A FAST AND ACCURATE FFT-BASED METHOD FOR PRICING EARLY-EXERCISE OPTIONS UNDER LÉVY PROCESSES[*]

R. LORD[†], F. FANG[‡], F. BERVOETS[§], AND C. W. OOSTERLEE[¶]

Abstract. A fast and accurate method for pricing early exercise and certain exotic options in computational finance is presented. The method is based on a quadrature technique and relies heavily on Fourier transformations. The main idea is to reformulate the well-known risk-neutral valuation formula by recognizing that it is a convolution. The resulting convolution is dealt with numerically by using the fast Fourier transform. This novel pricing method, which we dub the convolution method, is applicable to a wide variety of payoffs and requires only the knowledge of the characteristic function of the model. As such, the method is applicable within many regular affine models, among which is the class of exponential Lévy models. For an M-times exercisable Bermudan option, the overall complexity is $O(MN \log_2(N))$, with N grid points used to discretize the price of the underlying asset. American options are priced efficiently by applying Richardson extrapolation to the prices of Bermudan options.

Key words. option pricing, Bermudan options, American options, convolution, Lévy processes, fast Fourier transform

AMS subject classifications. 65Y20, 65T50, 62P05, 60E10, 91B28

DOI. 10.1137/070683878

1. Introduction. When valuing and risk-managing exotic derivatives, practitioners demand fast and accurate prices and sensitivities. As the financial models and option contracts used in practice are becoming increasingly complex, efficient methods have to be developed to cope with such models. Aside from nonstandard exotic derivatives, plain vanilla options in many stock markets are actually of the American type. As any pricing and risk-management system has to be able to calibrate to these plain vanilla options, it is important to be able to value these American options quickly and accurately.

By means of the risk-neutral valuation formula, the price of any option without early-exercise features can be written as an expectation of the discounted payoff of this option. By starting from this representation one can apply several numerical techniques to calculate the price itself: Monte Carlo simulation, numerical solution of the corresponding partial (integro)differential equation (P(I)DE) and numerical integration. While the treatment of early-exercise features within the first two techniques is relatively standard, the pricing of such contracts via quadrature pricing techniques has not been considered until recently; see [2, 40]. Each of these methods has its merits and demerits, though for the pricing of American options the PIDE approach currently seems to be the clear favorite [22, 43].

[†]Financial Engineering, Rabobank International, Thames Court, 1 Queenhithe, London EC4V 3RL, UK (roger.lord@rabobank.com).
[‡]Delft University of Technology, Delft Institute of Applied Mathematics, Delft, The Netherlands (f.fang@ewi.tudelft.nl).
[§]Modelling and Research (TC 2121), Rabobank International, Thames Court, 1 Queenhithe, London EC4V 3RL, UK (frank.bervoets@rabobank.com).
[¶]CWI – Center for Mathematics and Computer Science, Amsterdam, The Netherlands (c.w. oosterlee@cwi.nl).

In the past couple of years a vast body of literature has considered the modeling of asset returns as infinite activity Lévy processes, due to the ability of such processes to adequately describe the empirical features of asset returns and at the same time provide a reasonable fit to the implied volatility surfaces observed in option markets. Valuing American options in such models is, however, far from trivial, due to the weakly singular kernels of the integral terms appearing in the PIDE, as reported in, e.g., [4, 5, 12, 23, 36, 42].

In this paper we present a quadrature-based method for pricing options with early-exercise features. The method combines the recent quadrature pricing methods of [2, 40] with the methods based on Fourier transformation pioneered by [9, 37, 32]. Though the transform methods so far have been used mainly for the pricing of European options, we show how early-exercise features can be incorporated naturally. The requirements of the method are that the increments of the driving processes are independent of each other and that the conditional characteristic function of the underlying asset is known. This is certainly the case for many exponential Lévy models and models from the broader class of regular affine processes of [15], which also encompasses the exponentially affine jump-diffusion class of [14]. In contrast to the PIDE methods, processes of infinite activity, such as the variance gamma (VG) or CGMY (named after Carr, Geman, Madan, and Yor) models, can be handled with relative ease.

The present paper is organized as follows. We start with an overview of the recent history of transform and quadrature methods in option pricing. Subsequently we introduce the novel method called the convolution (CONV) method for early-exercise options. Its high accuracy and speed are demonstrated by pricing several Bermudan and American options under geometric Brownian motion (GBM), VG, CGMY, and Kou's model.

2. Overview of transform and quadrature methods. All transform methods start from the risk-neutral valuation formula that, for a European option, reads:

$$(1) \qquad V(t, S(t)) = e^{-r\tau}\mathbb{E}\left[V(T, S(T))\right],$$

where V denotes the value of the option, r is the risk-neutral interest rate, t is the current time point, T is the maturity of the option, and $\tau = T - t$. The variable S denotes the asset on which the option contract is based. The expectation is taken with respect to the risk-neutral probability measure. As (1) is an expectation, it can be calculated via numerical integration provided that the probability density is known in closed form.

This is not the case for many models which do, however, have a characteristic function in closed form.[1] A number of papers starting from Heston [21] have solved the problem differently. Focusing on a plain vanilla European call option, note that (1) can be written very generally as:

$$(2) \qquad V(t, S(t)) = e^{-r\tau}(F(t,T) \cdot \mathbb{S}(S(T) > K) - K \cdot \mathbb{P}(S(T) > K)),$$

where $F(t, T)$ is the forward price of the underlying asset at time T, as seen from t, and $\mathbb{P}$ and $\mathbb{S}$ indicate, respectively, the risk-neutral probability measure and the stock price measure, induced by taking the asset price itself as the numeraire asset. Note that (2) has the same form as the celebrated Black–Scholes formula. As such, both of the cumulative probabilities can be found by inverting the characteristic function, an

---

[1]Or the probability density involves complicated special functions, whereas the characteristic function is comparatively easier.

approach which in the form used here dates back to Gurland [19] and Gil-Pelaez [18].
We can write:

$$(3) \qquad \mathbb{P}(S(T) > K) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re} \frac{e^{-iuk}\phi(u)}{iu} du,$$

$$(4) \qquad \mathbb{S}(S(T) > K) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re} \frac{e^{-iuk}\phi(u-i)}{iu\phi(-i)} du,$$

where i is the imaginary unit, k is the logarithm of the strike price K, Re denotes taking the real part of the integral, and ϕ is the characteristic function of the log-underlying, i.e.,

$$\phi(u) = \mathbb{E}\left[e^{iu \ln S(T)}\right].$$

Carr and Madan [9] considered another approach. Note that L^1-integrability is a sufficient condition for the Fourier transform of a function to exist. A call option is not L^1-integrable with respect to the logarithm of the strike price, as:

$$\lim_{k \to -\infty} V(t, S(t)) = S(t).$$

Damping the option price with $\exp(\alpha k)$ for $\alpha > 0$ solves this, however, and Carr and Madan proposed the following solution:

$$\mathcal{F}\{e^{\alpha k}V(t,k)\} = e^{-r\tau} \int_0^\infty \mathrm{Re} \; e^{iuk}e^{\alpha k}\mathbb{E}\left[(S(T) - e^k)^+\right] dk$$

$$(5) \qquad\qquad = \frac{e^{-r\tau}\phi(u - (\alpha+1)i)}{-(u - \alpha i)(u - (\alpha+1)i)},$$

where we now consider the option price V as a function of time and k. Though this approach was new to mathematical finance, the idea of damping functions on the positive real line in order to be able to find their Fourier transform dates back to Dubner and Abate [13].

A necessary and sufficient condition for (5) to exist is that

$$|\phi(u - (\alpha+1)i)| \le \phi(-(\alpha+1)i) = \mathbb{E}[S(T)^{\alpha+1}] < \infty,$$

i.e., that the $(\alpha + 1)$th moment of the asset price exists. The option price can be recovered by inverting (5) and undamping

$$(6) \qquad V(t,k) = \frac{1}{2\pi} e^{-r\tau - \alpha k} \int_0^\infty \mathrm{Re} \; e^{-iuk} \frac{\phi(u - (\alpha+1)i)}{-(u - \alpha i)(u - (\alpha+1)i)} du.$$

The representation in (6) has two distinct advantages over (2). First, it requires only one numerical integration. Second, whereas (2) can suffer from cancellation errors, the numerical stability of (6) can be controlled by means of the damping coefficient α; see [30, 33]. Finally we note that, if we discretize (6) with Newton–Cotes quadrature, the option price can very efficiently be evaluated by means of the FFT, yielding option prices over a whole range of strike prices.

The methods discussed until here can handle only the pricing of European options. Before turning to methods that can handle early-exercise features, let us introduce some notation. We define the set of exercise dates as $\mathcal{T} = \{t_1, \ldots, t_M\}$ and $0 = t_0 \le t_1$.

For ease of exposure we assume the exercise dates are equally spaced, so that $t_{m+1} - t_m = \Delta t$. The best-known examples of options with early-exercise features are American and Bermudan options. American options can be exercised at any time prior to the option's expiry, whereas Bermudan options can be exercised only at certain dates in the future. If the option is exercised at some time $t \in \mathcal{T}$, the holder of the option obtains the exercise payoff $E(t, S(t))$. The Bermudan option price can then be found via backward induction as

$$
(7) \quad
\begin{cases}
V(t_M, S(t_M)) &= E(t_M, S(t_M)), \\
C(t_m, S(t_m)) &= e^{-r\Delta t} \mathbb{E}_{t_m}\left[V(t_{m+1}, S(t_{m+1})) \right], \\
V(t_m, S(t_m)) &= \max\{ C(t_m, S(t_m)), E(t_m, S(t_m)) \}, \\
V(t_0, S(t_0)) &= C(t_0, S(t_0)),
\end{cases}
\quad m = M-1, \ldots, 1,
$$

with C the continuation value of the option and V the value of the option immediately prior to the exercise opportunity. Note that we now have attached a subscript to the expectation operator to indicate that the expectation is being taken with respect to all information available at time t_m.

The dynamic programming problem in (7) is a successive application of the risk-neutral valuation formula, as we can write the continuation value as

$$
(8) \quad C(t_m, S(t_m)) = e^{-r\Delta t} \int_{-\infty}^{\infty} V(t_{m+1}, y) f(y | S(t_m)) dy,
$$

where $f(y | S(t_m))$ represents the probability density describing the transition from $S(t_m)$ at t_m to y at t_{m+1}. Based on (7) and (8), the quadrature (QUAD) method was introduced in [2]. The method requires the transition density to be known in closed form, which is the case in, e.g., the Black–Scholes model and Merton's jump-diffusion model. This requirement is relaxed in [40], where the QUAD-FFT method is introduced. The underlying idea is that the transition density can be recovered by inverting the characteristic function, so that the QUAD method can be used for a wider range of models. As such, the QUAD-FFT method, also applied in [11], effectively combines the QUAD method with the early transform methods. The overall complexity of both methods is $O(MN^2)$ for an M-times exercisable Bermudan option, with N grid points used to discretize the price of the underlying asset.

The complexity of this method can be improved to $O(MN \log_2(N))$ if the underlying is a monotone function of a Lévy process. We will demonstrate this shortly. In the remainder we assume, as is common, that the underlying process is modeled as an exponential of a Lévy process. Let $x_1, \ldots, x_N$ be a uniform grid for the log-asset price. If we discretize (8) by the trapezoidal rule, we can write the continuation value in matrix form as

$$
(9) \quad \mathbf{C}(t_m) \approx e^{-r\Delta t} \Delta x \left[\mathbf{FV} - \frac{1}{2} \left(V(t_{m+1}, x_1) \mathbf{f}_1 + V(t_{m+1}, x_N) \mathbf{f}_N \right) \right],
$$

where

$$
\mathbf{f}_i = \begin{pmatrix} f(x_i | x_1) \\ \vdots \\ f(x_i | x_N) \end{pmatrix}, \qquad \mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_N), \qquad \mathbf{V} = \begin{pmatrix} V(t_{m+1}, x_1) \\ \vdots \\ V(t_{m+1}, x_N) \end{pmatrix},
$$

and $f(y | x)$ now denotes the transition density in logarithmic coordinates. The key observation is that the increments of Lévy processes are independent, so that due to

the uniform grid

$$(10) \qquad \mathbf{F}_{j,\ell} = f(y_j|y_\ell) = f(y_{j+1}|y_{\ell+1}) = \mathbf{F}_{j+1,\ell+1}.$$

The matrix $\mathbf{F}$ is hence a Toeplitz matrix. A Toeplitz matrix can easily be represented as a circulant matrix, which has the property that the FFT algorithm can be employed to efficiently calculate matrix-vector multiplications. Therefore, an overall computational complexity of $O(MN\log_2(N))$ can be achieved. Though this method is significantly faster than [2] or [40], we do not pursue it in this paper, as the method we develop in the next section has the same complexity yet requires fewer operations.

The previous literature does not seem to have picked up on a presentation by Reiner [38], where it was recognized that for the Black–Scholes model the risk-neutral valuation formula in (8) can be seen as a convolution or correlation of the continuation value with the transition density. As convolutions can be handled very efficiently by means of the FFT, an overall complexity of $O(MN\log_2 N)$ can be achieved. By working forward instead of backward in time, a number of discrete path-dependent options can also be treated, such as lookbacks, barriers, Asian options, and cliquets. Building on Reiner's idea, Broadie and Yamamoto [7] reduced the complexity to $O(MN)$ for the Black–Scholes model by combining the double-exponential integration formula and the fast Gauss transform. Their technique is applicable to any model in which the transition density can be written as a weighted sum of Gaussian densities, which is the case in, e.g., Merton's jump-diffusion model.

As one of the defining properties of a Lévy process is that its increments are independent of each other, the insight of Reiner has a much wider applicability than only to the Black–Scholes model. This is especially appealing since the usage of Lévy processes in finance has become more established nowadays. By combining Reiner's ideas with the work of Carr and Madan, we introduce the CONV method. The complexity of the method is $O(MN\log_2 N)$ for an M-times exercisable Bermudan option.

Our method has similarities with both the quadrature pricing and the PIDE methods. Though the complexity of our method is smaller than that of the QUAD variants, we share the construct that time steps are required only on the exercise dates of the product. However, our application of the FFT to approximate convolution integrals bears more resemblance to the approximation of the integral term in the numerical solution of a PIDE. Here Andersen and Andreasen [1] were the first to suggest that for jump-diffusion models the integral term in the PIDE can be calculated efficiently via use of the FFT, rendering the complexity $O(MN\log_2 N)$ instead of $O(MN^2)$. Since then, similar ideas have been applied to various jump-diffusion and infinite activity Lévy models [20, 4, 5, 42]. We will compare our method in terms of accuracy and speed to two PIDE methods in Appendix C. Alternative methods for valuing options in Lévy models are the lattice-based approach of Këllezi and Webber [28], which is $O(MN^2)$, and the multinomial tree of Maller, Solomon, and Szimayer [35], which is $O(M^2)$.

3. The CONV method. The main premise of the CONV method is that the conditional probability density $f(y|x)$ in (8) depends on x and y only via their difference

$$(11) \qquad f(y|x) = f(y - x).$$

Note that x and y do not have to represent the asset price directly; they could be monotone functions of the asset price. The assumption made in (11) therefore cer-

tainly holds when the asset price is modeled as a monotone function of a Lévy process, since one of the defining properties of a Lévy process is that its increments are independent of each other. As mentioned earlier, we choose to work with exponential Lévy models in the remainder of this paper. In this case x and y in (11) represent the log-spot price. By including (11) in (8) and changing variables $z = y - x$, the continuation value can be expressed as

$$
(12) \qquad C(t_m, x) = e^{-r\Delta t} \int_{-\infty}^{\infty} V(t_{m+1}, x + z) f(z) dz,
$$

which is a cross correlation[2] of the option value at time t_{m+1} and the density $f(z)$ or, equivalently, a convolution of $V(t_{m+1})$ and the conjugate of $f(z)$. If the density function has a closed-form expression, it may be beneficial to proceed along the lines of (9). However, for many exponential Lévy models we either do not have a closed-form expression for the density (e.g., the CGMY/KoBoL (named after Koponen, Boyarchenko, and Levendorskii) model of [6, 8] and many regular affine models) or, if we have, it involves one or more special functions (e.g., the VG model). In contrast, the characteristic function of the log-spot price can typically be obtained in closed form or, in the case of regular affine models, via the solution of a system of ordinary (integro)differential equations (OIDEs).

We therefore take the Fourier transform of (12). The insight that the continuation value can be seen as a convolution is useful here, as the Fourier transform of a convolution is the product of the Fourier transforms of the two functions being convolved. In the remainder we will employ the following definitions for the continuous Fourier transform and its inverse:

$$
(13) \qquad \hat{h}(u) := \mathcal{F}\{h(t)\}(u) \quad = \int_{-\infty}^{\infty} e^{iut} h(t) dt,
$$

$$
(14) \qquad h(t) := \mathcal{F}^{-1}\{\hat{h}(u)\}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iut} \hat{h}(u) du.
$$

If we dampen the continuation value (12) by a factor $\exp(\alpha x)$ and subsequently take its Fourier transform, we obtain

$$
e^{r\Delta t} \mathcal{F}\{c(t_m, x)\}(u) = \int_{-\infty}^{\infty} e^{iux} e^{\alpha x} \int_{-\infty}^{\infty} V(t_{m+1}, x + z) f(z) dz dx
$$

$$
(15) \qquad\qquad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{iu(x+z)} v(t_{m+1}, x + z) e^{-iz(u-i\alpha)} f(z) dz dx,
$$

where in the first step we used the risk-neutral valuation formula from (12). We introduced the convention that small letters indicate damped quantities, i.e., $c(t_m, x) = e^{\alpha x} C(t_m, x)$ and $v(t_m, x + z) = e^{\alpha(x+z)} V(t_m, x + z)$. By changing the order of integration and remembering that $x = y - z$, we obtain

$$
e^{r\Delta t} \mathcal{F}\{c(t_m, x)\}(u) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{iuy} v(t_{m+1}, y) dy \, e^{-i(u-i\alpha)z} f(z) dz
$$

$$
= \int_{-\infty}^{\infty} e^{iuy} v(t_{m+1}, y) dy \int_{-\infty}^{\infty} e^{-i(u-i\alpha)z} f(z) dz
$$

$$
(16) \qquad = \mathcal{F}\{e^{\alpha y} V(t_{m+1}, y)\}(u) \, \phi(-(u - i\alpha)).
$$

---

[2]The cross correlation of two functions $f(t)$ and $g(t)$, denoted $f \star g$, is defined by $f \star g \equiv \bar{f}(-t) * g(t) = \int_{-\infty}^{\infty} f(\tau) g(t + \tau) d\tau$ where "$*$" denotes the convolution operator.

In the last step we used the fact that the complex-valued Fourier transform of the density is the extended characteristic function

$$(17) \qquad \phi\left(x + yi\right) = \int_{-\infty}^{\infty} e^{i(x+yi)z} f(z)dz,$$

which is well-defined when $\phi(yi) < \infty$, as $|\phi(x + yi)| \leq |\phi(yi)|$. As such, (16) puts a restriction on the damping coefficient α, because $\phi(\alpha i)$ must be finite.

The difference with the Carr–Madan approach in (5) is that we take a transform with respect to the log-spot price instead of the log-strike price, something which [32, 37] also consider for European option prices. The damping factor is again necessary when considering, e.g., a Bermudan put, as then $V(t_{m+1}, x)$ tends to a constant when $x \to -\infty$ and as such is not L^1-integrable. For the Bermudan put we must choose $\alpha > 0$. Though other values of α are allowed in principle, we need to know the payoff transform itself in order to apply Cauchy's residue theorem; see [32, 30, 37]. This restriction on α will *disappear* when we switch to a discretized version of (16) in the next section. The Fourier transform of the damped continuation value can thus be calculated as the product of two functions, one of which, the extended characteristic function, is readily available in exponential Lévy models. We now recover the continuation value by taking the inverse Fourier transform of the right-hand side of (16) and calculate $V(t_m)$ as the maximum of the continuation and the exercise value at t_m. This procedure, as outlined in (7), is repeated recursively until we obtain the option price at time t_0. In pseudocode the CONV algorithm is presented in Algorithm 1.

ALGORITHM 1: THE CONV ALGORITHM FOR BERMUDAN OPTIONS.

$V(t_M, x) = E(t_M, x)$ for all x
$E(t_0, x) = 0$ for all x
For $m = M - 1$ to 0
    Dampen $V(t_{m+1}, x)$ with $\exp(\alpha x)$ and take its Fourier transform
    Calculate the right-hand side of (16)
    Calculate $C(t_m, x)$ by applying Fourier inversion to (16) and undamping
    $V(t_m, x) = \max\{(E(t_m, x), C(t_m, x)\}$
Next m

In Appendix A we demonstrate how the hedge parameters can be calculated in the CONV method. As differentiation is exact in Fourier space, they will be more stable than when calculated via finite-difference-based approximations.

The following section deals with the implementation of the CONV algorithm. In particular we employ the FFT to approximate the continuous Fourier transforms that are involved.

4. Implementation details of the CONV method. The essence of the CONV method is the calculation of a convolution:[3]

$$(18) \qquad c(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \hat{v}(u) \phi\left(-(u - i\alpha)\right) du,$$

where $\hat{v}(u)$ is the Fourier transform of v:

$$(19) \qquad \hat{v}(u) = \int_{-\infty}^{\infty} e^{iuy} v(y)dy.$$

---

[3] For notational convenience we have dropped the discounting term out of the equation.

In the remainder of this section we will focus on (18) and (19) for notational ease. To be able to use the FFT means that we have to switch to logarithmic coordinates. For this reason the state variables x and y will represent $\ln S(t_m)$ and $\ln S(t_{m+1})$, respectively, up to a constant shift. This section is organized as follows. Section 4.1 deals with the discretization of the convolution in (18) and (19). Section 4.2 analyzes the error made by one step of the CONV method and provides guidelines on choosing the grids for u, x, and y. Section 4.3 considers the choice of grid further and investigates how to deal with points of discontinuity. This will prove to be important if we want to guarantee a smooth convergence of the algorithm. Finally, sections 4.4 and 4.5 deal with the pricing of Bermudan and American options, respectively, with the CONV method.

4.1. Discretizing the convolution. We approximate both integrals in (18) and (19) by a discrete sum, so that the FFT algorithm can be employed for their computation. This necessitates the use of uniform grids for u, x, and y:

$$(20) \qquad u_j = u_0 + j\Delta u, \qquad x_j = x_0 + j\Delta x, \qquad y_j = y_0 + j\Delta y,$$

where $j = 0, \ldots, N-1$. Though they may be centered around a different point, the x- and y-grids have the same mesh size: $\Delta x = \Delta y$. Further, the Nyquist relation must be satisfied, i.e.,

$$(21) \qquad \Delta u \cdot \Delta y = \frac{2\pi}{N}.$$

In principle we could use the fractional FFT algorithm (FrFT) which does not require the Nyquist relation to be satisfied. Numerical tests indicated, however, that this advantage of the FrFT does not outweigh the speed of the FFT, so we use the FFT throughout. Details about the exact location of x_0 and y_0 will be given in section 4.3. Inserting (19) into (18) and approximating (19) with a general Newton–Cotes rule and (18) with the left-rectangle rule yields:

$$(22) \qquad c(x_p) \approx \frac{\Delta u \Delta y}{2\pi} \sum_{j=0}^{N-1} e^{-iu_j x_p} \phi\left(-(u_j - i\alpha)\right) \sum_{n=0}^{N-1} w_n e^{iu_j y_n} v(y_n)$$

for $p = 0, \ldots, N-1$. When using the trapezoidal rule we choose the weights w_n as:

$$(23) \qquad w_0 = \frac{1}{2}, \qquad w_{N-1} = \frac{1}{2}, \qquad w_n = 1 \text{ for } n = 1, \ldots, N-2.$$

Though it may seem that the choice for the left-rectangle rule in (18) would cause the leading error term in (22) to be $O(du)$, the error analysis will show that the Newton–Cotes rule one uses to approximate (19) is the determining factor. Inserting the definitions of our grids into (22) yields:

$$(24) \quad c(x_p) \approx \frac{e^{-iu_0(x_0+p\Delta y)}}{2\pi} \Delta u \sum_{j=0}^{N-1} e^{-ijp2\pi/N} e^{ij(y_0-x_0)\Delta u} \phi\left(-(u_j - i\alpha)\right) \hat{v}(u_j),$$

where the Fourier transform of v is approximated by:

$$(25) \qquad \hat{v}(u_j) \approx e^{iu_0 y_0} \Delta y \sum_{n=0}^{N-1} e^{ijn2\pi/N} e^{inu_0\Delta y} w_n v(y_n).$$

Let us now define the DFT and its inverse of a sequence x_p, $p = 0, \dots, N-1$, as:

$$(26) \qquad \mathcal{D}_j\{x_n\} := \sum_{n=0}^{N-1} e^{ijn2\pi/N} x_n, \qquad \mathcal{D}_n^{-1}\{x_j\} = \frac{1}{N} \sum_{j=0}^{N-1} e^{-ijn2\pi/N} x_j.$$

Though the reason why will be described later, let us set $u_0 = -N/2\Delta u$. As $e^{inu_0\Delta y} = (-1)^n$, this finally leads us to write (24) and (25) as:

$$(27) \quad c(x_p) \approx e^{iu_0(y_0-x_0)}(-1)^p \mathcal{D}_p^{-1}\{e^{ij(y_0-x_0)\Delta u}\phi\left(-(u_j - i\alpha)\right)\mathcal{D}_j\{(-1)^n w_n v(y_n)\}\}.$$

4.2. Error analysis for Bermudan options. A first inspection of (27) suggests that errors will arise from two[4] sources:
- discretization of both integrals in (18) and (19);
- truncation of these integrals.

We will now consider both integrals in (18) and (19) separately and estimate both discretization and truncation errors by applying the error analysis of [3]. [30] recently combined an analysis similar to theirs with sharp upper bounds on European plain vanilla option prices to find a sharp error bound for the discretized Carr–Madan formula. Though it is possible to use parts of their analysis, we found that the resulting error bounds overestimated the true error of the discretized CONV formula. To be precise, the discretization of (18) does not contribute to the error of (27), which is why we can use the left-rectangle rule to approximate (18). Based on a Fourier series expansion of the damped continuation value $c(x)$, we will show why this is the case. This is natural, as the Fourier transform itself is generalized from Fourier series of periodic functions by letting their period approach infinity. We start from the risk-neutral valuation formula with damping and without discounting:

$$(28) \qquad c(x) = \int_{-\infty}^{\infty} v(x+z)e^{-\alpha z} f(z) dz.$$

Suppose that the density $f(z)$ is negligible outside $[-A/2, A/2]$ and that we are interested in $c(x)$ only for values of x in $[-B/2, B/2]$. According to (28), we require knowledge of $v(x)$ for x in $[-(A+B)/2, (A+B)/2]$. Truncating the integration range in (28) leads to

$$(29) \qquad c(x) \approx \widetilde{c}_1(x) = \int_{-A/2}^{A/2} v(x+z)e^{-\alpha z} f(z) dz.$$

We can replace v by its Fourier series expansion on $[-L/2, L/2]$, where we defined $L = A + B$:

$$\widetilde{c}_1(x) = \int_{-A/2}^{A/2} \sum_{j=-\infty}^{\infty} v_j e^{-ij(x+z)\frac{2\pi}{L}} e^{-\alpha z} f(z) dz$$

$$(30) \qquad = \sum_{j=-\infty}^{\infty} v_j e^{-ij\frac{2\pi}{L}x} \int_{-A/2}^{A/2} e^{-(\alpha+ij\frac{2\pi}{L})z} f(z) dz,$$

---

[4]If the spot price for which we want to calculate our option price does not lie on the grid, another error source will be added as we will have to interpolate between option prices.

and the Fourier series coefficients of v are given by:

$$(31) \qquad v_j = \frac{1}{L} \int_{-L/2}^{L/2} v(y) e^{ij\frac{2\pi}{L}y} dy.$$

Second, we can replace the integral in (30) by the known characteristic function:

$$\widetilde{c}_1(x) \approx \widetilde{c}_2(x) = \sum_{j=-\infty}^{\infty} v_j e^{ij\frac{2\pi}{L}x} \phi\left(-\left(j\frac{2\pi}{L} - i\alpha\right)\right).$$

The sum of both truncation errors now equals:

$$e_1(L) + e_2(L) = \widetilde{c}_2(x) - c(x)$$

$$(32) \qquad = \int_{\mathbb{R}\setminus[-A/2, A/2]} \left(v(x+z) - \sum_{j=-\infty}^{\infty} v_j e^{ij\frac{2\pi}{L}(x+z)}\right) e^{-\alpha z} f(z) dz.$$

Note that only the parameter L will appear in the final discretization. A general guideline for choosing L is to ensure that the mass of the density outside $[-L/2, L/2]$ is negligible. The function $\widetilde{c}_2$ can, at least on this interval, be interpreted as an approximate Fourier series expansion of $c(x)$.

The third error arises by truncating the infinite summation from $-N/2$ to $N/2-1$, leading to $\widetilde{c}_3$ and its associated error e_3:

$$\widetilde{c}_3 = \sum_{j=-N/2}^{N/2-1} v_j e^{-ij2\pi x/L} \phi\left(-\left(j\frac{2\pi}{L} - i\alpha\right)\right),$$

$$(33) \qquad |e_3(L, N)| = |\widetilde{c}_2(x) - \widetilde{c}_3(x)| \le \sum_{|j|=N/2}^{\infty} |v_j| \left|\phi\left(-\left(j\frac{2\pi}{L} - i\alpha\right)\right)\right|.$$

To further bound this error we require knowledge about the rate of decay of Fourier coefficients. It is well known that, even if v is only piecewise C^1 on $[-L/2, L/2]$, its Fourier series coefficients v_j tend to zero as $j \to \pm\infty$. The modulus of v_j can therefore be bounded as:

$$(34) \qquad |v_j| \le \frac{\eta_1(L)}{|j|^{\beta_1}}.$$

By $\eta_i(\cdot)$ we denote a bounding constant, which depends only on the quantities specified between the brackets. For functions that are piecewise continuous on $[-L/2, L/2]$ but whose L-periodic extension is discontinuous, we have $\beta_1 = 1$. The following example demonstrates that this is the case for a European put payoff.

Example 4.2.1 (European put). Suppose that we have a European put payoff and that $y = \ln S(t) - \ln K$. Then the payoff function equals $v(y) = e^{\alpha y} K(1 - e^y)^+$, and its Fourier series coefficients equal:

$$(35) \quad v_j = K\left(e^{-L\alpha/2}(-1)^j \frac{e^{-L/2} - 1}{L(\alpha+1) + 2\pi ij} - L\frac{e^{-L\alpha/2}(-1)^j - 1}{(L(\alpha+1) + 2\pi ij)(L\alpha + 2\pi ij)}\right).$$

Clearly, $\beta_1 = 1$ in (34), though when $L \to \infty$ and $j2\pi/L \to u$ it can be shown that the Fourier series coefficient converges to the Fourier transform of the payoff function, which can be seen to be $O(u^{-2})$ from (5).

The characteristic function can be assumed to have power decay:

$$(36) \qquad |\phi(x + yi)| \leq \frac{\eta_2(y)}{|x|^{\beta_2}}.$$

This is overly conservative for, e.g., the Black–Scholes model, where the characteristic function of the log-underlying $\phi(x + yi)$ decays as $\exp(-cx^2)$, or the Heston model, where the characteristic function has exponential decay. For the most popular Lévy models, however, the power decay assumption is appropriate. The VG model, for example, has $\beta_2 = 2\tau/\nu$, with τ being the time step between two exercise dates.

Remark. It should be noted that the error analysis here is valid for Bermudan options and *not* for American options in the limit $\tau \to 0$. In section 4.5 we will price American options by Richardson extrapolation on the prices of Bermudan options with a varying number of exercise opportunities. For problems where the time intervals are very small and the characteristic function decays slowly, we may encounter some numerical problems due to the oscillatoriness of the integrand. These problems are, however, well-known and can in part be overcome by choosing a proper value of parameter α.

Combining (34) and (36) yields:

$$
\begin{aligned}
|e_3(L, N)| &\leq \sum_{|j|=N/2}^{\infty} \frac{\eta_1(L)}{|j|^{\beta_1}} \frac{\eta_2(\alpha)}{\left(\frac{2\pi}{L}\right)^{\beta_2} |j|^{\beta_2}} \leq \eta_3(\alpha, L) \int_{N/2-1}^{\infty} x^{-\beta_1-\beta_2} dx \\
&= \eta_3(\alpha, L) \frac{(N/2-1)^{1-\beta_1-\beta_2}}{\beta_1 + \beta_2 - 1},
\end{aligned}
\tag{37}
$$

where $\eta_3(\alpha, L) = 2\eta_1(L)\eta_2(\alpha)(2\pi/L)^{-\beta_2}$. We finally arrive at the discretized CONV formula in (27) by approximating the Fourier series coefficients of v in (33) with a Newton–Cotes rule:

$$(38) \qquad \widetilde{v}(u_j) = \frac{1}{L} \Delta y \sum_{n=0}^{N-1} w_n e^{iu_j y_n} v(y_n).$$

This is equal to the right-hand side of (24) multiplied by $1/L$. It becomes clear that we can set $\Delta y = L/N$ and $y_0 = -L/2$.

Inserting (38) in $\widetilde{c}_3$ results in the final approximation:

$$(39) \qquad \widetilde{c}_4(x) = \sum_{j=-N/2}^{N/2-1} \widetilde{v}(u_j) e^{-ij2\pi/Lx} \phi\left(-\left(j\frac{2\pi}{L} - i\alpha\right)\right).$$

By assuming that the chosen Newton–Cotes rule is of $O(N^{-\beta_3})$, one can bound

$$(40) \qquad |v_j - \widetilde{v}(u_j)| \leq \frac{\eta_4(\alpha, L)}{N^{\beta_3}},$$

leading to the following error estimate for $\beta_2 \neq 1$:

$$|e_4(L,N)| = |\widetilde{c}_3(x) - \widetilde{c}_4(x)| \leq \frac{\eta_4(\alpha,L)}{N^{\beta_3}} \sum_{j=-N/2}^{N/2-1} \left| \phi\left(-\left(j\frac{2\pi}{L} - i\alpha \right) \right) \right|$$

$$\leq \frac{\eta_4(\alpha,L)}{N^{\beta_3}} \left(3\phi(i\alpha) + 2\eta_2(\alpha) \left(\frac{2\pi}{L} \right)^{-\beta_2} \sum_{j=2}^{N/2} \frac{1}{|j|^{\beta_2}} \right)$$

$$(41) \qquad = \frac{\eta_5(\alpha,L)}{N^{\beta_3}} + \frac{\eta_6(\alpha,L)}{(1-\beta_2)N^{\beta_3}} \left(\frac{2^{\beta_2-1}}{N^{\beta_2-1}-1} \right),$$

with $\eta_5(\alpha,L) = 3\eta_4(\alpha,L)\phi(i\alpha)$ and $\eta_6(\alpha,L) = 2\eta_2(\alpha)\eta_4(\alpha,L)(2\pi/L)^{-\beta_2}$. For $\beta_2 = 1$ the second error term should be $\eta_6(\alpha,L)\ln(N/2)/N^{\beta_3}$.

Summarizing, if we use a Newton–Cotes rule to discretize the Fourier transform of the payoff function v, the error in the discretized CONV formula can be bounded as:

$$|c(x) - \widetilde{c}_4(x)| \leq e_1(L) + e_2(L,N) + e_3(L,N) + e_4(L,N)$$

$$(42) \qquad = e_1(L) + e_2(L) + O(N^{-\beta_3 + \min(1-\beta_2,0)}).$$

As demonstrated, when we exercise into a European put or call, we will have $\beta_1 = 1$. The magnitude of β_3 will depend on the interplay between the chosen Newton–Cotes rule and the nature of the payoff function, something we investigate in the next section. However, let us assume that $\beta_3 \geq 2$, which we may expect if we use the trapezoidal rule or more sophisticated Newton–Cotes rules. This implies that, aside from the truncation error, the order of convergence will be:
- $O(N^{-\beta_3})$ for characteristic functions decaying faster than a polynomial;
- $O(N^{\min(-\beta_3-\min(0,\beta_2-1))})$ for characteristic functions having power decay.

For the Black–Scholes model this implies that the order of convergence will be fully dictated by the chosen Newton–Cotes rule, whereas in the VG model where $\beta_2 = 2\tau/\nu$ we can lose up to an order for sufficiently small time steps.

One final word should be mentioned on the damping coefficient α. In the continuous version of the algorithm in section 3, α was chosen such that the damped continuation value was L^1-integrable. The direct construction of the discretized CONV formula in section 4.2 via a Fourier series expansion of the continuation value replaces L^1-integrability on $(-\infty, \infty)$ with L^1-summability on $[-L/2, L/2]$, so that the restriction on α is removed. In principle any value of α is allowed as long as $\phi(i\alpha)$ is finite. Nevertheless it makes sense to adhere to the guidelines stated before, as the function will resemble its continuous counterpart more and more as L increases. The impact of α on the accuracy of the CONV algorithm is investigated in section 5.1.

This concludes the error analysis of one step of the CONV algorithm. It is easy to show that the error is not magnified further in the remaining time steps. The leading error of our algorithm is therefore dictated by the time step where the order of convergence in (42) is the smallest.

Remark. We explicitly mention that *aliasing*, a commonly observed feature when dealing with a convolution of sampled signals by means of the FFT, is not a problem in our application. We encounter a convolution of the characteristic function and the DFT of a vector with option values. The DFT is periodical, but this would make the convolution circular only if the characteristic function would also be obtained by a DFT. We can, however, work with the analytical characteristic function, which is not periodic.

4.3. Dealing with discontinuities. Our focus in this section lies on achieving smooth convergence for the CONV algorithm. As numerical experiments have shown that it is difficult to achieve smooth convergence with higher-order Newton–Cotes rules, we will from here on focus on the second-order trapezoidal rule in (23). Smooth convergence is desirable as we will be using extrapolation techniques later on to price American options in section 4.5.

The previous section analyzed the error in the discretized CONV formula when we use a Newton–Cotes rule to integrate the function V, the maximum of the continuation value, and the exercise value. If we focus on a simple Bermudan put, it is clear that already at the last time step this function will have a discontinuous first derivative. Certainly it is also possible that V itself is discontinuous; think of contracts with a barrier clause. This will affect the order of convergence.

It is well known that, if we numerically integrate a function with (a finite number of) discontinuities, we should split up the integration domain such that we are only integrating continuous functions. Appendix B demonstrates this for the trapezoidal rule. In particular, we show that the trapezoidal rule remains second-order if only the first derivative of the integrand is discontinuous, at the cost of nonsmooth convergence. If the integrand itself is discontinuous, the trapezoidal rule loses an order. Smooth second-order convergence can be restored by placing the discontinuities on the grid. This notion has often been utilized in lattice-based techniques, though the solutions have more often than not been payoff-specific. An approach that is more or less payoff-independent was recently proposed in [25], generalizing previous work in [26]. Unfortunately, we cannot use their methodology here, as our desire to use the FFT binds us to a uniform grid.

Before investigating how to handle discontinuities in the CONV algorithm, we collect the results from the previous sections and restate the grid choice for the basic CONV algorithm. By equating the grids for x and y for now we have:

$$u_j = \left(j - \frac{n}{2}\right)\Delta u, \qquad x_j = y_j = \left(j - \frac{1}{2}\right)\Delta y, \quad j = 0, \ldots, N - 1.$$

Here x and y represent, up to a constant shift, $\ln S(t_m)$ and $\ln S(t_{m+1})$, respectively. If in particular $x = \ln S(t_m) - \ln S(0)$ and $y = \ln S(t_{m+1}) - \ln S(0)$, so that x and y represent total log-returns, we will refer to this discretization as Discretization I. A convenient property of this discretization is that the spot price always lies on the grid, so that no costly interpolation is required to determine the desired option value. Note that we need to ensure that the mass of the density of x and y outside $[-L/2, L/2]$ is negligible. Though more sophisticated approximations can be devised, we use a rule of thumb from [40] which chooses L as a multiple of the standard deviation of $\ln S(t_m)$, i.e.,

$$(43) \qquad L = \delta \cdot \sqrt{-\left.\frac{\partial^2 \phi(t_m, u)}{\partial u^2}\right|_{u=0} + \left(\left.\frac{\partial \phi(t_m, u)}{\partial u}\right|_{u=0}\right)^2},$$

where $\phi(t_m, u)$ is the characteristic function of $\ln S(t_m)$ conditional upon $\ln S(0)$ and δ is a proportionality constant. Note that there is a trade-off in the choice of L: As we set $\Delta y = L/N$, the Nyquist relation implies $\Delta u = 2\pi/L$ and hence $[u_0, u_{N-1}] = [-N\pi/L, (N-2)\pi/L]$. While larger values of L imply smaller truncation errors, they also cause the range of the grid in the Fourier domain to be smaller, so that the error in turn will be larger initially.

A choice of grid that allows us to place one discontinuity on the grid is described here. Suppose that at time t_m the discontinuity we would like to place on the grid is d_m. We then shift our grid by a small amount to get

$$(44) \qquad x_j = \epsilon_x + \left(j - \frac{L}{2}\right)\Delta y, \qquad y_j = \epsilon_y + \left(j - \frac{L}{2}\right)\Delta y,$$

where $\epsilon_x = d_m - \lceil d_m/\Delta x\rceil \cdot \Delta x$ and ϵ_y is chosen in a similar fashion. This discretization will be referred to as Discretization II. Even for plain vanilla European options where only one time step is required, this is useful. By choosing $\epsilon_y = \ln K/S(0)$ and $\epsilon_x = 0$ we ensure that the discontinuity of the call or put payoff lies on the y-grid, and the spot price lies on the x-grid. When more discontinuities are present it seems impossible to guarantee smooth convergence while keeping the restriction of a uniform grid. In order to still be able to use the computational speed of the FFT, we will then have to resort to, e.g., the discontinuous FFT algorithm of [16] or a recent transform inversion technique in [27]. These directions are left for further research. Discretization II is, however, well-suited for the pricing of Bermudan and American options, as we will show in the following sections.

4.4. Pricing Bermudan options. It is well known that, in the case of American options, under Black–Scholes dynamics the derivative of the value function is continuous (smooth fit principle). This is, however, not the case anymore when pricing Bermudan options, for which the function V in (7) will have a discontinuous first derivative. Though at the final exercise time t_M the location of this discontinuity is known, this is not the case at previous exercise times. All we know after approximating V is that the discontinuity is contained in an interval of width Δx, say, $[x_\ell, x_{\ell+1}]$.

If we proceed with the CONV algorithm without placing the discontinuity on the grid, the algorithm will show a nonsmooth convergence. In the QUAD method [2] this is overcome by equating the exercise payoff and the continuation value and solving numerically for the location of the discontinuity. In our framework this can be quite costly, so that we propose an effective alternative. We can use a simple linear interpolation to locate the discontinuity, say, d_m:

$$(45) \qquad d_m \approx \frac{x_{\ell+1}(C(t_m, x_\ell) - E(t_m, x_\ell)) - x_\ell(C(t_m, x_{\ell+1}) - E(t_m, x_{\ell+1}))}{(C(t_m, x_\ell) - E(t_m, x_\ell)) - (C(t_m, x_{\ell+1}) - E(t_m, x_{\ell+1}))}.$$

We assume that the error made in determining d_m in (45) is negligible compared to the other error terms appearing (see also the discussion in Appendix B).

As in Discretization II we can now shift the grid such that d_m lies on it and recalculate both the continuation and the exercise value. In particular, note that the inner DFT of (27) does not have to be recalculated; the only term that is affected is the outer inverse DFT. Moreover, calculating d_m automatically gives us an approximation of the exercise boundary.

It is demonstrated in Appendix B that, if we choose the trapezoidal rule, a linear interpolation is sufficient to guarantee a smooth convergence. Obviously, if higher-order Newton–Cotes rules are used, higher-order interpolation schemes will have to be employed to locate the discontinuity. The resulting algorithm we use to value Bermudan call or put options with a fixed strike K is presented below in pseudocode.

ALGORITHM 2: DETAILS OF THE ALGORITHM FOR VALUING BERMUDAN
OPTIONS.

> Ensure that the strike K lies on the grid by setting $\epsilon_y = \ln K/S(0)$
> For $m = M - 1$ to 1
>   Equate the x-grid at t_m to the y-grid at t_{m+1}
>   Compute $C(t_m, x)$ through (27)
>   Locate x_ℓ and $x_{\ell+1}$ and approximate d_m, e.g., via (45)
>   Set $\epsilon_x = d_m$ and recompute $C(t_m, x)$
>   Calculate $V(t_m, x) = \max\left(E(t_m, x), C(t_m, x)\right)$
>   Set the y-grid at t_m to be equal to the x-grid at t_m
> Next m
> Set $\epsilon_x = 0$ such that the initial spot price lies on the grid
> Compute $V(0, x) = C(0, x)$ using (27)

4.5. Pricing American options. Within the CONV algorithm there are basically two approaches to value an American option. One way is to approximate an American option by a Bermudan option with many exercise opportunities, and the other is to use Richardson extrapolation on a series of Bermudan options with an increasing number of exercise opportunities. The method we use has been described in detail by Chang, Chung, and Stapleton [10], though the approach in finance dates back to Geske and Johnson [17]. The QUAD method in [2] also uses the same technique to price American options. We restrict ourselves to the essentials here. Let $V(\Delta t)$ be the price of a Bermudan option with a maturity of T years where the exercise dates are Δt years apart. It is assumed that $V(\Delta t)$ can be expanded as

$$(46) \qquad V(\Delta t) = V(0) + \sum_{i=1}^{\infty} a_i (\Delta t)^{\gamma_i},$$

with $0 < \gamma_i < \gamma_{i+1}$. $V(0)$ is the price of the American option. Classical extrapolation procedures assume that the exponents γ_i are known, which means that we can use $n + 1$ Bermudan prices with varying Δt in order to eliminate n of the leading order terms in (46). The only paper we are aware of that considers an expansion of the Bermudan option price in terms of Δt is Howison [24], who shows that $\gamma_1 = 1$ for the Black–Scholes model. Nevertheless, numerical tests indicate that the assumption $\gamma_i = i$ produces satisfactory results for the Lévy models we consider.

5. Numerical experiments. By various experiments we show the accuracy and speed of the CONV method. The method's flexibility is presented by showing results for three asset price processes: GBM, VG, and CGMY. In addition, we value a multiasset option to give an impression of the CPU times required to value a basket option of moderate dimension. The pricing problems considered are of European, Bermudan, and American style. We typically present the (positive or negative) error $V(0, S(0)) - V_{ref}(0, S(0))$, where the reference value $V_{ref}(0, S(0))$ is obtained either via another numerical scheme or via the CONV algorithm with 2^{20} grid points. In the tables to follow we will also present the error convergence defined as the absolute value of the ratio between two consecutive errors. A factor of 4 then denotes second-order convergence. All single-asset tests were performed in C^{++} on an Intel Xeon CPU 5160, 3.00 GHz with 2 GB RAM. The multiasset calculations were programmed in C on an Intel Core 2 CPU 6700, 2.66 GHz and 8 GB RAM. In Appendix C we compare the speed and accuracy of our method to that of two PIDE methods, one

for the VG model, a Lévy process with infinite activity, and a recent PIDE scheme for Kou's jump-diffusion model.

5.1. Characteristic function for Lévy price processes. The CONV method, as outlined in section 3, is particularly well suited for exponential Lévy models whose characteristic functions are available in closed form. We will briefly review some defining properties of these models before discussing the extended CGMY/KoBoL model (hereafter the extended CGMY model) of [6, 8] that will be used to analyze the performance of the CONV method. For more background information we refer you to [12] for the usage of Lévy processes in a financial context and to [39] for a detailed analysis of Lévy processes in general.

In exponential Lévy models the asset price is modeled as an exponential function of a Lévy process $L(t)$:

$$(47) \qquad S(t) = S(0)\exp(L(t)).$$

Though the CONV method can be adapted to cope with discrete dividend payments, for ease of exposure we assume that the asset pays a continuous stream of dividends, measured by the dividend rate q. In addition, we assume the existence of a bank account $B(t)$ which evolves according to $dB(t) = rB(t)dt$, r being the risk-free rate. Recall that a process $L(t)$ on $(\Omega, \mathcal{J}, P)$, with $L(0) = 0$, is a Lévy process if:

1. it has independent increments;
2. it has stationary increments;
3. it is stochastically continuous; i.e., for any $t \geq 0$ and $\epsilon > 0$ we have

$$(48) \qquad \lim_{s \to t} \mathbb{P}(|L(t) - L(s)| > \epsilon) = 0.$$

The first property (cf. (11)) is exactly the property we required to recognize a cross correlation in the risk-neutral valuation formula. Each Lévy process can be characterized by a triplet (μ, σ, ν), with $\mu \in \mathbb{R}, \sigma \geq 0$, and ν a measure satisfying $\nu(0) = 0$ and

$$(49) \qquad \int_{\mathbb{R}} \min\left(1, |x|^2\right)\nu(dx) < \infty.$$

In terms of this triplet the characteristic function of the Lévy process equals:

$$(50) \qquad \begin{aligned} \phi(u) &= \mathbb{E}[\exp\left(iuL(t)\right)] \\ &= \exp\left(t\right)\left(i\mu u - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} (e^{iux} - 1 - iux\mathbf{1}_{[|x|<1]})\nu(dx)\right), \end{aligned}$$

the celebrated Lévy–Khinchine formula. As is common in most models nowadays we assume that (47) is formulated directly under the risk-neutral measure. To ensure that the reinvested relative price $e^{qt}S(t)/B(t)$ is a martingale under the risk-neutral measure, we require that

$$(51) \qquad \phi(-i) = \mathbb{E}[\exp\left(L(t)\right)] = e^{(r-q)t},$$

which is satisfied if we choose the drift μ as:

$$(52) \qquad \mu = r - q - \frac{1}{2}\sigma^2 - \int_{\mathbb{R}} (e^x - 1 - x\mathbf{1}_{[|x|<1]})\nu(dx).$$

The motivation behind using more general Lévy processes than the Brownian motion with drift is the fact that the Black–Scholes model is not able to reproduce the volatility skew or smile present in most financial markets. Over the past few years it has been shown that several exponential Lévy models are, at least to some extent, able to reproduce the skew or smile. Most of our examples will stem from the CGMY model. Its underlying Lévy process is characterized by the triple $(\mu, \sigma, \nu_{CGMY})$, where the Lévy density is specified as:

(53)
$$\nu_{CGMY}(x) = \begin{cases} C\dfrac{\exp(-G|x|)}{|x|^{1+Y}} & \text{if } x < 0, \\[2ex] C\dfrac{\exp(-M|x|)}{|x|^{1+Y}} & \text{if } x > 0. \end{cases}$$

The parameters satisfy $C \geq 0$, $G \geq 0$, $M \geq 0$, and $Y < 2$. The condition $Y < 2$ is induced by the requirement that Lévy densities integrate x^2 in the neighborhood of 0. Conveniently, the characteristic function of the log-asset price can be found in closed form as:

(54)
$$\phi(u) = \exp\left(iu\mu t - \frac{1}{2}u^2\sigma^2 t + tC\Gamma(-Y)[(M-iu)^Y - M^Y + (G+iu)^Y - G^Y] \right),$$

where $\Gamma(x)$ is the gamma function. One can verify that the parameters G and M represent, respectively, the smallest and the largest finite moment in the model, as $\phi(-iu) = \mathbb{E}[S(t)^u]$ is infinite for $u < -G$ and for $u > M$. The model encompasses several models. When $\sigma = 0$ and $Y = 0$, we obtain the VG model, which is often parameterized slightly differently with parameters[5] σ, θ, and ν related to C, G, and M, respectively, through:

(55) $$C = \frac{1}{\nu}, \qquad G = \frac{\theta}{\sigma^2} + \sqrt{\frac{\theta^2}{\sigma^4} + \frac{2}{\nu\sigma^2}}, \qquad M = -\frac{\theta}{\sigma^2} + \sqrt{\frac{\theta^2}{\sigma^4} + \frac{2}{\nu\sigma^2}}.$$

Finally, when $C = 0$ the model collapses to the Black–Scholes model.

To conclude this section, Table 1 contains five parameter sets which will be used in various tests throughout this section. The only two parameters we have not specified yet are δ from (43), which determines the range of the grid, and the damping coefficient

TABLE 1
Parameter sets in the numerical experiments.

T1-GBM:	$S(0) = 100$,	$r = 0.1$,	$q = 0$,	$\sigma = 0.25$;
T2-VG:	$S(0) = 100$,	$r = 0.1$,	$q = 0$,	$\sigma = 0.12$,
	$\theta = -0.14$,	$\nu = 0.2$;		
T3-CGMY:	$S(0) = 1$,	$r = 0.1$,	$q = 0$,	$\sigma = 0$,
	$C = 1$,	$G = 5$,	$M = 5$,	$Y = 0.5$;
T4-CGMY:	$S(0) = 90$,	$r = 0.06$,	$q = 0$,	$\sigma = 0$
	$C = 0.42$,	$G = 4.37$,	$M = 191.2$,	$Y = 1.0102$;
T5-GBM:	$S(0) = 40$,	$r = 0.06$,	$q = 0.04$,	$\sigma_i = 0.2$,
	$\rho_{ij} = 0.25$.			

---

[5]The parameters σ and ν should not be confused with the volatility and Lévy density of the Lévy triplet.
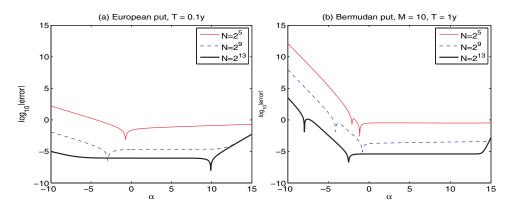
FIG. 1. *Error of CONV method under T2-VG and K = 110 for a European and a Bermudan put in dependence of parameter α.*

α. For all GBM tests we set $\delta = 20$; for the other Lévy models, which have fatter tails, we use $\delta = 40$.

Regarding the choice of α, Lord and Kahl [33] have demonstrated recently how to approximate the optimal damping coefficient when the payoff transform is known, which increases the numerical stability of the Carr–Madan formula. This is particularly effective for in/out-of-the-money options and options with short maturities. Though their rationale can to some extent be carried over to the pricing of European plain vanilla options (the difference being that now the payoff transform is also approximated numerically), the problem becomes much more opaque when dealing with Bermudan options. To see this, note that the continuation value of the Bermudan option at the penultimate exercise date equals that of a European option. At each grid point, the European option will have a different degree of moneyness, calling for a different value of α per grid point. Which single choice for α will be optimal is not clear at all, a problem which becomes more complex as the number of exercise dates increases. What is evident from Figure 1, where we present the error of the CONV algorithm as a function of α for a European and a Bermudan put under T2-VG, is that there is a relatively large range for which the error is stable. In all numerical experiments we will set $\alpha = 0$, which, at least for our examples, produces satisfactory results.

5.2. European call under GBM and VG. First of all, we evaluate the CONV method for pricing European options under VG. The parameters for the first test are from T2-VG with $T = 1$. Figure 2 shows that Discretizations I and II generate results of similar accuracy. What we notice from Figure 2 is that the only option with a stable convergence in Discretization I is the at-the-money option with $K = 100$. It is clear that placing the strike on the y-grid in Discretization II ensures a regular second-order convergence. The results are obtained in comparable CPU time. From the error analysis in section 4.2 it became clear that, for short maturities in the VG model, the slow decay of the characteristic function ($\beta_2 = 2\tau/\nu$) might impair the second-order convergence. To demonstrate this, we choose a call option with a maturity of 0.1 years, and $K = 90$. Table 2 presents the error of Discretization II for this option in models T1-GBM and T2-VG. The convergence under GBM is clearly of a regular second order. From the error analysis we expect the convergence under VG to be of first order. The nonsmooth convergence observed in Table 2 is caused
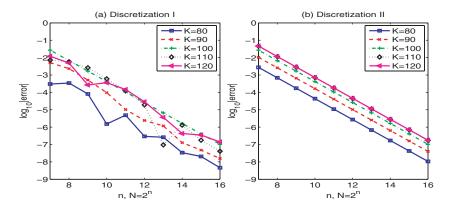
FIG. 2. *Convergence of the two discretization methods for pricing European call options at various K under T2-VG; left: Discretization* I; *right: Discretization* II.

TABLE 2
CPU time, error, and convergence rate for European call options under T1-GBM and T2-VG,
K = 90, and T = 0.1 (using Discretization II).

$(N = 2^n)$	GBM: $V_{ref}(0, S(0)) = 11.1352431$			VG: $V_{ref}(0, S(0)) = 10.9937032$		
n	Time (msec)	Error	Conv.	Time (msec)	Error	Conv.
7	0.095	$-2.08e-3$	–	0.15	$-2.91e-4$	–
8	0.20	$5.22e-4$	4.0	0.29	$-1.42e-4$	2.1
9	0.34	$-1.30e-4$	4.0	0.55	$-4.61e-5$	3.1
10	0.58	$-3.26e-5$	4.0	1.04	$-9.49e-6$	4.9
11	1.08	$-8.15e-6$	4.0	2.04	$-8.55e-7$	11.1
12	2.15	$-2.04e-6$	4.0	4.19	$7.97e-7$	1.1

by the highly oscillatory integrand. Note that all reference values are based on an adaptive integration of the Carr–Madan formula; all CPU times, in milliseconds, are determined after averaging the times of 1000 experiments.

In Appendix A the Greeks of the GBM call from Table 2 are computed.

5.3. Bermudan option under GBM and VG. Turning to Bermudan options, we compare Discretizations I and II for 10-times exercisable Bermudan put options under both T1-GBM and T2-VG. The reference values reported in Table 3 and 4 are found by the CONV method with 2^{20} grid points.

It is shown in Tables 3 and 4 that both Discretizations I and II give results of similar accuracy. Discretization I uses somewhat less CPU time, but Discretization

TABLE 3
CPU time, error, and convergence rate pricing a 10-times exercisable Bermudan put under
T1-GBM; K = 110, T = 1, and $V_{ref}(0, S(0)) = 11.98745352$.

$(N = 2^n)$	Discretization I			Discretization II		
n	Time (msec)	Error	Conv.	Time (msec)	Error	Conv.
7	0.13	$9.09e-3$	-	0.23	$-2.72e-2$	-
8	0.25	$-1.29e-3$	7.0	0.46	$-7.36e-3$	3.7
9	0.48	$1.80e-6$	717.8	0.90	$-2.00e-3$	3.7
10	1.09	$2.71e-5$	0.1	2.00	$-5.22e-4$	3.8
11	2.00	$-9.31e-6$	2.9	3.85	$-1.32e-4$	4.0
12	3.98	$-1.31e-5$	0.7	7.84	$-3.31e-5$	4.0

TABLE 4

CPU time, error, and convergence rate pricing a 10-times exercisable Bermudan put under T2-VG; K = 110 and T = 1 with reference value $V_{ref}(0, S(0)) = 9.040646119$.

$(N = 2^n)$	Discretization I			Discretization II		
n	Time (msec)	Error	Conv.	Time (msec)	Error	Conv.
7	0.18	$-8.45e-2$	-	0.28	$-9.63e-2$	-
8	0.35	$-9.02e-3$	9.4	0.55	$-1.07e-2$	9.0
9	0.68	$1.70e-4$	53.1	1.09	$-2.27e-3$	4.7
10	1.33	$2.04e-4$	0.8	2.15	$-6.06e-4$	3.8
11	2.67	$4.28e-5$	4.8	4.38	$-1.59e-4$	3.8
12	5.64	$1.11e-5$	3.8	9.29	$-4.08e-5$	3.9

II shows a regular second-order convergence, enabling the use of extrapolation. The computational speed of both discretizations is highly satisfactory.

5.4. American options under GBM, VG, and CGMY. Because Discretization II yields a regular convergence, we choose it in this section to price American options. We compare the accuracy and CPU time of the 2 approximation methods mentioned in section 4.5, i.e., the direct approximation via a Bermudan option and the repeated Richardson extrapolation technique. For the latter we opted for 2 extrapolations on 3 Bermudan options with 128, 64, and 32 exercise opportunities, which gave robust results. In our first test we price an American put under T1-GBM. The reference value was obtained by solving the Black–Scholes PDE on a very fine grid. The performance of both approximation methods is summarized in Table 5, where "$P(N/2)$" denotes that the American option is approximated by an $N/2$-times exercisable Bermudan option. "Richardson" denotes the results obtained by the 2-times repeated Richardson extrapolation scheme. It is evident that the extrapolation-based method converges fastest and costs far less CPU time than the direct approximation approach (e.g., to reach an accuracy of 10^{-4}, the extrapolation method is approximately 50 times faster).

In Appendix A the Greeks of the American put from Table 5 are computed.

In the remaining tests we demonstrate the ability of the CONV method to price American options accurately under alternative dynamics, by using the VG and both CGMY test sets. All reported reference values were generated with the CONV method on a mesh with 2^{20} points and 2-times Richardson extrapolation on 512-, 256-, and 128-times exercisable Bermudans. We have included one CGMY test with $Y < 1$ and one with $Y > 1$, as the latter is considered a hard test case when numerically solving the corresponding PIDE. Both CGMY tests stem from the PIDE literature, where reference values for the same American puts were reported as 0.112171 for

TABLE 5

CPU time and errors for an American put under T1-GBM, with $K = 110$, $T = 1$, and $V_{ref}(0, S(0)) = 12.169417$.

$(N = 2^n)$	$P(N/2)$			Richardson		
n	Time (msec)	Error	Conv.	Time (msec)	Error	Conv.
7	0.97	$-5.85e-2$	–	3.30	$-3.06e-2$	–
8	3.71	$-2.23e-3$	2.6	6.63	$-7.75e-3$	3.9
9	14.80	$-9.31e-3$	2.4	14.01	$-2.06e-3$	3.8
10	59.98	$-4.16e-3$	2.2	28.38	$-5.19e-4$	4.0
11	251.66	$-1.95e-3$	2.1	66.39	$-1.22e-4$	4.3
12	1108.09	$-9.39e-4$	2.1	151.85	$-2.10e-5$	5.8

TABLE 6
CPU time and errors for American puts under VG and CGMY.

	T2-VG		T3-CGMY		T4-CGMY	
	$K = 110, T = 1$		$K = 1, T = 1$		$K = 98, T = 0.25$	
$(N = 2^n)$	$V_{ref}(0, S(0)) = 10.0000$		$V_{ref}(0, S(0) = 0.112152$		$V_{ref}(0, S(0) = 9.225439$	
n	Time (msec)	Error	Time (msec)	Error	Time (msec)	Error
7	3.42	$-4.53e - 2$	3.82	$4.58e - 5$	3.83	$3.38e - 2$
8	6.85	$4.26e - 2$	7.60	$9.52e - 5$	7.68	$6.63e - 3$
9	14.29	$1.34e - 2$	15.87	$-1.03e - 4$	15.78	$-1.94e - 3$
10	28.99	$-5.00e - 3$	32.21	$-1.58e - 5$	33.37	$-5.41e - 6$
11	61.67	$-1.88e - 2$	68.16	$-1.09e - 5$	68.59	$-1.72e - 4$
12	135.09	$1.31e - 3$	148.16	$-3.73e - 6$	147.96	$-7.94e - 5$

T3-CGMY [4] and 9.2254842 for T4-CGMY [42]. The VG parameter set originally stems from [34] and is used in examples in [28, 40, 35]. In [35] the American option price is reported as 10. The reference values we use are calculated with the CONV method (using 2^{20} grid points and Richardson extrapolation on Bermudans with 512, 256, and 128 exercise opportunities) and agree up to four digits with the values from the literature. Though the convergence in Table 6 is less stable than for Bermudan options, the results in this section indicate that the CONV method is able to price American options under a wide variety of Lévy processes. A reasonable accuracy can be obtained quite quickly, so that it might be possible to calibrate a model to the prices of American options.[6]

5.5. 4D basket options under GBM. The CONV method can easily be generalized to higher dimensions. The only assumption that the multidimensional model is required to satisfy is the independent increments assumption in (11). We do not state the multidimensional version of Algorithm 1 here as it is a trivial generalization of the univariate case. Its ability to price options of a moderate dimension is demonstrated by considering a 4-asset basket put option. Upon exercise at time t_i, the payoff is

$$(56) \qquad V(t_i, \mathbf{S}(\mathbf{t_i})) = \max\left(K - \frac{1}{4}\sum_{p=1}^{4} S_p(t_i), 0\right).$$

The results of pricing a European and a 10-times exercisable Bermudan put under T5-GBM are summarized in Table 7. The CPU times on the tensor-product grids are very satisfactory, especially as the results on the coarse grids obtained in only a

TABLE 7
CPU time and prices for multiasset European and 10-times exercisable Bermudan basket put options under T5-GBM, $K = 40$, and $T = 1$.

	European		10-times exerc. Bermudan	
N	Result	Time (sec)	Result	Time (sec)
16^4	1.6428	0.02	1.7721	0.15
32^4	1.6537	0.51	1.7390	3.12
64^4	1.6539	7.0	1.7394	61.6
128^4	1.6538	159.2	1.7393	1511.7

---

[6]The majority of exchange-traded options in the equity markets are American.

few seconds seem to have converged within practical tolerance levels. In order to be able to price higher-dimensional problems, the multidimensional CONV method is combined with sparse grids in [31].

6. Conclusions. In this paper we have presented a novel FFT-based method for pricing options with early-exercise features, the CONV method. Like other FFT-based methods, it is flexible with respect to the choice of asset price process and the type of option contract, which has been demonstrated in numerical examples for European, Bermudan, and American options. Path-dependent exotics can in principle also be valued by a forward propagation in time, though this has not been demonstrated here. The crucial assumption of the method is that the underlying assets are driven by processes with independent increments, whose characteristic function is readily available. Though we have mainly focused on univariate exponential Lévy models, the techniques presented here certainly also extend to multivariate models, as section 5.5 has shown. The main strengths of the method are its flexibility and computational speed. By using the FFT to calculate convolutions we achieve a complexity of $O(MN \log_2 N)$, where N is the number of grid points and M is the number of exercise opportunities of the option contract. In comparison, the QUAD method of [2] is $O(MN^2)$. We have compared the CONV method to two PIDE schemes in Appendix C. The conclusion of this experiment is that we expect the CONV method to have an edge over PIDE schemes for the pricing of Bermudan options, in particular in exponential Lévy models with infinite activity. However, there will always be special cases, such as the Black–Scholes model and Kou's jump-diffusion model, for which highly efficient P(I)DE schemes can be designed. The speed of the method may make it possible to calibrate models to the prices of American options, as exchange-traded options are mainly of the American type. Future research will focus on the usage of more advanced quadrature rules, combined with speeding up the method for high-dimensional problems.

Appendix A. The hedge parameters. Here we present the CONV formulas for two important hedge parameters Δ and Γ, defined as

$$(57) \qquad \Delta = \frac{\partial V}{\partial S} = \frac{1}{S}\frac{\partial V}{\partial x}, \qquad \Gamma = \frac{\partial^2 V}{\partial S^2} = \frac{1}{S^2}\left(-\frac{\partial V}{\partial x} + \frac{\partial^2 V}{\partial x^2}\right).$$

As it is relatively easy to derive the corresponding CONV formulas, we merely present them here. For notational convenience we define:

$$(58) \qquad \mathcal{F}\{e^{\alpha x}V(t_0, x)\} = e^{-r\Delta t}A(u),$$

where $A(u) = \mathcal{F}\{e^{\alpha y}V(t_1, y)\} \cdot \phi(-u + i\alpha)$, and we assume that $t_1 > 0$. We now obtain the CONV formula for Δ, as

$$(59) \qquad \Delta = \frac{e^{-\alpha x}e^{-r\Delta t}}{S}\left[\mathcal{F}^{-1}\{-iuA(u)\} - \alpha\mathcal{F}^{-1}\{A(u)\}\right],$$

and for Γ:

$$\Gamma = \frac{e^{-\alpha x}e^{-r\Delta t}}{S^2}\left[\mathcal{F}^{-1}\{(-iu)^2 A(u)\} - (1 + 2\alpha)\mathcal{F}^{-1}\{-iuA(u)\}\right.$$
$$(60) \qquad\qquad\qquad \left. + \alpha(\alpha + 1)\mathcal{F}^{-1}\{A(u)\}\right].$$

Note that the only additional calculations occur at the final step of the CONV algorithm, where we calculate the value of the option given the continuation and exercise

TABLE 8
Accuracy of hedge parameters for a European call under T1-GBM; $K = 90$ and $T = 0.1$.

$(N = 2^n)$	European call				
	$\Delta_{ref} = 0.933029$		$\Gamma_{ref} = 0.01641389$		
d	Δ error	Conv.	Γ error	Conv.	
7	$-3.75e - 4$	–	$3.79e - 5$	–	
8	$-9.37e - 5$	4.0	$9.43e - 6$	4.0	
9	$-2.34e - 5$	4.0	$2.35e - 6$	4.0	
10	$-5.86e - 6$	4.0	$5.88e - 7$	4.0	
11	$-1.46e - 6$	4.0	$1.47e - 7$	4.0	
12	$-3.66e - 7$	4.0	$3.68e - 8$	4.0	

TABLE 9
Values of hedge parameters for an American put under T1-GBM; $K = 90$ and $T = 0.1$.

$(N = 2^n)$	American put	
d	$\Delta_{ref} = -0.62052$	$\Gamma_{ref} = 0.0284400$
7	-0.62170	0.028498
8	-0.62035	0.028687
9	-0.62050	0.028464
10	-0.62053	0.028463
11	-0.62054	0.028463
12	-0.62055	0.028463

values at time t_1. Since differentiation is exact in Fourier space, the rate of convergence of the Greeks will be the same as that of the value. To demonstrate this we evaluate the delta and gamma under T1-GBM of the European call from Table 2 and the American put from Table 5. For both tests we choose Discretization II. Tables 8 and 9 present the results. The reference values for the European call option are analytic solutions; for the American call these were found by numerically solving the Black–Scholes PDE on a very fine grid. Note that the delta and gamma of the American put converge to a slightly different value—this is due to our approximation of the American option via 2 Richardson extrapolations on 128-, 64-, and 32-times exercisable Bermudans. If we would increase the number of exercise opportunities of the Bermudan options, the delta and gamma would, at the cost of a longer computation time, converge to their true values.

Appendix B. Error analysis of the trapezoidal rule. Suppose that we are integrating $f \in C^\infty$ over an interval $[a, b]$. The discretization error induced by approximating this integral with the trapezoidal rule follows from the Euler–Maclaurin summation formula:

$$(61) \qquad \int_a^b f(x)dx - T(a, b, f, \Delta x) = \sum_{j=1}^{\infty} (\Delta x)^{2j} \frac{B_{2j}}{(2j)!} \left(f^{(2j-1)}(b) - f^{(2j-1)}(a) \right),$$

where B_j is the jth Bernoulli number and $T(a, b, f, \Delta x)$ is the trapezoidal sum:

$$(62) \qquad T(a, b, f, \Delta x) = \Delta x \left\{ \sum_{j=1}^{N-2} f(x_j) + \frac{1}{2}(f(a) + f(b)) \right\},$$

with $\Delta x = (b - a)/(N - 1)$ and $x_j = a + j\Delta x$. From (61) it is clear that, if the value of the first derivative is not the same in a and b, the trapezoidal rule is of order $1/N^2$.

The trapezoidal rule can obviously also be applied to functions that are piecewise continuously differentiable. The convergence may, however, be less stable if we do

not know the exact location of the discontinuities. To see this, suppose that f can be written as:

(63)
$$f(x) = \begin{cases} g(x), & x \le z, \\ h(x), & x > z. \end{cases}$$

Further, we define:

(64)
$$\ell = \max\{j | x_j \le z, \ j = 0, \ldots, N-1\},$$

so that the interval $[x_\ell, x_{\ell+1}]$ contains z. Placing the discontinuity on the grid would result in the same order of convergence as the trapezoidal rule itself:

$$\int_a^b f(x)dx \approx T(a, x_\ell, g, \Delta x) + T(x_{\ell+1}, b, h, \Delta x)$$

(65)
$$+ \frac{1}{2}(z - x_\ell)(g(x_\ell) + g(z)) + \frac{1}{2}(x_{\ell+1} - z)(h(z) + h(x_{\ell+1})).$$

A straightforward application of the trapezoidal rule would lead to $T(a, b, f, \Delta x)$. The difference with (65) is

$$\frac{1}{2}\Delta x g(x_\ell) + \frac{1}{2}\Delta x h(x_{\ell+1}) - \frac{1}{2}(z - x_\ell)(g(x_\ell) + g(z)) - \frac{1}{2}(x_{\ell+1} - z)(h(z) + h(x_{\ell+1})).$$

Expanding both g and h around the point of discontinuity z yields:

$$\frac{1}{2}(x_{\ell+1} + x_\ell - 2z)(g(z) - h(z)) + \frac{1}{2}(x_{\ell+1} - z)(z - x_\ell)(g^{(1)}(z) - h^{(1)}(z))$$

$$+ \frac{1}{2}(x_{\ell+1} - z)\sum_{j=1}^{\infty}\frac{1}{j!}g^{(j)}(z) + \frac{1}{2}(z - x_\ell)^j h^{(j)}(z).$$

If f is continuous, but the first derivatives of g and h do not match at z, the order of convergence is still $1/N^2$ since $(x_{\ell+1} - z)(z - x_\ell) \le (\Delta x)^2$. It is clear that, as N changes, the ratio of $(x_{\ell+1} - z)(z - x_\ell)$ to $(\Delta x)^2$ may vary strongly, leading to nonsmooth convergence. If f is discontinuous, i.e., if the values of g and h in z disagree, the order of convergence is $O(1/N)$.

Now suppose that we have computed g and h at grid points x_j, $j = 0, \ldots, N-1$. We know that $g(z) = h(z)$, though we do not know the exact location of z. All we know is that it is contained in $[x_\ell, x_{\ell+1}]$. This is a situation we encounter in the pricing of Bermudan options, as outlined in section 4.4. If we proceed to integrate f on this grid, we will not obtain smooth convergence. A simple approximation of the discontinuity can, however, be found by assuming a linear relationship between x and $g(x) - h(x)$. This leads to

(66)
$$z \approx \frac{x_{\ell+1}(g(x_\ell - h(x_\ell)) - x_\ell(g(x_{\ell+1}) - h(x_{\ell+1}))}{(g(x_\ell) - h(x_\ell)) - (g(x_{\ell+1}) - h(x_{\ell+1}))} + O\left(\Delta x^2\right),$$

where the error estimate follows from linear interpolation. Now suppose that we shift our grid (and recalculate g and h) such that either x_ℓ or $x_{\ell+1}$ coincides with this approximation of z and redo the numerical integration. It is easy to see that smooth convergence will be restored, as the contribution of the error term in (66) to the error term in (65) will be of $O\left(\Delta x^3\right)$. Note that, if we use higher-order Newton–Cotes rules, a higher-order interpolation step will be required.

Appendix C. Comparison of CONV with PIDE methods. In this section we will compare the speed and accuracy of the CONV method to two PIDE schemes, one for the VG model [5] and one recent scheme for Kou's model [41]. An advantage of the CONV method over various PIDE schemes is that it is flexible with respect to the choice of model, whereas the integral term in PIDEs typically requires a very careful treatment, for example, due to its weakly singular kernel for infinite activity Lévy models. Furthermore PIDE methods require a relatively fine discretization in the time direction to guarantee an accurate representation of the solution, whereas in the CONV method we require only as many time steps as exercise dates. For Bermudan options with few exercise dates this is advantageous, though for American options it works to our disadvantage.

At the end of the day, however, the only fair comparison is to compare two implementations in the same computer language, on the same CPU, in terms of speed and accuracy. First we compare the PIDE scheme from [5] to our method. Parameters for the problem solved in this section are given by T6-VG in Table 10. Code for the PIDE scheme from [5] was available in Matlab. As we wrote the CONV code in both Matlab and C^{++}, we were able to conduct a fair comparison. We found the CONV code, for large values of N, to be roughly three times as fast as the Matlab code, so we scaled CPU times in Figure 3 accordingly. For a Bermudan option with relatively few exercise dates, the CONV method is a clear winner. The advantage is reduced when pricing American options, as we price these by extrapolating the values of Bermudan options with a relatively large number of exercise dates. Nevertheless, in case of the VG model the CONV method still reaches a higher accuracy given the same computational budget as the PIDE scheme of [5].

TABLE 10
Parameter sets in the numerical experiments.

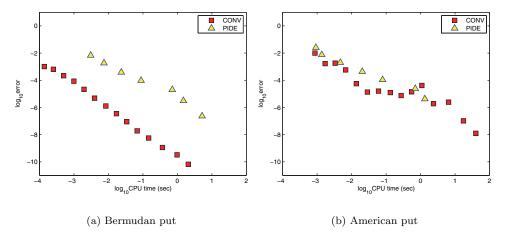| **T6-VG:**	$S(0) = 1,$	$r = 0.1,$	$q = 0,$	$\sigma = 0.282842,$
	$\theta = 0,$	$\nu = 1.$		
T7-Kou:	$S(0) = 100,$	$r = 0.05,$	$q = 0,$	$\sigma = 0.15,$
	$\lambda = 0.1,$	$p = 0.3445,$	$\eta_- = 3.0775,$	$\eta_+ = 3.0465.$

(a) Bermudan put   (b) American put

FIG. 3. *Comparison of the performance for a VG computation of the CONV method with a PIDE solver [5] for Bermudan and American options.*
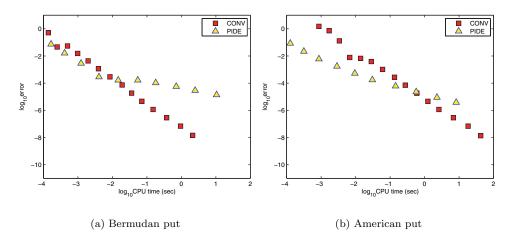
(a) Bermudan put   (b) American put

FIG. 4. *Comparison of performance for jump-diffusion computation (Kou's model) of the CONV method with a PIDE solver* [41] *for Bermudan and American options.*

One other model we consider when comparing the speed and accuracy of the CONV method to PIDE schemes is the Kou model [29]. Its characteristic function equals:

$$(67) \qquad \phi(u) = \exp\left(iu\mu t - \frac{1}{2}u^2\sigma^2 t + iut\lambda\left(\frac{p}{\eta_+ - iu} - \frac{1-p}{\eta_- + iu}\right)\right).$$

In Kou's jump-diffusion model, jumps arrive via a Poisson process with intensity λ. The logarithm of each jump follows a double-exponential density. Toivanen's [41] recent schemes for Kou's model utilize the log-double-exponential form of the jump density to derive efficient recursion formulas for evaluating the integral term in the PIDE. The benefits are clear: The complexity is reduced to $O(MN)$, and in addition his schemes are no longer bound to uniform grids. Therefore it is to be expected that this method outperforms ours, which is not tailored to any specific model. Code for the penalty method from [41] was available in C^{++}. We use parameter set T7-Kou in Table 10. In both examples in Figure 4, the PIDE scheme reaches a higher accuracy than the CONV method for small computational budgets. This is partially due to the fact that for this example we required a very wide grid for the CONV method ($\delta = 100$) in order to converge to the right solution, which implies a lower accuracy for small values of N. For the Bermudan option the CONV method is still competitive, yet for the American option the PIDE scheme is the clear favorite. Though the CONV method appears to converge faster than the penalty method, the PIDE scheme would be the method of choice for practical levels of accuracy.

Although we have compared only to two methods, we believe it is fair to say that the CONV method will compare favorably to most PIDE schemes for the pricing of Bermudan options under Lévy models. There will, however, always be special cases, such as Kou's jump-diffusion model, for which one can design highly efficient PIDE schemes that are faster and more accurate than the CONV method.

REFERENCES

[1] L. Andersen and J. Andreasen, *Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing*, Rev. Deriv. Res., 4 (2000), pp. 231–262.

[2] A. D. Andricopoulos, M. Widdicks, P. W. Duck, and D. P. Newton, *Universal option valuation using quadrature*, J. Financial Economics, 67 (2003), pp. 447–471.

[3] J. Abate and W. Whitt, *The Fourier-series method for inverting transforms of probability distributions*, Queueing Syst., 10 (1992), pp. 5–88.

[4] A. Almendral and C. W. Oosterlee, *Accurate evaluation of European and American options under the CGMY process*, SIAM J. Sci. Comput., 29 (2007), pp. 93–117.

[5] A. Almendral and C. W. Oosterlee, *On American options under the variance gamma process*, Appl. Math. Finance, 14 (2007), pp. 131–152.

[6] S. I. Boyarchenko and S. Z. Levendorskiĭ, *Non-Gaussian Merton-Black-Scholes Theory*, Adv. Ser. Stat. Sci. Appl. Probab. 9, World Scientific, River Edge, NJ, 2002.

[7] M. Broadie and Y. Yamamoto, *A double-exponential fast Gauss transform algorithm for pricing discrete path-dependent options*, Oper. Res., 53 (2005), pp. 764–779.

[8] P. P. Carr, H. Geman, D. B. Madan, and M. Yor, *The fine structure of asset returns: An empirical investigation*, J. Business, 75 (2002), pp. 305–332.

[9] P. P. Carr and D. B. Madan, *Option valuation using the fast Fourier transform*, J. Comput. Finance, 2 (1999), pp. 61–73.

[10] C.-C. Chang, S.-L. Chung, and R. C. Stapleton, *Richardson extrapolation technique for pricing American-style options*, J. Futures Markets, 27 (2007), pp. 791–817.

[11] K. Chourdakis, *Switching Lévy models in continuous time: Finite distributions and option pricing*, in Proceedings of the Quantitative Methods in Finance 2005, Sydney, Australia, 2005; available online from gemini.econ.umd.edu/cgi-bin/conference/download.cgi?db_name=QMF2005&paper_id=81.

[12] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman and Hall, Boca Raton, FL, 2004.

[13] H. Dubner and J. Abate, *Numerical inversion of Laplace transforms by relating them to the finite Fourier cosine transform*, J. ACM, 15 (1968), pp. 115–123.

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

[15] D. Duffie, D. Filipovic, and W. Schachermayer, *Affine processes and applications in finance*, Ann. Appl. Probab., 13 (2003), pp. 984–1053.

[16] G. Fan and G. H. Liu, *Fast Fourier transform for discontinuous functions*, IEEE Trans. Antennas and Propagation, 52 (2004), pp. 461–465.

[17] R. Geske and H. Johnson, *The American put valued analytically*, J. Finance, 39 (1984), pp. 1511–1542.

[18] J. Gil-Pelaez, *Note on the inverse theorem*, Biometrika, 37 (1951), pp. 481–482.

[19] J. Gurland, *Inversion formulae for the distribution of ratios*, Ann. Math. Stat., 19 (1948), pp. 228–237.

[20] Y. d'Halluin, P. A. Forsyth, and G. Labahan, *A penalty method for American options with jump diffusion processes*, Numer. Math., 97 (2004), pp. 321–352.

[21] S. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev. Financ. Stud., 6 (1993), pp. 327–343.

[22] D. J. Higham, *An Introduction to Financial Option Valuation*, Cambridge University Press, London, 2004.

[23] A. Hirsa and D. B. Madan, *Pricing American options under variance gamma*, J. Comput. Finance, 7, 2004.

[24] S. Howison, *A matched asymptotic expansions approach to continuity corrections for discretely sampled options. Part* 2: *Bermudan options*, Appl. Math. Finance, 14 (2007), pp. 91–104.

[25] Z. Hu, J. Kerkhof, P. McCloud, and J. Wackertapp, *Cutting edges using domain integration*, Risk, 19 (2006), pp. 95–99.

[26] P. Hunt, J. Kennedy, and A. A. J. Pelsser, *Markov-functional interest rate models*, Finance Stoch., 4 (2000), pp. 391–408.

[27] P. DEN ISEGER, *Numerical transform inversion using Gaussian quadrature*, Probab. Engrg. Inform. Sci., 20 (2006), pp. 1–44.

[28] E. KËLLEZI AND N. WEBBER, *Valuing Bermudan options when asset returns are Lévy processes*, Quant. Finance, 4 (2004), pp. 87–1000.

[29] S. G. KOU, *A jump diffusion model for option pricing*, Management Sci., 48 (2002), pp. 1086–1101.

[30] R. LEE, *Option pricing by transform methods: Extensions, unification, and error control*, J. Comput. Finance, 7 (2004), pp. 51–86.

[31] C. C. W. LEENTVAAR AND C. W. OOSTERLEE, *Multi-asset option pricing using a parallel Fourier-based technique*, J. Comput. Appl. Math., to appear.

[32] A. LEWIS, *A Simple Option Formula for General Jump-Diffusion and Other Exponential Lévy Processes*, SSRN working paper, 2001; also available online from http://ssrn.com/abstract=282110.

[33] R. LORD AND C. KAHL, *Optimal Fourier inversion in semi-analytical option pricing*, J. Comput. Finance, 10 (2007), pp. 1–30.

[34] D. B. MADAN, P. P. CARR, AND E. C. CHANG, *The variance gamma process and option pricing*, Eur. Finance Rev., 2 (1998), pp. 79–105.

[35] R. A. MALLER, D. H. SOLOMON, AND A. SZIMAYER, *A multinomial approximation for American option prices in Lévy process models*, Math. Finance, 16 (2006), pp. 613–633.

[36] A. M. MATACHE, P. A. NITSCHE, AND C. SCHWAB, *Wavelet Galerkin pricing of American options on Lévy driven assets*, Quantitative Finance, 5 (2005), pp. 403–424.

[37] S. RAIBLE, *Lévy Processes in Finance: Theory, Numerics and Empirical Facts*, Ph.D. thesis, Institute für Math. Stochastik, Albert-Ludwigs-Universität Freiburg, 2000.

[38] E. REINER, *Convolution methods for path-dependent options*, in Proceedings of the Financial Mathematics workshop, IPAM UCLA, 2001; available online from http://www.ipam.ucla.edu/publications/fm2001/fm2001_4272.pdf.

[39] K.-I. SATO, *Basic results on Lévy processes*, in Lévy Processes, Birkhäuser Boston, Boston, MA, 2001, pp. 3–37.

[40] C. O'SULLIVAN, *Path Dependent Option Pricing under Levy Processes*, EFA 2005 Moscow Meetings Paper; available online from http://ssrn.com/abstract=673424, 2005.

[41] J. TOIVANEN, *Numerical valuation of European and American options under Kou's jump-diffusion model*, SIAM J. Sci. Comput., to appear, 2008.

[42] I. WANG, J. W. WAN, AND P. FORSYTH, *Robust numerical valuation of European and American options under the CGMY process*, J. Comput. Finance, 10 (2007), pp. 31–70.

[43] P. WILMOTT, J. DEWYNNE, AND S. HOWISON, *Option Pricing*, Financial Press, Oxford, 1993.