J. Math. Anal. Appl. 450 (2017) 929–953

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

# Pricing Bermudan options under local Lévy models with default

A. Borovykh [a], A. Pascucci [a,*], C.W. Oosterlee [b,c]

[a] Dipartimento di Matematica, Università di Bologna, Bologna, Italy
[b] Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
[c] Delft University of Technology, Delft, The Netherlands

A R T I C L E   I N F O

*Article history:*
Received 27 April 2016
Available online 27 January 2017
Submitted by H.-M. Yin

*Keywords:*
Bermudan option
Local Lévy model
Defaultable asset
Asymptotic expansion
Fourier-cosine expansion

A B S T R A C T

We consider a defaultable asset whose risk-neutral pricing dynamics are described by an exponential Lévy-type martingale. This class of models allows for a local volatility, local default intensity and a locally dependent Lévy measure. We present a pricing method for Bermudan options based on an analytical approximation of the characteristic function combined with the COS method. Due to a special form of the obtained characteristic function the price can be computed using a fast Fourier transform-based algorithm resulting in a fast and accurate calculation. The Greeks can be computed at almost no additional computational cost. Error bounds for the approximation of the characteristic function as well as for the total option price are given.



## 1. Introduction

In financial mathematics, the fast and accurate pricing of financial derivatives is an important branch of research. Depending on the type of financial derivative, the mathematical task is essentially the computation of integrals, and this sometimes needs to be performed in a recursive way in a time-wise direction. For many stochastic processes that model the financial assets, these integrals can be most efficiently computed in the Fourier domain. However, for some relevant and recent stochastic models the Fourier domain computations are not at all straightforward, as these computations rely on the availability of the characteristic function of the stochastic process (read: the Fourier transform of the transitional probability distribution), which is not known. This is especially true for state-dependent asset price processes, and for asset processes that include the notion of default in their definition. With the derivations and techniques in the present paper we make available the highly efficient pricing of so-called Bermudan options to the above mentioned classes of state-dependent asset dynamics, including jumps in asset prices and the possibility of default. In this

* Corresponding author.
*E-mail addresses:* anastasia.borovykh2@unibo.it (A. Borovykh), andrea.pascucci@unibo.it (A. Pascucci), c.w.oosterlee@cwi.nl (C.W. Oosterlee).

http://dx.doi.org/10.1016/j.jmaa.2017.01.071

sense, the class of asset models for which Fourier option pricing is highly efficient increases by the contents of the present paper. Essentially, we approximate the characteristic function by an advanced Taylor-based expansion in such a way that the resulting characteristic function exhibits favorable properties for the pricing methods.

Fourier methods have often been among the winners in option pricing competitions such as BENCHOP [16]. In [5], a Fourier method called the COS method, as introduced in [4], was extended to the pricing of Bermudan options. The computational efficiency of the method was based on a specific structure of the characteristic function allowing to use the fast Fourier transform (FFT) for calculating the continuation value of the option. Fourier methods can readily be applied to solving problems under asset price dynamics for which the characteristic function is available. This is the case for exponential Lévy models, such as the Merton model developed in [13], the Variance-Gamma model developed in [12], but also for the Heston model [6]. However, in the case of local volatility, default and state-dependent jump measures there is no closed form characteristic function available and the COS method cannot be readily applied.

Recently, in [14] the so-called *adjoint expansion method* for the approximation of the characteristic function in local Lévy models is presented. This method is worked out in the Fourier space by considering the adjoint formulation of the pricing problem, that is using a backward parametrix expansion as was also later done in [1]. In this paper we generalize this method to include a defaultable asset whose risk-neutral pricing dynamics are described by an exponential Lévy-type martingale with a state-dependent jump measure, as has also been considered in [11] and in [7].

Having obtained the analytical approximation for the characteristic function we combine this with the COS method for Bermudan options. We show that this analytical formula for the characteristic function still possesses a structure that allows the use of a FFT-based method in order to calculate the continuation value. This results in an efficient and accurate computation of the Bermudan option value and of the Greeks. The characteristic function approximation used in the COS method is already very accurate for the 2nd-order approximation, meaning that the explicit formulas are simple and this makes method easy and quick to implement. We prove error bounds for the 0th- and 1st-order approximation, justifying the accuracy of the method and present a wide range of numerical examples, showing the flexibility, accuracy and speed of the method.

## 2. General framework

We consider a defaultable asset $S$ whose risk-neutral dynamics are given by:

$$\begin{aligned} S_t &= \mathbb{1}_{\{t<\zeta\}} e^{X_t}, \\ dX_t &= \mu(t, X_t)dt + \sigma(t, X_t)dW_t + \int_{\mathbb{R}} d\tilde{N}_t(t, X_{t-}, dz)z, \\ d\tilde{N}_t(t, X_{t-}, dz) &= dN_t(t, X_{t-}, dz) - \nu(t, X_{t-}, dz)dt, \\ \zeta &= \inf\{t \geq 0 : \int_0^t \gamma(s, X_s)ds \geq \varepsilon\}, \end{aligned} \tag{2.1}$$

where $\tilde{N}_t(t, x, dz)$ is a compensated random measure with state-dependent Lévy measure $\nu(t, x, dz)$. The default time $\zeta$ of $S$ is defined in a canonical way as the first arrival time of a doubly stochastic Poisson process with local intensity function $\gamma(t, x) \geq 0$, and $\varepsilon \sim \text{Exp}(1)$ and is independent of $X$. Thus the model features:

- a local volatility function $\sigma(t,x)$;
- a local Lévy measure: jumps in $X$ arrive with a state-dependent intensity described by the local Lévy measure $\nu(t,x,dz)$. The jump intensity and jump distribution can thus change depending on the value of $x$. A state-dependent Lévy measure is an important feature because it allows to incorporate stochastic jump-intensity into the modeling framework;
- a local default intensity $\gamma(t,x)$: the asset $S$ can default with a state-dependent default intensity.

This way of modeling default is also considered in a diffusive setting in [3] and for exponential Lévy models in [2].

We define the filtration of the market observer to be $\mathcal{G} = \mathcal{F}^X \vee \mathcal{F}^D$, where $\mathcal{F}^X$ is the filtration generated by $X$ and $\mathcal{F}^D_t := \sigma(\{\zeta \le u\}, u \le t)$, for $t \ge 0$, is the filtration of the default. We assume

$$\int_{\mathbb{R}} e^{|z|}\nu(t,x,dz) < \infty,$$

and by imposing that the discounted asset price $\tilde{S}_t := e^{-rt}S_t$ is a $\mathcal{G}$-martingale, we get the following restriction on the drift coefficient:

$$\mu(t,x) = \gamma(t,x) + r - \frac{\sigma^2(t,x)}{2} - \int_{\mathbb{R}} \nu(t,x,dz)(e^z - 1 - z).$$

Is it well-known (see, for instance, [8, Section 2.2]) that the price $V$ of a European option with maturity $T$ and payoff $\Phi(S_T)$ is given by

$$V_t = \mathbb{1}_{\{\zeta>t\}} e^{-r(T-t)} E\left[e^{-\int_t^T \gamma(s,X_s)ds}\phi(X_T)|X_t\right], \qquad t \le T, \tag{2.2}$$

where $\phi(x) = \Phi(e^x)$. Thus, in order to compute the price of an option, we must evaluate functions of the form

$$u(t,x) := E\left[e^{-\int_t^T \gamma(s,X_s)ds}\phi(X_T)|X_t = x\right]. \tag{2.3}$$

Under standard assumptions, $u$ can be expressed as the classical solution of the following Cauchy problem

$$\begin{cases} Lu(t,x) = 0, & t \in [0,T[, \ x \in \mathbb{R}, \\ u(T,x) = \phi(x), & x \in \mathbb{R}, \end{cases}$$

where $L$ is the integro-differential operator

$$\begin{aligned} Lu(t,x) &= \partial_t u(t,x) + r\partial_x u(t,x) + \gamma(t,x)(\partial_x u(t,x) - u(t,x)) + \frac{\sigma^2(t,x)}{2}(\partial_{xx} - \partial_x)u(t,x) \\ &\quad - \int_{\mathbb{R}} \nu(t,x,dz)(e^z - 1 - z)\partial_x u(t,x) + \int_{\mathbb{R}} \nu(t,x,dz)(u(t,x+z) - u(t,x) - z\partial_x u(t,x)). \end{aligned} \tag{2.4}$$

The function $u$ in (2.3) can be represented as an integral with respect to the transition distribution of the defaultable log-price process $\log S$:

$$u(t,x) = \int_{\mathbb{R}} \phi(y)\Gamma(t,x;T,dy). \tag{2.5}$$

Here we notice explicitly that $\Gamma(t,x;T,dy)$ is not necessarily a standard probability measure because its integral over $\mathbb{R}$ can be strictly less than one; nevertheless, with a slight abuse of notation, we say that its Fourier transform

$$\hat{\Gamma}(t,x;T,\xi) := \mathcal{F}(\Gamma(t,x;T,\cdot))(\xi) := \int_{\mathbb{R}} e^{i\xi y}\Gamma(t,x;T,dy), \qquad \xi \in \mathbb{R},$$

is the characteristic function of $\log S$.

### 2.1. Adjoint expansion of the characteristic function

In this section we generalize the results in [14] to our framework and develop an expansion of the coefficients

$$a(t,x) := \frac{\sigma^2(t,x)}{2}, \qquad \gamma(t,x), \qquad \nu(t,x,dz),$$

around some point $\bar{x}$. The coefficients $a(t,x)$, $\gamma(t,x)$ and $\nu(t,x,dz)$ are assumed to be continuously differentiable with respect to $x$ up to order $N \in \mathbb{N}$.

From now on for simplicity we assume that the coefficients are independent of $t$ (see Remark 2.2 for the general case). First we introduce the $n$th-order approximation of $L$ in (2.4):

$$\begin{aligned} L_n = L_0 + \sum_{k=1}^{n} \Big( &(x-\bar{x})^k a_k(\partial_{xx} - \partial_x) + (x-\bar{x})^k \gamma_k \partial_x - (x-\bar{x})^k \gamma_k \\ &- \int_{\mathbb{R}} (x-\bar{x})^k \nu_k(dz)(e^z - 1 - z)\partial_x + \int_{\mathbb{R}} (x-\bar{x})^k \nu_k(dz)(e^{z\partial_x} - 1 - z\partial_x) \Big), \end{aligned}$$

where

$$L_0 = \partial_t + r\partial_x + a_0(\partial_{xx} - \partial_x) + \gamma_0\partial_x - \gamma_0 - \int_{\mathbb{R}} \nu_0(dz)(e^z - 1 - z)\partial_x + \int_{\mathbb{R}} \nu_0(dz)(e^{z\partial_x} - 1 - z\partial_x),$$

and

$$a_k = \frac{\partial_x^k a(\bar{x})}{k!}, \qquad \gamma_k = \frac{\partial_x^k \gamma(\bar{x})}{k!}, \qquad \nu_k(dz) = \frac{\partial_x^k \nu(\bar{x},dz)}{k!}, \qquad k \geq 0.$$

The basepoint $\bar{x}$ is a constant parameter which can be chosen freely. In general the simplest choice is $\bar{x} = x$ (the value of the underlying at initial time $t$): we will see that in this case the formulas for the Bermudan option valuation are simplified.

Let us assume for a moment that $L_0$ has a fundamental solution $G^0(t,x;T,y)$ that is defined as the solution of the Cauchy problem

$$\begin{cases} L_0 G^0(t,x;T,y) = 0 & t \in [0,T[, \ x \in \mathbb{R}, \\ G^0(T,\cdot;T,y) = \delta_y. \end{cases}$$

In this case we define the $n$th-order approximation of $\Gamma$ as

$$\Gamma^{(n)}(t,x;T,y) = \sum_{k=0}^{n} G^k(t,x;T,y),$$

where, for any $k \geq 1$ and $(T, y)$, $G^k(\cdot, \cdot; T, y)$ is defined recursively through the following Cauchy problem

$$\begin{cases} L_0 G^k(t,x;T,y) = -\sum\limits_{h=1}^{k} (L_h - L_{h-1}) G^{k-h}(t,x;T,y) & t \in [0,T[,\ x \in \mathbb{R}, \\ G^k(T,x;T,y) = 0, & x \in \mathbb{R}. \end{cases}$$

Notice that

$$\begin{aligned} L_h - L_{h-1} &= (x-\bar{x})^h a_h (\partial_{xx} - \partial_x) + (x-\bar{x})^h \gamma_h \partial_x - (x-\bar{x})^h \gamma_h \\ &\quad - \int_{\mathbb{R}} (x-\bar{x})^h \nu_h(dz)(e^z - 1 - z)\partial_x + \int_{\mathbb{R}} (x-\bar{x})^h \nu_h(dz)(e^{z\partial_x} - 1 - z\partial_x). \end{aligned}$$

Correspondingly, the $n$th-order approximation of the characteristic function $\hat{\Gamma}$ is defined to be

$$\hat{\Gamma}^{(n)}(t,x;T,\xi) = \sum_{k=0}^{n} \mathcal{F}\left(G^k(t,x;T,\cdot)\right)(\xi) := \sum_{k=0}^{n} \hat{G}^k(t,x;T,\xi), \qquad \xi \in \mathbb{R}. \tag{2.6}$$

Now we remark that the operator $L$ acts on $(t,x)$ while the characteristic function is a Fourier transform taken with respect to $y$: in order to take advantage of such a transformation, in the following theorem we characterize $\hat{\Gamma}^{(n)}$ in terms of the Fourier transform of the adjoint operator $\tilde{L} = \tilde{L}^{(T,y)}$ of $L$, acting on $(T, y)$.

**Theorem 2.1** *(Dual formulation). For any $(t,x) \in ]0,T] \times \mathbb{R}$, the function $G^0(t,x;\cdot,\cdot)$ is defined through the following dual Cauchy problem*

$$\begin{cases} \tilde{L}_0^{(T,y)} G^0(t,x;T,y) = 0 & T > t,\ y \in \mathbb{R}, \\ G^0(T,x;T,\cdot) = \delta_x, \end{cases} \tag{2.7}$$

*where*

$$\tilde{L}_0^{(T,y)} = -\partial_T - r\partial_y + a_0(\partial_{yy} + \partial_y) - \gamma_0\partial_y - \gamma_0 + \int_{\mathbb{R}} \nu_0(dz)(e^z - 1 - z)\partial_y + \int_{\mathbb{R}} \bar{\nu}_0(dz)(e^{z\partial_y} - 1 - z\partial_y).$$

*Moreover, for any $k \geq 1$, the function $G^k(t,x;\cdot,\cdot)$ is defined through the dual Cauchy problem as follows:*

$$\begin{cases} \tilde{L}_0^{(T,y)} G^k(t,x;T,y) = -\sum\limits_{h=1}^{k} \left(\tilde{L}_h^{(T,y)} - \tilde{L}_{h-1}^{(T,y)}\right) G^{k-h}(t,x;T,y) & T > t,\ y \in \mathbb{R}, \\ G^k(T,x;T,y) = 0 & y \in \mathbb{R}, \end{cases} \tag{2.8}$$

*with*

$$\begin{aligned} \tilde{L}_h^{(T,y)} - \tilde{L}_{h-1}^{(T,y)} &= a_h h(h-1)(y-\bar{x})^{h-2} + a_h (y-\bar{x})^{h-1}\left(2h\partial_y + (y-\bar{x})(\partial_{yy} + \partial_y) + h\right) \\ &\quad - \gamma_h h(y-\bar{x})^{h-1} - \gamma_h (y-\bar{x})^h (\partial_y + 1) \\ &\quad + \int_{\mathbb{R}} \nu_h(dz)(e^z - 1 - z)\left(h(y-\bar{x})^{h-1} + (y-\bar{x})^h \partial_y\right) \\ &\quad + \int_{\mathbb{R}} \bar{\nu}_h(dz)\left((y+z-\bar{x})^h e^{z\partial_y} - (y-\bar{x})^h - z\left(h(y-\bar{x})^{h-1} - (y-\bar{x})^h \partial_y\right)\right), \end{aligned}$$

*where in defining the adjoint of the operator we use the notation*

$$e^{z\partial_y}f(y) := \sum_{n=0}^{\infty}\frac{z^n}{n!}\partial_y^n f(y) = f(y+z).$$

Notice that the adjoint Cauchy problems (2.7) and (2.8) admit a solution in the Fourier space and can be solved explicitly; in fact, we have

$$\mathcal{F}\left(\tilde{L}_0^{(T,\cdot)}G^k(t,x;T,\cdot)\right)(\xi) = \psi(\xi)\hat{G}^k(t,x;T,\xi) - \partial_T\hat{G}^k(t,x;T,\xi),$$

where $\psi(\xi)$ is the characteristic exponent of the Lévy process with coefficients $\gamma_0$, $a_0$ and $\nu_0(dz)$, that is

$$\psi(\xi) = i\xi(r+\gamma_0) + a_0(-\xi^2 - i\xi) - \gamma_0 - \int_{\mathbb{R}}\nu_0(dz)(e^z - 1 - z)i\xi + \int_{\mathbb{R}}\nu_0(dz)(e^{iz\xi} - 1 - iz\xi).$$

Thus the solution (in the Fourier space) to problems (2.7) and (2.8) is given by

$$\begin{aligned}
&\hat{G}^0(t,x;T,\xi) = e^{i\xi x+(T-t)\psi(\xi)},\\
&\hat{G}^k(t,x;T,\xi) = -\int_t^T e^{\psi(\xi)(T-s)}\mathcal{F}\left(\sum_{h=1}^{k}\left(\tilde{L}_h^{(s,\cdot)} - \tilde{L}_{h-1}^{(s,\cdot)}\right)G^{k-h}(t,x;s,\cdot)\right)(\xi)ds, \qquad k\geq 1.
\end{aligned}\tag{2.9}$$

Now we consider the general framework and in particular we drop the assumption on the existence of the fundamental solution of $L_0$: in this case, we define the $n$th-order approximation of the characteristic function $\hat{\Gamma}$ as in (2.6), with $\hat{G}^k$ given by (2.9). We also notice that

$$\begin{aligned}
&\mathcal{F}\left(\left(\tilde{L}_h^{(s,\cdot)} - \tilde{L}_{h-1}^{(s,\cdot)}\right)u(s,\cdot)\right)(\xi) =\\
&\quad\left(a_h h(h-1)(-i\partial_\xi - \bar{x})^{h-2} + a_h(-i\partial_\xi - \bar{x})^{h-1}\left(-2hi\xi + (-i\partial_\xi - \bar{x})(-\xi^2 - i\xi) + h\right)\right)\hat{u}(s,\xi)\\
&\quad - \left(\gamma_h h(-i\partial_\xi - \bar{x})^{h-1} - \gamma_h(-i\partial_\xi - \bar{x})^h(i\xi - 1)\right)\hat{u}(s,\xi)\\
&\quad + \int_{\mathbb{R}}\nu_h(dz)(e^z - 1 - z)\left(h(-i\partial_\xi - \bar{x})^{h-1} - (-i\partial_\xi - \bar{x})^h i\xi\right)\hat{u}(s,\xi)\\
&\quad + \int_{\mathbb{R}}\nu_h(dz)\left((-i\partial_y - z - \bar{x})^h e^{i\xi z} - (-i\partial_y - \bar{x})^h + z\left(h(-i\partial_\xi - \bar{x})^{h-1} - (-i\partial_\xi - \bar{x})^h i\xi\right)\right)\hat{u}(s,\xi).
\end{aligned}$$

**Remark 2.2.** In case the coefficients $\gamma$, $\sigma$, $\nu$ depend on time, the solutions to the Cauchy problems are similar:

$$\begin{aligned}
&\hat{G}^0(t,x;T,\xi) = e^{i\xi x}e^{\int_t^T \psi(s,\xi)ds},\\
&\hat{G}^k(t,x;T,\xi) = -\int_t^T e^{\int_s^T \psi(\tau,\xi)d\tau}\mathcal{F}\left(\sum_{h=1}^{k}\left(\tilde{L}_h^{(s,\cdot)}(s) - \tilde{L}_{h-1}^{(s,\cdot)}(s)\right)G^{k-h}(t,x;s,\cdot)\right)(\xi)ds,
\end{aligned}$$

with

$$\psi(s,\xi) = i\xi(r+\gamma_0(s)) + a_0(s)(-\xi^2 - i\xi) - \int_{\mathbb{R}} \nu_0(s,dz)(e^z - 1 - z)i\xi + \int_{\mathbb{R}} \nu_0(s,dz)(e^{iz\xi} - 1 - iz\xi),$$

$$\begin{aligned}\tilde{L}_h^{(s,y)}(s) - \tilde{L}_{h-1}^{(s,y)}(s) &= a_h(s)h(h-1)(y-\bar{x})^{h-2} + a_h(s)(y-\bar{x})^{h-1}\left(2h\partial_y + (y-\bar{x})(\partial_{yy}+\partial_y) + h\right)\\ &\quad - \gamma_h(s)h(y-\bar{x})^{h-1} - \gamma_h(s)(y-\bar{x})^h\left(\partial_y + 1\right)\\ &\quad + \int_{\mathbb{R}} \nu_h(s,dz)(e^z - 1 - z)\left(h(y-\bar{x})^{h-1} + (y-\bar{x})^h\partial_y\right)\\ &\quad + \int_{\mathbb{R}} \bar{\nu}_h(s,dz)\left((y+z-\bar{x})^h e^{z\partial_y} - (y-\bar{x})^h - z\left(h(y-\bar{x})^{h-1} - (y-\bar{x})^h\partial_y\right)\right).\end{aligned}$$

From these results one can already see that the dependency on $x$ comes in through $e^{i\xi x}$ and after taking derivatives the dependency on $x$ will take the form $(x-\bar{x})^m e^{i\xi x}$: this fact will be crucial in our analysis.

**Example 2.3.** To see the above dependency more explicitly for the second-order approximation of the characteristic function we consider, for ease of notation, a simplified model: a one-dimensional local Lévy model where the log-price solves the SDE

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + \int_{\mathbb{R}} d\tilde{N}_t(dz)z. \tag{2.10}$$

This model is a simplification of the original model, since we consider only a local volatility function, and no local default or state-dependent Lévy measure. Thus only a Taylor expansion of the local volatility coefficient is used. However, the dependency that we will see generalizes in the same way to the local default and state-dependent measure. By the martingale condition we have

$$\mu(x) = r - a(x) - \int_{\mathbb{R}} \nu(dz)(e^z - 1),$$

and therefore the Kolmogorov operator of (2.10) reads

$$\begin{aligned}Lu(t,x) &= \partial_t u(t,x) + r\partial_x u(t,x) + a(t,x)(\partial_{xx} - \partial_x)u(t,x)\\ &\quad - \int_{\mathbb{R}} \nu(dz)(e^z - 1) + \int_{\mathbb{R}} \nu(dz)\left(u(t,x+z) - u(t,x)\right).\end{aligned}$$

In this case, we have the following explicit approximation formulas for the characteristic function $\hat{\Gamma}(t,x;T,\xi)$:

$$\hat{\Gamma}(t,x;T,\xi) \;\approx\; \hat{\Gamma}^{(n)}(t,x;T,\xi) := e^{i\xi x + (T-t)\psi(\xi)} \sum_{k=0}^{n} \hat{F}^k(t,x;T,\xi), \qquad n \geq 0, \tag{2.11}$$

with

$$\psi(\xi) = ir\xi - a_0(\xi^2 + i\xi) - \int_{\mathbb{R}} \nu(dz)(e^z - 1)i\xi + \int_{\mathbb{R}} \nu(dz)\left(e^{iz\xi} - 1\right),$$

and

$$\hat{F}^k(t,x;T,\xi) = \sum_{h=0}^{k} g_h^{(k)}(T-t,\xi)(x-\bar{x})^h; \tag{2.12}$$

here, for $k = 0, 1, 2$, we have

$$
\begin{aligned}
g_0^{(0)}(s,\xi) &= 1,\\
g_0^{(1)}(s,\xi) &= a_1 s^2(\xi^2 + i\xi)\frac{i}{2}\psi'(\xi),\\
g_1^{(1)}(s,\xi) &= -\, a_1 s(\xi^2 + i\xi),\\
g_0^{(2)}(s,\xi) &= \frac{1}{2}s^2 a_2 \xi(i+\xi)\psi''(\xi) - \frac{1}{6}s^3\xi(i+\xi)(a_1^2(i+2\xi)\psi'(\xi) - 2a_2\psi'(\xi)^2 + a_1^2\xi(i+\xi)\psi''(\xi))\\
&\quad - \frac{1}{8}s^4 a_1^2\xi^2(i+\xi)^2\psi'(\xi)^2,\\
g_1^{(2)}(s,\xi) &= \frac{1}{2}s^2\xi(i+\xi)(a_1^2(1-2i\xi) + 2ia_2\psi''(\xi)) - \frac{1}{2}s^3 i a_1^2\xi^2(i+\xi)^2\psi''(\xi),\\
g_2^{(2)}(s,\xi) &= -\, a_2 s\xi(i+\xi) + \frac{1}{2}s^2 a_1^2\xi^2(i+\xi)^2.
\end{aligned}
$$

Using the notation from above, we can write in the same way the approximation formulas for the general case. Here we present the results for $k = 0, 1$, since higher-order formulas are too long to include. For the full formula we refer to Appendix B. We have:

$$
\begin{aligned}
g_0^{(0)}(s,\xi) =&1,\\
g_0^{(1)}(s,\xi) =&\frac{i}{2}a_1 s^2(\xi^2 + i\xi)\psi'(\xi) + \frac{1}{2}\gamma_1 s^2(i+\xi)\psi'(\xi) - \frac{1}{2}\int_{\mathbb{R}}\nu_1(dz)(e^z - 1 - z)s^2\xi\psi'(\xi)\\
&- \frac{1}{2}\int_{\mathbb{R}}\nu_1(dz)(ie^{i\xi z} - i + \xi z)s^2\psi'(\xi),\\
g_1^{(1)}(s,\xi) =& - a_1 s(\xi^2 + i\xi) + \gamma_1 s i(i+\xi) - \int_{\mathbb{R}}\nu_1(dz)(e^z - 1 - z)s\xi i\\
&+ \int_{\mathbb{R}}\nu_1(dz)(e^{i\xi z} - 1 - \xi i z)s.
\end{aligned}
\tag{2.13}
$$

**Remark 2.4.** From (2.11)–(2.12) we clearly see that the approximation of order $n$ is a function of the form

$$
\hat{\Gamma}^{(n)}(t,x;T,\xi) := e^{i\xi x}\sum_{k=0}^{n}(x-\bar{x})^k g_{n,k}(t,T,\xi), \tag{2.14}
$$

where the coefficients $g_{n,k}$, with $0 \le k \le n$, depend only on $t, T$ and $\xi$, but not on $x$. The approximation formula can thus always be split into a sum of products of functions depending only on $\xi$ and functions that are linear combinations of $(x-\bar{x})^m e^{i\xi x}$, $m \in \mathbb{N}_0$.

## 3. Bermudan option valuation

A Bermudan option is a financial contract in which the holder can exercise at a predetermined finite set of exercise moments prior to maturity, and the holder of the option receives a payoff when exercising. Consider a Bermudan option with a set of $M$ exercise moments $\{t_1, ..., t_M\}$, with $0 \le t_1 < t_2 < \cdots < t_M = T$. When the option is exercised at time $t_m$ the holder receives the payoff $\Phi(t_m, S_{t_m})$. Recalling (2.2), the no-arbitrage value of the Bermudan option at time $t$ is

$$v\left(t, X_t\right) = \mathbb{1}_{\{\zeta>t\}} \sup_{\tau\in\mathcal{T}_t} E\left[e^{-\int_t^\tau (r+\gamma(s,X_s))ds}\phi(\tau, X_\tau)|X_t\right],$$

where $\phi(t,x) = \Phi(t,e^x)$ and $\mathcal{T}_t$ is the set of all $\mathcal{G}$-stopping times taking values in $\{t_1,...,t_M\} \cap [t,T]$. For a Bermudan put option with strike price $K$, we simply have $\phi(t,x) = (K-e^x)^+$. By the dynamic programming approach, the option value can be computed by a backward recursion: we denote by

$$\tilde{v}(t_M,x) = \phi(t_M,x), \qquad v(t_M,x) = \mathbb{1}_{\{\zeta>t_M\}}\phi(t_M,x),$$

the pre-default value and the true value of the payoff respectively; moreover, we set

$$\begin{cases} \tilde{c}(t,x) = E\left[e^{-\int_t^{t_m}(r+\gamma(s,X_s))ds}\tilde{v}(t_m,X_{t_m})|X_t=x\right], & t\in[t_{m-1},t_m[ \\ \tilde{v}(t_{m-1},x) = \max\{\phi(t_{m-1},x),\tilde{c}(t_{m-1},x)\}, & m\in\{2,\dots,M\}. \end{cases} \tag{3.15}$$

In the above notation $\tilde{c}(t,x)$ is the so-called *pre-default continuation* value: the "true" continuation value is given by

$$c(t,x) = \mathbb{1}_{\{\zeta>t\}}\tilde{c}(t,x) = E\left[e^{-r(t_m-t)}\mathbb{1}_{\{\zeta>t_m\}}\tilde{v}(t_m,X_{t_m})|\mathcal{G}_t\right].$$

The option value is defined as

$$v(t,m) = \begin{cases} \mathbb{1}_{\{\zeta>t_{m-1}\}}\tilde{v}(t_{m-1},x), & \text{for } t=t_{m-1}, \\ c(t,m), & \text{for } t\in]t_{m-1},t_m[\text{, and, if } t_1>0\text{, also for } t\in[0,t_1[. \end{cases}$$

We notice explicitly that the pricing algorithm (3.15) requires, at each step, to evaluate explicitly only the *pre-default* continuation value: this makes the algorithm feasible since the item $\mathbb{1}_{\{\zeta>t\}}$ cannot be priced explicitly.

**Remark 3.5.** Since the payoff of a call option grows exponentially with the log-stock price, this may introduce significant cancelation errors for large domain sizes. For this reason we price put options only using our approach and we employ the well-known put-call parity to price calls via puts. This is a rather standard argument (see, for instance, [17]).

### *3.1. An algorithm for pricing Bermudan put options*

The COS method proposed by [5] is based on the insight that the Fourier-cosine series coefficients of $\Gamma(t,x;T,dy)$ (and therefore also of option prices) are closely related to the characteristic function of the underlying process, namely the following relationship holds:

$$\int_a^b e^{i\frac{k\pi}{b-a}}\Gamma(t,x;T,dy) \approx \hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right).$$

The COS method provides a way to calculating expected values (integrals) of the form

$$v(t,x) = \int_{\mathbb{R}} \phi(T,y)\Gamma(t,x;T,dy),$$

and it consists of three approximation steps:

1. In the first step we truncate the infinite integration range to $[a, b]$ to obtain approximation $v_1$:

$$v_1(t,x) := \int_a^b \phi(T,y)\Gamma(t,x;T,dy).$$

We assume this can be done due to the rapid decay of the distribution at infinity.

2. In the second step we replace the distribution with its cosine expansion and we get

$$v_1(t,x) := \frac{b-a}{2} \sum_{k=0}^{\infty}{}' A_k(t,x;T)V_k(T),$$

where $\sum'$ indicates that the first term in the summation is weighted by one-half and

$$A_k(t,x;T) = \frac{2}{b-a}\int_a^b \cos\left(k\pi\frac{y-a}{b-a}\right)\Gamma(t,x;T,dy),$$

$$V_k(T) = \frac{2}{b-a}\int_a^b \cos\left(k\pi\frac{y-a}{b-a}\right)\phi(T,y)dy,$$

are the Fourier-cosine series coefficients of the distribution and of the payoff function at time $T$ respectively. Due to the rapid decay of the Fourier-cosine series coefficients, we truncate the series summation and obtain approximation $v_2$:

$$v_2(t,x) := \frac{b-a}{2} \sum_{k=0}^{N-1}{}' A_k(t,x;T)V_k(T).$$

3. In the third step we use the fact that the coefficients $A_k$ can be rewritten using the truncated characteristic function:

$$A_k(t,x;T) = \frac{2}{b-a}\mathrm{Re}\left(e^{-ik\pi\frac{a}{b-a}}\int_a^b e^{i\frac{k\pi}{b-a}y}\Gamma(t,x;T,dy)\right).$$

The finite integration range can be approximated as

$$\int_a^b e^{i\frac{k\pi}{b-a}y}\Gamma(t,x;T,dy) \approx \int_{\mathbb{R}} e^{i\frac{k\pi}{b-a}y}\Gamma(t,x;T,dy) = \hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right).$$

Thus in the last step we replace $A_k$ by its approximation:

$$\frac{2}{b-a}\mathrm{Re}\left(e^{-ik\pi\frac{a}{b-a}}\hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right)\right),$$

and obtain approximation $v_3$:

$$v_3(t,x) := \sum_{k=0}^{N-1}{}' \mathrm{Re}\left(e^{-ik\pi\frac{a}{b-a}}\hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right)\right)V_k(T). \tag{3.16}$$

Next we go back to the Bermudan put pricing problem. Remembering that the expected value $c(t,x)$ in (3.15) can be rewritten in integral form as in (2.5), we have

$$c(t,x) = e^{-r(t_m-t)} \int_{\mathbb{R}} v(t_m,y)\Gamma(t,x;t_m,dy), \qquad t \in [t_{m-1},t_m[.$$

Then we use the Fourier-cosine expansion (3.16), so that we get the approximation:

$$\hat{c}(t,x) = e^{-r(t_m-t)} \sideset{}{'}\sum_{k=0}^{N-1} \mathrm{Re}\left(e^{-ik\pi\frac{a}{b-a}}\hat{\Gamma}\left(t,x;t_m,\frac{k\pi}{b-a}\right)\right)V_k(t_m), \qquad t \in [t_{m-1},t_m[ \tag{3.17}$$

$$V_k(t_m) = \frac{2}{b-a}\int_a^b \cos\left(k\pi\frac{y-a}{b-a}\right)\max\{\phi(t_m,y),c(t_m,y)\}dy,$$

with $\phi(t,x) = (K-e^x)^+$.

Next we recover the coefficients $(V_k(t_m))_{k=0,1,...,N-1}$ from $(V_k(t_{m+1}))_{k=0,1,...,N-1}$. To this end, we split the integral in the definition of $V_k(t_m)$ into two parts using the early-exercise point $x_m^*$, which is the point where the continuation value is equal to the payoff, i.e. $c(t_m,x_m^*) = \phi(t_m,x_m^*)$; thus we have

$$V_k(t_m) = F_k(t_m,x_m^*) + C_k(t_m,x_m^*), \qquad m = M-1, M-2, ..., 1,$$

where

$$\begin{aligned} F_k(t_m,x_m^*) &:= \frac{2}{b-a}\int_a^{x_m^*}\phi(t_m,y)\cos\left(k\pi\frac{y-a}{b-a}\right)dy, \\ C_k(t_m,x_m^*) &:= \frac{2}{b-a}\int_{x_m^*}^{b} c(t_m,y)\cos\left(k\pi\frac{y-a}{b-a}\right)dy, \end{aligned} \tag{3.18}$$

and $V_k(t_M) = F_k(t_M,\log K)$.

**Remark 3.6.** Since we have a semi-analytic formula for $\hat{c}(t_m,x)$, we can easily find the derivatives with respect to $x$ and use Newton's method to find the point $x_m^*$ such that $c(t_m,x_m^*) = \phi(t_m,x_m^*)$. A good starting point for the Newton method is $\log K$, since $x_m^* \leq \log K$.

The coefficients $F_k(t_m,x_m^*)$ can be computed analytically using $x_m^* \leq \log K$, so that we have

$$\begin{aligned} F_k(t_m,x_m^*) &= \frac{2}{b-a}\int_a^{x_m^*}(K-e^y)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\ &= \frac{2}{b-a}K\Psi_k(a,x_m^*) - \frac{2}{b-a}\chi_k(a,x_m^*), \end{aligned}$$

where

$$\begin{aligned} \chi_k(a,x_m^*) &= \int_a^{x_m^*} e^y\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\ &= \frac{1}{1+\left(\frac{k\pi}{b-a}\right)^2}\left(e^{x_m^*}\cos\left(k\pi\frac{x_m^*-a}{b-a}\right) - e^a + \frac{k\pi e^{x_m^*}}{b-a}\sin\left(k\pi\frac{x_m^*-a}{b-a}\right)\right), \end{aligned}$$

1. For $k = 0, 1, ..., N - 1$:
   - At time $t_M$, the coefficients are exact: $V_k(t_M) = F_k(t_M, \log K)$, as in (3.18).
2. For $m = M - 1$ to 1:
   - Determine the early-exercise point $x_m^*$ using Newton's method;
   - Compute $\hat{V}_k(t_m)$ using formula $\hat{V}_k(t_m) := F_k(t_m, x_m^*) + \hat{C}_k(t_m, x_m^*)$, (3.18) and (3.19). Use an FFT for the continuation value (see Section 3.2).
3. Final step: using $\hat{V}_k(t_1)$ determine the option price $\hat{v}(0, x) = \hat{c}(0, x)$ using (3.17).

**Fig. 1.** Algorithm 3.1: Bermudan option valuation.

$$\Psi_k(a, x_m^*) = \int_a^{x_m^*} \cos\left(k\pi \frac{y-a}{b-a}\right) dy = \begin{cases} \frac{b-a}{k\pi} \sin\left(k\pi \frac{x_m^*-a}{b-a}\right), & k \neq 0, \\ x_m^* - a, & k = 0. \end{cases}$$

On the other hand, by inserting the approximation (3.17) for the continuation value into the formula for $C_k(t_m, x_m^*)$ have the following coefficients $\hat{C}_k$ for $m = M - 1, M - 2, ..., 1$:

$$\hat{C}_k(t_m, x_m^*) = \frac{2e^{-r(t_{m+1}-t_m)}}{b-a} \sum_{j=0}^{N-1}{}' V_j(t_{m+1}) \int_{x_m^*}^{b} \mathrm{Re}\left(e^{-ij\pi\frac{a}{b-a}} \hat{\Gamma}\left(t_m, x; t_{m+1}, \frac{j\pi}{b-a}\right)\right) \cos\left(k\pi \frac{x-a}{b-a}\right) dx. \tag{3.19}$$

Thus the algorithm for pricing Bermudan options can then be summarized as in Fig. 1.

### 3.2. An efficient algorithm for the continuation value

In this section we derive an efficient algorithm for calculating $\hat{C}_k(t_m, x_m^*)$ in (3.19). When considering an exponential Lévy process with constant coefficients as done in [5], the continuation value can be calculated using a Fast Fourier Transform (FFT). This can be done due to the fact that the characteristic function $\hat{\Gamma}(t, x; T, \xi)$ can be split into a product of a function depending only on $\xi$ and a function of the form $e^{i\xi x}$. Note that we typically have $\xi = \frac{j\pi}{b-a}$. The integration over $x$ results in a sum of a Hankel and Toeplitz matrix (with indices $(j + k)$ and $(j - k)$ respectively). The matrix-vector product, with these special matrices, can be transformed into a circular convolution which can be computed using FFTs.

From (2.14) we know that the $n$th-order approximation of the characteristic function is of the form:

$$\hat{\Gamma}^{(n)}(t_m, x; t_{m+1}, \xi) = e^{i\xi x} \sum_{k=0}^{n} (x - \bar{x})^k g_{n,k}(t_m, t_{m+1}, \xi),$$

where the coefficients $g_{n,k}(t, T, \xi)$, with $0 \leq k \leq n$, depend only on $t, T$ and $\xi$, but not on $x$. Using (2.14) we write the continuation value as:

$$\hat{C}_k(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)} \sum_{j=0}^{N-1}{}' \mathrm{Re}\left(V_j(t_m) g_{n,h}\left(t_m, t_{m+1}, \frac{j\pi}{b-a}\right) M_{k,j}^h(x_m^*, b)\right),$$

where we have interchanged the sums and integral and defined:

$$M_{k,j}^h(x_m^*, b) = \frac{2}{b-a} \int_{x_m^*}^{b} e^{ij\pi\frac{x-a}{b-a}} (x - \bar{x})^h \cos\left(k\pi \frac{x-a}{b-a}\right) dx. \tag{3.20}$$

This can be written in vectorized form as:

$$\hat{\mathbf{C}}(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)} \mathrm{Re}\left(\mathbf{V}(t_{m+1})\mathcal{M}^h(x_m^*, b)\Lambda^h\right),$$

where $\mathbf{V}(t_{m+1})$ is the vector $[V_0(t_{m+1}), ..., V_{N-1}(t_{m+1})]^T$ and $\mathcal{M}^h(x_m^*, b)\Lambda^h$ is a matrix-matrix product with $\mathcal{M}^h$ being a matrix with elements $\{M_{k,j}^h\}_{k,j=0}^{N-1}$ and $\Lambda^h$ is a diagonal matrix with elements

$$g_{n,h}\Big(t_m, t_{m+1}, \frac{j\pi}{b-a}\Big), \qquad j = 0, \ldots, N-1.$$

We have the following theorem for calculating a generalized form of the integral in (3.20) which is used in the calculation of the continuation value.

**Theorem 3.7.** *The matrix $\mathcal{M}$ with elements $\{M_{k,j}\}_{k,j=0}^{N-1}$ such that:*

$$M_{k,j} = \int e^{jx}\cos(kx)x^m dx,$$

*consists of sums of Hankel and Toeplitz matrices.*

**Proof.** Using standard trigonometric identities we can rewrite the integral as:

$$\begin{aligned} M_{k,j} &= \int \cos(jx)\cos(kx)x^m dx + i\int \sin(jx)\cos(kx)x^m dx \\ &= M_{k,j}^H + iM_{k,j}^T, \end{aligned}$$

where we have defined:

$$\begin{aligned} M_{k,j}^H &= \frac{1}{2}\int \cos((j+k)x)x^m dx + \frac{1}{2}\int \sin((j+k)x)x^m dx, \\ M_{k,j}^T &= \frac{1}{2}\int \cos((j-k)x)x^m dx + \frac{1}{2}\int \sin((j-k)x)x^m dx. \end{aligned}$$

The following holds:

$$\begin{aligned} \int \cos(nx)x^m dx =& \frac{1}{n}x^m\sin(nx) + \sum_{i=1}^{\lceil m/2\rceil}(-1)^{i+1}\frac{\prod_{j=0}^{2i-2}(m-j)}{n^{2i}}\cos(nx)x^{m-(2i-1)} \\ &- \sum_{i=1}^{\lfloor m/2\rfloor}(-1)^{i+1}\frac{\prod_{j=0}^{2i-1}(m-j)}{n^{2i+1}}\sin(nx)x^{m-2i}, \\ \int \sin(nx)x^m dx =& -\frac{1}{n}x^m\cos(nx) + \sum_{i=1}^{\lceil m/2\rceil}(-1)^{i+1}\frac{\prod_{j=0}^{2i-2}(m-j)}{n^{2i}}\sin(nx)x^{m-(2i-1)} \\ &- \sum_{i=1}^{\lfloor m/2\rfloor}(-1)^{i+1}\frac{\prod_{j=0}^{2i-1}(m-j)}{n^{2i+1}}\cos(nx)x^{m-2i}. \end{aligned}$$

It follows that $\{M_{k,j}^H\}_{k,j=0}^{N-1}$ is a Hankel matrix with coefficient $(j+k)$ and $\{M_{k,j}^T\}_{k,j=0}^{N-1}$ is a Toeplitz matrix with coefficient $(j-k)$:

$$\mathcal{M}_H = \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_{N-1} \\ M_1 & M_2 & \dots & & M_N \\ \vdots & & & & \vdots \\ M_{N-2} & M_{N-1} & \dots & & M_{2N-3} \\ M_{N-1} & \dots & & M_{2N-3} & M_{2N-2} \end{pmatrix},$$

$$\mathcal{M}_T = \begin{pmatrix} M_0 & M_1 & \dots & M_{N-2} & M_{N-1} \\ M_{-1} & M_0 & M_1 & \dots & M_{N-2} \\ \vdots & & \ddots & & \vdots \\ M_{2-N} & \dots & M_{-1} & M_0 & M_1 \\ M_{1-N} & M_{2-N} & & M_{-1} & M_0 \end{pmatrix},$$

where we have defined

$$M_j = \frac{1}{2}\int \cos(jx)x^m dx + \frac{1}{2}\int \sin(jx)x^m dx. \quad \square$$

From Theorem 3.7 we see that $\mathcal{M}^h(x_m^*, b)$ with elements $M_{k,j}^h$ consists of a sum of a Hankel and Toeplitz matrix.

**Example 3.8.** We derive explicitly the Hankel and Toeplitz matrices for $m = 0$ and $m = 1$. We calculate the indefinite integral

$$M_{k,j} = \frac{2}{b-a}\int e^{ij\pi\frac{x-a}{b-a}} \cos\left(k\pi\frac{x-a}{b-a}\right)(x-\bar{x})^m dx.$$

Suppose $m = 0$, in this case we have $M_{k,j} = M_{k,j}^H + M_{k,j}^T$, with:

$$M_{k,j}^H = -\frac{i\exp\left(i\frac{(j+k)\pi(x-a)}{b-a}\right)}{\pi(j+k)},$$

$$M_{k,j}^T = -\frac{i\exp\left(i\frac{(j-k)\pi(x-a)}{b-a}\right)}{\pi(j-k)},$$

where $\{M_{k,j}^H\}_{k,j=0}^{N-1}$ is a Hankel matrix and $\{M_{k,j}^T\}_{k,j=0}^{N-1}$ is a Toeplitz matrix with

$$M_j = \begin{cases} \frac{x}{b-a}, & j = 0, \\ -\frac{i\exp\left(i\frac{j\pi(x-a)}{b-a}\right)}{\pi j}, & j \neq 0. \end{cases}$$

Suppose $m = 1$, in this case we have:

$$M_{k,j}^H = -\frac{a-b}{(j-k)^2\pi^2}\exp\left(i(j-k)\pi\frac{(x-a)}{b-a}\right) - \frac{x-\bar{x}}{(j-k)\pi} i\exp\left(i(j-k)\pi\frac{(x-a)}{b-a}\right),$$

$$M_{k,j}^T = -\frac{a-b}{(j+k)^2\pi^2}\exp\left(i(j+k)\pi\frac{(x-a)}{b-a}\right) - \frac{x-\bar{x}}{(j+k)\pi} i\exp\left(i(j+k)\pi\frac{(x-a)}{b-a}\right),$$

where $\{M_{k,j}^H\}_{k,j=0}^{N-1}$ is a Hankel matrix and $\{M_{k,j}^T\}_{k,j=0}^{N-1}$ is a Toeplitz matrix, with

$$M_j = \begin{cases} \frac{x(x-\bar{x})}{b-a}, & j = 0, \\ -\frac{a-b}{j^2\pi^2}\exp\left(ij\pi\frac{(x-a)}{b-a}\right) - \frac{x-\bar{x}}{j\pi} i\exp\left(ij\pi\frac{(x-a)}{b-a}\right), & j \neq 0. \end{cases}$$

1. For $h = 0, ..., n$:
   - Compute $M_j^h(x_1, x_2)$
   - Construct $\mathbf{m}_H^h$ and $\mathbf{m}_T^h$
   - Compute $\boldsymbol{u}^h(t_m) = \{u_j^h\}_{j=0}^{N-1}$
   - Construct $\mathbf{u}_T^h$ by padding $N$ zeros to $\boldsymbol{u}^h(t_m)$
   - $\mathbf{MTu}^h$ = the first $N$ elements of $\mathcal{D}^{-1}\{\mathcal{D}(\mathbf{m}_T^h) \cdot \mathcal{D}(\mathbf{u}_T^h)\}$
   - $\mathbf{MHu}^h$ = reverse{the first $N$ elements of $\mathcal{D}^{-1}\{\mathcal{D}(\mathbf{m}_H^h) \cdot \mathbf{sgn} \cdot \mathcal{D}(\mathbf{u}_T^h)\}$}
2. Compute the continuation value using $\hat{\mathbf{C}}(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)}\mathrm{Re}(\mathbf{MTu}^h + \mathbf{MHu}^h)$.

**Fig. 2.** Algorithm 3.2: Computation of $\hat{\mathbf{C}}(t_m, x_m^*)$.

**Remark 3.9.** If we take $\bar{x} = x$, which is most common in practice, the formulas are simplified significantly and only the case of $m = 0$ is relevant. In this case the characteristic function is simply $e^{i\xi x}$ times a sum of terms depending only on $t_m$, $t_{m+1}$ and $\xi = \frac{j\pi}{b-a}$:

$$\hat{\Gamma}^{(n)}(t_m, x; t_{m+1}, \xi) = e^{i\xi x} g_{n,0}(t_m, t_{m+1}, \xi).$$

Using the split into sums of Hankel and Toeplitz matrices we can write the continuation value in matrix form as:

$$\hat{\boldsymbol{C}}(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)}\mathrm{Re}\left((\mathcal{M}_H^h + \mathcal{M}_T^h)\boldsymbol{u}^h\right),$$

where $\mathcal{M}_H^h = \{M_{k,j}^{H,h}(x_m^*, b)\}_{k,j=0}^{N-1}$ is a Hankel matrix and $\mathcal{M}_T^l = \{M_{k,j}^{T,h}(x_m^*, b)\}_{k,j=0}^{N-1}$ is a Toeplitz matrix and $\boldsymbol{u}^h = \{u_j^h\}_{j=0}^{N-1}$, with $u_j^h = g_{n,h}\left(t_m, t_{m+1}, \frac{j\pi}{b-a}\right) V_j(t_{m+1})$ and $u_0^h = \frac{1}{2} g_{n,h}(t_m, t_{m+1}, 0)\, V_0(t_{m+1})$.

We recall that the circular convolution, denoted by $\circledast$, of two vectors is equal to the inverse discrete Fourier transform ($\mathcal{D}^{-1}$) of the products of the forward DFTs, $\mathcal{D}$, i.e.:

$$\mathbf{x} \circledast \mathbf{y} = \mathcal{D}^{-1}\{\mathcal{D}(\mathbf{x}) \cdot \mathcal{D}(\mathbf{y})\}.$$

For Hankel and Toeplitz matrices we have the following result:

**Theorem 3.10.** *For a Toeplitz matrix $\mathcal{M}_T$, the product $\mathcal{M}_T\mathbf{u}$ is equal to the first $N$ elements of $\mathbf{m}_T \circledast \mathbf{u}_T$, where $\mathbf{m}_T$ and $\mathbf{u}_T$ are $2N$ vectors defined by*

$$\mathbf{m}_T = [M_0, M_{-1}, M_{-2}, ..., M_{1-N}, 0, M_{N-1}, M_{N-2}, ..., M_1]^T,$$
$$\mathbf{u}_T = [u_0, u_1, ..., u_{N-1}, 0, ..., 0]^T.$$

*For a Hankel matrix $\mathcal{M}_H$, the product $\mathcal{M}_H\mathbf{u}$ is equal to the first $N$ elements of $\mathbf{m_H} \circledast \mathbf{u_H}$ in reversed order, where $\mathbf{m}_H$ and $\mathbf{u}_H$ are $2N$ vectors defined by*

$$\mathbf{m}_H = [M_{2N-1}, M_{2N-2}, ..., M_1, M_0]^T$$
$$\mathbf{u}_H = [0, ..., 0, u_0, u_1, ..., u_{N-1}]^T.$$

Summarizing, we can calculate the continuation value $\hat{\mathbf{C}}(t_m, x_m^*)$ using the algorithm in Fig. 2.

The continuation value requires five DFTs for each $h = 0, ..., n$, and a DFT is calculated using the FFT. In practice it is most common to have $\bar{x} = x$ and in this case we only need five FFTs. The computation of $F_k(t_m, x_m^*)$ is linear in $N$. The overall complexity of the method is dominated by the computation of $\hat{C}(t_m, x_m^*)$, whose complexity is $O(N \log_2 N)$ with the FFT. The complexity of the calculation for option value at time 0 is $O(N)$. If we have a Bermudan option with $M$ exercise dates, the overall complexity will be $O((M-1)N \log_2 N)$.

**Remark 3.11** *(American options).* The prices of American options can be obtained by applying a Richardson extrapolation (see, for instance, [9]) on the prices of a few Bermudan options with a small number of exercise dates. Let $v_M$ denote the value of a Bermudan option with maturity $T$ and a number $M$ of early exercise dates that are $\frac{T}{M}$ years apart. Then, for any $d \in \mathbb{N}$, the following 4-point Richardson extrapolation scheme

$$\frac{1}{21}\left(64 v_{2^{d+3}} - 56 v_{2^{d+2}} + 14 v_{2^{d+1}} - v_{2^d}\right)$$

gives an approximation of the corresponding American option price.

**Remark 3.12** *(The Greeks).* The approximation method can also be used to calculate the Greeks at almost no additional cost. In the case of $\bar{x} = x$, we have the following approximation formulas for Delta and Gamma:

$$\hat{\Delta} = e^{-r(t_1-t_0)} {\sum_{k=0}^{N-1}}' \operatorname{Re}\left(e^{ik\pi\frac{x-a}{b-a}}\left(\frac{ik\pi}{b-a} g_{n,0}\left(t_0, t_1, \frac{k\pi}{b-a}\right) + g_{n,1}\left(t_0, t_1, \frac{k\pi}{b-a}\right)\right)\right)\hat{V}_k(t_1),$$

$$\hat{\Gamma} = e^{-r(t_1-t_0)} {\sum_{k=0}^{N-1}}' \operatorname{Re}\Bigg(e^{ik\pi\frac{x-a}{b-a}}\Bigg(-\frac{ik\pi}{b-a} g_{n,0}\left(t_0, t_1, \frac{k\pi}{b-a}\right) - g_{n,1}\left(t_0, t_1, \frac{k\pi}{b-a}\right)$$
$$+ 2\frac{ik\pi}{b-a} g_{n,1}\left(t_0, t_1, \frac{k\pi}{b-a}\right) + \left(\frac{ik\pi}{b-a}\right)^2 g_{n,0}\left(t_0, t_1, \frac{k\pi}{b-a}\right) + 2g_{n,2}\left(t_0, t_1, \frac{k\pi}{b-a}\right)\Bigg)\Bigg)\hat{V}_k(t_1).$$

## 4. Error estimates

The error in our approximation consists of the error of the COS method and the error in the adjoint expansion of the characteristic function. The error of the COS method depends on the truncation of the integration range $[a, b]$ and the truncation of the infinite summation of the Fourier-cosine expansion by $N$. The density rapidly decays to zero as $y \to \pm\infty$. Then the overall error can be bounded as follows:

$$\epsilon_1(x; N, [a, b]) \le Q\left|\int_{\mathbb{R}\setminus[a,b]} \Gamma(t, x; T, dy)\right| + \left|\frac{P}{(N-1)^{\beta-1}}\right|,$$

where $P$ and $Q$ are constants not depending on $N$ or $[a, b]$ and $\beta \ge n \ge 1$, with $n$ being the algebraic index of convergence of the cosine series coefficients. For a sufficiently large integration interval $[a, b]$, the overall error is dominated by the series truncation error, which converges exponentially. The error in the backward propagation of the coefficients $V_k(t_m)$ is defined as $\epsilon_2(k, t_m) := V_k(t_m) - \hat{V}_k(t_m)$. With $[a, b]$ sufficiently large and a probability density function in $C^\infty([a, b])$, the error $\epsilon_1(k, t_m)$ converges exponentially in $N$. For a detailed derivation on the error of the COS method see [4] and [5].

We now present the error estimates for the adjoint expansion of the characteristic function at orders zero and one. We consider for simplicity a model with time-independent coefficients

$$X_t = x + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dW_s + \int_0^t \int_{\mathbb{R}} \eta(X_{s-})z d\tilde{N}(s, dz), \tag{4.21}$$

where we have defined as usual $d\tilde{N}(t, dz) = dN(t, dz) - \nu(dz)dt$. This model is similar to the model we considered initially in (2.1); only now we deal with slightly simplified version and assume that the dependency on $X_t$ in the measure can be factored out, which is often enough the case.

Let $\tilde{X}_t$ be the 0th-order approximation of the model in (4.21) with $\bar{x} = x$, that is

$$\tilde{X}_t = x + \int_0^t \mu(x)ds + \int_0^t \sigma(x)dW_s + \int_0^t \int_{\mathbb{R}} \eta(x)zd\tilde{N}(s, dz). \tag{4.22}$$

The characteristic exponent of $\tilde{X}_t - x$ is

$$\psi(\xi) = i\xi\mu(x) - \frac{\sigma(x)^2}{2}\xi^2 - \eta(x)\int_{\mathbb{R}} \nu(dz)(e^z - 1 - z)i\xi + \eta(x)\int_{\mathbb{R}} \nu(dz)(e^{iz\xi} - 1 - iz\xi). \tag{4.23}$$

**Theorem 4.13.** *Let $n = 0, 1$ and assume that the coefficients $\mu, \sigma, \eta$ are continuously differentiable with bounded derivatives up to order $n$. Let $\hat{\Gamma}^{(n)}(0, x; t, \xi)$ in (2.6) be the nth-order approximation of the characteristic function. Then, for any $T > 0$ there exists a positive constant $C$ that depends only on $T$, on the norms of the coefficients and on the Lévy measure $\nu$, such that*

$$\left|\hat{\Gamma}(0, x; t, \xi) - \hat{\Gamma}^{(n)}(0, x; t, \xi)\right| \leq C\left(1 + |\xi|^{1+3n}\right)t^{n+1}, \qquad t \in [0, T],\ \xi \in \mathbb{R}. \tag{4.24}$$

**Proof.** For the proof we refer to Appendix A. □

**Remark 4.14.** The proof of Theorem 4.13 can be generalized to obtain error bounds for any $n \in \mathbb{N}$: however, one can see that, for $n \geq 2$, the order of convergence improves only in the diffusive part, according to the results proved in [10].

## 5. Numerical experiments

In this section we apply the method developed in Section 4 to compute the European and Bermudan option values with various underlying stock dynamics. The computer used in the experiments has an Intel Core i7 CPU with a 2.2 GHz processor. We use the second-order approximation of the characteristic function. We have found this to be sufficiently accurate by numerical experiments and theoretical error estimates. The formulas for the second-order approximation are simple, making the method easy to implement.

For the COS method, unless otherwise mentioned, we use $N = 200$ and $L = 10$, where $L$ is the parameter used to define the truncation range $[a, b]$ as follows:

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

where $c_n$ is the $n$th cumulant of log-price process $\log S$, as proposed in [4]. The cumulants are calculated using the 0th-order approximation of the characteristic function. A larger $N$ and $L$ has little effect on the price, since a fast convergence is achieved already for small $N$ and $L$. We compare the approximated values to a 95% confidence interval computed with a Longstaff–Schwartz method with $10^5$ simulations and 250 time steps per year. Furthermore, in the expansion we always use $\bar{x} = x$.

**Table 1**
Prices for a European and a Bermudan put option (expiry $T = 0.25$ with 3 exercise dates, expiry $T = 1$ with 10 exercise dates and expiry $T = 2$ with 20 exercise dates) in the CEV-Merton model for the 2nd-order approximation of the characteristic function, and a Monte Carlo method.

| T | K | European | | Bermudan | |
|---|---|---|---|---|---|
| | | MC 95% c.i. | Value | MC 95% c.i. | Value |
| 0.25 | 0.6 | 0.001240–0.001433 | 0.001326 | 0.001243–0.001431 | 0.001307 |
| | 0.8 | 0.005218–0.005679 | 0.005493 | 0.005314–0.005774 | 0.005421 |
| | 1 | 0.04222–0.04321 | 0.04275 | 0.04274–0.04371 | 0.04304 |
| | 1.2 | 0.1923–0.1938 | 0.1935 | 0.1979–0.1989 | 0.1981 |
| | 1.4 | 0.3856–0.3872 | 0.3866 | 0.3948–0.3958 | 0.3955 |
| | 1.6 | 0.5812–0.5829 | 0.5825 | 0.5940–0.5950 | 0.5941 |
| 1 | 0.6 | 0.006136–0.006573 | 0.006579 | 0.006307–0.006729 | 0.006096 |
| | 0.8 | 0.02526–0.02622 | 0.02581 | 0.02617–0.02711 | 0.02520 |
| | 1 | 0.08225–0.08395 | 0.08250 | 0.08480–0.08640 | 0.08593 |
| | 1.2 | 0.1965–0.1989 | 0.1977 | 0.2097–0.2115 | 0.2132 |
| | 1.4 | 0.3560–0.3589 | 0.3574 | 0.3946–0.3957 | 0.3954 |
| | 1.6 | 0.5341–0.5385 | 0.5364 | 0.5930–0.5941 | 0.5932 |
| 2 | 0.6 | 0.01444–0.01513 | 0.01529 | 0.01528–0.01594 | 0.01365 |
| | 0.8 | 0.04522–0.04655 | 0.04613 | 0.04596–0.04719 | 0.04659 |
| | 1 | 0.1046–0.1067 | 0.1077 | 0.1149–0.1168 | 0.1171 |
| | 1.2 | 0.2054–0.2083 | 0.2065 | 0.2319–0.2341 | 0.2345 |
| | 1.4 | 0.3351–0.3386 | 0.3382 | 0.3968–0.3987 | 0.3991 |
| | 1.6 | 0.4904–0.4944 | 0.4919 | 0.5927–0.5938 | 0.5935 |

### *5.1. Tests under CEV-Merton dynamics*

Consider a process under the CEV-Merton dynamics:

$$dX_t = \left(r - a(X_t) - \lambda\left(e^{m+\delta^2/2} - 1\right)\right)dt + \sqrt{2a(X_t)}dW_t + \int_{\mathbb{R}} d\tilde{N}_t(t, dz)z,$$

with

$$\begin{aligned}
a(x) &= \frac{\sigma_0^2 e^{2(\beta-1)x}}{2}, \\
\nu(dz) &= \lambda \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(\frac{-(z-m)^2}{2\delta^2}\right) dz, \\
\psi(\xi) &= -a_0(\xi^2 + i\xi) + ir\xi - i\lambda\left(e^{m+\delta^2/2} - 1\right)\xi + \lambda\left(e^{mi\xi - \delta^2\xi^2/2} - 1\right).
\end{aligned}$$

We use the following parameters $S_0 = 1$, $r = 5\%$, $\sigma_0 = 20\%$, $\beta = 0.5$, $\lambda = 30\%$, $m = -10\%$, $\delta = 40\%$ and compute the European and Bermudan option values.

We present the results in Table 1. The option value for both the Bermudan options as well as the European options appears to be accurate. Since the COS method has a very quick convergence, already for $N = 64$ the error becomes stable. For at-the-money strikes we have $\log_{10}|\text{error}| \approx 3.5$. The use of the second-order approximation of the characteristic function is justified by the fact that the option value (and thus the error) stabilizes starting from the second-order approximation. Furthermore, it is noteworthy that the 0th-order approximation is already very accurate. The CPU time of the calculations depends on the number of exercise dates. Assuming we use the second-order approximation of the characteristic function, if we have $M$ exercise dates the CPU time will be $5 \cdot M$ ms.

**Remark 5.15.** The method can be extended to include time-dependent coefficients. The accuracy and speed of the method will be of the same order as for time-independent coefficients.

**Table 2**
Prices for a European and a Bermudan put option (10 exercise dates, expiry $T = 1$) in the CEV-VG model for the 2nd-order approximation of the characteristic function, and a Monte Carlo method.

| K | European | | Bermudan | |
|---|---|---|---|---|
| | MC 95% c.i. | Value | MC 95% c.i. | Value |
| 0.6 | 0.03090–0.03732 | 0.03546 | 0.03756–0.03876 | 0.03749 |
| 0.8 | 0.08046–0.08247 | 0.08029 | 0.08290–0.08484 | 0.08395 |
| 1 | 0.1507–0.1531 | 0.1511 | 0.1572–0.1600 | 0.1594 |
| 1.2 | 0.2501–0.2538 | 0.2522 | 0.2634–0.2668 | 0.2685 |
| 1.4 | 0.3831–0.3876 | 0.3847 | 0.4073–0.4108 | 0.4137 |
| 1.6 | 0.5430–0.5479 | 0.5436 | 0.5920–0.5938 | 0.5937 |

**Remark 5.16.** The Greeks can be calculated at almost no additional cost using the formulas presented in 3.12. Numerically, the order of convergence is algebraic and is the same for both the exact characteristic function as for the 2nd-order approximation.

### *5.2. Tests under the CEV-Variance-Gamma dynamics*

Consider the jump process to be a Variance-Gamma process. The VG process, is obtained by replacing the time in a Brownian motion with drift $\theta$ and standard deviation $\varrho$, by a Gamma process with variance $\kappa$ and unitary mean. The model parameters $\varrho$ and $\kappa$ allow to control the skewness and the kurtosis of the distribution of stock price returns. The VG density is characterized by a fat tail and is thus used as a model in situations where small and large asset values are more probable than would be the case for the lognormal distribution. The Lévy measure in this case is given by:

$$\nu(dx) = \frac{e^{-\lambda_1 x}}{\kappa x}\mathbb{1}_{\{x>0\}}dx + \frac{e^{\lambda_2 x}}{\kappa|x|}\mathbb{1}_{\{x<0\}}dx,$$

where

$$\lambda_1 = \left(\sqrt{\frac{\theta^2\kappa^2}{4} + \frac{\varrho^2\kappa}{2}} + \frac{\theta\kappa}{2}\right)^{-1}, \qquad \lambda_2 = \left(\sqrt{\frac{\theta^2\kappa^2}{4} + \frac{\varrho^2\kappa}{2}} - \frac{\theta\kappa}{2}\right)^{-1}.$$

Furthermore we have

$$\begin{aligned}
a(x) &= \frac{\sigma_0^2 e^{2(\beta-1)x}}{2},\\
\mu(t,x) &= r + \frac{1}{\kappa}\log\left(1 - \kappa\theta - \frac{\kappa\varrho^2}{2}\right) - a(x),\\
\psi(\xi) &= -a_0(\xi^2 + i\xi) + ir\xi + i\frac{1}{\kappa}\log\left(1 - \kappa\theta - \frac{\kappa\varrho^2}{2}\right)\xi - \frac{1}{\kappa}\log\left(1 - i\kappa\theta\xi + \frac{\xi^2\kappa\varrho^2}{2}\right).
\end{aligned}$$

We use the following parameters $S_0 = 1$, $r = 5\%$, $\sigma_0 = 20\%$, $\beta = 0.5$, $\kappa = 1$, $\theta = -50\%$, $\varrho = 20\%$. The results for the European and Bermudan option are presented in Table 2.

### *5.3. Tests under a CEV-like Lévy process with a state-dependent measure and default*

In this section we consider a model similar to the one used in [7]. The model is defined with local volatility, local default and a state-dependent Lévy measure as follows:

$$a(x) = \frac{1}{2}(b_0^2 + \epsilon_1 b_1^2\eta(x)),$$

**Table 3**
Prices for a European and a Bermudan put option (10 exercise dates, expiry $T = 1$) in the CEV-like model with state-dependent measure for the 2nd-order approximation characteristic function, and a Monte Carlo method.

| K | European | | Bermudan | |
|---|---|---|---|---|
| | MC 95% c.i. | Value | MC 95% c.i. | Value |
| 0.8 | 0.01025–0.01086 | 0.009385 | 0.01068–0.01125 | 0.01024 |
| 1 | 0.04625–0.04745 | 0.04817 | 0.05141–0.05253 | 0.05488 |
| 1.2 | 0.1563–0.1582 | 0.1564 | 0.1942–0.1952 | 0.1952 |
| 1.4 | 0.3313–0.3334 | 0.3314 | 0.3927–0.3934 | 0.3930 |
| 1.6 | 0.5207–0.5229 | 0.5218 | 0.5919–0.5926 | 0.5920 |
| 1.8 | 0.7103–0.7124 | 0.7122 | 0.7906–0.7913 | 0.7910 |

**Table 4**
Prices for a European and a Bermudan put option (10 exercise dates, expiry $T = 1$) in the CEV-like model with default for the 2nd-order approximation characteristic function, and a Monte Carlo method.

| K | European | | Bermudan | |
|---|---|---|---|---|
| | MC 95% c.i. | Value | MC 95% c.i. | Value |
| 0.8 | 0.002905–0.003175 | 0.003061 | 0.005876–0.006245 | 0.006361 |
| 1 | 0.01845–0.01918 | 0.01893 | 0.03419–0.03506 | 0.03520 |
| 1.2 | 0.08148–0.08296 | 0.08297 | 0.1820–0.1827 | 0.1824 |
| 1.4 | 0.2184–0.2205 | 0.2173 | 0.3793–0.3801 | 0.3792 |
| 1.6 | 0.3867–0.3892 | 0.3841 | 0.5752–0.5763 | 0.5763 |
| 1.8 | 0.5597–0.5638 | 0.5556 | 0.7727–0.7739 | 0.7733 |

$$\begin{aligned} \gamma(x) &= c_0 + \epsilon_2 c_1 \eta(x), \\ \nu(x, dz) &= \epsilon_3 \nu_N(dz) + \epsilon_4 \eta(x) \nu_N(dz), \\ \eta(x) &= e^{\beta x}. \end{aligned} \tag{5.25}$$

We will consider Gaussian jumps, meaning that

$$\nu_N(dz) = \lambda \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(\frac{-(z-m)^2}{2\delta^2}\right) dz.$$

The regular CEV model has several shortcomings; for instance, the volatility drops to zero as the underlying approaches infinity; also the model does not allow the underlying to experience jumps. This model tries to overcome these shortcomings, while still retaining CEV-like behavior through $\eta(x)$. The local volatility function $\sigma(x)$ behaves asymptotically like the CEV model, $\sigma(x) \sim \sqrt{\epsilon_1} b_1 e^{\beta x/2}$ as $x \to -\infty$, reflecting the fact that the volatility tends to increase as the asset price drops (the leverage effect). Jumps of size $dz$ arrive with a state-dependent intensity of $\nu(x, dz)$. Lastly, a default arrives with intensity $\gamma(x)$. The default function $\gamma(x)$ behaves asymptotically like $\epsilon_2 c_1 e^{\beta x}$ as $x \to -\infty$, reflecting the fact that a default is more likely to occur when the price goes down.

In Table 3 the results are presented for a model as defined in (5.25) without default, meaning that $c_0 = c_1 = 0$ and with a state-dependent jump measure, so $\nu(x, dz) = \eta(x)\nu_N(dz)$. In this case we have

$$\psi(\xi) = ir\xi - a_0(\xi^2 - i\xi) - \lambda\nu_0(e^{m+\delta^2/2} - 1)i\xi + \lambda\nu_0(e^{mi\xi - \delta^2\xi^2/2} - 1),$$

where $a_0 = \frac{1}{2} b_1^2 e^{\beta\bar{x}}$ and $\nu_0(dz) = e^{\beta\bar{x}}\nu_N(dz)$. The other parameters are chosen as: $b_1 = 0.15$, $b_0 = 0$, $\beta = -2$, $\lambda = 20\%$, $\delta = 20\%$, $m = -0.2$, $S_0 = 1$, $r = 5\%$, $\epsilon_1 = 1$, $\epsilon_3 = 0$, $\epsilon_4 = 1$, the number of exercise dates is 10 and $T = 1$. From the results for both the European option and the Bermudan option we see that the method performs very accurately, even for deeply in-the-money strikes.

In Table 4 the results are presented for the value of a defaultable put option. In case of default prior to exercise the put option payoff is 0, in case of no default the value is $(K - S_t)^+$, depending on the exercise time. We look at the model as defined in (5.25) with the possibility of default and consider state-independent jumps, meaning that we have $\gamma(x) = \eta(x)$ and $\nu(x, dz) = \nu_N(dz)$. We have

$$\psi(\xi) = ir\xi - a_0(\xi^2 - i\xi) + \gamma_0 i\xi - \gamma_0 - \lambda(e^{m+\delta^2/2} - 1)i\xi + \lambda(e^{mi\xi - \delta^2\xi^2/2} - 1),$$

where $a_0 = \frac{1}{2}b_1^2 e^{\beta\bar{x}}$ and $\gamma_0 = c_1 e^{\beta\bar{x}}$. The other parameters are $b_0 = 0$, $b_1 = 0.15$, $\beta = -2$, $c_0 = 0$, $c_1 = 0.1$, $S_0 = 1$, $r = 5\%$, $\epsilon_1 = 1$, $\epsilon_2 = 1$, $\epsilon_3 = 1$, $\epsilon_4 = 0$, the number of exercise dates is 10 and $T = 1$.

**Acknowledgments**

We thank an anonymous referee for some comments that improved the paper.

## Appendix A. Proof of Theorem 4.13

Let $X$ and $\tilde{X}$ be as in (4.21) and (4.22) respectively. We first prove that

$$E[|X_t - \tilde{X}_t|^2] \leq C\left(\kappa_2 t^2 + \kappa_1^2 t^3\right), \qquad t \in [0, T], \tag{A.1}$$

for some positive constant $C$ that depends only on $T$, on the Lipschitz constants of the coefficients $\mu$, $\sigma$, $\eta$ and on the Lévy measure $\nu$. Here $\kappa_1 = -\psi'(0)$ and $\kappa_2 = -\psi''(0)$ where $\psi$ in (4.23) is the characteristic exponent of the Lévy process $(\tilde{X}_t - x)$.

Using the Hölder inequality, the Itô isometry (see, for instance, [15]) and the Lipschitz continuity of $\eta$, $\mu$ and $\sigma$, the mean squared error is bounded by:

$$\begin{aligned} E\left[|X_t - \tilde{X}_t|^2\right] &\leq 3E\left[\left(\int_0^t (\mu(X_s) - \mu(x))ds\right)^2\right] + 3E\left[\left(\int_0^t (\sigma(X_s) - \sigma(x))dW_s\right)^2\right] \\ &\quad + 3E\left[\left(\int_0^t \int_{\mathbb{R}} (\eta(X_{s-}) - \eta(x))z d\tilde{N}(s, dz)\right)^2\right] \\ &\leq C\int_0^t E\left[|\tilde{X}_s - x|^2\right] ds + C\int_0^t E\left[|X_s - \tilde{X}_s|^2\right] ds, \end{aligned} \tag{A.2}$$

where

$$C = 6\left(\|\mu'\|_\infty^2 + \|\sigma'\|_\infty^2 + \|\eta'\|_\infty^2 \int_{\mathbb{R}} z^2 \nu(dz)\right).$$

Now we recall the following relationship between the first and second moment and cumulants

$$E[(\tilde{X}_s - x)] = c_1(s), \qquad E[(\tilde{X}_s - x)^2] = c_2(s) + c_1(s)^2,$$

where

$$c_n(s) = \frac{s}{i^n} \frac{\partial^n \psi(\xi)}{\partial \xi^n}\bigg|_{\xi=0},$$

and $\psi(\xi)$ is the characteristic exponent of $(\tilde{X}_s - x)$. Thus we have

$$E\left[|\tilde{X}_s - x|^2\right] = \kappa_2 s + \kappa_1^2 s^2. \tag{A.3}$$

Plugging (A.3) into (A.2) we get

$$E[|X_t - \tilde{X}_t|^2] \le C\left(\frac{\kappa_2}{2}t^2 + \frac{\kappa_1^2}{3}t^3\right) + C\int_0^t E\left[|X_s - \tilde{X}_s|^2\right] ds,$$

and therefore estimate (A.1) follows by applying the Gronwall inequality in the form

$$\phi(t) \le \alpha(t) + C\int_0^t \phi(s)ds \implies \phi(t) \le \alpha(t) + C\int_0^t \alpha(s)e^{C(t-s)}ds,$$

that is valid for any $C \ge 0$ and $\phi$, $\alpha$ continuous functions.

From (A.1) and (A.3) we can also deduce that

$$E\left[\left|X_t - x\right|^2\right] \le 2E\left[\left|X_t - \tilde{X}_t\right|^2\right] + 2E\left[\left|\tilde{X}_t - x\right|^2\right] \le C\left(\kappa_2 t + \kappa_1^2 t^2\right), \qquad t \in [0,T]. \tag{A.4}$$

Moreover, from (A.1) we also get the following error estimate for the expectation of a Lipschitz payoff function $v$:

$$\left|E\left[v(X_t)\right] - E[v(\tilde{X}_t)]\right| \le C\sqrt{\kappa_2 t + \kappa_1^2 t^2}, \qquad t \in [0,T],$$

where now $C$ also depends on the Lipschitz constant of $v$. In particular, taking $v(x) = e^{ix\xi}$, this proves (4.24) for $n = 0$.

Next we prove (4.24) for $n = 1$.

Proceeding as in the proof of Lemma 6.23 in [10] with $u(0,x) = \hat{\Gamma}(0,x;t,\xi)$ and $\bar{x} = x$, we find

$$\hat{\Gamma}(0,x;t,\xi) - \hat{\Gamma}^{(1)}(0,x;t,\xi) = \int_0^t E\left[(L - L_0)\hat{G}^1(s,X_s;t,\xi) + (L - L_1)\hat{G}^0(s,X_s;t,\xi)\right] ds,$$

where the 1st-order approximation is as usual

$$\hat{\Gamma}^{(1)}(s,X;t,\xi) = \hat{G}^0(s,X;t,\xi) + \hat{G}^1(s,X;t,\xi),$$

with

$$\begin{aligned}\hat{G}^0(s,X;t,\xi) &= e^{iX\xi + (t-s)\psi(\xi)},\\ \hat{G}^1(s,X;t,\xi) &= e^{iX\xi + (t-s)\psi(\xi)} g_0^{(1)}(t-s,\xi),\end{aligned}$$

and $g_0^{(1)}$ as in (2.13). Using the Lagrangian remainder of the Taylor expansion, we have

$$\begin{aligned}L - L_0 = {}& \gamma'(\varepsilon')(X - x)(\partial_X - 1) + a'(\varepsilon')(X - x)(\partial_{XX} - \partial_X) + \eta'(\varepsilon')(X - x)\int_{\mathbb{R}} \nu(dz)(e^z - 1 - z)\partial_X\\ &+ \eta'(\varepsilon')(X - x)\int_{\mathbb{R}} \nu(dz)(e^{z\partial_X} - 1 - z\partial_X),\end{aligned}$$

$$
\begin{aligned}
L - L_1 = {} & \frac{1}{2}\gamma''(\varepsilon'')(X-x)^2(\partial_X - 1) + \frac{1}{2}a''(\varepsilon'')(X-x)^2(\partial_{XX} - \partial_X) \\
& + \frac{1}{2}\eta''(\varepsilon'')(X-x)^2 \int_{\mathbb{R}} \nu(dz)(e^z - 1 - z)\partial_X + \frac{1}{2}\eta''(\varepsilon'')(X-x)^2 \int_{\mathbb{R}} \nu(dz)(e^{z\partial_X} - 1 - z\partial_X),
\end{aligned}
$$

for some $\varepsilon', \varepsilon'' \in [x, X]$. Now, $|\hat{G}^0| \le 1$ because $\hat{G}^0$ is the characteristic function of the process $\tilde{X}$ in (4.22); thus, we have

$$
\left|(L - L_1)\hat{G}^0(s, X_s; t, \xi)\right| \le C(1 + |\xi|^2)\, |X_s - x|^2 .
$$

On the other hand, from (2.13) we have

$$
\left|g_0^{(1)}(t-s, \xi)\right| \le C(t-s)^2 \left(1 + |\xi|^4\right),
$$

and therefore we get

$$
\left|(L - L_0)\hat{G}^1(s, X_s; t, \xi)\right| \le C(t-s)^2(1 + |\xi|^4)\, |X_s - x| .
$$

So we find

$$
\left|\hat{\Gamma}(0, x; t, \xi) - \hat{\Gamma}^{(1)}(0, x; t, \xi)\right| \le C(1 + |\xi|^4) \int_0^t \left((t-s)^2 E\left[|X_s - x|\right] + E\left[|X_s - x|^2\right]\right) ds
$$

The thesis then follows from estimate (A.4) and integrating.

## Appendix B. The 2nd-order approximation of the characteristic function

For completeness we present here the formulas of the characteristic function approximation in the general case up to the 2nd-order approximation for a process as in (2.1) with a local-volatility coefficient $a(t, x)$, a local default intensity $\gamma(t, x)$ and a state-dependent measure $\nu(t, x, dz)$. We expand the coefficients around $\bar{x} = x$. This choice of $\bar{x}$ is most common in practice and it simplifies the formulas significantly. We have

$$
\begin{aligned}
\hat{G}^{(0)}(t, x; T, \xi) = {} & e^{i\xi x + (T-t)\psi(\xi)} \\
\hat{G}^{(1)}(t, x; T, \xi) = {} & \hat{G}^{(0)}(t, x; T, \xi)\bigg(\frac{1}{2}i(T-t)^2\xi(i+\xi)\alpha_1\psi'(\xi) + \frac{1}{2}(T-t)^2(i+\xi)\gamma_1\psi'(\xi) \\
& - \frac{1}{2}\int_{\mathbb{R}} \nu_1(dz)z(T-t)^2\xi\psi'(\xi) - \frac{1}{2}\int_{\mathbb{R}} \nu_1(dz)(e^z - 1 - z)\xi\psi'(\xi) \\
& - \frac{1}{2}\int_{\mathbb{R}} i(e^{iz\xi} - 1)(T-t)^2\psi'(\xi)\bigg) \\
\hat{G}^{(2)}(t, x; T, \xi) = {} & \hat{G}^{(0)}(t, x; T, \xi)\big(G_1^{(2)}(t, x; T, \xi) + G_2^{(2)}(t, x; T, \xi) + G_3^{(2)}(t, x; T, \xi) \\
& + G_4^{(2)}(t, x; T, \xi) + G_5^{(2)}(t, x; T, \xi)\big),
\end{aligned}
$$

where we have defined:

$$
\begin{aligned}
G_1^{(2)}(t,x;T,\xi) &= \frac{1}{2}((T-t)^2 a_2 \xi(i+\xi)\psi''(\xi) - \frac{1}{8}(T-t)^4 a_1^2 \xi^2 (i+\xi)^2 \psi'(\xi)^2 \\
&\quad - \frac{1}{6}(T-t)^3 \xi(i+\xi)(a_1^2(i+2\xi)\psi'(\xi) - 2a_2\psi'(\xi)^2 + a_1^2\xi(i+\xi)\psi''(\xi)), \\
G_2^{(2)}(t,x;T,\xi) &= \frac{1}{8}(T-t)^2(i+\xi)^2\gamma_1^2\psi'(\xi)^2 + \frac{1}{2}(T-t)^2(1-i\xi)\gamma_2\psi''(\xi) \\
&\quad + \frac{1}{6}(T-t)^3(i+\xi)(\gamma_1^2\psi'(\xi) - 2i\gamma_2\psi'(\xi)^2 + (i+\xi)\gamma_1^2\psi''(\xi)), \\
G_3^{(2)}(t,x;T,\xi) &= \frac{1}{6}(T-t)^3\xi\psi'(\xi)\int_{\mathbb{R}^2} z\nu_1(dz) + \frac{1}{3}i(T-t)^3\xi\psi'(\xi)^2\int_{\mathbb{R}} z\nu_1(dz) \\
&\quad + \frac{1}{8}(T-t)^4\xi^2\psi'(\xi)^2\int_{\mathbb{R}^2} z\nu_1(dz) + \frac{1}{2}i\xi(T-t)^2\psi''(\xi)\int_{\mathbb{R}} z\nu_1(dz) \\
&\quad + \frac{1}{6}(T-t)^3\xi^2\psi''(\xi)\int_{\mathbb{R}^2} z\nu_1(dz), \\
G_4^{(2)}(t,x;T,\xi) &= -\frac{1}{6}i(T-t)^3\psi'(\xi)\int_{\mathbb{R}}(e^{iz\xi}-1)\nu_1(dz)\int_{\mathbb{R}} ze^{iz\xi}\nu_1(dz) \\
&\quad - \frac{1}{8}(T-t)^4\psi'(\xi)^2\int_{\mathbb{R}^2}(e^{iz\xi}-1)\nu_1(dz) - \frac{1}{3}(T-t)^3\psi'(\xi)\int_{\mathbb{R}}(e^{iz\xi}-1)\nu_2(dz) \\
&\quad - \frac{1}{6}(T-t)^3\psi''(\xi)\int_{\mathbb{R}^2}(e^{iz\xi}-1)\nu_1(dz) - \frac{1}{2}(T-t)^2\psi''(\xi)\int_{\mathbb{R}}(e^{iz\xi}-1)\nu_2(dz), \\
G_5^{(2)}(t,x;T,\xi) &= \frac{1}{6}(T-t)^3\xi\psi'(\xi)\int_{\mathbb{R}^2}(e^z-1-z)\nu_1(dz) + \frac{1}{8}(T-t)^4\xi^2\psi'(\xi)^2\int_{\mathbb{R}^2}(e^z-1-z)\nu_1(dz) \\
&\quad + \frac{1}{3}i(T-t)^3\xi\psi'(\xi)\int_{\mathbb{R}}(e^z-1-z)\nu_2(dz) + \frac{1}{6}(T-t)^3\xi^2\psi''(\xi)\int_{\mathbb{R}^2}(e^z-1-z)\nu_1(dz) \\
&\quad + \frac{1}{2}i(T-t)^2\xi\psi''(\xi)\int_{\mathbb{R}}(e^z-1-z)\nu_2(dz).
\end{aligned}
$$

Essentially $G_1^{(2)}$ corresponds to the Taylor expansion of the local volatility, $G_2^{(2)}$ results from the default function, $G_3^{(2)}$, $G_4^{(2)}$ and $G_5^{(2)}$ are related to the state-dependent measure.

## References


[1] V. Bally, A. Kohatsu-Higa, A probabilistic interpretation of the parametrix method, Ann. Appl. Probab. 25 (2015) 3095–3138.

[2] A. Capponi, S. Pagliarani, T. Vargiolu, Pricing vulnerable claims in a Lévy-driven model, Finance Stoch. 18 (2014) 755–789.

[3] P. Carr, V. Linetsky, A jump to default extended CEV model: an application of Bessel processes, Finance Stoch. 10 (2006) 303–330.

[4] F. Fang, C.W. Oosterlee, A novel pricing method for European options based on Fourier-cosine series expansions, SIAM J. Sci. Comput. 31 (2008/09) 826–848.

[5] F. Fang, C.W. Oosterlee, Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions, Numer. Math. 114 (2009) 27–62.

[6] 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.

[7] A. Jacquier, M. Lorig, The smile of certain Lévy-type models, SIAM J. Financial Math. 4 (2013) 804–830.

[8] V. Linetsky, Pricing equity derivatives subject to bankruptcy, Math. Finance 16 (2006) 255–282.

[9] R. Lord, F. Fang, F. Bervoets, C.W. Oosterlee, A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes, SIAM J. Sci. Comput. 30 (2008) 1678–1705.
[10] M. Lorig, S. Pagliarani, A. Pascucci, Analytical expansions for parabolic equations, SIAM J. Appl. Math. 75 (2015) 468–491.
[11] M. Lorig, S. Pagliarani, A. Pascucci, A family of density expansions for Lévy-type processes, Ann. Appl. Probab. 25 (2015) 235–267.
[12] D. Madan, E. Seneta, The variance gamma (VG) model for share market returns, J. Bus. 63 (1990) 511–524.
[13] R. Merton, Option pricing when underlying stock returns are discontinuous, J. Financ. Econ. 3 (1976) 125–144.
[14] S. Pagliarani, A. Pascucci, C. Riga, Adjoint expansions in local Lévy models, SIAM J. Financial Math. 4 (2013) 265–296.
[15] A. Pascucci, PDE and Martingale Methods in Option Pricing, Bocconi & Springer Series, vol. 2, Springer, Bocconi University Press, Milan, 2011.
[16] L. von Sydow, L.J. Höök, E. Larsson, et al., BENCHOP—the BENCHmarking project in option pricing, Int. J. Comput. Math. 92 (2015) 2361–2379.
[17] B. Zhang, C.W. Oosterlee, Fourier cosine expansions and put-call relations for Bermudan options, in: Numerical Methods in Finance, in: Springer Proc. Math., vol. 12, Springer, Heidelberg, 2012, pp. 323–350.