www.oeaw.ac.at

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

RICAM JOHANN-RADON-INSTITUTE FOR COMPUTATIONAL AND APPLIED MATHEMATICS

Efficient simulation of Lévy areas

K. Scheicher

RICAM-Report 2006-01

www.ricam.oeaw.ac.at

Efficient simulation of Lévy areas

Klaus Scheicher

Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A-4040 Linz, Austria.

Abstract

Discretization methods to simulate stochastic differential equations belong to the main tools in mathematical finance. For Itô processes, there exist several Euler- or Runge-Kutta-like methods which are analogues of well known approximation schemes in the non stochastic case. In the multidimensional case, there appear several difficulties, caused by the mixed second order derivatives. These mixed terms (or more precisely their differences) correspond to special random variables called *Lévy stochastic area* terms. In the present paper, we compare three approximation methods for such random variables with respect to computational complexity and the so called effective dimension.

Key words: Itô Integral, multidimensional stochastic differential equation, Lévy stochastic area, numerical approximation.
2000 MSC: 65C10,65C30

---

1 Introduction

We consider a process S satisfying a stochastic differential equation (SDE) of the form

$$dS(t) = a(t, S(t))dt + b(t, S(t))dW(t), \tag{1.1}$$

with $S(0)$ fixed. In the most general setting we consider that $S(\cdot) \in \mathbb{R}^d$, $W(\cdot) \in \mathbb{R}^m$ and $a : \mathbb{R}^{1+d} \to \mathbb{R}^d$, $b : \mathbb{R}^{1+d} \to \mathbb{R}^{d\times m}$ are measurable functions. The above equation is interpreted in the Itô sense. Since explicit solutions of (1.1) exist only in special cases, in general we have to confine ourselves with numerical approximations. A sequence of numerical approximations $\tilde{S}^h(t), 0 \le t \le T$ of a solution $S(t)$ is said to converge strongly at rate $O(h^\gamma)$, if

$$\mathbb{E}|S(T) - \tilde{S}^h(T)| = O(h^\gamma)$$

Email address: `klaus.scheicher@oeaw.ac.at` (Klaus Scheicher).

Preprint submitted to Elsevier Science *30 January 2006*

for $h \to 0$. Milshtein [11] (cf. also [7,8]) proposed a numerical scheme that converges strongly at rate $O(h)$ if $a \in C^{1,1}(\mathbb{R}^{1+d})$ and $b \in C^{1,2}(\mathbb{R}^{1+d})$. The kth component in this scheme is given by

$$\begin{aligned}\tilde{S}_k^h(t_n+h) = & \tilde{S}_k^h(t_n) + a_k h + \sum_{i=1}^m b_{ki} \Delta W_i(t_n, t_n+h) \\ & + \sum_{i=1}^m \sum_{j=1}^m L^i b_{kj} I_{ij}(t_n, t_n+h),\end{aligned}$$

with $\tilde{S}^h(t_0) = S(t_0)$, $\Delta W_i(t_n, t_n+h) = W_i(t_n+h) - W_i(t_n)$,

$$L^i = \sum_{\ell=1}^d b_{\ell i}(t,x) \frac{\partial}{\partial x_\ell} \quad \text{and} \quad I_{ij} = \int_{t_n}^{t_n+h} [W_i(s) - W_i(t_n)]\, dW_j(s).$$

Since the distributions of $I_{ij}(t_n, t_n+h)$, $\Delta W_i(t_n, t_n+h)$, $\Delta W_j(t_n, t_n+h)$ do not depend on t_n, we will write $I_{ij}(h)$, $\Delta W_i(h)$, $\Delta W_j(h)$. Note that $I_{ii}(h) = (\Delta W_i(h))^2/2 - h/2$.

The iterated Itô integrals $I_{ij}(h)$ are closely related to the so called *Lévy stochastic area* integrals, defined by

$$A_{ij}(h) = I_{ij}(h) - I_{ji}(h).$$

Remark 1.1 Sometimes the Levy area is defined as $A_i j(h)/2$. However, for convenience of notation we will omit the factor $\frac{1}{2}$ in the sequel.

Obviously $A_{ii}(h) = 0$, $A_{ji}(h) = -A_{ij}(h)$ and $\mathbb{E}[A_{ij}(h)] = 0$. Furthermore, it can be proved that

$$\mathbb{V}[A_{ij}(h)] = h^2. \tag{1.2}$$

Since

$$I_{ij}(h) + I_{ji}(h) = \Delta W_i(h) \Delta W_j(h) \quad \text{a.s. for } i \neq j,$$

it is clear that

$$I_{ij}(h) = \frac{A_{ij}(h) + \Delta W_i(h) \Delta W_j(h)}{2} \quad \text{a.s. for } i \neq j.$$

Therefore, the problem reduces to sample triples $(\Delta W_i(h), \Delta W_j(h), A_{ij}(h))$. Unfortunately, there is no exact method known for sampling $A_{ij}(h)$ conditional on $\Delta W_i(h)$ and $\Delta W_j(h)$. Lévy [9] obtained a formula for the characteristic function of the conditional distribution of $A_{ij}(h)$. However, the Fourier-inversion step is not analytically possible in the general case. Several methods to approximate $A_{ij}(h)$ are based on Lévy's formula (cf. [14,16]). In this paper, we analyze the difficulty of this approximation problem in terms of the so called effective dimension.

This paper is organized as follows. In section two, we review the main facts for simulating Brownian motions by decomposing the covariance matrix C. We show that all possible decompositions can be obtained by applying the Cholesky algorithm to an orthogonally similar matrix $P^\top CP$. In section three, we approximate $A_{ij}(h)$ by bilinear forms of normally distributed vectors and show how certain approximation methods correspond to special decompositions of C. In section four, we attempt to measure the quality of the different methods.

2 Decomposition of the Covariance matrix

In the remaining part of the paper, we will use the normalization $h = 1$ and we will use indices instead of arguments. Thus we will write W_t instead of $W(t)$ and so on.

Let $n \geq 1$ and for $i \geq 0$, set $t_i = \frac{i}{n}$. Let W_t be a standard Brownian motion with $W_0 = 0$. Since W_t is a Gaussian process, the random variable $(W_{t_1}, \ldots, W_{t_n})^\top \in \mathbb{R}^n$ is normally distributed, $(W_{t_1}, \ldots, W_{t_n})^\top \sim N(0, C)$. The covariance matrix is given by

$$C = (\min(t_i, t_j))_{i,j=1}^n = \frac{1}{n} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

The matrix C is positive definite. Let $I \in \mathbb{R}^{n \times n}$ be the identity matrix and $C = BB^\top$ be any decomposition of C. If $x' = (x'_1, \ldots, x'_n)^\top \sim N(0, I)$, then $\mathbb{E}[Bx'x'^\top B^\top] = BIB^\top = C$ and thus

$$x := Bx' \sim N(0, C). \tag{2.1}$$

Therefore, the problem of sampling $(W_{t_1}, \ldots, W_{t_n})^\top$ from $N(0, C)$ reduces to finding a matrix B for which $BB^\top = C$. The Cholesky decomposition of C is given by $C = RR^\top$, with

$$R = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 1 & \cdots & \cdots & 1 \end{pmatrix}.$$

For any orthogonal Q, we have $(RQ)(RQ)^\top = RQQ^\top R^\top = C$. Therefore, any orthogonal Q yields a decomposition of C. On the other hand, if $C = BB^\top$ with $R \neq B$ and $Q = R^{-1}B$, then $RR^\top = C = BB^\top = RQQ^\top R^\top$ and thus $QQ^\top = I$. Thus, Q must be orthogonal.

Let P be any orthogonal matrix and H be the lower triangular matrix obtained by Cholesky decomposition of $P^\top CP$. Then

$$P^\top CP = HH^\top \tag{2.2}$$

and thus

$$C = PHH^\top P^\top = (PH)(PH)^\top = (RQ)(RQ)^\top \tag{2.3}$$

with a certain orthogonal matrix Q. Therefore, any decomposition $C = BB^\top$ can be obtained by applying the Cholesky decomposition to a matrix $P^\top CP$. Furthermore, we have $Q = R^{-1}PB$.

Since $\mathbb{V}[Bx'] = \mathbb{E}[x'^\top B^\top Bx'] = \|C\|^2$, where $\|\cdot\|$ denotes the Frobenius norm, the variance does not depend of the choice of B. Therefore, for Monte Carlo simulations it is irrelevant, which matrix B is used. For Quasi Monte Carlo applications, the choice of B is crucial. In this case, the vectors x in (2.1) are transformed from low discrepancy points $\xi = (\xi_1, \ldots, \xi_n)^\top \in [0,1)^n$ by the inversion method or the Box-Muller method. However, all low discrepancy point sets have the property that their one dimensional components (i.e. their projections on the coordinate axes) have different quality. Therefore, one should use the good coordinates for generating the important components.

3 Simulation of the bilinear form

Let X_t, Y_t be two independent standard Brownian motions. Define

$$A_1 = \int_0^1 X_t dY_t - \int_0^1 Y_t dX_t. \tag{3.1}$$

We discretize X_t and Y_t and approximate (3.1) by a the corresponding finite sum. As we have mentioned in the last section, for Quasi Monte Carlo applications, the choice of the decomposition $C = BB^\top$ is crucial.

Let $x = (x_1, \ldots, x_n)^\top$ and $y = (y_1, \ldots, y_n)^\top$ with $x, y \sim N(0, C)$. Then A_1 can be approximated by

$$\begin{aligned} A_1 \approx A_1^{(n)} &= (x_1 y_2 - x_2 y_1) + \cdots + (x_{n-1} y_n - x_n y_{n-1}) && (3.2) \\ &= x^\top M y = x'^\top Q^\top R^\top M R Q y'. && (3.3) \end{aligned}$$

where $x', y' \sim N(0, I)$ and

$$M = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ -1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & -1 & 0 \end{pmatrix}.$$

Let

$$U^{(n)} := Q^\top R^\top M R Q.$$

3.1 The piecewise method

If $Q = I$, i.e. $x = Rx'$ and $y = Ry'$ with $x', y' \sim N(0, I)$, then

$$X_1 = \tfrac{1}{\sqrt{n}}(x'_1 + \cdots + x'_n) \tag{3.4}$$

$$Y_1 = \tfrac{1}{\sqrt{n}}(y'_1 + \cdots + y'_n) \tag{3.5}$$

$$A_1^{(n)} = x'^\top R^\top M R y' \tag{3.6}$$

and

$$U^{(n)} = R^\top M R = \frac{1}{n} \begin{pmatrix} 0 & 1 & \cdots & 1 \\ -1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ -1 & \cdots & -1 & 0 \end{pmatrix}. \tag{3.7}$$

If $Q = I$, i.e. $U^{(n)} = R^\top M R = (u_{ij})_{i,j=1}^n$, it follows that $\mathbb{E}[A_1^{(n)}] = 0$ and

$$\mathbb{V}[A_1^{(n)}] = \mathbb{E}[(x'^\top U^{(n)} y')^2] = \sum_{i,j,k,\ell=1}^n u_{ij} u_{k\ell} \mathbb{E}[x'_i y'_j x'_k y'_\ell] = \sum_{i,j=1}^n u_{ij}^2 = 1 - \tfrac{1}{n}. \tag{3.8}$$

Furthermore, an easy consideration implies that $\mathbb{E}[A_1^{(n)}]$ and $\mathbb{V}[A_1^{(n)}]$ must be independent of Q. Thus, also in the general case when $U^{(n)} = Q^\top R^\top M R Q$, we have

$$\mathbb{V}[A_1^{(n)}] = \sum_{i,j=1}^n u_{ij}^2 = 1 - \tfrac{1}{n} \tag{3.9}$$

whereas (1.2) implies that $\mathbb{V}[A_1] = 1$.

Remark 3.1 From the computational point of view, it is favorable to compute $A_1^{(n)}$ by (3.2) instead of (3.6) and (3.7). In the first case, the number of computational steps is $O(n)$, in the latter case, it is $O(n^2)$ to achieve a proportion of $1 - \frac{1}{n}$ of the variance $\mathbb{V}[A_1]$.

3.2 Simulation by Brownian Bridges

For $n = 2^k$, the Brownian Bridge algorithm might be applied. The number computational steps to generate one discrete sample path is again $O(n)$.

The algorithm works as follows: set $W_0 = W_{t_0} = 0$ and generate $W_{t_n} = W_1 \sim N(0,1)$. For $0 < i < n$, the intermediate values W_{t_i} are generated in permuted order, $(W_{t_{\pi(j)}})_{j=1}^{n-1}$, such that every new t_i is placed into the middle of one of the largest existing intervals. Therefore, the permutation $\pi : \{0,\ldots,n\} \mapsto \{0,\ldots,n\}$ giving the construction order must fulfill $\pi(0) = 0$, $\pi(1) = n$ and

$$\pi^{-1}(k \pm 1) < \pi^{-1}(k) \quad \text{for } 2 \leq k \leq n.$$

Given $W_{t_{i-1}}$ and $W_{t_{i+1}}$, the marginal distribution of W_{t_i} is normal, $W_{t_i} \sim N(\mu_{t_i}, \sigma_{t_i}^2)$ with

$$\mu_{t_i} = \tfrac{1}{2}(W_{t_{i-1}} + W_{t_{i+1}}) \quad \text{and} \quad \sigma_{t_i}^2 = \frac{(t_{i+1} - t_i)(t_i - t_{i-1})}{t_{i+1} - t_{i-1}}.$$

Since $t_i - t_{i-1} = \frac{1}{n}$, we set

$$W_{t_i} = \tfrac{1}{2}(W_{t_{i-1}} + W_{t_{i+1}}) + (2n)^{-\frac{1}{2}} x_i' \tag{3.10}$$

with $x_i' \sim N(0,1)$. Algorithmically, the well known Van der Corput sequence gives such a permutated ordering: if $i = \sum_{j=0}^{k} d_j 2^j$ is the binary representation of i, set $s_2(i) = \sum_{j=0}^{k} d_j 2^{-j-1}$. Thus, $s_2(i)$ is obtained by reflecting the binary digits of i at the comma. Then $t_{\pi(i)}$ is given by

$$t_{\pi(i)} = \begin{cases} 0 & \text{for } i = 0 \\ 1 & \text{for } i = 1 \\ s_2(i-1) & \text{for } i > 1. \end{cases} \tag{3.11}$$

Thus $t_{\pi(j)} = 0, 1, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \ldots$. In the notation of (2.3), this is equivalent to choosing P as a permutation matrix. In (2.2) we obtain that

$$P^\top C P = (\min\{t_{\pi(i)}, t_{\pi(j)}\})_{i,j=1}^{n},$$

and, for example, if $n = 2^3$, we have

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4}\sqrt{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 & \frac{1}{4}\sqrt{2} & 0 & 0 & 0 & 0 \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8}\sqrt{2} & 0 & \frac{1}{4} & 0 & 0 & 0 \\ \frac{5}{8} & \frac{3}{8} & 0 & \frac{1}{8}\sqrt{2} & 0 & \frac{1}{4} & 0 & 0 \\ \frac{3}{8} & \frac{3}{8} & \frac{1}{8}\sqrt{2} & 0 & 0 & 0 & \frac{1}{4} & 0 \\ \frac{7}{8} & \frac{1}{8} & 0 & \frac{1}{8}\sqrt{2} & 0 & 0 & 0 & \frac{1}{4} \end{pmatrix}.$$

Note that the sequence $t_{\pi(i)}$ appears in the first column of H. Then $U^{(n)} = Q^\top R^\top M R Q$ is given by

$$U^{(n)} = \begin{pmatrix} 0 & -\frac{1}{2} & -\frac{1}{8}\sqrt{2} & -\frac{1}{8}\sqrt{2} & -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{16} \\ \frac{1}{2} & 0 & -\frac{1}{8}\sqrt{2} & \frac{1}{8}\sqrt{2} & -\frac{1}{16} & \frac{1}{16} & -\frac{1}{16} & \frac{1}{16} \\ \frac{1}{8}\sqrt{2} & \frac{1}{8}\sqrt{2} & 0 & 0 & -\frac{1}{16}\sqrt{2} & 0 & \frac{1}{16}\sqrt{2} & 0 \\ \frac{1}{8}\sqrt{2} & -\frac{1}{8}\sqrt{2} & 0 & 0 & 0 & -\frac{1}{16}\sqrt{2} & 0 & \frac{1}{16}\sqrt{2} \\ \frac{1}{16} & \frac{1}{16} & \frac{1}{16}\sqrt{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{16} & -\frac{1}{16} & 0 & \frac{1}{16}\sqrt{2} & 0 & 0 & 0 & 0 \\ \frac{1}{16} & \frac{1}{16} & -\frac{1}{16}\sqrt{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{16} & -\frac{1}{16} & 0 & -\frac{1}{16}\sqrt{2} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

By taking higher powers $n = 2^k$, larger matrices $U^{(n)}$ can be generated. These matrices converge to a matrix in $U = (u_{ij})_{i,j=1}^\infty \in \mathbb{R}^{\infty\times\infty}$ corresponding to A_1.

For iid x'_i, $y'_j \sim N(0,1)$, it follows that $X_1 = x'_1$, $Y_1 = y'_1$ and $A_1 = \sum_{i,j=1}^\infty u_{ij} x'_i y'_j$. The largest entries of U are $u_{1,2}$ and $u_{2,1}$, which sum up to

$$u_{1,2}^2 + u_{2,1}^2 = \tfrac{1}{2} = \tfrac{1}{2}\mathbb{V}[A^{(1)}].$$

From the computational point of view, it is preferable to construct $(x_1, \dots, x_n)^\top$, $(y_1, \dots, y_n)^\top$ by (3.10) and $A_1^{(n)}$ by (3.2). In this case, the multiplications by zero are omitted and the number of computational steps is $O(n)$ to achieve a proportion of $1 - \frac{1}{n}$ of $\mathbb{V}[A_1]$.

Remark 3.2 Employing the Brownian Bridge method, any matrix $U^{(n)} \in \mathbb{R}^{n\times n}$ satisfying (3.9) and corresponding to $A_1^{(n)}$, can be extended to a matrix $U = (u_{ij})_{i,j=1}^\infty \in \mathbb{R}^{\infty\times\infty}$ corresponding to A_1, such that $U^{(n)}$ is a submatrix of U and $\sum_{i,j=1}^\infty u_{ij}^2 = 1$.

This construction can be used together with (3.9) to compute a bound for the Mallows 2-distance (cf. [10]) between the random vectors $V_1 := (X_1, Y_1, A_1)^\top$ and $V_1^{(n)} := (X_1, Y_1, A_1^{(n)})^\top$. For any $r > 0$, the Mallows r-distance between two probability distributions F_X and F_Y is defined by

$$d_r(F_X, F_Y) = \left(\inf_{(X,Y)} \mathbb{E}[\|X - Y\|]^r \right)^{\frac{1}{r}},$$

where the infimum is taken over all pairs (X, Y) whose marginal distribution functions are F_X and F_Y respectively. From (3.9) it follows that

$$\begin{aligned} d_2(F_{V_1}, F_{V_1^{(n)}}) &= \left(\inf_{(V_1, V_1^{(n)})} \mathbb{E}[\|V_1 - V_1^{(n)}\|]^2 \right)^{\frac{1}{2}} \leq \left(\mathbb{E}[|A_1 - A_1^{(n)}|]^2 \right)^{\frac{1}{2}} \\ &= \left(1 - \sum_{i,j=1}^{n} u_{ij}^2 \right)^{\frac{1}{2}} = \left(1 - (1 - \tfrac{1}{n})\right)^{\frac{1}{2}} = \frac{1}{\sqrt{n}}. \end{aligned}$$

3.3 Simulation by singular value decomposition

Since C is symmetric and positive definite, all of its eigenvalues λ_i are positive. If $V = (v_{ij})_{i,j=1}^n$ denotes the orthogonal matrix containing the eigenvectors of C as rows, then $CV^\top = V^\top D$ where D is the diagonal matrix containing the eigenvalues λ_i. If $S = (s_{ij})_{i,j=1}^n = V^\top D$, the singular value decomposition of C is given by

$$C = V^\top D V = (V^\top D^{\frac{1}{2}})(V^\top D^{\frac{1}{2}})^\top = SS^\top. \tag{3.12}$$

Lemma 3.3 (cf. [1,5]) *The eigenvalues of C are*

$$\lambda_i = \left(4n \sin^2\left(\tfrac{2i-1}{2n+1} \tfrac{\pi}{2} \right) \right)^{-1} \quad \textit{for } 1 \leq i \leq n. \tag{3.13}$$

The corresponding eigenvectors $\vec{v}_i$ are given by $\vec{v}_i = (v_{i,1}, \ldots, v_{in})^\top$ with

$$v_{ij} = \sin\left(\tfrac{2i-1}{2n+1} j\pi\right). \tag{3.14}$$

Furthermore,

$$\|\vec{v}_i\|_2 = \tfrac{\sqrt{2n+1}}{2}. \tag{3.15}$$

Then

$$s_{ij} = \left(\sqrt{2n^2+n}\sin\left(\tfrac{2j-1}{2n+1}\tfrac{\pi}{2}\right)\right)^{-1}\sin\left(\tfrac{2j-1}{2n+1} i\pi\right). \tag{3.16}$$

Thus, if $(x'_1, \ldots, x'_n) \sim N(0, I)$, a discrete Brownian sample path can be generated by

$$W_{t_i} = \sum_{j=1}^{n} s_{ij} x'_j. \tag{3.17}$$

Theorem 3.4 *If $x = Sx'$ and $y = Sy'$ with $x', y' \sim N(0, I)$, then*

$$X_1 = (2n^2+n)^{-\frac{1}{2}} \sum_{j=1}^{n} \sin\left(\tfrac{2j-1}{2n+1}\tfrac{\pi}{2}\right)^{-1} \sin\left(\tfrac{2j-1}{2n+1} n\pi\right) x'_j \tag{3.18}$$

$$Y_1 = (2n^2+n)^{-\frac{1}{2}} \sum_{j=1}^{n} \sin\left(\tfrac{2j-1}{2n+1}\tfrac{\pi}{2}\right)^{-1} \sin\left(\tfrac{2j-1}{2n+1} n\pi\right) y'_j \tag{3.19}$$

$$A_1^{(n)} = x'^\top S^\top M S y'. \tag{3.20}$$

Furthermore, if

$$T := (t_{ij})_{i,j=1}^{n} = S^\top M S = D^{\frac{1}{2}} V M V^\top D^{\frac{1}{2}},$$

then

$$t_{ij} = \frac{1}{n(2n+1)} \frac{\sin\left(\frac{i-j}{2n+1}\pi\right)^2 \sin\left(\frac{1-i-j}{2n+1}n\pi\right)^2 - \sin\left(\frac{i-j}{2n+1}n\pi\right)^2 \sin\left(\frac{1-i-j}{2n+1}\pi\right)^2}{\sin\left(\frac{2i-1}{2n+1}\frac{\pi}{2}\right)\sin\left(\frac{2j-1}{2n+1}\frac{\pi}{2}\right)\sin\left(\frac{i-j}{2n+1}\pi\right)\sin\left(\frac{1-i-j}{2n+1}\pi\right)}. \tag{3.21}$$

Proof. The equations (3.18) and (3.19) follow immediately from (3.16).

Let $(f_{ij})_{i,j=1}^{n} = VMV^\top$. In order to prove (3.21), matrix multiplication yields

$$\begin{aligned} f_{ij} &= v_{i1}v_{j2} + v_{i2}(v_{j3} - v_{j1}) + \cdots + v_{i,n-1}(v_{jn} - v_{j,n-2}) - v_{in}v_{j,n-1} \\ &= \sum_{k=1}^{n-1} (v_{ik}v_{j,k+1} - v_{i,k+1}v_{jk}). \end{aligned}$$

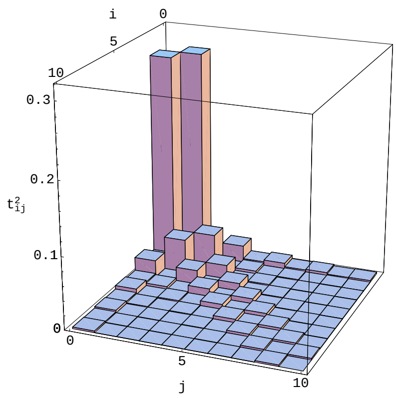

Fig. 1. A matrixplot of $(\bar{t}_{ij}^2)_{i,j=1}^n$ for $n = 10$.

By (3.14), a laborious but not complicated computation (which has been verified by Mathematica) yields

$$v_{ik}v_{j,k+1} - v_{i,k+1}v_{jk} = \frac{1}{2}\left(C_{ijk}^{(1)} + C_{ijk}^{(2)} + C_{ijk}^{(3)} + C_{ijk}^{(4)}\right)$$

with

$$C_{ijk}^{(1)} = \cos(\tfrac{1-2i+2(1-i-j)k}{2n+1}\pi) \qquad C_{ijk}^{(2)} = -\cos(\tfrac{1-2i-2(i-k)k}{2n+1}\pi)$$

$$C_{ijk}^{(3)} = -\cos(\tfrac{1-2j+2(1-i-j)k}{2n+1}\pi) \qquad C_{ijk}^{(4)} = \cos(\tfrac{1-2j+2(i-j)k}{2n+1}\pi).$$

From this follows

$$\begin{aligned}
\sum_{k=1}^{n-1} C_{ijk}^{(1)} &= \cos(\tfrac{i-j-(1-i-j)n}{2n+1}\pi)\frac{\sin(\frac{1-i-j}{2n+1}n\pi)}{\sin(\frac{1-i-j}{2n+1}\pi)} - \cos(\tfrac{2i-1}{2n+1}\pi)\\
\sum_{k=1}^{n-1} C_{ijk}^{(2)} &= -\cos(\tfrac{1-i-j-(i-j)n}{2n+1}\pi)\frac{\sin(\frac{i-j}{2n+1}n\pi)}{\sin(\frac{i-j}{2n+1}\pi)} + \cos(\tfrac{2i-1}{2n+1}\pi)\\
\sum_{k=1}^{n-1} C_{ijk}^{(3)} &= -\cos(\tfrac{i-j+(1-i-j)n}{2n+1}\pi)\frac{\sin(\frac{1-i-j}{2n+1}n\pi)}{\sin(\frac{1-i-j}{2n+1}\pi)} + \cos(\tfrac{2j-1}{2n+1}\pi)\\
\sum_{k=1}^{n-1} C_{ijk}^{(4)} &= \cos(\tfrac{1-i-j+(i-j)n}{2n+1}\pi)\frac{\sin(\frac{i-j}{2n+1}n\pi)}{\sin(\frac{i-j}{2n+1}\pi)} - \cos(\tfrac{2j-1}{2n+1}\pi).
\end{aligned}$$

Therefore

$$
\begin{aligned}
f_{ij} &= \tfrac{1}{2}\Bigg(\Big(\cos(\tfrac{i-j-(1-i-j)n}{2n+1}\pi) - \cos(\tfrac{i-j+(1-i-j)n}{2n+1}\pi)\Big)\frac{\sin(\frac{1-i-j}{2n+1}n\pi)}{\sin(\frac{1-i-j}{2n+1}\pi)} \\
&\quad + \Big(\cos(\tfrac{1-i-j+(i-j)n}{2n+1}\pi) - \cos(\tfrac{1-i-j-(i-j)n}{2n+1}\pi)\Big)\frac{\sin(\frac{i-j}{2n+1}n\pi)}{\sin(\frac{i-j}{2n+1}\pi)}\Bigg) \\
&= \frac{\sin(\frac{i-j}{2n+1}\pi)}{\sin(\frac{1-i-j}{2n+1}\pi)}\sin(\tfrac{1-i-j}{2n+1}n\pi)^2 - \frac{\sin(\frac{1-i-j}{2n+1}\pi)}{\sin(\frac{i-j}{2n+1}\pi)}\sin(\tfrac{i-j}{2n+1}n\pi)^2 \qquad (3.22) \\
&= \frac{\sin(\frac{i-j}{2n+1}\pi)^2\sin(\frac{1-i-j}{2n+1}n\pi)^2 - \sin(\frac{i-j}{2n+1}n\pi)^2\sin(\frac{1-i-j}{2n+1}\pi)^2}{\sin(\frac{i-j}{2n+1}\pi)\sin(\frac{1-i-j}{2n+1}\pi)}.
\end{aligned}
$$

Now (3.21) follows from (3.13) and (3.15) since

$$
t_{ij} = \frac{f_{ij}}{\sqrt{\lambda_i\lambda_j}\|\vec{v}_i\|_2\|\vec{v}_j\|_2}. \qquad (3.23)
$$

□

Remark 3.5 Once again, it is inefficient to compute $(X_1, Y_1, A_1^{(n)})$ by (3.18)–(3.21). The number of computational steps is again $O(n^2)$. Alternatively, one might employ fast Fourier methods to compute $x = Sx'$, $y = Sy'$ and afterwards apply (3.2). Since $2n+1 \neq 2^m$, binary FFT is not applicable. For instance, one can choose n such that $2n+1 = 3^m$.

Since we need two independent Brownian paths, simultaneous double real FFT (cf. [13, Chapter 12.3]) is promising. For $x' = (x'_1, \dots, x'_n)^\top \sim N(0, I)$, define

$$
\begin{aligned}
&\bar{x}_j = \left(\sqrt{2n^2+n}\sin\left(\tfrac{2j-1}{2n+1}\tfrac{\pi}{2}\right)\right)^{-1} x'_j, \quad \text{for } 1 \le j \le n \text{ and} \\
&X = (X_j)_{j=0}^{2n} = (0, \bar{x}_1, 0, \bar{x}_2, \dots, 0, \bar{x}_n, 0).
\end{aligned}
$$

Define $\bar{y}_j, Y$ analogously. If $\xi = \exp(\frac{\mathbf{i}\pi}{2n+1})$, then

$$
\sum_{j=1}^{n} \sin(\tfrac{2j-1}{2n+1}k\pi)\bar{x}_j = \Im\left[\sum_{j=0}^{2n}\xi^{jk}X_j\right]. \qquad (3.24)
$$

For $F = (F_j)_{j=0}^{2n}$, define

$$
\mathscr{F}(F)_k = \sum_{j=0}^{2n} F_j\xi^{kj}, \quad 1 \le k \le n.
$$

If $2n+1 = 3^m$, there is a ternary FFT algorithm to perform $\mathscr{F}(\cdot)$. If F is purely real, then $\mathscr{F}(F)^*_{2n+1-k} = -\mathscr{F}(F)_k$, where * denotes complex conjugation. Thus

$$\begin{aligned}\mathscr{F}(X+Y\mathbf{i})_k &= \mathscr{F}(X)_k + \mathscr{F}(Y)_k\,\mathbf{i}\\ \mathscr{F}(X+Y\mathbf{i})^*_{2n+1-k} &= -\mathscr{F}(X)_k + \mathscr{F}(Y)_k\,\mathbf{i}\end{aligned}$$

and therefore

$$\begin{aligned}\mathscr{F}(X)_k &= \tfrac{1}{2}\left(\mathscr{F}(X+Y\mathbf{i})_k - \mathscr{F}(X+Y\mathbf{i})^*_{2n+1-k}\right)\\ \mathscr{F}(Y)_k &= \tfrac{1}{2\mathbf{i}}\left(\mathscr{F}(X+Y\mathbf{i})_k + \mathscr{F}(X+Y\mathbf{i})^*_{2n+1-k}\right).\end{aligned}$$

The imaginary parts of $(\mathscr{F}(X)_k)_{k=1}^n$ and $(\mathscr{F}(Y)_k)_{k=1}^n$ have the desired form (3.24). Applying (3.2), we obtain $O(n\log n)$ computational steps.

Theorem 3.6 *For iid $x'_i, y'_j \sim N(0,1)$, we have*

$$A_1 = \sum_{i,j=1}^{\infty} \bar{t}_{ij} x'_i y'_j, \tag{3.25}$$

where

$$\bar{t}_{ij} := \lim_{n\to\infty} t_{ij} = \begin{cases} \frac{8}{\pi^2}\frac{i-j}{(2i-1)(2j-1)(1-i-j)} & \text{for } i+j \equiv 0 \pmod 2 \\ -\frac{8}{\pi^2}\frac{1-i-j}{(2i-1)(2j-1)(i-j)} & \text{for } i+j \equiv 1 \pmod 2). \end{cases}$$

Proof. Note that for $i, j, k \in \mathbb{Z}$,

$$\begin{aligned}\lim_{n\to\infty} \sin(\tfrac{kn\pi}{2n+1})^2 &= \sin(\tfrac{k\pi}{2})^2 = k \pmod 2 \quad \text{and}\\ i-j &\not\equiv 1-i-j \pmod 2.\end{aligned}$$

Taking a look at (3.22), this acts as a switch between the two cases. From (3.21) and (3.23), we see that $\lim_{n\to\infty} t_{ij}$ is either equal to

$$\lim_{n\to\infty} \frac{\sin(\frac{i-j}{2n+1}\pi)}{n(2n+1)\sin(\frac{2i-1}{2n+1}\frac{\pi}{2})\sin(\frac{2j-1}{2n+1}\frac{\pi}{2})\sin(\frac{1-i-j}{2n+1}\pi)} = \frac{8}{\pi^2}\frac{i-j}{(2i-1)(2j-1)(1-i-j)}$$

or

$$\lim_{n\to\infty} -\frac{\sin(\frac{1-i-j}{2n+1}\pi)}{n(2n+1)\sin(\frac{2i-1}{2n+1}\frac{\pi}{2})\sin(\frac{2j-1}{2n+1}\frac{\pi}{2})\sin(\frac{i-j}{2n+1}\pi)} = -\frac{8}{\pi^2}\frac{1-i-j}{(2i-1)(2j-1)(i-j)}.$$

$\square$

A view of Figure 1 yields, that the entries $\bar{t}_{ij}^2$ with small indices contribute the main portion to $\mathbb{V}[A_1]$. In particular,

$$\bar{t}_{1,2}^2 + \bar{t}_{2,1}^2 = \tfrac{512}{9\pi^4} = 0.58402\ldots.$$

This is slightly better than the value $\frac{1}{2}$, which is obtained by the Brownian Bridges algorithm. The price of this improvement is an asymptotically larger amount of computational steps.

4 The effective dimension

For $A \subset \mathbb{R}$, denote by $\mu(A)$ the Gaussian measure $\mu(A) = \frac{1}{\sqrt{2\pi}} \int_A \exp(-\frac{x^2}{2})dx$. Let $D = \{1, \ldots, n\}$. For $\varnothing \neq u \subset D$, let $\mathbb{R}^u$ be the $|u|$-dimensional projection of $\mathbb{R}^n$ to the coordinates indexed by u. Analogously, let A_u be the projection of $A \subset \mathbb{R}^n$ to $\mathbb{R}^u$. Let $\bar{u} = D \setminus u$. Note that $\mathbb{R}^n = \mathbb{R}^D = \mathbb{R}^{\bar{\varnothing}}$.

For $\varnothing \neq u \subset D$ and $A_u \subset \mathbb{R}^u$, we write

$$dx_u := \prod_{j \in u} dx_j \quad \text{and} \quad \mu_u(A_u) = \prod_{j \in u} \mu(A_j).$$

Any function $f \in \mathcal{L}^2(\mathbb{R}^n, \mu_D)$ can be expressed as the sum of its *ANOVA-effects* (cf. [2–4,6,12,15]),

$$f(x) = \sum_{u \subset D} f_u(x_u),$$

which are defined recursively by $f_\varnothing = \int_{\mathbb{R}^{\bar{\varnothing}}} f(x)\mu_{\bar{\varnothing}}(dx_{\bar{\varnothing}}) = \int_{\mathbb{R}^D} f(x)\mu_D(dx_D)$ and

$$f_u(x_u) = \int_{\mathbb{R}^{\bar{u}}} f(x)\mu_{\bar{u}}(dx_{\bar{u}}) - \sum_{v \subset u} f_v(x_v).$$

(By convention, we set $\int_{\mathbb{R}^\varnothing} f(x)dx_\varnothing := f(x)$.) Furthermore, the 2^n ANOVA-effects are orthogonal, i.e.

$$\int_{\mathbb{R}^D} f_u(x_u) f_v(x_v)\mu_D(dx_D) = 0 \quad \text{for} \quad u \neq v \tag{4.1}$$

and

$$\int_{\mathbb{R}^D} f_u(x_u)\mu_D(dx_D) = 0.$$

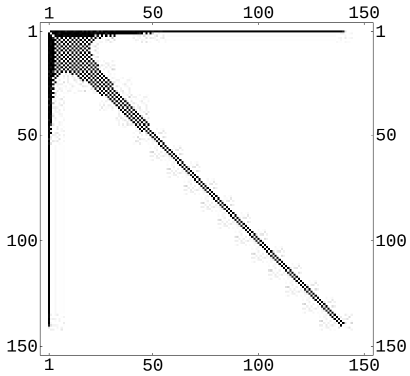

Fig. 2. The 900 largest elements of $(\bar{t}_{ij}^2)_{i,j=1}^{150}$ to obtain $0.99\sigma^2$.

The ANOVA-effect f_u is the part of f, depending only on x_j with $j \subset u$. From (4.1) follows

$$\int_{\mathbb{R}^D} (f(x_D) - f_\varnothing)^2 \mu_D(dx_D) = \sum_{\varnothing \neq u \subset D} \int_{\mathbb{R}^u} f_u(x_u)^2 \mu_u(dx_u).$$

Definition 4.1 *The effective dimension of f, in the superposition sense, is the smallest integer d_S, such that $\sum_{0<|u|\leq d_S} \sigma^2(f_u) \geq 0.99\sigma^2(f)$.*

Definition 4.2 *The effective dimension of f, in the truncation sense, is the smallest integer d_T, such that there is a set $v \subset \mathbb{N}$ with $|v| = d_T$ and $\sum_{u \subset v} \sigma^2(f_u) \geq 0.99\sigma^2(f)$.*

Remark 4.3 The constant 0.99 is an arbitrary choice and one might prefer other values in some settings.

Let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, $x_i, y_i \sim N(0,1)$. Then $A_1^{(n)} = A_1^{(n)}(x,y) = x^\top U y$ with $U = (u_{ij})_{ij=1}^n = Q^\top R^\top M R Q$. An easy consideration yields that $\mathbb{E}[A_1^{(n)}]$ and $\mathbb{V}[A_1^{(n)}]$ are independent of Q. However, the ANOVA-effects of $A_1^{(n)}$ are given by the entries of U and therefore, depend on Q:

$$A_1^{(n)} = x^\top U y = \sum_{i,k=1}^{n} u_{ik} x_i y_k.$$

Theorem 4.4 *The effective dimension in the superposition sense is $d_S = 2$. The effective dimension in the truncation sense is $d_T \leq 900$.*

Proof. The first statement is trivial, since $\sum_{|u|=2} \sigma^2(f_u) = \sigma^2(f)$, i.e. there are only ANOVA-effects of order two. The second statement results from a computation taking into account the submatrix $(\bar{t}_{ij})_{i,j=1}^{150}$. The entries which are needed are depicted in Figure 2. $\square$

Definition 4.5 *The variance proportion of a variable* x_i *is given by*

$$p(x_i) = \frac{1}{\sigma^2(f)} \sum_{\substack{u \subset \mathbb{N} \\ i \in u}} \frac{1}{|u|} \sigma^2(f_u).$$

The variance proportion of a set u *is defined by* $p_u = \sum_{i \in u} p_i$*. The effective dimension in the sense of variance proportion is the smallest integer* d_P*, such that there is a set* $v \subset \mathbb{N}$ *with* $|v| = d_P$ *and* $\sum_{i \in v} p_i \geq 0.99$.

Remark 4.6 The idea behind this notion is the following: every variable should be weighted in a natural way induced by the ANOVA-effects.

In the case of hybrid MC-QMC methods, where some variables are generated by Monte Carlo and others are generated by Quasi Monte Carlo, this notion allows precisely to measure the proportion of variance generated by MC and QMC respectively.

Theorem 4.7 *The variance proportion is given by*

$$p(x_i') = p(y_i') = \begin{cases} \frac{24\pi i - 12\pi + 32}{\pi^3(2i-1)^3} \text{ for } i \equiv 0 \pmod 2 \\ \frac{24\pi i - 12\pi - 32}{\pi^3(2i-1)^3} \text{ for } i \equiv 1 \pmod 2. \end{cases} \tag{4.2}$$

An easy computation yields that the effective dimension in the sense of variance proportion is $d_P = 122$*. Moreover*

$$\sum_{i=n}^{\infty} (p(x_i) + p(y_i)) = O(\tfrac{1}{n}).$$

Proof. From (3.25) clearly follows, since all $p(x_i') = p(y_i') = \sum_{j=1}^{\infty} {t_{ij}'}^2$. We have to consider two cases: if $i \equiv 0 \pmod 2$, then

$$\begin{aligned} p(x_i') &= \frac{32}{\pi^4(2i-1)^2} \left(\sum_{j=1}^{\infty} \left(\frac{i+2j-2}{(4j-3)(i-2j+1)} \right)^2 + \sum_{j=1}^{\infty} \left(\frac{i-2j}{(4j-1)(i+2j-1)} \right)^2 \right) \\ &= \frac{32}{\pi^4(2i-1)^2} \sum_{j=-\infty}^{\infty} \left(\frac{i+2j-2}{(4j-3)(i-2j+1)} \right)^2 = \frac{24\pi i - 12\pi + 32}{\pi^3(2i-1)^3}. \end{aligned}$$

The last equality can be verified by Mathematica. The case $i \equiv 1 \pmod 2$ is proved analogously. The value $d_P = 122$ follows from a straight foreward computation. □

Acknowledgement

The author thanks Hansjörg Albrecher for many valuable suggestions and remarks.

References

[1] F. Åkesson and J. P. Lehoczky, *Discrete eigenfunction expansion of the multi-dimensional brownian motion and the Ornstein-Uhlenbeck process*, preprint.

[2] R. E. Caflisch, W. J. Morokoff, and A. B. Owen, *Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension*, J. Comp. Finance **1** (1997), no. 1, 27–46.

[3] J. Dick, I. H. Sloan, X. Wang, and H. Woźniakowski, *Liberating the weights*, J. Complexity **20** (2004), no. 5, 593–623.

[4] B. Efron and C. Stein, *The jackknife estimate of variance*, Ann. Statist. **9** (1981), no. 3, 586–596.

[5] P. Glasserman, *Monte Carlo methods in financial engineering*, Applications of Mathematics (New York), vol. 53, Springer-Verlag, New York, 2004.

[6] F. J. Hickernell, *Quadrature error bounds with applications to lattice rules*, SIAM J. Numer. Anal. **33** (1996), no. 5, 1995–2016.

[7] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992.

[8] P. E. Kloeden, E. Platen, and H. Schurz, *Numerical solution of SDE through computer experiments*, Universitext, Springer-Verlag, Berlin, 1994.

[9] P. Lévy, *Wiener's random function, and other Laplacian random functions*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950 (Berkeley and Los Angeles), University of California Press, 1951, pp. 171–187.

[10] C. L. Mallows, *A note on asymptotic joint normality*, Ann. Math. Statist. **43** (1972), 508–515.

[11] G.N. Milshtein, *Approximate integration of stochastic differential equations.*, Theory Probab. Appl. **19** (1974), 557–562.

[12] A. B. Owen, *Orthogonal arrays for computer experiments, integration and visualization*, Statist. Sinica **2** (1992), no. 2, 439–452.

[13] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C*, second ed., Cambridge University Press, Cambridge, 1992.

[14] T. Rydén and M. Wiktorsson, *On the simulation of iterated Itô integrals*, Stochastic Process. Appl. **91** (2001), no. 1, 151–168.

[15] X. Wang and K. T. Fang, *The effective dimension and quasi-Monte Carlo integration*, J. Complexity **19** (2003), no. 2, 101–124.

[16] M. Wiktorsson, *Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions*, Ann. Appl. Probab. **11** (2001), no. 2, 470–487.