Research Article

Option Pricing by Willow Tree Method for Generalized Hyperbolic Lévy Processes

Hongying Wu,1 Zhiqiang Zhou,2 and Caijuan Kang2

1School of Mathematics and Information Science, Xiangnan University, Chenzhou 423000, China
2School of Economics and Management, Xiangnan University, Chenzhou 423000, China

Correspondence should be addressed to Zhiqiang Zhou; zq.zhou@xnu.edu.cn

Received 10 May 2023; Revised 11 September 2023; Accepted 28 September 2023; Published 17 October 2023

Copyright © 2023 Hongying Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a new approach is proposed to construct willow tree (WT) for generalized hyperbolic (GH) Lévy processes. There are two advantages of our proposed approach compared to the classical WT methods. Firstly, it avoids the moments matching from Johnson curve in the known WT construction. Secondly, the error of European option pricing is only determined by time partition $\Delta t$ under some conditions. Since the moments of Lévy measure are removed from this algorithm, our approach improves the stability and accuracy of WT in option pricing. Numerical experiments support our claims. Moreover, the new approach can be extended to other Lévy processes if their characteristic functions are expressed by explicit forms.

1. Introduction

To provide more modeling tools on jump types, the Lévy process is a generalization of the diffusion processes by allowing infinite jump activity. The standard form of the Lévy process assumes stationary increments, hence resulting in nice analytic tractability. The assumption of Lévy processes makes it a good choice for pricing equity derivatives. Generalized hyperbolic (GH, see [1]) process is a class of Lévy processes with wide application in financial field. Variance gamma (VG, see [2, 3]) and normal inverse gamma (NIG, see [4–7]) processes are two special cases of the GH model. VG and NIG models can be obtained from a randomized time-changed clock. The time-changed processes and general Lévy processes exhibit stochastic volatility such that they become capable of capturing volatility smile, smile skew, and term structure of the smile. Also, the hyperbolic (HYP) distribution and the Gauss normal (GN) distribution are viewed as subclasses of the GH model.

Currently, there are five popular numerical methods to price options under Lévy processes, binomial tree methods (BTMs, see [3, 8, 9]), finite difference methods (FDMs, see [10–13]), Monte Carlo methods (MCM, see [14]), FFT-based transformation methods (see [15–18]), and cosine-willow tree methods (see [19]). For BTMs and WTMs, computing transition probabilities (TPs) is an insurmountable barrier. Although a few literature studies discuss BTMs and WTMs for Lévy processes, the accuracy and computational efficiency of them are needed to be enhanced. To use FDMs to valuate options, corresponding partial integral-differential equations (PIDEs) should be established. Generally, PIDEs governed by Lévy processes are very complicated such that PIDEs are limited in several options. The Monte Carlo method is straightforward, but it is quite time-consuming to generate random samples for the Lévy process. The FFT-based transformation method is the most popular one in option pricing, such as the COS method (see [16]) and PROJ method (see [17]). The COS method employs a cosine series expansion on the risk-neutral return density and estimates the European option price based on the numerical integration on $(-\infty, +\infty)$. However, it is hard to determine a proper finite interval $[a, b]$ to truncate $(-\infty, +\infty)$ for the integration (see [15, 18, 20]) and is hard to be extended to path-dependent options. The PROJ method overcomes these shortcomings and is extendable to Asian options, variance swaps, and American options but its extendability is still limited compared to the Cosine-willow tree method.
The key of the WT method for option pricing is to construct willow tree structure (see [13]). In this paper, we propose WT algorithms for GH Lévy processes. The main contributions of this paper are fourfolds.

1. An numerical algorithm is proposed to compute the probability density functions (PDFs) and cumulative distribution functions (CDFs) from the characteristic functions (CDs) of GH processes. This algorithm is unified and suitable for all GH subclass models.

2. In FFT-based transformation methods (see [15, 18, 20, 21]), it is hard to find a proper finite interval \([a, b]\) to truncate the integration over \((-\infty, +\infty)\). While in the WT method, we propose an adaptive integration method in which the appropriate integral interval \([a, b]\) is automatically found. In determining \([a, b]\), the WT algorithm does not consume too much extra computational effort.

3. By setting appropriate \(m\) discrete stock prices at each time \(t_n\) and calculating transform probabilities \(p^{(m)}\) from \(t_n\) to \(t_{n+1}\), WT structure is constructed. Unlike Ma et al. [19], it is not needed to estimate the \(k\)-order moments of the GH Lévy model when selecting \(m\) nodes at each time \(t_n\). The determination of \(m\) nodes at each time \(t_n\) only relies on PDFs or CDFs of GH processes, which makes our WT programming run fast than those developed by Ma et al. (see [19]).

4. The convergence rate \(O(\Delta t)\) of European option on the WT structure is proved. Numerical experiments show that American options computed by our WT algorithm are also convergent for time partition \(\Delta t\) and the underlying partition number \(m\).

The remaining parts of this paper are arranged as follows. Basic conceptions are reviewed in Section 2. Calculation of PDFs and CDFs is illustrated in Section 3. The convergence analysis for European options is discussed in Section 4. Numerical examples of GH Lévy processes are carried out in Section 5. Some conclusions and remarks are given in the final section.

2. GH Lévy Processes

In option pricing with non-Gaussian processes, the asset price process \(S_t\) is defined as an exponential Lévy process \(X_t\), i.e.,

\[ S_t = S_0e^{(r+\omega_{GH})(t\Delta t)}X_t, \]

where \(r\) is the risk-free interest rate and \(\omega_{GH}\) is a martingal adjustment parameter under risk-neutral measure \(Q\). Under the GH Lévy processes (1), option pricing becomes more complicated than the classical BS model. For FDMs, partial integral-differential equations (PIDEs) are needed, and then numerical schemes should be designed to solve these PIDEs. For BTMs and WTM, calculating transform probabilities (TPs) cannot be avoided. Formulating PTMs, designing numerical schemes for PIDs, and calculating TPs for BTMs and WTM are not easy tasks.

We consider the generalized hyperbolic (GH) model (see [1, 22, 23]) whose characteristic function with five parameters \((\alpha, \beta, \mu, \lambda)\) is defned by the following equation:

\[ \phi^{GH}(u) = e^{iu\alpha - \frac{\alpha^2 - \beta^2}{2} u^2} \left( \frac{\delta^2 + (x - \mu)^2}{\delta^2 - \beta^2} \right)^{\lambda/2} K_1\left( \delta \sqrt{\alpha^2 - (\beta + iu)^2} \right) K_1\left( \delta \sqrt{\alpha^2 - (\beta - iu)^2} \right), \]

(2)

where \(K_1(\cdot)\) is the \(i^{th}\) order-modifed Bessel function of the second kind. The density function of the GH model \(\rho^{GH}(x)\) can be derived as follows:

\[ \rho^{GH}(x) = \frac{\left( \alpha^2 - \beta^2 \right)^{1/2} \delta^2 + (x - \mu)^2 \left( \delta^2 - \beta^2 \right)^{1/2} \left( \delta^2 + (x - \mu)^2 \right)^{1/2}}{\sqrt{2\pi (a\delta)^{1/2} \delta^{1/2} K_1(\delta \sqrt{\alpha^2 - \beta^2})}} e^{\beta(x-\mu)K_{\lambda-1/2}(\alpha \sqrt{\delta^2 + (x - \mu)^2})}. \]

(3)

In CFs (2), \(\mu \in \mathbb{R}\) and the other parameters satisfy the following constraints:

\[ \delta \geq 0, \alpha > 0, |\beta| < \alpha \text{ if } \lambda > 0, \]
\[ \delta > 0, \alpha > 0, |\beta| < \alpha \text{ if } \lambda = 0, \]
\[ \delta > 0, \alpha \geq 0, |\beta| \leq \alpha \text{ if } \lambda < 0. \]

The GH distribution embeds various distributions under special choices of the parameters. The parameter \(\lambda\) determines the subclass of the GH distribution. When \(\lambda = 1\), the GH distribution reduces to the hyperbolic distribution (HYP) whose logarithm of density is a hyperbolic. In addition, when \(\delta \rightarrow \infty\) and \(\delta/\alpha \rightarrow \sigma^2\), the GH distribution reduces to the Gauss normal (GN) distribution. Furthermore, when \(\delta = 0\) and \(\mu = 0\), the GH distribution becomes the variance gamma (VG) distribution; when \(\lambda = -1/2\), it becomes the normal inverse Gaussian (NIG) distribution. Table 1 gives the four different categories of the GH model.

Figure 1 plots the graphs of \(\rho^{GH}(x)\) with different values of parameters. From the Figure, we see the shape of tail, skewness, and kurtosis of the GH distribution which are controlled by \(\alpha, \beta, \mu, \sigma\) respectively.

Because the GH law is infinitely divisible, one can construct a GH Lévy process \(X_t\) whose distribution at fixed time \(t\) has characteristic function \(\phi_{X_t}\). The characteristic function is described by the following equation (see [22]):
\[ \phi_{X_t}(u) = E[e^{iuX_t}] = \int_{-\infty}^{\infty} e^{iux} \phi_{GH}(x) dx = \left[\phi_{GH}(u)\right]^t = e^{iu\mu + \int_{-\infty}^{\infty} \left(1 - e^{iux} + iux(\omega_{GH} \chi) + \log(\omega_{GH} \chi)\right) dx}. \tag{5} \]

where \( i = \sqrt{-1}, \) \( \phi_{GH}(x) \) is the density function of GH process at time \( t, \) and \( \phi_{GH}(u) \) is defined as (2). The function \( \phi_{GH}(u) = \phi_{GH}(u)e^{-i\mu u} \) is the part in which the oscillating factor \( e^{-i\mu u} \) is removed. The characteristic exponent of GH process is as follows:

\[ \psi_X(u) = \frac{1}{t} \log \phi_{X_t}(u) = \log \phi_{GH}(u). \tag{6} \]

Since a Lévy process is an infinitely divisible distribution, the Lévy-Khintchine theorem (see Theorem 8.1 in [24]) for an infinitely divisible distribution can be applied to establish the characteristic exponent \( \psi_X(u) \) of a Lévy process \( X_t. \) \( \psi_X(u) \) admits the Lévy-Khintchine representation, i.e.,

\[ \psi_X(u) = i\mu u - \frac{\sigma^2 u^2}{2} - \int_{\mathbb{R}\setminus\{0\}} \left(1 - e^{iux} + iux(\omega_{GH} \chi) + \log(\omega_{GH} \chi)\right) \Pi(dx). \tag{7} \]

The Lévy measure \( \Pi(dx) \) is defined on the real domain excluding zero with \( \Pi(\{0\}) = 0. \) The triplet \( (\mu, \sigma^2, \Pi) \) is called the Lévy characteristic of \( X_t, \) where \( \mu \in \mathbb{R} \) is the constant drift, \( \sigma > 0 \) is the constant volatility of the continuous component, and \( \Pi \) is the Lévy measure that represents the expected number of jumps per unit time. In GH model (1), the parameter \( \sigma \) is set as zero. The Lévy measure of the GH model has no explicit analytic form and it can be expressed only in terms of integrals (see [25]). To satisfy the martingale condition \( E_D[S_t|S_0] = S_0 \) in SDE (1), the adjustment parameter is chosen as \( \omega_{GH} = -\psi_X(-i). \)

The density function \( \phi_{GH}(x) \) can be restored from characteristic function \( \phi_X(u) \) by numerical inverse integral. Once the characteristic function (5) is given, the density function \( \phi_{GH}(x) \) is expressed as

\[ \phi_{GH}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \phi_{X_t}(u) du = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(x-u\mu)} \left[\phi_X(u)\right]^t du, \tag{8} \]

with definition \( \phi_X(u) \) in (5). We note that the integral factor \( e^{-i(x-u\mu)} \) is the part with high frequency oscillation, whereas \( \phi_X(u) \) is decaying quickly as \( |u| \to \infty. \) This observation is very important for computing the density function \( \phi_{GH}(x), \) numerically.

### 3. Willow Tree Algorithm

#### 3.1. Willow Tree Structure

A willow tree for the GH model is represented by \( \{X^n_t, S^n_i, p_{ij}^{[n]}\}, \) where \( X^n_t \) is discrete nodes of Lévy processes \( X_t \) at time \( t_n, \) \( S^n_i = S_0e^{\sigma^2\Delta t + \mu \Delta t + \omega_{GH} \chi \Delta t} \) is underlying prices, and \( p_{ij}^{[n]} \) is the transition probability. As shown in Figure 2, there are two main stages to construct a willow tree: (i) selecting the discrete tree nodes, \( X^n_i \) \((i = 1, 2, 3, \ldots, m \text{ and } n = 1, 2, \ldots, N), \) for \( X_t \) at each time \( t_n \) and (ii) determining the transition probability, \( p_{ij}^{[n]}, \) from \( X^n_i \) at \( t_n \) to \( X^{n+1}_j \) at \( t_{n+1} \) on a discrete time points \( 0 = t_0 < t_1 < t_2 < \cdots < t_N = T \) with \( t_n = n\Delta t \) and \( \Delta t = \frac{T}{N}, \) \( n = 1, 2, \ldots, N. \)

Firstly, the discrete pairs \( \{(X^n_i, q_i) \} \) \((i = 1, 2, \ldots, m; n = 1, 2, \ldots, N) \) are selected to approximate the distribution of \( X_t \) at time \( t_n. \) The cumulative distribution functions (CDFs) of \( X_t = X^n_i \) at \( t_n \) are computed by

\[ F^n(X^n_i) = \int_{-\infty}^{X^n_i} \phi_{GH}(x) dx = q_i, \quad i = 1, 2, \ldots, m; n = 1, 2, \ldots, N. \tag{9} \]

After setting discrete probabilities,

\[ q_i = \frac{(i - 0.5)}{m}, \quad i = 1, 2, \ldots, m, \tag{10} \]

we can numerically determine \( X^n_i \) by solving equation (9) if the integral can be computed efficiently and accurately. Very different from the classical WT method (see [19]), the determination of nodes \( X^n_i \) avoids computing \( k\)-order moments of \( X_t. \) This modification makes our algorithms more easily and widely available for Lévy processes.

Secondly, the transition probability between two consecutive discrete times \( t_n \) and \( t_{n+1} \) is computed as

\[ p_{ij}^{[n]} = P(\Delta_{(down)}^{[n]} \leq X^{n+1}_j - X^n_i \leq \Delta_{(up)}^{[n]}), \tag{11} \]

with increments of Lévy processes. Increments \( \Delta_{(down)}^{[n]} \) and \( \Delta_{(up)}^{[n]} \) are defined as

\[ \Delta_{(down)}^{[n]} = \frac{1}{2}\left(X^n_{j-1} + X^n_{j+1}ight) - X^n_i, \tag{12} \]

\[ \Delta_{(up)}^{[n]} = \frac{1}{2}\left(X^n_{j-1} + X^n_{j+1}ight) - X^n_i. \]

In computation (11), \( n = 0, 1, \ldots, N - 1 \) and \( i, j, 1, 2, \ldots, m, \) and an exception is that \( X^n_i = 0 \) with \( i = 1. \)
There are two most important aspects in willow tree construction. One is to calculate PDF $\rho_{GH}^{t_n}(x)$ and then generate willow tree nodes $X_n$ and another is to compute transition probability $p_{ij}^{\Delta t}$ via $\rho_{GH}^{\Delta t}(x)$ for given $\Delta t$. Once the PDFs $\rho_{GH}^{t_n}(x)$ and $\rho_{GH}^{\Delta t}(x)$ are computed, the willow tree algorithm can be applied in some types of option pricing, such as European, American, Asian options, and so on.

3.2. Numerical Computing of PDFs and CDFs for GH Processes. Firstly, we numerically compute the PDFs $\rho_{GH}^{t_n}(x)$ defined by (8). There is a challenge for the high frequency oscillation of integral kernel $e^{-iu(x-\mu t)}$ when the value of $(x-\mu t)$ is large. We consider the following truncation:

$$\rho_{GH}^{t_n}(x) \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iu(x-\mu t)} \left[ \Phi_X(u) \right]^t du = \frac{1}{2\pi} \int_{A}^{B} e^{-iu(x-\mu t)} \left[ \Phi_X(u) \right]^t du. \quad (13)$$

In the above expression, $A$ is a sufficiently large negative number and $B$ is a large enough positive constant. Let partition nodes $u_k = A + k\Delta u$ for $k = 0, 1, \cdots, M$ with $\Delta u = (B - A)/M$. Then, $\rho_{GH}^{t_n}(x)$ can be numerically integrated (NI), i.e.,

$$\text{NI}: \approx \frac{1}{2\pi} \sum_{k=0}^{M-1} \delta_k \int_{u_k}^{u_{k+1}} e^{-iu(x-\mu t)} du = \frac{1}{2\pi(\mu t - x)} \sum_{k=0}^{M-1} \delta_k C_k \quad (14)$$

Figure 1: Density function $\rho_{GH}^{t_n}(x)$ with parameters $\lambda = -1/2, \alpha = 4.5, \beta = 0.15, \mu = -0.1$, and $\delta = 0.2$: (a) density functions with different $\alpha$, (b) density functions with different $\beta$, and (c) density functions with different $\delta$.

Figure 2: Graphical depiction of the willow tree lattice with 5 possible asset prices and 4 discrete times.
with notation
\[
\begin{align*}
\delta_k &= \frac{1}{2} \left( \left[ \tilde{\phi}_X(u_k) \right]' + \left[ \tilde{\phi}_X(u_{k+1}) \right]' \right), \\
C_k &= i(\mu t - x) \int_{u_k}^{u_{k+1}} e^{-i[(\mu t - x)u + \omega(\omega + 1)u^2]} du = e^{-i(\mu t - x)u_k + \omega(\omega + 1)u_k^2}. 
\end{align*}
\]

(15)

Remark 1. The numerical formula (14) plays a key role for both \( t = t_n \) and \( t = \Delta t \). (i) Let \( t = t_n \), we can calculate CDFs \( F^n \) \((X^n)\) at each time \( t_n \). (ii) Nodes \( \{X^n_i\} \) can be determined by solving \( F^n(x) = q_i \) for \( i = 1, 2, \ldots, m \) (see expression (9)). (iii) Transition probabilities \( p^{(n)}_{ij} \) (see formula (11)) can be calculated by numerical approximation with \( t = \Delta t \). (iv) Using formula (14), we can expand the range of \([A, B]\) until the result \( \tilde{\varphi}^{GH} \) does not change much, which can be seen as an adaptive algorithm.

Remark 2. Since the absolute values of characteristic function \( \tilde{\varphi}_X(u) \) are decreasing to zero as \( u \to \pm \infty \), the nodes \( \{u_k\} \) can be selected as nonuniform, for example,
\[
\begin{align*}
u_k &= B \left( \frac{k}{M} \right)^x \quad \text{and} \quad u_k = A \left( \frac{k}{M} \right)^x, 
\end{align*}
\]
with \( x > 0 \) for \( k = 1, 2, \ldots, M \). Using nonuniform nodes \( \{u_k\} \), we can achieve more accurate \( \tilde{\varphi}^{GH} \) with less nodes.

Next, we compute transition probabilities \( p^{(n)}_{ij} \) with definition (11). Taking discrete values \( -\infty < x_0 < x_1 < \cdots < x_i = X^n_i \) with sufficiently large negative number \( x_0 \) and sufficiently small \( \Delta x := \max_{1 \leq j \leq i} \{x_j - x_{j-1}\} \), CDFs \( F^n \) \((X^n)\) are approximated by
\[
F^n(x) = \frac{1}{2} \sum_{j=1}^{i} \left[ \tilde{\varphi}^{GH}(x_{j+1}) + \tilde{\varphi}^{GH}(x_j) \right] \left(x_j - x_{j-1}\right),
\]
for \( i = 1, 2, \ldots, m \) and \( n = 1, 2, \ldots, N \). The transition probability \( p^{(n)}_{ij} \) is approximated by
\[
p^{(n)}_{ij} = F^{(n+1)}(\Delta_{n+1}^{up}) - F^{(n+1)}(\Delta_{n+1}^{down}),
\]
with \( \Delta_{n+1}^{up} \) and \( \Delta_{n+1}^{down} \) being denoted by expression (12). Using (9) and (14)–(18), Algorithm 1 describes the details of constructing willow tree \{\(X^n_i, S^n_i, \varphi^{(n)}_{ij}\}\} for GH Lévy processes.

3.3. European and American Options. After the willow tree \( \{X^n_i, S^n_i, \varphi^{(n)}_{ij}\}\} is constructed, European and American options can be valued on WT backward. Other options (Asian, Lookback, and so on) also can be computed from WT and we omit the details.

Define by \( V^n_t \) the option values at time \( t_n \) with underlying \( S^n_t = S_0e^{(x-r)\Delta t}+X^n_t \). Option values \( V \) at \( t = 0 \) with parameters \( (S_0, K, T, r) \) can be computed from willow tree as follows:
\[
\begin{align*}
V^n_i &= f(S^n_i), \quad \text{for } i = 1, 2, \ldots, m, \\
V^n_j &= e^{-r\Delta t} \sum_{j=1}^{m} p^{(n)}_{ij} V^{n+1}_j, \quad \text{for } n = N - 1, N - 2, \ldots, 1, \\
V(S_0, K, T, r) &= e^{-rT} \sum_{j=1}^{m} p^{(0)}_{ij} V_j,
\end{align*}
\]
where payoff function \( f(\xi) = (\xi - K)^+ \) for call options, whereas \( f(\xi) = (K - \xi)^+ \) for put options.

American option value \( V \) at time zero with parameters \( (S_0, K, T, r) \) can be determined backward on willow tree.

\[
V(S_0, K, T, r) = \max \left\{ e^{-rT} \sum_{j=1}^{m} p^{(0)}_{ij} V_j, f(S_0) \right\}.
\]

American options via willow tree are described in Algorithm 2.

3.4. A Simple Method to Determine \( m \) Nodes at Each Time \( t_n \). To generate \( m \) nodes \( X^n_i \) at each time \( t_n \), it is much time-consuming according to Step 6 in Algorithm 1. We give a simple algorithm to generate \( m \) nodes only depending on
\[
\sum_{j=1}^{m} F^{(n)}(\Delta_{n+1}^{up}) = (i - 0.5) = ma_{ij}, \quad \text{with } \Delta_{n+1}^{ij} X = X^{i+1}_j - X^n_j,
\]
\[
(21)
\]
### 4. Convergence Analysis

Errors of option pricing from willow tree have two aspects: the errors $\epsilon (\rho) = \rho^{\text{GH}} (x) - \bar{\rho}^{\text{GH}} (x)$ (see (14)) coming from the calculation for PDFs (or CDFs) and the errors from backward computation on willow tree (see (19) and (20)). We give some detailed analysis in this section.

#### 4.1. Errors of PDFs and CDFs

We know that the absolute value of characteristic function $\phi_X (u)$ is decreasing to zero as $u$ tends to $\pm \infty$. So, the error $\epsilon (\rho)$ can be estimated as follows.

**Theorem 3.** Assume that the characteristic function satisfies $|\phi_X (u)| \leq \epsilon / 2$ for $u \notin (A, B)$, then

$$
\epsilon (\rho) = |\rho^{\text{GH}} (x) - \bar{\rho}^{\text{GH}} (x)| \leq \frac{1}{2\pi |\mu - x|} \left[ \epsilon + o \left( \frac{1}{M^2} \right) \right],
$$

where PDFs $\rho^{\text{GH}} (x)$ and estimated PDFs $\bar{\rho}^{\text{GH}} (x)$ are defined by (14) and $M$ is the number of nodes $\{u_k\}$.

**Proof.** It is obvious that
Step 1. Give the first node \( X_0^n = 0 \) and \( S_0^n = S_0 \) at time zero.

\%\% generate CDFs \( F \) on nodes \( x \)

Step 2. Based on (17), generate discrete distribution \( F_j = P_{n+1}^{x_j} \) for \( j = 0, 2, \ldots, L \). Here, \( x_0 \) is a sufficient small number, \( x_L \) is a sufficient larger number and \( \max_{j \in I} |x_j - x_{j-1}| \) sufficient small.

\%\% generate \( X_{n+1} \) for \( n = 0, 1, \ldots, N - 1 \).

Step 3. For \( n = 0: N - 1 \), generate \( X_{n+1}^i \) by interpolation,

\[ X_{n+1}^i = \text{interp}(F, x, m_{iq}) + X_n^i, \quad i = 1, 2, \ldots, m. \]

END

Step 4. Compute discrete underlying \( S_n^i = S_n e^{(r + \sigma^n) t_n} X_n^i \) and then the nodes \( \{X_n^i, S_n^i\} \) of willow tree are generated.

**Algorithm 3:** Generate nodes \( \{X_n^i, S_n^i\} \).

\[
\epsilon(\rho) = \rho^{GH_i}(x) - \bar{\rho}^{GH_i}(x) = \frac{1}{2\pi} \left[ \int_{-\infty}^{A} e^{-iu(x-\mu)} du + \int_{B}^{\infty} e^{-iu(x-\mu)} du + \int_{A}^{B} \left[ g(u; t) - \bar{g}(u; t) \right] du \right],
\]

where \( g(u; t) = [\phi_X(u)]^t \) and \( \bar{g}(u; t) = 1/2 [\phi_X(u_k) + \phi_X(u_{k+1})] \) for \( u \in (u_k, u_{k+1}) \). From error estimation of composite trapezoidal rule (see Richard and Dougals Faires [26]), we have

\[
|\epsilon(\rho)| = \left| \rho^{GH_i}(x) - \bar{\rho}^{GH_i}(x) \right| \leq \frac{1}{2\pi} \left[ \epsilon \int_{-\infty}^{A} e^{-iu(x-\mu)} du + \epsilon \int_{B}^{\infty} e^{-iu(x-\mu)} du + \int_{A}^{B} \left[ g(u; t) - \bar{g}(u; t) \right] du \right]
\]

\[
\leq \frac{1}{2\pi |\mu - \xi|} \left[ \epsilon \left( \frac{\epsilon}{2} + \frac{B - A}{12\pi^2} |\xi| \right) \right], \quad \xi \in (A, B),
\]

\[
= \frac{1}{2\pi |\mu - \xi|} \left( \epsilon + O \left( \frac{1}{M^2} \right) \right), \quad \forall x \neq \mu,
\]

which is the result of (23).

**Remark 4.** Given \( x \in \mathbb{R} \), to ensure the error \( \epsilon(\rho) \) be small enough, \(|A|, B, M\) should be taken as large as enough. For given \( \epsilon(\rho) \), a simple choice is to select \(|A|, B, M\) such that

\[
\min_{A, B} \left| \left\{ \left| \phi_X(A) \right|, \left| \phi_X(B) \right| \right\} \right| \leq \pi \epsilon |\mu - \xi|,
\]

and then find \( M \) such that

\[
\frac{B - A}{12M^2} \max_{|\epsilon| \leq |A, B|} |g_{\mu}''(\xi, k)| \leq \pi \epsilon |\mu - \xi|.
\]

As an example, Table 2 gives the choices of \( A, B, M \) for different values of \( x \) and fixed \( \epsilon(\rho) = 10^{-6} \).

### 4.2. Convergence of Willow Tree for European Option

Denoted by \( E^{(k)}_{n,i} \) the conditional \( k^{th} \) moments of increment \( \Delta X = X_{n+1} - X_n \) with given \( X_n = X_i^n \), i.e.,

\[
E^{(k)}_{n,i} = \mathbb{E} \left[ \left( X_{n+1} - X_n \right)^k \right] = \int_{-\infty}^{\infty} u^k \rho^{GH_i}(u) du, \quad k = 1, 2, \ldots.
\]

where \( \rho^{GH_i}(u) \) is the PDFs of the increment of \( X_t \) at \( t_{n+1} \) given \( X^n = X_i^n \). For convenience, we give an assumption as follows.

**Assumption 5.** There exists a positive number \( H \) such that

\[
h := \max_{n,i,j} \left| \Delta_{n,i,j}^{(up)} - \Delta_{n,i,j}^{(down)} \right| \leq \frac{H}{m},
\]

for \( n = 1, 2, \ldots, N \) and \( i, j = 1, 2, \ldots, m \). Here, \( \Delta_{n,i,j}^{(up)} \) and \( \Delta_{n,i,j}^{(down)} \) are denoted by expression (12).

**Lemma 6.** Assume \( X_t \) as a GH Lévy process, then \( E[X_t] \) and \( \text{Var} [X_t] \) have the following results:
Parameters of the GH model are set as $\lambda = -2, \alpha = 15, \beta = 8, \delta = 0.3,$ and $\mu = 0.7$.

Given the characteristic exponent $\psi_X(u)$ of $X_t$, and under Assumption 5, i.e., $h$ is bounded by $H/m$ with constant $H$, and $\psi_X(u)$ is analytic for $|u| < \infty$, the error between $k^{th}$ moments, and their discrete approximations can be written as

$$E_{\Delta t}^{(k)} = O(\Delta t^k)$$

where $\psi_X(u)$ is analytic for $|u| < \infty$.

The following lemma gives an estimation of $R_{\Delta t,m}$, the error between $k^{th}$ moments, and their discrete approximations.

**Lemma 8.** Given $X_t$ following GH Lévy processes, under Assumption 5, i.e., $h$ is bounded by $H/m$ with constant $H$, and $\psi_X(u)$ is analytic for $|u| < \infty$, the errors between $E_{\Delta t}^{(k)}$ and its discrete approximations $\sum_{j=1}^{m} p_j^n (\Delta X^n_{ij})^k$ with $\Delta = X^n_{i+1} - X^n_i$ can be estimated by

$$R_{\Delta t,m}^{(k)} = \left| \sum_{j=1}^{m} p_j^n (\Delta X^n_{ij})^k - \int_{-\infty}^{\infty} u^k \rho_{GH} (u) du \right|$$

with $\zeta = \delta \sqrt{a^2 - \beta^2}$.

**Proof.** The proof can be seen in Cont and Tankov [27].

**Lemma 7.** Given $X_t$ following GH Lévy processes, the $k^{th}$ conditional moment of increment $\Delta X = X^n_{i+1} - X^n_i$ has estimation $E_{\Delta t}^{(k)} = O(\Delta t^k)$ for $k \geq 1$.

**Proof.** Given the characteristic exponent $\psi_X(u)$ of $X_t$ and characteristic function $\phi_{X_{n}}(u)$, it implies that

$$R_{\Delta t,m}^{(k)} = \left| \sum_{j=1}^{m} p_j^n (\Delta X^n_{ij})^k - E_{\Delta t}^{(k)} \right| = O\left( \frac{\sqrt{\Delta t}}{m} \right), k \geq 1.$$
backward induction (38) can be written as
\[ |R_{n,m}^{(k)}| \leq \sum_{\varepsilon=1}^{k-1} \frac{k!}{\varepsilon!(k-\varepsilon)!} \sum_{j=1}^{m} \mathcal{J} \left[ u^{\varepsilon} (\Delta X_{i,j} - u)^{k-\varepsilon} \rho^{GHw} (u) \right] \]
\[ \leq \sum_{\varepsilon=1}^{k-1} \frac{k!}{\varepsilon!(k-\varepsilon)!} \left\{ \sum_{j=1}^{m} \mathcal{J} (u^{\varepsilon} \rho^{GHw} (u)) \right\}^{1/2} \left\{ \sum_{j=1}^{m} \mathcal{J} \left[ (\Delta X_{i,j} - u)^{2(k-\varepsilon)} \rho^{GHw} (u) \right] \right\}^{1/2} \]
\[ \leq \sum_{\varepsilon=1}^{k-1} \frac{k!}{\varepsilon!(k-\varepsilon)!} \left[ E(2\varepsilon) \right]^{1/2} \left\{ \sum_{j=1}^{m} h_j^{2(k-\varepsilon)} \mathcal{J} \left[ \rho^{GHw} (u) \right] \right\}^{1/2} \]
\[ \leq \sum_{\varepsilon=1}^{k-1} \frac{k!}{\varepsilon!(k-\varepsilon)!} \left[ E(2\varepsilon) \right]^{1/2} \left( \frac{H}{m} \right)^{k-\varepsilon} = O \left( \frac{\sqrt{\Delta t}}{m}, \ k \geq 1, \right) \]

which is the result of (33).

The following theorem gives the error estimation of the willow tree algorithm for European options.

**Theorem 9.** Given the asset price \( S \), governed by the exponential Lévy model \( S_t = S_0 e^{(r-\omega) t+X_t} \), on discrete times \( 0 = t_0 < t_1 < \cdots < t_N = T \) with \( \Delta t = T/N \) and \( m \) discrete values \( X_n \) being generated by (9). If \( h \) satisfies conditions in Assumption 5, the error between the true value \( V(x,t) \) of the European option and the computed value \( \mathcal{V}(x,t) \) by the backward induction (19) in the willow tree is \( O(\Delta t) = O(1/N) \) when \( m \) is in \( O(\Delta t^{-3/2}) = O(N^{3/2}) \).

**Proof.** It is known that the European option \( V(y,t) \) with \( y = \log S(t) \) is the solution of the following partial integro-differential equation (PIDE):

\[ V_t(y,t) + (r + \omega) V_y(y,t) - r V(y,t) + \int_{-\infty}^{\infty} [V(y+\theta,t) - V(y,t)] \Pi (d\theta) = 0, \]

with the terminal condition \( V(y,T) = f(y) \), \( f(y) = (e^y - K)^+ \) for call option or \( f(y) = (K - e^y)^+ \) for put option, and \( \Pi (d\theta) \) being the Lévy measure (see Lévy representation Theorem 2.7 in [7]). Given willow tree, the European option \( \mathcal{V}(y_n,t_n) \) with \( y_n = \log S_n = (r + \omega) t_n + X_n \) is computed by the backward induction as in (19), i.e.,

\[ \mathcal{V}(y_n, t_n) = e^{-r \Delta t} \sum_{j=1}^{m} p_{ij} \mathcal{V}(y_{n+1}, t_{n+1}). \]

Expanding \( \mathcal{V}(y_{n+1}, t_{n+1}) \) at \( (y_n, t_n) \) by the Taylor series, we have

\[ \mathcal{V}(y_n, t_n) = \mathcal{V}(y_n, t_n) + \frac{1}{6} \frac{\partial^3 \mathcal{V}(\xi, t_n)}{\partial y^3} (\Delta Y_{ij})^3 + O((\Delta t)^3) + O(\Delta Y_{ij}^2 \Delta t), \]

where \( \Delta Y_{ij} = y_{n+1} - y_n \) and \( \xi \in (y_n, y_{n+1} + \Delta Y_{ij}) \). Thus, the backward induction (38) can be written as
\( \nabla(y^n_i, t_n) = e^{-r\Delta t} \sum_{j=1}^{m} p_{ij}^{(n)} \nabla(y^{n+1}_j, t_{n+1}) \)

\[
= (1 - r\Delta t) \sum_{j=1}^{m} p_{ij}^{(n)} \left[ \nabla(y^n_i, t_n) + \nabla_t(y^n_i, t_n) \Delta t + \nabla_y(y^n_i, t_n) \Delta y^n_{ij} \right] \\
+ (1 - r\Delta t) \sum_{j=1}^{m} p_{ij}^{(n)} \left[ \frac{1}{2} \nabla_{yy}(y^n_i, t_n)(\Delta y^n_{ij})^2 + \frac{1}{6} \frac{\partial^3 \nabla(\xi, t_n)}{\partial y^3}(\Delta y^n_{ij})^3 \right] \\
+ O(\Delta t^2) + O \left( \sum_{j=1}^{m} p_{ij}^{(n)} \Delta Y^n_{ij} \Delta t \right).
\]

From Lemma 8, the discrete approximation of the first- and second-order moments of the GH process can be estimated as

\[
\sum_{j=1}^{m} p_{ij}^{(n)} \Delta Y^n_{ij} = E^{(1)}_{\Delta t} + O \left( \frac{\sqrt{\Delta t}}{m} \right) \\
= (r + \omega_{GH}) \Delta t + \Delta t E[X_1] \\
+ O \left( \frac{\sqrt{\Delta t}}{m} \right). \tag{41}
\]

\[
\sum_{j=1}^{m} p_{ij}^{(n)} (\Delta Y^n_{ij})^2 = E^{(2)}_{\Delta t} + O \left( \frac{\sqrt{\Delta t}}{m} \right) \\
= \Delta t \text{Var}[X_1] + O \left( \frac{\sqrt{\Delta t}}{m} \right). \tag{42}
\]

\[
\sum_{j=1}^{m} p_{ij}^{(n)} (\Delta Y^n_{ij})^k = E^{(k)}_{\Delta t} + O \left( \frac{\sqrt{\Delta t}}{m} \right), \quad k \geq 3. \tag{43}
\]

From (40)–(43), we have

\[
\nabla_t(y^n_i, t_n) + \left[ (r + \omega_{GH}) + \frac{1}{\Delta t} E^{(1)}_{\Delta t} \right] \nabla_y(y^n_i, t_n) - r \nabla_y(y^n_i, t_n) + \frac{1}{2 \Delta t} E^{(2)}_{\Delta t} \nabla_{yy}(y^n_i, t_n) \\
+ \frac{1}{6 \Delta t} E^{(5)}_{\Delta t} \frac{\partial^3 \nabla(\xi, t_n)}{\partial y^3} + O(\Delta t) + O \left( \frac{\sqrt{\Delta t}}{m} \right) = 0 \text{ with } \xi \in (y^n_i, y^n_i + \Delta Y^n_{ij}). \tag{44}
\]

On the other hand, using the Taylor expansion of \( \nabla(\xi^n_i + \theta, t_n) \) at \( (y^n_i, t_n) \), we have

\[
\nabla(\xi^n_i + \theta, t_n) - \nabla(\xi^n_i, t_n) = \nabla_{\theta} (y^n_i, t_n) \theta + \frac{1}{2} \nabla_{\theta \theta} (y^n_i, t_n) \theta^2 + \frac{1}{6} \frac{\partial^3 \nabla(\xi, t_n)}{\partial y^3} \theta^3,
\]

with \( \xi \in (y^n_i, y^n_i + \Delta Y^n_{ij}) \). Applying the properties of Lévy measure, it is obtained that

\[
\int_{-\infty}^{\infty} \theta^k \Pi(d\theta) = \frac{1}{\Delta t} E^{(k)}_{\Delta t} + O \left( (\Delta t)^k \right), \quad k \geq 1. \tag{45}
\]
Therefore, using (45) and (46), it is yielded that

\[
\int_{-\infty}^{\infty} \left[ \nabla (y^n_j + \theta, t_n) - \nabla (y^n_j, t_n) \right] \Pi (d\theta)
= \frac{1}{\Delta t} E^{(1)}_{\Delta t} \nabla y_j (y^n_j, t_n) + \frac{1}{2 \Delta t} E^{(2)}_{\Delta t} \nabla y_j (y^n_j, t_n) + \frac{1}{6 \Delta t} E^{(3)}_{\Delta t} \frac{\partial^3 \nabla (\xi, t_n)}{\partial y^3} + O(\Delta t).
\]
Combining (44) and (47), we have

\[ \nabla_t(y^n_t, t_n) + (r + \omega_{GH})\nabla_x(y^n_t, t_n) - r\nabla(y^n_t, t_n) \\
+ \int_{-\infty}^{\infty} [\nabla(y^n_t + \theta, t_n) - \nabla(y^n_t, t_n)] \Pi(d\theta) + O(\Delta t) + O\left(\frac{1}{m \sqrt{\Delta t}}\right) = 0, \]

(48)

at all nodes \((y^n_t, t_n)\). Comparing (48) with (37), we see any European option value \(\nabla(y^n_t, t_n)\) computed by the backward induction (38) satisfies PIDE (37) with error term \(O(\Delta t) + O(1/m \sqrt{\Delta t})\). When \(m\) is in \(O(\Delta t^{-3/2})\), the error term \(O(\Delta t) + O(1/m \sqrt{\Delta t})\) is emerged as \(O(\Delta t) = O(1/N)\).

Remark 10. To obtain option values with good accuracy, the number of \(m\) should be taken as \(O(N^{1/2})\). Since European options are path-independent, the number of \(N\) should be selected as small as possible, whereas the number \(m\) should be taken as large enough since \(N\) should be taken as large as possible for American options.
Table 4: Option values calculated by WT method, MC method, and analytical formulas.

<table>
<thead>
<tr>
<th>K</th>
<th>ECWT</th>
<th>ECMC</th>
<th>ECA</th>
<th>ERR</th>
<th>K</th>
<th>APWT</th>
<th>APMC</th>
<th>ERR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GH model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>2.6349</td>
<td>2.6352</td>
<td>—</td>
<td>2.50e-04</td>
<td>10.00</td>
<td>0.4280</td>
<td>0.4365</td>
<td>8.53e-03</td>
</tr>
<tr>
<td>2.50</td>
<td>2.1497</td>
<td>2.1499</td>
<td>—</td>
<td>2.50e-04</td>
<td>10.50</td>
<td>0.7160</td>
<td>0.7260</td>
<td>9.96e-03</td>
</tr>
<tr>
<td>3.00</td>
<td>1.6645</td>
<td>1.6647</td>
<td>—</td>
<td>2.50e-04</td>
<td>11.00</td>
<td>1.0796</td>
<td>1.0888</td>
<td>9.19e-03</td>
</tr>
<tr>
<td>3.50</td>
<td>1.1793</td>
<td>1.1795</td>
<td>—</td>
<td>2.50e-04</td>
<td>11.50</td>
<td>1.5101</td>
<td>1.5159</td>
<td>5.76e-03</td>
</tr>
<tr>
<td>CPU(s)</td>
<td>3.14</td>
<td>32.53</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>3.14</td>
<td>311.05</td>
<td>—</td>
</tr>
<tr>
<td><strong>HYP model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>1.9754</td>
<td>1.9754</td>
<td>—</td>
<td>4.06e-06</td>
<td>10.00</td>
<td>0.7461</td>
<td>0.7486</td>
<td>2.47e-03</td>
</tr>
<tr>
<td>2.50</td>
<td>1.4901</td>
<td>1.4901</td>
<td>—</td>
<td>4.06e-06</td>
<td>10.50</td>
<td>1.0401</td>
<td>1.0424</td>
<td>2.31e-03</td>
</tr>
<tr>
<td>3.00</td>
<td>1.0049</td>
<td>1.0049</td>
<td>—</td>
<td>4.06e-06</td>
<td>11.00</td>
<td>1.3759</td>
<td>1.3788</td>
<td>2.89e-03</td>
</tr>
<tr>
<td>3.50</td>
<td>0.5198</td>
<td>0.5198</td>
<td>—</td>
<td>9.49e-06</td>
<td>11.50</td>
<td>1.7479</td>
<td>1.7511</td>
<td>3.16e-03</td>
</tr>
<tr>
<td>CPU(s)</td>
<td>2.78</td>
<td>23.02</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>2.90</td>
<td>290.50</td>
<td>—</td>
</tr>
<tr>
<td><strong>NIG model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>2.3485</td>
<td>2.3487</td>
<td>2.3487</td>
<td>3.27e-04</td>
<td>10.00</td>
<td>0.5662</td>
<td>0.5672</td>
<td>9.96e-04</td>
</tr>
<tr>
<td>2.50</td>
<td>1.8632</td>
<td>1.8636</td>
<td>1.8636</td>
<td>3.27e-04</td>
<td>10.50</td>
<td>0.8566</td>
<td>0.8584</td>
<td>1.87e-03</td>
</tr>
<tr>
<td>3.00</td>
<td>1.3780</td>
<td>1.3783</td>
<td>1.3784</td>
<td>3.27e-04</td>
<td>11.00</td>
<td>1.2025</td>
<td>1.2059</td>
<td>3.40e-03</td>
</tr>
<tr>
<td>3.50</td>
<td>0.8928</td>
<td>0.8931</td>
<td>0.8932</td>
<td>3.27e-04</td>
<td>11.50</td>
<td>1.5961</td>
<td>1.6000</td>
<td>3.91e-03</td>
</tr>
<tr>
<td>CPU(s)</td>
<td>0.82</td>
<td>18.48</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>9.77</td>
<td>169.23</td>
<td>—</td>
</tr>
<tr>
<td><strong>VG model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>8.0583</td>
<td>8.0587</td>
<td>8.0588</td>
<td>4.30e-04</td>
<td>10.00</td>
<td>0.1033</td>
<td>0.1016</td>
<td>1.73e-03</td>
</tr>
<tr>
<td>2.50</td>
<td>7.5731</td>
<td>7.5735</td>
<td>7.5735</td>
<td>4.30e-04</td>
<td>10.50</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>3.00</td>
<td>7.0879</td>
<td>7.0883</td>
<td>7.0884</td>
<td>4.30e-04</td>
<td>11.00</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>3.50</td>
<td>6.6026</td>
<td>6.6031</td>
<td>6.6032</td>
<td>4.30e-04</td>
<td>11.50</td>
<td>1.5000</td>
<td>1.5000</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>CPU(s)</td>
<td>0.82</td>
<td>18.48</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>9.59</td>
<td>169.23</td>
<td>—</td>
</tr>
</tbody>
</table>

European call options by WT, MC methods, and analytical formula are labeled by “ECWT,” “ECMC,” and “ECA,” respectively. American put options by WT and MC methods are labeled by “APWT” and “APMC,” respectively. ERRs are the corresponding errors between the WT method and Monte Carlo method.
Figure 6: European call and American put option values under. (a) GH model, (b) HYP model, (c) NIG model, and (d) VG model. All parameters of these models are listed in Table 3.

Figure 7: Convergence of willow tree with respect to \( m \). All parameters are listed in Table 3: (a) GH model, (b) HYP model, (c) NIG model, and (d) VG model.
5. Numerical Examples

To test the performance of the willow tree (WT) method for option valuating, we consider four classes of choices (GH, HYP, NIG, and VG models) with parameters being listed in Table 3. Parameters of risk-free interest, initial stock price, and maturity time are set as $r = 0.03, S_0 = 10, T = 1$. $(N = 1, m = 200)$ are set in the WT algorithm for European options, whereas $(N = 100, m = 200)$ for American options. In Monte Carlo (MC) simulation, $(N_{mc} = 100, M_{mc} = 5,000,000)$ are set with $N_{mc}$ representing the number of time partition and $M_{mc}$ represents the number of simulated paths. In numerical formula (14), we set numerical partition $M = 50,000$ for variable $u$ of characteristic functions.

All experiments are carried out by MATLAB R2012b running on a machine with Intel(R) Core(TM) i7-8550U CPU @ 1.80 GHz, 8 GB RAM under Windows 10.

Figure 3 plots probability density functions (PDFs) of GH, HYP, NIG, and VG models with $t = 1$. The figure shows that the PDFs computed from numerical formula (14) are very close to explicit expression (3). So, we believe that PDFs $\rho^{GH}(x)$ computed from (14) with time $t = \Delta t$ are also accurate enough. Figure 4 plots the shape of cumulative distribution functions (CDFs) with different numbers of $M$, from which we see the computed CDFs are convergent as $M$ becoming larger. Figure 5 plots the trajectories of nodes $X^n_t$ and $S^n_t$ for the NIG model.

Option values computed from WT algorithms, MC simulations, and analytical solutions (labeled by “ECA”) are listed in Table 4. Figure 6 plots values of European and American options, from which we see WT solutions are very close to those obtained by MC simulation. We see that the errors between WT solution and analytical solutions (or MC solutions) are about $10^{-3}$. In European options computation, the CPU time consumed from WT is less than 3 seconds whereas more than 18 s for MC simulation. In American option computation, the CPU time consumed from WT is less than 10 s, whereas more than 160 s for the MC method. The results in Table 4 illustrate the effectiveness of the proposed WT method. The analytical formula of European options under NIG and VG processes can be seen in literatures [2, 6, 28, 29]. Those pricing formulas are also listed in Appendix A and Appendix B.

To test the convergence of the willow tree method with respect to the number $m$ of space nodes and time partition number $N$, some experiments are carried out. Figure 7 plots the errors for different parameters $m$ with fixed $N = 20$. Figure 8: Convergence of willow tree with respect to $N$ with fixed $m = 50$. All parameters are listed in Table 3: (a) GH model, (b) HYP model, (c) NIG model, and (d) VG model.
Figure 8 plots the errors with different $N$ and corresponding $m = [N^{3/2}]$, from which we see the errors are decreasing as $N$ (and so $m$) increases. Table 5 lists the numerical convergent rates with respect to the values of $N$. These results in Figures 7 and 8 and Table 5 support the theoretical conclusion in Theorem 9.

### Table 5: Convergence rate for time partition $N$.  

<table>
<thead>
<tr>
<th>$N$</th>
<th>$m$</th>
<th>$\text{Err (Eu)}$</th>
<th>$\text{Conv. rate}$</th>
<th>$\text{Err (Am)}$</th>
<th>$\text{Conv. rate}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GH model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>0.1147</td>
<td>1.192</td>
<td>0.0302</td>
<td>1.019</td>
</tr>
<tr>
<td>10</td>
<td>158</td>
<td>0.0502</td>
<td>1.072</td>
<td>0.0149</td>
<td>1.033</td>
</tr>
<tr>
<td>15</td>
<td>290</td>
<td>0.0325</td>
<td>1.151</td>
<td>0.0098</td>
<td>0.976</td>
</tr>
<tr>
<td>20</td>
<td>447</td>
<td>0.0233</td>
<td>—</td>
<td>0.0074</td>
<td>—</td>
</tr>
<tr>
<td><strong>HYP model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>0.0823</td>
<td>1.099</td>
<td>0.0252</td>
<td>1.095</td>
</tr>
<tr>
<td>10</td>
<td>158</td>
<td>0.0409</td>
<td>1.089</td>
<td>0.0118</td>
<td>1.021</td>
</tr>
<tr>
<td>15</td>
<td>290</td>
<td>0.0263</td>
<td>1.058</td>
<td>0.0078</td>
<td>1.012</td>
</tr>
<tr>
<td>20</td>
<td>447</td>
<td>0.0194</td>
<td>—</td>
<td>0.0058</td>
<td>—</td>
</tr>
<tr>
<td><strong>NIG model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>0.0732</td>
<td>1.016</td>
<td>0.0145</td>
<td>0.998</td>
</tr>
<tr>
<td>10</td>
<td>158</td>
<td>0.0362</td>
<td>1.097</td>
<td>0.0073</td>
<td>1.062</td>
</tr>
<tr>
<td>15</td>
<td>290</td>
<td>0.0232</td>
<td>1.026</td>
<td>0.0047</td>
<td>1.069</td>
</tr>
<tr>
<td>20</td>
<td>447</td>
<td>0.0164</td>
<td>—</td>
<td>0.0035</td>
<td>—</td>
</tr>
<tr>
<td><strong>VG model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>0.1256</td>
<td>1.012</td>
<td>0.1423</td>
<td>0.985</td>
</tr>
<tr>
<td>10</td>
<td>158</td>
<td>0.0623</td>
<td>0.996</td>
<td>0.0719</td>
<td>1.054</td>
</tr>
<tr>
<td>15</td>
<td>290</td>
<td>0.0416</td>
<td>1.045</td>
<td>0.0469</td>
<td>1.037</td>
</tr>
<tr>
<td>20</td>
<td>447</td>
<td>0.0308</td>
<td>—</td>
<td>0.0348</td>
<td>—</td>
</tr>
</tbody>
</table>

The corresponding underlying discretization $m = [N^{3/2}]$. All parameters of these models are listed in Table 3.

6. Conclusions

In this paper, a unified and robust approach is proposed to construct the willow tree structure for GH Lévy processes. There are two advantages of our proposed approach compared to that in [19]. First, it avoids the moment matching failure by the Johnson curve under some circumstances in the willow tree construction. Second, the error of European call option pricing is only determined by $\Delta t$. The fifth-moment term of Lévy measure is removed from the error bound, so our approach improves the stability and accuracy of the willow tree in option pricing. Numerical experiments support our claims.

Moreover, we believe the proposed willow tree method can be extended to other option models, such as variance and volatility swaps with stochastic volatility and stochastic volatility model with regime switching stochastic mean reversion (see, e.g., [30–33]). We will discuss those models in the future.

### Appendix

#### A. Analytical Solution under VG Model

When the risk-neutral dynamics of the stock price is given by the VG process (for risk-neutral parameters $\sigma$, $\nu$, $\theta$), the European call option price on a stock is (see page 88 in [2]):

$$C(S; K, t) = SY \left( d \sqrt{1-C_2}, (\alpha + s) \sqrt{\frac{\nu}{1-C_1}}, \frac{t}{\nu} \right) - Ke^{-r t} \Psi \left( d \sqrt{1-C_2}, \alpha \sqrt{\nu}, \frac{t}{\nu} \right)$$

where $\alpha = \zeta s$,

$$\zeta = \frac{\theta}{\sigma^2},$$

$$s = \frac{\sigma}{\sqrt{1 + \nu (\theta/2)^2/2}},$$

$$C_1 = \frac{\nu (\alpha + s)^2}{2},$$

$$C_2 = \frac{\nu \alpha^2}{2},$$

and $\Psi (\cdot, \cdot, \cdot)$ is defined in terms of the modified Bessel function of the second kind.

#### B. Analytical Solution under NIG Model

When the risk-neutral dynamics of the stock price is given by NIG, the European call option price is (see Theorem 2.1, Theorem 2.2, and Corollary 2.1 in Ivanov [6]):

$$C = DAC - DCC,$$
The authors confirm that there are no conflicts of interest regarding the publication of this paper.

**Acknowledgments**

This work was supported by the National Natural Science Foundation of China (grant no. 12171409) and Key Project of Hunan Education Department (grant no. 238945).

**References**


