest rate r *and the reduced Lévy triplet* (a,b,ν) *under the risk-neutral measure* $\mathbb{Q}$.

For Example 2 discussed in Section 2, the PIDE reads

$$\begin{cases} v_t + \frac{1}{2}\sigma^2 v_{xx} + \big(r - \frac{1}{2}\sigma^2 - \lambda\zeta\big)v_x \\ \qquad -(r+\lambda)v + \lambda\int_{\mathbb{R}} v(t,x+y)\,\mathrm{d}F(y) = 0, \quad \forall (t,x) \in [0,T)\times\mathbb{R}, \\ v(T,x) = g(e^x), \quad \forall x \in \mathbb{R}, \end{cases} \tag{18}$$

where $\mathrm{d}F(x) = f(x)\,\mathrm{d}x$. In this paper we focus on f given by (6) and (8).

In terms of the prices s, and the function $w(t,s)$, the above problem transforms into

$$\begin{cases} w_t + \frac{1}{2}\sigma^2 s^2 w_{ss} + (r - \lambda\zeta)s w_s \\ \qquad - (r+\lambda)w + \lambda\int_0^\infty w(t,sz)\,\mathrm{d}\widetilde{F}(z) = 0, \quad \forall (t,s) \in [0,T)\times\mathbb{R}_+, \\ w(T,s) = g(s), \quad \forall s \in \mathbb{R}_+, \end{cases} \tag{19}$$

with $\mathrm{d}\widetilde{F}(z) := \mathrm{d}F(\ln(z)) = \tilde{f}(z)\,\mathrm{d}z$. This is the form given by Merton [26]. He studied this problem with $\tilde{f}$ given by the log-normal density

$$\tilde{f}_m(z) := \frac{1}{\sqrt{2\pi}\sigma_J z} e^{-(\ln z - \mu_J)^2/2\sigma_J^2} \mathbb{1}_{\{z>0\}}. \tag{20}$$

Analogously, for Kou's model [22], $\tilde{f}$ is the log-double-exponential density

$$\tilde{f}_{kw}(z) := p\alpha_1 z^{-\alpha_1 - 1}\mathbb{1}_{\{z\geqslant 1\}} + q\alpha_2 z^{\alpha_2 - 1}\mathbb{1}_{\{0<z<1\}}. \tag{21}$$

Remark 3.1. Only for a few particular cases, an analytical expression has been found for the solution of (19). E.g., for Merton's model and Kou's model, the solution is given in the form of an infinity series [22, 26].

4. Discretization with finite differences

In this section we concentrate on a straightforward numerical scheme for problem (18) with the normal density (6) and the exponential density (8). In the next section, a somewhat more elegant approach is given; a comparison of numerical results for both schemes is presented in Section 7.

The time variable is transformed to obtain a forward problem in time. If we set $u(\tau,\cdot) := v(T-\tau,\cdot)$ and $\tilde{g}(x) := g(e^x)$, (18) becomes

$$\begin{cases} u_\tau - \frac{1}{2}\sigma^2 u_{xx} - \left(r - \frac{1}{2}\sigma^2 - \lambda\zeta\right)u_x \\ \qquad + (r+\lambda)u - \lambda\int_{\mathbb{R}} u(t,x+y)\,\mathrm{d}F(y) = 0, \quad \forall(\tau,x)\in[0,T)\times\mathbb{R}, \\ u(0,x) = \tilde{g}(x), \quad \forall x\in\mathbb{R}. \end{cases} \tag{22}$$

To discretize the integral term, we change variables

$$\int_{\mathbb{R}} u(\tau,x+y)f(y)\,\mathrm{d}y = \int_{\mathbb{R}} u(\tau,z)f(z-x)\,\mathrm{d}z.$$

Next, we split the integral on the right-hand side as $\int_{\mathbb{R}} = \int_{\Omega_*} + \int_{\mathbb{R}\setminus\Omega_*}$, where $\Omega_* := (-x^*, x^*)$. In case we are computing a European call option, the integrand $u(\tau,z)$ over $\mathbb{R}\setminus\Omega_*$ must be replaced by the approximations:

$$u(\tau,x) \to e^x - Ke^{-r\tau}, \quad \text{as } x\to+\infty, \tag{23}$$

$$u(\tau,x) \to 0, \quad \text{as } x\to-\infty. \tag{24}$$

This motivates the introduction of the function

$$\varepsilon\left(\tau,x,x^*\right) := \int_{\Omega_*}\left(e^z - Ke^{-r\tau}\right)f(z-x)\,\mathrm{d}z. \tag{25}$$

After some computations we find for Merton's model (6), with $\mu_J = 0$, the following expression:

$$\varepsilon\left(\tau,x,x^*\right) = e^{x+\frac{\sigma_J^2}{2}}\Phi\left(\frac{x-x^*+\sigma_J^2}{\sigma_J}\right) - Ke^{-r\tau}\Phi\left(\frac{x-x^*}{\sigma_J}\right), \tag{26}$$

where $\Phi(y)$ is the cumulative normal distribution:

$$\Phi(y) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{x^2}{2}}\,\mathrm{d}x. \tag{27}$$

Consider a uniform mesh in space and in time, that is, let $x_i = -x^* + (i-1)h$ $(i=1,\ldots,n)$, and $\tau_m = (m-1)k$ $(m=1,\ldots,q)$. Let $u_i^m \approx u(\tau_m,x_i)$ and $f_{ij} := f(x_j - x_i)$. By the composite trapezoidal rule on $[-x^*,x^*]$, we have the following approximation of the integral:

$$\int_{\mathbb{R}} u(\tau_m,z)f(z-x_i)\,\mathrm{d}z$$

$$\approx \frac{h}{2}\left[u_1^m f_{i,1} + u_n^m f_{i,n} + 2\sum_{j=2}^{n-1} u_j^m f_{i,j}\right] + \varepsilon\left(\tau_m,x_i,x^*\right), \quad i=2,\ldots,n-1. \tag{28}$$

For the time variable and space variable we may write the following approximations:

$$u_\tau(\tau_m, x_i) \approx \begin{cases} \left(\frac{3}{2}u_i^m - 2u_i^{m-1} + \frac{1}{2}u_i^{m-2}\right)/k, & \text{if } m \geqslant 2, \\ \left(u_i^m - u_i^{m-1}\right)/k, & \text{if } m = 1, \end{cases} \tag{29}$$

$$u_{xx}(\tau_m, x_i) \approx \left(u_{i+1}^m - 2u_i^m + u_{i-1}^m\right)/h^2, \tag{30}$$

$$u_x(\tau_m, x_i) \approx \left(u_{i+1}^m - u_{i-1}^m\right)/2h. \tag{31}$$

That is, the time derivative is discretized by BDF2 for $m \geqslant 2$ and BDF1 for $m = 1$, whereas the first spatial derivative is approximated by the $O(h^2)$ central scheme and the second spatial derivative is estimated by the standard, $O(h^2)$-accurate 3-point difference scheme.

Define the vector $\mathbf{u}^m := (u_1^m, \ldots, u_n^m)^{\mathrm{T}}$. From the initial condition, the initial vector is $\mathbf{u}^1 := (\tilde{g}(x_1), \ldots, \tilde{g}(x_n))^{\mathrm{T}}$. With this notation and using the difference schemes (28)–(31), the finite difference discretization of (22) may be written in matrix form as

$$(\omega_0 \mathbf{I} + \mathbf{C} + \mathbf{D})\mathbf{u}^m = \mathbf{b}^m, \tag{32}$$

where

$$\omega_0 = \begin{cases} 1 & \text{if } m = 1, \\ 3/2 & \text{if } m \geqslant 2, \end{cases} \tag{33}$$

$\mathbf{I}$ is the identity matrix and the matrices $\mathbf{C} := [c_{ij}]_{i,j=1}^n$ and $\mathbf{D} := [d_{ij}]_{i,j=1}^n$ are given by

$$c_{ij} = \begin{cases} -k\sigma^2/2h^2 + k\left(r - \sigma^2/2 - \lambda\zeta\right)/2h & \text{if } i = j-1,\ 2 \leqslant i \leqslant n-1, \\ k\sigma^2/h^2 + (r+\lambda)k & \text{if } i = j,\ 2 \leqslant i \leqslant n-1, \\ -k\sigma^2/2h^2 - k\left(r - \sigma^2/2 - \lambda\zeta\right)/2h & \text{if } i = j+1,\ 2 \leqslant i \leqslant n-1, \\ 0 & \text{otherwise;} \end{cases}$$

$$d_{ij} = \begin{cases} -kh\lambda f_{ij}/2 & \text{if } 2 \leqslant i \leqslant n-1 \text{ and } j = 1, n, \\ -kh\lambda f_{ij} & \text{if } 2 \leqslant i \leqslant n-1 \text{ and } 2 \leqslant j \leqslant n-1, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, the right side $\mathbf{b}^m := (b_1, b_2, \ldots, b_{n-1}, b_n)^{\mathrm{T}}$ is given component-wise by

$$b_i = k\lambda\varepsilon\left(\tau_m, x_i, x^*\right) + \omega_1 u_i^{m-1} + \omega_2 u_i^{m-2}, \quad \text{for } i = 2, \ldots n-1, \tag{34}$$

where

$$\omega_1 = \begin{cases} 1 & \text{if } m = 1, \\ 2 & \text{if } m \geqslant 2, \end{cases} \tag{35}$$

$$\omega_2 = \begin{cases} 0 & \text{if } m = 1, \\ -1/2 & \text{if } m \geqslant 2, \end{cases} \tag{36}$$

and from the boundary conditions

$$b_1 = 0, \qquad b_n = \omega_0\left(e^{x^*} - Ke^{-r\tau_m}\right).$$

5. Variational formulation and discretization

In this section we borrow some of the ideas and notation from [25]. The method in [25] consists of the following three steps: Firstly, a problem similar to (18) is transformed into another with a homogeneous

initial condition. Secondly, the resulting equation is localized on a finite interval $\Omega_* := (-x^*, x^*)$ and zero boundary conditions at both boundary points are imposed. Finally, a Galerkin discretization in space together with a θ-method in time are proposed.

Our description will differ from the above in that we choose a two-step backward differentiation formula (BDF2), rather than a θ-method. BDF2 is a second order, A-stable scheme with a smoothing effect for the error.

Problem (22) may be written as

$$\begin{cases} u_\tau + \mathcal{L}u = 0, & \forall(\tau, x) \in (0, T] \times \mathbb{R}, \\ u(0, x) = \tilde{g}(x), & \forall x \in \mathbb{R}, \end{cases} \tag{37}$$

where $\mathcal{L} := \mathcal{D} + \mathcal{J}$, with

$$(\mathcal{D}\varphi)(x) := -\frac{1}{2}\sigma^2\varphi_{xx}(x) - \left(r - \frac{1}{2}\sigma^2 - \lambda\zeta\right)\varphi_x(x) + r\varphi(x),$$

$$(\mathcal{J}\varphi)(x) := -\lambda \int_{\mathbb{R}} \big[\varphi(\tau, x + y) - \varphi(x)\big] f(y)\,\mathrm{d}y.$$

Remark 5.1. As pointed out in [25], we may assume $r = 0$. The reason is the following: if u is a solution of (37) with $r = 0$, then $\hat{u}(\tau, x) = e^{-r\tau}u(\tau, x + r\tau)$ fulfills the same equation and boundary condition with $r \neq 0$. We may see this by looking at the asymptotic behavior of $\hat{u}$, that is, $\hat{u}(\tau, x) \to e^{-r\tau}(e^{x+r\tau} - K) = e^x - Ke^{-r\tau}$ as $x \to +\infty$.

Once we have simplified our problem to the case $r = 0$, the solution for each $\tau > 0$ asymptotically tends towards the payoff $\tilde{g}$. Hence, as in [25], the change of variables $\bar{u} = u - \tilde{g}$ reduces (37) into the following:

$$\begin{cases} \bar{u}_\tau + \mathcal{L}\bar{u} = -\mathcal{L}\tilde{g}, & \forall(\tau, x) \in (0, T] \times \mathbb{R}, \\ \bar{u}(0, x) = 0, & \forall x \in \mathbb{R}, \\ \bar{u}(\tau, x) \to 0, & \text{as } x \to \pm\infty, \text{ for all } \tau \in (0, T]. \end{cases} \tag{38}$$

In Merton's model [26] we find that the right side $\bar{g} := -\mathcal{L}\tilde{g}$ is given by

$$\begin{aligned} \bar{g}(x) = {} & K\frac{\sigma^2}{2}\delta_{\{\ln K\}} - \big[\lambda(\zeta + 1)e^x - K\lambda\big]\mathbb{1}_{\{x \geqslant \ln K\}} \\ & + \lambda\Phi\left(\frac{x + \sigma_J^2 - \ln K}{\sigma_J}\right)e^{x+\sigma_J^2/2} - \lambda K\Phi\left(\frac{x - \ln K}{\sigma_J}\right), \end{aligned} \tag{39}$$

with Φ as in (27) and $\delta_{\{c\}}$ is the Dirac's delta concentrated at the point c.

For Kou's model we have the expression

$$\bar{g}(x) = K\frac{\sigma^2}{2}\delta_{\{\ln K\}} + \frac{K\lambda p}{\alpha_1 - 1}\exp\big[\alpha_1(x - \ln K)\big]\mathbb{1}_{\{x \leqslant \ln K\}} + \frac{K\lambda q}{\alpha_2 + 1}\exp\big[\alpha_2(\ln K - x)\big]\mathbb{1}_{\{x > \ln K\}}.$$

The bilinear forms D and J, from $H^1(\mathbb{R}) \times H^1(\mathbb{R})$ into $\mathbb{R}$, associated with $\mathcal{D}$ and $\mathcal{J}$, respectively, are given by

$$D(\varphi, \chi) = \frac{\sigma^2}{2}\int_{\mathbb{R}} \varphi'(x)\chi'(x)\,\mathrm{d}x + \left(\lambda\zeta + \frac{\sigma^2}{2}\right)\int_{\mathbb{R}} \varphi'(x)\chi(x)\,\mathrm{d}x, \tag{40}$$

$$J(\varphi,\chi) = -\lambda \int_{\mathbb{R}}\int_{\mathbb{R}} \big[\varphi(x+y)-\varphi(x)\big]\chi(x) f(y)\,\mathrm{d}x\,\mathrm{d}y. \tag{41}$$

Let k denote the time step, $\tau_m := (m-1)k$, for $m = 1,\ldots,q$ and set $\bar{u}^m(x) \approx \bar{u}(\tau_m, x)$. The backward differentiation formula may be written as:

$$\bar{u}_\tau(\tau_m, x) \approx k^{-1}\big[\omega_0\bar{u}^m - \omega_1\bar{u}^{m-1} - \omega_2\bar{u}^{m-2}\big], \tag{42}$$

with ω_i $(i = 0,1,2)$ as in (33), (35) and (36). Now, setting $L := D + J$, we may write the following sequence of spatial problems associated to (38):

$$\omega_0\big(\bar{u}^m,\chi\big) + kL\big(\bar{u}^m,\chi\big) = \big\langle \rho^m,\chi\big\rangle \quad \text{for all } \chi \in H^1(\mathbb{R}), \tag{43}$$

where

$$\rho^m := k\bar{g} + \omega_1\bar{u}^{m-1} + \omega_2\bar{u}^{m-2}, \tag{44}$$

the inner product on the right side is in the sense $\langle \Lambda,\chi\rangle := \Lambda(\chi)$ and $(\cdot,\cdot)$ denotes the inner product in $L^2(\mathbb{R})$.

In practice, we restrict the bilinear form $L(\cdot,\cdot)$ to functions in $H_0^1(\Omega_*)$, where $\Omega_* = (-x^*, x^*)$. By doing this, we get an approximate problem which may be solved by Galerkin's method. Let the interval $[-x^*, x^*]$ be divided into $n-1$ equal parts, $-x^* = x_1 < x_2 < \cdots < x_{n-1} < x_n = x^*$ and let $S_h \subset H_0^1(\Omega_*)$ be the space of piece-wise linear functions with basis $\{\varphi_2,\ldots,\varphi_{n-1}\}$, where $\varphi_i(x_j)$ are piece-wise linear functions that are equal to 1 if $i = j$, and 0 otherwise.

The sequence of truncated spatial problems reads

$$\omega_0\big(\bar{u}_h^m,\chi_h\big) + kL\big(\bar{u}_h^m,\chi_h\big) = \big\langle \rho^m,\chi_h\big\rangle \quad \text{for all } \chi_h \in S_h. \tag{45}$$

The functions $\bar{u}_h^m$ are to be found for all m. This formulation translates into the systems

$$(\omega_0\mathbf{M} + k\mathbf{L})\bar{\mathbf{u}}^m = \mathbf{b}^m, \quad m = 1,\ldots,q. \tag{46}$$

Here, $\mathbf{M} := [(\varphi_i,\varphi_j)]_{i,j=2}^{n-1}$ is the mass matrix, $\mathbf{L} := [L(\varphi_i,\varphi_j)]_{i,j=2}^{n-1}$ is the stiffness matrix, the unknown $\bar{\mathbf{u}}^m$ is the vector of coefficients of $\bar{u}_h^m$ in the basis $\{\varphi_i\}_{i=2}^{n-1}$, and the right side vector is $\mathbf{b}^m := (b_2,\ldots,b_{n-1})^{\mathrm{T}}$, where $b_i := \langle \rho^m,\varphi_i\rangle$.

In practice, we compute $D(\varphi_i,\varphi_j)$ and the mass matrix by the Gauss–Legendre quadrature rule with weights $(1,1)$ at the points $\pm 1/\sqrt{3}$. This rule is exact for polynomials of degree at most 3. However, for the integral part, it suffices to evaluate at the central point to keep second order accuracy in space.

6. Iterative methods based on regular splittings

The purpose in this section is to present an iterative method to solve the systems (32) and (46). We give conditions for which the iterative method for the finite element system (46) converges. To achieve this, it is sufficient to apply the classical theory of convergence for regular splittings, as explained in [32]. Observe that in general, $\mathbf{A}$ is a dense matrix, with the property that its entries tend to zero as we move away from the diagonal.

A representation of the matrix $\mathbf{A}$ in the form

$$\mathbf{A} = \mathbf{Q} - \mathbf{R}, \tag{47}$$

is said to be a regular splitting if $\mathbf{Q}$ is a monotone matrix ($\mathbf{Q}^{-1} \geqslant 0$) and $\mathbf{R} \geqslant 0$.

To every splitting (47), there exists an associated iterative method:

$$\mathbf{v}^{l+1} = \mathbf{Q}^{-1}\mathbf{R}\mathbf{v}^{l} + \mathbf{Q}^{-1}\mathbf{b}, \quad l = 0, \ldots, \qquad \mathbf{v}^{0} = \mathbf{0}. \tag{48}$$

Iteration (48) converges ($\rho(\mathbf{Q}^{-1}\mathbf{R}) < 1$) iff $\mathbf{A}$ is monotone [32].

Denote by $\mathbf{A}^*$ the adjoint matrix of $\mathbf{A}$. A positive stable matrix is by a definition a matrix whose eigenvalues have positive real part. We need the following two statements:

(1) A sufficient condition for a matrix $\mathbf{A}$ to be positive stable is that $\mathbf{A} + \mathbf{A}^*$ is positive definite [19].
(2) A positive stable matrix with non-positive off-diagonal entries is a monotone matrix [19].

Let $\mathbf{A} = \omega_0\mathbf{M} + k\mathbf{L}$, from the finite element discretization (46). We claim that $\mathbf{A}$ is positive stable. Since the mass matrix $\mathbf{M}$ is S.P.D., it is sufficient to verify that $\mathbf{L}$ is positive stable.

Denote by $\mathbf{v}_\varphi$ the vector of coefficients of $\varphi \in H_0^1(\Omega_*)$ in the basis $\{\varphi_i\}_{i=2}^{n-1}$. Let B be some bilinear form and define the adjoint bilinear form B^* as $B^*(\varphi, \chi) := B(\chi, \varphi)$. Then we have

$$\mathbf{v}_\varphi^{\mathrm{T}}\big(\mathbf{B} + \mathbf{B}^*\big)\mathbf{v}_\chi = \big(B + B^*\big)(\varphi, \chi), \tag{49}$$

where $\mathbf{B} := [B(\varphi_i, \varphi_j)]_{i,j=2}^{n-1}$ and $\mathbf{B}^*$ is the adjoint matrix of $\mathbf{B}$.

Repeating the argument in [25], the following identity holds

$$(J + J^*)(\varphi, \chi) = \lambda \int_{\mathbb{R}} \int_{\mathbb{R}} \big[\varphi(x+y) - \varphi(x)\big]\big[\chi(x+y) - \chi(x)\big] f(y)\,\mathrm{d}y, \quad \text{for all } \varphi, \chi \in H_0^1(\Omega_*).$$

It follows that

$$\mathbf{v}_\varphi^{\mathrm{T}}\big(\mathbf{J} + \mathbf{J}^*\big)\mathbf{v}_\varphi \geqslant 0. \tag{50}$$

On the other hand, it is straightforward to check that

$$\mathbf{v}_\varphi^{\mathrm{T}}\big(\mathbf{D} + \mathbf{D}^*\big)\mathbf{v}_\varphi = \big(D + D^*\big)(\varphi, \varphi) \geqslant 0. \tag{51}$$

Inequalities (50) and (51) give the positive stability of $\mathbf{L}$.

We now propose two splittings:

(1) Jacobi: $\mathbf{A} = \mathbf{Q}_1 - \mathbf{R}_1$, where $\mathbf{Q}_1$ is the diagonal of $\mathbf{A}$.
(2) Tridiagonal: $\mathbf{A} = \mathbf{Q}_2 - \mathbf{R}_2$, where $\mathbf{Q}_2$ is made by extracting the main three diagonals of $\mathbf{A}$.

We have the following result:

Proposition 6.1. *If the following conditions are satisfied*:

(i) $-(k/h)\sigma^2/2 - k(\lambda\zeta + \sigma^2/2)/2 + h(\omega_0 + \lambda k)/6 \leqslant 0$,
(ii) $-(k/h)\sigma^2/2 + k(\lambda\zeta + \sigma^2/2)/2 + h(\omega_0 + \lambda k)/6 \leqslant 0$,

then splittings 1 *and* 2 *are regular. Moreover*

$$\rho\big(\mathbf{Q}_2^{-1}\mathbf{R}_2\big) \leqslant \rho\big(\mathbf{Q}_1^{-1}\mathbf{R}_1\big) < 1. \tag{52}$$

Proof. By conditions (i) and (ii) the off-diagonal elements of $\mathbf{A}$ are non-positive. Observe that the entries corresponding to the integral term in (22) are non-positive, so they are not taken into account in (i)–(ii). $\mathbf{A}$ is therefore an M-matrix (a monotone matrix with non-positive off-diagonal entries; see statement 2. above), which in particular implies that the diagonal entries of $\mathbf{A}$ are positive. If $\mathbf{A}$ is an M-matrix, then any splitting $\mathbf{A} = \mathbf{Q} - \mathbf{R}$, where $\mathbf{Q}$ is formed from $\mathbf{A}$ by replacing some of the off-diagonal elements of $\mathbf{A}$ by zero, is a regular splitting; see [32]. Inequality (52) follows from the same theorem since $\mathbf{R}_2 \leqslant \mathbf{R}_1$. □

Remark 6.1. In the finite difference situation, we indicate that from a practical point of view, the following two conditions:

(iii) $-(k/h)\sigma^2/2 - k(\lambda\zeta + \sigma^2/2)/2 \leqslant 0$,
(iv) $-(k/h)\sigma^2/2 + k(\lambda\zeta + \sigma^2/2)/2 \leqslant 0$,

are sufficient for an accurate stable solution, since in both the FEM and FD approach we compute an approximation of the integral term. Note that conditions (i)–(iv) are due to the choice of a central scheme for the convection term.

Remark 6.2. If we keep the quotient k/h fixed and let $h \to 0$, there exists a $h_0 > 0$ such that conditions (i)–(iv) are fulfilled for $h \leqslant h_0$. In practice we observe that iteration (48) converges even if the M-matrix property is violated, that is, the splitting need not be regular. Note that this assertion is related to the iterative method and not to the stability of the solution. Note also that conditions arising from explicit methods are in general worse, since they demand $k = \mathrm{O}(h^2)$; see [6].

Remark 6.3.

- The three diagonals in $\mathbf{Q}_2$ include some coefficients from the discretization of the integral.
- Tridiagonal splittings work extremely well for exponentially decaying kernels.
- Matrix $\mathbf{A}$ is Toeplitz. Hence, in principle, a fast Toeplitz solver (based, for example, on FFT) is also applicable.
- We aim already at varying σ: This affects only the main three diagonals of $\mathbf{A}$ and therefore the idea of using a tridiagonal splitting is to be preferred here.

We have partially solved the storage problem, since any of the two suggested splittings need only store a few vectors. However, we need to perform a multiplication of the dense matrix $\mathbf{R}_2$ by some vector, for each step of the iterative method (48). The fast Fourier transform (FFT) may be used in a fast algorithm to compute this matrix–vector product (for some particular type of matrices), without needing to store $\mathbf{R}_2$.

6.1. Connection with Toeplitz matrices

We verify now that a simple approximation of $\mathbf{R}_2 := [r_{ij}]_{i,j=2}^{n-1}$ becomes a Toeplitz matrix, i.e., a matrix that is constant along its diagonals. Recall that this matrix results from the discretization of the integral term in (22). Note that

$$r_{ij} = \begin{cases} -k\lambda \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi_j(x+y)\varphi_i(x) f(y)\,\mathrm{d}x\,\mathrm{d}y & \text{for } |i-j| > 1, \\ 0 & \text{otherwise.} \end{cases} \tag{53}$$

But

$$\eta(x) := \int_{\mathbb{R}} \varphi_j(x+y) f(y)\,\mathrm{d}y = \int_{\mathrm{supp}(\varphi_j)} \varphi_j(z) f(z-x)\,\mathrm{d}z \approx h f(x_j - x). \tag{54}$$

Therefore, for $|i-j|>1$

$$r_{ij} = -k\lambda \int_{\mathrm{supp}(\varphi_i)} \eta(x)\varphi_i(x)\,\mathrm{d}x \approx -k\lambda h \eta(x_i) = -k\lambda h^2 f(x_j - x_i). \tag{55}$$

Hence, r_{ij} is a function of the difference $(i-j)$ provided the spatial mesh size is constant. The approximated Toeplitz matrix $\mathbf{R}_2$ (denoted here with the same letter) is then completely determined by the vector:

$$\mathbf{a} = [r_{2,n-1}, r_{2,n-2}, \ldots, r_{2,4}, 0, 0, 0, r_{4,2}, r_{5,2}, \ldots, r_{n-1,2}]. \tag{56}$$

6.2. Convolutions and the FFT algorithm

The Discrete Fourier Transform (DFT) of a vector $\mathbf{d} = [d_0, d_2 \ldots, d_{R-1}]^{\mathrm{T}}$ is defined as:

$$D_k = \sum_{n=0}^{R-1} d_n e^{-i2\pi nk/R}, \quad k = 0, 1, \ldots, R. \tag{57}$$

The FFT is an algorithm designed to evaluate the DFT of a vector of length R in $\mathrm{O}(R \log R)$ operations. This is a significant improvement with respect to the direct evaluation, which has a computational cost of $\mathrm{O}(R^2)$ operations.

An important application of the DFT is in computing convolutions. Let $\{x_m\}$ and $\{y_m\}$ be two sequences with period R. The convolution sequence $z := x * y$ is defined component-wise as

$$z_n = \sum_{m=0}^{R-1} x_{m-n} y_m. \tag{58}$$

This is a so-called circulant convolution. We now use FFT to compute the vector $[z_0, \ldots, z_{R-1}]$. The periodic structure of x allows the derivation of the following simple relation:

$$Z_k = X_k \cdot Y_k, \tag{59}$$

where X, Y and Z denote the DFT of the sequences x, y and z, respectively. Now the vector $[z_0, \ldots, z_{R-1}]$ may be recovered by means of the Inverse Discrete Fourier Transform (IDFT):

$$z_n = \frac{1}{R} \sum_{k=0}^{R-1} Z_k e^{i2\pi kn/R}, \quad n = 0, 1, \ldots, R. \tag{60}$$

It is easy to see that, in the language of matrices and vectors, a circulant convolution may be represented as the product of a *circulant matrix* times a vector. Each row in a circulant matrix is by definition a circular shift of the previous row, a property that reflects the periodicity of the sequence in the convolution.

The next idea is to embed a Toeplitz matrix into a circulant matrix. As an example, let a Toeplitz matrix $\mathbf{T}(\mathbf{a})$ be determined by the vector $\mathbf{a} = [a_{-2}, a_{-1}, a_0, a_1, a_2]$ as follows

$$\mathbf{T}(\mathbf{a}) = \begin{bmatrix} a_0 & a_{-1} & a_{-2} \\ a_1 & a_0 & a_{-1} \\ a_2 & a_1 & a_0 \end{bmatrix}. \tag{61}$$

The matrix above may be embedded in a circulant matrix $\mathbf{C}$ of size 5 in the following way:

$$\mathbf{C} = \left[\begin{array}{ccc|cc} a_0 & a_{-1} & a_{-2} & a_2 & a_1 \\ a_1 & a_0 & a_{-1} & a_{-2} & a_2 \\ a_2 & a_1 & a_0 & a_{-1} & a_{-2} \\ \hline a_{-2} & a_2 & a_1 & a_0 & a_{-1} \\ a_{-1} & a_{-2} & a_2 & a_1 & a_0 \end{array}\right]. \tag{62}$$

If we define the vectors $\mathbf{d} := [d_0, d_1, d_2]^{\mathrm{T}}$ and $\tilde{\mathbf{d}} := [d_0, d_1, d_2, 0, 0]^{\mathrm{T}}$, then the product $\mathbf{T}(\mathbf{a})\mathbf{d}$ is the vector consisting of the first three elements of the product $\mathbf{C}\tilde{\mathbf{d}}$. As explained before, a product of a circulant matrix and a vector may be efficiently done by applying the FFT algorithm.

As a summary, following the ideas explained above the product of a Toeplitz matrix and a vector may be computed fast by "embedding" the Toeplitz matrix into a circulant matrix. The product of a circulant matrix and a vector is carried out in three FFT operations, namely, two DFT and one IDFT. We already verified that $\mathbf{R}_2$ is a Toeplitz matrix, so we may use the algorithm outlined here to multiply $\mathbf{R}_2$ by a vector for each step of iteration (48).

Note that the wrap-around effect normally observed when applying the DFT is not present in this framework, since the matrix–vector product is carried out exactly by embedding the Toeplitz matrix into a circular matrix. This would not be the case if we had applied DFT only to the Toeplitz matrix. Finally, for computational efficiency of the FFT algorithm, it is advisable to use a circulant matrix whose size is a power of 2. For further details on the computation of convolutions by FFT we refer to [31].

We summarize here the computational cost of the algorithm. For each time step we have the linear iteration (48) coupled with the FFT algorithm outlined in paragraph 6.2. For each iteration we need to solve a tridiagonal system and apply 3 times the FFT algorithm. The tridiagonal solver requires $\mathrm{O}(n)$ operations and each FFT requires $\mathrm{O}(n \log n)$. On the other hand, the number of iterations of splitting (48) is about 10 for the types of Lévy measures considered in this paper and for a tolerance of 10^{-8}. We conclude that the overall cost of the algorithm is $\mathrm{O}(n \log n)$ (also supported by the observed CPU times in Table 4 in the next section). As for the storage, only vectors of size $\mathrm{O}(n)$ need to be stored or updated.

7. Numerical experiments

For the numerical experiments in this section we assume $r = 0$ (see Remark 5.1).

Problem (19), for $\mu_J = 0$ has an analytic solution given by Merton's formula [26]:

$$w(t, s) = \sum_{m=0}^{\infty} \frac{e^{-\lambda'\tau}(\lambda'\tau)^m}{m!} C_{\mathrm{BS}}(\tau, s, K, r_m, \sigma_m) \tag{63}$$

where $\tau := T - t$, and recalling that $\zeta = e^{\sigma_J^2/2} - 1$, the rest of the parameters are given by

$$\lambda' = \lambda(1+\zeta), \qquad \sigma_m^2 = \sigma^2 + \frac{m\sigma_J^2}{\tau}, \qquad r_m = r - \lambda\zeta + \frac{m\ln(1+\zeta)}{\tau},$$

and C_{BS} denotes the Black–Scholes value of a call:

$$C_{\mathrm{BS}}(\tau, s, K, r, \sigma) = s\Phi(d_1) - Ke^{-r\tau}\Phi(d_2), \tag{64}$$

where

$$d_1 = \frac{\log(\frac{s}{K}) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}, \qquad d_2 = d_1 - \sigma\sqrt{\tau},$$

and Φ is the normal cumulative distribution function (27).

In general, for models where the characteristic function of the Lévy process is known, an analytical solution of PIDE (22) may be found using Fourier analysis [5,8,24]. In particular, for Kou's model, we have the following formula for a call option:

$$u(\tau, x) = \mathrm{Re}\left[-\frac{1}{\pi}\int_0^\infty \frac{\exp(-izx + \tau\psi(-z))}{z^2 - iz}\,\mathrm{d}y\right], \tag{65}$$

where $z = y + \rho i$, for $\rho > 1$, and ψ is the Lévy–Khintchine exponent of a double exponential Lévy process

$$\psi(z) = -\frac{1}{2}\sigma^2 z^2 - \left(\lambda\zeta + \frac{\sigma^2}{2}\right)iz + \lambda\left(\frac{p\alpha_1}{\alpha_1 - iz} + \frac{q\alpha_2}{\alpha_2 + iz} - 1\right). \tag{66}$$

We have carried out the following four experiments:

(1) Merton's model, using finite differences with BDF2. The integral is truncated and the non-homogeneity is kept; see Section 4.
(2) Merton's model, using finite elements with BDF2. Recall that the problem was transformed into a homogeneous problem; see Section 5.
(3) Kou's model, with finite elements and BDF2.
(4) Kou's model, with finite differences and BDF2.

Table 1 and Fig. 1 (first experiment) show that the second order is lost with the FD approximation for the whole interval, but not locally around the logarithm of the strike price which is most important, of course. The accuracy is not of $O(h^2)$ near the boundary of the computational domain. From Table 2 (second experiment) we see by contrast that the convergence is quadratic in the ℓ^∞-norm with finite elements. Table 3 shows the outcome of experiment 3, where the solution at x_K was found by integrating (65) with Simpson's rule on the interval $[0, 30]$ with 512 divisions and with $\rho = 1.5$. Observe that the convergence is also quadratic. The fourth experiment shows the quadratic convergence using finite differences and also the total CPU time with the FFT algorithm is incorporated, see Table 4. Note that the number of space and time steps is doubled each row. Therefore the total CPU time increases with about a factor of 4. In all experiments the stopping criterion for iteration (48) is given by the ℓ_∞-norm of the difference between two consecutive iterations, i.e.,

$$\left\|\mathbf{v}^{l+1} - \mathbf{v}^l\right\|_\infty < \varepsilon, \tag{67}$$

where the tolerance ε was set to 10^{-8}.

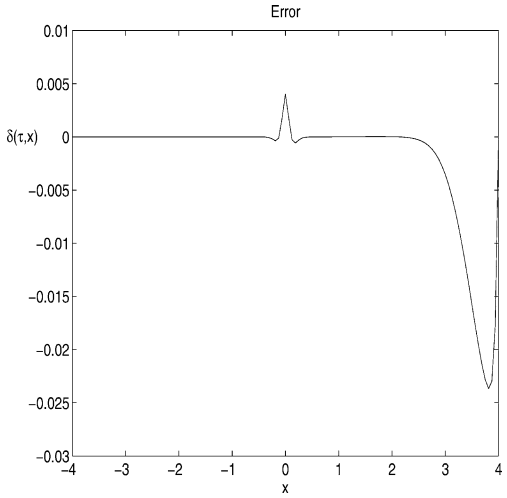

Fig. 1. Finite difference case. The function δ is the difference between the analytical solution and the numerical solution for Merton's model. The process parameters are the same as in Table 1. The number of spatial points is 129, $\tau = 0.2$ and the time step is $k = 0.1$.

Table 1
FD and BDF2 results for Merton's model. Point-wise and ℓ^∞ errors at maturity $T = 1$. Truncation point $x^* = 4$, volatility $\sigma = 0.2$, variance of the jumps $\sigma_J = 0.5$, intensity of the jumps $\lambda = 0.1$, strike price $K = 1$ and $x_K = \log(K)$

n	k	$T = 1$	
		l^∞ error	error at x_K
65	0.2	1.9569e–01	0.00442717
129	0.1	1.0011e–01	0.00102551
257	0.05	5.0416e–02	0.00025366
513	0.025	2.5272e–02	6.32903e–05
1025	0.0125	1.2646e–02	1.58379e–05

Table 2
FE and BDF2 results for Merton's model. Point-wise and ℓ^∞ errors at the two maturity times $T = 1$ and $T = 2$. Truncation point $x^* = 4$, volatility $\sigma = 0.2$, variance of the jumps $\sigma_J = 0.5$, intensity of the jumps $\lambda = 0.1$, strike price $K = 1$ and $x_K = \log(K)$

n	k	$T = 1$		$T = 2$	
		ℓ^∞ error	error at x_K	ℓ^∞ error	error at x_K
65	0.2	0.0020020	0.00147182	0.0013240	0.00091851
129	0.1	0.00048586	0.000369525	0.00032835	0.000228408
257	0.05	1.2167e–04	9.24111e–05	8.2418e–05	5.70452e–05
513	0.025	3.0282e–05	2.31407e–05	2.0550e–05	1.42507e–05
1025	0.0125	7.5252e–06	5.80396e–06	5.1032e–06	3.55107e–06
Solution at x_K			0.094135525		0.136963105

Table 3
FE and BDF2 results for Kou's model. Point-wise errors at maturity time $T = 0.2$. The truncation point $x^* = 6$ and the parameters of the process are: volatility $\sigma = 0.2$, $\alpha_1 = 3$, $\alpha_2 = 2$, $p = 0.5$, $\lambda = 0.2$, $K = 1$ and $x_K = \log(K)$

n	k	Num. sol. at x_K	error at x_K
65	0.2	0.0323466	0.0103295
129	0.1	0.0398864	0.0027897
257	0.05	0.0421572	0.0005189
513	0.025	0.0424579	0.0002182
Anal. sol. at x_K		0.0426761	

Table 4
FD and BDF2 results for Kou's model. Point-wise errors at maturity time $T = 0.2$ and CPU times on a Pentium IV, 1.70 GHz. The truncation point $x^* = 6$ and the parameters of the process are: volatility $\sigma = 0.2$, $\alpha_1 = 3$, $\alpha_2 = 2$, $p = 0.5$, $\lambda = 0.2$, $K = 1$ and $x_K = \log(K)$

n	k	Num. sol. at x_K	error at x_K	Total CPU-time
65	$T/10$	0.02438	0.01829	0.29 s
129	$T/20$	0.03407	0.00860	0.48 s
257	$T/40$	0.04086	0.00181	1.30 s
513	$T/80$	0.04240	0.00027	4.41 s
Anal. sol. at x_K		0.04267		

8. Conclusions

In this paper we investigated the numerical solution of a European option pricing problem in a market with a finite number of jumps, given that we a-priori know the Lévy–Khintchine representation of the underlying jump process. The option value is in general given as the solution of a partial integro-differential equation. We concentrated on two models, the classical model by Merton [26] and a more recent model by Kou [22,23].

Due to the non-locality of the integro-differential operator, the PIDE is numerically challenging if we discretize it by some fast converging implicit method. We showed that there is no need for "fully" explicit methods with severe time step constraints to treat the integral part. We found that a finite difference method combined with BDF2 in time gives second order accuracy close to the strike price while the homogenized finite element approach with BDF2 in time is second order accurate in the ℓ^∞-norm. Both FEM and FD are straightforwardly implemented and the resulting matrices differ only slightly due to the presence of the mass matrix in FEM. When it comes to the computation of the right side $\mathcal{L}\tilde{g}$ in FEM, it is essentially equivalent to the computation of the error function $\varepsilon(\tau, x, x^*)$ in the FD approach, but with the advantage that in FEM we have uniform convergence of the numerical solution to the analytical value.

For exponentially decaying kernels, like the two examples we considered, we observed that the suggested tridiagonal splitting has small spectral radius and that the fast Fourier transform may be used to speed up the iterative method, provided we have a uniform mesh in space. We have avoided the storage and inversion of a full matrix this way.

References

[1] L. Andersen, J. Andreasen, Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing, Rev. Derivatives Res. 4 (2000) 231–262.
[2] O.E. Barndorff, Processes of normal inverse Gaussian type, Finance Stochast. 2 (1998) 41–68.
[3] J. Bertoin, Lévy Processes, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996.
[4] F. Black, M.S. Scholes, The pricing of options and corporate liabilities, J. Political Economy 7 (1973) 637–654.
[5] S.I. Boyarchenko, S.Z. Levendorskiĭ, Non-Gaussian Merton–Black–Scholes Theory, Advanced Series on Statistical Science & Applied Probability, vol. 9, World Scientific, River Edge, NJ, 2002.
[6] M. Briani, C. La Chioma, R. Natalini, Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory, Numer. Mat., in press.
[7] P.P. Carr, H. Geman, D.B. Madan, The fine structure of asset returns: An empirical investigation, J. Business (2002).
[8] P.P. Carr, D.B. Madan, Option valuation using the fast fourier transform, J. Comput. Finance 2 (1999) 61–73.
[9] T.F. Coleman, Y. Li, A. Verma, Reconstructing the unknown local volatility function, J. Comput. Finance 2 (1999) 77–100.
[10] S.R. Das, S. Foresi, Exact solutions for bond and option prices with systematic jump risk, Rev. Derivatives Res. 1 (1996) 7–24.
[11] F. Delbaen, W. Schachermayer, A general version of the fundamental theorem of asset pricing, Math. Ann. 300 (1994) 463–520.
[12] Y. d'Halluin, P.A. Forsyth, G. Labahn, A penalty method for American options with jump diffusion processes, Numer. Math. 97 (2004) 321–352.
[13] B. Dupire, Pricing with a smile, RISK Magazine 1 (1999) 18–20.
[14] E. Eberlein, Application of generalized hyperbolic Lévy motions to finance, in: O. Barndorff-Nielsen, T. Mikoson, S. Resnick (Eds.), Lévy Processes—Theory and Applications, Birkhäuser, Boston, MA, 2001, pp. 319–336.
[15] H.U. Gerber, E.S.W. Shiu, Option pricing by Esscher transforms, Trans. Soc. Actuaries 46 (1995) 99–140.
[16] J.M. Harrison, D.M. Kreps, Martingales and arbitrage in multiperiod securities markets, J. Econom. Theory 20 (1979) 381–408.
[17] J.M. Harrison, S.R. Pliska, Martingales and stochastic integrals in the theory of continuous trading, Stochastic Process. Appl. 11 (1981) 215–260.
[18] S. Heston, A closed-form solution for options with stochastic volatility with applications to bond and currency options, Rev. Financ. Stud. 6 (1993) 327–343.
[19] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1994.
[20] J. Hull, A. White, The pricing of options with stochastic volatilities, J. Finance 42 (1987) 281–300.
[21] I. Karatzas, S.E. Shreve, Methods of Mathematical Finance, Applications of Mathematics, vol. 39, Springer, New York, 1998.
[22] S.G. Kou, A jump diffusion model for option pricing, Management Sci. 48 (2002) 1086–1101.
[23] S.G. Kou, H. Wang, Option pricing under a double exponential jump diffusion model, Working paper, Columbia University, 2001.
[24] A.L. Lewis, A simple option formula for general jump-diffusion and other exponential Lévy processes, in: 8th. Annual CAP Workshop on Derivative Securities and Risk Management, November 2001.
[25] A.M. Matache, T. von Petersdorff, C. Schwab, Fast deterministic pricing of options on Lévy driven assets, Working paper, ETH, Zürich, 2002.
[26] R.C. Merton, Option pricing when the underlying stocks are discontinuous, J. Financ. Econ. 5 (1976) 125–144.
[27] G.H. Meyer, The numerical valuation of options with underlying jumps, Acta Math. Univ. Comenian 67 (1998) 69–82.
[28] S. Raible, Lévy Processes in Finance: Theory, Numerics, and Empirical Facts, Ph.D. Thesis, Inst. für Mathematische Stochastik, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany, 2000.
[29] P.A. Samuelson, Rational theory of warrant pricing, Indust. Management Rev. 6 (1965) 13–32.
[30] K.-i. Sato, Basic results on Lévy processes, in: O. Barndorff-Nielsen, T. Mikoson, S. Resnick (Eds.), Lévy Processes—Theory and Applications, Birkhäuser, Boston, MA, 2001, pp. 3–37.
[31] C. Van Loan, Computational Frameworks for the Fast Fourier Transform, Frontiers in Applied Mathematics, vol. 10, SIAM, Philadelphia, PA, 1992.
[32] D.M. Young, Iterative Solution of Large Linear Systems, Academic Press, New York, 1971.