ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/263268910

A new sampling strategy willow tree method with application to path-dependent option pricing

Article *in* Quantitative Finance · June 2013
DOI: 10.1080/14697688.2012.762111

CITATIONS
8

READS
242

3 authors, including:

Wei Xu
Tongji University
43 PUBLICATIONS **237** CITATIONS
SEE PROFILE

Chenxiang Qin
Peking University
1 PUBLICATION **8** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Deriveative pricing and risk management View project

All content following this page was uploaded by Wei Xu on 25 August 2014.

The user has requested enhancement of the downloaded file.

This article was downloaded by: [Mr Zhiwu Hong]
On: 21 April 2013, At: 19:31
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Quantitative Finance
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rquf20

A new sampling strategy willow tree method with application to path-dependent option pricing

Wei Xu [a], Zhiwu Hong [b] & Chenxiang Qin [a]

[a] Department of Mathematics, Tongji University, Shanghai, 200092, China

[b] Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, Xiamen, 361005, China

Version of record first published: 20 Mar 2013.

To cite this article: Wei Xu, Zhiwu Hong & Chenxiang Qin (2013): A new sampling strategy willow tree method with application to path-dependent option pricing, Quantitative Finance, DOI:10.1080/14697688.2012.762111

To link to this article: http://dx.doi.org/10.1080/14697688.2012.762111

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Quantitative Finance, 2013
http://dx.doi.org/10.1080/14697688.2012.762111

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

A new sampling strategy willow tree method with application to path-dependent option pricing

WEI XU*†, ZHIWU HONG‡ and CHENXIANG QIN†

†Department of Mathematics, Tongji University, Shanghai 200092, China
‡Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, Xiamen 361005, China

(*Received 12 June 2012; in final form 18 December 2012*)

The willow tree algorithm, first developed by Curran in 1998, provides an efficient option pricing procedure. However, it leads to a large bias through Curran's sampling strategy when the number of points at each time step is not large. Thus, in this paper, a new sampling strategy is proposed. Compared with Curran's sampling strategy, the new strategy gives a much better estimation of the standard normal distribution with a small number of sampling points. We then apply the willow tree algorithm with the new sampling strategy to price path-dependent options such as American, Asian and American moving-average options. The numerical results illustrate that the willow tree algorithm is much more efficient than the least-squares Monte Carlo method and binomial tree method with higher precision.

Keywords: American options; Applied mathematical finance; Derivatives pricing; Option pricing; Numerical methods for option pricing

JEL Classification: G1, G12, G13

1. Introduction

Options are contracts whose future payoffs are determined by the price of another security, such as a common stock (Hull 2006). For example, for a European call option based on a stock with strike price K and exercise date T, the payoff function will be $V_T = \max\{0, S_T - K\}$ at time T, where S_T is the price of the stock at T. Since the underlying asset price at time t, S_t, changes stochastically over time, we assume that the asset price follows a geometric Brownian motion, that is

$$\frac{\mathrm{d}S_t}{S_t} = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t,$$

where $\mathrm{d}W_t$ is a Brownian motion increment, the parameter μ is the expected return rate and the parameter σ is the volatility of the asset price. When μ and σ are constant, the Black–Scholes model provides a closed-form solution for the European option price (Hull 2006).

If a closed-form solution does not exist, numerical methods have to be employed for option pricing. The Monte Carlo method is a commonly used numerical method for pricing. It is based on the analogy between probability and volume. It samples randomly from a universe of possible outcomes and takes the fraction of random draws that fall in a given set as an estimate of the set volume. The law of large numbers ensures that this estimate converges to the correct value as the number of draws increases. The Monte Carlo method is also suitable for multiple factor option pricing models, but it is too computationally expensive to price path-dependent options such as American, Asian and American moving-average options.

Another important numerical method is the binomial tree method first proposed by Cox *et al*. (1979). It is now one of the most popular approaches to pricing vanilla options due to its simplicity and flexibility. It can also be extended to the valuation of path-dependent options via a supplementary variable. The binomial tree method assumes that only one of two movements, up or down, occurs at each lattice node. It usually requires a large number n of time steps to produce an accurate value. Thus, it is quite expensive when n is large since the total number of nodes in a binomial tree is $O(n^2)$. Jiang and Dai (2005), using the notation of viscosity solutions, proved the uniform convergence of binomial tree methods for European and American options.

A new type of re-combining tree algorithm, called the willow tree method, was first proposed by Curran (1998). It approaches the Brownian motion directly through a discrete process with transition probabilities determined by

*Corresponding author. Email: wdxu@tongji.edu.cn

© 2013 Taylor & Francis

solving a sequence of linear programming problems. The beauty of the willow tree method is that these transition probabilities only need to be computed once and can be stored for further usage. Another distinct feature of this method is the constant number of nodes at each time step. As the number of time steps, n, increases, the number of tree nodes in a willow tree increases linearly, whereas it increases quadratically in the binomial tree. Thus, the willow tree is much more efficient than the binomial tree when the number of time steps is large. At each time step, a discrete process is used to approximate a standard normal distribution. When the number of nodes at each time step, m, is not large, the sampling strategy of Curran (1998) leads to a much lower variance and kurtosis than the standard normal distribution, which results in an underestimation of (deep) out-of-money option prices. Thus, in this paper, we propose a new sampling strategy to correct this problem and we apply the resulting willow tree method to pricing path-dependent options, such as arithmetic-average Asian options and American moving-average options.

The rest of the paper is structured as follows. The fundamentals of the willow tree method are described in section 2. We then give a new sampling strategy of discrete density pairs $\{(z_i, q_i) \mid i = 1, \ldots, m\}$ according to various sampling criteria in section 3. In section 4, we apply the willow tree to price path-dependent options such as American, Asian and American moving-average options. Numerical results are presented in section 5. Compared with the least-squares Monte Carlo method (Longstaff and Schwartz 2001), the willow tree method with the new sampling strategy is more efficient and powerful. Finally, we conclude with some remarks in section 6.

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

2. Fundamental of willow tree

Curran (1998) first proposed the new recombining tree algorithm, the willow tree algorithm. In the willow tree method, the number of spatial points at each time step is constant except for the initial step. Figure 1 depicts a willow tree lattice with 11 spatial points at each time step and four time steps in total. Between two contiguous time steps, a transition probability matrix is defined to describe the transition probability from one node at time t to another node at time $t + h$. The core idea for the willow tree is to approximate the Brownian motion by a discrete Markov process. The transition probability matrix can then be determined by a linear programming problem so that the discrete Markov process guarantees convergence to a Brownian motion. The following are the three main steps for the construction of a Markov process.

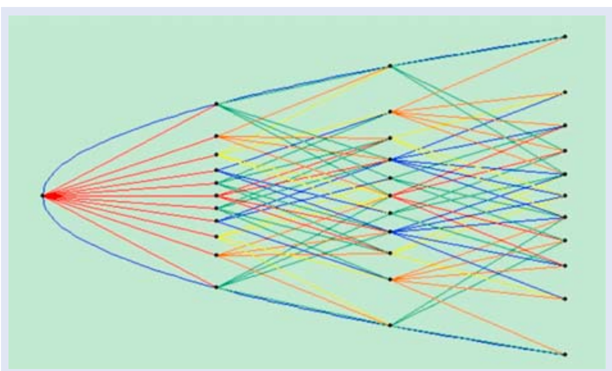

Figure 1. Graphical depiction of the willow tree lattice with 11 space nodes and four time steps.

(i) Generate a discrete approximation of the standard normal density function $\{(z_i, q_i) \mid i = 1, \ldots, m\}$, where $q_i = P(Z \le z_i)$. Curran (1998) suggested a selection of $\{z_i\}$ and $\{q_i\}$:

$$z_i = N^{-1}((i - 0.5)/m), \tag{1}$$

$$q_i = 1/m, \tag{2}$$

for $i = 1, \ldots, m$.

(ii) Let X_k be the embedded Markov chain on state space $\{1, 2, \ldots, m\}$ starting at 0 with a transition probability matrix $P^k = (p^k_{ij})$, where $p^k_{ij} = P(X_{k+1} = j \mid X_k = i)$, for $i, j = 1, \ldots, m$ and $k = 1, 2, \ldots, n - 1$. Since there is only one node at the initial time $t = 0$, the transition probabilities can be written as $p^0_{0i} = q_i$, for $i = 1, 2, \ldots, m$.

(iii) Let $z_1, z_2, \ldots, z_m$ be the discrete values for the standard normal distribution in the space direction and $0 = t_0 < t_1 < \cdots < t_n = T$ be a partition of $[0, T]$. Then, define $Y_{t_k} = \sqrt{t_k} z_{X_k}$ to be a discrete Markov process at time t_k.

To ensure that $\{Y_{t_k}, k = 0, 1, \ldots, n\}$ converges to a Brownian motion as $n \to \infty$ requires that $\{Y_{t_k}, k \ge 0\}$ is a martingale and the conditional variance of $Y_{t_{k+1}}$ given Y_{t_k} is h_k, where $h_k = t_{k+1} - t_k$, that is

$$E[Y_{t_{k+1}} \mid Y_{t_k}] = Y_{t_k}, \qquad \mathrm{Var}[Y_{t_{k+1}} | Y_{t_k}] = h_k, \tag{3}$$

for $k = 0, 1, \ldots, n - 1$. Since $\{X_k\}_{k=1,2,\ldots,n}$ is stationary, it also imposes that the conditional probability of ending at node i at time t_k is q_i, that is

$$\begin{aligned} P(Y_{t_k} = z_i\sqrt{t_k}) &= \sum_{j_1 j_2 \ldots j_{k-1}} q_{j_1} p^1_{j_1 j_2} p^2_{j_2 j_3} \ldots p^{k-1}_{j_{k-1} i} \\ &= \sum_{j_{k-1}} q_{j_{k-1}} p^{k-1}_{j_{k-1} i} \\ &= q_i. \end{aligned}$$

In order to ensure that the convergence is sufficiently fast, all transition probability matrices $\{P^k = (p^k_{ij}) \mid k = 1, 2, \ldots, n - 1\}$ should satisfy equation (3) with a uniform stationary distribution $\{q_i = 1/m, i = 1, 2, \ldots, m\}$. Therefore, this leads to the following linear programming problem (Ho 2000):

$$\min_{p_{ij}^k} \sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij}^k |\sqrt{t_{k+1}} z_j - \sqrt{t_k} z_i|^3, \tag{4}$$

s.t.

$$\sum_{j=1}^{m} p_{ij}^k = 1,$$

$$\sum_{j=1}^{m} p_{ij}^k \sqrt{t_{k+1}} z_j = \sqrt{t_k} z_i,$$

$$\sum_{j=1}^{m} p_{ij}^k t_{k+1} z_j^2 - t_k z_i^2 = h_k,$$

$$\sum_{i=1}^{m} q_i p_{ij}^k = q_j,$$

$$p_{ij}^k \geq 0, \quad \text{for } i, j = 1, 2, \ldots, m.$$

Although determining the transition probability matrix P from (4) is expensive, it is a one-time investment. The beauty of the willow tree method is that once the transition matrix P for some m and n is determined, this constructed willow tree can be used for option pricing with various S_0, K, μ and σ. Pricing options is then efficient and cheap. In practice, quite a few different sizes of willow trees are constructed and stored in the database in advance. When pricing an option, a suitable willow tree is extracted from the database.

From these constraints, it implies that the mean of the sequence $\{z_i, i = 1, 2, \ldots, m\}$ is zero and its variance is one, that is

$$\sum_{i=1}^{m} z_i q_i = 0, \qquad \sum_{i=1}^{m} z_i^2 q_i = 1. \tag{5}$$

The selection of $\{(z_i, q_i) \mid i = 1, 2, \ldots, m\}$ through equations (1) and (2) satisfies that the mean of $\{z_i\}$ is 0, but its variance fails to be 1. A simple compensation is to modify the two end points z_1 and z_m of $\{z_i, i = 1, 2, \ldots, m\}$ as $z_1 - \epsilon$ and $z_m + \epsilon$ so that equation (5) is tenable for an appropriate ϵ. Ho (2000) proved that the generated sequence $\{Y_{t_k}, k = 0, 1, \ldots, n\}$ with the above modified sampling strategy converges to a Brownian motion as m and n approach infinity. However, when the number of spatial nodes, m, is not large, the distribution of $\{z_i\}$ generated from the above sampling strategy usually has a much lower kurtosis than the standard normal distribution. Table 1 lists the two end points of $\{z_i\}$, z_1 and z_m, and the corresponding kurtosis with various m. It shows that Curran's modified sampling strategy does not approach a standard normal distribution very well for small m, which leads to less accuracy for (deep) out-of-money option pricing. On the other hand, since the number m determines the size of the transition probability matrix and the complexity of the linear programming problem, a large m means high computational and storage costs. Thus, it is important to find an effective $\{(z_i, q_i)\}$ sampling strategy with a small m to approximate the standard normal distribution well.

Table 1. Kurtosis, z_1, and z_m of Curran's sampling with $m = 30$, 50 and 100.

m	z_1	z_m	K
30	–2.2692	2.2692	2.8069
50	–2.4575	2.4575	2.8813
100	–2.6962	2.6962	2.9391

3. New sampling strategy

In Curran's sampling strategy, the sequence $\{z_i\}$ is restricted to have zero for its mean and one for its variance. Also, it assumes that the spatial nodes spread out in a uniform probability $q_i = 1/m$. In order to make $\{(z_i, q_i)\}$ approach the standard normal distribution, we first add two more restrictions on $\{(z_i, q_i)\}$. Thus, there are four restrictions for $\{(z_i, q_i)\}$ to determine the value of the pair sequence.

(i) The expected value of $\{z_i\}$ is zero, that is $\sum_{i=1}^{m} q_i z_i = 0$.

(ii) The variance of $\{z_i\}$ is one, that is $\sum_{i=1}^{m} q_i z_i^2 = 1$.

(iii) The kurtosis of $\{z_i\}$ is as close to three as possible, that is

$$\min \left(\sum_{i=1}^{m} q_i z_i^4 - 3 \right)^2.$$

(iv) The absolute values of the two end points z_1 and z_m are as large as possible.

The first three restrictions come directly from the properties of the standard normal distribution. The reason for the fourth restriction is as a standard normal distribution random variable z; the probability of z falling into $[-3, 3]$ is about 99.7%. The larger the absolute value of z_1 and z_m, the larger the interval of the standard normal distribution covered. Obviously, Curran's sampling strategy only satisfies the first two restrictions, not the third and fourth.

Another core idea of our sampling strategy is to spread out the node z_i non-uniformly with probabilities $\{q_i\}$. Since the inverse of the standard normal distribution function $N^{-1}(q)$ is symmetric around 0.5, we only need to generate half values of probability q_i and mirror the other half. First, let

$$q_i = (i - 0.5)^{\gamma}/m, \quad \text{for } i = 1, 2, \ldots, m/2, \tag{6}$$

where γ is a number between 0 and 1. Then, normalize $\{q_i\}$ by

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

$$q_i = q_i \Big/ \sum_{j=1}^{m} q_j, \tag{7}$$

where $q_i = q_{m+1-i}$. For $\gamma = 0$, our sampling strategy is the same as Curran's. For $\gamma > 0$, the probabilities for intervals close to the edge of the tree are decreased, while those for the center of the tree are increased. Thus, it leads to a further distribution away from the uniform distribution. Figure 2 shows the probability $\{q_i\}$ for $m = 30$ with various γ. The dotted line represents Curran's sampling strategy, that is $\gamma = 0$. The other two curves represents our sampling points with different γ. This illustrates that the larger γ, the further the curve is away from the uniform probability.

Given the number of space nodes m, and parameter γ, we can obtain m intervals $[Z_i, Z_{i+1}]$, where $Z_i = N^{-1}(\sum_{j=1}^{i} q_j)$, for $i = 1, 2, \ldots, m-1$, $Z_0 = -\infty$ and $Z_m = \infty$. Due to the symmetry of the standard normal distribution, we have $Z_i = -Z_{m-i}$ for $i = 0, 1, 2, \ldots, m/2$. Thus, a sequence $\{z_i\}$

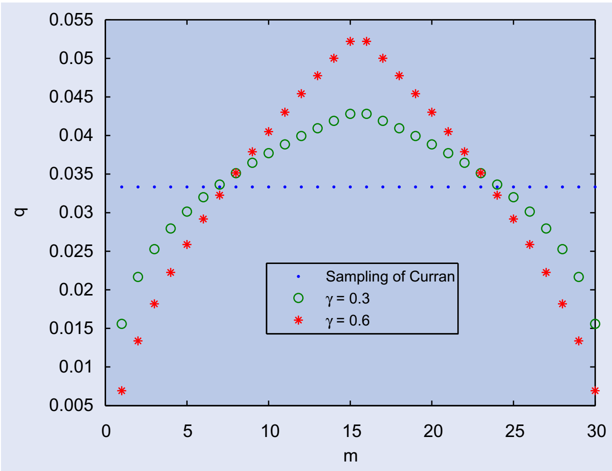

Figure 2. Comparison of $\{q_i\}_{i=1,2,\ldots,m}$ with different γ and $m = 30$.

satisfying the above four restrictions with a given probability sequence $\{q_i\}$ can be generated by solving the following constrained nonlinear least-squares problem:

$$\min_{z_i} \left[\sum_{i=1}^{m} q_i z_i^4 - 3 \right]^2, \tag{8}$$

s.t.

$$\sum_{i=1}^{m} q_i z_i = 0,$$

$$\sum_{i=1}^{m} q_i z_i^2 = 1,$$

$$Z_{i-1} \le z_i \le Z_i.$$

The first two constraints force the sequence $\{z_i\}$ to have zero mean and variance one. The last constraint for (8) guarantees z_i falling into $[Z_{i-1}, Z_i]$. The small value of q_i at the two ends implies that z_1 and z_{m-1} will be larger than or close to three. Table 2 lists $\{z_i, q_i\}$ from our optimization strategy (OP strategy) and the corresponding ends, Z_i with $m = 30$ and $\gamma = 0.3$, 0.6 and 1. Since z_i and q_i are symmetric around 0, only 15 points are recorded. The results of table 2 illustrate that all optimized z_i fall in the interval $[Z_{i-1}, Z_i]$ close to the middle point of the interval.

The minimization problem (8) is a constrained nonlinear least-squares problem that can be solved by many existing optimization algorithms (Dennis and Schnabel 1979, Xu *et al.* 2012). When m increases, the complexity of solving (8) increases significantly. However, the sampling points do not have to be computed every time for option pricing since these points are unchanged for a given m. Thus, several sets of sampling points for various m can be computed and stored in advance. When we price options or other derivatives, we just need to extract these points from the storage, rather than recomputing $\{z_i\}$.

Table 2. List of $\{z_i, q_i\}$ from our optimization strategy.

$\gamma = 0.3$			$\gamma = 0.6$			$\gamma = 1$		
	OP strategy			OP strategy			OP strategy	
Z_{i-1}	z_i	q_i	Z_{i-1}	z_i	q_i	Z_{i-1}	z_i	q_i
$-\infty$	−2.5908	0.0156	$-\infty$	−2.8821	0.0069	$-\infty$	−3.3500	0.0022
−2.1548	−1.9534	0.0217	−2.4613	−2.2081	0.0134	−2.8448	−2.5198	0.0067
−1.7834	−1.6549	0.0253	−2.0475	−1.8896	0.0182	−2.3702	−2.1700	0.0111
−1.5339	−1.4302	0.0279	−1.7685	−1.6486	0.0222	−2.0537	−1.9016	0.0156
−1.3379	−1.2463	0.0301	−1.5486	−1.4491	0.0259	−1.8048	−1.6777	0.0200
−1.1720	−1.0877	0.0320	−1.3619	−1.2749	0.0292	−1.5932	−1.4854	0.0244
−1.0253	−0.9461	0.0336	−1.1963	−1.1175	0.0323	−1.4051	−1.3182	0.0289
−0.8918	−0.8165	0.0351	−1.0449	−0.9717	0.0351	−1.2325	−1.1592	0.0333
−0.7675	−0.6955	0.0365	−0.9033	−0.8341	0.0379	−1.0704	−1.0003	0.0378
−0.6500	−0.5809	0.0377	−0.7685	−0.702	0.0405	−0.9154	−0.8466	0.0422
−0.5372	−0.4708	0.0389	−0.6384	−0.5737	0.0430	−0.7647	−0.6963	0.0467
−0.4277	−0.3637	0.0399	−0.5112	−0.4474	0.0454	−0.6162	−0.5475	0.0511
−0.3204	−0.2586	0.0409	−0.3852	−0.3216	0.0478	−0.4677	−0.3982	0.0556
−0.2139	−0.1541	0.0419	−0.2591	−0.1948	0.0500	−0.3172	−0.2463	0.0600
−0.1075	−0.0507	0.0428	−0.1312	−0.0655	0.0522	−0.1622	−0.0856	0.0644

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

The above proposed strategy is to fix $\{q_i\}$ and find a sequence of $\{z_i\}$ to satisfy the four restrictions. In contrast, we can let $z_1 = N^{-1}(q_1/2)$ and $z_i = N^{-1}(\sum_{j=1}^{i-1} q_j + q_i/2)$, for $i = 2, \ldots, m$ fixed and adjust $\{q_i\}$ so that the new adjusted sequences $\{\hat{q}_i\}$ and $\{z_i\}$ are also satisfied by the four restrictions. Apparently, the generated z_i falls into the interval $[Z_{i-1}, Z_i]$ and $z_i = -z_{m+1-i}$, for $i = 1, 2, \ldots, m$, due to the symmetry of the standard normal distribution. Thus, we propose another sampling strategy to generate $\{\hat{q}_i\}$ and $\{z_i\}$.

First, we generate a sequence of $\{q_i\}$, based on (6) and (7), and $\{z_i\}$. Due to the symmetry of the standard normal distribution, only half of the $\hat{q}_i$ values are considered, that is $\hat{q}_i$, for $i = 1, 2, \ldots, m/2$. Based on the four restrictions, $\{\hat{q}_i\}$ should satisfy

$$\sum_{j=1}^{m/2} \hat{q}_j = \frac{1}{2},$$
$$\sum_{j=1}^{m/2} \hat{q}_j z_j^2 = \frac{1}{2},$$
$$\sum_{j=1}^{m/2} \hat{q}_j z_j^4 = \frac{3}{2}.$$

Assuming $\hat{q}_1 = q_1 + a$, $\hat{q}_2 = q_2 + b$, $\hat{q}_{m/2} = q_{m/2} - a - b$ and $\hat{q}_i = q_i$, for $i = 3, \ldots, m/2 - 1$, we have

$$(z_1^2 - z_{m/2}^2)a + (z_2^2 - z_{m/2}^2)b = 0,$$
$$(z_1^4 - z_{m/2}^4)a + (z_2^4 - z_{m/2}^4)b = \frac{3}{2} - \sum_{j=1}^{m/2} q_j z_j^4.$$

Solving the above equation, we obtain

$$a = \frac{\frac{3}{2} - \sum_{j=1}^{m/2} q_j z_j^4}{(z_1^2 - z_2^2)(z_1^2 - z_{m/2}^2)},$$
$$b = \frac{\frac{3}{2} - \sum_{j=1}^{m/2} q_j z_j^4}{(z_2^2 - z_1^2)(z_2^2 - z_{m/2}^2)}.$$

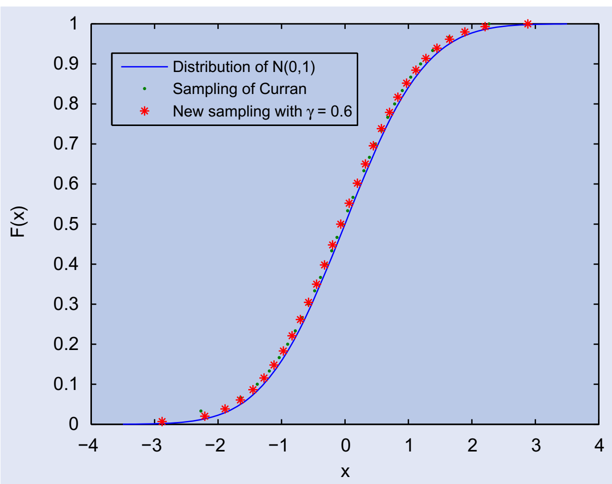

Figure 3. Distribution curve of $N(0, 1)$ and sampling points from Curran's strategy and our new strategy.

Therefore, the sets $\{\hat{q}_i\}$ and $\{z_i\}$ satisfy all four restrictions. Compared with the previously proposed sampling strategy, it is much cheaper to generate, but there is the possibility of $\hat{q}_1$, $\hat{q}_2$ or $\hat{q}_{m/2}$ being less than zero. We can adjust parameter γ to guarantee that $\hat{q}_1$, $\hat{q}_2$ and $\hat{q}_{m/2}$ are positive, or we can switch to the previously proposed sampling strategy when one of $\hat{q}_i$ is less than zero. Figure 3 depicts the curve of the standard normal distribution $N(0, 1)$ and the sampling points of Curran's sampling strategy and our second new sampling strategy with $\gamma = 0.6$ and $m = 30$. It indicates that our new sampling strategy has a wider spread for the standard normal distribution than Curran's sampling strategy.

Another sampling strategy is to match the first partial moment of $\{z_i, q_i\}$ with the standard normal distribution, rather than the high-order center moment, since matching higher-order moments is not so crucial for pricing near-the-money options. Given q_i generated from (6), $Z_i = N^{-1}(\sum_{j=1}^{i} q_i)$ for $i = 1, 2, \ldots, m - 1$, $Z_0 = -\infty$ and $Z_m = +\infty$, then z_i can be determined by the following constrained optimization problem:

$$\min_{z_i} \sum_{i=2}^{m} \left| \sum_{j=1}^{m} q_j (z_j - Z_{i-1})^+ - \int_{-\infty}^{+\infty} (z - Z_{i-1})^+ f(z) \mathrm{d}z \right|,$$

s.t.

$$\sum_{i=1}^{m} q_i z_i = 0,$$
$$\sum_{i=1}^{m} q_i z_i^2 = 1,$$
$$Z_{i-1} \le z_i \le Z_i,$$

where $(z_j - Z_i)^+ = \max(z_j - Z_i, 0)$ and $f(z)$ is the probability density function of the standard normal distribution. In fact, the objective function of the above optimization problem can be simplified as

$$\sum_{i=2}^{m} \left| \sum_{j=1}^{m} q_j (z_j - Z_{i-1})^+ - \left(\frac{1}{\sqrt{2}} \mathrm{e}^{-Z_{i-1}^2/2} - Z_{i-1}(1 - N(Z_{i-1})) \right) \right|, \tag{9}$$

since

$$\int_{-\infty}^{+\infty} (z - Z_{i-1})^+ f(z) \mathrm{d}z = \frac{1}{\sqrt{2\pi}} \int_{Z_{i-1}}^{+\infty} z \mathrm{e}^{-z^2/2} \mathrm{d}z - \frac{1}{\sqrt{2\pi}} \int_{Z_{i-1}}^{+\infty} Z_{i-1} \mathrm{e}^{-z^2/2} \mathrm{d}z = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-Z_{i-1}^2/2} - Z_{i-1}(1 - N(Z_{i-1})),$$

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

Table 3. Comparison of kurtosis, the objective function value of (9), z_1 and z_m for Curran's sampling strategy, the first partial moment matching strategy and our proposed strategy with different γ and m.

m		30	40	50
Curran's sampling	K	2.8069	2.8531	2.8813
	Obj	0.0864	0.0728	0.0646
	z_1	–2.2692	–2.3766	–2.4575
	z_m	2.2692	2.3766	2.4575
LP strategy	K	2.9271	2.9503	2.9632
$\gamma = 0.3$	Obj	0.0232	0.0175	0.0143
	z_1	–2.4653	–2.6931	–2.7886
	z_m	2.4653	2.6931	2.7886
LP strategy	K	3.0009	3.0042	3.0049
$\gamma = 0.6$	Obj	0.0057	0.0043	0.0034
	z_1	–2.8827	–3.0283	–3.1370
	z_m	2.8827	3.0283	3.1370
FPM strategy	K	2.9931	2.9972	2.9986
$\gamma = 0.3$	Obj	0.0042	0.0031	0.0023
	z_1	–2.6355	–2.7620	–2.8561
	z_m	2.6355	2.7620	2.8561
FPM strategy	K	3.0411	3.0102	2.9988
$\gamma = 0.6$	Obj	0.0029	0.0028	0.0031
	z_1	–2.9643	–2.9999	–3.0000
	z_m	2.9643	2.9999	3.0000

where $N(\cdot)$ is the standard normal distribution cumulative density function. Table 3 lists the two end points of $\{z_i\}$, the kurtosis and the objective function of (9) for Curran's sampling strategy, our new sampling strategy (LP strategy) and the first partial moment matching strategy (FPM strategy) with $\gamma = 0.3$ and $\gamma = 0.6$. It illustrates that the kurtosis values from our new sampling strategy and the first partial moment matching strategy are much closer to three and $\{z_i\}$ spreads more widely than Curran's sampling strategy, especially when the spatial number m is not large.

Since $\{q_i\}_{i=1,2,\ldots,m}$ in our new sampling strategy is not uniform, the objective function of the LP problem (4) should be modified as

$$\min_{p_{ij}^k} \sum_{i=1}^{m} q_i \sum_{j=1}^{m} p_{ij}^k |\sqrt{t_{k+1}} z_j - \sqrt{t_k} z_i|^3, \tag{10}$$

with all constraints unchanged to determine a new series of transition probability matrices $\{P^k = (p_{ij}^k) \mid k = 1, 2, \ldots, n-1\}$ by solving the new constrained LP problem. An interesting observation from solving the LP problem (4) or (10) is that when the time step h_k between t_{k+1} and t_k is relatively small, say less than 0.1, then the third equation of constraints (4) can be removed. In other words, the correct variance can be approximated for 'free' by virtue of the invariant unconditional distribution. Before giving a theoretical proof of this observation, we first introduce a lemma that will be used for the proof.

Lemma 3.1: *Given a continuous function g on $[a,b]$, $a = z_0 < z_1 < \cdots < z_m = b$ and $\omega_i \geq 0$, then there exists $\eta \in [a, b]$ such that*

$$\sum_{i=0}^{m} \omega_i g(z_i) = g(\eta) \sum_{i=0}^{m} \omega_i.$$

Proof: Assume z^* and $z_* \in [a, b]$, such that

$$g(z^*) = \max_{0 \leq i \leq m} g(z_i), \qquad g(z_*) = \min_{0 \leq i \leq m} g(z_i),$$

then we obtain

$$g(z_*) \sum_{i=0}^{m} \omega_i \leq \sum_{i=0}^{m} \omega_i g(z_i) \leq g(z_*) \sum_{i=0}^{m} \omega_i.$$

From the intermediate value theorem for the continuous function, there exists $\eta \in [a, b]$, such that $\sum_{i=0}^{m} \omega_i g(z_i) = g(\eta) \sum_{i=0}^{m} \omega_i$. □

We are now ready to prove the following proposition.

Propsition 3.2: *Given $\sum_{j=1}^{m} p_{ij}^k = 1$, $p_{ij}^k \geq 0$, and $\sum_{j=1}^{m} p_{ij}^k \sqrt{t_{k+1}} z_j = \sqrt{t_k} z_i$ from the constraints of the LP problem (4), then the constraint*

$$\sum_{j=1}^{m} p_{ij}^k t_{k+1} z_j^2 - t_k z_i^2 = h_k \tag{11}$$

can be approximated from the given conditions as h_k is relatively small.

Proof: Assume that function $g_1(z) = z$, from lemma 3.1 there exists $\eta_1 \in [z_1, z_m]$ such that

$$\sum_{j=1}^{m} p_{ij}^k \sqrt{t_{k+1}} z_j - \sqrt{t_k} z_i = \sqrt{t_{k+1}} \eta_1 \sum_{j=1}^{m} p_{ij}^k - \sqrt{t_k} z_i$$
$$= \sqrt{t_{k+1}} \eta_1 - \sqrt{t_k} z_i = 0,$$

i.e. $\eta_1^2 = (t_k/t_{k+1})z_i^2$. Similarly, define function $g_2(z) = z^2$, then there exists $\eta_2 \in [z_1, z_m]$ such that

$$h_k = \sum_{j=1}^{m} p_{ij}^k t_{k+1} z_j^2 - t_k z_i^2 = t_{k+1}\eta_2^2 - t_k z_i^2.$$

Thus, when h_k is relatively small, η_1^2 can be an approximation of η_2^2 so that

$$\sum_{j=1}^{m} p_{ij}^k t_{k+1} z_j^2 - t_k z_i^2 \approx t_{k+1}\eta_1^2 - t_k z_i^2 = 0,$$

i.e. (11) can be approximated by η_1^2 generated from the given conditions. Therefore, (11) can be removed when h_k is small. □

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

4. Applications to option pricing

In this section, we apply the willow tree method to price path-dependent options such as Asian arithmetic-average options and American moving-average options. Subsection 4.1 presents the Asian arithmetic-average pricing procedure and the American moving average option is priced in subsection 4.2.

4.1. Asian arithmetic-average option

The Asian arithmetic-average call option is one of the most commonly used path-dependent options, whose payoff function at time T is $\max\{0, (1/N)\sum_{i=1}^{N} S_i - K\}$, when there are N discrete time instances from time 0 to T. In this subsection, we construct a willow tree algorithm to price Asian options with our proposed sampling strategy.

Suppose that an arithmetic-average Asian call option with strike price K has $N+1$ monitoring time instants on a valid time interval $[0, T]$, that is

$$0 = t_0 < t_1 < \cdots < t_N = T.$$

Then, the willow tree lattice of the underlying asset can be set up as follows:

$$S_0 \begin{array}{cccc} S_{11} & S_{21} & \cdots & S_{N1} \\ S_{12} & S_{22} & \cdots & S_{N2} \\ \vdots & \vdots & \cdots & \vdots \\ S_{1m} & S_{2m} & \cdots & S_{Nm} \end{array}$$

where m is the number of spatial nodes at each monitoring time instant except the initial time t_0 and the underlying asset price at time t_i on node j is $S_{ij} = S_0 e^{(r-\sigma^2/2)t_i + \sigma\sqrt{t_i}z_j}$ for $i = 1, \ldots, N$ and $j = 1, \ldots, m$ (Ho 2000). The transition probabilities of the asset price from time t_0 to t_1 are set as $q = (q_1, q_2, \ldots, q_m)$ and the transition probability matrix from time t_i to t_{i+1}, for $i = 1, 2, \ldots, N-1$, is defined as

$$P^i = \begin{pmatrix} P_{11}^i & P_{12}^i & \cdots & P_{1m}^i \\ P_{21}^i & P_{22}^i & \cdots & P_{2m}^i \\ \vdots & \vdots & \vdots & \vdots \\ P_{m1}^i & P_{m2}^i & \cdots & P_{mm}^i \end{pmatrix},$$

where P_{kj}^i represents the transition probability from node k at time t_i to node j at time t_{i+1}. Similarly to the binomial tree method (Jiang 2008), we introduce a path variable F^i, an $m \times m^{i-1}$ matrix, at each monitoring instant t_i for $i = 1, \ldots, N$, that is

$$F^0 = S_0, \ F^1 = \begin{pmatrix} \frac{F^0}{2} \\ \frac{F^0}{2} \\ \vdots \\ \frac{F^0}{2} \end{pmatrix} + \begin{pmatrix} \frac{S_{11}}{2} \\ \frac{S_{12}}{2} \\ \vdots \\ \frac{S_{1m}}{2} \end{pmatrix},$$

and

$$F^i = \begin{pmatrix} \frac{i}{i+1} reshape(F^{i-1}) \\ \frac{i}{i+1} reshape(F^{i-1}) \\ \vdots \\ \frac{i}{i+1} reshape(F^{i-1}) \end{pmatrix} + \begin{pmatrix} \frac{S_{i1}}{i+1} \\ \frac{S_{i2}}{i+1} \\ \vdots \\ \frac{S_{im}}{i+1} \end{pmatrix} \cdot (1\ 1 \cdots 1)_{1\times m^{i-1}},$$

for $i = 2, \ldots, N$, where entry F_{kj}^i represents a path from time t_0 to t_i and the operation *reshape* reorganizes the matrix $A = [A_{ij}]_{m\times n}$ by columns, that is

$$reshape(A) = (A_{11}\ \cdots\ A_{m1}\ A_{12}\ \cdots A_{m2}\ \cdots\ A_{1n} \cdots\ A_{mn}).$$

The option value at t_i can then be evaluated as $V^N = \max(F^N - K, 0)$ or

$$V^i = e^{-rT/N} \begin{pmatrix} P^i(1,:)V^{i+1}(:, 1{:}m{:}m^i) \\ P^i(2,:)V^{i+1}(:, 2{:}m{:}m^i) \\ \vdots \\ P^i(m,:)V^{i+1}(:, m{:}m{:}m^i) \end{pmatrix},$$

for $i = N-1, \ldots, 1$. Finally, the Asian option price at time t_0 can be evaluated as

$$V^0 = e^{-rT/N} \sum_{j=1}^{m} q_j V_j^1.$$

The above procedure shows that only path variable F^N is required for option pricing, so it is not necessary to store $F^0, F^1, \ldots, F^{N-1}$ during the computation. The following example illustrates the computing procedure for the Asian option.

Assume that we use a willow tree method to price an Asian option with $m = 2$ and $N = 3$. The corresponding willow tree lattice for the underlying asset price can be constructed as

$$\begin{array}{cccc} & S_{11} & S_{21} & S_{31} \\ S_0 & & & \\ & S_{12} & S_{22} & S_{32} \end{array}$$

where S_{ij} is the underlying asset price on the jth space node at time t_i. The transition probability from t_i to t_{i+1} is denoted as

$$P^i = \begin{pmatrix} P^i_{11} & P^i_{12} \\ P^i_{21} & P^i_{22} \end{pmatrix}, \quad \text{for } i = 1, 2.$$

Thus, the path variable F^i can be written as

$$\begin{aligned} F^0 &= S_0, \\ F^1 &= \begin{pmatrix} \frac{S_0+S_{11}}{2} \\ \frac{S_0+S_{12}}{2} \end{pmatrix}, \\ F^2 &= \begin{pmatrix} \frac{S_0+S_{11}+S_{21}}{3} & \frac{S_0+S_{12}+S_{21}}{3} \\ \frac{S_0+S_{11}+S_{22}}{3} & \frac{S_0+S_{12}+S_{22}}{3} \end{pmatrix}, \\ F^3 &= \begin{pmatrix} \frac{S_0+S_{11}+S_{21}+S_{31}}{4} & \frac{S_0+S_{11}+S_{22}+S_{31}}{4} & \frac{S_0+S_{12}+S_{21}+S_{31}}{4} & \frac{S_0+S_{12}+S_{22}+S_{31}}{4} \\ \frac{S_0+S_{11}+S_{21}+S_{32}}{4} & \frac{S_0+S_{11}+S_{22}+S_{32}}{4} & \frac{S_0+S_{12}+S_{21}+S_{32}}{4} & \frac{S_0+S_{12}+S_{22}+S_{32}}{4} \end{pmatrix}, \end{aligned}$$

where each entry F^k_{ij} of F^k represents a path from t_0 to t_i. Based on the above pricing procedure, the option price V^i at time t_i can be computed using the following steps. First, the option price at the expiry time, V^3, is

$$\begin{aligned} V^3 &= \max(F^3 - K, 0) \\ &= \begin{pmatrix} (F^3_{11} - K)^+ (F^3_{12} - K)^+ (F^3_{13} - K)^+ (F^3_{14} - K)^+ \\ (F^3_{21} - K)^+ (F^3_{22} - K)^+ (F^3_{23} - K)^+ (F^3_{24} - K)^+ \end{pmatrix}. \end{aligned}$$

Secondly, since only paths F^3_{11} and F^3_{21} at time t_3 are generated from path F^2_{11} at time t_2, the option value for the path on F^2_{11} at time t_2 is $e^{-rT/3}(P^2_{11}V^3_{11} + P^2_{12}V^3_{21})$. Thus, with a similar analysis on the other six paths at time t_3, we have the option value for each path at time t_2:

$$V^2 = e^{-rT/3} \begin{pmatrix} P^2_{11}V^3_{11} + P^2_{12}V^3_{21} & P^2_{21}V^3_{13} + P^2_{22}V^3_{23} \\ P^2_{11}V^3_{12} + P^2_{12}V^3_{22} & P^2_{21}V^3_{14} + P^2_{22}V^3_{24} \end{pmatrix}.$$

Thirdly, the option values at time t_1 are

$$\begin{aligned} V^1_1 &= e^{-rT/3}(P^1_{11}V^2_{11} + P^1_{12}V^2_{21}), \\ V^1_2 &= e^{-rT/3}(P^1_{21}V^2_{12} + P^1_{22}V^2_{22}). \end{aligned}$$

Finally, the option price at time t_0 can be calculated as $V^0 = e^{-rT/3}(q_1V^1_1 + q_2V^1_2)$.

4.2. *American moving-average option (AMAO)*

American moving-average options are another path-dependent and American-type derivative whose payoff depends on the moving average of stock prices over a fixed length period. Kishimoto (2004), Dai *et al.* (2010), and Bernhart *et al.* (2011) have studied this type of option using the extended tree method, and the least-squares Monte Carlo and PDE method, respectively. In this subsection, we first give an example of this type of option and then apply the willow tree method to price these options.

Suppose that an American moving-average option contract has $N + 1$ monitoring time instants including the initial and terminal instants during the valid time period $[0, T]$. In other words, the valid time interval $[0, T]$ can be partitioned into N subintervals:

$$0 = t_0 < t_1 < \cdots < t_N = T.$$

If the moving-average length is over n monitoring time instants, then the option can be exercised at any time t_k where $n - 1 \le k \le N$, and the payoff can be determined by the average of the underlying asset prices $(S_{t_{k-n+1}}, \ldots, S_{t_{k-1}}, S_{t_k})$. Thus, we can define the payoff function at time t_k as $PF(\cdot)$ with respect to the path from t_{k-n+1} to t_k.

Similar to Asian options, a path variable F^i is introduced at each time instant t_i, for $i = 0, 1, \ldots, N$. Note that when $n = 1$ and $n = N + 1$ the American moving-average options are reduced to vanilla American options and Asian options, respectively, so we only consider situations with $2 \le n \le N$. Since the American moving-average option cannot be exercised before time t_n, the first n path variables F_i are the same as those for Asian options. For $n \le i \le N$, F^i can be generated through an $m \times m^{n-1}$ matrix as discussed in the previous subsection. Due to the characteristics of the American moving-average option, it only depends on the stock prices at time instants $t_{i-n+1}, \ldots, t_{i-1}, t_i$. The core of pricing American moving-average options is to find the relationship between F^i and F^{i+1}, and then the value of the option can be calculated by recursion from back to forth.

Similar to Asian options, we obtain

$$F^0 = S_0,$$

$$F^i = \begin{pmatrix} \frac{i}{i+1}\, reshape\ (F^{i-1}) \\ \frac{i}{i+1}\, reshape\ (F^{i-1}) \\ \vdots \\ \frac{i}{i+1}\, reshape\ (F^{i-1}) \end{pmatrix} + \begin{pmatrix} \frac{S_{i1}}{i+1} \\ \frac{S_{i2}}{i+1} \\ \vdots \\ \frac{S_{im}}{i+1} \end{pmatrix} \cdot (1\ 1 \cdots 1)_{1 \times m^{i-1}},$$

where $i = 1, 2, \ldots, n - 1$. For $n \le i \le N$, let

$$J^0_i = \begin{pmatrix} S_{i-n+1,1} \\ S_{i-n+1,2} \\ \vdots \\ S_{i-n+1,m} \end{pmatrix}$$

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

Table 4. Comparison of the relative errors and running times for different methods with $S_0 = 100$, $K = 95$, $T = 1$ and different r and σ of American options.

σ		0.1		0.2		0.4	
r	Method	Error	CPU time	Error	CPU time	Error	CPU time
0.03	B100	4.5e–3	0.2656	3.2e–3	0.2813	–6.1e–4	0.2813
	MC	3.3e–3	36.0938	9.1e–3	54.9219	4.1e–2	71.9063
	W	–4.6e–3	0.1094	6.1e–3	0.1094	3.3e–2	0.1094
	NW	9.0e–4	0.1250	–4.9e–3	0.1250	–8.0e–3	0.1094
	WFM	7.2e–4	0.1290	–3.2e–3	0.1324	–9.0e–3	0.1123
0.05	B100	2.1e–3	0.2500	1.8e–3	0.2813	6.8e–3	0.2969
	MC	1.2e–2	32.8906	3.4e–3	53.5156	–6.8e–3	75.6406
	W	–1.4e–2	0.1250	1.1e–2	0.1094	–7.3e–3	0.1250
	NW	1.2e–3	0.1250	4.0e–4	0.1250	–1.8e–3	0.1250
	WFM	1.1e–3	0.1023	3.1e–4	0.1011	–1.8e–3	0.1345
0.08	B100	2.8e–3	0.2656	2.8e–3	0.2656	2.6e–3	0.2969
	MC	–7.4e–3	30.7031	–9.5e–3	50.3750	1.2e–2	68.6094
	W	1.5e–2	0.1094	1.5e–2	0.1250	1.3e–2	0.1094
	NW	2.3e–4	0.1250	–1.2e–3	0.1094	–6.4e–5	0.1250
	WFM	3.2e–4	0.1235	–2.1e–3	0.1123	–7.2e–5	0.1267

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

and

$$J_i^j = \begin{pmatrix} \frac{j}{j+1}\, reshape\ (J_i^{j-1}) \\ \frac{j}{j+1}\, reshape\ (J_i^{j-1}) \\ \vdots \\ \frac{j}{j+1}\, reshape\ (J_i^{j-1}) \end{pmatrix} + \begin{pmatrix} \frac{S_{i-n+1+j,1}}{j+1} \\ \frac{S_{i-n+1+j,2}}{j+1} \\ \vdots \\ \frac{S_{i-n+1+j,m}}{j+1} \end{pmatrix} \cdot (1\ 1\ \cdots\ 1)_{1\times m^j}, \quad j = 1, 2, \ldots, n-1,$$

then $F^i = J_i^{n-1}$, where function $reshape(\cdot)$ is defined in subsection 4.1.

Note that V^i is the option value at time t_i, for $i = 0, 1, \ldots, N$, then we have $V_{jk}^N = PF(F_{jk}^N)$, where V_{jk}^N and F_{jk}^N are the (j, k) entries of V^N and F^N, respectively. When t_i is in the interval $[t_{n-1}, t_{N-1}]$, i.e. $t_{n-1} \le t_i \le t_{N-1}$, it can be exercised in advance, so we have

$$V_{jk}^i = \max\left(\mathrm{e}^{-rT/N} \sum_{l=1}^{m} P_{jl}^i V_{l,(k-1)m+j}^{i+1}, PF(F_{jk}^i)\right),$$

where P_{jl}^i is the (j, l) entry of the probability transition matrix P^i at time t_i, $c = k \bmod m^{n-2}$ and if $k \bmod m^{n-2} = 0$, then $c = m^{n-2}$. When t_i belongs to $[t_0, t_{n-1}]$, the American moving-average option can be treated as an Asian option whose terminal payoff is V^{n-1}. With the same algorithm for pricing Asian options, we can finally obtain the value of the American moving-average option V^0 as follows:

$$V_{jk}^i = \mathrm{e}^{-rT/N} \sum_{l=1}^{m} P_{jl}^i V_{l,j+(k-1)m}^{i+1},$$

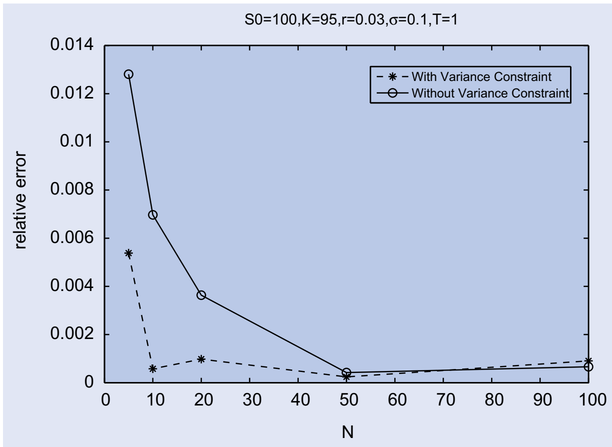

Figure 4. Relative errors on American option pricing with and without the variance constraint at $S_0 = 100$, $K = 95$, $r = 0.03$, $\sigma = 0.1$ and $T = 1$.

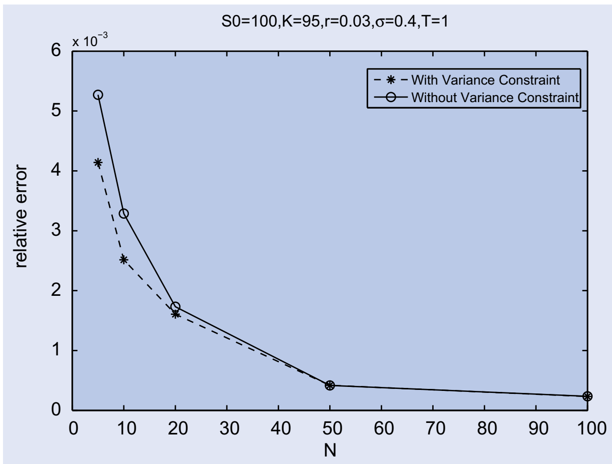

Figure 5. Relative errors on American option pricing with and without the variance constraint at $S_0 = 100$, $K = 95$, $r = 0.03$, $\sigma = 0.4$ and $T = 1$.

Table 5. Comparisons of the relative errors and running times for different methods with $S_0 = 100$, $K = 95$, $T = 1$ and different r and σ of Asian options.

σ		0.1		0.2		0.4	
r	Method	Error	CPU time	Error	CPU time	Error	CPU time
0.03	W	1.7e–3	4.9219	2.4e–3	4.8750	5.7e–3	4.7656
	NW	2.5e–4	4.8594	8.1e–5	4.6719	4.9e–5	4.5156
	WFM	3.0e–4	4.9921	7.2e–5	4.7001	3.2e–5	4.5111
	M2E+6	6.3e–4	4.8438	9.1e–4	4.8750	1.1e–3	4.9063
0.05	W	2.2e–3	5.0625	1.7e–3	5.0315	5.4e–3	5.0000
	NW	2.7e–4	4.8281	1.1e–4	4.6875	2.2e–5	4.5781
	WFM	3.0e–4	4.9821	2.1e–4	4.7623	3.2e–5	4.8211
	M2E+6	5.8e–4	4.8750	8.7e–4	4.8750	1.1e–3	4.9531
0.08	W	2.1e–3	4.5156	6.6e–4	4.9219	4.7e–3	4.8438
	NW	3.0e–4	4.8125	1.6e–4	4.7344	1.7e–4	4.6250
	WFM	2.2e–4	4.8233	2.2e–4	4.8012	1.1e–4	4.7012
	M2E+6	4.7e–4	4.7500	8.0e–4	4.9069	1.1e–3	4.9688

for $i = n-2, \ldots, 1$, $j = 1, \ldots, m$, and $k = 1, \ldots, m^{i-1}$, then the American moving-average option price at time t_0 is

$$V^0 = \mathrm{e}^{-rT/N} \sum_{j=1}^{m} q_j V_j^1.$$

5. Numerical results

In this section, we compare the performance of American, arithmetical-averaging Asian options and American moving-average put options with strike prices K by the (least-squares) Monte Carlo method (MC), the binomial tree method (B), the willow tree with Curran's sampling strategy (W), the willow tree method with the first partial moment matching (WFM) at $\gamma = 0.3$ and willow tree with our new LP sampling strategy (NW) at $\gamma = 0.6$. All numerical experiments were carried out on a machine with an Intel(R) Core(TM)2 Duo CPU T5750 running at 2.00 GHz, 2.00 GB memory and 250 GB hard drive running Windows XP and Matlab 7.0.

Compared with the binomial tree method, since the total number of willow tree lattice points increases linearly, rather than quadratically, it is feasible to price path-dependent options using the willow tree method in practice. Consider an American option with $S_0 = 100$, $K = 95$, $T = 1$ and various interest rates and volatilities. Then, five methods, the binomial tree with 100 time steps, the least-squares Monte Carlo method with 100 time steps 5000 simulations and the willow tree method with 30 spatial points and 100 time steps sampling with Curran's strategy, the first partial moment matching and our new proposed strategy, are employed for pricing. In this experiment, the option price determined by a binomial tree with 5000 time steps is set as the exact price. We then record the computational times and relative error between the computed option price V and the exact price V_{B5000}, i.e.

$$RE = \frac{V - V_{B5000}}{V_{B5000}}.$$

From the results in table 4, the willow tree method is the fastest method of the five examined. The accuracy of the results from the willow tree method is also improved when using our proposed sampling strategy. As is known, the number of nodes of the binomial tree is about 5000, while the number for the willow tree is 3000 in this experiment. If 200 time steps are taken, the number of nodes for the binomial tree increases to 20,000, while only 6000 nodes are required for the willow tree. Therefore, as the number of time steps increases, the advantage of the willow tree will become more significant. Figures 4 and 5 show the relative errors for American option pricing with and without the variance constraint, $\sum_{j=1}^{m} p_{ij}^k t_{k+1} z_j^2 - t_k z_i^2 = h_k$, in the transition probability matrices determination in (10). This illustrates that the relative errors without the constraint converge to the errors with the constraint as the number of time steps, N, increases. In other words, as the time interval h_k is small, that is a large N, the variance constraint can be removed from the LP problem (10).

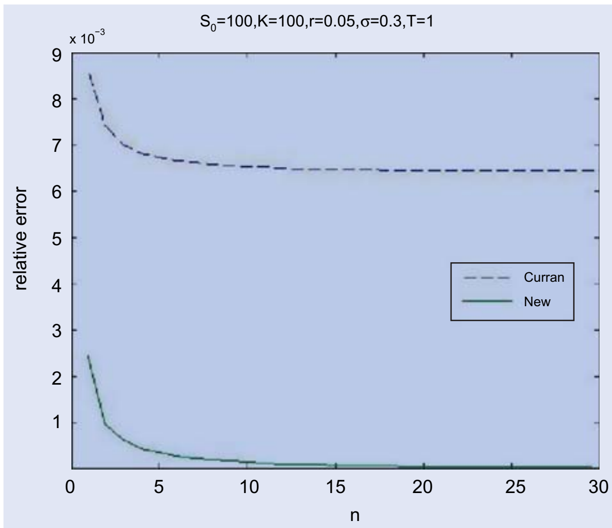

Figure 6. Relationship between the relative error and the subinterval number of time steps between two adjacent monitoring time instants k for Asian options.

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

Table 6. Option prices and CPU times of the least-squares Monte Carlo and willow tree methods with $S_0 = 100$, $T = 1$ and different r and σ for pricing American moving average options.

Method		W3050		NW3050		MC1000000		
r	σ	Value	CPU time	Value	CPU time	Value	CPU time	Confidence interval
0.03	0.1	2.1233	0.0156	2.1118	0.0156	2.1134	14.2813	[2.1115, 2.1153]
	0.2	4.0790	0.0313	4.0469	0.0313	4.0482	13.9063	[4.0442, 4.0522]
	0.4	7.8385	0.0313	7.7422	0.0313	7.7377	13.7031	[7.7287, 7.7467]
0.05	0.1	2.1968	0.0156	2.1867	0.0313	2.1890	14.3906	[2.1871, 2.1909]
	0.2	4.1493	0.0313	4.1184	0.0156	4.1202	13.9844	[4.1162, 4.1242]
	0.4	7.9031	0.0313	7.8075	0.0313	7.8033	13.7344	[7.7943, 7.8123]
0.08	0.1	2.3079	0.0313	2.3000	0.0156	2.3026	14.7188	[2.3007, 2.3045]
	0.2	4.2549	0.0313	4.2258	0.0156	4.2283	14.1563	[4.2243, 4.2323]
	0.4	8.0000	0.0156	7.9055	0.0313	7.9023	13.8906	[7.8933, 7.9113]

Next, suppose there are Asian options with six monitoring time instances including the initial and terminal time instances. The other parameters are set as $S_0 = 100$, $K = 95$, and $T = 1$. Then, Asian options are priced with various r and σ. The results from the Monte Carlo method with five million simulations are set as the exact values for the Asian options. In the willow tree method, the number of space nodes, m, is set as 30 while the number of time steps between two adjacent monitoring time instants, N, is 10, that is 50 time steps in total.

The results in table 5 show that the new sampling strategy provides a higher precision than the other two willow tree methods and the Monte Carlo method with two million simulations within similar running times. Also, we find that when subdividing the time interval $[t_i, t_{i+1}]$ into k subintervals, say

$$t_i = t_i^0 \le \cdots \le t_i^k = t_{i+1},$$

the transition probability matrix between t_i and t_{i+1} can be determined as a product of k transition matrices corresponding to k subintervals. Thus, it leads to a much higher precision when pricing Asian options. Obviously, this technique is quite expensive, but it only needs to be computed off-line once for a fixed m and n. Figure 6 shows that, as the number of subintervals between two adjacent monitoring time instants increases, the relative error of the willow tree method decreases. It also turns out that our new sampling strategy provides a much smaller relative error than Curran's strategy due to the good discrete approximation of the standard normal distribution. In addition, our proposed sampling strategy has a similar accuracy as the first partial moment matching strategy for American and Asian option pricing and both are much better than Curran's original strategy.

The last experiment was carried out for American moving-average put options with a floating strike price. The number of monitoring time instants of the valid time period is $N = 12$ and the moving-average period is $n = 3$. Table 6 shows the numerical results for American moving-average options with parameters $S_0 = 100$, $T = 1$ and various r and σ. Here we compare the results from the least-squares Monte Carlo method with one million simulations (MC1M) and 30 spatial nodes and 50 time steps willow tree method with Curran's and our new sampling strategy. From the table, the results from our new sampling strategy are similar to those for the least-squares Monte Carlo method and both sets of results are within the 95% confidence level. It also shows that the willow tree method is 1000 times faster than the Monte Carlo method.

6. Conclusion and further work

In this paper we introduce a new algorithm for option pricing, the willow tree method, and propose a new sampling strategy with a better discrete approximation of the standard normal distribution than those of the existing sampling strategies. We then apply the strategy to price two types of path-dependent options, Asian options and American moving-average options, and compare the results with those for the (least-squares) Monte Carlo method, the binomial tree method and the willow tree method with Curran's and the first partial moment matching sampling strategy. The numerical results show that our proposed method is much faster than the (least-squares) Monte Carlo method with similar precision. On the other hand, our strategy can price deep-out-of-the-money options precisely, whereas Curran's strategy exhibits a large bias.

The willow tree need not only be applied to path-dependent options, but can also be extended to other models, such as jump-diffusion models, GARCH models, multi-factor models, etc. Our future work will focus on multi-factor models and those options whose underlying asset prices are not standard geometric Brownian motion.

Acknowledgements

This work was supported by the Natural Science Foundation of China (project No. 11101310), Specified Research Fund for the Doctoral Program of Higher Education (SRFDP) (20110072120008) and the Fundamental Research Funds for the Central Universities. We are grateful to the anonymous reviewer for his/her valuable comments and suggestions that helped make the paper more precise and readable.

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

References

Bernhart, M., Tankov, P. and Warin, X., A finite dimentional approximation for pricing moving average options. *SIAM Financ. Math.*, 2011, **2**(1), 989–1013.

Cox, J.C., Ross, S.A. and Rubinstein, Rubinstein, Option pricing: a simplified approach. *J. Financ. Econ.*, 1979, **7**, 229–263.

Curran, M., Willow power. 1998, Technical report, Riskcare, London.

Dai, M., Li, P. and Zhang, J.E., A lattice algorithim for pricing average barrier options. RMI Working Paper No. 10/03, 2010.

Dennis, J.E. and Schnabel, R.B., *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, 1979 (SIAM: Philadelphia, PA).

Haussmann, U.G. and Yan, L., The modified willow tree algorithm. *J. Comput. Finance*, 2005, **8**(3), 63–80.

Ho, A., Willow tree. MSc thesis, University of British Columbia, 2000.

Hull, J.C., *Options, Futures, and Other Derivates*, 6th ed., 2006 (Pearson Education: New York).

Jiang, L., *Mathematical Modeling and Mehods of Option Pricing*, 2nd ed., 2008 (High Education Publishing: Beijing).

Jiang, L. and Dai, M., Convergence of binomial tree methods for European/American path dependent options. *SIAM Numer. Anal.*, 2005, **42**(3), 1094–1109.

Kishimoto, N., Pricing path-dependent securities by the extended tree method. *Mgmt Sci.*, 2004, **50**(9), 1235–1248.

Longstaff, F.A. and Schwartz, E.S., Valuing American options by simulation: a simple least-square approach. *Rev. Financ. Stud.*, 2001, **14**(1), 113–147.

Xu, W., Coleman, T.F. and Liu, G., A secant method for nonlinear least-squares minimization. *J. Comput. Optimiz. Applic.*, 2012, **51**, 159–173.

Downloaded by [Mr Zhiwu Hong] at 19:31 21 April 2013

View publication stats