Available online at www.sciencedirect.com
ScienceDirect
Journal of Complexity 24 (2008) 109 – 133
Journal of COMPLEXITY
www.elsevier.com/locate/jco

New Brownian bridge construction in quasi-Monte Carlo methods for computational finance

Junyi Lin*, Xiaoqun Wang

Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China

Received 4 December 2006; accepted 5 June 2007
Available online 13 July 2007

Abstract

Quasi-Monte Carlo (QMC) methods have been playing an important role for high-dimensional problems in computational finance. Several techniques, such as the Brownian bridge (BB) and the principal component analysis, are often used in QMC as possible ways to improve the performance of QMC. This paper proposes a new BB construction, which enjoys some interesting properties that appear useful in QMC methods. The basic idea is to choose the new step of a Brownian path in a certain criterion such that it maximizes the *variance explained* by the new variable while holding all previously chosen steps fixed. It turns out that using this new construction, the first few variables are more "important" (in the sense of explained variance) than those in the ordinary BB construction, while the cost of the generation is still linear in dimension. We present empirical studies of the proposed algorithm for pricing high-dimensional Asian options and American options, and demonstrate the usefulness of the new BB.

Keywords: Quasi-Monte Carlo methods; Brownian bridge; Effective dimension; Option pricing

1. Introduction

Monte Carlo (MC) methods are important tools for approximating high-dimensional integrals in computational finance [6]. If N is the number of function evaluations, then MC has an error of size $O(N^{-1/2})$, which is independent of the dimension. However, this convergence is very slow. Recently, alternative approaches, namely quasi-MC (QMC) methods, are widely used in pricing complex financial instruments. The basic idea of QMC is to use more uniformly distributed points

* Corresponding author.
E-mail addresses: linjy@mails.tsinghua.edu.cn, linjunyi82@gmail.com (J. Lin), xwang@math.tsinghua.edu.cn (X. Wang).

0885-064X/$ - see front matter
doi:10.1016/j.jco.2007.06.001

instead of random points. Let $I_d(f) = \int_{[0,1)^d} f(\mathbf{x})\,d\mathbf{x}$ and $Q_{N,d}(f) = \frac{1}{N}\sum_{i=1}^{N} f(\mathbf{x}_i)$ with the points $\mathbf{x}_1, \ldots, \mathbf{x}_N \in [0,1)^d$, the Koksma–Hlawka inequality yields that

$$|I_d(f) - Q_{N,d}(f)| \leqslant D_N^* V(f),$$

where D_N^* is the star discrepancy of the point set $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ and $V(f)$ is the variation of f in the sense of Hardy–Krause (see [15]).

Based on the Koksma–Hlawka inequality, the theoretical error bound of QMC is of size $O(N^{-1}\log^d N)$. This convergence is asymptotically better than $O(N^{-1/2})$ of MC. For high-dimensional problems, the advantage of QMC may be lost since the quantity $N^{-1}\log^d N$ is smaller than $N^{-1/2}$ only for extremely large N. However, many papers have shown great success of QMC for high-dimensional integrals arising in finance (see [3,16,20]).

One possible answer to this success is the concept of *effective dimensions* introduced in [4]. It is found that the effective dimensions of financial problems are generally small (see [4,25]). Another explanation for the success of QMC is given in [21], where it is shown that QMC may be especially suitable for functions in appropriate *weighted function spaces*. The superiority of QMC for some isotropic integrals is illustrated in [17,18].

QMC is applied in finance as a method to evaluate the expectation of a function of a random path generated by a stochastic process. For a continuous time process, this expectation is often expressed as a Feynman–Kac type integral over Brownian motion. In the numerical simulation, QMC (or MC) approximates the continuous process by means of discretizing it into small time steps and simulating the randomness with low discrepancy (or independent uniform random) points. The dimension can be large when the number of time steps is large. To speed up QMC, one may use variance reduction techniques borrowed from MC and dimension reduction techniques specially designed for QMC. Some strategies for dimension reduction are proposed, including the ordinary Brownian bridge (BB) (see [4,13,14]), the principal component analysis (PCA) (see [1]), the partial PCA (see [2]) and the linear transformation method (see [7]). These techniques may enjoy an advantage in a number of interesting cases. A special feature of most QMC point sets is their superior uniformity of the initial coordinates and of the low-dimensional projections (especially the one-dimensional projections). It is believed that these techniques may concentrate the variance of the function on the first few variables such that the better quality of the initial coordinates and the low-dimensional projections of QMC point sets can be fully used.

Each technique mentioned above has its own feature. The good performance of BB in pricing of financial derivatives is explained by the fact that BB uses the best coordinates of QMC points to determine most of the structure of a Brownian path. However, BB may be relatively less efficient compared with some other techniques. PCA outperforms BB in some examples, but the computational cost of PCA is larger than that of BB (see [1]). Linear transformation method takes into account the knowledge about the payoff functions, but it is required to solve complex optimization problems and is not easy to implement.

Motivated by the success of BB and PCA in some financial problems, in this paper we propose a new BB (NBB) construction, which enjoys some interesting properties and shares similar ideas to BB and PCA. Both BB and PCA can be considered as methods of variable transformation. It is desirable to find an optimal variable transformation such that it minimizes the actual error of QMC integration for a given QMC point set and a given class of functions, but this problem is complicated and we leave it as a future research topic. In this paper, we take a more convenient optimization criterion such that the approach picks the new step of a discrete time Brownian motion optimally (in the sense of *explained variance*) among all *permutation-based constructions*.

We recall that a permutation-based construction concerns generating the values of a Brownian motion according to a specified permutation of $(1, \ldots, d)^T$ and needs only $O(d)$ operations. Our new method is based on the evidence that in a permutation-based construction, the *variance explained* by the *i*th ($i = 1, 2, \ldots, d$) variable is totally determined by the first i components of the permutation (i.e., we need not know all d components of the permutation). It will be discussed in detail in Section 3. This feature enables us to find certain feasible generation order of a Brownian motion by a *step-by-step* procedure to maximize the variance explained by the new variable at each step. That is, we find the first step such that the variance explained by the first variable is maximized, then we fixed the first step, and find the next step, and so on. NBB is a permutation-based construction, therefore it can be implemented recursively and needs only $O(d)$ operations for each path.

This paper is organized as follows. In Section 2, we review various constructions of Brownian motions. In Section 3, we propose a step-by-step procedure to find NBB and illustrate its advantage over ordinary BB by comparing the cumulative explained variances. In Section 4, we apply NBB to problems of pricing Asian options and American–Burmuda–Asian options and show the usefulness of NBB. In Section 5, we make the conclusions and discuss some limitations of NBB.

2. Constructions of a discrete time Brownian motion

In many applications, it is required to generate Brownian motions (see Section 4 for examples). Assume that $\{B_t,\ 0 \leqslant t \leqslant T\}$ is a standard one-dimensional Brownian motion. We are interested in simulating the values of $B_{t_1}, \ldots, B_{t_d}$ at d discrete times, where $0 = t_0 < t_1 < \cdots < t_d = T$ and $B_{t_0} = 0$. For simplicity, we assume that $t_j - t_{j-1} = \Delta t = T/d,\ j = 1, 2, \ldots, d$, and denote $B_i = B_{t_i}$ in the following. Let $B = (B_1, \ldots, B_d)^T$, then $B \sim N(\mathbf{0}, \Sigma)$, where $N(\mathbf{0}, \Sigma)$ is a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix given by

$$\Sigma = (\min(t_i, t_j))_{i,j=1}^{d} = \begin{pmatrix} t_1 & t_1 & \ldots & t_1 \\ t_1 & t_2 & \ldots & t_2 \\ \vdots & \vdots & \ddots & \vdots \\ t_1 & t_2 & \ldots & t_d \end{pmatrix}.$$

According to the linear transformation property, if A is a $d \times d$ matrix and $Z = (Z_1, \ldots, Z_d)^T \sim N(\mathbf{0}, I_{d\times d})$, where $I_{d\times d}$ is the $d \times d$ identity matrix, then $AZ \sim N(\mathbf{0}, AA^T)$. Thus, the vector $B = (B_1, \ldots, B_d)^T$ can be generated as

$$\begin{pmatrix} B_1 \\ \vdots \\ B_d \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1d} \\ a_{21} & a_{22} & \ldots & a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \ldots & a_{dd} \end{pmatrix} \begin{pmatrix} Z_1 \\ \vdots \\ Z_d \end{pmatrix} = c_1 Z_1 + \cdots + c_d Z_d, \tag{1}$$

where $(a_{ij})_{d\times d} := A$ is a $d \times d$ matrix satisfying $AA^T = \Sigma$, and c_j is the jth column of A, and $(Z_1, \ldots, Z_d)^T \sim N(\mathbf{0}, I_{d\times d})$. For $1 \leqslant j \leqslant d$, we call $c_j = (a_{1j}, a_{2j}, \ldots, a_{dj})^T$ in the construction (1) the *coefficient vector* of $(B_1, B_2, \ldots, B_d)^T$ with respect to Z_j.

There are many decomposition matrices A satisfying $AA^T = \Sigma$. Several constructions of a discrete time Brownian motion are known in literature.

- The standard construction generates the Brownian motion sequentially by (given $B_0 = 0$)

$$B_k = B_{k-1} + \sqrt{\Delta t}\, Z_k, \quad k = 1, \ldots, d, \tag{2}$$

where Z_k are independent standard normal variables. The corresponding decomposition matrix is

$$A^{\mathrm{STD}} = \sqrt{\Delta t} \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{pmatrix},$$

which is the *Cholesky decomposition* of Σ.

- The ordinary BB (see [4,13,14]) first generates the final value B_d, then sample $B_{\lfloor d/2 \rfloor}$ conditional on the values of B_d and B_0, and proceed by progressively filling in intermediate values. Here, $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x. BB uses the first several coordinates of the low-discrepancy points to determine the general shape of the Brownian path, and the last few coordinates influence only the fine detail of the path. In particular, if d is a power of 2, then BB generates the Brownian motion as

$$\begin{aligned} B_d &= \sqrt{T} Z_1, \\ B_{d/2} &= \tfrac{1}{2}(B_0 + B_d) + \sqrt{T/4} Z_2, \\ B_{d/4} &= \tfrac{1}{2}(B_0 + B_{d/2}) + \sqrt{T/8} Z_3, \\ &\vdots \\ B_{d-1} &= \tfrac{1}{2}(B_{d-2} + B_d) + \sqrt{T/2d}\, Z_d, \end{aligned} \tag{3}$$

where Z_j are independent standard normal variables. BB construction corresponds to replacing the matrix A^{STD} in the standard construction by certain matrix A^{BB} satisfying $A^{\mathrm{BB}}(A^{\mathrm{BB}})^T = \Sigma$. For example, when $d = 8$,

$$A^{\mathrm{BB}} = \sqrt{T} \begin{pmatrix} 1/8 & 1/8 & \sqrt{2}/8 & 0 & 1/4 & 0 & 0 & 0 \\ 1/4 & 1/4 & \sqrt{2}/4 & 0 & 0 & 0 & 0 & 0 \\ 3/8 & 3/8 & \sqrt{2}/8 & 0 & 0 & 1/4 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5/8 & 3/8 & 0 & \sqrt{2}/8 & 0 & 0 & 1/4 & 0 \\ 3/4 & 1/4 & 0 & \sqrt{2}/4 & 0 & 0 & 0 & 0 \\ 7/8 & 1/8 & 0 & \sqrt{2}/8 & 0 & 0 & 0 & 1/4 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

- The PCA construction (see [1]) chooses the corresponding decomposition matrix A^{PCA} as

$$A^{\mathrm{PCA}} = (\sqrt{\lambda_1}\upsilon_1, \ldots, \sqrt{\lambda_d}\upsilon_d),$$

where $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$ are the eigenvalues of Σ in decreasing order, and $\upsilon_1, \ldots, \upsilon_d$ are the corresponding unit-length column eigenvectors. After obtaining the matrix A^{PCA}, PCA generates $(B_1, \ldots, B_d)^T$ by a full matrix product

$$(B_1, \ldots, B_d)^T = A^{\mathrm{PCA}} Z, \quad Z \sim N(\mathbf{0}, I_{d\times d}). \tag{4}$$

PCA maximizes the variance explained by $Z_1, \ldots, Z_k$ for all $k = 1, \ldots, n$ sequentially, and is optimal in the sense of *explained variance* (see next section).

Note that the values of $B_1, \ldots, B_d$ can be generated in any specified order, provided that at each step we generate the sample from the correct conditional distribution given the values already known. In other words, the values of B_i can be constructed in any permuted order $B_{\pi(1)}, \ldots, B_{\pi(d)}$, where $\Pi = \{\pi(1), \ldots, \pi(d)\}$ is an arbitrary permutation of $\{1, \ldots, d\}$. We call all such constructions the *permutation-based constructions*. The standard construction corresponds to $\pi(i) = i$, $i = 1, \ldots, d$. The BB construction (with d a power of 2) corresponds to $\pi(1) = d$, $\pi(2) = d/2$, $\pi(3) = d/4$, $\pi(4) = 3d/4$, …, $\pi(d) = d - 1$. Note that PCA does not belong to the class of permutation-based constructions because in PCA construction the values of B_i are not generated in any specified order.

3. An NBB construction

In the class of permutation-based constructions, it is natural to ask if we can find an "optimal" construction in some sense, or if we can find an "optimal" permutation Π^*. We need the following concept.

Definition 1. For a given i, $1 \leqslant i \leqslant d$, the *variance explained* by the ith variable in the construction (1) is the ratio $\|c_i\|^2 / \sum_{j=1}^d \|c_j\|^2$, where $\|\cdot\|$ is the L_2-norm of a vector.

For BB and PCA, although the total variance is the same as in the standard construction, the variance associated with each Z_i in construction (1) is different (see [1]). It has been rearranged such that more parts of the total variance are explained by the first few variables. Because the initial coordinates of most QMC points have superior uniformity compared with latter ones, this rearrangement would be useful in QMC.

Our purpose is to find an optimal permutation in the class of permutation-based constructions. We choose the explained variance as the optimization criterion, and this choice is mainly for two reasons: first, the criterion is the same to that of PCA; second, it will make the optimization problem easier to be solved. The chosen order should maximize the variance explained by the new variables. Totally, there are $d!$ permutations for $\{1, \ldots, d\}$ (or $d!$ different discretizations in the class of permutation-based constructions). Theorem 1 in Section 3.1 shows that c_1 is only determined by $\pi(1)$. The proofs of theorems in Section 3.2 conclude that for a given $i = 2, \ldots, d$, when the first $i - 1$ components of the permutation (i.e., $\pi(1), \ldots, \pi(i-1)$) are fixed, c_i in construction (1) is totally determined by $\pi(i)$. Therefore, we can write the squared norm of the ith column of the decomposition matrix A as $\|c_i\|^2_{\pi(i)}$.

Note that

$$\|A\|_F := \|c_1\|^2_{\pi(1)} + \cdots + \|c_d\|^2_{\pi(d)} = trace(AA^T) = trace(\Sigma) = \frac{T(1+d)}{2},$$

which is a constant when d is fixed. Thus in the first step we find $\pi^*(1) \in \{1, \ldots, d\}$ such that

$$\|c_1\|^2_{\pi^*(1)} = \max_{\pi(1)\in\{1,\ldots,d\}} \|c_1\|^2_{\pi(1)}. \tag{5}$$

When $\pi^*(1)$ is determined, we find $\pi^*(2) \in \{1, \ldots, d\}\backslash\{\pi^*(1)\}$ such that

$$\|c_2\|^2_{\pi^*(2)} = \max_{\pi(2)\in\{1,\ldots,d\}\backslash\{\pi^*(1)\}} \|c_2\|^2_{\pi(2)}$$

$$\vdots \tag{6}$$

We continue this procedure until all the components of the permutation $\Pi^* = \{\pi^*(1), \ldots, \pi^*(d)\}$ are determined. Once the optimal permutation is determined, we can generate $B_1, \ldots, B_d$ recursively by $O(d)$ operations.

Note that the optimization problems do not intend to minimize the integration error for a given class of functions, and it does not take into account the knowledge of the integrands. We will return to this in conclusion section.

The following conditional formula of multivariate normal distribution is important in our analysis.

Lemma 1 (*Conditional formula [6] Glasserman*). *Suppose the partitioned vector* $(B_{[1]}, B_{[2]})$ (*where each* $B_{[i]}$ *may itself be a vector*) *is multivariate normal with*

$$\begin{pmatrix} B_{[1]} \\ B_{[2]} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{[1]} \\ \mu_{[2]} \end{pmatrix}, \begin{pmatrix} \Sigma_{[11]} & \Sigma_{[12]} \\ \Sigma_{[21]} & \Sigma_{[22]} \end{pmatrix}\right)$$

and suppose $\Sigma_{[22]}$ *has full rank. Then*

$$(B_{[1]}|B_{[2]} = b) \sim N(\mu_{[1]} + \Sigma_{[12]}\Sigma^{-1}_{[22]}(b - \mu_{[2]}), \Sigma_{[11]} - \Sigma_{[12]}\Sigma^{-1}_{[22]}\Sigma_{[21]}). \tag{7}$$

3.1. Determination of the optimal first step $\pi^*(1)$

Given $B_0 = 0$, the standard construction generates B_1 as its first step, while BB generates B_d as its first step. In this subsection, we want to decide which point among $B_1, \ldots, B_d$ deserves to be generated first. That is, we study how to find the optimal first component $\pi^*(1)$ of the permutation Π^* such that the variance explained by the first variable Z_1 in construction (1) is maximized. For convenience, we assume $T = 1$ in the following part of this section.

Theorem 1. *In a permutation-based construction of the Brownian motion* $B_1, \ldots, B_d$, *the optimal first step* $\pi^*(1)$ *is the integer nearest to* $(6d + 3)/8$.

In particular, if $d = 2^n$ $(n \geqslant 2)$, *then the optimal first step is* $\pi^*(1) = 3d/4$, *and the corresponding squared norm of the first column of the decomposition matrix A is* $(\frac{9}{4}d^2 + \frac{9}{4}d + 1)/(6d)$.

Proof. Assume that the first generated point is B_i, i.e. $\pi(1) = i$, for some i with $1 \leqslant i \leqslant d$. Then we have

$$B_i = B_0 + \sqrt{t_i - t_0}Z_1 = \sqrt{t_i}Z_1, \quad Z_1 \sim N(0, 1).$$

Given the value of $B_i = \sqrt{t_i}\, Z_1$, according to Lemma 1, the conditional expectation of $(B_1, \ldots, B_d)^T$ is

$$E((B_1, \ldots, B_{i-1}, B_i, \ldots, B_d)^T | B_i) = 0 + t_i^{-1} \begin{pmatrix} t_1 \\ \vdots \\ t_{i-1} \\ t_i \\ \vdots \\ t_i \end{pmatrix} B_i = \frac{1}{\sqrt{t_i}} \begin{pmatrix} t_1 \\ \vdots \\ t_{i-1} \\ t_i \\ \vdots \\ t_i \end{pmatrix} Z_1. \tag{8}$$

On the other hand, we can write $(B_1, \ldots, B_d)^T$ in the form as in construction (1). Since $(Z_1, \ldots, Z_d)^T \sim N(\mathbf{0}, I_{d\times d})$ and $B_i = \sqrt{t_i}\, Z_1 = a_{i,1} Z_1$, we obtain the conditional expectation in terms of the elements of the decomposition matrix A

$$E((B_1, \ldots, B_{i-1}, B_i, \ldots, B_d)^T | B_i) = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{i-1,1} \\ a_{i,1} \\ \vdots \\ a_{d,1} \end{pmatrix} Z_1. \tag{9}$$

Comparing the coefficients in (8) and (9) leads to

$$\begin{pmatrix} a_{1,1} \\ \vdots \\ a_{i-1,1} \\ a_{i,1} \\ \vdots \\ a_{d,1} \end{pmatrix} = \frac{1}{\sqrt{t_i}} \begin{pmatrix} t_1 \\ \vdots \\ t_{i-1} \\ t_i \\ \vdots \\ t_i \end{pmatrix}. \tag{10}$$

Therefore, when $\pi(1) = i$, the first column of the decomposition matrix A is

$$\frac{1}{\sqrt{t_i}} \begin{pmatrix} t_1 \\ \vdots \\ t_{i-1} \\ t_i \\ \vdots \\ t_i \end{pmatrix}. \tag{11}$$

Note that $t_j = j\Delta t$, so the squared norm of the first column is

$$\begin{aligned}\|c_1\|^2_{\pi(1)} &= \frac{\Delta t}{i}\left[\sum_{j=1}^{i} j^2 + (d-i)i^2\right]\\ &= \frac{\Delta t}{6}[(i+1)(2i+1) + 6i(d-i)]\\ &= \frac{1}{6d}[-4i^2 + (6d+3)i + 1], \end{aligned} \tag{12}$$

which achieves its maximum value when $i = (6d+3)/8$. Since i should be an integer and the quantity in (12) is quadratic, the optimal first step $\pi^*(1)$ is the integer nearest to $(6d+3)/8$.

In particular, if $d = 2^n$ $(n \geqslant 2)$, then $\pi^*(1) = 3d/4$. Putting $i = 3d/4$ into (12), we get

$$\|c_1\|^2_{\pi^*(1)} = \frac{1}{6d}\left(\frac{9}{4}d^2 + \frac{9}{4}d + 1\right). \qquad \square$$

The variances explained by the first variable in NBB, BB and the standard constructions of the Brownian motion are different. Assume that $d = 2^n$ $(n \geqslant 2)$. We write c_1^{NBB}, c_1^{BB} and c_1^{STD} as the corresponding first columns of the decomposition matrices for NBB, BB and the standard constructions. We compare the effects of $\pi(1)$ in different constructions.

- NBB generates $B_{3d/4}$ in the first step. Putting $i = 3d/4$ into (11) and (12), we have

$$c_1^{\text{NBB}} = \sqrt{\frac{4}{3}}\begin{pmatrix} 1/d \\ 2/d \\ \vdots \\ \left(\frac{3}{4}d - 1\right)\Big/ d \\ 3/4 \\ 3/4 \\ \vdots \\ 3/4 \end{pmatrix},$$

  and the squared norm of c_1^{NBB} is

$$\|c_1^{\text{NBB}}\|^2 = \frac{1}{6d}\left(\frac{9}{4}d^2 + \frac{9}{4}d + 1\right).$$

- The ordinary BB generates B_d in the first step. Putting $i = d$ into (11) and (12), we have

$$c_1^{\text{BB}} = \begin{pmatrix} 1/d \\ 2/d \\ \vdots \\ (d-1)/d \\ 1 \end{pmatrix},$$

and the squared norm of c_1^{BB} is

$$\|c_1^{\mathrm{BB}}\|^2 = \frac{1}{6d}(2d^2 + 3d + 1).$$

- The standard construction generates B_1 in the first step. Putting $i = 1$ into (11) and (12), we have

$$c_1^{\mathrm{STD}} = \begin{pmatrix} 1/\sqrt{d} \\ 1/\sqrt{d} \\ \vdots \\ 1/\sqrt{d} \end{pmatrix},$$

and the squared norm of c_1^{STD} is

$$\|c_1^{\mathrm{STD}}\|^2 = 1.$$

The limiting explained variances of the first variable in different constructions are (as d trends to infinity)

$$\lim_{d\to\infty} \frac{\|c_1^{\mathrm{NBB}}\|^2}{\|A\|_F^2} = \lim_{d\to\infty} \frac{\frac{9}{24}d + \frac{9}{24} + \frac{1}{6d}}{\frac{d+1}{2}} = \frac{3}{4}, \tag{13}$$

$$\lim_{d\to\infty} \frac{\|c_1^{\mathrm{BB}}\|^2}{\|A\|_F^2} = \lim_{d\to\infty} \frac{\frac{1}{3}d + \frac{1}{2} + \frac{1}{6d}}{\frac{d+1}{2}} = \frac{2}{3}, \tag{14}$$

$$\lim_{d\to\infty} \frac{\|c_1^{\mathrm{STD}}\|^2}{\|A\|_F^2} = \lim_{d\to\infty} \frac{1}{\frac{d+1}{2}} = 0. \tag{15}$$

Comparing (13) and (14), we conclude that

$$\lim_{d\to\infty} \frac{\|c_1^{\mathrm{NBB}}\|^2 - \|c_1^{\mathrm{BB}}\|^2}{\|A\|_F^2} = \frac{1}{12}. \tag{16}$$

We see that in NBB, the first variable explains more variance than it does in BB as $d \to \infty$. Summarizing the results above, we have the following.

Corollary 1. *For* NBB, BB *and the standard constructions of the Brownian motion*, *the limiting variances explained by the first variable are given by* (13)–(15), *respectively*, *as* $d \to \infty$.

3.2. The generations of the subsequent points

After $\pi^*(1)$ has been determined and the point $B_{\pi^*(1)}$ has been generated, we need to determine the values of $\pi^*(2), \pi^*(3), \ldots, \pi^*(d)$ and to generate $B_{\pi^*(2)}, B_{\pi^*(3)}, \ldots, B_{\pi^*(d)}$. Take the determination of $\pi^*(2)$ for example. We want to choose the value of $\pi^*(2)$ such that it maximizes the variance explained by the second variable Z_2 in construction (1). We face two types of situations, that is, we may generate $B_{\pi(2)}$ just conditional on the past value $B_{\pi^*(1)}$ (in this case, $\pi(2) \in$

$\{\pi^*(1)+1,\ldots,d\}$), or we may generate $B_{\pi(2)}$ conditional on the past and the future values $B_0, B_{\pi^*(1)}$ (in this case, $\pi(2)\in\{1,\ldots,\pi^*(1)-1\}$). In the first situation, we need to know among $B_{\pi^*(1)+1},\ldots,B_d$, which step is the " local optimal" conditional on the value of $B_{\pi^*(1)}$. In the second situation, we also need to know among $B_1,\ldots,B_{\pi^*(1)-1}$, which step is the " local optimal" conditional on the values of B_0 and $B_{\pi^*(1)}$. When the local optimal steps in both situations are determined, we choose the one with the larger squared norm as the value of $\pi^*(2)$. This procedure maximizes the variance explained by the variable Z_2.

Similarly, in the generations of the subsequent points we always face two kinds of situations:

Situation 1. *The past value B_q (for some q with $0\leqslant q\leqslant d-1$) has been sampled, while $B_{q+1},\ldots,B_d$ are unknown.*

Situation 2. *The past value B_{q_1} and the future value B_{q_2} (for some q_1, q_2 with $0\leqslant q_1 < q_2\leqslant d$) have been sampled, while $B_{q_1+1},\ldots,B_{q_2-1}$ are unknown.*

In each situation, we need to find the local optimal step in the sense of explained variance. Note that the case of Theorem 1 belongs to Situation 1 since we only know the past value B_0. In Theorem 2, we study the local optimal step in Situation 1 in a general case.

Theorem 2. *In a permutation-based construction of the Brownian motion $B_1,\ldots,B_d$, assume that B_q has been sampled (for some q with $0\leqslant q\leqslant d-1$), while $B_{q+1},\ldots,B_d$ have not yet been generated. Among $B_{q+1},\ldots,B_d$ (conditional on the past value B_q), the local optimal new step q^* is the integer nearest to $(6d+2q+3)/8$.*

Proof. Assume that the desirable optimal step is i with $i\in\{q+1,\ldots,d\}$, conditional only on the past value B_q. Then we generate B_i as

$$B_i = B_q + \sqrt{t_i - t_q}\,z, \quad z\sim N(0,1). \tag{17}$$

Given the values of B_q and B_i, according to Lemma 1, the conditional expectation of $(B_1,\ldots,B_d)^T$ is

$$E((B_1,\ldots,B_q,B_{q+1},\ldots,B_i,B_{i+1},\ldots,B_d)^T|(B_q,B_i)^T)$$

$$= 0 + \begin{pmatrix} t_1 & t_1 \\ \vdots & \vdots \\ t_q & t_q \\ t_q & t_{q+1} \\ \vdots & \vdots \\ t_q & t_{i-1} \\ t_q & t_i \\ \vdots & \vdots \\ t_q & t_i \end{pmatrix} \begin{pmatrix} t_q & t_q \\ t_q & t_i \end{pmatrix}^{-1} \begin{pmatrix} B_q \\ B_i \end{pmatrix}$$

$$
= \frac{1}{t_i - t_q} \begin{pmatrix} (t_i - t_1)B_q \\ \vdots \\ (t_i - t_q)B_q \\ (t_i - t_{q+1})B_q + (t_{q+1} - t_q)B_i \\ \vdots \\ (t_i - t_{i-1})B_q + (t_{i-1} - t_q)B_i \\ (t_i - t_q)B_i \\ \vdots \\ (t_i - t_q)B_i \end{pmatrix}
$$

$$
= \begin{pmatrix} (t_i - t_1)/(t_i - t_q) \\ \vdots \\ (t_i - t_{q-1})/(t_i - t_q) \\ 1 \\ \vdots \\ 1 \end{pmatrix} B_q + \frac{1}{\sqrt{t_i - t_q}} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ t_{q+1} - t_q \\ \vdots \\ t_{i-1} - t_q \\ t_i - t_q \\ \vdots \\ t_i - t_q \end{pmatrix} z.
$$

Similarly as in the proof of Theorem 1, we can conclude that the coefficient vector of $(B_1, \dots, B_d)^T$ in construction (1) with respect to z is

$$
\frac{1}{\sqrt{t_i - t_q}} (\underbrace{0, \dots, 0}_{q}, t_{q+1} - t_q, \dots, t_{i-1} - t_q, \underbrace{t_i - t_q, \dots, t_i - t_q}_{d-i+1})^T. \tag{18}
$$

Note that $t_j = j\Delta t$, so the corresponding column of the decomposition matrix A in the construction (1) is

$$
\sqrt{\Delta t/(i-q)} (\underbrace{0, \dots, 0}_{q}, 1, 2, \dots, i-1-q, \underbrace{i-q, \dots, i-q}_{d-i+1})^T. \tag{19}
$$

Then the corresponding squared norm is

$$
\begin{aligned}
&\frac{\Delta t}{i-q} \left[\sum_{j=1}^{i-q} j^2 + (d-i)(i-q)^2 \right] \\
&\quad = \frac{\Delta t}{6} [(i-q+1)(2i-2q+1) + 6(d-i)(i-q)] \\
&\quad = \frac{1}{6d} [-4i^2 + (6d+2q+3)i + 2q^2 - (3+6d)q + 1],
\end{aligned} \tag{20}
$$

which achieves its maximum value when $i = (6d + 2q + 3)/8$. Since $i \in \{q + 1, \ldots, d\}$ should be an integer and the quantity in (20) is quadratic, the local optimal step q^*, conditional on the past value B_q, is the integer nearest to $(6d + 2q + 3)/8$. □

Remark 1. In Theorem 2, since $i \in \{q + 1, \ldots, d\}$, the quantity in (20) has its *minimum* value when $i = q + 1$, which corresponds to the standard construction of a Brownian motion. This means that the standard construction may be the least efficient one (in the sense of explained variance) in the class of permutation-based constructions.

Theorem 1 is a special case of Theorem 2 with $q = 0$. In the case of $d = 2^n$, according to Theorem 1, it is optimal to generate $B_{3d/4}$ first. When $B_{3d/4}$ is sampled, while $B_{3d/4+1}, \ldots, B_d$ have not yet been generated (this corresponds to $q = 3d/4$ in Theorem 2), there are $d' = d/4$ points after $B_{3d/4}$ that are unknown. According to Theorem 2, the local optimal new step among $B_{3d/4+1}, \ldots, B_d$ conditional on the past value $B_{3d/4}$ is $q^* = 15d/16 = q+3d'/4$, or equivalently $q^* - q = 3d'/4$. By deduction, if q has the form $\frac{4^{j-1}-1}{4^{j-1}}d$ for some j $(j = 1, \ldots, \lfloor \log_4 2^n \rfloor)$, then the local optimal new step among $B_{q+1}, \ldots, B_d$ conditional on the past value B_q is $q^* = \dfrac{4^j - 1}{4^j} d = q + 3(d - q)/4$, or equivalently $q^* - q = \dfrac{3}{4}(d - q)$.

So when the past value $B_{3d/4}$ is sampled, the local optimal new step among $B_{3d/4+1}, \ldots, B_d$ is

$$\tfrac{3}{4}d + \tfrac{3}{4}(d - \tfrac{3}{4}d) = \tfrac{15}{16}d.$$

When the past value $B_{15d/16}$ is sampled, the local optimal new step among $B_{15d/16+1}, \ldots, B_d$ is

$$\tfrac{15}{16}d + \tfrac{3}{4}(d - \tfrac{15}{16}d) = \tfrac{63}{64}d,$$

and so on. Summarizing these, we have the following principle which indicates how to choose the local optimal step in Situation 1 when $d = 2^n$.

Corollary 2 ($\frac{3}{4}$-*Rule*). *For the case in Theorem* 2, *assume that* $d = 2^n$ *and* q *has the form* $\frac{4^{j-1}-1}{4^{j-1}}d$ *for some* j $(j = 1, \ldots, \lfloor \log_4 2^n \rfloor)$. *Then among* $B_{q+1}, \ldots, B_d$ (*conditional on the past value* B_q), *the local optimal new step is* $q^* = \frac{4^j-1}{4^j}d = q + 3(d - q)/4$, *and the squared norm of the corresponding column of the decomposition matrix* A *is* $(\frac{36}{4^{2j}}d^2 + \frac{9}{4^j}d + 1)/(6d)$.

Now we turn to Situation 2. We study which step is the local optimal if we know both the past and the future values.

Theorem 3. *In a permutation-based construction of the Brownian motion* $B_1, \ldots, B_d$, *assume that* B_{q_1} *and* B_{q_2} *have been sampled (for some* q_1, q_2 *with* $0 \leqslant q_1 < q_2 \leqslant d$), *while* $B_{q_1+1}, \ldots, B_{q_2-1}$ *have not yet been generated. Among* $B_{q_1+1}, \ldots, B_{q_2-1}$ (*conditional on the past and the future values* B_{q_1}, B_{q_2}), *the optimal new step is the integer nearest to* $(q_1 + q_2)/2$.

In particular, if $(q_1 + q_2)/2$ *is an integer, then the local optimal new step is* $(q_1 + q_2)/2$, *and the squared norm of the corresponding column of the decomposition matrix* A *is* $[(q_2 - q_1)^2/2 + 1]/(6d)$. *If* $(q_1 + q_2)/2$ *is not an integer, then the local optimal new step is* $(q_1 + q_2 - 1)/2$ *or* $(q_1 + q_2 + 1)/2$, *and the corresponding squared norm is* $[(q_2 - q_1)^2/2 + 1/2]/(6d)$.

Proof. Assume that the desired optimal step is i with $i \in \{q_1+1, \ldots, q_2-1\}$, conditional on the past and the future values B_{q_1}, B_{q_2}. According to Lemma 1, we have

$$B_i = \frac{q_2 - i}{q_2 - q_1} B_{q_1} + \frac{i - q_1}{q_2 - q_1} B_{q_2} + \sqrt{\frac{(t_{q_2} - t_i)(t_i - t_{q_1})}{t_{q_2} - t_{q_1}}} z, \quad z \sim N(0, 1). \tag{21}$$

When B_i is fixed, using a similar approach as in the proof of Theorem 2, we can conclude that the coefficient vector of $(B_1, \ldots, B_d)^T$ in construction (1) with respect to z is

$$\begin{aligned}&(\underbrace{0, \ldots, 0}_{q_1}, \zeta_1(t_{q_1+1} - t_{q_1}), \zeta_1(t_{q_1+2} - t_{q_1}), \ldots, \zeta_1(t_i - t_{q_1}),\\ &\zeta_2(t_{q_2-1} - t_i), \ldots, \zeta_2(t_{i+1} - t_i), \underbrace{0, \ldots, 0}_{d-q_2+1})^T,\end{aligned} \tag{22}$$

where $\zeta_1 = \frac{1}{t_i - t_{q_1}} \sqrt{\frac{(t_{q_2}-t_i)(t_i-t_{q_1})}{t_{q_2}-t_{q_1}}}$ and $\zeta_2 = \frac{1}{t_{q_2}-t_i} \sqrt{\frac{(t_{q_2}-t_i)(t_i-t_{q_1})}{t_{q_2}-t_{q_1}}}$.

Note that $t_j = j\Delta t$, so the corresponding column of the decomposition matrix A in construction (1) is

$$(\underbrace{0, \ldots, 0}_{q_1}, \xi_1, 2\xi_1, \ldots, (i - q_1)\xi_1, (q_2 - i - 1)\xi_2, \ldots, 2\xi_2, \xi_2, \underbrace{0, \ldots, 0}_{d-q_2+1})^T, \tag{23}$$

where $\xi_1 = \sqrt{\frac{(q_2-i)\Delta t}{(i-q_1)(q_2-q_1)}}$ and $\xi_2 = \sqrt{\frac{(i-q_1)\Delta t}{(q_2-i)(q_2-q_1)}}$. Then the corresponding squared norm is

$$\begin{aligned}\xi_1^2 \sum_{j=1}^{i-q_1} j^2 + \xi_2^2 \sum_{j=1}^{q_2-1-i} j^2 &= \frac{\Delta t}{6(q_2 - q_1)}[(q_2 - i)(i - q_1 + 1)(2i - 2q_1 + 1)\\ &\quad +(i - q_1)(q_2 - i - 1)(2q_2 - 2i - 1)]\\ &= \frac{1}{6d}[-2i^2 + 2(q_2 + q_1)i + (1 - 2q_1q_2)],\end{aligned} \tag{24}$$

which achieves its maximum value when $i = (q_1+q_2)/2$. Since $i \in \{q_1+1, \ldots, q_2-1\}$ should be an integer and the quantity in (24) is quadratic, the local optimal new step, conditional on the past and the future values B_{q_1}, B_{q_2}, is the integer nearest to $(q_1 + q_2)/2$.

In particular, if $(q_1+q_2)/2$ is an integer, then the local optimal new step is $(q_1+q_2)/2$. Putting $i = (q_1+q_2)/2$ into (24), we get that the corresponding squared norm is $[(q_2-q_1)^2/2+1]/(6d)$. If $(q_1+q_2)/2$ is not an integer, then the local optimal new step is $(q_1+q_2-1)/2$ or $(q_1+q_2+1)/2$, and we get that the corresponding squared norm is $[(q_2 - q_1)^2/2 + 1/2]/(6d)$. □

Remark 2. According to the proof of Theorem 3, if the local optimal new step is $(q_1 + q_2)/2$, the squared norm of the corresponding column only depends on the value of $q_2 - q_1$. If the local optimal new step is $(q_1 + q_2 - 1)/2$ or $(q_1 + q_2 + 1)/2$, then the squared norm of the column is $[(q_1 - q_2)^2/2 + 1/2]/(6d)$, which only depends on the value of $q_2 - q_1$. Furthermore, if we have another pair of $\widehat{q}_1, \widehat{q}_2$ as prescribed in Theorem 3 satisfying $\widehat{q}_2 - \widehat{q}_1 = (q_2-q_1)+1$, where $(\widehat{q}_1 + \widehat{q}_2)/2$ is not an integer while $(q_1 + q_2)/2$ is an integer, we have

$$\frac{1}{6d}\left[\frac{1}{2}(\widehat{q}_2 - \widehat{q}_1)^2 + \frac{1}{2}\right] > \frac{1}{6d}\left[\frac{1}{2}(q_1 - q_2)^2 + 1\right]. \tag{25}$$

Therefore, among all possible pairs of q_1, q_2 prescribed in Situation 2, the local optimal new step in the largest time interval $[t_{q_1}, t_{q_2}]$ (with the largest value of $q_2 - q_1$) has the priority.

Remark 3. Theorem 3 and Remark 2 explain ordinary BB from a perspective of optimization problems. Actually, when B_0 is given and the end point B_d is already generated, subsequent values of ordinary BB are set at the successive midpoint, i.e., $B_{d/2}, B_{d/4}, B_{3d/4}, \ldots$. Since always at least one past value and one future value exist after B_d is generated, this kind of selection always maximizes the variance explained by the new variable according to Theorem 3 and Remark 2. We point out that in ordinary BB, only the choice of B_d as the first step does not maximize the variance explained by Z_1 (see Section 3.1).

Remark 4. Theorem 3 and Remark 2 may explain why ordinary BB usually performs better than the " golden section BB construction" in [8], which is built by using the golden section sequence for running through the time interval after B_d is generated as the first step.

Remark 5. The proofs of Theorems 2 and 3 imply that in a permutation-based construction, the variance explained by the ith ($i = 1, 2, \ldots, d$) variable is totally determined by the first i components of the permutation (i.e., we need not know all d components of the permutation). This fact enables us to determine the next step optimally by the information up to now.

3.3. Optimal generation order of the Brownian motion

Theorems 2 and 3 indicate how to choose the local optimal steps in Situations 1 and 2, respectively. In this subsection, we focus on finding the optimal generation order of the Brownian motion when $d = 2^n$ $(n \geqslant 2)$. The principle is that the new step should maximize the variance explained by the new variable.

We still take the determination of $\pi^*(2)$ for example. We first assume that $B_{3d/4}$ is generated as suggested in Theorem 1. If we generate the next new step conditional only on the past value $B_{3d/4}$, according to Corollary 2 (with $q = 3d/4$), the local optimal new step among $B_{3d/4+1}, \ldots, B_d$ would be the step $15d/16$. The squared norm of the corresponding column for this step is $(\frac{9}{64}d^2 + \frac{9}{16}d + 1)/(6d)$. Alternatively, if we generate the next new step conditional on the past and the future values $B_0, B_{3d/4}$, according to Theorem 3 (with $q_1 = 0$ and $q_2 = 3d/4$), the local optimal new step among $B_1, \ldots, B_{3d/4-1}$ would be the step $3d/8$. The squared norm of the corresponding column for this step is $(\frac{9}{32}d^2 + 1)/(6d)$. Now we determine the value of $\pi^*(2)$ between these two local optimal choices. Since $(9d^2)/32 > (9d^2)/64 + (9d)/16$, we choose the value of $\pi^*(2)$ to be $3d/8$, i.e., we generate $B_{3d/8}$ as the second step. Sequent steps are determined in a similar way.

When $d = 2^n$, we can find a more convenient rule to determine the superiority of these local optimal choices. We denote

$$\rho_j = \frac{1}{6d}\left(\frac{36}{4^{2j}}d^2 + \frac{9}{4^j}d + 1\right), \tag{26}$$

$$R_{j,1} = \frac{1}{6d}\left(\frac{18}{4^{2j}}d^2 + 1\right), \tag{27}$$

$$R_{j,2} = \frac{1}{6d}\left(\frac{72}{4^{2j}}d^2 + 1\right), \tag{28}$$

then for a fixed integer j ($j = 1, \ldots, \lfloor \log_4 d \rfloor$), we have $R_{j,1} < \rho_j < R_{j,2}$. The meanings of these quantities will be clear soon. As we discussed above, we face two types of situations. Situation 1 corresponds to the case where the past value B_q has been sampled, where q has the form $\frac{4^{j-1}-1}{4^{j-1}}d$ for some j ($j = 2, \ldots, \lfloor \log_4 d \rfloor$), while $B_{q+1}, \ldots, B_d$ are unknown; Situation 2 corresponds to the case where the past value B_{q_1} and the future value B_{q_2} have been sampled, while $B_{q_1+1}, \ldots, B_{q_2-1}$ are unknown. The generations of all values of $B_1, \ldots, B_d$ consist of two stages.

Stage 1: In Stage 1, at least one pair of $q_1,\ q_2$ prescribed in Situation 2 satisfies that $(q_1+q_2)/2$ is an integer.

We first generate $B_{3d/4}$ as suggested in Theorem 1. In the generations of the subsequent points, we determine the priority of several local optimal choices as follows by comparing the length of time intervals.

For Situation 1, according to Corollary 2, the local optimal new step among $B_{q+1}, \ldots, B_d$ would be $q^* = \dfrac{4^j-1}{4^j}d$ for some j ($j = 2, \ldots, \lfloor \log_4 d \rfloor$), and the squared norm of the corresponding column for the point B_{q^*} is ρ_j. The time interval $[t_q,\ t_{q^*}]$ has a length of $3/4^j$.

For Situation 2, according to Theorem 3, the local optimal new step among $B_{q_1+1}, \ldots, B_{q_2-1}$ would be $(q_1+q_2)/2$. If the length of time interval $[t_{q_1}, t_{(q_1+q_2)/2}]$ is $t_{(q_1+q_2)/2} - t_{q_1} = 3/4^j$ (which is the same length as for Situation 1), then the squared norm of the corresponding column for the point $B_{(q_1+q_2)/2}$ is $R_{j,1}$, which is smaller than ρ_j. Thus in this case, we would choose q^* in Situation 1 as the next new step of the Brownian motion. While, if $t_{(q_1+q_2)/2} - t_{q_1} = 2 \times 3/4^j$ (which is twice the length as for Situation 1), then the squared norm of the corresponding column for the point $B_{(q_1+q_2)/2}$ is $R_{j,2}$, which is larger than ρ_j. Thus in this case, we would choose $(q_1+q_2)/2$ in Situation 2 as the next new step of the Brownian motion.

In other words, if the length of the time interval $[t_q, t_{q^*}]$ in Situation 1 is the same as the length of the time interval $[t_{q_1}, t_{(q_1+q_2)/2}]$ in Situation 2, then the step q^* in Situation 1 has the priority. While if the length of the time interval $[t_q,\ t_{q^*}]$ in Situation 1 is half of the length of the time intervals $[t_{q_1}, t_{(q_1+q_2)/2}]$ in Situation 2, then the step $(q_1+q_2)/2$ in Situation 2 has the priority.

For example, after $B_{3d/4}$ is generated as the first step, the length of the time interval $[t_{3d/4},\ t_{15d/16}]$ in Situation 1 is half of the length of the time interval $[t_0,\ t_{3d/8}]$ in Situation 2, so $B_{3d/8}$ has the priority to $B_{15d/16}$. Next, the length of the time interval $[t_{3d/4},\ t_{15d/16}]$ in Situation 1 is the same to the length of the time intervals $[t_0,\ t_{3d/16}]$ and $[t_{3d/8},\ t_{9d/16}]$ in Situation 2, so $B_{15d/16}$ is generated as the third step.

Generally, after generating $B_{3d/4}$, we sample the subsequent points in the middle of one of the largest time intervals, whose two endpoints have been sampled, until the length of the largest time intervals are reduced into $2 \times 3/4^j$ for some j ($j = 2, \ldots, \lfloor \log_4 d \rfloor$). Then we generate the point $B_{\frac{4^j-1}{4^j}d}$ and continue the division process above.

Stage 2: Stage 1 will continue until for all pairs of $q_1,\ q_2$ prescribed in Situation 2, $(q_1+q_2)/2$ is not an integer. It is easy to see that when this comes true, the remaining points that need to be generated are

$$B_i, \quad 1 \leqslant i \leqslant d \text{ with } i \text{ not being a multiple of 3,}$$

i.e., the following points that are not in the parentheses:

$$(B_0), B_1, B_2, (B_3), B_4, B_5, (B_6), \ldots, (B_{d-4}), B_{d-3}, B_{d-2}, (B_{d-1}), B_d, \tag{29}$$

for n being even, or

$$(B_0), B_1, B_2, (B_3), B_4, B_5, (B_6), \ldots, (B_{d-3}), B_{d-4}, B_{d-3}, (B_{d-2}), B_{d-1}, B_d, \tag{30}$$

for n being odd.

Note that $B_0 = 0$, and other points in the parentheses are generated in Stage 1. The difference between (29) and (30) is at the tails, namely there is only one unknown point B_d after the generated point B_{d-1} in (29), while there are two unknown points B_{d-1} and B_d after the generated point B_{d-2} in (30).

If n is even, to generate the points in (29), according to Remark 2 we only need to compare the situation of generating the new point conditional on the past value B_{d-1} (this corresponds to Situation 1), and the situation of generating the new point conditional on the past and the future values B_0, B_3 (this corresponds to Situation 2). For the former one, according to Theorem 2 (or a straight calculation), the squared norm of the corresponding column for the point B_d is $1/d$. While for the latter one, according to Theorem 3, the squared norm of the corresponding column for the point B_1 (or the point B_2) is $5/(6d)$, which is smaller than $1/d$. Therefore, in (29) we first sample B_d conditional on the value of B_{d-1}. Note that after B_d is sampled we only need to consider Situation 2, then Theorem 3 and Remark 2 determine the generation order of the remaining points in (29).

Similarly, if n is odd, it can also be found that B_d should be sampled first in (30), and the other points are generated in the same way as above.

Now we know clearly about the optimal generation order of a discrete time Brownian motion $B_1, \ldots, B_{2^n}$. For example:

- for $n = 3$, we generate the Brownian motion as follows (given $B_0 = 0$):

B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8
4	7	**2**	5	8	**1**	6	3

  The numbers in the second line are the generation order of the Brownian path. B_6, B_3 are generated in Stage 1, while $B_8, B_1, B_4, B_2, B_5, B_7$ are generated in Stage 2. The optimal permutation Π^* is {6, 3, 8, 1, 4, 7, 2, 5}.
- for $n = 4$, we generate the Brownian motion as follows (given $B_0 = 0$):

B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9	B_{10}	B_{11}	B_{12}	B_{13}	B_{14}	B_{15}	B_{16}
7	12	4	8	13	**2**	9	14	5	10	15	**1**	11	16	3	6

The numbers in the second line are the generation order of the Brownian path. B_{12}, B_6, B_{15}, B_3, B_9 are generated in Stage 1, while the other points are generated in Stage 2. The optimal permutation Π^* is {12, 6, 15, 3, 9, 16, 1, 4, 7, 10, 13, 2, 5, 8, 11, 14}.

In the Appendix, we list the optimal permutation Π^* for $d = 2, 4, 8, 16, 32, 64, 128, 256$.

Remark 6. In this subsection, we only focused on finding the optimal generation order of a Brownian motion when d is a power of 2. Since the conclusions of Theorems 2 and 3 are also true for an arbitrary dimension d, the construction of NBB can be generalized to the problem with the dimension not being a power of 2. We only need to choose the new step such that it maximizes the variance explained by the new variable. Relations (18) and (22) are useful to get other optimal constructions in some other criterions.

3.4. Numerical comparisons of the explained variance

In this subsection, we compare the effects of NBB and BB by comparing their cumulative explained variances. In Table 1, we show $\sum_{i=1}^{k} \|c_i\|^2$, the sums of squared norm of the first k columns in (1), with $d = 2^n$ and k varying from 1 to 6 for NBB and BB, respectively. We observe that the cumulative variances explained by the first three dimensions in NBB are larger than those

Table 1
The sums of squared norms, $\sum_{i=1}^{k} \|c_i^{\text{NBB}}\|^2$ and $\sum_{i=1}^{k} \|c_i^{\text{BB}}\|^2$, with $k = 1, \ldots, 6$

	$\sum_{i=1}^{k} \|c_i^{\text{NBB}}\|$	$\sum_{i=1}^{k} \|c_i^{\text{BB}}\|$	$\lim_{d\to\infty} \sum_{i=1}^{k}(\|c_i^{\text{NBB}}\|^2 - \|c_i^{\text{BB}}\|^2)/\|A\|_F^2$
$k = 1$	$(\frac{9}{4}d^2 + \frac{9}{4}d + 1)/(6d)$	$(2d^2 + 3d + 1)/(6d)$	$\frac{1}{12}$
$k = 2$	$(\frac{81}{32}d^2 + \frac{9}{4}d + 2)/(6d)$	$(\frac{5}{2}d^2 + 3d + 2)/(6d)$	$\frac{1}{96}$
$k = 3$	$(\frac{171}{64}d^2 + \frac{45}{16}d + 3)/(6d)$	$(\frac{21}{8}d^2 + 3d + 3)/(6d)$	$\frac{1}{64}$
$k = 4$	$(\frac{351}{128}d^2 + \frac{45}{16}d + 4)/(6d)$	$(\frac{22}{8}d^2 + 3d + 4)/(6d)$	$-\frac{1}{384}$
$k = 5$	$(\frac{360}{128}d^2 + \frac{45}{16}d + 5)/(6d)$	$(\frac{89}{32}d^2 + 3d + 5)/(6d)$	$\frac{1}{96}$
$k = 6$	$(\frac{1449}{512}d^2 + \frac{45}{16}d + 6)/(6d)$	$(\frac{90}{32}d^2 + 3d + 6)/(6d)$	$\frac{3}{512}$

In the fourth column we show the limits as $d \to \infty$.

Table 2
The cumulative explained variance (%) from the first six dimensions for the standard, BB, NBB and PCA constructions with dimension $d = 64$

k	Standard	BB	NBB	PCA
1	3.08	67.19	75.01	81.07
2	6.11	83.61	84.25	90.08
3	9.09	87.72	89.16	93.32
4	12.02	91.83	91.47	94.98
5	14.90	92.86	93.79	95.99
6	17.74	93.89	94.37	96.66

in BB. However, under the optimization criterion we chosen, the procedure cannot guarantee that the sums of the squared norms of the decomposition matrix A in NBB are always larger than these in BB (see the case for $k = 4$).

In Table 2, we show for $d = 64$ the cumulative variance explained by the first six dimensions for the standard, BB, NBB and PCA constructions. It is clear that NBB outperforms ordinary BB in allocating variance to the initial dimensions, but not as well as PCA. For example, when $k = 1$, the variance explained by the first dimension for NBB is 75.01%, which is larger than that of BB (67.19%) and smaller than that of PCA (81.07%).

4. Numerical examples

Now we apply NBB to problems of valuing options, including European-type options and American-type options. For completeness, a brief review of ANOVA decomposition is given here. Let $\Upsilon = \{1, 2, \ldots, d\}$. For $u \subseteq \Upsilon$, let u^{c} denote the complementary set of u in Υ, $u - v$ denote the set difference $\{j | j \in u, j \overline{\in} v\}$ and $|u|$ denote the cardinality of u. Assume that $\int_{[0,1)^d} f^2(\mathbf{x})\, d\mathbf{x} < \infty$, then f has an ANOVA decomposition

$$f(\mathbf{x}) = \sum_{u \subseteq \Upsilon} f_u(\mathbf{x}), \tag{31}$$

where $f_u(\mathbf{x}) = \int_{[0,1)^d} f(\mathbf{x})\, d\mathbf{x}_{u^c} - \sum_{v \subsetneq u} f_v(\mathbf{x})$, and $f_\phi = I(f) = \int_{[0,1)^d} f(\mathbf{x})\, d\mathbf{x}$. Here $\mathbf{x}_{u^c} = (x_j)_{j \in u^c}$ denotes the vector of variable x_j whose index j is in u^c. From the definition we see that the term f_u depends only on the variable $\mathbf{x}_u$.

Define the variance of f

$$\sigma^2(f) = \int_{[0,1)^d} f^2(\mathbf{x})\, d\mathbf{x} - I^2(f), \tag{32}$$

and the variance of f_u

$$\sigma_u^2(f) = \int_{[0,1)^d} f_u^2(\mathbf{x})\, d\mathbf{x} \quad \text{for } u \neq \phi \text{ and } \sigma_\phi^2(f) = 0, \tag{33}$$

then we have

$$\sigma^2(f) = \sum_{u \subseteq \Upsilon} \sigma_u^2(f). \tag{34}$$

The quantity σ_u^2 denotes the contribution of f_u to σ^2, and thus σ_u^2/σ^2 measures the relative importance of f_u in the decomposition.

For a given subset u, define $D_u = \sum_{\upsilon \subseteq u} \sigma_\upsilon^2$ and $D_u^{\text{tot}} = \sum_{\upsilon \cap u \neq \phi} \sigma_\upsilon^2$. The truncation dimension of f is the smallest integer d_t such that

$$D_{\{1,\ldots,d_t\}} = \sum_{v \subseteq \{1,\ldots,d_t\}} \sigma_v^2 \geqslant p\, \sigma^2, \tag{35}$$

where $0 < p < 1$. Usually we take $p = 0.99$. The truncation dimension reveals how many important variables are there in the function f.

The dimension distribution in the superposition sense is a discrete probability with the mass function

$$\mu(j) = \frac{1}{\sigma^2} \sum_{|u|=j} \sigma_u^2, \quad j = 1, \ldots, d. \tag{36}$$

In other words, if we have a randomly chosen nonempty set $U \subseteq \{1, \ldots, d\}$ with $Pr(U = u) = \sigma_u^2/\sigma^2$, then $Pr(|U| = j) = \mu(j)$ (see [17]). It is shown in [11] that for $d \geqslant 2$,

$$\overline{d} = E(|U|) = \frac{1}{\sigma^2} \sum_{j=1}^{d} D_{\{j\}}^{\text{tot}}. \tag{37}$$

The quantities D_u and D_u^{tot} can be computed by the following formulas (see [22]):

$$D_u = \int f(\mathbf{x}) f(\mathbf{x}_u, \mathbf{y}_{u^c})\, d\mathbf{x}\, d\mathbf{y}_{u^c} - I^2(f), \tag{38}$$

$$D_u^{\text{tot}} = \frac{1}{2} \int (f(\mathbf{x}_u, \mathbf{x}_{u^c}) - f(\mathbf{y}_u, \mathbf{x}_{u^c}))^2\, d\mathbf{x}\, d\mathbf{y}_u. \tag{39}$$

4.1. Asian option

First, we consider the Asian option on the discrete arithmetic average. Assume that the strike price is K, the initial stock price is S_0, and the final expiration date is T. Let S_t denote the price of the stock at time t. The Black–Scholes model describes the evaluation of the stock price through the stochastic differential equation as

$$\frac{dS_t}{S_t} = \mu\, dt + \widehat{\sigma}\, dB_t, \tag{40}$$

where B_t is a standard Brownian motion. The parameter μ denotes the mean rate of return and $\widehat{\sigma}$ is the volatility of the stock price. Under the risk-neutral measure, taking the rate of return to be the same as the interest rate r, we have the solution of (40) as

$$S_t = S_0 \exp([r - \tfrac{1}{2}\widehat{\sigma}^2]t + \widehat{\sigma} B_t).$$

We now can simulate the path of the stock price at equally spaced times $t_1, \ldots, t_d$

$$S_{t_i} = S_0 \exp([r - \tfrac{1}{2}\widehat{\sigma}^2]t_i + \widehat{\sigma} B_{t_i}), \quad i = 1, 2, \ldots, d. \tag{41}$$

Denote $\overline{S} = \frac{1}{d}\sum_{i=1}^{d} S_{t_i}$. The terminal payoff of an Asian call option is

$$\max(\overline{S} - K, 0),$$

and the value of the option at time 0, based on the risk-neutral valuation principle, is

$$E_\beta[e^{-rT} \max(\overline{S} - K, 0)], \tag{42}$$

where $E_\beta(\cdot)$ is the expectation under the risk-neutral measure P_β (see [6]). We can express the option price as an integral

$$\begin{aligned} I_d(f) = \int_{[0,1)^d} e^{-rT} \max\left(\frac{1}{d}\sum_{i=1}^{d} S_0 \exp\left[\left(r - \frac{\widehat{\sigma}^2}{2}\right) t_i \right.\right. \\ \left.\left. + \widehat{\sigma} B_i(x_1, \ldots, x_d)\right] - K, 0\right) d\mathbf{x}. \end{aligned} \tag{43}$$

Here $B_i(x_1, \ldots, x_d)$ has different expressions in different constructions (see Section 2). For example, for the standard construction, $B_i = \sqrt{T/d}\sum_{j=1}^{i} \Phi^{-1}(x_j)$, where $\Phi(\cdot)$ is the standard normal distribution function.

We perform the experiments in two parts. In the first part, we empirically investigate the effect of NBB for dimension reduction; in the second part, we compare the efficiency of NBB with other dimension reduction techniques in practical computations. In this example, we use the following parameters: $K = 100, T = 1, \widehat{\sigma} = 0.2, r = 0.1$. Tables 3–5 present the truncation dimension (with $p = 0.99$), the first order indices (i.e., FOI $:= \sum_{i=1}^{d} \sigma_{\{i\}}^2/\sigma^2$) and the mean dimension for the standard, BB, NBB and PCA constructions. The results show that PCA, NBB and BB reduce the truncation dimension considerably compared with the standard construction, and in most cases NBB is more powerful than BB in dimension reduction. Comparing the first order indices (e.g., $S_0 = 90$), we can see that the contribution to σ^2 from the first order terms in NBB is about 89–90%, while in BB it is about 80–84%. It means that NBB allocates more variance to the one-dimensional terms. The comparisons of mean dimension demonstrate that in all cases NBB has lower mean dimension than BB, and in NBB the values are closer to 1. This indicates that in NBB the underlying integrand has stronger one-dimensional structure. Note that NBB does

Table 3
The truncation dimension, the sensitivity index of x_1 and the mean dimension for valuing Asian options with the initial stock price $S_0 = 90$

$S_0 = 90$	Standard			BB			NBB			PCA		
d	d_t	FOI	$\overline{d}$	d_t	FOI	$\overline{d}$	d_t	FOI	$\overline{d}$	d_t	FOI	$\overline{d}$
8	7	0.6322	1.4452	6	0.8421	1.1642	6	0.9054	1.0992	2	0.9909	1.0083
16	14	0.5860	1.5559	7	0.8243	1.1889	7	0.9009	1.1089	2	0.9903	1.0099
32	28	0.5540	1.6266	8	0.8071	1.2048	7	0.8885	1.1150	2	0.9888	1.0104
64	54	0.5443	1.6533	8	0.8078	1.1917	7	0.8943	1.1155	2	0.9902	1.0089
128	107	0.5245	1.6889	8	0.8063	1.2089	8	0.8939	1.1162	2	0.9916	1.0092

Table 4
The same as Table 3, but with the initial stock price $S_0 = 100$

$S_0 = 100$	Standard			BB			NBB			PCA		
d	d_t	FOI	$\overline{d}$	d_t	FOI	$\overline{d}$	d_t	FOI	$\overline{d}$	d_t	FOI	$\overline{d}$
8	7	0.8543	1.1725	5	0.9360	1.0645	5	0.9624	1.0396	2	0.9962	1.0039
16	14	0.8307	1.2066	7	0.9318	1.0725	5	0.9603	1.0419	2	0.9961	1.0042
32	27	0.8137	1.2245	7	0.9279	1.0777	5	0.9582	1.0446	2	0.9950	1.0069
64	53	0.8143	1.2380	7	0.9262	1.0786	5	0.9600	1.0378	2	0.9974	1.0047
128	105	0.8119	1.2368	7	0.9349	1.0773	6	0.9622	1.0429	2	0.9983	1.0042

not perform as well as PCA in dimension reduction, but NBB has an advantage in computational cost.

In Tables 6 and 7, we show the sample standard deviation for MC and QMC, and the variance reduction factors (in the parentheses). For QMC, we use the Sobol's sequence randomized by random digital shift, which preserves the (t, k, s)-net properties. We apply m independent randomizations ($m = 50$ in our computations) and compute the corresponding sample mean and sample variance, then the sample standard deviation is

$$\left[\frac{1}{m \cdot (m-1)} \sum_{j=1}^{m} (I_{j,N} - \overline{I}_m)^2\right]^{1/2}, \tag{44}$$

where $I_{j,N}$ is the estimated value in the jth randomization and $\overline{I}_m$ is the corresponding sample mean. The variance reduction factor of one estimate with respect to crude MC estimate is the inverse ratio of their sample variances. We see that when the dimension of the problem increases, the corresponding sample standard deviation under the standard construction also increases, while under BB, NBB and PCA, it is insensitive to the dimension. Furthermore, the variance reduction factors illustrate that NBB is more efficient than BB. QMC+NBB improves QMC + BB with variance reduction factor approximately 1.2, while the computational costs are the same. This improvement is more clear when N is large. In the case $S_0 = 110$, the efficiency improvement is satisfying (QMC + NBB improves QMC + BB with variance reduction factor approximately 1.44), and this is consistent with the result of effective dimension.

Table 5
The same as Table 3, but with the initial stock price $S_0 = 110$

$S_0 = 110$	Standard			BB			NBB			PCA		
d	d_t	FOI	$\bar{d}$	d_t	FOI	$\bar{d}$	d_t	FOI	$\bar{d}$	d_t	FOI	$\bar{d}$
8	7	0.9504	1.0579	5	0.9798	1.0214	5	0.9878	1.0132	2	0.9989	1.0012
16	14	0.9463	1.0644	6	0.9798	1.0225	5	0.9892	1.0129	2	0.9997	1.0013
32	26	0.9473	1.0661	6	0.9793	1.0227	5	0.9889	1.0134	2	0.9999	1.0018
64	52	0.9506	1.0739	7	0.9802	1.0236	5	0.9900	1.0126	2	1.0001	1.0014
128	104	0.9574	1.0671	7	0.9869	1.0228	5	0.9983	1.0125	2	1.0003	1.0010

Table 6
The estimated sample standard deviations and the variance reduction factors (in the parentheses) for $d = 64$ with 50 repetitions

	N	MC	QMC + STD	QMC + BB	QMC + NBB	QMC + PCA
$S_0 = 90$	2^8	4.41e − 2	2.11e − 2 (4)	7.50e − 3 (34)	7.20e − 3 (37)	6.70e − 3 (43)
	2^{10}	2.23e − 2	7.10e − 3 (9)	2.10e − 3 (112)	2.10e − 3 (112)	1.80e − 3 (153)
	2^{12}	1.04e − 2	3.00e − 3 (12)	7.03e − 4 (218)	6.32e − 4 (270)	5.67e − 4 (336)
$S_0 = 100$	2^8	6.40e − 2	2.46e − 2 (6)	1.07e − 2 (35)	1.03e − 2 (38)	9.10e − 3 (49)
	2^{10}	3.80e − 2	8.60e − 3 (19)	3.20e − 3 (141)	2.60e − 3 (213)	2.30e − 3 (273)
	2^{12}	1.98e − 2	2.60e − 3 (58)	9.27e − 4 (455)	8.43e − 4 (551)	7.19e − 4 (757)
$S_0 = 110$	2^8	9.15e − 2	2.90e − 2 (10)	1.48e − 2 (38)	1.39e − 2 (43)	1.30e − 2 (49)
	2^{10}	5.11e − 2	9.20e − 3 (31)	3.50e − 3 (213)	3.30e − 3 (239)	3.10e − 3 (271)
	2^{12}	2.42e − 2	2.50e − 3 (93)	1.20e − 3 (406)	1.00e − 3 (585)	9.34e − 4 (671)

Table 7
The same as Table 6, but for $d = 128$

	N	MC	QMC + STD	QMC + BB	QMC + NBB	QMC + PCA
$S_0 = 90$	2^8	4.63e − 2	3.16e − 2 (2)	7.30e − 3 (40)	7.00e − 3 (44)	6.30e − 3 (54)
	2^{10}	2.52e − 2	1.76e − 2 (2)	2.60e − 3 (94)	2.30e − 3 (120)	2.10e − 3 (144)
	2^{12}	1.17e − 2	7.20e − 3 (2)	6.59e − 4 (315)	5.09e − 4 (529)	5.06e − 4 (535)
$S_0 = 100$	2^8	9.04e − 2	3.73e − 2 (6)	1.00e − 2 (82)	9.30e − 3 (94)	8.70e − 3 (108)
	2^{10}	4.01e − 2	1.49e − 2 (7)	3.40e − 3 (139)	3.00e − 3 (178)	2.70e − 3 (220)
	2^{12}	1.88e − 2	7.20e − 3 (7)	7.86e − 4 (571)	7.21e − 4 (681)	6.59e − 4 (815)
$S_0 = 110$	2^8	9.03e − 2	2.81e − 2 (10)	1.45e − 2 (39)	1.34e − 2 (45)	1.10e − 2 (67)
	2^{10}	5.09e − 2	1.16e − 2 (19)	3.80e − 3 (179)	3.60e − 3 (200)	3.10e − 3 (269)
	2^{12}	2.46e − 2	2.8e − 3 (77)	9.62e − 4 (653)	8.72e − 4 (796)	7.83e − 4 (986)

4.2. American–Bermuda–Asian option

The valuation of American options remains one of the most challenging problems in computational finance. The computational cost of traditional valuation methods increases rapidly with

Table 8
The estimated values of American–Bermuda–Asian options, the corresponding sample standard deviations and the variance reduction factors (in the parentheses) for d=64 with 50 repetitions

		MC		QMC + STD		QMC + BB		QMC + NBB	
A	S_0	Value	S.D.	Value	S.D.	Value	S.D.	Value	S.D.
110	110	17.2228	2.48e − 2	17.2929	6.90e − 3 (13)	17.2921	4.40e − 3 (31)	17.2926	4.60e − 3 (29)
110	100	9.7532	2.27e − 2	9.7966	5.90e − 3 (15)	9.8051	4.70e − 3 (23)	9.8111	4.00e − 3 (32)
110	90	4.2358	1.76e − 2	4.2577	5.40e − 3 (10)	4.2551	2.30e − 3 (58)	4.2585	2.00e − 3 (77)
100	110	15.7553	3.38e − 2	15.7222	7.90e − 3 (18)	15.7151	3.10e − 3 (119)	15.7214	3.00e − 3 (127)
100	100	8.6920	2.63e − 2	8.7344	4.20e − 3 (39)	8.7292	2.20e − 3 (143)	8.7308	2.30e − 3 (130)
100	90	3.7566	1.75e − 2	3.7832	4.40e − 3 (16)	3.7774	1.60e − 3 (119)	3.7789	1.40e − 3 (156)
90	110	14.5469	3.13e − 2	14.5866	7.40e − 3 (18)	14.5793	2.70e − 3 (134)	14.5807	2.30e − 3 (185)
90	100	7.9437	2.56e − 2	7.9750	6.20e − 3 (17)	7.9704	1.90e − 3 (181)	7.9694	1.80e − 3 (202)
90	90	3.3484	1.78e − 2	3.3911	5.60e − 3 (10)	3.3870	1.40e − 3 (161)	3.3881	1.10e − 3 (261)

Here A is the initial average value of the stock and the initial stock price is S_0. We use $\widehat{\sigma} = 0.2$ and $r = 0.06$.

the number of underlying securities and other payoff-related variables. In recent years, several simulation-based methods have been proposed to estimate American option prices (see [6]). In this paper, we use the least-squares MC method proposed in [12].

We value a particular option considered in [12]: American–Bermuda–Asian option, which is a call option on the average price of a stock over some time horizon, where the call option can be exercised at any time after some initial lockout period. Define the current valuation date as time 0. We assume that the option has a final expiration date of $T = 2$ and that the holder may exercise the option at any time after $t = 0.25$ by payment of the strike price $K = 100$. The underlying average $A_t, 0.25 \leqslant t \leqslant T$, is the continuous arithmetic average of the underlying stock price during the period three months prior to the valuation date (a three-month lookback) to time t. Thus, the cash flow from exercising the option at time t is $\max(0, A_t - K)$.

We use the least-squares approach to evaluate the option. The choice of the basis functions in our tests is the set of simple powers of the state variables instead of Laguerre polynomials in [12]. We use a constant, the stock price, the average stock price, the squares of each stock price, the product of the two, and the product of the two up to degree three.

The first step in least-square method is to generate the paths of the asset prices. BB can be used to generate asset prices, and a significant increase in the rate of convergence can be obtained (see [5]). In this paper, we are interested in the rate of convergence of using NBB. We fix the dimension $d = 64$ and the number of paths $N = 2^{12}$. Table 8 illustrates the sample standard deviation for MC and QMC, and the variance reduction factors (in parentheses). We observe that NBB is powerful in valuing American–Bermuda–Asian options, and the variance reduction factors range from 29 to 261 for different parameters. In most cases, the sample standard deviation in NBB is smaller than that in BB, though the superiority is not so impressive as in Asian options. We also see that the results are impacted by options' characteristics like A and S_0. We believe that NBB may be more efficient when the first few steps of the Brownian path are more important.

5. Conclusion

BB is a permutation-based construction of Brownian motion and is often used to speed up QMC integration through dimension reduction. In this paper, we propose a new Brownian bridge

(NBB) construction. We find an optimal generation order of the Brownian motion in the sense of explained variance such that the first several dimensions are even more important than those in BB, while the computational costs are the same. We perform two empirical experiments: valuing Asian options and American–Bermuda–Asian options. For Asian options, we compare the truncation dimension and the mean dimension, and find that in most cases NBB is more powerful than BB in dimension reduction. We also show that compared with BB, the one-dimensional terms of NBB contributes more to the variance σ^2. For both types of options, we find that NBB has better performance than BB in QMC integration.

We point out that the NBB construction is based on the criterion of "explained variance", which is only related to the covariance matrix of the discrete Brownian motion. The advantage of this choice is that since the optimal permutation does not depend on the integrand, NBB is easy to use and can improve the performance of QMC in some financial problems, just as the ordinary BB and PCA. On the other hand, this criterion also has its disadvantage. The new construction does not take into account the function to be integrated, and it does not guarantee to give a consistent improvement for arbitrary examples. In fact, just as ordinary BB, NBB also performs worse than the standard method for problems studied in [19] (we did not give the results in this paper).

The possible disadvantage rises an interesting problem for future research: can we find an optimal variable transformation such that it minimizes the error of QMC integration for a given function or a class of functions? Or alternatively, by taking into account the knowledge of the integrand, can we find an optimal variable transformation such that it maximizes the first order index of the first variable or the first order indices in ANOVA decomposition? For example, though it is not reported here, we have numerical results for some simple functions, and we find that the construction of Brownian motion found by a similar idea has superiority over ordinary BB in approximating Wiener path integral considered in [23].

Here we also want to discuss the feasibility of extending this approach to the multidimensional case. It is difficult to directly generalize the NBB to case of multi-asset path dependent options. However, by using the two-stage dimension reduction technique developed in [24], we may easily combine NBB with other techniques in the case of multiple assets.

Appendix A.

The optimal permutation $\Pi^* = \{\pi(1), \pi(2), \ldots, \pi(d)\}$ for $d = 2, 4, 8, 16, 32, 64, 128, 256$ is given below:

$d = 2$	2	1														
$d = 4$	3	4	1	2												
$d = 8$	6	3	8	1	4	7	2	5								
$d = 16$	12	6	15	3	9	16	1	4	7	10	13	2	5	8	11	14
$d = 32$	24	12	30	6	18	3	9	15	21	27	32	1	4	7	10	
	16	19	22	25	28	31	2	5	8	11	14	17	20	23	26	29
$d = 64$	48	24	60	12	36	6	18	30	42	54	63	3	9	15	21	27
	33	39	45	51	57	64	1	4	7	10	13	16	19	22	25	28

	31	34	37	40	43	46	49	52	55	58	61	2	5	8	11	14
	17	20	23	26	29	32	35	38	41	44	47	50	53	56	59	
$d = 128$	96	48	120	24	72	12	36	60	84	108	126	6	18	30	42	54
	66	78	90	102	114	3	9	15	21	27	33	39	45	51	57	63
	69	75	81	87	93	99	105	111	117	123	128	1	4	7	10	13
	16	19	22	25	28	31	34	37	40	43	46	49	52	55	58	61
	64	67	70	73	76	79	82	85	88	91	94	97	100	103	106	109
	112	115	118	121	124	127	2	5	8	11	14	17	20	23	26	29
	32	35	38	41	44	47	50	53	56	59	62	65	68	71	74	77
	80	83	86	89	92	95	98	101	104	107	110	113	116	119	122	125
$d = 256$	192	96	240	48	144	24	72	120	168	216	252	12	36	60	84	108
	132	156	180	204	228	6	18	30	42	54	66	78	90	102	114	126
	138	150	162	174	186	198	210	222	234	246	255	3	9	15	21	27
	33	39	45	51	57	63	69	75	81	87	93	99	105	111	117	123
	129	135	141	147	153	159	165	171	177	183	189	195	201	207	213	219
	225	231	237	243	249	256	1	4	7	10	13	16	19	22	25	28
	31	34	37	40	43	46	49	52	55	58	61	64	67	70	73	76
	79	82	85	88	91	94	97	100	103	106	109	112	115	118	121	124
	127	130	133	136	139	142	145	148	151	154	157	160	163	166	169	172
	175	178	181	184	187	190	193	196	199	202	205	208	211	214	217	220
	223	226	229	232	235	238	241	244	247	250	253	2	5	8	11	14
	17	20	23	26	29	32	35	38	41	44	47	50	53	56	59	62
	65	68	71	74	77	80	83	86	89	92	95	98	101	104	107	110
	113	116	119	122	125	128	131	134	137	140	143	146	149	152	155	158
	161	164	167	170	173	176	179	182	185	188	191	194	197	200	203	206
	209	212	215	218	221	224	227	230	233	236	239	242	245	248	251	254

Acknowledgement

The authors are thankful for comments from Professor Art B. Owen and two referees. The support of the National Science Foundation of China is gratefully acknowledged.

References

[1] P. Acworth, M. Broadie, P. Glasserman, A comparison of some Monte Carlo and quasi-Monte Carlo techniques for option pricing, in: P. Hellekalek, G. Larcher, H. Niederreiter, P. Zinterhof (Eds.), Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, Springer, Berlin, 1998, pp. 1–18.

[2] F. Åkesson, J.P. Lehoczky, Path generation for quasi-Monte Carlo simulation of mortgage-backed securities, Management Sci. 46 (9) (2000) 1171–1187.

[3] P. Boyle, M. Broadie, P. Glassman, Monte Carlo methods for security pricing, J. Econom. Dynamics Control 21 (1997) 1267–1321.

[4] R.E. Caflisch, W. Morokoff, A.B. Owen, Valuation of Mortgage backed securities using Brownian bridges to reduce effective dimension, J. Comput. Finance 1 (1997) 27–46.

[5] S.K. Chaudhary, American options and the LSM algorithm: quasi-random sequences and Brownian bridges, J. Comput. Finance 8 (2005) 101–115.

[6] P. Glasserman, Monte Carlo Methods in Financial Engineering, Springer, New York, 2004.

[7] J. Imai, K.S. Tan, Minimizing effective dimension using linear transformation, in: H. Niederreiter (Ed.), Monte Carlo and Quasi-Monte Carlo Methods 2002, Springer, Berlin, 2004, pp. 275–292.

[8] G. Larcher, G. Leobacher, K. Scheicher, On the tractibility of the Brownian bridge algorithm, J. Complexity 19 (2003) 511–528.

[11] R. Liu, A.B. Owen, Estimating mean dimensionality of ANOVA decompositions, J. Amer. Statist. Assoc. 101 (474) (2006) 712–721.
[12] F.A. Longstaff, E.S. Schwartz, Valuing American options by simulation: a simple least-squares approach, Rev. Financial Stud. 14 (1) (2001) 113–147.
[13] W.J. Morokoff, Generating quasi-random paths for stochastic processes, SIAM Rev. 40 (1998) 765–788.
[14] B. Moskowitz, R.E. Caflisch, Smoothness and dimension reduction in quasi-Monte Carlo methods, Math. Comput. Modelling 23 (1996) 37–54.
[15] H. Niederreiter, Random number generation and quasi-Monte Carlo methods, SIAM, Philadelphia, 1992.
[16] S. Ninomiya, S. Tezuka, Toward real time pricing of complex financial derivatives, Appl. Math. Finance 3 (1996) 1–20.
[17] A.B. Owen, The dimension distribution and quadrature test functions, Statist. Sinica 13 (1) (2003) 1–17.
[18] A. Papageorgiou, Fast convergence of quasi-Monte Carlo for a class of isotropic integrals, Math. Comput. 70 (2001) 297–306.
[19] A. Papageorgiou, The Brownian bridge does not offer a consistent advantage in quasi-Monte Carlo integration, J. Complexity 18 (2002) 171–186.
[20] S.H. Paskov, J.F. Traub, Faster valuation of financial derivatives, J. Portfolio Management 22 (1995) 113–120.
[21] I.H. Sloan, H. Woźniakowski, When are quasi-Monte Carlo algorithms efficient for high dimensional integrals, J. Complexity 14 (1998) 1–33.
[22] I.M. Sobol, Sensitivity estimates for nonlinear mathematical models, Math. Modeling Comput. Exp. 1 (1993) 407–414.
[23] I.M. Sobol, S.S. Kucherenko, On global sensitivity analysis of quasi-Monte Carlo algorithms, Monte Carlo Methods Appl. 11 (1) (2005) 83–92.
[24] X. Wang, Dimension reduction techniques in quasi-Monte Carlo methods for option pricing, Working paper.
[25] X. Wang, K.T. Fang, The effective dimension and quasi-Monte Carlo integration, J. Complexity 19 (2003) 101–124.