Uniform Distribution Theory **6** (2011), no.1, 57–64

uniform distribution theory

DISCREPANCY BETWEEN QMC AND RQMC, II

Shu Tezuka

ABSTRACT. There are two types of randomization for (t, m, d)-nets: Owen scrambling and random digital shift. In the previous paper [Uniform Distribution Theory, **2** (2007), 93-105], we introduced a class of functions for which any Sobol' points have zero integration error, whereas Owen scrambling of Sobol' points has the same variance of integration error as that of simple Monte Carlo methods. In this paper, by using the same functions as the paper mentioned above, we construct an example of functions for which any Sobol' points have zero integration error, whereas not only Owen scrambling but also random digital shift of Sobol' points have variance of integration error no smaller than that of simple Monte Carlo methods.

Communicated by Oto Strauch

1. Introduction

In this paper, we study the difference between the integration errors of quasi-Monte Carlo (QMC) and randomized quasi-Monte Carlo (RQMC). In QMC, (t, m, d)-nets and (t, d)-sequences are one of the most popular methods for producing low-discrepancy points [4, 5, 9, 10], and among them it is widely accepted that Sobol' sequences with judiciously chosen direction numbers perform best for real practical applications (see, e.g., Jäckel [2]). As for RQMC [3], there are two major techniques for randomization of (t, m, d)-nets: Owen scrambling [6, 7] and random digital shift. Historically, random shift was considered by Cranley and Patterson [1] for good lattice rules. Random digital shift is an analogue of random shift for (t, m, d)-nets so as to preserve the net property with the same t-value.

In the previous paper [12], we introduced a class of functions for which Sobol' points with any direction numbers have zero integration error, whereas

2010 Mathematics Subject Classification: 65D30.
Keywords: Generalized Sobol' sequences, high dimensional integration, Monte Carlo and quasi-Monte Carlo methods, Owen's scrambling, random digital shift.
This research was supported by KAKENHI (18540128).

Owen scrambling of Sobol' points has the same variance of integration error as that of simple Monte Carlo methods. This result implies that the error of numerical integration based on Sobol' points with randomization can be totally different from the one without it. The focus of the present paper is on the other type of randomization, i.e., random digital shift. The organization of the paper is as follows: Section 2 summarizes necessary definitions and notations, and reminds us of a class of functions introduced in [12]. In Section 3, we first give the variance of integration error of random digital shift for a function in $L_2[0,1]^d$ in general, then derive a much simpler formula of the variance for the class of functions introduced in Section 2. Section 4 presents our main result, that is to say, an example of functions for which any Sobol' points have zero integration error, whereas not only Owen scrambling but also random digital shift of Sobol' points have variance of integration error no smaller than that of simple Monte Carlo methods. In the last section, we discuss the significance of this result and future research directions.

2. Definitions and notations

First, we recall the definition of Walsh functions:

$$\mathrm{wal}_0(x) = 1 \text{ for } x \in [0,1),$$

and for a nonnegative integer $m \geq 1$,

$$\mathrm{wal}_k(x) = (-1)^{k^{(1)}x^{(1)}+k^{(2)}x^{(2)}+\cdots} = (-1)^{(K,X)} \quad \text{for} \quad x \in [0,1),$$

where $k = k^{(1)} + k^{(2)}2 + \cdots$, and $x = x^{(1)}2^{-1} + x^{(2)}2^{-2} + \cdots$ in their canonical base 2 representations, and $K = (k^{(1)}, k^{(2)}, \ldots)$ and $X = (x^{(1)}, x^{(2)}, \ldots)$ are the binary vector representations of k and x, respectively. (K, X) denotes the inner product over GF(2) of K and X. The Rademacher functions are the subclass of the Walsh functions for which k is a power of 2.

Let t_ℓ be an integer with $2^{\ell-1} \leq t_\ell < 2^\ell$ for $\ell = 1, 2, \ldots$, and denote its binary representation by $t_\ell = t_\ell^{(1)} + t_\ell^{(2)}2 + \cdots + t_\ell^{(\ell)}2^{\ell-1}$ with $t_\ell^{(\ell)} = 1$. We define a nonsingular lower triangular infinite matrix T, where its (ℓ, j)-element for $j \leq \ell$ is equal to $t_\ell^{(j)}$. Hereafter, we denote

$$r_0^{(T)}(x) = \mathrm{wal}_0(x), \quad \text{for} \quad x \in [0,1),$$

and for $\ell = 1, 2, \ldots,$

$$r_\ell^{(T)}(x) = \mathrm{wal}_{t_\ell}(x). \tag{1}$$

Note that the matrix T specifies uniquely a subclass of the Walsh functions, and that the identity matrix I corresponds to the Rademacher functions.

From now on, we fix d matrices $T_1, \ldots, T_d$ which specify d subclasses of the Walsh functions. In the previous paper [12], we considered linear combination of $\prod_{i\in u} r_\ell^{(T_i)}(x_i)$, where $u \subseteq \{1, \ldots, d\}$. In this paper, by using the same functions we define a class $\mathfrak{F}_m$ of functions in d dimensions as follows:

Definition 1. For $m \geq 1$ and $d \geq 1$, we define an L_2 function on $[0, 1]^d$ by

$$f(x_1, \ldots, x_d) = c_0 + \sum_{\ell=m+1}^{\infty} \sum_{\emptyset \neq u \subseteq \{1, \ldots, d\}} c_{u,\ell} \prod_{i\in u} r_\ell^{(T_i)}(x_i),$$

where c_0 is constant and the coefficients $c_{u,\ell}$ satisfy the following condition:

$$\sum_{\ell=m+1}^{\infty} \sum_{\emptyset \neq u \subseteq \{1, \ldots, d\}} c_{u,\ell} = 0. \tag{2}$$

In this paper, we consider a class of (t, d)-sequences in base $b = 2$ whose generator matrices are written as $(T_i)^{-1}U_i, i = 1, \ldots, d$, where $T_i, i = 1, \ldots, d$, are matrices specifying subclasses of the Walsh functions, and $U_i, i = 1, \ldots, d$, are arbitrary nonsingular upper-triangular matrices. Hereafter, we denote this class by $\mathfrak{S}_d$. By definition [10], generalized Sobol' sequences with lower triangular matrices $(T_i)^{-1}, i = 1, \ldots, d$, are a subset of $\mathfrak{S}_d$, where $U_i, i = 1, \ldots, d$, are constructed based on irreducible polynomials with the so-called direction numbers. Note that Sobol' sequences [8] are the special case of generalized Sobol' sequences with $T_i = I, i = 1, \ldots, d$.

3. Integration error for random digital shift of Sobol' points

First, we give the variance of integration error with respect to random digital shift for any function $f(\mathbf{x})$ in $L_2[0, 1]^d$. Denote d-dimensional integral by

$$\mathrm{I}(f) = \int_{[0,1]^d} f(\mathbf{x}) d\mathbf{x}$$

and an equal weight cubature by

$$\mathrm{Q}_N(f) = N^{-1} \sum_{n=0}^{N-1} f(\mathbf{x}_n),$$

where $\mathbf{x}_n, n = 0, 1, \ldots, N-1$ are d-dimensional cubature points. Then the integration error is given by

$$\mathrm{Er}(f) = \big|\mathrm{I}(f) - \mathrm{Q}_N(f)\big|.$$

Hereafter, we denote d-dimensional Walsh functions by

$$\mathrm{wal}_{\mathbf{k}}(\mathbf{x}) = \prod_{i=1}^{d} \mathrm{wal}_{k_i}(x_i),$$

where $\mathbf{k} = (k_1, \ldots, k_d)$ and $\mathbf{x} = (x_1, \ldots, x_d)$.

Lemma 1. *Denote the d-dimensional Walsh series expansion of an L_2 function on $[0,1]^d$ by*

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{N}^d} w_{\mathbf{k}} \mathrm{wal}_{\mathbf{k}}(\mathbf{x}).$$

Then the variance of integration error with respect to random digital shift for a function $f(\mathbf{x})$ in $L_2[0,1]^d$ is given by

$$\mathrm{V}\big(\mathrm{Er}_N(f)\big) = \sum_{\mathbf{k}\in\mathbb{N}^d\setminus\{0\}} w_{\mathbf{k}}^2\ \mathrm{Q}_N^2(\mathrm{wal}_{\mathbf{k}}).$$

Proof. Note that $\mathrm{I}(f) = w_0$. Thus, the integration error is written as

$$\mathrm{Er}_N(f) = \big|\,\mathrm{I}(f) - \mathrm{Q}_N(f)\big| = \left|\sum_{\mathbf{k}\in\mathbb{N}^d\setminus\{0\}} w_{\mathbf{k}}\ \mathrm{Q}_N(\mathrm{wal}_{\mathbf{k}})\right|.$$

Since $\mathrm{Q}_N(f)$ with respect to random digital shift is an unbiased estimator to $\mathrm{I}(f)$, the variance of $\mathrm{Er}_N(f)$ with respect to random digital shift is given by

$$\mathrm{V}\big(\mathrm{Er}_N(f)\big) = \mathrm{E}\big(\mathrm{Er}_N(f)^2\big) = \mathrm{E}\left(\left(\sum_{\mathbf{k}\in\mathbb{N}^d\setminus\{0\}} w_{\mathbf{k}}\ \mathrm{Q}_N(\mathrm{wal}_{\mathbf{k}};\mathbf{s})\right)^2\right),$$

where

$$\mathrm{Q}_N(\mathrm{wal}_{\mathbf{k}};\mathbf{s}) = \frac{1}{N}\sum_{n=0}^{N-1} \mathrm{wal}_{\mathbf{k}}(\mathbf{x}_n \oplus \mathbf{s}),$$

and the expectation is taken over all $\mathbf{s}$ which are uniformly distributed in $[0,1)^d$. Here, the operation $\oplus$ means the bit-wise exclusive-or. Since

$$\mathrm{wal}_{\mathbf{k}}(\mathbf{x}\oplus\mathbf{s}) = \mathrm{wal}_{\mathbf{k}}(\mathbf{x})\mathrm{wal}_{\mathbf{k}}(\mathbf{s}),$$

we have

$$\begin{aligned}
\mathrm{V}\big(\mathrm{Er}_N(f)\big) &= \mathrm{E}\left(\sum_{\mathbf{k},\mathbf{h}\in\mathbb{N}^d\setminus\{0\}} w_{\mathbf{k}}w_{\mathbf{h}}\mathrm{wal}_{\mathbf{k}}(\mathbf{s})\mathrm{wal}_{\mathbf{h}}(\mathbf{s})\ \mathrm{Q}_N(\mathrm{wal}_{\mathbf{k}})\ \mathrm{Q}_N(\mathrm{wal}_{\mathbf{h}})\right)\\
&= \sum_{\mathbf{k},\mathbf{h}\in\mathbb{N}^d\setminus\{0\}} w_{\mathbf{k}}w_{\mathbf{h}}\ \mathrm{E}\big(\mathrm{wal}_{\mathbf{k}}(\mathbf{s})\mathrm{wal}_{\mathbf{h}}(\mathbf{s})\big)\ \mathrm{Q}_N(\mathrm{wal}_{\mathbf{k}})\ \mathrm{Q}_N(\mathrm{wal}_{\mathbf{h}})\\
&= \sum_{\mathbf{k}\in\mathbb{N}^d\setminus\{0\}} w_{\mathbf{k}}^2\ \mathrm{Q}_N^2(\mathrm{wal}_{\mathbf{k}}).
\end{aligned}$$

Note that $\mathrm{E}(\mathrm{wal}_{\mathbf{k}}(\mathbf{s})\mathrm{wal}_{\mathbf{h}}(\mathbf{s})) = 1$ if $\mathbf{k} = \mathbf{h}$; otherwise 0. The proof is complete. □

We prove our main result.

Theorem 1. *If the first N points of a generalized Sobol' sequence with any direction numbers are used for the cubature Q_N, then for any function f in $\mathfrak{F}_m$, the variance of integration error with respect to random digital shift is given by*

$$\mathrm{V}\big(\mathrm{Er}_N(f)\big) = \sum_{\ell=m+1}^{\infty}\ \sum_{\emptyset\neq u\subseteq\{1,\dots,d\}} c_{u,\ell}^2\,,$$

where $N = 2^m$.

Proof. The equation (1) implies that $\prod_{i\in u} r_\ell^{(T_i)}(x_i)$ is equivalent to $\mathrm{wal}_{\mathbf{k}}(\mathbf{x})$ such that

$$r_\ell^{(T_i)}(x_i) = \begin{cases} \mathrm{wal}_{k_i}(x_i) & \text{if } i\in u,\\ \mathrm{wal}_0(x_i) = 1, & \text{otherwise,}\end{cases}$$

where we denote the (ℓ, j)-element of T_i by $t_{i,\ell}^{(j)}$ and $k_i = t_{i,\ell}^{(1)} + t_{i,\ell}^{(2)}2 + \cdots + t_{i,\ell}^{(\ell)}2^{\ell-1}$.

From Lemma 4 of [11] and the fact that the first point of generalized Sobol' sequences is $(0, \dots, 0)$, we have

$$\mathrm{Q}_N\left(\prod_{i\in u} r_\ell^{(T_i)}(x_i)\right) = 1 \tag{3}$$

for $\ell > m$. Hence, Lemma 1 and Definition 1 give us that

$$\mathrm{V}\big(\mathrm{Er}_N(f)\big) = \sum_{\ell=m+1}^{\infty}\ \sum_{\emptyset\neq u\subseteq\{1,\dots,d\}} c_{u,\ell}^2\,,$$

since we have different Walsh functions $\prod_{i\in u} r_\ell^{(T_i)}(x_i)$ for different (u, ℓ). □

The following result is concerned with the integrations using Sobol' points with and without Owen scrambling.

Theorem 2. *Let $N = 2^m$. For any function f in $\mathfrak{F}_m$, the integration error using the first N points of a generalized Sobol' sequence with any direction numbers is zero, and the variance of integration error with respect to Owen scrambling of the first N points of a generalized Sobol' sequence with any direction numbers is*

$$V\big(\mathrm{Er}_N(f)\big) = \frac{1}{N} \sum_{\ell=m+1}^{\infty} \sum_{\emptyset \neq u \subseteq \{1,\dots,d\}} c_{u,\ell}^2 \,.$$

P r o o f. From Definition 1, we have

$$\mathrm{Er}_N(f) = \left| \sum_{\ell=m+1}^{\infty} \sum_{\emptyset \neq u \subseteq \{1,\dots,d\}} c_{u,\ell}\ \mathrm{Q}_N \left(\prod_{i \in u} r_\ell^{(T_i)}(x_i) \right) \right| .$$

If we use the first N points of a generalized Sobol' sequence with any direction numbers, we obtain $\mathrm{Er}_N(f) = 0$ because of the condition (2) of Definition 1 and the equation (3).

For the second part, first we should notice that if $\ell > m$, then the gain coefficient $\Gamma_{u,\kappa} = 1$ for the first $N = 2^m$ points of a generalized Sobol' sequence, where $\kappa = (\ell - 1, \dots, \ell - 1)$. Therefore, from Lemmas 2 and 3 of [12], we obtain the variance of integration error with respect to Owen scrambling as

$$\begin{aligned} V\big(\mathrm{Er}_N(f)\big) &= \frac{1}{N} \sum_{\ell=m+1}^{\infty} \sum_{\emptyset \neq u \subseteq \{1,\dots,d\}} \Gamma_{u,\kappa} \sigma_{u,\kappa}^2 \\ &= \frac{1}{N} \sum_{\ell=m+1}^{\infty} \sum_{\emptyset \neq u \subseteq \{1,\dots,d\}} c_{u,\ell}^2 \,, \end{aligned}$$

where $\sigma_{u,\kappa}^2$ indicates the variance of a function $c_{u,\ell} \prod_{i \in u} r_\ell^{(T_i)}(x_i)$. □

Note that we have the variance of integration error,

$$V\big(\mathrm{Er}_N(f)\big) = \frac{1}{N} \sum_{\ell=m+1}^{\infty} \sum_{\emptyset \neq u \subseteq \{1,\dots,d\}} c_{u,\ell}^2 \,,$$

for simple Monte Carlo methods with N samples. Therefore, we can conclude that for any function f in $\mathfrak{F}_m$, any Sobol' points have zero integration error, whereas not only Owen scrambling but also random digital shift of Sobol' points have variance of integration error no smaller than that of simple Monte Carlo methods.

4. Discussion

Originally, RQMC was introduced to give a more realistic estimate of the integration error, because the Koksma-Hlawka bound for QMC provides loose error estimates for many practical applications. However, this paper as well as the previous paper [12] showed that the consistency is not always guaranteed between QMC errors and RQMC errors. As pointed out in [12], the functions $\mathfrak{F}_m$ considered in this paper are artificial, but very simple. For example, in two dimensions, they contain chess-board functions and their linear combinations. Even for such simple functions, the integration errors of QMC and RQMC (not only Owen scrambling but also random digital shift) are shown to be totally different. It is reasonable to think that there exist some practical applications in which similar results can happen.

Acknowledgments. The author thanks the anonymous referee for valuable suggestions.

REFERENCES

[1] CRANLEY, R. – PATTERSON, T. N. L.: *Randomization of number theoretic methods for multiple integration*, SIAM J. Numer. Anal. **13** (1976), no. 6, 904–914.

[2] JÄCKEL, P.: *Monte Carlo Methods in Finance*, John Wiley and Sons, New York, 2002.

[3] LEMIEUX, C.: *Monte Carlo and Quasi-Monte Carlo Sampling*, *Springer Ser. Statist.* Springer, New York, 2009.

[4] NIEDERREITER, H.: *Random Number Generation and Quasi-Monte Carlo Methods*, *CBMS-NSF Regional Conference Series in Applied Mathematics* **63**, SIAM, Philadelphia, PA, 1992.

[5] NOVAK, E. – WOŹNIAKOWSKI, H.: *Tractability of Multivariate Problems, Volume II: Standard Information for Functionals*, *EMS Tracts in Mathematics* **12**, EMS, Zürich, 2010.

[6] OWEN, A. B.: *Monte Carlo variance of scrambled equidistribution quadrature*, SIAM Journal of Numerical Analysis **34** (1997), 1884–1910.

[7] OWEN, A. B.: *Scrambled net variance for integrals of smooth functions*, Ann. Stat. **25** (1997), no. 4, 1541–1562.

[8] SOBOL', I. M.: *The distribution of points in a cube and the approximate evaluation of integrals*, USSR Computat. Math. Math. Phys. **7** (1967), no. 4, 86–112.

[9] STRAUCH, O. – PORUBSKÝ, Š.: *Distribution of Sequences. A Sampler*, *Schriftenreihe der Slowakischen Akademie der Wissenschaften* **1**. Peter Lang, Bern, 2005.

[10] TEZUKA, S.: *Uniform Random Numbers. Theory and Practice*, *The Kluwer International Series in Engineering and Computer Science* **315**, Kluwer Academic Publishers, Dordrecht, 1995.
[11] TEZUKA, S.: *On the necessity of low-effective dimension*, J. Complexity **21** (2005), no. 5, 710–721.
[12] TEZUKA, S.: *Discrepancy between QMC and RQMC*, Unif. Distrib. Theory **2** (2007), no. 1, 93–105.

Received April 3, 2010
Accepted December 6, 2010

Shu Tezuka
Faculty of Mathematics
Kyushu University
744 Motooka, Nishi-ku,
Fukuoka-shi, Fukuoka-ken
JAPAN 819-0395
E-mail: tezuka@math.kyushu-u.ac.jp