

Brownian Meanders, Importance Sampling and Unbiased Simulation of Diffusion Extremes

Nan Chen* and Zhengyu Huang

Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong

Abstract

Computing expected values of functions involving extreme values of diffusion processes can find wide applications in financial engineering. Conventional discretization simulation schemes often converge slowly. We propose a Wiener-measure-decomposition based approach to construct unbiased Monte Carlo estimators. Combined with the importance sampling technique and the Williams path decomposition of Brownian motion, this approach transforms simulating extreme values of a general diffusion process to simulating two Brownian meanders. Numerical experiments show this estimator performs efficiently for diffusions with and without boundaries.

Key words: stochastic differential equation; exact simulation; importance sampling; extreme values; Brownian meanders

1. Introduction

In financial engineering applications, extreme values of diffusion processes are widely used to model the maximum or minimum of underlying asset prices over the life of some path-dependent options. For the purpose of pricing such options, we need to compute the expected values of functions involving these extremes efficiently. More specifically, consider a time horizon $[0, T]$ and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a standard Brownian motion $\{W_t : 0 \leq t \leq T\}$. Let $\{\mathcal{F}_t, 0 \leq t \leq T\}$ be the augmented filtration generated by $\{W_t\}$. Suppose that a stochastic process $\{S_t, 0 \leq t \leq T\}$ is defined by the following stochastic differential equation (SDE):

$$dS_t = \mu(S_t)dt + \sigma(S_t)dW_t, \quad S_0 = s, \quad (1)$$

where $\mu : \mathbb{R} \rightarrow \mathbb{R}$ and $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ are both Borel-measurable functions. Introduce two extremes of $\{S_t\}$, $M_T := \max_{0 \leq t \leq T} S_t$ and $m_T := \min_{0 \leq t \leq T} S_t$, respectively. We shall discuss Monte Carlo simulation

schemes in this paper to obtain efficient estimates to $\mathbb{E}[f(S_T, M_T, m_T)]$ for a given measurable function f .

One typical financial application of this formulation is to evaluate the *floating strike lookback option*. Its (non-arbitrage) price can be expressed as

$$e^{-rT} \mathbb{E}[M_T - S_T | S_0 = s],$$

where r is the risk-free interest rate (see, e.g., Chapter 7.4 of Shreve [32]). And another popular exotic option, the *up-and-out barrier call*, admits the following no-arbitrage price representation

$$e^{-rT} \mathbb{E}[(S_T - K)^+ 1_{\{M_T < B\}} | S_0 = s],$$

where K is the option strike price and B is a pre-specified barrier (see, e.g., Chapter 7.1 of Shreve [32]).

Applying the conventional discretization simulation schemes to SDEs, one can easily build up a naïve algorithm to estimate the preceding expectations. Fix a large integer N and let $h = T/N$. Discretize Eq. (1) into

$$\hat{S}_{i+1} = \hat{S}_i + \mu(\hat{S}_i)h + \sigma(\hat{S}_i)\Delta W_i, \quad (2)$$

for $0 \leq i \leq N - 1$, where $\Delta W_i \sim N(0, h)$. To get an approximate value to M_T or m_T , we simulate a sequence of i.i.d. Gaussian random variables ΔW_i , $1 \leq i \leq N$, and use the recursive relation

*Corresponding Author. 709A, William Mong Engineering Building, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. Tel: +852-39438237. Fax: +852-26035505. Email: nchen@se.cuhk.edu.hk, URL: <http://www.se.cuhk.edu.hk/nchen/>

(2) to generate sample paths for \hat{S} . Then, the Euler approximations to M_T and m_T are formed by $\hat{M}_N := \max_{0 \leq i \leq N} \hat{S}_i$ and $\hat{m}_N := \min_{0 \leq i \leq N} \hat{S}_i$. Repeat the above procedure to draw L independent samples of $(\hat{S}_N^j, \hat{M}_N^j, \hat{m}_N^j)$, $1 \leq j \leq L$. Averaging the results across all samples will lead to the following Monte Carlo estimates:

$$\frac{1}{L} \sum_{j=1}^L f\left(\hat{S}_N^j, \hat{M}_N^j, \hat{m}_N^j\right). \quad (3)$$

As N tends to infinity, $(\hat{S}_N, \hat{M}_N, \hat{m}_N)$ converge jointly to (S_T, M_T, m_T) under some regularity conditions. However, Asmussen, Glynn, and Pitman [4] point out that the convergence of \hat{M}_N and \hat{m}_N turns out to be the bottleneck of this naïve algorithm. They prove that \hat{M}_N converges to M_T in distribution at a rate of $1/\sqrt{N}$. In contrast, note that the convergence rate of \hat{S}_N to S_T is in the order of $1/N$ (see, e.g., Theorem 14.5.2 of Kloeden and Platen [26]).

As an alternative to overcome the difficulty of the conventional discretization methods, we propose an unbiased estimator in this paper on the basis of a combination of the Wiener measure decomposition and the importance sampling technique. In the case of Brownian motion, the joint distribution of $(W_T, \max_{0 \leq t \leq T} W_t, \min_{0 \leq t \leq T} W_t)$ is known explicitly so that we can simulate them exactly. In order to cope with a general diffusion governed by the SDE (1), we take a Wiener-measure-decomposition based approach to break the distribution law of (S_T, M_T, m_T) down to some quantities related to the Brownian motion $\{W_t\}$. This decomposition enables us to develop an importance sampling estimator to $\mathbb{E}[f(S_T, M_T, m_T)]$ through the exact simulation of $(W_T, \max_{0 \leq t \leq T} W_t, \min_{0 \leq t \leq T} W_t)$.

The aforementioned Wiener measure decomposition yields an importance sampling weight which includes

$$\exp\left(-\int_0^T \phi(W_u) du\right)$$

for some deterministic function ϕ . Generally speaking, this exponential does not admit any closed-form expressions and thus poses a technical obstacle for the implementation of our Monte Carlo procedure. A Poisson kernel is therefore introduced in this paper to estimate this exponential function, based on a key observation that it can be regarded as the probability of the event that no arrivals occur in $[0, T]$ for a doubly stochastic Poisson process with a random

intensity $\{\phi(W_t), 0 \leq t \leq T\}$. Combined with the celebrated Williams path decomposition of Brownian motion and a simulation of Brownian meanders, the Poisson kernel estimator demonstrates a high degree of efficiency and accuracy for diffusion processes, with or without boundaries. Notice that diffusions with boundaries are used popularly in financial engineering to model asset prices, interest rates, and volatilities. Furthermore, our estimators are also amenable to the application of some variance reduction techniques to improve the estimation quality further.

The mean-square error (MSE) of our estimators outperforms significantly the conventional biased estimators such as the ones presented in (3). The unbiasedness of our estimators leads to a fact that they suffer only from the simulation variance. Therefore, according to the central limit theorem, their MSE will be reduced at a rate of $1/\sqrt{L}$ as we increase the simulation size L . In contrast, the discretization estimators suffer from the discretization bias as well, which makes the MSE of such estimators converge to zero at a much slower rate even after an optimal arrangement of computational efforts. One may refer to Duffie and Glynn [14] or Chapter 6.3.3 of Glasserman [17] for a detailed discussion on the issue of convergence of MSE under the discretization schemes. From this point of view, our estimators should be more preferred.

This paper is related to the burgeoning literature of the exact simulation of diffusions. Beskos and Roberts [8] investigate an acceptance-rejection approach to simulate exactly from the marginal distribution of one-dimensional diffusions. Beskos, Paspiliopoulos, and Roberts [9, 10], Casella and Roberts [12], and Chen [11] extend the method to a more general class of diffusions. Giesecke and Smelov [16] apply the localization technique proposed by Chen [11] to generate exact samples of jump diffusion models. In contrast to the previous literature, the Poisson-kernel based estimator led by the current paper enjoys higher efficiency because we can evaluate the importance sampling weight fast with the help of the Wiener measure decomposition and Brownian meander simulation. However, the exact diffusion simulation algorithms spend a large amount of time generating samples that will be discarded eventually.

This paper contributes to the literature of SDE extremes simulation as well. Andersen and Brotherton-Ratcliffe [2], Beaglehole, Dybvig, and Zhou [5], and

Baldi [6] suggest interpolating over each interval $[ih, (i + 1)h]$ with a Brownian bridge and then using its maximum/minimum to approximate M_T or m_T . Kabaier [24] and Giles [17] propose the Richardson-Romberg extrapolation to generate SDE samples on multiple levels of time grids, using coarser grids to simulate a crude framework and finer grids to fine-tune the bias. In conjunction with the previous Brownian-bridge interpolation, the multi-level simulation method is proven to lead to a significant bias reduction and computational cost saving. Compared with these alternatives, our approach is more systematic and provides unbiased estimators to extreme-value-related payoffs. The numerical experiments in this paper also display that it has a more appealing performance.

Our method is also rooted in the literature of diffusion sample path decomposition in the probability theory. As noted by Williams [34], the sample path of a Brownian motion, if the maximum/minimum is given, can be decomposed into two “back-to-back” Brownian meanders. Imhof [21] shows that we can further express the Brownian meander as the square root of a sum of three squared independent Brownian bridges. These two facts play an essential role in the construction of the proposed unbiased estimators. Rogers and Williams [31] and Pitman and Yor [29] extend the Williams’ path decomposition to killed Brownian motions and a general class of one-dimensional diffusions, respectively. The results of this paper spell out a computational way to realize their theoretical discoveries.

The remainder of this paper is organized as follows. In Section 2, we present a Wiener measure decomposition to form the theoretical foundation for our Monte Carlo estimators. Section 3 discusses the implementation details, making use of $\mathbb{E}[f(S_T, M_T)]$ as an example. Section 4 extends the method to cover more general payoff functions and diffusions with boundaries. Section 5 summarizes some numerical experiments and related variance reduction techniques.

2. A Wiener Measure Decomposition

Denote $D_S := (\underline{s}, \bar{s})$ to be the domain of the diffusion process $\{S_t\}$. To facilitate the simulation procedure, we need to transform the original process $\{S_t\}$ into a more tractable one. For this purpose, assume that

Assumption 2.1. $\mu(x)$ is continuously differentiable on (\underline{s}, \bar{s}) and $\sigma(x)$ is twice continuously differentiable on (\underline{s}, \bar{s}) . $\sigma(x) > 0$ when $x \in (\underline{s}, \bar{s})$. In any compact subset of D_S , the function $1/\sigma(\cdot)$ should be locally integrable.

Introduce a transform, known as the Lamperti transform in the probability literature (see, e.g., Florens [15]), such that

$$F(x) = \int_{\underline{s}}^x \frac{1}{\sigma(u)} du$$

for any interior point $x \in (\underline{s}, \bar{s})$.

It is apparent that F should be well-defined due to Assumption 2.1. Furthermore, it is strictly increasing because the integrand $\sigma(u) > 0$. Let $Y_t := F(S_t)$. Ito’s lemma implies that the process $\{Y_t\}$ satisfies the following new SDE:

$$dY_t = b(Y_t)dt + dW_t, \quad (4)$$

where the new drift function b is given by

$$b(y) = \frac{\mu(F^{-1}(y))}{\sigma(F^{-1}(y))} - \frac{1}{2}\sigma'(F^{-1}(y)).$$

By the monotone property of F , we know that simulation of (S_T, M_T, m_T) is equivalent to simulation of $(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t)$ through the following relationship: $S_T = F^{-1}(Y_T)$, $M_T = F^{-1}(\max_{0 \leq t \leq T} Y_t)$, and $m_T = F^{-1}(\min_{0 \leq t \leq T} Y_t)$.

The Lamperti transform maps D_S into the domain of Y , $D_Y := (y, \bar{y})$. In this paper, we consider two cases: either $D_Y = (-\infty, +\infty)$ or $D_Y = (y, +\infty)$. In some models of financial interest, we will encounter $D_Y = (-\infty, \bar{y})$. However, it can be treated in a similar way as $D_Y = (y, +\infty)$. Let us investigate the case $D_Y = (-\infty, +\infty)$ first in this section and the next, and defer the other one to a later discussion in Section 4. We need an additional technical assumption to ensure the two boundaries $\pm\infty$ are unattainable.

Assumption 2.2 ($D_Y = (-\infty, +\infty)$). *There exist $E > 0$, $K > 0$ such that $b(y) \geq Ky$ for all $y \in (-\infty, -E)$ and $b(y) \leq Ky$ for all $y \in (E, +\infty)$.*

This assumption imposes a limit on the growth rate of function b : when the process Y approaches $\pm\infty$, $b(Y)$ is assumed to be bounded from above (or below) by a linear function. The sub-linearity of function b prevents Y from exploding, i.e., reaching $\pm\infty$ in a finite time horizon. Since we are interested in simulating extremes of the diffusion, this prevention

should be necessary to exclude the cases in which the maximum/minimum over $[0, T]$ is $\pm\infty$. It is worth pointing out that although Assumption 2.1 defines the constraints in terms of Y , it is simple to check whether it holds with the original process S . Observe that the unattainability of the boundaries of D_Y implies that the boundaries of D_S are also unattainable. Under this assumption, we can show that

Lemma 2.1. *Suppose that Assumption 2.2 holds. Let $\tau = \inf\{t \geq 0 : Y_t \notin (-\infty, \infty)\}$. Then, $\mathbb{P}[\tau = +\infty] = 1$.*

Proof. This conclusion is proved in Proposition 1 of Ait-Sahalia [1]. \square

The advantage of introducing Y is that we can find an explicit expression for the likelihood ratio of $(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t)$ with respect to $(W_T, \max_{0 \leq t \leq T} W_t, \min_{0 \leq t \leq T} W_t)$. This helps us build up a Wiener-measure based estimator later for the extreme-value-related option prices. Define

$$A(y) = \int_0^y b(u)du \quad \text{and} \quad \phi(y) = \frac{b^2(y) + b'(y)}{2}.$$

The following theorem presents the related result on the likelihood ratio.

Theorem 2.1. *Suppose that Assumptions 2.1 and 2.2 hold, and $h : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a Borel-measurable function. Then,*

$$\begin{aligned} & \mathbb{E}[h(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t)] \\ &= \mathbb{E} \left[h(W_T, \max_{0 \leq t \leq T} W_t, \min_{0 \leq t \leq T} W_t) \cdot B_T \right], \end{aligned}$$

where the likelihood ratio factor

$$B_T = \exp \left(A(W_T) - \int_0^T \phi(W_s)ds \right).$$

Proof. Under Assumption 2.2, it is easy to verify that b is locally square-integrable, i.e., for any $x \in \mathbb{R}$, there exists a $\delta > 0$ such that

$$\int_{x-\delta}^{x+\delta} b^2(y)dy < +\infty.$$

Using a generalized Girsanov formula (Karatzas and Shreve [22], Exercise 5.5.38, p. 352), we have

$$\begin{aligned} & \mathbb{E}[h(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t) \mathbf{1}_{\{\tau > T\}}] \\ &= \mathbb{E}[h(W_T, \max_{0 \leq t \leq T} W_t, \min_{0 \leq t \leq T} W_t) \\ & \cdot \exp \left(\int_0^T b(W_u)dW_u - \frac{1}{2} \int_0^T b^2(W_u)du \right)], \end{aligned}$$

for every finite $T > 0$.

On the other hand, applying Ito's lemma on $A(W_T)$ will lead to

$$A(W_T) = \int_0^T b(W_u)dW_u - \frac{1}{2} \int_0^T b'(W_u)du. \quad (5)$$

Substituting (5) into the above expectation,

$$\begin{aligned} & \mathbb{E}[h(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t) \mathbf{1}_{\{\tau > T\}}] \\ &= \mathbb{E}[h(W_T, \max_{0 \leq t \leq T} W_t, \min_{0 \leq t \leq T} W_t) \cdot B]. \end{aligned}$$

Lemma 2.1 excludes the possibility that Y explodes in finite time. That is, $\mathbb{P}[\tau = +\infty] = 1$. From this, it is easy to show the theorem holds. \square

3. Brownian Meanders and Importance Sampler of Diffusion Extremes

Theorem 2.1 lays out the theoretical foundation to our unbiased estimators for the extreme-value-related options. This section is devoted to illustrating the implementation details, using $\mathbb{E}[f(S_T, M_T)]$ as an example for notational simplicity. One can easily deal with $\mathbb{E}[f(S_T, m_T)]$ by exploiting the symmetric status of M_T and m_T . We defer the discussion of the more general case $\mathbb{E}[f(S_T, M_T, m_T)]$ to Section 4.

Denote $K_T := \max_{0 \leq t \leq T} W_t$ and let $\Theta_T := \inf\{u \in [0, T] : W_u = K_T\}$. Applying Theorem 2.1, we have

$$\begin{aligned} \mathbb{E}[f(S_T, M_T)] &= \mathbb{E} \left[f \left(F^{-1}(Y_T), F^{-1} \left(\max_{0 \leq t \leq T} Y_t \right) \right) \right] \\ &= \mathbb{E}[f(F^{-1}(W_T), F^{-1}(K_T)) \cdot \exp(A(W_T)) \cdot C], \end{aligned}$$

where

$$C = \mathbb{E} \left[\exp \left(- \int_0^T \phi(W_s)ds \right) \middle| \Theta_T, K_T, W_T \right]. \quad (6)$$

In view of this Wiener measure decomposition, we propose the following three-step algorithm to provide an unbiased estimator to $\mathbb{E}[f(S_T, M_T)]$:

1. generate exact samples of Θ_T, K_T and W_T ;
2. evaluate

$$\exp(A(W_T)) \cdot \mathbb{E} \left[\exp \left(- \int_0^T \phi(W_s)ds \right) \middle| \Theta_T, K_T, W_T \right];$$

3. evaluate $f(F^{-1}(W_T), F^{-1}(K_T))$.

Then we can form an unbiased estimator if we take a weighted average of $f(F^{-1}(W_T), F^{-1}(K_T))$ using the weight specified by the product in Step 2.

We explore how to implement the above three steps in the subsequent subsections.

3.1. Exact Simulation of (Θ_T, K_T, W_T)

The joint distribution of this triplet is explicitly known in the literature (see, e.g., Karatzas and Shreve [22], Problem 2.8.17, p. 102). Note that the distribution function can be inverted easily, which makes it convenient for us to simply use the inverse transform method to generate samples for this triplet. More specifically, generate three independent uniformly distributed random variables $(U, V, W) \sim \text{Unif}(0, 1)^3$ first; define sequentially that

$$\Theta_T = T \sin^2(\pi U/2), \quad K_T = \sqrt{-2\Theta_T \log(1 - V)},$$

and

$$W_T := K_T - \sqrt{2(T - \Theta_T) \left(-\log \left(\frac{W}{T - \Theta_T} \right) \right)}.$$

3.2. Simulating Importance Sampling Weight

As noted in the introduction, the difficulty in evaluating the importance sampling weight in Step 2 of our algorithm lies in the term

$$\mathbb{E} \left[\exp \left(- \int_0^T \phi(W_s) ds \right) \middle| \Theta_T, K_T, W_T \right] \quad (7)$$

because in most circumstances the expectation (7) does not yield any closed-form expressions. To circumvent it, we establish the following proposition to construct a Poisson-kernel estimator to the aforementioned conditional expectation.

Proposition 3.1. *Suppose that N is a Poisson random number with parameter ΛT for a positive constant Λ , and $\{\tau_1, \dots, \tau_N\}$ are N independent uniform random numbers in $[0, T]$. All of them are independent of the Brownian motion $\{W_t, 0 \leq t \leq T\}$. Then,*

$$\begin{aligned} & \mathbb{E} \left[\exp \left(- \int_0^T \phi(W_s) ds \right) \middle| W_T, K_T, \Theta_T \right] \\ &= \mathbb{E} \left[\prod_{i=1}^N \left(\frac{\Lambda - \phi(W_{\tau_i})}{\Lambda} \right) \middle| W_T, K_T, \Theta_T \right]. \end{aligned}$$

Proof. Conditional on the whole sample path of $\{W_t, 0 \leq t \leq T\}$ and N , the right-hand side of the equation in the proposition statement equals

$$\begin{aligned} & \mathbb{E} \left[\prod_{i=1}^N \left(\frac{\Lambda - \phi(W_{\tau_i})}{\Lambda} \right) \middle| W_t, 0 \leq t \leq T \right] \\ &= \sum_{n=0}^{+\infty} \mathbb{E} \left[\prod_{i=1}^n \left(\frac{\Lambda - \phi(W_{\tau_i})}{\Lambda} \right) \middle| \{W_t, 0 \leq t \leq T\}, N = n \right] \\ & \cdot \frac{e^{-\Lambda T} (\Lambda T)^n}{n!}. \end{aligned}$$

Furthermore, note that all τ_i 's are uniformly distributed in $[0, T]$. We have

$$\begin{aligned} & \mathbb{E} \left[\prod_{i=1}^n \left(\frac{\Lambda - \phi(W_{\tau_i})}{\Lambda} \right) \middle| \{W_t, 0 \leq t \leq T\}, N = n \right] \\ &= \left(\frac{1}{T} \int_0^T \left[\frac{\Lambda - \phi(W_t)}{\Lambda} \right] dt \right)^n. \end{aligned}$$

Combining the above two equations will yield that

$$\begin{aligned} & \mathbb{E} \left[\prod_{i=1}^N \left(\frac{\Lambda - \phi(W_{\tau_i})}{\Lambda} \right) \middle| W_t, 0 \leq t \leq T \right] \\ &= \sum_{n=0}^{+\infty} \left(\frac{1}{T} \int_0^T \left[\frac{\Lambda - \phi(W_t)}{\Lambda} \right] dt \right)^n \cdot \frac{e^{-\Lambda T} (\Lambda T)^n}{n!} \\ &= \exp \left(- \int_0^T \phi(W_t) dt \right). \end{aligned}$$

Taking expectations with respect to W_T, K_T, Θ_T , we prove the proposition. \square

Based on this proposition, we present the following procedure to estimate the expectation (7):

1. simulate $N \sim \text{Poisson}(\Lambda T)$;
2. generate independent τ_i , $1 \leq i \leq N$, each of which is from the distribution $\text{Unif}(0, T)$;
3. sort $\{\tau_1, \dots, \tau_N\}$ to obtain their order statistics: $\tau_{(1)} < \dots < \tau_{(j-1)} < \Theta_T \leq \tau_{(j)} < \dots < \tau_{(N)}$;
4. simulate $W_{\tau_{(i)}}$, $1 \leq i \leq N$, under the given W_T, K_T , and Θ_T ;
5. evaluate

$$\hat{C} = \prod_{i=1}^N \left(\frac{\Lambda - \phi(W_{\tau_{(i)}})}{\Lambda} \right). \quad (8)$$

It is easy to show from Proposition 3.1 that

$$f(F^{-1}(W_T), F^{-1}(K_T)) \cdot \exp(A(W_T)) \cdot \hat{C}$$

is an unbiased estimator to $\mathbb{E}[f(S_T, M_T)]$.

A technical problem remains open so far: how do we simulate $W_{\tau_{(i)}}$'s for given W_T, K_T , and Θ_T ? The answer is related to a classical result about the Brownian motion — the celebrated Williams path decomposition (see Williams [34] and Denisov [13]). Suppose that $W_0 = 0, \Theta_T = \theta, K_T = k$, and $W_T = y$. The decomposition asserts that $\{k - W_{\theta-u}, 0 \leq u \leq \theta\}$ and $\{k - W_{\theta+u}, 0 \leq u \leq T - \theta\}$ are two independent *Brownian meanders*. Figure 1 shows the relationship of the two meanders and the original Brownian motion. The two meanders sit back-to-back at

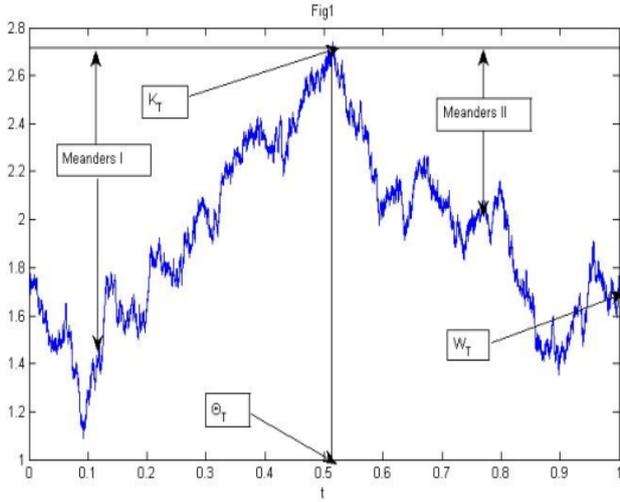


Figure 1: The Williams path decomposition of a Brownian motion. Given that the maximum is K_T , there are two Brownian meanders sitting back-to-back at Θ_T .

$\Theta_T = \theta$. Furthermore, as noted in Imhof [21], the law of a Brownian meander can be represented in terms of three independent Brownian bridges.

The Williams path decomposition of Brownian motion and the Imhof representation of Brownian meanders suggest a simulation scheme to generate W_{τ_i} 's when the triplet (Θ_T, K_T, W_T) is known. More specifically, denote $\{B_t^{1,j}, 0 \leq t \leq \theta\}$, $j = 1, 2, 3$, to be three independent Brownian bridges from 0 to 0 over $[0, \theta]$ and $\{B_t^{2,j}, \theta \leq t \leq T\}$, $j = 1, 2, 3$, to be three independent Brownian bridges from 0 to 0 over $[\theta, T]$. Given $W_0 = 0$, $\Theta_T = \theta$, $K_T = k$, and $W_T = y$, we can show that

$$\{W_u, 0 \leq u \leq \theta\} \stackrel{d}{=} k - \sqrt{\left(\frac{k \cdot (\theta - u)}{\theta} + B_u^{1,1}\right)^2 + (B_u^{1,2})^2 + (B_u^{1,3})^2} \quad (9)$$

and

$$\{W_u, \theta \leq u \leq T\} \stackrel{d}{=} k - \sqrt{\left(\frac{(k - y)(u - \theta)}{(T - \theta)} + B_u^{2,1}\right)^2 + (B_u^{2,2})^2 + (B_u^{2,3})^2}. \quad (10)$$

Now, the task of simulating $W_{\tau(i)}$, $1 \leq i \leq N$, is transformed into how to simulate Brownian bridges

$B^{l,j}$ at τ_i 's for $l = 1, 2$ and $j = 1, 2, 3$. It can be accomplished by some standard procedures in the existing literature (see, e.g., Glasserman [19], Section 3.1, pp. 82–86).

4. Extensions

4.1. General Payoff Functions

In this section, we consider a more general payoff function $f(S_T, M_T, m_T)$. Applying the Lamperti transform, Theorem 2.1, and Proposition 3.1 we obtain the following equality:

$$\begin{aligned} & \mathbb{E}[f(S_T, M_T, m_T)] \\ &= \mathbb{E} \left[f(F^{-1}(W_T), F^{-1}(K_T), F^{-1}(k_T)) \exp(A(W_T)) \hat{C} \right] \end{aligned}$$

where $k_T = \min_{0 \leq t \leq T} W_t$ and \hat{C} is defined in (8). In light of this representation, we need a small modification on the algorithm presented in the last section to construct an unbiased importance sampler for $\mathbb{E}[f(S_T, M_T, m_T)]$. More specifically, it can be done in the following six steps:

1. generate exact samples of (Θ_T, K_T, W_T) by following the procedure in Section 3.1;
2. simulate the Poisson random number N and independent uniform $\{\tau_1, \dots, \tau_N\}$ on the time interval $[0, T]$;
3. sort $\{\tau_1, \dots, \tau_N\}$ to obtain their order statistics: $\tau_{(1)} < \dots < \tau_{(j-1)} < \Theta_T \leq \tau_{(j)} < \dots < \tau_{(N)}$.
4. use the Brownian meander simulation to obtain $W_{\tau_{(1)}}, \dots, W_{\tau_{(N)}}$;
5. generate $\min_{\tau_{(i-1)} \leq t \leq \tau_{(i)}} W_t$ for all $1 \leq i \leq N$ and let

$$k_T = \min_i \min_{\tau_{(i-1)} \leq t \leq \tau_{(i)}} W_t;$$

6. construct the corresponding Poisson-kernel estimator.

Now we discuss the generation of $\min_{\tau_{(i-1)} \leq t \leq \tau_{(i)}} W_t$ in Step 5, the only technical issue that makes this algorithm different from the previous one in Section 3. Observe that

$$\min_{\tau_{(i-1)} \leq t \leq \tau_{(i)}} W_t = K_T - \max_{\tau_{(i-1)} \leq t \leq \tau_{(i)}} (K_T - W_t).$$

Recall that $\{K_T - W_t, t \in [0, \Theta_T] \cup [\Theta_T, T]\}$ consists of two pieces of Brownian meanders for given Θ_T , K_T , and W_T . Therefore, simulation of

$$me_i = \max_{\tau_{(i-1)} \leq t \leq \tau_{(i)}} (K_T - W_t)$$

can be done if we know how to sample the maximum of the Brownian meander over a fixed time interval.

Denote $\tau = \tau_{(i)} - \tau_{(i-1)}$ and

$$G_i(a; x, y, \tau) = \mathbb{P}[me_i < a | K_T - W_{\tau_{(i-1)}} = x, K_T - W_{\tau_{(i)}} = y]$$

We derive a closed-form expression of G_i in Lemma A.1 of the appendix. Then, we can generate samples of me_i by applying the standard inverse transform method. In other words, generate $U \sim \text{Unif}(0, 1)$ and let $me_i = G_i^{-1}(U)$, where G_i^{-1} is numerically inverted.

4.2. The cases of $D_Y = [\underline{y}, +\infty)$

In this subsection we consider the cases whose value domain of the transformed process Y is given by $[\underline{y}, +\infty)$. Just as in the case $D_Y = (-\infty, +\infty)$, we need a technical assumption to rule out the possibility that the process Y attains the boundaries \underline{y} or $+\infty$ in finite time.

Assumption 4.1 ($D_Y = [\underline{y}, +\infty)$). *There exist e, κ, α such that for all $\underline{y} < y < e$, $b(y) \geq \kappa(y - \underline{y})^{-\alpha}$, where either $\alpha > 1, \kappa > 0$ or $\alpha = 1, \kappa \geq 1$; There exist $E > 0, K > 0$ such that $b(y) \leq Ky$ for all $y \in (E, +\infty)$.*

The above assumption ensures the existence and uniqueness of the solution to SDE (4). When it is violated, the process Y may reach \underline{y} in $[0, T]$ with a positive probability. However, different specifications about the boundary behavior of Y on \underline{y} will lead to different solutions to the SDE. That would destroy the uniqueness of the solution — see, e.g. Chapter 15.8 of Karlin and Taylor [23] — and surely makes the simulation much more complicated. We leave the research on how to extend our method to simulate SDEs with different boundary specifications for future investigation.

Many models popularly used in financial engineering applications satisfy Assumption 4.1. Here we just mention one example, the Cox-Ingersoll-Ross (CIR) model, which is defined as

$$dS_t = \kappa(\alpha - S_t)dt + \sigma\sqrt{S_t}dW_t, S_0 = s > 0$$

where $\kappa > 0, \alpha > 0, \sigma > 0$ are constants. Its corresponding Lamperti transform is

$$F(x) = \frac{2(\sqrt{x} - \sqrt{s})}{\sigma}.$$

One of the most important characteristics of the model is its non-negativeness, i.e., $D_S = [0, +\infty)$. After the transformation, $D_Y = [-2\sqrt{s}/\sigma, \infty)$.

Under this assumption, we establish the corresponding Wiener measure decomposition in the following theorem:

Theorem 4.1. *Suppose that Assumptions 2.1 and 4.1 hold. Let $\tau = \inf\{t \geq 0 : Y_t \notin (y, +\infty)\}$. Then, $\mathbb{P}[\tau = +\infty] = 1$. Furthermore, if $h : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a Borel-measurable function, we have*

$$\begin{aligned} & \mathbb{E}[h(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t)] \\ &= \mathbb{E} \left[h(W_T, K_T, k_T) \cdot \mathbf{1}_{\{k_T > y\}} \cdot B_T \right], \end{aligned}$$

where B_T is defined in Theorem 2.1.

Proof. The first half of the theorem's conclusion is proved in Proposition 1 of Ait-Sahalia [1]. To show its second half, it is enough to prove the conclusion is true for the following h such that

$$\begin{aligned} & h(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t) \\ &= \mathbf{1}_{\{Y_T \in dx, \max_{0 \leq t \leq T} Y_t \in dy_1, \min_{0 \leq t \leq T} Y_t \in dy_2\}} \end{aligned}$$

for any $x, y_1, y_2 \in \mathbb{R}, y_2 < x < y_1$.

Define $\varsigma = \inf\{t \geq 0 : W_t = \underline{y}\}$ and two sequences of stopping times

$$\tau_k = T \wedge \inf \left\{ t \geq 0 : \int_0^t b^2(Y_u) du \geq k \right\}$$

and

$$\tau_k^W = T \wedge \inf \left\{ t \geq 0 : \int_0^t b^2(W_u) du \geq k \right\}.$$

for each integer $k \geq 0$. Furthermore, define four events such that

$$A := \{Y_T \in dx, \max_{0 \leq t \leq T} Y_t \in dy_1, \min_{0 \leq t \leq T} Y_t \in dy_2, \tau > T\};$$

$$\bar{A} := \{W_T \in dx, K_T \in dy_1, k_T \in dy_2, \varsigma > T\};$$

$$A^{(k)} := A \cap \{\tau_k = T\} \text{ and } \bar{A}^{(k)} := \bar{A} \cap \{\tau_k^W = T\}.$$

It is easy to show the following

$$\xi_t^{(k)} = \exp \left(- \int_0^{t \wedge \tau_k} b(Y_u) dW_u - \frac{1}{2} \int_0^{t \wedge \tau_k} b^2(Y_u) du \right)$$

is martingale for each k according to the Novikov condition. Introduce a sequence of new probability measures $\tilde{\mathbb{P}}_k$ such that $d\tilde{\mathbb{P}}_k/d\mathbb{P} = \xi_T^{(k)}$. Denote $\tilde{\mathbb{E}}^{(k)}$ to

be the corresponding expectation operator under the new measures. Then,

$$\mathbb{E}[\mathbf{1}_{A^{(k)}}] = \tilde{\mathbb{E}}^{(k)} \left[\mathbf{1}_{A^{(k)}} \cdot (\xi_T^{(k)})^{-1} \right]. \quad (11)$$

Using the fact $dW_t = dY_t - b(Y_t)dt$, we observe

$$(\xi_T^{(k)})^{-1} = \exp \left(\int_0^{T \wedge \tau_k} b(Y_t) dY_t - \frac{1}{2} \int_0^{T \wedge \tau_k} b^2(Y_t) dt \right).$$

The Girsanov theorem (Karatzas and Shreve [22], Corollary 3.5.13, p. 199) implies

$$Y_{t \wedge \tau_k} = \int_0^{t \wedge \tau_k} b(Y_s) ds + W_{t \wedge \tau_k}$$

is a standard Brownian motion stopped at the time τ_k under each new measure $\tilde{\mathbb{P}}_k$. Therefore, the right hand side of (11) equals to $\mathbb{E} \left[\mathbf{1}_{\bar{A}^{(k)}} B_T^{(k)} \right]$, where

$$\begin{aligned} B_T^{(k)} &:= \exp \left(\int_0^{T \wedge \tau_k^W} b(W_s) dW_s - \frac{1}{2} \int_0^{T \wedge \tau_k^W} b^2(W_s) ds \right) \\ &= \exp \left(A(W_{T \wedge \tau_k^W}) - \int_0^{T \wedge \tau_k^W} \phi(W_s) ds \right), \end{aligned}$$

where the second equality is due to Ito's formula. Therefore, we have

$$\mathbb{E}[\mathbf{1}_{A^{(k)}}] = \mathbb{E} \left[\mathbf{1}_{\bar{A}^{(k)}} \cdot B_T^{(k)} \right]. \quad (12)$$

Let $k \rightarrow +\infty$ on both sides of (12). Note that function b is continuous in the interior set of D_Y . Therefore, it must be square-integrable on every compact set in (\underline{y}, ∞) . We then have the integral $\int_0^T b^2(Y_u) du < +\infty$ on the event $\{\tau > T\}$, which implies that there exists a sufficiently large k such that $\tau^k = T$. By the dominated convergence theorem,

$$\lim_{k \rightarrow +\infty} \mathbb{E}[\mathbf{1}_{A^{(k)}}] = \mathbb{E}[\mathbf{1}_A].$$

On the other hand, following the same arguments as in the process Y , we also can show $\int_0^T b^2(W_u) du < +\infty$ on the event $\{\zeta > T\}$.

$$\lim_{k \rightarrow +\infty} \mathbb{E} \left[\mathbf{1}_{\bar{A}^{(k)}} \cdot B_T^{(k)} \right] = \mathbb{E}[\mathbf{1}_{\bar{A}} \cdot B_T].$$

In addition, noting that $\{\tau \leq T\}$ has zero probability under Assumptions 2.1 and 4.1 and the fact $\zeta > T \Leftrightarrow k_T > \underline{y}$, we have proved the result of the theorem. \square

From Theorem 4.1, we have

$$\begin{aligned} &\mathbb{E}[h(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t)] \\ &= \mathbb{E} \left[h(W_T, K_T, k_T) \cdot \mathbf{1}_{\{k_T > \underline{y}\}} \cdot \exp(A(W_T)) \cdot C \right], \end{aligned}$$

where C is given by the conditional expectation (6). Therefore in theory we still can utilize the Poisson kernel to obtain an unbiased estimator for $\mathbb{E}[h(Y_T, \max_{0 \leq t \leq T} Y_t, \min_{0 \leq t \leq T} Y_t)]$. The algorithm is summarized as follows:

1. generate exact samples of (θ_T, k_T, W_T) , where $\theta_T := \inf\{u \in [0, T] : W_u = k_T\}$;
2. if $k_T < \underline{y}$, we set the weight of this sample to be zero;
3. if $k_T > \underline{y}$, we further simulate the Poisson random number N , the order statistics $\{\tau_{(1)}, \dots, \tau_{(N)}\}$, the Brownian motion values $\{W_{\tau_{(1)}}, \dots, W_{\tau_{(N)}}\}$, and the maximum K_T ;
4. construct the corresponding Poisson-kernel estimator using \hat{C} .

Note that in the above algorithm, simulation of (θ_T, k_T, W_T) is similar to the simulation of (Θ_T, K_T, W_T) because of the symmetric property of Brownian motion.

However, different from the cases of $D_Y = (-\infty, +\infty)$, the evaluation of \hat{C} in this case sometimes suffers from a numerical instability. The reason is that $\lim_{y \rightarrow \underline{y}} b(y) = +\infty$ under Assumption 4.1, which leads to $\lim_{y \rightarrow \underline{y}} \phi(y) = +\infty$. Take the CIR process as an example. The corresponding function $b(y)$ is given by

$$\frac{4\kappa\alpha - \sigma^2}{2\sigma^2(y + 2\sqrt{s}/\sigma)} - \frac{\kappa}{2} \left(y + \frac{2\sqrt{s}}{\sigma} \right).$$

As we can see, when $y \rightarrow -2\sqrt{s}/\sigma$, $b(y)$ tends to infinity. Therefore, when we obtain some $W_{\tau_{(i)}}$'s that are very close to \underline{y} , some terms of $(\Lambda - \phi(W_{\tau_{(i)}}))/\Lambda$ could be large negative numbers and thus the absolute value of the weight \hat{C} on this sample path will be very large. This phenomenon introduces significant variance for the unbiased estimator.

To overcome this difficulty, we propose an approximation estimator: instead of calculating $\mathbb{E}[f(S_T, M_T, m_T)]$, we use the unbiased estimator to $\mathbb{E}[f(S_T, M_T, m_T) \mathbf{1}_{\{m_T > \underline{b}\}}]$, by choosing a suitable $\underline{b} \in D_S$, as an approximation to the former expectation. Since \underline{s} is unattainable for process S under Assumption 4.1, $\lim_{\underline{b} \rightarrow \underline{s}} \mathbb{P}[m_T > \underline{b}] = 1$. Therefore, if

\underline{b} is sufficiently close to \underline{s} , $\mathbb{E}[f(S_T, M_T, m_T)\mathbf{1}_{\{m_T > \underline{b}\}}]$ serves as a good approximation to the original objective $\mathbb{E}[f(S_T, M_T, m_T)]$. In the meanwhile, the new estimator avoids the issue of numerical instability because we exclude those sample paths with $m_T < \underline{b}$. The numerical examples in Section 5 corroborate this observation. It is also worth pointing out that this new estimator is still unbiased for some down-and-out options whose payoff function contains a part of $\mathbf{1}_{\{m_T > x\}}$.

5. Numerical Examples

5.1. Related Techniques of Variance Reduction

Our estimator is amenable to the application of some variance-reduction techniques. We also find that the Poisson intensity Λ has a direct influence on the estimator variance. In this subsection, we discuss how to use the stratified sampling and choose a suitable Λ to improve the algorithm efficiency.

As discussed in the previous sections, we simulate (Θ_T, K_T, W_T) by applying their respective inverted distribution functions on $(U, V, W) \sim \text{Unif}[0, 1]^3$. To make use of the stratified sampling, we stratify the three coordinates of the cube $[0, 1]^3$ into n_j , $j = 1, 2, 3$, intervals of equal length. Each stratum has the form of

$$A_{i_1, i_2, i_3} = \left[\frac{i_1 - 1}{n_1}, \frac{i_1}{n_1} \right) \times \left[\frac{i_2 - 1}{n_2}, \frac{i_2}{n_2} \right) \times \left[\frac{i_3 - 1}{n_3}, \frac{i_3}{n_3} \right)$$

for all $1 \leq i_1 \leq n_1$, $1 \leq i_2 \leq n_2$, and $1 \leq i_3 \leq n_3$ and the total number of strata is $n_1 n_2 n_3$.

Following the discussion at the beginning of Section 3, our simulation objective can be further represented as follows:

$$\sum_{i_1, i_2, i_3} \mathbb{E} \left[g(\Theta_T, K_T, W_T) \mathbf{1}_{(U, V, W) \in A_{i_1, i_2, i_3}} \right] \cdot \mathbb{P}((U, V, W) \in A_{i_1, i_2, i_3}),$$

where

$$\begin{aligned} g(\Theta_T, K_T, W_T) \\ = f(F^{-1}(W_T), F^{-1}(K_T)) \cdot \exp(A(W_T)) \cdot C. \end{aligned}$$

Note that the probability $\mathbb{P}((U, V, W) \in A_{i_1, i_2, i_3})$ is simply $1/n_1 n_2 n_3$.

The simplest way to construct a stratified unbiased estimator is proportional sampling, in which we ensure that the number of samples drawn from stratum

A_{i_1, i_2, i_3} is proportional to the theoretical probability $\mathbb{P}((U, V, W) \in A_{i_1, i_2, i_3}) = 1/n_1 n_2 n_3$. Compared with the vanilla unbiased estimator proposed in the last section, the new one eliminates sampling variability across strata and will lead to a significant variance reduction effect as shown in the numerical examples. For implementation details, one may refer to Chapter 4.3 of Glasserman [19] for a discussion on this issue.

Now, we turn to the choice of Λ . Here we face a tradeoff between the estimator variance and the computational time. A large Λ typically causes less variance, but it will lead to a large Poisson number N and require us to simulate many W_{τ_i} 's. On the other hand, we need only small computational cost for a small Λ . However, it is subject to a significant simulation variance. Figure 2 illustrates how a product, computational time \times variance, changes with Λ . This product is widely used in the literature to reflect the efficiency of Monte Carlo algorithms; see Chapter V of Asmussen and Glynn [3]. We can see this tradeoff very clearly from this figure. Our proposal is that we need to run some pilot experiments to suggest an optimal Λ to minimize the product before a larger scaled computation.

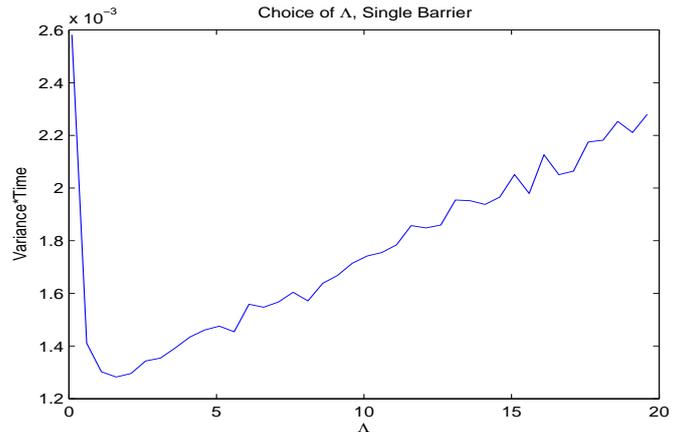


Figure 2: The dependence of Variance \times Time on Λ when we compute $\mathbb{P}[M_T < B]$ for the Ornstein-Uhlenbeck model in Section 5.2. The computational time is dominating in the right wing of the curve of the product and the variance is dominating in its left wing. We can easily find an optimal Λ to minimize the product from this figure.

5.2. Some Numerical Examples

In this subsection, we present some numerical results on four models: the geometry Brownian motion (GBM), the Ornstein-Uhlenbeck (OU) mean-

reverting process, the logistic growth model (LGM), and the CIR process. The domain of the OU process is $D_S = (-\infty, +\infty)$ and the GBM, LGM, and CIR models have $D_S = (0, +\infty)$. To illustrate the efficiency of the proposed Poisson-kernel unbiased estimators, we compare the performance of our unbiased estimator with the traditional discretization schemes such as the Euler method. The estimator proposed in this paper demonstrates significant numerical advantages over the traditional ones, whose performances are even enhanced by the Brownian interpolation technique. We also conduct comparisons in each example between our estimators and the multilevel estimators of Giles [17, 18]. The numerical experiments show that the MSE of our estimators converges very fast to zero.

Consider the GBM first. The model is described by the following SDE:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 = s_0.$$

Applying the corresponding Lamperti transform

$$F(x) = \frac{1}{\sigma} \log\left(\frac{x}{s_0}\right),$$

we have

$$dY_t = \left(\frac{\mu}{\sigma} - \frac{\sigma}{2}\right) dt + dW_t, \quad Y_0 = 0.$$

Observe the function $\phi(\cdot)$ in this model is constant. Therefore, the evaluation of C is straightforward: we do not even need the Poisson kernel at all! We omit the related numerical results for the interest of space.

The OU process is defined as a solution to the following SDE:

$$dS_t = \kappa(\alpha - S_t)dt + \sigma dW_t, \quad S_0 = s_0,$$

where κ and α are positive constants. Note that the drift is positive when $S_t < \alpha$ and negative when $S_t > \alpha$. Thus, the above process will be pulled toward the level α , a property generally referred to as mean reversion. A variation of this model is used by Vasicek [33] to model short rates in the interest rate market. Its corresponding Lamperti transform is given by $F(x) = (x - s)/\sigma$. Under this transform, we have

$$dY_t = \frac{\kappa}{\sigma} (\alpha - s - \sigma Y_t) + dW_t$$

if we let $Y_t = (S_t - s)/\sigma$.

We test the performance of different simulation schemes using two quantities $\mathbb{P}[M_T < B]$ and $\mathbb{P}[M_T < B_1, m_T > B_2]$ for some constants B, B_1 , and B_2 . The respective outcomes are reported in Tables I and II. It is apparent that the Poisson kernel estimator combined with the variance reduction technique of stratification yields the least root of mean squared errors (RMSE) in a comparable time horizon among all the numerical schemes under the test. This observation holds universally no matter which type of the payoff function we use.

To make the comparisons among various schemes clearer, we further illustrate their relative performances using the RMSE-Time plots. Figures 3 and 4 display the (log-)RMSE of each scheme shown in Tables I and II against their respective (log-)computation time. Consistent with the observation we make in the tables, we find that the Poisson-kernel estimators, with or without variance reduction, indeed enjoy faster convergence orders of RMSE against time than any others. Note that the slopes of the curves corresponding to our Poisson-kernel estimators are close to -0.5 . This observation verifies the unbiasedness of our estimators. Recall that the RMSE for an unbiased estimator is caused only by the simulation variance and therefore its slope of $\log(\text{RMSE})$ against $\log(\text{Time})$ should be -0.5 according to the central limit theorem.

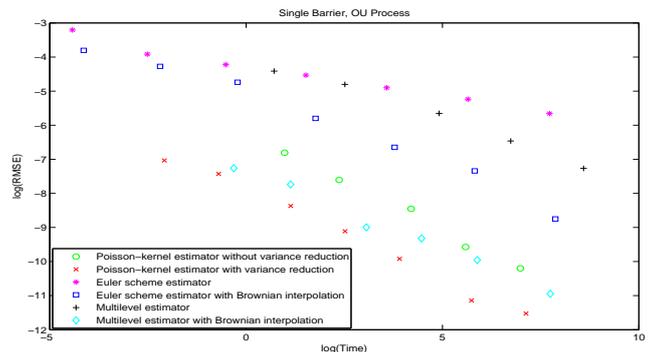


Figure 3: Performance of the six estimators in Table I. As we increase the computation time, all estimators generate more accurate results. The slopes of the curves corresponding to the Poisson-kernel estimators with or without variance reduction are -0.5735 and -0.4768 , respectively.

The third example is the CIR process. We calculate the quantities such as $\mathbb{P}[m_T > B_1]$ and $\mathbb{P}[M_T < B_2]$ for some positive B_1 and B_2 . By the discussion in Sec-

Table I: Single Barrier, Ornstein-Uhlenbeck Process					
POISSON			POISSON+VR		
Estimator	RMSE($\times 10^{-4}$)	Time(s)	Estimator	RMSE($\times 10^{-4}$)	Time(s)
0.35192290	2.1353	66.74	0.35192749	0.4914	49.76
0.35192522	0.6953	267.37	0.35192545	0.1448	311.01
0.35191967	0.2628	2163.03	0.35192596	0.0698	2481.79
EULER			EULER+BI		
Estimator	RMSE($\times 10^{-4}$)	Time(s)	Estimator	RMSE($\times 10^{-4}$)	Time(s)
0.35919531	74.4348	35.90	0.35121985	12.9748	43.88
0.35720850	53.1779	284.23	0.35202978	6.4823	337.30
0.35539874	34.8837	2262.67	0.35188983	1.5829	2630.53
MULTILEVEL			MULTILEVEL+BI		
Estimator	RMSE($\times 10^{-4}$)	Time(s)	Estimator	RMSE($\times 10^{-4}$)	Time(s)
0.35538212	35.0889	136.23	0.35185142	0.8967	86.98
0.35329272	15.5768	842.39	0.35190362	0.4734	359.57
0.35254149	6.9727	5389.18	0.35191730	0.1764	2315.27

Table 1: Numerical results of various schemes for the estimation of $\mathbb{P}[M_T < B]$ under the OU process. The parameters are $\kappa = 0.261$, $\alpha = 0.717$, $\sigma = 0.02237$, $s_0 = 0.6$, and $T = 1$. When $B = \alpha$, Yi [35] presents an explicit expression to the true probability, which equals 0.35192710 under this set of parameters. The RMSE is calculated on the basis of 10 trials. The first row documents the estimation results, the RMSE, and the computational time for our Poisson estimators with or without variance reduction (“POISSON+VR” and “POISSON”, respectively). The performance of the vanilla Euler scheme is reported under the category “EULER”. The columns with subtitles “EULER+BI” and “MULTILEVEL+BI” show the outcomes of the Euler and Multilevel combined with the Brownian interpolation.

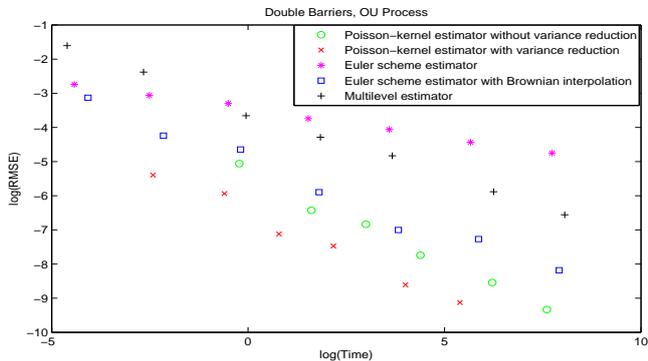


Figure 4: Performance of the five estimators in Table II. The slopes of the Poisson-kernel estimators with or without variance reduction are -0.5405 and -0.4798 , respectively. They are very close to -0.5 .

tion 4.2, our estimator is unbiased to the former one. And we use $\mathbb{P}[M_T < B_2, m_T > \underline{b}]$ to approximate the latter one by choosing a small \underline{b} . The respective numerical outcomes are reported in Tables III and IV. The numerical results in both tables suggest that our estimator is capable of producing sufficiently accurate outputs in a short time framework.

5.3. Comparison with Exact Simulation

In this subsection we compare our importance sampling method with Casella and Roberts [12]. This paper proposes an exact simulation algorithm for jump-diffusions on the basis of the acceptance-rejection method in Beskos and Roberts [8]. Our method has two obvious advantages in contrast to this literature. First, we do not need the boundedness assumption on the function ϕ as they require. Note that such an assumption rules out the OU and CIR processes, which find wide applications in financial engineering. Second, our estimator is much more efficient than theirs as shown in the following example.

To facilitate the comparison numerically, consider

Table II: Double Barriers, Ornstein-Uhlenbeck Process					
POISSON			POISSON+VR		
Estimator	RMSE($\times 10^{-3}$)	Time(s)	Estimator	RMSE($\times 10^{-3}$)	Time(s)
0.43798282	0.3649	106.19	0.43790779	0.3893	16.34
0.43798470	0.1807	665.53	0.43808808	0.1916	104.63
0.43801983	0.1003	2659.65	0.43797019	0.0891	417.07
EULER			EULER+BI		
Estimator	RMSE($\times 10^{-3}$)	Time(s)	Estimator	RMSE($\times 10^{-3}$)	Time(s)
0.45521582	17.2484	36.35	0.43797874	0.9091	45.86
0.44983521	11.8559	287.98	0.43773813	0.6950	352.72
0.44664099	8.6450	2290.27	0.43805220	0.2791	2748.02
MULTILEVEL					
Estimator	RMSE($\times 10^{-3}$)	Time(s)			
0.44458429	7.9713	39.31			
0.44070210	2.7847	518.27			
0.43931992	1.4159	3166.04			

Table 2: Numerical results of various schemes for the estimation of $\mathbb{P}[M_T < B_1, m_T > B_2]$ under the OU process. The parameters used in the table are $\kappa = 1$, $\alpha = 0$, $\sigma = \sqrt{2}$, $s_0 = 2.10$, $B_1 = -B_2 = 2.40$, and $T = 1$. The true value of $\mathbb{P}[M_T < B_1, m_T > B_2]$ is 0.4380 according to Keilson and Ross [25]. We skip the “MULTILEVEL+BI” part because we do not find an appropriate approach in the existing literature to incorporating the Brownian interpolation to the multilevel scheme for the double barrier case.

a sin model given by

$$dS_t = \sin(S_t)dt + dW_t, \quad S_0 = s,$$

and calculate the expectation of a single barrier payoff $\mathbb{E}[\max(0, S_T - K)\mathbf{1}_{\{\max_{0 \leq t \leq T} S_t < B\}}]$. Notice that the corresponding function ϕ in this model is bounded so that we can easily implement the Casella-Robert algorithm for the purpose of comparison. Table V illustrates that to achieve the same precision, the required time costs for the exact simulation are significantly larger than ours. The main bottleneck of the exact simulation is that it has to waste a large amount of time generating (and discarding) disqualified candidates.

6. Acknowledgement

This research is supported by Hong Kong Research Grant Council through the General Research Fund scheme (Research grant No. CUHK411108, CUHK411309, and CUHK411110).

A. Brownian Meanders

We give a brief introduction to Brownian meanders in this appendix. Following the terms of Pitman [28],

a Brownian meander, $\{B_t^{me}, 0 \leq t \leq T\}$, is defined as a standard Brownian motion conditioning on that it remains strictly positive for all $0 < t \leq T$. Given the ending point $B_T^{me} = r$ is fixed, the distribution law of B^{me} is identical to a three-dimensional Bessel bridge from 0 to r . Making use of this fact, we can construct the conditioned meander from three independent Brownian bridges. Namely,

$$(B_t^{me} | B_T^{me} = r) \stackrel{d}{=} \sqrt{\left(\frac{rt}{T} + B_t^{0 \rightarrow 0,1}\right)^2 + \left(B_t^{0 \rightarrow 0,2}\right)^2 + \left(B_t^{0 \rightarrow 0,3}\right)^2}$$

for $0 \leq t \leq T$, where $B_t^{0 \rightarrow 0,j}, j = 1, 2, 3$ are three independent copies of Brownian bridges from 0 to 0 on $[0, T]$. When $r = 0$, the conditioned Brownian meander becomes a Brownian excursion. One may refer to Imhof [21], Williams [34], and Revuz and Yor [30] for further discussion on this process.

Denote $G(a; x, y, \tau)$ to be the conditional distribution function of the maximum of a Brownian meander with its initial and final values being given, i.e.,

$$G(a; x, y, \tau) = \mathbb{P}\left[\max_{t \leq s \leq t+\tau} B_s^{me} < a | B_t^{me} = x, B_{t+\tau}^{me} = y\right].$$

The following lemma obtains an explicit expression for $G(a; x, y, \tau)$.

Table III: Single Barrier, CIR Process

POISSON			POISSON+VR		
Estimator	RMSE($\times 10^{-4}$)	Time(s)	Estimator	RMSE($\times 10^{-4}$)	Time(s)
0.64849668	1.9325	110.97	0.64812380	2.4896	7.73
0.64846448	1.0075	443.94	0.64845632	0.7309	30.97
0.64849171	0.5058	1175.80	0.64848162	0.4160	123.53
EULER			EULER+BI		
Estimator	RMSE($\times 10^{-4}$)	Time(s)	Estimator	RMSE($\times 10^{-4}$)	Time(s)
0.65462597	63.0412	36.23	0.64791852	14.1872	43.59
0.65297534	45.1689	285.95	0.64835165	5.8901	337.31
0.65166558	31.9592	2280.36	0.64846559	4.5837	2649.64
MULTILEVEL					
Estimator	RMSE($\times 10^{-4}$)	Time(s)			
0.65134686	54.3738	22.07			
0.64882261	14.7209	356.55			
0.64904630	9.1839	2480.91			

Table 3: Numerical results of various schemes for the estimation of $\mathbb{P}[m_T > B_1]$ under the CIR process. The parameter choices are $\kappa = 0.5$, $\alpha = 0.06$, $\sigma = 0.15$, $s_0 = 0.06$, $B_1 = 0.03$, and $T = 1$. We follow the procedure presented in Linetsky [27] to calculate the true value of the probability, which is given by 0.6484896. The ‘MULTILEVEL+BI’ part is skipped because this method seems not converge to the correct value.

Lemma A.1. For any $a > 0$, $y \in (0, a)$, and $\tau > 0$,

$G(a; x, y, \tau) =$

$$\begin{cases} 1 + \sum_{n=-\infty}^{\infty} \frac{e^{-2n^2 a^2 / \tau}}{1 - e^{-2xy / \tau}} \left(e^{\frac{2na(x-y)}{\tau}} - e^{-\frac{2na(x+y)}{\tau} - \frac{2xy}{\tau}} \right), & 0 < x < a; \\ 1 + \sum_{n=-\infty}^{\infty} \frac{e^{-2n^2 a^2 / \tau}}{2y / \sqrt{\tau}} \left(\frac{2na}{\sqrt{\tau}} e^{-\frac{2nay}{\tau}} + \frac{(2na+2y)}{\sqrt{\tau}} e^{-\frac{2nay}{\tau}} \right), & x = 0. \end{cases}$$

Proof. From the above definition of Brownian meanders and the scaling property of Brownian bridges, we can easily show that the scaling property also applies for the Brownian meanders; that is,

$$\begin{aligned} & (B_s^{me}, t \leq s \leq t + \tau | B_t^{me} = x, B_{t+\tau}^{me} = y) \\ & \stackrel{d}{=} \left(\sqrt{\tau} B_{(s-t)/\tau}^{me}, 0 \leq s-t \leq \tau | B_0^{me} = \frac{x}{\sqrt{\tau}}, B_1^{me} = \frac{y}{\sqrt{\tau}} \right) \end{aligned}$$

for all $\tau \geq 0$. From this, we can see that

$$G(a; x, y, \tau) = G(a/\sqrt{\tau}; x/\sqrt{\tau}, y/\sqrt{\tau}, 1).$$

Therefore, it suffices to calculate $G(a; x, y, 1)$.

First, we assume $x > 0, y \in (0, a)$. A key point of the proof is the following distributional equality implied by the definition of Brownian meanders:

$$\{B_t^{me}, 0 \leq t \leq 1\} \stackrel{d}{=} \{B_t, 0 \leq t \leq 1 | \min_{0 \leq t \leq 1} B_t > 0\},$$

where $\{B_t\}$ is a standard Brownian motion. Applying this equality and Bayes’ rule, we have

$$\begin{aligned} & \mathbb{P}[\max_{0 \leq t \leq 1} B_t^{me} < a | B_0^{me} = x, B_1^{me} = y] \\ & = \mathbb{P}[\max_{0 \leq t \leq 1} B_t < a | B_0 = x, B_1 = y, \min_{0 \leq t \leq 1} B_t > 0] \\ & = \frac{\mathbb{P}[\max_{0 \leq t \leq 1} B_t < a, \min_{0 \leq t \leq 1} B_t > 0 | B_0 = x, B_1 = y]}{\mathbb{P}[\min_{0 \leq t \leq 1} B_t > 0 | B_0 = x, B_1 = y]}. \end{aligned} \quad (13)$$

It is straightforward to verify that the denominator of the fraction on the right-hand side of (13) equals $1 - e^{-2xy}$. Let us focus on the numerator in the remainder of the proof. Denote T_z to be the first passage time of $\{B_t\}$ up to a level z . Using Bayes’ rule again to represent the numerator,

$$\begin{aligned} & \mathbb{P}[\max_{0 \leq t \leq 1} B_t < a, \min_{0 \leq t \leq 1} B_t > 0 | B_0 = x, B_1 = y] \\ & = \mathbb{P}[T_0 \wedge T_a > 1 | B_0 = x, B_1 = y] \\ & = \frac{\mathbb{P}[B_1 \in dy, T_0 \wedge T_a > 1 | B_0 = x]}{\mathbb{P}[B_1 \in dy | B_0 = x]}. \end{aligned} \quad (14)$$

Observe that

$$\mathbb{P}[B_1 \in dy | B_0 = x] = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-y)^2}{2}} \quad (15)$$

Table IV: Single Barrier, CIR Process

POISSON			POISSON+VR		
Estimator	RMSE($\times 10^{-4}$)	Time(s)	Estimator	RMSE($\times 10^{-4}$)	Time(s)
0.42401006	2.0739	224.14	0.42404864	1.1147	188.80
0.42403968	1.0984	896.55	0.42401198	0.5121	755.31
0.42399204	4.8528	3584.77	0.42400730	0.2130	4742.35
EULER			EULER+BI		
Estimator	RMSE($\times 10^{-4}$)	Time(s)	Estimator	RMSE($\times 10^{-4}$)	Time(s)
0.43239209	83.9981	72.16	0.42437003	11.7049	90.19
0.42995983	59.8195	571.83	0.42369934	6.1748	694.26
0.42815655	41.5452	4558.64	0.42381196	3.0290	5440.41
MULTILEVEL					
Estimator	RMSE($\times 10^{-4}$)	Time(s)			
0.42697681	42.3685	36.70			
0.42513243	18.8276	471.41			
0.42447096	8.8102	4359.42			

Table 4: Numerical results of various schemes for the estimation of $\mathbb{P}[M_T < B_2]$ under the CIR process. The parameter choices are $\kappa = 0.5$, $\alpha = 0.06$, $\sigma = 0.15$, $s_0 = 0.06$, $B_2 = 0.08$, and $T = 1$. The true value is 0.4240057 given by Linetsky [27]. We use $\mathbb{P}[M_T < B_2, m_T > \underline{b}]$ to approximate it and set \underline{b} to be 10^{-4} . The “MULTILEVEL+BI” part is skipped because this method seems not to converge to the correct value.

and

$$\begin{aligned} & \mathbb{P}[B_1 \in dy, T_0 \wedge T_a > 1 | B_0 = x] \\ &= \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} \left(e^{-\frac{(x-y-2na)^2}{2}} - e^{-\frac{(x+y+2na)^2}{2}} \right). \quad (16) \end{aligned}$$

Substituting (15) and (16) into (14) and (13), we can work out $G(a; x, y, 1)$ easily for the cases in which $x > 0$. As for the cases when $x = 0$, it follows when we take $x \rightarrow 0$ on both sides of (13). \square

References

- [1] Y.Ait-Sahalia, Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach, *Econometrica*. **70** (2002) 223–262.
- [2] L.Andersen, and R.Brotherton-Ratcliffe, Exact exotics, *Risk*. **9** (1996) 85–89.
- [3] S.Asmussen, and P.Glynn, *Stochastic Simulation: Algorithms and Analysis*, Springer-Verlag, New York, 2007.
- [4] S.Asmussen, P.Glynn, and J.Pitman, Discretization Error in Simulation of One-Dimensional Reflecting Brownian Motion, *Ann. Appl. Probab.* **5**(1995) 875–896.
- [5] D.R.Beaglehole, P.H.Dybvig, and G.Zhou, Going to extremes: correcting simulation bias in exotic option valuation, *Financial Analysts Journal*. **53**(1997) 62–68.
- [6] P.Baldi, Exact asymptotic for the probability of exit from a domain and applications to simulation, *Ann. Probab.* **23**(1995) 1644–1670.
- [7] P.Baldi, L.Caramellino, and M.G.Iovino, Pricing single and double barrier options via sharp large deviation techniques, *Math. Finance*. **9**(1999) 293–321.
- [8] A.Beskos, and G.O.Roberts, Exact simulation of diffusions, *Ann. Appl. Probab.* **15**(2005) 2422–2444.
- [9] A.Beskos, O.Papaspiliopoulos, and G.O.Roberts, Retrospective exact simulation of diffusion sample paths with applications, *Bernoulli*. **12**(2006) 1077–1098.
- [10] A.Beskos, O.Papaspiliopoulos, and G.O.Roberts, A factorisation of diffusion measure and finite sample path constructions, *Methodology and Computations in Applied Probability*. **10**(2008) 85–104.
- [11] N.Chen, Localization and exact simulation of stochastic differential equations driven by brownian motions. Working paper of the Chinese University of Hong Kong.(2010).
- [12] B. Casella and G.O. Roberts, Exact Simulation of Jump-Diffusion Processes with Monte Carlo Applications, *Methodology and Computing in Applied Probability*. **13**(2011) 449–473.
- [13] I.V.Denisov, A random walk and a Wiener process near a maximum, *Theory of Probability and its Applications*. **28**(1984) 821–824.
- [14] D.Duffie, and P.Glynn, Efficient Monte Carlo simulation of security prices, *Ann. Appl. Probab.* **5**(1995) 897–905.
- [15] D.Florens, Estimation of the diffusion coefficient from crossing, *Statistical Inference for Stochastic Processes*. **1**(1999) 175–195.
- [16] K.Giesecke, and D.Smelov, Exact sampling of jump-diffusions, Working Paper of Stanford University.(2010).
- [17] M.B.Giles, Multi-level Monte Carlo path simulation, *Operations Res.* **56**(2008) 607–617.
- [18] M.B.Giles, Improved multilevel Monte Carlo convergence

Table V: Single Barrier, sin Model

POISSON+VR			EXACT SIMULATION		
Estimator	SE($\times 10^{-5}$)	Time(s)	Estimator	SE($\times 10^{-5}$)	Time(s)
0.07620092	20.9054	5.607	0.07595906	19.8052	3521.695
0.07583637	9.8533	22.317	0.07576358	9.8927	13825.913
0.07593048	4.9784	87.466	0.07582331	4.9483	54833.95

Table 5: Comparison of two schemes for the sin model. The parameters used in the table are $s = 0.5$, $T = 2$, $K = 0.5$, and $B = 2$. We employ a large-scale Euler scheme to obtain a benchmark. The result from the Euler scheme has mean 0.07593238 and standard error 6.090×10^{-5} . The number of sample paths used in the Euler scheme is 1.6384×10^7 and the number of discretization steps in each path is 1.28×10^5 .

- using the Milstein scheme, in: A.Keller, S.Heinrich, and H. Niederreiter(Eds.), *Monte Carlo and Quasi-Monte Carlo Methods 2006*, Springer, New York,2008.
- [19] P.Glasserman, *Monte Carlo Methods in Financial Engineering*, Springer-Verlag, New York, 2004.
- [20] N.Ikeda, S.Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland mathematical library, vol 24, 2nd edn, North-Holland, Amsterdam, 1989.
- [21] J.P.Imhof, Density factorizations for Brownian motion, meander and the threedimensional Bessel process, and applications, *J. Appl. Probab.* **21**(1984) 500–510.
- [22] I.Karatzas, and S.E.Shreve, *Browinian Motion and Stochastic Calculus*, 2nd ed., Springer-Verlag, New York, 1991.
- [23] S. Karlin and H. M. Talor, *A Second Course in Stochastic Processes*. Academic Press, San Diego, 1981.
- [24] A.Kebaier, Statistical Romberg extrapolation: a new variance reduction method and applications to option pricing, *Ann. Appl. Probab.* **15**(2005) 2681–2705.
- [25] J.Keilson, and H.F.Ross, Passage time distribution for Gaussian Markov (Ornstein-Uhlenbeck) statistical process, in: *Selected Tables in Mathematical Statistics*, Vol. 3, American Mathematical Society, Rhode Island,1975.
- [26] P.E.Kloeden, and E.Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin,1992.
- [27] V. Linetsky, Computing hitting time densities for CIR and OU diffusions: applications to meanreverting models, *Journal of Computational Finance* **7**(2004) 1–22.
- [28] J.Pitman, Brownian motion, bridge, excursion, and meander characterized by sampling at independent uniform times, *Elec. J. Probab.* **4**(1999) 1–33.
- [29] J.Pitman, and M.Yor, Decomposition at the maximum for meanders and bridges of one-dimensional diffusions, in: *Ito's Stochastic Calculus and Probability Theory*. Springer Verlag, New York, Berlin, Heidelberg, 1996.
- [30] D.Revuz, and M.Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Springer-Verlag, Berlin,1999.
- [31] L.C.G.Rogers, and D.Williams, *Diffusions, Markov Processes, and Martingales*, Vol. 2, Wiley and Sons, New York,1987.
- [32] S.E.Shreve, *Stochastic Calculus for finance II*, Springer-Verlag, New York,2004.
- [33] O.A.Vasicek, An equilibrium characterization of the term structure, *J. Finan. Econ.* **5**(1977) 177–188.
- [34] D.Williams, Path decomposition and continuity of local time for one-dimensional diffusions I, *Proceedings of London Mathematical Society.* **28**(1974) 738–768.
- [35] C.Yi, On the first passage time distribution of an Ornstein-Uhlenbeck process, *Quant. Finance.* **10**(2010) 957–960.