Pricing Bermudan Options in Lévy Process Models*

Liming Feng† **Xiong Lin**‡

First draft: 2009-09-02; Accepted version: 2013-03-05
SIAM Journal on Financial Mathematics, 2013, 4(1), 474-493

Abstract

This paper presents a Hilbert transform method for pricing Bermudan options in Lévy process models. The corresponding optimal stopping problem can be solved using a backward induction, where a sequence of inverse Fourier and Hilbert transforms need to be evaluated. Using results from a sinc expansion based approximation theory for analytic functions, the inverse Fourier and Hilbert transforms can be approximated using very simple rules. The approximation errors decay exponentially with the number of terms used to evaluate the transforms for many popular Lévy process models. The resulting discrete approximations can be efficiently implemented using the fast Fourier transform. The early exercise boundary is obtained at the same time as the price. Accurate American option prices can be obtained by using Richardson extrapolation.

Keywords: Lévy process, Bermudan option, early exercise boundary, optimal stopping, Fourier transform, Hilbert transform, sinc methods, fast Fourier transform, analytic characteristic function

1 Introduction

Option contracts are actively traded both on exchanges and in the over-the-counter market. It is well known that the commonly used Black-Scholes-Merton option pricing model ([5], [34]) understates the likelihood of extreme price movements in financial markets. Lévy models relax the restrictive assumptions of the Black-Scholes-Merton model by allowing jumps in the underlying asset price and have become popular. Lévy models include finite activity jump-diffusion models ([28], [35]) as well as infinite activity pure jump models (e.g., [3], [11], [16], [33]).

Due to the Lévy-Khintchine formula for infinitely divisible distributions, the characteristic function of a Lévy process often admits a closed-form expression. Fourier transform based methods can thus be applied for pricing European style contracts (see, e.g., [13], [18], [20], [22], [29], [30]). However, most option contracts traded on exchanges and in the over-the-counter market are of American style and hence can be exercised early. It is thus of great interest to develop efficient methods for pricing options with early exercise features.

*This research was partially supported by the National Science Foundation under grants CMMI-0927367, CMMI-1029846.

†Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, 104 S Mathews Ave, Urbana, IL 61801, *E-mail:* fenglm@illinois.edu.

‡Gresham Investment Management, *E-mail:* afxlin9@gmail.com.

One class of options that allow early exercise are Bermudan options. In contrast to an American option, which can be exercised at any time before the option maturity, a Bermudan option can only be exercised on a discrete set of monitoring times. Although the valuation of Bermudan options is often of its own interest, it can also be considered as an approximation procedure for the valuation of American options, as the number of monitoring times increases. The pricing of perpetual Bermudan options was considered in [7]. In this paper, we consider pricing Bermudan options on a finite time horizon.

The valuation of Bermudan options corresponds to a discrete optimal stopping problem, which generally admits no closed-form solution and must be solved numerically. The optimal stopping problem can be implemented using a backward induction, where at each monitoring time, one takes the larger of a conditional expectation, representing the continuation value of the option, and the option payoff, representing the profit of immediate exercise. The main objective of the Bermudan option valuation problem is thus computing these conditional expectations.

[31] presents a least square Monte Carlo approach. This method is the most attractive for multi-dimensional applications. [10] presents a double exponential fast Gauss transform method that is very fast and accurate, with exponentially decaying pricing errors. It is however limited to Gaussian models. [19] proposes a Fourier-cosine series expansion approach that also exhibits exponentially decaying pricing errors (see also [32]). The discrete approximation is implemented using Toeplitz and Hankel matrix vector multiplications, which require five runs of the fast Fourier transform for each time step. [26] proposes a lattice method. This method is computationally intensive when the number of monitoring times is large. [21] presents an extrapolation approach, where the conditional expectation is computed by solving a partial integro-differential equation numerically. This method requires the existence of a diffusion term. [25] proposes a Fourier space time-stepping method. It uses the fact that the conditional expectation is a convolution, whose Fourier transform is a product. The algorithm proceeds as follows: starting from the option value function at a previous step, one computes its Fourier transform, multiply the result by the characteristic function of the Lévy process, compute an inverse Fourier transform to get the conditional expectation, and then compare it with the option payoff to get the option value at the current step. The method converges polynomially.

In this paper, we utilize the fact that, at each monitoring time, the option value function is divided into two parts by a critical value. On one side, it equals the option payoff. On the other side, it equals the continuation value. Taking a Fourier transform leads to a Hilbert transform representation. Our method thus proceeds as follows: knowing the Fourier transform of the option value at the previous time step, one determines the critical value, and then computes the Fourier transform of the option value at the current time step using the Hilbert transform representation. We thus do not have to recover the option value function in intermediate steps (unless desired). The implementation of this method involves inverse Fourier and Hilbert transforms of analytic functions. When a powerful approximation theory based on sinc methods is utilized, the transforms can be approximated using simple rules highly accurately, with errors decaying exponentially in $1/h$, where h is the step size used to discretize the transforms. For Lévy processes with characteristic functions with exponential tails, we might select h as a function of M, the number of terms used to evaluate the transforms. The resulting approximation error is then exponential in M, which represents the computational cost. At each step, a Toeplitz matrix vector multiplication needs to be computed. It can be implemented using only two runs of the fast Fourier transform. The total computational cost is then linear in the number of monitoring intervals, and $O(M \log(M))$ in terms of M. Finally, accurate American option prices can be obtained by using Richardson extrapolation.

The remaining of the paper is organized as follows. In Section 2, we review some basics of Lévy process models. In Section 3, we study the backward induction algorithm. In Section 4, we consider the discrete approximation of the resulting transforms. Numerical results on pricing Bermudan and American options are shown in Section 5. Section 6 concludes.

2 Lévy process models

2.1 Basics

In this section, we review some basics of Lévy processes that will be used in the paper. More details regarding Lévy processes can be found in [4], [38]. For financial modeling using Lévy processes, see [6], [14], [39].

Consider a complete and filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with the filtration $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ satisfying usual conditions. An adapted stochastic process $\{X_t, t \geq 0\}$ starting at zero is a Lévy process if it has independent and stationary increments (for $s < t$, $X_t - X_s$ is independent of $\mathcal{F}_s$ and has the same distribution as X_{t-s}), is stochastically continuous ($\forall t \geq 0$ and $\epsilon > 0$, $\lim_{s \to t} \mathbb{P}(|X_t - X_s| > \epsilon) = 0$), and has sample paths that are right continuous with left limits. By the Lévy-Khintchine theorem, the characteristic function of X_t has the following form:

$$\phi_t(\xi) = \mathbb{E}[e^{i\xi X_t}] = \exp\left\{-t\left(\frac{1}{2}\sigma^2\xi^2 - i\mu\xi + \int_{\mathbb{R}}(1 - e^{i\xi y} + i\xi y \mathbf{1}_{\{|y| \leq 1\}})\Pi(dy)\right)\right\}.$$

The triplet (μ, σ^2, Π) characterizes the drift, the Brownian motion and the pure jump components of a Lévy process. If $\Pi = 0$, X becomes a Brownian motion with drift μ and variance rate σ^2. Define

$$\mathcal{I}_X = \{\alpha \in \mathbb{R} : \mathbb{E}[e^{-\alpha X_t}] < \infty, \forall t > 0\}.$$

$\mathcal{I}_X$ is an interval containing the origin with endpoints $\lambda_- \leq 0$ and $\lambda_+ \geq 0$ (see [38]). If $\lambda_- < \lambda_+$, then the characteristic function $\phi_t(z)$, as a function of the complex variable z, is analytic in the strip $\mathcal{D}_{(\lambda_-, \lambda_+)} = \{z \in \mathbb{C} : \Im(z) \in (\lambda_-, \lambda_+)\}$. Here $\Im(z)$ denotes the imaginary part of z.

2.2 Esscher transform

It is often convenient to change the measure associated with a Lévy process. For a Lévy process X_t with $\mathbb{E}[e^{-\alpha X_t}] = \phi_t(i\alpha) < \infty$ (i.e., $\alpha \in \mathcal{I}_X$), define

$$Z_t = \frac{e^{-\alpha X_t}}{\phi_t(i\alpha)}. \tag{2.1}$$

Then Z_t is a positive martingale with constant expectation 1. It defines a new probability measure $\mathbb{P}^\alpha$ via the following Radon-Nikodým derivative:

$$\frac{d\mathbb{P}^\alpha}{d\mathbb{P}}|\mathcal{F}_t = Z_t. \tag{2.2}$$

The above measure change is known as an Esscher transform or exponential tilting. An Esscher transformed Lévy process is still a Lévy process (see [36]). The characteristic function of X_t under the new measure $\mathbb{P}^\alpha$ is given by

$$\phi_t^\alpha(\xi) = \mathbb{E}^\alpha[e^{i\xi X_t}] = \mathbb{E}[e^{i\xi X_t} Z_t] = \frac{1}{\phi_t(i\alpha)}\mathbb{E}[e^{i(\xi + i\alpha)X_t}] = \frac{\phi_t(\xi + i\alpha)}{\phi_t(i\alpha)}. \tag{2.3}$$

Here $\mathbb{E}^\alpha$ denotes the expectation under measure $\mathbb{P}^\alpha$. Moreover, for any $0 \le s < t$ and $Y_t \in \mathcal{F}_t$ (see [36]),

$$\mathbb{E}^\alpha[Y_t|\mathcal{F}_s] = \mathbb{E}[Y_t Z_t/Z_s|\mathcal{F}_s]. \tag{2.4}$$

2.3 Geometric Lévy models

We assume that the price of the asset underlying the Bermudan option is governed by a geometric Lévy process under a given equivalent martingale measure $\mathbb{P}$:

$$S_t = S_0 e^{X_t}.$$

Here the log return process $X_t = \ln(S_t/S_0)$ is a Lévy process starting at the origin at time 0. S_0 is the initial asset price. Under the equivalent martingale measure, the discounted gain process is a martingale. That is, the following martingale condition is satisfied:

$$\mathbb{E}[S_t] = S_0\mathbb{E}[e^{X_t}] = S_0\phi_t(-i) = S_0 e^{(r-q)t}. \tag{2.5}$$

Here r is the risk free interest rate, and q is the continuous yield the underlying asset is paying. This fixes the drift parameter μ of the Lévy process:

$$\mu = r - q - \frac{1}{2}\sigma^2 + \int_{\mathbb{R}} (1 - e^y + y\mathbf{1}_{\{|y|\le 1\}})\Pi(dy).$$

For (2.5) to hold, $\alpha = -1 \in \mathcal{I}_X$. Therefore, in the remaining of the paper, we assume that $[-1, 0] \in (\lambda_-, \lambda_+)$. For the analysis of our algorithm, we need to make further assumptions about the Lévy process. We summarize them in the following:

Assumption 2.1. *The Lévy process X satisfies the following conditions:*

1. $[-1, 0] \in (\lambda_-, \lambda_+)$.
2. *For any $\alpha \in (\lambda_-, \lambda_+)$ and $t > 0$, $\xi\phi_t(\xi + i\alpha) \in L^1(\mathbb{R})$. For any $[d_-, d_+] \subset (\lambda_-, \lambda_+)$,*

$$|\xi| \int_{d_-}^{d_+} |\phi_t(\xi + iy)|dy \to 0, \quad \xi \to \pm\infty.$$

3. *For any $t > 0$, the support of the transition probability density of the Lévy process is $\mathbb{R}$.*

Note that $\phi_t(\xi+i\alpha)$ is bounded on $\mathbb{R}$. So $\xi\phi_t(\xi+i\alpha) \in L^1(\mathbb{R})$ also implies that $\phi_t(\xi+i\alpha) \in L^1(\mathbb{R})$. It will become clear as we proceed when these conditions are used.

3 Bermudan options valuation

In this paper, we focus on Bermudan put options. Calls on assets with positive dividend yields can be handled similarly. Suppose the current time is $t_0 = 0$. Denote the set of monitoring times for the Bermudan option by $\mathcal{T} = \{t_0, t_1, \cdots, t_N\}$. Here $t_N = T$ is the option maturity. In general, the intervals between monitoring times can be different. However, to simplify notations and without loss of generality, we assume a constant monitoring interval Δ. That is, $t_j = j\Delta$, $j = 0, 1, \cdots, N$. A Bermudan option can be exercised at any time in $\mathcal{T}$. When exercised at time $t_n \in \mathcal{T}$, the payoff of the option is $G(S_{t_n})$, where $G(S) = (K - S)^+$ for a put. Here K is the strike price of the option.

3.1 Backward induction

The valuation of a Bermudan option corresponds to solving a discrete optimal stopping problem. The value at time 0 of the option is given by

$$V^0(S_0) = \sup_{\tau} \mathbb{E}[e^{-r\tau} G(S_\tau)],$$

where the supremum is taken over the set of all stopping times which take values in $\mathcal{T}$. Bermudan options pricing has been studied in [40] under very general settings. It is shown that the above optimal stopping problem can be solved with the following backward induction:

$$V^N(S) = G(S), \tag{3.1a}$$

$$V^j(S) = \max(G(S), e^{-r\Delta}\mathbb{E}_{j\Delta,S}[V^{j+1}(S_{(j+1)\Delta})]), \quad 0 \le j \le N-1, \tag{3.1b}$$

where $\mathbb{E}_{j\Delta,S}$ denotes the expectation conditional on $S_{j\Delta} = S$.

In the following, we first show that, for any $0 \le j \le N-1$, there exists a unique value $0 < S_j^* < K$ where $G(S)$ and $e^{-r\Delta}\mathbb{E}_{j\Delta,S}[V^{j+1}(S_{(j+1)\Delta})]$ cross. On the left of S_j^*, $V^j(S) = G(S)$. On the right, $V^j(S)$ is given by the discounted conditional expectation. Namely, when the underlying asset price at time $j\Delta$ is below S_j^*, it is optimal to exercise early. If the asset price is above S_j^*, it is optimal to further hold the option. The existence of a unique $S_j^* \in (0, K)$ is critical for us to develop the Hilbert transform approach for pricing Bermudan options. For $j = N$, we define $S_j^* = K$. That is, at option maturity, we either exercise if the option ends in the money or do not exercise and let the option expire. The collection $\{S_j^*, 0 \le j \le N\}$ is known as the early exercise boundary. In practice, in addition to determining the option value, we also want to determine the early exercise boundary.

Theorem 3.1. *Under Assumption 2.1, for any $0 \le j \le N-1$, there exists a unique $0 < S_j^* < K$ so that V^j defined in (3.1) satisfies*

$$V^j(S) = G(S) \cdot \mathbf{1}_{(0,S_j^*]}(S) + e^{-r\Delta}\mathbb{E}_{j\Delta,S}[V^{j+1}(S_{(j+1)\Delta})] \cdot \mathbf{1}_{(S_j^*,\infty)}(S). \tag{3.2}$$

Proof. For $0 \le j \le N-1$, denote

$$C^j(S) = e^{-r\Delta}\mathbb{E}_{j\Delta,S}[V^{j+1}(S_{(j+1)\Delta})].$$

We examine the derivative of $C^j(S)$ with respect to $S > 0$. Denote the transition probability density of the Lévy process X by $p_t(x)$. For $j = N-1$, we have

$$\begin{aligned}
\frac{d}{dS}C^{N-1}(S) &= \frac{d}{dS}\int_{\mathbb{R}} e^{-r\Delta}V^N(Se^x)p_\Delta(x)dx \\
&= \frac{d}{dS}\int_{-\infty}^{\ln\frac{K}{S}} e^{-r\Delta}(K - Se^x)p_\Delta(x)dx \\
&= -e^{-r\Delta}(K - Se^{\ln(K/S)})p_\Delta(\ln(K/S))/S - e^{-r\Delta}\int_{-\infty}^{\ln\frac{K}{S}} e^x p_\Delta(x)dx \\
&> -e^{-r\Delta}\int_{\mathbb{R}} e^x p_\Delta(x)dx \\
&= -e^{-r\Delta}\mathbb{E}[e^{X_\Delta}] \;=\; -e^{-q\Delta} \;\ge\; -1.
\end{aligned}$$

The strict inequality holds in the above since the support of $p_\Delta(x)$ is the whole real line by Assumption 2.1. We have also used the martingale condition (2.5): $\mathbb{E}[e^{X_\Delta}] = e^{(r-q)\Delta}$. It is also clear that $C^{N-1}(S) > 0$ for any $S > 0$. Since $\lim_{S\to 0+} C^{N-1}(S) = Ke^{-r\Delta}$, there must exist a unique $0 < S^*_{N-1} < K$ that solves $G(S) = C^{N-1}(S)$, and

$$
\begin{aligned}
V^{N-1}(S) &= \max(G(S), C^{N-1}(S)) \\
&= G(S)\cdot \mathbf{1}_{(0,S^*_{N-1}]}(S) + e^{-r\Delta}\mathbb{E}_{j\Delta,S}[V^N(S_{(j+1)\Delta})]\cdot \mathbf{1}_{(S^*_{N-1},\infty)}(S).
\end{aligned}
$$

Now suppose for some $0 \le j \le N-2$, $dC^{j+1}(S)/dS > -1$ for any $S > 0$, and there exists $0 < S^*_{j+1} < K$ that solves $G(S) = C^{j+1}(S)$ such that $V^{j+1}(S) = C^{j+1}(S)$ for $S > S^*_{j+1}$ and $V^{j+1}(S) = G(S)$ for $0 < S \le S^*_{j+1}$. We examine the derivative of $C^j(S)$ with respect to $S > 0$:

$$
\begin{aligned}
\frac{d}{dS}C^j(S) &= \frac{d}{dS}\int_{\mathbb{R}} e^{-r\Delta}V^{j+1}(Se^x)p_\Delta(x)dx \\
&= \frac{d}{dS}\int_{-\infty}^{\ln\frac{S^*_{j+1}}{S}} e^{-r\Delta}(K - Se^x)p_\Delta(x)dx + \frac{d}{dS}\int_{\ln\frac{S^*_{j+1}}{S}}^{\infty} e^{-r\Delta}C^{j+1}(Se^x)p_\Delta(x)dx \\
&= -e^{-r\Delta}(K - S^*_{j+1})p_\Delta(\ln\frac{S^*_{j+1}}{S})\frac{1}{S} - e^{-r\Delta}\int_{-\infty}^{\ln\frac{S^*_{j+1}}{S}} e^x p_\Delta(x)dx \\
&\quad + e^{-r\Delta}C^{j+1}(S^*_{j+1})p_\Delta(\ln\frac{S^*_{j+1}}{S})\frac{1}{S} + e^{-r\Delta}\int_{\ln\frac{S^*_{j+1}}{S}}^{\infty} \frac{dC^{j+1}}{dS}\Big|_{Se^x} e^x p_\Delta(x)dx \\
&> -e^{-r\Delta}\int_{-\infty}^{\ln\frac{S^*_{j+1}}{S}} e^x p_\Delta(x)dx - e^{-r\Delta}\int_{\ln\frac{S^*_{j+1}}{S}}^{\infty} e^x p_\Delta(x)dx \\
&= -e^{-r\Delta}\mathbb{E}[e^{X_\Delta}] \;=\; -e^{-q\Delta} \;\ge\; -1.
\end{aligned}
$$

Note that we have used $G(S^*_{j+1}) = C^{j+1}(S^*_{j+1})$ in the above. Again, $\lim_{S\to 0+} C^j(S) = Ke^{-r\Delta}$ and $C^j(S) > 0$ for any $S > 0$. Therefore, there must exist a unique $0 < S^*_j < K$ that solves $G(S) = C^j(S)$ such that (3.2) holds. The proof is finished by induction. □

We perform a change of variable $x = \ln(S/K)$ and consider the log process $X_t = \ln(S_t/K)$. Denote $f^0(x) = V^0(Ke^x)$. This is the option value function at time 0 in the new state variable x. Denote $g(x) = G(Ke^x)$, and $x^*_j = \ln(S^*_j/K) < 0$, $0 \le j \le N-1$. Then (3.1) and Theorem 3.1 lead to the following backward induction for solving f^0:

$$f^N(x) = g(x), \tag{3.3a}$$

for any $0 \le j \le N-1$, x^*_j solves

$$g(x) = e^{-r\Delta}\mathbb{E}_{j\Delta,x}[f^{j+1}(X_{(j+1)\Delta})], \tag{3.3b}$$

$$f^j(x) = g(x)\cdot \mathbf{1}_{(-\infty,x^*_j]}(x) + e^{-r\Delta}\mathbb{E}_{j\Delta,x}[f^{j+1}(X_{(j+1)\Delta})]\cdot \mathbf{1}_{(x^*_j,\infty)}(x). \tag{3.3c}$$

Here $\mathbb{E}_{t,x}$ denotes the expectation conditional on $X_t = x$. The option price at time 0 is then recovered by $V^0(S_0) = f^0(\ln(S_0/K))$.

We will implement the above backward induction using Fourier transform based methods. To guarantee that the Fourier transforms are well defined, we need to introduce an exponential dampening factor. For $\alpha \in \mathbb{R}$, we define $f_\alpha^j(x) = e^{\alpha x} f^j(x)$. Denote $g_\alpha(x) = e^{\alpha x} g(x)$. As the following theorem shows, for an appropriately selected α, $f_\alpha^j \in L^1(\mathbb{R})$ for any $0 \le j \le N$ and its Fourier transform is hence well defined.

Theorem 3.2. *Let $\alpha \in \mathcal{I}_X$ be such that $g_\alpha(x) \in L^1(\mathbb{R})$. Then $f_\alpha^j(x) \in L^1(\mathbb{R})$ for any $0 \le j \le N$ and solves the following backward induction:*

$$f_\alpha^N(x) = g_\alpha(x), \tag{3.4a}$$

for any $0 \le j \le N-1$, x_j^* solves

$$g_\alpha(x) = e^{-r\Delta} \phi_\Delta(i\alpha) \mathbb{E}^\alpha_{j\Delta,x}[f_\alpha^{j+1}(X_{(j+1)\Delta})], \tag{3.4b}$$

$$\begin{aligned} f_\alpha^j(x) &= g_\alpha(x) \cdot \mathbf{1}_{(-\infty, x_j^*]}(x) \\ &\quad + e^{-r\Delta} \phi_\Delta(i\alpha) \mathbb{E}^\alpha_{j\Delta,x}[f_\alpha^{j+1}(X_{(j+1)\Delta})] \cdot \mathbf{1}_{(x_j^*, \infty)}(x), \end{aligned} \tag{3.4c}$$

where $\mathbb{E}^\alpha$ denotes the expectation under measure $\mathbb{P}^\alpha$, which is defined in (2.2).

Proof. We first obtain the following from (2.1) and (2.4):

$$\begin{aligned} \mathbb{E}[f^{j+1}(X_{(j+1)\Delta}) | \mathcal{F}_{j\Delta}] &= \mathbb{E}[e^{-\alpha X_{(j+1)\Delta}} f_\alpha^{j+1}(X_{(j+1)\Delta}) | \mathcal{F}_{j\Delta}] \\ &= \phi_{(j+1)\Delta}(i\alpha) Z_{j\Delta}\ \mathbb{E}[Z_{(j+1)\Delta} f_\alpha^{j+1}(X_{(j+1)\Delta}) / Z_{j\Delta} | \mathcal{F}_{j\Delta}] \\ &= \frac{\phi_{(j+1)\Delta}(i\alpha)}{\phi_{j\Delta}(i\alpha)} e^{-\alpha X_{j\Delta}}\ \mathbb{E}^\alpha[f_\alpha^{j+1}(X_{(j+1)\Delta}) | \mathcal{F}_{j\Delta}]. \end{aligned}$$

Note that $\phi_{(j+1)\Delta}(i\alpha)/\phi_{j\Delta}(i\alpha) = \phi_\Delta(i\alpha)$. Therefore, conditional on $X_{j\Delta} = x$, we have

$$e^{\alpha x} \mathbb{E}_{j\Delta,x}[f^{j+1}(X_{(j+1)\Delta})] = \phi_\Delta(i\alpha) \mathbb{E}^\alpha_{j\Delta,x}[f_\alpha^{j+1}(X_{(j+1)\Delta})].$$

We therefore obtain the dampened backward induction (3.4). Note that $\mathbb{E}^\alpha_{j\Delta,x}[f_\alpha^{j+1}(X_{(j+1)\Delta})]$ as a function of x is simply a convolution of f_α^{j+1} and the density of the Lévy process. It is in $L^1(\mathbb{R})$ as long as $f_\alpha^{j+1} \in L^1(\mathbb{R})$. Since $f_\alpha^N = g_\alpha \in L^1(\mathbb{R})$, and the first term in (3.4c) is also in $L^1(\mathbb{R})$, we immediately obtain by induction that f_α^j is in $L^1(\mathbb{R})$ for all $0 \le j \le N$. $\square$

The above results enable us to implement the backward induction using a Fourier transform based method. We are then able to take advantage of the powerful approximation theory for analytic functions. In the following, we reformulate the backward induction (3.4) using inverse Fourier and Hilbert transforms.

3.2 Hilbert transform representations

The Hilbert transform of a function $f \in L^1(\mathbb{R})$ is well defined by the following Cauchy principal value integral (see [27]):

$$\mathcal{H}f(x) = \frac{1}{\pi}\ p.v. \int_{\mathbb{R}} \frac{f(y)}{x-y} dy.$$

Denote the Fourier transform of f by $\hat{f}$:

$$\hat{f}(\xi) = \mathcal{F}f(\xi) = \int_{\mathbb{R}} e^{i\xi x} f(x) dx.$$

If $\hat{f} \in L^1(\mathbb{R})$, we have the following for any $l \in \mathbb{R}$ (see [20]):

$$\mathcal{F}(\mathbf{1}_{(l,\infty)} \cdot f)(\xi) = \frac{1}{2}\hat{f}(\xi) + \frac{i}{2}e^{i\xi l}\mathcal{H}(e^{-i\eta l}\hat{f}(\eta))(\xi). \tag{3.5}$$

This relation can be used to establish the following backward induction for implementing (3.4) in terms of inverse Fourier and Hilbert transforms.

Theorem 3.3. *Under Assumption 2.1, the backward induction (3.4) can be solved as below:*

$$\hat{f}_\alpha^N(\xi) = \hat{g}_\alpha(\xi), \tag{3.6a}$$

for any $1 \le j \le N-1$, x_j^* *solves*

$$\frac{1}{2\pi}e^{-r\Delta}\int_{\mathbb{R}} e^{-i\xi x}\hat{f}_\alpha^{j+1}(\xi)\phi_\Delta(-\xi+i\alpha)d\xi - g_\alpha(x) = 0, \tag{3.6b}$$

$$\begin{aligned}\hat{f}_\alpha^j(\xi) \quad = \quad & \mathcal{F}(g_\alpha \cdot \mathbf{1}_{(-\infty,x_j^*]})(\xi) + e^{-r\Delta}\Bigg(\frac{1}{2}\hat{f}_\alpha^{j+1}(\xi)\phi_\Delta(-\xi+i\alpha) \\ & +\frac{i}{2}e^{i\xi x_j^*}\mathcal{H}(e^{-i\eta x_j^*}\hat{f}_\alpha^{j+1}(\eta)\phi_\Delta(-\eta+i\alpha))(\xi)\Bigg),\end{aligned} \tag{3.6c}$$

$$f_\alpha^0(x) = \max\left(g_\alpha(x), \frac{1}{2\pi}e^{-r\Delta}\int_{\mathbb{R}} e^{-i\xi x}\hat{f}_\alpha^1(\xi)\phi_\Delta(-\xi+i\alpha)d\xi\right). \tag{3.6d}$$

Proof. We first notice that the conditional expectation in (3.4b) and (3.4c) is essentially a convolution integral. Denote the transition density of X under measure $\mathbb{P}^\alpha$ by $p_t^\alpha(\cdot)$. Then

$$\mathbb{E}^\alpha_{j\Delta,x}[f_\alpha^{j+1}(X_{(j+1)\Delta})] = \int_{\mathbb{R}} f_\alpha^{j+1}(y)p_\Delta^\alpha(y-x)dy.$$

By the convolution theorem, the Fourier transform of the above is a product $\hat{f}_\alpha^{j+1}(\xi)\phi_\Delta^\alpha(-\xi)$, where ϕ_Δ^α is defined in (2.3). Note that $\hat{f}_\alpha^{j+1}$ is bounded on $\mathbb{R}$. By Assumption 2.1, $\phi_\Delta^\alpha \in L^1(\mathbb{R})$. We thus have $\hat{f}_\alpha^{j+1}(\xi)\phi_\Delta^\alpha(-\xi) \in L^1(\mathbb{R})$. This immediately leads to the inverse Fourier transform representation in (3.6b). Taking Fourier transform on both sides of (3.4c) and using (3.5), we obtain (3.6c). Finally, using the inverse Fourier transform representation of $\mathbb{E}^\alpha_{\Delta,x}[f_\alpha^1(X_\Delta)]$, we obtain the last equation (3.6d). $\square$

Note that in the last step, we compute f_α^0 from $\hat{f}_\alpha^1$ through an inverse Fourier transform. When needed, (3.6b) can be run one more time for $j=0$ to obtain x_0^*.

4 Discrete approximation

4.1 Sinc methods

In the backward induction (3.6), we need to evaluate the inverse Fourier transforms in (3.6b) and (3.6d), and the Hilbert transforms in (3.6c). When the analyticity of the characteristic function ϕ is utilized, the inverse Fourier and Hilbert transforms can be evaluated using very simple rules highly accurately. The rules are based on sinc expansion approximation of analytic functions [20, 41]. We call the resulting methods sinc methods.

We first introduce the analytic class previously mentioned. Let $\mathcal{D}_{(d_-,d_+)} = \{z \in \mathbb{C} : \Im(z) \in (d_-, d_+)\}$ for some $-\infty < d_- < 0 < d_+ < +\infty$. A function $\hat{f}$ is in $H(\mathcal{D}_{(d_-,d_+)})$ if it is analytic in $\mathcal{D}_{(d_-,d_+)}$, and satisfies

$$\int_{d_-}^{d_+} |\hat{f}(\xi + iy)|dy \to 0, \ \xi \to \pm\infty, \ ||\hat{f}||^{\pm} := \int_{\mathbb{R}} |\hat{f}(\xi + id_{\pm})|d\xi < +\infty.$$

Suppose $\hat{f} \in L^1(\mathbb{R}) \cap H(\mathcal{D}_{(d_-,d_+)})$. Let's consider the inverse Fourier transform of $\hat{f}$:

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi x} \hat{f}(\xi) d\xi.$$

We would like to numerically evaluate f. It turns out that the simplest trapezoidal sum approximation is highly accurate, with exponentially decaying discretization errors. More specifically, we approximate $f(x)$ by

$$f_{h,M}(x) = \frac{1}{2\pi} \sum_{m=-M}^{M} e^{-ixmh} \hat{f}(mh)h.$$

Here h is the step size, and M is the truncation level. Then we have the following result. The proof can be adapted from [20] and is therefore omitted.

Proposition 4.1. *Suppose $\hat{f} \in L^1(\mathbb{R}) \cap H(\mathcal{D}_{(d_-,d_+)})$. If $|\hat{f}(\xi)| \leq \kappa|\xi|^n \exp(-c|\xi|^{\nu})$ for any $\xi \in \mathbb{R}$ for some $\kappa, c, \nu > 0$ and $n \geq 0$, then for any $h > 0, M \geq 1$ such that $Mh > (\frac{n}{c\nu})^{1/\nu}$,*

$$\begin{aligned} |f(x) - f_{h,M}(x)| \ &\leq \ \frac{e^{-2\pi|d_-|/h + xd_-}}{2\pi(1 - e^{-2\pi|d_-|/h})}||\hat{f}||^- + \frac{e^{-2\pi d_+/h + xd_+}}{2\pi(1 - e^{-2\pi d_+/h})}||\hat{f}||^+ \\ &\quad + \frac{\kappa}{\pi\nu c^{(n+1)/\nu}} \Gamma\Big(\frac{n+1}{\nu}, c(Mh)^{\nu}\Big), \end{aligned}$$

where $\Gamma(\cdot, \cdot)$ is the incomplete gamma function. If $|\hat{f}(\xi)| \leq \kappa|\xi|^{-\nu-1}$ for any $\xi \in \mathbb{R}$ for some $\kappa, \nu > 0$, then

$$|f(x) - f_{h,M}(x)| \leq \frac{e^{-2\pi|d_-|/h + xd_-}}{2\pi(1 - e^{-2\pi|d_-|/h})}||\hat{f}||^- + \frac{e^{-2\pi d_+/h + xd_+}}{2\pi(1 - e^{-2\pi d_+/h})}||\hat{f}||^+ + \frac{\kappa}{\pi\nu}(Mh)^{-\nu}.$$

Note that the dominating term in the incomplete gamma function for large Mh is $\exp(-c(Mh)^{\nu})$. The above result therefore shows that the discretization error of approximating the inverse Fourier transform integral by an infinite trapezoidal sum with step size h converges exponentially in $1/h$. When $\hat{f}$ has exponential tails, the truncation error also decays exponentially with Mh. When the tail behaviors of $\hat{f}$ are different, one only needs to adjust the last term, which represents the truncation error. For example, if $\hat{f}$ has polynomial tails, then the truncation error decays to zero polynomially in Mh. Even when the tails of $\hat{f}$ are polynomial, a rather large h is often sufficient due to the exponential convergence of the discretization error. Consequently, one may be able to use a relatively smaller M when controlling the truncation error.

In this paper, we also need to compute the derivative of $f(x)$ with respect to x. We further assume that $\xi\hat{f}(\xi) \in L^1(\mathbb{R}) \cap H(\mathcal{D}_{(d_-,d_+)})$. Then

$$f'(x) = \frac{d}{dx}f(x) = -\frac{i}{2\pi}\int_{\mathbb{R}} e^{-i\xi x}\xi\hat{f}(\xi)d\xi.$$

The interchange of the order of the differentiation and integration is valid by the integrability of $\xi\hat{f}(\xi)$. We approximate $f'(x)$ similarly using the trapezoidal rule:

$$f'_{h,M}(x) = -\frac{i}{2\pi}\sum_{m=-M}^{M} e^{-ixmh}mh\hat{f}(mh)h.$$

Proposition 4.1 then still applies, with $||\hat{f}||^{\pm}$ replaced by $||\xi\hat{f}(\xi)||^{\pm}$, and n replaced by $n+1$.

We also need to evaluate Hilbert transforms of the following form:

$$\hat{g}(\xi) = \mathcal{H}(e^{-i\eta x}\hat{f}(\eta))(\xi).$$

Similarly, the above can be discretized using very simple rules with exponentially decaying discretization errors. Namely, we approximate $\hat{g}(\xi)$ by

$$\hat{g}_{h,M}(\xi) = \sum_{m=-M}^{M} e^{-ixmh}\hat{f}(mh)\frac{1-\cos(\pi(\xi-mh)/h)}{\pi(\xi-mh)/h}. \tag{4.1}$$

Then we have the following error estimates for the above approximation. Again, the proof can be adapted from [20] and is omitted.

Proposition 4.2. *Suppose $\hat{f} \in L^1(\mathbb{R}) \cap H(\mathcal{D}_{(d_-,d_+)})$. If $|\hat{f}(\xi)| \le \kappa\exp(-c|\xi|^{\nu})$ for any $\xi \in \mathbb{R}$ for some $\kappa, c, \nu > 0$, then for any $h > 0, M \ge 1$,*

$$\begin{aligned} |\hat{g}(\xi) - \hat{g}_{h,M}(\xi)| \;\; \le \;\; & \frac{e^{-\pi|d_-|/h + xd_-}}{\pi|d_-|(1-e^{-\pi|d_-|/h})}||\hat{f}||^- + \frac{e^{-\pi d_+/h + xd_+}}{\pi d_+(1-e^{-\pi d_+/h})}||\hat{f}||^+ \\ & + \frac{2\kappa}{\nu h c^{1/\nu}}\Gamma\Big(\frac{1}{\nu}, c(Mh)^{\nu}\Big). \end{aligned}$$

If $|\hat{f}(\xi)| \le \kappa|\xi|^{-\nu-1}$ for any $\xi \in \mathbb{R}$ for some $\kappa, \nu > 0$, then

$$|\hat{g}(\xi) - \hat{g}_{h,M}(\xi)| \le \frac{e^{-\pi|d_-|/h + xd_-}}{\pi|d_-|(1-e^{-\pi|d_-|/h})}||\hat{f}||^- + \frac{e^{-\pi d_+/h + xd_+}}{\pi d_+(1-e^{-\pi d_+/h})}||\hat{f}||^+ + \frac{2\kappa}{\nu h}(Mh)^{-\nu}.$$

In the following section, we utilize Propositions 4.1 and 4.2, and present a detailed algorithm for pricing Bermudan put options. We also discuss the selection of h in terms of M, and efficient implementation of the resulting discrete approximations using the fast Fourier transform.

4.2 Implementation

For a Bermudan put option, $g_\alpha(x) = Ke^{\alpha x}(1-e^x)^+ \in L^1(\mathbb{R})$ for any $\alpha > 0$. According to Theorem 3.2, the dampening factor should be $\alpha \in (0, \lambda_+)$. Our algorithm for (3.6) therefore goes as follows:

Step 1

We start with (3.6a), which becomes

$$\hat{f}^N_\alpha(\xi) = \frac{K}{(\alpha + i\xi)(\alpha + 1 + i\xi)}, \quad \xi = -Mh, \cdots, Mh.$$

Step 2

We solve (3.6b) using the Newton-Raphson method. An initial guess is needed when using the Newton-Raphson method. When solving for x^*_{N-1}, we start with the initial guess $x^*_N = 0$. When solving for $x^*_j, j = N-2, \cdots, 1$, we start with the initial guess x^*_{j+1}. Denote the initial guess by $\tilde{x}$. We evaluate the left hand side of (3.6b) and its derivatives with respect to x at the point $\tilde{x}$ using the trapezoidal rule with step size h and truncation level M. Denote these values by $f(\tilde{x})$ and $f'(\tilde{x})$:

$$\begin{aligned} f(\tilde{x}) &= \frac{1}{2\pi} \sum_{m=-M}^{M} e^{-i\tilde{x}mh} \hat{f}^{j+1}_\alpha(mh)\phi_\Delta(-mh + i\alpha)h - g_\alpha(\tilde{x}), \\ f'(\tilde{x}) &= -\frac{i}{2\pi} \sum_{m=-M}^{M} e^{-i\tilde{x}mh} mh \hat{f}^{j+1}_\alpha(mh)\phi_\Delta(-mh + i\alpha)h - g'_\alpha(\tilde{x}). \end{aligned}$$

where $g'_\alpha(x) = Ke^{\alpha x}(\alpha - (1+\alpha)e^x)$. We replace $\tilde{x}$ by $\tilde{x} - f(\tilde{x})/f'(\tilde{x})$ and repeat as long as $|f(\tilde{x})/f'(\tilde{x})| > \epsilon_{NR}$ for some predetermined tolerance level ϵ_{NR}.

Step 3

The first term in (3.6c) is given by

$$\mathcal{F}(g_\alpha \cdot \mathbf{1}_{(-\infty, x^*_j]})(\xi) = Ke^{(i\xi+\alpha)x^*_j}\Big(\frac{1}{i\xi + \alpha} - \frac{e^{x^*_j}}{i\xi + \alpha + 1}\Big).$$

Using (4.1), we compute $\hat{f}^j_\alpha(\xi)$ as follows:

$$\begin{aligned} \hat{f}^j_\alpha(\xi) &= \mathcal{F}(g_\alpha \cdot \mathbf{1}_{(-\infty, x^*_j]})(\xi) + e^{-r\Delta}\Bigg(\frac{1}{2}\hat{f}^{j+1}_\alpha(\xi)\phi_\Delta(-\xi + i\alpha) \\ &\quad + \frac{i}{2} e^{i\xi x^*_j} \sum_{m=-M}^{M} e^{-ix^*_j mh} \hat{f}^{j+1}_\alpha(mh)\phi_\Delta(-mh + i\alpha)\frac{1 - \cos(\pi(\xi - mh)/h)}{\pi(\xi - mh)/h}\Bigg), \quad \xi = -Mh, \cdots, Mh. \end{aligned}$$

Repeat Steps 2 and 3 for $j = N-1, \cdots, 1$

Step 4

The inverse Fourier transform in (3.6d) is again approximated using the trapezoidal rule:

$$f^0_\alpha(x) \approx \max\Bigg(g_\alpha(x), \frac{1}{2\pi}e^{-r\Delta} \sum_{m=-M}^{M} e^{-ixmh} \hat{f}^1_\alpha(mh)\phi_\Delta(-mh + i\alpha)h\Bigg).$$

End

In Step 2 and 4, we need to evaluate inverse Fourier transforms of $\hat{f}_\alpha^{j+1}(\xi)\phi_\Delta(-\xi+i\alpha)$ and $\hat{f}_\alpha^{j+1}(\xi)\xi\phi_\Delta(-\xi+i\alpha)$. In Step 3, we need to evaluate Hilbert transforms involving $\hat{f}_\alpha^{j+1}(\xi)\phi_\Delta(-\xi+i\alpha)$. To use Propositions 4.1 and 4.2, we need to show that $\hat{f}_\alpha^{j+1}(\xi)\phi_\Delta(-\xi+i\alpha)$ and $\hat{f}_\alpha^{j+1}(\xi)\xi\phi_\Delta(-\xi+i\alpha)$ are in our analytic class. This is shown in the following. Recall that ϕ_Δ is analytic in $\mathcal{D}_{(\lambda_-,\lambda_+)}$.

Theorem 4.3. *Let $\alpha \in (0,\lambda_+)$. Under Assumption 2.1, for any $1 \le j \le N$ and $\alpha-\lambda_+ < d_- < 0 < d_+ < \alpha$, both $\hat{f}_\alpha^{j}(z)\phi_\Delta(-z+i\alpha)$ and $\hat{f}_\alpha^{j}(z)z\phi_\Delta(-z+i\alpha)$ are in $H(\mathcal{D}_{(d_-,d_+)})$.*

Proof. Since $\phi_\Delta(z)$ is analytic in $\mathcal{D}_{(\lambda_-,\lambda_+)}$, we have that $\phi_\Delta(-z+i\alpha)$ is analytic in $\mathcal{D}_{(\alpha-\lambda_+,\alpha-\lambda_-)}$. As for $\hat{f}_\alpha^j$, according to Theorem 3.2, for any $\beta \in (\alpha-\lambda_+,\alpha)$ such that $\alpha-\beta \in (0,\lambda_+)$, we have $e^{-\beta x}f_\alpha^j(x) = f_{\alpha-\beta}^j(x) \in L^1(\mathbb{R})$. By analytic continuation of the Fourier transform [37], $\hat{f}_\alpha^j$ is analytic in $\mathcal{D}_{(\alpha-\lambda_+,\alpha)}$. Therefore, both $\hat{f}_\alpha^{j}(z)\phi_\Delta(-z+i\alpha)$ and $\hat{f}_\alpha^{j}(z)z\phi_\Delta(-z+i\alpha)$ are analytic in $\mathcal{D}_{(d_-,d_+)}$. Moreover,

$$\begin{aligned}
\int_{d_-}^{d_+} |\hat{f}_\alpha^j(\xi+iy)\phi_\Delta(-\xi-iy+i\alpha)|dy &= \int_{d_-}^{d_+} |\hat{f}_{\alpha-y}^j(\xi)\phi_\Delta(-\xi-iy+i\alpha)|dy \\
&\le \int_{d_-}^{d_+} ||f_{\alpha-y}^j||_{L^1(\mathbb{R})}|\phi_\Delta(-\xi-iy+i\alpha)|dy \\
&\le \sup_{y\in[d_-,d_+]} ||f_{\alpha-y}^j||_{L^1(\mathbb{R})} \int_{d_-}^{d_+} |\phi_\Delta(-\xi-iy+i\alpha)|dy \\
&= \sup_{y\in[d_-,d_+]} ||f_{\alpha-y}^j||_{L^1(\mathbb{R})} \int_{\alpha-d_+}^{\alpha-d_-} |\phi_\Delta(-\xi+iy)|dy.
\end{aligned}$$

Note that $[\alpha-d_+,\alpha-d_-] \subset (0,\lambda_+)$. According to Theorem 3.2, $\sup_{y\in[d_-,d_+]} ||f_{\alpha-y}^j||_{L^1(\mathbb{R})} < +\infty$. By the second condition in Assumption 2.1, $\int_{\alpha-d_+}^{\alpha-d_-} |\phi_\Delta(-\xi+iy)|dy \to 0$ as $\xi \to \pm\infty$. We thus have

$$\int_{d_-}^{d_+} |\hat{f}_\alpha^j(\xi+iy)\phi_\Delta(-\xi-iy+i\alpha)|dy \to 0, \quad \xi \to \pm\infty.$$

Similarly, by Assumption 2.1, $|\xi|\int_{\alpha-d_+}^{\alpha-d_-} |\phi_\Delta(-\xi+iy)|dy \to 0$ as $\xi \to \pm\infty$. It then follows that

$$\int_{d_-}^{d_+} |\hat{f}_\alpha^j(\xi+iy)(\xi+iy)\phi_\Delta(-\xi-iy+i\alpha)|dy \to 0, \quad \xi \to \pm\infty.$$

Note that $\alpha-d_\pm \in (0,\lambda_+)$. According to Theorem 3.2, $|\hat{f}_\alpha^j(\xi+id_\pm)| \le ||f_{\alpha-d_\pm}^j||_{L^1(\mathbb{R})} < +\infty$. From Assumption 2.1, $\phi_\Delta(-\xi+i(\alpha-d_\pm))$ is in $L^1(\mathbb{R})$. Therefore,

$$\int_{\mathbb{R}} |\hat{f}_\alpha^j(\xi+id_\pm)\phi_\Delta(-\xi-id_\pm+i\alpha)|d\xi < +\infty.$$

Similarly, $\xi\phi_\Delta(-\xi+i(\alpha-d_\pm))$ is in $L^1(\mathbb{R})$, and thus

$$\int_{\mathbb{R}} |\hat{f}_\alpha^j(\xi+id_\pm)(\xi+id_\pm)\phi_\Delta(-\xi-id_\pm+i\alpha)|d\xi < +\infty.$$

Therefore, both $\hat{f}_\alpha^{j}(z)\phi_\Delta(-z+i\alpha)$ and $\hat{f}_\alpha^{j}(z)z\phi_\Delta(-z+i\alpha)$ are in $H(\mathcal{D}_{(d_-,d_+)})$. □

Theorem 4.3 and Propositions 4.1, 4.2 show that the discretization errors of the discrete approximations in our algorithm all converge exponentially in $1/h$. The convergence rate depends on $d = \min(-d_-, d_+)$. To speed up the convergence in $1/h$, we make $(\alpha - \lambda_+, \alpha)$ and the corresponding $[d_-, d_+]$ symmetric. This suggests taking $\alpha = \lambda_+/2$. Note that the discretization errors converge faster for the inverse Fourier transforms than for the Hilbert transforms. According to Proposition 4.2, the dominating discretization error for our algorithm is thus of the order $\exp(-\pi d/h)$. If the characteristic function further satisfies

$$|\phi_\Delta(\xi + i\alpha)| \leq \kappa \exp(-c|\xi|^\nu) \tag{4.2}$$

for some $\kappa, c, \nu > 0$, then the dominating term in the truncation errors is $\exp(-c(Mh)^\nu)$. This suggests that we might take h as a function of M in the following way so that the discretization and truncation errors decay at the same order:

$$h(M) = \left(\frac{\pi d}{c}\right)^{1/(1+\nu)} M^{-\nu/(1+\nu)}. \tag{4.3}$$

The sum of the discretization and truncation error is then dominated by $\exp(-c^{1/(1+\nu)}(\pi d M)^{\nu/(1+\nu)})$. This is convenient in practical applications since one does not need to determine h and M separately, and the convergence is exponential in M, which represents the computational cost.

When the characteristic function only has polynomial tails with $|\phi_\Delta(\xi + i\alpha)| \leq \kappa|\xi|^{-\nu-1}$ for some $\nu > 1$, the discretization error is still of the order $\exp(-\pi d/h)$, while the dominating truncation error is of the order $(Mh)^{-\nu+1}$, resulting from implementing the Newton-Raphson step. One can thus select h and M according to $(Mh)^{-\nu+1} = e^{-\pi d/h}$, or equivalently, $M = (e^{-\pi d/h} h^{\nu-1})^{1/(1-\nu)}$.

4.3 The fast Fourier transform

The main computational cost comes from Step 3, where we must evaluate $\hat{f}^j_\alpha(\xi)$ for $\xi = -Mh, \cdots, Mh$. If implemented directly, the computational cost in this step is $O(M^2)$. Fortunately, the fast Fourier transform can be used to significantly reduce the computational cost. Note that the approximation of the Hilbert transform in Step 3 is of the following form:

$$\sum_{m=-M}^{M} f_m \frac{1 - \cos(\pi(kh - mh)/h)}{\pi(kh - mh)/h} = \sum_{m=-M, m\neq k}^{M} f_m \frac{1 - (-1)^{k-m}}{\pi(k - m)}, \quad k = -M, \cdots, M.$$

We immediately notice that the above is simply the multiplication of a matrix by a vector. In particular, the matrix has constant diagonals. Such matrices are known as Toeplitz matrices. A Toeplitz matrix can be easily embedded into a circulant matrix. It is well known that the multiplication of a circulant matrix by a vector can be implemented using the fast Fourier transform (see [15], [17]). The computational cost for the implementation of the above summations is thus $O(M \log(M))$. Moreover, the same matrix is used from step to step. Therefore, only two runs of the fast Fourier transform are required for each step.

In Step 2, the number of operations needed depends on the number of iterations in the Newton-Raphson method. The Newton-Raphson method converges fast. Numerical results show that often 4 to 5 iterations already lead to highly accurate solution. The computational cost in this step is thus approximately $O(M)$. In Step 4, if the option price is needed only at one initial asset price, the computational cost is $O(M)$. If one desires to get the option value function, one might again use the Fast Fourier transform to compute $2M + 1$ option prices simultaneously in $O(M \log(M))$ operations.

Note that the discrete approximation of the inverse Fourier transform in the last step reduces to the following form:

$$\sum_{m=-M}^{M} e^{-ixmh} f_m.$$

Then the method of [2] can be used to generate the above values for $2M+1$ evenly spaced x's. The main idea is again to transform the problem to a Toeplitz matrix vector multiplication. The advantage of such a method is that there is no restriction on the step size in x. This is in contrast to a direct fast Fourier transform implementation of the above summation, which require that the step size in x and h satisfy a certain equation. More details for the Toeplitz matrix vector multiplication can be found in [21, 20]. In summary, the computational cost of our method is approximately $O(NM\log(M))$. It is linear in the number of monitoring intervals, and $O(M\log(M))$ in the number of terms used to approximate the inverse Fourier and Hilbert transforms.

5 Numerical Results

In this section, we illustrate the performance of our method for pricing Bermudan and American options. All computations are conducted on a Lenovo T400 laptop with 2.53 GHz CPU and 2 GB memory using Matlab version R2012a (we expect further speed improvement in C/C++).

5.1 Bermudan puts in the NIG model

We first consider pricing Bermudan puts in the normal inverse Gaussian (NIG) model ([3]). The characteristic function of the log return process in the NIG model is given by

$$\phi_t(\xi) = \exp\left(i\mu t\xi - \delta_{NIG}\, t(\sqrt{\alpha_{NIG}^2 - (\beta_{NIG} + i\xi)^2} - \sqrt{\alpha_{NIG}^2 - \beta_{NIG}^2})\right),$$

where μ is fixed by the martingale condition (2.5):

$$\mu = r - q + \delta_{NIG}\left(\sqrt{\alpha_{NIG}^2 - (\beta_{NIG} + 1)^2} - \sqrt{\alpha_{NIG}^2 - \beta_{NIG}^2}\right).$$

For a NIG process, $(\lambda_-, \lambda_+) = (\beta_{NIG} - \alpha_{NIG}, \beta_{NIG} + \alpha_{NIG})$. Moreover, the exponential tail condition (4.2) is satisfied with $c = \delta_{NIG}\Delta$ and $\nu = 1$ for any $\alpha \in (\lambda_-, \lambda_+)$. Model parameters are the same as those in [20]:

$$\alpha_{NIG} = 15, \beta_{NIG} = -5, \delta_{NIG} = 0.5.$$

We consider a Bermudan put with maturity $T = 1$ and strike price $K = 100$. The risk free interest rate is $r = 5\%$, the continuous yield is $q = 2\%$. The Bermudan option is monitored daily with $N = 252$ and $\Delta = 1/252$. We implement the algorithm described in Section 4.2. According to Propositions 4.1, 4.2 and Theorem 4.3, α should satisfy $\alpha \in (0, \lambda_+)$. That is, $0 < \alpha < 10$ in our example. As discussed at the end of Section 4.2, we select $\alpha = 5$ so that the theoretical exponential convergence rate in $1/h$ is maximized. The discretization step size h is then selected according to (4.3) with c and ν as mentioned in the above, and $d = 5$. We compute the benchmark price for the at-the-money put by taking a large enough M and get 6.489580997740. The top left plot in Figure 1 shows the option value as a function of the initial asset price. The top right plot shows the early exercise boundary. In particular, the early exercise boundary at time 0 is $S_0^* = 81.1802638151$. According to the analysis in

Section 4.2, the errors for approximating the inverse Fourier and Hilbert transforms are exponential in $\sqrt{M}$. The left plot in the second row exhibits the exponential convergence of the pricing error in $\sqrt{M}$.

Deep out-of-the-money options

It is reported in [12] that the Fourier transform method may lead to significant pricing errors, or even negative option prices, particularly for deep out-of-the-money options (negative prices are not surprising for deep out-of-the-money options, where option prices are close to zero while errors are larger). This likely has occurred when the chosen step size h and truncation level M are not able to bound both the discretization error and the truncation error sufficiently in numerical discretization of the transforms. Selecting appropriate h and M is not necessarily easy without explicit error estimates in terms of h and M. With explicit error estimates given in this paper, our transform method turns out to be rather robust for deep out-of-the-money options. The left plot in the second row in Figure 1 also shows the convergence of the pricing error for a deep out-of-the-money put option where the initial asset price is $S_0 = 200$. The benchmark option price is computed to be 0.0180624066. Our method remains very accurate.

Monitoring frequency

We have considered a daily monitored Bermudan option. Daily monitoring is computationally very challenging. Note that the numerical performance of our method depends on the tail behavior of the characteristic function ϕ_Δ. For many popular Lévy process models, including the NIG model, ϕ_Δ satisfies 4.2 for some $c, \nu > 0$. However, the coefficient c usually depends on Δ linearly, as can be seen in the above in the case of NIG. For smaller Δ, ϕ_Δ decays slower and hence larger M is needed to bound the truncation error. However, numerical results show that our method works well even for daily monitored Bermudan options. When there are less monitoring, our method could work much faster. This is illustrated in the bottom right plot in Figure 1, where a weekly monitored put ($N = 52$, put price 6.4833874148) and a monthly monitored put ($N = 12$, put price 6.4574297377) are considered. It turns out that when pricing American options using Richardson extrapolation, only Bermudan options with small monitoring frequencies need to be priced. Our method will thus be very efficient for pricing American options.

Dampening factor

As previously mentioned, $\alpha = 5$ was selected to maximize the theoretical convergence rate in $1/h$ in our example. This rule of thumb often works well. In the bottom left plot in Figure 1, we also plot the pricing error when $\alpha = 2.5$. It shows that the convergence is much slower. To achieve an accuracy of 10^{-4}, $M = 3000$ is sufficient when $\alpha = 5$, and $M = 6000$ is required when $\alpha = 2.5$.

Newton-Raphson

As for the Newton-Raphson method, we stop the iteration when the difference between two consecutive approximations is less than 10^{-8}. That is, $\epsilon_{NR} = 10^{-8}$. The early exercise boundary obtained in a previous step is used as the initial value for Newton-Raphson iterations in a current step. For the daily monitored at-the-money put option with $M = 7000$, it takes 4.08 iterations on average for each time step for the Newton-Raphson method. However, we would like to comment that $\epsilon_{NR} = 10^{-8}$ is

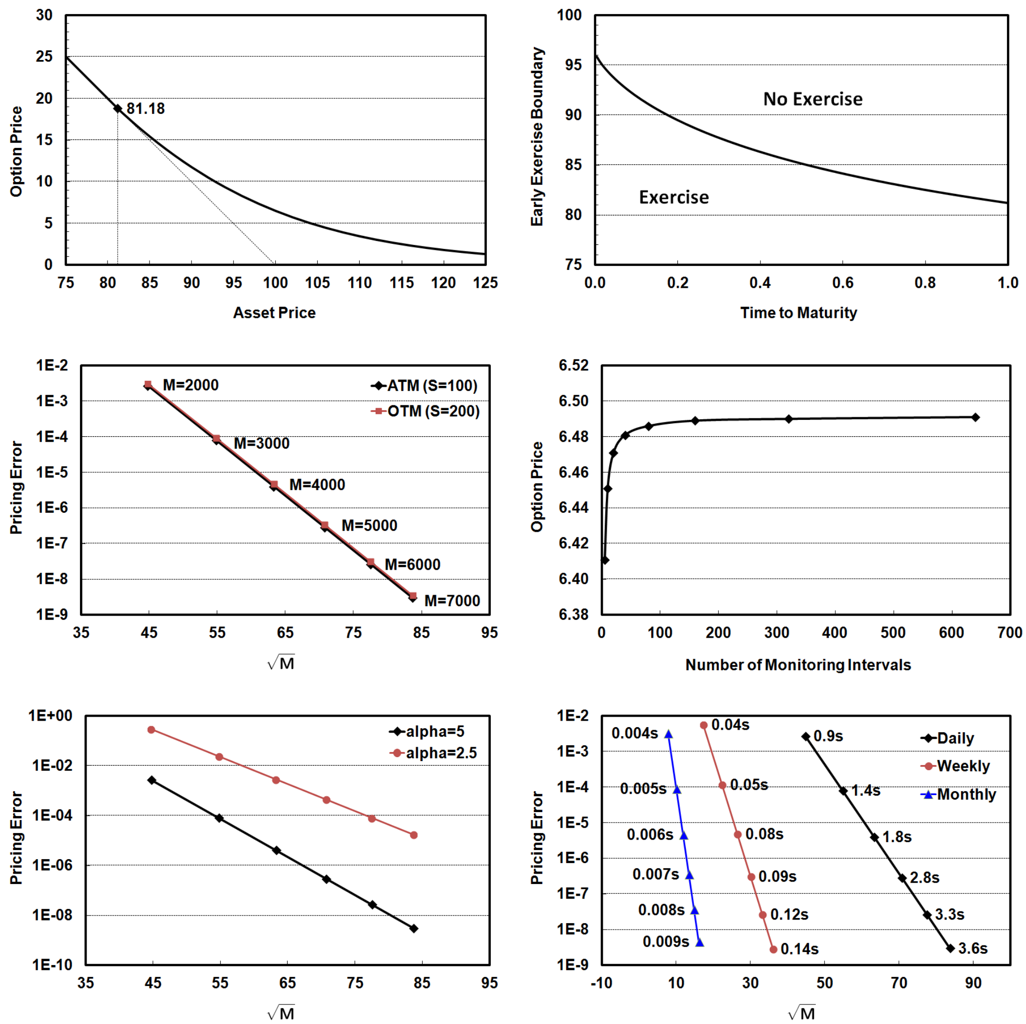

Figure 1: One year daily ($N = 252$) monitored Bermudan put in the NIG model (the bottom right plot also shows the cases with weekly ($N = 52$) and monthly ($N = 12$) monitoring.

rather conservative. For the daily monitored put option, with $M = 7000$, we produce option prices with $\epsilon_{NR} = 10^{-8}$ and with $\epsilon_{NR} = 10^{-4}$. The difference between the two prices is negligible. However, the average number of Newton-Raphson iterations drops from 4.08 to 2.22.

Finally, [26] considers pricing Bermudan options using a lattice method. As a double check, we reproduce their numerical results and are able to match very well the results in Table 10 in that paper for $N \leq 20$ in the NIG case. This further confirms the accuracy of our method. However, the lattice method becomes computationally rather intensive for large N.

5.2 American options

American option prices with continuous monitoring can be approximated by the prices of Bermudan options. In particular, Richardson extrapolation might be used to significantly improve the convergence. In this section, we price a sequence of Bermudan at-the-money put options with increasing number of monitoring intervals in the NIG, Kou's double exponential jump diffusion, the CGMY, and the Black-Scholes-Merton models. The parameters for these models are taken from [20]: for Kou's model, the characteristic function is given by

$$\phi_t(\xi) = \exp\Big(-\frac{1}{2}\sigma^2 t\xi^2 + i\mu t\xi - \lambda t \left(1 - \frac{p\eta_1}{\eta_1 - i\xi} - \frac{(1-p)\eta_2}{\eta_2 + i\xi} \right) \Big),$$

where $\sigma = 0.1, \lambda = 3, p = 0.3, \eta_1 = 40, \eta_2 = 12$, $(\lambda_-, \lambda_+) = (-\eta_1, \eta_2)$. For the CGMY model, the characteristic function is

$$\phi_t(\xi) = \exp\Big(i\mu t\xi + C\Gamma(-Y)t((\mathcal{M} - i\xi)^Y - \mathcal{M}^Y + (G + i\xi)^Y - G^Y) \Big),$$

where $C = 4, G = 50, \mathcal{M} = 60, Y = 0.7$, $(\lambda_-, \lambda_+) = (-\mathcal{M}, G)$. For the Black-Scholes-Merton model, the volatility is 20%. The strike price, maturity, risk free interest rate, dividend yield are the same as those in the previous section.

It is shown in [24] that, when the number of monitoring intervals increases, Bermudan option prices converge to the American option price at rate $1/N$ in the Black-Scholes-Merton model. Similar convergence has been observed in [19, 32] for more general Lévy models. Richardson extrapolation can therefore be used to speed up the convergence. More specifically, given two approximations P_1 (with N_1 monitoring intervals) and P_2 (with N_2 monitoring intervals) to the American put price (denoted by P_∞), we have the following extrapolated value as a new approximation:

$$P_\infty \approx \frac{N_1 P_1 - N_2 P_2}{N_1 - N_2}. \tag{5.1}$$

Table 1 reports the performance of Richardson extrapolation. The first column contains an increasing sequence of numbers of monitoring intervals. The following columns report the corresponding Bermudan option prices and the extrapolated values (RE) for the four models. As can be seen, American option prices can be obtained with one penny accuracy in all models using the prices of two Bermudan options, one with 5 monitoring intervals, and the other with 10 monitoring intervals. Therefore, for pricing American options in Lévy models, it may suffice to price a few Bermudan options with small numbers of monitoring intervals. Such contracts are computationally very easy to handle due to faster decaying of the characteristic function, as shown numerically in Section 5.1. An application of Richardson extrapolation then produces an accurate approximation to the American option price. The convergence of the Bermudan put prices as the number of monitoring intervals increases is also exhibited in the right plot in the second row in Figure 1 for the NIG model.

For a double check, we reproduce the American put option prices in Kou's model reported in [42]. We are able to match the prices reported in Table 6.2 of that paper. Similarly, we are able to match American put option prices in the CGMY model reported in Table 10 of [43] and Table 3 of [1]. In all of these papers, methods based on numerical solutions of partial integro-differential inequalities have been used. The valuation of American options in the Black-Scholes-Merton model has been intensively studied. See [8] and references cited therein (see also [9] for a recent survey). [23] proposes a two-phase approach for solving linear complementarity problems resulting from American options valuation in the Black-Scholes-Merton model. Again, we are able to reproduce prices reported in that paper.

6 Conclusions

This paper presents a Hilbert transform method for pricing Bermudan options in Lévy process models. The method involves a sequential evaluation of inverse Fourier and Hilbert transforms, which can be

N	NIG	RE	Kou	RE	CGMY	RE	BSM	RE
5	6.41114073		6.35854469		6.55002308		6.58462398	
10	6.45072757	6.49031	6.40861316	6.45868	6.58690113	6.62378	6.62146556	6.65831
20	6.47090173	6.49108	6.43434977	6.46009	6.60610853	6.62532	6.64073760	6.66001
40	6.48104498	6.49119	6.44745595	6.46056	6.61585279	6.62560	6.65061811	6.66050
80	6.48611979	6.49119	6.45408536	6.46071	6.62072205	6.62559	6.65562807	6.66064
160	6.48865570	6.49119	6.45742383	6.46076	6.62313530	6.62555	6.65815199	6.66068
320	6.48992286	6.49119	6.45910029	6.46078	6.62432849	6.62552	6.65941858	6.66069
640	6.49055616	6.49119	6.45994067	6.46078	6.62491923	6.62551	6.66005274	6.66069

Table 1: Pricing American puts using Richardson extrapolation.

approximated very efficiently with exponentially decaying errors. The computational cost of the method is approximately $O(NM\log(M)$, where N is the number of monitoring intervals, and M is the number of terms used to approximate the transforms. American options can be accurately priced using Richardson extrapolation. As for possible extensions, it will be interesting to apply at least some components of our method to the pricing of more general American style contracts (not only vanilla, but also barrier and lookback options) in Lévy models, as well as in alternative option pricing models (e.g., stochastic volatility and time-changed Lévy models), in particular, when the characteristic functions are known explicitly. We leave these for future research.

References

[1] A. Almendral and C. W. Oosterlee. Accurate evaluation of European and American options under the CGMY process. *SIAM Journal on Scientific Computing*, 29(1):93–117, 2007.

[2] D. H. Bailey and P. N. Swarztrauber. The fractional Fourier transform and applications. *SIAM Review*, 33(3):389–404, 1991.

[3] O. E. Barndorff-Nielsen. Process of normal inverse Gaussian type. *Finance and Stochastics*, 2(1):41–68, 1998.

[4] J. Bertoin. *Lévy Processes*. Cambridge University Press, Cambridge, UK, 1996.

[5] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[6] S. I. Boyarchenko and S. Z. Levendorskiǐ. *Non-Gaussian Merton-Black-Scholes Theory*. World Scientific, Singapore, 2002.

[7] S. I. Boyarchenko and S. Z. Levendorskiǐ. Pricing of perpetual Bermudan options. *Quantitative Finance*, 2(6):432–442, 2002.

[8] M. Broadie and J. Detemple. American option valuation: new bounds, approximations, and a comparison of existing methods. *Review of Financial Studies*, 9(4):1211–1250, 1996.

[9] M. Broadie and J. B. Detemple. Option pricing: valuation models and applications. *Management Science*, 50(9):1145–1177, 2004.

[10] M. Broadie and Y. Yamamoto. A double-exponential fast Gauss transform algorithm for pricing discrete path-dependent options. *Operations Research*, 53(5):764–779, 2005.

[11] P. Carr, H. Geman, D. B. Madan, and M. Yor. The fine structure of asset returns: an empirical investigation. *Journal of Business*, 75(2):305–332, 2002.

[12] P. Carr and D. Madan. Saddlepoint methods for option pricing. *Journal of Computational Finance*, 13(1):49–61, 2009.

[13] P. Carr and D. B. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999.

[14] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, Boca Raton, Florida, 2004.

[15] P. Davis. *Circulant Matrices*. Chelsea Publishing Company, New York, second edition, 1994.

[16] E. Eberlein, U. Keller, and K. Prause. New insights into smile, mispricing, and value at risk: the hyperbolic model. *Journal of Business*, 71(3):371–405, 1998.

[17] A. Eydeland. A fast algorithm for computing integrals in function spaces: financial applications. *Computational Economics*, 7(4):277–285, 1994.

[18] F. Fang and C. W. Oosterlee. A novel pricing method for European options based on Fourier-Cosine series expansions. *SIAM Journal on Scientific Computing*, 31(2):826–848, 2008.

[19] F. Fang and C. W. Oosterlee. Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions. *Numerische Mathematik*, 114:27–62, 2009.

[20] L. Feng and V. Linetsky. Pricing discretely monitored barrier options and defaultable bonds in Lévy process models: a fast Hilbert transform approach. *Mathematical Finance*, 18(3):337–384, 2008.

[21] L. Feng and V. Linetsky. Pricing options in jump-diffusion models: an extrapolation approach. *Operations Research*, 56(2):304–325, 2008.

[22] L. Feng and V. Linetsky. Computing exponential moments of the discrete maximum of a Lévy process and lookback options. *Finance and Stochastics*, 13(4):501–529, 2009.

[23] L. Feng, V. Linetsky, J. L. Morales, and J. Nocedal. On the solution of complementarity problems arising in American options pricing. *Optimization Methods and Software*, 26(4-5):813–825, 2010.

[24] S. Howison. A matched asymptotic expansions approach to continuity corrections for discretely sampled options. part 2: Bermudan options. *Applied Mathematical Finance*, 14(1):91–104, 2007.

[25] K. R. Jackson, S. Jaimungal, and V. Surkov. Fourier space time-stepping for option pricing with Lévy models. *Journal of Computational Finance*, 12(2):1–29, 2008.

[26] E. Kellezi and N. Webber. Valuing Bermudan options when asset returns are Lévy processes. *Quantitative Finance*, 4(1):87–100, 2004.

[27] F. W. King. *Hilbert Transforms*. Cambridge University Press, Cambridge, UK, 2009.

[28] S. G. Kou. A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101, 2002.

[29] R. W. Lee. Option pricing by transform methods: extensions, unification, and error control. *Journal of Computational Finance*, 7(3):51–86, 2004.

[30] A. L. Lewis. A simple option formula for general jump-diffusion and other exponential Lévy processes. Available at http://ssrn.com/abstract=282110, 2001.

[31] F. A. Longstaff and E. S. Schwartz. Valuing American options by simulation: a simple least-squares approach. *Review of Financial Studies*, 14(1):113–147, 2001.

[32] R. Lord, F. Fang, F. Bervoets, and C. Oosterlee. A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes. *SIAM Journal on Scientific Computing*, 30(4):1678–1705, 2008.

[33] D. Madan, P. Carr, and E. Chang. The variance gamma process and option pricing. *European Finance Review*, 2(1):79–105, 1998.

[34] R. C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[35] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.

[36] S. Raible. *Lévy Processes in Finance: Theory, Numerics and Empirical Facts*. PhD thesis, University of Freiburg, Germany, 2000.

[37] M. Reed and B. Simon. *Methods of Modern Mathematical Physics, II, Fourier Analysis, Self-Adjointness*. Academic Press, San Diego, California, 1975.

[38] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, UK, 1999.

[39] W. Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives*. John Wiley & Sons, Hoboken, New Jersey, 2003.

[40] M. Schweizer. On Bermudan options. In K. Sandmann and P. J. Schönbucher, editors, *Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann*, pages 257–269. Springer-Verlag, Berlin, Germany, 2002.

[41] F. Stenger. *Numerical Methods Based on Sinc and Analytic Functions*. Springer-Verlag, Berlin, Germany, 1993.

[42] J. Toivanen. Numerical valuation of European and American options under Kou's jump-diffusion model. *SIAM Journal on Scientific Computing*, 30(4):1949–1970, 2008.

[43] I. R. Wang, J. W. Wan, and P. A. Forsyth. Robust numerical valuation of European and American options under the CGMY process. *Journal of Computational Finance*, 10(4):31–70, 2007.